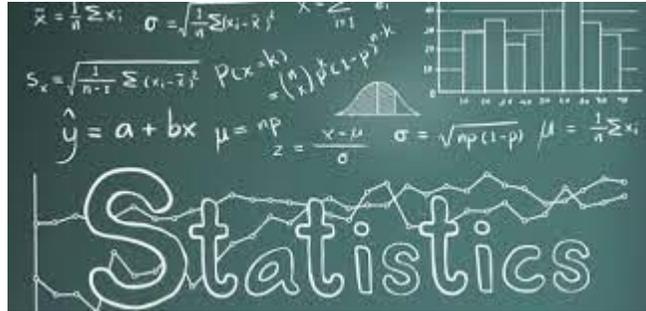


جامعة الفرات الأوسط التقنية
المعهد التقني كربلاء
قسم تقنيات المحاسبة



الحقيبة الدراسية لمادة الاحصاء
اعداد

مدرس المادة /م.م. حيدر جميل احمد
للعام الدراسي 2022 - 2023



مفردات مادة الإحصاء

الاسبوع	تفاصيل المفردات
الأول	علم الإحصاء, تعريفه , اهميته, علاقته بالعلوم الاخرى statistics تعريف الطريقة الاحصائية , مراجعة الطريقة الاحصائية .
الثاني والثالث	تصنيف وتبويب البيانات , تكوين الجداول التكرارية البسيطة والمزدوجة . Classification and Tabulation of data
الرابع والخامس	العرض البياني للبيانات المبوبة , المدرج التكراري , المصّلع التكراري , المنحنى التكراري , المنحنى التكراري للمتجمع الصاعد والمتجمع النازل
السادس	قياس النزعة المركزية , مفهومها واستخداماتها , الوسط الحسابي في البيانات غير المبوبة والبيانات المبوبة (طريقة مطولة) طريقة مختصرة . Averages or Measures of Central Tendency
السابع والثامن	الوسيط , تعريفه , طرق حسابه للبيانات غير المبوبة والمبوبة حسابيا وبيانيا . المنوال , مفهومه , حسابه للبيانات غير المبوبة والمبوبة (طريقة بيرسون حسابيا وبيانيا) . The Median
التاسع	مقاييس التشتت , مفهومها واستخداماتها , المدى للبيانات غير المبوبة والمبوبة , الانحراف الربيعي للبيانات غير المبوبة .
العاشر	الانحراف الربيعي للبيانات المبوبة حسابيا وبيانيا . Semi-Inter –Quartile – Range
الحادي عشر الثاني عشر	الانحراف المتوسط , مفهومه واهميته , طرق حسابه للبيانات غير المبوبة الانحراف القياسي (المعياري), مفهومه واهميته , طرق حسابه للبيانات غير المبوبة والمبوبة Standard Deviation
الثالث عشر و الرابع عشر والخامس عشر والسادس عشر والسابع عشر الثامن عشر والتاسع عشر	الارتباط البسيط , مفهومه , طرق حسابه للبيانات غير المبوبة (طريقة مطولة وطريقة مختصرة) . Simple Correlation Coefficient . معامل ارتباط person للبيانات المبوبة ارتباط الرتب Rank - Correlation , ارتباط سبيرمان للرتب sparman , sparmans rank correlation .coeff . ارتباط سبيرمان المعدل . ارتباط البيانات الصفات correlation between attributes معامل الاقتران . Coefficient of Association معامل التوافق . Coefficient of Contingency.
العشرون	السلاسل الزمنية - مفهومها . استخداماتها Time Series
الحادي والعشرون والثاني والعشرون	الاتجاه العام , مفهومه , طرق ايجاده Secular treand أ. طريقة المتوسطات المتحركة . ب. طريقة متوسطي نصفي السلسلة.
الثالث والعشرين	الارقام القياسية , مفهومها , واستخداماتها , index numbers

حساب الارقام القياسية البسيطة .simple index numbers	الرابع والعشرين
حساب الأرقام القياسية المرجحة Weighted index numbers	والخامس
- رقم الاسبير	والعشرين
- رقم باش	والسادس
- رقم فشر (الامتثل)	والعشرين
- بعض المواضيع التطبيقية	السابع والعشرون والثامن والعشرون والتاسع والعشرون والثلاثون

الهدف من دراسة مادة الإحصاء:

تهدف دراسة مادة الإحصاء الى:

- 1 (تعريف الطالب بطبيعة علم الإحصاء وأهميته ومجالات استخدامه).
- 2 (تعريف الطالب بأساليب جمع البيانات وطرق اخذ العينات ومن ثم طرائق عرضها بالطرق البيانية).

المقدمة

إنَّ علمَ الإحصاءِ الحديثَ بعيدٌ كلُّ البُعدِ عن كونه ذلك المجالَ الجافَّ والمُملِّ الذي يتصوَّرُه العامَّةُ؛ فهو في حقيقةِ الأمرِ موضوعٌ مُثيرٌ يستخدمُ نظريَّةً عميقةً وأدواتٍ برمجيةً قويةً لِيُسلِّطَ الضوءَ على كلِّ جوانبِ حياتنا تقريباً؛ من علمِ الفلكِ إلى الأبحاثِ الطبيَّةِ، ومن علمِ الاجتماعِ إلى السياساتِ الحزبيةِ والشركاتِ الكبرى.

الإحصاء هو أحد فروع الرياضيات الهامة ذات التطبيقات الواسعة. وهو علم جمع ووصف وتفسير البيانات وبمعنى آخر صندوق الأدوات الموضوع تحت البحث التجريبي. يهتم علم الإحصاء بجمع وتلخيص وتمثيل وإيجاد استنتاجات من مجموعة البيانات المتوفرة، محاولاً التغلب على مشاكل مثل عدم تجانس البيانات وتباعدها.

مفهوم الإحصاء اشتقَّ مصطلح الإحصاء باللغة الإنجليزية (Statistics) من الكلمة الإيطالية (Statista)، والكلمة الألمانية (Statistik)، والكلمة اللاتينية (Status)، والتي هي عبارة عن مصطلحات تُعنى بمعلومات الدولة (بالإنجليزية Political state):، حيث كانت بداية استخدام هذا المصطلح لجمع البيانات التي تخص أفراد الدولة، لغاية إنشاء قاعدة بيانات يتم من خلالها فرض الضرائب لتحسين الوضع المادي للدولة. [١] كما تم تعريف الإحصاء على أنه العلم الذي يهتم بجمع البيانات الرقمية، ومن ثم تنظيمها، وترتيبها، وتحليلها، بهدف الوصول إلى نتائج معينة لتوضيح ظاهرة أو حالة ما، أو بأنه العلم الذي يهتم بالطريقة التي يتم من خلالها جمع البيانات والمعلومات وتحويلها إلى صورة عددية، حيث تُجمَع البيانات من خلاله بشكل منتظم، وفيما يخص استخدامات علم الإحصاء فهي كثيرة؛ كاستخدامه في العلوم الطبية، وعلم الاجتماع، والاقتصاد، والصناعة، والكيمياء، والرياضة، والإدارة، وغيرها العديد من المجالات.

النظرة الشاملة-

□ الفئة المستهدفة

طلبة المرحلة الأولى / قسم تقنيات المحاسبة / المعهد التقني كربلاء

□ المبررات

صممت هذا الوحدة الفصل لتمكن الطالب من التعرف على مفهوم علم الاحصاء , استخداماته علاقته بالعلوم الأخرى, مراحل الطريقة الإحصائية (جمع البيانات.)

□ الفكرة المركزية

□ تعريف علم الاحصاء.

□ اقسام علم الاحصاء.

□ استخدامات علم الاحصاء.

□ علاقة علم الاحصاء بالعلوم الأخرى.

□ مراحل الطريقة الإحصائية في البحث العلمي.

□ مصادر جمع البيانات الإحصائية.

□ اساليب جمع البيانات الإحصائية.

□*الاسلوب الشامل.

□*الاسلوب النموذجي (العينة)

□ انواع العينات.

□ انواع البيانات الإحصائية.

طبيعة علم الاحصاء Nature Of Statistics .

كلمة (الإحصاء) في الماضي كانت تهدف الى العد والحصر حتى سمي الاحصاء بعلم العد The Science of canting) اما الاحصاء الآن فقد تطور كثيرا وخاصة في القرن العشرين واصبح علما مستقلا له أهميته كوسيلة واداة في البحث العلمي لجميع العلوم.

تعريف علم الاحصاء:-

لقد وضعت تعاريف عديدة لعلم الاحصاء تجمع على انه العلم الذي يختص بعمليات جمع البيانات الاحصائية وتصنيفها وتبويبها وعرضها وتحليلها بقصد التوصل الى استنتاجات يمكن الاعتماد عليها في اتخاذ القرارات المناسبة للظاهرة قيد الدراسة . ويمكن تقسيم علم الاحصاء الى قسمين رئيسيين هما:

1 الاحصاء الوصفي- .

2 الاحصاء الاستدلالي (الاستقرائي)

حيث يختص الاحصاء الوصفي بتنظيم وتلخيص وعرض البيانات الاحصائية في صورة جداول او رسوم بيانية او اشكال هندسية إضافة الى حساب بعض المقاييس الاحصائية مثل مقاييس النزعة المركزية ومقاييس التشتت . والغرض من الاحصاء الوصفي هو تقدير معالم المجتمع الاحصائي ووصفه تمهيدا للوصول الى استنتاجات عنه .

أما القسم الثاني فهو الاحصاء الاستنتاجي (الاستقرائي) فيختص بالتوصل الى استنتاجات عن المجتمع قيد الدراسة وذلك بالاعتماد على عينة مأخوذة من ذلك المجتمع وباستخدام العديد من الطرق والنظريات الاحصائية التي تعتمد على نظرية الاحتمالات واختبار الفرضيات والتقدير وغيرها .

اهمية علم الاحصاء وعلاقته بالعلوم الاخرى ومجالات تطبيقه

تكمن اهمية علم الاحصاء في كونه وسيلة) وليس غاية (تستخدم في مساعدة الباحثين في شتى مجالات المعرفة عن طريق تزويدهم بالادوات التي تساعدهم في تحليل المعطيات بشكل علمي واستخلاص النتائج ومن ثم اتخاذ القرارات المناسبة. ولقد اعتمد علم الاحصاء في بعض مراحل تطوره على العديد من النظريات الرياضية وخاصة نظرية الاحتمالات والتوزيعات الاحتمالية. ولقد تطور هذا العلم وتعددت استخداماته وتشعبت لتشمل كافة مجالات الحياة الاجتماعية والاقتصادية والزراعية والطبية والهندسية والامنبة وغيرها.

ونتيجة لكثرة استخدامات هذا العلم في المجالات اعلاه برزت تسميات مختلفة للإحصاء ومقرونة باسم اخر كالإحصاء الزراعي , والاحصاء الصناعي , والاحصاء الطبي , الاحصاء السكاني (الحيوي) , الاحصاء الاقتصادي... وغيرها من المسميات حيث ان الاسلوب واحد لكنها تختلف في مجال التطبيق.

ان تطبيق او استخدام الاسلوب الاحصائي في البحث العلمي مرهون بإمكانية التعبير عن الظاهرة المدروسة تعبيراً

(كميا) رقميا)

وفيما يلي المراحل الرئيسية للطريقة الإحصائية في البحث العلمي.

- 1- تحديد المشكلة او فرضية البحث , او الظاهرة المراد دراستها- .
- 2- جمع البيانات او المعلومات عن الظاهرة المدروسة- .
- 3- تصنيف البيانات وتبويبها وعرضها- .
- 4- حساب المؤشرات الإحصائية- .
- 5- تحليل معطيات الدراسة والتوصل الى النتائج- .
- 6- تفسير النتائج مع عملية اتخاذ القرار المناسب بشأن الظاهرة او الفرضية.

جمع البيانات الإحصائية :- Collection of data

تعرف عملية جمع البيانات بانها كيفية الحصول على البيانات المطلوبة من مصادرها المختلفة وهي تمثل نقطة الانطلاق لتصنيفها وتحليلها واستخلاص النتائج. حيث يتم الحصول على البيانات الإحصائية من مصدرين رئيسيين هما:

- 1- **المصادر غير المباشرة (التاريخية Historical sources)** وهي بيانات معدة مسبقا عن ظاهرة ما وباستطاعة الباحث الرجوع اليها واخذ المعلومات المطلوبة , دائرة الإحصاء العامة , الاحوال المدنية , الوثائق والسجلات , النشرات الإحصائية.
- 2- **المصادر الميدانية Field sources** ويقصد بها الحصول على المعلومات من مصادرها الاصلية وذلك عن طريق الاتصال بمفردات المجتمع مباشرة من خلال توجيه الاسئلة اما عن طريق المقابلة الشخصية , التلفون , المراسلة . مثل التعداد السكاني , الإحصاء الزراعي , والصناعي.

الاسبوع الثاني والثالث تصنيف وتبويب البيانات ، تكوين الجداول التكرارية البسيطة والمزدوجة طبيعة البيانات الاحصائية

عند جمع البيانات حول ظاهرة ما نرسم للظاهرة بالرمز y او x او اي رمز اخر وكل مفردة او مشاهدة ترمز لها y_i او x_i فمثلا عند دراسة اطوال الطلبة في قسم تقنيات المحاسبه فأنا نرسم لصفة الطول (الظاهرة y) ولطول (اي طالب) المفردة (بالرمز y_i . هذا وان قيمة y_i قد تختلف من طالب الى اخر لهذا نقول ان y متغير variables وعليه فان المتغير : هو اي ظاهرة تظهر اختلافات بين مفرداتها ويرمز له بالرمز y او اي رمز اخر . وتنقسم المتغيرات الى قسمين:

1- **متغيرات وصفية او نوعية** : هي الظواهر او الصفات التي لا يمكن قياسها مباشرة بالأرقام العددية مثل (صفة لون العين) ازرق ،اسود ،بني (او الحالة الاجتماعية) غني ،متوسط الحال ، فقير ، او النوع الاجتماعي (ذكر ،انثى).....الخ

2**متغيرات كمية** : هي الظواهر او الصفات التي يمكن قياسها مباشرة بأرقام عددية مثل صفة الطول ، الوزن ، العمر ، كمية المحصول . وتنقسم المتغيرات الكمية الى قسمين هما :

أ **متغيرات متصلة (مستمرة)**: المتغير المستمر هو المتغير الذي تأخذ المشاهدة او المفردة فيه اية قيمة رقمية في مدى معين ، فلو فرضنا ان اطوال طلبة المعهد تتراوح بين 145 سم و 190 سم اي ان المتغير y يمكن ان ياخذ اي قيمة بين 145 و 190 سم ومثال ذلك الوزن وكمية المحصول ودرجات الحرارة ، وبصورة عامة كل البيانات التي تقاس تعد بيانات لمتغير مستمر (تأخذ القيم عدد صحيح او كسر.)

ب **متغيرات غير مستمرة (منفصلة)** : تأخذ المشاهدة او المفردة فيها قيما متباعدة او منقطعة مثال ذلك عدد الثمار او عدد الوحدات الانتاجية او عدد الطلبة في السنة الأولى في قسم تقنيات المحاسبه ، فهي في الغالب تكون اعداد صحيحة .

المجتمع والعينة:

المجتمع : عبارة عن جميع القيم او المفردات التي يمكن ان يأخذها المتغير . او هو مجموعة متكاملة من الافراد او الاشياء او الاعداد او القياسات لها خاصية مشتركة يمكن ملاحظتها ويراد تحليلها , ويقسم المجتمع الى قسمين اما مجتمع محدود يمكن حصر جميع افراده , عدد الطلاب , عدد الموظفين , عدد الاسر .

أو مجتمع غير محدود والذي لا يمكن حصر افراده مثل ذرات الهواء , جزيئات الماء .
اما العينة : جزء من المجتمع ، وهي عبارة عن مجموعة من المشاهدات اختيرت بطريقة ما من المجتمع ففي بعض الاحيان دراسة المجتمع ككل قد يكون صعبا او يحتاج الى وقت وجهد ومال فيستعاض عن دراسة المجتمع بدراسة العينة وصفاتها ومنها يستنتج خواص المجتمع الاصيلي الذي اخذت منه العينة.

المعاينة : sampling - يقصد بها الطريقة التي يتم بموجبها اختيار مفردات العينة.

الحففيه التعليميه للإحصاء المعهد التقني كربلاء – قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
الوحدة الاحصائية : **unit** - هي اي عنصر او فرد في المجتمع الذي ندرسه.

المؤشر : **parameter** - وهو قيمة عددية تصف المجتمع مثل الوسط الحسابي او الانحراف المعياري للمجتمع.

المعلمة : **Statistics** - وهي قيمة عددية تصف العينة مثل الوسط الحسابي او الانحراف المعياري للعينة

مثال تطبيقي: كان في قسم المحاسبة في السنة الثانية 120 طالب يدرسون مادة التدقيق وكان عدد الناجحين 80
وعدد الراسبين 40 طالب وكان الناجحين تتراوح درجاتهم بين (50-95) والراسبين (10-49) .

م/ اجب عن الاتي:

1- هل المتغيرات وصفية ام كمية.

2- اين المجتمع ؟ هل ممكن اختيار عينة ؟ اختر عينة من الناجحين ؟ واختر عينة من الراسبين؟

اسلوب تصميم البحوث

هناك اعتبارات كثيرة يتوقف عليها تصميم البحث وعلى الباحث ان يأخذ بالاعتبار مسألة الحصول على البيانات
والمعلومات بأقصر وقت وأقل جهد واطىء كلفة وعليه يجب مراعاة الاتي عند تصميم البحث:

1 -تحديد الغرض من البحث :

2- الجانب العملي للبحث :

3تحديد اطار البحث :

اسلوب جمع البيانات والمعلومات : للوصول الى البيانات والمعلومات هناك اسلوبان يمكن من خلالهما جمع هذه
البيانات والمعلومات كل منهما له ميزاته وعيوبه وهذان الاسلوب هما :

1- اسلوب التسجيل الشامل : هو جمع البيانات من جميع المفردات التي يتكون منها المجتمع مجال البحث ومثال
ذلك التعداد العام للسكان من ميزات هذا الاسلوب يعطي بيانات كاملة حول الظواهر التي يتم البحث عنها اما عيوبه
فان هذا الاسلوب يحتاج الى وقت وجهد ومال كما لا يمكن استخدام هذا الاسلوب في المجتمعات غير المحددة.

2- اسلوب العينات : هو اخذ وحدات من المجتمع الاحصائي تسمى العينة sample والغرض من اخذ العينة ان
تكون بديلا عن المجتمع الاحصائي وعن طريق صفاتها يتمكن الباحث ان يصف خواص المجتمع بتعميم النتائج
التي حصل عليها من دراسة العينة . تفضل هذه الطريقة عن طريقة التسجيل الشامل للأسباب الاتية:

1- توفر المال والجهد والوقت اللازم لإجراء البحث.

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
2- صعوبة اجراء التسجيل الشامل بسبب طبيعة المجتمع فقد يكون المجتمع غير محدد او كبير جدا. ومن
عيوب هذا الاسلوب فان محاولة التعرف على خواص المجتمع عن طريق دراسة جزء منه ينطوي عليه
التضحية في دقة النتائج التي تستخرجها.

وتنقسم العينات الى قسمين رئيسيين هما العينات العشوائية والعينات الغير عشوائية

1العينات العشوائية : هي مجموعة المفردات المختارة من مجتمع الدراسة وليس للباحث دخل في اختيارها .
وللعينات العشوائية انواع عديدة منها

أ العينة العشوائية البسيطة : هي اختيار عينة عشوائية من مجتمع الدراسة بطريقة تعطي المفردات نفس الفرصة
في الظهور .ويشترط هنا ان يكون المجتمع متجانس) مشترك في الصفات (فمثلا دراسة اسباب عدم التحضير لدى
الاناث نلاحظ ان المجتمع متجانس حيث ان كافة مفردات هذا المجتمع هم من الاناث والصفة المشتركة هي عدم
التحضير اليومي .

ب العينة الطبقية العشوائية : يتم اختيار العينة عندما يكون المجتمع غير متجانس ،يقسم المجتمع الى طبقات كل
طبقة تعتبر مجتمع متجانس ومن كل مجتمع يتم اختيار عينة عشوائية بسيطة يتناسب حجمها مع حجم الطبقة ثم
تجمع هذه العينات ونحصل على الطبقة العشوائية.

مثلا لو كنا بصدد دراسة للمستوى العلمي لاحد طلبة المعهد التقني كربلاء هذا المجتمع غير متجانس من حيث
التخصص العلمي فهناك اختصاص صحة مجتمع واختصاص محاسبة واختصاص ومدني واختصاص سياحة
وهكذا.

ج العينة المتعددة المراحل : يتم تقسيم المجتمع الى وحدات اولية ثم يتم اختيار عينة عشوائية من هذه الوحدة
الاولية ثم تقسم كل وحدة من الوحدات الاولى الى وحدات ثانوية ثم تؤخذ عينة كمرحلة ثانية ثم تقسم الى وحدات
اصغر وتأخذ عينة منها الى ان نصل الى المفردة التي يتم جمع البيانات منها والتي تؤلف عينة البحث.

2- العينات غير العشوائية : يقصد بها مجموعة من المفردات المختارة من مجتمع الدراسة بطريقة يكون للباحث
دخل في اختيارها ومن هذه العينات.

أ المعاينة الحصصية : تقسيم مجتمع الدراسة الى طبقات استنادا الى معايير تقسيم معينة تتعلق بطبيعة الدراسة ثم
يتم اختيار عدد من المفردات من كل طبقة بشكل شخصي (غير عشوائي) بحيث ان عدد مفردات هذه العينات
يشكل حجم العينة المطلوبة لتلك الدراسة .فلو كنا بصدد استطلاع راي الجمهور في مواقع التواصل الاجتماعي
فانه يمكن تقسيم مجتمع الدراسة الى ذكور واناث ثم يتم اختيار عينة من الذكور واخرى من الاناث تتناسب كل
منهما مع عدد الذكور وعدد الاناث في مجتمع هذا الاستطلاع ومجموع مفردات هاتين العينتين تؤلفان حجم العينة
المطلوب للاستطلاع .

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
ب المعايينة العمديية : اختيار العينة بشكل متعمد يععتقد الباحث مسبقا بان مفردات هذه العينة هي خير من يمثل
مجتمع الدراسة.

مصادر جمع البيانات:

يتم الحصول على البيانات و المعلومات من احد المصدرين الآتيين:

1-المصادر التاريخية : هي البيانات المحفوظة لدى اجهزة الدولة المختلفة نتيجة الاستقصاءات او مسوحات قامت بها هذه الجهات لأغراض خاصة بها او تجمعت لديها بحكم وظائفها . مثال ذلك البيانات المتجمعة عن تعدادات السكان ،احصاءات الطلبة المتخرجين من الجامعات او احصاءات التجارة الداخلية والخارجية.

2مصادر الميدان : بيانات ومعلومات يمكن الحصول عليها من مصادر ها الاصلية بطريقة المراسلات (بالبريد) أو المواجهة (المقابلة الشخصية) أو عن طريق الهاتف أو اي وسيلة اتصال أخرى.

تصنيف وتبويب البيانات

لاحظنا ان عملية جمع البيانات تتم من خلال المصادر التاريخية او الميدانية باستخدام أسلوب التسجيل الشامل او اسلوب العينات حسب ما تتطلبه الدراسة ، ان البيانات المتحصل عليها بخصوص الظاهرة المعنية تسمى البيانات الاولية او البيانات غير المصنفة ،ان البيانات بشكلها الاولي تكون غير منظمة مما يتعذر على الباحث تكوين فكرة عن هذه الظاهرة او تلك التي جمعت البيانات ، كذلك يتعذر الاعتماد عليها بشكلها الغير المنظم لأغراض التحليل الاحصائي للوصول الى النتائج المطلوبة ، لذلك ان اولى الخطوات الهامة بعد عملية جمع البيانات هي **عملية تصنيف وتبويب البيانات.**

1- مراجعة البيانات : بعد اتمام عملية جمع البيانات وفق الوسيلة المناسبة لذلك البحث يتوجب الامر مراجعة

وتدقيق البيانات لغرض التأكد من مطابقتها وتكاملها لمتطلبات الدراسة.

2- تصنيف البيانات: بعد التأكد من دقة البيانات التي تم الحصول عليها يتم عملية تصنيف البيانات على

اساس الظواهر التي جمعت منها البيانات حيث يتم فرز بيانات كل ظاهرة على هيئة مجموعة فقد يكون التصنيف على ظاهرة العمر، الوزن، المهنة، الطول.

3- تبويب البيانات : بعد اتمام عملية تصنيف البيانات تبدأ عملية التبويب ، ويقصد بالتبويب عملية تفرغ

البيانات المصنفة في جداول خاصة بحيث ان كل جزء من البيانات المصنفة عن الظاهرة المعنية يعود الى مستوى معين لتلك الظاهرة ، الهدف من عملية التبويب هو ابراز البيانات وتوضيحها في أضيق حيز ممكن كي يتمكن الباحث من تكوين فكرة عنها ويختلف اسلوب تبويب البيانات تبعاً لطبيعتها . وفيما يلي عرض موجز لكل شكل من هذه الاشكال.

الحفبه التعليميه للإحصاء المعهد التقني كربلاء – قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
أ- التبويب الزمني : عبارة عن تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل مجموعة منها تعود

لوحة زمنية كاليوم ، الاسبوع ، الشهر ، السنة. والجدول الاتي يوضح عدد الطلبة الخريجين لعدد من السنوات



السنوات	عدد الخريجين
2018	800
2019	700
2020	500
2021	500
المجموع	2500

ب التبويب الجغرافي : تقسيم البيانات الى مجموعات كل منها خاص بوحدة جغرافية معينة او تقسيم اداري معين كالنواحي والاقضية والمحافظات والبلدان، الفارات ، عدد السكان حسب المحافظات العراقية.

اسم المحافظة	عدد السكان
بغداد	8000000
الموصل	7000000
كربلاء المقدسة	3000000
المجموع	18000000

ج التبويب الكمي : تقسيم البيانات الى مجموعات خاصة بوحدة معينة كوحدة الوزن والطول ، المساحة ، الحجم الخ ، المثال الاتي يوضح توزيع الاجور اليومية لعمال مصنع الخليج

الاجرة اليومية بالدينار	عدد العمال
اقل من 30000 دينار	185
اقل من 35000	95
اقل من 50000	70
من 50000 فاكثر	20
المجموع	370

د التبويب على اساس صفة معينة : تجميع البيانات وترتيبها في جداول على مجموعة منها يشترك بصفة معينة كالجنس ، الحالة الاجتماعية ، عنوان الوظيفة والمثال الاتي يوضح عدد المحاسبين في شركة امارون حسب نوعهم الاجتماعي.

النوع الاجتماعي	عدد المحاسبين
ذكور	17
الاناث	23
المجموع	40

الأسبوع الرابع والخامس : العرض البياني للبيانات (التوزيع التكراري) :-

النظرة العامة:-

-الفئة المستهدفة: طلبة المرحلة الاولى / قسم تقنيات المحاسبة/ المعهد التقني كربلاء

- المبررات :

صم هذا الفصل لتمكين الطالب من التعرف على مفهوم : تصنيف وتبويب البيانات وكيفية انشاء الجدول التكراري البسيط و المزدوج والجدول المتجمع الصاعد والنازل.

* الفكرة المركزية

- تعريف مفهوم التصنيف , والتبويب.
- انواع التصنيف.
- جدول التوزيع التكراري البسيط.
- جدول التوزيع التكراري المزدوج.
- جدول التوزيع المتجمع الصاعد والنازل.
- الجدول التكراري النسبي.
- الجداول المغلقة والمفتوحة.
- الجداول المنتظمة وغير المنتظمة

الاهداف الادائية- :

سيكون الطالب بعد انتهائه من دراسة هذه الوحدة النمطية قادرا على

- يعرف كل من تصنيف وتبويب البيانات.
- يميز بين انواع التبويب المختلفة.
- ينظم البيانات ويلخصها مستخدما الجدول التكراري البسيط.
- يلخص ويبوب البيانات مستخدما الجدول التكراري المزدوج.
- يميز بين الجدول المفتوح والمغلق.
- يميز بين الجدول المنتظم وغير المنتظم.

أ – المدرج التكراري . ب – المضع التكراري . ج – المنحني التكراري .

د – المنحني التكراري للتجمع الصاعد والنازل .

التوزيع التكراري Frequency Distribution

عبارة عن تلخيص وترتيب البيانات التي سبق ان جمعت وصنفت مقسمة الى عدد من المجاميع كل منها تسمى الفئة (class) هذه الفئات قد تكون مرتبة تصاعديا او تنازليا حسب طبيعة البيانات ويسمى توزيع عدد قيم x حسب الفئات بالتوزيع التكراري . وقد تكون فئات التوزيع التكراري متساوية في الطول ام غير متساوية وذلك يعتمد على طبيعة الدراسة ومتطلباتها. والاتي توضيح لبعض المصطلحات

البيانات غير المبوبة : هي البيانات الاولية التي جمعت ولم تبوب في جدول توزيع تكراري.

البيانات المبوبة : هي البيانات التي جمعت وبوبت ونظمت في جدول توزيع تكراري.

التوزيع التكراري : تقسيم البيانات او القيم الخاصة بظاهرة من الظواهر الاحصائية الى أصناف او فئات يطلق عليها بالتوزيع التكراري.

الفئة : هي المجاميع التي قسمت اليها قيم المتغير ، وكل فئة لها حدان ، حد ادنى ، وحد اعلى.

طول الفئة : هو مقدار المدى بين حدي الفئة.

المدى : هو الفرق بين اكبر قيمة واصغر قيمة + 1.

مركز الفئة : هي القيمة الواقعة عند منتصف الفئة.

تكرار الفئة : عدد المفردات او القيم التي تقع في مدى تلك الفئة ويرمز له ب f_i هذا وان مجموع التكرارات يجب ان يكون دائماً مساوي للعدد الكلي لقيم الظاهرة.

طول الفئة يرمز له ب L ويستخرج طول الفئة باستخدام احد القوانين الآتية:

$$L = x_L - x_S + 1$$

حيث ان L : طول الفئة :

x_L : الحد الاعلى للفئة :

x_S : الحد الادنى للفئة :

المدى

$$\text{او طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات}}$$

عدد الفئات

او طول الفئة = الفرق بين الحدين الادنى (او الحدين الاعلى (لفئتين متتاليتين)

او طول الفئة : الفرق بين مركزي فئتين متتاليتين

المدى : يرمز له $T.R$

$$T.R = L \cdot X - X_S + 1$$

حيث ان L : اكبر قيمة

X_S : اصغر قيمة

مركز الفئة يرمز له X

$$L.L + U.L$$

$$X = \frac{\text{-----}}{2}$$

2

حيث ان : الحد الاعلى للفئة $L.L$

الحد الادنى للفئة $U.L$

عدد الفئات يرمز له ب m , هناك عدة طرق تقريبية لإيجاد عدد الفئات اهمها:

$$m = 1 + 3.322 \log n$$

$$m = 2.5 \sqrt[4]{n}$$

هناك نوعان رئيسيان من الجداول الإحصائية:

1- الجدول البسيط : هو الجدول الذي توزع فيه البيانات حسب صفة واحدة ويتألف عادة من عمودين . الاول يمثل

تقسيمات صفة الظاهرة الى فئات او مجموعات والثاني بين عدد المفردات التابعة كل فئة او مجموعة والجدول

الآتي يمثل عدد من الطلبة حسب اوزانهم

فئات الوزن	عدد الطلبة
60-62	5
63-65	15
66-68	45
69-71	27
72-74	8
المجموع	100

2- الجدول المركب (المزدوج) : هو الجدول الذي توزع فيه البيانات حسب صفتين او ظاهرتين او اكثر في نفس الوقت و يتألف من:

الصفوف : تمثل فئات او مجاميع احدى الصفتين.

الاعمدة : تمثل فئات او مجاميع الصفة الاخرى.

والجدول التالي يبين توزيع عدد من الطلبة حسب صفتي الطول والوزن

المجموع	الوزن			الطول
	71-80	61-70	51-60	
30	4	6	20	121-140
52	10	40	2	141-160
18	10	6	2	161-180
100	24	52	24	المجموع

مثال (1) : لو اردنا عمل توزيع تكراري للأعداد الآتية التي تمثل الوزن بالكيلو غرامات لعشرين طالبا في المعهد التقني كربلاء (67,55,65,70,75,60,89,83,65,56,49,65,49,48,69,62,72,45,56,74,) نتبع الخطوات الآتية :

1- نبحث عن اكبر قيمة واصغر قيمة في المجموعة وهي (89-45) وذلك للتوصل الى المدى الكلي

2- نجد عدد الفئات. 3- نجد طول الفئة. 4- نكتب حدود الفئات ونستخرج عدد التكرارات لكل فئة

المدى الكلي = اكبر قيمة - اقل قيمة + 1

$$T.R = LX - XS + 1$$

$$= 89 - 45 + 1$$

$$= 45$$

$$عدد الفئات = M = 1 + 3.322 \log n$$

$$= 1 + 3.322 \log 20$$

نقربها الى 5.32

طول الفئة

$$T.R$$

$$L = \frac{T.R}{M}$$

$$= \frac{45}{5.32}$$

$$= 8.46$$

$$L = 9$$

كتابة حدود الفئات واستخراج عدد التكرارات لكل فئة

class	fi	fi
40-49		4
50-59		3
60-69		7
70-79		4
80-89		2
Total	20	20

ملاحظة : هناك عدة طرق لكتابة حدود الفئات:

- 1- اما ان تكون الاعداد لمتغيرات منفصلة كما في المثال السابق
- 2- او تكون الاعداد لمتغيرات متصلة وهو الذي يمثل بعدد صحيح او كسر مثل الاوزان والاطوال وتكتب

الفئات كالآتي:

من 40 الى اقل من 50

من 50 الى اقل من 60

من 60 الى اقل من 70

وللاختصار تكتب بالصيغة الآتية:

40-

50-

60-

70 -

تمتاز هذه الطريقة بالوضوح وتستخدم غالبا لأعداد التي تمثل متغيرات متصلة

3- وقد تكتب الفئات حسب الصيغة الآتية:

اكبر من 40 واقل من 50

اكبر من 50 واقل من 60

اكبر من 60 واقل من 70

وللاختصار تكتب

50-

60-

70-

قد يكون التوزيع في الجدول التكراري البسيط توزيعا منتظما كما في المثال السابق وذلك لتساوي طول الفئة ، او

يكون التوزيع غير منتظم اذا كان طول الفئة غير متساوي ، او يكون الجدول مغلقا اذا كان الحد الأدنى والحد

الاعلى للفئة معروف ، او يكون الجدول مفتوحا في الحالات الآتية:

أ- يكون مفتوحا من الطرف الأدنى فقط

ب- يكون مفتوحا من الطرف الاعلى فقط

ت- يكون مفتوحا من الطرفين (اذا كان الحد الأدنى والحد الاعلى للفئة غير معلوم)

تمرين واجب (1) البيانات الاتية تمثل عدد الوحدات المنتجة من (50) عاملا في احد الشركات الصناعية

57 43 46 24 44 38 19 54 49 57

29 53 45 47 41 37 49 56 47 29

31 32 51 52 42 45 28 43 49 34

24 42 28 39 21 18 37 34 29 23

35 43 39 41 26 27 32 37 28 33

المطلوب انشاء جدول توزيع تكراري مستخدما (6 فئات)

مثال (2) : البيانات الاتية لظاهرتين x و y المطلوب تفرغها في جدول تكراري مزدوج

X: 2 , 10 , 11 , 4 , 20 , 15 , 15 , 3

Y: 3 , 2 , 5 , 6 , 8 , 10 , 2 , 10 X 25 , 25 , 20 , 22 , 15 , 20 , 30 , 30 ,

Y: 2 , 6 , 5 , 9 , 15 , 12 , 3 , 9

X: 35 , 30 , 35 , 31

Y: 10 , 12 , 11 , 4

الحل : نستخرج معلومات لكل ظاهرة على حدى المتغير X

$$T.R = LX - XS + 1 \text{ : المدى}$$

$$= 35 - 2 + 1$$

$$= 34$$

$$m = 2.5^4 \sqrt{n} \text{ : عدد الفئات}$$

$$= 2.5^4 \sqrt{20}$$

$$= 2.5 * 2.114$$

$$= 5.28$$

وبالتقريب عدد الفئات = 5
طول الفئة:

$$T.R \quad 34$$

$$L = \frac{\quad}{M} = \frac{\quad}{5} = 6.8$$

$$M \quad 5$$

وبالتقريب = 7
المتغير Y

$$\text{المدى} : T.R = (LX - XS) + 1$$

$$= 15 - 2 + 1 = 14 \quad \text{عدد الفئات} : \text{Log}20 * 3.322 + 1$$

$$= 1 + 3.322 * 1.3 = 5.31$$

وبالتقريب = 5
طول الفئة:

$$T.R \quad 14$$

$$L = \frac{\quad}{M} = \frac{\quad}{5} = 2.8 \quad \text{بالتقريب} = 3$$

$$M \quad 5$$

fX	fy	fiX	fi y
2-8	2-4	3	6
9-15	5-7	5	4
16-22	8-10	4	6
23-29	11-13	2	3
30-36	14-16	6	1
		20	20

بعد ذلك نضع المعلومات في جدول مزدوج وكالاتي:

X \ Y	2-8	9-15	16-22	23-29	30-36	TOTAL
2-4						6
5-7						4
8-10						6
11-13						3
14-16						1
TOTAL	3	5	4	2	6	20

ملاحظات مهمة حول التوزيع التكراري

- 1-يجوز ان نبدأ الفئة الاولى باصغر قيمة في التوزيع او اصغر منها- .
- 2- طول الفئة = الفرق بين مركز الفئة اللاحق ومركز الفئة السابق لان طول الفئات متساوية- .
- 3 -قد يعبر عن كتابة الفئات بمراكزها- .
- 4 -يجب ان تكون حدود الفئات محددة بشكل واضح بحيث ان كل مفردة تقع في فئة واحدة من التوزيع.

التوزيع التكراري المتجمع

التوزيع التكراري البسيط يعطينا عن عدد المفردات في كل فئة لكن في بعض الاحيان نرغب في معرفة عدد المفردات التي قيمتها أقل او أكثر من قيمة معينة في التوزيع التكراري. ويعرف التوزيع التكراري المتجمع: بأنه التوزيع الذي يبين كمية التكرار المتجمع عند قيمة معينة من قيم المتغير العشوائي. وهناك نوعان من الجداول التكرارية المتجمعة . ويرمز له ب F_i

أ- التوزيع التكراري المتجمع الصاعد : وهو عبارة عن تجميع التكرارات من الفئة الاولى وانتهاء بالفئة الاخيرة منه ويتم حساب التكرارات المتجمعة على اساس الحدود العليا للفئات فلو رجعنا الى المثال رقم (1) فان جدول التوزيع التكراري المتجمع الصاعد يكون كالاتي:
نجمع التكرارات من الفئة الدنيا الى الفئة العليا

الفئات	التكرارات f_i	الحدود العليا للفئات	المتجمع الصاعد F
40-49	4	اقل من 49	4
50-59	3	اقل من 59	7
60-69	7	اقل من 69	14
70-79	4	اقل من 79	18
80-89	2	اقل من 89	20
المجموع	20		

ب- جدول التوزيع المتجمع النازل : عبارة عن تجميع التكرارات ابتداء من الفئات العليا وانتهاء بالفئات الدنيا ،بعبارة اخرى تناقص التكرارات ابتداء بالفئة الاولى في التوزيع وانتهاء بالفئة الاخيرة منه . ويتم حساب التكرارات المتجمعة النازلة على اساس الحدود الدنيا للفئات.
فلو رجعنا الى المثال رقم (1) فان التوزيع التكراري المتجمع النازل يكون كالاتي:

المتجمع النازل F	الحدود الدنيا للفئات	التكرارات f _i	الفئات
20	40 فاكثر	4	40-49
16	50 فاكثر	3	50-59
13	60 فاكثر	7	60-69
6	70 فاكثر	4	70-79
2	80 فاكثر	2	80-89
		20	المجموع

تمرين واجب (2) في تجربة لقياس السرعات الاتية للمركبات على طريق خارجي اعطيت اليك البيانات كما في الجدول الاتي :

المطلوب عمل جدول توزيع متجمع صاعد و جدول توزيع متجمع نازل

الفئات	30-	40-	50-	60-	70-	80-	90-	100-	110-	120-	المجموع
	39	49	59	69	79	89	99	109	119	129	
التكرارات	3	6	24	64	50	29	14	6	3	1	200

العرض الهندسي للبيانات او العرض البياني للبيانات المبوبة :

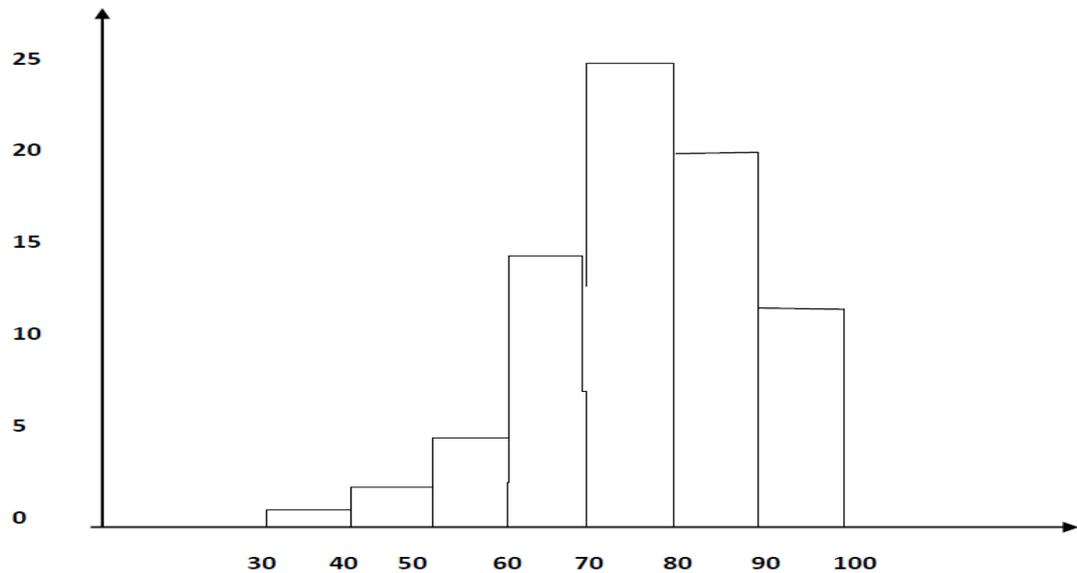
نواجه في الحياة العملية اعداد كبيرة من البيانات تتعلق بمختلف مجالات الحياة فاذا عرضت هذه البيانات بطريقة الجداول أو التقارير ستكون بلا شك مملة ويصعب استيعابها والمقارنة بين مفرداتها ، كما ان الرسوم والاشكال الهندسية ما هي الا تعبير وتوضيح للبيانات بطريقة جذابة وسهلة وفعالة تساعد القارئ على فهم واستيعاب الظاهرة ومقارنتها مع بعضها . ووسائل التمثيل البياني كثيرة ومتنوعة وسنكتفي بشرح العرض البياني للتوزيعات التكرارية فقط . وعادة نخصص المحور الافقي لتمثيل قيم او فئات المتغير بينما نخصص المحور العمودي لتمثل تكرارات هذا المتغير ويجب دائما ان يبدأ تدرج المحور العمودي من الصفر اما تدرج المحور الافقي فقد لا يبدأ بتدريجه من الصفر . ومن اشكال العرض البياني نذكر:

1 المدرج التكراري

عبارة عن مستطيلات رأسية تمتد قواعدها على المحور الافقي لتمثل اطوال الفئات بينما ارتفاعاتها تمثل تكرارات الفئات . ولرسم مدرج تكراري نتبع الخطوات الاتية

- 1- رسم المحور الافقي والمحور العمودي
 - 2- تدرج المحور الافقي الى اقسام متساوية بحيث يشمل جميع حدود الفئات ويفضل ترك مسافة صغيرة بين نقطة الصفر والحد الأدنى للفئة الاولى، ونقسم المحور العمودي الى اقسام متساوية بحيث تشمل اكبر التكرارات
 - 3- يرسم على كل فئة مستطيلا رأسيا تمثل قاعدته طول تلك الفئة وارتفاعه يمثل تكرارها
- مثال (3): الجدول الاتي يمثل توزيع تكراري لأطوال نباتات القطن /
المطلوب : تمثيل التوزيع التكراري بمدرج تكراري

الفئات	30-40	40-50	50-60	60-70	70-80	80-90	90-100	المجموع
التكرارات	1	2	5	15	25	20	12	80



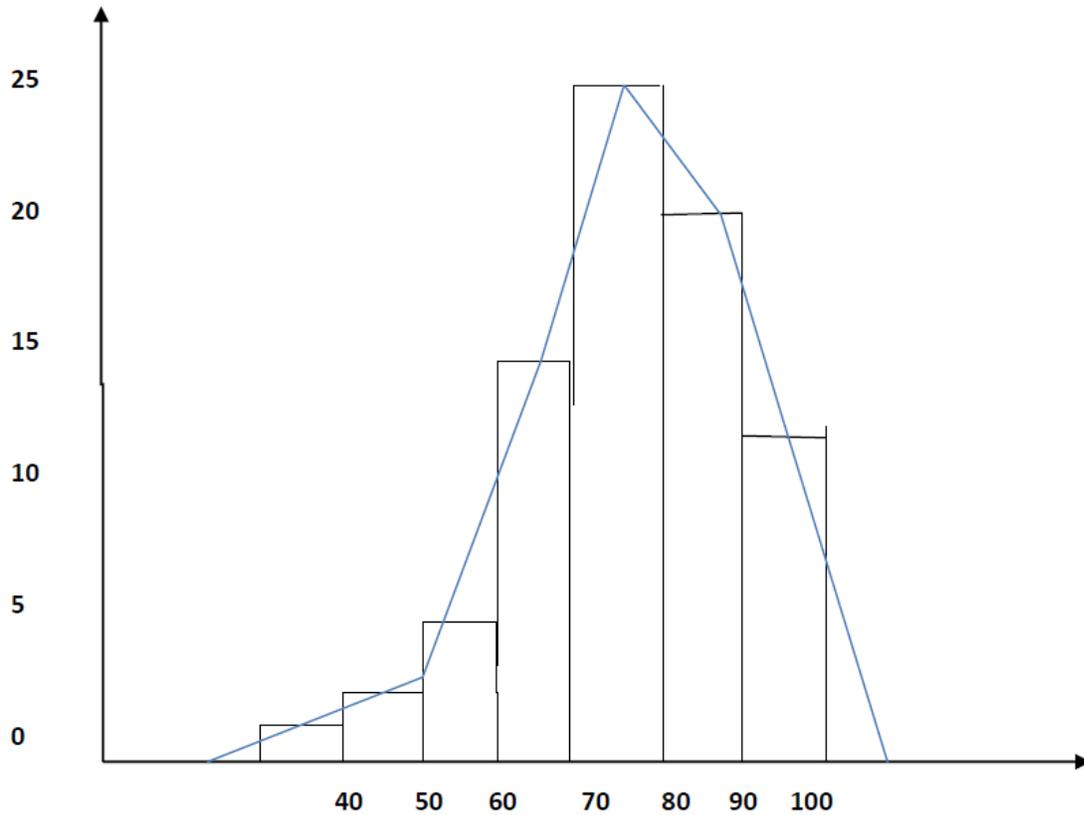
الحففيه التعليميه للإحصاء المعهد التقني كربلاء – قسم تفتيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
تمرين واجب (3) الاتي توزيع تكراري لأوزان عينة من طلبة احدى الكليات قوامها مئة طالب
المطلوب : رسم مدرج تكراري لهذا التوزيع

الفئات	46-53	53-60	60-67	67-74	74-81	81-88	88-95	95-102	المجموع
التكرارات	7	15	27	21	14	8	5	3	100

المضلع التكراري: هو وسيلة اخرى لتمثيل التوزيع التكراري بيانيا ويمكن رسمه بإحدى طريقتين:

اولهما: اذا كان المدرج التكراري معلوم . ويتم ذلك بتصنيف القواعد العليا لمستطيلات المدرج ثم نصل بين هذه النقط بمستقيمات ورسم فئة قبل الاولى تكرارها صفر وفئة بعد الاخيرة تكرارها صفر وتصنيف هاتين الفئتين وتوصيل بقية الخط نحصل على ما يسمى بالمضلع التكراري فلو عدنا الى المثال السابق(1) فالمضلع التكراري يكون بالشكل التالي:

الفئات	30-40	40-50	50-60	60-70	70-80	80-90	90-100	المجموع
التكرارات	1	2	5	15	25	20	12	80



الطريقة الثانية: رسم المضلع التكراري على مراكز الفئات مباشرة دون ضرورة لرسم المدرج التكراري اولا وعليه فان المحور الافقي يمثل مراكز الفئات والمحور العمودي يمثل التكرارات ثم نصل النقاط ببعضها ببعض وعليه فان خطوات رسم المضلع التكراري كما يأتي

- 1- ايجاد مراكز الفئات على المحور الافقي.
- 2- تحديد النقطة التي تقابل مركز كل فئة على المحور الرأسي.
- 3 وصل مستقيمات بين النقط التي حددناها بعضها ببعض.

وعليه فان المضلع التكراري يكون بالشكل الاتي:

مثال (4) : البيانات الآتية تمثل درجات 30 طالبا في الامتحان الفصلي . المطلوب رسم المضع التكراري

مراكز الفئات	الحدود الحقيقية	عدد الطلاب	فئات الدرجات
10	8.5 – 11.5	3	9 – 11
13	11.5 - 14.5	4	12 – 14
16	14.5 – 17.5	5	15 – 17
19	17.5 – 20.5	8	18 - 20

- ارسم المضع عن طريق مراكز الفئات

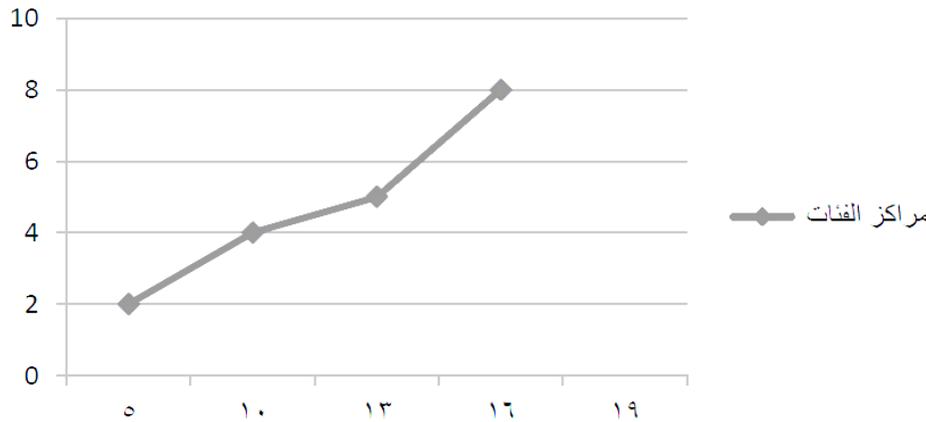
نقوم باتباع الخطوات الآتية

- نرسم محورين الافقي يمثل مراكز الفئات والعمودي يمثل التكرارات

- نجد مراكز الفئات

- . نعين النقاط على الرسم البياني حيث ان كل نقطة تمثل نقطة تقاطع مراكز الفئات مع التكرارات المقابلة للحصول على مضع تكراري مغلق نأخذ مركز فئة سابق بتكرار صفر ومركز فئة لاحق بتكرار صفر أيضا

مراكز الفئات



المنحنى التكراري Frequency curve

لرسم المنحنى التكراري نتبع نفس الخطوات التي اتبعناها في رسم المضع التكراري ولكن الفرق بينهما ان التوصيل بين النقاط يكون بخطوط منحنية وليس بخطوط مستقيمة كما هو الحال في المضع.

مثال (5): البيانات الآتية تمثل اوزان 30 موظفا المطلوب عرض البيانات ادناه مستخدما

أ المضع التكراري (الطريقتين-) ب المنحنى التكراري

فئات الاوزان	55-	60-	65-	70-
التكرار	6	8	12	4

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
تمثيل الجداول التكرارية المتجمعة بيانيا:

1 المنحني التكراري المتجمع الصاعد . (ك.م.ص)

2 المنحني التكراري المتجمع النازل . (ك. م. ن)

خطوات رسم المنحني المتجمع الصاعد:

1 ننشأ الجدول المتجمع الصاعد- .

2 رسم المحورين السيني والصادي , السيني يمثل الحدود العليا للفئات والصادي يمثل التكرار المتجمع الصاعد
(ك. م. ص.) .

3 نرصد النقاط على الرسم البياني والتي تمثل نقاط تقاطع الحدود العليا والتكرار المتجمع الصاعد- .

4 نصل بين النقاط اعلاه بخط منحني مرن بدلا من الخط المستقيم , نحصل على المنحني المتجمع الصاعد- .
خطوات رسم المنحني المتجمع النازل:

1 ننشأ الجدول المتجمع النازل- .

2 رسم المحورين السيني والصادي , السيني يمثل الحدود الدنيا للفئات والصادي يمثل التكرار المتجمع النازل-
(ك. م. ص.) .

- نرصد النقاط على الرسم البياني والتي تمثل نقاط تقاطع الحدود الدنيا والتكرار المتجمع النازل

3 . نصل بين النقاط اعلاه بخط منحني مرن , نحصل على المنحني المتجمع النازل 4

مثال (6): اعرض بيانات الجدول ادناه مستخدما المنحني المتجمع الصاعد ب المنحني المتجمع النازل

فئات الريح	15-	20-	25-	30- 35
التكرار	4	6	15	5

فئات الريح	عدد المساهمين f	الحدود العليا للفئات	ك . م . ص	الحدود الدنيا للفئات	ك . م . ن
15-	4	اقل من 20	4	اكثر من 15	30
20-	6	اقل من 25	10	اكثر من 20	24
25-	15	اقل من 30	25	اكثر من 25	9
30-35	5	اقل من 35	30	اكثر من 30	4

الدائرة البيانية *pie chart* :

تعد هذه الطريقة من افضل الطرق لتمثيل البيانات في حالة المقارنة بين الظاهرة الكلية واجزائها ويتم ذلك باتباع الخطوات الاتية:

1 -نستخرج زاوية القطاع من العلاقة الاتية:

$$\text{زاوية القطاع} = \frac{\text{قيمة الجزء المحدد}}{\text{المجموع الكلي للاجزاء}} \times 360^0$$

- تمثل الزاوية المركزية للدائرة باعتبار ان 360

- نقوم برسم دائرة معينة ونرسم عليها نصف قطر 2

- نرسم الزاوية المركزية التي ضلعها الابتدائي نصف القطر والممثلة للقطاع 3

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
مثال (7) : في جرد المكتبة المركزية بلغ عدد الكتب التي تم جردها ب(1000) بالألاف موزعة في اربعة ايام.

اليوم الاول 250 كتاب اليوم الثالث 220 كتاب

اليوم الثاني 300 كتاب اليوم الرابع 230 كتاب

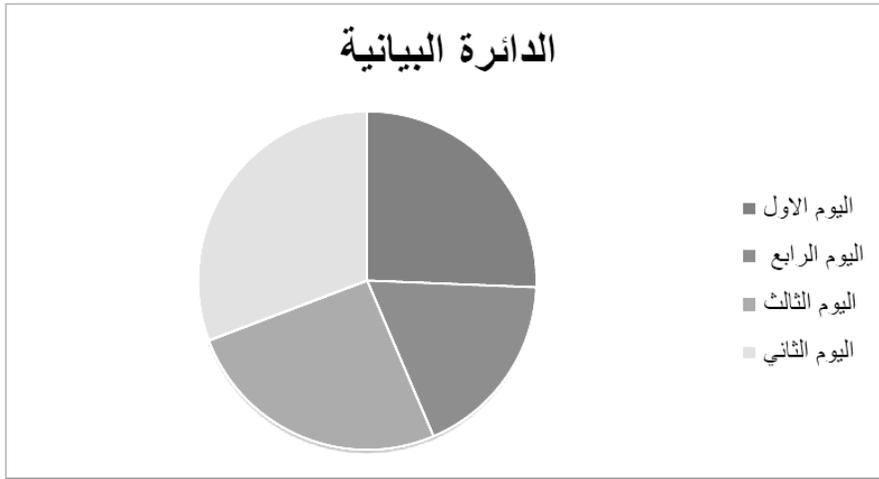
المطلوب / اعرض البيانات اعلاه مستخدما الدائرة البيانية

$$\text{الحل : زاوية قطاع اليوم الاول} = 360^\circ \times \frac{250}{1000} = 90^\circ$$

$$\text{زاوية قطاع اليوم الثاني} = 360^\circ \times \frac{300}{1000} = 108^\circ$$

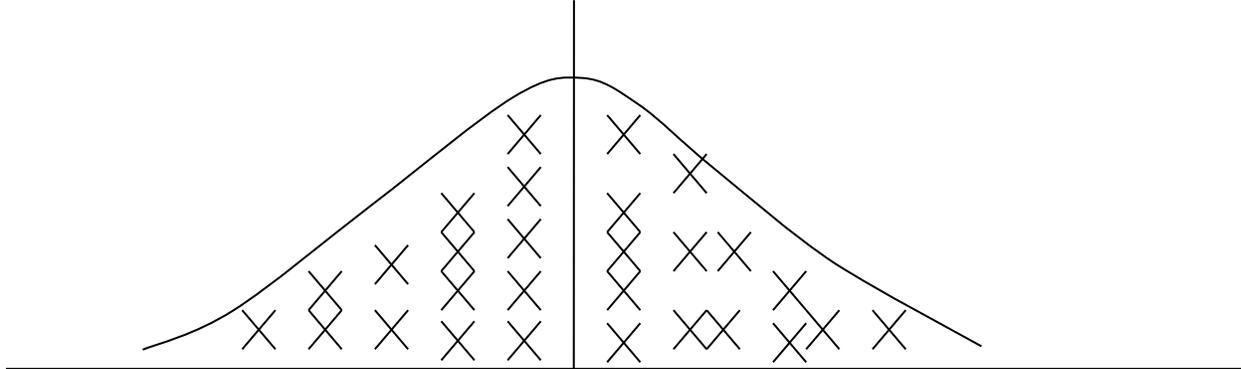
$$\text{زاوية قطاع اليوم الثالث} = 360^\circ \times \frac{220}{1000} = 79.2^\circ$$

$$\text{زاوية قطاع اليوم الرابع} = 360^\circ \times \frac{230}{1000} = 82.8^\circ$$



مقاييس النزعة المركزية :

معنى النزعة المركزية : هي ميل البيانات للتجمع حول المركز كما في الشكل الاتي:
 النزعة المركزية في حد ذاتها ظاهرة



في كثير من النواحي التطبيقية يكون الباحث في حاجة إلى حساب بعض المؤشرات التي يمكن الاعتماد عليها في وصف الظاهرة من حيث القيمة التي تتوسط القيم أو تنزع إليها القيم ، ومن حيث التعرف على مدى تجانس القيم التي يأخذها المتغير ، وأيضا ما إذا كان هناك قيم شاذة أم لا . كما انه في المحاضرات السابقة تعلمنا كيفية تمثيل البيانات بجداول ورسوم بغية تلخيصها وتوضيحها كذلك يمكن تمثيل البيانات بقيمة واحدة هي الوسط او المتوسط اي ان هذه البيانات تميل ان تقع في مركز البيانات المرتبة حسب الكبر لذلك تسمى مقاييس النزعة المركزية . والاوساط الاحصائية هي من اهم المقاييس الاحصائية الوصفية واكثرها استعمالا لدراسة البيانات ومقارنتها ، وهناك عدة انواع من المتوسطات واكثرها شيوعا واستعمالا هي:

1. الوسط الحسابي Arithmetic mean

2. الوسيط Median

3. المنوال Mode

أ- الوسط الحسابي: **Arithmetic Mean**: من أهم مقاييس النزعة المركزية ، وأكثرها استخداما في النواحي التطبيقية ، ويمكن حسابه للبيانات المبوبة وغير المبوبة ، كما يلي : للبيانات غير المبوبة: بشكل عام على أنه مجموع القيم مقسوما على عددها . فإذا كان لدينا n من القيم ، ويرمز لها بالرمز : x_1, x_2, \dots, x_n .

هناك طريقتين لحساب الوسط الحسابي

اولا : الطريقة المباشرة : الوسط الحسابي بموجب هذه الطريقة يمثل مجموع قياسات مفردات العينة مقسوما على عددها . فإن الوسط الحسابي لهذه القيم ، ونرمز له بالرمز \bar{x} يحسب بالمعادلة الاتية :

$$\text{الوسط الحسابي} = \frac{\text{مجموع القيم}}{\text{عدد القيم}}$$

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

حيث يدل الرمز Σ على المجموع .

مثال (8) : البيانات الآتية تمثل عينة من درجات الحرارة لمدة 10 أيام.

م/ إيجاد متوسط درجة الحرارة في هذه العينة (متغيرات مستمرة)

45 ، 49 ، 48 ، 42 ، 50 ، 51 ، 49.5 ، 50.5 ، 47.3 ، 48.2

/ الحل

$$\sum_{i=1}^n X_i$$

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

$$= \frac{45+49+48+42+50+51+49.5+50.5+47.3+48.2}{10}$$

$$= 48.05$$

لو قمنا بترتيب هذه البيانات تصاعديا لحصلنا على السلسلة الآتية:

42 ، 45 ، 47.3 ، 48 ، 48.2 ، 49 ، 49.5 ، 50 ، 50.5 ، 51

نلاحظ تمركز قيمة \bar{X} وسط هذه المجموعة هذا ما نقصده بمقاييس النزعة المركزية.

مثال (9) الآتي درجات 8 طلاب في مادة مبادئ المحاسبة .

34 32 42 37 35 40 36 40

المطلوب: إيجاد الوسط الحسابي لدرجة الطالب في الامتحان .

الحل لإيجاد الوسط الحسابي للدرجات تطبق المعادلة كما يلي:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

$$= \frac{34 + 32 + 42 + 37 + 35 + 40 + 36 + 40}{8} = \frac{296}{8} = 37$$

أي أن الوسط الحسابي لدرجة الطالب في امتحان مادة مبادئ المحاسبة يساوي 37 درجة

ثانياً: الوسط الحسابي للبيانات المبوبة

من المعلوم أن القيم الأصلية ، لا يمكن معرفتها من جدول التوزيع التكراري ، حيث أن هذه القيم موضوعة

في شكل فئات ، ولذا يتم التعبير عن كل قيمة من القيم التي تقع داخل حدود الفئة بمركز هذه الفئة ، ومن ثم يؤخذ

في الاعتبار أن مركز الفئة هو القيمة التقديرية لكل مفردة تقع في هذه الفئة.

فإذا كانت k هي عدد الفئات ، وكانت x_1, x_2, \dots, x_k هي مراكز هذه الفئات، وكانت f_1, f_2, \dots, f_k هي التكرارات ،

فإن الوسط الحسابي يحسب بالمعادلة التالية:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_k f_k}{f_1 + f_2 + \dots + f_k} = \frac{\sum_{i=1}^k x_i f_i}{\sum_{i=1}^k f_i}$$

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
مثال (10) : الجدول الاتي يعرض توزيع 40 تلميذ حسب أوزانهم .

فئات الوزن	32-34	34-36	36-38	38-40	40-42	42-44
عدد التلاميذ	4	7	13	10	5	1

المطلوب: إيجاد الوسط الحسابي.

الحل

لحساب الوسط الحسابي باستخدام المعادلة أعلاه يتم إتباع الخطوات الاتي :

- 1- إيجاد مجموع التكرارات $\sum f$.
- 2- حساب مراكز الفئات x .
- 3- ضرب مركز الفئة في التكرار المناظر له (Xf) ، وحساب المجموع $\sum xf$
- 4- حساب الوسط الحسابي بتطبيق المعادلة .

فئات الوزن (C)	التكرارات f	مراكز الفئات x	Xf
32-34	4	$(32+34) \div 2 = 33$	$4 \times 33 = 132$
34-36	7	35	$7 \times 35 = 245$
36-38	13	37	$13 \times 37 = 481$
38-40	10	39	$10 \times 39 = 390$
40-42	5	41	$5 \times 41 = 205$
42-44	1	43	$1 \times 43 = 43$
المجموع	40		1496

إذا الوسط الحسابي لوزن التلميذ هو :

$$\bar{x} = \frac{\sum_{i=1}^6 x_i f_i}{\sum_{i=1}^6 f_i} = \frac{1496}{40} = 37.4 \text{ k. g}$$

أي أن متوسط وزن التلميذ يساوي **37.4 kg**

خصائص الوسط الحسابي

يتصف الوسط الحسابي بعدد من الخصائص ، ومن هذه الخصائص ما يلي :

- 1- الوسط الحسابي للمقدار الثابت يساوى الثابت نفسه ، أي أنه إذا كانت قيم x هي : $x: a, a, \dots, a$ ، فإن الوسط الحسابي هو :

$$\bar{x} = \frac{a + a + \dots + a}{n} = \frac{na}{n} = a \quad (3-3)$$

ومثال على ذلك ، لو اخترنا مجموعة من 5 طلاب ، ووجدنا أن كل طالب وزنه 63 كيلوجرام ، فإن متوسط وزن الطالب في هذه المجموعة هو :

$$\bar{x} = \frac{63 + 63 + 63 + 63 + 63}{5} = \frac{315}{5} = 63 \text{ k. g}$$

- 2- مجموع انحرافات القيم عن وسطها الحسابي يساوى صفراً ، ويعبر عن هذه الخاصية بالمعادلة .

$$\sum (x - \bar{x}) = 0 \quad (4-3)$$

ويمكن التحقق من هذه الخاصية باستخدام بيانات مثال (10) ، نجد أن درجات الطلاب هي : 34, 32, 42, 40, 36, 35, 37 ، والوسط الحسابي للدرجة هو $\bar{x} = 37$ ، إذا :

x	34	32	42	37	35	40	36	40	296
$(x - \bar{x})$	34-37	32-37	42-37	37-37	35-37	40-37	36-37	40-37	0
$(x - 37)$	-3	-5	5	0	-2	3	-1	3	

أي أن : $\sum(x - 37) = 0$

3- إذا أضيف مقدار ثابت إلى كل قيمة من القيم ، فإن الوسط الحسابي للقيم المعدلة (بعد الإضافة) يساوى الوسط الحسابي للقيم الأصلية (قبل الإضافة) مضافاً إليها هذا المقدار الثابت . فإذا كانت القيم هي : x_1, x_2, \dots, x_n ، وتم إضافة مقدار ثابت (a) إلى كل قيمة من القيم ، ونرمز للقيم الجديدة بالرمز y ، أي أن $y = x + a$ ، فإن : الوسط الحسابي لقيم y (القيم بعد الإضافة) هو :

$$\bar{y} = \bar{x} + a$$

(٥-٣)

حيث أن \bar{y} هو الوسط الحسابي للقيم الجديدة ، ويمكن التحقق من هذه الخاصية باستخدام بيانات مثال رقم (10) : إذا قرر المصحح إضافة 5 درجات لكل طالب ، فإن الوسط الحسابي للدرجات المعدلة يصبح قيمته $\{(37+5)=42\}$ ، والجدول الآتي يبين ذلك .

x	34	32	42	37	35	40	36	40	296
y	34+5	32+5	42+5	37+5	35+5	40+5	36+5	40+5	
$= (x + 5)$	39	37	47	42	40	45	41	45	336

نجد أن مجموع القيم الجديدة هو : $\sum y = 336$ ، ومن ثم يكون الوسط الحسابي للقيم الجديدة هو

$$\bar{y} = \frac{\sum y}{n} = \frac{336}{8} = 42 \rightarrow (\bar{x} + 5 = 37 + 5 = 42)$$

4- إذا ضرب مقدار ثابت (a) في كل قيمة من القيم ، فإن الوسط الحسابي للقيم المعدلة (القيم الناتجة بعد الضرب) يساوى الوسط الحسابي للقيم الأصلية (القيم بعد التعديل) مضروباً في هذا المقدار الثابت . أي أنه إذا كان : $y = ax$ ، ويكون الوسط الحسابي للقيم الجديدة y هو :

$$\bar{y} = a \bar{x}$$

(٦-٣)

ويمكن للطالب أن يتحقق من هذه الخاصية باستخدام نفس بيانات المثال السابق . فإذا كان تصحيح الدرجة من 50 ، وقرر المصحح أن يجعل التصحيح من 100 درجة ، بمعنى أنه سوف يضرب كل درجة في قيمة ثابتة

($a=2$) ، ويصبح الوسط الحسابي الجديد هو : $\bar{y} = a \bar{x} = 2(37) = 74$:

5- مجموع مربعات انحرافات القيم عن وسطها الحسابي أقل ما يمكن ، أي أن :

$$\sum (x - \bar{x})^2 < \sum (x - a)^2 \text{ if } a \neq \bar{x}$$

(٧-٣)

وفي المثال السابق فإن : $\sum (x - 37)^2 < \sum (x - a)^2$ لجميع قيم $a \neq 37$

ثالثاً: الوسط الحسابي المرجح

في بعض الأحيان يكون لكل قيمة من قيم المتغير أهمية نسبية تسمى أوزن ، أو ترجيحات ، وعدم أخذ هذه الأوزان في الاعتبار عند حساب الوسط الحسابي ، تكون القيمة المعبرة عن الوسط الحسابي غير دقيقة ، فمثلاً لو أخذنا خمسة طلاب ، وسجلنا درجات هؤلاء الطلاب في مادة الإحصاء ، وعدد ساعات الاستذكار في الأسبوع .

تسلسل	1	2	3	4	5	sum
x (الدرجة)	23	40	36	28	46	173
w (عدد ساعات الاستذكار)	1	3	3	2	4	

نجد أن الوسط الحسابي غير المرجح للدرجة الحاصل عليها الطالب هي :

$$\bar{x} = \frac{\sum x}{n} = \frac{23 + 40 + 36 + 28 + 46}{5} = \frac{173}{5} = 34.6$$

وإذا أردنا أن نحسب الوسط الحسابي للدرجات x المرجحة بعدد ساعات الاستذكار w ، يتم تطبيق المعادلة الآتية :

$$\begin{aligned} (\bar{w}) &= \frac{\sum xw}{\sum w} = \frac{23 \times 1 + 40 \times 3 + 36 \times 3 + 28 \times 2 + 46 \times 4}{1 + 3 + 3 + 2 + 4} \\ &= \frac{23 + 120 + 108 + 56 + 184}{13} = \frac{491}{13} = 37.769 \end{aligned}$$

وهذا الوسط المرجح أكثر دقة من الوسط الحسابي غير المرجح .
إذا الوسط الحسابي المرجح (\bar{w}) يحسب بتطبيق المعادلة الآتية :

$$\boxed{(\bar{w}) = \frac{\sum xw}{\sum w}}$$

(أ-٣)

مزايا وعيوب الوسط الحسابي

يتميز الوسط الحسابي بالمزايا التالية :

- أنه سهل الحساب .
- يأخذ في الاعتبار كل القيم .
- أنه أكثر المقاييس استخداماً وفهماً .
- ومن عيوبه .
- أنه يتأثر بالقيم الشاذة والمتطرفة .
- يصعب حسابه في حالة البيانات الوصفية .
- يصعب حسابه في حالة الجداول التكرارية المفتوحة .

الأسبوع السابع والثامن : الوسيط، طرق حسابه للبيانات المبوبة وغير المبوبة حسابيا وبيانيا ، المنوال

مفهومه ، حسابه للبيانات المبوبة وغير المبوبة

الوسيط Median

هو أحد مقاييس النزعة المركزية، والذي يأخذ في الاعتبار رتب القيم ، ويعرف الوسيط بأنه القيمة التي يقل عنها نصف عدد القيم $(n/2)$ ، ويزيد عنها النصف الآخر $(n/2)$ ، أي أن 50% من القيم أقل منه، 50% من القيم أعلى منه. وفيما يلي كيفية حساب الوسيط في حالة البيانات غير مبوبة ، والبيانات المبوبة.

الوسيط لبيانات غير مبوبة

1 - إذا كان عدد القيم فرديا فيكون ترتيب الوسيط كما في الصيغة الآتية

$$T = \frac{n+1}{2}$$

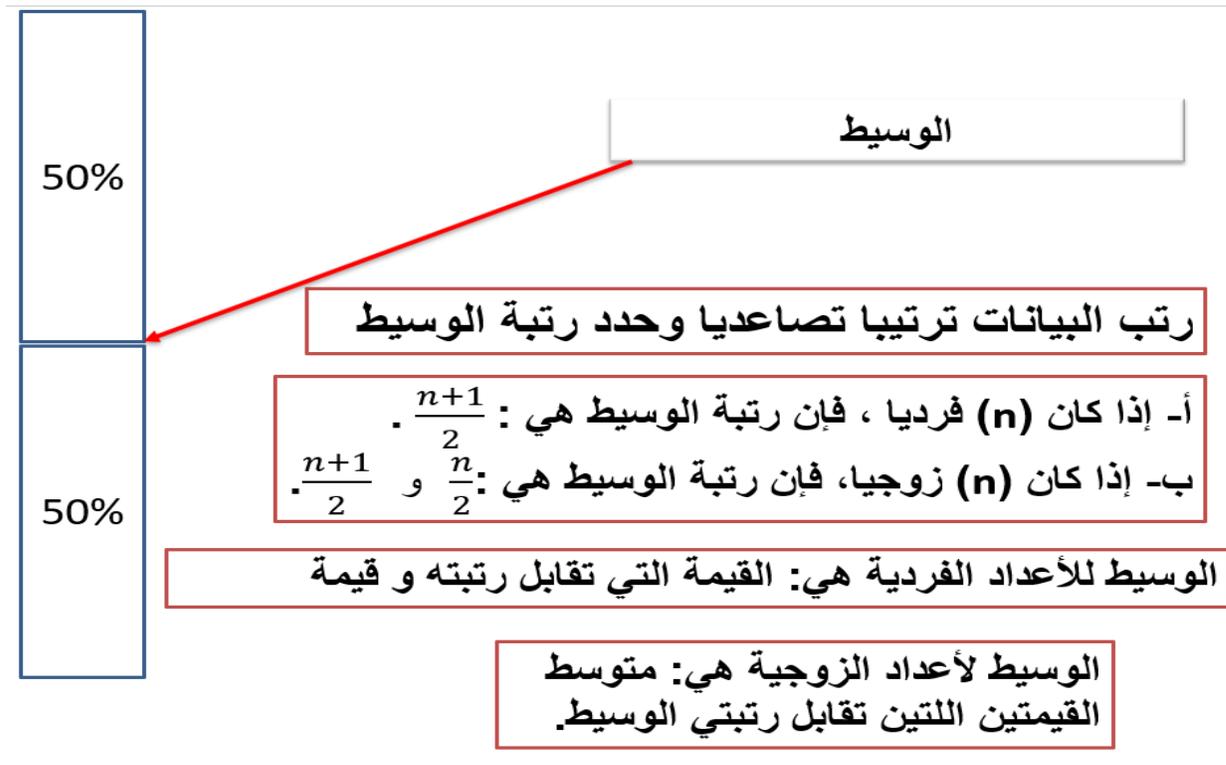
حيث ان T : تمثل ترتيب الوسيط وان n تمثل عدد القيم

خطوات ايجاد الوسيط

- 1 -ترتيب القيم اما تصاعديا او تنازليا
- 2 -نجد ترتيب الوسيط حسب الصيغة الآتية:

$$T = \frac{n+1}{2}$$

3 -تكون قيمة الوسيط هي القيمة الموجودة امام الترتيب الناتج في الخطوة (2)



مثال (11) اوجد الوسيط للبيانات الآتية :

27 , 12 , 15 , 300 , 22 , 18 , 14 , 24 , 17

الحل : 1- ترتيب البيانات : 12 , 14 , 15 , 17 , 18 , 22 , 24 , 27 , 300

الوسيط

2- نحدد موقع الوسيط (رتبته) كالتالي :

بما أن عدد البيانات : $n = 9$ وهو عدد فردي، إذن رتبة الوسيط هي :

$$\frac{n+1}{2} = 5 \quad \text{أي} \quad \frac{9+1}{2} = 5$$

3- إذن قيمة الوسيط هي القيمة الخامسة في البيانات المرتبة وهي :

$$\text{Med} = 18$$

مثال (12) اوجد الوسيط للبيانات الآتية :

3 , 9 , 12 , 16 , 14 , 33 , 10 , 15 , 20 , 7

الحل : 1- ترتيب البيانات : 3 , 7 , 9 , 10 , 12 , 14 , 15 , 16 , 20 , 33

القيمتين الوسيطيتين

2- تحديد موقع الوسيط: عدد البيانات $n = 10$ ، بما أن عدد البيانات عدد زوجي، فإن

$$\frac{10}{2} = 5 \quad \text{و} \quad \frac{10}{2} + 1 = 6$$

3- نحدد قيمة الوسيط، وهي متوسط القيمتين الوسيطيتين: الخامسة والسادسة، وهي :

$$\frac{12+14}{2} = 13$$

مثال (13) اوجد الوسيط للبيانات الآتية : . 134 , 78 , 204 , 63 , 12 , 189 , 152

الحل : نرتب القيم اما تصاعديا او تنازليا

ترتيب تصاعدي 12 , 63 , 78 , 134 , 152 , 189 , 204

او ترتيب تنازلي 204 , 189 , 152 , 134 , 78 , 63 , 12

$$\text{ترتيب الوسيط (T)} = \frac{n+1}{2}$$

$$7+1$$

$$T = \frac{7+1}{2} = 4$$

الوسيط هو الترتيب الرابع اي ان $Me = 134$
 س اوجد الوسيط لدرجات عينة من الطلبة قوامها 9 طلاب في امتحان معين
 الدرجات : 63 , 55 , 62 , 53 , 70 , 68 , 65 , 79 , 80
 2 - اذا كانت عدد القيم (n) زوجي فيكون الوسيط هو الوسط الحسابي لقيمتي الترتيبين
 اللتين تسلسلها على التوالي هو

$$N \quad n \\ \left(\frac{\quad}{2} + 1 \right) \text{ و } \left(\frac{\quad}{2} \right)$$

مثال (14) المطلوب ايجاد الوسيط للأعداد الآتية: 7 , 134 , 78 , 204 , 63 , 12 , 189 , 152
 الحل : نرتب القيم اما تصاعديا او تنازليا
 . الترتيب التصاعدي 7 , 12 , 63 , 78 , 134 , 152 , 189 , 204

$$N \quad 8 \\ 4 = \frac{\quad}{2} = \frac{8}{2} = T \quad 78 \text{ هو العدد الذي ترتيبه الرابع هو } T$$

$$N \quad 8 \\ 5 = 1 + \frac{\quad}{2} = 1 + \frac{8}{2} = T = 134 \text{ هو العدد الذي ترتيبه الخامس هو } T$$

الوسيط يمثل الوسط الحسابي لهاتين القيمتين اي ان

$$\frac{78+134}{2} = Me = 106$$

اما لو كان الترتيب تنازلي : 204 , 189 , 152 , 134 , 78 , 63 , 12 , 7

$$N \quad 8 \\ 4 = \frac{\quad}{2} = \frac{8}{2} = T \quad 78 \text{ هو العدد الذي ترتيبه الرابع هو } T$$

$$N \quad 8 \\ 5 = 1 + \frac{\quad}{2} = 1 + \frac{8}{2} = T = 134 \text{ هو العدد الذي ترتيبه الخامس هو } T$$

الوسيط يمثل الوسط الحسابي لهاتين القيمتين اي ان

$$\frac{134+78}{2} = Me = 106 \text{ الوسيط ايضا } 106$$

مثال (15): اوجد قيمة الدرجة الوسيطة 14 11 9 5 21 3 xi
 القيمة الرابعة
 21 14 11 9 5 3 الترتيب التصاعدي

$$T = \frac{n+1}{2} = \frac{7+1}{2} = 4$$

اذن قيمة الوسيط = 9

الحفيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
تمرين واجب (4): تم تقسيم قطعة أرض زراعية إلى 17 وحدة تجريبية متشابهة ، وتم
زراعتها بمحصول القمح ، وتم استخدام نوعين من التسميد هما : النوع (a) وجرب على
7 وحدات تجريبية ، والنوع (b) وجرب على 10 وحدات تجريبية ، وبعد انتهاء الموسم
الزراعي ، تم تسجيل إنتاجية الوحدة بالطن / هكتار ، وكانت بالجدول ادناه:

النوع
(a) 1.2 2.75 3.25 2 3 2.3 1.5

النوع
(b) 4.5 1.8 3.5 3.75 2 2.5 1.5 4 2.5 3

والمطلوب حساب وسيط الإنتاج لكل نوع من السماد المستخدم، ثم قارن بينها.

لحساب الوسيط من بيانات مبوبة في جدول توزيع تكراري ، يتم إتباع الخطوات الآتية .

- تكوين الجدول التكراري المتجمع الصاعد . $\left(\frac{n}{2}\right) = \left(\frac{\sum f}{2}\right)$
- تحديد رتبة الوسيط ملاحظة هنا التكرارات الاصلية وليس التكرار المتجمع الصاعد:
- تحديد فئة الوسيط كما في الشكل التالي :

تكرار متجمع صاعد سابق f_1 الحد الأدنى لفئة الوسيط (A)
 رتبة الوسيط $(n)/2$ ← الوسيط Med
 تكرار متجمع صاعد لاحق f_2 الحد الأعلى لفئة الوسيط

- ويحسب الوسيط ، بتطبيق المعادلة .

$$Med = A + \frac{\frac{n}{2} - f_1}{f_2 - f_1} \times L$$

(11-3)

حيث أن :

L هي طول فئة الوسيط، وتحسب بالمعادلة الآتية:
 طول الفئة = الحد الأعلى - الحد الأدنى

$$L = Upper - Lower$$

مثال (16) الاتي توزيع 50 عجل متوسط الحجم ، حسب احتياجاته اليومية من الغذاء الجاف بالكيلوجرام

فئات الاحتياجات اليومية	1.5 -	4.5 -	7.5 -	10.5 -	13.5 - 16.5
عدد العجول f	4	12	19	10	5

والمطلوب : حساب الوسيط : أ - حسابيا ب- بيانيا
 الحل

أولا : حساب الوسيط حسابيا

- رتبة الوسيط : $\frac{n}{2} = \frac{\sum f}{2} = \frac{50}{2} = 25$
- الجدول التكراري المتجمع الصاعد :

أقل من	تكرار متجمع صاعد
1.5	0
4.5	4
A 7.5	f_1 16
Med (الوسيط)	25
10.5	f_2 35
13.5	45
16.5	50

رتبة الوسيط

- تحديد فئة الوسيط : وهى الفئة التي تشمل قيمة الوسيط ، وهى قيمة أقل منها $\left(\frac{n}{2}\right)$ من القيم ، ويمكن معرفتها بتحديد التكرارين المتجمعين الصاعدين الذين يقع بينهما رتبة الوسيط $\left(\frac{n}{2}\right)$ ، وفى الجدول أعلاه نجد أن رتبة الوسيط (25) تقع بين التكرارين المتجمعين (16 , 35) ، ويكون الحد الأدنى لفئة الوسيط هو المناظر للتكرار المتجمع السابق 7.5 ، والحد الأعلى لفئة الوسيط هو المناظر للتكرار المتجمع الصاعد اللاحق 10.5 . أى أن فئة الوسيط هي : (7.5-10.5) .
- وبتطبيق معادلة الوسيط رقم (3-11) على هذا المثال نجد أن :

$$A = 7.5 , f_1 = 16 , f_2 = 35 , L = 10.5 - 7.5 = 3$$

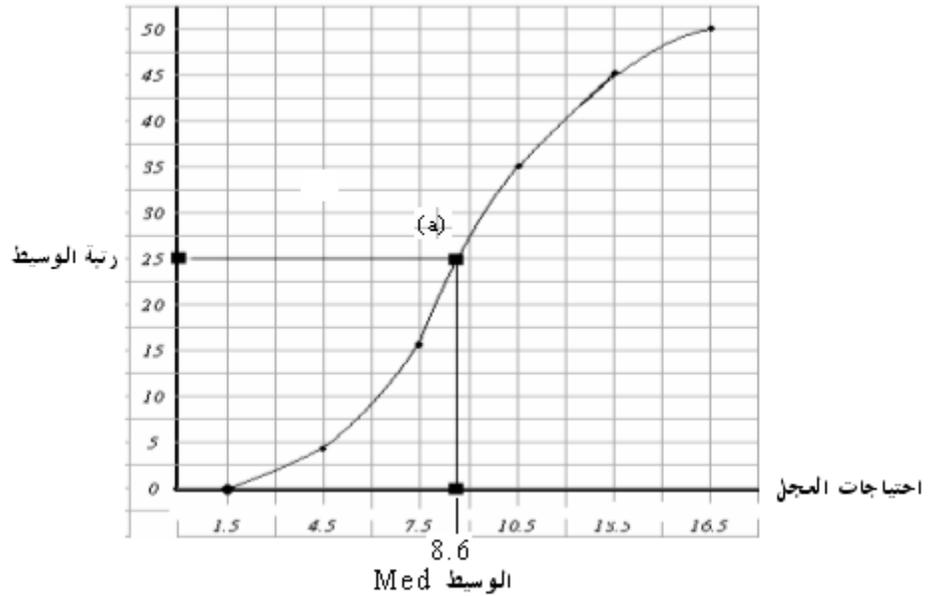
إذا الوسيط قيمته هي :

$$\begin{aligned} Med &= A + \frac{\frac{n}{2} - f_1}{f_2 - f_1} \times L = 7.5 + \frac{25 - 16}{35 - 16} \times 3 \\ &= 7.5 + \frac{9}{19} \times 3 = 7.5 + \frac{27}{19} = 7.5 + 1.421 = 8.921 \text{ k.g} \end{aligned}$$

ثانياً : حساب الوسيط بيانياً

- تمثيل جدول التوزيع التكراري المتجمع الصاعد بيانياً .

تكرار متجمع صاعد



- تحديد رتبة الوسيط (25) على المنحنى التكراري المتجمع الصاعد . ثم رسم خط مستقيم أفقي حتى يلقى المنحنى في النقطة (a) .
- إسقاط عمود رأسي من النقطة (a) على المحور الأفقي .
- نقطة تقاطع الخط الرأسي مع المحور الأفقي تعطى قيمة الوسيط .
- الوسيط كما هو مبين في الشكل $Med = 8.6$.

مزايا وعيوب الوسيط

من مزايا الوسيط

- 1- لا يتأثر بالقيم الشاذة أو المتطرفة .
 - 2- كما أنه سهل في الحساب .
 - 3- مجموع قيم الانحرافات المطلقة عن الوسيط أقل من مجموع الانحرافات المطلقة عن أي قيم أخرى .
- أي أن : $\sum |x - Med| \leq \sum |x - a| , a \neq Med$

ومن عيوب الوسيط

- 1- أنه لا يأخذ عند حسابه كل القيم في الاعتبار ، فهو يعتمد على قيمة أو قيمتين فقط .
- 2- يصعب حسابه في حالة البيانات الوصفية المقاسة بمعيار اسمي nominal

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
النوال Mode : هو القيمة التي تتكرر أكثر من غيرها ، أو القيمة الأكثر شيوعاً . وقد لا يكون للقيم منوال وقد يوجد أكثر من منوال واحد .

اولاً: حساب المنوال من بيانات غير مبوبة

المنوال (Mod) = القيمة (المستوى) الأكثر تكراراً

(٣-١٢)

مثال (17) : اوجد المنوال في البيانات الآتية 3 , 4 , 5 , 6 , 2 , 3

الحل : المنوال هو الرقم 3 لأنه تكرر اكثر من غيره من بين مفردات المجموعة.

مثال (18): اوجد المنوال في البيانات الآتية: 22,5,7,9,9,9,10,10,11,12,18

لها منوال واحد وهو 9

ملاحظة : بعض القيم تكون عديمة المنوال اذا لم يوجد رقم متكرر اكثر من غيره كما قد يكون هناك اكثر من منوال في المجموعة في حالة اكثر من رقم.

مثال (19) : احسب المنوال للبيانات الآتية : 2 , 4 , 6 , 8 , 10 , 12 , 14 , 19

لا يوجد منوال لهذه البيانات لأنه لا يوجد رقم متكرر

مثال (20) جد المنوال للبيانات الآتية : 2 , 4 , 8 , 10 , 12 , 2 , 5 , 4 , 6

المنوال هنا 4 منوال و 2 منوال لانهما تكررا بنفس المقدار

مثال (21) - جد المنوال للبيانات الآتية : 2,3,4,4,4,5,5,7,7,7,9

هنا منوالان وهما 4,7 وتسمى مجموعة ذات منوالين

مثال (22) اختيرت عينات عشوائية من طلاب بعض أقسام كلية الإدارة والاقتصاد ، وتم رصد درجات هؤلاء الطلاب في مادة مبادئ المحاسبة ، وكانت النتائج كالآتي:

قسم المحاسبة	80	77	75	77	77	77	65	70		58	67
قسم العلوم المالية	88	68	60	75	93	65	77	85		95	90
قسم الاقتصاد	80	65	69	80	65	88	76	65		86	80
قسم إدارة الاعمال	85	73	69	85	73	69	69	73		72	85

المطلوب : حساب منوال الدرجات لكل قسم من الأقسام :

الحل : هذه البيانات غير مبوبة ، لذا فإن :

المنوال = القيمة الأكثر تكراراً

والجدول ادناه يبين منوال الدرجة لكل قسم من الأقسام .

القسم	القيمة الأكثر تكراراً	القيمة المنوالية
قسم المحاسبة	الدرجة 77 تكررت 4 مرات	المنوال = 77 درجة
قسم العلوم المالية	جميع القيم ليس لها تكرار	لا يوجد منوال
قسم الاقتصاد	الدرجة 65 تكررت 3	يوجد منوالان هما :

	مرات الدرجة 80 تكرر 3 مرات	المنوال الأول = 65 المنوال الثاني = 80
قسم إدارة الاعمال	الدرجة 69 تكرر 3 مرات الدرجة 73 تكرر 3 مرات الدرجة 85 تكرر 3 مرات	يوجد ثلاث منوال هي : المنوال الأول = 69 المنوال الثاني = 73 المنوال الثالث = 85

ب- المنوال لبيانات مبوبة

يمكن ايجاد المنوال بعدة طرق بعد ايجاد الفئة المنوالية ، وتعرف الفئة المنوالية بانها : الفئة التي تحتوي على اكبر تكرار ، وذلك لان المنوال حسب التعريف هو القيمة التي تتكرر اكثر من غيرها . ومن هذه الطرائق :
- طريقة بيرسون وتسمى ايضا طريقة الفروق

$$Mod = A + \frac{d_1}{d_1 + d_2} \times L \quad (13-3)$$

حيث أن A : الحد الأدنى لفئة المنوال (الفئة المناظرة لأكبر تكرار) .

d_1 : الفرق الأول = (تكرار فئة المنوال - تكرار سابق)

d_2 : الفرق الثاني = (تكرار فئة المنوال - تكرار لاحق)

L : طول فئة المنوال .

فئة المنوال = الفئة المناظرة لأكبر تكرار

$$\begin{array}{l} \text{تكرار سابق} \\ \text{تكرار فئة المنوال} \\ \text{تكرار لاحق} \end{array} \left\{ \begin{array}{l} d_1 = (\text{تكرار فئة المنوال} - \text{تكرار سابق}) \\ d_2 = (\text{تكرار فئة المنوال} - \text{تكرار لاحق}) \end{array} \right.$$

ملخص هذه الطريقة (الخطوات)

- 1- نختار اكبر تكرار نفرضه والفئة التي تقابله هي الفئة المنوالية طولها (L)
- 2- نحدد التكرار الذي قبله ويرمز له d_1
- 3- نحدد التكرار الذي بعده ويرمز له d_2
- 4- نحدد الحد الادنى للفئة المنوالية (بداية الفئة المنوالية ويرمز لها (A)

$$Mod = A + \frac{d_1}{d_1 + d_2} \times L$$

(13-3)

5- نحسب قيمة المنوال بتطبيق صيغة القانون الاتي

مثال (23) ادناه جدول يبين توزيع 30 أسرة حسب الإنفاق الاستهلاكي الشهري لها بالألف دولار .

فئات الإنفاق	2 -	5 -	8 -	11 -	14 -
عدد الأسر f	4	7	10	5	4

المطلوب : حساب منوال الإنفاق الشهري للأسرة، باستخدام طريقة بيرسون (الفروق) .

الحل : لحساب المنوال لهذه البيانات يتم استخدام المعادلة رقم (3-12) ، ويتم إتباع الآتي :

• تحديد الفئة المنوالية : الفئة المنوالية هي الفئة المناظرة لأكبر تكرار : (8-11)

التكرارات	الفئات
4	2 -
7	5 -
10	8 -
5	11 -
4	14 - 17

فئة المنوال $A = 8$

$d_1 = 10 - 7 = 3$

أكبر تكرار

$d_2 = 10 - 5 = 5$

• حساب الفروق d ، حيث أن :

$$d_1 = (10 - 7) = 3 \quad d_2 = (10 - 5) = 5$$

• تحديد الحد الأدنى للفئة المنوالية ($A = 8$) ، وكذلك طول الفئة ($L = 3$)

• وبتطبيق المعادلة الخاصة بحساب المنوال في حالة البيانات المبوبة . نجد أن :

$$Mod = A + \frac{d_1}{d_1 + d_2} \times L$$

$$= 8 + \frac{3}{3 + 5} \times 3 = 8 + 1.125 = 9.125$$

مثال شامل (24) : قام مدير مراقبة الإنتاج بسحب عينة من 10 عبوات من المياه المعبأة للشرب ، ذات الحجم 5 لتر ، والمنتجة بواسطة إحدى شركات تعبئة المياه لفحص كمية الأملاح الذائبة، وكانت كالاتي : 115 123 119 123 124 119 123 121 123 121

المطلوب : حساب الوسط الحسابي، والوسيط، والمنوال

الحل : حساب الوسط الحسابي :

$$\bar{x} = \frac{\sum x}{n} = \frac{1211}{10} = 121.1$$

• حساب الوسيط :

$$\text{رتبة الوسيط} : \frac{n+1}{2} = \frac{10+1}{2} = 5.5$$

ترتيب القيم تصاعديا

					قيمة الوسط					
الطاقة	115	119	119	121	121	122	123	123	123	124
الرتبة	1	2	3	4	5	5.5	6	7	8	9
					رتبة الوسيط					

عدد القيم = 10 ، وهو عدد زوجي . الوسيط = الوسط الحسابي للقيمتين رقم (6 ، 5)

$$Med = \frac{121 + 123}{2} = \frac{244}{2} = 122$$

• حساب المنوال :

المنوال يساوي القيمة الأكثر تكرارا: القيمة 123 تكررت أكثر من غيرها ، إذا

$$Mod = 123$$

تمرين شامل (واجب) (5) : الجدول التكراري التالي يعرض توزيع 100 عامل في مزرعة حسب الأجر اليومي بالدولار .

الأجر	50	70	90	110	130	150	170 - 190
عدد العمال	8	15	28	20	15	8	6

المطلوب : حساب الوسط والوسيط والمنوال .

الحل : حساب الوسط والوسيط والمنوال .

الأسبوع التاسع : مقاييس التشتت مفهومها واستخدامها ، المدى للبيانات المبوبة وغير المبوبة ، الانحراف الربيعي للبيانات غير المبوبة

مقاييس التشتت Measures of Dispersion

مقاييس التشتت تشير إلى مدى تقارب أو تباعد أو تناثر أو اختلاف القيم عن بعضها البعض أو عن نقطة معينة كالتوسط مثلاً . فالتشتت يرجع إلى اختلاف قيم البيانات، لذلك فإن القيم المتشابهة لا يوجد لها تشتت، ولكن باختلاف قيم البيانات فلا بد من وجود تشتت بين هذه القيم، ويختلف مقدار التشتت من مجموعة إلى أخرى. وتبرز الحاجة إلى مقاييس التشتت لتقديم وصف دقيق حول التوزيعات التكرارية وخصوصاً عند إجراء المقارنات بين المجموعات أو بين الأفراد، فدالات مقاييس النزعة المركزية كالتوسط الحسابي قد تقود إلى اتخاذ قرارات غير صحيحة.

لقد ذكرنا في المحاضرات السابقة بعضاً من مقاييس النزعة المركزية والتي تمثل مقاييس عددية لموضع أو مكان تركيز البيانات لظاهرة ما. وقد ذكرنا بأن هذه المقاييس تستخدم لمقارنة مجموعات البيانات المختلفة. وفي الحقيقة فإن مقاييس النزعة المركزية غير كافية لإيجاد مقارنة شاملة بين مجموعات البيانات المختلفة. فقد تكون هناك مجموعات من البيانات لها نفس مقاييس النزعة المركزية (لها نفس الموضع) ولكنها تختلف في بعض الصفات الأخرى. فمثلاً المثال التالي يبين لنا مجموعتين من البيانات لهما نفس المتوسط ولكنهما مختلفتان في طبيعة تشتت البيانات.

مفهوم مقاييس التشتت: هي مقاييس عددية تستخدم لقياس اختلاف أو تشتت البيانات. والاختلاف أو التشتت لمجموعة من البيانات هو مقدار تفرق أو تباعد أو انتشار البيانات فيما بينها. فتشتت البيانات يكون صغيراً إذا كانت البيانات متقاربة فيما بينها والعكس بالعكس. وأما البيانات المتساوية فلا اختلاف ولا تشتت فيها. ومقاييس التشتت تستخدم لوصف مجموعة البيانات وكذلك لمقارنة مجموعات البيانات المختلفة إذ أن مقاييس النزعة المركزية لا تكفي وحدها لوصف مجموعة البيانات أو مقارنة مجموعات البيانات المختلفة. ومن أشهر مقاييس التشتت نذكر:

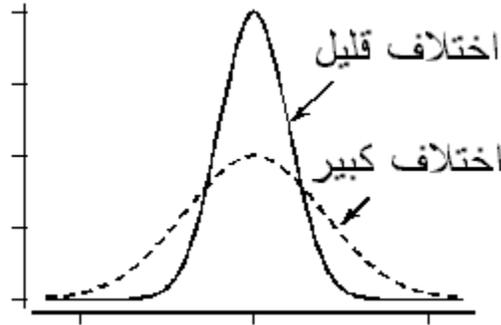
1. المدى: Range
2. الانحراف الربيعي او نصف المدى الربيعي Semi-Inter-quartile Range
3. التباين: Variance
4. الانحراف المعياري: Standard Deviation
5. معامل الاختلاف (أو التغير): Coefficient of Variation

مثال (25): الجدول ادناه يوضح مجموعتين لهما نفس المتوسط $\bar{x} = 60$ ونفس الوسيط $Med = 60$ ونفس المنوال $Mod = 60$ ولكنهما مختلفتان في طبيعة التشتت.

المجموعة	البيانات	شكل انتشار البيانات
الأولى	59, 61, 62, 60, 58, 60	
الثانية	50, 60, 66, 54, 60, 70	

بالرغم من تساوي مقاييس النزعة المركزية للمجموعتين إلا أن التشتت (أو الاختلاف) بين القيم في كل مجموعة غير متساو. فمن الواضح أن بيانات المجموعة الأولى أكثر تقارباً فيما بينها (أقل تشتتاً وتباعداً فيما بينها) من بيانات المجموعة الثانية. لذلك دعت الحاجة لإيجاد مقاييس تقيس طبيعة تشتت (أو تفرق أو اختلاف أو تباعد) البيانات فيما بينها. هذه المقاييس تسمى مقاييس التشتت أو الاختلاف.

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
الشكل ادناه يوضح المضلعين التكراريين لمجموعتين من البيانات لهما نفس مقاييس النزعة المركزية ولكنهما مختلفتان في طبيعة التشتت.



شكل (1) المضلعان التكراريان لتوزيعين لهما نفس مقاييس النزعة المركزية ولكنهما مختلفين في التشتت

(2-4) المدى: Range

يعد المدى من أسهل مقاييس التشتت تعريفًا وحسابًا ويعطينا فكرة سريعة عن مدى تفرق البيانات.
تعريف (1): نعرف المدى لمجموعة من البيانات على أنه الفرق بين أكبر قيمة وأصغر قيمة، أي أن المدى هو:

$$(1) \quad Range = X_{max} - X_{min}$$

حيث نعرف X_{max} و X_{min} كما يلي:

(أ) للبيانات المفردة:

$$X_{max} = \text{أكبر قيمة}$$

$$X_{min} = \text{أصغر قيمة}$$

(ب) للبيانات المبوبة:

$$X_{max} = \text{مركز الفئة العليا}$$

$$X_{min} = \text{مركز الفئة الدنيا}$$

ملاحظة (1): تعرف بعض الكتب أكبر قيمة وأصغر قيمة للبيانات المبوبة كما يلي:

$$X_{max} = \text{الحد الأعلى للفئة العليا}$$

$$X_{min} = \text{الحد الأدنى للفئة الدنيا}$$

مثال (26): أوجد المدى للمشاهدات ادناه وهي عبارة عن أوزان (بالكيلوجرام) مجموعة مكونة من سبعة أشخاص:

$$25, 30, 40, 45, 35, 55, 50$$

الحل:

$$X_{max} = 55$$

$$X_{min} = 25$$

$$Range = X_{max} - X_{min} = 55 - 25 = 30 \text{ (كيلوجرامًا)}$$

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
مثال (27): أوجد المدى لمستوى الهيموجلوبين في الدم لعينة مكونة من خمسين شخصا تم تلخيص مستوى الهيموجلوبين لهم.

مستوى الهيموجلوبين	التكرار f
12.95 – 13.95	3
13.95 – 14.95	5
14.95 – 15.95	15
15.95 – 16.95	16
16.95 – 17.95	10
17.95 – 18.95	1

الحل:

مستوى الهيموجلوبين	مركز الفترة x	التكرار f
12.95 – 13.95	13.45	3
13.95 – 14.95	14.45	5
14.95 – 15.95	15.45	15
15.95 – 16.95	16.45	16
16.95 – 17.95	17.45	10
17.95 – 18.95	18.45	1

$$X_{max} = \text{مركز الفئة العليا} = 18.45$$

$$X_{min} = \text{مركز الفئة الدنيا} = 13.45$$

$$\text{Range} = X_{max} - X_{min} = 18.45 - 13.45 = 5.00$$

بعض مميزات وعيوب المدى:

- أ- يتميز المدى بسهولة التعريف والحساب
- ب- يعيب المدى العيوب التالية:
 1. يتأثر المدى بالقيم الشاذة أو المتطرفة.
 2. لا يأخذ المدى في الاعتبار جميع البيانات.

ملاحظة (2):

1. وحدة المدى هي نفس وحدة البيانات الأصلية.
2. نظرًا لأن المدى يعتمد فقط على أكبر وأصغر قيمة ولا يأخذ في الاعتبار القيم الأخرى فهو مقياس غير جيد لقياس التشتت.

الأسبوع العاشر : الانحراف الربيعي للبيانات المبوبة حسابيا وبيانيا

الانحراف الربيعي

رأينا أن المدى يتأثر كثيراً بالقيم الشاذة أو المتطرفة. ولذلك دعت الحاجة لإيجاد مقاييس أخرى للتشتت لا تتأثر بالقيم المتطرفة. وأحد هذه المقاييس هو نصف المدى الربيعي. وحيث أن القيم المتطرفة هي تلك القيم الصغيرة جداً أو الكبيرة جداً فإنه عند حساب الانحراف الربيعي لا يؤخذ في الاعتبار ربع البيانات الصغيرة (25%) ولا ربع البيانات الكبيرة (25%). الشكل ادناه يوضح موضع القيم الشاذة والمتطرفة.



الرباعيات Quartiles

عند تقسيم القيم إلى أربع أجزاء متساوية، يوجد ثلاث إحصاءات ترتيبية تسمى بالرباعيات، وهي:

- الربع الأول: وهو القيمة التي يقل عنها ربع عدد القيم، أي يقل عنها 25% من القيم، ويرمز له بالرمز Q_1 .
 - الربع الثاني: وهو القيمة التي يقل عنها نصف عدد القيم، أي يقل عنها 50% من القيم، ويرمز له بالرمز Q_2 ، ومن ثم يعبر هذا الربع عن الوسيط.
 - الربع الثالث: وهو القيمة التي يقل عنها ثلاث أرباع عدد القيم، أي يقل عنها 75% من القيم، ويرمز له بالرمز Q_3 .
- والشكل (3-3) يبين أماكن الرباعيات الثلاث.

شكل (3-3)
الرباعيات



يرمز لنصف المدى الربيعي بالرمز Q ويعرف بالصيغة التالية:

حيث أن Q_1 هو الربع الأول و Q_3 هو الربع الثالث

ويعرف الانحراف الربيعي بأنه متوسط الفرق بين الربع الثالث والربع الأول لمجموعة من البيانات سواء كانت مبوبة أو غير مبوبة. فإذا رمزنا للربع الأول (الأدنى) Q_1 وللربع الثالث بالرمز Q_3 وللانحراف الربيعي Q وعليه فإن الانحراف الربيعي = الربع الثالث - الربع الأول

$$(2) \quad Q = \frac{Q_3 - Q_1}{2}$$

وطريقة حساب الربيعين الأدنى والأعلى هي تماما كطريقة حساب الوسيط

حساب الربيعين من بيانات غير مبوبة

لحساب قيمة الربيع الأعلى والربيع الأدنى يجب تحديد ترتيب (موقع) كل منهما مسبقا مثلما فعلنا عند حساب قيمة الوسيط.

ترتيب موقع (الربيع الأدنى) = عدد القيم / 4

وعليه فإن : $Q_1 = n / 4$ حيث ان Q_1 تمثل ترتيب الربيع الأول وان n هي عدد القيم

وترتيب الربيع الثالث (الأعلى) = (عدد القيم / 4) * 3

$Q_3 = 3 * (n / 4)$ (أي على بعد 75% من بداية البيانات ويجب هنا أيضا ان يعاد ترتيب

مجموعة القيم تصاعديا او تنازليا قبل حساب موقع او ترتيب الربيعين

مثال (28) احسب الربيعين الأعلى والأدنى والانحراف الربيعي لمجموعة البيانات الآتية:

3, 7, 5, 2, 8, 12, 10, 15

الحل: 1- نرتب القيم ترتيبا تصاعديا او ترتيبا تنازليا

ترتيب تصاعدي

2,3,5,7,8,10,12,15

2 نجد ترتيب الربيع الأدنى (الأول)

$$Q_1 = \frac{n}{4} = \frac{8}{4} = 2$$

أي ان ترتيب الربيع الأول هو الترتيب الثاني من بين البيانات ويساوي (3 اي ان $Q_1 = 3$)

3- نحدد ترتيب الربيع الأعلى (الثالث)

$$Q_3 = 3 * \frac{n}{4} = 3 * \frac{8}{4} = 6$$

بمعنى ان موقع الربيع الثالث هو الترتيب Q_3 السادس من البيانات ويساوي 10 اي ان $Q_3 = 10$

4- نجد الانحراف الربيعي والذي يمثل متوسط الفرق بين الربيع الأعلى والربيع الأدنى

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

$$Q = \frac{10 - 3}{2}$$

$$= 3.5$$

الانحراف الربيعي للبيانات المبوبة حسابيا وبيانيا

طريقة حساب الربيع الاعلى والادنى من بيانات مبوبة (جدول تكراري) تماثل تماما طريقة حساب الوسيط السابق شرحها.

خطوات ايجاد الربيعين لبيانات مبوبة كالآتي:

- 1- اعداد جدول تكراري تجمعي صاعد او نازل
- 2- تحديد ترتيب الربيع الادنى وذلك بتطبيق الصيغة الآتية:

$$Q_1 = \frac{\sum fi}{4}$$

هي مجموع التكرارات

- 3- تحديد ترتيب الربيع الاعلى بالصيغة الآتية:

$$Q_3 = 3 \times \left(\frac{\sum fi}{4} \right)$$

- 4- حساب قيمة كل من الربيعين من جدول تكراري متجمع صاعد او متجمع نازل وباستخدام الصيغة الآتية:

$$Q_1 = A + \left(\frac{R - F_1}{F_2 - F_1} \right) \times L$$

بداية فئة الربيع الادنى A

ترتيب الربيع الادنى R ويساوي $Q_1 = \frac{\sum fi}{4}$

التكرار المتجمع السابق الادنى F_1

الفرق بين التكرارين الصاعد السابق F_1 واللاحق F_2 وهو نفسه التكرار الاصلي لفئة الربيع الأدنى

طول فئة الربيع الأدنى L

لاستخراج الربيع الأعلى Q_3 نطبق المعادلة الثالثة

$$Q_3 = A^* + \left(\frac{R^* - F_1^*}{F_2^* - F_1^*} \right) \times L^*$$

بداية فئة الربيع الادنى A

ترتيب الربيع الاعلى R^* ويساوي $Q_3 = 3 \times \frac{\sum fi}{4}$

التكرار المتجمع السابق الأعلى F_1^*

الفرق بين التكرارين الصاعد السابق F_1^* واللاحق F_2^* وهو نفسه التكرار الاصلي لفئة الربيع الاعلى

طول فئة الربيع الاعلى L

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
مثال (29): أوجد الانحراف الربيعي لمستوى الهيموجلوبين في الدم لعينة مكونة من خمسين شخصا تم تلخيص
 مستوى الهيموجلوبين (كما في مثال 27) لهم باستخدام: الطريقة الحسابية. (ب) الطريقة البيانية.

مستوى الهيموجلوبين	التكرار f
12.95 – 13.95	3
13.95 – 14.95	5
14.95 – 15.95	15
15.95 – 16.95	16
16.95 – 17.95	10
17.95 – 18.95	1

الحل: (أ) الطريقة الحسابية:

مستوى الهيموجلوبين	التكرار المتجمع الصاعد
12.95 أقل من	0
13.95 أقل من	3
14.95 = A أقل من	8 = F_1
15.95 = A* أقل من	23 = F_2 = F_1^*
16.95 أقل من	39 = F_2^*
17.95 أقل من	49
18.95 أقل من	50

$Q_1 \Rightarrow$

$Q_3 \Rightarrow$

$$R = \frac{n}{4} = 12.5$$

$$R^* = \frac{3n}{4} = 37.5$$

حساب الربيع الأول Q_1 :

$$R = \frac{n}{4} = \frac{50}{4} = 12.5$$

$$A = 14.95, L = 1.00, F_1 = 8, F_2 = 23$$

$$Q_1 = A + \left(\frac{R - F_1}{F_2 - F_1} \right) \times L = 14.95 + \left(\frac{12.5 - 8}{23 - 8} \right) \times 1.00 = 15.25$$

حساب الربيع الثالث Q_3 :

$$R^* = \frac{3n}{4} = \frac{3 \times 50}{4} = 37.5$$

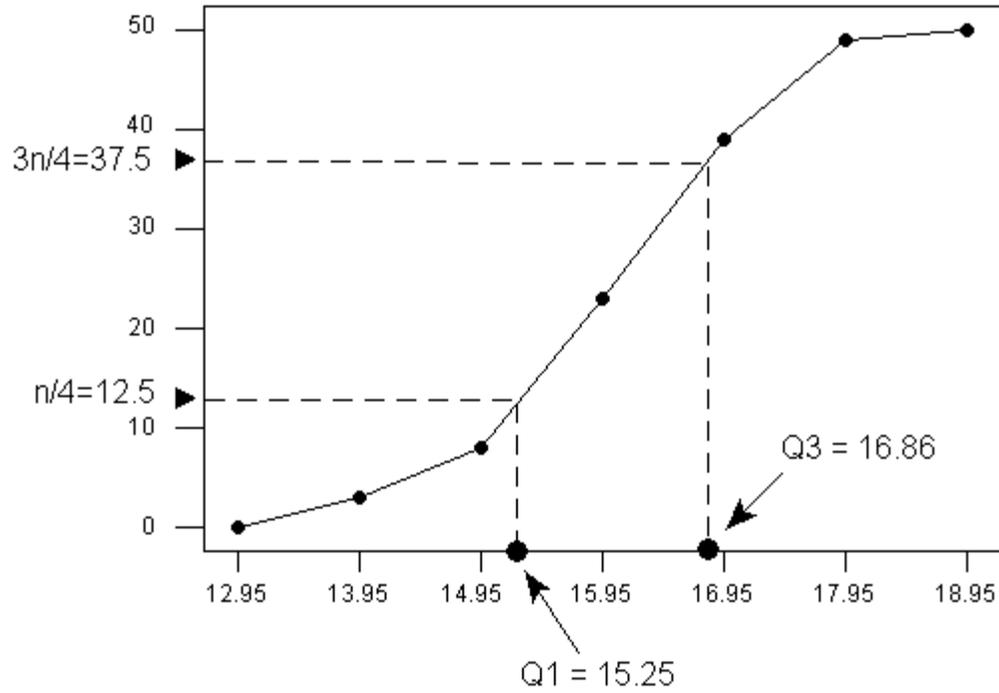
$$A^* = 15.95, L^* = 1.00, F_1^* = 23, F_2^* = 39$$

$$Q_3 = A^* + \left(\frac{R^* - F_1^*}{F_2^* - F_1^*} \right) \times L^* = 15.95 + \left(\frac{37.5 - 23}{39 - 23} \right) \times 1.00 = 16.86$$

وباستخدام الصيغة فإن نصف المدى الربيعي هو:

$$Q = \frac{Q_3 - Q_1}{2} = \frac{16.86 - 15.25}{2} = \frac{1.61}{2} = 0.805$$

(ب) الطريقة البيانية:



بعض مميزات وعيوب الانحراف الربيعي:

1. من المميزات أنه لا يتأثر بالقيم الشاذة أو المتطرفة.
2. من العيوب أنه لا يأخذ في الاعتبار جميع البيانات.

ملاحظة (3): وحدة نصف الانحراف الربيعي هي نفس وحدة البيانات الأصلية.

الأسبوع الحادي عشر والثاني عشر : الانحراف المعياري ، مفهومه واهميته ، طرق حسابه للبيانات المبوبة وغير المبوبة (الطريقة المطولة ، الطريقة المختصرة)

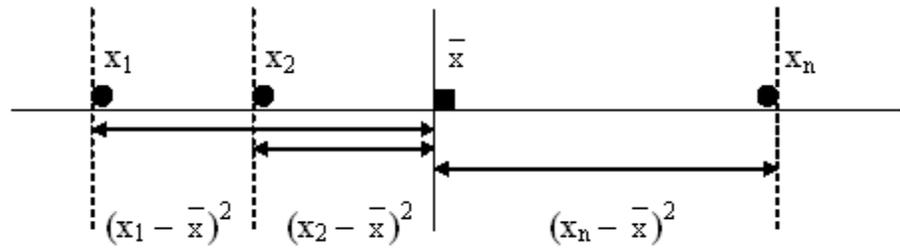
التباين (Variance) والانحراف المعياري (Standard Deviation):

يعد التباين والانحراف المعياري من أهم وأفضل مقاييس التشتت ومن أكثرها شيوعاً واستخداماً في التحليل الإحصائي وذلك لما يتمتعان به من خصائص وصفات إحصائية جيدة.

التباين: Variance

فكرة التباين تعتمد على تشتت أو تباعد البيانات عن متوسطها. فالتباين يكون كبيراً إذا كانت البيانات متباعدة عن متوسطها والعكس صحيح.

ويعرف التباين بأنه متوسط مربع انحرافات القيم عن الوسط الحسابي ويرمز له بالرمز S^2 . الشكل التالي يبين مربعات انحرافات القيم عن الوسط الحسابي.



والجدول ادناه يلخص طريقة حساب انحرافات ومربعات انحرافات القيم عن المتوسط.

x_1	x_2	...	x_n	القيم (البيانات)
$x_1 - \bar{x}$	$x_2 - \bar{x}$...	$x_n - \bar{x}$	انحرافات القيم عن المتوسط
$(x_1 - \bar{x})^2$	$(x_2 - \bar{x})^2$...	$(x_n - \bar{x})^2$	مربع انحرافات القيم عن المتوسط

تستخدم مربعات انحرافات القيم عن الوسط الحسابي لحساب التباين. فمن الواضح أن تشتت البيانات يزداد بزيادة مربعات الانحرافات والعكس صحيح. والسؤال الذي يتبادر للأذهان هو أي واحد من هذه المربعات ينبغي أن يستخدم لقياس التشتت؟ إن من المنطقي أن نبحث عن قيمة نموذجية تمثل هذه المربعات لاستخدامها لقياس التباين. وبناءً على دراستنا السابقة لمقاييس النزعة المركزية فإن المتوسط من أفضل المقاييس التي تمثل مجموعة البيانات.

وعليه فإن متوسط مربعات انحرافات القيم عن الوسط الحسابي $\sum_{i=1}^n (x_i - \bar{x})^2 / n$ يعتبر أحد المقاييس التي يمكن استخدامها لقياس التباين.

الانحراف المعياري: Standard Deviation

إن التباين من أهم وأفضل مقاييس التشتت ولكنه يقاس بوحدة البيانات الأصلية المربعة. وفي كثير من الأحيان نرغب في استخدام مقياس للتشتت يقاس بوحدة البيانات الأصلية ويتمتع بخصائص إحصائية جيدة مثل التباين. وأحد هذه المقاييس هو الانحراف المعياري. ويعرف الانحراف المعياري على أنه الجذر التربيعي الموجب للتباين ويرمز له بالرمز s .

حساب التباين والانحراف المعياري:

سنستعرض طرق حساب التباين والانحراف المعياري في حالة البيانات المفردة وفي حالة البيانات المبوبة.

أولاً: التباين والانحراف المعياري للبيانات المفردة (غير المبوبة):

إذا كانت x_1, x_2, \dots, x_n عينه حجمها n وكان متوسطها هو \bar{x} فإن تباين العينة يعرف كما يلي:

$$(3) \quad s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n - 1} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

لاحظ أننا قسمنا على المقدار $(n-1)$ ، وهو ما يسمى بدرجات الحرية، بدلاً من القسمة على عدد البيانات n في الصيغة السابقة وذلك لكي نحصل على مقياس يتمتع بصفات إحصائية جيدة. وأما الانحراف المعياري فإنه يعرف بالصيغة:

$$(4) \quad s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

ملاحظة (4):

1. $s^2 \geq 0$ (دائمًا) وكذلك $s \geq 0$ (دائمًا).
2. $s = 0 \Leftrightarrow s^2 = 0 \Leftrightarrow$ جميع قيم العينة متساوية (لا يوجد اختلاف بين القيم).
3. وحدة التباين، s^2 ، هي وحدة البيانات الأصلية المربعة.
4. وحدة الانحراف المعياري، s ، هي نفس وحدة البيانات الأصلية.

مثال (30): أوجد تباين العينة والانحراف المعياري لمجموعة الأوزان (بالكيلوجرام):

7.1, 2.5, 2.5, 5.4, 8.3

الحل: نحتاج عمل جدول:

x_i	الانحراف $(x_i - \bar{x}) = (x_i - 5.16)$	مربع الانحراف $(x_i - \bar{x})^2$
7.1	1.94	3.7636
2.5	-2.66	7.0756
2.5	-2.66	7.0756
5.4	0.24	0.0576
8.3	3.14	9.8596
$\sum_{i=1}^n x_i = 25.8$	0.00	$\sum_{i=1}^n (x_i - \bar{x})^2 = 27.832$

من هذا الجدول نوجد الكميات التالية:

$$\sum_{i=1}^n (x_i - \bar{x})^2 = 27.832, \quad \sum_{i=1}^n x_i^2 = 160.96, \quad \sum_{i=1}^n x_i = 25.8, \quad n = 5,$$

إن متوسط العينة هو:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{25.8}{5} = 5.16 \quad (\text{كيلوجراماً})$$

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} = \frac{27.832}{5-1} = 6.958 \quad (\text{كيلوجرامًا مربعًا})$$

الانحراف المعياري هو:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} = s = \sqrt{6.958} = 2.6378 \quad (\text{كيلوجرامًا})$$

ثانياً: التباين والانحراف المعياري للبيانات المبوبة:

الطريقة المطولة : (ساب انحرافات مراكز الفئات عن الوسط الحسابي) باستخدام الصيغة الآتية

$$\sqrt{\frac{\sum f_i(x_i - \bar{x})^2}{\sum F_i}}$$

$\sum F_i$ تمثل مجموع التكرارات اما $(x_i - \bar{x})$ تمثل انحرافات القيم عن الوسط الحسابي
خطوات الحل:

1- نجد مراكز الفئات X_i

2- نجد الوسط الحسابي باستخدام الصيغة $x = \frac{\sum f_i X_i}{\sum f_i}$

3- نجد انحرافات مراكز الفئات عن الوسط الحسابي $(x_i - \bar{x})$

4- نربع انحرافات مراكز الفئات عن الوسط الحسابي

5- نضرب التكرارات في مربعات الانحراف

6- نجمع حاصل الضرب

7- نطبق الصيغة القانون

$$\sqrt{\frac{\sum f_i(x_i - \bar{x})^2}{\sum F_i}}$$

مثال (31): الآتي جدول توزيع تكراري لدرجات مجموعة من الطلبة في امتحان فصلي المطلوب : ايجاد الانحراف المعياري بالطريقة المطولة

الفئات	0-2	2-4	4-6	6-8	8-10	المجموع
التكرارات	5	10	20	10	5	50

خطوات الحل:

الحل :
الخطوات (٦) (٥) (٤) (٣) (٢) (١)

الفئات	f_i	X_i	$F_i X_i$	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	$F_i(X_i - \bar{X})^2$
0-2	5	1	5*1=5	1-5=-4	4 ² =16	5*16=80
2-4	10	3	30	-2	4	40
4-6	20	5	100	0	0	0
6-8	10	7	70	2	4	40
8-10	5	9	45	4	16	80
\sum	50		250		40	240

$$x = \frac{\sum f_i X_i}{\sum f_i} = \frac{250}{50} = 5$$

الخطوة (7)

$$S = \sqrt{\frac{\sum f_i(x-\bar{x})^2}{\sum F_i}} =$$

$$s = \sqrt{\frac{240}{50}} = 2.190$$

مثال (32): أوجد التباين والانحراف المعياري لمستوى الهيموجلوبين في الدم لعينة مكونة من خمسين شخصا تم تلخيص مستوى الهيموجلوبين لهم كما في مثال (27).

مستوى الهيموجلوبين	التكرار f
12.95 – 13.95	3
13.95 – 14.95	5
14.95 – 15.95	15
15.95 – 16.95	16
16.95 – 17.95	10
17.95 – 18.95	1

الحل: نلخص إيجاد التباين باستخدام الجدول الآتي:

مستوى الهيموجلوبين	مركز الفئة x	التكرار f	xf	$x^2 f$	$f(x - \bar{x})^2$ $f(x - 16.01)^2$
12.95 – 13.95	13.45	3	40.35	542.708	19.6608
13.95 – 14.95	14.45	5	72.25	1044.013	12.1680
14.95 – 15.95	15.45	15	231.75	3580.538	4.7040
15.95 – 16.95	16.45	16	263.20	4329.640	3.0976
16.95 – 17.95	17.45	10	174.50	3045.025	20.7360
17.95 – 18.95	18.45	1	18.45	340.403	5.9536
المجموع		$\sum f$ = 50	$\sum xf$ = 800.5	$\sum x^2 f$ = 12882.33	$\sum_{i=1}^k f_i (x_i - \bar{x})^2$ = 66.320

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{\sum xf}{n} = \frac{800.5}{50} = 16.01$$

حساب التباين

$$s^2 = \frac{\sum_{i=1}^k f_i (x_i - \bar{x})^2}{n-1} = \frac{66.320}{50-1} = 1.3535$$

حساب الانحراف المعياري:

$$s = \sqrt{s^2} = \sqrt{1.3536} = 1.163$$

معامل الاختلاف (التغير): Coefficient of Variation:

ذكرنا سابقاً أن التباين والانحراف المعياري من المقاييس المفيدة لقياس التشتت لتوزيع متغير ما. ولكن في كثير من الأحيان نكون مهتمين بمقارنة التشتت والاختلاف لتوزيعي متغيرين مختلفين. وبما أن التباين والانحراف المعياري مقياسان يعتمدان على وحدة البيانات فإنه يصعب استخدامهما لمقارنة تجانس المجموعات المختلفة من البيانات وذلك لاختلاف الوحدة المستخدمة. وبشكل عام فإن مقاييس التشتت التي ذكرناها آنفاً تكون غير مناسبة لمقارنة تجانس مجموعات البيانات المختلفة في الحالتين التاليتين:

1. إذا كانت وحدتا المتغيرين مختلفتين حيث لا نستطيع مقارنة الوحدات المختلفة ببعض.
2. إذا كان متوسطا المتغيرين مختلفين وذلك لأن تباين توزيع المتغير ذي المتوسط الصغير ينزع لأن يكون صغيراً والعكس بالعكس.

لذلك دعت الحاجة إلى مقياس لا يعتمد على وحدة المتغير ويقاس ما يسمى بالتشتت النسبي. وأحد هذه المقاييس هو ما يسمى بمعامل الاختلاف أو معامل التغير. فمعامل الاختلاف هو أحد مقاييس التشتت النسبي وهو مقياس عديم الوحدة ويستخدم لمقارنة التشتت النسبي أو التجانس لمجموعات البيانات المختلفة. فمجموعة البيانات ذات معامل الاختلاف الأكبر يكون تشتتها النسبي أكبر أي أنها تكون أقل تجانساً والعكس بالعكس. ويعرف معامل الاختلاف للعينة التي متوسطها \bar{x} وانحرافها المعياري s بالصيغة التالية:

$$C.V. = \frac{s}{\bar{x}} \quad (10)$$

مثال (33): الجدول أدناه يتضمن بيانات إحدى الدراسات التي طبقت على خمسة أشخاص لقياس الوزن (بالكيلوجرام) والطول (بالسنتمتر). أي البيانات أكثر تشتتاً نسبياً (أقل تجانساً)؛ بيانات الأوزان أم بيانات الأطوال؟

رقم الشخص	1	2	3	4	5
الوزن	69	59	65	67	65
الطول	164	162	155	165	158

الحل: أولاً نوجد المتوسط \bar{x} والانحراف المعياري s لكل من بيانات الأوزان وبيانات الأطوال كما مر معنا سابقاً. نلخص الحسابات في الجدول التالي:

البيانات	المتوسط \bar{x}	الانحراف المعياري s	معامل الاختلاف $C.V. = \frac{s}{\bar{x}}$
الأوزان	65.0 kg	3.7417 kg	0.0576
الأطوال	160.8 cm	4.2071 cm	0.026

الدرجات (القيم) المعيارية : Standard Scores (Values)

نستطيع بكل يسر وسهولة استخدام قيمتي مشاهدتين في نفس المجموعة لمقارنتهما ببعض. فمثلاً، نستطيع أن نقول بأن أداء الطالب الحائز على الدرجة 85 في اختبار مقرر ما أفضل من أداء الطالب الحائز على الدرجة 80 في نفس الاختبار إذا كان الطالبان في نفس الشعبة. وفي المقابل، لا نستطيع القول بأن أداء الطالب الحائز على الدرجة 85 في اختبار مقرر ما في الشعبة التي يدرسها المدرس (أ) أفضل من أداء طالب آخر حائز على الدرجة 80 في نفس الاختبار ولكنه في شعبة أخرى يدرسها المدرس (ب). من هنا، نرى أنه من الضروري إيجاد قيم لا تعتمد على الوحدات ويمكن استخدامها لمقارنة البيانات في المجموعات المختلفة. هذه القيم التي لا تعتمد على الوحدات نسميها بالقيم (أو الدرجات) المعيارية.

لتكن x_1, x_2, \dots, x_n عينه من البيانات حجمها n ومتوسطها \bar{x} وانحرافها المعياري s . نعرف الدرجة (القيمة) المعيارية للملاحظة x_i بالصيغة التالية:

$$z_i = \frac{x_i - \bar{x}}{s}; i = 1, 2, \dots, n \quad (11)$$

أي أن الدرجات المعيارية للبيانات x_1, x_2, \dots, x_n هي:

$$z_1 = \frac{x_1 - \bar{x}}{s}, z_2 = \frac{x_2 - \bar{x}}{s}, \dots, z_n = \frac{x_n - \bar{x}}{s}$$

ملاحظة :

1. القيمة $z_i = \frac{x_i - \bar{x}}{s}$ هي الدرجة المعيارية للملاحظة الأصلية x_i .
2. الملاحظة الأصلية للدرجة المعيارية z_i هي $x_i = \bar{x} + S z_i$.
3. الدرجات المعيارية هي قيم عديمة الوحدة ولذلك فإنها تستخدم للمقارنة بين المشاهدات المختلفة في المجموعات المختلفة للبيانات.
4. متوسط الدرجات المعيارية = 0.
5. الانحراف المعياري للدرجات المعيارية يساوي 1.

مثال (34): إذا كان لدينا مجموعة من البيانات متوسطها $\bar{x} = 7$ وانحرافها المعياري $s = 5$ فأوجد:

1. الدرجة المعيارية للقيمة $x = 9$.
2. القيمة الأصلية للدرجة المعيارية $z = 0.1$.

الحل:

1. الدرجة المعيارية للقيمة $x = 9$ هي:

$$z = \frac{x - \bar{x}}{s} = \frac{9 - 7}{5} = \frac{2}{5} = 0.4$$

2. القيمة الأصلية للدرجة المعيارية $z = 0.1$ هي:

$$x = \bar{x} + s z = 7 + 5 \times 0.1 = 7 + 0.5 = 7.5$$

تمرين واجب (5): إذا كانت درجة أحد الطلاب في مقرر الإحصاء تساوي 82 ودرجته في مقرر الرياضيات تساوي 89، وإذا كان متوسط درجات الطلاب في مقرر الإحصاء يساوي 75 بانحراف معياري يساوي 10 ومتوسط درجات الطلاب في مقرر الرياضيات يساوي 81 بانحراف معياري يساوي 16، ففي أي المقرر كان أداء الطالب أفضل؟

الاسبوع الثالث عشر والرابع عشر

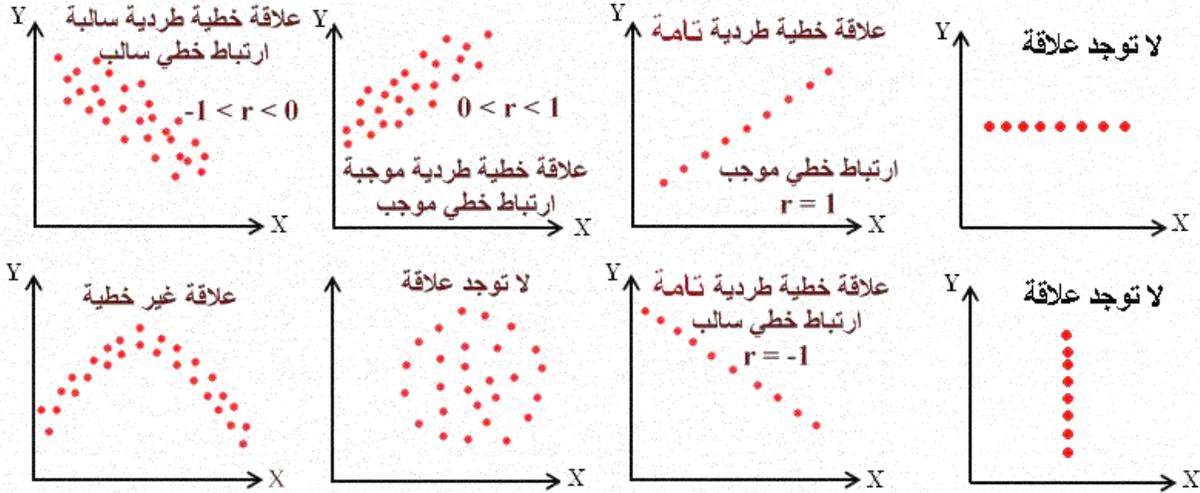
الارتباط البسيط، مفهومه، طرق حسابه للبيانات غير المبوبة (الطريقة المطولة والطريقة المختصرة)

الارتباط: Correlation

الارتباط : هو العلاقة بين ظاهرتين مثل العلاقة بين طول الشخص (سم) ووزنه (كغم) او العلاقة بين نسبة الشفاء من مرض معين وكمية الجرعة من الدواء المخصص للمريض وعمر المريض ، او العلاقة بين الدخل والاستهلاك والعلاقة بين درجات الطلبة وعدد ساعات الدراسة .

معامل الارتباط الخطي البسيط : هو المقياس الذي نقيس به درجة الارتباط.

شكل الانتشار يربط العلاقة بين المتغيرين وتستطيع من النقط أن تتعرف على العلاقة بين المتغيرين:



معامل الارتباط هو مؤشر هذه العلاقة

أول خطوه في تحديد طبيعة العلاقة هي رسم شكل الانتشار

إذا كان لدينا متغيران فقط . المتغير X وهو متغير يتم تحديده من قبل الباحث أو الشخص الذي يقوم

بالدراسة وهو يسمى بالمتغير المستقل Independent variable

يرافق المتغير X متغير آخر Y ويسمى بالمتغير التابع dependent variable وهو متغير عشوائي

لأن نتيجته غير محددة وتعتمد على قيم المتغير المستقل

معامل الارتباط : يعرف معامل الارتباط والذي يرمز له بالرمز r بأنه عبارة عن مقياس رقمي يقيس قوة

الارتباط بين متغيرين , حيث تتراوح قيمته بين (+1) و (-1) .

إذا كان الارتباط = صفر يعني لا يوجد ارتباط وإذا كان الارتباط = (1) موجب او سالب (يعني ارتباط تام .

ويرمز للارتباط بالرمز r x y

خصائص معامل الارتباط

1- معامل الارتباط مقياس وصفي

2- تتراوح قيمة معامل الارتباط بين -1 و +1.

3- معامل الارتباط يتأثر بالقيم الشاذة.

4- إذا كانت قيمة معامل الارتباط قريبة من الصفر فهذا دليل على عدم وجود علاقة خطية بين المتغيرين. أما إذا

كانت قيمة المعامل واحد صحيح فهذا دليل على أن العلاقة عكسية تامة ، أما إذا كانت قيمة معامل الارتباط

عند الواحد الصحيح الموجب فهذا يدل على العلاقة الموجبة الطردية التامة، وفيما عدا ذلك فإن العلاقة

توصف قوية أو متوسطة أو ضعيفة حسب الجدول التالي:

0.20	ضعيفة جدا
0.20 - أقل من 0.40	ضعيفة
0.40 - أقل من 0.60	متوسطة
0.60 - أقل من 0.80	قوية
0.80 - أقل من 1.00	قوية جدا
1.00	تام

عيوب معامل الارتباط

- 1- مقياس وصفي يقف عند حدود وصف العلاقة بين الظاهرتين ولا يسمح بالتنبؤ بقيمة أحد المتغيرين بمعلومية الآخر.
 - 2- لا يوضح العلاقة السببية بين المتغيرين أي أنه لا يميز بين المتغير المستقل والمتغير التابع
 - 3- لا يفرق بين العلاقة الحقيقية والعلاقة الناشئة من الصدفة.
- مثلا العلاقة بين درجة الاختبار ورقم السجل المدني تساوي 0.55

الارتباط لبيانات غير مبنوية معامل ارتباط بيرسون

1. الطريقة المختصرة (طريقة انحرافات القيم عن الوسط الحسابي) وذلك باستخدام الصيغة الآتية:

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

حيث ان r_{xy} تعني الارتباط بين X و y

X_i قيم مشاهدات المجموعة الاولى و y_i قيم مشاهدات المجموعة الثانية

حيث ان r_{xy} تعني الارتباط بين X و y

X_i قيم مشاهدات المجموعة الاولى و y_i قيم مشاهدات المجموعة الثانية

\bar{X} الوسط الحسابي للمجموعة الاولى وان \bar{y} الوسط الحسابي للمجموعة الثانية

مثال (35) احسب معامل الارتباط بين المتغيرين (X) و (y)

قيم, X : 2, 2, 5, 4, 5, 6, 3, 5, 4

قيم, y : 3, 5, 7, 8, 9, 11, 6, 8, 6

الحل : 1- نجد الوسط الحسابي لقيم X - والوسط الحسابي لقيم Y

2 نجد انحرافات القيم عن الوسط الحسابي ل \bar{X} - و ل \bar{Y}

3 نربع انحرافات القيم عن الوسط الحسابي ل X - و ل Y الوسط الحسابي للمجموعة الاولى وان \bar{y} الوسط الحسابي للمجموعة الثانية

$$\bar{x} = \frac{\sum xi}{.N} = \frac{36}{9} = 4$$

$$\bar{y} = \frac{\sum yi}{.N} = \frac{63}{9} = 7$$

Yi	Xi	(Yi - \bar{y})	(Xi - \bar{X})	(Yi - \bar{y}) ²	(Xi - \bar{X}) ²	(Yi - \bar{y}) (Xi - \bar{X})
3	2	-4	2-4 = -2	16	4	8
5	2	-2	-2	4	4	4
7	5	0	1	0	1	0
8	4	1	0	1	0	0
9	5	2	1	4	1	2
11	6	4	2	16	4	8
6	3	-1	-1	1	1	1
8	5	1	1	1	1	1
6	4	-1	0	1	0	0
—	—	—	—	—	—	—
Σ63	Σ36		0	Σ44	Σ16	Σ24

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

$$r_{xy} = \frac{24}{\sqrt{\sum(16)(44)}} = 0.905$$

الارتباط موجب (قوي جداً)

2- الطريقة المطولة باستخدام القيم الاصلية وحسب الصيغة الاتية:

$$r = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{\{n \sum x_i^2 - (\sum x_i)^2\} \{n \sum y_i^2 - (\sum y_i)^2\}}}$$

مثال (36) : البيانات الآتية تمثل الدخل (X) والاستهلاك (y) بالدولار الأمريكي

المطلوب : حساب العلاقة بين الدخل والاستهلاك

X : 200 , 300 , 400 , 600 , 900

Y : 180 , 270 , 320 , 480 , 700

الحل : 1 نجد مجموع - X 2 نجد مجموع - 3 y نضرب - X في 4 y نجد مجموع - x2 ومجموع y2

xi	Yi	xiyi	Xi ²	yi ²
200	180	36000	40000	32400
300	270	81000	90000	72900
400	320	128000	160000	102400
600	480	288000	360000	230400
900	700	630000	810000	490000
Σ 2400	Σ1950	Σ1163000	Σ1460000	Σ928100

$$r = \frac{n \sum xiyi - \sum xi \sum yi}{\sqrt{\{n \sum xi^2 - (\sum xi)^2\} \{n \sum yi^2 - (\sum yi)^2\}}}$$

$$r = \frac{5(1163000) - (2400)(1950)}{\sqrt{\{5 \sum 1460000 - (2400)^2\} \{5 \times 928100 - (1950)^2\}}}$$

$$r = \frac{5815000 - 4680000}{\sqrt{\{7300000 - 5760000\} \{4640500 - 3802500\}}}$$

$$r = \frac{1135000}{\sqrt{\{1540000\} \{838000\}}} = \frac{1135000}{1136010.5} = 0.999$$

العلاقة بين الدخل والاستهلاك طردية عالية جدا وقوية موجبة وقرابية من الواحد

مثال (37) : أوجد معامل ارتباط بيرسون بين المتغيرين (درجات الحرارة xi) و (مستوى الأداء في الامتحان yi) بالطريقتين (المختصرة والمطولة)

xi	yi
0	3
2	2
3	4
4	0
6	6
15	15

الحل / بالمختصرة

x	y	$(x - \bar{x})$	$(y - \bar{y})$	$(x - \bar{x})^2$	$(y - \bar{y})^2$	$(x - \bar{x})(y - \bar{y})$
0	3	-3	0	9	0	0
2	2	-1	-1	1	1	1
3	4	0	1	0	1	0
4	0	1	-3	1	9	-3
6	6	3	3	9	9	9
15	15	0	0	20	20	7

$$\bar{x} = 3 \quad \bar{y} = 3$$

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} = \frac{7}{\sqrt{20 \times 20}} = \frac{7}{20} = .35$$

الطريقة المطولة

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0	3	0	0	9
2	2	4	4	4
3	4	12	9	16
4	0	0	16	0
6	6	36	36	36
15	15	52	65	65

$$r_p = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{\{n \sum x_i^2 - (\sum x_i)^2\} \{n \sum y_i^2 - (\sum y_i)^2\}}}$$

$$r = \frac{5(52) - (15)(15)}{\sqrt{\{5 \sum 65 - (15)^2\} \{5 \times 65 - (15)^2\}}} = \frac{260 - 225}{\sqrt{(325)(225)}} = \frac{35}{\sqrt{(100)(100)}} = \frac{35}{\sqrt{100}} = 0.35$$

الارتباط طردي ضعيف

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
تمرين واجب (6) : احسب معامل الارتباط بيرسون للمتغيرين y و X باستخدام الطريقتين المطولة

X 2 , 2 , 5 , 4 , 5 , 6 , 3 , 5 , 4 ,

y 3 , 5 , 7 , 8 , 9 , 11 , 6 , 8 , 6

الاسبوع الخامس عشر والسادس عشر :ارتباط الرتب (سبيرمان وسبيرمان المعدل)

ان الصيغ السابقة الخاصة لحساب معامل الارتباط البسيط تستند بالحقيقة على ان المتغيرات المعتمدة في الحساب هي متغيرات من النوع الكمي . الا انه من الناحية العملية هناك الكثير من الحالات التي تكون فيها المتغيرات من النوع الوصفي (اي متغيرات غير قابلة للقياس كالمهنة ، الحالة الاجتماعية ، تقديرات درجات وغيرها). وبهدف حساب الارتباط بين متغيرات من هذا النوع فانه لا يمكن استخدام الصيغ السابقة مباشرة لهذا الغرض بل اجراء بعض التعديلات عليها بالشكل الذي يجعلها ممكنة الاستخدام . ان المعامل الذي يقيس درجة الترابط ما بين صفتين يدعى بمعامل ارتباط الرتب لسبيرمان.

معامل ارتباط الرتب : هو الذي يمثل درجات الارتباط بين المتغيرات من النوع الوصفي (متغيرات وصفية)

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

حيث ان $di = Xi - yi$ اي انحراف قيم الرتب Xi عن قيم الرتب yi ويعني هذا ان معامل الارتباط البسيط يساوي

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

مثال (38) : ادناه تقديرات 6 طلبة في امتحان الاقتصاد والاحصاء

المطلوب : حساب العلاقة بين تقديرات المادتين

تقديرات (الاقتصاد X):	ضعيف ، امتياز ، جيد ، متوسط ، مقبول ، جيد جدا
تقديرات:(الاحصاء y):	مقبول ، جيد جدا ، جيد ، ضعيف ، متوسط ، امتياز

خطوات الحل :

1- تحويل التقديرات الى ارقام نعطي لهذه التقديرات ارقام متسلسلة ونرتبها اما تصاعديا او تنازليا وكالاتي:

ضعيف	مقبول	متوسط	جيد	جيد جدا	امتياز
1	2	3	4	5	6

2- نرتب تقديرات X وتقديرات y ترتيبا تصاعديا او ترتيبا تنازليا

3- نجد di الفرق بين الترتيب المتناظرة من العلاقة $(Xi - yi)$

4- نجد تربيع di ثم نجمع di^2

5- نطبق صيغة القانون

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

ترتيب التقديرات	تسلسل الرتب	ترتيب X	ترتيب y	$d_i=(X_i - y_i)$	d_i^2
ضعيف	1	1	2	-1	1
مقبول	2	6	5	1	1
متوسط	3	4	4	0	0
جيد	4	3	1	2	4
جيد جدا	5	2	3	-1	1
امتياز	6	5	6	-1	1
					$\Sigma=8$

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

$$r_s = 1 - \frac{6(8)}{6(36-1)} = 1 - \frac{48}{210} = 1 - 0.228 = 0.775 \text{ الارتباط طردي قوي}$$

مثال (39) أوجد معامل ارتباط سبيرمان بين درجات الطلاب في مادتي الاحصاء و الادارة:

الاحصاء	جيد	متوسط	ضعيف	مقبول	جيد جدا	امتياز
الادارة	متوسط	جيد	مقبول	ضعيف	امتياز	جيد جدا

X	y	رتب x	رتب y	الفرق بين الرتبتين d	d^2
جيد	متوسط	4	3	1	1
متوسط	جيد	3	4	-1	1
ضعيف	مقبول	1	2	-1	1
مقبول	ضعيف	2	1	1	1
جيد جدا	امتياز	5	6	-1	1
امتياز	جيد جدا	6	5	1	1
					$6 \sum d^2 =$

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} = 1 - \frac{6(6)}{6(36 - 1)} = 1 - \frac{36}{210} = 1 - 0.11 = 0.829$$

ارتباط قوي طردي

امتياز	جيد جدا	جيد	متوسط	مقبول	ضعيف
6	5	4	3	2	1

بالنسبة لرتب x - y

تمرين واجب (7) : الاتي تقديرات 10 طلبة في مادتي القراءات المحاسبية و مبادئ المحاسبية

المطلوب : ايجاد معامل الارتباط بين تقديرات الطلبة

القراءات المحاسبية X : متوسط - جيد - مقبول - ضعيف - امتياز - جيد - جيد-مقبول - ضعيف- متوسط
مبادئ المحاسبية y : جيد-متوسط- ضعيف- مقبول-جيد جدا - امتياز - جيد جداً - متوسط - مقبول- جيد

ارتباط سبيرمان المعدل

في حالة تكرار بعض قيم احد المتغيرين او كليهما عندئذ لا يمكن استخدام صيغة معامل ارتباط الرتب لسبيرمان بل يستوجب اجراء بعض التعديلات عليها وعلى النحو الاتي:

بعد ترتيب قيم المتغير (الصفات) بشكل تصاعدي او تنازلي يتم تخصيص قيم سلسلة الاعداد الطبيعية كرتب لهذه الصفات ، ومن ثم يتم حساب معدل القيم المخصصة واعادة تخصيصه لتلك الصفات المكررة ، بعد ذلك يتم تعديل

$$\text{من خلال اضافة الكمية } \frac{m(m^2-1)}{12} \text{ الى } \sum d^2$$

، حيث m تمثل عدد مرات تكرار الصفة ، هذه الكمية تضاف الى $\sum d^2$ مقابل كل صفة مكررة . هذا الاجراء يتم لكلا المتغيرين X و y .

مثال (40): الاتي تقديرات لمشروع تخرج عشرة من طلبة قسم المحاسبة وقسم السياحة من حيث كتابتهم ومناقشتهم لمشروع البحث .

تقديرات طلبة المحاسبة X : جيد ، متوسط ، جيد جدا ، متوسط ، امتياز ، ضعيف ، مقبول ، متوسط ، جيد ، جيد .

تقديرات طلبة السياحة y : متوسط ، جيد ، جيد ، مقبول ، جيد جدا ، مقبول ، متوسط ، متوسط ، متوسط ، امتياز .

المطلوب : حساب معامل الارتباط البسيط

الحل:

- 1- نرتب تقديرات المتغير X تصاعديا او تنازليا ، ثم نضع لها ارقام متسلسلة
- 2- نجد معدل الرتب (رتب x)
- 3- نرتب تقديرات المتغير الاخر y اما تصاعديا او تنازليا ونضع لها ارقام متسلسلة
- 4- نجد معدل رتب y
- 5- نجد d_i من العلاقة $(X_i - y_i)$
- 6- نربع d_i ثم نجد مجموع d_i^2
- 7- نحسب عدد التكرارات للتقديرات المتكررة لكل متغير X و y عن طريق احتساب التعديلات وذلك باضافة الكمية $\frac{m(m^2-1)}{12}$
- 8- نجمع التعديلات الكلية للمتغيرين معا
- 9- نحسب معامل ارتباط الرتب لسبيرمان المعدل من الصيغة الاتية:

$$r_{xy} = 1 - \frac{6(\sum d^2 + m)}{n(n^2 - 1)}$$

تقديرات X	ترتيب X تصاعديا	معدل X رتب	تقديرات y	ترتيب y تصاعديا	معدل y رتب	di Xi-yi	di ²
جيد	1 ضعيف	7	متوسط	1 ضعيف	5	2	4
متوسط	2 مقبول	4	جيد	2 مقبول	7.5	-3.5	12.25
جيد جدا	3 متوسط	9	جيد	3 مقبول	7.5	1.5	2.25
متوسط	4 متوسط	4	مقبول	4 متوسط	2.5	1.5	2.25
امتياز	5 متوسط	10	جيد جدا	5 متوسط	9	1	1
ضعيف	6 جيد	1	مقبول	6 متوسط	2.5	-1.5	2.25
مقبول	7 جيد	2	ضعيف	7 جيد	1	1	1
متوسط	8 جيد	4	متوسط	8 جيد	5	-1	1
جيد	9 جيد جدا	7	متوسط	9 جيد جدا	5	2	4
جيد	10 امتياز	7	امتياز	10 امتياز	10	-3	9
						$\Sigma=39$	

نحسب عدد التكرارات وكالاتي :

تكرارات المتغير X :

$$\frac{m(m^2-1)}{12} = \frac{3(3^2-1)}{12} = \frac{24}{12} = 2 \quad - \quad m=3 \quad \text{لقد تكرر التقدير المتوسط 3 مرات فيعني ان } m=3$$

$$\frac{m(m^2-1)}{12} = \frac{3(3^2-1)}{12} = \frac{24}{12} = 2 \quad - \quad \text{عدد تكرارات التقدير جيد هو 3 فأذن}$$

تكرارات المتغير y :

$$\frac{m(m^2-1)}{12} = \frac{3(2^2-1)}{12} = \frac{6}{12} = 0.5 \quad - \quad m=2 \quad \text{التقدير مقبول تكرر مرتين وعليه فان } m=2$$

$$\frac{m(m^2-1)}{12} = \frac{3(3^2-1)}{12} = \frac{24}{12} = 2 \quad - \quad m=3 \quad \text{المتوسط تكرر 3 مرات فأذن } m=3$$

$$\frac{m(m^2-1)}{12} = \frac{3(2^2-1)}{12} = \frac{6}{12} = 0.5 \quad - \quad m=2 \quad \text{التقدير جيد تكرر مرتين فان } m=2$$

وعليه فان التعديل الكلي $2+2+0.5+2+0.5 = 7$ وبذلك معامل ارتباط الرتب لسبيرمان المعدل يكون

$$r_{xy} = 1 - \frac{6(\Sigma d^2 + m)}{n(n^2 - 1)} = 1 - \frac{6(39 + 7)}{10(10^2 - 1)} = 1 - \frac{6(46)}{990}$$

$$= 1 - \frac{267}{990} = 1 - 0.2787 = 0.721$$

الارتباط طردي موجب

لقياس ارتباط الصفات توجد عدة مقاييس 1

معامل الاقتران **AC** : تسمى العلاقة بين ظاهرتين تتصف كل منهما بصفتين فقط بالاقتران , فمثلا ظاهرة

الجنس- (ذكر , انثى) , الحرفة (زراعة , صناعة) .

فلو كان a يمثل عدد الاشخاص الذكور الذين يعملون في الزراعة , و b يمثل عدد الاناث اللاتي يشتغلن في الزراعة و c يمثل عدد الذكور الذين يعملون في الصناعة , و d عدد الاناث اللاتي يعملن في الصناعة فمن الممكن وضع هذه المعلومات في جدول مزدوج وكالاتي:-

	صناعة	زراعة	الحرفة الجنس
ذكر	c	a	
انثى	d	b	

عندئذ يكون معامل الاقتران بين الظاهرتين هو

$$AC = \frac{ad-bc}{ad+bc}$$

وقيمة معامل الاقتران كما هي الحال في معامل الارتباط البسيط بين الصفر والواحد بالسالب او الموجب .

مثال (41): أوجد معامل الاقتران بين الحقن باللقاح و المناعة ضد كورونا

	أصيب بكورونا	لم يصب بكورونا	الاصابة الوقاية
حقن باللقاح	4	192	
لم يحقن باللقاح	34	113	

$$Ac = \frac{(192)(34) - (113)(4)}{(192)(34) + (113)(4)} = 0.087$$

الاقتران طردي قوي

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
مثال (42) : من المعلوم ان عادة التدخين تؤثر تأثيراً سيئاً على الصحة العامة للفرد

المطلوب : ايجاد العلاقة بين الحالة الصحية وعادة التدخين

لا يدخن	يدخن	الحالة الصحية عادة التدخين
50	40	جيدة
60	50	غير جيدة

$$Ac = \frac{(40)(60) - (50)(50)}{(40)(60) + (50)(50)} = 0.02 \text{ - اقتران عكسي}$$

تمرين واجب (8) احسب معامل الاقتران

غير سعيدة	سعيدة	سعادة ثقافة
19	62	متقف
34	10	غير متقف



الأسبوع التاسع عشر : معامل التوافق Coefficient of contingency

درسنا سابقا معامل الاقتران الذي يكون مقصورا على الظواهر التي تنقسم المجموعتين ، اما اذا كانت احد الظاهرتين اللتين نبحث العلاقة بينهما او كليهما تنقسم الى اكثر من نوعين ، فاذا انقسمت كل من الظاهرتين الى اكثر من صيغتين فلا يمكن قياس العلاقة بينهما باستخدام معامل الاقتران وانما بمقياس من نوع اخر يدعى بمعامل التوافق ويستخرج وفق الصيغة الآتية حيث (c) تمثل المجموع الكلي لخارج قسمة مربع التكرار في كل خانة على حاصل ضرب مجموع التكرارين العمودي والافقي للخانة التي منها التكرار

$$CC = \sqrt{\frac{r-1}{r}}$$

المهنة الجنس	طيار	ضابط	محاسب	المجموع Tv
ذكر	2	3	4	9
انثى	1	5	7	13
المجموع TH	3	8	11	22

$$R1 \text{ (ذكر)} = \frac{2^2}{T.1T.1} + \frac{3^2}{T.1T.2} + \frac{4^2}{T.1T.3}$$

$$R1 \text{ (ذكر)} = \frac{2^2}{(9)(3)} + \frac{3^2}{(9)(8)} + \frac{4^2}{(9)(11)} = 0.1481 + 0.125 + 0.1616 = 0.4347$$

خطوات الحل

- 1- في العمود الأول يمثل النوع الاجتماعي (الجنس) الظواهر R وكل نوع نعطيه رقم بجانب R مثلاً الذكور R1 والاناث R2 .. الخ حسب عدد الظواهر المراد قياس معامل التوافق لها.
- 2- نربع رقم الخلية الذي تتقاطع الظاهرة مع الأخرى مثلاً (ذكر طيار) قد يكون الرقم مثلاً 2 بالبسط ثم نقسمه على (المقام) مجموع الذكور في جميع المهن (9) مضروب في مجموع كل مهنة (المجموع العمودي لمهنة الطيار) 3 ، ونستمر R2 و R3 ... الخ.
- 3- بعد ان استخرجنا (الخ R1 + R2 + R3 + ...) نقوم بجمع الناتج.
- 4- نطبق القانون ادناه

$$cc = \sqrt{\frac{r-1}{r}}$$

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
مثال (43) : الجدول الاتي يبين عدد حوادث الطرق التي تعرضت لها شاحنات لنقل البضائع خلال

فترة زمنية موزعة حسب نوع الحادث وحالة الجو

المطلوب : ايجاد معامل التوافق لهذا التوزيع

نوع الحادث حالة الجو	دهس	اصطدام	انقلاب	المجموع
صحو	25	12	9	46
ممطر	10	50	35	95
ضباب	20	45	40	105
المجموع	55	107	84	246

الحل: نستخرج R : نفترض ان : ضباب R3 ، ممطر R2 ، صحو R1

$$R1 = \frac{25^2}{(46)(55)} + \frac{12^2}{(46)(107)} + \frac{9^2}{(46)(84)} = \frac{625}{2530} + \frac{144}{4922} + \frac{81}{3864} = 0.296$$

$$R2 = \frac{10^2}{(95)(55)} + \frac{50^2}{(95)(107)} + \frac{35^2}{(95)(84)} = 0.419$$

$$R3 = \frac{20^2}{(105)(55)} + \frac{45^2}{(105)(107)} + \frac{40^2}{(105)(84)} = 0.431$$

$$R = R1 + R2 + R3$$

$$= 0.296 + 0.419 + 0.431 = 1.146$$

$$cc = \sqrt{\frac{1.146-1}{1.146}} = 0.358$$



تمرين واجب (9) : كانت نتائج مجموعة من الطلبة في الامتحانات النهائية كما في الجدول الاتي المطلوب : تقدير العلاقة بين تحصيل الطالب والالتزام بحضور المحاضرات

الحضور النتائج	جيد	متوسط	ضعيف	المجموع
الناجحين	8	5	2	15
المكملين	3	3	4	10
الراسبين	2	4	10	16
المجموع	13	12	16	41

الاسبوع (20-21-22) (السلاسل الزمنية ، مفهومها واستخداماتها)

النظرة الشاملة :

- **الفئة المستهدفة :** لطلبة المرحلة الاولى / قسم تقنيات المحاسبية / معهد التقني كربلاء
- **المبررات :** صمم هذا الفصل لتمكن الطالب من التعرف على مفهوم السلاسل الزمنية واستخدامها اضافة الى تحليل السلسلة الى مركباتها الاربعه.
- **الفكرة المركزية**



- ✓ مفهوم السلسلة الزمنية.
- ✓ تحليل السلسلة الزمنية الى مركباتها.
- ✓ طريقة ايجاد خط الاتجاه العام.
- ✓ طريقة متوسطي نصفي السلسلة.
- ✓ طريقة المعدلات (الايوساط المتحركة).
- ✓ طريقة المربعات الصغرى.

الاهداف الادائية- : سيكون الطالب بعد انتهائه من دراسة هذا الفصل قادرا على

- يعرف السلسلة الزمنية.
- يعرف مركبات السلسلة الزمنية.
- يميز بين الطرق المختلفة لتقدير معادلة خط الاتجاه العام.

Time Senes السلاسل الزمنية

السلسلة الزمنية هي مجموعة القياسات المسجلة لمتغير واحد أو أكثر مرتبة حسب زمن وقوعها.

او هي مجموعة من القيم المشاهدة لظاهرة معينة لمدة من الزمن (عدة سنوات) في فترات زمنية متساوية . من الأمور الطبيعية والواجبة للحكومات والمؤسسات والشركات التجارية منها والصناعية والتعليمية وغيرها بالتخطيط لمستقبلها لتحقيق الأهداف الخاصة والعامة وتقديم كافة الخدمات والوصول لحالة العدل والاستقرار للمجتمع والعمل على اتخاذ قرارات التنبؤ بوقوع الأحداث قبل وقوعها في كافة أوجه النشاط التي تخص المجتمع، وتعتبر السلاسل الزمنية من أهم أساليب التنبؤ حول المستقبل من خلال وقائع الأمس واليوم. من أهم السلاسل الزمنية تلك الخاصة بالمؤشرات الاقتصادية والمبيعات السنوية للشركات بكافة أوجه نشاطاتها والتعليم وحجم السكان وما شابه ذلك. والتغير الذي يحدث في قيم متغير السلسلة الزمنية أو قيم متغيراتها يعتبر دالة في الزمن يمكن تمثيلها بيانياً باتخاذ المحور الأفقي للزمن والرأسي لقيم المتغير ومن الامثلة عليها:

- المبيعات اليومية لمحل تجاري لمدة شهر - .
- قراءة درجة حرارة المريض كل ساعة خلال يوم واحد- .
- صادرات العراق من النفط الخام ولعدة سنوات- .
- اعداد الطلبة المقبولين في معهد الادارة التقني لعدة سنوات- .
- كمية الامطار الهاطلة خلال اشهر السنة- .

كل هذه القراءات وتتابعها الزمني يمثل سلسلة زمنية , ويمكن اعتبار السلسلة الزمنية علاقة دالية بين متغيرين هما قيمة الظاهرة والزمن $y = f(x)$ حيث y يشير الى قيمة الظاهرة x يشير الى الفترة الزمنية

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري وتهتم كثير من الدراسات ولاسيما الاقتصادية والاجتماعية بدراسة السلسلة وذلك لان كثير من الظواهر الاقتصادية كالصادرات والواردات اذا استعرضت لعدة سنوات تمكننا من معرفة التغيرات التي تطرأ على قيمة الظاهرة في المستقبل على ضوء ما حدث في الماضي ويمكن تمثيل السلسلة الزمنية بيانيا لان كل قراءة تمثل زوجا من النقاط

عندما نحصل على قيم مشاهدات السلسلة الزمنية لآبد من دراسة المؤثرات التي تؤثر على هذه القراءات وهذه المؤثرات تسمى بمركبات السلسلة الزمنية . والتي ناتج حاصل ضربها يعطي قيم المشاهدة الاصلية حيث Y هي قيمة المشاهدة الاصلية $Y=G \times S \times P \times E$

G مركبة الاتجاه العام

S التغيرات الموسمية

P التغيرات الدورية

E التغيرات العرضية

مركبة الاتجاه العام : وهي المركبة التي توضح مسيرة السلسلة بشكل عام وعلى مدى بعيد ويمكن استخراجها عن طريق معادلة انحدار Y/X والمتمثل بالعلاقة $Y = a + bx$

ومن الواضح ان قيمة y تزداد او تتناقص من خلال قيمة a, b وقد تحافظ على قيمتها وهناك عدة طرق لإيجاد هذه المركبة :

1- طريقة الانتشار .

2- طريقة متوسطي نصفي السلسلة .

3- طريقة الاوساط المتحركة .

1- طريقة الانتشار : يتم تمثيل قيم المشاهدات في سلسلة زمنية وايجاد مركبة الاتجاه العام عن طريق رسم انتشاري -وايجاد معادلة خط الاتجاه العام . وهذه الطريقة تختلف من شخص الى اخر.

2- طريقة متوسطي نصف السلسلة- :

تعد هذه الطريقة اكثر دقة من طريقة الانتشار وتتلخص خطواتها بالاتي:

- تقسم السلسلة الى قسمين متساويين- .
- نجد المتوسط الحسابي لكل قسم- .
- نصل بين النقطتين (المتوسطين) بعد تعيينهما على المستوى الاحداثي فتكون لدينا خط الاتجاه العام .

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
مثال (44): فيما يلي انتاج مصنع الالبسة الصوفية لعشرة سنوات بالألاف الاطنان المطلوب ايجاد معادلة خط الاتجاه العام بطريقة خط الاتجاه العام بطريقة متوسطي نصفي السلسلة

(السنة) x(عدد القطع المنتجة (y)	
2010	53	
2011	64	
2012	67	←
2013	60	
2014	69	
2015	74	
2016	67	
2017	79	←
2018	85	
2019	90	

$$\bar{x} = \frac{\sum xi}{n} = \frac{53+64+67+60+69}{5} = 62.6$$

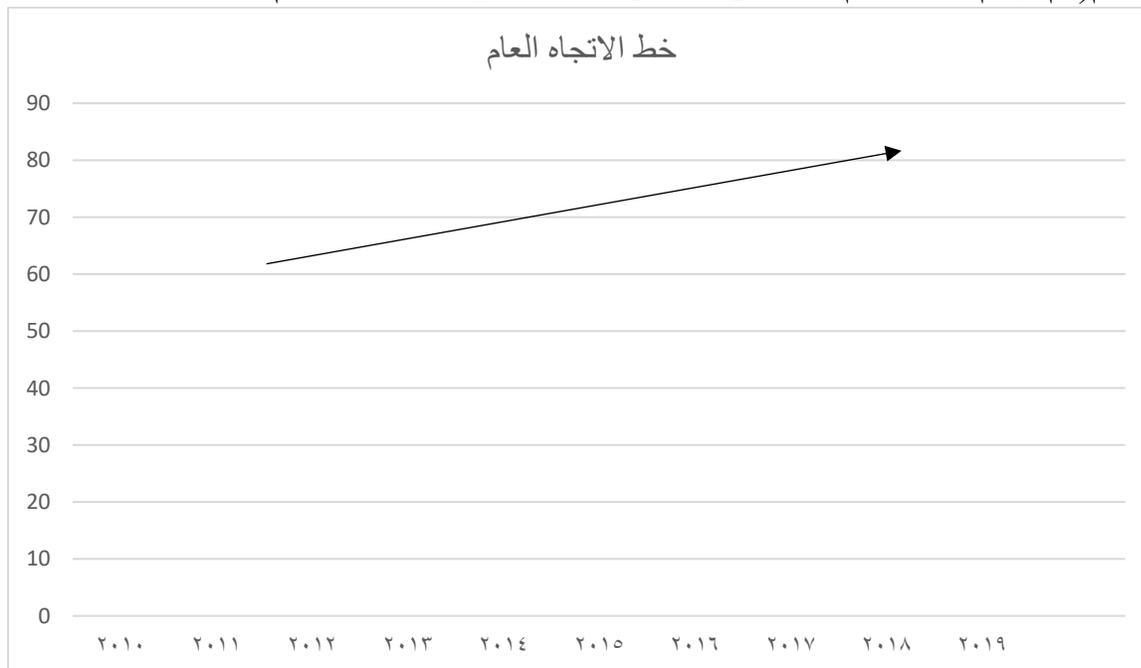
$$\bar{x} = \frac{\sum xi}{n} = \frac{74+67+79+85+90}{5} = 79$$

اذن النقطتين هما A(3,62.6)

ان الرقم الذي بجانب الوسط الحسابي 62.6 لنقطة A وهو رقم 3 (على المحور الافقي) يمثل نصف المجموعة المقسمة 5 من السلسلة 10.

B(8,79) : ان الرقم الذي بجانب الوسط الحسابي 79 لنقطة B وهو رقم 8 (المحور الافقي) يمثل نصف المجموعة المقسمة الثانية من السلسلة 10.

نرسم (المحور الافقي) ويمثل السنوات و(المحور العمودي) ويمثل الكميات بالألاف (ونحدد النقطتين) الوسط الحسابي لكل قسم) ثم نرسم خط مستقيم يصل بين النقطتين فنحصل على خط الاتجاه العام



تمرين (واجب) (9) : اوجد خط التجاه العام للسلسلة الزمنية الاتية باستخدام طريقة متوسطي نصفي السلسلة

السنوات : 2000 , 2001,2002 , 2003 , 2004 , 2005 , 2006 , 2007 , 2008 , 2009

الإنتاج (بالملايين): 5.93 , 6.09, 6.26 , 6.45 , 6.67 , 6.89 , 7.13 , 7.37 , 7.62 , 7.88

طريقة الاوساط المتحركة: method of moving averages

بموجب هذه الطريقة نختار عدد من السنين ولتكن ثلاثة سنوات او خمس سنوات ونحسب متوسط قيم الظاهرة لهذه السنين ثم نترك القيمة الاولى من القيم التي اخذناها ونأخذ بدلها القيمة التالية للمجموعة السابق اخذها فنحصل على قيم بنفس عدد القيم السابقة ونأخذ متوسطها ثم نترك القيمة الاولى من المجموعة الجديدة ونأخذ بدلها القيمة التالية للمجموعة ثم نستخرج متوسطها وهكذا نحصل على المتوسطات المتحركة لقيم الظاهرة ثم نضع المتوسط لكل مجموعة امام القيمة الوسطى من قيمها ان كان عددها فرديا وامام احدى القيمتين الوسطيتين ان كان عددا زوجيا ثم نثبت هذا المتوسط على شكل نقاط ونصل بينها بخط مستقيم فيكون خط الاتجاه العام.

مثال (45) : الجدول الاتي يبين قيم المبيعات الفعلية y_i لاحد مصانع السمنت وللسنوات من 1990 إلى 1998

المطلوب : رسم خط الاتجاه العام باستخدام طريقة المتوسطات المتحركة

السنوات : 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998

المبيعات بالملايين : 210, 240, 230, 220, 260, 250, 230, 220, 280

الحل : نفرض ان عدد السنين المناسب هو ثلاث سنوات

متوسط ثلاث سنوات (المتوسطات المتحركة)	مجموع ثلاث سنوات	المبيعات y_i بالملايين	السنوات
		210	1990
226.67	680	240	1991
230	690	230	1992
236.67	710	220	1993
243.33	730	260	1994
246.67	740	250	1995
233.33	700	230	1996
243.33	730	220	1997
		280	1998

السنوات

الاسبوع الثالث والعشرين الارقام القياسية /مفهومها ، واستخدامها

النظرة الشاملة

الفئة المستهدفة / طلبة المرحلة الاولى / قسم تقنيات قسم المحاسبة

المبررات : صممت هذا الفصل لتمكين الطالب من التعرف على مفهوم الارقام القياسية واستخداماتها اضافة الى التعرف على الانواع المختلفة للأرقام القياسية وكيفية قياسها.

الفكرة المركزية

- 1 . مفهوم الارقام القياسية
 - 2 . استخدامات الأرقام القياسية
 - 3 . انواع الارقام القياسية
- الارقام القياسية البسيطة
 - الارقام القياسية المرجحة
 - رقم لاسبيرز القياسي
 - رقم باش القياسي
 - رقم فيشر القياسي

الاهداف الادائية : سيكون الطالب بعد دراسته هذه الوحدة النمطية قادرا على ان

- يعرف الارقام القياسية
- يذكر استخدامات الارقام القياسية
- يحسب الارقام القياسية البسيطة
- يحسب الارقام القياسية المرجحة

مفهوم الرقم القياسي : الرقم القياسي هو اداة احصائية مصممة لبيان التغير في قيمة الظاهرة او مجموعة ظواهر مرتبطة مع بعضها بالنسبة الى قيمتها في زمن اخر او موقع جغرافي اخر.

او هو اداة لقياس التغير النسبي المؤي في ظاهرة ما او عدة ظواهر نظرا لاختلاف الزمان والمكان .

فترة الاساس : هي الفترة الزمنية التي نقيس فيها التغير في الظاهرة . وهذه الفترة قد تكون شهرا او سنة او غيرها ولكن الشائع ان تكون سنة واحدة وتسمى (سنة الاساس) كما يشترط ان تكون سنة طبيعية خالية من الشذوذ كالحروب والازمات او الكوارث.

فترة المقارنة : هي الفترة الزمنية التي حصل خلالها التغير في الظاهرة . او هي الفترة التي تنسب اسعارها او كمياتها الى اسعار او كميات فترة الاساس.

اما اذا اردنا مقارنة التغير في مكانين مختلفين فان المكان الذي نقيس فيه التغير يسمى مكان الاساس والمكان الذي حصل خلاله التغير يسمى مكان المقارنة.

- 1 -مقارنة اسعار السلع المختلفة.
- 2 -التنبؤ بأحوال الاعمال والاقتصاد.
- 3 -مقارنة المستوى التعليمي في بلد ما مع مستواه في نفس البلد في سنة أخرى.
- 4 -مقارنة عدد السكان في بلد وسنة ما مع عدد السكان في سنة أخرى.

خصائص سنة الاساس

- 1 -تحديد سنة الاساس بحيث لا تكون بعيدة عن سنة المقارنة.
- 2 -ان تكون سنة الاساس سنة طبيعية وذات هدوء نسبي من حيث انعكاساتها على الظاهرة المدروسة.

انواع الأرقام القياسية

هناك عدة انواع من الارقام القياسية

- 1 -الارقام القياسية البسيطة والتجميعية Simple index number
- 2 -الارقام القياسية المرجحة Weighted index number

الاسبوع الرابع والعشرين

اولا : **الارقام القياسية البسيطة Simple index number**

الارقام القياسية البسيطة

يحسب الرقم القياسي البسيط للأسعار بموجب هذه الطريقة بإيجاد الوسط الحسابي للأسعار خلال سنتي الاساس والمقارنة ثم ننسب الوسط الحسابي للأسعار لسنة المقارنة الى الوسط الحسابي لسنة الاساس ونضرب الناتج * 100 للحصول على نسبة مئوية.

مجموع الاسعار في سنة المقارنة

الوسط الحسابي للأسعار في سنة المقارنة

عدد الأسعار

$$\bar{p}_i = \frac{\sum p_i}{n}$$

p_i تمثل أسعار سنة المقارنة

مجموع الاسعار في سنة الاساس

الوسط الحسابي للأسعار في سنة الاساس

عدد الأسعار

$$\bar{p}_o = \frac{\sum p_o}{n}$$

p_o تمثل أسعار سنة الاساس

وعليه فإن الرقم القياسي يساوي = $\frac{\text{الوسط الحسابي للأسعار في سنة المقارنة}}{\text{الوسط الحسابي للأسعار في سنة الاساس}} \times 100$

$$p.i = \frac{\frac{\sum p_i}{n}}{\frac{\sum p_o}{n}} \times 100$$

وعليه فان الرقم القياسي للأسعار يأخذ الصيغة الآتية : بعد اختصار n من البسط والمقام

$$p.i = \frac{\sum p_i}{\sum p_o} \times 100$$

مثال (46): الجدول الاتي يبين اسعار بعض السلع في سنتي (2018)(2022) المطلوب : ايجاد الرقم القياسي البسيط للأسعار لسنة 2022 علما بان 2018 هي سنة الاساس

السلعة	السعر بالدينار	
	2022	2018
القمح	500	300
السكر	1000	700
اللحم	14000	10000
	15500	11000

$$p.i = \frac{\sum pi}{\sum po} \times 100$$

$$p.i = \frac{15500}{11000} \times 100 = 141\%$$

اي ان الاسعار ارتفعت بمقدار 1.41 مرة عما كانت عليه في سنة الاساس 2018

الاسبوع الخامس والعشرين والسادس والعشرين

حساب الارقام القياسية المرجحة ، رقم لاسبير ، رقم باش ، رقم فيشر (الامثل)
الارقام القياسية المرجحة:

ان ما يعاب على الارقام القياسية البسيطة كونها تعطي اهمية متساوية للسلع المختلفة لدى استخراج ارقامها القياسية اي انها تعطي صورة غير دقيقة للتغيرات الحاصلة في مستويات الاسعار او الكميات وغيرها.
لذلك نلجأ الى استخدام الارقام القياسية المرجحة اي التي تعطي لكل سلعة الوزن الحقيقي الخاص باهميتها وذلك من خلال ترجيح الاسعار وحيث يوجد لدينا نوعية من الاسعار (أسعار سنة الاساس واسعار سنة المقارنة) كذلك نوعين من الكميات (كميات سنة الاساس وكميات سنة المقارنة) لذلك يمكن ان نحصل على ثلاثة انواع من الارقام القياسية المرجحة تعرف بأسماء العلماء الذين توصلوا اليها وهي :

1-رقم لاسبير Laspeyres

$$p.i(las) = \frac{\sum pi qo}{\sum po qo} \times 100$$

2-رقم باش Paasche

$$p.i(paa) = \frac{\sum pi qi}{\sum po qi} \times 100$$

3-رقم فيشر (Fisher الامثل)

$$P.I(fis) = \sqrt{\frac{[\sum pi Q_0]}{[\sum po q_0]} \frac{[\sum pi q_i]}{[\sum po a_i]}} \times 100$$

1-رقم لاسبير: يعتمد على الترجيح بكميات سنة الاساس

مجموع اسعار سنة المقارنة × كميات سنة الاساس

رقم لاسبير = 100 ×

مجموع اسعار سنة الاساس × كميات سنة الاساس

$$p.i(las) = \frac{\sum pi qo}{\sum po qo} \times 100$$

مثال (47) الجدول الاتي يبين كميات السلع واسعارها في سنتين 2002 ، 2022 في احد المتاجر

المطلوب : ايجاد الرقم القياسي للأسعار في سنة 2022 باعتبار ان سنة 2002 هي سنة الأساس (بطريقة لاسبير)

السلعة	الاسعار بالدينار		الكميات	
	2002	2022	2002	2022
القمح	200	600	20	10
باقلاء	750	2250	25	5
ذرة	500	1250	10	20
رز	1000	800	20	8

الحل : نضرب أسعار سنة المقارنة 2022 × كميات سنة الأساس 2002 ثم نجمع حاصل الضرب
نضرب اسعار سنة الأساس 2002 × كميات سنة الأساس 2002 ثم نجمع حاصل الضرب

Pi 2022 qo 2002	Pi 2002 qo 2002
600× 20 =12000	200× 20 =4000
2250× 25=56250	750× 25=18750
1250 × 10= 12500	500 × 10= 5000
800 × 20 =16000	1000 × 20 =20000
المجموع 96750	المجموع 47750

$$p.i(las) = \frac{\sum pi qo}{\sum po qo} \times 100 = \frac{96750}{47750} \times 100 = 203 \%$$

يلاحظ ان هذا الرقم القياسي يمكن استخراجه في حالة عدم توفر البيانات الخاصة بكميات سنة المقارنة (لأنه يعتمد على كميات سنة الاساس)

2- رقم باش (الترجيح بكميات سنة المقارنة)

مجموع اسعار سنة المقارنة × كميات سنة المقارنة
الرقم القياسي لباش = $100 \times \frac{\text{مجموع اسعار سنة الاساس} \times \text{كميات سنة المقارنة}}{\text{مجموع اسعار سنة المقارنة} \times \text{كميات سنة الاساس}}$
ويأخذ الرقم القياسي لباش الصيغة الآتية:

$$p.i(paa) = \frac{\sum pi qi}{\sum po qi} \times 100$$

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
مثال (48): الجدول الآتي يبين كميات السلع وأسعارها في سنتين 2002 ، 2022 في أحد المتاجر

المطلوب : إيجاد الرقم القياسي للأسعار في سنة 2022 باعتبار ان سنة 2002 هي سنة الأساس (بطريقة باش)

السلعة	الاسعار بالدينار		الكميات	
	2002	2022	2002	2022
القمح	200	600	20	10
باقلاء	750	2250	25	5
ذرة	500	1250	10	20
رز	1000	800	20	8

الحل : نضرب أسعار سنة المقارنة 2022 × كميات سنة المقارنة 2022 ثم نجمع حاصل الضرب
نضرب اسعار سنة الأساس 2002 × كميات سنة المقارنة 2022 ثم نجمع حاصل الضرب

Pi 2022 qi 2022	Poi 2002 qi 2022
600× 10 =6000	200× 10 =2000
2250× 5=11250	750× 5=3750
1250 × 20= 25000	500 × 20= 10000
800 × 8 =6400	1000 × 8 =8000
المجموع 48650	المجموع 23750

$$p.i(paa) = \frac{\sum p_i q_i}{\sum p_0 q_i} \times 100 = \frac{48650}{23750} \times 100 = 205\%$$

الرقم القياسي الأمثل :- fis.

ويسمى بمعادلة فيشر وبموجب هذه المعادلة يمكن احتساب الرقم القياسي للأسعار من الجذر التربيعي لحاصل ضرب الناتج من معادلة باش × الرقم القياسي الناتج من معادلة لاسبيرر وعليه فان الرقم القياسي الأمثل =

$$P. I(fis) = \sqrt{\left[\frac{\sum p_i q_0}{\sum p_0 q_0} \right] \left[\frac{\sum p_i q_i}{\sum p_0 a_i} \right]} \times 100$$

الحففيه التعليميه للإحصاء المعهد التقني كربلاء - قسم تقنيات المحاسبه اعداد : م.م. حيدر جميل الجبوري
مثال (49): الجدول الاتي يبين كميات السلع واسعارها في سنتين 2002 ، 2022 في احد المتاجر
المطلوب : ايجاد الرقم القياسي للأسعار في سنة 2022 باعتبار ان سنة 2002 هي سنة الأساس (بطريقة فيشر)
الرقم القياسي الامثل

السلعة	الاسعار بالدينار		الكميات	
	2002	2022	2002	2022
القمح	200	600	20	10
باقلاء	750	2250	25	5
ذرة	500	1250	10	20
رز	1000	800	20	8

الحل : نضرب أسعار سنة المقارنة 2022 × كميات سنة الأساس 2002 ثم نجمع حاصل الضرب

نضرب اسعار سنة الأساس 2002 × كميات سنة الأساس 2002 ثم نجمع حاصل الضرب

Pi 2022 qo 2002	Pi 2002 qo 2002
600× 20 =12000	200× 20 =4000
25=562502250×	25=18750750×
× 10= 125001250	× 10= 5000500
× 20 =16000800	× 20 1000 =20000
المجموع 96750	المجموع 47750

$$p.i(las) = \frac{\sum pi qo}{\sum po qo} \times 100 = \frac{96750}{47750} \times 100 = 203 \%$$

نضرب أسعار سنة المقارنة 2022 × كميات سنة المقارنة 2022 ثم نجمع حاصل الضرب

نضرب اسعار سنة الأساس 2002 × كميات سنة المقارنة 2022 ثم نجمع حاصل الضرب

Pi 2022 qi 2022	Poi 2002 qi 2022
600× 10 =6000	200× 10 =2000
2250× 5=11250	750× 5=3750
1250 × 20= 25000	500 × 20= 10000
800 × 8 =6400	1000 × 8 =8000
المجموع 48650	المجموع 23750

$$p.i(paa) = \frac{\sum p_i q_i}{\sum p_0 q_i} \times 100 = \frac{48650}{23750} \times 100 = 205\%$$

بالنسبة لطريقة فيشر (الحل الأمثل)

$$P.I(fis) = \sqrt{\left[\frac{\sum p_i q_0}{\sum p_0 q_0} \right] \left[\frac{\sum p_i q_i}{\sum p_0 a_i} \right]} \times 100$$

$$P.I(fis) = \sqrt{\left[\frac{96750}{47750} \right] \left[\frac{48650}{23750} \right]} \times 100 = 204\%$$

تمرين شامل واجب (10) : الجدول الآتي يمثل أسعار السلع بالفلس وكميات المستهلكة (بالكغم) للعائلة الواحدة شهريا)

الكميات		الاسعار		السلعة
2017	2004	2017	2004	
8	7	750	500	السكر
12	10	250	100	الرز
2	2	100	50	الشاي
7	6	10000	5000	اللحوم
		11100	5650	

أحسب:

- 1- الرقم القياسي البسيط للأسعار - .
- 2- رقم لاسبيرز القياسي باعتبار - 2004 كسنة اساس.
- 3- رقم باش القياسي - .
- 4- رقم فيشر القياسي.

الاسبوع السابع والعشرين والثامن والعشرين والتاسع والعشرين والثلاثين
بعض المواضيع التطبيقية

- السيطرة النوعية
- جودة الإنتاج
- التعداد السكاني
- دراسات الجدوى الاقتصادية