مبادئ الإحصاء

كلية آداب

مستوى ثاني

أ.د. عبدالله بن عمر النجار

العام الدراسي ١٤٣٢-١٤٣٣هـ

ملتقى طلاب التعليم عن بعد - جامعة الملك فيصل

أحمد

المحاضرة الأولى

التعريف ببعض المفاهيم الإحصائية

سنتناول فى هذه المحاضرة المواضيع التالية

(١) مقدمة

(٢) مفهوم علم الإحصاء

(٣) المجتمع والعينة

(٤) البيانات

(٥) خطوات العملية الإحصائية

(٦) تمرينات محلولة

(٧) تدريبات للطالب

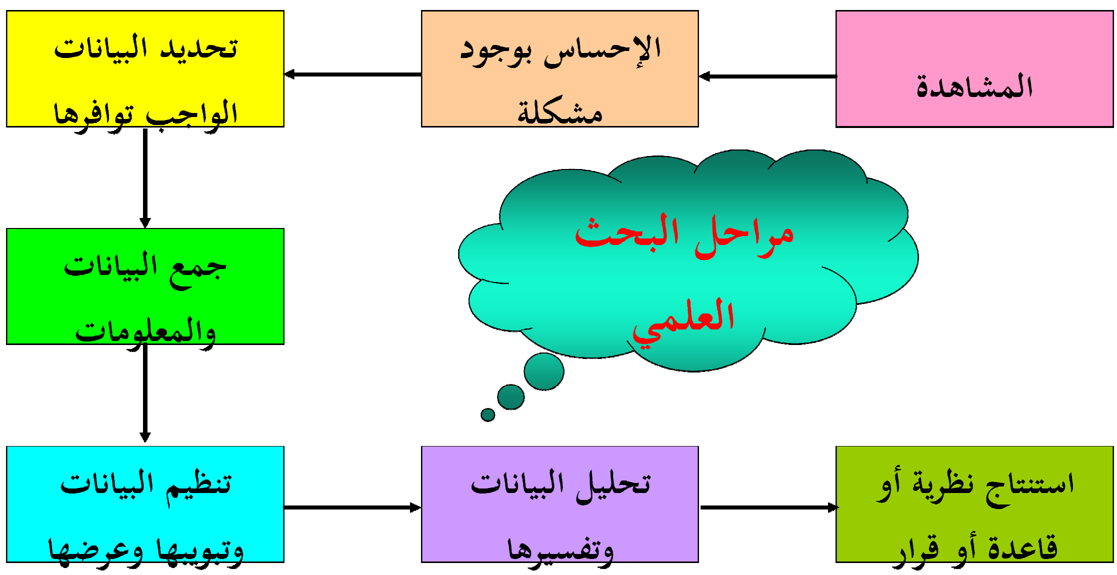
مقدمة

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

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

يأتي بعد ذلك جمع لتلك البيانات من مصادرها المختلفة وتنظيمها وتبويبها وعرضها في صور جدولية أو بيانية ، ثم يتم استخدامها في حساب بعض المقاييس الخاصة بهذه الظواهر وإجراء تحليل لتلك البيانات بما يساعد في تفسير النتائج المختلفة للبيانات واستخدامها في استنتاج نظرية أو قاعدة أو قانون أو المساعدة في اتخاذ القرارات أو التنبوء بنتائج مستوالشكل التالي يمكن أن يوضح الإطار العام لأي بحث علمي

قبلية



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 :

الإحصاء الوصفي

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

الإحصاء الاستقرائي

أو **الاستدلال الإحصائي** وهو يبحث في استقراء النتائج واتخاذ القرارات

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 نرمز له برمز مثل .. x , y , A , B ، فمثلاً :

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

والمتغير (أي الظاهرة تحت الدراسة) إما أن يكون :

تُسمى البيانات عندئذٍ ببيانات نوعية — تُسمى البيانات عندئذٍ ببيانات كمية

متغير نوعي أو **متغير كمي**

أي لا يمكن التعبير عنه **بعدد** [مثل لون العين أو تقدير الطلاب في الأمثلة (١) ، (٥) السابقة]

(١) لون العين	أخضر – أزرق – بني – .....
(٥) تقدير الطلبة	A – B – C – F – D – A – A - ..

أي يمكن التعبير عنه **بعدد** مثل الأطوال أو الأوزان أو أعداد الطلاب

متغير متصل

تُسمى البيانات عندئذٍ ببيانات (كمية) متصلة

وفيها يمكن أن يأخذ المتغير أي قيمة بين قيمتين معينتين كما في الأمثلة (٣) ، (٤) السابقة **[بتعبير آخر هو كمية يمكن أن تُقاس ولا تُعد]**

(٣) أطوال الطلاب	1.5 – 1.52 – 1.71 – 1.83 - ...
(٤) أوزان العاملات	55.2 – 60.1 – 63.25 - ......

أو

متغير متقطع

تُسمى البيانات عندئذٍ ببيانات (كمية) متقطعة

خلاف ذلك كما في المثال (٢) السابق **[أو بتعبير آخر هو كمية يمكن أن تُعد ولا تُقاس]**

عدد الطلاب	15 – 18 – 20 – 25 - ......

خطوات العملية الإحصائية

يمكن تلخيص خطوات أي عملية إحصائية في الآتي :

(أ) جمع البيانات

هي عملية الحصول على القياسات الخاصة بظاهرة معينة وعادةً ما تُسمى البيانات المجمعة **بالبيانات الخام**

(ب) تنظيم وعرض البيانات

هي عملية وضع البيانات السابقة في جداول خاصة وعرضها بطرق مناسبة

(ج) تحليل البيانات

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

(د) استقراء النتائج واتخاذ القرارات

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

تمارين محلولة

اختر الإجابة الصحيحة لكلٍ مما يأتي :

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

(أ) ● علم الإحصاء الوصفي
(ب) ○ علم الإحصاء الاستقرائي
(ج) ○ علم تقنية المعلومات
(د) ○ علم تكنولوجيا المعلومات

---

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

(أ) ○ تحليل البيانات
(ب) ○ استقراء النتائج واتخاذ القرارات
(ج) ○ تنظيم وعرض البيانات
(د) ● جمع 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 وتنبوءات 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 وتُعطي 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) مقاس الأحذية S هو :

(أ) ◯ متغير نوعي
(ب) ◯ متغير كمي متصل
(ج) ◯ متغير كمي متقطع
(د) ◯ خلاف ذلك

(٩) اللعبة الرياضية A التي يفضلها أفراد أسرتك 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 متقطعة　　(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 متقطعة　　(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 متقطعة　　(د) ◯ خلاف ذلك

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

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

(ج) ◯ بيانات كمية متقطعة　　(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

هناك سؤال مهم لابد من الإجابة عليه وهو:

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

في حالة اعتماد البحث على دراسة جميع مفردات المجتمع الإحصائي يسمى ذلك

أسلوب الحصر الشامل

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

أسلوب العينة

اسلوب الحصر الشامل

يمكنا 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

اسلوب العينات

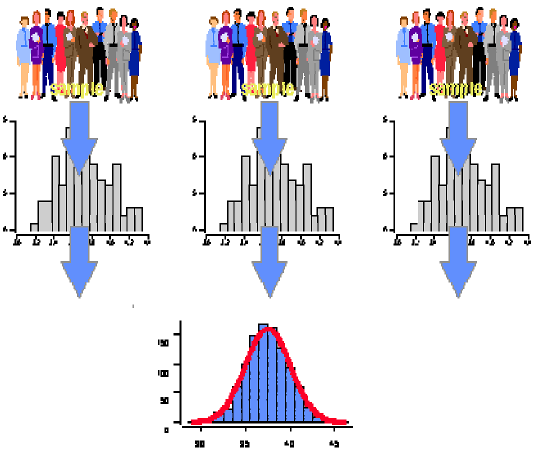
يبدوا هذا الأسلوب على العكس من أسلوب الحصر الشامل حيث تقتصر الدراسة فيه على جزء من المجتمع الإحصائي، لذا فهذا الأسلوب يوفر الوقت و الجهد و التكاليف ويصلح للمجتمعات غير المحدودة. إلا ان أهم عيوب هذا النوع هو ما يسمى بخطأ التحيز Sampling Bias .

أقسام مجتمع البحث

قسم بعض العلماء مجتمع البحث الى قسمين:

المجتمع الكلي للبحث

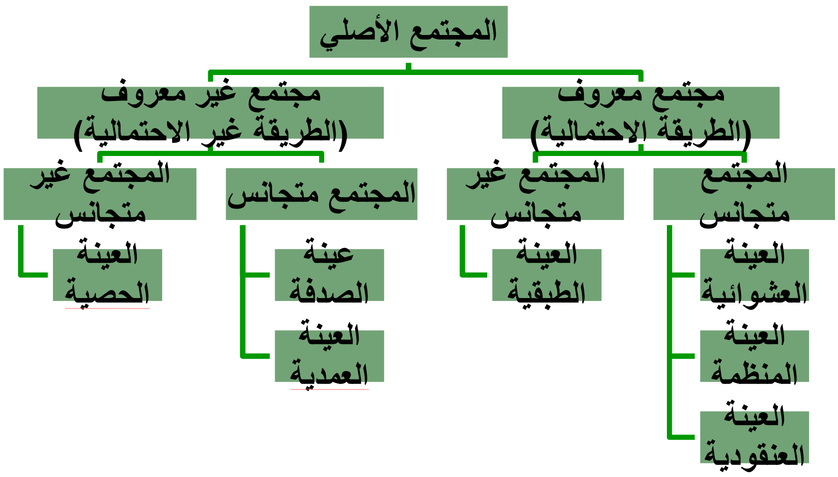
المجتمع الذي يمكن التعرف عليه



<u>مجتمع البحث</u> هو مصطلح علمي يراد به كل من يمكن أن تعمم عليه نتائج البحث

<u>عينة البحث</u> بأنها جزء من المجتمع اختير بطريقة علمية بشرط ان تمثل المجتمع ككل

<u>طرق اختيار العينات</u>

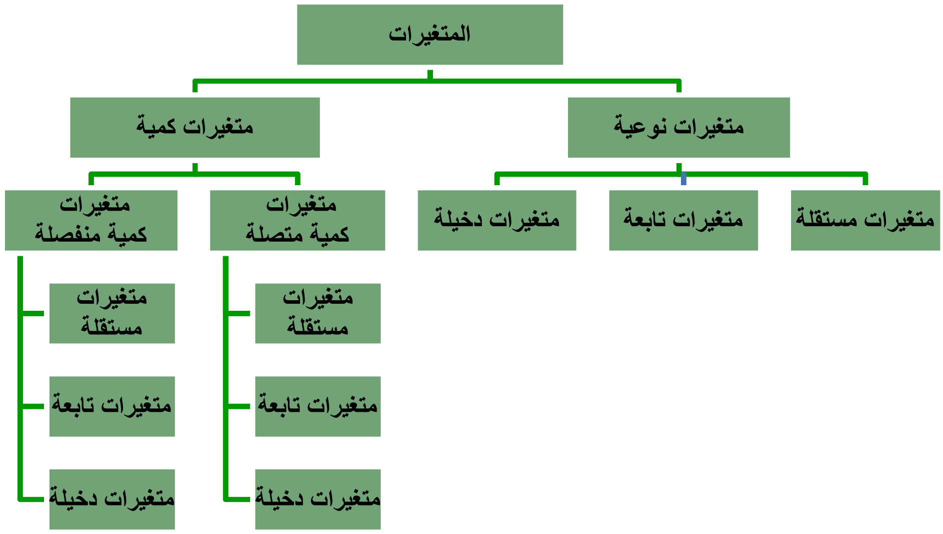


<u>المتغير والثابت في البحث العلمي</u>

<u>المتغير:</u> 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.

<u>الثابت:</u> 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 الغرفـــة. 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 الـ SPSS

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

مثال:

لو كنت تقوم بدراسة إحصائية حول موضوع "واقع استخدام الانترنت في البحث العلمي في الجامعات السعودية"، فإنك ستحتاجين إلى إعداد استبانة تحوي مجموعة من الاسئلة تتعلق بهذا الموضوع، ومن ثم توزيع هذه الاستبانة على عينة ممثلة لمجتمع البحث الذي تريدين أن تعممي نتائج دراستك عليه، وتطلبين من أفراد العينة الإجابة على جميع فقرات الاستبانة، والاستبانة التالية (والتي ستوزع عليكم) كمثال على ذلك.

ولغرض تفريغ 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 جزئيا

تمارين

أرادت باحث معرفة العلاقة بين حب الاستطلاع لدى الطلاب في السنوات الابتدائية وحل المسائل الرياضية، فاختار عشوائيا طلاب السنة الثالثة ثم اختار منهم عشوائيا 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 تصورتك المتغيرات الدخيلة التي من الممكن أن تؤثر على هذه الدراسة ؟**
- **حدد الفرضية التي يحاول الباحث اختبارها في هذه الدراسة ، وما نوعها ؟**
- **ما الوسيلة التي استخدمها الباحثة لجمع البيانات في هذه الدراسة ؟**

المحاضرة الثالثة

العرض الجدولى 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 هى أن تدمج هذه البيانات في صيغة كتابية إلا أن هذه الطريقة يشوبها الكثير من العيوب

أما الطرق الفنية فى عرض البيانات الاحصائية فهى:

العرض الجدولى للبيانات

العرض البيانى للبيانات

وسوف نتناول فى هذه المحاضرة العرض الجدولى للبيانات بينما نتعرض للعرض البيانى للبيانات فى المحاضرة التالية إن شاء الله تعالى.

ويقصد بالعرض الجدولى للبيانات أن يتم تلخيص البيانات محل الدراسة وتصنيفها فى صورة جداول تعبر عن القيم التى أخذها المتغير من خلال البيانات التى 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 بأكبر 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 هيك الجدول من أعمدة وصفوف، ويعتبر ترتيب المعلومات في الاعمدة والصفوف أهم خطوة في تكوين الجدول.

العمود: إن كل جدول يتكون من عمود أو اكثر 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 اما بترقيم الملاحظات او باستعمال علامة (*) .. الخ.

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

رقم الجدول — جدول رقم (٥)

عنوان الجدول — يوضح طلبة جامعة الملك فيصل للعام الجامعي ١٤٢٣هـ

عنوان توضيحي — (مصنفون حسب الجنس)

عنوان العمود

هيكل الجدول

عمود

المستوى*	طالب	طالبة	المجموع
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 تبعا لدرجة تعقيدها الى:

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

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 الى طرفة الآخر فنحصل على التكرار الكلي (مجموعة التكرارات)، (فاذا بدأ من أعلى الى أسفل الجدول) سمي جدول تكراري متجمع صاعد، (واذا بدأ من أسفل الى أعلى الجدول) سمي جدول تكرار متجمع نازل أو هابط.

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

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 والتى 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 العاملات 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 تُعد 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	بنفسجى	أحمر	أخضر
أبيض	أبيض	أحمر	أزرق	أبيض
أزرق	أحمر	أخضر	أحمر	بنفسجى
أخضر	أزرق	أبيض	بنفسجى	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 التى تعرضوا لها خلال العام الماضى فكانت اجاباتهم كما يلى:

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 فى 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 الجدولى للبيانات - ٢

ثانيا: 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 فى شركات الحاسبات الألية بالألف ريال:

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

المطلوب:

عرض البيانات السابقة فى صورة الجدول التكرارى المناسب

الحل تفصيلا فى الكتاب

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

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

أ.عدد المفردات محل الدراسة

ب.انتظام وتوزيع تلك البيانات

ج.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 تركز 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 هى:

١.الجداول التكرارية المنتظمة

٢.الجداول التكرارية غير المنتظمة

٣.الجداول التكرارية المفتوحة

أولا: الجداول التكرارية المنتظمة:

وهى الجداول التى يكون فيها أطوال كل الفئات متساوية

كما تم توضيحة فى المثال السابق

ثانيا: الجداول التكرارية غير المنتظمة:

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

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

ثالثا: الجداول التكرارية المفتوحة:

وتوضحها أشكال الجداول التالية:

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

جدول مفتوح من أسفل

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

جدول مفتوح من أعلى

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

جدول مفتوح من الطرفيين

الجداول التكرارية المتجمعة:

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

اولا- الجدول التكرارى المتجمع الصاعد

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

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

فئات مساحات الاراضى دونم	عدد قطع الاراضي
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 فى 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 الاراضي
1 –	14
3 –	29
5 –	18
7 – 10	9
المجموع	70

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

الحل تفصيلا فى الكتاب

الجدول التكرارى المزدوج:

عند دراستنا لمتغيرين لتحديد العلاقة بينهما مثل العلاقة بين عدد أفراد الأسرة والمستوى التعليمى أو العلاقة بين أجور العامل ودرجة الرضاء الوظيفى أو ماشابة ذلك، فى هذه الحالة لابد من تبويب البيانات بالطريقة التى تسمح باستنتاج أو تحديد العلاقة بين المتغيرين موضوع الدراسة ويتم ذلك من خلال الجدول التكرارى المزدوج كما يتضح من المثال التالى:

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

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

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

المطلوب: إعداد جدول تكرارى مزدوج

الحل تفصيلا فى الكتاب

تمارين محلولة

س ١ : المدى لمجموعة من البيانات المنفصلة هو :

- ○ أكثر القيم تكراراً في البيانات
- ○ أكبر قيمة في البيانات
- ○ أصغر قيمة في البيانات
- ● الفرق بين أكبر وأصغر قيمتين 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 :

● 0.35 ○ 35% ○ 4 ○ 7

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

○ 0.35 ● 35% ○ 4 ○ 7

هامش للإجابة

(٢-أ) $7 = 3 + 2 + 2$

(٢-ب) $4 = 2 + 2$

(٢-ج) $\frac{7}{20} = 0.35$

(٢-د) $0.35 \times 100 = 35\%$

خد بالك : المطلوب نسبة (وليس نسبة مئوية)

أيوه .. ، ده بقى نسبة مئوية

المحاضرة الخامسة

العرض البيانى للبيانات

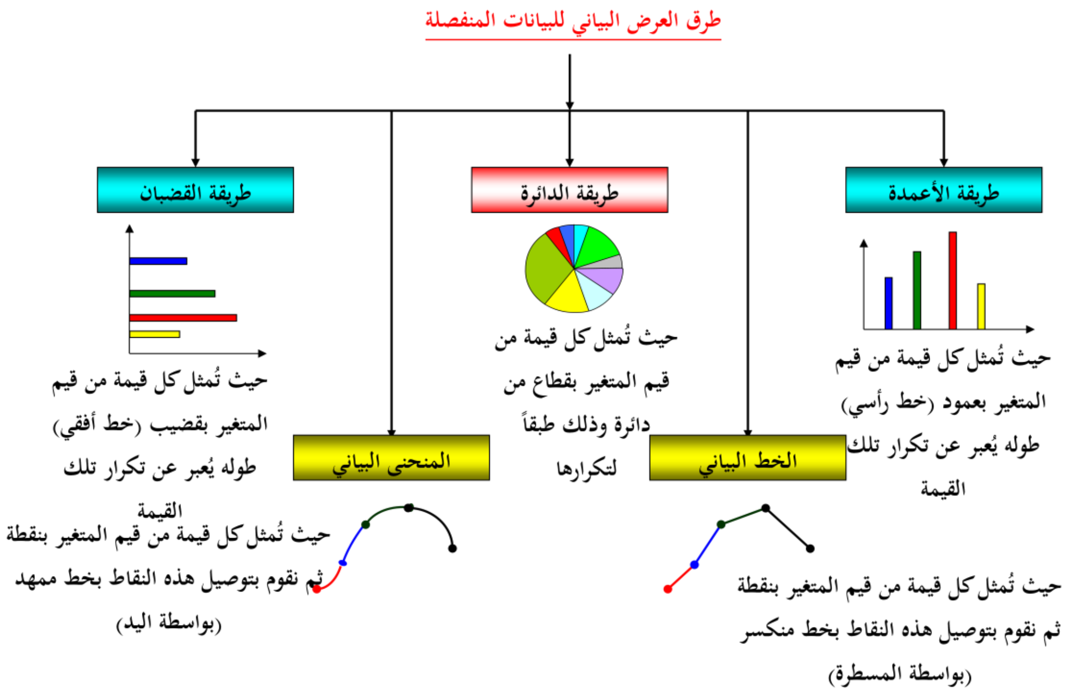
أولا: البيانات غير المبوبة

تعريف الرسوم البيانية:

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

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

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



١ ـ طريقة الأعمدة:

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

أ ـ الأعمدة البيانية البسيطة :

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

مثال: الجدول الآتي يوضح أعداد الطلاب المقيدين باحد الجامعات في السنوات الدراسية من ١٤٢٣ هـ حتي ١٤٢٧ هـ .

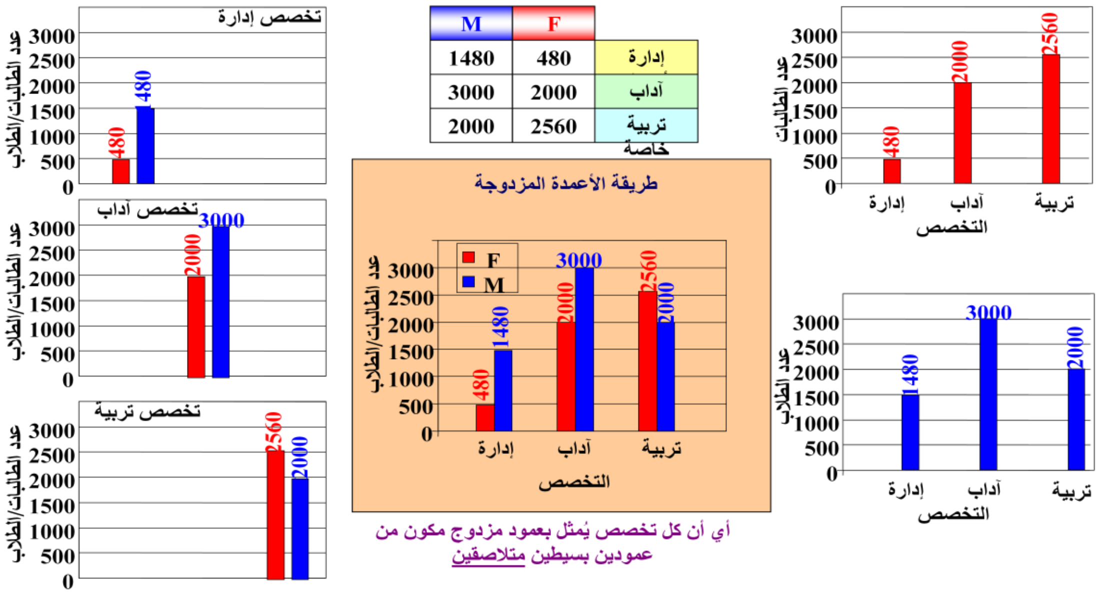
السنة الدراسية	١٤٢٣	١٤٢٤	١٤٢٥	١٤٢٦	١٤٢٧
عدد الطلاب بالالاف	٣	٥	٦	٨	١٠

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 .



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

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

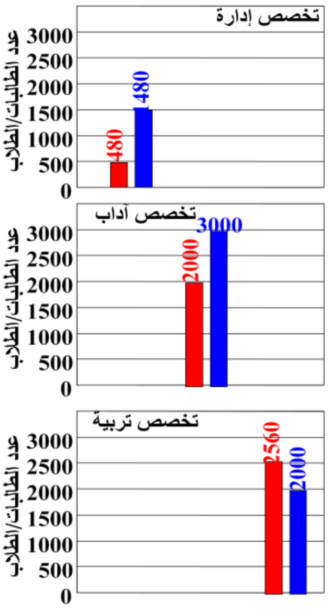
السنة الدراسية		١٤١٩	١٤٢٠	١٤٢١	١٤٢٢	١٤٢٣
عدد الطلبة بالألف	طلاب	٣١	٤٠	٤٥	٤٩	٥٤
	طالبات	٩	١٢	١٦	٢٠	٢٧

المطلوب:
مثل هذه البيانات بيانيا باستخدام الأعمدة البيانية المزدوجة ؟

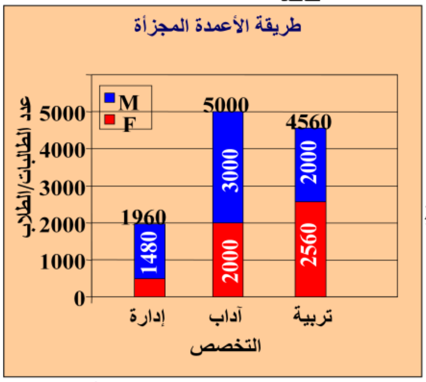
الحل تفصيلا في الكتاب

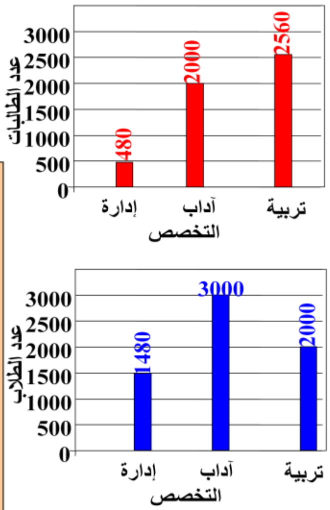
جـ – الأعمدة البيانية المجزأة:

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 .



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





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

مثال: اذا 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
الطــلاب	١٠٠	٢٠٠	٣٠٠	٤٠٠	٥٠٠	٦٠٠
الطالبـــات	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 البيــانية 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 لتميل 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:

$$\text{الزاوية المركزية لقيمة ما} = \frac{\text{تكرار القيمة}}{\text{مجموع التكرارات}} \times ٣٦٠$$

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

وفيما يلي تطبيق ذلك على بيانات إحصائية :

لوكانت لدينا البيانات التالية:

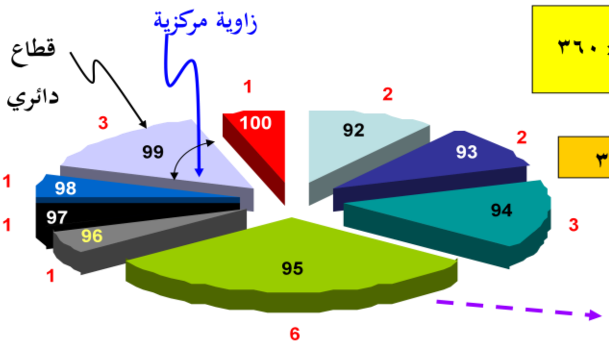
الدرجة x	92	93	94	95	96	97	98	99	100
التكرار f	2	2	3	6	1	1	1	3	1

تحدد زاويتة المركزية بالعلاقة :

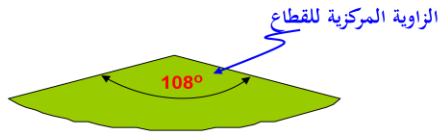
$$\text{الزاوية المركزية لقيمة ما} = \frac{\text{تكرار القيمة}}{\text{مجموع التكرارات}} \times ٣٦٠$$

أو

الزاوية المركزية لقيمة ما = التكرار النسبي للقيمة × ٣٦٠



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

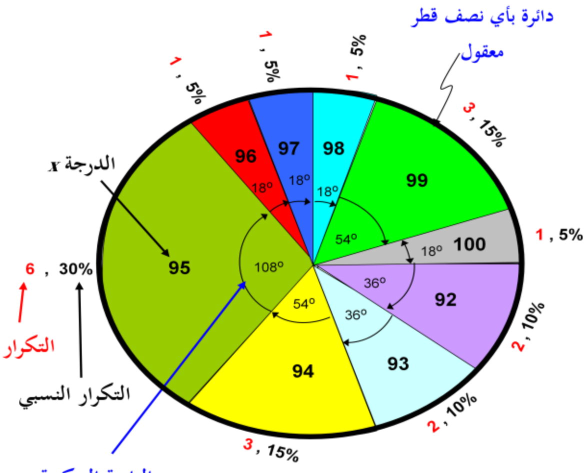


القطاع الخاص بالدرجة "95" ذات التكرار 6
قياس زاويته المركزية تساوي :

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

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

الزاوية المركزية	التكرار f	الدرجة x
(2/20)x360 = 36°	2	92
(2/20)x360 = 36°	2	93
(3/20)x360 = 54°	3	94
(6/20)x360 = 108°	6	95
(1/20)x360 = 18°	1	96
(1/20)x360 = 18°	1	97
(1/20)x360 = 18°	1	98
(3/20)x360 = 54°	3	99
(1/20)x360 = 18°	1	100
مجموع الزوايا = 360°		



طريقة الدائرة لتمثيل البيانات

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

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 بيانيا؟ وبماذا تختلف عن التمثيل البياني باستخدام الدائرة؟

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

١. عندما تكون الكميات المقارنة كثيرة العدد نسبيا.

٢. عند ما تكون الاجزاء المقارنة في فترات زمنية مختلفة.

٣. عندما نرغب في توضيح قيم الاجزاء المقارنة المختلفة للظاهرة موضع البحث وذلك من أجل ابراز المقارنة بين هذه الأجزاء أو توضيح التغير أو التطور عبر الزمن سواء لظاهرة واحدة أو عدة ظواهر بين فترات زمنية مختلفة.

٤. غالبا ما ينصح باستعمال الاعمدة البيانية (بأنواعها المختلفة) مع المتغيرات المنفصلة.

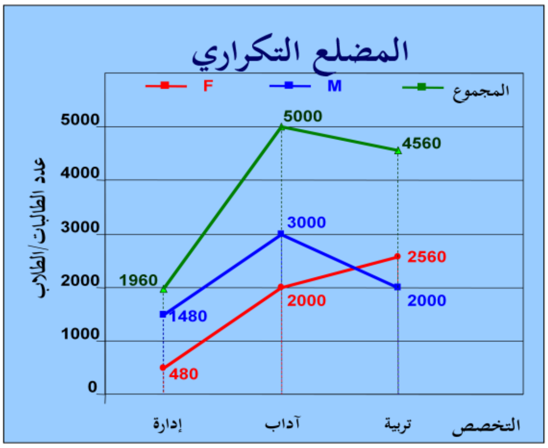
هـ – المنحنى أو الخط البياني:

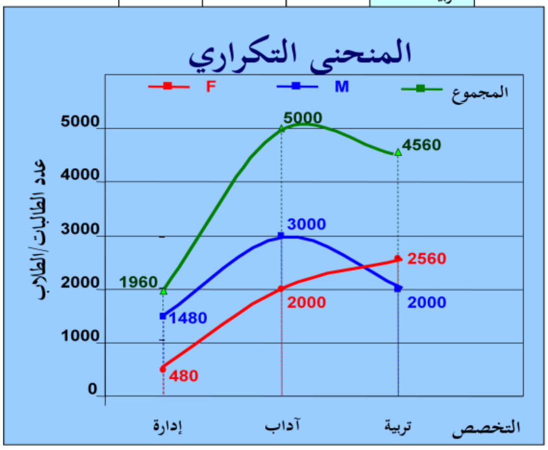
يستخدم المنحنى أو الخط البياني أساسا لتوضيح الاتجاه العام للظاهرة خلال فترة من الزمن، ويستخدم هذا النوع من الرسم البياني لتمثيل الظواهر ذات البيانات المتصلة (غالبا)، وكذلك ممكن استخدامه مع البيانات المنفصلة.

كما يمكن استخدام الخط أو المنحنى البياني لتمثيل أكثر من ظاهرة على نفس الرسم ومقارنتها ببعضها.

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

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





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

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

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

الحل تفصيلا في الكتاب

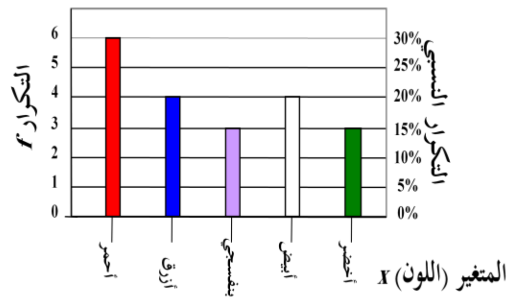
مثال: لوكانت لدينا البيانات التالية:

x	f	الزاوية
أحمر	6	108
أزرق	4	72
بنفسجي	3	54
أبيض	4	72
أخضر	3	54

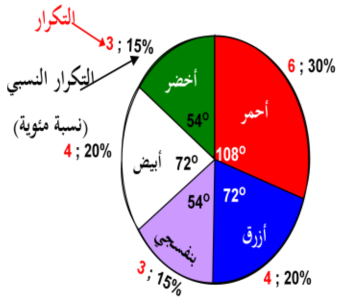
المطلوب: مثل البيانات السابقة بـ:

- الأعمدة البسيطة
- الخط البياني
- المنحنى البياني
- اللوحة الدائرية

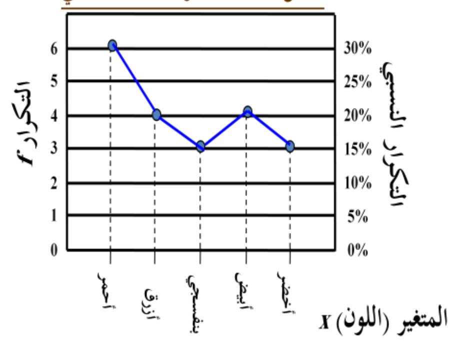
تمثيل البيانات بطريقة الأعمدة البسيطة



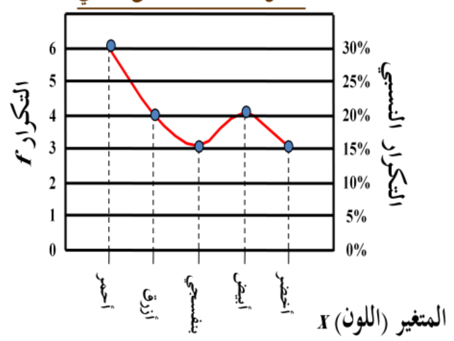
تمثيل البيانات بطريقة اللوحة الدائرية



تمثيل البيانات بطريقة الخط البياني



تمثيل البيانات المنحنى البياني



ملاحظات على المنحنى والخط البياني :

١. الرسم بالخط البياني أو المنحنى يتطلب جهدا 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 اذا كانت جيدة التصميم.

٢. توفر وقت المشاهدة اذ أن استنباط الحقائق من الرسوم البيانية أسرع من الوصول اليها بواسطة الأرقام الموضوعة في جداول.

٣. إمكانية معرفة الاتجاهات العامة للظواهر.

٤. سهولة فهم وتذكر العلاقات بين الظواهر محل الدراسة.

العيوب:

١. التضحية بدقة البيانات اذ 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 الى مقياس رسم كبير.

تمارين محلولة

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 :

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

(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

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

20	72°
?	36°

72 x ? = 36 x 20 , ∴ ? = 10

(١-ب) بنفس الأسلوب السابق

20	72°
30	?

72 x 30 = ? x 20 , ∴ ? = 108°

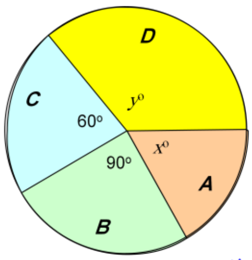
(١-ج) مجموع الزوايا المركزية يجب أن يكون 360°

∴ 72 + 36 + 108 + ? = 360 , ∴ ? = 144°

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

20	72°
	360°

$360 \times 20 = 72 \times \sum f$, $\therefore \sum f = 100$



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 التالية :

(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

(٢-أ)

100%	360°
?	90°

360 x ? = 90 x 100

? = 25%

(٢-ب)

$\frac{25}{100}$ x 5400 = 1350 $\rightarrow \sum f$

(٢-ج)

الزاوية المركزية المناظرة لمبيعات الشركتين معاً تساوي

360 – (90 + 60) = 210°

5400	360°
?	210°

360 x ? = 210 x 5400

? = 3150

المحاضرة السادسه

العرض البيانى للبيانات

ثانيا: البيانات المبوبة

يتم استخدام العديد من الاشكال للتعبير عن البيانات المبوبة فى صورة جداول توزيعات تكرارية وهى :

١. المدرج التكرارى

٢.المضلع التكرارى

٣.المنحنى التكرارى

٤.المنحنى التكرارى المتجمع الصاعد

٥.المنحنى التكرارى المتجمع الهابط (النازل)

المدرج التكرارى

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

يتم تقسيم المحور الرأسي (المحور الصادي) فى المدرج التكرارى حسب التكرار (فقد نستخدم التكرار الأصلى في حالة تمثيل التوزيع التكرارى، وكذلك يمكن أن نستخدم التكرار النسبى في حالة تمثيل التوزيع التكرارى النسبى).

ويتم تقسيم المحور الأفقى (المحور السينى) على أساس الفئات وهنا يظهر حالتين هما:

الحاله الأولى:- تساوى أطول الفئات

وفى هذه الحالة يكون ارتفاع المستطيل معبرا عن عدد مرات تكرار وجه الظاهرة محل الدراسة

الحاله الثانيه:- عدم تساوى أطوال الفئات

وفى هذه الحالة لابد من إجراء تعديل في التكرار الأصلى قبل رسم المدرج التكرارى، لذا فإننا نقوم بإيجاد التكرار المعدل والذى هو عبارة عن ناتج قسمه التكرار الأصلى لكل فئه على طول الفئة المقابلة

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 المدرج التكرارى.

الحل تفصيلا في الكتاب

بعض خصائص التوزيع التكرارى:

يمكن إستنتاج بعض خصائص التوزيع التكرارى من شكل المدرج التكرارى بدراسة الخصائص التاليه:

الخاصيه الأولى: التماثل

الخاصيه الثانية: الإلتواء

الخاصيه الثالثة: المنـوال

المضلع التكرارى

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

مثال: استخدم 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 التكرارى.

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 التكرارى.

الحل تفصيلا في الكتاب

التوزيعات التكرارية المتجمعة:

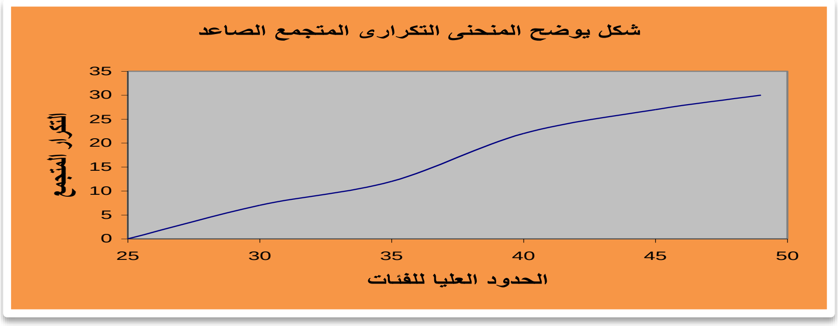
تستخدم المنحنيات المتجمعه لتمثيل التوزيعات التكراريه المتجمعه بيانياً بما يتلائم مع نوع التوزيع التكراري المتجمع، ونحصل على المنحنى المتجمع برصد التكرار المتجمع لأي فئة مقابل الحد الأعلى أو الحد الأدنى الفعلي لها ثم نوصل هذه النقاط فيما بينها بخطوط ممهدة.

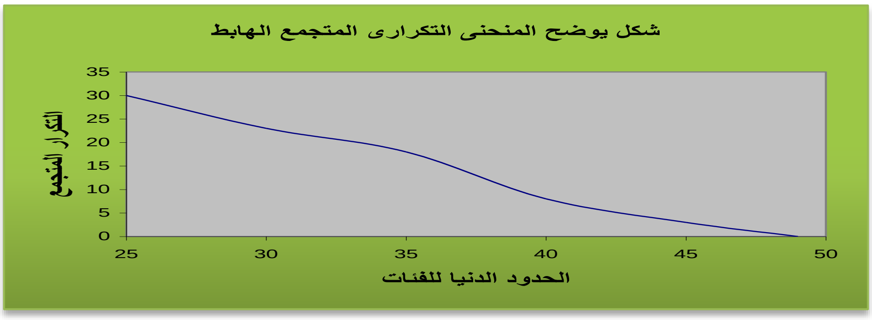
يستخدم المنحنى المتجمع الصاعد

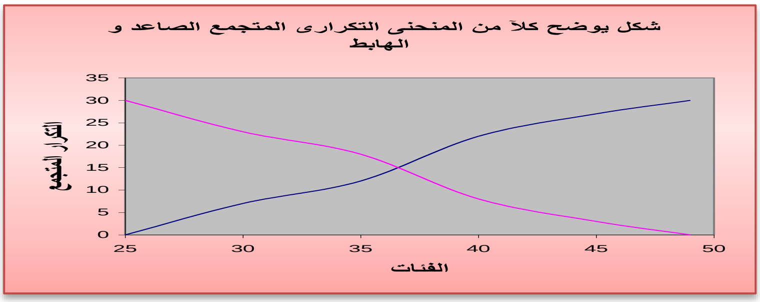
لتمثيل التوزيع التكرارى المتجمع الصاعد، سواء أكان بالقيم المطلقه للتكرارات، أو بالتكرار النسبي. ويراعي وضع النقاط الخاصه بالتكرارات في حالة المنحنى المتجمع الصاعد عند الحد الأعلى لكل فئه، لأنه يعبر عن العدد الاجمالى لأوجه الظاهرة الواقع أسفل الحد الأعلى للفئه.

ويستخدم المنحنى المتجمع الهابط (النازل)

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





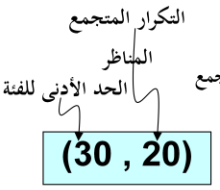


المنحنى المتجمع الصاعد

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 كالآتي :

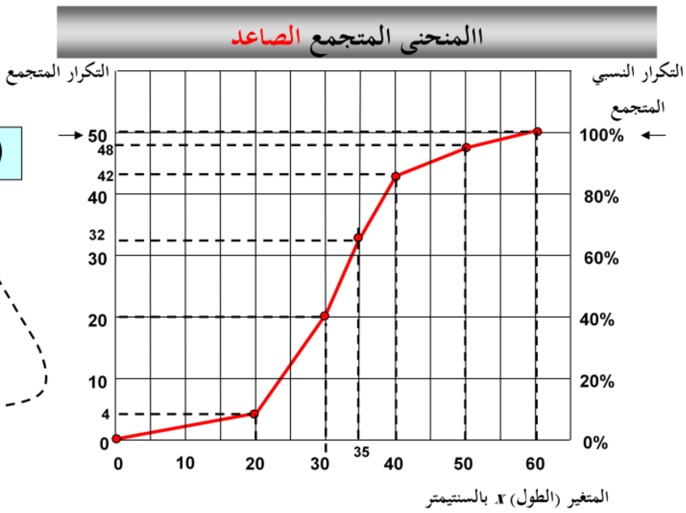
التوزيع التكراري الأصلي

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



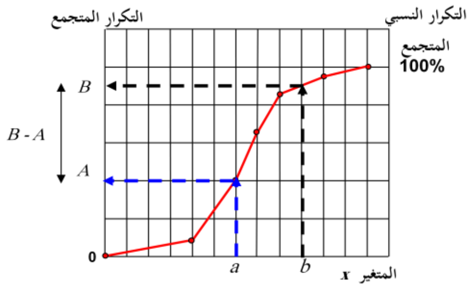
التوزيع التكراري المتجمع الصاعد

المتغير x	التكرار المتجمع	التكرار النسبي المتجمع	النقطة الموقعة 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"



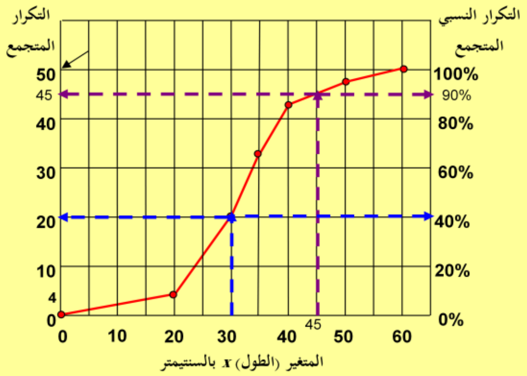
فلحساب قيمة التكرار المتجمع المناظر لـ "$a \leq x < b$"

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

فمثلاً في المثال التوضيحي السابق



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

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

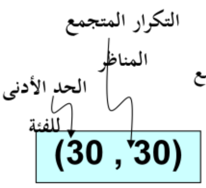
أو من الرسم : **90 – 40 = 50%**

المنحنى المتجمع النازل أو الهابط

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

التوزيع التكراري الأصلي

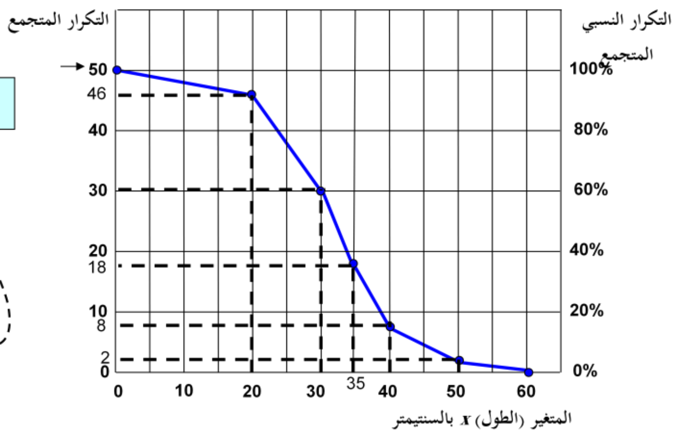
المتغير x	التكرار f
$0 \leq x < 20$	4
$20 \leq x < 30$	16
$30 \leq x < 35$	12
$35 \leq x < 40$	10
$40 \leq x < 50$	6
$50 \leq x < 60$	2



التوزيع التكراري المتجمع الهابط

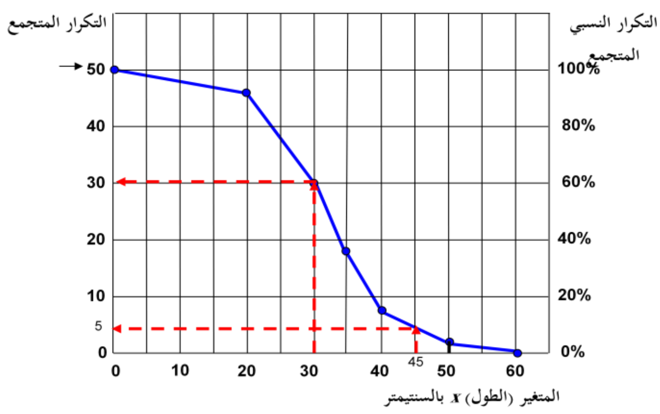
المتغير x	التكرار المتجمع	التكرار النسبي المتجمع	النقطة الموقعة على الرسم
≥ 0	50	100%	(0 , 50)
≥ 20	46	92%	(20 , 46)
≥ 30	30	60%	(30 , 30)
≥ 35	18	36%	(35 , 18)
≥ 40	8	16%	(40 , 8)
≥ 50	2	4%	(50 , 2)
≥ 60	0	0%	(60 , 0)

المنحنى المتجمع النازل أو الهابط



ويفيد المنحنى المتجمع النازل أو الهابط

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



أي أن المنحنيان التكراريان المتجمعان الصاعد والهابط يؤديان نفس الغرض تقريبا

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

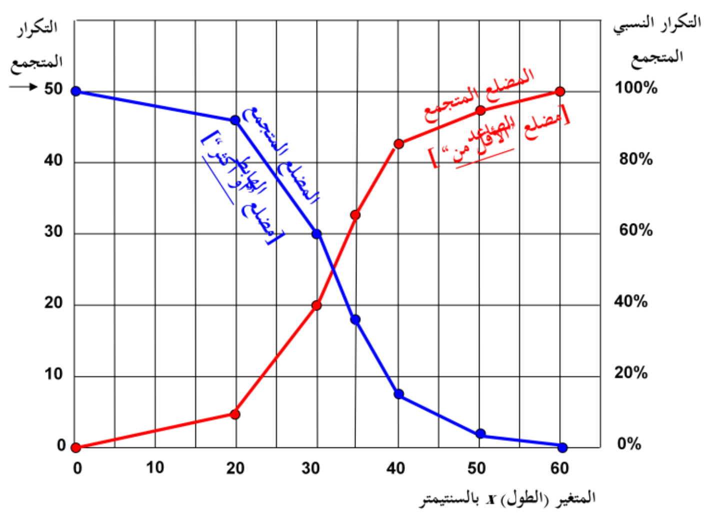
ويمكن رسم المضلعين التكراريين المتجمعين : الصاعد والهابط على رسمة واحدة كما هو مبين :

التوزيع التكراري المتجمع الصاعد

المتغير x	التكرار المتجمع	التكرار النسبي المتجمع	النقطة الموقعة 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)

التوزيع التكراري المتجمع الهابط

المتغير x	التكرار المتجمع	التكرار النسبي المتجمع	النقطة الموقعة 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)



الأشكال الشائعة للتوزيعات التكرارية

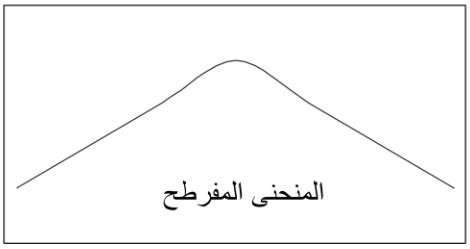
يعتبر التوزيع الطبيعى ذو شكل الجرس من التوزيعات التكراريه الهامه في دراستنا.

وفى أحيان أخرى يكون المنحنى التكرارى

مدبدب القمة بحيث تكون القمه ضيقة وذو طرفين واسعين نسبياً، فيسمى فى هذه الحالة منحنى قليل التفرطح أو المنحنى المدبدب.

وقد يكون المنحنى التكرارى

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



المحاضرة السابعة

المقاييس الإحصائية للبيانات غير المبوبة

أولا: مقاييس النزعة المركزية

فبعد جمع البيانات و المعلومات وعرضها يأتي بعد ذلك

تحليل البيانات Data Analysis

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

تساعدنا المقاييس الاحصائية

فى وصف المتغيرات المختلفة عن طريق معرفة القيم التى تتركز حولها البيانات ومدى التفاوت بين قيم المفردات محل الدراسة وتلك القيم.

كما تساعدنا فى المقارنة بين المتغيرات المختلفة من حيث مدى نزعتها نحو مراكز معينة وتحديد مدى تجانس البيانات بعضها مع بعض.

وتتمثل أهمية عملية وصف البيانات كميا

من خلال محاولة الوصول إلى فهم ورؤية أوضح للمعلومة المحتواة فى القيم الكمية للمتغيرات محل الدراسة،

تنقسم المقاييس الاحصائية إلى نوعين رئيسيين هما:

مقاييس النزعة المركزية Central Tendency Measures

مقاييس التشتت أو الأنتشار Dispersion Measures

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

اولا- مقاييس النزعة المركزية Central Tendency Measures

نقصد بمقاييس النزعة المركزية تلك القيم الوسطى التى توضح القيمة التي تجمع أكبر عدد من القيم الخاصة بمجموعة معينة عندها . ولتحديد 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 الحسابى (المتوسط) Mean

يُعرف المتوسط الحسابي بأنه قيمة التى 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:

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

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

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: حساب المتوسط الحسابى للمبيعات الشهرية.

الحل تفصيلا في الكتاب

ويجب ملاحظة عدة أمور فى الوسط الحسابى وهى:

١. انه لا يشترط أن يكون المتوسط الحسابي عددا صحيحا.

٢. إن 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 الحسابى هو تأثره بجميع العمليات الجبرية تجرى على البيانات من إضافة قيمة لجميع البيانات أو طرحها أو ضربها أو قسمتها.

مثال: بسؤال خمسة أشخاص عن أجرهم الشهرى فكانت إجاباتهم كما يلى بالألف ريال:
3 , 5 , 2, 7,3

المطلوب:

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

الحل تفصيلا في الكتاب

مزايا وعيوب 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 فردى أو زوجي كما يلي:

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

إيجاد قيمة الوسيط.

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

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

مزايا وعيوب الوسيط:

المزايا:

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 اكثر القيم شيوعا،

فى نفس المثال السابق للمبيعات الشهرية .

أحسب المنوال؟

نجد أن المبيعات الأكثر تكرارآ هنا هى ٣ ألف ريال لذلك

فان المنوال هنا = ٣

وقد يكون في التوزيع منوالين أو أكثر وذلك كالمثال الأتي:

٦ ، ٥ ، 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

مزايا وعيوب المنوال:

المزايا:

١.سهل الحساب سواء بالرسم أو بالحساب

٢.لا يتأثر كثيرا بالقيم الشاذة

٣.لا يتأثر كثيرا لو تغيرت قيم بعض مفردات البيانات

العيوب:

١.أقل مقاييس النزعة المركزية استعمالا

٢.عديم الفائدة في البيانات القليلة العدد

المحاضرة الثامنة

المقاييس الإحصائية للبيانات غير المبوبة

ثانيا: مقاييس التشتت أو الأنتشار

كما تميل القيم الى التمركز فانها تميل أيضا إلى التشتت أو الانتشار، فبالتالي فان أي توزيع من القيم له صفة التمركز، 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) : ١ ، ٢ ، ٣ ، ٥ ، ٦

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): (٤٥ ، ٥٠ ، ٥٥) المتوسط هنا = ٥٠

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

فلذا 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 مطلقا بما بينهما من قيم أخرى.

فالمدى لا يصلح الا اذا اراد الباحث أن يأخذ فكرة سريعة عن مدى تشتت بينات التوزيع موضع الدراسة، الا أن استخدامه والاعتماد عليه قد يؤديان الى نتائج خادعة، وخاصة اذا كان هناك انفصال بين الدرجات المتطرفة وباقي الدرجات موضع البحث.

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

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:
حساب المدى للمبيعات الشهرية.

الحل:

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

Range=12-3=9

عيوب المدى:

نجد أن من أهم عيوب المدى أنة يتم حسابة بناءا على أكبر و أصغر قيميتين وبالتالى فى حالة كونهما أو أحدهما متطرفتين أو قيم شاذة فأن 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 جميعا (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 الى تطبيق متوسط الانحرافات المطلقة AAD الذي يقيس تباعد كافة القيم عن متوسطها الحسابي.

ويمكن حساب متوسط الأنحرافات 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 بالألف ريال كما يلى:

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 متوسط الأنحرافات المطلقة للمبيعات الشهرية.

الحل تفصيلا في الكتاب

التباين والانحراف المعياري:

التباين Variance هو متوسط مربعات انحرافات القيم عن وسطها الحسابى. ويرمز له بالرمز σ^2 (تقراء سيجما تربيع) وذلك إذا كان محسوب لبيانات المجتمع أما فى حالة حسابة لبيانات عينة من المجتمع فيرمز له بالرمز s^2.

الانحراف المعياري Standard Deviation وهو الجذر التربيعى لمتوسط مربعات انحرافات القيم عن وسطها الحسابى أى هو جذر التباين لذلك يرمز له بالرمز σ (تقراء سيجما) و ذلك إذا كان محسوب لبيانات المجتمع أما فى حالة حسابة لبيانات عينة من المجتمع فيرمز له بالرمز s.

ويعتبر الانحراف المعياري والتباين من أهم مقاييس التشتت جميعا أو اكثرها استعمالا، وهما قريبين في خطوات ايجادهما من الانحراف عن المتوسط.

فالتباين واللإنحراف المعياري يختلف عن الإنحراف عن المتوسط في طريقة التخلص من اشارات الفروق بين القيم والمتوسط الحسابي، فبينما نتخلص من هذه الاشارات في طريقة الانحراف عن المتوسط بإهمال 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:

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

وبالتالى يكون حساب الإنحراف المعياري كما يلي:

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

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

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 التباين وقيمة الإنحراف المعياري للمبيعات الشهرية.

الحل تفصيلا في الكتاب

ملاحظة هامة:

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

فنلاحظ عدم تغير قيمة الانحراف المعيارى في حالة الجمع أو الطرح وإنما تظل قيمتة كما هى بالرغم من طرح مقدار ثابت من جميع قيم التوزيع.

أما في حالة الضرب أو القسمة فنلاحظ تغير قيمة الانحراف المعيارى وهى نفس قيمة الانحراف المعيارى القديمة مضروبة فى القيمة التي ضرب فيها أو قسم عليها.

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

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 المعيارى الجديد؟

الحل تفصيلا في الكتاب

نلاحظ عدم تغير قيمة الانحراف المعيارى وإنما ظلت قيمتة كما هى 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 يلى:

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 المعيارى للمبيعات الشهرية إذا تم زيادة المبيعات الشهرية إلى ثلاث أمثال الموجود حالياً ؟

الحل تفصيلا في الكتاب

نلاحظ تغير قيمة الانحراف المعيارى وهى نفس قيمة الانحراف المعيارى القديمة مضروبة فى ٣

وبالتالى يمكن أن نكون حصلنا عاى كافة المقاييس الإحصائية الوصفية التى تصف المبيعات الشهرية فكانت كما يلى:

المتوسط	الوسيط	المنوال	الوسط الهندسي
5.75	5	3	5.20114

المدى	متوسط الانحرافات المطلقة	التباين	الانحراف المعيارى
9	2.20833	7.840909	2.80016

المحاضرة التاسعة

المقاييس الإحصائية للبيانات المبوبة

أولا: الوسط الحسابي والتشتت حوله

يقصد بالبيانات المبوبة تلك البيانات التى تم وضعها فى صورة جداول تكرارية.

والجداول التكرارية للمتغير الكمى المتقطع يمكن تحويلها لتكون بيانات غير مبوبة و نتعامل معها كما سبق توضيح ذلك في المحاضرة السابقة، الا أن الأمر يختلف بالنسبة للمتغير الكمى المتصل حيث يصعب ذلك ولابد من التعامل معها كما هي على صورتها الجدولية وهذا ما سوف نتناوله فى هذه المحاضرة إن شاء الله

وسيتم عرض لكيفية حساب كلا من مقاييس النزعة المركزية ومقاييس التشتت فى ثلاث حالات للجداول التكرارية وهى

١.الجداول المنتظمة

٢.الجداول غير المنتظمة

٣.الجداول المفتوحة

الجداول المنتظمة:

وهى تلك الجداول التى تكون فيها أطوال الفئات جميعها متساوية .

اولا- الوسط الحسابى والتشتت حولة:

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

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

$\overline{x}$ الوسط الحسابى

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

f_i تكرار الفئة **i**

l عدد الفئات

ويتم حساب التشتت حول المتوسط الحسابى من خلال الآتي:

أ- متوسط الانحرافات المطلقة AAD:

وهو يقيس إنحراف القيم عن وسطها الحسابى بغض النظر عن اشارة ذلك الانحراف حيث يتم حسابه من خلال المعادلة التالية :

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

ب – التباين σ^2 :

وهو متوسط مجموع مربع إنحرافات القيم عن وسطها الحسابى. ويتم حسابه من خلال المعادلة التالية

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

جـ - الانحراف المعيارى σ :

هو الجذر التربيعى للتباين، ويمكن حسابه من خلال المعادلة التالية:

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

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

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

المطلوب: حساب التالى:

- الوسط الحسابى
- التباين
- الانحراف المعيارى
- متوسط الانحرافات المطلقة

الحل تفصيلا في الكتاب

الوسط الحسابى:

يمكن إيجاد الوسط الحسابى بتطبيق المعادلة الخاصة بذلك كما يلى:

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

التباين:

يمكن الحصول على التباين باستخدام المعادلة الخاصة بذلك كما يلى:

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

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

الانحراف المعيارى:

يمكن حساية باستخدام المعادلة الخاصة بذلك كما يلى :

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

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

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

المحاضرة العاشرة

المقاييس الإحصائية للبيانات المبوبة

ثانيا: الوسيط والتشتت حوله

الوسيط هو القيمة التى يصغرها عدد من القيم يتساوى مع العدد الذى يكبر هذه القيمة

ولحساب الوسيط من البيانات المبوبة هناك ثلاث خطوات يجب إتباعها وهى:

1. إيجاد الجدول التكرارى المتجمع الصاعد
2. 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 الوسيطية

k_{Med} ترتيب الوسيط

F_a التكرار المتجمع السابق للفئة الوسيطية

F_b التكرار المتجمع اللاحق للفئة الوسيطية

I طول الفئة الوسيطية

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 الوسيط؟

الحل تفصيلا في الكتاب

الرُبيع الادنى (الأول):

يُعبر الرُبيع الأول Q1 عن القيمة التى يكون قبلها عدد المشاهدات ربع العدد الكلى للمشاهدات والمشاهدات بعده تمثل ثلاث ارباع العدد الكلى للمشاهدات محل الدراسة.

لذلك يتم حسابة كما فى حالة الوسيط مع اختلاف أن ترتيب الرُبيع الاول Q1 هو (n / 4)

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

الرُبيع الاعلى (الثالث):

يُعبر الرُبيع الثالث Q3 عن القيمة التى يكون قبلها عدد المشاهدات ثلاث ارباع العدد الكلى للمشاهدات والمشاهدات بعده تمثل ربع العدد الكلى للمشاهدات محل الدراسة.

لذلك يتم حسابه كما فى حالة الوسيط مع اختلاف أن ترتيب الرُبيع 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 الرُبيع الادنى (الاول) Q1 و الرُبيع الاعلى (الثالث) Q3 بنفس خطوات حساب الوسيط الا أن الامر المختلف هنا هو الترتيب حيث يكون كالتالي:

	Q1	Q3
الترتيب		

$k_{Q1} = n/4$ $\qquad$ $k_{Q3} = 3n/4$

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 كل من:

- قيمة الرُبيع الاول
- قيمة الرُبيع الثالث

الحل تفصيلا في الكتاب

حساب قيمة العُشير $P_{0.10}$:

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

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

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 العُشير؟

الحل تفصيلا في الكتاب

حساب قيمة المئين $P_{0.01}$:

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

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

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 المئين؟

الحل تفصيلا في الكتاب

وعلى ذلك نكون قد حصلنا على مقاييس النزعة المركزية التى تصف تركز البيانات عند أي نسبة من مفردات البيانات محل الدراسة فى حالة البيانات المبوبة والتى كانت كما يلي

المقياس	$P0.01$	$P_{0.10}$	Q1	Med	Q3
القيمة	٢١,١	٣٠,٣٣٣	٣٥,٨٣٣٣	٤٣	٤٨,٥

نصف المدى الرُبيعى Inter Quartile Range

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

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

مثال: 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 نصف المدى الرُبيعي؟

الحل تفصيلا في الكتاب

ثالثا: المنوال

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

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

Mod قيمة المنوال

L_{Mod} الحد الأدنى لفئة المنوال

$D1$ يساوى تكرار فئة المنوال – تكرار الفئة السابقة

$D2$ يساوى تكرار فئة المنوال – تكرار الفئة اللاحقة

I طول الفئة المنوالية

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 في الكتاب

الجداول غير المنتظمة:

وهي الجدوال التى يكون فيها أطوال الفئات غير متساويه ويكفى وجود فئة واحدة فقط طولها غير متساوى مع باقى الفئات لجعل الجدول غير منتظم.

ويتم حساب المقاييس الإحصائية التى 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 في أعمدة فئات التوزيع ولذلك يتم التخلص من تأثير طول الفئة بإيجاد التكرار المعدل، ويتم ذلك من خلال المعادلة التالية:

التكرار المعدل = التكرار الأصلى للفئة ÷ 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 الموظفيين	٢٠	٥٠	١٥	١٥

المطلوب حساب:

١ – الوسط الحسابى
٢ – متوسط الانحرافات المطلقة
٣ – التباين
٤ – الانحراف المعيارى
٥ – الوسيط
٦ – الرُبيع الأول
٧ – الرُبيع الثالث
٨ – العُشير
٩ – المئويين
١٠ – نصف المدى الرُبيعي
١١ – المنوال

الحل تفصيلا في 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 الذى طريقته تحتاج إلى تعديل فى الحساب في حالة الجداول غير المنتظمة.

الجداول المفتوحة:

وهى ذلك النوع من الجداول التى يكون فيها الحد الادنى للفئة الاولى غير محدد أو الحد الأعلى للفئة الأخيره 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 فيعيبر من أنسب المقاييس الإحصائية فى تلك الحالة هى المقاييس الوسيطية والتى يقصد بها الوسيط والرُبيع الادنى والرُبيع الاعلى والعُشير والمئويين وكذلك لقياس التشتت يتم من خلال نصف المدى الربييعى.

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 النسبى والدرجة المعيارية

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

حيث سيتم فى هذه المحاضرة استعراض كلا من:

مقاييس التشتت النسبى

القيمة المعيارية

اولا – مقاييس التشتت النسبى Coefficient of Variation

يستخدم هذا النوع من المقاييس لمقارنة تشتت مجموعتين من البيانات او ظاهرتين او توزيعين حيث يتم الاعتماد فى عملية المقارنة على مقاييس التشتت النسبى Coefficient of variations (c.v.) والتى يعبر عنها من خلال معامل الاختلاف المعيارى والذي يحسب من خلال المعادلات التالية:

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

أو

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

معادلة حساب الربيع الأول Q_1

الربيع الأول Q_1:

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

Q_1 قيمة الربيع الأدنى أو الأول

L_{Q_1} الحد الادنى لبداية الفئة الربيعية الأولى

k_{Q_1} ترتيب الربيع الأول

F_a التكرار المتجمع السابق للفئة الربيعية الأولى

F_b التكرار المتجمع اللاحق للفئة الربيعية الأولى

I_{Q_1} طول الفئة الربيعية الأولى

معادلة حساب الربيع الثالث Q_3

الربيع الثالث Q_3:

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

Q_3 قيمة الربيع الأدنى أو الثالث

L_{Q_3} الحد الادنى 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 السنوى بأحد الاحياء:

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

المطلوب: حساب :

- معامل الاختلاف للإيجار السنوى
- معامل الاختلاف الربيعى للإيجار السنوى

الحل تفصيلا فى 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 أنة يفضل استخدام المعادلة الثانية فى حالة االجداول 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:

$$z = \frac{x - \overline{x}}{S}$$

مثال: حصل أحد الطلاب فى مقرر المحاسبة على (٨٠) درجة حيث بلغ متوسط درجات الطلاب فى أختبار المحاسبة (٨٣) درجة بإنحراف معيارى (٥). بينما حصل فى اختبار مقرر الرياضيات على (٧٠) درجة حيث بلغ متوسط درجة الطلاب فى اختبار الرياضيات (٦٥) درجة بأنحراف معيارى قدرة (٥) درجات .

المطلوب:

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

الحل تفصيلا في الكتاب

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

المحاضرة الثانية عشر

مقاييس الإلتواء والتفلطح

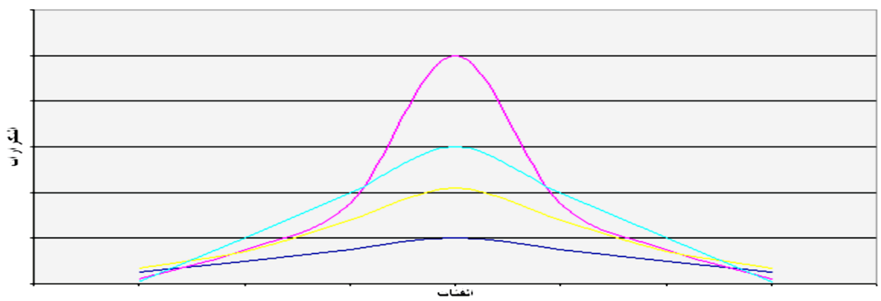
أولا: مقاييس الإلتواء Skewdness Measures

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

المنحنى المتماثل Symmetrical Curve

هو المنحنى الذى اذا قسمناه إلى نصفين إنطبق هذان النصفان على بعضهما البعض تماما

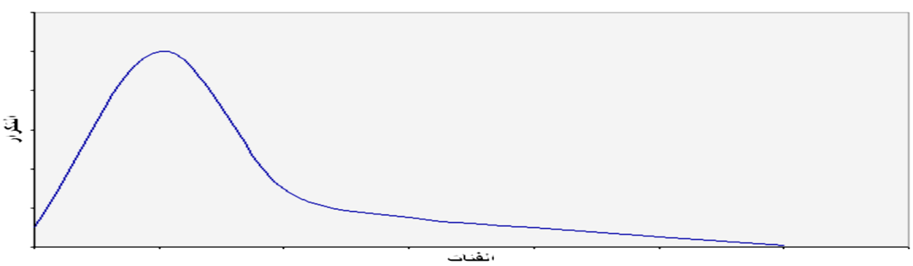
شكل يوضح منحنيات التوزيع المتماثل



المنحنيات الملتوية Skewed

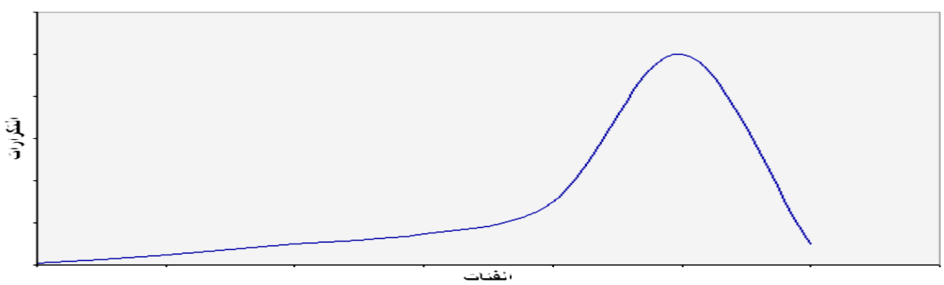
إن الكثير من التوزيعات الإحصائية تبتعد عن التماثل بتركز تكراراتها إما عند أصغر القيم فيصبح المنحنى ملتويا جهة اليمين أو إلتواء موجب كما يظهر فى الشكل التالي:

شكل يوضح منحنى ملتوى جهة اليمين



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

شكل يوضح منحنى ملتوى جهة اليسار



ويمكن قياس الإلتواء من خلال معامل **الإلتواء SK** والذى يفيدنا فى الحكم على مدى تماثل أو إلتواء التوزيع

تتعدد مقاييس الإلتواء إلا أن من أهمها:

معامل الإلتواء لبيرسون والذى يكون فى أحد الصورتين التاليتين:

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

أو

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

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

لذلك يمكن الاعتماد على مقياس الإلتواء لباولى SkB الذى يعرف كما يلى:

$$SK_B = \frac{Q_3 - 2Med + Q1}{Q_3 - Q1}$$

معادلة حساب الوسيط Med

الوسيط Med:

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

Med قيمة الوسيط

L_{Med} الحد الادنى لبداية الفئة الوسيطية

k_{Med} ترتيب الوسيط

F_a التكرار المتجمع السابق للفئة الوسيطية

F_b التكرار المتجمع اللاحق للفئة الوسيطية

I_{Med} طول الفئة الوسيطية

معادلة حساب الربيع الأول Q1

الربيع الأول Q1:

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

Q_1 قيمة الربيع الأدنى أو الأول

L_{Q_1} الحد الادنى لبداية الفئة الربيعية الأولى

k_{Q_1} ترتيب الربيع الأول

F_a التكرار المتجمع السابق للفئة الربيعية الأولى

F_b التكرار المتجمع اللاحق للفئة الربيعية الأولى

I_{Q_1} طول الفئة الربيعية الأولى

معادلة حساب الربيع الثالث Q3

الربيع الثالث Q3:

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

Q_3 قيمة الربيع الأدنى أو الثالث

L_{Q_3} الحد الادنى 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 السنوى بأحد الاحياء في أحد المدن:

الايجار بالألف ريال	٦–	١٠ –	١٢–	١٤ – ١٨
عدد الوحدات السكنية	١٥	٢٠	١٢	١٣

المطلوب:

حساب معامل الإلتواء لتوزيع الإيجار السنوى للوحدات السكنية.

الحل تفصيلا في الكتاب

ويظهر لنا من النتيجة لجميع المعادلات الخاصة بحساب معامل الإلتواء وجود التواء موجب جهة اليمين الا أن قيمة معامل الإلتواء صغيرة تقترب من الصفر مما يدل ايضا على أن التوزيع قريب من التماثل.

ونتيجة لوجود اختلاف فى الاصل الرياضى لكل من المعادلات الثلاث السابقة لذا نجد أن قيمة معامل الإلتواء تختلف. إلا أنة كما سبق وذكرنا بأنه يفضل استخدام معامل الإلتواء لبيرسون فى أي من صيغتيه فى حالة البيانات غير المبوبة وكذلك الجداول التكرارية المغلقة أما في حالة الجداول التكرارية المفتوحة فيفضل استخدام معامل الإلتواء لباولي.

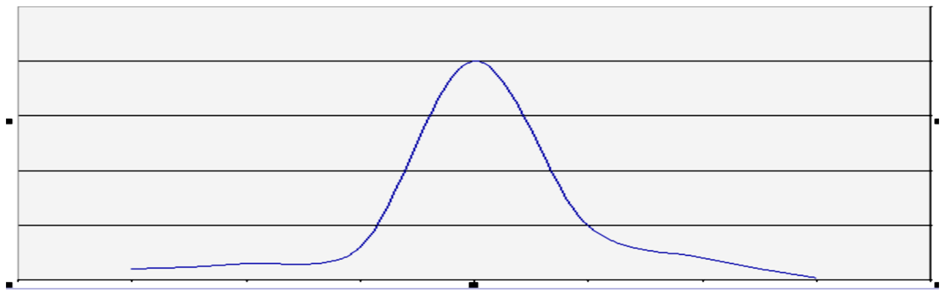
ثانيا: التفلطح Kurtosis

يقصد بالتفلطح مقدار التدبب (الارتفاع أوالإنخفاض) فى قمة المنحنى مقارنة بقمة منحنى التوزيع الطبيعى.

وتكون قيمة معامل التفلطح صفر فى حالة التوزيع الطبيعى المعيارى.

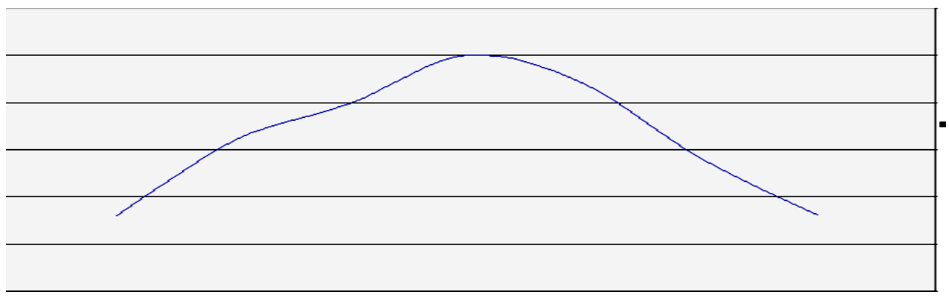
ففى حالة ما يكون معامل التفلطح للبيانات الاصلية أكبر من ٣ يكون المنحنى مدبب لأعلى كما بالشكل التالي:

شكل يوضح المنحنى المدبب



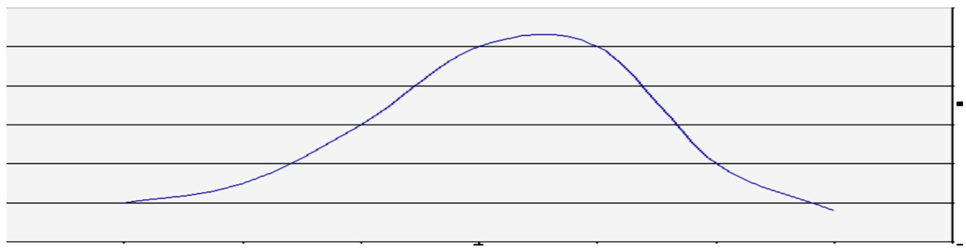
أما فى حالة ما يكون معامل التفلطح للبيانات الأصلية أقل من ٣ يعنى ذلك أن المنحنى مفلطح كما يتضح من الشكل التالي:

شكل يوضح المنحنى المفلطح



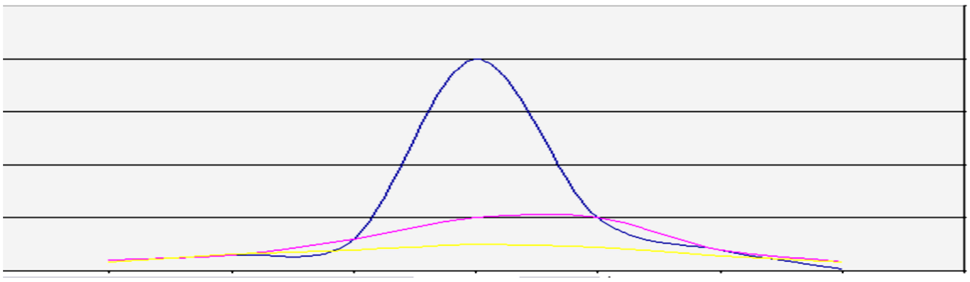
أما فى حالة ما يكون معامل التفلطح يساوى ثلاثة يكون المنحنى متوسط التفلطح و يكون بالشكل التالي:

شكل يوضح المنحنى متوسط التفلطح



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

شكل يوضح المنحنيات الثلاث معاً المدبب و متوسط التفلطح و المفلطح



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:

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

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

الايجار 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 تركز بدرجة كبيرة فى أحد الفئات على حساب باقى الفئات الأخرى.

المحاضرة الثالثة عشر

تحليل الإرتباط – 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، وتتراوح 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 درجـة 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:

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

إن معامل الارتباط التام الموجب (+١) يعنى التغير في اتجاه واحد في كلا الظاهرتين مع بقاء الأوضاع النسبية لوحدات 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 :

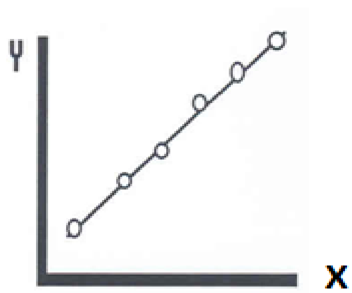
هناك وسيلة مبدئية يعرف الباحث من خلالها نوع الارتباط بين المتغيرين وما إذا كان الارتباط قوياً وضعيفاً أو منعدماً، وما إذا كانت العلاقة خطية أو غير خطية، موجبة أو سالبة. هذه الوسيلة هي " شكل الانتشار " والتي تصلح إذا كان المتغيران كميين. وجدير بالذكر أن هذه وسيلة مبدئية تساعد فقط في معرفة نوع الارتباط 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 النقط 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 :

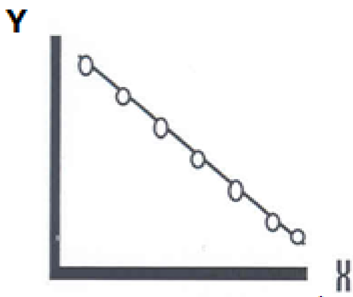
الشكل الأول :

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



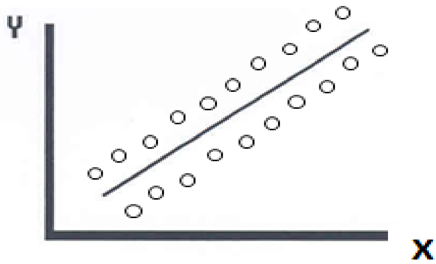
الشكل الأول (أ) ارتباط طردي تام (موجب)



الشكل الأول (ب) ارتباط عكسي تام (سالب)

الشكل الثاني :

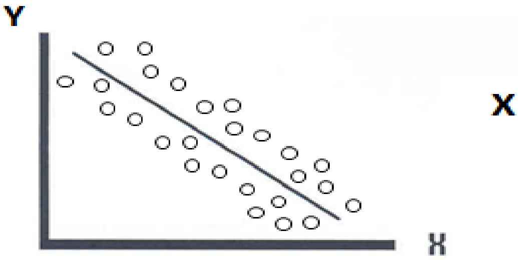
أما إذا كانت النقاط تأخذ شكل خط مستقيم ولكن لا تقع جميعها على الخط قيل أن العلاقة خطية (موجبة أو سالبة) كما في الشكل الثاني أ، ب.



الشكل الثاني (أ)

ارتباط موجب قوي

(ارتباط خطي طردي)



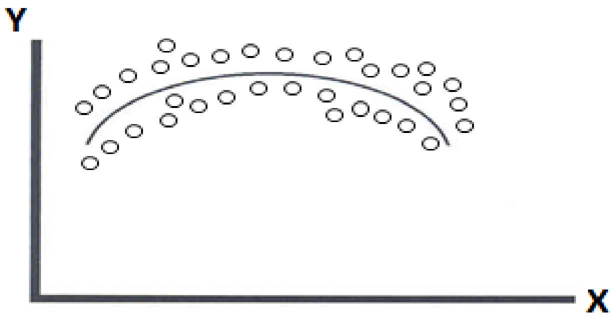
الشكل الثاني (ب)

ارتباط سالب قوي

(ارتباط خطي عكسي)

الشكل الثالث :

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

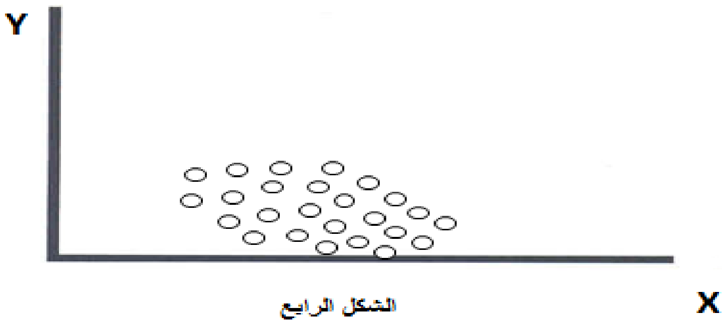


الشكل الثالث

(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)

ثانيا: معامل الارتباط Correlation Coefficient :

يقاس الارتباط بين متغرين 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.

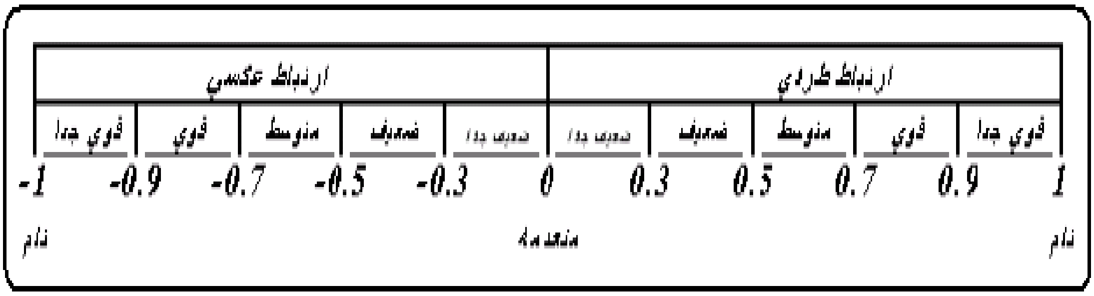
وإذا كانت قيمة معامل الارتباط تساوي – ١ فمعنى ذلك أن الارتباط بين المتغيرين عكسي تام، وهو أقوى أنواع الارتباط العكسي بين متغيرين.

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 منعدماً 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:



معامل بيرسون للارتباط الخطى البسيط

Person's Correlation Coefficient

يعتبر معامل الارتباط الخطى البسيط لبيرسون Person's Correlation Coefficient والذى سنرمز له بالرمز r_p من أكثر الأدوات الإحصائية استخداما فى تحديد قوة العلاقة بين متغيرين كما يستعمل لتحديد مدى وجود علاقة خطية بين متغيرين.

وهناك أكثر من صيغة يمكن الاعتماد عليها فى حساب معامل الارتباط الخطى البسيط لبيرسون منها:

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

وكذلك المعادلة الرياضية التالية والتي تعتبر اسهل وابسط:

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

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

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

والاتباط غالبا قيمته كسر أي اقل من الواحد الصحيح

ولتحديد نوع العلاقة نعتمد على اشارة معامل الارتباط فإذا كانت الإشارة:
- موجبة فإن العلاقة تكون طرديه
- سالبة فإن العلاقة تكون عكسية

ولتحديد قوة العلاقة نعتمد 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 خطيه او ارتباط بينهما أى يكون المتغيرين مستقلين عن بعضهما البعض وتكون العلاقة منعدمة

فمثلا إذا كانت قيمة معامل الارتباط r_p كالتالي فإن تفسيرة يكون:

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

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

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 إجراءها على بيانات اى من المتغيرين أو أحدهما من جمع أو طرح أو ضرب أو قسمة

مثال: في 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:

نجد 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 الإرتباط - ٢

سيتم في هذه المحاضرة استعراض المواضيع التالية:

أولا: معامل إرتباط سبيرمان

ثانيا: معامل الإقتران

ثالثا: معامل التوافق

معامل أرتباط الرتب لسبيرمان Spearman's Rank Correlation Coefficient r_s

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

أما فى حالة المتغيرات الوصفية فنستخدم معامل ارتباط الرتب لـ سبيرمان، والذى يتم أستخدامة فى قياس الارتباط خاصة فى حالة البيانات الوصفية الترتيبية مثل تقديرات الطلاب (ممتاز – جيدجدآ – جيد – مقبول – ضعيف) وكذلك قوة المركز المالى (جيد - متوسط - ضعيف) ودرجة الموافقة على الرأى فى اسئلة الاستبانة (موافق تماماً – موافق – محايد – غير موافق – غير موافق على الاطلاق).

ويتم حساب معامل الارتباط الرتب لسبيرمان r_s بأستخدام 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 تصاعديا أو تنازليا ولكن أهم شىء هو اذا كان ترتيب x

٢. تصاعدى لابد ان يكون ترتيب y تصاعدى ايضا والعكس صحيح.فى حالة الترتيب التصاعدى مثلا يتم اعطاء أقل قيمة الرتبة ١ والقيمة التى هى أكبر منها الرتبة ٢ و هكذا

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

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

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

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

الحل تفصيلا في الكتاب

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

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

المطلوب:
أحسب معامل الارتباط المناسب.

الحل تفصيلا في الكتاب

معامل الإقتران Conjunction Coefficient

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

النوع (ذكر – انثى)، والحالة التعليمية (متعلم - غير متعلم)

وعلى ذلك إذا كان لدينا متغيران لدي كلاً منهما زوج من الصفات فيكون جدول تكرارات الصفات المشتركة بينهما على الصورة التالية:

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

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

ويمكن حساب معامل الاقتران فى هذه الحالة كما يلي:

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

مثال: فى دراسة اجريت لمعرفة هل هناك علاقة بين العمل والتعليم تم سؤال ٢٠٠ شخص سؤالين هما:
هل انت متعلم ؟ نعم لا
هل انت ملتحق بأى عمل ؟ نعم لا
وبتجميع الاجابات تم عمل جدول الاقتران التالي:

العمل \ التعليم	متعلم	أمى
يعمل	113	23
لايعمل	49	15

المطلوب: أحسب معامل الاقتران ؟

معامل التوافق Concordance Coefficient

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

وحتى يمكن حسابه يتم إعداد الجدول المزدوج بين صفات المتغيرين ومنه يتضح لنا التكرارات المشتركة بين الصفات التى نعتمد عليها فى حساب مقدار يطلق علية " M "

ويتم حساب معامل التوافق من خلال المعادلة التالية:

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

حيث أن:

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

أى يتم إيجاد: مربع تكرار كل خلية مشتركة
مجموع الصف × مجموع العمود
ثم نجمعهم كلهم

وعلى ذلك يتم حساب معامل التوافق كما يلى:

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

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

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

الحل تفصيلا في الكتاب

وفي النهاية اتمنى للجميع التوفيق والنجاح

اخوكم أحمد