



جامعة الملك فيصل

عمادة التعلم الإلكتروني والتعليم عن بعد

ملخص مبادئ الإحصاء

د. عبدالله عمر النجار

هذا الملخص مطابق للمحتوى بدون اضافة أو حذف

ولا يغني عن الاستماع للمحاضرات وتدوين حل المسائل والجداول والرجوع أن لزم الأمر للكتاب

الاستماع للمحاضرات جدا مهم وذلك للنقاط المهمة التي يذكرها الدكتور

والقوانين المستخدمة في الحل

المحاضرة التمهيدية

تعريف بمقرر مبادئ الإحصاء ومحتوياته

هدف المقرر العام:

يهدف المقرر إلى إلمام الطالب / الطالبة من طلبة كلية الآداب بأساسيات علم الإحصاء مع التركيز بصفة خاصة على أساسيات الإحصاء الوصفي الخاصة بجمع البيانات وتبويبها وعرضها وتحليلها بما يخدم طلبة كلية الآداب.

أهداف المقرر الإجرائية:

- التعرف ببعض المفاهيم الإحصائية
- فهم عملية جمع البيانات وترميزها
- التعرف بالعرض الجدولي للبيانات (تبويب البيانات)
- التعرف على الطريقة المثلى للعرض البياني للبيانات
- التعرف بالمقاييس الإحصائية للبيانات غير المبوبة والبيانات المبوبة ومقاييس الالتواء والتفرطح وتحليل الارتباط

محتويات المنهج:

- مفاهيم إحصائية أساسية
- جمع البيانات وترميزها
- العرض الجدولي للبيانات
- العرض البياني للبيانات
- المقاييس الإحصائية للبيانات غير المبوبة:
- مقاييس النزعة المركزية (المتوسط – الوسيط – المنوال)
- مقاييس التشتت (المدى – الإنحراف عن المتوسط – التباين – الإنحراف المعياري)

المقاييس الإحصائية للبيانات المبوبة

- مقاييس النزعة المركزية (المتوسط – الوسيط – المنوال)
- مقاييس التشتت (المدى – الإنحراف عن المتوسط – التباين – الإنحراف المعياري)

مقاييس الالتواء والتفرطح

تحليل الارتباط :

معامل ارتباط بيرسون

معامل ارتباط سبيرمان

المحاضرة الأولى

التعريف ببعض المفاهيم الإحصائية

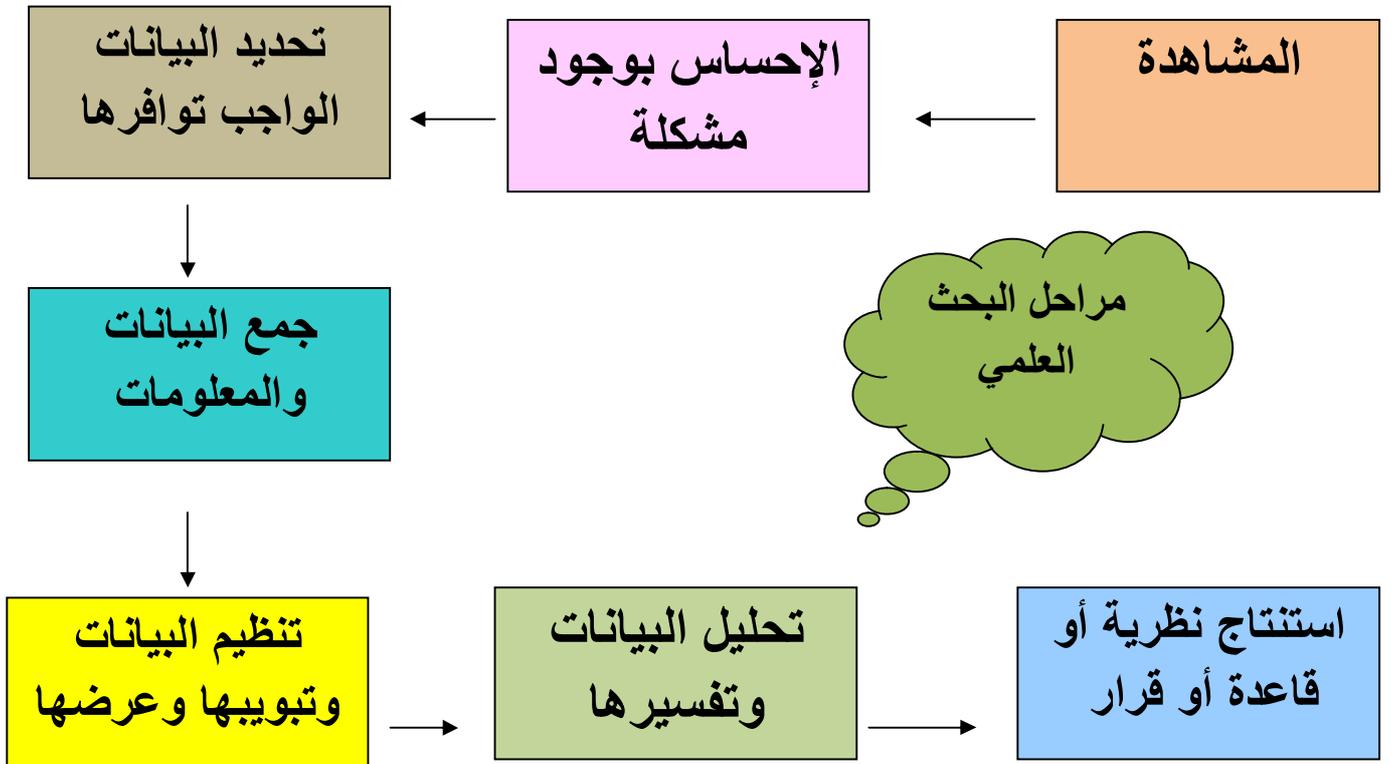
مقدمة

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

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

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

والشكل التالي يمكن أن يوضح الإطار العام لأي بحث علمي



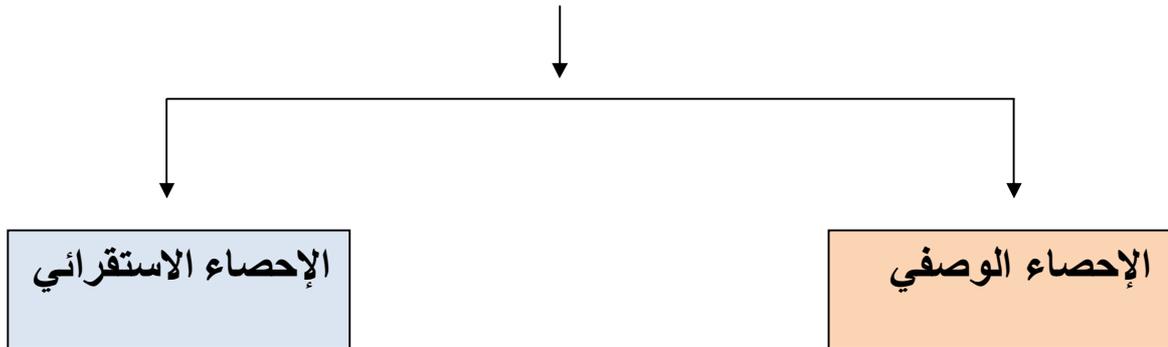
مفهوم علم الإحصاء

يختص علم الإحصاء بالطرق العملية لجمع وتنظيم وعرض وتحليل البيانات وكذلك الوصول إلى نتائج مقبولة وقرارات سليمة على ضوء هذا التحليل

وقديماً عُرف علم الإحصاء على أنه جمع البيانات عن ظاهرة معينة وترتيبها في جداول أو عرضها في صورة رسومات وأشكال بيانية بسيطة، ومن ثم استخدم اصطلاح "علم الإحصاء" للتعبير عن البيانات والمقاييس المستخرجة من تلك البيانات (مثل المتوسطات)، وعلى هذا الأساس نتحدث عن إحصاءات البطالة والحوادث والمواليد والوفيات ، ... إلخ

لكن في حقيقة الأمر هذا استخدام ذي معنى ضيق لاصطلاح "علم الإحصاء"، لكن مع تقدم العلوم بدأ علم الإحصاء يلعب دوراً متزايداً في حياتنا اليومية بحيث أصبح يشغل حيزاً كبيراً بين بقية العلوم الأخرى، فأصبح يبحث في جمع البيانات وتنظيمها وعرضها وتحليلها واستنتاج وتوقع نتائج واتخاذ قرارات

وينقسم علم الإحصاء إلى قسمين رئيسيين :



أو الاستدلال الإحصائي وهو يبحث في استقراء النتائج واتخاذ القرارات

وهو يهتم بجمع وتبويب وعرض ووصف البيانات وحساب بعض المقاييس الخاصة بها دون الوصول إلى نتائج أو استدلالات خاصة

المجتمع والعينة

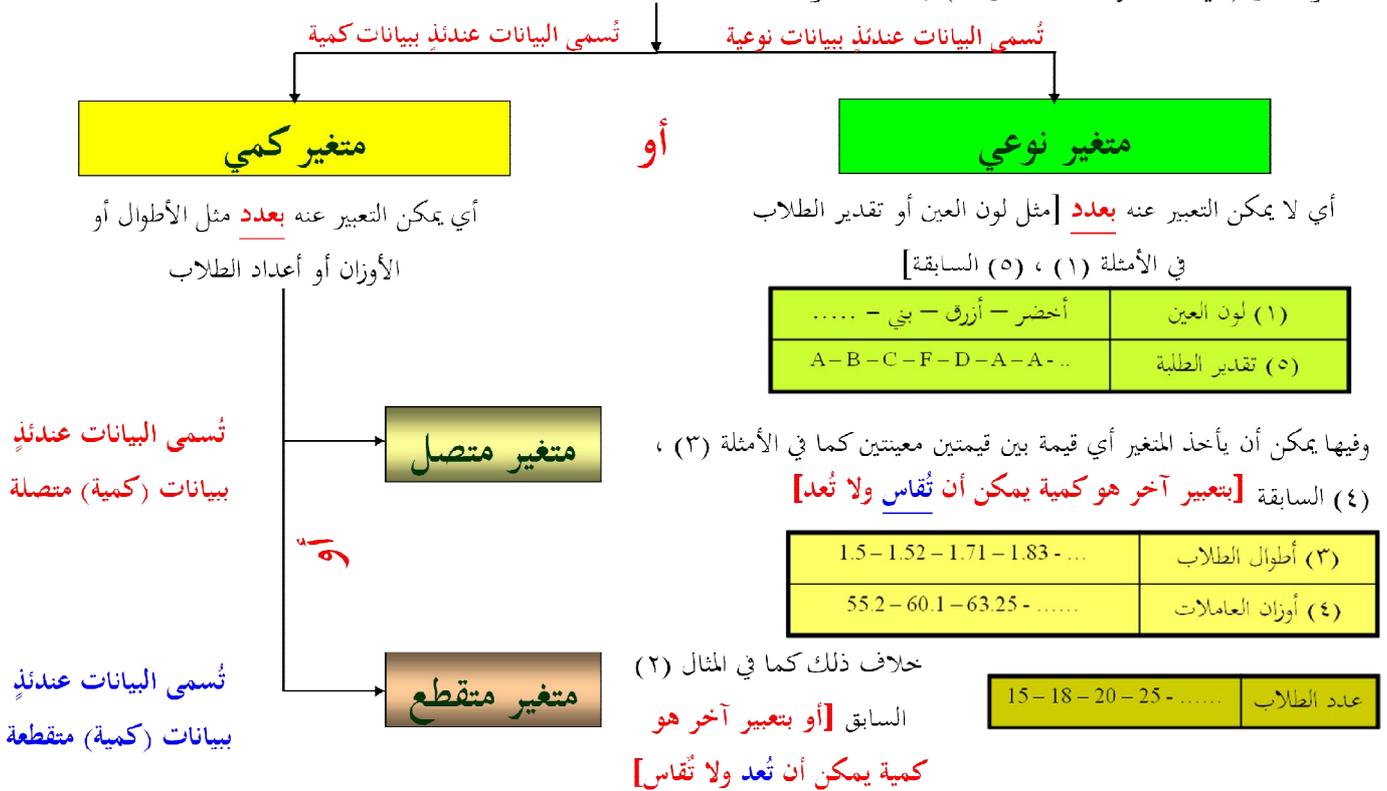
مثلاً لتحليل نتائج طلاب المملكة في مقرر اللغة الإنجليزية لطلاب وطالبات الثانوية العامة، فمن المستحيل أو غير العملي أن نقوم بجمع درجات جميع الطلاب في هذا المقرر على مستوى المملكة وتنظيمها وتحليلها ثم نستنتج بعض النتائج من هذا التحليل، هنا يكون **المجتمع** هو جميع طلاب المملكة بدلاً من ذلك نقوم باختيار **عينة** من هؤلاء الطلاب (تحت شروط معينة حتى تكون ممثلة للمجتمع) ونقوم بتحليل بيانات هذه العينة ونخرج من هذا التحليل باستدلالات تخص المجتمع ككل

البيانات

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

المتغير x	البيانات (القياسات أو المشاهدات)	العملية الإحصائية : دراسة
لون العين	أخضر - أزرق - بني -	لون العين لبعض الأطفال حديثي الولادة
عدد الطلاب	15 - 18 - 20 - 25 - 17 -	عدد الطلاب في فصول مدرسة
طول الطالب	1.5 - 1.52 - 1.71 - 1.83 -	أطوال مجموعة من الطلاب في فصل ما (بالمتر)
وزن العاملة	55.2 - 60.1 - 63.35 - 70.52 -	أوزان بعض العاملات بمصنع معين (بالكيلوجرام)
تقدير الطالب	A - B - C - D - F - A - C - B -	تقديرات عدد من الطلاب في مقرر الإحصاء

والمتغير (أي الظاهرة تحت الدراسة) إما أن يكون :



خطوات العملية الإحصائية

يمكن تلخيص خطوات أي عملية إحصائية في الآتي :

(أ) جمع البيانات

هي عملية الحصول على القياسات الخاصة
السابقة في بظاهرة معينة وعادةً ما تُسمى

البيانات المجمعة بالبيانات الخام

(ج) تحليل البيانات

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

(ب) تنظيم وعرض البيانات

هي عملية وضع البيانات
جداول خاصة وعرضها بطرق مناسبة

(د) استقراء النتائج واتخاذ القرارات

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

تمارين محلولة

اختر الإجابة الصحيحة لكل مما يأتي :

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

(ب) علم الإحصاء الاستقرائي

(أ) علم الإحصاء الوصفي

(د) علم تكنولوجيا المعلومات

(ج) علم تقنية المعلومات

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

(ب) استقراء النتائج واتخاذ القرارات

(أ) تحليل البيانات

(د) جمع البيانات

(ج) تنظيم وعرض البيانات

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

(أ) تحليل البيانات (ب) استقراء النتائج واتخاذ القرارات

(ج) تنظيم وعرض البيانات (د) جمع البيانات

(٤) عدد الأيام N في كل شهر هو :

(أ) متغير نوعي (ب) متغير كمي متصل

(ج) متغير كمي متقطع (د) خلاف ذلك

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

(أ) متغير نوعي (ب) متغير كمي متصل

(ج) متغير كمي متقطع (د) خلاف ذلك

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

(أ) بيانات نوعية (ب) بيانات كمية متصلة

(ج) بيانات كمية متقطعة (د) خلاف ذلك

(٧) البيانات المجمعة عن الدخل السنوي لمنسوبي إحدى الهيئات الحكومية هي :

(أ) بيانات نوعية (ب) بيانات كمية متصلة

(ج) بيانات كمية متقطعة (د) خلاف ذلك

تدريبات للطالب

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

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

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

(٢) هي عملية الوصول إلى استنتاجات وتوقعات وتنبؤات خاصة بظاهرة معينة

(أ) تحليل البيانات (ب) استقراء النتائج واتخاذ القرارات

(ج) تنظيم وعرض البيانات (د) جمع البيانات

(٣) هي عملية إيجاد قيم لمقاييس تتحدد قيمها من البيانات الخاصة بظاهرة معينة وتُعطى بعض الدلالات عن تلك الظاهرة

(أ) تحليل البيانات (ب) استقراء النتائج واتخاذ القرارات

(ج) تنظيم وعرض البيانات (د) جمع البيانات

(٤) المسافة d (بالكيلومتر) التي يقطعها شخص يومياً من بيته لمكان عمله هي :

(أ) متغير نوعي (ب) متغير كمي متصل

(ج) متغير كمي متقطع (د) خلاف ذلك

(٥) وزن البطاطس W (بالكيلوجرام) التي تنتجها مزارع مختلفة في سنة معينة هو :

(أ) متغير نوعي (ب) متغير كمي متصل

(ج) متغير كمي متقطع (د) خلاف ذلك

(٦) عدد حبات البطيخ N التي تبيعها محلات سوبر ماركت مختلفة يوم الجمعة هو :

(أ) متغير نوعي (ب) متغير كمي متصل

(ج) متغير كمي متقطع (د) خلاف ذلك

(٧) الزمن t الذي يأخذه كل طالب في كليتك لحل اختبار مقرر الإحصاء هو :

(أ) متغير نوعي (ب) متغير كمي متصل

(ج) متغير كمي متقطع (د) خلاف ذلك

(٨) مقياس الأحذية S هو :

(أ) متغير نوعي (ب) متغير كمي متصل

(ج) متغير كمي متقطع (د) خلاف ذلك

(٩) اللعبة الرياضية A التي يفضلها أفراد أسرته هي :

(أ) متغير نوعي (ب) متغير كمي متصل

(ج) متغير كمي متقطع ○ (د) خلاف ذلك ○

(١٠) البيانات المجمعة عن نوع السيارات في موقف ما ، هي :

(أ) بيانات نوعية ○ (ب) بيانات كمية متصلة ○

(ج) بيانات كمية متقطعه ○ (د) خلاف ذلك ○

(١١) البيانات المجمعة عن النسبة المئوية لدرجات الطلاب في أحد المقررات الدراسية هي :

(أ) بيانات نوعية ○ (ب) بيانات كمية متصلة ○

(ج) بيانات كمية متقطعه ○ (د) خلاف ذلك ○

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

(أ) بيانات نوعية ○ (ب) بيانات كمية متصلة ○

(ج) بيانات كمية متقطعه ○ (د) خلاف ذلك ○

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

(أ) بيانات نوعية ○ (ب) بيانات كمية متصلة ○

(ج) بيانات كمية متقطعه ○ (د) خلاف ذلك ○

الإجابة :

١) ب (٢) ب (٣) أ (٤) ب (٥) ب (٦) ج (٧) ب (٨) ج (٩) أ

١٠) أ (١١) ب (١٢) ب (١٣) أ

المحاضرة الثانية

أساليب إجراء البحث الميداني

أساليب إجراء البحث الميداني :

هناك سؤال مهم لا بد من الإجابة عليه وهو:

هل تشمل الدراسة جميع مفردات المجتمع الإحصائي أم سيطبق على جزء منه؟

في حالة اعتماد البحث على دراسة جميع مفردات المجتمع الإحصائي يسمى ذلك أسلوب **(الحصص الشامل)**

أما إذا اعتمد البحث على دراسة جزء فقط من مفردات المجتمع الإحصائي يسمى ذلك

(أسلوب العينة)

أسلوب الحصر الشامل :

يمكننا هذا الأسلوب من الحصول على كافة البيانات والمعلومات عن كافة مفردات المجتمع الإحصائي وبالتالي فإن النتائج التي نحصل عليها لا يوجد بها تحيز ولا تحتاج لتعديل لكنها تحتاج إلى وقت وجهد كبيرين

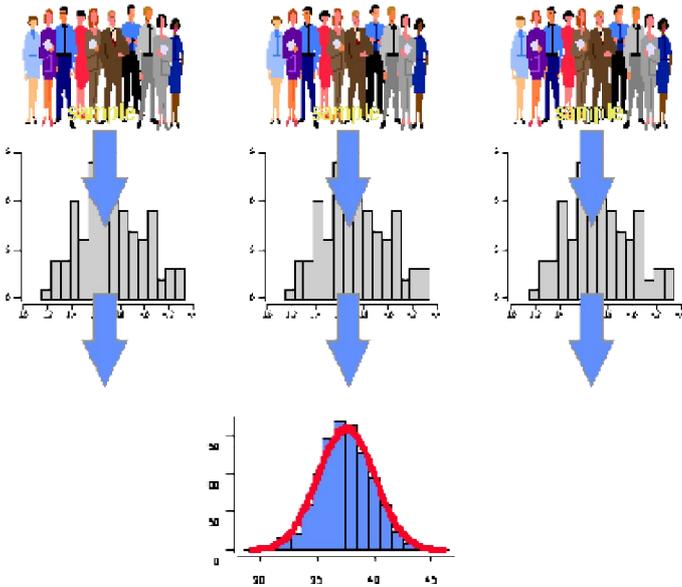
أسلوب العينات:

يبدوا هذا الأسلوب على العكس من أسلوب الحصر الشامل حيث تقتصر الدراسة فيه على جزء من المجتمع الإحصائي، لذا فهذا الأسلوب يوفر الوقت و الجهد و التكاليف ويصلح للمجتمعات غير المحدودة. إلا ان أهم عيوب هذا النوع هو ما يسمى بخطأ التحيز Sampling Bias .

أقسام مجتمع البحث:

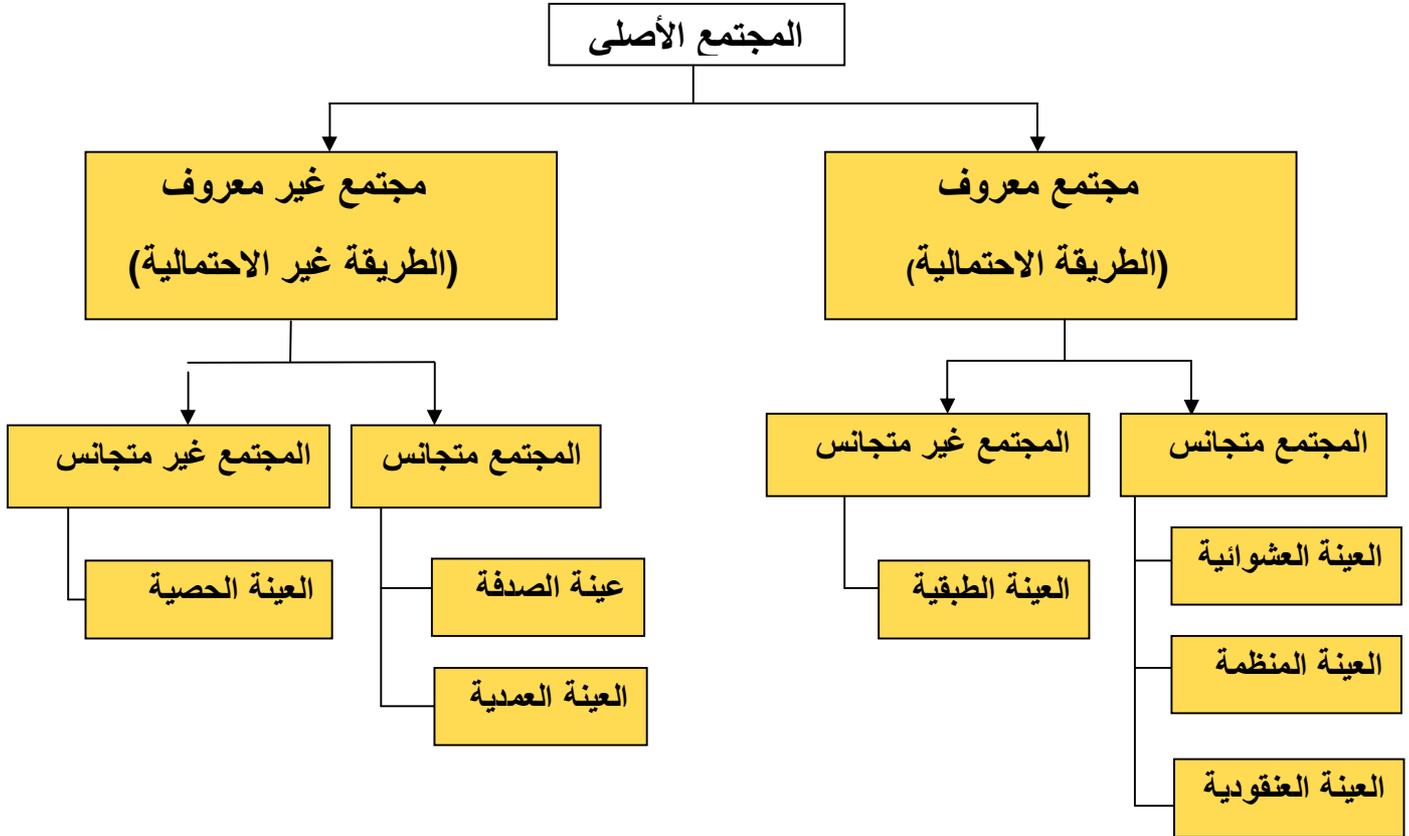
قسم بعض العلماء مجتمع البحث الى قسمين:

- المجتمع الكلي للبحث
- المجتمع الذي يمكن التعرف عليه



مجتمع البحث : هو مصطلح علمي يراد به كل من يمكن أن تعمم عليه نتائج البحث
عينة البحث : بأنها جزء من المجتمع اختير بطريقة علمية بشرط ان تمثل المجتمع ككل

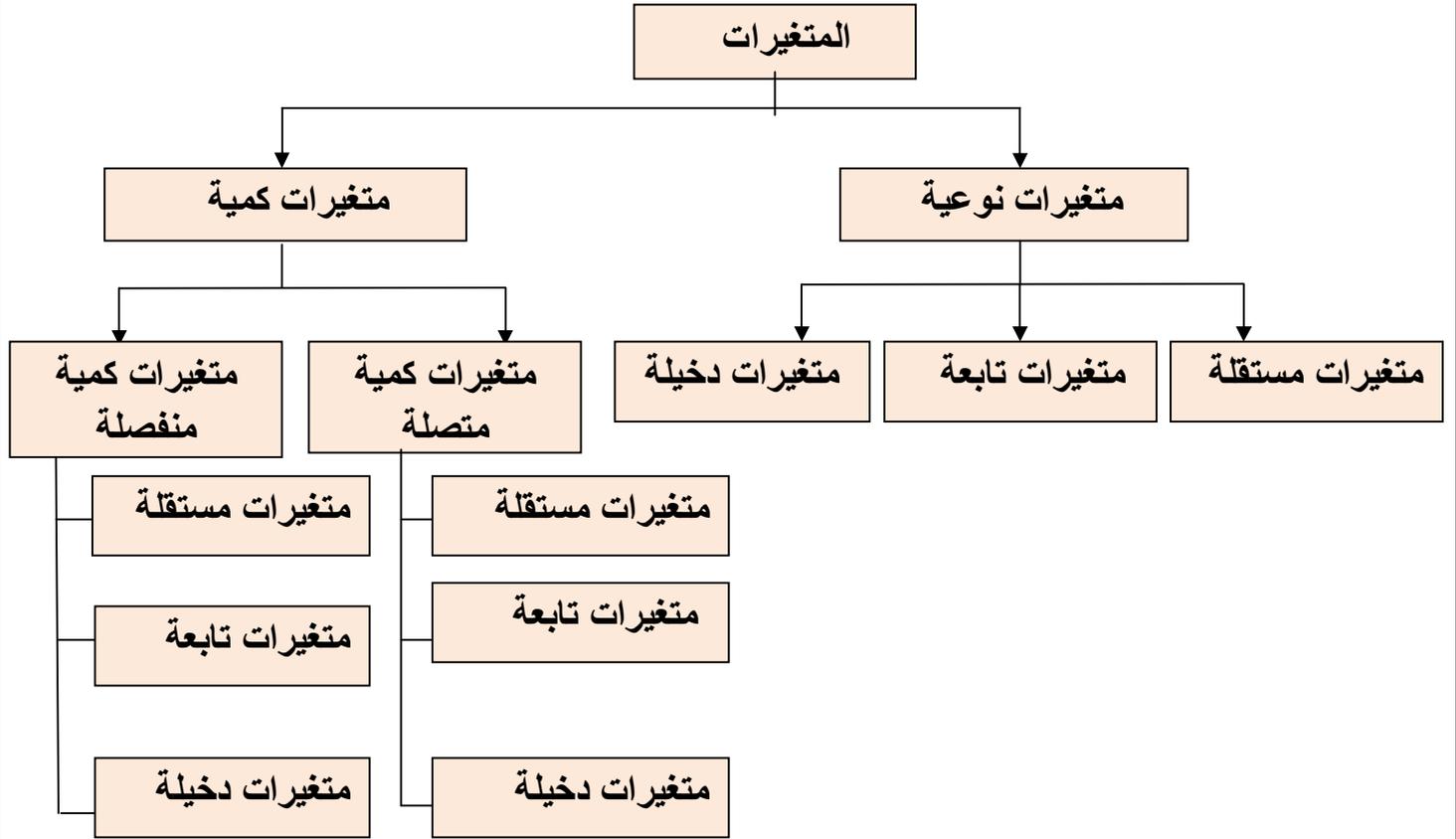
طرق اختيار العينات



المتغير والثابت في البحث العلمي:

- ▶ **المتغير:** هو أي خاصية أو صفة سواء للأفراد أو الأشكال والتي تختلف من شخص لآخر ومن وقت لآخر مثل الطول، الذكاء ، التحصيل ويعمل الباحث على دراستها وقياسها.
- ▶ **الثابت:** هي الصفات أو الظواهر التي لا تتغير، أو أي صفة أو خاصية تأخذ صفة واحدة ومن الممكن أخذ متغير وتحويله الى ثابت مثل درجة الحرارة في الغرفة. والباحث يسعى الى تثبيت عدد من المتغيرات في دراسته للتخلص من تأثيرها .

تصنيف المتغيرات:



الخطوات الواجب مراعاتها بعد جمع البيانات:

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

- تسجيل البيانات

- ترميز البيانات

١- الترميز الرقمي أو العددي

٢- الترميز الأبجدي أو الحرفي

٣- الترميز الأبجدي الرقمي

- تصنيف البيانات

- مراجعة وتنقية البيانات

ترميز بيانات الاستبانة وجعلها متاحة لبرنامج ال-SPSS:

تعتبر الاستبانة من أكثر وسائل جمع البيانات البحثية استخداماً، لذلك سوف نقوم الآن بالتعرف على كيفية توييب 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 ستحتاجين إلى إعداد استبانة تحوي مجموعة من الاسئلة تتعلق بهذا الموضوع، ومن ثم توزيع هذه الاستبانة على عينة ممثلة لمجتمع البحث الذي تريد أن تعممي نتائج دراستك عليه، وتطلبين من أفراد العينة الإجابة على جميع فقرات الاستبانة، والاستبانة التالية (والتي ستوزع عليكم) كمثال على ذلك. ولغرض تفرغ البيانات المجموعة من خلال هذه الاستبانة بطريقة مناسبة يفهمها برنامج الـ SPSS لابد من توضيح التالي :

◦ الأفراد الذين يقومون بالإجابة على أسئلة الاستبانة يطلق عليهم اسم حالات Cases

◦ كل سؤال (فقرة) في الاستبانة تمثل متغير Variable

◦ تسمى إجابات الافراد على الاسئلة (الفقرات) بقيم المتغيرات Variable values

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

١- سؤال يسمح باختيار إجابة واحدة فقط

٢- سؤال يسمح باختيار أكثر من إجابة واحدة

٣- سؤال مفتوح جزئياً

تمارين

أرادت باحث معرفة العلاقة بين حب الاستطلاع لدى الطلاب في السنوات الابتدائية وحل المسائل الرياضية، فاختار عشوائياً طلاب السنة الثالثة ثم اختار منهم عشوائياً ٢٠٠ 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 ؟
- حدد مجتمع البحث في هذه الدراسة ، وما نوعه ؟
- حدد عينة الدراسة في هذه الدراسة ، وما نوعها؟
- حدد المتغير المستقل في هذه الدراسة ، وما نوعه ؟

- حدد المتغير التابع في هذه الدراسة ، وما نوعه ؟
- حدد في صورتك 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 الثالثة

العرض الجدولي للبيانات - ١

إن الصورة التي يعرض بها الباحث بياناته تعكس لدرجة كبيرة مدى امكانية فهمها وسهولة تتبعها والاستفادة منها.

وهناك عدة طرق لعرض وتبويب البيانات الا أن من أبسط تلك الطرق للتعبير عن البيانات هي أن تدمج هذه البيانات في صيغة كتابية إلا أن هذه الطريقة يشوبها الكثير من العيوب

أما الطرق الفنية في عرض البيانات الاحصائية فهي:

- العرض الجدولي للبيانات
- العرض البياني للبيانات

وسوف نتناول في هذه المحاضرة العرض الجدولي للبيانات بينما نتعرض للعرض البياني للبيانات في المحاضرة التالية إن شاء الله تعالى.

ويقصد بالعرض الجدولي للبيانات أن يتم تلخيص البيانات محل الدراسة وتصنيفها في صورة جداول تعبر عن القيم التي أخذها المتغير من خلال البيانات التي جمعها و تكرار كل قيمة من تلك القيم. أهمية الجداول الاحصائية:

- تعبر عن الحقائق الكمية المعروضة بعدد كبير من الارقام في جداول بطريقة منظمة
- تلخيص المعلومات الرقمية الكثيرة العدد، المتغيرة القيم، مما يسهل التعرف عليها.
- الاستيعاب وبسهولة عدد كبير من الموضوعات
- اظهار البيانات بأكثر وضوح ممكن وأصغر حيز مستطاع

تكوين الجداول:

تتكون اجزاء الجدول مما يلي:

- **رقم الجدول:** يجب ان يرقم كل جدول حتى تسهل الاشارة اليه.
- **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.

- **الهيكل الرئيسي:** ويتكون هيكل الجدول من أعمدة وصفوف، ويعتبر ترتيب المعلومات في الأعمدة والصفوف أهم خطوة في تكوين الجدول.
- **العمود:** إن كل جدول يتكون من عمود أو أكثر ويوجد لكل عمود عنوان يوضح محتوياته.
- **الحواشي:** قد يحتوي الجدول على مفردات بيانات لا ينطبق عليها عنوان الجدول أو عنوان العمود، ففي هذه الحالة تستعمل الحواشي لتوضيح ذلك وذلك إما بترقيم الملاحظات أو باستعمال علامة (*) .. الخ.
- **المصدر:** قد تؤخذ بيانات الجدول من مصادر جاهزة لذلك يجب إظهار المصدر في أسفل الجدول حتى يمكن الرجوع إليه عند الحاجة.

رقم الجدول جدول رقم (٥)

عنوان الجدول يوضح طلبة جامعة الملك فيصل للعام الجامعي ١٤٢٣ هـ

عنوان توضيحي (مصنفون حسب الجنس)

عنوان العمود	المستوى*	طالب	طالبة	المجموع
هيكل الجدول	الأول	٢٠٠	٢٥٠	٤٥٠
	الثاني	١٠٠	١٢٠	٢٢٠
	الثالث	٨٠	١١٠	١٩٠
	الرابع	١٠٠	١٢٠	٢٢٠
	المجموع	٤٨٠	٦٠٠	١٠٨٠

المصدر: جامعة الملك فيصل، احصائية الجامعة حسب الكليات
* يحدد المستوى بالسنة الدراسية التي يدرس فيها الطالب .

أنواع الجداول الاحصائية:

تقسم الجداول تبعاً لدرجة تعقيدها إلى:

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 المكانية (أي توزيع الظاهرة حسب المكان) أو مؤشرات وصفية بسيطة وأرقام بسيطة أيضاً.

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

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

الجدول المزدوجة أو المركبة: وهي الجداول التي تتكون من متغيرين أو أكثر، وهذه المتغيرات قد توزع على أعمدة وحقول الجدول بصورة نظامية، تعبر عن الأفكار العلمية التي يريد الباحث توضيحها توضيحاً عددياً.

وقد أوضحنا في المحاضرة السابقة ما هي البيانات وعرفناها بأنها **[هي مجموعة المشاهدات أو القياسات التي تخص ظاهرة معينة تحت الدراسة]**

وعرفنا كذلك المتغير على أنه تلك الكمية التي نقوم بمشاهدتها أو قياسها ، كما ذكرنا أن البيانات إما أن تكون نوعية أو كمية ، حيث :

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

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

- لون (أو نوع) السيارات الموجودة في موقف ما [أحمر – أبيض – أسود -]
- الحالة الاجتماعية للسيدات في محافظة معينة [متزوجة – عزباء – مطلقة – أرملة – منفصلة]
- وغيره من مثل هذه الأمثلة .

(ب) البيانات الكمية : هي تلك البيانات التي يُعبر فيها عن المتغير بعدد (أي بيانات رقمية) ، وهذه البيانات بدورها تنقسم إلى :

(ب - ١) بيانات كمية متصلة : وفيها يمكن أن يأخذ المتغير أي قيمة بين قيمتين (أي بيانات يمكن أن تُقاس ولا تُعد ، مثل :

أطوال الطلاب في إحدى المدارس

أوزان العوامل بإحدى المصانع .

الدخل السنوي لمنسوبي مؤسسة معينة

وغيره من مثل هذه الأمثلة .

(ب - ٢) بيانات كمية متقطعة : وفيها يمكن أن يأخذ المتغير قيمة رقم صحيح بدون كسور (مثلا إما ١٠ أو ١١ وليس أي قيمة بينهما) ، وبتعبير آخر هي بيانات يمكن أن تعد ولا تُقاس ، مثل

عدد طلاب الفصول المختلفة في مدرسة ما

والبيانات المنفصلة إما أن تكون نوعية أو كمية متقطعة

أولا: البيانات النوعية 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 على البيانات النوعية:

مثال: في دراسة قام بإجرائها أحد الأطباء لطفل معرض لأحد الأمراض النفسية ،

تم سؤاله عن لون مجموعة من الأشياء فكانت إجاباته كما يلي :

أحمر أزرق بنفسجي أحمر أخضر

أبيض أبيض أحمر أزرق أبيض

أزرق أحمر أخضر أحمر بنفسجي

أخضر أزرق أبيض بنفسجي أحمر

المطلوب: عرض البيانات السابقة بطرق مختلفة

الحل تفصيلا في الكتاب

مثال على البيانات الكمية المتقطعة:

مثال: تم سؤال عدد من طلاب كليتي الآداب وإدارة الأعمال عن عدد حوادث

السيارات التي تعرضوا لها خلال العام الماضي فكانت اجاباتهم كما يلي:

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

المطلوب:

١. عرض البيانات السابقة في صورة جدول تكراري

٢. أحسب الاحتمالات التالية:

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

الحل تفصيلا في الكتاب

المحاضرة الرابعة

العرض الجدولي للبيانات - ٢

ثانياً: البيانات الكمية المتصلة:

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

الخطوة الأولى: تحديد عدد الفئات

الخطوة الثانية: تحديد طول الفئة

الخطوة الثالثة: تعيين حدود الفئات

الخطوة الرابعة: توزيع التكرارات على الفئات

مثال: البيانات التالية تعبر عن رأس المال المستثمر في شركات الحاسبات الآلية بالآلاف ريال:

٢٥	٢٦	٤١	٣٦	٤٤	٢٣	١٥	٧	١٢	٢
١٣	٢١	٣٣	٣٥	٤٥	٢٢	٢٦	١٢	٢٢	٣
٤٣	٤١	٣٠	٣٢	٤٨	١٨	٢٤	٢٣	٣٢	٥
٢٣	١٦	١	٩	٢٣	١١	٢٣	٣٢	٣٦	٦
١٨	١٧	٢٠	٢١	٢٦	٢٠	٣٩	٣٦	٣٥	٧

المطلوب:

عرض البيانات السابقة في صورة الجدول التكراري المناسب

الحل تفصيلاً في الكتاب

وهناك عدة ملاحظات يجب الإنتباه إليها عند عمل جدول التوزيع التكراري لبيانات المتغير الكمي المتصل:

١- إن تحديد عدد الفئات يتوقف على أمور عدة منها:

• عدد المفردات محل الدراسة

• انتظام وتوزيع تلك البيانات

• طبيعة بيانات المشكلة محل الدراسة

٢- 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 تركيز البيانات بتلك الفئة بقدر الإمكان حيث يعبر مركز الفئة عن قيمة كل مفردة من المفردات التي تنتمي لتلك الفئة

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

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

- الجداول التكرارية المنتظمة
- الجداول التكرارية غير المنتظمة
- الجداول التكرارية المفتوحة

أولاً: الجداول التكرارية المنتظمة:

وهي الجداول التي يكون فيها أطوال كل الفئات متساوية كما تم توضيحه في المثال السابق

ثانياً: الجداول التكرارية غير المنتظمة:

وفيها تكون أطوال الفئات غير متساوية، ومثال ذلك البيانات التالية والتي توضح توزيع عدد من العمال وفقاً للاجر الذي يحصل عليه كل منهم:

فئات الاجر	- ١٠	- ٢٠	- ٤٠	٥٠ - ٥٥	المجموع
عدد العمال (التكرار)	١٠	٤٠	١٥	٥	٧٠

ثالثاً: الجداول التكرارية المفتوحة:

وتوضحها أشكال الجداول التالية:

فئات العمر	عدد الطلاب
أقل من ٦	٢٠
-٦	٣٥
-١٢	٢٥
-١٥	١٨
١٨ فأكثر	٢٢

فئات العمر	عدد الطلاب
-٦	٢٠
-١٢	٣٥
-١٥	٢٥
١٨ فأكثر	١٨

فئات العمر	عدد الطلاب
أقل من ٦	٢٠
-٦	٣٥
-١٢	٢٥
١٨ - ١٥	١٨

جدول مفتوح من الطرفين

جدول مفتوح من أعلى

جدول مفتوح من أسفل

الجداول التكرارية المتجمعة:

وهي جداول يتم إعدادها لإعطاء نتيجة تراكمية لمجموعة من الفئات والتي يمكن أن تكون بشكل تصاعدي أو تنازلي ولكل منهما أهمية في 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 الاراضى دونم	عدد قطع الاراضى
1 -	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

الحل تفصيلا في الكتاب

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).

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 الاراضى
1 -	14
3 -	29
5 -	18
7 - 10	9
المجموع	70

المطلوب:

إعداد الجدول التكرارى المتجمع الهابط مع بيان نسبة قطع الأراضى التى تزيد أو تساوى 5 دونم

الحل تفصيلا في الكتاب

الجدول التكرارى المزدوج:

عند دراستنا لمتغيرين لتحديد العلاقة بينهما مثل العلاقة بين عدد أفراد الأسرة والمستوى التعليمى أو العلاقة بين أجزء العامل ودرجة الرضاء الوظيفى أو ماشابه ذلك، فى هذه الحالة لا بد من تبويب البيانات بالطريقة

التي تسمح باستنتاج أو تحديد العلاقة بين المتغيرين موضوع الدراسة ويتم ذلك من خلال الجدول التكراري المزدوج كما يتضح من المثال التالي:

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

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

المطلوب: إعداد جدول تكراري مزدوج

تمارين محلولة

س ١: المدى لمجموعة من البيانات المنفصلة هو :

أكبر قيمة في البيانات

أكثر القيم تكراراً في البيانات

الفرق بين أكبر وأصغر قيمتين في البيانات

أصغر قيمة في البيانات

س ٢: الجدول المرافق يبين درجات ٢٠ طالباً في أحد المقررات الدراسية :

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

(أ) عدد الطلاب الحاصلين على 94 فأقل هو :

3 0.15 4 7

(ب) عدد الطلاب الحاصلين على درجة أقل من 94 هو :

3 0.15 4 7

(ج) نسبة الطلاب الحاصلين على 94 فأقل هي :

7 4 35% 0.35

(د) النسبة المئوية للطلاب الحاصلين على 94 فأقل هي :

7 4 35% 0.35

هامش للإجابة

$$7 = 3 + 2 + 2 \quad (أ-٢)$$

$$4 = 2 + 2 \quad (ب-٢)$$

$$\frac{7}{20} = 0.35 \quad (ج-٢)$$

$$0.35 \times 100 = 35\% \quad (د-٢)$$

خذ بالك : المطلوب

نسبة (وليس نسبة مئوية)

أيوه ... ده بقى

نسبة مئوية

المحاضرة الخامسة

العرض البياني للبيانات

أولاً: البيانات غير المبوبة

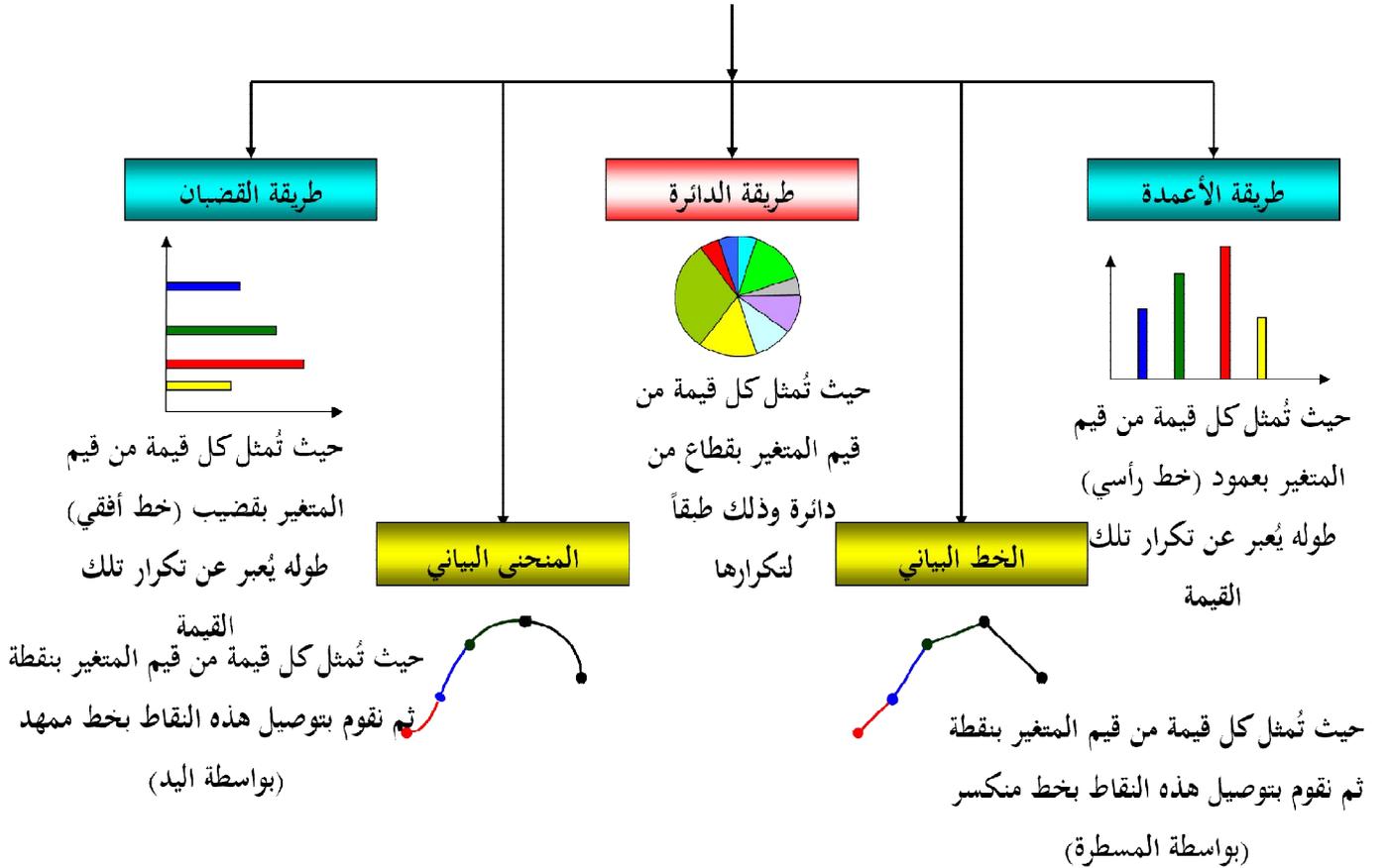
تعريف الرسوم البيانية:

هي وسيلة مفيدة وفعالة لتوضيح وشرح الحقائق الرقمية وإبراز العلاقة بين المتغيرات، واستقراء اتجاهاتها العامة بأسلوب يسهل فهمه وتذكره بمجرد النظر .

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

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

طرق العرض البياني للبيانات المنفصلة



١ - طريقة الأعمدة:

ويتم عرض البيانات من خلال هذا الأسلوب من خلال عدة أنواع من الأعمدة البيانية وهي:

أ- الأعمدة البيانية البسيطة :

وهي عبارة عن مجموعة من الأعمدة الرأسية أو المستطيلات المتساوية القاعدة والتي تتناسب ارتفاعاتها مع البيانات التي تمثلها

مثال: الجدول الآتي يوضح أعداد الطلاب المقيدون بأحد الجامعات في السنوات الدراسية من ١٤٢٣ هـ حتى ١٤٢٧ هـ.

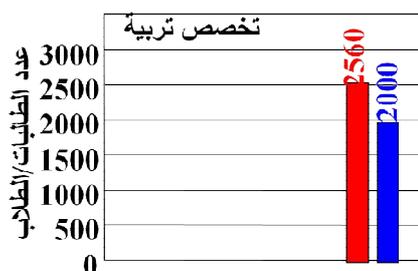
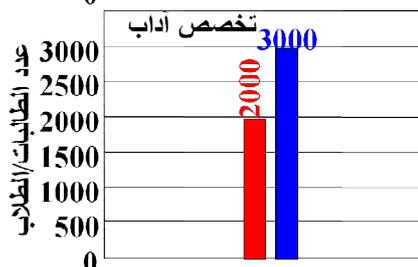
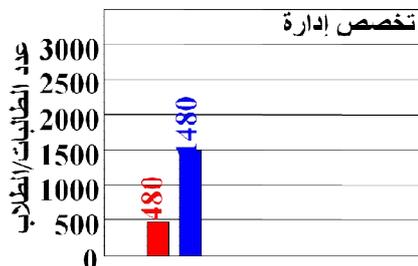
السنة الدراسية	١٤٢٣	١٤٢٤	١٤٢٥	١٤٢٦	١٤٢٧
عدد الطلاب بالالف	٣	٥	٦	٨	١٠

المطلوب:

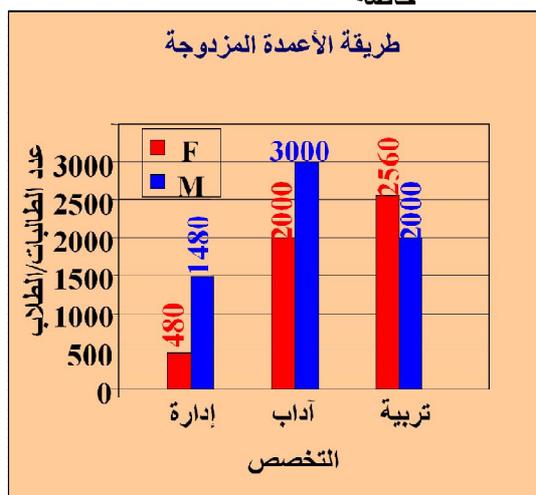
تمثيل البيانات باستخدام الرسم البياني المناسب

الحل تفصيلا في الكتاب

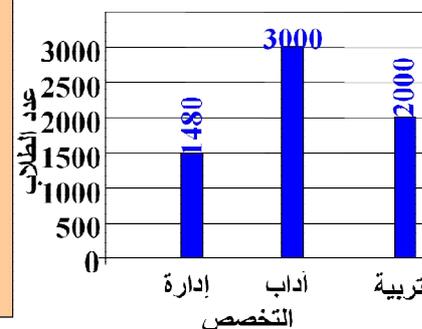
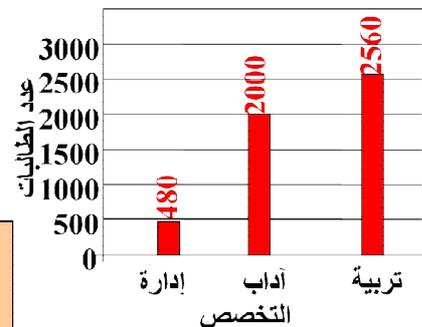
ب - الأعمدة البيانية المزدوجة: وهو ذلك النوع من الرسوم البيانية الذي يستخدم اذا كان الهدف من الرسم هو مقارنة ظاهرتين أو اكثر لعدة سنوات، أو اذا كان لدينا بيانات مزدوجة لخواص مختلفة .



	M	F	
	1480	480	إدارة
	3000	2000	أداب
	2000	2560	تربية خاصة



أي أن كل تخصص يُمثل بعمود مزدوج مكون من عمودين بسيطين متلاصقين



مثال: الجدول الآتي يوضح أعداد الطلبة المسجلين بإحد الجامعات السعودية في السنوات الدراسية ١٤١٩هـ حتى ١٤٢٣هـ

السنة الدراسية		١٤٢٣	١٤٢٢	١٤٢١	١٤٢٠	١٤١٩
عدد الطلبة بالآلاف	طلاب	٥٤	٤٩	٤٥	٤٠	٣١
	طالبات	٢٧	٢٠	١٦	١٢	٩

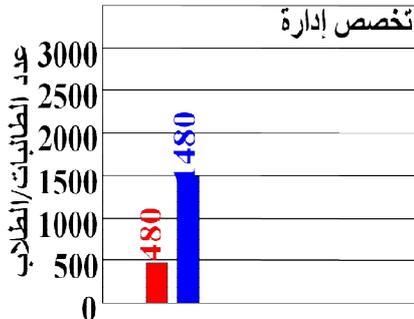
المطلوب:

مثل هذه البيانات بيانياً باستخدام الأعمدة البيانية المزدوجة ؟

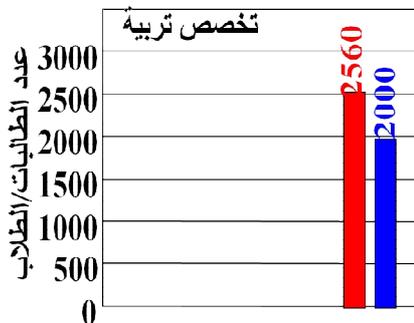
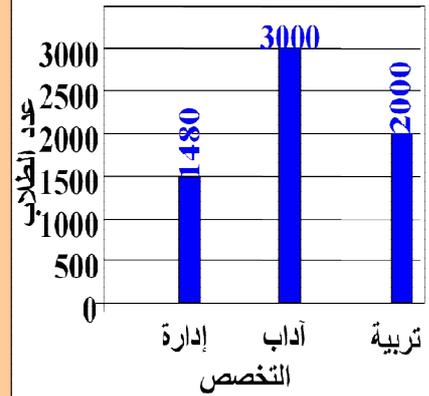
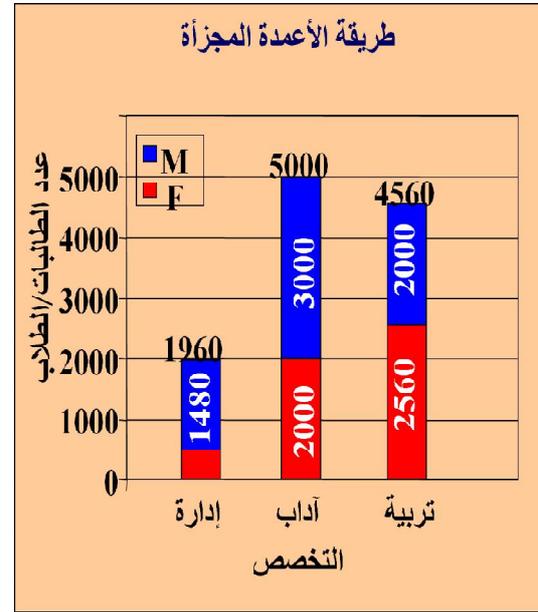
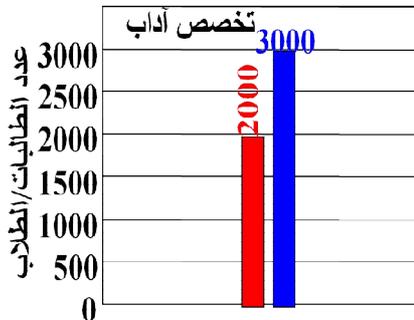
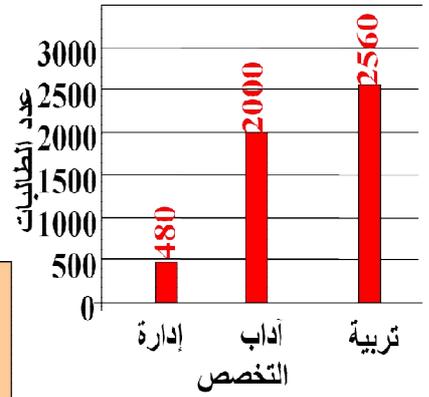
الحل تفصيلاً في الكتاب

ج - الأعمدة البيانية المزدوجة:

وهو ذلك النوع من الرسوم البيانية الذي يستخدم إذا كان الهدف من الرسم هو مقارنة ظاهرتين أو أكثر لعدة سنوات، أو إذا كان لدينا بيانات مزدوجة لخواص مختلفة .



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



أي أن كل تخصص يُمثل بعمود طوله يُعبر عن مجموع عدد طالباته وطلابها معاً ثم يتم تجزئته إلى عمودين كل منهما يمثل فئة من الفئات

مثال: إذا كانت اعداد الطلاب والطالبات المسجلين في كلية التربية بجامعة الملك فيصل بالاحساء تزداد كما هو موضح في الجدول الآتي:

السنوات الدراسية	١٤٢١	١٤٢٢	١٤٢٣	١٤٢٤	١٤٢٥	١٤٢٦
الطلاب	١٠٠	٢٠٠	٣٠٠	٤٠٠	٥٠٠	٦٠٠
الطالبات	١٥٠	٣٠٠	٤٥٠	٦٠٠	٧٥٠	٩٠٠

المطلوب:

مثل هذه البيانات بيانيا باستخدام الأعمدة المجزأة؟

الحل تفصيلا في الكتاب

ويمكن ابداء الملاحظات التالية على الرسومات بالاعمدة البيانية بأنواعها المختلفة:

١. تعتبر الاعمدة البيانية من اكثر الرسومات البيانية انتشارا،
٢. يفضل تظليل الاعمدة أو تخطيطها بواسطة خطوط متوازية أو ابرازها بألوان مختلفة وخاصة عند مقارنة ظواهر مختلفة.
٣. يستحسن اختيار مقياس رسم مناسب وثابت.
٤. يفضل عدم كتابة القيم التي تمثلها الاعمدة فوق الاعمدة وذلك لتلافي المبالغة في طول الاعمدة.
٥. يمكن استخدام العمود الواحد لتمثيل اكثر من نوع واحد من البيانات، وذلك باستخدام مفهوم الاعمدة المجزأة.
٦. تصلح الاعمدة البيانية لتمثيل البيانات ذات المتغيرات المنفصلة، كما تصلح بشكل خاص لتمثيل البيانات الوصفية (النوعية).

د - اللوحة الدائرية:

تستخدم الدائرة أو اللوحة الدائرية لتمثيل البيانات في الحالات التالية:
عندما يكون الهدف منها مقارنة الاجزاء المختلفة بالنسبة للمجموع الكلي
أن تكون الاجزاء المقارنة قليلة العدد نسبيا وفي فترة زمنية واحدة.

وفيما يلي خطوات رسم الدائرة وتقسيمها الى قطاعات:

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

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

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

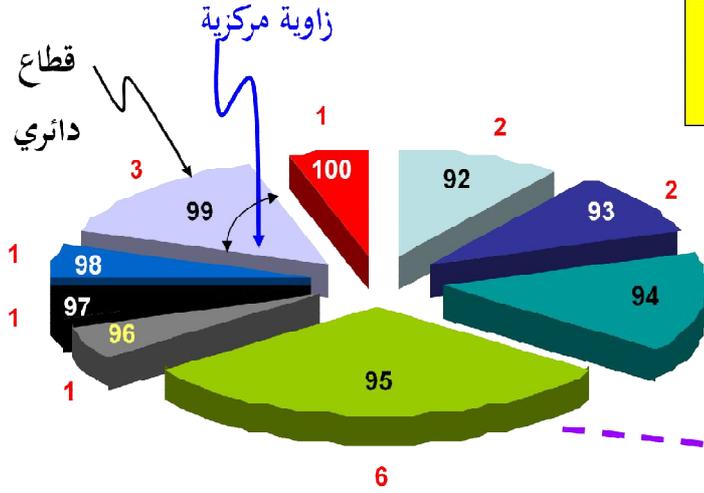
وفيما يلي تطبيق ذلك على بيانات إحصائية :

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

لو كانت لدينا البيانات التالية:

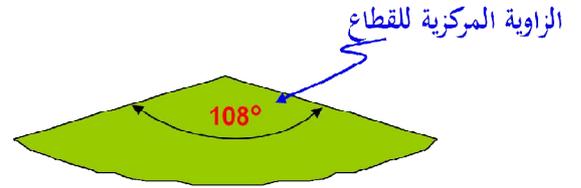
تحدد زاويته المركزية بالعلاقة :

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



أو

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



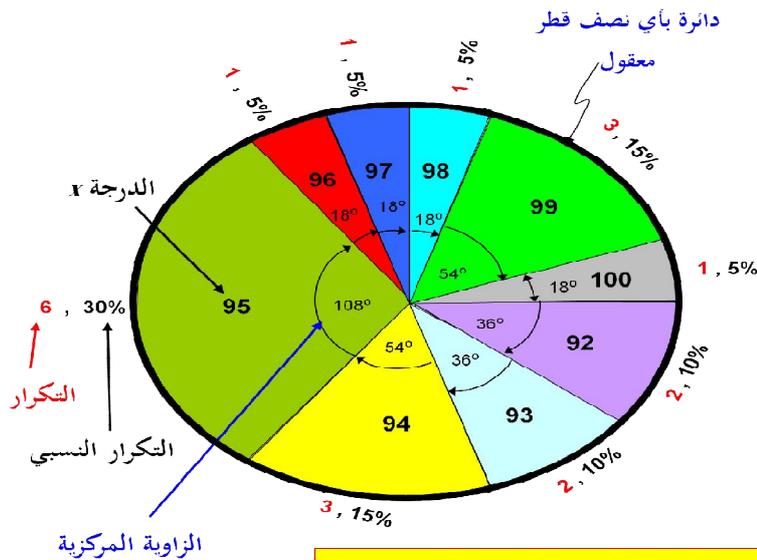
القيم داخل القطاعات تمثل الدرجة (المتغير) x والقيم المكتوبة خارج القطاعات باللون الأحمر تمثل التكرار f

القطاع الخاص بالدرجة "95" ذات التكرار 6 قياس زاويته المركزية تساوي :

$$\frac{6}{20} \times 360 = 108^\circ$$

إذن لابد من حساب الزاوية المركزية المناظرة لكل قيمة من قيم المتغير x (الدرجة) ، وهذه القيم مبينة

بالتالي :



الدرجة x	التكرار f	الزاوية المركزية
92	2	$(2/20) \times 360 = 36^\circ$
93	2	$(2/20) \times 360 = 36^\circ$
94	3	$(3/20) \times 360 = 54^\circ$
95	6	$(6/20) \times 360 = 108^\circ$
96	1	$(1/20) \times 360 = 18^\circ$
97	1	$(1/20) \times 360 = 18^\circ$
98	1	$(1/20) \times 360 = 18^\circ$
99	3	$(3/20) \times 360 = 54^\circ$
100	1	$(1/20) \times 360 = 18^\circ$
		مجموع الزوايا = 360°

طريقة الدائرة لتمثيل البيانات

مثال: فيما يلي احصائية لطلاب البكالوريوس في كلية العلوم الإدارية موزعين حسب السنة الدراسية للعام الجامعي ١٤٢٦ هـ.

السنة الدراسية	عدد الطلبة
السنة الأولى	٢٢٦
السنة الثانية	٢٧٦
السنة الثالثة	٢٦٦
السنة الرابعة	١٦٧
المجموع	٩٣٥

المطلوب:

عرض هذه البيانات باستخدام اللوحة الدائرية؟

الحل تفصيلا في الكتاب

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

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

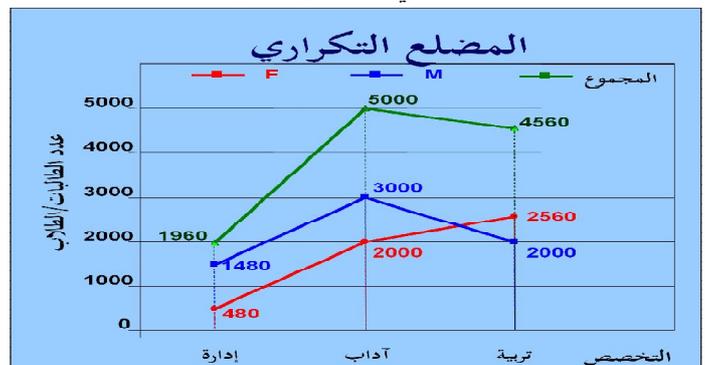
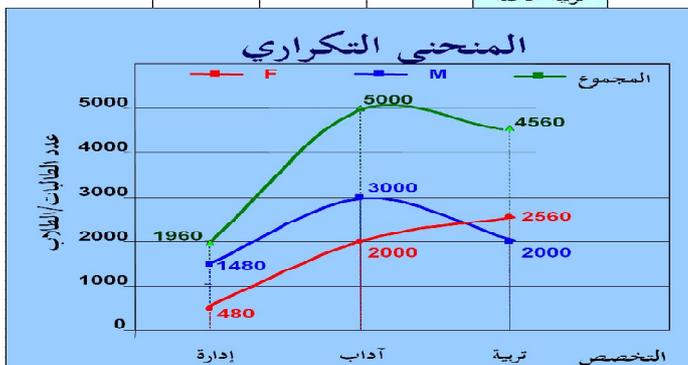
١. عندما تكون الكميات المقارنة كثيرة العدد نسبياً.
٢. عندما تكون الأجزاء المقارنة في فترات زمنية مختلفة.
٣. عندما نرغب في توضيح قيم الأجزاء المقارنة المختلفة للظاهرة موضع البحث وذلك من أجل إبراز المقارنة بين هذه الأجزاء أو توضيح التغير أو التطور عبر الزمن سواء لظاهرة واحدة أو عدة ظواهر بين فترات زمنية مختلفة.
٤. غالباً ما ينصح باستعمال الأعمدة البيانية (بأنواعها المختلفة) مع المتغيرات المنفصلة.

هـ - المنحنى أو الخط البياني:

يستخدم المنحنى أو الخط البياني أساساً لتوضيح الاتجاه العام للظاهرة خلال فترة من الزمن، ويستخدم هذا النوع من الرسم البياني لتمثيل الظواهر ذات البيانات المتصلة (غالباً)، وكذلك ممكن استخدامه مع البيانات المنفصلة. كما يمكن استخدام الخط أو المنحنى البياني لتمثيل أكثر من ظاهرة على نفس الرسم ومقارنتها ببعضها.

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

أيضاً نود التنويه أنه يمكن تمثيل جميع البيانات بطريقة الخط البياني أو المنحنى البياني كما هو مبين



مثال: البيانات التالية لدرجات عشر طلاب بكلية العلوم الإدارية في 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).

الحل تفصيلا في الكتاب

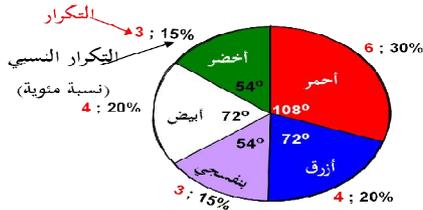
مثال : لو كانت لدينا البيانات التالية

X	F	الزاوية
أحمر	٦	١٠٨
أزرق	٤	٧٢
بنفسجي	٣	٥٤
أبيض	٤	٧٢
أخضر	٣	٥٤

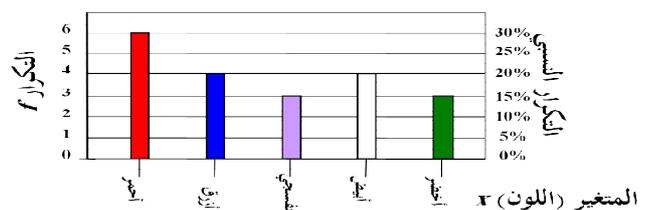
المطلوب : مثل البيانات السابقة بـ :

الأعمدة البسيطة الخط البياني المنحنى البياني اللوحة الدائرية

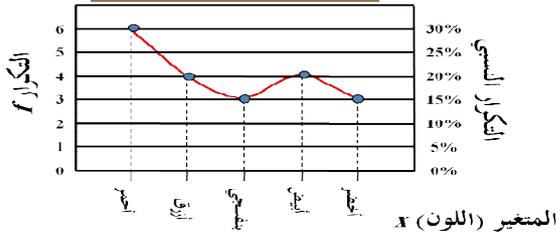
تمثيل البيانات بطريقة اللوحة الدائرية



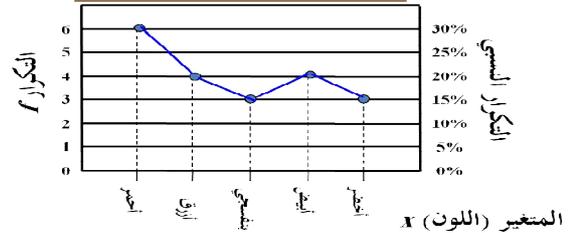
تمثيل البيانات بطريقة الأعمدة البسيطة



تمثيل البيانات المنحنى البياني



تمثيل البيانات بطريقة الخط البياني



ملاحظات على المنحنى والخط البياني :

١. الرسم بالخط البياني أو المنحنى يتطلب جهداً أقل من الجهد والوقت اللذين يتطلبهما رسم الأعمدة البيانية بأنواعها المختلفة.
٢. يسهل الخط البياني أو المنحنى المقارنة على القارئ.
٣. يمكن استخدام الخط البياني أو المنحنى (كما في الأعمدة البيانية) لتمثيل أكثر من ظاهرة على نفس الرسم ومقارنتها ببعضها.

مزايا وعيوب الرسوم البيانية :

المزايا:

١. تثير انتباه المشاهد خاصة إذا كانت جيدة التصميم.
٢. توفر وقت المشاهدة إذ أن استنباط الحقائق من الرسوم البيانية أسرع من الوصول إليها بواسطة الأرقام الموضوعية في جداول.
٣. إمكانية معرفة الاتجاهات العامة للظواهر.
٤. سهولة فهم وتذكر العلاقات بين الظواهر محل الدراسة.

العيوب:

١. التضحية بدقة البيانات إذ أن الرسوم توضح فقط التغيرات العامة للظواهر ولا تبين التفاصيل الدقيقة لها.
٢. أحيانا تكون الرسوم معقدة، خاصة إذا كانت تشتمل على مجموعات من البيانات المتباينة.
٣. كثرة التكاليف خاصة إذا كانت البيانات تحتاج إلى مقياس رسم كبير.

تمارين محلولة

المتغير (العمر) x	التكرار (العدد) f	الزاوية المركزية
20	20	72°
25	?	36°
30	30	?
35	?	?
Σf		

س ١ : الجدول المقابل يبين الجدول التكراري لأعمار عدد

من الممرضات (لأقرب سنة) اللاتي تعملن في أحد أقسام إحدى المستشفيات ، من هذا الجدول أجب

على الأسئلة التالية :

(أ) عدد الممرضات ذات العمر 25 سنة هو :

40 30 20 10

(ب) الزاوية المركزية المناظرة للعمر 30 سنة هي :

144° 108° 72° 36°

(ج) الزاوية المركزية المناظرة للعمر 35 سنة هي :

144° 108° 72° 36°

(د) عدد الممرضات الكلي [أي مجموع التكرارات] هو :

110 105 100 95

هامش للإجابة

(أ-١) هناك تناسب بين التكرار والزاوية المركزية ، إذن :

20	72°
?	36°

 $72 \times x = 36 \times 20$ ، $\therefore ? = 10$

(ب-١) بنفس الأسلوب السابق

20	72°
30	?

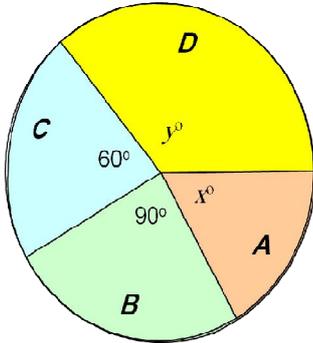
 $72 \times 30 = ? \times 20$ ، $\therefore ? = 108^\circ$

(ج-١) مجموع الزوايا المركزية يجب أن يكون 360°
 $\therefore 72 + 36 + 108 + ? = 360$ ، $\therefore ? = 144^\circ$

(د-١) هناك أكثر من طريقة أميزها الأسلوب المتبع في الجزئين (أ) ، (ب) :

20	72°
Σf	360°

 $360 \times 20 = 72 \times \Sigma f$ ، $\therefore \Sigma f = 100$



س ٢ : الشكل المقابل يبين مبيعات أربع شركات A, B, C, D (لبيع لعب

الأطفال) وذلك خلال عيد الفطر المبارك ، فإذا كان عدد اللعب الكلي

التي تم بيعها بواسطة هذه الشركات هو 5400 لعبة ، أجب على الأسئلة

(أ) النسبة المئوية لمبيعات الشركة B هي

60% 40% 30% 25%

(ب) عدد اللعب التي باعتها الشركة B هو :

1350 900 2250 2700

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

1350 3150 2250 900

هامش للإجابة

(أ-٢)

$360 \times ? = 90 \times 100$
 $? = 25\%$

100%	360°
?	90°

(ب-٢) $\frac{25}{100} \times 5400 = 1350 \rightarrow \Sigma f$

(ج-٢)

الزاوية المركزية المناظرة لمبيعات الشركتين معاً تساوي

$360 - (90 + 60) = 210^\circ$

5400	360°
?	210°

 $360 \times ? = 210 \times 5400$
 $? = 3150$

المحاضرة السادسة
العرض البياني للبيانات
ثانياً: البيانات المبوبة

يتم استخدام العديد من الأشكال للتعبير عن البيانات المبوبة في صورة جداول توزيعات تكرارية وهي :

١. المدرج التكراري
٢. المضلع التكراري
٣. المنحنى التكراري
٤. المنحنى التكراري المتجمع الصاعد
٥. المنحنى التكراري المتجمع الهابط (النازل)

المدرج التكراري

هو عبارة عن أعمدة مستطيلة متلاصقة يعبر ارتفاع العمود فيها على التكرار المناظر للفئة. ويستخدم هذا النوع من الرسوم البيانية لتمثيل البيانات التي تم عرضها في جدول توزيع تكراري، وفيه يمثل كل مستطيل فئة من فئات التوزيع التكراري.

يتم تقسيم المحور الرأسي (المحور الصادي) في المدرج التكراري حسب التكرار (فقد نستخدم التكرار الأصلي في حالة تمثيل التوزيع التكراري، وكذلك يمكن أن نستخدم التكرار النسبي في حالة تمثيل التوزيع التكراري النسبي).

يتم تقسيم المحور الرأسي (المحور الصادي) في المدرج التكراري حسب التكرار (فقد نستخدم التكرار الأصلي في حالة تمثيل التوزيع التكراري، وكذلك يمكن أن نستخدم التكرار النسبي في حالة تمثيل التوزيع التكراري النسبي).

ويتم تقسيم المحور الأفقي (المحور السيني) على أساس الفئات وهنا يظهر حالتين هما:

الحالة الأولى:- تساوى أطول الفئات
وفي هذه الحالة يكون ارتفاع المستطيل معبراً عن عدد مرات تكرار وجه الظاهرة محل الدراسة
الحالة الثانية:- عدم تساوى أطوال الفئات
وفي هذه الحالة لا بد من إجراء تعديل في التكرار الأصلي قبل رسم المدرج التكراري، لذا فإننا نقوم بإيجاد التكرار المعدل والذي هو عبارة عن ناتج قسمه التكرار الأصلي لكل فئة على طول الفئة المقابلة

مثال: البيانات التالية تعبر عن رأس المال المستثمر في شركات الحاسبات الآلية بالآلاف ريال

فئات رأس المال	٠-	١٠-	٢٠-	٣٠-	٤٠-٥٠	المجموع
عدد الشركات	٨	٩	١٦	١١	٦	٥٠

المطلوب : عرض البيانات في شكل المدرج التكراري

الحل تفصيلا في الكتاب

بعض خصائص التوزيع التكراري:

يمكن إستنتاج بعض خصائص التوزيع التكراري من شكل المدرج التكراري بدراسة الخصائص التالية:

الخاصية الأولى: التماثل

الخاصية الثانية: الإنتواء

الخاصية الثالثة: المنوال

المضلع التكراري

هو ذلك النوع من الرسوم البيانية الذي يمكن الحصول عليه من خلال حساب مراكز الفئات أو بتصنيف الأضلاع العلوية للمستطيلات في المدرج التكراري، ثم نوصل هذه النقاط بعضها مع بعض، كما يبدو لنا في المثال التالي:

مثال : استخدام البيانات في المثال السابق والتي تعبر عن رأس المال المستثمر في شركات الحاسبات الآلية بالألف ريال

فئات رأس المال	-٠	-١٠	-٢٠	-٣٠	٤٠ - ٥٠	المجموع
عدد الشركات	٨	٩	١٦	١١	٦	٥٠

المطلوب : عرض البيانات السابقة في شكل المضلع التكراري

الحل تفصيلا في الكتاب

المنحنى التكراري

ونحصل عليه إذا مهدنا المضلع التكراري وجعلناه منحنى بدلا من خطوط منكسرة فإننا نحصل على المنحنى التكراري.

مثال : البيانات التالية تعبر عن رأس المال المستثمر في شركات الحاسبات الآلية بالألف ريال

فئات رأس المال	-٠	-١٠	-٢٠	-٣٠	٤٠ - ٥٠	المجموع
عدد الشركات	٨	٩	١٦	١١	٦	٥٠

المطلوب : عرض البيانات السابقة في شكل المنحنى التكراري

الحل تفصيلا في الكتاب

التوزيعات التكرارية المتجمعة:

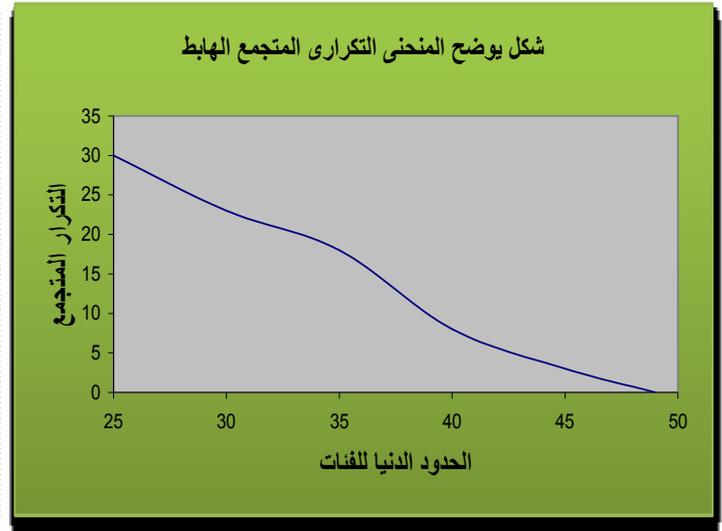
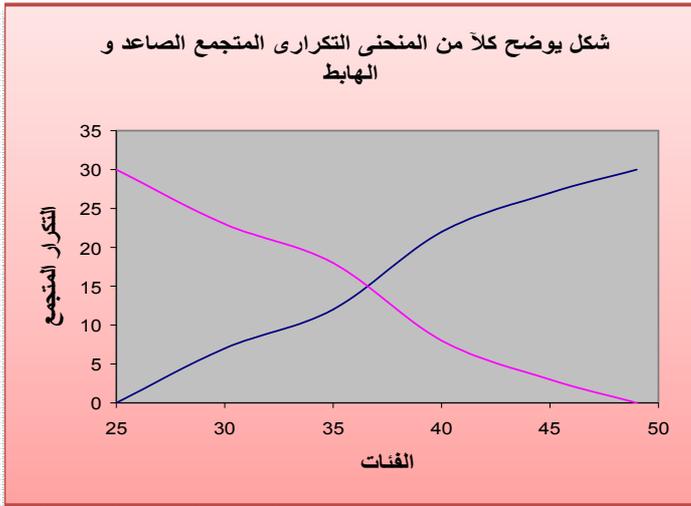
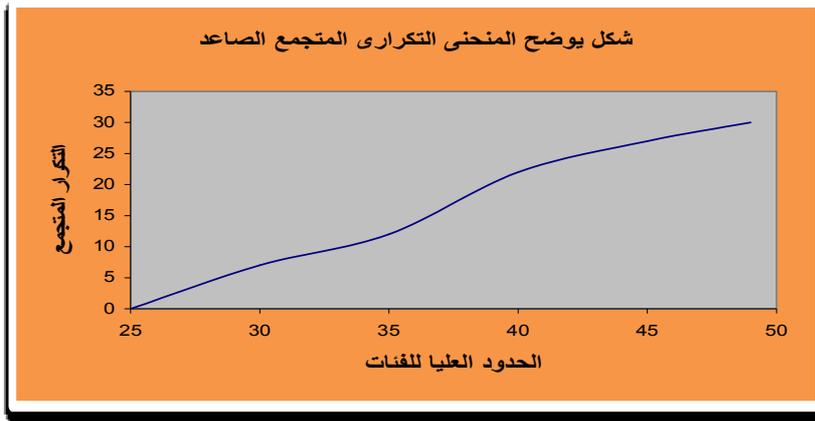
تستخدم المنحنيات المتجمعة لتمثيل التوزيعات التكرارية المتجمعة بيانياً بما يتلائم مع نوع التوزيع التكراري المتجمع، ونحصل على المنحنى المتجمع برصد التكرار المتجمع لأي فئة مقابل الحد الأعلى أو الحد الأدنى الفعلي لها ثم نوصل هذه النقاط فيما بينها بخطوط ممهدة.

يستخدم المنحنى المتجمع الصاعد

لتمثيل التوزيع التكراري المتجمع الصاعد، سواء أكان بالقيم المطلقة للتكرارات، أو بالتكرار النسبي، ويراعي وضع النقاط الخاصة بالتكرارات في حالة المنحنى المتجمع الصاعد عند الحد الأعلى لكل فئة، لأنه يعبر عن العدد الإجمالي لأوجه الظاهرة الواقعة أسفل الحد الأعلى للفئة.

ويستخدم المنحنى المتجمع الهابط (النازل)

لتمثيل التوزيع التكراري المتجمع الهابط (النازل) أيضاً بالقيم المطلقة للتكرارات أو بالتكرار النسبي، ويراعي وضع النقاط الخاصة بالتكرارات المتجمعة الهابطه (النازلة) عند الحد الأدنى لكل فئة، لأنه يعبر عن العدد الإجمالي لأوجه الظاهرة الواقعة أعلى الحد الأدنى للفئة.



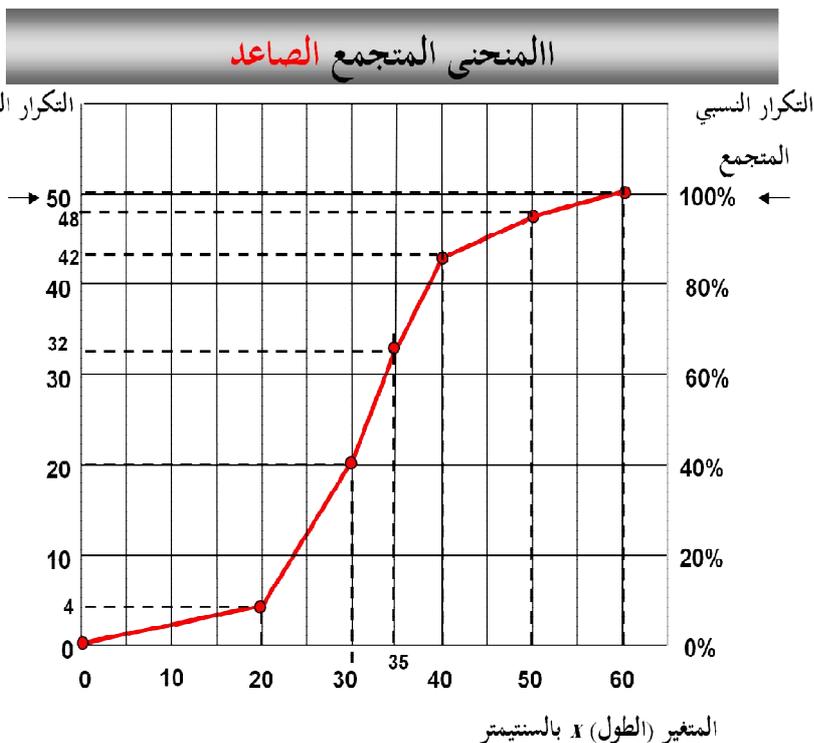
المنحنى المتجمع الصاعد

ذكرنا سابقاً عند عرضنا للبيانات عن طريق الجداول أنه يمكن عرض البيانات عن طريق التوزيع التكراري المتجمع الصاعد أو النازل، ويمكن الاستفادة من هذه الجداول في رسم المنحنى المتجمع الصاعد أو النازل كالاتي :

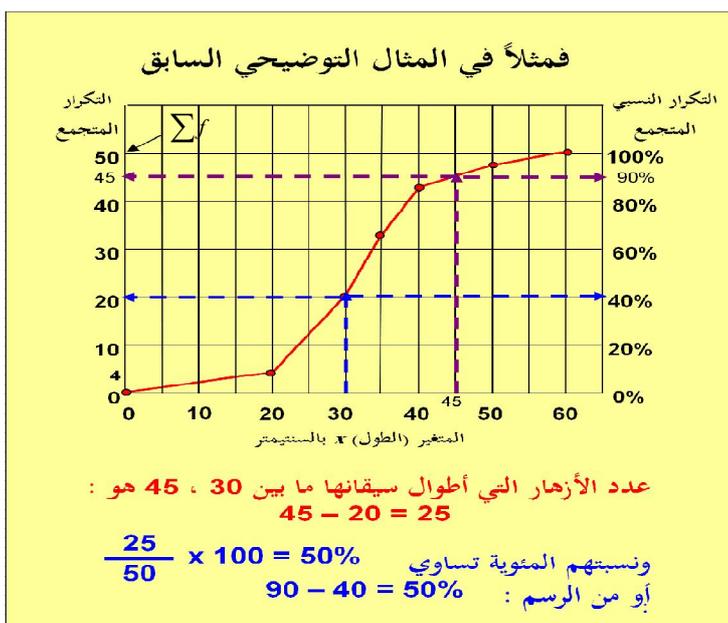
التوزيع التكراري الأصلي	
المتغير x	التكرار f
$0 \leq x < 20$	4
$20 \leq x < 30$	16
$30 \leq x < 35$	12
$35 \leq x < 40$	10
$40 \leq x < 50$	6
$50 \leq x < 60$	2



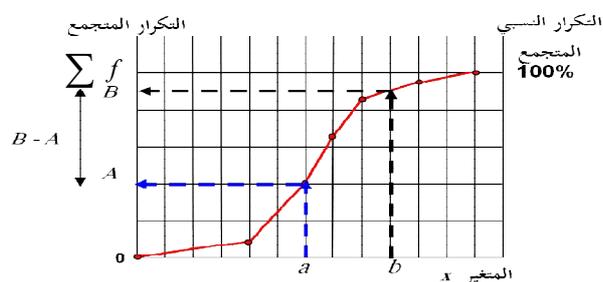
التوزيع التكراري المتجمع الصاعد			
المتغير x	التكرار المتجمع	التكرار النسبي	النقطة الموقفة على الرسم
< 0	0	0%	(0, 0)
< 20	4	8%	(20, 4)
< 30	20	40%	(30, 20)
< 35	32	64%	(35, 32)
< 40	42	84%	(40, 42)
< 50	48	96%	(50, 48)
< 60	50	100%	(60, 50)



تحديد التكرار المتجمع المناظر لـ



" x محصورة بين قيمتين"



فلحساب قيمة التكرار المتجمع المناظر لـ " $a \leq x < b$ " نحدد قيمتي a, b على المحور الأفقي [محور المتغير] ونحدد قيم التكرارات المتجمعة المناظرة [لتكن A, B على الترتيب] ، فيكون الحل المطلوب هو :

الفرق بين القيمتين A, B

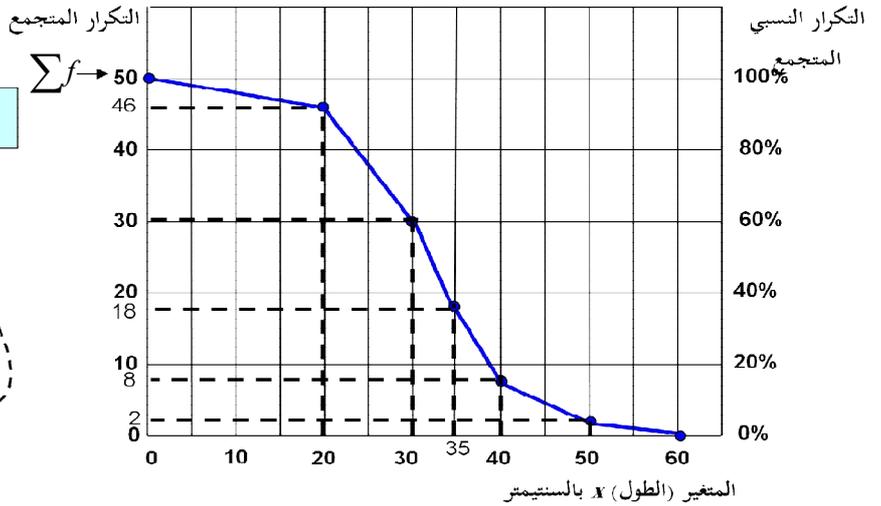
المنحنى المتجمع النازل أو الهابط

وبنفس طريقة المنحنى المتجمع الصاعد يمكن رسم المنحنى المتجمع النازل أو الهابط كالآتي

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

التكرار المتجمع
المناظر
الحد الأدنى
للفئة
(30, 30)

المنحنى المتجمع النازل أو الهابط



المتغير	التكرار المتجمع	التكرار النسبي	النقطة الواقعة على الرسم
≥ 0	50	100%	(0, 50)
≥ 20	46	92%	(20, 46)
≥ 30	30	60%	(30, 30)
≥ 35	18	36%	(35, 18)
≥ 40	8	16%	(40, 8)
≥ 50	2	4%	(50, 2)
≥ 60	0	0%	(60, 0)

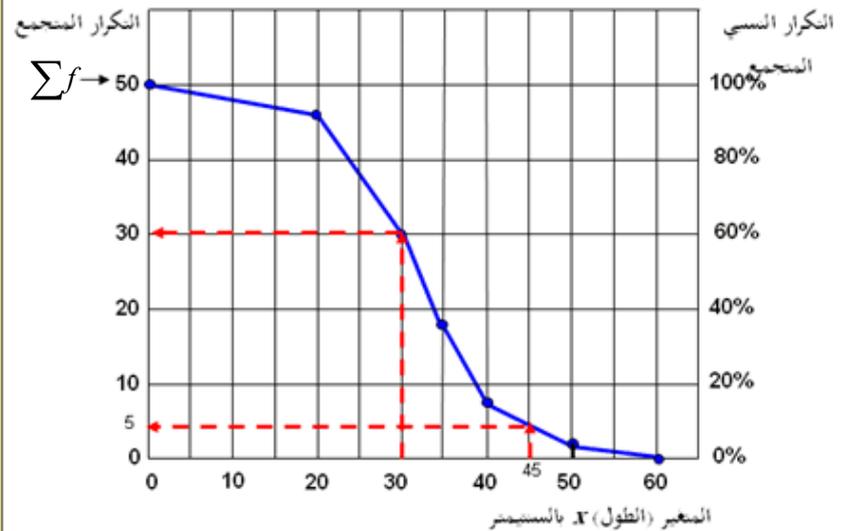
ويفيد المنحنى المتجمع النازل أو الهابط

في الرد على نفس الأسئلة التي يرد عليها المنحنى المتجمع الصاعد مع الأخذ في الاعتبار أن التدرج الرأسي [التكرار المتجمع] يمثل التكرار المناظر لـ " x أكبر من أو تساوي"

فمثلاً في المثال التوضيحي السابق

- عدد الأزهار التي أطوال سيقانها 30 فأكثر هو 30 بينما عدد الأزهار التي أطوال سيقانها أقل من 30 هو : $50 - 30 = 20$
- عدد الأزهار التي أطوال سيقانها 45 فأكثر هو 5 بينما عدد الأزهار التي أطوال سيقانها أقل من 45 هو : $50 - 5 = 45$
- عدد الأزهار التي أطوال سيقانها ما بين 30 , 45 هو : $30 - 5 = 25$

قارن النتائج السابقة بالنتائج التي سبق وحصلنا عليها باستخدام المضلع التكراري المتجمع المتصاعد



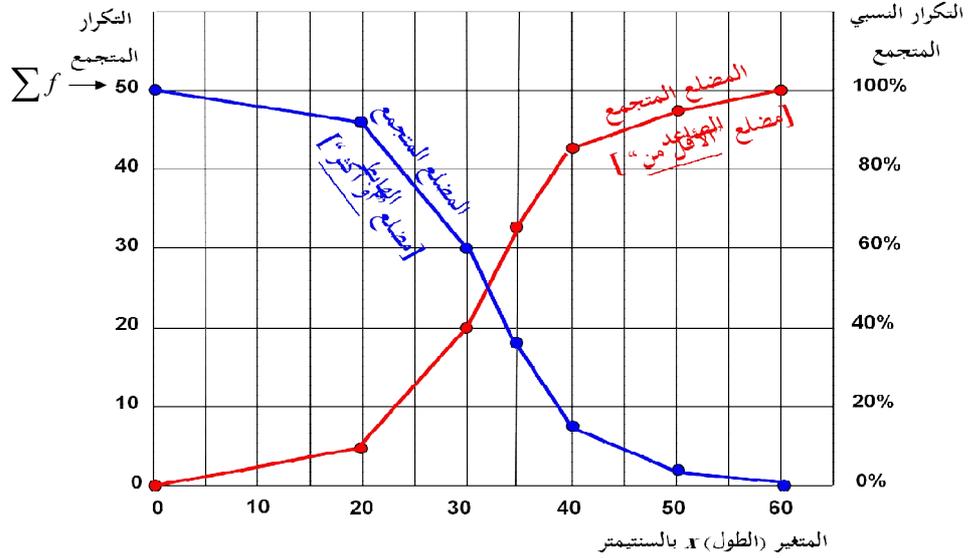
أي أن المنحنيان التكراريان المتجمعان الصاعد والهابط يؤديان نفس

الغرض تقريباً

ويمكن رسم المضلعين التكراريين المتجمعين : الصاعد والهابط على رسمة واحدة كما هو مبين :

التوزيع التكراري المتجمع الصاعد			
النقطة الموقعة على الرسم	التكرار النسبي	التكرار المتجمع	المتغير x
(0, 0)	0%	0	< 0
(20, 4)	8%	4	< 20
(30, 20)	40%	20	< 30
(35, 32)	64%	32	< 35
(40, 42)	84%	42	< 40
(50, 48)	96%	48	< 50
(60, 50)	100%	50	< 60

التوزيع التكراري المتجمع الهابط			
النقطة الموقعة على الرسم	التكرار النسبي	التكرار المتجمع	المتغير x
(0, 50)	100%	50	≥ 0
(20, 46)	92%	46	≥ 20
(30, 30)	60%	30	≥ 30
(35, 18)	36%	18	≥ 35
(40, 8)	16%	8	≥ 40
(50, 2)	4%	2	≥ 50
(60, 0)	0%	0	≥ 60



الأشكال الشائعة للتوزيعات التكرارية

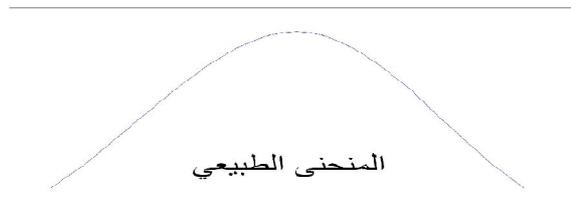
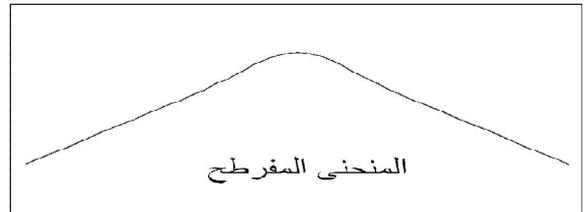
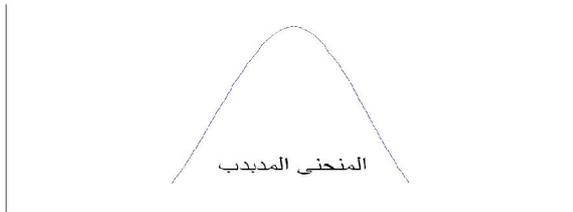
يعتبر التوزيع الطبيعي ذو شكل الجرس من التوزيعات التكرارية الهامة في دراستنا.

وفي أحيان أخرى يكون المنحنى التكراري

مدبب القمة بحيث تكون القمة ضيقة وذو طرفين واسعين نسبياً، فيسمى في هذه الحالة منحنى قليل التفرطح أو المنحنى المدبب.

وقد يكون المنحنى التكراري

مسطح القمة بحيث تكون القمة واسعة وذو طرفين ضيقين نسبياً، فيسمى منحنى كبير التفرطح أو المنحنى المفرطح، وفيما يلي رسم بياني يوضح كلا المنحنيين المدبب والمفرطح.



المحاضرة السابعة

المقاييس الإحصائية للبيانات غير المبوبة

أولاً: مقاييس النزعة المركزية

فبعد جمع البيانات و المعلومات و عرضها يأتي بعد ذلك

تحليل البيانات Data Analysis

والتى فيها يتم استخدام الأدوات الإحصائية المختلفة لوصف البيانات من خلال حساب المقاييس الإحصائية المختلفة و التى سوف نستعرضها في هذه المحاضرة بمشيئة الله.

تساعدنا المقاييس الإحصائية

في وصف المتغيرات المختلفة عن طريق معرفة القيم التى تتركز حولها البيانات ومدى التفاوت بين قيم المفردات محل الدراسة و تلك القيم.

كما تساعدنا في المقارنة بين المتغيرات المختلفة من حيث مدى نزعتها نحو مراكز معينة و تحديد مدى تجانس البيانات بعضها مع بعض.

وتمثل أهمية عملية وصف البيانات كميًا

من خلال محاولة الوصول إلى فهم و رؤية أوضح للمعلومة المحتواة في القيم الكمية للمتغيرات محل الدراسة،

تنقسم المقاييس الإحصائية إلى نوعين رئيسيين هما:

مقاييس النزعة المركزية Central Tendency Measures

مقاييس التشتت أو الأنتشار Dispersion Measures

في هذه المحاضرة سنتعرض لكيفية حساب مقاييس النزعة المركزية و مقاييس التشتت في حالة استخدام البيانات الخام غير المبوبة، أي تلك التى لم يتم تصنيفها في صورة جداول تكرارية

أولاً- مقاييس النزعة المركزية Central Tendency Measures

نقصد بمقاييس النزعة المركزية تلك القيم الوسطى التى توضح القيمة التى تجمع أكبر عدد من القيم الخاصة بمجموعة معينة عندها . و لتحديد القيمة المتوسطة للتوزيع يوجد هناك عدة مقاييس أهمها :

١. المتوسط الحسابي

٢. الوسيط

٣. المنوال (الشائع)

أهمية حساب مقاييس النزعة المركزية :

حساب مقاييس النزعة المركزية يساعد على التالي:

ايجاد ذلك الرقم المتوسط الذي يدل على خصائص أرقام مجموعة من المجموعات فيكفي أن ننظر الى ذلك الرقم المتوسط لنعرف الكثير عن خصائص هذه المجموعة من الأرقام

المقارنة بين عدة مجموعات في وقت واحد ، فنقول أن هذه المجموعة أقوى من تلك، وذلك اعتمادا على مقارنة هذه المتوسطات بعضها ببعض

الوسط الحسابي (المتوسط) Mean

يعرف المتوسط الحسابي بأنه قيمة التي اذا أعطيت لكل مفرد من مفردات الظاهرة لكان مجموع القيم الجديدة مساويا للمجموع الفعلي للقيم الأصلية للظاهرة، ويتم حساب المتوسط الحسابي للبيانات غير المبوبة من خلال المعادلة التالية:

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

مثال: البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام ١٤٢٧ هـ بالألف ريال كما يلي:

الشهر	محرم	صفر	ربيع أول	ربيع ثان	جمادى أول	جمادى الآخر	رجب	شعبان	رمضان	شوال	ذي القعدة	ذي الحجة
المبيعات	٣	٥	٨	٣	٦	٤	١٢	٥	٤	٣	٧	٩

المطلوب: حساب المتوسط الحسابي للمبيعات الشهرية.

الحل تفصيلا في الكتاب

ويجب ملاحظة عدة أمور في الوسط الحسابي وهي:

١. انه لا يشترط أن يكون المتوسط الحسابي عددا صحيحا.
٢. ان المتوسط الحسابي دائما محصور بين أقل القيم وأعلاها، ولكن هذا لا يعني أنه يقع في الوسط تماما بين هذين الحدين.
٣. إن المجموع الجبري لانحراف القيم عن المتوسط يكون دائما صفر.
٤. ومن أهم خصائص الوسط الحسابي هو تأثيره بجميع العمليات الجبرية تجرى على البيانات من إضافة قيمة لجميع البيانات أو طرحها أو ضربها أو قسمتها.

مثال: بسؤال خمسة أشخاص عن أجرهم الشهري فكانت إجاباتهم كما يلي بالألف ريال:

3 , 5 , 2 , 7,3

المطلوب:

- أحسب متوسط الأجر الشهري
- وإذا قررت إدارة الشركة زيادة أجورهم أحسب متوسط الأجر الجديد في الحالتين التاليتين
- ١. زيادة اجور العاملين بمقدار 2000 ريال
- ٢. زيادة أجور العاملين بنسبة 5 %

الحل تفصيلا في الكتاب

مزاي و عيوب المتوسط الحسابي:

المزايا:

١. يعد المتوسط الحسابي من اكثر مقاييس النزعة المركزية استخداما، واسهلها فهما وذلك نتيجة لسهولة حسابه

٢. يدخل في حسابه كل القيم دون اهمال أي قيمة منها.

العيوب:

١. يتأثر بالقيم المتطرفة الشاذة قلة أو كثرة، فقد يرتفع لمجرد وجود قيمة مرتفعة، وقد يقل كثيرا لمجرد وجود قيمة واحدة صغيرة وهذا بالتالي يؤدي الى عدم تمثيل المتوسط لواقع المعلومات.

٢. لا يمكن ايجاده من خلال الرسم.

الوسيط Median

يعرف الوسيط بأنه الدرجة التي تتوسط مجموعة من الدرجات المرتبة ترتيبا تصاعديا أو تنازليا، ويمكن حساب الوسيط باتباع الخطوات التالية:

١. ترتيب الدرجات تصاعديا أو تنازليا

٢. إيجاد ترتيب الوسيط ويقصد به إيجاد مكان الوسيط، ويختلف ترتيب الوسيط إذ كان عدد المشاهدات فردي أو زوجي كما يلي:

عدد المشاهدات n	ترتيب الوسيط
فردى	$(n+1)/2$
زوجى	يوجد ترتيبين هما $n/2$, $(n/2)+1$

إيجاد قيمة الوسيط.

مثال: البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام ١٤٢٧ هـ بلألف ريال كما يلي:

الشهر	محرم	صفر	ربيع أول	ربيع ثان	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذى القعدة	ذى الحجة
المبيعات	٣	٥	٨	٣	٦	٤	١٢	٥	٤	٣	٧	٩

المطلوب: إيجاد قيمة الوسيط للبيانات السابقة.

الحل تفصيلا في الكتاب

مزاياء و عيوب الوسيط:

المزاياء:

١. لا يتأثر بالقيم الشاذة.
٢. يمكن استخدام الوسيط في البيانات الناقصة.
٣. يمكن الحصول على الوسيط وحسابه من خلال الرسم.
٤. يمكن استخدام الوسيط في البيانات التي يعرف ترتيبها ولا تعرف قيمتها.

العيوب:

١. لا يعتمد على جميع القيم، حيث أنه لا يدخل في حسابه سوى قراءة واحدة أو قراءتين من البيانات كلها.

المنوال Mode

يعرف المنوال بأنه القيمة التي تعتبر أكثر القيم شيوعاً، في نفس المثال السابق للمبيعات الشهرية .

أحسب المنوال؟

نجد أن المبيعات الأكثر تكراراً هنا هي ٣ ألف ريال لذلك

فان المنوال هنا = ٣

وقد يكون في التوزيع منوالين أو أكثر وذلك كالمثال الآتي:

٦ ، ٥ ، ٥ ، ٤ ، ٤ ، ٤

فالمنوال هنا = ٤ ، ٥ أي أنه يوجد منوالين .

وقد لا يكون في التوزيع منوال وذلك كالمثال الآتي:

٢ ، ٥ ، ٧ ، ٩ ، ١١

مزاياء و عيوب المنوال:

المزاياء:

١. سهل الحساب سواء بالرسم أو بالحساب
٢. لا يتأثر كثيراً بالقيم الشاذة
٣. لا يتأثر كثيراً لو تغيرت قيم بعض مفردات البيانات

العيوب:

١. أقل مقاييس النزعة المركزية استعمالاً
٢. عديم الفائدة في البيانات القليلة العدد

المحاضرة الثامنة

المقاييس الإحصائية للبيانات غير المبوبة

ثانياً: مقاييس التشتت أو الانتشار

كما تميل القيم الى التركز فانها تميل أيضا إلى التشتت أو الانتشار، فبالتالي فان أي توزيع من القيم له صفة التركز، وصفة التشتت. فمقاييس التشتت هي تلك المقاييس التي تعبر عن مدى تباعد القيم أو تقاربها في المجموعات التي يشملها البحث

مثال

مجموعة (أ) : ٨ ، ٨ ، ٨ ، ٨ ، ٨

مجموعة (ب) : ٦ ، ٥ ، ٣ ، ٢ ، ١

نلاحظ أن المجموعة الأولى (أ) لا يوجد بها تشتت، فهذه المجموعة متجانسة.

في حين نلاحظ ان المجموعة الثانية (ب) يوجد بها تشتت

يمكن ان يقاس تشتت البيانات عن طريق مقاييس التشتت المختلفة، وأهم هذه المقاييس:

١. المدى

٢. المدى الربيعي

٣. الإنحراف عن المتوسط

٤. التباين

٥. الإنحراف المعياري

لماذا نستخدم مقاييس التشتت؟

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

مجموعة (أ) : (٥٥ ، ٥٠ ، ٤٥) المتوسط هنا = ٥٠

مجموعة (ب) : (٧٠ ، ٥٠ ، ٣٠) المتوسط هنا = ٥٠

فإذا لا نستطيع ان نقول هنا ان المجموعتين متساويتين لأننا إذا رجعنا الى المجموعتين وجدنا انهما مختلفتين في الدرجات رغم تساوي المتوسطين حيث أن المتوسط الحسابي في المجموعتين يساوي (٥٠) . لكن اذا استخدمنا

احد مقاييس التشتت مثل المدى والذي يحسب من خلال العلاقة التالية: المدى = أعلى درجة - أقل درجة وعلى ذلك فإن:

مدى مجموعة (أ) : ٥٥ - ٤٥ = ١٠

مدى مجموعة (ب) : ٧٠ - ٣٠ = ٤٠

نرى ان درجة التشتت في المجموعة (أ) أقل منها في المجموعة (ب)، أي ان المجموعة (أ) تكون أكثر تجانسا من المجموعة (ب)

المدى Range

المدى هو الفرق بين أعلى درجة وأقل درجة في التوزيع.

ويعتبر المدى الوسيلة المباشرة لمعرفة مدى تقارب القيم أو تباعدها في أي توزيع، وهو وسيلة سهلة، إلا أنها أقل الوسائل دقة وذلك لأن حسابه يتوقف على قيمتين فقط من قيم المجموعة، ولا يهتم مطلقاً بما بينهما من قيم أخرى. **فالمدى** لا يصلح إلا إذا أراد الباحث أن يأخذ فكرة سريعة عن مدى تشتت بيانات التوزيع موضع الدراسة، إلا أن استخدامه والاعتماد عليه قد يؤديان إلى نتائج خادعة، وخاصة إذا كان هناك انفصال بين الدرجات المتطرفة وباقي الدرجات موضع البحث.

مثال: البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام ١٤٢٧ هـ بلألف ريال كما يلي:

الشهر	محرم	صفر	ربيع أول	ربيع ثان	جمادى أول	جمادى الآخر	رجب	شعبان	رمضان	شوال	ذى القعدة	ذى الحجة
المبيعات	٣	٥	٨	٣	٦	٤	١٢	٥	٤	٣	٧	٩

المطلوب:

حساب المدى للمبيعات الشهرية.

الحل:

نلاحظ أن أكبر قيمة هي ١٢ وأقل قيمة للمبيعات الشهرية هي ٣ لذلك يكون المدى ٩

$$\text{Range} = 12 - 3 = 9$$

عيوب المدى: نجد أن من أهم عيوب المدى أنه يتم حسابة بناء على أكبر و أصغر قيمتين وبالتالي في حالة كونهما أو أحدهما متطرفتين أو قيم شاذة فإن المدى يعطى نتائج مضللة.

متوسط الانحرافات المطلقة Average Absolute Deviation

متوسط الانحرافات المطلقة AAD

هو ذلك المقياس الذي يقيس تباعد كافة القيم عن المتوسط الحسابي .

وعلى الرغم من أن حساب نصف المدى الربيعي يقضي على أثر القيم المتطرفة، والتي تؤثر على حساب المدى المطلق، إلا أنها جميعاً (**المدى، ونصف المدى الربيعي**) يتناولان التباعد بين قيمتين فقط (**أعلى قيمة وأدنى قيمة**) في المدى، (**وقيمة الربيع الأدنى وقيمة الربيع الأعلى**) في نصف المدى الربيعي، وذلك من بين القيم موضع الدراسة، أما بقية القيم تبقى مهملة .

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

ويمكن حساب متوسط الانحرافات المطلقة من خلال المعادلة التالية :

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

مثال: البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام ١٤٢٧ هـ بلآلف ريال كما يلي:

الشهر	محرم	صفر	ربيع أول	ربيع ثان	جمادى أول	جمادى الآخر	رجب	شعبان	رمضان	شوال	ذى القعدة	ذى الحجة
المبيعات	٣	٥	٨	٣	٦	٤	١٢	٥	٤	٣	٧	٩

المطلوب:

أحسب متوسط الانحرافات المطلقة للمبيعات الشهرية.

الحل تفصيلا في الكتاب

التباين والانحراف المعياري:

التباين Variance: هو متوسط مربعات انحرافات القيم عن وسطها الحسابي و يرمز لها بالرمز σ^2 (تقراء سيجمما تربيع) وذلك اذا كان محسوب لبيانات المجتمع اما في حالة حسابة لبيانات عينة من المجتمع فيرمز له بالرمز S^2

الانحراف المعياري Standard Deviation: وهو الجذر التربيعي لمتوسط مربعات انحرافات القيم عن وسطها الحسابي اي هو جذر التباين لذلك يرمز له بالرمز σ (تقراء سيجمما) وذلك اذا كان محسوب لبيانات المجتمع اما في حالة حسابة لبيانات S عينة من المجتمع فيرمز له بالرمز S ويعتبر **الانحراف المعياري والتباين** من أهم مقاييس التشتت جميعا أو أكثرها استعمالا، وهما قريبين في خطوات ايجادهما من الانحراف عن المتوسط.

فالتباين والانحراف المعياري يختلف عن الانحراف عن المتوسط في طريقة التخلص من اشارات الفروق بين القيم والمتوسط الحسابي، فبينما نتخلص من هذه الاشارات في طريقة الانحراف عن المتوسط باهمال الاشارات كلية، نحتال على ذلك في طريقة التباين والانحراف المعياري بتربيع هذه الفروق (**أي نضربها في نفسها**) فتصبح بالتالي جميع الاشارات موجبة.

حساب التباين والانحراف المعياري:

يمكن حساب التباين من خلال المعادلة التالية:

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

وبالتالي يكون حساب الانحراف المعياري كما يلي:

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

مثال: البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام ١٤٢٧ هـ بلآلف ريال كما يلي:

الشهر	محرم	صفر	ربيع أول	ربيع ثان	جمادى أول	جمادى الآخر	رجب	شعبان	رمضان	شوال	ذى القعدة	ذى الحجة
المبيعات	٣	٥	٨	٣	٦	٤	١٢	٥	٤	٣	٧	٩

المطلوب:

أحسب قيمة التباين وقيمة الانحراف المعياري للمبيعات الشهرية.

الحل تفصيلا في الكتاب

ملاحظة هامة:

يعتبر من أهم خصائص الانحراف المعياري هو عدم تأثره بعمليات **الجمع والطرح** وإنما يتأثر فقط بعمليات **الضرب والقسمة**.

فلاحظ عدم تغير قيمة الانحراف المعياري في حالة **الجمع أو الطرح** وإنما تظل قيمة كما هي بالرغم من طرح مقدار ثابت من جميع قيم التوزيع.

أما في حالة **الضرب أو القسمة** فلاحظ تغير قيمة الانحراف المعياري وهي نفس قيمة الانحراف المعياري القديمة مضروبة في القيمة التي ضرب فيها أو قسم عليها.

مثال: البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام ١٤٢٧ هـ بلألف ريال كما يلي:

الشهر	محرم	صفر	ربيع أول	ربيع ثان	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذى القعدة	ذى الحجة
المبيعات	٣	٥	٨	٣	٦	٤	١٢	٥	٤	٣	٧	٩

المطلوب:

فإذا تم طرح ٢ من جميع بيانات المبيعات الشهرية أى تم تخفيض المبيعات الشهرية بمقدار ٢ أحسب قيمة الانحراف المعياري الجديد؟

الحل تفصيلا في الكتاب

نلاحظ عدم تغير قيمة الانحراف المعياري وإنما ظلت قيمة كما هي بالرغم من طرح مقدار ثابت ٢ من جميع قيم المبيعات الشهرية.

مثال: البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام ١٤٢٧ هـ بلألف ريال كما يلي:

الشهر	محرم	صفر	ربيع أول	ربيع ثان	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذى القعدة	ذى الحجة
المبيعات	٣	٥	٨	٣	٦	٤	١٢	٥	٤	٣	٧	٩

المطلوب:

أحسب قيمة الانحراف المعياري للمبيعات الشهرية إذا تم زيادة المبيعات الشهرية إلى ثلاث أمثال الموجود حالياً؟

الحل تفصيلا في الكتاب

نلاحظ تغير قيمة الانحراف المعياري وهي نفس قيمة الانحراف المعياري القديمة مضروبة في ٣ وبالتالي يمكن أن نكون حصلنا على كافة المقاييس الإحصائية الوصفية التي تصف المبيعات الشهرية فكانت كما يلي:

المتوسط	الوسيط	المتوال	الوسط الهندسي
5.75	5	3	5.20114

المدى	متوسط الانحرافات المطلقة	التباين	الانحراف المعياري
9	2.20833	7.840909	2.80016

المحاضرة التاسعة

المقاييس الإحصائية للبيانات المبوبة

أولاً: الوسط الحسابي والتشتت حوله

يقصد بالبيانات المبوبة تلك البيانات التي تم وضعها في صورة جداول تكرارية. والجداول التكرارية للمتغير الكمي المتقطع يمكن تحويلها لتكون بيانات غير مبوبة و نتعامل معها كما سبق توضيح ذلك في المحاضرة السابقة، الا أن الأمر يختلف بالنسبة للمتغير الكمي المتصل حيث يصعب ذلك ولا بد من التعامل معها كما هي على صورتها الجدولية وهذا ما سوف نتناوله في هذه المحاضرة إن شاء الله

وسيتم عرض كيفية حساب كلا من مقاييس النزعة المركزية ومقاييس التشتت في ثلاث حالات للجداول التكرارية وهي

١. الجداول المنتظمة

٢. الجداول غير المنتظمة

٣. الجداول المفتوحة

الجداول المنتظمة:

وهي تلك الجداول التي تكون فيها أطوال الفئات جميعها متساوية .

أولاً- الوسط الحسابي والتشتت حولة:

الوسط الحسابي كما سبق أن تم تعريفه في الفصل السابق هو القيمة التي إذا أخذها جميع المفردات لكان مجموعها يساوي مجموع القيم الأصلية، ويمكن حساب الوسط الحسابي او المتوسط الحسابي كما يلي:

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

\bar{x} الوسط الحسابي

x_i مركز الفئة i وهي تساوي (الحد الأعلى للفئة + الحد الأدنى للفئة) ÷ ٢

f_i تكرار الفئة i

l عدد الفئات

ويتم حساب التشتت حول المتوسط الحسابي من خلال الآتي:

أ- متوسط الانحرافات المطلقة **AAD**:

وهو يقيس إنحراف القيم عن وسطها الحسابي بغض النظر عن اشارة ذلك الانحراف حيث يتم حسابه 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 مربعات انحرافات القيم عن وسطها الحسابي. 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: هو الجذر التربيعي 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:

٦٠ - ٥٠	-٤٠	-٣٠	-٢٠	فئات العمر
٢٠	٥٠	٣٠	١٠	عدد العمال

المطلوب: حساب التالي:

- الوسط الحسابي
- التباين
- الانحراف المعياري
- متوسط الانحرافات المطلقة

الحل تفصيلا في الكتاب

الوسط الحسابي:

يمكن إيجاد الوسط الحسابي بتطبيق المعادلة الخاصة بذلك كما يلي:

$$\bar{x} = \frac{\sum_{i=1}^l x_i f_i}{\sum_{i=1}^l f_i} = \frac{4650}{110} = 42.2727$$

التباين:

يمكن الحصول على التباين باستخدام المعادلة الخاصة بذلك كما يلي:

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

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

الانحراف المعياري:

يمكن حسابة باستخدام المعادلة الخاصة بذلك كما يلي :

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

ويمكن الحصول على متوسط الانحرافات المطلقة **AAD** بتطبيق المعادلة الخاصة بذلك كما يلي:

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

المحاضرة العاشرة

المقاييس الإحصائية للبيانات المبوبة

ثانيا: الوسيط والتشتت حوله

الوسيط هو القيمة التي يصغرها عدد من القيم يتساوى مع العدد الذى يكبر هذه القيمة ولحساب الوسيط من البيانات المبوبة هناك ثلاث خطوات يجب إتباعها وهى:

١. إيجاد الجدول التكرارى المتجمع الصاعد

٢. إيجاد ترتيب الوسيط من خلال المعادلة التالية:

$$k_{Med} = n / 2$$

3. إيجاد قيمة الوسيط من خلال المعادلة التالية:

$$Med = L_{Med} + \frac{k_{Med} - F_a}{F_b - F_a} \times I$$

حيث أن :

Med قيمة الوسيط

L_{Med} الحد الادنى لبداية الفئة الوسيطة

k_{Med} ترتيب الوسيط

F_a التكرار المتجمع السابق للفئة الوسيطة

F_b التكرار المتجمع اللاحق للفئة الوسيطة

I طول الفئة الوسيطة

مثال: فى بيانات المثال السابق توزيع مجموعة من المدرسين العاملين فى مجال التربية وفقا لفئات أعمارهم،

٦٠ - ٥٠	-٤٠	- ٣٠	- ٢٠	فئات العمر
٢٠	٥٠	٣٠	١٠	عدد العمال

المطلوب: حساب قيمة الوسيط؟

الحل تفصيلا في الكتاب

الرُّبِيعِ الادنى (الأول):

يُعبّر الرُّبِيعِ الأول Q1 عن القيمة التي يكون قبلها عدد المشاهدات ربع العدد الكلي للمشاهدات والمشاهدات بعده تمثل ثلاث ارباع العدد الكلي للمشاهدات محل الدراسة. لذلك يتم حسابه كما في حالة الوسيط مع اختلاف أن ترتيب الرُّبِيعِ الاول Q1 هو (n / 4)

$$Q_1 = L_{Q_1} + \frac{\frac{n}{4} - F_a}{F_b - F_a} \times I_{Q_1}$$

الرُّبِيعِ الاعلى (الثالث):

يُعبّر الرُّبِيعِ الثالث Q3 عن القيمة التي يكون قبلها عدد المشاهدات ثلاث ارباع العدد الكلي للمشاهدات والمشاهدات بعده تمثل ربع العدد الكلي للمشاهدات محل الدراسة. لذلك يتم حسابه كما في حالة الوسيط مع اختلاف أن ترتيب الرُّبِيعِ الثالث Q3 هو (3 n / 4)

$$Q_3 = L_{Q_3} + \frac{\frac{3(n)}{4} - F_a}{F_b - F_a} \times I_{Q_3}$$

ويمكن إيجاد كلا من الرُّبِيعِ الادنى (الاول) Q1 و الرُّبِيعِ الاعلى (الثالث) Q3 بنفس خطوات حساب الوسيط الا أن الامر المختلف هنا هو الترتيب حيث يكون كالتالي:

Q3	Q1	
$k_{Q_3} = 3n/4$	$k_{Q_1} = n/4$	الترتيب

مثال: في بيانات المثال السابق توزيع مجموعة من المدرسين العاملين في مجال التربية وفقا لفئات أعمارهم،

فئات العمر	- ٢٠	- ٣٠	- ٤٠	٥٠ - ٦٠
عدد العمال	١٠	٣٠	٥٠	٢٠

المطلوب: حساب كل من:

- قيمة الرُّبِيعِ الاول
- قيمة الرُّبِيعِ الثالث

الحل تفصيلا في الكتاب

حساب قيمة العُشير $P_{0.10}$:

وبنفس الطريقة السابقة يمكن الحصول على العشير $P_{0.10}$ وهو القيمة التي يكون قبلها ١٠ % من مفردات المجتمع و ٩٠ % منها أكبر منه. والاختلاف يكون فقط في الترتيب حيث أن ترتيب العشير هو:

$$k_{P_{0.10}} = n/10$$

مثال: في بيانات المثال السابق توزيع مجموعة من المدرسين العاملين في مجال التربية وفقا لفئات أعمارهم،

فئات العمر	- ٢٠	- ٣٠	- ٤٠	٥٠ - ٦٠
عدد العمال	١٠	٣٠	٥٠	٢٠

المطلوب: حساب قيمة العُشير؟

الحل تفصيلا في الكتاب

حساب قيمة المنين $P_{0.01}$:

وبنفس الطريقة السابقة يمكن الحصول على المنوي $P_{0.01}$ وهو القيمة التي يكون قبلها 1 % من مفردات المجتمع و 99 % منها أكبر منه، والاختلاف بينه وبين ما سبق حسابه من الوسيط والرابع الأول أو الربع الثالث أو العُشير يكون فقط في الترتيب حيث أن ترتيب المنويين هو :

$$k_{P_{0.01}} = n/100$$

مثال: في بيانات المثال السابق توزيع مجموعة من المدرسين العاملين في مجال التربية وفقا لفئات أعمارهم،

فئات العمر	- ٢٠	- ٣٠	- ٤٠	٥٠ - ٦٠
عدد العمال	١٠	٣٠	٥٠	٢٠

المطلوب: حساب قيمة المنين؟

الحل تفصيلا في الكتاب

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

المقياس	$P_{0.01}$	$P_{0.10}$	Q_1	Med	Q_3
القيمة	٢١,١	٣٠,٣٣٣	٣٥,٨٣٣٣	٤٣	٤٨,٥

نصف المدى الربيعي Inter Quartile Range

بسبب العيب الموجود في مقياس التشتت (المدى) وتأثرة بالقيم الشاذة أدى ذلك للجوء إلى مقياس آخر يسمى (نصف المدى الربيعي IQR) والذي يستبعد القيم المتطرفة من الطرفين، حيث يعتمد في حسابه على كلا من الربع الأول Q_1 والربع الثالث Q_3 ويتم حسابه من خلال المعادلة التالية:

$$IQR = \frac{Q_3 - Q_1}{2}$$

مثال: في بيانات المثال السابق توزيع مجموعة من المدرسين العاملين في مجال التربية وفقا لفئات أعمارهم،

فئات العمر	- ٢٠	- ٣٠	- ٤٠	٥٠ - ٦٠
عدد العمال	١٠	٣٠	٥٠	٢٠

المطلوب: حساب قيمة نصف المدى الربيعي؟

الحل تفصيلا في الكتاب

ثالثا: المنوال

المنوال هو القيمة الأكثر شيوعا أو تكرارا. وفي حالة البيانات المبوبة يمكن حسابه باستخدام المعادلة التالية:

$$Mod = L_{Mod} + \frac{D_1}{D_1 + D_2} \times I$$

قيمة المنوال Mod

الحد الأدنى لفئة المنوال L_{Mod}

يساوى تكرار فئة المنوال - تكرار الفئة السابقة D_1

يساوى تكرار فئة المنوال - تكرار الفئة اللاحقة D_2

طول الفئة المنوالية I

مثال: فى بيانات المثال السابق توزيع مجموعة من المدرسين العاملين فى مجال التربية وفقا لفئات أعمارهم،

فئات العمر	٢٠ -	٣٠ -	٤٠ -	٥٠ - ٦٠
عدد العمال	١٠	٣٠	٥٠	٢٠

المطلوب: حساب قيمة المنوال؟

الحل تفصيلا فى الكتاب

الجداول غير المنتظمة:

وهي الجدوال التي يكون فيها أطوال الفئات غير متساوية ويكفى وجود فئة واحدة فقط طولها غير متساوى مع باقى الفئات لجعل الجدول غير منتظم. ويتم حساب المقاييس الإحصائية التي سبق عرضها فى حالة الجداول المنتظمة بنفس الطريقة فيما عدا المنوال،

ويتعين علينا عند حساب المنوال تعديل التكرارات قبل حسابة وكذلك قبل رسم المدرج التكرارى وذلك لأن حجم التكرارات فى تلك الحالة قد يسبب اتساع أو ضيق فى أعمدة فئات التوزيع ولذلك يتم التخلص من تأثير طول الفئة بإيجاد التكرار المعدل، ويتم ذلك من خلال المعادلة التالية:

$$\text{التكرار المعدل} = \text{التكرار الأصلي} \div \text{طول الفئة}$$

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

فئات الدخل	٣ -	٥ -	٨ -	١٠ - ١٥
عدد الموظفين	٢٠	٥٠	١٥	١٥

المطلوب حساب:

- ١- الوسط الحسابى
- ٢- متوسط الانحرافات المطلقة
- ٣- التباين
- ٤- الانحراف المعيارى
- ٥- الوسيط
- ٦- الربيع الأول
- ٧- الربيع الثالث
- ٨- العشير
- ٩- المتوسين
- ١٠- نصف المدى الربيعى
- ١١- المنوال

الحل تفصيلا فى الكتاب

يمكن حساب المطلوب من ١ إلى ١٠ بنفس طريقة حسابها فى حالة الجداول المنتظمة بدون أي تعديل. اما المطلوب رقم ١١ فيطلب حساب المنوال، وهو الذي طريقتة تحتاج إلى تعديل فى الحساب فى حالة الجداول غير المنتظمة.

الجداول المفتوحة:

وهي ذلك النوع من الجداول التي يكون فيها الحد الادنى للفئة الاولى غير محدد أو الحد الأعلى للفئة الأخير غير محدد أو كلاهما. وفى هذا النوع من الجداول يصعب حساب الوسط الحسابى والتباين والانحراف المعيارى، حيث لا يمكن تحديد مركز الفئة للفئات المفتوحة، لذا فيعبر من أنسب المقاييس الإحصائية فى تلك الحالة هي المقاييس الوسيطة والتي يقصد بها الوسيط والربيع الادنى والربيع الاعلى والعشير والمتوسين وكذلك لقياس التشتت يتم من خلال نصف المدى الربيعى.

مثال: البيانات تعبر عن أوزان مجموعة من الطلاب بالكيلوجرام في المرحلة الجامعية فكانت كما يلي:

فئات الوزن	أقل من ٥٠	- ٥٠	- ٦٠	- ٧٠	٨٠ فأكثر
عدد الطلاب	٥	١٠	٣٥	١٥	١٠

المطلوب:

حساب مقاييس النزعة المركزية والتشتت المناسبة ؟

الحل تفصيلا في الكتاب

المحاضرة الحادية عشر

مقاييس التشتت النسبي والدرجة المعيارية

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

حيث سيتم في هذه المحاضرة استعراض كلا من:

- مقاييس التشتت النسبي
- القيمة المعيارية

أولاً - مقاييس التشتت النسبي Coefficient of Variation

يستخدم هذا النوع من المقاييس لمقارنة تشتت مجموعتين من البيانات أو ظاهرتين أو توزيعين حيث يتم الاعتماد في عملية المقارنة على مقاييس التشتت النسبي (c.v.) Coefficient of variations والتي يعبر عنها من خلال معامل الاختلاف المعياري والذي يحسب من خلال المعادلات التالية:

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

أو

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

معادلة حساب الربيع الأول Q1

الربيع الأول Q1:

$$Q_1 = L_{Q_1} + \frac{\frac{n}{4} - F_a}{F_b - F_a} \times I_{Q_1}$$

F_a التكرار المتجمع السابق للفئة الربيعية الأولى

Q_1 قيمة الربيع الأدنى أو الأول

F_b التكرار المتجمع اللاحق للفئة الربيعية الأولى

L_{Q_1} الحد الأدنى لبداية الفئة الربيعية الأولى

I_{Q_1} طول الفئة الربيعية الأولى

k_{Q_1} ترتيب الربيع الأول

$$Q_3 = L_{Q_3} + \frac{3(n) - F_a}{F_b - F_a} \times I_{Q_3}$$

الربيع الثالث Q3:

Q_3 قيمة الربيع الأدنى أو الثالث F_a التكرار المتجمع السابق للفئة الربيعية الثالثة

L_{Q_3} الحد الأدنى لبداية الفئة الربيعية الثالثة F_b التكرار المتجمع اللاحق للفئة الربيعية

k_{Q_3} ترتيب الربيع الثالث I_{Q_3} طول الفئة الربيعية الثالثة

مثال: البيانات التالية تعبر عن توزيع الوحدات السكنية حسب الإيجار السنوى بأحد الأحياء:

١٨ - ١٤	-١٢	-١٠	-٦	الإيجار بالألف ريال
١٣	١٢	٢٠	١٥	عدد الوحدات السكنية

المطلوب:

حساب:

معامل الاختلاف للإيجار السنوى

معامل الاختلاف الربيعى للإيجار السنوى

الحل تفصيلا في الكتاب

ويتضح لنا من الحل السابق أن:

• معامل الاختلاف للإيجار السنوي للوحدات السكنية بلغ ٢٤%

• معامل الاختلاف الربيعى للإيجار السنوى للوحدات السكنية بلغ ١٥,٤٩٤%

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

ثانيا: القيمة المعيارية Standardized values

وهي تلك القيمة التي تقيس مدى انحراف قيمة مفردة ما من مفردات الدراسة عن الوسط الحسابى لها وذلك بوحدات من الانحراف المعيارى، ويرمز للقيمة المعيارية بالرمز Z حيث أن:

$$Z = \frac{x - \bar{x}}{S}$$

مثال: حصل أحد الطلاب في مقرر المحاسبة على (٨٠) درجة حيث بلغ متوسط درجات الطلاب في اختبار المحاسبة (٨٣) درجة بانحراف معياري (٥). بينما حصل في اختبار مقرر الرياضيات على (٧٠) درجة حيث بلغ متوسط درجة الطلاب في اختبار الرياضيات (٦٥) درجة بانحراف معياري قدرة (٥) درجات .

المطلوب:

هل يمكن القول بأن درجات الطالب في مقرر المحاسبة أفضل من درجته في مقرر الرياضيات ؟

الحل تفصيلا في الكتاب

يتضح لنا من الحل أن القيمة المعيارية لدرجة الطالب في مقرر الرياضيات هي (+1) مما يعنى أن الدرجة التي حصل عليها الطالب أكبر من متوسط درجات الطالب بينما بلغت القيمة المعيارية للدرجة التي حصل عليها الطالب في مقرر المحاسبة (-0.6) مما يدل على أن الدرجة التي حصل عليها الطالب أقل من متوسط الدرجات التي حصل عليها الطلاب .

المحاضرة الثانية عشر

مقاييس الإلتواء والتفلطح

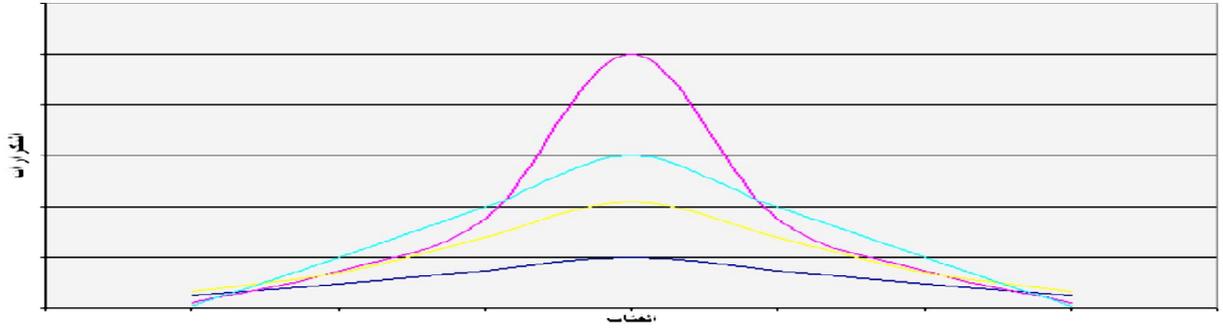
أولاً: مقاييس الإلتواء Skewness Measures

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

المنحنى المتماثل Symmetrical Curve

هو المنحنى الذى إذا قسمناه إلى نصفين إنطبق هذان النصفان على بعضهما البعض تماماً

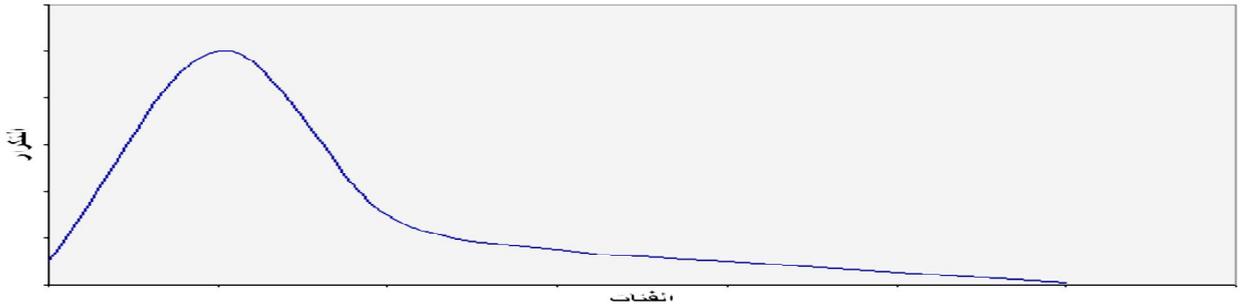
شكل يوضح منحنيات التوزيع المتماثل



المنحنيات الملتوية Skewed

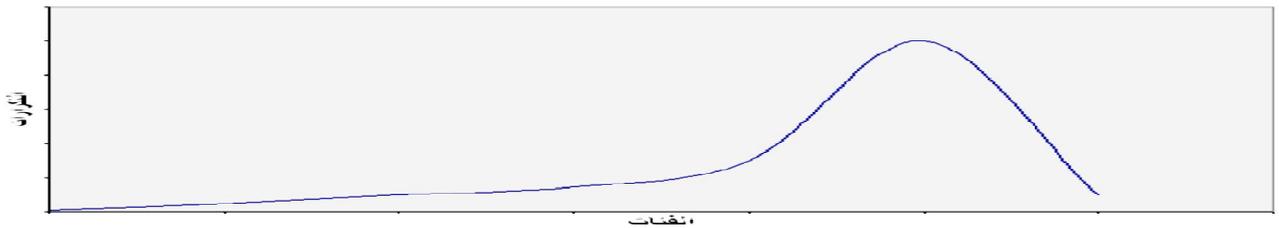
إن الكثير من التوزيعات الإحصائية تبتعد عن التماثل بتركز تكراراتها إما عند أصغر القيم فيصبح المنحنى ملتويًا جهة اليمين أو إلتواء موجب كما يظهر فى الشكل التالي:

شكل يوضح منحنى ملتوى جهة اليمين



أما فى حالة تركز التكرارات عند أكبر القيم فيسمى المنحنى فى تلك الحالة منحنى ملتوي جهة اليسار (إلتواء سالب) كما يظهر من الشكل التالي:

شكل يوضح منحنى ملتوى جهة اليسار



ويمكن قياس الإلتواء من خلال معامل الإلتواء **SK** والذي يفيدنا في الحكم على مدى تماثل أو إلتواء التوزيع **تتعدد مقاييس الإلتواء إلا أن من أهمها:**
معامل الإلتواء لبيرسون والذي يكون في أحد صورتين التاليتين:

$$SK = \frac{3(\bar{x} - Med)}{S}$$

أو

$$SK = \frac{\bar{x} - Mod}{S}$$

وحيث أنه لا يمكن حساب معامل الإلتواء لبيرسون في حالة المنحنيات التي تكون شديدة الإلتواء أو في حالة التوزيعات التكرارية المفتوحة.

لذلك يمكن الاعتماد على مقياس الإلتواء لباولي **SKB الذي يعرف كما يلي:**

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

معادلة حساب الوسيط **Med**

$$Med = L_{Med} + \frac{k_{Med} - F_a}{F_b - F_a} \times I_{Med} : \text{الوسيط } Med$$

Med قيمة الوسيط
 L_{Med} الحد الأدنى لبداية الفئة الوسيطة
 k_{Med} ترتيب الوسيط
 F_a التكرار المتجمع السابق للفئة الوسيطة
 F_b التكرار المتجمع اللاحق للفئة الوسيطة
 I_{Med} طول الفئة الوسيطة

معادلة حساب الربع الأول **Q1**

$$Q_1 = L_{Q_1} + \frac{\frac{n}{4} - F_a}{F_b - F_a} \times I_{Q_1} : \text{الربع الأول } Q_1$$

Q_1 قيمة الربع الأدنى أو الأول
 L_{Q_1} الحد الأدنى لبداية الفئة الربعية الأولى
 k_{Q_1} ترتيب الربع الأول
 F_a التكرار المتجمع السابق للفئة الربعية الأولى
 F_b التكرار المتجمع اللاحق للفئة الربعية الأولى
 I_{Q_1} طول الفئة الربعية الأولى

معادلة حساب الربيع الثالث Q₃

$$Q_3 = L_{Q_3} + \frac{\frac{3(n)}{4} - F_a}{F_b - F_a} \times I_{Q_3} \quad \text{الربيع الثالث } Q_3:$$

Q_3 قيمة الربيع الأدنى أو الثالث

L_{Q_3} الحد الأدنى لبداية الفئة الربيعية الثالثة

k_{Q_3} ترتيب الربيع الثالث

F_a التكرار المتجمع السابق للفئة الربيعية الثالثة

F_b التكرار المتجمع اللاحق للفئة الربيعية الثالثة

I_{Q_3} طول الفئة الربيعية الثالثة

مثال: البيانات التالية تعبر عن توزيع الوحدات السكنية حسب الإيجار السنوي بأحد الأحياء في أحد المدن:

١٨-١٤	-١٢	-١٠	-٦	الإيجار بالآلاف ريال
١٣	١٢	٢٠	١٥	عدد الوحدات السكنية

المطلوب:

حساب معامل الإلتواء لتوزيع الإيجار السنوي للوحدات السكنية.

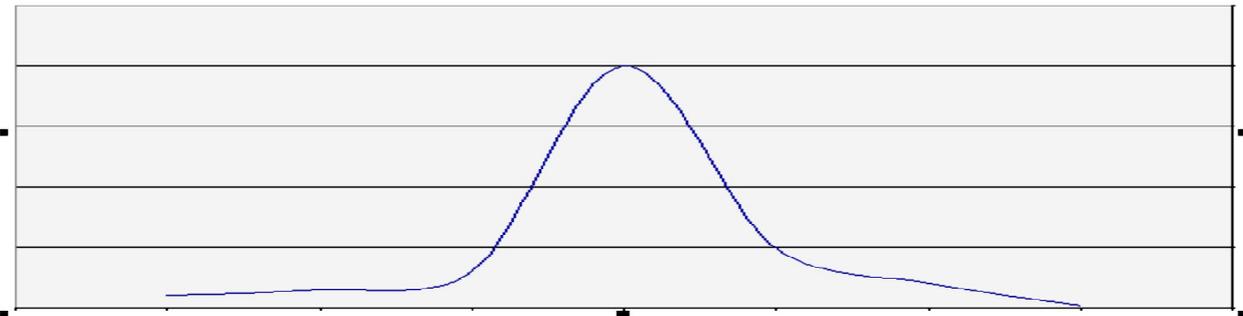
الحل تفصيلا في الكتاب

ويظهر لنا من النتيجة لجميع المعادلات الخاصة بحساب معامل الإلتواء وجود التواء موجب جهة اليمين إلا أن قيمة معامل الإلتواء صغيرة تقترب من الصفر مما يدل أيضا على أن التوزيع قريب من التماثل. ونتيجة لوجود اختلاف في الأصل الرياضي لكل من المعادلات الثلاث السابقة لذا نجد أن قيمة معامل الإلتواء تختلف. إلا أنه كما سبق وذكرنا بأنه يفضل استخدام معامل الإلتواء لبيرسون في أي من صيغتيه في حالة البيانات غير المبوبة وكذلك الجداول التكرارية المغلقة أما في حالة الجداول التكرارية المفتوحة فيفضل استخدام معامل الإلتواء لباولي.

ثانيا: التفلطح Kurtosis

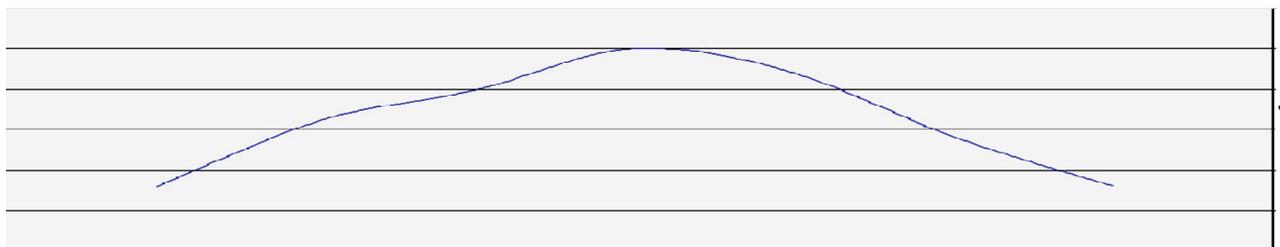
يقصد بالتفلطح مقدار التدبب (الارتفاع أو الإنخفاض) في قمة المنحنى مقارنة بقمة منحنى التوزيع الطبيعي. وتكون قيمة معامل التفلطح **صفر** في حالة التوزيع الطبيعي المعياري. ففي حالة ما يكون معامل التفلطح للبيانات الأصلية **أكبر من ٣** يكون المنحنى مدبب لأعلى كما بالشكل التالي:

شكل يوضح المنحنى المدبب



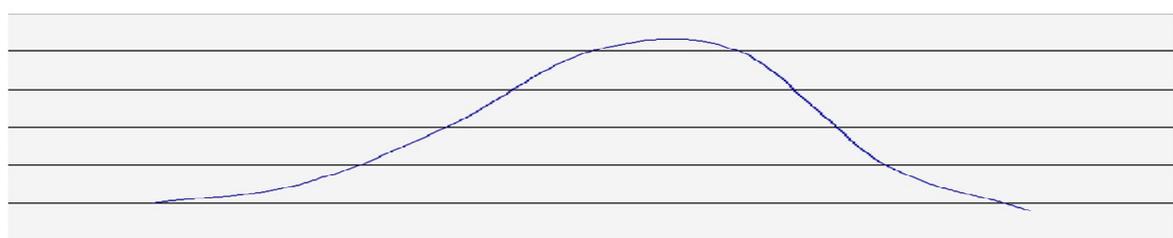
أما في حالة ما يكون معامل التفلطح للبيانات الأصلية **أقل من ٣** يعني ذلك أن المنحنى مفلطح كما يتضح من الشكل التالي:

شكل يوضح المنحنى المفلطح



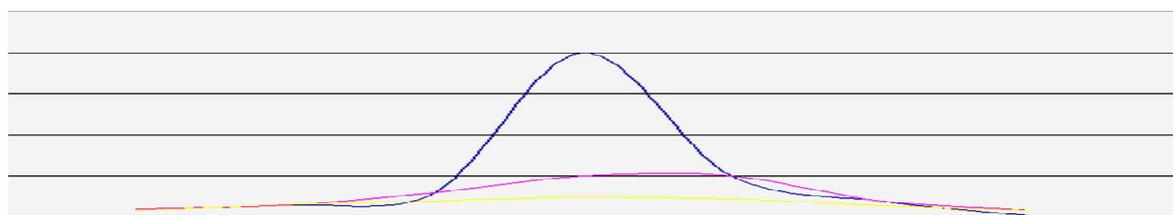
أما في حالة ما يكون معامل التفلطح **يساوي ثلاثة** يكون المنحنى متوسط التفلطح و يكون بالشكل التالي:

شكل يوضح المنحنى متوسط التفلطح



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

شكل يوضح المنحنيات الثلاث معا لتمديد و متوسط التفلطح و المنفلطح



ويتم قياس معامل التفرطح **KU** باستخدام الربيعات والمئينيات من خلال المعادلة التالية:

$$KU = \frac{Q_3 - Q_1}{2 (P_{0.90} - P_{0.10})}$$

حيث يشير:

إلى المئين التسعين والذي يعبر عن ٩٠% من المفردات تكون أقل منه و ١٠% منها أكبر منه	$P_{0.90}$
إلى المئين العاشر (العشير) والذي يعبر عن ١٠% من المفردات تكون أقل منه و ٩٠% منها أكبر منه	$P_{0.10}$

مثال: البيانات التالية تعبر عن توزيع الوحدات السكنية حسب الإيجار السنوى بأحد الأحياء في أحد المدن:

الإيجار بالألف ريال	-٦	-١٠	-١٢	١٤-١٨
عدد الوحدات السكنية	١٥	٢٠	١٢	١٣

المطلوب:

حساب معامل التفلطح لتوزيع الإيجار السنوى للوحدات السكنية.

الحل تفصيلا في الكتاب

ويتضح لنا أن **معامل التفلطح أقل من ٣** مما يدل على أن **المنحنى مفلطح**

أى أن المشاهدات (التكرارات) موزعة على الفئات المختلفة للإيجار السنوى ولا يوجد تركيز بدرجة كبيرة فى أحد الفئات على حساب باقى الفئات الأخرى.

المحاضرة الثالثة عشر تحليل الارتباط - ١

بالرغم من أن مقاييس العلاقة تختلف عما سبقها من مقاييس ، فهي تتعلق بدراسة العلاقة بين متغيرين **(الإنتاجية والجودة)** مثلا، بينما المقاييس السابقة فتهتم بدراسة الفروق بين المتغيرات .

وعندما نقول **مقاييس العلاقة** نعني بذلك تلك المقاييس التي تبين درجة العلاقة والارتباط بين متغيرين أو أكثر مثلا، كأن يكون الهدف معرفة هل هناك علاقة بين مستوى الإنتاجية وجودة المنتج في مصنع ما؟، أي هل كلما زادت الإنتاجية تقل جودة المنتج أو العكس .

معامل الارتباط: هو تعبير يشير إلى المقياس الإحصائي الذي يدل على مقدار العلاقة بين المتغيرات سلبية كانت أم إيجابية، وتراوح قيمته بين الارتباط الموجب التام (+١) وبين الارتباط السالب التام (-١) .

العلاقة الطردية بين المتغيرات: هو تعبير يشير إلى تزايد المتغيرين المستقل والتابع معا، فإذا كانت الإنتاجية مرتفعة، ومستوى الجودة مرتفع، يقال حينئذ أن بينهما ارتباط موجب، وأعلى درجة تمثله هي (+١) .

العلاقة العكسية بين المتغيرات: هو تعبير يشير إلى تزايد في متغير يقابله تناقص في المتغير الآخر، فإذا كانت الإنتاجية منخفضة ومستوى الجودة مرتفع، يقال حينئذ أن بينهما ارتباط سالب، وأعلى درجة تمثله هي (-١) .

ومن الطبيعي ملاحظة أن الارتباط الكامل لا وجود له في الظواهر الطبيعية، وأن معامل الارتباط الناتج في الأبحاث والدراسات الإنسانية والاجتماعية يكون عادة كسرا موجبا أو سالبا.

والجدول التالي يوضح أنواع العلاقات بين المتغيرات كما يصفها معامل الارتباط:

قيمة معامل الارتباط	نوع العلاقة
+١	طردية كاملة
+ كسر (قيمة موجبة)	طردية ناقصة
صفر	صفرية
- كسر (قيمة سالبة)	عكسية ناقصة
-١	عكسية كاملة

إن معامل الارتباط التام الموجب (+١) يعنى التغير في اتجاه واحد في كلا الظاهرتين مع بقاء الأوضاع النسبية لوحداث الظاهرة ثابتة، سواء كان هذا التغير في اتجاه الزيادة **(أي زيادة قيم الظاهرة الأولى تتبعها زيادة في قيم الظاهرة الأخرى)**، أو في اتجاه النقص **(أي نقص قيم الظاهرة الأولى يتبعها نقص في قيم الظاهرة الأخرى)** .

طرق التعرف على العلاقة بين متغيرين وحسابها

أولا: طريقة شكل الانتشار Scatter Diagram :

هناك وسيلة مبدئية يعرف الباحث من خلالها نوع الارتباط بين المتغيرين وما إذا كان الارتباط قوياً وضعيفاً أو منعدماً، وما إذا كانت العلاقة خطية أو غير خطية، موجبة أو سالبة. هذه الوسيلة هي " **شكل الانتشار** " والتي تصلح إذا كان المتغيران كميّين. وجدير بالذكر أن هذه وسيلة مبدئية تساعد فقط في معرفة نوع الارتباط ولا تعتبر بديلاً عن الطرق الإحصائية التي سوف نتناولها بالتفصيل في هذه المحاضرة.

والمقصود **بشكل الانتشار** هو تمثيل قيم الظاهرتين بيانياً على المحورين، المتغير الأول X على المحور الأفقي، والمتغير الثاني Y على المحور الرأسي، حيث يتم تمثيل كل زوج $Pair$ من القيم بنقطة، فنحصل على شكل يمثل كيفية انتشار القيم على المستوى، وهو الذي يسمى شكل الانتشار. وطريقة انتشار القيم تدل على وجود أو عدم وجود علاقة بين المتغيرين ومدى قوتها ونوعها. فإذا كانت تتوزع بشكل منتظم دل ذلك على وجود علاقة (**يمكن استنتاجها**)، أما إذا كانت النقاط مبعثرة ولا تنتشر حسب نظام معين دل ذلك على عدم وجود علاقة بين المتغيرين أو أن العلاقة بينهما ضعيفة. والأشكال التالية تظهر بعض أشكال الانتشار المعروفة :

الشكل الأول :

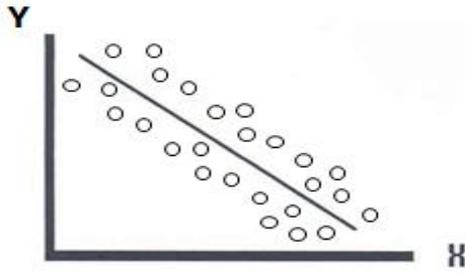
إذا وقعت جميع النقاط على خط مستقيم، دل ذلك على أن العلاقة بينهما خطية وأنها ثابتة أو تامة. وهذه تمثل أقوى أنواع الارتباط بين المتغيرين " ارتباط تام ". فإذا كانت العلاقة طردية فإن " الارتباط طردي تام " كما في الشكل الأول (أ). ومثاله العلاقة بين الكمية المشتراة من سلعة والمبلغ المدفوع لشراء هذه الكمية. أما إذا كانت العلاقة عكسية (وجميع النقاط تقع على خط مستقيم واحد فإن " الارتباط عكسي تام " كما في الشكل الأول (ب).

ومثال على ذلك العلاقة بين السرعة والزمن

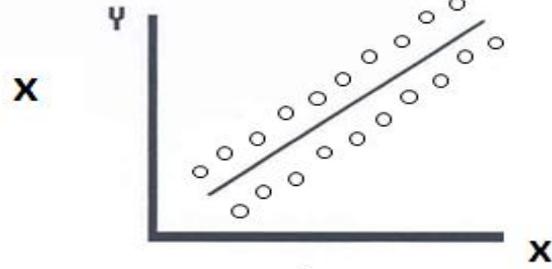


الشكل الثاني :

أما إذا كانت النقاط تأخذ شكل خط مستقيم ولكن لا تقع جميعها على الخط قيل أن العلاقة خطية (موجبة أو سالبة) كما في الشكل الثاني أ، ب.



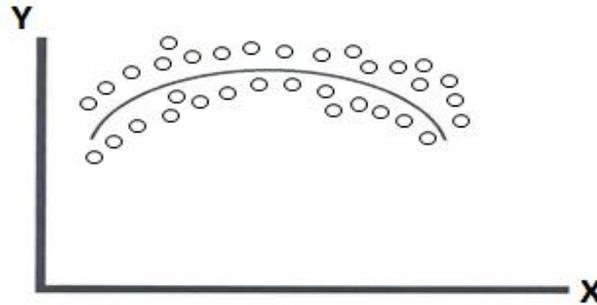
الشكل الثاني (ب)
ارتباط سالب قوي
(ارتباط خطي عكسي)



الشكل الثاني (أ)
ارتباط موجب قوي
(ارتباط خطي طردي)

الشكل الثالث :

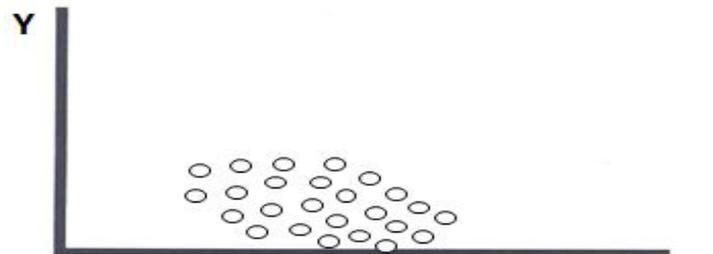
وإذا كانت العلاقة تأخذ شكل منحنى فإن الارتباط لا يكون خطياً "ارتباط غير خطي" Non Linear " Correlation كما في الشكل الثالث :



الشكل الثالث
(ارتباط غير خطي)

الشكل الرابع :

أما إذا كانت النقاط تتبعثر بدون نظام معين فإن ذلك يدل على عدم وجود علاقة بين المتغيرين (أو أن العلاقة بينهما ضعيفة جداً) كالعلاقة مثلاً بين دخل الشخص وطوله كما في الشكل الرابع :



الشكل الرابع
(لا توجد علاقة)

ثانياً: معامل الارتباط Correlation Coefficient :

يقاس الارتباط بين متغيرين بمقياس إحصائي يسمى "معامل الارتباط" ويعكس هذا المقياس درجة أو قوة العلاقة بين المتغيرين واتجاه هذه العلاقة. وتنحصر قيمة معامل الارتباط بين + ١ ، - ١ .

فإذا كانت قيمة معامل الارتباط تساوي + ١ فمعنى ذلك أن الارتباط بين المتغيرين طردي تام، وهو أقوى أنواع الارتباط الطردي بين متغيرين.

وإذا كانت قيمة معامل الارتباط تساوي - ١ فمعنى ذلك أن الارتباط بين المتغيرين عكسي تام، وهو أقوى أنواع الارتباط العكسي بين متغيرين.

وإذا كانت قيمة معامل الارتباط تساوي صفر، فمعنى ذلك أنه لا يوجد ارتباط بين المتغيرين.

وكلما اقتربت قيمة معامل الارتباط من + ١ أو - ١ كلما كان الارتباط قوياً، وكلما اقترب من الصفر كلما كان الارتباط ضعيفاً.

والخلاصة :

أنه كلما كانت العلاقة قوية بين المتغيرين كلما اقترب معامل الارتباط من + ١ أو - ١ فإذا وصلت قيمة المعامل إلى + ١ أو - ١ كان الارتباط تاماً بين المتغيرين. وأنه كلما كانت العلاقة ضعيفة بين المتغيرين كلما اقترب معامل الارتباط من الصفر، فإذا وصلت قيمة المعامل إلى الصفر كان الارتباط منعماً بين المتغيرين. ومعنى ذلك أيضاً أنه لا يوجد ارتباط بين متغيرين تكون قيمة المعامل فيه أكبر من + ١ ولا أصغر من - ١. ويمكن تمثيل قوة العلاقة بالشكل التالي:

ارتباط عكسي					ارتباط طردي					
قوي جداً	قوي	متوسط	ضعيف	ضعيف جداً	ضعيف جداً	ضعيف	متوسط	قوي	قوي جداً	
-١	-٠,٩	-٠,٧	-٠,٥	-٠,٣	٠	٠,٣	٠,٥	٠,٧	٠,٩	١
تام					منعدم					تام

معامل بيرسون للارتباط الخطي البسيط Person's Correlation Coefficient

يعتبر معامل الارتباط الخطي البسيط لبيرسون Person's Correlation Coefficient والذي سنرمز له بالرمز r_p من أكثر الأدوات الإحصائية استخداماً في تحديد قوة العلاقة بين متغيرين كما يستعمل لتحديد مدى وجود علاقة خطية بين متغيرين. وهناك أكثر من صيغة يمكن الاعتماد عليها في حساب معامل الارتباط الخطي 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 أسهل وأبسط:

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

وتتراوح قيمة معامل الارتباط بين الواحد الصحيح الموجب و الواحد الصحيح السالب أى أن قيمة معامل تكون كالتالي:

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

والارتباط غالبا قيمته كسر أى اقل من الواحد الصحيح

ولتحديد نوع العلاقة نعلم على اشارة معامل الارتباط فإذا كانت الإشارة:
 • موجبة فإن العلاقة تكون **طردية**
 • سالبة فإن العلاقة تكون **عكسية**

ولتحديد قوة العلاقة نعلم على قيمة معامل الارتباط فإذا كانت القيمة:
 • من أكبر من صفر إلى أقل من 0.3 فتكون **علاقة ضعيفة جدا**
 • من أكبر من 0.3 إلى أقل من 0.5 تكون **علاقة ضعيفة**
 • من أكبر من 0.5 إلى أقل من 0.7 تكون **علاقة متوسطة**
 • من أكبر من 0.7 إلى أقل من 0.9 تكون **علاقة قوية**
 • من أكبر من 0.9 إلى أقل من 1.00 تكون **علاقة قوية جدا**
 • الواحد الصحيح تكون **علاقة تامة**

• إما إذا كانت قيمة معامل الارتباط تساوى صفر **فلا توجد علاقة خطية** او ارتباط بينهما أى يكون المتغيرين مستقلين عن بعضهما البعض وتكون **العلاقة منعدمة**

فمثلا إذا كانت قيمة معامل الارتباط r_p كالتالى فإن تفسيره يكون:

تفسير معامل الارتباط	قيمة
ارتباط طردى قوى جدا	0.91
ارتباط عكسى قوى	-0.87
ارتباط عكسى ضعيف جدا	-0.21
ارتباط طردى ضعيف	0.43
ارتباط طردى تام	1
ارتباط عكسى متوسط	-0.51

مثال: فيما يلى بيان بالمنفق على الاعلان والمبيعات لأحد المنتجات فكانت بالمليون ريال كما يلى:

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:

ارسم شكل الانتشار يوضح العلاقة بين المنفق على الاعلان و المبيعات ؟
 احسب معامل الارتباط الخطى البسيط (بيرسون)، مع التعليق

الحل تفصيلا في الكتاب

ومن أهم خصائص معامل الارتباط الخطى البسيط لبيرسون أنه لا يعتمد على قيم المتغيران نفسها عند حساب قيمته وإنما يعتمد على مقدار التباعد بين هذه القيم بعضها البعض. لذلك لا يتأثر معامل الارتباط الخطى البسيط بأى عمليات جبرية يتم إجراؤها على بيانات أى من المتغيرين أو أحدهما من جمع أو طرح أو ضرب أو قسمة

مثال: فى بيانات المثال السابق إذا أكتشفت إدارة الشركة أن البيانات تم تجميعها وحسابها بطريقة خاطئة حيث يجب إضافة ٥ مليون ريال إلى جميع قيم المنفق على الإعلان. كما أن المبيعات يجب مضاعفة قيمتها لجميع القيم.
المطلوب:
أحسب معامل الارتباط فى هذه الحالة بين المنفق على الإعلان والمبيعات.

الحل تفصيلا في الكتاب

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:

وجد أن المنفق على الاعلان يفسر نسبة (0.8756^2) أى 76.675 % من التغير فى قيمة المبيعات بينما 23.32 % من التغير فى المبيعات ترجع إلى عوامل أخرى منها الخطاء العشوائى .

المحاضرة الرابعة عشر

تحليل الارتباط - ٢

سيتم في هذه المحاضرة استعراض المواضيع التالية:

أولاً: معامل ارتباط سبيرمان

ثانياً: معامل الإقتران

ثالثاً: معامل التوافق

معامل ارتباط الرتب لسبيرمان Spearman's Rank Correlation Coefficient

معامل الارتباط لبيرسون r_s لا يمكن استخدامه في حساب قوة العلاقة بين متغيرين الا اذا كانت البيانات المتوافره عنهما في صورة كمية فقط، أما اذا كانت البيانات في صورة وصفية فلا يمكن تطبيق معامل ارتباط بيرسون وحساب الارتباط بين المتغيرين محل الدراسة.

أما في 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 - جيد جداً - جيد - مقبول - ضعيف) وكذلك قوة المركز المالي (جيد - متوسط - ضعيف) ودرجة الموافقة على الرأي في اسئلة الاستبانة (موافق تماماً - موافق - محايد - غير موافق - غير موافق على الاطلاق).

ويتم حساب معامل الارتباط الرتب لسبيرمان r_s باستخدام المعادلة التالية:

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

حيث أن:

d الفرق بين رتبة المتغيرين

n عدد المشاهدات

ملاحظات يجب مراعاتها عند ترتيب المتغيرات:

١. يتم ترتيب قيم مشاهدات المتغير x وتسمى القيم الترتيبية للمتغير x "رتب x " وكذلك الامر للمتغير y تسمى بـ "رتب y ". والترتيب يكون تصاعدياً أو تنازلياً ولكن أهم شيء هو اذا كان ترتيب x تصاعدي لا بد ان يكون ترتيب y تصاعدي ايضاً والعكس صحيح.

٢. في حالة الترتيب التصاعدي مثلاً يتم اعطاء أقل قيمة الرتبة ١ والقيمة التي هي أكبر منها الرتبة ٢ وهكذا

٣. في حالة تكرار أو تساوى بعض القيم لأي متغير تعطى كل منهم رتبة كما لو كانت القيم غير متساوية ثم نحسب الوسط الحسابي (مجموع الرتب ÷ عددها) لتلك الرتب ويعطى الوسط الحسابي كرتبة تلك القيم المتساوية

مثال: فيما يلي بيان بالمنفق على الاعلان والمبيعات لأحد المنتجات فكانت بالمليون ريال كمايلي:

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

المطلوب:

أحسب معامل الارتباط لسيرمان بين المنفق على الاعلان و المبيعات ؟

الحل تفصيلا في الكتاب

مثال: البيانات التالية تمثل التقديرات التي حصل عليها عشر طلاب في مقررى المحاسبة والقانون:

المحاسبة	ممتاز	جيد جدا	جيد	مقبول	ضعيف	جيد	مقبول	جيد	جيد جدا	ممتاز	جيد جدا	مقبول
القانون	جيد	جيد	مقبول	ممتاز	جيد جدا	جيد جدا	جيد	مقبول	جيد	جيد جدا	جيد جدا	مقبول

المطلوب:

أحسب معامل الارتباط المناسب.

الحل تفصيلا في الكتاب

معامل الإقتران Conjunction Coefficient

ويستخدم معامل الإقتران في حساب العلاقة الارتباطية بين المتغيرات الوصفية التي ليس في طبيعتها صفة الترتيب أى الوصفية الأسمية التي يكون لها زوج من الصفات مثل:

النوع (ذكر - انثى)، والحالة التعليمية (متعلم - غير متعلم) وعلى ذلك إذا كان لدينا متغيران لدي كلا منهما زوج من الصفات فيكون جدول تكرارات الصفات المشتركة بينهما على الصورة التالية:

		Y	X
الصفة الثانية لـ y	الصفة الأولى لـ y	الصفة الأولى لـ x	الصفة الثانية لـ x
B	A		
D	C		

حيث أن A , B , C , D تشير إلى التكرارات المشتركة بين صفات المتغيرين،

ويمكن حساب معامل الإقتران 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 ٢٠٠ شخص سؤاليين هما:

هل انت متعلم؟ نعم لا

هل انت ملتحق بأى عمل؟ نعم لا

وبتجميع الاجابات تم عمل جدول الإقتران التالي:

	العمل	التعليم
يعمل	113	متعلم
لايعمل	49	أمى
		23
		15

المطلوب: أحسب معامل الإقتران ؟

معامل التوافق Concordance Coefficient

ويستخدم معامل التوافق لحساب الارتباط بين المتغيرات الوصفية الاسمية والتي يكون لصفاتها قيم أكثر من ٢، مثل الحالة الاجتماعية (اعزب - متزوج - متزوج ويعول - أرمل - مطلق)

وحتى يمكن حسابه يتم إعداد الجدول المزدوج بين صفات المتغيريين ومنه يتضح لنا التكرارات المشتركة بين الصفات التي نعتمد عليها في حساب مقدار يطلق عليه " M "

ويتم حساب معامل التوافق من خلال المعادلة التالية:

$$M = \sum \frac{(f_{ij})^2}{f_{i.} \cdot f_{.j}}$$

حيث أن:

التكرار المشترك بين الصفة i والصفة j	f_{ij}
مجموع صف الصفة i	$f_{i.}$
مجموع عمود الصفة j	$f_{.j}$

أي يتم إيجاد: مربع تكرار كل خلية مشتركة
مجموع الصف \times مجموع العمود
ثم نجمعهم كلهم

الحل تفصيلا في الكتاب

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:

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

الحل تفصيلا في الكتاب

تم بحمد الله وفضله

تمنياتي لكم بالتوفيق

لا تنسوني من صالح دعائكم