محاضرات الاحصاء في الادارة

أ.د. عبدالله بن عمر النجار

1. هذه الملخص عبارة عن جمع لمحتوى المحاضرات مضاف عليه حلول التمارين من الكتاب وكذلك طرق الحل للامور التي تقبل الحل بالالة الحاسبة و بعض الملاحظات الجانبية .
2. تذكر عزيزي قارئ هذا المحتوى انه اجتهاد شخصي وانني بشر اصيب واخطئ فإن اصبت فيما كتبت وجمعت فمن الله وإن اخطأت فمن نفسي والشيطان ، ولهذا وجب التنبيه على ذلك ، وإن تحميلك لهذا الملف هو بناءا على ارداتك الشخصية وانك تشهد الله بأن ذمتي بريئة من اي خطاء او ضرر قد يصيبك جراء اتباعك لما كتب في هذا الملف لا سمح الله .
3. المجهود المبذول في هذا الملخص ليس مجاني فقيمته ريال واحد تصدق به عني .

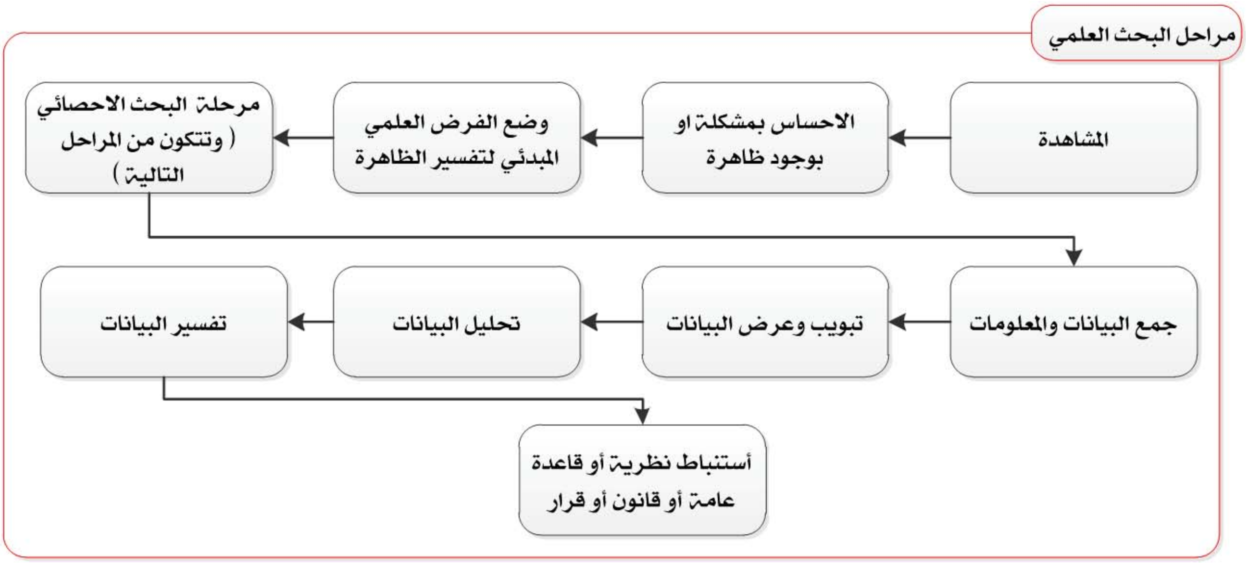
DR. JEKYLL

(المحاضرة الاولـــى)

علم الإحصاء ودوره في خدمة المجتمع

❖ تمهيد :

ان الغرض من العلم هو البحث عن الحقيقة، وأن البحث العلمي هو الوسيلة للوصول إلى حقائق الأشياء والظواهر ومعرفة كل العلاقات التي تربط بينها



❖ تاريخ علم الإحصاء وتطوره : لقد مر علم الإحصاء بثلاث مراحل للتطور وهذه المراحل هي:

1. مرحلة التعداد : وقت اهتمت بفكرة الجرد شبه الدائم للسكان والخيرات المتوافرة في البلاد ، وكان ذلك في مرحلة ما قبل التاريخ ومرحلة التاريخ الاسلامي .
2. مرحلة الحساب السياسي : تعدت هذه المرحلة عملية الوصف الى عملية الوصول الى القوانين التي تفسر مختلف الاحداث والعمليات الاجتماعية ومن هذه المرحلة بدأ الاحصاء كعلم ، (بدأت مع مطلع القرن السادس عشر الميلادي)
3. مرحلة الإحصاء وحساب الاحتمالات : تم فيها استخدام الاساليب الاحصائية المتقدمة ، والتعرف على التوزيعات الاحصائية بأنواعها ، (بدأت خلال القرن الثامن عشر الميلادي)

❖ مجالات استعمال علم الإحصاء:

1. يستخدم الإحصاء في تطوير التعليم وخططه .
2. يستعمل الإحصاء في دراسة مختلف العلوم . (الصحية والصيدلة)
3. يستعمل الإحصاء في مجال الدعاية والإعلانات التجارية .
4. يستعمل الإحصاء بشكل كبير من قبل شركات التأمين .
5. يستعمل الإحصاء في حساب الأرقام القياسية . (سوف ندرسها في المحاضرات القادمة)
6. يستعمل الإحصاء في اختبارات الذكاء والتحصيل والقدرات .
7. يستعمل الإحصاء بشكل كبير في القطاع الصناعي .

❖ تعريف الاحصاء :

1. تعريف الإحصاء في اللغة : يعرف الإحصاء في اللغة بأنه العدد الشامل .

2. تعريف الإحصاء في الاصطلاح: هو فرع من فروع الرياضيات يهدف الى جمع وعرض وتنظيم ووصف وتحليل البيانات المقاسة رقمياً مما يساعد على اتخاذ قرارات واستنتاجات وتوصيات مبنية على نظرية الاحتمالات .

❖ أهداف علم الاحصاء :

1- جمع البيانات عن الظواهر المختلفة التي تهم الباحث بطرق علمية محددة تحديدا دقيقا وبشكل منسق

2- تبويب البيانات : طبقا لاسلوب تصنيف محدد مسبقا

3- عرض البيانات باستخدام الجداول او الاشكال البيانيه او الرسوم البيانيه

4- وصف البيانات : وذلك عن طريق ابراز الخصائث الاساسية لها والتي يمكن التعبير عنها بمقاييس معينة ومعينة ومحدده وتقاس الخصائص الاساسية لاي مجموعة من البيانات بمقاييس النزعة المركزية او التشتت او الالتواء والاعتدال

5- تحليل البيانات : وذلك عن طريق استعمال خصائصها الاساسية التي تم ابرازها للوصلو الى الارقام ذات العلاقة بالمشكلة والتي يهم الباحث الحصول عليها للوصول الى نتائج محددة .

6- استخدام النتائج وتفسيرها تفسيراً منطقيا مناسبا لطبيعة المشكلة التي يبحثها حتى يتسنى للباحث الاستفادة منها وتطبيقها في الحياة الواقعية .

❖ أهمية علم الإحصاء للباحث والبحوث العلمية :

يعتبر علم الإحصاء وسيلة لا غاية يساعد استخدامه على التالي :

1. الوصف بدقة إلى اكبر حد ممكن : وتزيد الدقه كلما زادت قدرة الباحث في استعمال الاساليب الرياضية الاحصائية .

2. التزام التحديد والدقة في أساليبنا العملية وفي تفكيرنا : عند كثرة استخدامنا للاحصاء تتولد لدنيا القدرة على الدقه في اغلب الامور ويتطور تفكيرنا .

3. تلخيص نتائجنا في شكل ملائم ذو معنى واضح : فالاحصاء يساعدنا على تلخيص وتصنيف المعلومات التي نستمدها بالمشاهدة من الظواهر المحطية بنا وكذلك يساعد على تجنب الاصطرابات والارتباك الناتج عن تجميع البيانت بدون نظام او ترتيب .

4. استخلاص النتائج في الدراسات والبحوث : بعد عملية تلخيص النتائج في شكل ملائم لا بد ان يتلو هذه الخطوة خطوة مهمه يهتم بها الاحصاء الاستنتاجي والذي يساعد على استخلاص النتائج من العينة الماخوذة من المجتمع وبناء على الاساليب المستخدمة يمكن تعميم هذه النتائج على المجتمع الاصلي مع تحديد درجة الذقه التي يمكن اعطاؤها للتعميم .

5. التنبؤ بالمدى الذي تحصل فيه ظاهرة تحت ظروف نعرفها ونقيسها : كالتنبؤ بمقدرة الطالب على اجتياز الدراسة الجامعية بناءا على اختبار القدرات الخاص به .

6. تحليل بعض العوامل المعقدة والمتشابكة التي تؤثر في حادث من الحوادث : ان اجراء دراسة لعينة من الناس يساعد على الكشف عن عوامل اي ظاهرة موضوعه للدراسة .

❖ من خلال العرض السابق يتبين لنا أن الإحصاء ينقسم إلى قسمين:

1. الإحصاء الوصفي Descriptive Statistics : هو ذلك القسم من الاحصاء الذي يهتم بجمع بيانات المشكلة وتصنيفها وعرضها ثم إجراء الحسابات المختلفة للوصلو الى النتائج المختلفة التي تبرز خصائصها الاساسية . والغرض من الاحصاء الوصفي هو تقدير معالم المجتمع الاحصائي (المتوسط ، الانحراف المعياري ...) ووصفه تمهيداً للوصول الى استنتاجات عنه فهو عادة خطوة تسبق الاحصاء الاستنتاجي .

2. الإحصاء الاستنتاجي أو الاستدلالي(التحليلي) Inferential Statistics : هو العلم الذي يدرس الظروف والظواهر الاجتماعية والتربوية متعديا العرض الوصفي للبيانات الاحصائية الى تحليل

هذه الحقائق والبيانات باستعمال عدد من الاساليب والطرق الاحصائية الاستنتاجية وذلك باستنتاج معلومات جديدة ، واتخاذ قرارات وتوصيات في ضوء تلك النتائج .
وهناك تعريف اخر للاحصاء الاستنتاجي يقول بأن الاحصاء هو علم اتخاذ القرارات في ظل عدم التأكد .

ويلاحظ ان ان الاحصاء الاستنتاجي يبدأ حيث ينتهي الاحصاء الوصفي فبعد ابراءز الخاصئث الاساسية للبيانات يبدأ الاحصاء الاستنتاجي حيث يتم تحليل البيانات واستخدام نتائج التحليل في الاستنتاج ثم تفسير تلك النتائج منطقيا واتخاذ قرارات في ضوء ذلك .

(المحاضرة الثانية)

جمع البيانات وترميزها

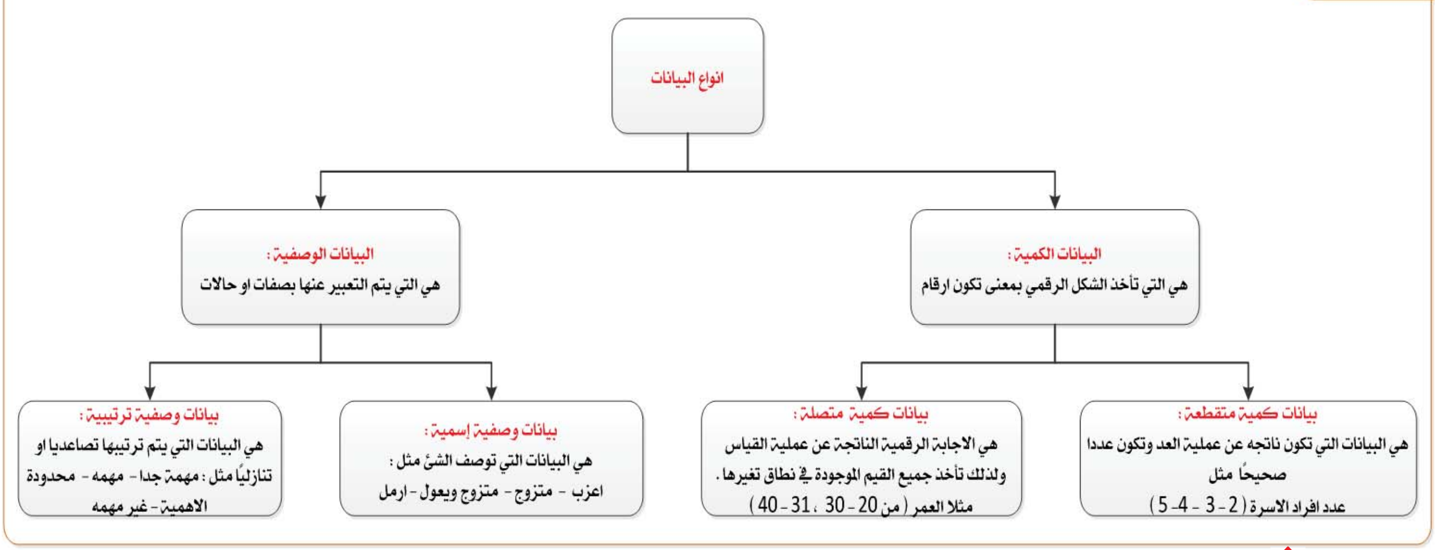
❖ مصطلحات علم الإحصاء :

1. المجتمع Population : ويقصد به المجتمع الاحصائي للظاهرة محل الدراسة ويعرف بأنه جميع المفردات التي يجمعها في اطار عام واحد او مجموعة خصائص عامة واحدة . مثال (عند اجراء دراسة على نوع من اللمبات الكهربائي فان كل انتاج المنصع من هذا النوع يمثل مجتمع الدراسة في حين تمثل اللمبة مفردة الدراسة)
2. العينة Sample : هي جزء من المجتمع الاحصائي محل الدراسة يتم اختياره بطريقة علمية ليتم اجراء الدراسة عليه .
3. المتغير variable : هو خاصية عن المجتمع الاحصائي والتي يتم اختبارها من خلال التحليل الاحصائي . فهي اي صفة او خاصية تتغير من شخص لاخر ومن وقت لاخر ويعمد الباحث لدراستها .
4. المعلمة Parameter : هي قياس وصفي لاحد المتغيرات يتم باستخدام بيانات المجتمع الاحصائي كله
5. الإحصائية Statistic : هي قياس وصفي لاحد المتغيرات يتم باستخدام بيانات العينة والتي تكون تقدير لمعلمة المجتمع .
6. البيانات Data : هي القيمة الوصفية او الرقمية التي نحتاج اليها لمساعدتنا في جعل القرارات التي نتخذها اكثر معلوماتية في موقف محدد .

❖ قبل جمع البيانات لا بد من الإجابة على السؤال التالي:

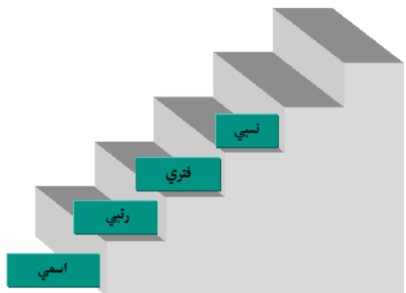
1. ما البيانات الواجب أو المطلوب جمعها؟
2. وما البيانات المرفوضة والتي يجب استبعادها لعدم الحاجة إليها ؟

انواع البيانات



❖ مستويات قياس البيانات هي:

في العلوم الانسانية (ادارة الاعمال)	العلوم الطبيعية
الاسمي Nominal	النسبي Ratio
الرتبي Ordinal	
الفتري Interval	



نوع المقياس	تعريفه	مثال	مزاياه	عيوبه	اهميته ودقته
الاسمي Nominal	هو ذلك النوع من المقاييس الذي يستخدم للدلالة على الشيء	المقايس الاسمي : الجنس ◄ ذكر (1) ◄ انثى (2) فالارقام هنا للدلالة على الشيء وليس لتحديد قيمته	◄ الارقام في للدلالة على الشيء ◄ وهي حصريه فلا تكرر بمعنى لا نستطيع اجراء عمليات حسابية على هذه الارقام ولكن نحسب التكرار فقط		
الرتبي Ordinal	هو مقياس يرتقي قليلا عن الاسمي ويستخدم التصنيف الرقمي لغرض ترتيب الاشياء بدلا عن عملية الدلالة على الشيء كما في الاسمي	لو وجد لدينا اربع متسابقين فعند انتهاء السباق يتم تحديد مستواهم بناءاً على ترتيبهم . مثلا على الاول وخالد المرتبة الثانية وحمد المرتبة الثالثة وهكذا فافضليتهم حسب الترتيب ولكن لا نستطيع تحديد مقدار الافضلية ان على افضل من خالد بكم			
الفترى (الفئوى) Interval	مجموعة من الاعداد او القيم التي ياخذها المتغير الكمي وليس للصفر معنى حقيقي فيها .	درجة الذكاء لايعني ان درجة الذكاء صفر ان الانسان غبي ولكن هي مقياس رمزي	يحدد مقدار الافضلية ويحدد القيمة لكل شخص من الاشخاص	عدم وجود صفر حقيقي في هذا المقايس أي عدم وجود صفه مقاسه من الصفات التي نرغب في قياسها	افضل من ما قبله
النسبي Ratio	مجموعة من الاعداد او القيم التي ياخذها المتغير الكمي والصفر له معنى حقيقي اي يعني انعدام الخاصية محل الدراسة	مثل الوزن و الطول			افضل الانواع وادقهم

❖ تتمثل مصادر البيانات فى ثلاث مصادر أساسية وهى:

1. المصادر التاريخية للبيانات : كالاحصاءات والنشرات والبيانات التي تنشرها الوزارات والشركات ، وبالتالي تكون منظمة معدة بطريقة معينة يسهل التعامل معها واستخدامها مباشرة في عملية التحليل الاحصائي وتوفر الجهد والتكاليف مقارنة بالمصادر الاخرى .
2. الملاحظة : تعتبر من اقدم وسائل جمع المعلومات عن ظاهرة معينة ، حيث تستخدم لجمع معلومات عن سلوك معين سواء من خلال المشاهدة بالعين المجردة او من خلال استخدام بعض الوسائل التكنولوجية مثل كاميرا الفيديو ونحوها ، وتفيد الملاحظة بشكل عام لدراسة سلوك الافراد في اماكنها الطبيعية .
3. المصادر الميدانية : وفيها يقوم البحث بالنزول الى مجتمع الدراسة ليجمع من البيانات التي يحتاج اليها من اجل دراسة المشكلة محل البحث ودراسة ظواهرها المختلفة ، اي يقوم بجمع البيانات المطلوبة من مفردات المجتمع الاحصائي محل الدراسة . ولايتم اللجوء الى المصادر الميدانية الا في حال عدم وجود البيانات المطلوبة لدى المصادرة التاريخية او تكون قديمة او غير دقيقة .

❖ ادوات جمع البيانات للمصادر الميدانية :

يقصد بأداة جمع البيانات الوسيلة التي تتم بواسطتها عملية جمع البيانات بهدف اختبار فرضيات البحث او الاجابة عن تساؤلاته / ويتوقف اختيار الاداة المناسبة لجمع البيانات اللازمة والتي ستستخدم في اجراء بحث معين على الامور التالية :

1. نوعية البحث نفسه .
2. طبيعته
3. الهدف من تطبيق البحث .
4. نوعية المفحوصين وخصائصهم... الخ .

مع الاشارة الى امكانية استخدام الباحث لاداة واحدة فقط او اكثر من اداة في جمع البيانات اذا وجد مبررا لذلك ، وتجدر الاشارة الى خطوة جمع البيانات في البحث تعتبر من الخطوات الاساسية التي يبدأ منها عمل الباحث .

لذا فالهدف النهائي من إعداد وسائل وأدوات جمع البيانات هو الحصول على تلك المعلومات التي تخدم في تحقيق أغراض البحث ودراسة مشكلته، وإيجاد الحلول المناسبة له .

❖ الأدوات الأساسية شائعة الاستعمال من قبل الباحثين لجمع البيانات في المصادر الميدانية :

اسم الاداة	تعريفها	ميزاتها	عيوبها	اهميتها
الاستبانة	مجموعة من الاسئلة المكتوبة تغطي جميع جوانب البحث مكان الدراسة والتي يطلب من المفحوص (افراد العينة) الاجابة عليها بأنفسهم .	1. يمكن تطبيق الاستبانة من خلال الاتصال المباشر بالمفحوص او بارسالها بالبريد اذا كانوا بعيدين 2. اتاحة فرصة كبيرة للمفحوصين لقراءة بنود الاستبانة والتمعن بها 3. اتاحة الفرصة للمحفوص للاجابة بدون خجل وبلا حساسية وبصراحه لعدم طلب الاسم او معلوماته الشخصية 4. تعتبر اكثر موضوعية من غيرها من الاداوات لانها لا تتأثر بتحيزات ذاتية او شخصية من الباحث 5. يوفر استخدام الاستبانة الجهد والوقت والمال حيث انها تحتاج الى قله من المساعدين ويمكن جمع كمية كبيرة من البيانات من المفحوصين في وقت قصير	1. لايمكن استخدام الاستبانة مع الافراد الاميين والصغار في السن . 2. زيادة نسبة الفاقد من الاستجابات وذلك لتسليم الكثير من المفحوصين استبياناتهم دون الاجابة على جميع الاسئلة او عدم ارسالها نهائيا 3. يتطلب بناء الاستبانة مهارة فائقة في الاعداد من حيث اختيار البنود التي تغطي كافة مجالات الظاهرة المراد قياسها .	اهم وسيلة من وسائل جمعي البيانات في المصادر الميدانية
المقابلة Interview	هي مجموعة من الاسئلة المقرؤة ويتم الاجابة عليها من قبل المستجيب شفهيا	1. امكانية الحصول على استجابات لكل البنود التي تتضمنها استمارة المقابلة . 2. امكانية الحصول على المعلومات المراد معرفتها وفقا لتسلسل البنود الواردة في القائمة ووفقا لترتيب الباحث 3. امكانية الحصول على المعلومات بدرجة تكون اكثر دقة وذلك لعدم تاثر المفحوص بمشاورة الاخرين 4. قلة نسبة الفاقد في الاجابات للبنود المستفسر عنها وتصلح للاميين والاطفال	1. قد لا تتصف البيانات المتحصل عليها من المفحوصين بالموضعية حيث قد تتاثر بالتحييز الشخصي من قبل الباحث نفسه او مساعديه 2. تتطلب كثير من الجهد والوقت والمال .	

(المحاضرة الثالثة)

أساليب إجراء البحث الميداني

❖ أساليب إجراء البحث الميداني :

عند القيام بالبحث والاعماد على المصادر الميدانية في الحصول على البيانات يواجهناتسأول هاملابد من الاجابة عليه من قبل الباحث :

⇐ هل ستشمل الدراسة جميع مفردات (افراد) المجتمع الاحصائي ام ستطبق على جزء منه ؟

في حال اعتماد البحث على دراسة جميع مفردات المجتمع الاحصائي يسمى ذلك اسلوب الحصر الشامل .

ام اذا اعتمد على دراسة جزء فقط من مفردات المجتمع الاحصائي فإن ذلك يسمى باسلوب العينة

ان كلا من الاسلوبين يمكن تطبيقه لجميع الحالات وهناك حالات ترغمنا الاسباب الى تطبيق احد هذين الاسلوبين .

الاسلوب	تعريفه	الحالات التي يجب استخدامه فيها	عيوبه
الحصر الشامل	يمكننا من الحصول على كافة البيانات والمعلومات عن طافة مفردات المجتمع الاحصائي وبالتالي فان النتائج التي نحصل عليها لايوجد بها تحيز ولا تحتاج لتعديل .	1. التعدادات : السكانية والمناطق الصناعية والمؤسسات. 2. الحالات التي اذا تركت بعض مفرداتها دون فحص قد تؤدي الى الحاق الضرر بالمجتمع كله : مثل المرضى المصابين بمرض انفلونزا الطيور – التطعيم	يتطلب وقت وجهد كثير وتكلفة كبيرة ولا يصلح في حالات المجتمعات غير المحددة
اسلوب العينات	عكس اسلوب الحصر الشامل وتقتصر فيه الدراسة على جزء من مفردات المجتمع الاحصائي ولهذا فهو يوفر الوقت والجهد والتكاليف ويصلح للمجتمعات الغير محدودة	1. ما يميزه هو ادراسة المجتمعات التي ينتج عن فحص ودراسة مفرداتها هلاك تلك المفردات مثل : فحص اللمبات الكهربائية المنتجة ، فحص دم الانسان ، فحص البيض في مزارع الدواجن	اهم عيوبه هو ما يسمى بخطاء التحيز sampling bais وقد يقع فيه الباحث بقصد او بدون قصد نتيجة عدم تمثيل العينة تمثيلا صادقا وكاملا لمفردات المجتمع الاحصائي محل الدراسة والذي قد يرجع الى تحيز الباحث لفكرة او رأي معين او التحيز لمفردات العينة .

في حال اتباع الباحث لاسلوب العينات في اجراء دراسته فيتوجب عليه ان يختار ما يناسب طبيعة البحث الذي يجريه من انواع العينات المختلفة بحيث تكون تلك العينات تعبر عن جميع خصائص المجتمع محل الدراسة وكذلك تحديد حجم العينة المناسب و الذي يمكن الاعتماد عليه في الوثوق بالنتائج التي يتم التوصل اليها من خلال تلك العينة وامكانية تعميمها على المجتمع باكمله .

أنواع العينات

المجتمع الأصلي
(كل ما تعمم عليه نتائج الدراسة)

المجتمع المعروف
(المجتمع الذي نستطيع ان نحيط بجميع افراده مثل طلاب التعليم المطور بالانتساب حيث اننا نعرف عددهم)
الطريقة الاحتمالية

العينات الاحتمالية او الطريقة الاحتمالية
وهي الاكثر استخداما واهمية حيث يكون فيها اختيار مفردات العينة يحقق مبداء العشوائي وان لكل مفردة يكون لها احتمال الظهر ضمن مفردات العينة المختارة من المجتمع الاصلي الاحصائي محل الدراسة

الجتمع المتجانس المعروف
هو المجتمع الذي لا يكون فيه صفات تميز بين افراد المجتمع والباحث يرغب في دراسة هذه الصفات مثلا طلاب التعليم المطور لو تم النظر اليهم كوحده واحده متماثله بغض النظر عن تخصصاتهم فهذا مجتمع معروف متجانس

العينة العشوائية
وفيها يتم ترقيم مفردات المجتمع الاحصائي ثم نستخدم الحاسب الالي او جداول الاعداد العشوائية لاختيار المفردات وهناك الطريقة البسيطه منها لو كان عدد مفردات المجتمع قليل كـ 200 شخص مثلا واردنا ان نستخرج 50 بطرق القرعة . مع الاشارة الى ان نسبة اختيار اي فرد من المجتمع في هذه العينة تكون متساوية للاخرين

العينة المنظمة
وفيها يتم تقسيم مفردات المجتمع الى شرائح متساويه الطول بعدد مفردات العينة المطلوب اختيارها ويتم ترقيم مفردات كل شريحة ، مثال لو لدينا مائة عامل ونريد ان اختار 5 عمال فانه يتم تقسيم العمل الى 5 شرائح طول كل منها 20 عامل ونرقم كل شريحه ، فاذا سحبنا مفردة من الشريحة الاولى وكان الرقم 9 مثلا فان ذلك يعني ان مفردات العينة هي الارقام التالية 9 ، 29 (9 + 20 طول الشريحة) ،49 (29+ 20) - 69 ، 89 -- للاستزاده ص 29

المجتمع المعروف الغير متجانس
هو البحث في الصفات التي تميز بين افراد المجتمع من قبل الباحث كطلاب التعليم المطور وفق الكليات التي يدرسونها فهذه صفات تميزهم واصبح المجتمع معروف غير متجانس

العينة الطبقية
وفيها يكون المجتمع مقسم الى طبقات غير متجانسة لذلك يتم اختيار مفردات العينة من كل طبقة على حده باستخدام العينة العشوائية البسيطة

العينة العنقودية
يكون فيها المجتمع مقسم لطبقات وكل طبقة مقسمة الى طبقات فرعية وكل فرع مقسم الى طبقات صغيرة وفي النهاية نأخذ جميع المجتمع الاخير الناتج عن العملية : كاجراء دراسة على طلاب المدارس الابتدائية فيمكن اختيار العينة كالتالي : يتم اختيار احد المناطق في المملكة عشوائياً و بعد ذلك نختار احد محافظات تلك المنطقة عشوائيًا وبعد ذلك نختار حي عشوائيًا من هذه المحافظة و بعد ذلك نختار مدرسة ابتدائية واحده من هذا الحي عشوائيًا وهنا يتم دراسة **جميع** الطلاب الموجودين داخل هذه المدرسة .

المجتمع الغير معروف
(المجتمع الذي الذي لا نستطيع ان نحيط فيه بجميع افراد المجتمع مثل العمالة الغير نظامية في السعودية والتي لم يتم القبض عليها او المدخنين مثلا)

العينات غير الاحتمالية
هي العينات التي لا يتم فيها تطبيق مبدأ العشوائية عن اختيار مفرداتها

المجتمع المتجانس غير المعروف
كالعمالة الغير نظامية والنظر اليهم بغض النظر عن الجنسية .

العينات العمدية
وفيها يقوم الباحث باختيار المفردات بنفسه وفقًا لمعيار معين لتحقيق غرض معين يخدم اهداف الدراسة كان تذهب مثلا الى عيادة المدخنين لاختيار العينات

عينة الصدفة
الاختيار يتم من خلال المصادفة مثل أن تجد شخصاً بالصدفة تنطبق عليه شروط العينة التي تريدها وتطلب منه ان يكون عينة لدراستك
(غير موجوده في الكتاب)

المجتمع الغير متجانس غير المعروف
كالعمالة الغير نظامية عندما نميزهم بجنسياتهم

العينة الحصية
ويتم تحديد العينة فيها بناءا على حصص محددة لكل مساعد من مساعدين الباحث على ان تترك الحرية كامله لكل مساعد في اختيار مفردات عينته

❖ اقسام مجتمع البحث :

بعض العلماء قسم مجتمع البح الى القسمين :

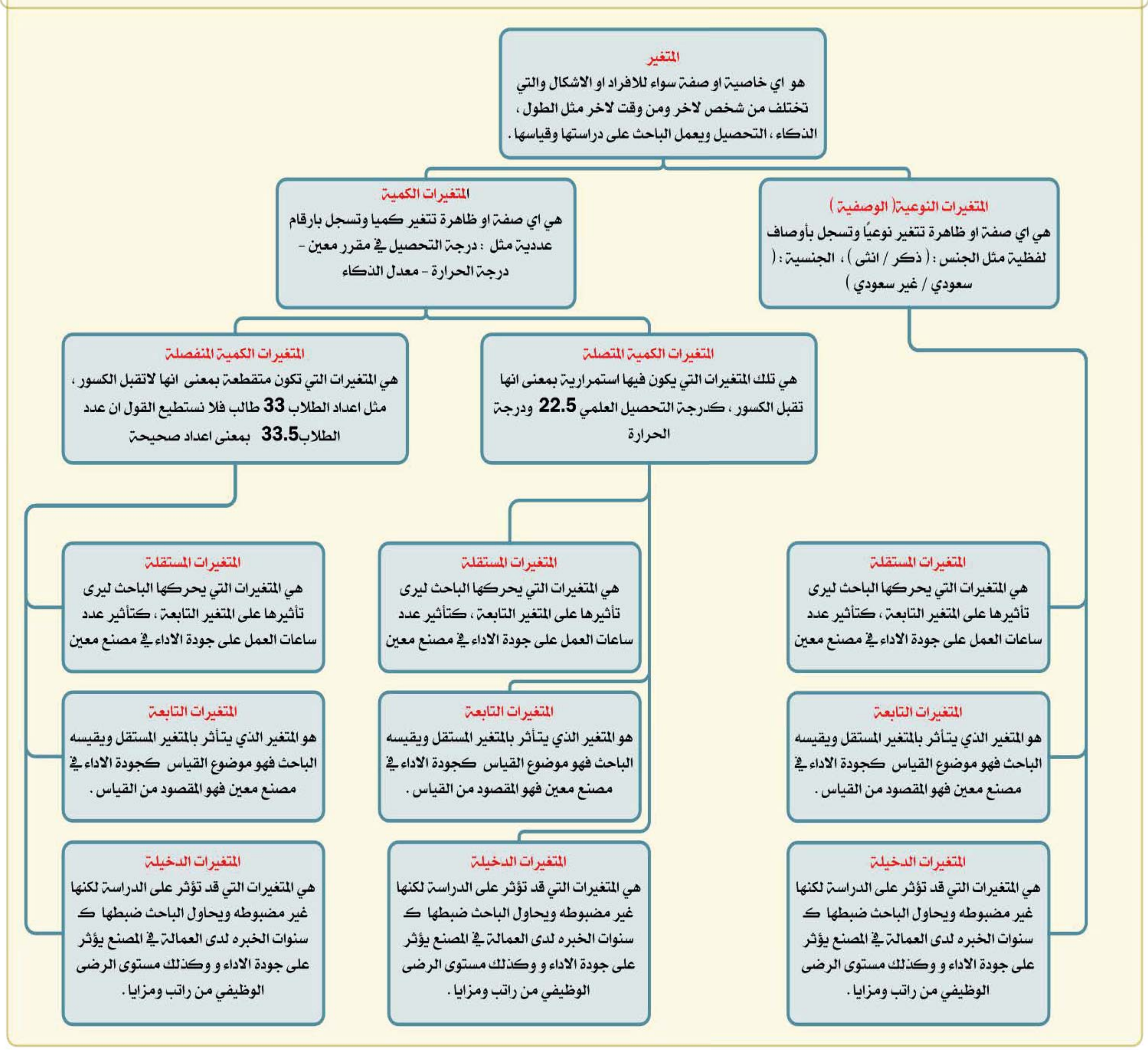
⇐ المجتمع الكلي للبحث : ويعني كل من يمكن ان تعمم عليه نتائج البحث .

⇐ المجتمع الذي يمكن التعرف عليه : يعني القائمة التي يمكن للباحث ان يتعرف عليها .

مثال : دراسة تقويمية لمباني المدارس الحكومية في المملكة العربية السعودية ، مع التطبيق على بعض المدارس الحكومية في المنطقة الشرقية.

- تعريف مجتمع البحث : هو مصطلح علمي منهجي يراد به كل من يمكن ان تعمم عليه نتائج البحث . فمجتمع البحث في المثال اعلاه يشمل كل مبنى مدرسي حكومي في المملكة .
- تعريف عينة البحث : هي جزء من المجتمع اختير بطريقة علمية بشرط ان تمثل المجتمع ككل . فعينة البحث في المثال اعلاه تشمل بعض المباني المدرسية الحكومية في المنطقة الشرقية

تصنيف المتغيرات



❖ المتغير و الثابت في البحث العلمي :

⇐ المتغير : هو اي خاصية او صفة سواء للافراد او الاشكال والتي تختلف من شخص لاخر ومن وقت لاخر مثل الطول ، الذكاء ، التحصيل ويعمل الباحث على دراستها وقياسها .

⇐ الثابت : هي الصفات او الظواهر التي لا تتغير ، او اي صفة او خاصية تأخذ صفة واحدة ، ومن الممكن اخذ متغير وتحويله الى ثابت مثل درجة الحرارة في الغرفة ، والباحث يسعى الى تثبيت عدد من المتغيرات في دراسته للتخلص من تأثيرها .

❖ الخطوات الواجب مراعاتها بعد جمع البيانات:

هناك عدد من الخطوات يجب على الباحث مراعاتها بعد جمع البيانات منها :

1- تسجيل البيانات : وذلك من خلال استخدام الطرق المناسبة لهذا الامر ، مثل استخدام التقنيات الحديثة كالحاسب الالي .

2- ترميز البيانات : هي عملية تحويل البيانات من الصيغة اللفظية الى الصيغ الرمزية او الرقمية و التي تساعد على عملية تحليل البيانات وجعل البيانات اكثر ملائمة للمعالجة والتشغيل ، واختصار وتبسيط كمية البيانات المطلوب تسجيلها ، وهناك عدة نظم للترميز (وتسمى التكويد) منها :

I. الترميز الرقمي أو العددي: ويقصد بالترميز العددي استخدام الارقام بصورة متتالية لتمييز مفردات البيانات ، فمثلا يستخدم الرقم (1) للذكور والرقم (2) للإناث لتميز الجنس في البيانات الشخصية .

II. الترميز الأبجدي أو الحرفي : ويقصد بالترميزد الأبجدي هو الحرف استخدام الحروف بدلا من الارقام لتمييز مفردات البيانات ، فمثلا استخدم الحرف M للذكور والحرف F للإناث لتمييز الجنس في نظام البيانات الشخصية .

III. الترميز الأبجدي الرقمي : ويقصد بالترميز الابجدي الرقمي استخدام الحروف والارقام لتمييز مفردات البيانات ، فمثلا استخدام الحرف والرقم L1 للمستوى الدراسي الاولى و L2 للثاني وهكذا لتمييز المستويات الدراسية الجامعية للطلاب والطالبات .

3- تصنيف البيانات : وتعني عملية تقسيم البيانات الى مجموعات نوعية ذات خواص مشتركة وذلك لغرض تسهيل وتيسير عملية التعامل معها ، مثلا عند تقسيم الطلبة عند دخولهم الجامعة بعد مرحلة الثانوية يتم تقسيمهم الى علمي وادبي .

4- مراجعة وتنقية البيانات : تهدف هذه الخطوة الى التحقق من صحة البيانات واكتمالها وخلوها من الاخطاء وان عملية التسجيل في الحاسب تمت بدقة ، ولهذا فنستطيع القول ان هناك اخطاء حدثت اثناء تسجيل البيانات واخطاء حدثت نتيجة بيانات خاطئة لم يتم مراجعتها جيدا . وهي من اهم الامور في عملية نجاح التحليل والوصول الى نتائج دقيقة .

❖ ترميز بيانات الاستبانة وجعلها متاحة لبرنامج الـ SPSS:

تعتبر الاستبانة من أكثر وسائل جمع البيانات البحثية استخداما، لذلك سوف نقوم الآن بالتعرف على كيفية تبويب البيانات التي يتم الحصول عليها من خلال الاستبانة، وطريقة إدخالها في برنامج الـ SPSS

⇐ مثال: لو كنت تقوم بدراسة إحصائية حول موضوع "واقع استخدام الانترنت في البحث العلمي في الجامعات السعودية"، فإنك ستحتاج إلى إعداد استبانة تحوي مجموعة من الاسئلة تتعلق بهذا الموضوع، ومن ثم توزيع هذه الاستبانة على عينة ممثلة لمجتمع البحث الذي تريد أن تعمم نتائج دراستك عليه، وتطلب من أفراد العينة الإجابة على جميع فقرات الاستبانة .

وهناك مفاهيم مهمه يجب عليك ان تعرفها وهي :

⇐ حتى تستطيع تفريغ البيانات المجموعة من خلال هذه الاستبانة بطريقة مناسبة يفهمها برنامج الـ SPSS يتوجب عليك معرفة الامور التالية :

1) الأفراد الذين يقومون بالإجابة على أسئلة الاستبانة يطلق عليهم اسم **حالات** Cases
2) كل سؤال (فقرة) في الاستبانة تمثل **متغير** Variable
3) تسمى إجابات الافراد على الاسئلة (الفقرات) **بقيم المتغيرات** Variable values

⇐ إن كل استبانة تحوي عدة أنواع من الاسئلة والفقرات، وهذه الانواع هي:

1- سؤال يسمح باختيار إجابة واحدة فقط : مثال :

(1) عدد سنوات الخبرة في العمل الاكاديمي :

1- ( ) اقل من سنة .	2- ( ) من 1 – 5 سنوات	3- ( ) من 6 – 10 سنوات
4- ( ) من 11 – 15 سنه	5- ( ) اكثر من 16 سنه .	

ففي هذا السؤال هناك خمس احتمالات فتعطى كل اجابة رقم يمثلها فعلى سبيل المثال :

أقل من سنة القيمة (1)	من 1 – 5 سنوات القيمة (2)	من 6 – 10 سنوات القيمة (3)
من 11 – 15 القيمة (4)	اكثر من 16 سنه القيمة (5)	

وبالامكان ان تعطى هذه الاجابات رموزا حرفية اذا تم تعريف المتغير على انه متغير من نوع حرفي ولكن يفضل عدم استخدام مثل هذا الاجراء وذلك لان ادخال البيانات الرقمية في برنامج SPSS اسهل .

2- سؤال يسمح باختيار أكثر من إجابة واحدة :

وهو ذلك النوع من الاسئلة التي تتاح من خلالها الفرصة للمستجيب لاختيار اكثر من اجابة ، وفي هذا النوع من الاسئلة قد يختار الفرد اكثر من اجابة على السؤال ، ولذلك فان متغيرا واحدا لايكفي لتمثيل هذا السؤال بل يحتاج هذا السؤال الى تسعة متغيرات وكل متغير منها له احتمال اجابتين (نعم وتأخذ القيمة " 1 " ، و لا وتأخذ القيمة " 0 " مثلا) ، مثال :-

(1) ما اهم المعوقات التي قد تحول دون استخدامك للانترنت في البحث العلمي ؟ (يمكن اختيار اكثر من عائق) :

1- ( ) عدم الاهتمام بالانترنت	2- ( ) عدم توفر التدريس المناسب لاستخدام الانترنت	3- ( ) عدم وجود الوقت الكافي .
4- ( ) عدم توفر اجهزة الحاسب	5- ( ) عدم توفر المتصفح المناسب للانترنت	6- ( ) عدم توفر الحوافز الخارجية لاستخدام النترنت في البحث العلمي
7- ( ) عدم توفر المعلومات والمهارات الاساسية الانترنت	8- ( ) الاهتمام بحقوق	9- ( ) الخوف من العولمة

3- سؤال مفتوح جزئيا :

هو النوع الذي يسمح للمستجيب باختيار اجابة موجوده من ضمن الخيارات او كتابة اجابة خرى مثال :

(1) الدرجة العلمية التي تحملها ؟

1- ( ) دكتوراه	2- ( ) ماجستير
3- ( ) بكالورويوس	4- ( ) غير ذلك ، حدد ....

فهذا النوع من الاسئلة يتم تمثيله بمتغير واحد فقط ، لان المطلوب من المستجيب اختيار اجابة واحده ، الا ان المشكلة في هذا النوع من السئلة تكمن في الخيار ذو الاجابة المفتوحة ، ففي هذا السؤال هناك اربعة احتمالات ، فتعطى كل اجابة رقم يمثلها كالدكتوراه رقم 1 و الماجستير رقم 2 و البكالوريوس رقم 3 اما الخيار الرابع فيتم التعامل معه باكثر من طريقة منها :

- تعيين قيمة محددة لهذا الاحتمال : وليكن القيمة (4) بغض النظر عما ذكر من درجات علمية في داخله (ثانوية – متوسطة – دبلوم ، الخ ...) وهذا الاجراء يسهل التعامل مع هذا الخيار الا انه يُفقد الباحث معلومات كثيرة .

- حصر جميع الاجابات ومن ثم تحديد قيمة لكل درجة علمية غير تلك التي ذكرت في السؤال : وهنا يتم تحديد عدد الاحتمالات المتاحة للسؤال بعدد الاجابات المذكورة في الاستبانات ، وهذا الاجراء يحتاج الى وقت كبير لانه يتم معالجة كل استبانة بشكل منفرد ليتم جمع كل الاجابات الممكنة .
- عدم التعامل مع هذا المتغير على انه متغير رقمي Numeric والتعامل معه على انه متغير حرفي string : لذا لايتم تعيين قيم تصف الاجابات بل يتم كتابة الاجابة كاملة لكل درجة علمية / وهذه الطريقة تؤدي الى حصر جميع الاجابات الا انها تزيد العبء على الباحث من خلال ادخال بيانات اكثر في الحاسب مما قد يؤدي الى زيادة اخطاء الادخال .

تمارين الكتاب

أراد باحث معرفة العلاقة بين حب الاستطلاع لدى الطلاب في السنوات الابتدائية وحل المسائل الرياضية، فاختار عشوائيا طلاب السنة الثالثة ثم اختار منهم عشوائيا 200 طالب، ثم قام بصياغة الفرضية التالية:

"لا توجد علاقة ذات دلالة إحصائية بين حب الاستطلاع وحل المسائل الرياضية"

ثم قام بتطبيق اختبار عليهم وذلك للحصول على البيانات اللازمة لاستنتاج العلاقة واتخاذ قرارات في ضوء ذلك

المطلوب :

- ما نوع الإحصاء الذي استخدمه الباحث في هذه الدراسة؟ علل ذلك ؟ نوع الاحصائي الاحصائي التحليل لوجود الرغبة في اتخاذ القرار في السؤال
- حدد مجتمع البحث في هذه الدراسة ، وما نوعه ؟ مجتمع البحث : جميع الطلاب في السنوات الابتدائية ، نوعه مجتمع معروف
- حدد عينة الدراسة في هذه الدراسة ، وما نوعها؟ عينة الدراسة : 200 من طلاب السنة الثالثة ابتدائي ، نوع العينة : عنقودية
- حدد المتغير المستقل في هذه الدراسة ، وما نوعه ؟
- حدد المتغير التابع في هذه الدراسة ، وما نوعه ؟
- حدد في تصورتك المتغيرات الدخيلة التي من الممكن أن تؤثر على هذه الدراسة ؟
- حدد الفرضية التي يحاول الباحث اختبارها في هذه الدراسة ، وما نوعها ؟

ما الوسيلة التي استخدمها الباحثة لجمع البيانات في هذه الدراسة ؟

(المحاضرة الرابعة)

العرض الجدولي للبيانات (تبويب البيانات)

الجزء الأول

❖ تمهيد :

هناك عدة طرق لعرض وتبويب البيانات الا ان من ابسط تلك الطرق للتعبير عن البيانات هي ان تدمج هذه البيانات في صيغة كتابية كأن نقول مثلا :

بلغ عدد المتقدمين بكلمية العلوم الادارية بجامعة الملك فيصل في الفصل الدراسي الاول من عام 1427 هـ 1800 طالب وطالبة ، منهم 700 طالب و 1100 طالبة ، في حين بلغ عدد المتقدمين للكلية في نفس الوقت من العام الماضي 1200 طالب وطالبه ، منهم 500 طالب و 700 طالبة ، اي ان عدد الطلاب في السنة الحالية يزيد بمقدار 600 طالب وطالبة عن السنة الماضية ويرجع السبب في ذلك اللاى زيادة اعداد الخريجين من الثانوية العامة.

وعلى الرغم من ان طريقة العرض هذه تمتاز على غيرها من الطرق بالسهولة التامة بسبب امكان توضيح الارقام بعبارات تفسية كلما دعت الحاجة الى ذلك الا ان هذه الطريقة يشوبها الكثير من العيوب منها :

- انها طريقة مطولة وعقيمة
- تتطلب وقتا طويلا في القراءة الشيء الذي يجعل الملل يتسرب الى القارئ .
- قلما يكلف الانسان نفسه مشقة الاطلا على احصاءات معروضة بهذه الكيفية .
- انه يتعذر عرض البيانات الخاصة بعدد كبير من السنين بهذه الطريقة .

ونتيجة لهذه العيوب فان هذه الطريقة لا تعد من الطرق الفنية في العرض الاحصائي ، اما الطرق الفنية في عرض البيانات الاحصائي والتي لا تحتوي هذه العيوب فهي :

1- العرض الجدولي للبيانات (تبويب البيانات) : وفيه يتم تلخيص البيانات محل الدراسة وتصنيفها في صورة جداول تعبر عن القيم التي اخذها المتغير من خلال البيانات التي جمعها وتكرار كل قيمة من تلك القيم . (هي محل المحاضرتين الرابع والخامسة)

2- العرض البياني للبيانات : هي وسيلة مفيدة وفعالة لتوضيح وشرح الحقائق الرقمية وابراز العلاقات بين المتغيرات ، واستقراء اتجاهاتها العامة بأسلوب يسهل فهمه وتذكره بمجرد النظر اليه .

❖ أهمية الجداول الاحصائية :

- تعبر عن الحقائق الكمية المعروضة بعدد كبير من الارقام عن طريق عرضها في جداول بطريقة منظمة تسهل اكتشاف اهمية هذه الارقام والاستفادة منها.
- هي وسيلة سهله لتلخيص المعلومات الرقمية الكثيرة العدد ،و المتغيرة القيم ، مما يسهل التعرف عليها.
- تستوعب الجداول بهسولة عدد كبير من الموضوعات لان تفريغ الارقام فيها يقلل كثيرا من تكرار الكلمات التي تصف البيانات وبالتالي فهي طريقة اقتصادية في الوقت والحيز والمجهود .
- اظهار البيانات بأكبر وضوح ممكن وأصغر حيز مستطاع .

❖ تكوين الجداول :

1) رقم الجدول: يجب ان يرقم كل جدول حتى تسهل الاشارة اليه.

2) العنوان: يجب أن يعطي كل جدول عنوانا كاملا لتسهيل مهمة استخراج المعلومات منه، ويجب أن يكون هذا العنوان واضحا قصيرا بقدر الامكان، ويستخدم في بعض الاحيان عنوان توضيحي لبعض الجداول وذلك من أجل إعطاء معلومات إضافية عن بيانات الجدول.

3) الهيكل الرئيسي: ويتكون هيكل الجدول من أعمدة وصفوف، ويعتبر ترتيب المعلومات في الاعمدة والصفوف أهم خطوة في تكوين الجدول.

4) العمود: إن كل جدول يتكون من عمود أو اكثر ويوجد لكل عمود عنوان يوضح محتوياته.

5) الحواشي: قد يحتوي الجدول على مفردات بيانات لا ينطبق عليها عنوان الجدول أو عنوان العمود، ففي هذه الحالة تستعمل الحواشي لتوضيح ذلك وذلك اما بترقيم الملاحظات او باستعمال علامة (❖) .. الخ.

6) المصدر: قد تؤخذ بيانات الجدول من مصادر جاهزة لذلك يجب إظهار المصدر في أسفل الجدول حتى يمكن الرجوع اليه عند الحاجة.

رقم الجدول ← جدول رقم (□)

عنوان الجدول ← يوضح طلبة جامعة الملك فيصل للعام الجامعي□□□□هـ

عنوان توضيحي ← (مصنفون حسب الجنس)

عنوان العمود

هيكل الجدول (الاطار العام للجدول : المستوى - طالب - طالبة - المجموع)

عمود

المستوى❖	طالب	طالبة	المجموع
الأول	200	250	450
الثاني	100	120	220
الثالث	80	110	190
الرابع	100	120	220
المجموع	480	600	1080

المصدر: جامعة الملك فيصل، احصائية الجامعة حسب الكليات

❖ يحدد المستوى بالسنة الدراسية التي يدرس فيها الطالب .

❖ أنواع الجداول الاحصائية : تقسم الجداول تبعا لدرجة تعقيدها الى:

1) جداول بسيطة: وفيها يتكون كل من موضوع الجدول ومادته من بضع أسطر وخانات تتعلق بالتقسيمات الزمانية (أي الأمور التي يتناولها الجدول أمور تتسلسل حسب السنوات) أو المكنية (أي توزيع الظاهرة حسب المكان) أو مؤشرات وصفية بسيطة وبأرقام بسيطة أيضا.

2) جداول التوزيع التكراري: وفيها تكون المعطيات مجمعة في فئات بمؤشر أو متغير واحد، ولكل فئة تكراراتها الخاصة عند ذلك المؤشر

3) جدول التوزيع التكراري المتجمع: وفيه تجمع التكرارات على التوالي من أحد طرفي الجدول الى طرفة الآخر فنحصل على التكرار الكلي (مجموعة التكرارات)، (فاذا بدأ من أعلى الى أسفل الجدول) سمي **جدول تكراري متجمع صاعد**، (واذا بدأ من أسفل الى أعلى الجدول) سمي **جدول تكرار متجمع نازل أو هابط**.

4) الجداول المزدوجة أو المركبة: وهي الجداول التي تتكون من متغيرين أو أكثر، وهذه المتغيرات قد توزع على أعمدة وحقول الجدول بصورة نظامية، تعبر عن الافكار العلمية التي يريد الباحث توضيحها توضيحا عدديا.

وتتوقف عملية تبويب وتصنيف البيانات على نوع البيانات الإحصائية المراد التعامل معها ودراستها والتى يمكن تقسيمها من حيث طريقة إعداد الجداول إلى مجموعتين:

1. مجموعة البيانات الوصفية والكمية المتقطعة
2. مجموعة البيانات الكمية المتصلة

للتذكير !! تم التطرق الى هذه الانواع في المحاضرة الثالثة:

- البيانات الوصفية او النوعية : هي اي صفة او ظاهرة تتغير نوعيًا وتسجل بأوصاف لفظية مثل الجنس : (ذكر / انثى) ، الجنسية : (سعودي / غير سعودي) .
- البيانات الكمية المتقطعة : هي المتغيرات التي تكون متقطعة بمعنى انها لاتقبل الكسور ، مثل اعداد الطلاب 33 طالب فلا نستطيع القول ان عدد الطلاب 33.5 بمعنى اعداد صحيحة .
- المتغيرات الكمية المتصلة : هي تلك المتغيرات التي يكون فيها استمرارية بمعنى انها تقبل الكسور ، كدرجة التحصيل العلمي 22.5 ودرجة الحرارة

أولا: البيانات الوصفية والكمية المتقطعة:

وفيها يتم تصنيف وحساب تكرار كل عنصر من العناصر الواردة في بيانات المتغير الذي يتم دراسته كما يمكن حساب التكرار النسبي لكل عنصر من خلال حساب نسبة تكراره إلي مجموع التكرارات.

⇐ مثال (لمتغير وصفي):

فى دراسة قام بإجرائها أحد الأطباء لطفل معرض لأحد الأمراض النفسية فتم سؤاله عن لون مجموعة من الأشياء فكانت إجاباته كما يلى:

أحمر	أزرق	بنفسجي	أحمر	اخضر
أبيض	أبيض	أحمر	اخضر	أبيض
أزرق	أحمر	اخضر	أحمر	بنفسجى
اخضر	أزرق	أبيض	بنفسجى	أحمر

المطلوب: عرض البيانات السابقة فى صورة جدول التوزيع التكرارى ؟

✓ طريقة الحل :

1- نرسم جدول ونحصر فيه الالوان المستخدمة في الاستبيان او الاختبار

2- نقوم بعملية العد بجانب عمود الالوان عن طريق العد بالشرطات (كل خمس شرطات تكون مجموعة في العد في نظام العد بالشرطات) وتسمى حزمه

اللون	التفريغ (العد بالشرطات – كل خمس شرطات تكون ما يسمى بالحزمة)	التكرارات (بالارقام)	التكرار النسبي التكرارات / اجمالي التكرارات
الاحمر	~~IIIII~~ I	6	0.3 = 20 ÷ 6
الازرق	IIII	4	0.2 = 20 ÷ 4
البنفسجي	III	3	0.15 = 20 ÷ 3
الابيض	IIII	4	0.2 = 20 ÷ 4
الاخضر	III	3	0.15 = 20 ÷ 3
المجموع		20	1

نلاحظ : مجموع التكرار النسبي يجب ان يكون 1 صحيح

⇐ مثال (لمتغير كمى متقطع)

تم سؤال عدد من طلاب كليتي الآداب وإدارة الأعمال عن عدد حوادث السيارات التى تعرضوا لها خلال العام الماضى فكانت اجاباتهم كما يلى:

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

◄ المطلوب :

1- عرض البيانات السابقة فى صورة جدول تكرارى

2- أحسب الأحتمالات التالية:

- أن لا يتعرض أى شخص لحادث
- أن يكون هناك حادث واحد على الأكثر
- أن يكون هناك حادث واحد على الأقل

✓ طريقة الحل :

عدد الحوادث	التفريغ (العد بالشرطات – كل خمس شرطات تكون ما يسمى بالحزمة)	التكرارات (بالارقام)	التكرار النسبي التكرارات / اجمالي التكرارات
صفر	𝍸 IIII	9	9 ÷ 30 = 0.30
1	𝍸 𝍸 I	11	11 ÷ 30 = 0.36667
2	𝍸 II	7	7 ÷ 30 = 0.23333
3	III	3	3 ÷ 30 = 0.10
المجموع	30		1

◄ احتمال الا يتعرض اي شخص لحادث ؟

✓ صفر = 0.30

◄ احتمال ان يكون هناك حادث واحد على الاكثر ؟

✓ بمعنى انه سوف يكون اكبر عدد للحوادث هو 1

وذلك بجمع نسبة تكرار الصفر + 1 = 0.30 + 0.36667 = 0.66667

◄ احتمال ان يكون هناك حادث واحد على الاقل ؟

✓ بمعنى انه سوف يكون اقل عدد للحوادث هو 1

لذلك لا بد من استبعاد نسبة تكرار الصفر ونحسب باقي الاعداد ونختصرها في هذا المثال بطرح قيمة تكرار الصفر من اجمالي التكرار النسبي 1

1 - 0.30 = 0.70

❖ ثانيا: البيانات الكمية المتصلة:

وفيها يتم توزيع البيانات في جدول تكراري ذوفئات، ويتم ذلك من خلال اتباع الخطوات التالية:

1- الخطوة الأولى: **تحديد عدد الفئات:**

ويمكننا إتباع قاعدة Sturge's Rule كأساس عند تحديد عدد الفئات, وتنص القاعدة علي وجود علاقة بين عدد المفردات المتاحة عن الظاهرة محل الدراسة (عينه البحث) وبين عدد الفئات, وتستخدم القاعدة الرقم 2 أساس مرفوع للقوه K . وجدير بالذكر هنا أن بتطبيق قاعدة "Sturge" علي عينات بأحجام مختلفة نحصل علي الجدول التالي:

حجم العينة	عدد الفئات
16 – 11	4 – 3
32 – 16	5 – 4
64 – 32	6 – 5
128 – 64	7 – 6
256 – 128	8 – 7
512 – 256	9 – 8
512 فأكثر	10 فاكثر

2- الخطوة الثانية: **تحديد طول الفئة:**

بعد قيامنا بتحديد عدد الفئات في الخطوة السابقة، فإن الخطوة الحالية هي قيامنا بتحديد طول الفئة، ويفضل أن تكون الفئات كلها ذات أطوال متساوية, إلا فى بعض الحالات التى تحتم علينا الظاهرة التالية لتحديد طول الفئة: **طول الفئة = المدى ÷ عدد الفئات**

ويمثل المدى الفرق بين اكبر مفرده واصغر مفرده في البيانات الأولية مثال :

لو افترضنا ان لدينا بيانات اقل قيمه فيها صفر واعلي قيمه فيها 20 **فإذا المدى هنا** = 20 – صفر = 20

لو افترضنا ان لدينا 6 فئات **فإن طول الفئة في هذا المثال** = 20 ÷ 6 = 3.33

ملاحظة بالنسبة للتقريب في الاحصاء : اقل من النصف 0.5 نقرب الى الرقم السابق ، اكبر او يساوي 0.5 نجبر الكسر

ففي المثال السابق 20 ÷ 6 = 3.33 ≅ 3

3- الخطوة الثالثة: **تعيين حدود الفئات:**

نبدأ بتعيين الحد الأدنى للفئة الأولى وهو قيمه اصغر مفرده في البيانات الأولية للظاهرة محل الدراسة، ويجوز أن نختار قيمه أقل من أصغر مفرده ليبدأ الحد الأدنى للفئة الأولى بقيمه صحيحة, ونقوم بتحديد الحد الأعلى للفئة الأولى بإضافة طول الفئة الذى حصلنا عليه من الخطوة الثانية. يعتبر الحد الأدنى للفئة الثانية هو الحد الأعلى للفئة الأولى وبإضافة طول الفئة نصل إلى الحد الأعلى للفئة الثانية، ونستمر في تكرار هذه الطريقة حتى يتم تكوين عدد الفئات المطلوبة المحدد فى الخطوة الأولى.
يجب علينا التأكد من عدم وجود تداخل فيما بينها الفئات بعضها البعض، حيث أن الفئة تحتوى على كل المفردات التى تساوى حدها الأدنى تماماً وما يزيد عنه حتى يصل إلى حدها الأعلى.

4- الخطوة الرابعة: **توزيع التكرارات على الفئات:**

نبدأ الآن في توزيع مفردات العينة بحسب الفئات المقابلة كى نصل إلى التوزيع التكرارى، وهو عبارة عن جدول مكون من عمودين، يحتوى العمود الأول على فئات المتغير العشوائي ويحتوى العمود الثانى على عدد مرات تكرار كل مفرده أمام الفئة الخاصة بها ويسمى التكرار الاصلي. ويجب أن يكون مجموع التكرارات ألا صليه مساويا لحجم عينه الدراسة

⇐ **مثال :**

البيانات التالية تعبر عن رأس المال المستثمر فى شركات الحاسبات الألية بالألف ريال:

25	26	41	36	44	23	15	7	12	2
13	21	33	35	45	22	26	12	22	3
43	41	30	32	48	18	24	23	32	5
23	16	1	9	23	11	23	32	36	6
18	17	20	21	26	20	39	36	35	7

◄ **المطلوب:** عرض البيانات السابقة فى صورة الجدول التكرارى المناسب ؟

✓ **طريقة الحل :**

1- **الخطوة الاولى :** نحدد عدد الفئات كما هو موضح في الجدول :

عدد الاعمدة 10 × عدد الصفوف 5 = 50 (تقع بين 32 – 64)

64 – 32	6 – 5

ناخذ عدد الفئات = 5 (اخذنا الاقل مثل ما فعل الدكتور والكتاب)

2- **الخطوة الثانية:** تحديد طول الفئة : المدى ÷ عدد الفئات

∴ المدى : اعلى قيمة – اقل قيمه = 48 – 1 = 47

∴ طول الفئة = 47 ÷ 5 = 9.4 ≅ 10 (في الكتاب قربها الى 10 و يشير الدكتور الى انها ترجع الى تقدير الباحث نفسه ولكن المطلوب منا ان نستخدم التقريب لما هو 0.5 او اكثر مع الاشارة الى ان تقريبها الى 10 في هذا المثال سوف يساعد على تقليل عدد الفئات)

3- **الخطوة الثالثة:** تعيين حدود الفئات :

نقوم باخذ اقل قيمه او اقل منها كما هو واضح في الخطوة الثالثة فاقل قيمة هنا هو 1 ولكن في الكتاب اخذ صفر كا اقل قيمه ونقوم بجمع اقل قيمه + طول الفئة

الفئة الاولى	من 1 الى اقل من 11 وتكتب 1 – 10 ونستطيع البدء من صفر فنقول 0 الى اقل من 10 فتصبح 0 – 10
الفئة الثانية	10 الى اقل من 20 10 – 20
الفئة الثالثة	20 – 30
الفئة الرابع	30 – 40
الفئة الخامسة	40 – 50

4- **الخطوة الرابعة :** تفريغ البيانات في الجدول الاحصائي :

فئات رأس المال	التفريغ (العد بالشرطات – كل خمس شرطات تكون ما يسمى بالحزمة)	التكرارات (بالارقام)	التكرار النسبي التكرارات / اجمالي التكرارات
0-	卌 III	8	8 ÷ 50 =0.16
10-	卌 IIII	9	9 ÷ 50 = 0.18
20-	卌 卌 卌 I	16	16 ÷ 50 = 0.32
30-	卌 卌 I	11	11 ÷ 50 = 0.22
40 – 50	卌 I	6	6 ÷ 50 = 0.12
المجموع	50		1

(المحاضرة الخامسة)

العرض الجدولي للبيانات (تبويب البيانات)

الجزء الثاني

هناك عدة ملاحظات يجب الإنتباه إليها عند عمل جدول التوزيع التكرارى لبيانات المتغير الكمى المتصل:

- إن تحديد عدد الفئات يتوقف على أمور عدة منها:
  - عدد المفردات محل الدراسة
  - انتظام وتوزيع تلك البيانات
  - طبيعة بيانات المشكلة محل الدراسة

- طول الفئة لا بد أيضاً من تحديده بعناية حيث يمثل الوجه الآخر للعملة مع عدد الفئات، فمن الأفضل أن يكون تحديده بطريقة تجعل مركز الفئة قريباً من تركز البيانات بتلك الفئة بقدر الإمكان حيث يعبر مركز الفئة عن قيمة كل مفردة من المفردات التي تنتمي لتلك الفئة

- أن تكون حدود الفئات واضحة بحيث لا يكون هناك أي تداخل فيما بينها.

❖ ومن هنا يمكن إعداد جداول التوزيعات التكرارية للمتغيرات المتصلة بثلاث صور هى:

1) الجداول التكرارية المنتظمة
2) الجداول التكرارية غير المنتظمة
3) الجداول التكرارية المفتوحة

أولا: الجداول التكرارية المنتظمة: وهى الجداول التى يكون فيها أطوال كل الفئات متساوية كما تم توضيحة فى المثال السابق في المحاضرة السابقة .

ثانيا: الجداول التكرارية غير المنتظمة: وفيها تكون أطوال الفئات غير متساوية، ومثال ذلك البيانات التالية والتي توضح توزيع عدد من العمال وفقا للاجر الذى يحصل عليه كل منهم:

فئات الاجر	10 – (20 – 10)	20 – (40 – 20)	40 – (50 – 40)	50 – 55 (50 – 55)	المجموع
عدد العمال (التكرار)	10	40	15	5	70

ويتضح لنا من الجدول السابق أن أطوال الفئات غير متساوية حيث يكون طول الفئة للفئة الأولى " 10 – " هو 10 بينما فى الفئة الثانية " 20 – " بلغ 20 وفى الفئة الثالثة " 40 – " كان 10 والفئة الأخيرة " 50 – 55 " بلغ طول الفئة فيها 5

ثالثا: الجداول التكرارية المفتوحة: وتوضحها أشكال الجداول التالية:

فئات العمر	عدد الطلاب
أقل من 6	20
6–	35
12–	25
15– 18	18

جدول مفتوح من اسفل

فئات العمر	عدد الطلاب
6 –	20
12–	35
15–	25
18 فأكثر	18

جدول مفتوح من اعلى

فئات العمر	عدد الطلاب
أقل من 6	20
6–	35
12–	25
15–	18
18 فأكثر	22

جدول مفتوح من الطرفين

❖ الجداول التكرارية المتجمعة:

وهى جداول يتم إعدادها لإعطاء نتيجة تراكمية لمجموعة من الفئات والتى يمكن أن تكون بشكل تصاعدى أو تنازلى ولكل منهما أهمية فى تفسير النتائج والظواهر المختلفة.

اولا- الجدول التكرارى المتجمع الصاعد : يعطى جدول التكرار المتجمع الصاعد الحدود العليا للفئات وعدد المفردات التى تقل عن الحدود العليا لكل فئة (وتكتب بصيغة أقل من الحد الأعلى).

⇐ مثال :

فى دراسة جغرافية لعدد من مساحات مجموعة من قطع الأراضى لمنطقة سكنية معينة تبين أن التوزيع التكرارى لها كما يلى:

فئات مساحات الاراضي دونم	عدد قطع الاراضي
1 – (معناها من الى اقل من 3 وهكذا)	14
3 –	29
5 –	18
7 – 10	9
المجموع	**70**

◄ **المطلوب:** إعداد جدول تكرارى متجمع صاعد مع بيان نسبة الأراضى التى تقل مساحتها عن 5 دونم .

✓ **طريقة الحل :**

لاعداد الجدول التكراري المتجمع الصاعد يعد جدولا من خانتين : الاولى للحدود العليا للفئات ويستخدم معها كلمة " اقل من " والثانية يتم تخصصيها لتجميع التكرارات حسب ترتيب ورودها

الحدود العليا للفئات	التكرار المتجمع الصاعد
اقل من 1	صفر
اقل من 3	14
اقل من 5	43
اقل من 7	61
اقل من 10	70

◄ بيان نسبة الاراضي التي تقل مساحتها عن 5 دونم :

✓ 43 قطعة وتمثل نسبتها $43 \div 70 = 0.61428$ ∴ 61.428 ٪ (نضربها في 100 للحصول على النسبة المئوية)

نلاحظ ان الجدول التكراري المتجمع الصاعد **يبدأ بتكرار اول فئة** ثم يتزايد حتى يصل الى مجموع التكرارات بالكامل عند اخر فئة .

ثانيا - الجدول التكرارى المتجمع الهابط (النازل): ويعطى الجدول المتجمع الهابط (النازل) الحدود الدنيا للفئات وعدد المفردات التى تكون أكثر من أو تساوى الحدود الدنيا لكل فئة (وتكتب بصيغة الحد الأدنى فأكثر). أي يتم اعداد الجدول التكراري المتجمع الهابط اذا كان المطلوب معرفة عدد المفردات التي تزيد او تساوي قيمة معينة .

⇐ **مثال :**

فى نفس المثال السابق والذي يتعلق بدراسة جغرافية لعدد من مساحات مجموعة من قطع الأراضى لمنطقة سكنية معينة تبين أن التوزيع التكرارى لها كما يلى:

فئات مساحات الاراضى دونم	عدد قطع الاراضي
1 – (معناها من الى اقل من 3 وهكذا)	14
3 –	29
5 –	18
7 – 10	9
المجموع	**70**

◄ **المطلوب:** إعداد الجدول التكراري المتجمع الهابط مع بيان نسبة قطع الأراضى التى تزيد أو تساوى 5 دونم ؟

✓ **طريقة الحل :**

الحدود العليا للفئات	التكرار المتجمع الهابط
1 فاكثر (نلاحظ ان التكرار المتجمع الهابط الذي يظهر اما الفئة 1 فاكثر هو عبارة عن اجمالي مجموع التكرارات بالكامل)	70
3 فاكثر (70 – 14 وهكذا)	56
5 فاكثر (56 – 29 وهكذا)	27
7 فاكثر	9
10 فاكثر	صفر

◄ بيان نسبة الاراضي التي تزيد او تساوي مساحتها 5 دونم ؟

✓ 27 ÷ 70 = 0.3857 اي 38.57 ٪

نلاحظ ان الجدول التكراري المتجمع الهابط يبدأ بمجموع التكرارات بالكامل ثم يتناقص حتى يصل الى الصفر عند اخر فئة.

ثالثا – الجدول التكرارى المزدوج: الجداول التكرارية البسيطة التى اشرنا إليها سابقاً تساعد فى تحليل البيانات التى تخص وتعبر عن متغير واحد فقط مثل قيمة المبيعات ومعدل التحصيل الدراسى ونسبة الذكاء ومعدل الإنجاب وغيرها من المتغيرات. الا أننا عند دراستنا لمتغيرين لتحديد العلاقة بينهما مثل العلاقة بين عدد أفراد الأسرة والمستوى التعليمى أو العلاقة بين أجور العامل ودرجة الرضاء الوظيفى أو ماشابة ذلك، فى هذه الحالة لابد من تبويب البيانات بالطريقة التى تسمح باستنتاج أو تحديد العلاقة بين المتغيرين موضوع الدراسة ويتم ذلك من خلال الجدول التكرارى المزدوج كما يتضح من المثال التالى:

⇐ **مثال:**

فيما يلى بيانات 20 طالب يعانون أحد صعوبات التعلم مع نوع كل طالب كما يلى:

النوع	صعوبة التعلم
ذكر	سمعية
أنثى	بصرية
ذكر	سمعية
ذكر	بصرية
ذكر	ذهنية
أنثى	ذهنية
أنثى	تخاطب
أنثى	بصرية
ذكر	سمعية
أنثى	ذهنية

النوع	صعوبة التعلم
ذكر	بصرية
أنثى	سمعية
ذكر	ذهنية
ذكر	تخاطب
أنثى	تخاطب
ذكر	سمعية
ذكر	تخاطب
أنثى	بصرية
أنثى	سمعية
ذكر	سمعية

◀ **المطلوب:** إعداد جدول تكراري مزدوج ؟

✓ **طريقة الحل :**

لاعداد الجدول التكراري المزدوج لابد من مراعاة مايلي :

1- يتم التعامل مع كثر متغير على حدى من حيث تحديد قيمة او تحديد فئاته وطول كل فئة في حالة المتغيرات المتصلة .

2- يعد الجدول التكراري المزدوج ويكون عدد الخانات الافقية بعدد فئات او قيم المتغير الاول وعدد الخانات الرأسية بعدد فئات او قيم المتغير الثاني ويجب تخصيص خانة لمجوع التكرارات الافقية و الرأسية .

3- التكرارات الكلية الافقية او الرأسية تكون ما يسمى بالتوزيع الهامشي للمتغير والذي يمثل التوزيع التكراري للمتغير .

اولا: يتم تفريغ البيانات :

النوع / الصعوبة	سمعية	بصرية	تخاطب	ذهنية
ذكر	~~IIII~~	II	II	II
انثى	II	III	II	II

ثانيا : حساب التكرارات والتوزيع الهامشي لكل من المتغيرين :

النوع / الصعوبة	سمعية	بصرية	تخاطب	ذهنية	المجموع
ذكر	5	2	2	2	11
انثى	2	3	2	2	9
المجموع	7	5	4	4	20

تمارين الكتاب

1- **البيانات التالية لتقديرات مجموعة من الطلاب في مقرر الاحصاء :**

جيد جدا	جيد	ممتاز	ضعيف	مقبول
ضعيف	مقبول	مقبول	جيد	جيد
مقبول	جيد جدا	جيد جدا	مقبول	ممتاز
مقبول	مقبول	ضعيف	جيد	ضعيف
جيد جدا	ضعيف	جيد	جيد	جيد

◀ **المطلوب :**

أ) إعداد الجدول التكراري للبيانات السابقة

ب) حساب التكرار النسبي

ت) ماهي نسبة النجاح في مقرر الاحصاء

✓ **طريقة الحل :**

اعداد الجدول التكراري وحساب التكرار النسبي طبعا هذه متغيرات وصفية لانها تصف تقدير كل طالب

التقدير	التفريغ (العد بالشرطات – كل خمس شرطات تكون ما يسمى بالحزمة)	التكرارات (بالارقام)	التكرار النسبي التكرارات / اجمالي التكرارات
ضعيف	~~IIII~~	5	0.20 = 25 ÷ 5
مقبول	~~IIII~~ II	7	0.28
جيد	~~IIII~~ II	7	0.28
جيد جدا	IIII	4	0.16
ممتاز	II	2	0.08
المجموع	25		1

◀ نسبة النجاح في مقرر الاحصاء :

✓ 1 – 0.20 (تمثل نسبة تكرار تقدير ضعيف) = 0.80 ، اي ان نسبة النجاح هي 80 ٪

2- في احدى الدراسات الاجتماعية لبعض الاسر في احد المناق السكنية تم السؤال عن عدد الاطفال في هذه الاسر فكانت اجابتهم كما يلي :

3	4	7	2	2
4	3	6	2	3
5	4	6	3	6
9	4	3	4	5
3	5	9	5	5
7	3	8	8	6

◀ المطلوب : اعداد الجدول التكراري لهذه البيانات ؟

✓ طريقة الحل :

الجدول التكراري للبيانات طبعا هذه متغيرات كمية منقطعة

عدد الاطفال	التفريغ (العد بالشرطات – كل خمس شرطات تكون ما يسمى بالحزمة)	التكرارات (بالارقام)	التكرار النسبي التكرارات / اجمالي التكرارات
2	III	3	0.10 = 30 ÷ 3
3	~~IIII~~ II	7	0.2333
4	~~IIII~~	5	0.1666
5	~~IIII~~	5	0.1666
6	IIII	4	0.1333
7	II	2	0.0666
8	II	2	0.0666
9	II	2	0.0666
المجموع	30		1

3- البيانات التالية تخص مساحة مجموعة من قطع الاراضي بالفدان التي شملتها احد الدراسات الجغرافية فكانت كما يلي :

3	6	7	14	25	33	45	27	38	40
28	45	42	44	30	41	32	56	51	43
22	51	11	32	42	38	51	51	23	47
38	49	34	13	27	53	18	22	36	31
14	39	47	39	31	25	40	34	32	16

◄ المطلوب :

أ) كون جدول التوزيع التكراري المناسب (بفرض وجود 6 فئات متساوية الطول)

ب) كون الجدول التكراري المتجمع الصاعد .

ت) كون الجدول التكراري المتجمع الهابط .

ث) ماهي نسبة القطع التي مساحتها اقل من 40 ؟

ج) ماهي نسبة القطع التي مساحتها 32 فاكثر ؟

✓ طريقة الحل :

أ) تكوين الجدول التكراري المناسب : بيانات كمية متصلة :

1) الخطوة الأولى: تحديد عدد الفئات : ويمكننا إتباع قاعدة Sturge's Rule وهنا نضرب عدد الاعمدة في الصفوف $10 \times 5 = 50$ ، فإذا النتيجة تقع وفق جدول Sturge بين 32 و 64 :

64 – 32	6 – 5

الان نختار اما 5 او 6 وسوف نختار 6 لان السؤال طلب ذلك

2) الخطوة الثانية : تحديد طول الفئة بالقانون طول الفئة = المدى ÷ عدد الفئات

المدى = اعلى قيمة – اقل قيمة = $56 - 3 = 53$

اذا طول الفئة = $53 \div 6 \cong 8.833 = 9$

3) الخطوة الثالثة : تعيين حدود الفئات: نبدأ بتعيين الحد الأدنى للفئة الأولى وهو قيمة اصغر مفرده في البيانات الأولية للظاهرة محل الدراسة + طول الفئة

الفئة الاول	3 – 12
الفئة الثانية	12 –21
الفئة الثالثة	21 – 30
الفئة الرابع	30 – 39
الفئة الخامسة	39 – 48
الفئة السادسة	48 –57

4) الخطوة الرابعة : تفريغ البيانات في الجدول الاحصائي :

فئات مساحات الاراضي	التفريغ (العد بالشرطات – كل خمس شرطات تكون ما يسمى بالحزمة)	التكرارات (بالارقام)	التكرار النسبي التكرارات / اجمالي التكرارات
3–	IIII	4	0.08
12–	~~IIII~~	5	0.10
21–	III ~~IIII~~	8	0.16
30–	III ~~IIII~~ ~~IIII~~	13	0.26
39–	III ~~IIII~~ ~~IIII~~	13	0.26
48–	II ~~IIII~~	7	0.14
المجموع		50	1

ب) الجدول التكراري المتجمع الصاعد : لاعداد الجدول التكراري المتجمع الصاعد يعد جدولا من خانتين : الاولى للحدود العليا للفئات ويستخدم معها كلمة " اقل من " والثانية يتم تخصصيها لتجميع التكرارات حسب ترتيب ورودها

الحدود العليا للفئات	التكرار المتجمع الصاعد
اقل من 3	صفر
اقل من 12	4
اقل من 21	9
اقل من 30	17
اقل من 39	30
اقل من 48	43
اقل من 57	50

ت) الجدول التكراري المتجمع الهابط : ويعطى الجدول المتجمع الهابط (النازل) الحدود الدنيا للفئات وعدد المفردات

الحدود العليا للفئات	التكرار المتجمع الهابط
3 فاكثر	50
12 فاكثر	46
21 فاكثر	41
30 فاكثر	33
39 فاكثر	20
48 فاكثر	7
57 فاكثر	صفر

ث) ماهي نسبة القطع التي مساحتها اقل من 40 ؟

60 ٪ (اجمالي نسب اول 4 بنود في الجدول الاحصائي في فقرة أ الجزء d) ونستخرجها من الجدول الصاعد .

ج) ماهي نسبة القطع التي مساحتها 32 فاكثر ؟ 40 ٪ وسنتخرجها من الجدول الهابط .

4- في احد الدراسات عن هل يوجد علاقة بين النوع والتدخين فتم اخذ عينة من 15 شخص تم سؤالهم عن نوعهم وهل الشخص يدخن ام لا ؟ فكانت اجابتهم كما يلي :

رقم الشخص	النوع	التدخين
1	ذكر	لا
2	انثى	لا
3	ذكر	نعم
4	ذكر	نعم
5	انثى	نعم
6	ذكر	لا
7	انثى	نعم
8	ذكر	لا
9	ذكر	نعم
10	ذكر	نعم
11	ذكر	نعم
12	انثى	لا
13	انثى	نعم
14	ذكر	نعم
15	انثى	لا

◄ المطلوب : اعداد الجدول التكراري المناسب لعرض البيانات السابقة ؟

✓ طريقة الحل : جدول تكراري مزدوج

اولا: يتم تفريغ البيانات :

النوع / التدخين	يدخن	لايدخن
ذكر	~~IIII~~ I	III
انثى	III	III

ثانيا : حساب التكرارات والتوزيع الهامشي لكل من المتغيرين :

النوع / التدخين	يدخن	لا يدخن	المجموع
ذكر	6	3	9
انثى	3	3	6
المجموع	9	6	15

(المحاضرة السادسة)

العرض البياني للبيانات (الجزء الأول)

❖ **تعريف الرسوم البيانية:** هي وسيلة مفيدة وفعالة لتوضيح وشرح الحقائق الرقمية وابراز العلاقة بين المتغيرات، واستقراء اتجاهاتها العامة بأسلوب يسهل فهمه وتذكره بمجرد النظر .

وتنطبق القواعد التي ذكرناها في العرض الجدولي على الرسوم البيانية، اذ يجب أن يرقم كل رسم ، ويعنون، ويمكن أن يستعمل الحواشي والمصدر وغيرها ..

البيانات غير المبوبة تكون فيها كل قيمة بشكل منفرد لوحدها وهي البيانات الخام قبل وضعها في صيغة جداول تكرارية .

❖ **اولا : البيانات غير المبوبة :**

والمقصود بالبيانات غير المبوبة هي البيانات الاسمية او الرتبية او الكمية المتقطعة (المنفصلة) (هي موضوع الجزء الاول من المحاضرة السادسة)

البيانات المبوبة تكون على شكل فئات (كل فئة تضم قسم أو مجموعة من البيانات) والفئات يجب تكون شاملة (بمعنى كل قيمة لازم تدخل في فئة) ولا يكون هناك تداخل بين الفئات (بمعنى لا يمكن لقيمة تسجل لفئتين أو نحتار هي تتبع لأي فئة) ، ولمزيد من التوضيح فالبيانات المبوبة هي تلك البيانات التي تم وضعها في صورة جداول تكرارية .

❖ **ثانيا : البيانات المبوبة :**

والمقصود بالبيانات المبوبة هي البيانات الكمية المتصلة (سوف نتطرق لها في الجزء الثاني من المحاضرة السادسة)

◄ اولا : البيانات غير المبوبة :

تختلف الرسوم البيانية حسب طبيعة ونوع البيانات المراد عرضها فاذا كانت البيانات اسمية أو رتبية (أي منفصلة) فإننا نستخدم أحد الأشكال البيانية التالية:

أ. الاعمدة البيانية البسيطة : وهي عبارة عن مجموعة من الأعمدة الرأسية أو المستطيلات المتساوية القاعدة والتي تتناسب ارتفاعاتها مع البيانات التي تمثلها ، وتستخدم لاظهار التطور الذي يطرأ على ظاهرة ما على مدار عدة سنوات ، وعادة ما يؤخذ المحور الرأسي لتمثيل قيم الظاهرة .

⇐ مثال :

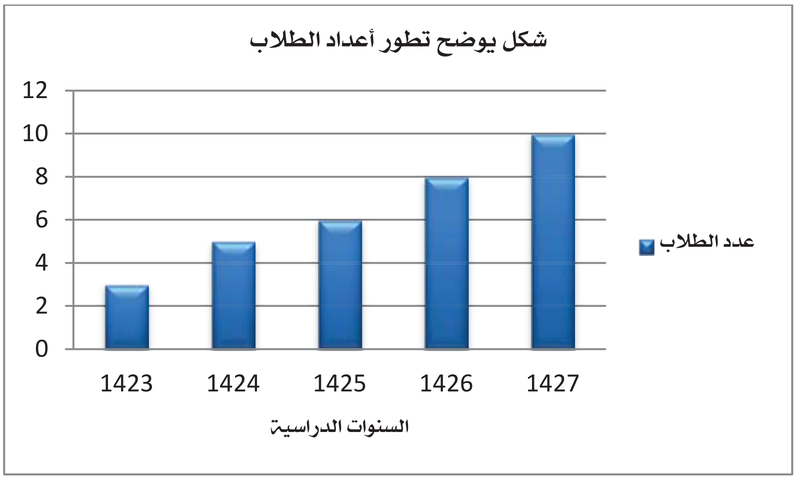
الجدول الأتي يوضح أعداد الطلاب المقيدين باحد الجامعات في السنوات الدراسية من 1423هـ حتي 1427هـ .

السنة الدراسية	1423	1424	1425	1426	1427
عدد الطلاب بالالاف	3	5	6	8	10

◄ المطلوب: تمثيل البيانات باستخدام الرسم البياني المناسب ؟

✓ طريقة الحل :

يمكن تمثيل هذه الظاهر بيانيا باستخدام الاعمدة البيانية البسيطة لان الظاهر التي بين ايدينا توضح مدى التطور في اعداد الطلاب مع تقدم السنوات مستخدمة متغير واحد فقط .



ب. الأعمدة البيانية المزدوجة: وهو ذلك النوع من الرسوم البيانية الذي يستخدم اذا كان الهدف من الرسم هو مقارنة ظاهرتين أو اكثر لعدة سنوات، أو اذا كان لدينا بيانات مزدوجة لخواص مختلفة .

⇐ مثال :

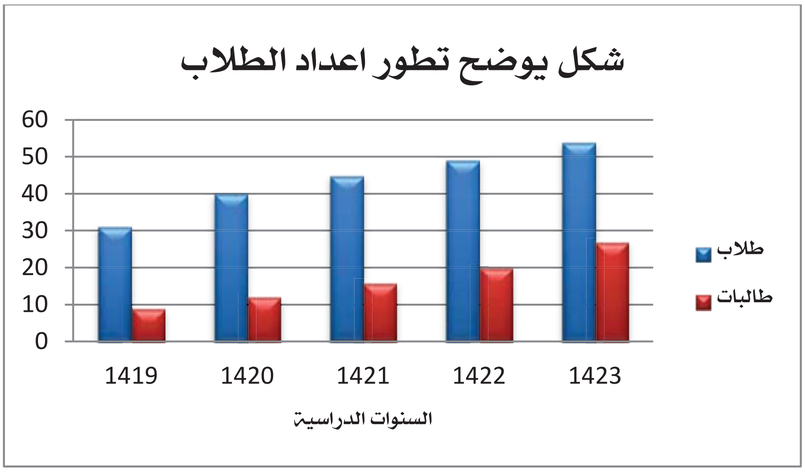
الجدول الأتي يوضح أعداد الطلبة المسجلين باحد الجامعات السعودية في السنوات الدراسية 1419هـ حتى 1423هـ

السنة الدراسية		1419	1420	1421	1422	1423
عدد الطلبة بالألف	طلاب	31	40	45	49	54
	طالبات	9	12	16	20	27

◄ المطلوب: مثل هذه البيانات بيانيا باستخدام الأعمدة البيانية المزدوجة ؟

✓ طريقة الحل :

يمكن تمثيل هذه الظاهر بيانيا باستخدام الاعمدة البيانية المزدوجة لان الظاهر التي بين ايدينا توضح مدى التطور في اعداد الطلاب مع تقدم السنوات مستخدمة متغيرين (ذكور و اناث) .



ج. الأعمدة البيانية المجزأة : يستخدم ها النوع من الرسوم البيانية في تمثيل نفس الحالات التي تستخدم فيها الاعمدة المزدوجة ، ويتم رسم هذا النوع من الاعمدة كالتالي :

- نقوم برسم عمود واحد يمثل جملة الظواهر محل الدراسة في كل سنة كما في حالة الاعمدة البيانية البسيطة .
- نقسم كل عمود الى مكوناته بحيث يتناسب كل جزء مع العدد الذي يمثله .
- نميز بين هذه الاجزاء بالتظليل او بالالوان المختلفة ، ونوضح ذلك على الرسم.

⇐ مثال :

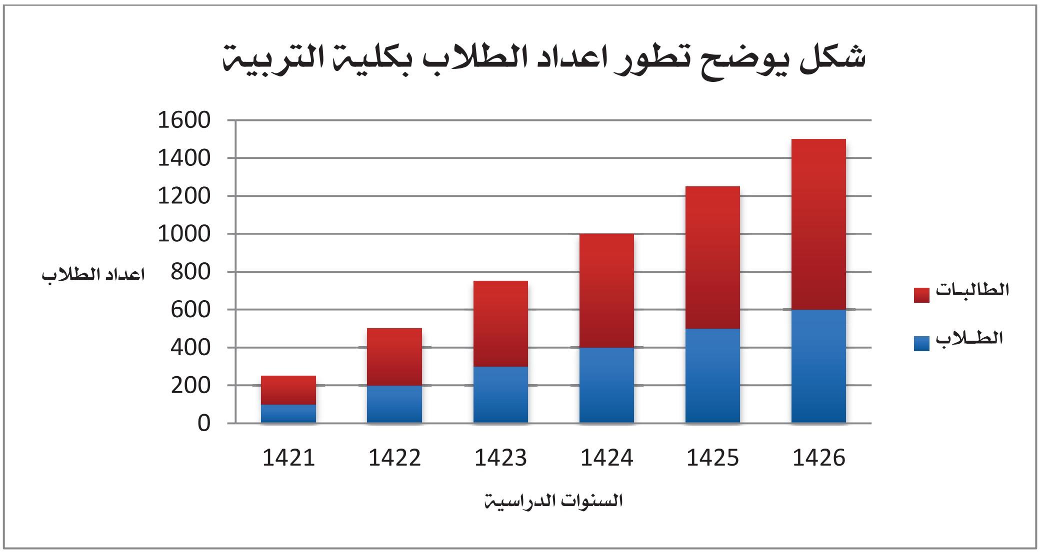
اذا كانت اعداد الطلاب والطالبات المسجلين في كلية التربية بجامعة الملك فيصل بالاحساء تزداد كما هو موضح في الجدول الأتي:

السنوات الدراسية	1421	1422	1423	1424	1425	1426
الطلاب	100	200	300	400	500	600
الطالبات	150	300	450	600	750	900

◄ المطلوب: مثل هذه البيانات بيانيا باستخدام الأعمدة المجزأة؟

✓ طريقة الحل :

يمكن تمثيل هذه الظاهر بيانيا باستخدام الاعمدة البيانية المزدوجة لان الظاهر التي بين ايدينا توضح مدى التطور في اعداد الطلاب مع تقدم السنوات مستخدمة متغيرين (ذكور و اناث) .



❖ ملاحظات على استخدام الاعمدة البيانية (بأنواعها المختلفة): يمكن ابداء الملاحظات التالية على الرسومات بالاعمدة البيانية بأنواعها المختلفة :

- تعتبر الاعمدة البيانية من اكثر الرسومات البيانية انتشارا.
- يفضل تظليل الاعمدة أو تخطيطها بواسطة خطوط متوازية أو ابرازها بألوان مختلفة وخاصة عند مقارنة ظواهر مختلفة.
- يستحسن اختيار مقياس رسم مناسب وثابت. ولهذا يتوجب على مصمم الرسم التعرف على القيمة الكبرى والصغرى لتحديد مقياس الرسم المناسب ، كما يجب البدء بالصفر على المحور الرأسي الذي يدل على القيمة الرقمية حتى تكون المقارنة سهلة وسليمة وغير مضللة .

- يفضل عدم كتابة القيم التي تمثلها الاعمدة فوق الاعمدة وذلك لتلافي المبالغة في طول الاعمدة، ولتجنب جعل الرسم مكتظا مما ينفر القارئ .
- يمكن استخدام العمود الواحد لتمثيل اكثر من نوع واحد من البيانات، وذلك باستخدام مفهوم الاعمدة المجزأة، ويفضل عدم عرض اكثر من ثلاثة ظواهر في العمود الواحد حتى لايفقد الرسم الهدف الاساسي منه.
- تصلح الاعمدة البيانية لتمثيل البيانات ذات المتغيرات المنفصلة، كما تصلح بشكل خاص لتميل البيانات الوصفية (النوعية) ، (اي غير الرقمية) وذلك كما في تمثيل الحالة الاجتماعية (متزوج ، مطلق ، ارمل) .

د. اللوحة الدائرية: تستخدم الدائرة أو اللوحة الدائرية لثميل البيانات في الحالات التالية:

1. عندما يكون الهدف منها مقارنة الاجزاء المختلفة بالنسبة للمجموع الكلي
2. أن تكون الاجزاء المقارنة قليلة العدد نسبيا وفي فترة زمنية واحدة.

- وفيما يلي خطوات رسم الدائرة وتقسيمها الى قطاعات:

1) اختيار نصف قطر مناسب لها.
2) تحسب الزاوية المقابلة لكل قطاع من خلال العلاقة التالية:

$$\text{زاوية القطاع} = \frac{\text{قيمة القطاع}}{\text{المجموع العام}} \times \text{الزاوية المركزية الدائرة } (360)$$

3) تقسم الدائرة الى قطاعاتها المختلفة بتحديد مساحة كل قطاع على الدائرة وذلك بتقسم الزاوية المركزية للدائرة الى زوايا القطاعات المختلفة.

⇐ مثال :

فيما يلي احصائية لطلاب البكالوريوس في كلية العلوم الإدارية موزعين حسب السنة الدراسية للعام الجامعي 1426هـ .:

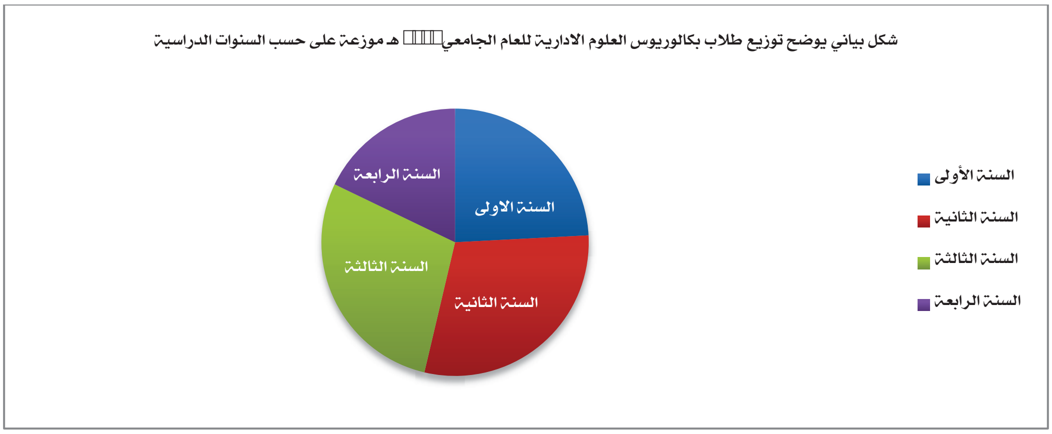
السنة الدراسية	عدد الطلبة
السنة الاولى	226
السنة الثانية	276
السنة الثالثة	266
السنة الرابعة	167
المجموع	935

◄ المطلوب: مثل هذه البيانات بيانيا باستخدام الأعمدة المجزأة؟

✓ طريقة الحل :

لمعرفة زاوية كل قطاع(سنه) في المثال نطبق القانون مثلا عهلى السنه الاولى : $226 \div 935 = 0.2417$

اذا : $0.2417 \times 360 = 87$ درجة زاوية السنة الاولى ، (وهكذا على جميع السنوات حتى يتجمع لنا زوايه اجماليها 360 درجة والتي تمثل محيط الدائرة)



هذا ويستحن تظليل القطاعات الدائرة او تلوينها وذلك زيادة في قيمة الرسم البياني وبالتالي زيادة جاذبيته ووضوحه ، وكذلك ينصح كتابة الجزء (السنة) داخل كل قطاع دائري .

وعند الحاجة الى مقارنة بين مجموعتين او اكثر باستخدام اللوحة الدائرية فاننا نرسم عددا من الدوائر يتناسب مع عدد البيانات المطلوب مقارنتها ، ونتبع فيها نفس الخطوات السابقة لرسم اللوحة الدائرية .

س: متى نستخدم الأعمدة البيانية (بأنواعها المختلفة) في تمثيل البيانات الاحصائية بيانيا؟ وبماذا تختلف عن التمثيل البياني باستخدام الدائرة؟

يرى غالبية المختصين أن الاعمدة البيانية يفضل استخدامها في الحالات التالية:

1. عندما تكون الكميات المقارنة كثيرة العدد نسبيا ، حيث يصعب تمثيلها بالدائرة وذلك ان كثرة الكميات المقارنة تجعل الدائرة مكتظة لدرجة يصعب مقارنة التوزيع النسبي للظاهرة المدروسة .
2. عندما تكون الاجزاء المقارنة في فترات زمنية مختلفة ، وهذا لايمنع من استعمالها في فترة زمنية واحدة ، الا ان الدائرة لايمكن استخدامها لمقارنة الاجزاء بالكل في فترات زمنية مختلفة .
3. عندما نرغب في توضيح قيم الاجزاء المقارنة المختلفة للظاهرة موضع البحث وذلك من اجل ابراز المقارنة بين هذه الاجزاء او توضيح التغير او التطور عبر الزمن سواء لظاهرة واحدة او عدة ظواهر بين فترات زمنية مختلفة .
4. غالباً ما ينصح باستعمال الاعمدة البيانية (بانواعها المختلفة) مع المتغيرات المنفصلة (وهي التي تأخذ قيما او اعداد صحيحة) كما في عدد الطلبة او افراد الاسرة او عدد الكتب في المكتبة ... الخ .

٥. المنحنى أو الخط البياني: يستخدم المنحنى أو الخط البياني أساسا لتوضيح الاتجاه العام للظاهرة خلال فترة من الزمن، ويستخدم هذا النوع من الرسم البياني لتمثيل الظواهر ذات البيانات المتصلة (غالبا)، وكذلك ممكن استخدامه مع البيانات المنفصلة.

ويتم رسم المنحنى او الخط البياني باتباع الاتي :

- نرسم محورين افقي ورأسي بحيث يمثل المحور الافقي الزمن مثلا والمحور الرأسي قيم الظاهرة .
- نستخدم نفس المبدأ الذي اتبعناه في رسم الاعمدة البيانية المختلفة اللهم بدلا من رسم الاعمدة ذاتها نستعيض عنها بتعيين نقطة (احداثية النقطة) فقط لكل منها .
- توصيل هذه النقط ببعضها بمعنى ممهد متصل فنحصل على خط متصل يسمى المنحنى ، او القيام بتوصيل كل نقطتين متجاورتين بخط مستقيم فنحصل عندئذ على الخط البياني .

⇐ **مثال :**

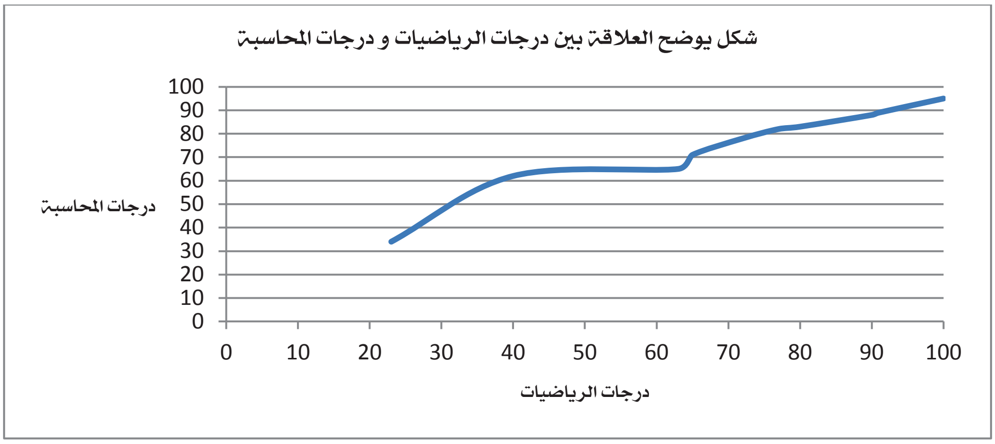
البيانات التالية لدرجات عشر طلاب بكلية العلوم الإدارية فى مقررى الرياضيات والمحاسبة، فكانت كما يلي:

رقم الطالب	1	2	3	4	5	6	7	8	9	10
درجات الرياضيات	23	40	63	65	72	77	80	90	91	100
درجات المحاسبة	34	62	65	71	78	82	83	88	89	95

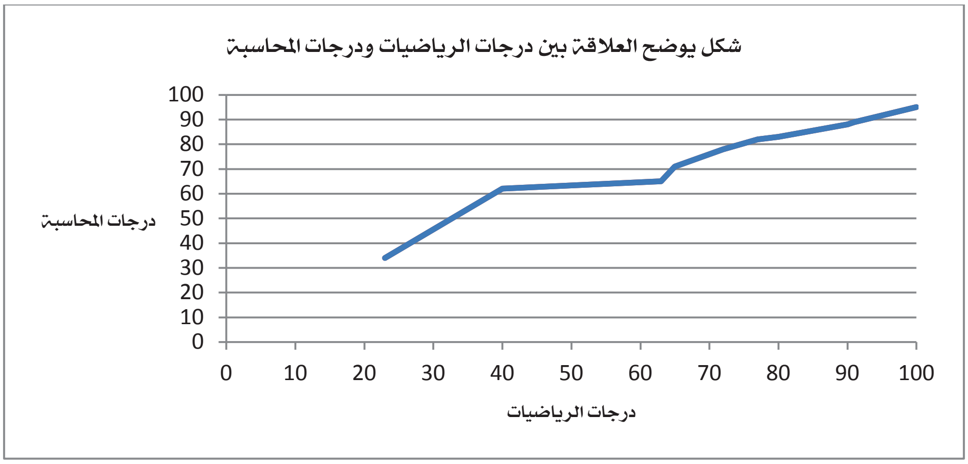
◄ **المطلوب:** استخدام المنحنى او الخط البياني لتمثيل هذه البيانات (درجات مقرر الرياضيات ودرجات مقرر المحاسبة) ؟

✓ **طريقة الحل :**

اولا :باستخدام المنحنى



ثانيا :باستخدام الخط البياني



❖ **ملاحظات على المنحنى والخط البياني :**

- الرسم بالخط البياني أو المنحنى يتطلب جهدا أقل من الجهد والوقت اللذين يتطلبهما رسم الأعمدة البيانية بأنواعها المختلفة.
- يسهل الخط البياني أو المنحنى المقارنة على القارئ. من بداء ان العين تدرك الاشياء المتصلة اكثر من الاشياء المنفصلة .
- يمكن استخدام الخط البياني أو المنحنى (كما في الأعمدة البيانية) لتمثيل أكثر من ظاهرة على نفس الرسم ومقارنتها ببعضها، مع ملاحظة تمييز الخط البياني لكل ظاهرة بخطوط متصلة او منفصلة او اعطائها الوانًا مختلفة وتوضيح ذلك في مفتاح الرسم .

❖ مزايا وعيوب الرسوم البيانية :

المزايا	العيوب
▪ تثير انتباه المشاهد خاصة اذا كانت جيدة التصميم. ▪ توفر وقت المشاهد اذ أن استنباط الحقائق من الرسوم البيانية أسرع من الوصول اليها بواسطة الأرقام الموضوعة في جداول. ▪ إمكانية معرفة الاتجاهات العامة للظواهر. ▪ سهولة فهم وتذكر العلاقات بين الظواهر محل الدراسة.	▪ التضحية بدقة البيانات اذ أن الرسوم توضح فقط التغيرات العامة للظواهر ولا تبين التفاصيل الدقيقة لها. ▪ أحيانا تكون الرسوم معقدة، خاصة إذا كانت تشتمل على مجموعات من البيانات المتباينة. ▪ كثرة التكاليف خاصة إذا كانت البيانات تحتاج الى مقياس رسم كبير.

(المحاضرة السادسة)

العرض البياني للبيانات (الجزء الثاني)

ثانيا : البيانات المبوبة :

يتم استخدام العديد من الاشكال للتعبير عن البيانات المبوبة فى صورة جداول توزيعات تكرارية (بما اننا تكلمنا عن التوزيعات التكرارية فهذا معناه ان المتغيرات التي نتكلم عنها هنا هي متغيرات كمية متصلة) وهى :

- المدرج التكرارى
- المضلع التكرارى
- المنحنى التكرارى
- المنحنى التكرارى المتجمع الصاعد
- المنحنى التكرارى المتجمع الهابط (النازل)

أ. **المدرج التكرارى:** المدرج التكرارى هو عبارة عن أعمدة مستطيلة متلاصقة يعبر ارتفاع العمود فيها على التكرار المناظر للفئة. ويستخدم هذا النوع من الرسوم البيانية لتمثيل البيانات التى تم عرضها في جدول توزيع تكرارى، وفيه يمثل كل مستطيل فئه من فئات التوزيع التكرارى.

ويتم تقسيم المحور الرأسي (المحور الصادي) فى المدرج التكرارى حسب التكرار (فقد نستخدم التكرار الأصلى في حالة تمثيل التوزيع التكرارى، وكذلك يمكن أن نستخدم التكرار النسبى في حالة تمثيل التوزيع التكرارى النسبى). أما المحور الأفقى (المحور السيني) على أساس الفئات وهنا يظهر حالتين هما:

1) الحاله الأولى:- تساوى أطول الفئات

وفي هذه الحالة يكون ارتفاع المستطيل معبرا عن عدد مرات تكرار وجهه الظاهرة محل الدراسة حيث انه يتناسب مع مساحة المستطيل ، وذلك لان طول الفئة هو عرض المستطيل ، وحيث ان اطوال الفئات متساوية فإن مساحة المستطيل تتناسب مع طوله فقط .

2) الحاله الثانيه:- عدم تساوى أطوال الفئات

وفي هذه الحالة لابد من اجراء تعديل في التكرار الاصلي قبل رسم المدرج التكراري ، لذا فاننا نقوم بايجاد التكرار المعدل والذي هو عبارة عن ناتج قسمة التكرار الاصلي لكل فئة على طول الفئة المقابلة ، وهنا تكون مساحة المستطيل معبرة عن وجه الظاهرة المقابلة لها ، وليس ارتفاع المستطيل .

❖ خطوات رسم المدرج التكراري :

❖ نرسم محورين افقي وراسي بحيث يمثل المحور الافقي الفئات والمحور الرأسي التكرارات .

❖ نمثل بيانات الدراسة من خلال مجموعة من المستطيلات المتلاصقة بحيث يعبر ارتفاع المستطيل عن عدد مرات تكرار وجه الظاهرة محل الدراسة .

⇐ مثال :

فيما يلي بيان بتوزيع لعينة من 40 عامل على اساس فئات العمر للعمال:

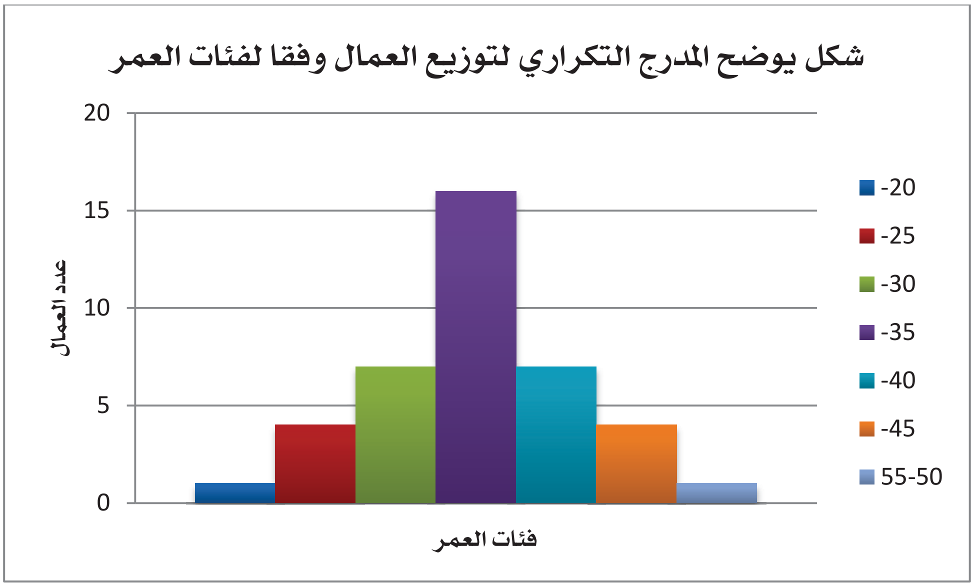
فئات العمر	20-	25-	30-	35-	40-	45-	50-55	المجموع
عدد العمال	1	4	7	16	7	4	1	40

◄ المطلوب: عرض البيانات السابقة في شكل المدرج التكراري ؟

✓ طريقة الحل :

حالة فئات العمر المتساوية :

1. يتم رسم المدرج التكراري على اساس التكرار الاصلي (فئات عمر العمال) .
2. نرسم محورين افقي ورأسي بحيث يمثل المحور الافقي فئات العمر والمحور الرأسي تكرار عدد العمال في كل فئة عمرية .
3. نمثل بيانات الدراسة من خلال مجموعة من المستطيلات المتلاصقة بحيث يعبر ارتفاع المستطيل عن عدد العمال في كل فئة .

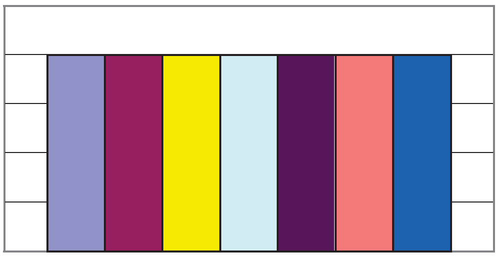


❖ بعض خصائص التوزيع التكراري :- يمكن استنتاج بعض خصائص التوزيع التكراري من شكل المدرج التكراري بدراسة الخصائص التالية :

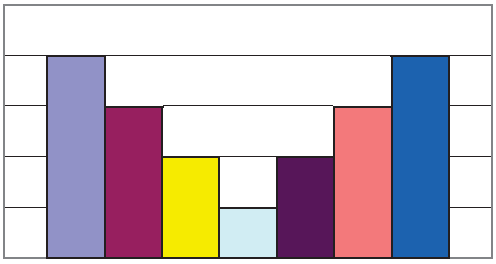
1) الخاصية الاولى : التماثل

يسمى المدرج التكراري متماثلا عندما نقوم برسم خط مستقيم في منتصف المدرج التكراري فيظهر لنا التطابق التام بين الجانبين هو الخط المستقيم ، وذلك يظهر في الرسم السابق مباشرة حيث يكون الجانب الايمن كخيال للجانب الايسر في المرآه ، وكذلك قد يكون شكل المدرج التكراري متماثل كما هو واضح في الشكلين التاليين :

شكل يوضح التوزيع المتماثل



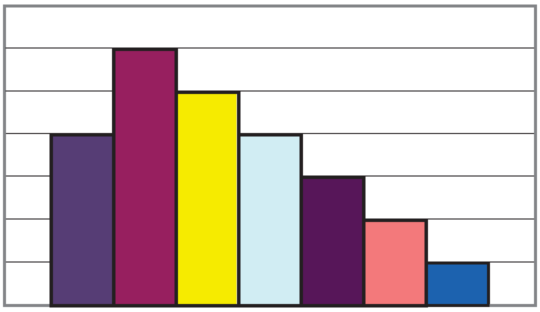
شكل يوضح التوزيع المتماثل



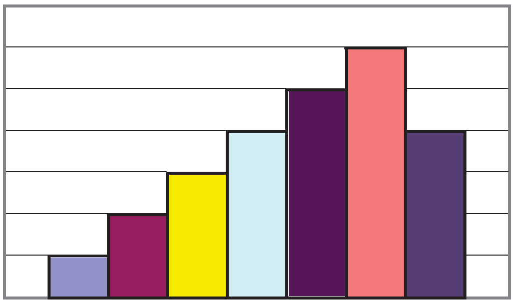
2) الخاصية الثانية :- الالتواء

وعندما يكون ذيل التوزيع جهة اليسار – بمعنى ان الطرف الايسر للتوزيع اطول من طرفه الايمن – يكون الالتواء باتجاه اليسار ويسمى توزيع سالب الالتواء ، فمثلا توزيع الوقت اللازم لاجابة الامتحان بالنسبة لعدد الطلاب يكون في الغالب سالب الالتواء ويرجع ذلك لقيام عدد قليل من الطلاب بتسليم اوراق الاجابة قبل موعد انتهاء الامتحان ، وفي المقابل يفضل الكثير من الطلاب تسليم اوراق الاجابة مع نهاية وقت الامتحان وفيما يلي توضيح الالتواء بنوعيه في الشكلين التاليين :

شكل يوضح الالتواء الموجب



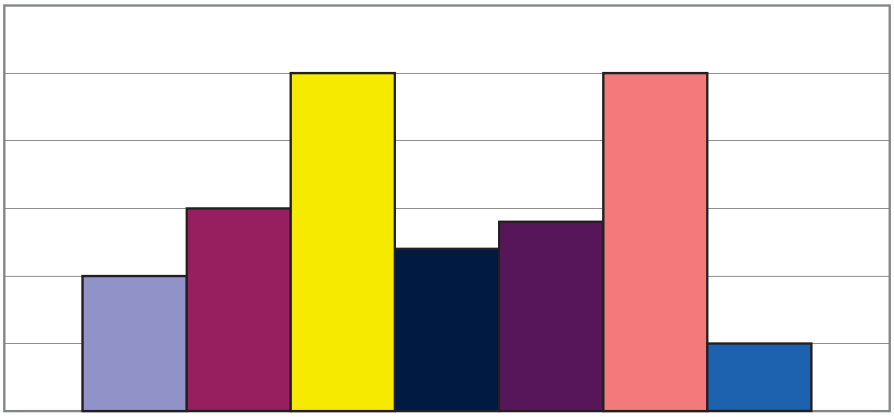
شكل يوضح الالتواء السالب



3) الخاصية الثالثة : المنوال

المنوال هو القيمة الاكثر تكرراً (شيوعاً) في الظاهرة محل الدراسة ، وفي بعض الاحيان يكون المدرج التكراري أحادي المنوال عندما تقع معظم البيانات داخل فئة في منتصف التوزيع التكراري وتسمى الفئة المنواليه وهي تمثل قمة واحده للتوزيع ، مع وجود بعض البيانات قبل وبعد هذه الفئة ، وفي احيان اخرى يكون المدرج التكراري ثنائي المنوال ، وذلك في حالة وجود قيمتين في التوزيع ويشترط تساوي القمتين معًا ، فمثلا اذا نظرنا الى التوزيع التكراري للدخول في احد البلدان التي يعيش فيها كثير من الاغنياء وكثير من الفقراء وقلة من الطبقة المتوسطة ، فإن شكل المدرج التكراري لسكان هذا البلد يكون ثنائي المنوال كما في الشكل التالي :

شكل يوضح توزيع ثنائي المنوال



ب. **المضلع التكراري :** التكراري هو المضلع مغلق نحصل عليه من خلال حساب مراكز الفئات او بتصنيف الاضلاع العلوية للمستطيلات في المدرج التكراري ، ثم نوصل هذه النقاط بعضها بع بعض ، ولكي نغلق الخط المنكسر الذي حصلنا عليه نعتبر ان هناك فئتين متطرفتين واحدة في اقصى اليمين والثانية في اقصى اليسار وتكرار كل منهما صفر ، نأخذ مركز كل من هاتين الفئتين ، ونغلق المضلع كما يبدوا لنا في المثال التالي :

❖ خطوات رسم المضلع التكراري :

❖ نرسم محورين افقي وراسي بحيث يمثل المحور الافقي الفئات والمحور الرأسي التكرارت .

❖ لكي نرسم المضلع من خلال المدرج التكراري نقوم بتمثيل بيانات الدراسة من خلال مجموعة من المستطيلات المتلاصقة بحيث يعبر ارتفاع المستطيل عن عدد مرات تكرار وجه الظاهرة محل الدراسة .

❖ نقوم بتقسيم هذه المستطيلات من اعلى (مركز الفئة) ، ثم بعد ذلك نوصل نقاط التقسيم هذه بعضها مع بعض بخط مستقيم من خلال المسطرة لنحصل بالتالي على المضلع التكراري من خلال المدرج التكراري .

$$\text{مركز الفئة} = \frac{\text{الحد الأدنى للفئة} + \text{الحد الأعلى لنفس الفئة}}{2}$$

❖ ولرسم المضلع من خلال مراكز الفئات نقوم بايجاد مركز الفئه لجميع فئات التوزيع التكراري ، ثم نقوم بتمثيل التكرار الاصلي المقابل لكل فئه بنقطه تناظر مركز هذه الفئه .

❖ نقوم برسم خط بإستخدام المسطره يصل كل نقطتين متتاليتين ، فنحصل على المضلع التكراري .

❖ لاغلاق المضلع من الطرفين نقوم بإنشاء فئه سابقه عند النقطه الأولى في التوزيع التكراري يقابلها تكرار أصلي يساوي الصفر ، وكذلك إنشاء فئه لاحقه للفئه الأخيرة في التوزيع التكراري يقابلها تكرار أصلي يساوي الصفر أيضًا ، ونحسب مركز الفئة لكل منهما .

⇐ مثال :

فيما يلي بيان بتوزيع لعينة من 40 عامل على اساس فئات العمر للعمال:

فئات العمر	20-	25-	30-	35-	40-	45-	55-50	المجموع
عدد العمال	1	4	7	16	7	4	1	40

◄ المطلوب: عرض البيانات السابقة في شكل المضلع التكراري ؟

✓ طريقة الحل :

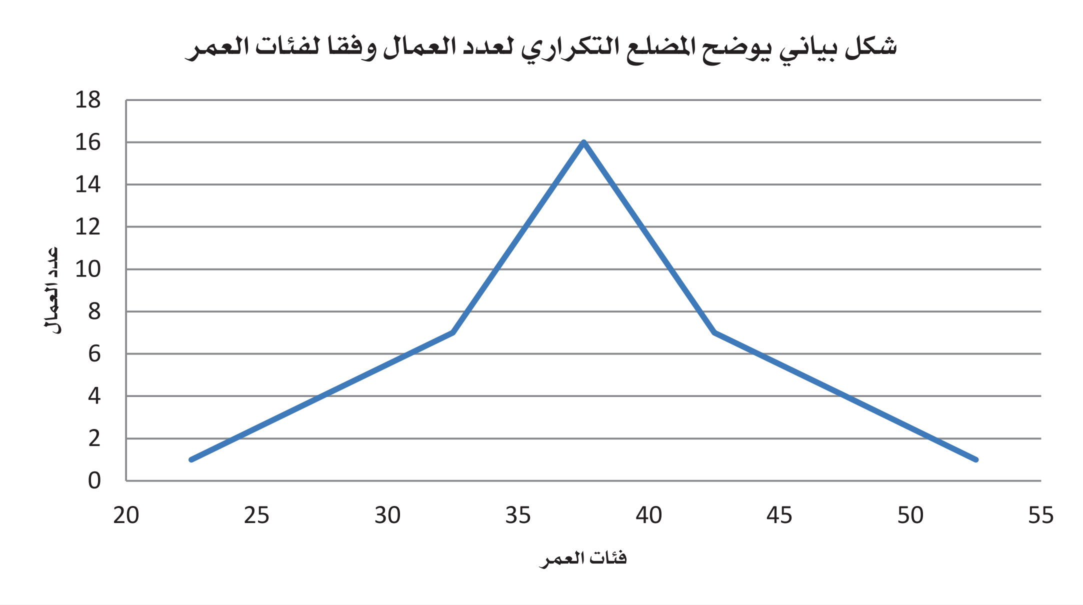
(1) نحصل على مراكز الفئات حسب القانون الموضح أعلاه في خطوات رسم الشكل :

فئات العمر	20-	25-	30-	35-	40-	45-	55-50	المجموع
مركز الفئات	22.5	27.5	32.5	37.5	42.5	47.5	52.5	
عدد العمال	1	4	7	16	7	4	1	40

(2) استحداث فئتين سابقة ولاحقة للتوزيع وحساب مركز الفئة لكل منهما ، وذلك باعتبار الفئة السابقة سوف تكون 15 – 20 و اللاحقة 55-60 وتكرارهما صفر ونقوم بحساب مراكزهما كالتالي :

فئات العمر	15-	20-	25-	30-	35-	40-	45-	55-50	60-55	المجموع
مركز الفئات	17.5	22.5	27.5	32.5	37.5	42.5	47.5	52.5	57.5	
عدد العمال	0	1	4	7	16	7	4	1	0	40

(3) نرسم محورين افقي وراسي بحيث يمثل المحور الافقي الفئات والمحور الرأسي التكرارات ، ثم بعد ذلك نرسم شكل المدرج التكراري ومن ثم نقوم بتقسيم مستطيلات المدرج التكراري من أعلى (مركز الفئة) (مركز الفئات وعدد العمال هذا ما سوف نستخدمه في الرسم وهذا هو سبب استخراج وتواجد مركز الفئات)، ثم بعد ذلك نوصل نقاط التقسيم هذه بعضها مع بعض بخط مستقيم من خلال المسطرة لنحصل بالتالي على المضلع التكراري من خلال المدرج التكراري .



(الجدول التالي هو الذي تم استخدامه في الرسم البياني في برنامج الاكسل)

فئات العمر (مركز الفئات)	22.5	27.5	32.5	37.5	42.5	47.5	52.5
عدد العمال	1	4	7	16	7	4	1

ج. **المنحنى التكراري:** اذا مهدنا المضلع التكراري وجعلناه منحنى بدلا من خطوط منكسرة فإننا نحصل على المنحنى التكراري ، ويلاحظ أنه ينبغي عدم رسم المنحنى التكراري إلا إذا كانت الفئات كثيرة العدد ، وذات طول صغير وكان عدد البيانات كبيراً وكانت هذه البيانات من النوع **المتصل مثل الزمن و الوزن** .

❖ **خطوات رسم المنحنى التكراري :**

❖ نرسم محورين افقي ورأسي بحيث يمثل المحور الافقي الفئات والمحور الرأسي التكرارات .
❖ نقوم بإنشاء فئة سابقة عند النقطة الاولى في التوزيع التكراري يقابلها تكرار اصلي يساوي صفر .
❖ نقوم بإنشاء فئة لاحقه للفئة الاخيرة في التوزيع التكراري يقابلها تكرار اصلي يساوي الصفر أيضًا .
❖ إيجاد مركز الفئة لجميع فئات التوزيع التكراري ، ثم نقوم بتمثيل التكرار الأصلي المقابل لكل فئة بنقطة تناظر مركز هذه الفئة .
❖ نقوم برسم خط باليد دون استخدام المسطره يصل كل نقطتين متتاليتين ، فنحصل على المنحنى التكراري .

⇐ **مثال :**

فيما يلي بيان بتوزيع لعينة من 40 عامل على اساس فئات العمر للعمال:

فئات العمر	-20	-25	-30	-35	-40	-45	55-50	المجموع
عدد العمال	1	4	7	16	7	4	1	40

◄ **المطلوب:** عرض البيانات السابقة في شكل المنحنى التكراري ؟

✓ **طريقة الحل :**

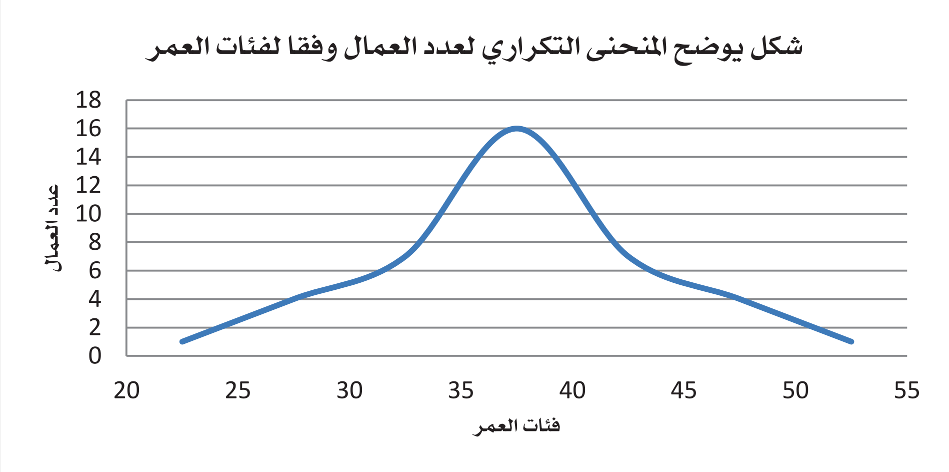
(1) نحصل على مراكز الفئات

فئات العمر	-20	-25	-30	-35	-40	-45	55-50	المجموع
مركز الفئات	22.5	27.5	32.5	37.5	42.5	47.5	52.5	
عدد العمال	1	4	7	16	7	4	1	40

(2) استحداث فئتين سابقة ولاحقة للتوزيع وحساب مركز الفئة لكل منهما ، وذلك باعتبار الفئة السابقة سوف تكون 15 – 20 و اللاحقة 55-60 وتكرارهما صفر ونقوم بحساب مراكزهما كالتالي :

فئات العمر	15-	20-	25-	30-	35-	40-	45-	50-55	55-60	المجموع
مركز الفئات	17.5	22.5	27.5	32.5	37.5	42.5	47.5	52.5	57.5	
عدد العمال	0	1	4	7	16	7	4	1	0	40

(3) نقوم برسم خط باليد دون استخدام المسطره يصل كل نقتطين متتاليتين مركز الفئات وعدد العمال هذا ما سوف نستخدمه في الرسم وهذا هو سبب استخراج وتواجد مركز الفئات ، فنحصل على المنحنى التكراري.



(الجدول التالي هو الذي تم استخدامه في الرسم البياني في برنامج الاكسل)

فئات العمر (مركز الفئات)	22.5	27.5	32.5	37.5	42.5	47.5	52.5
عدد العمال	1	4	7	16	7	4	1

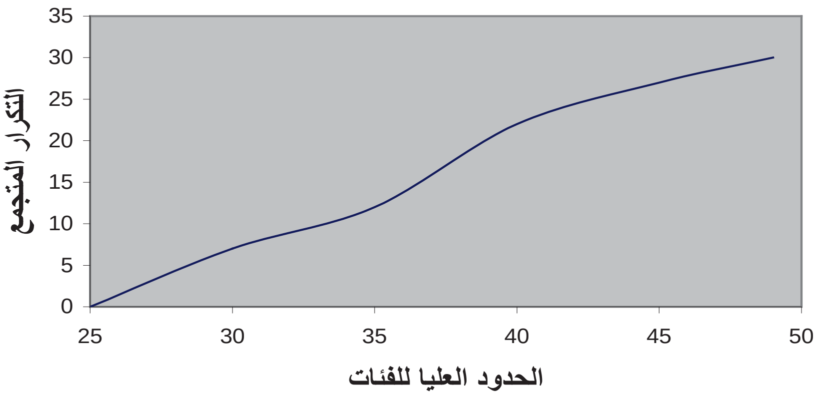
ملاحظة هامه : الدكتور في الشرح اثناء الرسم اليدوي قام باستخدم النقاط المضافة مثل 17.5 و 57.5 ذات القيم صفر في رسم المنحنى و المضلع التكراري وجعل الخط الموصل بينها متقطع حتى نميزها ولكن في الكتاب والرسمة التي اوضحها الدكتور ببرنامج الاكسل لم يتم استخدام هاتين النقطتين في الرسم البياني .

د. التوزيعات التكرارية المتجمعة : تستخدم المنحنيات المتجمعه لتمثيل التوزيعات التكرارية المتجمعة بيانيا بما يتلائم مع نوع التوزيع التكراري المتجمع ، ومحصل على المنحنى المتجمع برصد التكرار المتجمع لاي فئة مقابل الحد الاعلى او الحد الادنى الفعلي لها ثم نوصل هذه النقاط فيما بينها بخصوص ممهدة و نستطيع توزيع هذه المنحنيات الى نوعين كالتالي : (ملاحظة نفس فكرة الجداول التكرارية المجمعة)

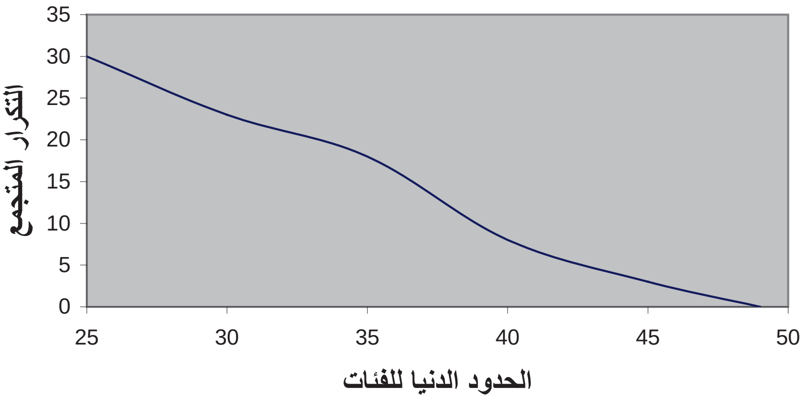
1) فيستخدم المنحنى المتجمع الصاعد لتمثيل التوزيع التكرارى المتجمع الصاعد، سواء أكان بالقيم المطلقه للتكرارات, أو بالتكرار النسبي. ويراعي وضع النقاط الخاصه بالتكرارات في حالة المنحنى المتجمع الصاعد عند الحد الأعلى لكل فئه، لأنه يعبر عن العدد الاجمالى لأوجه الظاهرة الواقع أسفل الحد الأعلى للفئه.

2) ويستخدم المنحنى المتجمع الهابط (النازل) لتمثيل التوزيع التكرارى المتجمع الهابط (النازل) أيضاً بالقيم المطلقه للتكرارات أو بالتكرار النسبى، ويراعى وضع النقاط الخاصه بالتكرارات المتجمعه الهابطه (النازلة) عند الحد الأدنى لكل فئه، لأنه يعبر عن العدد الاجمالى لأوجه الظاهرة الواقع أعلى الحد الأدنى للفئه.

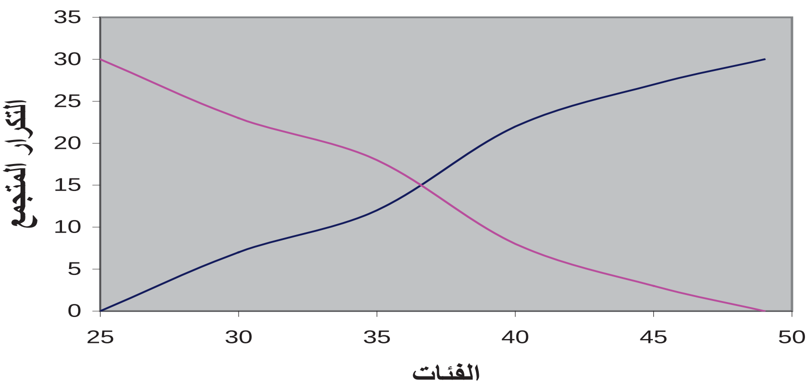
شكل يوضح المنحنى التكرارى المتجمع الصاعد



شكل يوضح المنحنى التكرارى المتجمع الهابط



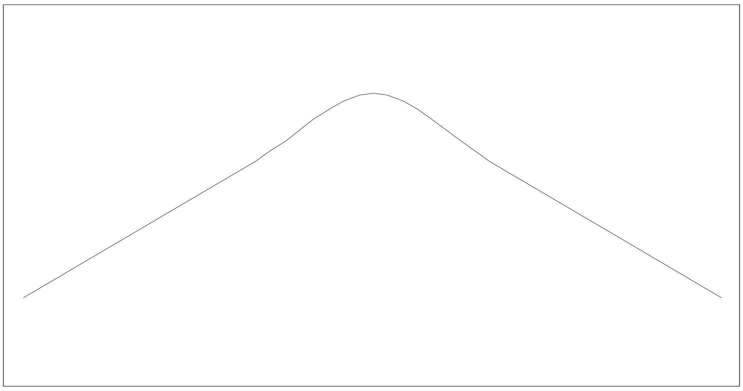
شكل يوضح كلاّ من المنحنى التكرارى المتجمع الصاعد و الهابط



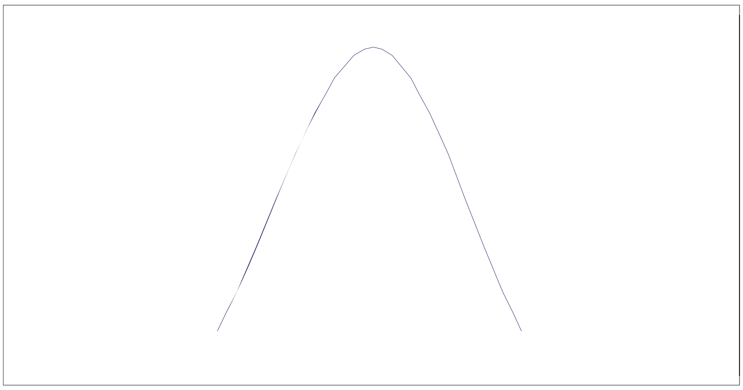
❖ الأشكال الشائعة للتوزيعات التكرارية :

يعتبر التوزيع الطبيعي ذو شكل الجرس من التوزيعات التكراريه الهامه في دراستنا. وفى أحيان أخرى يكون المنحنى التكرارى مدبدب القمة بحيث تكون القمه ضيقة وذو طرفين واسعين نسبيا، فيسمى فى هذه الحالة منحنى قليل التفرطح أو المنحنى المدبدب.

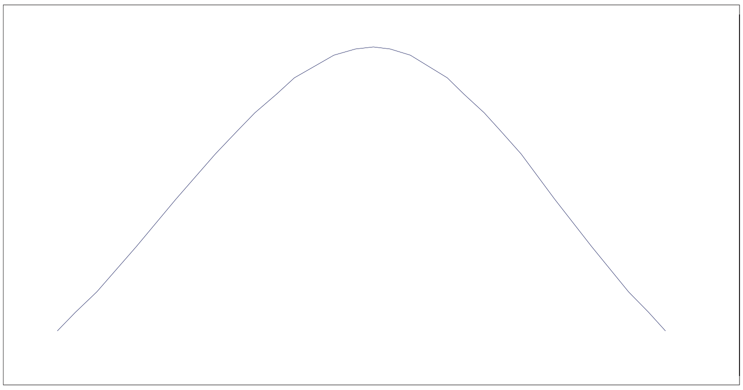
وقد يكون المنحنى التكرارى مسطح القمه بحيث تكون القمه واسعه ذو طرفين ضيقين نسبياً، فيسمى منحنى كبير التفرطح أو المنحنى المفرطح. وفيما يلي رسم بيانى يوضح كلا المنحنين المدبدب والمفرطح.



المنحنى المفرطح



المنحنى المدبب



المنحنى الطبيعي

تمارين الكتاب

1) البيانات التالية تعبر عن الجدول التكراري للحالة التعليمية لمجموعة من الاشخاص التي شملتهم أحد الدراسات :

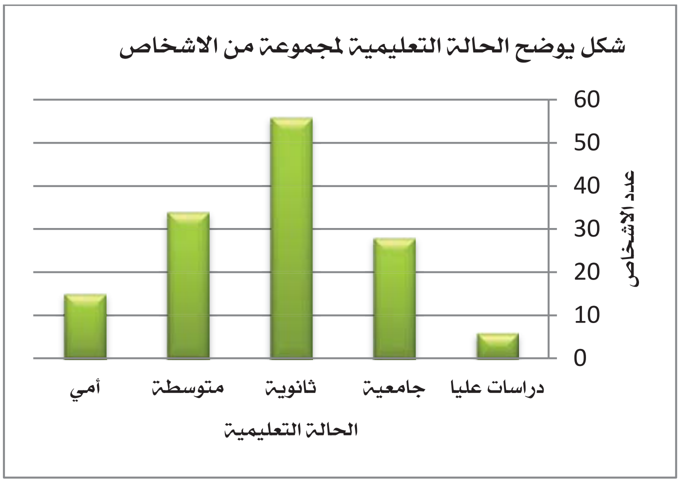
الحالة التعليمية	دراسات عليا	جامعية	ثانوية	متوسطة	أمي
عدد الاشخاص	6	28	56	34	15

◄ المطلوب : التمثيل البياني للجدول السابق من خلال

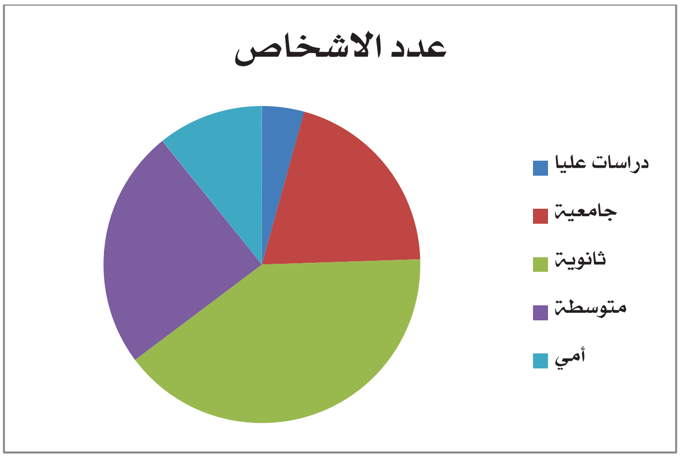
1. الأعمدة
2. اللوحة البيانية

✓ طريقة الحل :

1- الاعمدة :



2- اللوحة البيانية :



2) البيانات التالية تمثل توزيع عدد من الشركات وفقاً لأرباحها في العام الماضي بالمليون ريال :

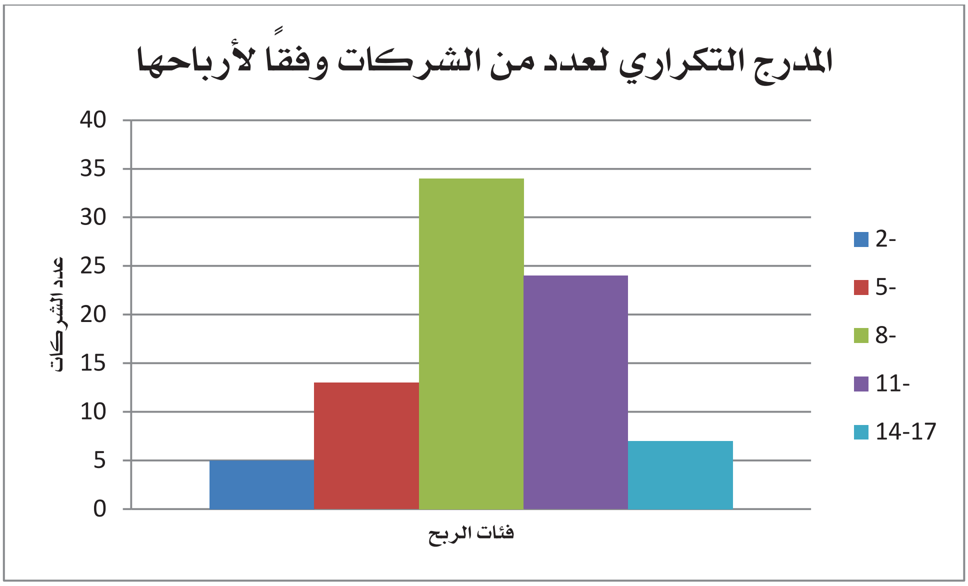
فئات الربح	-2	-5	-8	-11	17-14
عدد الشركات	5	13	34	24	7

المطلوب : التمثيل البياني للجدول السابق من خلال

1. اعداد المدرج التكراري واستنتج منه قيمة المنوال ؟
2. اعداد المضلع التكراري
3. اعداد المنحنى التكراري
4. اعداد المنحنى التكراري المتجمع الصاعد والهابط واستنتج منه قيمة الوسيط ؟
5. ماهو عدد الشركات التي ربحها اقل من 10 مليون ؟
6. ماهي نسبة الشركات التي ربحها 8 مليون فاكثر ؟
7. ماهي نسبة الشركات التي ربحها 9 مليون فاكثر ؟

✓ طريقة الحل :

1- اعداد المدرج التكراري واستنتج منه قيمة المنوال :



الجدول المستخدم في الاكسل (النسخه الانجليزية ويبدأ التضليل من الفئات)

فئات الربح	عدد الشركات
-2	5
-5	13
-8	34
-11	24
17-14	7
مجموع التكرارات	83

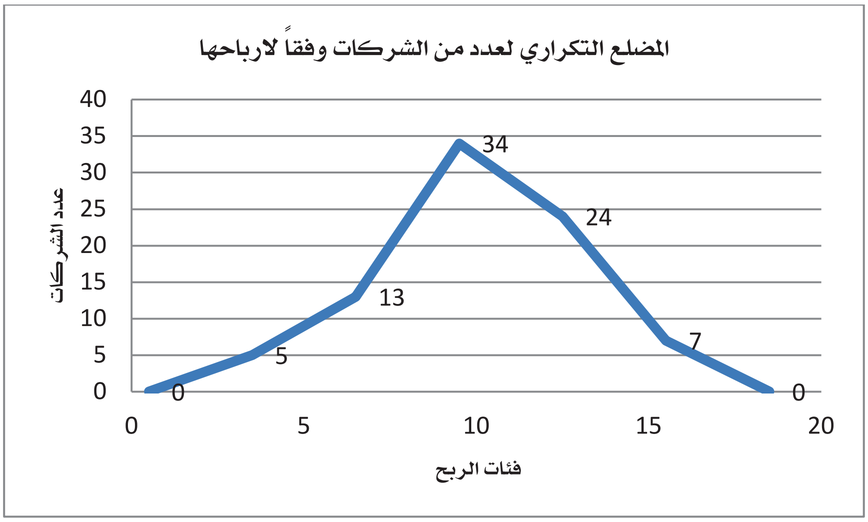
2- اعداد المضلع التكراري:

ا. نحصل على مراكز الفئات

فئات الربح	-2	-5	-8	-11	17-14	المجموع
مركز الفئات	22.5	27.5	32.5	37.5	42.5	
عدد الشركات	5	13	34	24	7	83

II. استحداث فئتين سابقة ولاحقة للتوزيع وحساب مركز الفئة لكل منهما ، وذلك باعتبار الفئة السـابقة سـوف تكـون (-1) - 2 و اللاحقة 17-20 وتكرارهما صفر ونقوم بحساب مراكزهما كالتالي :

فئات الربح	-(1-)	-2	-5	-8	-11	17-14	20-17	المجموع
مركز الفئات	0.5	3.5	6.5	9.5	12.5	15.5	18.5	
عدد الشركات	0	5	13	34	24	7	0	83

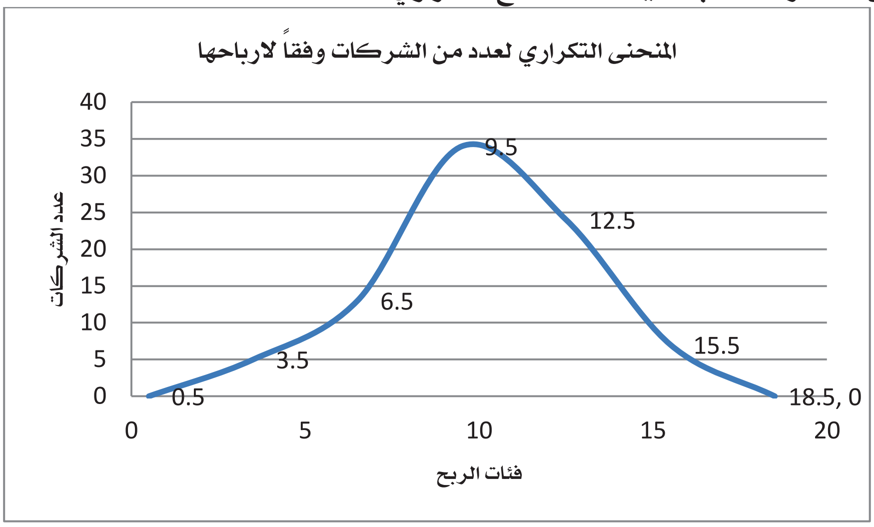


الجدول المستخدم في الاكسل

فئات الربح (مركز الفئات)	0.5	3.5	6.5	9.5	12.5	15.8	18.5
عدد الشركات	0	5	13	34	24	7	0

3- اعداد المنحنى التكراري:

نفس الخطوات السابقة في اعداد المضلع التكراري

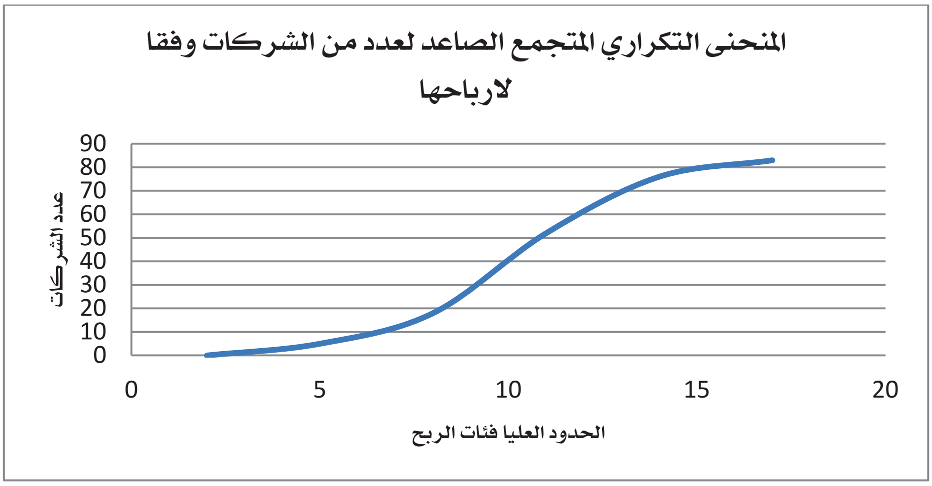


4- اعداد المنحنى التكراري المتجمع الصاعد والهابط واستنتج منه قيمة الوسيط ؟:

أ. اولا : المنحنى التكراري المتجمع الصاعد: يجب علينا اعداد الجدول المتجمع التكراري الصاعد لكي نتمكن من رسم هذا الشكل البياني (راجع الحاضرة الخامسة)

الحدود العليا للفئات	التكرار المتجمع الصاعد
اقل من 2	صفر
اقل من 5	5
اقل من 8	18
اقل من 11	52
اقل من 14	76
اقل من 17	83

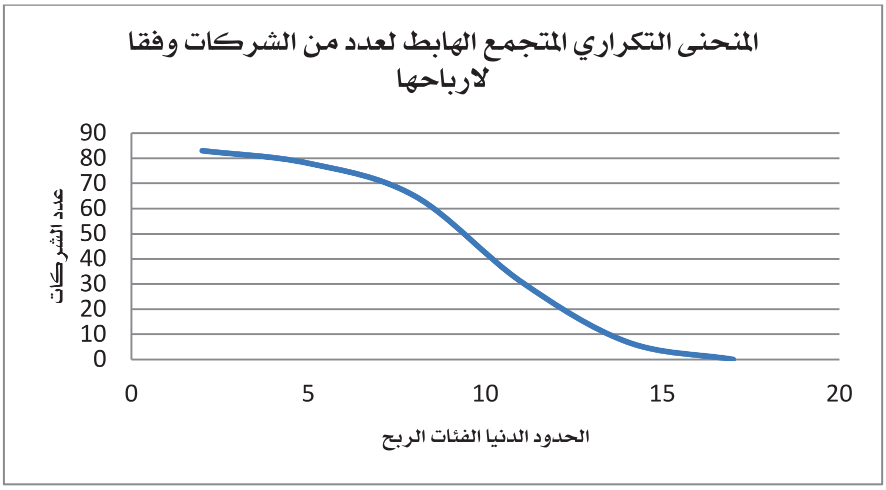
رسم الشكل البياني ببرنامج الاكسل باستخدام الجدول اعلاه



II. ثانيا : المنحنى التكراري المتجمع الهابط : يجب علينا اعداد الجدول المتجمع التكراري الهابط لكي نتمكن من رســـم هذا الشكل البياني (راجع الحاضرة الخامسة)

الحدود الدنيا لفئات الربح	التكرار المتجمع الهابط
2 فاكثر	83
5 فاكثر	78
8 فاكثر	65
11 فاكثر	31
14 فاكثر	7
17 فاكثر	صفر

رسم الشكل البياني ببرنامج الاكسل باستخدام الجدول اعلاه



قيمة الوسيط ؟؟؟؟؟؟ و المنوال

1. ماهو عدد الشركات التي ربحها اقل من 10 مليون ؟ 18 شركة
2. ماهي نسبة الشركات التي ربحها 8 مليون فاكثر ؟ $65 \div 83 = 0.7831$ ، $\therefore$ $0.7831 \times 100 = 78.31$ ٪
3. ماهي نسبة الشركات التي ربحها 9 مليون فاكثر ؟ $31 \div 83 = 0.3734$ ، $\therefore$ $0.3734 \times 100 = 37.34$ ٪

(المحاضرة السابعة)

المقاييس الإحصائية للبيانات غير المبوبة (الجزء الاول)

أولا: مقاييس النزعة المركزية

سبق واستعرضنا مراحل ابحث العلمي واتضح لنا ان البحث الاحصائي له نفس المراحل فبعد جمع البيانات والمعلومات data collection لابد من عرض هذه البيانات في شكل جدولي او في شكل الرسومات البيانية data presentation and tabulation مما يسهل من فهل واستيعاب مضمونها ، وتأتي بعد ذلك المرحلة التالية وهي تحليل البيانات data analysis والتي فيها يتم استخدام الاداوات الاحصائية المختلفة لوصف البيانات من خلال حساب المقاييس الاحصائية المختلفة التي سوف نستعرضها في هذه المحاضرة بمشيئة الله .

❖ المقاييس الاحصائية :

تتمثل اهمية عملية وصف البيانات كميًا من خلال محاولة الوصول الى فهم ورؤية اوضح للمعلومة المحتواه في القيم الكمية للمتغيرات محل الدراسة ومحاولة التعبير عن تلك البيانات الكمية بقيم تصف طبيعة وشكل المتغيرات محل الدراسة بالطريقة التي تمكننا من التعامل معها بشكل ادق وافضل ويطلق على تلك القيم المقاييس الاحصائية .

المقاييس الاحصائية لم توجد من تلقاء نفسها وانمها دعت الحاجة الى وجودها حيث تساعدنا في وصف المتغيرات المختلفة عن طريق معرفة القيم التي تتركز من حولها البيانات ومدى التفاوت بين قيم المفردات محل الدراسة وتلك القيم ، كما تساعدنا في المقارنة بين المتغيرات المختلفة من حيث مدى نزعتها نحو مراكز معينة وتحديد مدى تجانس البيانات بعضها مع بعض .

وتنقسم المقاييس الاحضائي الى نوعين رئيسيين هما :

1- مقياس النزعة المركزية central tendencu measures
2- مقاييس التشتت او الانتشار dispersion measures

في هذه المحاضرة سنتعرض لكيفية حساب مقاييس النزعة المركزية ومقاييس التشتت في حالة استخدام البيانات الخام غير المبوبة ، أي تلك التي لم يتم تصنيفها في صورة جداول تكرارية ، وذلك هو الاصل في التحليل الاحصائي للبيانات ، لانه يعطي الصورة الحقيقة للنتائج بدون اي تدخل شخصي فيها ، الا ان ذلك لا يقلل ايضا من اهمية الحاجة لدراسة كيفية حساب المقاييس الاحصائية المختلفة من البيانات المبوبة والتي سنتعرض لها في المحاضرات التالية ان شاء الله .

اولا : مقاييس النزعة المركزية CENTRAL TENDENCY MEASURES : نقصد بمقاييس النزعة المركزية تلك القيم الوسطى التي توضح القيمة التي تجمع اكبر عدد من القيم الخاصة بمجموعة معينة عندها ، او هي قيمة تلك الدرجة التي يمكن ان تعتبر ممثلة لكافة الدرجات الموجوده في تلك المجموعة ، ولتحديد القيمة المتوسطة للتوزيع يوجد هناك عدة مقاييس اهمها :

- المتوسط الحاسبي
- الوسيط
- المنوال او الشائع
- الربيع الادنى (الاول) – مذكور في الكتاب فقط
- الربيع الاعلى (الثالث) – مذكور في الكتاب فقط

كما يوجد عدة مقاييس اخرى اقل شيوعا مثل :

- الوسط الهندسي
- الوسط التوافقي
- العشير
- المئين

❖ اهمية حساب مقاييس النزعة المركزية : حساب مقاييس النزعة المركزية يساعد على التالي :

- ايجاد الرقم المتوسط والذي يدل على خصائص ارقام مجموعة من المجموعات فيكفي ان ننظر الى ذلك الرقم المتسوط لنعرف الكثير عن خصائص هذه المجموعة من الارقام .
- يمكننا ان قارن بين عدة مجموعات في وقت واحد فنقول ان هذه المجموعة اقوى من تلك وذلك اعتماداً على مقارنة هذه المتوسطات بعضها ببعض .

1) الوسط الحسابي (المتوسط) Mean

يعرف المتوسط الحاسبي بأنه القيمة التي اذا اعطيت لكل مفرد من مفردات الظاهرة لكان مجموع القيم الجديدة مساويا للمجموع الفعلي للقيم الاصلية للظاهرة ، أي ان الوسط الحسابي يساوي مجموع القيم مقسوما على عددها .

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

الحل بالالة الحاسبة CASIO 991ES PLUS للبيانات الغير مبوبة

1- مرحلة تجهيز الالة وادخال البيانات :

$MODE \rightarrow 3\,(STAT) \rightarrow 1(1-VAR)$ او $AC \rightarrow SHIFT \rightarrow 1(STAT) \rightarrow 2(DATA)$

$\rightarrow AC$ (لانهاء الادخال البيانات والتجهيز للعمليات المطلوبة)

2- مرحلة الحصول على الوسط الحسابي :

الوسط الحسابي $= SHIFT \rightarrow 1\,(STAT) \rightarrow 4(VAR) \rightarrow 2\,(\bar{x}) \rightarrow =$

⇐ مثال :

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام 1427 هـ بألف ريال كما يلي:

الشهر	محرم	صفر	ربيع أول	ربيع ثان	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذى القعدة	ذى الحجة
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

◄ المطلوب: حساب المتوسط الحسابى للمبيعات الشهرية.

✓ طريقة الحل :

يمكن استخدام العلاقة السابقة في حساب المتوسط الحسابي او متوسط المبيعات الشهرية كما يلي :

$$\text{متوسط المبيعات} = \frac{\text{اجمالي المبيعات الشهرية}}{\text{عدد الاشهر}} = \frac{69}{12} = 5.75$$

ويجب ملاحظة عدة امور في الوسط الحسابي وهي :

- انه لا يشترط ان يكون المتوسط الحسابي عددا صحيحا .
- ان المتوسط الحسابي دائما يكون محصورا بين اقل القيم واعلاها ، ولكن هذا لا يعني انه يقع في الوسط تماما بين هذين الحدين .
- ان المجموع الجبري لانحراف القيم عن المتوسط يكون دائما صفر . (بمعنى لو طرحنا المتوسط الحسابي من قيمة مبيعات كل شهر في المثال السابق (مثلا الشهر الاول 3-5.75 = -2.75) وهكذا

لجميع الاشهر ومن ثم جمعنا الناتج (طبعا كل رقم وطبيعته بدون تقيير -2.75 تجمع كما هي سالبة دون تغيير اشارتها) فيكون الناتج للجمع هو صفر .

لشهر	محرم	صفـر	ربيع أول	ربيع ثان	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذى القعدة	ذى الحجة	المجموع
لبيعات	3	5	8	3	6	4	12	5	4	3	7	9	69
راف القيم	-2.75	-0.75	2.25	-2.75	0.25	-1.75	6.25	-0.75	-1.75	-2.75	1.25	3.25	0

- ومن أهم خصائص الوسط الحسابي هو تأثره بجميع العمليات الجبرية تجري على البيانات من اضافة قيمة لجميع البيانات او طرحها او ضربها او قسمتها . **والمثال التالي يوضح المقصود من هذه النقطة** :

⇐ **مثال**

بسؤال خمسة أشخاص عن أجرهم الشهرى فكانت إجاباتهم كما يلى بالألف ريال:

7,3 ,2 , 5 , 3

◄ **المطلوب:**

1- أحسب متوسط الأجر الشهرى

2- وإذا قررت إدارة الشركة زيادة أجورهم أحسب متوسط الأجر الجديد فى الحالتين التاليتين :

- زيادة اجور العاملين بمقدار 2000 ريال
- زيادة أجور العاملين بنسبة 5 ٪

✓ **الحل :**

1- متوسط الاجر الشهري $4 = \frac{20}{5} = \frac{3+5+2+3+7}{5} = \bar{x}$ بمعنى ان متوسط الرواتب الشهرية 4 ريال .

2- اذا قررت الشركة زيادة اجورهم حسب الحالتين التاليتين :

- زيادة اجور العاملين بمقدار 2 الفين ريال $6 = \frac{30}{5} = \frac{5+7+4+5+9}{5} = \bar{x}$ بمعنى اصبح متوسط الرواتب 6 الاف ريال بعد زيادة كل راتب 2000 ريال
- زيادة اجور العاملين بنسبة 5 ٪ : $= \frac{21}{5} = \frac{3.15+5.25+2.10+3.15+7.35}{5} = \bar{x}$ 4.20 بمعنى اصبح متوسط الرواتب 4200 ريال (نقدر نختصر هالقصة بضرب المتوسط الاصلي 4 × 5 ٪ = 4.20 المتوسط الجديد .

❖ مزايا وعيوب المتوسط الحاسبي :

المزايا	العيوب
▪ يعد المتوسط الحسابي من اكثر مقاييس النزعة المركزية استخدما ، واسهلها فهما وذلك نتيجة لسهولة حسابه . ▪ يدخل في حسابه كل القيم دون اهمال اي قيمة منها .	▪ يتأثر بالقيم المتطرفة الشاذة سواء كانت قليلة او كثيرة ، فقد يرتفع لمجرد وجود قيمة مرتفعة ، او يقل كثيرا لمجرد وجود قيمة واحدة صغيرة وهذا بالتالي يؤدي الى عدم تمثيل المتوسط لواقع المعلومات . ▪ لايمكن ايجاده من خلال الرسم .

2) الوسيط Median

هو الدرجة التي تتوسط مجموعة من الدرجات المرتبة ترتيبا تصاعديا او تنازليا ، اي هو القيمة التي يصغرها عدد من القيم يساوي العدد الذي يكبر هذه القيمة ويمكن حساب الوسيط Me باتباع الخطوات التالية :

- ترتيب الدرجات تصاعديا او تنازليا

- ايجاد ترتيب الوسيط ويقصد به ايجاد مكان الوسيط ويختلف ترتيب الوسيط اذا كان عدد المشاهدات فردي او زوجي كما يلي :

عدد المشاهدات n	قانون ايجاد ترتيب الوسيط Me
فردي	$\frac{n+1}{2}$
زوجي	$\frac{n}{2}$ و $\left(\frac{n}{2}\right)+1$ بمعنى نبحث عن القيمة الموجوده في الترتيب الناتج عن القانون الاول ومن ثم القيمة الموجوده في الترتيب الناتج عن القانون الثاني

- ايجاد قيمة الوسيط

عدد المشاهدات n	قانون ايجاد الوسيط Me
فردي	$\frac{n+1}{2}$ بمعنى قيمة المتغير الموجود في الترتيب السابق مثلا لوكان الناتج هو 5 فنبحث عن المتغير الموجود في البند الخامس بعد ترتيب البيانات تصاعديا او تنازليا
زوجي	$\frac{n}{2}$ و $\left(\frac{n}{2}\right)+1$ بعد ايجاد بنود الوسيط ناخذ القيم ونجمعها على و نقسمها على 2 $$\frac{value\left(\left(\frac{n}{2}\right)+1\right)+value\left(\frac{n}{2}\right)}{2}$$

⇐ مثال :

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام 1427 هـ بالألف ريال كما يلي:

الشهر	محرم	صفر	ربيع أول	ربيع ثان	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذى القعدة	ذى الحجة
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

◄ المطلوب: إيجاد قيمة الوسيط للبيانات السابقة.

✓ طريقة الحل :

اولا : نرتب المبيعات تصاعديا او تنازليات : فتصبح كالتالي :

3 , 3 , 3 , 4 , 4 , 5 , 5 , 6 , 7 , 8 , 9 , 12

ثانيا : نوجد ترتيب الوسيط وهنا عدد القيم زوجي فمعنا ذلك لدينا ترتيبين $\frac{12}{2}=6$ و $\left(\frac{12}{2}\right)+1=7$ بمعنى القيم الموجودة في البند السادس والسابع من القيم وفي حالتنا سوف يكون 5 و 5

ثالثا : ايجاد الوسيط باستخدام القانون (جمع القيمتين المتواجدين في البند السادس والسابع وقسمتهما على 2 فاذا الحل سيكون $\frac{5+5}{2}=\frac{10}{2}=5$ اذا قيمة الوسيط هي 5

❖ مزايا وعيوب الوسيط :

المزايا	العيوب
▪ لا يتأثر بالقيم الشاذة. ▪ يمكن استخدام الوسيط في البيانات الناقصة. ▪ يمكن الحصول على الوسيط وحسابه من خلال الرسم. ▪ يمكن استخدام الوسيط في البيانات التي يعرف ترتيبها ولا تعرف قيمتها.	▪ لايعتمد على جميع القيم حيث انه لا يدخل في حسابه سوى قراءة واحدة او قراءتين من البيانات كلها.

3) المنوال Mode

هو القيم التي تعتبر اكثر القيم شيوعاً. وعلى ذلك فتحديده يتوقف على تكرار القيم في المجموعة.

⇐ مثال :

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام 1427 هـ بالألف ريال كما يلي:

الشهر	محرم	صفر	ربيع أول	ربيع ثان	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذى القعدة	ذى الحجة
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

◄ المطلوب: احسب المنوال.

✓ طريقة الحل :

نجد ان المبيعات الاكثر تكرارا هنا هي 3 الاف ريال لذلك فان المنوال هنا هو = 3

ويجب ملاحظة عدة امور في المنوال وهي :

- قد يكون في التوزيع منوالين او اكثر وذلك كالمثال التالي : 6 ، 5 ، 5، 5، 4، 4، 4
  فالمنوال هنا = 4 ، 5 اي انه يوجد منوالين .
- وقد لا يكون في التوزيع منوال وذلك كالمثال التالي : 2 ، 5، 7، 9 ، 11
- يتضح لنا ان تحديد المنوال في مجموعة بيانات غير مبوبة سهل

❖ مزايا وعيوب المنوال :

المزايا	العيوب
▪ سهل الحساب سواء بالرسم او بالحساب. ▪ لا يتأثر كثيرا بالقيم الشاذة . ▪ لايتأثر كثيرا لو تغيرت قيم بعض مفردات البيانات.	▪ اقل مقاييس النزعة المركزية استعمالا. ▪ عدم الفائدة في البيانات القليلة العدد.

4) الوسط الهندسي :

نتيجة ان الوسط الحسابي يتأثر بالقيم الشاذة دعت الحاجة الى وجود مقاييس لا تتأثر بقدر الامكان بالقيم الشاذة والمتطرفة ومن تلك المقاييس الوسط الهندسي GM والذي يكون مفيد في بعض التطبيقات الاقتصادية ودراسات نمو الظواهر الديموجرافية وكذلك في حساب الارقام القياسية ، فالوسط الهندسي هو الجذر النوني لحاصل ضرب القيم محل الدراسة ويمكن حسابه من خلال المعادلة التالية :

$$GM = \sqrt[n]{x_1 \times x_2 \times x_3 \times \ldots \times x_n}$$

⇐ مثال :

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام 1427 هـ بالألف ريال كما يلي:

الشهر	محرم	صفر	ربيع أول	ربيع ثان	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذى القعدة	ذى الحجة
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

◄ المطلوب: احسب الوسط الهندسي .

✓ طريقة الحل :

يمكن تطبيق المعادل السابقة على البيانات الموجوده في المثال ولكن قد يكون الامر صعبا في حالة ما تكون المشاهدات محل الدراسة n كبيرة الحجم .

لذا يمكن حسابه كما يلي :

$$GM = \sqrt[12]{3 \times 5 \times 8 \times 3 \times 6 \times 4 \times 12 \times 5 \times 4 \times 3 \times 7 \times 9}$$

$$GM = \sqrt[12]{391910400} = 5.2014$$

خواص الوسط الهندسي :

1- يعطي نتائج اكثر اعتدالا من المتوسط الحسابي .
2- تتوقف قيمته على سائر القيم دون استثناء او استبعاد ، شأنه شأن الوسط الحسابي .
3- أقل تأثيرا بالقيم المتطرفة عن الوسط الحسابي .

❖ مزايا وعيوب الوسط الهندسي :

المزايا	العيوب
▪ اكثر تمثيلا للقيم عن الوسط الحسابي باعتبار انه لايتأثر بالقيم المتطرفة بنفس درجة الوسط الحسابي . ▪ يعتبر من انسب المقاييس لحساب متوسطات النسب ومعدلات النمو . ▪ يعتبر من اكثر مقاييس النزعة المركزية ملائمة لحساب الارقام القياسية للمناسيب.	▪ لايمكن حسابه اذا كانت احدى القيم صفر . ▪ لايمكن استخدامه اذا كان ناتج حاصل ضرب قيم المشاهدات محل الدراسة سالب . ▪ صعوبة حسابه يدويا وانما يمكن ذلك باستخدم الحاسب الالي او الحاسبة.

5) الوسط التوافقي Harmonic mean (يوجد في الكتاب فقط ووضعته من باب الاحتياط)

يعتبر الوسط التوافقي HM من المقاييس التي تحد من تأثير القيم المتطرفة وخاصة حالة التطرف نحو الكبر وهو عبارة عن مقلوب الوسط الحسابي لمقلوبات القيم .

لذالك يمكن حساب الوسط التوافقي من خلال المعادلة التالية :

$$\frac{1}{\mathrm{HM}} = \frac{1}{n}\left(\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}\right)$$

ويعتبر الوسط التوافقي من اكثر المتوسطات صلاحية في حالة الظاهرة التي تقاس بالنسبة لوحدة ثابته كوحدة الزمن مثلا .

⇐ مثال :

(جدول المبيعات الشهرية المستخدم في جميع الامثلة السابقة):

◄ المطلوب: احسب الوسط التوافقي .

✓ طريقة الحل :

يمكن تطبيق المعادلة السابقة على البيانات الموجودة في الجدول ولكن قد يكون الامر صعب في حالة ما تكون المشاهدات محل الدراسة n كبير الحجم :

$$\frac{1}{\mathrm{HM}} = \frac{1}{12}\left(\frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{9}\right) = 4.74502$$

اذا الوسط التوافقي للمبيعات الشهرية هو 4.74502

(المحاضرة السابعة الجزء الثاني)

المقاييس الإحصائية للبيانات غير المبوبة

ثانيا: مقاييس التشتت أو الأنتشار DISPERSION MEASURES

كما تميل القيم الى التمركز فانها تميل ايضا الى التشتت او الانتشار ، وبالتالي فإن أي توزيع من القيم له صفة التمركز . وصفة التشتت .

فمقاييس التشتت : هي تلك المقاييس التي تعبر عن مدى تباعد القيم او تقاربها في المجموعات التي يشملها البحث .

مثال:

مجموعة (أ) : 8 ، 8 ، 8 ، 8 ، 8

مجموعة (ب) : 1 ، 2 ، 3 ، 5 ، 6

نلاحظ أن المجموعة الأولى (أ) لا يوجد بها تشتت، فهذه المجموعة متجانسة ، في حين نلاحظ ان المجموعة الثانية (ب) يوجد بها تشتت

بمعنى في المجموعة الاولى لا يوجد تباعد بين ارقامها فكلها رقم 8 اما في المجموعة ب فهناك تباعد بين قيمها مثلا 1 تبعد عن رقم 2 بمقدار 1 و رقم 2 يبعد عن رقم 3 بمقدار 1 و3 يبعد عن رقم 5 بمقدار 2 وهكذا فهذا هو مقياس التشت ولهذا اسمينها مجموعة غير متجانسة .

ويمكن ان يقاس تشتت البيانات عن طريق مقاييس التشتت المختلفة واهم هذه المقاييس :

- المدى
- المدى الربيعي
- الانحراف عن المتوسط
- التباين
- الانحراف المعياري

لماذا نستخدم مقاييس التشتت : نستخدم هذه المقاييس اذا كان عندنا مجموعتين ونريد ان نقارن بينهما ، وكان المتوسط فيما بينهم متساوي ، كما في المثال التالي :

- مجموعة (أ) : (45 ، 50 ، 55) المتوسط هنا = 50 (قانون الوسيط لمجموعة فردية $\frac{n+1}{2}$)
- مجموعة (ب) : (30 ، 50 ، 70) المتوسط هنا = 50

فلذا لانستطيع ان نقول هنا ان المجموعتين متساويتين لاننا اذا رجعنا الى المجموعتين وجدنا انهما مختلفتين في الدرجات رغم تساوي المتوسطين حيث ان المتوسط الحسابي في المجموعتين يساوي (50) ، لكن اذا استخدمنا احد مقاييس التشتت مثل المدى مثلا والذي يحسب من خلال العلاقة التالية :

قانون مقياس المدى = أعلى درجة – اقل درجة

وعلى ذلك فإن مقياس المدى هو :

- مدى المجموعة (أ) = 55 – 45 = 10
- مدى المجموعة (ب) = 70 – 30 = 40

نرى ان درجة التشتت في المجموعة (أ) اقل منها في المجموعة (ب) أي ان المجموعة (أ) تكون اكثر تجانسا من المجموعة (ب)

الحل بالالة الحاسبة CASIO 991ES PLUS للبيانات الغير مبوبة

1- مرحلة تجهيز الالة وادخال البيانات :

$MODE \rightarrow 3\ (STAT) \rightarrow 1(1-VAR)$ او $AC \rightarrow SHIFT \rightarrow 1(STAT) \rightarrow 2(DATA)$
$\rightarrow AC$ (لانهاء الادخال البيانات والتجهيز للعمليات المطلوبة)

2- مرحلة الحصول على المدى للبيانات الغير مبوبة:

المدى $= SHIFT \rightarrow 1\ (STAT) \rightarrow 6(MAXMIN) \rightarrow 2\ (MAX) \rightarrow - \rightarrow SHIFT$
$\rightarrow 1\ (STAT) \rightarrow 6(MAXMIN) \rightarrow 1\ (MIN) =$

1) المدى Range :

المدى هو الفرق بين اعلى درجة واقل درجة في التوزيع ، ويعتبر المدة الوسيلة المباشرة لمعرفة مدى تقارب القيم او تباعدها في اي توزيع ، وهو وسيلة سهلة ، الا انها اقل الوسائل دقة وذلك لان حسابه يتوقف على قيمتين فقط من قيم المجموعة ، ولايهتم مطلقا بما بينهما من قيم اخرى . فالمدى لا يصلح الا اذا اراد الباحث ان يأخذ فكرة سريعة عن مدة تشتت بيانات التوزيع موضع الدراسة ، الا ان استخدامه والاعتماد عليه قد يؤديان الى نتائج خادعة ، وخاصة اذا كان هناك انفصال بين الدرجات المتطرفة وباقي الدرجات موضع البحث.

⇐ مثال :

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام 1427 هـ بالألف ريال كما يلي:

الشهر	محرم	صفـر	ربيع أول	ربيع ثان	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذى القعدة	ذى الحجة
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

◄ المطلوب: احسب المدى للمبيعات الشهرية.

✓ طريقة الحل :

نلاحظ ان اكبر قيمة للمبيعات هي 12 واقل قيمة للمبيعات الشهرية هي 3 لذلك يكون المدى 9

Range = 12-3 = 9

❖ عيوب المدى :

نجد ان من اهم عيوب المدى انه يتم حسابه بناء على اكبر واصغر قيمتين وبالتالي في حالة كونهما او احدهما متطرفتين او قيم شاذة فان المدى يعطي نتائج مضللة .

2) متوسط الانحرافات المطلقة Average absolute Deviation :

متوسط الانحرافات المطلقة AAD هو ذلك المقياس الذي يقيس تباعد كافة القيم عن المتوسط الحسابي .

وعلى الرغم من ان حساب نصف المدى الربيعي يقضي على اثر القيم المتطرفة ، والتي تؤثر على حساب المدى المطلق ، الا انها جميعا (المدى ، ونصف المدى الربيعي) يتناولان التباعد بين قيمتين فقط (اعلى قيمة وادنى قيمة) في المدى ، (وقيمة الربيع الادنى وقيمة الربيع الاعلى) في نصف المدى الربيعي ، وذلك من بين القيم موضع الدراسة ام ابقية القيم تبقى مهملة .

وهذا ما ادى الى تطبيق متوسط الانحرافات المطلقة AAD الذي يقيس تباعد كافة القيم عن متوسطها الحسابي .

ويمكن حساب متوسط الانحرافات المطلقة من خلال المعادلة التالية :

$$AAD = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

تفسير القانون : نطرح كل قيمة من الوسط الحسابي المحسوب بالقانون التالي :

$5.75 = \frac{69}{12} = \frac{\text{اجمالي المبيعات الشهرية}}{\text{عدد الاشهر}} =$ متوسط المبيعات ، بعد ذلك نستخرج القيم المطلقة لكل ناتج مع عملية الطرح ونقوم بجمع جميع النتائج وبعد ذلك نقسمها على عدد الفئات .

⇐ مثال :

(جدول المبيعات الشهرية المستخدمة في جميع الامثلة السابقة)

◄ المطلوب: احسب متوسط الانحرافات المطلقة لبيانات المبيعات الشهرية.

✓ طريقة الحل :

في هذا المثال سبق وان حسبنا قيمة الوسط الحساب 5.75 والان نطبق القانون بالطريقة التالية :

المبيعات (المتغير X)	الانحراف $(x - \bar{x})$	القيمة المطلقة للناتج $\lvert x - \bar{x} \rvert$
3	2.75-=5.75-3	2.75
5	0.75-=5.75-5	0.75
8	2.25=5.75-8	2.25
3	2.75-=5.75-3	2.75
6	0.25=5.75-6	0.25
4	1.75-=5.75-4	1.75
12	6.25=5.75-12	6.25
5	0.75-=5.75-5	0.75
4	1.75-=5.75-4	1.75
3	2.75-=5.75-3	2.75
7	1.25=5.75-7	1.25
9	3.25=5.75-9	3.25
المجموع	0	26.5

وعلى ذلك يمكن حساب متوسط الانحرافات المطلقة كما يلي : $AAD = \frac{26.5}{12} = 2.2083$

3) التباين والانحراف المعياري :

i. التباين variance : هو متوسط مربعات القيم عن وسطها الحسابي ، ويرمز لها بالرمز (σ^2 تقرا سيجما تربيع) وذلك اذا كان محسوب لبيانات المجتمع اما في حالة حساب لبيانات عينة من المجتمع فيرمز له بالرمز S^2 .

ii. الانحراف المعياري Standerd Deviation : هو الجذر التربيعي لمتسوط مربعات انحرافات القيم عن وسطها الحسابي ، اي هو جذر التباين لذلك يرمز له بالرمز σ (و تقرأ سيجما) و ذلك اذا كان محسوب لبيانات المجتمع اما في حالة حسابه لبيانات عينة من المجتمع فيرمز له بالرمز S. ويعتبر الانحراف المعياري والتباين من اهم مقاييس التشتت جميعا او اكثرها استعمالا ، وهما قريب في خطوات ايجاده من الانحراف عن المتوسط ، فهو يختلف عنه في طريقة التخلص من اشارات الفروق بين القيم والمتوسط الحسابي ، فبينما نتخلص من هذه الاشارات في طريقة الانحراف عن المتوسط باعمال الاشارات كلية ، نحتال على ذلك في طريقة الانحراف المعياري بتربيع هذه الفروق (اي بضربها في نفسها) فتصبح بالتالي جميع الاشارات موجبة .

طريقة حساب التباين والانحراف المعياري :

1- في حالة البيانات من المجتمع : يمكن حساب التباين من خلال المعادلة التالية :

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

الحل بالالة الحاسبة CASIO 991ES PLUS للبيانات الغير مبوبة

1- مرحلة تجهيز الالة وادخال البيانات :

$MODE \rightarrow 3\,(STAT) \rightarrow 1(1-VAR)$ او $AC \rightarrow SHIFT \rightarrow 1(STAT) \rightarrow 2(DATA)$

$\rightarrow AC$ (لانهاء الادخال البيانات والتجهيز للعمليات المطلوبة)

2- مرحلة الحصول على التباين للبيانات الغير مبوبة:

التباين $= SHIFT \rightarrow 1\,(STAT) \rightarrow 4(VAR) \rightarrow 4\,(SX) \rightarrow x^2 \rightarrow =$

و بالتالي يكون حساب الإنحراف المعياري كما يلي: $S = \sqrt{S^2}$

الحل بالالة الحاسبة CASIO 991ES PLUS للبيانات الغير مبوبة

1- مرحلة تجهيز الالة وادخال البيانات :

$MODE \rightarrow 3\,(STAT) \rightarrow 1(1-VAR)$ او $AC \rightarrow SHIFT \rightarrow 1(STAT) \rightarrow 2(DATA)$

$\rightarrow AC$ (لانهاء الادخال البيانات والتجهيز للعمليات المطلوبة)

2- مرحلة الحصول على الانحراف المعياري للبيانات الغير مبوبة:

الانحراف المعياري $= SHIFT \rightarrow 1\,(STAT) \rightarrow 4(VAR) \rightarrow 4\,(SX) \rightarrow =$

⇐ **مثال :**

(جدول المبيعات الشهرية المستخدمة في جميع الامثلة السابقة)

◄ **المطلوب:** احسب قيمة التباين وقيمة الانحراف المعياري للمبيعات الشهرية.

✓ **طريقة الحل :**

بكل سهولة باستخدام الالة الحاسبة يكون :

- الانحراف المعياري هو : 2.80016
- اما التباين فهو مربع الانحراف المعياري : $(2.80016^2) = 7.8409$

❖ **ملاحظة هامة :**

يعتبر من اهم خصائص الانحراف المعياري هو عدم تأثره بعمليات الجمع والطرح وانما يتأثر بعمليات الضرب والقسمة .

فنلاحظ عدم تغير قيمة الانحراف المعياري في حالة الجمع او الطرح وانما تظل قيمته كما هي بالرغم من طرح مقدار ثابت من جميع قيم التوزيع .

اما في حالة الضرب او القسمة فنلاحظ تغير قيمة الانحراف المعياري وهي نفس قيمة الانحراف المعياري القديمة مضروبة في القيمة التي ضرب فيها او قسم عليها .

⇐ **مثال (توضيحي) :**

(جدول المبيعات الشهرية المستخدمة في جميع الامثلة السابقة)

◄ **المطلوب:** اذا تم طرح 2 من جميع بيانات المبيعات الشهرية اي تم تخفيض المبيعات الشهرية بمقدار 2 احسب قيمة الانحراف المعياري الجديد.

✓ طريقة الحل :

خلاصة هذا المثال في صفحة 118 في الكتاب ان الطرح او الجمع لا يؤثر على قيمة الانحراف او التباين وانما ظلت قيمته كما هي بالرغم من طرح مقدار ثابت 2 من جميع قيم المبيعات الشهرية .

⇐ مثال :

(جدول المبيعات الشهرية المستخدم في جميع الامثلة السابقة)

◄ المطلوب: الانحراف المعياري للمبيعات الشهرية اذا تم زيادة المبيعات الشهرية الى 3 اضعاف الموجود حاليا ؟.

✓ طريقة الحل :

افضل واسرع طريقة هي استخراج الناتج كالتباين او الانحراف بالالة الحاسبة واجراء العملية المطلوبة على الناتج سواء كانت ضرب او قسمة وفي حالتنا هذه (بكل سهولة باستخدام الالة الحاسبة) يكون :

- الانحراف المعياري هو : 2.80016 × 3 (تمثل الزيادة المطلوبة) = 8.40048
- اما التباين فهو مربع الانحراف المعياري : $(8.40048^2) = 70.56818$

❖ خاتمة المحاضرة :

في نهاية هذه المحاضرة يمكن أن نكون حصلنا على كافة المقاييس الإحصائية الوصفية التي تصف المبيعات الشهرية فكانت كما يلي:

المتوسط	الوسيط	المنوال	الوسط الهندسي
5.75	5	3	5.20114

المدى	متوسط الانحرافات	التباين	الانحراف المعياري
9	2.20833	7.840909	2.80016

(المحاضرة الثامنة الجزء الاول)

المقاييس الإحصائية للبيانات المبوبة

❖ تمهيد :

يقصد بالبيانات المبوبة هي تلك البيانات التي تم وضعها في صورة جداول تكرارية . ونكون في حاجة لحساب المقاييس الاحضائية لذلك النوع من البيانات حيث ان الاصل هو حساب المقاييس الاحصائية للبيانات الخام اي الغير مبوبة ولكن في بعض الاحيان تكون البيانات الخام مفقودة او يصعب الحصول عليها لذا كان من الواجب علينا عرض لكيفية حساب المقاييس الاحصائية في تلك الحالة .

والجداول التكرارية للمتغير الكمي المتقطع يمكن تحويلها لتكون بيانات غير مبوبة ونتعامل معها كما سبق في الفصل السابق الا ان الامر يختلف بالنسبة للمتغير الكمي المتصل حيث يصعب ذلك ولا بد من التعامل معه كما هو على صورته الجدولية ، وهذا ما سوف نتناوله في هذا الفصل ان شاء الله ، وسيتم عرض لكيفية حساب كلا من مقاييس النزعة المركزية ومقاييس التشتت في ثلاث حالات للجداول التكرارية وهي :

1- الجداول المنتظمة   2- الجداول غير المنتظمة   3- الجداول المفتوحة

1- **الجداول المنتظمة :** وهي تلك الجداول التي تكون فيها اطوال الفئات جميعها متساوية .

اولا: الوسط الحسابي والتشتت حوله :

الوسط الحسابى كما سبق أن تم تعريفة فى الفصل السابق هو القيمة التى إذا أخذها جميع المفردات لكان مجموعها يساوى مجموع القيم الأصلية، ويمكن حساب الوسط الحسابى او المتوسط الحسابى من خلال البيانات المبوبة كما يلي:

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

حيث $\bar{x}$ = الوسط الحسابي ، x_i = مركز الفئة i ، f_i = تكرار الفئة i ، l = عدد الفئات

ويتم حساب التشتت حول المتوسط الحسابي من خلال كلا من :

(أ) متوسط الانحرافات المطلقة : وهو يقيس انحراف القيم عن وسطها الحسابي بغض النظر عن اشارة ذلك الانحراف حيث يقاس من خلال المعادلة التالية :

$$AAD = \frac{\sum_{i=1}^{l} |x_i - \bar{x}| f_i}{\sum_{i=1}^{l} f_i}$$

(ب) التباين σ^2 : وهو متوسط مجموع مربع إنحرافات القيم عن وسطها الحسابى. ويتم حسابه من خلال المعادلة التالية:

$$\sigma^2 = \frac{\sum_{i=1}^{l} (x_i - \bar{x})^2 f_i}{\sum_{i=1}^{l} f_i}$$

(ج) الانحراف المعياري σ : هو الجذر التربيعى للتباين، ويمكن حسابه من خلال المعادلة التالية:

$$\sigma = \sqrt{\sigma^2}$$

⇐ **مثال :**

البيانات التالية توضح توزيع مجموعة من المدرسين العاملين في مجال التربية وفقا لفئات أعمارهم فكانت النتائج كما يلي:

فئات العمر	20-	30-	40-	60-50
عدد العمال	10	30	50	20

◄ **المطلوب:** حساب مقاييس التشتت التالية :

(أ) الوسط الحسابي

(ب) التباين

(ج) الانحراف المعياري

(د) متوسط الانحرافات المطلقة

✓ **طريقة الحل :**

يتم اعداد الجدول التالي حتى يمكن حساب كلا من الوسط الحسابي والتباين والانحراف المعياري :

فئات العمر	التكرار f	مركز الفئة x	xf	x^2f
20-	10	25=2÷(30+20)	250= 10 × 25	6250 =10×$(25)^2$
30-	30	35	1050	36750
40-	50	45	2250	101250
60 – 50	20	55	1100	60500
المجموع	$\sum f = 110$		$\sum xf = 4650$	$\sum x^2f = 204750$

(أ) **الوسط الحسابي** : يمكن ايجاد الوسط الحسابي بتطبيق المعادلة الخاصة بذلك والموضحه سابقا وبالطريقة التالية :

$$\bar{x} = \frac{4650}{110} = 42.2727$$

الحل بالالة الحاسبة CASIO 991ES PLUS للبيانات المبوبة

1- ماقبل مرحلة التجهيز يجب ان تتأكد ان الحقل $FREQ$ يظهر لك في حالة ادخال البيانات في نمط الاحصاء واذا لم يكن ظاهر فيمكن اظاهره بالطريقة التالية :

$$SHIFT \rightarrow MODE \rightarrow \downarrow \rightarrow 4(STAT) \rightarrow 1(ON)$$

2- مرحلة تجهيز الالة وادخال البيانات :

$MODE \rightarrow 3\ (STAT) \rightarrow 1(1-VAR)$ او $AC \rightarrow SHIFT \rightarrow 1(STAT) \rightarrow 2(DATA)$

3- في البيانات المبوبة عند ادخال البيانات يتوجب علينا ادخل مركز الفئات اولا في خانة المتغير x واما القيم فيتم ادخالها في العمود $FREQ$ مع العلم بان هذا العمود لو كان ظاهرا في حال استخدامك للالة في البيانات الغير مبوبة فلن يسبب لك اي ضرر فقط لاتغير القيم الموجودة بداخله . ويكن ادخال في خانة المتغير طبقا لقانون مركز الفئات $($حد الفئة الاعلى $+$ حد الفئة الادنى$) \div 2$ لكل فئة مثال : $(10+20) \div 2$ او احسبها على ورقة خارجية . وبعد ذلك نقوم بانهاء الادخال بالطريقة التالية :

$\rightarrow AC$ (لانهاء الادخال البيانات والتجهيز للعمليات المطلوبة)

4- مرحلة الحصول على الوسط الحسابي للبيانات المبوبة:

الوسط الحسابي $= SHIFT \rightarrow 1\ (STAT) \rightarrow 4(VAR) \rightarrow 2\ (\bar{x}) \rightarrow =$

(ب) التبـايـن : يمكـن الحصـول علـى التبـايـن باسـتخدام المعادلـة المخصصـة بـذلك والمذكـورة سـابقاً بالشكل التالي :

$$\sigma^2 = \frac{204750}{110} - (42.2727)^2 = 74.3801$$

الحل بالالة الحاسبة CASIO 991ES PLUS للبيانات المبوبة

1- ماقبل مرحلة التجهيز يجب ان تتأكد ان الحقل $FREQ$ يظهر لك في حالة ادخال البيانات في نمط الاحصاء واذا لم يكن ظاهر فيمكن اظاهره بالطريقة التالية :

$$SHIFT \rightarrow MODE \rightarrow \downarrow \rightarrow 4(STAT) \rightarrow 1(ON)$$

2- مرحلة تجهيز الالة وادخال البيانات :

$$MODE \rightarrow 3\,(STAT) \rightarrow 1(1-VAR) \text{ او } AC \rightarrow SHIFT \rightarrow 1(STAT) \rightarrow 2(DATA)$$

3- في البيانات المبوبة عند ادخال البيانات يتوجب علينا ادخل مركز الفئات اولا في خانة المتغير x واما القيم فيتم ادخالها في العمود $FREQ$ مع العلم بان هذا العمود لو كان ظاهرا في حال استخدامك للالة في البيانات الغير مبوبة فلن يسبب لك اي ضرر فقط لاتغير القيم الموجودة بداخله . ويكن ادخال في خانة المتغير طبقا لقانون مركز الفئات $\left(\text{حد الفئة الاعلى} + \text{حد الفئة الادنى}\right) \div 2$ لكل فئة مثال : $(10+20) \div 2$ او احسبها على ورقة خارجية . وبعد ذلك نقوم بانهاء الادخال بالطريقة التالية :

$$\left(\text{لانهاء الادخال البيانات والتجهيز للعمليات المطلوبة}\right) AC \leftarrow$$

4- مرحلة الحصول على التباين للبيانات المبوبة:

$$\text{التباين} = SHIFT \rightarrow 1\,(STAT) \rightarrow 4(VAR) \rightarrow 3\,(\sigma x) \rightarrow x^2 \rightarrow =$$

(ج) الانحـراف المعيـاري : يمكـن الحصـول علـى الانحـراف الميعـاري باسـتخدام المعادلـة المخصصـة بذلك والمذكورة سابقاً بالشكل التالي :

$$\sigma = \sqrt{74.3801} = 8.62439$$

الحل بالالة الحاسبة CASIO 991ES PLUS للبيانات المبوبة

1- ماقبل مرحلة التجهيز يجب ان تتأكد ان الحقل $FREQ$ يظهر لك في حالة ادخال البيانات في نمط الاحصاء واذا لم يكن ظاهر فيمكن اظاهره بالطريقة التالية :

$$SHIFT \rightarrow MODE \rightarrow \downarrow \rightarrow 4(STAT) \rightarrow 1(ON)$$

2- مرحلة تجهيز الالة وادخال البيانات :

$$MODE \rightarrow 3\,(STAT) \rightarrow 1(1-VAR) \text{ او } AC \rightarrow SHIFT \rightarrow 1(STAT) \rightarrow 2(DATA)$$

3- في البيانات المبوبة عند ادخال البيانات يتوجب علينا ادخل مركز الفئات اولا في خانة المتغير x واما القيم فيتم ادخالها في العمود $FREQ$ مع العلم بان هذا العمود لو كان ظاهرا في حال استخدامك للالة في البيانات الغير مبوبة فلن يسبب لك اي ضرر فقط لاتغير القيم الموجودة بداخله . ويكن ادخال في خانة المتغير طبقا لقانون مركز الفئات $\left(\text{حد الفئة الاعلى} + \text{حد الفئة الادنى}\right) \div 2$ لكل فئة مثال : $(10+20) \div 2$ او احسبها على ورقة خارجية . وبعد ذلك نقوم بانهاء الادخال بالطريقة التالية :

$$\left(\text{لانهاء الادخال البيانات والتجهيز للعمليات المطلوبة}\right) AC \leftarrow$$

4- مرحلة الحصول على الانحراف المعياري للبيانات المبوبة:

$$\text{الانحراف المعياري} = SHIFT \rightarrow 1\,(STAT) \rightarrow 4(VAR) \rightarrow 3\,(\sigma x) \rightarrow =$$

(د) متوسط الانحرافات المطلقة AAD: حتى يمكن ايجاد متوسط الانحرافات المطلقة لابد اولا من ايجاد الانحرافات عن الوسط الحسابي ثم استخدامها في الحساب كما يتضح ذلك من الجدول التالي :

فئات العمر	التكرار f	مركز الفئة x	$x-\bar{x}$	$(x-\bar{x})f$	$\lvert x-\bar{x}\rvert f$
20–	10	25=2÷(30+20)	-17.2727	-172.727	172.7273
30–	30	35	-7.27273	-218.182	218.1819
40–	50	45	2.727273	136.3636	136.3636
60 – 50	20	55	12.72727	254.5455	254.5455
المجموع	$\sum f = 110$		تذكر اوجدنا في اول خطوة $\bar{x} = 42.2727$	$\sum (x-\bar{x})f = 0$	$\sum \lvert x-\bar{x}\rvert f = 781.8182$

ويمكن الحصول على متوسط الانحرافات المطلقة AAD بتطبيق المعادلة التالية :

$$AAD = \frac{\sum_{i=1}^{l} \lvert x_i - \bar{x}\rvert f_i}{\sum_{i=1}^{l} f_i} = \frac{781.8182}{110} = 7.1074$$

ويتضح لنا من الجدول السابق أن :

مجموع انحرافات القيم عن وسطها الحسابي تسوي صفر حيث ان $\sum (x-\bar{x})f = 0$
كما يمكن الاعتماد على هذه الانحرافات في حساب التباين بتطبيق المعادلة الخاصة بذلك .

ملاحظة : للحل بالالة *TI-84 plus* فقط ادخال الجدول اعلاه و استخرج اجمالي الخانة الاخيرة بتطبيق المعادلة في راس الجدول وبعد ذلك نقسم الناتج على n وهو اجمالي التكرارات

(المحاضرة الثامنة الجزء الثاني)

تابع المقاييس الإحصائية للبيانات المبوبة

❖ تمهيد :

استكمال لما سبق وتحدثنا عنه في الجزء الاول من المحاضرة الثامنة والخاصة المقاييس الإحصائية للبيانات المبوبة ما زلنا نتحدث عن مقاييس النزعة المركزية ومقاييس التشتت في الجداول المنتظمة والتي تعد الحالة الاولى من حالات الجداول التكرارية .

ثانيا : الوسيط والتشتت حوله :

الوسيط هو القيمة التي يصغرها عدد من القيم يتساوى مع العدد الذي يكبر هذه القيمة ، (بمعنى انه العنصر او الرتبة الموجوده في منتصف القيم) ، وهو يعتبر احد مقاييس النزعة المركزية التي نلجأ إليها لتحليل الظواهر وفقا لشكل التوزيع الاحصائي محل الدراسة .

ولحساب الوسيط من البيانات المبوبة هناك ثلاث خطوات يجب اتباعها وهي :

- ايجاد الجدول التكراري المتجمع الصاعد .
- ايجاد ترتيب الوسيط من خلال المعادلة التالية : $k_{med} = n \div 2$
- ايجاد قيمة الوسيط من خلال المعادلة التالية : $Med = L_{Med} + \frac{k_{med} - f_a}{f_b - f_a} \times I$

حيث ان (تعريف المصطلحات) :

- قيمة الوسيط Med
- ترتيب الوسيط k_{med}
- التكرار المتجمع اللاحق للفئة الوسيطية f_b
- الحد الادني لبداية الفئة الوسيطية L_{Med}
- التكرار المتجمع السابق للفئة الوسيطية f_a
- طول الفئة الوسيطية I ونحصل عليها بطرح الحد الادنى للفئة الوسيطة من الحد الاعلى

⇐ **مثال :**

البيانات التالية توضح توزيع مجموعة من المدرسين العاملين في مجال التربية وفقا لفئات أعمارهم فكانت النتائج كما يلي:

فئات العمر	20-	30-	40-	60-50
عدد العمال	10	30	50	20

◄ **المطلوب:** احسب قيمة الوسيط :

✓ **طريقة الحل :**

يتم اتباع الخطوات الثلاث المذكورة اعلاه :

1- ايجاد الجدول التكراري المتجمع الصاعد كما يلي :

الحدود العليا للفئات	التكرار المتجمع الصاعد
اقل من 20	0
اقل من 30	10
اقل من 40	40
اقل من 50	90
اقل من 60	110

$f_a(40)$

L_{Med}

$f_b(90)$

ترتيب الوسيط 55 بمعنى انه يقع بين التكرار المتجمع الصاعد 40 و الـ 90

2- ايجاد ترتيب الوسيط بناء على القانون التالي : $k_{med} = n \div 2$

$$k_{med} = 110 \div 2 = 55$$

اذا ترتيب الوسيط هو 55

3- ايجاد قيمة الوسيط : نلاحظ ان ترتيب الوسيط [55] مما يعني ان ترتيب الوسيط يقع بين التكرار المتجمع الصاعد [$f_a(40)$] وهو المقابل للحد الاعلى للفئة [40] وبين التكرار المتجمع الصاعد [$f_b(90)$] و هو المقبل للحد الاعلى للفئة [50] ، أي ان الحد الادنى للفئة هو [$L_{Med} = 40$]

وبالتالي يكون طول الفئة الوسيطية هو $I = 50 - 40 = 10$

والان وبعد توفر جميع عناصر المعادلة لدينا نستطيع تعويضها في القانون ايجاد قيمة الوسيط للوصول الى قيمة الوسيط كالتالي :

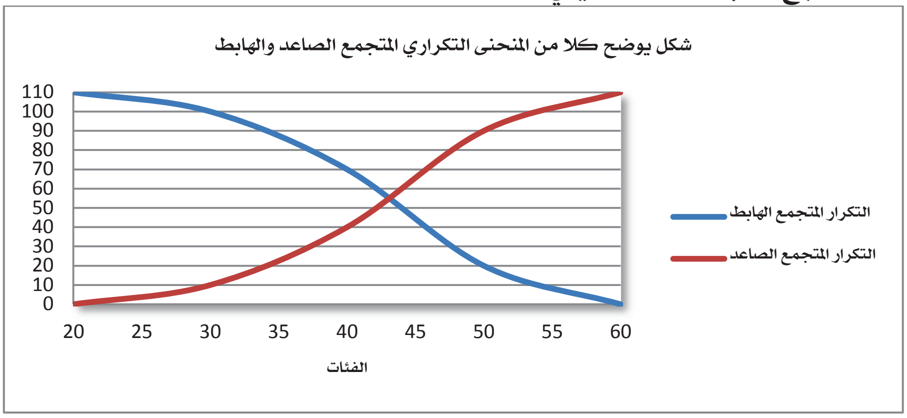
$$Med = 40 + \frac{55 - 40}{90 - 40} \times 10 = 43$$

كما يمكن ايجاد الوسيط عن طريق رسم كلا من المنحنى التكراري المتجمع الهابط والمنحنى التكراري المتجمع الصاعد كما يلي :

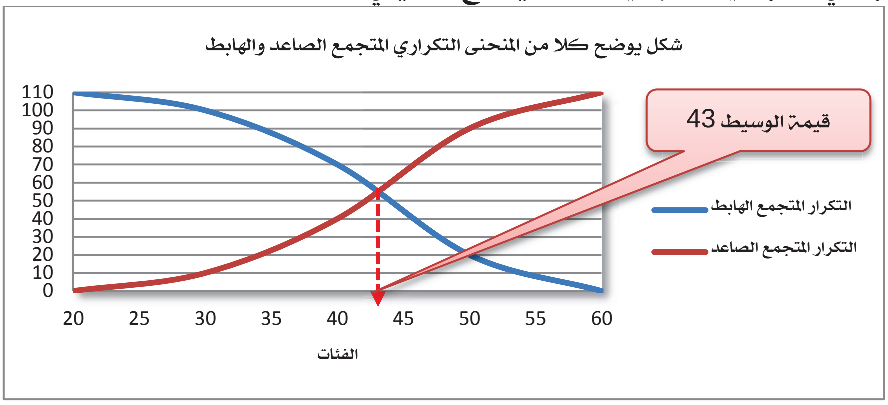
1- اولا يجب علينا ايجاد الجدول التكراري المتجمع الهابط كما يلي :

الحدود الدنيا للفئات	التكرار المتجمع الهابط
20 فاكثر	110
30 فاكثر	100
40 فاكثر	70
50 فاكثر	20
60 فاكثر	0

2- ثم بعد ذلك نقوم برسم كلا من المنحنى التكراري المتجمع الصاعد والمنحنى التكراري المتجمع الهابط معا كما يلي :



ويمكن الحصول على قيمة الوسيط من خلال الرسم بأن نسقط عمود رأسي من نقطة تقاطع كلا من المنحنى التكراري المتجمع الصاعد مع المنحنى التكراري المتجمع الهابط على المحور الرأسي لنقرأ قيمة الوسيط كما يتضح مما يلي :



ويتضح لنا من الشكل السابق ان الوسيط تبلغ قيمته 43 تقريبا.

مقاييس النزعة المركزية للبيانات المبوبة في الجداول المنتظمة :

(أ) الرُبيع الادنى (الاول) : يعبر الربيع الادنى (الاول) Q1 عن القيمة التي يكون قبلها عدد المشاهدات ربع العدد الكلي للمشاهدات والمشاهدات بعده تمثل ثلاث ارباع العدد الكلي للمشاهدات محل الدراسة .
لذا يتم حسابه كما في حالة الوسيط مع اختلاف ان قانون ايجاد ترتيب الربيع الاول Q1 هو كالتالي:

قانون ايجاد ترتيب الربيع الادنى الاول Q1	قانون ايجاد قيمته
$\frac{n}{4}$	$Q_1 = L_{Q_1} + \frac{\frac{n}{4} - F_a}{F_b - F_a} \times I_{Q_1}$

(ب) الربيع الاعلى (الثالث) :يعبر الربيع الثالث Q3 عن القيمة التي يكون قبها عدد المشاهدات ثلاثة ارباع العدد الكلي للمشاهدات والمشاهدات بعده تمثل ربع العدد الكلي للمشاهدات محل الدراسة . لذا يتم حسابه كما هو في حالة الوسيط مع اختلاف ان قانون ايجاد ترتيب الربيع الاول Q3 هو كالتالي:

قانون ايجاد ترتيب الربيع الادنى الاول Q3	قانون ايجاد قيمته
$\frac{3n}{4}$	$Q_3 = L_{Q_3} + \frac{\frac{3(n)}{4} - F_a}{F_b - F_a} \times I_{Q_3}$

لذا ومما سبق يتضح لنا انه يمكن ايجاد كلا من الربيع الادنى (الاول) Q1 والربيع الاعلى (الثالث)Q3 بنفس خطوات حساب الوسيط الا ان الامر المختلف هنا هو الترتيب كما استنتجنا يكون كالتالي :

المقياس	Q1	Q3
قانون ايجاد الترتيب	$k_{Q1} = \frac{n}{4}$	$k_{Q3} = \frac{3n}{4}$

⇐ مثال :

البيانات التالية توضح توزيع مجموعة من المدرسين العاملين في مجال التربية وفقا لفئات أعمارهم فكانت النتائج كما يلي:

فئات العمر	20-	30-	40-	60-50
عدد العمال	10	30	50	20

◄ المطلوب: احسب كل من الربيع الاول والربيع الثالث :

✓ طريقة الحل :

1. سواء كان الربيع الاول او الثالث نحتاج الى الجدول المتجمع الصاعد ، لذا يجب علينا اعداده كالتالي :

الحدود العليا للفئات	التكرار المتجمع الصاعد
اقل من 20	0
اقل من 30	10
اقل من 40	40
اقل من 50	90
اقل من 60	110

2. الان نبدأ بحساب الطلوب منا في السؤال :

i. حساب الربيع الادنى (الاول) $Q1$:

- ايجاد ترتيب الربيع الاول : $k_{Q1} = \frac{n}{4} = \frac{110}{4} = 27.5$
- ايجاد قيمة الربيع الادنى (الاول) Q1 : نلاحظ ان ترتيب الربيع الادنى هو [27.5] مما يعني ان الربيع الادنى يقع بين التكرار المتجمع الصاعد [$f_a(10)$] و هو المقابل للحد الاعلى للفئة 30 والتكرار المتجمع الصاعد [$f_b(40)$] وهو المقابل للحد الاعلى للفئة 40 والحد الادنى للفئة هو [$L_{Q1} = 30$]

وبالتالي يكون طول فئة الربيع الادنى (الاول) Q1 : $I = 40 - 30 = 10$ (نطرح الفئات وليس التكرارات)

والان نستطيع حساب قيمة الربيع الادنى (الاول) Q1 كما يلي :

$$Q1 = 30 + \frac{27.5 - 10}{40 - 10} \times 10 = 35.8333$$

ii. حساب الربيع الاعلى (الثالث)$Q3$:

- ايجاد ترتيب الريع الاعلى (الثالث) Q3 $k_{Q3} = \frac{3n}{4} = \frac{3(110)}{4} = 82.5$
- ايجاد قيمة الربيع الاعلى (الثالث)Q3 : ملاحظ ان ترتيب الربيع الاعلى الثالث [82.5] مما يعني ان الربيع الاعلى الثالث يقع بين التكرار المتجمع الصاعد [$f_a(40)$] (يخضع للفئة : اقل من 40) ، و التكرار المتجمع الصاعد [$f_b(90)$] (يخضع للفئة : اقل من 50) ، اذا فالحد الادنى للفئة يصبح [$L_{Q3} = 40$] .

وبالتالي يكون طول فئة الربيع الاعلى (الثالث) Q3 : $I = 40 - 50 = 10$ (نطرح الفئات وليس التكرارات)

وبناءا على هذه المعطيات فيمكننا حساب قيمة الربيع الاعلى (الثالث) Q3 بالتعويض في القانون الخاص بذلك :

$$Q3 = 40 + \frac{82.5 - 40}{90 - 40} \times 10 = 48.5$$

iii. ايجاد قيمة العشير $P_{0.10}$

- بداية نعرف العشير : وهو القيمة التي يكون قبلها 10 ٪ من مفردات المجتمع و 90 ٪ منها اكبر منه ، والاختلاف يكون فقط في الترتيب حيث ان ترتيب العشير هو $k_{P_{0.10}} = \frac{n}{10}$ ، لذا فنستطيع القول انه يمكن الحصول على العشير بنفس الطرق السابقة .
- ففي مثالنا هذا يكون ترتيب العشير $P_{0.10}$ هو $k_{P_{0.10}} = \frac{110}{10} = 11$
- قيمة العشير : نلاحظ ان ترتيب العشير $P_{0.10}$ هو [11] مما يعني انه يقع بين التكرار الصاعد [10] و [40] أي بين الفئتين (اقل من 30 ، و اقل من 40) وبناءا عليه فان الحد الادنى للفئة هو [$L_{P_{010}} = 30$] ، و الحد الادنى [$f_a(10)$] ، و الحد الاعلى [$f_b(40)$]

وبالتالي يكون طول فئة العشير $P_{0.10}$: $I = 30 - 40 = 10$ (نطرح الفئات وليس التكرارات)

بناءا على هذه المعطيات فيمكننا حساب قيمة العشير $P_{0.10}$ بالتعويض في القانون الخاص بذلك :

$$P_{0.10} = 30 + \frac{11 - 30}{40 - 30} \times 10 = 30.333$$

iv. ايجاد قيمة المئوين $P_{0.01}$

- بداية نعرف المئوين $P_{0.01}$: وهو القيمة التي يكون قبلها 1 ٪ من مفردات المجتمع و 99٪ منها اكبر منه ، والفرق بينها وبين العشير هو الترتيب فقط فترتيب المئوين هو $k_{P_{0.01}} = \frac{n}{100}$
- ففي مثالنا هذا يكون ترتيب المئوين $P_{0.01}$ هو $k_{P_{0.01}} = \frac{110}{100} = 1.1$
- قيمة المئوين $P_{0.01}$: نلاحظ ان ترتيب المئوين $P_{0.01}$ هو [1.1] مما يعني انه يقع بين التكرار الصاعد [20] و [30] أي بين الفئتين (اقل من 20 ، و اقل من 30) وبناءا عليه فان الحد الادنى للفئة هو [$L_{P_{010}} = 20$] ، و الحد الادنى [$f_a(0)$] ، و الحد الاعلى [$f_b(10)$]

وبالتالي يكون طول فئة المئوين $P_{0.01}$: $I = 20 - 30 = 10$ (نطرح الفئات وليس التكرارات)

بناء على هذه المعطيات فيمكننا حساب قيمة المئوين $P_{0.01}$ بالتعويض في القانون الخاص بذلك :

$$P_{0.01} = 20 + \frac{1.1 - 0}{10 - 0} \times 10 = 21.1$$

ملاحظة : نلاحظ اننا نستخدم نفس الطريقة لاستخراج الربيع الاول والثالث و العشير والمئوين فقط الاختلاف في قانون استخراج ترتيب المقياس المراد ايجاد ترتيبه

وعلى ذلك نكون حصلنا على مقاييس النزعة المركزية التي تصف تركيز البيانات عند اي نسبة من مفردات البيانات محل الدراسة في حالة البيانات المبوبة والتي كانت كما يلي :

المقياس	$P_{0.10}$	$P_{0.01}$	$Q1$	Med	$Q3$
القيمة	21.1	30.333	35.8333	43	48.5

مقاييس التشتت للبيانات المبوبة في الجداول المنتظمة :

(أ) نصف المدى الرُبيعي *Inter Quartile Range*: بسبب العيب الموجود فى مقياس التشتت (المدى) وتأثره بالقيم الشاذة أدى ذلك للجوء إلى مقياس آخر يسمى (نصف المدى الربيعي IQR) والذى يستبعد القيم المتطرفة من الطرفين. حيث يعتمد فى حسابه على كلا من الرُبيع الأول Q1 والرُبيع الثالث Q3 ويتم حسابه من خلال المعادلة التالية:

$$IQR = \frac{Q3 - Q1}{2}$$

⇐ **مثال :**

(باستخدام بيانات جدول الاعمار المذكور في المثال السابق)

◄ **المطلوب:** احسب نصف المدى الربيعي:

✓ **طريقة الحل :**

تطبيق القانون بعد ايجاد الربيع الادنى الاول والثالث كما في المثال السابق فلا نستطيع ايجاد المدى الربيعي من دون ايجادهما :

$$IQR = \frac{48.5 - 35.8333}{2} = 6.33335$$

ثالثا : المنوال :

المنوال كما سبق تعريفه في الفصل السابق هو القيمة الاكثر شيوعا او تكراراً ، وفي حالة البيانات المبوبة يمكن حسابه باستخدام المعادلة التالية :

$$Mod = L_{Mod} + \frac{D1}{D1 + D2} \times I$$

حيث ان (تعريف المصطلحات) :

- قيمة المنوال Mod
- الحد الادنى لفئة المنوال L_{Mod}
- يساوي تكرار فئة المنوال – تكرار الفئة السابقة $D1$
- يساوي تكرار فئة المنوال – تكرار الفئة اللاحقة $D2$
- طول الفئة المنوالية I ونحصل عليها بطرح الحد الادنى للفئة المنوالية من الحد الاعلى للفئة المنوالية

⇐ **مثال :**

(باستخدام بيانات جدول الاعمار المذكور في المثال السابق)

◄ **المطلوب:** احسب المنوال لاعمار مجموعة المدرسين العاملين في مجال التربية في المثال السابق؟

✓ **طريقة الحل :**

بيانات توزيع العمال وفقا لفئات العمر كانت كما يلي :

الحد الادنى L_{Mod}

فئات العمر	-20	-30	-40	60-50
عدد العمال	10	30	50	20

D1=30

الفئة المنوالية

D2=20

نلاحظ ان اكبر تكرار هو 50 ويكون مقابل للفئة [$40 - 50$] لذلك يطلق عليها الفئة المنوالية ومن ثم فان الحد الادنى لها هو [$L_{Mod} = 40$] وطول الفئة هو [$I = 10$]

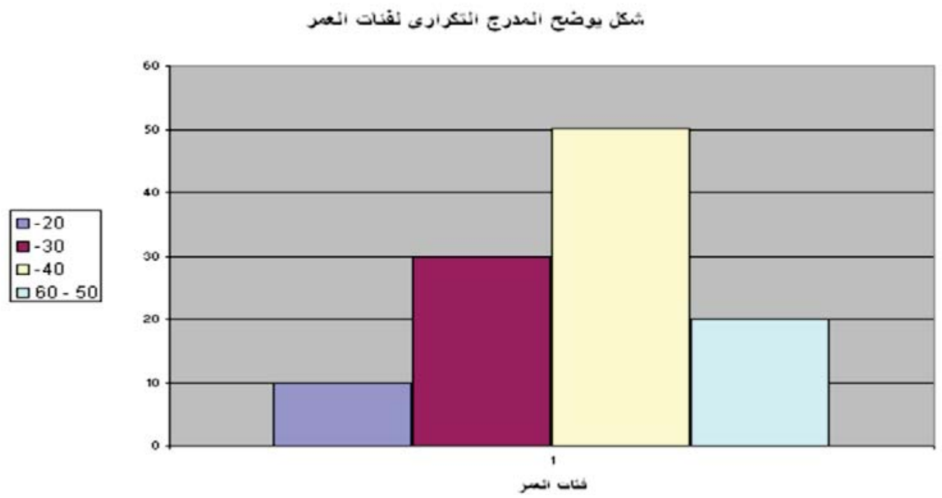
كما يمكن حساب كلا من D1 و D2 كالتالي :

D1= 50-30 =20

D2= 50-20 =30

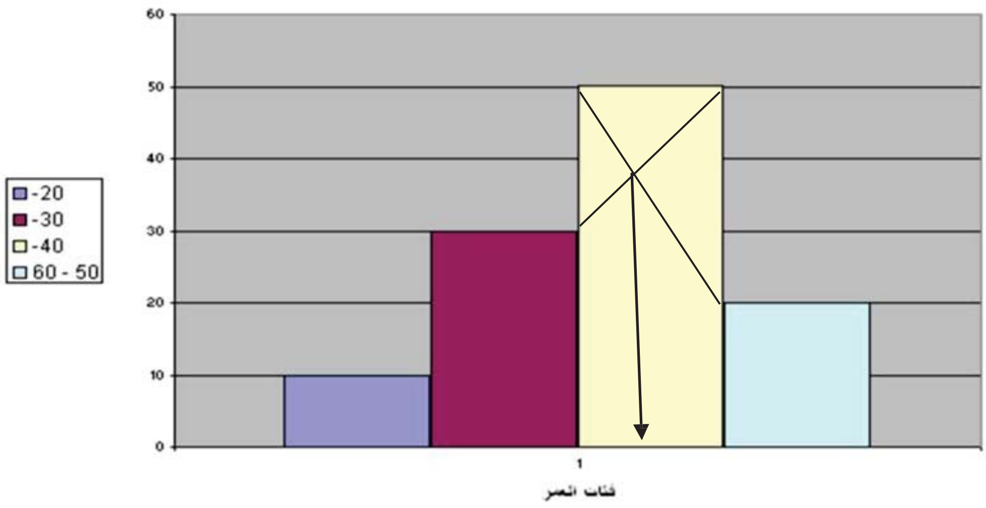
وبالتالي يمكن حساب قيمة المنوال Mod كالتالي : $Mod = 40 + \frac{20}{20+30} \times 10 = 44$

كما يمكن ايجاد المنوال بيانيا ، و يتم ذلك عن طريق رسم المدرج التكراري كما يلي :



وبعد رسم المدرج التكراري نأتي على اعلى مستطيل الذي يمثل اكبر تكرار ونصل بداية الفئة المنوالية ببداية الفئة التالية ونهاية الفئة المنوالية بنهاية الفئة السابقة عليها فيتقاطع الخطان عند نقطة نسقط منها عامودا على المحور الافقي فيلتقي عند نقطة تكون هي قيمة المنوال كما يتضح من الشكل التالي :

شكل يوضح المدرج التكراري لفئات العمر



2- **الجداول غير المنتظمة :** وهي الجداول التي يكون فيها اطوال افئات غير متساوية ويكفي وجود فئة واحده فقط طولها غير متساوي مع باقي الفئات لجعل الجدول غير منتظم .

ويتم حساب المقاييس الاحصائية التي سبق عرضها في حالة الجداول المنتظمة بنفس الطريقة **ما عدا المنوال** .

ويتعين علينا عند حساب المنوال تعديل التكرارات قبل حسابه وكذلك قبل رسم المدرج التكراري وذلك لان حجم التكرارات في تلك الحالة قد يسبب اتساع او ضيق في اعمدة فئات التوزيع ولذلك يتم التخلص من تأثير طول الفئة بإيجاد التكرار المعدل ، ويتم ذلك من خلال المعادلة التالية :

التكرار المعدل = التكرار الاصلي للفئة ÷ طول الفئة

⇐ **مثال :**

البيانات التالية توضح توزيع مجموعة من الموظفين وفقا لفئات دخلهم الشهري بالالف ريال فكانت كما يلي :

فئات الدخل	3-	5-	8-	15-10
عدد الموظفين	20	50	15	15

المطلوب: المطلوب حساب التالي :

1- الوسط الحسابي
2- متوسط الانحرافات المطلقة
3- التباين
4- الانحراف المعياري
5- الوسيط
6- الربيع الاول
7- الربيع الثالث
8- العشير
9- المئوين
10- نصف المدى الربيعي
11- المنوال

✓ **طريقة الحل :**

يمكن حساب المطلوب من 1 الى 10 بنفس طريقة حسابها في حالة الجداول المنتظمة بدون اي تعديل ، اما المطلوب رقم 11 فيطلب حساب **المنوال** ، وهو الذي طريقته تحتاج الى تعديل في الحساب في حالة الجداول غير المنتظمة ، ويتم ذلك وفق التالي :

لحساب المنوال في هذه الحالة لايتم الاعتماد على بيانات الفئات الاصلية وانما يتم ايجاد التكرار المعدل بقمة تكرار كل فئة على طولها كما يلي :

فئات الدخل	التكرار f	طول الفئة (حد الفئة الاعلى – حد الفئة الادنى)	التكرار المعدل التكرار ÷ طول الفئة
3-	20	5 – 3 = 2	10
5-	50	8 – 5 = 3	16.6667
8-	15	2	7.5
15-10	15	5	3
المجموع	100		37.1667

نلاحظ ان اكبر تكرار معدل هو 16.6667 ويكون مقابل للفئة [5 – 8] لذلك يطلق عليها الفئة المنوالية ومن ثم فا الحد الادنى لها هو هو [$L_{Mod} = 5$] وطول الفئة هو [$I = 3$]

والان يمكننا حساب كلا من D1 و D2 (راجع تعريف المصطلحات في الصفحة 66) كالتالي:

D1 = 16.66667 – 10 = 6.66667

D2 = 16.66667 – 7.5 = 9.16667

وبالتالي يمكن حساب قيمة المنوال Mod :

$$Mod = 5 + \frac{6.66667}{6.66667 + 9.16667} \times 3 = 6.263158$$

3- **الجداول المفتوحة :** وهي الجداول التي يكون فيها الحد الادنى للفئة الاولى غير محدد او الحد الاعلى للفئة الاخيرة او كلاهما ، وفي هذا النوع من الجداول يصعب حساب الوسط الحسابي او التباين او الانحراف المعياري حيث لايمكن تحديد مركز الفئة للفئات المفتوحة ، لذا فيعتبر من انسب المقاييس الاحصائية في هذه الحالة هي المقاييس الوسيطية والتي يقصد بها الوسيط والربيع الادنى والربيع الاعلى والعشير و المئوين وكذلك لقياس التشتت يتم من خلال نصف المدى الربيعي .

⇐ **مثال :**

البيانات التالية تعبر عن اوزان مجموعة من الطلاب بالكيلو جرام في المرحلة الجامعية فكانت كما يلي :

فئات الوزن	اقل من 50	50-	60-	70-	80 فاكثر
عدد الطلاب	3	10	35	15	10

جدول مفتوح لانه لم يختم الفئات بقيمة وانما تركها الى مالا نهاية (فاكثر)

◄ **المطلوب:** حساب مقاييس النزعة المركزية والتشتت المناسبة ؟

✓ **طريقة الحل :**

اولا : يتم اعداد الجدول التكراري المتجمع الصاعد كما يلي :

الحدود العليا للفئات	التكرار المتجمع الصاعد
اقل من 50	5
اقل من 60	15
اقل من 70	50
اقل من 80	65
اقل من ∞	75

ثانيا : ايجاد الرتبة :

المقياس	Q1	Med	Q3
الرتبة	$k_{Q1} = \frac{n}{4} = \frac{75}{4} = 18.75$	$k_{Med} = \frac{n}{2} = \frac{75}{2} = 37.5$	$k_{Q3} = \frac{3n}{4} = \frac{3(75)}{4} = 56.25$

ثالثا : ايجاد القيمة وذلك بالتعويض بالقوانين المعروفة لدينا :

المقياس	القانون الخاص به	التعويض والنتيجة
الوسيط	$Med = L_{Med} + \frac{k_{Med} - F_a}{F_b - F_a} \times I$	$= 60 + \frac{37.5 - 15}{50 - 15} \times 10 = 66.4285$
الربيع الادنى	$Q1 = L_{Q1} + \frac{k_{Q1} - F_a}{F_b - F_a} \times I$	$= 60 + \frac{18.75 - 15}{50 - 15} \times 10 = 61.071$
الربيع الاعلى	$Q3 = L_{Q3} + \frac{k_{Q3} - F_a}{F_b - F_a} \times I$	$= 70 + \frac{56.25 - 50}{65 - 50} \times 10 = 74.1667$
نصف المدى الربيعي	$IQR = \frac{Q3 - Q2}{2}$	$= \frac{74.1667 - 61.071}{2} = 6.5478$

(المحاضرة التاسعة – الجزء الاول)

مقاييس التشتت النسبي والالتواء والتفلطح

❖ تمهيد :

هناك مقاييس أخرى لابد من دراستها غير تلك التي تم التعرض لها في المحاضرات السابقة لمساعدة الباحث فى الحكم على البيانات محل التحليل والدراسة من حيث درجة التشتت والمقارنة فيما بينها وكذلك مقاييس التوزيع والتى تتمثل فى دراسة الإلتواء والتفلطح للمنحنيات التكرارية لتوزيعات المتغيرات المختلفة .

لذا سيتم في هذا الفصل دراسة كل من التالي :

- مقاييس التشتت النسبي
- القيمة المعيارية
- الالتواء
- التفلطح

اولا – مقاييس التشتت النسبي Coefficient of Variation

يستخدم هذا النوع من المقاييس لمقارنة تشتت مجموعتين من البيانات او ظاهرتين توزيعيتين ، وفي تلك الحالة لا يصلح مقارنة التباين الانحراف المعياري لكلا المجموعتين حيث يكون لها وحدات قياس تختلف على حسب طبيعة الظاهرة محل الدراسة ، لذا في حالة الرغبة في المقارنة بين التشتت لظاهرتين او اكثر فانه يتم الاعتماد في عملية المقارنة على مقاييس التشتت النسبي Coefficient of Variation (c. v.) والتي يعبر عنها من خلال معامل الاختلاف المعياري والذي يمكن حسابه بالاعتماد على كلاً من الوسط الحسابي و الانحراف المعياري حيث ان :

معامل الاختلاف = الانحراف المعياري ÷ الوسط الحسابي

مع الاخذ بالملاحظات :

1) في حال الاعتماد على بيانات العينة يتم حساب معامل الاختلاف النسبي من خلال المعادلة التالية (وهذه هي الحالة المتكررة في مقررنا) :

$$c.v.=\frac{S}{\bar{x}}=\times 100$$

طبعا في الالات الحاسبة تكتب هذه المتغيرات هكذا
sx , σx

2) اي انه في حال الاعتماد على بيانات المجتمع يتم حساب معامل الاختلاف النسبي (ونادرا ما نستخدم بيانات المجتمع في مقررنا) :

$$c.v.=\frac{\sigma}{\bar{x}}=\times 100$$

3) اما اذا كانت البيانات المتاحه من جداول تكرارية (بيانات مبوبة) فيمكن الاعتماد على معامل الاختلاف الربيعي المعياري (وليس النسبي) والذي يعتمد في حسابه على الوسيط والربيع الادنى والربيع الاعلى وخاصة في حالة الجداول المفتوحة ، حيث ان :

$$c.v.=\frac{Q_3-Q_1}{Q_3+Q_1}=\times 100$$

ملاحظة : قام الدكتور في المحتوى بالتذكير بطريقة احتساب الوسيط Med و الربيع الادنى والربيع الاعلى (يرجى مراجعة صفحة 61 و 64 لتذكرها)

⇐ مثال :

البيانات التالية تعبر عن توزيع الوحدات السكنية حسب الايجار السنوي بأحد الاحياء :

الايجار بالالف الريال	-6	-10	-12	18-14
عدد الوحدات السكنية	15	20	12	13

بناءا على اخر ملاحظة في نهاية المثال فان هذا الجدول جدول مغلق لانه ختم اخر الفئات بقيمة فيفضل معه استخدام المعادلة في البند 2

المطلوب: حساب التالي :

1) معامل الاختلاف للايجار السنوي

2) معامل الاختلاف الربيعي للايجار السنوي

✓ طريقة الحل :

1- حساب معامل الاختلاف للايجار السنوي:

حتى يمكن حساب معامل الاختلاف لابد من حساب كلا من:

- الوسط الحسابي
- الإنحراف المعياري
- وأخيرا معامل الإختلاف

لذا يتم اعداد الجدول التالي حتى يمكن حساب كلا من الوسط الحسابي والتباين والانحراف المعياري :

فئات الايجار	التكرار f	مركز الفئة x	xf	$\bar{x}$	$x-\bar{x}$	$(x-\bar{x})^2$	$f(x-\bar{x})^2$
6 -	15	8	120	11.73	-3.73	13.91	208.69
10 -	20	11	220	11.73	-0.73	0.533	10.66
12 -	12	13	156	11.73	1.27	1.61	19.32
14 – 18	13	16	208	11.73	4.27	18.23	236.99
المجموع	$\sum f = 60$		$\sum xf = 704$				$\sum f(x-\bar{x})^2 = 475.66$

مادام حدد فئات (من كذا الى كذا) معناها بيانات مجتمع وناخذ σx في الالة الحاسبة

$\sigma x \approx 2.816$,

$\sigma^2 \approx 7.93$

- الوسط الحسابي :

يمكن ايجاد الوسط الحسابي بالقانون التالي او بالالة الحاسبة :

$$\bar{x} = \frac{704}{60} \approx 11.733$$

- التباين :

يمكن ايجاد التباين بالقانون التالي او بالالة الحاسبة : $\sigma^2 = \frac{475.66}{60} \approx 7.93$

- الانحراف المعياري :

يمكن ايجاد الانحراف المعياري بالقانون التالي او بالالة الحاسبة :

$$\sigma = \sqrt{\sigma^2} \approx 2.816$$

والان نستطيع حساب معامل الاختلاف المعياري للمجتمع طبقا للقانون $\boxed{c.v. = \frac{\sigma}{\bar{x}} = \times 100}$

$$\boxed{c.v. = \frac{2.816}{11.733} = \times 100 = 24\%}$$

اي ان معامل الاختلاف للايجار السنوي للوحدات السكنية بلغ 24 ٪

2- حساب معامل الاختلاف الربيعي:

حتى يمكن حساب معامل الاختلاف الربيعي لابد من حساب كلا من:

- الوسيط
- الربيع الاعلى
- الربيع الادنى

لذا يتم اعداد الجدول التكراري المتجمع الصاعد التالي:

اولا : يتم اعداد الجدول التكراري المتجمع الصاعد كما يلي :

الحدود العليا للفئات	التكرار المتجمع الصاعد
اقل من 6	صفر
اقل من 10	15
اقل من 12	35
اقل من 14	47
اقل من 18	60

ثانيا : ايجاد الرتبة :

المقياس	Q1	Med	Q3
الرتبة	$k_{Q1}=\frac{n}{4}=\frac{60}{4}=15$	$k_{Med}=\frac{n}{2}=\frac{60}{2}=30$	$k_{Q3}=\frac{3n}{4}=\frac{3(60)}{4}=45$

ثالثا : ايجاد القيمة وذلك بالتعويض بالقوانين المعروفة لدينا :

المقياس	القانون الخاص به	التعويض والنتيجة
الوسيط	$Med=L_{Med}+\frac{k_{Med}-F_a}{F_b-F_a}\times I$	$=10+\frac{30-15}{35-15}\times 2=11.5$
الربيع الادنى	$Q1=L_{Q1}+\frac{k_{Q1}-F_a}{F_b-F_a}\times I$	سبب تجاهل القانون ان رتبة الربيع 15 ويوجد متجمع صاعد يساوي 15 بالضبط اما الفئة 10 لهذا اخترنا $=10$
الربيع الاعلى	$Q3=L_{Q3}+\frac{k_{Q3}-F_a}{F_b-F_a}\times I$	$=12+\frac{45-35}{47-35}\times 12=13.6667$
نصف المدى الربيعي	$IQR=\frac{Q3-Q2}{2}$	$=\frac{13.6667-10}{2}=1.8333$

والان نستطيع حساب معامل الاختلاف الربيعي للجدول التكراري طبقا للقانون $c.v.=\frac{Q_3-Q_1}{Q_3+Q_1}=\times 100$

$$c.v.=\frac{13.6667-10}{13.6667+10}=\times 100=15.493\%$$

اي ان معامل الاختلاف الربيعي للايجار السنوي للوحدات السكنية بلغ 15.493 ٪

ونلاحظ وجود أختلاف بين قيمتى معامل الاختلاف بأستخدام كلا من المعادلة الأولى والثانية وذلك لأختلاف الأساس الرياضى فى كل من التعريفين المعادلتين. الا أنة يفضل استخدام المعادلة الثانية فى حالة الجداول التكرارية المفتوحة أما غير ذلك فيفضل استخدام المعادلة الأولى.

⇐ مثال : (هذا المثال نذاله يتوجب عليك حفظ القوانين لان الالة مارح تفيدك)

في دراسة اجريت على اطوال و اوزان مجموعة من الطلاب فكانت النتائج كما يلي :

1- اطوال الطلاب : $\sum x=3034$ و $\sum x^2=463040$

2- اوزان الطلاب : تم تلخيص البيانات المتاحة عن الطلاب في صورة الجدول التكراري التالي :

فئات الوزن	50-	60-	70-	80-90
عدد الطلاب	3	5	8	4

◄ المطلوب: قارن بين تشتت اطوال الطلاب و اوزانهم ؟

✓ طريقة الحل :

للمقارنة بين تشتت اطوال اطلاب واوزانهم يتم حساب معاملا الاختلاف لكلاً منهما ، لذا يلزم الامر حساب كلا من الوسط الحسابي والانحراف المعياري كما يلي :

1- اطوال الطلاب :

نلاحظ ان المتاع لدينا من البيانات هو $\sum x = 3034$ و $\sum x^2 = 463040$ الا ان اعداد الطلاب اعدد الطلاب لمتغير الوزن لذا يكون عدد الطلاب محل الدراسة هو 20 طالب .

- يمكن حساب الوسط الحسابي كما يلي :

$$\bar{x} = \frac{\sum x}{n} = \frac{3034}{20} = 151.7$$

- ويمكن حساب التباين كما يلي : (صفحة 55 : بيانات من المجتمع)

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - n(\bar{x}\,)^2}{n-1} = \frac{463040 - 20(151.7)^2}{19} = 146.43$$

- وبالتالي يكون الانحراف المعياري كما يلي :

$$s = \sqrt{s^2} = \sqrt{146.43} = 12.1008916$$

- وعلى ذلك يكون معامل الاختلاف المعياري للمجتمع (صفحة 69 بند 2) لاوزان الطلاب هو :

$$c.v. = \frac{12.1008916}{151.7} \times 100 = 7.9768\%$$

2- اوزان الطلاب :

نلاحظ ان البيانات المتاحة لدينا هي جدول لتوزيع تكراري لاوزان الطلاب لذلك يكمن حساب الوسط الحسابي والانحراف المعياري من خلال تكوين الجدول التالي :

فئات اوزان الطلاب	التكرار f	مركز الفئة x	xf	$x^2 f$
50-	3	55	165	9075
60-	5	65	325	21125
70-	8	75	600	45000
80-90	4	85	340	28900
المجموع	$\sum f = 20$		$\sum xf = 1430$	$\sum x^2 f = 104100$

- يمكن حساب الوسط الحسابي كما يلي :

$$\bar{x} = \frac{\sum_{i=1}^{l} x_i f_i}{\sum_{i=1}^{l} f_i} = \frac{1430}{20} = 71.5$$

- ويمكن حساب التباين كما يلي : (صفحة 55 : بيانات من المجتمع)

$$\sigma^2 = \frac{\sum_{i=1}^{l} x_i^2 f_i}{\sum_{i=1}^{l} f_i} - \overline{x^2} = \frac{104100}{20} - (71.5\,)^2 = 92.75$$

- وبالتالي يكون الانحراف المعياري كما يلي :

$$\sigma = \sqrt{\sigma^2} = \sqrt{92.75} = 9.6306$$

- وعلى ذلك يكون معامل الاختلاف المعياري للمجتمع (صفحة 69 بند 2) لاوزان الطلاب هو :

$$c.v. = \frac{9.6306}{71.5} \times 100 = 13.469\%$$

ويتضح لنا ان معامل الاختلاف لاوزان الطلاب هو 13.46948 ٪ ، وبذلك يكون وزن الطلاب اكثر تشتتا من اطوالهم لان قيمة معامل الاختلاف لاوزان الطلاب اكبر منه في اطوالهم .

ثانيا – القيمة المعيارية Standardized values

هي تلك القيمة التي تقيس مدى انحراف قيمة مفردة ما من مفردات الدراسة عن الوسط الحسابي لها وذلك بوحدات من الانحراف المعياري ، ويشار الى المتغير الذي يعبر عن القيم المعيارية بالمتغير المعياري ويرمز للقيمة المعيارية بالرمز Z حيث انه :

$$z = \frac{x - \bar{x}}{s}$$

تذكر s تعني الانحراف المعياري

وبالتالي يمكن الاعتماد على القيمة المعيارية في المقارنة بين القيم المطلقة للظواهر المختلفة .

⇐ مثال :

حصل احد الطلاب في مقرر المحاسبة على 80 درجة حيث بلغ متوسط درجات الطلاب في اختبار المحاسبة 83 درجة بانحراف معياري قدره 5 . بينما حصل في اختبار الرياضيات على 60 درجة وبلغ متوسط درجات الطلاب في الختبار الرياضيات 65 درجة بانحراف معياري قدره 5 درجات .

◄ المطلوب: هل يمكن القول بان درجات الطالب في مقرر المحاسبة افضل من درجته في مقرر الرياضيات ؟

✓ طريقة الحل :

للحكم على مدى افضلي الدرجة التي حصل عليها الطالب في اي من المقررين يجب حساب القيمة المعيارية لكل منهما كما يلي :

- القيمة المعيارية لدرجة الطالب في مقرر المحاسبة هي :

$$z = \frac{x - \bar{x}}{s} = \frac{80 - 83}{5} = -0.6$$

- القيمة المعيارية لدرجة الطالب في مقرر الرياضيات هي :

$$z = \frac{x - \bar{x}}{s} = \frac{70 - 65}{5} = 1$$

يتضح لنا مما سبق ان القيمة المعيارية لدرجة الطالب في مقرر الرياضيات هي (+1) مما يعني ان الدرجة التي حصل عليها الطالب اكبر من متوسط درجات الطلاب ، بينما بلغت القيمة المعيارية للدرجة التي حصل عليها الطالب في مقرر المحاسبة (-0.6) مما يدل على ان الدرجة التي حصل عليها الطالب اقل من متوسط الدرجات التي حصل عليها الطلاب .

ويدل ذلك على انه من الناحية الظاهرية قد تبدو درجة الطالب في مقرر المحاسبة افضل الا انه في حقيقة الامر ان مستوى الطالب في مقرر الرياضيات هو الافضل .

(المحاضرة التاسعة – الجزء الثاني)

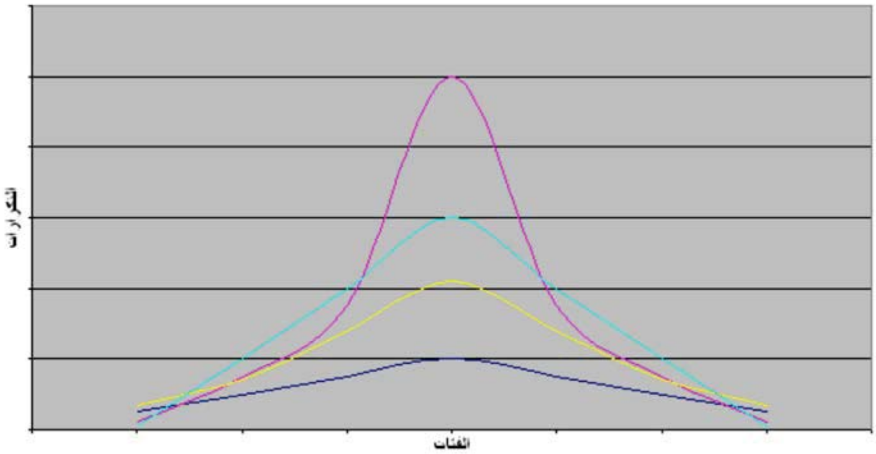
تابع ..... مقاييس التشتت النسبي والالتواء والتفلطح

ثالثا – مقاييس الإلتواء Skewdness Measures

عند دراسة أشكال منحنيات التوزيعات التكرارية المختلفة نجد أن منها ماهو متماثل Symmertical ومنها الغير متماثل أى يوجد به ما يسمى بالإلتواء Skewed كما يتضح من أشكال منحنيات التوزيعات التالية:

1- **المنحنى المتماثل SYMMETRICAL CURVE :** هو المنحنى الذى اذا قسمناه إلى نصفين إنطبق هذان النصفان على بعضهما البعض تماما

شكل يوضح منحنيات التوزيع المتماثل



ونلاحظ فيما يخص مقاييس التشتت النسبى Coefficient of Variation بالنسبة للمنحنيات المتماثلة فى الشكل السابق تختلف قممها ارتفاعا أو تفلطحآ و تدببا حسب حجم التكرارات على جانبى القمة. ويتميز التوزيع المتماثل بأن:

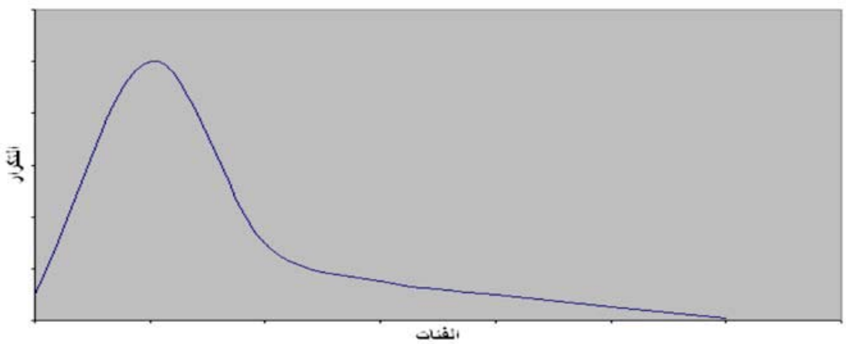
الوسط الحسابى = الوسيط = المنوا ل

اي ان الوسط الحسابي يساوي الوسيط والمنوال

2- **المنحنيات الملتوية SKEWED :** إن الكثير من التوزيعات الإحصائية تبتعد عن التماثل بتركز تكراراتها عند احد الجهات ونستطيع حصرها في الانواع التالية :

أ) توزيع ملتوي جهة اليمين (الإلتواء الموجب): وفيه تركز التوزيعه الاحصائية تكرارتها عند اصغر القيم فيصبح المنحنى ملتويا جهة اليمين او التواء موجب كما هو موضع في الشكل التالي:

شكل يوضح منحنى ملتوى جهة اليمين

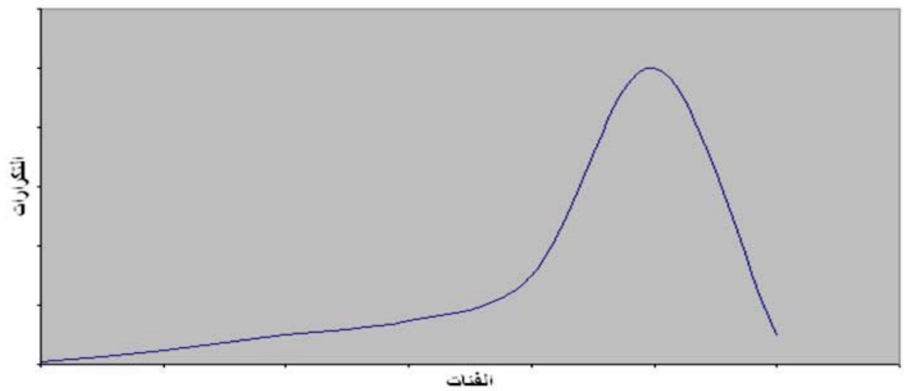


ويكون فيه : **الوسط الحسابي > الوسيط > المنوال**

اي ان الوسط الحسابي اكبر من الوسيط و الوسيط اكبر من المنوال

ب) توزيع متلوي جهة اليسار(إلتواء سالب) : وفيه تركز التوزيعه الاحصائية تكرارتها عند اكبر القيم فيصبح المنحنى ملتويا جهة اليسار او التواء سالب كما هو موضع في الشكل التالي:

شكل يوضح منحنى ملتوى جهة اليسار



ويكون فيه : المنوال > الوسيط > الوسط الحسابي

اي ان المنوال اكبر من الوسيط و الوسيط اكبر من الوسط الحسابي

ويمكن قياس الالتواء من خلال معامل الالتواء SK والذي يفيدنا في الحكم على مدى تماثل او التواء التوزيع حيث يكون متماثل اذا كان يساوي صفر او يكون التواء موجب اذا كانت قيمته موجبه او سالب اذا كانت قيمته سالبه .

الا ان بعض الاحيان يكون التوزيع قريب من التماثل في حالة ما تقترب قيمة معامل الالتواء من الصفر ، وتتعدد مقاييس الالتواء الا ان من اهمها :

I. معامل الالتواء لبيرسون :ذوالذي كون في احد الصورتين التاليتين حسب المتوفر لدينا (المنوال او الوسيط) :

1) معامل الالتواء $= \frac{\text{المنوال} - \text{الوسط الحسابي}}{\text{الانحراف المعياري}}$	$sK = \frac{\overline{x} - Mod}{S}$
2) معامل الالتواء $= \frac{3(\text{الوسيط} - \text{الوسط الحسابي})}{\text{الانحراف المعياري}}$	$sK = \frac{3(\overline{x} - Med)}{S}$

II. مقياس الالتواء لباولي SK_B : بما الا انه لايمكن حساب معامل الالتواء لبيرسون في حالة المنحنيات لتي تكون شديدة الالتواء او في حالة التوزيعات التكرارية المفتوحة ، لذلك يمكن الاعتماد على مقياس الالتواء لباولي SK_B والذي يعرف كما يلي :

$$SK_B = \frac{(Q_3 - Med\ ) - (Med - Q_1)}{(Q_3 - Med\ ) + (Med - Q_1)}$$

او يمكن اعادة صياغة معامل الالتواء لباولي SK_B على الصورة التالية :

$$SK_B = \frac{Q_3 - 2Med + Q_1}{Q_3 - Q_1}$$

ملاحظة : قام الدكتور في المحتوى بالتذكير بطريقة احتساب الوسيط Med و الربيع الادنى والربيع الاعلى (يرجى مراجعة صفحة 61 و 64 لتذكرها)

⇐ مثال :

البيانات التالية تعبر عن توزيع الوحدات السكنية حسب الايجار السنوي بأحد الاحياء :

الايجار بالالف الريال	6-	10-	12-	14-18
عدد الوحدات السكنية	15	20	12	13

◄ المطلوب: حساب معامل الالتواء لتوزيع الايجار السنوي للوحدات السكنية ؟ :

✓ طريقة الحل :

تم استخراج اغلب متطلبات حساب معامل الالتواء في صفحة 70 و 71 ولكن يبقى علينا حساب المنوال حتى تكون جميع المقاييس الاحصائية التي نحتاجها موجوده ، لذا يمكن حساب المنوال كما يلي :
نلاحظ ان اطوال الفئات للايجار السنوي غير متساوي لذا لحساب المنوال يلزم ايجاد التكرار المعدل و من ثم يتم اعداد الجدول التالي :

فئات الايجار	التكرار f	طول الفئة (حد الفئة الاعلى – حد الفئة الادنى)	التكرار المعدل التكرار ÷ طول الفئة
6 -	15	4	3.75
10 -	20	2	10
12 -	12	2	6
14 – 18	13	4	3.25
المجموع	$\sum f = 60$		23

ويمكن حساب المنوال بتطبيق المعادلة التالية كما سبق ان وضحنا (صفحة 66) بناء على القانون التالي :

$$Mod = L_{Mod} + \frac{D1}{D1 + D2} \times I$$

حيث ان:

$$D1 = 10 - 3.75 = 6.25 \qquad D2 = 10 - 6 = 4$$

$$L_{Mod} = 10 \qquad I = 2$$

وعلى ذلك يمكن حساب المنوال كما يلي :

$$Mod = 10 + \frac{6.25}{6.25 + 4} \times 2 = 11.21951$$

- وبناءا على ما سبق يمكن حساب معامل الالتواء لبيرسون باستخدام المعادلة $sK = \frac{\bar{x} - Mod}{S}$ كما يلي :

$$sK = \frac{\bar{x} - Mod}{S} = \frac{11.73333 - 11.2195122}{2.8158} = 0.18247$$

تفسير النتيجة : يعبر ذلك على وجود التواء موجب جهة اليمين الا ان قيمة معامل الالتواء صغيرة تقترب من الصفر مما يدل ايضا على ان التوزيع قريب من التماثل .

- كما يمكن تطبيق المعادلة $sK = \frac{3(\bar{x} - Med)}{S}$ ايضا لحساب معامل الالتواء كما يلي :

$$sK = \frac{3(\bar{x} - Med)}{S} = \frac{3(11.73333 - 11.5)}{2.8158} = 0.24859$$

تفسير النتيجة : يعبر ذلك ايضا على وجود التواء موجب جهة اليمين كما حددته النتيجة في المعادلة السابقة .

- وايضا يمكن تطبيـــــق معادلة معامـــــــــل الالتـواء لبـاولي SK_B طبقـا للقـانون التالي $SK_B = \frac{Q_3-2Med+Q_1}{Q_3-Q_1}$ وذلك بالشكل التالي :

$$SK_B = \frac{Q_3 - 2Med + Q_1}{Q_3 - Q_1} = \frac{13.6667 - 2(11.5) + 10}{13.6667 - 10} = 0.1816$$

تفسير النتيجة : يشير ايضا معامل الالتواء لباولي بوجود التواء موجب .

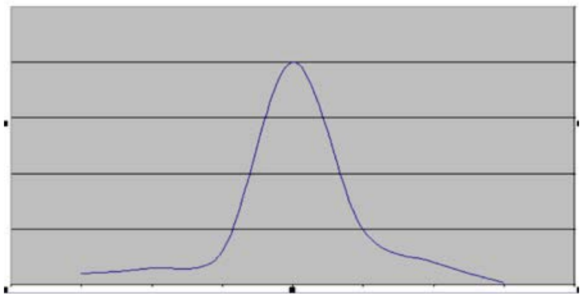
❖ خلاصة استخدام معامل الالتواء :

نتيجة لوجود اختلاف في الاصل الرياضي لكل من المعادلات الثلاث السابقة ، لذا نجد ان قيمة معامل الالتواء تختلف . الا انه كما سبق وذكرنا بأنه يفضل استخدام معامل الالتواء لبيرسون في اي من صيغتيه في حالة البيانات الغير مبوبة وكذلك الجداول التكرارية المغلقة ، اما في حالة الجداول التكرارية المفتوحة فيفضل استخدام معامل الالتواء لباولي .

رابعًا – التفلطح Kurtosis

يقصد بالتفلطح مقدار التدبب (الارتفاع او الانخفاض) في قمة المنحنى مقارنة بقمة منحنى التوزيع الطبيعي .
وتكون قيمة معامل التفلطح صفر في حالة التوزيع الطبيعي المعياري . لذا تقوم الكثير من البرامج الاحصائية بحساب معامل التفلطح للقيم المعيارية للبيانات ، فإذا كان الناتج :

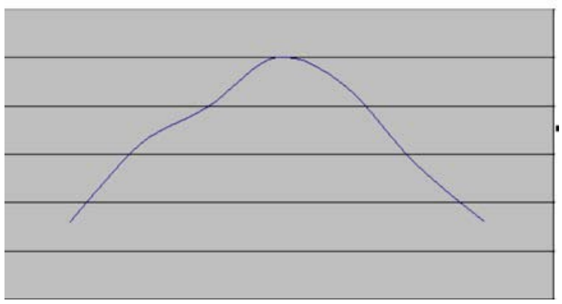
شكل يوضح المنحنى المدبب



1- موجب اي ان قيمة معامل التفلطح للبيانات الاصلي اكثر من 3 يكون المنحنى مدبب الى اعلى

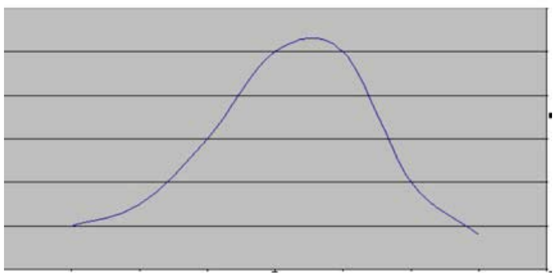
ويتضح لنا من الشكل ان هناك فئة معينة تتركز بها التكرارات والتي تجعل المنحنى مدبب الى اعلى

شكل يوضح المنحنى المفلطح



2- سالب اي ان قيمة معامل التفلطح للبيانات الاصلية اقل من 3 يكون المنحنى مفلطح او اكثر انبطاحا من قمة منحنى التوزيع الطبيعي .

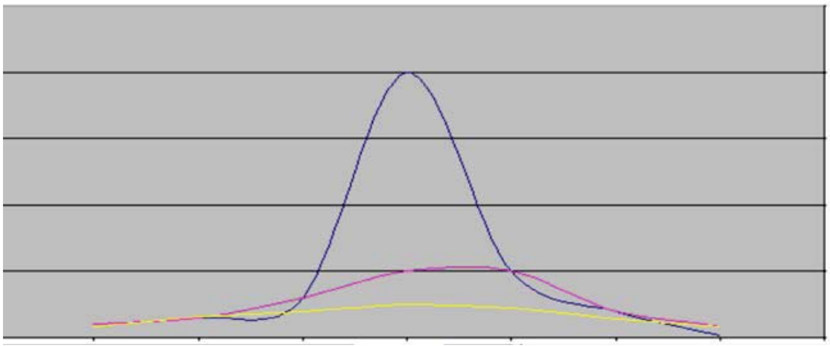
شكل يوضح المنحنى متوسط التفلطح



3- صفر أي ان قيمة معامل التفلطح للبيانات الاصلية تساوي 3 ويكون المنحنى متوسط التفلطح.

وحتى يتضح الفرق بين المنحنيات الثلاث يمكن رسمها معا كما يلي:

شكل يوضح المنحنيات الثلاث معاً المدبب و متوسط التفلطح و المفلطح



ويتم قياس معامل التفاطح KU باستخدام الربيعات والمئينيات بالقانون التالي :

$$KU = \frac{Q3 - Q1}{2(P_{0.90} - P_{0.10})}$$

حيث أن :

$P_{0.10}$	إلى المئين العاشر (العشير) والذى يعبر عن 10 ٪ من المفردات تكون أقل منه و 90٪ منها أكبر منه.
$P_{0.90}$	إلى المئين التسعين والذى يعبر عن 90 ٪ من المفردات تكون أقل منه و 10٪ منها أكبر منه.

⇐ مثال :

البيانات التالية تعبر عن توزيع الوحدات السكنية حسب الايجار السنوي بأحد الاحياء :

الايجار بالالف الريال	6-	10-	12-	14-18
عدد الوحدات السكنية	15	20	12	13

◀ المطلوب: حساب معامل التفلطح لتوزيع الإيجار السنوى للوحدات السكنية؟ :

✓ طريقة الحل :

تم سابقا حساب Q1 و Q3 ، ولكن يتبقى علينا حساب كلا من $P_{0.10}$ و $P_{0.90}$ بنفس طريقة حساب الوسيط والربيع الاعلى والادنى كما تم شرح ذلك من قبل بالطريقة التالية :

اولا : اعداد الجدول التكراري المتجمع الصاعدة كما يلي :

الحدود العليا للفئات	التكرار المتجمع الصاعد
أقل من 6	صفر
أقل من 10	15
أقل من 12	35
أقل من 14	47
أقل من 18	60

ثانيا :إيجاد الرتبة كالتالي:

المقياس	$P_{0.10}$	$P_{0.90}$
الرتبة	$K_{P_{0.10}} = \frac{n}{10} = \frac{60}{10} = 6$	$K_{P_{0.90}} = \frac{9n}{10} = \frac{9 \times 60}{10} = 54$

ثالثا: ايجاد القيم :

المقياس	قانون ايجاد القيمة	التعويض في القانون
$P_{0.10}$	$P_{0.10} = L_{P_{0.10}} + \frac{K_{P_{0.10}} - F_a}{F_b - F_a} \times I$	$P_{0.10} = 6 + \frac{6-0}{15-0} \times 4 = 7.6$
$P_{0.90}$	$P_{0.90} = L_{P_{0.90}} + \frac{K_{P_{0.90}} - F_a}{F_b - F_a} \times I$	$P_{0.10} = 14 + \frac{54-47}{60-47} \times 4 = 16.153$

رابعا: الربيع Q1 و Q3 : لقد قمنا بحسابهما مسبقا عندما ذكرنا نفس المثال في صفحة 71 لذا فنوردهما هنا مباشرة :

المقياس	الربيع الأول أو الأدنى Q1	الربيع الثالث أو الأعلى Q3
قيمته	10	13.667

خامسا : والان يمكننا حساب معامل التفلطح كالتالي بعد توفر جميع المتطلبات لدينا :

المقياس	قانون ايجاد القيمة	التعويض في القانون
معامل التفلطح KU	$KU = \frac{Q3 - Q1}{2(P_{0.90} - P_{0.10})}$	$KU = \frac{13.6667 - 10}{2(16.15385 - 7.6)} = 0.2143$

ويتضح لنا ان معامل التفلطح اقل من 3 مما يدل على ان المنحنى مفلطح ، أي ان المشاهدات (التكرارات) موزعة على الفئات المختلفة للايجار السنوي ولا يوجد تركز بدرجة كبيرة في احد الفئات على حساب باقي الفئات الاخرى .

(المحاضرة العاشرة)

تحليل الارتباط

❖ تمهيد :

بعد جمع البيانات وتبويبها ومحاولة وصف التطورات التي حدثت للظاهرة محل الدراسة خلال فترة زمنية معينة ، الا ان اشتقاق بعض العلاقات بين المتغيرات التي تكون الظاهرة محل الدراسة يكون من الاهمية بمكان لمعرفة تطورات الظاهرة في المستقبل وكيفية التأثير فيها من خلال التأثير في المتغيرات المكونة لها .

لذلك فمن الاساليب الاحصائية المناسبة لتقييم العلاقات بين المتغيرات هو تحليل الارتباط Correlation Analysis ، وكذلك نحتاج لعملية التنبؤ باداء الظاهرة في المستقبل الاعتماد على تحليل الانحدار Regression Analysis .

اولا : تحليل الارتباط CORRELATION ANALYSIS : يستخدم معامل الارتباط البسيط Correlation coefficient في تحديد ما اذا كان هناك علاقة بين المتغيرين ، وكذلك تحديد نوع وقوة العلاقة ان وجدت .

اما في حالة دراسة مدى وجود علاقة ارتباطيه بين اكثر من متغيرين فانه يتم الاعتماد على معامل الارتباط المتعدد Multiple Correlation coefficient.

اما في حالة وجود اكثر من متغير ويرغب الباحث في تثبيت تأثير احد المتغيرات كما في الدراسات الاقتصادية حيث يتم فيها دراسة تأثير السعر على الكمية المطلوبة بفرض ثبات الجودة ومستوى الذوق كما هو ، فهنا يتم الاعتماد في هذه الحالة على معامل الارتباط الجزئي Partial Correlation coefficient.

مع الاشارة الى انه يتم استخدام معامل الارتباط في الحكم على نوع العلاقة بين المتغيرين حيث تكون العلاقة اما طرديه او عكسية ، وكذلك بالنسبة لقوة العلاقة فقد تكون علاقة قوية او متوسطة او ضعيفة .

❖ عادة ما يتم تقسيم محل الدراسة كما ذكرنا سابقاً الى :

أ) متغيرات مستقلة Independent Variables: وهي المتغيرات التي بتغير قيمتها تؤثر في تغيير قيمة متغير او متغيرات اخرى ، اي هي المتغيرات التي تتغير اولا . وسنرمز للمتغير المستقل بالرمز x .

ب) المتغيرات التابعة Dependent Variables : وهي تلك المتغيرات التي تتغير قيمتها بتغير المتغيرات المستقلة او احدها ، اي هي المتغيرات التي تتغير تالية للمتغيرات المستقلة ، وسنرمز للمتغير التابع بالرمز y .

❖ وسيتم قياس الارتباط البسيط من خلال كلا من:

- معامل الارتباط الخطي البسيط لبيرسون Person's Correlation Coefficient
- معامل ارتباط الرتب لسبيرمان Spearman's Rank Correlation Coefficient

اولا : معامل الارتباط الخطي البسيط لبيرسون Person's Correlation Coefficient : يعتبر معامل الارتباط الخطي البسيط لبيرسون Person's Correlation Coefficient والذي سنرمز له بالرمز r_p من أكثر الأدوات الإحصائية استخداما في 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 منها:

$$r_p = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2}\sqrt{\sum(y-\bar{y})^2}}$$

وكذلك المعادلة الرياضية التالية والتي تعتبر اسهل وابسط :

$$r_p = \frac{n\sum xy - \sum x \sum y}{\sqrt{\sum x^2 - (\sum x)^2}\sqrt{\sum y^2 - (\sum y)^2}}$$

ما علينا من هالكلام كله نقدر نطلها بالالة الحاسبة

الحل بالالة الحاسبة CASIO 991ES PLUS

3. مرحلة تجهيز الالة وادخال البيانات :

$MODE \rightarrow 3\ (STAT) \rightarrow 2(A+BX)$ او $AC \rightarrow SHIFT \rightarrow 1(STAT) \rightarrow 2(DATA) \rightarrow AC$ (لانهاء الادخال البيانات والتجهيز للعمليات المطلوبة)

4. مرحلة الحصول على معاملة الارتباط الخطي البسيط لبيرسون:

$r_p = SHIFT \rightarrow 1\ (STAT) \rightarrow 5(RED) \rightarrow 3\ (r) \rightarrow =$

- وتتراوح قيمة معامل الارتباط بين الواحد الصحيح الموجب والواحد الصحيح السالب اي ان قيمة معامل الارتباط تكون كالتالي :

$$1 \geq r_p \geq -1$$

- والارتباط غالبا قيمته كسر اي اقل من الواحد الصحيح
- ولتحديد نوع العلاقة نعتمد على اشارة معامل الارتباط فاذا كانت الاشارة :
  - i. موجبة فان العلاقة تكون طردية
  - ii. سالبة فان العلاقة تكون عكسية
- ولتحديد قوة العلاقة نعتمد على قيمة معامل الارتباط فاذا كانت القيمة :
  - i. اكبر من صفر الى اقل من 0.3 فتكون علاقة ضعيفة
  - ii. من 0.3 الى اقل من 0.7 تكون علاقة متوسطة
  - iii. من 0.7 الى الواحد الصحيح تكون علاقة قوية
  - iv. اما اذا كانت قيمة معامل الارتباط تساوي صفر فلا توجد علاقة خطية او ارتباط بينهما اي يكون المتغيرين مستقلين عن بعضهما البعض
- فمثلا اذا كانت قيمة معامل الارتباط r_p كالتالي فان تفسيره يكون :

قيمة معامل الارتباط	تفسير معامل الارتباط
0.91	ارتباط طردي قوي جدا
-0.87	ارتباط عكسي قوي
-0.21	ارتباط عكسي ضعيف
0.43	ارتباط طردي متوسط
1	ارتباط طردي تام
-0.51	ارتباط عكسي متوسط

⇐ مثال :

فيما يلي بيان بالمنفق على الاعلان والمبيعات لاحد المنتجات فكانت بالمليون ريال كما يلي :

المنفق على الاعلان	2	3	2	7	6	5	10	15	4	11	9	8
المبيعات	10	12	9	22	18	19	26	33	18	22	15	17

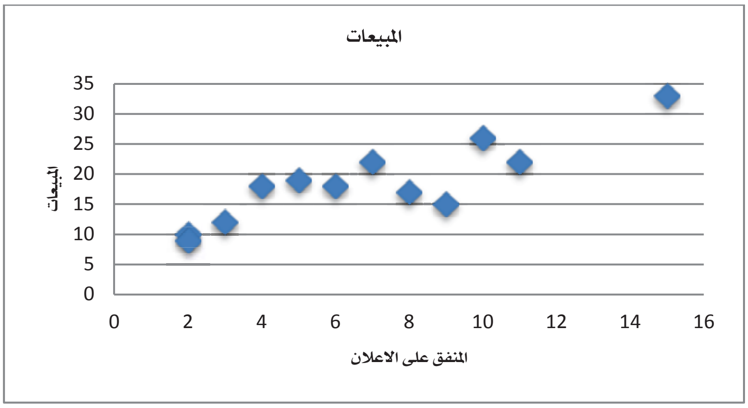
◄ المطلوب:

1- ارسم شكل الانتشار يوضح العلاقة بين المنفق على الاعلان و المبيعات ؟

2- احسب معامل الارتباط الخطى البسيط (بيرسون)، مع التعليق

✓ طريقة الحل :

1- رسم شكل الانتشار يوضح العلاقة بين المنفق على الاعلان و المبيعات :



نستنتج من شكل الانتشار ان قيم كلا من المنفق على الاعلان والمبيعات ياخذ اتجاه تصاعدي جهة اليمين مما يدل على وجود علاقة طرديه بينهما .

2- احسب معامل الارتباط الخطى البسيط (بيرسون)، مع التعليق :

واذا اردنا استخدام المعادلة الرياضية في حسابنا معامل الارتباط بين المنفق على الاعلان والمبيعات لابد اولا من حساب الوسط الحسابي $\bar{x}$ للمنفق على الاعلان والوسط الحسابي للمبيعات $\bar{y}$

حيث ان : $r_p = \dfrac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2}\sqrt{\sum(y-\bar{y})^2}}$

واذا اردنا استخدام المعادلة السابقة في حساب معامل الارتباط بين المنفق على الاعلان والمبيعات لابد اولا من حساب الوسط الحسابي $\bar{x}$ للمنفق على الاعلان والوسط الحسابي للمبيعات $\bar{y}$:

$\bar{x} = \frac{\sum x}{n} = \frac{82}{12} = 6.83333$ ، و $\bar{y} = \frac{\sum y}{n} = \frac{221}{12} = 18.41667$

نختصر الموضوع والحل الطويل نحلها بالالة

الحل بالالة الحاسبة CASIO 991ES PLUS

- مرحلة تجهيز الالة وادخال البيانات :

$MODE \rightarrow 3\,(STAT) \rightarrow 2(A+BX)$ او $AC \rightarrow SHIFT \rightarrow 1(STAT) \rightarrow 2(DATA)$

$\rightarrow AC$ (لانهاء الادخال البيانات والتجهيز للعمليات المطلوبة)

- مرحلة الحصول على معاملة الارتباط الخطي البسيط لبيرسون:

$r_p = SHIFT \rightarrow 1\,(STAT) \rightarrow 5(RED) \rightarrow 3\,(r) \rightarrow =$

حل الكتاب وحل ملخص الزامل خطاء للوسط الحسابي $\bar{x}$ خلاص المثال معامل الارتباط $r_p = 0.8756$ وتدل قيمة معامل الارتباط على وجود علاقة قوية وطردية بين المنفق على الاعلان والمبيعات .

ومن اهم خصائص معامل الارتباط الخطي البسيط لبيرسون انه لايعتمد على قيم المتغيران نفسها عند حساب قيمته وانما يعتمد على مقدار التباعد بين هذه القيم بعضها البعض .

مهمه هذه النقطة لذلك لا يتأثر معامل الارتباط الخطي البسيط باي عمليات جبرية يتم اجراءها على بيانات اي من المتغيرين او احدهما من جمع او طرح او ضرب او قسمة .

⇐ مثال على هذه الملاحظة :

في بيانات المثال السابق اذا اكتشفت ادارة الشركة ان البيانات تم تجميعها وحسابها بطريقة خاطئة حيث يجب اضافة 5 مليون ريال الى جميع قيم المنفق على الاعلان ، كما ان المبيعات يجب مضاعفة قيمتها لجميع القيم .

◀ المطلوب: حسب معامل الارتباط في هذه الحالة بين المنفق على الاعلان والمبيعات ؟

✓ طريقة الحل :

1- اوضح لنا ان ه يجب اضافة 5 ملايين على جميع قيم المنفق بمعنى ناخذ كل قيمة من x ونضيف عليها 5 .

2- اوضح لنا ان المبيعات يجب ان نضاعفها ففي هذه الحالة y نضربها في 2 لكل القيم .

طبعا بدون ما نتعب نفسنا ونعيد الحسابات للاوساط الحسابي نقدر نقوم بالعملية بالالة .
بعد التعويض بالالة يطلع نفس معامل الارتباط (حاولت اضرب واطرح في نفس الوقت ومع هذه خرجت بنفس معامل الارتباط) $r_p = 0.8756$

ثانيا : معامل التحديد **Determination Coefficient** : وهو مربع معامل الارتباط لذلك يرمز له بالرمز R^2 أو R-Square و هو يشير إلى نسبة تفسير المتغير أو المتغيرات المستقلة للتغير فى المتغير التابع .

⇐ فمثلا:

نجد أن المنفق على الاعلان يفسر نسبة $\left(0.8756^2\right)$ أى 76.675 ٪ من التغير فى قيمة المبيعات بينما 23.32 ٪ من التغير فى المبيعات ترجع إلى عوامل أخرى منها الخطاء العشوائى . (يحل بالالة طبعا عبارة عن تربيع معامل الارتباط سواء بالكاسيو او ti)

ثالثا : معامل أرتباط الرتب لسبيرمان r_s **Spearman's Rank Correlation Coefficient** :معامل الارتباط لبيرسون لايمكن استخدمه في حساب قوة العلاقة بين متغيرين الا اذا كانت البيانات المتوافرة عنها في صورة كمية فقط . اما اذا كانت البيانات في صورة وصفية فلا يمكن تطبيق معامل ارتباط بيرسون وحساب الارتباط بين المتغيرين محل الدراسة .

لذا ففي حالة المتغيرات الوصفية فنستخدم معامل ارتباط الرتب لـ سبيرمان ، والذي يتم استخدامه في قياس الارتباط خاصة في حالة البيانات الوصفية الترتيبية مثل تقديرات الطلاب (ممتاز - جيدجدا – جيد ....الخ) وكذلك قوة المركز المالي (جيد – متوسط – ضعيف) ودرجة الموافقة على الرأي في اسئلة الاستبانة (موافق تماما – موافق – محايد - غير موافق – غير موافق على الاطلاق)

ويتم حساب معامل ارتباط الرتب لسبير مان r_s باستخدام المعادلة التالية :

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

حيث ان :

d الفرق بين رتبة المتغيرين

n عدد المشاهدات

ملاحظة يجب مراعاتها عند ترتيب المتغيرات :

1- يتم ترتيب قيم مشاهدات المتغير x وتسمى القيم الترتيب للمتغير x " رتب x " وكذلك الامر للمتغير y وتسمى بـ y " رتب y " ، و الترتيب يكون تصاعديا او تنازليا ولكن اهم شيء هو اذا كان ترتيب x تصاعدي لابد ان يكون ترتيب y تصاعدي ايضا والعكس صحيح .

2- في حال الترتيب التصاعدي مثلا يتم اعطاء اقل قيمة الرتبة 1 والقيمة التي هي اكبر منها الرتبة 2 وهكذا .

3- في حالة تكرار او تساوي بعض القيم لا يمتغير تعطى كل منهم رتبة كما لو كانت القيم غير متساوي ثم نحسب الوسط الحسابي (مجموع الرتب ÷ عددها) لتلك الرتب ويعطى الوسط الحسابي كرتبة تلك القيم المتساوية .

⇐ مثال :

فيما يلي بيان بالمنفق على الاعلان والمبيعات لاحد المنتجات فكانت بالمليون ريال كما يلي :

المنفق على الاعلان	2	3	2	7	6	5	10	15	4	11	9	8
المبيعات	10	12	9	22	18	19	26	33	18	22	15	17

◀ المطلوب: أحسب معامل الارتباط لسبيرمان بين المنفق على الاعلان و المبيعات؟

✓ طريقة الحل :

1- يتم اولا ترتيب قيم كلا من x و y كما يتضح من الجدول التالي :

المنفق على الاعلان x	المبيعات y	رتب x	رتب y	d	$d2$
2	10	1.5	2	-0.5	0.25
3	12	3	3	0	0
2	9	1.	1	0.5	0.25
7	22	6	9.5	-3.5	12.25
9	18	8.5	6.5	2	4
5	19	5	8	-3	9
10	26	10	11	-1	1
15	33	12	12	0	0
4	18	4	6.5	-02.5	6.25
11	22	11	9.5	1.5	2.25
9	15	8.5	4	4.5	20.25
8	17	7	5	2	4
				0	59.5

ونلاحظ من هذا الجدول التالي :

1- تم ترتيب المتغيران تصاعديا

2- عند ترتيب قيم المتغير المنفق على الاعلان x وجدنا ان القيمة 2 تكررت مرتان لتأخذ الرتب 1 و 2 لذلك نحسب المتوسط لهما وهو (1+2) ÷ 2 ليكونا 1.5 لذلك وضعنا اماما

القيمة 2 الرتبة 1.5 وكذلك الامر بالنسبة للقيمة 9 فانها تاخذ الرتبة 8 و 9 لذلك وضعنا اما القيمة 9 الرتبة 5.8 (8 + 9) ÷2= 8.5 .

3- عند ترتيب قيم المتغير " المبيعات " y وجدنا ان القيمة 18 اخذت الرتبة 6 و 7 ووضعنا اما القيمة 18 الرتبة 6.5 وكذلك القيمة 22 اخذت الرتبة 9 و 10 لذلك وضعنا امامها الرتبة 9.5

4- ثم نحسب الفرق بين رتب المتغير x و رتب المتغير y والتي نعطي لها الرمز d ونلاحظ من الجدول السابق ان مجموع الفروق d لابد ان يكون صفر والا يكون هناك خطاء في الترتيب لاحد المتغيرين او كلاهما ولا بد من مراجعة الترتيب مرة اخرى .

والان وبالعودة الى مثالنا : بلغ $\sum d^2 = 59.5$ وحيث ان عدد المشاهدات $n = 12$ فانه يمكن حساب معامل ارتباط الرتب لسبيرمان كما يلي :

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(59.5)}{12(144-1)} = 0.7919$$

بلغ معامل ارتباط الرتب لسبيرمان 0.7919 مما يدل على وجود ارتباط طردي قوي بين المنفق على الاعلان والمبيعات ، وهي قيمة قريبة من التي تم حسابها باستخدام معامل الارتباط لسبيرمان حيث بلغ 0.8756 .

⇐ مثال :

البيانات التالية تمثل التقديرات التى حصل عليها عشر طلاب فى مقررى المحاسبة والقانون:

المحاسبة	ممتاز	جيدجدا	جيد	مقبول	ضعيف	جيد	مقبول	جيدجدا	جيد	مقبول
القانون	جيد	جيد	مقبول	جيد	جيدجدا	جيدجدا	ممتاز	مقبول	جيد	جيدجدا

◄ المطلـــوب:
أحســـب معامـــل الارتباط المناسب ؟

✓ طريقة الحل :

يتم ترتيب المشاهدات وحساب الفروق بين الرتب ومربعاتها كما يتضح من الجدول التالي: (ملاحظة بما ان البيانات رتبية فهنا نستخدم معامل سبيرمان)

المحاسبة	القانون	رتب المحاسبة	رتب قانون	d	d^2
ممتاز	جيد	10	4.5	5.5	30.25
جيد جيدا	جيد	8.5	4.5	4	16
جيد	مقبول	6	1.5	4.5	22.25
مقبول	جيد	3	4.5	-1.5	2.55
ضعيف	جيد جيدا	1	8	-7	49
جيد	جيد جدا	6	8	-2	4
مقبول	ممتاز	3	10	-7	49
جيد جدا	مقبول	8.5	1.5	7	49
جيد	جيد	6	4.5	1.5	2.25
مقبول	جيد جدا	3	8	-5	25
				0	247

- نلاحظ عند ترتيب تقديرات مقرر المحاسبة ان التقدير " مقبول " اخذ الرتب 2 و3 و4 لذلك تم وضع 3 اما التقدير مقبول في مقرر المحاسبة(تذكر القانون ايجاد ترتيب العناصر المتشابهه: (2+3+4)÷3 = 3)
- كما ان تقدير جيد في مقرر القانون اخذ الرتب 3 و4 و5 و6 لذلك توضح الرتبة 4.5 اما التقدير جيد في مقرر القانون . (نفس الطريقة السابقة)
- من الجدول السابق يتضح لنا ان مجموع الفروق d لابد ان يكون صفر .

ومن خلال الجدول اعلاه نستنتج ان $\sum d^2 = 247$ و عدد المشاهدات $n = 10$ ، وبناءا عليه نستطيع حساب معامل ارتباط الرتب لسبيرمان كما يلي :

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(247)}{10(100-1)} = -0.4969$$

ونلاحظ ان معامل ارتباط رتب سبيرمان بلغ -0.4969 مما يدل على وجود ارتباط عكسي متوسط بين تقدير مقرر المحاسبة وتقدير مقرر القانون .

رابعاً : معامل الاقتران Conjunction Coefficient ويستخدم في حساب العلاقة الارتباطية بين المتغيرات الوصفية التي ليس في طبيعتها صفة الترتيب اي الوصفية الاسمية التي يكون لها زوج من الصفات مثل : النوع (ذكر ، انثى) او الحالة التعليمية (متعلم ، غير متعلم) .

وعلى ذلك اذا كان لدينا متغيران لدى كلا منها زوج من الصفات فيكون جدول تكرارات الصفات المشتركة بينهما على الصورة التالية :

x \ Y	الصفة الاولى لـ y	الصفة الثانية لـ y
الصفة الاولى لـ x	A	B
الصفة الثانية لـ x	C	D

حيث A,B,C,D تشير الى التكرارات المشتركة بين صفات المتغيرين

ويمكن حساب معامل الاقتران في هذه الحالة كما يلي :

$$r_c = \frac{AD - BC}{AD + BC}$$

⇐ مثال :

في دراسة اجريت لمعرفة هل هناك علاقة بين العمل والتعليم تم سؤال 200 شخص سؤالين هما :

- هل انت متعلم؟ نعم لا
- هل انت ملتحق باي عمل ؟ نعم لا

وبتجميع الاجابات تم عمل الجدول الاقتران التالي :

العمل \ التعليم	متعلم	أمي
يعمل	113	24
لايعمل	49	15

◄ المطلوب: احسب معامل الاقتران ؟

✓ طريقة الحل :

يمكن حساب معامل الاقتران في هذه الحالة كما :

العمل \ التعليم	متعلم	أمي
يعمل	A=113	B=23
لايعمل	C=49	D=15

$$r_c = \frac{AD - BC}{AD + BC} = \frac{(113)(15) - (23)(49)}{(113)(15) + (23)(49)} = \frac{568}{2822} = 0.20$$

اي يوجد ارتباط ضعيف بين العمل والتعليم

خامساً : معامل التوافق Contingency Coefficient ويستخدم لحساب الارتباط بين المتغيرات الوصفية الاسمية والتي يكون لصفاتها قيم اكثر من 2 مثل الحالة الاجتماعية (اعزب – متزوج – متزوج ويعول – ارمل – مطلق)

وحتى يمكن حسابه يتم اعداد الجدول المزدوج بين صفات المتغيرين ومنه يتضح لنا التكرارات المشتركة بين الصفات التي نعتمد عليها في حساب مقدار يطلق عليه **M**

ويتم حساب معامل التوافق من خلال المعادلة التالية:

$$M = \sum \frac{(f_{ij})^2}{f_i . f_j}$$

حيث ان :

f_{ij} التكرار المشترك بين الصفة i و الصفة j

f_i مجموع صف الصفة i

f_j مجموع عمود الصفة j

اي يتم ايجاد مربع تكرار كل خلية مشتركة (مجموع الصف × مجموع العمود) ثم نجمعهم جميعا .

وعلى ذلك يتم حساب معامل التوافق كما يلي :

$$r_T = \sqrt{\frac{M-1}{M}}$$

⇐ مثال :

أوجد معامل التوافق بين تخصص الطالب ودرجة الرضا عن الدراسة بالكلية الملتحق بها إذا كانت البيانات كما يلي:

الرضا \ التخصص	لغة عربية	جغرافيا	تربية خاصة	المجموع
عالي	30	15	45	90
متوسط	20	30	20	70
منخفض	10	5	5	20
المجموع	**60**	**50**	**70**	**180**

✓ طريقة الحل :

1- يتم اولا ايجاد قيمة M كما يلي :

$$M = \frac{(30)^2}{60 \times 90} + \frac{(15)^2}{50 \times 90} + \frac{(45)^2}{70 \times 90} + \frac{(20)^2}{60 \times 70} + \frac{(30)^2}{50 \times 70} + \frac{(20)^2}{70 \times 70} + \frac{(10)^2}{60 \times 20} + \frac{(5)^2}{50 \times 20} + \frac{(5)^2}{70 \times 20}$$

$$M = 0.166 + 0.05 + 0.32 + 0.095 + 0.257 + 0.081 + 0.083 + 0.025 + 0.017 = 1.094$$

2- وعلى ذلك يتم حساب معامل التوافق كما يلي :

$$r_T = \sqrt{\frac{M-1}{M}} = \sqrt{\frac{1.094-1}{1.094}} = 0.293$$

يوجد ارتباط ضعيف بين تخصص الطالب ودرجة الرضا عن الدراسة بالكلية الملتحق بها .

(المحاضرة الحادية عشر)

تحليل الانحدار

❖ تمهيد :

استكمال لما ذكرنا في المحاضرة السابقة عن الاساليب الاحصائية المناسبة لتقييم العلاقات بين المتغيرات وذلك عن طريق تحليل الارتباط ، والتنبؤ باداء الظاهر في المستقبل وذلك عن طريق تحليل الانحدار موضوع هذه المحاضرة .

ثانيا : تحليل الانحدار REGRESSION ANALYSIS يعتبر تحليل الانحدار اكثر طرق التحليل الاحصائي استخداما ، حيث يتم من خلاله التنبؤ بقيمة احد المتغيرات (المتغير التابع) عند قيمة محددة لمتغير او متغيرات اخرى (المتغيرات المستقلة)
وتسمى العلاقة الرياضية التي تصف سلوك المتغيرات محل الدراسة والتي من خلالها يتم التنبؤ بسلوك احد المتغيرين عند معرفة الاخر بمعادلة خط الانحدار.

❖ هناك صورتان اساسيتان لمعادلة الانحدار وهما :

1- معادلة انحدار $x|y$ (يطلق عليها معادلة انحدار y على x)
2- معادلة انحدار $y|x$ (يطلق عليها معادلة انحدار x على y)

اولا : معادل انحدار $x|y$ (يطلق عليها معادلة انحدار y على x)

في هذه المعادلة تتحدد قيمة المتغير y تبعاً لقيمة المتغير x لذلك يمكن التعبير عن تلك العلاقة الخطية بالمعادلة التالية :

$$\hat{y} = b_0 + b_1 x$$

حيث ان

b_0 ثابت الانحدار او الجزء المقطوع من محور الصادات
b_1 معامل الانحدار او معدل التغير في الدالة

ولتحديد المعادلة الدالة على العلاقة بين المتغيرين y و x لابد من تقدير للثابتين b_0 و b_1 والذين يمكن تقديرهما من خلال تطبيق طريقة المربعات الصغرى كالتالي :
تقوم نظرية المربعات الصغرى على تدنية مجموع مربعات الأخطاء في التقدير إلى اقل حد ممكن.

$$b_1 = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$b_0 = \frac{\sum y}{n} - b_1 \frac{\sum x}{n}$$ او باختصار هذه المعادلة تعني $$b_0 = \bar{y} - b_1 \bar{x}$$

⇐ **مثال :**

عند دراسة العلاقة بين عدد غرف المسكن وكمية الكهرباء المستهلكة بالألف كيلو وات فكانت كما يلي:

عدد الغرف	12	9	14	6	4	7	10	10	5	8
استهلاك كهرباء	9	7	10	5	3	7	8	10	4	6

◄ المطلوب: اوجد التالي :

1- معادلة انحدار y على x ؟
2- تحديد معدل التزايد أو التناقص فى استهلاك الكهرباء؟
3- ماهو الاستهلاك المتوقع لمسكن مكون من 8 غرف؟

✓ طريقة الحل :

1- معادلة انحدار y على x نقوم بعمل الجدول التالي :

يوجد خربطه بين المحتوى والكتاب ضد كلام الدكتور في المحاضرة حيث ان الكتاب والمحتوى موضوح ان x هو الاستهلاك والاصح ان x هي عدد الغرف لانها المتغير المستقل

y^2	x^2	xy	y	x
81	144	108	9	12
49	81	63	7	9
100	196	140	10	14
25	36	30	5	6
9	16	12	3	4
49	49	49	7	7
64	100	80	8	10
100	100	100	10	10
16	25	20	4	5
36	64	48	6	8
529	811	650	69	85

وبالتالي يمكن تقدير b_1 من خلال تطبيق المعادلة التالية :

$$b_1 = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} = \frac{10(650) - (85)(69)}{10(811) - (85)^2} = \frac{635}{885} = 0.717$$

وكذلك يمكننا تقدير b_0 من خلال تطبيق المعادلة التالية :

$$b_0 = \frac{\sum y}{n} - b_1 \frac{\sum x}{n} = \frac{69}{10} - (0.717)\frac{85}{10} = 6.9 - 6.095 = 8.055$$

وعلى ذلك يمكن كتابة معادلة الانحدار y على x على الشكل التالي :

$$\hat{y} = b_0 + b_1 x = 8.055 + 0.717\,x$$

الالة بالنسبة لـ b_0 تطلعة 8.011 بسبب احتساب b_1 باكمل الارقام بعد الفاصلة (لغاية 6 ارقام تقريبا وبدون تقريب)

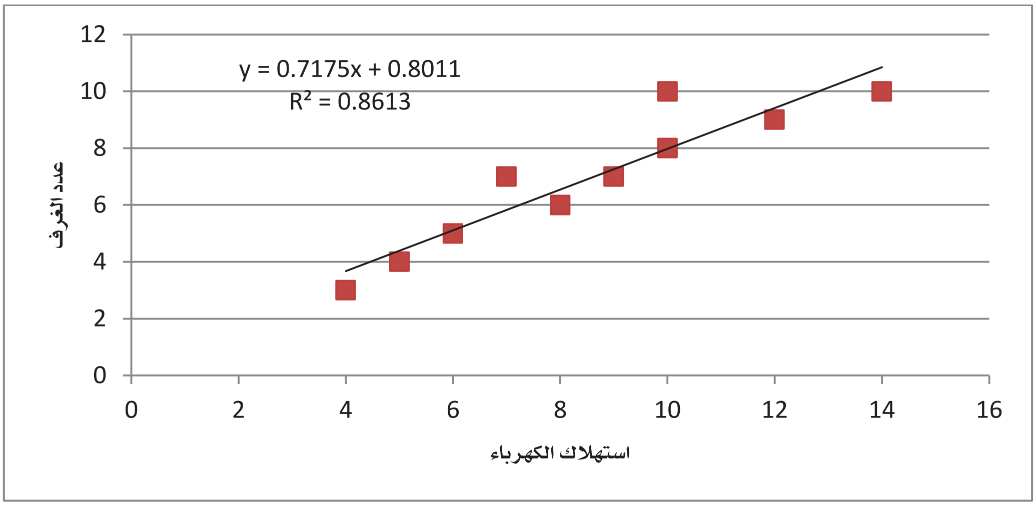
2- وبالتالي يكون معدل التزايد في استخلاك الكهرباء هو b_1 (في الالة سوف يكون نتـــــيجة الـ b) لانها موجبة ويساوي 0.717 اي ان كل غرفة بالمسكن تعمل على زيادة استهلاك الكهرباء بمقدار 717 كيلو وات لكل غرفة بالمنزل .

3- الاستهلاك المتوقع لمسكن مكون من 8 غرف :يتم التعويض في معادلة الانحدار التي سبق ايجادها عندما $x = 8$ كما يلي :

$$\hat{y} = b_0 + b_1 x = 8.055 + (0.717 \times 8) = 6.541$$

اي ان الاستهلاك المتوقع لمسكن مكون من 8 غرف هو 6541 كيلو وات .

ويمكن رسم بيانات المثال السابق وخط معادلة الانحدار y على x كما يلي:



الجدول المستخدم في الرسم : ونوع السم البياني *scatter*

عدد الغرف x	12	9	14	6	4	7	10	10	5	8
استهلاك كهرباء y	9	7	10	5	3	7	8	10	4	6

ويتضح لنا من الشكل السابق ان خط الانحدار لا يمر بجميع النقاط حيث تكون هناك نقاط مشتته حول الخط ، وبالرغم من ذلك يعد هذا الخط من افضل الخطوط التي حصلنا عليها للتعير عن العلاقة بين المتغيرين محل الدراسة ولكن بخطأ معين يسمى خطأ التقدير Standard Error

يمكن حلها بالالة الحاسبة وتريح بالنا :

الحل بالالة الحاسبة CASIO 991ES PLUS

1- مرحلة تجهيز الالة وادخال البيانات :

$MODE \rightarrow 3\ (STAT) \rightarrow 2(A+BX)$ او $AC \rightarrow SHIFT \rightarrow 1(STAT) \rightarrow 2(DATA)$

$\rightarrow AC$ (لانهاء الادخال البيانات والتجهيز للعمليات المطلوبة)

2- بعد ادخال البيانات يمكننا استخراج قيمة a و b حتى نكمل معادلة الانحدار $\hat{y} = b_0 + b_1 x$ هي نفسها معادلة الخط المستقيم التي درسنها في الرياضيات 2 $y = a + bx$ ويكون ذلك بالطريقة التالية :

$SHIFT \rightarrow 1\ (STAT) \rightarrow 5(RED) \rightarrow 1\ (a) \rightarrow =$

$SHIFT \rightarrow 1\ (STAT) \rightarrow 5(RED) \rightarrow 1\ (b) \rightarrow =$

وعندها نكتب المعادلة بالشكلة التالي $\hat{y} = value(a) + value(b)x$

ملاحظة هامة :

- اذا طلب انحدار y على x نستخدم الطريقة اعلاه ولكن اذا طلب انحدار x على y فقط عند ادخال البيانات نقوم بادخال بيانات y في خانة x وبالمثل لقيم x .
- اذا طلب استخراج امر ما بتغيير متغير مستقل كـ x نقوم في المعادلة:

$\hat{y} = value(a) + value(b) \times value(x)$

ثانيا : معادلة انحدار $x|y$ (يطلق عليها معادلة انحدار x على y) : وهنا تتحدد قيمة المتغير x تبعا لقيمة المتغير y لذلك يمكن التعبير عن تلك العلاقة الخطية بالمعادلة التالية :

$$\hat{x} = c_0 + c_1 y$$

حيث ان

c_0 ثابت الانحدار او الجزء الثابت

c_1 معامل الانحدار او معدل التغير في الدالة

ولتحديد المعادلة الدالة على العلاقة بين المتغيرين y و x لابد من تقدير للثابتين c_0 و c_1 والذين يمكن تقديرهما من خلال تطبيق طريقة المربعات الصغرى كالتالي :

تقوم نظرية المربعات الصغرى على تدنية مجموع مربعات الأخطاء فى التقدير إلى اقل حد ممكن.

$$c_1 = \frac{n\sum xy - \sum x \sum y}{n\sum y^2 - (\sum y)^2}$$

$$c_0 = \frac{\sum x}{n} - c_1 \frac{\sum y}{n}$$ او باختصار هذه المعادلة تعني $$c_0 = \bar{x} - c_1 \bar{y}$$

⇐ مثال :

عند دراسة العلاقة بين عدد غرف المسكن وكمية الكهرباء المستهلكة بالألف كيلو وات فكانت كما يلي:

عدد الغرف x	12	9	14	6	4	7	10	10	5	8
استهلاك كهرباء y	9	7	10	5	3	7	8	10	4	6

◄ المطلوب: اوجد التالي :

1- معادلة انحدار x على y

2- ماهو عدد الغرف المتوقع لاستهلاك 25000 كيلو وات ؟

✓ طريقة الحل :

1- معادلة انحدار x على y :

$$c_1 = \frac{n\sum xy - \sum x \sum y}{n\sum y^2 - (\sum y)^2} = \frac{10(650) - (85)(69)}{10(529) - (69)^2} = \frac{635}{529} = 1.2004$$

وكذلك يمكننا تقدير b_0 من خلال تطبيق المعادلة التالية :

$$c_0 = \frac{\sum x}{n} - c_1 \frac{\sum y}{n} = \frac{85}{10} - (1.2004)\frac{69}{10} = 0.217$$

وعلى ذلك يمكن كتابة معادلة الانحدار y على x على الشكل التالي :

$$\hat{x} = c_0 + c_1 y = 0.217 + 1.2004\, y$$

2-اذا كان الاستهلاك للمنزل 25000 كيلووات :

فان عدد الغرف المتوقعه هو : يتم التعويض في معادلة الانحدار التي سبق ايجادها عندما تكون $y = 25$ كمايلي:

$$\hat{x} = 0.217 + (1.2004 \times 25) = 0.217 + 30.01 = 30.227$$

أي انه اذا كان استهلاك الكهرباء في الحد المنازل 25000 كيلووات فان عدد الغرف المتوقع في هذا المنزل = 30 غرفة تقريبا

العلاقة بين معاملي معادلتي الانحدار y على x و معادلة انحدار x على y :

اذا علم معامل انحدار y على x (b_1) ومعامل انحدار انحدار x على y (c_1) فانه يمكن تقدير كلا من معامل التحديد ومعامل الارتباط كما يلي :

$$r^2 = b_1 \times c_1$$

فكما يبدوا معام التحديد هو عبارة عن حاصل ضرب معاملي الانحدار b_1 و c_1 وبالتالي يمكن الحصول على معامل الارتباط باخذ الجذر التربيعي لمعامل التحديد كما يلي :

$$r = \sqrt{r^2}$$

مع ملاحظة ان اشارة معامل الارتباط تكون موجبة او سالبة بما يتفق واشارة كلا من b_1 و c_1 حيث ان اشارتهم جميعا واحده ، لان الاشارة لاي منهم تتوقف على البسط نفسه وهو التغاير بين المتغيرين x و y

كما يمكن معرفة قيمة اي معامل انحدار بمعلومية الاخر كما يلي :

$$b_1 = r \times \frac{\sigma y}{\sigma x} \qquad c_1 = r \times \frac{\sigma x}{\sigma y}$$

حيث ان :

σx الانحراف المعياري للمتغير x

σy الانحراف المعياري للمتغير y

⇐ مثال :

احسب معامل الارتباط بين عدد الغرف والمستهلك من الكهرباء اذا علمت ان :

$$b_1 = 0.717 \quad c_1 = 1.2004$$

✓ طريقة الحل :

ايجاد معامل التحديد كالتالي :

$$r^2 = b_1 \times c_1 = 1.2004 \times 0.717 = 0.8606$$

اي ان عدد الغرف يفسر 86.06 ٪ من التغير في استهلاك الكهرباء .

ايجاد معامل الارتباط كالتالي :

$$r = \sqrt{r^2} = \sqrt{0.8606} = 0.9276$$

مما يدل على وجود ارتباط طردي قوي بين عدد الغرف واستهلاك الكهرباء

(المحاضرة الثانية عشر – الجزء الاول)

السلاسل الزمنية

❖ تمهيد :

يحتاج التخطيط الفعال الى ادوات تنبؤ متقدمة نظريا وتطبيقيا في مجالات احصائية عديدة ومنها تحليل السلاسل الزمنية والتي تقوم على دراسة التطور التاريخي لقيم الظواهر المختلفة لمعرفة خصائثها واستخدامها في استخلاص النتائج النهاية .

وتبرز اهمية تحليل السلاسل الزمنية في حالات كثيرة مرتبطة بالجوانب الاقتصادية والادارية والاجتماعية والبيئية ، ومن ضمن الحالات المتعلقة بالجوانب الاقتصادية والادارية مايلي :

- الناتج المحلي الاجمالي
- اسعار النفط الخام والمنتجات النفطية
- متسوى المخزون
- معدلات التضخم
- كمية وقيمة المبيعات
- اجمالي التكاليف
- اجمالي الودائع
- ميزانية الاعلان
- متسوى الدائنون والمدينون

وفي جميع هذه الحلات يحتاج متخذ القرار الى دراسة البيانات التاريخية كما وكيف ما يشاء ، ومن ثم تحديد الفروق الجوهرية بين الظروف التي احاطت هذه البيانات التاريخية والظروف الحالية من اجل دمجها في مراحل عملية التحليل النهائي المساعدة في اتخاذ القرار .

تعريف السلسلة الزمنية : السلة الزمنية عبار عن مجموعة من المشاهدات الاحصائية تصف الظاهرة مع مرور الزمن ، او هي البيانات الاحصائية التي تجمع او تشاهد او تسجل فترات متتالية من الزمن .
وقد تكون السلسة الزمنية بالارقام المطلقة (وتسمى سلسلة قيم مطلقة) ، او قد تكون السلسلة الزمنية بالقيم النسبية مثل تلك الجداول التي تبين معدلات الزيادة الطبيعية للسكان في الالف ونحوها ، او قد تكون السلسلة الزمنية بالمتوسطات مثل السلسلة الزمنية التي تبين متوسط انتاج الكيلو متر مربع من القمح .

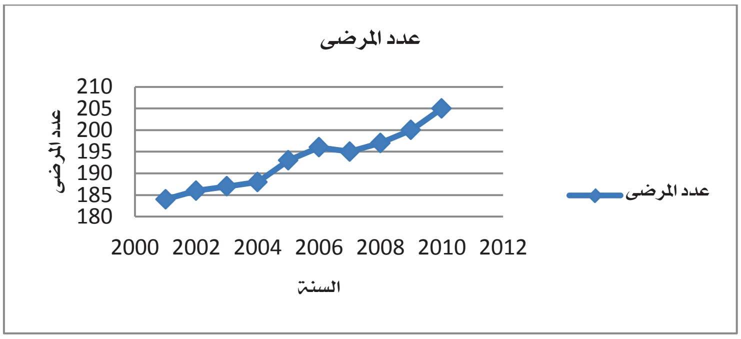
▪ امثلة متنوعة على السلاسل الزمنية :

- مرضى العيادات النفسية المترددين شهريا
- عدد الاطفال المرضى الجدد المصابين بالتوحد شهريات
- قراءة درجات حرارة المريض في ساعة لمدة يوم واحد
- عد المتعطلين سنويا عن العمل
- معدلات الطلاق السنوية
- قراءة الانتاج الشهري لمدة سنة في شركة للادوية
- معدلات الانجاب السنوية
- المبيعات اليومية في مركز لبيع الكتب لمدة شهر
- الانتاج الشهري من البترول للسعودية ولعدة سنوات

كل هذه القراءات وتتابعها الزمني جميعها تمثل سلسلة زمنية .
مثال لجدول سلسلة زمنية يوضح عدد مرضى الفصام المترددين على احد العيادات خلال العشر سنوات الماضية :

السنة	عدد المرضى
2001	184
2002	186
2003	187
2004	188
2005	193
2006	196
2007	195
2008	197
2009	200
2010	205

ويمكن رسم الشكل البياني للسلسلة الزمنية للجدول اعلاه على الشكل التالية scatter حيث يتم الرسم من خلال رسم محورين سيني ويوضح السنوات وصادي يوضح عدد مرضى الفصام ومن ثم تحديد احداثيات النقاط فيظهر لنا الشكل التالي :



- انواع السلسلة الزمنية: السلسلة الزمنية نوعان هما :

1- سلسلة زمنية فترية : وهي السلسلة التي تتكون من بيانات كمية لمستوى الظاهرة عن فترات محددة من الزمن (شهر ، ربع سنة ، او ما شابه ذلك)
2- سلسلة زمنية لحظية : وهي السلسلة التي تتكون من مستويات للظاهرة مقاسة في لحظات (تواريخ معينة ومحددة)

- تحليل السلسلة الزمنية: لغرض فهم السلسلة الزمنية لابد من تحليلها الى عناصرها ومركباتها الاساسية مما يمكننا من معرفة تطور الظاهرة مع الزمن والتنبؤ بمعالمها خلال الفترات المقبلة لتتخذ اساسا للتخطيط الاقتصادي او الاداري الطويل الاجل ، وتتألف السلسلة الزمنية من اربعة عناصر اساسية هي :

1- الاتجاه العام ويرمز لقيمه بالرمز (T) وتسمى " القيم الاتجاهية "
2- التغيرات الموسمية ويرمز لقيمها بالرمز (S) وتسمى " القيم الموسمية "
3- التغيرات الدورية ويرمز لقيمها بالرمز (C) وتسمى " القيم الدورية "
4- التغيرات العشوائية او الفجائية ويرمز لقيمها بالرمز (R) وتسمى " القيم العشوائية "

اي ان القيمة الاصلية للظاهرة (Yt) في كل سنة من السنوات يمكن وصفها بالعلاقة التالية :

1. نموذج الجمع :

ويستخدم عندما يكون مدى التغيرات الموسمية ثابت من سنة الى اخرى ومستقل عن الاتجاه العام ، ويتم فرض ان السلسلة الزمنية مكونة من مجموع مكوناتها من الاربع عناصر السابق ذكرها ، اي كون النموذج بالصورة التالية :

$$y_t = T_t + C_t + S_t + R_t$$

2. نموذج الضرب :

ويستخدم هذا النموذج في الحالات المعاكسة لحالات استخدام نموذج الجمع ، ويتم فرض ان السلسلة الزمنية مكونة من حاصل ضرب مكوناتها من الاربع عناصر ، اي يكون النموذج على الصورة التالية :

$$y_t = T_t \times C_t \times S_t \times R_t$$

حيث ان:

- Y_t = قيمة الظاهرة المدروسة في الفترة t (القيمة الحقيقة)
- T_t = قيمة الاتجاه في الفترة t
- C_t = قيمة التغيرات الموسمية (القيمة الموسمية) في الفترة t
- S_t = قيمة التغيرات الدورية (القيم الدورية) في الفترة t
- R_t = قيمة التغيرات العشوائية (القيم العشوائية) في الفترة t

ويجب ملاحظة ان قيم المتغيرات الموسمية وكذلك المتغيرات الدورية عبارة عن نسب مئوية في نموذج الضرب

- **عناصر السلسلة الزمنية** :ان دراسة اي سلسلة زمنية وتحليلها يستدعي دراسة كل عنصر من هذه العناصر على حده ، وهذه العناصر هي :

1) الاتجاه العام THE SECULAR TREND:

تغيرات الاتجاه العام تعني الزيادة او الانخفاض طويل الاجل في البيانات عبر الزمن ، ويتم التعرف على ذلك من خلال تمثيل السلسلة الزمنية بيانيا فنحصل بالتالي على خط بياني ، واتجاع خط السلسلة الزمنية صعودا او هبوطا يسمى الاتجاع العام للسلسلة ، فاذا نظرنا للخط ووجدناع يتجع من الاسفل الى الاعلى دل ذلك على نمو الظاهرة مع مرور الزمن ، اما اذا كان الخط يهبط من لاعلى الى الاسفل دل ذلك على ان الظاهرة تنتقص مع مرور الزمن ، اما اذا كان الخط افقيا دل ذلك على ثبات الظاهرة .

طرق حساب الاتجاه العام :

أ. **طريقة الانتشار (التمهيد باليد) :**

يتم بهذه الطريقة رسم شكل الانتشار للظاهرة موضوع الدراسة ، وشكل الانتشار عبارة عن رسم بياني لمتغيرين بحيث يكون الزمن على المحور السيني وقيم الظاهرة على المحور الصادي ، وعند توصيل نقاط شكل الانتشار ببعضها البعض نحصل على الخط البياني للظاهرة عبر الزمن ، ويعطي شكل الانتشار فكرة سريعة عن طبيعة الاتجاه العام للظاهرة ومدة ارتباطه بالزمن ومدى تأثير التقلبات الدورية او الموسمية او التغيرات العشوائية ، وبالامكان من خلال شكل الانتشار القيام بعملية مقارنة بين سلسلتين او اكثر عبر فترات مختلفة من الزمن.

وعملية التمهيد باليد (شكل الانتشار) عادة لاتكون دقيقة مما يقلل الاعتماد عليها وذلك لان التمهيد باليد يتم بطريقة تقديرية تختلف من شخص لاخر وتعتمد على مهارة الشخص في رسم خط يمر باكبر عدد ممكن من النقاط ويمثل السلسلة افضل تمثيل .

⇐ **مثال :**

اذا كان لدينا بيانات ربع سنوية لاجمالي الودائع في المصارف السعودية (الاف الملايين من الريالات) في الفترة من النصف الاخير من عام 2005 م الى عام 2007 م والموضحة في الجدول التالي :

السنة	2005	2005	2006	2006	2006	2006	2007	2007	2007	2007
الفصل	3	4	1	2	3	4	1	2	3	4
الودائع	195.6	196.9	200.1	205.3	207.4	215.4	223.3	222.3	222.07	226.1

◄ **المطلوب:** رسم شكل الانتشار لهذه البيانات ومن ثم تفسيره وابراز معالم الاتجاه العام للظاهرة موضع الدراسة

✓ **طريقة الحل :**

يتم رسم شكل الانتشار من خلال رسم محورين سيني ويوضح الفترات الزمنية بربع السنة وصادي يوضح الودائع ومن ثم تحديد احداثيات النقاط.

ويمكن رسم شكل الانتشار من خلال ادخال البيانات السابقة الى برنامج الاكسل لتكون بالشكل التالي:

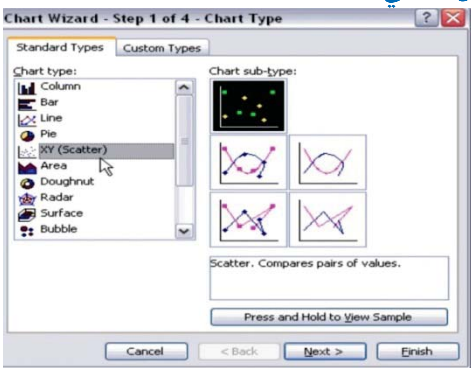
	A	B	C	D
1	السنة	الفصل	الفترة الزمنية	الودائع
2	2005	3	1	195.64
3		4	2	196.97
4	2006	1	3	200.11
5		2	4	205.33
6		3	5	207.49
7		4	6	215.46
8	2007	1	7	223.36
9		2	8	222.31
10		3	9	222.07
11		4	10	226.18

نلاحظ اننا بالاضافة الى البيانات التي كانت موجودة بالتمرين وهي السنة والفصل والودائع تم اضافة عمود يوضح الفترة الزمنية وهي تأخذ القيم (1،2،3،.....10) ، ويمكن رسم الشكل الانتشاري للبيانات الربع سنوية الخاصة باجمالي الودائع في المصارف السعودية باتباع الخطوات التالية :

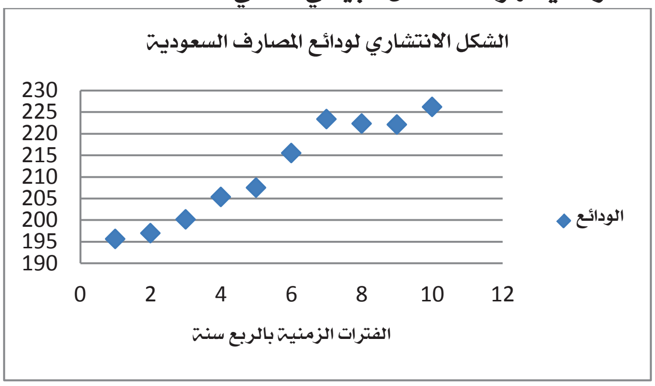
1- يتم تحديد العمودين الخاصين بالفترات الزمنية والودائع المطلوب رسم الشكل الانتشاري لهما كما بالشكل التالي :

	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

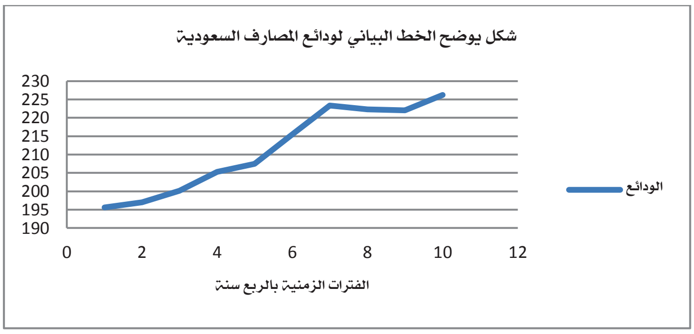
2- ثم نختار من قائمة الرسومات البيانية chart wizard رسم الشكل الانتشاري من خلال xy (scatter) كما بالشكل التالي :



وباستكمال باقي الخطوات يظهر لنا الشكل البياني التالي:



كما يمكن رسم الخط البياني الخاص ببيانات اجمالي الودائع في المصارف السعودية ليكون كما يلـي :



فعند رسم شكل الانتشار لهذه البيانات كما يبدوا في الشكل السابق ، نستطيع من خلال هذا الرسم لتوضيح التالي :

1- يتبين لنا ان هناك ارتفاع مستمر في اجمالي الودائع عبر الزمن
2- الاتجاه العام لبيانات اجمالي الودائع يمكن وصفه بدالة خطية .
3- ميل خط الاتجاه العام لبيانات اجمالي الودائع سيكون موجبا .

ب. **طريقة المتوسطات المتحركة :**

تعتمد هذه الطريقة على اخذ متوسطات متتابعة لمجموعات متتابعة ومتداخلة من البيانات، والهدف الاساسي من ذلك هو ازالة التعرجات من خط السلسلة الزمنية . وهذه الطريقة اكثر دقة في تحديد خط الاتجاه العام من طريقة شكل الانتشار (التمهيد باليد).

ويتم حساب المتوسط المتحرك من خلال تطبيق قانون المتوسط الحاسبي بشكل متتابع لعدد من المشاهدات المعطاة لدينا، مع الاخذ في الاعتبار طول المجموعة التي يتم تقسيم البيانات اليها فمثلا اذا كان طول المجموعة 5 يتم ايجاد متوسط المشاهدات $(x1, x2, x3, x4, x5)$ وذلك بايجاد مجموعهم والقسمة على عددهم كما يبدو ذلك من خلال العلاقة التالية :

$$\frac{x1, x2, x3, x4, x5}{5}$$

ونضع المتوسط الذي تم الحصول عليه امام الفئة التي في المنتصف وهي امام المشاهدة $x3$ ثم نحسب المتوسط من جديد للمشاهدات $(x2, x3, x4, x5, x6)$ ونضع المتوسط الجديد الذي تم حسابه امام المشاهدة $x4$. وهكذا حتى نصل الى المتوسط الاخير 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 الاولى ويرسم الخط من النقاط التي تمثل المتوسطات المتحركة الثانية لانها تعطي خطا اكثر تمهيدا ، ويكون اسلوب المتوسط المتحرك فعلا عندما تكون بيانات السلسلة الزمنية مستقرة عبر الزمن .

⇐ مثال :

اوجد المتوسطات المتحركة بطول 5 للسلسلة الزمنية التالية :

المشاهدة	$x1$	$x2$	$x3$	$x4$	$x5$	$x6$	$x7$
قيمتها	7	13	21	23	27	19	17

✓ طريقة الحل :

يتم اولا ايجاد متوسط الخمس مشاهدات والتي يكون مركزها هو $x3$ وكان الناتج هو 18.2 ، ثم نحسب المتوسط مره اخرى بداية من $x2$ حتى $x6$ والتي يكون مركزها $x4$ وكان الناتج هو 20.6 وهكذا ونتوقف حين لايمكن لنا تكوين سلسلة طولها 5 مشاهدات وتظهر لنا النتيجة كما يبدو ذلك في الجدول التالي :

المشاهدات	القيمة	المتوسط المتحرك
$x1$	7	
$x2$	13	
$x3$	21	18.2
$x4$	23	20.6
$x5$	27	21.4
$x6$	19	
$x7$	17	

وبعد حساب المتوسطات المتحركة نقوم برسم خط الاتجاه العام من هذه المتوسطات المتحركة.

ج. طريقة متوسط نصف السلسلة :

تعتبر هذه الطريقة ادق من طرق شكل الانتشار وطريقة المتوسطات المتحركة ، ويمكن حسابها من خلال اتباع الخطوات التالية :

- نقسم السلسلة الى مجموعتين وفق تسلسل السنوات.
- لتعيين الاحداثي الصادي للنقطتين نوجد المتوسط الحسابي لنصف السلسلة الاول اذا كان عدد المشاهدات زوجي ، اما اذا كان عدد المشاهدات فردي فتهمل المشاهدة الوسطى ثم نجد المتوسط الحسابي للنصف الثاني .
- لتحديد الاحداثي السيني نعطي قيم المشاهدات ترقيم متسلسل سواء كانت المشاهدات قيما او غير ذلك ، ثم نجد المتوسط الحسابي للنصف الاول من القيم سواء كان عددها زوجي او فردي فيكون الامتسوط هو الاحداثي السيني ، وكذلك حساب المتوسط الحسابي للنصف الثاني والذي يمثل الاحداثي السيني وبذلك تتعين النقطتين.
- نصل بين النقطتين بعد تعيينهما على مستوى الاحداثي فيكون لدينا خط الاتجاه العام.
- نوجد معادلة خط الاتجاه العام من خلال العلاقة التالية :

$$\frac{y - y_1}{x - x_1} = \frac{y_{21} - y_1}{x_2 - x_1}$$

⇐ **مثال :**

اذا كان انتاج مصنع سيارات (بالالاف) خلال عشر سنوات كالتالي :

السنة x	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
عدد السيارات y	53	64	67	60	69	74	67	79	85	90

◄ **المطلوب:** ايجاد معادلة خط الاتجاه العام بطريقة متوسط نصف السلسلة؟

✓ **طريقة الحل :**

نكون الجدول التالي من الجدول الرئيسي :

السنة	السنة بالترقيم x	عدد السيارات المنتجة y	متوسط نصف x بالترقيم	متوسط نصف y
1998	1	53	$x_1 = 3$	$y_1 = 62.6$
1999	2	64		
2000	3	67		
2001	4	60		
2002	5	69		
2003	6	74	$x_1 = 8$	$y_1 = 79$
2004	7	67		
2005	8	79		
2006	9	85		
2007	10	90		

المتوسط الاول لنصف x_1	$x_1 = \frac{1+2+3+4+5}{5} = \frac{15}{5} = 3$
المتوسط الثاني لنصف x_2	$x_2 = \frac{6+7+8+9+10}{5} = \frac{40}{5} = 8$
المتوسط الاولى لنصف y_1	$y_1 = \frac{53+64+67+60+69}{5} = \frac{313}{5} = 62.6$
المتوسط الاولى لنصف y_2	$y_2 = \frac{74+67+79+85+90}{5} = \frac{395}{5} = 79$

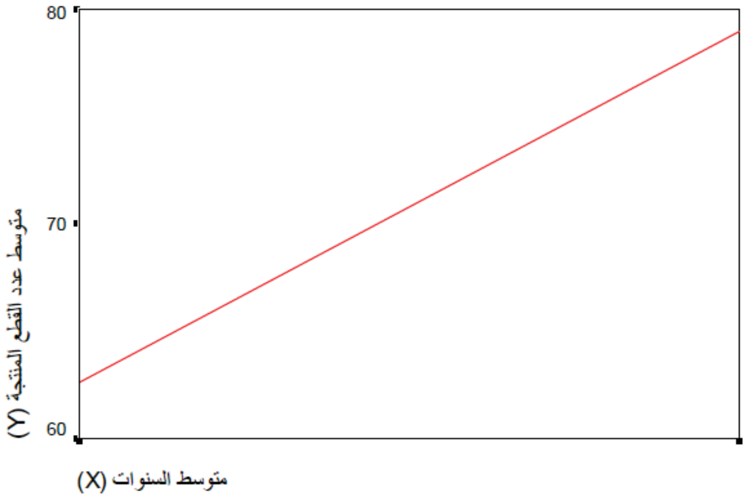
اذا النقطتين المطلوبتين لتحديد الاحداثي السيني والصادي هما :

(3، 62.6) ونسميها النقطة (أ) ، و (8، 79) ونسميها النقطة (ب)

نعين النقطتين على الرسم البياني بحيث يكون احداثي النقطة الاولى هو (3 ، 62.6) و احداثي النقطة الثنانية هو و (8 ، 79) ثم نصل بين النقطتين بخط مستيقيم فيكون هو خط الاتجاه العام كما يبوا ذلك في الشكل التالي :

خط الاتجاه العام

من خلال طريقة متوسط نصف السلسلة



والان نجد معادلة خط الاتجاه العام من خلال تطبيق العلاقة التالية :

$$\frac{y-y_1}{x-x_1}=\frac{y_2-y_1}{x_2-x_1}=\frac{y-62.6}{x-3}=\frac{79-62.6}{8-3}=\frac{16.4}{5}=\frac{y-62.6}{x-3}=\frac{16.4}{5}$$

وبضرب طرفي المعادلة كالتالي :

$$5y-(62.6\ \times 5)=16.4x-(164.\times 3)$$
$$5y-313=16.4x-49.2$$
$$5y=16.4x-(49.2)+(313)$$
$$5y=16.4x+263.8$$
$$\mathrm{y}=\frac{16.4}{5}x+\frac{263.8}{5}$$
$$\mathrm{y}=3.28\mathrm{x}+52.76$$

وهذه هي معادلة خط الاتجاه العام من خلال طريقة متوسط نصف السلسلة

د. **طريقة المربعات الصغرى :**

تعتبر طريقة المربعا الصغرى اكثر دقة من الطرق السابقة لحساب خط الاتجاه العام وذلك من خلال استخدام استلوب الانحدار الخطي البسيط المعتمد على طريقة المربعات الصغرى التي تجعل مجموع مربعات انحرافات القيم المقدرة عن القيم الفعلية اقل ما يمكن وذلك من خلال العلاقات التالية :

$$\hat{y}_t=b_0+b_1t$$

حيث ان :

$\hat{y}_t$	القيمة الاتجاهية للسلسلة الزمنية في الفترة t
b_0	نقطة تقاطع خط الاتجاه العام مع المحور الصادي او الجزء الثابت
b_1	ميل خط الاتجاه العام
t	الزمن

ولغرض حساب b_0 و b_1 نقوم بتطبيق المعادلتين التاليتين :

$$b_1 = \frac{n\sum ty_t - \sum t \sum y_t}{n\sum t^2 - (\sum t)^2}$$

$$b_0 = \frac{\sum y_t}{n} - b_1 \frac{\sum t}{n}$$

حيث ان :

y_t	القيمة الفعلية للسلسلة الزمنية في الفترة t
n	عدد الفترات

⇐ مثال :

بدراسة احد الظواهر الاجتماعية والمتمثلة في العنف الاسري لاحد المدن ، تبين ان تطور اعداد الاسر التي يوجد بها عنف اسري كانت كما يلي خلال مدة الدراسة :

السنة	2004	2005	2006	2007	2008	2009	2010
عدد الاسر	17	25	33	41	39	48	53

◄ المطلوب: التالي:

1. تقدير معادلة الاتجاه العام لتطور اعداد الاسر المعرضة لظاهرة العنف الاسري بهذه المدينة ؟
2. ماهو عدد الاسر المتوقع المعرضون لهذه الظاهرة في عام 2013 ؟

✓ طريقة الحل :

1. حتى يمكن تقدير معادلة الاتجاه العام لتطور اعداد الاسر المعرضة لظاهرة العنف الاسري بهذه المدينة لابد من اعداد الجدول التالي على اعتبار ان السنة الاولى تكون قيمة t تساوي 1 والسنة الثانية تكون قيمتها 2 وهكذا كما يلي :

السنوات	y	t	y_t	t^2
2004	17	1	17	1
2005	25	2	50	4
2006	33	3	99	9
2007	41	4	164	16
2008	39	5	195	25
2010	53	7	371	49
المجموع	256	28	1184	140

كما يتضح لنا ان عدد المشاهدات $n = 7$

ولغرض حساب b_0 و b_1 نقوم بتطبيق المعادلتين التاليتين :

$$b_1 = \frac{n\sum ty_t - \sum t \sum y_t}{n\sum t^2 - (\sum t)^2}$$

$$b_1 = \frac{7(1184) - (28 \times 256)}{7(140) - (28)^2} = \frac{1120}{196} = 5.714$$

ويدل ذلك علي ان معدل التزايد السنوي في الاسر المعرضة للعنف الاسرى 5.714 اسرة

$$b_0 = \frac{\sum y_t}{n} - b_1 \frac{\sum t}{n} = \frac{256 - (5.714 \times 28)}{7} = 13.715$$

وعلى ذلك تكون معادلة الاتجاه العام كما يلي :

$$\hat{y}_t = b_0 + b_1 t = 13.715 + 5.714t$$

2. ماهو عدد الاسر المتوقع المعرضون لهذه الظاهرة في عام 2013 ؟

حتى يمكن تقدر عدد الاسر المتوقع المعرضون لهذه الظاهرة في عام 2013 لابد من تحديد قيمة t في هذه السنة كما يلي :

السنة	t
2010	7
2011	8
2012	9
2013	10

وبناء على ذلك نعوض في معادلة الاتجاه العام عن قيمة t تساوي 10 كالتالي :

$$\hat{y}_t = 13.715 + (5.714 \times 10) = 70.855$$

ويتضح لنا مما سبق ان العدد المتوقع للاسر المعرضة للعنف الاسري يبلغ 70.855 اي ما يقرب من 71 اسرة في عام 2013 .

(المحاضرة الثانية عشر- الجزء الثاني)

السلاسل الزمنية

2) التغيرات الموسمية SEASONAL VARIATIONS

التغيرات الموسمية هي نتيجه طبيعيه لاختلاف الظروف بشكل منتظم مما يؤثر على اختلاف رغبات الناس تبعا لعوامل عديدة منها الزمان والمكان ، ويمكن تعريفها بانها التغيرات التي تطرأ على الظاهرة على مدار المواسم المختلفة للفترة الزمنية موضوع القياس (الموسم) ، فهي قد تكون يومية ، وقد تكون اسبوعية ، وقد تكون شهرية .

مما سبق نرى ان التغيرات الموسمية تحدث في مواعيد زمنية محددة لا يلبث هذا التغير ان يستعيد سيرته الاولى في نفس المواعيد وعلى مدار نفس الفترة الزمنية .

والتغير الموسمي يعتبر ابسط انواع 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 اعلى .

وتكمن اهمية دراسة التغيرات الموسمية في تحليل السلسلة الزمنية للظاهرة خاصة فيما يتعلق بالتخطيطي لعمليات الانتاج او الاوقات المناسبة للاعلانات عن السلع او التوسع في المشاريع ، فالتغيرات الموسمية بشكل عام تساعد على الكشف على :

- الاوقات المناسبة للتغيير
- مسببات التغيير
- الاستعدادات المناسبة لمواجعة التغيير

ويتم قياس التغيرات الموسمية عن طريق ايجاد قيمة الظاهرة في كل موسم من المواسم التي تتعرض لها الظاهرة للتغير ثم تنسب كل قيمة للمتوسط العام لقيم هذه الظاهرة ، اذ يتم اعتبر المتوسط العام (100 ٪) فنحصل على ارقام تدل على مدى التغيرات للظاهرة هل هي فوق المتوسط او دونه ، مثال على ذلك ما يذاع عن درجات الحرارة المتوقعة في النشرات الجوية من انها فوق المتوسط او ادون المتوسط ، ولحساب الاثار الموسمية هناك طريقتان :

1- طريقة النسب للمتوسط المتحرك من خلال المعادلة التالية:

$$y_t = T_t \times C_t \times S_t \times R_t$$

2- طريقة ايجاد القيم المخلصة من اثر الاتجاه العام وذلك بقسمة طرفية المعادلة على (Tt) والتي تمثل الاتجاه العام فنحصل بالتالي على المعادلة التالية :

$$\frac{y_t}{T_t} = C_t \times S_t \times R_t$$

⇐ مثال :

اذا كان لدينا انتاج احدى الشركات خلال ثلاث سنوات ، وكانت كمية الانتاج ماخوذه كل ثلاثة شهور (السنة مقسمة الى اربعة ارباع) والانتاج بالالاف الوحدات كما يبدو ذلك في الجدول التالي :

ربع السنة	2008	2009	2010
الاول	3	4	8
الثاني	7	5	10
الثالث	9	6	12
الرابع	2	4	6

◄ المطلوب: التالي:

1) تقدير معادلة الاتجاه العام للعلاقة بين الانتاج والزمن ؟
2) تقدير القيم الاتجاهيه المقابلة للقيم الاصلية للانتاج ؟
3) ايجاد القيم المخلصة من اثر الاجاه العام ؟
4) تحديد تأثير كل موسم ؟
5) تقدير الانتاج المتوقع سنة 2012 ؟

✓ **طريقة الحل :**

1. **تقدير معادلة الاتجاه العم للعلاقة بين الانتاج والزمن :**

يتم اولا ادخال البيانات السابقة مع اضافة عنصر الزمن t، ثم يتم حساب العمود y_t والعمود t^2 وايجاد المجاميع اللازمة لحساب معامل الانحدار b_1 كمايلي :

السنة	الربع	الانتاج y	الزمن t	y_t	t^2
2008	الاول	3	1	3	1
	الثاني	7	2	14	4
	الثالث	9	3	27	9
	الرابع	2	4	8	16
2009	الاول	4	5	20	25
	الثاني	5	6	30	36
	الثالث	6	7	42	49
	الرابع	4	8	32	64
2010	الاول	8	9	72	81
	الثاني	10	10	100	100
	الثالث	12	11	132	121
	الرابع	6	12	72	144
المجموع		76	78	552	650

حيث n هي الفترات الزمنية تساوي 12

نحسب قيمة b_1 من خلال العلاقة التالي :

$$b_1 = \frac{n\sum ty - \sum t \sum y}{n\sum t^2 - (\sum t)^2} = \frac{12(552) - (78 \times 76)}{12(650) - (78)^2} = 0.40559$$

وبالتالي يكون معدل التزايد كل فترة ربع سنة هو 0.40559 الف وحدة .

نحسب قيمة b_0 من خلال المعادلة التالية :

$$b_0 = \frac{\sum y_t}{n} - b_1 \frac{\sum t}{n} = \frac{76 - (0.40559 \times 78)}{12} = 3.69697$$

وعلى هذا تتحدد قيمة معادلة خط الاتجاه العام من خلال العلاقة التالية

$$\hat{y}_t = b_0 + b_1 t = 3.69697 + 0.40559\, t$$

2. **تقدير القيم الاتجاهيه المقابلة للقيم الاصلية للانتاج :**

يمكن ايجاد القيم الاتجاهية بالتعويض في معادلة الانحدار السابق الحصول عليها بقيم t بداية من 1،2،3،....،12 وبذلك تكون القيم الاتجاهية هي :

$$\hat{y}_t = 3.69697 + 0.40559\, t$$

السنة	الربع	الانتاج y	الزمن t	القيم الاتجاهية
2008	الاول	3	1	4.10256
	الثاني	7	2	4.50815
	الثالث	9	3	4.91374
	الرابع	2	4	5.31933
2009	الاول	4	5	5.72492
	الثاني	5	6	6.13051
	الثالث	6	7	6.5361
	الرابع	4	8	6.94169
2010	الاول	8	9	7.34728
	الثاني	10	10	7.75287
	الثالث	12	11	8.15846
	الرابع	6	12	8.56405
المجموع		76	78	

3. ايجاد القيم المخلصة من اثر الاتجاه العام:

ويتم حساب القيم المخلصة من اثر الاتجاع العام بقسمة قيم الظاهرة الاصلية على القيم الاتجاهية فتكون النتيجة كما بالجدول التالي :

$$\hat{y}_t = 3.69697 + 0.40559\ t$$

السنة	الربع	الانتاج y	الزمن t	القيم الاتجاهية	القيم المخلصة من اثر الاتجاه العام
2008	الاول	3	1	4.10256	0.7313
	الثاني	7	2	4.50815	1.5527
	الثالث	9	3	4.91374	1.8316
	الرابع	2	4	5.31933	0.376
2009	الاول	4	5	5.72492	0.6987
	الثاني	5	6	6.13051	0.8156
	الثالث	6	7	6.5361	0.918
	الرابع	4	8	6.94169	0.5762
2010	الاول	8	9	7.34728	1.0888
	الثاني	10	10	7.75287	1.2898
	الثالث	12	11	8.15846	1.4709
	الرابع	6	12	8.56405	0.7006
المجموع		76	78		

4. ايجاد تأثير كل موسم :

حتى يمكن ايجاد تأثير كل موسم نعيد ترتيب القيم المخلصة من اثر الاتجاه العام السابق الحصول عليها كما يلي :

الموسم	2008	2009	2010
الاول	0.7313	0.6987	1.0888
الثاني	1.5527	0.8156	1.2898
الثالث	1.8316	0.918	1.4709
الرابع	0.376	0.5762	0.7006

ثم ايجاد متوسط القيم المخلصة من اثر الاتجاه العام لكل ربع للتعبير عن اثر ذلك الموسم فمثلا :

تأثير الربع الاول: $\frac{0.7313+0.6987+1.0888}{3} = 0.8396$

وبهكذا لباقي المواسم فتكون النتيجة كما يلي :

الموسم	2008	2009	2010	تأثير الموسم
الاول	0.7313	0.6987	1.0888	0.8396
الثاني	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 الموسمي) 4.0167 اي 401.67 ٪ وحيث يوجد 4 مواسم لذا فان مجموع تأثيرات المواسم لابد ان تساوي 400 ٪ .

لذا لابد من تعديل قيم الدليل الموسمي بمعامل تصحيح قدره $\frac{4}{4.0167}$

الموسم	تأثير الموسم	تأثير الموسم المعدل
الاول	0.8396	0.836109
الثاني	1.2194	1.21433
الثالث	1.4068	1.400951
الرابع	0.5509	0.54861
المجموع	4.0167	4

5. تقدير الانتاج المتوقع سنة 2012 :

نلاحظ ان قيم t في الربع الاخير من سنة 2010 بلغت 12 لذلك يتم الزيادة عليها سنة 2011 لتكون 13 ، 14 ، 15 ، 16 خلال المواسم الاربع ولذل تكون القيم خلال سنة 2012 هي 17 ، 18 ، 19 ، 20 والتي يتم التعويض بها في معادلة الاتجاه العام للحصول على القيم الاتجاهيه ويمكن تقدير القيم المتنبئ بها لكل ربع كما يلي :

القيم المتنبئ بها للموسم = القيمة الاتجاهية × تأثير الموسم المعدل

$$\hat{y}_t = 3.69697 + 0.40559\, t$$

وعلى ذلك يكمن تقدير الانتاج المتوقع سنة 2012 كما يلي :

الموسم	t	القيم الاتجاهية	تأثير الموسم المعدل	الانتاج المتوقع
الاول	17	10.592	0.836109	8.856069
الثاني	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 ان الانتاج المتوقع سنة 2012 هو 44664.48 وحدة

3) التغيرات الدورية CYCLICAL VARIATIONS

ويعرف هذا النوع من التغيرات بدورات الاعمال ، وهذا يمتد لفترة زمنية اطول من سنة ، وتنشأ هذه التغيرات عن ظروف عامة تعزي الى العوامل التي تتحكم في الحياة الاقتصادي للبلاد ويهتم الباحثون الاقتصاديون ورجال الاعمال بالتغيرات الدورية لغايات التخطيط لمواجهة المشاكل التي قد تنشأ عن حدوثها ، وقد تمتد بعض التغيرات الدورية الى 50 سنة وهذه دورة طويلة ، اما الدورة المتوسطة فقد تمتد بين 8 -12 سنة ، اما الدورة القصيرة فتكون بين 3-4 سنوات ، وتقع التقلبات الدورية اعلى واسفل خط الاتجاه العام .

4) التغيرات العشوائية أو الفجائية (RANDOM IRREGULAR)

تؤثر هذه التغيرات على السلسلة الزمنية بشكل عشوائي او مفاجئ وغير منتظم ، فقد تكون هذه التغيرات ناتجة عن حدوث ظواهر طبيعية مثل الزلازل و البراكين او حروب و نحوها ، 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 :

- انها لا تحدث وفقا لقاعدة او قانون .
- قد تتكرر او لا تتكرر
- تأثيرها غير ثابت فتارة تأثر بالنقص وتارة بالزيادة .
- لاتستمر طويلا لذا يطلع عليها اسم التغيرات قصيرة الاجل .

(المحاضرة الثالثة عشر)

الارقام القياسية

❖ تعريف الارقام القياسية

الرقم القياسي هو مؤشر احصائي (رقم نسبي) يستخدم في قياس التغير النسبي الذي طرأ على ظاهرة من الظواهر الاقتصادية او الاجتماعية ، فهو يستخدم لقياس التغير في اسعار السلع او في حجم انتاجها او في كميات المبيعات منها او في حجم السكان او اجور العمال (وفقا لاساس معين) سواء كان هذا الاساس فترة زمنية معينة او مكانا معينا .

❖ فترة الاساس :

الاساس هو فترة زمنية معينة او مكانا معينا ، وعادة تكون فترة الاساس فترة سابقة للفترة التي نريد مقارنتها (وفي حالات نادرة جدا قد تكون فترة الاساس فترة لاحقة لفترة المقارنة) ويجب ان تمتاز فترة الاساس بما يلي :

- الاستقرار الاقتصادي
- خلوها من العوامل المؤثرة على الاسعار (الحروب مثلا)
- ان تكون بعيدة جدا عن سنوات المقارنة .
- اما عند اختيار مكان الاساس لابد ان يكون لهذا المكان اهمية خاصة وان يكون مركزا اساسيا لانتاج السلعة المراد استخراج الرقم القياسي لها .

❖ الأرقام القياسية للأسعار Price Index Numbers

تعتبر الارقام القياسية للاسعار من اهم انواع الارقام القياسية واكثرها شيوعا ، فهي (اي الارقام القياسية للاسعار) تساهم في قياس التغير في المستوى العام للاسعار او التغير في تكاليف المعيشة في فترة زمنية معينة مقارنة بفترة زمنية اخرى ومن اشهرها :

1) مؤشر أسعار المستهلكين Consumer Price Index ويرمز له (CPI)
2) مخفض الناتج القومي الإجمالي Gross National Product Deflator
3) مؤشر اسعار المنتجين Producer Price Index ويرمز له (PPI)
4) مخفض الناتج المحلي الإجمالي Gross Domestic Product Deflator
5) مؤشر اسعار الاسهم

❖ امثلة على بعض الارقام القياسية للاسعار في النظام الاقتصادي السعودي :

يهتم النظام الاقتصادي السعودي بنشر الارقام القياسية للاسعار وتكاليف المعيشة على شكل تقارير شهرية ، ومن هذه الارقام ما يلي :

أ. الرقم القياسي لتكاليف المعيشة لمتوسطي الدخل : ويشمل هذا الرقم المواد الغذائية ، السكن وتوابعه ، الاقمشة والملابس ، الاثاث المنزلية ، الرعاية الطبية ، النقل والمواصلات ، التعليم والترفية ، النفقات والخدمات الاخرى ، والرقم القياسي العام.

ب. الرقم القياسي لتكاليف المعيشة لجميع السكان : ويشمل المواد الغذائية ، السكن توابعه ، الاقمشة والملابس ، الاثاث المنزلي ، الرعاية الطبية ، النقل والاتصالات ، التعليم والترفيه ، النفقات والخدمات الاخرى ، الرقم القياسي العام .

ج. الرقم القياسي لاسعار الجملة : ويشمل المواد الغذائية ، المشروبات ، المواد الخام ما عادا الوقود ، الوقود المعدني وزيوت التشحيم ، الدهون والزيوت الحيوانية والنباتية ، الكيماويات والمواد ذات الصلة ، السلع المصنعة مصنفة حسب المادة ، الالات ومعدات النقل والاتصالات ، التعليم والترفيه ، النفقات والخدمات الاخرى ، الرقم القياسي العام .

دور الارقام القياسية في حساب معدلات التضخم :

المقصود بالتضخم هو الارتفاع المستمر في المستوى العام للاسعار والذي على ضوئه تنخفض القيمة الشرائية للوحدة النقدية (الريال مثلا)، وتقوم الجهات الاقتصادية في الدول باستخدام الارقام القياسية للاسعار لايجاد معدلات التضخم السنوية ، وفي معظم الاحيان يستخدم مؤشر اسعار المستهلكين (CPI) لسنتين متتاليتن لحساب معدل التضخم السنوي في السنة الاخيرة وذلك من خلال العلاقة التالية :

$$i_{2010} = \frac{CPI_{2010} - CPI_{2009}}{CPI_{2009}} \times 100$$

حيث :

i_{2010}	معدل التضخم في سنة 201 م
CPI_{2009}	مؤشر اسعار المستهلكين في سنة 2009 م (وهي سنة القياس ايضا)
CPI_{2010}	مؤشر اسعار المستهلكين في سنة 2010 م

⇐ مثال :

اذا افتراضنا ان مؤشر اسعار المستهلكين في المملكة لسنة 2006 م = 120 وسنة 2007م = 123 ، فما هو معدل التضخم في سنة 2007 م

✓ طريقة الحل :

معدل التضخم في سنة 2007 م يتم حسابه من خلال العلاقة التالية :

$$i_{2010} = \frac{CPI_{2010} - CPI_{2009}}{CPI_{2009}} \times 100 = \frac{123 - 120}{120} \times 100 = 2.5\%$$

اي ان معدل التضخم سنة 2007 يساوي 2.5 ٪

فوائد الارقام القياسية واستعمالاتها :

تستخدم الارقام القياسية عادة لقياس التغير الذي يطرأ على الحياة بمجملها بشكل عام والجوانب الاقتاصدية بشكل خاص ، كما تساعد الارقام القياسية على تحليل العوامل التي تساهم في تغير الظاهرة فتبين مدى مساهمة كل من هذه العوامل في احداث التغير الكلي ، وتستخدم كذلك في الرقابة على تنفيذ الخطط .

الرقم القياسي المرجح :

هو ذلك الرقم الذي يأخذ الاهمية النسبية للسلعة او الاجر بعين الاعتبار فيعطي كل سلعة (اجر) وزنا يتلائم مع اهميته ، فعند تركيب رقم قياسي للكميات يجب ترجيحه بالاسعار ، وعند تركيب رقم قياسي للاسعار يجب ترجيحه بالكميات وبالتالية يكون الناتج رقما قياسيا مرجحا .

❖ منسوب السعر لسلعة واحدة (ظاهرة بسيطة) :

يمكن ايجار رقم قياسي لسعر سلة بمفردها (حيث يمثل هذا الرقم القياسي التغير في سعر السلعة او الخدمة في سنة معينة مقارنة بسنة الاساس) ، ويسمى الرقم القياسي للسعر بمنسوب السعر ويرمز له بالرمز P_r ويمكن حسابه بالطريقة التالية :

$$P_r = \frac{P_1}{P_0} \times 100$$

حيث ان :

P_r منسوب السعر

P_1 السعر سنة المقارنة

P_0 السعر سنة الاساس

⇐ مثال :

اذا كانت لدينا البيانات التالية والممثلة لسعر سلعة معينة من الفترة 2006 م وحتى 2010م :

السنة	سعر السلعة بالريال
2006	25
2007	30
2008	24
2009	32
2010	36

◄ المطلوب: ايجاد منسوب السعر لهذه السلعة للفترة من سنة 2006 م وحتى سنة 2010 م باعتبار سنة الاساس هي 2006م ، مع تفسير النتائج التي يتم الحصول عليها .؟

✓ طريقة الحل :

يتم حساب قيمة منسوب السعر لهذه السلعة من خلال العلاقة التالية :

$$P_r = \frac{P_1}{P_0} \times 100$$

مع اعتبار ان P_0 هو سعر السلعة في سنة 2006 (سنة الاساس) ، وبتطبيق المعادلة السابقة يمكن بالتالي تلخيص النتائج في الجدول التالي :

السنة	سعر السلعة بالريال	منسوب السعر
2006	25	$P_r = \frac{25}{25} \times 100 = 100\ \%$
2007	30	$P_r = \frac{30}{25} \times 100 = 120\ \%$
2008	24	$P_r = \frac{24}{25} \times 100 = 96\ \%$
2009	32	$P_r = \frac{32}{25} \times 100 = 128\ \%$
2010	36	$P_r = \frac{36}{25} \times 100 = 144\ \%$

تفسير النتائج : ان منسوب السعر لسنة 2007م والذي يساوي 120 ٪ يوضح ان هناك زيادة في سعر السلعة بنسبة 20 ٪ في سنة 2007م مقارنة بسنة 2006 م ، كما ان منسوب السعر لسنة 2008م والذي يساوي 96 يعني ان سعر السلعة انخفض في سنة 2008م بنسبة 4٪ مقارنة بسنة 2006م (سنة الاساس) .

❖ منسوب السعر لمجموعة من السلع – التجميعية (ظاهرة معقدة) :

الرقم القياسي في المثال السابق يوضح منسوب السعر لسلعة واحدة ، الا ان كثير من الحالات تكون اكثر تعقيدا فقد يكون لدينا عدة سلع متغيرة ونرغب في حساب منسوب السعر او الرقم القياسي لها ، ففي حالة استخراج الرقم القياسي لمثل هذا الوضع فانه يدخل في الحساب جميع قيم السلع التي تتألف منها الظاهرة ويتم ذلك من خلال استخدام الطرق التالية :

1) الرقم القياسي التجميعي البسيط للاسعار
2) الرقم القياسي التجميعي للاسعار المرجح بكميات سنة الاساس (رقم لاسبير)
3) القرم القياسي التجميعي للاسعار المرجح بكميات سنة المقارنة (رقم باش)
4) القرم القياسي التجميعي المرجح بكميات سنة الاساس وسنة المقارنة (رقم فيشر)

❖ حساب الارقام القياسية التجميعية (مجموعة من السلع) :

1) الرقم القياسي التجميعي البسيط للاسعار : نرمز لهذا الرقم القياسي بالرمز I_s ويتم حسابه من خلال العلاقة التالية :

$$I_s = \frac{\sum P_1}{\sum P_0} \times 100$$

حيث ان :

$\sum P_1$	مجموع اسعار السلع والخدمات في سنة المقارنة .
$\sum P_0$	مجموع اسعار السلع والخدمات في سنة الاساس

وتكمن مشكلة الرقم القياسي التجميعي البسيط للاسعار في انه لا يعطي للكميات المستهلكة من السلع والخدمات اوزانا ، فبالتالي يكون حساسا عندما يكون هناك تباين في الكميات المستهلكة .

2) الرقم القياسي التجميعي للاسعار المرجحج بكميات سنة الاساس (رقم لاسبير) : يرمز له بالرمز I_r وهذا القرم يعبر عن اثر التغير في السعر كما لو بقيت الكميات المشترة في سنة الاساس هي نفسها في سنة المقارنة ، ويتم حسابه بنفس الطريقة السابقة مع ترجيح وزن كل سعر بكميته المستهلكة في سنة الاساس ، ويتم ذلك من خلال تطبيق العلاقة التالية :

$$I_r = \frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times 100$$

حيث ان :

$\sum P_0 Q_0$	مجموع اسعار السلع والخدمات سنة المقارنة مرجحة بكميات سنة الاساس
$\sum P_1 Q_0$	مجموع اسعار السلعة والخدمات سنة الاساس مرجحة بكميات سنة الاساس

ويفضل استخدام هذه الطريقة عند حساب مؤشر اسعار المستهلكين (CPI) وذلك للاقتصاد في الجهد والقوت والمال ، لان كمية سنة الاساس ثابتة عند ايجاد رقم لاسبير لأي سنة لاحقة لسنة الاساس .

3) الرقم القياسي التجميعي للاسعار المرجح بكميات سنة المقارنة (رقم باش) : و يرمز له بالرمز I_p
وهذا الرقم يعبر عن اثر التغير في السعر كما لو ان الكميات المشتراة في سنة المقارنة كانت قد اشتريت في سنة الاساس . وتختلف طريقة حساب هذا الرقم من حيث انه يرجح كل سعر بكميته المستهلكة في سنة المقارنة ، ويتم ذلك من خلال تطبيق العلاقة التالية :

$$I_p = \frac{\sum P_1 Q_1}{\sum P_0 Q_1} \times 100$$

حيث ان :

$\sum P_1 Q_1$	مجموع اسعار السلع والخدمات سنة المقارنة مرجحة بكميات سنة المقارنة
$\sum P_0 Q_1$	مجموع اسعار السلع والخدمات سنة الاساس مرجحة بكميات سنة المقارنة

المشكلة الاساسية في هذه الطريقة هي الحاجة لتحديد الكميات المستهلكة من كل سلعة سنويا حتى يتسنى لنا حساب هذا الرقم .

4) الرقم القياسي التجميعي للاسعار المرجح بكميات سنة الاساس وسنة المقارنة (رقم فيشر) :
ويرمز له بالرمز I_f وهو عبارة عن الوسط الهندسي لكل من رقمي لاسبير وباش ، اي انه الجذر التربيعي لحاصل ضرب رقم لاسبير برقم باش ، ()، ويتم ذلك من خلال العلاقة التالية :

$$I_f = \sqrt{I_r \, I_p}$$

$$I_f = \sqrt{\frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times \frac{\sum P_1 Q_1}{\sum P_0 Q_1}}$$

اهم عيوبه انه هذا الرقم يهتم بالناحية الرياضية ولكنه لا معنى له اقتصاديا

⇐ مثال لحساب الارقام القياسية التجميعية :

يبين الجدول التالي اسعار وكميات ثلاث منتجات استهلاكية للسنتين 2007م و 2010م على اعتبار ان سنة 2007م هي سنة الاساس :

المنتجات / السنوات	سنة 2007م (سنة الاساس)		سنة 2010 م (سنة المقارنة)	
	الكمية Q_0	السعر P_0	الكمية Q_1	السعر P_1
السلعة الاولى	5000	9	8500	12
السلعة الثانية	8000	25	15000	31
السلعة الثالثة	9000	14	19000	17

◄ المطلوب: اوجد التالي :

1- حساب الرقم التجميعي البسيط للأسعار.
2- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير).
3- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش).
4- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر).
5- تفسير نتائج الفقرات السابقة.

✓ طريقة الحل :

يمكن من خلال بيانات الجدول السابق اعداد الجدول التالي :

المنتجات / السنوات	سنة 2007م (سنة الاساس)		سنة 2010 م (سنة المقارنة)		P_0Q_0	P_1Q_0	P_0Q_1	P_1Q_1
	الكمية Q_0	السعر P_0	الكمية Q_1	السعر P_1				
السلعة الاولى	5000	9	8500	12	45000	60000	76500	102000
السلعة الثانية	8000	25	15000	31	200000	248000	375000	465000
السلعة الثالثة	9000	14	19000	17	126000	153000	266000	323000
المجموع		48		60	371000	461000	717500	890000

1- الرقم القياسي التجميعي البسيط : ويتم حساب الرقم القياسي التجميعي البسيط من خلال تطبيق العلاقة التالية :

$$I_s = \frac{\sum P_1}{\sum P_0} \times 100 = \frac{60}{48} \times 100 = 125\ \%$$

التفسير : هذا يعني ان المستوى العام لاسعار المنتجات الثلاث قد ارتفع في سنة 2010 م بمعدل 25٪ وذلك مقارنة بسنة 2007 م

2- الرقم القياسي التجميعي للاسعار المرجح بكميات سنة الاساس (رقم لاسبير) : يمكن ايجاد مجموع الاسعار سنة المقارنة المرجحة بكميات سنة الاساس $\sum P_1Q_0$ من خلال ضرب خلايا العمود P_1 في خلايا العمود Q_0 ثم نجمع الناتج وهو 461000 وكذلك يمكن ايجاد مجموع اسعار سنة الاساس المرجحة بكميات سنة الاساس $\sum P_0Q_0$ وذلك من خلال ضرب خلايا العمود P_0 في خلايا العمود Q_0 ثم نجمع الناتج وهو 371000 .

ويتم حساب الرقم القياسي التجميعي للاسعار المرجح بكميات سنة الاساس (رقم لاسبير) من خلال تطبيق العلاقة التالية :

$$I_r = \frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times 100 = \frac{461000}{371000} \times 100 = 124.2588\%$$

التفسير : هذا يدل على ان المستوى العام لاسعار المنتجات الثلاث قد ارتفع في سنة 2010 م بمعدل 24.25 ٪ وذلك مقارنة بسنة 2007 م .

3- الرقم القياسي التجميعي للاسعار المرجح بكميات سنة المقارنة (رقم باش) : يمكن ايجاد مجموع الاسعار سنة المقارنة المرجحة بكميات سنة المقارنة $\sum P_1 Q_1$ من خلال ضرب خلايا العمود P_1 في خلايا العمود Q_1 ثم نجمع الناتج وهو 890000 وكذلك يمكن ايجاد مجموع اسعار سنة الاساس المرجحة بكميات سنة المقارنة $\sum P_0 Q_1$ من خلال ضرب خلايا العمود P_0 في خلايا العمود Q_1 ثم نجمع الناتج وهو 717500 ويتم حساب الرقم القياسي التجميعي للاسعار المرجح بكميات سنة المقارنة (رقم باش) من خلال :

$$I_p = \frac{\sum P_1 Q_1}{\sum P_0 Q_1} \times 100 = \frac{890000}{717500} \times 100 = 124.0418\%$$

التفسير : هذا يدل على ان المستوى العام لاسعار المنتجات الثلاث قد ارتفع في سنة 2010م بمعدل 24.0418 ٪ مقارنة بسنة 2007م.

4- الرقم القياسي التجميعي للاسعار المرجح بكميات سنة الاساس وسنة المقارنة (رقم فشر) الرقم القياسي الامثل : ويتم حساب هذا الرقم من خلال حساب الوسط الهندسي لرقم لاسبير و باش من خلال العلاقة التالية :

$$I_f = \sqrt{124.2588 \times 124.0418} = 124.1502\%$$

التفسير: هذا يعني ان المستوى العام لاسعار المنتجات الثلاث قد ارتفع في سنة 2010 م بمعدل 24.15 ٪ وذلك مقارنة بسنة 2007 م .

❖ ملاحظات عامة على الارقام القياسية :

هناك مجموعة من الملاحظات المتعلقة بتفسير الارقام القياسية لسنوات الاساس والمقارنة وهذه الملاحظات كالتالي :

1) الرقم القياسي للظاهرة في سنة الاساس يساوي 100 .
2) اذا كان الرقم القياسي للظاهرة في سنة المقارنة اكبر من 100 فهذا يعني ان هناك ارتفاع في المستوى العام للظاهرة مقارنة بسنة الاساس .
3) اذا كان الرقم القياسي للظاهرة في سنة المقارنة اصغر من 100 فهذا يعني ان هناك انخفاض في المستوى العام للظاهرة مقارنة بسنة الاساس .