بسم الله الرحمن الرحيم

ملزمة مبادئ الإحصاء ..

أ . د / عبدالله عمر النجار ..

المحتوى إضافة لشرح الدكتور ..

شرح وتنسيق .. oOo al maznei oOo .. حلم المشاعر ..

كلية الآداب – المستوى الثاني

المحـاضـرة الأولـــى ..

التعريف ببعض المفاهيم الإحصائية ..

المقدمة :

الغرض من العلم (بوجهٍ عام) هو البحث عن الحقيقة ، والبحث العلمي هو الوسيلة للوصول إلى حقائق الأشياء والظواهر ومعرفة كلٍ 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 جدوليه أو بيانية ، ثم يتم استخدامها في حساب بعض المقاييس الخاصة بهذه الظواهر وإجراء تحليل لتلك البيانات بما يساعد في تفسير النتائج المختلفة للبيانات واستخدامها في استنتاج نظرية أو قاعدة أو قانون أو المساعدة في اتخاذ القرارات أو التنبوء بنتائج مستقبلية ....

والشكل التالي يمكن أن يوضح الإطار العام لأي بحث علمي :

مفهوم علم الإحصاء : هو ذلك العلم الذي يختص بالطرق العلمية والعملية لجمع وتنظيم وعرض وتحليل البيانات وكذلك الوصول إلى نتائج مقبولة وقرارات سليمة على ضوء هذا التحليل .

وقديماً عُرِف 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 . لا يتعمق في استخدام الإحصاء

الإحصاء الاستقرائي :

أو الاستدلال الإحصائي وهو يبحث في استقراء النتائج واتخاذ القرارات . أكثر عمقاً من الوصفي

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 ،، إذن المتغير هو أي صفه أو ظاهرة تختلف من شخص لآخر ومن وقت لآخر ويعمد الباحث لدراستها وقياسها .. وعادةً نرمز له برمز مثل .. x , y , A , B ، فمثلاً :

مثال	العملية الإحصائية : دراسة	البيانات (المشاهدات – القياسات)	المتغير X
(١)	لون العين لبعض الأطفال حديثي الولادة	أخضر – أزرق – بني -	لون العين
(٢)	عدد الطلاب في فصول المدرسة	١٥ – ١٨ – ٢٠ – ٢٥ – ١٧ -	عدد الطلاب
(٣)	أطوال مجموعة من الطلاب في فصل ما (بالمتر)	١.٥ – ١.٥٢ – ١.٧١ – ١.٨٣ -	طول الطلاب
(٤)	أوزان بعض العاملات بمصنع معين (بالكيلوجرام)	٥٥.٢ – ٦٠.١ – ٦٣.٣٥ – ٧٠.٥٢ -	وزن العاملة
(٥)	تقديرات عدد من الطلاب في مقرر الإحصاء	B – C – A – F – D – C – B – A -	تقدير الطالب

والمتغير (أي الظاهر تحت الدراسة) إما أن تكون :

تُسمى البيانات عندئذٍ ببيانات نوعية

متغير نوعي

هو أي صفه أو ظاهرة تتغير من شخص لآخر ومن وقت لآخر وتسجل بأوصاف لفضيه . أي لا يمكن التعبير عنه بعدد [مثل لون العين أو تقدير الطلاب في الأمثلة (١) ، (٥) السابقة]

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

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	١.٥ – ١.٥٢ – ١.٧١ – ١.٨٣ - ...
(٤) أوزان العاملات	٥٥.٢ – ٦٠.١ – ٦٣.٢٥ - ......

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

(٢) عدد الطلاب	١٥ – ١٨ – ٢٠ – ٢٥ - ......

أو

تُسمى البيانات عندئذٍ ببيانات كمية

متغير كمي

هو أي صفه أو ظاهرة تتغير من شخص لآخر ومن وقت لآخر وتسجل بأرقام عدديه . أي يمكن التعبير عنه بعدد مثل الأطوال أو الأوزان أو أعداد الطلاب .

متغير متصل (قابل للكسور) مثل طول الطلاب ← تُسمى البيانات عندئذٍ ببيانات (كمية) متصلة

أو

متغير متقطع (منفصلة) (غير قابل للكسور) مثل عدد الطلاب ← تُسمى البيانات عندئذٍ ببيانات (كمية) متقطعة

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 . يتم الأستفادة منها بطرق علمية .

(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 وحلول :

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

الإجــابـــات :

(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

المحـــــاضرة الثـــانيـــة ...

أساليب إجراء البحث الميداني ...

هناك سؤال مهم لابد 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 الحصر الشامل

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

أسلوب العينة

أسلوب 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 .

أسلوب العينات :

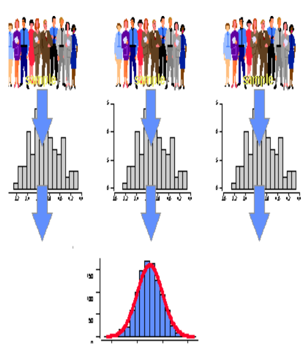
يبدوا هذا الأسلوب على العكس من أسلوب الحصر الشامل حيث تقتصر الدراسة فيه على جزء من المجتمع الإحصائي، لذا فهذا الأسلوب يوفر الوقت و الجهد و التكاليف ويصلح للمجتمعات غير المحدودة. إلا أن أهم عيوب هذا النوع هو ما يسمى بخطأ التحيز Sampling Bias .

اقسام مجتمع البحث :

قسم بعض العلماء مجتمع البحث الى قسمين :

- المجتمع الكلي للبحث . هو الإطار الكبير(البيانات) لهذا المجتمع.

- المجتمع الذي يمكن التعرف عليه . هو الإطار الداخلي أو القوائم التي تحوي من يشملهم المجتمع .



مجتمع البحث هو مصطلح علمي يراد به كل من يمكن أن تعمم عليه نتائج البحث

عينة البحث بأنها جزء من المجتمع اختير بطريقة علمية

طرق اختيار العينات

المجتمع الأصلي

مجتمع معروف (الطريقة الاحتماليه) هوالمجتمع الذي نستطيع أن نحيط بجميع أفراده

المجتمع متجانس: أهتمامات أو دراسات الباحث لاتتركز على صفات معينه بل تشمل المجتمع ككل .

العينة العشوائية: أن تكون الفرصه واحده لجميع أفراد المجتمع(أن يكونوا ضمن العينة) يتم أختيارها بطريقة بسيطة(إذا كان المجتمع محدود أو قليل) أو من خلال أستخدام الجداول العشوائية(في الأعداد الكبيرة مجموعة من الأرقام مرتبه بطريقة غير منظمة) .

العينة المنظمة: أن يقوم الباحث بإختيار تنظيم معين يقوم من خلاله بإختيار افراد العينة وفقاً لهذا التنظيم .

العينة العنقودية: يستخدم في حالة المجتمعات الكبيرة جداً . العنقود وحدة التعامل مع المجموعة وليس المفردة .

المجتمع غير متجانس: الباحث لديه هدف وهو النظر للمجتمع وفقاً لتقسيماته الداخلية .

العينة الطبقية: يتم فيها تحديد نسبة معينة لكل طبقة من طبقات المجتمع .

مجتمع غير معروف (الطريقة غير الاحتماليه) هو المجتمع الذي لا نستطيع أن نحيط بجميع أفراده

المجتمع متجانس

عينة الصدفة: أختيار الأفراد بناءً على المصادفة، أي أنه لا يوجد هناك ترتيبات مسبقة لعملية الأختيار .

العينية العمدية: متعمد الذهاب إلى مكان محدد لتطبيق الدراسة عليهم .

المجتمع غير متجانس

العينة الحصية : نفس فكرة العينة الطبقية، سميت حصية تميزاً لها في المجتمع غير المعروف .

ملاحظه ذكرها الدكتور :

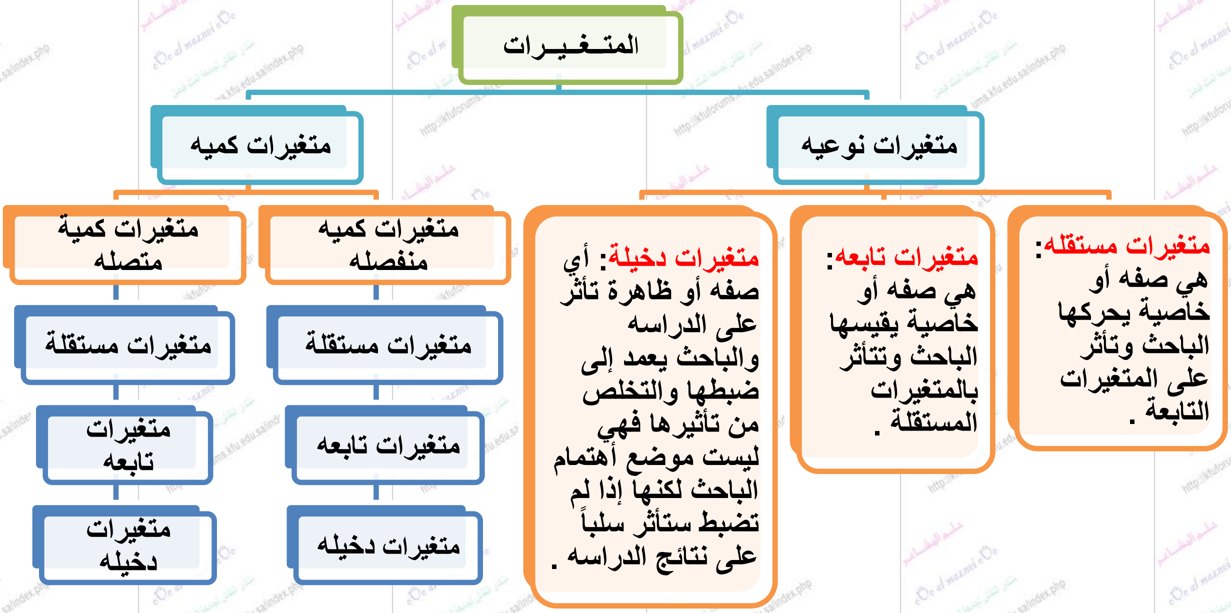
الأختيار الأمثل يترتب عليه جميع الخصائص في المجتمع .

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 إلى ثابت مثل درجة الحرارة في الغرفة. والباحث يسعى إلى 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 . تسجيلها بطريقة علمية ومنظمة بحيث يسهل التعامل معها أثناء التحليل.**
- **ترميز البيانات . نحول الألفاظ والمصطلحات إلى أرقام أو حروف حتى يسهل التعامل معها.**

١- الترميز الرقمي أو العددي . أفضل أنواع الترميز >> يسهل على الباحث في أثناء التعامل معه عملية التحليل الإحصائي .

٢- الترميز الأبجدي أو الحرفي .

٣- الترميز الأبجدي الرقمي .

- تصنيف البيانات . إلى مجاميع بالنسبة للباحث .
- مراجعة وتنقية البيانات . من الشوائب والأخطاء .

- ترميز بيانات الإستبانة وجعلها متاحة لبرنامج الـ SPSS : (SPSS تعني حزمة البرامج الإحصائية للعلوم الإجتماعية أو الإنسانية)

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 الأسئلة 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 الأسئلة والفقرات، وهذه الأنواع هي:

١- سؤال يسمح باختيار إجابة واحدة فقط .. وهو ذلك النوع من الأسئلة التي تلزم المستجيب بإختيار إجابه واحده فقط ويتم التعامل مع هذا النوع من الأسئلة من خلال تمثيلة بمتغير واحد يحوي جميع الإجابات الممكنه .. مثال ،،

- عدد سنوات الخبرة في العمل الأكاديمية :

١. (   ) أقل من سنه .
٢. (   ) من ١ – ٥ سنوات .
٣. (   ) من ٦ – ١٠ سنوات .
٤. (   ) من ١١ – ١٥ سنوات .
٥. (   ) أكثر من ١٦ سنه .

٢- سؤال يسمح باختيار أكثر من إجابة واحدة .. وهو ذلك النوع من الأسئلة التي تتاح من خلالها الفرصه للمستجيب لاختيار أكثر من إجابه ويتم التعامل مع هذا النوع من الأسئلة والفقرات من خلال تمثيلة بعدد من المتغيرات يماثل عدد الإجابات أو الاحتمالات المتاحة للسؤال أو الفقرة .. مثال ،،

- ما أهم المعوقات التي تحول دون استخدامك للإنترنت في البحث العلمي :
يمكن أختيار أكثر من عائق

١. ( ) عدم الإهتمام بالإنترنت .
٢. ( ) عدم وجود الوقت الكافي .
٣. ( ) عدم توفر أجهزة الحاسب الآلي .
٤. ( ) عدم توفر المتصفح المناسب للإنترنت .
٥. ( ) الإهتمام بحقوق النشر .
٦. ( ) الخوف من العولمة .

٣- سؤال مفتوح جزئيا .. وهوذلك النوع من الأسئلة التي تسمح للمستجيب باختيار إجابه موجوده ضمن الخيارات أو كتابه إجابه أخرى غير موجودة ضمن الخيارات المتاحه في السؤال .. مثال ،،

- الدرجه العلمية التي تحملها :

١. ( ) دكتوراه .
٢. ( ) ماجستير .
٣. ( ) بكالوريوس .
٤. ( ) غير ذلك ,, حدد ...

تمـــــاريــــن :

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 ؟

الحــــل :::

اجابة السؤال الاول: احصاء استدلالي لأن الباحث بعد جمع البيانات قام باتخاذ القرارات ضوء ذلك.
التعليل: ان الباحث لم يكتفي بجمع البيانات ولم يكتفي للتعرض للخصائص الاساية للبيانات بل قام باتخاذ القرارات من هذه البيانات التي قام بجمعها.

اجابة السؤال الثاني: مجتمع البحث:هو جميع طلاب السنوات الابتدائية في منطقة ما. اي في المنطقة التي اجريت عليها الدراسة.
نوعه: مجتمع معروف.

اجابة السؤال الثالث: عينه الدراسة: طلاب السنة الثالثة الابتدائية.
نوعها: عينة عنقودية لماذا؟ لأن الباحث لم يصل الى ٢٠٠ طالب من خلال هذه الاعداد المهولة من اعداد السنوات الاولى الابتدائية الامن خلال استخدام وحدة العينة العنقودية. ونحن ذكرنا عشوائيا للتموية فقط

اجابة السؤال الرابع: المتغير المستقل: هو حب الاستطلاع
نوعه:هو متغير وصفي او نوعي او كيفي يعني كيفي وصفي نوعي وهو ذلك النوع من المتغيرات الذي يتغير في الصفة ويسجل باوصاف لفظية فهذا نوعه كيفي او نوعي او وصفي.

اجابة السؤال الخامس: المتغير التابع: حل المسائل الرياضية.
نوعه: متغير تابع كمي لأن الباحث أجرى عليهم اختبار يعني عليه درجات.

اجابة السؤال السادس: المتغيرات الدخيلة: هي تلك المتغيرات التي تؤثر على الدراسة لكنها غير مضبوطة مثل: الناحية الاقتصادية للطلبة، الحالة الاجتماعية لهم، وقت اجراء الاختبار، البيئة الصفيّة.

اجابة السؤال السابع:
الفرضية هي (لاتوجد علاقة ذات دلالة احصائية بين حب الاستطلاع وحل المسائل الرياضية)
نوعها: فرضية صفرية.
قلنا أن الفرضية هي حلول ممكنة لمشكلة الدراسة فنوع الفرضية اذا كانت فرضية صفرية فهي التي تبدأ ((بلا)) فأي فرضية تبدأ ((بلا)) تكون فرضية صفرية وهي تنفي وجود علاقة بين المتغيرات او تنفي وجود فروق بين المتغيرات.

اجابة السؤال الثامن:
الوسيلة هي: استخدام الاختبارات المقننة.

المحــــاضـــرة الـــثــالثـــة

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

أما الطرق الفنية في عرض البيانات الاحصائية فهي:

- العرض الجدولى للبيانات
- العرض البيانى للبيانات

يقصد بالعرض الجدولى للبيانات أن يتم تلخيص البيانات محل الدراسة وتصنيفها فى صورة جداول تعبر عن القيم التى أخذها المتغير من خلال البيانات التى جمعها الباحث و تكرار كل قيمة من تلك القيم.

أهمية الجداول الإحصائية:

- 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 تكل 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.

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 الإحصائية:

تقسم الجداول تبعا 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 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)، (فإذا بدأ من أعلى الى أسفل الجدول) سمى جدول تكرارى متجمع صاعد، (واذا بدأ من أسفل الى أعلى الجدول) سمى 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.

تعريف البيانات : هى مجموعة المشاهدات أو القياسات التى تخص ظاهرة معينة تحت الدراسة .

تعريف المتغير : هو أى خاصية أو صفة سواء للأفراد أو الأشكال والتى تختلف من شخص لآخر ومن وقت لآخر مثل الطول، الذكاء ، التحصيل ويعمل الباحث على دراستها وقياسها.

البيانات إما أن تكون : نوعيه أو كميه . حيث : عملية تبويب وتصنيف البيانات تعتمد على نوع البيانات الإحصائية المراد التعامل معها ودراستها والتى يمكن تقسيمها من حيث طريقة إعداد الجداول إلى التالى :

(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]
- 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) أي تقبل الكسور ، 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: كيف يمكننا عرض البيانات النوعيه والكميه 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: فى دراسة قام بإجرائها أحد الأطباء لطفل معرض لأحد الأمراض النفسية ، تم سؤاله عن لون مجموعة من الأشياء فكانت إجاباته كما يلى :

أحمر	أزرق	بنفسجى	أحمر	أخضر
أبيض	أبيض	أحمر	أزرق	أبيض
أزرق	أحمر	أخضر	أحمر	بنفسجى
أخضر	أزرق	أبيض	بنفسجى	أحمر

المطلوب: عرض البيانات السابقة من خلال جدول تكراري .

اللون	العلامات	التكرارات
أحمر	~~////~~ /	6
أزرق	////	4
أبيض	////	4
أخضر	///	3
بنفسجي	///	3

$\sum = 20$ ($\sum$ هذا الرمز يعني مجموع)

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:

٠	١	٢	٢	٣
١	١	١	٢	١
٢	٢	١	٠	٠
٠	٠	١	٣	١
٢	٠	١	٢	١
١	٠	٠	٠	٣

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

عدد الحوادث	العـلامـات	التكرارات
0	~~////~~ ////	9
1	~~////~~ ~~////~~ /	11
2	~~////~~ //	7
3	///	3

حل المطلوب ١ $\sum = 30$

حل المطلوب٢ :

- أن لايتعرض أي شخص لحادث . عدد الحوادث (٠)
= ٩ ÷ ٣٠ = ٠,٣٠
- ان يكون هناك حادث واحد على الأكثر .

إحتمالات (٠) و (١) = ٩ + ١١ = ٢٠ >>> ٢٠ ÷ ٣٠ = ٠,٦٦٦

- أن يكون هناك حادث واحد على الأقل .

إحتمالات (١) و (٢) و (٣) = ١١ + ٧ + ٣ = ٢١

>>> ٢١ ÷ ٣٠ = ٠,٧

المحـــــاضـــرة الرابـــعة

العرض الجدولي 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 . شرح بسيط لهذه الخطوة ،،

هناك قاعدة تسمى Strugs Ruls يعني ضوابط تحديد عدد الفئات ،، عدد الفئات يرمز له بالحرف K ..

قانون
2^k

كيفية حساب الفئات .. مثال ١٦ – ١٠

هذا الرمز (-) يقصد به أقل من ،، الحل

$2 \times 2 \times 2 \times 2$

4 ، 8 ، 16

إذن هناك 4 فئات K= 4

نطبق بالقانون 2^4

أكبر عدد الفئات ١٥ ،،، أقل عدد الفئات ٤ ،،،

الخطوة الثانية: تحديد طول الفئة .

قـــانـــون :

$$\text{طول الفئة} = \frac{\text{المـدى}}{\text{عدد الفئات المقترح}}$$

المدى = أعلى درجة في التوزيع – أقل درجه

المدى (Rang) يرمز له R

مثال : أحسب المدى لهذه الدرجات ؟

2 , 5 , 7 , 3 , 8 , 12 , 15 , 19 , 22 , 13 , 19 , 24 , 16 ,

أعلى قيمه 24 ، أقل قيمة 2 .. المدى = ٢٤ – ٢ = ٢٢

المدى = ٢٢ ، عدد الفئات المقترح = ٤ << من المثال بالخطوة الأولى .

$$\text{طول الفئة} = \boxed{\frac{٢٢}{٤}} = ٥,٥$$ رقم غير صحيح من الأفضل

نحوله صحيح نقربه ،، ٦ ح .

الخطوة الثالثة: تعيين حدود الفئات . لكل فئة حدين ،، حد أدنى – حد أعلى ..

بداية الفئة — نهاية الفئة

حد أدنى — حد أعلى

مثال : 5 – 10 ،، تعني أن هذه الفئة تحتوي على القيم التالية ،، (5 , 6 , 7 , 8 ,9) إذن طول الفئة 5 على حسب الأعداد اللي بين الأقواس وتوقفنا عند 9 لأنها أقل من 10 (- هذا الرمز يعني أقل) .

لابد عندما نقوم بتعيين حدود الفئات لإنشاء جدول تكراري أن نشمل أقل قيمة بالبيانات موضع الدراسة وأعلى قيمة بحيث أنها تكون جميعها مشمولة في الجدول ..

مثال : 4 – 49 ،، **الحل :**

0 - 5	5 – 10	10 – 15	15 – 20
20 - 25	25 - 30	30 - 35	35 - 40
40 – 45	45 – 50		

أقل قيمة 4 ،،،، أعلى قيمة 49 ،،،،

الخطوة الرابعة: توزيع التكرارات على الفئات .

نقوم برسم الجدول ثم نبدأ بتفريغ البيانات ...

مثال : 4 – 49

الفئات	العلامات	التكرارات
0 -	#### ///	8
10 -	#### ////	9
20 -	#### #### #### /	16
30 -	#### #### /	11
40 - 50	#### /	6

مجموع الفئات = $\sum F$ 50

الفئات يرمز لها بحرف (F) ....

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 للمثال السابق :

اولاً : نحدد عدد الفئة : عدد الفئة هنا = ٦

ثانياً : نحدد طول الفئة : هو المدى مقسوم على عدد الفئات ؟

وعدد طول الفئة المقترح هو : ٦

س/ كيف تم حساب عدد الفئة المقترح ؟

ج/ نضرب العدد ٢ × ٢ × ٢ حتى نوصل عدد الفئة اللي هو : ٦ أو نوصل أعلى قيمه ٤٨ ..

س/ كيف يحسب المدى ؟

ج / المدى= اعلى قيمة في التوزيع ناقص اقل قيمة في التوزيع .

طبعاً في المثال السابق اعلى قيمة ٤٨ واقل قيمة ١ .

نقول : ٤٨ ـ ١ = ٤٧

إذاً يتم تحديد طول الفئة : ٤٧ ÷ ٦ = ٧.٨

اذاً نقربها الى اقرب عدد ممكن هو : ٨

ثالثاً : تعيين حدود الفئات .

يتم تعيين حدود الفئات بـ اقل قيمة في البيانات = ١

واكبر قيمة في البيانات = ٤٨

لذا لابد من شمول هذي القيم في حدود الفئات المقترح .

كيفية الشمول هكذا :

من ١ الى اقل من ٩

من ٩ الى اقل من ١٧

من ١٧ الى اقل ٢٥

من ٢٥ الى اقل من ٣٣

من ٣٣ الى اقل من ٤١

من ٤١ الى اقل من ٤٩

أحياناً تكون الفئة أكثر من المقترح بسبب التقريب إلى أقرب رقم صحيح ،، وهذا يمكن أن يشار له من خلال (*)
ويوضح سبب التقريب زاد عدد الفئات عن عدد الفئات المقترح وهذا لايضير في الأمر .

رابعاً : توزيع التكرارات ع الفئات .

يتم توزيع التكرارات على الفئات برسم جدول كتالي :

ملاحظة : الرجوع الى المحاضرة المسجلة لتعرف اكثر الى شرح مفصل وكيفية عمل الجدول.

وهناك عدة ملاحظات يجب الانتباه 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 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 هي:**

1. **الجداول التكرارية المنتظمة : يقصد بها ان يكون اطوال الفئات متساوية . كمافي المثال السابق .**
2. **الجداول التكرارية غير المنتظمة : يقصد بها ان اطوال الفئات غير متساوية .**

مثال :

طول الفئة = الحد الأعلى للفئة - الحد الأدنى

5 ٥٥-٥٠ ١٠

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.

٥٠-٤٠

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

١٠-٢ ٢٠-٤

١٠ ٢٠

ملاحظة : من الأفضل إننا نتعامل مع بيانات ذات جداول منتظمة ، حتى يسهل المقارنة والتعامل معها بسهولة .

٣. الجداول التكرارية المفتوحة : يقصد بها انه يحتمل ان تكون بدايتها غير معروفة ونهايتها غير معروفة .

مثال :

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 فاكثر	١٨

جدول مفتوح من أعلى

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

جدول مفتوح من الطرفين

الجداول التكرارية المتجمعة:

وهى 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 التكراري المتجمع الصاعد :

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

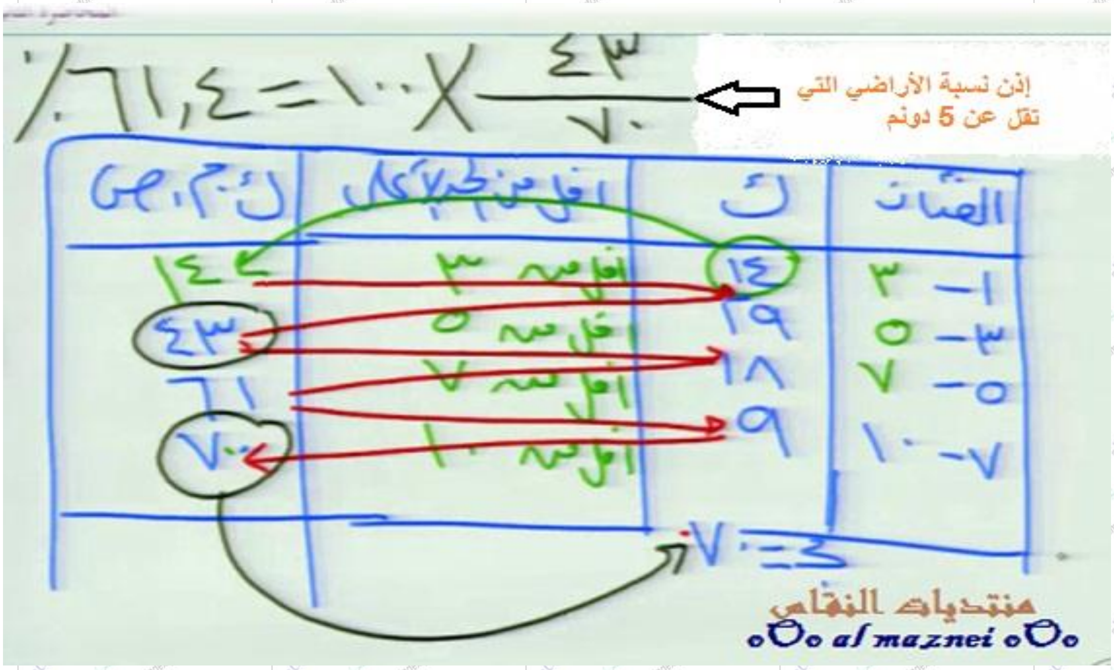
مثال :

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

فئات مساحات الأراضي دونم	عدد قطع الأراضي
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 :



الفئات	ك	أقل من الحد الأعلى	ك.م.ص
١ – ٣	١٤	أقل من ٣	١٤
٣ – ٥	٢٩	أقل من ٥	٤٣
٥ – ٧	١٨	أقل من ٧	٦١
٧ – ١٠	٩	أقل من ١٠	٧٠
مج = ٧٠			

ملاحظة : الرجوع للمحاضرة المسجلة لتعرف اكثر ع طريقة الحل .

ثانيا : الجدول التكراري المتجمع الهابط (النازل):

- **ويعطى الجدول المتجمع الهابط (النازل) الحدود الدنيا للفئات وعدد المفردات التي تكون أكثر من أو تساوى الحدود الدنيا لكل فئة (وتكتب بصيغة الحد الأدنى فأكثر).**

مثال :

منتديات النقاش
oOo al maznei oOo

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 الأراضي دونم	عدد قطع الأراضي
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 :

منتديات النقاش
oOo al amznei oOo

الحل تفصيلا في الكتاب

الفئات	التكرار	الحد الأدنى فأكثر	ك.م.ه.ن
١–	١٤	١ فأكثر	٧٠
٣–	٢٩	٣ فأكثر	٥٦
٥–	١٨	٥ فأكثر	٢٧
٧–١٠	٩	٧ فأكثر	٩
مج = ٧٠			

ملاحظة : الرجوع للمحاضرة المسجلة لتعرف اكثر ع طريقة الحل .

الجدول التكراري المزدوج:

● عند دراستنا لمتغيرين لتحديد العلاقة بينهما مثل العلاقة بين عدد أفراد الأسرة والمستوى التعليمي أو العلاقة بين أجور العامل ودرجة الرضاء الوظيفي أو ما شابة ذلك، في هذه الحالة لابد 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 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 في الكتاب

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

ملاحظة : الرجوع للمحاضرة المسجلة لتعرف اكثر ع طريقة الحل .

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) $\frac{7}{20} = 0.35$

خد بالك : المطلوب نسبة (وليس نسبة مئوية)

(٢-د) $0.35 \times 100 = 35\%$

أيوه . . ده بقى نسبة مئوية

المحــــاضرة الخامسـة

العرض البيانى للبيانات

أولا: البيانات غير المبوبة

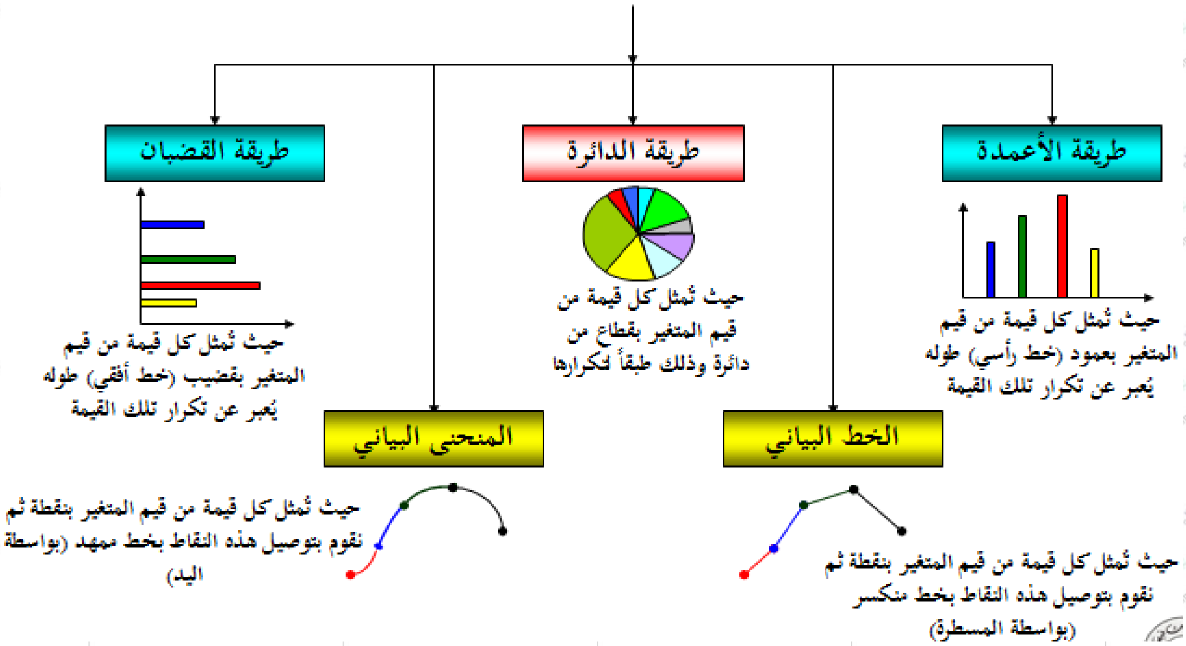
تعريف الرسوم البيانية:

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 ..

خصص هذا الجزء للبيانات المنفصلة لأن الرسوم البيانية تختلف حسب طبيعة ونوع البيانات المراد عرضها فاذا كانت البيانات اسمية أو رتبية (أي منفصلة) فإننا نستخدم أحد الأشكال البيانية التالية

طرق العرض البياني للبيانات المنفصلة



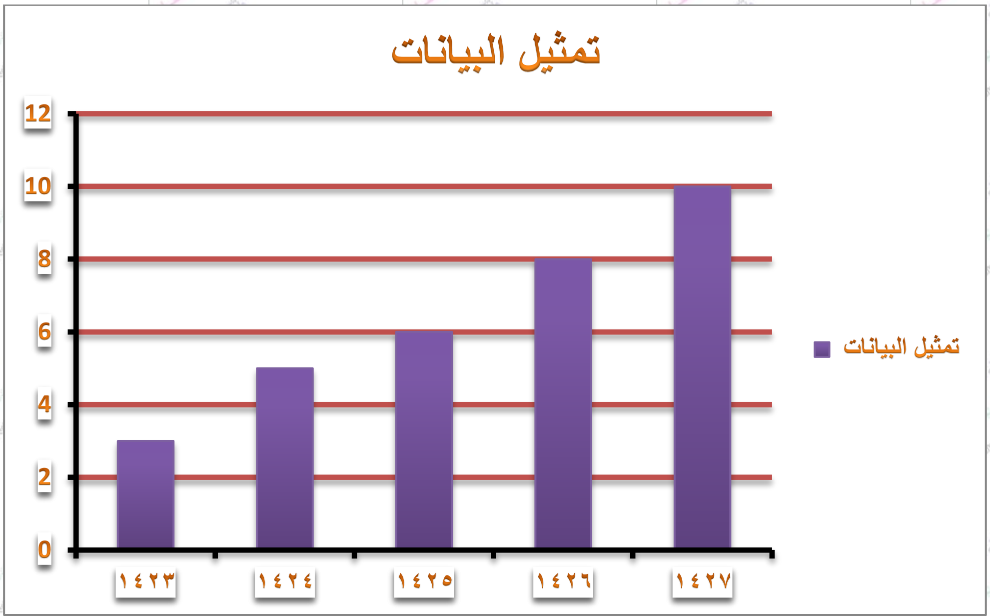
١- **طريقة الأعمدة:** ويتم عرض البيانات من خلال هذا الأسلوب من خلال عدة أنواع من الأعمدة البيانية وهي:

أ- الأعمدة البيانية البسيطة : وهي عبارة عن مجموعة من الأعمدة الرأسية أو المستطيلات المتساوية القاعة والتي تتناسب ارتفاعاتها مع البيانات التي تمثلها .

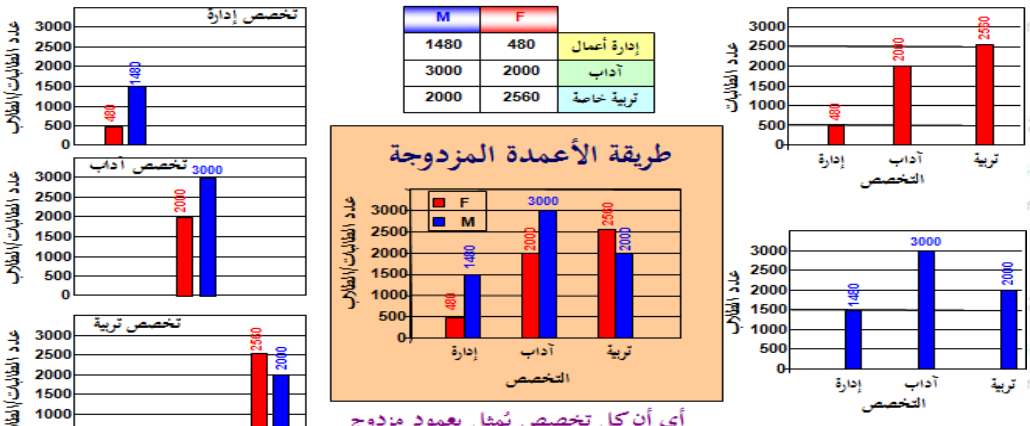
مثال: الجدول الآتي يوضح أعداد الطلاب المقيدين بأحد الجامعات في السنوات الدراسية من ١٤٢٣ هـ حتى ١٤٢٧ هـ.

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

المطلوب :
تمثيل البيانات باستخدام الرسم البياني المناسب



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 .

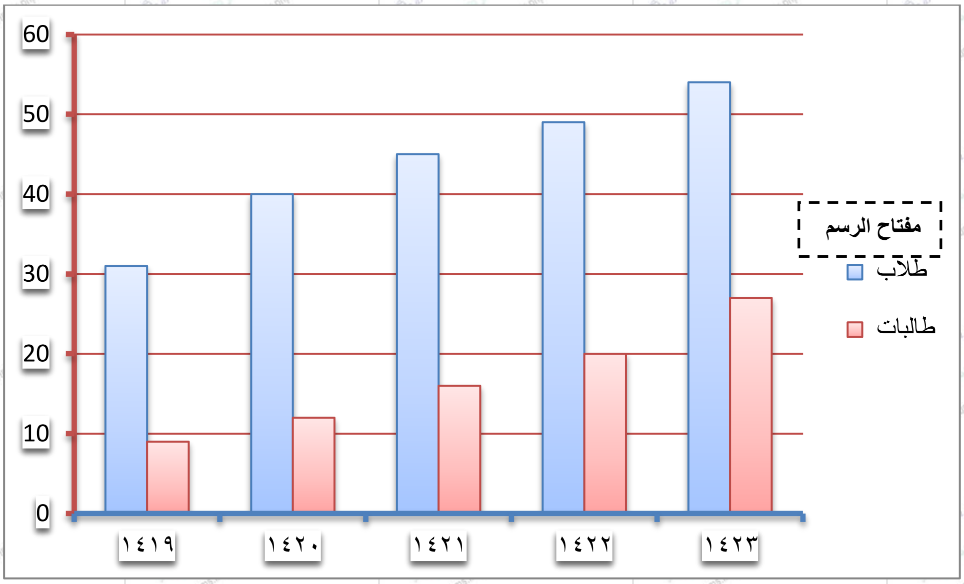


أي أن كل تخصص يُمثل بعمود مزدوج مكون من عمودين بسيطين متلاصقين

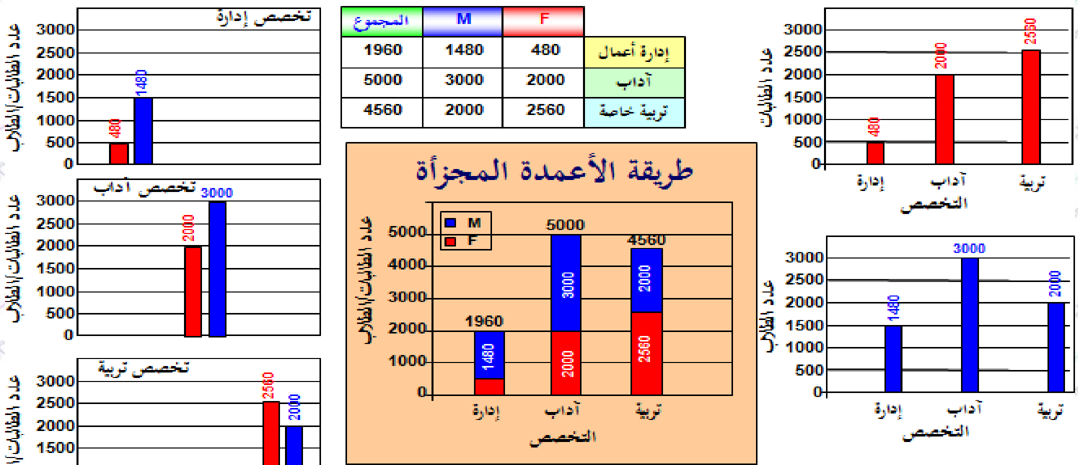
مثال: الجدول الآتي يوضح أعداد الطلبة المسجلين بأحد الجامعات السعودية في السنوات الدراسية ١٤١٩ هـ حتى ١٤٢٣ هـ

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 ؟



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 .

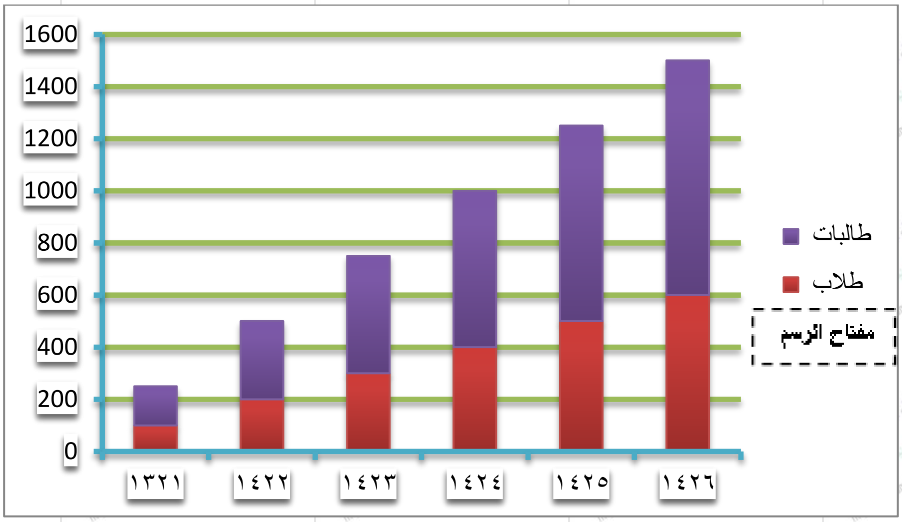


أي أن كل تخصص يُمثل بعمود طوله يُعبر عن مجموع عدد طالباته وطلابه معاً ثم يتم تجزئته إلى عمودين كل منهما يمثل فئة من الفئات

مثال: اذا كانت اعداد الطلاب والطالبات المسجلين في كلية التربية بجامعة الملك فيصل بالإحصاء 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؟



ويمكن إبداء الملاحظات التالية على الرسومات بالاعمدة البيــانية بأنواعها المختلفة :

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

د – اللوحة الدائرية: تستخدم الدائرة أو اللوحة الدائرية لثميل 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:

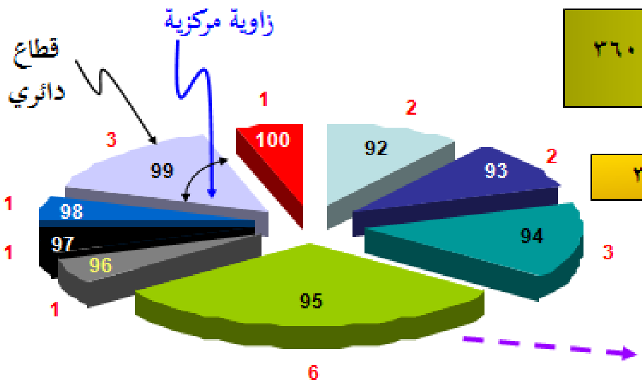
الدرجة x	92	93	94	95	96	97	98	99	100
التكرار f	2	2	3	6	1	1	1	3	1

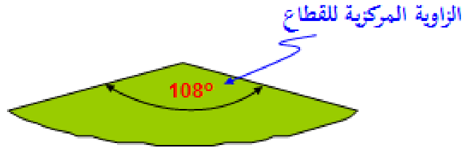
تحدد زاوية المركزية بالعلاقة :

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

أو

الزاوية المركزية لقيمة ما = التكرار النسبي للقيمة × ٣٦٠





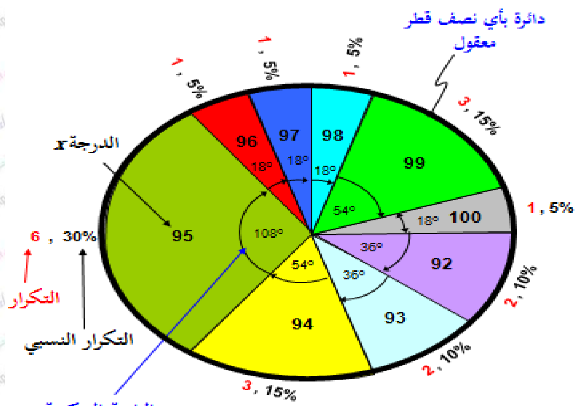
القيم داخل القطاعات تمثل الدرجة (المتغير) x والقيم المكتوبة خارج القطاعات باللون الأحمر تمثل التكرار f

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

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

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

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

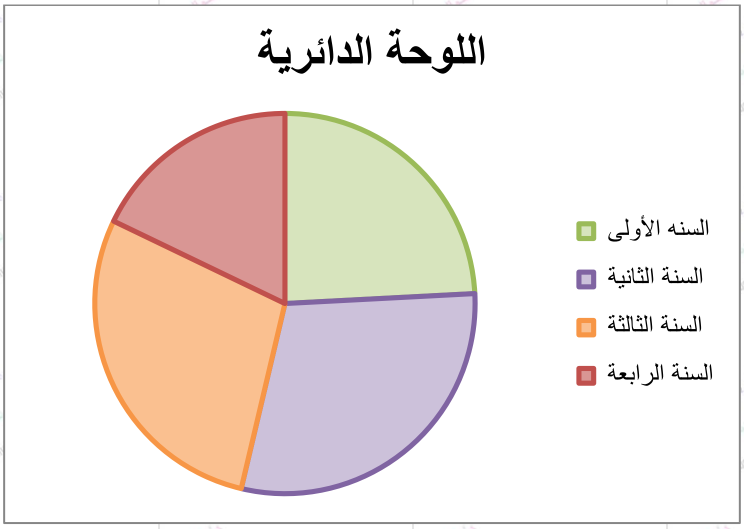
الحل : نطبق بالقاعدة :

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

$$\left[\frac{٢٢٦}{٩٣٥}\right] \times ٣٦٠ = ٨٧ \qquad \left[\frac{٢٧٦}{٩٣٥}\right] \times ٣٦٠ = ١٠٦{،}٣$$

$$\left[\frac{٢٦٦}{٩٣٥}\right] \times ٣٦٠ = ١٠٢{،}٤ \qquad \left[\frac{١٦٧}{٩٣٥}\right] \times ٣٦٠ = ٦٤{،}٣$$

نرسم دائرة بنصف قطر مناسب :



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

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

- عندما تكون الكميات المقارنة كثيرة العدد نسبيا.
- عند ما تكون الاجزاء المقارنة في فترات زمنية مختلفة.

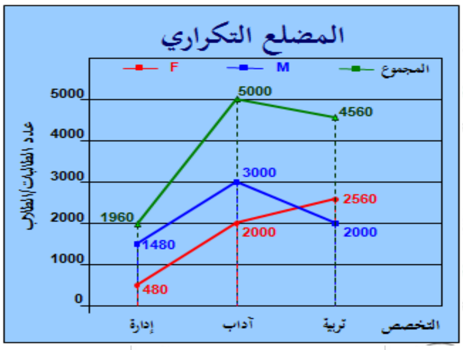
- عندما نرغب في توضيح قيم الاجزاء المقارنة المختلفة للظاهرة موضع البحث وذلك من أجل ابراز المقارنة بين هذه الأجزاء أو توضيح التغير أو التطور عبر الزمن سواء لظاهرة واحدة أو عدة ظواهر بين فترات زمنية مختلفة.
- غالبا ما ينصح باستعمال الاعمدة البيانية (بأنواعها المختلفة) مع المتغيرات المنفصلة.

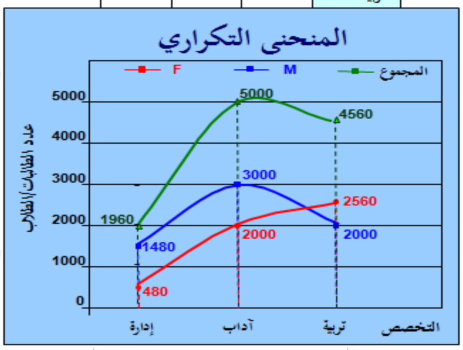
هـ – المنحنى أو الخط البياني:

- 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

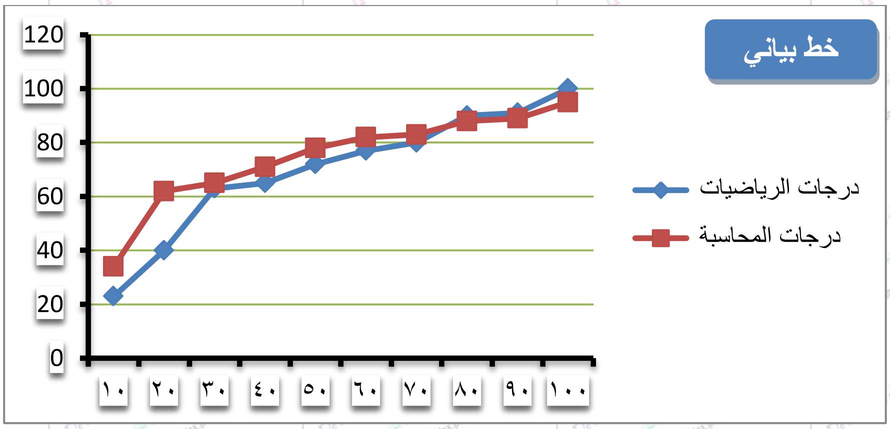


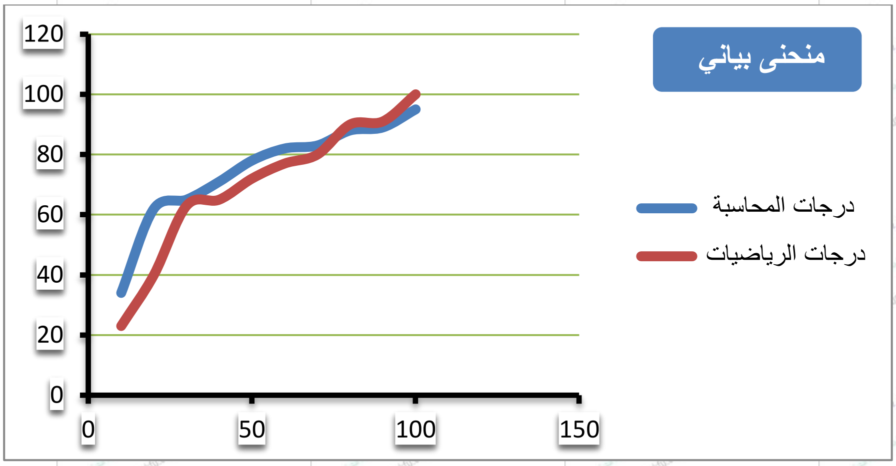


مثال: البيانات التالية لدرجات عشر طلاب بكلية العلوم الإدارية في مقرري الرياضيات والمحاسبة، فكانت كمايلي :

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).



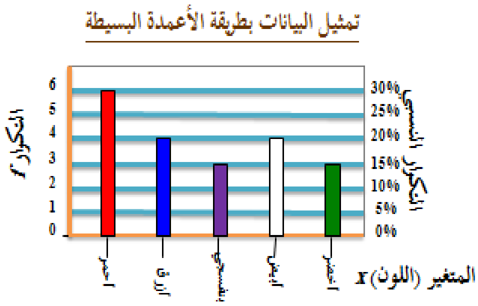


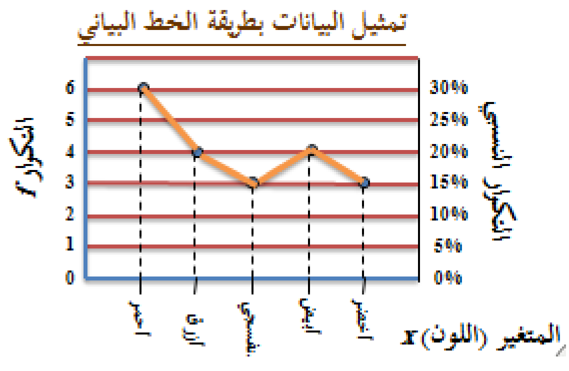
مثال: لوكانت لدينا البيانات التالية :

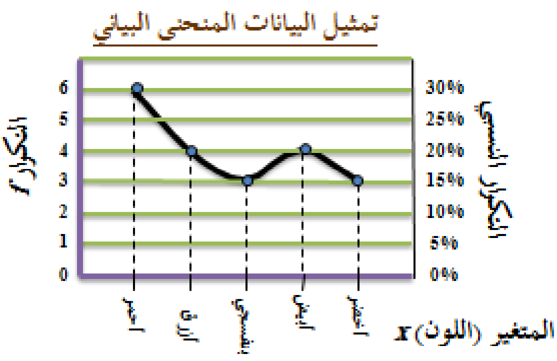
x	f	الزاوية
أحمر	6	108
أزرق	4	72
بنفسجي	3	54
أبيض	4	72
أخضر	3	54

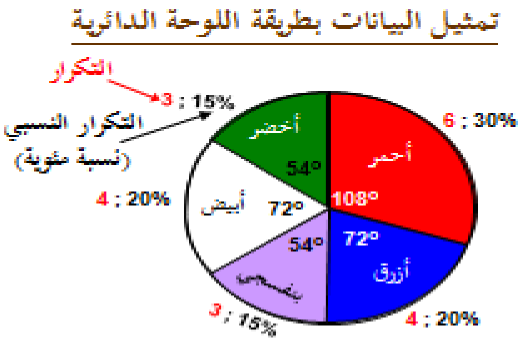
المطلوب: مثل البيانات السابقة بـ:

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









ملاحظات على المنحنى والخط البياني :

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

مزايا وعيوب الرسوم البيانية :

المزايا:

- **تثير انتباه المشاهد خاصة اذا كانت جيدة التصميم.**
- **توفر وقت المشاهدة اذ أن استنباط الحقائق من الرسوم البيانية أسرع من الوصول اليها بواسطة الأرقام الموضوعة في جداول.**
- **إمكانية معرفة الاتجاهات العامة للظواهر.**
- **سهولة فهم وتذكر العلاقات بين الظواهر محل الدراسة.**

العيوب:

- **التضحية بدقة البيانات اذ أن الرسوم توضح فقط التغيرات العامة للظواهر ولا تبين التفاصيل الدقيقة لها.**
- **أحيانا تكون الرسوم معقدة، خاصة إذا كانت تشتمل على مجموعات من البيانات المتباينة.**
- **كثرة التكاليف خاصة إذا كانت البيانات تحتاج الى مقياس رسم كبير.**

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	?	?
	$\sum f$	

(أ) عدد الممرضات ذات العمر 25 سنة هو :

● 10 ○ 20 ○ 30 ○ 40

(ب) الزاوية المركزية المناظرة للعمر 30 سنة هي :

○ 36° ○ 72° ● 108° ○ 144°

(ج) الزاوية المركزية المناظرة للعمر 35 سنة هي :

○ 36° ○ 72° ○ 108° ● 144°

(د) عدد الممرضات الكلي [أي مجموع التكرارات $\sum f$] هو :

○ 95 ● 100 ○ 105 ○ 110

هامش للإجابة

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

20	72°
?	36°

$72 \times ? = 36 \times 20 \quad , \quad \therefore \; ? = 10$

(١-ب) بنفس الأسلوب السابق

20	72°
30	?

$72 \times 30 = ? \times 20 \quad , \quad \therefore \; ? = 108^\circ$

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

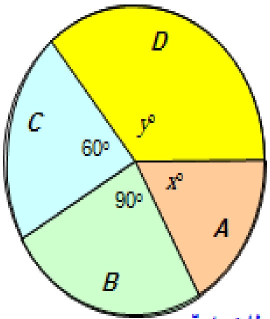
$\therefore \; 72 + 36 + 108 + ? = 360 \quad , \quad \therefore \; ? = 144^\circ$

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

20	72°
$\sum f$	360°

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

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 :

● 25% ○ 30% ○ 40% ○ 60%

(ب) عدد اللعب التي باعتها الشركة B هو :

○ 2700 ○ 2250 ○ 900 ● 1350

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

○ 900 ○ 2250 ● 3150 ○ 1350

هامش للإجابة

(٢-أ)

100%	360°
?	90°

$360 \times ? = 90 \times 100$

$? = 25\%$

(٢-ب)

$\frac{25}{100} \times 5400 = 1350$

$\sum f$

(٢-ج)

الزاوية المركزية المناظرة لمبيعات الشركتين معاً تساوي

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

5400	360°
?	210°

$360 \times ? = 210 \times 5400$

$? = 3150$

المحـــاضــرة الســـادسة

العرض البيانى للبيانات

ثانيا: البيانات المبوبة

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

- المدرج التكرارى
- المضلع التكرارى
- المنحنى التكرارى
- المنحنى التكرارى المتجمع الصاعد
- 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 السيني) على أساس الفئات وهنا يظهر حالتين هما:

الحالة الأولى:- تساوى أطول الفئات (الجداول المنتظمة)

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

الحالة الثانيه:- عدم تساوى أطوال الفئات (الجداول الغير منتظمة)

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	٠ـ	١٠ـ	٢٠ـ	٣٠ـ	٤٠ـ٥٠	المجموع
عدد الشركات	٨	٩	١٦	١١	٦	٥٠

المطلوب:

عرض البيانات السابقة في شكل المدرج التكرارى.

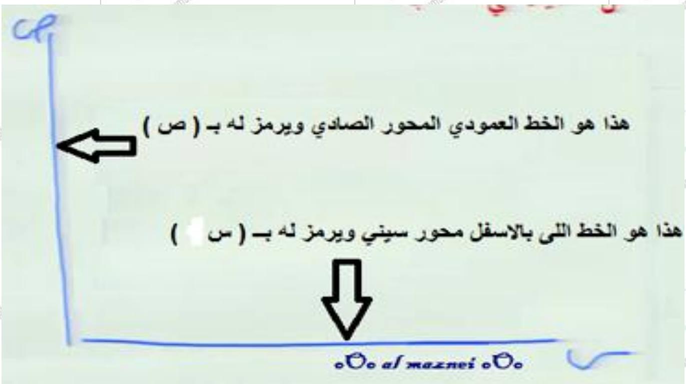
شرح مفصل :

المدرج التكراري : هي اعمدة مستطيلة متلاصقة تعبر عن مستوي الفئة.

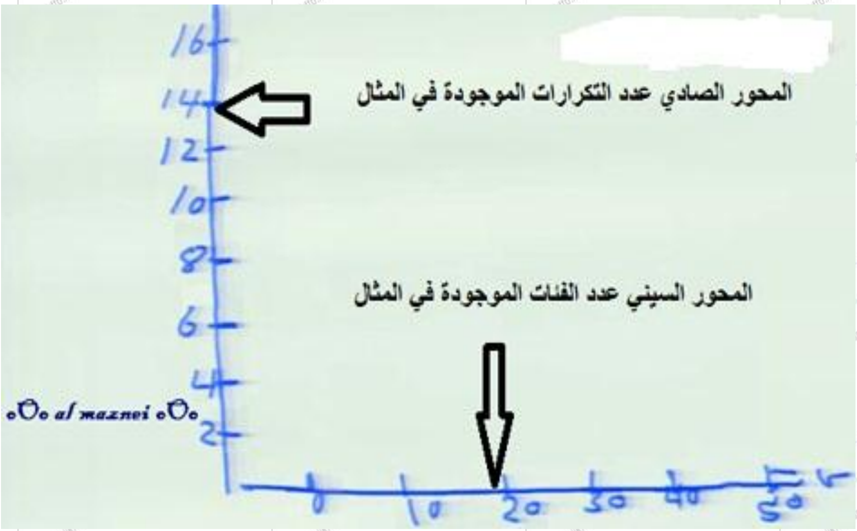
بعطيك بالعامية <<شكل المدرج التكراري يجي ع شكل الدرج >>
طبعاً : كل مستطيل يمثل فئة والارتفاع يمثل التكرار لهذي الفئة (التكرار هو العدد للفئة)

نرسم خطين:
الخط اللى في الاسفل يكون في عدد الفئات اللى في المثال ويرمز له بـ (المحور السيني)
خط عمودي يكون التكراراي العدد الناتج عن الفئات ويرمز له بـــ (المحور الصادي)

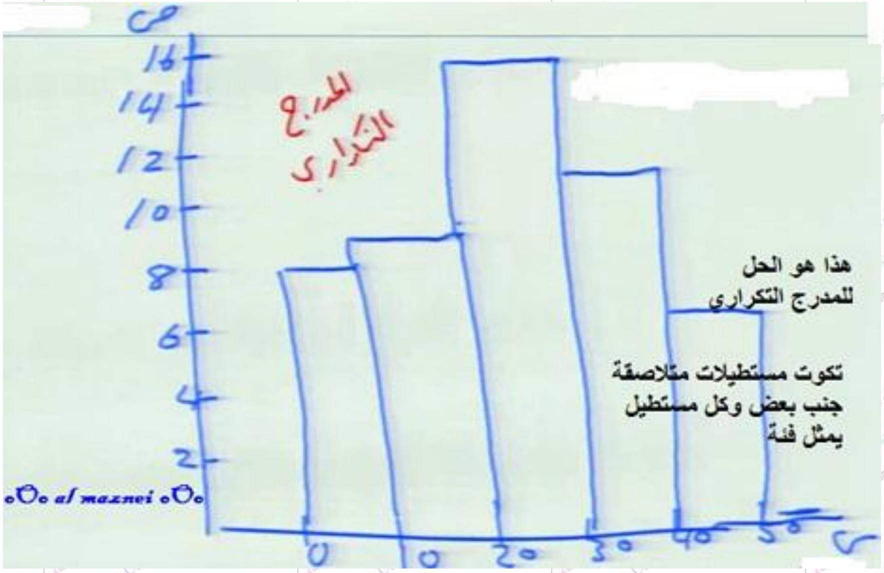
مثل كذا:



بعد ما نرسم الخطوط:
نقسمها المحور السيني راح يكون عدد الفئات
والمحور الصادي يكون التكرارت:



بعد كذا نقوم برسم المستطيلات:
نشوف اول فئة كم عدد تكراراها:
اول فئة في المثال تكرارها (٨) ،،،،، ثاني فئة (٩) ولا بد ان يكون التكرار ملاصق للاول ،،،،
ثالث فئة (١٦) ،،،،،،،،،،، رابع فئة (١١) ،،،،،،،،،،، خامس فئه (٦)
هذا الحل برسم المدرج التكراري:



بعض خصائص التوزيع التكراري:

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

- **الخاصية الأولى : التماثل (بحيث أنه إذا تم تقسيم الشكل من النصف لأنطبق الجزئين)**
- **الخاصية الثانية : الالتواء (قد يكون الالتواء موجب لأن اتجاه يقع في جهة اليسار ،، بمعنى ان أعلى قيمة في اليسار وأقل مستطيل يقع في اليمين) أو (قد يكون الالتواء سالب لأن اتجاه التوزيع في جهة اليمين ،، بمعنى ان أعلى قيمة في اليمين وأقل مستطيل يقع في اليسار)**
- **الخاصية الثالثة : المنوال (هو أعلى قيمة في التوزيع)**

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

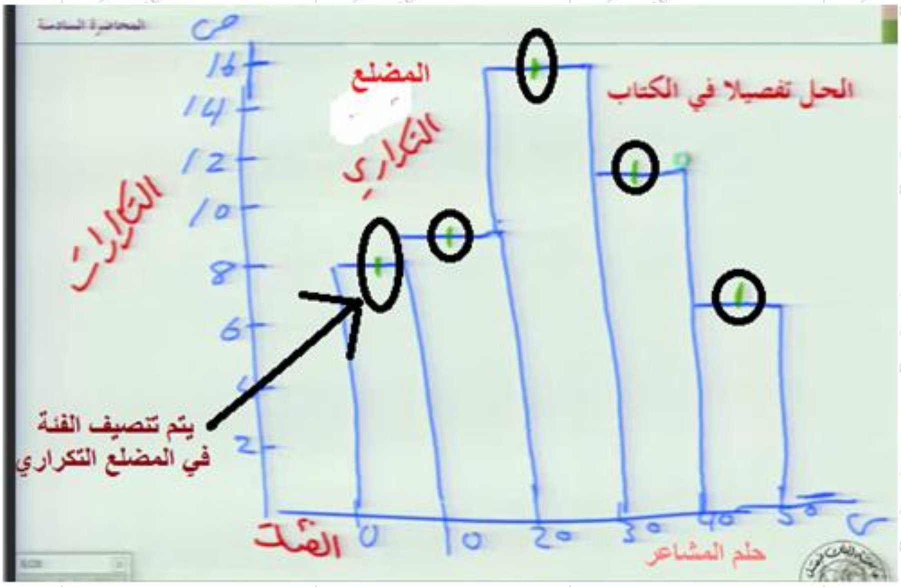
فئات رأس المال	-٠	-١٠	-٢٠	-٣٠	٥٠-٤٠	المجموع
عدد الشركات	٨	٩	١٦	١١	٦	٥٠

المطلوب:

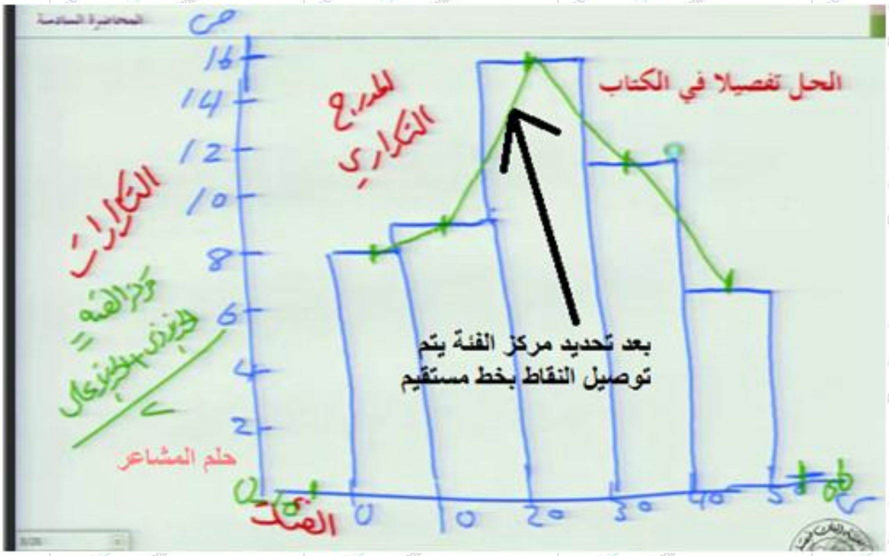
عرض البيانات السابقة في شكل مضلع التكرارى.

الطريقة الأولى : عن طريق المضلع التكراري ..

الخطوة الأولى : تحديد مركز الفئة عن طريق تنصيف كل الفئات .

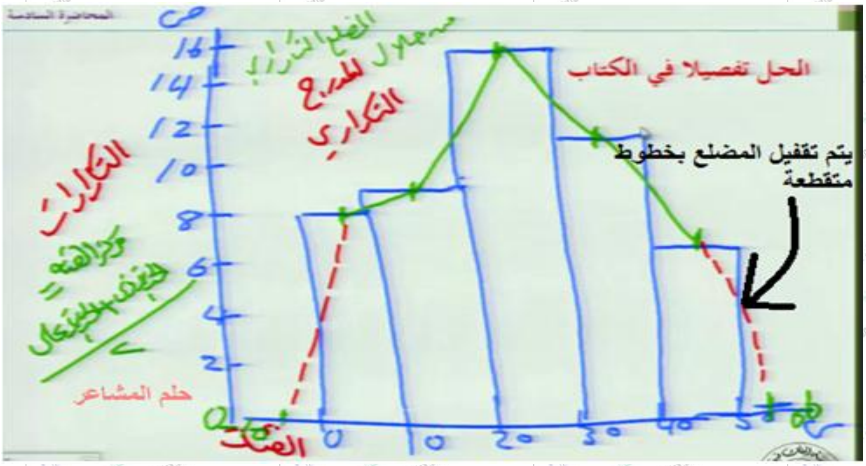


الخطوة الثانية : توصيل النقاط

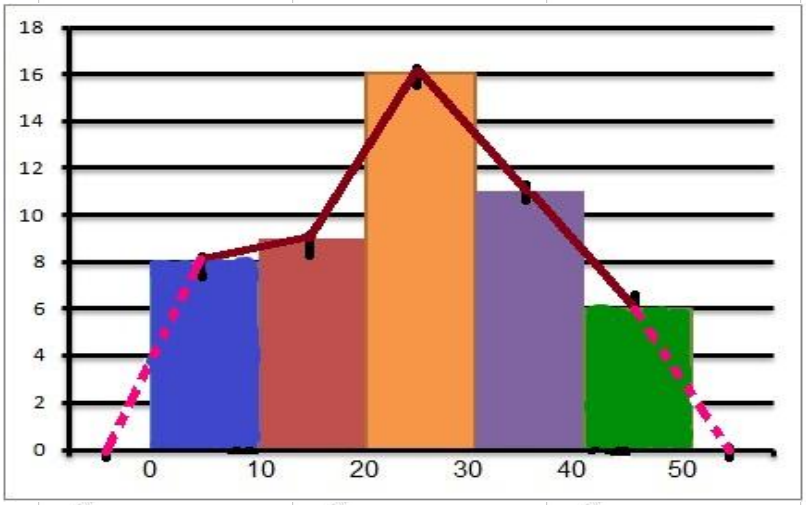


الخطوة الثالثة : إغلاق المضلع . ويتم تحديد مراكز الفئة = (الحد الأدنى + الحد الأعلى) ÷ 2

(-10 – 0) ÷ 2 = -5 ،،،،،،، (60 – 50) ÷ 2 = 55



توضيح الشكل اكثر



الطريقة الثانية : تحديد النقاط عن طريق مركز الفئات ..

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 مضلع التكرارى.

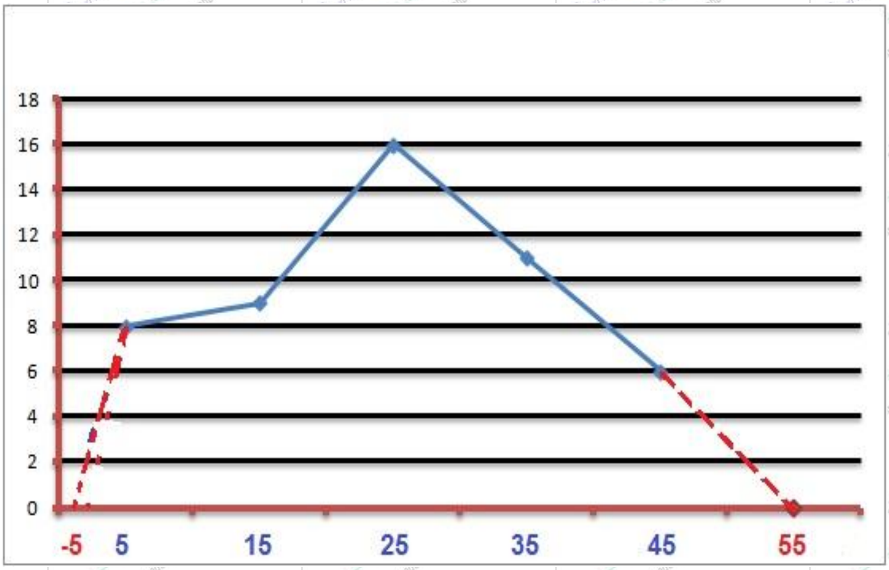
نطبق القانون مركز الفئة = (الحد الأدنى + الحد الأعلى)÷ 2

مركز الفئة = (-10 + -0) ÷ 2 = 5

مركز الفئة = (-20 + -10) ÷ 2 = 15

نحسب بقية الفئات بنفس الطريقة

الرسم :



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 منحنى التكرارى.

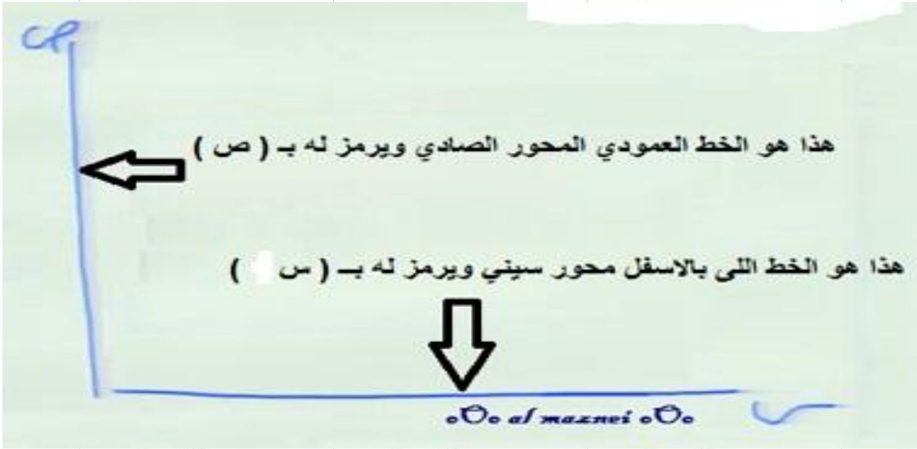
شرح تفصيلي :

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

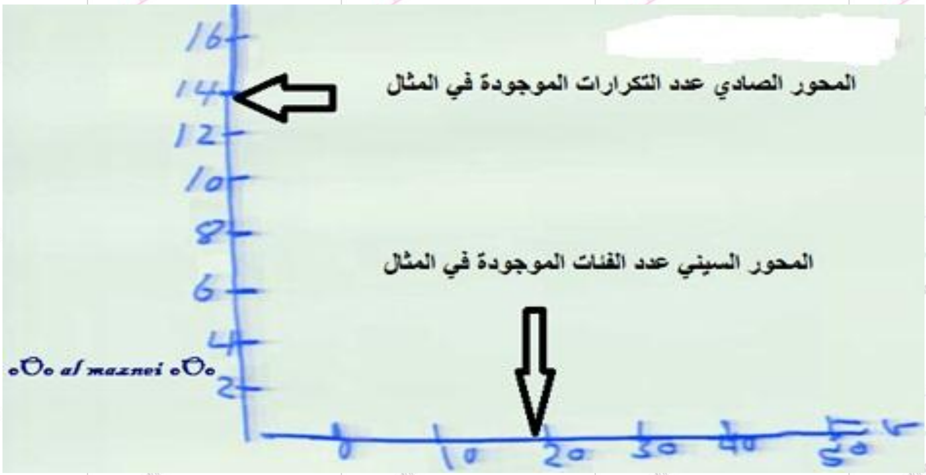
المعني من التعريف : يعني يكون رسم الخطوط باليد ويكون ممهد <<يعني ما نستخدم المسطره يكون خط عادي .

ملاحظة : القيم الشاذة في المنحني التكراري يتم تجاهلها..

اولاً : نقوم برسم المحور الصادي والسيني بهذا الشكل:



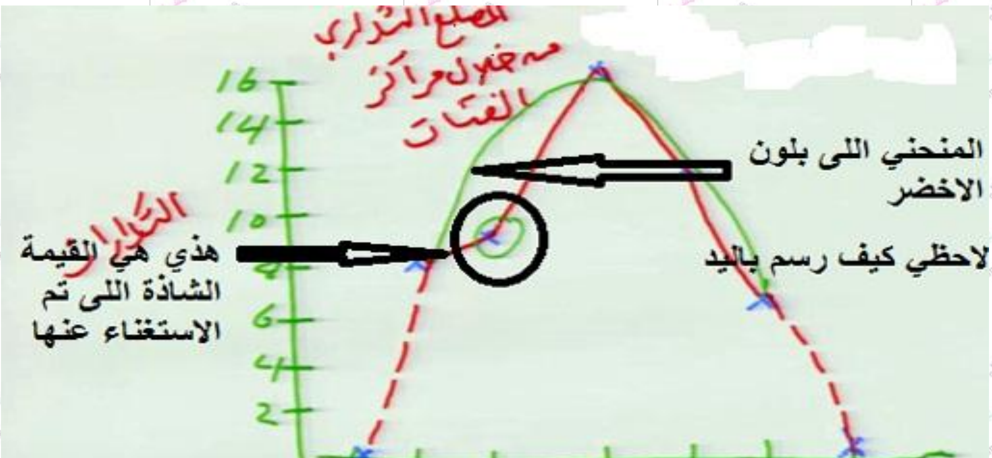
ثانياً : نقوم بعملية الترقيم للفئات والتكرارات بهذا الشكل:



ثالثا : الرسم النهائي

لا حظوا ان الحل للمنحنى التكراري في الصورة باللون الاخضر.

وكمان القيمة الشاذة اللى استغنينا عنها عليها دائرة:

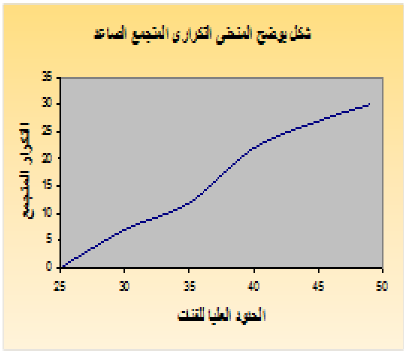


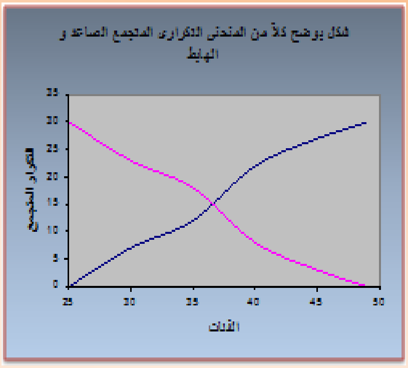
التوزيعات التكرارية المتجمعة:

تستخدم المنحنيات المتجمعة 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 المطلقة 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 المطلقة للتكرارات أو بالتكرار النسبى، ويراعى وضع النقاط الخاصه بالتكرارات المتجمعه الهابطه (النازلة) عند الحد الأدنى لكل فئه، لأنه يعبر عن العدد الاجمالى لأوجه الظاهرة الواقع أعلي الحد الأدنى للفئه.



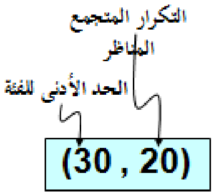


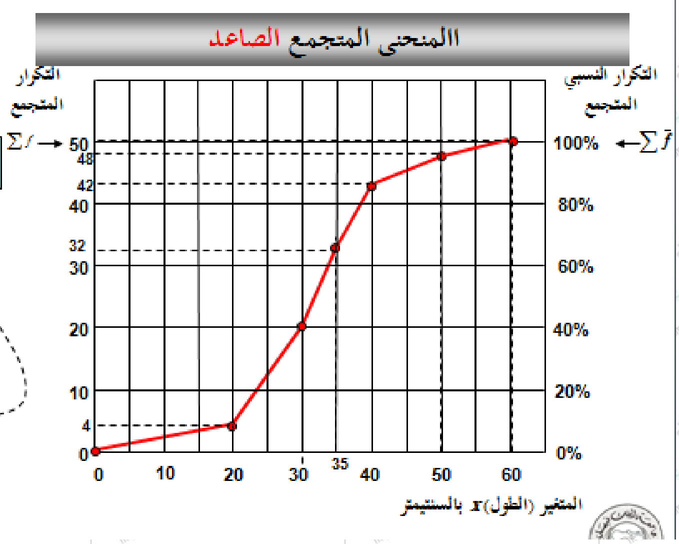
المنحنى المتجمع الصاعد :

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
	$\sum f = 50$



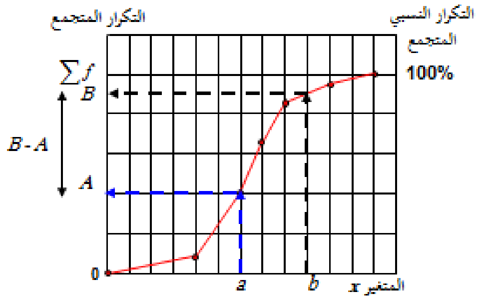


التوزيع التكراري المتجمع الصاعد

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 :

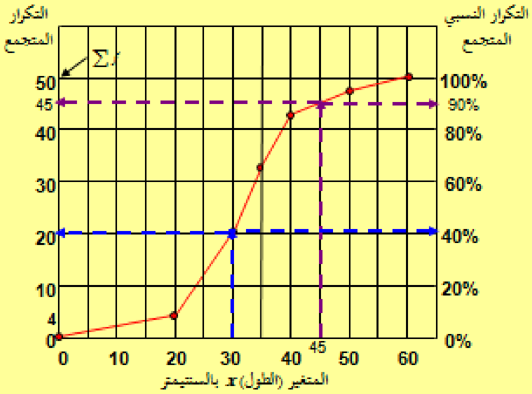
"x محصورة بين قيمتين"



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

الفرق بين القيمتين A , B

فمثلاً في المثال التوضيحي السابق



عدد الأزهار التي أطوال سيقانها ما بين 30 ، 45 هو :

45 – 20 = 25

ونسبتهم المئوية تساوي $\frac{25}{50}$ x 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
$0 \le x < 20$	4
$20 \le x < 30$	16
$30 \le x < 35$	12
$35 \le x < 40$	10
$40 \le x < 50$	6
$50 \le x < 60$	2
	$\sum f = 50$

التوزيع التكراري المتجمع الهابط

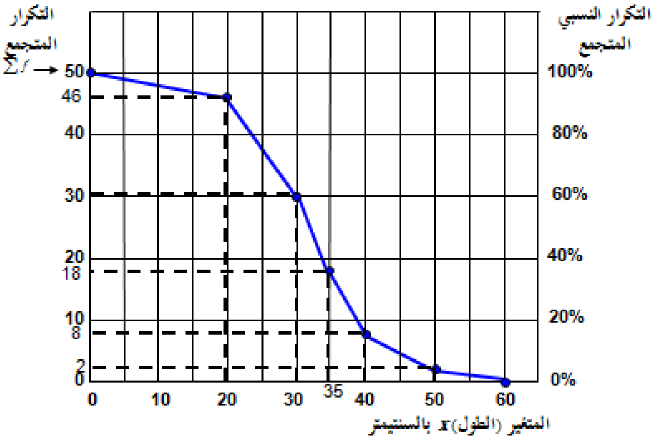
المتغير x	التكرار المتجمع	التكرار النسبي المتجمع	النقطة الموقعة على الرسم
≥ 0	50	100%	(0 , 50)
≥ 20	46	92%	(20 , 46)
≥ 30	30	60%	(30 , 30)
≥ 35	18	36%	(35 , 18)
≥ 40	8	16%	(40 , 8)
≥ 50	2	4%	(50 , 2)
≥ 60	0	0%	(60 , 0)

التكرار المتجمع المناظر

الحد الأدنى للفئة

(30 , 30)

المنحنى المتجمع النازل أو الهابط

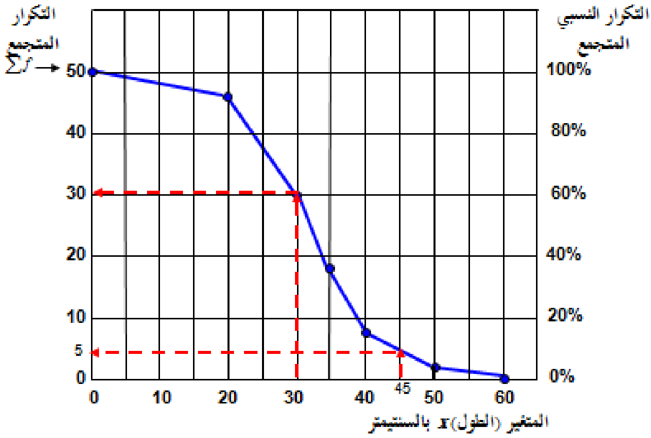


ويفيد المنحنى المتجمع النازل أو الهابط في الرد على نفس الأسئلة التي يرد عليها المنحنى المتجمع الصاعد مع الأخذ في الاعتبار أن التدريج الرأسي [التكرار المتجمع] يمثل التكرار المناظر لـ "x أكبر من أو تساوي"

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 تقريبا

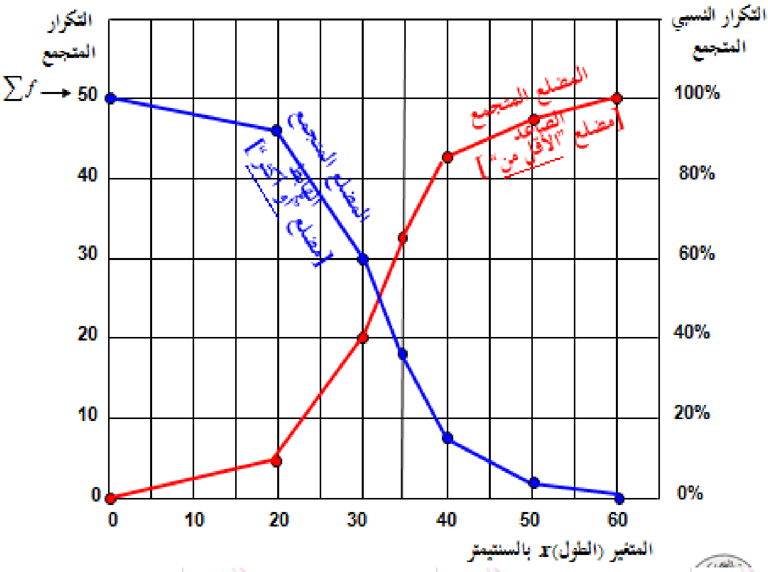
ويمكن رسم المضلعين التكراريين المتجمعين : الصاعد والهابط على رسمة واحدة كما هو مبين :

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)

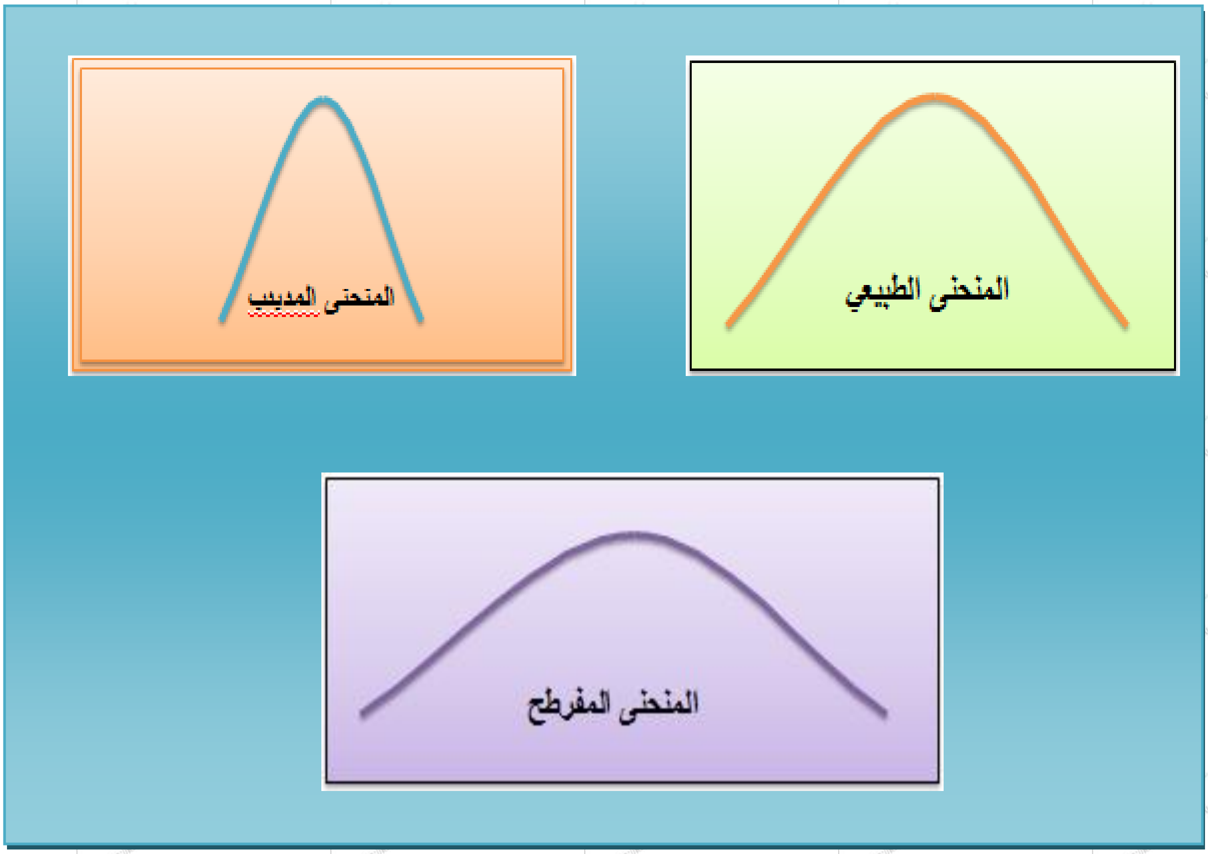


الأشكال الشائعة للتوزيعات التكرارية :

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

وفى أحيان أخرى يكون المنحنى التكرارى مدبب القمة بحيث تكون القمه ضيقة وذو طرفين واسعين نسبياً، فيسمى فى هذه الحالة منحنى قليل التفرطح أو المنحنى المدبب.

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 المدبب والمفرطح.



المحـــاضــرة 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

تساعدنا المقاييس الإحصائية 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 الإحصائية إلى نوعين رئيسيين هما:

- مقاييس النزعة المركزية Central Tendency Measures ، تقيس مدى تمركز الدرجات نحو درجة معينة .
- مقاييس التشتت أو الأنتشار Dispersion Measures ، تقيس مدى تباعد الدرجات بعضها عن بعض .

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 التى تجمع أكبر عدد من القيم الخاصة بمجموعة معينة عندها . ولتحديد القيمة المتوسطة للتوزيع يوجد هناك عدة مقاييس أهمها :

- المتوسط الحسابي
- الوسيط
- 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 الحسابى بأنه قيمة التى اذا أعطيت لكل مفرد من مفردات الظاهرة لكان مجموع القيم الجديدة مساويا للمجموع الفعلى للقيم الأصلية للظاهرة، ويتم حساب المتوسط الحسابي للبيانات غير المبوبة من خلال المعادلة التالية:

$\bar{x} = \frac{\sum_{i=1}^{n} xi}{n}$ >> يمكن استبدال هذا القانون بقانون أسهل <<<< $\bar{x} = \frac{\sum x}{n}$

توضيح للقانون :

هذا الرمز $\sum x$ يعني مجموع مفردات الظاهرة الأصلية ..

هذا الرمز n يعني عدد المجموعات ..

هذا الرمز $\bar{x}$ يعني الوسط الحسابي أو المتوسط الحسابي ..

مثال توضيحي للتعريف : مفردات الظاهرة ، 3 , 7 , 5 . ..

نطبق بالقانون $\bar{x} = \frac{\sum x}{n}$ $= \frac{3+7+5}{3} = \frac{15}{3} = 5$ >> إذن $\bar{x} = 5$

لو وزعنا قيمة المتوسط الحسابي (5) إلى مفردات الظاهرة وجمعناها لأعطى نفس الناتج للوسط الحسابي أي :

3 , 7 , 6
↓ ↓ ↓
5 , 5 , 5

3 + 7 + 6 = 15

5 + 5 + 5 = 15

من شرح الدكتور ....

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

نطبق بالقانون : $\overline{x} = \frac{\sum x}{n}$

للتوضيح : $\sum x$ مجموع المبيعات .. n عدد الأشهر .. $\overline{x}$ الناتج (المتوسط أو الوسط الحسابي) ..

$$\frac{3 + 5 + 8 + 3 + 6 + 4 + 12 + 5 + 4 + 3 + 7 + 9}{12} = \frac{69}{12} = 5,75$$

$$\therefore \overline{x} = 5,75$$

ويجب ملاحظة عدة أمور في الوسط الحسابي وهي:

- انه لا يشترط أن يكون المتوسط الحسابي عددا صحيحا.
- ان المتوسط الحسابي دائما محصور بين أقل القيم وأعلاها، ولكن هذا لا يعني أنه يقع في الوسط تماما بين هذين الحدين.
- إن المجموع الجبري لانحراف القيم عن المتوسط يكون دائما صفر.

شرح لهذه الملاحظة : مثال : 3,7,4,6

نعوض بالقانون السابق $\frac{3 + 7 + 4 + 6}{4} = \frac{20}{4} = 5$

ننقص الناتج (المتوسط الحسابي) من مفردات الظاهرة الأصلية .

3 – 5 = -2 ,,,,, 7 – 5 = +2 ,,,,, 4 – 5 = -1 ,,,,, 6- 5 = + 1

نجمع الناتج : -2 + 2 -1 + 1 = 0

إذن انطبقت الملاحظة .. الناتج صفر ..

- 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 لهذه الملاحظة :

السؤال : لديك متوسط حسابي للبيانات (5) تم زيادة درجتين لكل قيمة من قيم البيانات الموجودة في هذه الدراسة ، فماذا سيحصل للمتوسط الحسابي ، هل سيرتفع أم يبقى كما هو ؟

الإجابة : سيرتفع .. لأنه يتأثر بجمع البيانات .

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

1. زيادة اجور العاملين بمقدار 2000 ريال
2. زيادة أجور العاملين بنسبة 5 %

الإجابه : حساب المتوسط ،، نطبق بالقانون $\overline{x} = \frac{\sum x}{n}$ << $\overline{x} = \frac{3+5+2+7+3}{5}$ << $\overline{x} = \frac{20}{5} = 4$

- زيادة اجور العاملين بمقدار 2000 ريال . يعني زيادة 2 للمفردات الظاهرة .

... 3+2 = 5 ,,,, 7+2 = 9 ,,,, 2+2 = 4 ,,,, 5+2 = 7 ,,,, 2+ 3 = 5

نجمع الناتج ليسهل تطبيقه بالقانون :: 30 = 5+7+4+9+5

نطبق بالقانون $6 = \frac{30}{5} = \overline{x}$ ،، إذن ناتج المتوسط الحسابي الجديد 6 .

ملاحظه : لاختصار الوقت والجهد نجيب على هذه الفقرة من السؤال بجمع الزيادة المذكورة بالمتوسط الحسابي الأصلي

تطبيق لكلامنا : 6 = 2 + 4 ،،، 6 الوسط الحسابي الجديد .. إذن تطابق ناتج الحل التفصيلي مع الحل المختصر .

- زيادة أجور العاملين بنسبة 5 % .

.. 3+5% = 3,15 ,,,, 7+5% = 7,35 ,,,, 2+5% = 2,1 ,,,, 5+5% = 5,25 ,,,, 3 + 5% = 3,15

نجمع كل النواتج : 21 = 3,15 + 7.35 + 2,1 + 5,25 + 3,15 ، ،،، ، نطبق بالقانون :

$$\overline{x} = \frac{21}{5} = 4,2$$

ولو طبقنا الطريقة المختصرة 4,2 = 5% + 4 إذن تطابق ناتج الحل التفصيلي مع المختصر .

مزايا وعيوب المتوسط الحسابي:

المزايا :

- يعد المتوسط الحسابي من اكثر مقاييس النزعة المركزية استخداما، واسهلها فهما وذلك نتيجة لسهولة حسابه .
- يدخل في حسابه كل القيم دون اهمال أي قيمة منها.

العيوب :

- 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 .

مثال توضيحي لهذه النقطة : عملنا حملة تبرع لجهه معينه فتبرع عدة أشخاص ، الأول تبرع بريال واحد ، الثاني تبرع بريال ، الثالث تبرع بريال ، الرابع تبرع بـ ٩٧ ريال .
أحسب المتوسط الحسابي وما القيمة الشاذة ؟

$$\bar{x} = \frac{1+1+1+97}{4} = 25$$

٩٧ تسمى القيمة الشاذة لأنها رفعت قيمة المتوسط الحسابي .

- 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:

- ترتيب الدرجات تصاعديا أو تنازليا
- 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
زوجي	يوجد ترتيبين هما $n/2$, $(n/2)+1$

- إيجاد قيمة الوسيط.

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:

إيجاد قيمة الوسيط للبيانات السابقة.

الحل نتبع الخطوات :

١- ترتيب الدرجات تصاعدياً أو تنازلياً .

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

٢- إيجاد ترتيب الوسيط .

الأرقام هنا زوجيه : نطبق قانون الأعداد ازوجية بصيغتين ،،

n\2 << توضيح n عدد المجموعات ،،، $6 = \frac{12}{2}$

أو n\2)+1) << 7 =(12\2)+1

٣- إيجاد قيمة الوسيط ،، الوسيط عرفناه من الخطوة الثانية أن الوسيط هو 7,6 من ترتيب الدرجات أي

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

1 , 2 , 3 , 4 , 5 , 6 , 7 , 8 , 9 , 10 , 11 , 12

الوسيط = $\frac{5+5}{2} = \frac{10}{2} = 5$ ،،،، إذن قيمة الوسيط 5 .

مثال للأرقام الفردية : 2 , 9 , 7 , 3 , 8 أحسب قيمة الوسيط لهذه البيانات ؟

١- ترتيب الدرجات . 9 , 8 , 7 , 3 , 2

٢- ترتيب الوسيط . نطبق قانون الأعداد الفردية ، n + 1) \ 2) << 3 = 2 \ (1 + 5)

٣- إيجاد قيمة الوسيط من الخطوة السابقة عرفنا أن ترتيب الوسيط 3

2 , 3 , 7 , 8 ,9

1 , 2 , 3 , 4 , 5

إن قيمة الوسيط = 7

مزايا وعيوب الوسيط:

المزايا:

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

العيوب:

- 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 في التوزيع ، أي التي تتكرر في التوزيع .

فى نفس المثال السابق للمبيعات الشهرية . أحسب المنوال؟

نجد أن المبيعات الأكثر تكراراً 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:

- سهل الحساب سواء بالرسم أو بالحساب
- لا يتأثر كثيرا بالقيم الشاذة
- لا يتأثر كثيرا لو تغيرت قيم بعض مفردات البيانات

العيوب:

- أقل مقاييس النزعة المركزية استعمالا
- عديم الفائدة في البيانات القليلة العدد

المحــاضـرة الثـامنـة

المقاييس الإحصائية للبيانات غير المبوبة

ثانيا: مقاييس التشتت أو الانتشار

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 التى تعبر عن مدى تباعد القيم أو تقاربها فى المجموعات التى 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:

- المدى
- المدى الربيعي
- الإنحراف عن المتوسط
- التباين
- الإنحراف المعياري

- لماذا نستخدم مقاييس التشتت؟

نستخدم هذه المقاييس اذا كان عندنا مجموعتين ونريد ان نقارن بينهما، وكان المتوسط فيما يبينهما 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)

كلما صغر التشتت كلما دل على أن المجموعة أكثر تجانس

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.

إذا كان هناك قيم شاذة فلنحذر من استخدام المدى لأنه يؤدي إلى نتائج خادعه

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 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 حسابه بناءا على أكبر و أصغر قيمتان 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 الذى يقيس تباعد كافة القيم عن المتوسط الحسابى .

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 :

$$\frac{\sum|X - \bar{X}|}{n}$$

توضيح للقانون فروق المجموعة لكل متوسط حسابي مقسوم على عدد المجموعات .

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:
أحسب متوسط الانحرافات المطلقة للمبيعات الشهرية .

الحل :

X	$\overline{X}$	$X - \overline{X}$	$\lvert X - \overline{X} \rvert$
3	5,75	-2,75	2,75
5	5,75	-0,75	0,75
8	5,75	+2,25	2,25
3	5,75	-2,75	2,75
6	5,75	+0,25	0,25
4	5,75	-1,75	1,75
12	5,75	+6,25	6,25
5	5,75	-0,75	0,75
4	5,75	- 1,75	1,75
3	5,75	- 2,75	2,75
7	5,75	+ 1,25	1,25
9	5,75	+3,25	3,25

للتوضيح نتائج الجدول : X المبيعات ..

$$5,75 = \frac{69}{12} = \frac{\sum X}{n} = \overline{X}$$

$$x - \bar{x} = 3 - 5,75 = -2,75$$

وتم تطبيق نفس الطريقة على باقي الجدول

$|x - \bar{x}|$ نزلنا نفس الأعداد ولكن بحذف الإشارات

نطبق بقانون متوسط افنحرافات المطلقة : $\frac{\sum|x-\bar{x}|}{\text{n}}$

$\sum|x - \bar{x}| = 26,5$ << نتيجة جمع x و $\bar{x}$ بالجدول = 26,5

n عدد الأشهر

نعوض بالقانون : $,21 = \frac{26,5}{12}$ ... إذن قيمة متوسط الانحرافات المطلقة = 2,21

التباين والانحراف المعياري:

التباين Variance هو متوسط مربعات انحرافات القيم عن وسطها الحسابى. ويرمز له بالرمز σ^2 (تقرأ 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 التربيعي لمتوسط مربعات انحرافات القيم عن وسطها الحسابى أى هو جذر التباين لذلك يرمز له بالرمز σ (تقرأ سيجما) و ذلك إذا كان محسوب لبيانات المجتمع أما فى حالة حسابة لبيانات S عينة من المجتمع فيرمز له بالرمز S .

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

فالتباين و الانحراف 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 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:

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

أو و هذا أسهل

$$S^2 = \frac{\sum(x - \bar{x})^2}{n}$$

وبالتالى يكون حساب الإنحراف المعياري كما يلي:

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

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 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 :

X	$\overline{X}$	$X-\overline{X}$	$(x-\overline{x})^2$
3	5,75	-2,75	7,56
5	5,75	-0,75	0,56
8	5,75	+2,25	5,06
3	5,75	-2,75	7,56
6	5,75	+0,25	0,06
4	5,75	-1,75	3,06
12	5,75	+6,25	39,06
5	5,75	-0,75	0,56
4	5,75	- 1,75	3,06
3	5,75	- 2,75	7,56
7	5,75	+ 1,25	1,56
9	5,75	+3,25	10,56

لتوضيح الجدول :

X المبيعات

$$\overline{X} = \frac{\sum X}{n} = \frac{69}{12} = 5,75$$

$$x-\overline{x} = 3 - 5,75 = -2,75$$

وتم تطبيق نفس الطريقة على باقي الجدول

$$\sum(x-\overline{x})^2 = 86,22$$

نطبق بالقانون : $s^2 = \frac{(x-\overline{x})^2}{n}$ >>>> $7,19 = \frac{86,22}{12} = s^2$ >> هذا التباين

الانحراف المعياري : $S = \sqrt{s^2}$ >> = $\sqrt{7,19}$ >> = 2,68

ملاحظة : عادة تكون قيمة الانحراف المعياري أصغر من التباين

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.

لأنه يهتم بالتباعد بين القيم فبتالي لو انطرحت قيمة من جميع القيم فلا يتأثر من ذلك وإذا اضيفت قيمة لجميع القيم فلا يتأثر من ذلك ...

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

يتم تنقيص 2 من القيم الأساسية للمبيعات حسب ما هو مطلوب بالسؤال ..

ونرمز له بـ X ونبدأ بترتيب الجدول وتطبيق الطرق السابقة ...

X	$\bar{X}$	X - $\bar{X}$	$(X-\bar{X})^2$
1	3,75	- 2,75	7,56
3	3,75	- 0,75	0,56
6	3,75	+ 2,25	5,06
1	3,75	- 2,75	7,56
4	3,75	+ 0,25	0,06
2	3,75	- 1,75	3,06
10	3,75	+ 6,25	39,06
3	3,75	- 0,75	0,56
2	3,75	- 1,75	3,06
1	3,75	- 2,75	7,56
5	3,75	+ 1,25	1,56
7	3,75	+ 3,25	10,56

الطرق بنفس الطرق السابقة ..

نطبق بالقانون : $S^2 = \frac{\sum(X-\bar{X})^2}{n}$ <<<< = $\frac{86,22}{12}$ <<<< = 7,19 << هذا التباين ..

الانحراف : S = $\sqrt{S^2}$ <<<< = $\sqrt{7,19}$ <<<< = 2,68

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 بالألف ريال كما يلي :

الشهر	محرم	صفر	ربيع أول	ربيع ثان	جمادى أول	جمادى الأخر	رجب	شعبان	رمضان	شوال	ذي القعدة	ذي 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 ؟

X	$\overline{X}$	X - $\overline{X}$	$(X-\overline{X})^2$
9	17,25	- 8,25	68,0625
15	17,25	- 2,25	5,0625
24	17,25	+ 6,75	45,5625
9	17,25	- 8,25	68,0625
18	17,25	+ 0,75	0,5625
12	17,25	- 5,25	27,5625
36	17,25	+ 18,75	351,5625
15	17,25	- 2,25	5,0625
12	17,25	- 5,25	27,5625
9	17,25	- 8,25	68,0625
21	17,25	+ 3,75	14,0625
27	17,25	+ 9,75	95,0625

نطبق بالقانون :

$$64,5 = \frac{776,25}{12} = \frac{\sum(x-\bar{x})^2}{n} = s^2$$ <<< هذا التباين

نطبق بقانون الانحراف المعياري :

$s = \sqrt{s^2}$ <<< = $s = \sqrt{64,5}$ = << 8,03 << هذا الانحراف المعياري

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

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

مقياس التشتت

المدى	متوسط الانحرافات المطلقة	التباين	الانحراف المعيارى
9	2.20833	7.840909	2.80016

مقياس النزعه المركزية

المتوسط	الوسيط	المنوال	الوسط الهندسي
5.75	5	3	5.20114

المحـــاضرة التـاسـعة

المقاييس الإحصائية للبيانات المبوبة

أولا: الوسط الحسابي والتشتت حوله

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

البيانات الغير مبوبة هي البيانات الخام التى نحصل عليها من وسائل جمع البيانات مباشرة دون عملية تبويب أو تصنيف لها .

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.

سيتم عرض لكيفية حساب كلا من مقايس النزعة المركزية ومقاييس التشتت فى ثلاث حالات للجداول التكرارية وهي:

- الجداول المنتظمة
- الجداول غير المنتظمة
- الجداول المفتوحة

الجداول المنتظمة : هي تلك الجداول التى تكون فيها أطوال الفئات جميعها متساوية .

أولاً : الوسط الحسابي والتشتت حوله :

الوسط الحسابي هو القيمة التى إذا أخذها جميع المفردات لكان مجموعها يساوى مجموع القيم الأصلية، ويمكن حساب الوسط الحسابي او المتوسط الحسابى كما يلي:

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

أو

$$\bar{x} = \frac{\sum \times f}{\sum \times f}$$

$\bar{x}$ الوسط الحسابي

l عدد الفئات

f_i تكرار الفئة i

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

ويتم حساب التشتت حول المتوسط الحسابى من خلال الآتي:

أ - متوسط الانحرافات المطلقة *AAD*:

وهو يقيس انحراف القيم عن وسطها الحسابى بغض النظر عن اشارة ذلك الانحراف حيث يتم حسابه من خلال المعادلة التالية : ملاحظة جميع القوانين متشابهه فقط وضحناها بصيغ مختلفة ..

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

أو

$$AAD = \frac{\sum |x - \bar{x}| f}{\sum f}$$

أو

$$AAD = \frac{\sum (f |x - \bar{x}|)}{\sum f}$$

ب - التباين σ^2 :

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

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

أو

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f}$$

أو

$$\sigma^2 = \frac{\sum (f (x - \bar{x})^2)}{\sum f}$$

ملاحظة جميع القوانين متشابهه فقط وضحناها بصيغ مختلفة ...

جـ - الانحراف المعياري σ :

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

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

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

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

المطلوب: حساب التالي:

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

الفئات	التكرارات f	مركز الفئة x	$x \times f$	الوسط الحسابي $\bar{x}$	$(x - \bar{x})$	$(x - \bar{x})^2$	$f\,(x - \bar{x})^2$	$(x - \bar{x})$f
20 -	10	25	250	42,27	- 17,27	298,25	298,25	172,7
30 -	30	35	1050	42,27	-7,27	52,58	1582,5	218,1
40 -	50	45	2250	42,27	+ 2,73	7,45	372,5	136,5
50 – 60	20	55	1100	42,27	+12,73	162,05	3241	254,6
	$\sum f = 110$		$\sum xf = 4650$		40 نجمعها بدون إشارات		8181,5	781,9

توضيح للجدول : الفئات متساوية إذن الجدول منتظم (كيف عرفنا أن الفئات متساوية ؟؟ نشوف الفرق بين ٢٠ و ٣٠ – ٣٠ و ٤٠ – ٤٠ و ٥٠)

الفئات والتكرارات نأخذها من الجدول و نطع مجموع التكرارات ،،

مركز الفئة نطبق عليه القانون :

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

$= \frac{40+30}{2} = 35 \qquad = \frac{30+20}{2} = 25$

$= \frac{60+50}{2} = 55 \qquad = \frac{50+40}{2} = 45$

$\bar{x} = \frac{\sum xf}{\sum f}$ الوسط الحسابي

$= \frac{4650}{110} = 42,27$

،، $(x - \bar{x})$

24,27 – 25 = -17,27

24,27 – 35 = - 7,27

24,27 – 45 = + 2,73

24,27 – 55 = + 12,73

،، $(x - \bar{x})^2$

$298,25 = (-17,27)^2$

$52,85 = (-7,27)^2$

$7,45 = (2,73)^2$

$162,05 = (12,73)^2$

(f) × (x)

250 = 10 × 25

1050 = 30 × 35

2250 = 50 × 45

1100 = 20 × 55

،،، $f\,(x - \bar{x})^2$

2982,5 = 298,25 × 10

1585,5 = 52,82 × 20

372,5 = 7,45 × 50

3241 = 162,05 × 20

،،، $(x - \bar{x})$f   172,7 = 17,27 × 10

218,1 = 7,27 × 20   ،،،،   136,5 = 2,73 × 50

254.6 = 12.73 × 20

.. التباين .. $\sigma^2 = \frac{\sum(f\,|x-\bar{x}|^2)}{\sum f}$

$74{,}38 = \frac{8181{,}5}{110}$

الانحراف المعياري : $\sigma = \sqrt{\sigma^2}$

$8{,}62 = \sqrt{74{,}38}$

متوسط الانحرافات المطلقة $AAD = \frac{\sum(f\,|x-\bar{x}|)}{\sum f}$

$\frac{781{,}9}{110} = 7{,}108$

المحاضـرة العـاشـرة

المقاييس الإحصائية للبيانات المبوبة

ثانيا: الوسيط والتشتت حوله

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

(أي أنه القيمة التي تقع في الوسط عند ترتيب البيانات تصاعدياً أو تنازلياً) ..

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

- إيجاد الجدول التكرارى المتجمع الصاعد
- إيجاد ترتيب الوسيط من خلال المعادلة التالية:

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

- إيجاد قيمة الوسيط من خلال المعادلة التالية:

$$Med = \mathrm{L}_{\mathrm{Med}} + \frac{k_{Med} - F_a}{F_b - F_a} \times I$$

حيث أن :

Med قيمة الوسيط ..

L_{Med} الحد الأدنى 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 اعمارهم ،

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

المطلوب: حساب قيمة الوسيط؟

الحل :

الخطوة الأولى : إيجاد الجدول التكراري المتجمع الصاعد .

الفئات	التكرارات F	التكرار المتجمع الصاعد
20 -	10	10
30 -	30	40
40 -	50	90
50 – 60	20	110
	$\sum f = 110$	

الخطوة الثانية : إيجاد ترتيب الوسيط . من خلال المعادلة $k_{Med} = n/2$

55 = 2 ÷ 110

الخطوة الثالثة : إيجاد قيمة الوسيط من خلال المعادلة . $Med = \mathrm{L}_{\mathrm{Med}} + \frac{k_{Med} - F_a}{F_b - F_a} \times I$

توضيح :

L_{Med} الحد الأدنى لبداية الفئة الوسيطية عرفنا أن الفئة الوسيطية المحدده بالبني بالجدول يعني قيمتها 40

K_{Med} ترتيب الوسيط عرفناه من الخطوة الثانية ، ومنه عرفنا الفئة الوسيطية 55 بين 90 – 40

$\mathrm{f_a}$ التكرار المتجمع السابق للفئة الوسيطية يعني 40

f_b التكرار المتجمع اللاحق للفئة الوسيطية يعني 90

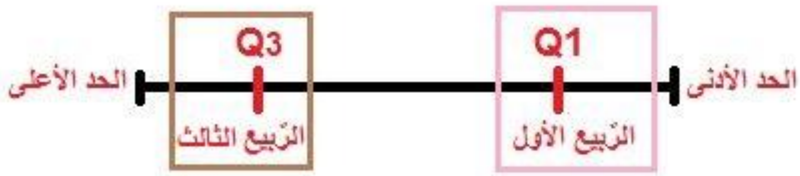
I طول الفئة الوسيطية يعني 10

نبدأ نعوض بالمعادلة .. $Med = 40 + \left(\frac{55-40}{90-40}\right) \times 10$

$Med = 40 + \frac{15}{50} \times 10$ <<< $Med = 40 + 130 \times 10$ <<< $Med = 40 + 3$

$\therefore Med = 43$

رسم توضيحي لحساب نصف المدى الربيعي :



الرُبيع الادنى (الأول):

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

لذلك يتم حسابه 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 الأعلى (الثالث):

يُعبر الرُبيع الثالث Q3 عن تلك 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)

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

يعني لو رسمنا رسم توضيحي ..

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

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

<u>مثال:</u> 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 الربيع الاول
- قيمة الربيع الثالث

الفئات	التكرارات f	التكرار المتجمع الصاعد (ك . م . ص)
20 -	10	10
30 -	30	40
40 -	50	90
50 – 60	20	110
	$\sum f = 110$	

قيمة الرُبيع الأول : الفئة محدده باللون البني بالجدول

$K_{Q_1} = n \div 4$ << 110 ÷ 4 = 27,5 .. نشوف موقعها بين الفئات ..

$L_{Q_1} = 10$ ،،،،، $I_{Q_3} = 10$ لأن الفئات متساوية ..

$F_a = 10$ ،،،،، $F_b = 40$ ،،،

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

نعوض بالقانون :

$$Q_1 = 10 + \left(\frac{27,5 - 10}{40 - 10}\right) \times 10 ..$$

$$Q_1 = 35,83$$

قيمة الرّبيع الثالث : الفئة محدده باللون الأزرق بالجدول ..

$K_{Q_3} = 3\,(n) \div 4$ << 3 (110) / 4 = 82,5 .. نشوف موقعها بين الفئات ..

$L_{Q_3} = 10$ ،،،،، $I_{Q_3} = 10$ لأن الفئات متساوية ..

$F_a = 10$ ،،،،، $F_b = 40$ ،،،

نعوض بالقانون :

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

$$Q_3 = 10 + \left(\frac{82,5 - 40}{90 - 40}\right) \times 10 \quad ،،، \quad Q_3 = 48,5 ..$$

حساب قيمة العُشير $P_{0.10}$:

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

$$P_{0,10} = \mathrm{L}_{P0,10} + \frac{K_{P0,10} - F_a}{F_b - F_a} \times I_{P_{0,10}} \qquad k_{P_{0.10}} = n/10$$

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؟

ك . م . ص .	التكرارات f	الفئات
10	10	20 -
40	30	30 -
90	50	40 -
110	20	50 - 60
	$\sum f = 110$	

$K_{P0.10} = n \div 10$ >> 110 ÷ 10 = >> 11 .. نشوف مكانها بالجدول (حددتها باللون الوردي)

$L_{P0,10} = 30$ ،،،، $F_a = 10$ ،،،، $F_b = 40$ ،،،، $I_{P_{0,10}} = 10$ لأن الفئات متساوية ....

نعوض بالقانون :

$$P_{0,10} = \mathrm{L}_{P0,10} + \frac{\frac{n}{10} - F_a}{F_b - F_a} \times I_{P_{0,10}}$$

$$P_{0,10} = 30 + \left(\frac{11-10}{40-10}\right) \times 10 \quad <<< \quad P_{0,10} = 30{,}33$$

حساب قيمة المئين $P_{0.01}$:

وبنفس الطريقة السابقة يمكن الحصول على المئويي $P_{0.01}$ وهو تلك 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 :

$$P_{0,01} = \mathrm{L}_{P0,01} + \frac{\frac{n}{100} - F_a}{F_b - F_a} \times I_{P_{0,01}} \qquad k_{P_{0.01}} = n/100$$

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؟

الفئات	التكرارات f	ك . م . ص
20 -	10	10
30 -	30	40
40 -	50	90
50 - 60	20	110
	$\sum f$ = 110	

$K_{P0,01} = n \div 100$ ،،، 1,1 = 100 ÷ 110 .. نشوف ترتيبها بالجدول محدده باللون الأحمر بالجدول

$L_{P0,01} = 20$ ،،،، $F_a = 0$ لماذا 0 لأننا فرضنا فئة سابقة للفئة الأولى تكرارها صفر في التكرار المتجمع الصاعد نستطيع افتراض فئة . الفئة تكون الفئات 10 ، التكرارات صفر ، التكرار المتجمع الصاعد صفر ،،،، $F_b = 10$ ،،،، $I_{P0,01} = 10$ لأن الفئات متساوية ....

نعوض بالقانون :

$$P_{0,01} = L_{P0,01} + \frac{\frac{n}{100} - F_a}{F_b - F_a} \times I_{P_{0,01}}$$

$P_{0,01} = 20 + \left(\frac{1,1 - 0}{10 - 0}\right) \times 10$ ،، $P_{0,01}$ = 21,1

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 الرُبيعي Inter Quartile Range :

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

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

مثال: 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؟

نطبق الخطوات السابقة :

التكرار المتجمع الصاعد (ك م ص)	التكرارات f	الفئات
10	10	20 -
40	30	30 -
90	50	40 -
110	20	50 – 60
	$\sum f = 110$	

قيمة الرُبيع الأول : الفئة محدده باللون البني بالجدول

$K_{Q_1} = n \div 2$ << 110 ÷ 4 = 27,5 .. نشوف موقعها بين الفئات ..

$L_{Q_1} = 10$ ،،،،، I_{Q_3} = 10 لأن الفئات متساوية ..

$F_a = 10$ ،،،،، F_b = 40 ،،،

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

نعوض بالقانون :

$$Q_1 = 10 + \left(\frac{27{,}5 - 10}{40 - 10}\right) \times 10 \;..$$

$$Q_1 = 35{,}83$$

قيمة الرّبيع الثالث : الفئة محدده باللون الأزرق بالجدول ..

$K_{Q_3} = 3\,(n) \div 4$ << 3 (110) / 4 = 82,5 .. نشوف موقعها بين الفئات ..

$L_{Q_3} = 10$ ،،،،، I_{Q_3} = 10 لأن الفئات متساوية ..

$F_a = 10$ ،،،،، F_b = 40 ،،،

نعوض بالقانون :

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

$$Q_3 = 10 + \left(\frac{82{,}5 - 40}{90 - 40}\right) \times 10 \;،،،\; Q_3 = 48{,}5 \;..$$

ثم نحسب نصف المدى الرُبيعي : نعوض بالقانون مباشرة ،،

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

$$IQR = \frac{48,5 - 35,83}{2} = \frac{12,67}{2} = 6,335$$

ثالثاً : المنوال :

المنوال هو تلك القيمة الأكثر شيوعا أو تكرارا. وفي حالة البيانات المبوبة يمكن حسابه 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: 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 المنوال ؟

بالبداية نطلع الفئة الأكثر تكراراً بالجدول التكرارات عدد العمال

الفئة (-40) تكرارها (50) هذي الفئة الأكثر تكراراً .. يعني 40 هي المنوال ..

تكرار الفئة السابقة لها 30 ،، وتكرار الفئة اللاحقة 20 ،،

نطلع D1 .. $D1$ يساوى تكرار فئة المنوال – تكرار الفئة السابقة

D1 = 50 – 30 = 20

نطلع D2 .. $D2$ يساوى تكرار فئة المنوال – تكرار الفئة اللاحقة

D2 = 50 – 20 = 30

نشوف طول الفئة المنوالية .. الفرق بين ٢٠ و ٣٠ .. ٣٠ و ٤٠ .. ٤٠ و ٥٠ .. ٥٠ و ٦٠ ..

الفرق بينهم 10 << I طول الفئة المنوالية = 10

L_{Mod} الحد الأدنى لفئة المنوال = 40

نطبق بالقانون :

$$Mod = L_{Mod} + \frac{D1}{D1 + D2} \times I \qquad Mod = 40 + \frac{20}{20 + 30} \times 10$$

$$Mod = 40 + \frac{20}{50} \times 10 \;<<\; Mod = 40 + 0,4 \times 10 \;<<\; Mod = 40 + 4 \;<<\; Mod = 44$$ المنوال = 44

الجداول غير المنتظمة:

وهي تلك الجدوال التى 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 التى سبق عرضها فى حالة الجداول المنتظمة بنفس الطريقة فيما عدا المنوال،

ويتعين علينا عند حساب المنوال تعديل التكرارات قبل حسابه 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 الموظفين	٢٠	٥٠	١٥	١٥

المطلوب حساب:

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

الفئات	التكرارات f	مركز الفئة x	X f	الوسط الحسابي	$(x-\bar{x})$	$(x-\bar{x})^2$	$f(x-\bar{x})^2$	$f(x-\bar{x})$	ك.م.ص
3 -	20	4	80	6,61	2,61	6,81	136,2	52,2	20
5 -	50	6,5	325	6,61	- 0,11	0,01	0,5	5,5	70
8 -	15	9	135	6,61	2,39	5,71	85,65	35,8	85
15 - 10	15	12,5	187,5	6,61	5,89	34,69	526,35	88,35	100
	$\sum f = 100$		$\sum xf = 727,5$		11	47,22	748,7	181,9	

١- الوسط الحسابي : نطبق القانون ،، $\bar{x} = \frac{\sum xf}{\sum f}$ ،، $\frac{727,5}{100} = 7,275$

٢- متوسط الانحرافات المطلقة : نطبق القانون ،، $AAD = \frac{\sum(f\,|x-\bar{x}|)}{\sum f}$ ،، $\frac{181,9}{100} = 1,82$

٣- التباين : نطبق القانون ،، $\sigma^2 = \frac{\sum(f\,|x-\bar{x}|^2)}{\sum f}$ ،، $\frac{748,7}{100} = 7,487$

٤- الانحراف المعياري : نطبق القانون ،، $\sigma = \sqrt{\sigma^2}$ ،، $\sqrt{7,489} = 2,74$

٥- الوسيط : اول خطوة إيجاد الجدول التكراري المتجمع الصاعد .. واضح بالجدول ،، ك.م.ص ..
الخطوة الثانية إيجاد ترتيب الوسيط من خلال المعادله : $k_{Med} = n/2$ >> 100 ÷ 2 = 50
الـ 50 موجودة بالفئة الـ 5- تكرارها المتجمع الصاعد 70

الخطوة الثالثة إيجاد قيمة الوسيط من خلال المعادلة $Med = L_{Med} + \frac{k_{Med} - F_a}{F_b - F_a} \times I$

$= 5 + \left(\frac{50-20}{85-20}\right) \times 2$ >> $5 + \left(\frac{30}{65}\right) \times 2$ >> $5 + 0,46 \times 2$ >> 5,92 = 0,92 + 5

٦- الرّبيع الأول : $K_{Q_1} = n \div 4$ >> 25 = 4 ÷ 100

نعوض بالقانون

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

$5 + \left(\frac{25-20}{70-20}\right) \times 2$ >> 5,2 = 2 × 0,1 + 5

٧- الرّبيع الثالث : $K_{Q_3} = 3(n) \div 4$ >> 75 = 4 ÷ 3 (100)

نعوض بالقانون :

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

$5 + \left(\frac{75-20}{70-20}\right) \times 2$ >> 7,2 = 5 + 2,2 = 5 + 1,1 × 2

٨- العشير : $K_{P0.10} = n \div 10$ >> 10 = 100 ÷ 10

نفرض فئة تكرارها صفر ..

2-	0	0

$I_{P0,10} = 2$ $\quad f_b = 0$,,, $\quad f_a = 0$,,,, $\quad L_{P0.10} = 2$,,,,,

نعوض بالقانون :

$$P_{0,10} = L_{P0,10} + \frac{\frac{n}{10} - F_a}{F_b - F_a} \times I_{P_{0,10}}$$

$2 + \left(\frac{10-0}{0-0}\right) \times 2 \gg 2 + \cdots \times 2 = \cdots$ لم استطع التوصل للحل الصحيح اعتذر

٩- المئين : $K_{P0,1} = 100 \div n$ >> 1 = 100 ÷ 100

نعوض بالقانون :

$$P_{0,01} = L_{P0,01} + \frac{\frac{n}{100} - F_a}{F_b - F_a} \times I_{P_{0,01}}$$

$2 + \left(\frac{1-0}{0-0}\right) \times 2 \gg 2 + \cdots \times 2 =$ لم استطع التوصل للحل الصحيح اعتذر

١٠- نصف المدى الربيعي : نعوض بالقانون : $IQR = \frac{Q3 - Q1}{2}$

$$\frac{7,2 - 5,2}{2} = \frac{2}{2} = 1$$

١١- المنوال : نطبق قانون المنوال

التكرار المعدل = التكرار الأصلى للفئة ÷ طول الفئة

طول الفئة :	التكرار المعدل :	نحدد الفئة المنواليه الأكثر تكراراً : هي 5
٣ - ٥ = ٢	٢٠ ÷ ٢ = ١٠	D1 = 16,67 – 10 = 6,67
٥ – ٨ = ٣	٥٠ ÷ ٣ = ١٦,٦٧	D2 = 16,67 – 7,5 = 9,17
٨ – ١٠ = ٢	١٥ ÷ ٢ = ٧,٥	$I = 3$
١٠ – ١٥ = ٥	١٥ ÷ ٥ = ٣	

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

$$5 + \left(\frac{6,67}{6,67+9,17}\right) \times 3 \gg 5 + \left(\frac{6,67}{15,84}\right) \times 3 \gg 5 + 0,42 \times 3 \gg 5 + 1,26 = 6,26$$

الجداول المفتوحة:

وهى ذلك النوع من الجداول التى يكون فيها الحد الادنى للفئة الاولى غير محدد أو الحد الأعلى للفئة الأخيرة غير محدد أو كلاهما. وفى هذا النوع من الجداول يصعب حساب الوسط الحسابى والتباين والانحراف المعياري حيث لا يمكن تحديد مركز الفئة للفئات المفتوحة ، لذا فيعتبر 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 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 ؟

نوجد الوسيط :

اول خطوة نوجد الجدول التكراري المتجمع الصاعد ..

التكرار المتجمع الصاعد	التكرارات f	الفئات
5	5	50 -
15	10	60 -
50	35	70 -
65	15	80 -
75	10	80 فأكثر
	$\sum f = 75$	

ثانياً إيجاد ترتيب الوسيط $k_{Med} = n/2$ >> 37,5 = 75 ÷ 2

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

$$60 + \frac{37,5 - 15}{50 - 15} \times 10 \gg 60 + 0,64 \times 10 \gg 60 + 6,4 = 66,4$$

الربيع الأول : $K_{Q_1} = n \div 4$ >> 18,75 = 75 ÷ 4

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

$$60 + \frac{18,75 - 15}{50 - 15} \times 10 \gg 60 + 0,11 \times 10 \gg 1,1 + 60 = 61,1$$

الربيع الثالث : $K_{Q_3} = 3(n) \div 4$ ‹‹ 56,25 = 225 ÷ 4 = 3(75) ÷ 4

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

$$70 + \frac{56,25-50}{65-50} \times 10 \gg 70 + 0,42 \times 10 \gg 70 + 4,2 = 74,2$$

نصف المدى الربيعي : نطبق بالقانون مباشرة .. $$IQR = \frac{Q3 - Q1}{2}$$

$$\frac{47,2 - 61,1}{2} = 6,95$$

المحـــاضـرة الحاديـة عشـر

مقاييس التشتت النسبى والدرجة المعيارية

سيتم فى هذه المحاضرة استعراض كلا من:

- مقاييس التشتت النسبى
- القيمة المعيارية

أولاً – 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:

تستخدم في حال وجود المتوسط الحسابي و الانحراف المعياري $c.v.=\frac{s}{\bar{x}}\times 100$

تستخدم في حال وجود الرُبيعات (الربيع الأول والثاني) $c.v.=\frac{Q_3-Q_1}{Q_3+Q_1}\times 100$

للتذكير :

معادلة حساب الربيع الأول Q1 $$Q_1=L_{Q_1}+\frac{\frac{n}{4}-F_a}{F_b-F_a}\times I_{Q_1}$$

معادلة حساب الربيع الثالث Q1 $$Q_3=L_{Q_3}+\frac{\frac{3(n)}{4}-F_a}{F_b-F_a}\times I_{Q_3}$$

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

الفئات	التكرارات f	X	X f	$\bar{x}$	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$	ك . م . ص
6 -	15	8	120	11,73	- 3,73	13,91	208,65	15
10 -	20	11	220	11,73	-0,73	0,533	10,66	35
12 -	12	13	150	11,73	-1,27	1,81	19,32	47
14 - 18	13	16	208	11,73	-4,27	18,23	236,99	60
	$\sum f = 60$		$\sum xf = 704$				475,62	

ملاحظة : النواتج بالجدول طبقنا عليها القوانين السابقة ،، بنفس طريقة إيجاد الجداول السابقة

بالبداية نحسب التباين بهذا القانون : $S^2 = \frac{\sum f(x-\bar{x})^2}{\sum f}$

نعوض ،، $\frac{475,62}{60} = 7,93 = S^2$

الانحراف المعياري : $S = \sqrt{S^2}$ << $S = \sqrt{7,93}$ << S = 2,82

- نحسب معامل الاختلاف للإيجار السنوي .

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

$c.v = \frac{2,82}{11,73} \times 100 = 24\%$

∴ معامل الاختلاف للإيجار السنوي يبلغ ٢٤%

- معامل الاختلاف الربيعي للإيجار السنوي .

أولا نوجد الربيع الأول والثالث ،،

نوجد الربيع الأول . $K_{Q1} = \frac{n}{4} = 60 \div 4 = 10$ .. نوجد مكانها بالجدول بـ ت . م . ص ... قيمة Q1 = 10

ملاحظة مهمه :

عندما تكون قيمة التكرار المتجمع الصاعد مطابقة لترتيب الرُبيع فإن الحد الأعلى للفئة الرُبيعية تمثل قيمة الربيع المطلوب

نوجد الرُبيع الثالث . $K_{Q1} = \frac{3(n)}{4} = \frac{3(60)}{4} = 45$

نوجد مكانها بالجدول بـ ت . م . ص .. $L_{Q3} = 12$ ... $F_b = 47$ .... $F_a = 35$ ..... $I_{Q3} = 2$

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

$$Q_3 = 12 + \frac{45-35}{47-45} \times 2 \gg Q_3 = 13,49$$

نعوض بالقانون لنوجد معامل الاختلاف الرُبيعي .. $c.v. = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$

$C.V = \frac{13,67-15}{13,67+15} \times 100 \gg C.V = 15,49$

ويتضح لنا من الحل السابق أن:

- معامل الاختلاف للإيجار السنوي للوحدات السكنية بلغ ٢٤%
- معامل الاختلاف الربيعي للإيجار السنوى للوحدات السكنية بلغ ١٥.٤٩٤%

ونلاحظ وجود اختلاف بين قيمتى معامل الاختلاف باستخدام كلا من المعادلة الأولى والثانية وذلك لاختلاف الأساس الرياضى فى كل من التعريفين المعادلتين. إلا أنة 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:

X الدرجة المتحصل عليها

$\overline{X}$ المتوسط الحسابي

S الانحراف المعياري

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

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

المطلوب:

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

X المحاسبة = 80 .. $\overline{X}$ للمحاسبة = 83 .. S للمحاسبة = 5

نعوض بالقانون : $z = \frac{x - \overline{x}}{S}$

$= \frac{80-83}{5} = \frac{-3}{5} = -0,6$

X الرياضيات = 76 .. $\overline{X}$ للرياضيات = 65 .. S للرياضيات = 5

نعوض بالقانون : $z = \frac{x - \overline{x}}{S}$

$= \frac{70-65}{5} = \frac{5}{5} = 1$

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 .

المحـاضـرة الثانية عشرة

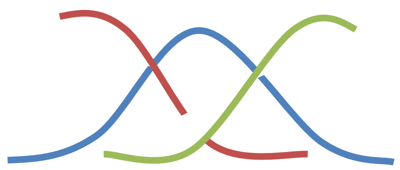
مقاييس الالتواء والتفلطح

أولا : مقاييس الالتواء Skewdness Measures

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

مقاييس الالتواء هي تلك الأشكال التي تأخذها التوزيعات المختلفة للبيانات عندما يكون هناك اتجاه سواء إلى الالتواء السالب او الالتواء الموجب ، بدلا من أن يكون التوزيع طبيعي أو متماثل .

وعادة يكون التوزيع طبيعي ..



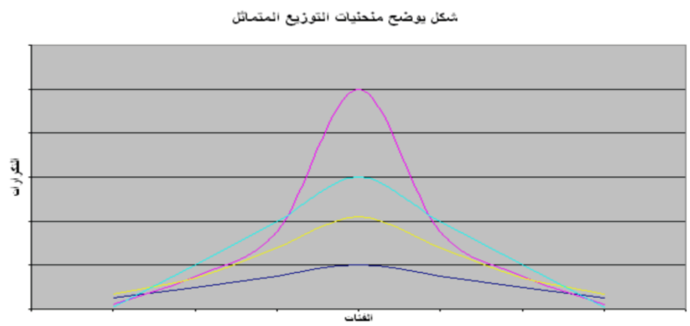
اللون الأزرق متماثل

اللون الأحمر التواء موجب

اللون الأخضر التواء سالب

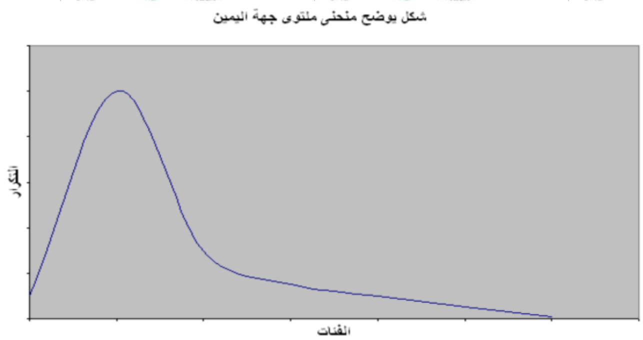
المنحنى المتماثل Symmetrical Curve

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



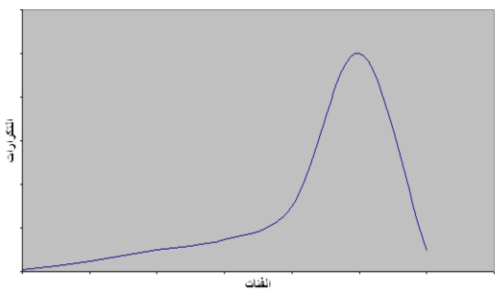
المنحنيات الملتوية Skewed

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



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

شكل يوضح منحنى ملتوى جهة اليسار



ويمكن قياس الالتواء من خلال معامل الالتواء SK والذى يفيدنا فى الحكم على مدى تماثل أو التواء التوزيع

تتعدد مقاييس الالتواء إلا أن من أهمها:

معامل الالتواء لبيرسون والذى يكون فى أحد الصورتين التاليتين:

$$SK = \frac{\bar{x} - Mod}{S} \quad \text{أو} \quad SK = \frac{3(\bar{x} - Med)}{S}$$

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

لذلك يمكن الاعتماد على مقياس الالتواء لباولى SkB الذى يعرف كما يلي :

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

للتذكير :

الوسيط **Med**:
$$Med = \mathrm{L}_{\mathrm{Med}} + \frac{k_{Med} - F_a}{F_b - F_a} \times I_{Med}$$

الربيع الأول **Q1**:
$$Q_1 = \mathrm{L}_{Q_1} + \frac{\frac{n}{4} - F_a}{F_b - F_a} \times I_{Q_1}$$

الربيع الثالث **Q3**:
$$Q_3 = \mathrm{L}_{Q_3} + \frac{\frac{3(n)}{4} - F_a}{F_b - F_a} \times I_{Q_3}$$

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.

لنفس هذا المثال بالمحاضرة السابقة حسبنا المتوسط الحسابي و الانحراف المعياري والرُبيع الأول والرُبيع الثالث و الوسيط ،، نأخذ الناتج الأخير لها ..

$$\bar{x} = 11,73 \quad ,,, \quad Med = 11,5 \quad ,,, \quad Q_1 = 10 \quad ,,,$$
$$Q_3 = 13,67 \quad ,,, \quad S = 2,82$$

نطبق مباشرة في المعادلة من خلال استخدام الوسيط ،، $SK = \frac{3(\bar{x} - Med\ )}{S}$

نعوض .. $\frac{3(11,73-11,5)}{2,82} = 0,249$

لنطبق القانون الآخر يجب ان نعرف المنوال ..

قانون المنوال $Mod = L_{Mod} + \frac{D_1}{D_1 + D_2} \times I_{Mod}$

نجد أن الجدول غير منتظم .. لذلك نستخدم التكرار المعدل ،، نأخذ كل تكرار ونقسمه على طول الفئة ..

$15 \div 4 = 3,75$,, $20 \div 2 = 10$,, $12 \div 2 = 6$,, $13 \div 4 = 3,25$

أكبر تكرار اللي هو 10 وبكذا يكون الفئة المنواليه ..

نطلع D1 .. $D1$ يساوى تكرار فئة المنوال – تكرار الفئة السابقة

D1 = 3,75 – 10 = 6,25

نطلع D2 .. $D2$ يساوى تكرار فئة المنوال – تكرار الفئة اللاحقة

D2 = 6 – 10 = 4

نطبق بقانون المنوال ،، $Mod = 10 + \frac{6,25}{6,25 + 4} \times 2 \gg 11,22$

نطبق مباشرة بالمعادلة من خلال استخدام المنوال .. $SK = \frac{\bar{x} - Mod}{S}$

$$= \frac{11,73 - 11,22}{2,82} = 0,18$$

ويظهر لنا من النتيجة لجميع المعادلات الخاصة بحساب معامل الالتواء وجود التواء موجب جهة اليمين إلا أن قيمة معامل الالتواء 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 الالتواء تختلف. إلا أنة كما سبق وذكرنا بأنه يفضل استخدام معامل الالتواء لبيرسون فى أي من صيغتيه فى حالة البيانات غير المبوبة وكذلك الجداول التكرارية المغلقة أما في حالة الجداول التكرارية المفتوحة فيفضل استخدام معامل الالتواء لباولي.

لو طبقنا قانون لباولى SkB

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

$$= \frac{13{,}67 - 2\,(11{,}5) + 10}{13{,}67 - 10} = -0{,}235$$

هذا القانون يستخدم للجداول التكرارية المفتوحة التي لا نستطيع أن نحسب من خلالها طول الفئة ..

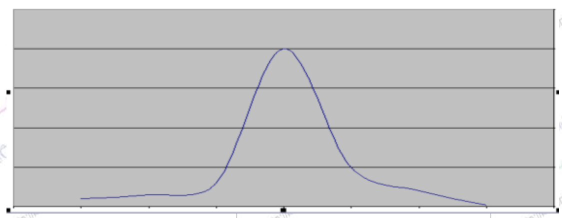
ثانيا: التفلطح Kurtosis

يقصد بالتفلطح مقدار التدبب (الارتفاع أو الانخفاض) فى قمة المنحنى مقارنة بقمة منحنى التوزيع الطبيعي.

وتكون قيمة معامل التفلطح صفر فى حالة التوزيع الطبيعى المعياري .

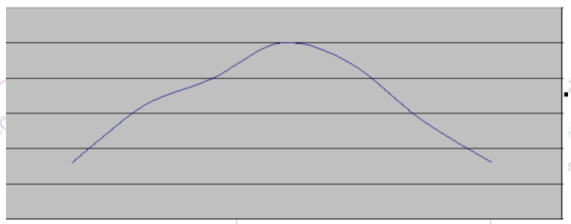
ففى حالة ما يكون معامل التفلطح للبيانات الاصلية أكبر من ٣ يكون المنحنى مدبب لأعلى كما بالشكل التالي:

شكل يوضح المنحنى المدبب



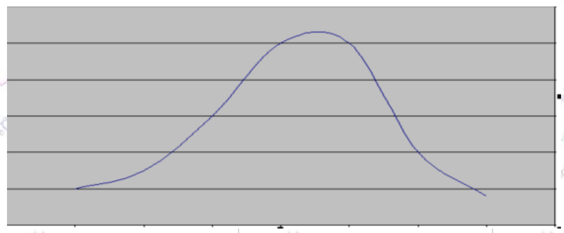
أما فى حالة ما يكون معامل التفلطح للبيانات الأصلية أقل من ٣ يعنى ذلك أن المنحنى مفلطح كما يتضح من الشكل التالي:

شكل يوضح المنحنى المفلطح



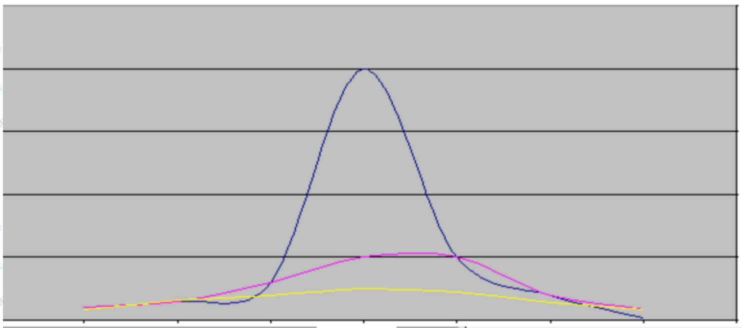
أما فى حالة ما يكون معامل التفلطح يساوى ثلاثة يكون المنحنى متوسط التفلطح و يكون بالشكل التالي :

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.

نحسبه من خلال هذا القانون $KU = \frac{Q_3 - Q_1}{2(P_{0.90} - P_{0.10})}$

الرُبيع الأول والرُبيع الثالث طلعنا قيمها بالأمثلة السابقة $Q_1 = 10$,, $Q_3 = 13,67$

بقي لنا $P_{0,10}$ ونطبق عليه هذا القانون $P_{0,10} = L_{P0,10} + \frac{K_{P0,10} - f_a}{f_b - f_a} \times I_{P0,10}$

وأيضا بقي لنا $P_{0,90}$ ونطبق عليه هذا القانون $P_{0,90} = L_{P0,90} + \frac{K_{P0,90} - f_a}{f_b - f_a} \times I_{P0,90}$

الـ $K_{P0,10} = n \div 10$ **وأيضاً الـ** $K_{P0,90} = (n \times 9) \div 10$

نوجد أولا الجدول التكراري

الفئات	التكرارات	ت . م . ص
6 -	15	15
10 -	20	35
12 -	12	47
14 – 18	13	60
	$\sum f = 60$	

نعوض $K_{P0,10} = n \div 10$ << $K_{P0,10} = 60 \div 10 = 6$

و الـ $K_{P0,90} = (n \times 9) \div 10$ << $K_{P0,90} = (60 \times 9) \div 10 = 540 \div 10 = 54$

نبدأ نعوض بالقوانين ،،

$$P_{0,10} = L_{P0,10} + \frac{K_{P0,10} - f_a}{f_b - f_a} \times I_{P0,10}$$

$$P_{0,10} = 6 + \frac{6 - 0}{15 - 0} \times 4 = 7,6$$

$$P_{0,90} = L_{P0,90} + \frac{K_{P0,90} - f_a}{f_b - f_a} \times I_{P0,90}$$

$$P_{0,90} = 14 + \frac{54 - 47}{60 - 47} \times 4 = 16,15$$

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

$$KU = \frac{13,67 - 10}{2(16,15 - 7,6)} = \frac{3,67}{2(8,55)} = \frac{3,67}{17,1} = 0,215$$

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 .

الـمـحـاضـرة الـثـالـثـة عشر

تحليل الإرتباط - ١

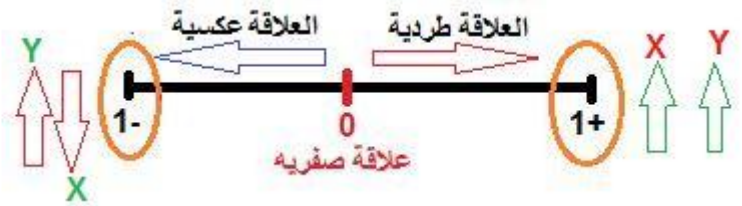
بالرغم من أن مقاييس العلاقة تختلف عما سبقها من مقاييس ، فهي تتعلق بدراسة العلاقة بين متغيرين (الإنتاجية والجودة) مثلا، بينما المقاييس السابقة فتهتم بدراسة الفروق بين المتغيرات .

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) .

معظم الأحيان يكون معامل الارتباط كسر بمعنى الحصول على ارتباط تام نادر جدا ،، إذا كان تام وصل إلى +1 و -1 أي يكون عدد صحيح .. هذا يعطينا مؤشر أنه عندما يكون ناتج معامل الارتباط أكبر من +1 أو أصغر من –1 يكون الحل خطأ ..

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 أو نادر الوجود 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
طرديه كاملة	١+
طرديه 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

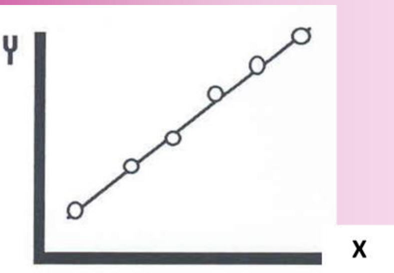
أولا: طريقة شكل الانتشار Scatter Diagram :

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.

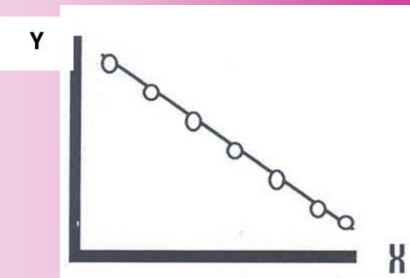
والمقصود بشكل الانتشار هو تمثيل قيم الظاهرتين بيانياً على المحورين، المتغير الأول X على المحور الأفقي والمتغير الثانى Y على المحور الرأسي، حيث يتم تمثيل كل زوج Pair من القيم بنقطه ، فنحصل على شكل يمثل كيفية انتشار القيم على المستوى، وهو الذي يسمى شكل الانتشار. وطريقة انتشار القيم تدل على وجود أو عدم 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.



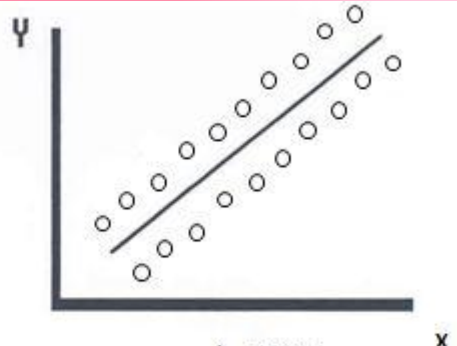
الشكل الأول (أ) ارتباط طردي تام (موجب)



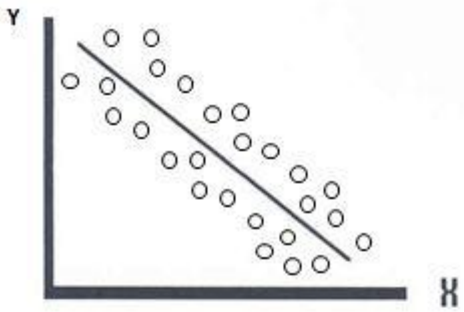
الشكل الأول (ب) ارتباط عكسي تام (سالب)

الشكل الثانى :

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.



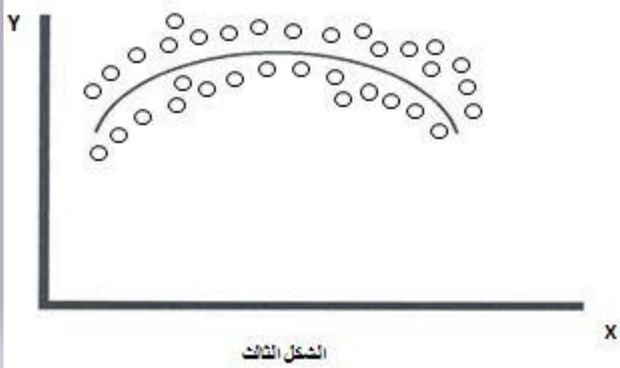
الشكل الثاني (أ)
ارتباط موجب قوي
(ارتباط خطي طردي)



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

الشكل الثالث :

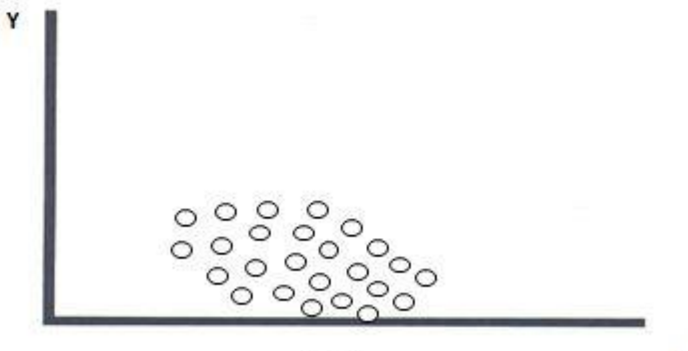
وإذا كانت العلاقة تأخذ شكل منحنى فإن الارتباط لا يكون خطياً "ارتباط غير خطى" Non Linear Correlation كما في الشكل الثالث :



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)

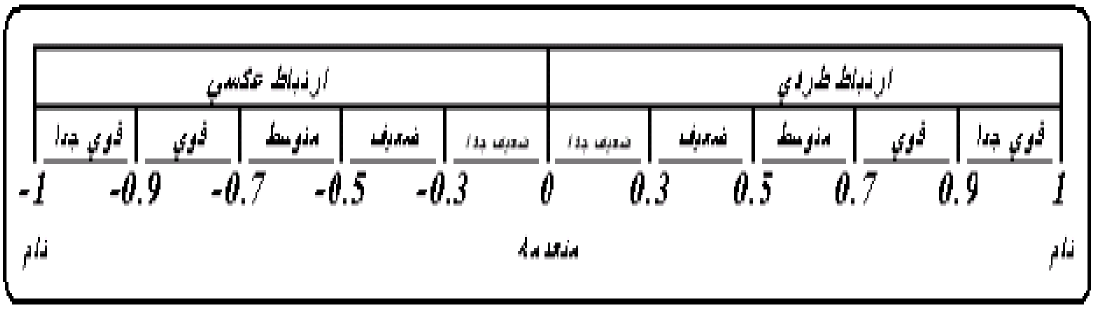
ثانيا: معامل الارتباط Correlation Coefficient :

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.**

الخلاصة :

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:

تستخدم هذه المعادلة من خلال إيجاد المتوسطات

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

وكذلك المعادلة الرياضية التالية والتي تعتبر اسهل وابسط:

تستخدم هذه المعادلة من خلال الدرجات الخام

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

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 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 عكسي قوى
-0.21	ارتباط عكسي ضعيف جدا
0.43	ارتباط طردى ضعيف
1 +	ارتباط طردى تام
-0.51	ارتباط عكسي متوسط

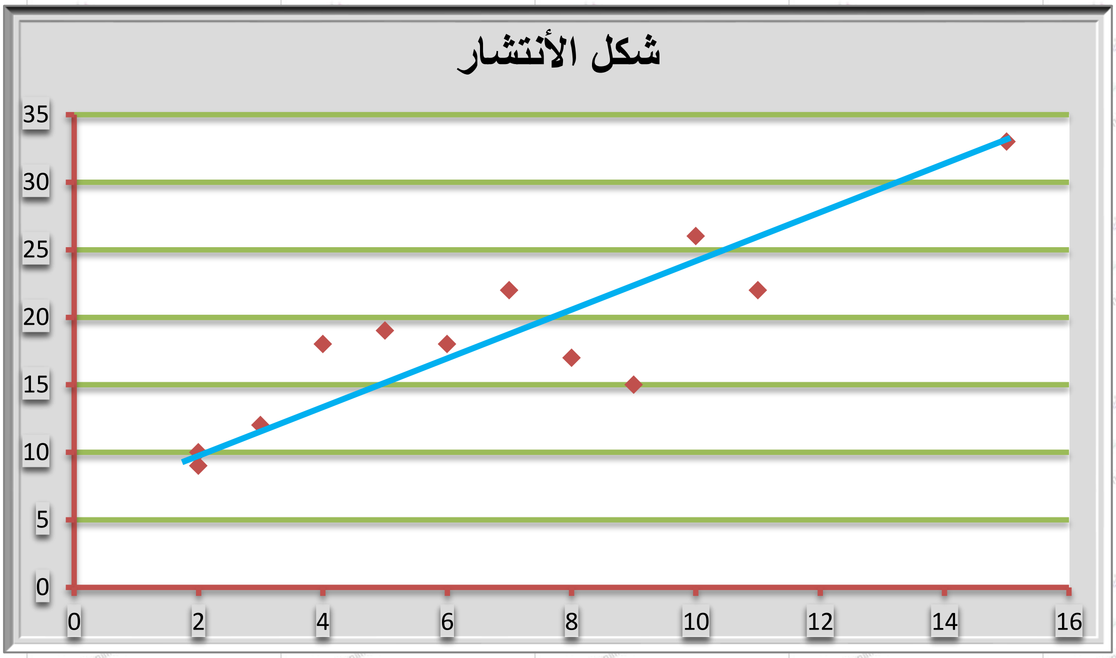
مثال: فيما يلي بيان بالمنفق على الاعلان والمبيعات لأحد المنتجات فكانت بالمليون ريال كمايلي:

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



١- نستنتج بأن المؤشر يتجه تصاعدياً جهة اليمين مما يدل على وجود علاقة طرديه وقوية إلى حد ما ..

٢- نحسب معامل الارتباط ،، بداية نصلح الجدول ..

بيانات X المنفق	بيانات Y المبيعات	X Y	X^2	Y^2
2	10	20	4	100
3	12	36	9	144
2	9	18	4	81
7	22	154	49	484
6	18	108	36	324
5	19	95	25	361
10	26	260	100	676
15	33	495	225	1089
4	18	72	16	324
11	22	242	121	484
9	15	135	81	225
8	17	136	64	289
82	221	1771	734	4581

نطبق بالقانون :

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

$$r_P = \frac{12\,(1771) - \,[(82)(221)]}{\sqrt{12\,(734) - \,(82)^2}\,\sqrt{12\,(4581) - \,(221)^2}}$$

$$= \frac{21252 - 18122}{\sqrt{8808 - 6724}\,\sqrt{54972 - 48841}} = \frac{3130}{\sqrt{2084}\,\sqrt{6131}}$$

$$= \frac{3130}{(45,65)\,(78,30)} = \frac{3130}{3574,395} = 0,87567$$

التعليق على النتيجة : توجد علاقة طرديه بين المنفق على الإعلان وبين المبيعات

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 : فى بيانات المثال السابق إذا اكتشفت 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.

أولا نصلح الجدول نفس السابق ولكن نضيف 5 لجميع القيم المنفق .. أما المبيعات مضاعفة القيم ..

بيانات X المنفق	بيانات Y المبيعات	X Y	X^2	Y^2
7	20	140	49	400
8	24	192	64	576
7	18	126	49	324
12	44	528	144	1936
11	63	396	121	1296
10	38	380	100	1444
15	52	780	225	2704
20	66	1320	400	4356
9	36	324	81	1296
16	44	704	256	1936
14	30	420	196	900
13	34	442	169	1156
142	442	5752	1854	18324

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

$$r_p = \frac{12\,(5752) - (142 \times 442)}{\sqrt{12\,(1854) - (142)^2}\ \sqrt{12\,(18324) - (442)^2}}$$
$$= \frac{69024 - 62764}{\sqrt{22248 - 20164}\ \sqrt{219888 - 195364}} = \frac{6260}{\sqrt{2084}\ \sqrt{24524}}$$
$$= \frac{6260}{45,65 \times 156,60} = \frac{6260}{7148,79} = 0,87567$$

معامل التحديد Determination Coefficient

وهو مربع معامل الارتباط لذلك يرمز له بالرمز R^2 أي 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 .

المــحاضــرة الرابـعة عشـر

تحليل الإرتباط – ٢

معامل ارتباط الرتب لسبيرمان Spearman's Rank Correlation Coefficient

معامل الارتباط لبيرسون r_s لا يمكن استخدامه فى حساب قوة العلاقة بين متغيرين الا اذا كانت البيانات المتوافرة 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 في حالة المتغيرات الوصفية فنستخدم معامل ارتباط الرتب لـ سبيرمان ، والذى يتم استخدامه فى قياس الارتباط خاصة فى حالة البيانات الوصفية الترتيبية مثل تقديرات الطلاب (ممتاز – جيد جدآ – جيد – مقبول – ضعيف) وكذلك قوة المركز المالى (جيد - متوسط - ضعيف) ودرجة الموافقة على الرأى فى اسئلة الإستبانه (موافق تماما – 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 المتغير

n عدد المشاهدات

ملاحظات يجب مراعاتها عند ترتيب المتغيرات:

- 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 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 وهكذا .
- 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 ؟

ملاحظه : تم تعديل الحل وفقاً للمثال لأن الدكتور أخطأ في أحد الأعداد واستبدله بعدد آخر . فغير مسار الحل جميعه ..

X	Y	رتبت X	رتبت Y	d	d^2
١ ← 2	٢ ← 10	1,5	2	-0,5	0,25
٣ ← 3	٣ ← 12	3	3	0	0
٢ ← 2	١ ← 9	1,5	1	+0,5	0,25
٧ ← 7	٩ ← 22	7	0,86	6,14	37,69
٦ ← 6	٦ ← 18	6	0,72	5,28	27,88
٥ ← 5	٨ ← 19	5	8	-3	9
١٠ ← 10	١١ ← 26	10	11	-1	1
١٢ ← 15	١٢ ← 33	12	12	0	0
٤ ← 4	٧ ← 18	4	0,72	3,28	10,76
١١ ← 11	١٠ ← 22	11	0,86	10,14	102,81
٩ ← 9	٤ ← 15	9	4	5	25
٨ ← 8	٥ ← 17	8	5	+3	9
					223,69

ملاحظه : الأعداد اللي باللون الأحمر طريقة تحديدنا للرتب ..

الفئات المتشابهه في المنفق والمبيعات طبقنا عليها هذا القانون مجموع الرتب ÷ عددها

مثلاً 3 ÷ 2 = 1,5 الناتج نضعه ترتيب لتلك الأعداد المتشابهه ..

وكذلك 19 ÷ 22 = 0,86

وأيضاً 13 ÷ 18 = 0,72

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

$$r_s = 1 - \frac{6\,(223,69)}{12\,(12^2-1)} = 1 - \frac{1342,14}{1716} = 0,782$$

قيمة عامل ارتباط سبيرمان (لو لا حظتوا الناتج لا يختلف عن الناتج اللي حسبه الدكتور)

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.

نرسم الجدول ونوزع الدرجات من 0 إلى 10 (الحل من الكتاب)

المحاسبة	القانون	رتب المحاسبة	رتب القانون	d	d^2
ممتاز	جيد	10	4,5	5,5	30,25
جيد جيدا	جيد	8,5	4,5	4	16
جيد	مقبول	6	1,5	4,5	20,25
مقبول	جيد	3	4,5	-1,5	2,25
ضعيف	جيد جدا	1	8	-7	49
جيد	جيد جدا	6	8	-2	4
مقبول	ممتاز	3	10	-7	49
جيد جدا	مقبول	8,5	1,5	7	49
جيد	جيد	6	4,5	1,5	2,25
مقبول	جيد جدا	3	8	-5	25
				0	247

- نلاحظ عند ترتيب تقديرات مقرر المحاسبه أن التقدير مقبول اخذ الرتب ٢ و ٣ و ٤ لذلك تم وضع ٣ أمام التقدير مقبول في مقرر المحاسبة .

- كما أن تقدير جيد في مقرر القانون أخذ الرتب ٣ و ٤ و ٥ و ٦ لذلك تم وضع الرتبه ٤.٥ امام التقدير جيد في مقرر القانون .

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

$$r_s = 1 - \frac{6\,(247)}{10\,(10^2-1)} = 1 - \frac{1482}{10\,(100-1)} = 1 - \frac{1482}{990} = 1 - 1,4969 = -0,4969$$

معامل الاقتران Conjunction Coefficient

ويستخدم معامل الاقتران فى حساب العلاقة الإرتباطية 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	أمي
يعمل	113	23
لا يعمل	49	15

المطلوب:

أحسب معامل الاقتران ؟

نطبق بالقانون : $r_C = \dfrac{AD - BC}{AD + BC}$

$$r_c = \frac{[(113)\ (15)] - [(23)\ (49)]}{[(113)\ (15)] + [(23)\ (49)]} = \frac{1695 - 1127}{1695 + 1127} = \frac{568}{2822} = +0,20$$

يظهر لنا في حل المثال أن العلاقة ضعيفة جداً بين العمل والتعليم ..

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

أي يتم إيجاد: مربع تكرار كل خلية مشتركة

مجموع الصف × مجموع العمود

ثم نجمعهم كلهم

وعلى ذلك يتم حساب معامل التوافق كما يلي : $r_T = \sqrt{\dfrac{M-1}{M}}$

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

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

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

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

$$M = \frac{900}{5400} + \frac{225}{4500} + \frac{2025}{6300} + \frac{400}{4200} + \frac{900}{3500} + \frac{400}{4900} + \frac{100}{1200} + \frac{25}{1000} + \frac{25}{1400}$$

$$M = 0,166 + 0,05 + 0,32 + 0,095 + 0,257 + 0,081 + 0,083 + 0,025 + 0,017$$

$$M = 1,094$$

أوجدنا قيمة M نبدأ نعوض بالقانون : $r_T = \sqrt{\frac{M-1}{M}}$

$$r_T = \sqrt{\frac{1,094 - 1}{1,094}} = \sqrt{\frac{0,094}{1,094}} = \sqrt{0,0859} = +\ 0,293$$

نجد ان هناك علاقة ضعيفة إلى حد ما بين التخصص و الرضا عن الدراسة ..

تـمــت بـحـمـد الله

نعتذر عن التقصير والأخطاء ،، وأي ملاحظه نسعد بتقبلها ،، فهذا جهد شخصي قد لا يخلو من الخطأ ..

نتمنى لكم التوفيق و الاستفادة من هذه الملزمه ،، و لا نريد منكم غير الدعاء ..

شرح وتنسيق .. oOo al maznei oOo .. حلم المشاعر ..