جامعة حماة

كلية التربية

السنة الأولى (معلم صف)

مقرر الإحصاء في التربية

د. أسماء الحسن

العام الدراسي: 2021/2020

المحاضرة الأولى: مقدمة في الإحصاء

أولاً–معنى الإحصاء وتعريفه:

تعد كلمة الإحصاء من الكلمات الشائعة الاستخدام، سواء في الحياة اليومية، أم في ميادين العلوم المختلفة، ولكن الحديث عن الإحصاء في التربية يختلف عنه في الاقتصاد عنه في البيولوجية أو في الرياضيات بالرغم من وجود قاسم مشترك بينها. وقد اقتصر استخدام الإحصاء في البداية على الأعمال الخاصة بشؤون الدولة، كما يدل على ذلك الأصل اللغوي في اسم هذا العلم، وهو بالإنجليزية "Statistics" وهو مشتق من كلمة "State" أي الدولة ومعناه "مجموعة الحقائق الخاصة بشؤون الدولة".

وإن الإحصاء كعلم فهو **أحد فروع العلوم الرياضية**، وذلك أن أسسه وأصول قوانينه هي أسس وقوانين رياضية صرفة، إلا أنه يتميز بكونه **علماً تطبيقياً**، يتناول الناحية الكمية للظواهر في مختلف العلوم، لذلك فلا عجب أن يعده العلماء في العلوم المختلفة أداة للبحث العلمي وأسلوباً للدراسة الكمية العلمية، وطالما أن هناك حاجة للدراسات الميدانية فإن هناك حاجة للإحصاء، فلدراسة ظاهرة معينة أو واقع معين لا بد من الحصول على قراءات واحتمالات وتوقعات لا يقدمها إلا علم الإحصاء الذي وضع عملياً في خدمة العلوم الأخرى.

ومنه يعرف الإحصاء بأنه "**مجموعة من الطرائق العلمية القياسية التي يمكن توظيفها لجمع المعطيات (البيانات والمعلومات) عن الظواهر، وتبويبها وتلخيصها وتقييمها، والخروج منها باستنتاجات حول مجموع وحدات المجتمع من خلال اعتماد جزء صغير من هذا المجتمع**".

ثانياً–استخدام علم الإحصاء في العلوم الإنسانية المختلفة:

يستخدم الإحصاء في شتى ميادين الحياة في الاقتصاد وإدارة الأعمال والمحاسبة ومجالات التأمين وأسواق المال وفي الدراسات السكانية والطب والزراعة والصيدلة والوراثة وفي العلوم الإنسانية المختلفة، وقد أسهم استخدام علم الإحصاء في هذه الميادين في تطورها تطوراً كبيراً. ونظراً لاستخداماته الكبيرة والهامة في مجال العلوم الإنسانية فقد أصبح فرعاً من فروعها فيسمى في علم الاجتماع بالإحصاء الاجتماعي، وفي الجغرافيا بالتحليل الكمي وفي التربية وعلم النفس بالإحصاء التطبيقي، وانطلاقاً من أهمية الأساليب الإحصائية في تحليل ودراسة ظواهر هذه العلوم التي تعتمد على الفروض والاختبارات الإحصائية فقد أصبح تطور العلوم الإنسانية مرتبطاً بتطبيق الأساليب والأدوات الإحصائية، وقد أصبح الإحصاء ركيزة هامة من ركائز البحث العلمي.

فالباحث التربوي إذا أراد دراسة ظاهرة لدى مجتمع ما فإنه يعتمد على مجموعة صغيرة تمثل عينة من المجتمع ليدرسها ويعمم نتائجها، وكل ذلك لا يتم إلا باستخدام مفاهيم الإحصاء من جداول لتنظيم القيم ورسوم بيانية والاعتماد على مقاييس النزعة المركزية للمقارنة بين الأفراد والفئات، وكذلك استخدام مقاييس التشتت، أو إيجاد علاقة بين متغيرين من خلال حساب معامل الارتباط بينهما أو إجراء تحليل الانحدار وكذلك حساب الفروق بين القيم اعتماداً على متغير مستقل من خلال استخدام ت ستيودنت أو تحليل التباين فهي مقاييس إحصائية تسهل مهمة الباحث في العلوم الإنسانية وتزيل العقبات التي تعترضه أثناء قيامه بالبحث وتمكنه من إيجاد الحلول العلمية المنطقية السليمة.

وبهذا يسهم علم الإحصاء في مساعدتنا على فهم أعمق للسلوك الإنساني، واستنباط البراهين التجريبية لقبول الفرضيات أو رفضها، كما أنه يسهم في عملية التخطيط واتخاذ القرارات، ومن الأهمية أن تكون التحليلات الإحصائية ذات نتائج صادقة إذ يتوقف اتخاذ القرار السليم على سلامة النتائج الإحصائية التي جرى التوصل إليها وتزداد أهمية اختيار الإحصائيات المناسبة كلما زادت أهمية القرارات المترتبة على النتائج، إذ أن كثيراً من الدراسات والبحوث ليست ذات قيمة نظرية أو تطبيقية نتيجة لضعف في خطوات البحث العلمي، وترجع معظم مصادر الضعف إلى نوعية التحليلات الإحصائية المستخدمة.

ثالثاً-أقسام الإحصاء:

ينقسم علم الإحصاء إلى قسمين أساسيين هما:

1- الإحصاء الوصفي (Descriptive Statistics):

هو ذلك الإحصاء الذي يصف الواقع كما هو، ويهتم بعرض حقيقة الظاهرة المدروسة ووصفها بشكل ملخص فهو يهتم بجمع البيانات عن الظاهرة وتنظيمها وعرضها عن طريق الجداول أو الرسوم البيانية وغيرها، بهدف التعرف إلى خصائصها وصفاتها، بغية الوصول لنتائج دون محاولة تعميمها من أفراد العينة التي درست خصائصها على عينة أخرى أو على المجتمع.

2- الإحصاء الاستدلالي (Inferential statistics):

وهو ذلك الإحصاء الذي يهدف للتوصل إلى صورة للمجتمع الإحصائي من خلال دراسة جزء من المجتمع والمسمى (العينة)، ويهتم بتحليل البيانات المتوفرة حول عينة مسحوبة من مجتمع إحصائي محدد بغرض التنبؤ والاستقراء، فالاستدلال هو التوصل من الخاص إلى العام، أي من الجزء (العينة) إلى العام (المجتمع

الإحصائي)، والإحصاء الاستدلالي هو "مجموعة الطرائق المستخدمة في التوصل لنتائج المجتمع بالاعتماد على بيانات العينة المسحوبة من ذلك المجتمع".

ولا بد من الإشارة إلى أنه من الممكن إجراء **دراسة إحصائية وصفية** على عينة مسحوبة من مجتمع إحصائي تماماً كما نجريها على ذلك المجتمع، ولكن عندما نقوم بتعميم نتائج العينة على المجتمع الإحصائي تصبح دراستنا **إحصائية استدلالية.**

والإحصاء الاستدلالي أحدث من الإحصاء الوصفي فقد بدأت أهم التطورات في علم الإحصاء الاستدلالي تظهر مع أبحاث كارل بيرسون (Karl Person, 1857–1936) ورونالد فيشر (Ronald Fisher, 1890–1962) وقد نشرا نتائج دراساتهما مع بداية القرن العشرين، ومع بداية أعمالهما بدأ علم الإحصاء الاستدلالي يتطور بسرعة وأصبح يطبق على نطاق واسع في مجالات مختلفة.

رابعاً–مراحل البحث الإحصائي:

إذا أردنا القيام ببحث ما مستخدمين الأساليب الإحصائية، فإننا سنقوم القيام بالخطوات الآتية:

1–تحديد المشكلة المدروسة.

2–جمع البيانات الإحصائية بإحدى الطرائق المناسبة.

3–تصنيف البيانات وتبويبها.

4–عرض البيانات بواحدة أو أكثر من طرائق العرض المناسبة.

5–دراسة هذه البيانات وتحليلها باستخدام الأساليب الإحصائية كمقاييس النزعة المركزية ومقاييس التشتت والارتباط والانحدار والفروق...إلخ.

6–استخلاص النتائج.

إن كل مرحلة مهمة ويجب أن تتم دون أخطاء، وهذه المراحل مرتبطة ببعضها بعضاً، إذ أن كل مرحلة تبدأ من نهاية المرحلة السابقة لها وتعتمد عليها وعلى صحة نتائجها، لذلك فأي خطأ في أي مرحلة سيؤثر على المراحل التي تليها.

خامساً–المتغيرات الإحصائية (Variables):

تسمى المعلومات الإحصائية المتعلقة بالظواهر الاجتماعية والاقتصادية والثقافية والصحية والتربوية والمالية والسكانية **بالبيانات**، وتختلف هذه البيانات من حيث نوعها وطبيعتها باختلاف الظاهرة المدروسة. وفي المجال

التربوي والنفسي يتركز الاهتمام على خصائص الأفراد وسماتهم الجسمية والشخصية والعقلية، وهذه الخصائص تأخذ قيماً مختلفة بالنسبة للأفراد المختلفين، وهذه الخصائص تدعى بالمتغير "Variable"، وبالمقابل إذا كانت هذه الخصائص لها القيمة نفسها عند جميع الأفراد، فإن هذه الخصائص تدعى بالثابت "Constant".

ومنه يعرف **المتغير أنه: "الصفة أو الخاصية التي تأخذ قيماً مختلفة للأفراد المختلفين في المجموعة قيد الدراسة"** أما **الثابت: "فهو الخاصية التي لها القيمة نفسها عند جميع أفراد المجموعة قيد الدراسة".**

يمكن توضيح الفرق بين الثابت والمتغير من خلال المثال، إذا قمنا بقياس ذكاء مجموعة من طلبة الصف الثالث الثانوي بأحد اختبارات الذكاء فإنه من الطبيعي أن يأخذ الطلبة درجات متفاوتة في اختبار الذكاء، هنا تسمى درجة الذكاء المتغير، أما الثابت بين جميع هؤلاء الطلبة هو أنهم من طلبة الثالث الثانوي، أي أن المستوى الصفي يعد ثابتاً.

ولا بد من التعرف إلى المتغيرات وتحديدها من أجل جمع البيانات عن الأفراد أو المفردات.

تصنيف المتغيرات:

تصنف المتغيرات الإحصائية في العلوم النفسية والتربوية تصنيفات عدة ومن هذه التصنيفات:

1-متغيرات إحصائية نوعية وكمية:

أ-متغيرات إحصائية نوعية أو وصفية Qualitative Variables:

تسمى أيضاً متغيرات تصنيفية "Categorized Variable"" فهي تصنف الأفراد أو الأشياء ضمن أصناف أو فئات معينة دون أن تتضمن هذه المتغيرات قيماً كمية كتصنيف الطلاب وفق متغير الجنس إلى ذكور وإناث، والأرقام في هذا المتغير لا تعبر عن كميات من الخصائص أو السمات، وتتميز المتغيرات النوعية أنه لا يوجد تداخل بين تلك الفئات أو التصنيفات، وهذه التصنيفات قد تكون ترتيبية (كالصف الدراسي، المؤهل العلمي) أو غير ترتيبية (كالجنس، مكان الإقامة).

ب-متغيرات إحصائية كمية Quantitative Variables:

إن المتغيرات الكمية هي المتغيرات التي تتعامل مع قيم رقمية من مثل: التحصيل، الذكاء، الطول.

ويستند هذا التصنيف إلى مدلول القيمة الممثلة للسمة أو الخاصية المقاسة، فإذا كانت هذه القيمة تشير إلى مقدار ما في الفرد من خاصية مقارناً بأفراد مجموعته، فإن هذه القيمة تحمل معنى كمياً والمتغير متغير كمي، وإذا كانت القيمة لا تعبر عن مقدار الخاصية عند فرد معين، وإنما تعبر فقط عما إذا كان يمتلك تلك الخاصية

أم لا، أو أنها تشير إلى فئة أو مجموعة مثل الجنس، المرحلة الدراسية، المستوى التحصيلي، فإن هذه المتغيرات هي متغيرات نوعية.

وتصنف المتغيرات الكمية إلى:

-متغيرات كمية متصلة أو مستمرة "Continuous Variables":

هي المتغيرات التي تأخذ أي قيمة ضمن مجال معين، وتأخذ قيماً لا نهائية من القيم المحتملة، وهي المتغيرات التي تأخذ قيمها أرقاماً صحيحة أو كسرية، مثل متغيرات: الطول والوزن والعمر والزمن والحجم والزمن.

-متغيرات كمية منفصلة Discrete Variables:

هي المتغيرات التي تأخذ قيماً محددة، أي المتغير الذي تأخذ قيمه أرقاماً عددية صحيحة قابلة للعد ولا يمكن استخدام الكسور مثل: عدد الطلاب، عدد العمال، عدد الغرف في الفندق، عدد الأبنية في أحد الأحياء، عدد أفراد الأسرة.

2-متغيرات إحصائية مستقلة وتابعة:

يقوم هذا التصنيف على أساس وجود علاقة بين متغيرين وهذه العلاقة تمكن الإحصائي من التنبؤ بقيمة أحد المتغيرين (المتغير التابع أو المتنبأ به) من معرفته لقيمة المتغير الآخر وهو المتغير المستقل أو المتنبئ.

أ-المتغيرات المستقلة" Independent Variable":

المتغير المستقل: هو المتغير الذي يستطيع الباحث أن يتحكم به ويغيره وفقاً لطبيعة البحث، ويؤثر في المتغير التابع، وغالباً ما يكون متغير نوعي أو وصفي ويمكن أن يكون كمياً، فمثلاً إذا كان الباحث يدرس تأثير طرائق تدريس مختلفة في التعليم (طريقة المحاضرة، طريقة النقاش، طريقة التعليم المبرمج) على التحصيل عند عينة من طلبة الصف السابع، فإن طريقة التعليم تعد متغيراً مستقلاً واحداً، والطرائق الثلاث تمثل مستويات المتغير المستقل.

ب-المتغيرات التابعة "Dependent Variable":

المتغير التابع هو: المتغير الذي يتأثر بالمتغير المستقل، والمتغير التابع متغير كمي متصل غالباً، وعلى الباحث ملاحظة تأثير التغير في قيم المتغير المستقل على قيم المتغير التابع، ولا يكون ذلك إلا في حال وجود علاقة ارتباطية بين المتغيرين المستقل والتابع.

والمتغيرات المستقلة تدعى بالعوامل وتباينـاتها تدعى بالمستويات أو الفئات ففي المثال السابق طريقة التعليم عبارة عن عامل أو متغير مستقل والتصنيفات الواقعة ضمن طرائق التدريس (طريقة المحاضرة، طريقة النقاش، طريقة التعليم المبرمج) تدعى بالمستويات أو فئات المتغير، والجدول الآتي يوضح المزيد من الأمثلة عن المتغيرات المستقلة ومستوياتها:

الجدول (1) بعض المتغيرات المستقلة (العوامل) ومستوياتها (فئاتها)

المتغير المستقل أو العامل	فئات المتغير المستقل أو مستوياته
الجنس	ذكور–إناث
المؤهل العلمي	سادس–تاسع–ثانوي–إجازة–دبلوم–دراسات
الحالة الاجتماعية	أعزب–مطلق–متزوج–أرمل
مكان الإقامة	ريف–مدينة
مستوى الدخل	منخفض–متوسط–مرتفع
الصف الدراسي	سابع–ثامن–تاسع

المحاضرة الثانية: جمع البيانات الإحصائية

يمر البحث العلمي بخطوات عدة إذ لا بد بداية من التحديد الدقيق لمشكلة البحث ومعرفة جوانبها كافة وصياغة الفروض حولها، وجمع البيانات الإحصائية عنها، وتصنيفها وعرضها بإحدى طرائق العرض الملائمة وتحليلها باستخدام أساليب إحصائية متنوعة واستخلاص النتائج والتعميمات، وسنتناول جمع البيانات الإحصائية بشيء من التفصيل. إذ تعد عملية جمع البيانات من أهم مراحل البحث الإحصائي وتحتاج إلى كثير من الوقت والجهد والنفقات مقارنة بالمراحل الأخرى، وهذه المرحلة ستبنى عليها كل العمليات التالية، وبالتالي لابد أن تكون هذه البيانات صحيحة ودقيقة ومفصلة وشاملة لكل جوانب الموضوع من أجل التوصل إلى قرارات سليمة.

أولاً-مصادر جمع البيانات:

تقسم البيانات حسب **مصدرها** إلى قسمين هما:

1-المصادر الثانوية لجمع البيانات:

وتسمى أيضاً "بالمصادر التاريخية" وتتضمن البيانات التي جمعت من قبل الباحثين بشكل مسبق، كما تتضمن البيانات الصادرة عن الوزارات والمؤسسات الحكومية والخاصة والهيئات والشركات المختلفة كل في مجال نشاطها، وما يتوفر في المكتبات من مؤلفات ومطبوعات تضم معطيات إحصائية، مثل سجلات الولادات والوفيات، أو السجلات الخاصة بطلبة الجامعات والمدارس وغيرها (الفهارس، الدوريات).

ومن ميزاتها أنها توفر الوقت والجهد والكلفة إلا أن من مساوئها هو مقدار الثقة التي تتصف بها هذه البيانات وأحياناً اختلاف متغيرات هذه البيانات عن المتغيرات التي حددها الباحث.

2-المصادر الأولية لجمع البيانات:

وتسمى أيضاً "بالمصادر الميدانية" ويتم هنا الحصول على البيانات من وحدات المجتمع بصورة مباشرة عن طريق أدوات كالاستبانات والاستمارات والاختبارات المعدة لهذا الغرض والتي يجب أن تتوافر فيها شروط الصلاحية وهي الصدق والثبات. وتتميز هذه المصادر بالدقة والثقة، ما يساعد على الحصول على نتائج موثوقة بدرجة كبيرة، إلا أن هذه المصادر تحتاج إلى وقت وجهد وتكلفة كبيرة حتى يستطيع الباحث الحصول على بيانات وافية ودقيقة.

ثانياً-أساليب جمع البيانات الميدانية:

هناك أسلوبان لجمع البيانات الميدانية:

1-أسلوب الحصر (المسح) الشامل Census:

يجري في هذا الأسلوب جمع البيانات من جميع عناصر أو مفردات المجتمع الإحصائي، وبالتالي تكون النتائج على درجة عالية من الدقة والتفصيل والوضوح، وهي تحتاج إلى تكلفة ووقت كبير، وتحتاج أيضاً إلى وجود العديد من الباحثين، وفي حال المجتمع الكبير جداً قد يصعب أو يستحيل جمع البيانات بهذا الشكل.

2-أسلوب العينة Sample Method:

يجري وفق هذا الأسلوب جمع البيانات من جزء من المجتمع الإحصائي يسمى بالعينة، التي يجب أن تكون ممثلة في خصائصها لخصائص المجتمع قيد الدراسة تمثيلاً صادقاً وسليماً وصحيحاً، ومن ثم يجري تعميم النتائج التي تم الحصول عليها من العينة على المجتمع التي سحبت منه العينة.

ثالثاً-المجتمع الإحصائي والعينة Statistical Population and Sample:

1-المجتمع الإحصائي:

يعرف **المجتمع الإحصائي: جميع الأفراد (أو الأشياء أو العناصر أو المفردات) الذين لهم خصائص واحدة يمكن ملاحظتها، أي أنه المجموعة الشاملة التي يجري اختيار العينة منها.** ولابد من تعريف المجتمع وتحديده تحديداً دقيقاً، أي تحديد ما يشمله المجتمع من مفردات (أفراد، أشياء...)، ويجب أيضاً تحديد خصائص المجتمع من حيث الجنس، والعمر، والمنطقة الجغرافية، والوضع الاقتصادي الاجتماعي. وقد يتسع أو يضيق مفهوم المجتمع طبقاً لتعريف الباحث، فعلى سبيل المثال: تحديد المجتمع بأنه جميع طلبة كلية التربية، أو قد يكون المجتمع أوسع من ذلك فنقول جميع طلبة جامعة دمشق أو حماه، أو جميع طلبة الجامعات السورية الحكومية.

وتجدر الإشارة إلى أنه إذا كان المجتمع صغيراً فإنه يمكن دراسته وتناول جميع عناصره وأفراده، وفي هذه الحالة نقوم بالمسح الشامل. أما إذا كان المجتمع كبيراً ومكون من مجموعة كبيرة جداً من العناصر والأفراد، ولا يمكن تناول جميع عناصره بعملية المسح الشامل، فإنه لابد من الاقتصار على عينة منهم.

2-العينات:

تعرف العينة بأنها: مجموعة جزئية من المجتمع الإحصائي يجري اختيارها بطريقة وحجم نستطيع من خلالها أن نقوم بعملية تعميم النتائج على المجتمع. ويجب أن تكون العينة ممثلة تماماً في خصائصها

لخصائص المجتمع حتى نستطيع القيام بعملية التعميم.

ويجري اختيار العينة من المجتمع الإحصائي نظراً لضخامة المجتمع وصعوبة الوصول إلى جميع أفراده، وبالتالي يسهم سحب العينة في توفير الوقت والجهد والتكاليف التي يمكن أن تنفق في حال دراسة المجتمع بأكمله، وقد تحدد العينة بصورة ضيقة محدودة مثل (النساء المتزوجات من عمر 35-45 سنة اللائي يقطن في مدينة دمشق) أو يمكن أن يحدد بصورة واسعة مثل (النساء السوريات من عمر 18 سنة فما فوق).

رابعاً-خطوات اختيار العينة:

لاختيار عينة من المجتمع لا بد من اتباع أربع خطوات هي:

1-تعريف المجتمع

2-تحديد خصائص المجتمع

3-تحديد حجم عينة كاف لتمثيل خصائص المجتمع

4-اختيار العينة

1-تعريف المجتمع:

فالمجتمع هو الهدف الأساسي من البحث إذ أن الباحث يعمم نتائج العينة في النهاية عليه، وما العينة إلا وسيلة لدراسة خصائص المجتمع، لذلك فالخطوة الأولى في اختيار العينة هي تعريف المجتمع. ويتضمن تعريف المجتمع تحديد خاصية واحدة على الأقل تميزه عن بقية المجتمعات، إذ يمكن أن يكون المجتمع طلاب كلية التربية في جامعة حماه، أو طلاب المرحلة الثانوية في مدارس مدينة حماه.

2-تحديد خصائص المجتمع:

عند تعريف المجتمع نعطي وصفاً عاماً له، وفي هذه الخطوة نضع قائمة بخصائص المجتمع بصورة أكثر تحديداً، أي من وجهة نظر المتغيرات التي تشملها الدراسة، وفيما يأتي عدد من الخصائص (المتغيرات) التي يمكن استخدامها عند القيام ببحث ما (الجنس، العمر، مكان الإقامة، الحالة الاجتماعية، مهنة الأب، مهنة الأم، المستوى التعليمي للأب، المستوى التعليمي للأم، مستوى الدخل...الخ).

فعند اختيار مجتمع البحث وهم طلاب كلية التربية في جامعة حماه للعام الدراسي (2020/2019) يمكن تحديد الخصائص الآتية لهذا المجتمع (الأعداد افتراضية)، كما في الجدول الآتي:

الجدول (1) خصائص المجتمع طلاب كلية التربية في جامعة حماه

المتغير	فئات (مستويات) المتغير	العدد	النسبة المئوية
الجنس	ذكور	600	2500/600 × 100 = 24%
	إناث	1900	76%
	المجموع	2500	100%
السنة الدراسية	الأولى	350	14%
	الثانية	400	16%
	الثالثة	750	30%
	الرابعة	1000	40%
	المجموع	2500	100%

وكذلك عن تحديد أعداد للمجتمع أو العينة لا بد من تحديد النسب المئوية وفق قانون النسبة المئوية:

$$\frac{\text{عدد الأفراد}}{\text{عدد الأفراد الكلي}} \times 100$$

3-تحديد حجم عينة كاف لتمثيل خصائص المجتمع

هناك مجموعة من العوامل التي يجب مراعاتها عند تحديد حجم العينة وهي:

- إن التباين (الاختلاف) بين أفراد المجتمع وليس حجم المجتمع هو العامل الحاسم في تقدير حجم العينة، فكلما كبر حجم التباين بين أفراد المجتمع، كلما زاد حجم العينة بغض النظر عن حجم ذلك المجتمع.
- عند تقدير حجم العينة لابد من الآخذ في الحسبان التكلفة المادية والزمن لإنجاز البحث، قد يضطر الباحث إلى خفض حجم العينة عند ارتفاع تكاليف جمع البيانات.
- مدى التفصيل المطلوب في نتائج العينة فكلما زادت درجة التفصيل المطلوبة زاد حجم العينة المسحوبة.
- مدى الخطأ الذي يسمح به في نتائج العينة، فكلما قل مدى الخطأ الذي يمكن السماح به زاد حجم العينة.
- درجة الثقة المطلوب توفرها في تحقيق السمات السابقة، فكلما زادت درجة الثقة المطلوبة زاد حجم العينة.
- حجم المجتمع: يجب أن يتناسب حجم العينة مع حجم المجتمع الأصلي.
- عندما يكون هناك توقع أن للمعالجة التجريبية أثراً قليلاً على المتغير التابع، أو عندما يتوقع أن تكون العلاقة بين المتغيرات قيد الدراسة فيها ضعيفة فإنه من المرغوب فيه أن تكون العينة كبيرة الحجم.
- عندما يكون هناك توقع لإهدار (فقدان) كبير في عدد المفحوصين، أو هناك عدداً من الأفراد لن يرد الاستبيانات، لابد من اختيار عينة كبيرة الحجم، حتى لا تتأثر النتائج بهذا الفقدان أو الهدر.

- مع الإشارة إلى استخدام عينات ذات حجوم صغيرة في بعض البحوث ربما كان أفضل من استخدام عينات ذات أحجام كبيرة مثل الدراسات التي تعتمد على التحليل النفسي وأدوات القياس الإسقاطي. كما أن عينة كبيرة الحجم في التصميم التجريبي تعد ضعيفة لأن هناك عدد من المتغيرات غير المضبوطة.
- حجم العينة الصغير مقبول في الدراسات الاستطلاعية، إلا أنه في الدراسات التي يترتب عليها اتخاذ قرارات هامة كتوزيع الطلاب في مجموعات أو اتخاذ قرار يتطلب أموال طائلة فمن الواجب الحصول على عينة كبيرة بشكل كاف للتقليل من الخطأ في النتائج قدر الإمكان.
- إذا لم تكن طرائق أو وسائل جمع البيانات دقيقة بدرجة عالية، فلا بد من الحصول على عينة كبيرة لتعويض الخطأ المتحصل في جمع البيانات، كما يتأثر حجم العينة بنوع الأداة المستخدمة في جمع البيانات، فإذا كان الباحث يستخدم المقابلة أو الملاحظة أو الاختبار الفردي فإنه يقلل حجم العينة حتى يقلل من الوقت والجهد والنفقات، أما إذا كان يستخدم الاستبانة أو الاختبار الجمعي فإنه يستخدم أعداد كبيرة لسهولة جمع البيانات باستخدام الاستبانة أو الاختبارات الجمعية.

وباختصار هناك اتجاهان لتحديد حجم العينة، الاتجاه الأول: يعتمد على خبرات الباحث وخبرات الباحثين الآخرين في هذا المجال، ووفقاً لهذا الأسلوب فإننا نجد أن الإحصائيين ينصحون بأن يكون الحد الأدنى للأفراد كما يأتي (50–100) فرد في الدراسات الارتباطية "دراسات العلاقة"، بينما يفضل أن يكون حجم العينة في الدراسات التجريبية ذات المعالجة الواحدة 30 فرداً، أما في الدراسات التجريبية ذات المعالجتين أو أكثر فيفضل ألا يقل عدد الأفراد في الخلية الواحدة عن 5 أفراد. أما في **الدراسات المسحية:** فالعدد المناسب 20% من أفراد المجتمع إذا كان عدد المجتمع (500 إلى 1000) وتنخفض النسبة إلى 5% في المجتمعات الكبيرة جداً.

والاتجاه الثاني: يعتمد على القواعد الاحتمالية الرياضية لتحديد حجم العينة، فهناك عدد من المعادلات الإحصائية التي يمكن استخدامها لتحديد حجم العينة المناسب، ويذكر كريستنسن (Christensen, 1997) أن قسم البحوث بالرابطة الأمريكية للتربية (NEA) وضع معادلة لتحديد حجم العينة حتى تكون ممثلة للمجتمع، وقد قام كيريسي ومورجان (Kerjcie & Morgan, 1970) بحساب حجم العينة اللازمة لمجتمعات يصل حجمها إلى مليون فرد عندما يكون مستوى الدلالة 0.05.

وتبين الجداول الآتية حجم العينة المطلوب عندما يكون حجم المجتمع محدداً:

حجم العينة عند معرفة حجم مجتمع الدراسة عند مستوى دلالة (0.05)

حجم مجتمع الدراسة	حجم العينة	حجم مجتمع الدراسة	حجم العينة	حجم مجتمع الدراسة	حجم العينة
10	10	220	140	1200	291
15	14	230	144	1300	297
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	169	2000	322
55	48	320	175	2200	327
60	52	340	181	2400	331
65	56	360	186	2600	335
70	59	380	191	2800	338
75	63	400	196	3000	341
80	66	420	201	3500	346
85	70	440	205	4000	351
90	73	460	210	4500	354
95	76	480	214	5000	357
100	80	500	217	6000	361
110	86	550	226	7000	364
120	92	600	243	8000	367
130	97	650	242	9000	368
140	103	700	248	10000	370
150	108	750	254	15000	375
160	113	800	260	20000	377
170	118	850	265	30000	379
180	123	900	269	40000	380
190	127	950	274	50000	381
200	132	1000	278	75000	382
210	136	1100	285	100000	383

حجم العينة عند معرفة حجم مجتمع الدراسة عند مستويات دلالة مختلفة

حجم مجتمع الدراسة	حجم العينة عند مستوى 0.05	حجم العينة عند مستوى 0.04	حجم العينة عند مستوى 0.03	حجم العينة عند مستوى 0.02	حجم العينة عند مستوى 0.01
200	132	150	168	185	196
300	166	200	234	276	291
400	196	240	291	343	384
500	217	273	340	414	475
600	234	300	384	480	565
700	248	323	423	542	652
800	260	343	457	600	738
900	269	360	488	655	823
1000	278	375	516	706	906
2000	322	462	696	1091	1655
3000	341	500	787	1334	2286
4000	350	522	842	1500	2824
5000	357	536	879	1622	3288
6000	361	546	906	1715	3693
7000	364	553	926	1788	4049
8000	367	558	942	1847	4364
9000	368	563	954	1895	4646
10000	370	566	964	1936	4899
15000	375	577	996	2070	5855
20000	377	583	1013	2144	4688
25000	378	586	1023	2191	6938
30000	379	588	1030	2223	7275
40000	381	591	1039	2265	7745
50000	381	593	1045	2291	8056
75000	382	595	1052	2327	8514
100000	383	597	1056	2354	8762
500000	384	600	1065	2390	9423
1000000	384	600	1066	2395	9513

وحجم العينة الموضح في الجدولين السابقين يمثل الحد الأدنى المقبول لحجم العينة المناسب لمجتمع دراسة، وعند اختيار عينة البحث يجب ألا يقل حجم العينة عن هذا الحد، مع ملاحظة أنه زيادة حجم العينة عن هذا الحد يمكن أن يوفر تمثيلاً أعلى لخصائص المجتمع والتالي تعميماً أصدق لنتائج البحث.

4-اختيار العينة:

والتي تعد خطوة مهمة للغاية إذ يتوقف على مدى سلامة اختيار العينة إمكانية تعميم النتائج على المجتمع، والعينة السليمة هي العينة الممثلة للمجتمع الذي اختيرت منه، وعملية اختيار العينة عملية منظمة فهناك عدة أساليب أو طرائق لاختيار العينة، وبعض هذه الأساليب أنسب لمواقف معينة، لأن كل أسلوب لا يعطي نفس نتائج الأساليب الأخرى ولا يضمن تمثيل العينة بنفس الدقة، وعليه يجب أن نوازن بين الأساليب المختلفة لاختيار العينة ونختار الأسلوب الأفضل والأنسب لظروف بحثنا.

وتقسم العينات إلى نوعين رئيسين:

أ-العينات العشوائية أو الاحتمالية Probabilistic samples

ب-العينات غير العشوائية أو غير الاحتمالية Non Probability Sample

أ-العينات العشوائية أو الاحتمالية Probabilistic samples:

وهي العينات التي يتم اختيارها باتّباع طرائق علمية غير متحيزة تستند إلى مبادئ الاحتمالات بغرض تمثيل المجتمع.

وللعينات العشوائية أنواع عدة هي:

-العينة العشوائية البسيطة Simple Random sample:

وهي العينة التي تتاح فيها الفرصة لكل مفردة من مفردات المجتمع بالظهور في العينة وتكون الفرص متكافئة لا انحياز فيها، أي أن احتمال سحب أي مفردة من مفردات المجتمع يكون ثابتاً ويساوي

عدد مرات تكرار المفردة (عدد أفراد العينة)

عدد أفراد المجتمع

فمثلاً إذا أردنا سحب 50 طالباً من أصل ألف طالب فإن احتمال سحب كل طالب= $\frac{50}{1000} = \frac{5}{100} = 5\%$

وهناك طرائق عدة لاختيار العينة العشوائية البسيطة منها: **طريقة القرعة:** وفي هذه الطريقة تكتب أسماء جميع أفراد المجتمع في بطاقات صغيرة وتخلط مع بعضها، ويجري سحب المطلوب من هذه البطاقات دون تحيز. **وهناك جداول الأعداد العشوائية، ودواليب الحظ أو اليانصيب.**

-العينة العشوائية المنتظمة Systematic Random Sample:

وفيها يتم اختيار الفرد أو العنصر على مسافات متساوية من قائمة أسماء المجتمع. فإذا توافرت قوائم بأسماء

جميع أفراد المجتمع وكان هؤلاء مرتبين بصورة عشوائية فإن الباحث يمكن أن يختار عينة دراسته بصورة دورية منتظمة، ولكن بداية لا بد من تحديد طول فترة السحب K من خلال قسمة حجم المجتمع Nعلى حجم العينة n: أي $K=\frac{N}{n}$ ثم نختار مفردة بشكل عشوائي من المفردات ضمن فترة السحب مثلاً عن طريق القرعة، وتكون هذ المفردة هي المفردة الأولى في العينة، وللحصول على باقي مفردات العينة نضيف إليها فترة السحب وهكذا.

وعند عدم توفر قوائم بأسماء الأفراد وليكن (طلبة المدارس) فإن الباحث يمكن له الوقوف أمام المدرسة واختيار كل عاشر طالب يدخل المدرسة إذا كانت فترة السحب 10 أو كل خامس طالب حسب فترة السحب.

ومن **مزايا هذه العينة** قلة التكاليف وسهولة سحبها، ولكن من **مساوئها** أنه لا يكن سحبها إلا إذا كان المجتمع محدوداً ومرتباً ومرقماً ومعلوماً جميع مفرداته.

خطوات العينة العشوائية المنتظمة:

1-تعريف وتحديد المجتمع: طلبة السنة الرابعة في كلية التربية جامعة حماه للعام الدراسي 2020/2019 والبالغ عددهم (1000) طالباً وطالبة.

2-تحديد حجم العينة المرغوب فيها: 200 طالب وطالبة.

3-الحصول على قائمة بأسماء أفراد المجتمع.

4-تحديد المسافة بين أفراد العينة (طول فترة السحب) وذلك بقسمة عدد أفراد المجتمع على عدد أفراد العينة: 200/1000=5، أي أن المسافة بين الفرد والفرد الذي تليه هي 5.

5-اختيار رقم عشوائي في حدود المسافة التي حصلنا عليها في الخطوة السابقة: وليكن 3.

6-البدء من هذا الرقم واختيار الأفراد على مسافات متساوية حتى نحصل على عدد العينة كاملاً: 3، 8، 13، 18، 23، 28، 33، 38، 43، 48، 53، 58، 63، 68...الخ.

مثال آخر: إذا أردنا اختيار عينة عشوائية منتظمة مؤلفة من 50 فرداً من أصل 500 فرداً فإن طول فترة السحب سوف تتحدد بقسمة حجم المجتمع إلى حجم العينة أي $\frac{500}{50}=10$، وهذا يعني أننا سوف نأخذ فرداً من بين كل عشر أفراد متتاليين، على أن نحدد رقم البداية عشوائياً، فإذا حددنا رقم البداية بالرقم 5 فإن الأرقام التالية التي يجب سحبها هي 5، 15، 25، 35، 45،...

-العينة العشوائية الطبقية Satisfied Random Sample:

عندما يكون المجتمع مقسماً إلى مجموعات غير متجانسة في الخصائص التي يقوم الباحث بدراستها ومن أجل ضمان التمثيل الجيد لكل مجموعة من هذه المجموعات ضمن العينة، نقوم بتقسيم المجتمع إلى طبقات على أساس متغير واحد مثل (الجنس أو العمر، المستوى الاقتصادي الاجتماعي، المهنة ...) أو أكثر من متغير وذلك حسب موضوع البحث على أن تكون كل طبقة متجانسة فيما بينها قدر الإمكان، ومن ثم يسحب عينة من كل طبقة إما بطريقة عشوائية بسيطة أم بطريقة عشوائية منتظمة.

ومن ميزات العينة العشوائية الطبقية أنها تساعد في الحصول على عينة أكثر تمثيلاً لمجتمع البحث الذي اختيرت منه، إضافة إلى الدقة العالية في النتائج المتحصلة منها، وتستخدم في المجتمعات الكبيرة الحجم.

ولكن من مساوئ هذه الطريقة أنه لا يمكن سحبها إلا إذا كان معلوماً جميع مفردات أو وحدات المجتمع وعددها وكذلك مفردات كل طبقة وعددها.

فمثلاً إذا كان المجتمع هم طلبة كلية التربية في جامعة حماه وقد تم تقسيم الطلبة إلى أربع طبقات (حسب متغير السنة الدراسية) وكذلك تقسيم الطلبة في كل طبقة أو سنة دراسية إلى طبقتين (حسب متغير الجنس)، ويمكننا بعد ذلك الحصول على عينة عشوائية بسيطة من أصغر خلية، فنحصل في النهاية على عينة طبقية عشوائية ممثلة لخصائص المجتمع، وعند سحب العينة من كل خلية فإننا نتبع أحد الأسلوبين:

-**المعاينة النسبية:** نسحب من كل طبقة عدداً يتناسب مع نسبة عدد أفراد الطبقة إلى المجتمع، وذلك على أساس المعادلة الآتية:

حجم العينة الطبقية = (حجم الطبقة ÷ حجم المجتمع) × حجم العينة

-**المعاينة المتساوية:** نحصل على أعداد متساوية من كل طبقة وبطريقة عشوائية، ورغم بساطة هذا الأسلوب إلا أنه يؤدي إلى اختلال التوزيع النسبي للخصائص.

خطوات اختيار العينة العشوائية الطبقية:

1-تحديد وتعريف المجتمع: وهم طلبة كلية التربية في جامعة حماه والبالغ عددهم (2500) طالباً وطالبة من السنوات الدراسية الأربع.

2-تحديد حجم العينة: 400 طالبٍ وطالبة.

3-تحديد المجموعات الفرعية بناء على خصائص المجتمع (السنة الدراسية والجنس) التي تم تحديدها في الخطوة 1، وقد تم تقسيم عينة البحث إلى طبقتين:

* طبقة السنة الدراسية (أولى، ثانية، ثالثة، رابعة).

* طبقة الجنس (ذكور، إناث).

4-تصنيف أفراد المجتمع وفقاً للمجموعات الفرعية السابق تحديدها، بحيث ينتمي كل فرد لمجموعة واحدة فقط، وبالتالي عدم تداخل المجموعات.

5-اختيار عينة عشوائية بسيطة من كل مجموعة فرعية.

ويبين الجدول الآتي اختيار **العينة العشوائية الطبقية المتساوية:**

الجدول (2) العينة العشوائية الطبقية المتساوية

المجتمع (طلبة كلية التربية وعددهم 2500 طالبٍ وطالبة)							
التصنيف حسب السنة الدراسية							
السنة الأولى		**السنة الثانية**		**السنة الثالثة**		**السنة الرابعة**	
350		**400**		**750**		**100**	
اختيار عشوائي		اختيار عشوائي		اختيار عشوائي		اختيار عشوائي	
100 طالب		**100 طالب**		**100 طالب**		**100 طالب**	
تعيين عشوائي		تعيين عشوائي		تعيين عشوائي		تعيين عشوائي	
ذكور	**إناث**	**ذكور**	**إناث**	**ذكور**	**إناث**	**ذكور**	**إناث**
50	**50**	**50**	**50**	**50**	**50**	**50**	**50**

كما يمكن اختيار **العينة العشوائية الطبقية النسبية**، ويبين الجدول الآتي أعداد المجتمع والنسب المئوية:

الجدول (3) توزع المجتمع والنسب المئوية في كل طبقة

المتغير	**فئات (مستويات) المتغير**	**العدد**	**النسبة المئوية**	**متغير الجنس**	**العدد**	**النسبة المئوية**
السنة الدراسية	**الأولى**	**350**	**%14**	**ذكور**	**75**	**%3**
				إناث	**275**	**%11**
	الثانية	**400**	**%16**	**ذكور**	**125**	**%5**
				إناث	**275**	**%11**
	الثالثة	**750**	**%30**	**ذكور**	**150**	**%6**

				إناث	600	%24
	الرابعة	1000	%40	ذكور	250	%10
				إناث	750	%30
	المجموع	2500	%100	ذكور	600	%24
				إناث	1900	%76

ومنه يتم اختيار **العينة العشوائية الطبقية النسبية**، باستخدام القانون الآتي:

حجم العينة الطبقية = (حجم الطبقة ÷ حجم المجتمع) × حجم العينة

ويبين الجدول الآتي أعداد العينة ونسبها المئوية المطابقة لنسب المجتمع:

الجدول (4) توزع العينة والنسب المئوية في كل طبقة

المتغير	فئات (مستويات) المتغير	العدد	النسبة المئوية	متغير الجنس	العدد	النسبة المئوية
السنة الدراسية	الأولى	56=400×2500/350	%14	ذكور	12=400×2500/75	%3
				إناث	44	%11
	الثانية	64	%16	ذكور	20	%5
				إناث	44	%11
	الثالثة	120	%30	ذكور	24	%6
				إناث	96	%24
	الرابعة	160	%40	ذكور	40	%10
				إناث	120	%30
	المجموع	400	%100	ذكور	96	%24
				إناث	304	%76

مثال: إذا أراد باحث أن يستطلع آراء طلبة كلية التربية في جامع دمشق حول جدول دوام المحاضرات، فلا بد له من أخذ آراء الطلاب من جميع التخصصات، وكان أعداد الطلاب في التخصصات المختلفة كما يأتي:

الجدول (5) أعداد الطلاب في التخصصات المختلفة

الاختصاص	عدد الطلبة
تربية	2000
علم النفس	1500
رياض أطفال	200

إرشاد نفسي	1000
معلم صف	2300
المجموع	7000

ولتحقيق هذا الغرض اختار عينة مكونة من 100 طالب، وأخذ في الحسبان عدد الطلبة في كل اختصاص فيكون الاختيار كما يأتي:

من طلبة التربية $\frac{2000}{7000} \times 100 = 28.57 \approx 29 =$ طالباً

علم نفس = $21.42 \approx 21$ طالباً

معلم صف $\approx$ 33 طالباً

إرشاد نفسي $\approx$ 14 طالباً

رياض أطفال $\approx$ 3

-العينة العشوائية العنقودية Cluster Random Sample:

تستند طريقة سحب هذه العينة العنقودية إلى تقسيم مجتمع البحث إلى عدد من العناقيد (مجموعات) وفق معيار معين غالباً ما يكون جغرافياً وهذه المجموعات (ليست طبقات ذات خصائص مشتركة) تسمى كل مجموعة عنقوداً وتكون وحدة الاختيار في هذا النوع من العينات هي العنقود، ويتم اختيار هذا العنقود عشوائياً وبهذا تكون **وحدة السحب في هذه العينة هي المجموعة وليس الفرد كما هو الحال في العينات السابقة**. وفي بعض الأحيان نختار العينة من جميع أفراد هذه الوحدات وتسمى حينها عنقودية ذات المرحلة الواحدة وفي حالات أخرى تقسم هذه الوحدات إلى وحدات جزئية، ثم تسحب العينة بمرحلة إضافية وتسمى عملية سحب العينة هنا بالعينة العنقودية ذات المراحل المتعددة.

ومن الأنسب استخدام العينة العشوائية العنقودية عندما تكون المجتمعات كبيرة جداً أو منتشرة على مساحة جغرافية واسعة. وعادة ما تستغرق العينة العنقودية وقتاً قصيراً وتكاليفاً محدودة وسهولة في الحصول عليها.

خطوات اختيار العينة العشوائية العنقودية:

بفرض أن باحثاً يريد إجراء بحث على مجتمع المدرسين في أحد المناطق التعليمية، ومكون من 5000 مدرس، ويريد الحصول على عينة مكونة من 500 مدرس باستخدام العينة العشوائية العنقودية، فإن ذلك يتم وفق الخطوات الآتية:

1-تعريف وتحديد خصائص المجتمع: وهو جميع المدرسين (5000) مدرس في المنطقة التعليمية.

2-تحديد حجم العينة المرغوب فيه: وهو 500 مدرس.

3-تعريف وتحديد العنقود: وهو المدرسة.

4-عمل حصر لكل العناقيد، أو وضع قائمة بالعناقيد التي يتكون منها المجتمع: 100 مدرسة أو عنقود.

5-تقدير عدد الأفراد في كل عنقود: وسطياً 50 مدرس في المدرسة الواحدة.

6-تحديد عدد العناقيد بقسمة عدد أفراد العينة على العدد التقديري للأفراد في كل عنقود: 50/500= 10 أي عدد المدارس المطلوبة 10 مدارس.

7-اختيار عدد من العناقيد عشوائياً.

8-عدد الأفراد في العينة هم جميع الأفراد الذين تشملهم العناقيد المختارة عشوائياً: أي هم جميع المدرسين في هذه المدارس العشر، وبذلك نحصل على عينة عنقودية ذات المرحلة الواحدة.

والمثال التالي يوضح كيفية سحب عينة عنقودية ذات المراحل المتعددة:

لنفرض أن مدينة دمشق مقسمة إلى عدة مناطق جغرافية ولتكن ثمان مناطق (عناقيد) وجرى اختيار أربع مناطق منها عشوائياً، ولتكن هذه المناطق (الجنوب، الشمال، الشرق، الغرب).

يتم اختيار عدد من المدارس من كل عنقود (منطقة) باستخدام إحدى الطرائق العشوائية، ولنفرض أن حجم العينات الجزئية كما يأتي:

الجدول (6) عدد المدارس المسحوبة من كل منطقة

المنطقة	عدد المدارس
الشرق	3
الجنوب	5
الشمال	5
الغرب	4
المجموع	17

وفي مرحلة لاحقة يمكن أن يختار عدد من الصفوف من كل مدرسة ولتكن على النحو الآتي:

الجدول (7) عدد المدارس والصفوف المسحوبة من كل منطقة

المنطقة	عدد المدارس	عدد الصفوف
الشرق	3	2
الجنوب	5	3
الشمال	5	3

الغرب	4	2
المجموع	17	10

وبالتالي يصبح عدد الوحدات الجزئية (العناقيد) الجزئية عشر صفوف ثم يتبع أسلوب الحصر الشامل أي يختار جميع الطلبة في تلك الصفوف.

-العينة العشوائية المتعددة المراحل Multi-Stage Sample:

عندما تكون المجتمعات كبيرة الحجم وعندما يكون الوقت المتاح قصيراً والإمكانيات المادية المتاحة قليلة، فإنه يصعب سحب عينة عشوائية بسيطة أو منتظمة أو طبقية، ولكن بدلاً منها تستخدم العينة العشوائية متعددة المراحل، ويناسب هذا النوع من العينات الدراسات السكانية أو الدراسات في مجال الجغرافيا الاقتصادية.

فإذا أردنا سحب عينة متعدد المراحل لدراسة متوسط دخل الأسر أو متوسط إنفاقها في الجمهورية العربية السورية نختار محافظات عدة بشكل عشوائي (مرحلة أولى)، بعد ذلك من ضمن المحافظات المختارة في المرحلة الأولى نختار مناطق عدة (مرحلة ثانية)، ثم نختار قرى وأحياء بشكل عشوائي من المناطق المختارة في المرحلة الثانية (مرحلة ثالثة)، وثم نختار عدداً من الأسر في تلك القرى والأحياء المختارة في المرحلة الثالثة عشوائياً (مرحلة رابعة) ونجمع البيانات من هذه الأسر فقط المختارة بالمرحلة الرابعة.

ومن عيوب هذه العينة كثرة عدد المراحل التي قد تتضمنها، مما يضعف العلاقة بين معالم المجتمع الأصلي وخصائص العينة، ويؤثر في تقدير معالم المجتمع من بيانات العينة المتحصل عليها في آخر مرحلة. كما يتطلب هذا النوع من العينات جهداً كبيراً في تحديد حدود أو إطار كل مرحلة وتحديد حجم العينة الفرعية المطلوبة من كل منها.

ب-العينات غير العشوائية (غير الاحتمالية) Non Probability Sample:

يضطر الباحث أحياناً لاتّباع التحيز الشخصي أو القصدي في اختيار العينة، أو ما يعرف بأسلوب المعاينة غير الاحتمالية، إما لعدم توفر الإمكانات المادية أو لوجود صعوبة في الحصول على مفردات المجتمع، وفي حالة العينات غير الاحتمالية يجري السحب دون مراعاة تكافؤ الفرص لمفردات المجتمع للظهور في العينة.

وعادة تستخدم العينات غير العشوائية، في الأبحاث الاستطلاعية، بهدف التأكد من وضوح وسلامة الأدوات المستخدمة في الأبحاث، أو في مرحلة تجريب الاختبارات القبلية للتأكد من وضوح التعليمات والأسئلة حتى يمكن إجراء التعديلات اللازمة في الأسئلة قبل سحب العينة الأساسية.

والعينات غير الاحتمالية على أنواع هي:

-العينة المصادفة أو العرضية Accidental sample:

في بعض الظروف قد تتوافر للباحث عينة عن طريق الصدفة ممن يصادفهم من الأشخاص، ويستخدم الباحث الحالات على أساس سهولة الحصول عليها، مثل الإعلامي الذي يستطلع آراء بعض الناس الذين يصادفهم في طريقه. أو يأخذ الباحث عينته من طلبة المدرسة القريبة من بيته ويجري الدراسة عليهم، وتفيد هذه العينة في تكوين فكرة أولية عن موضوع البحث، وبالتالي الاستناد إلى هذه الفكرة في صياغة أهداف البحث وفرضياته، إلا أن نتائج هذه العينة لا يمكن تعميمها وبناء أحكام عليها لأن نتائجها لا تتعدى الإطار الضيق الذي خضع له الباحث.

-العينة الغرضية أو المقصودة Purposeful:

في بعض الأحيان يلزم الحصول على معلومات من فئات محددة، بمعنى أنه لا يمكن الحصول على المعلومات المطلوبة إلا من هؤلاء الأشخاص أو أنهم أناس تتوافر لديهم المعايير المحددة من قبل الباحث، أو الأفراد المناسبين بناء على غرض خاص في عقل الباحث، من مثل سحب عينة من مدراء المدارس في خاصية معينة للكشف عن مشكلات التعليم في هذه الناحية، أو دراسة واقع التعليم الجامعي من وجهة نظر المشرفين عليه في الجامعات الرسمية في بلد ما، الأمر الذي يستلزم سحب عينة من رؤساء الجامعات ونوابهم إضافة إلى عمداء الكليات ونوابهم، أو دراسة عينة من ذوي الحاجات الخاصة، أو المدمنين.

ومن العينات غير العشوائية أيضاً: (عينة الحصص، عينة الحالات المتطرفة، العينة المتتابعة، عينة الكتلة، العينة الاجتهادية، العينة النظرية).

المحاضرتين الثالثة والرابعة: أساليب تنظيم البيانات وعرضها

مقدمة:

بعد القيام بعملية جمع البيانات من مصادر مختلفة لابد من تنظيمها وعرضها بصورة ملخصة يسهل فهمها وقراءتها ومن ثم تحليلها واستخراج النتائج.

وتنظيم البيانات هذه يأخذ أشكالاً عدة منها الجداول أو التوزيعات التكرارية ومنها الأشكال والرسوم البيانية. وسيتم في هذا الفصل توضيح كل من هذه الأساليب.

أولاً-أساليب تبويب بيانات المتغيرات النوعية (الوصفية):

التبويب (Tabulaion): هو وضع البيانات الإحصائية في جداول.

يتم في المتغيرات النوعية أو التصنيفية تصنيف الأفراد أو الأشياء في فئات أو أصناف، ويمكن أن تكون هذه المتغيرات ترتيبية (كالمؤهل العلمي أو الصف الدراسي) ويمكن أن تكون غير ترتيبية (الجنس، المحافظة، الحالة الاجتماعية)، ويتم عرض المتغيرات النوعية وفق:

1-الجداول أو التوزيعات التكرارية للمتغيرات النوعية:

يعني العرض الجدولي للبيانات تنظيم هذه البيانات في صورة جداول تسمى بالجداول التكرارية، وهذه الجداول قد تكون جداول بسيطة إذا تناولت ظاهرة واحدة، وقد تكون جداول "مزدوجة" إذا تناولت ظاهرتين أو دراسة العلاقة بين ظاهرتين، أو قد تكون الجداول "مركبة" إذا تناولت أكثر من ظاهرتين، وتفيد هذه الجداول (البسيطة أو المزدوجة أو المركبة) في تعرف خصائص البيانات وتكون المنطلق والأساس للعملية الإحصائية المختلفة للبيانات.

يتألف الجدول التكراري البسيط من عمودين: العمود الأول قائمة بتصنيفات المتغير، والثاني بتكرار كل تصنيف من تصنيفات المتغير، ويمكن إضافة عمود ثالث يمثل النسبة المئوية للأوجه المختلفة للبيانات.

مثال (1): فيما يأتي جدول بالحالة الاجتماعية (متغير نوعي غير ترتيبي) لـ 30 فرداً:

الجدول (1) الحالة الاجتماعية لعينة من 30 فرداً

أعزب	أعزب	أعزب	أعزب	أعزب	أعزب	أعزب	أعزب	أعزب			
مطلق	مطلق	مطلق	مطلق	مطلق							
أرمل	أرمل	أرمل									
متزوج	متزوج	متزوج	متزوج	متزوج	متزوج	متزوج	متزوج	متزوج	متزوج	متزوج	متزوج

ويتم إنشاء جدول تكراري لبيانات المتغير السابق وفق الآتي:

الجدول (2) توزيع عينة من الأفراد وفق متغير الحالة الاجتماعية ل 30 فرداً

الحالة الاجتماعية	الإشارات	عدد الأفراد (التكرار) Frequency	النسبة % Percentage
أعزب	//// ///	9	%30 =100 ×30/9
مطلق	/ ///	6	%20
أرمل	///	3	%10
متزوج	// /// ///	12	%40
المجموع		30	%100

كل إشارة / تشير إلى فرد وعندما يبلغ عدد الإشارات 5 نشكل منها حزمة بوضع الإشارة الخامسة بشكل مائل ///، وعدد الإشارات هو عدد الأفراد في عمود

النسبة المئوية = (عدد الأفراد / العدد الكلي للأفراد) × 100

مثال (2): فيما يأتي جدول بالمستوى التعليمي (متغير نوعي ترتيبي) لعينة من 50 فرداً:

الجدول (3) جدول بالمستوى التعليمي لعينة من 50 فرداً

										دكتوراه	دكتوراه	دكتوراه	دكتوراه	دكتوراه
								ماجستير	ماجستير	ماجستير	ماجستير	ماجستير	ماجستير	ماجستير
				إجازة	إجازة	إجازة	إجازة	إجازة	إجازة	إجازة	إجازة	إجازة	إجازة	إجازة
معهد	معهد	معهد	معهد	معهد	معهد	معهد	معهد	معهد	معهد	معهد	معهد	معهد	معهد	معهد
			ثانوي	ثانوي	ثانوي	ثانوي	ثانوي	ثانوي	ثانوي	ثانوي	ثانوي	ثانوي	ثانوي	ثانوي

الجدول (4) توزيع عينة من الأفراد وفق متغير المؤهل التعليمي ل 50 فرداً

المستوى التعليمي	عدد الأفراد (التكرار) Frequency	النسبة % Percentage
دكتوراه	5	%10
ماجستير	7	%14
إجازة	11	%22
معهد	15	%30
ثانوي	12	%24
المجموع	50	%100

ملاحظة: يمكن إضافة عمود للإشارات للجدول السابق

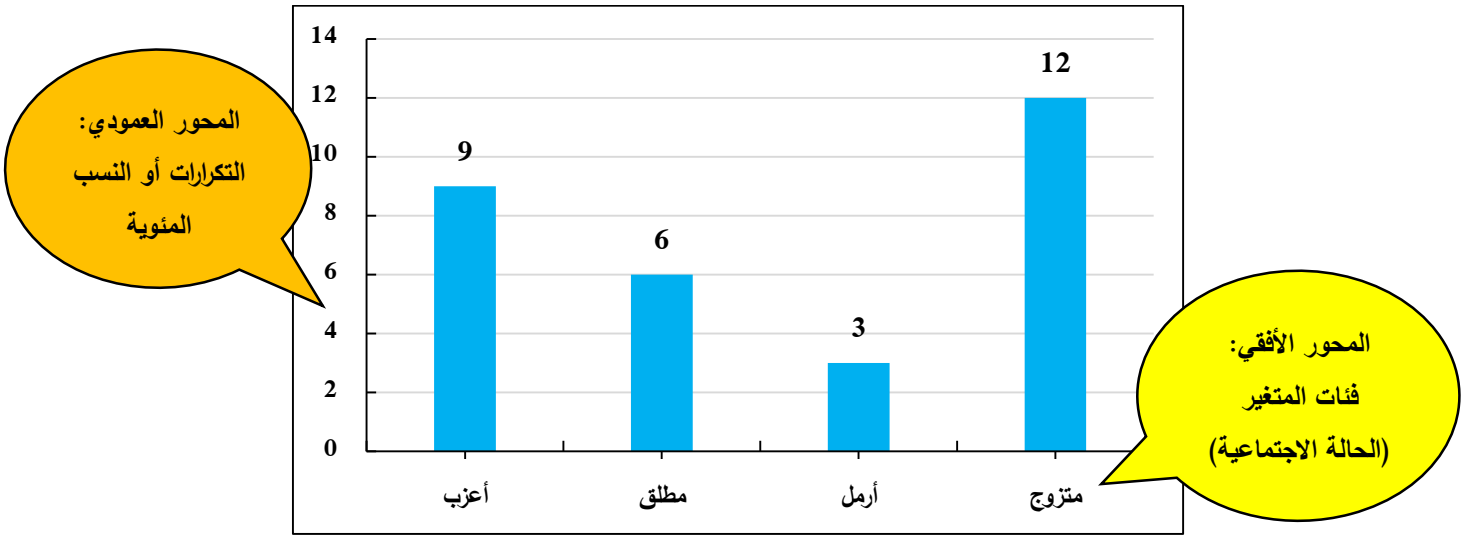
2-العرض البياني للمتغيرات النوعية Graphical Presentation:

يقصد بالعرض البياني استخدام الرسوم للتعبير عن البيانات العددية، تعد أساليب العرض البياني وسيلة إضافية في وصف المتغيرات إلى جانب الجداول التكرارية، إذ يساعد العرض البياني في إعطاء صورة سريعة أقرب إلى الذهن والتذكر عن الظاهرة المدروسة، ومن أساليب عرض المتغيرات النوعية بيانياً:

أ-طريقة الأعمدة البيانية Bar Chart:

تُمثل البيانات باستخدام الأعمدة برسم محورين متعامدين المحور الأفقي والمحور العمودي، وتوضع فئات المتغير على المحور الأفقي والتكرارات (أو النسب المئوية) على المحور العمودي.

والشكل الآتي يمثل رسماً بيانياً لمتغير الحالة الاجتماعية الذي ذكر في المثال (1).



الشكل (1) التمثيل البياني وفق طريقة الأعمدة البيانية لبيانات نوعية (الحالة الاجتماعية)

من الشكل البياني السابق يلاحظ أن عدد الأفراد المتزوجين أكبر من عدد بقية الأفراد ونظراً إلى أن البيانات نوعية فهناك فجوات بين الأعمدة.

ب–طريقة الدوائر Circle Chart:

تستخدم طريقة الدوائر للعرض البياني من أجل إظهار **نسب** توزيع الظاهرة والمقارنة بينها، حيث تمثل الظاهرة بدائرة وتقسم إلى قطاعات تتناسب مساحتها مع حجم الأجزاء المكونة للظاهرة، وتلون القطاعات بألوان مختلفة لتمييزها عن بعضها بعضاً، ويتم تحديد مساحة كل جزء عن طريق تقسيم الزوايا المركزية للدائرة والبالغة 360 درجة إلى أجزاء، حيث مجموع أجزاء الظاهرة يساوي 360، وبالتناسب نستطيع حساب زاوية كل قطاع، فمثلاً يمكن إنشاء رسم بياني بطريقة الدوائر للبيانات الواردة في المثال (2)، ويمكن تحديد مساحة كل جزء أو كل فئة وفق الآتي: عدد الأفراد ككل (50) فرداً، ومجموع الزوايا (360)، وعدد الأفراد في فئة الدكتوراه (5) أفراد ومنه نحسب زاوية فئة الدكتوراه:

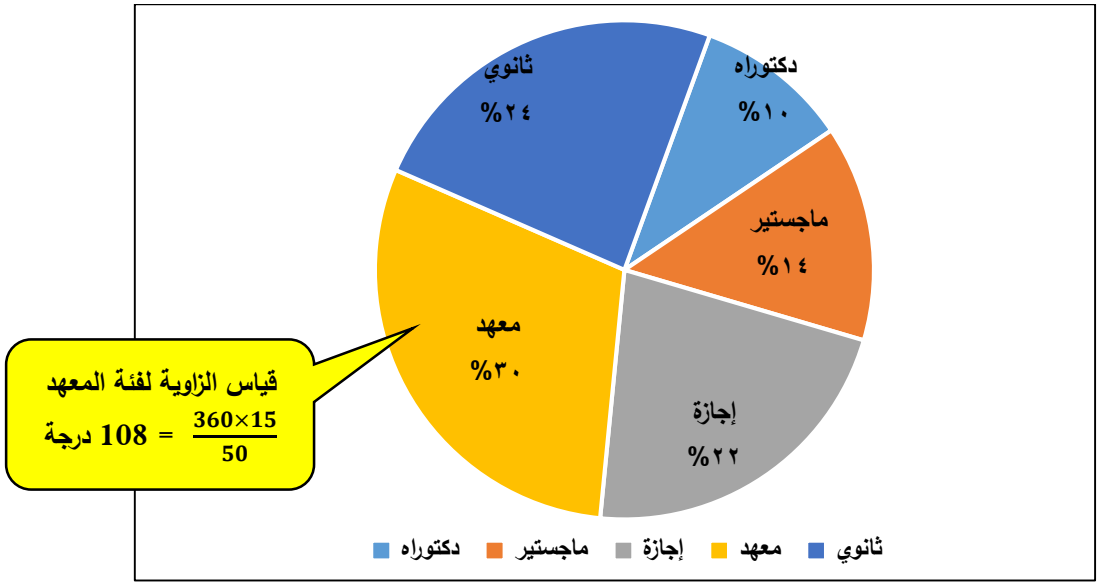
50 تقابل 360

5 تقابل X

ومنه $X = \frac{360\times5}{50}$ = 36 درجة.

أي أن زاوية القطاع الممثل لفئة الدكتوراه تبلغ (36 درجة) وهكذا نحسب الزوايا كلها.

ويتم تمثيل الأجزاء بشكل مئوي (أي ذكر النسبة المئوية المقابلة لك فئة بعد تحديد زاوية رسمها).



الشكل (2) التمثيل البياني وفق طريقة الدوائر لبيانات نوعية (المؤهل التعليمي)

ثانياً-أساليب تبويب بيانات المتغيرات الكمية:

الحالة الأولى: في حال المتغيرات الكمية المنفصلة:

فالمتغير الكمي المنفصل هو المتغير الذي تأخذ قيمه أرقاماً صحيحة قابلة للعد ولا يمكن استخدام الكسور فيها مثل عدد الطلاب، عدد أفراد الأسرة، عدد أيام الغياب...إلخ.

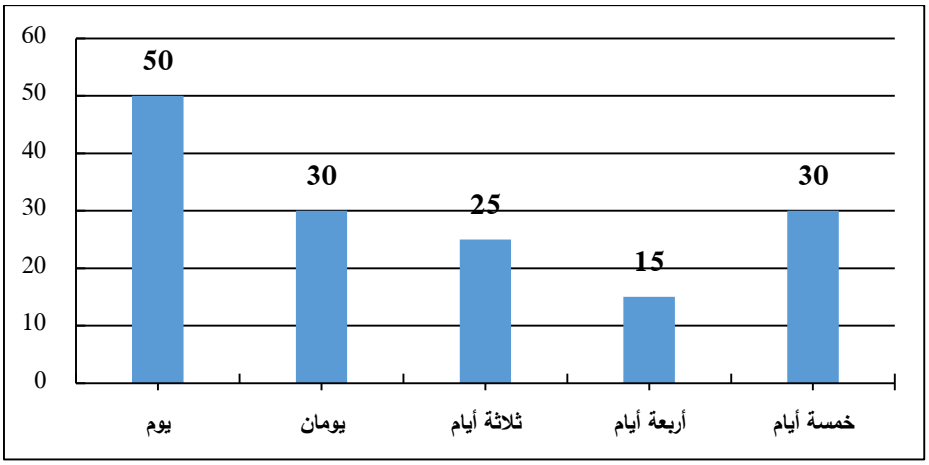
ويتم عرض بيانات المتغيرات الكمية المنفصلة بنفس طريقة عرض بيانات المتغيرات النوعية وفق أسلوبي الجداول التكرارية والرسوم البيانية.

مثال (3): فيما يأتي جدول تكراري بسيط لمتغير كمي منفصل وهو عدد أيام الغياب لمجموعة من الموظفين وعددهم (150) موظفاً.

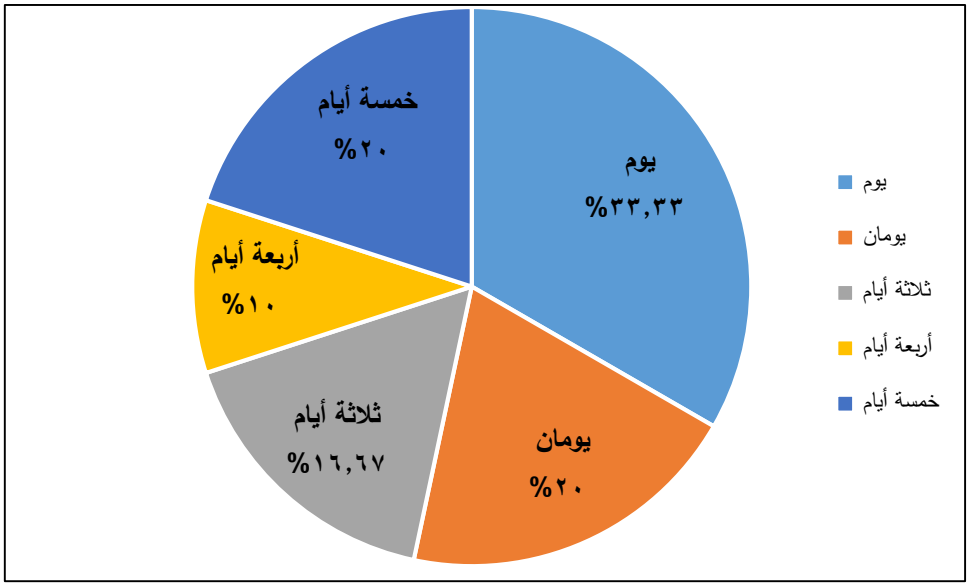
الجدول (4) توزيع عينة من الأفراد وفق متغير عدد أيام الغياب اليومي

عدد أيام الغياب	عدد الأفراد (التكرار) Frequency	النسبة % Percentage
يوم	50	%33.33
يومان	30	%20
ثلاثة أيام	25	%16.67
أربعة أيام	15	%10
خمسة أيام	30	%20
المجموع	150	%100

ويمكن تمثيل البيانات السابقة وفق طريقة الأعمدة ووفق طريق الدوائر:



الشكل (3) التمثيل البياني وفق طريقة الأعمدة البيانية لمتغير كمي منفصل (عدد أيام الغياب)



الشكل (2) التمثيل البياني وفق طريقة الدوائر لمتغير كمي منفصل (عدد أيام الغياب)

الحالة الثانية: في حال المتغيرات الكمية المتصلة:

فالمتغيرات الكمية المتصلة هي المتغيرات التي تأخذ قيماً لا نهائية من القيم المحتملة، وتأخذ أرقاماً صحيحة أو كسرية مثل: الطول، الوزن، العمر، الزمن...إلخ. ويعد تبويب هذه المتغيرات أكثر تعقيداً من المتغيرات الكمية المنفصلة، فعدد القيم التي سوف يأخذها المتغير غير محدود، ويمكن إنشاء جدول تكراري لمتغير كمي متصل بالطريقة نفسها التي يتم فيها إنشاء جدول تكراري لمتغير نوعي أو كمي منفصل عندما يكون عدد البيانات قليل، أما إذا كان عدد البيانات كبيراً فإنه يلزم تجميع المفردات ضمن فئات والتعامل معها بدلاً من التعامل معها مفردة وذلك لتسهيل العمليات الإحصائية، ومن أجل تبسيط عرض البيانات.

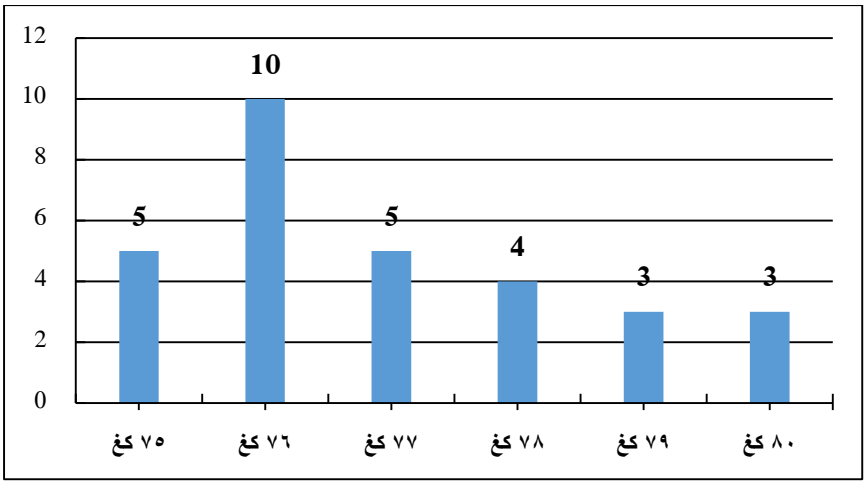
مثال (4): فيما يأتي جدول تكراري بسيط لمتغير كمي متصل وهو أوزان 30 لاعباً من (75–80) كغ.

الجدول (4) توزيع عينة من الأفراد وفق متغير الوزن

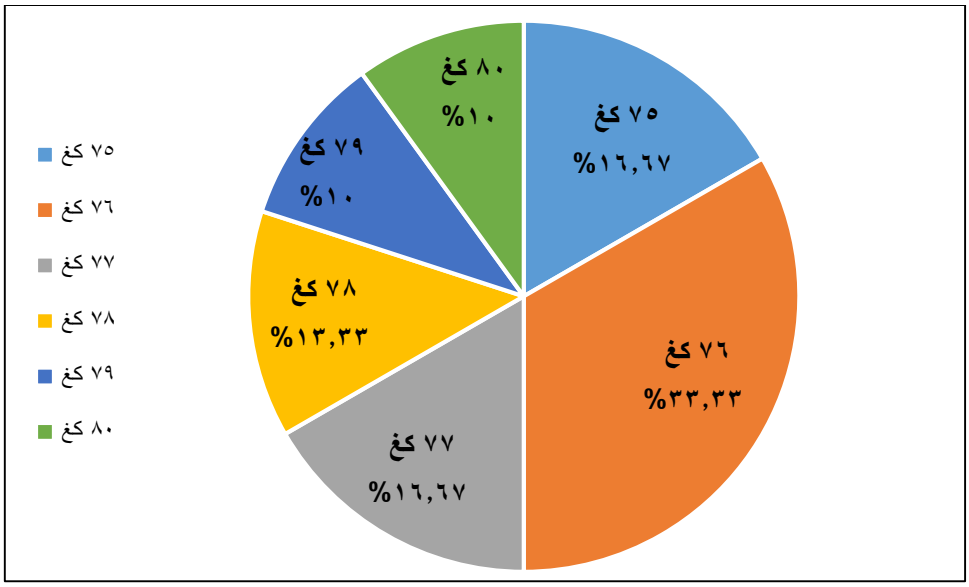
الوزن	عدد الأفراد (التكرار) Frequency	النسبة % Percentage
75 كغ	5	%16.67
76 كغ	10	%33.33
77 كغ	5	%16.67
78 كغ	4	%13.33
79 كغ	3	%10
80 كغ	3	%10
المجموع	30	%100

لاحظ أن عدد بيانات متغير (الوزن) قليلة

ويمكن تمثيل البيانات السابقة وفق طريقة الأعمدة ووفق طريق الدوائر:



الشكل (3) التمثيل البياني وفق طريقة الأعمدة البيانية لمتغير كمي متصل (الوزن)



الشكل (2) التمثيل البياني وفق طريقة الدوائر لمتغير كمي متصل (الوزن)

في المثال السابق كان عدد بيانات المتغير الكمي المتصل قليلاً من وزن 75 كغ إلى 80 كغ، ولكن في حال كان عدد بيانات المتغير الكمي المتصل كبيراً مثلاً من وزن 70 كغ إلى 100 كغ فإنه يلزم تجميع المفردات ضمن فئات والتعامل معها بدلاً من التعامل معها مفردة لتسهيل العمليات الإحصائية.

1-الجداول التكرارية للفئات:

يتكون الجدول التكراري للفئات من فئات القيم إضافة إلى عمود التكرارات.

تفريغ البيانات: ويمثل الأرقام الواقعة ضمن الفئة ونرمز لها بالإشارة (/)

التكرارات: نحول الإشارات في عمود تفريغ البيانات إلى أرقام، ويجب أن يكون حاصل التكرارات مساوياً حجم العينة وإلا كان هناك خلل.

وفيما يأتي تلخيص لخطوات إنشاء جدول تكراري للفئات:

1- نرتب البيانات الخام تصاعدياً أو تنازلياً: وذلك لسهولة تعرف أكبر مفردة وأصغر مفردة، وسهولة تفريغ البيانات.

2- حساب المدى الكلي: وهو الفرق بين أكبر قيمة وأصغر قيمة للبيانات، وذلك وفق العلاقة الآتية:

المدى = أكبر قيمة – أصغر قيمة

3- تحديد عدد الفئات المناسبة لعرض البيانات، ويفضل ألا يقل العدد عن (5) وألا يزيد عن (15) فئة.

4- تحديد طول الفئة وفق العلاقة:

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

ويرمز لطول الفئة بـ C

ويجري التقريب إلى الأعلى عندما يكون الناتج عدداً كسرياً.

5- بعد تحديد الحد الأدنى للفئة الأولى والذي يمثل أصغر قيمة من قيم المفردات، نحدد الحد الأعلى للفئة الأولى وذلك بإضافة طول الفئة إلى الحد الأدنى، والحد الأعلى للفئة الأولى هو حد أدنى للفئة التي تليها وهكذا لباقي الفئات وبهذا يمكن تعريف **طول الفئة أنه الفرق بين الحدين الأعلى والأدنى.**

6- إيجاد عدد القيم في كل فئة من فئات الجدول التكراري ويسمى هذا العدد بالتكرارات (Frequencies) ومجموع التكرارات يساوي عدد البيانات ويرمز للتكرار بـ f ومجموع التكرارات Σf. نقوم بعملية العد، وفي كل مرة نصادف مفردة نضع إشارة هكذا (/) في العمود الخاص بالإشارات ونسجل عدد الإشارات في عمود التكرار، والأرقام في عمود التكرار تظهر لنا كم مرة تكررت المفردات داخل كل فئة.

ملاحظات:

- يجب تحديد عدد الفئات المناسبة لعرض البيانات، ويفضل ألا يقل هذا العدد عن (5) فئات وألا يزيد عن (15) فئة.
- يجب أن تكون الفئات شاملة وغير متداخلة ولا توجد بينها فجوات.
- يستحسن أن يكون أطوال الفئات متساوية.
- ينبغي الابتعاد عن استخدام الفئات المفتوحة وتعني الفئات المفتوحة الفئات الحدية (الفئة الأولى أو الفئة الأخيرة) التي يكون حدها الأعلى أو الأدنى مجهولاً، مثل القول عن الفئة الأولى أقل من 18 أو أكثر من (42) للفئة الأخيرة.
- يحسب مركز كل فئة، إذ إن مركز الفئة يمثل تلك الفئة، ومركز الفئة يحسب بجمع الحد الأدنى للفئة والحد الأعلى للفئة نفسها وقسمة المجموع على 2

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

مثال (5): فيما يأتي بيانات لدرجات تلاميذ صف سادس في مادة اللغة العربية، والمطلوب إنشاء جدول توزيع تكراري ذو فئات متساوية عددها (5)

10	15	19	25	34	30	35	36	54	22
32	34	45	50	55	45	60	50	36	16

1-نرتب البيانات الخام تصاعدياً:

10	15	16	19	22	25	30	32	34	34
35	36	36	45	45	50	50	54	55	60

2-نحسب المدى الكلي = أكبر قيمة – أصغر قيمة = 60 – 10 = 50

3-تحديد عدد الفئات المناسبة لعرض البيانات وهو 5 فئات.

4-نحسب طول الفئة (C) = $\frac{\text{المدى الكلي}}{\text{عدد الفئات}} = \frac{50}{5} = 10$

5-نحدد الحد الأدنى للفئة الأولى ويساوي (10) وهو أصغر قيمة موجودة في البيانات، ومن ثم نحسب الحد الأعلى للفئة الأولى وذلك بإضافة طول الفئة (10) إلى الحد الأدنى ليصبح الحد الأعلى (20). ويبدأ الحد الأدنى للفئة الثانية بالحد الأعلى للفئة الأولى (20) وهكذا لبقية الفئات حتى نحصل على 5 فئات.

6-نحسب التكرارات، عن طريق وضع إشارة (/) لكل درجة في العمود الخاص بالإشارات ونسجل عدد الإشارات في العمود الخاص بالتكرار.

7-هذا يمكن حساب مراكز الفئات من خلال العلاقة الآتية: مركز الفئة = $\frac{\text{الحد الأدنى+الحد الأعلى}}{2}$

وبذلك يكون مركز الفئة الأولى $\frac{10+20}{2} = 15$

ويبين الجدول الآتي الجدول التكراري للفئات

الجدول (5) الجدول التكراري للفئات

الفئات	الإشارات	التكرار f	مركز الفئة xi
10-20	////	4	15
20-30	///	3	25
30-40	/ ///	6	35
40-50	////	4	45
50-60	///	3	55
		Σf=20	

الحد الأعلى للفئة الأولى

الحد الأدنى للفئة الأولى

الحد الأدنى للفئة الثانية هو نفسه الحد الأعلى للفئة الأولى

Σf : مجموع التكرارات ويساوي عدد البيانات

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

ملاحظة: عدد الفئات يفرضه الباحث على ألا يقل عن 5 فئات ولا يزيد عن 15 فئة.

التوزيع المنتظم والتوزيع غير المنتظم:

عندما تكون أطوال الفئات متساوية فإنه يدعى التوزيع التكراري في هذه الحالة توزيعاً منتظماً، كما في المثال السابق فأطوال جميع الفئات كان 10، ولكن في بعض الحالات يمكن أن تكون هناك فئات فارغة من المفردات أو قليلة المفردات فندمجها مع بعضها، فنحصل على جدول بأطوال فئات غير متساوية الطول، وبهذه الحالة يدعى التوزيع غير المنتظم. ويفضل أن تكون أطوال الفئات متساوية لتسهيل الحسابات وتسهيل التمثيل البياني للتوزيع.

الجدول المغلق والجدول المفتوح:

يدعى الجدول التكراري مغلقاً إذا علمت حدود كل فئاته، كما في الجدول السابق. أما إذا وجد في الجدول التكراري أن الفئة الأولى بلا حد أدنى فإنه يدعى بالجدول المفتوح من الأدنى، ويدعى الجدول مفتوحاً من الأعلى إذا كانت آخر فئة بدون حد أعلى، ويدعى الجدول مفتوحاً من الطرفين إذا كانت الفئة الأولى والأخيرة مفتوحتين.

يترك الباحث الفئة مفتوحة لأحد السببين:

1- المحافظة على سرية البيانات.

2- أو أن الدراسة لم تنته بعد.

ومن عيوب الجداول التكرارية المفتوحة هو أننا لا نستطيع إيجاد بعض المقاييس الإحصائية، كما لا نستطيع إيجاد التمثيل البياني للتوزيع.

الجداول التكرارية المتجمعة أو التجميعية:

يبين الجدول التكراري عدد المفردات في كل فئة، ولكن إذا رغبنا بمعرفة عدد المفردات التي تقل عن حد معين أو عدد المفردات التي تزيد عن حد معين، فإننا نحصل على هذا بإيجاد التكرار التجميعي أو التكرار المتجمع، وهناك تكرار تجميعي صاعد وتكرار تجميعي هابط.

-**فالتكرار التجميعي الصاعد:** يعرف بأنه عدد المفردات في التوزيع التي تقل عن الحد الأعلى لكل فئة، ونحصل عليه بإضافة تكرار الفئات إلى بعضها على التتالي، بحيث تبدأ بتكرار أول فئة وتنتهي بالمجموع الكلي للتكرارات.

-**أما التكرار التجميعي الهابط:** يعرف بأنه عدد المفردات في التوزيع التي قيمها أكبر أو تساوي الحد الأدنى لكل فئة، ونحصل عليه بطرح تكرار كل فئة في المجموع بدءاً من أول فئة. إذ تبدأ التكرارات المتجمعة الهابطة بالمجموع الكلي للتكرارات ثم نطرح منه التكرار الذي يقابله.

-**التكرار التجميعي الصاعد النسبي:** يعطي نسبة المفردات التي تقل قيمتها عن الحد الأعلى لكل فئة، ونحصل عليه من قسمة:

$$\frac{\text{التكرار المتجمع الصاعد}}{\text{مجموع التكرارات}} \times 100$$

-**أما التكرار التجميعي الهابط النسبي:** يعطي نسبة المفردات التي قيمتها تساوي أو أكبر من الحد الأدنى لكل فئة، ونحصل عليه من قسمة:

$$\frac{\text{التكرار المتجمع الهابط}}{\text{مجموع التكرارات}} \times 100$$

-**التكرار النسبي:** وهو يعطي نسبة المفردات التي قيمها تنتمي إلى كل فئة، ويمكن التعبير عنه بشكل عشري أو بشكل مئوي% من خلال ضرب التكرار النسبي بـ 100، ونحصل على **التكرار النسبي** بقسمة

$$\frac{\text{تكرار كل فئة}}{\text{مجموع التكرارات}} \times 100$$

ويمكن حساب التكرار التجميعي الصاعد والهابط والتكرار التجميعي النسبي الصاعد والهابط والتكرار النسبي، من خلال العودة إلى المثال السابق:

التكرار النسبي = $\frac{\text{تكرار كل فئة}}{\text{مجموع التكرارات}} \times 100$

الجدول (6) التكرار التجميعي الصاعد والتكرار التجميعي الصاعد النسبي والتكرار النسبي

الفئات	التكرار f	التكرار النسبي% f	التكرار المجتمع الصاعد $f\uparrow$	التكرار المجتمع الصاعد النسبي %$f\uparrow$
10−20	4	$\frac{4}{20} \times 100 = 20$	4	$\frac{4}{20} \times 100 = 20$
20−30	3	$\frac{3}{20} \times 100 = 15$	4 + 3 = 7	$\frac{7}{20} \times 100 = 35$
30−40	6	$\frac{6}{20} \times 100 = 30$	7 + 6 = 13	$\frac{13}{20} \times 100 = 65$
40−50	4	$\frac{4}{20} \times 100 = 20$	13 + 4 = 17	$\frac{17}{20} \times 100 = 85$
50−60	3	$\frac{3}{20} \times 100 = 15$	17 + 3 = 20	$\frac{20}{20} \times 100 = 100$
	$\Sigma f = 20$			

التكرار المتجمع الصاعد النسبي = $\frac{\text{التكرار المتجمع الصاعد}}{\text{مجموع التكرارات}} \times 100$

إنّ التكرارات المتجمعة الصاعدة تبدأ بتكرار أول فئة وتنتهي بالمجموع الكلى للتكرارات

الجدول (7) التكرار التجميعي الهابط والتكرار التجميعي الهابط النسبي

الفئات	التكرار f	التكرار المتجمع الهابط $f\downarrow$	التكرار المتجمع الهابط النسبي % $f\downarrow$
10−20	4	20	$\frac{20}{20} \times 100 = 100$
20−30	3	20 − 4 = 16	$\frac{16}{20} \times 100 = 80$
30−40	6	16 − 3 = 13	$\frac{13}{20} \times 100 = 65$
40−50	4	13 − 6 = 7	$\frac{7}{20} \times 100 = 35$
50−60	3	7 − 4 = 3	$\frac{3}{20} \times 100 = 15$
	$\Sigma f = 20$		

التكرار المتجمع الهابط النسبي = $\frac{\text{التكرار المتجمع الهابط}}{\text{مجموع التكرارات}} \times 100$

إنّ التكرارات المتجمعة الهابطة تبدأ بالمجموع الكلي للتكرارات ثم نطرح منه التكرار الذى يقابله

2-التمثيل البياني للفئات (المتغيرات الكمية المتصلة):

أ-المدرج التكراري Frequency Histogram:

المدرج التكراري هو عرض بيانات الجدول التكراري على المحاور الإحداثية، ويمكن إتباع الخطوات الآتية في رسم المدرج التكراري:

1- نرسم محورين متعامدين بحيث تمثل حدود كل فئة على المحور الأفقي، في حين تمثل التكرارات على المحور العمودي.

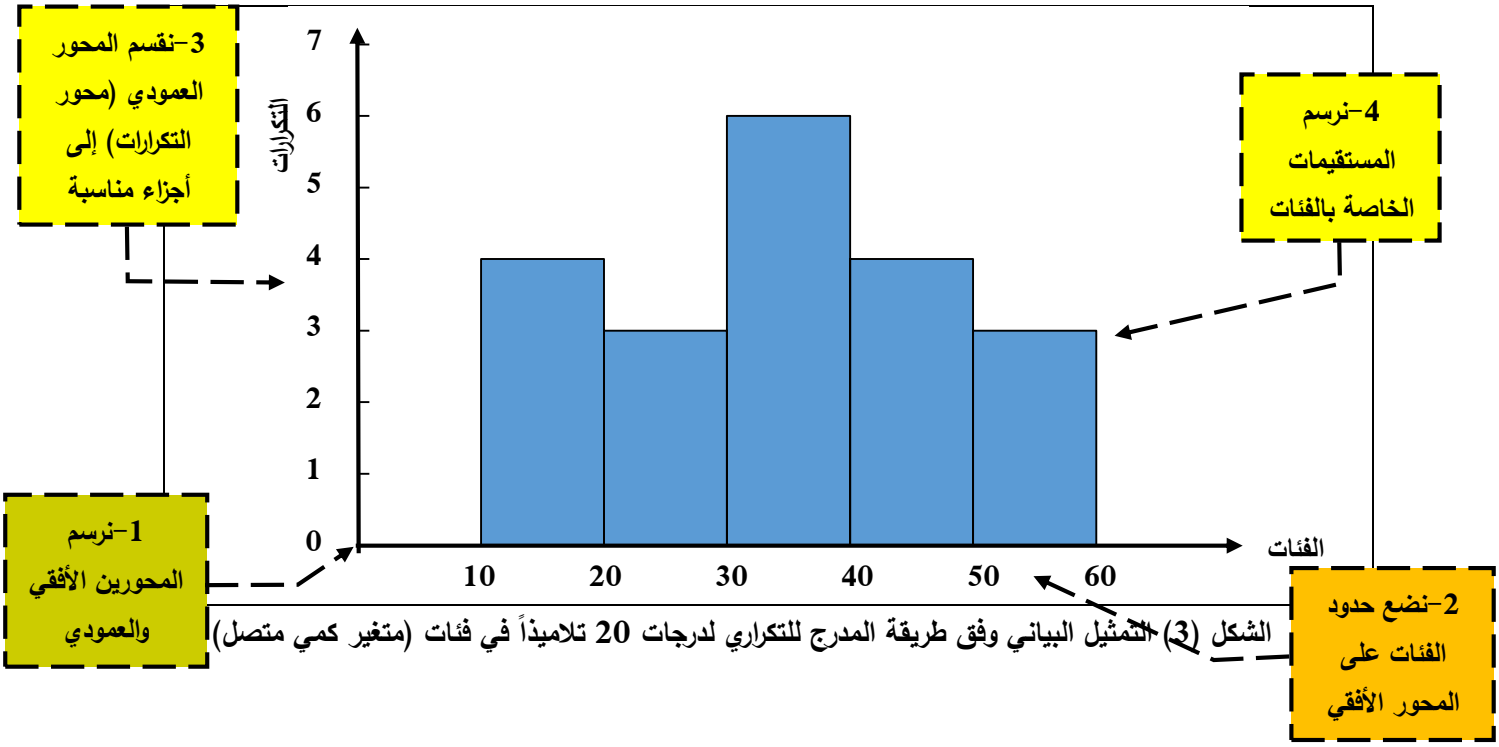
2- نضع حدود الفئات على المحور الأفقي بحيث تكون الفئات متلاصقة ولا مسافات بينها والمسافات متساوية داخل الفئات، ولا حاجة لتكرار الحد الأعلى والأدنى الحقيقي لكل فئة بل يكتب مرة واحدة.

3- نقسم المحور العمودي (محور التكرارات) إلى أجزاء مناسبة باستخدام مقياس رسم مناسب، فإذا كان أعلى تكرار هو (20) فإنه يمكن أن نقسم المحور إلى 5، 10، 15، 20 أو 1، 2، 3... إلى 20.

4- نرسم المستقيم الخاص بالفئة التي لها أكبر تكرار للمفردات بحيث يكون عرض المستطيل مساوياً لطول الفئة وارتفاعه متناسب مع تكرار الفئة، ثم نرسم باقي المستطيلات للفئات الأخرى.

وهكذا نحصل على مجموعة مستطيلات متلاصقة تشبه الدرج.

مثال على رسم المدرج التكراري: لو أخذنا البيانات السابقة من الجدول (5) والتي تمثل فئات لبيانات درجات تلاميذ صف سادس في مادة اللغة العربية، فإننا أولاً نبدأ برسم المحورين الأفقي والعمودي، ونضع على المحور الأفقي حدود الفئات وعلى المحور العمودي التكرارات، ثم نرسم المستطيلات التي عرضها يمثل طول الفئات وارتفاعها يمثل تكرار الفئات وفق الشكل الآتي:

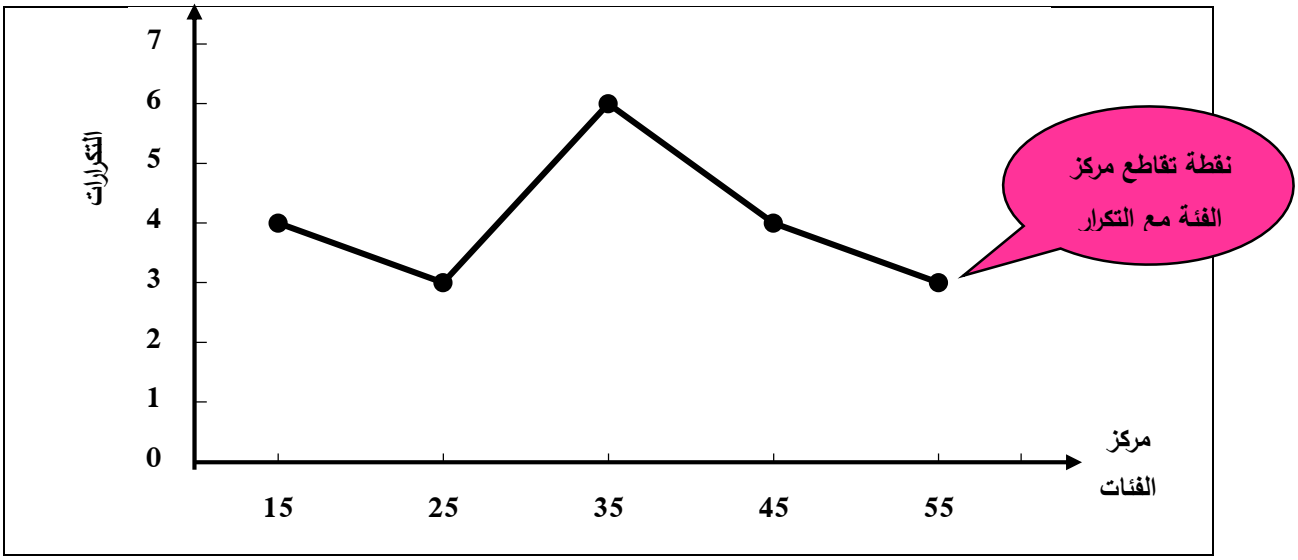


الشكل (3) التمثيل البياني وفق طريقة المدرج للتكراري لدرجات 20 تلاميذاً في فئات (متغير كمي متصل)

ب-المضلع التكراري Frequency Polygon:

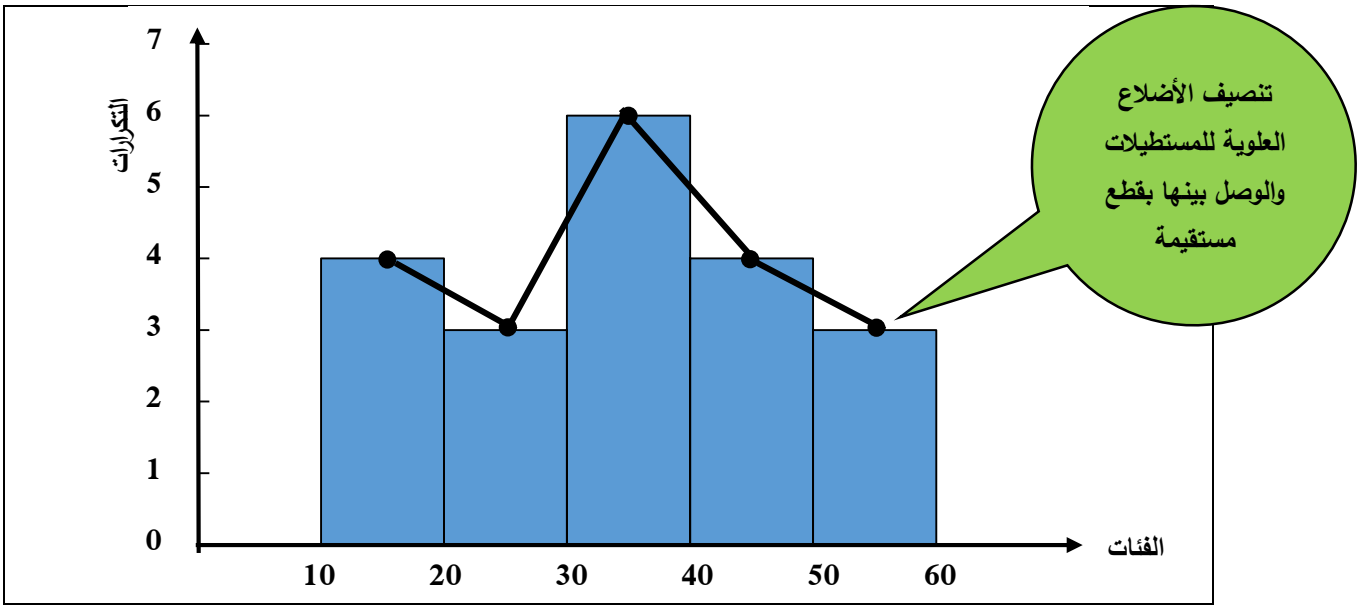
لرسم المضلع التكراري نستخدم مراكز الفئات بدلاً من حدود الفئات على المحور الأفقي، ويمكن أن نحصل على المضلع التكراري أيضاً بتنصيف الأضلاع العلوية للمستطيلات في المدرج التكراري ثم وصل هذه النقاط بعضها مع بعض بقطع مستقيمة، ولإغلاق الفئات نفترض وجود فئتين على الأطراف واحدة إلى أقصى اليسار وأخرى إلى أقصى اليمين بحيث يكون تكرار كل فئة مساوياً الصفر.

-الحالة الأولى: من خلال تمثيل مراكز الفئات بدلاً من حدود الفئات على المحور الأفقي، ووصل نقاطع تقاطع كل مركز فئة مع التكرار المقابل لها، وفق الشكل الآتي وذلك اعتماداً على المثال السابق:



الشكل (4) التمثيل البياني وفق طريقة المضلع للتكراري لدرجات 20 تلاميذاً مبوبة في فئات (متغير كمي متصل)

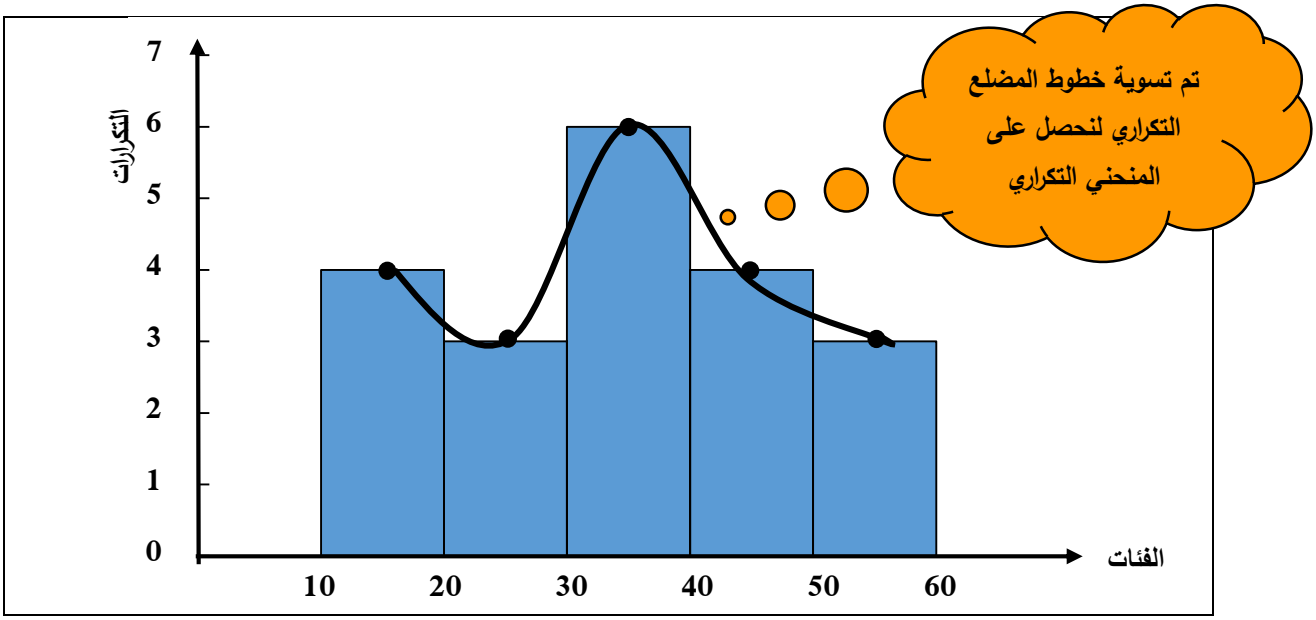
-الحالة الثانية: نحصل على المضلع التكراري أيضاً بتنصيف الأضلاع العلوية للمستطيلات في المدرج التكراري ثم وصل هذه النقاط بعضها مع بعض بقطع مستقيمة، وفق الشكل الآتي:



الشكل (5) التمثيل البياني وفق طريقة المضلع للتكراري لدرجات 20 تلاميذاً مبوبة في فئات (متغير كمي متصل)

ج-المنحنى التكراري Frequency Curve:

نحصل على المنحنى التكراري بتسوية خطوط المضلع التكراري، ويفيد المنحنى التكراري في الدراسات الاستدلالية أكثر من الوصفية، وذلك للاستدلال على شكل التوزيع في المجتمع الذي أخذت منه عينة الدراسة.



الشكل (6) التمثيل البياني وفق طريقة المنحني للتكراري لدرجات 20 تلاميذاً مبوبة في فئات (متغير كمي متصل)

المحاضرتين الخامسة والسادسة: مقاييس النزعة المركزية
Measures of central tendency

مقدمة:

بعد جمع البيانات وتنظيمها في جداول تكرارية وعرضها بيانياً، سنتناول في هذا الفصل موضوع أساليب التحليل الإحصائي وإيجاد المقياس المناسب الذي يمثل الظاهرة موضع الدراسة، واستخدامه للمقارنة بين تلك الظاهرة وظواهر أخرى.

فإذا نظرنا إلى مفردات أي ظاهرة مدروسة نجد أن أكثر هذه المفردات تتجمع حول قيمة ما تمثل متوسط الظاهرة، ونجد أن هناك تناقصاً تدريجياً في عدد المفردات كلما بعدت هذه القيم عن المتوسط ومن الجانبين، وهذا التجمع حول القيمة المتوسطة يسمى بـ **(النزعة المركزية) أي ميل المفردات إلى الاقتراب والتجمع حول قيمة معينة هي القيمة المتوسطة، وكأن القيمة المتوسطة تحاول شد المفردات لكي تتراكم حولها،** ولكن من الطبيعي ألا تنجح جميع المفردات في الاقتراب من القيمة المتوسطة، فهناك مفردات تبتعد بشكل بسيط عن المتوسط وهناك مفردات عددها أقل تبتعد بشكل أكبر عن المتوسط، فمثلاً الأقزام والعمالقة لا نشاهدهم إلا نادراً في حياتنا اليومية ومعظم الناس هم متوسطو الطول، وعدد قليل منهم يختلف طولهم عن متوسط الطول بالزيادة أو النقصان.

وهناك أنوع عدة من مقاييس النزعة المركزية أهمها: المتوسط الحسابي (الوسط الحسابي)، الوسيط، المنوال، الوسط الهندسي، الوسط التوافقي، الوسط التربيعي وهو نادر الاستخدام.

وأشهر مقاييس النزعة المركزية هي: المتوسط والوسيط والمنوال، ويجب أن تتوافر في هذه المقاييس بعضاً من الصفات الآتية:

- أن يعتمد المقياس في حسابه على كل المشاهدات
- أن يكون المقياس سهل الحساب والفهم
- أن يتوفر فيه القابلية للتعامل الجبري
- ألا يتأثر بالقيم الشاذة أو المتطرفة

ولا تتوافر كل الصفات السابق ذكرها في مقياس واحد ولكن كل مقياس من هذه المقاييس يفضل استخدامه في حالات معينة ولا يفضل استخدامه في حالات أخرى، والجدول الآتي يبين توفر الصفات السابقة في كل مقياس من مقاييس النزعة المركزية:

الجدول (1) الصفات المتوفرة في كل مقياس من مقاييس النزعة المركزية

المقياس / الخاصية	المتوسط	الوسيط	المنوال
يعتمد في حسابه على كل المشاهدات	✓	×	×
المقياس سهل الفهم والحساب	✓	✓	✓
تتوفر فيه القابلية للتعامل الجبري	✓	✓	✓
لا يتأثر بالقيم الشاذة أو المتطرفة	×	✓	✓

وفيما يأتي وصف وتوضيح لكل مقياس من هذه المقاييس:

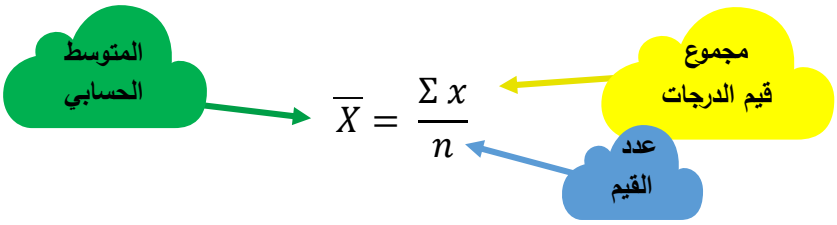
أولاً-المتوسط الحسابي (Mean):

هو أكثر مقاييس النزعة المركزية استخداماً ويساوي مجموع القيم على عددها ويرمز له بـ $\overline{X}$ (اكس بار) للعينة و M للمجتمع، وتعتمد طريقة حساب المتوسط الحسابي على نوعية البيانات إن كانت مبوبة في فئات أو غير مبوبة في فئات.

-حساب المتوسط الحسابي:

1-حساب المتوسط الحسابي من الدرجات مباشرة للبيانات غير المبوبة:

يحسب المتوسط الحسابي من الدرجات مباشرة وذلك بقسمة مجموع هذه الدرجات على عددها. ولا يشترط أن يكون المتوسط الحسابي عدداً صحيحاً، وذلك وفق المعادلة الآتية:



مثال 1: احسب المتوسط الحسابي لعلامات أحد الطلاب في خمس مواد دراسية وهذه العلامات هي: 70، 75، 88، 81، 86.

$$\overline{X} = \frac{\Sigma X}{n} = \frac{70 + 75 + 88 + 81 + 86}{5} = \frac{400}{5} = 80$$

أي متوسط علامات الطالب في المواد الخمسة هو 80.

2-حساب المتوسط من تكرار الدرجات غير المبوبة في فئات:

ويُحسب وفق القانون الآتي:

$$\overline{X} = \frac{\Sigma fx}{\Sigma f}$$

حيث Σf: مجموع التكرارات ويساوي عدد الأفراد

Σfx : مجموع حاصل ضرب الدرجة بالتكرار المقابل لها.

لحساب المتوسط في هذه الحالة لابد من حساب عدد الأفراد Σf، كذلك لابد من ضرب كل درجة بتكرارها، ثم جمع حاصل ضرب الدرجة في التكرار.

مثال 2: احسب المتوسط الحسابي لدرجات 30 طالباً في مادة الإحصاء وفق الآتي:

الجدول (2) درجات مجموعة من الطلبة في مادة الإحصاء

الدرجات X	التكرار f	الدرجة × التكرار fX
90	4	360
88	5	440
82	3	246
80	4	320
77	6	462
75	5	375
65	3	195
	$\Sigma f = 30$	Σfx =2398

1 2 3

4 $\overline{X} = \frac{\Sigma fx}{\Sigma f} = \frac{2398}{30} = 79.93$

3-حساب المتوسط من فئات القيم:

لحساب المتوسط الحسابي بهذه الطريقة يتم إتباع الخطوات الآتية:

أ- إيجاد مراكز الفئات x_i.

ب-ضرب مركز كل فئة بالتكرار المناظر.

ت-إيجاد مجموع حاصل ضرب مراكز الفئات بالتكرارات المناظرة.

ث-التعويض بالقانون الآتي:

$$\overline{X} = \frac{\Sigma(x_i f)}{\Sigma f}$$

حيث $\overline{X}$: المتوسط الحسابي

$\Sigma(xif)$: مجموع حاصل ضرب مركز الفئة بتكرارها

Σf : مجموع التكرارات (حجم العينة)

مثال 3: فيما يأتي بيانات لدرجات (40) طالباً في مادة اللغة العربية والمطلوب حساب المتوسط الحسابي لدرجات الطلاب:

1-إيجاد مركز الفئة

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

2-ضرب مركز كل فئة بالتكرار المقابل

الجدول (3) التوزيع التكراري لفئات درجات الطلاب في مادة اللغة العربية

الفئات	التكرار f	مركز الفئة X_i	مركز الفئة ×التكرار fX_i
10-20	1	15	15
20-30	3	25	75
30-40	4	35	140
40-50	3	45	135
50-60	11	55	605
60-70	7	65	455
70-80	5	75	375
80-90	4	85	340
90-100	2	95	190
المجموع	$\Sigma f = 40$		$\sum x_i f = 2330$

3-مجموع حاصل ضرب مراكز الفئات بالتكرارات المقابلة

4-نعوض بالقانون الآتي:

$$\overline{X} = \frac{\Sigma(x_i f)}{\Sigma f} = \frac{2330}{40} = 58.25$$

وتسمى الطريقة السابقة في حساب المتوسط الحسابي بالطريقة أو الأسلوب المباشر، ومن طرائق حساب المتوسط الحسابي أيضاً أسلوب الانحرافات، وأسلوب الانحرافات المختصرة.

-خواص المتوسط الحسابي:

1- يستخدم المتوسط للبيانات الكمية (فئوية أو نسبية) كالذكاء والتحصيل والوزن والطول ولا يستخدم للبيانات الوصفية (اسمية، ترتيبية) كالجنس ومكان الإقامة والمؤهل العلمي.

2- قابل للعمليات الجبرية، ويدخل في حساب مقاييس أخرى كثيرة.

3- تدخل جميع القيم في حسابه.

4- يعطى قيمة وحيدة لمجموعة القيم.

5- مجموع انحرافات القيم عن متوسطها الحسابي يساوي صفراً.

وللتحقق من هذه الخاصية نأخذ المثال الآتي:

مثال 4: ليكن لدينا البيانات الآتية: (2، 4، 4، 9، 1)

خطوات الحل:

أ-نحسب المتوسط وفق القانون الآتي:

$$\overline{X} = \frac{\Sigma X}{n} = \frac{2+4+4+9+1}{5} = \frac{20}{5} = 4$$

ب-نحسب انحراف كل درجة عن متوسطها ثم نحسب مجموع الانحراف كما يأتي:

الجدول (4) انحراف كل درجة عن المتوسط

X	d
2	2 − 4 = −2
4	4 −4 = 0
4	4 −4 = 0
9	9 − 4 = 5
1	1 − 4 = −3
$\overline{X} = 4$	5 − 5 = 0

d: انحراف الدرجة عن المتوسط

لاحظ أن: مجموع انحرافات الدرجات عن المتوسط يساوي صفراً

6- يتأثر المتوسط بالقيم المتطرفة أو الشاذة، ويقال بأن مفردة ما متطرفة إذا كانت كبيرة جداً أو صغيرة جداً مقارنة ببقية المفردات.

مثال 5: لتكن لدينا القيم الآتية والتي تمثل درجات خمس طلاب في مادة التاريخ:

16	14	11	10	9

$$\overline{X} = \frac{\Sigma X}{n} = \frac{60}{5} = 12$$

ولتكن لدينا القيم الآتية والتي تمثل درجات خمس طلاب في مادة الجغرافية:

40	14	11	10	9

قيمة متطرفة

$$X = \frac{\Sigma X}{n} = \frac{84}{5} = 16.8$$

ولتكن لدينا القيم الآتية والتي تمثل درجات خمس طلاب في مادة التربية الوطنية:

16	14	11	10	1

قيمة متطرفة

$$X = \frac{\Sigma X}{n} = \frac{52}{5} = 10.4$$

يلاحظ أن المتوسط في مادة الجغرافية قد ازداد زيادة واضحة عن متوسط درجات الطلاب في مادة التاريخ نظراً لوجود قيمة متطرفة في الجغرافية وهي الدرجة 40، كما يلاحظ أن المتوسط في مادة التربية الوطنية قد نقص عن متوسط درجات الطلاب في مادة التاريخ نظراً لوجود قيمة متطرفة في التربية الوطنية وهي الدرجة 1 في حين كانت باقي الدرجات متطابقة، لهذا في حال وجود قيم متطرفة يفضل استخدام الوسيط.

7- لا يمكن حساب المتوسط في حال الجداول ذات الفئات الحدية المفتوحة، والفئة المفتوحة هي الفئة التي لا يعرف أحد حديها (الأعلى أو الأدنى).

ثانياً-الوسيط Median

وهو المقياس الثاني من مقاييس النزعة المركزية، **ويعرّف الوسيط بأنه: تلك القيمة التي تتوسط توزيع مجموعة قيم مرتبة تصاعدياً بحيث يسبقها نصف القيم ويليها النصف الآخر، أو القيمة التي يقع تحتها أو يقل عنها (50%) من القيم، وذلك بعد أن يتم ترتيبها تصاعدياً أو تنازلياً**. يعتمد حساب الوسيط على **عدد الدرجات** سواء كان فردياً أم زوجياً. ويرمز له اختصاراً **(Med)**.

-حساب الوسيط:

1-حساب الوسيط للبيانات غير المبوبة:

أ-حساب الوسيط عندما يكون عدد القيم فردياً:

مثال 6: لدينا القيم الآتية: 9، 8، 12، 7، 13 والمطلوب أوجد الوسيط.

خطوات الحل:

1-ترتب القيم ترتيباً تصاعدياً كما يأتي: 7، 8، 9، 12، 13

2-نبحث بعد ذلك عن النقطة التي تنصف هذه الدرجات فنرى أنها تقع تماماً عند الدرجة 9 لأن عدد الدرجات التي تسبقها 2 وهي 7،8 وعدد الدرجات التي تليها 2 وهي 12،13.

3-يمكن أن نحدد الوسيط أيضاً من خلال حساب رتبة الوسيط أو ترتيب الوسيط وفق القانون الآتي:

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

ويرمز لعدد الدرجات بـ n

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

والدرجة الثالثة هي **9**.

ونلخص الخطوات السابقة كما يأتي:

- نرتب الدرجات ترتيباً تصاعدياً.
- عدد الدرجات 5
- ترتيب الوسيط = $\frac{5+1}{2}$ = 3
- إذن الدرجة الوسطى هي 9

ب-حساب الوسيط عندما يكون عدد القيم زوجياً:

مثال 7: لدينا القيم الآتية: 9، 8، 12، 7، 13، 5 والمطلوب أوجد الوسيط

خطوات الحل:

1-ترتب القيم ترتيباً تصاعدياً كما يأتي: 5، 7، 8، 9، 12، 13

2-نحدد الوسيط أيضاً من خلال حساب رتبة الوسيط وذلك وفق القانون الآتي:

$$\textbf{رتبة الوسيط} = \frac{n}{2}$$

ويرمز لعدد الدرجات بـ n

$$\text{رتبة الوسيط} = \frac{6}{2} = 3$$

الدرجة التي ترتيبها الثالث من الطرف الأول اليمين لتدريج الدرجات هي 8.

الدرجة التي ترتيبها الثالث من الطرف الثاني اليسار لتدريج الدرجات هي 9.

فيكون الوسيط هو متوسط العددين 8 و9.

$$\text{الوسيط (med)} = \frac{8+9}{2} = 8.5$$

2-حساب الوسيط من فئات القيم (الفئات المغلقة والفئات المفتوحة):

ويتم حساب الوسيط وفق الخطوات الآتية:

1- نحسب التكرار المتجمع الصاعد.

2- نحسب ترتيب الوسيط وهو يساوي $\frac{\text{مجموع التكرارات}}{2}$ أي $\frac{\Sigma f}{2}$

3- نبحث في حقل **التكرار التجميعي الصاعد** عن قيمة ترتيب الوسيط فإذا لم نجدها نأخذ الأكبر منها مباشرة.

4- نحسب الوسيط وفق المعادلة الآتية:

$$\textbf{الوسيط=الحد الأدنى لفئة الوسيط+}\left(\frac{\text{رتبة الوسيط-التكرار التجميعي لفئة ما قبل الوسيط}}{\text{تكرار فئة الوسيط}}\right)\times \textbf{طول فئة الوسيط}$$

أي $$Med=Lme+\left(\frac{\frac{\Sigma F}{2}-F\uparrow}{Fme}\right) x\ Cme$$

حيث Lme الحد الأدنى لفئة الوسيط

Σf مجموع التكرارات (عدد الأفراد)

$f\uparrow$ التكرار التجميعي الصاعد لفئة ما قبل الوسيط

$f\ me$: تكرار فئة الوسيط

Cme: طول فئة الوسيط

ملاحظة: نظراً لأن الوسيط لا يتأثر بالفئات المفتوحة فإنه يمكن أن نحسب الوسيط من الفئات المفتوحة.

مثال 8: لتكن لدينا بيانات لدرجات 20 تلميذاً في مادة الرياضيات، والمطلوب أوجد الوسيط.

الجدول (5) حساب الوسيط لدرجات 20 تلميذاً في مادة الرياضيات

فئات القيم	f	$f\uparrow$
6−10	2	2
10−14	3	5
14−18	4	9
18−22	5	14
22−26	3	17
26−30	3	20
	Σf =20	

فئة الوسيط

الحل:

لحساب الوسيط من الفئات نتبع الخطوات الآتية:

1- نحسب التكرار المتجمع الصاعد f↑.

2- حسب ترتيب الوسيط وهو يساوي $\frac{\text{مجموع التكرارات}}{2}$ أي $\frac{\Sigma f}{2} = \frac{20}{2} = 10$

3- نبحث في حقل التكرار التجميعي الصاعد عن قيمة ترتيب الوسيط (10) فإذا لم نجدها نأخذ الأكبر منها مباشرة وهي هنا (14) فنجد أن فئة الوسيط هي الفئة (18−22) المقابلة للتكرار التجميعي الصاعد 14 لأن التكرار التجميعي للفئة التي تسبقه 9 وهو أقل من 10.

4- نحسب الوسيط وفق المعادلة الآتية

$$\text{الوسيط}=\text{الحد الأدنى لفئة الوسيط}+\left(\frac{\text{رتبة الوسيط}-\text{التكرار التجميعي لفئة ما قبل الوسيط}}{\text{تكرار فئة الوسيط}}\right)\times \text{طول فئة الوسيط}$$

أي أن $Med = Lme + \left(\frac{\frac{\Sigma F}{2} - F\uparrow}{Fme}\right) XCme$

حيث Lme الحد الأدنى لفئة الوسيط = 18

Σf مجموع التكرارات (عدد الأفراد) = 20

$f\uparrow$ التكرار التجميعي الصاعد لفئة ما قبل الوسيط = 9

$f\ me$: تكرار فئة الوسيط = 5

C : طول فئة الوسيط = 4

وبتطبيق هذه المعادلة نحصل على

$$Med = 18 + \frac{10-9}{5} \times 4$$

$$Med = 18 + \frac{1}{5} \times 4$$

$$Med = 18 + \frac{4}{5}$$

$$Med = 18 + 0.80$$

$$\mathbf{Med = 18.80}$$

خواص الوسيط:

1- يفضل استخدام الوسيط في حالة المتغيرات الترتيبية (مثل الصف الدراسي، المؤهل العلمي) لذلك يناسب المستوى الترتبي من مستويات القياس.

2- لا يتأثر الوسيط بالقيم المتطرفة أو الشاذة لذلك يفضل استخدامه في وصف المتغيرات الكمية التي توجد فيها مثل هذه القيم، كما هي الحال مع بيانات الدخل.

3- يستخدم في حال وجود الفئات المفتوحة.

4- يتحدد بعدد البيانات وليس بقيمتها.

5- يستخدم في حال الفئات غير المتساوية في أطوالها.

6- يعتمد الوسيط في حسابه على قيمة واحدة أو قيمتين، وعلى فئة واحدة في حال الفئات، وبالتالي لا يأخذ جميع القيم في حسابه.

ثالثاً-المنوال Mode

يعرّف المنوال أنّه: القيمة الأكثر تكراراً أو الأكثر شيوعاً بين القيم، أي القيمة التي تتكرر أكثر من غيرها، ويرمز له بـ Mod.

يستخدم المنوال في جميع مستويات القياس (الاسمي الترتيبي، الفئوي، النسبي)، ويستخدم كذلك في الفئات المغلقة والفئات المفتوحة.

-حساب المنوال:

1-حساب المنوال للبيانات غير المبوبة:

مثال 9: أوجد المنوال للقيم الآتية التي تمثل أعمار مجموعة من الموظفين:

27 – 26 – 26 – 29 – 29 – 29 – 28 – 28 – 31 – 30

يلاحظ أن المنوال (العمر الأكثر تكراراً) هو 29، بمعنى أن غالبية الموظفين هم من عمر 29

في حال كانت البيانات على الشكل الآتي:

27 – 26 – 26 – 26 – 29 – 29 – 29 – 28 – 28 – 31 – 30

يلاحظ وجود منوالين (29، 26).

ملاحظة: يمكن أن يوجد أكثر من منوال في مجموعة القيم، إذا تكرر بعض الفئات بالدرجة نفسها. فيسمى عندها **بالتوزيع متعدد المنوال.**

2-حساب المنوال من فئات القيم (الفئات المغلقة والفئات المفتوحة):

هناك أساليب عدة لحساب المنوال وسوف نقتصر هنا على أسلوب الفروق.

يتم حساب المنوال وفق الخطوات الآتية:

1-نحدد فئة المنوال وهي الفئة المقابلة لأعلى تكرار.

2-نحسب المنوال وفق المعادلة الآتية:

$$Mod=Lmo+\frac{\Delta 1}{\Delta 1+\Delta 2}\times C_{m0}$$

حيث Mod: المنوال

Lmo : الحد الأدنى لفئة المنوال

$\Delta 1$: تكرار فئة المنوال – تكرار الفئة التي قبلها

$\Delta 2$: تكرار فئة المنوال – تكرار الفئة التي تليها

C_{m0}: طول فئة المنوال

بالعودة إلى المثال السابق (8):

الجدول (6) حساب المنوال لدرجات 20 تلميذاً في مادة الرياضيات

فئات القيم	f
6−10	2
10−14	3
14−18	4
18−22	5
22−26	3
26−30	3
	$\Sigma f = 20$

فئة المنوال (ذات التكرار الأعلى)

الحل:

1−نحدد فئة المنوال وهي في المثال السابق الفئة (18−22) المقابلة لأعلى تكرار وهو 5

2−نحسب المنوال وفق المعادلة الآتية:

$$Mo = Lmo + \frac{\Delta 1}{\Delta 1 + \Delta 2} \times C_{m0}$$

حيث Mo : المنوال

Lmo : الحد الأدنى لفئة المنوال وهو 18

$\Delta 1$: تكرار فئة المنوال – تكرار الفئة التي قبلها أي $5 - 4 = 1$

$\Delta 2$: تكرار فئة المنوال – تكرار الفئة التي تليها أي $5 - 3 = 2$

C_{m0} طول فئة المنوال = 4

نعوض بالقيم:

$$Mod = 18 + \frac{1}{1+2} \times 4$$

$$= 18 + \frac{4}{3}$$

$$= 18 + 1.33$$

$$\mathbf{Mod = 19.33}$$

<u>−خواص المنوال:</u>

1− يستخدم المنوال في جميع أنواع المتغيرات سواء كانت كمية أو نوعية، وهو المقياس الوحيد المستخدم في وصف البيانات التي يمكن قياسها على المستوى الاسمي.

2− يفضل استخدام المنوال في حال وجود فئات مفتوحة.

3- لا يتأثر المنوال بالقيم المتطرفة أو الشاذة.

4- يمتاز بسهولة حسابه.

5- قد يوجد أكثر من منوال في مجموعة القيم، إذا تكرر بعض الفئات بالدرجة نفسها. فيسمى عندها بالتوزيع متعدد المنوال.

مقاييس النزعة المركزية ومستويات القياس:

إن تحديد مستويات القياس للبحث يعد أمراً هاماً للباحث الذي يقوم بتنفيذه ليتسنى له اختيار الأسلوب الإحصائي المناسب لوصف المتغير من جهة، وللإجابة عن أسئلة بحثه من جهة أخرى.

ويبين الجدول الآتي مدى ملائمة كل من مقاييس النزعة المركزية لكل مستوى من مستويات القياس.

الجدول (8) مقاييس النزعة المركزية ومستويات القياس

	الاسمي	الترتيبي	الفئوي	النسبي
المتوسط			******	******
الوسيط		******	******	******
المنوال	******	******	******	******

وتجدر الملاحظة أن المنوال هو المقياس الوحيد المستخدم في جميع مستويات القياس، في حين أن المستويين الفئوي والنسبي يمكن وصفهما بجميع مقاييس النزعة المركزية.

فوائد مقاييس النزعة المركزية:

تتلخص أهم الفوائد التطبيقية لمقاييس النزعة المركزية في:

1- المعايير: تعتمد المعايير على المتوسط، ولهذا ينسب طول الفرد مع متوسط طول من هم في عمره، كذلك يقاس ذكاء الفرد بالنسبة لمتوسط ذكاء أقرانه وبمدى انحرافه عن هذا المعيار زيادة أو نقصاناً.

2- المقارنة: تستخدم المتوسطات لمقارنة مجموعة من الأفراد بمجموعة أخرى، من مثل مقارنة متوسط درجات الطلاب في مادة ما بمتوسط درجات نظرائهم في صف آخر. على أن تتم المقارنة للمجموعات المتجانسة من حيث بعض الصفات من مثل: المهنة، العمر، الصف الدراسي، حتى تصبح نتائجها صحيحة.

3- هذا ويصلح الوسيط والمنوال للأغراض نفسها التي يصلح لها المتوسط ويجري استخدام الوسيط بدلاً عن المتوسط في حال البيانات الكمية الترتيبية، أما المنوال فيستخدم بدلاً من المتوسط في البيانات الاسمية.

وفيما يأتي جدول يبين حالات استخدام كل من المتوسط والوسيط والمنوال:

الجدول (7) حالات استخدام كل من المتوسط والوسيط والمنوال

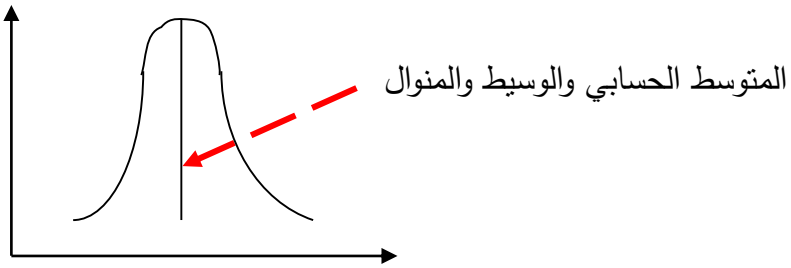
المتوسط	الوسيط	المنوال
-يستخدم إذا كان التوزيع متماثلاً تقريباً. -إذا كان الاهتمام منصباً على القيمة العددية لجميع البيانات أي المجموع الكلي.	- يستخدم إذا كان التوزيع ملتوياً. - يستخدم في حال الرغبة في إيجاد قيمة ممثلة، أي قيمة نموذجية بدلاً من الاهتمام بالمجموع الكلي. - في حال فقدان بعض القيم وعرف ترتيبها، بحيث لا يمكن حساب المتوسط مباشرة.	- يستخدم إذا كان التوزيع متعدد المنوالات

العلاقة بين المتوسط الحسابي والوسيط والمنوال (علاقة بيرسون):

1- إذا كان التوزيع التكراري متماثلاً وله قيمة واحدة وشكله يشبه شكل الجرس (التوزيع الطبيعي أو الاعتدالي) وتتساوى فيه قيم المتوسط والوسيط والمنوال، أي:

المتوسط = الوسيط = المنوال

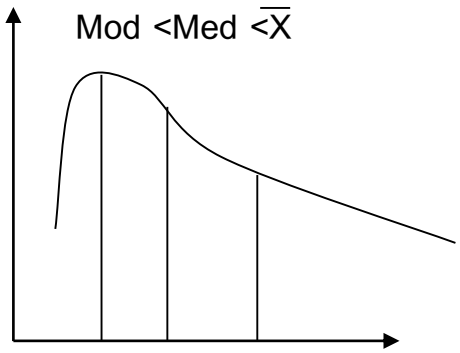
$$\overline{X} = Med = Mod$$



الشكل (1) توزيع تكراري متماثل (المتوسط الحسابي = الوسيط = المنوال)

2- ولكن عندما يختل التماثل يبقى المنوال واقعاً تحت قمة التوزيع كونه يمثل التكرار الأعلى، ويبتعد المتوسط عن المنوال باتجاه الالتواء ويقع الوسيط بينهما، وقد يكون المنحني ملتوي نحو اليمين (التواء موجب)، ويكون المتوسط أكبر من الوسيط والمنوال، كما يتضح من الشكل الآتي.

المتوسط > الوسيط> المنوال



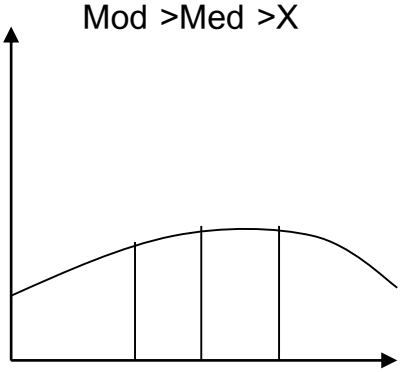
المتوسط > الوسيط> المنوال

الشكل (2) توزيع ملتوٍ ناحية اليمين

3- وقد يكون المنحني ملتوي نحو اليسار (التواء سالب)، ويكون المتوسط أصغر من الوسيط والمنوال، كما يتضح من الشكل الآتي:

المنوال > الوسيط> المتوسط

$Mod > Med > \overline{X}$



المنوال > الوسيط> المتوسط

الشكل (3) توزيع ملتوٍ ناحية اليسار

وقد تمكن الإحصائيون من إيجاد علاقة تقريبية بين مقاييس النزعة المركزية الثلاثة: المتوسط والوسيط والمنوال. وتستخدم هذه العلاقة عندما يتعذر استخراج إحداها، كما يحدث عندما يراد إيجاد المتوسط الحسابي في جدول تكراري مفتوح، وهذه العلاقة تسمى علاقة بيرسون ويمكن توضيحها على النحو الآتي:

المتوسط– المنوال =3 (المتوسط – الوسيط)

$$\overline{X} - Mod = 3(\overline{X} - Med)$$

أي أن الفرق بين المتوسط الحسابي والمنوال يعادل ثلاثة أمثال الفرق بين المتوسط الحسابي والوسيط، ويمكن من هذه العلاقة الحصول على أي قيمة من هذه المقاييس إذا عرف اثنان.

في المثال السابق رقم (8) عند حساب الوسيط كانت قيمته (18.80)، وكان المنوال (19.33)، والمطلوب حساب المتوسط الحسابي وفق علاقة بيرسون يساوي:

المتوسط– المنوال =3 (المتوسط – الوسيط)

$$\text{المتوسط الحسابي} = \frac{(3\ \text{الوسيط}) - (\text{المنوال})}{2} = \frac{(19.33) - (18.80 \times 3)}{2} = \frac{19.33 - 56.4}{2} = \frac{37.07}{2} = \mathbf{18.53}$$

نجد أن المتوسط أصغر من الوسيط والمنوال ومنه فإن التوزيع ملتوٍ نحو اليسار (التواء سالب).

ملاحظات هامة:

- المتوسط الحسابي أدق المقاييس الثلاثة لأنه يأخذ في الحسبان جميع فئات التوزيع، يليه الوسيط، وأخيراً المنوال لأن طرائق حسابه تقريبية.
- عادة ما تكون قيم مقاييس النزعة المركزية قريبة من بعضها البعض.

المحاضرة السابعة: مقاييس التشتت

Measures of Dispersion

يشير مفهوم مقاييس النزعة المركزية إلى ميل البيانات للتمركز حول قيمة ممثلة أو نموذجية في التوزيع، وتستخدم مقاييس النزعة المركزية لغايات المقارنة بين مجموعتين من البيانات، ولوصف توزيع البيانات إن كان متماثلاً أم غير متماثل.

وهناك مقياس مكملة لمقاييس النزعة المركزية وهي مقاييس التشتت، إذ أن **مقاييس التشتت تشير إلى مدى تقارب أو تباعد أو تناثر أو اختلاف القيم عن بعضها البعض أو عن نقطة معينة كالمتوسط مثلاً.** فالتشتت يرجع إلى اختلاف قيم البيانات، لذلك فإن القيم المتشابهة لا يوجد لها تشتت، ولكن باختلاف قيم البيانات فلا بد من وجود تشتت بين هذه القيم، ويختلف مقدار التشتت من مجموعة إلى أخرى.

وتبرز الحاجة إلى مقاييس التشتت لتقديم وصف دقيق حول التوزيعات التكرارية وخصوصاً عند إجراء المقارنات بين المجموعات أو بين الأفراد، فدلالات مقاييس النزعة المركزية كالمتوسط الحسابي قد تقود إلى اتخاذ قرارات غير صحيحة.

مثال 1: لتكن لدينا البيانات الآتية لثلاث مجموعات:

الجدول (1) بيانات ثلاث مجموعات

المجموعات	الأولى	الثانية	الثالثة
القيم (البيانات)	5 , 5, 5	4, 5,7, 9,2,3	1, 9
المتوسط	5	5	5

يلاحظ أن قيمة المتوسط الحسابي لكل مجموعة يساوي (5) فإذا اكتفينا بهذا المقياس فإننا نقرر أن المجموعات الثلاث متشابهة ولكن في الحقيقة أن قيم المجموعة الثالثة أكثر تباعداً من قيم المجموعة الثانية والتي هي أكثر تباعداً من قيم المجموعة الأولى.

مثال 2: بفرض أننا اخترنا وبشكل عشوائي مجموعتين من الطلبة الذكور (10 طلاب) والإناث (10 طالبات)، قد امتحنوا في مادة اللغة العربية، وكانت علاماتهم وفق الآتي:

علامات الطلاب الذكور: 77 – 77 – 77 –74 – 73 – 71 – 68 – 67 – 66 – 65

علامات الطالبات الإناث: 100 – 94 – 84 – 77 – 77 – 67 – 63 – 57 – 55 – 41

وعند حساب مقاييس النزعة المركزية لكل مجموعة كانت وفق الآتي:

مجموعة الطلاب الذكور: المتوسط= 71.5 الوسيط = 72 المنوال = 77

مجموعة الطالبات الإناث: المتوسط= 71.5 الوسيط = 72 المنوال = 77

مناقشة: بالنظر إلى علامات المجموعتين نلاحظ للوهلة الأولى أن علاماتهم متماثلة، ولكن لو تفحصنا علامات كل طالب وطالبة لوجدنا أن هناك فرقاً كبيراً بين الطلاب والطالبات، وأن علامات الطالبات تنتشر وتتباعد عن بعضها البعض بشكل أكبر وبمدى أوسع من علامات الطلاب، وبالتالي فالاختلاف بين المفردات صفة لا يمكن أن تظهرها مقاييس النزعة المركزية بل تظهرها مقاييس التشتت.

ومنه فإن مقاييس التشتت تفيد في:

1-تزودنا بمعلومات حول تبعثر (تشتت) أو تجمع البيانات داخل التوزيع، وحول المتوسط الحسابي.

2-تعد مقاييس التشتت مكملاً لمقياس النزعة المركزية في وصف البيانات.

3-تعد مؤشراً على مدى كفاءة وفاعلية مقياس النزعة المركزية في تمثيل البيانات، فمثلاً كلما تجمعت القيم حول المتوسط الحسابي كان هذا المتوسط يمثل بشكل جيد هذه القيم والعكس صحيح، بمعنى أنه كلما كان مقياس التشتت صغيراً كان هذا مؤشراً على أن مقياس النزعة المركزية يمثل بياناته أصدق تمثيل.

وهناك مقاييس عدة للتشتت: المدى والمدى الربيعي والانحراف الربيعي والانحراف المتوسط والانحراف المعياري والتباين. وسيتم التركيز على كل من المدى والانحراف المعياري والتباين.

أولاً-المدى Range:

يعد المدى أسهل وأبسط مقياس من مقاييس تشتت البيانات الكمية، ويعرف **المدى بأنه الفرق بين أكبر قيمة وأصغر قيمة في البيانات**، ويفيد المدى في إعطاء صورة مبسطة للتشتت في الدرجات والتفاوت بينها، ويرمز له اختصاراً بالرمز (R).

-حساب المدى:

1-حساب المدى لبيانات غير مبوبة:

المدى (R) = أكبر قيمة – أصغر قيمة

مثال 3: احسب المدى للبيانات الآتية: 5، 10، 9، 15، 14، 3، 7

المدى (R) = أكبر قيمة – أصغر قيمة = 15 – 3 = 12

بالعودة إلى المثال 2 إذ كانت:

علامات الطلاب الذكور: 77 – 77 – 77 –74 – 73 – 71 – 68 – 67 – 66 – 65

علامات الطالبات الإناث: 100 – 94 – 84 – 77 – 77 – 67 – 63 – 57 – 55 – 41

وعند حساب المدى لعلامات الطلاب الذكور = 77 – 65 = 12

وعند حساب المدى لعلامات الطالبات الإناث = 100 – 41 = 59

مناقشة: نجد أن المدى لعلامات الطلاب أصغر منه لعلامات الطالبات، وهذا يعني أن تشتت علامات الطلاب أصغر من تشتت علامات الطالبات، بمعنى أوضح أن الطلاب كانت علاماتهم متقاربة أكثر من علامات الطالبات التي كانت أكثر تباعداً أو تناثراً.

مثال 4: أوجد المدى لكل مجموعة من المجموعتين الآتيتين:

المجموعة الأولى: 7 – 4 – 6 – 7 – 7 – 6 – 6 – 20

المجموعة الثانية: 20 – 14 – 16 – 5 – 9 – 4 – 18 – 12

المدى للمجموعة الأولى = 20 – 4 = 16 المدى للمجموعة الثانية = 20 – 4 = 16

مناقشة: إن المدى لكلا المجموعتين متساوٍ ولكن عند النظر إلى بيانات المجموعة الثانية نجد أن الاختلاف بين بياناتها أكبر من الاختلاف بين بيانات المجموعة الأولى، فأغلب بيانات مجموعة الأولى متقاربة تتألف من الرقمين 6 و7 في معظمها، لذلك يمكن أن يكون المدى مقياساً مضللاً في بعض الأحيان لاعتماده على القيمتين المتطرفتين اللتين غالباً ما تكونان شاذتين أو متطرفتين.

2–حساب المدى لفئات القيم:

يحسب المدى لفئات القيم من خلال: **المدى (R) = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى**

مثال 5: احسب المدى لبيانات 20 طالباً في مادة العلوم:

الجدول (2) بيانات 20 طالباً في مادة العلوم

الفئات	التكرار
10 – 14	2
14 – 18	3
18 – 22	5
22 – 26	6
26 – 30	4
	$\Sigma f = 20$

الحد الأدنى للفئة الأولى

الحد الأعلى للفئة الأخيرة

المدى (R) = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى = 30 – 10 = 20

-خواص المدى:

- يعد المدى أسهل وأبسط مقاييس التشتت، ولذلك فهو يستخدم في الأوساط العامة.
- يعطي فكرة سريعة عن طبيعة البيانات ويستخدم كثيراً في مراقبة جودة الإنتاج، وكذلك في وصف الأحوال الجوية.
- يتأثر بالقيم المتطرفة أو الشاذة، لذلك لا يفضل استخدامه في حال وجود قيم متطرفة.
- لا يمكن حسابه من الجداول التكرارية المفتوحة.
- لا يصلح استخدامه إذا كان التوزيع البياني شديد الالتواء.

ثانياً-الانحراف المعياري Standard deviation:

يعد الانحراف المعياري أهم مقاييس التشتت وهو يقوم في جوهره على حساب انحرافات الدرجات عن متوسطها كما تدل تسميته عليه. **فالانحراف المعياري هو الجذر التربيعي لمجموع مربعات انحرافات القيم عن متوسطها مقسوماً على عددها.** في حين أن **مجموع مربعات انحرافات القيم عن متوسطها مقسوماً على عددها يدعى التباين.** ومنه فإن الانحراف المعياري هو الجذر التربيعي للتباين، أو بمعنى آخر إن التباين هو مربع الانحراف المعياري، ويرمز للانحراف المعياري اختصاراً بالرمز (S)، ويرمز للتباين بالرمز (S^2).

-حساب الانحراف المعياري:

هناك طرائق عدة لحساب الانحراف المعياري منها الطريقة أو الأسلوب المباشر، أسلوب المربعات، أسلوب الانحرافات، وأسلوب الانحرافات المختصرة، وسنقتصر على حساب الانحراف المعياري للبيانات المبوبة وغير المبوبة وفق الأسلوب المباشر.

1-حساب الانحراف المعياري للبيانات غير المبوبة (وفق الأسلوب المباشر):

نحسب الانحراف المعياري وفق الأسلوب المباشر للبيانات غير المبوبة وفق الخطوات الآتية:

- نحسب المتوسط الحسابي للقيم $\overline{X}$.
- نحسب انحراف القيم عن المتوسط وذلك بطرح كل قيمة من متوسط القيم.
- نربع كل انحراف من انحرافات القيم.

- نحسب مجموع مربع انحرافات القيم.
- نحسب الانحراف المعياري وفق المعادلة الآتية:

التباين يساوي $S = \frac{\Sigma(X-\overline{X})^2}{n}$

$$S = \sqrt{\frac{\Sigma(X-\overline{X})^2}{n}}$$

مثال 6: لتكن لدينا مجموعة من القيم والمطلوب حساب الانحراف المعياري.

الجدول (3) حساب الانحراف المعياري لمجموعة من القيم

الدرجات X	$X-\overline{X}$	$(X-\overline{X})^2$
5	−2	4
7	0	0
9	+2	4
4	−3	9
10	+3	9
6	−1	1
8	+1	1
Σx= 49 $\overline{X}$ = 7	0	28

2-نحسب انحراف القيم عن المتوسط

3-نربع انحرافات القيم

1-نحسب المتوسط= مجموع القيم/عددها

4-نجمع مربع انحرافات القيم

5-نطبق معادلة الانحراف المعياري

لاحظ أن انحراف القيم عن المتوسط يساوي الصفر

الحل:

- نحسب المتوسط الحسابي للقيم وهو مجموع القيم على عددها ويساوي 7.
- نحسب انحراف القيم عن المتوسط $(X-\overline{X})$ وذلك بطرح كل قيمة من متوسط القيم.
- نربع كل انحراف من انحرافات القيم.
- نحسب مجموع مربع انحرافات القيم.
- نحسب الانحراف المعياري وفق المعادلة الآتية:

$$S = \sqrt{\frac{\Sigma(X-\overline{X})^2}{n}} = \sqrt{\frac{28}{7}} = \sqrt{4} = 2$$

بما أن الانحراف المعياري يساوي 2، فإن التباين يساوي 4.

2-حساب الانحراف المعياري من فئات الدرجات (بيانات مبوبة):

نحسب الانحراف المعياري وفق الأسلوب المباشر للبيانات المبوبة في فئات وفق الخطوات الآتية:

- نحسب مركز كل فئة.
- نحسب المتوسط الحسابي للفئات السابقة وفق المعادلة الآتية:

$$\bar{X} = \frac{\Sigma(x_i f)}{\Sigma f}$$

- نطرح مركز كل فئة من المتوسط $(X - \bar{X})$.
- نربع انحراف مركز كل فئة عن المتوسط $(X - \bar{X})^2$.
- نضرب مربع الانحراف بتكرار كل فئة $f(X - \bar{X})^2$ ثم نجمع العمود .
- نحسب الانحراف المعياري وفق المعادلة الآتية:

$$S = \sqrt{\frac{\Sigma f(X_i - \bar{X})^2}{\Sigma f}}$$

مثال 7: لتكن لدينا مجموعة من فئات القيم والتي تمثل درجات 16 طالباً في مادة دراسية والمطلوب حساب الانحراف المعياري

1-نحسب مركز الفئة

4-نربع الانحرافات

الجدول (4) حساب الانحراف المعياري لمجموعة من الفئات

الفئات	f	X_i	fX_i	$X_i - \bar{X}$	$(X - \bar{X})^2$	$f(X - \bar{X})^2$
2 – 4	4	3	12	−2.87	8.23	32.92
4 – 6	5	5	25	−0.87	0.75	3.75
6 – 8	3	7	21	1.13	1.27	3.81
8 – 10	4	9	36	3.13	9.79	39.16
$\bar{X}$=5.87	∑f=16		94		20.04	79.64

5-نضرب مربع الانحراف بالتكرار

2-نوجد المتوسط

6-نطبق معادلة الانحراف المعياري

3-نطرح مركز الفئة من المتوسط

الحل:

- نحسب مركز كل فئة.
- نحسب المتوسط الحسابي للفئات السابقة وفق المعادلة الآتية:

$$\bar{X} = \frac{\Sigma(x_i f)}{\Sigma f} = \frac{94}{16} = 5.87$$

نطرح مركز كل فئة من المتوسط $(X - \bar{X})$.

- نربع انحراف مركز كل فئة عن المتوسط $(X - \bar{X})^2$.

- نضرب مربع الانحراف بتكرار كل فئة $f(X-\bar{X})^2$ ثم نجمع العمود .
- نحسب الانحراف المعياري وفق المعادلة الآتية:

$$S=\sqrt{\frac{\Sigma f(X_i-\bar{X})^2}{\Sigma f}}=\sqrt{\frac{79.64}{16}}=\sqrt{4.97}=2.22$$

-الانحراف المعياري النسبي (معامل الاختلاف) Coefficient of Variation:

يعد معامل الاختلاف (C.V) من أهم مقاييس التشتت النسبية، ويستخدم بشكل واسع في التحليل الإحصائي **لمقارنة مقدار تشتت مجموعتين (أو أكثر) من القيم** إذا كانت تختلف في متوسطها أو في وحدات قياسها، وحتى عندما تكون وحدات القياس غير مختلفة. **وهو عبارة عن النسبة المئوية للانحراف المعياري منسوبة إلى المتوسط الحسابي** وبالطبع كلما كانت قيمة هذا المعامل كبيرة دل ذلك على وجود تشتت كبير بين مفردات التوزيع والعكس صحيح، ويحسب وفق المعادلة الآتية:

$$\text{معامل الاختلاف النسبي}=\frac{\text{الانحراف المعياري}}{\text{المتوسط الحسابي}}\times 100$$

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

مثال 8: قارن بين تشتت درجات مجموعتين من الطلاب في مادة دراسية، وحدد أيهما أكثر تشتتاً.

العينة الأولى: المتوسط 50، الانحراف المعياري 10

العينة الثانية: المتوسط 45، الانحراف المعياري 8

لا نستطيع الاعتماد على قيمة الانحراف المعياري فقط للقول إن تشتت درجات المجموعة الأولى أكبر من تشتت درجات المجموعة الثانية، وذلك بسبب اختلاف المتوسطات الحسابية في المجموعتين، لذلك سيصبح من الضروري استخدام معامل الاختلاف النسبي لغرض المقارنة، كما يأتي:

معامل الاختلاف النسبي (للمجموعة الأولى) $=\frac{10}{50}\times 100=20\%$

معامل الاختلاف النسبي (للمجموعة الثانية) $=\frac{8}{45}\times 100=17.77\%$

وبالاعتماد على معامل الاختلاف النسبي نستطيع القول بأن تشتت درجات المجموعة الأولى أكبر من تشتت درجات المجموعة الثانية لأن معامل اختلافها أكبر، في حين أنه كان من الممكن أن نصل إلى استنتاج خاطئ لو اعتمدنا على قيم الانحراف المعياري فقط.

-ملاحظات على الانحراف المعياري:

- الانحراف المعياري حساس لبعد أو قرب العلامات من المتوسط ولذلك كلما صغرت قيمته دل ذلك على طبيعة البيانات مقاربة ومتراكمة حول المتوسط، وبالتالي التشتت قليل والعكس هو الصحيح.
- يأخذ في الحسبان جميع القيم في حسابه وبالتالي يستخدم عند المقارنة بين المجموعات والعينات الإحصائية والاستنتاجات الإحصائية المتعلقة بالفرضيات.
- يتصف بالكثير من الخواص الجبرية، ولذلك فاستخداماته كثيرة أكثر من أي مقياس آخر للتشتت.
- يتأثر الانحراف المعياري بالقيم الشاذة أو المتطرفة.
- قيمته دوماً أكبر من أي مقياس تشتت آخر ما عدا المدى.
- قيمته دوماً أصغر في العينة عنه في المجتمع الذي سحبت منه هذه العينة.
- يصعب إيجاد الانحراف المعياري في حالة الفئات المفتوحة.

-الدرجة المعيارية Z Scores:

إن حساب الدرجات المعيارية مهم جداً عندما نريد مقارنة تحصيل متعلّم واحد في أكثر من مادة، ومقارنة القيم التي تنتمي لأكثر من مجموعة إحصائية والتي لا تصلح للمقارنة أساساً، كأن نقارن تحصيل طالبين من شعبتين مختلفتين في مادة ما، أو المقارنة بين دخل موظفين اثنين في بلدين.

هنا المقارنة المباشرة غير عادلة ولا يمكن إجراؤها من خلال القيم الظاهرة، ولابد من إجراء عملية تحويلية لهذه القيم، وتحويلها إلى درجات معيارية يمكن أن نقيم المقارنة من خلالها. إن **الدرجة المعيارية للعلامة الأصلية: هي عدد الانحرافات التي تنحرفها العلامة الأصلية عن الوسط الحسابي لعلامات المجموعة، ويرمز لها الرمز (ز) أو (Z).**

ويمكن تحويل الدرجة الخام إلى درجة معيارية من خلال المعادلة الآتية:

$$Z = \frac{X-\overline{X}}{S}$$

إذ Z: الدرجة المعيارية X: الدرجة الخام $\overline{X}$: المتوسط S: الانحراف المعياري.

مثال 9: حصل طالب في مادة الإحصاء على درجة قدرها 85، علماً أن متوسط درجات جميع الطلاب في مادة الإحصاء 71 درجة وانحراف معياري 10، وحصل هذا الطالب في مادة مفاهيم هندسية على درجة قدرها

82 علماً أن متوسط درجات جميع الطلاب في مادة مفاهيم هندسية 77 درجة وانحراف معياري 12، في أي من المقررين كان تحصيله أفضل؟

الحل: عند إجراء المقارنة للدرجات الخام نجد أن تحصيل الطالب في مادة الإحصاء أفضل من تحصيله في مادة مفاهيم هندسية، ولكن حتى تكون المقارنة عادلة لا بد من تحويل الدرجات الخام إلى درجات معيارية وفق الآتي:

مادة الإحصاء $Z1 = \frac{X1-X1}{S1} = \frac{85-71}{10} = 1.4$

مادة مفاهيم هندسية $Z2 = \frac{X2-X2}{S2} = \frac{82-77}{12} = 0.41$

يتضح من خلال مقارنة الدرجات المعيارية أن تحصيل الطالب أعلى في مادة الإحصاء مقارنة مع مادة مفاهيم هندسية، لأن الدرجة المعيارية لمادة الإحصاء Z1 أكبر من الدرجة المعيارية لمادة مفاهيم هندسيةZ2 .

المحاضرة الثامنة: الارتباط

Correlation

معنى الارتباط:

سوف نتعرف في هذه المحاضرة على أساليب تحليل جديدة نستطيع استخدامها في تحليل متغير ما، من خلال علاقته بمتغير آخر أو بمتغيرات عدة، إذ يوجد عدد كبير من الظواهر أو المتغيرات التي يوجد بينها علاقة، مثل العلاقة بين دخل الأسرة وانفاقها، الذكاء والتحصيل، القلق والتحصيل، مساحة الدائرة ونصف القطر، مساحة المربع وطول ضلعه، وبالتالي نستخدم أسلوب إحصائي يسمى **تحليل الارتباط**، لدراسة العلاقة بين متغيرين اثنين أو أكثر، ويجب التنبه إلى أن الارتباط بين متغيرين معناه أن التغير في أحدهما يكون عادة مصحوباً بتغير في الآخر إما إيجاباً أوسلباً.

يعرف الارتباط أنه: "علاقة بين متغيرين أو أكثر"، ويقاس الارتباط بمعامل الارتباط، إذ تتراوح قيمة معامل الارتباط بين (−1، +1) مروراً بالصفر.

أنواع الارتباط:

بناء على **عدد المتغيرات** التي تدخل في حساب الارتباط فإنه يمكن حصر أنواع الارتباط بما يأتي:

1-الارتباط الثنائي Simple corrolation:

ويدرس العلاقة بين متغير واحد مستقل (وهو المتغير الذي يؤثر في المتغير الآخر، أو الذي يحدث أولاً) ومتغير واحد تابع (وهو المتغير الذي يتأثر بالمتغير المستقل، أو الذي يحدث لاحقاً) وتمثل المعادلة على شكل خط مستقيم أي معادلة من الدرجة الأولى.

2-الارتباط المتعدد Multiple Corrolation:

وهو الارتباط الذي يدرس العلاقة بين متغير تابع واحد من جهة وعدد من المتغيرات المستقلة (اثنان أو أكثر) من جهة أخرى تدرس جملة واحدة، ويمثل بمعادلة من الدرجة الثانية فأكثر، أي خط منحني أو قطع كامل أو قطع ناقص.

مثلاً الراتب الشهري للموظف يعد متغيراً مستقلاً عند دراسة علاقته مع المصروف (ارتباط ثنائي) لأنه هو الذي يحدد بدرجة أساسية مقدار المصروف اليومي، بينما يعد الراتب الشهري متغيراً تابعاً يتأثر بمتغيرات مستقلة عدة كالمؤهل العلمي، سنوات الخبرة، طبيعة العمل (ارتباط متعدد إذا تم دراستها معاً).

ويمكن أن يتم دراسة الارتباط أو العلاقة بين متغيرات عدة من جهة (تحصيل الطالب في العلوم العامة) ومتغيرات عدة من جهة أخرى (تحصيل الطالب في المواد الاجتماعية)، أي دراسة العلاقة بين متغيرات عدة مستقلة ومتغيرات عدة تابعة، ويسمى الارتباط في هذه الحالة **بالارتباط القانوني**.

مؤشرات الارتباط:

يتم الكشف عن درجة الارتباط بين متغيرين بواسطة مؤشرين هما شكل الانتشار، ومعامل الارتباط.

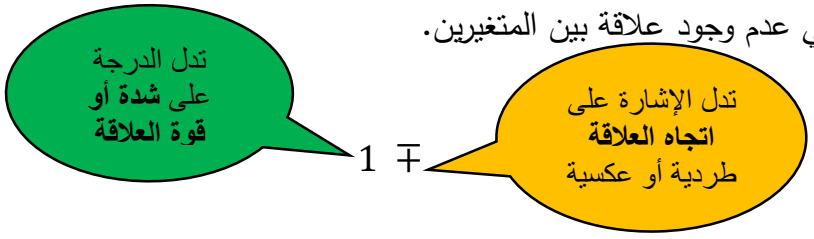
1-معامل الارتباط:

تتراوح قيمة معامل الارتباط بين (-1 و+1) **إذ تدل القيمة المطلقة (الدرجة) على قوة الارتباط/ العلاقة، وتدل الإشارة على اتجاه العلاقة**. ويمكن أن يكون الارتباط سالباً (عكسياً) أو موجباً (طردياً)، وكلما اقتربت قيمة معامل الارتباط من الواحد كان الارتباط أقوى.

+1 تعني وجود ارتباط موجب (طردي)، أي أن الشخص الذي حاز على المرتبة الأولى في الاختبار الأول حصل على المرتبة الأولى في الاختبار الثاني (فالارتباط الموجب يعني زيادة قيم أحد المتغيرين مع زيادة قيم المتغير الآخر والعكس صحيح).

-1 تعني وجود ارتباط سالب (عكسي)، أي أن الشخص الذي حصل على المرتبة الأولى في الاختبار الأول حصل على المرتبة الأخيرة في الاختبار الثاني، والشخص الذي حصل على الثانية في الاختبار الأول كان ترتيبه الثاني قبل الأخير في الاختبار الثاني وهكذا... (فالارتباط السالب يعني نقصان قيم أحد المتغيرين مع زيادة قيم المتغير الآخر).

أما **الصفر** فيعني عدم وجود علاقة بين المتغيرين.



ويبين الجدول الآتي قيم معامل الارتباط:

الجدول (1) قوة معامل الارتباط بدلالة القيمة العددية له

قوة معامل الارتباط	مدى معامل الارتباط
تام	+1 أو -1
عال جداً	من 0.80 إلى 0.99
عال	من 0.60 إلى 0.79
متوسط	من 0.40 إلى 0.59
ضعيف	من 0.20 إلى 0.39
ضعيف جداً	من 0.01 إلى 0.19
لا توجد علاقة	صفر

2-شكل الانتشار Scatter Plot:

يقصد بشكل الانتشار تمثيل البيانات عن الظاهرتين أو المتغيرين على المحاور الإحداثية، إذ يمثل المتغير المستقل X على محور السينات الأفقي، ويمثل المتغير التابع Y على محور العينات العمودي، وتمثل كل ثنائية (X, Y) بنقطة على المستوى الإحداثي، ويمكن الحصول على شكل الانتشار مباشرة من جدول المتغيرين لنفس المفحوصين.

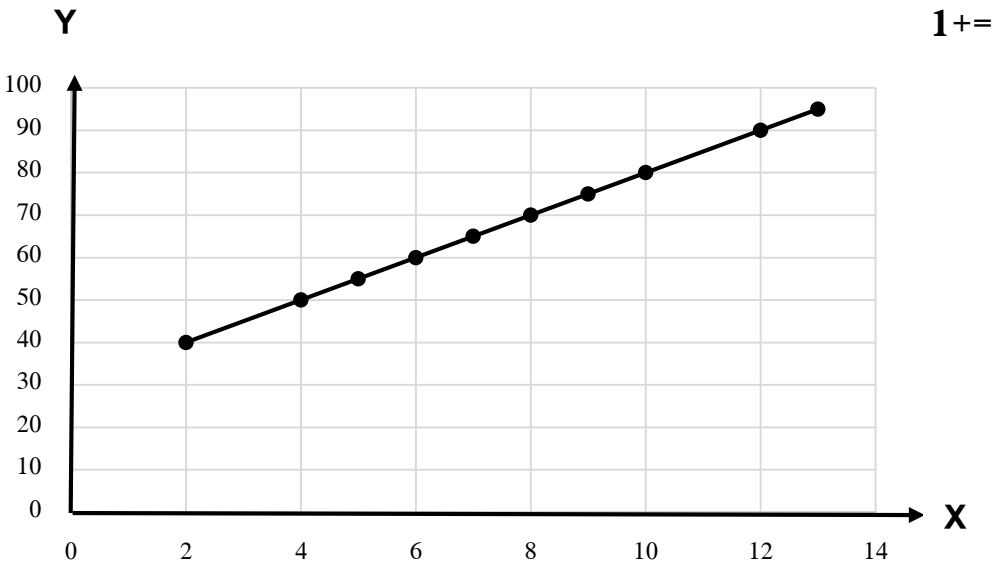
مثال 1: يمثل الجدول الآتي درجات 10 أطفال في كل من العمر والذكاء.

الجدول (2) درجات 10 أطفال في كل من العمر والذكاء

الأطفال	1	2	3	4	5	6	7	8	9	10
العمر X	2	4	5	6	7	8	9	10	12	13
الذكاء Y	40	50	55	60	65	70	75	80	90	95

ويمكن تحويل درجات الأطفال في هذا الجدول إلى شكل انتشار كما في الشكل الآتي:

ر=+1



الشكل (1) علاقة موجبة كاملة بين متغيرين

يلاحظ من بيانات الجدول والرسم البياني السابقين أن نقاط الانتشار تقع على مستقيم واحد، وبالتالي فالارتباط خطي تام، وهذا الخط يدل على علاقة موجبة كاملة تدل على أن الزيادة في المتغير x تناظرها زيادة بنفس النسبة في المتغير (y) والنقص في المتغير (x) يناظره نقص بنفس النسبة في المتغير (y) وهذه العلاقة يعبر عنها بمعامل الارتباط +1 وهو أقصى معامل يمكن الحصول عليه. فلا يمكن لمعامل الارتباط أن يتجاوز الواحد الصحيح سلباً أو إيجاباً، ويعبر عن علاقة حتمية، أي أن المفحوص الذي يكون له عمر معين فإن من المحتم أن يكون ذكاؤه محدداً. وكذلك يمكن استنتاج العمر من الذكاء. إلا أنه تجدر

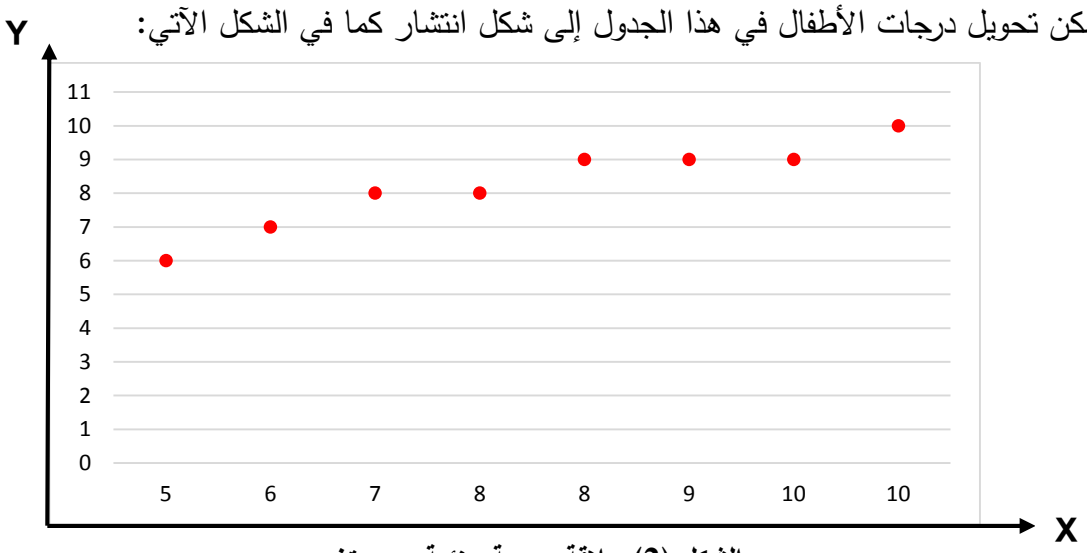
الإشارة أن هذا المثال هو مثال افتراضي ويندر وجوده في العلوم الإنسانية والاجتماعية، فمعاملات الارتباط دائماً فيها أقل من الواحد الصحيح.

مثال 2: يمثل الجدول الآتي درجات 8 تلاميذ في مادتي الحساب والعلوم.

الجدول (3) درجات 10 تلاميذ في كل من مادتي الحساب والعلوم

التلميذ	1	2	3	4	5	6	7	8
الحساب X	5	6	7	8	8	9	10	10
العلوم Y	6	7	8	8	9	9	9	10

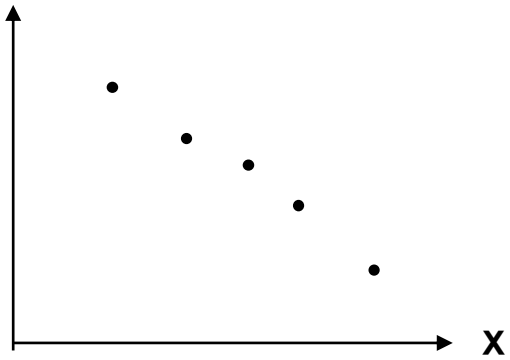
ويمكن تحويل درجات الأطفال في هذا الجدول إلى شكل انتشار كما في الشكل الآتي:



الشكل (2) علاقة موجبة جزئية بين متغيرين

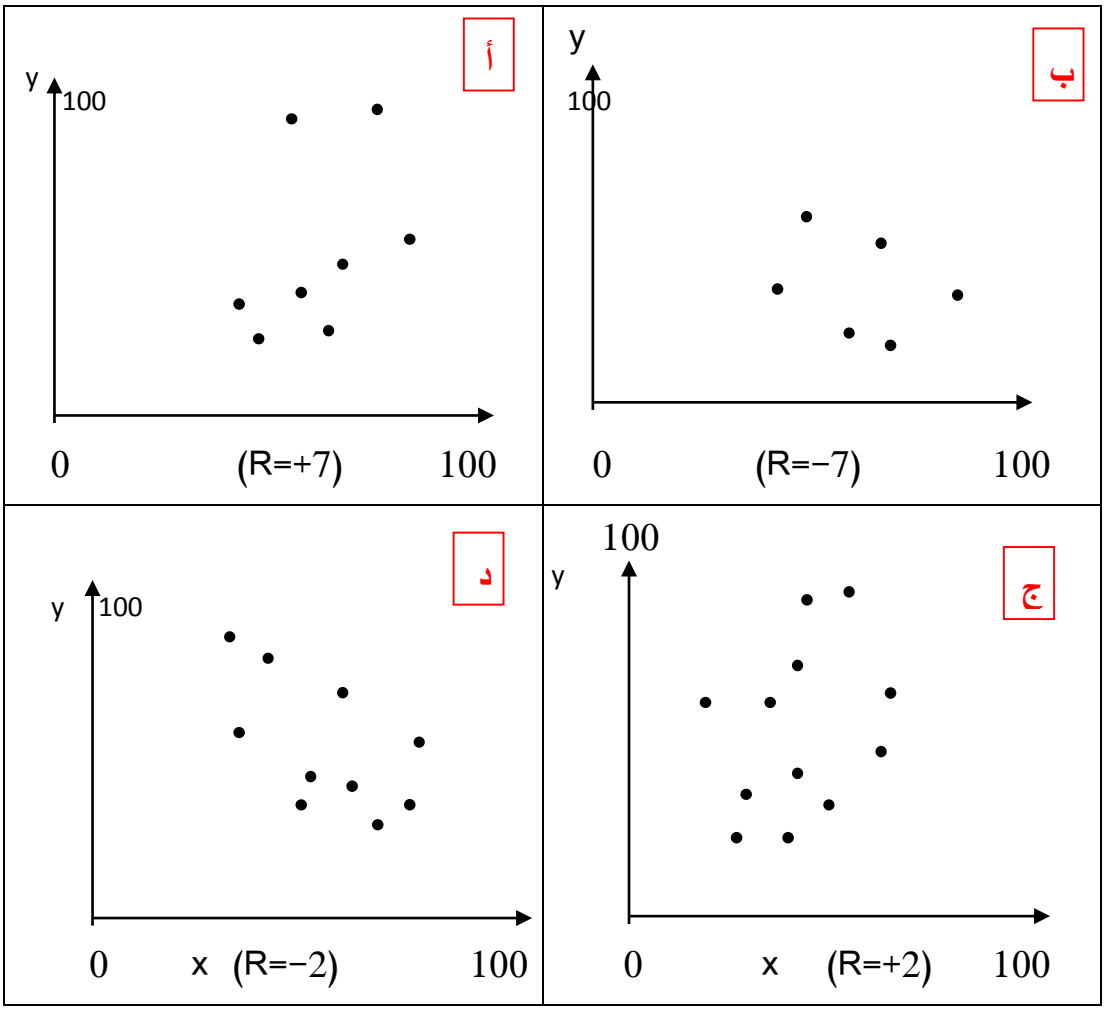
يتبين من الرسم السابق أن شكل الانتشار للمتغيرين ليس خطاً مستقيماً كما هو الحال في الشكل السابق. ومع ذلك نلاحظ بصفة عامة أن الشخص الذي يحصل على درجة عالية في المتغير x أي الحساب يحتمل أن يحصل على درجة عالية في المتغير y أي العلوم، كما أن الشخص الذي يحصل على درجة منخفضة في x يحتمل أن يحصل على درجة منخفضة في y أيضاً . وهذا النوع من العلاقات يسمى العلاقة الجزئية الموجبة، وحين تحسب كمعامل ارتباط يكون مقدارها كسراً عشرياً أعلى من الصفر وأقل من الواحد الصحيح وتكون إشارته الجبرية موجبة .

وبالتأكيد يمكن للعلاقات أن تكون سالبة، فمعامل الارتباط السالب التام أو الكامل −1 يعبر عن علاقة حتمية بين المتغيرين في اتجاه عكس العلاقة الموجبة التامة. وهو نادر الحدوث في العلم عادة، إن لم يكن مستحيلاً في العلوم الإنسانية والاجتماعية. ولعل أقرب الأمثلة التي توضحه العلاقة بين حجم الغاز x وضغطه y، فمن المعروف أن الزيادة في حجم الغاز x تؤدي إلى قلة الضغط y، والنقص في الحجم يؤدي إلى زيادة الضغط، ويمكن التعبير عن هذه العلاقة في الشكل الآتي:



الشكل (3) علاقة سالبة كاملة

إن العلاقات الجزئية فقد تكون موجبة أو سالبة، وقد تكون كبيرة أو صغيرة، ويتحدد اتجاه العلاقة بالإشارة الجبرية لمعامل الارتباط عند حسابه، أما حجم العلاقة فيتحدد في ضوء مقدار الكسر العشري المحسوب، ويوضح الشكل التالي علاقات جزئية موجبة وسالبة، كبيرة وصغيرة.

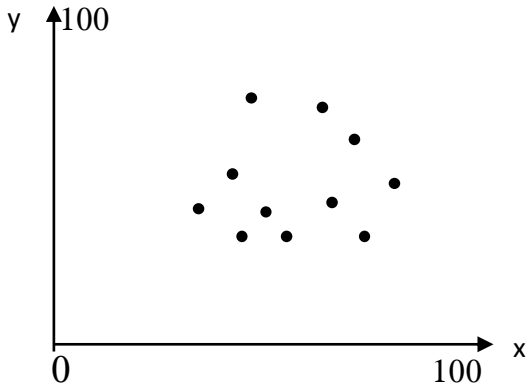


الشكل (4) علاقات جزئية

يبين الشكل السابق علاقات جزئية مختلفة: (أ) علاقة جزئية موجبة كبيرة (R=+7) (ب) علاقة سالبة كبيرة (R=−7)، (ج) علاقة جزئية موجبة صغيرة (R=+2)، (د) علاقة جزئية سالبة صغيرة (R=−2).

وهذه العلاقات الجزئية هي الأكثر شيوعاً في البحوث العلمية عامة، والبحوث النفسية والتربوية والاجتماعية خاصة.

وقد تكون العلاقة بين المتغيرين صفرية، والتي تعني عدم وجود علاقة بين المتغيرين ومن أمثلة ذلك العلاقة بين طول القامة والذكاء. والصفر هنا لا يقصد به الصفر بمعناه الرياضي المعتاد، وإنما يدل على أن معامل الارتباط ليست له دلالة إحصائية مهما كان مقداره. وحسبنا أن ننتبه هنا إلى خطأ فادح شائع في كثير من البحوث الحديثة حين يفسر بعض الباحثين معامل الارتباط بأنه موجب أو سالب بينما هو غير دال إحصائياً. إن معامل الارتباط في هذه الحالة ليس إلا صفراً، ولا يحمل معنى العلاقة الموجبة أو السالبة الحال، وتكون نقاط الانشار مبعثرة دون أن يكون لها أي اتجاه معين.



الشكل (5) علاقة صفرية (R=0.09)

خلاصة لأنواع العلاقات:

في ضوء ما سبق يمكن تصنيف العلاقات في ضوء الجدول الآتي:

الجدول (4) خلاصة لأنواع العلاقات

قيمة معامل الارتباط	التفسير	مثال العلاقة بين:
+1	علاقة موجبة كاملة	محيط الدائرة وقطرها مساحة المربع وطول ضلعه
0.01 ، 0.99	علاقة موجبة/طردية جزئية	الذكاء والتحصيل الدافعية والإنجاز
0	لا توجد علاقة	طول الجسم والذكاء
−0.01 ، −0.99	علاقة سلبية/عكسية جزئية	القلق والتحصيل
−1	علاقة سالبة كاملة	حجم الغاز وضغطه

شروط حساب معامل الارتباط:

هناك شروط يجب أن تتوفر بين الظاهرتين المدروستين حتى نستطيع دراسة علاقتهما المشتركة وهي:

- أن تكون بين الظاهرتين المدروستين علاقة جدلية واضحة.
- أن تكون كل من الظاهرتين قابلتين للقياس بواسطة وحدة قياس معينة لكل منهما (مستقلة أو مشتركة).
- أن تكون القياسات المأخوذة متقابلة من حيث الزمان أو المكان أو كلاهما معاً.

طرائق حساب معامل الارتباط:

يمكن استخدام معامل الارتباط بين متغيرين بطرائق عدة منها:

1. **معامل بيرسون** (Pearson): يستخدم إذا كان كلا المتغيرين مقاساً بمقياس **كمي** مثل إيجاد معامل الارتباط بين الدخل والإنفاق أو بين الذكاء والتحصيل.
2. **معامل سبيرمان** (Spearman): يستخدم إذا كان كلاً من المتغيرين مقاساً **بمقياس ترتيبي** مثل إيجاد علاقة المستوى التحصيلي (مرتفع – متوسط – منخفض) والمستوى التعليمي للأم (ابتدائي – إعدادي – ثانوي– جامعة – دراسات عليا)، أو كان أحد المتغيرين ترتيبياً والآخر كمياً كالعلاقة بين الذكاء ومستوى التحصيل، كما يمكن استخدام معامل سبيرمان في حال كان **المتغيرين كميين**.

1-معامل ارتباط بيرسون:

سمي معامل ارتباط بيرسون (Person) نسبة للعالم الإحصائي الذي استخدمه لأول مرة بشكل منهجي ويستخدم لدراسة العلاقة بين **متغيرين كميين**. ولحساب هذا المعامل لا بد من توفر قيمة لكل متغير عند نفس الفرد، بمعنى أن هذه القيم عبارة عن قيم مزدوجة وهذا يعني أيضاً تساوي عدد هذه القيم في المتغيرين.

ويحسب معامل ارتباط بيرسون وفق طرائق مختلفة منها باستخدام المعادلة الأساسية، أو باستخدام انحرافات المتغيرين عن متوسطيهما، أو وفق الطريقة العامة، وسيتم حساب معامل ارتباط بيرسون باستخدام انحرافات المتغيرين عن متوسطيهما وهي من الطرائق المعروفة والمستخدمة كثيراً وفق المعادلة الآتية:

$$R = \frac{\Sigma(X - \overline{X})(Y - \overline{Y})}{\sqrt{\Sigma(X - \overline{X})^2 \, \Sigma(Y - \overline{Y})^2}}$$

مثال 3: احسب معامل ارتباط بيرسون بين درجات خمسة تلاميذ من الصف الرابع في مادتي اللغة العربية (**X**) والعلوم (**Y**)، وبين طبيعة الارتباط:

الجدول (6) معامل ارتباط بيرسون بين درجات خمسة تلاميذ في مادتي اللغة العربية (X) والعلوم (Y)

X	Y	$X-\bar{X}$	$Y-\bar{Y}$	$(X-\bar{X})(Y-\bar{Y})$	$(X-\bar{X})^2$	$(Y-\bar{Y})^2$
2	3	−2	−2	+4	4	4
4	5	0	0	0	0	0
5	8	+1	+3	+3	1	9
7	5	+3	0	0	9	0
2	4	−2	−1	+2	4	1
$\Sigma X = 20$	$\Sigma Y = 25$	$\Sigma = 0$	$\Sigma = 0$	$\Sigma = 9$	$\Sigma = 18$	$\Sigma = 14$

2-نحسب انحراف كل درجة عن متوسطها

4-نربع كل انحراف

1-نوجد متوسط X و Y

3-نحسب ناتج ضرب الانحرافات

$$\bar{X} = \frac{\Sigma X}{n} = \frac{20}{5} = 4$$

$$\bar{Y} = \frac{\Sigma Y}{n} = \frac{25}{5} = 5$$

$$R = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\sqrt{\Sigma(X-\bar{X})^2\,\Sigma(Y-\bar{Y})^2}}$$

$$R = \frac{9}{\sqrt{18*14}} = \frac{9}{\sqrt{252}} = \frac{9}{15.87} = 0.56$$

5-نطبق قانون الارتباط

إن الارتباط بين درجات التلاميذ في المادتين بلغ (0.56) وهو ارتباط طردي (موجب) ومتوسط الشدة.

2-معامل ارتباط سبيرمان (للرتب):

يستخدم معامل ارتباط الرتب لسبيرمان (نسبة إلى واضعه)، لقياس العلاقة بين متغيرين من المستوى الترتيبي، أو أحدهما ترتيبي والآخر كمي، أو كلاهما كميان. وبمقارنة كفاءة هذا المعامل مع كفاءة معامل بيرسون فإن معامل سبيرمان يقترب 91% من معامل بيرسون، ويحسب معامل سبيرمان وفق المعادلة:

$$R_s = 1 - \frac{6\Sigma D^2}{n(n^2-1)}$$

حيث **D**: فرق رتبة **X** من رتبة **Y** لكل ثنائية **n**: عدد الأفراد أو المشاهدات.

مثال 4: احسب معامل ارتباط سبيرمان بين عدد ساعات الدراسة **(X)** والتحصيل الدراسي **(Y)** لخمسة تلاميذ وفق الجدول الآتي:

الجدول (7) معامل ارتباط سبيرمان بين عدد ساعات الدراسة (X) والتحصيل الدراسي (Y)

X	Y	رتبة x	رتبة y	فرق الرتب D	D^2
2	3	5	5	0	0
4	5	4	3.5	0.5	0.25
5	6	3	2	1	1
7	5	2	3.5	−1.5	2.25
10	8	1	1	0	0
					$\sum D^2=3.5$

1-نوجد رتب X و Y

3-نربع فرق الرتب

2-نحسب فرق الرتب X-Y

ملاحظة: نرتب القيم، فالمفردة أو القيمة الأكبر تأخذ المرتبة الأولى وهكذا. أما إذا تساوت قيمتان في الرتبة فإنهما تأخذان مرتبتين متتاليتين ويتم جمع المرتبتين وتقسيمهما على 2 فنحصل على رتبة كل منهما. وإذا تساوت ثلاثة قيم في الرتب فإنها تأخذ ثلاث مرات متتالية ويتم جمع المراتب الثلاث وتقسيمها على ثلاثة فنحصل على رتبة كل منهم.

4-نطبق القانون

$$R_s = 1 - \frac{6\Sigma D^2}{n(n^2-1)}$$

$$R_s = 1 - \frac{6*3.5}{5(5^2-1)}$$

$$R_s = 1 - \frac{21}{120} = 1 - 0.175 = 0.825$$

ومنه فإن الارتباط بين رتب x ورتب y قد بلغ (0.825) وهو ارتباط طردي (موجب) ومرتفع/عالٍ جداً.

مثال 5: يبين الجدول الآتي نتائج تقييم مجموعة من الطلاب مكونة من 10 طلاب، في مادتي الجغرافية والتاريخ، والمطلوب احسب معامل ارتباط سبيرمان لإيجاد العلاقة بين أداء الطلاب في المادتين:

الجدول (8) تقييم الطلاب في مادتي الجغرافية والتاريخ

الطالب	1	2	3	4	5	6	7	8	9	10
التاريخ X	ممتاز	ممتاز	ضعيف	جيد	جيد جداً	جيد	وسط	ضعيف جداً	ضعيف	جيد
الجغرافية Y	جيد جداً	ممتاز	وسط	جيد	جيد	وسط	وسط	ضعيف	ضعيف جداً	وسط

الجدول (9) معامل ارتباط سبيرمان بين تقييم الطلاب في مادتي الجغرافية والتاريخ

تأخذ "ممتاز" الرتب 1 +2 ويقسم الناتج على 2 = 1.5

تأخذ "الوسط" الرتب 5+6+7+8 ويقسم الناتج على 4 = 6.5

التاريخ X	الجغرافية Y	رتبة x	رتبة y	فرق الرتب D	D^2
ممتاز	جيد جداً	1.5	2	0.5-	0.25
ممتاز	ممتاز	1.5	1	0.5	0.25
ضعيف	وسط	8.5	6.5	2	4
جيد	جيد	5	3.5	1.5	2.25
جيد جداً	جيد	3	3.5	0.5-	0.25
جيد	وسط	5	6.5	1.5-	2.25
وسط	وسط	7	6.5	0.5	0.25
ضعيف جداً	ضعيف	10	9	1	1
ضعيف	ضعيف جداً	8.5	10	1.5-	2.25
جيد	وسط	5	6.5	1.5-	2.25
					$\sum D^2=15$

$$R_s = 1 - \frac{6\sum D_i^2}{n(n^2-1)} = 1 - \frac{6*15}{10(100-1)} = 1 - \frac{90}{990} = 1 - 0.09$$

$$R_s = 0.91$$

ومنه فإن الارتباط بين رتب x ورتب y قد بلغ (0.91) وهو ارتباط طردي (موجب) ومرتفع/عالٍ جداً.