عناصر الدرس

الباب الأول : مفاهيم أساسية

- مقدمة
- مفهوم علم الإحصاء
- المجتمع والعينة
- المعلم والإحصاءة
- خطوات العملية الإحصائية

الباب الأول : مفاهيم أساسية

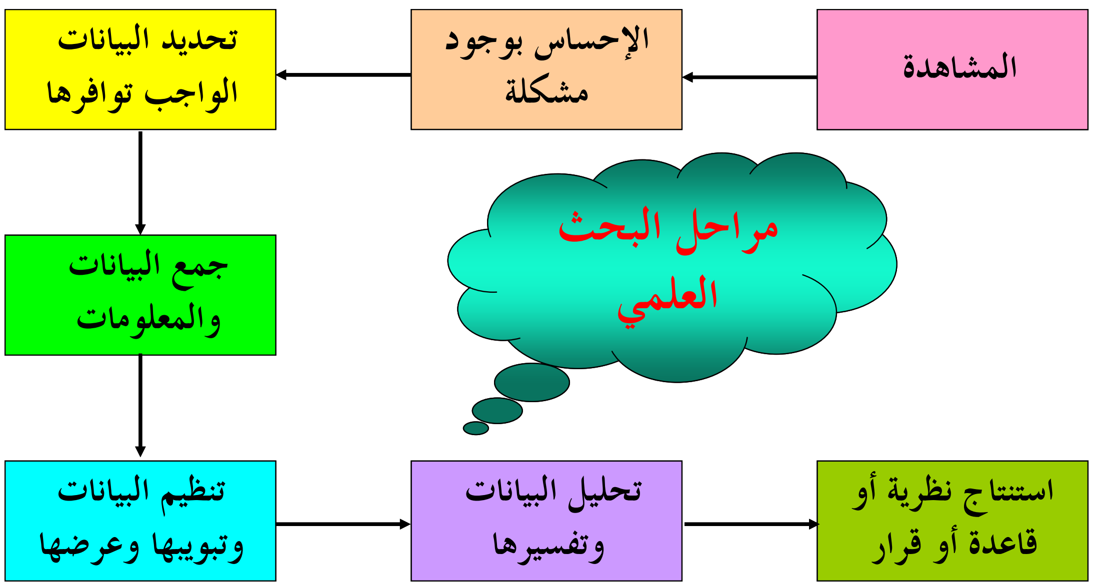
(1-1) مقدمة :

الغرض من العلم (بوجهٍ عامٍ) هو البحث عن الحقيقة ، والبحث العلمي هو الوسيلة للوصول إلى حقائق الأشياء والظواهر ومعرفة كل العلاقات التي تربط بينها وبعضها البعض ، سواء كانت هذه الظواهر اجتماعية أو اقتصادية أو طبيعية أو غير ذلك ، لذا يستخدم البحث العلمي العلم بقصد دراسة ظاهره معينة لاكتشاف حقائقها ومعرفة القواعد العامة التي تحكمها .

والإحساس بوجود مشكلة (أو ظاهرة) ما يمثل شرطاً أساسياً للقيام ببحث علمي ، وهذا الإحساس لا يأتي إلا من خلال المشاهدة للظواهر المختلفة ، وهذا يتطلب تحديد البيانات الواجب توافرها حتى يمكن إجراء البحث والوصول إلى نتائج مقبولة يمكن الاعتماد عليها في تفسير تلك الظواهر المختلفة التي قد تثير الاهتمام .

يأتي بعد ذلك جمع لتلك البيانات من مصادرها المختلفة وتنظيمها وتبويبها وعرضها في صور جدولية أو بيانية ، ثم يتم استخدامها في حساب بعض المقاييس الخاصة بهذه الظواهر وإجراء تحليل لتلك البيانات بما يساعد في تفسير النتائج المختلفة للبيانات واستخدامها في استنتاج نظرية أو قاعدة أو قانون أو المساعدة في اتخاذ القرارات أو التنبوء بنتائج مستقبلية .

والشكل التالي [شكل (1-1)] يمكن أن يوضح **الإطار العام** لأي بحث علمي :



شكل (1-1) : الإطار العام لأي بحث علمي

(1–2) مفهوم علم الإحصاء :

يختص علم الإحصاء بالطرق العملية لجمع وتنظيم وعرض وتحليل البيانات وكذلك الوصول إلى نتائج مقبولة وقرارات سليمة على ضوء هذا التحليل ، ويُستخدم اصطلاح "علم الإحصاء" في معناه الضيق للتعبير عن البيانات والمقاييس المستخرجة من تلك البيانات (مثل المتوسطات) ، وعلى هذا نتحدث عن إحصاءات البطالة والحوادث والمواليد والوفيات وغيرها .

والإحصاء من أقدم العلوم التي استخدمها الإنسان ، فقد استخدمه قدماء المصريين والرومان وغيرهما بغرض التعداد السكاني تمهيداً لتوزيع الأراضي الزراعية على السكان وإعداد الجيوش وفرض الضرائب إلى غير ذلك من الأمور . وفي الحضارة الإسلامية استخدمه المسلمون في حالات كثيرة حيث أجريت في مراحل مختلفة من الخلافة الإسلامية إحصائيات على أراضي العراق ومصر والشام من حيث المساحة والصلاحية للزراعة وصنفت حسب ملاكها وما تنتجه من محاصيل . وحديثاً امتدت تطبيقات الإحصاء لجميع الميادين العلمية كالزراعة والطب والصناعة والهندسة وغيرها ، واعتمدت عليها الإدارات والمصالح الحكومية والهيئات والمنظمات في حل الكثير من المشاكل واتخاذ القرارات في الكثير من الموضوعات .

وقديماً عُرِف علم الإحصاء على أنه جمع البيانات عن ظاهرة معينة وترتيبها في جداول أو عرضها في صورة رسومات وأشكال بيانية بسيطة ، لكن مع تقدم العلوم بدأ علم الإحصاء يلعب دوراً متزايداً في حياتنا اليومية بحيث أصبح يشغل حيزاً كبيراً بين بقية العلوم الأخرى ، فأصبح يبحث في جمع البيانات وتنظيمها وعرضها وتحليلها واستنتاج وتوقع نتائج واتخاذ قرارات . وينقسم علم الإحصاء إلى قسمين :

الأول (ويُسمى الإحصاء الوصفي) : وهو يهتم بجمع وتبويب وعرض ووصف البيانات وحساب بعض المقاييس الخاصة بها دون الوصول إلى نتائج أو استدلالات خاصة .

الثاني (ويُسمى الإحصاء الاستقرائي أو الاستدلال الإحصائي) : وهو يبحث في استقراء النتائج واتخاذ القرارات .

(1–3) المجتمع والعينة :

عند جمع بيانات تخص خاصية (أو ظاهرة) معينة من خصائص مجموعة من الأفراد أو الأشياء (مثل أطوال أو أوزان طلبة جامعيين أو عدد الوحدات المعيبة في إنتاج مصنع في يوم معين) فإنه قد يكون من المستحيل أو غير العملي ملاحظة المجموعة بأكملها خاصةً إذا كانت كبيرة ، وبدلاً من اختيار المجموعة كلها (والتي تُسمى **بالمجتمع**) فإنه يمكن اختيار جزء صغير من المجموعة يُسمى **بالعينة** . فمثلاً لتحليل نتائج طلاب المملكة في مقرر اللغة الإنجليزية لطلاب وطالبات الثانوية العامة ، فمن المستحيل أو غير العملي أن نقوم بجمع درجات جميع الطلاب في هذا المقرر على مستوى المملكة وتنظيمها وتحليلها ثم نستنتج بعض النتائج من هذا التحليل ، هنا

يكون المجتمع هو جميع طلاب المملكة . بدلاً من ذلك نقوم باختيار عينة من هؤلاء الطلاب (تحت شروط معينة حتى تكون ممثلة للمجتمع) ونقوم بتحليل بيانات هذه العينة ونخرج من هذا التحليل باستدلالات تخص المجتمع ككل . وحيث أن هذا النوع من الاستدلال لا يمكن أن يكون مؤكداً فإن **لغة الاحتمال تُستخدم عند عرض النتائج** .

أي أن **المجتمع** يتكون من جميع المشاهدات التي تهتم بها ، أما **العينة** فهي فئة جزئية من المجتمع

(1–4) المعلم *Parameter* والإحصاءة *Statstic* :

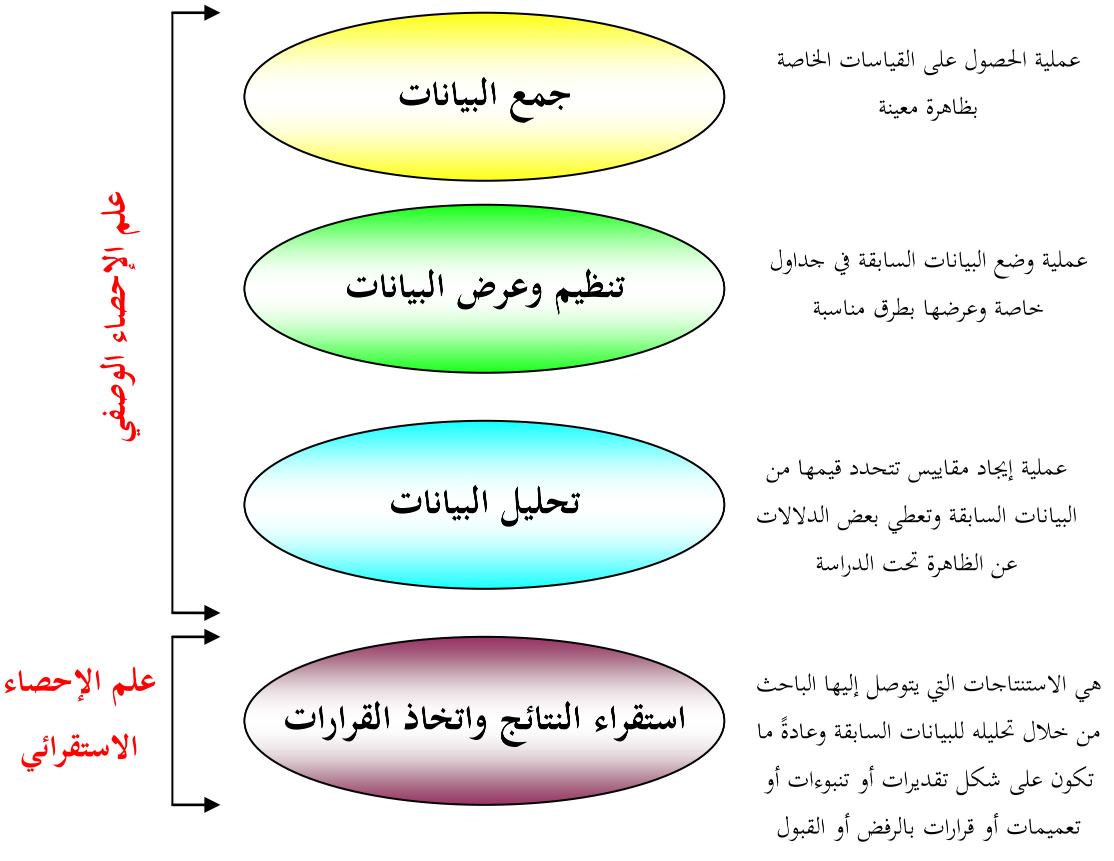
إذا أردنا أن نوجد (مثلاً) متوسط درجات 1000 في مقرر ما بإحدى المدارس فإن هذا المتوسط يُسمى أحد **معالم** المجتمع [مجتمعنا هنا هو درجات جميع الطلاب] ، لكن إذا أخذنا **عينة** [ممثلة للمجتمع] من هؤلاء الطلاب وحسبنا لهم متوسط درجاتهم ، فإن هذا المتوسط (للعينة) يُسمى **إحصاءة** . أي أن **المعلم** هي مقياس للمجتمع تعتمد في حسابها على جميع قيم المجتمع ، أما **الإحصاءة** فهي إحدى مقاييس العينة (الممثلة للمجتمع) والتي تعتمد في حسابها على جميع قيم العينة .

(1–5) خطوات العملية الإحصائية :

يمكن تلخيص الخطوات الأساسية لأي عملية إحصائية في الآتي :

- **جمع البيانات :** هي عملية الحصول على القياسات والبيانات الخاصة بظاهرة معينة ، وعادةً ما تُسمى البيانات المجمعة **بالبيانات الخام** .
- **تنظيم وعرض البيانات :** هي عملية وضع البيانات الخاصة بظاهرة معينة في جداول منسقة وعرضها بطرق مناسبة .
- **تحليل البيانات :** هي عملية إيجاد قيم لمقاييس تتحدد قيمها من البيانات الخاصة بظاهرة معينة وتُعطي بعض الدلالات عن تلك الظاهرة .
- **استقراء النتائج واتخاذ القرارات :** هي أهم أهداف علم الإحصاء وهي ببساطة تمثل الاستنتاجات التي يتوصل إليها الباحث من خلال تحليله للبيانات الخاصة بظاهرة معينة وغالباً ما تكون على شكل تقديرات أو تنبوءات أو تعميمات أو قرارات بالرفض أو القبول .

ويبين شكل (1–2) الخطوات الأساسية السابقة لأي عملية إحصائية .



عملية الحصول على القياسات الخاصة بظاهرة معينة

عملية وضع البيانات السابقة في جداول خاصة وعرضها بطرق مناسبة

عملية إيجاد مقاييس تتحدد قيمها من البيانات السابقة وتعطي بعض الدلالات عن الظاهرة تحت الدراسة

هي الاستنتاجات التي يتوصل إليها الباحث من خلال تحليله للبيانات السابقة وعادةً ما تكون على شكل تقديرات أو تنبوءات أو تعميمات أو قرارات بالرفض أو القبول

شكل (1–2) : الخطوات الأساسية لأي عملية إحصائية

(1–6) البيانات :

يمكن ببساطة تعريف البيانات على أنها مجموعة من "**المشاهدات أو القياسات**" التي تخص الظاهرة تحت الدراسة ، والكمية (أو الخاصية) التي نقوم بمشاهدتها أو قياسها تُسمى **بالمتغير** وعادةً نرمز له برمز مثل $x , y , A , B , \dots$ ، فمثلاً :

مثال رقم	المتغير x	القياسات أو المشاهدات
(1)	لون العين	أخضر – أزرق – بني – ....
(2)	عدد الطلاب في فصول مدرسة	15 – 18 – 20 – 25 – 17 - ....
(3)	أطوال مجموعة من الطلاب في أحد الفصول (بالمتر)	1.5 – 1.52 – 1.71 – 1.83 - ....
(4)	أوزان بعض العاملات بمصنع معين (بالكيلوجرام)	55.2 – 60.1 – 63.25 – ....
(5)	تقديرات عدد من الطلاب في مقرر الإحصاء	A – B – C – D – F – A - ....

والمتغير (أي الظاهرة تحت الدراسة) إما أن يكون :

(أ) **نوعياً** [أي لا يمكن التعبير عنه بعدد كما في الأمثلة (1) ، (5)] مثل لون العين أو تقدير الطلاب في مقرر ما ، وتُسمى البيانات عندئذٍ بـ **بيانات نوعية** .

(ب) أو **كمياً** [أي يمكن التعبير عنه بعدد كما في الأمثلة (2) ، (3) ، (4)] ، وفي هذه الحالة قد يأخذ المتغير أي قيمة بين قيمتين معينتين [كما في الحالات (3) ، (4)] ، عندئذٍ يُسمى **متغير متصل** وتُسمى البيانات عندئذٍ بـ **بيانات متصلة** ، وخلاف ذلك يُسمى **متغير متقطع** [كما في الحالة (2)] ، وعندئذٍ تُسمى البيانات بـ **بيانات متقطعة** .

وفي بعض الأحيان يُعرف المتغير الكمي المتقطع بأنه المتغير الكمي الذي يمكن أن **يُعد** ولا **يُقاس** ، أما المتغير الكمي المتصل فيوصف بأنه المتغير الكمي الذي يمكن أن **يُقاس** ولا **يُعد** ، وهذا التعريف قد يكون مفيداً في كثير من الأحيان [**لكنه ليس بالدقة الكافية**] .

أمثلة لذلك :

- المتغير N (الذي يمثل عدد الأطفال في أسرة معينة) يمكن أن يأخذ القيم ..... , 3 , 2 , 1 , 0 ولا يمكن أن يأخذ قيمة مثل 1.5 ، وبالتالي فإن N متغير متقطع .
- العمر A لشخص من الممكن أن يكون 42 سنة أو 43 سنة أو أي قيمة بين 42 , 43 حسب درجة الدقة في القياس ، لذا فإن A متغير متصل .

ملخص للدرس الأول [الباب الأول : مفاهيم أساسية]

- **علم الإحصاء الوصفي :** وهو يهتم بجمع وتبويب وعرض ووصف البيانات وحساب بعض المقاييس الخاصة بها دون الوصول إلى نتائج أو استدلالات خاصة .
- **علم الإحصاء الاستقرائي (أو الاستدلال الإحصائي) :** وهو يبحث في استقراء النتائج واتخاذ القرارات
- **المجتمع :** هو المجموعة الكاملة من العناصر موضع الدراسة .
- **العينة :** هي مجموعة جزئية من المجتمع موضع الدراسة .
- **المعلم :** هو قياس يوضح أو يصف خاصية معينة عن المجتمع .
- **الإحصاءة :** هي قياس يوضح أو يصف خاصية معينة عن العينة .
- **المتغير :** خاصية كميه أو وصفيه تأخذ مفرداتها قيم مختلفة عند قياسها أو مشاهدتها
- **البيانات :** هي مجموعة من " **المشاهدات أو القياسات**" التي تخص الظاهرة تحت الدراسة ، والكمية التي نقوم بمشاهدتها أو قياسها تُسمى **بالمتغير** ، وعادةً نرمز له برمز مثل $x , y , A , B , ..$.
- **المتغير النوعي :** هو الكمية التي نقوم بقياسها ولا يمكن التعبير عنه **بعدد** [مثل لون العين أو تقدير الطلاب] ، وتُسمى البيانات التي يكون فيها المتغير نوعياً بـ **البيانات النوعية** .
- **المتغير الكمي :** هو الكمية التي نقوم بقياسها ويمكن التعبير عنه **بعدد** [مثل الأطوال أو الأوزان أو أعداد الطلاب] ، وتُسمى البيانات التي يكون فيها المتغير كمياً بـ **البيانات الكمية** . والمتغير الكمي إما أن يكون :

متغير متصل وفيها يمكن أن يأخذ المتغير أي قيمة بين قيمتين معينتين [أو بتعبير آخر هو كمية يمكن أن تُقاس ولا تُعد] ، أو

متغير متقطع وفيها يمكن أن يأخذ المتغير قيماً محددة دون أي قيمة بينها [أو بتعبير آخر هو كمية يمكن أن تُعد ولا تُقاس] .

- **خطوات أي عملية إحصائية**

(أ) **جمع البيانات :** وهي عملية الحصول على القياسات الخاصة بظاهرة معينة وعادةً ما تُسمى البيانات المجمعة بالبيانات الخام .

(ب) **تنظيم وعرض البيانات :** وهي عملية وضع البيانات السابقة في جداول خاصة وعرضها بطرق مناسبة

(ج) **تحليل البيانات :** وهي عملية إيجاد مقاييس تتحدد قيمها من البيانات السابقة وتعطي بعض الدلالات عن الظاهرة تحت الدراسة

(د) **استقراء النتائج واتخاذ القرارات :** وهي الاستنتاجات التي يتوصل إليها الباحث من خلال تحليله للبيانات السابقة وعادةً ما تكون على شكل تقديرات أو تنبوءات أو تعميمات أو قرارات بالرفض أو القبول

تدريبات (1)

الإجابة النهائية لجميع التمرينات موجودة في نهاية التدريب

اختر الإجابة الصحيحة

(1) هو العلم الذي يهتم بجمع وتبويب وعرض ووصف البيانات وحساب بعض المقاييس الخاصة بها دون الوصول إلى نتائج أو استدلالات خاصة

(أ) علم الإحصاء الوصفي (ب) علم الإحصاء الاستقرائي

(ج) علم تقنية المعلومات (د) علم تكنولوجيا المعلومات

(2) هو العلم الذي يبحث في استقراء النتائج واتخاذ القرارات

(أ) علم الإحصاء الوصفي (ب) علم الإحصاء الاستقرائي

(ج) علم تقنية المعلومات (د) علم تكنولوجيا المعلومات

(3) .... هي عملية الحصول على القياسات والبيانات الخاصة بظاهرة معينة

(أ) جمع البيانات (ب) تنظيم وعرض البيانات

(ج) تحليل البيانات (د) استقراء النتائج واتخاذ القرارات

(4) .... هي عملية وضع البيانات الخاصة بظاهرة معينة في جداول منسقة وعرضها بطرق مناسبة

(أ) جمع البيانات (ب) تنظيم وعرض البيانات

(ج) تحليل البيانات (د) استقراء النتائج واتخاذ القرارات

(5) .... هي عملية إيجاد قيم لمقاييس تتحدد قيمها من البيانات الخاصة بظاهرة معينة وتُعطي بعض الدلالات عن تلك الظاهرة .

(أ) جمع البيانات (ب) تنظيم وعرض البيانات

(ج) تحليل البيانات (د) استقراء النتائج واتخاذ القرارات

(6) .... هي عملية الوصول إلى استنتاجات وتوقعات وتنبوءات الخاصة بظاهرة معينة .

(أ) جمع البيانات (ب) تنظيم وعرض البيانات

(ج) تحليل البيانات (د) استقراء النتائج واتخاذ القرارات

(7) عدد الأيام N في كل شهر هو :

(أ) متغير نوعي (ب) متغير كمي متقطع

(ج) متغير كمي متصل (د) خلاف ما سبق

(8) لون السيارات C في أحد مواقف السيارات هو :

(أ) متغير نوعي (ب) متغير كمي متقطع

(ج) متغير كمي متصل (د) خلاف ما سبق

(9) المسافات d التي بقطعها شخص خلال ساعات يوم معين

(أ) متغير نوعي (ب) متغير كمي متقطع

(ج) متغير كمي متصل (د) خلاف ما سبق

(10) وزن البطاطس W التى تنتجها مزارع مختلفة في أحد المواسم هو :

(أ) متغير نوعي (ب) متغير كمي متقطع

(ج) متغير كمي متصل (د) خلاف ما سبق

(11) الزمن t الذي يأخذه طالب في حل عدد من مسائل الإحصاء هو :

(أ) متغير نوعي (ب) متغير كمي متقطع

(ج) متغير كمي متصل (د) خلاف ما سبق

(12) عدد حبات البطيخ N الذي تبيعه إحدى محلات السوبر ماركت في يوم معين هو :

(أ) متغير نوعي (ب) متغير كمي متقطع

(ج) متغير كمي متصل (د) خلاف ما سبق

(13) اللعبة الرياضية G المفضلة لدى مجموعة من الطلاب هي :

(أ) متغير نوعي (ب) متغير كمي متقطع

(ج) متغير كمي متصل (د) خلاف ما سبق

(14) البيانات المجمعة عن تقديرات الطلبة في أحد المقررات الدراسية هي :

(أ) بيانات نوعية (ب) بيانات كمية متقطعة

(ج) بيانات كمية متصلة (د) خلاف ما سبق

(15) البيانات المجمعة عن درجات الطلبة (مقربة لأقرب عدد صحيح) في أحد المقررات الدراسية هي :

أ) بيانات نوعية	(ب) بيانات كمية متقطعة
(ج) بيانات كمية متصلة	(د) خلاف ما سبق

(16) البيانات الخاصة بالمعدلات التراكمية لطلاب التعليم المطور للانتساب هي :

أ) بيانات نوعية	(ب) بيانات كمية متقطعة
(ج) بيانات كمية متصلة	(د) خلاف ما سبق

هل لاحظت الفرق بين الأسئلة (14) ، (15) ، (16)

(17) البيانات المجمعة عن الحالة الاجتماعية لسكان منطقة معينة هي :

أ) بيانات نوعية	(ب) بيانات كمية متقطعة
(ج) بيانات كمية متصلة	(د) خلاف ما سبق

(18) البيانات المجمعة عن درجة الحرارة ساعة الظهيرة في عدد من مدن المملكة هي :

أ) بيانات نوعية	(ب) بيانات كمية متقطعة
(ج) بيانات كمية متصلة	(د) خلاف ما سبق

(19) البيانات المجمعة عن ماركات (نوع) السيارات في أحد المواقف هي :

أ) بيانات نوعية	(ب) بيانات كمية متقطعة
(ج) بيانات كمية متصلة	(د) خلاف ما سبق

(20) هو المجموعة الكاملة موضع الدراسة :

أ) العينة	(ب) المجتمع
(ج) المعلم	(د) الإحصاءة

(21) مجموعة جزئية من المجتمع محل الدراسة :

أ) العينة	(ب) المجتمع
(ج) المعلم	(د) الإحصاءة

(22) مقياس يوضح أو يصف خاصية معينة في المجتمع .

أ) العينة	(ب) المجتمع
(ج) المعلم	(د) الإحصاءة

(23) مقياس يوضح أو يصف خاصية معينة في العينة .

(أ) العينة　　　　(ب) المجتمع

(ج) المعلم　　　　(د) الإحصاءة

الإجابة :

(1) أ　(2) ب　(3) أ　(4) ب　(5) ج

(6) د　(7) ب　(8) أ　(9) ج　(10) ج

(11) ج　(12) ب　(13) أ　(14) أ　(15) ب

(16) ج　(17) أ　(18) ج　(19) أ　(20) ب

(21) أ　(22) ج　(23) د

الباب الثاني : التوزيعات التكرارية

(2–1) البيانات المنفصلة والبيانات المتصلة :

ذكرنا في الباب السابق (الباب الأول) "**ما هي البيانات ؟**" [هي مجموعة المشاهدات أو القياسات التي تخص ظاهرة معينة تحت الدراسة] وعرفنا **المتغير** على أنه تلك الكمية التي نقوم بمشاهدتها أو قياسها .

كما ذكرنا أن البيانات إما أن تكون نوعية أو كمية ، حيث :

أ) البيانات النوعية :

وهي تلك البيانات التي لا يمكن التعبير عن متغيرها بعدد (أي بيانات غير رقمية) ، مثل :

- لون (أو نوع) السيارات الموجودة في موقف ما [أحمر/أبيض/أسود/ ....]
- الحالة الاجتماعية للسيدات في محافظة معينة [متزوجة/عزباء/مطلقة/أرملة/منفصلة]
- رأيك في قرار خاص بالمؤسسة التي تعمل بها [أوافق بشدة/أوافق/أعترض/أتحفظ – ...]
- وغيره من مثل هذه الأمثلة .

ب) البيانات الكمية :

وهي تلك البيانات التي يُعبر فيها عن المتغير بعدد (أي بيانات رقمية) ، وهذه البيانات بدورها تنقسم إلى

ب – 1) بيانات كمية متصلة :

وفيها يمكن أن يأخذ المتغير أي قيمة بين قيمتين (أي بيانات يمكن أن تُقاس ولا تُعد ، مثل :

- أطوال الطلاب في إحدى المدارس .
- أوزان العاملات بإحدى المصانع .
- الدخل السنوي لمنسوبي مؤسسة معينة .
- وغيره من مثل هذه الأمثلة .

ب – 2) بيانات كمية متقطعة :

وفيها يمكن أن يأخذ المتغير قيمة على (إما ..... أو ...... وليس أي قيمة بينهما) ، وبتعبير آخر هي بيانات يمكن أن تُعد ولا تُقاس ، مثل عدد طلاب الفصول المختلفة في مدرسة ما

والبيانات المنفصلة إما أن تكون نوعية أو كمية متقطعة

ويمكن تقسيم هذا الباب إلى جزئين : الأول ويخص البيانات المنفصلة ، والثاني ويخص البيانات المتصلة

(2–2) تنظيم وعرض البيانات المنفصلة :

ذكرنا في البند السابق أن البيانات المنفصلة إما أن تكون بيانات نوعية أو بيانات كمية متقطعة يأخذ فيها المتغير (الخاصية تحت الدراسة) قيماً محددة ولا يأخذ قيماً موزعة على فترة ، وهذه البيانات يمكن عرضها بطرق مختلفة منها **الجداول** ومنها **الأشكال البيانية** ، ولتوضيح ذلك دعنا نتعامل مع المثال التوضيحي التالي :

مثال (2–1) : قام أحد الباحثين بجمع البيانات التالية عن درجة الطلاب في مادة اللغة الإنجليزية بالثانوية العامة بأحد الفصول **المتميزة** بإحدى المدارس وكانت الدرجات كالتالي (الدرجة العظمى 100) :

100	95	99	99	95	93	94	99	98	92
97	95	94	95	93	96	95	92	95	94

المطلوب تنظيم وعرض النتائج السابقة بطرق مختلفة .

أولاً : العرض الجدولي للبيانات :

البيانات المعطاة في هذا المثال تمثل الخطوة **الأولى** في أي عملية إحصائية وهي عملية "**جمع البيانات**" ، والبيانات هنا معطاة على صورة "**بيانات خام**" أي بيانات كاملة لكن في صورة غير منظمة ، **ولتنظيم** هذه البيانات نحاول تكوين ما يُسمى **بالتوزيع التكراري** [أو **الجدول التكراري**] لهذه البيانات ، ويتم ذلك كالآتي

- **تحديد المدى :**

من المناسب أولاً أن نحدد ما يُسمى **بالمدى** (ونرمز له بالرمز R) وهو "**الفرق بين أكبر قيمة وأقل قيمة**" في البيانات المعروضة ، ويمكن بسهولة ملاحظة أن أكبر قيمة = 100 وأن أقل قيمة = 92 وبالتالي يكون المدى :

$$\boxed{R = 100 - 92 = 8}$$

- **تفريغ البيانات :**

الظاهرة تحت الدراسة هنا هي درجة الطلاب وهي تمثل المتغير هنا (ولنرمز له بالرمز x) ، ثم نقوم بعمل جدول مكون من عمودين [كما هو مبين بجدول (2–1أ)] حيث يمثل العمود الأول (من اليسار) القيم التي يمكن أن يأخذها المتغير x (من أقل قيمة 92 إلى أكبر قيمة 100) ، والعمود الثاني يمثل كيفية تفريغ البيانات السابقة حيث نقوم بالمرور على البيانات السابقة قيمة تلو الأخرى ، وعند مرورنا بقيمة معينة نضع شرطة (علامة) في عمود "تفريغ البيانات أو عمود العلامات" المقابل لتلك القيمة ؛ فمثلاً عند مرورنا بالقيمة (92) نضع في الجدول أمام (92) شرطة رأسية ، ثم عند انتقالنا لقيمة أخرى مثل (98) نضع في الجدول شرطة أمام (98) ، وهكذا حتى ننتهي من جميع البيانات ، مع مراعاة أنه عند تكرار القيمة للمرة

الخامسة نضع الشرطة الخامسة مائلة أو أفقية (كما في حالة القيمة 95) بحيث تُسمى𝍸 **بحزمة** (وعددها خمسة) حتى يسهل حساب التكرار فيما بعد .

جدول (2–1أ) تفريغ البيانات [مثال (2–1)]

تفريغ البيانات (العلامات)	المتغير (الدرجة) x
||	92
||	93
|||	94
𝍸 |	95
|	96
|	97
|	98
|||	99
|	100

- **التوزيع (الجدول) التكراري للبيانات** : نضيف للجدول السابق [جدول (2–1أ)] عموداً ثالثاً [نسميه **عمود التكرارات**] وفيه نقوم بحساب تكرار كل قيمة من القيم [وسنرمز له بالرمز f] وهو عدد العلامات المناظرة لكل قيمة [جدول (2–1ب)] ، يُسمى هذا الجدول الأخير (جدول (2–1ب)] **بالتوزيع (أو الجدول) التكراري** للبيانات وهي أولى طرق عرض البيانات وأسهلها . لاحظ أن مجموع التكرارات [المرموز له بالرمز $\sum f$] يجب أن يساوي عدد الطلاب ، ومن هذا الجدول يمكن الرد على عدد من الأسئلة مثل :

  - عدد الطلبة الحاصلين على درجات أعلى من 94 [**الإجابة : 13 طالباً**]
  - عدد الطلبة الحاصلين على درجات أعلى من 95 وأقل من 99 [**الإجابة : 3 طلاب**] .

وهكذا .

جدول (2–1ب) : التوزيع (الجدول) التكراري للبيانات [مثال (2–1)]

التكرار f	تفريغ البيانات	المتغير (الدرجة) x
2	||	92
2	||	93
3	|||	94
6	𝍸 |	95
1	|	96
1	|	97
1	|	98
3	|||	99
1	|	100
$\sum f = 20$		

$\sum f$ تعني مجموع التكرارات (التي تساوي مجموع الطلاب) وتُقرأ "سيجما **f**"

وإذا أضفنا للجدول السابق [جدول (2-1ب)] عموداً رابعاً [نسميه عمود **التكرارات النسبية** كما هو مبين بجدول (2-1ج)] حيث يتم في هذا العمود الرابع قسمة كل تكرار على مجموع التكرارات فنحصل على **التكرار النسبي** لكل قيمة [ونرمز له بالرمز $\bar{f}$] وهو "**نسبة تكرار تلك القيمة إلى مجموع التكرارات**" ، يُسمى هذا الجدول [(2-1ج)] عندئذٍ **بالتوزيع (أو الجدول) التكراري النسبي** للبيانات ، ويمكن أن يُكتب كل تكرار نسبي كنسبة مئوية وذلك بضرب قيمته في 100 .

جدول (2-1ج) : التوزيع (الجدول) التكراري النسبي للبيانات [مثال (2-1)]

التكرار النسبي $\bar{f}$ $\left[=\frac{f}{\sum f}\right]$	التكرار f	العلامات	المتغير (الدرجة) x
$\frac{2}{20}=0.1$ or 10%	2	||	92
$\frac{2}{20}=0.1$ or 10%	2	||	93
$\frac{3}{20}=0.15$ or 15%	3	|||	94
$\frac{6}{20}=0.3$ or 30%	6	𝍸 |	95
$\frac{1}{20}=0.05$ or 5%	1	|	96
$\frac{1}{20}=0.05$ or 5%	1	|	97
$\frac{1}{20}=0.05$ or 5%	1	|	98
$\frac{3}{20}=0.15$ or 15%	3	|||	99
$\frac{1}{20}=0.05$ or 5%	1	|	100
$\sum \bar{f}=1$ or 100%	$\sum f=20$		

لاحظ أن مجموع التكرارات النسبية $\sum \bar{f}$ يجب أن تساوي الواحد الصحيح [أو 100% كنسبة مئوية] .

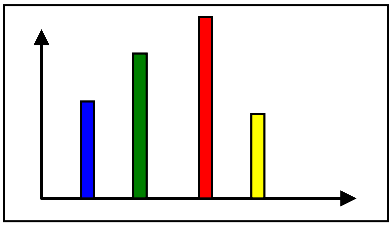
أيضاً لاحظ أن جدول (2-1ج) هو تجميع للجدولين : **التكراري** ، **التكراري النسبي** معاً .

ثانياً : العرض البياني للبيانات :

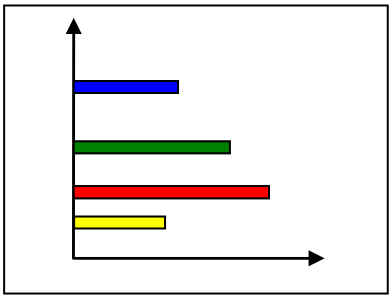
بعد تنظيم البيانات (المنفصلة) في صور جداول (توزيعات) تكرارية يمكن عرضها بيانياً (بالرسم) بطرق مختلفة ، ومن أساسيات عرض أي بيانات بيانياً هو وضوح وبساطة طريقة العرض ولا مانع من أن تكون أيضاً جاذبة .

وهناك طرق مختلفة لعرض البيانات المنفصلة بيانياً ، منها :

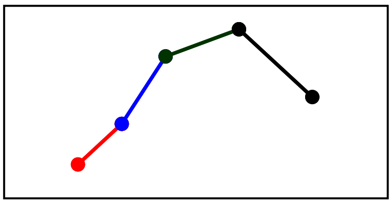
- **الأعمدة البيانية البسيطة** : وهي مجموعة من الأعمدة الرأسية والتي تتناسب ارتفاعاتها مع قيم البيانات التي تمثلها ، ويمكن أن تُستبدل الأعمدة بخطوط رأسية .



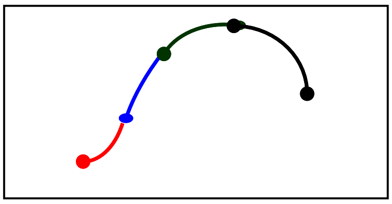
- **القضبان البيانية البسيطة :** وهي مجموعة من القضبان الأفقية والتي تتناسب أطوالها مع قيم البيانات التي تمثلها ، ويمكن أن تُستبدل القضبان بخطوط أفقية .



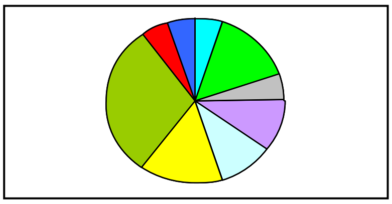
- **الخط البياني :** وهو خط منكسر يصل بين نقاط إحداثياتها الرأسية تتناسب مع قيم البيانات التي تمثلها .



- **المنحنى البياني :** وهو مثل الخط المنكسر ولكن يتم توصيل النقاط بخط ممهد (باليد) بدلاً من استخدام السطرة .



- **اللوحة الدائرية :** حيث تُمثل البيانات بقطاعات دائرية من دائرة نصف قطرها اختياري بحيث تتناسب الزوايا المركزية لهذه القطاعات مع قيم البيانات التي تمثلها .



وفيما يلي عرض بيانات مثال (2-1) بالطرق المختلفة السابقة :

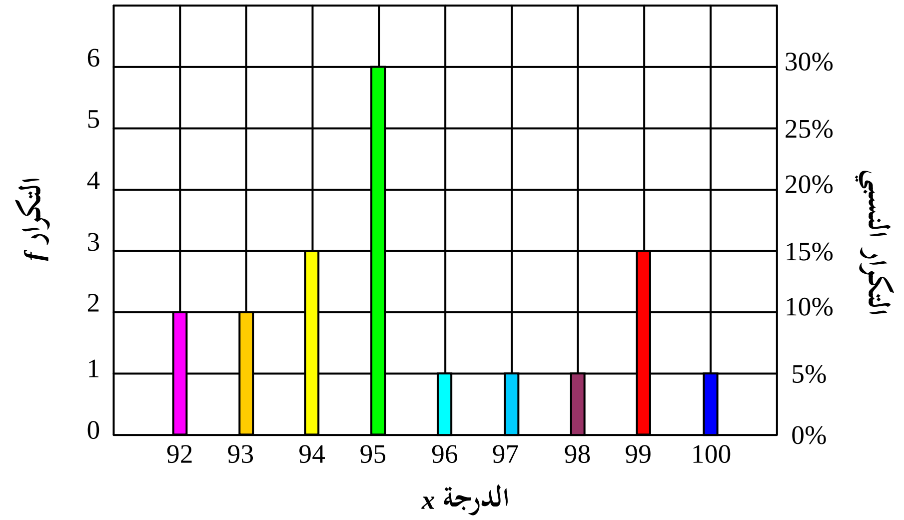
* طريقة الأعمدة البيانية البسيطة :

حيث نقوم برسم محورين متعامدين : الأفقي يمثل قيم المتغير x (الدرجة) والرأسي يمثل التكرار f [أو التكرار النسبي)] ، ثم نرسم عند كل قيمة من قيم المتغير مستطيلاً ارتفاعه يمثل تكرار القيمة (وعرضه غير مهم فنحصل على العرض المبين بشكل (2-1) ، لاحظ في هذا الرسم الآتي :

1. سُمك كل عمود اختياري ومن المناسب أن يكون سمك الأعمدة متساوي حتى يكون العرض البياني متناسقاً وجذاباً [لكن تساوي السمك لكل الأعمدة ليس شرطاً ضرورياً] ، لكن الضروري جداً أن تكون الأعمدة منفصلة عن بعضها لأن التمثيل هنا **لبيانات منفصلة** عن بعضها .
2. مقياس الرسم الخاص بالمحور الرأسي (التكرار أو التكرار النسبي) مهم وضروري [أخذناه في الرسم : كل وحدة رأسية تمثل وحدة تكرار أو 5% تكرار نسبي مئوي] ، في حين لا يهم هنا أخذ مقياس رسم للمحور الأفقي (قيم المتغير) ، حيث من الممكن أن نبدأ من أي مكان ونعتبر هذا الموضع ممثلاً للقيمة الأولى (92) ، ثم أي موضع آخر ليمثل القيمة الثانية (93) ، .. وهكذا ، لكن من

المناسب أن نأخذ مواضع القيم المختلفة للمتغير بحيث تكون مريحة للنظر [تذكر أن العرض البياني يجب أن يكون جذاباً وليس منفراً] .

3. يمكن أن تُستبدل الأعمدة بخطوط رأسية ، وعندئذٍ يُسمى العرض البياني بـ طريقة الخطوط الرأسية .



شكل (2–1) : عرض بيانات مثال (2–1) بيانياً بطريقة الأعمدة البسيطة

* طريقة القضبان البيانية البسيطة :

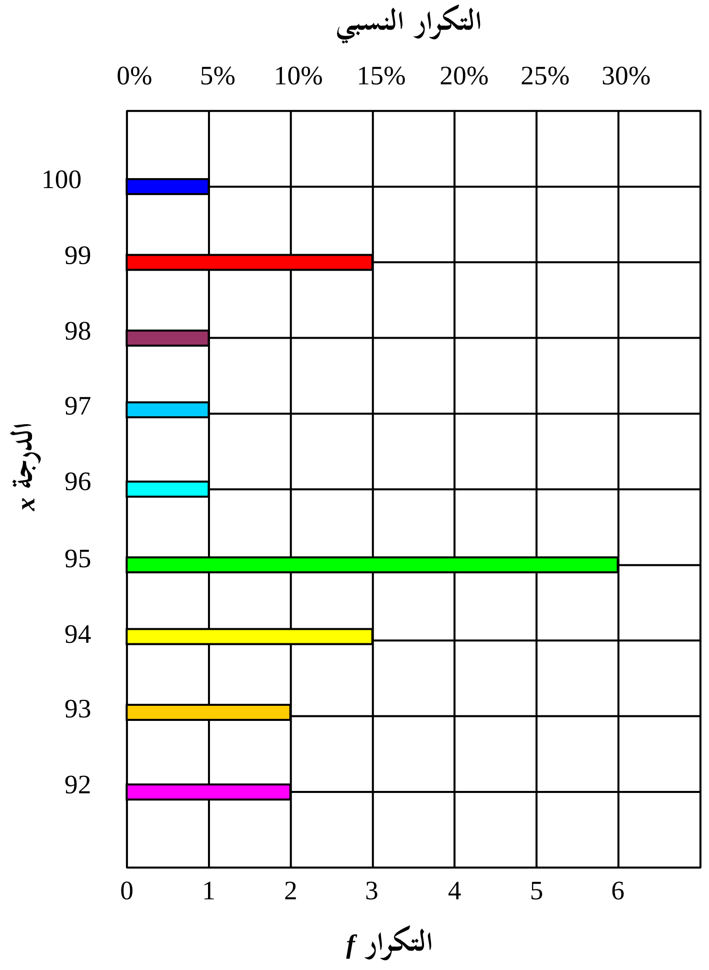
وهي نفس طريقة الأعمدة الرأسية مع تبديل المحورين الأفقي وارأسي معاً بحيث يكون المحور الأفقي هو الممثل للتكرار والمحور الرأسي هو الممثل للدرجة [شكل (2–2)] .

* طريقة الخط البياني والمنحنى البياني :

وفيها نقوم أيضاً برسم نفس المحورين السابقين : الأفقي (ويمثل الدرجة) ، والرأسي (ويمثل التكرار)، وعند كل قيمة من قيم المتغير نتحرك رأسياً بقدر التكرار فنحدد نقطة ، ثم نقوم بتوصيل هذه النقاط معاً بالمسطرة فنحصل على خط منكسر كالمبين بشكل (2–3) ، يُسمى هذا الخط المنكسر **بالخط البياني للبيانات** .

أما إذا قمنا بتوصيل النقاط السابقة باليد بدلاً من المسطرة ، فإننا نحصل على منحنى (بدلاً من خط منكسر [شكل (2–3) أيضاً] ، هذا المنحنى يُسمى **بالمنحنى البياني للبيانات** .

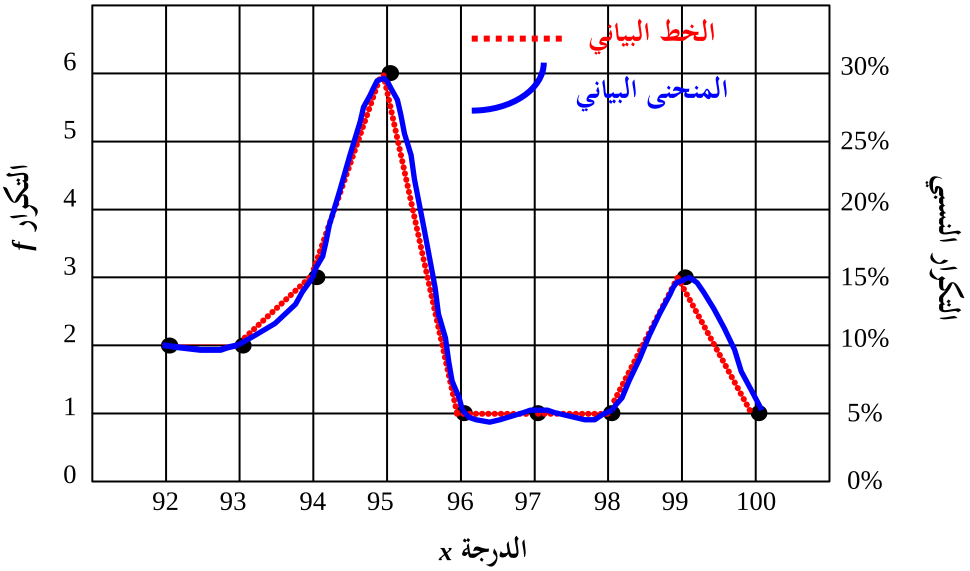
ملحوظة هامة : توصيل النقاط بخط لا يعني أن البيانات متصلة ، بمعنى أن الخط بين النقاط لا يعني أننا نستطيع أن نحسب من هذا الرسم (مثلاً) عدد الطلاب الحاصلين على درجة 95.5 ، أي أن توصيل النقاط بخط هذا لزوم العرض فقط .



شكل (2−2)

عرض بيانات مثال (2−1)

بيانياً بطريقة القضبان البسيطة



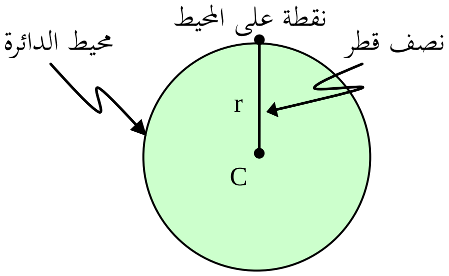
شكل (2−3) : عرض بيانات مثال (2−1) بيانياً بطريقة الخط البياني والمنحنى البياني

* طريقة اللوحة الدائرية [أو ببساطة طريقة الدائرة] :

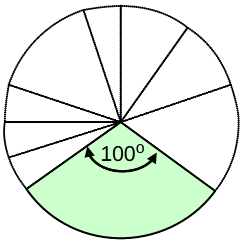
حيث نقوم برسم دائرة (بأي نصف قطر) ثم نحسب لكل درجة ما يُسمى **بقياس زاويتها** ثم نقوم بتمثيل كل قيمة **بقطاع** من هذا الدائرة فنحصل على ما يُسمى بطريقة اللوحة الدائرية للبيانات .

وقبل أن نستكمل عرض البيانات بطريقة اللوحة الدائرية ، دعنا نلقي الضوء على بعض المعلومات الخاصة بهذه الطريقة .

* نفرض أن لدينا دائرة نصف قطرها r ومركزها C ، أي خط يصل بين مركزها وأية نقطة على محيطها يُسمى بنصف قطر للدائرة [وهو الذي طوله r]



* والجزء (المساحة) من الدائرة المحصور بين أي نصفين قطرين فيها ومحيطها يُسمى **بقطاع دائري زاويته المركزية** θ [وتُقرأ "ثيتا"]هي الزاوية المحصورة بين نصفي قطر هذا القطاع . والنسبة بين مساحة أي قطاع من دائرة ومساحة الدائرة كالنسبة بين الزاوية المركزية للقطاع والزاوية المركزية للدائرة [وهي 360^o] .



$$\frac{\text{مساحة القطاع المظلل}}{\text{مساحة الدائرة}} = \frac{100}{360}$$

* نخرج مما سبق بأنه لتمثيل درجة طالب بقطاع دائري يجب أن نحسب أولاً الزاوية المركزية المناظرة لتلك الدرجة ، وطبيعي سوف تختلف الزوايا المركزية للدرجات 100 , 99 , …… , 94 , 93 , 92 طبقاً لتكراراتها ، وحيث أنه في النهاية يجب أن تعطي مساحات القطاعات الدائرية الممثلة للدرجات مساحة الدائرة ، فإن مجموع الزوايا المركزية يجب أن يساوي 360^o ، وهذا يمكننا من حساب الزاوية المركزية لكل درجة من الدرجات بأن نقوم بتقسيم الـ 360^o على مجموع التكرارات ثم نقوم بضرب الناتج في قيمة التكرار الخاص بهذه الدرجة ، أي أن :

$$\text{الزاوية المركزية لقيمة للمتغير} = \frac{360}{\text{مجموع التكرارات}} \times \text{تكرار القيمة}$$

أو بإعادة ترتيب للمعادلة السابقة :

$$\text{الزاوية المركزية لقيمة للمتغير} = \frac{\text{تكرار القيمة}}{\text{مجموع التكرارات}} \times 360$$

* وحيث أن تكرار أي قيمة مقسوماً على مجموع التكرارات يعطي التكرار النسبي لتلك القيمة ، إذن يمكن أيضاً حساب الزاوية المركزية من :

$$\text{الزاوية المركزية لقيمة للمتغير} = \text{التكرار النسبي لتلك القيمة} \times 360$$

نعود الآن للمثال : إذن لعرض البيانات بيانياً عن طريق اللوحة الدائرية ، يجب أولاً أن نحسب الزاوية المركزية المناظرة لكل قيمة من قيم المتغير وذلك من العلاقة :

$$\textbf{الزاوية المركزية لقيمة للمتغير} = \frac{\textbf{تكرار القيمة}}{\textbf{مجموع التكرارات}} \times 360^{o}$$

أو

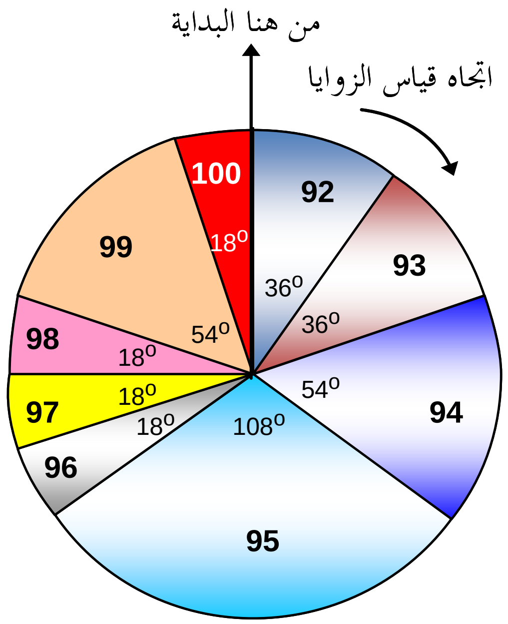
$$\textbf{الزاوية المركزية لقيمة للمتغير} = \textbf{التكرار النسبي لتلك القيمة} \times 360^{o}$$

فنحصل على جدول (2-1) ، ثم نمثل كل قيمة من قيم المتغير x بقطاع من الدائرة زاويته المركزية هي المحسوبة طبقاً لتكرارها فنحصل على شكل (2-4)] .

جدول (2-1) قياس الزوايا المناظرة لبيانات مثال (2-1)

المجموع	100	99	98	97	96	95	94	93	92	x
20	1	3	1	1	1	6	3	2	2	f
360^{o}	18^{o}	54^{o}	18^{o}	18^{o}	18^{o}	108^{o}	54^{o}	36^{o}	36^{o}	الزاوية

لاحظ أن مجموع الزوايا يجب أن يساوي 360^{o}



من نصف قطر اختياري [الرأسي مثلاً ولنعتبره خط البداية] نبدأ برسم الزاوية الأولى 36 الممثلة للدرجة 92 ، ثم الزاوية 36 الممثلة للدرجة 93 ، ثم الزاوية 54 الممثلة للدرجة 94 ، … وهكذا حتى نصل للقطاع الأخير الممثل للدرجة 100 ، فيجب أن تكون زاويته المركزية 18 حيث أن مجموع الزوايا يجب أن يكون مساوياً لـ 360

شكل (2–4) : تمثيل بيانات مثال (2–1) بيانياً بواسطة اللوحة الدائرية (الدائرة)

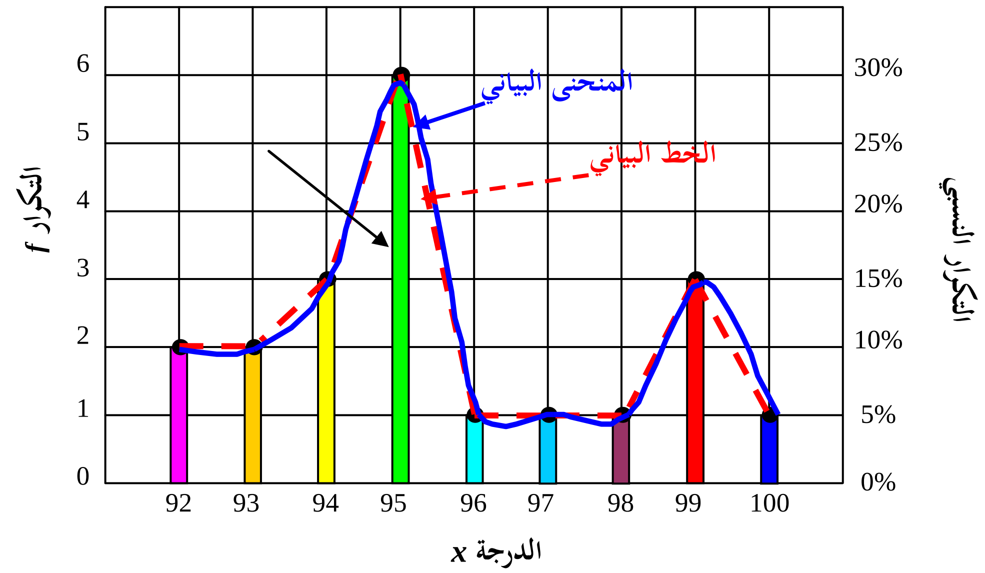
ملاحظات على الطرق المختلفة لعرض البياني (على ضوء المثال السابق) :

الرسومات البيانية (كما ذكرنا سابقاً) وسيلة مفيدة وفعالة لتوضيح وشرح الحقائق الرقمية بأسلوب يسهل فهمه بمجرد النظر ، ويمكن إبداء الملاحظات التالية على الرسومات :

1. تعتبر الأعمدة البيانية من أكثر الرسومات البيانية انتشاراً وهي عبارة عن مستطيلات قواعدها متساوية [أو يُفضل ذلك] وارتفاعاتها تتناسب مع القيم التي تمثلها ، وتكون منفصلة عن بعضها البعض بمسافة يقدرها الباحث .
2. يفضل تمييز الأعمدة عن بعضها بتظليلها أو تخطيطها بواسطة خطوط متوازية أو إبرازها بألوان مختلفة وخاصة عند مقارنة ظواهر مختلفة .
3. يستحسن اختيار مقياس رسم مناسب للمحور الرأسي، ولهذا لابد لمصمم الرسم من التعرف على القيمة الكبرى والقيمة الصغرى لتحديد مقياس الرسم المناسب ، هذا ويجب البدء بالصفر على المحور الرأسي الذي يدل على القيم الرقمية حتى تكون المقارنة سهلة وسليمة وغير مضللة .
4. يفضل عدم كتابة القيم التي تمثلها الأعمدة فوق الأعمدة وذلك لتلافي المبالغة في طول الأعمدة ، وبالتالي تجنب إظهار الرسم مزدحماً او مكتظاً مما ينفر القارئ ، إلا إذا كان ذلك هدفاً في حد ذاته .
5. تصلح الأعمدة البيانية لتمثيل البيانات ذات المتغيرات المنفصلة ، كما تصلح بشكل خاص لتميل البيانات الوصفية (النوعية) [كما سيتضح من المثال التالي] .
6. وتستخدم الدائرة أو اللوحة الدائرية لتمثيل البيانات في الحالات التالية :

- مقارنة الأجزاء المختلفة بالنسبة للمجموع الكلي .
- عندما تكون الأجزاء المقارنة قليلة العدد نسبياً .

7. يمكن الجمع بين أكثر من طريقة للعرض في رسمة واحدة كما هو مبين بشكل (2-5) .



شكل (2-5) : عرض بيانات مثال (2-1) بطرق بيانية مختلفة

مثال (2-2) : في دراسة قام بإجرائها أحد الأطباء لطفل معرض لأحد الأمراض النفسية تم سؤال الطفل عن لون مجموعة من الأشياء ، فكانت إجاباته كما يلي :

أحمر	أزرق	بنفسجي	أحمر	أخضر
أبيض	أبيض	أحمر	أزرق	أبيض
أزرق	أحمر	أخضر	أحمر	بنفسجي
أخضر	أزرق	أبيض	بنفسجي	أحمر

المطلوب تنظيم وعرض النتائج السابقة بطرق مختلفة .

أولاً : العرض الجدولي للبيانات :

هنا يكون المتغير x هو اللون ويأخذ القيم :

الأحمر **الأزرق** **البنفسجي** **الأبيض** **الأخضر**

[ليس المقصود بالقيم هنا قيم عددية ، ولكن المقصود بها "ما يأخذه المتغير من وصف"] ، أما التكرار f فهو عدد مرات ظهور كل قيمة في الإجابات ، وبالتالي يمكن إعداد الجدول التكرارى كما يلى :

جدول (2–2) : التوزيع (الجدول) التكراري (النسبي) للبيانات [مثال (2–2)]

المتغير (اللون) x	العلامات	التكرار f	التكرار النسبي $\bar{f}$ $\left[=\frac{f}{\sum f}\right]$
أحمر	𝍸 𝍷	6	$\frac{6}{20}=0.3$ or 30%
أزرق	𝍷𝍷𝍷𝍷	4	$\frac{4}{20}=0.2$ or 20%
بنفسجي	𝍷𝍷𝍷	3	$\frac{3}{20}=0.15$ or 15%
أبيض	𝍷𝍷𝍷𝍷	4	$\frac{4}{20}=0.2$ or 20%
أخضر	𝍷𝍷𝍷	3	$\frac{3}{20}=0.15$ or 15%
		$\sum f=20$	$\sum \bar{f}=1$ or 100%

ثانياً : العرض البياني للبيانات :

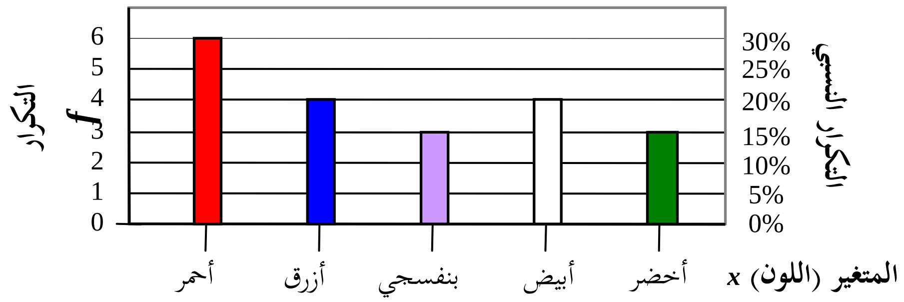
لاحظ هنا أن البيانات هنا بيانات نوعية وهي نوع من البيانات المنفصلة التي يمكن تمثيلها بيانياً بأي طريقة من الطرق التي استعرضناها في المثال السابق ، وقيم المتغير x هنا (اللون) هي :

الأحمر – الأزرق – البنفسجي – الأبيض . والأخضر

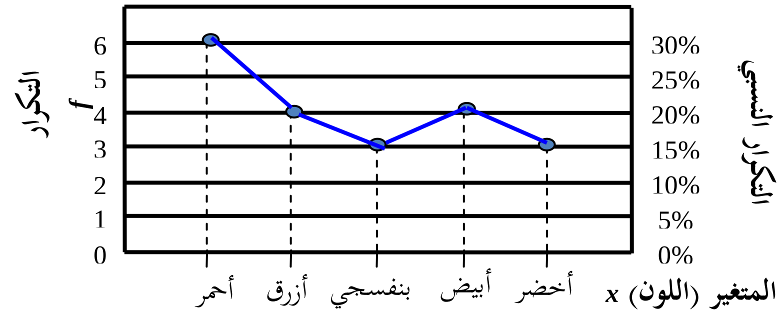
شكل (2–6) يبين طرق مختلفة لتمثيل بيانات هذا المثال حيث يمثل الشكل الأول طريقة عرض البيانات بطريقة الأعمدة البسيطة ، ويمثل الشكل الثاني والثالث طريقة عرض البيانات بطريقة الخط البياني والمنحنى البياني ، أما الشكل الأخير فيوضح طريقة عرض البيانات عن طريق اللوحة الدائرية والتي رُسِمت باستخدام الجدول (2–3) الذي يُحدد قياس الزاوية المركزية لكل قيمة من قيم المتغير x .

جدول (2–3) : حساب الزاوية المركزية المناظرة لكل قيمة من قيم المتغير [مثال (2–2)]

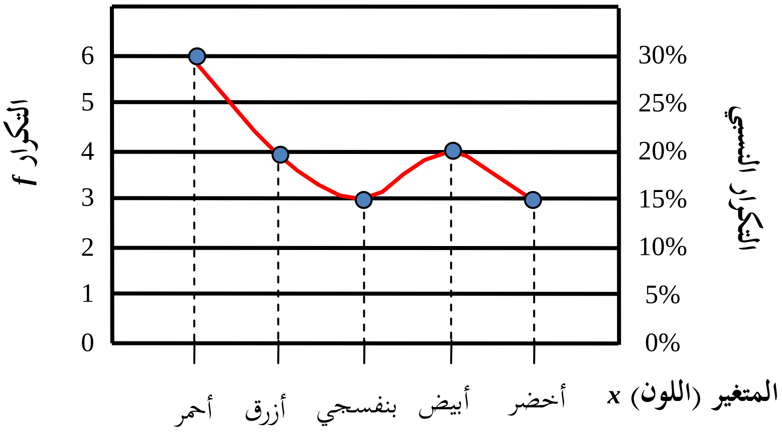
المتغير (اللون) x	التكرار f	قياس الزاوية المركزية
أحمر	6	$\frac{6}{20}\times 360=108^{\circ}$
أزرق	4	$\frac{4}{20}\times 360=72^{\circ}$
بنفسجي	3	$\frac{3}{20}\times 360=54^{\circ}$
أبيض	4	$\frac{4}{20}\times 360=72^{\circ}$
أخضر	3	$\frac{3}{20}\times 360=54^{\circ}$
	$\sum f=20$	مجموع الزوايا = 360^{0}



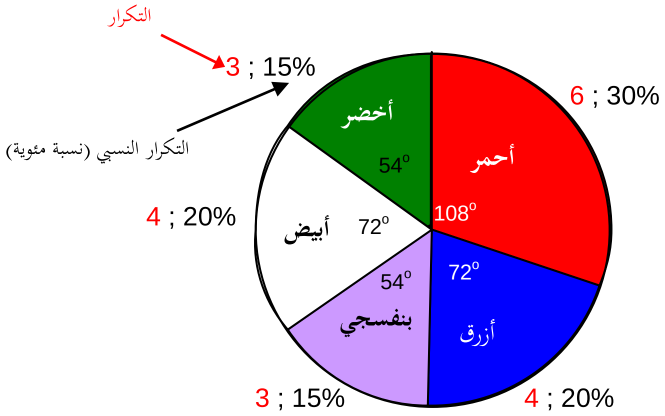
تمثيل البيانات بطريقة الأعمدة البسيطة



تمثيل البيانات بطريقة الخط البياني



تمثيل البيانات بطريقة المنحنى البياني



تمثيل البيانات بطريقة الدائرة

شكل (2–6) : طرق بيانية مختلفة لعرض بيانات مثال (2–2)

مثال (2-3) : تم سؤال عدد من طلاب كليتي الآداب والتربية عن عدد حوادث السيارات التي تعرضوا لها خلال العام الماضي ، فكانت إجاباتهم كما يلي :

3 2 2 1 0 1 2 1 1 1
0 0 1 2 2 1 3 1 0 0
1 2 1 0 2 3 0 0 0 1

المطلوب عرض البيانات السابقة في صورة جدول تكراري وأيضاً بطرق بيانية مختلفة .

أولاً : العرض الجدولي للبيانات :

هنا يكون المتغير x هو عدد الحوادث ويأخذ القيم : $\{0 , 1 , 2 , 3\}$ ، أما التكرار f فهو عدد مرات ظهور كل قيمة في الإجابات ، وبالتالي يمكن إعداد الجدول التكراري كما يلي :

جدول (2-3) : التوزيع (الجدول) التكراري النسبي للبيانات [مثال (2-3)]

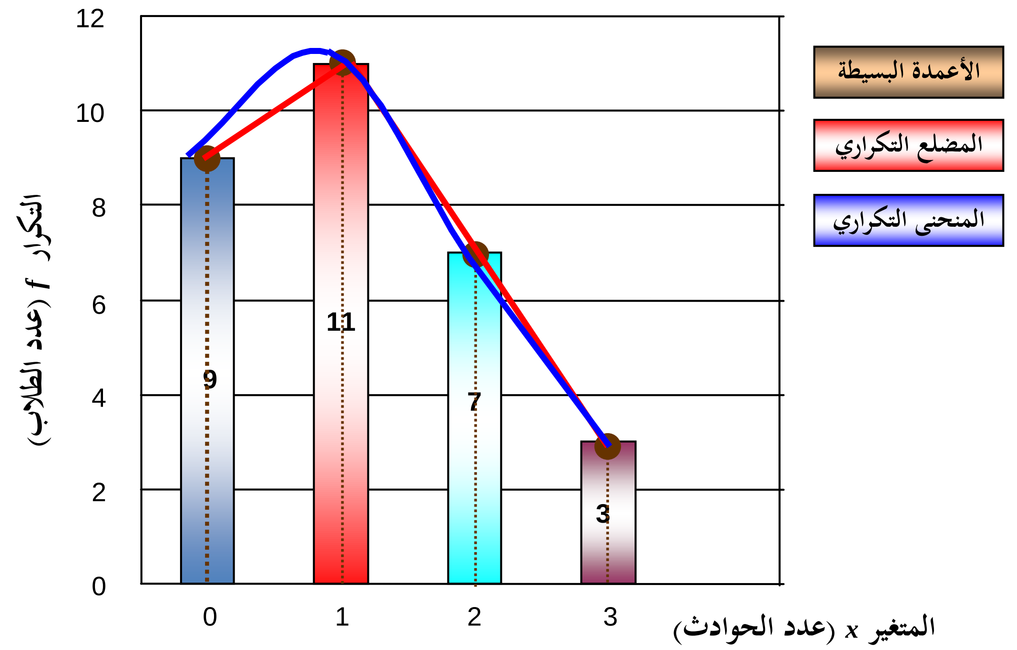
التكرار النسبي $\bar{f} = \left[\frac{f}{\sum f}\right]$	التكرار f	العلامات	المتغير (عدد الحوادث) x
$\frac{9}{30} = 0.3$ or 30%	9	𝍸 ||||	0
$\frac{11}{30} \cong 0.37$ or 37%	11	𝍸 𝍸 |	1
$\frac{7}{30} \cong 0.23$ or 23%	7	𝍸 ||	2
$\frac{3}{30} = 0.1$ or 10%	3	|||	3
$\sum \bar{f} = 1$ or 100%	$\sum f = 30$		

ثانياً : العرض البياني للبيانات :

بنفس الأسلوب المتبع في المثالين السابقين يمكن تمثيل البيانات بطريقة الأعمدة البسيطة والخط البياني والمنحنى البياني واللوحة الدائرية حيث يأخذ المتغير x [هنا عدد الحوادث] القيم : 0 , 1 , 2, 3 بالتكرارات 9 , 11 , 7 , 3 على التوالي ، شكل (2-7) يوضح تمثيل تلك البيانات بواسطة الأعمدة البسيطة ، الخط البياني ، المنحنى البياني (على رسمة واحدة) ، في حين يبين الشكل (2-8) طريقة اللوحة الدائرية لتمثيل هذه البيانات بعد حساب الزوايا المناظرة لكل

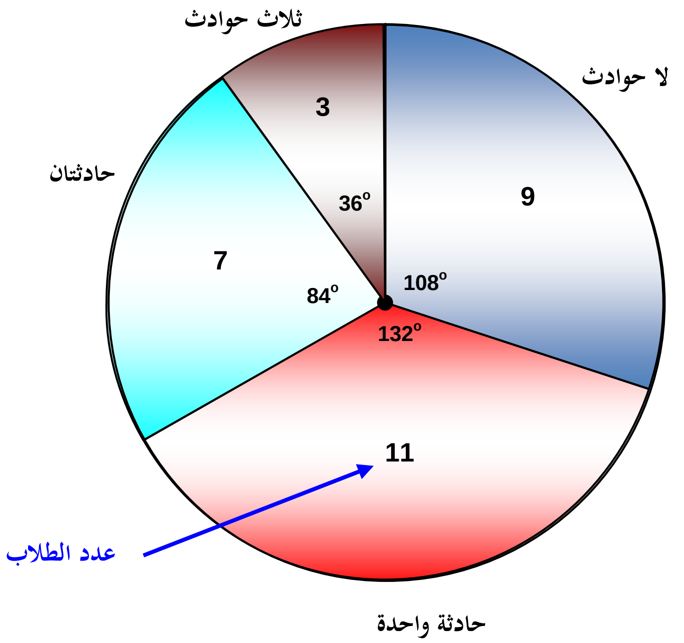
الزاوية المركزية	$\bar{f}$	f	x
(9 ÷ 30) x 360 = 108°	**30%**	**9**	**0**
(11 ÷ 30) x 360 = 132°	**37%**	**11**	**1**
(7 ÷ 30) x 360 = 84°	**23%**	**7**	**2**
(3 ÷ 30) x 360 = 36°	**10%**	**3**	**3**
360°	**100%**	**30**	
↑ مجموع الزوايا	↑ $\sum \bar{f}$	↑ $\sum f$	

قيمة من قيم المتغير x والتي يعطيها الجدول المقابل .



شكل (2–7)

تمثيل بيانات مثال (2–3) بيانياً بواسطة الأعمدة البسيطة ، الخط البياني ، المنحنى البياني



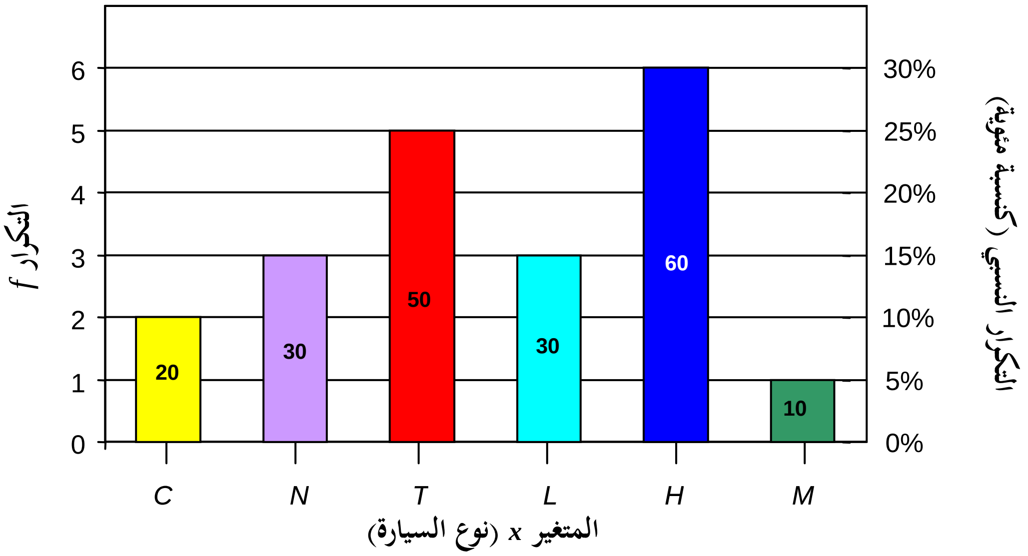
شكل (2–8) : طريقة اللوحة الدائرية لتمثيل بيانات مثال (2–3)

مثال (2–4) : الجدول التالي يبين عدد السيارات الموجودة في أحد المواقف طبقاً لنوع (ماركة) السيارة ، المطلوب عرض هذه البيانات بطرق مختلفة .

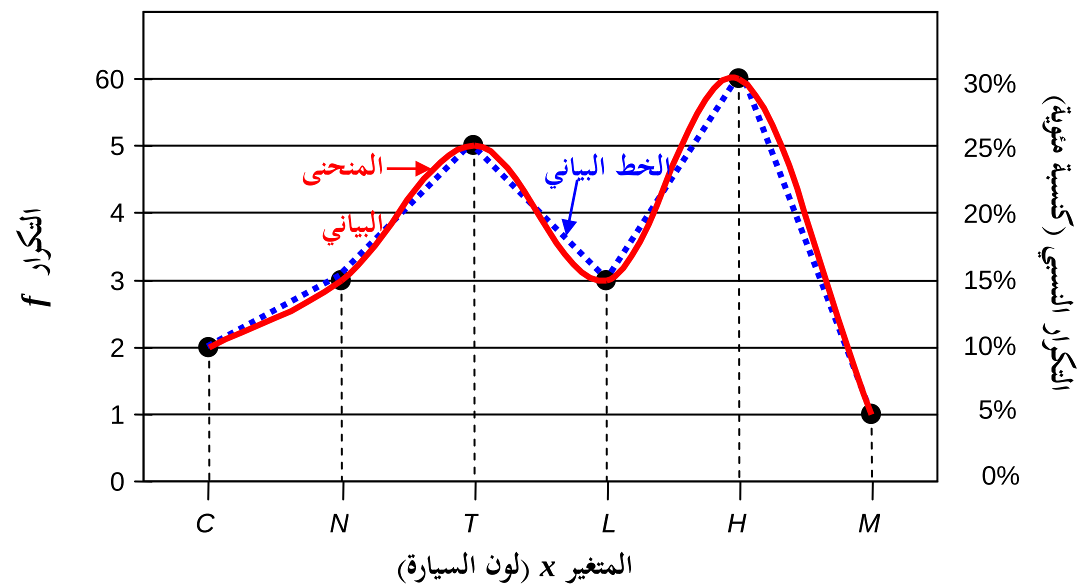
النوع	شيفروليه C	نيسان N	تويوتا T	لانسر L	هيونداي H	مرسيدس M
العدد	20	30	50	30	60	10

شكل (2–9) يبين كلاً من الجدول التكراري وطريقة الأعمدة البسيطة لتمثيل البيانات السابقة ، ويبين شكل (2–10) كلاً من طريقة الخط البياني والمنحنى البياني لتمثيل هذه البيانات ، أم شكل (2–11) فيبين طريقة الدائرة (اللوحة الدائرية) لتمثيل البيانات بعد حساب الزوايا المركزية المناظرة لقيم المتغير المختلفة .

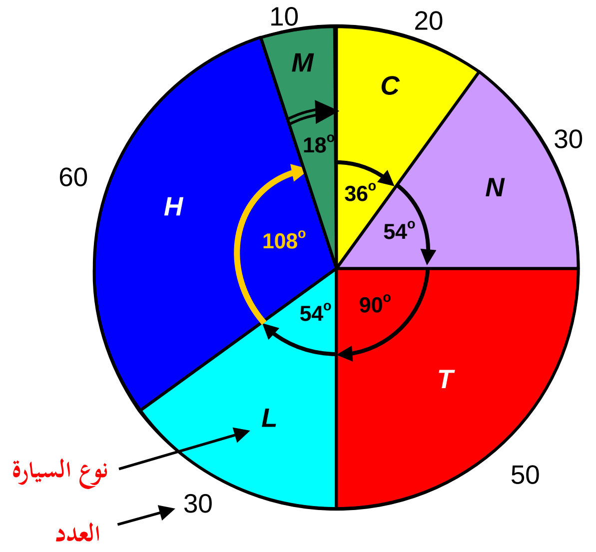
التكرار النسبي	التكرار	x
10%	20	C
15%	30	N
25%	50	T
15%	30	L
30%	60	H
5%	10	M



شكل (2–9) : طريقة الأعمدة البسيطة لتمثيل بيانات مثال (2–4)



شكل (2–10) : طريقة الخط البياني والمنحنى البياني لتمثيل بيانات مثال (2–4)



الزاوية المركزية	التكرار	x
36°	20	C
54°	30	N
90°	50	T
54°	30	L
108°	60	H
18°	10	M

شكل (2–11) : طريقة اللوحة الدائرية لتمثيل بيانات مثال (2–4)

مثال توضيحي (2–5) : في دراسة قامت بها عمادة التعلم الإلكتروني والتعليم عن بُعد بجامعة الملك فيصل عن أعداد الطلاب والطالبات الذين تقدموا لاختبارات نهاية الفصل الدراسي الثاني للعام الجامعي 1431/1430 في تخصصات إدارة الأعمال والآداب والتربية الخاصة كانت البيانات كالتالي :

تخصص إدارة أعمال : 480 (طالبة) ، 1480 (طالب)

تخصص آداب : 2000 (طالبة) ، 3000 (طالب)

تخصص تربية خاصة : 2560 (طالبة) ، 2000 (طالب)

المطلوب عرض هذه البيانات عن طريق جداول مناسبة **وأيضاً** بطرق بيانية مناسبة ، ثم **أجب** على السؤال التالي : ما هي النسبة المئوية للطلاب (الذكور) تخصص آداب بالقياس لـ :

(أ) لجميع الطلبة (طالبات + طلاب) في كل التخصصات

(ب) لجميع الطلاب (ذكور فقط) في كل التخصصات

(ج) لطلبة (طالبات + طلاب) تخصصهم

العرض الجدولي :

يمكن وضع البيانات السابقة في صورة جداول ذات أشكال مختلفة [وهذا يتوقف على الغرض من العرض الجدولي] ، وقبل أن نستعرض الأشكال المختلفة للجداول سنعطي الرموز التالية لتبسيط العرض :

- F : للدلالة على الطالبات (الإناث) [Female]
- M : للدلالة على الطلاب (الذكور) [Male]
- A : للدلالة على جميع الطلبة (طالبات + طلاب) [All]

وبالتالي يمكن وضع البيانات السابقة في جداول مختلفة مثل ما هو مبين حيث يعطي جدول (2-4) تفريغاً مباشراً للبيانات ، بينما يعطي جدول (2-5) تفريغاً للبيانات مع إضافة عمود يبين العدد الكلي لطلبة كل تخصص (طالبات + طلاب) ، ويعطي جدول (2-6) تفريغاً للبيانات مع إضافة صف يوضح العدد الكلي لكلٍ من الطالبات (على حده) والطلاب (على حده) في جميع التخصصات ، أما جدول (2-7) فهو تجميع لكل بيانات الجداول الأخرى .

جدول (2–4)

	F	M
إدارة أعمال	480	1480
آداب	2000	3000
تربية خاصة	2560	2000

جدول (2–5)

	F	M	A
إدارة أعمال	480	1480	1960
آداب	2000	3000	5000
تربية خاصة	2560	2000	4560

جدول (2–6)

	F	M
إدارة أعمال	480	1480
آداب	2000	3000
تربية خاصة	2560	2000
المجموع	5040	6480

جدول (2–7)

	F	M	A
إدارة أعمال	480	1480	1960
آداب	2000	3000	5000
تربية خاصة	2560	2000	4560
المجموع	5040	6480	11520

ومن هذا الجدول الأخير [جدول (٢–٧)] يمكن الرد على السؤال المعطى في المثال قبل أن نستعرض العرض البياني للبيانات بالطرق المختلفة كالآتي :

لحساب النسبة المئوية لأي متغير ، يجب تحديد قيمتين : القيمة الأولى وهي قيمة المتغير ، والثانية وهي الكمية المنسوب إليها ، فتكون النسبة المئوية المطلوبة هي :

$$\text{النسبة المئوية للمتغير } \% = \frac{\text{قيمة المتغير}}{\text{القيمة المنسوب لها}} \times 100$$

هنا المطلوب حساب النسبة المئوية للطلاب (الذكور) تخصص آداب [وعددهم 3000] :

(أ) بالقياس لجميع الطلاب (ذكور وإناث) في كل التخصصات [وعددهم 11520] ، فتكون النسبة المئوية المطلوبة هي :

$$\frac{3000}{11520} \times 100 \cong 26\%$$

(ب) بالقياس لجميع الطلاب (ذكور فقط) في كل التخصصات [وعددهم 6480] ، فتكون النسبة المئوية المطلوبة هي :

$$\frac{3000}{6480} \times 100 \cong 46.3\%$$

(ج) بالقياس لطلبة (ذكور وإناث) تخصص آداب [وعددهم 5000] ، فتكون النسبة المئوية المطلوبة هي :

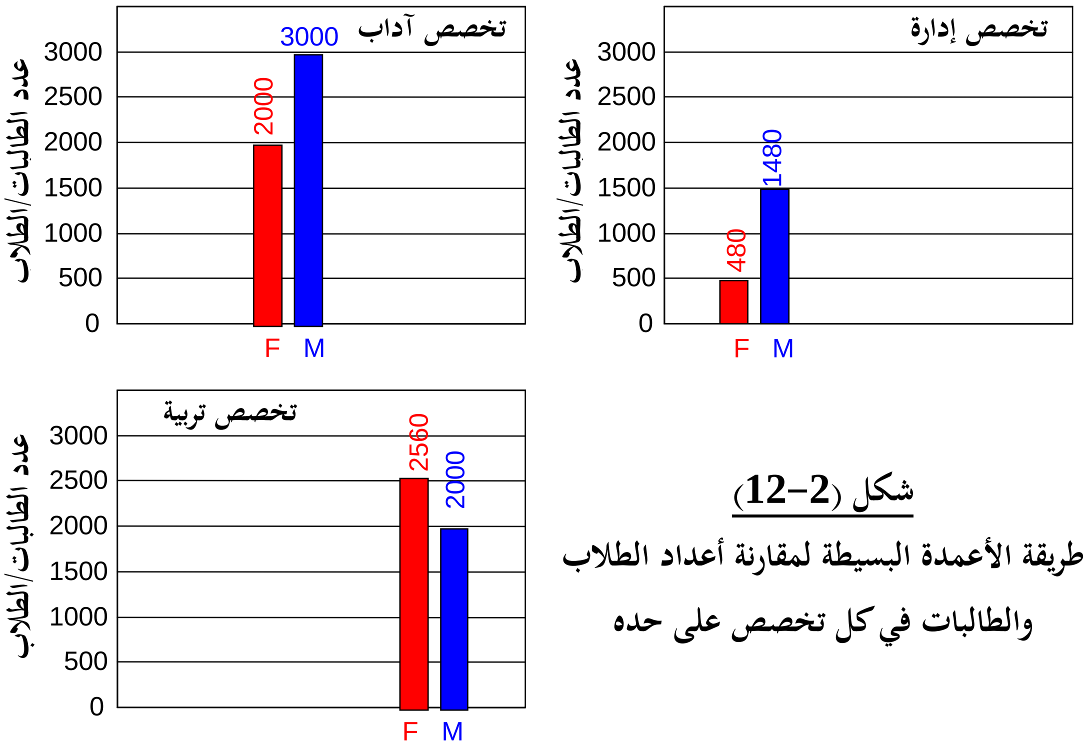
$$\frac{3000}{5000} \times 100 = 60\%$$

العرض البياني :

يمكن أن نقوم بعرض هذه البيانات بيانياً بطرق مختلفة [وهذا أيضاً يتوقف على الغرض من العرض] ، ومن هذه الطرق الآتي :

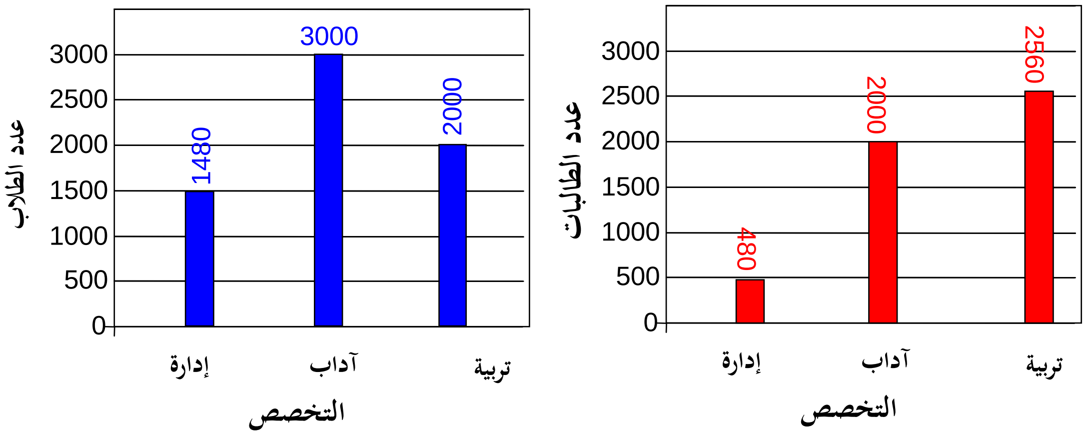
- باستخدام جدول (2-4) يمكن أن نقوم بعرض أعداد الطالبات والطلاب لكل تخصص من التخصصات على حده في ثلاثة رسومات منفصلة باستخدام طريقة الأعمدة البسيطة (مثلاً) كما هو مبين بشكل (2-12) .
- باستخدام جدول (2-4) أيضاً يمكننا عرض بيانات الطالبات في كل التخصصات على رسمة ، وبيانات الطلاب على رسمة أخرى كما هو مبين بشكل (2-13) .

وهنا يتبادر إلى الذهن السؤال التالي : **"أليس من الممكن تجميع الرسومات السابقة في رسمة واحدة ؟"** ، الإجابة : **نعم** ، وذلك عن طريق **الأعمدة المزدوجة** [شكل (2-14)] أو **الأعمدة المجزأة** [شكل (2-15)] .



شكل (2-12)

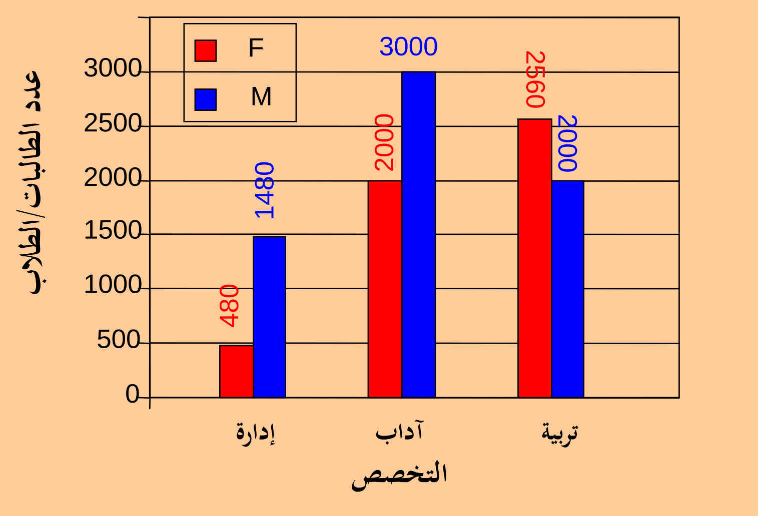
طريقة الأعمدة البسيطة لمقارنة أعداد الطلاب والطالبات في كل تخصص على حده



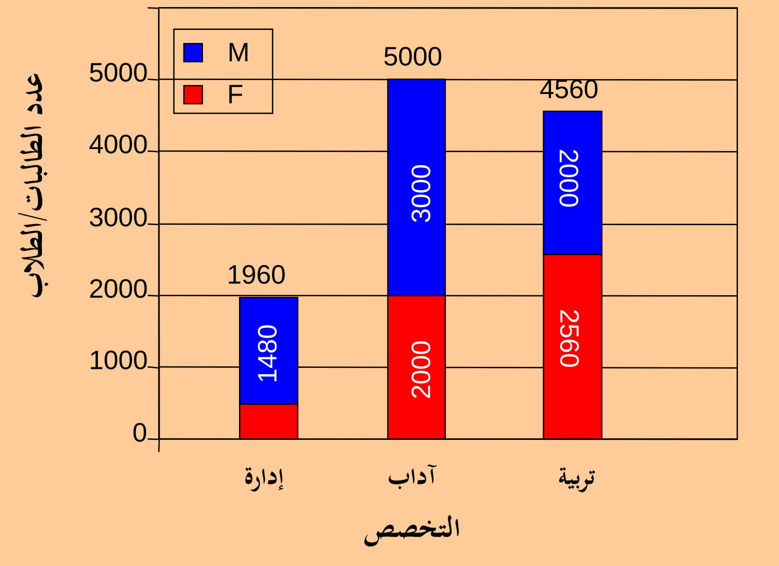
شكل (2–13)

طريقة الأعمدة البسيطة لمقارنة أعداد الطلاب (على حده) والطالبات (على حده) في جميع التخصصات

ففي شكل (2–14) تم الاستفادة من جدول (2–4) في تمثيل كل تخصص بعمود **مزدوج** مكون من عمودين بسيطين متلاصقين يمثلا أعداد الطالبات والطلاب في هذا التخصص ، أما في شكل (2–15) فإن كل تخصص يُمثل بعمود واحد طوله يعبر عن العدد الكلي للطلبة (طالبات + طلاب) بهذا التخصص ثم يتم **تجزئته** إلى عمودين كل واحد منهما يمثل فئة من الفئات [وهنا نحتاج للجدول (2–5)] .



شكل (2-14) : طريقة الأعمدة المزدوجة

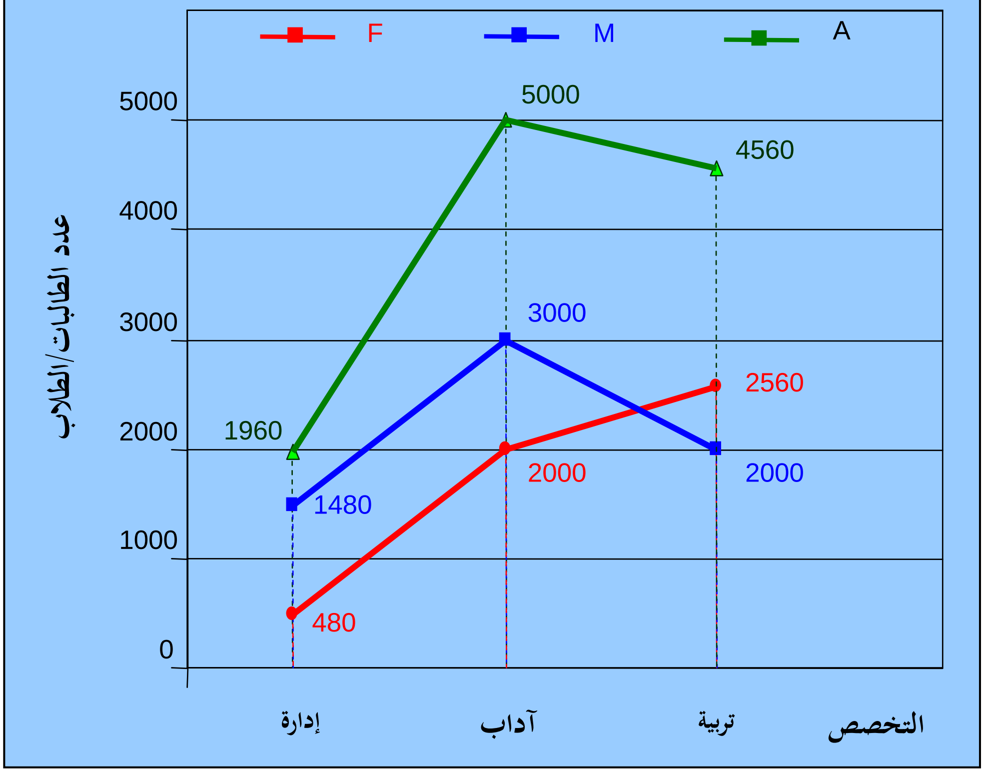


شكل (2-15) : طريقة الأعمدة المجزأة

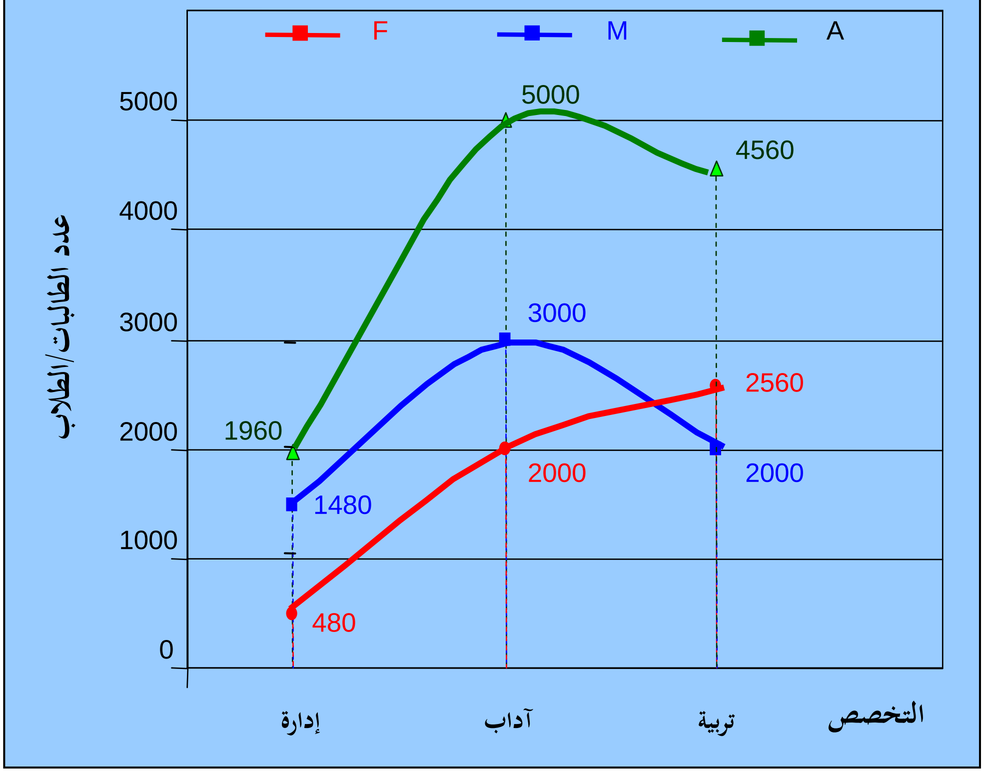
أيضاً نود التنويه أنه يمكن تمثيل جميع البيانات السابقة بطريقة الخط البياني [شكل (2-16)] أو المنحنى البياني [شكل (2-17)] وهنا نحتاج لجدول (2-5)] .

كما نود أن ننبه أيضاً إلى أنه يمكن تمثيل جميع البيانات بطريقة **اللوحة الدائرية** [أو الدائرة] ، وهنا يمكن أن نتعامل مع العرض بأكثر من طريقة [كما في حالة الأعمدة] ، من هذه الطرق أن نقوم برسم دائرة لكل تخصص على حده كما هو موضح بشكل (2-18) ، أيضاً يمكن العرض باستخدام طريقة الدائرة وذلك برسم دائرتين : الأولى خاصة بطالبات جميع التخصصات والأخرى خاصة بطلاب جميع التخصصات كما هو موضح بشكل (2-19) ، مع مراعاة كيفية حساب الزوايا المركزية في كل حالة .

شكل (2–16) : طريقة الخط البياني

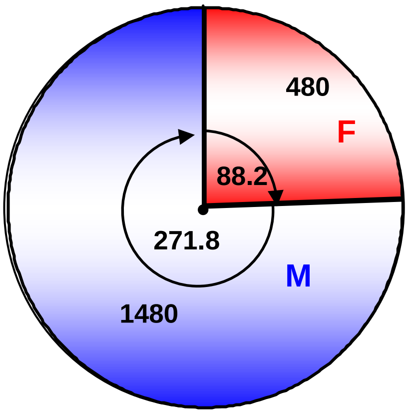


شكل (2–17) : طريقة المنحنى البياني



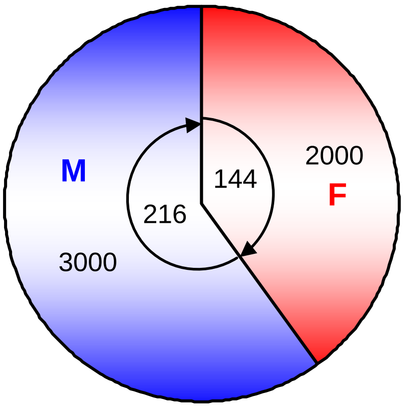
تخصص إدارة	F	M	A
العدد (التكرار)	480	1480	1960
الزاوية المركزية	88.2°	271.8°	360°

لاحظ أن قيم الزوايا المركزية هنا مبني على أساس العدد الكلي لطلبة تخصص إدارة أعمال (طالبات + طلاب)



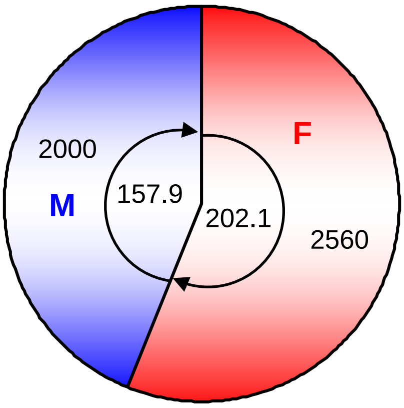
تخصص آداب	F	M	A
العدد (التكرار)	2000	3000	5000
الزاوية المركزية	144°	216°	360°

لاحظ أن قيم الزوايا المركزية هنا مبني على أساس العدد الكلي لطلبة تخصص آداب (طالبات + طلاب)



تخصص تربية	F	M	A
العدد (التكرار)	2560	2000	4560
الزاوية المركزية	202.1°	157.9°	360°

لاحظ أن قيم الزوايا المركزية هنا مبني على أساس العدد الكلي لطلبة تخصص تربية (طالبات + طلاب)



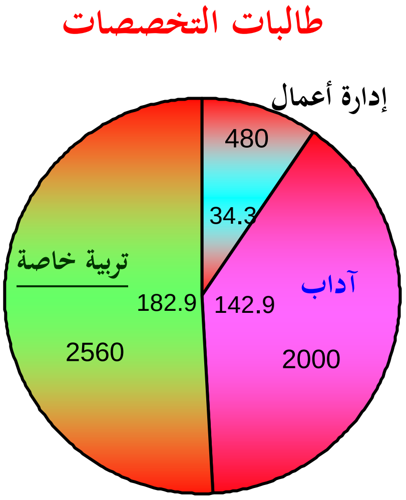
شكل (2–18) : طريقة اللوحة الدائرية لتمثيل أعداد طلبة كل تخصص على حده

وهنا يتبادر إلى الذهن السؤال التالي [وهو مشابه للسؤال الذي سألناه عند تعرضنا لطرق الأعمدة المزدوجة والمجزأة] ، ألا وهو "أليس من الممكن تجميع الرسومات السابقة في دائرة واحدة ؟" ، الإجابة **نعم** ، لكن لابد أن ننتبه إلى أن الزوايا المركزية هنا يجب أن تُحسب على أساس العدد الكلي للطلبة (طالبات + طلاب كل التخصصات) ، وبالتالي يجب تكوين جدول (2–8) التالي أولاً وهو يعتمد كليةً على جدول (2–7) . في هذا الجدول (2–8) لاحظ أن الأعداد الموجودة داخل المستطيلات تمثل التكرار [أعداد الطالبات والطلاب] ، أما

الأعداد الموجودة داخل الدوائر فتمثل قيم الزوايا المركزية المناظرة لتلك التكرارات والمحسوبة على أساس العدد الكلي للطلبة في جميع التخصصات [أي العدد 11520] .

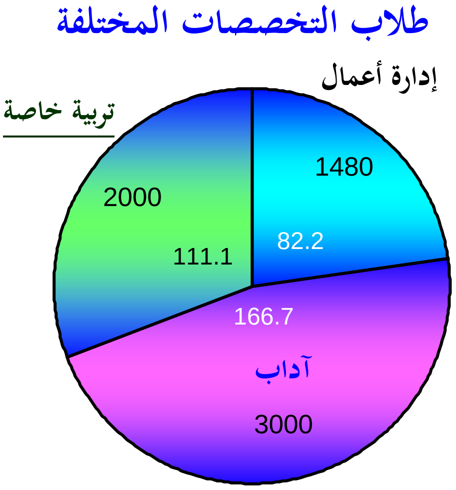
الطالبات F	العدد (التكرار)	الزاوية المركزية
تخصص إدارة	480	34.3^{o}
تخصص آداب	2000	142.9^{o}
تخصص تربية	2560	182.9^{o}
المجموع	**5040**	**$\cong 360^{o}$**

لاحظ أن قيم الزوايا المركزية هنا مبني على أساس العدد الكلي
لطالبات جميع التخصصات



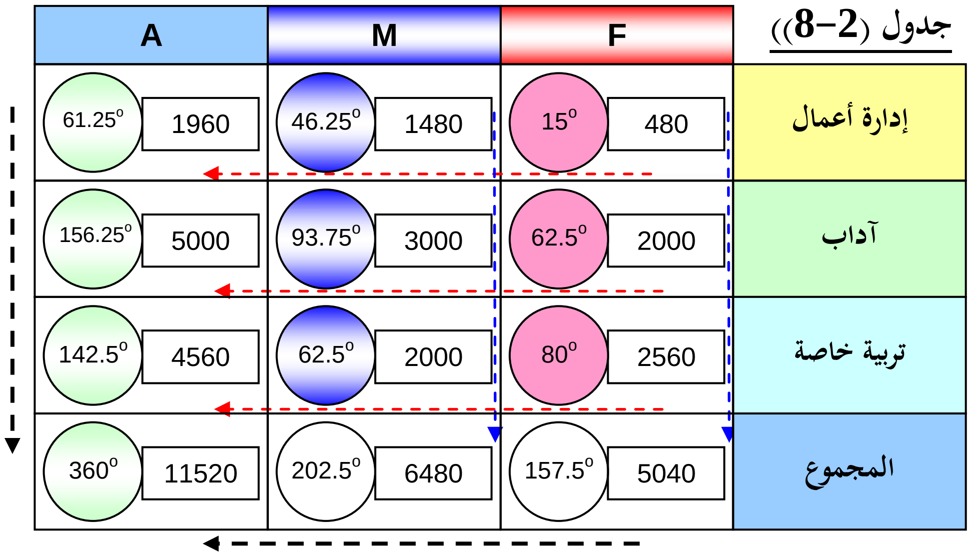
الطلاب M	العدد (التكرار)	الزاوية المركزية
تخصص إدارة	1480	82.2^{o}
تخصص آداب	3000	166.7^{o}
تخصص تربية	2000	111.1^{o}
المجموع	**6480**	360^{o}

لاحظ أن قيم الزوايا المركزية هنا مبني على أساس العدد الكلي
لطلاب جميع التخصصات



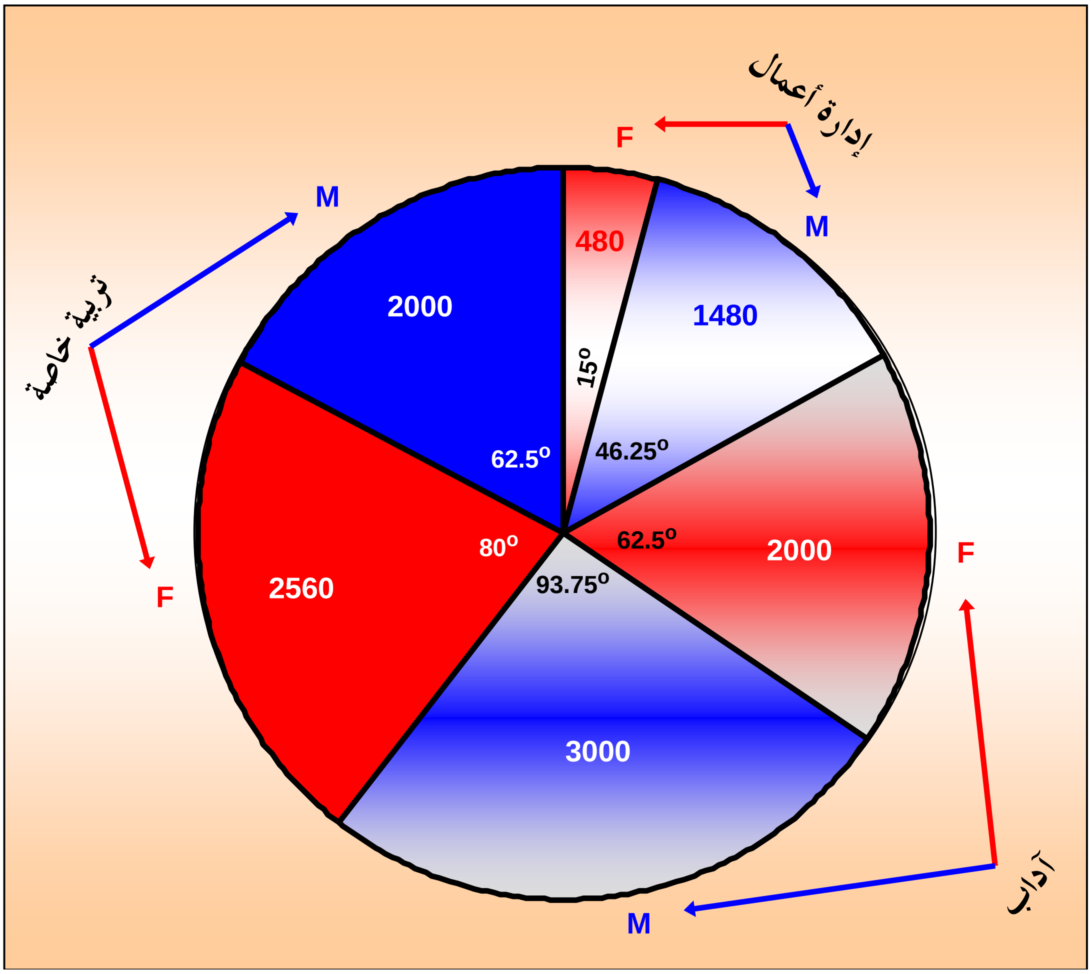
شكل (2–19) : طريقة اللوحة الدائرية لتمثيل أعداد كلٍ من الطالبات (على حده) والطلاب (على حده) في جميع التخصصات

جدول (8–2)



	F	M	A
إدارة أعمال	480 15°	1480 46.25°	1960 61.25°
آداب	2000 62.5°	3000 93.75°	5000 156.25°
تربية خاصة	2560 80°	2000 62.5°	4560 142.5°
المجموع	5040 157.5°	6480 202.5°	11520 360°

ومن الجدول (2–8) يمكن تمثيل البيانات السابقة عن طريق لوحة دائرية واحدة كما هو مبين بشكل (20–2)



شكل (20–2) : عرض بيانات مثال (5–2) بلوحة دائرية واحدة

ملخص للدرس الثاني [الباب الثاني : التوزيعات التكرارية (بيانات منفصلة)]

- **البيانات المنفصلة :** هي بيانات إما أن تكون بيانات نوعية [تلك البيانات التي لا يمكن التعبير عن متغيرها بعدد] أو بيانات كمية متقطعة [تلك البيانات التي يأخذ فيها المتغير قيماً عددية معينة دون أي قيمة بينها ، أي بيانات كمية تُعد ولا تُقاس] .
- **التكرار النسبي** $\bar{f}$ لأي قيمة : هو خارج قسمة تكرار تلك القيمة على مجموع التكرارات ويمكن أن يُوضع كنسبة عادية أو نسبة مئوية [بضرب النسبة العادية في 100] .

$$\text{التكرار النسبي لقيمة ما} = \frac{\text{تكرار القيمة}}{\text{مجموع التكرارات}}$$

- **مجموع التكرارات النسبة** $\sum \bar{f}$: يجب أن يساوي 1 [أو 100%]
- **الزاوية المركزية المناظرة لقيمة معينة لـ x :** نقوم بقسمة تكرار القيمة على مجموع التكرارات ثم نضرب الناتج في 360 [أو نضرب التكرار النسبي (كنسبة) في 360] .

$$\text{الزاوية المركزية لقيمة ما} = \frac{\text{تكرار القيمة}}{\text{مجموع التكرارات}} \times 360$$

$$\text{الزاوية المركزية لقيمة ما} = \text{التكرار النسبي للقيمة} \times 360$$

- **مجموع الزوايا المركزية :** يجب أن يساوي 360
- **المدى R لمجموعة من البيانات :** هو الفرق بين أكبر قيمة في البيانات وأصغر قيمة فيها
- **عرض البيانات المنفصلة بيانياً :**

طريقة الأعمدة البسيطة : حيث تُمثل كل قيمة من قيم المتغير بعمود (خط رأسي) طوله يُعبر عن تكرار تلك القيمة [لا يهم عرض الأعمدة أو المسافات بينها ولكن المهم جداً أن تكون الأعمدة منفصلة عن بعضها] .

طريقة القضبان البسيطة : حيث تُمثل كل قيمة من قيم المتغير بقضيب (خط أفقي) طوله يُعبر عن تكرار تلك القيمة [لا يهم سمك القضبان أو المسافات بينها ولكن المهم جداً أن تكون القضبان منفصلة عن بعضها] .

طريقة الخط البياني : حيث تُمثل كل قيمة من قيم المتغير وتكرارها بنقطة ثم نقوم بتوصيل هذه النقاط بخط منكسر (بواسطة المسطرة)

طريقة المنحنى البياني : حيث تُمثل كل قيمة من قيم المتغير وتكرارها بنقطة ثم نقوم بتوصيل هذه النقاط بخط ممهد (باليد)

طريقة اللوحة الدائرية (أو الدائرة) : حيث تُمثل كل قيمة من قيم المتغير بقطاع من دائرة وذلك طبقاً لتكرارها

تدريب (2)

الإجابة النهائية لجميع التمرينات موجودة في نهاية التدريب

اختر الإجابة الصحيحة

(1) البيانات المنفصلة هي :

(أ) بيانات نوعية فقط　　(ب) بيانات كمية متقطعة فقط

(ج) أي بيانات كمية يمكن أن تُقاس　　(د) بيانات نوعية أو كمية متقطعة

(2) البيانات المتصلة هي :

(أ) بيانات نوعية فقط　　(ب) بيانات كمية متقطعة فقط

(ج) أي بيانات كمية يمكن أن تُقاس　　(د) بيانات نوعية أو كمية متقطعة

(3) المدى R يمكن تحديده لـ :

(أ) البيانات النوعية فقط　　(ب) البيانات الكمية المتقطعة فقط

(ج) أي بيانات كمية　　(د) أي بيانات

(4) المدى R لمجموعة من البيانات هو :

(أ) أكثر القيم تكراراً في البيانات　　(ب) أكبر قيمة في البيانات

(ج) أصغر قيمة في البيانات　　(د) الفرق بين أكبر وأصغر قيمة من البيانات

(5) المدى R لمجموعة القيم 2 , 10 , 4 , 5 , 5 , 7 هو :

(أ) 5　　(ب) 8　　(ج) 2　　(د) 10

(6) التكرار النسبي $\bar{f}$ لأي قيمة في مجموعة من القيم هو :

(أ) خارج قسمة القيمة على مجموع القيم .

(ب) خارج قسمة تكرار القيمة على مجموع التكرارات .

(ج) خارج قسمة مجموع التكرارات على تكرار القيمة

(د) خارج قسمة القيمة على مجموع التكرارات .

(7) الزاوية المركزية لأي قيمة في مجموعة من القيم هو :

(أ) (القيمة ÷ مجموع القيم) × 360　　(ب) تكرار القيمة × 360

(ج) تكرار القيمة ÷ 360　　(د) التكرار النسبي للقيمة × 360

(8) في طريقة **الأعمدة البسيطة** لعرض البيانات المنفصلة تُمثل كل قيمة من قيم المتغير x بـ :

(أ) بعمود (خط رأسي) طوله يُعبر عن تكرار تلك القيمة .

(ب) بقضيب (خط أفقي) طوله يُعبر عن تكرار تلك القيمة .

(ج) بنقطة إحداثياتها هي قيمة المتغير وتكرارها ثم نقوم بتوصيل هذه النقاط بخط منكسر (بواسطة المسطرة)

(د) بقطاع من دائرة طبقاً لتكرارها .

(9) في طريقة **القضبان البسيطة** لعرض البيانات المنفصلة تُمثل كل قيمة من قيم المتغير x بـ :

(أ) بعمود (خط رأسي) طوله يُعبر عن تكرار تلك القيمة .

(ب) بقضيب (خط أفقي) طوله يُعبر عن تكرار تلك القيمة .

(ج) بنقطة إحداثياتها هي قيمة المتغير وتكرارها ثم نقوم بتوصيل هذه النقاط بخط منكسر (بواسطة المسطرة)

(د) بقطاع من دائرة طبقاً لتكرارها .

(10) في طريقة **الخط البياني** لعرض البيانات المنفصلة تُمثل كل قيمة من قيم المتغير x بـ :

(أ) بعمود (خط رأسي) طوله يُعبر عن تكرار تلك القيمة .

(ب) بقضيب (خط أفقي) طوله يُعبر عن تكرار تلك القيمة .

(ج) بنقطة إحداثياتها هي قيمة المتغير وتكرارها ثم نقوم بتوصيل هذه النقاط بخط منكسر (بواسطة المسطرة)

(د) بقطاع من دائرة طبقاً لتكرارها .

(11) في طريقة **المنحنى البياني** لعرض البيانات المنفصلة تُمثل كل قيمة من قيم المتغير x بـ :

(أ) بعمود (خط رأسي) طوله يُعبر عن تكرار تلك القيمة .

(ب) بقضيب (خط أفقي) طوله يُعبر عن تكرار تلك القيمة .

(ج) بنقطة إحداثياتها هي قيمة المتغير وتكرارها ثم نقوم بتوصيل هذه النقاط بخط ممهد (باليد)

(د) بقطاع من دائرة طبقاً لتكرارها .

(12) في طريقة **الدائرة** لعرض البيانات المنفصلة تُمثل كل قيمة من قيم المتغير x بـ :

(أ) بعمود (خط رأسي) طوله يُعبر عن تكرار تلك القيمة .

(ب) بقضيب (خط أفقي) طوله يُعبر عن تكرار تلك القيمة .

(ج) بنقطة إحداثياتها هي قيمة المتغير وتكرارها ثم نقوم بتوصيل هذه النقاط بخط منكسر (بواسطة المسطرة)

(د) بقطاع من دائرة طبقاً لتكرارها .

خاص بالأسئلة من (13) إلى (18) : الجدول التالي يبين الجدول التكراري لأعمار 10 ممرضات تعملن في أحد أقسام إحدى المستشفيات ، من هذا الجدول :

خاص بالمسائل من 13 إلى 18

التكرار f	المتغير (العمر) x
2	22
3	25
2	28
1	31
1	32
1	35
$\sum f$	

(13) مجموع التكرارات $\sum f$ تساوي :

(أ) 3 (ب) 2 (ج) 10 (د) 18

(14) المدى R للعمر هو :

(أ) 3 (ب) 2 (ج) 10 (د) 13

(15) زاوية القياس المناظرة للعمر 31 تساوي :

(أ) 36^o (ب) 360^o (ج) 72^o (د) 108^o

(16) التكرار النسبي للعمر "25 سنة" هو :

(أ) *0.2* (ب) 0.3 (ج) 0.1 (د) 1

(17) عدد الممرضات اللاتي يزيد أعمارهن عن 32 سنة هو :

(أ) 1 (ب) 2 (ج) 3 (د) 5

(18) النسبة المئوية للممرضات اللاتي أعمارهن *31* سنة فأقل هي :

(أ) 0.8 (ب) 0.7 (ج) 70% (د) 80%

خاص بالأسئلة من (19) إلى (25) : الجدول التكراري المعطى يبين عدد السيارات الموجودة في أحد المواقف طبقاً لنوع السيارة $[C , N , T , L , H , M]$

(19) شكل (1) يبين طريقة .... لتمثيل هذه البيانات بيانياً .

(أ) الخط البياني (ب) المنحنى البياني (ج) الأعمدة البسيطة (د) الدائرة

(20) بينما شكل (2) يبين طريقة .... لتمثيل هذه البيانات بيانياً .

(أ) الخط البياني (ب) المنحنى البياني (ج) الأعمدة البسيطة (د) الدائرة

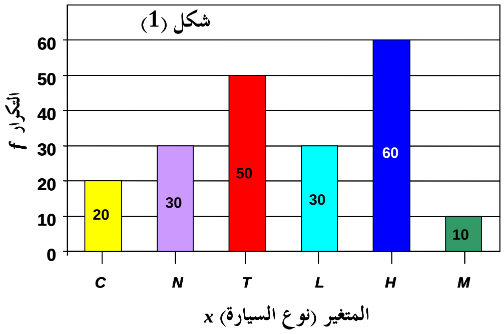
(21) شكل (3) يبين طريقة .... لتمثيل هذه البيانات بيانياً .

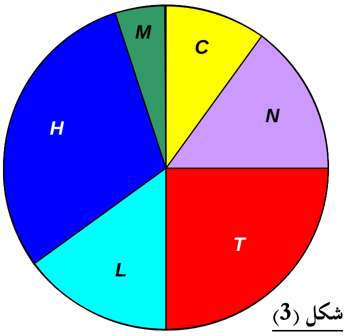
(أ) الخط البياني (ب) المنحنى البياني (ج) الأعمدة البسيطة (د) الدائرة

خاص بالمسائل من 19 إلى 25

الجدول التكراري	
التكرار	*x*
20	***C***
30	***N***
50	***T***
30	***L***
60	***H***
10	***M***

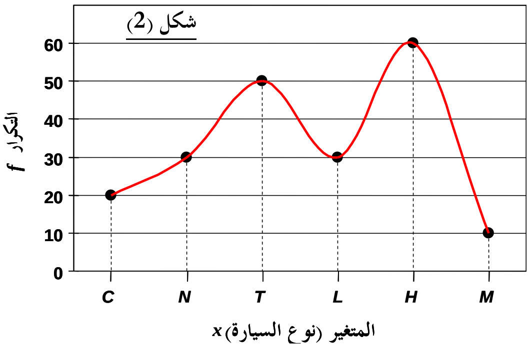
شكل (1)





شكل (3)

شكل (2)



(22) عدد السيارات الموجودة بالموقف هو :

(أ) 100 (ب) 150 (ج) 200 (د) 250

(23) التكرار النسبي للسيارات من النوع *C* هو :

(أ) 10 (ب) 10% (ج) 0.1 (د) 0.2

(24) النسبة المئوية للسيارات من النوع T هي :

(أ) 50 (ب) 50% (ج) 0.25 (د) 25%

(25) الزاوية المركزية للسيارات من النوع H تساوي

(أ) 108^o (ب) 36^o (ج) 90^o (د) 18^o

خاص بالأسئلة من (26) إلى (29) : الجدول المرافق يبين درجات 20 طالباً في أحد المقررات الدراسية :

الدرجة	92	93	94	95	96	97	98	99	100
التكرار	2	2	3	6	1	1	1	3	1

(26) عدد الطلاب الحاصلين على 94 فأقل هو

(أ) 3 (ب) 0.15 (ج) 4 (د) 7

(27) عدد الطلاب الحاصلين على درجة أقل من *94* هو

(أ) 3 (ب) 0.15 (ج) 4 (د) 7

(28) نسبة الطلاب الحاصلين على 94 فأقل هي

(أ) 0.35 (ب) 35% (ج) 4 (د) 7

(29) النسبة المئوية للطلاب الحاصلين على 94 فأقل هي

(أ) 0.35 (ب) 35% (ج) 4 (د) 7

خاص بالأسئلة من (30) على (33) : الجدول المرافق يبين أعمار عدد من العاملات في إحدى المؤسسات (لأقرب سنة) :

الزاوية المركزية	التكرار (العدد) f	المتغير (العمر) x
72^o	20	20
36^o	?	25
?	30	30
?	?	35
	$\sum f$	

(30) عدد العاملات ذات العمر 25 سنة هو :

(أ) 10 (ب) 20

(ج) 30 (د) 40

(31) الزاوية المركزية المناظرة للعمر 30 سنة تساوي

(أ) 36^o (ب) 72^o (ج) 108^o (د) 144^o

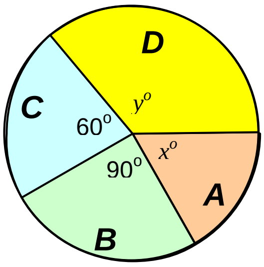
(32) الزاوية المركزية المناظرة للعمر 35 سنة تساوي

(أ) 36^o (ب) 72^o (ج) 108^o (د) 144^o

(33) عدد العاملات الكلي [أي مجموع التكرارات]

(أ) 95 (ب) 100 (ج) 105 (د) 110

خاص بالأسئلة من (34) إلى (37) : الشكل المقابل يبين مبيعات أربع شركات A , B , C , D لبيع لعب الأطفال وذلك خلال أحد الأعياد ، فإذا كان عدد اللعب الكلي التي تم بيعها بواسطة هذه الشركات هو 5400 لعبة ، فإن :



(34) النسبة المئوية لمبيعات الشركة B هي

(أ) 25% (ب) 30% (ج) 40% (د) 60%

(35) عدد اللعب التي باعتها الشركة C هو

(أ) 900 (ب) 2250 (ج) 3150 (د) 1350

(36) عدد اللعب التي باعتها الشركتان A , D معاً هو

(أ) 900 (ب) 2250 (ج) 3150 (د) 1350

(37) نسبة مبيعات الشركة B إلى مبيعات الشركة C هي كالنسبة بين

(أ) 4 إلى 3 (ب) 2 إلى 3 (ج) 3 إلى 4 (د) 3 إلى 2

خاص بالأسئلة من (38) إلى (42) : في إحصائية لعمادة التعلم الإلكتروني والتعليم عن بُعد بجامعة الملك فيصل عن أعداد الطلاب والطالبات الذين تقدموا لاختبارات التعليم المطور للانتساب في الفصل الدراسي الثاني للعام الجامعي 1431/1430 هـ في تخصصات إدارة أعمال وتربية خاصة وآداب كانت البيانات كما هو موضح بالجدول المزدوج التالي :

	طالبات F	طلاب M
إدارة أعمال	480	1480
آداب	2000	3000
تربية خاصة	2560	2000

(38) عدد الطالبات اللاتي تقدمن للاختبارات هو

(أ) 480 (ب) 2000 (ج) 2580 (د) 5040

(39) عدد الطلبة (طالبات وطلاب) الذين تقدموا للاختبارات في تخصص تربية خاصة

(أ) 4560 (ب) 11520 (ج) *6480* (د) 5000

(40) عدد الطلبة (طالبات وطلاب) الذين تقدموا للاختبارات

(أ) 5040 (ب) 5000 (ج) 5040 (د) 11520

(41) نسبة الطلاب (الذكور) تخصص آداب الذين تقدموا للاختبارات وذلك بالقياس لجميع المتقدمين للاختبارات هي (تقريباً)

(أ) 60% (ب) 46.3% (ج) 26% (د) 59.5%

(42) نسبة الطالبات (الإناث) تخصص تربية الذين تقدمن للاختبارات وذلك بالقياس لجميع المتقدمين للاختبارات من تخصص تربية هي (تقريباً)

(أ) 56.1% (ب) 50.8% (ج) 22.2% (د) 39.5%

الإجابة :

(1) د (2) ج (3) ج (4) د (5) ب
(6) ب (7) د (8) أ (9) ب (10) ج
(11) ج (12) د (13) ج (14) د (15) أ
(16) ب (17) أ (18) د (19) ج (20) ب
(21) د (22) ج (23) ج (24) د (25) أ
(26) د (27) ج (28) أ (29) ب (30) أ
(31) ج (32) د (33) ب (34) أ (35) أ
(36) ج (37) د (38) د (39) أ (40) د
(41) ج (42) أ

عناصر الدرس

الباب الثاني : التوزيعات التكرارية [تابع]

- **تنظيم وعرض البيانات المتصلة**
  - ✓ **العرض الجدولي**
    - ❖ **الجدول التكراري والتكراري النسبي**
    - ❖ **الجدول التكراري المتجمع الصاعد والمتجمع الهابط**
  - ✓ **العرض البياني**
    - ❖ **طريقة اللوحة الدائرية**
    - ❖ **طريقة المدرج التكراري**
    - ❖ **طريقة المضلع التكراري**
    - ❖ **طريق المنحنى التكراري**
    - ❖ **طريقة المضلع التكراري المتجمع الصاعد والمتجمع الهابط**

الباب الثاني : التوزيعات التكرارية [تابع]

(3–2) تنظيم وعرض البيانات المتصلة :

كما ذكرنا سابقاً ، فإن البيانات **المتصلة** هي تلك البيانات التي يمكن أن يأخذ فيها المتغير (الخاصية تحت الدراسة) أية قيمة بين قيمتين محددتين [مثل الأطوال ، الأوزان ، درجات الحرارة ، الدخل الشهري أو السنوي ، وغيرها] . ويمكن عرض هذه البيانات أيضاً عن طريق الجداول أو بيانياً . ولتوضيح ذلك دعنا ننظر للمثال التوضيحي التالي :

مثال (2–6) : في تجربة على أطوال سيقان زهور معينة في أحد المعامل البحثية بكلية الزراعة بجامعة الملك فيصل، قيست سيقان 50 زهرة فكانت البيانات كالتالي :

x	$0 \le x < 20$	$20 \le x < 30$	$30 \le x < 35$	$35 \le x < 40$	$40 \le x < 50$	$50 \le x < 60$
f	4	16	12	10	6	2

حيث x هو طول الساق (بوحدات السنتيمتر) ، f هو عدد الأزهار . **المطلوب** عرض هذه البيانات بطرق مختلفة .

قبل أن نبدأ في عرض البيانات لابد من التذكير والتنويه للتالي :

1. البيانات هنا بيانات **كمية متصلة** فيها المتغير x (طول الزهرة) متغير كمي متصل ، وعدد الأزهار f هو تكرار المتغير x [وهذا واضح] .

الفئة	المتغير x (الطول)
الأولى	$0 \le x < 20$
الثانية	$20 \le x < 30$
الثالثة	$30 \le x < 35$
الرابعة	$35 \le x < 40$
الخامسة	$40 \le x < 50$
السادسة	$50 \le x < 60$

2. قيم المتغير x هنا معطاة على صورة ست **فترات** أو ما يُسمى بـ **الفئات** حيث :

الفئة الأولى : $[0 \le x < 20]$ يكون المتغير أكبر من أو يساوي 0 إلى ما قبل 20

الفئة الثانية : $[20 \le x < 30]$ يكون المتغير أكبر من أو يساوي 20 إلى ما قبل 30

الفئة الثالثة : $[30 \le x < 35]$ يكون المتغير أكبر من أو يساوي 30 إلى ما قبل 35

الفئة الرابعة : $[35 \le x < 40]$ يكون المتغير أكبر من أو يساوي 35 إلى ما قبل 40

الفئة الخامسة : $[40 \le x < 50]$ يكون المتغير أكبر من أو يساوي 40 إلى ما قبل 50

الفئة السادسة : $[50 \le x < 60]$ يكون المتغير أكبر من أو يساوي 50 إلى ما قبل 60

انتبه للفرق بين المتباينات وطريقة كتابتها ومعناها :

* فالمتباينة $x < 10$ [وتقرأ x **أقل من** 10] تعني أن x لا تأخذ القيمة 10 ولكن تأخذ جميع القيم الأقل من 10

* والمتباينة $x \le 10$ [وتقرأ x **أقل من أو تساوي** 10] تعني أن x تأخذ القيمة 10 وجميع القيم أقل من تأخذ جميع القيم الأقل من 10

* والمتباينة $x > 10$ [وتقرأ x **أكبر من** 10] تعني أن x لا تأخذ القيمة 10 ولكن تأخذ جميع القيم الأكبر من 10

* واخيراًالمتباينة $x \ge 10$ [وتُقرأ x **أكبر من أو تساوي** 10] تعني أن x تأخذ القيمة 10 وجميع القيم الأكبر من 10

3. لكل فئة حدان : حد **أدنى** ، وحد **أعلى** :

فالفئة الأولى $[0 \le x < 20]$: حدها الأدنى 0 وحدها الأعلى 20 [وهو الحد الأدنى للفئة الثانية] .

والفئة الثانية $[20 \le x < 30]$: حدها الأدنى 20 [الحد الأعلى للفئة الأولى] وحدها الأعلى 30 [وهو الحد الأدنى للفئة الثالثة] .

والفئة الثالثة $[30 \le x < 35]$: حدها الأدنى 30 [الحد الأعلى للفئة الثاية] وحدها الأعلى 35 [وهو الحد الأدنى للفئة الرابعة] .

والفئة الرابعة $[35 \le x < 40]$: حدها الأدنى 35 [الحد الأعلى للفئة الثالثة] وحدها الأعلى 40 [وهو الحد الأدنى للفئة الخامسة] .

والفئة الخامسة $[40 \le x < 50]$: حدها الأدنى 40 [الحد الأعلى للفئة الرابعة] وحدها الأعلى 50 [وهو الحد الأدنى للفئة السادسة] .

والفئة السادسة والأخيرة $[50 \le x < 60]$: حدها الأدنى 50 [الحد الأعلى للفئة الخامسة] وحدها الأعلى 60 .

أي أن الفئات **متصلة** ولا فراغات بينها ، والحد الأدنى لكل فئة من الفئات الوسطى [غير الأولى والأخيرة] هو الحد الأعلى للفئة السابقة ، والحد الأعلى لها هو الحد الأدنى للفئة التالية لها ، وعليه يمكن كتابة الفئات على الصورة :

الفئة	المتغير x (الطول)	
الأولى	**0 -**	من 0 إلى ما قبل 20
الثانية	**20 -**	من 20 إلى ما قبل 30
الثالثة	**30 -**	من 30 إلى ما قبل 35
الرابعة	**35 -**	من 35 إلى ما قبل 40
الخامسة	**40 -**	من 40 إلى ما قبل 50
السادسة	**50 - 60**	من 50 إلى ما قبل 60

يمكنك التعرف على المزيد فيما يخص الفئات يوذلك مع نهاية هذا الدرس

4. لكل فئة **طول** [سنرمز له بالرمز c] وهو يساوي الفرق بين حدها الأعلى وحدها الأدنى ؛

طول أي فئة = حدها الأعلى – حدها الأدنى

فالفئة الأولى $[0 \le x < 20]$ طولها يساوي **20** والثانية $[20 \le x < 30]$ طولها يساوي **10** والثالثة $[30 \le x < 35]$ والرابعة $[35 \le x < 40]$ طول كلٍ منه يساوي **5** والخامسة $[40 \le x < 50]$ والسادسة $[50 \le x < 60]$ طول كلٍ منها يساوي **10** ، أي أن الفئات هنا غير متساوية الأطوال [عندئذٍ نقول ببساطة أن الفئات **غير متساوية**] .

5. لكل فئة **مركز** [وسنرمز له بالرمز x_0] وهي قيمة المتغير x الواقعة في منتصف تلك الفئة ، وتُحسب ببساطة على أنها متوسط حديها الأدنى والأعلى ، أي أن :

$$\text{مركز أي فئة} = \frac{\text{حدها الأدنى} + \text{حدها الأعلى}}{2}$$

فالفئة الأولى $[0 \le x < 20]$ مركزها **10** والثانية $[20 \le x < 30]$ مركزها **25** والثالثة $[30 \le x < 35]$ مركزها **32.5** والرابعة $[35 \le x < 40]$ مركزها **37.5** والخامسة $[40 \le x < 50]$ مركزها **45** والسادسة (الأخيرة) $[50 \le x < 60]$ مركزها **55** .

لاحظ أن مركز أي فئة هو أيضاً حدها الأدنى مضافاً إليه نصف طول الفئة [تحقق بنفسك]

ويمكن تجميع كل ما تقدم في جدول واحد كالتالي :

الفئة	المتغير x (الطول)	طول الفئة c	مركز الفئة x_0
الأولى	$0 \le x < 20$	20 – 0 = 20	(0 + 20) ÷ 2 = 10
الثانية	$20 \le x < 30$	30 – 20 = 10	(20 + 30) ÷ 2 = 25
الثالثة	$30 \le x < 35$	35 – 30 = 5	(30 + 35) ÷ 2 = 32.5
الرابعة	$35 \le x < 40$	40 – 35 = 5	(35 + 40) ÷ 2 = 37.5
الخامسة	$40 \le x < 50$	50 – 40 = 10	(40 + 50) ÷ 2 = 45
السادسة	$50 \le x < 60$	60 – 50 = 10	(50 + 60) ÷ 2 = 55

وبعد هذا التمهيد الضروري ، نعود إلى مثالنا ، حيث يمكن عرض البيانات **جدولياً** ، **وبيانياً** كالتالي :

أولاً : العرض الجدولي للبيانات الكمية المتصلة :

كما في حالة البيانات المنفصلة ، يمكن عرض البيانات الكمية المتصلة عن طريق جداول مختلفة سنستعرض بعضها فيما يلي :

(1) الجدول (التوزيع) التكراري (أو التكراري النسبي) :

يُعتبر الجدول المعطى في المثال إحدى طرق العرض الجدولي للبيانات ، يُسمى هذا الجدول [جدول (2–9)] بـ "**الجدول التكراري**" أو "**التوزيع التكراري**" للبيانات وهو في هذا المثال معطى مباشرةً [وسوف نرى فيما بعد كيف يمكن تكوينه من بيانات خام] ، ومن المفضل أن نقوم بحساب مجموع التكرارات $\sum f$.

جدول (2–9) : الجدول التكراري [مثال (2–6)]

التكرار f	المتغير x (الطول)
4	$0 \le x < 20$
16	$20 \le x < 30$
12	$30 \le x < 35$
10	$35 \le x < 40$
6	$40 \le x < 50$
2	$50 \le x < 60$
$\sum f = 50$	

أيضاً من الممكن أن نضيف لهذا الجدول [إذا دعت الحاجة لذلك] عموداً يوضح التكرار النسبي $\bar{f}$ لكل فئة من الفئات [تذكر : التكرار النسبي هو النسبة بين التكرار f ومجموع التكرارات] فنحصل على الجدول (2–10) الذي يُسمى **بالجدول** (أو **التوزيع**) **التكراري النسبي للبيانات** .

جدول (2–10) : الجدول التكراري النسبي لبيانات مثال (2–6)		
التكرار النسبي $\bar{f}$	**التكرار (العدد) f**	**المتغير x (الطول)**
$4 \div 50 = 0.08$ or 8%	4	$0 \le x < 20$
$16 \div 50 = 0.32$ or 32%	16	$20 \le x < 30$
$12 \div 50 = 0.24$ or 24%	12	$30 \le x < 35$
$10 \div 50 = 0.20$ or 20%	10	$35 \le x < 40$
$6 \div 50 = 0.12$ or 12%	6	$40 \le x < 50$
$2 \div 50 = 0.04$ or 4%	2	$50 \le x < 60$
$\sum \bar{f} = 1$ or 100%	$\sum f = 50$	

تذكر أن مجموع التكرارات النسبية $\sum \bar{f}$ **يجب** أن يساوي الواحد الصحيح (أو 100%) أي أن $\left[\sum \bar{f} = 1 \quad (=100\%)\right]$.

(2) الجدول (التوزيع) التكراري المتجمع الصاعد و الهابط (أو النازل) :

وهي جداول يتم إعدادها لإعطاء نتيجة تراكمية لمجموعة من الفئات والتي يمكن أن تكون بشكل تصاعدي أو تنازلي ، ولكن منها أهمية في تفسير النتائج والظواهر المختلفة .

ويُعرف **التكرار المتجمع الصاعد** المناظر لقيمة معينة a للمتغير x على أنها مجموع التكرارات لجميع القيم الأقل من a [أي لجميع القيم $x < a$] ، فمثلاً بالرجوع لجدول (2–9) [أو جدول (2–10)] فإن :

- التكرار المتجمع الصاعد المناظر للقيمة 0 هو 0 [ليست هناك مشاهدات لـ $x < 0$] ،
- والتكرار المتجمع الصاعد المناظر للقيمة 20 هو 4 ،
- والتكرار المتجمع الصاعد المناظر للقيمة 30 هو : $4+16=20$ ،
- والتكرار المتجمع الصاعد المناظر للقيمة 35 هو : $\underbrace{4+16}_{=20}+12=32$ ،
- والتكرار المتجمع الصاعد المناظر للقيمة 40 هو : $\underbrace{4+16+12}_{=32}+10=42$ ،
- والتكرار المتجمع الصاعد المناظر للقيمة 50 هو : $\underbrace{4+16+12+10}_{=42}+6=48$
- والتكرار المتجمع الصاعد المناظر للقيمة 60 هو :

$$\underbrace{4+16+12+10+6}_{=48}+2=50=\sum f$$

أي بدأنا بتكرار متجمع صاعد 0 ثم أخذنا في إضافة تكرارات الفئات على التوالي حتى وصلنا إلى تكرار متجمع صاعد يساوي مجموع التكرارات $\sum f$.

وعلى هذا الأساس يمكن من الجدول التكراري (2-9) تكوين ما يُسمى **بالجدول التكراري المتجمع الصاعد** [جدول (2-11)] والذي يتكون من عمودين :

- الأول ويعبر عن قيم المتغير x الأقل من الحدود الدنيا للفئات (مع إضافة الحد الأعلى للفئة الأخيرة لهذا العمود) ،
- والعمود الثاني يُعطي التكرار المتجمع الصاعد المناظر لتلك القيم .

لاحظ أننا في هذا الجدول بدأنا بالتكرار المتجمع 0 (أعلى الجدول) وانتهينا بالتكرار المتجمع $\sum f$ (أسفل الجدول) .

جدول (2-11)
الجدول التكراري المتجمع الصاعد

	التكرار المتجمع	المتغير x	
البداية	0	$x<0$	أقل من 0
	4	$x<20$	أقل من 20
	20	$x<30$	أقل من 30
	32	$x<35$	أقل من 35
	42	$x<40$	أقل من 40
	48	$x<50$	أقل من 50
النهاية	$\sum f=50$	$x<60$	أقل من 60

منه يمكن تكوين

الجدول التكراري

المتغير x	التكرار f
$0\le x<20$	4
$20\le x<30$	16
$30\le x<35$	12
$35\le x<40$	10
$40\le x<50$	6
$50\le x<60$	2
	$\sum f=50$

وبالمثل ويُعرف **التكرار المتجمع الهابط (أو النازل)** المناظر لقيمة معينة a للمتغير x على أنها مجموع التكرارات لجميع القيم الأكبر من أو تساوي a [أي لجميع القيم $x\ge a$] وعليه [أنظر جدول (2-9)] يكون :

- التكرار المتجمع الهابط المناظر للقيمة 60 هو 0 [ليست هناك مشاهدات لـ $x\ge 60$]
- والتكرار المتجمع الهابط المناظر للقيمة 50 هو 2 ،

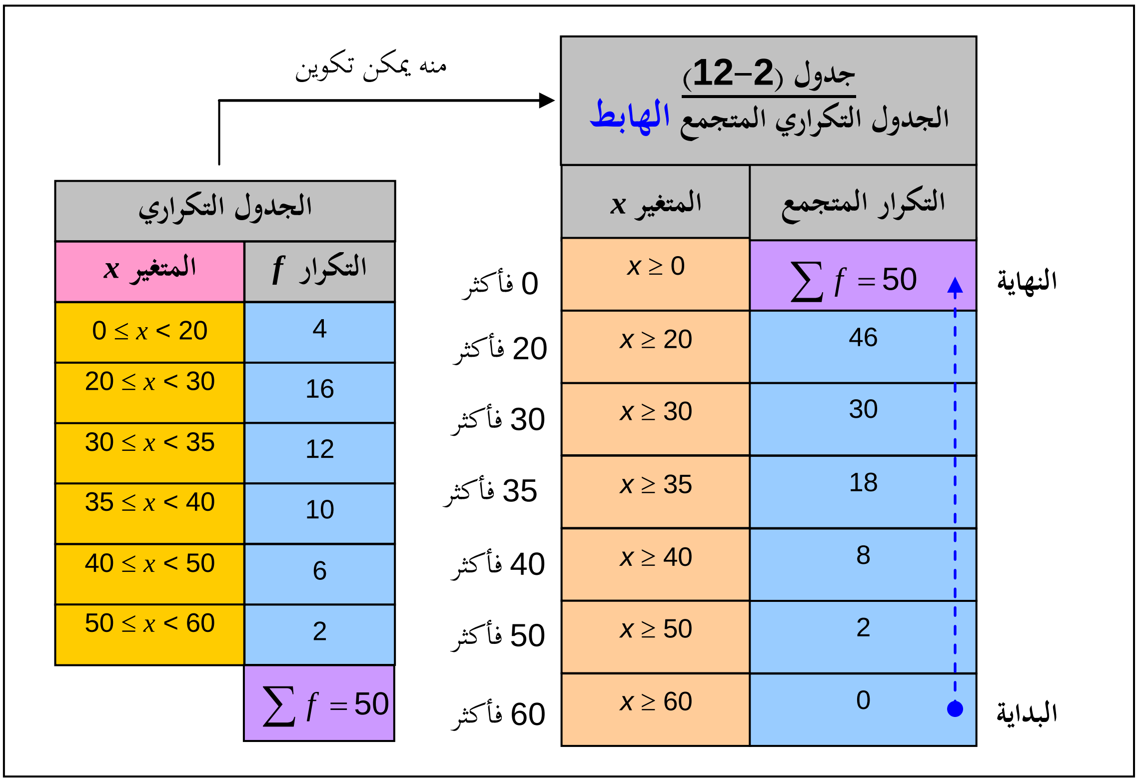
- والتكرار المتجمع الهابط المناظر للقيمة 40 هو : $2+6=8$ ،
- والتكرار المتجمع الهابط المناظر للقيمة 35 هو : $\underbrace{2+6}_{=8}+10=18$ ،
- والتكرار المتجمع الهابط المناظر للقيمة 30 هو : $\underbrace{2+6+10}_{=18}+12=30$ ،
- والتكرار المتجمع الهابط المناظر للقيمة 20 هو : $\underbrace{2+6+10+12}_{=30}+16=46$
- والتكرار المتجمع الهابط المناظر للقيمة 0 هو :

$$\underbrace{2+6+10+12+16}_{=46}+4=50=\sum f$$

ويمكن ترجمة ما تم في صورة جدول يُسمى بالجدول التكراري المتجمع الهابط (أو النازل) [جدول (2−12)] والذي يتكون أيضاً من عمودين :

- الأول ويعبر عن قيم المتغير x الأكبر من أو المساوية للحدود الدنيا للفئات (مع إضافة الحد الأعلى للفئة الأخيرة لهذا العمود) ،
- والعمود الثاني يُعطي التكرار المتجمع الهابط المناظر لتلك القيم .

لاحظ أننا في هذا الجدول بدأنا بالتكرار المتجمع 0 (أسفل الجدول) وانتهينا بالتكرار المتجمع $\sum f$ (أعلى الجدول) .



الجدول التكراري

التكرار f	المتغير x
4	$0 \le x < 20$
16	$20 \le x < 30$
12	$30 \le x < 35$
10	$35 \le x < 40$
6	$40 \le x < 50$
2	$50 \le x < 60$
$\sum f = 50$	

جدول (2−12)

الجدول التكراري المتجمع الهابط

	التكرار المتجمع	المتغير x	
النهاية	$\sum f = 50$	$x \ge 0$	0 فأكثر
	46	$x \ge 20$	20 فأكثر
	30	$x \ge 30$	30 فأكثر
	18	$x \ge 35$	35 فأكثر
	8	$x \ge 40$	40 فأكثر
	2	$x \ge 50$	50 فأكثر
البداية	0	$x \ge 60$	60 فأكثر

ملاحظات هامة :

1. هناك بعض المراجع تعرف التكرار المتجمع الصاعد والتكرار المتجمع الهابط كالآتي

التكرار المتجمع الصاعد إلى فئة معينة هو مجموع التكرارات المقابلة لجميع القيم الأقل من الحد الأعلى لتلك الفئة والمتضمن تكرارها أيضاً .

التكرار المتجمع الهابط إلى فئة معينة هو مجموع التكرارات المقابلة لجميع القيم الأكبر من أو تساوي الحد الأدنى لتلك الفئة .

وليس هناك تعارض بين هذه التعريفات والتعريفات التي قدمناها سابقاً ، ولكننا فضلنا التعريفات المقدمة لأنها (من وجهة نظرنا) أبسط ، وأيضاً تجعلنا نتعامل مع الحدود الدنيا للفئات فقط ، وليس الحدود العليا مرة (كما في حالة التكرار المتجمع الصاعد) والحدود الدنيا مرة أخرى (كما في حالة التكرار المتجمع الهابط) ، لذا نعيد التعريفات التي سننتهجها في هذا الكتاب :

التكرار المتجمع الصاعد المناظر لقيمة معينة *a* للمتغير *x* هو مجموع التكرارات لجميع القيم الأقل من *a* أي لجميع القيم [$x < a$]

التكرار المتجمع الهابط (أو النازل) المناظر لقيمة معينة *a* للمتغير *x* هو مجموع التكرارات لجميع القيم الأكبر من أو تساوي *a* أي لجميع القيم [$x \geq a$]

2. عند تكوين الجدول التكراري المتجمع الصاعد تكون عناصر العمود الأول [عمود المتغير] هي : الحدود الدنيا للفئات $x <$ عدا العنصر الأخير فهو :

الحد الأعلى للفئة الأخيرة $x <$

3. ونفس الأمر بالنسبة للجدول التكراري المتجمع الهابط مع الاختلاف الوحيد في علامة التباين ، أي تكون عناصر العمود الأول هي : الحدود الدنيا للفئات $x \geq$ عدا العنصر الأخير فهو :

الحد الأعلى للفئة الأخيرة $x \geq$

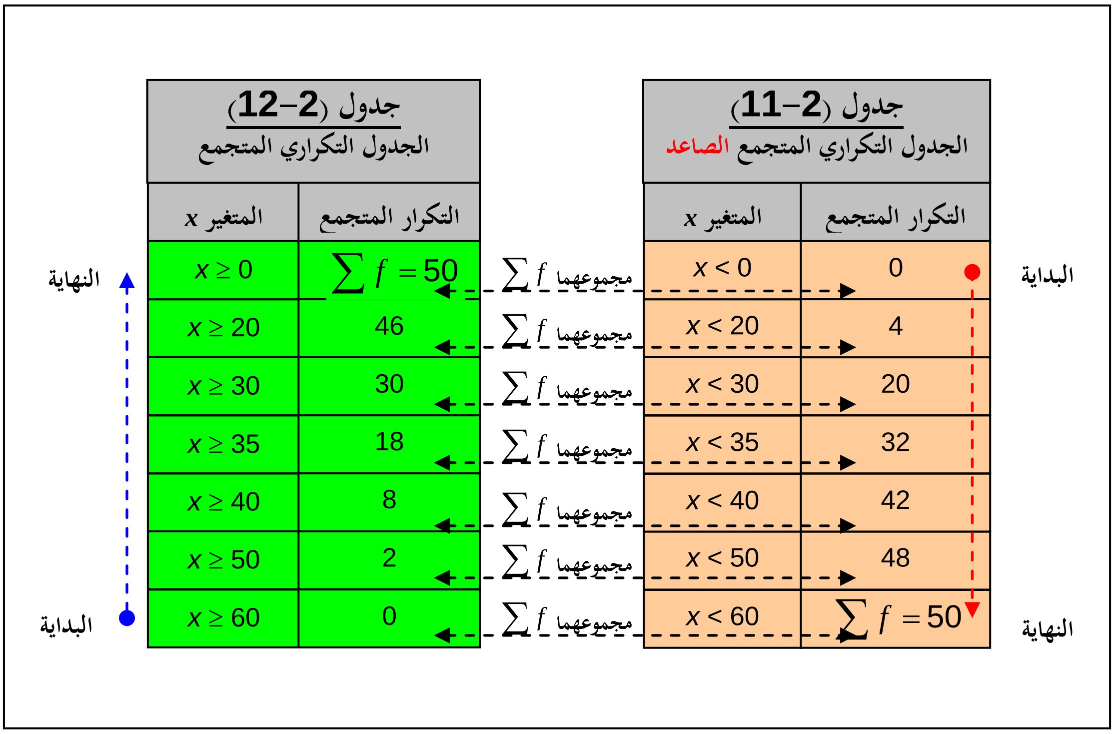
لذا يُسمى الجدول التكراري المتجمع الصاعد في بعض الأحيان بجدول الـ "الأقل من" ، في حين يُسمى الجدول التكراري المتجمع الهابط في بعض الأحيان جدول الـ "أكبر من أو يساوي" أو " .... فأكثر" .

4. يمكن إضافة عمود لكلٍ من الجدوليين التكراريين المتجمعين الصاعد والهابط ليعطيا التكرار المتجمع النسبي (الصاعد أو الهابط) وذلك بقسمة التكرار المتجمع على مجموع التكرارات [جداول (3-13) ، (3-14)] .

جدول (13–2) الجدول التكراري (أو التكراري النسبي) المتجمع الصاعد		
التكرار النسبي المتجمع	التكرار المتجمع	المتغير x
0 ÷ 50 = 0 [0%]	0	$x < 0$
4 ÷ 50 = 0.08 [8%]	4	$x < 20$
20 ÷ 50 = 0.40 [40%]	20	$x < 30$
32 ÷ 50 = 0.64 [64%]	32	$x < 35$
42 ÷ 50 = 0.84 [84%]	42	$x < 40$
48 ÷ 50 = 0.96 [96%]	48	$x < 50$
$\sum \bar{f} =$ 50 ÷ 50 = 1 [100%]	$\sum f = 50$	$x < 60$

جدول (14–2) الجدول التكراري (أو التكراري النسبي) المتجمع الهابط		
التكرار النسبي المتجمع	التكرار المتجمع	المتغير x
$\sum \bar{f} =$ 0 ÷ 50 = 1 [100%]	$\sum f = 50$	$x \geq 0$
46 ÷ 50 = 0.92 [92%]	46	$x \geq 20$
30 ÷ 50 = 0.60 [60%]	30	$x \geq 30$
18 ÷ 50 = 0.36 [36%]	18	$x \geq 35$
8 ÷ 50 = 0.16 [16%]	8	$x \geq 40$
2 ÷ 50 = 0.04 [4%]	2	$x \geq 50$
0 ÷ 50 = 0 [0%]	0	$x \geq 60$

5. بمقارنة الجدولين التكراريين المتجمعين الصاعد (11–2) والهابط (12–2) نلاحظ الآتي :



جدول (2–11)
الجدول التكراري المتجمع الصاعد

التكرار المتجمع	المتغير x
0	$x < 0$
4	$x < 20$
20	$x < 30$
32	$x < 35$
42	$x < 40$
48	$x < 50$
$\sum f = 50$	$x < 60$

جدول (2–12)
الجدول التكراري المتجمع

التكرار المتجمع	المتغير x
$\sum f = 50$	$x \geq 0$
46	$x \geq 20$
30	$x \geq 30$
18	$x \geq 35$
8	$x \geq 40$
2	$x \geq 50$
0	$x \geq 60$

- في الجدول التكراري المتجمع **الصاعد** (2–11) ، يزداد التكرار المتجمع كلما اتجهنا لأسفل الجدول ، [حيث أننا نبدأ بالقيمة 0 (أعلى الجدول) ثم نضيف التكرارات على التوالي ونحن متجهون لأسفل الجدول حتى ننتهي بالقيمة $\sum f$ (أسفل الجدول)] ، أما في الجدول التكراري المتجمع الهابط (2–12) فإن التكرار المتجمع يقل كلما اتجهنا لأعلى الجدول [حيث أننا نبدأ بالقيمة 0 (أسفل الجدول) ثم نضيف التكرارات على التوالي ونحن متجهون لأعلى الجدول حتى ننتهي بالقيمة $\sum f$ (أعلى الجدول)].

- مجموع كل تكراريين متجمعين متناظرين في الجدولين ثابت ويساوي $\sum f$ وبالتالي [وعلى هذا الأساس] يمكن تكوين أحد الجدولين من الآخر ، بتعبير آخر فإن أحد الجدولين من الممكن أن يغنينا عن الآخر .

- تفيد الجداول السابقة في الرد على بعض الأسئلة مثل :
  - عدد القيم الأقل من قيمة معينة [مثلاً 30] فهي تساوي 20 [مباشرةً من الجدول الصاعد] وهي أيضاً 30 – 50 أي 20 [من الجدول الهابط] .
  - عدد القيم الأكبر من أو تساوي قيمة معينة [مثلاً 40] فهي تساوي 8 [مباشرةً من الجدول الهابط] وهي أيضاً 42 – 50 أي 8 [من الجدول الصاعد] .
  - عدد القيم المحصورة بين قيمتين معينتين [مثلاً 20 ، 40] فهي 4 – 42 أي 38 [من الجدول الصاعد] وهي أيضاً 8 – 46 أي 38 أيضاً [من الجدول الهابط] .

هذا بالإضافة لفوائد أخرى سنتعرف عليها في البنود والفصول القادمة بإذن الله .

بهذا نكون قد تعرضنا على الطرق شائعة الاستخدام لعرض البيانات المتصلة عن طريق الجداول ، وتبقى سؤال واحد فيما يختص بالفئات ، وهو :

ما جدوى وضع البيانات على صورة فئات ؟ وما هو الأسلوب الذي يمكن اتباعه لتحويل البيانات من صورتها الخام إلى فئات (كما هو معطى في المثال السابق) **، وهل وضع البيانات على صورة فئات يحافظ على جميع المعلومات الخاصة بالظاهرة تحت الدراسة أم هناك بعض المعلومات قد تُفقد نتيجة لذلك ؟**

هذا السؤال المركب يمكن الرد عليه من خلال المثال التالي :

مثال (2–7) : في نفس المدرسة السابق ذكرها في مثال (2–1) قام نفس المدرس بجمع البيانات عن درجة الطلاب في مادة اللغة الإنجليزية بالثانوية العامة ليس بالفصل **المتميز** السابق ولكن بجميع فصول المدرسة وعددها 5 فصول بكل منها 20 طالباً فكانت البيانات المجمعة هي كالتالي (الدرجة العظمى 100) :

100	95	99	99	95	93	94	99	98	92
97	95	94	95	93	96	95	92	95	94
66	72	73	87	66	83	72	68	70	75
66	75	79	91	88	65	69	75	77	72
66	66	63	57	68	90	75	82	72	86
73	77	77	62	87	82	61	65	77	72
55	73	89	75	70	73	73	73	73	75
66	66	81	58	68	80	75	82	72	86
73	68	52	53	61	79	52	53	82	63
53	72	66	70	85	84	65	63	62	55

المطلوب عرض النتائج السابقة بطرق مختلفة .

دعنا نفترض أننا أردنا التعامل مع هذا المثال بنفس الأسلوب الذي اتبعناه مع مثال (2–1) ، إذن للبيانات الكمية المتقطعة المعطاة على الصورة الخام السابقة وبإتباع نفس الخطوات السابق إتباعها مع مثال (2–1) سيكون لدينا الآتي :

- **المدى *R*** :

وهو "الفرق بين أكبر قيمة وأقل قيمة" في البيانات المعروضة ، بالمرور على البيانات الخام السابقة يمكن ملاحظة أن أكبر قيمة = 100 وأن أقل قيمة = 52 ، إذن :

$$R = 100 - 52 = 48$$

- **تفريغ البيانات والجدول التكراري والعرض البياني :**

إذا قمنا بتفريغ البيانات [كما في مثال (2-1)] وحددنا التكرار لكل قيمة من قيم x (درجة الطالب) فسوف نجد أن الجدول التكراري سيتكون من 49 صفاً [المدى + 1 أي 48 + 1 = 49] وسنحصل على الجدول التكراري (2-15) [وقد أضفنا إليه عموداً يبين زوايا القياس لكل قيمة وذلك إذا أردنا عرض البيانات عن طريق اللوحة الدائرية] .

- **عرض البيانات عن طريق الرسم [أي بيانياً] :**

ويمكن عرض البيانات بيانياً كما بنفس الأساليب التي اتبعناها في مثال (2-1) ، وشكل (2-21) يبين العرض البياني للبيانات الواردة بجدول (3-13) وذلك باستخدام طريقة الأعمدة البسيطة وطريقة الخط البياني وطريقة اللوحة الدائرية .

لكن أم تلاحظ الآتي :

(1) أن تكوين الجدول التكراري (2-15) السابق يتطلب جهداً ووقتاً طويلاً وأن هذا الجهد وهذا الوقت سيزدادان بشكل كبير جداً عندما يزيد حجم البيانات مما يجعل عملية تنظيم البيانات بهذه الطريقة شبه مستحيلة .

جدول (2-15) : الجدول التكراري لبيانات مثال (2-7)

زاوية القياس	f	تفريغ البيانات	x		زاوية القياس	f	تفريغ البيانات	x
0	0		76		7.2°	2	||	52
14.4°	4	||||	77		10.8°	3	|||	53
0	0		78		0	0		54 *
7.2°	2	||	79		7.2°	2	||	55
3.6°	1	|	80		0	0		56
3.6°	1	|	81		3.6°	1	|	57
14.4°	4	||||	82		3.6°	1	|	58
3.6°	1	|	83		0	0		59
3.6°	1	|	84		0	0		60
3.6°	1	|	85		7.2°	2	||	61
7.2°	2	||	86		7.2°	2	||	62
7.2°	2	||	87		10.8°	3	|||	63
3.6°	1	|	88		0	0		64
3.6°	1	|	89		10.8°	3	|||	65
3.6°	1	|	90		28.8°	8	𝍸 |||	66

67		0	0
68	||||	4	14.4°
69	|	1	3.6°
70	|||	3	10.8°
71		0	0
72	𝍸 ||	7	25.2°
73	𝍸 |||	8	28.8°
74		0	0
75	𝍸 ||	7	25.2°

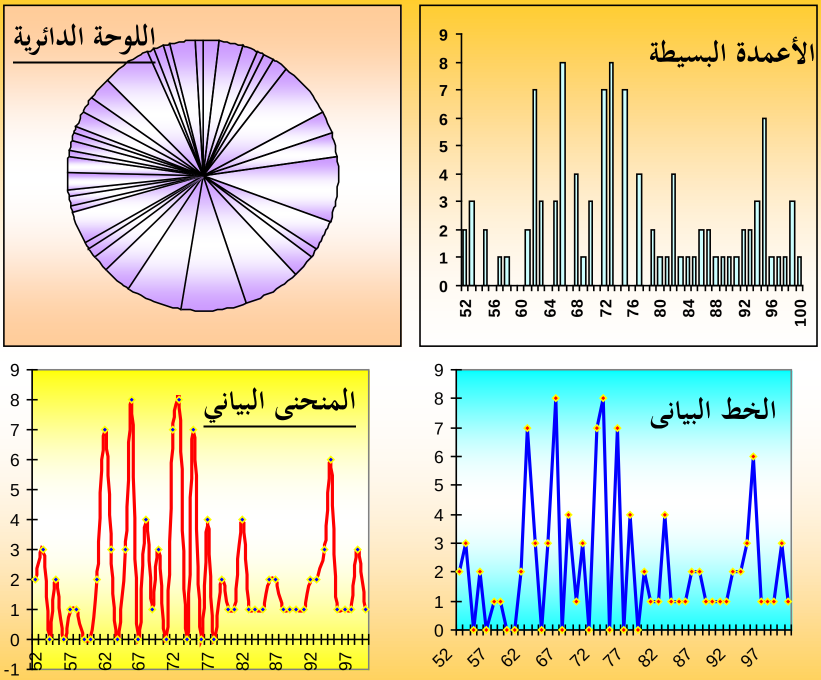
91	|	1	3.6°
92	||	2	7.2°
93	||	2	7.2°
94	|||	3	10.8°
95	𝍸 |	6	21.6°
96	|	1	3.6°
97	|	1	3.6°
98	|	1	3.6°
99	|||	3	10.8°
100	|	1	3.6°
		$\sum f = 100$	

* القيم ذات التكرار 0 يمكن إسقاطها من الجدول بعد تكوينه .

(2) أن طريقة عرض البيانات في هذا المثال [شكل (2-21)] بدأ يفقد أهميته ومعناه عندما بدأ حجم البيانات في الزيادة ، فما بالك عندما يصبح حجم البيانات 1000 طالب أو أكثر وليس 100 طالب .

من هنا كانت الحاجة إلى تلخيص البيانات الخام في صورة فئات .

وبوجهٍ عام لتحديد عدد الفئات التي يتم عرض البيانات من خلالها أو طول الفئة ، فإن ذلك يتوقف على الظاهرة محل الدراسة وهل الظاهرة مقسمة بالطبيعة إلى فئات أم تحتاج إلى تقسيم ، فمثلاً عند التحدث عن تقديرات الطلاب نجد أن لكل تقدير يوجد له فئة محددة مقابلة له كما يوضحها جدول (2-16) [وذلك طبقاً لجامعة الملك فيصل] ، عندئذٍ يُسمى الجدول التكراري الناتج بـ "**جدول تكراري غير منتظم**" [أي أطوال الفئات غير متساوية] .



شكل (2–21) : التمثيل البياني لبيانات مثال (2–7) كبيانات متقطعة

جدول (2–16) : الدرجات مقابل التقديرات

التقدير	الدرجة x	التقدير	الدرجة x
جيد جداً B	$80 \le x < 85$	راسب F	$0 \le x < 60$
جيد جداً مرتفع $B+$	$85 \le x < 90$	مقبول D	$60 \le x < 65$
ممتاز A	$90 \le x < 95$	مقبول مرتفع $D+$	$65 \le x < 70$
ممتاز مرتفع $A+$	$x \ge 95$	جيد C	$70 \le x < 75$
		جيد مرتفع $C+$	$75 \le x < 80$

أما إذا كانت البيانات السابقة تمثل أي ظاهرة أخرى ليس لها تقسيم مسبق ، فمن المناسب وضع البيانات على صورة "**جدول تكراري منتظم**" حيث يتم تقسيم البيانات إلى فئات **متساوية** في الطول كالتالي :

- **نحدد المدى R** : [وهو في مثالنا هذا يساوي 52] .
- **تقسيم المدى R إلى عدد مناسب من الفئات** : حيث يأخذ الباحث عدداً مناسباً من الفئات (يتوقف على حجم البيانات) ، فإذا استعمل عدداً كبيراً من الفئات (أكثر من 20) فسوف يخسر الباحث البساطة المطلوبة ، وإذا استعمل عدداً قليلاً من الفئات (أقل من 5) فسوف يفقد الكثير من تفصيلات البيانات الأصلية (البيانات الخام) . نفرض مثلاً أننا اخترنا 10 فئات .

- **تحديد طول الفئة** : [وذلك بقسمة المدى على عدد الفئات] :

$$\frac{48}{10} = 4.8$$

ومن المناسب أن يكون طول الفئة عدداً صحيحاً أكبر قليلاً من الناتج [ليكن 5]

- نبدأ الفئة الأولى من أقل قيمة في البيانات [أي 52] أو أي قيمة أنسب [أقل من 52 ولتكن 50] ، وبالتالي تكون الفترات كالتالي :

50 – , 55 – , 60 – , 65 – , 70 –
75 – , 80 – , 85 – , 90 – , 95

لكن في مثالنا هذا [لنعود إليه مرةً ثانيةً بعد هذا التنويه البسيط عن كيفية تحديد الفئات] فمن المناسب تقسيم البيانات وفقاً للتقديرات كما هو موضح بجدول (2–16) ، ثم نقوم بتفريغ البيانات كما في مثال (2–1) وتجميع التكرارات لنحصل على الجدول (التوزيع) التكراري المبين بجدول (2–17) .

جدول (2–17) : التوزيع (الجدول) التكراري للبيانات [مثال (2–7)]

التكرار f	تفريغ البيانات (العلامات)	المتغير (الدرجة) x	الفئة
9	𝍸 ||||	$0 \le x < 60$	الأولى
7	𝍸 ||	$60 \le x < 65$	الثانية
16	𝍸 𝍸 𝍸 |	$65 \le x < 70$	الثالثة
18	𝍸 𝍸 𝍸 |||	$70 \le x < 75$	الرابعة
13	𝍸 𝍸 |||	$75 \le x < 80$	الخامسة
8	𝍸 |||	$80 \le x < 85$	السادسة
7	𝍸 ||	$85 \le x < 90$	السابعة
9	𝍸 ||||	$90 \le x < 95$	الثامنة
13	𝍸 𝍸 |||	$95 \le x < 100.0001$	التاسعة
$\sum f = 100$	حيث في الفئة الأخيرة اعتبرنا الحد الأعلى للفئة الأخيرة أكثر (بقيمة لا تُذكر) من 100 بدلاً من كتابتها $x \ge 95$ وهذا لن يؤثر في شيء .		

إذن بمقارنة الجدولين (2–15) ، (2–17) يتضح لنا درجة البساطة في العرض الجدولي للبيانات من خلال وضع البيانات على صورة فئات ، والجدول (2–17) يسمح بالرد على عدد الطلاب الحاصلين على الدرجات الواقعة داخل كل فئة [مثلاً عدد الطلاب الحاصلين على درجة أكبر من أو تساوي 70 وأقل من 75 يساوي 18 (تكرار الفئة الرابعة] ، كما يسمح بمعرفة عدد الطلاب الحاصلين على درجات بين حدود الفئات [مثلاً عدد الطلاب الحاصلين على درجات أكبر من أو تساوي 80 وأقل من 95 يساوي 24 وهو مجموع تكرارات الفئات السادسة والسابعة والثامنة] ، **لكن** لا يسمح (مثلاً) بالرد على سؤال مثل : ما هو عدد الطلاب الحاصلين على

درجة 83 ، في حين يسمح جدول (2-15) بالرد على مثل هذا السؤال . إذن نتيجة التبسيط الذي حصلنا عليه من خلال وضع البيانات على صورة فئات :

يكون هناك ضياع لبعض التفاصيل الخاصة بالبيانات

وعلى الرغم من أن عملية تجميع البيانات على صورة فئات تؤدي بشكل عام إلى ضياع كثير من تفصيلات البيانات الأصلية فإن الفائدة الهامة منها هي الصورة العامة التي يمكن الحصول عليها والعلاقات الأساسية التي تظهر بالتالي أكثر وضوحاً ويمكن تقليص ضياع التفاصيل إلى حد كبير بالاختيار المناسب لعدد الفئات .

بهذا نكون قد أجبنا على السؤال المركب السابق :

ما جدوى وضع البيانات على صورة فئات ؟ وما هو الأسلوب الذي يمكن اتباعه لتحويل البيانات من صورتها الخام إلى فئات ، وهل وضع البيانات على صورة فئات يحافظ على جميع المعلومات الخاصة بالظاهرة تحت الدراسة أم هناك بعض المعلومات قد تُفقد نتيجة لذلك ؟

ولمزيد من المعلومات الخاصة بالفئات ، يمكن للقارئ الاطلاع على مرفق (2) من هذا الكتاب ، لكن نود أن نشير هنا إلى أننا سنتبع الأسلوب الموضح بمثال (2-6) فيما يخص فئات الكميات المتصلة .

ثانياً : العرض البياني للبيانات الكمية المتصلة :

كما في حالة البيانات المنفصلة ، يمكن عرض البيانات الكمية المتصلة بطرق بيانية مختلفة سنستعرض بعضها فيما يلي :

(1) طريقة اللوحة الدائرية
(2) المدرج التكراري
(3) المضلع التكراري
(4) المنحنى التكراري
(5) المضلع (أو المنحنى) المتجمع الصاعد
(6) المضلع (أو المنحنى) المتجمع الهابط .

(1) طريقة اللوحة الدائرية : وهي مشابهة تماماً لطريقة عرض البيانات المنفصلة ، لذا فقد بدأنا طرق العرض بها حيث تُمثل كل فئة بقطاع دائري من دائرة [لها نصف قطر اختياري مناسب] زاويته المركزية تُعطى بـ :

$$\text{الزاوية المركزية لفئة ما} = \frac{\text{تكرار الفئة}}{\text{مجموع التكرارات}} \times 360^o$$

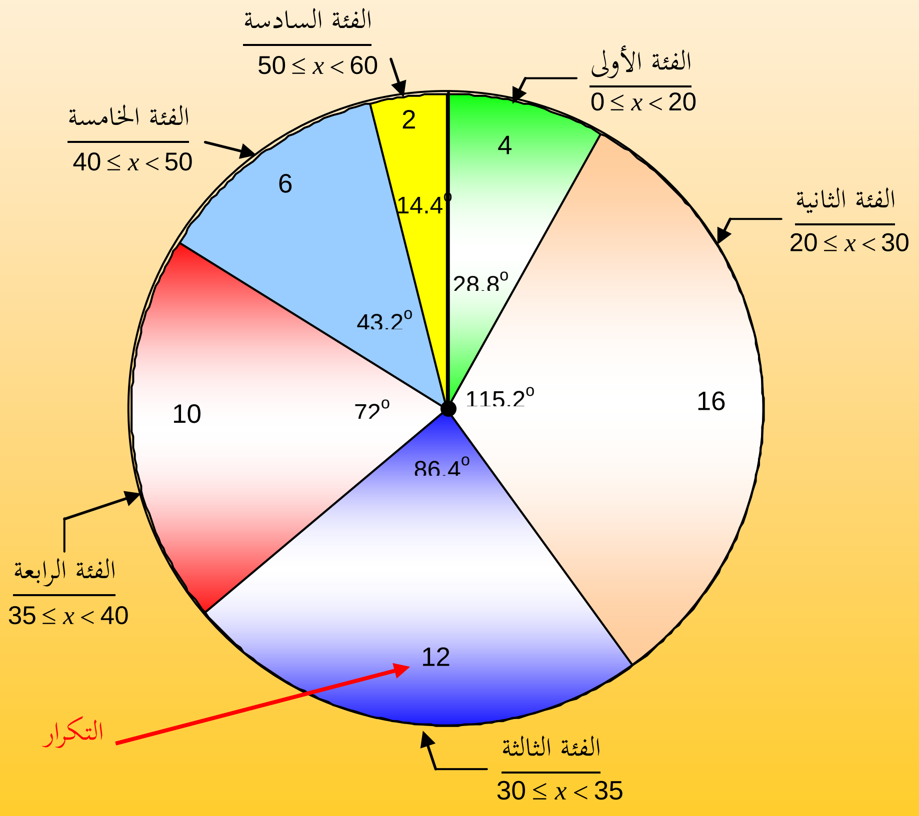
أو

$$\text{الزاوية المركزية لفئة ما} = \text{التكرار النسبي لتلك الفئة} \times 360^o$$

فمثلاً البيانات المعطاة عن سيقان الأزهار [مثال (2–6)] يمكن تمثيلها باللوحة الدائرية المبينة بشكل (2–22) حيث الزوايا المركزية للقطاعات الدائرية الممثلة للفئات يعطيها الجدول (2–18) .

جدول (2–18) : الزوايا المركزية لبيانات مثال (2–6)

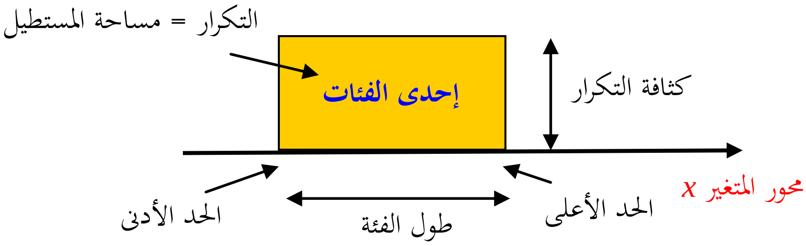
المتغير x	التكرار f	الزاوية المركزية
$0 \le x < 20$	4	$(4 \div 50) \text{ x } 360 = 28.8^o$
$20 \le x < 30$	16	$(16 \div 50) \text{ x } 360 = 115.2^o$
$30 \le x < 35$	12	$(12 \div 50) \text{ x } 360 = 86.4^o$
$35 \le x < 40$	10	$(10 \div 50) \text{ x } 360 = 72^o$
$40 \le x < 50$	6	$(6 \div 50) \text{ x } 360 = 43.2^o$
$50 \le x < 60$	2	$(2 \div 50) \text{ x } 360 = 14.4^o$
	$\sum f = 50$	مجموع الزوايا = 360^o



شكل (2–22) : اللوحة الدائرية لبيانات مثال (2–6)

(2) طريقة المدرج التكراري : وهي طريقة مشابهة لطريقة الأعمدة البسيطة [للبيانات المنفصلة] مع الاختلافات التالية :

- تُمثل كل فئة بمستطيل قاعدته تقع على المحور الأفقي [الذي يمثل المتغير x] وعرضه يساوي طول الفئة ومساحته تساوي تكرار الفئة .



وحيث أن مساحة أي مستطيل تساوي عرض المستطيل مضروباً في ارتفاعه ، فإن ارتفاع أي مستطيل يكون مساوياً لـ ”تكرار الفئة مقسوماً على طول الفئة“ ، يُسمي خارج القسمة هذا بـ ”**كثافة التكرار**“ .

- المحور الرأسي هنا يمثل كثافة التكرار [وليس التكرار كما في حالة الأعمدة البسيطة] ، إلا في حالة الفئات المتساوية الطول فإن المحور الرأسي يمكن أن يمثل التكرار ، وحيث أننا سنتعامل مع فئات قد تكون متساوية الطول وقد تكون أطوالها مختلفة ، لذا سنعتبر دائماً أن المحور الرأسي يمثل كثافة التكرار .
- لا فراغات موجودة بين المستطيلات [حيث أن البيانات هنا بيانات متصلة] بخلاف طريقة الأعمدة في حالة البيانات المنفصلة حيث يجب ألا تكون الأعمدة متلاصقة .

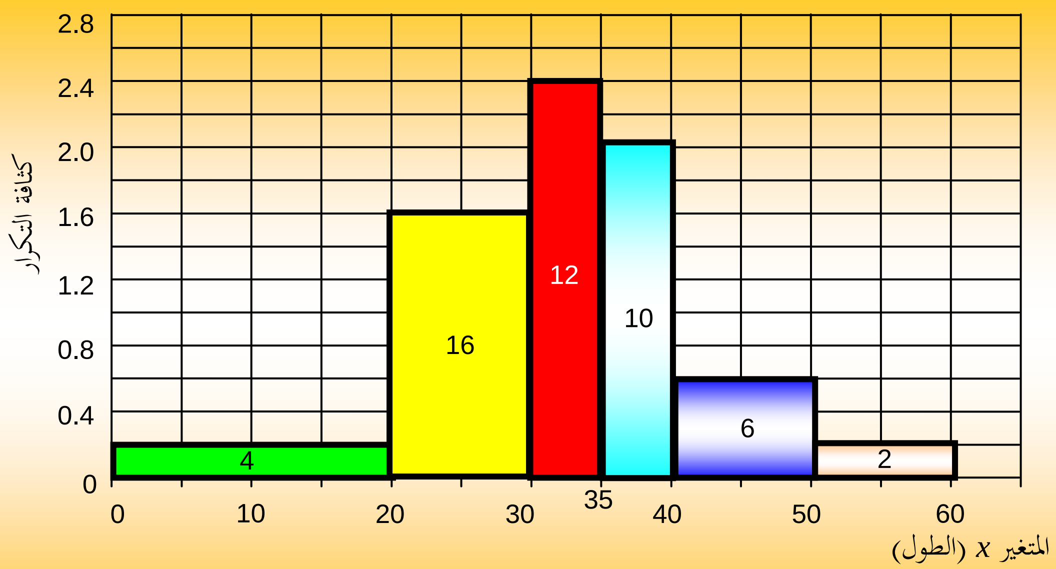
وبالتالي لرسم المدرج التكراري لابد أن نضيف للجدول التكراري أعمدة تبين **طول كل فئة** و **كثافة تكراراها** [جدول (2–19)] .

جدول (2–19) : البيانات اللازمة لرسم المدرج التكراري لبيانات مثال (2–6)

الفئة	المتغير x (الطول)	التكرار f	طول الفئة c	كثافة التكرار f/c
الأولى	$0 \le x < 20$	4	20	$4 \div 20 = 0.2$
الثانية	$20 \le x < 30$	16	10	$16 \div 10 = 1.6$
الثالثة	$30 \le x < 35$	12	5	$12 \div 5 = 2.4$
الرابعة	$35 \le x < 40$	10	5	$10 \div 5 = 2$
الخامسة	$40 \le x < 50$	6	10	$6 \div 10 = 0.6$
السادسة	$50 \le x < 60$	2	10	$2 \div 10 = 0.2$

← الجدول التكراري →

والآن يمكن رسم المدرج التكراري بأخذ محورين متعامدين : الأفقي ويمثل المتغير x [وهنا مقياس الرسم له أو تدريجه مهم] ، والرأسي يمثل كثافة التكرار ونقوم بتمثيل كل فئة بمستطيل قاعدته على المحور الأفقي (وطولها = طول الفئة) وارتفاعه يمثل كثافة تكرار الفئة (وبالتالي مساحته تساوي تكرار الفئة) [شكل (2–23)].



شكل (2-23) : المدرج التكراري لبيانات مثال (2-6)

تذكر أنه في المدرج التكراري :

1. مساحة كل مستطيل هي تكرار الفئة ، وبالتالي فإن مجموع مساحات المستطيلات في المدرج التكراري هي التكرار الكلي [50 هنا]
2. ارتفاع كل مستطيل هو كثافة التكرار وليس التكرار أو التكرار النسبي [كقاعدة عامة] .
3. طول قاعدة أي مستطيل هي طول الفئة .
4. المستطيلات متلاصقة وليس بينها فراغات .

(3) المضلع التكراري :

وهو أسلوب مشابه لطريقة الخط البياني للبيانات المنفصلة ، إلا أن كل فئة تُمثل بنقطة إحداثيها الأفقي هو مركز الفئة ، وإحداثيها الرأسي هو كثافة تكرارها ، ثم نقوم بتوصيل هذه النقاط بالمسطرة فنحصل على خط منكسر يكون هو المضلع التكراري للبيانات ، وبالتالي لرسم المضلع التكراري لابد أن نضيف للجدول التكراري أعمدة تبين طول كل فئة وكثافة تكراراها [كما في حالة المدرج التكراري] إلى جانب عمود يبين **مركز الفئة** [جدول (2-20)] ، ثم نقوم بتحديد النقاط الممثلة للفئات ونوصل بينها بالمسطرة فنحصل على المضلع التكراري [شكل (2-24أ)] .

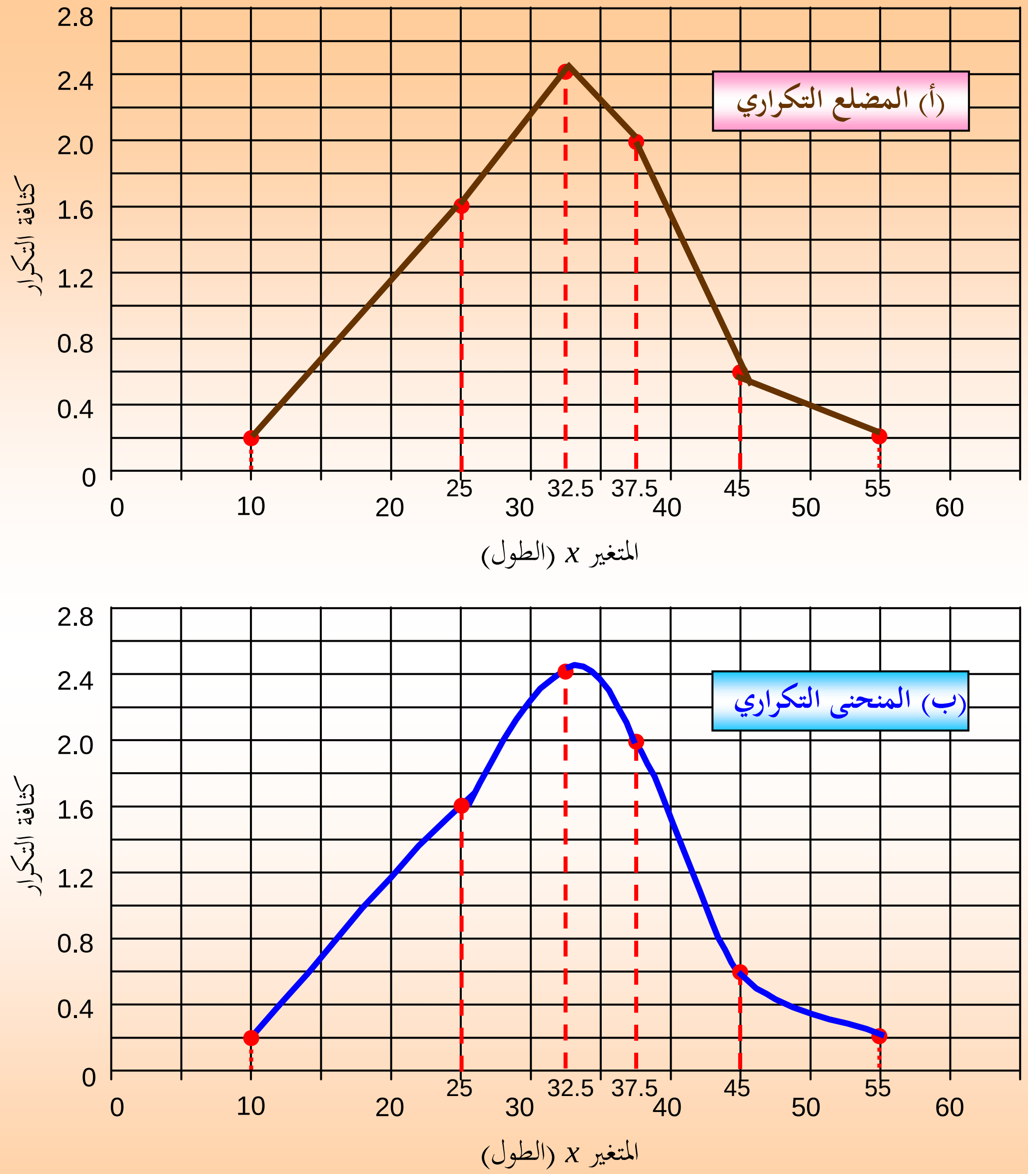
(4) المنحنى التكراري :

وفيه نتبع نفس الخطوات السابق إتباعها مع المضلع التكراري لكن يتم توصيل النقاط الممثلة للفئات باليد (بدلاً من المسطرة) فنحصل على خط ممهد هو المنحنى التكراري للبيانات [شكل (2-24ب)] .

جدول (2–20):البيانات اللازمة لرسم المضلع (أو المنحنى) التكراري [مثال (2–6)]

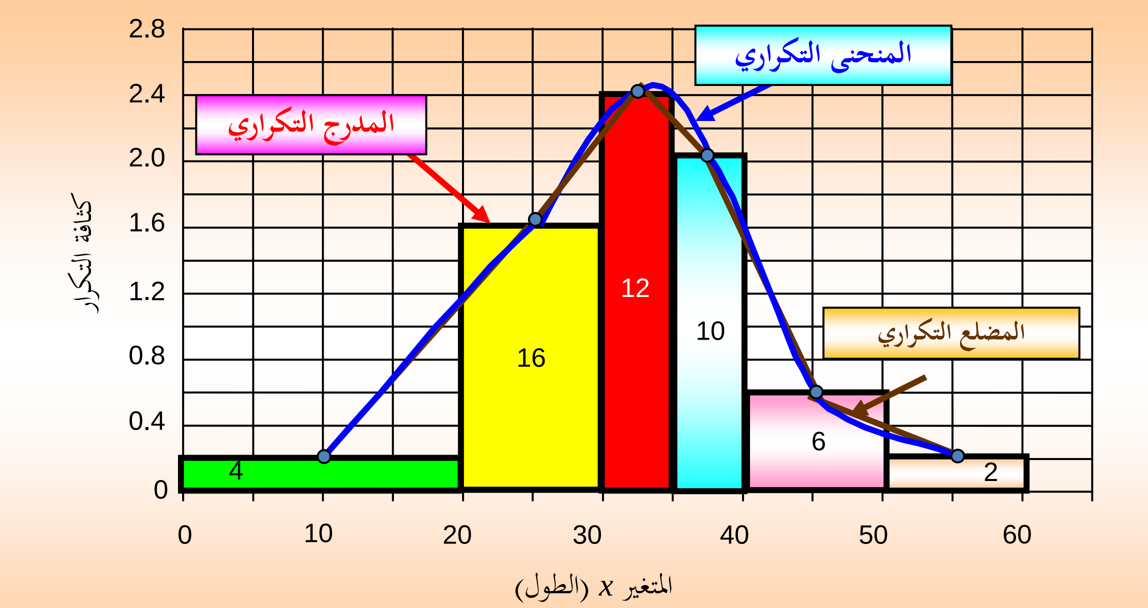
النقطة الممثلة للفئة	كثافة	مركز الفئة	طول الفئة	التكرار f	المتغير x
$A(10,0.2)$	0.2	10	20	4	$0 \le x < 20$
$B(25,1.6)$	1.6	25	10	16	$20 \le x < 30$
$(C(32.5,2.4)$	2.4	32.5	5	12	$30 \le x < 35$
$D(37.5,2)$	2	37.5	5	10	$35 \le x < 40$
$E(45,0.6)$	0.6	45	10	6	$40 \le x < 50$
$F(55,0.2)$	0.2	55	10	2	$50 \le x < 60$

الجدول التكراري



شكل (2–24) : المضلع التكراري والمنحنى التكراري لبيانات مثال (2–6)

وتجدر الإشارة إلى أنه يمكن رسم المدرج التكراري للبيانات والمضلع التكراري والمنحنى التكراري لنفس البيانات على رسمة واحدة كما هو مبين بشكل (2–25) ، لاحظ أن نقطة منتصف القاعدة العليا من كل مستطيل في المدرج التكراري هي النقطة الممثلة للفئة عند رسم كلٍ من المضلع التكراري والمنحنى التكراري .



شكل (2−25) : المدرج التكراري والمضلع التكراري والمنحنى التكراري لبيانات مثال (2−6)

(5) المضلع (أو المنحنى) التكراري المتجمع الصاعد :

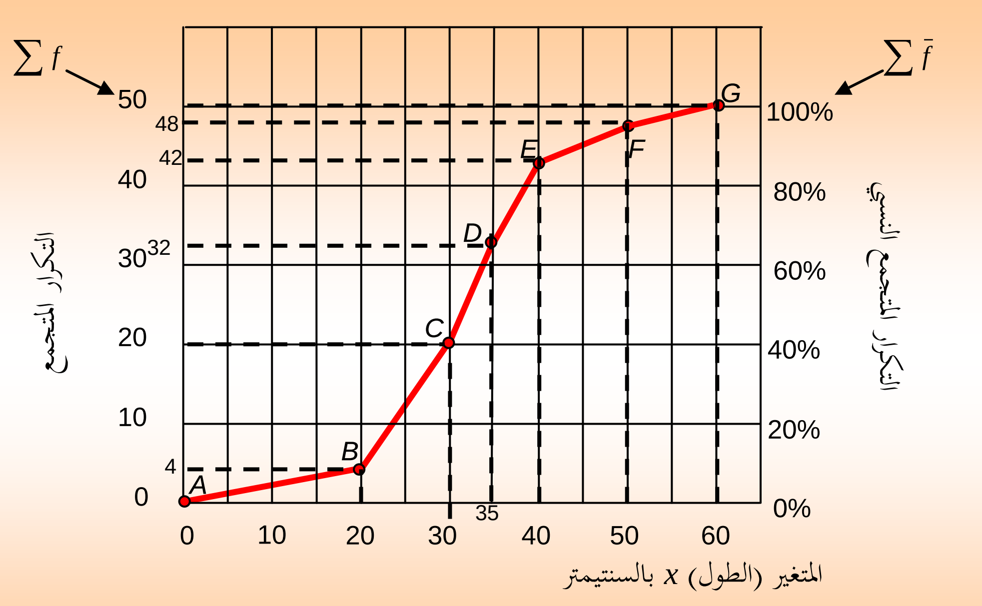
ذكرنا سابقاً عند عرضنا للبيانات عن طريق الجداول أنه يمكن عرض البيانات عن طريق التوزيع التكراري (أو التكراري النسبي) المتجمع الصاعد [جدول (2−13)] ، ويمكن الاستفادة من هذه الجداول في رسم المضلع التكراري المتجمع الصاعد حيث نختار محورين :

- الأفقي ويمثل المتغير x (حيث نقوم بتوقيع حدود الفئات عليه) ،
- والرأسي ويمثل التكرار المتجمع (أو التكرار المتجمع النسبي) .

ثم نقوم بتمثيل البيانات بنقاط إحداثياتها الأفقية هي الحدود الدنيا للنقاط وإحداثياتها الرأسية هي التكرارات المتجمعة الصاعدة المناظرة لهذه الحدود كما هو مبين بالجدول (2−21) [تذكر أن آخر نقطة إحداثيها الأفقي هو الحد الأعلى للفئة الأخيرة وإحداثيها الرأسي هو مجموع التكرارات $\sum f$] ، وبتوصيل هذه النقاط بالمسطرة نحصل على خط منكسر يُسمى **بالمضلع التكراري المتجمع الصاعد** وأحياناً يُسمى بمضلع "**الأقل من**" [شكل (2−26)] . وإذا قمنا بتوصيل النقاط السابقة باليد [بدلاً من المسطرة] نحصل على خط ممهد يُسمى **بالمنحنى التكراري المتجمع الصاعد** .

جدول (2–**21**) : المعلومات اللازمة لرسم المضلع (أو المنحنى) المتجمع الصاعد [مثال (2–**6**)]

النقطة الموقعة على الرسم	التكرار المتجمع النسبي	التكرار المتجمع	المتغير x
$A(0\ ,\ 0)$	0%	0	< 0
$B(20\ ,\ 4)$	8%	4	< 20
$C(30\ ,\ 20)$	40%	20	< 30
$D(35\ ,\ 32)$	64%	32	< 35
$E(40\ ,\ 42)$	84%	42	< 40
$F(50\ ,\ 48)$	96%	48	< 50
$G(60\ ,\ 50)$	100%	50	< 60

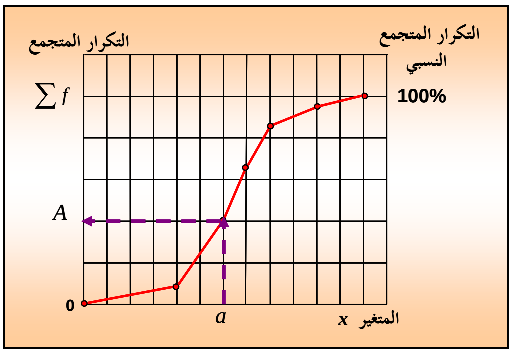


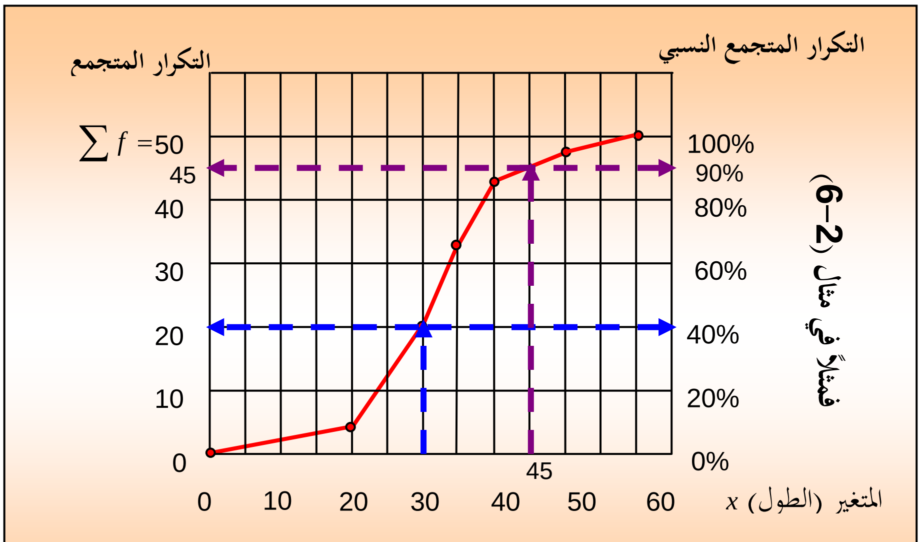
شكل (2–**26**) : المضلع التكراري المتجمع الصاعد لبيانات مثال (2–**6**)

ويفيد المضلع التكراري المتجمع الصاعد في الرد على العديد من الأسئلة نستعرض بعضها في التالي :

- **تحديد التكرار المتجمع المناظر لـ "x أقل من قيمة معينة"**

فلحساب قيمة التكرار المتجمع المناظر لـ "$x < a$" نحدد قيمة a على المحور الأفقي [محور المتغير] ونرسم خطاً رأسياً حتى يتقاطع مع المضلع في نقطة ، فيكون التكرار المتجمع المطلوب هي **القراءة الأفقية A** [على محور التكرار المتجمع] المناظرة لنقطة التقاطع .





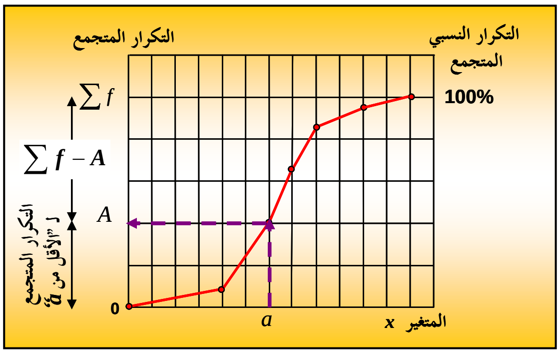
فمثلاً في مثال (2-6)

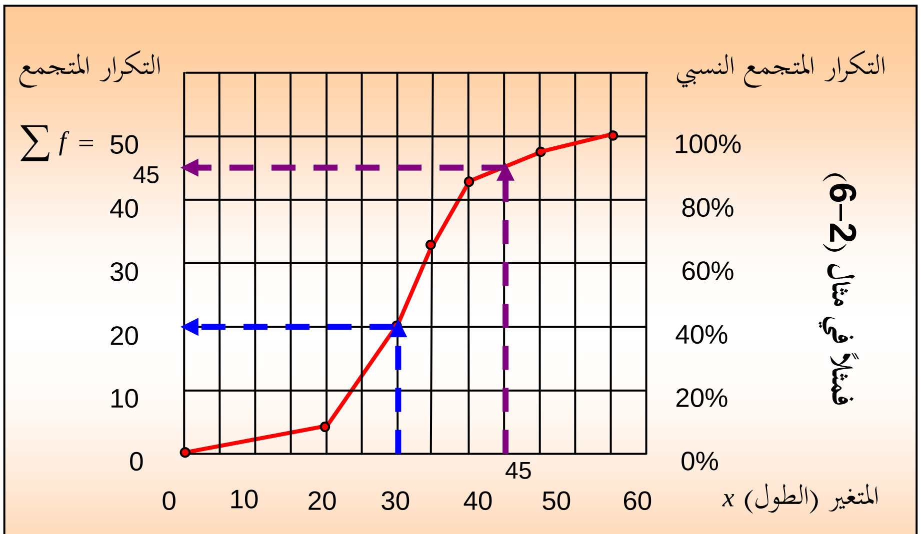
عدد الأزهار التي أطوال سيقانها أقل من 30 هو 20 ونسبتهم المئوية

عدد الأزهار التي أطوال سيقانها أقل من 45 هو 45 ونسبتهم المئوية 90%

- **تحديد التكرار المتجمع المناظر لـ ”x أكبر من أو تساوي قيمة معينة“**

فلحساب قيمة التكرار المتجمع المناظر لـ “$x \geq a$” نحدد قيمة a على المحور الأفقي [محور المتغير] ونرسم خطاً رأسياً حتى يتقاطع مع المضلع في نقطة ، ونحدد القراءة A [على محور التكرار المتجمع] ، فيكون الحل المطلوب هو ”المجموع الكلي للتكرارات – القيمة A“ .

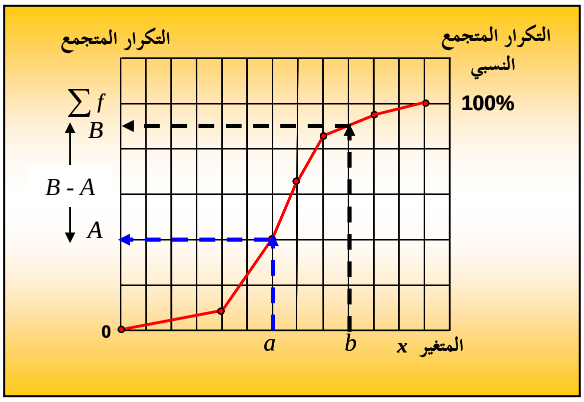




فمثلاً في مثال (2-6)

عدد الأزهار التي أطوال سيقانها 30 فأكثر هو 30 [= 20 – 50]

عدد الأزهار التي أطوال سيقانها 45 فأكثر هو 5 [= 45 – 50]

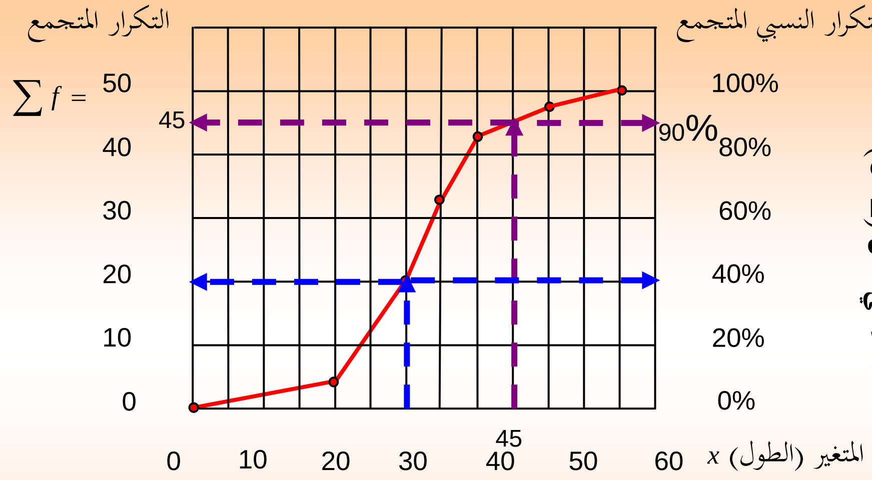


- **تحديد التكرار المتجمع المناظر لـ ”x محصورة بين قيمتين“**

فلحساب قيمة التكرار المتجمع المناظر لـ

“$a \le x < b$”

نحدد قيمتي a , b على المحور الأفقي [محور المتغير] ونحدد قيم التكرارات المتجمعة المناظرة [لتكن A , B على الترتيب] ، فيكون الحل المطلوب هو **الفرق بين القيمتين A , B** .



فمثلاً في مثال (2–9)

عدد الأزهار التي أطوال سيقانها ما بين 30 ، 45 هو 25 = 20 – 45

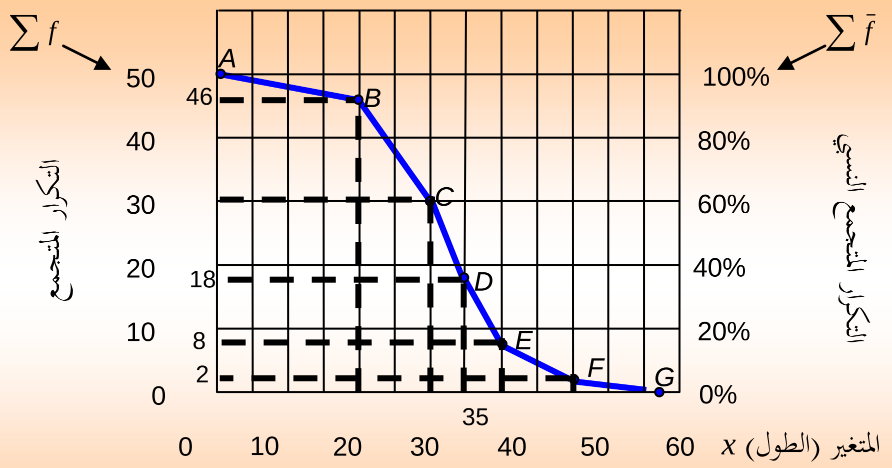
ونسبتهم المئوية تساوي : $\frac{25}{50}$ x 100 = 50%

أو من الرسم : 90 – 40 = 50%

(6) المضلع (أو المنحنى) التكراري المتجمع الهابط : بنفس طريقة رسم المضلع التكراري المتجمع الصاعد يمكن رسم المضلع (المنحنى) التكراري المتجمع الهابط وذلك بالاستعانة بجدول (2–14) مع الاختلاف الوحيد أن الإحداثيات الرأسية للنقاط هي التكرارت المتجمعة الهابطة المناظرة للحدود الدنيا للفئات كما هو مبين بالجدول (2–22) [تذكر أن آخر نقطة إحداثيها الأفقي هو الحد الأعلى للفئة الأخيرة وإحداثيها الرأسي صفر] ، وبتوصيل هذه النقاط بالمسطرة نحصل على خط منكسر يُسمى **بالمضلع التكراري المتجمع الهابط** وأحياناً يُسمى بمضلع "**الأكبر من أو يساوي**" [شكل (2–27)] . وإذا قمنا بتوصيل النقاط السابقة باليد [بدلاً من المسطرة] نحصل على خط ممهد يُسمى **بالمنحنى التكراري المتجمع الهابط** .

جدول (2–22) : المعلومات اللازمة لرسم المضلع (أو المنحنى) المتجمع الهابط [مثال (2–6)]

النقطة الموقعة على الرسم	التكرار المتجمع النسبي	التكرار المتجمع	المتغير x
$A(0 , 50)$	100%	50	≥ 0
$B(20 , 46)$	92%	46	≥ 20
$C(30 , 30)$	60%	30	≥ 30
$D(35 , 18)$	36%	18	≥ 35
$E(40 , 8)$	16%	8	≥ 40
$F(50 , 2)$	4%	2	≥ 50
$G(60 , 0)$	0%	0	≥ 60

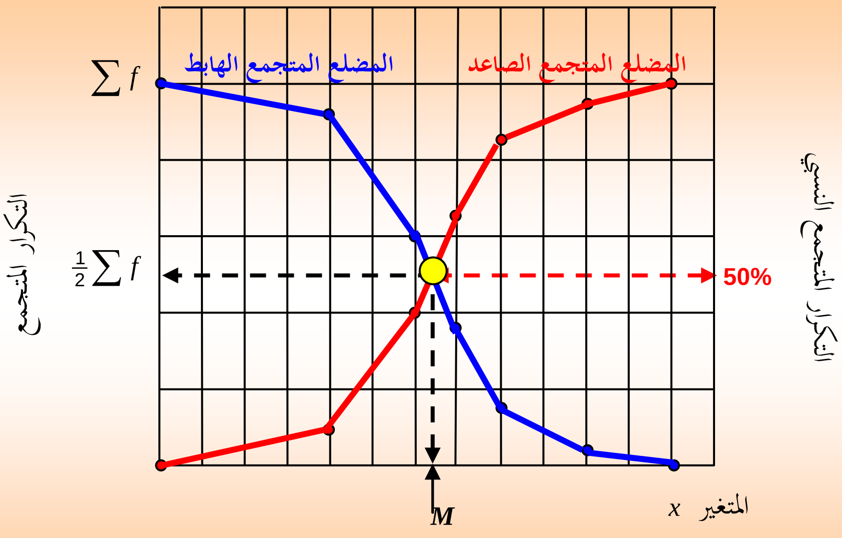


شكل (2–27) : المضلع التكراري المتجمع الهابط لبيانات مثال (2–6)

ويفيد الضلع التكراري المتجمع الهابط في الرد على نفس الأسئلة التي يرد عليها المضلع التكراري المتجمع الصاعد مع الأخذ في الاعتبار أن التدريج الرأسي يمثل التكرار المناظر لـ " x أكبر من أو تساوي" ، أي أن المضلعان التكراريان المتجمعان الصاعد والهابط يؤديان نفس الغرض ، لذا سنوجه اهتمامنا لأحدهما فقط [وليكن الصاعد]

.

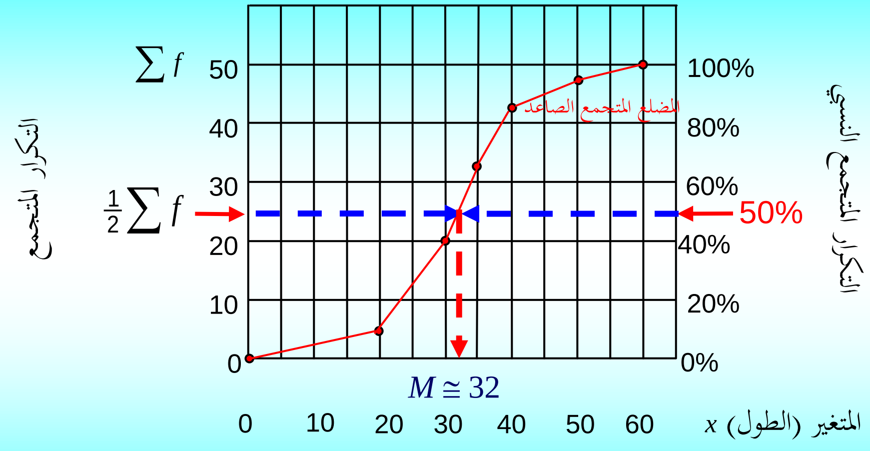
ومن الجدير بالذكر أنه يمكن ويمكن رسم المضلعين التكراريين المتجمعين : **الصاعد** و **الهابط** على رسمة واحدة ، وعندئذٍ سنلاحظ أن المضلعين يتقاطعان في نقطة يناظرها تكرار متجمع قدره $\frac{1}{2}\sum f$ [أو تكرار متجمع نسبي قدره 50%] ، أي أن **القيمة** ***M*** المناظرة لتلك النقطة تقسم البيانات إلى مجموعتين متساويتين في العدد بحيث تكون 50% من قيم المتغير أقل من M ، الـ 50% الأخرى من القيم أكبر منها [شكل (2–28)] وهذه النقطة ذات أهمية كبيرة وتسمى بالـ "**الوسيط**" للبيانات [وسوف نتناولها بالتفصيل في الفصل القادم] ، أي أن :



شكل (2–28) : تحديد الوسيط لمجموعة من البيانات المتصلة بيانياً

أي أن وسيط مجموعة من البيانات المرتبة ترتيباً تصاعدياً أو تنازلياً هي قيمة في وسط المجموعة وتقسمها إلى مجموعتين متساويتين في العدد

فمثلاً في مثال (2–6)



تحديد الوسيط لبيانات مثال (2–6)

والآن يمكن تلخيص طرق العرض البياني للبيانات الكمية المتصلة من خلال المثال التالي :

مثال (2–8) : الجدول التالي يبين الأجر السنوي [بآلاف الريالات السعودية] لـ 60 عاملاً في إحدى الشركات :

120 – 180	100 –	90 –	80 –	70 –	60 –	50 –	x
3	6	9	12	15	9	6	f

حيث x هو الأجر [بالألف ريال سعودي] ، f هو عدد العمال .

(أ) أوجد المدى R للأجور .

(ب) اعرض البيانات السابقة باستخدام طريقة اللوحة الدائرية ، المدرج التكراري ، المضلع التكراري .

(ج) كون كلاً من الجدولين التكراري المتجمع الصاعد والتكراري المتجمع الهابط .

(د) ارسم المضلع التكراري المتجمع الصاعد ومنه قدر عدد العاملين الذين يحصلون على أجر :

(1) أقل من 88 ألف سنوياً

(2) 96 ألف سنوياً أو أكثر

(3) لا يقل عن 63 ألف سنوياً ولا يزيد عن 75 ألف سنوياً

(هـ) قدر قيمة الوسيط M للأجور .

في هذا المثال يمثل الأجر (بآلاف الريالات) المتغير x ويمثل عدد العمال f التكرار .

(أ) المدى R للأجور :

هو الفرق بين أكبر قيمة في البيانات [الحد الأعلى للفئة الأخير = 180] وأصغر قيمة فيها [الحد الأدنى للفئة الأولى = 50] ، وبالتالي يكون :

$$R = 180 - 50 = \underline{\underline{130}}$$

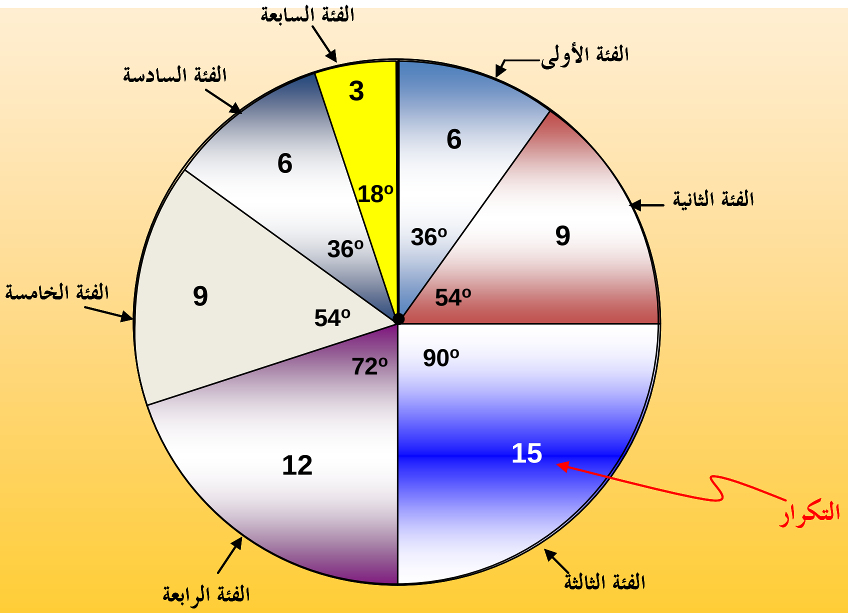
(ب) العرض البياني للبيانات : شكل (2–29) يبين طريقة اللوحة الدائرية لعرض البيانات السابقة وذلك بالاستعانة بالجدول المرافق للرسم .

ويبين جدول (2–23) البيانات اللازمة لرسم كلٍ من المدرج التكراري والمضلع التكراري (على رسمة واحدة) لتلك البيانات [شكل (2–30)] .

(ج) الجدول التكراري المتجمع الصاعد والجدول التكراري المتجمع الهابط :

جدول (2–24أ) يبين الجدول التكراري للبيانات السابقة ، ومنه يمكن تكوين كلٍ من الجدول التكراري المتجمع الصاعد [جدول (2–24ب)] وأيضاً الجدول التكراري المتجمع الهابط [جدول (2–24ج)] .

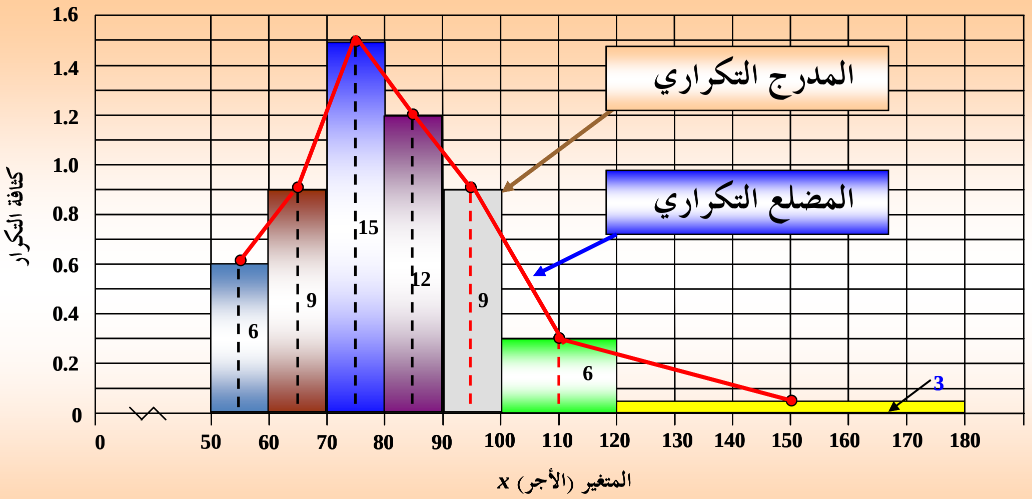
	الجدول التكراري		
الفئة	المتغير (الأجر) x	التكرار f	الزاوية المركزية
الأولى	$50 \le x < 60$	6	36^{o}
الثانية	$60 \le x < 70$	9	54^{o}
الثالثة	$70 \le x < 80$	15	90^{o}
الرابعة	$80 \le x < 90$	12	72^{o}
الخامسة	$90 \le x < 100$	9	54^{o}
السادسة	$100 \le x < 120$	6	36^{o}
السابعة	$120 \le x < 180$	3	18^{o}
		$\sum f = 60$	المجموع = 360^{o}



شكل (2–29) : العرض البياني بطريقة اللوحة الدائرية لبيانات مثال (2–8)

جدول (2–23) : البيانات اللازمة لرسم كلٍ من المدرج التكراري والمضلع التكراري لبيانات مثال (2–8)

	الجدول التكراري					
الفئة	المتغير (الأجر) x	التكرار f	طول الفئة c	مركز الفئة	كثافة التكرار	النقطة
الأولى	$50 \le x < 60$	6	10	55	0.6	(55 , 0.6)
الثانية	$60 \le x < 70$	9	10	65	0.9	(65 , 0.9)
الثالثة	$70 \le x < 80$	15	10	75	1.5	(75 , 1.5)
الرابعة	$80 \le x < 90$	12	10	85	1.2	(85 , 1.2)
الخامسة	$90 \le x < 100$	9	10	95	0.9	(95 , 0.9)
السادسة	$100 \le x < 120$	6	20	110	0.3	(110 , 0.3)
السابعة	$120 \le x < 180$	3	60	150	0.05	(150 , 0.05)
		$\sum f = 60$				



شكل (2–30) : المدرج التكراري والمضلع التكراري لبيانات مثال (2–8)

جدول (2–24) : الجدول التكراري (أ) والتكراري المتجمع الصاعد (ب) والتكراري المتجمع الهابط (ج) لبيانات مثال (2–8)

جدول (2–24أ)

الجدول التكراري		
التكرار f	المتغير (الأجر) x	الفئة
6	$50 \le x < 60$	الأولى
9	$60 \le x < 70$	الثانية
15	$70 \le x < 80$	الثالثة
12	$80 \le x < 90$	الرابعة
9	$90 \le x < 100$	الخامسة
6	$100 \le x < 120$	السادسة
3	$120 \le x < 180$	السابعة
$\sum f = 60$		

جدول (2–24ب)

الجدول التكراري المتجمع الصاعد		
التكرار المتجمع النسبي	التكرار المتجمع	المتغير x
0%	0	< 50
10%	6	< 60
25%	15	< 70
50%	30	< 80
70%	42	< 90
85%	51	< 100
95%	57	< 120
100%	60	< 180

جدول (2–24ج)

الجدول التكراري المتجمع الهابط		
التكرار المتجمع النسبي	التكرار المتجمع	المتغير x
100%	60	≥ 50
90%	54	≥ 60
75%	45	≥ 70
50%	30	≥ 80
30%	18	≥ 90
15%	9	≥ 100
5%	3	≥ 120
0%	0	≥ 180

(د) المضلع التكراري المتجمع الصاعد :

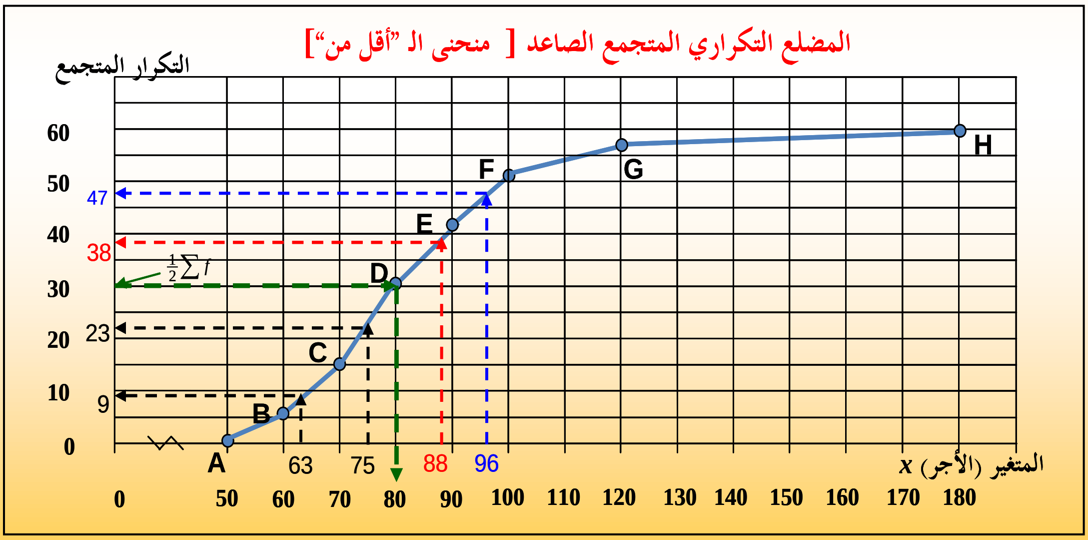
من الجدول التكراري المتجمع الصاعد للبيانات [جدول (2-24ب)] يمكن رسم المضلع التكراري المتجمع الصاعد وذلك باختيار محورين متعامدين : الأفقي (ويمثل المتغير x) والرأسي (ويمثل التكرار المتجمع) ، وباختيار مقياس رسم (تدريج) مناسب لكلٍ من المحورين يمكن توقيع النقاط :

$$A(50,0) \quad , \quad B(60,6) \quad , \quad C(70,15) \quad , \quad D(80,30)$$
$$E(90,42) \quad , \quad F(100,51) \quad , \quad G(120,57) \quad , \quad H(180,60)$$

ثم توصيلها بالمسطرة فنحصل على خط منكسر يكون هو المضلع التكراري المتجمع الصاعد للبيانات [شكل (2-31)] .

الجدول التكراري المتجمع الصاعد (2-24ب) موضحاً عليه النقاط اللازمة لرسم المضلع التكراري المتجمع الصاعد

الجدول التكراري المتجمع **الصاعد**		
النقطة	التكرار المتجمع	المتغير x
A(50 , 0)	0	< 50
B(60 , 6)	6	< 60
C(70 , 15)	15	< 70
D(80 , 30)	30	< 80
E(90 , 42)	42	< 90
F(100 , 51)	51	< 100
G(120 , 57)	57	< 120
H(180 , 60)	60	< 180



شكل (2-31) : المضلع التكراري المتجمع الصاعد لبيانات مثال (2-8)

ومن هذا المضلع يمكن تقدير الآتي :

(1) عدد العاملين الذين يحصلون على أقل من 88 ألف سنوياً حوالي : **38** عاملاً .

(2) عدد العاملين الذين يحصلون على 96 ألف سنوياً أو أكثر حوالي :

$60 - 47 = \mathbf{13}$

(3) عدد العاملين الذين يحصلون على أجر لا يقل عن *63* ألف ولا يزيد عن *75* ألف سنوياً حوالي :

$23 - 9 = \mathbf{14}$

(هـ) الوسيط M : هي قيمة x المناظرة لتكرار متجمع قدره $\frac{1}{2}\sum f$ [أي 30] : $M = \mathbf{80}$

(2–4) المزيد عن الفئات

يمكن إلقاء الضوء عن بعض المعلومات الخاصة بالفئات التي توضح أهمية تنظيم البيانات ووضعها على صورة فئات وكيفية اختيار تلك الفئات وذلك من خلال المثال التالي :

مثال (2–9) :

تم جمع البيانات التالية عن أوزان 200 طالب (لأقرب كيلوجرام) في إحدى الكليات بالجامعة :

67	72	74	77	67	73	72	68	70	75
55	74	79	75	70	73	74	73	73	76
66	75	78	81	88	65	69	75	76	72
78	81	67	74	84	71	80	71	81	78
67	66	63	48	68	80	75	82	71	86
73	78	77	62	77	82	60	65	77	72
68	80	75	57	75	64	78	70	70	78
81	65	75	73	64	59	75	76	67	72
74	76	72	69	69	67	71	69	83	70
68	74	68	78	73	69	77	75	50	76
67	70	66	77	73	79	65	63	64	66
72	67	78	67	77	71	80	74	78	74
71	71	73	70	66	58	76	72	66	74
62	69	74	75	70	68	72	73	73	72
71	77	83	79	71	74	74	63	67	73
75	72	69	66	75	79	79	94	82	64
73	73	73	58	70	73	73	84	76	76
85	74	73	61	68	66	85	78	65	73
75	78	62	78	78	71	76	74	68	72
73	69	76	80	69	81	64	75	74	87

المطلوب تكوين جدول تكراري لهذه البيانات .

في البيانات السابقة (**البيانات الخام**) ، المتغير x هو الوزن (لأقرب كيلوجرام) وهي بالصورة المعطاة على هيئة بيانات كمية متقطعة قد يصعب من خلالها الرد على بعض الأسئلة مثل " **ما هو عدد الطلاب الذين يتراوح وزنهم بين 51 ، 83 كيلوجرام ؟**" أو "**ما هي نسبة الطلاب الذين تزيد أوزانهم على 63 كيلوجرام ، وتقل عن 75 كيلوجرام ؟**" ، واضح أنه من الصعب الرد على مثل هذه الأسئلة ما لم يتم وضع البيانات في صورة أبسط من

الصورة المعطاة ، وتعتمد إحدى الطرق الإحصائية المستخدمة في ترتيب البيانات الخام بصورة بسيطة على تكوين التوزيع (أو الجدول) التكراري بتتبع الخطوات التالية :

- تحديد المدى R للبيانات وهو الفرق بين أكبر قيمة وأصغر قيمة في البيانات الخام ، ومن الممكن استنتاج أن أكبر قيمة في البيانات هي 94 وأقل قيمة 48 ، وبالتالي يكون المدى :

$$R = 94 - 48 = 46$$

- تقسيم هذا المدى إلى عدد مناسب من الفئات على أن يكون ما بين 5 ، 20 [فإذا استعملنا أكثر من 20 فئة افتقدنا إلى البساطة التي ننشدها عند عمل التوزيع التكراري ، وإذا استعملنا أقل من 5 فئات ضاعت الكثير من تفصيلات البيانات الأصلية] ، لنفرض أننا قررنا استخدام 10 فئات متساوية الطول ، فبالتالي يكون طول كل فئة c يساوي المدى مقسوماً على عدد الفئات ، أي أن :

$$c = \frac{46}{10} = 4.6$$

ويتوجب علينا (في هذا المثال) أن نأخذ أول عدد صحيح تالي ليكون طولاً للفئة ، أي سنعتبر أن $c = 5$ ، وبالتالي تكون لدينا الفئات المبينة بالجدول (2–25) .

جدول (2–25) : الجدول التكراري للبيانات الخام لمثال (2–9)

الفئة	المتغير (الوزن) x	تفريغ البيانات (العلامات)	التكرار f	طول الفئة c	مركز الفئة x_0
الأولى	45 – 49	|	1	5	47
الثانية	50 – 54	|	1	5	52
الثالثة	55 – 59	||||	4	5	57
الرابعة	60 – 64	卌 卌 |||	13	5	62
الخامسة	65 – 69	卌 卌 卌 卌 卌 卌 卌 卌	40	5	67
السادسة	70 – 74	卌 卌 卌 卌 卌 卌 卌 卌 卌 卌 卌 卌 卌	65	5	72
السابعة	75 – 79	卌 卌 卌 卌 卌 卌 卌 卌 卌 卌 ||	52	5	77
الثامنة	80 – 84	卌 卌 卌 |||	18	5	82
التاسعة	85 – 89	卌	5	5	87
العاشرة	90 - 94	|	1	5	92
			$\sum f = 200$		

- نبدأ في تفريغ البيانات طبقاً للفئات التي حددناها ونحدد تكرار كل فئة كما هو مبين بالجدول (2–25) فنحصل على التوزيع (أو الجدول) التكراري للبيانات الخام المعطاة سابقاً ، ومن الجدول

التكراري (2–25) نلاحظ أن أي فئة [مثل الفئة الأولى] لها حدان : الحد الأدنى (= 45) وهو القيمة الصغرى في الفئة ، والحد الأعلى (= 49) وهو القيمة العظمى في الفئة ، ويكون **"طول كل فئة هو الحد الأعلى للفئة – الحد الأدنى لها + وحدة قياس"** ، و **"مركز الفئة هو متوسط حديها الأدنى والأعلى"** [جدول (2–25)] [وحدة القياس تتوقف على درجة التقريب المستخدمة في قياس البيانات ، ففي مثالنا السابق الأوزان مقربة لأقرب عدد صحيح ، إذن وحدة القياس هنا هي الواحد الصحيح].

- ولكن من الجدير بالذكر أن أوزان الطلاب معطاة على شكل بيانات كمية متقطعة ، إلا أنه في حقيقة الأمر فقد تكون قيمة الوزن 70 (مقربة لأقرب عدد صحيح) هي أي قيمة أكبر من أو تساوي 69.5 إلى ما قبل 70.5 [فعلى سبيل المثال ، القيم 70.49 , 70.2 , 69.91 , 69.5 كلها تُقرب إلى 70 وذلك لأقرب عدد صحيح] ، لذا يُفضل في الإحصاء أن نعبر عن الفئات **بحدودها الحقيقية** وذلك بإضافة نصف وحدة قياس للحد الأعلى وإنقاص نصف وحدة قياس من الحد الأدنى للفئة فتكون الحدود الحقيقية للفئة السادسة هي 69.5 – 74.5 ، وللفئة السابعة 74.5 – 79.5 وهكذا ، وبالتالي يأخذ الجدول التكراري (2–25) الشكل الموضح بالجدول (2–26) ، ويكون **"طول أي فئة هو الفرق بين حدها الأعلى الحقيقي وحدها الأدنى الحقيقي"** و **"مركز كل فئة هو متوسط حديها الحقيقيين الأعلى والأدنى"** [جدول (2–26)] .

جدول (2–26) : الجدول التكراري للبيانات الخام لمثال (2–9)

مركز الفئة x_0	طول الفئة c	التكرار f	تفريغ البيانات (العلامات)	المتغير (الوزن) x	الفئة
47	5	1	|	44.5 – 49.5	الأولى
52	5	1	|	49.5 – 54.5	الثانية
57	5	4	||||	54.5 – 59.5	الثالثة
62	5	13	𝍸 𝍸 |||	59.5 – 64.5	الرابعة
67	5	40	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸	64.5 – 69.5	الخامسة
72	5	65	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸	69.5 – 74.5	السادسة
77	5	52	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 ||	74.5 – 79.5	السابعة
82	5	18	𝍸 𝍸 𝍸 |||	79.5 – 84.5	الثامنة
87	5	5	𝍸	84.5 – 89.5	التاسعة
92	5	1	|	89.5 – 94.5	العاشرة
		$\sum f = 200$			

ويعيب الأسلوب السابق أن تساوي الحد الحقيقي الأعلى لإحدى الفئات [الثالثة مثلاً] مع الحد الأدنى للفئة التالية [الرابعة] قد يحدث بعض الغموض ، فإذا كان لدينا طالب وزنه 59.5 فإنه يكون من الصعب تقرير ما إذا كانت تلك القيمة منتمية للفئة الثالثة أم الرابعة ، لكن هذا الغموض يمكن أن يتم تلافيه بأن نكون حريصين على ألا تتطابق الحدود الحقيقية للفئات مع إحدى القيم الفعلية (المشاهدات) .

- أيضاً بالإمكان صياغة الفئات السابقة على الصورة المبينة بالجدول (2–27) [وهو ما استخدمناه في الجزء الأول من هذا الفصل] ، حيث يتطابق تعريف طول الفئة ومركز الفئة مع التعريف المستخدم في جدول (2–26) ، إلا أن هناك اختلافات بسيطة في قيم مراكز الفئات .

جدول (2–27) : الجدول التكراري بالأسلوب المتبع في الجزء الأول من هذا الفصل

الفئة	المتغير x (الوزن)	التكرار f (العدد)	طول الفئة c	مركز الفئة x_0
الأولى	$45 \le x < 50$	1	5	47.5
الثانية	$50 \le x < 55$	1	5	52.5
الثالثة	$55 \le x < 60$	4	5	57.5
الرابعة	$60 \le x < 65$	13	5	62.5
الخامسة	$65 \le x < 70$	40	5	67.5
السادسة	$70 \le x < 75$	65	5	72.5
السابعة	$75 \le x < 80$	52	5	77.5
الثامنة	$80 \le x < 85$	18	5	82.5
التاسعة	$85 \le x < 90$	5	5	87.5
العاشرة	$90 \le x < 95$	1	5	92.5
		$\sum f = 200$		

- وأخيراً تجدر الإشارة هنا إلى أنه قد يصعب علينا استخدام فئات متساوية الطول فنضطر لاستخدام فئات ذات أطوال مختلفة [كما في حالة تقديرات الطلاب] .

ملخص للدرس الثالث [الباب الثاني : التوزيعات التكرارية (بيانات متصلة)]

البيانات الكمية المتصلة :

- في حالة البيانات الكمية المتصلة تكون قيم المتغير x هنا معطاة على صورة فترات أو ما يُسمى بـ **الفئات** .
- لكل فئة حدان : حد **أدنى** وحد **أعلى** ، والحد الأدنى لفئة هو الحد الأعلى للفئة السابقة لها ، والحد الأعلى لفئة هو الحد الأدنى للفئة التالية ، أي أن الفئات متصلة ولا فراغات بينها
- لكل فئة **طول** c حيث : **طول الفئة = حدها الأعلى – حدها الأدنى ،** وأطوال الفئات (بوجهٍ عام) تكون غير متساوية .
- لكل فئة **مركز** x_0 وهو قيمة المتغير الواقعة في منتصف تلك الفئة ويتحدد كالآتي :

مركز أي فئة = $\dfrac{\text{حد الفئة الأدنى + حدها الأعلى}}{2}$

عرض البيانات المتصلة بواسطة الجداول :

أولاً : الجدول (التوزيع) التكراري : وهو جدول يوضح فئات المتغير x مع تكرار كل فئة [أي عدد القيم الواقعة في تلك الفئة .

ثانياً : الجدول (التوزيع) التكراري النسبي: حيث يُضاف للجدول التكراري السابق عمود ثالث يوضح التكرار النسبي لكل فئة (كنسبة عادية أو نسبة مئوية) حيث :

التكرار النسبى لفئة ما = $\dfrac{\text{تكرار الفئة}}{\text{مجموع التكرارات}}$

ثالثاً : الجدول (التوزيع) التكراري المتجمع الصاعد والمتجمع الهابط(النازل) :

يُعرف **التكرار المتجمع الصاعد** المناظر لقيمة معية a لمتغير x هو مجموع تكرارات جميع قيم المتغير **الأقل** من a .

ويُعرف **التكرار المتجمع الهابط** المناظر لقيمة معية a لمتغير x هو مجموع تكرارات جميع قيم **الأكبر من أو تساوي** a .

وعلى هذا الأساس يمكن تكوين ما يُسمى **بالجدول التكراري المتجمع الصاعد** **والجدول التكراري المتجمع الهابط** كما هو مبين ، مع مراعاة الآتي :

- في الجدول التكراري المتجمع **الصاعد** تكون عناصر العمود الأول [عمود المتغير x] هي **الحدود الدنيا للفئات** مسبوقة بعلامة "**أقل من**" والعنصر الأخير هو الحد الأعلى للفئة الأخيرة مسبوقاً أيضاً بعلامة "**أقل من**" ، نفس الشيء في الجدول التكراري المتجمع **الهابط** لكن العلامة تكون "**أكبر من أو تساوي**" بدلاً من "**أقل من**" .

- في الجدول التكراري المتجمع **الصاعد** يزداد التكرار المتجمع [العمود الثاني] كلما اتجهنا لأسفل الجدول بادئين بالقيمة 0 (أعلى الجدول) ثم نضيف تكرارات الفئات فئة تلو الأخرى كلما اتجهنا للأسفل حتى ننتهي بالقيمة $\sum f$ [مجموع التكرارات (أسفل الجدول)] ، أما في الجدول التكراري المتجمع **الهابط** يزداد التكرار المتجمع [العمود الثاني] كلما اتجهنا لأعلى الجدول بادئين بالقيمة 0 (أسفل الجدول) ثم نضيف تكرارات الفئات فئة تلو الأخرى كلما اتجهنا لأعلى الجدول حتى ننتهي بالقيمة $\sum f$ [مجموع التكرارات (أعلى الجدول)] .
ويمكن إضافة [لأي من الجدولين] عمود يمثل "**التكرار المتجمع النسبي**" حيث :

$$\text{التكرار المتجمع النسبي} = \frac{\text{التكرار المتجمع}}{\text{مجموع التكرارات}}$$

- **ملحوظة الجداول (التوزيعات) التكرارية المفتوحة :**
هي جداول إما أن تكون مفتوحة من أسفل أو من أعلى أو من الطرفين .

مفتوح من أسفل

f	x
20	$x < 6$
25	$6 \le x < 12$
35	$12 \le x < 15$
18	$15 \le x < 18$

الحد الأدنى للفئة الأولى غير معلوم

مفتوح من أعلى

f	x
20	$6 \le x < 12$
25	$12 \le x < 15$
35	$15 \le x < 18$
18	$x \ge 18$

الحد الأعلى للفئة الأخيرة غير معلوم

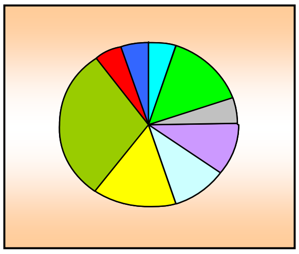
مفتوح من الطرفين

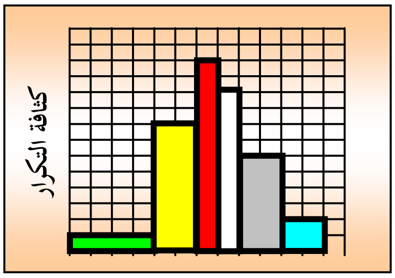
f	x
20	$x < 6$
25	$6 \le x < 12$
35	$12 \le x < 15$
18	$x \ge 15$

الحدان الأدنى (للفئة الأولى) والأعلى (للفئة الأخيرة) غير معلومين

عرض البيانات المتصلة بيانياً :

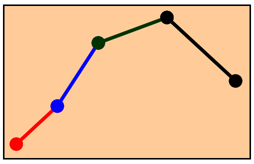
طريقة الدائرة : مثل حالة البيانات المنفصلة حيث تُمثل كل **فئة** من الفئات بقطاع من دائرة وذلك طبقاً لتكرارها .



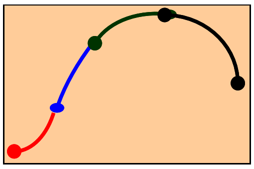


طريقة المدرج التكراري : حيث تُمثل الفئات بمستطيلات متلاصقة بحيث يمثل كل مستطيل إحدى الفئات ، بحيث تقع قاعدة المستطيل (الممثل لفئة ما) على المحور الأفقي [محور المتغير] وممتدة بين الحد الأدنى للفئة وحدها الأعلى [أي **طول قاعدة المستطيل يساوي طول الفئة**] **وارتفاعه هو كثافة تكرار الفئة** **ومساحته هي تكرار الفئة** .

$$\text{كثافة التكرار لفئة} = \frac{\text{تكرار الفئة}}{\text{طول الفئة}}$$

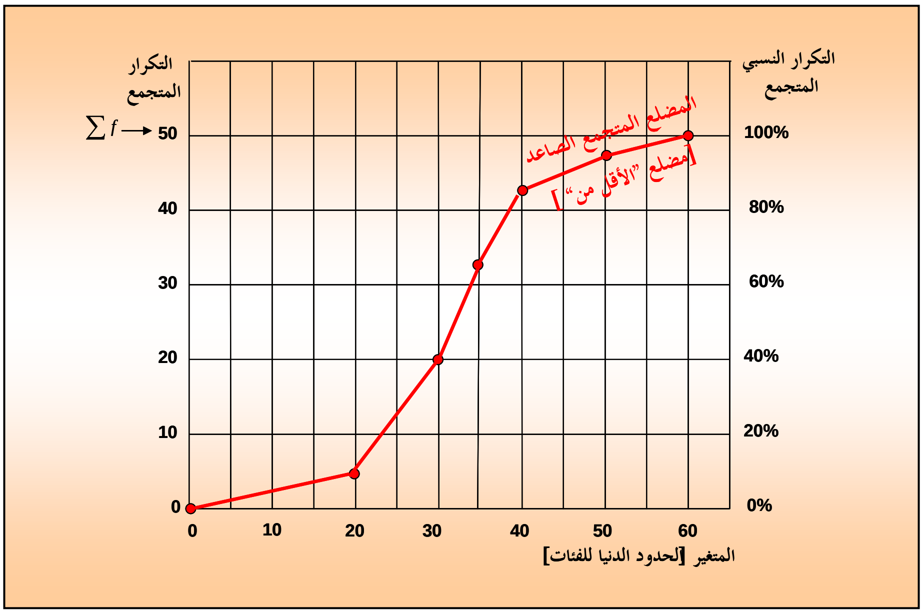


طريقة المضلع التكراري : حيث تُمثل كل فئة من الفئات بنقطة إحداثياتها هي **مركز الفئة وكثافة تكرارها** ثم نقوم بتوصيل هذه النقاط بخط منكسر (بواسطة المسطرة) .

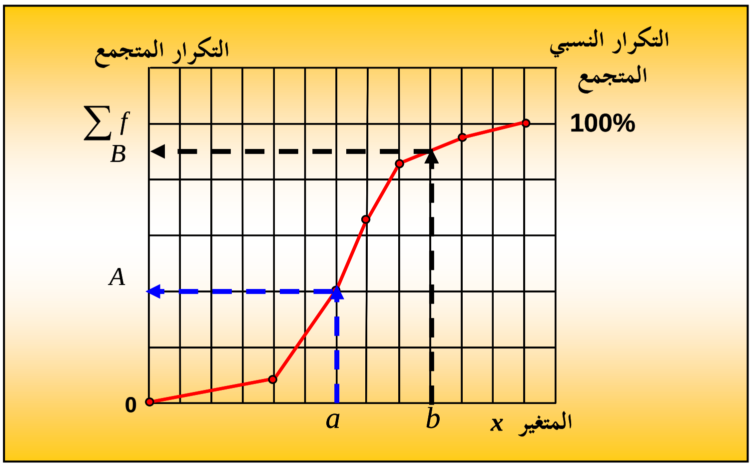


طريقة المنحنى التكراري : حيث تُمثل كل فئة من الفئات بنقطة إحداثياتها هي **مركز الفئة وكثافة تكرارها** ثم نقوم بتوصيل هذه النقاط بخط ممهد (باليد) .

طريقة المضلع التكراري المتجمع الصاعد: حيث تُمثل كل فئة من الفئات بنقطة إحداثياتها هي **الحد الأدنى للفئة والتكرار المتجمع الصاعد المناظر** ثم نقوم بتوصيل هذه النقاط بخط منكسر (بواسطة المسطرة) ، وهنا لابد من تكوين الجدول التكراري المتجمع الصاعد . وأحياناً يُسمى هذا المضلع بـ "**مضلع الأقل من**" .



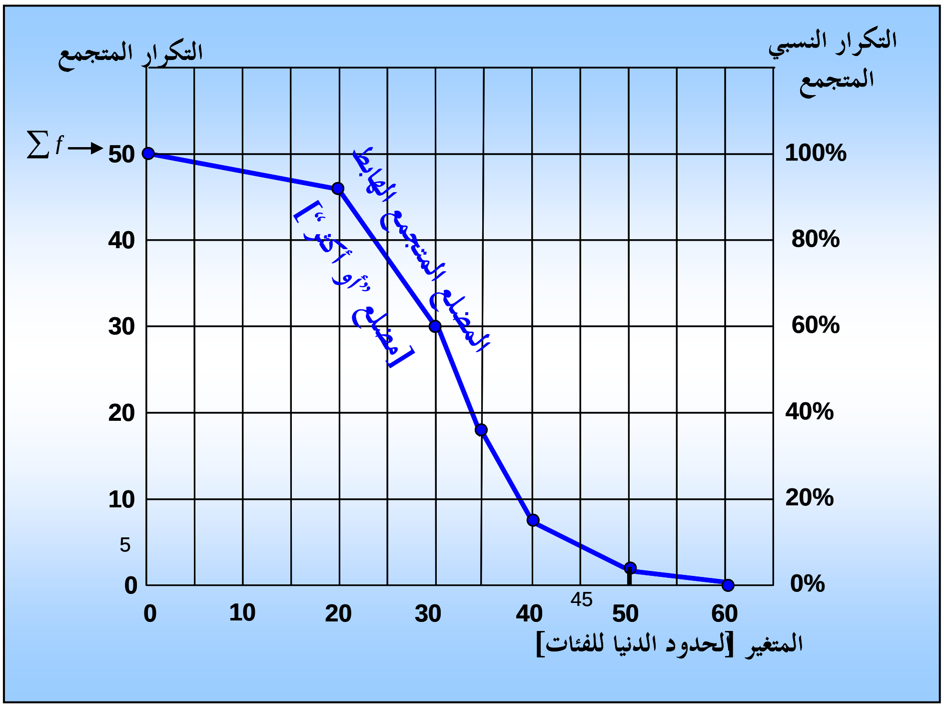
ويفيد المضلع التكراري المتجمع الصاعد في الرد على العديد من الأسئلة مثل :



- حسب التكرار المتجمع المناظر لـ "$x < a$" . الإجابة A
- احسب التكرار المتجمع المناظر لـ "$x \geq a$" : الإجابة : مجموع التكرارات مطروح منها A .
- احسب التكرار المتجمع المناظر لـ "$a \leq x < b$" : الإجابة B مطروح منها A .

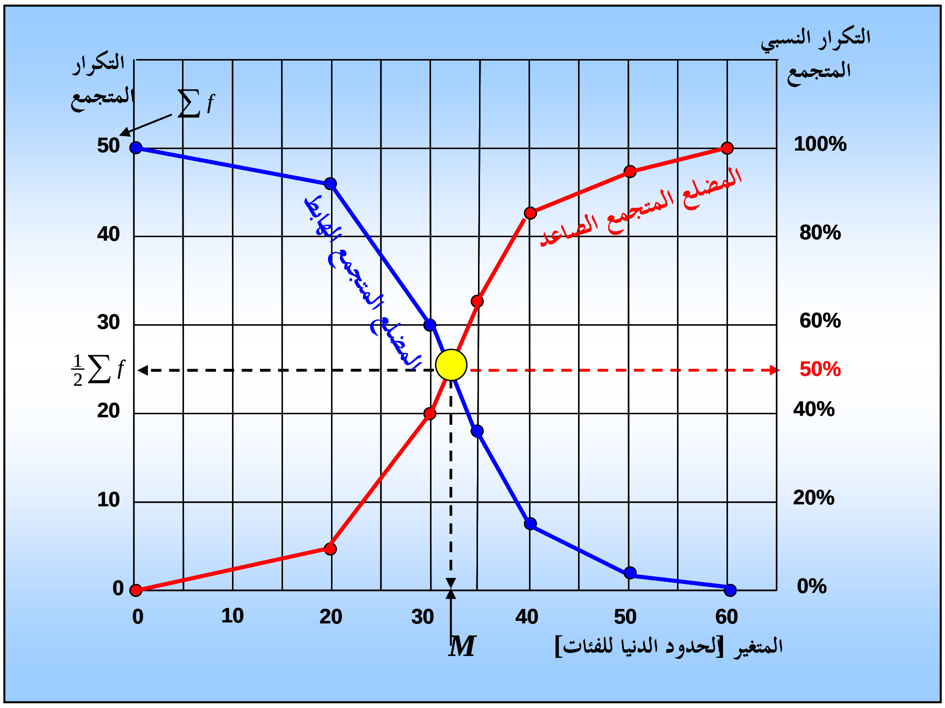
طريقة المضلع التكراري المتجمع الهابط: وهو مشابه للمضلع التكراري المتجمع الصاعد مع الاختلاف أن كل فئة تُمثل بنقطة إحداثياتها هي **الحد الأدنى للفئة والتكرار المتجمع الهابط المناظر** ثم نقوم بتوصيل هذه النقاط بخط منكسر (بواسطة المسطرة) وهنا لابد من تكوين الجدول التكراري المتجمع الهابط .

وأحياناً يُسمى هذا المضلع بـ "**مضلع الأكبر من أو يساوي**" .



ويتقاطع المضلعان التكراريان الصاعد والهابط في نقطة تكون قيمة المتغير المناظرة لها هي **الوسيط** M للبيانات ، وهي قيمة :

- **تقسم مجموعة البيانات إلى مجموعتين متساويتين في العدد .**
- **يناظرها تكرار متجمع قدره** $\frac{1}{2}\sum f$
- **يناظرها تكرار متجمع نسبي قدره** 50% .



تدريب عملي

البيانات التالية تمثل توزيع عدد من الشركات وفقاً لأرباحها في العام الماضي [بالمليون ريال سعودي] :

فئات الربح	2 -	5 -	8 -	11 -	14 - 17
عدد الشركات	5	13	34	24	7

المطلوب : رسم :

* كلٍ من المدرج التكراري ، المضلع التكراري ، والمنحنى التكراري [على نفس الرسمة] .
* كلٍ من المضلع التكراري المتجمع الصاعد والمضلع التكراري المتجمع الهابط [على نفس الرسمة] .
* اللوحة الدائرية .

أيضاً **مطلوب** حساب :

* وسيط الأرباح .
* عدد الشركات التي حققت أرباحاً أقل من 10 مليون .
* نسبة الشركات التي حققت أرباحاً 8 مليون فأكثر .
* النسبة المئوية للشركات التي حققت أرباحاً بين 6 ، 9 مليون .

تدريبات (3)

الإجابة النهائية لجميع التمرينات موجودة في نهاية التدريب

اختر الإجابة الصحيحة

(1) التكرار النسبي لفئة من الفئات هو :
- (أ) النسبة بين الحد الأعلى للفئة ومجموع التكرارات
- (ب) خارج قسمة تكرار الفئة على طولها
- (ج) نسبة تكرار الفئة إلى مجموع التكرارات
- (د) النسبة بين الحد الأدنى للفئة ومجموع التكرارات

(2) في المدرج التكراري لبيانات متصلة ذات فئات غير متساوية تكون مساحة أي مستطيل من المستطيلات هي :
- (أ) تكرار الفئة التي يمثلها المستطيل
- (ب) التكرار النسبي للفئة التي يمثلها المستطيل
- (ج) كثافة تكرار الفئة التي يمثلها المستطيل
- (د) طول الفئة التي يمثلها المستطيل

(3) في المدرج التكراري لبيانات متصلة ذات فئات غير متساوية تكون طول قاعدة أي مستطيل من المستطيلات هي :
- (أ) تكرار الفئة التي يمثلها المستطيل
- (ب) التكرار النسبي للفئة التي يمثلها المستطيل
- (ج) كثافة تكرار الفئة التي يمثلها المستطيل
- (د) طول الفئة التي يمثلها المستطيل

(4) في المدرج التكراري لبيانات متصلة ذات فئات غير متساوية يكون ارتفاع أي مستطيل من المستطيلات هو :
- (أ) تكرار الفئة التي يمثلها المستطيل
- (ب) التكرار النسبي للفئة التي يمثلها المستطيل
- (ج) كثافة تكرار الفئة التي يمثلها المستطيل
- (د) طول الفئة التي يمثلها المستطيل

(5) في المدرج التكراري لبيانات متصلة تكون المستطيلات الممثلة للفئات :

(أ) متلاصقة تماماً (أي لا مسافات بينها)

(ب) منفصلة عن بعضها

(ج) متداخلة

(د) فوق بعضها

(6) في المضلع التكراري تُمثل كل فئة بنقطة إحداثياتها :

(أ) الحد الأدنى للفئة والتكرار المتجمع لجميع قيم المتغير الأقل من هذا الحد .

(ب) الحد الأدنى للفئة والتكرار المتجمع لجميع قيم المتغير الأكبر من أو تساوي هذا الحد .

(ج) مركز الفئة وكثافة تكرارها .

(د) مركز المستطيل الممثل لتلك الفئة

(7) في المضلع التكراري المتجمع الصاعد تُمثل كل فئة بنقطة إحداثياتها :

(أ) الحد الأدنى للفئة والتكرار المتجمع لجميع قيم المتغير الأقل من هذا الحد .

(ب) الحد الأدنى للفئة والتكرار المتجمع لجميع قيم المتغير الأكبر من أو تساوي هذا الحد .

(ج) مركز الفئة وكثافة تكرارها .

(د) مركز المستطيل الممثل لتلك الفئة

(8) في المضلع التكراري المتجمع الهابط تُمثل كل فئة بنقطة إحداثياتها :

(أ) الحد الأدنى للفئة والتكرار المتجمع لجميع قيم المتغير الأقل من هذا الحد .

(ب) الحد الأدنى للفئة والتكرار المتجمع لجميع قيم المتغير الأكبر من أو تساوي هذا الحد .

(ج) مركز الفئة وكثافة تكرارها .

(د) مركز المستطيل الممثل لتلك الفئة

خاص بالأسئلة من (9) إلى (14) : في التوزيع التكراري المبين :

الفئة	المتغير x	التكرار f
الأولى	$0 \le x < 20$	10
الثانية	$\ldots \le x < \ldots$	15
الثالثة	$30 \le x < \ldots$	20
الرابعة	$50 \le x < 60$	5

خاص بالمسائل من 9 إلى 14

(9) مجموع التكرارات $\sum f$ يساوي

(أ) 100 (ب) 200

(ج) 1 (د) 50

(10) التكرار النسبي للفئة الرابعة يساوي :

(أ) 0.2 (ب) 0.3

(ج) 0.1 (د) 0.4

(11) مركز الفئة الأولى عند x تساوي :

(أ) 0 (ب) 10

(ج) 15 (د) 20

(12) كثافة تكرار الفئة الرابعة تساوي :

(أ) 0.1 (ب) 0.5

(ج) 5 (د) 55

(13) الحد الأعلى للفئة الثالثة هو :

(أ) 20 (ب) 30

(ج) 40 (د) 50

(14) مركز الفئة الثانية عند x تساوي :

(أ) 25 (ب) 30

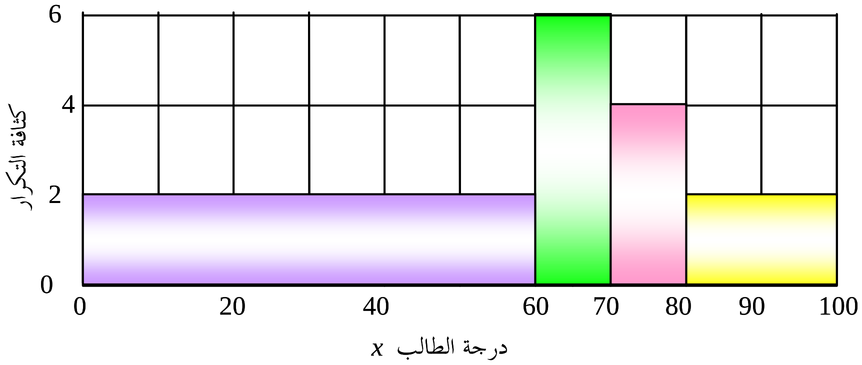
(ج) 35 (د) 15

خاص بالأسئلة من (15) إلى (20) :

المدرج التكراري المبين يوضح الدرجة x لعدد من الطلاب في مقرر مبادئ الإحصاء مقسمين على 4 فئات ، من هذا المدرج يمكن استنتاج الآتي :

خاص بالمسائل من 15 إلى 20

الفئة	(1)	(2)	(3)	(4)
الدرجة	$0 \le x < 60$	$60 \le x < 70$	$70 \le x < 80$	$80 \le x < 100$



خاص بالمسائل من

من 15 إلي 20

(15) العدد الكلي للطلاب :

(أ) 120 (ب) 180

(ج) 220 (د) 260

(16) عدد الطلاب الراسبين [الحاصلين على درجة أقل من 60] :

(أ) 40 (ب) 60

(ج) 100 (د) 120

(17) عدد الطلاب الحاصلين على 80 فأكثر :

(أ) 40 (ب) 60

(ج) 100 (د) 120

(18) عدد الطلاب الحاصلين على تقدير $C+$ [أكثر من 75 وأقل من 80] :

(أ) 120 (ب) 60

(ج) 40 (د) 20

(19) عدد الطلاب الناجحين والحاصلين على تقدير B على الأكثر [أكثر من 60 وأقل من 80] :

(أ) 40 (ب) 60

(ج) 100 (د) 120

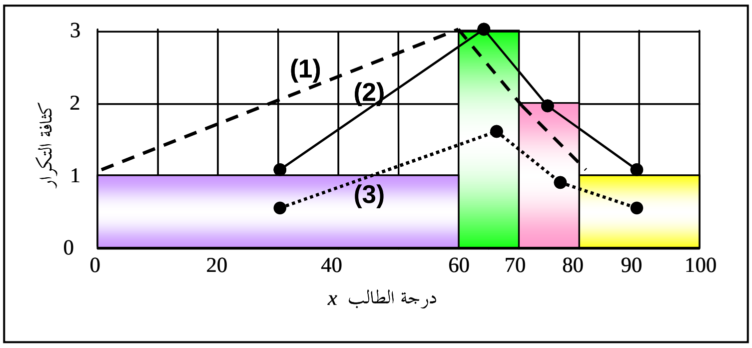
(20) الخط المنكسر الذي يمثل المضلع التكراري للبيانات السابقة :

(أ) الخط المنكسر (1)

(ب) الخط المنكسر (2)

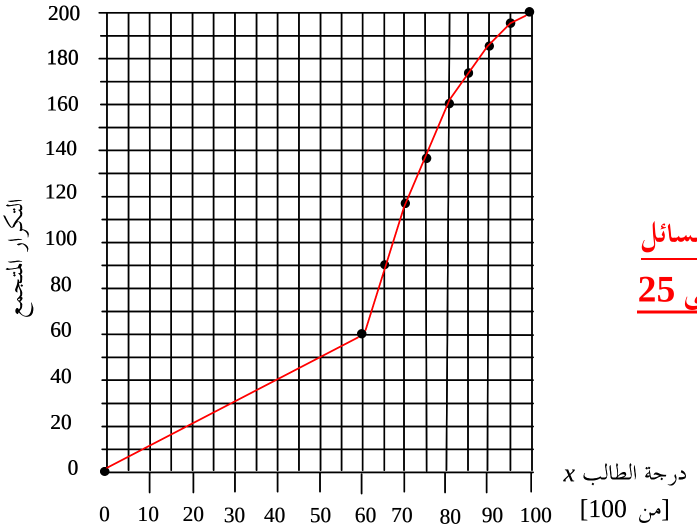
(ج) الخط المنكسر (3)

(د) ليس أي خط مما سبق .



خاص بالأسئلة من (21) إلى (25) :

الشكل المرافق يبين المضلع التكراري المتجمع الصاعد لدرجات عدد من الطلاب في مقرر مبادئ الإدارة ، من هذا الشكل يمكن أن نستنتج ن :



خاص بالمسائل

من 21 إلي 25

(21) العدد الكلي للطلاب هو :

(أ) 50 (ب) 100

(ج) 150 (د) 200

(22) الوسيط M لدرجات الطلاب يقع بين :

(أ) 40 , 45 (ب) 50 , 55

(ج) 65 , 70 (د) 75 , 80

(23) عدد الطلاب الحاصلات على درجة أقل من 40 هو :

(أ) 20% (ب) 40

(ج) 160 (د) 80%

(24) النسبة المئوية للطلاب الحاصلين على تقدير $D+$ على الأقل [أي على درجة 65 فأكثر] هي :

(أ) 55% (ب) 45

(ج) 40% (د) 65%

(25) عدد الطلاب الناجحين والحاصلين على درجة أقل من 80 هو :

(أ) 60 (ب) 80

(ج) 100 (د) 120

الإجابة : (1) ج (2) أ (3) د (4) ج (5) أ
(6) ج (7) أ (8) ب (9) د (10) ج
(11) ب (12) ب (13) د (14) أ (15)
د
(16) د (17) أ (18) د (19) ج (20) ب
(21) د (22) ج (23) ب (24) أ (25) ج

عناصر الدرس

الباب الثالث : مقاييس النزعة المركزية

- المتوسطات ومقاييس النزعة المركزية
- أهمية مقاييس النزعة المركزية
- الوسط الحسابي

الباب الثالث : مقاييس النزعة المركزية

(3–1) المتوسطات ومقاييس النزعة المركزية

المتوسط هو قيمة نموذجية يمكن أن تمثل مجموعة من البيانات بحيث تعطي دلالة معينة لتلك البيانات ، بمعنى أنه عندما ينظر الباحث (أو القارئ لتلك البيانات) ويريد أن يبحث عن شيء يربط هذه البيانات فإن تلك المتوسطات يمكن أن تعطيه بعضاً مما يريده .

وحيث أن مثل هذه القيم (المتوسطات) تميل إلى الوقوع في المركز داخل مجموعة البيانات (عند ترتيبها حسب قيمها) ، فإن هذه المتوسطات تُسمى أيضاً **بمقاييس النزعة المركزية** .

وهناك صور عديدة من هذه المقاييس وإن كان الأكثر شيوعاً :

- **الوسط الحسابي (أو باختصار الوسط) .**
- **الوسيط**
- **المنوال (أو الشائع)**

وغيرها من المقاييس مثل الوسط الهندسي والوسط التوافقي ، وكل منها له مميزاته وعيوبه وهذا يعتمد على البيانات والهدف من استخدامه .

وإلى جانب كونه ممثلاً لمجموعة البيانات يجب أيضاً أن تتوافر في المتوسط عدة شروط ، منها :

- أن يمكن تحديد قيمته بالضبط وتكون عملية حسابه سهلة إلى حد كبير .
- أن يأخذ في الاعتبار جميع البيانات .

ومن الجدير بالذكر أن بعض هذه المقاييس يمكن تحديدها حسابياً بسهولة ، وبعضها يمكن تحديدها **بيانياً** بسهولة ، والبعض يمكن تحديده **حسابياً وبيانياً** بسهولة ، لكننا في هذا المقرر سنكتفي بالطريقة **الأبسط** (للطالب) عند تحديد هذه المقاييس ، وهذه الطريقة الأبسط ستختلف من مقياس لآخر .

(3–2) أهمية حساب مقاييس النزعة المركزية

عند معرفتنا بتلك المتوسطات (مقاييس النزعة المركزية) يصبح أمامنا فرصة كبيرة لأن :

- ننظر لمتوسط مجموعة من البيانات لنعرف الكثير عن خصائص تلك المجموعة .
- نعقد مقارنة بين عدة مجموعات من البيانات في وقت واحد وذلك من خلال مقارنة متوسطات تلك المجموعات بعضها ببعض .

(3-3) الوسط الحسابي

(3-3-1) تعريف الوسط الحسابي

يُعرف الوسط الحسابي [وسنرمز له بالرمز $\bar{x}$] لمجموعة من البيانات $x_1 , x_2 , \cdots\cdots , x_n$ [قيم المتغير x وعددها n] كالآتي :

$$\text{الوسط الحسابي} = \frac{\text{مجموع قيم البيانات}}{\text{عددها}}$$

أي أن $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ ، ويمكن (للتبسيط) كتابة هذه العلاقة على الصورة

$$\boxed{\bar{x} = \frac{\sum x}{n}} \qquad (3\text{-}1)$$

حيث $\sum x$ تمثل مجموع القيم [وهي اختصار للتعبير الرياضي الدقيق $\sum_{i=1}^{n} x_i$] ، n هو عدد القيم .

مثال (3-1) : درجات خمسة طلاب في مقرر ما [الدرجة العظمى 20] هي : 10 , 12 , 7 , 2 , 9 . **أوجد** الوسط الحسابي لدرجاتهم .

الحل : باستخدام العلاقة (3-1) :

$$\bar{x} = \frac{\sum x}{n} = \frac{9+2+7+12+10}{5} = \frac{40}{5} = 8$$

أي أن الوسط الحسابي لدرجات الطلاب الخمسة هو 8 درجات [من 20] .

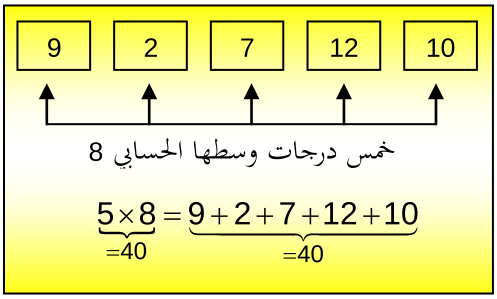
وعلى ضوء هذا المثال البسيط السابق يمكن ملاحظة الملاحظات العامة التالية :

- ملحوظة (1) : يمكن تحديد قيمة الوسط الحسابي بالضبط .
- ملحوظة (2) : طريقة تحديدة سهلة .
- ملحوظة (3) : يأخذ في الاعتبار جميع البيانات .
- ملحوظة (4) : لا يتأثر بترتيب البيانات .
- ملحوظة (5) : لا يُشترط أن يكون الوسط الحسابي عدداً صحيحاً ولا يُشترط أن يكون إحدى قيم البيانات ولكنه قيمة تقع بين أقل قيمة في البيانات وأكبر قيمة فيها .
- ملحوظة (6) : حاصل ضرب قيمة الوسط الحسابي في عدد البيانات = مجموع قيم البيانات ، وهذا واضح من تعريف الوسط الحسابي حيث أن :

$$\bar{x} = \frac{\sum x}{n} \quad \Rightarrow \quad n\bar{x} = \sum x$$

فمثلاً في مثال (1-3) :

وبالتالي يكون الوسط الحسابي هو تلك قيمة التي إذا أعطيت لكل مفرد من مفردات الظاهرة لكان مجموع القيم الجديدة مساويا للمجموع الفعلي للقيم الأصلية للظاهرة



$$\underbrace{5\times8}_{=40}=\underbrace{9+2+7+12+10}_{=40}$$

- ملحوظة (7) : يتأثر **بالقيم المتطرفة** في البيانات [كما يتضح من السؤالين التاليين] .

س 1 : احسب الوسط الحسابي للقيم :

10 , 15 , 12 , 13 , **9**

ج 1 :

$$\frac{10+15+12+13+9}{5}=\frac{59}{5}=11.8$$

س 2 : احسب الوسط الحسابي للقيم :

10 , 15 , 12 , 13 , **900**

ج 2 :

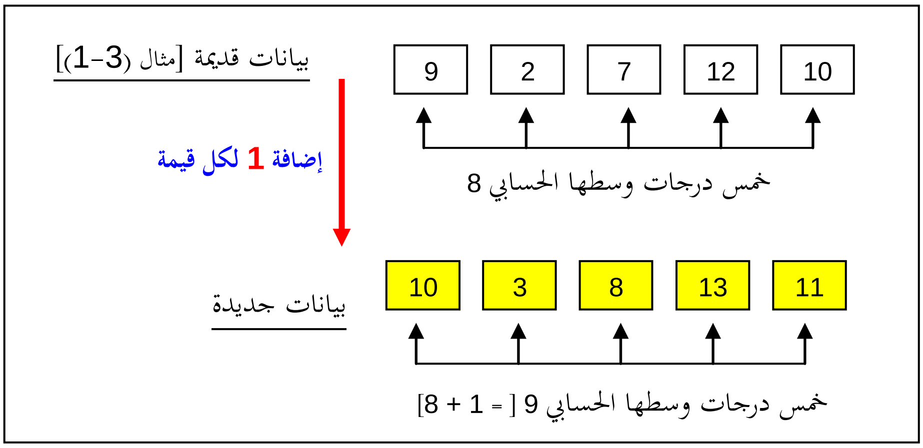
$$\frac{10+15+12+13+900}{5}=\frac{950}{5}=190$$

هل لاحظت تأثر قيمة الوسط الحسابي عندما تغيرت القيمة 9 إلى قيمة متطرفة 900 [بعيدة عن القيم الأخرى] .

- ملحوظة (8) : إذا أضفنا عدد ثابت c لكل قيمة من قيم البيانات ، فإن :

الوسط الحسابي الجديد = الوسط الحسابي القديم + العدد الثابت c

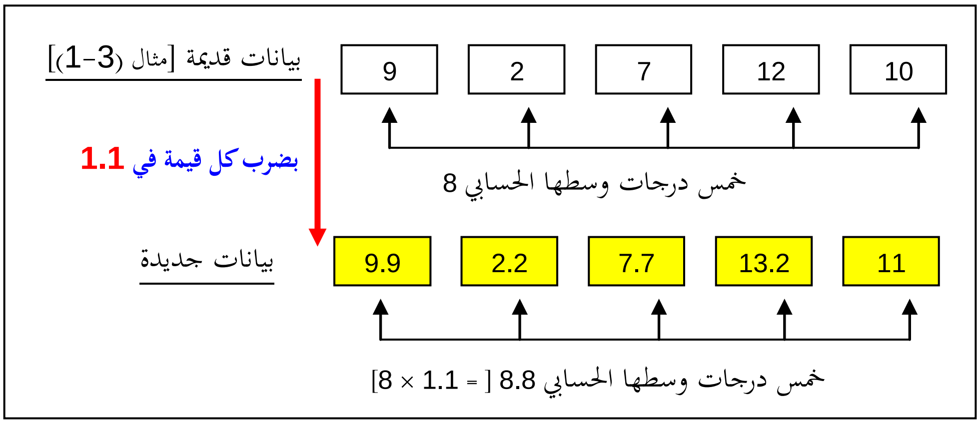
فمثلاً في مثال (1-3) إذا أضفنا العدد الثابت 1 لكل قيمة من القيم ، فإن الوسط الحسابي الجديد يصبح 9 [= 1 + 8] ، ويمكنك التأكد بنفسك .



- **ملحوظة (9) :** أما إذا ضربنا عدد ثابت c في كل قيمة من قيم البيانات ، فإن :

الوسط الحسابي الجديد = الوسط الحسابي القديم × العدد الثابت c

فمثلاً في مثال (3-1) إذا ضربنا كل قيمة في العدد الثابت 1.1 ، فإن الوسط الحسابي الجديد يصبح 8.8 [= 1.1 × 8] ، ويمكنك التأكد بنفسك .



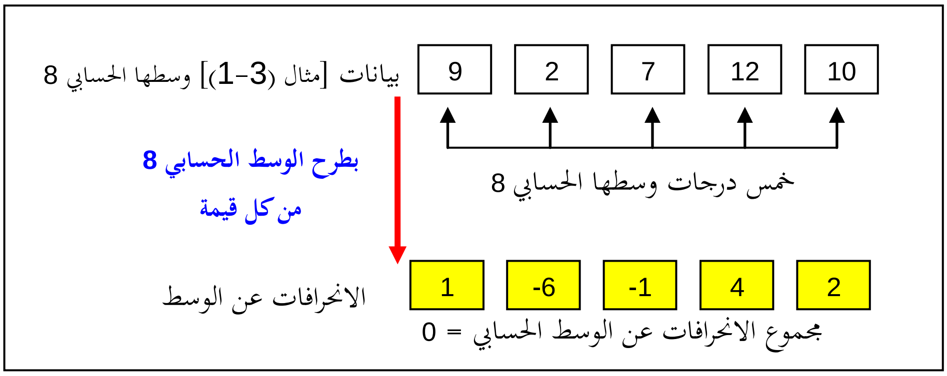
- **ملحوظة (10) :** المجموع الجبري لانحرافات القيم عن الوسط الحسابي يساوي صفراً

المقصود بانحراف قيمة عن الوسط الحسابي هو الفرق بين تلك القيمة والوسط الحسابي

فإذا رمزنا لانحرافات القيم بالرمز d فإن :

$$\boxed{\sum d = 0} \qquad (3\text{-}2)$$

فمثلاً في مثال (3-1) إذا طرحنا من كل قيمة من القيم الوسط الحسابي للبيانات [= 8] ، فإن مجموع الانحرافات الناتجة يجب أن يساوي صفراً ، ويمكنك التأكد بنفسك .



مثال (3-2) : يسؤال خمسة أشخاص عن أجرهم الشهري فكانت إجاباتهم كما يلي (بالألف ريال) :

3   5   2   7   3

(أ) **احسب** الوسط الحسابي للأجور الشهرية .

(ب) وإذا قررت إدارة الشركة زيادة أجورهم **فأيهما** يحسن الوسط الحسابي للأجور بصورة أفضل : زيادة أجر كل شخص بمقدار 2000 ريال أم زيادة الأجور بنسبة 5% .

الحل :

(أ) الوسط الحسابي للأجر الشهري : $\bar{x} = \frac{3+5+2+7+3}{5} = \frac{20}{5} = 4$

أي أن الوسط الحسابي للأجر الشهري هو أربعة آلاف ريال .

(ب) إذا قررت إدارة الشركة زيادة الأجور بمقدار 2000 ريال ، إذن بإضافة 2 [ألفان] لكل أجر فإن الأجور الجديدة تصبح 5 9 4 7 5 ، وبالتالي يكون الوسط الحسابي (الجديد) لها هو :

$$\bar{x} = \frac{5+7+4+9+5}{5} = \frac{30}{5} = 6$$

أي يصبح الوسط الجديد ستة آلاف . لاحظ أن النتيجة السابقة يمكن الحصول عليها مباشرة بإضافة 2 للوسط الحسابي القديم [ملحوظة (8) السابقة] : $\bar{x} = 4+2 = 6$

أما إذا قررت الشركة زيادة الأجور بنسبة 5% فهذا يعني أن كل أجر جديد سيصبح مساوياً للأجر القديم مضروباً في 1.05 لتصبح الأجور الجديدة كالتالي :

3.15 7.35 2.1 5.25 3.15

وبالتالي يكون الوسط الحسابي (الجديد) لها هو :

$$\bar{x} = \frac{3.15+5.25+2.1+7.35+3.15}{5} = \frac{21}{5} = 4.2$$

أي يصبح الوسط الجديد أربعة آلاف ومائتان ريال . لاحظ أن النتيجة السابقة يمكن الحصول عليها مباشرة بضرب الوسط الحسابي القديم أيضاً في 1.05 [ملحوظة (9) السابقة] :

$$\bar{x} = 4 \times 1.05 = 4.2$$

ومن السابق يتضح أن زيادة الأجر الشهري ألفان ريال سيحسن الوسط الحسابي للأجور الشهرية بصورة أفضل من زيادة الأجور بنسبة 5% .

(3-3-2) : الوسط الحسابي للبيانات المبوبة

الأصل عند حساب أية مقاييس أن تكون هذه المقاييس للبيانات الخام ، لكن في بعض الأحيان يكون من المناسب وضع هذه البيانات الخام في صورة جداول تكرارية [كما رأينا في الفصل الثاني] ، وعندئذٍ تكون لدينا بيانات مبوبة .

- فإذا كانت البيانات المبوبة تخص **متغير كمي متقطع** ، فإن الوسط الحسابي $\bar{x}$ لهذه البيانات سيكون :

$$\boxed{\bar{x}=\frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}}$$

حيث f_i هو تكرار القيمة x_i ، $\sum_{i=1}^{n} f_i$ هو مجموع التكرارات ، والعلاقة السابقة يمكن [للتبسيط] كتابتها على الصورة :

$$\boxed{\bar{x}=\frac{\sum fx}{\sum f}} \qquad (3\text{-}3)$$

حيث $\sum fx$ تعني مجموع [حاصل ضرب كل قيمة في تكرارها] ، $\sum f$ تعني مجموع التكرارات .

لاحظ أن العلاقة (3-3) هي نفسها العلاقة (3-1) لكن بعد استبدال x [في العلاقة (3-1)] بـ fx [في العلاقة (3-3)] ، واستبدال n [في العلاقة (3-1)] بـ $\sum f$ [في العلاقة (3-3)] وسوف نلاحظ تكرار تلك الملحوظة دائماً عند تعميم كثير من التعريفات الخاصة ببيانات غير مبوبة إلى التعريفات المماثلة والخاصة بالبيانات المبوبة .

- وإذا كانت البيانات المبوبة تخص **متغير كمي متصل** ، فإن الوسط الحسابي $\bar{x}$ لهذه البيانات سيتم حسابه أيضاً عن طريق العلاقات (3-3) حيث يمثل f_i هنا تكرار الفئة رقم i ، وتمثل القيمة x_i مركز تلك الفئة ، ومن المناسب [حتى نتذكر أن x_i هنا تعني مركز الفئة] إعادة كتابة العلاقة (3-3) على الصورة :

$$\boxed{\bar{x}=\frac{\sum fx_0}{\sum f}} \qquad (3\text{-}4)$$

حيث f هو تكرار الفئة التي مركزها x_0 .

- وفي حالة البيانات المبوبة [متقطعة أو متصلة] يكون مجموع الانحرافات عن الوسط الحسابي هو $\sum fd$ ، وبالتالي فإن :

$$\sum fd = 0$$

أما $\sum d$ فلا يُشترط أن يكون صفراً ، حيث d هو الانحراف عن الوسط الحسابي والذي يُعطى بـ $x-\bar{x}$ [في حالة البيانات الكمية المتقطعة] أو $x_0-\bar{x}$ [في حالة البيانات الكمية المتصلة] .

مثال (3–3) :

أوجد الوسط الحسابي للأرقام 8 , 8 , 8 , 2 , 2 , 4 , 4 , 4 , 4 , 4 , 6 , 6 , 3 , 3 , 5 , 5 , 5 , 5 , 5 , 5

الحل :

بتطبيق مباشر للتعريف (3-1) باعتبار البيانات بيانات غير مبوبة :

$$\overline{x} = \frac{(5+...+5)+(3+3)+(6+6)+(4+...+4)+(2+2)+(8+..+8)}{20} = \frac{96}{20} = 4.8$$

لاحظ أن الرقم **5** متكرر **6** مرات ، الرقم **3** **مرتان** ، والرقم **6** **مرتان** ، والرقم **4** متكرر *5* مرات ، والرقم **2** **مرتان** ، والرقم **8** **ثلاث** مرات ، وبالتالي يمكن عمل العملية الحسابية السابقة كالآتي :

$$\overline{x} = \frac{(6 \times 5)+(2 \times 3)+(2 \times 6)+(5 \times 4)+(2 \times 2)+(3 \times 8)}{6+2+2+5+2+3} = 4.8$$

لاحظ أيضاً أن العملية الحسابية السابقة هي نفسها العلاقة (3-3) حيث يمثل البسط مجموع حاصل ضرب كل قيمة في تكرارها [أي $\sum fx$] ، ويمثل المقام مجموع التكرارات [أي $\sum f$] .

ومن الممكن [لكن ليس بشرط] أن نقوم بعمل جدول تكراري للبيانات المعطاة [والمكون من عمود للقيم وعمود آخر لتكرارات تلك القيم] ثم نضيف عموداً ثالثاً يمثل حاصل ضرب كل قيمة في تكرارها *fx* ثم نقسم مجموع قيم هذا العمود الثالث على مجموع التكرارات كما هو مبين بالجدول التالي :

fx	التكرار *f*	المتغير *x*
30	6	5
6	2	3
12	2	6
20	5	4
4	2	2
24	3	8
96	20	

$\sum f = 20$ $\quad$ $\sum fx = 96$

الجدول التكراري

$$\overline{x} = \frac{\sum fx}{\sum f} = \frac{96}{20} = 4.8$$

مثال (3–4) :

من مائة رقم يتكرر الرقم 4 عشرون مرة ، والرقم 5 أربعون مرة ، والرقم 6 ثلاثون مرة ، والباقي كانوا الرقم 7 .

احسب الوسط الحسابي للمائة رقم .

الحل :

بتكوين الجدول التكراري للأرقام المذكورة ، ثم بضرب كل قيمة في تكرارها والتجميع [عمود *fx*] يكون الوسط الحسابي للأرقام المذكورة هو :

	الجدول التكراري	
fx	التكرار *f*	المتغير *x*
80	20	4
200	40	5
180	30	6
70	10	7
530	100	
$\sum fx$	$\sum f$	

$$\overline{x} = \frac{\sum fx}{\sum f} = \frac{530}{100} = 5.3$$

مثال (5–3) : **احسب** الوسط الحسابي لأطوال سيقان الأزهار (بوحدات السنتيمتر) لبيانات مثال (6–2) من الباب الثاني .

الحل :

الجدول التكراري

الفئة	المتغير x (الطول)	التكرار f	مركز الفئة x_0	fx_0
الأولى	$0 \le x < 20$	4	10	40
الثانية	$20 \le x < 30$	16	25	400
الثالثة	$30 \le x < 35$	12	32.5	390
الرابعة	$35 \le x < 40$	10	37.5	375
الخامسة	$40 \le x < 50$	6	45	270
السادسة	$50 \le x < 60$	2	55	110
		50		1585
		$\sum f$		$\sum f x_0$

من الجدول التكراري للبيانات يمكن تحديد مركز كل فئة [عمود x_0] ومن ثم حاصل ضرب مركز كل فئة في تكرارها [عمود fx_0] وبالتالي يكون الوسط الحسابي لأطوال سيقان الأزهار هو :

$$\bar{x} = \frac{\sum fx_0}{\sum f} = \frac{1585}{50} = 31.7$$

أي أن الوسط الحسابي لأطوال السيقان هو 31.7 سنتيمتر .

مثال (6–3) :

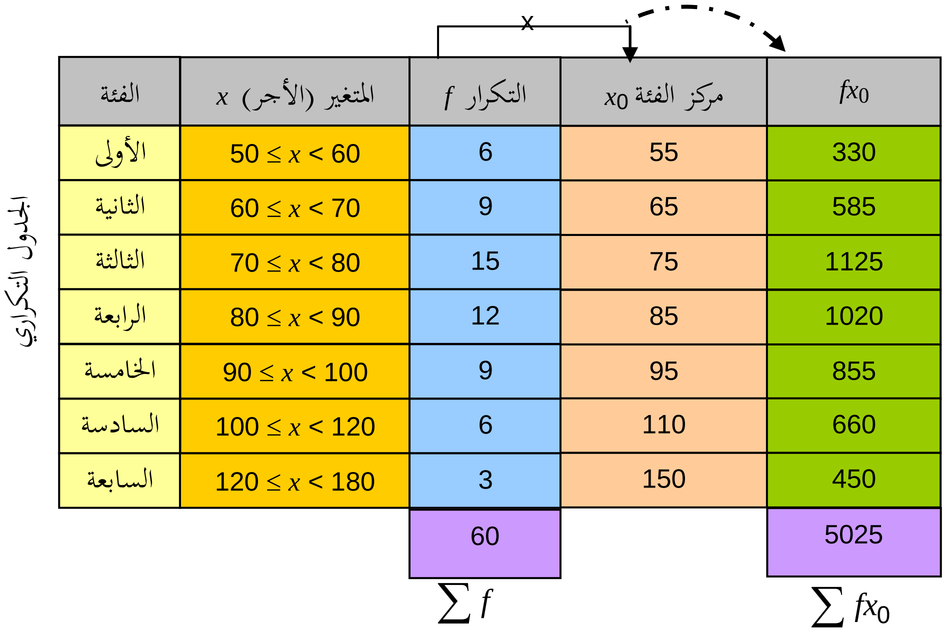
احسب الوسط الحسابي للأجر السنوي (بالألف ريال) لبيانات العاملين بمثال (8–2) من الباب الثاني .

الحل :

من الجدول التكراري المرافق للبيانات يمكن تحديد مركز كل فئة [عمود x_0] ومن ثم حاصل ضرب مركز كل فئة في تكرارها [عمود fx_0] وبالتالي يكون الوسط الحسابي لأطوال سيقان الأزهار هو :

$$\bar{x} = \frac{\sum fx_0}{\sum f} = \frac{5025}{60} = 83.75$$

أي أن الوسط الحسابي للدخل السنوي هو 83.75 ألفاً من الريالات .



fx_0	مركز الفئة x_0	التكرار f	المتغير (الأجر) x	الفئة
330	55	6	$50 \le x < 60$	الأولى
585	65	9	$60 \le x < 70$	الثانية
1125	75	15	$70 \le x < 80$	الثالثة
1020	85	12	$80 \le x < 90$	الرابعة
855	95	9	$90 \le x < 100$	الخامسة
660	110	6	$100 \le x < 120$	السادسة
450	150	3	$120 \le x < 180$	السابعة
5025		60		
$\sum fx_0$		$\sum f$		

ملخص للدرس الرابع [الباب الثالث : مقاييس النزعة المركزية (الوسط الحسابي)]

- **مقاييس النزعة المركزية (أو المتوسطات) :** هي قيم نموذجية يمكن أن تمثل مجموعة البيانات بحيث تعطي دلالات معينة لتلك البيانات . **أمثلة** : الوسط الحسابي – الوسيط – المنوال .

- الوسط الحسابي = $\frac{\sum x}{n}$ أو $\frac{\sum fx}{\sum f}$ أو $\frac{\sum fx_0}{\sum f}$ [لبيانات مفردة أو بيانات مبوبة (كمية متقطعة) ، أو بيانات مبوبة (كمية متصلة)] .

- مجموع انحرافات القيم عن الوسط الحسابي يساوي صفراً ، أي أن :
  $\sum d = 0$ [للبيانات المفردة] ، $\sum fd = 0$ [في حالة البيانات المبوبة]

- إذا كان لدينا مجموعة من القيم وحسبنا لها الوسط الحسابي ثم أضفنا لكل قيمة من القيم العدد الثابت c فإن الوسط الحسابي الجديد = الوسط القديم + c .

- أما إذا ضربنا كل قيمة من القيم في عدد ثابت c فإن الوسط الجديد = القديم مضروباً في c .

- وللوسط الحسابي المزايا والعيوب التالية :
  - المزايا : سهولة الحساب – جميع البيانات تُؤخذ في الاعتبار – لا يحتاج إلى ترتيب معين للبيانات عند حسابه .
  - العيوب : يتأثر بشدة بالقيم المتطرفة – لا يمكن إيجاده بالرسم – لا يمكن حسابه في حالة التوزيعات التكرارية المفتوحة – لا يمكن تحديده للبيانات النوعية [أي يمكن حسابها فقط للبيانات الكمية]

تدريب عملي

في كل حالة من الحالات التالية ، **أوجد** الوسط الحسابي $\bar{x}$ للبيانات المعطاه :

(أ) 7 , 4 , 10 , 9 , 15 , 12 , 7 , 9 , 7

(ب) 8 , 11 , 4 , 3 , 2 , 5 , 10 , 6 , 4 , 1 , 10 , 8 , 12 , 6 , 5 , 7

(ج) 85 , 76 , 93 , 82 , 96

(د) 0.53 , 0.46 , 0.50 , 0.49 , 0.52 , 0.53 , 0.44 , 0.55

(هـ)

x	462	480	498	516	534	552	570	588	606	624
f	98	75	56	42	30	21	15	11	6	2

(و)

x	10 –	15 –	20 –	25 –	30 –	35 –	40 – 45
f	3	7	16	12	9	5	2

تدريبات (4)

الإجابة النهائية لجميع التمرينات موجودة في نهاية التدريب

اختر الإجابة الصحيحة

(1) **مقاييس النزعة المركزية** هي

(أ) مقاييس ترصد درجة التدبب في قمة المنحنى مقارنة بقمة منحنى التوزيع الطبيعي

(ب) مقاييس ترصد الدرجة التي تتجه بها البيانات الكمية للانتشار حول قيمة متوسطة

(ج) مقاييس تحدد النسبة المئوية للتشتت المطلق بالنسبة لقيمة متوسطة

(د) هي مقاييس ترصد درجة تماثل أو البعد عن التماثل لتوزيع ما

(هـ) قيم نموذجية يمكن أن تمثل مجموعة البيانات

(2) **الوسط الحسابي** هو أحد مقاييس

(أ) النزعة المركزية (ب) التشتت

(ج) الالتواء (د) التفرطح

(3) لعدد من القيم ، يُعرف **مجموع هذه القيم مقسوماً على عددها** على أنه

(أ) تباين تلك القيم (ب) الانحراف المتوسط للقيم

(ج) الوسط الحسابي للقيم (د) الانحراف المعياري للقيم

خاص بالأسئلة (4) ، (5) : إذا كان $\sum x$ هو مجموع عدد قدره n من القيم ، وكان $\sum d$ هو مجموع انحرافات هذه القيم عن وسطها الحسابي ، فإن :

(4) $\frac{\sum x}{n}$ هو :

(أ) الانحراف المتوسط للقيم (ب) تباين تلك القيم

(ج) الوسط الحسابي للقيم (د) صفر

(5) $\frac{\sum d}{n}$ هو :

(أ) الوسط الحسابي للقيم (ب) الانحراف المتوسط للقيم

(ج) تباين تلك القيم (د) صفر

---

(6) إذا كان الوسط الحسابي لمجموعة من القيم هو 20 وأضفنا لكل قيمة من القيم 2 ، فإن الوسط الحسابي للقيم الجديدة يكون :

(أ) 20 (ب) 22

(ج) 40 (د) 18

(7) إذا كان الوسط الحسابي لمجموعة من القيم هو 20 وضربنا كل قيمة من القيم في العدد 2 ، فإن الوسط الحسابي للقيم الجديدة يكون :

(أ) 20 (ب) 22

(ج) 40 (د) 18

(8) إذا كان الوسط الحسابي لمجموعة من القيم هو 20 وضربنا كل قيمة من القيم في العدد 2- ، فإن الوسط الحسابي للقيم الجديدة يكون :

(أ) 20 (ب) 22

(ج) 40 (د) 40-

(9) لمجموعة القيم 4 5 8 9 4 ، الوسط الحسابي يساوي :

(أ) 8 (ب) 5

(ج) 4 (د) 6

(10) لمجموعة القيم 16 4 8 2 3 9 ، الوسط الحسابي يساوي :

(أ) 6 (ب) 8

(ج) 7 (د) 9

(11) مجموعة من القيم عددها 10 ولها البيانات التالية : $\sum x = 60 , \sum | d | = 22 , \sum d^2 = 76$ حيث $\sum x$ هو مجموع القيم ، d هو الانحراف عن الوسط الحسابي للقيم ، $|d|$ هو القيمة المطلقة لهذا الانحراف ، إذن الوسط الحسابي للبيانات السابقة هو :

(أ) 2.2 (ب) 7.6

(ج) 6 (د) 2.76

خاص بالأسئلة (12) ، (13) ، (14) :

في الجدول التكراري المبين [غير مهم البيانات المرصود لها ....] ، إذا كان d يمثل الانحراف [لكل قيمة x] عن الوسط الحسابي ، فإن :

x	f	fx	d	$\lvert d \rvert$	$f \lvert d \rvert$	fd
......	......	......	......	......	......	......
......	......	......	......	......	......	......
$\sum x = 20$	$\sum f = 100$	$\sum fx = 450$	$\sum d = -3$	......	$\sum f \lvert d \rvert = 185$	$\sum fd = ?$

(12) الوسط الحسابي للبيانات السابقة هو :

(أ) 4.5 (ب) 1.85

(ج) 2.18 (د) 4.75

(13) $\sum fd$ يساوي

(أ) 185 (ب) -300

(ج) 0 (د) قيمة أخرى غير ما سبق

(13) $\sum x \times \sum f$ يساوي

(أ) 2000 (ب) 450

(ج) 0 (د) قيمة أخرى غير ما سبق

الإجابة :

(1) هـ (2) ب (3) ج (4) ج (5) د

(6) ب (7) ج (8) د (9) د (10) ج

(11) ج (12) أ (13) ج (14) ب

عناصر الدرس

الباب الثالث : مقاييس النزعة المركزية [تابع]

- الوسيط

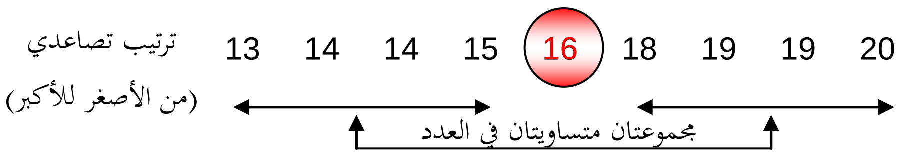
الباب الثالث : مقاييس النزعة المركزية [تابع]

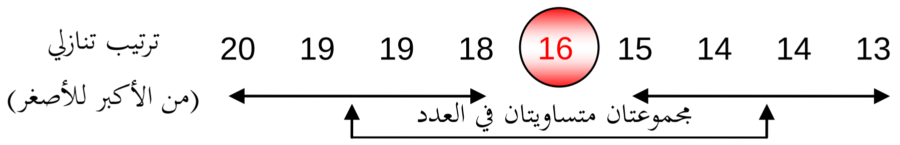
(3–4) الوسيط Median

(3–4–1) تعريف الوسيط

(ببساطة) يُعرف **الوسيط** [ويرمز له بالرمز M] لمجموعة من القيم (المرتبة تصاعدياً أو تنازلياً حسب قيمها) على أنه القيمة التي تقسم مجموعة القيم إلى مجموعتين متساويتين في العدد ، أو بتعبير آخر هي القيمة التي في المنتصف .

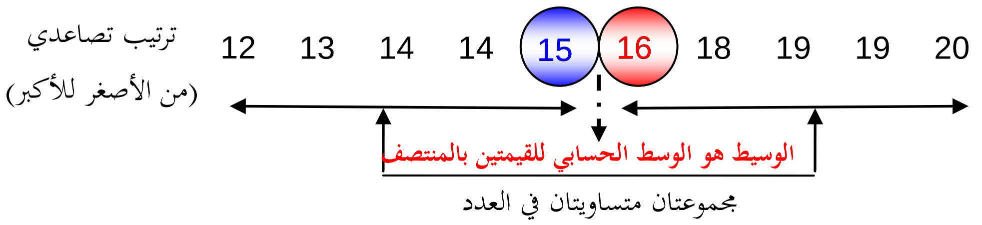
فمثلاً لمجموعة القيم : 13 , 14 , 19 , 16 , 20 , 15 , 18 , 14 , 19 [عددها 9 قيم] ، إذا قمنا بترتيبها تصاعدياً أو تنازلياً ، يكون الوسيط هو العدد **الخامس** وقيمته **16** [فرق بين رتبة الوسيط (والتي تعني ترتيبه بين القيم) وقيمته] .

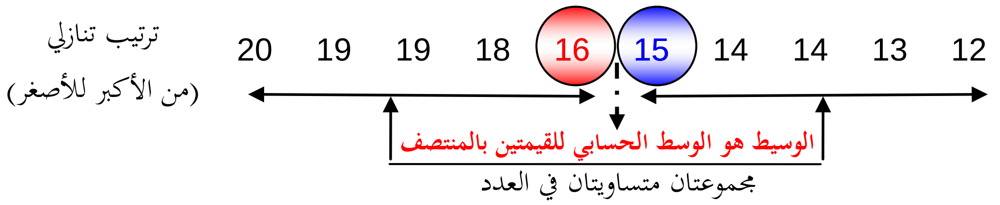




لاحظ في المثال السابق أن عدد القيم n [هنا $= 9$] فردي وبالتالي هناك قيمة واحدة في منتصف المجموعة .

أما لمجموعة القيم : 12 , 13 , 14 , 19 , 16 , 20 , 15 , 18 , 14 , 19 [عددها 10 قيم (أي زوجي) حيث أضفنا القيمة 12 للمجموعة السابقة] ، فإذا قمنا بترتيبها تصاعدياً أو تنازلياً سنجد الآتي :





في هذه الحالة توجد قيمتان بالمنتصف وهما القيمة **الخامسة** والقيمة **السادسة** [وهما القيمتان 16 , 15] ، عندئذ يكون الوسيط هو الوسط الحسابي لهاتين القيمتين ، أي :

$$\frac{15 + 16}{2} = 15.5$$

إذن من السابق يمكن استنتاج **طريقة حساب الوسيط لمجموعة من القيم** كالآتي :

- قم أولاً بترتيب البيانات تصاعدياً أو تنازلياً .
- حدد ما إذا كانت هناك قيمة واحدة بالمنتصف أم قيمتين ، وهذا يتوقف على قيمة n حيث n عدد القيم :

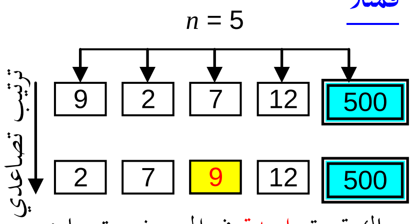
فإذا كانت n فردية

كانت هناك قيمة واحدة في المنتصف رتبتها $\frac{n+1}{2}$ وتكون هذه القيمة هي **الوسيط**

وإذا كانت n زوجية

كانت هناك قيمتان في المنتصف رتبتهما $\frac{n}{2}, \frac{n}{2}+1$ ويكون **الوسيط** هو الوسط الحسابي لهاتين القيمتين

فمثلاً



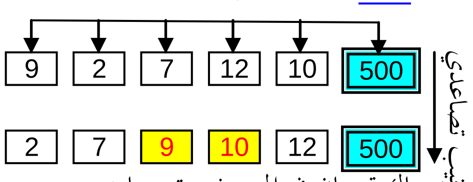
هناك قيمة واحدة في المنتصف رتبتها :

$$\frac{n+1}{2} = \frac{5+1}{2} = 3$$

أي القيمة الثالثة ، وتكون تلك القيمة هي الوسيط . أي أن :

الوسيط = 9

فمثلاً



هناك قيمتان في المنتصف رتبتهما :

$$\frac{n}{2} = \frac{6}{2} = 3 \quad , \quad 3+1 = 4$$

أي القيمتان الثالثة والرابعة ، وتكون قيمة الوسيط هي الوسط الحسابي لهاتين القيمتين ، أي :

$$\frac{9+10}{2} = 9.5$$

هل لاحظت أن الوسيط لم يتأثر بالقيمة المتطرفة 500

مثال (3−7) : أوجد الوسط الحسابي $\bar{x}$ والوسيط M لكل مجموعة القيم :

(أ) 9 6 5 3 2 7 6 4 3

(ب) 18 12 9 5 15 11 7 5

الحل :

(أ) الوسط الحسابي :

$$\bar{x} = \frac{2+3+3+4+5+6+6+7+9}{9} = 5$$ [لا نحتاج لترتيب القيم]

الوسيط : نحتاج أولاً لترتيب القيم [تصاعدياً مثلاً] :

عدد القيم فردي [$n = 9$]   2   3   3   4   **5**   6   6   7   9

هناك قيمة واحدة في المنتصف رتبتها $\frac{9+1}{2}=5$ (أي القيمة الخامسة) وقيمتها **5** وهي الوسيط .

(ب) الوسط الحسابي :

$$\bar{x}=\frac{5+5+7+9+11+12+15+18}{8}=10.25$$

الوسيط : نحتاج أولاً لترتيب القيم [تصاعدياً مثلاً] :

عدد القيم زوجي [$n = 8$]   5   5   7   **9**   **11**   12   15   18

هناك قيمتان في المنتصف رتبتهما :

$$\frac{8}{2}=\underline{\underline{4}} \quad , \quad \frac{8}{2}+1=\underline{\underline{5}}$$

(أي القيمتان الرابعة والخامسة) ووسطهما الحسابي يساوي $\frac{9+11}{2}=10$ ، أي أن الوسيط = **10** .

مثال (3–8) : الأجر (بالريال) في الساعة لخمسة عاملين في مكتب هو : 37 , 92 , 32 , 39 , 25 . **احسب** الوسط الحسابي للأجور ووسيط هذه الأجور . **أيهما** تفضل كمقياس لمتوسط أجر الساعة ؟ **ولماذا** ؟

الحل :

الوسط الحسابي للأجور هو :

$$\bar{x}=\frac{\sum x}{n}=\frac{25+39+32+92+37}{5}=\frac{225}{5}=45$$

أما لتحديد الوسيط ، فلابد أولاً من ترتيب القيم (تصاعدياً مثلاً) :

25 , 32 , **37** , 39 , 92

وحيث أن عدد القيم فردي [$n = 5$] ، إذن هناك قيمة واحدة في المنتصف رتبتها $\frac{5+1}{2}=3$ (أي القيمة الثالثة **37**) وهي الوسيط ، إذن الوسيط = **37** .

لاحظ في السؤال السابق أن الوسط الحسابي (بالرغم من عدم احتياجه لترتيب القيم وفي نفس الوقت يأخذ في الاعتبار جميع قيم البيانات) إلا أنه تأثر جداً بالقيمة المتطرفة **92** ، في حين لم يتأثر بها الوسيط لأنه يعتمد على البيانات في المنتصف . لذا يُفضل هنا استخدام الوسيط كمقياس للنزعة المركزية حيث يعطي دلالة أفضل لمتوسط الأجور من الوسط الحسابي .

(3–4–2) : الوسيط للبيانات المبوبة [بيانات كمية متقطعة]

في حالة البيانات المبوبة يقوم مجموع التكرارات $\sum f$ بدور عدد القيم n ، وبالتالي :

- إذا كان $\sum f$ فردياً تكون هناك قيمة واحدة في المنتصف رتبتها $\frac{(\sum f)+1}{2}$ وتكون هذه القيمة هي الوسيط
- وإذا كان $\sum f$ زوجياً تكون هناك قيمتان في المنتصف رتبتهما $\frac{\sum f}{2}+1$, $\frac{\sum f}{2}$ ويكون الوسط الحسابي لهاتين القيمتين هو الوسيط ؟

مثال (3–9) : ما هو وسيط مجموعة القيم المعطاة في مثال (3–4) ؟

الحل :

الجدول المرافق يبين الجدول التكراري للقيم [بعد ترتيبها تصاعدياً] ، وحيث أن عدد القيم هنا 100 [= $\sum f$] إذن هناك قيمتان في المنتصف رتبتهما :

التكرار f	المتغير x
20	4
40	5
30	6
10	7
100	$\sum f$

$$\frac{100}{2}+1=51 \quad , \quad \frac{100}{2}=50$$

وهاتان القيمتان هما 5 , 5 [حيث أن الـ 20 قيمة الأولى هي الرقم 4 ذات التكرار 20 ثم يليها الرقم 5 بتكرار 40 [وبذلك يحتل الرقم 5 الترتيب من 21 حتى 60] ، وبالتالي يكون الوسيط هو الوسط الحسابي للقيمتين 5 , 5 .. أي **5** [تذكر أن الوسط الحسابي لمجموعة الأرقام كلها (ارجع لمثال (3–4) هو **5.3**] .

(3–4–3) : الوسيط للبيانات المبوبة [بيانات كمية متصلة]

والآن ماذا عن الوسيط لبيانات كمية متصلة ؟

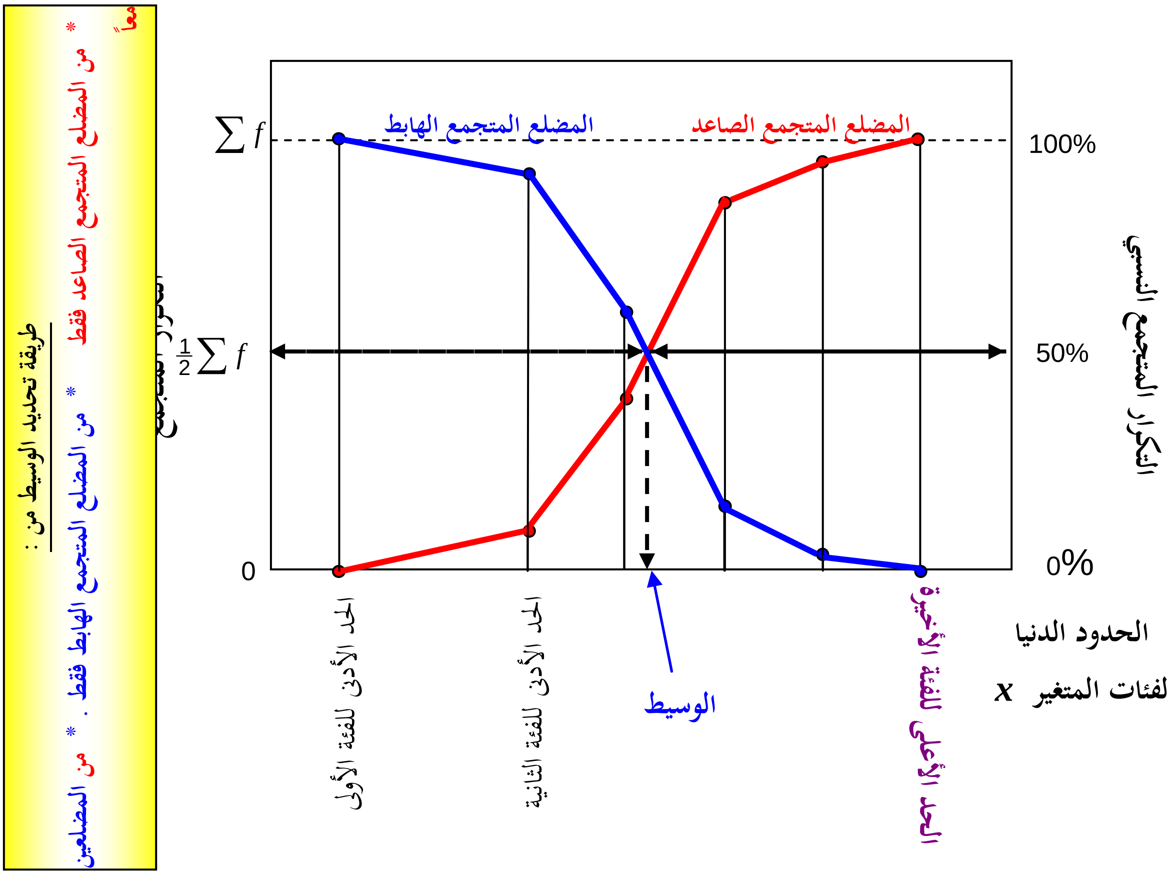
نستطيع تحديده بسهولة [حيث نوهنا لذلك في الباب الثاني] فهو :

قيمة المتغير المناظرة التي تقسم مجموعة القيم لمجموعتين متساويتين في العدد [التعريف الأساسي]

أو **قيمة المتغير المناظرة لنقطة تقاطع المضلعين : المتجمع الصاعد والمتجمع الهابط للبيانات .**

أو **القيمة التي يناظرها تكرار متجمع = نصف مجموع التكرارات**

أو **القيمة التي يناظرها تكرار نسبي متجمع = 50%**



طريقة تحديد الوسيط من :

* من المضلع المتجمع الصاعد فقط * من المضلع المتجمع الهابط فقط . * من المضلعين معاً

مثال (3–10) : في دراسة جغرافية لعدد من مساحات مجموعة من الأراضي لمنطقة سكنية بالرياض تبين أن التوزيع التكراري لها كما هو مبين .

المساحة (بالفدان)	عدد قطع الأراضي
1 -	14
3 -	29
5 -	18
7 - 10	9

المطلوب حساب الوسط الحسابي والوسيط لمساحة الأراضي .

الحل :

المتغير x هنا هو مساحة الأرض (بالفدان) ، في حين يمثل عدد قطع الأراضي **التكرار** f .

أولاً : الوسط الحسابي : نستكمل الجدول التكراري كما هو مبين بالجدول المرافق ، ومنه يمكن تحديد الوسط الحسابي كالتالي :

$$\therefore \quad \bar{x} = \frac{\sum f\, x_0}{\sum f} = \frac{328.5}{70} = 4.692857143 \cong \underline{\underline{4.7}}$$

الجدول التكراري

x

الفئة	المتغير (المساحة) x	التكرار f	المركز x_0	$f\,x_0$
الأولى	$1 \le x < 3$	14	2	28
الثانية	$3 \le x < 5$	29	4	116
الثالثة	$5 \le x < 7$	18	6	108
الرابعة	$7 \le x < 10$	9	8.5	76.5
		$\sum f = 70$		$\sum f\, x_0 = 328.5$

ثانياً : الوسيط : نكون الجدول التكراري المتجمع الصاعد [أو الهابط]

إحدى الطرق البيانية التي يمكن استخدامها في حساب الوسيط هي رسم المضلع التكراري المتجمع الصاعد ، ويتم ذلك كالآتي :

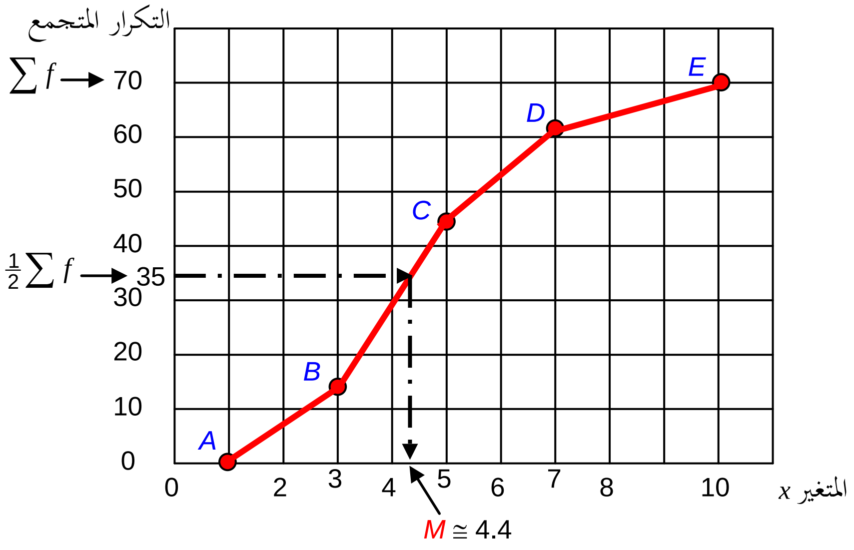
- من الجدول التكراري ، يمكن تكوين الجدول التكراري المتجمع الصاعد وتحديد النقاط A , B , C , D , E التي منها يمكن رسم المضلع التكراري المتجمع الصاعد [كما أشرنا لذلك في الباب الثاني] .

الجدول التكراري المتجمع الصاعد

المتغير x	التكرار المتجمع	النقطة على المضلع
< 1	0	$A\,(1 . 0)$
< 3	14	$B\,(3 . 14)$
< 5	43	$C\,(5 . 43)$
< 7	61	$D\,(7 . 61)$
< 10	$\sum f = 70$	$E\,(10 , 70)$

الجدول التكراري

الفئة	المتغير (المساحة) x	التكرار f
الأولى	$1 < x < 3$	14
الثانية	$3 < x < 5$	29
الثالثة	$5 < x < 7$	18
الرابعة	$7 < x < 10$	9
		$\sum f = 70$

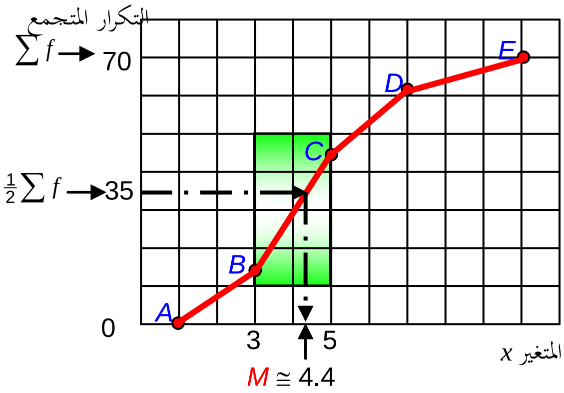


- وحيث أن الوسيط هو قيمة المتغير المناظرة لتكرار متجمع صاعد = $\frac{1}{2}\sum f$ [أي المناظرة لتكرار متجمع صاعد يساوي 35] يمكن تحديد الوسيط M بصورة تقريبية ليكون :

$$M \cong 4.4$$

تدريب للقارئ : أعد الحل باستخدام المضلع التكراري المتجمع الهابط [مرة] ، والمضلعين التكراريين الصاعد والهابط [مرة أخرى] ، وقارن بالطرق البيانية المختلفة لحساب الوسيط .

ملحوظة هامة جداً [طريقة الاستكمال لحساب الوسيط] :



الوسيط يقع بين النقطتين $B(3\ ,\ 14)\ ,\ C(5\ ,\ 43)$ [أي داخل الفئة $3 \le x < 5$] ، هذه الفئة تُسمى بـ **الفئة الوسيطية** [أو **فئة الوسيط**] وهي تلك الفئة التي يقع داخلها الوسيط ، أي أن الفئة الوسيطية [أو فئة الوسيط] هي تلك الفئة التي يقع داخلها الوسيط .

وهنا نسأل أنفسنا سؤالين :

السؤال الأول : هل من الممكن تحديد الفئة الوسيطية من الجدول التكراري مباشرةً أم لابد من تكوين الجدول التكراري المتجمع الصاعد ورسم المضلع التكراري المتجمع الصاعد ، ثم نحدد تلك الفئة ؟ .

السؤال الثاني : هل من الممكن [بعد تحديد الفئة الوسيطية] تحديد الوسيط من الجدول التكراري مباشرة دون الحاجة للجدول التكراري المتجمع الصاعد أو المضلع التكراري المتجمع الصاعد ؟ .

والإجابة على السؤالين : نعم يمكن تحديد الفئة الوسيطية من الجدول التكراري مباشرةً ، ثم بعد ذلك **يمكن** أيضاً من هذا الجدول التكراري تحديد قيمة الوسيط دون أن نحتاج لعمل جدول تكراري متجمع صاعد ورسم المضلع التكراري المتجمع الصاعد ، ولنرى ذلك معاً :

بالنسبة للسؤال الأول [تحديد الفئة الوسيطية]

- نقوم أولاً بحساب نصف مجموع التكرارات [أي $\frac{1}{2}\sum f$] .
- مبتدئين بالرقم صفر (في ذهننا) نقوم بإضافة تكرار الفئة الأولى :

(1) فإن كان الناتج أكبر من أو يساوي $\frac{1}{2}\sum f$ تكون الفئة المضاف تكرارها هي الفئة الوسيطية .

(2) وإذا كان الناتج أقل من $\frac{1}{2}\sum f$ لا تكون الفئة السابقة هي الفئة الوسيطية ، فنقوم بإضافة تكرار الفئة التالية للناتج السابق (الناتج من الخطوة السابقة) ثم نعود للخطوة (1) السابقة مرة أخرى .

- ونستمر في تكرار الخطوتين (1) ، (2) حتى نصل للفئة الوسيطية .

الجدول التكراري		
الفئة	المتغير (المساحة) x	التكرار f
الأولى	$1 \le x < 3$	14
الثانية	$3 \le x < 5$	29
الثالثة	$5 \le x < 7$	18
الرابعة	$7 \le x < 10$	9
		$\sum f = 70$

فمثلاً في المثال السابق [مثال (3–10)] :

$$\frac{1}{2}\sum f = \frac{1}{2}\times 70 = 35$$

- مبتدئين بالصفر (في ذهننا) نقوم بإضافة تكرار الفئة الأولى (= 14) فنحصل على :

$$0+14=14 \quad (<35)$$

وبالتالي فإن الفئة الأولى ليست الفئة الوسيطة

- نضيف للناتج السابق (= 14) تكرار الفئة الثانية (= 29) فنحصل على :

$$14+29=43 \quad (>35)$$

إذن الفئة **الثانية** هي الفئة الوسيطية ، بتعبير آخر يكون الوسيط واقع داخل الفئة الثانية .

وبالنسبة للسؤال الثاني [تحديد الوسيط (بعد ما حددنا الفئة الوسيطية)]

- للفئة الوسيطية ، حدد الحد الأدنى لها ، وتكرارها ، وطولها .
- احسب ما يُسمى بـ "**التكرار المتجمع السابق**" والذي يساوي مجموع تكرار الفئات السابقة للفئة الوسيطية .
- احسب الوسيط من العلاقة :

$$\text{الوسيط } M = \text{الحد الأدنى للفئة الوسيطية} + \left[\frac{(\text{نصف مجموع التكرارات} - \text{التكرار المتجمع السابق})}{\text{تكرار الفئة الوسيطية}} \times \text{طول الفئة الوسيطية}\right]$$

تُسمى هذه الطريقة **بطريقة الاستكمال** .

فمثلاً في المثال السابق [مثال (3–10)] ، حددنا الفئة الوسيطية [الفئة الثانية] ومنها نستنتج أن :

- حدها الأدنى 3 ، وطولها 2 ، وتكرارها 29
- والتكرار المتجمع السابق يساوي مجموع تكرارات الفئات السابقة للفئة الوسيطية [أي تكرار الفئة الأولى فقط] = 14

الجدول التكراري		
التكرار f	المتغير (المساحة) x	الفئة
14	$1 \le x < 3$	الأولى
29	$3 \le x < 5$	الثانية
18	$5 \le x < 7$	الثالثة
9	$7 \le x < 10$	الرابعة
$\sum f = 70$		

إذن يمكن حساب الوسيط M من :

$$M = 3 + \left[\frac{(35-14)}{29} \times 2 \right] = 4.44827.... \cong 4.4$$

مثال (3–11) : باستخدام طريقة الاستكمال ، **قدر** قيمة الوسيط للبيانات الخاصة بالأجور السنوية للعاملين المعطاة في مثال (2–8) من الباب الثاني [الدرس الثالث] ، **وقارن** تلك النتيجة بما سبق وحصلت عليه بيانياً في مثال (2–8) .

الحل :

من الجدول التكراري لأجور العاملين يمكن استنتاج الآتي :

- نصف مجموع التكرارات $\frac{1}{2}\sum f \;=\; 30$
- مبتدئين بالصفر (في ذهننا) نقوم بإضافة تكرار الفئة الأولى (= 6) فنحصل على :

الجدول التكراري		
التكرار f	المتغير (الأجر) x	الفئة
6	$50 \le x < 60$	الأولى
9	$60 \le x < 70$	الثانية
15	$70 \le x < 80$	الثالثة
12	$80 \le x < 90$	الرابعة
9	$90 \le x < 100$	الخامسة
6	$100 \le x < 120$	السادسة
3	$120 \le x < 180$	السابعة
$\sum f = 60$		

$$0 + 6 = 6 \quad (< 30)$$

وبالتالي فإن الفئة الأولى ليست الفئة الوسيطية .

- نضيف للناتج السابق (= 6) تكرار الفئة الثانية (= 9) فنحصل على :

$$6 + 9 = 15 \quad (< 30)$$

وبالتالي فإن الفئة الثانية ليست الفئة الوسيطة .

- نضيف للناتج السابق (= 15) تكرار الفئة الثانية (= 15) فنحصل على :

$$15 + 15 = 30 \quad (= \tfrac{1}{2}\sum f)$$

إذن الفئة **الثالثة** هي الفئة الوسيطة .

- الحد الأدنى للفئة الوسيطة = 70 ، وطولها = 10 ، وتكرارها = 15 .
- والتكرار المتجمع السابق يساوي مجموع تكرارات الفئات السابقة للفئة الوسيطة [أي مجموع تكرارات الفئتين الأولى والثانية] = 9 + 6 = 15

إذن يمكن حساب الوسيط M من :

$$M = 70 + \left[\frac{(30-15)}{15} \times 10 \right] = 80$$

وهي نفس النتيجة التي حصلنا عليها سابقاً في مثال (2–8) .

مثال (3–12) : **قدر** قيمة الوسيط للبيانات الخاصة بأطوال سيقان الأزهار المعطاة في مثال (2–6) من الباب الثاني [الدرس الثالث] .

الحل :

من الجدول التكراري لأطوال السيقان يمكن استنتاج الآتي :

- نصف مجموع التكرارات $\frac{1}{2}\sum f$ = 25
- مبتدئين بالصفر (في ذهننا) نقوم بإضافة تكرار الفئة الأولى (= 4) فنحصل على :

$$0 + 4 = 4 \quad (< 25)$$

وبالتالي فإن الفئة الأولى ليست الفئة الوسيطة .

- نضيف للناتج السابق (= 4) تكرار الفئة الثانية (= 16) فنحصل على :

$$4 + 16 = 20 \quad (< 25)$$

وبالتالي فإن الفئة الثانية ليست الفئة الوسيطة .

- نضيف للناتج السابق (= 20) تكرار الفئة الثانية (= 12) فنحصل على :

$$20 + 12 = 32 \quad (> 25)$$

إذن الفئة **الثالثة** هي الفئة الوسيطة .

- الحد الأدنى للفئة الوسيطة = 30 ، وطولها = 5 ، وتكرارها = 12 .
- والتكرار المتجمع السابق يساوي مجموع تكرارات الفئات السابقة للفئة الوسيطة [أي مجموع تكرارات الفئتين الأولى والثانية] = 16 + 4 = 20

إذن يمكن حساب الوسيط M من :

	الجدول التكراري	
الفئة	المتغير x	التكرار f
الأولى	$0 \le x < 20$	4
الثانية	$20 \le x < 30$	16
الثالثة	$30 \le x < 35$	12
الرابعة	$35 \le x < 40$	10
الخامسة	$40 \le x < 50$	6
السادسة	$50 \le x < 60$	2
		$\sum f = 50$

$$M=30+\left[\frac{(25-20)}{12}\times 5\right]\cong 32.1$$

تدريب للقارئ : أعد الحل باستخدام إحدى الطرق البيانية وقارن نتيجتك بالنتيجة السابقة

مثال (3–13) : طُلِب من 3 مشرفين بإحدى المدارس تقسيم طلبة المدرسة إلى 3 مجموعات متساوية على أن يقوم كل مشرف بتقديم بيان عن فئات العمر المختلفة لطلبة مجموعته وعدد الطلبة في كل فئة من فئات العمر ، فكانت الجداول التكرارية المبينة :

المجموعة (1)

العدد f	العمر x	الفئة
20	$x<6$	الأولى
25	$6\ \le x<12$	الثانية
35	$12\ \le x<15$	الثال
18	$15\le x<18$	الرابعة
$\sum f=98$		

المجموعة (2)

العدد f	العمر x	الفئة
20	$6\le x<12$	الأولى
25	$12\ \le x<15$	الثانية
35	$15\ \le x<18$	الثالثة
18	$x\ge 18$	الرابعة
$\sum f=98$		

المجموعة (3)

العدد f	العمر x	الفئة
20	$x<6$	الأولى
25	$6\ \le x<12$	الثانية
35	$12\ \le x<15$	الثالثة
18	$x\ge 15$	الرابعة
$\sum f=98$		

هل يمكن من خلال البيانات السابقة حساب **الوسط الحسابي** لعمر الطلبة في كل مجموعة ؟ **علل إجابتك** .

وإذا كانت الإجابة بـ "**لا**" احسب **مقياساً مناسباً** يُعطي دلالة لمتوسط العمر في كل مجموعة .

الحل :

الإجابة هي " **لا** " للمجموعات الثلاث [أي لا يمكن حساب الوسط الحسابي للعمر] ، وها هي الأسباب :

- في المجموعة الأولى : الحد الأدنى للفئة الأولى غير معروف [يُقال للجدول عندئذٍ أنه مفتوح من أسفل]
- في المجموعة الثانية : الحد الأعلى للفئة الأخيرة غير معروف [يُقال للجدول عندئذٍ أنه مفتوح من أعلى]
- في المجموعة الثالثة : الحد الأدنى للفئة الأولى غير معروف وأيضاً الحد الأعلى للفئة الأخيرة غير معروف [يُقال للجدول عندئذٍ أنه مفتوح من الطرفين]

مثل هذه الجداول تُسمى **جداول تكرارية مفتوحة** :

- من **أسفل** [إذا كان الحد الأدنى للفئة الأولى غير معروف]
- من **أعلى** [إذا كان الحد الأعلى للفئة الأخيرة غير معروف]

- من **الطرفين** [إذا كان الحد الأدنى للفئة الأولى والحد الأعلى للفئة الأخيرة غير معروفين]

وحيث أن الوسيط لأي مجموعة من البيانات يعتمد في حسابه على البيانات الموجودة في المنتصف ، إذن يمكن استخدام **الوسيط** كمتوسط للدلالة على متوسط العمر في كل مجموعة :

بالنسبة للمجموعة الأولى من الطلبة :

المجموعة (**1**)		
الفئة	العمر x	العدد f
الأولى	$x < 6$	20
الثانية	$6 \le x < 12$	25
الثالثة	$12 \le x < 15$	35
الرابعة	$15 \le x < 18$	18
		$\sum f = 98$

- نصف مجموع التكرارات $49 = \frac{1}{2}\sum f$
- مبتدئين بالصفر (في ذهننا) نقوم بإضافة تكرار الفئة الأولى (= 20) فنحصل على :

$$0 + 20 = 20 \quad (< 49)$$

وبالتالي فإن الفئة الأولى ليست الفئة الوسيطة

- نضيف للناتج السابق (= 20) تكرار الفئة الثانية (= 25) فنحصل على :

$$20 + 25 = 45 \quad (< 49)$$

وبالتالي فإن الفئة الثانية ليست الفئة الوسيطة .

- نضيف للناتج السابق (= 45) تكرار الفئة الثانية (= 35) فنحصل على :

$$45 + 35 = 80 \quad (> 49)$$

إذن الفئة **الثالثة** هي الفئة الوسيطة .

- الحد الأدنى للفئة الوسيطة = 12 ، وطولها = 3 ، وتكرارها = 35 .
- والتكرار المتجمع السابق يساوي مجموع تكرارات الفئات السابقة للفئة الوسيطة [أي مجموع تكرارات الفئتين الأولى والثانية] = 25 + 20 = 45

إذن يمكن حساب الوسيط M من :

$$M = 12 + \left[\frac{(49 - 45)}{35} \times 3\right] \cong 12.3$$

وبنفس الطريقة يمكن التعامل مع المجموعتين (2) ، (3) [وعليك التأكد من صحة الحل]

للمجموعة الثانية : $M \cong 15.3$ للمجموعة الثالثة : $M \cong 12.3$

ملخص للدرس الخامس [الباب الثالث : مقاييس النزعة المركزية (الوسيط)]

- **الوسيط M لمجموعة من القيم [وهو أحد مقاييس النزعة المركزية] :** هو قيمة تقسم مجموعة القيم [بعد ترتيبها تصاعدياً] إلى مجموعتين متساويتين في العدد ، أي بحيث تقع 50% من القيم تحتها (أي أقل منها) و 50% من القيم فوقها (أي أكبر منها) .

- ولتحديد الوسيط لقيم المفردة [عددها n] يجب أولاً ترتيب القيم [تصاعدياً مثلاً] فيكون الوسيط هو القيمة التي تقع في المنتصف (إذا كانت n فردية) أو الوسط الحسابي للقيمتين بالمنتصف (إذا كانت n زوجية) .

- ويمكن للبيانات الكمية المتصلة أن نحدد الوسيط بيانياً من المضلع التكراري المتجمع الصاعد [أو الهابط] والتي هي قيم المتغير المناظر لتكرار متجمع قدره $\left[\frac{1}{2}\sum f\right]$ أو تكرار متجمع نسبي قدره 50% ، أيضاً الوسيط هو نقطة تقاطع المضلعين التكراريين الصاعد والهابط .

- كما يمكن [في حالة البيانات الكمية المتصلة] تحديد الوسيط حسابياً بطريقة الاستكمال بأن :

☐ نحدد أولاً الفئة الوسيطية : بأن نحسب نصف مجموع التكرارات $\left[\frac{1}{2}\sum f\right]$ ثم نبدأ (من الصفر في ذهننا) في إضافة تكرارات الفئات الواحدة تلو الأخرى ، ومع كل إضافة نقارن الناتج بـ $\left[\frac{1}{2}\sum f\right]$ ، فإن كان أقل نستمر في إضافة التكرارات حتى نصل إلى قيمة أكبر من أو تساوي $\left[\frac{1}{2}\sum f\right]$ فتكون آخر فئة أضفنا تكرارها هي الفئة الوسيطية .

☐ نحدد لهذه الفئة الوسيطية الآتي : حدها الأدنى – طولها – تكرارها .

☐ نحسب ما يُسمى بالتكرار المتجمع السابق [وهو مجموع تكرارات الفئات السابقة للفئة الوسيطية]

☐ نحسب قيمة الوسيط من :

$$\text{الوسيط } M = \text{الحد الأدنى للفئة الوسيطية} + \left[\frac{\text{(نصف مجموع التكرارات} - \text{التكرار المتجمع السابق)}}{\text{تكرار الفئة الوسيطية}} \times \text{طول الفئة الوسيطية}\right]$$

- وللوسيط المزايا والعيوب التالية :

– المزايا : سهولة حسابه بيانياً وحسابياً – لا يتأثر بالقيم المتطرفة – يمكن استخدامه كمقياس للنزعة المركزية وذلك للتوزيعات التكرارية المفتوحة .

– العيوب : يحتاج لترتيب القيم إما تصاعدياً أو تنازلياً – لا يأخذ في الاعتبار جميع القيم – لا يمكن تحديده للبيانات النوعية [أي يمكن حسابها فقط للبيانات الكمية]

تدريب عملي

في كل حالة من الحالات التالية ، **أوجد** الوسيط M للبيانات المعطاه :

(أ) 7 , 9 , 7 , 12 , 15 , 9 , 10 , 4 , 7

(ب) 7 , 5 , 6 , 12 , 8 , 10 , 1 , 4 , 6 , 10 , 5 , 2 , 3 , 4 , 11 , 8

(ج) 96 , 82 , 93 , 76 , 85

(د) 0.55 , 0.44 , 0.53 , 0.52 , 0.49 , 0.50 , 0.46 , 0.53

(هـ)

x	462	480	498	516	534	552	570	588	606	624
f	98	75	56	42	30	21	15	11	6	2

(و)

x	10 –	15 –	20 –	25 –	30 –	35 –	40 – 45
f	3	7	16	12	9	5	2

تدريبات (5)

الإجابة النهائية لجميع التمرينات موجودة في نهاية التدريب

اختر الإجابة الصحيحة

(1) **الوسيط** هو أحد مقاييس

(أ) النزعة المركزية (ب) التشتت (ج) الالتواء (د) التفرطح

(2) **الوسيط** لمجموعة من القيم المرتبة ترتيباً تصاعدياً أو تنازلياً هو :

(أ) القيمة الأكثر تكراراً

(ب) متوسط أكبر وأقل قيمتين

(ج) مجموع القيم مقسوماً على عددها .

(د) القيمة التي تقسم مجموعة القيم إلى مجموعتين متساويتين في العدد

(3) **الوسيط** لمجموعة القيم 4 5 8 9 4 يساوي

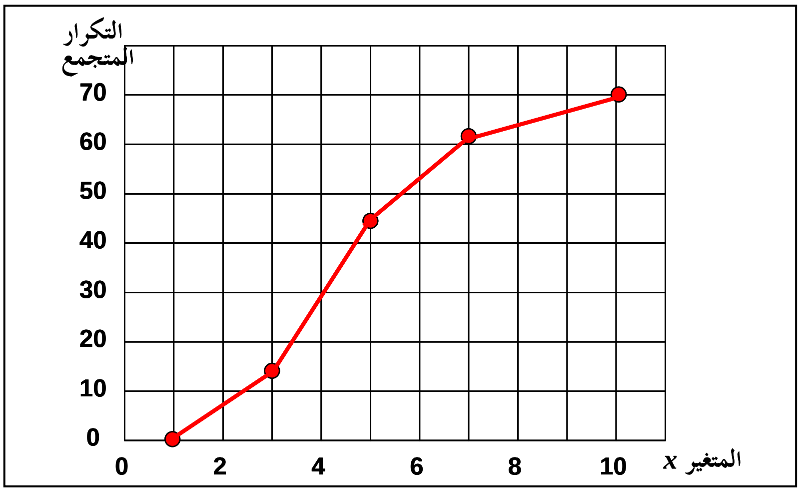
(أ) 8 (ب) 5 (ج) 4 (د) 6

(4) **الوسيط** لمجموعة القيم 16 4 8 2 3 9 يساوي

(أ) 6 (ب) 8 (ج) 7 (د) غير موجود

(5) **الوسيط** لمجموعة البيانات التي يمثلها المضلع التكراري المتجمع الصاعد لمتغير متصل x يقع بين :

(أ) 2 , 3 (ب) 3 , 4 (ج) 4 , 5 (د) 5 , 6



الإجابة : (1) أ (2) د (3) ب (4) أ (5) ج

عناصر الدرس

الباب الثالث : مقاييس النزعة المركزية [تابع]

- المنوال
- العلاقة بين المتوسطات الثلاثة : الوسط ، الوسيط ، والمنوال

الباب الثالث : مقاييس النزعة المركزية [تابع]

(3–5) المنوال (الشائع) *Mode*

(3–5–1) تعريف المنوال

يُعرف المنوال لمجموعة من القيم على أنه القيمة التي تتكرر أكثر من غيرها أو القيمة الأكثر شيوعاً [لذا يُسمى في بعض الأحيان بال ''**الشائع**''] . وأحياناً يُرمز للمنوال بالرمز $\hat{X}$

- فمثلاً مجموعة القيم

2 2 5 7 **9 9 9** 10 10 11 12 18

لها منوال **9**

- ومجموعة القيم

9 3 5 8 10 12 15 16

ليس لها منوال [أو **عديمة المنوال**]

- ومجموعة القيم

2 3 **4 4 4** 5 5 **7 7 7** 9

لها منوالان **7** , **4**

> أي أن مجموعة القيم قد تكون **وحيدة المنوال** [لها منوال واحد] ، وقد تكون **عديدة المنوال** [منوالان أو أكثر] وقد تكون **عديمة المنوال** [لا يوجد لها منوال]

- أما مجموعة القيم : 7 7 6 6 5 5 4 4 ، فقد تتسرع وتقول أنها رباعية المنوال ومناولها : 4 , 5 , 6 , 7 ، لكن [حيث أن جميع القيم لها نفس التكرار] هذه المجموعة الأخيرة **عديمة المنوال**

والمنوال [مقارنةً بالوسط الحسابي والوسيط] به العديد من **العيوب** منها :

- أنه لا يأخذ في الاعتبار جميع البيانات ولكنه يهتم فقط بالقيم الأكثر تكراراً .
- أنه قد لا يتواجد أو قد يكون هناك أكثر من منوال للبيانات .

إلا أنه أيضاً يتميز ببعض **المزايا** منها :

- أنه أسرع في تحديده من الوسط والوسيط
- من الممكن تحديده للتوزيعات التكرارية للبيانات المنفصلة سواء كانت تلك البيانات كمية متقطعة أو نوعية [والبيانات الأخيرة (النوعية) لا يمكن حساب الوسط الحسابي لها أو الوسيط] .

والأمثلة التالية توضح ذلك :

درجات طلاب في مقرر الإحصاء

عدد الطلاب	درجة الطالب
28	12
24	14
39	16
9	18

بيانات كمية متقطعة

لها منوال وحيد وهو "الدرجة "16

درجات طلاب في مقرر الإنجليزي

عدد الطلاب	درجة الطالب
23	12
30	14
30	16
17	18

بيانات كمية متقطعة

لها منوالان وهما

"14 , 16"

بيانات منفصلة

درجات طلاب في مقرر الفقه

عدد الطلاب	درجة الطالب
25	12
25	14
25	16
25	18

بيانات كمية متقطعة

ليس لها منوال

سيارات في أحد المواقف

عدد السيارات	لون السيارة
10	أحمر R
23	أزرق B
12	أبيض W
5	أصفر Y

بيانات نوعية

لها منوال وهو "اللون الأزرق"

وماذا عن تحديد المنوال للتوزيعات التكرارية للبيانات الكمية المتصلة ؟

في هذه الحالة يتم تحديد المنوال [تقريبياً] كالآتي :

حدد **الفئة المنوالية** [وهي الفئة التي يناظرها **أكبر كثافة تكرار**] ، فيكون **المنوال** (تقريباً) هو **مركز تلك الفئة**

فمثلاً في التوزيع التكراري المبين [مثال (2-6)/الدرس الثالث] يكون المنوال تقريباً 32.5

مثال (6–2)/الدرس الثالث

الجدول التكراري

	المتغير x	التكرار f	طول الفئة c	مركز الفئة x_0	كثافة التكرار
الفئة الأولى	$0 < x < 20$	4	20	10	0.2
الفئة الثانية	$20 < x < 30$	16	10	25	1.6
الفئة الثالثة	$30 < x < 35$	12	5	32.5	2.4
الفئة الرابعة	$35 < x < 40$	10	5	37.5	2
الفئة الخامسة	$40 < x < 50$	6	10	45	0.6
الفئة السادسة	$50 < x < 60$	2	10	55	0.2
		$\sum f = 50$			

أكبر كثافة تكرار

الفئة المنوالية هى الفئة الثالثة

المنوال = 32.5

وبالرغم من هذه البساطة إلا أن هناك **تنبيه** و **تحذير** :

بالنسبة للتنبيه : فالطريقة السابقة [اعتبار أن مركز الفئة المنوالية هو المنوال] هي طريقة **تقريبية** ، لكن هناك طرق أخرى **حسابية وبيانية** تعطي تقريباً أكثر دقة ، لكننا لن نتطرق لهذه الطرق في هذا المكان وذلك للتبسيط

وبالنسبة للتحذير : فالفئة المنوالية هي الفئة التي يناظرها **أكبر كثافة تكرار** وليس **أكبر تكرار** ، إلا فقط عندما تكون أطوال الفئات واحدة .

فمثلاً في المثال السابق [مثال (6–2)] ، إذا قلنا أن الفئة المنوالية هي الفئة المناظرة لأكبر تكرار فستكون تلك الفئة هي الفئة الثانية [وهذا خطأ] في حين أن الفئة المنوالية هي الفئة الثالثة [كما سبق وبينا] ، وإليك بعض الحالات التي قد تخدع :

حالة (1) [أنظر الجدول التكراري (1)] :

إذا اعتمدنا في ردنا على التكرار فستكون الإجابة : "التوزيع : ثنائي المنوال ، والفئات المنوالية هي الثانية والثالثة ، والمنوالان هما 25 و 37.5 [مراكز فئات المنوال]" . أما الإجابة **الصحيحة** فهي التي تعتمد على كثافة التكرار وليس التكرار : "التوزيع : وحيد المنوال ، والفئة المنوالية هي الثانية ، والمنوال هو 25 [مركز الفئة المنوالية]" .

الجدول التكراري (1)

	x	f	طول الفئة	مركز الفئة	كثافة التكرار
الفئة لأولى	$0 \le x < 20$	4	20	10	0.2
الفئة الثانية	$20 \le x < 30$	18	10	25	1.8
الفئة الثالثة	$30 \le x < 45$	18	15	37.5	1.2
الفئة الرابعة	$45 \le x < 55$	8	10	50	0.8

حالة (2) [أنظر الجدول التكراري (2)] :

إذا اعتمدنا في ردنا على التكرار فستكون الإجابة : "التوزيع : وحيد المنوال ، والفئة المنوالية هي الثانية ، والمنوال هو 40 [مركز الفئة المنوالية]" . أما الإجابة **الصحيحة** فهي التي تعتمد على كثافة التكرار وليس التكرار : "التوزيع : عديم المنوال [أي ليس هناك منوال]" .

الجدول التكراري (2)

كثافة التكرار	مركز الفئة	طول الفئة	f	x	
0.2	10	20	4	$0 \le x < 20$	الفئة لأولى
0.2	40	40	8	$20 \le x < 60$	الفئة الثانية
0.2	75	10	2	$60 \le x < 70$	الفئة الثالثة
0.2	72.5	5	1	$70 \le x < 75$	الفئة الرابعة

حالة (3) [أنظر الجدول التكراري (3)] :

إذا اعتمدنا في ردنا على التكرار فستكون الإجابة : "التوزيع : وحيد المنوال ، والفئة المنوالية هي الرابعة ، والمنوال هو 40 [مركز الفئة المنوالية]" . أما الإجابة **الصحيحة** فهي التي تعتمد على كثافة التكرار وليس التكرار : "التوزيع : ثنائي المنوال ، والفئات المنوالية هي الثانية و الثالثة ، والمنوالان هما 10 و 17.5 [مراكز الفئات المنوالية]" .

الجدول التكراري (3)

كثافة التكرار	مركز الفئة	طول الفئة	f	x	
0.8	2.5	5	4	$0 \le x < 5$	الفئة لأولى
1.6	10	10	16	$5 \le x < 15$	الفئة الثانية
1.6	17.5	5	8	$15 \le x < 20$	الفئة الثالثة
0.5	40	40	20	$20 \le x < 60$	الفئة الرابعة

أما في الحالة (4) [أنظر الجدول التكراري (4)] :

الجدول التكراري (4)

كثافة التكرار	مركز الفئة	طول الفئة	f	x	
1.8	5	10	18	$0 \le x < 10$	الفئة لأولى
2	15	10	20	$10 \le x < 20$	الفئة الثانية
2.5	25	10	25	$20 \le x < 30$	الفئة الثالثة
1.2	35	10	12	$30 \le x < 40$	الفئة الرابعة

لأن الفئات متساوية الأطوال [وبالتالي الفئة ذات أكبر كثافة تكرار هي نفسها الفئة ذات أكبر تكرار] فالفئة المنوالية هي الفئة ذات أكبر تكرار [أي الفئة الثالثة] ويكون المنوال هو 25 [مركز تلك الفئة] .

مثال (3–14) : في كل حالة من الحالات الآتية **أوجد** الوسط الحسابي $\bar{x}$ ، الوسيط M ، والمنوال $\hat{X}$:

(1) 3 , 5 , 2 , 6 , 5 , 9 , 5 , 2 , 8 , 6

(2) 5 , 4 , 8 , 3 , 7 , 2 , 9

(3) 18.3 , 20.6 , 19.3 , 22.4 , 20.2 , 18.8 , 19.7 , 20.0

(4) 85 , 76 , 93 , 82 , 94

(5) 6 [ست مرات] ، 7 [سبع مرات] ، 8 [ثماني مرات] ، 9 [تسع مرات] ، 10 [عشر مرات]

الحل :

(1) الوسط الحسابي :

$$\bar{x}=\frac{3+5+2+6+5+9+5+2+8+6}{10}=\underline{\underline{5.1}}$$

الوسيط :

البيانات الأصلية : 3 , 5 , 2 , 6 , 5 , 9 , 5 , 2 , 8 , 6

البيانات بعد ترتيبها تصاعدياً : 2 , 2 , 3 , 5 , **5 , 5** , 6 , 6 , 8 , 9 [عددها 10 (زوجي) ، إذن الوسيط هو الوسط الحسابي للقيمتين الخامسة و السادسة] ، إذن :

$$M=\frac{5+5}{2}=\underline{\underline{5}}$$

المنوال : هو القيمة الأكثر تكراراً [3 , **5** , 2 , 6 , **5** , 9 , **5** , 2 , 8 , 6] ، أي أن :

$$\hat{X}=\underline{\underline{5}}$$

—

(2) الوسط الحسابي :

$$\bar{x}=\frac{5+4+8+3+7+2+9}{7}=5.42857\cong\underline{\underline{5.43}}$$

الوسيط :

البيانات بعد ترتيبها تصاعدياً : 2 , 3 , 4 , **5** , 7 , 8 , 9 [عددها 7 (فردي) ، إذن الوسيط هو القيمة الرابعة] ، إذن : $M=5$

المنوال : هو القيمة الأكثر تكراراً [لا يوجد] .

—

(3) الوسط الحسابي :

$$\bar{x}=\frac{18.3+20.6+19.3+22.4+20.2+18.8+19.7+20}{8}=19.9125\cong\underline{\underline{19.91}}$$

الوسيط :

البيانات بعد ترتيبها تصاعدياً : 22.4 , 20.6 , 20.2 , **20 , 19.7** , 19.3 , 18.8 , 18.3 [عددها 8 (زوجي) ، إذن الوسيط هو الوسط الحسابي للقيمتين الرابعة و الخامسة] ، أي :

$$M=\frac{19.7+20}{2}=\underline{\underline{19.85}}$$

المنوال : هو القيمة الأكثر تكراراً [لا يوجد] .

—

(4) الوسط الحسابي :
$$\bar{x}=\frac{85+76+93+82+94}{5}=\underline{\underline{86}}$$

الوسيط :

البيانات بعد ترتيبها تصاعدياً : 94 , 93 , **85** , 82 , 76 [عددها 5 (فردي) ، إذن الوسيط هو القيمة الثالثة] ، إذن :
$$M=85$$

المنوال : هو القيمة الأكثر تكراراً [لا يوجد] .

—

(5) الوسط الحسابي :

$$\bar{x}=\frac{\sum fx}{\sum f}=\frac{(6\times 6)+(7\times 7)+(8\times 8)+(9\times 9)+(10\times 10)}{6+7+8+9+10}$$
$$=\frac{36+49+64+81+100}{40}=\frac{330}{40}=\underline{\underline{8.25}}$$

ويمكنك الحل بعمل جدول تكراري [لكن لا يستدعي الأمر ذلك] .

x	f	fx
6	6	36
7	7	49
8	8	64
9	9	81
10	10	100
	40	330
	$\sum f$	$\sum fx$

$$\boxed{\bar{x}=\frac{\sum fx}{\sum f}=\frac{330}{40}=\underline{\underline{8.25}}}$$

الوسيط : عدد القيم n هنا هو مجموع التكرارات [أي $\sum f=40$] زوجي ، وبالتالي هناك قيمتان في المنتصف ترتيبهما 21 , 20 ، وبالتالي يكون الوسيط هو الوسط الحسابي للقيمتين العشرين و الواحد وعشرين ، وهاتان القيمتان هما **8 , 8** ، أي أن :
$$M=\frac{8+8}{2}=\underline{\underline{8}}$$

توضيح

القيم مرتبة	6	6	6	6	6	6	7	7	7	7	7	7	7	8	8	8	8	8	8	8	8	9	9	....
رتبتها	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	....

الـ 6 قيم الأولى هي الرقم 6 [ذات التكرار 6] ثم يليها الرقم 7 بتكرار 7 [وبذلك يحتل الرقم 7 الترتيب من 7 حتى 13] ، ثم يليها الرقم 8 بتكرار 8 [وبذلك يحتل الرقم 8 الترتيب من 14 حتى 21] ، أي أن الرقم 8 يحتل الترتيبين 20 , 21 ، وبالتالي يكون الوسيط هو الوسط الحسابي للقيمتين 8 , 8

المنوال : هو القيمة الأكثر تكراراً ، إذن : $\hat{X} = \underline{\underline{10}}$

مثال (3−15) : الجدول التكراري المرافق يبين توزيع أقطار رؤوس مسامير x [بالمليمتر] منتجة بواسطة شركة ما ، **احسب** الوسط الحسابي $\bar{x}$ ، الوسيط M ، والمنوال $\hat{X}$ للأقطار .

الجدول التكراري		
f	x	
3	$10 \le x < 15$	الفئة لأولى
7	$15 \le x < 20$	الفئة الثانية
16	$20 \le x < 25$	الفئة الثالث
12	$25 \le x < 30$	الفئة الرابعة
9	$30 \le x < 35$	الفئة الخامسة
5	$35 \le x < 40$	الفئة السادسة
2	$40 \le x < 45$	الفئة السابعة

الحل :

fx_0	x_0	f	x	
37.5	12.5	3	$10 \le x < 15$	الفئة لأولى
122.5	17.5	7	$15 \le x < 20$	الفئة الثانية
360	22.5	16	$20 \le x < 25$	الفئة الثالثة
330	27.5	12	$25 \le x < 30$	الفئة الرابعة
292.5	32.5	9	$30 \le x < 35$	الفئة الخامسة
187.5	37.5	5	$35 \le x < 40$	الفئة السادسة
85	42.5	2	$40 \le x < 45$	الفئة السابعة
1415		54		

من الجدول التكراري المرافق للبيانات يمكن تحديد مركز كل فئة [عمود x_0] ومن ثم حاصل ضرب مركز كل فئة في تكرارها [عمود fx_0] وبالتالي يكون الوسط الحسابي لأقطار المسامير هو :

$$\bar{x} = \frac{\sum f\, x_0}{\sum f} = \frac{1415}{54} = 26.2037037 \cong \underline{\underline{26.20}}$$

أي أن **الوسط الحسابي** للدخل السنوي هو 83.75 ألفاً من الريالات .

وبالنسبة للمنوال :

	f	x	
	3	10 ≤ x < 15	الفئة لأولى
	7	15 ≤ x < 20	الفئة الثانية
الفئة المنوالية	16	20 ≤ x < 25	الفئة الثالثة
	12	25 ≤ x < 30	الفئة الرابعة
	9	30 ≤ x < 35	الفئة الخامسة
	5	35 ≤ x < 40	الفئة السادسة
	2	40 ≤ x < 45	الفئة السابعة
	54		

هنا الفئات متساوية الطول وبالتالي تكون الفئة المنوالية هي الفئة التي يناظرها أكبر تكرار [لأن ''أكبر تكرار'' يناظره في هذا السؤال ''أكبر كثافة تكرار''] ، وبالتالي تكون إذن **الفئة المنوالية هنا هي الفئة الثالثة** ومركزها [22.5] هو تقريباً **المنوال** .

أما بالنسبة للوسيط ، فيمكن تحديده حسابياً وتخطيطياً كالآتي :

f	x	الفئة
3	10 ≤ x < 15	لأولى
7	15 ≤ x < 20	الثانية
16	20 ≤ x < 25	الثالثة
12	25 ≤ x < 30	الرابعة
9	30 ≤ x < 35	الخامسة
5	35 ≤ x < 40	السادسة
2	40 ≤ x < 45	السابعة
54		

(أ) الوسيط حسابياً [طريقة الاستكمال]

- $\frac{1}{2}\sum f = \frac{54}{2} = 27$
- نبدأ بالصفر [في ذهننا]
- نزود على الصفر السابق تكرار الفئة الأولى [=3] ينتج 3 [وهي أقل من 27]
- نزود على ال 3 السابقة تكرار الفئة الثانية [=7] ينتج 10 [ما زال أقل من 27]
- نزود على ال 10 السابقة تكرار الفئة الثالثة [=16] ينتج 26 [ما زال أقل من 27]
- نزود على ال 26 السابقة تكرار الفئة الرابعة [=12] ينتج 38 [أكبر من 27] ،

إذن الفئة الرابعة هي الفئة الوسيطية

حدها الأدنى 25 وطولها 5 [25 – 30 =] ، وتكرارها 12 .

- التكرار المتجمع السابق = مجموع تكرارات الفئات الأولى والثانية والثالثة ، أي

$$3+7+16=26$$

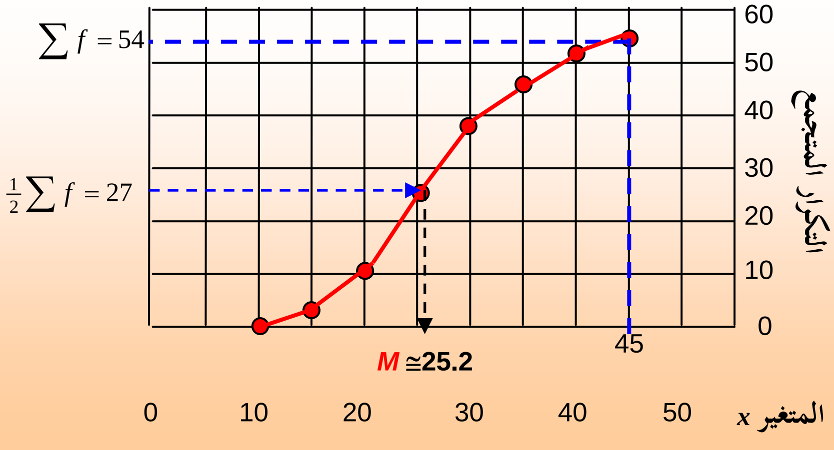
وبالتالي يكون الوسيط M مساوياً لـ :

$$M=25+\left[\frac{(27-26)}{12}\times 5\right]=25+\left[\frac{1}{12}\times 5\right]=25+0.41666=25.416666\cong 25.42$$

(ب) الوسيط بيانياً [المضلع التكراري المتجمع الصاعد]

نكون الجدول التكراري المتجمع الصاعد ، ومنه نرسم المضلع المتجمع الصاعد ، فتكون قيمة الوسيط هي قيمة المتغير x المناظرة لتكرار متجمع قدره $\frac{1}{2}\sum f=27$.

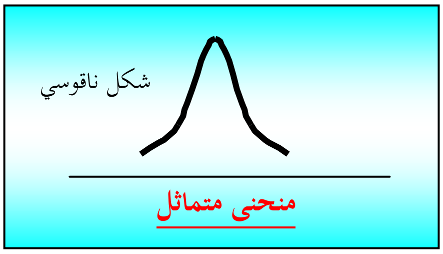
النقطة	متجمع f	x
(10 , 0)	0	$x<10$
(15 , 3)	3	$x<15$
(20 , 10)	10	$x<20$
(25 , 26)	26	$x<25$
(30 , 38)	38	$x<30$
(35 , 47)	47	$x<35$
(40 , 52)	52	$x<40$
(45 , 54)	54	$x<45$



(6–3) العلاقة بين المتوسطات الثلاثة : الوسط ، الوسيط ، والمنوال

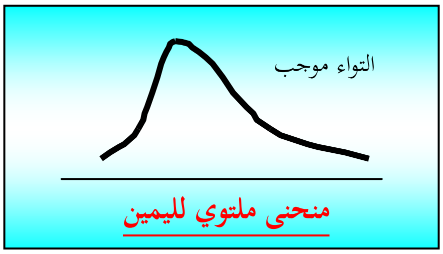
قبل أن نتعرف على وجود علاقة بين المتوسطات الثلاثة السابق التعرض لها ، وهي الوسط الحسابي ، والوسيط ، والمنوال من عدم وجودها ، دعنا نلقي الضوء على بعض أشكال التوزيعات التكرارية شائعة الظهور ، فالمنحنيات التكرارية التي تظهر في الناحية العملية تأخذ أشكالاً مميزة كالأشكال التالية :

- منحنى متماثل (ناقوسي) : وفيه تكون النهاية العظمى في المنتصف وتكون المشاهدات المتساوية البعد عن مركز النهاية العظمى لها نفس التكرارات



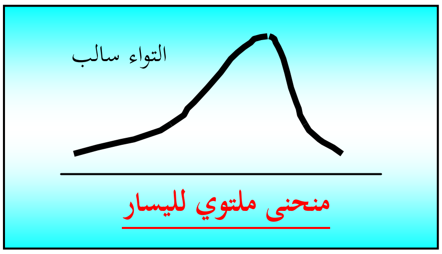
منحنى متماثل

- منحنى ملتوي لليمين (التواء موجب) : وفيه تكون يكون المنحنى قريباً من التماثل لكن طرفه على الجانب **الأيمن** من مركز النهاية العظمى يكون ممتداً أكثر من طرفه على الجانب الأيسر من هذا المركز ، وفي هذه الحالة يُقال للمنحنى أن له **التواء موجب** .



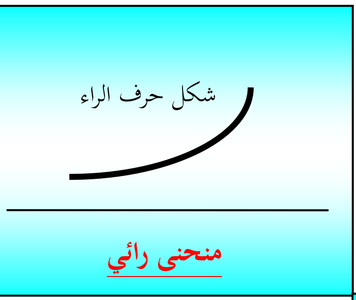
منحنى ملتوي لليمين

- منحنى ملتوي لليسار (التواء سالب) : وفيه تكون يكون المنحنى قريباً من التماثل لكن طرفه على الجانب **الأيسر** من مركز النهاية العظمى يكون ممتداً أكثر من طرفه على الجانب الأيمن من هذا المركز ، وفي هذه الحالة يُقال للمنحنى أن له **التواء سالب** .

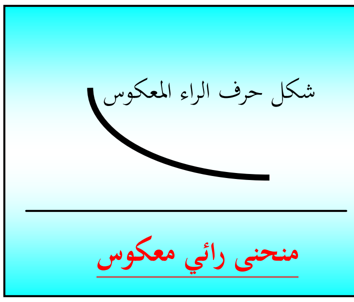


منحنى ملتوي لليسار

- المنحنى الرائي والرائي المعكوس : في المنحنيات ذات الشكل الرائي أو الرائي المعكوس فإن نقطة النهاية العظمى للمنحنى تقع عند أحد طرفي المنحنى .

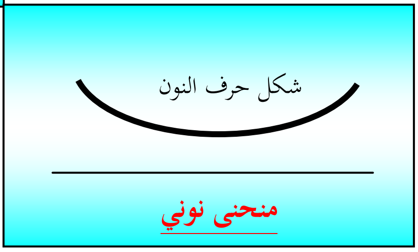


منحنى رائي



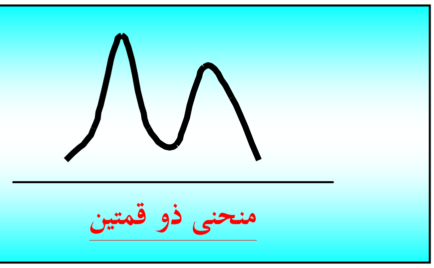
منحنى رائي معكوس

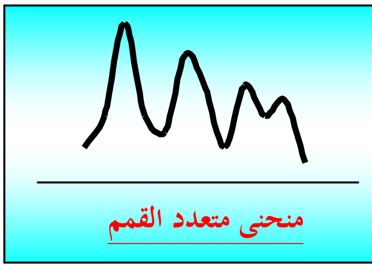
- المنحنى النوني : وله نهاية عظمى عند كلٍ من طرفيه .



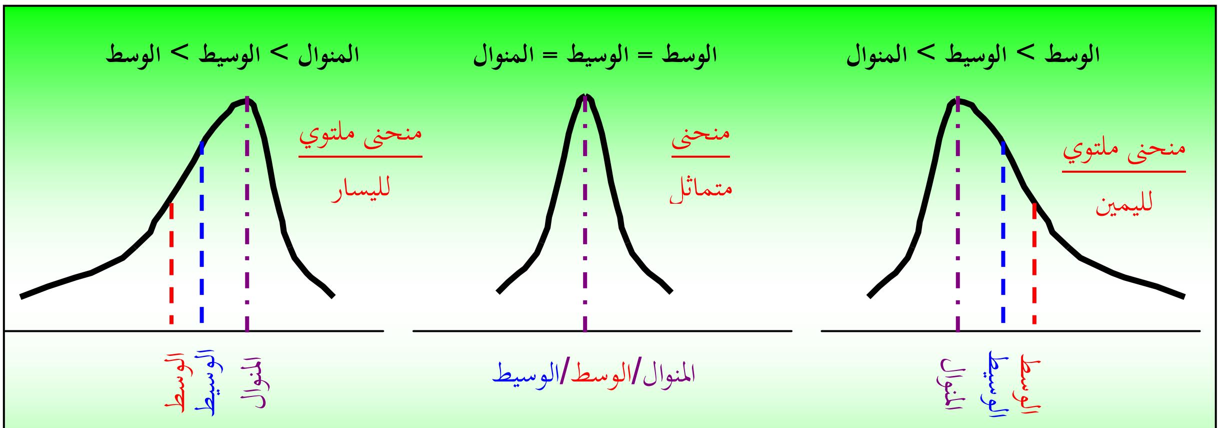
منحنى نوني

- المنحنى متعدد القمم : وله أكثر من نهاية عظمى .


منحنى ذو قمتين


منحنى متعدد القمم

والمنحنيات التكرارية **وحيدة المنوال والبسيطة الالتواء** تحقق العلاقة الاعتبارية التالية :



في جميع المنحنيات المبينة ، لاحظ أن المنوال هو القيمة المناظرة لأعلى نقطة في المنحنى وأن الوسيط يقع دائماً بين الوسط و المنوال

الوسط – المنوال = 3 × (الوسط – الوسيط)

والتي تعني أن المسافة بين الوسط والمنوال هي ثلاثة أمثال المسافة بين الوسط والوسيط

وهذه العلاقة يمكن وضعها على أي صورة من الصور التالية :

$$\text{الوسط} = \frac{(3 \times \text{الوسيط}) - \text{المنوال}}{2}$$

وهذه الصورة مفيدة عندما يكون الوسيط و المنوال معلومان ونريد معرفة الوسط الحسابي

أو

$$\text{المنوال} = (3 \times \text{الوسيط}) - (2 \times \text{الوسط})$$

وهذه الصورة مفيدة عندما يكون الوسط الحسابي و الوسيط معلومان ونريد معرفة المنوال

أو

$$\text{الوسيط} = \frac{(2 \times \text{الوسط}) + \text{المنوال}}{3}$$

وهذه الصورة مفيدة عندما يكون الوسط الحسابي و المنوال معلومان ونريد معرفة الوسيط

- فمثلاً إذا كان المنوال $\hat{X}$ لمجموعة من القيم = 95 ، والوسيط M لها = 85 ، فإن الوسط الحسابي $\bar{x}$ لها يكون :

$$\bar{x} = \frac{3M - \hat{X}}{2} = \frac{3 \times 85 - 95}{2} = \frac{255 - 95}{2} = \frac{160}{2} = 80$$

- وإذا كان الوسط الحسابي $\overline{x}$ لمجموعة من القيم = 80 ، والوسيط M لها = 85 ، فإن المنوال $\hat{X}$ لها يكون :

$$\hat{X} = 3M - 2\overline{x} = 3\times 85 - 2\times 80 = 255 - 160 = 95$$

- وإذا كان الوسط الحسابي $\overline{x}$ لمجموعة من القيم = 80 ، والمنوال $\hat{X}$ لها = 95 ، فإن الوسيط M لها يكون :

$$M = \frac{2\overline{x} + \hat{X}}{3} = \frac{2\times 80 + 95}{3} = \frac{160 + 95}{3} = \frac{255}{3} = 85$$

هذا بالطبع [في كل ما سبق] باعتبار أن التوزيع للبيانات هو توزيع وحيد المنوال وبسيط الالتواء ، والالتواء في كل الأمثلة السابقة هو التواء سالب (أي لليسار) ، [لماذا ؟] .

ملخص للدرس السادس [الباب الثالث : مقاييس النزعة المركزية (المنوال)]

- يُعرف **المنوال** لمجموعة من القيم على أنه القيمة التي تتكرر أكثر من غيرها أو القيمة الأكثر شيوعاً [لذا يُسمى في بعض الأحيان بال ''**الشائع**''] . وأحياناً يُرمز للمنوال بالرمز $\hat{X}$ ، فمثلاً

❖ لمجموعة القيم : 18 12 11 10 10 **9** **9** **9** 7 5 2 2 لها منوال وحيد هو 9 [مجموعة وحيدة المنوال]

❖ والمجموعة : 16 15 12 10 8 5 3 9 ليس لها منوال [عديمة المنوال]

❖ والمجموعة : 9 **7** **7** **7** 5 5 **4** **4** **4** 3 2 لها منوالان هما 4 , 7 [ثنائية المنوال]

❖ والمجموعة : 7 7 6 6 5 5 4 4 ليس لها منوال [وليست رباعية المنوال]

❖ أما المجموعة : **8** 7 7 6 6 5 5 4 4 فهي رباعية المنوال [هل لاحظت الفرق بين المجموعتين الأخيرتين]

- وللتوزيعات التكرارية المبينة يكون المنوال هو القيمة المناظرة لأكبر تكرار :

عدد الطلاب	الدرجة
28	12
24	14
39	16
9	18

بيانات كمية متقطعة

لها منوال وحيد وهو "الدرجة **16**"

عدد الطلاب	الدرجة
23	12
30	14
30	16
17	18

بيانات كمية متقطعة

لها منوالانوهما "**14 , 16**"

عدد الطلاب	الدرجة
25	12
25	14
25	16
25	18

بيانات كمية متقطعة

ليس لها منوال

عدد السيارات	لون السيارة
10	أحمر R
23	أزرق B
12	أبيض W
5	أصفر Y

بيانات نوعية

لها منوال وهو "اللون الأزرق"

- وللتوزيعات التكرارية المتصلة يكون المنوال (تقريباً) هو مركز **الفئة المنوالية** [وهي **الفئة التي لها أكبر كثافة تكرار**] ، وفي حالة تساوي أطوال الفئات تكون هي الفئة المناظرة لأكبر تكرار .

- وللمنوال **المزايا التالية :** سهل تحديده (إن وُجِد) – لا يتأثر بالقيم المتطرفة – لا يحتاج لترتيب البيانات – يمكن تحديده (في حالة وجوده) للبيانات النوعية .

- وله **العيوب التالية :** قد لا يتواجد – قد يكون غير وحيد .

- وفي حالة المنحنيات وحيدة المنوال وبسيطة الالتواء ، هناك علاقة اعتبارية (تقريبية) بين مقاييس النزعة المركزية : الوسط الحسابي ، الوسيط ، المنوال على الصورة :

الوسط – المنوال = 3 (الوسط – الوسيط)

- وإذا كان المنحنى :

- متماثلاً يكون : الوسط = الوسيط = المنوال
- ملتوياً قليلاً لليمين (التواء موجب) يكون : الوسط أكبر من الوسيط أكبر من المنوال
- ملتوياً قليلاً لليسار (التواء سالب) يكون : الوسط أصغر من الوسيط أصغر من المنوال

تدريب عملي

في كل حالة من الحالات التالية ، أوجد المنوال $\hat{X}$ للبيانات المعطاه :

(أ) 7 , 4 , 10 , 9 , 15 , 12 , 7 , 9 , 7

(ب) 8 , 11 , 4 , 3 , 2 , 5 , 10 , 6 , 4 , 1 , 10 , 8 , 12 , 6 , 5 , 7

(ج) 85 , 76 , 93 , 82 , 96

(د) 0.53 , 0.46 , 0.50 , 0.49 , 0.52 , 0.53 , 0.44 , 0.55

(هـ)

x	462	480	498	516	534	552	570	588	606	624
f	98	75	56	42	30	21	15	11	6	2

(و)

x	10 –	15 –	20 –	25 –	30 –	35 –	40 – 45
f	3	7	16	12	9	5	2

تدريبات (6)

الإجابة النهائية لجميع التمرينات موجودة في نهاية التدريب

اختر الإجابة الصحيحة

(1) **المنوال** هو أحد مقاييس

(أ) النزعة المركزية (ب) التشتت (ج) الالتواء (د) التفرطح

(2) في **المنحنى المتماثل** يكون

(أ) الوسط أكبر من المنوال (ب) الوسط ضعف المنوال

(ج) المنوال أكبر من الوسط (د) الوسط = المنوال

(3) في **التوزيعات وحيدة المنوال وبسيطة الالتواء لليمين** يكون

(أ) الوسط أكبر من المنوال (ب) الوسط ضعف المنوال

(ج) المنوال أكبر من الوسط (د) الوسط = المنوال

(4) في **التوزيعات وحيدة المنوال وبسيطة الالتواء لليسار** يكون

(أ) الوسط أكبر من المنوال (ب) الوسط ضعف المنوال

(ج) المنوال أكبر من الوسط (د) الوسط = المنوال

(5) **المنوال** لمجموعة من القيم هو :

(أ) القيمة التي تقسم مجموعة القيم إلى مجموعتين متساويتين في العدد

(ب) القيمة الأكثر تكراراً

(ج) متوسط أكبر وأقل قيمتين

(د) مجموع القيم مقسوماً على عددها .

(6) لمجموعة من القيم ، فإن القيمة الأكثر تكراراً (إن وُجِدت) تُسمى :

(أ) الوسط الحسابي (ب) الوسيط

(ج) المنوال (د) المدى

(7) لمجموعة من البيانات الكمية المتصلة (فئات غير متساوية الطول) تكون الفئة المنوالية هي الفئة :

(أ) الأكبر طولاً (ب) الأكثر تكراراً

(ج) الفئة الوسطى (د) الأكثر كثافة تكرار

(8) أحد مقاييس النزعة المركزية الذي قد يمكن تحديده للبيانات النوعية :

(أ) الوسط الحسابي (ب) المنوال

(ج) الوسيط (د) المدى

(9) للمنحنيات التكرارية وحيدة المنوال وبسيطة الالتواء يكون :

(أ) الوسط − الوسيط = 3 × (الوسط − المنوال)

(ب) الوسيط − المنوال = 3 × (الوسيط − الوسط)

(ج) الوسط − المنوال = 3 × (الوسط − الوسيط)

(د) المنوال − الوسيط = 3 × (المنوال − الوسط)

(10) لمجموعة القيم 4 5 8 9 4 ، المنوال:

(أ) 8 (ب) 5 (ج) 4 (د) 6

(11) لمجموعة القيم 16 4 8 2 3 9 ، المنوال

(أ) 6 (ب) 8 (ج) 7 (د) غير موجود

خاص بالأسئلة من (12) إلى (19) :

الشكل المرافق يبين عدة توزيعات لمتغير متصل x :

الجدول التكراري (1)		
f	x	
4	$0 \le x < 20$	الفئة لأولى
18	$20 \le x < 30$	الفئة الثانية
18	$30 \le x < 45$	الفئة الثالثة
8	$45 \le x < 55$	الفئة الرابعة

الجدول التكراري (2)		
f	x	
4	$0 \le x < 20$	الفئة لأولى
8	$20 \le x < 60$	الفئة الثانية
2	$60 \le x < 70$	الفئة الثالثة
1	$70 \le x < 75$	الفئة الرابعة

الجدول التكراري (3)		
f	x	
4	$0 \le x < 5$	الفئة لأولى
16	$5 \le x < 15$	الفئة الثانية
8	$15 \le x < 20$	الفئة الثالثة
20	$20 \le x < 60$	الفئة الرابعة

الجدول التكراري (4)		
f	x	
18	$0 \le x < 10$	الفئة لأولى
20	$10 \le x < 20$	الفئة الثانية
25	$20 \le x < 30$	الفئة الثالثة
12	$30 \le x < 40$	الفئة الرابعة

(12) للتوزيع التكراري (1) ، الفئة المنوالية هي :

(أ) الأولى (ب) الثانية

(ج) الثانية والثالثة (د) غير موجودة

(13) للتوزيع التكراري (2) ، الفئة المنوالية هي :

(أ) الأولى (ب) الثانية

(ج) الثانية والثالثة (د) غير موجودة

(14) للتوزيع التكراري (3) ، الفئة المنوالية هي :

(أ) الأولى (ب) الثانية

(ج) الثانية والثالثة (د) الرابعة

(15) للتوزيع التكراري (4) ، الفئة المنوالية هي :

(أ) الأولى (ب) الثانية

(ج) الثالثة (د) الرابعة

(16) للتوزيع التكراري (1) ، المنوال هو (تقريباً) :

(أ) 10 (ب) 25

(ج) 37.5 , 25 (د) غير موجود

(17) للتوزيع التكراري (2) ، المنوال هو (تقريباً) :

(أ) 10 (ب) 25

(ج) 37.5 , 25 (د) غير موجود

(18) للتوزيع التكراري (3) ، المنوال هو (تقريباً) :

(أ) 5 (ب) 10

(ج) 17.5, 10 (د) 17.5

(19) للتوزيع التكراري (4) ، المنوال هو (تقريباً) :

(أ) 5 (ب) 15

(ج) 25 (د) 35

الإجابة : (1) أ (2) د (3) أ (4) ج (5) ب

(6) ج (7) د (8) ب (9) ج (10) ج

(11) د (12) د (13) ب (14) ج (15) د

(16) د (17) ب (18) ج (19) د

عناصر الدرس

الباب الرابع : مقاييس التشتت

- تعريف التشتت
- المدى
- الانحراف المتوسط (متوسط الانحرافات)
- الانحراف المعياري

الدرس السابع

الباب الرابع : مقاييس التشتت

(4–1) تعريف التشتت

الدرجة التي تتجه بها البيانات الكمية للانتشار حول قيمة متوسطة (أحد مقاييس النزعة المركزية) تُسمى تشتت أو تغير البيانات

فمثلاً إذا كان لدينا 3 مجموعات من الطلاب ، كل مجموعة مكونة من خمسة طلاب ، وكانت لها الدرجات التالية (من 10 درجات) في أحد المقررات :

المجموعة الأولى	المجموعة الثانية	المجموعة الثالثة
5 , 5 , 5 , 5 , 5	3 , 4 , 5 , 6 , 7	1 , 2 , 5 , 8 , 9
وسطها الحسابي 5	وسطها الحسابي 5	وسطها الحسابي 5

المجموعات الثلاثة لها وسط حسابي 5 ، لكن في المجموعة الأولى : جميع القيم متساوية وتساوي الوسط 5 ، في حين تنتشر البيانات في المجموعة الثانية حول هذا الوسط بقدرٍ ما ، وفي المجموعة الثالثة تنتشر البيانات حول الوسط بقدرٍ آخر ، أي أن الوسط الحسابي وحده [وهو ممثل لمقياس نزعة مركزية ، أي قيمة نموذجية ممثلة للبيانات] ليس كافياً وحده لوصف البيانات ، ولكن لابد من وجود نوع آخر من المقاييس لرصد مدى تشتت البيانات عن تلك القيمة المتوسطة الممثلة للبيانات .

وهناك العديد من المقاييس التي يمكن استخدامها لقياس هذا التشتت ولكن أكثرها شيوعاً :

المدى – الانحراف المتوسط – الانحراف المعياري – الانحراف الربيعي – الانحراف المئيني

ولنتعرف على كلٍ منها الآن .

(4–2) المدى

يُعرف المدى R لمجموعة من البيانات الكمية على أنه **الفرق** بين **أكبر** قيمة في البيانات **وأقل** قيمة فيها .

فمثلاً لمجموعة القيم [مجموعة (1)] :

15 13 3 5 18 12 6 7 3 15

يكون المدى :

$$R = 18 - 3 = 15$$

ولمجموعة القيم [مجموعة (2)] :

16 14 13 17 18 17 15 14 3 16

يكون المدى أيضاً :

$$R = 18 - 3 = 15$$

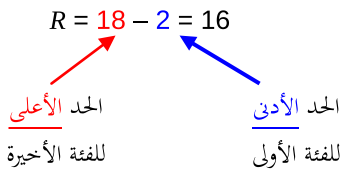
أي أن المدى واحد للمجموعتين في حين يبدو للعين المجردة أن هناك تشتت للبيانات أكبر في المجموعة (1) السابقة عنه في المجموعة (2) السابقة ، مما يعني أن المدى هنا لا يظهر هذا الفارق ، لذا يُعد المدى مقياساً للتشتت لكنه **غير جيد** في كثير من الأحيان .

وبالرغم من بساطة تحديده إلا أن بعض العيوب [مثل تأثره بالقيم المتطرفة كما اتضح من المثال السابق عند حسابه للمجموعة الثانية حيث تأثر بالقيمة المتطرفة 3] ، فإذا استبعدنا تلك القيمة يكون المدى (للمجموعة الثانية) مساوياً لـ :

$$R = 18 - 13 = 5$$

أيضاً من بين عيوبه أنه لا يمكن تحديده في حالة التوزيعات التكرارية المفتوحة ، كما يتضح من التوزيعات المبينة :

الفئة	العمر x
الأولى	$2 \le x < 6$
الثانية	$6 \le x < 12$
الثالثة	$12 \le x < 15$
الرابعة	$15 \le x < 18$



مفتوح من أسفل

الفئة	العمر x
الأولى	$x < 6$
الثانية	$6 \le x < 12$
الثالثة	$12 \le x < 15$
الرابعة	$15 \le x < 18$

مفتوح من أعلى

الفئة	العمر x
الأولى	$6 \le x < 12$
الثانية	$12 \le x < 15$
الثالثة	$15 \le x < 18$
الرابعة	$x \ge 18$

مفتوح من الطرفين

الفئة	العمر x
الأولى	$x < 6$
الثانية	$6 \le x < 12$
الثالثة	$12 \le x < 15$
الرابعة	$x \ge 15$

لا يمكن تحديد مدى البيانات

(4–3) الانحراف المتوسط [أو متوسط الانحرافات] $M.D$

يُعرف **الانحراف المتوسط** (أو **متوسط الانحرافات**) [وسنرمز له بالرمز $M.D$] على أنه متوسط القيم المطلقة للانحرافات عن قيمة متوسطة للبيانات [عادةً تكون الوسط الحسابي أو الوسيط] .

فإذا اعتبرنا أن القيمة المتوسطة للبيانات هي **الوسط الحسابي** ، فإن الانحراف المتوسط لمجموعة من البيانات $x_1 , x_2 , \cdots\cdots , x_n$ عددها n يُعطى بـ :

$$M.D = \frac{\sum_{i=1}^{n} | d_i |}{n} = \frac{\sum_{i=1}^{n} | x_i - \bar{x} |}{n}$$

حيث $d_i = x_i - \bar{x}$ هي انحراف القيمة x_i عن الوسط الحسابي $\bar{x}$ لمجموعة القيم ، أما $| d_i |$ فهي القيمة المطلقة للانحراف d_i .

ملحوظة هامة : القيمة المطلقة لأي عدد y هي القيمة العددية له دون إشارة ، ونرمز له بنفس الرمز y لكن بين خطين رأسيين | | ، أي نكتب القيمة المطلقة لـ y على الصورة $|y|$. فمثلاً :

$|3| = 3$, $|- 3| = 3$, $|2.5| = 2.5$, $|- 3.25| = 3.25$

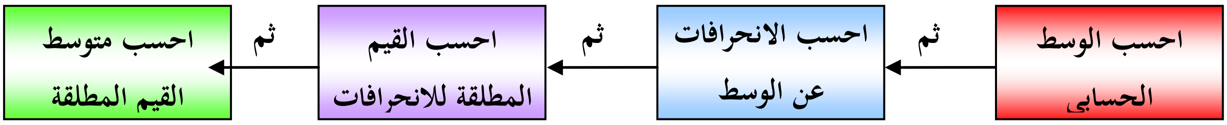
وهكذا .

وللتبسيط سنكتب الانحراف المتوسط على الصورة :

$$\boxed{M.D = \frac{\sum | d |}{n} = \frac{\sum | x - \bar{x} |}{n}} \qquad (4\text{-}1)$$

حيث يكون مفهوماً أن $\sum | d |$ تعني $\sum_{i=1}^{n} | d_i |$ [أي مجموع القيم المطلقة لانحرافات القيم عن وسطها الحسابي $\bar{x}$] .

إذن لحساب الانحراف المتوسط (أو متوسط الانحرافات) $M.D$ لمجموعة من القيم يلزم حساب الوسط الحسابي أولاً ، ثم نحسب انحرافات كل قيمة من هذه القيم عن الوسط الحسابي ، ثم القيم المطلقة لهذه الانحرافات ، ثم متوسط هذه القيم المطلقة كما هو مبين بالمخطط التالي :



فمثلاً : لمجموعتي القيم التي تعاملنا معها عندما تعرفنا على "المدى" :

للمجموعة (1) من القيم :

15	3	7	6	12	18	5	3	13	15

وهي مجموعة من القيم وسطها الحسابي :

$$\bar{x} = \frac{15+13+3+5+18+12+6+7+3+15}{10} = 9.7$$

بطرح هذا الوسط من كل قيمة من القيم السابقة نحصل على الانحرافات عن هذا الوسط :

5.3	-6.7	-2.7	-3.7	2.3	8.3	-4.7	-6.7	3.3	5.3

لاحظ أن مجموع الانحرافات = صفر

وبالتالي تكون القيم المطلقة لهذه الانحرافات هي :

5.3	6.7	2.7	3.7	2.3	8.3	4.7	6.7	3.3	5.3

إذن الانحراف المتوسط هو متوسط القيم المطلقة للانحرافات ، أي :

$$M.D = \frac{5.3+3.3+6.7+4.7+8.3+2.3+3.7+2.7+6.7+5.3}{10} = \underline{\underline{4.9}}$$

وللمجموعة (2) من القيم :

16	3	14	15	17	18	17	13	14	16

وهي مجموعة من القيم وسطها الحسابي :

$$\bar{x} = \frac{16+14+13+17+18+17+15+14+3+16}{10} 14.3$$

وبطرح هذا الوسط من كل قيمة من القيم السابقة نحصل على الانحرافات عن هذا الوسط :

1.7	**-11.3**	**-0.3**	**0.7**	**2.7**	**3.7**	**2.7**	**-1.3**	**-0.3**	**1.7**

لاحظ أيضاً أن مجموع الانحرافات = صفر

وبالتالي تكون القيم المطلقة لهذه الانحرافات هي :

1.7	11.3	0.3	0.7	2.7	3.7	2.7	1.3	0.3	1.7

إذن الانحراف المتوسط هو متوسط القيم المطلقة للانحرافات ، أي :

$$M.D = \frac{1.7+0.3+1.3+2.7+3.7+2.7+0.7+0.3+11.3+1.7}{10} = \underline{\underline{2.64}}$$

أي أن الانحراف المتوسط للمجموعة (2) من القيم أقل من الانحراف المتوسط للمجموعة (1) من القيم مما يعني أن المجموعة (2) أقل تشتتاً من المجموعة (1) وهذا ما لم يمكن ملاحظته عند استخدام المدى كمقياس للتشتت .

ويمكن أن يتم حل السؤال السابق وذلك بتنظيم خطواتنا من خلال جدول كالتالي :

المجموعة (1) [$n = 10$]

$\|d\|$	$d = x - \bar{x}$	$\bar{x}$	x
5.3	15 – 9.7 = 5.3	9.7	15
3.3	13 – 9.7 = 3.3	9.7	13
6.7	3 – 9.7 = -6.7	9.7	3
4.7	5 – 9.7 = -4.7	9.7	5
8.3	18 – 9.7 = 8.3	9.7	18
2.3	12 – 9.7 = 2.3	9.7	12
3.7	6 – 9.7 = -3.7	9.7	6
2.7	7 – 9.7 = -2.7	9.7	7
6.7	3 – 9.7 = -6.7	9.7	3
5.3	15 – 9.7 = 5.3	9.7	15
49	**0**	**97**	**97**

$\sum x = n\bar{x}$ $\quad$ $\sum d$ $\quad$ $\sum |d|$

ويمكن الاستغناء عن هذا العمود

$$\bar{x} = \frac{\sum x}{n} = \frac{97}{10} = 9.7$$

$$M.D = \frac{\sum |d|}{n} = \frac{49}{10} = 4.9$$

ويمكن إتباع نفس الأسلوب للمجموعة (2) من القيم .

وفي حالة **التوزيعات التكرارية لبيانات كمية متقطعة** ، يمكن حساب الانحراف المتوسط $M.D$ من العلاقة :

$$M.D = \frac{\sum_{i=1}^{n} f_i \,|\, d_i \,|}{\sum_{i=1}^{n} f_i} = \frac{\sum_{i=1}^{n} f_i \,|\, x_i - \bar{x} \,|}{\sum_{i=1}^{n} f_i}$$

حيث f_i هو تكرار القيمة x_i ، $d_i = x_i - \bar{x}$ هو انحراف القيمة x_i عن الوسط الحسابي $\bar{x}$ لمجموعة القيم ، $|\, d_i \,|$ هي القيمة المطلقة للانحراف d_i ، أما $\sum_{i=1}^{n} f_i$ فهو مجموع التكرارات ، n هو عدد القيم .

وللتبسيط أيضاً سنكتب تلك العلاقة على الصورة :

$$\boxed{M.D=\frac{\sum f\,|\,d\,|}{\sum f}} \qquad (4\text{-}2)$$

حيث يكون مفهوماً أن $\sum f\,|\,d\,|$ تعني $\sum_{i=1}^{n} f_i\,|\,d_i\,|$ ، وأن $\sum f$ تعني $\sum_{i=1}^{n} f_i$. لاحظ أن العلاقة (4-2) للتوزيعات التكرارية هي نفسها العلاقة (4-1) لكن بعد استبدال n بـ $\sum f$ واستبدال $|\,d\,|$ بـ $f\,|\,d\,|$.

ونفس العلاقة السابقة يمكن استخدامها أيضاً في حالة **التوزيعات التكرارية لبيانات كمية متصلة** مع الفارق الوحيد [عن التوزيعات التكرارية لبيانات كمية متقطعة] أن الانحراف d هنا يُحسب لكل فئة من الفئات كالآتي :

$$d=x_0-\bar{x}$$

حيث x_0 هو مركز تلك الفئة .

مثال (1−4) : احسب الانحراف المتوسط للبيانات المبينة بالجدول التكراري المعطى [بيانات مثال (4−3)] :

7	6	5	4	x
	10	30	40	20

الحل :

بتكوين الجدول التكراري للأرقام المذكورة ، ثم بضرب كل قيمة في تكرارها والتجميع [عمود fx] يكون الوسط الحسابي للأرقام المذكورة هو [أنظر مثال (4−3)] :

$$\bar{x}=\frac{\sum fx}{\sum f}=\frac{530}{100}=5.3$$

	الجدول التكراري	
fx	التكرار f	المتغير x
80	20	4
200	40	5
180	30	6
70	10	7
530	100	
$\sum fx$	$\sum f$	

وبعد ذلك نقوم استكمال الجدول المبين لحساب انحراف كل قيمة من تلك القيم عن هذا الوسط الحسابي ، ثم القيمة المطلقة لتلك الانحرافات ، ثم متوسط هذه القيم المطلقة فنحصل على الانحراف المتوسط المطلوب ، وهذا ما يعطيه الجدول المتكامل التالي . لاحظ في هذا الجدول أن مجموع الانحرافات عن الوسط الحسابي أيضاً يجب أن يساوي صفراً ، لكن **انتبه** : فهذا المجموع (مجموع الانحرافات عن الوسط الحسابي) لا يمثله المجموع $\sum d$ ولكن يمثله هنا $\sum f\times d$ نظراً لوجود تكرارات للقيم . أي أن $\sum d$ **لا يُشترط أن يساوي صفراً** لكن $\sum f\times d$ **يجب أن يساوي صفراً** .

الجدول التكراري

المتغير x	التكرار f	fx
4	20	80
5	40	200
6	30	180
7	10	70
	100	530

$\sum f = 100$ $\quad \sum fx = 530$

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{530}{100} = 5.3$$

خاص بحساب الوسط الحسابي

$f\times\|d\|$	$\|d\|$	$d = x - \bar{x}$
20 × 1.3 = 26	1.3	4 − 5.3 = -1.3
40 × 0.3 = 12	0.3	5 − 5.3 = -0.3
30 × 0.7 = 21	0.7	6 − 5.3 = 0.7
10 × 1.7 = 17	1.7	7 − 5.3 = 1.7
$\sum f\|d\| = 76$		

انتبه : مجموع الانحرافات هنا [والذي يجب أن يساوي صفراً] هو $\sum f \times d$ وليس $\sum d$

وهذا الجزء يُضاف إذا كان مطلوباً حساب الانحراف المتوسط

$$M.D = \frac{\sum f\,|d\,|}{\sum f} = \frac{76}{100} = \underline{\underline{0.76}}$$

مثال (2−4) : احسب الانحراف المتوسط للبيانات المعطاة بمثال (2−6) − الباب الثاني .

الحل :

الفئة	المتغير x (الطول)	التكرار f	مركز الفئة x_0	fx_0
الأولى	$0 \le x < 20$	4	10	40
الثانية	$20 \le x < 30$	16	25	400
الثالثة	$30 \le x < 35$	12	32.5	390
الرابعة	$35 \le x < 40$	10	37.5	375
الخامسة	$40 \le x < 50$	6	45	270
السادسة	$50 \le x < 60$	2	55	110
		50		1585
		$\sum f$		$\sum fx_0$

سبق وحسبنا الوسط الحسابي لهذه البيانات [مثال (3−5) − الباب الثالث] ووجدنا أن هذا الوسط يُعطى بـ :

$$\bar{x} = \frac{\sum fx_0}{\sum f} = \frac{1585}{50} = 31.7$$

ولحساب الانحراف المتوسط لهذه البيانات يمكن استكمال هذا الجدول بأن نكون أعمدة تبين انحرافات كل فئة عن هذا الوسط الحسابي [العمود المسمى بعمود d حيث $d = x_0 - \bar{x}$] ، ثم تحديد القيم المطلقة لهذه الانحرافات [وهو العمود المسمى بعمود $|\,d\,|$] ، ثم بعمود يختص بحاصل ضرب تكرار كل فئة في القيمة المطلقة لانحراف الفئة المناظرة [العمود المسمى بعمود $f\,|\,d\,|$] ، وأخيراً حساب الانحراف المتوسط للبيانات من العلاقة :

$$M.D = \frac{\sum f\,|\,d\,|}{\sum f}$$

المتغير x (الطول)	التكرار f	المركز x_0	fx_0	$d = x_0 - \bar{x}$	$\|d\|$	$f \times \|d\|$
$0 \le x < 20$	4	10	40	10 – 31.7 = -21.7	21.7	86.8
$20 \le x < 30$	16	25	400	25 – 31.7 = -6.7	6.7	107.2
$30 \le x < 35$	12	32.5	390	32.5 – 31.7 = 0.8	0.8	9.6
$35 \le x < 40$	10	37.5	375	37.5 – 31.7 = 5.8	5.8	58
$40 \le x < 50$	6	45	270	45 – 31.7 = 13.3	13.3	79.8
$50 \le x < 60$	2	55	110	55 – 31.7 = 23.3	23.3	46.6
	50		1585			388
	$\sum f$		$\sum fx_0$			$\sum f\|d\|$

لحساب الوسط الحسابي

$$M.D = \frac{\sum f\,|\,d\,|}{\sum f} = \frac{388}{50} = 7.76$$

مثال (3–4) : احسب الانحراف المتوسط للبيانات المعطاة بمثال (8–2) – الباب الثاني .

الحل :

بنفس أسلوب حل المثال السابق [مثال (2–4)] يمكن التعامل مع هذا المثال حيث نقوم أولاً بإيجاد الوسط الحسابي [وهذا ما فعلناه بالفعل من خلال مثال (6–3) – الباب الثالث] ، ثم نكون أعمدة توضح الكميات $d, |d|, f|d|$ ومنها نوجد الانحراف المتوسط للبيانات من العلاقة $M.D = \frac{\sum f\,|\,d\,|}{\sum f}$. ويمكن تجميع كل الحسابات السابقة في الجدول التالي :

المتغير x (الأجر)	التكرار f	المركز x_0	fx_0	$d = x_0 - \bar{x}$	$\|d\|$	$f \times \|d\|$
$50 \le x < 60$	6	55	330	55 – 83.75 = -28.75	28.75	172.5
$60 \le x < 70$	9	65	585	65 – 83.75 = -18.75	18.75	168.75
$70 \le x < 80$	15	75	1125	75 – 83.75 = -8.75	8.75	131.25
$80 \le x < 90$	12	85	1020	85 – 83.75 = 1.25	1.25	15
$90 \le x < 100$	9	95	855	95 – 83.75 = 11.25	11.25	101.25
$100 \le x < 120$	6	110	660	110 – 83.75 = 26.25	26.25	157.5
$120 \le x < 180$	3	150	450	150 – 83.75 = 66.25	66.25	198.75
	60		5025			945
	$\sum f$		$\sum fx_0$			$\sum f\|d\|$

لحساب الوسط الحسابي

$$\bar{x} = \frac{\sum fx_0}{\sum f} = \frac{5025}{60} = 83.75$$

$$M.D = \frac{\sum f\,|\,d\,|}{\sum f} = \frac{945}{60} = 15.75$$

ملحوظة :

من تعريف الانحراف المتوسط يتضح لنا أن الانحراف المتوسط يعتمدا تماماً في حسابه على الوسط الحسابي ، وبالتالي يكون له نفس مزايا وعيوب الوسط الحسابي . أي :

المزايا : من السهل حسابه – يأخذا في الاعتبار جميع البيانات – لا يحتاج لترتيب معين للبيانات

العيوب : يتأثر بشدة بالقيم المتطرفة – لا يمكن إيجاده بالرسم (بيانياً) – لا يمكن حسابه للتوزيعات التكرارية المفتوحة

(4–4) الانحراف المعياري s

يُعرف **متوسط مربعات الانحرافات عن الوسط الحسابي** على أنه **تباين** مجموعة البيانات [ويُرمز له بالرمز s^2] ، ويُعرف **الجذر التربيعي للتباين** على أنه **الانحراف المعياري** للبيانات [ويُرمز له بالرمز s] ، أي أنه لمجموعة من القيم $x_1 , x_2 , \cdots\cdots , x_n$ عددها n يكون :

التباين $$= s^2 = \frac{\sum_{i=1}^{n} d_i^2}{n}$$

ومنه يكون :

الانحراف المعياري $$= s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n}}$$

حيث $d_i = x_i - \bar{x}$ هي انحراف القيمة x_i عن الوسط الحسابي $\bar{x}$ لمجموعة القيم . وللتبسيط سنكتب العلاقات السابقة على الصور :

التباين $$= s^2 = \frac{\sum d^2}{n} \qquad (4\text{-}3)$$

الانحراف المعياري $$= s = \sqrt{s^2} = \sqrt{\frac{\sum d^2}{n}} \qquad (4\text{-}4)$$

حيث يكون مفهوماً أن $\sum d^2$ تعني $\sum_{i=1}^{n} {d_i}^2$ [أي مجموع مربعات انحرافات القيم عن وسطها الحسابي $\bar{x}$] .

وفي حالة التوزيعات التكرارية **لبيانات كمية متقطعة** ، تأخذ العلاقات (4-4) , (4-3) السابقة الصور :

التباين $$= s^2 = \frac{\sum fd^2}{\sum f} \qquad (4\text{-}5)$$

الانحراف المعياري $$= s = \sqrt{s^2} = \sqrt{\frac{\sum fd^2}{\sum f}} \qquad (4\text{-}6)$$

حيث $\sum fd^2$ تعني $\sum_{i=1}^{n} f_i d_i^2$ ، $\sum f$ تعني $\sum_{i=1}^{n} f_i$. لاحظ أن العلاقات (4-6) , (4-5) للتوزيعات التكرارية هي نفسها العلاقات (4-4) , (4-3) لكن بعد استبدال n بـ $\sum f$ واستبدال d^2 بـ fd^2 .

ونفس العلاقات (4-6) , (4-5) السابقة يمكن استخدامها أيضاً في حالة **التوزيعات التكرارية لبيانات كمية متصلة** مع الفارق الوحيد [عن التوزيعات التكرارية لبيانات كمية متقطعة] أن الانحراف d يُحسب لكل فئة من الفئات كالآتي :

$$d = x_0 - \bar{x}$$

حيث x_0 هو مركز تلك الفئة .

مثال (4–4) : في كل حالة من الحالتين التاليتين **احسب** الانحراف المعياري للبيانات المعطاة .

مجموعة (1) 15 3 7 6 12 18 5 3 13 15

مجموعة (2) 16 3 14 15 17 18 17 13 14 16

الحل : لاحظ أن البيانات المعطاة هي البيانات الخاصة بالمجموعتين (1) ، (2) التي تناولناها عند التعرف على المدى R في البند (4–1) من هذا الباب

في كل حالة من الحالتين يجب أولاً حساب الوسط الحسابي للبيانات ، ثم حساب انحرافات البيانات عن هذا الوسط ، ثم مربعات هذه الانحرافات ، ثم متوسط هذه المربعات [أي استخدام العلاقة (4-3)] ، فنكون بالتالي حصلنا على تباين البيانات ، ويكون الانحراف المعياري هو الجذر التربيعي لهذا التباين . هذه الخطوات تلخصها الجداول التالية :

المجموعة (1) [n = 10]

d^2	$d = x - \bar{x}$	x
28.09	15 – 9.7 = 5.3	15
10.89	13 – 9.7 = 3.3	13
44.89	3 – 9.7 = -6.7	3
22.09	5 – 9.7 = -4.7	5
68.89	18 – 9.7 = 8.3	18
5.29	12 – 9.7 = 2.3	12
13.69	6 – 9.7 = -3.7	6
7.29	7 – 9.7 = -2.7	7
44.89	3 – 9.7 = -6.7	3
28.09	15 – 9.7 = 5.3	15
274.1	$\sum d = 0$	97
$\sum d^2$		$\sum x$

المجموعة (2) [n = 10]

d^2	$d = x - \bar{x}$	x
2.89	16 – 14.3 = 1.7	16
0.09	14 – 14.3 = -0.3	14
1.69	13 – 14.3 = -1.3	13
7.29	17 – 14.3 = 2.7	17
13.69	18 – 14.3 = 3.7	18
7.29	17 – 14.3 = 2.7	17
0.49	15 – 14.3 = 0.7	15
0.09	14 – 14.3 = -0.3	14
127.69	3 – 14.3 = -11.3	3
2.89	16 – 14.3 = 1.7	16
164.1	$\sum d = 0$	143
$\sum d^2$		$\sum x$

❖ **بالنسبة للمجموعة (1) :**

$$\bar{x} = \frac{\sum x}{n} = \frac{97}{10} = 9.7$$

وبالتالي يمكن حساب انحرافات القيم عن هذا الوسط الحسابي ، ومنها يمكن معرفة مربعات هذه الانحرافات ، وباستخدام العلاقة (4-3) يكون التباين هو :

$$s^2 = \frac{\sum d^2}{n} = \frac{274.1}{10} = 27.41$$

ومن التباين يمكن حساب الانحراف المعياري s [العلاقة (4-4)] ليكون :

$$s = \sqrt{s^2} = \sqrt{27.41} \cong 5.24$$

❖ **وبالنسبة للمجموعة (2) :** وبنفس الأسلوب تماماً يمكن التعامل مع بيانات المجموعة (2) لنحصل على الآتي :

$$\bar{x} = \frac{\sum x}{n} = \frac{143}{10} = 14.3$$

$$s^2 = \frac{\sum d^2}{n} = \frac{164.1}{10} = 16.41 \quad \rightarrow \quad s = \sqrt{s^2} = \sqrt{16.41} \cong 4.05$$

مثال (4-5) : احسب الانحراف المعياري للبيانات المعطاة بمثال (4-1) .

الحل :

سبق وحسبنا كلاً من الوسط الحسابي لهذه البيانات ووجدنا أن هذا الوسط يُعطى بـ :

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{530}{100} = 5.3$$

وأن الانحراف المتوسط يُعطى بـ :

$$M.D = \frac{\sum f\,|\,d\,|}{\sum f} = \frac{76}{100} = 0.76$$

ولحساب الانحراف المعياري لهذه البيانات نتبع تقريباً نفس الخطوات التي اتبعناها عند حسابنا للانحراف المتوسط مع الاهتمام بحساب مربعات انحرافات القيم عن الوسط الحسابي بدلاً من القيم المطلقة لهذه الانحرافات ، ثم نقوم بتطبيق العلاقات (4-6) , (4-5) لتحديد كلٍ من التباين والانحراف المعياري فنحصل على [أنظر إلى الجدول التالي] :

$$s^2 = \frac{\sum fd^2}{\sum f} = \frac{81}{100} = 0.81$$

$$\Rightarrow \quad s = \sqrt{s^2} = \sqrt{0.81} = 0.9$$

الجدول التكراري					
المتغير x	التكرار f	fx	$d = x - \bar{x}$	d^2	fd^2
4	20	80	4 – 5.3 = -1.3	1.69	20 × 1.69 = 33.8
5	40	200	5 – 5.3 = -0.3	0.09	40 × 0.09 = 3.6
6	30	180	6 – 5.3 = 0.7	0.49	30 × 0.49 = 14.7
7	10	70	7 – 5.3 = 1.7	2.89	10 × 2.89 = 28.9
	100	530			81

$\sum f = 100$ $\sum fx = 530$ $\sum fd^2$

مثال (6–4) : احسب الانحراف المعياري للبيانات المعطاة بمثال (2–4) .

الحل :

سبق وحسبنا كلاً من الوسط الحسابي لهذه البيانات والانحراف المتوسط لهذه البيانات ووجدنا أن هذا الوسط يُعطى بـ :

$$\bar{x} = \frac{\sum fx_0}{\sum f} = \frac{1585}{50} = 31.7$$

وأن الانحراف المتوسط يُعطى بـ :

$$M.D = \frac{\sum f\,|\,d\,|}{\sum f} = \frac{388}{50} = 7.76$$

والآن لحساب الانحراف المعياري لهذه البيانات نتبع تقريباً نفس الخطوات التي اتبعناها عند حسابنا للانحراف المتوسط مع الاهتمام بحساب مربعات انحرافات القيم عن الوسط الحسابي بدلاً من القيم المطلقة لهذه الانحرافات ، ثم نقوم بتطبيق العلاقات (4-6) , (4-5) لتحديد كلٍ من التباين والانحراف المعياري كالآتي :

$$s^2 = \frac{\sum fd^2}{\sum f} = \frac{5093}{50} = 101.86 \qquad \Rightarrow \qquad s = \sqrt{s^2} = \sqrt{101.86} \cong 10.09$$

المتغير x (الطول)	التكرار f	المركز x_0	fx_0	$d = x_0 - \bar{x}$	d^2	$f \times d^2$
$0 \le x < 20$	4	10	40	10 – 31.7 = -21.7	470.89	4 × 470.89 = 1883.56
$20 \le x < 30$	16	25	400	25 – 31.7 = -6.7	44.89	16 × 44.89 = 718.24
$30 \le x < 35$	12	32.5	390	32.5 – 31.7 = 0.8	0.64	12 × 0.64 = 7.68
$35 \le x < 40$	10	37.5	375	37.5 – 31.7 = 5.8	33.64	10 × 33.64 = 336.4
$40 \le x < 50$	6	45	270	45 – 31.7 = 13.3	176.89	6 × 176.89 = 1061.34
$50 \le x < 60$	2	55	110	55 – 31.7 = 23.3	542.89	2 × 542.89 = 1085.78
	50		1585			5093

$\sum f$ $\sum fx_0$ $\sum fd^2$

مثال (4–7) : احسب الانحراف المعياري للبيانات المعطاة بمثال (4–3) .

الحل :

سبق وحسبنا كلاً من الوسط الحسابي لهذه البيانات والانحراف المتوسط لهذه البيانات ووجدنا أن هذا الوسط يُعطى بـ :

$$\bar{x} = \frac{\sum f x_0}{\sum f} = \frac{5025}{60} = 83.75$$

وأن الانحراف المتوسط يُعطى بـ :

$$M.D = \frac{\sum f \,|\, d \,|}{\sum f} = \frac{945}{60} = 15.75$$

وبنفس أسلوب المثال السابق يمكنا حساب كلٍ من التباين والانحراف المعياري فنحصل على :

$$s^2 = \frac{\sum f d^2}{\sum f} = \frac{27731.12}{60} \cong 462.19 \qquad \Rightarrow \qquad s = \sqrt{s^2} = \sqrt{462.19} \cong 21.5$$

$f \times d^2$	d^2	$d = x_0 - \bar{x}$	fx_0	المركز x_0	f	المتغير x (الطول)
4959.38	798.06	55 – 83.75 = -28.75	330	55	6	$50 \le x < 60$
3164.04	351.56	65 – 83.75 = -18.75	585	65	9	$60 \le x < 70$
1148.4	76.56	75 – 83.75 = -8.75	1125	75	15	$70 \le x < 80$
18.72	1.56	85 – 83.75 = 1.25	1020	85	12	$80 \le x < 90$
1139.04	126.56	95 – 83.75 = 11.25	855	95	9	$90 \le x < 100$
4134.36	689.06	110 – 83.75 = 26.25	660	110	6	$100 \le x < 120$
13167.18	4389.06	150 – 83.75 = 66.25	450	150	3	$120 \le x < 180$
27731.12			5025		60	
$\sum fd^2$			$\sum fx_0$		$\sum f$	

ملحوظة :

كما في حالة الانحراف المتوسط يتضح لنا من تعريف الانحراف المعياري لبيانات ما أن الانحراف المعياري يعتمدا تماماً في حسابه على الوسط الحسابي ، وبالتالي يكون له نفس مزايا وعيوب الوسط الحسابي . أي :

المزايا : من السهل حسابه – يأخذ في الاعتبار جميع البيانات – لا يحتاج لترتيب معين للبيانات

العيوب : يتأثر بشدة بالقيم المتطرفة – لا يمكن إيجاده بالرسم (بيانياً) – لا يمكن حسابه للتوزيعات التكرارية المفتوحة

(5–4) خاصتان هامتان للانحراف المتوسط والانحراف المعياري

الخاصية الأولى :

إضافة عدد ثابت c لكل قيمة من قيم البيانات لا يؤثر على قيمة الانحرافين المتوسط والمعياري .

الانحراف المتوسط (أو المعياري) الجديد = الانحراف المتوسط (أو المعياري) القديم

الخاصية الثانية : ضرب كل قيمة من قيم البيانات في عدد ثابت c يجعل :

الانحراف المتوسط (أو المعياري) الجديد = الانحراف المتوسط (أو المعياري) القديم × القيمة المطلقة للثابت c

مثال (4–8) : درجات 5 طلاب بأحد الفصول في مقرر اللغة العربية كانت كالتالي :

9 , 2 , 7 , 12 , 10

(أ) **احسب** الوسط الحسابي والانحراف المتوسط والانحراف المعياري لدرجات الطلاب .

(ب) وإذا أضفنا 5 درجات لكل طالب وذلك لتحسين درجاتهم ، **ما قيمة** الوسط الحسابي والانحراف المتوسط والانحراف المعياري للدرجات الجديدة ؟.

(ج) وإذا أردنا تحسين درجات الطلاب بزيادة كل درجة 50% من قيمتها الأصلية ، **ما قيمة** الوسط الحسابي والانحراف المتوسط والانحراف المعياري للدرجات الجديدة ؟.

الحل :

(أ) للدرجات الأصلية :

$$\bar{x} = \frac{\sum x}{n} = \frac{40}{5} = 8 =$$ الوسط الحسابي

$$M.D = \frac{\sum |d|}{n} = \frac{14}{5} = 2.8 =$$ الانحراف المتوسط

$$s^2 = \frac{\sum d^2}{n} = \frac{58}{5} = 11.6 =$$ التباين

$$s = \sqrt{s^2} = \sqrt{11.6} \cong 3.4 =$$ الانحراف المعياري

الدرجات الأصلية			
d^2	$\|d\|$	d	x
1	1	1	9
36	6	-6	2
1	1	-1	7
16	4	4	12
4	2	2	10
58	14		40
$\sum d^2$	$\sum \|d\|$		$\sum x$

(ب) عند زيادة كل درجة 5 درجات :

الوسط الحسابي = $\bar{x} = \frac{\sum x}{n} = \frac{65}{5} = 13$

الانحراف المتوسط = $M.D = \frac{\sum |d|}{n} = \frac{14}{5} = 2.8$

التباين = $s^2 = \frac{\sum d^2}{n} = \frac{58}{5} = 11.6$

الانحراف المعياري = $s = \sqrt{s^2} = \sqrt{11.6} \cong 3.4$

أي أن :

الوسط الجديد = الوسط القديم + **5**

الانحراف المتوسط الجديد = الانحراف المتوسط القديم

الانحراف المعياري الجديد = الانحراف المعياري القديم

الدرجات الجديدة			
d^2	$\|d\|$	d	x
1	1	1	14
36	6	-6	7
1	1	-1	12
16	4	4	17
4	2	2	15
58	14		65
$\sum d^2$	$\sum \|d\|$		$\sum x$

(ج) عند زيادة كل درجة 50% من قيمتها الأصلية :

[أي أن كل درجة تُضرب في 1.5]

الوسط الحسابي = $\bar{x} = \frac{\sum x}{n} = \frac{60}{5} = 12$

الانحراف المتوسط = $M.D = \frac{\sum |d|}{n} = \frac{21}{5} = 4.2$

التباين = $s^2 = \frac{\sum d^2}{n} = \frac{130.5}{5} = 26.1$

الانحراف المعياري = $s = \sqrt{s^2} = \sqrt{26.1} \cong 5.1$

أي أن :

الوسط الجديد = الوسط القديم × **1.5**

الانحراف المتوسط الجديد = الانحراف المتوسط القديم × **1.5**

الانحراف المعياري الجديد = الانحراف المعياري القديم × **1.5**

الدرجات الجديدة			
d^2	$\|d\|$	d	x
2.25	1.5	1.5	13.5
81	9	-9	3
2.25	1.5	-1.5	10.5
36	6	6	18
9	3	3	15
130.5	21		60
$\sum d^2$	$\sum \|d\|$		$\sum x$

ملحوظة هامة :

علاقة التباين (3-4) للبيانات غير المبوبة يمكن كتابتهما على الصورة :

$$s^2 = \frac{\sum x^2}{n} - \bar{x}^2 \qquad (4\text{-}7)$$

وللبيانات المبوبة فإن العلاقة (4-5) للتباين يمكن كتابتها على الصورة :

$$s^2 = \frac{\sum fx^2}{\sum f} - \bar{x}^2 \qquad (4\text{-}8)$$

ويمكن إثبات ذلك بسهولة كالتالي :

للبيانات غير المبوبة ، ومن المعادلة (4-3) يمكن كتابة :

$$s^2 = \frac{\sum d^2}{n} = \frac{\sum (x-\bar{x})^2}{n} = \frac{\sum (x^2 - 2x\bar{x} + \bar{x}^2)}{n}$$

$$= \frac{\sum x^2}{n} - \frac{\sum 2x\bar{x}}{n} + \frac{\sum \bar{x}^2}{n} = \frac{\sum x^2}{n} - \frac{2\bar{x}\sum x}{n} + \frac{\bar{x}^2 \sum 1}{n}$$

وحيث أن :

$$\frac{\sum x}{n} = \bar{x} \quad , \quad \sum 1 = n$$

فإن :

$$s^2 = \frac{\sum x^2}{n} - 2\bar{x}(\bar{x}) + \bar{x}^2(1) = \frac{\sum x^2}{n} - \bar{x}^2$$

وهي العلاقة (4-7) ، وبالمثل ، في حالة البيانات المبوبة ومن المعادلة (4-5) يمكن كتابة :

$$s^2 = \frac{\sum fd^2}{\sum f} = \frac{\sum f(x-\bar{x})^2}{\sum f} = \frac{\sum f(x^2 - 2x\bar{x} + \bar{x}^2)}{\sum f}$$

$$= \frac{\sum fx^2}{\sum f} - \frac{\sum f(2x\bar{x})}{\sum f} + \frac{\sum f\bar{x}^2}{\sum f} = \frac{\sum fx^2}{\sum f} - \frac{2\bar{x}\sum fx}{\sum f} + \frac{\bar{x}^2 \sum f}{\sum f}$$

وحيث أن $\frac{\sum fx}{\sum f} = \bar{x}$ ، إذن :

$$s^2 = \frac{\sum fx^2}{\sum f} - 2\bar{x}(\bar{x}) + \bar{x}^2(1) = \frac{\sum fx^2}{\sum f} - \bar{x}^2$$

وهي العلاقة (4-8) .

وتذكر الآتي :

- في حالة البيانات الكمية المتصلة المبوبة نستبدل x [في العلاقة (4-8)] بـ x_0 .
- الانحراف المعياري s [في جميع الحالات] هو **الجذر التربيعي للتباين** .

مثال (4–9) :

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام 2010 [بالألف ريال] :

Month	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Sales	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب حساب الانحراف المعياري للمبيعات الشهرية .

الحل :

المبيعات x	x^2
3	9
5	25
8	64
3	9
6	36
4	16
12	144
5	25
4	16
3	9
7	49
9	81
$\sum x = 69$	$\sum x^2 = 483$

$$n = 12$$

$$\sum x = 69$$

$$\overline{x} = \frac{\sum x}{n} = \frac{69}{12} = 5.75$$

$$s^2 = \frac{\sum x^2}{n} - \overline{x}^2$$

$$= \frac{483}{12} - (5.75)^2 = 7.1875$$

$$s = \sqrt{s^2} = \sqrt{7.1875} \cong 2.68$$

مثال (4–10) : في دراسة أُجريت على أطوال 20 طالباً كانت النتائج كالتالي :

$$\sum x = 3034 \qquad , \qquad \sum x^2 = 463040$$

حيث x هو الطول (بالسنتيمتر) ، **احسب** الوسط الحسابي والانحراف المعياري للأطوال .

الحل :

الوسط الحسابي للأطوال :

$$n = 20 \quad , \quad \sum x = 3043 \quad \Rightarrow \quad \overline{x} = \frac{\sum x}{n} = \frac{3034}{20} = 151.7$$

التباين للأطوال :

$$s^2 = \frac{\sum x^2}{n} - \overline{x}^2 = \frac{463040}{20} - (151.7)^2 = 139.11$$

الانحراف المعياري للأطوال :

$$s = \sqrt{s^2} = \sqrt{139.11} \cong 11.8$$

مثال (4–11) :

على نفس الـ 20 طالباً في المثال السابق أجريت دراسة على أوزانهم [وليس أطوالهم] كانت النتائج كالتالي :

الوزن x	50 -	60 -	70 -	80 - 90
عدد الطلاب	3	5	8	4

احسب الوسط الحسابي والانحراف المعياري للأوزان .

الحل :

بتكوين الجدول المبين واستخدام العلاقة (3-4) [لتحديد الوسط الحسابي] والعلاقة (4-8) [لتحديد التباين ومن ثم الانحراف المعياري] ، يمكن الحصول على :

فئات الأوزان x	التكرار f	مراكز الفئات x_0	fx_0	fx_0^2
50 -	3	55	165	9075
60 -	5	65	325	21125
70 -	8	75	600	45000
80 - 90	4	85	340	28900
	$\sum f == 20$		$\sum fx_0 = 1430$	$\sum fx_0^2 = 104100$

$$\bar{x} = \frac{\sum fx_0}{\sum f} = \frac{1430}{20} = 71.5$$

$$s^2 = \frac{\sum fx_0^2}{\sum f} - \bar{x}^2 = \frac{104100}{20} - (71.5)^2 = 92.75$$

$$s = \sqrt{s^2} = \sqrt{92.75} \cong 9.63$$

ملخص للدرس السابع [الباب الرابع : مقاييس التشتت]

المدى – الانحراف المتوسط – الانحراف المعياري

- **التشتت** هو الدرجة التي تتجه بها البيانات الكمية للانتشار حول قيمة متوسطة (أحد مقاييس النزعة المركزية) ، **ومقاييس التشتت** هي مقاييس يمكن استخدامها لقياس هذا التشتت .

- ولبيانات ما ، يُعرف :
  - ❖ **المدى** على أنه الفرق بين أكبر قيمة في البيانات وأصغر قيمة فيها .
  - ❖ **الانحراف المتوسط** (متوسط الانحرافات) على أنه متوسط القيم المطلقة للانحرافات عن الوسط الحسابي للبيانات
  - ❖ **التباين** على أنه متوسط مربعات الانحرافات عن الوسط الحسابي للبيانات .
  - ❖ **الانحراف المعياري** على أنه الجذر التربيعي للتباين .

- والانحراف المتوسط والتباين والانحراف المعياري مقاييس للتشتت ترتبط بالوسط الحسابي (أحد مقاييس النزعة المركزية) ، ويمكن حسابها للقيم المفردة وللتوزيعات التكرارية .

- **للقيم المفردة :**

مربع الانحرافات	القيم المطلقة للانحرافات	الانحرافات عن	قيم عددها **n**
d^2	$\|d\|$	$d = x - \bar{x}$	x
…	…	…	…
…	…	…	…
$\sum d^2$	$\sum \|d\|$		$\sum x$

$$\bar{x} = \frac{\sum x}{n} = \text{الوسط الحسابي}$$

$$M.D = \frac{\sum |d|}{n} = \text{الانحراف المتوسط}$$

$$s^2 = \frac{\sum d^2}{n} = \frac{\sum x^2}{n} - \bar{x}^2 = \text{التباين}$$

$$s = \sqrt{s^2} = \text{الانحراف المعياري}$$

- **للتوزيعات التكرارية لبيانات كمية متقطعة :**

x	f	fx	$d = x - \bar{x}$	$\lvert d\rvert$	d^2	$f\lvert d\rvert$	fd^2
	$\sum f$	$\sum fx$				$\sum f\lvert d\rvert$	$\sum fd^2$

الوسط الحسابي = $\bar{x} = \dfrac{\sum fx}{\sum f}$

الانحراف المتوسط = $M.D = \dfrac{\sum f\,|d|}{\sum f}$

التباين = $s^2 = \dfrac{\sum fd^2}{\sum f} = \dfrac{\sum fx^2}{\sum f} - \bar{x}^2$

الانحراف المعياري = $s = \sqrt{s^2}$

- **للتوزيعات التكرارية لبيانات كمية متصلة :**

x	x_0 مراكز الفئات	f	fx_0	$d = x_0 - \bar{x}$	$\lvert d\rvert$	d^2	$f\lvert d\rvert$	fd^2
فئات								
		$\sum f$	$\sum fx$				$\sum f\lvert d\rvert$	$\sum fd^2$

الوسط الحسابي = $\bar{x} = \dfrac{\sum fx_0}{\sum f}$

الانحراف المتوسط = $M.D = \dfrac{\sum f\,|d|}{\sum f}$

التباين = $s^2 = \dfrac{\sum fd^2}{\sum f} = \dfrac{\sum fx_0^2}{\sum f} - \bar{x}^2$

الانحراف المعياري = $s = \sqrt{s^2}$

- وإذا أضفنا العدد الثابت c لكل قيمة من قيم بيانات ما فإن :
  - ❖ الوسط الحسابي الجديد = الوسط الحسابي القديم + c
  - ❖ الانحراف المتوسط الجديد = الانحراف المتوسط القديم
  - ❖ الانحراف المعياري الجديد = الانحراف المعياري القديم

- أما إذا ضربنا كل قيمة من القيم في العدد الثابت c فإن :
  - ❖ الوسط الحسابي الجديد = الوسط الحسابي القديم × c
  - ❖ الانحراف المتوسط الجديد = الانحراف المتوسط القديم × $|c|$
  - ❖ الانحراف المعياري الجديد = الانحراف المعياري القديم × $|c|$

- للوسط الحسابي والانحراف المتوسط والانحراف المعياري والتباين نفس المزايا ونفس العيوب :

– المزايا : سهولة الحساب – جميع البيانات تُؤخذ في الاعتبار – لا تحتاج إلى ترتيب معين للبيانات عند حسابها

– العيوب : تتأثر بشدة بالقيم المتطرفة – لا يمكن إيجادها بالرسم – لا يمكن حسابها في حالة التوزيعات التكرارية المفتوحة – لا يمكن تحديدها للبيانات النوعية [أي يمكن حسابها فقط للبيانات الكمية]

تدريب عملي

في كل حالة من الحالات التالية ، **أوجد** المدى R ، الانحراف المتوسط $M.D$ ، التباين s^2 ، الانحراف المعياري s :

(أ) 7 , 4 , 10 , 9 , 15 , 12 , 7 , 9 , 7

(ب) 8 , 11 , 4 , 3 , 2 , 5 , 10 , 6 , 4 , 1 , 10 , 8 , 12 , 6 , 5 , 7

(ج) 85 , 76 , 93 , 82 , 96

(د) 0.53 , 0.46 , 0.50 , 0.49 , 0.52 , 0.53 , 0.44 , 0.55

(هـ)

x	462	480	498	516	534	552	570	588	606	624
f	98	75	56	42	30	21	15	11	6	2

(و)

x	10 –	15 –	20 –	25 –	30 –	35 –	40 – 45
f	3	7	16	12	9	5	2

تدريبات (7)

الإجابة النهائية لجميع التمرينات موجودة في نهاية التدريب

اختر الإجابة الصحيحة

(1) **مقاييس التشتت** هي

(أ) قيم نموذجية يمكن أن تمثل مجموعة البيانات

(ب) مقاييس ترصد الدرجة التي تتجه بها البيانات الكمية للانتشار حول قيمة متوسطة

(ج) مقاييس تحدد النسبة المئوية للتشتت المطلق بالنسبة لقيمة متوسطة

(د) هي مقاييس ترصد درجة تماثل أو البعد عن التماثل لتوزيع ما

(هـ) مقاييس ترصد درجة التدبب في قمة المنحنى مقارنة بقمة منحنى التوزيع الطبيعي

(2) **الانحراف المتوسط** هو أحد مقاييس

(أ) النزعة المركزية (ب) التشتت

(ج) الالتواء (د) التفرطح

ويمكن أن يُستبدل الانحراف المتوسط في رأس السؤال بالانحراف المعياري أو المدي الربيعي أو الانحراف الربيعي أو الانحراف المئيني

(3) لعدد من القيم ، يُعرف **متوسط القيم المطلقة للانحرافات عن الوسط الحسابي** على أنه

(أ) الوسط الحسابي للقيم (ب) الانحراف المتوسط للقيم

(ج) تباين تلك القيم (د) الانحراف المعياري للقيم

(4) لعدد من القيم ، يُعرف **متوسط مربعات الانحرافات عن الوسط الحسابي** على أنه

(أ) الوسط الحسابي للقيم (ب) الانحراف المتوسط للقيم

(ج) تباين تلك القيم (د) الانحراف المعياري للقيم

(5) لعدد من القيم ، يُعرف **الجذر التربيعي المتوسط مربعات الانحرافات عن الوسط الحسابي** على أنه

(أ) الوسط الحسابي للقيم (ب) الانحراف المتوسط للقيم

(ج) تباين تلك القيم (د) الانحراف المعياري للقيم

خاص بالأسئلة من (**6**) إلى (**9**) :

إذا كان $\sum x$ هو مجموع عدد قدره n من القيم ، وكان $\sum d$ هو مجموع انحرافات هذه القيم عن وسطها الحسابي ، $\sum |d|$ هو مجموع القيم المطلقة لتلك الانحرافات ، $\sum d^2$ هو مجموع مربعات تلك الانحرافات ، فإن

(6) $\dfrac{\sum x}{n}$ هو :

(أ) الوسط الحسابي للقيم   (ب) الانحراف المتوسط للقيم

(ج) تباين تلك القيم   (د) صفر

(7) $\dfrac{\sum d}{n}$ هو :

(أ) الوسط الحسابي للقيم   (ب) الانحراف المتوسط للقيم

(ج) تباين تلك القيم   (د) صفر

(8) $\dfrac{\sum |d|}{n}$ هو :

(أ) الوسط الحسابي للقيم   (ب) الانحراف المتوسط للقيم

(ج) تباين تلك القيم   (د) صفر

(9) $\dfrac{\sum d^2}{n}$ هو :

(أ) الوسط الحسابي للقيم   (ب) الانحراف المتوسط للقيم

(ج) تباين تلك القيم   (د) صفر

خاص بالأسئلة من (**10**) إلى (**13**) :

إذا كان الوسط الحسابي لمجموعة من القيم هو 20 وانحرافها المتوسط 4 وانحرافها المعياري 5 وأضفنا لكل قيمة من القيم 2 ، فإن :

(10) الوسط الحسابي للقيم الجديدة يكون :

(أ) 20   (ب) 22   (ج) 40   (د) 18

(11) الانحراف المتوسط للقيم الجديدة يكون :

(أ) 4   (ب) 6   (ج) 8   (د) 2

(12) الانحراف المعياري للقيم الجديدة يكون :

(أ) 5 (ب) 7 (ج) 10 (د) 3

(13) التباين للقيم الجديدة يكون :

(أ) $\sqrt{5}$ (ب) 25 (ج) 7 (د) 49

خاص بالأسئلة من (14) إلى (17) :

إذا كان الوسط الحسابي لمجموعة من القيم هو 20 وانحرافها المتوسط 4 وانحرافها المعياري 5 وضربنا كل قيمة من القيم في العدد 2 ، فإن :

(14) الوسط الحسابي للقيم الجديدة يكون :

(أ) 20 (ب) 22 (ج) 40 (د) 18

(15) الانحراف المتوسط للقيم الجديدة يكون :

(أ) 4 (ب) 6 (ج) 8 (د) 2

(16) الانحراف المعياري للقيم الجديدة يكون :

(أ) 3 (ب) 5 (ج) 7 (د) 10

(17) التباين للقيم الجديدة يكون :

(أ) $\sqrt{5}$ (ب) 25 (ج) 10 (د) 100

خاص بالأسئلة من (18) إلى (21) :

إذا كان الوسط الحسابي لمجموعة من القيم هو 20 وانحرافها المتوسط 4 وانحرافها المعياري 5 وضربنا كل قيمة من القيم في العدد 2- ، فإن :

(18) الوسط الحسابي للقيم الجديدة يكون :

(أ) 20 (ب) 22 (ج) 40 (د) -40

(19) الانحراف المتوسط للقيم الجديدة يكون :

(أ) 4 (ب) 6 (ج) 8 (د) -8

(20) الانحراف المعياري للقيم الجديدة يكون :

(أ) 5 (ب) 7 (ج) 10 (د) -10

21) التباين للقيم الجديدة يكون :

أ) $\sqrt{5}$ ب) 25 ج) 100 د) -100

---

22) التباين لمجموعة من القيم هو

أ) الانحراف المعياري للقيم ب) مربع الانحراف المعياري للقيم

ج) الجذر التربيعي للانحراف المعياري د) نصف الانحراف المعياري

23) الانحراف المعياري لمجموعة من القيم هو

أ) تباين هذه القيم ب) نصف التباين للقيم

ج) الجذر التربيعي لتباين هذه القيم د) مربع تباين هذه القيم

24) مقياس لا يتأثر بالقيم المتطرفة

أ) الوسط الحسابي ب) الانحراف المعياري

ج) المدى د) الوسيط

خاص بالأسئلة من (25) إلى (28) :

مجموعة من القيم عددها 10 ولها البيانات التالية :

$$\sum x = 60 \quad , \quad \sum |d| = 22 \quad , \quad \sum d^2 = 76$$

حيث $\sum x$ هو مجموع القيم ، d هو الانحراف عن الوسط الحسابي للقيم ، $|d|$ هو القيمة المطلقة لهذا الانحراف ، إذن :

25) الوسط الحسابي للبيانات السابقة هو :

أ) 2.2 ب) 7.6 ج) 6 د) 2.76

26) الانحراف المتوسط للبيانات السابقة هو :

أ) 2.2 ب) 7.6 ج) 6 د) 2.76

27) التباين للبيانات السابقة هو :

أ) 2.2 ب) 7.6 ج) 6 د) 2.76

28) الانحراف المعياري للبيانات السابقة هو :

أ) 2.2 ب) 7.6 ج) 6 د) 2.76

خاص بالأسئلة من (29) إلى (32) :

في الجدول التكراري المبين [غير مهم البيانات المرصود لها ....] ، إذا كان d يمثل الانحراف [لكل قيمة x] عن الوسط الحسابي ، فإن :

x	f	fx	d	$\|d\|$	$f\|d\|$	d^2	fd^2
......	......	......	......	......	......	......	.....
......	......	......	......	......	......	......	......
2	....	....	....	......	......	......	......
	$\sum f = 100$	$\sum fx = 450$			$\sum f\|d\| = 185$		$\sum fd^2 = 475$

(29) الوسط الحسابي للبيانات السابقة هو :

(أ) 4.5 (ب) 1.85 (ج) 2.18 (د) 4.75

(30) الانحراف المتوسط للبيانات السابقة هو :

(أ) 4.5 (ب) 1.85 (ج) 2.18 (د) 4.75

(31) التباين للبيانات السابقة هو :

(أ) 4.5 (ب) 1.85 (ج) 2.18 (د) 4.75

(32) الانحراف المعياري للبيانات السابقة هو :

(أ) 4.5 (ب) 1.85 (ج) 2.18 (د) 4.75

الإجابة :

(1) ب (2) ب (3) ب (4) ج (5) د
(6) أ (7) د (8) ب (9) ج (10) ب
(11) أ (12) أ (13) ب (14) ج (15) ج
(16) د (17) د (18) د (19) ج (20) ج
(21) ج (22) ب (23) ج (24) د (25) ج
(26) أ (27) ب (28) د (29) أ (30) ب
(31) د (32) ج

عناصر الدرس

الباب الرابع : مقاييس التشتت [تابع]

- الانحراف الربيعي [نصف المدى الربيعي]
- المدى المئيني
- علاقات اعتبارية بين مقاييس التشتت
- التشتت النسبي ومقاييسه
- الدرجات المعيارية

الباب الرابع : مقاييس التشتت [تابع]

(4–6) الانحراف الربيعي [نصف المدى الربيعي] Q

لمجموعة من البيانات يُعرف **المدى الربيعي** على أنه الفرق بين الربيع الثالث والربيع الأول ، أي أن :

$$\text{المدى الربيعي} = Q_3 - Q_1$$

حيث Q_1 هو **الربيع الأول** ، Q_3 هو **الربيع الثالث** للبيانات .

ويُعرف **الانحراف الربيعي** [أو **نصف المدى الربيعي**] وسنرمز له بالرمز **Q** على أنه **نصف المدى الربيعي** ، أي أن :

$$\text{الانحراف الربيعي [نصف المدى الربيعي] Q} = \frac{1}{2}(Q_3 - Q_1)$$

الربيع الأول

الربيع الثالث

ويفضل استخدام هذا المقياس [**الانحراف الربيعي**] في الكثير من الحالات خاصة تلك الحالات التي يستعصي فيها حساب الانحراف المتوسط أو المعياري [مثل حالة التوزيعات التكرارية المفتوحة أو حالة وجود قيم متطرفة في البيانات] ، وفي بعض الأحيان يُستخدم **المدى الربيعي** $Q_3 - Q_1$ كمقياس للتشتت بدلاً من نصف المدى الربيعي .

س : ما هي الربيعات ؟

ج : إذا رتبنا مجموعة من القيم ترتيباً تصاعدياً فإن القيمة التي تقسم المجموعة إلى مجموعتين متساويتين في العدد تُسمى بالوسيط M ، وبتعميم هذه الفكرة ، يمكن أن نقسم مجموعة القيم إلى أربعة أجزاء متساوية في العدد وذلك بثلاثة قيم [سنرمز لها بالرموز Q_1 , Q_2 , Q_3] ، هذه القيم تُسمى بالربيعات حيث :

Q_1 تُسمى بالربيع الأول ، Q_2 تُسمى بالربيع الثاني ، Q_3 تُسمى بالربيع الثالث

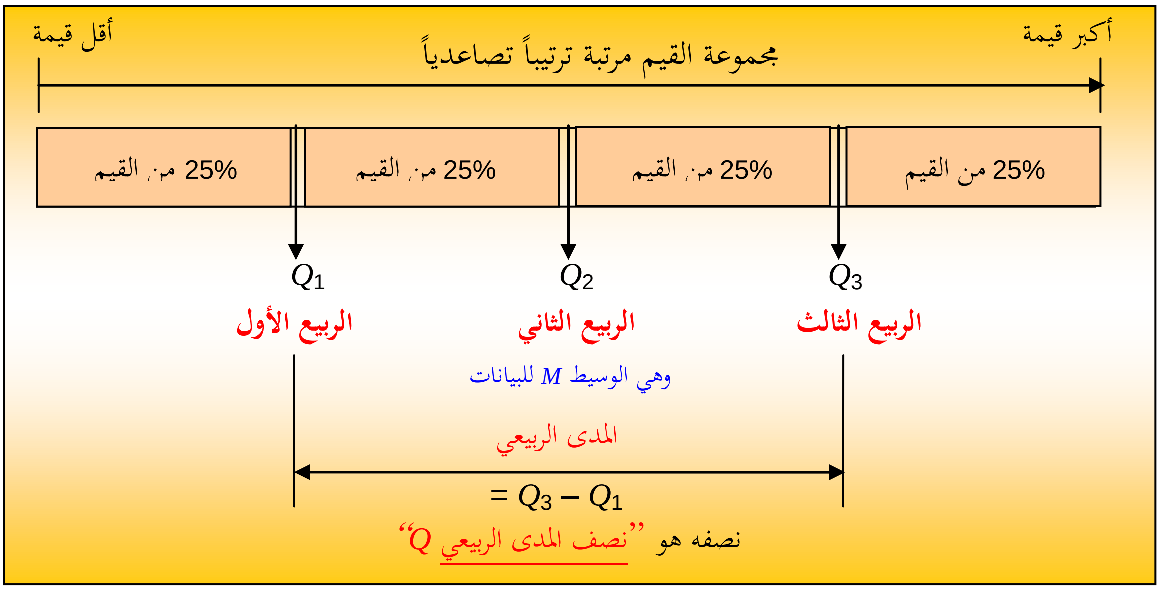
وفيها يكون :

Q_1 [**الربيع الأول**] هي تلك القيمة التي يقع تحتها 25% من القيم [وبالطبع وفوقها 75% من القيم] .

Q_2 [**الربيع الثاني**] هي تلك القيمة التي يقع تحتها 50% من القيم [وبالطبع فوقها 50% من القيم ، أي] **الوسيط M**] .

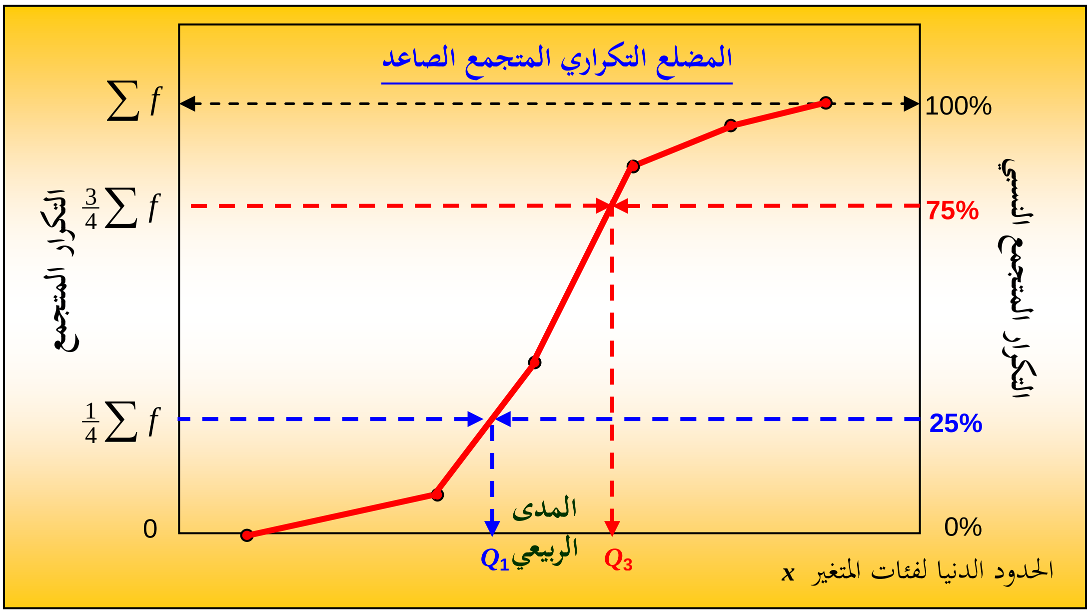
Q_3 [**الربيع الثالث**] هي تلك القيمة التي يقع تحتها 75% من القيم [وبالطبع فوقها 25% من القيم] .

والشكل التالي يعطي مخططاً لذلك :



ويمكن تحديد الربيعين Q_1 (الأول) ، Q_3 (الثالث) بنفس الطريقة التي حددنا بها الوسيط M [الربيع الثاني Q_2] ، إلا أننا سنكتفي هنا بتحديدها [ومن ثم تحديد نصف المدى الربيعي Q] للبيانات الكمية المتصلة **بيانياً** باستخدام المضلع التكراري المتجمع الصاعد **وحسابياً** باستخدام طريقة الاستكمال كالتالي :

أولاً : بيانياً :



❖ حدد قيمة المتغير المناظرة لتكرار متجمع قدره $\frac{25}{100}\sum f = \frac{1}{4}\sum f$ [أو تكرار متجمع نسبي قدره 25%] فتكون تلك القيمة هي Q_1 [الربيع الأول] .

❖ حدد قيمة المتغير المناظرة لتكرار متجمع قدره $\frac{75}{100}\sum f = \frac{3}{4}\sum f$ [أو تكرار متجمع نسبي قدره 75%] فتكون تلك القيمة هي Q_3 [الربيع الثالث] .

فيكون : المدى الربيعي $= Q_3 - Q_1$

ويكون : الانحراف الربيعي (أو نصف المدى الربيعي) $= \frac{1}{2}(Q_3 - Q_1) = Q$

ثانياً : حسابياً [طريقة الاستكمال] :

كما يمكن حساب كلٍ من Q_1 , Q_3 [ومن ثم المدى الربيعي والانحراف الربيعي] بطريقة الاستكمال كما في حالة الوسيط مع تعميم ما تم تنفيذه عند حساب الوسيط كالتالي :

لتحديد أيٍ من الربيعات الثلاثة Q_1 , Q_2 , Q_3 ، نتبع الآتي :

(1) نقوم أولاً بحساب تكرار [سنرمز له بالرمز f_0] قيمته تساوي :

$$f_0 = A \times \sum f$$

حيث A ثابت قيمته تساوي التكرار النسبي المتجمع الصاعد المناظر للقيمة المطلوب تحديدها [هنا الربيع المطلوب] ، وبالتالي فهو يساوي 0.25 [أي $\frac{25}{100}$] في حالة الربيع الأول ، ويساوي 0.5 [أي $\frac{50}{100}$] في حالة الربيع الثاني أو الوسيط ، ويساوي 0.75 [أي $\frac{75}{100}$] في حالة الربيع الثالث .

(2) مبتدئين بالرقم صفر (في ذهننا) نقوم بإضافة تكرار الفئات بصورة متتالية الواحدة تلو الأخرى مع مقارنة التكرار المتجمع الناتج بالتكرار f_0 الذي سبق وحددناه ، حتى نحصل على تكرار متجمع أكبر من أو يساوي f_0 فتكون آخر فئة أضفنا تكرارها هي **فئة الربيع** المطلوب تحديده .

(2) نحدد لتلك الفئة [فئة الربيع] كلاً من حدها الأدنى ، وتكرارها ، وطولها ، والتكرار المتجمع السابق [وهو مجموع تكرارات الفئات السابقة لتلك الفئة] .

(3) نحسب **قيمة الربيع** من :

$$\text{قيمة الربيع} = \text{الحد الأدنى لفئة هذا الربيع} + \left[\frac{(f_0 - \text{التكرار المتجمع السابق})}{\text{تكرار فئة هذا الربيع}} \times \text{طول فئة هذا الربيع}\right]$$

وعلى القارئ مقارنة العلاقة السابقة بالعلاقة التي سبق واستخدمناها لتحديد قيمة الوسيط [الربيع الثاني] وذلك في الباب الثالث .

مثال (12−4) : للتوزيع التكراري المبين ، **احسب** الوسيط ، الانحراف الربيعي .

x	$1 \le x < 3$	$3 \le x < 5$	$5 \le x < 7$	$7 \le x < 10$
f	14	29	18	9

الحل :

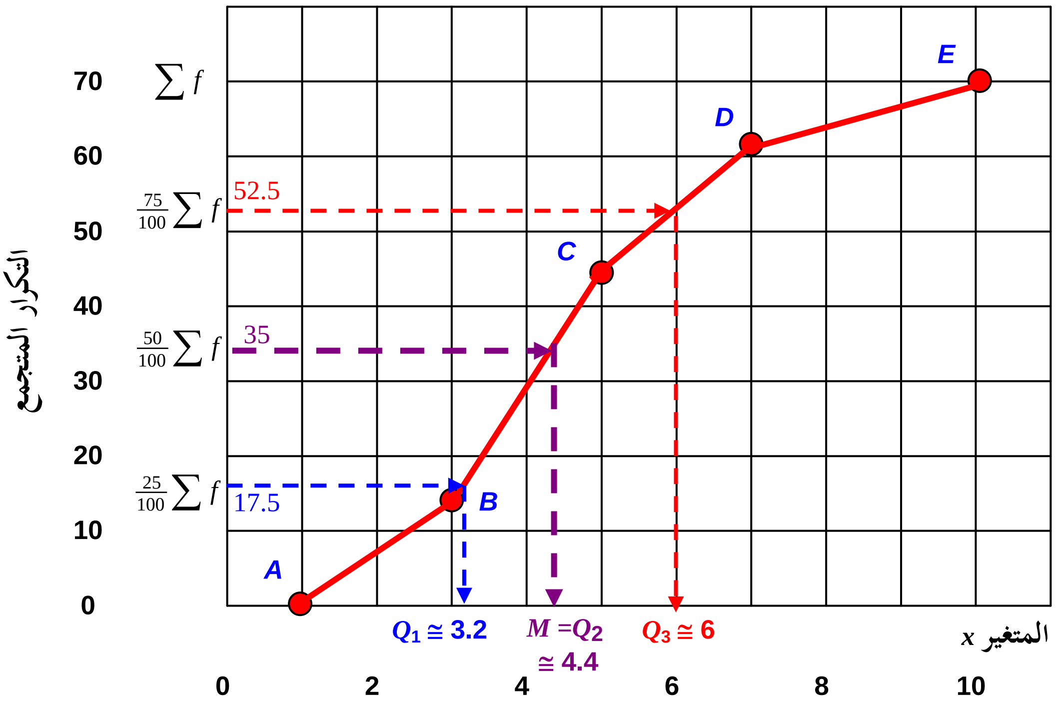
أولاً : بيانياً

الجدول التكراري المتجمع الصاعد

النقطة على المضلع	التكرار المتجمع	المتغير x
$A\ (1 , 0)$	0	< 1
$B\ (3 , 14)$	14	< 3
$C\ (5 , 43)$	43	< 5
$D\ (7 , 61)$	61	< 7
$E\ (10 , 70)$	$\sum f = 70$	< 10

من الجدول التكراري المتجمع الصاعد يمكن رسم المضلع التكراري المتجمع الصاعد ، ومنه يمكن تحديد الربيعات : الأول Q_1 ، الثاني (الوسيط) Q_2 أو M ، والربيع الثالث Q_3 حيث :

- ❖ Q_1 هي قيمة المتغير x التي تناظر تكرار متجمع صاعد قدره $17.5 = 0.25 \times \sum f$
- ❖ $M = Q_2$ هي قيمة المتغير x التي تناظر تكرار متجمع صاعد قدره $35 = 0.5 \times \sum f$
- ❖ Q_3 هي قيمة المتغير x التي تناظر تكرار متجمع صاعد قدره $52.5 = 0.75 \times \sum f$



من هذا المضلع يمكن الحصول على :

الوسيط [أو الربيع الثاني] :

$$M = Q_2 \cong 4.4$$

المدى الربيعي :

$$Q_1 \cong 3.2 \quad , \quad Q_3 \cong 6 \quad \Rightarrow \quad Q_3 - Q_1 \cong 2.8$$

الانحراف الربيعي [أو نصف المدى الربيعي] :

$$Q = \tfrac{1}{2}(Q_3 - Q_1) = 1.4$$

ثانياً : حسابياً

الجدول التكراري		
التكرار f	المتغير (المساحة) x	الفئة
14	$1 \le x < 3$	الأولى
29	$3 \le x < 5$	الثانية
18	$5 \le x < 7$	الثالثة
9	$7 \le x < 10$	الرابعة
$\sum f = 70$		

من الجدول التكراري

$$\sum f = 70$$

$$\Rightarrow \ \tfrac{1}{4}\sum f = 17.5 \, , \tfrac{1}{2}\sum f = 35 \, , \tfrac{3}{4}\sqrt{f} = 52.5$$

وبالتالي [مبتدئين بالصفر في ذهننا] :

- نقوم بإضافة تكرار الفئة الأولى (= 14) فنحصل على : $0 + 14 = 14$ وهذه القيمة أقل من جميع قيم الثلاثة السابقة $[17.5 \, , 35 \, , 52.5]$ ، وبالتالي هذه الفئة ليست فئة أي من الربيعات الثلاث .
- نضيف للناتج السابق (= 14) تكرار الفئة الثانية (= 29) فنحصل على : $14 + 29 = 43$ وهذه القيمة أكبر من 17.5 , 35 ، لكنها أقل من 52.5 ، وبالتالي هذه الفئة هي فئة كلٍ من الربيعين الأول والثاني ولكنها ليست فئة الربيع الثالث .
- نضيف للناتج السابق (= 43) تكرار الفئة الثالثة (= 18) فنحصل على : $43 + 18 = 61$ وهذه القيمة أكبر من 52.5 وبالتالي هذه الفئة [الفئة الثالثة] هي فئة الربيع الثالث .

إذن من السابق يمكن القول بأن :

الربيع	الفئة	الحد الأدنى للفئة	طول الفئة	تكرار الفئة	التكرار المتجمع السابق
الأول	الثانية	3	$5 - 3 = 2$	29	14
الثاني	الثانية	3	$5 - 3 = 2$	29	14
الثالث	الثالثة	5	$7 - 5 = 2$	18	$14 + 29 = 43$

ومن هذا الجدول يمكن حساب :

$$\text{الربيع الأول} = Q_1 = 3 + \left[\frac{0.25 \times \sum f - 14}{29}\right] \times 2 \cong 3.2$$

$$\text{[أو الوسيط] الربيع الثاني} = Q_2 = M = 3 + \left[\frac{0.5 \times \sum f - 14}{29}\right] \times 2 \cong 4.4$$

$$\text{الربيع الثالث} = Q_3 = 5 + \left[\frac{0.75 \times \sum f - 43}{18}\right] \times 2 \cong 6.1$$

$$\text{المدى الربيعي} = Q_3 - Q_1 = 6.1 - 3.2 = 2.9$$

$$\text{[نصف المدى الربيعي] الانحراف الربيعي} = Q = \tfrac{1}{2}(Q_3 - Q_1) = 1.45$$

ملحوظة : من الطبيعي أن تكون هناك فوارق بين نتائج الحل التخطيطي والحل الحسابي وهذه الفوارق تتوقف على درجة الدقة في الرسم ، فكلما زادت الدقة قلت الفوارق .

(7-4) المدى المئيني P

لمجموعة من البيانات يُعرف **المدى المئيني** على أنه الفرق بين المئين التسعين والمئين العاشر ، أي أن :

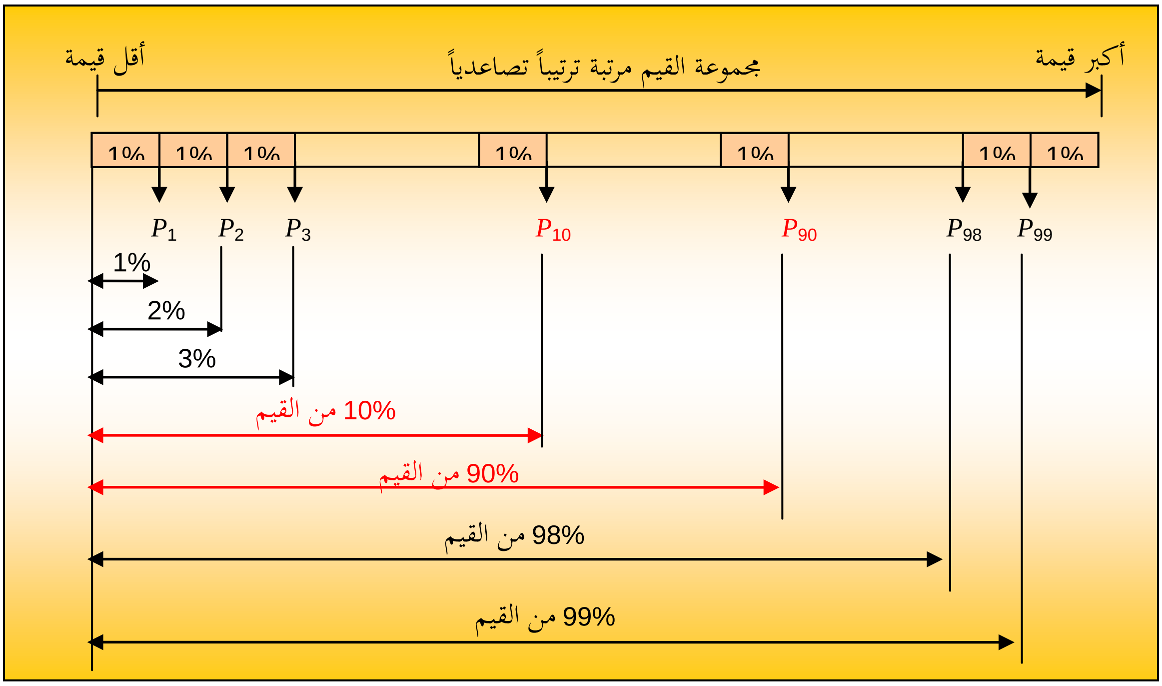
$$\text{المدى المئيني} = P_{90} - P_{10}$$

حيث P_{10} هو **المئين العاشر** ، P_{90} هو **المئين التسعون** للبيانات ، ويفضل أيضاً استخدامه في الحالات التي يستعصي فيها حساب الانحراف المتوسط أو المعياري [مثل حالة التوزيعات التكرارية المفتوحة أو حالة وجود قيم متطرفة في البيانات] .

س : ما هي المئينات ؟

ج : بنفس الطريقة التي تم بها تقسيم مجموعة من القيم إلى مجموعتين متساويتين في العدد [عن طريق الوسيط M] أو تقسيمها إلى أربعة مجموعات متساوية في العدد [عن طريق الربيعات Q_1 , Q_2 , Q_3] ، يمكن تقسيم مجموعة القيم إلى 100 مجموعة متساوية في العدد عن طريق قيم عددها 99 سنرمز لها بالرموز :

$$P_1 , P_2 , \ldots\ldots , \boldsymbol{P_{10}} , \ldots\ldots\ldots , \boldsymbol{P_{90}} , \ldots\ldots , P_{98} , P_{99}$$



تُسمى هذه القيم بـ **المئينات** ، حيث :

P_1 [**المئين الأول**] : هو قيمة يقع تحتها 1% من مجموع القيم [وبالطبع يقع فوقها 99% من القيم]

P_2 [**المئين الثاني**] : هو قيمة يقع تحتها 2% من مجموع القيم [وبالطبع يقع فوقها 98% من القيم]

....

P_{10} [المئين العاشر] : هو قيمة يقع تحتها 10% من مجموع القيم [وبالطبع يقع فوقها 90% من القيم]

P_{90} [المئين التسعون] : هو قيمة يقع تحتها 90% من مجموع القيم [وبالطبع يقع فوقها 10% من القيم]

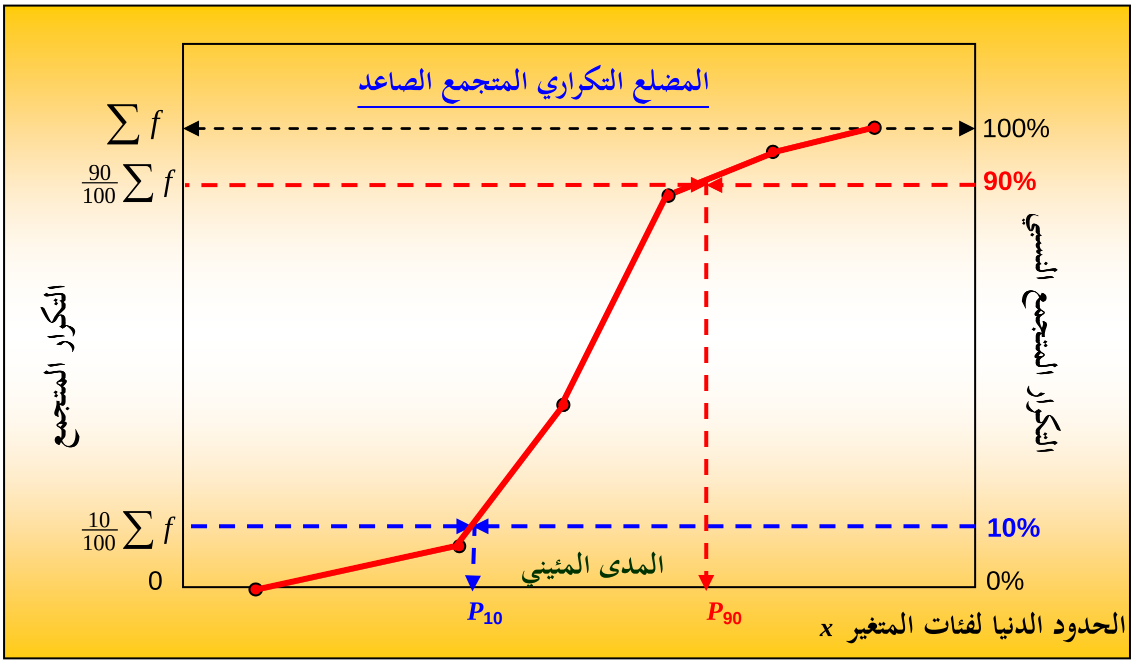
وهكذا ....

من هذه التعريفات ، هل لاحظت أن :

- ❖ المئين الخامس والعشرين P_{25} هو نفسه الربيع الأول Q_1 .
- ❖ المئين الخمسين P_{50} هو نفسه الربيع الثاني Q_2 وهو نفسه الوسيط M .
- ❖ المئين الخامس والسبعين P_{75} هو نفسه الربيع الثالث Q_3 .

ويكن تحديد المئين P_{10} (العاشر) ، P_{90} (التسعين) بنفس الطريقة التي حددنا بها الوسيط M والربيعات ، إلا أننا سنكتفي هنا بتحديدها [ومن ثم تحديد المدى المئيني P] للبيانات الكمية المتصلة **بيانياً** [باستخدام المضلع التكراري المتجمع الصاعد] وحسابياً [بطريقة الاستكمال] كالتالي :

أولاً : بيانياً



- ❖ حدد قيمة المتغير المناظرة لتكرار متجمع قدره $\frac{10}{100}\sum f = 0.1\sum f$ [أو تكرار متجمع نسبي قدره 10%] فتكون تلك القيمة هي P_{10} [المئين العاشر] .
- ❖ حدد قيمة المتغير المناظرة لتكرار متجمع قدره $\frac{90}{100}\sum f = 0.9\sum f$ [أو تكرار متجمع نسبي قدره 90%] فتكون تلك القيمة هي P_{90} [المئين التسعون] .

فيكون : المدى المئيني $= P_{90} - P_{10}$

ثانياً : حسابياً [طريقة الاستكمال]

كما يمكن حساب كلٍ من P_{90} ، P_{10} [ومن ثم المدى المئيني] بطريقة الاستكمال كما في حالة الوسيط والربيعات ، حيث :

لتحديد أيٍ من المئينات نتبع الآتي :

(1) نقوم أولاً بحساب التكرار f_0 السابق تعريفه بالصورة :

$$\boxed{f_0 = A \times \sum f}$$

حيث A ثابت قيمته تساوي التكرار النسبي المتجمع الصاعد المناظر للقيمة المطلوب تحديدها [هنا المئينات] ، وبالتالي فهو يساوي 0.01 [أي $\frac{1}{100}$] في حالة المئين الأول ، ويساوي 0.1 [أي $\frac{10}{100}$] في حالة المئين العاشر ، ويساوي 0.34 [أي $\frac{34}{100}$] في حالة المئين الرابع والثلاثين الثاني ، ويساوي 0.9 [أي $\frac{90}{100}$]، .. وهكذا .

(2) مبتدئين بالرقم صفر (في ذهننا) نقوم بإضافة تكرار الفئات بصورة متتالية الواحدة تلو الأخرى مع مقارنة التكرار المتجمع الناتج بالتكرار f_0 الذي سبق وحددناه ، حتى نحصل على تكرار متجمع أكبر من أو يساوي f_0 فتكون آخر فئة أضفنا تكرارها هي **فئة المئين** المطلوب تحديده .

(3) نحدد لتلك الفئة [فئة المئين] كلاً من حدها الأدنى ، وتكرارها ، وطولها ، والتكرار المتجمع السابق [وهو مجموع تكرارات الفئات السابقة لتلك الفئة] .

(4) نحسب **قيمة المئين** من :

$$\text{قيمة المئين} = \text{الحد الأدنى لفئة هذا المئين} + \left[\frac{(f_0 - \text{التكرار المتجمع السابق})}{\text{تكرار فئة هذا المئين}} \times \text{طول فئة هذا المئين}\right]$$

وعلى القارئ مقارنة العلاقة السابقة بالعلاقة التي سبق واستخدمناها لتحديد قيمة الوسيط والربيعات .

مثال (4-13) : للتوزيع التكراري المبين بمثال (4-8) ، **احسب** المدى المئيني .

الحل :

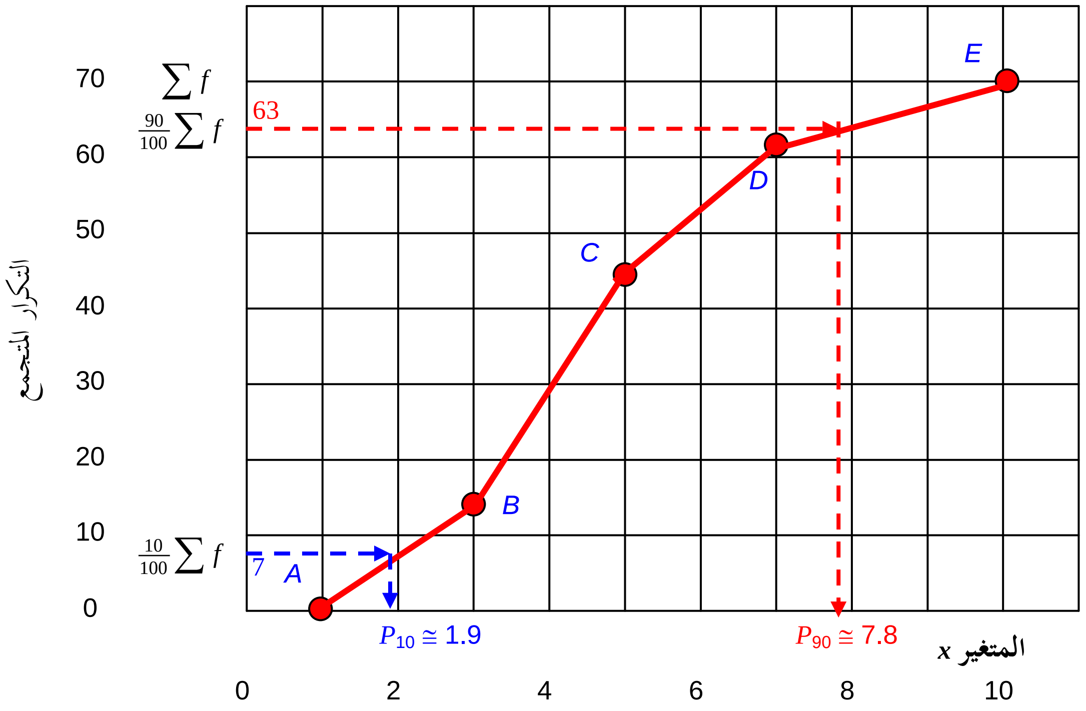
أولاً : تخطيطياً

من الجدول التكراري المتجمع الصاعد [أنظر مثال (4–12)] يمكن رسم المضلع التكراري المتجمع الصاعد ، ومنه يمكن تحديد المئينات : العاشر P_{10} ، والتسعين P_{90} حيث :

❖ P_{10} هي قيمة المتغير x التي تناظر تكرار متجمع صاعد قدره $7 = 0.1 \times \sum f$

❖ P_{90} هي قيمة المتغير x التي تناظر تكرار متجمع صاعد قدره $63 = 0.9 \times \sum f$

من هذا المضلع يمكن الحصول على :



$$P_{10} \cong 1.9 \quad , \quad P_{90} \cong 7.8 \quad \Rightarrow \quad P = P_{90} - P_{10} = 5.9 = \text{المدى المئيني}$$

ثانياً : حسابياً

الجدول التكراري		
الفئة	المتغير (المساحة) x	التكرار f
الأولى	$1 \le x < 3$	14
الثانية	$3 \le x < 5$	29
الثالثة	$5 \le x < 7$	18
الرابعة	$7 \le x < 10$	9
		$\sum f = 70$

من الجدول التكراري

$$\sum f = 70 \quad \Rightarrow \quad \frac{10}{100}\sum f = 7 \, , \, \frac{90}{100}\sum f = 63$$

وبالتالي [مبتدئين بالصفر في ذهننا] :

- نقوم بإضافة تكرار الفئة الأولى (= 14) فنحصل على : $0 + 14 = 14$ ، وهذه القيمة أكبر من 10 وأقل من 63 ، وبالتالي فهي فئة المئين العاشر P_{10} وليست فئة المئين التسعين P_{90} .
- نضيف للناتج السابق (= 14) تكرار الفئة الثانية (= 29) فنحصل على : $14 + 29 = 43$ ، وهذه القيمة ما زالت أقل من 63 ، وبالتالي هذه الفئة ليست أيضاً فئة المئين التسعين P_{90} .

- نضيف للناتج السابق (= 43) تكرار الفئة الثالثة (= 18) فنحصل على : $43+18=61$ وهذه القيمة ما زالت أقل من 63 ، وبالتالي هذه الفئة ليست أيضاً فئة المئين التسعين P_{90} .
- نضيف للناتج السابق (= 61) تكرار الفئة الرابعة (= 9) فنحصل على : $61+9=70$ وهذه القيمة أصبحت أكبر من 63 ، وبالتالي هذه الفئة هي فئة المئين التسعين P_{90} .

إذن من السابق يمكن القول بأن :

المئين	الفئة	الحد الأدنى للفئة	طول الفئة	تكرار الفئة	التكرار المتجمع السابق
العاشر	الأولى	1	$3-1=2$	14	0
التسعون	الرابعة	7	$10-7=3$	9	61

ومن هذا الجدول يمكن حساب :

$$\text{المئين العاشر} = P_{10} = 1 + \left[\frac{0.10 \times \sum f - 0}{14} \right] \times 2 = 2$$

$$\text{المئين التسعون} = P_{90} = 7 + \left[\frac{0.90 \times \sum f - 61}{9} \right] \times 3 \cong 7.7$$

$$\text{المدى المئيني} = P = P_{90} - P_{10} = 7.7 - 2 = 5.7$$

(4–8) علاقات اعتبارية بين مقاييس التشتت

في التوزيعات **متوسطة الالتواء** هناك علاقات اعتبارية (تقريبية) بين مقاييس التشتت السابقة كالآتي :

الانحراف المتوسط = $\frac{4}{5}$ × الانحراف أي $M.D = \frac{3}{4} \times s$

الانحراف الربيعي = $\frac{2}{3}$ × الانحراف أي $Q = \frac{2}{3} \times s$

الانحراف المتوسط = $\frac{6}{5}$ × الانحراف أي $M.D = \frac{6}{5} \times Q$

وبتعبير آخر فإن النسبة بين الانحراف المتوسط والانحراف المعياري والانحراف الربيعي في التوزيعات متوسطة الالتواء كالنسبة 12 إلى 15 إلى 10 ، أي أن

$$\begin{array}{ccccc} M.D & : & s & : & Q \\ 12 & : & 15 & : & 10 \end{array}$$

هذه العلاقات الاعتبارية تمكننا من حساب قيم تقريبية لبقية مقاييس التشتت متى عُلم أحدها [وذلك في حالة صلاحيتها .. أي في حالة التوزيعات التكرارية متوسطة الالتواء] ، فمثلاً :

- **إذا كان s = 30 فإن :**

$$Q = \tfrac{10}{15} \times s = \tfrac{2}{3} s = \frac{2}{3} \times 30 = \underline{\underline{20}} \qquad , \qquad M.D = \tfrac{12}{15} \times s = \tfrac{4}{5} s = \frac{4}{5} \times 30 = \underline{\underline{24}}$$

- **وإذا كان *Q* = 20 فإن :**

$$s = \tfrac{15}{10} \times Q = \tfrac{3}{2} Q = \frac{3}{2} \times 20 = \underline{\underline{30}} \qquad , \qquad M.D = \tfrac{12}{10} \times Q = \tfrac{6}{5} Q = \frac{6}{5} \times 20 = \underline{\underline{24}}$$

- **وإذا كان *M.D* = 24 فإن :**

$$Q = \tfrac{10}{12} \times M.D = \tfrac{5}{6} \times M.D = \frac{5}{6} \times 24 = \underline{\underline{20}} \qquad , \qquad s = \tfrac{15}{12} \times M.D = \tfrac{5}{4} \times M.D = \frac{5}{4} \times 24 = \underline{\underline{30}}$$

(4–9) التشتت النسبي ومقاييسه

التغير الفعلي أو التشتت كما نحصل عليه من الانحراف المتوسط أو المعياري أو الربيعي أو غيره من مقاييس التشتت يُسمى **بالتشتت المطلق** ، ولكن تشتت قدره 10 درجات عن قيمة متوسطة 50 درجة (مثلاً) يختلف عن تشتت قدره 10 درجات عن قيمة متوسطة 200 ، لذا من المناسب تعريف ما يُسمى بـ **التشتت النسبي** وهو :

$$\textbf{التشتت النسبي (كنسبة مئوية)} = \frac{\textbf{التشتت المطلق}}{\textbf{المتوسط}} \times 100$$

وبالتالي يكون التشتت النسبي لتشتت مطلق قدره 10 لبيانات متوسطها 50 هو :

$$\frac{10}{50} \times 100 = 20\%$$

بينما التشتت النسبي لتشتت مطلق قدره 10 لبيانات متوسطها 200 فهو :

$$\frac{10}{200} \times 100 = 5\%$$

ومن أكثر مقاييس التشتت النسبي استخداماً ما يُسمى **بمعامل الاختلاف** [أو **معامل التشتت**] و **معامل الاختلاف الربيعي** ، حيث :

معامل الاختلاف = $\dfrac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}}$ × 100

معامل الاختلاف الربيعي = $\dfrac{Q_3 - Q_1}{Q_3 + Q_1}$ × 100

مثال (14–4) : البيانات الموضحة بالجدول المرافق توضح توزيع الوحدات السكنية حسب الإيجار السنوي بأحد الأحياء [x يمثل الإيجار بالألف ريال ، f يمثل عدد الوحدات السكنية] .

x	$6 \le x < 10$	$10 \le x < 12$	$12 \le x < 14$	$14 \le x < 18$
f	8	20	12	10

المطلوب : تحديد كلٍ من معامل الاختلاف للإيجار و معامل الاختلاف الربيعي له .

الحل :

أولاً : بالنسبة لمعامل الاختلاف

لابد أولاً من تحديد كلٍ من الوسط الحسابي والانحراف المعياري كالتالي :

x	f	x_0	fx_0	$d = x_0 - \bar{x}$	d^2	fd^2
$6 \le x < 10$	8	8	64	8 – 12 = -4	16	128
$10 \le x < 12$	20	11	220	11 – 12 = -1	1	20
$12 \le x < 14$	12	13	156	13 – 12 = 1	1	12
$14 \le x < 18$	10	16	160	16 – 12 = 4	16	160
	50		600			320
	$\sum f$		$\sum fx_0$			$\sum fd^2$

$$\bar{x} = \frac{\sum fx_0}{\sum f} = \frac{600}{50} = 12 \quad , \quad s^2 = \frac{\sum fd^2}{\sum f} = \frac{320}{50} = 6.4 \rightarrow s = \sqrt{s^2} = \sqrt{6.4} \cong 2.53$$

وبالتالي يكون معامل الاختلاف $21.1\% \cong 100 \times \frac{2.53}{12} = 100 \times \frac{s}{\bar{x}} =$ ، أي أن الإيجار يتغير بنسبة 21.1% تقريباً .

ثانياً : بالنسبة لمعامل الاختلاف الربيعي :

لابد أولاً من تحديد الربيعين الأول والثالث كالتالي :

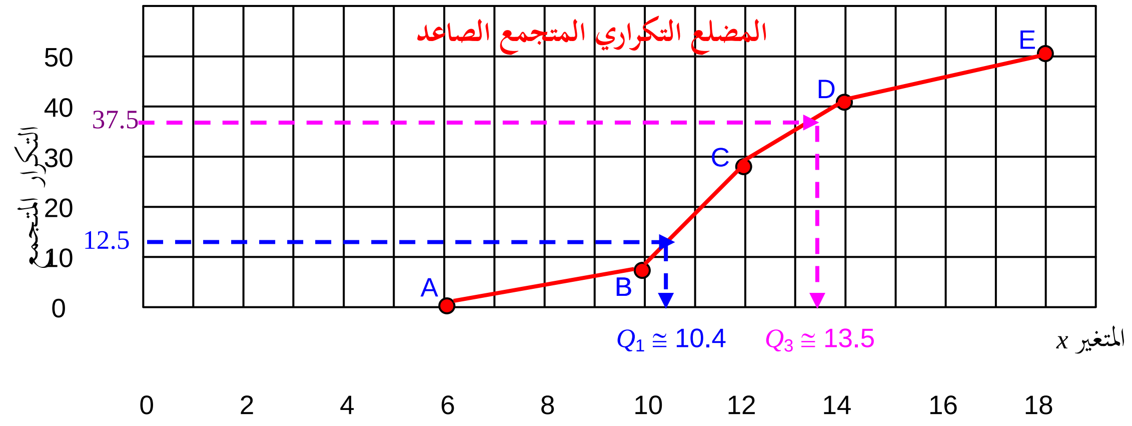
تخطيطياً :

الجدول التكراري المتجمع الصاعد

النقطة على المضلع	التكرار المتجمع	x
A (6 , 0)	0	< 6
B (10 , 8)	8	< 10
C (12 , 28)	28	< 12
D (14 , 40)	40	< 14
E (18 , 50)	$\sum f = 50$	< 18

من الجدول التكراري يمكن تكوين الجدول التكراري المتجمع الصاعد كما هو مبين ، ومن هذا الجدول التكراري المتجمع الصاعد يمكن رسم المضلع التكراري المتجمع الصاعد ، ومنه يمكن تحديد الربيع الأول Q_1 [قيمة المتغير x المناظرة لتكرار متجمع صاعد قدره $\frac{1}{4}\sum f = 12.5$] والربيع الثالث Q_3 [قيمة المتغير x المناظرة لتكرار متجمع صاعد قدره $\frac{3}{4}\sum f = 37.5$] ، وبالتالي نستطيع تحديد معامل الاختلاف الربيعي من العلاقة :

$$\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$$



إذن من المضلع التكراري المتجمع الصاعد نستنتج أن :

$$\left.\begin{array}{l} Q_1 \cong 10.4 \\ Q_3 \cong 13.5 \end{array}\right\rangle \Rightarrow \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100 = \frac{13.5 - 10.4}{13.5 + 10.4} \times 100 \cong 13\%$$

أي أن معامل الاختلاف الربيعي تقريباً يساوي 13% .

حسابياً :

من الجدول التكراري

$$\sum f = 50 \Rightarrow \frac{1}{4}\sum f = 12.5 \,, \; \frac{3}{4}\sqrt{f} = 37.5$$

التكرار f	المتغير x
8	$6 \le x < 10$
20	$10 \le x < 12$
12	$12 \le x < 14$
10	$14 \le x < 18$

$\sum f = 50$

وبالتالي [مبتدئين بالصفر في ذهننا] :

- نقوم بإضافة تكرار الفئة الأولى (= 8) فنحصل على : $0 + 8 = 8$ ، وهذه القيمة أقل من 12.5 [وبالتالي هذه الفئة ليست فئة الربيع الأول] .
- نضيف للناتج السابق (= 8) تكرار الفئة الثانية (= 20) فنحصل على : $8 + 20 = 28$ وهذه القيمة أكبر من 12.5 ولكنها أقل من 37.5 ، وبالتالي هذه الفئة [الثانية] هي فئة الربيع الأول .
- نضيف للناتج السابق (= 28) تكرار الفئة الثالثة (= 12) فنحصل على : $28 + 12 = 40$ وهذه القيمة أكبر من 37.5 وبالتالي هذه الفئة [الفئة الثالثة] هي فئة الربيع الثالث .

إذن من السابق يمكن القول بأن :

الربيع	الفئة	الحد الأدنى للفئة	طول الفئة	تكرار الفئة	التكرار المتجمع السابق
الأول	الثانية	10	12 – 10 = 2	20	8
الثالث	الثالثة	12	14 – 12 = 2	12	28

ومن هذا الجدول يمكن حساب :

$$\text{الربيع الأول} = Q_1 = 10 + \left[\frac{0.25\times\sum f - 8}{20}\right]\times 2 = 10.45$$

$$\text{الربيع الثالث} = Q_3 = 12 + \left[\frac{0.75\times\sum f - 28}{12}\right]\times 2 \cong 13.58$$

$$\text{معامل الاختلاف الربيعي} = \frac{Q_3 - Q_1}{Q_3 + Q_1}\times 100 = \frac{13.58 - 10.45}{13.58 + 10.45}\times 100 \cong 13.03\%$$

(10–4) الدرجات المعيارية

لمجموعة من القيم $x_1 , x_2 , \ldots\ldots , x_n$ وسطها الحسابي $\bar{x}$ وانحرافها المعياري s تُسمى الكمية :

$$\boxed{z = \frac{x - \bar{x}}{s}}$$

بالدرجة المعيارية للقيمة x .

مثال (15–4) : احسب الدرجات المعيارية لكل قيمة من القيم 8 3 4 12 6 7 4 8 3 5

الحل :

لابد أولاً أن نحدد كلاً من الوسط الحسابي $\bar{x}$ والانحراف المعياري s لمجموعة القيم المعطاة . من الجدول المقابل نستنتج أن :

	x	d	d^2	
	5	-1	1	
	3	-3	9	
	8	2	4	
	4	-2	4	
	7	1	1	
	6	0	0	
	12	6	36	
	4	-2	4	
	3	-3	9	
$\sum x$	8	2	4	$\sum d^2$
	60		72	

$$\bar{x} = \frac{\sum x}{n} = \frac{60}{10} = 6$$

$$s^2 = \frac{\sum d^2}{n} = \frac{72}{10} = 7.2$$

$$s = \sqrt{s^2} = \sqrt{7.2} \cong 2.68$$

وباستخدام العلاقة :

$$z = \frac{x - \bar{x}}{s}$$

يمكن إيجاد الدرجة المعيارية لكل قيمة من مجموعة القيم :

القيم	8	3	4	12	6	7	4	8	3	5
الدرجات المعيارية	0.75	-1.12	-0.75	2.24	0	0.37	-0.75	0.75	-1.12	-0.37
	$\frac{8-6}{2.68}$			$\frac{12-6}{2.68}$			$\frac{4-6}{2.68}$			$\frac{5-6}{2.68}$

وللدرجات المعيارية للقيم أهمية كبيرة في مقارنة نتائج بيانات مختلفة ببعضها حيث قد يؤدي الاعتماد على القيم الحقيقة إلى استنتاجات غير سليمة أو مضللة . لتوضيح ذلك دعنا نعتبر المثال التالي

مثال (16-4) : في الاختبار النهائي لمقرر الإحصاء حصل طالب على 82 درجة [حيث كان الوسط الحسابي للدرجات 76 بانحراف معياري 10] وحصل في مقرر الصحة واللياقة على 90 درجة [حيث كان الوسط الحسابي للدرجات 82 بانحراف معياري 16] . هل يمكن القول أن الطالب درجة استيعابه لمقرر الصحة واللياقة كانت أعلى من درجة استيعابه لمقرر الإحصاء ؟

الحل : الاعتماد على درجات الطالب في المقررين [82 في الإحصاء ، 90 في الصحة واللياقة] تجعل الإجابة : نعم درجة استيعاب الطالب لمقرر الصحة واللياقة أعلى من درجة استيعابه لمقرر الإحصاء ، ولكن الإجابة الصحيحة تعتمد على الدرجة المعيارية للطالب في كل من المقررين :

في مقرر الإحصاء	في مقرر الصحة واللياقة
$x=84$, $\bar{x}=76$, $s=10$ $\therefore \quad z=\frac{x-\bar{x}}{s}=\frac{84-76}{10}=\frac{8}{10}=\underline{\underline{0.8}}$	$x=90$, $\bar{x}=82$, $s=16$ $\therefore \quad z=\frac{x-\bar{x}}{s}=\frac{90-82}{16}=\frac{8}{16}=\underline{\underline{0.5}}$

أي أن الدرجة المعيارية للطالب في مقرر الإحصاء أعلى من نظيرتها في مقرر الصحة واللياقة ، مما يعني أن درجة استيعاب الطالب لمقرر الإحصاء أعلى من درجة استيعابه لمقرر الصحة واللياقة .

ملخص للدرس الثامن [الباب الرابع : مقاييس التشتت]

الانحراف الربيعي – المدى المئيني – التشتت النسبي – الدرجات المعيارية

- **الربيع الأول Q_1 لمجموعة من القيم:** هو قيمة تقسم مجموعة القيم [بعد ترتيبها تصاعدياً] إلى مجموعتين بحيث تقع 25% من القيم تحتها (أي أقل منها) ، 75% من القيم فوقها (أي أكبر منها) [وبالتالي هي قيمة المتغير التي يناظرها تكرار متجمع صاعد قدره $\frac{1}{4}\sum f$ في حالة القيم ذات التكرارات أو البيانات المتصلة ، أو تكرار متجمع نسبي قدره 25%] .
- **الربيع الثاني Q_2 [وهو الوسيط M] لمجموعة من القيم :** هو قيمة تقسم مجموعة القيم [بعد ترتيبها تصاعدياً] إلى مجموعتين متساويتين في العدد ، أي بحيث تقع 50% من القيم تحتها (أي أقل منها) ، 50% من القيم فوقها (أي أكبر منها) [وبالتالي هي قيمة المتغير التي يناظرها تكرار متجمع صاعد قدره $\frac{1}{2}\sum f$ في حالة القيم ذات التكرارات أو البيانات المتصلة ، أو تكرار متجمع نسبي قدره 50%] .
- **الربيع لثالث Q_3 لمجموعة من القيم:** هو قيمة تقسم مجموعة القيم [بعد ترتيبها تصاعدياً] إلى مجموعتين بحيث تقع 75% من القيم تحتها (أي أقل منها) ، 25% من القيم فوقها (أي أكبر منها) [وبالتالي هي قيمة المتغير التي يناظرها تكرار متجمع صاعد قدره $\frac{3}{4}\sum f$ في حالة القيم ذات التكرارات أو البيانات المتصلة ، أو تكرار متجمع نسبي قدره 75%] .
- **المئين العاشر P_{10} لمجموعة من القيم:** هو قيمة تقسم مجموعة القيم [بعد ترتيبها تصاعدياً] إلى مجموعتين بحيث تقع 10% من القيم تحتها (أي أقل منها) ، 90% من القيم فوقها (أي أكبر منها) [وبالتالي هي قيمة المتغير التي يناظرها تكرار متجمع صاعد قدره $\frac{10}{100}\sum f$ (أو $\frac{1}{10}\sum f$) في حالة القيم ذات التكرارات أو البيانات المتصلة ، أو تكرار متجمع نسبي قدره 10%] .
- **المئين التسعون P_{90} لمجموعة من القيم:** هو قيمة تقسم مجموعة القيم [بعد ترتيبها تصاعدياً] إلى مجموعتين بحيث تقع 90% من القيم تحتها (أي أقل منها) ، 10% من القيم فوقها (أي أكبر منها) [وبالتالي هي قيمة المتغير التي يناظرها تكرار متجمع صاعد قدره $\frac{90}{100}\sum f$ (أو $\frac{9}{10}\sum f$) في حالة القيم ذات التكرارات أو البيانات المتصلة ، أو تكرار متجمع نسبي قدره 90%] .

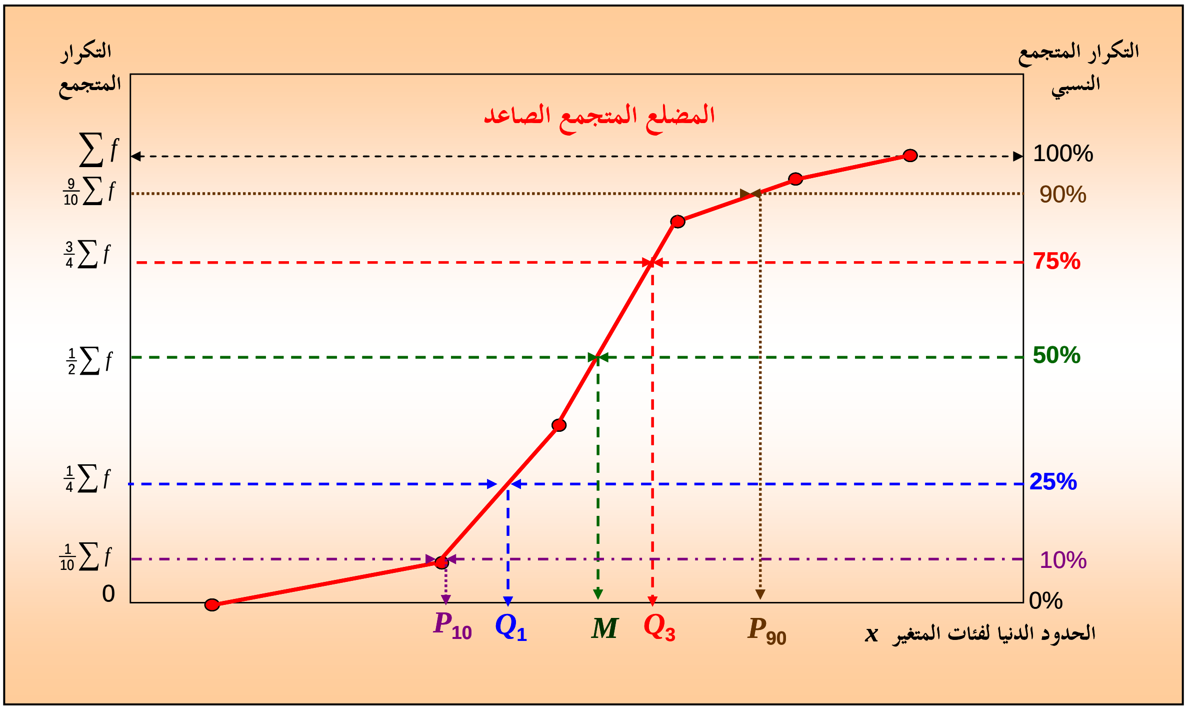
تذكر أن : الوسيط هو نفسه الربيع الثاني هو نفسه المئين الخمسين [أي $M = Q_2 = P_{50}$] ، وأن الربيع الأول هو نفسه المئين 25 [أي $Q_1 = P_{25}$] وأن الربيع الثالث هو نفسه المئين 75 [أي $Q_3 = P_{75}$]

- وللبيانات الكمية المتصلة تتحدد القيم السابقة بيانياً من المضلع التكراري المتجمع الصاعد بنفس الأسلوب كالتالي :

- ❖ احسب قيمة $f_0 = A \times \sum f$ حيث A ثابت يساوي التكرار النسبي المتجمع الصاعد المناظر للقيمة المطلوب تحديدها ، والجدول التالي يحدد قيمة A لبعض القيم .

القيمة المطلوب تحديدها	المئين العاشر	الربيع الأول أو المئين الخامس والعشرون	الوسيط أو الربيع الثاني أو المئين الخمسون	الربيع الثالث أو المئين الخامس والسبعون	المئين التسعون
قيمة A	$\frac{10}{100} = 0.1$	$\frac{25}{100} = \frac{1}{4} = 0.25$	$\frac{50}{100} = \frac{2}{4} = \frac{1}{2} = 0.5$	$\frac{75}{100} = \frac{3}{4} = 0.75$	$\frac{90}{100} = \frac{9}{10} = 0.9$

- ❖ حدد على المحور الرأسي التكرار المتجمع الصاعد f_0 ثم ارسم خطاً أفقياً حتى المضلع ثم اتجه رأسياً لأسفل نحو المحور الأفقي وارصد قيمة المتغير فتكون هي القيمة المطلوب .



- كما يمكن تحديد أي قيمة من قيم الربيعات أو المئينات بنفس الأسلوب التالي :

- ❖ قم بحساب $f_0 = A \times \sum f$ كما سبق وبينا .
- ❖ مبتدئين بالرقم صفر (في ذهننا) قم بإضافة تكرار الفئات بصورة متتالية الواحدة تلو الأخرى مع مقارنة التكرار المتجمع الناتج بالتكرار f_0 الذي سبق وحددناه ، حتى تحصل على تكرار متجمع أكبر من أو يساوي f_0 فتكون آخر فئة قمت بإضافة تكرارها هي **فئة القيمة** المطلوب تحديدها [الربيع أو المئين المطلوب] .

❖ حدد لتلك الفئة [فئة القيمة المطلوب حسابها] كلاً من حدها الأدنى ، وتكرارها ، وطولها ، والتكرار المتجمع السابق [وهو مجموع تكرارات الفئات السابقة لتلك الفئة] .

❖ نحسب **القيمة المطلوبة** من :

$$\text{القيمة المطلوب} = \text{الحد الأدنى لفئة هذه القيمة} + \left[\frac{(f_0 - \text{التكرار المتجمع السابق})}{\text{تكرار فئة هذه القيمة}} \times \text{طول فئة هذه القيمة}\right]$$

- **ويرتبط بالكميات السابقة الآتي :**

المدى الربيعي [وهو أحد مقاييس التشتت] = $Q_3 - Q_1$

المدى المئيني [وهو أحد مقاييس التشتت] = $P_{90} - P_{10}$

الانحراف الربيعي [وهو أحد مقاييس التشتت] = **نصف المدى الربيعي** = $\frac{1}{2}(Q_3 - Q_1)$

معامل الاختلاف الربيعي [وهو أحد مقاييس التشتت النسبي ويُكتب كنسبة مئوية] = $\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$

- وفي التوزيعات متوسطة الالتواء فإن النسبة بين الانحراف المتوسط والانحراف المعياري والانحراف الربيعي هي كالآتي :

$$M.D : s : Q$$
$$12 : 15 : 10$$

- **مقاييس التشتت النسبي :** هي مقاييس تحدد النسبة المئوية للتشتت المطلق بالنسبة لقيمة متوسطة . **أمثلة :** معامل الاختلاف [أو معامل التشتت] – معامل الاختلاف الربيعي ، حيث :

❖ معامل الاختلاف (أو معامل التشتت) : وهو مقياس للتشتت النسبي [ويُعبر عنه كنسبة مئوية كالآتي] ويرتبط بالوسط الحسابي والانحراف المعياري بالعلاقة :

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}} \times 100$$

أي $\frac{s}{\bar{x}} \times 100$

❖ **معامل الاختلاف الربيعي** [وهو أحد مقاييس التشتت النسبي ويُكتب كنسبة مئوية] =

$$\boxed{\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100}$$

- **الدرجة المعيارية z لقيمة x** : وهي ذات أهمية كبيرة في مقارنة نتائج مختلفة ببعضها $\boxed{z = \frac{x - \bar{x}}{s}}$.

تدريب عملي

في كل حالة من الحالات التالية ، **احسب** مقياساً مناسباً للنزعة المركزية ، مقياساً مناسباً للتشتت ، ومقياساً مناسب للتشتت النسبي :

(أ)

x	$60-$	$70-$	$80-$	$90-100$
f	98	75	56	42

(ب)

x	< 3	$3-$	$6-$	> 9
f	3	7	16	12

سؤال للقارئ

بافتراض أنك حصلت في مقرر الإحصاء على 75 درجة [حيث بلغ متوسط درجات الطلاب في هذا الاختبار 72 درجة بانحراف معياري 8] ، وبافتراض أنك حصلت في اختبار مقرر علم النفس على 68 درجة [حيث بلغ متوسط درجة الطلاب في هذا الاختبار 63 درجة بانحراف معياري 6 درجات] ، هل يمكن القول بأن استيعابك لمقرر الإحصاء أفضل من استيعابك لمقرر علم النفس ؟

تدريبات (8)

الإجابة النهائية لجميع التمرينات موجودة في نهاية التدريب

اختر الإجابة الصحيحة

(1) **مقاييس التشتت النسبي** هي

(أ) قيم نموذجية يمكن أن تمثل مجموعة البيانات

(ب) مقاييس ترصد الدرجة التي تتجه بها البيانات الكمية للانتشار حول قيمة متوسطة

(ج) مقاييس تحدد النسبة المئوية للتشتت المطلق بالنسبة لقيمة متوسطة

(د) هي مقاييس ترصد درجة تماثل أو البعد عن التماثل لتوزيع ما

(هـ) مقاييس ترصد درجة التدبب في قمة المنحنى مقارنة بقمة منحنى التوزيع الطبيعي

(2) **معامل الاختلاف** هو أحد مقاييس

(أ) النزعة المركزية	(ب) التشتت
(ج) الالتواء	(د) التشتت النسبي

ويمكن أن يُستبدل "معامل الاختلاف" في رأس السؤال بـ "معامل الاختلاف الربيعي"

(3) معامل الاختلاف (أو معامل التشتت) يساوي :

(أ) [الوسط الحسابي ÷ الانحراف المعياري] × 100

ب) الوسط الحسابي – الانحراف المعياري

(ج) [الانحراف المعياري ÷ الوسط الحسابي] × 100

د) الانحراف المعياري – الوسط الحسابي

(4) هو قيمة تقسم مجموعة القيم [بعد ترتيبها تصاعدياً] إلى مجموعتين بحيث تقع 25% من القيم تحتها (أي أقل منها) ، 75% من القيم فوقها (أي أكبر منها) .

(أ) الربيع الأول	(ب) الوسيط
(ج) الربيع الثالث	(د) المئين العاشر

(5) هو قيمة تقسم مجموعة القيم [بعد ترتيبها تصاعدياً] إلى مجموعتين بحيث تقع 75% من القيم تحتها (أي أقل منها) ، 25% من القيم فوقها (أي أكبر منها) .

(أ) الربيع الأول (ب) الوسيط
(ج) الربيع الثالث (د) المئين العاشر

(6) هو قيمة تقسم مجموعة القيم [بعد ترتيبها تصاعدياً] إلى مجموعتين بحيث تقع 10% من القيم تحتها (أي أقل منها) ، 90% من القيم فوقها (أي أكبر منها) .

(أ) المئين التسعون (ب) الوسيط
(ج) الربيع الثالث (د) المئين العاشر

(7) هو قيمة تقسم مجموعة القيم [بعد ترتيبها تصاعدياً] إلى مجموعتين بحيث تقع 90% من القيم تحتها (أي أقل منها) ، 10% من القيم فوقها (أي أكبر منها) .

(أ) المئين التسعون (ب) الوسيط
(ج) الربيع الثالث (د) المئين العاشر

(8) الوسيط لمجموعة من القيم هو نفسه

(أ) المئين العاشر (ب) الربيع الأول
(ج) الربيع الثاني (د) الربيع الثالث

(9) الوسيط لمجموعة من القيم هو نفسه

(أ) المئين العاشر (ب) الربيع الأول
(ج) المئين الخمسون (د) الربيع الثالث

(10) الربيع الأول لمجموعة من القيم هو نفسه

(أ) المئين رقم 25 (ج) المئين رقم 75
(ج) نصف الوسيط (د) الوسيط

(11) الربيع الثالث لمجموعة من القيم هو نفسه

(أ) المئين رقم 25 (ج) المئين رقم 75
(ج) نصف الوسيط (د) الوسيط

(12) المدى الربيعي يساوي

(أ) ضعف الانحراف الربيعي

(ب) نصف الانحراف الربيعي

(ج) الانحراف الربيعي

(د) المدى المئيني

خاص بالأسئلة من (13) إلى (18) :

إذا كان [مجموعة من القيم] Q_1 هو الربيع الأول ، Q_3 هو الربيع الثالث ، P_{10} هو المئين العاشر ، P_{90} هو المئين التسعون ، M هو الوسيط ، فإن :

(13) المدى الربيعي لمجموعة القيم يساوي :

(أ) $\frac{1}{2}(Q_3 - Q_1)$

(ب) $\frac{1}{2}(P_{90} - P_{10})$

(ج) $(Q_3 - Q_1)$

(د) $(P_{90} - P_{10})$

(14) المدى المئيني لمجموعة القيم يساوي :

(أ) $\frac{1}{2}(Q_3 - Q_1)$

(ب) $\frac{1}{2}(P_{90} - P_{10})$

(ج) $(Q_3 - Q_1)$

(د) $(P_{90} - P_{10})$

(15) الانحراف الربيعي لمجموعة القيم يساوي :

(أ) $\frac{1}{2}(Q_3 - Q_1)$

(ب) $\frac{1}{2}(P_{90} - P_{10})$

(ج) $(Q_3 - Q_1)$

(د) $(P_{90} - P_{10})$

(16) معامل الاختلاف الربيعي لمجموعة القيم يساوي :

(أ) $\frac{P_{90} - P_{10}}{2(Q_3 - Q_1)} \times 100$

(ب) $\frac{Q_3 - Q_1}{2(P_{90} - P_{10})} \times 100$

(ج) $\frac{Q_3 + Q_1}{Q_3 - Q_1} \times 100$

(د) $\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$

(17) للمنحنيات التكرارية وحيدة المنوال وبسيطة الالتواء يكون الانحراف المتوسط مساوياً (تقريباً) لـ :

(أ) $\frac{4}{5}$ × الانحراف المعياري

(ب) $\frac{3}{2}$ × الانحراف المعياري

(ج) $\frac{5}{4}$ × الانحراف المعياري

(د) $\frac{2}{3}$ × الانحراف المعياري

(18) للمنحنيات التكرارية وحيدة المنوال وبسيطة الالتواء يكون الانحراف الربيعي مساوياً (تقريباً) لـ :

(أ) $\frac{4}{5}$ × الانحراف المعياري

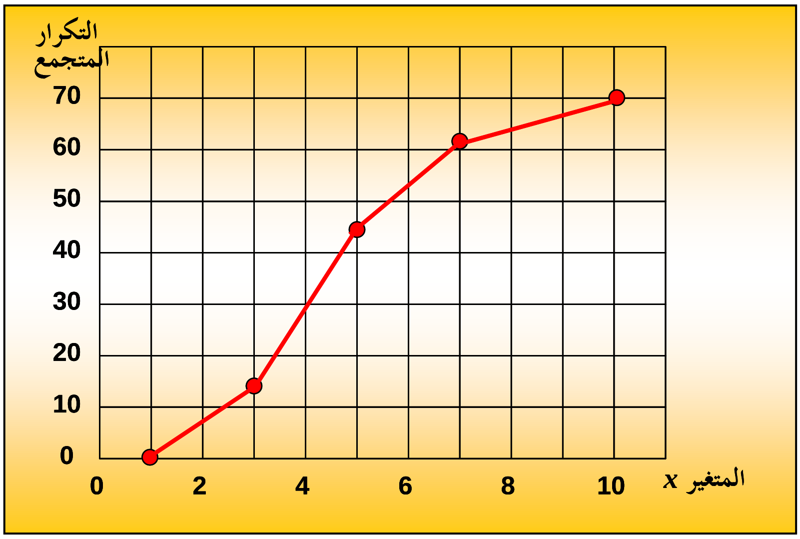
(ب) $\frac{3}{2}$ × الانحراف المعياري

(ج) $\frac{5}{4}$ × الانحراف المعياري

(د) $\frac{2}{3}$ × الانحراف المعياري

خاص بالأسئلة من (19) إلى (25) :

الشكل المرافق يبين المضلع التكراري المتجمع الصاعد لمتغير متصل x :



(19) مجموع التكرارات يساوي :

(أ) 5 (ب) 10 (ج) 35 (د) 70

(20) الربيع الأول يقع بين :

(أ) 2 , 3 (ب) 3 , 4 (ج) 4 , 5 (د) 5 , 6

(21) الربيع الثاني يقع بين :

(أ) 2 , 3 (ب) 3 , 4 (ج) 4 , 5 (د) 5 , 6

(22) الربيع الثالث يقع بين :

(أ) 2 , 3 (ب) 3 , 4 (ج) 4 , 5 (د) 5 , 6

(23) المئين العاشر يقع بين :

(أ) 1 , 2 (ب) 4 , 5 (ج) 7 , 8 (د) 9 , 10

(24) المئين الخمسون يقع بين :

(أ) 1 , 2 (ب) 4 , 5 (ج) 7 , 8 (د) 9 , 10

(25) المئين التسعون يقع بين :

(أ) 1 , 2 (ب) 4 , 5 (ج) 7 , 8 (د) 9 , 10

الإجابة :

(1) ج (2) ب (3) ج (4) أ (5) ج
(6) د (7) أ (8) ج (9) ج (10) أ
(11) ب (12) أ (13) ج (14) د (15) أ
(16) د (17) أ (18) د (19) د (20) ب
(21) ج (22) د (23) أ (24) ج (25) ج

عناصر الدرس

الباب الخامس : الالتواء والتفرطح

- تعريف الالتواء
- مقاييس الالتواء
- تعريف التفرطح
- مقاييس التفرطح

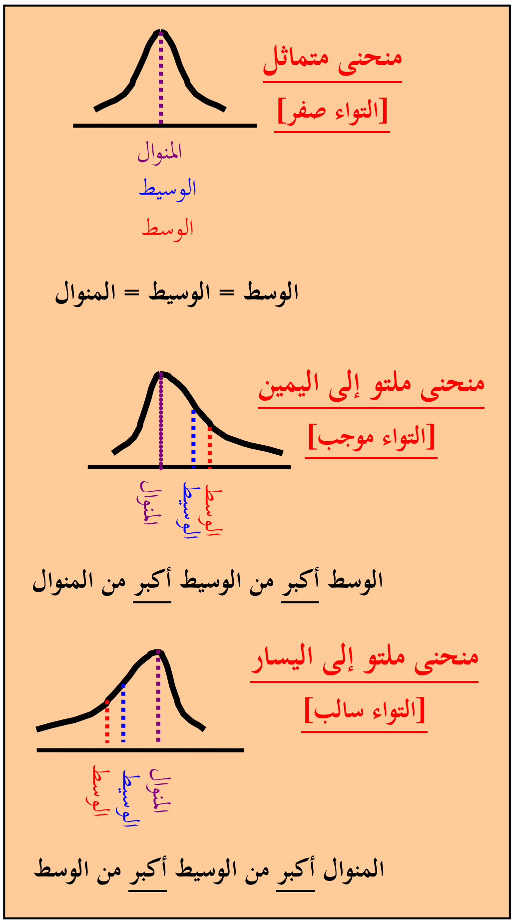
الدرس التاسع

الباب الخامس : الالتواء والتفرطح

(5–1) تعريف الالتواء

ذكرنا سابقاً أن المنحنيات التكرارية التي تظهر في الناحية العملية تأخذ أشكالاً مميزة منها الآتي :

- **منحنى متماثل (ناقوسي) :** وهو المنحنى الذي إذا قسمناه إلى نصفين انطبق هذان النصفان على بعضهما البعض تماماً ، وفيه تكون النهاية العظمى في المنتصف وتكون المشاهدات المتساوية البعد عن مركز النهاية العظمى لها نفس التكرارات .
- **منحنى ملتوي لليمين (التواء موجب) :** وفيه تكون يكون المنحنى قريباً من التماثل لكن طرفه على الجانب **الأيمن** من مركز النهاية العظمى يكون ممتداً أكثر من طرفه على الجانب الأيسر من هذا المركز ، وفي هذه الحالة يُقال للمنحنى أن له **التواء موجب** .
- **منحنى ملتوي لليسار (التواء سالب) :** وفيه تكون يكون المنحنى قريباً من التماثل لكن طرفه على الجانب **الأيسر** من مركز النهاية العظمى يكون ممتداً أكثر من طرفه على الجانب الأيمن من هذا المركز ، وفي هذه الحالة يُقال للمنحنى أن له **التواء سالب** .



ويُعرف **الالتواء** على أنه **درجة تماثل أو البعد عن التماثل لتوزيع ما :**

- فإذا كان المنحنى له ذيل أكبر إلى يمين النهاية العظمى للمنحنى عنه إلى يسارها يُسمى التوزيع ملتوي إلى اليمين [أو موجب الالتواء] وعندئذٍ يقع الوسط الحسابي يمين المنوال [أي الوسط يكون أكبر من المنوال] .
- وإذا كان المنحنى له ذيل أكبر إلى يسار النهاية العظمى للمنحنى عنه إلى يمينها يُسمى التوزيع ملتوي إلى اليسار [أو سالب الالتواء] وعندئذٍ يقع الوسط الحسابي يسار المنوال [أي المنوال يكون أكبر من الوسط] .

(5–2) مقاييس الالتواء :

يُقاس الالتواء بعدة مقاييس [كل منها يُسمى بـ معامل الالتواء] منها :

- **معامل بيرسون للالتواء :**

$$\text{معامل بيرسون الأول للالتواء } sk_1 = \frac{\text{الوسط} - \text{المنوال}}{\text{الانحراف المعياري}} \qquad (5\text{-}1)$$

وتُستخدم إذا علمنا الوسط الحسابي والمنوال (ويكون وحيداً) وكذلك الانحراف المعياري

وبالاستفادة من العلاقة بين مقاييس النزعة المركزية : الوسط الحسابي ، الوسيط ، والمنوال ، يمكن كتابة العلاقة *(5-1)* على الصورة :

$$\text{معامل بيرسون الثاني للالتواء } sk_2 = \frac{3(\text{الوسط} - \text{الوسيط})}{\text{الانحراف المعياري}} \qquad (5\text{-}2)$$

وتُستخدم إذا علمنا الوسط الحسابي والوسيط وكذلك الانحراف المعياري

- **معامل باولي للالتواء [أو معامل الالتواء الربيعي] :**

$$\text{معامل الالتواء الربيعي } sk_3 = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1} \qquad (5\text{-}3)$$

وتُستخدم إذا علمنا الربيعات الأول والثالث وأيضاً الربيع الثاني (الوسيط)

- **معامل الالتواء المئيني :**

$$\text{معامل الالتواء المئيني } sk_4 = \frac{P_{90} - 2P_{50} + P_{10}}{P_{90} - P_{10}} \qquad (5\text{-}4)$$

وتُستخدم إذا علمنا المئنات العاشر والتسعين وأيضاً المئين الخمسين (الوسيط)

تذكر أن :

$$\underbrace{P_{50}}_{\text{المئين الخمسون}} = \underbrace{M}_{\text{الوسيط}} = \underbrace{Q_2}_{\text{الربيع الثاني}}$$

وتحدد إشارة معامل الالتواء على نوع الالتواء ، فالإشارة الموجبة لمعامل الالتواء تعني التواءً موجباً [أي جهة اليمين] ، والإشارة السالبة تعني التواءً سالباً [أي جهة اليسار] .

مثال (5-1) :

في كل حالة من الحالات التالية **احسب** معامل الالتواء المناسب للتوزيع المعطى بياناته مع **توضيح نوع الالتواء** (لليمين/لليسار) :

(أ) الوسط الحسابي $\bar{x} = 80$ ، المنوال $\hat{X} = 82$ ، الانحراف المعياري $s = 20$

(ب) الوسط الحسابي $\bar{x} = 80$ ، الوسيط $M = 79$ ، الانحراف المعياري $s = 10$

(ج) الربيع الأول $Q_1 = 68$ ، الوسيط $M = 79$ ، الربيع الثالث $Q_3 = 91$

(د) المئين العاشر $P_{10} = 58$ ، الوسيط $M = 79$ ، المئين التسعون $P_{90} = 99$

الحل :

تتوقف العلاقة المناسبة لحساب معامل الالتواء على معطيات المسألة :

(أ) في الحالة الأولى نعلم الوسط الحسابي ، المنوال ، والانحراف المعياري :

هنا العلاقة *(5-1)* هي العلاقة المناسبة لحساب معامل الالتواء sk :

$$sk = sk_1 = \frac{\bar{x} - \hat{X}}{s} = \frac{80 - 82}{20} = \frac{-2}{20} = -0.1$$

ومعامل الالتواء هنا **سالب** مما يعني أن الالتواء جهة اليسار .

(ب) في الحالة الثانية نعلم الوسط الحسابي ، الوسيط ، والانحراف المعياري :

هنا العلاقة *(5-2)* هي العلاقة المناسبة لحساب معامل الالتواء sk :

$$sk = sk_2 = \frac{9(\bar{x} - M)}{s} = \frac{3(80 - 79)}{10} = \frac{3}{10} = 0.3$$

ومعامل الالتواء هنا **موجب** مما يعني أن الالتواء جهة اليمين .

(ج) في الحالة الثالثة نعلم الربيعات الأول ، الثاني (= الوسيط) ، والثالث :

هنا العلاقة *(5-3)* هي العلاقة المناسبة لحساب معامل الالتواء sk [معامل الالتواء الربيعي] :

$$sk = sk_3 = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1} = \frac{91 - 2 \times 79 + 68}{91 - 68} = \frac{1}{23} \cong 0.04$$

ومعامل الالتواء هنا **موجب** مما يعني أن الالتواء جهة اليمين [التواء بسيط جداً].

(د) في الحالة الرابعة نعلم المئينات العاشر ، الخمسون (= الوسيط) ، والتسعون :

هنا العلاقة *(5-4)* هي العلاقة المناسبة لحساب معامل الالتواء sk [معامل الالتواء المئيني] :

$$sk = sk_4 = \frac{P_{90} - 2P_{50} + P_{10}}{P_{90} - P_{10}} = \frac{99 - 2 \times 79 + 58}{99 - 58} = \frac{-1}{41} \cong -0.02$$

ومعامل الالتواء هنا **سالب** مما يعني أن الالتواء جهة اليسار [التواء بسيط جداً].

مثال (2–5) [مثال عام] :

البيانات التالية توضح توزيع مجموعة من المدرسين العاملين في مجال التربية في إحدى المدارس وفقاً لفئات أعمارهم :

فئات العمر x	20 -	30 -	40 -	50 - 60
عدد المدرسين	10	30	50	20

المطلوب حساب :

(1) الوسط الحسابي ، الانحراف المتوسط ، والانحراف المعياري للأعمار .

(2) الوسيط ، الانحراف الربيعي ، والانحراف المئيني للأعمار

(3) منوال الأعمار .

(4) معامل الالتواء .

الحل :

(1) الوسط الحسابي – الانحراف المتوسط – الانحراف المعياري

يمكن حساب هذه الكميات معاً وذلك بعمل جدول يبين الفئات ، وتكراراتها ، ومراكزها ، وانحراف هذه المراكز عن الوسط الحسابي [بالطبع بعد حساب الوسط الحسابي] ، ومربعات هذه الانحرافات ، ثم استخدام العلاقات :

$$\bar{x} = \frac{\sum fx_0}{\sum f} \quad , \quad M.D = \frac{\sum f\,|\,d\,|}{\sum f} \quad , \quad s^2 = \frac{\sum fd^2}{\sum f}$$

لحساب الوسط الحسابي ، والانحراف المتوسط ، والتباين لأعلى الترتيب] ، ويكون الجذر التربيعي للتباين هو الانحراف المعياري كما يلي :

x	**f**	x_0	fx_0	$d = x - x_0$	$\|d\|$	$f\,\|d\|$	fd^2
20 -	10	25	250	-17.27	17.27	172.7	2982.53
30 -	30	35	1050	-7.27	7.27	218.1	1585.59
40 -	50	45	2250	2.73	2.73	136.5	372.65
50 - 60	20	55	1100	12.73	12.73	254.6	3241.06

	$\sum f = 110$		$\sum fx_0 = 4650$		$\sum f\,|d| = 781.9$	$\sum fd^2 = 8181.83$

$$\bar{x} = \frac{\sum fx_0}{\sum f} = \frac{4650}{110} \cong 42.27$$

$$M.D = \frac{\sum f\,|d|}{\sum f} = \frac{781.9}{110} \cong 7.11$$

$$s^2 = \frac{\sum fd^2}{\sum f} = \frac{8181.83}{110} \cong 74.38 \quad \Rightarrow \quad s = \sqrt{s^2} = \sqrt{74.38} \cong 8.62$$

أيضاً كان من الممكن حساب التباين باستخدام العلاقة :

$$s^2 = \frac{\sum fx_0^2}{\sum f} - \bar{x}^2$$

وذلك من الجدول التالي :

x	f	x_0	fx_0	fx_0^2
20 -	***10***	***25***	***250***	***6250***
30 -	***30***	***35***	***1050***	***36750***
40 -	***50***	***45***	***2250***	***101250***
50 - 60	***20***	***55***	***1100***	***60500***
	$\sum f = 110$		$\sum fx_0 = 4650$	$\sum fx_0^2 = 204750$

$$\bar{x} = \frac{\sum fx_0}{\sum f} = \frac{4650}{110} \cong 42.27$$

$$s^2 = \frac{\sum fx_0^2}{\sum f} - \bar{x}^2 = \frac{204750}{110} - (42.27)^2 \cong 74.61$$

$$s = \sqrt{s^2} = \sqrt{74.61} \cong 8.64$$

ملحوظة : الفارق في الحسابات بين الطريقتين راجع إلى التقريبات التي يتم استخدامها

(2) الوسيط – الانحراف الربيعي – المدى المئيني

كما رأينا أن الكميات الثلاث : الوسط الحسابي ، الانحراف المتوسط ، والانحراف المعياري كميات مرتبطة ببعضها ، كذلك الوسيط ، الربيعات [ومن ثم الانحراف الربيعي] ، والمئينات [ومن ثم المدى المئيني] كميات تتفق في طريقة حسابها ، فكلٍ من الوسيط والربيعات والمئينات تتحدد بيانياً عن طريق المنحنى التكراري المتجمع الصاعد (أو الهابط) وتتحدد حسابياً بطريقة الاستكمال ، وسوف نكتفي هنا بتحديدها حسابياً وعلى القارئ العزيز محاولة تحديدها بيانياً ومقارنة نتائج الحل البياني بالحل التحليلي الموضح بالأسفل :

طريقة الاستكمال لتحديد الوسيط ، الربيعات ، والمئينات :

إذا فرضنا أننا سمينا المطلوب تحديده [الوسيط – الربيعات – المئينات] بال "**قيمة**" فإنه لابد أولاً من تحديد "**ثابت القيمة**" A والذي يساوي التكرار النسبي المتجمع الصاعد المناظر لتلك القيمة ، وبعد تحديد قيمة الثابت A نقوم بحساب قيمة التكرار f_0 المعرف بالعلاقة *(5-13)* حيث :

$$f_0 = A \times \sum f$$

جدول (1) يبين قيم كلٍ من الثابت A والتكرار f_0 للقيم المختلفة المطلوب تحديدها .

جدول (1)		
القيمة المطلوب تحديدها	A	$f_0 = A \times \sum f$
المئين العاشر P_{10}	*0.10*	$0.10 \times 110 = 11$
الربيع الأول Q_1	*0.25*	$0.25 \times 110 = 27.5$
الوسيط (الربيع الثاني) M	*0.5*	$0.50 \times 110 = 55$
الربيع الثالث Q_3	*0.75*	$0.75 \times 110 = 82.5$
المئين التسعون P_{90}	*0.90*	$0.90 \times 110 = 99$

الجدول التكراري		
f	x	الفئة
10	*20 -*	الأولى
30	*30 -*	الثانية
50	*40 -*	الثالثة
20	*50 - 60*	الرابعة
$\sum f = 110$		

وبنفس الأسلوب الذي اتبعناه عند تحديد الوسيط ، الربيعات ، والمئينات يمكن تحديد قيم كلٍ من الوسيط M ، الربيعات Q_1 , Q_3 ، والمئينات P_{10} , P_{90} كالتالي :

- [مبتدئين بالصفر في ذهننا] نقوم بإضافة تكرار الفئة الأولى (= *10*) فنحصل على :

$$0 + 10 = 10$$

وهذه القيمة أقل من *11* فهي فئة المئين العاشر P_{10} [ومن ثم ليست فئة أي من القيم الأخرى] .

- نضيف للناتج السابق (أي نضيف لـ *10*) تكرار الفئة الثانية (= *30*) فنحصل على :

$$10 + 30 = 40$$

وهذه القيمة أكبر من 11 وأكبر من 27.5 ولكنها أقل من 99 , 82.5 , 55 [أنظر للجدول السابق] ، إذن الفئة **الثانية** هي فئة كلٍ من Q_1 , P_{10} .

- نضيف للناتج السابق (أي نضيف لـ 40) تكرار الفئة الثالثة (= 50) فنحصل على : $40 + 50 = 90$

وهذه القيمة أكبر من 55 وأكبر من 82.5 ولكنها أقل من 99 ، إذن الفئة **الثالثة** هي فئة كلٍ من Q_3 , M .

- نضيف للناتج السابق (أي نضيف لـ 90) تكرار الفئة الرابعة (= 20) فنحصل على :

$$90 + 20 = 110$$

وهذه القيمة أكبر من 99 ، إذن الفئة **الرابعة** هي فئة كلٍ من P_{90} .

من السابق نستنتج أن فئة القيم المختلفة أصبحت معروفة ، وبالتالي أصبح معروفاً تكرار فئة كل قيمة وطول تلك الفئة والتكرار المتجمع السابق لتلك الفئة ، إذن بالتعويض في العلاقة :

القيمة المطلوبة = الحد الأدنى لفئة هذه القيمة +

[(f_0 − التكرار المتجمع السابق) / تكرار فئة هذه القيمة × طول فئة هذه القيمة]

لاحظ أن العلاقة السابقة هي تعميم للعلاقات التي سبق واستخدمناها مع كلٍ من الوسيط والربيعات والمئينات جميعها

جدول (2) يبين جميع الكميات اللازمة لتحديد جميع القيم المطلوبة وهي :

جدول (2)						
القيمة المطلوب تحديدها	f_0	**فئة القيمة**	**طول الفئة**	**تكرار الفئة**	**حد الفئة الأدنى**	**التكرار المتجمع السابق**
المئين العاشر P_{10}	*11*	**الثانية**	*10*	*30*	*30*	*10*
الربيع الأول Q_1	*27.5*	**الثانية**	*10*	*30*	*30*	*10*
الوسيط (الربيع الثاني) M	*55*	**الثالثة**	*10*	*50*	*40*	*10 + 30 = 40*
الربيع الثالث Q_3	*82.5*	**الثالثة**	*10*	*50*	*40*	*40*

المئين التسعون P_{90}	99	الرابعة	10	20	50	10 + 30 + 50 = 90

$$P_{10} = 30 + \left[\frac{11-10}{30}\right] \times 10 = 30.33 \quad , \quad Q_1 = 30 + \left[\frac{27.5-10}{30}\right] \times 10 = 35.83$$

$$M = 40 + \left[\frac{55-40}{50}\right] \times 10 = 43 \quad , \quad Q_3 = 40 + \left[\frac{82.5-40}{50}\right] \times 10 = 48.5$$

$$P_{90} = 50 + \left[\frac{99-90}{20}\right] \times 10 = 54.5$$

وبالتالي يكون :

$$Q = \tfrac{1}{2}(Q_3 - Q_1) = \tfrac{1}{2}(48.5 - 35.83) = 6.335 = \text{الانحراف الربيعي}$$

$$P = P_{90} - P_{10} = 59.5 - 30.33 = 29.17 = \text{المدى المئيني}$$

(3) المنوال :

حيث أن الفئات متساوية الطول [طول كل منها = *10*] ، يكون المنوال هو (تقريباً) مركز الفئة المنوالية (الفئة المناظرة لأكبر تكرار) ، أي مركز الفئة الثالثة [أي تقريباً *45*] .

(4) معامل الالتواء :

فيما سبق حسبنا كلٍ من الوسط الحسابي ، الوسيط ، المنوال ، الانحراف المعياري ، الربيع الأول والربيع الثالث ، المئين العاشر والمئين التسعون ، ويمكن تجميع تلك النتائج السابقة في الجدول التالي :

الكمية	رمزها	قيمتها
الوسط الحسابي	$\bar{x}$	**42.27**
الوسيط	**M**	**43**
المنوال	$\hat{X}$	**44**
الانحراف المعياري	**s**	**8.62**
الربيع الأول	Q_1	**35.38**
الربيع الثالث	Q_3	**48.5**

المئين العاشر	P_{10}	**30.33**
المئين التسعون	P_{90}	**54.5**

وبالتالي يمكن استخدام أي من العلاقات *(5-4)* , *(5-3)* , *(5-2)* , *(5-1)* في حساب معامل الالتواء :

$$\xrightarrow{\text{using (5-1)}} \quad sk_1 = \frac{\overline{x} - \hat{X}}{s} = \frac{42.27 - 44}{8.62} \cong -0.2$$

$$\xrightarrow{\text{using (5-2)}} \quad sk_2 = \frac{3(\overline{x} - M)}{s} = \frac{3(42.27 - 43)}{8.62} \cong -0.25$$

$$\xrightarrow{\text{using (5-3)}} \quad sk_3 = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1} = \frac{48.5 - 2 \times 43 + 35.38}{48.5 - 35.38} \cong -0.16$$

$$\xrightarrow{\text{using (5-4)}} \quad sk_4 = \frac{P_{90} - 2P_{50} + P_{10}}{P_{90} - P_{10}} = \frac{54.5 - 2 \times 43 + 30.33}{54.5 - 30.33} \cong -0.05$$

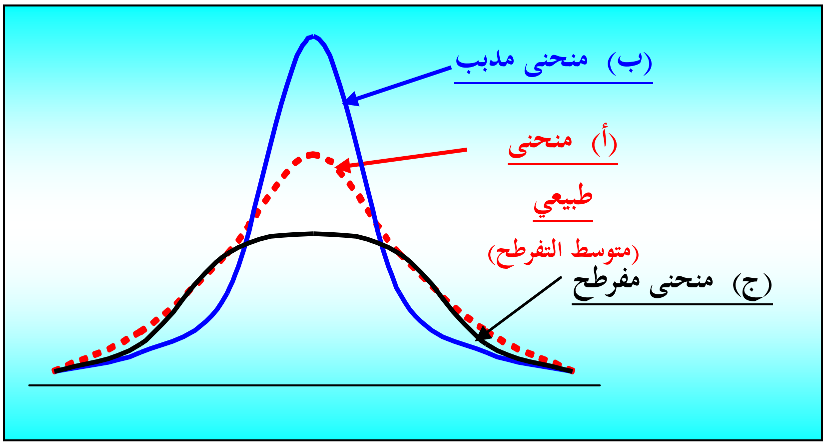
ملحوظة :

نتيجة وجود اختلاف فى الأصل الرياضى لكل من المعادلات *(5-4)* , *(5-3)* , *(5-2)* , *(5-1)* السابقة ، تختلف قيمة معامل الالتواء من معادلة لأخرى ، إلا أنه يفضل استخدام معامل الالتواء لبيرسون [المعادلات *(5-1)* , *(5-2)*] فى حالة **البيانات غير المبوبة** و كذلك **الجداول التكرارية المغلقة** ، أما فى حالة **الجداول التكرارية المفتوحة** فيفضل استخدام معامل الالتواء الربيعي [المعادلة *(5-3)*] أو معامل الالتواء المئيني [المعادلة *(6-4)*] .

(3-6) تعريف التفرطح

يُقصد بالتفرطح **درجة تدبب** (الارتفاع أو الانخفاض) في قمة المنحنى مقارنةً بقمة منحنى التوزيع الطبيعي الذي يُعد متوسط التفرطح :

- فإذا كانت قمة المنحنى أعلى من مثيلتها في التوزيع الطبيعي يُسمى المنحنى **مدبب** .
- وإذا كانت قمة المنحنى أدنى من مثيلتها في التوزيع الطبيعي يُسمى المنحنى **مفرطح** [تكون قمته مسطحة لحد ما] .
- أما إذا كانت القمة ليست مدببة أو مسطحة [أي قريبة من المنحنى الطبيعي] يُسمى المنحنى **متوسط التفرطح** .



ويُقاس تفرطح أي توزيع بعدة مقاييس ، أحد هذه المقاييس يعتمد على الربيعات والمئينات ويُسمى بـ **معامل التفرطح المئيني** ، وهذا المعامل يساوي (تقريباً) 0.263 في حالة التوزيع الطبيعي ، وبالتالي إذا كان معامل التفرطح لأي توزيع :

- أكبر من *0.263* كان التوزيع **مدبباً** .
- أقل من *0.263* كان التوزيع **مفرطحاً** .

ويتحدد معامل التفرطح المئيني لأي توزيع من العلاقة :

$$\boxed{\text{معامل التفرطح المئيني } ku = \frac{\text{نصف المدى الربيعي}}{\text{المدى المئيني}} = \frac{\text{الانحراف الربيعي}}{\text{المدى المئيني}}}$$

أي :

$$\boxed{ku = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})}} \qquad (5\text{-}5)$$

مثال (3–5) : حدد ما إذا التوزيع مفرطحاً أم مدبباً وذلك إذا كان للتوزيع :

(أ) ربيع أول *69* وربيع ثالث *91* ، ومئين عاشر *59* ، ومئين تسعون *94*

(ب) إنحراف ربيعي *20* ، ومدى مئيني *100* .

الحل :

(أ) باستخدام مباشر العلاقة *(5-5)* نجد أن :

$$ku = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})} = \frac{91 - 69}{2(94 - 59)} = \frac{22}{70} \cong 0.3 > 0.263$$

مما يعني أن المنحنى قريباً من التوزيع الطبيعي لكنه **مدبباً** قليلاً .

(ب) أيضاً باستخدام مباشر العلاقة *(5-5)* نجد أن :

$$ku = \frac{20}{100} = 0.2 < 0.263$$

مما يعني أن المنحنى قريباً من التوزيع الطبيعي لكنه **مفرطحاً** قليلاً .

مثال (5-4) : **حدد** معامل التفرطح لتوزيع مثال (5-2) مبيناً ما إذا كان التوزيع مدبباً أم مفرطحاً .

الحل :

في مثال (5-2) كانت النتائج (التي تهمنا هنا) كالتالي :

الكمية	قيمتها
الربيع الأول Q_1	***35.38***
الربيع الثالث Q_3	***48.5***

الكمية	قيمتها
المئين العاشر P_{10}	***30.33***
المئين التسعون P_{90}	***54.5***

وباستخدام مباشر للعلاقة *(5-5)* ، نحصل على :

$$ku = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})} = \frac{48.5 - 35.38}{2(54.5 - 30.33)} = \frac{13.12}{48.34} \cong 0.271 > 0.263$$

مما يعني أن المنحنى قريباً من التوزيع الطبيعي لكنه **مدبباً** قليلاً .

ملخص للدرس التاسع [الباب الخامس : الالتواء والتفرطح]

- يُعرف **الالتواء** على أنه درجة تماثل أو البعد عن التماثل لتوزيع ما .
- **مقاييس الالتواء :** هي مقاييس ترصد درجة تماثل أو البعد عن التماثل لتوزيع ما .
- وهناك عدة **مقاييس** لقياس الالتواء منها :

$$sk_1 = \frac{\overline{x} - \hat{X}}{s} \quad , \quad sk_2 = \frac{3(\overline{x} - M)}{s}$$

$$sk_3 = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1} \quad , \quad sk_4 = \frac{P_{90} - 2P_{50} + P_{10}}{P_{90} - P_{10}}$$

- وتحدد إشارة معامل الالتواء على نوع الالتواء ، **فالإشارة الموجبة** لمعامل الالتواء تعني **التواءً موجباً** [أي جهة **اليمين**] ، **والإشارة السالبة** تعني **التواءً سالباً** [أي جهة **اليسار**] .
- ويُقصد بالتفرطح **درجة تدبب** (الارتفاع أو الانخفاض) في قمة المنحنى مقارنةً بقمة منحنى التوزيع الطبيعي الذي يُعد متوسط التفرطح .
- **مقاييس التفرطح :** هي مقاييس ترصد درجة التدبب (الارتفاع أو الانخفاض) في قمة المنحنى مقارنةً بقمة منحنى التوزيع الطبيعي .
- وهناك عدة **مقاييس** لقياس التفرطح منها معامل التفرطح المئيني :

$$ku = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})}$$

الذي يساوي 0.263 حالة التوزيع الطبيعي ، وبالتالي إذا كان معامل التفرطح لأي توزيع أكبر من 0.263 كان التوزيع **مدبباً** ، وإذا كان أقل من 0.263 كان التوزيع **مفرطحاً** .

تدريب عملي

في كل حالة من الحالات التالية ، **احسب** معامل الالتواء ومعامل التفرطح

(أ)

x	60 –	70 –	80 –	90 – 100
f	98	75	56	42

(ب)

x	< 3	3 –	6 –	> 9
f	3	7	16	12

تدريبات (9)

الإجابة النهائية لجميع التمرينات موجودة في نهاية التدريب

اختر الإجابة الصحيحة

(1) **مقاييس الالتواء** هي

(أ) قيم نموذجية يمكن أن تمثل مجموعة البيانات

(ب) مقاييس ترصد الدرجة التي تتجه بها البيانات الكمية للانتشار حول قيمة متوسطة

(ج) مقاييس تحدد النسبة المئوية للتشتت المطلق بالنسبة لقيمة متوسطة

(د) هي مقاييس ترصد درجة تماثل أو البعد عن التماثل لتوزيع ما

(هـ) مقاييس ترصد درجة التدبب في قمة المنحنى مقارنة بقمة منحنى التوزيع الطبيعي

(2) **مقاييس التفرطح** هي

(أ) قيم نموذجية يمكن أن تمثل مجموعة البيانات

(ب) مقاييس ترصد الدرجة التي تتجه بها البيانات الكمية للانتشار حول قيمة متوسطة

(ج) مقاييس تحدد النسبة المئوية للتشتت المطلق بالنسبة لقيمة متوسطة

(د) هي مقاييس ترصد درجة تماثل أو البعد عن التماثل لتوزيع ما

(هـ) مقاييس ترصد درجة التدبب في قمة المنحنى مقارنة بقمة منحنى التوزيع الطبيعي

خاص بالأسئلة (3) ، (4) :

إذا كان [لمجموعة من القيم] Q_1 هو الربيع الأول ، Q_3 هو الربيع الثالث ، P_{10} هو المئين العاشر ، P_{90} هو المئين التسعون ، M هو الوسيط ، فإن :

(3) معامل الالتواء الربيعي لمجموعة القيم يساوي :

(أ) $\frac{Q_3 - 2M + Q_1}{Q_3 - Q_1}$ (ب) $\frac{P_{90} - 2M + P_{10}}{Q_3 - Q_1}$

(ج) $\frac{Q_3 - 2M + Q_1}{P_{90} - P_{10}}$ (د) $\frac{P_{90} - 2M + P_{10}}{P_{90} - P_{10}}$

(4) معامل التفرطح المئيني لمجموعة القيم يساوي :

(أ) $\frac{Q_3 - Q_1}{P_{90} + P_{10}}$ (ب) $\frac{P_{90} - P_{10}}{Q_3 - Q_1}$

(ج) $\dfrac{Q_3 - Q_1}{2(P_{90} - P_{10})}$ (د) $\dfrac{Q_3 - Q_1}{P_{90} - P_{10}}$

(5) لتحديد معامل بيرسون الأول للالتواء يلزم معرفة

(أ) الوسط والوسيط (ب) الوسط والمنوال

(ج) الربيعات Q_1 , Q_3 (د) المئينات P_{10} , P_{90}

(6) لتحديد معامل بيرسون الثاني للالتواء يلزم معرفة

(أ) الوسط والوسيط (ب) الوسط والمنوال

(ج) الربيعات Q_1 , Q_3 (د) المئينات P_{10} , P_{90}

(7) لتحديد معامل الالتواء الربيعي يلزم معرفة

(أ) الوسط والوسيط (ب) الوسط والمنوال

(ج) الربيعات Q_1 , Q_3 (د) المئينات P_{10} , P_{90}

(8) لتحديد معامل الالتواء المئيني يلزم معرفة

(أ) الوسط والوسيط (ب) الوسط والمنوال

(ج) الربيعات Q_1 , Q_3 (د) المئينات P_{10} , P_{90}

إجابة تدريبات (9) :

(1) د (2) هـ (3) أ (4) ج (5) ب

(6) أ (7) ج (8) د

عناصر الدرس

الباب السادس : تحليل الارتباط

- تمهيد
- الارتباط الخطي وشكل الانتشار
- قوة ونوع الارتباط الخطي
- معامل الارتباط البسيط
    - ❖ معامل الارتباط الخطي البسيط لبيرسون
    - ❖ معامل ارتباط الرتب لسبيرمان

الباب السادس : تحليل الارتباط

(6–1) تمهيد :

في دراستنا للفصول السابقة كنا نتعامل مع بيانات ذات متغير واحد [كنا نرمز له بالرمز x] ورأينا كيف نتعامل مع هذه البيانات من حيث :

- جمع البيانات .
- تنظيمها وعرضها [عن طريق جداول أو رسومات بيانية] .
- استخراج مقاييس خاصة بهذه البيانات مثل
  - ❖ مقاييس نزعة مركزية [الوسط الحسابي – الوسيط – المنوال – وغيرها]
  - ❖ مقاييس تشتت [المدى – الانحراف المتوسط – الانحراف المعياري – الانحراف الربيعي – المدى المئيني – وغيرها]
- مقاييس التواء .
- مقاييس تفرطح .

وكل ذلك من خلال القسم الأول من علم الإحصاء وهو "**علم الإحصاء الوصفي**" .

إلا أن اشتقاق بعض العلاقات بين المتغيرات التي تكون الظاهرة محل الدراسة يكون من الأهمية بمكان لمعرفة تطورات الظاهرة فى المستقبل و كيفية التأثير فيها من خلال التأثير فى المتغيرات المكونة لها ، لذلك فمن الأساليب الإحصائية المناسبة لتقييم العلاقات بين المتغيرات المختلفة هو "تحليل الارتباط Correlation Analysis" لمعرفة ما إذا كان هناك ارتباط بين ظاهرتين أو أكثر ، وأيضاً التنبؤ بأداء الظاهرة فى المستقبل وذلك من خلال ما يُسمى بـ "تحليل الانحدار Regression Analysis" .

وفي هذا الباب سنركز على "**تحليل الارتباط**" ، أما تحليل الانحدار فسوف سنتعرض له في الباب القادم [بإذن الله] ، أي أننا في هذا الفصل سنتعامل مع بيانات يمثلها متغير [ليكن x] وبيانات أخرى يمثلها متغير آخر [ليكن y] ونبحث في الآتي :

(1) هل هناك علاقة بين هاتين المجموعتين من البيانات أم لا : فإذا كانت هناك علاقة نقول أن المتغيرين x , y مرتبطان وإلا فهما غير مرتبطين .

(2) مدى قوة هذه العلاقة [إن وُجِدت] ، وهل هي قوية جداً أم قوية أم متوسطة أم ضعيفة أم ضعيفة جداً .

(3) نوع هذه العلاقة [إن وُجِدت] ، وهل هي طردية أم عكسية .

العلاقة الطردية : كلما زادت قيمة x زادت أيضاً قيمة y

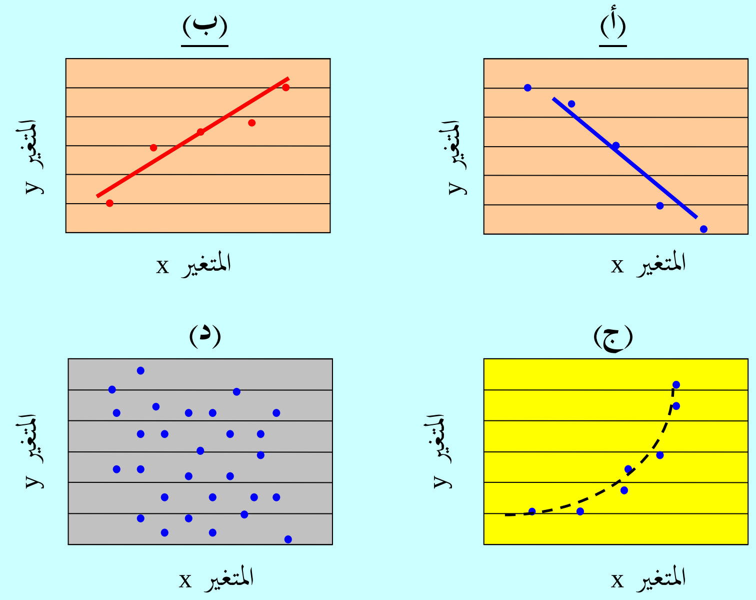
العلاقة العكسية : كلما زادت قيمة x نقصت قيمة y

وكأمثلة لهذا النوع من الدراسة إيجاد العلاقة بين :

- البيانات عن حجم الإعلانات عن منتج معين وحجم المبيعات ، فكلما زادت الإعلانات عن منتج معين زاد حجم المبيعات [ارتباط طردي] .
- البيانات عن الكمية المعروضة في السوق من منتج معين وسعر هذه السلعة ، فكلما زادت الكمية المعروضة في السوق من منتج معين قل سعر المنتج [ارتباط عكسي] .

(6-2) الارتباط الخطي وشكل الانتشار

نفرض أن لدينا بيانات $x_1 , x_2 ,\cdots, x_n$ تخص ظاهرة معينة X [مثل الطول x لمجموعة من الطلبة] ، ويناظرها بيانات $y_1 , y_2 ,\cdots, y_n$ تخص ظاهرة أخرى Y [مثل الوزن y لهؤلاء الطلبة] واخترنا محورين متعامدين : الأفقي (ويخص المتغير x) والرأسي (ويخص المتغير y) وقمنا بتوقيع النقاط $(x_1,y_1) , (x_2,y_2) , \cdots , (x_n,y_n)$ ، فإننا نحصل بذلك على ما يُسمى بـ "**شكل الانتشار**" لبيانات المتغيرين [شكل (6-1)] ، ومن شكل الانتشار يمكن بدايةً معرفة ما إذا كان هناك ارتباط بين الظاهرتين أم لا ، فإذا أمكن تمهيد (باليد) منحنى [يُسمى **بالمنحنى التقريبي**] بين النقاط الموقعة [كما في الحالات (أ) ، (ب) ، (ج) من شكل (6-1)] يمكن القول بأن هناك ارتباط بين الظاهرتين ، أما إذا لم نتمكن من تمهيد منحنى بين هذه النقاط فإننا نقول أنه ليس هناك ارتباط بين الظاهرتين [كما في الحالة (6-1د)] .



شكل (6–1) : شكل الانتشار

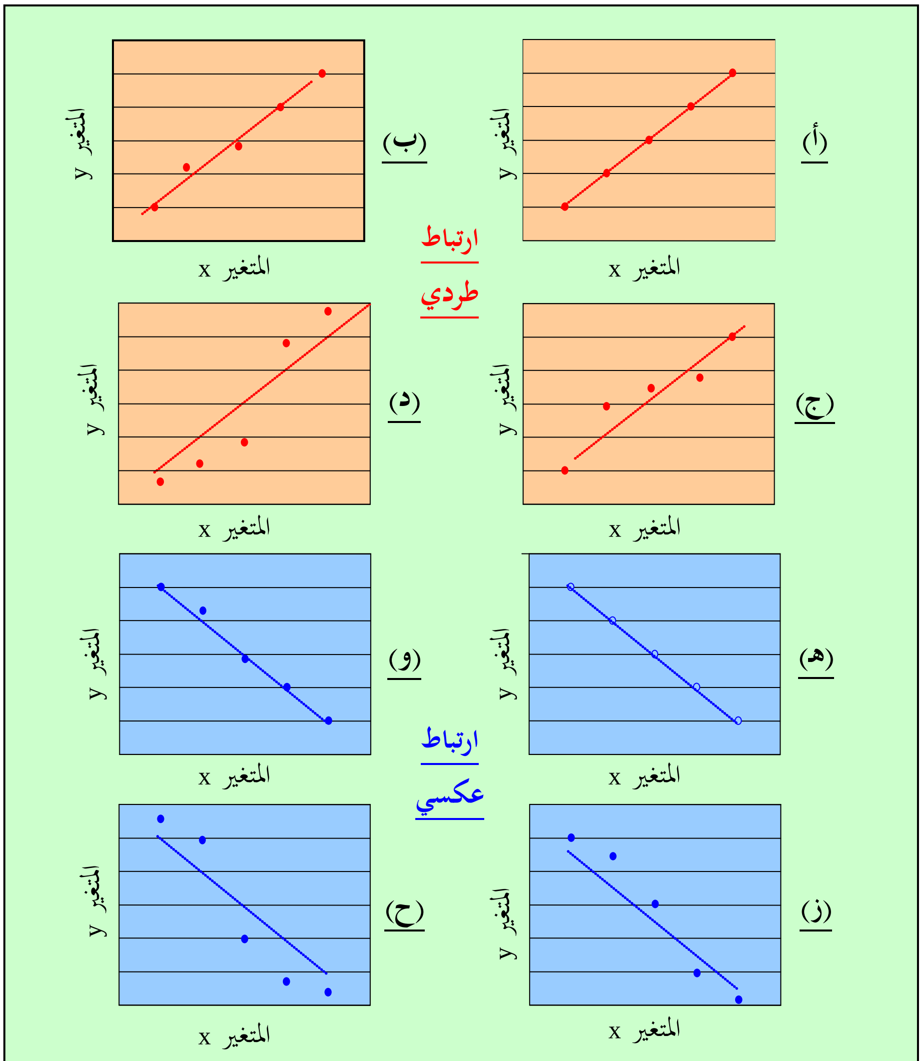
وإذا كانت الظاهرتان مرتبطتين [كما في الحالات (أ) ، (ب) ، (ج)] وكان المنحنى التقريبي الذي أمكن تمهيده **خطاً مستقيماً** [الحالات (أ) ، (ب)] سُمي الارتباط بين الظاهرتين **ارتباطاً خطياً** ، وإلا يكون الارتباط **غير خطي** [الحالة (ج)] .

أيضاً إذا كانت الظاهرتان مرتبطتين وكانت الظاهرة Y تتجه للزيادة كلما ازدادت الظاهرة X [الحالات (ب) ، (ج)] فإن الارتباط يُسمى ارتباطاً **موجباً** أو ارتباطاً **طردياً** ، أما إذا كانت الظاهرة Y تتجه للنقصان كلما ازدادت الظاهرة X [الحالة (أ)] فإن الارتباط يُسمى ارتباطاً **سالباً** أو ارتباطاً **عكسياً** .

وسوف نقصر دراستنا في هذا الفصل على **الارتباط الخطي** .

(6–3) قوة ونوع الارتباط الخطي

كما ذكرنا في البند السابق ، إذا أمكن تمهيد خط مستقيم بين نقاط شكل الانتشار ، سُمي الارتباط بين الظاهرتين X , Y بالارتباط الخطي [شكل (6–2)] .



شكل (7−2) : نوع وقوة الارتباط الخطي

نوع الارتباط :

- ❖ إذا كانت الظاهرة Y تتجه للزيادة كلما ازدادت الظاهرة X [الحالات من (أ) إلى (د)] فإن الارتباط يُسمى ارتباطاً **موجباً** أو ارتباطاً **طردياً** ،
- ❖ أما إذا كانت الظاهرة Y تتجه للنقصان كلما ازدادت الظاهرة X [الحالات من (هـ) على (ح)] فإن الارتباط يُسمى ارتباطاً **سالباً** أو ارتباطاً **عكسياً** .

قوة الارتباط :

- ❖ إذا أمكن رسم خط مستقيم يمر **بجميع** نقاط شكل الانتشار سُمي الارتباط "**ارتباط تام**" [طردي مثل الحالة (أ) أو عكسي مثل الحالة (هـ)] .
- ❖ وإذا أمكن رسم خط مستقيم بحيث تكون انحرافات النقاط عنه ضعيفة جداً ، سُمي الارتباط "**ارتباط قوي**" [طردي كالحالة (ب) أو عكسي كالحالة (و)] .

- أما إذا زادت الانحرافات عن الخط المستقيم ولكن بشكل معقول ، سُمي الارتباط ''**ارتباط متوسط**'' [طردي كالحالة (ج) أو عكسي كالحالة (ز)] .
- وإذا زادت الانحرافات عن الخط المستقيم بشكل كبير إلى حد ما ، سُمي الارتباط ''**ارتباط ضعيف**'' [طردي كالحالة (د) أو عكسي كالحالة (ح)] .

(6–4) معامل الارتباط البسيط

يستخدم **معامل الارتباط البسيط** Correlation coefficient فى تحديد ما إذا كان هناك علاقة بين المتغيرين ، وكذلك تحديد نوع وقوة العلاقة إن وجدت ، أما فى حالة دراسة مدى وجود علاقة ارتباطية بين أكثر من متغيرين فإنه يتم الاعتماد على **معامل الارتباط المتعدد** Multiple Correlation coefficient و/أو **معامل الارتباط الجزئي** Partial Correlation Coefficient حيث يتم دراسة تأثير أحد المتغيرات مع تثبيت المتغيرات الأخرى [كما في حالة الكثير من المشكلات الاقتصادية حيث يتم دراسة تأثير السعر على الكمية المطلوبة بفرض ثبات الجودة ومستوى الذوق] ، ويتم استخدام معامل الارتباط فى الحكم على نوع العلاقة بين المتغيرين [حيث تكون علاقة طرديه أو عكسية] ، وكذلك بالنسبة لقوه العلاقة [فقد تكون علاقة قويه ، أو متوسطة أو ضعيفة] .

و عادة ما يتم تقسيم المتغيرات محل الدراسة إلى :

- **متغيرات مستقلة Independent Variables**

وهى المتغيرات التى بتغير قيمتها تؤثر فى تغيير قيمة متغير أو متغيرات أخرى ، أى هى المتغيرات التى تتغير أولاً ، و سنرمز للمتغير المستقل بالرمز x .

- **متغيرات تابعة Dependent Variables**

وهى المتغيرات التى تتغير قيمتها بتغير المتغيرات المستقلة أو إحداها ، أى هى المتغيرات التى تتغير تالية للمتغيرات المستقلة ، وسنرمز للمتغير التابع بالرمز y .

وسيتم قياس الارتباط البسيط من خلال كلٍ من :

- معامل الارتباط الخطى البسيط لبيرسون Person's Correlation Coefficient
- معامل ارتباط الرتب لسبيرمان Spearman's Rank Correlation Coefficient

(6–4–1) معامل الارتباط الخطى البسيط لبيرسون *Person's Correlation Coefficient* :

يعتبر معامل الارتباط الخطى البسيط لبيرسون Person's Correlation Coefficient (والذى سنرمز له بالرمز r_p) من أكثر الأدوات الإحصائية استخداماً فى تحديد قوة العلاقة بين متغيرين ، كما يستعمل لتحديد مدى

وجود علاقة خطية بين متغيرين ، وهناك أكثر من صيغة يمكن الاعتماد عليها فى حساب معامل الارتباط الخطى البسيط لبيرسون r_p كما يلى :

$$\boxed{r_p = \frac{Cov(x,y)}{s_x s_y}} \qquad (6\text{-}1)$$

حيث :

$$Cov(x,y) = \frac{\sum (x-\bar{x})(y-\bar{y})}{n} = x,y \text{ التغاير بين المتغيرين}$$

$$s_x = \sqrt{\frac{\sum (x-\bar{x})^2}{n}} = x \text{ الانحراف المعياري للمتغير}$$

$$s_y = \sqrt{\frac{\sum (y-\bar{y})^2}{n}} = y \text{ الانحراف المعياري للمتغير}$$

$\bar{x}$ هو الوسط الحسابي للمتغير x ، $\bar{y}$ هو الوسط الحسابي للمتغير y .

ويمكن إعادة كتابة المعادلة *(6-1)* لتأخذ الشكل :

$$\boxed{r_p = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2}\sqrt{\sum (y-\bar{y})^2}}} \qquad (6\text{-}2)$$

كما يمكن تبسيط المعادلة (6-2) لتأخذ الصورة :

$$\boxed{r_p = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{n\sum x^2 - \left(\sum x\right)^2}\sqrt{n\sum y^2 - \left(\sum y\right)^2}}} \qquad (6\text{-}3)$$

وتتراوح قيمة معامل الارتباط بين *1 , -1* [أي أن $-1 \le r_p \le 1$] ، بتعبير آخر فإن قيمة معامل الاتباط دائما كسر قيمتة أقل من الواحد الصحيح ، وطبقاً لتلك القيمة تكون قوة العلاقة على النحو التالي :

- أكبر صفر إلى أقل من *0.3* فتكون العلاقة **ضعيفة** .
- من *0.3* إلى أقل من *0.7* تكون العلاقة **متوسطة** .
- من *0.7* إلى الواحد الصحيح تكون العلاقة **قوية** .
- قيمة معامل الارتباط الخطي = *1* [الواحد الصحيح] فهذا يعني أن المتغيرين مرتبطان ارتباطاً تاماً .
- وإذا كانت قيمة معامل الارتباط تساوى صفراً فهذا يعني **عدم وجود علاقة خطيه أو ارتباط** بين المتغيرين x , y ، أى يكون المتغيران مستقلين عن بعضهما البعض .
- أما إذا كانت قيمة معامل الارتباط أكبر من الواحد الصحيح فهذا لا يعني إلا "**وجود خطأ في الحسابات**" .

هذا بالنسبة لقوة العلاقة ، أما نوع العلاقة (طردية أو عكسية) فتعتمد على إشارة معامل الارتباط ، فإذا كانت الإشارة :

- **موجبة** فإن العلاقة تكون **طردية** .
- **سالبة** فإن العلاقة تكون **عكسية** .

فمثلاً

معامل بيرسون للارتباط r_p	المعنى
0.91	ارتباط طردى قوى جدآ
-0.87	ارتباط عكسى قوى
-0.21	ارتبلط عكسى ضعيف
0.43	ارتباط طردى متوسط
1	ارتباط طردى تام
-0.51	ارتباط عكسى متوسط

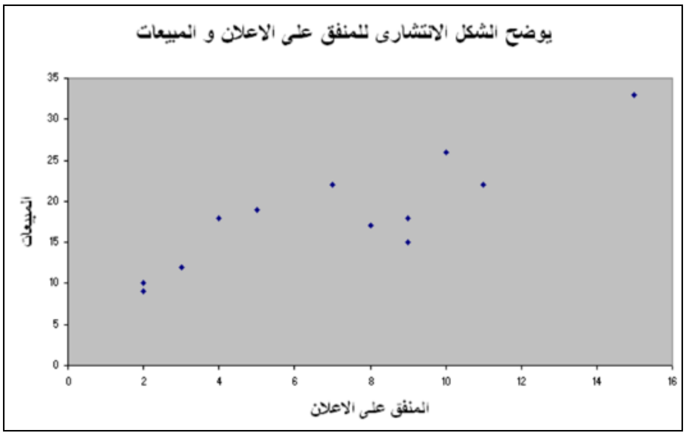
مثال (6–1) : فيما يلى بيان بالمنفق على الاعلان *(x)* [بالمليون ريال] والمبيعات *(y)* [أيضاً بالمليون ريال] :

8	9	11	4	15	10	5	6	7	2	3	2	*x*
17	15	22	18	33	26	19	18	22	9	12	10	*y*

المطلوب : رسم شكل الانتشار موضحاً العلاقة بين المنفق على الاعلان والمبيعات ، **وأيضاً حساب** معامل الارتباط الخطى البسيط (معامل بيرسون للارتباط) ، **مع التعليق** .

الحل :

الشكل المرافق يوضح شكل الانتشار بين المنفق على الاعلان و المبيعات



ومن شكل الانتشار نستنتج أن قيم كلٍ من المنفق على الاعلان والمبيعات يأخذ [بوجهٍ عام] اتجاه تصاعدى جهة اليمين مما يدل على وجود علاقة طرديه بينهما .

و اذا اردنا استخدام المعادلة (*6-2*) فى حساب معامل الارتباط بين المنفق على الاعلان والمبيعات لابد أولاً من حساب الوسط الحسابى $\bar{x}$ للمنفق على الاعلان والوسط الحسابى $\bar{y}$ للمبيعات حيث :

$$\bar{x} = \frac{\sum x}{n} = \frac{82}{12} = 7.08333$$

$$\bar{y} = \frac{\sum y}{n} = \frac{221}{12} = 18.41667$$

ومنهما يكون معامل ارتباط بيرسون r_p هو :

$$r_p = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2}\sqrt{\sum (y-\bar{y})^2}}$$

و على ذلك يمكن لنا أعداد الجدول التالي :

x	y	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$
2	10	-4.83333	-8.41667	40.68056	23.36111	70.84028
3	12	-3.83333	-6.41667	24.59722	14.69444	41.17361
2	9	-4.83333	-9.41667	45.51389	23.36111	88.67361
7	22	0.16666	3.58333	0.597222	0.027778	12.84028
6	18	-0.83333	-0.41667	0.347222	0.694444	0.173611
5	19	-1.83333	0.58333	-1.06944	3.361111	0.340278
10	26	3.16666	7.58333	24.01389	10.02778	57.50694
15	33	8.16666	14.5833	119.0972	66.69444	212.6736
4	18	-2.83333	-0.41667	1.180556	8.027778	0.173611
11	22	4.16666	3.58333	14.93056	17.36111	12.84028
9	15	2.16666	-3.41667	-7.40278	4.694444	11.67361
8	17	1.16666	-1.41667	-1.65278	1.361111	2.006944
82	221	0	0	260.8333	173.6667	510.9167
$\sum x$	$\sum y$	$\sum (x-\bar{x})$	$\sum (y-\bar{y})$	$\sum (x-\bar{x})(y-\bar{y})$	$\sum (x-\bar{x})^2$	$\sum (y-\bar{y})^2$

وبالتالي يكون :

$$r_p = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2}\sqrt{\sum (y-\bar{y})^2}} = \frac{260.8333}{\sqrt{173.667}\sqrt{510.9167}} = 0.8756$$

وتدل قيمة معامل الارتباط على وجود **علاقة قوية وطردية** بين المنفق علي الإعلان والمبيعات .

كما يمكن حساب معامل الارتباط من خلال المعادلة (6-3) حتى تكون هناك سهولة فى حساب معامل الارتباط التى تكون بالصورة التالية :

$$r_p = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{n\sum x^2 - \left(\sum x\right)^2}\sqrt{n\sum y^2 - \left(\sum y\right)^2}}$$

و حتى يمكن تطبيق هذه المعادلة على بيانات المثال السابق لحساب معامل الاتباط بين المنفق على الاعلان و المبيعات يتم تكوين الجدول التالى :

y^2	x^2	xy	y	x
100	4	20	10	2
144	9	36	12	3
81	4	18	9	2
484	49	154	22	7
324	36	108	18	6
361	25	95	19	5
676	100	260	26	10
1089	225	495	33	15
324	16	72	18	4
484	121	242	22	11
225	81	135	15	9
289	64	136	17	8
4581	734	1771	221	82
$\sum y^2$	$\sum x^2$	$\sum xy$	$\sum y$	$\sum x$

و بالتالى يمكن تطبيق المعادلة (6-3) كما يلى :

$$r_p = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{n\sum x^2 - \left(\sum x\right)^2}\sqrt{n\sum y^2 - \left(\sum y\right)^2}}$$

$$= \frac{12(1771) - (82)(221)}{\sqrt{12(734) - (82)^2}\sqrt{12(4581) - (221)^2}} = 0.8756$$

وهى نفس النتيجة التى حصلنا عليها بتطبيق المعادلة (7-2) مما يدل علي وجود علاقة طردية وقوية بين المنفق على الإعلان والمبيعات .

ومن أهم خصائص معامل الارتباط الخطى البسيط لبيرسون **أنه لا يعتمد على قيم المتغيران نفسها عند حساب قيمته وإنما يعتمد على مقدار التباعد بين هذه القيم وبعضها البعض** ، لذلك لا يتأثر معامل الارتباط

الخطى البسيط بأى عمليات جبرية يتم إجراءها على بيانات اى من المتغيرين أو أحدهما من جمع أو طرح أو ضرب أو قسمة ، كما يتضح من المثال التالي :

مثال (8-2) : فى بيانات المثال السابق ،إذا اكتشفت إدارة الشركة أن البيانات تم تجميعها وحسابها بطريقة خاطئة حيث يجب إضافة 5 مليون ﷼ إلى جميع قيم المنفق على الإعلان ، كما أن المبيعات يجب مضاعفة قيمتها لجميع القيم ، **احسب** معامل الارتباط فى هذه الحالة بين المنفق على الإعلان والمبيعات .

الحل :

يتم اولاً تعديل البيانات لكل من المنفق على الأعلان والمبيعات لتكون النتائج كما يلى :

x	y	xy	x^2	y^2
7	20	140	49	400
8	24	192	64	576
7	18	126	49	324
12	44	528	144	1936
11	36	396	121	1296
10	38	380	100	1444
15	52	780	225	2704
20	66	1320	400	4356
9	36	324	81	1296
16	44	704	256	1936
14	30	420	196	900
13	34	442	169	1156
142	**442**	**5752**	**1854**	**18324**

$\sum x \quad \sum y \quad \sum xy \quad \sum x^2 \quad \sum y^2$

وبالتالى يمكن تطبيق المعادلة (3-6) كما يلى :

$$r_p = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{n\sum x^2 - \left(\sum x\right)^2}\sqrt{n\sum y^2 - \left(\sum y\right)^2}}$$

$$= \frac{12(15752) - (142)(442)}{\sqrt{12(1754) - (142)^2}\sqrt{12(18324) - (442)^2}} = 0.8756$$

وهى نفس النتيجة التى حصلنا عليها سابقاً مما يدل على أن معامل الأرتباط لم تتأثر قيمته بالعمليات الجبرية من جمع (5 مليون) أو الضرب (× 2). وبالمثل لايتأثر بالطرح أو القسمة .

ملحوظة هامة : معامل التحديد Determination Coefficient

هو مربع معامل الارتباط لذلك يرمز له بالرمز r^2 وهو يشير إلى نسبة تفسير المتغير أو المتغيرات المستقلة للتغير فى المتغير التابع ، فمثلاً نجد أن $r^2 = 0.8756^2 = 0.76675$ ، مما يعني أن المنفق

على الاعلان يفسر نسبة 76.675% من التغير فى قيمة المبيعات ، بينما 23.32% من التغير فى المبيعات ترجع إلى عوامل أخرى منها الخطاء العشوائى .

(6-4-2) معامل ارتباط الرتب لسبيرمان

Spearman's Rank Correlation Coefficient

نلاحظ مما سبق أن معامل الارتباط لبيرسون لا يمكن استخدامه فى حساب قوة العلاقة بين متغيرين إلا إذا كانت البيانات المتوافره عنهما فى صورة كمية فقط ، أما إذا كانت فى صورة وصفية فلا يمكن تطبيقة في حساب الارتباط بين المتغيرين محل الدراسة ، ونتيجة أنه فى العديد من الدراسات يوجد متغيرات وصفية يكون مطلوباً دراسة الارتباط بينها ، نشأت الحاجة إلى مقياس يعمل على حساب الارتباط فى تلك الحاجة ، لذا ظهر لنا **معامل الارتباط الرتب لسبيرمان** r_s والذى يمكن استخدامه فى قياس الارتباط خاصة فى حالة البيانات الوصفية الترتيبية مثل تقديرات الطلاب : ممتاز/جيدجداً/جيد/مقبول/ ضعيف ، وكذلك قوة المركز المالى : جيد/متوسط/ضعيف ، ودرجة الموافقة على الرأى فى أسئلة الاستبانة : موافق تماماً/موافق/محايد/غير موافق/غير موافق على الاطلاق .

و يتم حساب معامل الارتباط الرتب لسبيرمان r_s باستخدام المعادلة التالية :

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} \qquad (6\text{-}4)$$

حيث d هو **الفرق** (في **الرتب**) بين المتغيرين ، n هو عدد المشاهدات .

ملاحظات يجب مراعاتها عند ترتيب المتغيرات:

- يتم ترتيب قيم مشاهدات المتغير x [وتسمى القيم الترتيبية للمتغير x] وكذلك الامر للمتغير y ، ويكون الترتيب تصاعدياً أو تنازلياً ولكن أهم شيء هو أنه إذا كان ترتيب x تصاعدياً لابد أن يكون ترتيب y أيضاً تصاعدياً ، والعكس صحيح .
- فى حالة الترتيب التصاعدى [مثلاً] يتم إعطاء أقل قيمة الرتبة *1* و القيمة التى هى أكبر منها الرتبة *2* ، .... ، وهكذا .
- فى حالة تكرار أو تساوى بعض القيم لأى متغيير تعطى كل منهم رتبة كما لو كانت القيم غير متساوية ثم نحسب الوسط الحسابى (مجموع الرتب ÷ عددها) لتلك الرتب و يعطى الوسط الحسابى كرتبة تلك القيم المتساوية .

مثال (6-3) :

فى المثال السابق **احسب** معامل الارتباط لسبيرمان بين المنفق على الاعلان و المبيعات ؟

الحل :

يتم أولاً ترتيب قيم كلٍ من x و y كما يتضح من الجدول التالى :

المنفق x	المبيعات y	رتب x		رتب y		d	d^2
2	10	~~1~~	1.5	~~2~~	2	-0.5	0.25
3	12	~~3~~	3	~~3~~	3	0	0
2	9	~~2~~	1.5	~~1~~	1	0.5	0.25
7	22	~~6~~	6	~~10~~	9.5	-3.5	12.25
9	18	~~8~~	8.5	~~7~~	6.5	2	4
5	19	~~5~~	5	~~8~~	8	-3	9
10	26	~~10~~	10	~~11~~	11	-1	1
15	33	~~12~~	12	~~12~~	12	0	0
4	18	~~4~~	4	~~6~~	6.5	-2.5	6.25
11	22	~~11~~	11	~~9~~	9.5	1.5	2.25
9	15	~~9~~	8.5	~~4~~	4	4.5	20.25
8	17	~~7~~	7	~~5~~	5	2	4
الأرقام المشطوبة تمثل الترتيب المبدئي وليس النهائي (والذي يمثل الرتب)						0	**59.5**
						$\sum d$	$\sum d^2$

نلاحظ من ذلك الجدول أنه :

- تم ترتيب المتغيران تصاعدياً .
- عند ترتيب قيم المتغير المنفق على الاعلان x نجد ان القيمة *2* تكررت مرتان لتأخذ الرتب *1* و *2* لذلك نحسب المتوسط لهما وهو $\frac{1+2}{2}=1.5$ ، لذلك وضعنا أمام القيمة *2* الرتبة ***1.5*** ، وكذلك الامر بالنسبة للقيمة *9* فأنها تأخذ الرتبة *8* و *9* لذلك وضعنا أمام القيمة *9* الرتبة ***8.5*** .
- عند ترتيب قيم المتغير y "المبيعات" نجد أن القيمة *18* تأخذ الرتبة *6* و *7* لذلك وضعنا أمام القيمة *18* الرتبة ***6.5*** ، وكذلك القيمة *22* تأخذ الرتبة *9* و *10* ، لذلك وضعنا أمامها الرتبة ***9.5*** .
- عند حساب الفرق d بين رتب المتغير x و رتب المتغير y ، نلاحظ أن مجموع الفروق d لابد أن يكون صفراً و الا يكون هناك خطاء فى الترتيب لأحد المتغيرين أو كلاهما ولابد مراجعة الترتيب مرة أخرى .

وحيث أن عدد المشاهدات $n=12$ فإنه يمكن حساب معامل ارتباط الرتب لسبيرمان كما يلي :

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(59.5)}{12(144-1)} = 0.7919$$

أي أن معامل ارتباط الرتب لسبيرمان *0.7919* مما يدل على وجود ارتباط طردى قوى بين المنفق على الاعلان و المبيعات ، وهى قيمة قريبة من التى تم حسابها باستخدام معامل الارتباط لسبيرمان التي بلغت *0.8756*

مثال (4-6) :

البيانات التالية تمثل التقديرات التى حصل عليها عشر طلاب فى مقررى المحاسبة x والقانون y :

x	*A*	*B*	*C*	*D*	*F*	*C*	*D*	*B*	*C*	*D*
y	*C*	*C*	*D*	*C*	*B*	*B*	*A*	*D*	*C*	*B*

حيث A تعني "ممتاز" ، B "جيد جداً" ، C "جيد" ، D "مقبول" ، F "راسب"

المطلوب حساب معامل الارتباط المناسب ؟

الحل :

هنا لا يمكن حساب معامل بيرسون للارتباط لأن البيانات وصفية وليست كمية ، لذا يكون معامل الاتباط الذي يمكن حسابه هو معامل ارتباط الرتب "معامل سبيرمان " r_s ، لذا يتم ترتيب المشاهدات و حساب الفروق بين الرتب و مربعاتها كما يتضح من الجدول التالى :

d^2	d	رتب y		رتب x		القانون y	المحاسبة x
30.25	*5.5*	*4.5*	~~*3*~~	*10*	~~*10*~~	C	A
16	*4*	*4.5*	~~*4*~~	*8.5*	~~*8*~~	C	B
20.25	*4.5*	*1.5*	~~*1*~~	*6*	~~*5*~~	D	C
2.25	*-1.5*	*4.5*	~~*5*~~	*3*	~~*2*~~	C	D
49	*-7*	*8*	~~*7*~~	*1*	~~*1*~~	B	F
4	*-2*	*8*	~~*8*~~	*6*	~~*6*~~	B	C
49	*-7*	*10*	~~*10*~~	*3*	~~*3*~~	A	D
49	*7*	*1.5*	~~*2*~~	*8.5*	~~*9*~~	D	B
2.25	*1.5*	*4.5*	~~*6*~~	*6*	~~*7*~~	C	C
25	*-5*	*8*	~~*9*~~	*3*	~~*4*~~	B	D
247	*0*						

الأرقام المشطوبة تمثل الترتيب المبدئي وليس النهائي (والذي يمثل الرتب)

- نلاحظ عند ترتيب تقديرات مقرر المحاسبة x أن التقدير "D" اخذ الرتب *2* و *3* و *4* ، لذلك تم وضع ***3*** [الوسط الحسابي لـ *4* , *3* , *2*]أمام التقدير D فى مقرر المحاسبة .
- كما أن تقدير C فى مقرر القانون y أخذ الرتب *3* و *4* و *5* و *6* ، لذلك تم وضع الرتبة ***4.5*** [الوسط الحساي لـ *6* , *5* , *4* , *3*]أمام التقدير C فى مقرر القانون .
- من الجدول السابق يتضح لنا أن مجموع الفروق d لابد أن يكون صفراً .

وحيث أن عدد المشاهدات $n = 10$ ، $\sum d^2 = 247$ يكون معامل ارتباط الرتب r_s مساوياً لـ :

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(247)}{10(100-1)} = -0.4969$$

ونلاحظ أن معامل ارتباط الرتب لسبيرمان بلغ *-0.4969* مما يدل على وجود ارتباط **عكسى متوسط** بين تقديرات مقرر المحاسبة وتقديرات مقرر القانون .

ملخص للدرس العاشر [الباب السادس : تحليل الارنباط]

- المقصود بـ "**تحليل الارتباط**" هو تحديد وجود علاقة بين ظاهرتين [يمثلهما متغيرين x , y] من عدم وجودها وتحديد كلٍ من نوع هذه العلاقة وقوتها [إن وُجِدت] ، ويمكننا ذلك عن طريق "**شكل الانتشار**" أو **معامل الارتباط** .

- **معامل الارتباط لبيرسون :**

$$\boxed{\begin{aligned} r_p &= \frac{Cov(x,y)}{s_x s_y} \\ &= \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2}\sqrt{\sum(y-\bar{y})^2}} \\ &= \frac{n\sum xy-\left(\sum x\right)\left(\sum y\right)}{\sqrt{n\sum x^2-\left(\sum x\right)^2}\sqrt{n\sum y^2-\left(\sum y\right)^2}} \end{aligned}}$$

وهو لا يصلح للاستخدام إلا إذا كانت البيانات المتوفرة للمتغيرين بيانات في صورة كمية فقط .

- **معامل الارتباط لسبيرمان [معامل ارتباط الرتب] :**

$$\boxed{r = 1-\frac{6\sum d^2}{n(n^2-1)}}$$

[حيث d هو الفرق في الرتب بين قيم x , y] ، وهذا المعامل يصلح استخدامه عندما تكون البيانات المتوفرة كمية أو وصفية .

- **معامل التحديد** هو مربع معامل الارتباط وهو يشير إلى نسبة تفسير المتغير أو المتغيرات المستقلة للتغير فى المتغير التابع .

- وتنحصر قيمة معامل الارتباط بين *+1* , *-1* ، فإذا كانت قيمته :
  - ❖ **موجبة** ، دل ذلك على أن هناك ارتباط **طردي** بين المتغيرين x , y
  - ❖ **سالبة** ، دل ذلك على أن هناك ارتباط **عكسي** بين المتغيرين x , y
  - ❖ **صفراً** ، دل ذلك على أنه ليس هناك ارتباط بين المتغيرين x , y

أما قوة الارتباط فتحددها القيمة المطلقة لمعامل الارتباط طبقاً للجدول التالي :

القيمة المطلقة لمعامل الارتباط	قوة الارتباط
0	لا يوجد ارتباط
$0 < r < 0.3$	ارتباط ضعيف
$0.3 \leq r < 0.7$	ارتباط متوسط
$0.7 \leq r < 1$	ارتباط قوي
1	ارتباط تام
> 1	خطأ غي الحسابات

تدريب عملي

في كل حالة من الحالات التالية ، **احسب** معامل الارتباط بين كلٍ من x , y :

(أ)

x	2	5	4	2	6	7
y	12	17	15	13	18	22

(ب)

	التقدير							
الإحصاء *x*	C	C	F	D	F	C	D	B
مبادىء علم الاجتماع *y*	A	C	D	C	F	C	A	D

تدريبات (10)

الإجابة النهائية لجميع التمرينات موجودة في نهاية التدريب

اختر الإجابة الصحيحة

(1) إذا كان معامل الارتباط r بين المتغيرين x , y يساوي 0.45 فهذا يعني أن x , y :

(أ) مرتبطان ارتباطاً عكسياً متوسطاً
(ب) مرتبطان ارتباطاً طردياً قوياً
(ج) غير مرتبطين
(د) مرتبطان ارتباطاً طردياً متوسطاً

(2) إذا كان معامل الارتباط r بين المتغيرين x , y يساوي 0.84 فهذا يعني أن x , y :

(أ) مرتبطان ارتباطاً عكسياً متوسطاً
(ب) مرتبطان ارتباطاً طردياً قوياً
(ج) غير مرتبطين
(د) مرتبطان ارتباطاً طردياً متوسطاً

(3) إذا كان معامل الارتباط r بين المتغيرين x , y يساوي -0.92 فهذا يعني أن x , y :

(أ) مرتبطان ارتباطاً عكسياً قوياً
(ب) مرتبطان ارتباطاً طردياً قوياً
(ج) مرتبطان ارتباطاً عكسياً تاماً
(د) مرتبطان ارتباطاً طردياً متوسطاً

(4) إذا كان معامل الارتباط r بين المتغيرين x , y يساوي -0.22 فهذا يعني أن x , y :

(أ) مرتبطان ارتباطاً عكسياً قوياً
(ب) مرتبطان ارتباطاً عكسياً متوسطاًً
(ج) مرتبطان ارتباطاً عكسياً تاماً
(د) مرتبطان ارتباطاً عكسياً ضعيفاً

(5) إذا كان معامل الارتباط r بين المتغيرين x , y يساوي -1 فهذا يعني أن x , y :

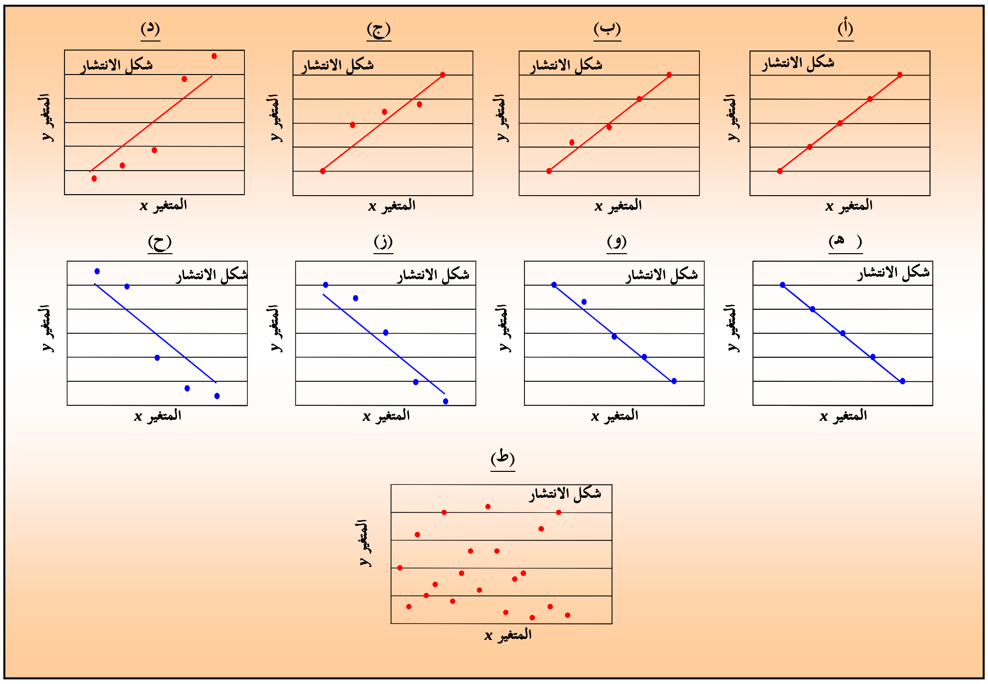
(أ) مرتبطان ارتباطاً عكسياً قوياً
(ب) مرتبطان ارتباطاً عكسياً متوسطاًً
(ج) مرتبطان ارتباطاً عكسياً تاماً
(د) مرتبطان ارتباطاً عكسياً ضعيفاً

(6) إذا كان معامل الارتباط r بين المتغيرين x , y يساوي -2 فهذا يعني أن x , y :

(أ) مرتبطان ارتباطاً عكسياً قوياً
(ب) مرتبطان ارتباطاً طردياً قوياً
(ج) مرتبطان ارتباطاً عكسياً تاماً
(د) هناك خطأ في الحسابات

خاص بالأسئلة من (7) إلى (15) :

في الشكل المرافق ،



(7) في شكل (أ) ، شكل الانتشار المعطى يوضح أن المتغيرين x , y :

(أ) مرتبطان عكسياً ارتباطاً قوياً (ب) مرتبطان طردياً ارتباطاُ قوياً

(ج) غير مرتبطين (د) مرتبطان ارتباطاً طردياً تاماً

(8) في شكل (ب) ، شكل الانتشار المعطى يوضح أن المتغيرين x , y :

(أ) مرتبطان عكسياً ارتباطاً قوياً (ب) مرتبطان طردياً ارتباطاُ قوياً

(ج) غير مرتبطين (د) مرتبطان ارتباطاً طردياً تاماً

(9) في شكل (ج) ، شكل الانتشار المعطى يوضح أن المتغيرين x , y :

(أ) مرتبطان عكسياً ارتباطاً متوسطاً (ب) مرتبطان طردياً ارتباطاُ قوياً

(ج) غير مرتبطين (د) مرتبطان ارتباطاً طردياً متوسطاً

(10) في شكل (د) ، شكل الانتشار المعطى يوضح أن المتغيرين x , y :

(أ) مرتبطان عكسياً ارتباطاً متوسطاً (ب) مرتبطان طردياً ارتباطاُ ضعيفاً

(ج) غير مرتبطين (د) مرتبطان ارتباطاً طردياً متوسطاً

(11) في شكل (هـ) ، شكل الانتشار المعطى يوضح أن المتغيرين x , y :

(أ) مرتبطان عكسياً ارتباطاً قوياً (ب) مرتبطان عكسياً ارتباطاُ تاماً

(ج) غير مرتبطين (د) مرتبطان ارتباطاً عكسياً ضعيفاً

(12) في شكل (و) ، شكل الانتشار المعطى يوضح أن المتغيرين x , y :

(أ) مرتبطان عكسياً ارتباطاً متوسطاً (ب) مرتبطان عكسياً ارتباطاً قوياً

(ج) غير مرتبطين (د) مرتبطان عكسياً ارتباطاً ضعيفاً

(13) في شكل (ز) ، شكل الانتشار المعطى يوضح أن المتغيرين x , y :

(أ) مرتبطان عكسياً ارتباطاً متوسطاً (ب) مرتبطان عكسياً ارتباطاً قوياً

(ج) غير مرتبطين (د) مرتبطان عكسياً ارتباطاً ضعيفاً

(14) في شكل (ح) ، شكل الانتشار المعطى يوضح أن المتغيرين x , y :

(أ) مرتبطان عكسياً ارتباطاً متوسطاً (ب) مرتبطان عكسياً ارتباطاً قوياً

(ج) غير مرتبطين (د) مرتبطان عكسياً ارتباطاً ضعيفاً

(15) في شكل (ط) ، شكل الانتشار المعطى يوضح أن المتغيرين x , y :

(أ) مرتبطان عكسياً ارتباطاً ضعيفاً (ب) مرتبطان طردياً ارتباطاً ضعيفاً

(ج) غير مرتبطين (د) مرتبطان عكسياً ارتباطاً متوسطاً

(16) إذا كانت d تمثل الفرق في الرتب [بين قيم x , y] ، n هو عدد أزواج القيم (x,y) ، فإن معامل ارتباط الرتب r_s بين x , y هو

(أ) $r_s = \frac{1-6\sum d^2}{n(n^2-1)}$ (ب) $r_s = 1-\frac{6\sum d^2}{n(n^2-1)}$

(ج) $r_s = \frac{1-6\sum d^2}{n(n-1)}$ (د) $r_s = 1-\frac{6\sum d^2}{n(n-1)}$

(17) إذا كانت $x_1, x_2, \cdots, x_n$ تمثل n قيمة يمكن أن يأخذها متغير x ، $y_1, y_2, \cdots, y_n$ تمثل n قيمة يمكن أن يأخذها متغير آخر y ، وكانت $\bar{x}, \bar{y}$ هي الأوساط الحسابية للمتغيينر x , y ، وكانت d_x, d_y هي انحرافات قيم المتغيرين x , y [على الترتيب] عن أوساطهما الحسابية ، إذن يمكن التعبير عن معامل بيرسون للارتباط بين المتغيرين x , y على الصورة :

(أ) $r_p = \frac{\sum d_x d_y}{\left(\sum d_x^2\right)\left(\sum d_y^2\right)}$ (ب) $r_p = \frac{\left(\sum d_x\right)\left(\sum d_y\right)}{\left(\sum d_x^2\right)\left(\sum d_y^2\right)}$

(ج) $r_p = \frac{\sum d_x d_y}{\left(\sum d_x\right)\left(\sum d_y\right)}$ (د) $r_p = \frac{\sum d_x d_y}{\sum d_x^2 d_y^2}$

(18) وإذا كانت s_x , s_y هي الانحرافات المعيارية للمتغيرين x , y [على الترتيب] ، فإنه يمكن أيضاً التعبير عن معامل بيرسون للارتباط بين المتغيرين x , y على الصورة :

(أ) $$r_p = \frac{\left(\sum d_x\right)\left(d_y\right)}{n s_x s_y}$$

(ب) $$r_p = \frac{\sum d_x d_y}{\sqrt{n s_x^2 s_y^2}}$$

(ج) $$r_p = \frac{\sum d_x d_y}{n s_x s_y}$$

(د) $$r_p = \frac{\sum d_x d_y}{s_x s_y}$$

(19) لعدد من المشاهدات $n = 10$ لظاهرتين x , y ، كانت $\sum d^2 = 250$ ، حيث d تمثل الفرق في الرتب بين قيم x , y ، يكون معامل ارتباط الرتب r_s مساوياً لـ :

(أ) -1.52 (ب) -0.52 (ج) -16.66 (د) -14.15

(20) إذا كانت البيانات الخاصة بقيم ظاهرتين x , y على الصورة :

x	2	5	8	12
y	1	7	8	5

وكان r_p هو معامل بيرسون للارتباط بين المتغيرين x , y ، r_p هو معامل ارتباط سبيرمان (الرتب) بينهما ، فإنه [في هذا السؤال] :

(أ) يمكن حساب r_p فقط (ب) يمكن حساب r_s فقط

(ج) يمكن حساب كلٍ من r_p , r_s (د) لا يمكن حساب أيٍ من r_p , r_s

(21) إذا كانت البيانات الخاصة بقيم ظاهرتين x , y على الصورة :

x	A	B	C	D
y	1	7	8	5

[حيث A , B , C , D قيم غير كمية] وكان r_p هو معامل بيرسون للارتباط بين المتغيرين x , y ، r_p هو معامل ارتباط سبيرمان (الرتب) بينهما ، فإنه [في هذا السؤال] :

(أ) يمكن حساب r_p فقط (ب) يمكن حساب r_s فقط

(ج) يمكن حساب كلٍ من r_p , r_s (د) لا يمكن حساب أيٍ من r_p , r_s

---

خاص بالأسئلة من (22) إلى (27) :

لمجموعتين من القيم $(x_1 , x_2 , \cdots , x_n)$ ، $(y_1 , y_2 , \cdots , y_n)$ عدد كلٍ منهما n كانت هناك النتائج التالية :

$$n = 5 \quad , \quad \sum x = 30 \quad , \quad \sum y = 50 \quad , \quad \sum xy = 364 \quad , \quad \sum x^2 = 220 \quad , \quad \sum y^2 = 604$$

لهذه المجموعة يكون :

22) الوسط الحسابي للمتغير x يساوي :

أ) *6* ب) *10* ج) *44* د) *120.8*

23) الوسط الحسابي للمتغير y يساوي :

أ) *6* ب) *10* ج) *44* د) *120.8*

24) تباين المتغير x يساوي :

أ) *8* ب) *2.83* ج) *20.8* د) *4.56*

25) تباين المتغير y يساوي :

أ) *8* ب) *2.83* ج) *20.8* د) *4.56*

26) الانحراف المعياري للمتغير x يساوي :

أ) *8* ب) *2.83* ج) *20.8* د) *4.56*

27) الانحراف المعياري للمتغير y يساوي :

أ) *8* ب) *2.83* ج) *20.8* د) *4.56*

28) معامل التحديد للمتغيرين x , y يساوي :

أ) *0.985* ب) *-0.985* ج) *-0.993* د) *0.993*

29) معامل الارتباط بين x , y يساوي :

أ) *0.985* ب) *-0.985* ج) *-0.993* د) *0.993*

30) العلاقة بين x , y علاقة :

أ) طردية متوسطة ب) عكسية قوية جداً

ج) طردية قوية جداً د) طردية ضعيفة

أجوبة تدريبات (10)

1) د 2) ب 3) أ 4) د 5) ج 6) د 7) د 8) ب

9) د 10) ب 11) ب 12) ب 13) أ 14) د 15) ج 16) ب

17) أ 18) ج 19) ب 20) ج 21) ب 22) أ 23) ب 24) أ

25) ج 26) ب 27) د 28) أ 29) د 30) ج

عناصر الدرس

الباب السابع : تحليل الانحدار

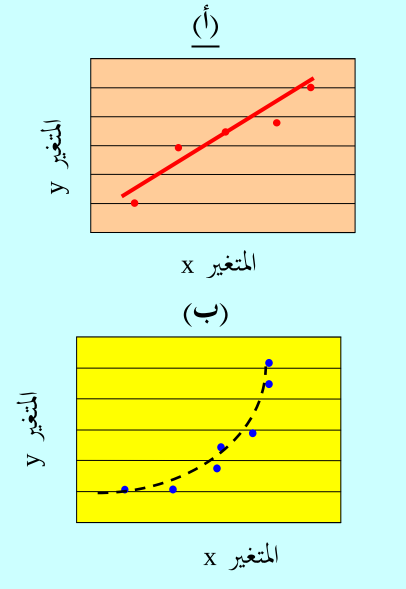
- **تمهيد رياضي**
  - ❖ **توفيق المنحنيات**
  - ❖ **طريقة المربعات الصغرى**
- **تحليل الانحدار**
  - ❖ **معادلة انحدار y على x**
  - ❖ **خطأ التقدير**
  - ❖ **معادلة انحدار x على y**
- **علاقات هامة**

الباب السابع : تحليل الانحدار

(7–1) تمهيد رياضي

(7–1–1) توفيق المنحنيات

في كثير من النواحي العملية نجد أن هناك علاقة بين متغيرين (أو أكثر) ، على سبيل المثال نجد أن أوزان مجموعة من الطلبة البالغين تعتمد بدرجة معينة على أطوالهم ، وفي أغلب الأحيان يكون من المرغوب فيه التعبير عن هذه العلاقة بصورة رياضية وذلك بتحديد المعادلة التي تربط بين هذه المتغيرات .



شكل (7–1) : شكل الانتشار

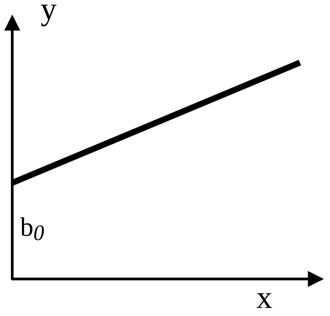
وللمساعدة في تحديد تلك المعادلة التي تربط بين المتغيرات ، نقوم [كخطوة أولى] بجمع بيانات تظهر القيم المتقابلة للمتغيرات محل الدراسة ، فعلى سبيل المثال نفرض أن المتغير x يمثل أطوال مجموعة من الطلبة ، وأن المتغير y يمثل أوزانهم ، إذن بأخذ عينة حجمها n [عدد الطلبة] يمكن تجميع بيانات عن أطوال الطلبة ممثلة بالقيم $x_1 , x_2 ,\cdots, x_n$ وبيانات عن أوزان نفس الطلبة ممثلة بالقيم $y_1 , y_2 ,\cdots, y_n$ ، ثم نقوم [كخطوة ثانية] برسم شكل الانتشار (7–1) [كما سبق وبينا في الفصل السابع] ومنه [وبالنظر] يمكن أن نقوم بتمهيد منحنى تقريبي يمثل العلاقة بين المتغيرين x , y [وذلك إذا كانت هناك علاقة بينهما] ، والعلاقة بين هذين المتغيرين قد تكون خطية [أي يمكن أن يكون المنحنى التقريبي خطاً مستقيماً كما في شكل (7–1أ)] ، وقد تكون غير خطية [أي لا يكون المنحنى التقريبي خطاً مستقيماً كما في شكل (7–1ب)] .

والمشكلة العامة في الحصول على معادلة المنحنيات التقريبية والتي تعطي أحسن توفيق لمجموعة من البيانات تُسمى "**توفيق المنحنيات**" ، وهناك العديد من الأشكال الشائعة للمنحنيات التقريبية ومعادلاتها منها [على سبيل المثال لا الحصر] الآتي :

خط مستقيم [منحنى من الدرجة الأولى]	$y = b_0 + b_1 x$
قطع مكافئ [منحنى من الدرجة الثانية]	$y = b_0 + b_1 x + b_2 x^2$

منحنى من الدرجة الثالثة	$y = b_0 + b_1 x + b_2 x^2 + b_3 x^3$
منحنى أسي	$y = b_0 e^{b_1 x}$
منحنى لوغاريتمي	$y = b_0 + b_1 \log x$
منحنى هندسي	$\log y = b_0 + b_1 \log x$

حيث جميع الأحرف [عدا y , x] تمثل ثوابت .

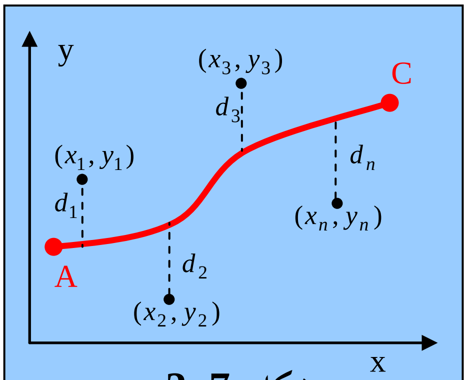


وأبسط صورة للمنحنى التقريبي هو الخط المستقيم :

$$y = b_0 + b_1 x \qquad (7\text{-}1)$$

حيث يمثل الثابت b_1 ميل الخط المستقيم ، في حين يمثل الثابت b_0 قيمة y عندما $x = 0$ [أي طول الجزء المقطوع من محور y] .

(2–1–8) طريقة المربعات الصغرى :



شكل (2–7)

لتلافي الخطأ الشخصي في تكوين المنحنى التقريبي ، فمن الضروري الاتفاق على تعريف "أفضل منحنى تقريبي" ، وبهدف الحصول على تعريف ممكن ، نفرض أننا رسمنا شكل الانتشار للبيانات [شكل (2–7)] ونفرض أن المنحنى AC هو المنحنى التقريبي الممثل للبيانات ، لقيمة معينة من قيم المتغير x [ولتكن x_1] سيكون هناك فارق بين القيمة **الفعلية** y_1 التي تعطيها البيانات والقيمة **المقدرة** من خلال المنحنى ، ليكن هذا الفارق d_1 (مثلاً) والذي يُسمى **بالانحراف** أو **الخطأ** أو **الباقي** [وقد يكون موجباً أو سالباً أو صفراً] ، وبنفس الأسلوب لقيم $x_2, \cdots, x_n$ على الانحرافات المقابلة $d_2, \cdots, d_n$.

ولقياس "جودة التوفيق" للمنحنى AC للبيانات المعطاة نحسب الكمية

$$E = d_1^2 + d_2^2 + \cdots\cdots + d_n^2 \qquad (7\text{-}2)$$

فإذا كانت هذه الكمية E صغيرة فإن التوفيق يكون جيداً ، وإذا كانت كبيرة فإن التوفيق يكون سيئاً ، ويُعتبر المنحنى الذي يجعل هذه الكمية "**أصغر ما يمكن**" هو **أفضل** منحنى يمكن توفيقه .

تُسمى الطريقة السابقة في توفيق المنحنيات بطريقة "**المربعات الصغرى**" ويُسمى المنحنى الناتج عندئذٍ **بمنحنى المربعات الصغرى** .

(2–7) تحليل الانحدار Regression Analysis

يعتبر تحليل الانحدار أكثر طرق التحليل الإحصائى استخداماً ، حيث يتم من خلاله التنبؤ بقيمة أحد المتغيرات (المتغير التابع) عند قيمة محددة لمتغير أو متغيرات أخرى (المتغيرات المستقلة) ، تسمى العلاقة الرياضية التى تصف سلوك المتغيرات محل الدراسة والتى من خلالها يتم التنبؤ بسلوك احد المتغيرين عند معرفة الاخر **بمعادلة خط الانحدار** . وهناك صورتين أساسيتين لمعادلة الانحدار وهما : **معادلة انحدار x | y** (التى يطلق عليها معادلة انحدار y على x) ، والأخرى معادلة انحدار y | x (والتى يطلق عليها معادلة انحدار x على y) .

(1–2–7) معادلة انحدار y على x $[y/x]$:

وهى التى تحدد قيمة المتغير y تبعاً لقيمة المتغير x [أي أننا نعتبر أن x هو المتغير المستقل ، y هو المتغير التابع] ، لذلك يمكن التعبير عن تلك العلاقة الخطية بالمعادلة التالية :

$$\boxed{\hat{y} = b_0 + b_1 x} \qquad (7\text{-}3)$$

حيث $\hat{y}$ هي قيمة y **التقديرية** والمناظرة للقيمة x ، b_0 , b_1 ثوابت . المعادلة (7-3) السابقة تمثل معادلة خط مستقيم يقطع من المحور الرأسي طولاً قدره b_0 وميله b_1 ، لذا يُسمى b_0 **بثابت الانحدار** (وهو يمثل طول الجزء المقطوع من المحور الرأسي) بينما b_1 يطلق عليه اسم **معامل الانحدار** أو **معدل التغير فى الدالة** .

ولتحديد المعادلة الدالة على العلاقة بين المتغيرين y و x [أي لتحديد العلاقة (7-3)] لابد من تقدير قيمة للثابتين b_0 , b_1 والذين يمكن تقديرهما من خلال تطبيق طريقة المربعات الصغرى كما يلى :

عند كل قيمة من قيمة x تكون القيمة الفعلية للمتغير التابع هي y ، بينما القيمة المقدرة من خلال الخط المستقيم [المنحنى التقريبي] تتحدد من العلاقة (7-3) ، وبالتالي يكون هناك فارق [أو خطأ] قدره :

$$d = y - \hat{y} = y - (b_0 + b_1 x) = y - b_0 - b_1 x$$

وذلك عند كل قيمة من قيم المتغير المستقل x ، وبالتالى يكون مجموع مربعات الأخطاء على الصورة التالية :

$$\boxed{E = \sum d^2 = \sum (y - b_0 - b_1 x)^2} \qquad (7\text{-}4)$$

وبالتالى يكون لدينا كمية E **دالة** في مجهولين b_0 , b_1 . ويتم تقدير قيمة للثابتين b_0 , b_1 التى تجعل مجموع مربعات الأخطاء فى المعادلة السابقة أقل ما يمكن من خلال إيجاد تفاضل المعادلة بالنسبة لكلٍ من b_0 , b_1 ومساواة التفاضل بالصفر ليكون كما يلى :

$$\frac{\partial E}{\partial b_0} = -2\sum (y - b_0 - b_1 x) = 0 \quad \Rightarrow \quad \sum y = n b_0 + b_1 \sum x \qquad (7\text{ - }5)$$

$$\frac{\partial E}{\partial b_1} = -2\sum x(y - b_0 - b_1 x) = 0 \quad \Rightarrow \quad \sum xy = b_0 \sum x + b_1 \sum x^2 \qquad (7\text{ - }6)$$

وبحل المعادلتين (7-5) ، (7-6) يمكن الحصول على قيمة كلٍ من b_0 , b_1 كالآتي :

$$\boxed{b_1 = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum x^2 - \left(\sum x\right)^2}} \qquad (7\text{-}7)$$

$$\boxed{b_0 = \frac{\sum y}{n} - b_1\frac{\sum x}{n} = \bar{y} - b_1\bar{x}} \qquad (7\text{-}8)$$

مثال (7−1) :

عند دراسة العلاقة بين عدد غرف المسكن x وكمية الكهرباء المستهلكة y بالألف كيلو وات ، كانت كما يلي :

عدد الغرف x	9	7	10	5	3	7	8	10	4	6
استهلاك الكهرباء y	12	9	14	6	4	7	10	10	5	8

المطلوب ، إيجاد :

(1) معادلة انحدار y على x ؟

(2) تحديد معدل التزايد أو التناقص فى استهلاك الكهرباء .

(3) الاستهلاك المتوقع لمسكن مكون من 8 غرف .

الحل :

من الجدول التالي يمكن حساب كلٍ من الآتي :

x	y	xy	x^2	y^2
9	12	108	81	144
7	9	63	49	81
10	14	140	100	196
5	6	30	25	36
3	4	12	9	16
7	7	49	49	49
8	10	80	64	100
10	10	100	100	100
4	5	20	16	25
6	8	48	36	64
$\sum x = 69$	$\sum y = 85$	$\sum xy = 650$	$\sum x^2 = 529$	$\sum y^2 = 811$

$$\bar{x} = \frac{\sum x}{n} = \frac{69}{10} = 6.9 \quad , \quad \bar{y} = \frac{\sum y}{n} = \frac{85}{10} = 8.5$$

$$b_1 = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum x^2 - \left(\sum x\right)^2} = \frac{10(650) - (69)(850}{10(529) - (69)^2} = 1.2003$$

$$b_0 = \bar{y} - b_1\bar{x} = 8.5 - (1.2003)(6.9) = 0.21793$$

وعلى ذلك يمكن كتابة معادلة الانحدار yعلى x على الشكل التالى :

$$\boxed{\hat{y} = 0.21793 + 1.2003x} \qquad (a)$$

وهو الرد على الجزء (1) من السؤال .

أما الجزء (2) من السؤال والخاص بمعدل التزايد أو التناقص في استهلاك الكهرباء فيمثله الثابت $b_1 = 1.2003$ وهو موجب مما يعني أن معدل استهلاك الكهرباء هو **معدل تزايد** وليس تناقص وبالمقدار *1.2003* ، أى أن كل غرفة بالمسكن تعمل على زيادة استهلاك الكهرباء بمقدار *1200.3* كيلو وات .

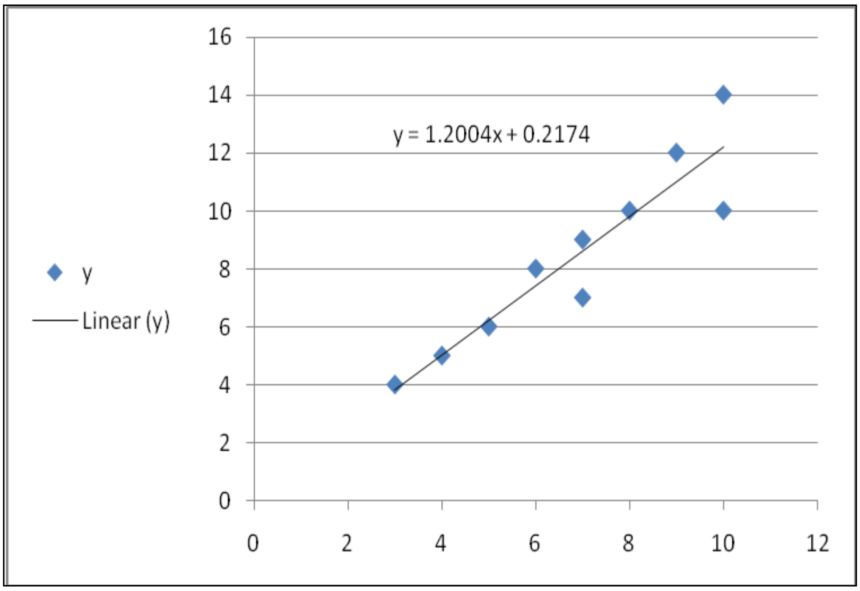
أما بالنسبة للجزء (3) من السؤال ، فإن الاستهلاك المتوقع لمسكن مكون من *8* غرف [أي قيمة تقديرية لـ y عندما x = *8*] فيمكن الحصول عليها بالتعويض فى المعادلة *(a)* [معادلة الانحدار التى سبق إيجادها] فنحصل على

$$\hat{y} = 0.21793 + 1.2003(8) = 9.8203$$

أى أن الاستهلاك المتوقع لمسكن مكون من *8* غرف هو *9820.3* كيلو وات.

(2-2-7) خطأ التقدير *Standard Error*

على ضوء المثال السابق ، يمكن لنا رسم بيانات التمرين السابق وخط معادلة الانحدار y على x كما هو موضح بالشكل التالي والذي منه يتضح لنا أن خط الانحدار لا يمر بجميع النقاط حيث يكون هناك تشتت حول الخط ، وبالرغم من ذلك يعد هذا الخط من أفضل الخطوط التى حصلنا عليها للتعبير عن العلاقة بين المتغيرين محل الدراسة ولكن بخطأ معين يسمى **خطأ التقدير Standard Error** والذي يُرمز له فى حالة معادلة انحدار y على x بـ $S_{y|x}$ والذي يشير إلى الجذر التربيعي لمتوسط مربعات الخطأ .



ويُعرف خطأ التقدير $S_{y/x}$ من العلاقة :

$$\boxed{S_{y/x} = \sqrt{\frac{\sum (y-\hat{y})^2}{n-2}}} \qquad (7\text{-}9)$$

والمعادلة *(7-9)* يمكن كتابتها على الصورة :

$$\boxed{S_{y/x}=\sqrt{\frac{\sum y^2-b_1\sum xy-b_0\sum y}{n-2}}}\qquad (7\text{-}10)$$

والتي قد تكون أكثر ملائمة للحساب [أنظر مثال (7–2)] .

مثال (7–2) :

من تعريف خطأ التقدير $S_{y/x}$ [المعادلة (7-9)] **بين** صحة المعادلة (7-10) .

الحل :

من معادلة انحدار y/x :

$$\hat{y}=b_0+b_1x$$

وبالتالي [ومن تعريف خطأ التقدير] يكون :

$$\begin{aligned}
S^2_{y/x}&=\frac{\sum(y-\hat{y})^2}{n-2}=\frac{1}{n-2}\sum(y-b_0-b_1x)^2\\
&=\frac{1}{n-2}\sum(y-b_0-b_1x)(y-b_0-b_1x)\\
&=\frac{1}{n-2}\sum y(y-b_0-b_1x)\\
&-\frac{b_0}{n-2}\underbrace{\sum(y-b_0-b_1x)}_{=\sum y-nb_0-b_1\sum x=0}-\frac{b_1}{n-2}\underbrace{\sum x(y-b_0-b_1x)}_{=\sum xy-b_0\sum x-b_1\sum x^2=0}
\end{aligned}$$

[وذلك بالاستفادة من المعادلات (7-5) ، (7-6)] ، إذن :

$$S^2_{y/x}=\frac{1}{n-2}\sum y(y-b_0-b_1x)=\frac{\sum y^2-b_0\sum y-b_1\sum xy}{n-2}$$

ومنها يكون خطأ التقدير $S_{y/x}$ هو :

$$S_{y/x}=\sqrt{S^2_{y/x}}=\sqrt{\frac{\sum y^2-b_1\sum xy-b_0\sum y}{n-2}}$$

مثال (7–3) :

احسب خطأ التقدير $S_{y/x}$ نتيجة لاستخدام معادلة انحدار y/x في مثال (7–1) .

الحل :

باستخدام بيانات ونتائج مثال (7–1) :

$$\sum y^2=811\quad,\quad\sum xy=650\quad,\quad\sum y=85$$

$$b_0=0.21\quad,\quad b_1=1.2003$$

يمكن حساب خطأ التقدير من المعادلة *(7-10)* كالآتي :

$$S_{y/x}=\sqrt{\frac{811-(1.2003)(650)-0.21793(85)}{10-2}}\cong 1.239$$

(3−2−7) معادلة انحدار x على y $[x/y]$:

وهي التي تحدد قيمة المتغير x تبعاً لقيمة المتغير y [أي أننا نعتبر أن y هو المتغير المستقل ، x هو المتغير التابع] ، لذلك يمكن التعبير عن تلك العلاقة الخطية بالمعادلة التالية :

$$\boxed{\hat{x}=c_0+c_1y} \qquad (7\text{-}11)$$

حيث $\hat{x}$ هي قيمة ء التقديرية والمناظرة للقيمة y ، c_0 هو ثابت الانحدار ، c_1 هو معامل الانحدار ، والثابتان c_0 , c_1 يمكن تحديدهما بنفس أسلوب تحديد الثوابت b_0 , b_1 فتكون النتيجة :

$$\boxed{c_1=\frac{n\sum xy-\left(\sum x\right)\left(\sum y\right)}{n\sum y^2-\left(\sum y\right)^2}} \qquad (7\text{-}12)$$

$$\boxed{c_0=\frac{\sum x}{n}-b_1\frac{\sum y}{n}=\bar{x}-c_1\bar{y}} \qquad (7\text{-}13)$$

لاحظ أن العلاقات *(7-12)* , *(7-13)* [المستخدمة في تحديد الثوابت c_0 , c_1 هي نفس العلاقات *(7-8)* , *(7-7)* [المستخدمة في تحديد الثوابت b_0 , b_1] بعد استبدال x بـ y والعكس .

وكما في حالة معادلة انحدار y على x يكون خطأ التقدير في حالة استخدام معادلة انحدار x على y هو :

$$\boxed{S_{x/y}=\sqrt{\frac{\sum(x-\hat{x})^2}{n-2}}=\sqrt{\frac{\sum x^2-c_1\sum xy-c_0\sum x}{n-2}}} \qquad (7\text{-}14)$$

وهي معادلة شبيهة بالمعادلتين *(7-10)* , *(7-9)* ، وبوجهٍ عام فإن $\boxed{S_{y/x}\neq S_{x/y}}$.

مثال (4−7) :

باستخدام بيانات مثال (1−7) السابق ، **أوجد** :

(1) معادلة انحدار x على y ؟

(2) عدد الغرف المتوقع لاستهلاك *25000* وات .

(3) خطأ التقدير $S_{x/y}$ نتيجة لاستخدام معادلة انحدار x على y .

الحل :

(1) معادلة انحدار x على y :

$$\hat{x}=c_0+c_1y$$

حيث :

$$c_1 = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum y^2 - (\sum y)^2} = \frac{10((650)-(69)(85)}{10(811)-(85)^2} \cong 0.718$$

$$c_0 = \bar{x} - c_1\bar{y} = 6.9 - 0.718(8.5) \cong 0.797$$

وبالتالي تكون معادلة انحدار x على y $[x/y]$ على الصورة :

$$x = c_0 + c_1 y \quad \Rightarrow \quad \hat{x} = 0.797 + 0.718y$$

(2) <u>عدد الغرف المتوقع لاستهلاك *25000* وات :</u>

أي قيمة $\hat{x}$ [قيمة x المقدرة] عندما $y = 25$ [لأن y بالألف كيلو وات] ، إذن بالتعويض في معادلة انحدار x على y عن $y = 25$ نحصل على :

$$\hat{x} = 0.797 + 0.718(25) = 18.747 \cong 19$$

أي تقريباً *19* غرفة [عدد الغرف يجب أن يكون عدداً صحيحاً] .

(3) <u>خطأ التقدير نتيجة استخدام معادلة انحدار x على y</u>

باستخدام بيانات مثال (1–8) ، ونتائج الجزء الأول من هذا المثال :

$$\sum x^2 = 529 \quad , \quad \sum xy = 650 \quad , \quad \sum x = 69$$

$$c_0 = 0.797 \quad , \quad c_1 = 0.718$$

يمكن حساب خطأ التقدير $S_{x/y}$ من المعادلة *(8-14)* كالآتي :

$$S_{x/y} = \sqrt{\frac{529 - (0.718)(650) - 0.0.797(69)}{10-2}} \cong 0.956$$

(3–7) علاقات هامة

- من العلاقات *(7-12) , (7-7)* يمكن استنتاج أن :

$$\boxed{r^2 = b_1 c_1} \qquad (7-15)$$

أي أن معامل التحديد [مربع معامل بيرسون للارتباط ، العلاقة *(6-3)*/الباب السادس/الدرس العاشر] هو حاصل ضرب معاملي الانحدار b_1 [معامل انحدار y/x] ، c_1 [معامل انحدار x/y] ، وبالتالي يمكن الحصول على معامل الارتباط r بأخذ الجذر التربيعي لمعامل التحديد :

$$\boxed{r = \sqrt{r^2} = \sqrt{b_1 c_1}} \qquad (7-16)$$

مع مراعاة أن إشارة معامل الارتباط [موجبة أو سالبة] هي نفسها إشارة كلٍ من معاملي التحديد b_1 , c_1 [أي أن إشارات r , b_1 , c_1 واحدة ؛ كلها سالبة أو كلها موجبة] .

- أيضاً يمكن استنتاج أن :

$$\boxed{b_1 = r \times \frac{s_y}{s_x}} \qquad (7\text{-}17)$$

$$\boxed{c_1 = r \times \frac{s_x}{s_y}} \qquad (7\text{-}18)$$

حيث s_x هو الانحراف المعياري للمتغير x ، s_y هو الانحراف المعياري للمتغير y .

والعلاقات (7-17) , (7-18) السابقة تمكننا من معرفة معاملي التحديد b_1 , c_1 من معامل الارتباط r والانحرافات المعيارية لكلٍ من x , y .

مثال (7−5) : **أثبت** صحة العلاقات (7-16) , (7-17) , (7-18) السابقة .

الإثبات :

- العلاقة (7-16) :

من العلاقات (7-7) , (7-12) نجد أن :

$$b_1c_1 = \underbrace{\frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum x^2 - \left(\sum x\right)^2}}_{\text{from } (8-7)} \times \underbrace{\frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum y^2 - \left(\sum y\right)^2}}_{\text{from } (8-12)}$$

$$= \frac{\left(n\sum xy - \left(\sum x\right)\left(\sum y\right)\right)^2}{\left(n\sum x^2 - \left(\sum x\right)^2\right)\left(n\sum y^2 - \left(\sum y\right)^2\right)}$$

ومن العلاقة (6-3) [الباب السادس/الدرس العاشر] :

$$r^2 = r_p^2 = \frac{\left(n\sum xy - \left(\sum x\right)\left(\sum y\right)\right)^2}{\left(n\sum x^2 - \left(\sum x\right)^2\right)\left(n\sum y^2 - \left(\sum y\right)^2\right)}$$

مما يعني صحة العلاقة (7-16) .

- العلاقات (7-17) , (7-18) :

من الباب الرابع [العلاقة (4-7)] بينا أن التباين s_x^2 للمتغير x يمكن كتابته على الصورة :

$$s_x^2 = \frac{\sum x^2}{n} - \bar{x}^2$$

وحيث أن $\bar{x} = \frac{\sum x}{n}$ ، فإن الصورة أيضاً يمكن كتابتها كالآتي :

$$s_x^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2 = \frac{\sum x^2}{n} - \frac{\left(\sum x\right)^2}{n^2} = \frac{n\sum x^2 - \left(\sum x\right)^2}{n^2}$$

وبالمثل ، يمكن كتابة التباين s_y^2 للمتغير y يمكن كتابته على الصورة :

$$s_y^2 = \frac{n\sum y^2 - \left(\sum y\right)^2}{n^2}$$

ومن ثم يكون :

$$\frac{s_y^2}{s_x^2} = \frac{\dfrac{n\sum y^2 - \left(\sum y\right)^2}{n^2}}{\dfrac{n\sum x^2 - \left(\sum x\right)^2}{n^2}} = \frac{n\sum y^2 - \left(\sum y\right)^2}{n\sum x^2 - \left(\sum x\right)^2}$$

ومنها :

$$r^2 \times \frac{s_y^2}{s_x^2} = \frac{\left(n\sum xy - \left(\sum x\right)\left(\sum y\right)\right)^2}{\left(n\sum x^2 - \left(\sum x\right)^2\right)\left(n\sum y^2 - \left(\sum y\right)^2\right)} \times \frac{n\sum y^2 - \left(\sum y\right)^2}{n\sum x^2 - \left(\sum x\right)^2}$$

$$= \frac{\left(n\sum xy - \left(\sum x\right)\left(\sum y\right)\right)^2}{\left(n\sum x^2 - \left(\sum x\right)^2\right)^2} = b_1^2$$

أي أن :

$$b_1 = r \times \frac{s_y}{s_x}$$

وهي العلاقة *(7-17)* .

وبنفس الطريقة :

$$\frac{s_x^2}{s_y^2} = \frac{n\sum x^2 - \left(\sum x\right)^2}{n\sum y^2 - \left(\sum y\right)^2}$$

ومنها :

$$r^2 \times \frac{s_x^2}{s_y^2} = \frac{\left(n\sum xy - \left(\sum x\right)\left(\sum y\right)\right)^2}{\left(n\sum y^2 - \left(\sum y\right)^2\right)^2} = c_1^2$$

أي أن :

$$c_1 = r \times \frac{s_y}{s_x}$$

وهي العلاقة *(7-18)* .

مثال (7−6) :

لبيانات مثال (7−1) ، **أوجد** معامل الارتباط بين عدد الغرف x واستهلاك الكهرباء y .

الحل :

الطريقة الأولى [الطريقة المباشرة] : من بيانات مثال (7−1) يمكن تكوين الجدول التالي ، ثم من العلاقة

$$r=\frac{n\sum xy-\left(\sum x\right)\left(\sum y\right)}{\sqrt{n\sum x^2-\left(\sum x\right)^2}\sqrt{n\sum y^2-\left(\sum y\right)^2}}$$

يمكن تحديد معامل الارتباط المطلوب كالتالي :

x	y	xy	x^2	y^2
9	12	108	81	144
7	9	63	49	81
10	14	140	100	196
5	6	30	25	36
3	4	12	9	16
7	7	49	49	49
8	10	80	64	100
10	10	100	100	100
4	5	20	16	25
6	8	48	36	64
$\sum x=69$	$\sum y=85$	$\sum xy=650$	$\sum x^2=529$	$\sum y^2=811$

$$r=\frac{n\sum xy-\left(\sum x\right)\left(\sum y\right)}{\sqrt{n\sum x^2-\left(\sum x\right)^2}\sqrt{n\sum y^2-\left(\sum y\right)^2}}$$

$$=\frac{10(650)-(69)(85)}{\sqrt{10(529)-(69)^2}\sqrt{10(811)-(85)^2}}=\frac{635}{\sqrt{529}\sqrt{875}}\cong 0.93$$

الطريقة الثانية : من معاملي التحديد b_1 [مثال (7−1)] ، c_1 [مثال (7−4)] نجد أن :

$$b_1=1.2003\quad,\quad c_1=0.718\quad\Rightarrow\quad r=+\sqrt{b_1c_1}=\sqrt{(1.2003)(0.718)}\cong 0.93$$

لاحظ أننا أخذا الإشارة الموجبة للجذر التربيعي حيث أن إشارة r هي نفسها إشارة كلٍ من b_1 , c_1 [وهما موجبتان في هذا المثال] .

ملخص للدرس الحادي عشر [الباب السابع : تحليل الانحدار]

- "**تحليل الانحدار**" هو تحليل إحصائي يتم من خلاله التنبؤ بقيمة أحد المتغيرات (المتغير التابع) عند قيمة محددة لمتغير أو متغيرات أخرى مرتبطة به .

- **المعادلة y/x** [معادلة انحدار y على x] هي معادلة تحدد قيمة المتغير y (المتغير التابع) تبعاً لقيمة المتغير x (المتغير المستقل) ، وهي تأخذ الصورة :

$$\hat{y} = b_0 + b_1 x$$

- **المعادلة x/y** [معادلة انحدار x على y] حيث يكون المتغير x هو المتغير المستقل ، y هو المتغير التابع ، وهي تكون على الصورة :

$$\hat{x} = c_0 + c_1 y$$

- يُسمى b_0 , c_0 بثوابت الانحدار ، ويُسمى الثابتان b_1 , c_1 بمعاملي الانحدار حيث :

$$b_1 = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum x^2 - \left(\sum x\right)^2} \quad , \quad c_1 = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum y^2 - \left(\sum y\right)^2}$$

$$b_0 = \bar{y} - b_1 \bar{x} \quad , \quad c_0 = \bar{x} - c_1 \bar{y}$$

أيضاً :

$$b_1 c_1 = r^2 \quad , \quad b_1 = r \times \frac{s_y}{s_x} \quad , \quad c_1 = r \times \frac{s_x}{s_y}$$

حيث r هو معامل الارتباط بين المتغيرين x , y ، s_x هو الانحراف المعياري للمتغير x ، s_y هو الانحراف المعياري للمتغير y .

- **خطأ التقدير $S_{y/x}$** هو الخطأ الناتج من استخدام معادلة انحدار y على x :

$$S_{y/x} = \sqrt{\frac{\sum (y - \hat{y})^2}{n-2}} = \sqrt{\frac{\sum y^2 - b_1 \sum xy - b_0 \sum y}{n-2}}$$

- **خطأ التقدير $S_{x/y}$** هو الخطأ الناتج من استخدام معادلة انحدار x على y :

$$S_{x/y} = \sqrt{\frac{\sum (x - \hat{x})^2}{n-2}} = \sqrt{\frac{\sum x^2 - c_1 \sum xy - c_0 \sum x}{n-2}}$$

وبوجهٍ عام فإن $S_{y/x} \neq S_{x/y}$.

تدريب عملي

البيانا

x	2	5	4	2	6	7
y	12	17	15	13	18	22

المطلوب :

- تقدير معادلة انحدار y/x وتقدير قيمة y عندما $x = 10$.
- تقدير معادلة انحدار x/y وتقدير قيمة x عندما $y = 18$.
- حساب الخطأ المعياري للتقدير $S_{y/x}$.

تدريبات (11)

الإجابة النهائية لجميع التمرينات موجودة في نهاية التدريب

اختر الإجابة الصحيحة

خاص بالأسئلة من (1) إلى (7) :

إذا كانت $x_1, x_2, \cdots, x_n$ تمثل n قيمة يمكن أن يأخذها متغير x ، $y_1, y_2, \cdots, y_n$ تمثل n قيمة يمكن أن يأخذها متغير آخر y ، وكانت $\bar{x}, \bar{y}$ هي الأوساط الحسابية للمتغيين x , y ، وكانت s_x, s_y هي الانحرافات المعيارية للمتغيرين ، وكان b_0 هو ثابت خط انحدار y على x ، b_1 هو معامل خط انحدار y على x ، c_0 هو ثابت خط انحدار x على y ، c_1 هو معامل خط انحدار x على y ، فإن :

(1) الانحراف المعياري للمتغير x يساوي :

(أ) $\dfrac{\sum x^2}{n} - \bar{x}^2$

(ب) $\sqrt{\dfrac{\sum x^2}{n} - \bar{x}^2}$

(ج) $\sqrt{\dfrac{\sum x^2}{n}} - \bar{x}^2$

(د) $\dfrac{\sum x}{n} - \bar{x}$

(2) الانحراف المعياري للمتغير y يساوي :

(أ) $\sqrt{\dfrac{(y-\bar{y})^2}{n}}$

(ب) $\dfrac{(y-\bar{y})^2}{n}$

(ج) $\dfrac{y^2-\bar{y}^2}{n}$

(د) $\sqrt{\dfrac{y^2-\bar{y}^2}{n}}$

(3) معادلة خط انحدار y على x هي :

(أ) $\hat{y} = b_1 + b_0 x$

(ب) $\hat{x} = c_1 + c_0 y$

(ج) $\hat{y} = b_0 + b_1 x$

(د) $\hat{x} = c_0 + c_1 y$

(4) معادلة خط انحدار x على y هي :

(أ) $\hat{y} = b_1 + b_0 x$

(ب) $\hat{x} = c_1 + c_0 y$

(ج) $\hat{y} = b_0 + b_1 x$

(د) $\hat{x} = c_0 + c_1 y$

(5) معامل الارتباط بين x , y يساوي :

(أ) $\sqrt{b_1 c_1}$

(ب) $b_1 c_1$

(ج) $\dfrac{b_1}{c_1}$

(د) $\sqrt{\dfrac{b_1}{c_1}}$

(6) معامل التحديد بين x , y يساوي :

(أ) $\sqrt{b_1c_1}$ (ب) b_1c_1

(ج) $\frac{b_1}{c_1}$ (د) $\sqrt{\frac{b_1}{c_1}}$

(7) معامل خط انحدار y على x يُعطى بـ :

(أ) $\frac{s_x}{rs_y}$ (ب) $\frac{rs_x}{s_y}$

(ج) $\frac{rs_y}{s_x}$ (د) $\frac{s_y}{rs_x}$

خاص بالأسئلة من (8) إلى (19) :

لمجموعتين من القيم $(x_1, x_2, \cdots, x_n)$ ، $(y_1, y_2, \cdots, y_n)$ عدد كلٍ منهما n كانت هناك النتائج التالية :

$$n = 5 \quad , \quad \sum x = 15 \quad , \quad \sum y = 32 \quad , \quad \sum xy = 118 \quad , \quad \sum x^2 = 55 \quad , \quad \sum y^2 = 254$$

لهذه المجموعة يكون :

(8) الوسط الحسابي للمتغير x يساوي :

(أ) 3 (ب) 6.4 (ج) 2.83 (د) 3.14

(9) الوسط الحسابي للمتغير y يساوي :

(أ) 3 (ب) 6.4 (ج) 2.83 (د) 3.14

(10) الانحراف المعياري للمتغير x يساوي :

(أ) 8 (ب) 9.84 (ج) 2.83 (د) 3.14

(11) الانحراف المعياري للمتغير y يساوي :

(أ) 8 (ب) 9.84 (ج) 2.83 (د) 3.14

(12) معامل خط انحدار y على x يساوي :

(أ) -0.2 (ب) 2.2 (ج) 0.139 (د) 0.447

(13) معامل خط انحدار x على y يساوي :

(أ) -0.2 (ب) 2.2 (ج) 0.139 (د) 0.447

(14) ثابت خط انحدار y على x يساوي :

(أ) -0.2 (ب) 2.2 (ج) 0.139 (د) 0.447

(15) ثابت خط انحدار x على y يساوي :

أ) -0.2   ب) 2.2   ج) 0.139   د) 0.447

16) معامل الارتباط بين x , y يساوي :

أ) 0.992   ب) 0.203   ج) -0.992   د) 0.451

17) العلاقة بين x , y علاقة :

أ) طردية متوسطة   ب) عكسية قوية جداً

ج) طردية قوية جداً   د) طردية ضعيفة

18) خطأ التقدير في الحسابات نتيجة استخدام خط انحدار y على x في حساب القيم المقدرة يساوي :

أ) 0.267   ب) 0.1.446   ج) 1.20   د) 0.52

19) خطأ التقدير في الحسابات نتيجة استخدام خط انحدار x على y في حساب القيم المقدرة يساوي :

أ) 0.267   ب) 0.446   ج) 1.20   د) 0.52

خاص بالأسئلة من (20) إلى (24) :

عند تحديد خط انحدار y على x ، وخط انحدار x على y لظاهرتين x , y كانت لنا النتائج التالية :

$$b_0 = 3.9 \quad , \quad b_1 = 2.2 \quad , \quad c_0 = -2.5 \quad , \quad c_1 = 0.4$$

حيث b_0 هو ثابت خط انحدار y على x ، b_1 هو معامل خط انحدار y على x ، c_0 هو ثابت خط انحدار x على y ، c_1 هو معامل خط انحدار x على y ، من هذه البيانات يكون :

20) معادلة انحدار y على x هي :

أ) $\hat{y} = 3.9 + 2.2x$ ب) $\hat{y} = 2.2 + 3.9x$

ج) $\hat{x} = -2.5 + 0.4y$ د) $\hat{x} = 0.4 - 2.5y$

21) معادلة انحدار x على y هي :

أ) $\hat{y} = 3.9 + 2.2x$ ب) $\hat{y} = 2.2 + 3.9x$

ج) $\hat{x} = -2.5 + 0.4y$ د) $\hat{x} = 0.4 - 2.5y$

22) قيمة y المقدرة عند $x = 2$ هي :

أ) 10   ب) 8.3   ج) -1.7   د) -4.6

23) قيمة x المقدرة عند $y = 5$ هي :

أ) -12.1   ب) 14.9   ج) -0.5   د) 21.7

24) معامل الارتباط بين المتغيرين x , y يساوي

أ) *0.942*   ب) *0.88*   ج) *-0.942*   د) *-0.88*

أجوبة تدريبات (11)

1) ب   2) أ   3) ج   4) د   5) أ   6) ب   7) ج   8) أ

9) ب   10) ج   11) د   12) ب   13) د   14) أ   15) ج   16) أ

17) ج   18) د   19) ج   20) أ   21) ج   22) ب   23) ج   24) أ

عناصر الدرس

الباب الثامن : السلاسل الزمنية

- تعريف السلسلة الزمنية
- الرسم البياني للسلاسل الزمنية
- أنواع السلاسل الزمنية
- تحليل السلسلة الزمنية
- الاتجاه العام
- التغيرات الموسمية
- التغيرات الدورية
- التغيرات العشوائية (أو الفجائية)

الباب الثامن : السلاسل الزمنية

(8–1) تعريف السلسلة الزمنية

السلسلة الزمنية عبارة عن مجموعة من المشاهدات الإحصائية التي تصف الظاهرة مع مرور الزمن ، أو هي البيانات الإحصائية التي تجمع أو تشاهد أو تسجل لفترات متتالية من الزمن .

وقد تكون السلسلة الزمنية بالأرقام المطلقة [وتسمى بالتالي **سلسلة قيم مطلقة**] ، أو قد تكون بالقيم النسبية مثل تلك الجداول التي تبين معدلات الزيادة الطبيعية للسكان في الألف ونحوها ، أو قد تكون بالمتوسطات [مثل السلسلة الزمنية التي تبين متوسط إنتاج الكيلومتر مربع من القمح] .

ومن أمثلة هذه السلاسل الزمنية :

- مرضى العيادات النفسية المترددين شهرياً على إحدى المستشفيات .
- عدد المتعطلين سنوياً عن العمل .
- معدلات الإنجاب السنوية .
- المبيعات اليومية في مركز لبيع الكتب لمدة شهر .
- قراءة درجات حرارة المريض في ساعة لمدة يوم واحد .
- قراءة الإنتاج الشهري لمدة سنة في شركة للأدوية .
- الإنتاج الشهري من البترول للسعودية ولعدة سنوات .

وتُعرف السلسلة الزمنية رياضياً بالقيم $y_1, y_2, \cdots\cdots$ التي يأخذها المتغير y [درجات الحرارة ، عدد المرضى ، أسعار الأسهم ، ... ، وغيرها] عند الزمن $t_1, t_2, \cdots\cdots$ على الترتيب ، ونعبر عن ذلك رياضياً بالصورة :

$$y = F(t)$$

والتي تعني أن y دالة في t .

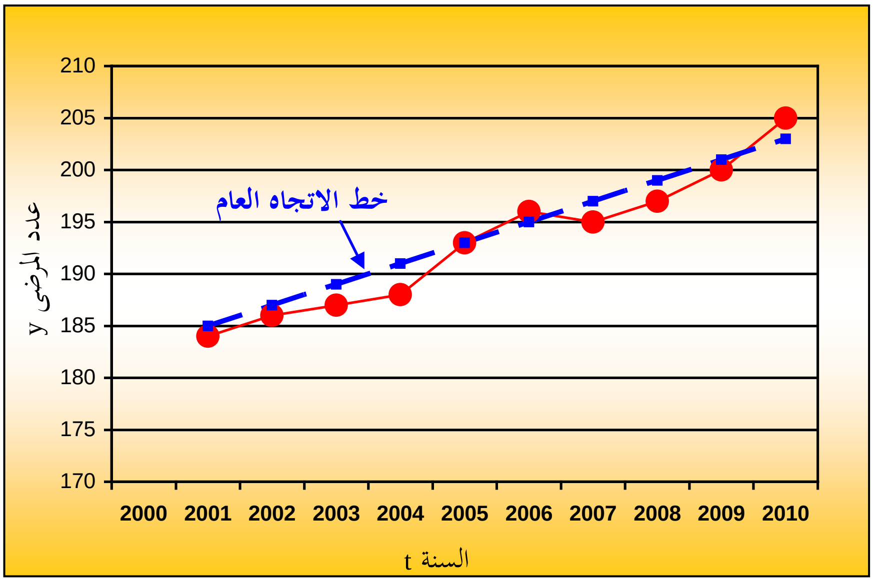
(8–2) الرسم البياني للسلاسل الزمنية :

تُمثل السلسلة الزمنية المتضمنة المتغير y بيانياً برسم الشكل البياني "y مقابل t" ، فمثلاً البيانات المبينة بجدول (8–1) التالي يوضح عدد مرض الفصام المترددين على إحدى العيادات خلال عشر سنوات [من سنة *2001* حتى سنة *2010*] ، هذه البيانات يمكن تمثيلها بيانياً بالشكل (8–1) حيث

جدول (8–1)

السنة	عدد المرضى
2001	184
2002	186
2003	187
2004	188
2005	193
2006	196
2007	195
2008	197
2009	200
2010	205

يُمثل الزمن t [السنة] على المحور الأفقي ، ويُمثل المتغير y [عدد المرضى] على المحور الرأسي .



شكل (8–1) : السلسلة الزمنية لبيانات جدول (9–1)

(8–3) أنواع السلاسل الزمنية :

- **سلسلة زمنية فترية** وهي السلسلة التي تتكون من بيانات كمية لمستوى الظاهرة عن فترات محددة من الزمن (شهر ، ربع سنة ، أو ما شابه ذلك) .
- **سلسة زمنية لحظية** وهي السلسلة التي تتكون من مستويات للظاهرة مقاسة في لحظات (تواريخ معينة ومحددة) .

(8–4) تحليل السلسلة الزمنية :

لغرض فهم السلسلة الزمنية لابد من تحليلها إلى عناصرها ومركباتها الأساسية بطريقة تمكننا من معرفة تطور الظاهرة مع الزمن والتنبؤ بمعالمها خلال الفترات المقبلة لتتخذ أساساً للتخطيط الاقتصادي أو الإداري الطويل الأجل .

وتتألف السلسلة الزمنية من أربعة عناصر أساسية هي :

- **التحركات طويلة المدى [الاتجاه العام]** ويرمز لقيمته بالرمز (T) وتسمى "**القيم الإتجاهية**" ، وهي تشير إلى الاتجاه العام الذي يظهر به الشكل البياني للسلسلة الزمنية على مدى فترة طويلة من الزمن .

فمثلاً في شكل (9-1) أعلاه فإن الخط المتصل يمثل الشكل البياني للسلسلة الزمنية ، في حين يمثل الخط المتقطع الاتجاه العام والذي يُسمى منحنى الاتجاه العام ، ولبعض السلاسل الزمنية قد يكون خط الاتجاه العام أكثر ملاءمة وقد سبق دراسة تحديد مثل هذه الخطوط والمنحنيات بطريقة المربعات الصغرى في الفصل الثامن .

- **التغيرات الموسمية** [ويرمز لقيمها بالرمز (S) وتسمى "**القيم الموسمية**"] وهي تشير إلى النمط المتماثل لحركة السلسلة الزمنية في الأشهر المتقابلة خلال السنوات المتتالية ، مثل الزيادة المفاجئة في مبيعات المحلات في الفترة السابقة للأعياد .
- **التغيرات الدورية** [ويرمز لقيمها بالرمز (C) وتسمى "**القيم الدورية**"] وهي تشير إلى الذبذبات طويلة المدى حول خط (أو منحنى) الاتجاه العام ، وهذه الدورات قد تكون وقد لا تكون على فترات ، بمعنى أنها قد تتبع وقد لا تتبع نفس النمط بعد كل فترة زمنية متساوية ، مثل دورات الأعمال التي تمثل فترات الرخاء ، الركود ، الكساد .
- **التغيرات العشوائية** أو **الفجائية** [ويرمز لقيمها بالرمز (R) وتسمى "**القيم العشوائية**"] مثل الفيضانات ، الإضرابات ، الانتخابات وغيرها .

أي أن القيمة الأصلية y_t للظاهرة [أو المتغير] عند أي فترة زمنية t [سنة من السنوات مثلاً] يمكن وصفها بأحد النموذجين التاليين :

- **نموذج الجمع :** وفيه يتم فرض أن السلسلة الزمنية مكونة من مجموع مكوناتها من الأربع عناصر السابق ذكرها ، أى يكون النموذج على الصورة التالية :

$$\boxed{y_t = T_t + C_t + S_t + R_t} \qquad (8\text{-}1)$$

حيث y_t يمثل قيمة الظاهرة (القيمة الحقيقية للمتغير أو الظاهرة) ، T_t يمثل قيمة الاتجاه العام ، C_t يمثل قيمة التغيرات الموسمية (القيم الموسمية) ، S_t يمثل قيمة التغيرات الدورية (القيم الدورية) ، R_t يمثل قيمة التغيرات العشوائية (القيم العشوائية) ، وجميع تلك الكميات عند اللحظة المدروسة t .

ويستخدم هذا النموذج عندما يكون مدى التغيرات الموسمية ثابت من سنة إلى أخرى ومستقل عن الاتجاه العام

.

- **نموذج الضرب :** وفيه يتم فرض أن السلسلة الزمنية مكونة من حاصل ضرب مكوناتها من الأربع عناصر السابق ذكرها ، أى يكون النموذج على الصورة التالية :

$$\boxed{y_t = T_t \times C_t \times S_t \times R_t} \qquad (8\text{-}2)$$

مع مراعاة أن قيم المتغيرات الموسمية وكذا المتغيرات الدورية عبارة عن نسب مئوية في نموذج الضرب

ودراسة أي سلسلة زمنية وتحليلها يستدعي دراسة كل عنصر من العناصر الأربعة السابقة على حده والتي يُشار إليها بتفكيك السلسلة الزمنية إلى عناصرها الأساسية ، وقرار إتخاذ أيٍ من النموذجين السابقين في التفطيط [نموذج الجمع أو نموذج الضرب] يعتمد على درجة النجاح المتحقق في تطبيق النموذج ، وسوف نفترض **نموذج الضرب** في هذا الفصل .

(8–5) الاتجاه العام *The Secular Trend*

تغيرات الاتجاه العام تعني الزيادة أو الانخفاض طويل الأجل في البيانات عبر الزمن ، ويتم التعرف على ذلك من خلال تمثيل السلسلة الزمنية بيانياً فنحصل بالتالي على خط بياني ، واتجاه خط السلسلة الزمنية صعوداً أو هبوطاً يسمى **الاتجاه العام للسلسلة** ، فإذا نظرنا للخط ووجدناه يتجه من الأسفل إلى الأعلى دل ذلك على نمو الظاهرة مع مرور الزمن ، أما إذا كان الخط يهبط من الأعلى إلى الأسفل دل ذلك على أن الظاهرة تنقص مع مرور الزمن ، أما إذا كان الخط أفقياً دل ذلك على ثبات الظاهرة .

ويمكن تقدير الاتجاه العام بعدة طرق منها :

- طريقة الانتشار (التمهيد باليد) .
- طريقة المتوسطات المتحركة .
- طريقة متوسط نصف السلسلة .
- طريقة المربعات الصغرى .

ولنتناول كل طريقة من الطرق التالية :

(8–5–1) طريقة الانتشار (التمهيد باليد)

يتم من خلال هذه الطريقة رسم شكل الانتشار Scatter Diagram للظاهرة موضع الدراسة ، وشكل الانتشار عبارة عن رسم بياني لمتغيرين بحيث يكون الزمن على المحور السيني وقيم الظاهرة على المحور الصادي ، وعند توصيل نقط شكل الانتشار ببعضها البعض نحصل على الخط البياني للظاهرة عبر الزمن [كما بينا في شكل (9–1)] .

ويعطي شكل الانتشار فكرة سريعة عن طبيعة الاتجاه العام للظاهرة ومدى ارتباطه بالزمن ومدى تأثير التقلبات الدورية أو الموسمية أو التغيرات العشوائية ، وبالامكان [ومن خلال شكل الانتشار] القيام بعملية مقارنة بين سلسلتين أو أكثر عبر فترات مختلفة من الزمن .

وعملية التمهيد باليد عادةً لا تكون دقيقة مما يقلل الاعتماد عليها وذلك لأن التمهيد باليد يتم بطريقة تقديرية تختلف من شخص لآخر وتعتمد على مهارة الشخص في رسم خط يمر بأكبر عدد ممكن من النقاط ويمثل السلسلة أفضل تمثيل .

مثال (1-8) :

إذا كان لدينا بيانات ربع سنوية لإجمالي الودائع في المصارف السعودية [بآلاف الملايين من الريالات] في الفترة من النصف الأخير من عام *2005* إلى نهاية عام *2007* كالتالي :

السنة	*2005*	*2005*	*2006*	*2006*	*2006*	*2006*	*2007*	*2007*	*2007*	*2007*
الفصل	*3*	*3*	*1*	*2*	*3*	*4*	*1*	*2*	*3*	*4*
الودائع	*195.6*	*196.9*	*200.1*	*205.3*	*207.4*	*215.4*	*232.3*	*222.3*	*222.7*	*226.1*

المطلوب :

رسم شكل الانتشار لهذه البيانات ومن ثم تفسيره وإبراز معالم الاتجاه العام للظاهرة موضع الدراسة .

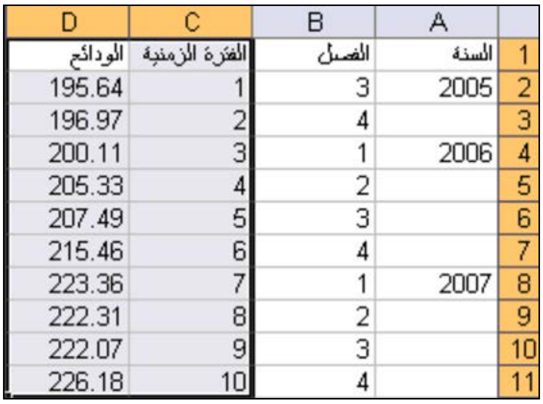
الحل : يتم إدخال البيانات السابقة إلى برنامج الإكسل Excell لتكون بالشكل التالى :

	A	B	C	D
1	السنة	الفصل	الفترة الزمنية	الودائع
2	2005	3	1	195.64
3		4	2	196.97
4	2006	1	3	200.11
5		2	4	205.33
6		3	5	207.49
7		4	6	215.46
8	2007	1	7	223.36
9		2	8	222.31
10		3	9	222.07
11		4	10	226.18

نلاحظ أننا بالإضافة إلى البيانات التى كانت موجودة بالمثال [وهى السنة ، الفصل ، والودائع] تم إضافة عمود يوضح الفترة الزمنية و هى تأخذ القيم : *1 , 2 , ... , 10* .

ويمكن رسم الشكل الانتشارى للبيانات الربع سنوية الخاصة بإجمالي الودائع في المصارف السعودية باتباع الخطوات التالية :

- يتم تحديد العمودين الخاصين بالفترات الزمنية والودائع المطلوب رسم الشكل الانتشارى لهما كما بالشكل التالى :



	A	B	C	D
1	السنة	الفصل	الفترة الزمنية	الودائع
2	2005	3	1	195.64
3		4	2	196.97
4	2006	1	3	200.11
5		2	4	205.33
6		3	5	207.49
7		4	6	215.46
8	2007	1	7	223.36
9		2	8	222.31
10		3	9	222.07
11		4	10	226.18

- نختار من قائمة الرسومات البيانية Wizard Chart رسم الشكل الانتشارى من خلال (xy Scatter) كما بالشكل التالى :

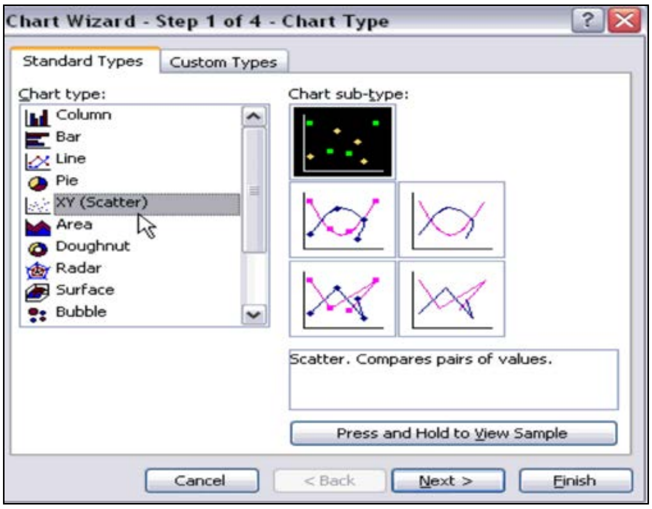


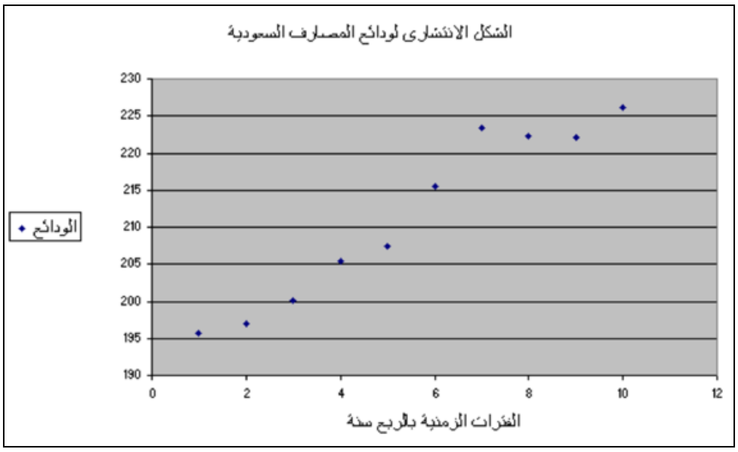
Chart Wizard - Step 1 of 4 - Chart Type

Standard Types | Custom Types

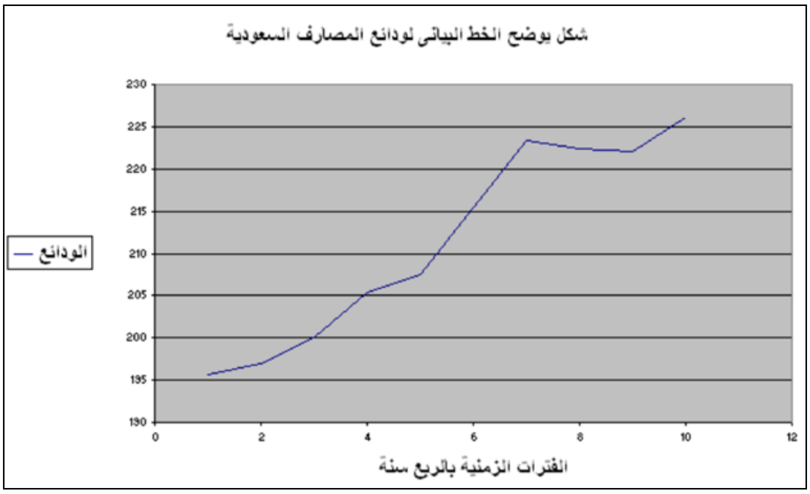
Chart type: XY (Scatter)

Scatter. Compares pairs of values.

- و باستكمال باقى الخطوات يظهر لنا الشكل البيانى التالى :



- كما يمكن رسم الخط البيانى الخاص ببيانات إجمالي الودائع في المصارف السعودية ليكون كما يلى :



وعند رسم شكل الانتشار لهذه البيانات [كما يبدو في الشكل السابق] نستطيع من خلال هذا الرسم رؤية التالي :

- ❖ يتبين لنا أن هناك ارتفاع مستمر في إجمالي الودائع عبر الزمن .
- ❖ الاتجاه العام لبيانات إجمالي الودائع يمكن وصفه بدالة خطية .
- ❖ ميل خط الاتجاه العام لبيانات إجمالي الودائع سيكون موجبا .

(8-5-2) طريقة المتوسطات المتحركة :

تعتمد هذه الطريقة على أخذ متوسطات متتابعة لمجموعات متتابعة ومتداخلة من البيانات [والهدف الأساسي من ذلك هو إزالة التعرجات من خط السلسلة الزمنية] ، وهذه الطريقة أكثر دقة في تحديد خط الاتجاه العام من طريقة شكل الانتشار (التمهيد باليد) .

ويتم حساب المتوسط المتحرك من خلال تطبيق قانون الوسط الحسابي بشكل متتابع لعدد المشاهدات المعطاة لدينا ، مع الأخذ في الاعتبار طول المجموعة التي يتم تقسيم البيانات إليها ، فمثلاً إذا كان لدينا مجموعة من المشاهدات $t_1,t_2,\dots\dots$ واعتبرنا أن طول المجموعة 5 ، إذن يتم إيجاد متوسط المشاهدات t_1,t_2,t_3,t_4,t_5 ، أي :

$$\frac{t_1+t_2+t_3+t_4+t_5}{5}$$

ونضع هذا المتوسط [الذى تم الحصول عليه] أمام المشاهدة التى فى المنتصف [أي أمام المشاهدة t_3 ، ثم نحسب المتوسط من جديد للمشاهدات t_2,t_3,t_4,t_5,t_6 ونضع المتوسط الجديد الذى تم حسابه أمام المشاهدة t_4 ، وهكذا حتى نصل إلى المتوسط الأخير في البيانات المعطاة ، وبعد حساب المتوسطات المتحركة نقوم برسم خط الاتجاه العام من هذه المتوسطات المتحركة ، وقد ينتج عن ذلك خط غير ممهد كما يجب ، وفي هذه الحالة لا

يرسم الخط ، بل تؤخذ متوسطات ثانية للمتوسطات المتحركة الأولى ويرسم الخط من النقاط التي تمثل المتوسطات المتحركة الثانية [لأنها تعطي خطا أكثر تمهيداً] .

ويكون أسلوب المتوسط المتحرك فعالاً عندما تكون بيانات السلسلة الزمنية مستقرة عبر الزمن .

مثال (8–2) : **أوجد** المتوسطات المتحركة بطول (5) للسلسلة الزمنية التالية :

المشاهدة	t_1	t_2	t_3	t_4	t_5	t_6	t_7
قيمتها	7	13	21	23	27	19	17

الحل :

يتم أولاً إيجاد متوسط الخمس مشاهدات الأولى $\left[=\frac{t_1+t_2+t_3+t_4+t_5}{5}\right]$ ونضع الناتج [= 18.2] أمام مركز هذه المشاهدات [وهي المشاهدة t_3] ، ثم نحسب متوسط الخمس مشاهدات التالية [بداية من t_2 حتى t_6] ونضع الناتج [= 20.6] أمام مركز هذه المشاهدات [وهي المشاهدة t_4] ، ثم نحسب متوسط الخمس مشاهدات التالية [بداية من t_3 حتى t_7] ونضع الناتج [= 21.4] أمام مركز هذه المشاهدات [وهي المشاهدة t_5] ، وبذلك نكون قد كونا ما يُسمى بـ "**5 سنوات متوسط متحرك**" .

المشاهدات	القيمة	5 سنوات متوسط متحرك
t_1	7	
t_2	13	
t_3	21	$\frac{7+13+21+23+27}{5}=\frac{91}{5}=18.2$
t_4	23	$\frac{13+21+23+27+19}{5}=\frac{103}{5}=20.6$
t_5	27	$\frac{21+23+27+19+17}{5}=\frac{107}{5}=21.4$
t_6	19	
t_7	17	

مثال (8–3) : البيانات التالية تبين أرباح أحد البنوك [بمليون الريال] للسنوات من 1989 إلى 1999 ميلادية :

السنة	89	90	91	92	93	94	95	96	97	98	99
الأرباح	50.0	36.5	43.0	44.5	38.9	38.1	32.6	38.7	41.7	41.1	33.8

المطلوب :

(أ) حساب 5 سنوات متوسط متحرك مع رسم شكل يوضح هذا المتوسط المتحرك والبيانات الأصلية . علق على الشكل الذي تحصل عليه .

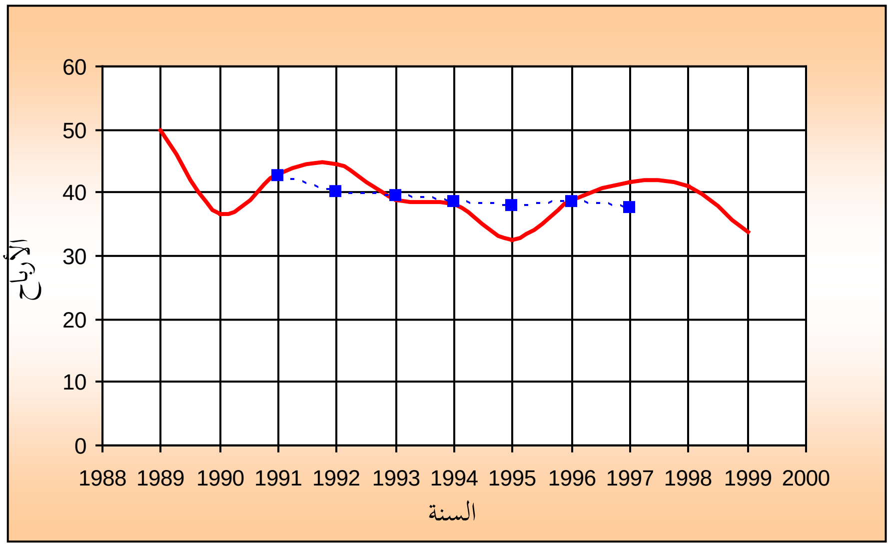
(ب) حساب 4 سنوات متوسط متحرك .

الحل :

(أ) حساب 5 سنوات متوسط متحرك

بنفس طريقة مثال (2-8) يمكن حساب 5 سنوات متوسط متحرك ، والرسم البياني للبيانات الأصلية موضح بالشكل السابق بالخط المتصل ، في حين الرسم البياني للمتوسط المتحرك موضح بالخط المتقطع .

السنة	الأرباح	5 سنوات متوسط متحرك
1989	50.0	
1990	36.5	
1991	43.0	$\frac{50+36.5+43+44.5+38.9}{5}=\frac{212.9}{5}=42.58$
1992	44.5	$\frac{36.5+43+44.5+38.9+38.1}{5}=\frac{201}{5}=40.2$
1993	38.9	$\frac{43+44.5+38.9+38.1+32.6}{5}=\frac{197.1}{5}=39.42$
1994	38.1	$\frac{44.5+38.9+38.1+32.6+38.7}{5}=\frac{192.8}{5}=38.56$
1995	32.6	$\frac{38.9+38.1+32.6+38.7+41.7}{5}=\frac{190}{5}=38$
1996	38.7	$\frac{38.1+32.6+38.7+41.7+41.1}{5}=\frac{192.2}{5}=38.44$
1997	41.7	$\frac{32.6+38.7+41.7+41.1+33.8}{5}=\frac{187.9}{5}=37.58$
1998	41.1	
1999	33.8	



تعليق على الرسم السابق :

من الشكل السابق يمكن ملاحظة كيف أن المتوسط المتحرك قد مهد الخط البياني للبيانات الأصلية مبيناً بشكل واضح خط الاتجاه العام . أيضاً يوضح الشكل السابق كيف أننا فقدنا البيانات عند بداية ونهاية السلسلة الزمنية وهذا يُعد أحد عيوب طريقة المتوسط المتحرك .

(ب) حساب *4* سنوات متوسط متحرك :

بنفس أسلوب الجزء (أ) يمكن حساب *4* سنوات متوسط متحرك فيما عدا أننا نجمع أربعة عناصر في كل مرة بدلاً من خمسة عناصر .

السنة	الأرباح	4 سنوات متوسط متحرك
1989	50.0	
1990	36.5	
		43.5
1991	43.0	
		40.73
1992	44.5	
		41.13
1993	38.9	
		38.53
1994	38.1	
		37.08
1995	32.6	
		37.78

1996	38.7	
		38.53
1997	41.7	
		38.08
1998	41.1	
1999	33.8	

لاحظ أن المجاميع المتحركة تتمركز بين السنوات المتتالية وذلك بخلاف الحالة (أ) وهذه الحالة دائماً تظهر فيما لو أخذنا عدداً زوجياً من السنوات عند حساب المتوسط المتحرك ، بمعنى أننا إذا اعتبرنا أن سنة *1989* (على سبيل المثال) تعبر عن أول يوليو *1989* ، فإن المتوسط المتحرك الأول يتمركز عند أول يناير *1990* .

(8−5−3) طريقة متوسط نصف السلسلة

تعتبر هذه الطريقة أدق من طريقة شكل الانتشار وطريقة المتوسطات المتحركة ، ويمكن حسابها من خلال إتباع الخطوات التالية :

- أعطي للمشاهدات ترقيماً تسلسلياً : *1 , 2 , ...* [سواء كانت المشاهدات قيماً أو غير ذلك] .
- قم بتقسيم السلسلة إلى جزئين متساويين ، فإذا كان عدد المشاهدات الكلي زوجياً يمكن تقسيم السلسة إلى جزئين متساويين ، وإذا كان عدد المشاهدات فردياً أهمل القيمة الوسطى وقم بالتقسيم إلى جزئين متساويين .
- حدد مركز كل جزء [وهو الترقيم الخاص بالقيمة الوسطى لكل جزء] ، لتكن t_1 هي مركز الجزء الأول ، t_2 هي مركز الجزء الثاني .
- احسب متوسط كل جزء [الوسط الحسابي أو الوسيط لقيم هذا الجزء] ، لتكن y_1 هي متوسط الجزء الأول ، y_2 هي متوسط الجزء الثاني .

بذلك نكون قد حصلنا على نقطتين : $A(t_1, y_1)$ [تمثل الجزء الأول] ، $B(t_2, y_2)$ [تمثل الجزء الثاني] ، ويكون خط الاتجاه العام هو الخط AB الواصل بين هاتين النقطتين ، هذا الخط يمكن تمثيله بيانياً ، كما يمكن تحديد معادلته من :

$$\boxed{\frac{y - y_1}{t - t_1} = \frac{y_2 - y_1}{t_2 - t_1}} \qquad (8\text{-}3)$$

ومن معادلة هذا الخط يمكن تحديد القيم الاتجاهية كما يتضح من الأمثلة التالية :

مثال (8−4) :

إذا كان إنتاج مصنع سيارات (بالآلاف) خلال عشر سنوات كالتالي :

السنة t	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
العدد y	53	64	67	60	69	74	67	79	85	90

المطلوب : إيجاد معادلة خط الاتجاه العام بطريقة متوسط نصف السلسلة .

الحل :

نكون الجدول التالي من الجدول الرئيسي :

السنة	السنة بالترقيم (t)	عدد السيارات المنتجة (y)	مركز النصف [متوسط t بالترقيم لكل جزء]	متوسط نصف القيم [متوسط القيم y للجزء]
1998	1	53	$t_1=\frac{1+2+3+4+5}{5}$ $=3$	$y_1=\frac{53+64+67+60+69}{5}$ $=\frac{313}{5}=62.6$
1999	2	64		
2000	3	67		
2001	4	60		
2002	5	69		
2003	6	74	$t_2=\frac{6+7+8\_9+10}{5}$ $=8$	$y_2=\frac{74+67+79+85+90}{5}$ $=\frac{395}{5}=79$
2004	7	67		
2005	8	79		
2006	9	85		
2007	10	90		

وبالتالي يكون خط الاتجاه العام هو الخط AB حيث $A(3,62.6)$ ، $B(8,79)$ ، ومعادلة هذا الخط تتحدد من المعادلة *(8-3)* كالآتي :

$$\frac{y-62.6}{t-3}=\frac{79-62.6}{8-3} \quad \Rightarrow \quad \frac{y-62.5}{t-3}=\frac{16.4}{5}=3.28$$

$$\Rightarrow \quad y-62.6=3.28t-9.84$$

أي أن معادلة خط الاتجاه العام هي :

$$y=3.28t+52.76$$

وهي معادلة خط مستقيم ميله *3.28* ويقطع من المحور الرأسي [محور y] طولاً قدره *52.76* .

مثال (5-8) : أوجد القيم الاتجاهية لبيانات مثال (3-9) [أرباح البنك] مستخدماً طريقة متوسط نصف السلسلة .

الحل : بتقسيم البيانات إلى جزئين متساويين [مع حذف السنة الوسطى *1994* كما هو موضح ، وحساب مركز كل جزء ومتوسط نصف القيم [كما في مثال (8–4)] نحصل على الجدول المبين :

السنة	السنة بالترقيم (t)	الأرباح (بالمليون ريال) (y)	مركز النصف [متوسط t بالترقيم لكل جزء]	متوسط نصف القيم [متوسط القيم y للجزء]
1989	*1*	*50.0*	$t_1 = \frac{1+2+3+4+5}{5}$ $= 3$	$y_1 = \frac{50+36.5+43+44.5+38.9}{5}$ $= \frac{212.9}{5} = 42.58$
1990	*2*	*36.5*		
1991	*3*	*43.0*		
1992	*4*	*44.5*		
1993	*5*	*38.9*		
1994	*6*	*38.1*		
1995	*7*	*32.6*	$t_2 = \frac{7+8+9+10+11}{5}$ $= 9$	$y_2 = \frac{32.6+38.7+41.7+41.1+33.8}{5}$ $= \frac{187,9}{5} = 37.58$
1996	*8*	*38.7*		
1997	*9*	*41.7*		
1998	*10*	*41.1*		
1999	*11*	*33.8*		

وبالتالي يكون خط الاتجاه العام هو الخط AB حيث $A(3,42.58)$ ، $B(9,37.58)$ ، ومعادلة هذا الخط تتحدد من المعادلة *(8-3)* كالآتي :

$$\frac{y-42.58}{t-3} = \frac{37.58-42.58}{9-3} \quad \Rightarrow \quad \frac{y-42.58}{t-3} = \frac{-5}{6} = -0.833$$

$$\Rightarrow \quad y - 42.58 = -0.833t + 2.5$$

أي أن معادلة خط الاتجاه العام هي :

$$y = -0.833t + 45.08$$

وبالتعويض [في معادلة خط الاتجاه العام] عن $t = 1$ نحصل على القيمة الاتجاهية لسنة *1989* ، عن $t = 2$ نحصل على القيمة الاتجاهية لسنة *1990* ، ..... ، عن $t = 6$ نحصل على القيمة الاتجاهية لسنة *1994* ، .... ، عن $t = 11$ نحصل على القيمة الاتجاهية لسنة *1999* ، كما هو مبين بالجدول التالي :

السنة	السنة بالترقيم (t	الأرباح (بالمليون ريال) (y) [القيم الفعلية]	الأرباح (المليون ريال) (y) [القيم الاتجاهية]
1989	*1*	*50.0*	*44.25*
1990	*2*	*36.5*	*43.41*
1991	*3*	*43.0*	*42.58*
1992	*4*	*44.5*	*41.75*
1993	*5*	*38.9*	*40.92*
1994	*6*	*38.1*	*40.08*
1995	*7*	*32.6*	*39.25*
1996	*8*	*38.7*	*38.42*
1997	*9*	*41.7*	*37.58*

1998	10	41.1	36.75
1999	11	33.8	35.92

مثال (8–6) :

أعد حل المثال السابق [مثال (8–5)] باستخدام الوسيط كمتوسط بدلاً من الوسط الحسابي .

الحل :

بترتيب كلٍ من قيم t ، y [تصاعدياً] يمكن معرفة كلٍ من t_1 , t_2 , y_1 , y_2 [حيث تمثل كل قيمة الوسيط بدلاً من الوسط الحسابي] كما هو مبين :

السنة	السنة بالترقيم (t)	الأرباح (بالمليون ريال) (y)	مركز النصف [وسيط t بالترقيم لكل جزء]	متوسط نصف القيم [وسيط القيم y للجزء]
1989	1	50.0	$t_1 = 3$	$36.5 , 38.9 , \underline{\underline{43}} , 44.5 , 50$ $\therefore \quad y_1 = 43$
1990	2	36.5		
1991	3	43.0		
1992	4	44.5		
1993	5	38.9		
1994	6	38.1		
1995	7	32.6	$t_2 = 9$	$32.6 , 33.8 , \underline{\underline{38.7}} , 41.1 , 41.7$ $\therefore \quad y_2 = 38.7$
1996	8	38.7		
1997	9	41.7		
1998	10	41.1		
1999	11	33.8		

وبالتالي يكون خط الاتجاه العام هو الخط AB حيث $A(3,43)$ ، $B(9,38.7)$ ، ومعادلة هذا الخط تتحدد من المعادلة *(8-3)* كالآتي :

$$\frac{y-43}{t-3} = \frac{38.7-43}{9-3} \Rightarrow \frac{y-43}{t-3} = \frac{-4.3}{6} = -0.717$$

أي أن معادلة خط الاتجاه العام يمكن كتابتها على الصورة :

$$y = -0.717\, t + 45.15$$

وبالتعويض [في معادلة خط الاتجاه العام] عن قيم t المناظرة للسنوات المختلفة نحصل على القيم الاتجاهية لكل سنة ، فمثلاً القيمة الاتجاهية لسنة *1997* [حيث $t = 9$] هي :

$$-0.717(9) + 45.15 = 38.697$$

وهكذا لجميع السنوات .

ملحوظة : إن لم يُذكر نوع المتوسط المستخدم فإن هذا يعني استخدام الوسط الحسابي

(8–5–4) طريقة المربعات الصغرى

وتعتبر طريقة المربعات الصغرى أكثر دقة من الطرق السابقة لحساب خط الاتجاه العام وذلك من خلال استخدام أسلوب الانحدار الخطي البسيط المعتمد على طريقة المربعات الصغرى التي تجعل مجموع مربعات انحرافات القيم المقدرة عن القيم الفعلية أقل ما يمكن وذلك من خلال العلاقة التالية [معادلة *(8-3)*/الفصل الثامن] :

$$\boxed{\hat{y} = b_0 + b_1 t} \qquad (8\text{-}4)$$

حيث $\hat{y}$ هي القيمة الاتجاهية للسلسلة الزمنية في الفترة t ، b_0 يمثل طول الجزء المقطوع من المحور الرأسي [محور y] وهو ما يمثل ثابت انحدار y على t ، b_1 الذي يمثل ميل خط الاتجاه العام وهو أيضاً ميا انحدار y على t . وتتحدد قيم الثوابت b_0 , b_1 من علاقات مشابهة تماماً للعلاقات *(7-8)* , *(7-7)* [الدرس الحادي عشر/الباب السابع] وهي :

$$\boxed{b_1 = \frac{n\sum ty - \left(\sum t\right)\left(\sum y\right)}{n\sum t^2 - \left(\sum t\right)^2}} \qquad (8\text{-}5)$$

$$\boxed{b_0 = \frac{\sum y}{n} - b_1 \frac{\sum t}{n} = \bar{y} - b_1 \bar{t}} \qquad (8\text{-}6)$$

حيث y هي القيمة الفعلية للسلسلة الزمنية في الفترة t ، n هي عدد الفترات .

مثال (7–8) :

بدراسة إحدى الظواهر الاجتماعية والمتمثلة فى العنف الأسرى بإحدى المدن ، تبين أن تطور أعداد الأسر التي تعاني من هذه الظاهرة كانت كالتالي [خلال مدة الدراسة] :

السنة	2004	2005	2006	2007	2008	2009	2010
عدد الأسر	17	25	33	41	39	48	53

المطلوب تقدير معادلة الاتجاه العام لتطور أعداد الأسر المعرضة لظاهرة العنف الأسرى بهذه المدينة . **ما هو** عدد الأسر (المتوقع) المعرضة لهذه الظاهرة فى عام *2013* ؟ .

الحل:

حتى يمكن تقدير معادلة الاتجاه العام لتطور أعداد الأسر المعرضة لظاهرة العنف الأسرى بهذه المدينة لابد من إعداد الجدول التالى علي اعتبار أن السنة الأولى تكون قيمة t فيها تساوى *1* والسنة الثانية تكون قيمتها *2* ، وهكذا كما يلى :

السنوات	t	عدد الأسر y	ty	t^2
2004	1	17	17	1
2005	2	25	50	4
2006	3	33	99	9

2007	4	41	164	16
2008	5	39	195	25
2009	6	48	288	36
2010	7	53	371	49
	$\sum t=28$	$\sum y=256$	$\sum ty=1184$	$\sum t^2=140$

وحيث أن عدد المشاهدات n = 7 ، إذن [من العلاقات (8-5) , (8-6)] :

$$b_1=\frac{n\sum ty-\sum t\sum y}{n\sum t^2-\left(\sum t\right)^2}=\frac{7(1184)-(28\times 256)}{7(140)-(28^\wedge 2)}=\frac{1120}{196}=5.714$$

ويدل ذلك علي أن معدل التزايد السنوى في الأسر المعرضة للعنف الأسرى 5.714 أسرة [أي ما بين 5 ، 6 أسر]

$$b_0=\frac{\sum y}{n}-b_1\frac{\sum t}{n}=\frac{256-(5.714\times 28)}{7}=13.715$$

وعلى ذلك تكون معادلة الاتجاه العام على الصورة :

$$\hat{y}=b_0+b_1t=13.715+5.714\,t$$

عدد الأسر (المتوقع) المعرضة لهذه الظاهرة فى عام 2013 ؟

وحتى يمكن تقدير عدد الأسر (المتوقع) المعرضة لهذه الظاهرة فى عام 2013 لابد من تحديد قيمة t المناظرة لهذه السنة وهذا يمكن تحديده بسهولة حيث أن السنة 2004 تناظر القيمة $t=1$ ، إذن بإضافة 9 [الذي يساوي الفرق بين 2013 ، 2004] على القيمة $t=1$ تكون قيمة t المناظرة للسنة 2013 هي :

$$t=1+9=10$$

وبالتعويض في معادلة الاتجاه العام السابقة عن $t=10$ نحصل على :

$$\hat{y}=13.715+5.714\,(10)=70.855$$

أي أن عدد الأسر المتوقع تعرضها للعنف الأسرى في عام 2013 يبلغ 70.855 [أى ما يقرب من 71 أسرة] .

(8–6) التغيرات الموسمية Seasonal Variations

في الواقع أننا لايمكن تحديد التغير في قيم ظاهرة ما فقط من الاتجاه العام وحده ، وذلك لأن هناك أنواع أخرى من التغيرات تتعرض لها هذه الظاهرة ومنها التغيرات الموسمية .

والتغيرات الموسمية هي نتيجة طبيعية لاختلاف الظروف بشكل منتظم مما يؤثر على اختلاف رغبات الناس تبعاً لعوامل عديدة منها الزمان والمكان ، ويمكن تعريفها بأنها التغيرات التي تطرأ على الظاهرة على مدار المواسم

المختلفة للفترة الزمنية موضوع القياس (الموسم) ، فهي قد تكون يومية [مثل درجات الحرارة حيث تبدأ منخفضة ثم تأخذ في الارتفاع مع تقدم النهار حتى تبلغ الحد الأعلى لها ثم بعد ذلك تأخذ في الهبوط تدريجياً] فهذه التغيرات مداها يوم واحد ، وقد تكون أسبوعية [مثل ايداع النقود في البنك حيث يزداد يوم الخميس قبل انتهاء ساعات دوام البنك عنه في باقي أيام الأسبوع] ، وقد تكون شهرية [مثل حركة السحب من البنك ، فهي تزداد في نهاية الشهر وذلك عند صرف رواتب الموظفين] .

مما سبق نرى أن التغيرات الموسمية تحدث في مواعيد زمنية محددة ولا يلبث هذا التغير أن يستعيد سيرته الأولى في نفس المواعيد وعلى مدار نفس الفترة الزمنية ، والتغير الموسمي يعتبر أبسط أنواع التغيرات في السلاسل الزمنية ؛ حيث يشتمل على نماذج متكررة بانتظام ، وهي تغيرات تتميز بالطبيعة الدورية بشرط أن لايزيد طول الدورة المتكررة عن سنة واحدة كحد أعلى .

وتكمن أهمية دراسة التغيرات الموسمية في تحليل السلسلة الزمنية للظاهرة خاصةً فيما يتعلق بالتخطيط لعمليات الإنتاج أو الأوقات المناسبة للإعلانات عن السلع أو التوسع في المشاريع ، فالتغيرات الموسمية بشكل عام تساعد على الكشف عن :

- الأوقات المناسبة للتغيير .
- مسببات التغيير .
- الاستعدادات المناسبة لمواجهة التغيير .

ويتم قياس التغيرات الموسمية عن طريق إيجاد قيمة الظاهرة في كل موسم من المواسم التي تتعرض لها الظاهرة للتغير ثم تنسب كل قيمة للمتوسط العام لقيم هذه الظاهرة ، إذ يتم اعتبار المتوسط العام *(100%)* فنحصل على أرقام تدل على مدى التغيرات للظاهرة ، هل هي فوق المتوسط أو دونه [ومثال على ذلك ما يذاع عن درجات الحرارة المتوقعة في النشرات الجوية من أنها فوق المتوسط أو دون المتوسط]

ولحساب الآثار الموسمية هناك طريقتان :

- الأولى : طريقة النسب للمتوسط المتحرك .
- الثانية : من خلال العلاقة *(8-2)* :

$$y_t = T_t \times C_t \times S_t \times R_t$$

والتي منها يمكن إيجاد **القيم المخلصة من أثر الاتجاة العام** وذلك بقسمة طرفي المعادلة السابقة على T_t [والتي تمثل تأثير الاتجاه العام] فنحصل بالتالي على المعادلة التالية [وهي التي تعطي القيم المخلصة من أثر الاتجاه العام]

$$\boxed{\frac{y_t}{T_t} = C_t \times S_t \times R_t} \qquad (8\text{-}7)$$

لكن هذه المعادلة تستلزم معرفة قيمة الاتجاه العام والقيم الدورية والقيم العشوائية ، ومن خلال معرفة هذه القيم نستطيع معرفة قيم التغيرات الموسمية ، ولكون هذه الطريقة تحتاج إلى معرفة جميع هذه القيم سنستخدم الطريقة الأولى (طريقة النسب للمتوسط المتحرك) لسهولة التعامل معها . والمثال التالي يوضح ذلك :

مثال (8-8) :

الجدول التالي يبين إنتاج إحدى الشركات خلال ثلاث سنوات [كمية الإنتاج مأخوذة كل ثلاثة شهور ، والسنة مقسمة إلى أربعة أرباع] حيث الإنتاج بآلاف الوحدات :

ربع السنة	2008	2009	2010
الأول	3	4	8
الثاني	7	5	10
الثالث	9	6	12
الرابع	2	4	6

المطلوب :

(أ) تقدير معالة الاتجاه العام للعلاقة بين الانتاج و الزمن ؟

(ب) تقدير القيم الإتجاهية المقابلة للقيم الأصلية للأنتاج ؟

(ج) إيجاد القيم المخلصة من أثر الأتجاة العام ؟

(د) تحديد تأثير كل موسم ؟

(هـ) تقدير الانتاج المتوقع سنة 2012 ؟.

الحل :

(أ) تقدير معادلة الاتجاة العام للعلاقة بين الانتاج والزمن

يتم أولاً إدخال البيانات السابقة مع إضافة عنصر الزمن t ، ثم يتم حساب الأعمدة ty , t^2 وإيجاد المجاميع الازمة لحساب معامل الانحدار b_1 كما يلي :

السنة	الربع	الزمن t	الانتاج y	ty	t^2
2008	الأول	1	3	3	1
	الثاني	2	7	14	4
	الثالث	3	9	27	9
	الرابع	4	2	8	16
2009	الأول	5	4	20	25
	الثاني	6	5	30	36

	الثالث	7	6	42	49
	الرابع	8	4	32	64
2010	الأول	9	8	72	81
	الثانى	10	10	100	100
	الثالث	11	12	132	121
	الرابع	12	6	72	144
		$\sum t = 78$	$\sum y = 76$	$\sum ty = 552$	$\sum t^2 = 650$

وعدد الفترات الزمنية n هنا يساوي 18 ، إذن من هذه البيانات يمكن [من خلال المعادلات (8-6) , (8-5)] حساب الآتي :

$$b_1 = \frac{n\sum ty - \sum t \sum y}{n\sum t^2 - \left(\sum t\right)^2} = \frac{12(552) - (78\times 76)}{12(650) - (78\text{^}2)} = 0.40559$$

$$b_0 = \frac{\sum y}{n} - b_1 \frac{\sum t}{n} = \frac{76 - (0.40559\times 78)}{12} = 3.69697$$

وبالتالي تكون معادلة خط الاتجاه العام بين الإنتاج y والزمن t على الصورة :

$$\hat{y} = b_0 + b_1 t = 3.69697 + 0.40559\ t$$

[لاحظ أن معدل التزايد كل فترة ربع سنة هو *0.40559* ألف وحدة]

(ب) تقدير القيم الإتجاهية المقابلة للقيم الأصلية للإنتاج

يمكن إيجاد القيم الاتجاهية بالتعويض فى معادلة خط الاتجاه العام السابقة عن قيم t من 1 حتى 12 فنحصل على القيم الاتجاهية المبينة بالجدول التالي :

السنة	الربع	الزمن t	الانتاج [القيم الحقيقية] y	القيم الاتجاهية $\hat{y} = T$	القيم المخلصة من أثر الاتجاه العام $\frac{y}{\hat{y}} = \frac{y}{T}$
2008	الأول	1	3	4.10256	$\frac{3}{4.10256} \cong 0.7313$
	الثاني	2	7	4.50815	$\frac{7}{4.50815} \cong 1.5527$
	الثالث	3	9	4.91374	1.8316

	الرابع	4	2	5.31933	0.376
2009	الأول	5	4	5.72492	0.6987
	الثانى	6	5	6.13051	0.8156
	الثالث	7	6	6.5361	0.918
	الرابع	8	4	6.94169	0.5762
2010	الأول	9	8	7.34728	1.0888
	الثانى	10	10	7.75287	1.2898
	الثالث	11	12	8.15846	1.4709
	الرابع	12	6	8.56405	0.7006

(ج) إيجاد القيم المخلصة من أثر الاتجاه العام :

ويتم حساب القيم المخلصة من أثر الاتجاه العام بقسمة قيم الظاهرة الاصلية على القيم الاتجاهية ، فتكون النتيجة كما بالجدول السابق .

(د) إيجاد تأثير كل موسم :

حتى يمكن إيجاد تأثير كل موسم نعيد ترتيب القيم المخلصة من أثر الاتجاه العام السابق الحصول عليها كما يلى :

الموسم	2008	2009	2010
الأول	0.7313	0.6987	1.0888
الثانى	1.5527	0.8156	1.2898
الثالث	1.8316	0.918	1.4709
الرابع	0.376	0.5762	0.7006

ثم يتم إيجاد متوسط القيم المخلصة من أثر الاتجاه العام لكل ربع للتعبير عن أثر ذلك الموسم ، فمثلاً تأثيرالربع الأول هو :

$$\frac{0.7313+0.6987+1.0888}{3}=0.8396$$

وهكذا لباقى المواسم فتكون النتيجة كما يلى :

الموسم	2008	2009	2010	تأثير الموسم
الأول	0.7313	0.6987	1.0888	0.8396
الثانى	1.5527	0.8156	1.2898	1.2194

الثالث	1.8316	0.918	1.4709	1.4068
الرابع	0.376	0.5762	0.7006	0.5509
			المجموع	4.0167

ونلاحظ أن مجموع تأثيرات المواسم (الدليل الموسمى) 4.0167 اى 401.67 % وحيث أنه يوجد 4 مواسم ، لذا فإن مجموع تأثيرات المواسم لابد أن تساوى 400% ، وبالتالي لابد من تعديل قيم الدليل الموسمى بمعامل تصحيح قدره :

$$\frac{4}{4.0167}$$

فنحصل على تأثير الموسم المعدل كما هو مبين :

الموسم	2008	2009	2010	تأثير الموسم	تأثير الموسم المعدل
الأول	0.7313	0.6987	1.0888	0.8396	$0.8396\times\frac{4}{4.0167}$ $=0.836109$
الثانى	1.5527	0.8156	1.2898	1.2194	1.21433
الثالث	1.8316	0.918	1.4709	1.4068	1.400951
الرابع	0.376	0.5762	0.7006	0.5509	0.54861
			المجموع	4.0167	4

(هـ) تقدير الانتاج المتوقع سنة 2012 :

نلاحظ أن قيم t فى الربع الاخير سنة 2010 بلغت 12 لذلك يتم الزيادة عليها سنة 2011 لتكون 13 , 14 , 15 , 16 خلال المواسم الاربع ، ولذلك تكون القيم خلال سنة 2012 هى 17 ، 18 , 19 , 20 ، والتى يتم التعويض بها معادلة الاتجاه العام للحصول على القيم الاتجاهيه ، كما يمكن تقدير القيم المتنبأ بها لكل ربع كما يلى :

القيم المتنبأ بها للموسم = القيمة الاتجاهية x تأثير الموسم المعدل

وعلى ذلك يمكن تقدير الانتاج المتوقع سنة 2012 كما يلى :

الموسم	t	القيمة الاتجاهية	تأثير الموسم المعدل	الانتاج المتوقع
الأول	17	10.592	0.836109	8.856069
الثانى	18	10.99759	1.21433	13.35471

الثالث	19	11.40318	1.400951	15.9753
الرابع	20	11.80877	0.54861	6.478404
			المجموع	44.66448

ويتضح لنا أن الانتاج المتوقع سنة *2012* هو *44664.48* وحدة .

(7–8) التغيرات الدورية Cyclical Variations

ويعرف هذا النوع من التغيرات **بدورات الأعمال** ، وهذا النوع يختلف عن التغيرات الموسمية في أنها تمتد لفترة زمنية أطول من سنة ، وتنشأ هذه التغيرات عن ظروف عامة تعزى إلى العوامل التي تتحكم في الحياة الاقتصادية للبلاد ، بينما تعزى التغيرات الموسمية لأسباب مردها الجو أو العادات الاجتماعية .

ويهتم الباحثون الاقتصاديون ورجال الأعمال بالتغيرات الدورية لغايات التخطيط لمواجهة المشاكل التي قد تنشأ عن حدوثها ، وقد تمتد بعض التغيرات الدورية إلى 50 سنة وهذه دورة طويلة ، أما الدورة المتوسطة فتمتد بين 8 إلى 12 سنة ، أما الدورة القصيرة فتكون بين *3* إلى *4* سنوات ، وتقع التقلبات الدورية أعلى و أسفل خط الاتجاه العام .

ويتم حساب التغيرات الدورية عن طريق تخليص السلسلة الزمنية من أثر الاتجاه العام ومن أثر التغيرات الموسمية وذلك من خلال العلاقة *(9-2)* :

$$y_t = T_t \times C_t \times S_t \times R_t$$

وذلك بقسمة طرفي المعادلة السابقة على $T_t S_t$ [والتي تمثل تأثير كلٍ من الاتجاه العام والتغيرات الموسمية] فنحصل بالتالي على المعادلة التالية :

$$\boxed{\frac{y_t}{T_t S_t} = C_t R_t} \qquad (8\text{-}8)$$

(8–8) التغيرات العشوائية أو الفجائية

Random (Irregular) Variations

وهي أحداث تنتج تغيرات تستمر لفترة قصيرة من الزمن تؤثر على السلسلة الزمنية بشكل عشوائي أو مفاجئ وغير منتظم ، فقد تكون هذه التغيرات ناتجة عن حدوث ظواهر طبيعية مثل الزلازل والبراكين أو حروب ونحوها ، لذا فهذا النوع من التغيرات من الصعب التنبؤ بها ومن الصعب كذلك تحديد حجم هذه التغيرات ومدى تأثيرها على الظاهرة موضع الدراسة . وتمتاز هذه التغيرات بعدد من المميزات منها :

- إنها لا تحدث وفقا لقاعدة أو قانون .
- قد تتكرر أو لا تتكرر .
- تأثيرها غير ثابت فمرة تأثر بالنقص ومرة بالزيادة .

- لا تستمر طويلا لذا يطلق عليها اسم التغيرات قصيرة الأجل .

ويمكن حساب التغيرات العشوائية من خلال العلاقة *(9-2)* :

$$y_t = T_t \times C_t \times S_t \times R_t$$

وذلك بقسمة طرفي المعادلة السابقة على $T_tS_tC_t$ [والتي تمثل تأثير كلٍ من الاتجاه العام والتغيرات الموسمية والتغيرات الدورية] فنحصل بالتالي على المعادلة التالية :

$$\boxed{\frac{y_t}{T_tS_tC_t} = R_t} \qquad (8\text{-}9)$$

ومن الناحية العملية وُجِد أن هذه التغيرات العشوائية (أو الفجائية) تتجه دائماً على أن تكون ذات حجم صغير وأنها غالباً تتجه إلى أن تتبع نمط التوزيع الطبيعي .

ملخص للدرس الثاني عشر [الباب الثامن : السلاسل الزمنية]

- **السلسلة الزمنية** هي مجموعة من المشاهدات أُخِذت في فترات زمنية محددة ، عادةً على فترات متساوية ، وتمثل السلسلة الزمنية المتضمنة لمتغير y بتكوين الشكل البياني y مقابل الزمن t [شكل الانتشار] .
- وتُصنف السلسلة الزمنية إلى **سلسلة زمنية فترية** وهي السلسلة التي تتكون من بيانات كمية لمستوى الظاهرة عن فترات محددة من الزمن ، و**سلسة زمنية لحظية** وهي السلسلة التي تتكون من مستويات للظاهرة مقاسة في لحظات .
- وتتألف السلسلة الزمنية من أربعة عناصر أساسية هي :
  - ❖ **التحركات طويلة المدى[الاتجاه العام] T** [وتسمى "**القيم الإتجاهية**"] وهي تشير إلى الاتجاه العام الذي يظهر به الشكل البياني للسلسلة الزمنية على مدى فترة طويلة من الزمن .
  - ❖ **التغيرات الموسمية S** [وتسمى "**القيم الموسمية**"] وهي تشير إلى النمط المتماثل لحركة السلسلة الزمنية في الأشهر المتقابلة خلال السنوات المتتالية .
  - ❖ **التغيرات الدورية C** [وتسمى "**القيم الدورية**"] وهي تشير إلى الذبذبات طويلة المدى حول خط (أو منحنى) الاتجاه العام .
  - ❖ **التغيرات العشوائية** أو **الفجائية R** [وتسمى "**القيم العشوائية**"] ، وهي أحداث تنتج تغيرات تستمر لفترة قصيرة من الزمن ومن الصعب التنبؤ بها ومن الصعب كذلك تحديد حجم هذه التغيرات ومدى تأثيرها على الظاهرة موضع الدراسة .
- ويمكن وصف السلسلة الزمنية بأحد النموذجين :
  - ❖ نموذج الجمع : $y_t = T_t + C_t + S_t + R_t$
  - ❖ نموذج الضرب : $y_t = T_t \times C_t \times S_t \times R_t$
- ويمكن تقدير الاتجاه العام بعدة طرق منها :
  - ❖ طريقة الانتشار (التمهيد باليد) .
  - ❖ طريقة المتوسطات المتحركة .
  - ❖ طريقة متوسط نصف السلسلة .
  - ❖ طريقة المربعات الصغرى .

تدريبات عملية

تدريب عملي (1) :

أوجد المتوسطات المتحركة بطول (3) للسلسلة الزمنية التالية :

المشاهدة	t_1	t_2	t_3	t_4	t_5	t_6	t_7
قيمتها	12	18	25	28	33	26	18

تدريب عملي (2) :

إذا كان عدد الطلاب الملتحقين بكلية الأداب (بالآلاف) خلال عشر سنوات [من عام 1422 إلى 1431 هجري] كالتالي :

السنة t	22	23	24	25	26	27	28	29	30
العدد y	1.5	2	3	`3	3.3	4.1	5	5.2	5.6

المطلوب : إيجاد معادلة خط الاتجاه العام بطريقة متوسط نصف السلسلة ؟

تدريب عملي (3) :

بدراسة ميزانية الأسرة تبين أن متوسط الأنفاق الشهرى للأسرة (بالألف ريال) فى أحد المناطق كانت كما يلى خلال مدة الدراسة :

السنة	2004	2005	2006	2007	2008	2009	2010
متوسط الإنفاق الشهري	5	7.3	7.7	8.1	8.7	9.3	10.4

المطلوب: تقدير معادلة الاتجاه العام لتطور لمتوسط الأنفاق الشهرى للأسرة بهذه المنطقة. **ما هو** متوسط الانفاق الشهرى للأسرة المتوقع فى عام 2015 ؟ .

تديب عملي (4) :

إذا كان لدينا مبيعات إحدى الشركات خلال ثلاث سنوات ، وكانت كمية المبيعات مأخوذة كل ثلاثة شهور [السنة مقسمة إلى أربعة أرباع] والمبيعات بآلاف الوحدات كما يبدوا ذلك من الجدول التالي :

ربع السنة	2008	2009	2010
الأول	5	8	9
الثانى	6	11	10
الثالث	4	7	8
الرابع	3	5	7

المطلوب :

(أ) تقدير معادلة الاتجاة العام للعلاقة بين المبيعات و الزمن .

(ب) تقدير القيم الإتجاهية المقابلة للقيم الأصلية للمبيعات .

(ج) إيجاد القيم المخلصة من أثر الأتجاة العام .

(د) تحديد تأثير كل موسم .

(هـ) تقدير المبيعات المتوقع سنة 2013 .

تدريبات (12)

الإجابة النهائية لجميع التمرينات موجودة في نهاية التدريب

اختر الإجابة الصحيحة

(1) .... تشير إلى الاتجاه العام الذي يظهر به الشكل البياني للسلسلة الزمنية على مدى فترة طويلة من الزمن .

(أ) التحركات طويلة المدى (ب) التغيرات الموسمية

(ج) التغيرات الدورية (د) التغيرات العشوائية

(2) .... تشير إلى النمط المتماثل لحركة السلسلة الزمنية في الأشهر المتقابلة خلال السنوات المتتالية .

(أ) التحركات طويلة المدى (ب) التغيرات الموسمية

(ج) التغيرات الدورية (د) التغيرات العشوائية

(3) .... تشير إلى الذبذبات طويلة المدى حول خط (أو منحنى) الاتجاه العام .

(أ) التحركات طويلة المدى (ب) التغيرات الموسمية

(ج) التغيرات الدورية (د) التغيرات العشوائية

(4) .... تشير إلى الاتجاه العام الذي يظهر به الشكل البياني للسلسلة الزمنية على مدى فترة طويلة من الزمن .

(أ) التحركات طويلة المدى (ب) التغيرات الموسمية

(ج) التغيرات الدورية (د) التغيرات العشوائية

(5) "اشتعال النار في مصنع أدى إلى تأخير الإنتاج ثلاثة أسابيع" يعتبر من :

(أ) التحركات طويلة المدى (الاتجاه العام) (ب) التغيرات موسمية .

(ج) التغيرات دورية . (د) التغيرات عشوائية (فجائية) .

(6) "عهد من الرفاهية" يعتبر من :

(أ) التحركات طويلة المدى (الاتجاه العام) (ب) التغيرات موسمية .

(ج) التغيرات دورية . (د) التغيرات عشوائية (فجائية) .

(7) "مبيعات ما بعد عيد الأضحى المبارك في أحد المتاجر" يعتبر من :

(أ) التحركات طويلة المدى (الاتجاه العام) (ب) التغيرات موسمية .

(ج) التغيرات دورية . (د) التغيرات عشوائية (فجائية) .

(8) "الحاجة إلى زيادة إنتاج القمح في المملكة نتيجة للزيادة المستمرة في عدد السكان" يعتبر من :

(أ) التحركات طويلة المدى (الاتجاه العام) (ب) التغيرات موسمية .

(ج) التغيرات دورية . (د) التغيرات عشوائية (فجائية) .

(9) "عدد ملليمترات الأمطار التي تهبط في الشهر على مدينة معينة خلال فترة 5 سنوات" يعتبر من :

(أ) التحركات طويلة المدى (الاتجاه العام) (ب) التغيرات موسمية .

(ج) التغيرات دورية . (د) التغيرات عشوائية (فجائية) .

(10) "كساد مؤقت" يُعتبر من :

(أ) التحركات طويلة المدى (الاتجاه العام) (ب) التغيرات موسمية .

(ج) التغيرات دورية . (د) التغيرات عشوائية (فجائية) .

(11) "زيادة العمالة خلال أشهر الصيف" يُعتبر من :

(أ) التحركات طويلة المدى (الاتجاه العام) (ب) التغيرات موسمية .

(ج) التغيرات دورية . (د) التغيرات عشوائية (فجائية) .

(12) "انخفاض معدل الوفيات الراجع للتقدم العلمي" يُعتبر من :

(أ) التحركات طويلة المدى (الاتجاه العام) (ب) التغيرات موسمية .

(ج) التغيرات دورية . (د) التغيرات عشوائية (فجائية) .

(13) "إضراب في أحد المصانع" يُعتبر من :

(أ) التحركات طويلة المدى (الاتجاه العام) (ب) التغيرات موسمية .

(ج) التغيرات دورية . (د) التغيرات عشوائية (فجائية) .

(14) "الزيادة المستمرة في الطلب على سيارات الركوب الصغيرة" يُعتبر من :

(أ) التحركات طويلة المدى (الاتجاه العام) (ب) التغيرات موسمية .

(ج) التغيرات دورية . (د) التغيرات عشوائية (فجائية) .

(15) إذا كان لدينا الأرقام 2 , 6 , 1 , 5 , 3 , 7 , 2 فإن الوسط المتحرك بطول 3 يُعطى بـ :

(أ) $\frac{26}{7}$ (ب) المتتابعة 2 , 5 , 2

(ج) المتتابعة 1 , 7   (د) المتتابعة 3 , 4 , 3 , 5 , 4

(16) عند حساب متوسط متحرك بطول 5 للسلسلة $t_1 ,t_2 ,\cdots\cdots,t_{11}$ ، فإن أول قيمة في متتابعة المتوسط تُوضع :

(أ) تحت القيمة t_1 (ب) تحت القيمة t_3

(ج) تحت القيمة t_5 (د) تحت القيمة t_6

(17) عند حساب متوسط متحرك بطول 5 للسلسلة $t_1 ,t_2 ,\cdots\cdots,t_{10}$ ، فإن أول قيمة في متتابعة المتوسط تُوضع :

(أ) تحت القيمة t_1 (ب) بين القيمتين t_5 ,t_6

(ج) تحت القيمة t_5 (د) بين القيمتين t_2 ,t_3

خاص بالأسئلة من (18) إلى (21) :

في دراسة لتحديد خط الاتجاه العام لانتاج أحد المصانع من السيارات بواسطة طريقة نصف متوسط السلسلة كانت البيانات التالية خلال الفترة من 2005 إلى 2010 :

السنة	السنة بالترقيم (t)	y	متوسط t	متوسط y
2005	1	50	$t_1 = ?$	$y_1 = 58$
2006	2	?		
2007	3	64		
2008	4	65	$t_2 = 5$	$y_2 = ?$
2009	5	65		
2010	6	80		

من هذا الجدول ، أجب على التالي :

(18) عدد السيارات المنتجة خلال سنة 2006 يساوي :

(أ) 55   (ب) 57   (ج) 60   (د) 62

(19) قيمة t_1 المبينة بالجدول تساوي :

(أ) 1   (ب) 2   (ج) 3   (د) 2006

(20) قيمة y_2 المبينة بالجدول تساوي :

(أ) 58　　(ب) 65　　(ج) 70　　(د) 80

(21) معادلة خط الاتجاه العام بطريقة متوسط نصف السلسلة هي :

(أ) $\frac{y-58}{t-t_1}=\frac{y_2-58}{5-t_1}$　　(ب) $\frac{y}{t}=\frac{y_2-58}{5-t_1}$

(ج) $\frac{y-58}{5-t_1}=\frac{y_2-58}{t-t_1}$　　(د) $\frac{y}{5-t_1}=\frac{y_2-58}{t}$

خاص بالأسئلة من (22) إلى (25) :

إذا كان لدينا مبيعات لإحدى الشركات خلال سنتين ، وكانت كمية المبيعات مأخوذة كل ثلاثة شهور [السنة مقسمة إلى أربعة أرباع] والمبيعات بآلاف الوحدات ، وبعد تخليص المبيعات من أثر الاتجاه العام للعلاقة بين المبيعات والزمن كانت النتائج التالية :

الموسم	القيم المخلصة من أثر الاتجاه العام		تأثير الموسم	تأثير الموسم المعدل
	2009	2010		
الأول	0.6	0.8	A	
الثاني	1.4	B	1.1	
الثالث	1.7	0.9	1.3	D
الرابع	0.4	0.6	0.5	
			C	

من هذا الجدول [غير المكتمل] أجب على التالي :

(22) قيمة A بالجدول المرافق تساوي :

(أ) 0.6　　(ب) 0.7　　(ج) 0.8　　(د) 1

(23) قيمة B بالجدول المرافق تساوي :

(أ) 0.8　　(ب) 1　　(ج) 1.2　　(د) 1.4

(24) قيمة C بالجدول المرافق تساوي :

(أ) 2.8　　(ب) 3.2　　(ج) 3.6　　(د) 4

25) قيمة C بالجدول المرافق تساوي :

أ) *0.56*　　ب) *0.78*　　ج) *1.22*　　د) *1.44*

إجابة تدريبات (12) :

1) أ	2) ب	3) ج	4) د	5) د
6) ج	7) ب	8) أ	9) ب	10) ج
11) ب	12) أ	13) د	14) أ	15) د
16) ج	17) د	18) ج	19) ب	20) ج
21) أ	22) ب	23) أ	24) ج	25) د

عناصر الدرس

الباب التاسع : الأرقام القياسية

- تعريف الرقم القياسي
- فترة الأساس
- الأرقام القياسية للأسعار
- دور الأرقام القياسية في حساب معدلات التضخم
- فوائد الأرقام القياسية واستعمالاتها
- الرقم القياسي المرجح
- منسوب السعر لسلعة واحدة (ظاهرة بسيطة)
- منسوب السعر لمجموعة من السلع (ظاهرة معقدة)
- الرقم القياسي للكميات
- الرقم القياسي للدخل الحقيقي
- ملاحظات عامة على الأرقام القياسية

الباب التاسع : الأرقام القياسية

(9–1) تعريف الرقم القياسي

الرقم القياسي هو مؤشر إحصائي (رقم نسبي) يستخدم في قياس التغير النسبي الذي يطرأ على ظاهرة من الظواهر الاقتصادية أو الاجتماعية ، فهو يستخدم لقياس التغير في أسعار السلع أو في حجم انتاجها أو في كميات المبيعات منها أو في حجم السكان أو أجور العمال ، وذلك وفقاً لأساس معين سواء كان هذا الأساس فترة زمنية معينة أو مكاناً معيناً ، وتُسمى أحياناً المجموعة من الأرقام القياسية لسنوات أو أماكن مختلفة ، وما إلى ذلك ، **بالسلسلة القياسية** .

(9–2) فترة الأساس :

الأساس هو فترة زمنية معينة أو مكاناً معيناً [أو أي خاصية أخرى مثل الدخل ، الوظيفة ، وغير ذلك] ، وعادة تكون فترة الأساس فترة سابقة للفترة التي نريد مقارنتها ، وفي حالات نادرة جداً قد تكون فترة الأساس فترة لاحقة لفترة المقارنة .

ويجب أن تمتاز فترة الأساس بما يلي : الاستقرار الاقتصادي – خلوها من العوامل المؤثرة على الأسعار (مثل الحروب) – وأن تكون بعيدة جدًا عن سنوات المقارنة .

أما عند اختيار مكان الأساس ، فلابد أن يكون لهذا المكان أهمية خاصة وأن يكون مركزاً أساسياً لإنتاج السلعة المراد استخراج الرقم القياسي لها .

(9–3) الأرقام القياسية للأسعار *Price Index Numbers* :

تعتبر الأرقام القياسية للأسعار من أهم أنواع الأرقام القياسية وأكثرها شيوعاً ، فهي (أي الأرقام القياسية للأسعار) تساهم في قياس التغير في المستوى العام للأسعار أو التغير في تكاليف المعيشة في فترة زمنية معينة مقارنة بفترة زمنية أخرى .

ومن أشهر الأرقام القياسية للأسعار ما يلي :

- مؤشر أسعار المستهلكين Consumer Price Index ويرمز له (CPI) .
- مخفض الناتج القومي الإجمالي Gross National Product Deflator .
- مؤشر اسعار المنتجين Producer Price Index ويرمز له (PPI) .
- مخفض الناتج المحلي الإجمالي Gross Domestic Product Deflator .

- مؤشر أسعار الأسهم .

ويختلف مؤشر اسعار المستهلكين عن مخفض الناتج القومي في أن مؤشر أسعار المستهلكين يتم حسابه باستخدام السلعة والخدمات التي لها علاقة بالمستهلك العادي كالمواد الغذائية والملابس والسكن والمواصلات والاتصالات والرعاية الصحية والتعليم ، وما إلى ذلك ، بينما يتم حساب مخفض الناتج القومي الإجمالي باستخدام جميع السلع والخدمات التي يستهلكها النظام الاقتصادي في الدولة من مستهلكين عاديين وقطاع خاص وقطاع حكومي .

ويهتم النظام الاقتصادي السعودي بنشر الأرقام القياسية للأسعار وتكالبف المعيشة على شكل تقارير شهرية ، ومن هذه الأرقام مايلي :

- **الرقم القياسي لتكاليف المعيشة لجميع السكان :** ويشمل المواد الغذائية ، السكن وتوابعه ، الأقمشة والملابس ، الأثاث المنزلي ، الرعاية الطبية ، النقل والاتصالات ، التعليم والترفيه ، النفقات والخدمات الأخرى .
- **الرقم القياسي لأسعار الجملة :** ويشمل المواد الغذائية ، المشروبات ، المواد الخام ، الوقود المعدني وزيوت التشحيم ، الدهون والزيوت الحيوانية والنباتية ، الكيماويات والمواد ذات الصلة ، السلع المصنعة مصنفة حسب المادة ، الألات ومعدات النقل والاتصالات ، التعليم والترفيه ، النفقات والخدمات الأخرى .

(4–9) دور الأرقام القياسية في حساب معدلات التضخم :

المقصود **بالتضخم** هو الارتفاع المستمر في المستوى العام للأسعار والذي على ضوئه تنخفض القيمة الشرائية للوحدة النقدية (الريال مثلاً) ، وتقوم الجهات الاقتصادية في الدول باستخدام الأرقام القياسية للاسعار لإيجاد معدلات التضخم السنوية ، وفي معظم الأحيان يستخدم **مؤشر اسعار المستهلكين (CPI)** لسنتين متتاليتين لحساب معدل التضخم السنوي في السنة الأخيرة وذلك من خلال العلاقة التالية :

$$\boxed{i_r = \frac{CPI_r - CPI_{r-1}}{CPI_{r-1}} \times 100} \qquad (9\text{-}1)$$

حيث : i_r هو معدل التضخم للسنة r [وهو نسبة مئوية] .

CPI_r هو مؤشر أسعار المستهلكين للسنة r .

CPI_{r-1} هو مؤشر أسعار المستهلكين للسنة $r-1$ [أي السنة السابقة] .

فمثلاً إذا افترضنا أن مؤشر أسعار المستهلكين في المملكة لسنة *2006* هو *120* ومؤشر أسعار المستهلكين لسنة *2007* هو *123* ، فإن معدل التضخم في سنة *2007* يكون :

$$i_{2007} = \frac{CPI_{2007} - CPI_{2006}}{CPI_{2006}} \times 100 = \frac{123-120}{120} \times 100 = 2.5\%$$

أي أن معدل التضخم في سنة *2007* يساوي *2.5%* .

(9–5) فوائد الأرقام القياسية واستعمالاتها :

تستخدم الأرقام القياسية عادة لقياس التغير الذي يطرأ على الحياة بمجملها بشكل عام والجوانب الاقتصادية بشكل خاص ، كما تساعد الأرقام القياسية على تحليل العوامل التي تساهم في تغير الظاهرة فتبين مدى مساهمة كلٍ من هذه العوامل في إحداث التغير الكلي ، وتستخدم كذلك في الرقابة على تنفيذ الخطط .

(9–6) الرقم القياسي المرجح

هو ذلك الرقم الذي يأخذ الأهمية النسبية للسلعة (أو الأجر) بعين الاعتبار فيعطي كل سلعة (أجر) وزناً يتلاءم مع أهميته

فعند تركيب رقم قياسي للكميات يجب ترجيحه بالأسعار ، وعند تركيب رقم قياسي للأسعار يجب ترجيحه بالكميات وبالتالي يكون الناتج **رقماً قياسياً مرجحاً** .

(9–7) منسوب السعر لسلعة واحدة (ظاهرة بسيطة) :

يمكن إيجاد رقم قياسي لسعر سلعة بمفردها (حيث يمثل هذا الرقم القياسي التغير في سعر السلعة أو الخدمة في سنة معينة مقارنة بسنة الأساس) ، ويسمى الرقم القياسي للسعر **بمنسوب السعر** ويرمز له بالرمز Pr ويُعطى بـ :

$$\boxed{P_r = \frac{P_1}{P0} \times 100} \qquad (9\text{-}2)$$

حيث : P_r هو منسوب السعر للسلعة [وهو نسبة مئوية] .

P_0 هو السعر خلال فترة الأساس .

P_1 هو السعر خلال فترة المقارنة .

مثال (9–1) :

البيانات التالية تمثل سعر سلعة معينة من الفترة من *2006* إلى *2010* .

السنة	2006	2007	2008	2009	2010
السعر (بالريال)	25	30	24	32	36

المطلوب إيجاد منسوب السعر لهذه السلعة للفترة من سنة *2006* حتى سنة *2010* باعتبار سنة *2006*سنة أساس ، **مع تفسير** النتائج التي يتم الحصول عليها .

الحل :

يتم مباشرةً حساب قيمة منسوب السعر لهذه السلع من خلال العلاقة (10-2) [باعتبار سنة 2006 سنة الأساس] كما يوضحه الجدول التالي :

السنة	سعر السلعة بالريال	منسوب السعر	تفسير النتائج
2006	25	$\frac{25}{25}\times 100 = 100\%$	سنة الأساس
2007	30	$\frac{30}{25}\times 100 = 120\%$	زيادة في السعر بنسبة 20%
2008	24	$\frac{24}{25}\times 100 = 96\%$	إنخفاض في السعر بنسبة 4%
2009	32	$\frac{32}{25}\times 100 = 128\%$	زيادة في السعر بنسبة 28%
2010	36	$\frac{36}{25}\times 100 = 144\%$	زيادة في السعر بنسبة 44%

والعمود الأخير من الجدول السابق يفسر النتائج التي حصلنا عليها .

(9–8) منسوب السعر لمجموعة من السلع (ظاهرة معقدة) :

الرقم القياسي السابق يوضح منسوب السعر لسلعة واحدة ، إلا أن كثيراً من الحالات تكون أكثر تعقيداً ، فقد يكون لدينا عدة سلع متغيرة ونرغب في حساب منسوب السعر (أو الرقم القياسي) لها ، ففي حالة استخراج الرقم القياسي لمثل هذا الوضع فإنه يدخل في الحساب جميع قيم السلع التي تتألف منها الظاهرة .

ويتم ذلك من خلال استخدام الطرق التالية :

- الرقم القياسي التجميعي البسيط للأسعار .
- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير)
- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش)
- الرقم القياسي التجميعي المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر)

وسنلقي الضوء على كلٍ منها في البنود التالية :

(9–8–1) الرقم القياسي التجميعي البسيط للأسعار :

يرمز لهذا الرقم القياسي بالرمز " I_s " ويتم حسابه من خلال العلاقة التالية :

$$I_s = \frac{\sum P_1}{\sum P_0}\times 100 \qquad (9\text{-}3)$$

حيث : $\sum P_1$ هو مجموع أسعار السلع والخدمات خلال فترة المقارنة .

$\sum P_0$ هو مجموع أسعار السلع والخدمات خلال فترة الأساس .

وتكمن مشكلة الرقم القياسي التجميعي البسيط للأسعار في أنه لا يعطي للكميات المستهلكة من السلع والخدمات أوزاناً ، وبالتالي يكون حساساً عندما يكون هناك تبايناً في الكميات المستهلكة .

(9-8-2) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير) :

ويرمز له بالرمز " I_r " ، وهذا الرقم يعبر عن أثر التغير في السعر كما لو بقيت الكميات المشتراة في سنة الأساس هي نفسها في سنة المقارنة ، ويتم حسابه بنفس الطريقة السابقة مع ترجيح وزن كل سعر بكميته Q_0 خلال فترة الأساس ، ويتم ذلك من خلال تطبيق العلاقة التالية :

$$\boxed{I_r = \frac{\sum P_1Q_0}{\sum P_0Q_0} \times 100} \qquad (9\text{-}4)$$

حيث : $\sum P_1Q_0$ هو مجموع أسعار السلع والخدمات خلال فترة المقارنة مرجحة بكميات فترة الأساس .

$\sum P_0Q_0$ هو مجموع أسعار السلع والخدمات خلال فترة الأساس مرجحة بكميات فترة الأساس .

ويفضل استخدام هذه الطريقة عند حساب مؤشر أسعار المستهلكين (CPI) وذلك للاقتصاد في الجهد والوقت والمال ، لأن كمية فترة الأساس ثابتة عند إيجاد رقم لاسبير لأي فترة لاحقة لفترة الأساس .

(9-8-3) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش):

ويرمز له بالرمز " I_p " ، وهذا الرقم يعبر عن اثر التغير في السعر كما لو أن الكميات المشتراة في سنة المقارنة كانت قد اشتريت في سنة الأساس . وتختلف طريقة حساب هذا الرقم من حيث أنه يرجح كل سعر بكميته Q_1 خلال فترة المقارنة ، ويتم ذلك من خلال تطبيق العلاقة التالية :

$$\boxed{I_p = \frac{\sum P_1Q_1}{\sum P_0Q_1} \times 100} \qquad (9\text{-}5)$$

حيث : $\sum P_1Q_1$ هو مجموع أسعار السلع والخدمات خلال فترة المقارنة مرجحة بكميات فترة المقارنة .

$\sum P_0Q_1$ هو مجموع أسعار السلع والخدمات خلال فترة الأساس مرجحة بكميات فترة المقارنة .

والمشكلة الأساسية في هذه الطريقة هي الحاجة لتحديد الكميات المستهلكة من كل سلعة سنوياً حتى يتسنى لنا حساب هذا الرقم .

(9-8-4) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر – أو الرقم الأمثل):

ويرمز له بالرمز " I_f " ، وهو عبارة عن الوسط الهندسي لكل من رقمي لاسبير وباش [أي أنه الجذر التربيعي لحاصل ضرب رقم لاسبير برقم باش] ، وهذا الرقم يهتم بالناحية الرياضية ولكنه لا معنى اقتصادي له ، وهذا هو أهم عيوبه . أي أن رقم فيشر يُعطى بـ :

$$I_f = \sqrt{I_r I_p} = \sqrt{\frac{\sum P_1Q_0}{\sum P_0Q_0} \times \frac{\sum P_1Q_1}{\sum P_0Q_1}} \times 100 \qquad (9\text{-}6)$$

مثال (2-9) :

يبين الجدول التالي أسعار وكميات ثلاث منتجات استهلاكية للسنتين 2007 ، 2010 :

	سنة 2007		سنة 2010	
المنتج	الكمية	السعر	الكمية	السعر
السلعة الأولى	5000	9	8500	12
السلعة الثانية	8000	25	15000	31
السلعة الثالثة	9000	14	19000	17

على اعتبار أن سنة 2007 هي سنة الأساس ، **المطلوب** :

(أ) حساب الرقم التجميعي البسيط للأسعار .

(ب) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير) .

(ج) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش) .

(د) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر) .

مع تفسير نتائج الفقرات السابقة .

الحل :

يمكن إعادة كتابة بيانات الجدول المعطى بعد إضافة الرموز المختلفة وأعمدة مختلفة توضح الكميات المطلوبة [P_0 : السعر سنة الأساس 2007 ، Q_0 : الكمية سنة الأساس 2007 ، P_1 : السعر سنة المقارنة 2010 ، Q_1 : الكمية سنة المقارنة 2010] ، يمكن حساب الأرقام القياسية المطلوبة كالتالي :

	سنة الأساس 2007		سنة المقارنة 2010					
السلعة	**كمية**	**سعر**	**كمية**	**سعر**	P_0Q_0	P_0Q_1	P_1Q_0	P_1Q_1

	Q_0	P_0	Q_1	P_1				
الأولى	5000	9	8500	12	45000	76500	60000	102000
الثانية	8000	25	15000	31	200000	375000	248000	465000
الثالثة	9000	14	19000	17	126000	266000	153000	323000
		48		60	371000	717500	461000	890000
		$\sum P_0$		$\sum P_1$	$\sum P_0Q_0$	$\sum P_0Q_1$	$\sum P_1Q_0$	$\sum P_1Q_1$

(أ) الرقم القياسي التجميعي البسيط :

من العلاقة *(9-3)* يكون الرقم التجميعي البسيط I_s هو :

$$I_s = \frac{\sum P_1}{\sum P_0} \times 100 = \frac{60}{48} \times 100 = 125\%$$

وهذا يعني أن المستوى العام لأسعار المنتجات الثلاث قد ارتفع في سنة *2010* بمعدل *25%* وذلك مقارنة بسنة *2007* .

(ب) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير) :

من العلاقة *(9-4)* يكون الرقم القياسي التجميعي المرجح بكميات سنة الأساس [رقم سبير] I_r هو :

$$I_r = \frac{\sum P_1Q_0}{\sum P_0Q_0} \times 100 = \frac{461000}{371000} \times 100 \cong 124.26\%$$

وهذا يدل على أن المستوى العام لأسعار المنتجات الثلاث قد ارتفع في سنة *2010* بمعدل *24.26%* وذلك مقارنة بسنة *2007* .

(ج) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش) :

من العلاقة *(9-5)* يكون الرقم القياسي التجميعي المرجح بكميات سنة المقارنة [رقم باش] I_p هو :

$$I_p = \frac{\sum P_1Q_1}{\sum P_0Q_1} \times 100 = \frac{890000}{717500} \times 100 \cong 124.04\%$$

وهذا يدل على أن المستوى العام لأسعار المنتجات الثلاث قد ارتفع في سنة *2010* بمعدل *24.04%* وذلك مقارنة بسنة *2007* .

(د) الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر) الرقم القياسي الأمثل :

من العلاقة (9-6) يكون الرقم القياسي التجميعي المرجح بكميات كلٍ من سنة الأساس وسنة المقارنة [رقم فيشر أو الرقم الأمثل] I_f هو الوسط الهندسي لكلٍ من رقمي سبير وباش ، أي أن :

$$I_f = \sqrt{I_r \times I_p} = \sqrt{124.26 \times 124.04} \cong 124.15\%$$

وهذا يدل على أن المستوى العام لأسعار المنتجات الثلاث قد ارتفع في سنة 2010 بمعدل 24.15% وذلك مقارنة بسنة 2007 .

(9-9) الرقم القياسي للكميات :

بالإضافة إلى الحاجة لحساب الرقم القياسي للأسعار قد يحتاج المرء لحساب الرقم القياسي لكميةٍ ما مثل كمية إنتاج أو كمية استهلاك [وسوف نرمز له بالرمز I_q] والذي يُعرف بـ :

$$\boxed{I_q = \frac{Q_1}{Q_0} \times 100} \qquad (9\text{-}7)$$

حيث : Q_0 هي الكمية خلال فترة الأساس .

Q_1 هي الكمية خلال فترة المقارنة .

وكما في حالة الأسعار يمكن تعريف الرقم القياسي التجميعي للكميات كالآتي :

$$\boxed{I_q = \frac{\sum Q_1}{\sum Q_0} \times 100} \qquad (9\text{-}8)$$

حيث : $\sum Q_0$ هي مجموع الكميات خلال فترة الأساس .

$\sum Q_1$ هي مجموع الكميات خلال فترة المقارنة .

مثال (9-3) :

الجدول التالي يمثل كمية الإنتاج (بالطن) من القمح بإحدى المزارع وذلك خلال الفترة من سنة 1998 إلى سنة 2007 .

السنة	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
الإنتاج	86	87	84	90	88	93	100	110	105	120

المطلوب إيجاد الرقم القياسي لإنتاج هذه المزرعة لعام *2007* على اعتبار أن فترة الأساس (*5* سنوات) من عام *1998* إلى عام *2002* .

الحل :

من العلاقة *(9-7)* يكون الرقم القياسي لكمية الإنتاج لعام *2007* هو :

$$I_q = \frac{Q_1}{Q_0} \times 100$$

حيث : Q_1 هي كمية الإنتاج في سنة *2007* [$120 =$] ،

Q_0 هي كمية الإنتاج خلال فترة الأساس .

وهنا لا توجد سنة أساس ولكن توجد فترة أساس [*5* سنوات من عام *1998* حتى *2002*] لذا فإن الكمية Q_0 تُمثل بالوسط الحسابي $\overline{Q}_0$ خلال فترة الأساس والتي تساوي :

$$\overline{Q}_0 = \frac{86+87+84+90+88}{5} = 87$$

وعلى ذلك يمكن حساب الرقم القياسي لكمية الأنتاج لسنة *2007* هو :

$$I_q = \frac{Q_1}{\overline{Q}_0} \times 100 = \frac{120}{87} \times 100 \cong 137.39\%$$

أي إن مقدار الزيادة في كمية الإنتاج في سنة *2007* تساوي *37.39%* مقارنةً بمتوسط كمية الإنتاج في فترة الأساس [*5* سنوات من عام *1998* حتى *2002*] .

(9–10) الرقم القياسي للدخل الحقيقي :

يمثل الدخل الحقيقي للفرد (القوة الشرائية للنقود) مؤشراً أساسياً في النظام الاقتصادي لأي دولة وذلك بالإضافة للأرقام القياسية الأخرى ، ويتم حساب الرقم القياسي للدخل الحقيقي [وسنرمز له بالرمز I_i] من خلال العلاقة التالية :

$$\text{الرقم القياسي للدخل الحقيقي } I_i = \frac{\text{الرقم القياسي للدخل النقدي}}{\text{الرقم القياسي لنفقات المعيشة}} \qquad (9\text{-}8)$$

مثال (9–4) :

إذا كان الرقم القياسي لدخل الفرد عام *2009* بالنسبة لعام *2003* يساوي *1.7* بينما الرقم القياسي لتكاليف المعيشة لعام *2009* بالنسبة لعام *2003* يساوي *3.4* ، **المطلوب** حساب الرقم القياسي للدخل الحقيقي مع التعليق على النتائج المتحصل عليها .

الحل :

من العلاقة *(8-9)* يكون الرقم القياسي للدخل الحقيقي I_i هو :

$$\mathbf{I_i} = \frac{\text{الرقم القياسي للدخل النقدي}}{\text{الرقم القياسي لنفقات المعيشة}} = \frac{1.7}{3.4}\times 100 = 50\%$$

أي أن الدخل الحقيقي (القوة الشرائية للنقود) قد انخفضت إلى النصف ، وهذه تسمى ظاهرة الانكماش للأرقام القياسية للدخل الحقيقي .

(9–11) ملاحظات عامة على الأرقام القياسية :

هناك مجموعة من الملاحظات المتعلقة بتفسير الأرقام القياسية لسنوات الأساس والمقارنة ، وهذه الملاحظات يمكن تلخيصها في التالي :

- الرقم القياسي للظاهرة في سنة الأساس يساوي *100* .
- إذا كان الرقم القياسي للظاهرة في سنة المقارنة أكبر من *100* فهذا يعني أن هناك ارتفاع في المستوى العام للظاهرة مقارنة بسنة الأساس .
- إذا كان الرقم القياسي للظاهرة في سنة المقارنة أصغر من *100* فهذا يعني أن هناك انخفاض في المستوى العام للظاهرة مقارنة بسنة الأساس

ملخص للفصل التاسع

- **الرقم القياسي** هو مؤشر إحصائي (رقم نسبي) يستخدم في قياس التغير النسبي الذي يطرأ على ظاهرة من الظواهر الاقتصادية أو الاجتماعية .

- **الأساس** هو فترة زمنية معينة أو مكاناً معيناً ، وعادة تكون فترة الأساس فترة سابقة للفترة التي نريد مقارنتها ، وفي حالات نادرة جداً قد تكون فترة الأساس فترة لاحقة لفترة المقارنة .

- **الأرقام القياسية للأسعار** من أهم أنواع الأرقام القياسية وأكثرها شيوعاً ، وهي تساهم في قياس التغير في المستوى العام للأسعار أو التغير في تكاليف المعيشة في فترة زمنية معينة مقارنة بفترة زمنية أخرى ، ومن أشهرها : مؤشر أسعار المستهلكين ، مخفض الناتج القومي الإجمالي ، مؤشر اسعار المنتجين ، مخفض الناتج المحلي الإجمالي ، ومؤشر أسعار الأسهم .

- المقصود **بالتضخم** هو الارتفاع المستمر في المستوى العام للأسعار ، وفي معظم الأحيان يستخدم مؤشر اسعار المستهلكين (CPI) لسنتين متتاليتين لحساب معدل التضخم السنوي i_r في السنة r من خلال العلاقة التالية :

$$\boxed{i_r = \frac{CPI_r - CPI_{r-1}}{CPI_{r-1}} \times 100}$$

- **منسوب السعر P_r لسلعة واحدة** هو رقم قياسي لسعر سلعة بمفردها وهو النسبة المئوية بين سعر السلعة P_1 في سنة المقارنة وسعرها P_0 في سنة الأساس :

$$\boxed{P_r = \frac{P_1}{P0} \times 100}$$

- **الرقم القياسي التجميعي البسيط للأسعار :** هو النسبة المئوية بين مجموع أسعار السلع والخدمات في سنة المقارنة ومجموع الأسعار والخدمات في سنة الأساس :

$$\boxed{I_s = \frac{\sum P_1}{\sum P_0} \times 100}$$

- **الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير) :** وهو يعبر عن أثر التغير في السعر كما لو بقيت الكميات المشتراة في سنة الأساس هي نفسها في سنة المقارنة :

$$\boxed{I_r = \frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times 100}$$

- **الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش) :** وهو يعبر عن اثر التغير في السعر كما لو أن الكميات المشتراة في سنة المقارنة كانت قد اشتريت في سنة الأساس :

$$I_p = \frac{\sum P_1 Q_1}{\sum P_0 Q_1} \times 100$$

- **الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر [أو الرقم الأمثل] :** وهو عبارة عن الوسط الهندسي لكل من رقمي لاسبير وباش :

$$I_f = \sqrt{I_r I_p} = \sqrt{\frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times \frac{\sum P_1 Q_1}{\sum P_0 Q_1}} \times 100$$

- **الرقم القياسي لكمية :** وهو النسبة المئوية بين الكمية Q_1 في سنة المقارنة والكمية Q_0 في سنة الأساس :

$$I_q = \frac{Q_1}{Q_0} \times 100$$

- **الرقم القياسي التجميعي للكميات :** وهو النسبة المئوية بين الكمية Q_1 في سنة المقارنة والكمية Q_0 في سنة الأساس :

$$I_q = \frac{\sum Q_1}{\sum Q_0} \times 100$$

- **الرقم القياسي للدخل الحقيقي I_i :** وهو النسبة المئوية بين الرقم القياسي للدخل النقدي والرقم القياسي لتكاليف المعيشة .

- والرقم القياسي للظاهرة في سنة الأساس يساوي *100* ، فإذا كان الرقم القياسي للظاهرة في سنة المقارنة أكبر من *100* فهذا يعني أن هناك ارتفاع في المستوى العام للظاهرة مقارنة بسنة الأساس ، أما إذا كان الرقم القياسي للظاهرة في سنة المقارنة أصغر من *100* فهذا يعني أن هناك انخفاض في المستوى العام للظاهرة مقارنة بسنة الأساس .

تدريبات عملية

تدريب عملي (1) :

إذا كان مؤشر اسعار المستهلكين في المملكة لسنة 2007 هو 125 ، وسنة 2010 هو 134 ، ما **هو** معدل التضخم في سنة 2010 ؟ .

تدريب عملي (2) :

إذا كان الرقم القياسي لدخل الفرد عام 2009 بالنسبة لعام 2003 يساوي 2.1 بينما الرقم القياسي لتكاليف المعيشة لعام 2009 بالنسبة لعام 2003 يساوي 5.3 ، **احسب** الرقم القياسي للدخل الحقيقي مع التعليق على النتائج المتحصل عليها .

تدريب عملي (3) :

الجدول التالي يمثل كمية الإنتاج (بالطن) من القمح بأحد المزارع خلال الفترة من سنة 2000 إلى سنة 2009 :

السنة	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
الإنتاج	56	62	73	64	65	83	87	92	98	96

أوجد الرقم القياسي لإنتاج هذه المزرعة لعام 2009 على اعتبار أن فترة الأساس (6 سنوات) من عام 2000 إلى عام 2005 .

تدريب (4) :

يبين الجدول التالي أسعار وكميات ثلاث منتجات استهلاكية للسنتين 2007 و 2010 :

	سنة الأساس 2007		سنة المقارنة 2010	
	الكمية Q_0	السعر P_0	الكمية Q_1	السعر P_1
السلعة (1)	2500	11	3750	18
السلعة (2)	3000	25	5600	33
السلعة (3)	4500	17	7240	23

باعتبار أن سنة 2007 هي سنة الأساس ، **المطلوب** :

- حساب الرقم التجميعي البسيط للأسعار .
- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير) .
- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش) .
- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر) .

تدريبات (13)

الإجابة النهائية لجميع التمرينات موجودة في نهاية التدريب

اختر الإجابة الصحيحة

(1) .... هو مؤشر إحصائي (رقم نسبي) يستخدم في قياس التغير النسبي الذي يطرأ على ظاهرة من الظواهر الاقتصادية أو الاجتماعية .

(أ) الأساس (ب) الرقم القياسي (ج) التضخم

(2) .... هو فترة زمنية معينة أو مكاناً معيناً يُستخدم في عملية المقارنة .

(أ) الأساس (ب) الرقم القياسي (ج) التضخم

(3) .... هو الارتفاع المستمر في المستوى العام للأسعار .

(أ) الأساس (ب) الرقم القياسي (ج) التضخم

(4) .... هو النسبة المئوية بين مجموع أسعار السلع والخدمات في سنة المقارنة ومجموع الأسعار والخدمات في سنة الأساس .

(أ) الرقم القياسي التجميعي البسيط للأسعار

(ب) الرقم القياسي التجميعي المرجح بكميات سنة الأساس .

(ج) الرقم القياسي التجميعي المرجح بكميات سنة المقارنة .

(د) الرقم القياسي التجميعي الأمثل .

(5) .... يعبر عن اثر التغير في السعر كما لو أن الكميات المشتراة في سنة المقارنة كانت قد اشتريت في سنة الأساس .

(أ) الرقم القياسي التجميعي البسيط للأسعار

(ب) الرقم القياسي التجميعي المرجح بكميات سنة الأساس .

(ج) الرقم القياسي التجميعي المرجح بكميات سنة المقارنة .

(د) الرقم القياسي التجميعي الأمثل .

(6) .... يعبر عن اثر التغير في السعر كما لو أن الكميات المشتراة في سنة المقارنة كانت قد اشتريت في سنة الأساس .

(أ) الرقم القياسي التجميعي البسيط للأسعار

(ب) الرقم القياسي التجميعي المرجح بكميات سنة الأساس .

(ج) الرقم القياسي التجميعي المرجح بكميات سنة المقارنة .

(د) الرقم القياسي التجميعي الأمثل .

خاص بالأسئلة من (7) إلى (11) :

إذا كان P_1 يمثل سعر السلعة ، Q_1 هو كميتها [وذلك خلال فترة المقارنة] ، وكان P_0 يمثل سعر السلعة ، Q_0 هو كميتها [وذلك خلال فترة الأساس] ، فإن :

(7) الرقم القياسي البسيط التجميعي للأسعار يُعطى بـ :

(أ) $\frac{\sum P_1}{\sum P_0}\times 100$ (ب) $\frac{\sum P_1Q_0}{\sum P_0Q_0}\times 100$

(ج) $\frac{\sum P_1Q_1}{\sum P_0Q_1}\times 100$ (د) $\sqrt{\frac{\sum P_1Q_0}{\sum P_0Q_0}\times\frac{\sum P_1Q_1}{\sum P_0Q_1}}$

(8) رقم سبير يُعطى بـ :

(أ) $\frac{\sum P_1}{\sum P_0}\times 100$ (ب) $\frac{\sum P_1Q_0}{\sum P_0Q_0}\times 100$

(ج) $\frac{\sum P_1Q_1}{\sum P_0Q_1}\times 100$ (د) $\sqrt{\frac{\sum P_1Q_0}{\sum P_0Q_0}\times\frac{\sum P_1Q_1}{\sum P_0Q_1}}$

(9) رقم باش يُعطى بـ :

(أ) $\frac{\sum P_1}{\sum P_0}\times 100$ (ب) $\frac{\sum P_1Q_0}{\sum P_0Q_0}\times 100$

(ج) $\frac{\sum P_1Q_1}{\sum P_0Q_1}\times 100$ (د) $\sqrt{\frac{\sum P_1Q_0}{\sum P_0Q_0}\times\frac{\sum P_1Q_1}{\sum P_0Q_1}}$

(10) رقم فيشر (الرقم الأمثل) يُعطى بـ :

(أ) $\frac{\sum P_1}{\sum P_0}\times 100$ (ب) $\frac{\sum P_1Q_0}{\sum P_0Q_0}\times 100$

(ج) $\frac{\sum P_1Q_1}{\sum P_0Q_1}\times 100$ (د) $\sqrt{\frac{\sum P_1Q_0}{\sum P_0Q_0}\times\frac{\sum P_1Q_1}{\sum P_0Q_1}}$

(11) الرقم القياسي لكمية الإنتاج يُعطى بـ :

(أ) $\frac{Q_0P_0}{Q_1P_1}\times 100$ (ب) $\frac{Q_0}{Q_0}\times 100$

(ج) $\frac{Q_1P_1}{Q_0P_0}\times 100$ (د) $\frac{Q_0}{Q_1}\times 100$

خاص بالأسئلة من (12) إلى (16) :

الجدول التالي يبين أسعار وكميات سلعتين خلال سنتي أساس ومقارنة ، من هذا الجدول يمكن استنتاج الآتي :

	سنة الأساس		سنة المقارنة					
	Q_0	P_0	Q_1	P_1	P_0Q_0	P_0Q_1	P_1Q_0	P_1Q_1
السلعة الأولى	100	15	125	18	1500	1875	1800	2250
السلعة الثانية	150	20	200	30	3000	4000	4500	6000
المجموع	250	35	325	48	4500	5875	6300	8250

(12) منسوب السعر للسلعة الأولى يساوي :

(أ) 137.1% (ب) 140% (ج) 120% (د) 140.4%

(13) الرقم التجميعي البسيط للسلع يساوي :

(أ) 137.1% (ب) 140% (ج) 120% (د) 140.4%

(14) رقم سبير القياسي للأسعار يساوي :

(أ) 137.1% (ب) 140% (ج) 120% (د) 140.4%

(15) رقم باش القياسي للأسعار يساوي :

(أ) 137.1% (ب) 140% (ج) 120% (د) 140.4%

(16) الرقم الأمثل للأسعار يساوي :

(أ) 140.2% (ب) 138.5% (ج) 129.6% (د) 129.8%

(17) الرقم القياسي لكمية السلعة الثانية يساوي :

أ) 125%   ب) 133.3%   ج) 130%   د) 130.6%

18) الرقم القياسي التجميعي لكميات السلع يساوي :

أ) 125%   ب) 133.3%   ج) 130%   د) 130.6%

إجابة تدريبات (13) :

1) ب   2) أ   3) ج   4) أ   5) ب   6) ج

7) أ   8) ب   9) ج   10) د   11) ب   12) ج

13) أ   14) ب   15) د   16) أ   17) ب   18) ج