

تقنية البيئة

مقدمة في الإحصاء

١٧١ إحص



تمهيد

تهدف هذه الحقيبة الى عرض الموضوعات الرئيسية المتعلقة بمبادئ الإحصاء لتكون في متناول المتدربين الذين يدرسون مقرر مبادئ الإحصاء في فصل دراسي واحد.

لقد انطلقنا في تأليف هذه الحقيبة إفتراض أن المتدرب لم يدرس مقرر مبادئ الإحصاء من قبل ولهذا حاولنا أن تتسم هذه الحقيبة بالبساطة والوضوح في أسلوبها ومنهجها , ولمساعدة المتدربين في التركيز على الجوانب المهمة في دراستهم , فقد قمنا خلال كل وحدة دراسية بشرح الدرس مع التوضيح بالأمثلة والعرض البياني والإشارة إلى المصطلحات الإحصائية المهمة باللغتين العربية والإنجليزية مع وجود بعض الأسئلة في نهاية كل وحدة.

وقد قسمنا الحقيبة إلى ستة وحدات تدريبيه. فالوحدة الأولى تشمل تعريف بعلم الإحصاء وأهميته وأقسامه وأهدافه الإحصائية، والوحدة الثانية تشتمل على مرحلة جمع البيانات، والوحدة الثالثة تشتمل على عرض البيانات مبوبة في شكل جداول، والوحدة الرابعة تشتمل على تمثيل ورسم البيانات الإحصائية، والوحدة الخامسة تشتمل على ايجاد وتطبيق مقاييس النزعة المركزية والوحدة السادسة الأخيرة تشتمل على استخراج مقاييس التشتت.

وأخيرا نرجوا ان تكون هذه الحقيبة بادرة لخدمة المتدربين في الكليات التقنية ومشاركة فاعلة في تطوير المناهج داخل المؤسسة العامة للتعليم الفني والتدريب المهني بما يواكب تطورات العصر الحديث.

مقدمة في الإحصاء

ماهية علم الإحصاء



الوحدة الأولى

اسم الوحدة : ماهية علم الإحصاء.

الجدارة : معرفة علم الإحصاء و أهميته.

الأهداف :

- ١ - المقدرة على تعريف الإحصاء.
- ٢ - الإلمام بمعرفة الأقسام و الأهداف وبعض المفاهيم الإحصائية.

مستوى الأداء المطلوب : ٩٠٪.

الوقت المتوقع للتدرب على الجدارة : ٤ ساعات.

الوسائل المساعدة :

١. الرجوع إلى مراجع أخرى.
٢. المناقشة.
٣. حل التمارين.
٤. واجبات منزلية.
٥. استخدام وسائل تعليمية حديثة (جهاز العرض).

متطلبات الجدارة : بدون.

ماهية علم الإحصاء

مقدمة:

تهدف هذه الوحدة إلى إعطاء الطالب فكرة موجزة عن ماهية علم الإحصاء وعلاقته بالعلوم الأخرى. فعلم الإحصاء من العلوم التطبيقية المساعدة التي أصبحت من العلوم الملازمة للكثير من العلوم التجريبية وذلك بهدف صياغة النظريات الخاصة بهذه العلوم في قالب رقمي يتخذ من الأساليب الإحصائية أداة لتفسير ظواهره المختلفة واختبار نظرياته المتشعبة. وحتى تتضح الصورة، نسعى في هذه الوحدة إلى إبراز كل ما من شأنه إيضاح أهمية علم الإحصاء مع إبراز بعض المفاهيم والمصطلحات الأساسية التي نحرص على عرضها للطالب كقاعدة نظرية لهذا المنهج من خلالها نعرض أقسام و بعض أهداف علم الإحصاء.

١. تعريف علم الإحصاء: Statistics

كان الإحصاء قديماً يشار إليه على أنه علم العد أو الحصر ويقصد به حصر المجتمع المراد دراسته، قوله تعالى (لقد أحصاهم وعدهم عداً) آية ٩٤ سورة مريم وقوله تعالى (وأحاط بما لديهم وأحصى كل شيء عدداً) آية ٤ سورة المجادلة. وتختلف طريقة الحصر باختلاف الهدف من الدراسة، فقد يكون الحصر شاملاً لجميع مفردات المجتمع كما قد يقتصر على بعض مفردات المجتمع كعينة، بحيث يتم دراسة هذه العينة بهدف تعميم النتائج على المجتمع كله ومع مرور الزمن أصبحت المهمة الإحصائية أكثر مهاماً من كونها عداً وحصرًا لمفردات المجتمع حتى أصبح يطلق على علم الإحصاء علم اتخاذ القرارات الحكيمة وحتى في ظل عدم توفر المعلومات الكافية. ونعرض فيما يلي بعض التعريفات التي أطلقها الإحصائيون على علم الإحصاء:

- تعريف روملين لعلم الإحصاء على أنه ذلك العلم الذي يتبع كافة العلوم التجريبية التي تدرس العالم الإنساني. ويقصد بهذا التعريف إبراز أهمية علم الإحصاء في خدمة العلوم التجريبية كالاقتصاد والاجتماع والصحة والبيئة وغيرها من العلوم التجريبية.

- تعريف هايس لعلم الإحصاء على أنه ذلك العلم الذي ينطوي على الجانبين التاليين:

- مجموعة الوقائع المعبر عنها بالأرقام والأعداد كإحصاءات المحلات التجارية والمصانع والسيارات والحوادث والوفيات وغيرها.

- الطرق المختلفة المستخدمة في تحليل البيانات الإحصائية من ناحية جمعها وتبويبها وتبسيطها وتحليلها وما إلى ذلك من العمليات الإحصائية.
- التعريف العام: بالنظر إلى العمليات الإحصائية التي تتم عند الرغبة في دراسة مشكلة ما، فإنه يمكن تعريف علم الإحصاء على أنه ذلك العلم الذي يهتم بجمع البيانات وتبويبها وعرضها وتحليلها بغية التعرف على مدلولها وإمكانية الاستفادة من ذلك في وضع الخطط المستقبلية المتعلقة بالمشكلة موضع الدراسة.

ويلاحظ أن هذا التعريف يعطي تعريفاً وصفاً للوظيفة الإحصائية، حيث تبدأ أولاً بمرحلة جمع البيانات وتنتهي بتحليلها مروراً بالعمليات التبسيطية المتمثلة في تبويب وعرض تلك البيانات. ومن هذا التعريف نستطيع أن نتعرف على علم الإحصاء البيئي أو الصحي أو الاقتصادي أو الاجتماعي وذلك من خلال التعرف على ماهية البيانات الإحصائية التي يرغب الباحث في جمعها حول المشكلة موضع الدراسة، فالباحث في مجال البيئة يقوم بجمع البيانات المتعلقة بالظواهر البيئية مستعيناً في دراستها بالأساليب الإحصائية المختلفة، وعلى ذلك يمكن القول بأن الإحصاء البيئي هو ذلك العلم الذي يهتم بجمع البيانات البيئية وتبويبها وعرضها ووصفها وتحليلها.. الخ. كما قد يستعين الطالب بالأساليب الإحصائية لدراسة البيئة وعلاقتها بالمتغيرات الصحية، ومن ثم فإنه يقوم بجمع البيانات المتعلقة بالظواهر الصحية ثم تبويبها وعرضها وتحليلها وتفسير النتائج والاستفادة منها في التعرف على الواقع ووضع الخطط المستقبلية.

ونخلص من ذلك أن علم الإحصاء علم مستقل له مبادئه وقواعده وأساليبه التي يمكن الاستفادة بها في مختلف المجالات العلمية، فإذا ما طبق في مجال الصحة أصبح إحصاءاً صحياً، وفي حالة تطبيقه في المجال البيئي يصبح إحصاءاً بيئياً وهكذا.

٢. أقسام علم الإحصاء:

يمكن تقسيم علم الإحصاء إلى قسمين رئيسيين هما:

- الإحصاء الوصفي Descriptive Statistics: ويختص بتجميع وعرض البيانات كما تشاهد في الواقع بغرض توفير المعلومات عن الاتجاهات المختلفة للظاهرة، وقد يصاحب هذا استخراج بعض المؤشرات الإحصائية المتوسطات أو النسب التي تفسر وتشرح تلك البيانات.

- الإحصاء الاستدلالي Inference Statistical: ويعتمد على نظرية الاحتمالات ويختص بمحاولة اتخاذ قرارات دقيقة من البيانات المتاحة. كما يساعد في تقرير كيفية الحصول على البيانات المطلوبة بأقصى كفاءة ممكنة وبأقل تكلفة ممكنة.

٣. أهداف علم الإحصاء:

يمكن إيجاز أهم أهداف علم الإحصاء في النقاط التالية:

- التعبير عن الحقائق: نظراً لكون علم الإحصاء يتعامل أصلاً مع البيانات الرقمية والوصفية، فإن ذلك يمثل أدق أنواع القياس والتعبير عن الواقع وذلك لكون الأرقام بطبيعتها لا تعرف الميول والكذب.

- الاهتمام بالمجموعات: يهتم علم الإحصاء بدراسة الظواهر المتكررة ذات الصبغة التجميعية، ذلك أن علم الإحصاء لا يهتم بدراسة حادثة منفردة وقعت في مكان ما، بل أن الظاهرة التي تتكرر مكانياً أو زمنياً هي التي تستأثر باهتمام علم الإحصاء.

- تبسيط البيانات: سبق أن ذكرنا أن جمع البيانات بشكلها الخام غير مجدي لإتمام الدراسة الإحصائية، لذلك فإنه يتم تبسيط هذه البيانات وذلك بتقسيمها وتصنيفها بالشكل الذي يسهل على الباحثين والدارسين تحليلها، مثال ذلك أن نقوم بتصنيف الأغذية المستوردة وفقاً للسبب أو وفقاً لصفاتها (ضرورية وغير ضرورية)، كما يمكن تصنيف المواليد وفقاً للجنس أو الحالة العامة للمريض وما إلى ذلك من الصفات.

- إجراء المقارنات: يمكن عن طريق البيانات الرقمية أو العددية والوصفية إجراء المقارنة بين الحالات المتشابهة، فلو أردنا أن نقارن عدد الأغذية المستوردة في مدينتين مختلفتين، كأن نقول مثلاً أن عدد الأغذية المستوردة في المدينة الأولى أكبر من عددها في المدينة الثانية، كما يمكن أن نقول أن عدد الأغذية المستوردة في المدينة الأولى أكبر من عددها في المدينة الثانية، كما يمكن أن نقول أن عدد الأغذية المستوردة في المدينة الأولى ٢٠٠ طن في السنة وعدد الأغذية المستوردة في المدينة الثانية ١٠٠ طن في السنة، وإذا ما نظرنا إلى العبارتين نلاحظ أن العبارة الأولى مطلقة تحتل الصدق أو الكذب ولكن عندما نذكر عدد الأغذية المستوردة رقمياً فإن المقارنة تكون أكثر دقة وموضوعية.

- تعميم النتائج: عند إجراء الدراسة الإحصائية، فإننا قد نستنتج بعض النتائج التي يمكن تعميمها على المجتمع بشكل عام عن طريق أساليب إحصائية. فلو قمنا مثلاً بدراسة بعض مفردات المجتمع

والتي تم اختبارها إحصائياً كما سنتناول ذلك لاحقاً، يمكن لنا أن نعمم النتائج على المجتمع الذي سحبت منه هذه العينة موضع الدراسة.

- **التوقعات:** يهتم علم الإحصاء بعمل التوقعات اللازمة عن المشكلة أو الظاهرة محل الدراسة، فعن طريق تحليل البيانات الرقمية والوصفية التي تم الحصول عليها والمتعلقة بسلوك الظاهرة في الماضي والحاضر والتوقع بالسلوك المستقبلي لتلك الظاهرة وفقاً للنتائج المستخلصة فلو افترضنا أن عدد حوادث السيارات في تزايد مستمر خلال السنوات العشر الماضية، فإنه يمكن القول بأنه مع افتراض بقاء الأشياء الأخرى المؤثرة على الظاهرة على ما هي عليه، فإن عدد الحوادث سوف يزداد في السنة المقبلة عما هو عليه في السنة الحالية ولا يمكن لنا أن نقول غير ذلك ما لم يكن لدينا معلومات جديدة يتوقع أن تؤدي إلى نقص في عدد حوادث السيارات مثل فرض عقوبات صارمة جديدة على السرعة أو وضع كمرات مراقبة أو الفحص الدوري على السيارات.. وما إلى ذلك من العوامل التي قد تؤدي إلى الإقلال من عدد الحوادث.

٤. بعض المفاهيم الإحصائية المستخدمة:

نتناول فيما يلي بعض المفاهيم الإحصائية التي يتعين على الدارس الإلمام بها قبل تعرضه لدراسة محتويات هذا المقرر وذلك حتى يتييسر للطلاب الفهم الواضح للأساليب الإحصائية:

- **المجتمع الإحصائي Population:** هو عبارة عن مجموعة من الوحدات (مثل الأفراد أو الأشياء) تشترك فيما بينها في صفة متغيرة أو أكثر. فالمجتمع الإحصائي عبارة عن جميع القيم أو المفردات التي يمكن أن يأخذها المتغير. وتمثل المفردة وحدة جمع البيانات والتي قد تكون إنساناً أو حيواناً أو نباتاً.. الخ وفقاً لطبيعة المجتمع. فمثلاً إذا كانت دراستنا متعلقة بأطوال طلاب كلية التقنية مثلاً، فإن المجتمع في هذه الحالة هو أطوال جميع طلبة الكلية وقد يكون المجتمع الإحصائي:

- **مجتمعاً منتهي Finite Population:** وهو المجتمع الذي يمكن حصر عدد مفرداته كما هو الحال في أطوال طلبة الكلية مثلاً أو عدد الوحدات الإنتاجية في مصنع ما في يوم معين.. الخ.
- **مجتمعاً غير منتهي Infinite Population:** وهو المجتمع الذي يصعب أو يستحيل حصر عدد مفرداته مثل درجة الحرارة في مدينة معينة خلال سنة كاملة.

- **العينة Sample:** هي جزء من المجتمع، فهي عبارة عن مجموعة من المشاهدات اختيرت بطريقة ما من المجتمع، حيث أن دراسة المجتمع ككل قد يكون صعباً أو يحتاج إلى وقت وجهد وتكلفة، لذا فقد

استعويض عن دراسة المجتمع بدراسة العينة وصفاتها، ومنها نستطيع أن نستنتج خواص المجتمع الأصلي الذي سحبت منه العينة. ويشترط في العينة أن تكون ممثلة للمجتمع.

- المتغير Variable: هو الصفة أو الخاصية القابلة للتغير في النوع أو الكم من فرد إلى آخر بين أفراد العينة أو المجتمع. فأفراد مجموعة ما (طلبة الكلية مثلاً) يختلفون في السن، الطول ولون العين والذكاء ومستوى الدخل.. الخ من الصفات. فكل صفة من هذه الصفات وغيرها تعتبر متغيراً. ويمكن تقسيم المتغيرات إلى قسمين رئيسيين هما:

أ - متغيرات وصفية Qualitative Variables: وهي المتغيرات التي لا يمكن قياسها مباشرة بأرقام عددية، بل يمكن تقسيم المفردات في المتغير الوصفي إلى عدد قليل من الفئات لا تداخل بينها ومن أمثلة تلك المتغيرات الشكل، القوام، الذكاء، الطعم، الرائحة... الخ.

ب - متغيرات كمية Quantitative Variables: وهي تشمل جميع المتغيرات التي يمكن التعبير عنها بصورة كمية (أرقام عددية) منها الصفات الاقتصادية والأطوال والأوزان وعدد السكان... إلخ، ويمكن تقسيم المتغيرات الكمية بدورها إلى:

■ متغيرات متصلة Continuous Variables: وفيها يأخذ كل فرد أي قيمة في مجال التغير مثل الأطوال والأوزان، حيث يمكن قياسها تدريجياً وليس هناك حداً فاصلاً بين قياس ما والقياس التالي، فهي قد تمثل بعدد صحيح أو كسر من العدد الصحيح للمقياس المستعمل.

■ متغيرات منفصلة أو متقطعة Discrete Variables: وفيها لا يقل معدل التغير عن وحدة القياس مثل عدد طلبة الكلية أو عدد سكان المملكة، حيث تأخذ هذه الأعداد قيماً عددية صحيحة.

- الثابت Constant: إذا كان هناك صفة ما لا تختلف من فرد إلى آخر من أفراد العينة أو المجتمع، ولهذا فإن الإحصاء لا يمكن تطبيقه على الثوابت لأن مجرد معرفة قيمة مفردة واحدة من مفردات المجتمع يمكننا معرفة كل القيم في المجتمع فهي متساوية تماماً، بل تختص الطرق الإحصائية بدراسة المتغيرات. بعد أن يأخذ الباحث هذه المبادئ في اعتباره، فإنه يتبع معالجة البيانات الإحصائية التالية:

١ - مرحلة جمع البيانات، ٢ - مرحلة تبويب البيانات، ٣ - مرحلة تمثيل (رسم) البيانات، ٤ - مرحلة تحليل البيانات. وسوف نتناول ذلك في الوحدات القادمة بإذن الله.

تمارين

السؤال الأول:

ضع تعريفاً مناسباً لعلم الإحصاء، موضحاً أهم أهدافه؟

السؤال الثاني:

أجب بنعم أم بلا مع تصحيح العبارات الخاطئة:

- ١ - يعتمد الإحصاء الوصفي على نظرية الاحتمالات.
- ٢ - لا يمكن تحويل المتغيرات الكمية إلى متغيرات وصفية.

السؤال الثالث:

افتراض أنك بصدد دراسة ظاهرة اجتماعية معينة، فما هي الخطوات التي تقوم بها لدراسة الظاهرة والتعرف على طبيعتها وسلوكها ودوافعها؟

السؤال الرابع:

ناقش مع التوضيح بالأمثلة كلا مما يأتي:

- ١ - الاستدلال الإحصائي.
- ٢ - المجتمع الإحصائي والعينة.
- ٣ - المتغير والثابت.
- ٤ - المتغير الكمي والمتغير الوصفي.
- ٥ - المتغير المتصل والمتغير المنفصل.

مقدمة في الإحصاء

جمع البيانات الإحصائية



الوحدة الثانية

اسم الوحدة : جمع البيانات الإحصائية.

الجدارة : معرفة كيفية جمع البيانات الإحصائية.

الأهداف :

- ١ - المقدرة في تحديد الدراسة الإحصائية والغرض منها على مبادئ أساسية.
- ٢ - المقدرة في معرفة طرق الحصول على البيانات الإحصائية.
- ٣ - المقدرة على معرفة الأسلوب الحصائي وبعض العينات الإحصائية.

مستوى الأداء المطلوب : ٩٥٪.

الوقت المتوقع للتدرب على الجدارة : ٤ ساعات.

الوسائل المساعدة على تحقيق الجدارة :

١. الرجوع إلى مراجع أخرى.
٢. المناقشة.
٣. حل التمارين.
٤. واجبات منزلية.

متطلبات الجدارة : الوحدة الأولى.

جمع البيانات الإحصائية

مقدمة:

إنه لمن الأهمية بعد أن تعرفنا على أهمية علم الإحصاء أن نتناول في هذه الوحدة مرحلة جمع البيانات. ولقد بدئ واضحاً من تعريف علم الإحصاء أن دور الإحصاء يبرز من خلال جمع البيانات، وتبويبها وعرضها ثم تحليلها، وهذا الدور يمثل الخطوات الرئيسية التي يتعين على الإحصائي اتباعها حتى تستطيع الحصول على نتائج يمكن الاعتماد عليها من خلال دراسته الإحصائية، غير أنه يتعين عليه قبل ذلك مراعاة المبادئ الأساسية التالية:

١ - **تحديد المشكلة:** من الضروري تحديد الغرض من الدراسة وتحديد المشكلة التي سيتناولها الباحث بالبحث والتقصي، فليس من المنطقي أن يبدأ الباحث مهمته الإحصائية دون أن يعرف ما هو الغرض من هذه المهمة وما هي المشكلة التي يتعين عليه القيام بجمع البيانات الإحصائية حولها، كذلك تبويبها وعرضها تم تحليلها.

٢ - **تحديد المجتمع المراد دراسته:** يتعين على الباحث تحديد وتوصيف نطاق المجتمع الإحصائي الذي يرغب دراسته. فلو أردنا مثلاً إجراء دراسة حول طلبة الدبلوم بكلية التقنية، فإن المجتمع الإحصائي هنا هو جميع طلبة الدبلوم وكل طالب في هذا القسم يمثل مفردة من مفردات هذا المجتمع الإحصائي.

٣ - **تحديد البيانات المطلوب جمعها:** يتعين على الباحث أيضاً المعرفة الدقيقة بالمتغيرات التي يرغب جمعها من مجتمع البحث، فلو أردنا مثلاً جمع بيانات من طلاب الدبلوم حول الوضع داخل القسم فإنه يلزم طرح الأسئلة المتعلقة بالوضع القائم، دون طرح تساؤلات أخرى لا تتطلبها المشكلة موضع الدراسة.

٤ - **طريقة الحصول على البيانات:** يفترض معرفة الباحث بالطرق التي سوف تمكنه من الحصول على البيانات المطلوبة، ومن أهم هذه الطرق:

أ - **المقابلة الشخصية:** حيث يتم مقابلة الأشخاص وجهاً لوجه لتعبئة الاستمارة والتساؤلات المطلوبة وإعادة كافة الاستمارات (الاستبانات). وهي أفضل الطرق حفاظاً للبيانات واسترجاعها كاملة وبأسرع مدة.

- ب - المراسلة بريدياً: حيث يتم إرسال استمارة تحتوي على بعض التساؤلات التي يرغب الباحث الإجابة عليها، وفي الغالب يتم إرسال مظروف مع الاستبانة حتى يسهل على المبحوث إعادة الاستبانة دون كلفة مادية.
- ج - الاتصال التليفوني: وهذه الطريقة تتسم بالسرعة في الحصول على الإجابات المطلوبة، إلا أن الأمر يصبح غاية في الصعوبة عندما تشتمل الاستبان على أسئلة محرّجة.
- د - التسجيل: وهي الطريقة المتبعة لدى الهيئات والمؤسسات الحكومية وغير الحكومية، كأن يقوم شئون الطلاب بتسجيل الملاحظات المتبعة التي يتعرض لها الطالب، حيث يمكن من هذه السجلات الحصول على البيانات المطلوبة.
- هـ - البريد الإلكتروني (e-mail): حيث ترسل الاستمارة بالكمبيوتر في الإنترنت (الشبكة الإلكترونية الحديثة) ويقوم بإستقبالها الأشخاص الذين أرسلت لهم. وهي أسهل طريقة على الباحث في إرسال وإستقبال تلك الإستمارات وأقل كلفة مادية، ولكن غير مضمون في استرجاعها كاملة وبأسرع مدة.
- و - الإستمارة (الإستبانة) الإحصائية: هي عبارة عن ورقة أو أكثر تحتوي على أسئلة يراد الإجابة عنها وكل سؤال يترك له مكان للإجابة عنه، ثم توزع على أفراد العينة من المجتمع محل الدراسة، ويتم جمعها بعد وقت كافٍ للإجابة عنها. ويجب أن تكون الأسئلة في الإستمارة واضحة ودقيقة، وغير مملّة، وغير محرّجة وغير متحيّزة.

١. الأساليب الإحصائية:

- يلي تحديد المشكلة البحثية، تحديد المجتمع المراد دراسته وطريقة الحصول على البيانات، فإن الباحث ينهج أحد الأسلوبين التاليين للحصول على البيانات الإحصائية اللازمة لإجراء بحثه.
- ١ - أسلوب الحصر الشامل للمجتمع Population: ويقصد به حصر جميع مفردات المجتمع وجمع كافة البيانات المطلوبة حول هذه المفردات.. ويتبع هذا الأسلوب في عمليات التعداد السكاني.

- ٢ - أسلوب الحصر الجزئي أو العينات Samples: ويقصد به سحب مجموعة صغيرة من مفردات المجتمع الإحصائي بهدف الحصول منها على استنتاجات يمكن تعميمها على المجتمع الإحصائي. ويسعى الباحث عند استخدامه لهذا الأسلوب إلى تحقيق الأهداف التالية:
- أن تكون العينة ممثلة Representative للمجتمع الأصلي.
 - أن يكون حجم العينة كافياً بمعنى أن تتناسب مع حجم المجتمع الأصلي.
 - أن يتحقق مفهوم العشوائية Randomness عند اختيار العينة ويقصد بذلك عدم التحيز لبعض مفردات المجتمع في عملية الاختيار.

٢. أنواع العينات:

استناداً إلى الأهداف السابقة التي يحرص الباحث على تحقيقها عند اختيار العينة، يمكن عرض أنواع العينات الشائعة الاستخدام على النحو التالي:

أولاً: العينات الاحتمالية: وهي العينات التي يتم سحبها وفقاً لقوانين الاحتمالات وأهما:

- (١) العينة العشوائية البسيطة.
- (٢) العينة العشوائية المنتظمة.
- (٣) العينة العشوائية الطبقية.
- (٤) العينة متعددة المراحل.

وتوجد هناك عينات احتمالية (عشوائية) مثل العينة العنقودية Cluster Random Sample وغيرها. ولكن نكتفي بالاربع السابقة.

ثانياً: العينات غير الاحتمالية: وهي العينات التي لا يتم سحبها وفقاً لقوانين الاحتمالات وأهما:

- (١) العينة العمدية.
- (٢) العينة كبيرة الحجم.

وفيما يلي نعرض العينات الاحتمالية الأربع بجانب التركيز على العينة العمدية من العينات غير الاحتمالية، حيث من الشائع استخدام العينات الاحتمالية في الدراسات الإحصائية.

■ العينة العشوائية البسيطة Simple Random Sample:

وهي العينة التي يتم اختيارها بشكل يضمن تكافؤ الفرص في الاختيار لجمع مفردات المجتمع وهي تتطلب معرفة نطاق المجتمع.

■ العينة العشوائية المنتظمة Systematic Random Sample:

هي العينة التي يتم اختيار المفردة الأولى منها عشوائياً ثم يتم إضافة رقم ثابت بعد ذلك حتى نحصل على باقي مفردات العينة بالحجم المطلوب. وهي من أهم الطرق الشائعة في اختيار العينات. وتعتمد فكرتها الأساسية على تساوي الفترات بين الوحدات المختارة في العينة، وهي لا تتطلب معرفة نطاق المجتمع.

■ العينة العشوائية الطبقية Stretified Random Sample:

وهي العينة التي يتم اللجوء لاستخدامها عندما يكون المجتمع غير متجانس بحيث يمكن تمييزه إلى طبقات مختلفة.

■ العينة متعددة المراحل Multistage Sample:

وهي العينة التي تم اختيارها على مراحل مختلفة. وتستخدم هذه الطريقة خاصة إذا كان حجم المجتمع الأصلي كبيراً ويتعذر على الباحث استخدام الأساليب السابقة لاختيار مفردات العينة نظراً لانتشار وحدات العينة في مساحات جغرافية متشعبة وكذلك لصعوبة إعداد إطار لجميع وحدات المجتمع في بعض الحالات. وتجدر الإشارة إلى أنه من الممكن أخذ عينة منتظمة أو طبقية في المرحلة الأولى أو عشوائية في الثانية وهكذا ويتوقف ذلك على طبيعة موضوع الدراسة والغرض من البحث ونوع الوحدات.

■ العينة العمدية Purposive Sample:

تبدو فائدة الاختبار العمدية عندما يتكون المجتمع من وحدات قليلة نسبياً معروفة الخصائص لدى الباحث القائم بتصميم خطة العينة، ومن ثم يمكن اختيار العينة على النحو الذي يضمن توافر مجموعة من خصائص المجتمع الهامة ذات الصلة بموضوع الدراسة.

تمارين

- ١ - أذكر ثلاث طرق للحصول على البيانات ؟
- ٢ - أعط تعريفاً مختصراً لكل من المجتمع والعينة مع الأمثلة لكل منهما ؟
- ٣ - قارن بين العينة العمدية والعينة متعددة المراحل من حيث التعريف، كيفية الاختيار ودواعي الاستعمال.
- ٤ - أجب بنعم أم لا:
 - أ - تستخدم العينة العشوائية الطبقية في حالة تجانس المجتمع.
 - ب - لا تتطلب العينة العشوائية المنتظمة معرفة نطاق المجتمع.
 - ج - مكن استخدام أسلوب العينة العمدية في حالة عدم المعرفة بخصائص المجتمع.

مقدمة في الإحصاء

تبويب البيانات الإحصائية



الوحدة الثالثة

اسم الوحدة : تبويب البيانات الإحصاء.

الجدارة : معرفة تبويب (جدولة) البيانات الإحصائية.

الأهداف :

- ١ - المقدرة على تبويب البيانات الإحصائية.
- ٢ - معرفة توزيع الجداول التكرارية المختلفة للبيانات الإحصائية.

مستوى الأداء المطلوب : ٩٥٪.

الوقت المتوقع للتدرب على الجدارة : ٦ ساعات.

الوسائل المساعدة على تحقيق الجدارة :

١. الرجوع إلى مراجع أخرى.
٢. المناقشة.
٣. حل التمارين.
٤. واجبات منزلية.

متطلبات الجدارة :

فهم الوحدات السابقة.

تبويب البيانات الإحصائية

مقدمة:

يعرف تبويب (جدولة) البيانات بأنه عملية تنظيم حشد كبير من البيانات والمعلومات الأولية في شكل ملخص بحيث تكون منظمة ومصنفة ومرتبعة، وتمثل عملية التبويب الخطوة الأولى لإجراء عملية التحليل الإحصائي. ويتمثل الغرض من تبويب البيانات في تسهيل مهمة المحلل الإحصائي، إذ أن البيانات في شكلها الخام تكون في صورة غير معبرة يصعب معها استخلاص النتائج المطلوبة، لذا فإن الباحث يلجأ دائماً إلى تصنيف هذه البيانات الخام وفقاً لصفة هذه البيانات أو الهدف المبتغى من الدراسة.

١. العرض الجدولي للبيانات Tabular Presentation:

يعرف الجدول الإحصائي بأنه عبارة عن الترتيب المنظم للبيانات في صورة صفوف وأعمدة حتى يمكن مقارنة البيانات ببعضها وإبراز أهميتها مثل الجداول التكرارية التالية:

٢. التوزيع التكراري Frequency Distribution:

بالرغم من كثرة وتنوع أساليب تبويب البيانات، إلا أن السائد هو تبويبها في صورة فئات أو مجموعات، بحيث تحتوي كل فئة على مجموعة متشابهة من البيانات سواء من حيث الصفات الوصفية أو الكمية أو الزمنية أو المكانية. ولتوضيح ذلك افترض أن الباحث قد قام بجمع بيانات عدد الموظفين على مستوى ٥٠ مطعم في بعض مدن المملكة العربية السعودية خلال سنة ما وذلك على النحو المبين بالجدول التالي:

مثال (١ - ٣) عدد الموظفين على مستوى ٥٠ مطعم بالمملكة العربية السعودية

٣٤	٢٣	١٢	٥	١٦	٧٧	٥٠	١٠	١٥	٥
٥	٣٤	٢٣	٨١	٨٠	٦٦	٣٣	٢٥	٤٥	٤٠
١٦	٤٢	٤٢	٥٣	٢٠	٣٠	٢٨	٣٨	١٨	٦
٢٥	٨٢	٧٣	٤٠	٢٥	٢٢	٤٣	١٠	٩	٧
١١	٣٢	٤٣	٣٢	٦٧	١٦	٥٦	٢٦	١٩	١٢

من الملاحظ أن هذه البيانات لا تعطي دلالة واضحة، حيث أن كل ما يمكن معرفته من خلالها أن المطعم الأول يوجد بها ٥ موظفين، بينما يوجد في المطعم الخمسين ١١ موظف. كما يصعب علينا وضع هذه البيانات في جدول يوضح المطعم وعدد الموظفين المقابل لها، حيث يتطلب ذلك عدداً من الصفحات. وللتغلب على هذه المشكلة فإنه يمكن إعداد هذه البيانات في صورة جدول تكراري تبرز خصائص المتغير ذو الأهمية، والذي يمثل التوزيع التكراري للبيانات والموضح بالجدول التالي:

جدول (١): جدول التوزيع التكراري لعدد الموظفين على مستوى ٥٠ مطعم بالسعودية:

الفئة	العلامات	التكرار F	مركز الفئة X
١٩- ٥	/// /// /// /// /// //	١٧	١٢,٥
٢٠ -	/// /// /// /// ///	١٥	٢٧,٥
٣٥ -	/// /// /// /// ///	٨	٤٢,٥
٥٠ -	/// /// /// /// ///	٣	٥٧,٥
٦٥ -	/// /// /// /// ///	٤	٦٢,٥
٨٠ -	/// /// /// /// ///	٣	٧٧,٥
المجموع		٥٠	

وتبين من الجدول (١) أننا قمنا بتقسيم البيانات إلى فئات متقاربة، حيث تضم الفئة الأولى عدد المطاعم التي تتراوح عدد موظفيهم ما بين ٥ موظفين و ١٩ موظف والتي بلغ عددها ١٧ مطعم. أما الفئة الثانية (٢٠ - ٣٤) فإنها تضم عدد المطاعم التي تتراوح عدد الموظفين بها ما بين ٢٠ موظف وأقل من ٣٥ موظف، والتي بلغ عدد ١٥ مطعم، وهكذا بالنسبة لباقي الفئات.

ومن الملاحظ أيضاً أن المجموع للتكرارات يساوي ٥٠ مطعم وهذا يعني أن مطاعم العينة قد تم توزيعها على الفئات المختلفة وفقاً لعدد الموظفين الذي يقع في نطاق كل فئة. ولتسهيل عملية إعداد هذا الجدول نسجل الفئات في العمود الأول، ثم نقوم بحصر الأرقام التي تشتمل عليها كل فئة من الفئات المحددة، ونعبر عنها في شكل حزم، بحيث تشتمل الحزمة الواحدة على خمس علامات على أن توضع العلامة الخامسة بشكل متقاطع مع العلامات الأربع الأخرى وبعد توصيف جدول التوزيع التكراري، يتعين توضيح بعض المفاهيم المرتبطة بالتوزيع التكراري وكيفية حسابها وذلك بهدف التوصل إلى الخطوات الواجب اتباعها لإعداد جدول التوزيع التكراري سالف الذكر.

■ المدى Range: يعرف المدى بأنه الفرق بين القيمة الكبرى والقيمة الصغرى من القيم المعطاة أي أن:

$$\text{المدى} = \text{القيمة الكبرى} - \text{القيمة الصغرى}.$$

ومن واقع بيانات الجدول (١) نجد أن: المدى = ٨٢ - ٥ = ٧٧.

ويمكننا أن نقسم هذا المدى إلى عدد معين من الفئات وهو في مثالنا فرضنا أن طول الفئة = ١٥

$$\therefore \text{عدد الفئات} = \frac{\text{المدى}}{\text{طول الفئة}} = \frac{٧٧}{١٥} = ٥,١, \text{ أي أن عدد الفئات يساوي } ٦ \text{ فئات.}$$

■ **حدود الفئة Class Limits:** من الواضح أن لكل فئة من الفئات التي يشتمل عليها جدول التوزيع

التكراري حدان هما: الحد الأدنى والذي يمثل القيمة التي تبدأ بها الفئة والحد الأعلى وهو القيمة التي تحدد نهاية الفئة، ومن واقع بيانات الجدول (١) نجد أنه بالنسبة للفئة الأولى: الحد الأدنى يساوي ٥، الحد الأعلى يساوي ٢٠.

■ **طول الفئة Class Length:** يعبر طول الفئة عن الفرق بين الحد الأعلى للفئة والحد الأدنى لها، أي

$$\text{أن: طول الفئة} = \text{الحد الأعلى للفئة} - \text{الحد الأدنى للفئة}.$$

ومن واقع بيانات الجدول (١) نجد أن: طول الفئة الأولى مثلاً: ٢٠ - ٥ = ١٥ .

ومن الملاحظ أن طول الفئة في مثالنا هذا متساوي لجميع الفئات وهي الحالة الأكثر شيوعاً، وإن كان هنالك بعض الحالات التي يكون فيها طول الفئة مختلفاً، خاصة إذا كان المدى الكلي كبيراً جداً كأن تكون القيمة الصغرى ٥٠ والقيمة الكبرى ٢٠٠٠٠ مثلاً. ففي مثل هذه الحالات يلجأ الباحث عادة إلى فئات ذات أطوال مختلفة، غير أننا سوف نركز في دراستنا هذه على الفئات ذات الأطوال المتساوية.

■ **مركز الفئة Mid Point-Class:** يعبر مركز الفئة عن القيمة التي تقع في المنتصف بين الحد

الأدنى للفئة والحد الأعلى لها ويمكن إيجاده من الصيغة الرياضية التالية:

$$\text{مركز الفئة} = \frac{\text{الحد الأدنى للفئة} + \text{الحد الأعلى للفئة}}{٢}$$

ويعرض الجدول (١) السابق القيم المقدرة لمراكز الفئات وفقاً لهذه المعادلة، غير أنه من الملاحظ وجود

$$\text{بعض الكسور في مراكز الفئات، حيث أن: مركز الفئة الأولى} = \frac{٥ + ٢٠}{٢} = ١٢,٥$$

ويعرض الجدول (١) السابق القيم المقدرة، لذا يجب تعديل هذه القيمة إلى ١٢ أو ١٣، وهكذا بالنسبة لمثل هذه الفئات، وذلك لأن عدد الموظفين من المتغيرات غير المستمرة (أي لاتقبل الكسور) أو منفصلة (متقطعة).

وعلى ذلك فإنه يمكننا الآن تحديد الخطوات التي يجب اتباعها عند إعداد جدول التوزيع التكراري وذلك على النحو التالي:

١. تحديد المدى الكلي للبيانات المعطاة كما سبق توضيحه.

٢. تقسيم المدى الكلي إلى فئات مناسبة، أي تحديد عدد الفئات التي يجب أن يشتمل عليها

الجدول التكراري. ونستطيع في هذا الصدد أن نميز بين حالتين:

أ - إذا كان عدد المفردات المعطاة قليلاً فإنه يترك للباحث حرية تحديد عدد الفئات التي يتضمنها الجدول التكراري.

ب - إذا كان عدد المفردات كبيراً، فإن عدد الفئات يتحدد وفقاً لقاعدة سترجس Sturges :

$$\text{عدد الفئات} = ١ + (٣,٣ \times \text{لون}).$$

حيث: لون = لوغاريتم عدد المفردات.

ومن المفضل ألا يزيد عدد الفئات عن ٢٠ فئة ولا يقل عن ٥ فئات.

■ حدود الفئات في حالة المتغيرات المستمرة وغير المستمرة:

لقد سبق أن تعرفنا في الباب الأول على الفرق بين المتغيرات المستمرة والتي يمكن أن تأخذ أرقاماً كسرية (عشرية)، والمتغيرات غير المستمرة أو (المنفصلة) والتي لا تأخذ قيماً كسرية (عشرية) بل تتغير بوحدات صحيحة. ويعتبر الموظفين من المتغيرات غير المستمرة، حيث أنه ليس من المعقول عملياً أن نقول أن عدد الموظفين يساوي ٣٠,٥ موظف مثلاً وعلى ذلك نكون بصدد تحديد حدود الفئات المناسبة لنوع المتغير، وهنا يمكن كتابة الفئة بطريقتين:

- الطريقة الأولى:

$$\blacksquare ١٩ - ٥$$

$$\blacksquare ٢٠ - ٣٤$$

تستخدم في حالة المتغيرات غير المستمرة المنفصلة (الموظفين)، لن نجد هنا الفئة التي يقع خلالها الرقم

١٩,٥، حيث أنه أكبر من نهاية الفئة الأولى وأقل من بداية الفئة الثانية، لذلك تأخذ القيمة ١٩.

- الطريقة الثانية:

$$\blacksquare 5 - 19,5$$

$$\blacksquare 20 - 34,5$$

تستخدم في حالة المتغيرات المستمرة ويقع الرقم 19,5 في الفئة الأولى.

ومما تقدم يتضح أنه إذا كانت الأرقام غير مستمرة , أي منفصلة (متقطعة) كالموظفين مثلاً فإنه يمكن استخدام الطريقة الأولى. أما إذا كانت الأرقام مستمرة كالأعمار والأوزان والأطوال مثلاً، فإنه يمكن استخدام الطريقة الثانية بصفة دائمة.

وحيث أن الطريقة الثانية تتناسب مع كلا الحالتين، فإننا نرى الاستمرار في استخدامها بغض النظر عن طبيعة المتغيرات لملاءمتها الدائمة للمتغيرات المختلفة.

٣. النسبة (النسب) والنسبة المئوية Ratio and Percentage:

ج (الجزء)

يعرف النسب Ratio على أنه حاصل قسمة الجزء على الكل

ك (الكل)

أما النسبة المئوية Percentage فهي حاصل ضرب النسب $\times 100$ أي أنها تساوي

$$100 \times \frac{\text{ج (الجزء)}}{\text{ك (الكل)}}$$

مثال (٢ - ٣): افترض أن إجمالي عدد المواليد في الأسبوع هو ١٠٠٠ مولود، وأن عدد المواليد

الذكور ٦٠٠ مولوداً، عدد المواليد الإناث ٤٠٠ مولودة، أحسب نسب المواليد الذكور والإناث.

الحل:

$$\text{نسبة (منسوب) المواليد الذكور} = \frac{60}{100} = 0.6$$

$$\text{نسبة (منسوب) المواليد الإناث} = \frac{40}{100} = 0.4$$

ويمكن تحويل هذه النسب المقدرة إلى نسب مئوية وذلك بضرب كل منها في ١٠٠ كالآتي:

$$\text{نسبة المئوية للمواليد الذكور} = 100 \times \frac{60}{100} = 60\%$$

$$\text{نسبة المئوية للمواليد الإناث} = 100 \times \frac{40}{100} = 40\%$$

٤. **التوزيع التكراري النسبي** Relative Frequency Distribution: يعبر التوزيع النسبي عن ذلك التوزيع الذي يحدد الأهمية النسبية لتكرار كل فئة من إجمالي عدد التكرارات في العينة. ويمكن الاستعانة بجدول التوزيع التكراري الخاص بعدد الموظفين في تقدير ذلك التوزيع النسبي وذلك كما هو الحال بالجدول التالي:

جدول (٢): التوزيع التكراري النسبي لعدد الموظفين على مستوى

٥٠ مطعم في بعض مدن المملكة العربية السعودية

الفئة	التكرار	التكرار النسبي	التكرار المئوي %
٥ -	١٧	٠,٣٤	٣٤
٢٠ -	١٥	٠,٣٠	٣٠
٣٥ -	٨	٠,١٦	١٦
٥٠ -	٣	٠,٠٦	٦
٦٥ -	٤	٠,٠٨	٨
٨٠ -	٣	٠,٠٦	٦
المجموع	٥٠	١,٠٠	١٠٠

$$١ - \text{التكرار النسبي للفئة} = \frac{\text{التكرار المناظر للفئة}}{\text{المجموع الكلي للتكرارات}}$$

$$٢ - \text{التكرار المئوي} = \frac{\text{التكرار المناظر للفئة}}{\text{المجموع الكلي للتكرارات}} \times ١٠٠$$

ويلاحظ أن التوزيع النسبي يفيد في المقارنة ما بين التوزيعات التكرارية المختلفة، حيث لا يمكن إجراء المقارنة اعتماداً على التكرارات المطلقة.

٥. **التوزيع التكراري المتجمع** Cumulative Frequency Distribution:

أحياناً ما يهتم الباحث بمجموع عدد المفردات الأقل أو الأكبر من قيمة معينة. وعلى ذلك فإنه يمكن الحصول على توزيع تكراري متجمع صاعد وتوزيع تكراري متجمع هابط وفيما يلي كيفية إعدادهما:

- التوزيع التكراري المتجمع الصاعد Ascending C.F.D:

التوزيع التكراري المتجمع الصاعد هو التوزيع الذي يهتم بتحديد التكرارات الأقل من قيمة معينة، ويحسب من جدول التوزيع التكراري البسيط وذلك بجمع التكرارات من جهة الفئات الصغيرة إلى الفئات الكبيرة.

- التوزيع التكراري المتجمع الهابط Descending C.F.D:

يهتم هذا التوزيع بتحديد التكرارات الأكبر من قيمة معينة، ويحسب من جدول التوزيع التكراري البسيط وذلك بجمع التكرارات من جهة الفئات الكبيرة إلى الفئات الصغيرة.

ويمكن الاستعانة ببيانات الجدول التكراري (٢) في إعداد التوزيعين وذلك على النحو المبين

بالجدول التالي:

جدول (٣): التوزيع التكراري المتجمع الصاعد والهابط لعدد الموظفين من مثال (١ - ١)

التكرار المتجمع الهابط		التكرار المتجمع الصاعد		التكرار	الفئة
التكرار المتجمع الهابط	الحدود الدنيا للفئات	التكرار المتجمع الصاعد	الحدود العليا للفئات		
٥٠	٥ فأكثر	صفر	أقل من ٥		
٣٣	٢٠ فأكثر	١٧	أقل من ٢٠	١٧	٥ -
١٨	٣٥ فأكثر	٣٢	أقل من ٣٥	١٥	٢٠ -
١٠	٥٠ فأكثر	٤٠	أقل من ٥٠	٨	٣٥ -
٧	٦٥ فأكثر	٤٣	أقل من ٦٥	٣	٥٠ -
٣	٨٠ فأكثر	٤٧	أقل من ٨٠	٤	٦٥ -
صفر	٩٥ فأكثر	٥٠	أقل من ٩٥	٣	٨٠ -
				٥٠	المجموع

ومن الجدول (٣) يتضح أن التوزيع التكراري المتجمع الصاعد يهتم بتحديد التكرارات الأقل من الحد الأعلى لكل فئة. فمثلاً تجد أن الحد الأعلى للفئة الأولى هو ٢٠ وعلى ذلك فإن القيم التي تصغر القيمة ٢٠ تدرج تحت الفئة الأولى، أي أن هناك ١٧ مطعم تقل عدد موظفيهم عن ٢٠، وبالنسبة للفئة الثالثة فإن حدها الأقصى هو ٥٠، وعلى ذلك فإن عدد التكرارات (المطاعم) التي تقل عدد موظفيهم عن ٥٠ يساوي ٤٠ مطعم (١٧ + ١٥ + ٨).. وهكذا حتى نصل إلى الفئة الأخيرة، حيث أن عدد التكرارات الأقل من ٩٥ هو إجمالي التكرارات أي ٥٠ مطعم. بقى أن نشير هنا أننا بدأنا بالفئة (أقل من ٥)، وهي

القيمة التي تمثل الحد الأدنى للفئة الأولى، وذلك بغرض تحديد نقطة بدء المنحنى التكراري المتجمع الصاعد كما سوف نوضح فيما بعد.

كما يتضح من نفس الجدول أن التوزيع التكراري المتجمع الهابط يهتم بتحديد التكرارات التي تساوي أو تكبر الحد الأدنى للفئة، فمثلاً بالنسبة الأولى، نجد أن حدها الأدنى هو ٥، فإذا سألنا عن القيم التي تساوي ٥ أو تكبرها من القيم المعطاة عن الموظفين في الخمسين مطعم لوجدنا أن هنالك ٥٠ قيمة، إما تساوي أو تكبر القيمة ٥. وبالنسبة للقيم التي تساوي ٢٠ أو تكبرها (في الفئة الثانية)، فإنها تساوي إجمالي عدد القيم (٥٠) مخصوماً منها عدد القيم التي تقل عن ٢٠ (٥٠ - ١٧) أي ٣٣ قيمة. وإذا انتقلنا إلى الفئة الثالثة وأردنا أن نعرف عدد القيم التي تساوي أو تكبر القيمة ٣٥، لوجدنا أن جميع القيم التي تقع في الفئة الأولى وعددها ١٧ والقيمة التي تقع في الفئة الثانية وعددها ١٥، كلها قيم أقل من ٣٥، إذا لدينا ٥٠ قيمة منها ٣٢ قيمة أقل من ٣٥، وبذلك يكون لدينا ١٨ قيمة إما تساوي أو تكبر القيمة ٣٥. وهكذا بالنسبة لباقي الفئات، حيث نقوم بخصم التكرارات السابقة للفئة من المجموع الكلي للتكرارات.

بقي أن نلاحظ أنه بالنسبة للفئة الأخيرة والتي حدها الأدنى (٨٠)، هناك ٣ قيم فقط تساوي أو تكبر هذا الحد الأدنى ونظراً لعدم وجود فئات تالية لتلك الفئة، فإننا نسأل عن القيمة التي تساوي أو تكبر حدها الأعلى، ومن الملاحظ عدم وجود أي قيمة من القيم الخمسين تساوي أو تكبر الحد الأعلى للفئة الأخيرة (٩٥).

مثال (٢ - ١): البيانات التالية تبين عدد العمال لـ ٣٠ مصنع في مدينة ما:

١٥	٢٠	٧٠	٢٣	١٨	٣٢	٢٢	٦٣	٥٤	٧١
٣٠	٤٠	٨٠	٣٢	١٩	٧٣	٤٢	٢٧	٥٥	٨٠
٥٠	٦٠	٨٢	٤٣	٦٥	٨٤	٥٢	٤٤	٧٣	٥٢

والمطلوب: إعداد جدول التوزيع التكراري المتجمع الصاعد والهابط.

الحل

تقوم أولاً بإعداد جدول التوزيع التكراري وفقاً للخطوات المتبعة في المثال الأول وذلك على النحو

التالي:

(١) تحديد المدى الكلي: وهو الفرق بين أصغر قيمة وأكبر قيمة في الجدول السابق. إذاً

$$\text{فالمدى} = ٨٤ - ١٥ = ٦٩.$$

- (٢) تحديد عدد الفئات: نظراً لأن عدد القيم المعطاة في هذا المثال صغيره (٣٠ مفردة)، حيث سنقوم بتوزيع القيم المعطاة على ٥ فئات، ويمكن زيادتها إلى ست أو سبع فئات، إلا أننا سوف نقوم بتوزيع القيم على خمس فئات، وذلك كما هو مبين بالجدول التالي.
- (٣) تحديد التكرارات التي تقل عن الحد الأعلى لكل فئة للحصول على التوزيع التكراري المتجمع الصاعد.
- (٤) تحديد التكرارات التي تساوي أو تكبر الحد الأدنى لكل فئة للحصول على التوزيع التكراري المتجمع الهابط.

جدول (٤): التوزيع التكراري المتجمع الصاعد والهابط لعدد العمال لـ ٣٠ مصنع في مدينة ما

الفئة	التكرار	التكرار المتجمع الصاعد		التكرار المتجمع الهابط	
		الحدود العليا للفئات	التكرار المتجمع الصاعد	الحدود الدنيا للفئات	التكرار المتجمع الهابط
		أقل من ١٥	صفر	١٥ فأكثر	٣٠
١٥ -	٧	أقل من ٣٠	٧	٣٠ فأكثر	٢٣
٣٠ -	٧	أقل من ٤٥	١٤	٤٥ فأكثر	١٦
٤٥ -	٥	أقل من ٦٠	١٩	٦٠ فأكثر	١١
٦٠ -	٧	أقل من ٧٥	٢٦	٧٥ فأكثر	٤
٧٥ -	٤	أقل من ٩٠	٣٠	٩٠ فأكثر	صفر
المجموع	٣٠	_____	_____	_____	_____

تمارين

س١: ذكر أهمية تبويب البيانات في مهمة الباحث الإحصائي؟

س٢: أذكر أهم أساليب تبويب البيانات؟

س٣: أفترض أن عدد الطلاب في ٤٠ معهد هي كالتالي:

١٥٧	١٤٩	١٣٨	١٤٤	١٢٣	١٥٠	١٦٤	١٣٨
١٤٤	١٥٢	١٤٨	١٤٦	١٤٧	١٤٠	١٥٨	١٤٦
١٦٥	١٥٤	١٢٩	١٦٣	١٧٦	١٣٨	١٢٧	١٦٨
١٣٥	١٤٠	١٥٣	١٣٥	١٤٧	١٤٢	١٧٣	١٢٦
١٢٨	١٤٥	١٥٣	١٥٠	١٤٢	١٣٥	١٥٥	١٦١

المطلوب: إعداد جدول التوزيع التكراري.

س٤: إذا كانت أعمار طلبة الشعبة الأولى من القسم البيئي على النحو التالي:

٢١	٢٠	٢١	٢١	٢٠	٢٠	٢١	٢٦	١٨	١٧
١٩	٢٠	٢٢	٢٢	٢٣	١٨	١٧	٢٦	١٧	١٨
١٨	٢١	٢٤	٢٣	٢٣	١٩	١٨	٢٥	٢٠	١٩
١٨	٢٦	٢٦	٢٤	٢٠	٢٦	١٩	٢٤	٢٢	٢٠
١٧	٢٢	٢٥	٢٥	٢٤	٢٥	٢١	٢٢	٢٤	٢١

المطلوب:

أ - إعداد التوزيع التكراري المتجمع الصاعد.

ب - إعداد التوزيع التكراري المتجمع النازل.

ج - إعداد التوزيع التكرار النسبي.

مقدمة في الإحصاء

التمثيل البياني للبيانات



الوحدة الرابعة

اسم الوحدة : التمثيل البياني للبيانات.

الجدارة : معرفة رسم و تمثيل البيانات الإحصائية بيانياً.

الأهداف :

- ١ - معرفة الأنواع المختلفة من التمثيل (العرض) البياني.
- ٢ - المقدرة على التمثيل أو الرسم البياني.
- ٢ - معرفة طرق الحصول على الرسوم البيانية للبيانات الإحصائية إما بالرسم باليد (باستخدام الأدوات الهندسية) أو عن طريق الإكسل أو spss وغيرها من البرامج الإحصائية.

مستوى الأداء المطلوب : ٩٥٪.

الوقت المتوقع للتدرب على الجدارة : ١٠ ساعات.

الوسائل المساعدة على تحقيق الجدارة :

١. الرجوع إلى مراجع أخرى.
٢. المناقشة.
٣. حل التمارين.
٤. واجبات منزلية.
٥. استخدام بعض برامج الكمبيوتر الإحصائية.
٦. استخدام الهدبروجيكتور.

متطلبات الجدارة :

فهم الوحدات السابقة.

التمثيل البياني للبيانات

Graphical Representation of Data

مقدمة :

يمثل العرض البياني وسيلة من وسائل تلخيص البيانات الإحصائية وهي أكثر من الجداول لفتاً للنظر وأيسر منها في الفهم والوضوح، ولكنها لا يمكن أن تقوم بكل ما تقوم به الجداول الإحصائية كترتيب وتنظيم مجموعة كبيرة من الأرقام في شكل جدول ليسهل رسمه وتحليل بياناته. وحيث أن الرسوم البيانية تشير إلى أعداد تقريبية فإنها تعطي صورة سريعة للموقف العام. ولقد شاع استخدام العرض البياني وتنوعت أساليبه. وهناك أنواع عديدة من الرسوم البيانية والتي يمكن ذكرها فيما يلي:

١. العرض البياني للتوزيعات التكرارية :

يمكن عرض التوزيع التكراري بيانياً لإعطاء صورة واضحة عن خصائص هذا التوزيع الرئيسية بحيث نقوم بوضع المتغير (الفئات) على المحور الأفقي ونضع التكرارات على المحور الرأسي ويلاحظ أن تكون وحدات القياس على كل محور متناسقة مع بعضها البعض فمثلاً إذا كان ابتعادنا عن نقطة الصفر على المحور الأفقي بمسافة ٥ سم فيجب أن يكون ابتعاد كل نقطة عن النقطة التي تليها مساوياً لـ ٥ سم وهكذا.

ويمكن عرض التوزيع التكراري بيانياً عن طريق إحدى الطرق التالية:

- المدرج التكراري Histogram
- المضلع التكراري Frequency Polygon
- المنحنى التكراري المتجمع الصاعد Cumulative Frequency Curve
- المنحنى التكراري المتجمع النازل Decumulative Frequency Curve

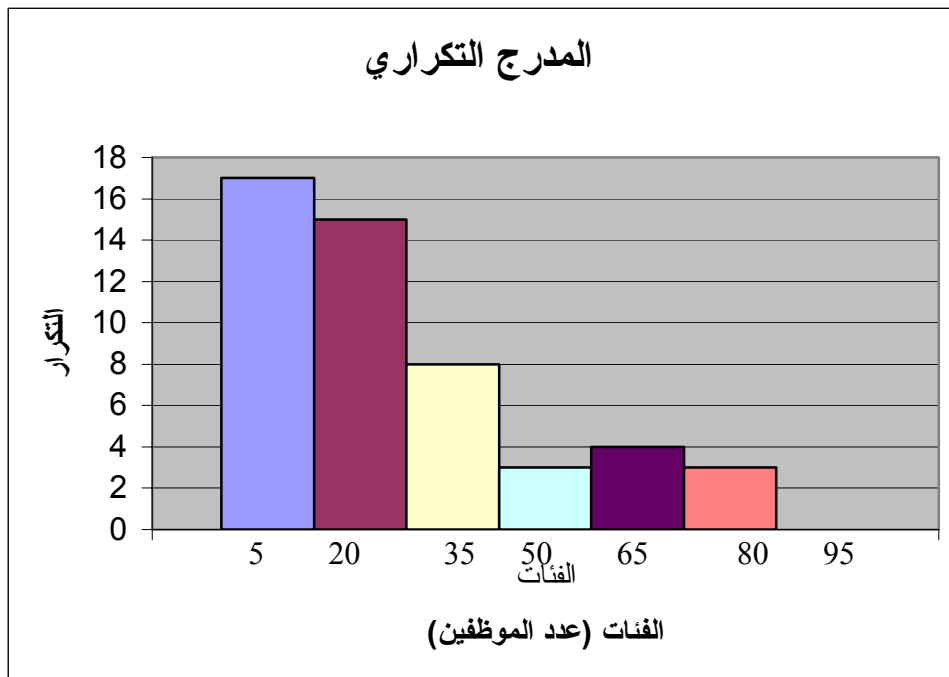
أولاً: المدرج التكراري:

المدرج التكراري عبارة عن شكل بياني يحتوي على مستطيلات متلاصقة تتناسب مساحتها مع تكرارات الفئات وبما أن طول الفئات متساوي فإن قواعد المستطيلات متساوي وارتفاع المستطيلات يتناسب مع التكرارات.

ولرسم المدرج التكراري نرسم محورين متعامدين ثم نضع الفئات على المحور الأفقي والتكرارات على المحور الرأسي ويلاحظ كما أشرنا سابقاً أنه يجب تقسيم كل من المحورين إلى تقسيمات متساوية ويغطي المحور الرأسي أكبر تكرار موضح بالجدول التكراري ونقيم بعد ذلك على طول كل فئة مستطيلاً يتناسب مع تكرارها أو نقيم أعمدة من الحدود الدنيا للفئات ارتفاعها يساوي عدد التكرار المقابل لكل فئة ثم نقوم بتوصيل هذه الأعمدة بخطوط أفقية. ولإيضاح كيفية رسم المدرج التكراري دعنا نستعيد التوزيع التكراري في الجدول رقم (١) من مثال (١ - ١)

الفئات	التكرار
٥ - ١٩	١٧
٢٠ - ٣٤	١٥
٣٥ - ٤٩	٨
٥٠ - ٦٤	٣
٦٥ - ٧٩	٤
٨٠ - ٩٤	٣
المجموع	٥٠

شكل رقم (١) المدرج التكراري لعدد الموظفين في خمسين مطعم



يلاحظ أن لدينا ست فئات لكل فئة مستطيل يغطي هذه الفئة حيث نجد أن مساحة كل مستطيل تساوي الفئة (١٥) × عدد التكرار المقابل لها أي أن:

- مساحة المستطيل (١) = 17×15

- مساحة المستطيل (٢) = 15×15

- مساحة المستطيل (٣) = 8×15

- مساحة المستطيل (٤) = 3×15

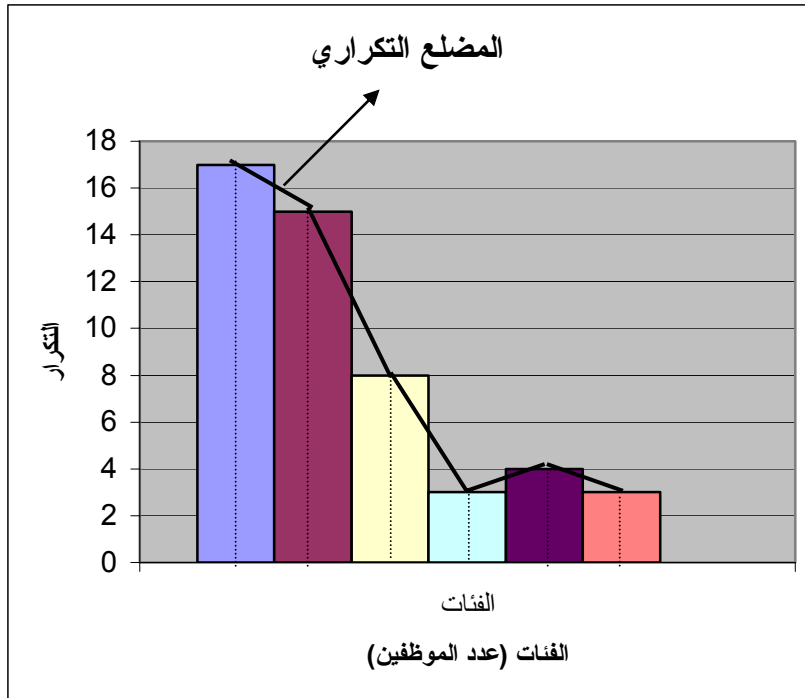
- مساحة المستطيل (٥) = 4×15

- مساحة المستطيل (٦) = 3×15

ثانياً: المضلع التكراري:

المضلع التكراري عبارة عن خط بياني يتكون من مستقيمات تصل بين النقاط التي تأتي في منتصف القاعدة العليا للمستطيل التي تتصف الفئات في المدرج التكراري. ولإيضاح كيفية رسم المضلع التكراري وتلافياً للتشويش الذي قد يحدث سوف نعيد رسم المدرج التكراري السابق هنا.

شكل رقم (٢) المضلع التكراري لعدد الموظفين في خمسين مطعم



يلاحظ من الشكل المضلع التكراري عبارة عن مستقيمات تصل بين منتصف القاعدة العليا لكل مستطيل ولكن يلاحظ أننا نبدأ من نقطة على المحور الأفقي تكرارها صفر وننتهي بنقطة على ذات المحور تكرارها صفر أيضاً.

ويمكن بدلاً من رسم المدرج التكراري تحديد مركز كل فئة = $\frac{\text{الحد الأدنى} + \text{الحد الأعلى}}{2}$

٢

ثم نحدد النقط عند مركز الفئة ثم نقوم بتوصيل تلك النقط بمستقيمات تمثل في مجملها المضلع التكراري.

مثال (١ - ٤): افترض أن لدينا جدول تكراري يمثل عدد العمال الذين يعملون في ٥٠ محل تجاري على النحو التالي:

جدول (٥) عدد العمال الذين يعملون في ٥٠ محل تجاري

مراكز الفئات	التكرار	الفئات
١٥	١٢	١٠ -
٢٥	١٥	٢٠ -
٣٥	١٠	٣٠ -
٤٥	٨	٤٠ -
٥٥	٥	٥٠ -
- - -	٥٠	المجموع

المطلوب: رسم المضلع التكراري باستخدام مركز الفئة
الحل:

أولاً: يتم إيجاد مركز الفئة كالتالي:

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

وبالتالي فإن

$$15 = \frac{20 + 10}{2} = \text{مركز الفئة (١)}$$

$$25 = \frac{30 + 20}{2} = \text{مركز الفئة (٢)}$$

$$35 = \frac{40 + 30}{2} = \text{مركز الفئة (٣)}$$

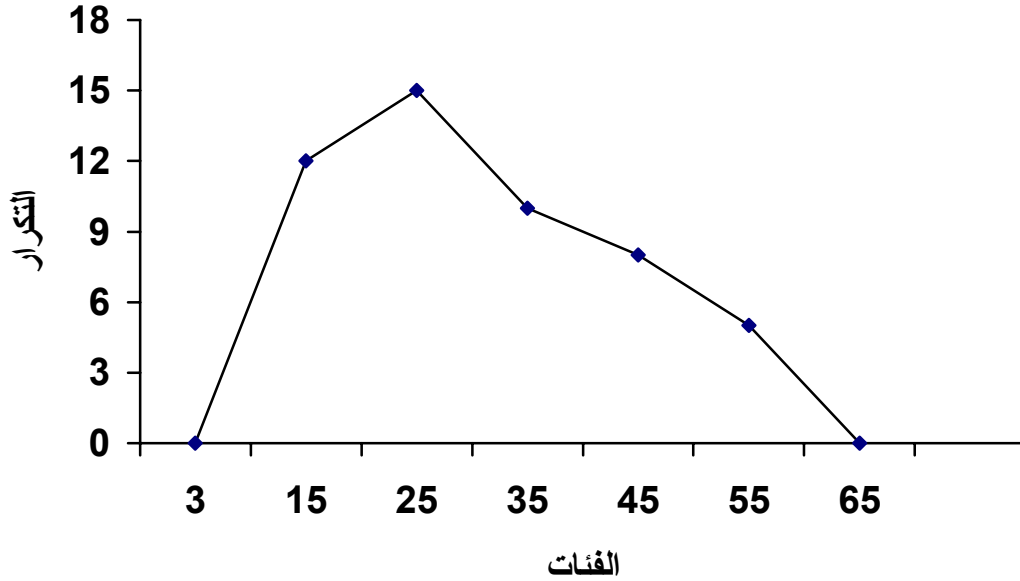
$$45 = \frac{50 + 40}{2} = \text{مركز الفئة (٤)}$$

$$55 = \frac{60 + 50}{2} = \text{مركز الفئة (٥)}$$

ثانياً: تمثيل نقاط مراكز الفئات.

ثالثاً: توصيل النقاط مع بعضها لنحصل على المضلع التكراري

شكل رقم (٣) المضلع التكراري لعدد العمال في ٥٠ محل تجاري



ثالثاً: المضلع التكراري المتجمع الصاعد:

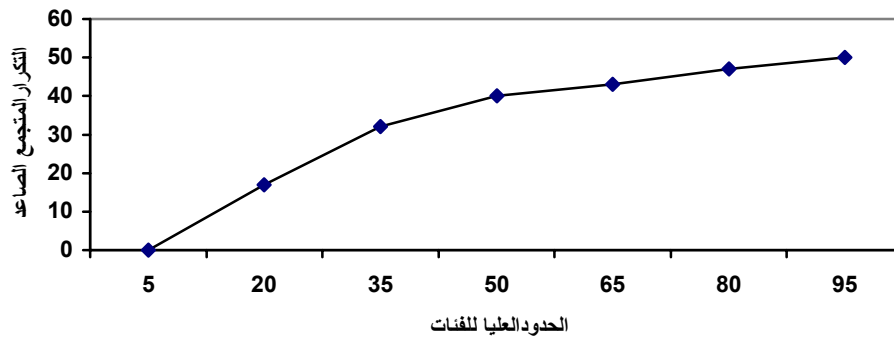
المنحنى التكراري المتجمع الصاعد هو الرسم البياني الذي يمثل التوزيع التكراري المتجمع الصاعد ويمكن توضيح كيفية رسم المنحنى التكراري المتجمع الصاعد بالنظر إلى جدول (١) لعدد الموظفين في ٥٠ مطعم كالتالي:

جدول (٦) جدول تكراري متجمع صاعد يمثل عدد العمال الذين يعملون في ٥٠ محل تجاري

الفئات	التكرار	الحدود العليا للفئات	التكرار المتجمع الصاعد
٥ -	١٧	أقل من ٥	صفر
٢٠ -	١٥	أقل من ٢٠	١٧
٣٥ -	٨	أقل من ٣٥	٣٢
٥٠ -	٣	أقل من ٥٠	٤٠
٦٥ -	٤	أقل من ٦٥	٤٣
٨٠ -	٣	أقل من ٨٠	٤٧
		أقل من ٩٥	٥٠

المنحنى التكراري المتجمع الصاعد هو الرسم البياني لجدول التوزيع التكراري المتجمع الصاعد ويمكن رسم المنحنى التكراري المتجمع الصاعد عن طريق رسم محورين متعامدين على المحور الأفقي تكتب الفئات وعلى المحور الرأسي تكتب التكرارات ويتم تحديد نقاط التقاء الحدود العليا للفئات مع التكرارات المقابلة لها إلا أننا نحتاج إلى أن نبدأ بالحد الأدنى للفئة الأولى حتى نحدد انطلاق المنحنى من على المحور الأفقي.

شكل رقم (٤) المنحنى التكراري المتجمع الصاعد لعدد الموظفين في خمسين مطعم



نجد أن:

- النقطة (١) تمثل أقل من ٥ وتكرارها صفر فيلتقيان على المحور الأفقي.
- النقطة (٢) تمثل أقل من ٢٠ وتكرارها ١٧ فيلتقيان عند النقطة ١٧ على المحور الرأسي والنقطة التي تلاصق ولكن تقل عند العشرين على المحور الأفقي.
- النقطة (٣) تمثل أقل من ٣٥ وتكرارها ٣٢ ويلتقيان عند النقطة ٣٢ على المحور الرأسي والنقطة التي تلاصق ولكن تقل عن ٣٥ على المحور الأفقي وهكذا لجميع النقط على المنحنى التكراري المتجمع الصاعد.

رابعاً: المنحنى التكراري المتجمع النازل:

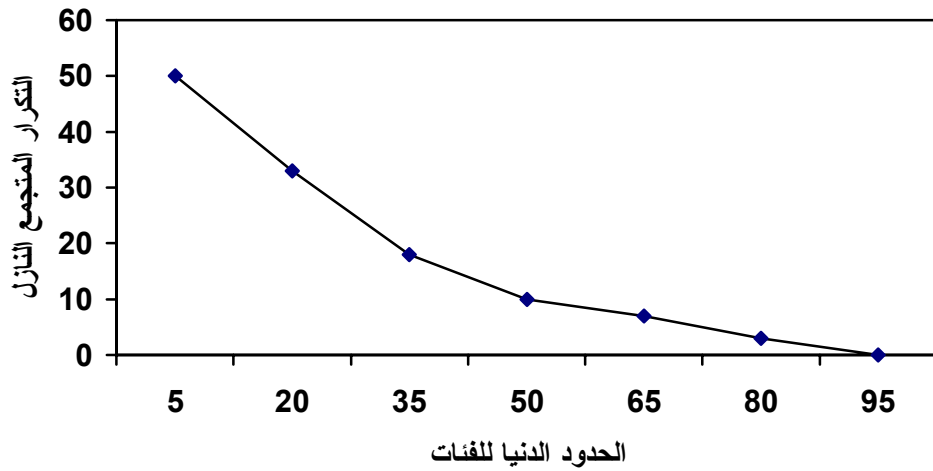
المنحنى التكراري المتجمع النازل هو الرسم البياني الذي يمثل التوزيع التكراري المتجمع النازل ويمكن توضيح كيفية رسم المنحنى التكراري المتجمع النازل بالنظر إلى الجدول التالي:

جدول (٧) جدول تكراري متجمع نازل يمثل عدد العمال الذين يعملون في ٥٠ محل تجاري

الفئات	التكرار	الحدود الدنيا للفئات	التكرار المتجمع النازل
٥ -	١٧	من ٥ فأكثر	٥٠
٢٠ -	١٥	من ٢٠ فأكثر	٣٣
٣٥ -	٨	من ٣٥ فأكثر	١٨
٥٠ -	٣	من ٥٠ فأكثر	١٠
٦٥ -	٤	من ٦٥ فأكثر	٧
٨٠ -	٣	من ٨٠ فأكثر	٣
		من ٩٥ فأكثر	صفر

وبتمثيل هذا الجدول بيانياً نستطيع الحصول على المنحنى التكراري المتجمع النازل

شكل رقم (٥) المنحنى التكراري المتجمع النازل لعدد الموظفين في خمسين مطعم



حيث يلاحظ أن النقطة (١) على المنحنى التكراري المتجمع النازل توضح عدد التكرارات التي من ٥ فأكثر بينما توضح النقطة (٢) عدد المطاعم التي تضم ٣٣ موظفاً فأكثر وهكذا حتى نصل إلى النقطة (٧) على نفس المنحنى والتي تدل على أنه لا يوجد أي مطعم بها ٩٥ موظف فأكثر أي أن تكرارها صفر في هذه الحالة.

مثال (٢-٤): يمثل الجدول التالي الأجر الشهري لـ ٥٨ موظف بألف الريالات في إحدى الشركات. والمطلوب رسم المنحى التكراري المتجمع الهابط.

جدول (٨) تكراري يمثل الأجر الشهري لـ ٥٨ موظف في إحدى الشركات

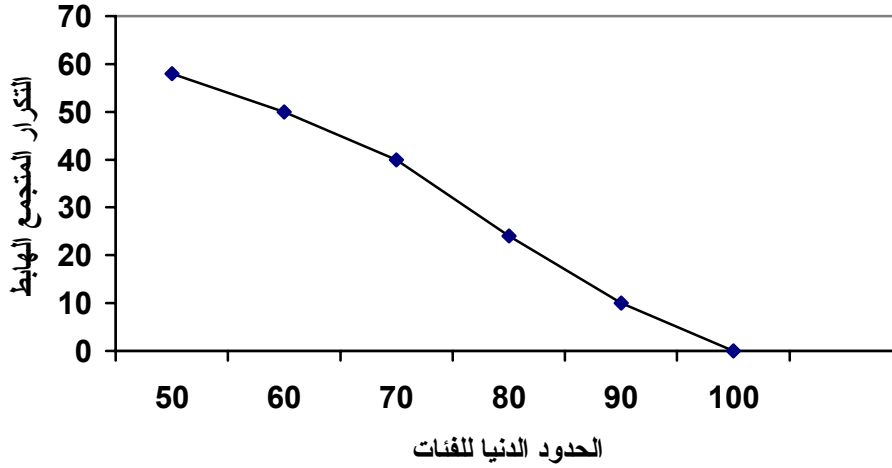
التكرار	الفئات (الأجر الشهري)
٨	٥ -
١٠	٦ -
١٦	٧ -
١٤	٨ -
١٠	٩ -
٥٨	المجموع

ومعنى هذا الجدول أن هنالك ٨ موظفين تقع رواتبهم من ٥ إلى أقل من ٦ ألف ريال وهكذا لباقي الفئات حيث يوضح التكرار عدد الموظفين الذين تقع رواتبهم داخل الفئة المقابلة. من الجدول السابق نستطيع اشتقاق الجدول الذي يوضح التوزيع التكراري المتجمع الهابط.

جدول (٩) التوزيع التكراري المتجمع (النازل) الهابط

الحدود الدنيا للفئات	التكرار المتجمع الهابط
من ٥ فأكثر	٥٨
من ٦ فأكثر	٥٠
من ٧ فأكثر	٤٠
من ٨ فأكثر	٢٤
من ٩ فأكثر	١٠
من ١٠ فأكثر	صفر

شكل رقم (٦) المنحنى التكراري المتجمع النازل لرواتب الموظفين

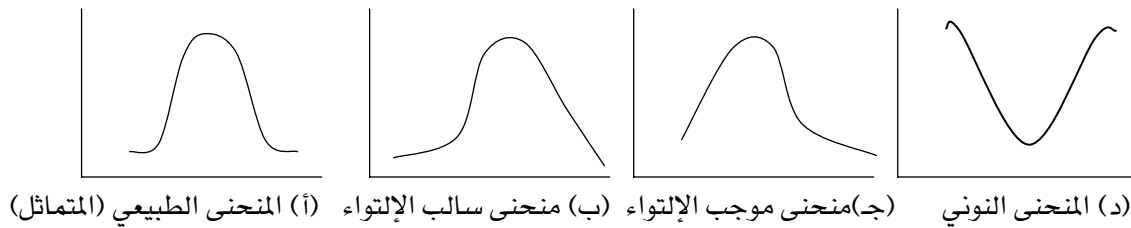


يلاحظ هنا أننا عند رسم المنحنى التكراري المتجمع الهابط لا نحتاج إلى توزيع المحور الأفقي إلى مسافات تساوي المسافة التي بين نقطة الصفر وأول نقطة على هذا المحور. بل نحتاج إلى كتابة بداية الفئات بحيث نبدأ بالحد الأدنى للفئة الأولى.

٢. الخطوط والمنحنيات:

والتي تستخدم لتمثيل التوزيعات التكرارية والسلاسل الزمنية. ومن بعض المنحنيات:

- بعض المنحنيات التكرارية المشهورة Some Frequency Curve:



شكل رقم (٧) بعض أنواع المنحنيات التكرارية

■ المنحنى الطبيعي أو المتماثل symmetric curve: شكل (٧ - أ) يمثل المنحنى الطبيعي كثير من الظواهر الطبيعية مثل أطوال الناس، ونسب الذكاء ... كما يمثل بعض الظواهر التجريبية في كثير من الدراسات العملية، ومن أهم مميزات المنحنى الطبيعي (المتماثل) أن

العمود النازل من أعلى تكرر في قمته على محور الفئات (الأفقي) يقسم المساحة الواقعة تحت المنحنى إلى قسمين متساويين. وهذا بسبب القيم الظاهرة تتركز في الوسط وتقل بنفس المعدل كلما اتجهنا نحو الطرفين.

■ المنحنى النوني أو منحنى (U): شكل (٧ - د) يشبه المنحنى الطبيعي (المتماثل) مقلوبا أي أن التكرارات تكون كبيرة في اتجاه الطرفين وتقل كلما اتجهنا إلى وسط القراءات. وأهم الظواهر التي تبدو بهذا الشكل اعداد الوفيات حسب العمر حيث تكثر الوفيات في الطفولة والشيخوخة وتقل في متوسط العمر.

■ المنحنيات الملتوية: شكل (٧ - ب - ج) والمنحنيات الملتوية تشبه المنحنى الطبيعي من حيث أن لها قمة واحدة ، ولكن إذا كان طرف المنحنى الأيمن أطول من طرفه الأيسر يسمى بالمنحنى الموجب الالتواء (أي ملتوي جهة اليمين) *positively skewed curve* أما إذا كان طرف المنحنى الأيسر أطول من طرفه الأيمن فيسمى سالب الالتواء (أي ملتوي نحو اليسار) *negatively skewed curve*.

٣. الأعمدة البيانية Bar Charts :

وتستعمل هذه الطريقة لتمثيل المتغيرات النوعية بحيث تمثل كل فئة من فئات المتغير بمستطيلات أو أعمدة بيانية ذات قواعد متماثلة ، بحيث تتناسب مساحة كل مستطيل أو عمود مع القيمة التي يمثلها كل من قاعدته مع طول الفئة. ويكون هدف تمثيل البيانات هو إظهار قيمة كل فئة من البيانات مع توضيح نسب الفئات بعضها إلى بعض وإلى جملة قيمة المتغير. وتشمل الأعمدة البيانية ثلاثة أنواع هي: الأعمدة المنفردة - الأعمدة المزدوجة - الأعمدة المجزأة.

- الأعمدة المنفردة :

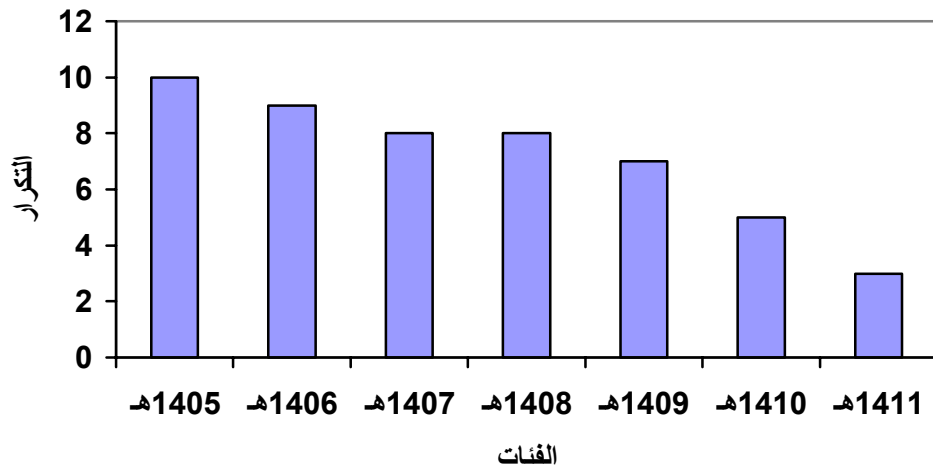
مثال (٣ - ٤): إذا كان عدد المخالفات في عدة محلات تجارية في أحد الأحياء في مدينة ما سنة ١٤٠٥ هـ حتى ١٤١١ هـ وكانت كالتالي:

جدول (١٠) عدد المخالفات في عدة محلات تجارية في أحد الأحياء في مدينة ما لعدة سنوات

الفئات (السنوات)	التكرار (عدد المخالفات)
١٤٠٥ هـ	١٠
١٤٠٦ هـ	٩
١٤٠٧ هـ	٨
١٤٠٨ هـ	٨
١٤٠٩ هـ	٧
١٤١٠ هـ	٥
١٤١١ هـ	٣
المجموع	٥٠

وبتمثيل هذا الجدول بيانياً نستطيع الحصول على الأعمدة المنفردة كالتالي:

شكل رقم (٨) الأعمدة المنفردة لعدد المخالفات من ١٤٠٥ - ١٤١١ هـ



- الأعمدة المزدوجة والأعمدة المجزأة:

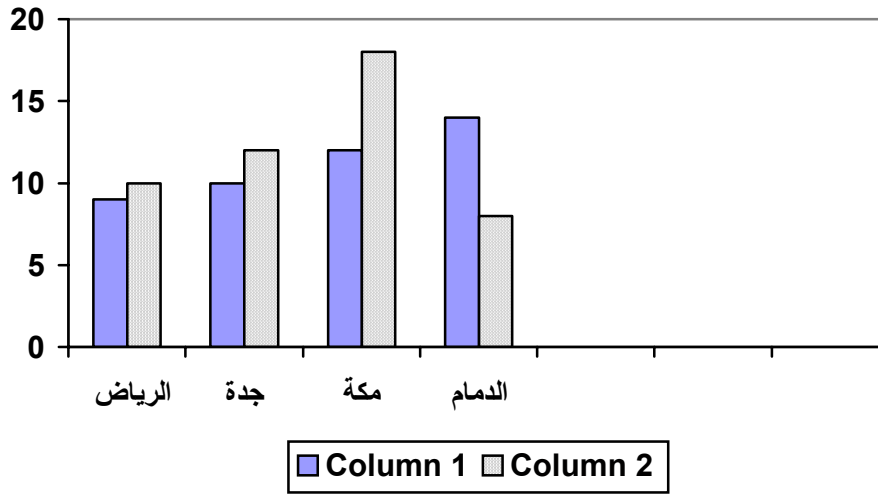
مثال (٤ - ٤): إذا كان عدد بعض المصانع المخالفة وغير المخالفة في عدة مدن كالتالي:

جدول (١١) عدد بعض المصانع المخالفة وغير المخالفة في عدة مدن

الحالة \ المدينة	الرياض	جدة	مكة	الدمام
غير مخالفة	٩	١٠	١٢	١٤
مخالفة	١٠	١٢	١٨	٨

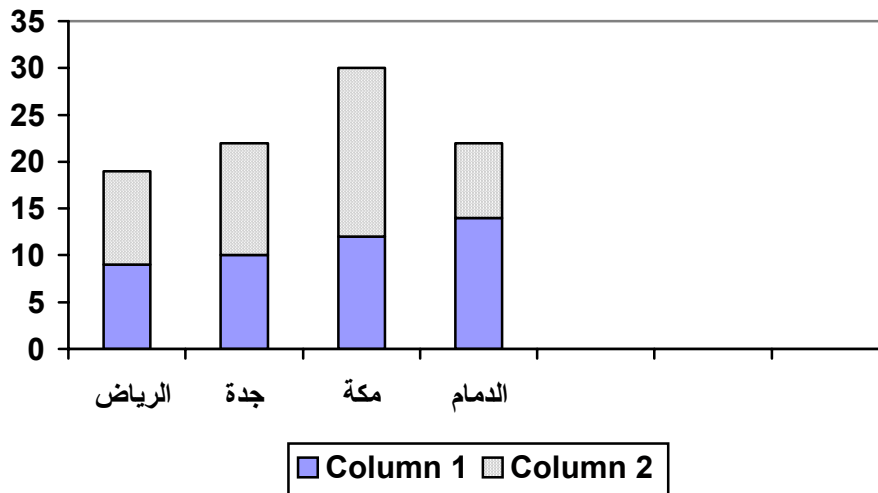
وبتمثيل هذا الجدول بيانياً نستطيع الحصول على الأعمدة المزدوجة والمجزئة كالتالي:

شكل رقم (٩) الأعمدة المزدوجة للمصانع المخالفة وغير المخالفة في بعض المدن



حيث أن العمود ١ يمثل عدد المصانع الغير المخالفة. وأن العمود ٢ يمثل عدد المصانع المخالفة.

شكل رقم (١٠) الأعمدة المزدوجة للمصانع المخالفة وغير المخالفة في بعض المدن



حيث أن الجزئ السفلي ١ يمثل عدد المصانع الغير المخالفة. وأن الجزئ العلوي ٢ يمثل عدد المصانع المخالفة.

٤. العرض البياني بواسطة الدوائر Pie Charts:

أحياناً ما تمثل الإحصاءات بقطاعات دائرة. ويمكن هنا أن نميز ما بين نوعين من الدوائر المستخدمة في عرض البيانات وذلك على النحو التالي:

(١) الدوائر المختلفة. (٢) الدائرة المقسمة.

وسوف نستعرض فيما يلي كل منهما على حدة.

(١) الدوائر المختلفة:

لتوضيح كيفية استخدام الدوائر المختلفة في عرض البيانات من المثال الآتي.

مثال (٥ - ٤): افترض أن باحثاً قام بجمع البيانات عن عدد المخالفات للمحلات في خمس مدن مختلفة خلال أسبوع وذلك على النحو التالي:

جدول (١٢) عدد المخالفات للمحلات على مستوى خمس مدن مختلفة

المدينة	عدد مخالفات المحلات
الرياض	١٠٠
بريدة	٢٥
جدة	٨١
أبها	١٦
الدمام	٦٤

والمطلوب: تمثيل هذه البيانات بطريقة الدوائر المختلفة.

لتمثيل هذه البيانات نتبع الخطوات التالية:

أولاً: نوجد الجذر التربيعي للبيانات المقابلة لكل مدينة وذلك على النحو التالي:

$$\sqrt{100} = 10 \quad \text{الرياض :}$$

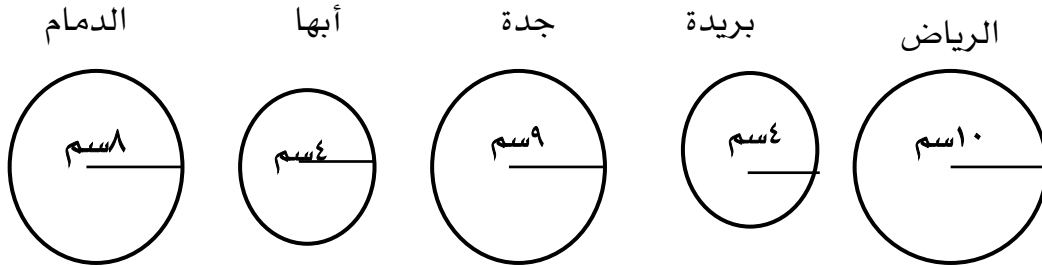
$$\sqrt{25} = 5 \quad \text{بريدة :}$$

$$\sqrt{81} = 9 \quad \text{جدة :}$$

$$\sqrt{16} = 4 \quad \text{أبها :}$$

$$\sqrt{64} = 8 \quad \text{الدمام :}$$

ثانياً: رسم دوائر مختلفة بحيث يكون لكل مدينة دائرة مستقلة نصف قطرها يساوي الجذر التربيعي المقابل لها.



يلاحظ أنه ليس من الضروري أن تكون الدائرة ذات نصف قطر مساوياً تماماً للجذر التربيعي المقابل، ولكن يجب أن تكتب قيمة نصف القطر داخل الدائرة كما يجب أن تكون مساحة الدائرة ذات النصف قطر الأكبر أكبر من مساحة الدائرة ذات النصف قطر الأقل، وهكذا إذ ليس من المعقول أن تكون الدائرة التي كتب داخلها ١٠ اسم أصغر حجماً للمشاهد من الدائرة التي كتب بداخلها ٤ اسم.

(٢) الدائرة المقسمة:

تمثل الإحصاءات بقطاعات دائرة، بحيث توضح الأجزاء المكونة للمجموع بقطاعات تتناسب مع نسبة كل فئة لجزء من المجموع الذي تشير إليه الدائرة بأكملها. ويمثل هذا النوع من الدوائر النوع الأكثر انتشاراً واستعمالاً في عرض البيانات ولتوضيح كيفية تمثيل البيانات بطريقة الدوائر المقسمة.

مثال (نعرض الجدول رقم (١١) إذا تقدم عدد من الأشخاص لفحص فصيلة الدم وكان ٧٠ شخص منهم فصيلة الدم لديهم (O) وكانت فصيلة الدم (A) لـ ٢٠ شخص، وكان ٣٠ شخص فصيلة دمهم (AB).

للأشخاص المتقدمين للفحص (AB, A, O) جدول (١٣): فصيلة الدم

فصيلة الدم	الأشخاص المتقدمين	
	عدد	النسبة
O	٧٠	٥٨,٣
A	٢٠	١٦,٧
AB	٣٠	٢٥
المجموع	١٢٠	%١٠٠

عند الرغبة في تمثيل هذه البيانات بطريقة الدائرة المقسمة يلزم مراعاة التالي:

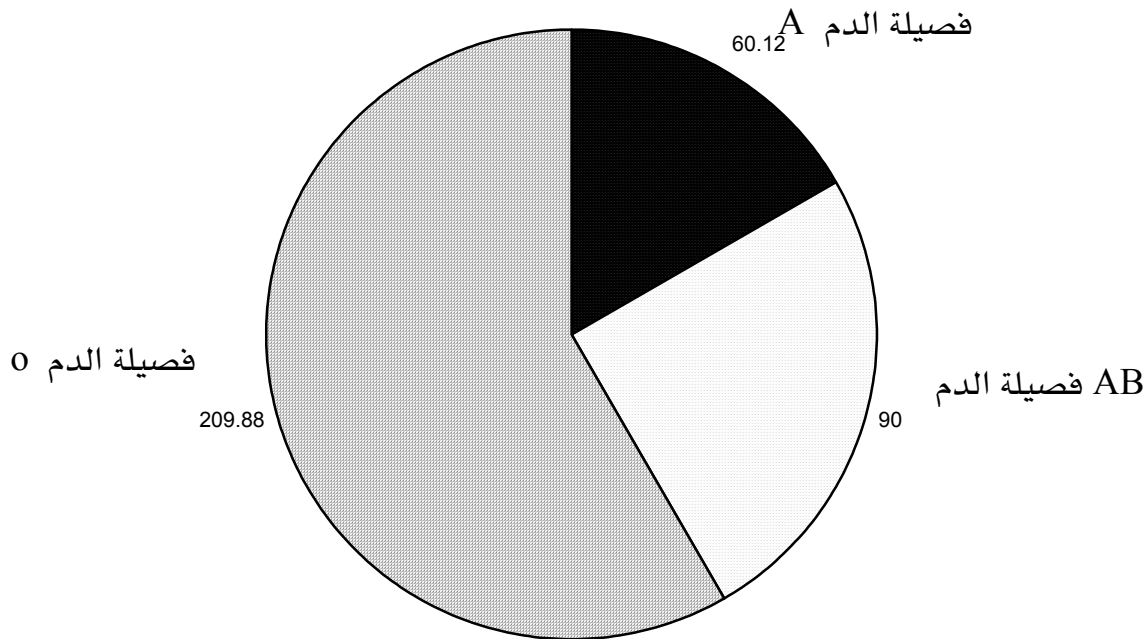
بما أن مقدار الزاوية التي رأسها عند مركز الدائرة يساوي 360° درجة وبما أن إجمالي مساحة الدائرة تساوي 100% من المكان الذي تشغله الدائرة، فإن لكل 1% من مساحة الدائرة $3,6^\circ$ درجة وبالتالي فإن حصة كل نوع من فصائل الدم من الدائرة يمكن تقديره على النحو التالي:

(أ) مقدار الزاوية للأشخاص ذو فصيلة الدم O $= 3,6 \times 58,3 = 209,88^\circ$ درجة.

(ب) مقدار الزاوية للأشخاص ذو فصيلة الدم A $= 3,6 \times 16,7 = 60,12^\circ$ درجة.

(ج) مقدار الزاوية للأشخاص ذو فصيلة الدم AB $= 3,6 \times 31,6 = 90^\circ$ درجة.

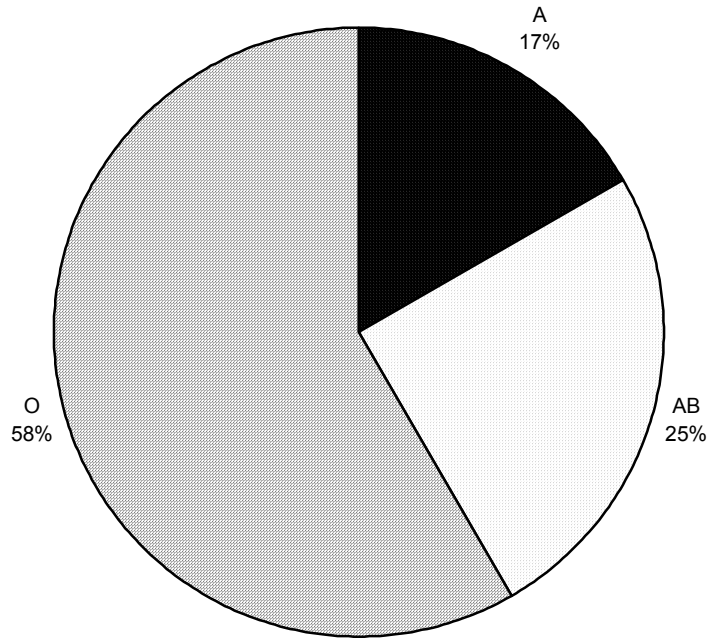
ويمكن بعد ذلك تقسيم الدائرة ذات 360° درجة بحيث يعطي كل نوع من فصائل الدم ما يقابله من الدرجات المخصصة.



شكل رقم (١١) الدائرة المقسمة لفصائل الدم لـ ١٢٠ شخص حسب الدرجة

ويمكن بدلاً من توزيع الدائرة عن طريق الدرجات توزيعها عن طريق النسب المئوية حيث مساحة الدائرة هي 100% من الحيز الذي تشغله وعليه يتم توزيع الدائرة الكاملة على فصائل الدم كل على حسب النسبة المقابلة له.

شكل رقم (١٢) الدائرة المقسمة لفصائل الدم لـ ١٢٠ شخص حسب النسبة المئوية



- فصيلة الدم O تشغل ٥٨ ٪ من مساحة الدائرة.
- فصيلة الدم A تشغل ١٧ ٪ من مساحة الدائرة.
- فصيلة الدم AB تشغل ٢٥ ٪ من مساحة الدائرة.

تمارين

س١: إذا كانت درجات ٤٠ طالباً في مادة الإحصاء هي:

٢٧	٢٥	١٧	١٩	٢٩	٢٦	٢٥	١٦	٣٦	٤٢
٣٠	١٥	٣٩	٣٦	٤٤	٣٨	٤١	٢٠	٢١	٢٩
٣٨	٤٤	٣٢	٢٤	٢٢	٣١	٣٣	٢٦	٢٢	٣٤
٣١	٢٣	٢٧	٢٨	٢٣	٣٣	٢٧	٣٤	٢٣	٣٦

المطلوب: تمثيل البيانات في:

- أ - جدول تكراري يتكون من ٦ فئات.
- ب - رسم المدرج التكراري.
- ج - رسم المضلع التكراري.
- د - رسم المنحنى التكراري المتجمع الصاعد.
- هـ - رسم المنحنى التكراري المتجمع النازل.

س٢: إذا كانت تقديرات طلبة السنة الأولى كالتالي:

التقدير	ممتاز	جيد جداً	جيد
عدد الطلبة	١٦	٢٥	٢٠

المطلوب: تمثيل تقادير الطلبة باستخدام:

- ١ - الأعمدة المنفردة.
- ٢ - الدوائر المختلفة.
- ٣ - الدائرة المقسمة حسب الدرجات (٣٦٠ درجة).
- ٤ - الدائرة المقسمة حسب النسبة (الدائرة ١٠٠٪).

س٣: أين يقع أكثر تكرار في كل من المنحنى الطبيعي (المتماثل) والمنحنى النوني (U) مع ذكر مثال يطبق لكلاهما؟

س٤: الجدول التالي يبين عدد السعوديين وغير السعوديين في أحد مزارع الألبان:

السنة الهجرية	١٤٠٠	١٤٠١	١٤٠٢	١٤٠٤	١٤٠٥
العاملين					
سعوديين	٦	١٠	٢٠	٤٠	٦٠
غير سعوديين	٣٠	٤٠	٥٠	٥٥	٤٥

المطلوب: تمثيل البيانات في:

أ - رسم الأعمدة المزدوجة.

ب - رسم الأعمدة المجزئة.

مقدمة في الإحصاء

مقاييس النزعة المركزية



الوحدة الخامسة

اسم الوحدة : مقاييس النزعة المركزية.

الجدارة : معرفة مقاييس النزعة المركزية وتطبيقاتها.

الأهداف :

- ١ - المقدرة على إيجاد مقاييس النزعة المركزية (الوسط , والوسيط , والمنوال).
- ٢ - الإلمام بمعرفة الخصائص الهامة لمقاييس النزعة المركزية والترابط بينهم.
- ٣ - دراسة طرق إستخراج مقاييس النزعة المركزية.

مستوى الأداء المطلوب : ٩٥٪.

الوقت المتوقع للتدرب على الجدارة : ٩ ساعات.

الوسائل المساعدة :

١. الرجوع إلى مراجع أخرى.
٢. المناقشة.
٣. حل التمارين.
٤. واجبات منزلية.
٥. استخدام وسائل تعليمية حديثة (جهاز العرض, بعض برامج الكمبيوتر الإحصائية).

متطلبات الجدارة :

فهم الوحدات السابقة.

مقاييس النزعة المركزية

Measures of Central Tendency

مقدمة:

لقد تعرفنا في المراحل السابقة على طرق عرض البيانات الإحصائية في صورة جداول وطرق تمثيل هذه البيانات بأشكال هندسية بهدف التعرف على خصائص المجتمعات الإحصائية التي جمعت عنها البيانات، غير أن هاتين الوسيلتين لا تكفيان لتحقيق الأغراض الإحصائية الأساسية في وصف البيانات خاصة بالنسبة للمتغيرات الكمية، إذ يهدف الإحصائي إلى تلخيص البيانات الإحصائية المبوبة في جداول في عدد قليل من المقاييس، يمكن التعرف من خلالها على خصائص المجتمع الإحصائي.

ومن هذا المنطلق تتجلى أهمية هذه المرحلة من مراحل البحث الإحصائي، وذلك من خلال تطبيق الطرق والأساليب الإحصائية المناسبة لتحليل العلاقات الدالية بين المتغيرات موضع الدراسة. وهناك العديد من الطرق والأساليب الإحصائية التي يمكن استخدامها وفقاً لطبيعة البيانات الإحصائية المتاحة والمتغيرات الإحصائية موضع التحليل وكذلك وفقاً للأهداف البحثية المحددة. لذا فإننا سوف نتناول في الأجزاء التالية بعض هذه الطرق والأساليب الإحصائية الأولية، وخاصة تلك التي تناسب طبيعة البحث في مجال البيئة، ومنها مقاييس النزعة المركزية، ومقاييس التشتت.

بالرغم من أن التوزيعات التكرارية تعد تلخيصاً كبيراً للبيانات الإحصائية والتي يمكن منها استنتاج بعض الحقائق عن طبيعة هذه البيانات إلا أنه عادة ما تظهر الحاجة إلى تلخيص هذه البيانات بدرجة أكبر، بحيث يلخص التوزيع التكراري مثلاً أو توزيع البيانات في صورتها الأصلية في شكل رقم واحد يعتبر معبراً عن أحد الخصائص أو الظواهر الهامة التي يمتاز بها توزيع البيانات الإحصائية الخاصة بالمجتمع أو العينة موضع البحث، والذي يمكن استخدامه في إجراء المقارنات واختبار العلاقات مع مجتمعات وعينات أخرى. وتعتبر المتوسطات بأنواعها المختلفة من وسط حسابي، وسيط، منوال، وسط هندسي والوسط التوافقي مثلاً على هذه الأنواع من الأرقام التي توضح بعض الخصائص الهامة لتوزيعات البيانات الإحصائية. وبالرغم من أن المتوسطات المذكورة تقوم بخدمة كثير من التحليلات الإحصائية الأكثر تقدماً مثل الارتباط والانحدار وتحليل السلاسل الزمنية.. الخ، إلا أنه سوف ينظر إليها في هذه الوحدة من ناحية اخراج هذه المقاييس ودراسة أبسط استخداماتها أي باعتبارها مقاييس النزعة المركزية للبيانات الإحصائية.

ويمكن تعريف النزعة المركزية الإحصائية، على أنه ميل هذه البيانات نحو التراكم حول قيمة متوسطة معينة حولها البيانات، وتسمى هذه القيمة التي تتراكم حولها البيانات باسم القيمة المتوسطة. فمن الملاحظ أن عدد القيم (المفردات) التي تأخذ قيمة تختلف عن القيمة المتوسطة يقل عادة كلما بعدنا عن هذه القيمة المتوسطة حتى تصل إلى مرحلة نجد فيها أن عدد المفردات التي تأخذ قيمة تبعد بعداً كبيراً عن القيمة المتوسطة يكاد أن يتلاشى، ويتضح ذلك من دراسة الجداول التكرارية والرسوم البيانية التكرارية التي سبق عرضها، حيث يلاحظ أن أعداداً كبيرة من التكرارات يتراكم حول نقطة متوسطة في المدى الموزع عليه التكرار الكلي، ثم يتناقص عدد التكرار باستمرار كلما بعدنا عن هذه النقطة المتوسطة، حتى يأخذ التكرار أقل قيم له عند طرفي التوزيع (الفئتين الأولى والأخيرة). وبالطبع فإن هذا لا يحدث في جميع الحالات، ولكنه يحدث في أغلب الحالات خاصة التي تتميز بتوزيع طبيعي أو قريب من الطبيعي مثل التوزيعات الملتوية والمفلطحة. ومن ناحية أخرى توجد بعض التوزيعات التي لا تتميز بظاهرة النزعة المركزية، ومن ثم لا يصح في مثل هذه الحالات حساب متوسطات لها، وذلك مثل التوزيعات ذات القمم المتعددة، فمثلاً أخذ عينة من مجتمع يتكون من مجموعتين مختلفتين من الظواهر، فإن حساب متوسط هذه العينة لن يمثل في الواقع أي من المجموعتين الموجودتين في المجتمع الأصلي، إذ أنه من المفترض لا يوجد متوسط واحد يمكن أن يمثل مثل هذه المجتمعات المكونة من مجموعات مختلفة من الظواهر، مثلاً أن مجتمع منطقة الأشجار يشتمل أشكال مختلفة وفقاً لأنواع الأشجار، فليس من الصحيح أن نحسب متوسط مدة طول الشجرة على مستوى مجتمع منطقة الأشجار، فلا معنى لهذا المتوسط، فقد يكون من المناسب حساب متوسط لكل نوع على حده.

١. الشروط التي يجب توافرها في مقاييس النزعة المركزية:

يفضل أن يتوافر في المتوسط الشروط التالية:

- (١) ألا يتأثر بالقيم الشاذة أو المتطرفة.
- (٢) يسهل حسابه ومعالجته جبرياً.
- (٣) أن يأخذ في الاعتبار جميع المفردات التي تتكون منها الظاهرة.
- (٤) يمكن حسابه بسرعة وسهولة.
- (٥) أن يكون له معنى وخواص مميزة.

ويوجد عدة مقاييس للنزعة المركزية أهمها: الوسط الحسابي - الوسيط - المنوال - الوسط الهندسي والوسط التوافقي. وسوف نركز في هذا المقرر على المقاييس الثلاث الأولى.

٢. الوسط (المتوسط) الحسابي (Arithmetic Mean (or Means):

يعتبر المتوسط أو الوسط الحسابي أكثر مقاييس النزعة المركزية انتشاراً وذلك لسهولة تطبيقه وعظم فائدته. ويعرف الوسط الحسابي لقيم متغير ما على أنه القيمة الناتجة عن قسمة مجموع تلك القيم على عددها. ويمكن حسابه من البيانات غير المبوبة والمبوبة وذلك على النحو التالي:

٢.١. الوسط الحسابي من البيانات غير المبوبة:

يُعرف الوسط الحسابي لأي مجموعة من القيم غير المبوبة بأنه عبارة عن حاصل جمع هذه القيم مقسوماً على عددها.

فإن كان لدينا (n) من القيم أو المشاهدات:

$$X_1, X_2, \dots, X_n$$

فإن المتوسط أو الوسط الحسابي ($\bar{X}$) يمكن تقديره من المعادلة التالية:

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

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

من المشاهدات $i = 1, 2, \dots, n$

مثال (١ - ٥): إذا كانت درجات خمسة طلبة في مادة الإحصاء هي: 75, 95, 99, 85, 80 فما هو الوسط الحسابي لدرجات الطلبة؟

الحل

$$\bar{X} = \frac{\sum X}{n} = \frac{75 + 95 + 99 + 85 + 80}{5} = \frac{434}{5} = 86.8 \approx 87$$

مثال (٢ - ٥): افترض أن عدد المخالفات المرورية المسجلة بإدارة المرور خلال أيام أسبوع معين هي على النحو التالي:

اليوم	السبت	الأحد	الاثنين	الثلاثاء	الأربعاء	الخميس	الجمعة
عدد الحوادث	٢٠٠	٣٠٠	١٥٠	١٨٠	٣٥٠	٥٠٠	٤٥٠

أوجد المتوسط الحسابي لعدد المخالفات المرورية.

الحل

$$\bar{X} = \frac{\sum X}{n} = \frac{200 + 300 + 150 + 180 + 350 + 500 + 450}{7} = \frac{2130}{7} = 304.29$$

$$\therefore \bar{X} = 304$$

٢.٢. حساب الوسط الحسابي من بيانات مبوبة:

عندما تكون البيانات مبوبة في جدول توزيع تكراري فإنه يمكن حساب المتوسط الحسابي باتباع

الخطوات التالية:

(١) حساب مراكز الفئات (X).

(٢) إيجاد حاصل ضرب مركز كل فئة في التكرار المناظر لها (FX)

(٣) حساب مجموع حواصل ضرب مركز الفئة في تكرارها لجميع فئات التوزيع التكراري أي

$$\sum FX \quad \text{حساب .}$$

(٤) قسمة حاصل الجمع الناتج من الخطوة السابقة على المجموع الكلي للتكرارات.

وعلى ذلك يمكن حساب الوسط الحسابي من جدول توزيع تكراري وفقاً للمعادلة التالية:

$$\bar{X} = \frac{F_1 X_1 + F_2 X_2 + \dots + F_n X_n}{F_1 + F_2 + \dots + F_n}$$

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

مثال (٣-٥): من مثال سابق (١-١) احسب الوسط الحسابي لعدد الموظفين على مستوى ٥٠ مطعم في بعض مدن السعودية.

الفرقة	التكرار F	مركز الفرقة X	FX
٥ -	١٧	١٢,٥	٢١٢,٥
٢٠ -	١٥	٢٧,٥	٤١٢,٥
٣٥ -	٨	٤٢,٥	٣٤٠
٥٠ -	٣	٥٧,٥	١٧٢,٥
٦٥ -	٤	٧٢,٥	٢٩٠
٨٠ -	٣	٨٧,٥	٢٦٢,٥
المجموع	$\Sigma F=50$		$\Sigma FX=1690$

الحل:

من هذا الجدول يمكن حساب الوسط الحسابي وذلك على النحو التالي:

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

$$= \frac{1690}{50} = 33.8$$

مثال (٤-٥): إذا كان لديك الجدول التكراري الذي يمثل عدد آلاف الدواجن في ٢٤ مزرعة خلال أسبوعين كالتالي:

جدول (١٤) جدول تكراري لعدد الدواجن في ٢٤ مزرعة خلال اسبوعين

الفرقة	التكرار
١٠ -	٣
٢٠ -	١٢
٣٠ -	٤
٤٠ - ٥٠	٥
المجموع	٢٤

احسب الوسط الحسابي لهذا التوزيع ؟

الحل

الفترة	التكرار F	مركز الفترة X	FX
١٠ -	٣	١٥	٤٥
٢٠ -	١٢	٢٥	٣٠٠
٣٠ -	٤	٣٥	١٤٠
٤٠ - ٥٠	٥	٤٥	٢٢٥
المجموع	$\Sigma F = 24$		$\Sigma FX = 710$

يمكن حساب الوسط الحسابي لهذا التوزيع كالآتي:

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

$$\bar{X} = \frac{710}{24} = 29.9$$

وفي حالة ما إذا كان المتغير غير مستمر فإنه يمكن تقريب الوسط الحسابي ليساوي:

$$\bar{X} = 30$$

٣.٢. خصائص الوسط الحسابي:

أ - مجموع الانحرافات عن المتوسط تساوي صفراً لإيضاح هذه الخاصية افترض أن لدينا درجات ثلاث طلاب في مادة الإحصاء هي ٥٠ ، ٦٠ ، ٧٠.

$$\text{المتوسط الحسابي} = \frac{٧٠ + ٦٠ + ٥٠}{٣} = \frac{١٨٠}{٣} = ٦٠$$

وانحرافات القيمة الأصلية عن المتوسط هي:

$$١٠ - = ٦٠ - ٥٠$$

$$٦٠ - ٦٠ = \text{صفر}$$

$$٧٠ - ٦٠ = ١٠$$

$$\text{ومجموع الانحرافات} = (١٠-) + (\text{صفر}) + (١٠) = \text{صفر}$$

ب - إذا أضيف عدد ثابت لكل قيمة من مجموعه من القيم متوسطها X فإن القيمة التي سوف نحصل عليها سوف يكون لها متوسط $X +$ العدد الثابت لإثبات ذلك افترض أن لدينا القيمة التالية:

$$٣ ، ٤ ، ٥ ، ٧ ، ١ \quad \text{المجموع} = ٢٠ \quad \text{المتوسط} = ٤$$

لو تم إضافة عدد ثابت يساوي ٣ لكل القيم السابقة بحيث نحصل على:

$$(٣+٣) ، (٣+٤) ، (٣+٥) ، (٣+٧) ، (٣+١) \text{ أي}$$

$$٦ ، ٧ ، ٨ ، ١٠ ، ٤ \quad \text{فإن المجموع يصبح يساوي } ٣٥ \text{ والمتوسط } ٧.$$

وعليه نجد أن المتوسط الجديد (٧) يساوي المتوسط القديم ٤ + العدد الثابت (٣).

ج - إذا ضربنا كل قيمة من مجموعة من القيم متوسطها (m) بعدد ثابت فإن القيمة التي سوف نحصل عليها سوف يكون لها متوسط يساوي (m) x العدد الثابت لإثبات ذلك افترض أننا ضربنا القيمة السابقة في العدد (٢) فإننا نحصل على:

$$(٢ \times ٣) ، (٢ \times ٤) ، (٢ \times ٥) ، (٢ \times ٧) ، (٢ \times ١) \text{ أو}$$

$$٦ ، ٨ ، ١٠ ، ١٤ ، ٢$$

$$\text{المتوسط الحسابي الجديد} = \frac{٤٠}{٥} = ٨$$

حيث نلاحظ أن المتوسط الحسابي الجديد = المتوسط الحسابي القديم (٤) × العدد الثابت (٢).

٤.٢. عيوب الوسط الحسابي:

نظراً لارتباط قيمة الوسط الحسابي بقيم جميع المفردات فإن هناك بعض العيوب المترتبة على تلك

الخاصية وهي:

أ - تتأثر قيمة الوسط الحسابي بالقيم المتطرفة والشاذة مما يجعله غير ممثل لحقيقة اتجاه توزيع القيم وبذلك يتأثر المتوسط الحسابي بالقيمة القريبة منه تأثيراً قليلاً ويتأثر بالقيمة البعيدة عند تأثيراً كبيراً.

هذه الخاصية تمثل أهم عيوب المتوسط الحسابي حيث أن القيمة المتطرفة عن المتوسط الحسابي تؤثر في المتوسط تأثيراً كبيراً وقد تجعله في بعض الأحيان غير صالح كمقياس للنزعة المركزية حيث يعطى صورة خاطئة في حالة المقارنة بين متوسط حسابي لمجموعة من القيم وآخر لمجموعة أخرى من القيم. ولإيضاح ذلك افترض:

القيم ٢ ، ٣ ، ٥ ، ٦ المجموع = ١٦ المتوسط = ٤

القيم ٢ ، ٣ ، ٥ ، ٦٢ المجموع = ٧٢ المتوسط = ١٨

حيث نجد أن المتوسط الحسابي في الحالة الأولى = ٤ والحالة الثانية = ١٨ ولكن القيمة أكثر تطرفاً في المجموعة الثانية عنه في الأولى. وبذلك يتأثر الوسط الحسابي بالقيم المتطرفة.

ب - البيانات المبوبة ذات الفئات المفتوحة يصعب حساب الوسط الحسابي، بحيث إذا قدر بفرض حد للفئة المفتوحة فإنه يكون ممثل غير دقيق لمجموعة البيانات.

٣. الوسيط The Median:

الوسيط هو القيمة التي تقسم القيم المعطاة إلى قسمين بحيث يسبقها نصف عدد القيم ويتلوها النصف الآخر.

٣.١ إيجاد الوسيط للبيانات غير المبوبة:

إذا كانت البيانات غير مبوبة فإنه لإيجاد الوسيط يتعين علينا الآتي:

١ - ترتيب القيم المعطاة تصاعدياً أو تنازلياً.

٢ - إذا كان عدد القيم المعطاة فردياً فإن ترتيب القيمة الوسيطة يتحدد وفقاً للقاعدة التالية:

$$\text{ترتيب الوسيط} = \frac{n + 1}{2} \quad \text{حيث أن } n \text{ هي عدد القيم.}$$

٣ - إذا كان عدد القيم المعطاة زوجياً فإن الوسيط هو القيمة التي تقع بين القيمتين المركزيتين عندما

$$\text{نرتب القيمة تصاعدياً أو تنازلياً، أي القيمتين اللتين ترتبهما} = \left(\frac{n}{2}\right) \text{ و } \left(\frac{n}{2} + 1\right)$$

مثال (٥-٥): افترض أن لدينا الدرجات التالية لخمس طلبة في مادة الإحصاء وهي ١٨، ١٣، ١٩، ١٢، ٢٠.

الحل

- لإيجاد الوسيط نقوم أولاً بترتيب القيمة تصاعدياً: ١٢، ١٣، ١٨، ١٩، ٢٠

- نظراً لكون عدد القيم هنا ٥ قيم أي أن عدد القيم فردياً فإن ترتيب الوسيط يتحدد

على النحو التالي:

$$\text{ترتيب الوسيط} = \frac{n+1}{2} = \frac{5+1}{2} = 3$$

وهذا يعني أن القيمة الوسيطة هي القيمة التي ترتبها الثالثة من القيم المرتبة تصاعدياً وهي في

مثالنا الدرجة ١٨.

افترض الآن أننا أضفنا درجة طالب آخر ليصبح عدد الطلبة ٦ ويصبح لدينا ٦ قيم مختلفة ١٢،

١٣، ١٨، ١٩، ١٩، ٢٠ حيث نلاحظ أن درجة الطالب السادس هنا هي ١٩ وهي القيمة الجديدة عن

مثالنا لإيجاد الوسيط نوجد ترتيب الوسيط وفقاً للقاعدة $\frac{n}{2} + 1 = \frac{6}{2} + 1 = 3$ ، $\frac{n}{2} = \frac{6}{2} = 3$

أي أن ترتيب القيمة الوسيطة يقع في منتصف المسافة بين القيمة الثالثة والرابعة وعليه فإن قيمة

الوسيط يمكن حسابها من المعادلة التالية:

$$M = \frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2} = \frac{18+19}{2} = 18.5$$

مثال (٦-٥): افترض إن قمنا بإجراء دراسة حول أعمار ١٢ موظفاً في التعليم بالكلية ووجدناها

على النحو التالي:

٣٠، ٢٥، ٣٤، ٤٢، ٣٦، ٢٢، ٣٣، ٥٢، ٤١، ٢٣، ٥٤، ٤٤.

المطلوب: إيجاد الوسيط:

الحل

١ - ترتيب القيم تصاعدياً.

٢٢، ٢٣، ٢٥، ٣٠، ٣٣، ٣٤، ٣٦، ٤١، ٤٢، ٤٤، ٥٢، ٥٤.

٢ - إيجاد ترتيب الوسيط:

$$\frac{n}{2} = \frac{12}{2} = 6 \quad , \quad \frac{n}{2} + 1 = \frac{12}{2} + 1 = 7$$

أي أن قيمة الوسيط هي القيمة التي تقع في المنتصف بين القيمة السادسة والقيمة السابعة.

- القيمة التي ترتيبها السادس هي ٣٤

- القيمة التي ترتيبها السابع هي ٣٦

- وعليه فإن قيمة الوسيط $M = \frac{34 + 36}{2} = 35$

افترض أن الموظف الذي يبلغ عمره ٥٤ غير موجود وبالتالي لم تشمله الدراسة. وفي هذه الحالة نجد أن عدد القيم أصبح فردياً وبالتالي فإن:

$$\text{ترتيب الوسيط: } \frac{n + 1}{2} = \frac{11 + 1}{2} = 6$$

وهذا يعني أن قيمة الوسيط هي القيمة التي ترتيبها السادس

∴ قيمة الوسيط = ٣٤

٢.٣. إيجاد الوسيط للبيانات المبوبة:

يمكن إيجاد الوسيط من البيانات المبوبة في جدول توزيع تكراري بطريقتين هما: طريقة الرسم والطريقة الحسابية من الجدول ولكي نتعرف على طريقة تطبيقهما نعرض المثال التالي:

من مثال (١ - ١) السابق: الجدول التالي يوضح التوزيع التكراري المتجمع الصاعد لعدد الموظفين على مستوى ٥٠ مطعم في أحد مدن السعودية.

الفئة	التكرار	الحدود العليا للفئات	التكرار المتجمع الصاعد
		أقل من ٥	صفر
٥ -	١٧	أقل من ٢٠	١٧
٢٠ -	١٥	أقل من ٣٥	٣٢
٣٥ -	٨	أقل من ٥٠	٤٠
٥٠ -	٣	أقل من ٦٥	٤٣
٦٥ -	٤	أقل من ٨٠	٤٧
٨٠ -	٣	أقل من ٩٥	٥٠
المجموع	٥٠		

واستناداً إلى بيانات هذا الجدول سوف نعرض فيما يلي لإيجاد الوسيط بطرقتي الرسم والحساب:

٢.٣.١. الوسيط بطريقة الحساب:

يمكن إيجاد الوسيط من البيانات المبوبة بهذه الطريقة بتطبيق القانون التالي:

$$M = L_1 + \left[\frac{\sum F / 2 - F_1}{f_1} \right] W$$

حيث:

L_1 = الحد الأدنى للفئة

$\sum F / 2$ = ترتيب الوسيط

F_1 = التكرار المتجمع السابق للفئة الوسيطة

f_1 = تكرار الفئة الوسيطة

W = طول الفئة الوسيطة

ومن الجدول السابق نستطيع الحصول على التقديرات التالية:

$$\frac{\sum F}{2} = \frac{50}{2} = 25 \quad \text{ترتيب الوسيط:}$$

وهذا يعني أن الفئة الوسيطة هي الفئة التي يقابلها التكرار المتجمع الذي يقع فيه الرقم ٢٥ وهي الفئة

(٢٠ - ٣٥).

$L_1 = 20$ الحد الأدنى للفئة الوسيطة (L_1):

$W = 15$ طول الفئة الوسيطة

$F_1 = 17$ التكرار المتجمع السابق للفئة الوسيطة:

$f_1 = 15$ تكرار الفئة الوسيطة:

وبتطبيق القانون السابق نحصل على قيمة الوسيط على النحو التالي:

$$M = 20 + \left[\frac{25 - 17}{15} \right] 15 = 20 + 8 = 28$$

٣.٢.٢.١. إيجاد الوسيط بطريقة الرسم:

لإيجاد الوسيط بطريقة الرسم تتبع الخطوات التالية:

١ - من الجدول التكراري نوجد التوزيع التكراري المتجمع الصاعد.

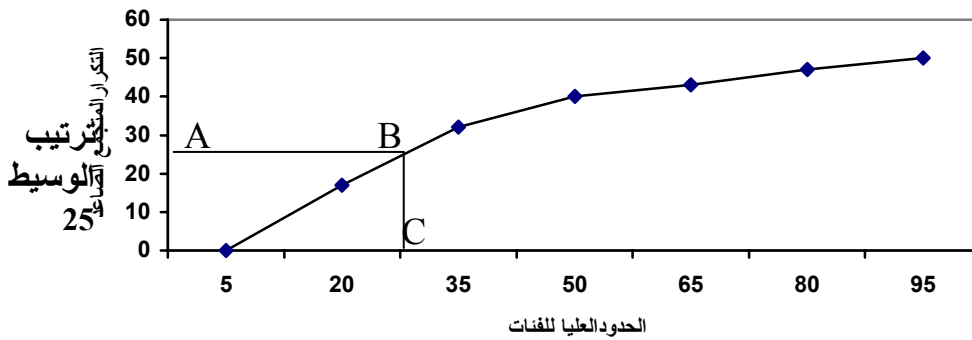
$$٢ - \text{نقوم بتحديد ترتيب الوسيط} = \frac{\text{مجموع التكرارات}}{٢} = \frac{\sum F}{2}$$

٣ - نقوم برسم المنحنى التكراري المتجمع الصاعد.

٤ - من نقطة ترتيب الوسيط نسقط عموداً أفقياً من المحور الرأسي حتى يلاقى المنحنى التكراري المتجمع الصاعد ومن نقطة الالتقاء هذه نسقط عموداً رأسياً حتى يلاقى المحور الأفقي عند قيمة معينة هي الوسيط.

لإيضاح ذلك سوف نعيد رسم المنحنى التكراري المتجمع الصاعد لعدد الموظفين.

شكل رقم (١٣) المنحنى المتجمع الصاعد لعدد الموظفين



$$\text{ترتيب الوسيط:} \quad \frac{\sum F}{2} = \frac{50}{2} = 25$$

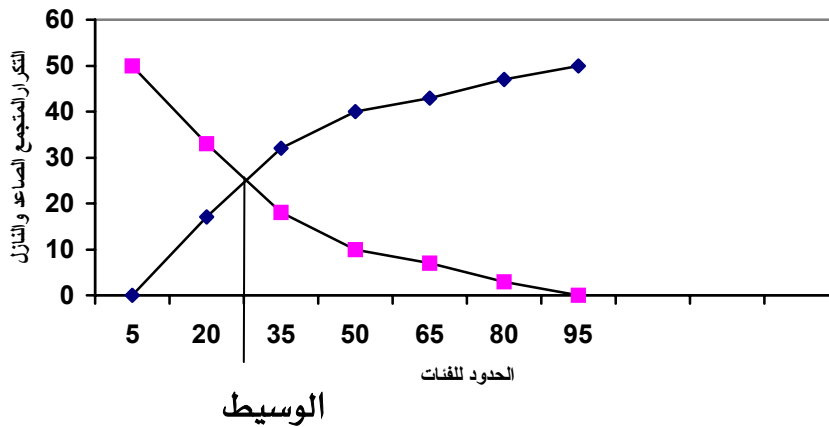
يلاحظ هنا أن ترتيب الوسيط بتحديد بناءً على قسمة مجموع التكرارات على ٢ بغض النظر عن كون التكرارات زوجياً أو فردياً.

من على المحور الرأسي وعند نقطة ترتيب الوسيط (٢٥) نسقط العمود الأفقي (AB) حتى يلاقى المنحنى التكراري المتجمع الصاعد عند النقطة (B). ومن النقطة (B) على المنحنى نسقط عموداً رأسياً BC حتى يلاقى المحور الأفقي عند النقطة (C) وعندها يتحدد الوسيط حيث يساوي تقريباً ٢٨ موظف.

٣.٢.٢.٣. طريقة أخرى من خلال الرسم نوجد الوسيط:

نقوم برسم المنحنى التكراري المتجمع الصاعد والمنحنى المتجمع الهابط (النازل) في رسمة واحدة ونقطة التقاء (تقاطع) المنحنيان نسقط عامود على المحور الأفقي ويكون الوسيط.

مثال: من مثال (١- ١) السابق من خلال جدول التكرار المتجمع الصاعد والنازل نرسم منحنى التكرار المتجمع الصاعد ومنحنى التكرار المتجمع النازل في رسما واحدة كالتالي:



شكل رقم (١٤) المنحنى المتجمع الصاعد والنازل لعدد الموظفين

نسقط عموداً من نقطة التقاء (تقاطع) المنحنيان على المحور الأفقي حتى يلاقى المحور الأفقي وعندها يتحدد الوسيط حيث يساوى تقريباً ٢٨ موظف.

مثال (٧- ٥): من الجدول التكراري التالي الذي يوضح أعمار طلبة السنة الأولى بالكلية أوجد الوسيط باستخدام طريقة الرسم وطريقة الحساب .

جدول (١٥) جدول تكراري لأعمار طلبة السنة الأولى بالكلية

الفئة	التكرار	التكرار المتجمع
-------	---------	-----------------

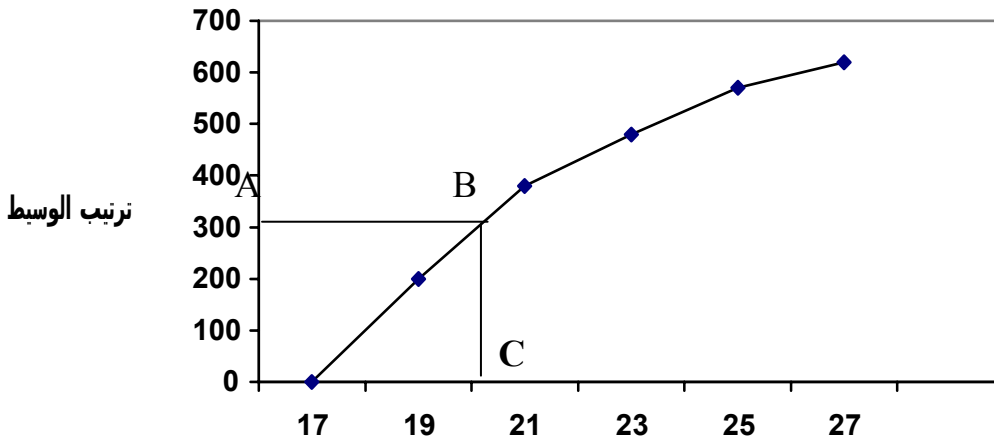
الصاعد		
٢٠٠	٢٠٠	- ١٧
٣٨٠	١٨٠	- ١٩
٤٨٠	١٠٠	- ٢١
٥٧٠	٩٠	- ٢٣
٦٢٠	٥٠	- ٢٥
	٦٢٠	المجموع

الحل

الطريقة الأولى: طريقة الرسم:

شكل رقم (١٥) المنحنى التكراري المتجمع الصاعد لأعمار الطلبة

التكرارات



$$\frac{\sum F}{2} = \frac{620}{2} = 310$$

ترتيب الوسيط:

العمود (A-B) من المحور الرأسي إلى المنحنى.

العمود (B-C) من المنحنى إلى المحور الأفقي

الوسيط هو القيمة التي تقع عند النقطة C على المحور الأفقي ويساوي ٢٠ عاماً تقريباً.

الطريقة الثانية: الوسيط بطريقة الحساب:

$$M = L_1 + \left[\frac{\sum F / 2 - F_1}{f_1} \right] W$$

يطبق القانون التالي:

- الفئة الوسيطة هي الفئة الثانية (١٩ - ٢١) حيث أن ترتيب الوسيط ٣١٠ يقع في هذه الفئة (أنظر التوزيع التكراري المتجمع الصاعد).

$$L_1 = 19$$

$$\frac{\sum F_1}{2} = \frac{620}{2} = 310$$

$$F_1 = 200$$

$$F_1 = 180$$

$$W = 2$$

$$M = 19 + \left[\frac{310 - 200}{180} \right] 2 = 19 + 1.2 = 20.2$$

٣.٣. خصائص الوسيط:

يمكن إيجاز أهم خصائص الوسيط في النقاط التالية:

- (١) يمتاز الوسيط بأنه مقياس للمركز أو الموقع المنصف لعدد القيم بعد ترتيبها تصاعدياً أو تنازلياً، حيث يكفي لحسابه فقط معرفة عدد القيم التي تزيد عن قيمة الوسيط وعدد القيم التي تقل عن قيمته. وهنا عكس الوسط الحسابي الذي يتطلب لحسابه تحديد قيمة كل مفردة من مفردات العينة.
- (٢) يمكن حساب قيمة الوسيط من الجداول التكرارية المفتوحة من طرفها الأعلى أو الأسفل أو المفتوحة من الطرفين.
- (٣) إن قيمة الوسيط تكون أقل تأثراً بالقيم المتطرفة التي تأخذها الظاهرة موضع الدراسة وذلك بخلاف الوسط الحسابي الذي يتأثر بالقيم المتطرفة.

فإن كان لدينا مجموعتان من البيانات الخاصة بظاهرة معينة:

المجموعة (أ): ٥، ٧، ٩، ١١، ١٣.

المجموعة (ب): ٥، ٧، ٩، ١١، ٢٣.

نجد أن:

قيمة الوسط قيمة الوسط الحسابي

المجموعة (أ):	٩	٩
المجموعة (ب):	٩	١١

ومن هذه التقديرات نلاحظ أن قيمة الوسيط لم تتأثر بالقيمة المتطرفة (٢٣)، في حين أن الوسط الحسابي قد تأثر بها بدرجة واضحة.

(٤) مجموع الانحرافات المطلقة للقيم المجموعة عن الوسيط أصغر من مجموع هذه الانحرافات المطلقة عن أي متوسط أو قيمة أخرى.

لذا فإنه يتعين على الباحث خاصة بالنسبة للظواهر البيئية أن يتخير في ضوء هذه الخصائص نوع المتوسط التي يتناسب مع طبيعة البيانات الإحصائية المجموعة عن تلك الظاهرة، فقد يكون استعمال الوسيط أكثر واقعية إذا ما كنا بصدد ظاهرة تحتوي مفردات بياناتها على بعض القيم الشاذة أو المتطرفة كما تبين سلفاً.

٤. المنوال The Mode:

المنوال لقيم معطاة هو القيمة الأكثر شيوعاً أو بعبارة أخرى هو القيمة التي يرتبط بها أعلى تكرار.

ويمكن إيجاد المنوال للقيم غير المبوبة كما يمكن إيجاده للقيم المبوبة.

٤. ١. حساب المنوال من بيانات غير مبوبة:

مثال (٨ - ٥): لإيضاح كيفية إيجاد المنوال للقيم الغير مبوبة افترض أن درجات ٢٠ طالباً في مادة الإحصاء هي:

١٧	١٧	١٦	٥	١٤	٢٣	٢٢	١٣	٢٠	٥
١٠	٩	٦	١٤	٥	٢٢	٥	١٣	١٩	١٨

أوجد المنوال:

الحل

لحساب المنوال تتبع الخطوات التالية:

أولاً: نرتب الدرجات تصاعدياً:

٥، ٥، ٥، ٥، ٦، ٩، ١٠، ١٣، ١٣، ١٤، ١٤، ١٦، ١٧، ١٧، ١٨، ١٩، ٢٠، ٢٢، ٢٢، ٢٣.

ثانياً: إيجاد تكرار كل درجة:

الدرجة	٥	٦	٩	١٠	١٣	١٤	١٦	١٧	١٨	١٩	٢٠	٢٢	٢٣
التكرار	٤	١	١	١	٢	٢	١	٢	١	١	١	٢	١

يلاحظ هنا أن الدرجة ٥ لها أكثر تكرار وهذا يعني أنها أكثر شيوعاً وعليه فإن المنوال لهذه الدرجات هو ٥.

ولكن لاحظ لو أننا أضفنا طالبين للمجموعة بحيث يكون لدينا ٢٢ طالباً وكانت درجات الطالبين الجديدين ١٧، ١٧ في هذه الحالة سوف نجد أن لدينا قيمتين لهما نفس التكرار (تكرار ٤ مرات) وهي الدرجة ٥ والدرجة ١٧ وفي هذه الحالة نجد للدرجات أكثر من منوال وهذا من عيوب المنوال بحيث يمكن أن يكون هنالك أكثر من منوال.

٢.٤. حساب المنوال من بيانات مبوبة:

يمكن إيجاد المنوال للبيانات المبوبة بطريقتين:

■ الطريقة الأولى: المنوال هو مركز الفئة المنوالية:

ولإيجاد المنوال بهذه الطريقة تتبع الخطوات التالية:

- ١ - إيجاد الجدول التكراري للبيانات بحيث يشتمل على الفئة والتكرار المقابل لها.
- ٢ - تحديد الفئة المنوالية وهي الفئة التي لها تكرار أكثر.
- ٣ - إيجاد مركز الفئة المنوالية واعتباره المنوال للبيانات الخام.

مثال: من مثال (١ - ١) سابق: الجدول التكراري الذي يوضح عدد الموظفين لخمسین مطعم أوجد المنوال:

الحل:

الفئة	التكرار
٥ -	١٧
٢٠ -	١٥
٣٥ -	٨
٥٠ -	٣
٦٥ -	٤
٨٠ -	٣

نلاحظ أن الفئة الأولى (٥ - ٢٠) لها تكرار أكثر من غيرها وعليه تكون هي الفئة المنوالية، وبالتالي فإن المنوال هو مركز الفئة المنوالية ويحسب من القانون التالي:

$$M_O = \frac{L_1 + L_2}{2}$$

$$M_O = \frac{5 + 20}{2}$$

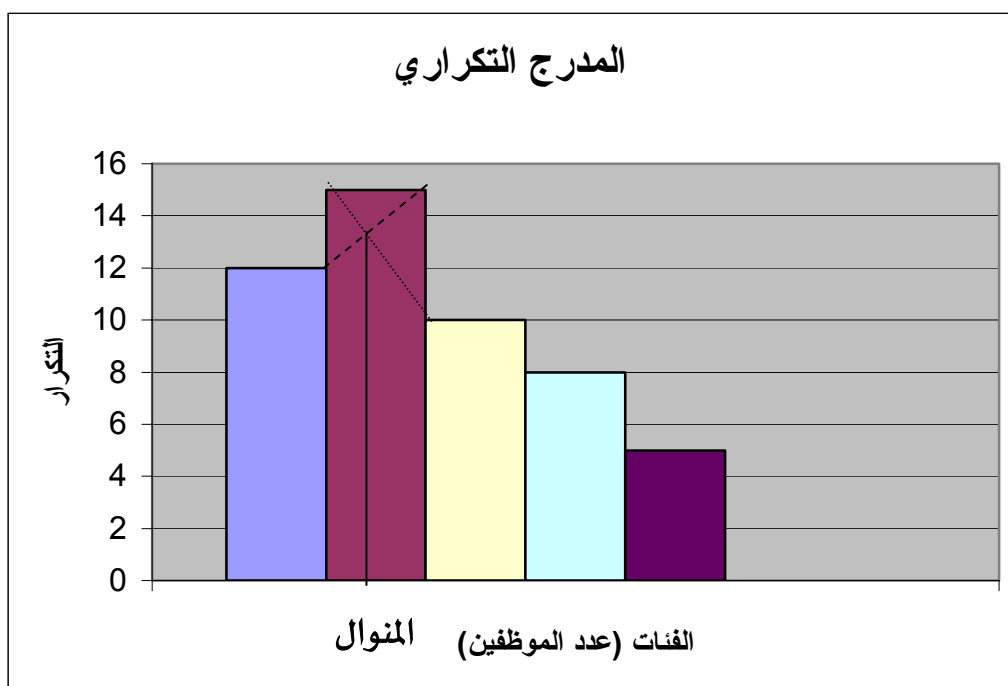
حيث: الحد الأدنى للفئة المنوالية = L_1
الحد الأعلى للفئة المنوالية = L_2

■ الطريقة الثانية: إيجاد المنوال بطريقة الرسم:

لإيجاد المنوال بطريقة الرسم تتبع الخطوات التالية:

- ١ - نرسم المدرج التكراري.
- ٢ - نحدد المستطيل الخاص بالفئة المنوالية.
- ٣ - نحدد المستطيل الخاص بالفئة السابقة لها.
- ٤ - نحدد المستطيل الخاص بالفئة اللاحقة لها.
- ٥ - توصيل الزاوية اليمين العليا (أ) من المستطيل الخاص بالفئة السابقة للفئة المنوالية بالزاوية اليمين العليا (ب) من المستطيل الخاص بالفئة المنوالية.
- ٦ - توصيل الزاوية اليسار العليا من مستطيل الفئة المنوالية (ج) بالزاوية اليسار العليا من مستطيل الفئة اللاحقة لها (د).
- ٧ - من نقطة التقاطع (هـ) نسقط عاموداً رأسياً على المحو الأفقي.
- ٨ - نقطة التقاء هذا العامود مع المحور الأفقي (و) تحدد المنوال وهو في مثالنا السابق حوالي ٢٤.

شكل رقم (١٦) المنوال بطريقة الرسم



تنبيه: هنالك صعوبة في إيجاد المنوال بطريقة الرسم في حالة كون الفئة المنوالية هي الفئة الأولى.

٤.٣. خصائص المنوال:

يمكن إيجاز خصائص المنوال في النقاط التالية:

- (١) يعتبر المنوال مقياس موضعي مثله في ذلك مثل الوسيط، بمعنى أنه لا يتأثر بالقيم الشاذة المتطرفة التي تأخذها الظاهرة، إلا أن تأثير هذه القيم الشاذة (المتطرفة) يكون أقل بالنسبة للمنوال عنه بالنسبة للوسيط.
- (٢) يمكن حسابه من الجداول التكرارية المفتوحة من أحد طرفيها أو من الطرفين.
- (٣) سهولة حسابه، حيث أن تقدير قيمته يحتاج فقط إلى معرفة القيمة التي تكرر عدداً أكبر من تكرارات غيرها من القيم.
- (٤) يستخدم المنوال في دراسات معينة، خاصة تلك الظواهر التي تتركز قياساتها حول قيم معينة ولا يمكن ترتيبها تصاعدياً وتنزلياً مثل فصائل الدم.
- (٥) يفضل كل من المنوال والوسيط عن الوسط الحسابي في الحالات التي يكون فيها التوزيع ملتوياً التواءً شديداً سواءً لليسار أو لليمين.

٥. العلاقة بين الوسط والوسيط والمنوال :

هناك علاقة بين الوسط والوسيط والمنوال في التوزيعات التكرارية ويمكن توضيحها كما يلي:

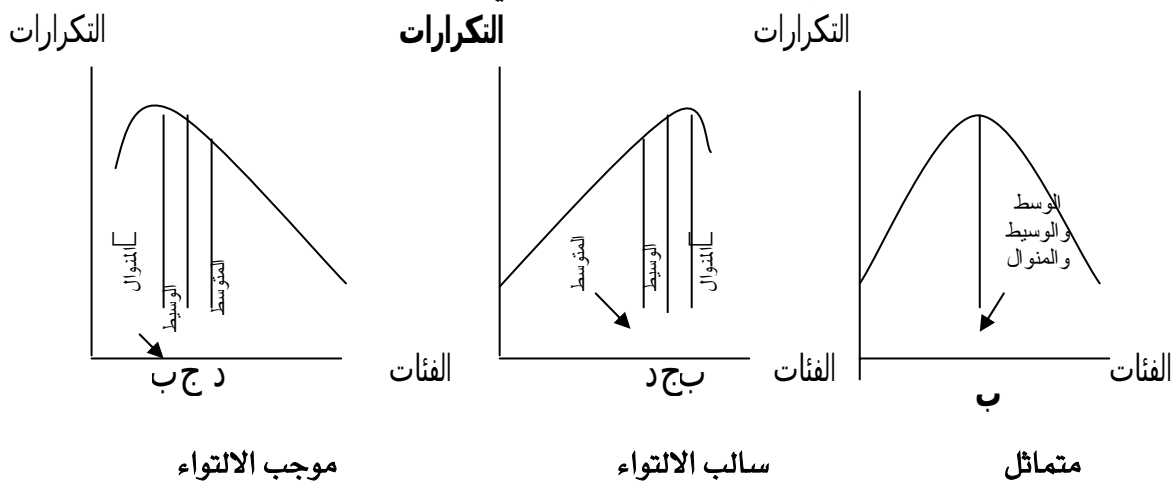
١ - في التوزيع التكراري المتماثل يكون:

الوسط = الوسيط = المنوال.

٢- في التوزيع التكراري غير المتماثل (الملتوي) يبعد الوسيط عن المتوسط بمقدار ثلث المسافة

بين المتوسط والنوال فيكون أقرب إلى المتوسط من النوال فالوسيط يقع دائماً بين

المتوسط والمنوال، كما يتضح في الشكل (١٧) التالي:



شکل رقم (۱۷)

في المنحنى المتماثل نقطة ب تمثل الوسط والوسيط والنوال في المنحنى المتلوي:

- نقطة ب تمثل المنوال
- نقطة ج تمثل الوسيط
- نقطة د تمثل المتوسط

من الشكل يلاحظ أن:

- ١ - المنوال يكون أكبر من كل من الوسيط والمتوسط في حالة التوزيع سالب الالتواء.
- ٢ - المنوال يكون أصغر من كل من الوسيط والمتوسط في حالة التوزيع موجب الالتواء.
- ٣ - المسافة (ج د) دوماً تساوى ١/٣ المسافة (ب د) في التوزيع غير المتماثل.

ويمكن وضع العلاقة بين الوسط والوسيط والمنوال على شكل معادلة رياضية هي:

$$\text{الوسط الحسابي} - \text{المنوال} = 3 (\text{الوسط الحسابي} - \text{الوسيط})$$

وبموجب هذه العلاقة يمكن تقدير أي من المنوال أو المتوسط أو الوسيط إذا ما توفرت المعلومات اللازمة.

مثال (٩ - ٥):

إذا كان متوسط توزيع ما ١,٦٣

ووسيطه ١,٦١

استخرج المنوال.

الحل:

نفرض أن المنوال س

من العلاقة:

$$\text{المنوال} = \text{الوسط} - 3 (\text{الوسط} - \text{الوسيط})$$

$$س = ١,٦٣ - 3 (١,٦٣ - ١,٦١)$$

$$س = ١,٦٣ - 3 (٠,٠٢)$$

$$س = ١,٦٣ - ٠,٠٦$$

$$= ١,٥٧$$

مثال (١٠ - ٥):

إذا كان متوسط توزيع ما ١,٧٩

ومنواله ١,٥٨

استخرج وسيطة

الحل:

نفرض أن الوسيط = س

$$\text{المتوسط} - \text{المنوال} = 3 (\text{المتوسط} - \text{الوسيط})$$

$$١,٧٩ - ١,٥٨ = 3 (س - ١,٧٩)$$

$$٠,٢١ = 3س - ٥,٣٧$$

$$س^3 = ٥,١٦$$

$$س = \frac{٥,١٦}{٣} = ١,٧٢$$

مثال (١١ - ٥):

استخرج المتوسط للجدول التكراري رقم (٥) إذا علمت أن:

منواله ١٦٧,٦٧٧

وسيطه ١٧٥,٥٥٦

الحل:

نفرض أن المتوسط = س

المتوسط - المنوال = ٣ (المتوسط - الوسيط)

$$س - ١٧٦,٦٦٧ = ٣ (س - ١٧٥,٥٥٦)$$

$$س - ١٧٦,٦٦٧ = ٣س - ٥٢٦,٦٦٨$$

$$٣س - ٣٥٠ =$$

$$س = ١٧٥ \text{ المتوسط}$$

تمارين

س١: أذكر أهمية قياس النزعة المركزية للبيانات ؟

س٢: عرف المتوسط الحسابي، الوسيط، المنوال ؟

س٣: أحسب المتوسط الحسابي لعدد المبيعات الغذائية التالية والتي سجلت خلال شهر في ١٥ مدينة.

٥٠	٦٥	٩٧	٩١	٩٢	٩٤	٩٧	٩٨
-	٨٤	٥٥	٤٧	٣٦	٥٢	٤٣	٨٠

ثم تحدث عن فائدة الرقم التي حصلت عليه كمتوسط حسابي.

س٤: بين كيف يتم حساب الوسيط من البيانات التالية:

١٢ ١٣ ١٥ ٧ ٩ ١٠ ١٤

س٥: بين كيف يتم حساب الوسيط من البيانات التالية وما هو الفرق بين كيفية إيجادها في السؤال الرابع

والسؤال الخامس: ١٢ ١٣ ١٥ ٧ ٩ ١٠ ١٤ ١١

س٦: البيانات التالية تمثل خسائر المواد الغذائية الفاسدة السنوية لعشر متاجر في إحدى مدن المملكة

(الأرقام التالية تتمثل بآلاف الريالات).

١٥ ١٧ ١٦ ١٤ ١٩ ١٧ ١١ ١٨ ١٥ ١٧

والمطلوب حساب المنوال وتوضيح مدلوله وأهميته الإحصائية.

س٧: البيانات التالية تمثل النسبة المئوية لتلوث المياه لـ ١٠ مدن في العالم:

٥٠ ٧٠ ٧٠ ٢٢ ٣٠ ٢٢ ٧٠ ٢٢ ٨٠ ٥٠

والمطلوب حساب قيمة المنوال وتوضيح مدلوله الإحصائي والبيئي.

س٨: من الجدول التكراري التالي الذي يوضح أعمار الطلاب في كلية التقنية أوجد المتوسط الحسابي

لهذه الأعمار:

الفئة	التكرار
١٨ -	١٠
٢٠ -	١٥
٢٢ -	١٧
٢٤ -	١٣
٢٦ -	٥
المجموع	٦٠

س٩: من الجدول التكراري الذي يوضح أعمار طلاب كلية التقنية في السؤال الثامن، بين كيف يتم إيجاد الوسيط بطريقة الرسم ثم بطريقة الحساب. ثم تحدث قليلاً عن أهمية إيجاد الوسيط في الدراسة الإحصائية.

س١٠: من الجدول التكراري الذي يوضح أعمار طلاب كلية التقنية أوجد المنوال وذلك:

- ١ - بطريقة الرسم.
- ٢ - باعتباره مركز الفئة المنوالية.
- ٣ - بطريقة الحساب.

س١١: أحسب أفضل مقياس للنزعة المركزية لفصائل الدم لعشرة مصابين بتسمم غذائي:

O , AB , B , O , AB , O , O , A , B , A

مقدمة في الإحصاء

مقاييس التشتت



الوحدة السادسة

اسم الوحدة : مقاييس التشتت.

الجدارة : معرفة مقاييس التشتت وتطبيقاتها.

الأهداف :

- ١ - القدرة في استخراج مقاييس التشتت.
- ٢ - الإلمام بمعرفة الخصائص الهامة لمقاييس التشتت.
- ٣ - دراسة طرق استخراج مقاييس التشتت.

مستوى الأداء المطلوب : ٩٥٪.

الوقت المتوقع للتدريب على الجدارة : ٧ ساعات.

الوسائل المساعدة على تحقيق الجدارة :

١. الرجوع إلى مراجع أخرى.
٢. المناقشة.
٣. حل التمارين.
٤. واجبات منزلية.
٥. استخدام وسائل تعليمية حديثة (جهاز العرض، بعض برامج الكمبيوتر الإحصائية).

متطلبات الجدارة :

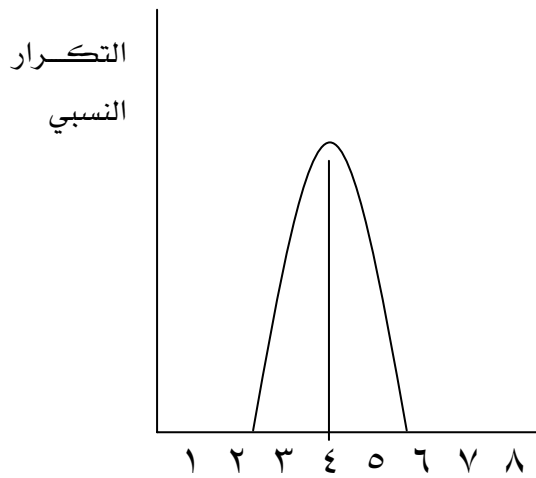
فهم الوحدات السابقة.

مقاييس التشتت

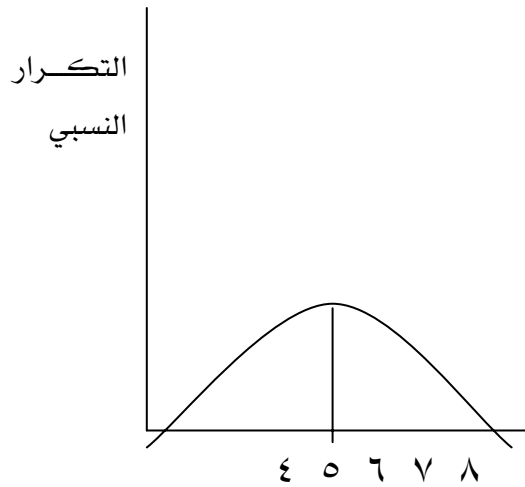
Measures of Dispersion

مقدمة :

لقد تبين من دراسة المتوسطات إمكانية تكوين فكرة مبدئية عن التوزيعات التكرارية بالاعتماد على المتوسطات المقدرة لها. غير أن مقاييس النزعة المركزية لا تعطي وحدها فكرة عامة عن حقيقة المجموعة التي يمثلها التوزيع، حيث لا تبين طبيعتها أو كيفية توزيع مفرداتها. وبعبارة أخرى فإن مقارنة المجموعات على أساس متوسطاتها فقط مقارنة غير كاملة وقد تكن مضللة، إذ كثيراً ما تتساوى متوسطات عدد من المجاميع، بينما تتباين مفرداتها تبايناً كاملاً. ولكي نصف مجموعتين ونقارنهما بدقة، يجب ألا نقتصر على مقارنة متوسطي المجموعتين، بل يجب أن نصف درجة اختلاف مفردات المجموعتين بعضها عن بعض، أو عن متوسطاتها، وبعبارة أخرى لكي نضمن دقة المقارنة يجب أن نستخدم مع مقاييس التوسط مقاييس أخرى تبين مدى تشتت أو تباين مفردات المجموعات عن بعضها أو عن متوسطها ولتوضيح هذه الفكرة دعنا نتأمل التوزيعين الموضحين بالشكل التالي:



شكل (أ)



شكل (ب)

شكل رقم (١٨)

يلاحظ أن كل من التوزيعين يتركز حول القيمة (٤)، إلا أن تشتت قيم كل توزيع حول هذه القيمة المركزية يختلف بالنسبة لكل توزيع. فبينما تنحصر قيم التوزيع الأول بين القيمتين ٣، ٥ فإن التوزيع الثاني تتوزع بين القيمتين صفر، ٨. لذا فإنه من الضروري التعرف على بعض المقاييس الهامة للتشتت. سوف نقتصر في هذا المقرر على دراسة المقاييس الأربع التالية:

- ١ - المدى. 1- The Range
- ٢ - الانحراف المتوسط. 2- Mean Deviation
- ٣ - الانحراف المعياري. 3- Standard Deviation
- ٤ - معامل الاختلاف. 4- Coefficient of Variation
- ٥ - مقاييس التشتت النسبي. 5- Relative Dispersion

١. المدى The Range:

يعرف المدى بأنه الفرق بين أكبر وأصغر قيمة من قيم المجموعة. ويعتبر أسهل مقاييس التشتت تطبيقاً ولكن أقلها دقة حيث تبرز أهميته في أنه يفيد بأن قيم المجموعة تقع بين هاتين القيمتين ويمكن حساب المدى سواء من البيانات غير المبوبة أو المبوبة وذلك على النحو التالي:

١.١ حساب المدى من البيانات غير المبوبة:

ليبان كيفية حساب المدى من البيانات غير المبوبة نفترض أن لدينا إحصائية كالتالي:

مثال (١-٦): عدد المخلفات اليومية في ثلاث مطاعم وذلك على النحو التالي بالجدول التالي:

جدول (١٤) المخالافات اليومية في ثلاث مطاعم

المجموع	السبت	الأحد	الاثنين	الثلاثاء	الأربعاء	الخميس	الجمعة
المطعم الأول	٥	١٠	٢٥	٢٠	٤٠	٥٠	٥٠
المطعم الثاني	٢٠	٣٠	٤٥	٣٠	٣٠	٢٠	٢٥
المطعم الثالث	١	٢	٣	٤	٥	٥	١٨٠

فالمطلوب هو حساب المتوسط الحسابي والمدى على مستوى كل مطعم من المطاعم الثلاثة.

بحساب المتوسط الحسابي لعدد الخالفات في المطاعم الثلاث نجد أن هذه المطاعم الثلاثة متساوية

في المتوسط الحسابي، أي أن:

$$\overline{X}_1 = \overline{X}_2 = \overline{X}_3$$

$$\overline{X}_1 = \frac{\sum X}{n} = \frac{200}{7} = 28.6$$

لاحظ أن هذا المتوسط يعطي دلالة مركزية عن المخالفات اليومية في المطعم الثاني دلالة شبه معقولة للمطعم الأول ولكنه لا يعطي أي دلالة بالنسبة للمطعم الثالث ويمكننا معرفة التشتت لكل مطعم عن طريق حساب المدى وذلك من القانون التالي:

حيث أن: -

- أصغر قيمة تساوي $X_{\min}$ و أكبر قيمة تساوي $X_{\max}$

$$R_1 = 50 - 5 = 45$$

- المدى المطعم الأول

$$R_2 = 45 - 20 = 25$$

- المدى المطعم الثاني

$$R_3 = 180 - 1 = 179$$

- المدى المطعم الثالث

لاحظ الاختلاف الكبير في المدى في الحالات الثلاث على الرغم من تساوي قيمة المتوسط.

٢.١. حساب المدى من البيانات المبوبة:

لا يمكن إيجاد المدى للبيانات المبوبة عن طريق الفرق بين القيمة الكبرى والقيمة الصغرى إلا إذا أعطينا البيانات الخام. إما من خلال الجدول التكراري فإننا لا نعرف قطعاً القيمة الصغرى والقيمة الكبرى من القيم المعطاة وبالتالي فإن إيجاد المدى للبيانات المبوبة يعتمد على افتراض كون الحد الأدنى للفئة الأولى هو القيمة الصغرى وأن الحد الأعلى للفئة الأخيرة هو القيمة الكبرى وبالتالي فإن المدى يصبح:

للبيانات المبوبة = الحد الأعلى للفئة الأخيرة - الحد الأدنى للفئة الأولى

$$R = C_{\max} - C_{\min}$$

حيث: -

- الحد الأعلى للفئة الأخيرة = $C_{\max}$

- الحد الأدنى للفئة الأولى = $C_{\min}$

مثال (٢-٦): أوجد المدى الخاص بالتوزيع التكراري لعدد مبيعات ٤٢ متجر للمواد الغذائية بآلاف الريالات كالتالي:

جدول (١٥) تكراري لمبيعات المواد الغذائية في ٤٢ متجر

الفئات	التكرار
١٠ -	١٢
٢٠ -	١٤
٣٠ -	٩
٤٠ -	٥
٥٠ -	٢
المجموع	٤٢

الحل:

$$C_{\max} = 60$$

الحد الأعلى للفئة الأخيرة

$$C_{\min} = 10$$

الحد الأدنى للفئة الأولى

$$R = 60 - 10 = 50$$

المدى

ومن الملاحظ أن تقدير المدى في هذه الحالة يعطي قيمة أكبر من المدى الحقيقي.

ويمكن إيجاد المدى للبيانات المبوبة عن طريق إيجاد الفرق بين مركز الفئة الأخيرة ومركز الفئة

الأولى وذلك وفقاً للقانون التالي:

المدى = مركز الفئة الأخيرة - مركز الفئة الأولى

$$R = X_{\max} - X_{\min}$$

حيث أن: -

- مركز الفئة الأولى $X_{\min}$ ومركز الفئة الأخيرة $X_{\max}$

مثال (٢-٦): من التوزيع التكراري السابق أوجد المدى باستخدام مركز الفئة ٩

الحل:

يطبق القانون الخاص بمركز الفئة

$$X_{\min} = \frac{10 + 20}{2} = 15 \quad \text{- مركز الفئة الأولى:}$$

$$X_{\max} = \frac{50 + 60}{2} = 55 \quad \text{- مركز الفئة الأخيرة}$$

$$R = 55 - 15 = 40 \quad \text{وعليه فإن قيمة المدى}$$

لاحظ هنا أننا استخدمنا مركز الفئة ولم نستخدم القيمة المتطرفة (الحد الأدنى، الحد الأعلى) وبالتالي فإن التشتت أصبح أقل منه فيما لو استخدمنا الحدود.

هذا وعلى الرغم من بساطة المدى وسهولة حسابه فإنه لا يمثل مقياساً مرضياً للتشتت فقد يتساوى المدى لتوزيعين رغم اختلاف طبيعتهما. كما أن من عيوبه أنه لا يصلح للمقارنة بين تشتت مجموعتين لأنه يعتمد في حسابه على مشاهدتين فقط من مشاهدات العينة، ويعطي نتائج مضللة عند وجود قيم شاذة أو متطرفة، كذلك لا يمكن حسابه في حالة التوزيعات التكرارية المفتوحة مما يستلزم الرجوع للقيم الأصلية.

ورغم ما يؤخذ على المدى من عيوب، فإن هناك بعض الحالات التي يستخدم فيها المدى وخاصة خرائط ضبط المواصفات، مدى التغير في درجات الحرارة اليومية ومدى التغير في الأسعار في خلال فترة زمنية معينة.

٢. الانحراف المتوسط Mean Deviation:

يتميز مقياس الانحراف المتوسط أو متوسط الانحرافات بأنه يأخذ في الاعتبار جميع القيم حيث يعتمد على انحراف كل قيمة عن المتوسط الحسابي للقيم بينما نجد أن المدى يعتمد على القيم المتطرفة ونجد أن الانحراف الربيعي يعتمد على قيم معينة وهو الربيع الأول والربيع الثالث وهذا يعني أن مقياس الانحراف المتوسط أكثر دقة وشمولاً.

١.٢. الانحراف المتوسط للبيانات غير المبوبة:

مثال (٤ - ٦): افترض أن لدينا درجات مادة الإحصاء لخمسة طلبة ونرغب في إيجاد الانحراف المتوسط (أو متوسط الانحرافات) والدرجات هي: ١٠، ١٢، ١٣، ٢٤، ١٩.

لإيجاد الانحراف المتوسط نتبع الخطوات التالية:

$$\bar{X} = \frac{\sum X}{n} = \frac{68}{5} = 13.6$$

أ - نوجد المتوسط الحسابي للدرجات:

ب - نوجد انحراف كل درجة عن المتوسط الحسابي:

$$10 - 13.6 = -3.6$$

$$12 - 13.6 = -1.6$$

$$13 - 13.6 = -0.6$$

$$24 - 13.6 = 10.4$$

$$19 - 13.6 = 5.4$$

$$\sum (X - \bar{X}) = 0$$

ج - نوجد مجموع الانحرافات:

هذه الخطوة للتأكيد ومنها علاقة بحساب الانحراف المعياري والانحراف المتوسط.

يلاحظ هنا أن مجموع الانحرافات يساوي صفراً كما تعرفنا على ذلك عند الحديث عن المتوسط الحسابي ولذلك لإيجاد الانحراف المتوسط سوف نهمل إشارة السالب ونأخذ القيمة المطلقة وعليه سوف يكون مجموع الانحرافات بعد إهمال إشارة السالب يساوي ٢٣,٦ ولإيجاد الانحراف المتوسط نقسم مجموع الانحرافات المطلقة على عدد القيم كما يلي:

$$M.D = \frac{\sum |X - \bar{X}|}{N}$$

تعني الانحرافات المطلقة لإهمال الإشارة

$$M.D = \frac{23.6}{5} = 4.72$$

وبشكل عام إذا كان لدينا عدد (n) مفردة تتكون من $X_1, X_2, X_3, \dots, X_n$ فإنه يمكن إيجاد الانحراف المتوسط كالتالي:

أ - إيجاد المتوسط الحسابي: $\bar{X} = \frac{\sum X}{n}$

ب - إيجاد الانحرافات المطلقة عن المتوسط:

$$|X_1 - \bar{X}|, |X_2 - \bar{X}|, \dots, |X_n - \bar{X}|$$

ج - إيجاد الانحراف المتوسط عن طريق قسمة مجموع الانحرافات على عدد القيم:

$$M.D = \frac{\sum |X - \bar{X}|}{n}$$

٢.٢. حساب الانحراف المتوسط من البيانات المبوبة:

يمكن إيجاد الانحراف المتوسط للبيانات المبوبة بتطبيق القانون التالي:

$$M.D = \frac{\sum F |X - \bar{X}|}{\sum F}$$

X = مركز الفئة

$\bar{X}$ = المتوسط الحسابي

F = التكرار

مثال (٥ - ٦): أوجد الانحراف المتوسط لعدد الموظفين في الخمسين مطعم كما في الجدول (١٦)

جدول (١٦)

الفئة	التكرار F	مركز الفئة X	F x X	$ X - \bar{X} $	F $ X - \bar{X} $
٥ -	١٧	١٢,٥	٢١٢,٥	٢١,٣	٣٦٢,١
٢٠ -	١٥	٢٧,٥	٤١٢,٥	٦,٣	٩٤,٥
٣٥ -	٨	٤٢,٥	٣٤٠	٨,٦	٦٩,٦
٥٠ -	٣	٥٧,٥	١٧٢,٥	٢٤	٧٢
٦٥ -	٤	٧٢,٥	٢٩٠	٣٨,٧	١٥٤,٨
٨٠ -	٣	٨٧,٥	٢٦٢,٥	٥٣,٧	١٦١,١
المجموع	٥٠		١٦٩٠		٩١٤,١

إيجاد المتوسط الحسابي للتوزيع:

$$\bar{X} = \frac{\sum FX}{\sum F} = \frac{1690}{50} = 33.8$$

$$M.D = \frac{\sum F |X - \bar{X}|}{\sum F} = \frac{914.1}{50} = 18.28$$

يمكن استخدام الوسيط بدلاً من المتوسط الحسابي لإيجاد الانحراف المتوسط وبذلك يصبح

قانون إيجاد الانحراف المتوسط :

- في حالة البيانات غير المبوبة:

$$M.D = \frac{\sum |X - M|}{n}$$

- في حالة البيانات المبوبة

$$M.D = \frac{\sum F |X - M|}{\sum F}$$

حيث:

الوسيط $M =$ ، وأيضاً تهمل الإشارة السالبة عن حساب الانحرافات . | .

٣.٢. خصائص الانحراف المتوسط:

يمكن تلخيص خصائص الانحراف المتوسط في النقاط التالية:

- (١) يدخل في حسابه جمع القيم التي تأخذها الظاهرة، كما يراعى حجم العينة عند حسابه.
- (٢) يتأثر بالقيم الشاذة أو المتطرفة التي تأخذها الظاهرة.
- (٣) لا يمكن حسابه من الجداول المفتوحة.
- (٤) لا يمكن تقدير قيمته بيانياً.
- (٥) لا يخضع للعمليات الجبرية، حيث لا يوجد مبرر رياضي معقول لإهمال إشارات الانحرافات المستخدمة في حسابه.

٣. الانحراف المعياري Standard Deviation

يعتبر الانحراف المعياري من أكثر مقاييس التشتت استعمالاً خاصة إذا كان يتبع قياس التشتت عمليات إحصائية أخرى. ويعتبر الإحصائي (Karl Pearson). كارل بيرسون أول من قدم فكرة الانحراف المعياري ويستعمل الانحراف المعياري في كثير من الأحيان ملازماً للتباين Variance وهو مربع الانحراف المعياري أي أن:

$$\text{التباين} = (\text{الانحراف المعياري})^2$$

$$\sqrt{\text{التباين}} = \text{الانحراف المعياري}$$

٣.١. الانحراف المعياري للبيانات غير المبوبة:

إذا كان لدينا مجموعة من القيم $X_1, X_2, \dots, X_n$

فإن الانحراف المعياري لهذه القيم يتم عن طريق:

١ - إيجاد الانحرافات عن المتوسط الحسابي:

$$(X_1 - \bar{X}), (X_2 - \bar{X}), \dots, (X_n - \bar{X})$$

٢ - إيجاد مربع الانحرافات وذلك لتلافي إشارة السالب التي كانت تهمل عند حساب الانحراف المتوسط:

$$(X_1 - \bar{X})^2, (X_2 - \bar{X})^2, \dots, (X_n - \bar{X})^2$$

٣ - إيجاد مجموع مربعات الانحرافات عن المتوسط الحسابي

$$\sum (X - \bar{X})^2 = (X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2$$

٤ - قسمة مجموع مربعات الانحرافات على عدد القيم (N) للحصول على التباين:

ويرمز للتباين Variance بالرمز σ^2 (Sigma Square)، ويقدر وفقاً للمعادلة التالية:

$$\sigma^2 = \frac{\sum (X - \bar{X})^2}{N}$$

وقلنا سابقاً أن الانحراف المعياري هو الجذر التربيعي للتباين أي أن:

$$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{N}}$$

مثال (٦-٦): افترض أن عدد المصابين بالتسمم التي سجلت في أحد أيام الشهر في ١٢ مدينة كانت كالتالي:

$$٤٢, ٣١, ٢٠, ١٧, ٢٥, ١٤, ٣٢, ٩, ١٤, ١٣, ١٥, ١٢$$

والمطلوب إيجاد الانحراف المعياري لعدد المصابين بالتسمم.

الحل:

نتبع الخطوات التالية:

أ - تقدير المتوسط الحسابي:

$$\bar{X} = \frac{\sum X}{n} = \frac{244}{12} = 20.3$$

ب - نوجد انحرافات كل قيمة عن المتوسط:

٤٢	٣١	٢٠	١٧	٢٥	١٤	٣٢	٩	١٤	١٣	١٥	١٢	القيمة
٢١,٧	١٠,٧	-	-	٤,٧	-	١١,٧	-	-	-	-	-	الانحراف
		٠,٣	٣,٣		٦,٣		١١,٣	٦,٣	٧,٣	٥,٣	٨,٣	

ج - نوجد مربع الانحرافات:

٢١,٧	١٠,٧	٠,٣-	٣,٣-	٤,٧	٦,٣-	١١,٧	١١,٣-	٦,٣-	٧,٣-	٥,٣-	٨,٣-	القيمة
٤٧٠,٨٩	١١٤,٤٩	٠,٠٩	١٠,٨٩	٢٢,٠٩	٣٩,٦٩	١٣٦,٨٩	١٢٧,٦٩	٣٩,٦٩	٥٣,٢٩	٢٨,٠٩	٦٨,٨٩	الانحراف

د - قسمة مجموع مربع الانحرافات $(X - \bar{X})^2$ على عدد القيم وبذلك نحصل على التباين:

$$\sigma^2 = \frac{\sum (X - \bar{X})^2}{N} = \frac{1112.68}{12} = 92.7$$

هـ - إيجاد الانحراف المعياري عن طريق إيجاد الجذر التربيعي للتباين:

$$\sigma = \sqrt{\sigma^2} = \sqrt{92.7} = 9.6$$

٢.٣. الانحراف المعياري للبيانات المبوبة:

يمكن توضيح كيفية إيجاد الانحراف المعياري للبيانات المبوبة باستخدام القانون التالي:

$$\sigma = \sqrt{\frac{\sum F (X - \bar{X})^2}{\sum F}}$$

مثال (٧-٦): أوجد الانحراف المعياري لعدد الموظفين في الخمسين مطعم كما في الجدول التالي:

الحل:

الفئة	F	X	FX	$X - \bar{X}$	$(X - \bar{X})^2$	$F(X - \bar{X})^2$
٥ -	١٧	١٢,٥	٢١٢,٥	-٢١,٣	٤٥٣,٦٩	٧٧١٢,٧٣
٢٠ -	١٥	٢٧,٥	٤١٢,٥	-٦,٣	٣٩,٦٩	٥٩٥,٣٥
٣٥ -	٨	٤٢,٥	٣٤٠	٨,٦	٧٣,٩٦	٥٩١,٦٨
٥٠ -	٣	٥٧,٥	١٧٢,٥	٢٣,٧	٥٦١,٦٩	١٦٨٥,٠٧
٦٥ -	٤	٧٢,٥	٢٩٠	٣٨,٧	١٤٩٧,٦٩	٥٩٩٠,٧٦
٨٠ -	٣	٨٧,٥	٢٦٢,٥	٥٣,٧	٢٨٨٣,٦٩	٨٦٥١,٠١
المجموع	٥٠		١٦٩٠			٢٥٢٢٦,٦

نلاحظ أننا عند إيجاد الانحراف المعياري نتبع الخطوات التالية:

(أ) إيجاد المتوسط الحسابي لعدد الأشخاص وذلك بتطبيق القانون:

$$\bar{X} = \frac{\sum FX}{\sum F} = \frac{1690}{50} = 33.8$$

(ب) إيجاد الانحرافات عن المتوسط وذلك بطرح المتوسط الحسابي من مركز كل فئة (X).

(ج) تربيع $(X - \bar{X})$ لكل فئة لنحصل على $(X - \bar{X})^2$ لكل فئة.

(د) ضرب $(X - \bar{X})^2$ لكل فئة في عدد التكرار المقابل لها (F) للحصول على $F(X - \bar{X})^2$

(هـ) إيجاد مجموع مربع الانحرافات بعد ضربه في عدد التكرارات (العمود السابع) لنحصل على

$$\sum F(X - \bar{X})^2$$

(و) إيجاد التباين عن طريق القانون التالي:

$$\sigma^2 = \frac{\sum F(X - \bar{X})^2}{\sum F}$$

$$\sigma^2 = \frac{25226 \cdot 6}{50} = 504 \cdot 532$$

(ر) إيجاد الانحراف المعياري وذلك بأخذ الجذر التربيعي للتباين:

$$\sigma = \sqrt{504 \cdot 532} = 22 \cdot 46$$

مثال (٨ - ٦): من الجدول التكراري التالي الذي يوضح أعمار العمال بأحد المطاعم، أوجد الانحراف المعياري؟

جدول (١٨) اعمار العمال بأحد المطاعم

الفئة	F	X	FX	$X - \bar{X}$	$(X - \bar{X})^2$	$F(X - \bar{X})^2$
٢١ -	٥	٢٢	١١٠	-٥	٢٥	١٢٥
٢٣ -	٧	٢٤	١٦٨	-٣	٩	٦٣
٢٥ -	٨	٢٦	٢٠٨	-١	١	٨
٢٧ -	٧	٢٨	١٩٦	١	١	٧
٢٩ -	٩	٣٠	٢٧٠	٣	٩	٨١
٣١ -	٤	٣٢	١٢٨	٥	٢٥	١٠٠
المجموع	٤٠		١٠٨٠			٣٨٤

الحل:

(أ) نوجد أولاً المتوسط الحسابي (X) للعمر

$$\bar{X} = \frac{\sum FX}{\sum F} = \frac{1080}{40} = 27$$

(ب) نوجد الانحراف عن المتوسط

$$(X - \bar{X})$$

(ج) تربيع الانحرافات لنحصل على

$$(X - \bar{X})^2$$

(د) نوجد حاصل ضرب تكرار كل فئة في مربع الانحراف المقابل لها لنحصل على:

$$F (X - \bar{X})^2$$

(هـ) نوجد مجموع ضرب التكرار في مربع الانحرافات لجميع الفئات لنحصل على

$$\sum F (X - \bar{X})^2 = 384$$

$$\sigma^2 = \frac{\sum F (X - \bar{X})^2}{\sum F} = \frac{384}{40} = 9.6 \quad \text{(و) نوجد التباين:}$$

(ر) إيجاد الانحراف المعياري وذلك بأخذ الجذر التربيعي للتباين:

$$\sigma = \sqrt{\frac{\sum F (X - \bar{X})^2}{\sum F}} = \sqrt{\sigma^2} = \sqrt{9.6} = 3.09$$

٤. الطريقة الثانية لإيجاد الانحراف المعياري:

يمكن إيجاد الانحراف المعياري بطريقة أخرى تعتبر أكثر سهولة من الطريقة السابقة ويمكن توضيح ذلك في الآتي:

٤.١. حساب الانحراف المعياري من البيانات غير المبوبة:

قلنا سابقاً أن الانحراف المعياري للبيانات الغير مبوبة يمكن إيجاده عن طريق القانون التالي:

$$\sigma = \sqrt{\frac{\sum F (X - \bar{X})^2}{N}}$$

وهذا القانون يمكن استبداله بالقانون التالي:

$$\sigma = \sqrt{\frac{\sum X^2}{N} - \left(\frac{\sum X}{N}\right)^2}$$

مثال (٩-٦): افترض أن لدينا درجات عشرة طلاب في مادة الإحصاء على النحو التالي:

١٤ ١٧ ١٩ ٢٠ ١٢ ١٨ ١٥ ١٣ ١٢ ١٠

أوجد الانحراف المعياري بالطريقة الثانية

نوجد (X^2) على النحو التالي:

$$(10)^2 = 100, (12)^2 = 144, (13)^2 = 169, (15)^2 = 225, (18)^2 = 324$$

$$(12)^2 = 144, (20)^2 = 400, (19)^2 = 361, (17)^2 = 289, (14)^2 = 196$$

$$\sum X^2 = 2352$$

- إيجاد مجموع مربعات القيم

$$\frac{\sum X^2}{n} = \frac{2352}{10} = 235.2$$

ومنه نحصل على

$$\sum X = 150$$

- إيجاد المجموع الجبري للقيم

ومنه نحصل على

$$\left\{\frac{\sum X}{N}\right\}^2 = \left\{\frac{150}{10}\right\}^2 = (15)^2 = 225$$

وبتطبيق القانون السابق نحصل على الانحراف المعياري:

$$\sigma = \sqrt{\frac{\sum X^2}{N} - \left(\frac{\sum X}{N}\right)^2} = \sqrt{235.2 - 225} = \sqrt{10.2} = 3.19$$

٢.٤. البيانات المبوبة باستخدام الطريقة الثانية:

هنا يطبق القانون التالي:

$$\sigma = \sqrt{\frac{\sum F (X - \bar{X})^2}{\sum F}}$$

كما يمكن تطبيق الصورة التالية من القانون:

$$\sigma = \sqrt{\frac{\sum FX^2}{\sum F} - \left\{ \frac{\sum FX}{\sum F} \right\}^2}$$

مثال (١٠ - ٦): ولإيضاح كيفية تطبيق القانون السابق نعرض الجدول التكراري التالي:

FX	FX ²	X ²	x	F	الفئة
٢١٢,٥	٢٦٥٦,٢٥	١٥٦,٢٥	١٢,٥	١٧	٥ -
٤١٢,٥	١١٣٤٣,٧٥	٧٥٦,٢٥	٢٧,٥	١٥	٢٠ -
٣٤٠	١٤٤٥٠	١٨٠٦,٢٥	٤٢,٥	٨	٣٥ -
١٧٢,٥	٩٩١٨,٧٥	٣٣٠٦,٢٥	٥٧,٥	٣	٥٠ -
٢٩٠	٢١٠٢٥	٥٢٥٦,٢٥	٧٢,٥	٤	٦٥ -
٢٦٢,٥	٢٢٩٦٨,٧٥	٧٦٥٦,٢٥	٨٧,٥	٣	٨٠ -
١٦٩٠	٨٢٣٦٢,٥			٥٠	المجموع

وبتطبيق القانون السابق نحصل على:

$$\sigma = \sqrt{\frac{\sum FX^2}{\sum F} - \left\{ \frac{\sum FX}{\sum F} \right\}^2} = \sqrt{1647.25 - 1142.44} = \sqrt{504.81} = 22.46$$

مثال (١١ - ٦): من التوزيع التكراري لأعمار العمال ، أوجد الانحراف المعياري باستخدام الطريقة الثانية:

الفئة	F	x	X ²	FX	FX ²
٢١ -	٥	٢٢	٤٨٤	١١٠	٢٤٢٠
٢٣ -	٧	٢٤	٥٧٦	١٦٨	٤٠٣٢
٢٥ -	٨	٢٦	٦٧٦	٢٠٨	٥٤٠٨
٢٧ -	٧	٢٨	٧٨٤	١٩٦	٥٤٨٨
٢٩ -	٩	٣٠	٩٠٠	٢٧٠	٨١٠٠
٣١ -	٤	٣٢	١٠٢٤	١٢٨	٤٠٩٦
المجموع	٤٠		٤٤٤٤	١٠٨٠	٢٩٥٤٤

$$\sigma = \sqrt{\frac{\sum FX^2}{\sum F} - \left\{ \frac{\sum FX}{\sum F} \right\}^2}$$

$$\left\{ \frac{\sum FX}{\sum F} \right\}^2 = \left\{ \frac{1080}{40} \right\}^2 = (27)^2 = 729$$

وبتطبيق القانون نحصل على الانحراف المعياري الآتي:

$$\sigma = \sqrt{738.6 - 729} = \sqrt{9.6} = 3.09$$

ويلاحظ عدم اختلاف النتيجة باختلاف الطريقة إيجاد الانحراف المعياري ففي كلا الحالتين حصلنا على انحراف معياري يساوي (3.09).

٥. خصائص الانحراف المعياري:

يتصف الانحراف المعياري بالخصائص التالية:

(١) تدخل في حسابه جميع القيم التي تأخذها الظاهرة.

- (٢) تتأثر قيمته بالقيم المتطرفة أو الشاذة التي تأخذها الظاهرة.
- (٣) لا يمكن حسابه من الجداول التكرارية المفتوحة (الذي يحوي فئات مفتوحة).
- (٤) لا يمكن حسابه بيانياً.
- (٥) لا تتأثر قيمة الانحراف المعياري المحسوبة من بيانات ظاهرة معينة إذا تم جمع طرق مقدار ثابت من قيمة كل مفردة من مفردات الظاهرة.
- (٦) إذا ضربت قيمة كل مفردة من مفردات الظاهرة في مقدار ثابت، فإن قيمة الانحراف المعياري للقيم الجديدة للظاهرة تساوي قيمة الانحراف المعياري المحسوب من البيانات الأصلية مضروباً في القيمة المطلقة للمقدار الثابت الذي ضرب في قيمة كل مفردة من مفردات الظاهرة.
- (٧) يخضع الانحراف المعياري للعمليات الجبرية. لذا فإنه يفيد من تحليل التباين والارتباط والانحدار.
- (٨) تقل أخطاء المعاينة له عن مقاييس التشتت الأخرى.

٦. مقاييس التشتت النسبي Relative Dispersion :

تستخدم مقاييس التشتت المطلق التي سبق ذكرها (الانحراف الربيعي، الانحراف المتوسط، الانحراف المعياري) لقياس تشتت قيمة متغير حول قيمتها المتوسطة فكلما كان هذا التشتت صغيراً كلما أمكن الاعتماد على نتيجة مقاييس النزعة المركزية. كما أنه باستخدام مقاييس التشتت المطلق يمكن مقارنة أكثر من مجموعة من القيم مع بعضها البعض وهذا ممكن فقط في حالة كون وحدة القياس للمتغيرين المراد المقارنة بينهما واحدة أما في حالة اختلاف وحدة القياس أو اختلاف المتوسطات، فإن مقاييس التشتت المطلق تصبح غير مجدية. للتغلب على هذه الصعوبة قدم الإحصائيون مقياساً نسبياً لا يتأثر بتغير وحدة المقياس ويلغي تأثير المتوسط وهو ما يعرف بمعامل الاختلاف Coefficient of variation وهذا المقياس يعتمد على نسبة مقياس التشتت المطلق إلى المتوسط وعرضه كنسبة مئوية أي أن:

$$C.V. = \frac{\sigma}{X} \times 100$$

حيث أن :

$$= \sigma \quad \text{الانحراف المعياري.}$$

$$= \overline{X} \quad \text{المتوسط الحسابي.}$$

ويمكن إيجاده عن طريق استخدام الربيع الأول والربيع الثالث كما في هذه المعادلة التالية :

$$C.V. = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$$

مثال (١٢ - ٦): من التوزيع التكراري السابق لأعمار الأفراد أوجد معامل الاختلاف.

الحل:

$$\sigma = 3.09 \quad \text{مما سبق أوجدنا الانحراف المعياري}$$

$$\overline{X} = \frac{\sum FX}{\sum F} = 27 \quad \text{المتوسط الحسابي:}$$

$$C.V. = \frac{\sigma}{\overline{X}} \cdot 100 = \frac{3.09}{27} \cdot 100 = 11.44 \% \quad \text{معامل الاختلاف:}$$

ومما هو جدير بالذكر أنه يمكن مقارنة تشتت توزيعين تكراريين في حالة تقارب متوسط كل من التوزيعين وعند استعمال نفس وحدات القياس في كل منهما. ولكن عند مقارنة التشتت لتوزيعين تكراريين لظاهرتين مختلفتين أو أكثر في نفس الفترة أو لظاهرة واحدة على فترتين من الزمن أو أكثر، نلاحظ الآتي:

أ - اختلاف وحدات القياس (الطول بالسنتيمتر والوزن بالكيلو جرام مثلاً).

ب - اختلاف المتوسطات الحسابية.

ففي هذه الحالة لا يصلح الانحراف المعياري لمقارنة التوزيعات المختلفة، ولهذا نلجأ إلى استعمال مقياس التشتت النسبي أو معامل الاختلاف.

ويعتبر معامل الاختلاف من مقاييس التشتت الهامة التي يمكن تطبيقها في مجال البيئة، حيث تتطلب دراسة بعض الظواهر المقارنة بين توزيعات مختلفة في وحدات القياس أو المتوسطات، كما قد تفيد في البحوث البيئية، فقد يكون هناك معامل اختلاف تقريبي متوقع لظاهرة ما، يتم مقارنته بمعامل الاختلاف المقدر من الاختبار المعلمي للظاهرة، فإذا زاد المقدار كثيراً عن نظيره المتوقع دل ذلك على عدم دقة الاختبار أو حدوث خطأ في حسابات الناتج.

تمارين

س١: أذكر أهمية دراسة مقاييس التشتت.

س٢: عرف المدى ثم أذكر كيف يتم إيجاد المدى للبيانات التالية:

٥ ٢ ٣٣ ٢٧ ٤ ٥٠ ٦٦ ٧٧

س٣: قارن بين المدى للمجموعات التالية ثم أذكر أي مجموعة يمكننا الاعتماد على متوسطها الحسابي كقيمة نموذجية لبقية المجموعة.

- المجموعة الأولى ٣٢، ٤٢، ٥٧، ٦٣، ٤٤، ٦٧، ٣١، ٣٠، ٨٥.
- المجموعة الثانية: ١٧، ٨٢، ٤٧، ٥٣، ٣٢، ٢٧، ٤٢، ٥٠، ٥١.
- المجموعة الثالثة: ١٠٠، ٩٩، ٩٨، ٩٧، ٩٥، ٩٠، ٩٤، ٩٣، ١٠١.

س٤: من الجدول التكراري التالي بين كيف يتم إيجاد المدى:

الفئات	التكرار
١٢ -	٥
١٤ -	١٢
١٦ -	١٧
١٨ -	٩
٢٠ -	٨

س٥: يعتبر الانحراف الربيعي أكثر دقة لقياس التشتت من المدى صح أو خطأ ثم لماذا ؟.

س٦: بين كيف يتم إيجاد الانحراف الربيعي للبيانات التالية:

١، ٤، ١٣، ١٢، ٢١، ٧، ١٧، ١٩، ١٥، ١١، ٢٠، ٢٣، ٢٥، ٨، ٩.

س٧: أوجد الانحراف الربيعي للمجموعات الثلاث في السؤال الثالث ثم بين أيهما أكثر تشتتاً؟ قارن النتيجة هنا بالنتيجة التي حصلت عليها للسؤال الثالث.

س٨: بين كيف يتم إيجاد الانحراف الربيعي من الجدول التكراري الذي يوضح عدد العمال في خمسين مصنع:

الفئة	التكرار
١٢٠ -	١٠
١٣٠ -	١٢
١٤٠ -	٩
١٥٠ -	٨
١٦٠ -	٩
١٧٠ -	٢
المجموع	٥٠

س٩: بماذا يتميز مقياس الانحراف المتوسط عن مقياس الانحراف الربيعي عند قياس التشتت للبيانات.

س١٠: بين كيف يتم إيجاد الانحراف المتوسط للمجموعات الثلاث المذكورة في السؤال الثالث؟ قارن إجابتك بتلك التي تحصلت عليها في السؤال الثالث والسؤال السابع.

س١١: من الجدول التكراري الذي يوضح عدد العمال في خمسين مصنع المذكورة في السؤال الثامن؟ بين كيف يتم إيجاد الانحراف المتوسط؟

س١٢: أوجد التباين والانحراف المعياري للمجموعات الثلاثة المذكورة في السؤال الثالث.

س١٣: من الجدول التكراري الموضح في السؤال الثامن أوجد التباين والانحراف المعياري وضع كيف يمكن إيجاد الانحراف المعياري بأكثر من طريقة.

س١٤: أذكر أهمية مقاييس التشتت النسبي؟

س١٥: من الجدول التكراري الموضح في السؤال الثامن بين كيف يتم إيجاد معامل الاختلاف.

المراجع

مراجع باللغة العربية:

- (١) أحمد صفى الدين عوض (دكتور)، الإحصاء العام، دار العلوم، المملكة العربية السعودية، ١٤٠٤هـ.
- (٢) أحمد محمد توفيق الفيل (دكتور)، محاضرات في أساسيات علم الإحصاء، المملكة العربية السعودية، ١٤٠٠هـ.
- (٣) جلال الصياد (دكتور)، ومحمد حبيب (د)، مقدمة في الطرق الإحصائية، دار عكاظ للطباعة والنشر، الطبعة الثالثة ١٤١٥هـ، جدة، جامعة الملك عبد العزيز.
- (٤) جلال الصياد (دكتور)، وعادل سمرة (أ)، مبادئ الإحصاء، تهامة - جدة - المملكة العربية السعودية ١٤٠٤هـ.
- (٥) عبد الحميد محمد نجم (دكتور)، محمد عبد الهادي المحميد (دكتور)، الإحصاء الوصفي والتحليلي في استخدام البرامج الجاهزة، جامعة الكويت، الطبعة الثانية، ١٩٩١م.
- (٦) عبد الرحمن أبوعمه (دكتور)، د. أنور عبد الله، د. محمود هندي، الإحصاء التطبيقي، مطابع جامعة الملك سعود ١٤١٥هـ.
- (٧) عبد المنعم محمد الأعسر (دكتور)، الإحصاء لعلوم الحياة والبيئة، الدار السعودية للنشر والتوزيع الطبعة الأولى ١٤٢٣هـ.
- (٨) فارس عياد شاكر (دكتور)، محاضرات في الإحصاء الاقتصادي والاجتماعي، قسم الاقتصاد الزراعي، كلية الزراعة، جامعة القاهرة، ١٩٩٤م.
- (٩) محمد فتحي محمد (دكتور)، المفاهيم الأساسية للإحصاء، مكتبة عين شمس، القاهرة، ١٩٩٠/١٩٩١م.
- (١٠) مصطفى زايد (دكتور)، الإحصاء ووصف البيانات، دار العلوم، المملكة العربية السعودية، ١٤٠٤هـ.
- (١١) نادية محمد عبد السلام (دكتور)، الإحصاء الوصفي في العلوم النفسية والتربوية، مكتبة الأنجلو المصرية، القاهرة، ١٩٨٧م.

مراجع باللغة الإنجليزية:

- (1) Dominick Salvatore, Statistical & Econometrics, Schaum's Outline series in Economic, Mc Graw-Hill Book Company, 1982.
- (2) Iver E. Bradley, John B. South. Introductory Statistics For Business and Economics, The Dryden Press Hinsdale, Illinois, 1981.
- (3) Joy Devore and Roxy peck. Statistics ; The Exploration and Analysis of Data, West Publishing Company, L.A., 1986.
- (4) Robert B. Miller and Dean W. Wichern Intermediate Business Statistics, Hoit, Rinchart and Winston, U.S.A. 1977.

الصفحة	العنوان
	الوحدة الأولى
١	ماهية علم الإحصاء
	الوحدة الثانية
٨	جمع البيانات الإحصائية
	الوحدة الثالثة
١٤	تبويب البيانات الإحصاء
	الوحدة الرابعة
٢٥	التمثيل البياني للبيانات
	الوحدة الخامسة
٤٤	مقاييس النزعة المركزية
	الوحدة السادسة
٧٠	مقاييس التشتت
٩٢	المراجع