المملكة العربية السعودية

جامعة الملك فيصل بالإحساء

كلية الاداب

مبادئ الاحصاء

كلية الاداب – المستوى الثاني

دعاء قبل المذاكرة:

اللهم انى أسألك فهم النبيين وحفظ الملائكة المقربين اللهم اجعل لسانى عامرا بذكرك وقلبى بخشيتك وسرى بطاعتك انك على كل شيء قدير وحسبنا الله ونعم الوكيل.

ادعية اخرى تعين على تيسير المذاكرة:

رب اشرح لى صدرى ويسر لى امرى واحلل عقدة من لسانى يفقهوا قولى. اللهم لا سهل الا ما جعلته سهلا وأنت تجعل الحزن ما شئت سهلا .

ربنا أتنا من لدنا رحمة وهيئ لنا من أمرنا رشدا.

دعاء بعد المذاكرة:

اللهم انى استودعك ما قرأت وما حفظت وما فهمت فرده عند حاجتى اليه انك على كل شىء قدير وحسبنا الله ونعم الوكيل.

عند التوجه للامتحان:

اللهم انى توكلت عليك وسلمت امرى اليك لا ملجأ ولا منجى منك الا اليك.

عند الدخول للجنة الامتحان:

رب ادخلنى مدخل صدق واخرجنى مخرج صدق واجعل لى من لدنك سلطانا نصيرا.

عند بداية الاجابة:

رب اشرح لى صدرى ويسر لى امرى وأحلل عقدة من لسانى يفقهوا قولى بسم الله الفاتح..اللهم لا سهل الا ما جعلته سهلا يا أرحم الراحمين.

عند تعسير الاجابة:

لا اله الا انت سبحانك انى كنت من الظالمين يا حي يا قيوم برحمتك أستغيث.رب انى مسنى الضر وأنت ارحم الراحمين.

عند النسيان:

اللهم يا جامع الناس ليوم لا ريب فيه اجمع على ضالتى.

عند النهايه:

الحمد لله الذى هدانا لهذا وما كنا لنهتدى لولا ان هدانا الله.

المحـــــــــــاضره الاولى

(علم الإحصاء ودوره في خدمة المجتمع)

ـ البحث العلمي :

إن الغرض من العلم هو البحث عن الحقيقة، وأن البحث العلمي هو الوسيلة للوصول إلى حقائق الأشياء والظواهر ومعرفة كل العلاقات التي تربط بينها وبعضها البعض، سواء كانت هذه الظواهر اجتماعية أو اقتصادية أو طبيعية أو غير ذلك. لذلك يستخدم البحث العلم لتحرى غموض موضوع معين تحرياً منظماً دقيقاً بقصد اكتشاف حقائقه ومعرفة القواعد العامة التي تحكمه.

ـ مراحل البحث العلمي :

1- المشاهده.
2- الاحساس بمشكلة أو بوجود ظاهرة .
3- وضع الفرض العلمى المبدئى اللازم لتفسير الظاهرة .
4- مراحل البحث الاحصائى .
5- جمع البيانات و المعلومات .
6- تبويب و عرض البيانات .
7- تحليل البيانات .
8- تفسير البيانات .
9- أستنباط نظرية أو قاعدة عامة أو قانون أو قرار.

ـ تاريخ علم الاحصاء وتطوره :

لقد مر علم الإحصاء بثلاث مراحل للتطور ساير من خلالها حاجات الإنسان ورافق في تقدمه تقدم الحضارة وسد حاجاتها حتى أصبح اليوم يحتل مكانة رفيعة وهذه المراحل هي:

- مرحلة التعداد
- مرحلة الحساب السياسي
- مرحلة الإحصاء وحساب الاحتمالات

- مجالات استعمال علم الاحصاء في الحياه اليوميه :

لم يعد علم الإحصاء في الوقت الراهن مقتصرا على مجالات محددة بل امتد ليشمل معظم القطاعات في مختلف ميادين الحياة ، وفيما يلي سنورد أمثلة لبعض المجالات التي يستعمل فيها الإحصاء والتي كان له دور بارز في حل كثير من مشاكلها وبالتالي تقدمها وتطورها :

- يستخدم الإحصاء في تطوير التعليم وخططه.
- يستعمل الإحصاء في دراسة مختلف العلوم.
- يستعمل الإحصاء في مجال الدعاية والإعلانات التجارية
- يستعمل الإحصاء بشكل كبير من قبل شركات التأمين
- يستعمل الإحصاء في حساب الأرقام القياسية
- يستعمل الإحصاء في اختبارات الذكاء والتحصيل والقدرات
- يستعمل الإحصاء بشكل كبير في القطاع الصناعي

- تعريف علم الاحصـــاء :

الإحصاء في اللغة :

يعرف الإحصاء في اللغة بأنه العدد الشامل

الإحصاء في الاصطلاح :

ويعرف الإحصاء في الاصطلاح بأنه فرع من فروع الرياضيات يهدف إلى جمع وعرض وتنظيم ووصف وتحليل البيانات المقاسة رقميا مما يساعد على اتخاذ قرارات واستنتاجات وتوصيات مبنية على نظرية الاحتمالات .

ـ أهداف علم الاحصـــــــــــاء :

- جمع البيانات عن الظواهر المختلفة التي تهم الباحث بطرق علمية محددة تحديدا دقيقا وبشكل مسبق .
- تبويب البيانات طبقا لأسلوب تصنيف محدد مسبقا .
- عرض البيانات باستخدام أحد الأساليب التالية: الجداول، الأشكال البيانية، الرسوم البيانية
- وصف البيانات عن طريق إبراز الخصائص الأساسية لها والتي يمكن التعبير عنها بمقاييس معينة ومحددة، والخصائص الأساسية لأي مجموعة من البيانات تقاس بمقاييس النزعة المركزية، أو مقاييس التشتت، أو مقاييس الالتواء والاعتدال .
- تحليل البيانات المبوبة عن طريق استعمال خصائصها الأساسية التي تم إبرازها للوصول إلى الأرقام ذات العلاقة بالمشكلة والتي يهم الباحث الحصول عليها للوصول إلى نتائج محددة .
- استخدام النتائج وتفسيرها تفسيرا منطقيا مناسبا لطبيعة المشكلة التي يبحثها ، حتى يتسنى للباحث الاستفادة منها وتطبيقها في الحياة الواقعية.

ـ أهمية علم الاحصاء للباحث والبحوث العلمية :

يعتبر علم الإحصاء وسيلة لا غاية يساعد استخدامه على التالي:

- الوصف بدقة إلى اكبر حد ممكن .
- التزام التحديد والدقة في أساليبنا العملية وفي تفكيرنا .
- تلخيص نتائجنا في شكل ملائم ذو معنى واضح .
- استخلاص النتائج في الدراسات والبحوث.
- التنبؤ بالمدى الذي تحصل فيه ظاهرة تحت ظروف نعرفها ونقيسها .
- تحليل بعض العوامل المعقدة والمتشابكة التي تؤثر في حادث من الحوادث .

ـ أقســـــــــــام علم الإحصاء :

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

1- الإحصاء الوصفي Descriptive Statistics

2- الإحصاء الاستنتاجي أو الاستدلالي (التحليلي) Inferential Statistics

ويلاحظ من التعريفين السابقين بأن الإحصاء الاستنتاجي (التحليلي) يبدأ بالفعل حيث ينتهي الإحصاء الوصفي، فبعد إبراز الخصائص الأساسية للبيانات يبدأ الإحصاء الاستنتاجي (التحليلي)، حيث يتم تحليل البيانات واستخدام نتائج التحليل في الاستنتاج ثم تفسير تلك النتائج منطقيا واتخاذ قرارات في ضوء ذلك .

المحـــــــــــــــــــــاضره الثانيه

(جمع البيانات وترميزها)

- مصطلحات علم الاحصـــــــــاء :

المجتمع Population

ويقصد به المجتمع الإحصائي للظاهرة محل الدراسة. ويعرف بأنه جميع المفردات التي يجمعها إطار عام واحد أو مجموعة خصائص عامة واحدة.

العينة Sample

هى جزء من المجتمع الإحصائى محل الدراسة اخير بطريقة علمية ليتم إجراء الدراسة عليه

المتغير variable

هو خاصية عن المجتمع الإحصائى والتى يتم أختبارها من خلال التحليل الإحصائى. فهو أي صفة أو خاصية تتغير من شخص لآخر ومن وقت لآخر ويعمد الباحث لدراستها.

المعلمة Parameter

هى قياس وصفى لأحد المتغيرات يتم بأستخدام بيانات المجتمع الإحصائي كله.

الإحصائية Statistic

هى قياس وصفى لأحد المتغيرات يتم بأستخدام بيانات العينة والتى تكون تقدير لمعلمة المجتمع

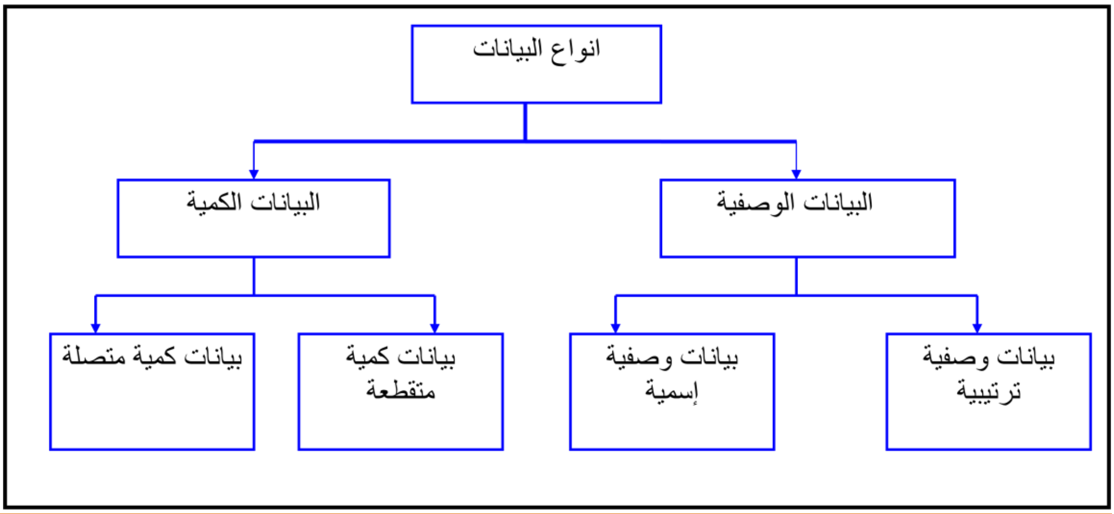
البيانات Data

هى القيمة الوصفية أو الرقمية التى نحتاج إليها لمساعدتنا فى جعل القرارات التى نتخذها أكثر معلوماتية فى موقف محدد

قبل جمع البيانات لا بد من الإجابة على السؤال التالي:

- ما البيانات الواجب أو المطلوب جمعها؟
- وما البيانات المرفوضة والتي يجب استبعادها لعدم الحاجة إليها ؟

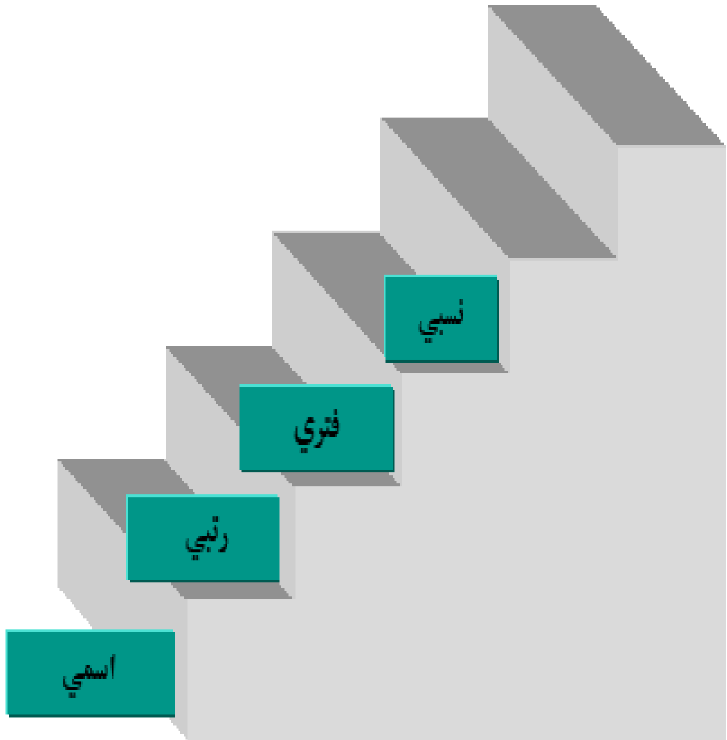
ـ أنواع البيانات الاحصائيه:



أنواع مستويات القياس للبيانات :

يوجد أربعة أنواع من مستوى القياس للبيانات هي :

- الأسمى (Nominal)
- الرتبي (Ordinal)
- الفترى (Interval)
- النسبي (Ratio)



ـ مصادر البيانات :

تتمثل مصادر البيانات فى ثلاث مصادر أساسية وهى:

- المصادر التاريخية للبيانات
- الملاحظة
- المصادر الميدانية

ـ أدوات جمع البيانات للمصادر الميدانية:

يقصد بأداة جمع البيانات الوسيلة التي تتم بواسطتها عملية جمع البيانات بهدف اختبار فرضيات البحث أو الإجابة عن تساؤلاته .

ويتوقف اختيار الأداة المناسبة لجمع البيانات اللازمة والتي ستستخدم في إجراء بحث معين على:

- نوعية البحث نفسه
- طبيعته
- الهدف من تطبيق البحث
- نوعية المفحوصين وخصائصهم ...الخ

- وقد يستخدم الباحث أداة واحدة فقط لجمع البيانات التي يحتاج إليها في بحثه، وقد يستخدم أكثر من أداة إذا وجد مبررا لذلك.
- لذا فالهدف النهائي من إعداد وسائل وأدوات جمع البيانات هو الحصول على تلك المعلومات التي تخدم في تحقيق أغراض البحث ودراسة مشكلته، وإيجاد الحلول المناسبة له .

الأدوات الأساسية شائعة الاستعمال من قبل الباحثين لجمع البيانات:

أولا: الاستبانة

ثانيا: المقابلة

المحاضره الثالثه

- أساليب إجراء البحث الميداني

عند القيام بالبحث والاعتماد على المصادر الميدانية فى الحصول على البيانات يواجهنا تساؤل هام لابد من الإجابة علية من قبل الباحث

" 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).

***** إن كلا من الأسلوبيين يمكن تطبيقه لجميع الحالات، وهناك من الأسباب التي تدعونا لتطبيق أسلوب الحصر الشامل أو أسلوب العينات.**

- 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. ويعتبر الحصر الشامل مناسب في الحالات التالية :

- التعدادات: مثل تعداد السكان وتعداد المناطق الصناعية والمؤسسات التجارية
- الحالات التي إذا تركت بعض مفرداتها دون فحص قد تؤدى إلى إلحاق الضرر بالمجتمع كلة: مثل المرضى المصابين بمرض أنفلونزا الطيور – تطعيم الاطفال من مرض معين .

***** إلا أن الحصر الشامل يتطلب وقت و جهد كثير وكذلك تكلفة كبيرة بالإضافة أنه لا يصلح فى حالات المجتمعات غير المحدودة .**

- 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. كما أن أم ما يميز أسلوب العينات أنه يصلح في دراسة المجتمعات التي ينتج عن فحص ودراسة مقراتها هلاك تلك المفردات، مثل: فحص اللمبات الكهربائية المنتجة – فحص دم الأنسان – فحص البيض المنتج في أحد مزارع الدواجن .

*****إن أهم عيوب اسلوب العينات هو** ما يسمى بخطأ التحيز Sampling Bias و هو ذلك النوع من الأخطاء التي قد يقع فيها الباحث بقصد أو بدون قصد نتيجة عدم تمثيل العينة تمثلا صادقا و كاملا لمفردات المجتمع الإحصائي محل الدراسة والذى قد يرجع إلى تحيز الباحث لفكرة أو رأى معين أو التحيز لمفردات العينة .

ـ أقسام مجتمع البحث :ـ

قسم بعض العلماء مجتمع البحث الى قسمين:

-المجتمع الكلي للبحث : ويعني كل من يمكن أن تعمم عليه نتائج البحث .

-المجتمع الذي يمكن التعرف عليه : يعني القائمة التي يمكن للباحث أن يتعرف عليها .

- تعريف مجتمع البحث هو مصطلح علمي منهجي يراد به كل من يمكن أن تعمم عليه نتائج البحث .

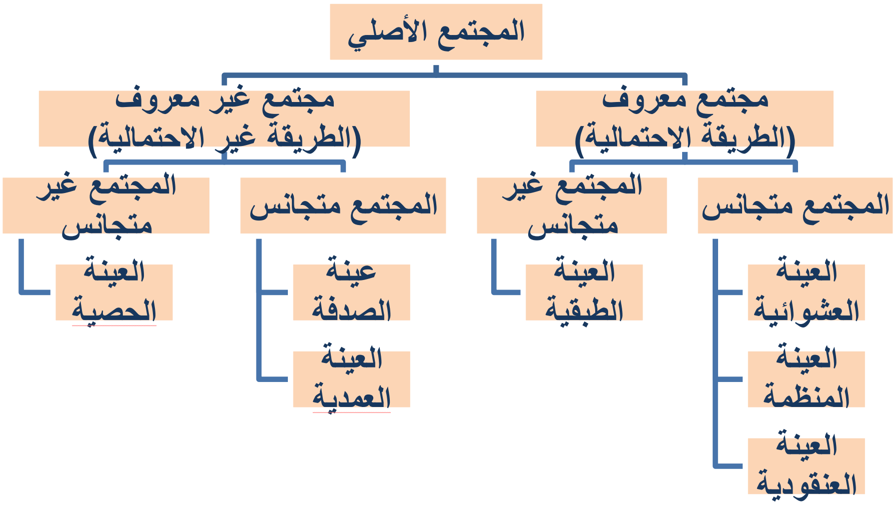
فمجتمع البحث هنا يشمل كل مبنى مدرسى حكومى فى المملكة

- تعريف عينة البحث بأنها جزء من المجتمع اختير بطريقة علمية بشرط ان تمثل المجتمع ككل.

وعينة البحث تشمل بعض المبانى المدرسية الحكومية فى المنطقة الشرقية

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

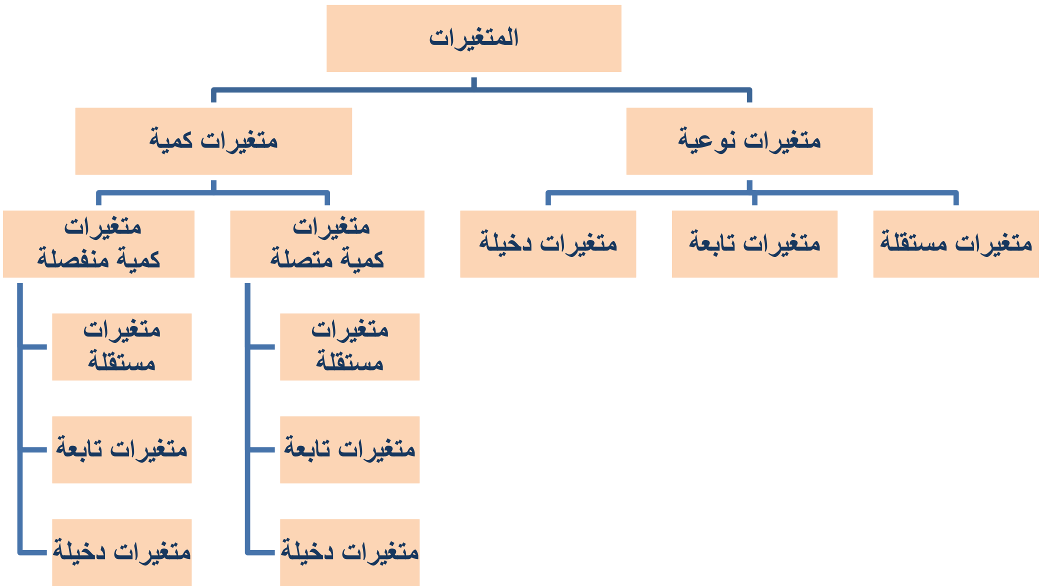
ـ طرق اختيار العينات :ـ



ـ المتغير والثابت في البحث العلمي:ـ

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

ـ تصنيف المتغيرات :ـ



ـ الخطوات الواجب مراعاتها بعد جمع البيانات :ـ

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

أولا : تسجيل البيانات

ثانيا : ترميز البيانات

ـ طرق ترميز البيانات :ـ

1- الترميز الرقمي أو العددي:

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

2- الترميز الأبجدي أو الحرفي:

ويقصد بالترميز الأبجدي أو الحرفي استخدام الحروف بدلا من الأرقام لتمييز مفردات البيانات، فمثلا استخدام الحرف M للذكور والحرف F للإناث لتمييز الجنس في نظام البيانات الشخصية.

3- الترميز الأبجدي الرقمي:

ويقصد بالترميز الأبجدي الرقمي استخدام الحروف والأرقام لتمييز مفردات البيانات، فمثلا استخدام الحرف والرقم L1 للمستوى الدراسي الأول و L2 للمستوى الدراسي الثاني و L3 للمستوى الدراسي الثالث و L4 للمستوى الدراسي الرابع وذلك لتمييز المستويات الدراسية الجامعية للطلاب والطالبات.

ثالثا : تصنيف البيانات

رابعا : مراجعة وتنقية البيانات

- ترميز بيانات الاستبانة وجعلها متاحة لبرنامج الـSPSS :-

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

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

***ولغرض تفريغ البيانات المجموعة من خلال هذه الاستبانة بطريقة مناسبة يفهمها برنامج الـ SPSS لابد من توضيح التالي :

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

- كل سؤال (فقرة) في الاستبانة تمثل متغير Variable .

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

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

1- سؤال يسمح باختيار إجابة واحدة فقط :

وهي ذلك النوع من الأسئلة التي تلزم المستجيب باختيار إجابة واحدة فقط، ويتم التعامل مع هذا النوع من الاسئلة من خلال تمثيله بمتغير واحد يحوي جميع الاجابات الممكنة، فمثلا السؤال رقم (2) في الاستبانة السابقة والذي يقول:

عدد سنوات الخبرة في العمل الأكاديمي:

1- ( ) أقل من سنة

2- ( ) من 1-5 سنوات

3- ( ) من6-10 سنوات

4- ( ) من11-15 سنة

5- ( ) أكثر من16 سنة

ففي هذا السؤال هناك خمس احتمالات

فتعطى كل إجابة رقم يمثلها، فعلى سبيل المثال يعطى:

أقل من سنة القيمة (1)،

و من 1-5سنوات القيمة (2)،

ومن 6-10سنوات القيمة (3)،

ومن 11-15 سنة القيمة (4)،

وأكثر من 16 سنة تعطى القيمة (5).

وبالإمكان أن تعطى هذه الإجابات رموزا حرفية إذا تم تعريف المتغير على أنه متغير من نوع حرفي String، لكن يفضل عدم استخدام مثل هذا الإجراء وذلك لأن إدخال البيانات الرقمية في برنامج الـ SPSS أسهل .

2- سؤال يسمح باختيار أكثر من إجابة واحدة :

وهي ذلك النوع من الأسئلة التي تتاح من خلالها الفرصة للمستجيب لاختيار أكثر من إجابة، ويتم التعامل مع هذا النوع من الاسئلة والفقرات من خلال تمثيله بعدد من المتغيرات يماثل عدد الاجابات أو الاحتمالات المتاحة للسؤال أو الفقرة، مثال على ذلك السؤال رقم (7) في الاستبانة الآنفة الذكر والذي يقول :

- ما أهم المعوقات التي قد تحول دون استخدامك للإنترنت في البحث العلمي؟ (يمكن اختيار اكثر من عائق)

1- ( ) عدم الاهتمام بالإنترنت .

2- ( ) عدم توفر التدريس المناسب لاستخدام الإنترنت.

3- ( ) عدم وجود الوقت الكافي .

4- ( ) عدم توفر أجهزة الحاسب الآلي.

5- ( ) عدم توفر المتصفح المناسب للإنترنت .

6- ( ) عدم توفر الحوافز الخارجية لاستخدام الإنترنت في البحث العلمي

7- ( ) عدم توفر المعلومات والمهارات الأساسية لاستخدام الإنترنت.

8- ( ) الاهتمام بحقوق النشر.

9- ( ) الخوف من العولمة.

ففي هذا النوع من الأسئلة نلاحظ أن المستجيب قد يختار أكثر من إجابة على هذا السؤال، لذلك فإن متغيرا واحدا لا يكفي لتمثيل هذا السؤال، بل يحتاج هذا السؤال إلى تسعة متغيرات، كل متغير له احتمال إجابتين ("نعم" وتأخذ القيمة " 1" ، و "لا" وتأخذ القيمة صفر "0" مثلا)

3- سؤال مفتوح جزئيا:

وهي ذلك النوع من الأسئلة التي تسمح للمستجيب باختيار إجابة موجودة ضمن الخيارات أو كتابة إجابة أخرى غير موجودة ضمن الخيارات المتاحة في السؤال، ومثال على ذلك السؤال رقم (1) في الاستبانة والذي يقول :

الدرجة العلمية التي تحملها:

1- ( ) دكتوراه

2- ( ) ماجستير

3- ( ) بكالوريوس

4- ( ) غير ذلك ، حدد ....

- فهذا النوع من الاسئلة يتم تمثيله بمتغير واحد فقط، لأن المطلوب من المستجيب اختيار إجابة واحدة، إلا أن المشكلة في هذا النوع من الأسئلة تكمن في الخيار ذو الإجابة المفتوحة، ففي هذا السؤال هناك أربع احتمالات، فتعطى كل إجابة رقم يمثلها ، فعلى سبيل المثال يعطى دكتوراه القيمة (1)، وماجستير القيمة (2)، وبكالوريوس القيمة (3)، أما الخيار الرابع "غير ذلك" فيتم التعامل معه بأكثر من طريقة منها:

تعيين قيمة محددة لهذا الاحتمال وليكن القيمة (4) بغض النظر عما ذكر من درجات علمية للمستجيبين، وهذا الإجراء يسهل التعامل مع هذا الخيار إلا أنه يفقد الباحث معلومات كثيرة.

حصر جميع الاجابات ومن ثم تحديد قيمة لكل درجة علمية غير تلك التي ذكرت في السؤال، وهنا يتم تحديد عدد الاحتمالات المتاحة للسؤال بعدد الإجابات المذكورة في الإستبانات، وهذا الاجراء يحتاج إلى وقت كبير لأنه سيتم معالجة كل استبانة بشكل منفرد ليتم جمع كل الإجابات الممكنة.

عدم التعامل مع هذا المتغير على أنه متغير رقمي Numeric والتعامل معه على أنه متغير حرفي String، لذا لا يتم تعيين قيم تصف الإجابات، بل يتم كتبة الاجابة كاملة لكل درجة علمية. وهذه الطريقة تؤدي إلى حصر جميع الإجابات إلا أنها تزيد العبء على الباحث من خلال إدخال بيانات أكثر في الحاسب مما قد يؤدي إلى زيادة أخطاء الإدخال.

- تمرين

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

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

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

المطلوب :

- ما نوع الإحصاء الذي استخدمه الباحث في هذه الدراسة؟ علل ذلك ؟
- 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 ؟

المحاضره الرابعه

العرض الجدولي للبيانات (تبويب البيانات)

كما تعرضنا في المحاضرات السابقة ان البيانات هي المستهدفة في الاحصاء فهي التي تجمع وهي التي تعرض وهي التي تحلل. وعرض البيانات صار في الآونة الاخيرة علما وفنا قائما بذاته، فالصورة 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 منها:

- طريقة مطولة وعقيمة.
- تتطلب وقتا طويلا في القراءة في الأمر الذي يجعل الملل يتسرب الى القارئ.
- قلما يكلف الانسان نفسه مشقة الاطلاع على احصاءات معروضة بهذه الكيفية.
- انه يتعذر عرض بيانات خاصة بعدد كبير من السنين بهذه الطريقة.

وبالتالي تعتبر الطريقة السابقة غير فنية في عرض البيانات ، أما الطرق الفنية في عرض البيانات الاحصائية فهي:

- العرض الجدولى للبيانات (تبويب البيانات)
- العرض البيانى للبيانات

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

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

ـ أهمية الجداول الاحصائية:

- تعبر عن الحقائق الكمية المعروضة بعدد كبير من الارقام، وعن طريق عرض هذه الارقام في جداول بطريقة منظمة فانه يمكن بالتالي اكتشاف أهميتها والاستفادة من.
- تعتبر الجداول وسيلة يمكن بواسطتها تلخيص المعلومات الرقمية الكثيرة العدد، المتغيرة القيم، مما يسهل التعرف عليها.
- ان الجداول تستوعب بسهولة عدد كبير من الموضوعات، فتفريغ الارقام في جداول يقلل كثيرا من تكرار الكلمات التي تصف البيانات، لأن عنوان كل عمود في الجدول ينطبق على الارقام فيه، فهي بالتالي طريقة اقتصادية في الوقت والحيز والمجهود.
- تساعد الجداول على 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.**

رقم الجدول   جدول رقم (5)

عنوان الجدول   يوضح طلبة جامعة الملك فيصل للعام الجامعي1423هـ

عنوان توضيحي   (مصنفون حسب الجنس)

عنوان العمود

هيكل الجدول

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

المصدر: جامعة الملك فيصل، احصائية الجامعة حسب الكليات

* يحدد المستوى بالسنة الدراسية التي يدرس فيها الطالب .

ـ أنواع الجداول الاحصائية:ـ

تقسم الجداول تبعا لدرجة تعقيدها الى:

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

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

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

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

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

1. مجموعة البيانات الوصفية والكمية المتقطعة

البيانات الوصفيه (هي أي معلومات يجمعها الباحث وتتغير في الصفات)

الكميه المتقطعه (هي المعلومات الغير قابله للكسور أي تكون اعداد صحيحه)

2. مجموعة البيانات الكمية المتصلة

الكميه المتصله (هي أي معلومات قابله للكسور أي تكون فيها قيمه كسريه)

أولا : البيانات الوصفية والكمية المتقطعة:

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

– مثال (لمتغير وصفي):

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

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

المطلوب: عرض البيانات السابقة فى صورة جدول التوزيع التكرارى

الحل مفصلا في الكتاب صفحة 45

الحل /

- أرسم جدول واحصر فيه جميع الالون المتشابهه

- أضع شرطه بجانب كل لون في عامود العلامات اذا وصل عدد الشرطات الى خمسه اضع عليها علامه عكسيه واسميها حزمه

الجدول كالتالي :-

التكرار النسبي	التكرارات	العلامات	اللون
0,3	6	~~////~~ /	أحمر
0,2	4	////	أبيض
0,15	3	///	أزرق
0,2	4	////	أخضر
0,15	3	//	بنفسجي
1,00	20	مج ك =	

- نترجم العلامات الى ارقام كما في العامود الثالث .

- اذا اردت ان استخرج قانون تكرار النسبي استخدم المعادله التاليه =

$$\text{التكرار النسبي} = \frac{\text{تكرار الدرجه}}{\text{مج ك}}$$

ملاحظه / لابد ان يكون مجموع التكرار النسبي في الأخير 1 صحيح .

- مثال (لمتغير كمى متقطع): ان تكون الارقام غير قابله للكسور

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

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

المطلوب :

١. عرض البيانات السابقة فى صورة جدول تكرارى
٢. أحسب الأحتمالات التالية:

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

الحل مفصلا في الكتاب صفحة 46-47

وهذا حلي الشخصي من شرح الدكتوور

الحل :

- نحصر القيم المتكرره في السؤال .

- اقوم برسم جدول يحوي هذه المعلومات .

للتذكير قانون

التكرار النسبي = $\frac{\text{تكرار الدرجه}}{\text{مج ك}}$

عدد الحوادث	العلامات	التكرارات	التكرار النسبي
صفر	~~////~~ ////	9	0,30
1	~~////~~ ~~////~~ /	11	0,366
2	~~////~~ //	7	0,233
3	///	3	0,10
المجموع		30	

احتمال (الا يتعرض أي شخص لحادث)

يعني صفر والصفر = 0,30

احتمال (أن يكون هناك حادث واحد على الاكثر)

بمعنى انه سوف يكون اكبر عدد للحودات هو 1

وذلك بجمع نسبة تكرار الصفر + 1 =

0,30 + 0,366 = 0,666

احتمال (أن يكون هناك حادث واحد على الأقل)

بمعنى انه سوف يكون أقل عدد للحوادث هو 1

لذلك لا بد من استبعاد نسبة تكرار الصفر ونحسب باقي الاعداد

1 + 2 + 3 =

0,366 + 0,233 + 0,10 = 0,7

ثانيا : البيانات الكمية المتصلة: (و هي معلومات قابله للكسور أي تكون فيها قيمه كسريه)

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

الخطوة الأولى: تحديد عدد الفئات -:

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

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

الخطوة الثانية: تحديد طول الفئة:

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

طول الفئة = المدى ÷ عدد الفئات

ويمثل المدى الفرق بين اكبر مفرده واصغر مفرده في البيانات الأولية.

الخطوة الثالثة: تعيين حدود الفئات:-

نبدأ بتعيين الحد الأدنى للفئة الأولى وهو قيمه اصغر مفرده في البيانات الأولية للظاهرة محل الدراسة، ويجوز أن نختار قيمه أقل من أصغر مفرده ليبدأ الحد الأدنى للفئة الأولى بقيمه صحيحة، ونقوم بتحديد الحد الأعلى للفئة الأولى بإضافة طول الفئة الذى حصلنا عليه من الخطوة الثانية. يعتبر الحد الأدنى للفئة الثانية هو الحد الأعلى للفئة الأولى وبإضافة طول الفئة نصل إلى الحد الأعلى للفئة الثانية، ونستمر في تكرار هذه الطريقة حتى يتم تكوين عدد الفئات المطلوبة المحدد فى الخطوة الأولى.

يجب علينا التأكد من عدم وجود تداخل فيما بينها الفئات بعضها البعض، حيث أن الفئة تحتوى على كل المفردات التى تساوى حدها الأدنى تماماً وما يزيد عنه حتى يصل إلى حدها الأعلى.

الخطوة الرابعة: توزيع التكرارات على الفئات:

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

مثال :

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

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

المطلوب :

عرض البيانات السابقة فى صورة الجدول التكرارى المناسب

الحل مفصلا في الكتاب صفحة 49

المحاضره الخامسه

العرض الجدولي للبيانات (تبويب البيانات)

الجزء الثاني

في بداية هذه المحاضره يتعرض الدكتور لبعض من شرائح المحاضره السابقه من صفحه23- الى نهاية المحاضره السابقه

بالاضــــــــــــافه :

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

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

- عدد المفردات محل الدراسة
- انتظام وتوزيع تلك البيانات
- طبيعة بيانات المشكلة محل الدراسة

2- 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

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

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

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

أولا: الجداول التكرارية المنتظمة:

وهى الجداول التى يكون فيها أطوال كل الفئات متساوية

كما تم توضيحة فى المثال السابق

ثانيا: الجداول التكرارية غير المنتظمة:

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

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

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

ثالثا: الجداول التكرارية المفتوحة:

وتوضحها أشكال الجداول التالية:

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

جدول مفتوح من أسفل

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

جدول مفتوح من أعلى

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

جدول مفتوح من الطرفين

ـ الجداول التكرارية المتجمعة :

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

اولا- الجدول التكرارى المتجمع الصاعد

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

مثال:

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

فئات مساحات الاراضي دونم	عدد قطع الاراضي
1 –	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 مفصلا في الكتاب صفحة 52

ثانيا - الجدول التكرارى المتجمع الهابط (النازل):

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

مثال:

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

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

المطلوب:

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

الحل مفصلا في الكتاب صفحة 53

ـ الجدول التكرارى المزدوج:

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

مثال:

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

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

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

المطلوب:

إعداد جدول تكراري مزدوج

الحل مفصلا في الكتاب صفحة 55

المحاضـــــــــــــره السادسه

العرض البيانى للبيانات

أولا: البيانات غير المبوبة

البيانات الاسمية أو الرتبية أو الكمية المتقطعة (أي المنفصلة)

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 ..

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

أ ـ الأعمدة البيانية البسيطة :

وهي عبارة عن مجموعة من الأعمدة الرأسية أو المستطيلات المتساوية القاعة والتي تتناسب ارتفاعاتها مع البيانات التي تمثلها، وتستخدم لاظهار التطور الذي يطرأ على ظاهرة ما على مدار عدة سنوات، وعادة يؤخذ المحور الرأسي لتمثيل قيم الظاهرة، والمحور الأفقي يمثل الزمن بحيث يتناسب طول كل عمود مع العدد الذي يمثله.

ويجب مراعاة ان يقسم المحور الرأسي بحيث يسمح مقياس الرسم باظهار جميع قيم الظاهرة، كذلك يجب أن تكون المسافات بين الاعمدة متساوية.

مثال:

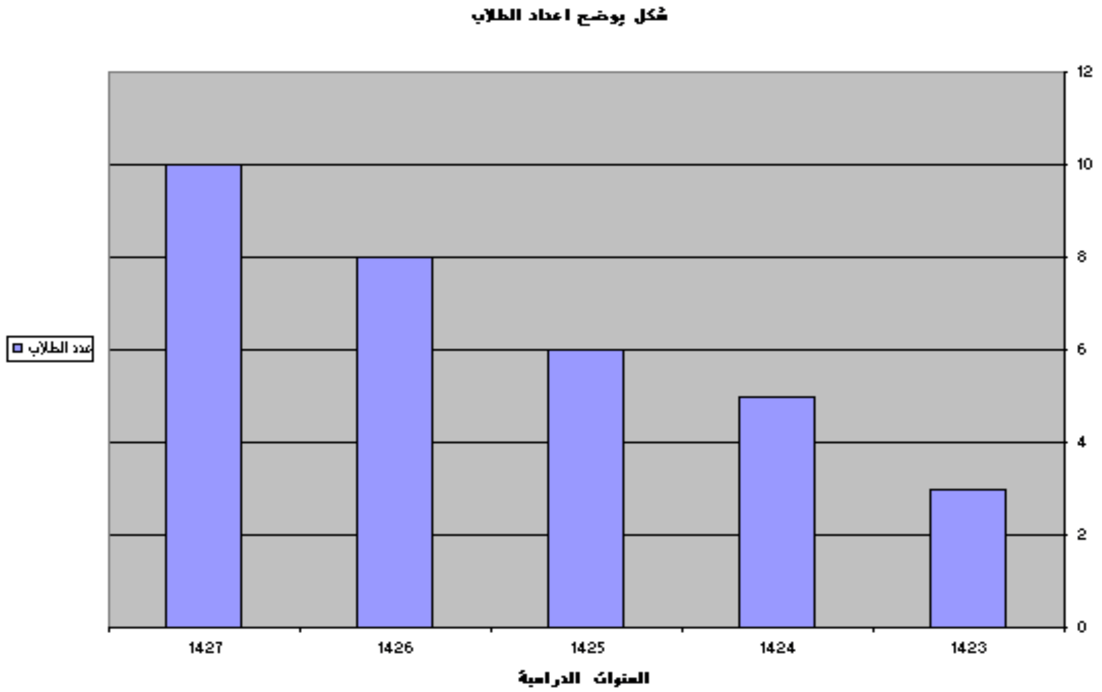
الجدول الأتي يوضح أعداد الطلاب المقيدين باحد   الجامعات في السنوات الدراسية من 1423هـ حتي 1427هـ .

السنه الدراسيه	1423	1424	1425	1426	1427
عدد الطلاب بالألف	3	5	6	8	10

المطلوب:

تمثيل البيانات باستخدام الرسم البياني المناسب

الحل :



ب – الأعمدة البيانية المزدوجة:

يستخدم هذا النوع اذا كان الهدف من الرسم هو مقارنة ظاهرتين أو اكثر لعدة سنوات، أو اذا كان لدينا بيانات مزدوجة لخواص مختلفة .

ويتم رسم الأعمدة الزدوجة بإتباع ما يلي :

- رسم عمودين متلاصقين يمثلان قيم الظاهرتين محل الدراسة في كل سنة، بحيث يتناسب طول كل عمود مع العدد الذي يمثله .
- نفرق بين الأعمدة بالتظليل أو بالالوان المختلفة ونوضح ذلك على الرسم وذلك بوضع مفتاح للرسم .
- ضرورة مراعاة أن تكون قواعد المستطيلات متساوية والمسافات بينهما متساوية.

مثال:

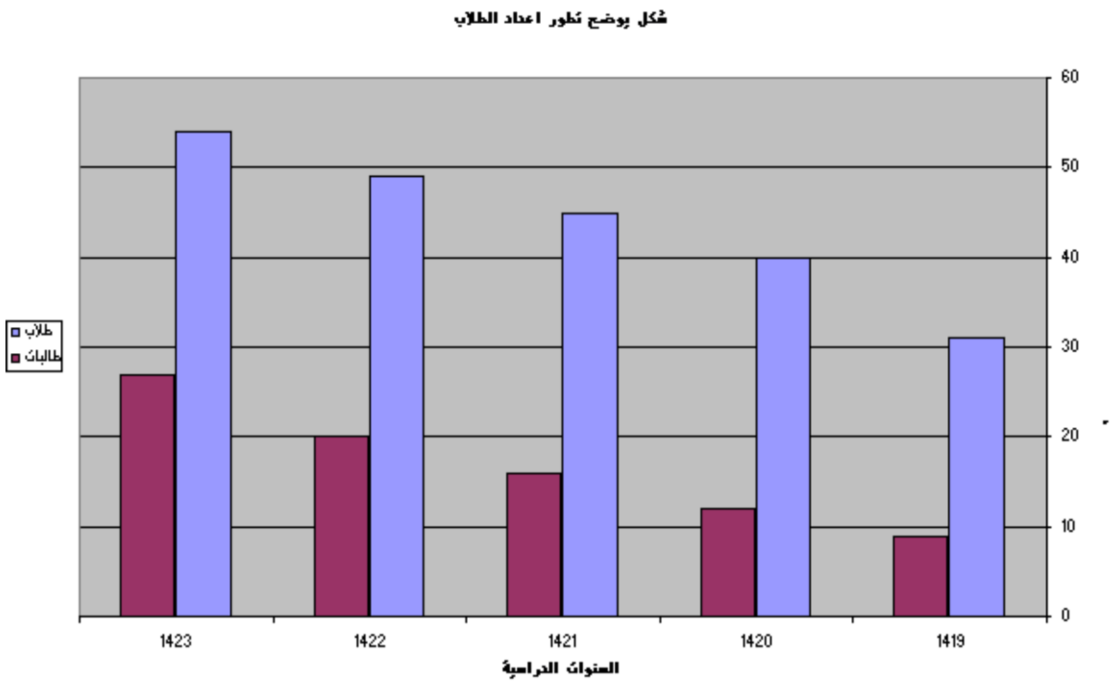
الجدول الأتي يوضح أعداد الطلبة المسجلين باحد الجامعات السعودية في السنوات الدراسية 1419هـ حتى 1423هـ

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

المطلوب:-

مثل هذه البيانات بيانيا باستخدام الأعمدة البيانية المزدوجة ؟

الحل :



جـ ـ الأعمدة البيانية المجزأة :

يستخدم هذا النوع من الرسوم البيانية في تمثيل نفس الحالات التي تستخدم فيها الأعمدة البيانية المزدوجة .

ويتم رسم هذا النوع من الأعمدة كالآتي :

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

مثال :ـ

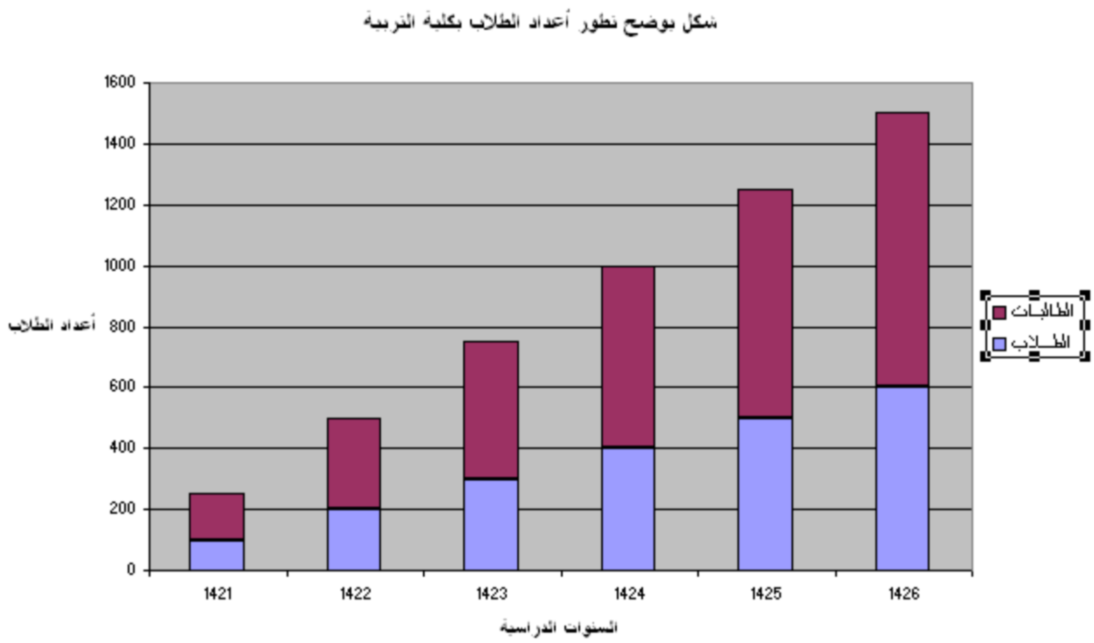
اذا كانت اعداد الطلاب والطالبات المسجلين في كلية التربية بجامعة الملك فيصل   بالاحساء تزداد كما هو موضح في الجدول الأتي:

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

المطلوب :ـ

مثل هذه البيانات بيانيا باستخدام الأعمدة المجزأة؟

الحل :



ملاحظات على استخدام الاعمدة البيانية (بأنواعها المختلفة) :

يمكن ابداء الملاحظات التالية على الرسومات بالاعمدة البيـانية بأنواعها المختلفة :

- تعتبر الاعمدة البيانية من اكثر الرسومات البيانية انتشارا، وهي عبارة عن مستطيلات قواعدها متساوية وأطوالها (ارتفاعاتها) مختلفة تتناسب مع القيم التي تمثلها، وتكون منفصلة عن بعضها البعض بمسافة يقدرها الباحث.

يفضل تظليل الاعمدة أو تخطيطها بواسطة خطوط متوازية أو ابرازها بألوان مختلفة وخاصة عند مقارنة ظواهر مختلفة.

- يستحسن اختيار مقياس رسم مناسب وثابت، ولهذا لا بد لمصمم الرسم من التعرف على القيمة الكبرى والقيمة الصغرى لتحديد مقياس الرسم المناسب. هذا ويجب البدء بالصفر على المحور الرأسي الذي يدل على القيم الرقمية حتى تكون المقارنة سهلة وسليمة وغير مضللة.

- يفضل عدم كتابة القيم التي تمثلها الاعمدة فوق الاعمدة وذلك لتلافي المبالغة في طول الاعمدة، وبالتالي تجنب اظهار الرسم مزدحما او مكتظا مما ينفر القارئ، الا اذا كان ذلك هدفا في حد ذاته .

- يمكن استخدام العمود الواحد لتمثيل اكثر من نوع واحد من البيانات، وذلك باستخدام مفهوم الاعمدة المجزأة، هذا ويفضل أن لا نعرض اكثر من ثلاثة ظواهر في العامود حتى لا يفقد الرسم البياني الهدف الأساسي منه.

- تصلح الاعمدة البيانية لتمثيل البيانات ذات المتغيرات المنفصلة، كما تصلح بشكل خاص لتميل البيانات الوصفية (النوعية) (أي غير الرقمية) وذلك كما في تمثيل الحالة الاجماعية (متزوج، مطلق، أرمل)

د – اللوحة الدائرية:

تستخدم الدائرة أو اللوحة الدائرية لتميل 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.

مثال: يمكن استخدام الدائرة لبيان توزيع طلبة جامعة الملك فيصل حسب الكليات (التربية – الزراعة – الادارة - الطب البيطري) أو توزيع طلبة كلية العلوم الإدارية (أو أي كلية أخرى حسب السنة الاكاديمية (أولى – ثانية – ثالثة – رابعة) .

- وتمثل المساحة الكلية للدائرة المجموع الكلي، ثم تقسم الدائرة الى قطاعات دائرية تتناسب مساحة كل منها مع نسبة كل جزء الى المجموع الكلي، وتميز بين هذا القطاعات بالتظليل أو بالألوان المختلفة.

- وفيما يلي خطوات رسم الدائرة وتقسيمها الى قطاعات:

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

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

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

فمثلا : مساحة القطاع الأول تحدد بوضع قاعدة المنقلة على نصف القطر ونقيس زاوية مساوية لزاوية القطاع، نسقط من عندها عمودا على مركز الدائرة، فنحصل على القطاع الأول، ثم نقيس من عند نهاية مساحة القطاع الأول زاوية مساوية لزاوية القطاع الثاني، نسقط من عندها عمودا على مركز الدائرة فنحصل على القطاع الثاني. وهكذا بالنسبة لباقي القطاعات.

مثال:

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

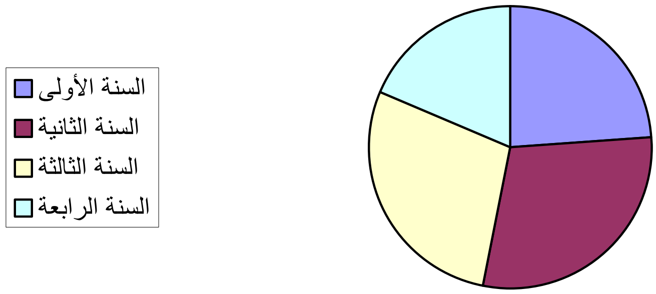
السنة الدراسية	عدد الطلبة
السنة الأولى	226
السنة الثانية	276
السنة الثالثة	266
السنة الرابعة	167
المجموع	935

المطلوب :-

عرض هذه البيانات باستخدام اللوحة الدائرية؟

الحل :

شكل بيانى يوضح توزيع طلاب بكالوريوس العلوم الإدارية للعام الجامعى 1426 هـ موزعة على حسب السنوات الدراسية



هذا ويستحسن تظليل القطاعات الدائرية أوتلوينها وذلك زيادة في قيمة الرسم البيان وبالتالي زيادة جاذبيته ووضوحه، وكذلك ينصح كتابة الجزء (السنة) داخل كل قطاع دائري .

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

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

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

- عندما تكون الكميات المقارنة كثيرة العدد نسبيا، حيث يصعب تمثيلها بالدائرة وذلك أن كثرة الكميات المقارنة تجعل الدائرة مكتظة لدرجة يصعب مقارنة التوزيع النسبي للظاهرة المدروسة.
- عند ما تكون الاجزاء المقارنة في فترات زمنية مختلفة، وهذا لا يمنع من استعمالها في فترة زمنية واحدة، الا أن الدائرة لايمكن استخدامها لمقارنة الاجزاء بالكل 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 (وهي التي تأخذ قيما أو أعداد صحيحة) كما في عدد الطلبة أو أفراد الأسرة أو عدد الكتب في المكتبة .. الخ.

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) كما في التحصيل الدراسي أو الذكاء والأعمار وكذا اسعار السلع ...الخ، وكذلك ممكن استخدامه مع البيانات المنفصلة كعدد الطلاب .. الخ .

ويتم رسم المنحنى أو الخط البياني بإتباع الآتي:

- نرسم محورين أفقي ورأسي بحيث يمثل المحور الأفقي الزمن مثلا والمحور الرأسي قيم الظاهرة.
- نستخدم نفس المبدأ الذي اتبعناه في رسم الأعمدة ا لبيانية المختلفة اللهم بدلا من رسم الأعمدة ذاتها نستعيض عنها بتعيين نقطة (إحداثية النقطة) فقط لكل منها.
- توصيل هذه النقط ببعضها بمنحنى ممهد متصل فنحصل على خط متصل يسمى المنحنى، أو القيام بتوصيل كل نقطتين متجاورتين بخط مستقيم فنحصل عندئذ على الخط البياني.

مثال :ـ

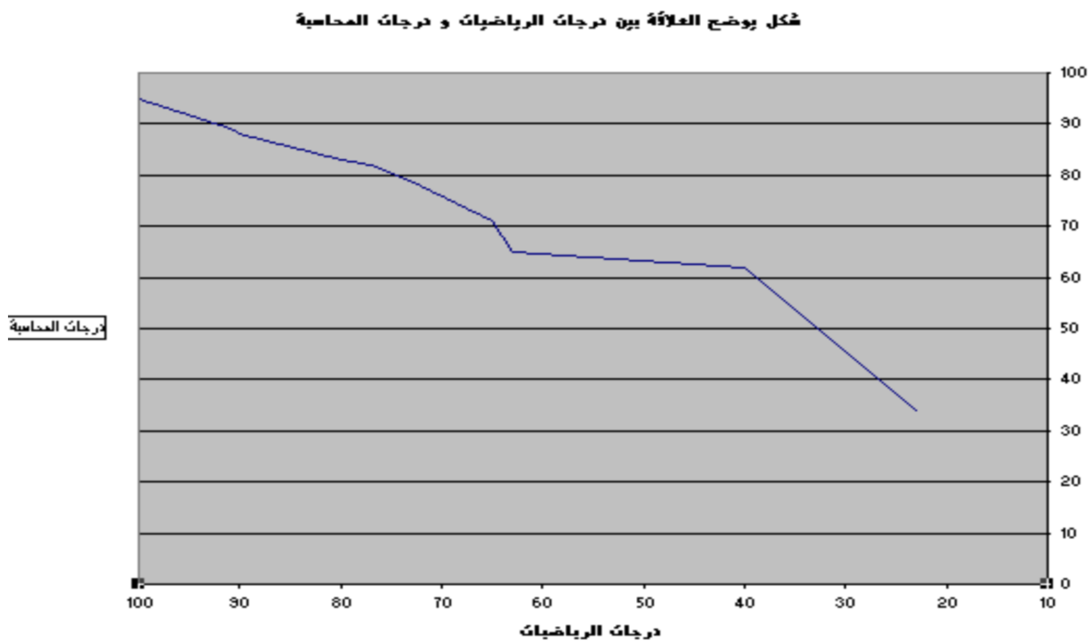
البيانات التالية لدرجات عشر طلاب بكلية العلوم الإدارية فى مقررى الرياضيات والمحاسبة، فكانت كما يلي:

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

المطلوب :ـ

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

الحل :



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 وذلك انطلقا من المبدأ الذي يرى أن العين تدرك الأشياء المتصلة بسهولة ويسر أكثر من ادراكها للأشياء المنفصلة، وبالتالي يستطيع الشخص استخلاص بعض النتائج او المدلولات الرقمية بطريقة أسهل، كما يسهل عليه معرفى الاتجاه العام للظاهرة.

- يمكن استخدام الخط البياني أو المنحنى (كما في الأعمدة البيانية) لتمثيل أكثر من ظاهرة على نفس الرسم ومقارنتها ببعضها، مع ملاحظة تمييز الخط البياني لكل ظاهرة إما بخطوط متصلة أو متقطعة أو إعطائها الوانا مختلفة وتوضيح ذلك في مفتاح الرسم.

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.

العرض البيانى للبيانات

ثانيا: البيانات المبوبة البيانات الكمية المتصلة

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: المدرج التكرارى هو عبارة عن أعمدة مستطيلة متلاصقة يعبر ارتفاع العمود فيها على التكرار المناظر للفئة.

ويستخدم المدرج التكرارى 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 الأفقي (المحور السيني) على أساس الفئات و هنا يظهر حالتين هما:

الحالة الأولى:- تساوى أطوال الفئات

وفى هذه الحالة يكون ارتفاع المستطيل معبرا عن عدد مرات تكرار وجه الظاهرة محل الدراسة حيث انه يتناسب مع مساحه المستطيل، وذلك لان طول الفئه هو عرض المستطيل، وحيث أن أطوال الفئات متساوية فإن مساحه المستطيل تتناسب مع طوله فقط.

الحالة الثانية:- عدم تساوى أطوال الفئات

وفى هذه الحالة لابد من إجراء تعديل في التكرار الأصلى قبل رسم المدرج التكرارى، لذا فإننا نقوم بإيجاد التكرار المعدل والذى هو عبارة عن ناتج قسمه التكرار الأصلى لكل فئه على طول الفئة المقابلة، وهنا تكون مساحه المستطيل معبره عن وجه الظاهرة المقابل لها، وليس ارتفاع المستطيل.

خطوات رسم المدرج التكراري:

- نرسم محورين أفقي ورأسي بحيث يمثل المحور الأفقي الفئات والمحور الرأسي التكرارات.
- نمثل بيانات الدراسة من خلال مجموعة من المستطيلات المتلاصقة بحيث يعبر ارتفاع المستطيل عن عدد مرات تكرار وجه الظاهرة محل الدراسة.

مثال:

فيما يلى بيان بتوزيع لعينة من 40 عامل على أساس فئات العمر للعمال.

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

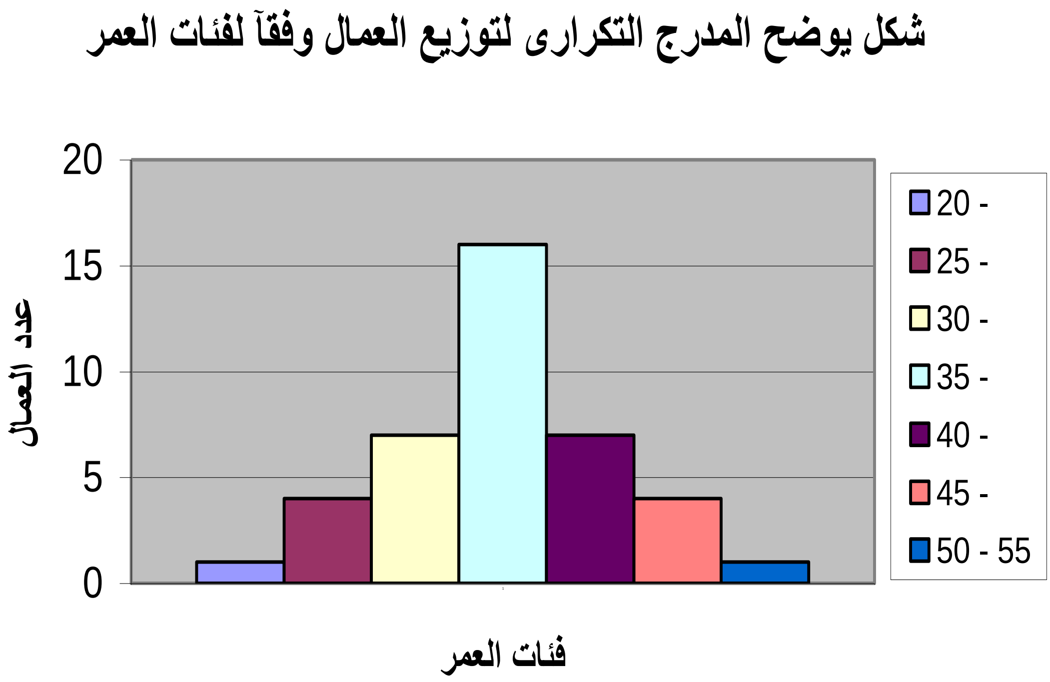
المطلوب:-

عرض البيانات السابقة في شكل المدرج التكرارى.

الحل:-

حالة فئات العمر المتساوية:

- يتم رسم المدرج التكرارى على أساس التكرار الأصلى (فئات عمر العمال).
- نرسم محورين أفقي ورأسي بحيث يمثل المحور الأفقي فئات العمر والمحور الرأسي تكرار عدد العمال في كل فئة عمرية.
- نمثل بيانات الدراسة من خلال مجموعة من المستطيلات المتلاصقة بحيث يعبر ارتفاع المستطيل عن عدد العمال في كل فئة.



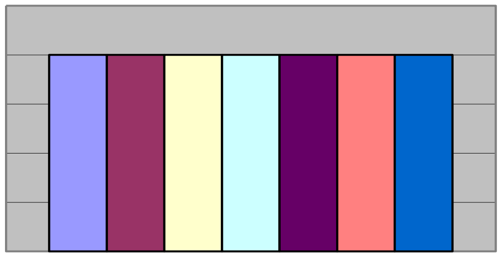
- بعض خصائص التوزيع التكراري: -

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

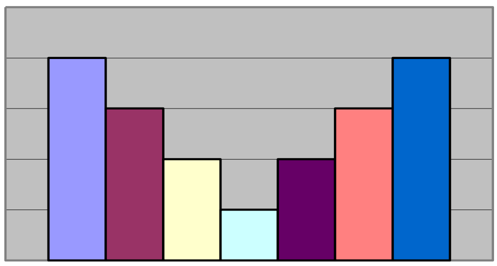
الخاصيه الأولى: التماثل

يسمى المدرج التكرارى متماثلا عندما نقوم برسم خط مستقيم في منتصف المدرج التكرارى فيظهر لنا التطابق التام بين الجانبين حول الخط المستقيم. وذلك يظهر في الرسم السابق مباشرة حيث يكون الجانب الأيمن كخيال للجانب الأيسر في المرآه، وكذلك قد يكون شكل المدرج التكراى متماثل كما هو واضح فى الشكلين التاليين:-

شكل يوضح التوزيع المتماثل



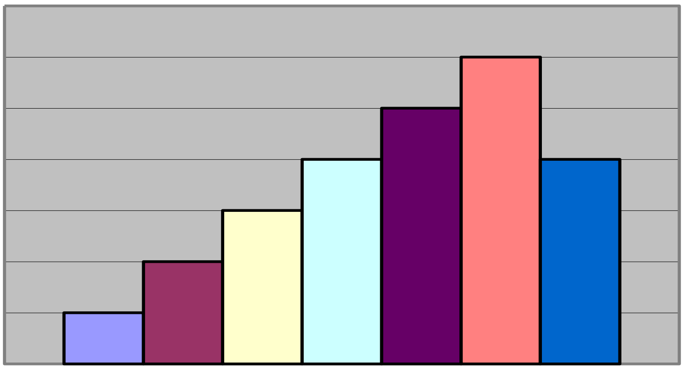
شكل يوضح التوزيع المتماثل



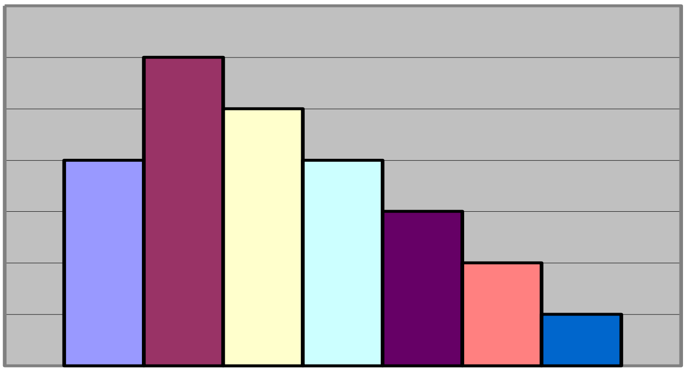
الخاصيه الثانيه: الإلتواء

وعندما يكون ذيل التوزيع جهة اليسار - بمعنى أن الطرف الايسر للتوزيع أطول من طرفه الإيمن – يكون الإلتواء بإتجاه اليسار ويسمى توزيع سالب الالتواء، فمثلا توزيع الوقت اللازم لإجابة الإمتحان بالنسبة لعدد الطلاب يكون فى الغالب سالب الالتواء ويرجع ذلك لقيام عدد قليل من الطلاب بتسليم أوراق الإجابه قبل موعد إنتهاء الامتحان، وفى المقابل يفضل الكثير من الطلاب تسليم أوراق الاجابة مع نهاية وقت الامتحان وفيما يلى توضيح الالتواء بنوعيه فى الشكلين التاليين:-

شكل يوضح الإلتواء السالب



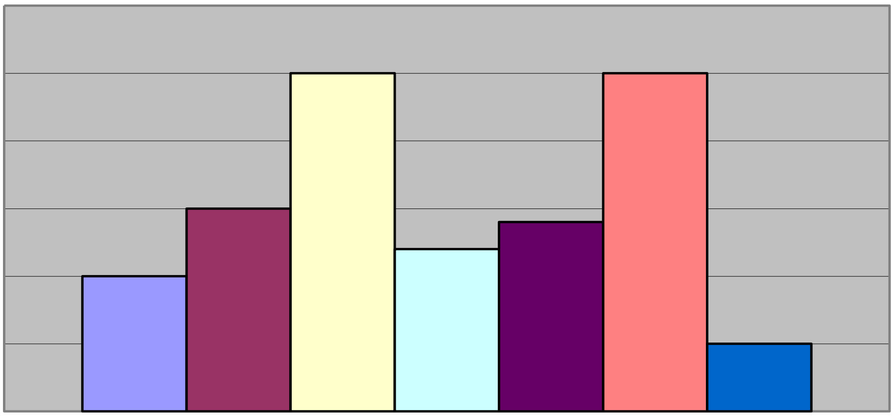
شكل يوضح الإلتواء الموجب



الخاصيه الثالثة: المنـوال

المنوال هو القيمة الأكثر تكرراً (شيوعاً) في الظاهره محل الدراسه، وفى بعض الأحيان يكون المدرج التكرارى أحادى المنوال عندما تقع معظم البيانات داخل فئه فى منتصف التوزيع التكرارى وتسمى الفئه المنواليه وهي تمثل قمه واحده للتوزيع، مع وجود بعض البيانات قبل وبعد هذه الفئه، وفى أحيان أخرى يكون المدرج التكرارى ثنائي المنوال، وذلك في حالة وجود قيمتين في التوزيع ويشترط تساوى القمتين معاً، فمثلاً إذا نظرنا إلى التوزيع التكرارى للدخول فى إحدى البلدان التى يعيش فيها كثير من الاغنياء وكثير من الفقراء وقله من الطبقه المتوسطه، فإن شكل المدرج التكرارى لسكان هذا البلد يكون ثنائى المنوال كما فى الشكل التالي:

شكل يوضح توزيع ثنائى المنوال



المضلع التكرارى : المضلع التكراري هو مضلع مغلق نحصل عليه من خلال حساب مراكز الفئات أو بتنصيف الأضلاع العلوية للمستطيلات في المدرج التكراري، ثم نوصل هذه النقاط بعضها مع بعض، ولكي نغلق الخط المنكسر الذي حصلنا عليه نعتبر أن هناك فئتين متطرفتين واحدة في أقصى اليمين والثانية في أقصى اليسار وتكرار كل منهما صفر، نأخذ مركز كل من هاتين الفئتين، ونغلق المضلع كما يبدوا لنا في المثال التالي:

خطوات رسم المضلع التكرارى:

- نرسم محورين أفقي ورأسي بحيث يمثل المحور الأفقي الفئات والمحور الرأسي التكرارات.
- **لكي نرسم المضلع من خلال المدرج التكراري** نقوم بتمثيل بيانات الدراسة من خلال مجموعة من المستطيلات المتلاصقة بحيث يعبر ارتفاع المستطيل عن عدد مرات تكرار وجه الظاهرة محل الدراسة.
- نقوم بتقسيم هذه المستطيلات من أعلى (مركز الفئة)، ثم بعد ذلك نوصل نقاط التقسيم هذه بعضها مع بعض بخط مستقيم من خلال المسطرة لنحصل بالتالي على المضلع التكراري من خلال المدرج التكراري.
- **ولرسم المضلع من خلال مراكز الفئات** نقوم بإيجاد مركز الفئه لجميع فئات التوزيع التكرارى، ثم نقوم بتمثيل التكرار الأصلى المقابل لكل فئه بنقطه تناظر مركز هذه الفئه.
- نقوم برسم خط بإستخدام المسطره يصل كل نقتطتين متتاليتين، فنحصل على المضلع التكرارى.
- لإغلاق المضلع من الطرفين نقوم بإنشاء فئه سابقه عند النقطه الأولى في التوزيع التكرارى يقابلها تكرار أصلى يساوى الصفر، وكذلك إنشاء فئه لاحقه للفئه الأخيرة في التوزيع التكرارى يقابلها تكرار أصلى يساوى الصفر أيضاً، ونحسب مركز الفئة لكل منهما.

مثال:ـ

استخدم بيانات المثال السابق لرسم المضلع التكراري

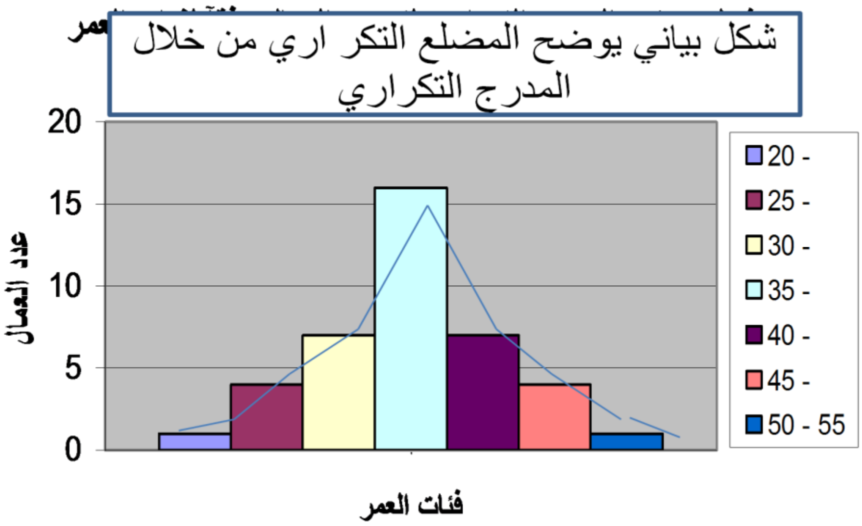
الحل:ـ

(١) نحصل على مراكز الفئات

فئات العمر	-20	-25	-30	-35	-40	-45	55-50	المجموع
مراكز الفئات	22.5	27.5	32.5	37.5	42.5	47.5	52.5	
عدد العمال	1	4	7	16	7	4	1	٤٠

(٢) استحداث فئتين سابقة ولاحقة للتوزيع وحساب مركز الفئه لكل منها، فمركز الفئة السابقه عن الفئه الاولى للتوزيع هو 17.5، وذلك باعتبار الفئة السابقة هى 15-20، وكذلك مركز الفئه اللاحقه للفئه الأخيره للتوزيع هو 57.5، وذلك باعتبار الفئة اللاحقة هى 60-55، والتكرار المقابل لكل مركز منهما يساوى الصفر.

(٣) نرسم محورين أفقي ورأسي بحيث يمثل المحور الأفقي الفئات والمحور الرأسي التكرارات، ثم بعد ذلك نرسم شكل المدرج التكراري ومن ثم نقوم بتقسيم مستطيلات المدرج التكراري من أعلى (مركز الفئة)، ثم بعد ذلك نوصل نقاط التقسيم هذه بعضها مع بعض بخط مستقيم من خلال المسطرة لنحصل بالتالي على المضلع التكراري من خلال المدرج التكراري.

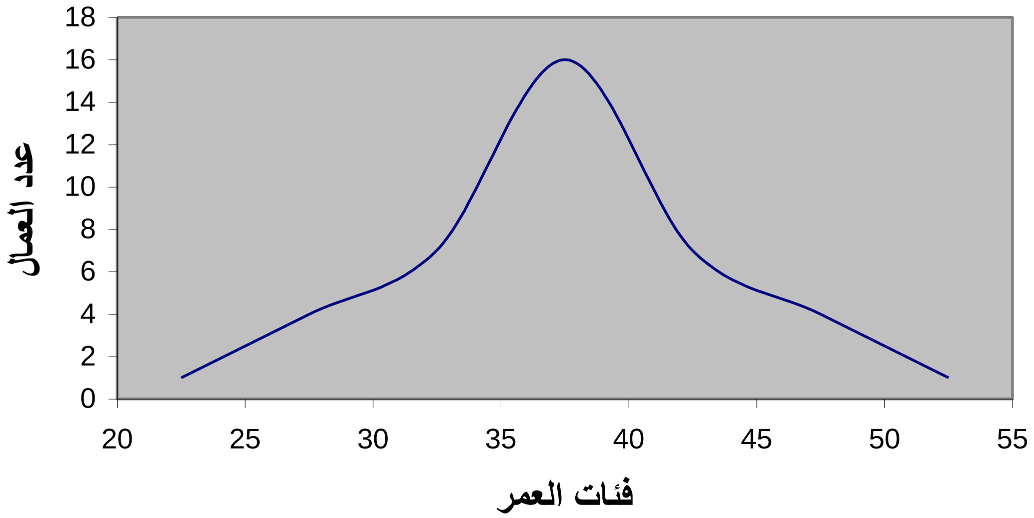


المنحى التكراري : إذا مهدنا المضلع التكراري وجعلناه منحنى بدلا من خطوط منكسرة فإننا نحصل على المنحنى التكراري، ويلاحظ أنه ينبغي عدم رسم المنحنى التكراري إلا إذا كانت الفئات كثيرة العدد، وذات طول صغير وكان عدد البيانات كبيرا وكانت هذه البيانات من النوع المتصل مثل الزمن والوزن.

خطوات رسم المنحنى التكراري:

- نرسم محورين أفقي ورأسي بحيث يمثل المحور الأفقي الفئات والمحور الرأسي التكرارات.
- نقوم بإنشاء فئه سابقه عند النقطه الأولى في التوزيع التكرارى يقابلها تكرار أصلى يساوى الصفر.
- نقوم بإنشاء فئه لاحقه للفئه الأخيرة في التوزيع التكرارى يقابلها تكرار أصلى يساوى الصفر أيضاً.
- إيجاد مركز الفئه لجميع فئات التوزيع التكرارى، ثم نقوم بتمثيل التكرار الأصلى المقابل لكل فئه بنقطه تناظر مركز هذه الفئه.
- نقوم برسم خط باليد دون إستخدام المسطره يصل كل نقتطتين متتاليتين، فنحصل على المنحنى التكرارى.

شكل يوضح المنحنى التكرارى لعدد العمال وفقآ لفئات العمر



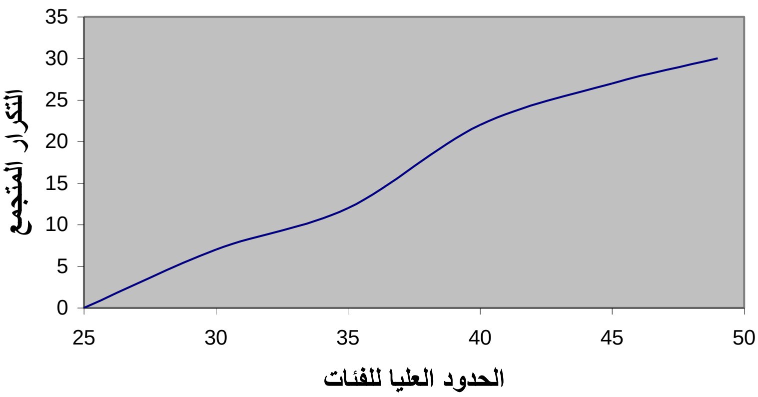
التوزيعات التكرارية المتجمعه:

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

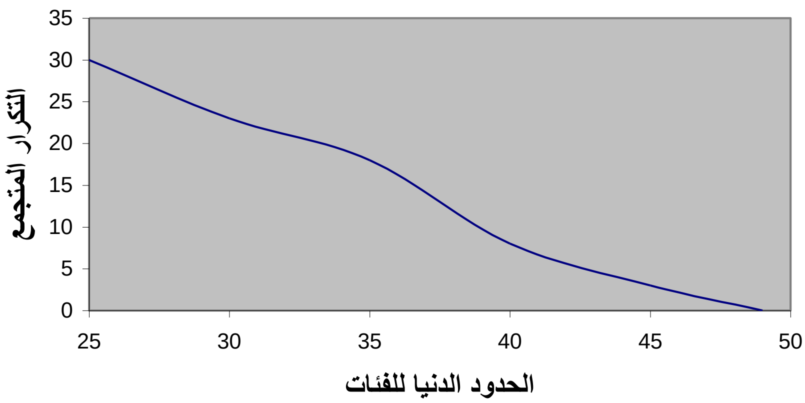
يستخدم المنحنى المتجمع الصاعد لتمثيل التوزيع التكرارى المتجمع الصاعد، سواء أكان بالقيم المطلقه للتكرارات، أو بالتكرار النسبي. ويراعي وضع النقاط الخاصه بالتكرارات في حالة المنحنى المتجمع الصاعد عند الحد الأعلى لكل فئه، لأنه يعبر عن العدد الاجمالى لأوجه الظاهرة الواقع أسفل الحد الأعلى للفئه.

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

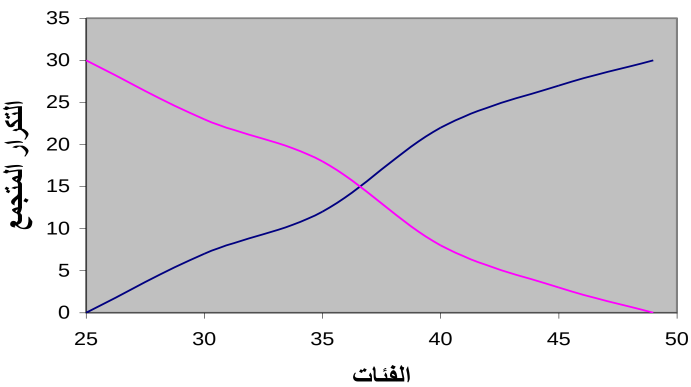
شكل يوضح المنحنى التكرارى المتجمع الصاعد



شكل يوضح المنحنى التكرارى المتجمع الهابط



شكل يوضح كلاّ من المنحنى التكرارى المتجمع الصاعد و الهابط

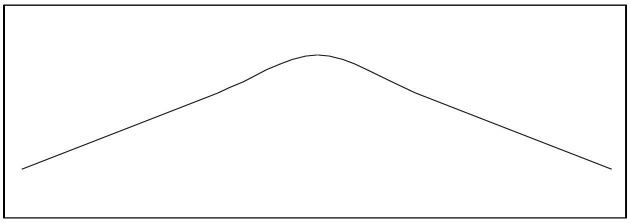


الأشكال الشائعة للتوزيعات التكرارية:-

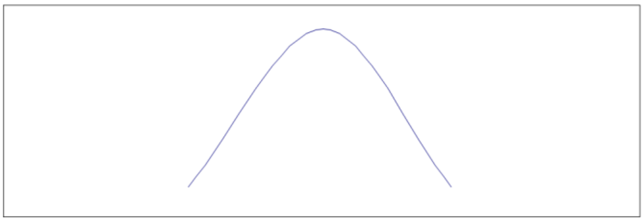
يعتبر **التوزيع الطبيعي** ذو شكل الجرس من التوزيعات التكراريه الهامه في دراستنا.

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

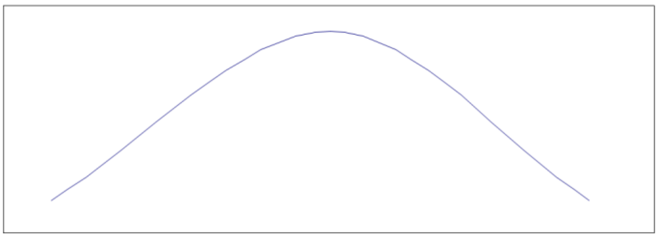
وقد يكون المنحنى التكرارى مسطح القمه بحيث تكون القمه واسعه ذو طرفين ضيقين نسبياً، فيسمى منحنى كبير التفرطح أو **المنحنى المفرطح**، وفيما يلي رسم بيانى يوضح كلا المنحنين المدبدب والمفرطح.



المنحنى المفرطح



المنحنى المدبدب



المنحنى الطبيعي

المحاضـــــره السابعه (الجزء الاول)

المقاييس الإحصائية للبيانات غير المبوبة

أولا: مقاييس النزعة المركزية

سبق و أن أستعرضنا مراحل البحث العلمى و أتضح لنا أن البحث الإحصائى له نفس المراحل فبعد جمع البيانات و المعلومات Data Collection لا بد من عرض هذة البيانات فى شكل جدولى او فى شكل الرسومات بيانية Data Presentation and Tabulation مما يسهل من فهم و أستيعاب مضمونها . و تأتى بعد ذلك المرحلة التالية وهى تحليل البيانات Data Analysis و التى فيها يتم أستخدام الأدوات الإحصائية المختلفة لوصف البيانات من خلال حساب المقاييس الإحصائية المختلفة التى سوف نستعرضها في هذه المحاضرة بمشيئة الله.

المقاييس الإحصائية:

تتمثل أهمية عملية وصف البيانات كمياً من خلال محاولة الوصول إلى فهم ورؤية أوضح للمعلومة المحتواه فى القيم الكمية للمتغيرات محل الدراسة ومحاولة التعبير عن تلك البيانات الكمية بقيم تصف طبيعة وشكل المتغيرات محل الدراسة بالطريقة التي تمكننا من التعامل معها بشكل أدق وأفضل ويطلق على تلك القيم المقاييس الإحصائية.

المقاييس الإحصائية لم توجد من تلقاء نفسها وأنما دعت الحاجة إلى وجودها حيث تساعدنا 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.

أقسام المقاييس الإحصائية :

تنقسم المقاييس الإحصائية إلى نوعيين رئيسيين هما:

- مقاييس النزعة المركزية Central Tendency Measures
- مقاييس التشتت أو الأنتشار Dispersion Measures_

فى هذه المحاضرة سنتعرض لكيفية حساب مقاييس النزعة المركزية ومقاييس التشتت فى حالة أستخدام البيانات الخام غير المبوبة، أي تلك التى لم يتم تصنيفها فى صورة جداول تكرارية، وذلك هو الأصل فى التحليل الإحصائي للبيانات، لأنة يعطى الصورة الحقيقية للنتائج بدون أي تدخل شخصى فيها، إلا ان ذلك لا يقلل إيضا من أهمية الحاجة لدراسة كيفية حساب المقاييس الإحصائية المختلفة من البيانات المبوبة والتى سنتعرض لها فى المحاضرات التالية إن شاء الله.

اولا- مقاييس النزعة المركزية Central Tendency Measures

نقصد بمقاييس النزعة المركزية تلك القيم الوسطى التى توضح القيمة التي تجمع أكبر عدد من القيم الخاصة بمجموعة معينة عندها . أو هي قيمة تلك الدرجة التي يمكن أن تعتبر ممثلة لكافة الدرجات الموجودة في تلك المجموعة . ولتحديد القيمة المتوسطة للتوزيع يوجد هناك عدة مقاييس أهمها :

- المتوسط الحسابي
- الوسيط
- المنوال (الشائع)

كما يوجد عدة مقاييس أخرى أقل شيوعا مثل:

- الوسط الهندسى
- الوسط التوافقى
- العشير
- المئين

أهمية حساب مقاييس النزعة المركزية :

حساب مقاييس النزعة المركزية يساعد على التالي:

- ايجاد ذلك الرقم المتوسط الذي يدل على خصائص أرقام مجموعة من المجموعات فيكفي أن ننظر الى ذلك الرقم المتوسط لنعرف الكثير عن خصائص هذه المجموعة من الأرقام
- المقارنة بين عدة مجموعات في وقت واحد ، فنقول أن هذه المجموعة أقوى من تلك ، وذلك اعتمادا على مقارنة هذه المتوسطات بعضها ببعض

الوسط الحسابى (المتوسط) Mean

يُعرف المتوسط الحسابي بأنه قيمة التى اذا أعطيت لكل مفرد من مفردات الظاهرة لكان مجموع القيم الجديدة مساويا للمجموع الفعلي للقيم الأصلية للظاهرة، أي أن الوسط الحسابي يساوي مجموع القيم مقسوما على عددها، ويتم حساب المتوسط الحسابي للبيانات غير المبوبة من خلال المعادلة التالية:

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

مثال:

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

الشهر	محرم	صفر	ربيع أول	ربيع ثاني	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذي القعده	ذي الحجه
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب:

حساب المتوسط الحسابى للمبيعات الشهرية.

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.
- ومن أهم خصائص الوسط الحسابى هو تأثره بجميع العمليات الجبرية تجرى على البيانات من إضافة قيمة لجميع البيانات أو طرحها أو ضربها أو قسمتها.

مثال:

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

3 , 5 , 2, 7,3

المطلوب:

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

1- زيادة اجور العاملين بمقدار 2000 ريال

2- زيادة أجور العاملين بنسبة 5 %

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.
- لايمكن ايجاده من خلال الرسم

الوسيط Median

هو الدرجة التي تتوسط مجموعة من الدرجات المرتبة ترتيبا تصاعديا أو تنازليا، أى هو القيمة التى يصغرها عدد من القيم يساوى العدد الذى يكبر هذة القيمة ويمكن حساب الوسيط باتباع الخطوات التالية:

- ترتيب الدرجات تصاعديا أو تنازليا
- إيجاد ترتيب الوسيط و يقصد بة إيجاد مكان الوسيط، ويختلف ترتيب الوسيط إذ كان عدد المشاهدات فردى أو زوجي كما يلي:

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

- - إيجاد قيمة الوسيط.

مثال:

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام 1427 هـ بلألف ريال كما يلى:

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

المطلوب:

إيجاد قيمة الوسيط للبيانات السابقة.

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

هو القيمة التي تعتبر اكثر القيم شيوعا، وعلى ذلك فتحديده يتوقف على تكرار القيم في المجموعة.

فى نفس المثال السابق للمبيعات الشهرية . أحسب المنوال؟

نجد أن المبيعات الأكثر تكراراً هنا هى 3 ألف ريال لذلك

فان المنوال هنا = 3

وقد يكون في التوزيع منوالين أو أكثر وذلك كالمثال الأتي:

٦ ، 5 ، 5 ، 5 ، 4 ، 4 ، 4

فالمنوال هنا = 4 ، 5 أي أنه يوجد منوالين .

وقد لا يكون في التوزيع منوال وذلك كالمثال الآتي:

2 ، 5 ، 7 ، 9 ، 11

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

الوسط الهندسى Geometric Mean

نتيجة أن الوسط الحسابى يتأثر بالقيم الشاذة دعت الحاجة إلى وجود مقاييس لا تتأثر بقدر الإمكان بالقيم الشاذة والمتطرفة ومن تلك المقاييس الوسط الهندسى GM والذى يكون مفيد فى بعض التطبيقات الاقتصادية ودراسات نمو الظواهر الديموجرافية وكذلك فى حساب الأرقام القياسية، فالوسط الهندسى هو الجذر النونى لحاصل ضرب القيم محل الدراسة ويمكن حسابة من خلال المعادلة التالية:

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

ويحسب الجذر النوني من خلال استخدام الآلة الحاسبة العلمية بكتابة حاصل ضرب القيم محل الدراسة ثم الضغط على الجذر النوني ثم إدخال قيمة n ثم الضغط على يساوي فتظهر بالتالي قيمة الوسط الهندسي.

مثال :

البيانات تعبر عن المبيعات الشهرية لأحد المحال التجارية خلال عام1427 هـ بلألف ريال كما يلى:

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

المطلوب:

إيجاد قيمة الوسط الهندسي للبيانات السابقة.

الحل:

يمكن تطبيق المعادلة السابقة على البيانات الموجودة بالمثال ولكن قد يكون الأمر صعب فى حالة ما تكون المشاهدات محل الدراسة (n) كبيرة الحجم . لذا يمكن حساب الوسط الهندسي كما يلى:

$$GM = \sqrt[12]{3\times5\times8\times3\times6\times4\times12\times5\times4\times3\times7\times9} =$$

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

خواص الوسط الهندسى:

١. يعطى نتائج أكثر اعتدالا من المتوسط الحسابى
٢. تتوقف قيمتة على سائر القيم دون استثناء أو استبعاد، شأنه شأن الوسط الحسابى
٣. أقل تأثرا بالقيم المتطرفة عن الوسط الحسابى

مزايا وعيوب الوسط الهندسي:

المزايا :

١. أكثر تمثيلا للقيم عن الوسط الحسابى بأعتبار أنة لا يتأثر بالقيم المتطرفة بنفس درجة الوسط الحسابى

٢. يعتبر من أنسب المقاييس لحساب متوسطات النسب ومعدلات النمو

٣. يعتبر من أكثر مقاييس النزعة المركزية ملائمة لحساب الأرقام القياسية للمناسيب

العيوب :

١. لا يمكن حسابه اذا كانت احدى القيم صفر

٢. لا يمكن استخدامه اذا كان ناتج حاصل ضرب قيم المشاهدات محل الدراسة سالب

٣. صعوبة حسابه يدويا وإنما يمكن ذلك بأستخدام الحاسب الألى (الآلة الحاسبة).

تابع المحاضره السابعه (الجزء الثاني)

تابع ... المقاييس الإحصائية للبيانات غير المبوبة

ثانيا: مقاييس التشتت أو الأنتشار Dispersion Measures

كما تميل القيم الى التمركز فانها تميل أيضا إلى التشتت أو الانتشار، فبالتالي فان أي توزيع من القيم له صفة التمركز، وصفة التشتت.

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

مثال

مجموعة (أ) : 8 ، 8 ، 8 ، 8 ، 8

مجموعة (ب) : 1 ، 2 ، 3 ، 5 ، 6

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

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

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): (45 ، 50 ، 55) المتوسط هنا = 50

مجموعة (ب): (30 ، 50 ، 70) المتوسط هنا = 50

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 (50) .

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) = 55 - 45 = 10

مدى مجموعة (ب) = 70 – 30 = 40

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

المدى Range

المدى هو الفرق بين أعلى درجة وأقل درجة في التوزيع.

ويعتبر المدى الوسيلة المباشرة لمعرفة مدى تقارب القيم أو تباعدها في أي توزيع، وهو وسيلة سهلة، إلا أنها أقل الوسائل دقة وذلك لأن حسابه يتوقف على قيمتين فقط من قيم المجموعة، ولا يهتم مطلقا بما بينهما من قيم أخرى.

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

مثال:

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

الشهر	محرم	صفر	ربيع أول	ربيع ثاني	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذي القعده	ذي الحجه
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب:

حساب المدى للمبيعات الشهرية.

الحل:

نلاحظ أن أكبر قيمة هى 12 وأقل قيمة للمبيعات الشهرية هى 3 لذلك يكون المدى 9

$$Range=12-3=9$$

عيوب المدى:

نجد أن من أهم عيوب المدى أنة يتم حسابة بناءا على أكبر و أصغر قيمتين وبالتالى فى حالة كونهما أو أحدهما متطرفتين أو قيم شاذة فأن المدى يعطى نتائج مضللة.

- متوسط الانحرافات المطلقة Average Absolute Deviation

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

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

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

ويمكن حساب متوسط الأنحرافات المطلقة من خلال المعادلة التالية :

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

مثال :

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

الشهر	محرم	صفر	ربيع أول	ربيع ثاني	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذي القعده	ذي الحجه
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب:

أحسب متوسط الأنحرافات المطلقة للمبيعات الشهرية.

- التباين والانحراف المعياري:

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

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

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

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

حساب التباين والانحراف المعياري :

- يمكن حساب التباين من خلال المعادلة التالية:

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

- بالتالى يكون حساب الإنحراف المعياري كما يلي:

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

مثال :

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

الشهر	محرم	صفر	ربيع أول	ربيع ثاني	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذو القعده	ذو الحجه
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب:

أحسب قيمة التباين وقيمة الإنحراف المعياري للمبيعات الشهرية.

ملاحظة هامة:

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

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

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

مثال :

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

الشهر	محرم	صفر	ربيع أول	ربيع ثاني	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذو القعده	ذو الحجه
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب:

فإذا تم طرح 2 من جميع بيانات المبيعات الشهرية أى تم تخفيض المبيعات الشهرية بمقدار 2 أحسب قيمة الانحراف المعيارى الجديد؟

نلاحظ عدم تغير قيمة الانحراف المعيارى وإنما ظلت قيمتة كما هى بالرغم من طرح مقدار ثابت 2 من جميع قيم المبيعات الشهرية.

مثال :

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

الشهر	محرم	صفر	ربيع أول	ربيع ثاني	جمادى أول	جمادى الاخر	رجب	شعبان	رمضان	شوال	ذو القعده	ذو الحجه
المبيعات	3	5	8	3	6	4	12	5	4	3	7	9

المطلوب:

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

نلاحظ تغير قيمة الانحراف المعيارى وهى نفس قيمة الانحراف المعيارى القديمة مضروبة فى 3

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

المتوسط	الوسيط	المنوال	الوسط الهندسى
5.75	5	3	5.20114

المدى	متوسط الانحرافات المطلقة	التباين	الانحراف المعيارى
9	2.20833	7.840909	2.80016

المحاضره الثامنه " الجزء الاول "

المقاييس الإحصائية للبيانات المبوبة

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

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

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

- الجداول المنتظمة
- الجداول غبر المنتظمة
- الجداول المفتوحة

الجداول المنتظمة:

- وهى تلك الجداول التى تكون فيها أطوال الفئات جميعها متساوية .

اولا- الوسط الحسابي والتشتت حولة:

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

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

$\overline{x}$ الوسط الحسابي

x_i مركز الفئة i وهى تساوى (الحد الأعلى للفئة + الحد الأدنى للفئة) ÷ 2

f_i تكرار الفئة i

l عدد الفئات

ويتم حساب التشتت حول المتوسط الحسابى من خلال الآتي:

أ- متوسط الانحرافات المطلقة *AAD*:

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

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

ب – التباين σ^2 :

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

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

جـ - الانحراف المعياري σ :

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

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

مثال:

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

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

المطلوب: حساب التالي:

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

الحل:

يتم إعداد الجدول التالي حتى يمكن حساب كلاً من الوسط الحسابي والتباين والانحراف المعيارى:

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

- الوسط الحسابي:

يمكن إيجاد الوسط الحسابى بتطبيق المعادلة الخاصة بذلك كما يلى:

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

- التباين:

يمكن الحصول على التباين باستخدام المعادلة الخاصة بذلك كما يلى:

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

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

- الانحراف المعياري:

يمكن حسابة باستخدام المعادلة الخاصة بذلك كما يلى :

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

- متوسط الانحرافات المطلقة *AAD*:

حتى يمكن إيجاد متوسط الانحرافات المطلقة لابد اولاً من إيجاد الانحرافات عن الوسط الحسابى ثم استخدامها فى الحساب كما يتضح ذلك من الجدول التالى:

$\lvert x-\bar{x}\rvert f$	$(x-\bar{x})f$	$x-\bar{x}$	مركز الفئة x	التكرار f	فئات العمر
172.7273	-172.727	-17.2727	25	10	20 -
218.1818	-218.182	-7.27273	35	30	30 -
136.3636	136.3636	2.727273	45	50	40 -
254.5455	254.5455	12.72727	55	20	50 - 60
781.8182	0			110	المجموع
$\sum\lvert x-\bar{x}\rvert f$	$\sum(x-\bar{x})f$				

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

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

فيتضح لنا من الجدول السابق أن:

مجموع انحرافات القيم عن وسطها الحسابي تسـاوى صفر حيث أن

$$\sum(x-\bar{x})f=0$$

كما يمكن الاعتماد على هذه الانحرافات فى حساب التباين بتطبيق المعادلة الخاصة بذلك.

تابع ... المقاييس الإحصائية للبيانات المبوبة

ثانيا – الوسيط و التشتت حوله:

الوسيط هو القيمة التى يصغرها عدد من القيم يتساوى مع العدد الذى يكبر هذه القيمة، و هو يعتبر أحد مقاييس النزعة المركزية التى نلجأ إليها لتحليل الظواهر وفقا لشكل التوزيع الإحصائى محل الدراسة.

ولحساب الوسيط من البيانات المبوبة هناك ثلاث خطوات يجب إتباعها و هى:

- 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 الادنى لبداية الفئة الوسيطة
k_{Med}	ترتيب الوسيط
F_a	التكرار المتجمع السابق للفئة الوسيطة
F_b	التكرار المتجمع اللاحق للفئة الوسيطة
I	طول الفئة الوسيطة

مثال :ـ

فى بيانات المثال السابق توزيع مجموعة من المدرسين العاملين فى مجال التربية وفقا لفئات أعمارهم، أحسب قيمة الوسيط؟

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

الحل :ـ

- إيجاد الجدول التكرارى المتجمع الصاعد كما يلي:

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

- إيجاد ترتيب الوسيط كالتالي:

$$k_{Med} = 110/2 = 55$$

- إيجاد قيمة الوسيط من خلال التالي:

نلاحظ أن ترتيب الوسيط = 55 ، مما يعنى أن ترتيب الوسيط يقع بين التكرار المتجمع الصاعد (40) وهو المقابل للحد الأعلى للفئة 40 والتكرار المتجمع الصاعد (90) وهو المقابل للحد الأعلى للفئة 50

أي أن الحد الادنى للفئة هو L Med = 40

وبالتالى يكون طول الفئة الوسيطية هو:

$$I = 50 - 40 = 10$$

وعلى ذلك يمكن حساب قيمة الوسيط كما يلي:

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

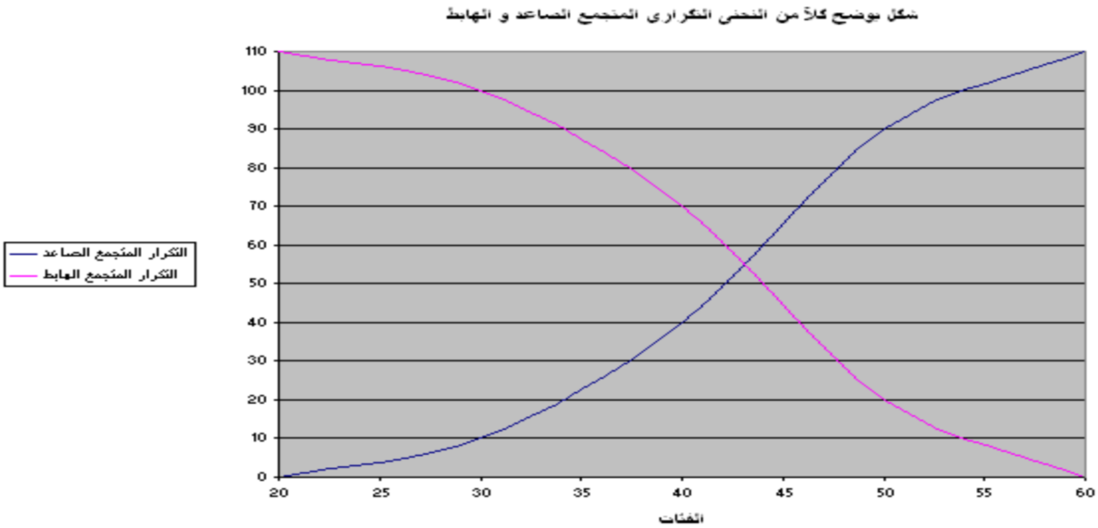
$$Med = 40 + \frac{55-40}{90-40} \times 10 = 43$$

كما يمكن إيجاد الوسيط عن طريق رسم كلا من المنحنى التكراري المتجمع الهابط والمنحنى التكراري المتجمع الصاعد كما يلى :

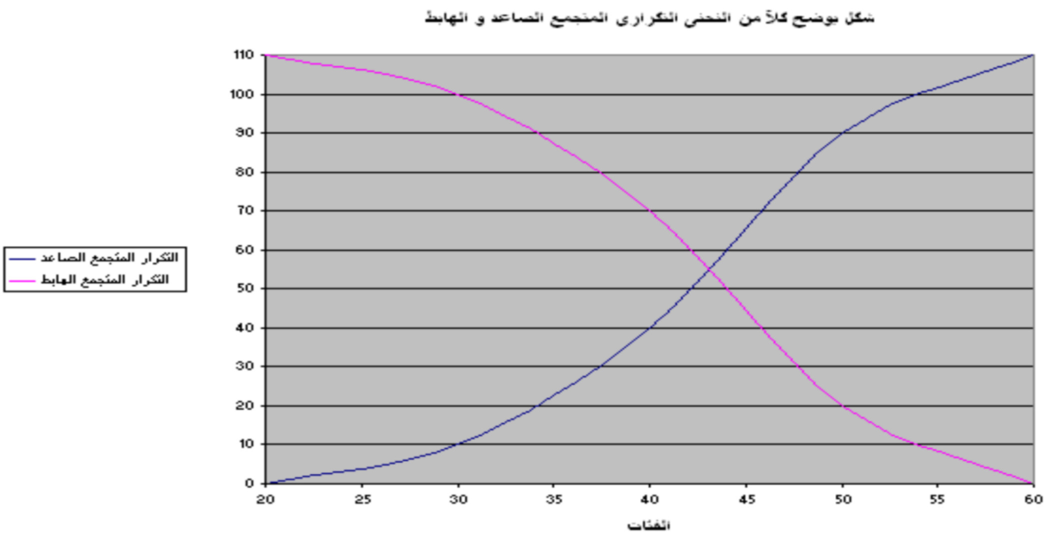
لقد قمنا في أثناء حل المثال السابق بحساب التكرار المتجمع الصاعد، ونقوم الآن بإيجاد التكرار المتجمع الهابط كما يلى:

التكرار المتجمع الهابط	الحدود الدنيا للفئات
110	20 فأكثر
100	30 فأكثر
70	40 فأكثر
20	50 فأكثر
صفر	60 فأكثر

ثم بعد ذلك نقوم برسم كلا من المنحنى التكراري المتجمع الصاعد والمنحنى التكرارى المتجمع الهابط معا كما يلى:



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



و يتضح لنا من الشكل السابق أن الوسيط تبلغ قيمتة 43 تقريبا

الرُبيع الادنى (الأول):

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

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

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

الرُبيع الاعلى (الثالث):

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

لذلك يتم حسابه كما فى حالة الوسيط مع اختلاف أن ترتيب الرُبيع الثالث Q3 هو (3 n / 4)

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

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

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

مثال :ـ

من بيانات المثال السابق أحسب كلا من الرُبيع الاول والرُبيع الثالث؟

حساب الرُبيع الادنى (الاول) Q1 :

الجدول المتجمع الصاعد تم إعداده سابقا

- إيجاد ترتيب الرُبيع الاول كالتالي:

$$k_{Q1} = n/4 = 110/4 = 27.5$$

- إيجاد قيمة الرُبيع الادنى (الاول) Q1 كالتالي:

نلاحظ أن ترتيب الرُبيع الادنى (الاول) Q1 27.5 مما يعنى أن الرُبيع الادنى (الاول) Q1 يقع بين التكرار المتجمع الصاعد (10 F_a) وهو المقابل للحد الأعلى للفئة 30 و التكرار المتجمع الصاعد (40 F_b) وهو المقابل للحد الأعلى للفئة 40 والحد الادنى للفئة هو = (30 L_{Q1}) .

وبالتالى يكون طول فئة الرُبيع الادنى (الاول) Q1

I_{Q_1} = 40 – 30 = 10

وعلى ذلك يمكن حساب قيمة الرُبيع الادنى (الاول) Q1 من خلال المعادلة التالية كما يلى :

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

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

حساب الرُبيع الاعلى (الثالث) Q3 :

الجدول المتجمع الصاعد تم إعداده سابقا

نقوم بإيجاد ترتيب الرُبيع الاعلى (الثالث) Q3 كالتالي:

$$k_{Q3} = 3(n)/4 = (3)110/4 = 82.5$$

إيجاد قيمة الرُبيع الاعلى (الثالث) Q3 ، ونلاحظ أن ترتيب الرُبيع الاعلى (الثالث) Q3 82.5 مما يعنى أن الرُبيع الاعلى (الثالث) Q3 يقع بين التكرار المتجمع الصاعد (40 F_a) وهو المقابل للحد الأعلى للفئة 40 والتكرار المتجمع الصاعد (90 F_b) وهو المقابل للحد الأعلى للفئة 50 والحد الادنى للفئة هو = (40 L_{Q3}) .

وبالتالى يكون طول فئة الربيع الاعلى (الثالث)

I_{Q_3} = 50 – 40 = 10

وعلى ذلك يمكن حساب قيمة الربيع الاعلى (الثالث) Q3 كما يلي:

$$Q_3 = \mathrm{L}_{\mathrm{Q}_3} + \frac{\frac{3(n)}{4} - F_a}{F_b - F_a} \times I_{Q_3}$$

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

حساب قيمة العُشير :

وبنفس الطريقة السابقة يمكن الحصول على العشير $P_{0.10}$ وهو القيمة التى يكون قبلها 10 % من مفردات المجتمع و 90 % منها أكبر منه. والاختلاف يكون فقط فى الترتيب حيث أن ترتيب العشير هو:

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

ففى مثالنا هذا يكون ترتيب العشير هو :

$$k_{P_{0.10}} = n/10 = 110/10 = 11$$

ونلاحظ أن ترتيب العُشير $P_{0.10}$ هو 11 مما يعنى أن العُشير $P_{0.10}$ يقع بين التكرار المتجمع الصاعد (10 F_a) وهو المقابل للحد الأعلى للفئة 30 والتكرار المتجمع الصاعد (40 F_b) وهو المقابل للحد الأعلى للفئة 40 والحد الادنى للفئة هو = (30 $L_{P_{0.10}}$) .

وبالتالى يكون طول فئة العُشير :

$$I_{P_{0.10}} = 40 - 30 = 10$$

وعلى ذلك يمكن حساب قيمة العُشير $P_{0.10}$ كما يلى :

$$P_{0.10} = L_{P_{0.10}} + \frac{\frac{n}{10} - F_a}{F_b - F_a} \times I_{P_{0.10}}$$

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

حساب قيمة المئويين أو المئين $P_{0.01}$:

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

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

ففى مثالنا هذا يكون ترتيب المئويين $P_{0.01}$ هو :

$$k_{P_{0.01}} = n/100 = 110/100 = 1.1$$

ونلاحظ أن ترتيب المئويين $P_{0.01}$ هو 1.1 مما يعنى أن المئويين $P_{0.01}$ يقع بين التكرار المتجمع الصاعد (صفر F_a) وهو المقابل للحد الأعلى للفئة 20 والتكرار المتجمع الصاعد (10 F_b) وهو المقابل للحد الأعلى للفئة 30 والحد الادنى للفئة هو = (20 $L_{P_{0.01}}$).

وبالتالى يكون طول فئة المئويين

$$I_{P_{0.01}} = 30 - 20 = 10$$

وعلى ذلك يمكن حساب قيمة المئويين كما يلى :

$$P_{0.01} = \mathrm{L}_{P_{0.01}} + \frac{\frac{n}{100} - F_a}{F_b - F_a} \times I_{P_{0.01}}$$

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

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	21.1	30.333	35.8333	43	48.5

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 الربيعى IQR) والذى يستبعد القيم المتطرفة من الطرفين، حيث يعتمد فى حسابه على كلا من الرُبيع الأول Q1 والرُبيع الثالث Q3 ويتم حسابه من خلال المعادلة التالية:

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

وبالتالى يكون نصف المدى الرُبيعي فى مثالنا هو:

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

ثالثا: المنوال :

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

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

قيمة المنوال	Mod
الحد الأدنى لفئة المنوال	L_{Mod}
يساوى تكرار فئة المنوال – تكرار الفئة السابقة	$D1$
يساوى تكرار فئة المنوال – تكرار الفئة اللاحقة	$D2$
طول الفئة المنوالية	I

أحسب المنوال للأعمار مجموعة من المدرسين العاملين فى مجال التربية فى المثال السابق؟

الحل:

بيانات توزيع العمال وفقا لفئات العمر كانت كما يلى :

الحد الادنى L_{Mod}

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

D1 = 30

الفئه المنواليه

D2 = 20

نلاحظ أن أكبر تكرار هو 50 و يكون مقابل للفئة " 40 – 50 " لذلك يطلق عليها الفئة المنوالية و من ثم فأن الحد الادنى لها هو (40 L_{Mod}) وطول الفئة هو (10 I).

كما يمكن حساب كلا D1 و D2 كالتالي:

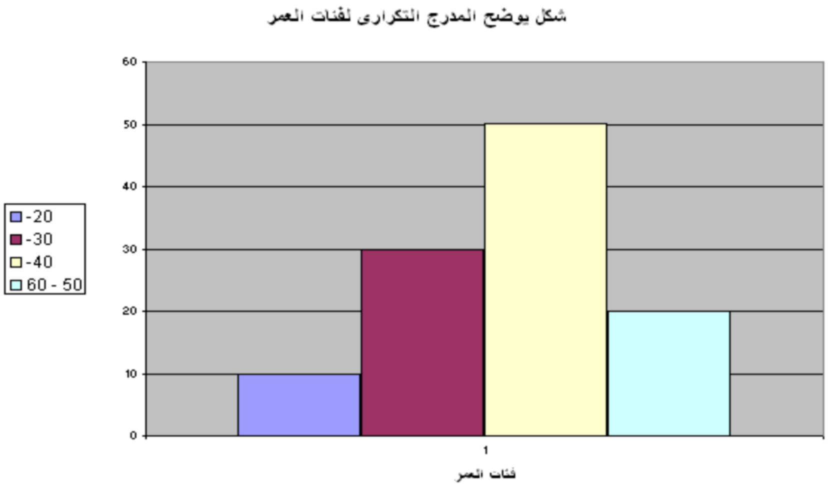
$$D1 = 50 - 30 = 20$$

$$D2 = 50 - 20 = 30$$

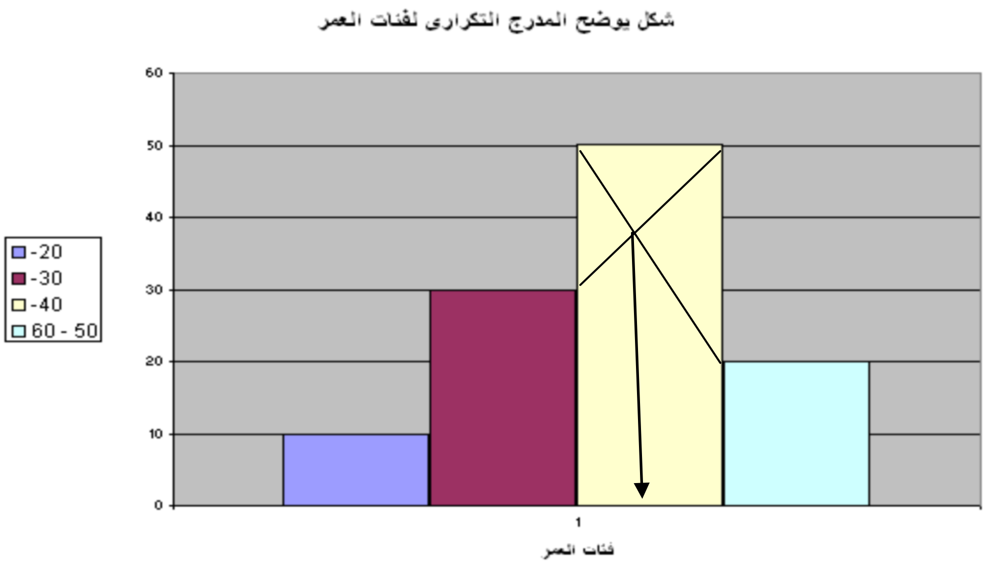
وبالتالى يمكن حساب قيمة المنوال Mod كالتالي:

$$Mod = 40 + \frac{20}{20+30} \times 10 = 44$$

كما يمكن إيجاد المنوال بيانيا، ويتم ذلك عن طريق رسم المدرج التكرارى كما يلي:



ثم نأتى على أعلى مستطيل الذى يمثل أكبر تكرار ونصل بداية الفئة المنوالية ببداية الفئة التالية ونهاية الفئة المنوالية بنهاية الفئة السابقة عليها فيتقاطع الخطان عند نقطة نسقط منها عامودا على المحور الأفقى فيلتقى عند نقطة تكون هي قيمة المنوال كما يتضح من الشكل التالى:



الجداول غير المنتظمة:

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

ويتم حساب المقاييس الإحصائية التى سبق عرضها فى حالة الجداول المنتظمة بنفس الطريقة فيما عدا المنوال،

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

التكرار المعدل = التكرار الأصلى للفئة ÷ طول الفئة

مثال:

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

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

المطلوب حساب:

1- الوسط الحسابي
2- متوسط الانحرافات المطلقة
3- التباين
4- الانحراف المعياري
5- الوسيط
6- الرُبيع الأول
7- الرُبيع الثالث
8- العُشير
9- المئويين
10- نصف المدى الرُبيعى
11- المنوال

الحل

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

ولحساب المنوال فى هذه الحالة فإنه لا يتم الاعتماد على بيانات الفئات الاصلية وإنما يتم إيجاد التكرار المعدل بقسمة تكرار كل فئة على طولها كما يلي:

فئات الدخل	التكرار f	طول الفئة	التكرار المعدل
3 -	20	2	10
5 -	50	3	16.6667
8 -	15	2	7.5
10 -15	15	5	3
المجموع			

نلاحظ أن أكبر تكرار معدل هو 16.6667 و يكون مقابل للفئة " 5 - 8 " لذلك يطلق عليها الفئة المنوالية و من ثم فإن الحد الادنى لها L_{Mod} هو 5 طول الفئة I هو 3.

كما يمكن حساب كلا من D1 و D2 كالتالي:

$$D1 = 16.66667 - 10 = 6.66667$$

$$D2 = 16.66667 - 7.5 = 9.16667$$

وبالتالى يمكن حساب قيمة المنوال Mod من خلال تطبيق معادلة حساب المنوال مع الأخذ في الاعتبار التكرار المعدل كالتالي:

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

الجداول المفتوحة :

وهى ذلك النوع من الجداول التى يكون فيها الحد الادنى للفئة الاولى غير محدد أو الحد الأعلى للفئة الأخيره غير محدد أو كلاهما. وفى هذا النوع من الجداول يصعب حساب الوسط الحسابى والتباين والانحراف المعيارى، حيث لا يمكن تحديد مركز الفئة للفئات المفتوحة، لذا فيعبير 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 50	50 -	60 -	70 -	80 فأكثر
عدد الطلاب	5	10	35	15	10

المطلوب:

حساب مقاييس النزعة المركزية والتشتت المناسبة ؟

الحل :

موجود في الكتاب صفحة 144 و 145

المحـــــــاضره التاسعه (الجزء الأول)

مقاييس التشتت النسبى والإلتواء والتفلطح

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

لذلك سيتم فى هذا الفصل دراسة كلا من:

- مقاييس التشتت النسبى
- القيمة المعيارية
- الإلتواء
- التفلطح

اولا – مقاييس التشتت النسبى Coefficient of Variation

يستخدم هذا النوع من المقاييس لمقارنة تشتت مجموعتين من البيانات او ظاهرتين او توزيعين، وفى تلك الحالة لا يصلح مقانة التباين او الانحراف المعيارى لكلا المجموعتين، حيث يكون لها وحدات قياس تختلف على حسب طبيعة الظاهرة محل الدراسة.

لذا فى حالة الرغبة فى المقارنة بين التشتت لظاهرتين أو أكثر فإنة يتم الاعتماد فى عملية المقارنة على مقاييس التشتت النسبى Coefficient of variations (c.v.) والتى يعبر عنها من خلال معامل الاختلاف المعيارى والذى يمكن حسابة بالاعتماد على كلا من الوسط الحسابى والانحراف المعيارى حيث أن

معامل الاختلاف = الانحراف المعيارى ÷ الوسط الحسابى

في حالة الاعتماد على بيانات العينة يتم حساب معامل الاختلاف من خلال المعادلة:

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

أما اذا كانت البيانات المتاحة من جداول تكرارية (بيانات مبوبة) فيمكن الاعتماد على معامل الاختلاف الربيعى المعيارى والذى يعتمد فى حسابة على الوسيط والربيع الادنى والربيع الأعلى وخاصة فى حالة الجداول المفتوحة حيث أن:

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

معادلات حساب كل من الوسيط Med والربيع الأول Q1 والربيع الثالث Q3

الوسيط *Med*:

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

Med	قيمة الوسيط
L_{Med}	الحد الادنى لبداية الفئة الوسيطية
k_{Med}	ترتيب الوسيط
F_a	
F_b	التكرار المتجمع اللاحق للفئة الوسيطية
I_{Med_1}	طول الفئة الوسيطية

معادلات حساب كل من الوسيط Med والربيع الأول Q1 والربيع الثالث Q3

الربيع الأول Q1:

$$Q_1 = \mathrm{L}_{\mathrm{Q}_1} + \frac{\frac{n}{4} - F_a}{F_b - F_a} \times I_{Q_1}$$

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

Q_3

معادلات حساب كل من الوسيط Med والربيع الأول Q1 والربيع الثالث Q3

الربيع الثالث Q3:

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

Q_3	قيمة الربيع الأدنى أو الثالث
L_{Q3}	الحد الادنى لبداية الفئة الربيعية الثالثة
k_{Q_3}	ترتيب الربيع الثالث
F_a	التكرار المتجمع السابق للفئة الربيعية الثالثة
F_b	التكرار المتجمع اللاحق للفئة الربيعية الثالثة
I_{Q_3}	طول الفئة الربيعية الثالثة

مثال :

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

الايجار بالألف ريال	6-	10 -	12-	14- 18
عدد الوحدات السكنية	15	20	12	13

المطلوب:-

حساب :

- معامل الاختلاف للإيجار السنوى
- معامل الاختلاف الربيعى للإيجار السنوى

الحل :

- حساب معامل الاختلاف للإيجار السنوي

يتم إعداد الجدول التالي حتى يمكن حساب كلاً من الوسط الحسابى والتباين والانحراف المعيارى.

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

الوسط الحسابى:

يمكن إيجاد الوسط الحسابى كما يلي:

$$\bar{x} = \frac{\sum_{i=1}^{l} x_i f_i}{\sum_{i=1}^{l} f_i} = \frac{704}{60} = 11.733$$

التباين:

يمكن الحصول على التباين كما يلي:

$$S^2 = \frac{\sum_{i=1}^{l} \left[f(x - \bar{x})^2 \right]}{\sum_{i=1}^{l} f_i}$$

$$S^2 = \frac{475.66}{60} = 7.93$$

الانحراف المعيارى

يمكن حسابة كما يلي:

$$S = \sqrt{7.93} = 2.816$$

وبذلك يمكن حساب معامل الاختلاف V. C. ليكون كما يلي:

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

$$= \frac{2.816}{11.733} \times 100 = 24\%$$

أى أن معامل الاختلاف للإيجار السنوى للوحدات السكنية بلغ 24 %

حساب معامل الاختلاف الربيعي:

وحتى يمكن حسابة لابد من حساب كلا من:

- الربيع الاعلى
- الربيع الادنى
- الوسيط

وحتى يتسنى لنا حساب ذلك لابد من إعداد جدول التكرار المتجمع الصاعد كما يلي:

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

إيجاد الرتبة:

	Q1	**Med**	**Q3**
الرتبة	$k_{Q1}=n/4$ $=60/4$ $=15$	$k_{Med}=n/2$ $=60/2$ $=30$	$k_{Q3}=3n/4$ $=3(60)/4$ $=45$

إيجاد القيمة:

أ- الوسيط

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

$$Med = 10 + \frac{30-15}{35-15} \times 2 = 11.5$$

ب- الربيع الأدنى (الأول):

نلاحظ أن رتبة الربيع الأول 15 ويوجد تكرار متجمع صاعد نفسه 15 أمام الحد الأعلى للفئة 10 لذلك لا يتم تطبيق قانون الربيع الأول وإنما نحصل على قيمة الربيع الأول مباشرة وهي:

$$Q_1 = 10$$

ج- الربيع الأعلى (الثالث):

$$Q_3 = 12 + \frac{45-35}{47-35} \times 2 = 13.6667$$

وبذلك يمكن حساب معامل الاختلاف الربيعى كما يلى:

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

$$= \frac{13.6667\text{-}10}{13.667 + 10} \times 100 = 15.494\%$$

ويتضح لنا مما سبق أن معامل الاختلاف الربيعى للإيجار السنوى للوحدات السكنية بلغ 15.494 % .

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

ثانيا :القيمة المعيارية Standardized values

وهي تلك القيمة التي تقيس مدى انحراف قيمة مفردة ما من مفردات الدراسة عن الوسط الحسابى لها وذلك بوحدات من الانحراف المعيارى. ويشار إلى المتغير الذى يعبر عن القيم المعيارية بالمتغير المعيارى. ويرمز للقيمة المعيارية بالرمز Z حيث أن:

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

وبالتالى يمكن الاعتماد على القيمة المعيارية فى المقارنة بين القيم المطلقة للظواهر المختلفة

مثال:-

حصل أحد الطلاب فى مقرر المحاسبة على (80) درجة حيث بلغ متوسط درجات الطلاب فى أختبار المحاسبة (83) درجة بإنحراف معيارى (5). بينما حصل فى اختبار مقرر الرياضيات على (70) درجة حيث بلغ متوسط درجة الطلاب فى اختبار الرياضيات (65) درجة بأنحراف معيارى قدرة (5) درجات . هل يمكن القول بأن درجات الطالب فى مقرر المحاسبة أفضل من درجته فى مقرر الرياضيات ؟

الحل :

للحكم على مدى أفضلية الدرجة التى حصل عليها الطالب فى أى من المقررين يجب حساب القيمة المعيارية لكل منهما كما يلى:

القيمة المعيارية لدرجة الطالب فى مقرر المحاسبة هى:

$$z_1 = \frac{x-\bar{x}}{S} = \frac{80-83}{5} = -0.6$$

القيمة المعيارية لدرجة الطالب فى مقرر الرياضيات هي:

$$z_2 = \frac{70-65}{5} = 1$$

يتضح لنا مما سبق 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 .

ويدل ذلك على أنه من الناحية الظاهرية قد تبدو درجة الطالب فى مقرر المحاسبة أفضل إلا أنه فى حقيقة الأمر أن مستوى الطالب فى مقرر الرياضيات هو الأفضل.

المحاضره التاسعه (الجزء الثاني)

تابع .... مقاييس التشتت النسبى والإلتواء والتفلطح

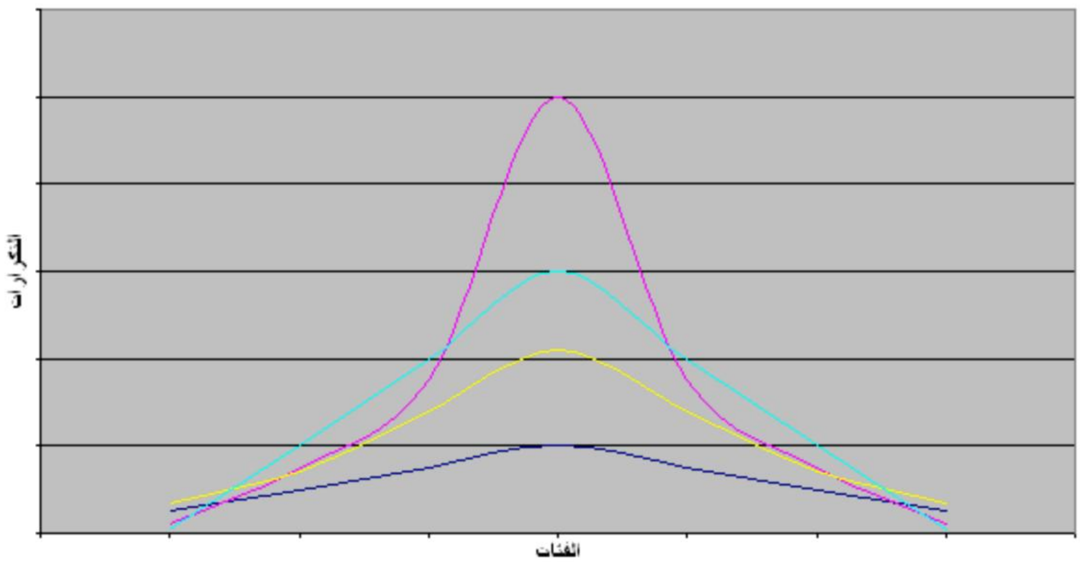
ثالثا : مقاييس الإلتواء Skewdness Measures

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

المنحنى المتماثل Symmetrical Curve

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

شكل يوضح منحنيات التوزيع المتماثل



مقاييس التشتت النسبى Coefficient of Variation

ونلاحظ ان المنحنيات المتماثلة فى الشكل السابق تختلف قممها ارتفاعا أو تفلطحا و تدببا حسب حجم التكرارات على جانبى القمة.

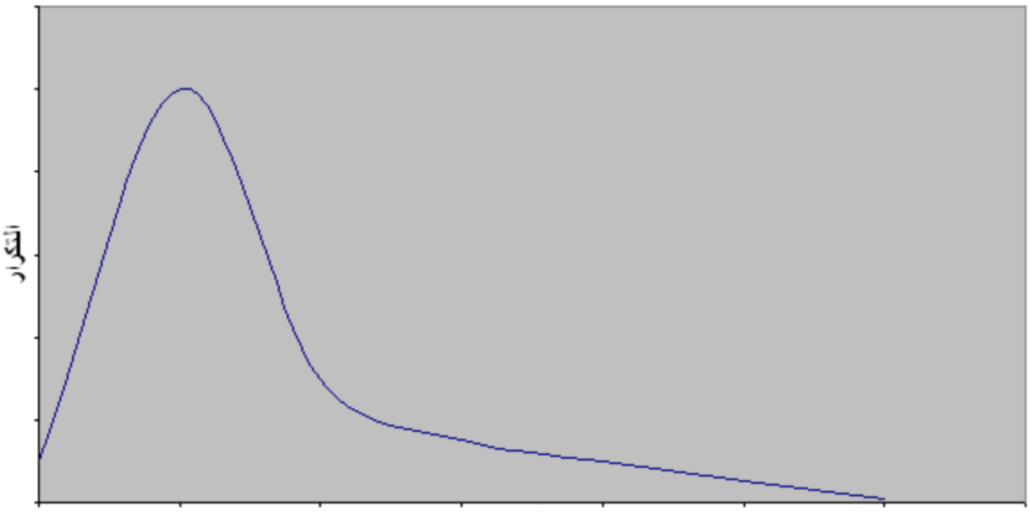
ويتميز التوزيع المتماثل بأن:

الوسط الحسابى = الوسيط = المنوال

المنحنيات الملتوية Skewed

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

شكل يوضح منحنى ملتوى جهة اليمين



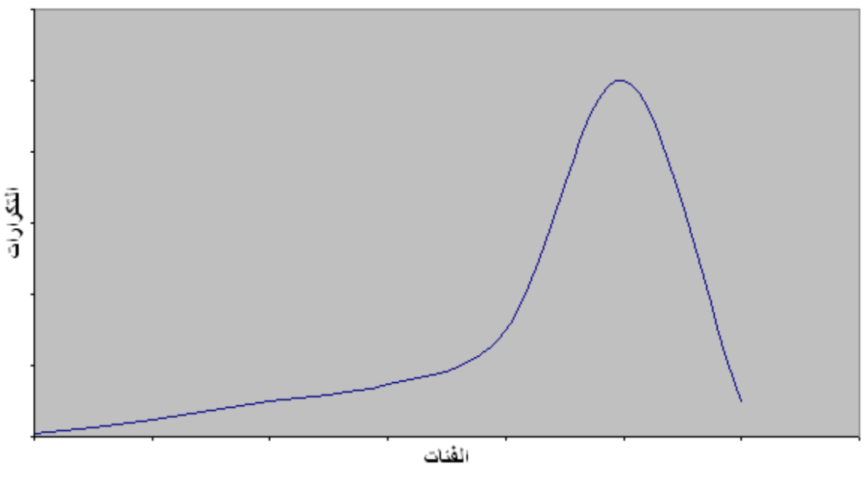
والتوزيع الملتوى جهة اليمين (الإلتواء الموجب) يكون فيه :

الوسط الحسابى > الوسيط > المنوال

أى أن الوسط الحسابى أكبر من الوسيط والوسيط أكبر من المنوال

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

شكل يوضح منحنى ملتوى جهة اليسار



والتوزيع الملتوى جهةاليسار (الإلتواء السالب) يكون فيه:

المنوال > الوسيط > الوسط الحسابى

أى أن المنوال أكبر من الوسيط والوسيط أكبر من الوسط الحسابى

ويمكن قياس الإلتواء من خلال معامل الإلتواء SK والذى يفيدنا فى الحكم على مدى تماثل أو إلتواء التوزيع حيث يكون التوزيع متاثل اذا كان معامل الإلتواء يساوى صفر او يكون إلتواء موجب إذا كانت قيمة معامل الإلتواء موجبة و يكون سالب اذا كانت قيمة معامل الإلتواء سالبة

إلا أنه فى بعض الاحيان يكون التوزيع قريب من التماثل ففى حالة ما تقترب قيمة معامل الإلتواء من الصفر، وتتعدد مقاييس الإلتواء إلا أن من أهمها:

معامل الإلتواء لبيرسون والذى يكون فى أحد الصورتين التاليتين:

الصورة الأولى: معامل الإلتواء = (الوسط الحسابي − المنوال) / الانحراف المعيارى

أي :

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

إلا أنه فى بعض الاحيان يكون التوزيع قريب من التماثل ففى حالة ما تقترب قيمة معامل الإلتواء من الصفر، وتتعدد مقاييس الإلتواء إلا أن من أهمها:

معامل الإلتواء لبيرسون والذى يكون فى أحد الصورتين التاليتين:

الصورة الثانية: معامل الإلتواء = 3 (الوسط الحسابي − الوسيط) / الانحراف المعيارى

أي

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

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

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

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

أو يمكن أعادة صياغة معامل الإلتواء لباولى SkB على الصورة التالية:

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

<u>معادلات حساب كل من الوسيط Med و الربيع الأول Q1 و الربيع الثالث Q3</u>

الوسيط *Med*

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

Med	قيمة الوسيط
L_{Med}	الحد الادنى لبداية الفئة الوسيطية
k_{Med}	ترتيب الوسيط
F_a	التكرار المتجمع السابق للفئة الوسيطية
F_b	التكرار المتجمع اللاحق للفئة الوسيطية
I_{Med_1}	طول الفئة الوسيطية

معادلات حساب كل من الوسيط Med و الربيع الأول Q1 و الربيع الثالث Q3

الربيع الأول Q1:

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

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

معادلات حساب كل من الوسيط Med و الربيع الأول Q1 و الربيع الثالث Q3

الربيع الثالث Q3:

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

Q_3	**قيمة الربيع الأدنى أو الثالث**
L_{Q3}	**الحد الادنى لبداية الفئة الربيعية الثالثة**
k_{Q_3}	**ترتيب الربيع الثالث**
F_a	**التكرار المتجمع السابق للفئة الربيعية الثالثة**
F_b	**التكرار المتجمع اللاحق للفئة الربيعية الثالثة**
I_{Q_3}	

مثال:

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

الايجار بالألف ريال	6-	10 -	12-	14- 18
عدد الوحدات السكنية	15	20	12	13

المطلوب:

حساب معامل الإلتواء لتوزيع الإيجار السنوى للوحدات السكنية.

الحل :

تم من قبل حساب المقاييس التالية:

المقياس	الوسط الحسابي	الانحراف المعيارى	الربيع الأول	الوسيط	الربيع الثالث
قيمته	11.733	2.8158	10	11.5	13.667

ولكن يبقى لنا حساب المنوال حتى تكون جميع المقاييس الإحصائية التى نحتاج إليها موجودة لذا يمكن الحصول على المنوال كما يلى:

نلاحظ أن أطوال الفئات للإيجار السنوى غير متساوية لذا لحساب المنوال يلزم إيجاد التكرار المعدل ومن ثم يتم إعداد الجدول التالي:

فئات الإيجار	التكرار f	طول الفئة	التكرار المعدل
6 -	15	4	3.75
10 -	20	2	10
12 -	12	2	6
14 - 18	13	4	3.25
المجموع			

ويمكن حساب المنوال بتطبيق المعادلة التالية كما سبق أن بينا ذلك فى الفصل السابق:

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

حيث أن:

$$D1 = 10 - 3.75 = 6.25$$

$$D2 = 10 - 6 = 4$$

$$L_{Mod} = 10$$

$$I_{Mod} = 2$$

وعلى ذلك يمكن حساب المنوال كما يلى:

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

- وعلى ذلك يمكن حساب معامل الإلتواء لبيرسون بإستخدام المعادلة $SK = \frac{\bar{x} - Mod}{S}$ كما يلي:

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

تفسير النتيجة:

ويعبر ذلك على وجود التواء موجب جهة اليمين الا أن قيمة معامل الإلتواء صغيرة تقترب من الصفر مما يدل ايضا على أن التوزيع قريب من التماثل.

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

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

تفسير النتيجة:

ويعبر ذلك ايضا على وجود التواء موجب جهة اليمين كما حددته النتيجة في المعادلة السابقة.

- كما يمكن تطبيق المعادلة $SK_B = \frac{Q_3 - 2Med + Q1}{Q_3 - Q1}$ لحساب معامل الإلتواء كما يلي:

$$SK_B = \frac{Q_3 - 2Med + Q1}{Q_3 - Q1} = \frac{13.6667 - 2(11.5) + 10}{13.6667 - 10} = .1816$$

تفسير النتيجة:

ويشير معامل الإلتواء لباولى بوجود التواء موجب.

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

رابعا: التفلطح Kurtosis

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

وتكون قيمة معامل التفلطح صفر فى حالة التوزيع الطبيعى المعيارى.

- لذا تقوم الكثير من البرامج الإحصائية بحساب معامل التفلطح للقيم المعيارية للبيانات فإذا كان الناتج:

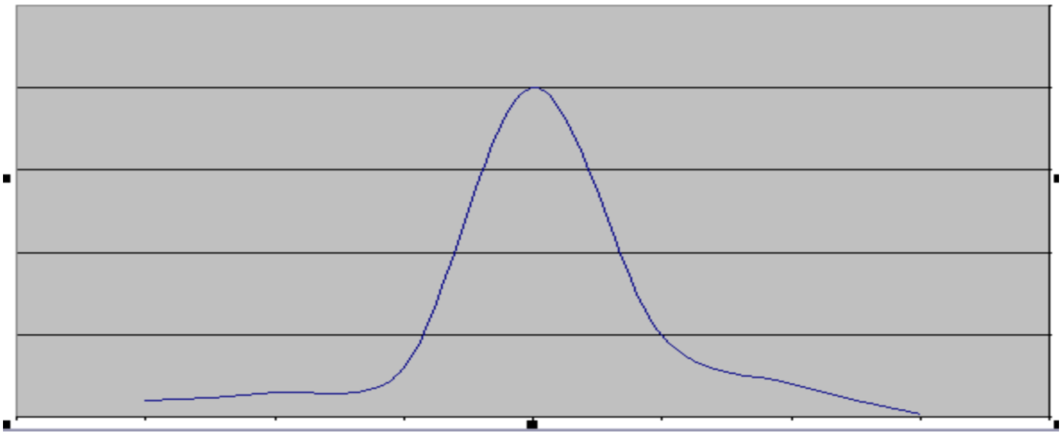
موجب أى قيمة معامل التفلطح للبيانات الأصلية أكثر من 3 فيكون بالتالي المنحنى مدبب إلي أعلى.

سالب أى قيمة معامل التفلطح للبيانات الأصلية أقل من 3 فيكون بالتالي المنحنى مفلطح أو أكثر إنبطاحا من قمة منحنى التوزيع الطبيعي.

صفر أى قيمة معامل التفلطح للبيانات الأصلية تساوى 3 فيكون بالتالي المنحنى متوسط التفلطح.

- ففى حالة ما يكون معامل التفلطح للبيانات الاصلية أكبر من 3 يكون المنحنى مدبب لأعلى كما بالشكل التالي:

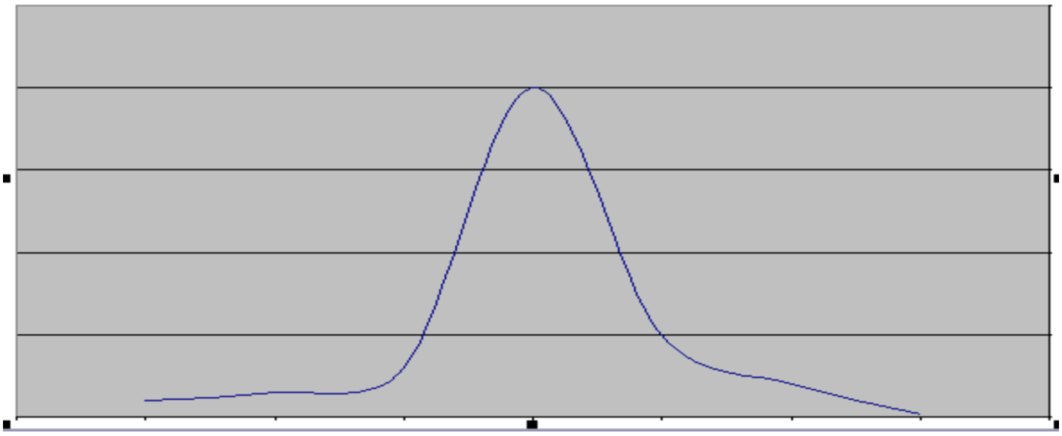
شكل يوضح المنحنى المدبب



وكما يتضح من الشكل السابق أن هناك فئة معينة من البيانات تتركز بها التكرارات والتى تجعل المنحنى مدبب إلى أعلى.

أما فى حالة ما يكون معامل التفلطح للبيانات الأصلية أقل من 3 يعنى ذلك أن المنحنى مفلطح كما يتضح من الشكل التالي:

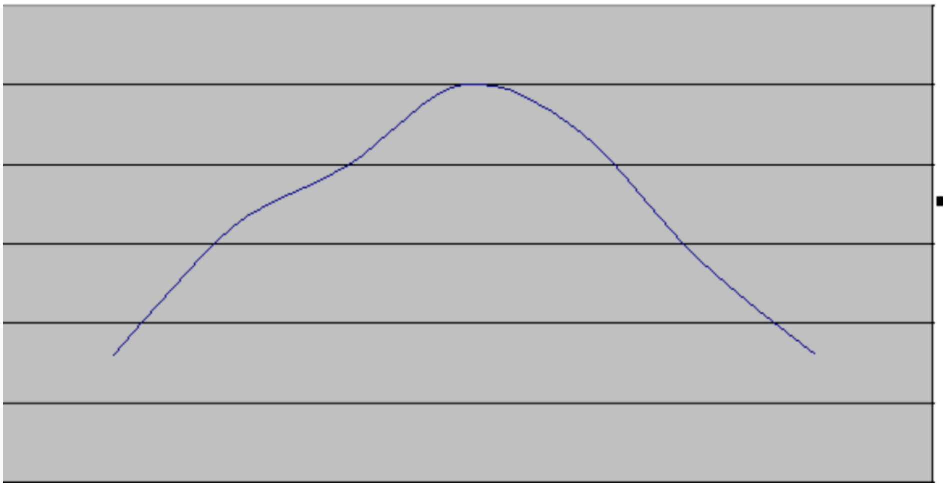
شكل يوضح المنحنى المدبب



وكما يتضح من الشكل السابق أن هناك فئة معينة من البيانات تتركز بها التكرارات والتى تجعل المنحنى مدبب إلى أعلى.

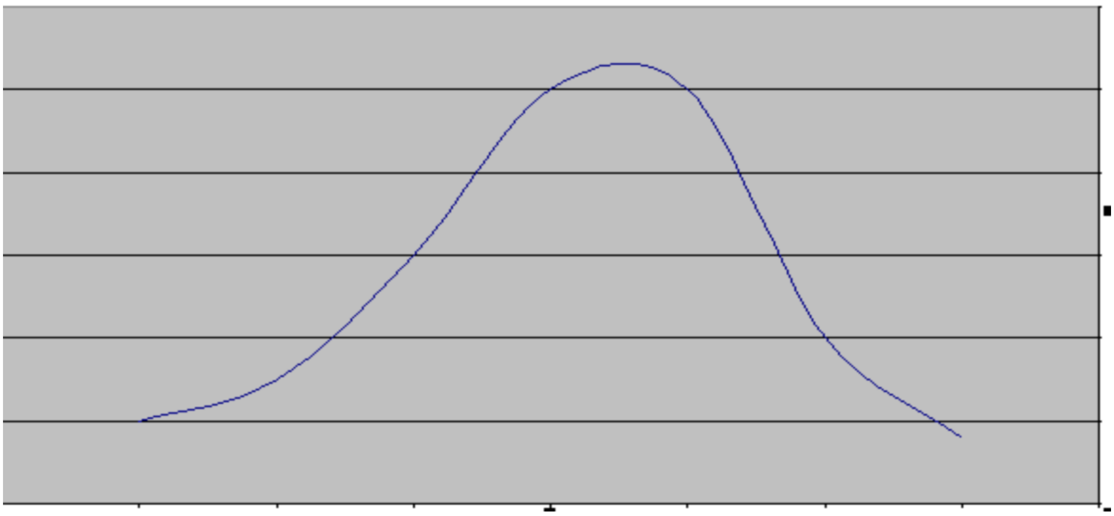
- أما فى حالة ما يكون معامل التفلطح للبيانات الأصلية أقل من 3 يعنى ذلك أن المنحنى مفلطح كما يتضح من الشكل التالي:

شكل يوضح المنحنى المفلطح



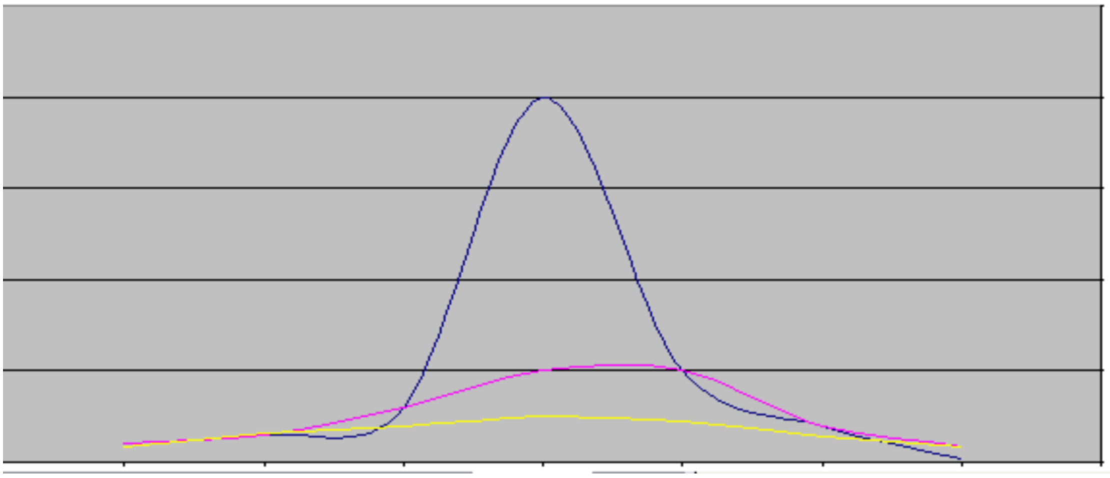
أما فى حالة ما يكون معامل التفلطح يساوى ثلاثة يكون المنحنى متوسط التفلطح و يكون بالشكل التالي:

شكل يوضح المنحنى متوسط التفلطح



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

شكل يوضح المنحنيات الثلاث معآ 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 المئين التسعين والذى يعبر عن 90 % من المفردات تكون أقل منه و 10% منها أكبر منه
$P_{0.10}$	إلى المئين العاشر (العشير) والذى يعبر عن 10 % من المفردات تكون أقل منه و 90% منها أكبر منه

مثال :

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

الايجار بالألف ريال	6-	10 -	12-	14- 18
عدد الوحدات السكنية	15	20	12	13

المطلوب:

حساب معامل التفلطح لتوزيع الإيجار السنوى للوحدات السكنية.

الحل :

تم سابقا حساب Q1 و Q3

ولكن يبقى علينا حساب كلا من $P_{0.10}$ و $P_{0.90}$ بنفس طريقة حساب الوسيط والربيع الأعلى والأدنى كما تم شرح ذلك من قبل كما يلي:

- إعداد الجدول التكراري المتجمع الصاعد كما يلي:

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

- إيجاد الرتبة كالتالي:

	$P_{0.10}$	$P_{0.90}$
الرتبة	$k_{P_{0.10}} = n/10$ $= 60/10$ $= 6$	$k_{P_{0.90}} = (n\times 9)/10$ $= (60\times 9)/10$ $= 54$

- إيجاد القيمة كالتالي:

المئين العاشر $P_{0.10}$

$$P_{0.10} = L_{P_{0.10}} + \frac{k_{P_{0.10}} - F_a}{F_b - F_a} \times I$$

$$P_{0.10} = 6 + \frac{6-0}{15-0} \times 4 = 7.6$$

إيجاد القيمة كالتالي:

المئين التسعين $P_{0.90}$

$$P_{0.90} = L_{P_{0.90}} + \frac{k_{P_{0.90}} - F_a}{F_b - F_a} \times I$$

$$P_{0.90} = 14 + \frac{54-47}{60-47} \times 4 = 16.153$$

- وقد تم حساب الربيعات Q1 و Q3 سابقا:

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

وعلى ذلك يمكن حساب معامل التفلطح كالتالي:

$$KU = \frac{Q_3 - Q_1}{2(P_{0.90} - P_{0.10})} = \frac{13.6667 - 10}{2(16.15385 - 7.6)} = 0.2143$$

ويتضح لنا أن معامل التفلطح أقل من 3 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.

المحاضره العاشـــــــــــــــــــره

تحليل الارتباط

يعتبر تحليل الارتباط Correlation Analysis من الاساليب الإحصائية المناسبة لتقييم العلاقات بين المتغيرات المختلفة.

ويتم استخدام معامل الارتباط فى الحكم على نوع العلاقة بين المتغيرين حيث تكون علاقة طرديه أو عكسية، وكذلك بالنسبة لقوه العلاقة فقد تكون علاقة قويه، أو متوسطه أو ضعيفة.

يستخدم معامل الارتباط البسيط Correlation coefficient فى تحديد ما إذا كان هناك علاقة بين المتغيرين، وكذلك تحديد نوع وقوة العلاقة إن وجدت.

أما فى حالة دراسة مدى وجود علاقة ارتباطية بين أكثر من متغيرين فإنة يتم الاعتماد على معامل الارتباط المتعدد Multiple Correlation coefficient.

أما فى حالة وجود أكثر من متغير ويرغب الباحث في تثبيت تأثير أحد المتغيرات كما فى الدراسات الاقتصادية يتم دراسة تأثير السعر على الكمية المطلوبة بفرض ثبات الجودة ومستوى الذوق كما هو يتم الاعتماد على معامل الارتباط الجزئى

Partial Correlation coefficient

تنقسم المتغيرات محل الدراسة كما أوضحنا ذلك سابقا إلى:

- متغيرات مستقلة Independent Variables

وهى المتغيرات التى بتغير قيمتها تؤثر فى تغيير قيمة متغير أو متغيرات أخرى، أى هى المتغيرات التى تتغير أولا. وسنرمز للمتغير المستقل بالرمز x

- المتغيرات التابعة Dependent Variables

وهى تلك المتغيرات التى تتغير قيمتها بتغير المتغيرات المستقلة أو إحداها، أى هى المتغيرات التى تتغير تالية للمتغيرات المستقلة. وسنرمز للمتغير التابع بالرمز y

وسيتم قياس الارتباط البسيط من خلال كلا من:

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

معامل الارتباط الخطى البسيط لبيرسون Person's Correlation Coefficient

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

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

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

وكذلك المعادلة الرياضية التالية والتي تعتبر اسهل وابسط:

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

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

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

والاتباط غالبا قيمته كسر أي اقل من الواحد الصحيح

- ولتحديد نوع العلاقة نعتمد على اشارة معامل الارتباط فإذا كانت الإشارة:

- موجبة فإن العلاقة تكون طرديه
- سالبة فإن العلاقة تكون عكسية

- ولتحديد قوة العلاقة نعتمد على قيمة معامل الارتباط فإذا كانت القيمة:

- أكبر من صفر إلى أقل من 0.3 فتكون علاقة ضعيفة
- من 0.3 إلى أقل من 0.7 تكون علاقة متوسطه
- من 0.7 إلى الواحد الصحيح تكون علاقة قويه
- إما إذا كانت قيمة معامل الارتباط تساوى صفر فلا توجد علاقة خطيه او ارتباط بينهما أى يكون المتغيرين مستقلين عن بعضهما البعض .

- فمثلا إذا كانت قيمة معامل الارتباط r p كالتالي فإن تفسيرة يكون:

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

مثال :

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

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

المطلوب:

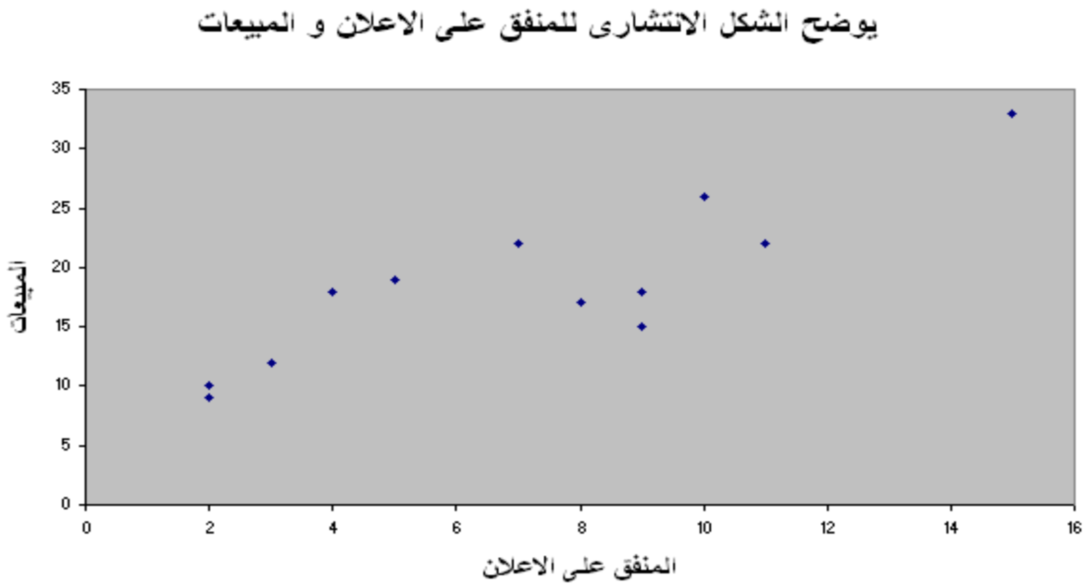
ارسم شكل الانتشار يوضح العلاقة بين المنفق على الاعلان و المبيعات ؟

احسب معامل الارتباط الخطى البسيط (بيرسون)، مع التعليق

الحل :

ارسم شكل الانتشار والذي يوضح العلاقة بين المنفق على الاعلان والمبيعات

حيث يتم الرسم من خلال رسم محورين سيني ويوضح المبالغ المنفقة على الإعلان وصادي يوضح المبيعاتومن ثم تحديد إحداثيات النقاط فيظهر لنا الشكل التالي:



نستنتج من شكل الانتشار أن قيم كلا من المنفق على الاعلان والمبيعات يأخذ اتجاه تصاعدى جهة اليمين مما يدل على وجود علاقة طرديه بينهما .

- و اذا اردنا أستخدام المعادلة الرياضية فى حساب معامل الارتباط بين المنفق على الاعلان والمبيعات لابد أولا من حساب الوسط الحسابى $\bar{x}$ للمنفق على الاعلان والوسط الحسابى للمبيعات $\bar{y}$ حيث أن:

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

واذا اردنا أستخدام المعادلة السابقة فى حساب معامل الارتباط بين المنفق على الاعلان والمبيعات لابد من حساب الوسط الحسابى للمنفق على الاعلان $\bar{x}$ والوسط الحسابى للمبيعات $\bar{y}$ كما يلي:

$$\bar{x} = \frac{\sum x}{n} = \frac{82}{12} = 7.08333$$

$$\bar{y} = \frac{\sum y}{n} = \frac{221}{12} = 18.41667$$

- و على ذلك يمكن لنا أعداد الجدول التالي :

$(y-\bar{y})^2$	$(x-\bar{x})^2$	$(x-\bar{x})(y-\bar{y})$	$(y-\bar{y})$	$(x-\bar{x})$	y	x
70.84028	23.36111	40.68056	-8.41667	-4.83333	10	2
41.17361	14.69444	24.59722	-6.41667	-3.83333	12	3
88.67361	23.36111	45.51389	-9.41667	-4.83333	9	2
12.84028	0.027778	0.597222	3.58333	0.16666	22	7
0.173611	0.694444	0.347222	-0.41667	-0.83333	18	6
0.340278	3.361111	-1.06944	0.58333	-1.83333	19	5
57.50694	10.02778	24.01389	7.58333	3.16666	26	10
212.6736	66.69444	119.0972	14.5833	8.16666	33	15
0.173611	8.027778	1.180556	-0.41667	-2.83333	18	4
12.84028	17.36111	14.93056	3.58333	4.16666	22	11
11.67361	4.694444	-7.40278	-3.41667	2.16666	15	9
2.006944	1.361111	-1.65278	-1.41667	1.16666	17	8
510.9167	173.6667	260.8333	0	0	221	82

- ويمكن بالتالي تطبيق المعادلة الرياضية لحساب معامل الارتباط كما يلي:

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

$$= \frac{260.8333}{\sqrt{173.667}\sqrt{510.9167}} = 0.8756$$

وتدل قيمة معامل الارتباط على وجود علاقة قوية وطردية بين المنفق على الإعلان والمبيعات

- كما يمكن حساب معامل الارتباط من خلال المعادلة الثانية والتى تكون بالصورة التالية:

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

وحتى يمكن تطبيق هذه المعادلة على بيانات المثال السابق لحساب معامل الاتباط بين المنفق على الاعلان والمبيعات لابد من تكوين الجدول التالي:

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

- وبالتالى يمكن تطبيق المعادلة السابقة كما يلي:

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

$$= \frac{12(1771)\text{-}(82\times 221)}{\sqrt{12(734)\text{-}(82)^2}\sqrt{12(4581)-(221)^2}}$$

$$= \frac{3130}{\sqrt{2084}\sqrt{6131}} = 0.8756$$

وهى نفس النتيجة التى حصلنا عليها بتطبيق المعادلة السابقة مما يدل علي وجود علاقة طردية وقوية بين المنفق على الإعلان والمبيعات .

***ومن أهم خصائص معامل الارتباط الخطى البسيط لبيرسون أنه لا يعتمد على قيم المتغيران نفسها عند حساب قيمته وإنما يعتمد على مقدار التباعد بين هذه القيم 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 أكتشفت إدارة الشركة أن البيانات تم تجميعها وحسابها بطريقة خاطئة حيث يجب إضافة 5 مليون ريال إلى جميع قيم المنفق على الإعلان. كما أن المبيعات يجب مضاعفة قيمتها لجميع القيم.

المطلوب:

أحسب معامل الارتباط فى هذه الحالة بين المنفق على الإعلان والمبيعات.

الحل :

يتم اولاً تعديل البيانات لكل من المنفق على الأعلان والمبيعات لتكون النتائج كما يلي:

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

- وبالتالى يمكن تطبيق واحدة من المعادلات السابقة كما يلي:

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

$$= \frac{12(5752)-(142\times 442)}{\sqrt{12(1854)-(142)^2}\sqrt{12(18324)-(442)^2}}$$

$$= \frac{6260}{\sqrt{2084}\sqrt{24524}} = 0.8756$$

وهى نفس النتيجة التى حصلنا عليها سابقاً مما يدل على أن معامل الأرتباط لم تتأثر قيمته بالعمليات الجبرية من جمع (5 مليون) أو الضرب (× 2). وبالمثل لايتأثر بالطرح أو القسمة.

معامل التحديد Determination Coefficient

وهو مربع معامل الارتباط لذلك يرمز له بالرمز R^2 أو R-Square و هو يشير إلى نسبة تفسير المتغير أو المتغيرات المستقلة للتغير فى المتغير التابع

فمثلا:

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

معامل أرتباط الرتب لسبيرمان r_s Spearman's Rank Correlation Coefficient

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

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

- ويتم حساب معامل الارتباط الرتب لسبيرمان r_s بأستخدام المعادلة التالية:

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

حيث ان :

d الفرق بين رتبة المتغيرين

n عدد المشاهدات

ملاحظات يجب مراعاتها عند ترتيب المتغيرا ت:

- 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 شيء 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 1 والقيمة التى هى أكبر منها الرتبة 2 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 ؟

الحل :

يتم اولا ترتيب قيم كلا من x و y كما يتضح من الجدول التالى:

المنفق على الإعلان X	المبيعات y	رتب x	رتب y	d	d2
2	10	1.5	2	-0.5	0.25
3	12	3	3	0	0
2	9	1.5	1	0.5	0.25
7	22	6	9.5	-3.5	12.25
9	18	8.5	6.5	2	4
5	19	5	8	-3	9
10	26	10	11	-1	1
15	33	12	12	0	0
4	18	4	6.5	-2.5	6.25
11	22	11	9.5	1.5	2.25
9	15	8.5	4	4.5	20.25
8	17	7	5	2	4
				0	59.5

نلاحظ من ذلك الجدول أن:

١. تم ترتيب المتغيران تصاعدياً

٢. عند ترتيب قيم المتغير المنفق على الاعلان x وجدنا ان القيمة 2 تكررت مرتان لتأخذ الرتب 1 و 2 لذلك نحسب المتوسط لهما وهو (2+1) ÷ 2 ليكون 1,5 لذلك وضعنا امام القيمة 2 الرتبة 1.5 . وكذلك الامر بالنسبة للقيمة 9 فإنها تأخذ الرتبة 8 و 9 لذلك وضعنا أمام القيمة 9 الرتبة 8,5

٣. عند ترتيب قيم المتغير " المبيعات " y وجدنا أن القيمة 18 أخذت الرتبة 6 و 7 لذلك وضعنا أمام القيمة 18 الرتبة 6,5 وكذلك القيمة 22 أخذت الرتبة 9 و 10 لذلك و ضعنا أمامها الرتبة 9,5

ثم نحسب الفرق بين رتب المتغير x ورتب المتغير y والتى نعطى لها الرمز d ، ونلاحظ من الجدول السابق أن مجموع الفروق d لابد أن يكون صفر والا يكون هناك خطأ فى الترتيب لأحد المتغيرين أو كلاهما، ولابد من مراجعة الترتيب مرة أخرى والتأكد من ذلك.

كما بلغ $\sum d^2 = 59.5$ وحيث أن عدد المشاهدات $n = 12$ فإنة يمكن حساب معامل ارتباط الرتب لسبيرمان كما يلي:

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

وقد بلغ معامل ارتباط الرتب لسبيرمان 0.7919 مما يدل على وجود ارتباط طردى قوى بين المنفق على الاعلان والمبيعات، وهى قيمة قريبة من التى تم حسابها بإسخدام معامل الارتباط لـ بيرسون حيث بلغ 0.8756

مثال :

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

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

المطلوب:

أحسب معامل الارتباط المناسب.

الحل :

يتم ترتيب المشاهدات وحساب الفروق بين الرتب ومربعاتها كما يتضح من الجدول التالي:

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

- نلاحظ عند ترتيب تقديرات مقرر المحاسبة أن التقدير " مقبول " اخذ الرتب 2 و 3 و 4 لذلك تم جمع (2+3+4) ÷ 3 = 9÷3 = 3 فوضع 3 أمام التقدير مقبول فى مقرر المحاسبة.
- كما أن تقدير جيد فى مقرر القانون أخذ الرتب 3 و 4 و 5 و 6 لذلك تم جمع (3+4+5+6) ÷ 4 = 18 ÷ 4 = 4,5 فوضع الرتبة 4.5 امام التقدير جيد فى مقرر القانون.

- من الجدول السابق يتضح لنا أن مجموع الفروق d لابد أن يكون صفر.

كما بلغ $\sum d^2 = 247$ وحيث أن عدد المشاهدات $n = 10$ فإنة يمكن حساب معامل ارتباط الرتب لسبيرمان كما يلي:

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

ونلاحظ أن معامل ارتباط الرتب لسبيرمان بلغ -0.4969 مما يدل على وجود ارتباط عكسى متوسط بين تقدير مقرر المحاسبة وتقدير مقرر القانون.

معامل الاقتران Conjunction Coefficient

ويستخدم فى حساب العلاقة الاتباطية 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:

	الصفة الأولى لـــ y	الصفة الثانية لـــ y
الصفة الأولى لـــ x	A	B
الصفة الثانية لـــ 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 200 شخص سؤالين هما:

هل انت متعلم ؟ نعم لا

هل انت ملتحق بأى عمل ؟ نعم لا

وبتجميع الاجابات تم عمل جدول الاقتران التالي:

العمل / التعليم	متعلم	أمى
يعمل	**113**	**23**
لايعمل	**49**	**15**

المطلوب :

أحسب معامل الاقتران ؟

الحل :

يمكن حساب معامل الاقتران فى هذه الحالة كما يلي:

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

	متعلم	أمى
يعمل	113 = A	23 = B
لايعمل	49 = C	15 = D

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

- أى يوجد ارتباط ضعيف بين العمل والتعليم

معامل التوافق Contingency C0efficient

ويستخدم لحساب الارتباط بين المتغيرات الوصفية الاسمية والتى يكون لصفاتها قيم أكثر من 2، مثل الحالة الاجتماعية (اعزب - متزوج – متزوج ويعول – أرمل – مطلق)

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

- ويتم حساب معامل التوافق من خلال المعادلة التالية:

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

حيث أن:

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

أى يتم إيجاد:

مربع تكرار كل خلية مشتركة

مجموع الصف × مجموع العمود

ثم نجمعهم كلهم

- وعلى ذلك يتم حساب معامل التوافق كما يلى:

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

مثال :

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

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

الحل :

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

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

- وعلى ذلك يتم حساب معامل التوافق كما يلى:

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

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

المحاضره الحادية عشر

تحليل الإنحدار

يعتبر تحليل للانحدار أكثر طرق التحليل الإحصائى استخداما، حيث يتم من خلاله التنبؤ بقيمة احد المتغيرات (المتغير التابع) عند قيمة محددة لمتغير أو متغيرات أخرى (المتغيرات المستقلة).

وتسمى العلاقة الرياضية التى تصف سلوك المتغيرات محل الدراسة والتى من خلالها يتم التنبؤ بسلوك احد المتغيرين عند معرفة الاخر بمعادلة خط الانحدار.

- وهناك صورتان أساسيتان لمعادلة الانحدار وهما:

الصورة الأولي: معادلة انحدار x|y (التى يطلق عليها معادلة انحدار y على x)

الصورة الثانية: معادلة انحدار y|x (التى يطلق عليها معادلة انحدار x على y)

معادلة انحدار y على x

وهى تلك المعادلة التي يطلق عليها معادلة انحدار y | x. أي تتحدد قيمة المتغير y تبعا لقيمة المتغير x لذلك يمكن التعبير عن تلك العلاقة الخطية بالمعادلة التالية:

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

حيث يسمى b_0 ثابت الانحدار او الجزء الثابت او الجزء المقطوع من محور الصادات بينما يطلق عليها معامل الانحدار أو معدل التغير فى الدالة .

ويمكن تقدير قيمة للثابتين b_0 و b_1 كما يلي، وهما المعادلات التي نستخدمها لحساب معامل الانحدار باستخدام طريقة المربعات الصغرى:

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

$$b_0 = \frac{\sum y}{n} - b_1 \frac{\sum x}{n}$$

$$= \bar{y} - b_1 \bar{x}$$

مثال :

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

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

المطلوب أوجد:

١. معادلة انحدار y على x ؟

٢. تحديد معدل التزايد أو التناقص فى استهلاك الكهرباء؟

٣. ماهو الاستهلاك المتوقع لمسكن مكون من 8 غرف؟

٤. **الحل :**

نقوم بعمل الجدول التالى:

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

وبالتالى يمكن تقدير b_1 من خلال تطبيق المعادلة التالية:

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

$$b_1 = \frac{10(650) - (69)(85)}{10(529) - (69)^2} = \frac{635}{529} = 1.2003$$

وكذلك يمكننا تقدير b_0 من خلال تطبيق المعادلة التالية:

$$b_0 = \frac{\sum y}{n} - b_1 \frac{\sum x}{n}$$
$$= \frac{85}{10} - (1.2003)\frac{69}{10}$$
$$= 8.5 - 8.28207$$
$$= 0.21793$$

وعلى ذلك يمكن كتابة معادلة الانحدار yعلى x على الشكل التالي:

$$\hat{y} = 0.21793 + 1.2003x$$

وبالتالى يكون معدل التزايد فى استهلاك الكهرباء هو لأنها موجبة ويساوى 1.2003 أى أن كل غرفة بالمسكن تعمل على زيادة استهلاك الكهرباء بمقدار 1200.3 كيلو وات.

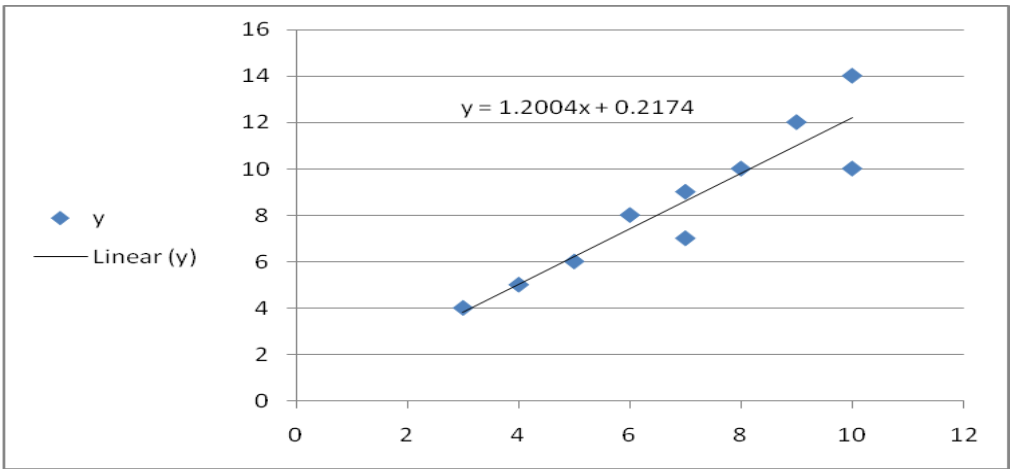
- الاستهلاك المتوقع لمسكن مكون من 8 غرف:

يتم التعويض فى معادلة الانحدار التى سبق إيجادها عندما تكون x = 8 كما يلي:

$$y = 0.21793 + 1.2003(8) = 9.8203$$

أى أن الاستهلاك المتوقع لمسكن مكون من 8 غرف هو 9820.3 كيلو وات

- ويمكن لنا رسم بيانات المثال السابق وخط معادلة الانحدار y على x كما يلي:



ويتضح لنا من الشكل السابق ان خط الانحدار لا يمر بجميع النقاط حيث تكون هناك نقاط مشتته حول الخط، وبالرغم من ذلك يعد هذا الخط من أفضل الخطوط التى حصلنا عليها للتعبير عن العلاقة بين المتغيرين محل الدراسة ولكن بخطـــأ

معين يسمى خطأ التقدير Standard Error

معادلة انحدار x على y

وهى التي يطلق عليها معادلة انحدار x | y. أي تتحدد قيمة المتغير x تبعا لقيمة المتغير y لذلك يمكن التعبير عن تلك العلاقة الخطية بالمعادلة التالية:

$$\hat{x} = c_0 + c_1 y$$

حيث يسمى c_0 ثابت الانحدار او الجزء الثابت بينما c_1 يطلق عليها معامل الانحدار أو معدل التغير فى الدالة

ولتحديد المعادلة الدالة على العلاقة بين المتغيرين x و y لابد من تقدير قيمة للثابتين c_0 و c_1 الذين يمكن تقديرهما من خلال تطبيق طريقة المربعات الصغرى فتكون النتيجة كما يلي:

$$c_1 = \frac{n\sum xy - \sum x \sum y}{n\sum y^2 - (\sum y)^2}$$

$$c_0 = \frac{\sum x}{n} - c_1 \frac{\sum y}{n}$$
$$= \bar{x} - c_1 \bar{y}$$

مثال :

باستخدام بيانات المثال السابق لعدد الغرف واستهلاك الكهرباء أوجد التالي:

١. معادلة انحدار x على y ؟

٢. ماهو عدد الغرف المتوقع لأستهلاك 25000 كيلو وات ؟

الحل :

نقوم بعمل الجدول التالي:

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

من خلال الجدول السابق يمكن تقدير معادلة انحدار x على y كما يلي:

اولاً- يتم تقدير قيمة معامل الانحدار c_1

$$c_1 = \frac{n\sum xy - \sum x \sum y}{n\sum y^2 - (\sum y)^2}$$

$$c_1 = \frac{10(650) - (69)(85)}{10(811) - (85)^2}$$

$$= \frac{635}{885} = 0.717$$

ثانياً – تقدير قيمة c_0

$$c_0 = \frac{\sum x}{n} - c_1 \frac{\sum y}{n}$$

$$c_0 = \frac{69}{10} - 0.717\frac{85}{10}$$
$$= 6.9 - 6.0945 = 0.8055$$

معادلة انحدار x على y هي:

$$\hat{x} = c_0 + c_1 y$$

$$\hat{x} = 0.8055 + 0.717y$$

ماهو عدد الغرف المتوقع لأستهلاك 25000 كيلو وات

يتم التعويض فى المعادلة السابقة عن قيمة y تساوى 25 كما يلي:

$$\hat{x} = 0.8055 + 0.717(25) = 18.7305$$

- العلاقة بين معاملي معادلتي الانحدار y على x و معادلة انحدار x على y

اذا علم معامل معادلة انحدار y على x b_1 ومعامل معادلة انحدار x على y c_1 فإنة يمكن تقدير كلاً من معامل التحديد ومعامل الارتباط كما يلي:

$$r^2 = b_1 \times c_1$$

فكما يبدوا معامل التحديد هو عبارة عن حاصل ضرب معاملى الانحدار b_1 و c_1

وبالتالي يمكن الحصول على معامل الارتباط بأخذ الجذر التربيعى لمعامل التحديد كما يلي:

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

مع ملاحظة أن اشارة معامل الارتباط تكون موجبة أو سالبة بما يتفق وإشارة كلا من b_1 c_1 حيث أن إشارتهم جميعا واحدة، لأن الاشارة لأي منهم تتوقف على البسط نفسه وهو التغاير بين المتغيرين x و y .

كما يمكن معرفة قيمة أي معامل انحدار بمعلومية الأخر كما يلي:

$$b_1 = r \times \frac{\sigma_y}{\sigma_x} \qquad\qquad c_1 = r \times \frac{\sigma_x}{\sigma_y}$$

حيث ان :

σ_x	الانحراف المعيارى للمتغير x
σ_y	الانحراف المعيارى للمتغير y

مثال :

أحسب معامل الارتباط بين عدد الغرف والمستهلك من الكهرباء إذا علمت أن:

$$b_1 = 1.2003 \qquad\qquad c_1 = 0.717$$

الحل :

إيجاد معامل التحديد كالتالي:

$$\begin{aligned} r^2 &= b_1 \times c_1 \\ &= 1.2003 \times 0.717 \\ &= 0.8606 \end{aligned}$$

أى أن عدد الغرف يفسر 86.06 % من التغير فى استهلاك الكهرباء

إيجاد معامل الارتباط كالتالي:

$$r = \sqrt{r^2} = \sqrt{0.8606} = 0.9276$$

مما يدل على وجود ارتباط طردى قوى بين عدد الغرف و استهلاك الكهرباء

المحاضره الثانيت عشر (الجزء الأول)

السلاسل الزمنية

يحتاج التخطيط الفعال إلى أدوات تنبؤ متقدمة نظريا وتطبيقيا في مجالات إحصائية عديدة ومنها تحليل السلاسل الزمنية والتي تقوم على دراسة التطور التاريخي لقيم الظواهر المختلفة لمعرفة خصائصها واستخدامها في استخلاص النتائج النهائية.

وتبرز أهمية تحليل السلاسل الزمنية في حالات كثيرة مرتبطة بالجوانب الاقتصادية والإدارية والاجتماعية والبيئية، ومن ضمن الحالات المتعلقة بالجوانب الاقتصادية والإدارية مايلي:

- الناتج المحلي الإجمالي
- معدلات التضخم
- إجمالي الودائع
- أسعار النفط الخام والمنتجات النفطية
- كمية وقيمة المبيعات
- ميزانية الإعلان
- مستوى المخزون
- إجمالي التكاليف
- مستوى الدائنون والمدينون

- وفي جميع هذه الحالات يحتاج متخذ القرار إلى دراسة البيانات التاريخية كما وكيفا، ومن ثم تحديد الفروق الجوهرية بين الظروف التي أحاطت هذه البيانات التاريخية والظروف الحالية من أجل دمجها في مراحل عملية التحليل النهائي المساعدة في اتخاذ القرار.

- تعريف السلسلة الزمنية :

السلسة الزمنية عبارة عن مجموعة من المشاهدات الاحصائية تصف الظاهرة مع مرور الزمن، أو هي البيانات الاحصائية التي تجمع أو تشاهد أو تسجل لفترات متتالية من الزمن .

وقد تكون السلسلة الزمنية بالارقام المطلقة (وتسمى بالتالي سلسلة قيم مطلقة).

أو قد تكون السلسلة الزمنية بالقيم النسبية مثل تلك الجداول التي تبين معدلات الزيادة الطبيعية للسكان في الألف ونحوها.

أو قد تكون السلسلة الزمنية بالمتوسطات مثل السلسلة الزمنية التي تبين متوسط إنتاج الكيلومتر مربع من القمح.

- أمثلة متنوعة على السلاسل الزمنية :

- مرضى العيادات النفسية المترددين شهريا
- عدد الأطفال المرضى الجدد المصابين بالتوحد شهرياً
- عدد المتعطلين سنوياً عن العمل
- معدلات الإنجاب السنوية
- معدلات الطلاق السنوية
- المبيعات اليومية في مركز لبيع الكتب لمدة شهر
- قراءة درجات حرارة المريض في ساعة لمدة يوم واحد
- قراءة الإنتاج الشهري لمدة سنة في شركة للأدوية
- الإنتاج الشهري من البترول للسعودية ولعدة سنوات

كل هذه القراءات وتتابعها الزمني جميعها تمثل سلسلة زمنية

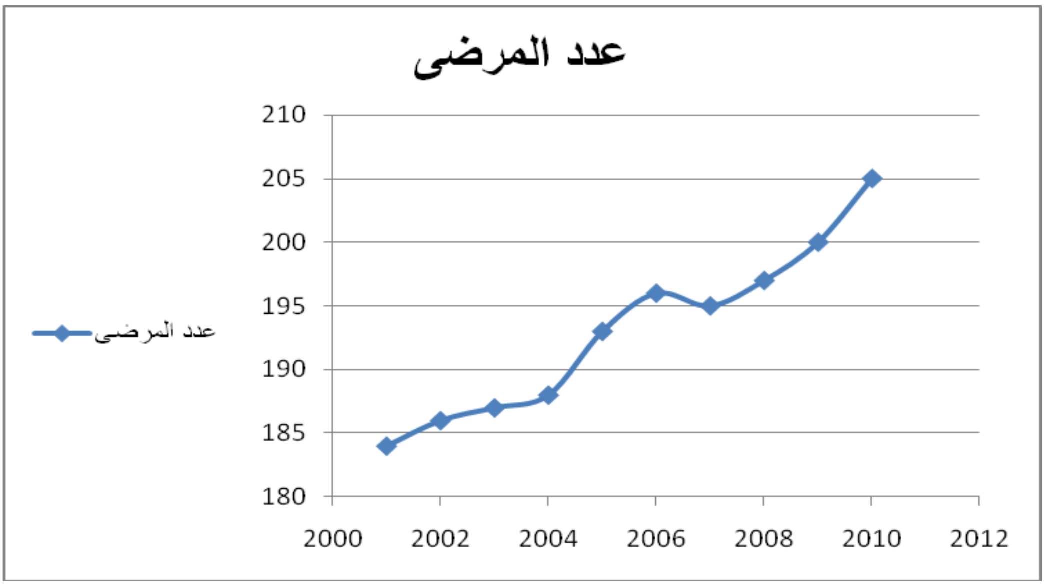
مثال :

الجدول التالي يوضح عدد مرض الفصام المترددين على احد العيادات خلال العشر سنوات الماضية:

عدد المرضى	السنة
184	2001
186	2002
187	2003
188	2004
193	2005
196	2006
195	2007
197	2008
200	2009
205	2010

- ويمكن رسم الشكل البيانى للسلسلة الزمنية على الشكل التالي:

حيث يتم الرسم من خلال رسم محورين سيني ويوضح السنوات وصادي يوضح عدد مرضى الفصام ومن ثم تحديد إحداثيات النقاط فيظهر لنا الشكل التالي:



- أنواع السلسلة الزمنية :

السلسلة الزمنية نوعان هما:

١. سلسلة زمنية فترية وهي السلسلة التي تتكون من بيانات كمية لمستوى الظاهرة عن فترات محددة من الزمن (شهر، ربع سنة، أو ما شابه ذلك)

٢. السلسة الزمنية اللحظية وهي السلسلة التي تتكون من مستويات للظاهرة مقاسة في لحظات (تواريخ معينة ومحددة)

- تحليل السلسلة الزمنية :

لغرض فهم السلسلة الزمنية لابد من تحليلها إلى عناصرها ومركباتها الأساسية مما يمكننا من معرفة تطور الظاهرة مع الزمن والتنبؤ بمعالمها خلال الفترات المقبلة لتتخذ اساسا للتخطيط الاقتصادي أو الإداري الطويل الأجل، وتتألف السلسة الزمنية من أربعة عناصر اساسية هي:

١. الاتجاه العام ويرمز لقيمه بالرمز (T) وتسمى "القيم الإتجاهية"

٢. التغيرات الموسمية ويرمز لقيمها بالرمز (S) وتسمى "القيم الموسمية"

٣. التغيرات الدورية ويرمز لقيمها بالرمز (C) وتسمى "القيم الدورية"

٤. التغيرات العشوائية أو الفجائية ويرمز لقيمها بالرمز (R) وتسمى "القيم العشوائية"

أي أن القيمة الأصلية للظاهرة (Yt) في كل سنة من السنوات يمكن وصفها بالعلاقة التالية :

١. نموذج الجمع :

ويستخدم عندما يكون مدى التغيرات الموسمية ثابت من سنة إلى أخرى ومستقل عن الاتجاه العام , و يتم فرض أن السلسلة الزمنية مكونة من مجموع مكوناتها من الاربع عناصر السابق ذكرها، أى يكون النموذج بالصورة التالية:

$$y_t = T_t + C_t + S_t + R_t$$

2. نموذج الضرب:

ويستخدم هذا النموذج في الحالات المعاكسة لحالات استخدام نموذج الجمع. ويتم فرض أن السلسلة الزمنية مكونة من حاصل ضرب مكوناتها من الاربع عناصر، أى يكون النموذج على الصورة التالية :

$$y_t = T_t \times C_t \times S_t \times R_t$$

ويجب ملاحظة أن قيم المتغيرات الموسمية وكذا المتغيرات الدورية عبارة عن نسب مئوية في نموذج الضرب.

حيث أن :

Yt = قيمة الظاهرة المدروسة في الفترة t (القيمة الحقيقية)

Tt = قيمة الاتجاه العام في الفترة t

Ct = قيمة التغيرات الموسمية (القيم الموسمية) في الفترة t

St = قيمة التغيرات الدورية (القيم الدورية) في الفترة t

Rt = قيمة التغيرات العشوائية (القيم العشوائية) في الفترة t

- عناصر السلسلة الزمنية:

إن دراسة أي سلسلة زمنية وتحليلها يستدعي دراسة كل عنصر من هذه العناصر على حدة، وهذه العناصر هي:

1- الاتجاه العام The Secular Trend :

تغيرات الاتجاه العام تعني الزيادة أو الإنخفاض طويل الأجل في البيانات عبر الزمن، ويتم التعرف على ذلك من خلال تمثيل السلسلة الزمنية بيانيا فنحصل بالتالي على خط بياني، واتجاه خط السلسلة الزمنية صعودا أو هبوطا يسمى الاتجاه العام للسلسلة، فإذا نظرنا للخط ووجدناه يتجه من الأسفل إلى الأعلى دل ذلك على نمو الظاهرة مع مرور الزمن، أما إذا كان الخط يهبط من الأعلى إلى الأسفل دل على ان الظاهرة تنقص مع مرور الزمن، أما إذا كان الخط أفقيا دل ذلك على ثبات الظاهرة .

- طرق حساب الإتجاه العام:

أ- طريقة الانتشار (التمهيد باليد):

يتم بهذه الطريقة رسم شكل الانتشار للظاهرة موضع الدراسة، **وشكل الانتشار عبارة عن رسم بياني لمتغيرين بحيث يكون الزمن على المحور السيني، وقيم الظاهرة على المحور الصادي،** وعند توصيل نقط شكل الانتشار ببعضها البعض نحصل على الخط البياني للظاهرة عبر الزمن، ويعطي شكل الانتشار فكرة سريعة عن طبيعة الاتجاه العام للظاهرة ومدى ارتباطه بالزمن ومدى تأثير التقلبات الدورية أو الموسمية أو التغيرات العشوائية، وبالامكان ومن خلال شكل الانتشار القيام بعملية مقارنة بين سلسلتين أو أكثر عبر فترات مختلفة من الزمن.

- وعملية التمهيد باليد (شكل الانتشار) عادة لا تكون دقيقة مما يقلل الاعتماد عليها وذلك لأن التمهيد باليد يتم بطريقة تقديرية تختلف من شخص لآخر وتعتمد على مهارة الشخص في رسم خط يمر بأكبر عدد ممكن من النقاط ويمثل السلسلة أفضل تمثيل

مثال :

إذا كان لدينا بيانات ربع سنوية لإجمالي الودائع في المصارف السعودية (آلاف الملايين من الريالات) في الفترة النصف الأخير من عام 2005م إلى عام 2007م والموضحة في الجدول التالي:

السنة	2005	2005	2006	2006	2006	2006	2007	2007	2007	2007
الفصل	3	4	1	2	3	4	1	2	3	4
الودائع	195.6	196.9	200.1	205.3	207.4	215.4	223.3	222.3	222.07	226.1

المطلوب :

رسم شكل الانتشار لهذه البيانات ومن ثم تفسيره وإبراز معالم الاتجاه العام للظاهرة موضع الدراسة ؟

الحل :

يتم رسم شكل الانتشار من خلال رسم محورين سيني ويوضح الفترات الزمنية بربع السنة وصادي يوضح الودائع ومن ثم تحديد إحداثيات النقاط.

ويمكن كذلك رسم شكل الانتشار من خلال إدخال البيانات السابقة إلى برنامج الإكسل لتكون بالشكل التالى:

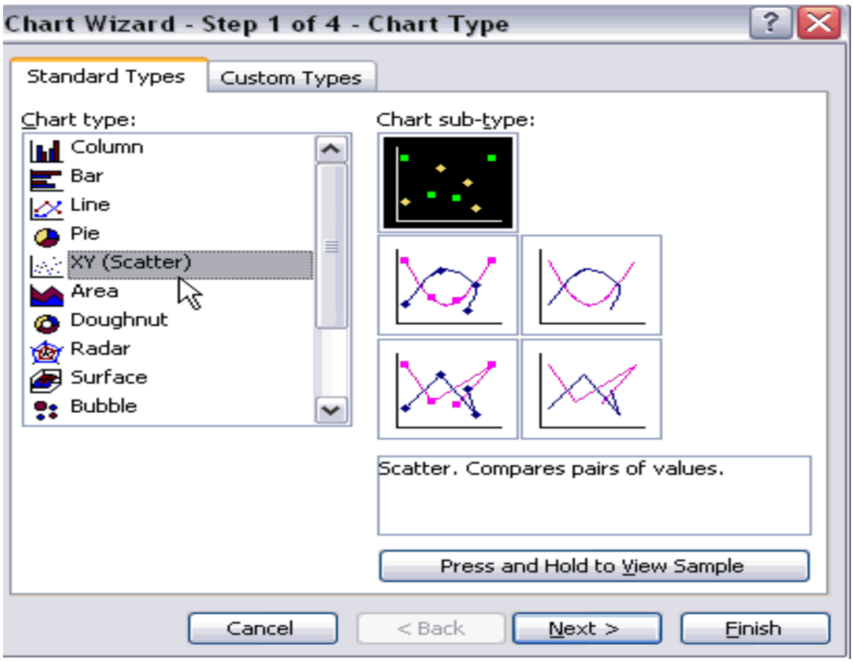
	A	B	C	D
1	السنة	الفصل	الفترة الزمنية	الودائع
2	2005	3	1	195.64
3		4	2	196.97
4	2006	1	3	200.11
5		2	4	205.33
6		3	5	207.49
7		4	6	215.46
8	2007	1	7	223.36
9		2	8	222.31
10		3	9	222.07
11		4	10	226.18

نلاحظ أننا بالإضافة إلى البيانات التى كانت موجودة بالتمرين وهى السنة والفصل والودائع تم إضافة عمود يوضح الفترة الزمنية وهى تأخذ القيم (1 و 2 و 3 و ... و 10)، ويمكن رسم الشكل الانتشارى للبيانات الربع سنوية الخاصة بإجمالي الودائع في المصارف السعودية باتباع الخطوات التالية :-

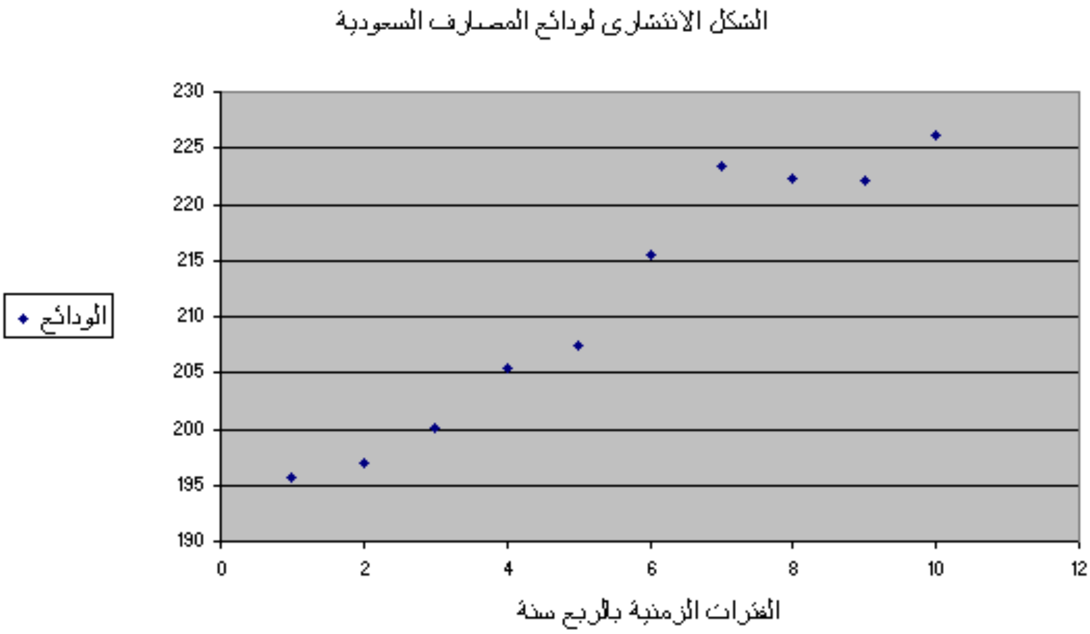
١. يتم تحديد العموديين الخاصين بالفترات الزمنية والودائع المطلوب رسم الشكل الإنتشارى لهما كما بالشكل التالى:

	A	B	C	D
1	السنة	الفصل	الفترة الزمنية	الودائع
2	2005	3	1	195.64
3		4	2	196.97
4	2006	1	3	200.11
5		2	4	205.33
6		3	5	207.49
7		4	6	215.46
8	2007	1	7	223.36
9		2	8	222.31
10		3	9	222.07
11		4	10	226.18

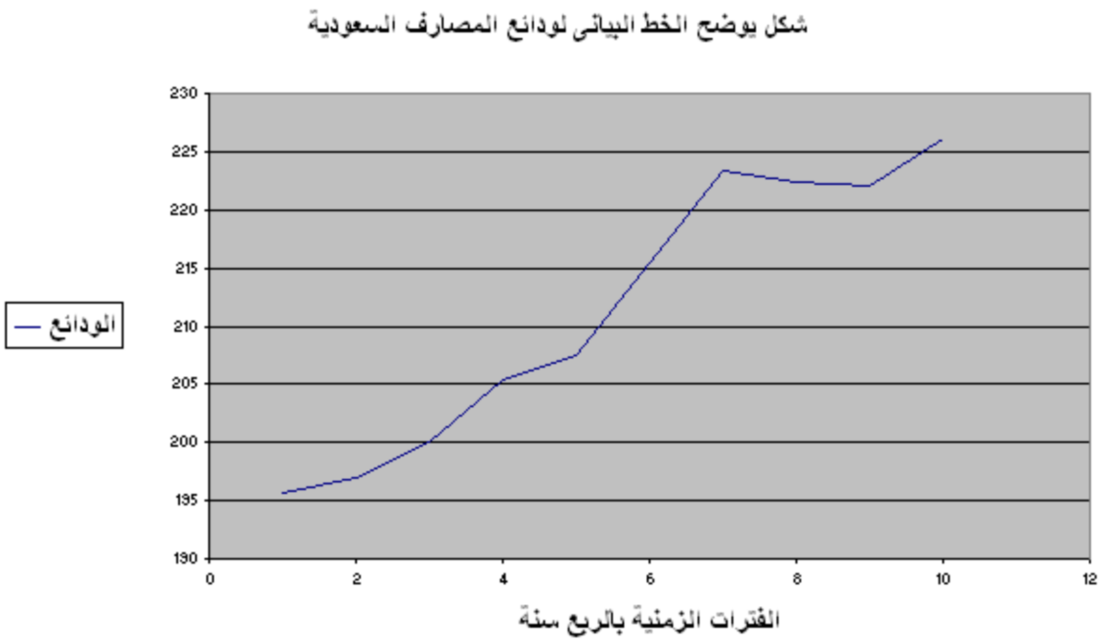
2. ثم نختار من قائمة الرسومات البيانية Wizard Chart رسم الشكل الانتشارى من خلال xy (Scatter) كما بالشكل التالى:



وباستكمال باقى الخطوات يظهر لنا الشكل البيانى التالي:



- كما يمكن رسم الخط البيانى الخاص ببيانات إجمالي الودائع في المصارف السعودية ليكون كما يلى :-



- فعند رسم شكل الانتشار لهذه البيانات كما يبدوا في الشكل السابق، نستطيع من خلال هذا الرسم لتوضيح التالي:

١. يتبين لنا أن هناك ارتفاع مستمر في إجمالي الودائع عبر الزمن
٢. الاتجاه العام لبيانات إجمالي الودائع يمكن وصفه بدالة خطية
٣. ميل خط الاتجاه العام لبيانات إجمالي الودائع سيكون موجبا

ب- طريقة المتوسطات المتحركة:

تعتمد هذه الطريقة على اخذ متوسطات متتابعة لمجموعات متتابعة ومتداخلة من البيانات، والهدف الأساسي من ذلك هو إزالة التعرجات من خط السلسلة الزمنية. وهذه الطريقة أكثر دقة في تحديد خط الاتجاه العام من طريقة شكل الانتشار (التمهيد باليد) .

ويتم حساب المتوسط المتحرك من خلال تطبيق قانون المتوسط الحسابي بشكل متتابع لعدد المشاهدات المعطاة لدينا، مع الأخذ في الاعتبار طول المجموعة التي يتم تقسيم البيانات إليها فمثلا اذا كان طول المجموعة 5 يتم إيجاد متوسط المشاهدات (x1 , x2 , x3 , x4 , x5) وذلك بإيجاد مجموعهم والقسمة على عددهم كما يبدوا ذلك من خلال العلاقة التالية:

$$\frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}$$

ونضع المتوسط الذى تم الحصول علية أمام الفئة التى فى المنتصف وهى امام المشاهدة x3 ثم نحسب المتوسط من جديد للمشاهدات (x2 , x3 , x4 , x5, x6) **ونضع المتوسط الجديد الذى تم حسابة أمام المشاهدة x4** . وهكذا حتى نصل إلى المتوسط الأخير في البيانات المعطاة، وبعد حساب المتوسطات المتحركة نقوم برسم خط الاتجاه العام من هذه المتوسطات المتحركة، وقد ينتج عن ذلك خط غير ممهد كما يجب، وفي هذه الحالة لا يرسم الخط، بل تؤخذ متوسطات ثانية للمتوسطات المتحركة الأولى ويرسم الخط من النقاط التي تمثل المتوسطات المتحركة الثانية لأنها تعطي خطا أكثرا تمهيدا. ويكون أسلوب المتوسط المتحرك فعالآ عندما تكون بيانات السلسلة الزمنية مستقرة عبر الزمن

مثال :

أوجد المتوسطات المتحركة بطول (5) للسلسلة الزمنية التالية :

المشاهدة	X1	X2	X3	X4	X5	X6	X7
قيمتها	7	13	21	23	27	19	17

الحل :

يتم اولاً إيجاد متوسط الخمس مشاهدات والتى يكون مركزها هو x3 وكان الناتج هو 18.2 . ثم نحسب المتوسط مرة اخرى بداية من x2 حتى x6 والتى يكون مركزها x4 وكان الناتج هو 20.6 وهكذا ونتوقف حين لا يمكن لنا تكوين سلسلها طولها 5 مشاهدات، وتظهر لنا النتيجة كما يبدوا ذلك في الجدول التالي:

المشاهدات	القيمة	المتوسط المتحرك
X1	7	
X2	13	
X3	21	18.2
X4	23	20.6
X5	27	21.4
X6	19	
X7	17	

- وبعد حساب المتوسطات المتحركة نقوم برسم خط الاتجاه العام من هذه المتوسطات المتحركة.

ج- طريقة متوسط نصف السلسلة:

تعتبر هذه الطريقة أدق من طريقة شكل الانتشار وطريقة المتوسطات المتحركة، ويمكن حسابها من خلال إتباع الخطوات التالية:

- نقسم السلسلة إلى مجموعتين وفق تسلسل السنوات.
- لتعيين الإحداثي الصادي للنقطتين نوجد المتوسط الحسابي لنصف السلسلةالأول إذا كان عدد المشاهدات زوجي، أما إذا كان عدد المشاهدات فردي فتهمل المشاهدة الوسطى ثم نجد المتوسط الحسابي للنصف الثاني.
- لتحديد الإحداثي السيني نعطي قيم المشاهدات ترقيم متسلسل سواء كانت المشاهدات قيما أو غير ذلك، ثم نجد المتوسط الحسابي للنصف الأول من القيم سواء كان عددها زوجي أو فردي فيكون المتوسط هو الإحداثي السيني، وكذلك حساب المتوسط الحسابي للنصف الثاني والذي يمثل الإحداثي السيني وبذا تتعين النقطتين.
- نصل بين النقطتين بعد تعيينهما على مستوى الإحداثي فيكون لدينا خط الاتجاه العام
- نوجد معادلة خط الاتجاه العام من خلال العلاقة التالية:

$$\frac{Y - Y_1}{X - X_1} = \frac{Y_2 - Y_1}{X_2 - X_1}$$

مثال :

إذا كان إنتاج مصنع سيارات (بالآلاف) خلال عشر سنوات كالتالي:

السنة (X)	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
عدد السيارات (Y)	53	64	67	60	69	74	67	79	85	90

المطلوب:

إيجاد معادلة خط الاتجاه العام بطريقة متوسط نصف السلسلة

الحل :

نكون الجدول التالي من الجدول الرئيسي:

السنة	السنة بالترقيم (X)	عدد السيارات المنتجة (Y)	متوسط نصف بالترقيم (X)	متوسط نصف (Y)
1998	1	53	$X_1 = 3$	$Y_1 = 62.6$
1999	2	64		
2000	3	67		
2001	4	60		
2002	5	69		
2003	6	74	$X_2 = 8$	$Y_2 = 79$
2004	7	67		
2005	8	79		
2006	9	85		
2007	10	90		

المتوسط الأول لنصف	X_1	$X_1 = \frac{1+2+3+4+5}{5} = \frac{15}{5} = 3$
المتوسط الثاني لنصف	X_2	$X_2 = \frac{6+7+8+9+10}{5} = \frac{40}{5} = 8$
المتوسط الأول لنصف	Y_1	$Y_1 = \frac{53+64+67+60+69}{5} = \frac{313}{5} = 62.6$
المتوسط الثاني لنصف	Y_2	$Y_2 = \frac{74+67+79+85+90}{5} = \frac{395}{5} = 79$

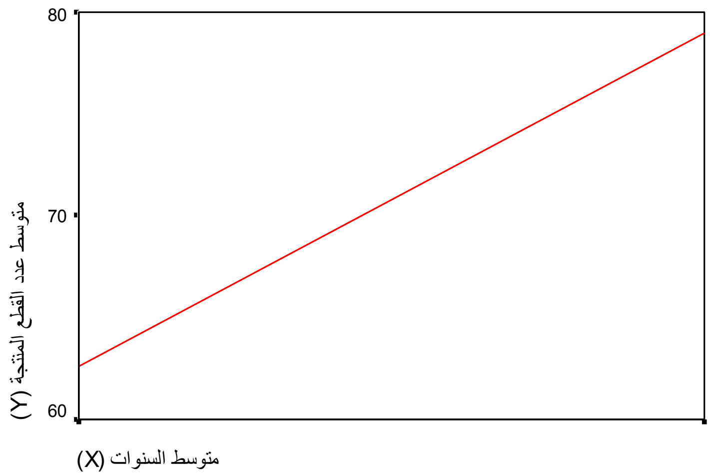
إذا النقطتين المطلوبتين لتحديد الإحداثي السيني والصادي هما :

(3 ، 62.6) ونسميها بالنقطة (أ) ، و (8 ، 79) ونسميها بالنقطة (ب)

نعين النقطتين على الرسم البياني بحيث يكون إحداثي النقطة الأولى هو (3 ، 62.6) وإحداثي النقطة الثانية هو (8 ، 79) ثم نصل بين النقطتين بخط مستقيم فيكون هو خط الاتجاه العام كما يبدوا ذلك في الشكل التالي :

خط الاتجاه العام

من خلال طريقة متوسط نصف السلسلة



- نجد معادلة خط الاتجاه العام من خلال تطبيق العلاقة التالية:

$$\frac{Y-Y_1}{X-X_1}=\frac{Y_2-Y_1}{X_2-X_1}=\frac{Y-62.6}{X-3}=\frac{79-62.6}{8-3}=\frac{16.4}{5}=\frac{Y-62.6}{X-3}=\frac{16.4}{5}$$

وبضرب طرفي المعادلة كالتالي:

$$5Y-(62.6*5)=16.4X-(16.4*3)$$

$$5Y-313=16.4X-49.2$$

$$5Y=16.4X-(49.2)+(313)$$

$$5Y=16.4X+263.8$$

$$Y=\frac{16.4}{5}X+\frac{263.8}{5}$$

$$Y=3.28X+52.76$$

وهذه هي معادلة خط الاتجاه العام من خلال طريقة متوسط نصف السلسلة

د ـ طريقة المربعات الصغرى :

وتعتبر طريقة المربعات الصغرى أكثر دقة من الطرق السابقة لحساب خط الاتجاه العام وذلك من خلال استخدام أسلوب الانحدار الخطي البسيط المعتمد على طريقة المربعات الصغرى التي تجعل مجموع مربعات انحرافات القيم المقدرة عن القيم الفعلية أقل ما يمكن وذلك من خلال العلاقة التالية :

$$\hat{y}_t = b_0 + b_1 t$$

حيث أن :

القيمة الاتجاهية للسلسلة الزمنية في الفترة t

نقطة تقاطع خط الاتجاه العام مع المحور الصادي أو الجزء الثابت

ميل خط الاتجاه العام

الزمن

- ولغرض حساب b_1 و b_0 نقوم بتطبيق المعادلتين التاليتين:

$$b_1 = \frac{n\sum t y_t - \sum t \sum y_t}{n\sum t^2 - (\sum t)^2}$$

$$b_0 = \frac{\sum y_t}{n} - b_1 \frac{\sum t}{n}$$

حيث أن :

القيمة الفعلية للسلسلة الزمنية في الفترة t

عدد الفترات

مثال :

بدراسة احد الظواهر الاجتماعية والمتمثلة فى العنف الأسرى لأحد المدن. تبين أن تطور أعداد الأسر التي يوجد بها عنف أسرى كانت كما يلى خلال مدة الدراسة

السنة	2004	2005	2006	2007	2008	2009	2010
عدد الأسر	17	25	33	41	39	48	53

المطلوب:

١. تقدير معادلة الاتجاه العام لتطور أعداد الأسر المعرضة لظاهرة العنف الأسرى بهذه المدينة
٢. ما هو عدد الأسر المتوقع المعرضون لهذه الظاهرة فى عام 2013 ؟

الحل :

حتى يمكن تقدير معادلة الاتجاه العام لتطور أعداد الأسر المعرضة لظاهرة العنف الأسرى بهذه المدينة لابد من أعداد الجدول التالى علي اعتبار أن السنة الأولى تكون قيمة t فيها تساوى 1 والسنة الثانية تكون قيمتها 2 وهكذا كما يلي:

السنوات	y	t	y t	t^2
2004	17	1	17	1
2005	25	2	50	4
2006	33	3	99	9
2007	41	4	164	16
2008	39	5	195	25
2009	48	6	288	36
2010	53	7	371	49
المجموع	256	28	1184	140

كما يتضح لنا أن عدد المشاهدات n = 7

- ولغرض حساب b_1 و b_0 نقوم بتطبيق المعادلتين التاليتين:

$$b_1 = \frac{n\sum ty_t - \sum t \sum y_t}{n\sum t^2 - (\sum t)^2}$$

$$= \frac{7(1184)-(28\times 256)}{7(140)-(28\wedge 2)}$$

$$= \frac{1120}{196} = 5.714$$

ويدل ذلك علي أن معدل التزايد السنوى في الأسر المعرضة للعنف الأسرى 5.714 اسرة

$$b_0 = \frac{\sum y_t}{n} - b_1 \frac{\sum t}{n}$$

$$= \frac{256-(5.714\times 28)}{7} = 13.715$$

- وعلى ذلك تكون معادلة الاتجاه العام كما يلى:

$$\hat{y}_t = b_0 + b_1 t$$

$$\hat{y}_t = 13.715 + 5.714\, t$$

ما هو عدد الأسر المتوقع المعرضون لهذه الظاهرة فى عام 2013 ؟

حتى يمكن تقدير عدد الأسر المتوقع المعرضون لهذه الظاهرة فى عام 2013 لابد من تحديد قيمة t فى هذه السنة كما يلى:

t	السنة
7	**2010**
8	**2011**
9	**2012**
10	**2013**

وعلي ذلك يتم التعويض في معادلة الاتجاه العام عن قيمة t تساوي 10

$$\hat{y}_t = 13.715 + 5.714(10) = 70.855$$

ويتضح لنا مما سبق أن العدد المتوقع للأسر المعرضة للعنف الأسرى يبلغ 70.855 أى ما يقرب من 71 أسرة فى عام 2013

المحاضره الثانيه عشر (الجزء الثاني)

السلاسل الزمنية

2- التغيرات الموسمية Seasonal Variations

التغيرات الموسمية هي نتيجة طبيعية لاختلاف الظروف بشكل منتظم مما يؤثر على اختلاف رغبات الناس تبعا لعوامل عديدة منها الزمان والمكان، ويمكن تعريفها بأنها التغيرات التي تطرأ على الظاهرة على مدار المواسم المختلفة للفترة الزمنية موضوع القياس (الموسم) ، فهي قد تكون يومية، وقد تكون اسبوعية، وقد تكون شهرية.

مما سبق نرى أن التغيرات الموسمية تحدث في مواعيد زمنية محددة ولا يلبث هذا التغير أن يستعيد سيرته الأولى في نفس المواعيد وعلى مدار نفس الفترة الزمنية

والتغير الموسمي يعتبر ابسط انواع التغيرات في السلاسل الزمنية حيث يشتمل على نماذج متكررة بانتظام، وهي تغيرات تتميز بالطبيعة الدورية بشرط أن لايزيد طول الدورة المتكررة عن سنة واحدة كحد أعلى .

وتكمن أهمية دراسة التغيرات الموسمية في تحليل السلسلة الزمنية للظاهرة خاصة فيما يتعلق بالتخطيط لعمليات الإنتاج أو الأوقات المناسبة للإعلانات عن السلع أو التوسع في المشاريع، فالتغيرات الموسمية بشكل عام تساعد على الكشف عن:

- الأوقات المناسبة للتغيير
- مسببات التغيير
- الاستعدادات المناسبة لمواجهة التغيير

- ويتم قياس التغيرات الموسمية عن طريق إيجاد قيمة الظاهرة في كل موسم من المواسم التي تتعرض لها الظاهرة للتغير ثم تنسب كل قيمة للمتوسط العام لقيم هذه الظاهرة، إذ يتم اعتبار المتوسط العام (100%) فنحصل على أرقام تدل على مدى التغيرات للظاهرة هل هي فوق المتوسط أو دونه، مثال على ذلك ما يذاع عن درجات الحرارة المتوقعة في النشرات الجوية من أنها فوق المتوسط أو دون المتوسط، ولحساب الآثار الموسمية هناك طريقتان:

- طريقة النسب للمتوسط المتحرك
- طريقة إيجاد القيم المخلصة من أثر الاتجاة العام وذلك بقسمة طرفي المعادلة على (Tt)

- طريقة النسب للمتوسط المتحرك، ويتم ذلك من خلال المعادلة التالية:

$$y_t = T_t \times C_t \times S_t \times R_t$$

- طريقة إيجاد القيم المخلصة من أثر الاتجاة العام وذلك بقسمة طرفي المعادلة على (Tt) والتي تمثل تأثير الاتجاه العام فنحصل بالتالي على المعادلة التالية :

$$\frac{y_t}{T_t} = C_t \times S_t \times R_t$$

مثال :

إذا كان لدينا إنتاج إحدى الشركات خلال ثلاث سنوات، وكانت كمية الإنتاج مأخوذة كل ثلاثة شهور (السنة مقسمة إلى أربعة أرباع) والإنتاج بآلاف الوحدات كما يبدوا ذلك في الجدول التالي :

ربع السنة	2008	2009	2010
الأول	3	4	8
الثاني	7	5	10
الثالث	9	6	12
الرابع	2	4	6

المطلوب :

١. تقدير معادلة الاتجاه العام للعلاقة بين الانتاج و الزمن؟
٢. تقدير القيم الإتجاهية المقابلة للقيم الأصلية للأنتاج؟
٣. إيجاد القيم المخلصة من أثر الأتجاة العام؟
٤. تحديد تأثير كل موسم؟
٥. تقدير الانتاج المتوقع سنة 2012 ؟

الحل :

1- تقدير معادلة الاتجاة العام للعلاقة بين الانتاج و الزمن:

يتم اولا إدخال البيانات السابقة مع أضافة عنصر الزمن t ، ثم يتم حساب العمود y t والعمود t^2 وإيجاد المجاميع الازمة لحساب معامل الانحدار b1 كما يلي:

السنة	الربع	الانتاج y	الزمن t	y t	t^2
2008	الأول	3	1	3	1
	الثانى	7	2	14	4
	الثالث	9	3	27	9
	الرابع	2	4	8	16
2009	الأول	4	5	20	25
	الثانى	5	6	30	36
	الثالث	6	7	42	49
	الرابع	4	8	32	64
2010	الأول	8	9	72	81
	الثانى	10	10	100	100
	الثالث	12	11	132	121
	الرابع	6	12	72	144
المجموع		76	78	552	650

حيث n هى الفترات الزمنية تساوى 12

- نحسب قيمة b_1 من خلال العلاقة :

$$b_1 = \frac{n\sum ty_t - \sum t \sum y_t}{n\sum t^2 - \left(\sum t\right)^2}$$

$$= \frac{12(552)-(78\times 76)}{12(650)-(78\text{^}2)} = 0.40559$$

وبالتالى يكون معدل التزايد كل فترة ربع سنة هو0.40559 ألف وحدة

-نحسب قيمة b_0 من خلال العلاقة

$$b_0 = \frac{\sum y_t}{n} - b_1 \frac{\sum t}{n}$$

$$= \frac{76-(0.40559\times 78)}{12} = 3.69697$$

وعلى هذا تتحدد قيمة معادلة خط الاتجاه العام من خلال العلاقة:

$$\hat{y}_t = b_0 + b_1 t$$

$$\hat{y}_t = 3.69697 + 0.40559\ t$$

2- تقدير القيم الإتجاهية المقابلة للقيم الأصلية للأنتاج:

يمكن إيجاد القيم الاتجاهية بالتعويض فى معادلة الانحدار السابق الحصول عليها بقيم t بداية من 1 و 2 و 3 و ... و 12 وبذلك تكون القيم الاتجاهية هى :

$$\hat{y}_t = 3.69697 + 0.40559\ t$$

السنة	الربع	الانتاج y	الزمن t	القيم الاتجاهية
2008	الأول	3	1	4.10256
	الثاني	7	2	4.50815
	الثالث	9	3	4.91374
	الرابع	2	4	5.31933
2009	الأول	4	5	5.72492
	الثاني	5	6	6.13051
	الثالث	6	7	6.5361
	الرابع	4	8	6.94169
2010	الأول	8	9	7.34728
	الثاني	10	10	7.75287
	الثالث	12	11	8.15846
	الرابع	6	12	8.56405
المجموع		76	78	

3- إيجاد القيم المخلصة من أثر الاتجاه العام :

ويتم حساب القيم المخلصة من أثر الاتجاه العام بقسمة قيم الظاهرة الاصلية على القيم الاتجاهية فتكون النتيجة كما بالجدول السابق.

$$\hat{y}_t = 3.69697 + 0.40559\ t$$

السنة	الربع	الانتاج y	الزمن t	القيم الاتجاهية	القيم المخلصة من أثر الاتجاه العام
2008	الأول	3	1	4.10256	0.7313
	الثاني	7	2	4.50815	1.5527
	الثالث	9	3	4.91374	1.8316
	الرابع	2	4	5.31933	0.376
2009	الأول	4	5	5.72492	0.6987
	الثاني	5	6	6.13051	0.8156
	الثالث	6	7	6.5361	0.918
	الرابع	4	8	6.94169	0.5762
2010	الأول	8	9	7.34728	1.0888
	الثاني	10	10	7.75287	1.2898
	الثالث	12	11	8.15846	1.4709
	الرابع	6	12	8.56405	0.7006
المجموع		76	78		

4- إيجاد تأثير كل موسم:

حتى يمكن إيجاد تأثير كل موسم نعيد ترتيب القيم المخلصة من أثر الاتجاه العام السابق الحصول عليها كما يلي:

الموسم	2008	2009	2010
الأول	0.7313	0.6987	1.0888
الثانى	1.5527	0.8156	1.2898
الثالث	1.8316	0.918	1.4709
الرابع	0.376	0.5762	0.7006

ثم يتم إيجاد متوسط القيم المخلصة من أثر الاتجاه العام لكل ربع للتعبير عن أثر ذلك الموسم

فمثلا:

تأثيرالربع الأول =

$$0.8396=\frac{0.7313+0.6987+1.0888}{3}$$

وهكذا لباقى المواسم فتكون النتيجة كما يلي:

الموسم	2008	2009	2010	تأثير الموسم
الأول	0.7313	0.6987	1.0888	0.8396
الثانى	1.5527	0.8156	1.2898	1.2194
الثالث	1.8316	0.918	1.4709	1.4068
الرابع	0.376	0.5762	0.7006	0.5509
المجموع				4.0167

<u>ونلاحظ</u> أن مجموع تأثيرات المواسم (الدليل الموسمى) 4.0167 اى 401.67 % وحيث يوجد 4 مواسم لذا فأن مجموع تأثيرات المواسم لابد أن تساوى 400 %

- لذا لابد من تعديل قيم الدليل الموسمى بمعامل تصحيح قدرة $\frac{4}{4.0167}$

تأثير الموسم المعدل	تأثير الموسم	الموسم
0.836109	0.8396	الأول
1.21433	1.2194	الثانى
1.400951	1.4068	الثالث
0.54861	0.5509	الرابع
4	4.0167	المجموع

5- تقدير الانتاج المتوقع سنة 2012

نلاحظ أن قيم t فى الربع الاخير سنة 2010 بلغت 12 لذلك يتم الزيادة عليها سنة 2011 لتكون 13 ، 14 ، 15 ، 16 خلال المواسم الاربع ولذلك تكون القيم خلال سنة 2012 هى 17و 18 ، 19 ، 20 والتى يتم التعويض بها معادلة الاتجاه العام للحصول على القيم الاتجاهيه ويمكن تقدير القيم المتنبئ بها لكل ربع كما يلي:

القيم المتنبئ بها للموسم = القيمة الاتجاهية × تأثير الموسم المعدل

$$\hat{y}_t = 3.69697 + 0.40559\ t$$

- وعلى ذلك يمكن تقدير الانتاج المتوقع سنة 2012 كما يلي:

الانتاج المتوقع	تأثير الموسم المعدل	القيمة الاتجاهية	t	الموسم
8.856069	0.836109	10.592	17	الأول
13.35471	1.21433	10.99759	18	الثانى
15.9753	1.400951	11.40318	19	الثالث
6.478404	0.54861	11.80877	20	الرابع
44.66448				المجموع

ويتضح لنا أن الانتاج المتوقع سنة 2012 هو 44664.48 وحدة

3- التغيرات الدورية Cyclical Variations

ويعرف هذا النوع من التغيرات بدورات الأعمال، وهذا يمتد لفترة زمنية أطول من سنة، وتنشأ هذه التغيرات عن ظروف عامة تعزى إلى العوامل التي تتحكم في الحياة الاقتصادية للبلاد.
ويهتم الباحثون الاقتصاديون ورجال الأعمال بالتغيرات الدورية لغايات التخطيط لمواجهة المشاكل التي قد تنشأ عن حدوثها، وقد تمتد بعض التغيرات الدورية إلى 50 سنة وهذه دورة طويلة، أما الدورة المتوسطة فتمتد بين 8-12 سنة، أما الدورة القصيرة فتكون بين 3-4 سنوات، وتقع التقلبات الدورية أعلى وأسفل خط الاتجاه العام .

4- التغيرات العشوائية أو الفجائية Random (Irregular) Variations:

تؤثر هذه التغيرات على السلسلة الزمنية بشكل عشوائي أو مفاجئ وغير منتظم، فقد تكون هذه التغيرات ناتجة عن حدوث ظواهر طبيعية مثل الزلازل والبراكين أو حروب ونحوها، لذا فهذا النوع من التغيرات من الصعب التنبؤ بها ومن الصعب كذلك تحديد حجم هذه التغيرات ومدى تأثيرها على الظاهرة موضع الدراسة، وتمتاز هذه التغيرات بعدد من المميزات منها :

- إنها لا تحدث وفقا لقاعدة أو قانون
- قد تتكرر أو لا تتكرر
- تأثيرها غير ثابت فمرة تأثر بالنقص ومرة بالزيادة
- لا تستمر طويلا لذا يطلق عليها اسم التغيرات قصيرة الأجل

المحاضرة الثالثة عشر

الأرقام القياسية

- تعريف الأرقام القياسية:

الرقم القياسي هو مؤشر إحصائي (رقم نسبي) يستخدم في قياس التغير النسبي الذي طرأ على ظاهرة من الظواهر الاقتصادية أو الاجتماعية، فهو يستخدم لقياس التغير في أسعار السلع أو في حجم انتاجها أو في كميات المبيعات منها أو في حجم السكان أو أجور العمال (وفقا لأساس معين) سواء كان هذا الأساس فترة زمنية معينة أو مكانا معينا

- فترة الأساس:

الأساس هو فترة زمنية معينة أو مكانا معينا، وعادة تكون فترة الأساس فترة سابقة للفترة التي نريد مقارنتها (وفي حالات نادرة جدا قد تكون فترة الأساس فترة لاحقة لفترة المقارنة) **ويجب أن تمتاز فترة الأساس بما يلي :**

- الاستقرار الاقتصادي
- خلوها من العوامل المؤثرة على الأسعار (الحروب)
- أن تكون بعيدة جدا عن سنوات المقارنة

أما عند **اختيار مكان الأساس** لابد أن يكون لهذا المكان أهمية خاصة وأن يكون مركزا أساسيا لإنتاج السلعة المراد استخراج الرقم القياسي لها .

- الأرقام القياسية للأسعار Price Index Numbers

تعتبر الأرقام القياسية للاسعار من أهم أنواع الأرقام القياسية وأكثرها شيوعا، فهي (أي الأرقام القياسية للأسعار) تساهم في قياس التغير في المستوى العام للأسعار أو التغير في تكاليف المعيشة في فترة زمنية معينة مقارنة بفترة زمنية أخرى ومن أشهرها:

- مؤشر أسعار المستهلكين Consumer Price Index ويرمز له (CPI)
- مخفض الناتج القومي الإجمالي Gross National Product Deflator
- مؤشر اسعار المنتجين Producer Price Index ويرمز له (PPI)
- مخفض الناتج المحلي الإجمالي Gross Domestic Product Deflator
- مؤشر أسعار الأسهم

- أمثلة على بعض الأرقام القياسية للاسعار في النظام الاقتصادي السعودي :

يهتم النظام الاقتصادي السعودي بنشر الأرقام القياسية للاسعار وتكاليف المعيشة على شكل تقارير شهرية، ومن هذه الأرقام مايلي:

- الرقم القياسي لتكاليف المعيشة لمتوسطي الدخل: ويشمل هذا الرقم المواد الغذائية، السكن وتوابعه، الأقمشة والملابس، الأثاث المنزلي، الرعاية الطبية، النقل والمواصلات، التعليم والترفيه، النفقات والخدمات الأخرى، والرقم القياسي العام)
- الرقم القياسي لتكاليف المعيشة لجميع السكان: ويشمل المواد الغذائية ، السكن وتوابعه، الأقمشة والملابس، الأثاث المنزلي، الرعاية الطبية، النقل والاتصالات، التعليم والترفيه، النفقات والخدمات الأخرى، والرقم القياسي العام)
- الرقم القياسي لأسعار الجملة: ويشمل المواد الغذائية، المشروبات، مواد الخام ماعدا الوقود، الوقود المعدني وزيوت التشحيم، الدهون والزيوت الحيوانية والنباتية، الكيماويات والمواد ذات الصلة، السلع المصنعة مصنفة حسب المادة، الآلات ومعدات النقل والاتصالات، التعليم والترفيه، النفقات والخدمات الأخرى، والرقم القياسي العام

- دور الأرقام القياسية في حساب معدلات التضخم :

المقصود بالتضخم هو الارتفاع المستمر في المستوى العام للاسعار والذي على ضوئه تنخفض القيمة الشرائية للوحدة النقدية (الريال مثلا)، وتقوم الجهات الاقتصادية في الدول باستخدام الأرقام القياسية للاسعار لإيجاد معدلات التضخم السنوية، وفي معظم الأحيان يستخدم مؤشر اسعار المستهلكين (CPI) لسنتين متتاليتين لحساب معدل التضخم السنوي في السنة الأخيرة وذلك من خلال العلاقة التالية:

$$i_{2010} = \frac{CPI_{2010} - CPI_{2009}}{CPI_{2009}}(100)$$

حيث :

i_{2010} = معدل التضخم في سنة2010م

CPI_{2009} =مؤشر اسعار المستهلكين في سنة2009م

CPI_{2010} =مؤشر اسعار المستهلكين في سنة2010م

مثال :

إذا افترضنا أن مؤشر اسعار المستهلكين في المملكة لسنة 2006م=120 وسنة 2007م=123 ، ما هو معدل التضخم في سنة 2007م

الحل :

معدل التضخم في سنة 2007م يتم حسابه من خلال العلاقة التالية:

$$i_{2007} = \frac{CPI_{2007} - CPI_{2006}}{CPI_{2006}}(100) \qquad = \frac{123-120}{120}(100) = 2.5\%$$

أي أن معدل التضخم في سنة 2007 يساوي 2.5 %

- فوائد الأرقام القياسية واستعمالاتها :

تستخدم الأرقام القياسية عادة لقياس التغير الذي يطرأ على الحياة بمجملها بشكل عام والجوانب الاقتصادية بشكل خاص. كما تساعد الأرقام القياسية على تحليل العوامل التي تساهم في تغير الظاهرة فتبين مدى مساهمة كل من هذه العوامل في إحداث التغير الكلي، وتستخدم كذلك في الرقابة على تنفيذ الخطط .

الرقم القياسي المرجح :

وهو ذلك الرقم الذي يأخذ الأهمية النسبية للسلعة أو الأجر بعين الاعتبار فيعطي كل سلعة (أجر) وزنا يتلاءم مع أهميته، فعند تركيب رقم قياسي للكميات يجب ترجيحه بالأسعار، وعند تركيب رقم قياسي للأسعار يجب ترجيحه بالكميات وبالتالي يكون الناتج رقما قياسيا مرجحا.

- منسوب السعر لسلعة واحدة (ظاهرة بسيطة):

يمكن إيجاد رقم قياسي لسعر سلعة بمفردها (حيث يمثل هذا الرقم القياسي التغير في سعر السلعة أو الخدمة في سنة معينة مقارنة بسنة الأساس)، ويسمى الرقم القياسي للسعر بمنسوب السعر ويرمز له بالرمز Pr ويمكن حسابه بالطريقة التالية :

$$P_r = \frac{P_1}{P0}(100)$$

حيث أن :

P_r = منسوب السعر

P_1 =السعر سنة المقارنه

P_0 =السعر سنة الاساس

مثال :

إذا كانت لدينا البيانات التالية والممثلة لسعر سلعة معينة من الفترة 2006م وحتى 2010م .

السنة	سعر السلعة بالريال
2006	25
2007	30
2008	24
2009	32
2010	36

المطلوووب :

إيجاد منسوب السعر لهذه السلعة للفترة من سنة 2006م حتى سنة 2010م باعتبار سنة 2006م سنة أساس، مع تفسير النتائج التي يتم الحصول عليها .

الحل في الكتاب صفحه 240

- منسوب السعر لمجموعة من السلع-التجميعية (ظاهرة معقدة) :

الرقم القياسي السابق يوضح منسوب السعر لسلعة واحدة، إلا أن كثيرا من الحالات تكون أكثر تعقيدا فقد يكون لدينا عدة سلع متغيرة ونرغب حساب منسوب السعر أو الرقم القياسي لها، ففي حالة استخراج الرقم القياسي لمثل هذا الوضع فإنه يدخل في الحساب جميع قيم السلع التي تتألف منها الظاهرة ويتم ذلك من خلال استخدام الطرق التالية:

- الرقم القياسي التجميعي البسيط للأسعار
- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير)
- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش)
- الرقم القياسي التجميعي المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر)

- حساب الأرقام القياسية التجميعية (مجموعة من السلع):

1- الرقم القياسي التجميعي البسيط للأسعار:

لهذا الرقم القياسي بالرمز "Is " ويتم حسابه من خلال العلاقة التالية:

$$I_s = \frac{\sum P_1}{\sum P_0}(100)$$

حيث أن :

$\sum P_1$ = مجموع اسعار السلع والخدمات في سنة المقارنه .

$\sum P_0$ = مجموع اسعار السلع والخدمات في سنة الاساس .

- وتكمن مشكلة الرقم القياسي التجميعي البسيط للأسعار في أنه لا يعطي للكميات المستهلكة من السلع والخدمات أوزانا، فبالتالي يكون حساسا عندما يكون هناك تباينا في الكميات المستهلكة .

2- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير):

ويسمى برقم لاسبير ويرمز له بالرمز Ir وهذا الرقم يعبر عن أثر التغير في السعر كما لو بقيت الكميات المشتراة في سنة الأساس هي نفسها في سنة المقارنة. ويتم حسابه بنفس الطريقة السابقة مع ترجيح وزن كل سعر بكميته المستهلكة في سنة الأساس، ويتم ذلك من خلال تطبيق العلاقة التالية:

$$I_r = \frac{\sum P_1 Q_0}{\sum P_0 Q_0}(100)$$

حيث أن :

I_r =الرقم القياسي التجميعي للاسعار المرجح بكميات سنة الاساس(رقم لاسبير)

$\sum P_1Q_0$ =مجموع اسعار السلع والخدمات سنة المقارنه مرجحه بكميات سنة الاساس

$\sum P_0Q_0$ =مجموع اسعار السلع والخدمات سنة الاساس مرجحه بكميات سنة الاساس

- ويفضل استخدام هذه الطريقة عند حساب مؤشر اسعار المستهلكين (CPI) وذلك للاقتصاد في الجهد والوقت والمال، لأن كمية سنة الأساس ثابتة عند إيجاد رقم لاسبير لأي سنة لاحقة لسنة الأساس .

3- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش):

ويسمى برقم باش ويرمز له بالرمز Ip وهذا الرقم يعبر عن اثر التغير في السعر كما لو أن الكميات المشتراة في سنة المقارنة كانت قد اشتريت في سنة الأساس. وتختلف طريقة حساب هذا الرقم من حيث أنه يرجح كل سعر بكميته المستهلكة في سنة المقارنة ، ويتم ذلك من خلال تطبيق العلاقة التالية:

$$I_p = \frac{\sum P_1Q_1}{\sum P_0Q_1}(100)$$

حيث أن :

I_p =الرقم القياسي التجميعي للاسعار المرجح بكميات سنة الاساس (رقم باش)

$\sum P_1Q_1$ =مجموع اسعار السلع والخدمات سنة المقارنه مرجحه بكميات سنة المقارنه

$\sum P_0Q_1$ =مجموع اسعار السلع والخدمات سنة الاساس مرجحه بكميات سنة المقارنه

والمشكلة الأساسية في هذه الطريقة هي الحاجة لتحديد الكميات المستهلكة من كل سلعة سنويا حتى يتسنى لنا حساب هذا الرقم .

4- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر):

ويسمى برقم فيشر ويرمز له بالرمز If، وهو عبارة عن الوسط الهندسي لكل من رقمي لاسبير وباش، أي أنه الجذر التربيعي لحاصل ضرب رقم لاسبير برقم باش، (وهذا الرقم يهتم بالناحية الرياضية ولكنه لا معنى اقتصادي له) وهذا هو أهم عيوبه . ويتم ذلك من خلال العلاقة التالية:

$$I_f = \sqrt{I_r \quad \mathrm{I_P}}$$

$$I_f = \sqrt{\frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times \frac{\sum p_1 Q_1}{\sum p_0 Q_1}}$$

مثال لحساب الأرقام القياسية التجميعية :

يبين الجدول التالي أسعار وكميات ثلاث منتجات استهلاكية للسنتين 2007م و 2010م على اعتبار أن سنة 2007م هي سنة الأساس .

السعر P1	الكمية Q1	السعر P0	الكمية Q0	السنوات / المنتجات
سنة **2010م** (سنة المقارنة)		سنة **2007م** (سنة الأساس)		
12	8500	9	5000	**السلعة الأولى**
31	15000	25	8000	**السلعة الثانية**
17	19000	14	9000	**السلعة الثالثة**

المطلوب :

- حساب الرقم التجميعي البسيط للأسعار.
- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس (رقم لاسبير).
- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة المقارنة (رقم باش).
- الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأساس وسنة المقارنة (رقم فيشر).
- تفسير نتائج الفقرات السابقة.

الحل في الكتاب صفحة 243 - 246:

- ملاحظات عامة على الأرقام القياسية :

هناك مجموعة من الملاحظات المتعلقة بتفسير الأرقام القياسية لسنوات الأساس والمقارنة، وهذه الملاحظات كالتالي:

- الرقم القياسي للظاهرة في سنة الأساس يساوي 100 .
- إذا كان الرقم القياسي للظاهرة في سنة المقارنة أكبر من 100 فهذا يعني أن هناك ارتفاع في المستوى العام للظاهرة مقارنة بسنة الأساس .
- إذا كان الرقم القياسي للظاهرة في سنة المقارنة أصغر من 100 فهذا يعني أن هناك انخفاض في المستوى العام للظاهرة مقارنة بسنة الأساس

بسم الله الرحمن الرحيم

المحآضره الآولى

س / **الاحصاء في اللغه هو ؟**

١- البحث عن الحقيقه

٢- العدد الشامل

٣- فرع من فروع الرياضيات يهدف الى جمع وعرض ووصف وتحليل البيانات المقاسه رقميا مما يساعد في اتخاذ القرارات

٤- ابراز الخصائص الاساسية

الآجآبـه ((١))

س / **الغرض من العلم هو ؟**

١- البحث عن الحقيقه

٢- العدد الشامل

٣- فرع من فروع الرياضيات يهدف الى جمع وعرض ووصف وتحليل البيانات المقاسه رقميا مما يساعد في اتخاذ القرارات

٤- ابراز الخصائص الاساسية

الآجآبـه ((١))

س / **فرع من فروع الرياضيات يهدف الى جمع وعرض ووصف وتحليل البيانات المقاسه رقميا مما يساعد في اتخاذ القرارات ؟**

١- الاحصاء الوصفي

٢- الاحصاء الاستنتاجي

٣- الاحصاء في الاصطلاح

٤- الاحصاء في اللغه

الآجآبـه ((٣))

المحآضره الثانية

س / المجتمع ؟
١- هي جزء من المجتمع الاحصائي محل دراسه اخير بطريقة عملية ليتم اجراء الدراسة عليه
٢- جميع المفردات التي يجمعها اطار عام واحد او مجموعة خصائص عامه واحده
٣- هي القيمه الوصفية او الرقمية التي تحتاج اليها لمساعدتنا في جعل القرارات التي نتخذها اكثر معلوماتية في موقف محدد
٤- هي خاصية عن المجتمع الاحصائي والتي يتم اختبارها من خلال التحليل الاحصائي

الآجآبـه ((٢))

س / العينة ؟
١- هي جزء من المجتمع الاحصائي محل دراسه اخير بطريقة عملية ليتم اجراء الدراسة عليه
٢- جميع المفردات التي يجمعها اطار عام واحد او مجموعة خصائص عامه واحده
٣- هي القيمه الوصفية او الرقمية التي تحتاج اليها لمساعدتنا في جعل القرارات التي نتخذها اكثر معلوماتية في موقف محدد
٤- هي خاصية عن المجتمع الاحصائي والتي يتم اختبارها من خلال التحليل الاحصائي

الآجآبـه ((١))

س / البيانات ؟
١- هي جزء من المجتمع الاحصائي محل دراسه اخير بطريقة عملية ليتم اجراء الدراسة عليه
٢- جميع المفردات التي يجمعها اطار عام واحد او مجموعة خصائص عامه واحده
٣- هي القيمه الوصفية او الرقمية التي تحتاج اليها لمساعدتنا في جعل القرارات التي نتخذها اكثر معلوماتية في موقف محدد
٤- هي خاصية عن المجتمع الاحصائي والتي يتم اختبارها من خلال التحليل الاحصائي
الآجآبـه ((٣))

س / المتغير ؟
١- هي جزء من المجتمع الاحصائي محل دراسه اخير بطريقة عملية ليتم اجراء الدراسة عليه
٢- جميع المفردات التي يجمعها اطار عام واحد او مجموعة خصائص عامه واحده
٣- هي القيمه الوصفية او الرقمية التي تحتاج اليها لمساعدتنا في جعل القرارات التي نتخذها اكثر معلوماتية في موقف محدد
٤- هي خاصية عن المجتمع الاحصائي والتي يتم اختبارها من خلال التحليل الاحصائي

الآجآبـه ((٤))

س / المعلمة ؟
١- هي قياس وصفى لآحد المتغيرات يتم بآستخدام بيانات العينة والتي تكون تقدير لمعلمة المجتمع
٢- هي قياس وصفى لآحد المتغيرات يتم بآستخدام بيانات المجتمع الاحصائي كله
٣- هي القيمه الوصفية او الرقمية التي تحتاج اليها لمساعدتنا في جعل القرارات التي نتخذها اكثر معلوماتية في موقف محدد
٤- هي خاصية عن المجتمع الاحصائي والتي يتم اختبارها من خلال التحليل الاحصائي

الآجآبـه ((٢))

س / الاحصائية ؟
١- هي قياس وصفى لآحد المتغيرات يتم بآستخدام بيانات العينة والتي تكون تقدير لمعلمة المجتمع
٢- هي قياس وصفى لآحد المتغيرات يتم بآستخدام بيانات المجتمع الاحصائي كله
٣- هي القيمه الوصفية او الرقمية التي تحتاج اليها لمساعدتنا في جعل القرارات التي نتخذها اكثر معلوماتية في موقف محدد
٤- هي خاصية عن المجتمع الاحصائي والتي يتم اختبارها من خلال التحليل الاحصائي
الآجآبـه ((١))

س / هدف ادوات جمع البيانات للمصادر الميدانية ؟
١- اختبار فرضيات البحث
٢- الاجابة عن تساؤلاته
٣- لاشي مما سبق
٤- الاجابه ١-٢

الآجآبـه ((٤))

س / قابله للكسر مثل درجة التحصيل ودرجة الذكآء ؟
١- بيانات وصفية ترتيبية
٢- بيانات وصفية اسمية
٣- بيانات كمية متقطعه
٤- بيانات كمية متصلة
الآجآبـه ((٤))

س / غير قآبله للكسور وتعتمد على الارقام الصحيحة مثلآ عدد الطلآب في الفصل ؟
١- بيانات وصفية ترتيبية
٢- بيانات وصفية اسمية
٣- بيانات كمية متقطعه
٤- بيانات كمية متصلة
الآجآبـه ((٣))

س / بيانات للدلاله على الشي مثل تحديد الجنس ذكر ، آنثى هي ؟
١- بيانات وصفية ترتيبية
٢- بيانات وصفية اسمية
٣- بيانات كمية متقطعه
٤- بيانات كمية متصلة
الآجآبـه ((٢))

س / آجرآء ممتآز ' جيد جدآ ، جيد ، مقبول مثآل على ؟

١- بيانات وصفية ترتيبية

٢- بيانات وصفية اسمية

٣- بيانات كمية متقطعه

٤- بيانات كمية متصلة

الآجآبـه ((١))

س / يوجد اربعه من مستويات القياس للبيانات منهى مايستخدم للعلوم الطبيعية وهو ؟

١- الاسمي

٢- الرتبي

٣- الفتري

٤- النسبي

٥- الآجآبة ١-٢-٣

الآجآبـه ((٤))

س / يوجد اربعه من مستويات القياس للبيانات منهى مايستخدم للعلوم الآسآسية وهو ؟

١- الاسمي

٢- الرتبي

٣- الفتري

٤- النسبي

٥- الآجآبة ١-٢-٣

الآجآبـه ((٥))

١- الاسمي : يستخدم للدلاله على شي مثال .. تحديد الجنس لايمكن اجراء عمليات حسابية .
٢- الرتبي : يرتقى قليلا على المقياس الاسمي ويستخدم الترقيم لترتيب الاشياء ويحدد الافضليه لكن لايحدد المسافات بين القيم لايمكن اجراء عمليات حسابية .
٣- الفتري :(يسمى بالفئوي) يحدد الفروق ويحدد القيمه وعيبه عدم وجود صفر حقيقي والصفر الحقيقي معناه هو عدم وجود صفه مقاسه من الصفات التي نرغب بقياسه
مثآل متغير الذكاء من غير الصفر بحيث اذا اتى صفر لاتسمى بالفترية
٤- النسبي : يعتبر من آفضل انواع المقاييس . يستخدم للعلوم الطبيعية . آرتبآطه بالصفر الحقيقي
مثآل : الوزن

الاستبانه هي مجموعه من الاسئله المكتوبه والتي يطلق من المفحوص الاجابه عليها بانفسهم ومميزاتها السرعه و من خلالها لاتوجد بها حرج وفرصه لمراجعة الاستبانه ومن عيوبها لابد من الشخص القراءه والكتابه عدم التقيد بالترتيب

المقابله : هي مجموعه من الاسئله المقروءه ويتم الاجابه عليه من قبل المستجيب شفهيآ ومميزاتها تصلح للاشخاص الذين لايكتبون ولا يقرءون والاطفال عيوبها بطيئه جدا وفيها بعض الحرج

المحاضرة الثآلثـه

س / يعتبر من الاسلوب المناسب للتعداد ؟
١- اسلوب العينات
٢- اسلوب الحصر الشامل
٣- عينة البحث
٤- مجتمع البحث
الآجآبـه ((٢))

س / من نتائجها التي نحصل عليها لايوجد بها تحيز ولا تحتاج لتعديل. وتآخذ وقت طويل وجهد كبير ؟
١- اسلوب العينات
٢- اسلوب الحصر الشامل
٣- عينة البحث
٤- مجتمع البحث
الآجآبـه ((٢))

س / يوفر الوقت والجهد والتكاليف ويصلح للمجتمعات غير المحدوده ؟
١- اسلوب العينات
٢- اسلوب الحصر الشامل
٣- عينة البحث
٤- مجتمع البحث
الآجآبـه ((١))

س / اهم عيوب اسلوب العينات يسمى بـ ؟
١- خطآ المرجعي
٢- الخطا الشامل
٣- خطا التحيز
٤- الجابه ١-٢
الآجآبـه ((٣))

س / جزء من المجتمع اختير بطريقة عملية بشرط ان تمثل المجتمع ككل ؟
١- اسلوب العينات
٢- اسلوب الحصر الشامل
٣- عينة البحث
٤- مجتمع البحث
الآجآبـه ((٣))

س / هو مصطلح منهجي يراد به كل من يمكن ان تعمم عليه نتائج البحث ؟
١- اسلوب العينات
٢- اسلوب الحصر الشامل
٣- عينة البحث
٤- مجتمع البحث
الآجآبـه ((٤))

س / من طرق اختيار العينات في المجتمع المعروف الغير متجانس؟
١- العينه الطبقية
٢- العينه الحصية
٣- العينه العشوائية
٤- العينه العمديه
الآجآبـه ((١))

س / من طرق اختيار العينات في المجتمع المعروف متجانس؟
١- العينه الطبقية
٢- العينه الحصية
٣- العينه العشوائية
٤- العينه العمديه
الآجآبـه ((٣))

س / من طرق اختيار العينات في المجتمع غير المعروف الغير متجانس؟
١- العينه الطبقية
٢- العينه الحصية
٣- العينه العشوائية
٤- العينه العمديه
الآجآبـه ((٢))

س / من طرق اختيار العينات في المجتمع غير المعروف متجانس؟
١- العينه الطبقية
٢- العينه الحصية
٣- العينه العشوائية
٤- العينه العمديه
الآجآبـه ((٤))

س / spss تعني؟
١- حزمه البرامج الاحصائيه للعلوم الفكرية
٢- حزمه البرامج الاحصائيه للعلوم الفلكية
٣- حزمه البرامج الاحصائيه للعلوم الطبيعية
٤- حزمه البرامج الاحصائيه للعلوم الانسانية
الآجآبـه ((٤))

س / تسمى اجابات الافراد على الاسئله (الفقرآت)؟
١- حالات
٢- تمثل متغير
٣- بقيم المتغير
٤- اسباب
الآجآبـه ((٣))
س / كل سؤوال في (فقره) في الاستبانه؟
١- حالات
٢- تمثل متغير
٣- بقيم المتغير
٤- اسباب
الآجآبـه ((٢))

س / الافراد الذين يقومون بالاجابة على اسئله الاستبانه يطلق عليهم اسم ؟

١- حالات
٢- تمثل متغير
٣- بقيم المتغير
٤- اسباب

الآجآبـه ((١))

عندما تكون لدينا البيانات التالية :

المستوى الأكاديمي	عدد الطالبات
المستوى الأول	٢٢٠
المستوى الثاني	١٨٠
المستوى الثالث	١٦٠
المستوى الرابع	١٤٠
المجموع	٧٠٠

١- إذا أردنا أن نمثل هذه البيانات باستخدام اللوحة الدائرية، فإن قيمة زاوية القطاع للمستوى الثالث هي:

أ- ٩٦,٥

ب- ٦٧,٤

ج- ٧٩,٥

د- ٨٢,٣

القانون : زاوية القطاع (١٦٠) ÷ مجموع القطاعات (٧٠٠)×٣٦٠