



وزارة التعليم العالي والبحث العلمي

جامعة الفرات الأوسط التقنية

المعهد التقني الديوانية

قسم تقنيات الإدارة الصحية

محاضرات مادة

مبادئ الإحصاء الحيوي

BIO- STATISTICS

2025-2024

الإحصاء:

الإحصاء سابقاً هو أحد فروع الرياضيات ذات التطبيقات الواسعة في مختلف المجالات، اما في الوقت الحاضر فقد أصبح علماً مستقلاً بذاته. يهتم علم الإحصاء بجمع وتلخيص وتصنيف وتمثيل البيانات ومن ثم التوصل الى استنتاجات من مجموعة البيانات المتوفرة وأخيراً يساعدنا في اتخاذ القرارات المبنية على أسس علمية منطقية. كل هذا يجعله ذو أهمية تطبيقية واسعة في شتى مجالات العلوم من الفيزياء والحاسوب إلى العلوم الاجتماعية وحتى الإنسانية، كما يلعب دوراً أساسياً في السياسة والأعمال.

تتضوي المصطلحات المفتاحية لعلم الإحصاء بشكل أساسي على مفاهيم نظرية الاحتمالات منها:

- مجتمع إحصائي **population**
- عينة **sample**
- وحدة معاينة **sampling unit**
- احتمال **probability**

الخطوة الأولى في أي عملية إحصائية هي جمع البيانات **data** من خلال عملية المعاينة **sampling** من ضمن المجتمع الإحصائي الضخم أو من خلال تسجيل الاستجابات لمعالجة ما في تجربة (تصميم تجريبي **experimental design**)، أو عن طريق ملاحظة عملية متكررة مع الزمن (سلاسل زمنية **time series**)، ومن ثم وضع خلاصات رقمية وتمثيلية (مخططات) **graphical** باستخدام ما يدعى الإحصاء الوصفي **descriptive statistics**.

الأنماط الموجودة ضمن البيانات يتم نمذجتها **modeling** لأخذ استدلالات حول مجتمعات كبيرة، لذلك يجب دراسة حجم العينة بحيث تكون ممثلة للمجتمع الإحصائي المسحوبة منه أفضل تمثيل. تتم هذه العملية ضمن ما يدعى الإحصاء الاستدلالي **inferential statistics** ليأخذ بعين الاعتبار عشوائية الملاحظات (القياسات).

الإحصاء: نشأته وعلاقته بالعلوم الأخرى:

استخدم المسلمون الأوائل الإحصاء في اللغة والذي سمي فيما بعد بالإحصاء اللغوي وهو عملية إحصاء تواتر الحروف في لغة ما، وهي فكرة رياضية على غاية من الأهمية، هي فكرة قانون الأعداد الكبيرة. ولعل الكندي هو أول من أجرى ذلك الإحصاء في تاريخ الدراسات الكمية على اللغة، ولا شك في أنه أفاد من إحصائيات حروف القرآن الكريم التي سبقت، ومن آياته تعالى

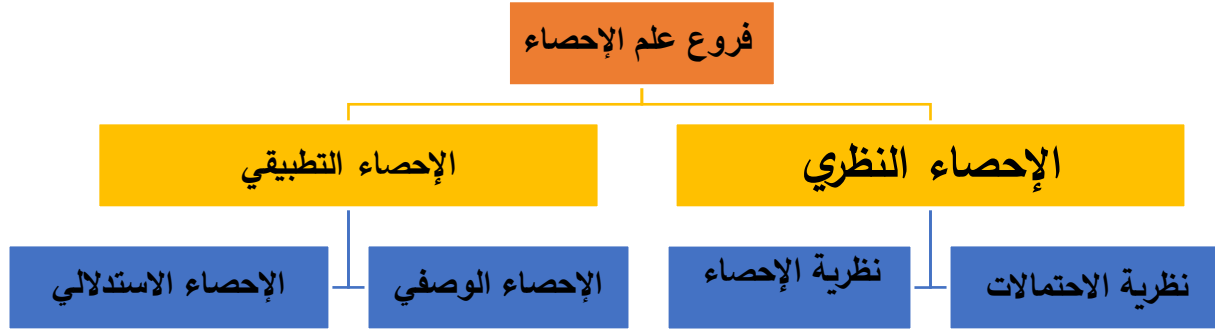
- وَإِنْ تَعُدُّوا نِعْمَةَ اللَّهِ لَا تُحْصُوهَا إِنَّ اللَّهَ لَعَفُورٌ رَحِيمٌ... (النحل 18)
- ثُمَّ بَعَثْنَاهُمْ لِنَعْلَمَ أَيُّ الْجَرْبَيْنِ أَحْصَى لِمَا لَبِثُوا أَمْدًا ... (الكهف 12)
- إِنَّ كُلَّ مَنْ فِي السَّمَاوَاتِ وَالْأَرْضِ إِلَّا آتَى الرَّحْمَنَ عَبْدًا لَقَدْ أَحْصَاهُمْ وَعَدَّهُمْ عَدًّا. (مريم 93 / 94)
- لِيَعْلَمَ أَنْ قَدْ أَبْلَغُوا رَسُولَاتِ رَبِّهِمْ وَأَحَاطَ بِمَا لَدَيْهِمْ وَأَحْصَى كُلَّ شَيْءٍ عَدَدًا.. (الجن 28)

كما كان للعرب في الإحصاء الاجتماعي أيضاً أثرٌ يجدر ذكره، وهو أن المفكر العربي ابن خلدون ربما كان أول من عالج قضايا السكان معالجة علمية، فبحث في عمران الدول واتساعها وتأخرها، وربط كل ذلك بنمو عدد السكان ونقصانهم.

وفي بداية القرن التاسع عشر الميلادي دخل الإحصاء مرحلة من مراحل تطوره على يد العالم لاپلاس (Laplace) الذي يجب أن يوضع في مقدمة أولئك الذين جعلوا من حساب الاحتمالات الأداة الأساسية لدراسة التحليل الإحصائي، وقد أوضح في كتابه «النظرية التحليلية للاحتتمالات» عام 1812 الفوائد والميزات التي يمكن أن تستخلص من دراسة الظواهر الطبيعية التي أسبابها معقدة جداً إلى حد لا يمكن معه معرفتها جميعاً. ثم وسّع كيتلي (1796-1874) حقل تطبيقات الإحصاء مسترشداً بأعمال لاپلاس وقد توسعت منذ نهاية القرن التاسع عشر وإلى يومنا هذا طرائق التحليل الإحصائي فوصلت إلى كل مجالات التحريات والدراسات العلمية، وأدت المسائل العلمية الحديثة المدروسة وفق هذا الأسلوب إلى تطوير سريع وكبير للنظرية الإحصائية. فبعد أعمال كيتلي أنشأ بيرسون (1857-1936) Pearson فرعاً جديداً للإحصاء يحمل اسم الإحصاء الحيوي (Biostatistics) الذي امتد حالياً إلى ميادين الاختبارات المتعلقة بعلم المداواة والطب العلاجي. كذلك فقد توطدت الصلة بين الإحصاء والاقتصاد بابتداع فرع علمي جديد هو الاقتصاد القياسي Econometrics ويدعوه بعضهم الإحصاء الاقتصادي. أما أعمال ماكسويل J. C. Maxwell (1831-1879) التي أوصلت إلى النظرية الحركية للغازات فقد كانت نقطة البدء للميكانيك الإحصائي وللفيزياء النووية. وقد امتدت استخدامات الإحصاء إلى الزراعة على يد فيشر (1890-1962) وإلى الصناعة على يد شيوهارت (1891-1967) W. Shewhart بدراسة المراقبة الإحصائية للجودة والوثوق، أما في ميادين العلوم الإنسانية فإن دراسة سبيرمان. (1863-1945) Spearman حول سلوك الأفراد، التي طوّرت بعد ذلك في علم النفس التطبيقي فقد أدت إلى وجود طرائق التحليل العائلي Factor Analysis الذي هو امتداد منطقي لدراسة الارتباط. Correlation. وفي إدارة الأعمال والمشاريع غدت الطرائق الإحصائية عاملاً مساعداً لا بدّ منه في دراسة حالة السوق ومراقبة الميزانية وإدارة المخزون الاحتياطي، وهي إضافة إلى نظرية المباراة game theory، ونظرية القرار Decision theory وإلى الطرائق الحديثة في الحساب قد مهدت لولادة بحوث العمليات Operations research والذي يدخل في العلوم العسكرية وعلوم الحاسبات والمحاكاة وغيرها من العلوم المهمة.

تعريف علم الإحصاء:

هو العلم الذي يهتم بوصف طرق متعددة لجمع البيانات والمشاهدات ومن ثم يتم تنظيمها وتبويبها وعرضها باستخدام الأساليب العلمية لتحليلها واستخلاص النتائج منها . كما يشار للإحصاء إلى أنه مجموعة من الأساليب العلمية القياسية التي يمكن توظيفها لجمع المعطيات (البيانات والمعلومات) عن الظواهر وتبويبها وتلخيصها وتقييمها والخروج منها باستنتاجات حول مجموعة وحدات المجتمع من خلال اعتماد جزء صغير من هذا المجتمع.



الإحصاء الوصفي: Descriptive Statistics

يمثل الطرق الرقمية أو الحاسوبية لجمع المعلومات والبيانات لتلخيصها أو اختصارها وعرضها في الصور المناسبة (رسومات – جداول – مؤشرات) وذلك سواء من العينة أو المجتمع.

الإحصاء الاستدلالي: Inferential Statistics

ويعنى بتحليل البيانات المتوافرة عن العينة كأساس لتحليل ووصف بيانات المجتمع من خلال أساليب (التقدير - التنبؤ - اختبارات الفرضيات).

المجتمع: Population

مجموعة من الوحدات Units أو المفردات Elements أو المشاهدات Observations التي تتكون منها أي ظاهرة معينة، سواء كانت وحدات العد على شكل مفردة كالأشخاص أو الحيوان أو النبات، أو على شكل مجموعات كالأسر أو المستشفيات أو المدن والقرى وغيرها، ويتحدد حجم المجتمع الإحصائي بعدد الوحدات ويرمز لها بالرمز N .

وهناك نوعين من المجتمعات هما:

المجتمع الحقيقي: المجتمع الحقيقي هو المجتمع الذي يتوافر له إطار كامل ومتكامل مثل عدد الموظفين في دائرة معينة أو عدد الطلاب في مرحلة دراسية معينة بإحدى الكليات .

المجتمع الافتراضي: وهو ذلك المجتمع اللانهائي الذي يفترض الباحث أن جميع مفرداته تجمعها خصائص معينة محل اهتمام الدراسة كرواد النوادي أو الزوار لأحد مراكز التسوق ... الخ.

تدريب: أمثلة للمجتمعات الإحصائية

على كل طالب/طالبة اختيار أو تحديد أحد المجتمعات التي يمكن دراستها.

العينة: Sample

أما العينة فهي جزء من المجتمع يتم اختيارها بأحد طرق المعاينة المعروفة ويجب ان تمثل المجتمع المأخوذة منه افضل تمثيل، ويتحدد حجم العينة بناء على حجم المجتمع ودرجة الدقة المطلوبة، ويرمز لها بالرمز n .

البيانات الإحصائية: عبارة عن معلومات كمية (رقمية) أو نوعية (وصفية) تجمع من مصادر محددة وبطريقة علمية سليمة.

أمثلة للمتغيرات الكمية الطول بالسنتيمتر، الوزن بالكيلوغرام، العمر بالأشهر، الدخل بالدينار، درجة الحرارة المئوية، عدد أفراد الأسرة الخ.

أمثلة للبيانات الوصفية (النوعية) نوع الجنس، القومية، لون الشعر، الجنسية

أسئلة للمناقشة:

1. هل جميع الدراسات تتم على المجتمعات؟
2. لماذا نضطر إلى استخدام العينة في الكثير من الدراسات؟
3. ما هي المزايا التي يمكن تحقيقها من استخدام العينة؟
4. ما هي مزايا دراسة المجتمع على العينة؟
5. هل تتشابه طرق اختيار العينة؟

أنواع العينات

يمكن تقسيم العينات وفقا لأسلوب اختيارها إلى نوعين هما:

- أ- العينات الاحتمالية
- ب- العينات غير الاحتمالية

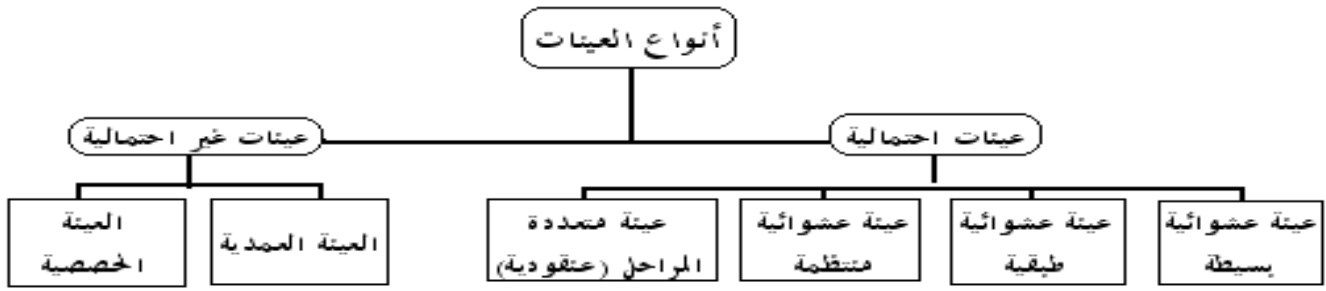
شكل رقم (1) الفرق بين المجتمع والعينة



عينة الدراسة

مجتمع الدراسة

شكل رقم (2) أنواع العينات



أولاً: العينات الاحتمالية

هي العينات التي يتم اختيار مفرداتها وفقاً لقواعد الاحتمالات، بمعنى آخر هي التي يتم اختيار مفرداتها من مجتمع الدراسة بطريقة عشوائية، بهدف تجنب التحيز الناتج عن اختيار المفردات، ومن أهم أنواع العينات الاحتمالية، ما يلي:

1. العينة العشوائية البسيطة Simple Random Sample.
2. العينة العشوائية الطبقية Stratified Random Sample.
3. العينة العشوائية المنتظمة Systematic Random Sample.
4. العينة العقودية أو المتعددة المراحل Cluster Sample.

ثانياً: العينات غير الاحتمالية

هي العينات التي يتم اختيار مفرداتها بطريقة غير عشوائية، حيث يقوم الباحث باختيار مفردات العينة بالصورة التي تحقق الهدف من المعاينة، مثل اختيار عينة من مرضى السكري في فئة عمر معينة، وأهم أنواع العينات غير الاحتمالية:

1. العينة العمدية Judgmental Sample.
2. العينة الحصصية Quota Sample.

ممارسة عملية: يتوزع الطلبة الى ستة مجاميع كل مجموعة تقوم بدراسة أحد أنواع العينات مع إعطاء مثال

مراحل الطريقة الاحصائية

الطريقة الاحصائية Statistical Method هي طريقة علمية تختص بالمعالجة والتحليل الكمي القياسي للظاهرة او الظواهر قيد الدراسة ولذلك فان تطبيق هذه الطريقة يكون مرهونا بإمكانية التعبير الكمي للبيانات. وتمتاز بكونها تعمل بمثابة الخوارزمية الرياضية التي توفر دليلا ومنهجية علمية واضحة يمكن السير عليها للوصول الى البحث العلمي الرصين , ويمكن تلخيص مراحل الطريقة الاحصائية كما يلي:

1- مرحلة تحديد الهدف او مشكلة الدراسة

ان هذه الخطوة تعتبر الخطوة الاساسية في مراحل الطريقة الاحصائية لانها ترسم الطريق امام الباحث وترسي القواعد وتحدد له المسار الذي يجب ان يسير عليه, وتشمل هذه المرحلة دراسة المشكلة قيد البحث وتحديد اطار البحث وهذا يتطلب من الباحث الاطلاع على البحوث والدراسات السابقة والتي تتعلق بهذه المشكلة والجوانب التي تطرق اليها الباحثون السابقون والمشاكل التي واجهتهم ونقاط الضعف والقوة في المقترحات التي تطرقوا اليها .

2- مرحلة جمع البيانات.

بعد ان ينتهي الباحث من تحديد مشكلة الدراسة يبدأ في عملية جمع البيانات الخاصة بالدراسة وتحديد نوع البيانات التي يمكن الاستفادة منها في دراسته وفضل الطرق التي يمكن استخدامها للحصول على هذه البيانات. كذلك يجب تحديد حجم العينة الذي يتناسب مع حجم المجتمع قيد الدراسة.

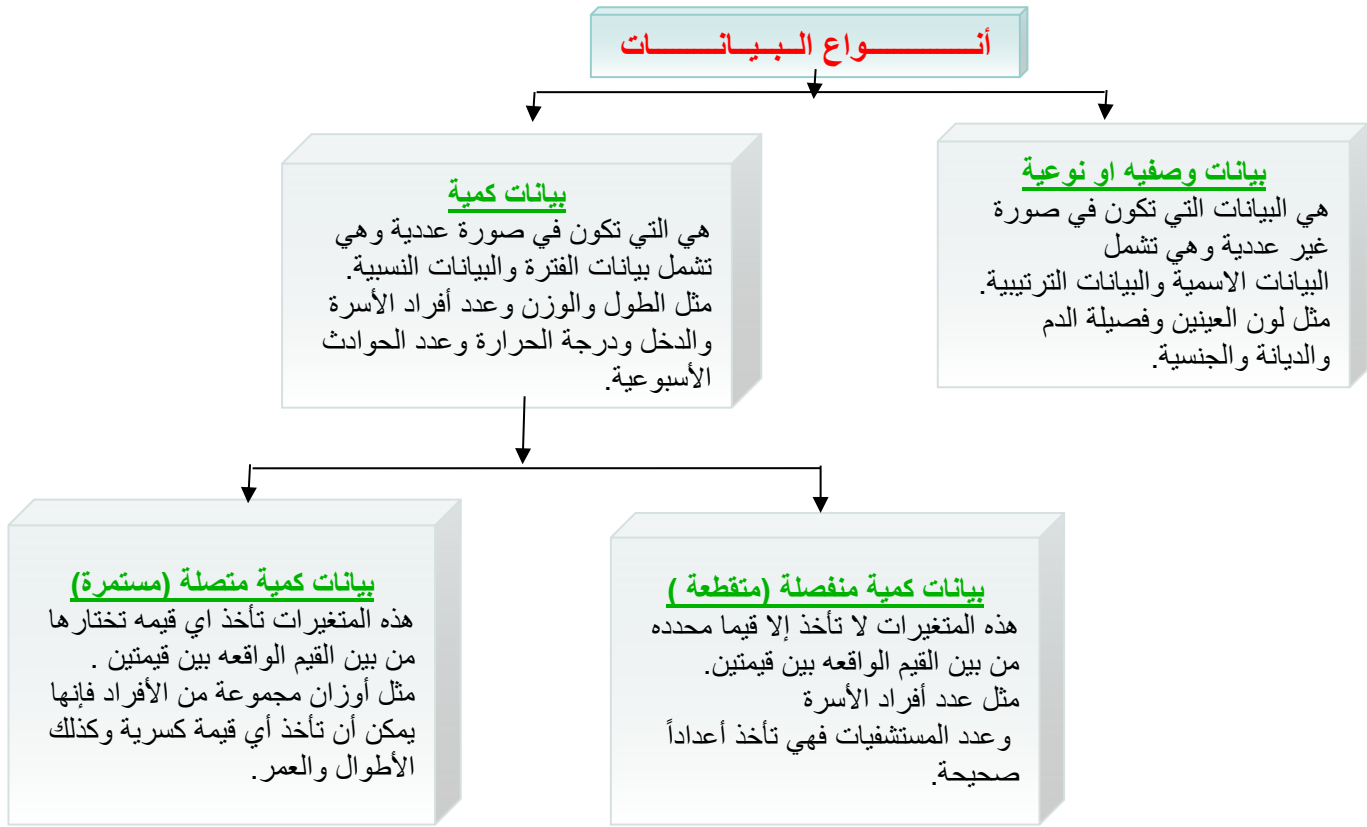
3- مراجعة وتدقيق البيانات

بعد اتمام عملية جمع البيانات يتوجب على الباحث مراجعة وتدقيق البيانات لغرض التأكد من صحتها وملائمتها لمتطلبات الدراسة, فعلى سبيل المثال لو كانت الدراسة هي دراسة استقصائية تعتمد على استمارة استبيان خاصة اعدت من قبل الباحث, عندئذ يتوجب عليه مراجعة وتدقيق الاستمارات الاحصائية التي تم جمعها واستبعاد ما هو غير متكامل او واضح او دقيق وعزل الاستمارات التي يعتقد الباحث انها غير صالحة للدراسة لعدم مطابقتها معايير الاستبانة.

4- تصنيف وتبويب البيانات

بعد اجراء عملية مراجعة وتدقيق البيانات تبدأ عملية تصنيف وتبويب هذه البيانات. يتم التصنيف على اساس الظواهر التي جمعت عنها البيانات فقد يكون التصنيف على اساس الفترة الزمنية او الرقعة الجغرافية او على اساس الجنس او اللون او الطول وغيرها وبعد اتمام عملية تصنيف البيانات نبدأ بعملية التبويب ويقصد بالتبويب هو عملية تفريغ البيانات المصنفة في جداول خاصة بحيث ان كل جزء من البيانات المصنفة عن الظاهرة المعينة يعود الى مستوى معين لتلك الظاهرة. ان الهدف من عملية التبويب هو ابراز البيانات

وتوضيحا في اضيق حيز ممكن كي نتمكن من تكوين فكرة عنها . ويختلف اسلوب تبويب البيانات تبعا لطبيعتها ومن هذه الاساليب:



جمع وتصنيف البيانات

ان عملية جمع البيانات تتم من خلال مجموعة من المصادر وبعتماد عدة اساليب لجمع البيانات، وبالتالي فان البيانات التي يتم الحصول عليها تسمى بالبيانات الخام او البيانات الاولية elementary data وهذه النوعية من البيانات لا يمكن الاستفادة منها بشكلها الحالي لذلك يجب تصنيفها وتبويبها باستخدام الاساليب العلمية .

مصادر جمع البيانات

1- المصادر التاريخية Historical Sources

وهي البيانات المحفوظة لدى اجهزة ومؤسسات ودوائر الدولة وكذلك لدى الشركات والهيئات ذات العلاقة بموضوع الدراسة مثل المصارف والمستشفيات ودوائر النفوس والوزارات ... الخ .

2- مصادر الميدان Field Sources

وهي البيانات والمعلومات التي يمكن الحصول عليها من مصادرها الاصلية بطريقة المقابلة او المراسلة او أي طريقة اتصال اخرى مثل التعدادات السكانية او الاستفتاءات او استمارات الاستبيان.

اساليب جمع البيانات: ان اختيار مصدر واسلوب جمع البيانات يعتمد بشكل كبير على طبيعة الدراسة والنتائج المتوخاة منها وكذلك تعتمد على الامكانية المتاحة والوقت المتيسر وبشكل عام هناك اسلوبين لجمع البيانات هما:

1- اسلوب التسجيل الشامل Census

نقصد بالتسجيل الشامل هو جمع البيانات عن كافة افراد مجتمع الدراسة بلا استثناء والذي يجب ان يكون مجتمعاً محدوداً كما هو الحال في عملية التعداد السكاني العام او مجموعة البنوك الحكومية. ويتميز أسلوب الحصر الشامل بالشمول وعدم التحيز، ودقة النتائج ولكن هذا الأسلوب يتطلب الكثير من الجهد والوقت والتكلفة العالية.

2- اسلوب العينات Samples

يعتبر هذا الاسلوب من الاساليب الشائعة في الدراسات والبحوث ويعتمد على جمع البيانات والمعلومات من جزء من مجتمع الدراسة ويسمى هذا الجزء بالعيينة (Sample) حيث يتم دراسة هذه العينة ويتم تعميم النتائج التي يتوصل لها على كافة مجتمع الدراسة , وهنا يجب الاشارة الى اختيار هذه العينة يجب ان يكون بشكل علمي وان تكون ممثلة للمجتمع المسحوبة منه افضل تمثيل . ويمتاز هذا الاسلوب بوقت وجهد وتكلفة اقل مما هو عليه في اسلوب التسجيل الشامل .

تصنيف وتبويب البيانات Classification of data

بعد عملية جمع البيانات تبدأ عملية التصنيف على اساس الظواهر التي جمعت منها هذه البيانات , حيث يتم فرز بيانات كل ظاهرة على هيئة مجموعة فقد يكون التصنيف على اساس صفة العمر او الجنس او اللون , الوزن , الطول ... الخ .

تبويب البيانات :

بعد اتمام عملية تصنيف البيانات تبدأ عملية التبويب والتي يقصد بها عملية تفريغ البيانات المصنفة في جداول بحيث ان كل جزء من البيانات المصنفة يعود الى مستوى معين لتلك الظاهرة, والهدف من هذه العملية هو ابراز البيانات وتوضيحها ووضعها بشكل يسهل عملية فهمها وتكوين فكرة عنها . واهم انواع التبويب هي

1- التبويب الزمني: هو عملية تجميع البيانات المصنفة وترتيبها في جداول على اساس وحدات زمنية معينة مثل السنة او الفصل او الشهر او اليوم ... الخ .

- 2- **التبويب الجغرافي** : هو عملية تجميع البيانات المصنفة وترتيبها في جداول على اساس وحدات جغرافية معينة كالمدن والنواحي والاقضية الخ.
- 3- **التبويب الكمي** : هو عملية تجميع البيانات المصنفة وترتيبها في جداول على اساس وحدات كمية معينة مثل وحدات الطول او الوزن او المسافة .
- 4- **التبويب على اساس صفة معينة** : هو عملية تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل مجموعة منها تشترك بصفة معينة كالجنس او الحالة الاجتماعية, العمر , القومية.

الأخطاء الشائعة في جمع البيانات

في بعض الاحيان يقع الباحث بمجموعة من الاخطاء لدى جمعه البيانات والمعلومات وهذه الاخطاء تحدث نتيجة سوء استخدام الطريقة الاحصائية . وهناك نوعين شائعين من هذه الاخطاء .

- 1- **خطأ التحيز Bias error** : تحدث هذه الاخطاء نتيجة جمع البيانات والمعلومات من غير مصادرها الاصلية . ويقصد بالتحيز انحراف متوسط جميع التقديرات الممكنة عن قيمتها الحقيقية، وسبب حصول هذا الخطأ هو الإخلال بتطبيق أسس نظرية العينات وبالتالي نظرية الاحتمالات التي تنص على ضرورة إعطاء جميع مفردات المجتمع نفس فرصة أو احتمال الظهور في العينة
- 2- **خطأ الصدفة Chance error** : ينتج هذا الخطأ نتيجة إتباع أسلوب العينات الذي يقضي بدراسة جزء من المجتمع وليس المجتمع بأكمله وهذا الخطأ ناتج على ان الصدفة قد حالفت مفردة معينة في العينة ولن تحالف مفردة أخرى، أو قد يحصل خطأ الصدفة نتيجة عدم الدقة في تدوين المعلومات المطلوبة من قبل أفراد العينة الناتج أصلاً عن اختلاف العدادين أو اختلاف الدافع الشخصي للإجابة على الأسئلة أو الحالة النفسية لإفراد المجتمع... الخ علماً بأن هذا الخطأ موجود في كل من أسلوب الحصر الشامل وأسلوب العينة.

المتغيرات العشوائية Random Variables

- المتغير العشوائي هو دالة ذات قيمة حقيقية معرفة في فضاء العينة لأي تجربة عشوائية وعادة ما يرمز لها بالأحرف الانكليزية الكبيرة مثل X, Y, Z وتقسّم المتغيرات العشوائية الى قسمين هما:
- 1- **متغيرات وصفية او نوعية: Qualitative variables** وهي المتغيرات التي تكون مفرداتها غير قابلة للقياس بالأرقام بل تكون على شكل صفات مثل لون العين او تقديرات الطلاب .
- 2- **متغيرات كمية: Quantitative variables** وهي المتغيرات التي تكون مفرداتها قابلة للقياس بالأرقام مثل اعداد طلبة الجامعات العراقية، الطول، العمر... الخ وتقسّم الى قسمين هما:
- أ- **متغيرات منقطعة او منفصلة: Discrete variables** وهي المفردات التي تأخذ قيما متميزة عن بعضها وتكون قابلة للعد والحساب ولا تتقبل الكسور مثل عدد الطلاب في كلية معينة او عدد العوائل في منطقة معينة.
- ب- **متغيرات مستمرة: continuous variable** وهي المفردات التي تأخذ مدى معين او مجال معين وتكون غير قابلة للعد والحساب وتتقبل الكسور مثل اعمار الافراد او اطوالهم او درجاتهم.

كذلك يمكن تصنيف المتغيرات من حيث العلاقة الى نوعين هما:

المتغير المستقل: Independent variable والمتغير المعتمد dependent variable فعند وجود علاقة رياضية بين متغيرين فان المتغير المستقل هو المتغير الذي لا يخضع تغيره في تلك العلاقة لتغير قيم المتغير الاخر في حين ان المتغير المعتمد تتاثر قيمه بشكل مباشر عند أي تغير يطرأ على المتغير المستقل. فعند دراسة العلاقة بين نسبة كريات الدم البيضاء ومناعة الجسم, نجد ان المناعة (متغير معتمد) تزداد بزيادة نسبة الكريات البيضاء في الدم (متغير مستقل).

موازين قياس المتغيرات Measurement scale of variables

هي القياسات والملاحظات التي يتم رصدها من افراد العينه. وهناك أربع مجموعات من لبيانات:

1- البيانات الاسمية Nominal Scale: وهي تتضمن المتغيرات التي يتم تصنيف فئاتها إلى فئات

اسمية. مثل: الحالة الزوجية تقسم إلى فئات وهي: متزوج – أعزب – مطلق – أرمل.

2- البيانات الترتيبية Ordinal Scale: يتضمن التغيرات التي يتم تصنيف فئاتها إلى وحدات مرتبة

من أسفل إلى أعلى أو العكس. مثل: الحالة التعليمية تقسم فئاتها إلى: أمي – ابتدائي – متوسط – ثانوي – جامعي – دراسات عليا.

3- بيانات الفترات Interval scale: يتضمن المتغيرات التي يتم تصنيف فئاتها إلى وحدات مرتبة

ومحددة رقمياً من أسفل إلى أعلى أو العكس وبيانات الفترة تمتلك خاصية المسافات المتساوية التي تفصل بين فئة وأخرى مجاورة لها وإن الصفر في هذا المقياس يعتبر نسبياً وليس مطلقاً. مثل: مقياس الذكاء – مقياس الحرارة – درجات جميع الاختبارات التحصيلية.

4- البيانات النسبية Ratio scale: يتضمن كل خصائص البيانات السابقة إضافة إلى إمكانية

التعبير عن المستويات المختلفة للمتغير بعلاقات نسب والصفر في هذه البيانات هو حقيقي. مثل: العمر – الدخل.

التوزيعات التكرارية Frequency distribution

عادة ما يكون عدد البيانات كبيراً بحيث يصعب الاستفادة منها بشكل مباشر. في مثل هذه الحالات يكون من الضروري تلخيص البيانات وتصنيفها وتبويبها بشكل جداول لتعطينا صورة أكثر وضوحاً عن البيانات المدروسة. وتعتبر التوزيعات التكرارية إحدى الطرق المستخدمة في تلخيص البيانات سواءً كانت هذه البيانات وصفية أم كمية.

التوزيع التكراري في أبسط أشكاله يتكون من حصر لجميع القيم التي يمكن أن يأخذها المتغير أو الظاهرة وعدد مرات تكرار كل قيمة من قيم هذه القيم. وتوضع هذه المعلومات في جدول يسمى الجدول التكراري لأنه يبين توزيع القيم و تكراراتها وهناك نوعين من الجداول التكرارية هما :

أولاً: جدول التوزيع التكراري البسيط:

وهو الجدول الذي توزع فيه البيانات حسب صفة واحدة ويتألف عادة من عمودين العمود الأول للفئات والعمود الثاني للتكرارات المقابلة لكل فئة .

ثانياً: جدول التوزيع التكراري المزدوج أو المركب:

وهو الجدول الذي توزع فيه البيانات حسب صفتين أو أكثر في نفس الوقت, فالجدول المزدوج مثلاً يتألف من صفوف واعمدة عناوين الصفوف تمثل تقسيمات أو فئات الصفة الأولى وعناوين الاعمدة تمثل تقسيمات أو فئات الصفة الأخرى والخلايا الداخلية تمثل التكرارات المشتركة لكلا الصفتين .

كيفية عمل جدول التوزيع التكراري البسيط في حالة البيانات الوصفية

- 1- تحديد نوع البيانات : بيانات وصفية
- 2- ننشئ جدولاً مكوناً من ثلاثة أعمدة على النحو التالي:
 - أ- العمود الأول: عمود الفئات
 - ب- العمود الثاني: عمود العلامات.
 - ج- العمود الثالث: عمود التكرار ” عدد الحالات“
- 3- نكتب الفئات في العمود الأول.
- 4- في العمود الثاني نرصد البيانات واحدة تلو أخرى وذلك بوضع خط مائل (/) لكل نمط أمام الفئة المطابقة لها وتسهيلاً لعملية العد نضع الخط الخامسة على صورة خط أفقي يقطع الخطوط الأربعة السابقة فنحصل على ما يسمى بالحزمة.
- 5- بعد الفراغ من تعبئة العمود الثاني نقوم بجمع علامات كل فئة ونضع الناتج في عمود عدد الحالات المقابل لكل فئة (عمود التكرار) ، وهذا الجدول يسمى جدول التفرغ .

مثال: البيانات التالية توضح مكان الإقامة الأصلية لعينة مكونة من 50 طالباً:

مدينة كبيرة	مدينة كبيرة	قرية	عشوائيات	مدينة كبيرة
مدينة صغيرة	قرية	مدينة متوسطة	مدينة كبيرة	قرية
قرية	مدينة متوسطة	مدينة كبيرة	مدينة كبيرة	مدينة صغيرة
مدينة متوسطة	عشوائيات	مدينة صغيرة	عشوائيات	مدينة متوسطة
مدينة صغيرة	مدينة متوسطة	قرية	مدينة متوسطة	مدينة كبيرة
قرية	مدينة متوسطة	مدينة كبيرة	قرية	مدينة متوسطة
مدينة كبيرة	قرية	مدينة متوسطة	مدينة كبيرة	مدينة صغيرة
مدينة متوسطة	مدينة كبيرة	مدينة متوسطة	قرية	مدينة متوسطة
مدينة متوسطة	مدينة متوسطة	عشوائيات	مدينة متوسطة	مدينة كبيرة
عشوائيات	مدينة كبيرة	قرية	مدينة متوسطة	قرية

المطلوب: تنظيم هذه البيانات في جدول توزيع تكراري
الحل:

مكان الإقامة	العلامات	(التكرار) عدد الطلاب
عشوائيات	###	5
قرية	### -##	11
مدينة صغيرة	###	5
مدينة متوسطة	### ## -##	16
مدينة كبيرة	## -##	13
المجموع		50

تنبيه: لابد أن يكون مجموع التكرارات مساوياً لعدد أفراد العينة.

6- نحذف العمود الثاني "عمود العلامات" ونبقي على العمود الأول والثالث لنحصل على جدول التوزيع التكراري البسيط:

مكان الإقامة	عدد الطلاب
عشوائيات	5
قرية	11
مدينة صغيرة	5
مدينة متوسطة	16
مدينة كبيرة	13
المجموع	50

عمل جدول التوزيع التكراري البسيط في حالة البيانات الكمية

لعمل جدول التوزيع التكراري للبيانات الكمية سواء كانت منفصلة او متصلة يجب استخراج وحساب المؤشرات التالية :

- 1- المدى الكلي Total Range
- 2- عدد الفئات Number of classes
- 3- طول الفئة Length of class
- 4- الحد الادنى والحد الاعلى للفئة Lower and Upper bound of class
- 5- مركز الفئة Center of class
- 6- تكرار الفئة Class Frequency

ويتم حساب المؤشرات أعلاه بالشكل الاتي:

1- المدى الكلي: نرمل له بالرمز $T.R$

وهو الفرق بين اكبر قيمة واصغر قيمة في مجموعة البيانات وحسب الصيغة التالية:
حيث ان:

$$T.R = x_L - x_S$$

x_L : اكبر قيمة في البيانات

x_S : اصغر قيمة في البيانات

2- عدد الفئات: نرمل له بالرمز m

ونقصد به عدد المجاميع التي يتالف منها التوزيع التكراري وهناك صيغ تقريبية يمكن من خلالها تحديد عدد الفئات اهمها صيغة يول (yell) :
حيث ان n عدد المفردات (حجم العينة)

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

ملاحظة: عند التطبيق يقرب الناتج الى أقرب عدد صحيح (لماذا)

3- طول الفئة: نرمل له بالرمز L

ويمثل مقدار سعة الفئة أي مقدار المسافة بين الحد الادنى والحد الاعلى للفئة ويحسب من الصيغة التالية

4- الحد الادنى (L.L) والحد الاعلى (U.L):

لكل فئة حد ادنى والذي يمثل بداية الفئة وحد اعلى والذي يمثل نهاية الفئة والجدول التالي يوضح الحدود الدنيا والحدود العليا للفئات في حالة تساوي اطوال الفئات .
حيث:

x_S : اصغر قيمة في البيانات m : عدد الفئات L : طول الفئة

5- مركز الفئة: نرمز له بالرمز C.

مركز الفئة هو القيمة التي تتوسط المسافة بين الحد الأدنى والحد الأعلى للفئة ويحسب من الصيغة التالية

$$C_i = \frac{UL + LL}{2}$$

حيث ان UL (Upper Limit) هو الحد الأعلى للفئة
حيث ان LL (Lower Limit) هو الحد الأدنى للفئة

6- تكرار الفئة: يرمز له بالرمز f

يمثل تكرار الفئة هو عدد القيم المحصورة بين حدي الفئة وان مجموع التكرارات يجب ان يساوي حجم العينة n اي ان

$$f_1 + f_2 + f_3 + \dots + f_m = n$$

مثال:

في عينة من 20 قرية تم اختيارها من إحدى المدن كان عدد المصابين بمرض معين في كل قرية كما يلي:

المطلوب: وضع هذه البيانات في صورة جدول تكراري بسيط .

0	3	8	5	3
2	2	9	1	2
7	6	1	1	4
4	2	3	0	1

الحل:

لتنظيم البيانات في جدول توزيع تكراري نتبع الخطوات التالية :

- 1- تحديد نوع البيانات : بيانات كمية منفصلة.
- 2- ننشئ جدولاً مكوناً من ثلاثة أعمدة على النحو التالي:
أ- العمود الأول: عمود الفئات " عدد المصابين " .
ب- العمود الثاني: عمود العلامات .
ج- العمود الثالث: عمود التكرار " عدد القرى "
- 3- نستخرج المدى الكلي من القانون التالي.

$$T.R = x_L - x_S \\ = 9 - 0 = 9$$

4- نستخرج عدد الفئات من القانون التالي

$$m = 2.5 \sqrt[4]{n} \\ = 2.5 \sqrt[4]{20} \cong 5 \\ L = \frac{T.R}{m} \\ = \frac{9}{5} \cong 2$$

5- نستخرج طول الفئة

سؤال (لماذا تم تقريب عدد الفئات وطول الفئة الى اقرب عدد صحيح) ؟

6- نقوم بعمل الفئات ووضعها في الجدول حيث ان عدد الفئات هو (5) وأصغر قيمة ($X_s=0$) وطول الفئة ($L=2$)

الحد الأدنى للفئة الأولى = أصغر رقم من البيانات او اقل من منه بقليل = 0

الحد الأعلى لأي فئة $UL = LL + L - 1$

الحد الأعلى للفئة الأولى $UL1 = 0 + 2 - 1 = 1$

وبذلك تكون الفئة الأولى (0-1) وبنفس الطريقة نستمر لبقية الفئات

7- في العمود الثاني نرصد بيانات عدد المصابين واحدة تلو أخرى وذلك بوضع خط مائل (/) لكل عدد أمام الفئة المطابقة لها، وتسهيلا لعملية العد نضع العلامة الخامسة على صورة خط أفقي يقطع الخطوط الأربعة السابقة فنحصل على ما يسمى بالحزمة.

8- بعد الانتهاء من تفرغ البيانات في عمود التفرغ نقوم بجمع علامات كل فئة ونضع الناتج في عمود عدد القرى المقابل لكل فئة (عمود التكرار)، وبذلك تم عمل الجدول التكراري.

التكرارات (عدد الأسر)	الفئات (عدد الأطفال)
5	0 - 1
8	2 - 3
3	4 - 5
2	6 - 7
2	8 - 9
20	المجموع

مثال(3):

البيانات التالية تمثل عدد الراقين في إحدى المستشفيات حسب الايام لمدة شهر:

37	15	7	17	34
41	23	26	46	24
57	44	37	27	53
22	28	44	28	20
32	34	60	52	39
32	30	5	31	40

المطلوب تصنيف هذه البيانات في جدول توزيع تكراري

الحل:

$$\begin{aligned} 1) \quad T.R &= x_L - x_S \\ &= 60 - 5 = 55 \\ 2) \quad m &= 2.5 \sqrt[4]{n} \\ &= 2.5 \sqrt[4]{30} \cong 6 \\ 3) \quad L &= \frac{T.R}{m} \\ &= \frac{55}{6} = 9.2 \cong 10 \end{aligned}$$

4- تكوين الفئات:

الحد الأدنى للفئة الأولى = أصغر رقم من البيانات أو أقل منه بقليل = 5
الحد الأدنى للفئة الثانية = الحد الأدنى للفئة الأولى + طول الفئة
الحد الأدنى للفئة الثالثة = الحد الأدنى للفئة الثانية + طول الفئة

ونستمر بنفس الطريقة لإيجاد الحدود الدنيا لبقية الفئات

ملاحظة: في البيانات من النوع المستمر نستخدم الجداول المفتوحة من الأعلى (الفئات تحوي الحدود الدنيا فقط ما عدا الفئة الأخيرة يتم غلقها بالحد الأعلى للفئة)

الحد الأعلى لأي فئة = الحد الأدنى للفئة + طول الفئة (للبيانات من النوع المستمر)

5- جدول التفرغ

الفئات	العلامات	عدد الراقدين (التكرار)
5 -	//	2
15 -	/ ###	6
25 -	### ###	10
35 -	// ###	7
45 -	///	3
55- 65	//	2
المجموع		30

6- يمكن ان يضاف إلى هذا الجدول عمود "مركز الفئة"

عند الحصول على مركز الفئة الأولى يمكن الحصول على مراكز الفئات التالية بإضافة طول الفئة إلى مركز الفئة السابقة لها.

(طول الفئة في هذا المثال = 10)

الفئات	عدد الراقدين (التكرار)	مركز الفئة
5-	2	$(5+15)/2=10$
15-	6	$10 + 10 = 20$
25-	10	$20 + 10 = 30$
35-	7	40
45-	3	50
55-65	2	60
المجموع	30	

سؤال واجب رقم (1)

البيانات التالية تمثل عدد الزيارات الميدانية التي قام بها (20) مركز صحي في مدينة البصرة
 $X_i: 65, 56, 88, 83, 75, 60, 70, 65, 55, 67, 45, 65, 74, 72, 62, 48, 69, 65, 49,$
 49

المطلوب تمثيل هذه البيانات في جدول توزيع تكراري

سؤال واجب رقم (2)

البيانات التالية تمثل اوزان (16) طالبا في قسم الإدارة الصحية
 $X_i: 50, 82.5, 75, 60.5, 76.6, 80, 74, 66.5, 78.5, 72.5, 84, 71.5, 70.5, 70, 72,$
 58

المطلوب تمثيل هذه البيانات في جدول توزيع تكراري.

جدول التوزيع التكراري المزدوج (Double Frequency Distribution Table):

هو الجدول الذي يضم (أو تفرغ فيه) البيانات لظاهرتين أو متغيرين في نفس الوقت. ويتكون من الصفوف والتي تمثل فئات أحد المتغيرين والأعمدة والتي تمثل فئات المتغير الآخر. إما المربعات أو الخانات والتي تقابل الصفوف والأعمدة فأنها تحتوي على المفردات أو التكرارات المشتركة في فئات ومجاميع كلا الصفتين.

مثال:- البيانات التالية تمثل عينة لقياسات الطول والوزن لعشرين من المرضى الراقدين في إحدى المستشفيات : المطلوب تفرغ البيانات في جدول توزيع تكراري مزدوج

تسلسل المريض	1	2	3	4	5	6	7	8	9	10
الطول	141	149.8	150	180	158	157.6	159.5	160	166	172
الوزن	40	57.7	59.9	89	62	61.5	60.2	67.9	70.3	60

تسلسل المريض	11	12	13	14	15	16	17	18	19	20
الطول	180	177	163.3	157.5	179.8	185	173.3	172	170	169.9
الوزن	87.7	89.8	75.8	60	89	77	78	79.9	85	64

الحل: نستخرج اطوال وعدد الفئات لكل صفة من الصفات ثم نقوم بعمل الجدول التكراري المزدوج بوضع فئات احد الصفات في الاعمدة وفئات الصفة الثانية في الصفوف وأخيرا نفرغ البيانات في الجدول. نفرض ان صفة الطول (X) و صفة الوزن (Y)

$$TR_x = 180 - 141 = 39$$

$$M_x = 5$$

$$L_x = 39 / 5 \sim 8$$

$$Tr_y = 89 - 40 \sim 50$$

$$M_y = 5$$

$$L_y = 50 / 5 = 10$$

X الطول Y وزن	(140 -)	(148 -)	(156 -)	(164 -)	(172 - 180)	المجموع
40 -	1					1
50 -	1	1				2
60 -		4	2	1		7
70 -			2	2		4
80 - 90				3	3	6
المجموع	2	5	4	6	3	20

سؤال واجب (3):-

البيانات التالية تمثل الطول (X) بالسنتيمتر والوزن (y) بالكيلو غرام لعدد من طلبة قسم الإدارة الصحية، والمطلوب تفرغ هذه البيانات في جدول توزيع تكراري مزدوج.

(X) : 152, 162, 155, 144, 174, 170, 165, 160, 167, 159, 152, 164, 166, 178, 149, 175.

(y) : 62, 47, 65, 40, 72, 69, 49, 56, 65, 46, 56, 65, 49, 89, 60, 75

التكرارات المتجمعة الصاعدة والنازلة

جدول التكرار المتجمع (Cumulative Frequency Table):

وهو الجدول الذي تتجمع فيه التكرارات اما تصاعديا او تنازليا، وهو على نوعين هما:-

1 . جدول التكرار المتجمع الصاعد (Upward (Less than) Cumulative Frequency Table):

وهو الجدول الذي تتجمع فيه التكرارات ابتداءً من الفئة الدنيا (الفئة الأولى) إلى أن نصل إلى التكرار الكلي عند الفئة العليا (الفئة الأخيرة)، ويرمز له بالرمز F_i

2 . جدول التكرار المتجمع النازل (Downward (More than) Cumulative Frequency Table):

وهو الجدول الذي تتجمع فيه التكرارات ابتداءً من الفئة العليا (الأخيرة) إلى أن نصل إلى التكرار الكلي عند الفئة الدنيا (الأولى) ويرمز له بالرمز F^{\sim} .

مثال: لبيانات الجدول التالي المطلوب عمل:-

1 . جدول تكراري متجمع صاعد.

2 . جدول تكراري متجمع نازل.

فئات الطول	Fi
(162 -)	3
(168 -)	4
(174 -)	4
(180 -)	3
(186 -192)	1
المجموع	15

الحل:

1- التكرار المتجمع الصاعد

فئات الطول Classes	التكرارات fi	الحدود العليا للفئات Less than	التكرار المتجمع الصاعد Fi
162 -	3	less than 168	3
168 -	4	less than 174	7
174 -	4	less than 180	11
180 -	3	less than 186	14
186 - 192	1	less than 192	15

1- التكرار المتجمع النازل

فئات الطول Classes	التكرارات fi	الحدود الدنيا للفئات More than	التكرار المتجمع النازل Fi`
162 -	3	162 and More than	15
168 -	4	168 and More than	12
174 -	4	174 and More than	8
180 -	3	180 and More than	4
186 - 192	1	186 and More than	1

سؤال واجب (4):-

البيانات التالية تمثل الوزن (X) بالكيلو غرام لمجموعة من الطلبة والمطلوب:-

1 . تفرغ هذه البيانات في جدول توزيع تكراري بسيط.

2 . للجدول التكراري الناتج اعمل جدول تكراري متجمع صاعد وآخر نازل.

الوزن (X) : 75, 60, 89, 83, 65, 56, 46, 65, 49, 48, 69, 62, 72, 45, 56, 74.

العرض البياني للبيانات غير المبوبة

التمثيل البياني:

في الدروس السابقة درسنا كيفية تمثيل البيانات باستخدام الجداول التكرارية، في هذه المحاضرة سوف نقوم بدراسة نوع اخر من تمثيل البيانات باستخدام الاشكال البيانية. ان الرسوم والاشكال الهندسية ما هي الا تعبير وتوضيح للبيانات بطريقة جذابة وسهلة وفعالة تساعد القارئ على فهم واستيعاب قيم الظاهرة ومقارنتها مع بعضها وكما تجدر الإشارة هنا الى إمكانية تمثيل البيانات المبوبة وغير المبوبة بيانيا:

أولا: تمثيل البيانات الوصفية (غير مبوبة في جداول تكرارية)

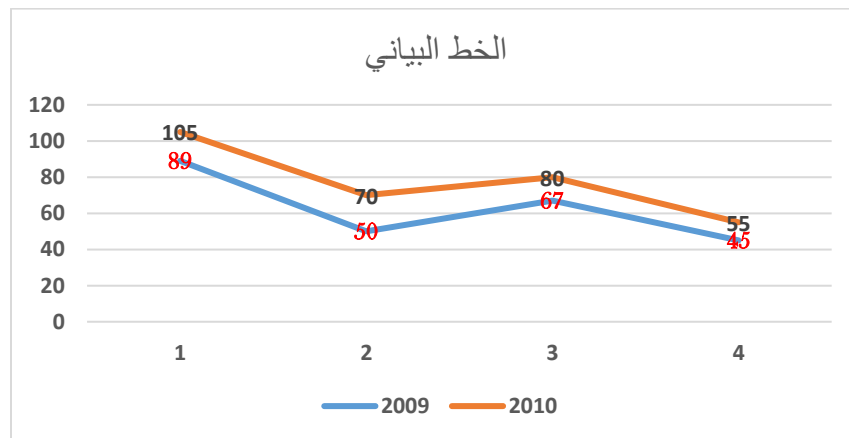
الخط البياني Line chart

عبارة عن شكل بياني يوضح التغيرات الحاصلة في ظاهرة معينة عبر فترة زمنية محددة ويستخدم عادة في اجراء المقارنة بين ظاهرتين او أكثر مقاسة بنفس وحدات القياس، مثل المقارنة في كميات الانتاج الشهري في مصنع معين خلال عدة اعوام.

مثال: البيانات الاتية تمثل عدد الاصدارات الفصلية من الكتب في احدى دور النشر، المطلوب تمثيل هذه البيانات باستخدام الخط البياني.

عدد الإصدارات				السنة
الفصل الرابع	الفصل الثالث	الفصل الثاني	الفصل الاول	
45	67	50	89	2009
40	75	70	105	2010

الحل:



2- الأشرطة البيانية البسيطة Bar chart

هي عبارة عن مجموعة من الأعمدة الرأسية أو الأفقية المتساوية القاعدة والتي تتناسب ارتفاعها مع البيانات التي تمثلها وتستخدم لإظهار التطور الذي يطرأ على ظاهرة ما على مدار عدة سنوات.

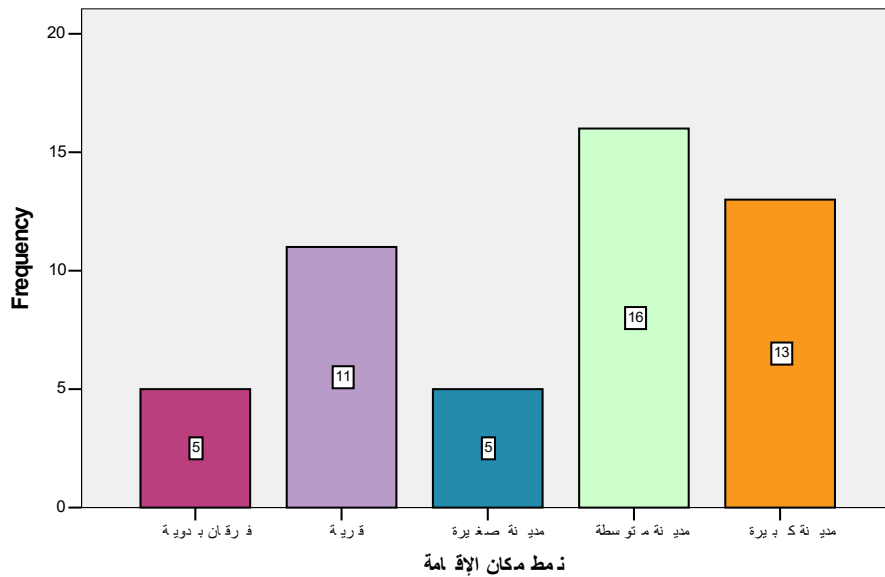
ويجب مراعاة أن يقسم المحور الرأسي بحيث يسمح بقياس الرسم بإظهار جميع قيم الظاهرة وكذلك يراعى أن تكون المسافات بين الأعمدة متساوية وقاعدات الأعمدة متساوية .

مثال: من الجدول التالي المطلوب تمثيل البيانات باستخدام الأشرطة (الأعمدة) البيانية:

عدد الطلاب	مكان الإقامة
5	عشوائيات
11	قرية
5	مدينة صغيرة
16	مدينة متوسطة
13	مدينة كبيرة
50	المجموع

الحل

نمط مكان الإقامة



3- الدائرة البيانية pie chart

تستخدم الدائرة إذا كانت بيانات الظاهرة موضوع الدراسة عبارة عن مجموع عام مقسم إلى أجزائه المختلفة، وتمثل المساحة الكلية للدائرة المجموع الكلي ثم تقسم الدائرة إلى قطاعات وتميز بينها بالتظليل أو الألوان على الرسم.

خطوات رسم الدائرة وتقسيمها إلى قطاعات:

1- يتم رسم الدائرة بعد اختيار نصف قطر مناسب لها.

2- تحسب زاوية كل قطاع باستخدام القانون:

$$\text{زاوية القطاع} = \frac{\text{قيمة القطاع}}{\text{المجموع الكلي للبيانات}} \times (360^\circ)$$

ويراعى أن يكون مجموع زوايا القطاعات المختلفة مساوياً للزاوية المركزية للدائرة (360 درجة).

3- تقسم الدائرة إلى قطاعاتها وذلك بتقسيم الزاوية المركزية للدائرة إلى زوايا القطاعات المختلفة.

وبالعودة إلى المثال السابق

جدول التوزيع التكراري

عدد الطلاب	مكان الإقامة
5	عشوائيات
11	قرية
5	مدينة صغيرة
16	مدينة متوسطة
13	مدينة كبيرة
50	المجموع

$$36^\circ = 360^\circ \times \frac{5}{50} = \text{زاوية القطاع الأول (عشوائيات)}$$

$$79.2^\circ = 360^\circ \times \frac{11}{50} = \text{زاوية القطاع الثاني (قرية)}$$

$$36^\circ = 360^\circ \times \frac{5}{50} = \text{زاوية القطاع الثالث (مدينة صغيرة)}$$

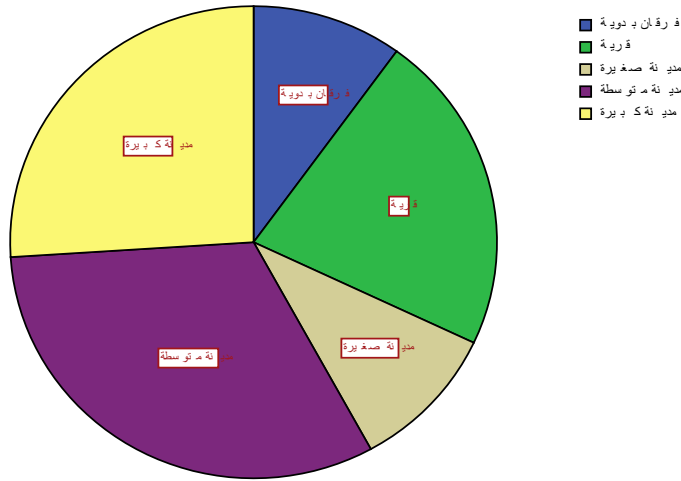
$$115.2^\circ = 360^\circ \times \frac{16}{50} = \text{زاوية القطاع الرابع (مدينة متوسطة)}$$

$$93.6^\circ = 360^\circ \times \frac{13}{50} = \text{زاوية القطاع الخامس (مدينة كبيرة)}$$

ملاحظة:

$$360 = 93.6 + 115.2 + 36 + 79.2 + 36 = \text{مجموع الزوايا}$$

نمط مكان الإقامة



4- المستطيل البياني Rectangular chart

هو عبارة عن شكل هندسي يستخدم في تمثيل بيانات ظاهرة معينة يمكن تجزئتها الى عدد من الاصناف القابلة للتجميع مثل تكاليف سلعة معينة, عدد الطلبة موزعين حسب المراحل الدراسية .
تتلخص فكرة رسم المستطيل البياني باختيار مستطيل ذو قاعدة مناسبة يمثل هذا المستطيل مجموع البيانات الكلية وبعد ذلك يتم تمثيل كل صنف من البيانات بمستطيل جزئي داخل المستطيل الكبير بحيث ان مجموع مساحات المستطيلات الجزئية تساوي مساحة المستطيل الكبير. ويتم تحديد مساحة المستطيلات الجزئية من القانون الاتي

$$\text{مساحة المستطيل الجزئي} = \frac{\text{بيانات القطاع}}{\text{المجموع الكلي للبيانات}} \times (\text{طول قاعدة المستطيل الكبير})$$

مثال: البيانات التالية تمثل اعداد طلبة قسم الاحصاء موزعين حسب المراحل الدراسية , المطلوب تمثيل البيانات باستخدام المستطيل البياني .

المرحلة الاولى : 120 المرحلة الثانية : 100 المرحلة الثالثة : 100 المرحلة الرابعة : 80

الحل: نفرض ان طول قاعدة المستطيل الكبير 10 سم وان مجموع الطلبة الكلي هو 120 + 100 +

$$400 = 80 + 100$$

نستخرج اطوال قواعد المستطيلات الجزئية

$$\text{طول قاعدة المستطيل للمرحلة الاولى} = \frac{120}{400} \times 10 = 3 \text{ c. m}$$

$$\text{طول قاعدة المستطيل للمرحلة الثانية} = \frac{100}{400} \times 10 = 2.5 \text{ c. m}$$

$$\text{طول قاعدة المستطيل للمرحلة الثالثة} = \frac{100}{400} \times 10 = 2.5 \text{ c. m}$$

$$\text{طول قاعدة المستطيل للمرحلة الرابعة} = \frac{80}{400} \times 10 = 2 \text{ c. m}$$

الأولى 3سم	الثانية 2.5سم	الثالثة 2.5 سم	الرابعة 2سم
------------	---------------	----------------	-------------

المستطيل البياني

العرض البياني للبيانات المبوبة

ثانياً: التمثيل البياني للبيانات المبوبة:

ان اهم الاشكال البيانية المستخدمة مع الجداول التكرارية هي:

- المدرج التكراري Histogram
- المضلع التكراري Frequency Polygon
- المنحنى التكراري Frequency Curve

1 - المدرج التكراري Histogram

وهو عبارة عن مجموعة من المستطيلات الراسية (ويمكن ان تكون عمودية) قواعدها تمثل اطوال الفئات وارتفاعها يمثل التكرار المقابل لكل فئة أي ان هناك علاقة بين الفئات والتكرارات، ولرسم المدرج التكراري نتبع الخطوات التالية.

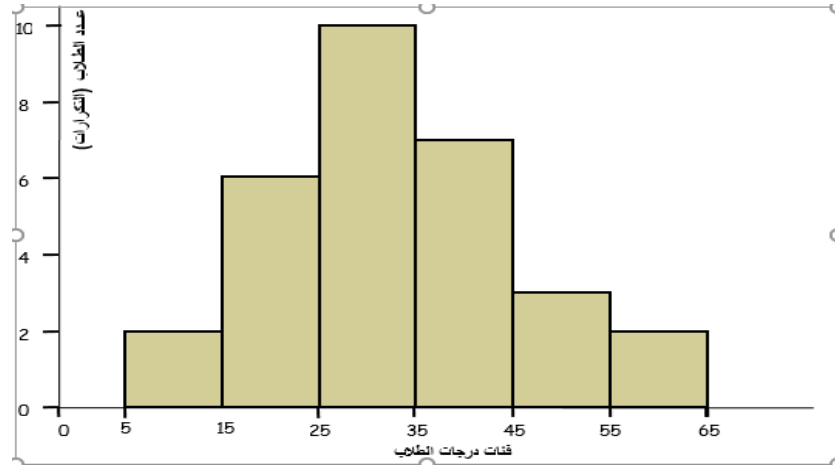
1-نقسم المحور الأفقي(السيني) إلى أجزاء مناظرة للفئات، أما المحور العمودي فيكون للتكرارات.

2-نرسم لكل فئة عموداً قاعدته هي طول الفئة وارتفاعه هو تكرار هذه الفئة.

مثال: مثل بيانات الجدول التكراري التالي باستخدام المدرج التكراري

فئات الدرجات	عدد الطلاب(التكرار)
5 -	2
15 -	6
25 -	10
35 -	7
45 -	3
55 -65	2
المجموع	30

الحل :



2- المضلع التكراري Frequency Polygon

هو عبارة عن خطوط مستقيمة متكسرة تصل بين مجموعة من النقاط، الاحداثي الاول لكل نقطة يمثل مركز الفئة والاحداثي الثاني يمثل التكرار المقابل لتلك الفئة، ويفضل غلق المضلع التكراري وذلك بفرض مراكز فئة وهمية قبل مركز الفئة الاولى وبعد مركز الفئة الاخيرة وان تكرر هاتين الفئتين الوهميتين مساوي للصفر.

لرسم المضلع التكراري: (علاقة بين مراكز الفئات والتكرارات) نتبع الخطوات التالية:

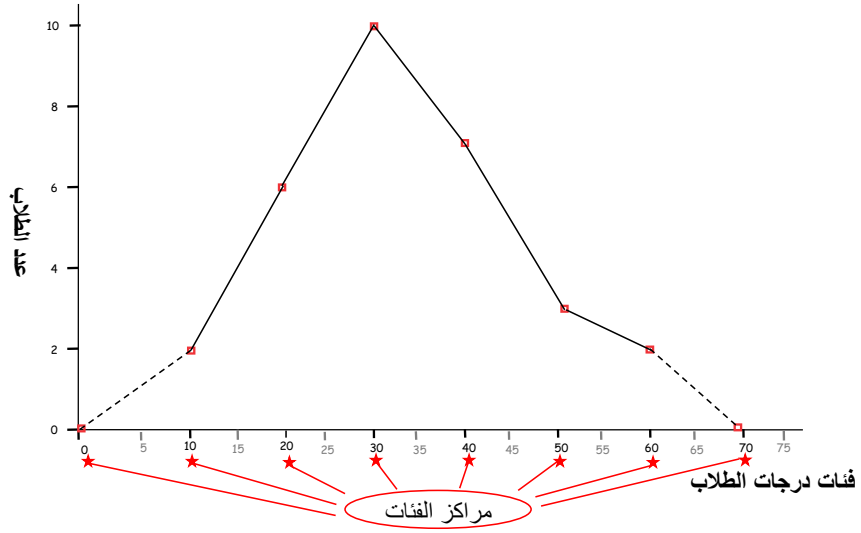
- 1- نرسم محورين كما هو الحال في المدرج التكراري.
- 2- نمثل البيانات بنقاط احداثيها هما مركز الفئة وتكرار كل فئة.
- 3- نضيف مركز فئة في البداية وفي النهاية ويكون تكراره بصفر وذلك لإغلاق المضلع والمنحنى.
- 4- إذا وصلنا بين هذه النقاط بخطوط مستقيمة نحصل على المضلع التكراري

مثال: بالرجوع الى بيانات المثال السابق (درجات الطلبة) المطلوب تمثيل البيانات باستخدام المضلع التكراري.

الحل: نستخرج اولاً مراكز الفئات

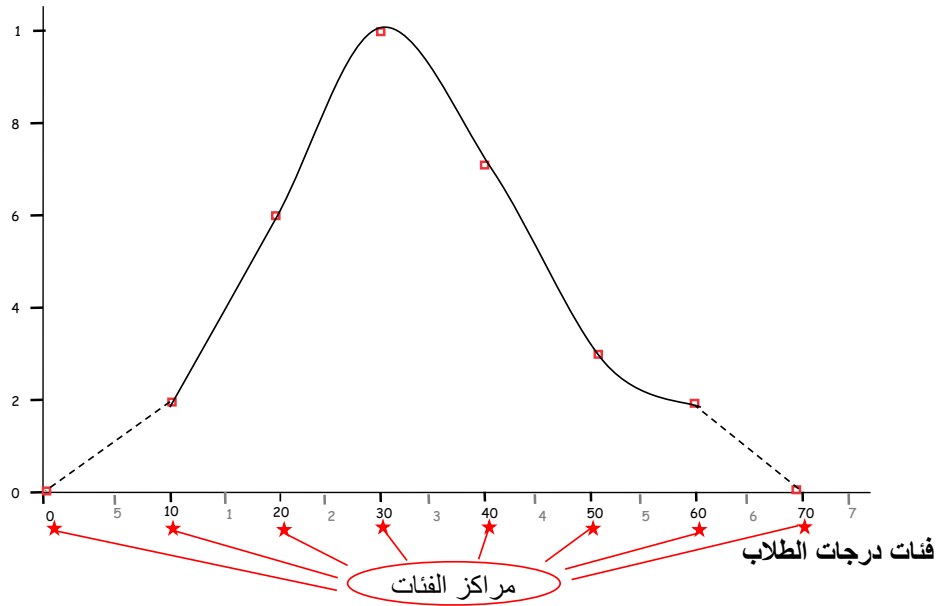
مركز الفئة	عدد الطلاب(التكرار)	فئات الدرجات
10	2	5 -
20	6	15 -
30	10	25 -
40	7	35 -
50	3	45 -
60	2	55 -65
#	30	المجموع

وبذلك يكون المضلع التكراري بالشكل التالي:



3- المنحنى التكراري Frequency Curve

وهو عبارة عن منحنى يمر بمعظم النقاط الواقعة على مراكز الفئات وارتفاعها يمثل تكرارات تلك الفئات ويفضل غلق المنحنى التكراري مع المحور الافقي كما هو الحال مع المضلع التكراري. ولرسم المنحنى التكراري للمثال السابق نقوم برسم منحنى يمر بمعظم هذه النقاط.



مقاييس النزعة المركزية

مقدمة

في كثير من النواحي التطبيقية يكون الباحث في حاجة إلى حساب بعض المؤشرات التي يمكن الاعتماد عليها في وصف الظاهرة وكذلك التعرف على مدى تجانس القيم التي يأخذها المتغير، وأيضاً ما إذا كان هناك قيم شاذة في البيانات. ان الاعتماد على العرض البياني او الجدولي للبيانات يكون غير كافي في مثل هذه الحالات، لذلك يجب علينا استخدام أساليب أخرى في وصف البيانات، من هذه الأساليب المهمة هي مقاييس النزعة المركزية. يتناول هذا الفصل، والذي يليه عرض بعض المقاييس الإحصائية التي يمكن من خلالها التعرف على خصائص الظاهرة محل البحث، وكذلك إمكانية مقارنة ظاهرتين أو أكثر، ومن أهم هذه المقاييس، مقاييس النزعة المركزية والتشتت.

مقاييس النزعة المركزية

تسمى مقاييس النزعة المركزية بمقاييس الموضع أو المتوسطات، وهي القيم التي تتمركز القيم من حولها ، ومن هذه المقاييس ، الوسط الحسابي ، والمنوال ، والوسيط ، والوسط الهندسي ، والوسط التوافقي ، والرباعيات ، والمئينات ، وفيما يلي عرض لأهم هذه المقاييس قبل الشروع بدراسة طرق حساب المؤشرات الاحصائية لابد لنا ان نعرف على اهم الرموز المستخدمة في حساب هذه المؤشرات.

أولاً : رمز الجمع \sum

يستخدم رمز الجمع \sum للتعبير عن جمع مجموعة من الاشياء وهو حرف اغريقي يقرأ **Summation**

فعلى فرض ان لدينا مجموعة من المفردات مثل $x_1, x_2, x_3, x_4, x_5, x_6$ فان :

$$\sum_{i=1}^6 x_i = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \quad ; \quad i = 1,2,3,4,5,6$$

حيث ان n هو حجم العينة او عدد المفردات وفي هذا المثال $n=6$

i . يمثل الدليل (subscript) لتسلسل العدد عند عملية الجمع فعندما $i=1$ هذا يعني المشاهدة x_1 ,

وعندما $i=2$ هذا يعني المشاهدة x_2 وهكذا نستمر الى المشاهدة الاخيرة x_n عندما $i= n$.

مثال :

على فرض ان لدينا قياسات عينة بحجم (3) وهي x_1, x_2, x_3 المطلوب :

1- حساب $\sum_{i=1}^n x_i$

2- حساب $\sum_{i=1}^n x_i^2$

3- حساب $(\sum_{i=1}^n x_i)^2$

4- حساب $\sum_{i=1}^n \frac{1}{x_i}$

5- حساب $\frac{1}{\sum_{i=1}^n x_i}$

الحل

$$1- \sum_{i=1}^n x_i = \sum_{i=1}^3 x_i \\ = x_1 + x_2 + x_3$$

$$2- \sum_{i=1}^n x_i^2 = x_1^2 + x_2^2 + x_3^2$$

$$3- (\sum_{i=1}^n x_i)^2 = (x_1 + x_2 + x_3)^2$$

$$4- \sum_{i=1}^n \frac{1}{x_i} = \frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3}$$

$$5- \frac{1}{\sum_{i=1}^n x_i} = \frac{1}{x_1 + x_2 + x_3}$$

بعض خصائص رمز المجموع

إذا كان حجم العينة n وان a كمية ثابتة فان

$$1- \text{ حاصل جمع الكمية الثابتة يساوي } n \text{ مضروباً في الكمية الثابتة } a \text{ أي ان } \sum_{i=1}^n a = n \cdot a$$

2- حاصل ضرب الثابت a بقيم عناصر السلسلة x_i يمثل حاصل ضرب الثابت a بمجموع قيم عناصر السلسلة أي

$$\sum_{i=1}^n a x_i = a \sum_{i=1}^n x_i \quad \text{ان :}$$

3- إذا كان x_i , y_i متغيرين فان

$$\sum_{i=1}^n (x_i \mp a) = a \sum_{i=1}^n x_i$$

$$\sum_{i=1}^n (x_i \mp y_i) = \sum_{i=1}^n x_i + \sum_{i=1}^n y_i$$

$$\sum_{i=1}^n x_i y_i = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

ثانياً: رمز الضرب π

يرمز الحرف اللاتيني $\prod_{i=1}^n$ والذي يقرأ (باي) الى حاصل ضرب عناصر سلسلة من الكميات ببعضها ابتداءً بالعنصر

الاول عندما $i=1$ وانتهاءً بالعنصر الاخير عندما $i=n$ وفيما يلي بعض الخواص المهمة واستخدامات رمز الضرب

$$1- \prod_{i=1}^n a = a^n$$

$$2- \prod_{i=1}^n a x_i = a^n \prod_{i=1}^n x_i$$

$$3- \prod_{i=1}^n a b x_i y_i = (a b)^n \prod_{i=1}^n x_i \prod_{i=1}^n y_i :$$

$$4- \prod_{i=1}^n \frac{1}{x_i} = \frac{1}{\prod_{i=1}^n x_i}$$

$$5- \log \prod_{i=1}^n x_i = \sum_{i=1}^n \log x_i$$

$$6- \prod_{i=1}^n \sqrt{x_i} = \sqrt{\prod_{i=1}^n x_i}$$

ان من اهم مقاييس النزعة المركزية التي سوف يتم دراستها هي :-

1- الوسط الحسابي Mean

2- الوسيط Median

3- المنوال Mode

سوف يتم دراسة هذه المقاييس في حالتين

أ- في حالة البيانات غير المبوبة

ب- في حالة البيانات المبوبة

1- الوسط الحسابي The mean

الوسط الحسابي لمجموعة من القيم هو القيمة الناتجة من حاصل قسمة مجموع تلك القيم على عددها ويرمز له

بالرمز \bar{X} اذا كان محسوباً من العينة وقرأ ((**X Bar**))

أ- حساب الوسط الحسابي من بيانات غير مبوبة .

اذا كان لدينا مجموعة من المشاهدات هي x_1, x_2, \dots, x_n , حيث ان n هو حجم العينة فان الوسط الحسابي لها هو

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

مثال 1: البيانات التالية تمثل عدد الراقدين اليومي في احدى المستشفيات في مدينة بغداد خلال خمس ايام

400 , 380 , 450 , 350 , 520

المطلوب إيجاد معدل عدد الراقدين خلال هذه الفترة .

الحل :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_5}{5}$$

$$= \frac{520 + 350 + 450 + 380 + 400}{5} = 420$$

أي ان معدل عدد الراقدين اليومي خلال الفترة هو 420 شخص يوميا

مثال 2 :

احسب متوسط درجات الطالب حيث كانت درجاته هي 70 , 55 , 80 , 70 , 75 , 50

الحل :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_6}{6}$$

$$= \frac{70 + 55 + 80 + 70 + 75 + 50}{6} = \frac{400}{6} = 66.66$$

ب- حساب الوسط الحسابي من بيانات مبوبة

اذا كان لدينا جدول توزيع تكراري عدد فئاته m وان x_1, x_2, \dots, x_m تمثل مركز الفئات وان f_1, f_2, \dots, f_m تمثل التكرارات المقابلة لمراكز الفئات فان الوسط الحسابي يحسب وفق الصيغة التالية:

$$\bar{x} = \frac{\sum_{i=1}^m f_i x_i}{\sum_{i=1}^m f_i} = \frac{f_1 x_1 + f_2 x_2 + \dots + f_m x_m}{f_1 + f_2 + \dots + f_m}$$

حيث ان :

m	عدد الفئات
x_i	مركز الفئة i
f_i	تكا، الفئة i

مثال 3 : من جدول التوزيع التكراري التالي احسب الوسط الحسابي

الفئات	31-40	41-50	51-60	61-70	71-80
f_i	1	2	5	15	25

الحل: نقوم باستخراج مراكز الفئات وثم تطبيق قانون الوسط الحسابي

الفئات classes	التكرار f_i	مركز الفئة x_i	$f_i x_i$
31-40	1	35.5	35.5
41-50	2	45.5	91
51-60	5	55.5	277.5
61-70	15	65.5	982.5
71-80	25	75.5	1887.5
	$\sum f_i = 48$		$\sum f_i x_i = 3274$

$$\bar{x} = \frac{\sum_{i=1}^m f_i x_i}{\sum_{i=1}^m f_i} = \frac{f_1 x_1 + f_2 x_2 + \dots + f_m x_m}{f_1 + f_2 + \dots + f_m}$$

$$= \frac{3274}{48} = 68.20$$

سؤال واجب : جد الوسط الحسابي لجدول التوزيع التكراري التالي :

الفئات	(0 -)	(1 -)	(2 -)	(3 -)	(4 - 5)
f_i	0.5	1	0.7	1.2	0.88

مزايا وعيوب الوسط الحسابي

يتميز الوسط الحسابي بالمزايا التالية:

- أنه سهل الحساب.
- يأخذ في الاعتبار كل القيم.
- أنه أكثر المقاييس استخداما وفهما.

ومن عيوبه .

- أنه يتأثر بالقيم الشاذة والمتطرفة.
- يصعب حسابه في حالة البيانات الوصفية
- يصعب حسابه في حالة الجداول التكرارية المفتوحة

الوسيط Median

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

أولاً: الوسيط للبيانات غير المبوبة

لبيان كيف يمكن حساب الوسيط للبيانات غير المبوبة ، نتبع الخطوات التالية:

- ترتيب القيم تصاعدياً .

- تحديد رتبة الوسيط، وهي : رتبة الوسيط = $\left(\frac{n+1}{2}\right)$

- إذا كان عدد القيم (n) فردي فإن الوسيط هو:

- الوسيط هو القيمة التي ترتيبها $\left(\frac{n+1}{2}\right)$

- إذا كان عدد القيم (n) زوجي، فإن الوسيط يقع بين القيمة رقم $(n/2)$ ، والقيمة رقم $((n/2)+1)$ ،

ومن ثم يحسب الوسيط بتطبيق المعادلة التالي:

$$\frac{\left(\frac{n}{2} + 1\right) \text{ القيمة رقم} + \left(\frac{n}{2}\right) \text{ القيمة رقم}}{2} = \text{الوسيط}$$

مثال: تم تقسيم قطعة أرض زراعية إلى 17 وحدة تجريبية متشابهة ، وتم زراعتها بمحصول القمح ، وتم استخدام نوعين من التسميد هما : النوع (a) وجرب على 7 وحدات تجريبية ، والنوع (b) وجرب على 10 وحدات تجريبية ، وبعد انتهاء الموسم الزراعي ، تم تسجيل إنتاجية الوحدة بالطن / هكتار ، وكانت على النحو التالي :

النوع (a) 1.5 2.3 3 2 3.25 2.75 1.2

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

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

الحل

أولا : حساب وسيط الإنتاج للنوع الأول (a)

- ترتيب القيم تصاعديا :

	قيمة الوسيط						
الإنتاج	1.2	1.5	2	2.3	2.75	3	3.25
الرتبة	1	2	3	4	5	6	7
	رتبة الوسيط						

- عدد القيم فردى ($n = 7$)
- إذا رتبة الوسيط هي: $((n+1)/2 = (7+1)/2 = 4)$.
- ويكون الوسيط هو القيمة التي تسلسلها 4 ، أي أن وسيط الإنتاج للنوع a هو:

طن / هكتار $Med = 2.3$

ثانيا : حساب وسيط الإنتاج للنوع الثاني (b) :

- ترتيب القيم تصاعديا .

	قيمة الوسيط = $\frac{2.5 + 3}{2}$									
الإنتاج	1.5	1.8	2	2.5	2.5	3	3.5	3.75	4	4.5
الرتبة	1	2	3	4	5	6	7	8	9	10
	الوسيط يقع بين القيمة الخامسة والسادسة									

بما ان عدد القيم زوجي ($n = 10$)

إذا الوسيط يقع بين القيمة رقم $(n/2)$ ، والقيمة رقم $((n/2) + 1)$ ، ومن ثم يحسب الوسيط بتطبيق المعادلة

التالي:

$$n/2 = 10/2 = 5 \quad \text{رتبة الوسيط الاولى}$$

$$(n/2+1) = (10/2+1) = 5+1 = 6 \quad \text{رتبة الوسيط الثانية}$$

- الوسيط = معدل القيمتين رقم 5 و 6

$$Med = \frac{2.5 + 3}{2} = 2.75 \text{ طن / هكتار}$$

و بمقارنة النوعين من السماد ، نجد أن وسيط إنتاجية النوع (a) أقل من وسيط إنتاجية النوع (b) ، أي أن :
 $Med_b > Med_a$.

ثانياً: الوسيط للبيانات المبوبة

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

- تكوين الجدول التكراري المتجمع الصاعد .

- تحديد رتبة الوسيط : $\left(\frac{n}{2}\right) = \left(\frac{\sum f}{2}\right)$

- تحديد فئة الوسيط على ضوء رتبة الوسيط

- عندئذ يحسب الوسيط ، بتطبيق المعادلة .

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

حيث أن :

$n/2$ تمثل ترتيب الوسيط

A تمثل الحد الأدنى للفئة الوسيطة

f_1 تكرار الفئة السابقة لفئة الوسيط

f_2 تكرار الفئة اللاحقة لفئة الوسيط

L هي طول فئة الوسيط، وتحسب بالمعادلة التالية:

$L = \text{Upper} - \text{Lower}$

طول الفئة = الحد الأعلى - الحد الأدنى

مثال : فيما يلي توزيع 50 مريض حسب مدة الرقود في المستشفى على اعتبار ان المريض الذي يخرج في اليوم الأول يعتبر نصف يوم

مدة الرقود في المستشفى	1.5 -	4.5 -	7.5 -	10.5 -	13.5 – 16.5
عدد المرضى	4	12	19	10	5

والمطلوب : حساب الوسيط : أ - حسابيا ب - بيانيا

الحل

أولا : حساب الوسيط حسابيا

$$\frac{n}{2} = \frac{\sum f}{2} = \frac{50}{2} = 25$$

• رتبة الوسيط :

• الجدول التكراري المتجمع الصاعد :

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

رتبة الوسيط

• تحديد فئة الوسيط : وهى الفئة التي تشمل قيمة الوسيط ، ويمكن معرفتها بتحديد التكرارين المتجمعين الصاعدين الذين يقع بينهما رتبة الوسيط $(n/2)$ ، وفي الجدول أعلاه نجد أن رتبة الوسيط (25) تقع بين

التكرارين المتجمعين (16 , 35) ، ويكون الحد الأدنى لفئة الوسيط هو المناظر للتكرار المتجمع الصاعد السابق 7.5 ، والحد الأعلى لفئة الوسيط هو المناظر للتكرار المتجمع الصاعد اللاحق 10.5 . أى أن فئة الوسيط هي : (7.5-10.5) .

• وتطبيق معادلة الوسيط رقم (3-11) على هذا المثال نجد أن :

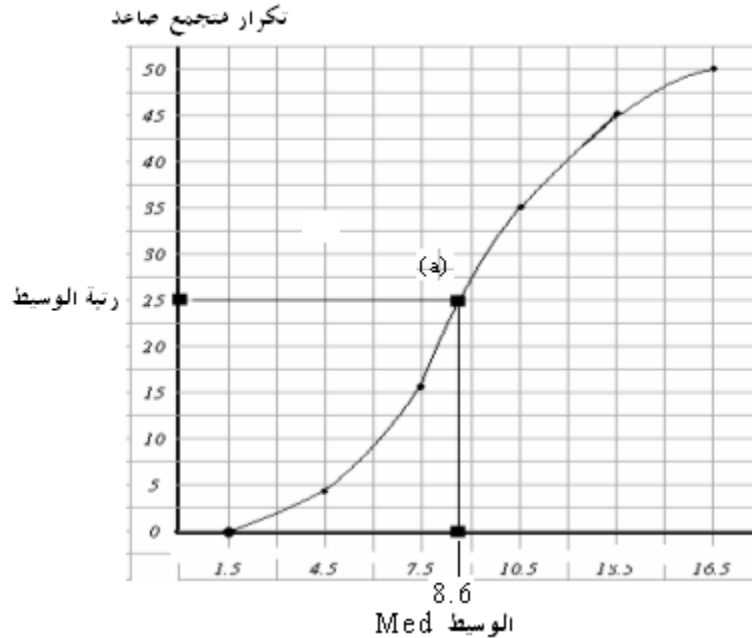
$$A = 7.5 , f_1 = 16 , f_2 = 35 , L = 10.5 - 7.5 = 3$$

إذا الوسيط قيمته هي :

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

ثانياً: حساب الوسيط بيانياً

• تمثيل جدول التوزيع التكراري المتجمع الصاعد بيانياً .



• تحديد رتبة الوسيط (25) على المنحنى التكراري المتجمع الصاعد . ثم رسم خط مستقيم أفقي حتى يلقى

المنحنى في النقطة (a) .

- إسقاط عمود رأسي من النقطة (a) على المحور الأفقي .
- نقطة تقاطع الخط الرأسي مع المحور الأفقي تعطى قيمة الوسيط .
- الوسيط كما هو مبين في الشكل $Med = 8.6$ ويساوي تقريبا 9 .

مزايا وعيوب الوسيط

من مزايا الوسيط

- 1- لا يتأثر بالقيم الشاذة أو المتطرفة .
- 2- كما أنه سهل في الحساب .
- 3- مجموع قيم الانحرافات المطلقة عن الوسيط أقل من مجموع الانحرافات المطلقة عن أي قيم أخرى . أي أن :

$$\sum |x - Med| \leq \sum |x - a| , a \neq Med$$

ومن عيوب الوسيط

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

ثالثا: المنوال Mode

يعرف المنوال بأنه القيمة الأكثر شيوعا أو تكرارا، ويكثر استخدامه في حالة البيانات الوصفية، لمعرفة النمط (المستوى) الشائع، ويمكن حسابه للبيانات المبوبة وغير المبوبة كما يلي:

أولا: حساب المنوال في حالة البيانات غير المبوبة

المنوال (Mod) = القيمة (المستوى) الأكثر تكرارا

ثانيا: حساب المنوال في حالة البيانات المبوبة (طريقة الفروق)

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

حيث أن :

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

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

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

L : طول فئة المنوال .

فئة المنوال = الفئة المناظرة لأكبر تكرار

مثال

اختيرت عينات عشوائية من طلاب بعض أقسام المعهد التقني ديوانية ، وتم رصد درجات هؤلاء الطلاب في

مقرر إحصاء التطبيقي ، وكانت النتائج كالتالي:

قسم الصحة	80	77	75	77	77	77	65	70	58	67
قسم التمريض	88	68	60	75	93	65	77	85	95	90
قسم الإدارة الصحية	87	95	69	80	97	88	76	95	86	80
قسم المحاسبة	85	73	69	85	73	69	69	73	72	85

والمطلوب حساب منوال الدرجات لكل قسم من الأقسام :

الحل هذه البيانات غير مبوبة ، لذا فإن :

المنوال هو القيمة الأكثر تكرارا

والجدول التالي يبين منوال الدرجة لكل قسم من الأقسام .

القسم	القيمة الأكثر تكرار	القيمة المنوالية
قسم الصحة	الدرجة 77 تكررت 4 مرات	المنوال = 77
قسم التمريض	جميع القيم ليس لها تكرار	لا يوجد منوال
قسم الإدارة الصحية	الدرجة 65 تكررت 2 مرات الدرجة 80 تكررت 2 مرات	يوجد منوالان هما : المنوال الأول = 95 المنوال الثاني = 80
قسم المحاسبة	الدرجة 69 تكررت 3 مرات الدرجة 73 تكررت 3 مرات الدرجة 85 تكررت 3 مرات	يوجد ثلاث منوال هي : المنوال الأول = 69 المنوال الثاني = 73 المنوال الثالث = 85

مثال

فيما يلي توزيع 30 أسرة حسب الإنفاق الاستهلاكي الشهري لها بالألف الدينارين .

فئات الإنفاق	2 -	5 -	8 -	11 -	14 - 17
عدد الأسر f	4	7	10	5	4

والمطلوب حساب منوال الإنفاق الشهري للأسرة، باستخدام طريقة الفروق .

الحل

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

- تحديد الفئة المنوالية

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

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

$d_1 = 10 - 7 = 3$
 أكبر تكرار
 $d_2 = 10 - 5 = 5$
 فئة المنوال
 $A = 8$

- حساب الفروق d ، حيث أن :

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

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

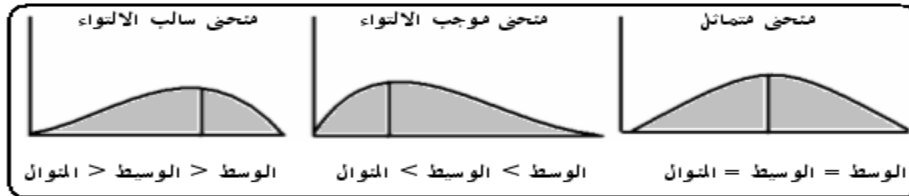
- وتطبيق المعادلة الخاصة بحساب المنوال في حالة البيانات المبوبة . نجد أن :

$$\begin{aligned}
 Mod &= A + \frac{d_1}{d_1 + d_2} \times L \\
 &= 8 + \frac{3}{3 + 5} \times 3 = 8 + 1.125 = 9.125
 \end{aligned}$$

استخدام مقاييس النزعة المركزية في تحديد شكل توزيع البيانات

يمكن استخدام الوسط الحسابي والوسيط والمنوال في وصف المنحنى التكراري، والذي يعبر عن شكل توزيع

البيانات ، كما يلي :



- يكون المنحنى متماثل إذا كان :

$$\text{الوسط} = \text{الوسيط} = \text{المنوال} .$$

- يكون المنحنى موجب الالتواء إذا كان:

$$\text{الوسط} < \text{الوسيط} < \text{المنوال}$$

- يكون المنحنى سالب الالتواء إذا كان :

الوسط > الوسيط > المنوال

مثال : الجدول التكراري التالي يعرض توزيع 100 عامل في مزرعة حسب الأجر اليومي بالدينار .

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

والمطلوب :

- حساب الوسط والوسيط والمنوال .
- بيان شكل توزيع الأجور في هذه المزرعة .

الحل

- حساب الوسط والوسيط والمنوال .

أولاً : الوسط الحسابي \bar{x}

فئات الأجر	التكرارات (f)	مراكز الفئات (x)	f x
50 - 70	8	60	480
70 - 90	15	80	1200
90 - 110	28	100	2800
110 - 130	20	120	2400
130 - 150	15	140	2100
150 - 170	8	160	1280
170 - 190	6	180	1080
المجموع	100		11340

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{11340}{100} = 113.4 \text{ R.S}$$

ثانياً : الوسيط Med

رتبة الوسيط : $(n/2 = 100/2 = 50)$

تكوين التوزيع التكراري المتجمع الصاعد .

أقل من	تكرار متجمع صاعد
أقل من 50	0
أقل من 70	8
أقل من 90	$\leftarrow f_1 23$
أقل من 110	$\leftarrow f_2 51$
أقل من 130	71
أقل من 150	86
أقل من 170	94
أقل من 190	100

رتبة الوسيط (50)

من الجدول أعلاه نجد أن :

$$\frac{n}{2} = 50 , f_1 = 23 , f_2 = 51 , A = 90 , L = 110 - 90 = 20$$

إذا الوسيط قيمته هي :

$$\begin{aligned} Med &= A + \frac{\frac{n}{2} - f_1}{f_2 - f_1} \times L = 90 + \frac{50 - 23}{51 - 23} \times 20 \\ &= 90 + \frac{27}{28} \times 20 = 90 + \frac{540}{28} = 90 + 19.286 = 109.3 \text{ R.S} \end{aligned}$$

ثالثاً : المنوال Mod

الفئة المنوالية ، هي الفئة المناظرة لأكبر تكرار

أكبر تكرار = 28 ، وهو يناظر الفئة التقريبية (90 - 110) .

$$d_2 = 28 - 20 = 8 \quad , \quad d_1 = 28 - 15 = 13 \quad \text{حساب الفروق :}$$

$$L = 110 - 90 = 20 \quad \text{الحد الأدنى للفئة : } A = 90 \quad \text{طول الفئة}$$

إذا المنوال يحسب بتطبيق المعادلة التالية :

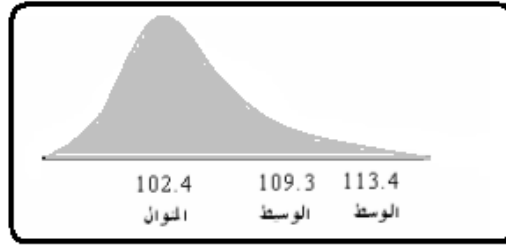
$$Mod = A + \frac{d_1}{d_1 + d_2} \times L = 90 + \frac{13}{13 + 8} \times 20 = 90 + \frac{260}{21} = 102.4 \quad R.S$$

• بيان شكل التوزيع .

من النتائج السابقة ، نجد أن :

$$\bar{x} = 113.4 \quad \text{الوسط الحسابي} \quad Med = 109.3 \quad \text{الوسيط} \quad Mod = 102.4 \quad \text{المنوال}$$

أي أن : الوسط < الوسيط < المنوال ، إذا "توزيع بيانات الأجرور موجب الالتواء. كما هو مبين في الشكل التالي:



مثال واجب

قام مدير مراقبة الإنتاج بسحب عينة من 10 عبوات من المياه المعبأة للشرب ، ذات الحجم 5 لتر ،

والمنتجة بواسطة إحدى شركات تعبئة المياه لفحص كمية الأملاح الذائبة، وكانت كالتالي :

115 123 119 123 124 119 123 121 123 121

والمطلوب : حساب الوسط الحسابي، والوسيط، والمنوال، ثم حدد شكل الالتواء لهذه البيانات .