نبذة تاريخية عن علم الاحصاء:

ويرجع الاهتمام بالإحصاء إلى عصور قديمه ، وان تعداد السكان عند القدماء المصرين وفي الصين أمثلة توضح اهتمام الحكومات منذ القدم بالمعلومات الاجتماعية وذلك لأغراض التنظيم والتخطيط فى أحوال السلم والحرب .ويبدو أن كلمه إحصاء (statistics) قد ظهرت لأول مره عام 1749 وهى مشتقه من الكلمة اللاتينية (status) أو الايطالية (statista) وتعني كلاهما الدولة السياسية . ومن الطبيعي أن تكون الدولة أول من اهتم بجمع البيانات وذلك لإدارة شؤون البلاد خاصة عن السكان لأغراض حربية وضريبية ، وامتدت بعد ذلك لتشمل إحصاءات حجم السكان والمواليد والوفيات والإنتاج والاستهلاك والثروة ....الخ . وهكذا بداء العلم وتطوره باعتباره علم الدولة أو علم الملوك .ولقد تطور علم الإحصاء من مجرد فكره الحصر والعد إلى أن أصبح الآن علما له قواعده ونظرياته ويرجع الفضل فى ذلك إلى كثير من العلماء من أمثال عائله برونلي Bernoulli وفردريك جاوس F.gauss وكيتليه Quetlet وجولتونF.galton وأخيرا كارل بيرسون Karl.pearson وبولي A.bowley وبول U.yule فيشر L.fisher و...الخ .

وجاء التطور في علم الإحصاء بصفه عامه ملازما وموازيا للتطور في نظرية الاحتمالات . فقد نشأت نظريه الاحتمالات على أساس رياضى فى (1494) بواسطة باسيولي Lucapacidi . ومن الدراسات الفلكية لكل من كبلر (1630-1517) Keplr وجاليليو (1642-1564) Galilio قاما بتطوير نماذج الاحتمالات . غير أن التاريخ الحقيقي لنظريه الاحتمالات بدء فى القرن السابع عشر حيث وضعت أسسها في عام 1654 بواسطة كلا من العالمين : باسكان Pascal,B. (1623 1662) عالم الرياضيات والفيزياء والفيلسوف الفرنسى – وكذا العالم فرمات Fermat (1608 – 1665) . وبعد ذلك بثلاث سنوات قام هينجينز Huygens (1629 – 1695) بنشر كتيب صغير في موضوع المعالجة الرياضية لفرص الفوز في مباريات ورق اللعب وزهرة النرد ،وفى نفس الوقت تقريبا قام جرونت grunt (1620 – 1674) بنشر ملاحظاته عن معالجة البيانات المتعلقة بالحكومة خاصة في النواحي الطبيعية والسياسية والتجارية والنمو والوفيات والأمراض.

وقد كان العمل الذى قام به هيجيتر دافعا للكثيرين لدراسة النظريات والمشاكل المتعلقة بمباريات الصدفة ومنهم برنوللى (1654 – 1705) ودي موافر De Moivre(1667 – 1754) واربوثنوت Arbuthnott ولابلاس laplace (1749 – 1827) وجاوس Gauss(1777 – 1855). ويعد العالم البلجيكى كتيليه (1796 – 1874) أول من وضع قواعد محددة لعلم الإحصاء ، وكلمة إحصاء في الوقت الحاضر ذات معان متعددة فمنها يفهم جمع المعلومات التى تبين الحالة في الدولة مثل عدد المواليد والوفيات وبيانات عن المحاصيل والتجارة الخارجية... الخ ويسمى نشر الأجهزة الحكومية لمثل هذه المعلومات في شكل كتب وتقارير " بالإحصاء الرسمي، وأخيرا يفهم بالإحصاء فرع من العلم له نظريته الخاصة . وعلم الإحصاء ، شأنه في ذلك شأن أى فرع آخر من فروع العلم له أسلوبه وموضوعات البحث الخاص به .

تعريف علم الإحصاء: هو العلم الذي يختص بجمع وتصنيف وتبويب البيانات وعرض وتحليل البيانات واستخلاص النتائج منها. او هو الطريقة العلمية التي تختص بجمع البيانات والحقائق بشكل يسهل عملية تحليلها وتفسيرها ومن ثم استخلاص النتائج واتخاذ القرار على ضوء ذلك.

يقسم الإحصاء بصورة عامة إلى قسمين رئيسين هما:

1- الإحصاء الوصفي Descriptive Statistic

يتضمن طرق جمع البيانات وتصنيفها وعرضها في جداول ورسومات بيانية وحساب بعض المؤشرات الإحصائية.

2- الإحصاء الاستدلالي Inference Statistic

هو الجزء من علم الإحصاء الذي يهتم بتحليل البيانات واستقراء النتائج واتخاذ القرارات بناء عليها ، اي يستخدم العينة للاستدلال عن خصائص المجتمع او هو الجزء الذي يتضمن طرق اختبار الفرضيات وتقدير المعالم.

اهمية علم الاحصاء ومجالات استخدامه:

يعتبر علم الإحصاء احد الوسائل المهمة في البحث العلمي من خلال استخدام قواعده وقوانينه وطرقه في جمع البيانات و المعلومات اللازمة في البحث العلمي وتحليل هذه المعلومات لغرض الوصول إلى النتائج التي يهدف لها البحث كما وان للإحصاء دورا بارزا في وضع الخطط المستقبلية عن طريق التنبؤ بالنتائج ففي قصة نبينا يوسف عليه السلام مثال عظيم لدور الاحصاء في التخطيط فيبين ان هناك سنوات عجاف يقل فيها المحصول وسنوات سمان يزيد فيها المحصول ويبين انه يجب الاحتفاظ لسني القحط بادخار جزء من انتاج سني الرخاءلكافة القطاعات حيث يمكن تطبيق علم الإحصاء في كافة المجالات وكما يلي:

* الاقتصاد :اذ يستخدم لمعرفة حجم التجارة – مصادر القوى العاملة – مستوى المعيشة – تحليل سلوك المنتج والمستهلك– مدى تأثر الاسواق بالأنظمة والسياسات – اختبار كفاءة الانتاج.... الخ

* الادارة : اذ يستخدم في ادارة جودة الانتاج والمقارنة بين السياسات التسويقية والإدارية.

* الاجتماع :اذ يستخدم في وصف الهجرة الداخلية ومدى تأثيرها على سلوك المجتمع – تحليل تكاليف المساعدات والتأمينات الاجتماعية.... الخ

* علم النفس :اذ يستخدم في نظريات التربية – المشاكل المتعلقة بقياس القدرة على التعلم والذكاء والصفات الشخصية والسلوك الطبيعي وغير الطبيعي للأشخاص....الخ

* التعليم :اذ يستخدم في وضع خطط التعليم الحالية والمستقبلية – تقدير احتياجاتها من القوى البشرية والمباني والمعامل والأجهزة.

*السكان :اذ يستخدم لدراسة تطور مجتمع السكان – معدل النمو – معرفة معدلات المواليد والوفيات – خصائص المجتمع الاجتماعية والاقتصادية والمهنية.

* الاحياء : اذ يستخدم في الابحاث الاساسية والتجارب المعملية المتعلقة بتطوير الحياة الوراثية

* الطب اذ يستخدم لمعرفة واختبار فعالية الادوية الجديدة – التعرف على العلاقة بيم الامراض ومسبباتها .

الطريقة الاحصائية في البحث العلمي:

استخدام الاسلوب الاحصائي في البحث العلمي يعني توفير البيانات والمعلومات عن الظاهرة المطلوب دراستها في ذلك البحث وهذا يعني ان امكانية تطبيق الطريقة الاحصائية مرهوناً بإمكانية التعبير عن هذه الظاهرة أو تلك تعبيرا كميا (رقمياً) ، وتمتاز الطريقة الاحصائية بكونها تهيئ اسلوباً موضوعيا محايدا للبحث له قواعده وأصوله التي يجب أن يلتزم بها الباحث حتى يتجنب التحيز الشخصي والوقوع في بعض الاخطاء . كما يساعد استخدام الطريقة الإحصائية الى وصول الباحث الى النتائج الدقيقة بأقصر طريق وأقل كلفة.

المراحل الرئيسية للطريقة الاحصائية:

1- تحديد مشكلة البحث بشكل دقيق كي نتمكن من تحديد ماهية البيانات المطلوبة.

2- صياغة الفرضيات باعتبارها عبارات محددة حول معالم المجتمع بهدف قبولها او رفضها.

3- تحديد مجتمع الدراسة ووحدة المجتمع اي تحديد الاطار الذي يحتوي على كافة المفردات التي تشملها الدراسة.

4- جمع البيانات والمعلومات الاحصائية ويتم ذلك بتحديد مصادر البيانات والمعلومات اللازمة للظاهرة او الموضوع المراد دراسته ومن ثم يتم تحديد اسلوب جمع هذه البيانات في ضوء طبيعة البيانات المطلوبة وحجم ونوع مجتمع الدراسة والإمكانات المادية والزمنية والفنية.

5- فرز وتصنيف البيانات وتبويبها.

6-عرض البيانات الاحصائية حسب طرق العرض المناسبة سواء كانت بيانية او اشكال هندسية.

7- حساب المؤشرات أو المعالم للبيانات واختبار الفرضيات واستخلاص النتائج.

8- التفسير والتنبؤ.

مفاهيم اساسية:

المجتمع الإحصائي هو عبارة عن مجموعة الوحدات او المفردات التي تشترك بصفة او صفات معينة والتي غالباً ما يتم الحصول منها على البيانات والمعلومات المطلوبة ،ويقسم المجتمع الإحصائي من حيث العدد الى:

1-مجتمع منتهي (محدود):هو المجتمع الذي يمكن حصر مفرداته خلال فتره محدده.

مثال :حجم أعداد الطلاب في جامعة بابل

2-مجتمع غير منتهي (غير محدود) : هو المتجمع الذي لا يمكن حصر مفرداته.

مثال : عدد قطرات المطر

العينة:هي جزء من المجتمع الاحصائي او هي مجموعه جزئيه من مفردات المجتمع محل الدراسة يتم اختيارها بحيث تكون ممثله للمجتمع تمثيل صحيح.

البيانات الإحصائية: هي مجموعة من الأرقام أو المقاييس أو الصفات التي تم جمعها عن المجتمع الإحصائي أو العينة التي قصد معالجتها وتحليلها.

الظاهرة: الموضوع الذي نخص به المجتمع الإحصائي..

المعلمة : هي شيء يميز المجتمع ككل مثل متوسط الدخل الشهري في دولة معينة او متوسط طول الطلاب في مدرسة معينة او نسبة في مجتمع ما او نسبة المدخنين في في دولة ما ...الخ.اوهي قيمه عددية تصف المجتمع مثل الوسط الحسابي والانحراف المعياري للمجتمع.

الاحصاءة: هي شيء يميز العينة مثل متوسط الدخل الشهري لعينة متكونة من 100 اسرة في دولة معينة او متوسط الطول لعينة متكونة من 30 طالب في مدرسة معينة ما وهكذا أوهي قيمة عددية تصف العينة مثل الوسط الحسابي والانحراف المعياري للعينة.

المشاهدة (Observation) :هي قيمة المتغير الخاصة بوحدة الدراسة.

البيانات (Data) :هي مجموعة من المشاهدات أو الملاحظات التي تؤخذ أثناء دراسة معينة ، وقد تكون بيانات رقمية (كمية)، أو بيانات غير رقمية(وصفية).

مصادر جمع البيانات:

1 -المصادر المباشرة (الميدانية) :جمع البيانات عند ظاهرة أثناء حدوثها في ميدان العمل مثل(المشاهدة ،الملاحظة ، والتسجيل ،الاتصال الهاتفي، المقابلة الشخصية، الاستبيان)

2 -المصادر غير مباشرة (تاريخية): جمع البيانات من خلال سجلات سبق نشرها وتكون معدة مسبقا عن ظاهرة ما ويستطيع الباحث الرجوع اليها واخذ المعلومات المطلوبة من مصادر رسمية مثل (دائرة الاحصاءات العامة ،الاحوال المدنية ،هيئات دولية)

طرق جمع البيانات بناءاً على عدد المجتمع (اساليب جمع البيانات):

هناك أسلوبين لجمع البيانات :

أ- أسلوب المسح الشامل

يتناول هذا الاسلوب دراسة كافة مفردات المجتمع الاحصائي مثل (التعداد السكاني)

مزاياه

1 - دقة النتائج

2 -عدم وجود أخطاء عشوائية

عيوبه

1 -ارتفاع التكاليف (جهد ، المال،الوقت)

2 -عدم إمكانية تطبيقه عمليا على المجتمعات ذات المفردات الكبيرة الحجم.

ب- أسلوب المسح العيني:

يتناول هذا الاسلوب اختيار جزء من المجتمع لتمثيله بطريقة العينة العشوائية بشرط أن تكون

العينة ممثلة تمثيلاً صادقاً دون تحيز.

مزاياه:

1 -انخفاض التكاليف (جهد ، المال،الوقت)

2 -إمكانية تطبيقه مهما كان حجم المجتمع

عيوبه :

1 -نتائجه تقريبية

2 -إمكانية الخطأ نتيجة التحيز

ويعتمد اختيار الاسلوبين على طبيعة الظاهرة والبيانات الاحصائية المطلوبة.

أنواع البيانات الاحصائية :

تصنيف البيانات طبقا للأسس الأربعة التالية:

أ .نوع البيانات Type of Data

البيانات المنفصلة او المتقطعة :(Discrete data) وهي البيانات الكمية التي تأخذ قيما عددية محددة صحيحة ولا تحتوي علي قيم كسرية مثل عدد موظفي جامعة البصرة ، عدد المصانع في كل مدينة من مدن دولة ما ، الاشخاص، التفاح، والقوميات.

المتغيرات الكمية المتصلة (Continuous data) : هي البيانات الكمية التي تأخذ قيما تكون عددا صحيحا وكسرا من وحدة القياس مثل : بيانات الدخل اليومي (الدولار)لعينة من الاشخاص، الطول(مقاسا بالسنتمترات) ، الوزن(مقاسا بالكيلوغرام) ، العمر(مقاسا بالسنوات)درجة حرارة الجو(مقاسة بالدرجات.) وغيرها.

ب .مستويات القياس:

القياس Measurement : التعبير عن الظاهرة كميا أو رقميا.

1 -المقياس الاسمي(Nominal Scale) :

وهي تتميز عن غيرها من البيانات بإعطاء اسم لها كالأطفال والمزارعين ... الخ.

واستخدام الارقام هنا للتصنيف والترميز مثال تصنيف الجنس: الذكر 1 والانثى 2

وفي هذا النوع من البيانات لا نستطيع المفاضلة فالعيون الخضراء ليست افضل من السوداء مثلاً

2-المقياس الرتبي:(Ordinal Scale)

يتميز بخاصية الترتيب بالإضافة للتصنيف(الترتيب للصفات تصاعديا أو تنازليا)لا يشترط تساوي المسافة بين الترتيبات(ترتيب المتسابق:(...،3 ، 2،1)---دون معرفة الفرق بينهم مثال على ذلك: الرتب العسكرية

3-المقياس الفئوي:(Interval Scale) :

يتميز بان المسافة بين النقاط المتتالية متساوية، درجات التحصيل، نسب الذكاء--- والصفر هنا نسبي وليس مطلق

4- المقياس النسبي:(Ratio Scale) :

يتميز بانه يحوي جميع الخصائص السابقة، بالإضافة للصفر المطلق (درجة الحرارة، المسافة بين بعدين،)-----

ج .عدد الجماعات Groups التي يتم جمع البيانات عنها:

•تنشأ البيانات عن **جماعة واحدة One – group data** من خلال جماعة منفردة من المبحوثين ، مثل إجراء دراسة قبلية أو عمدية على بعض المتغيرات خاصة جماعة واحدة من الأفراد.

•تنشأ البيانات عن **جماعتين Two – group data** عند دراسة جماعتين كما هو الحال بالنسبة للجماعة الضابطة أو الجماعة التجريبية

•تنشأ البيانات عن **جماعات كثيرة Many groups data** من خلال دراساتنا لجماعات سكنية متعددة Multi group populationsوالتي يتم فيها دراسة متغيرات متعارضة في مواقف متباينة.

د .عدد المتغيرات Variables :

•هناك بيانات قائمة على متغير واحد Univariate عن جماعة من الجماعات ويعد هذا النوع من الدراسات من ابسط الأنواع وهو ينتمي إلى نموذج بحث كلاسيكي .وفي هذه الحالة يتم تثبيت جميع المتغيرات ما عدا متغير واحد عند القيام بإجراءات الدراسة.

•هناك بيانات تشمل على متغيرين bivariate اثنين .ومثال هذا النوع من الدراسات تتضمن قياس التحصيل النسبي في مادتين لكل فرد من أفراد العينة.

•هناك بيانات تشمل على متغيرات متعددة Multi variates عادة ما يتم عزلها ثم دراستها باستخدام التحليل المتعدد للمتغيرات.

المعاينة:

هي عملية اختيار جزء من المجتمع الاحصائي للاستدلال حول خواص المجتمع بأكمله عن طريق تعميم نتائج العينة وتقوم على علم وفن التحكم وقياس دقة المعلومات الاحصائية عن طريق استخدام بعض النظريات الرياضية وليس مجرد استخدام جزء من المجتمع بدلاً من كله.

مميزات اسلوب المعاينة:

يتميز أسلوب المعاينة عن أسلوب الحصر الشامل بمزايا عديده منها:

1- يؤدي استخدام العينات العشوائية إلى خفض تكاليف الدراسات الميدانية بسبب صغر حجم العينة بالنسبة إلى حجم المجتمع وهو ما يؤدي إلى تخفيض الأعباء الإدارية والفنية التي تتطلبها أي دراسة ميدانية.

2- يتحقق وفر واضح في الوقت الذي ينفق في دراسة ميدانية على أساس عينة بدلاً من الحصر الشامل وتتضح أهمية الوقت عندما نقوم بدراسة ظاهرة تتغير بمرور الوقت ، وعينة قد يترتب على دراسة تلك الظاهرة في المجتمع كله بجمع البيانات من جميع مفردات المجتمع أن يمر وقت بديل فتكون البيانات والنتائج وقت ظهورها غير مطابقة لواقع المجتمع وتصبح النتائج ذات قيمه محدودة بعد أن فقدت عنصر المطابقة مع واقع الظاهرة وتوزيعها الحالي لمجتمع ، والتعدادات الدورية للسكان وبسبب ضخامة حجم العمل بها تستغرق وقتاً طويلاً حتى تصبح نتائجها جاهزة ومنشورة وقد يطول هذا الوقت إلى أكثر من ثلاث أو أربع سنوات حتى مع استخدام أحدث أجهزة الحاسبات الأليه الضخمة ، ويكون على الباحثين مستخدمي هذه النتائج مراعاة الوقت الذي ينقض بين تاريخ إجراء التعداد وتاريخ نشر نتائجه وتعديل هذه النتائج في حدود ذلك .. وهذا دفع الكثير من الدول إلى تعزيز نتائج التعدادات الدورية للسكان بنتائج تعدادات تجري بين كل تعدادين متتاليين على أساس العينة.

3- في المجتمعات غير المحدودة (اللانهائية) مثل مجتمع الكائنات الحية في البحار والمحيطات لا يمكن أن تتم الدراسة على أساس الحصر الشامل ولكن لابد وأن تتم الدراسة بأسلوب المعاينة.

4- أيضاً هناك بعض الاختبارات لابد وأن تتم بأسلوب المعاينة لأن إجراء مثل هذه الاختبارات على أساس الحصر الشامل يؤدي إلى تلف المادة المختبرة أو هلاكها.. فاختبار صلاحية شحنه من المفرقعات مثلاً لابد وأن يتم على أساس العينة وبالمثل تحليل دم المرضى يتم على أساس عينه.

كيفية اختيار عينة البحث بدقة :

لاختيار عينة الدراسة بدقة وعناية على الباحث أن يلتزم بالخطوات الآتية :

1- **تحديد الهدف من البحث: يعد تحديد أهداف البحث من أهم الخطوات التي على الباحث الالتزام بها، إذ إن نجاح هذه الخطوة يؤدي إلى نجاح البحث بأكمله.**

2- **تحديد المجتمع الأصلي: يعد مجتمع الدراسة المجتمع الذي يتم تعميم عينة الدراسة عليه، فعينة الدراسة ليست سوى نموذج عن هذا المجتمع، والتي تمكننا من دراسة خصائصه، وبالتالي إسقاط نتائجها عليه.**

3- **تحديد خصائص مجتمع الدراسة: يجب على الباحث أن يقوم بدراسة المجتمع الذي يريد دراسته، وتصنيف خصائصه وفقاً لفهمه الخاص ودراسته، والتي قد تتغير بحسب أهداف الدراسة، كما عليه دراسة المتغيرات التي تشملها الدراسة.**

4- **تحديد حجم عينة الدراسة: لا يوجد حجم معين لعينة الدراسة، ولكن يفضل أن تكون العينة كبيرةً لتعطي نتائج دقيقة بشكل أكبر، وقد تتسبب العينة الكبيرة في تصعيب مهمة الباحث في ضبط المتغيرات، وذلك نظراً لكثرة تغيراتها، وتلعب العديد من العوامل دوراً في تحديد حجم عينة الدراسة، ومن أهم هذه العوامل ما يأتي:**

أ- **نوع المجتمع الأصلي : يلعب تجانس المجتمع دوراً كبيراً في نوع المجتمع الأصلي، فكلما كان المجتمع متجانساً كان دراسته أسهل، ويتطلب عينة أصغر.**

ب- **نوع البحث : قام الباحثون بتحديد عدد أفراد العينة بحسب نوع البحث، ففي البحوث العاملية يكون عدد الأفراد من شخص إلى خمسة، وفي البحوث الوصفية يكون العدد 20% للمجتمعات الصغيرة والتي يبلغ تعدادها بالمئات، و10% للمجتمعات الكبيرة التي يبلغ تعدادها بالآلاف، و5 % للمجتمعات التي يبلغ تعداد سكانها عشرات الآلاف.**

ت- **فروض البحث : يجب على الباحث أن لا يحصل على فروق صغيرة، وفي حال شعر بأنه قد يحصل على فروق صغيرة فعليه بتكبير عينة البحث.**

ث- **تكاليف البحث : يجب أن يختار الباحث عدد أفراد العينة بطريقة تتوافق مع التكاليف التي يرصدها للبحث.**

ج- **أهمية النتائج : في الدراسات الاستطلاعية يتم القبول بنتائج العينة الصغيرة، أما في حال ترتب على الدارسة اتخاذ قرارات مهمة فيفضل استخدام العينة الكبيرة.**

ح- **طرق جمع البيانات : يجب أن يكون الباحث دقيقاً في اختيار أدوات بحثه، وكلما زاد حجم العينة زادت نسبة إمكانية تعميمها.**

خ- **الدقة المطلوبة : تتوقف دقة العينة على حجمها، فكلما كانت العينة أكبر كانت دقتها أكثر.**

<u>انواع العينات :</u>

هناك نوعان رئيسيان من العينات هي :

<u>اولاً/ العينات الغير أحتمالية (الغير عشوائية nonrandom Samples):</u>

تستخدم هذه العينات في حالة عدم القدرة على تحديد مجتمع الدراسة بشكل دقيق مثل دراسة تاريخ الأردن في مرحلة الإمارة الأردنية على سبيل المثال، وتتصف هذه العينات بأنها لا تعطي نفس الفرصة لجميع أفراد مجتمع الدراسة بالظهور في العينة. ومن أنواع هذه العينات ما يلي :

1- العينة الصدفة (العرضية) Accidental Sample : وهذا النوع من العينة يتم اختياره بالصدفة مثلما تستطلع صحيفة معينة الرأي العام حول قضية معينة أو مرشح ما، وغالبا ما يكون هذا النوع من العينات غير ممثلا لمجتمع الدراسة ، وتستخدم هذه العينة في الدراسات الاستطلاعية المسحية المبدئية.

2- العينة القصدية Purposive Sample : ينتقي الباحث أفراد عينته بما يخدم أهداف دراسته وبناءا على معرفته دون أن يكون هناك قيود أو شروط غير التي يراها هو مناسبة من حيث الكفاءة أو المؤهل العلمي أو الاختصاص أو غيرها، وهذه عينة غير ممثلة لكافة وجهات النظر ولكنها تعتبر أساس متين للتحليل العلمي ومصدر ثري للمعلومات التي تشكل قاعدة مناسبة للباحث حول موضوع الدراسة.

3- عينة القطعة أو الكسرة Chunk Sample : ويقوم الباحث باقتطاع عدد معين من المجتمع كأن يأخذ أول عشرة أفراد ويطبق عليهم الدراسة، وهي اضعف أنواع العينات على الإطلاق، لعدم قدرتها على تمثيل المجتمع.

4- عينة التطوع Volunteer Sample : تحتاج بعض الدراسات إلى متطوعين لإجرائها مثل التحدث مع البث المباشر حول موضوع محدد، أو لإجراء التجارب التربوية أو النفسية، و غالبا لا تمثل هذه العينة مجتمع الدراسة، ولكنها تسهل على الباحث التعاون من قبل أفراد العينة وسرعة الإنجاز.

5- العينة الحصصية Quota Sample : وتشبه العينة الطبقية ولكن الاختلاف أن مجتمع الدراسة غير محدد. وفيها يتم اختيار المبحوثين بنسبة توزيعهم فى المجتمع الاحصائى مثال اختيار20% من الإناث 40 % من الذكور وهكذا . ولكن الاختيار الاعتباطي والاختيار بالحصة يعد اختيارا غير اهتمامي ، بمعنى أنه لا يوفر فرصة متكافئة لكل مفردات المجتمع الاحصائى لتظهر فى العينة مما يؤدى إلى إخفاق العينة فى أن تمثل المجتمع ككل وتستخدم أحيانا فى المسوح اللاحتمالية للرأي العام وتكون فى هذه الحالة أشبه بالعينة الطبقية . ففي هذه الحالة يعطي القائم بالمقابلة حصة معينة يجب استيفاء بياناتها كأن يلتزم بعدد كبير من الإناث فمن يزيد أعمارهن عن أربعين عاما وأيضا يلزم بعدد كبير من الأشخاص تقل دخولهم السنوية عن (300) دولار . أو أن يخصص له نسبة معينة من الأطباء فى مجتمع ما وهكذا بحيث يكون الباحث قادرا على أن يتم الحصة المطلوبة منه .

ثانياً/ العينات الاحتمالية : Probability Samples :

لقد طور العلماء أساليب المعاينة الاحتمالية لتجنب المخاطر التى تترتب على اختيار عينة غير ممثلة لمجتمع الدراسة وهذه المخاطر لا يمكن تجنبها تماما ولكن هذه الأساليب تمكننا على الأقل من تحديد نسبة الخطأ المحتمل وتعرف العينة الاحتمالية بأنها العينة التى يتم سحبها بحيث يكون لكل مفردة من مفردات المجتمع فرصة معلومة ومتكافئة فى أن يكون جزءا من العينة . ويتسم هذا النوع من العينات بالخصائص التالية :-

أ - لكل مفردة فى العينة درجة احتمالات معروفة يفترض وجودها بين باقى مفردات تلك العينة

ب - لجميع مفردات المجتمع الأصلي فرص متساوية للظهور فى العينة .

يلزم أن تكون الاحتمالات معروفة لدى الباحث حتى يمكن التوصل إلى الثقل الصحيح للعينة أما إذا لم يعرف الباحث تلك الاحتمالات فإنه قد يستحيل عليه أن يستخدم بنجاح الاستنتاج الإحصائي المعتمد على دلالات بحثية . ومن أنواع هذه العينات ما يلي :

1- العينة العشوائية البسيطة: Simple Random sample

العينة العشوائية هي العينة التى تختار بحيث تعطي جميع مفردات المجتمع المراد بحثه نفس الفرصة فى الاختيار وهذا يعنى عدم الاهتمام ببعض المفردات أكثر من البعض الآخر وإتاحة الفرصة المتكافئة أمام كل مفردة للظهور فى العينة ويمكن أن نحقق ذلك بأن نحضر عدا من البطاقات المتشابهة (فى اللون والحجم والوزن وكل شيء) ونكتب على كل بطاقة رقماً يمثل مفرده من مفردات المجتمع وتسحب عددا من هذه البطاقات (بعد خلطها) فنجد أن الأرقام المدونة عليها تعطي لنا المفردات التى تم اختيارها بطريقة عشوائية . وتعرف العينة العشوائية البسيطة بأنها اختياراً بسيطاً بطريقة تتصف بخاصيتين أساسيتين هما :-

أ - أن يتحقق لكل عضو أو مفرده من المجتمع الأصلي درجة احتمال متساوية فى الاختيار

ب – أن يكون اختيار كل مفردة من مفردات العينة بصورة مستقلة عن الأخرى

لو تصورنا أن أحد الأساتذة بقسم الاجتماع يود إجراء دراسة عن اتجاهات طلاب القسم نحو إدمان المخدرات ثم وضع أسماء هؤلاء الطلاب وعددهم 4000 فى حقيبة كبيرة ثم سحب منها 400 اسم أو أنه أعطى رقما مسلسلاً لكل من هؤلاء الأربعة آلاف طالب تم اختيار 400 رقما من جدول الأرقام العشوائية وقام بعد ذلك باختيار الطلاب الذين يتطابق رقمهم المسلسل مع الأرقام العشوائية المختارة له فإنه يكون بذلك قد أعطى لكل طالب من الطلاب فرصة متكافئة لكى يكون من أحد أفراد العينة .

2- العينة المنتظمة : Systematic sample : العينة المنتظمة هي نوع من المعاينة العشوائية بمقتضاها يمكن أن يختار الباحث لو أخذنا فى الاعتبار المثال السابق نسبة 10% من عدد الطلاب (400 طالب) ويستطيع الباحث أن يختار هؤلاء الطلاب بطريقة عشوائية فيبدأ بالطالب رقم 8 ثم بعد كل عشر طلاب يقوم باختيار طالب آخر وهكذا أي أنه فى هذه الحالة سيختار الطالب رقم 8 ، 18 ، 28 ، 38 وهكذا . وهذه الطريقة فى الاختيار مقبولة ما لم يكن اختيار الأرقام من البداية يخفض وراءه تحيز الباحث نحو اختيار طلاب بعينهم . والواقع أن الطريقتين السابقتين من طرق اختيار العينات تلائم الباحثين المبتدئين وغيرهم ممن يريدون تجنب التعقيدات الإحصائية وهناك بالإضافة إلى تلك الطرق أساليب أخرى أكثر تطوراً لسحب العينات توفر للعينة صفات أساسية كأن تكون ممثلة ومقبولة ومناسبة من حيث التكاليف . وتعتبر العينة المنتظمة أكثر أفضلية من العينة العشوائية البسيطة وذلك فى حالة توفر قوائم تضم جميع مفردات المجتمع الأصلى غير أن السهولة فى العينة المنتظمة يناظر بعض العيوب من أهمها .

أ- توقع نتائج خاطئة إذا تم استخدام هذا النوع من العينات فى مجتمعات تتسم بتكرار ظواهر دورية .

ب - اقتصار العشوائية فقط في تحديد الرقم الأول فى بداية اختيار العينة .

مثال : يريد مدير مدرسة أن يعرف رأي طلبة مدرسته في مستوى تدريس أحد المعلمين. إذا كان عدد طلبة المدرسة 10000 طالب، وحجم العينة المطلوبة يساوي 200.

للحصول على العينة بسرعة، يختار من كل $50 = \frac{10000}{200}$ طالبا. فمن أول خمسين طالباً ، يختار طالبا عشوائيا ، وليكن رقم 39. وعليه فإن العينة هي الأشخاص : 39 ، 89 ، 139 ، ... ، 9989.

3- العينات الطبقية : Stratified Samples :

تتميز العينات الطبقية على غيرها من العينات بأنها بالإضافة إلى كونها دراسة للمجتمع ككل فإنها تتيح لنا دراسة كل طبقة من الطبقات على حده وهذا قد يكون مرغوباً فيه فى كثير من الأحيان ففى دراسة لبحث ميزانية الأسرة

نحصل على نتائج البحث لكل من الريف والحضر على حده وهما الطبقتان اللتان يتكون منهما المجتمع ، وبذلك تمكننا العينة الطبقية من دراسة كل من الريف والحضر إلى جانب دراسة المجتمع المصري ككل .

تعتمد هذه الطريقة على تقسيم المجتمع الإحصائي إلى فئات أو طبقات ثم اختيار عينة من كل طبقة ففى المثال السابق يمكن لباحث أن يقسم الأربعة آلاف طالب بحسب أصولهم الحضرية إلى طلاب من المدينة، وطلاب من الريف ، ثم يقوم باختيار عدد من الطلاب الذين ينتمون إلى كل من هذه التقسيمات بطريقة عشوائية ويتحدد عدد الطلاب الذين سيتم اختيارهم من كل طبقة بحسب نسبة تلك الطبقة إلى المجموع الكلي للمجتمع الأصلي فلو فرضنا على سبيل المثال أن 50% من جملة عدد الطلاب وهم 4000 طالب ، من المدن فإن معنى هذا أن 50% من العينة التى حجمها 400 طالب يتم اختيارهم من المدن وهكذا . وعموما يمكن صياغة تلك العلاقة فى القانون التالي :

$$\text{عدد الافراد المراد سحبهم من الطبقة} = \text{حجم العينة} \; X \; \frac{\text{عدد افراد الطبقة}}{\text{حجم المجتمع}}$$

فى هذه الحالة من المعتقد أن خطأ المعاينة من المحتمل أن يتناقص ليصل إلى الصفر . فتوزيع الطلاب بحسب موطنهم الأصلي فضلا عما يعكسه من تباين ثقافي بين الطلاب فإنه يقترب كثيراً من الواقع .

وتقوم العينة الطبقية على تقسيم المجتمع الأصلي إلى مجموعات يطلق عليها طبقات فرعية أو شرائح Strata ثم نأخذ عينة من كل شريحة على حده بحيث يتكون لدينا عينة ذات حجم كلي (n) ومن الأهمية بمكان أن يتحدد تعريف الشريحة الطبقية بضرورة ظهور كل فرد من شريحة واحدة فقط ولا يتكرر فى غيرها . وفى الطريقة البسيطة والشائعة من حيث الاستخدام للعينة الطبقية أن تستخدم فى الاختيار وعند بداية تصميم نموذج العينة الطبقية على الباحث اتخاذ الخطوات التالية :

أ- حساب تقديري للمتوسطات الحسابية لكل شريحة على حده .

ب- حساب تقديري للانحراف المعياري لكل شريحة على حده .

ج- بعد تقدير قيمة (n_i) لكل شريحة نبدأ فى وضع أوزان تبعا لحجم الشريحة ونسبة هذا الحجم للمجتمع الأصلي .

مثال: ولنفترض انه بالرجوع إلى السجلات الرسمية في الجامعة وجدت البيانات المبينة في الجدول التالي:

فئات الطلبة	سنه أولى	سنة ثانية	سنة ثالثة	سنة رابعة	دراسات عليا	المجموع
عدد الطلبة	700	800	700	600	200	3000

في هذه الحالة لابد من عينة طبقية ولنفترض عدد أفراد العينة المطلوبة هو 200 مفردة

الطبقة الأولى هم طلبة السنة الأولى ونختار منهم بإحدى الطرق السالفة الذكر، العشوائية البسيطة أو المنتظمة عددا من المفردات مقداره :

$$\text{عدد مفردات العينة المطلوبة} = \frac{\text{عدد طلبة السنة الاولى}}{\text{مجموع الطلبة الكلي}} X \text{حجم العينة}$$

لذافان الطبقة الاولى(سنة اولى) يمثلهم عينة حجمها هو $47 = 200 \; X \; \frac{700}{3000}$

وان الطبقة الثانية(سنة ثانية) يمثلهم عينة حجمها هو $53 = 200\ X\ \frac{800}{3000}$ وهكذا لبقية الطبقات.

4- العينة غير المتناسبة : Disproportionate Sample يلجأ الباحث عادة إلى مثل هذا النوع من العينات إذا كان يريد أن يرفع نسبة عينة جماعة فرعية معينة . فلو أراد الباحث فى مثلنا السابق أن يعرف رأى الطلاب الذين من أصل قروي فى قضية الإدمان لما يتميزون به من وازع ديني وأخلاقي فإنه فى هذه الحالة يزيد من نسبة تمثيل الطلاب القرويين لأن طبيعة مشكلة البحث تقتضي ذلك فيختار الباحث 200 طالب من المناطق الريفية وباقي الطلاب من المدينة. ولكن فى هذه الحالة ينبغي على الباحث أن يظهر فى تحليله العوامل التى دفعته لمثل هذا النوع من الاختيار .

5- العينات العنقودية ذات المرحلة الواحدة ومتعددة المراحل . Single , stage and Multi stage cluster Samples

فى حالة العينات كبيرة الحجم يلجأ الباحث إلى هذا الأسلوب من أساليب المعاينة لتخفيض نفقات اختيار العينة والعينة العنقودية ذات المرحلة الواحدة تتمثل فيما يقرره احد الباحثين من اختيار حى سكنى معين من إحدى المدن كعينة للدراسة ثم يختار مجموعة من الأسر التى تقطن ذلك الحي لإجراء مقابلة معهم . معنى هذا أن المقابلات التى سيقوم بها الباحث سوف تتجمع فى حي معين الأمر الذي ساعد على تخفيض الوقت والنفقات ونلاحظ هنا أن اختيار العينة تم على مرحلة واحدة .

أما العينة العنقودية متعددة المراحل فيلجأ إليها الباحث عند اختيار عينة أكبر حجما . فلو أردنا أن ندرس اتجاهات الشباب نحو الإدمان فإنه يمكن أن تحصل على خريطة بأحياء المدينة ثم تختار من بينها عددا من الأحياء الشعبية وعددا أخر من الأحياء الراقية ثم تختار عددا من القطاعات داخل الأحياء وبعد ذلك يتم اختيار من تتم مقابلتهم كأفراد داخل العينة . من ذلك يتضح لنا أن أسلوب العينة العنقودية متعددة المراحل وإن كان يحقق الدقة ويرفع درجة تمثيل العينة للمجتمع الأصلي إلا أنه أسلوب يكتنفه التعقيد ولا يستطيع كثير من الباحثين ذوى الإمكانيات المحدودة الاستعانة به .

مثال/ مدير مدرسة يريد أن يعرف رأي طلبة مدرسته في مستوى تدريس أحد المعلمين . فإذا كان عدد طلبة المدرسة 200 طالب، كيف يتم سحب عينة منتظمة بحجم 10 من هذا المجتمع.

الحل: من احل سحب عينة منتظمة نتبع الخطوات التالية :

1-ترقيم عناصر المجتمع من 1 إلى N (اي من 1 الى 200).
2-تحديد حجمِ العيّنةَ n التي نحتاجها بالدراسة وهي n=10.
3-نجد قيمة الزيادة المنتظمة k اذ ان : K = N/n
4-يتم اختيارُ اول مفردة باختيار اي عدد صحيحُ بشكل عشوائي بين 1 إلى K .
5-ثمّ نأخذ العناصر الاخرى عن طريق الزيادة بانتظام بمقدار k الى ان يصبح عدد الارقام المختارة يساوي n اي 10 . وكالتالي:

بما ان حجم المجتمع (N=200) يتم ترقيم عناصر المجتمع من 1 الى 200 وحسب ترميز معين او خاصية محددة ثم نحدد حجم العينة ولتكن (n=10) ،نحسب: $K = \frac{N}{n} = \frac{200}{10} = 20$ ، ثم نختار رقم من 1 الى 20 بصورة عشوائية وليكن 12 عندها نختار العناصر المرقمة التالية الى ان نصل الى 10 ارقام:

12,32,52,72,92,112,132,152,172,192

ثم نختار العناصر في المجتمع التي تحمل هذه الارقام وعددها 10.

مثال/ كلية تحتوي على 1000 طالب ومكونة من 4 اقسام حجم كل قسم على التوالي

من اليسار الى اليمين هي :

(150 ، 350 ، 400 ،100) يراد سحب عينة طبقية حجمها 40 طالب من هذه الكلية ،

كيف يتم ذلك بحيث تمثل هذه العينة تمثيلاً سليماً.

الحل : نلاحظ ان : $N = 1000$ ، $n = 40$

$N4 = 150$, $N3 = 350$, $N2 = 400$, $N1 = 100$

$$n_i = \frac{N_i}{N} x\, n$$

$$n_1 = \frac{N_1}{N} x\, n = \frac{100}{1000} x40 = 4$$

$$n_2 = \frac{N_2}{N} x\, n = \frac{400}{1000} x40 = 16$$

$$n_3 = \frac{N_3}{N} x\, n = \frac{350}{1000} x40 = 14$$

$$n_4 = \frac{N_4}{N} x\, n = \frac{150}{1000} x40 = 6$$

$$n = n1 + n2 + n3 + n4$$

$$= 4 + 16 + 14 + 6 = 40$$

المتغيرات العشوائية:

هي دالة ذات قيمة حقيقية معرفة في فضاء العينة لأي تجربة عشوائية كما يعرف المتغير العشوائي ايضاً بأنه مقدار له خصائص رقمية(كمية)وغير رقمية(وصفية)تتغير قيمته من عنصر إلى آخر من عناصر المجتمع أو العينة وعادة ما يرمز لها بالأحرف الانكليزية الكبيرة مثل X,Y,Z وتقسم المتغيرات العشوائية الى قسمين هما:

1- **المتغيرات الوصفية:** وهي المتغيرات التي تكون مفرداتها غير قابلة للقياس بالأرقام بل تكون على شكل صفات او مسميات مختلفة وتنقسم إلى متغير اسمي(لون العين، الجنس، الجنسية)ومتغير رتبي(ممتاز، جيد جداً، جيد) او مثل (موافق بشدة ، موافق، محايد ، غير موافق ، غير موافق بشدة).

2- **المتغيرات الكمية:** وهي المتغيرات التي تكون مفرداتها قابلة للقياس بالأرقام مثل اعداد طلبة الجامعات العراقية ،الطول ،العمر ...الخ وتقسم الى قسمين هما:

ا) **المتغيرات المتقطعة :** 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) المتغيرات المستمرة: وهي المفردات التي تأخذ مدى معين او مجال معين وتكون غير قابلة للعد والحساب و تتقبل الكسور مثل اعمار الافراد او اطوالهم او درجاتهم.

تبويب البيانات :

يقصد بتبويب البيانات عرض هذه البيانات (البيانات الخام) فى جداول مناسبة وذلك حتى يمكن تلخيصها وفهمها واستيعابها واستنتاج النتائج منها ومقارنتها بغيرها من البيانات ، كما يسهل الرجوع إليها فى صورة جداول دون الاطلاع على الاستمارات الأصلية التى قد تحمل أسماء أصحابها مما يخل بمبدأ سرية البيانات الإحصائية .

كما يعتبر عرض وتبويب البيانات الإحصائية الخطوة الثانية (بعد تجميع هذه البيانات الخام) في مفهوم التحليل الإحصائي، ويلجأ الباحث إلى حصر وتصنيف هذه البيانات وعرضها بطريقة مختصرة تساعد على فهمها وتحليلها إحصائياً للتعرف عليها ووصفها ومقارنتها بغيرها من الظواهر ، والخروج ببعض المدلولات الإحصائية عن مجتمع الدراسة .

وفيما يلي عرض موجز لاربعة انواع من التبويب:

1- التبويب الزمني:

عبارة عن تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل جمع منها يعود لوحدة زمنية معينة كاليوم ،الاسبوع ، الشهر ، السنة ،...الخ.

مثال/اذا كان لدينا 30 موظف وان توزيع عدد الموظفين حسب سنوات التخرج من الجامعة موضح كما يلي:

سنة التخرج	العدد
2000	7
2001	9
2002	6
2003	5
2004	3
المجموع	30

2- التبويب الجغرافي:

عبارة عن تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل جمع منها خاص بوحدة جغرافية معينة او تقسيم اداري معين كالنواحي ،الاقضية ، المحافظات ، البلدان ،...الخ.

مثال/ اذا كان لدينا 30 موظف وان توزيع عدد الموظفين حسب الجامعة موضح كما يلي:

الجامعة المانحة للشهادة	العدد
جامعة بغداد	5
جامعة الموصل	3
الجامعة المستنصرية	3
جامعة البصرة	10
جامعة الكوفة	9
المجموع	30

3- التبويب الكمي:

عبارة عن تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل جمع منها خاص بوحدة كمية معينة كوحدات الوزن ،الطول، المسافة ، الحجم ،...الخ

مثال/ اذا كان لدينا 30 موظف وان توزيع عدد الموظفين حسب سنوات الخدمة في الوظيفة موضح كما يلي:

سنوات الخدمة	العدد
اقل من 5 سنوات	7
5-10	10
10-15	8
15-20	5
المجموع	30

4- التبويب على اساس صفة معينة:

عبارة عن تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل جمع منها يشترك بصفة معينة كالجنس،الحالة الاجتماعية، عنوان الوظيفة ، القومية ،...الخ.

مثال/ اذا كان لدينا 30 موظف وان توزيع عدد الموظفين حسب الجنس موضح كما يلي:

الجنس	العدد
ذكر	17
انثى	13
المجموع	30

عرض البيانات الاحصائية:

ان البيانات الخام التي يتم جمعها عن ظاهرة معينة لا يمكن وصفها وتفسيرها وهي في هيئتها الاولى (كبيانات اولية) الامر الذي يتطلب تلخيص ووضع هذه البيانات في جداول خاصة او رسوم بيانية هندسية بهدف تسهيل دراستها واجراء التحليل الاحصائي عليها.

يوجد طريقتان لعرض البيانات المصنفة احصائيا وهذه الطريقتان هما:

الطريقة الاولى : العرض الجدولي للبيانات:

وهو عبارة عن تمثيل ووصف البيانات وتنظيمها في جداول خاصة حسب اشتراكها في صفة معينة.

التوزيع التكراري: و هو عبارة عن ترتيب بيانات المتغير العشوائي الكمي بنوعيها (المستمرة والمتقطعة) حسب اشتراكها بصفة معينة في جداول خاصة تسمى بجداول التوزيع التكراري (Frequency Distribution Tables) مقسمة الى فئات (Classes) وتكرارات (Frequencies) وتكون الفئات على نوعين فئات مغلقة او مفتوحة حسب طبيعة البيانات .

خطوات وضع البيانات في جدول توزيع تكراري

1- تحديد اكبر قيمة في البيانات ويرمز لها ب (V_L) واصغر قيمة فيها ويرمز لها ب (V_S).

2- ايجاد المدى الكلي للبيانات (Total Range) ويرمز له ب (T.R) وفق الصيغة التالية:

$$T.R=V_L-V_S+1$$

3- ايجاد عدد الفئات وفق احدى الصيغ التالية

$$K=1+3.322*Log(n)$$

$$K = 2.5 * \sqrt[4]{n}$$

4-ايجاد طول الفئة وهو عبارة عن مقدار سعة الفئة من الارقام وفق الصيغة التالية:

L= T.R/K

5 – تحديد الحد الادنى (Lower Limit) لكل فئة والحد الاعلى (Upper Limit) لها كما يأتي:

أ- في حالة المتغير المتقطع:

L.L=V_S

U.L=L.L + L - 1

ب- في حالة المتغير المستمر

L.L=VS

U.L=L.L+L

6- تحديد تكرار الفئة (Class Frequency) وهو عبارة عن جميع المفردات التي تقع ضمن حدود الفئة من حيث القيمة العددية ويرمز له ب (fi) .

7- ايجاد مركز الفئة (Class Midpoint) وهو القيمة التي تقع في منتصف الفئة ويرمز له ب (Mi) وفق الصيغة التالية: Mi=(L.L+U.L)/2

فيكون شكل جدول التوزيع التكراري كالتالي:

1- في حالة المتغير المتقطع يسمى بالتوزيع المغلق أي ان جميع الفئات لها حدود دنيا وعليا كما يأتي:

تسلسل الفئة	الفئات	طول الفئة	التكرار
1	$(V_S) – (V_S+L-1)$	L	f_1
2	$(V_S+L) – (V_S+2L-1)$	L	f_2
3	$(VS+2L) – (VS+3L-1)$	L	f_3
.	.	L	
.	.		
.	.		
.			f_K
K	$(VS+(K-1)L) – (VS+KL-1)$	L	
المجموع			N

2- في حالة المتغير المستمر ويسمى بالتوزيع المفتوح حيث ان جميع الفئات لديها حدود دنيا فقط ماعدا الفئة الاخيرة التي تكون مغلقة أي ان لها حد ادنى واعلى كما ياتي:

تسلسل الفئة	الفئات	طول الفئة	التكرار

1	(V_S) –	L	f_1
2	(V_S+L) –	L	f_2
3	(VS+2L) –	L	f_3
.	.	L	
.	.		
.	.		
.			
K	(VS+(K-1)L) – (VS+KL)	L	f_K
المجموع			N

مثال (1) : حالة المتغير المتقطع

البيانات التالية تمثل عدد اشجار النخيل التي تمتلكها (20) عائلة فلاحية في مدينة البصرة

Xi: 65 , 56 , 89 , 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

الحل:

(1) VS=45,VL=89

(2) T.R=VL-VS+1=89-45+1=45

(3) K=1+3.3*Log(n)=1+3.3*Log(20)=5.29=5

(4) L= T.R/K= 45/5=9

(5)

لحساب الفئة الاولى : $(V_S) – (V_S+L-1)$

VS=45الحد الادنى للفئة الاولى ، V_S+L-1=45+9-1= 53 الحد الاعلى للفئة الاولى

اذا الفئة الاولى : (45-53)

لحساب الفئة الثانية: $(V_S+L) – (V_S+2L-1)$

VS+L=45 +9 = 54 الحد الادنى للفئة الثانية ، V_S+2L-1=45+18-1= 62 الحد الاعلى للفئة الثانية

اذا الفئة الثانية : (54-62)

وهكذا لبقية الفئات.... الى ان نصل الاخيرة: (VS+(K-1)L) – (VS+KL-1)

(45+(5-1)*9) – (45+(5*9) -1)

(45+36) –(45+45-1)

(81–89)

والآن نضع البيانات في جدول تكراري كالتالي:

عدد الاشجار(الفئات)	التكرار بالإشارات	التكرار	مراكز الفئات
45 –53	////	4	M1=(45+53)/2=49
54–62	𝍸//	4	M2=(54+62)/2=58
63–71	𝍸// /	7	M3=(63+71)/2=67

M4=(72+80)/2=76	3	///	72–80
M5=(81+89)/2=85	2	//	81–89
	20	n=20	المجموع

يتضح من الجدول أعلاه أن كل فئة تمثل عدد الأشجار وان التكرارات تمثل عدد العوائل التي تمتلكها.

مثال (2):حالة المتغير المستمر

البيانات التالية تمثل درجات (16) طالبا في احدى المواد
Xi: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.

(1)VS=50 , VL=84
(2)T.R=VL-VS+1=84-50+1=35
(3)K=2.5*$(n)^{1/4}$=2.5*$(16)^{1/4}$= 2.5*2 = 5
(4)L= T.R/K= 35/5=7
(5)

VS=50الحد الادنى للفئة الاولى ، V_S+L=50+7= 57الحد الادنى للفئة الثانية
V_S+2L=50+14= 64الحد الاعلى للفئة الثالثة

وهكذا لبقية الفئات الاخرى الى ان نصل للفئة الاخيرة: (VS+(K-1)L) – (VS+KL)

VS+kL=50 + (5)(7) = 85 الحد الادنى للفئة الاخيرة ، V_S+(k-1)L=50+(4)(7)= 78 الحد الاعلى للفئة الاخيرة

مع ملاحظة ان مركز الفئات يحسب كالتالي : $M_i = \frac{\text{الحد الادنى للفئة الحالية+الحد الادنى للفئة اللاحقة}}{2}$

والان نضع البيانات في جدول تكراري كالتالي:

مراكز الفئات	التكرار	التكرار بالإشارات	الدرجات (الفئات)
M1=(50+57)/2=53.5	1	/	50 –
M2=(57+64)/2=60.5	2	//	57 –
M3=(64+71)/2=67.5	3	///	64 –
M4=(71+78)/2=74.5	6	~~////~~ /	71–
M5=(78+85)/2=81.5	4	////	78 –85
	16	n=16	المجموع

يتضح من الجدول اعلاه ان كل فئة تمثل الدرجات وان التكرارات تمثل عدد الطلاب الذين حصلوا عليها.

التوزيع التكراري التراكمي او المتجمع:

يهتم هذا النوع من التوزيعات بتحديد القيم التي تقل او تزيد عن قيمة معينة مقابل كل فئة من فئات التوزيع وتكون التوزيعات التكرارية المتجمعة على نوعين هما:

ا) التكرار المتجمع الصاعد:

يمكن الحصول على التكرار المتجمع الصاعد من خلال تجميع او تراكم تكرارات الجدول الاصلي بدءا بتكرار الفئة الاولى وانتهاء بتكرار الفئة الاخيرة منه الى ان نحصل على مجموع التكرارات كتكرار متجمع صاعد للفئة الاخيرة والشكل التالي يوضح التصميم العام لهذا النوع من التوزيع ويرمز له ب (Fi)

تسلسل الفئة	التكرار	الحدود العليا للفئات	التكرار المتجمع الصاعد
1	f_1	(V_S+L-1)	f_1
2	f_2	(V_S+2L-1)	f_1+f_2
3	f_3	$(VS+3L-1)$	$f_1+f_2+f_3$
.		.	
.		.	
.		.	
.	f_K		
K		$(VS+(K-1)L)$	$f_1+f_2+f_3+\ldots..+f_k=n$

مثال (3) : بالعودة لبيانات مثال رقم (1) جد التكرار المتجمع الصاعد.

عدد الاشجار(الفئات)	التكرار	الحدود العليا للفئات	التكرار المتجمع الصاعد
45 –53	4	اقل او يساوي 53	4
54–62	4	اقل او يساوي 62	8
63–71	7	اقل او يساوي 71	15
72–80	3	اقل او يساوي 80	18
81–89	2	اقل او يساوي 89	20
المجموع	20		

يتضح من الجدول أعلاه إن عدد العوائل التي تمتلك اقل من 71 نخلة هو 15 وهكذا.

ب) التكرار المتجمع النازل:

يمكن الحصول على التكرار المتجمع النازل من خلال طرح تكرارات الجدول الاصلي من مجموع التكرارات على التوالي بدءا بتكرار الفئة الاولى وانتهاء بتكرار الفئة الاخيرة منه الى ان نحصل على التكرار الاخير كتكرار متجمع نازل للفئة الاخيرة والشكل التالي يوضح التصميم العام لهذا النوع من التوزيع ويرمز له ب (Fi)

تسلسل الفئة	التكرار	الفئات	التكرار المتجمع النازل
1 2 3 . . . K	f_1 f_2 f_3 f_K	(V_S) (V_S+L) (VS+2L) . . . (VS+(K-1)L)	n n-f_1 n-f_1-f_2 n-f_1-f_2-f_3-…-f_{k-1}=f_k
المجموع	n		

مثال (4) : بالعودة لبيانات مثال رقم (1) جد التكرار المتجمع النازل.

عدد الأشجار(الفئات)	التكرار	الحدود الدنيا للفئات	التكرار المتجمع النازل
45 –53	4	45 فاكثر	20
54–62	4	54 فاكثر	16
63–71	7	63 فاكثر	12
72–80	3	72 فاكثر	5
81–89	2	81 فاكثر	2
المجموع	20		

يتضح من الجدول اعلاه ان عدد العوائل التي تمتلك اكثر من 63 نخلة هو 12 وهكذا.

التوزيع التكراري النسبي المئوي

وهو توزيع اعتيادي تكون التكرارات فيه على شكل تكرارات نسبية مئوية ويرمز له ب (F^*i) ويمكن الحصول على التكرار النسبي المئوي وفق الصيغة التالية:

$$F_i = \frac{f_i}{n} * 100\%$$

مثال (5) : بالعودة لبيانات مثال رقم (2) جد التكرار النسبي المئوي.

الدرجات (الفئات)	التكرار	التكرار النسبي المئوي
50 –	1	6.25%
57 –	2	12.5%
64 –	3	18.75%
71–	6	37.5%
78 –85	4	25%
المجموع	16	%100

يتضح من الجدول أعلاه إن نسبة 25% من الطلاب تتراوح درجاتهم بين 78 و 85 .

الطريقة الثانية :العرض الهندسي للبيانات

وهو عبارة عن تمثيل ووصف البيانات التي يتم جمعها عن ظاهرة معينة بواسطة أشكال بيانية أو رسوم هندسية بهدف إعطاء فكرة واضحة وسهلة وسريعة عن بيانات الظاهرة المدروسة ويمكن استخدام الرسوم البيانية للبيانات الاعتيادية والبيانات المبوبة على حد سواء.

أولا: العرض الهندسي للبيانات الاعتيادية:

إن البيانات الاعتيادية تعني البيانات التي لا تكون معروضة بشكل جدول توزيع تكراري ويتم تمثيلها بيانيا بالأشكال التالية:

1- المستطيل البياني Rectangle Chart

تتلخص فكرة هذا الشكل باختيار مستطيل ذو قاعدة مناسبة يتم افتراضها ثم يقسم المستطيل الى مستطيلات جزئية تمثل بيانات الصفة المدروسة وفق الصيغة التالية:

طول قاعدة المستطيل الجزئي =(البيانات الجزئية/ البيانات الكلية) * طول قاعدة المستطيل الكلي

مثال (6):بلغت التكاليف الإنتاجية لإنتاج سلعة معينة (300) دولار كما موضحة بالجدول التالي:

مستلزمات الإنتاج	التكاليف/دولار
أجور	120
مواد خام	60
مصاريف مباشرة	90
مصاريف غير مباشرة	30

الحل: نفرض أن طول قاعدة المستطيل الكلي = 10

طول قاعدة المستطيل الجزئي = (البيانات الجزئية/ البيانات الكلية) * طول قاعدة المستطيل الكلي

طول قاعدة المستطيل الأول (الأجور) = (300/120) *10= 4

طول قاعدة المستطيل الثاني(المواد الخام) =(300/60) *10= 2

طول قاعدة المستطيل الثالث (المصاريف المباشرة)= (300/ 90) *10= 3

طول قاعدة المستطيل الرابع (المصاريف غير المباشرة)= (300/30) *10= 1

الأجور	المواد الخام	المصاريف المباشرة	المصاريف غير المباشرة

2- الاشرطة البيانية Bar- Charts

وهي عبارة عن مجموعة من المستطيلات الافقية او العمودية قواعدها متساوية وتمثل الصفة التي تم على اساسها التبويب (سنة ،محافظة ، وهكذا) وارتفاعاتها تمثل البيانات المقابلة لتلك الصفة والاشرطة البيانية على نوعين هما:

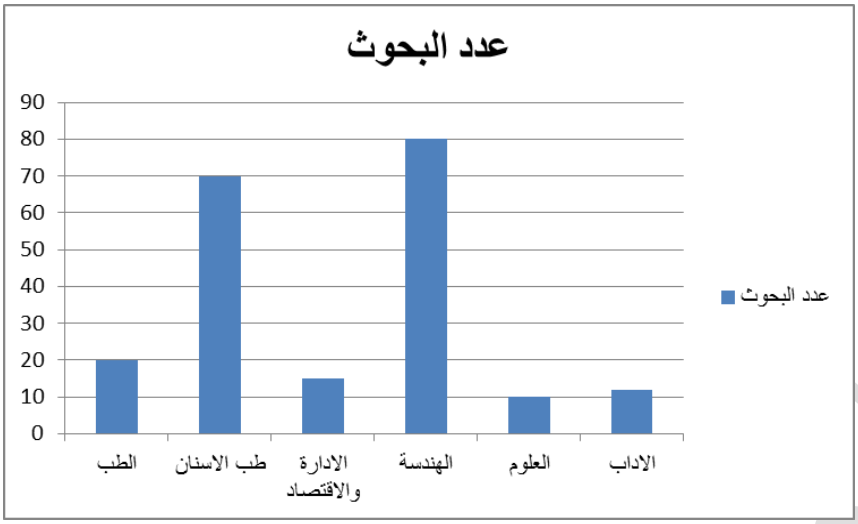
ا) الاشرطة البيانية المفردة

وهي اشرطة بيانية تخص صنف واحد للبيانات مثل عدد الطلبة المقبولين او تطور عدد سكان العراق حسب التعداد السكاني

مثال(7): البيانات الواردة في الجدول التالي تمثل عدد البحوث العلمية المنجزة من قبل اعضاء هيئة التدريس في جامعة بابل موزعين حسب كلياتهم لعام 2006

الكليات	عدد البحوث
الطب	**20**
طب الاسنان	**70**
الادارة والاقتصاد	**15**
الهندسة	**80**
العلوم	**10**
الآداب	**12**

الحل: يتم رسم الاشرطة البيانية بعد وضع المحور الافقي(السيني) X والذي يمثل الكليات في هذا المثال والمحور العمودي (الصادي)Y والذي يمثل عدد البحوث بعد ان يؤخذ تقسيم مناسب للمحور (Y) كما يأتي :



ب) الاشرطة البيانية المركبة

وهي اشرطة بيانية تخص صنفين او اكثر للبيانات مثل عدد الكتب الموجودة في احدى المكتبات مصنفة حسب انواعها الى كتب علمية وادبية وتاريخية وغيرها او عدد الموظفين في احدى الدوائر مصنفين حسب درجاتهم الوظيفية.

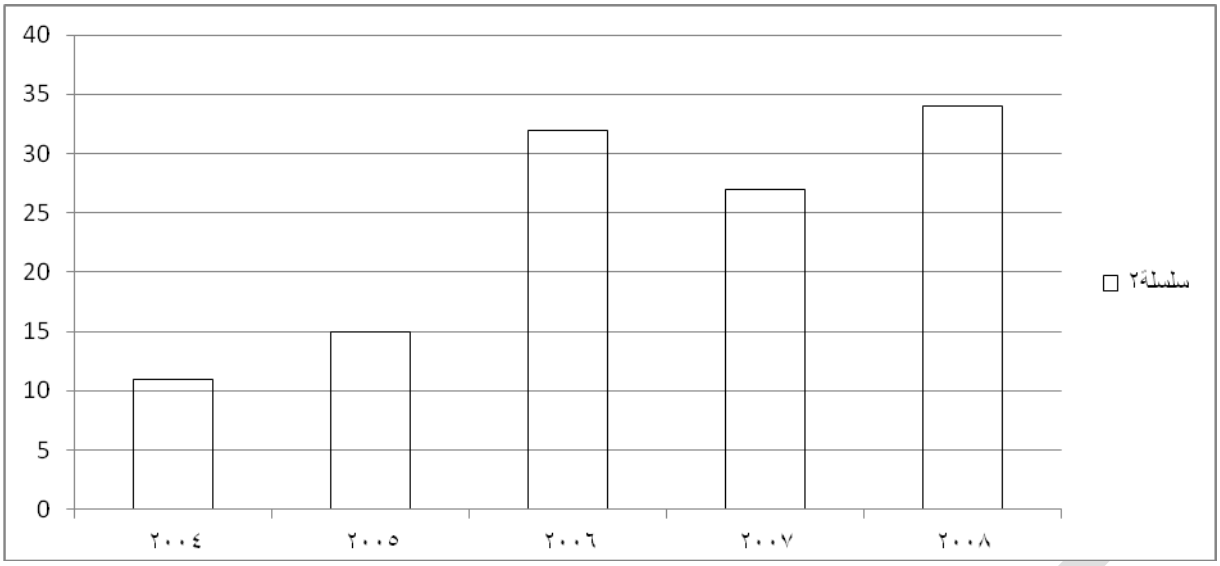
مثال (8) : البيانات التالية تمثل عدد الندوات والمؤتمرات العلمية التي عقدتها كليات جامعة معينة:

السنوات	2004	2005	2006	2007	2008
عدد الندوات	9	14	23	15	18
عدد المؤتمرات	2	1	9	12	16

الحل : نعمل الجدول التالي

السنوات	2004	2005	2006	2007	2008
عدد الندوات	9	14	23	15	18
عدد المؤتمرات	2	1	9	12	16
المجموع	11	15	32	27	34

يتم رسم الاشرطة البيانية بعد وضع المحور الافقي(السيني) X والذي يمثل السنوات في هذا المثال والمحور العمودي (الصادي)Y والذي يمثل المجموع بعد ان يؤخذ تقسيم مناسب للمحور (Y) ومن ثم يرسم شريط ضخم قاعدته السنة وارتفاعه المجموع وبداخله شريطين صغيرين احدهما لعدد الندوات والاخر لعدد المؤتمرات كما يأتي:



3- الدائرة البيانية Pie-chart

وهي عبارة عن شكل هندسي مثل المستطيل البياني ولكن يتم هنا تمثيل البيانات بقطاعات داخل دائرة بحيث ان مجموع هذه القطاعات تمثل مساحة الدائرة الكلية ويتم تحديد زاوية القطاع وفق الصيغة التالية:

زاوية القطاع = (البيانات الجزئية/ البيانات الكلية) * 360

مثال (9): بالعودة الى بيانات مثال رقم 6 ارسم الدائرة البيانية

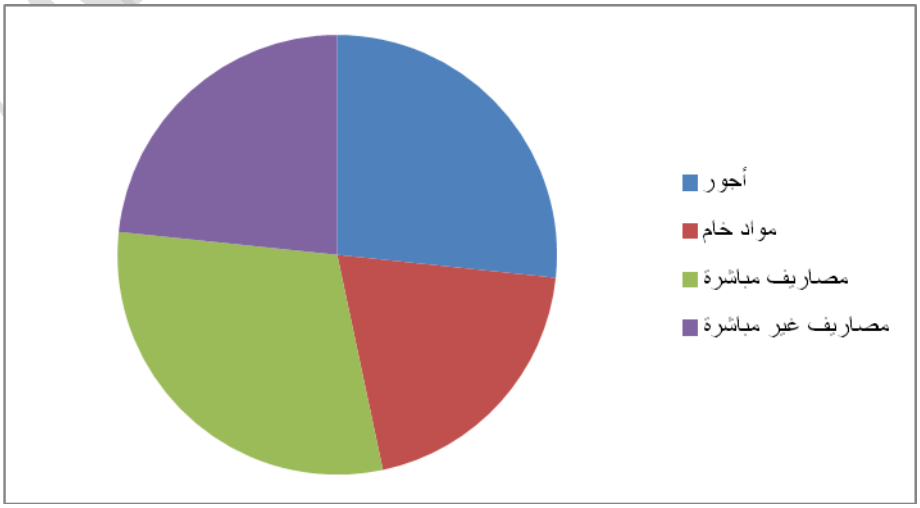
الحل:

زاوية قطاع الاجور = (120/300)* 360° =360°* 0.4 =144°

زاوية قطاع المواد الخام = (60/300)* 360° = 360°*0.2 =72°

زاوية قطاع المصاريف المباشرة = (90/300)* 360° =360°*0.3= 108 °

زاوية قطاع المصاريف غير المباشرة = (30/300)* 360° =360°*0.1= 36°

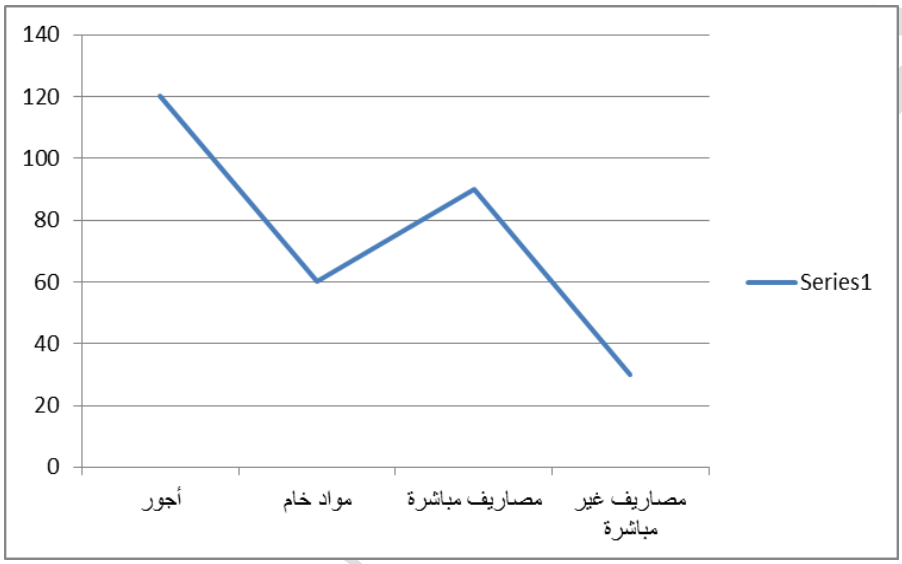


4- الخط البياني Line-chart

عبارة عن شكل بياني يوضح التغيرات الحاصلة في ظاهرة معينة عبرة فترة محددة من الزمن وهو شكل نافع في اجراء مقارنة بين ظاهرتين او اكثر مقاسة بنفس وحدات القياس على سبيل المثال مقارنة التغيرات الحاصلة بين كميات النفط المنتجة والمصدرة او مقارنة تكاليف انتاج سلعة معينة والارباح المتحققة من مبيعاتها خلال مدة زمنية معينة وغيرها من الامثلة الاخرى.

مثال (10): بالعودة الى بيانات مثال رقم 6 ارسم الخط البياني

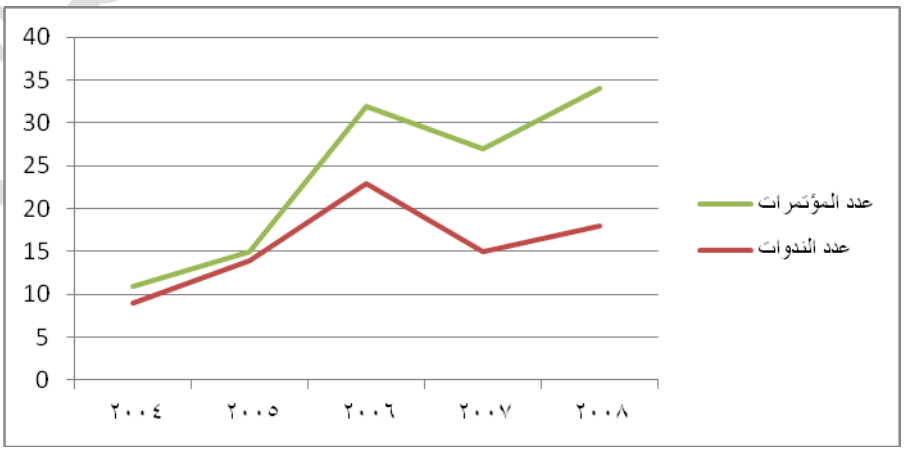
الحل: يتم رسم الخط البياني بعد وضع المحور الافقي(السيني) X والذي يمثل مستلزمات الإنتاج في هذا المثال والمحور العمودي (الصادي)Y والذي يمثل التكاليف بعد ان يؤخذ تقسيم مناسب للمحور (Y)



ملاحظة يمكن رسم الخط البياني لصنفين او اكثر من البيانات كما يأتي:

السنوات	2004	2005	2006	2007	2008
عدد الندوات	9	14	23	15	18
عدد المؤتمرات	2	1	9	12	16

بالعودة لبيانات مثال (8) / يكون الخط البياني كالتالي :



:

ثانيا: العرض الهندسي للبيانات المبوبة:

إن البيانات المبوبة تعني البيانات التي تكون معروضة بشكل جدول توزيع تكراري ويتم تمثيلها بيانيا بالأشكال التالية:

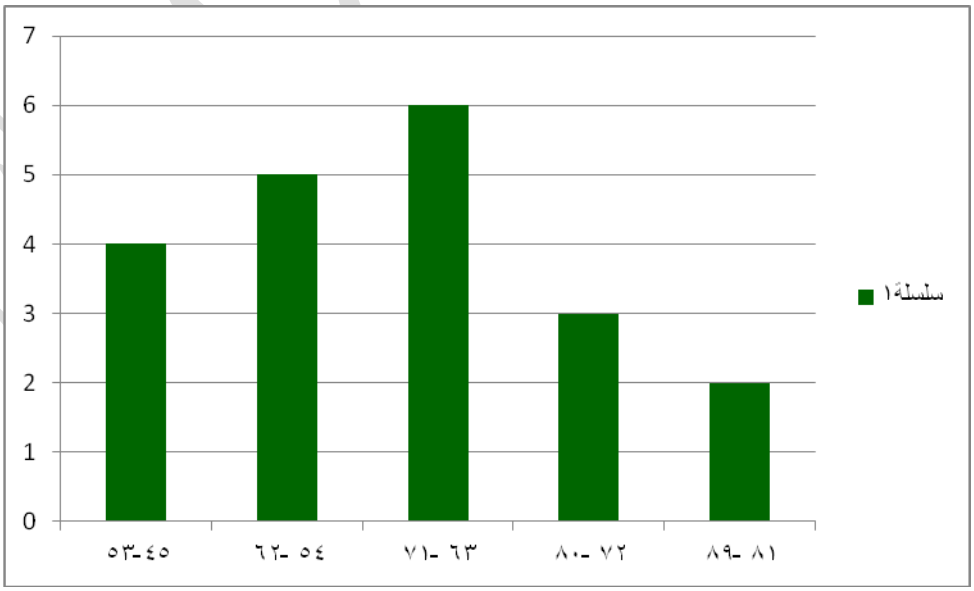
1- المدرج التكراري Histogram

وهو عبارة عن مجموعة من المستطيلات قاعدة كل منها تمثل طول الفئة في التوزيع التكراري وارتفاعها يمثل التكرار المقابل لتلك الفئة أي ان المحور السيني (X)تستقر فيه الفئات والمحور الصادي (Y) تستقر فيه التكرارات هذه المستطيلات تكون منفصلة عن بعضها في حالة المتغير المتقطع ومتصلة مع بعضها في حالة المتغير المستمر وحسب تسلسل فئات التوزيع.

مثال (11) للتوزيع التكراري التالي ارسم المدرج التكراري

عدد الأشجار	التكرار
45 –53	4
54 –62	5
63 –71	6
72 –80	3
81 –89	2

الحل: المحور السيني (X) تستقر فيه الفئات (عدد اشجار النخيل)والمحور الصادي (Y) تستقر فيه التكرارات (عدد العوائل التي تملكها).



2- المضلع التكراري Polygon

وهو عبارة عن عدد من المستقيمات التي تتصل ببعضها بواسطة نقاط هذه النقاط تمثل مراكز الفئات أي ان المحور السيني (X) تستقر فيه مراكز الفئات والمحور الصادي (Y) تستقر فيه التكرارات مع مراعاة غلق المضلع بمركزي فئة وهميين (مركز الفئة الاولى – طول الفئة ومركز الفئة الاخيرة + طول الفئة) وبتكرارين مساويين للصفر.

مثال (12) للتوزيع التكراري في مثال (11) ارسم المضلع التكراري

الحل: المحور السيني (X)تستقر فيه مراكز الفئات والمحور الصادي (Y) تستقر فيه التكرارات

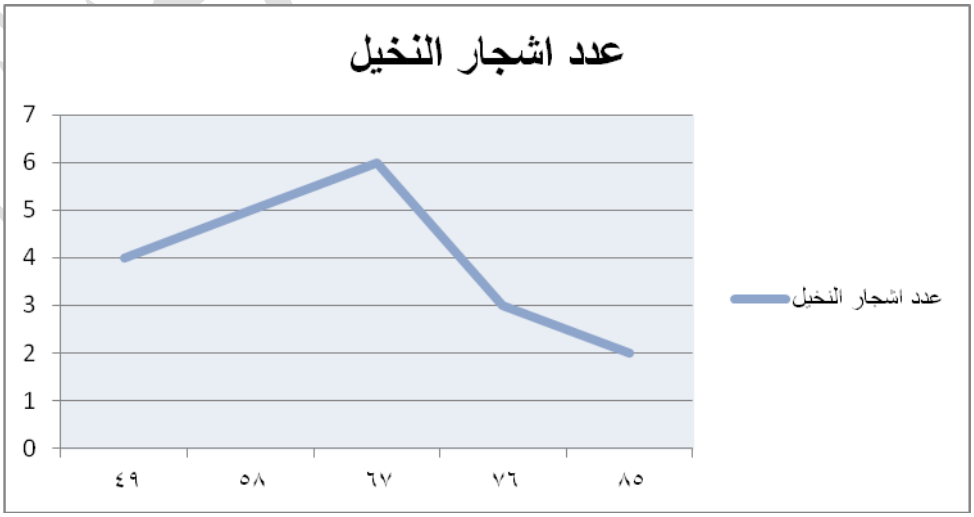
مراكز الفئات	التكرار	عدد الأشجار
49	4	45 –53
58	5	54 –62
67	6	63 –71
76	3	72 –80
85	2	81 –89

بعد ذلك نحدد مركزي فئة وهميين من اجل غلق المضلع وذلك بطرح طول الفئة من مركز الفئة الاولى واضافة طول الفئة لمركز الفئة الاخيرة:

Mo=49-9=40

M6=85+9=94

فيكون المضلع بالشكل التالي:



3- المنحنى التكراري Curve:

هو طريقة شائعة في الرسم البياني وهو عبارة عن منحنى بدلا من الخطوط المستقيمة يمر بمعظم النقاط الواقعة على مراكز الفئات والتي ارتفاعها يمثل تكرارات تلك الفئة.

خطوات رسم المنحنى التكراري :

1- نجد مراكز الفئات

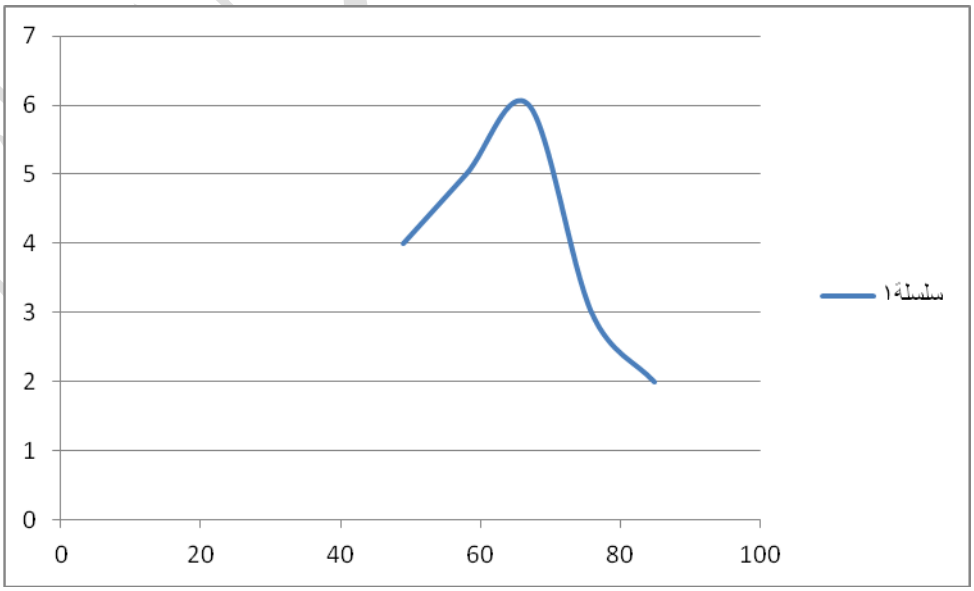
2- نرسم الاحداثيين الافقي (الفئات) والعمودي (التكرارات) ثم نعين النقاط فوق مراكز الفئات ونصل بينها بمنحنى مستمر

مثال (13) للتوزيع التكراري في مثال (11) ارسم المضلع التكراري

الحل: المحور السيني (X) تستقر فيه مراكز الفئات والمحور الصادي (Y) تستقر فيه التكرارات

عدد الأشجار	التكرار	مراكز الفئات
45 –53	4	49
54 –62	5	58
63 –71	6	67
72 –80	3	76
81 –89	2	85

فيكون المنحنى بالشكل التالي:

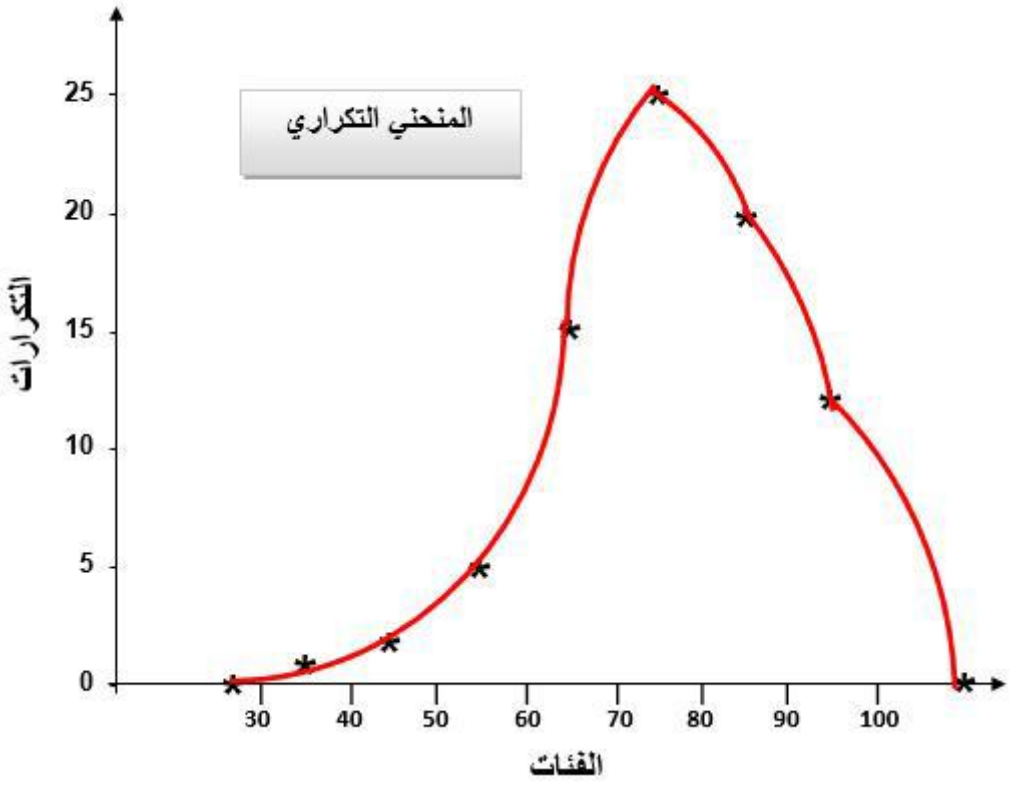


مثال/ ارسم المنحنى التكراري للبيانات التالية:

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

الحل/نستخرج مراكز الفئات

الفئة	30-40	40-50	50-60	60-70	70-80	80-90	90-100	المجموع
التكرار	1	2	5	15	25	20	12	80
مراكز الفئات	35	45	55	65	75	85	95	



أ- المنحنى التكراري المتجمع الصاعد:

لرسم هذا المنحنى نتبع الخطوات الاتية:

1-نكون جدولا تكراريا متجمعا صاعدا من الجدول التكراري البسيط

2-نرصد نقطا احداثياتها الافقية الحدود العليا للفئات واحداثياتها العمودية التكرار المتجمع الصاعد ونصل هذه النقاط ببعضها بخط منحني يكون هو المنحنى المتجمع الصاعد وتسري هذه الخطوات على الجداول الغير منتظمة بدون ان نعدل التكرارات وذلك لان رسم المنحنى المتجمع الصاعد او النازل لتوزيع فئات غير متساوية لا يستدعي تعديل التكرارات.

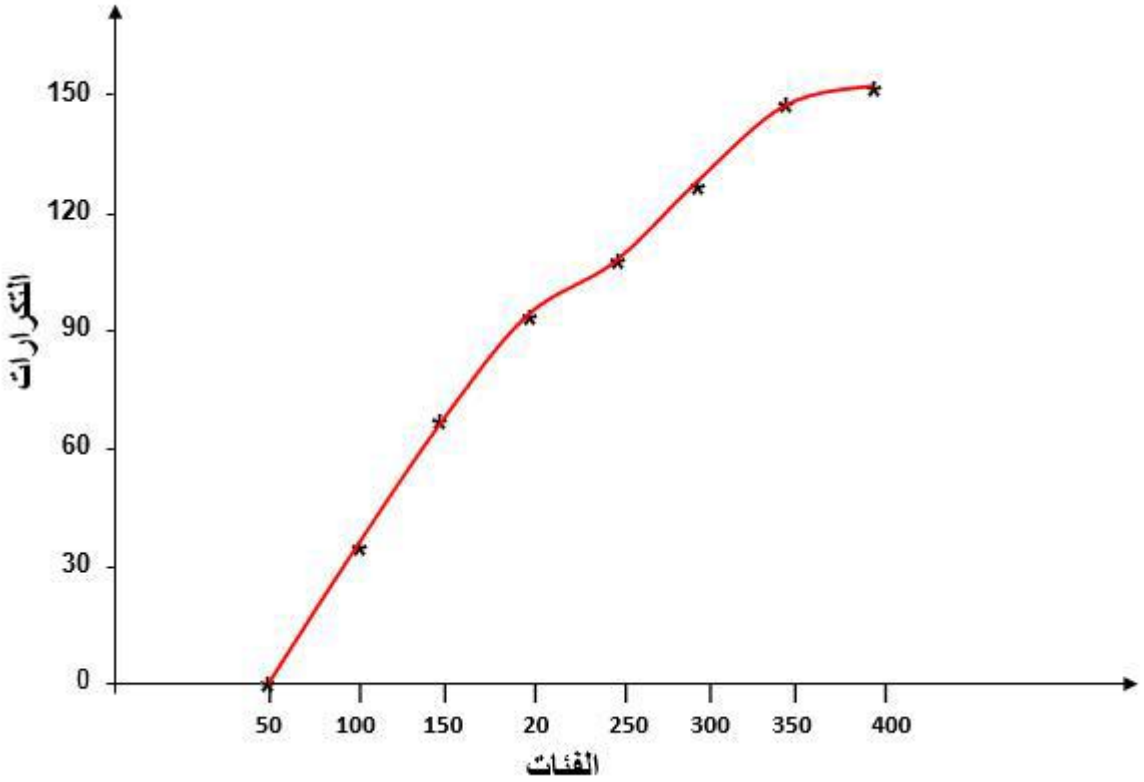
مثال : (4) التوزيع الاتي يمثل ما تدفعه 150 عائلة فلاحية للإيجار سنويا . المطلوب /رسم منحنى متجمع صاعد لهذا التوزيع.

الفئات	50-100	100-150	150-200	200-250	250-300	300-350	350-400	المجموع
التكرارات	32	35	25	20	17	13	8	150

الحل/نكون جدول التكرار المتجمع الصاعد

(الفئات)	التكرار	الحدود العليا للفئات	التكرار المتجمع الصاعد
50-100	32	اقل من 100	32
100-150	35	اقل من 150	67
150-200	25	اقل من 200	92
200-250	20	اقل من 250	112
250-300	17	اقل من 300	129
300-350	13	اقل من 350	142
350-400	6	اقل من او يساوي 400	150
المجموع	150		

نرسم المنحنى التكراري الصاعد



المنحنى التكراري المتجمع النازل:

لرسم المنحنى من الجدول البسيط المنتظم وغير المنتظم نتبع الخطوات الاتية:

1-نكون جدولا تكراريا متجمعا نازلا من الجدول التكراري البسيط

2-نرصد نقاطا احداثياتها الأفقية الحدود الدنيا للفئات واحداثياتها العمودية التكرارات المتجمعة النازلة ثم نصل هذه النقاط بعضها ببعض بخط منحني فيكون هو المنحنى التكراري المتجمع النازل.

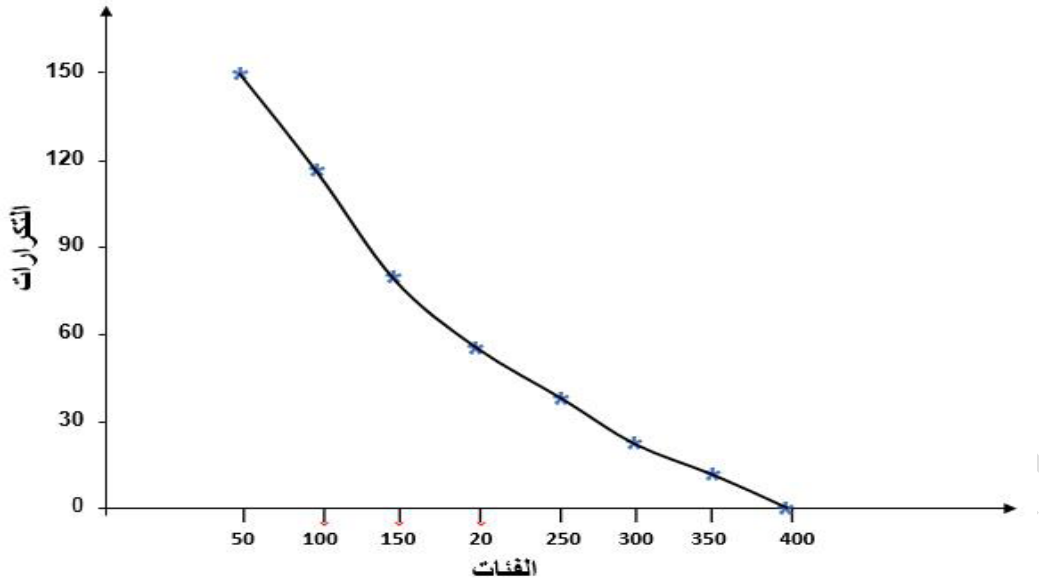
مثال : (5) التوزيع الاتي يمثل ما تدفعه 150 عائلة فلاحية للإيجار سنويا . المطلوب /رسم منحنى متجمع نازل لهذا التوزيع.

الفئات	50-100	100-150	150-200	200-250	250-300	300-350	350-400	المجموع
التكرارات	32	35	25	20	17	13	8	150

الحل/نكون جدول التكرار المتجمع النازل

(الفئات)	التكرار	الحدود العليا للفئات	التكرار المتجمع الصاعد
50-100	32	اقل من 100	150
100-150	35	اقل من 150	118
150-200	25	اقل من 200	83
200-250	20	اقل من 250	58
250-300	17	اقل من 300	38
300-350	13	اقل من 350	21
350-400	6	اقل من او يساوي 400	8
المجموع	150		

نرسم المنحنى التكراري النازل كالتالي:



مفهوم المتوسطات والهدف من احتسابها:

يمكن تمثيل مجموعة من البيانات بقيمة واحدة فقط الهدف من ذلك اعطاء صورة سريعة عن ماهية تلك المجموعة من خلال ايجاد عدد يمثل قيمها.ان المقياس الذي يختص بتحديد هذا العدد يسمى مقياس نزعة مركزية او مقياس توسط (متوسط). هذا العدد يميل ان يقع في وسط تلك المجموعة من البيانات في حال ترتيبها حسب صغرها او كبرها ، اي ما نعنيه ان هذا العدد يؤول لان يتمركز وسط المجموعة التي احتسب منها.هذا الامر جعلنا ان نطلق على هذا النوع من المقاييس ب "مقاييس نزعة مركزية" . ان للمتوسطات اهمية كبرى في موضوع الاستدلال الاحصائي من خلال تقدير قيم عددية لبعض مؤشرات المجتمع تحت الدراسة والبحث التي غالباً ما تكون غير معلومة ، اي ما نعنيه دراسة خصائص مجتمع الدراسة من خلال خصائص العينة التي تعتبر المتوسطات واحدة منها . كذلك تبرز اهمية المتوسطات من الناحية العملية في الكثير من جوانب الحياة المختلفة ،فالخطط الاقتصادية للدولة التي من شأنها رفع مستوى رفاهية الشعب غالباً ما يكون الهدف منها رفع "متوسط" دخل الغرد الى حالة افضل ،والخطط والبرامج التدريبية وإدخال الاساليب الحديثة في التصنيع تهدف اجمالاً الى رفع "متوسط " انتاجية الفرد العامل في قطاع الصناعة.، وغيرها من الامثلة الاخرى. وفيما يلي اهم مقاييس المتوسطات او مقاييس النزعة المركزية :

مقاييس النزعة المركزية او ما تسمى بمقاييس التمركز او التوسط Measures of central tendency

يشير مفهوم مقاييس النزعة المركزية الى ميل البيانات للتمركز حول قيمة ممثلة او نموذجية في التوزيع وتستخدم مقاييس النزعة المركزية لغايات المقارنة بين مجموعتين من البيانات ولوصف توزيع المشاهدات وتساعد هذه المقاييس في فهم وتفسير سلوك الظواهر وهذه المقاييس :-

1- **الوسط الحسابي Arithmetic Mean**
2- **الوسيط Median and similar Measures**
3- **المنوال Mode**
4- **الوسط الهندسي Geometric Mean**
5- **الوسط التربيعي Root Mean Squares**
6- **الوسط التوافقي Harmonic Mean**

ومن اهم مقاييس النزعة المركزية التي يمكن ان نستفاد منها في دراستنا هي :-

1- الوسط الحسابي Arithmetic Mean

هو عبارة عن القيمة التي يحصل عليها من خلال قسمة المجموع الكلي للقيم على عددها

أ- الوسط الحسابي للبيانات الغير مبوبة

$$\bar{y} = \frac{\sum yi}{n}$$

حيث $\bar{y}$ = الوسط الحسابي

n = عدد المشاهدات

مثال1 :- اوجد الوسط الحسابي للبيانات التالية:

yi = 11 ,12 , 13 , 12 , 13 , 11

1- الطريقة الاعتيادية

$$\bar{y} = \frac{\sum yi}{n}$$

$$\bar{y} = \frac{11+12+13+12+13+11}{6} = \frac{72}{6} = 12$$

مثال2 :- اذا كان متوسط مستوى الهرمون المحفز لنمو الحويصلات يساوي 18 Mg/dL حيث كان مستوى الهرمون المحفز لنمو الحويصلات في انثى الارنب الاولى هو 18 وفي الانثى الثانية هو 19 وفي الانثى الثالثة هو 17 والانثى الرابعة هو 19 اوجد مستوى الهرمون في انثى الارنب الخامسة :

$$\bar{y} = \frac{\sum yi}{n}$$

$$\bar{y} = \frac{y1+y2+y3+y4+y5}{5}$$

$$18 = \frac{18+19+17+19+y5}{5}$$

$$90 = 73 + y5$$

$$y\,5 = 90 - 73 = 17$$

2- طريقة الوسط الفرضي : تستخدم هذه الطريقة عندما تكون قيم مفردات العينة اعداد كبيرة ويصعب التعامل معها وخصوصاً عند عدم توفر الحاسبة تفي هذه الطريقة بالغرض

$$\bar{y} = a + \frac{\sum di}{n}$$

حيث ان : a : الوسط الفرضي

$\sum di$: مجموع الانحرافات عن الوسط الفرضي

n: عدد المشاهدات

مثال // اذا كانت اوزان ستة طلاب من طلبة كلية الاداب كالآتي :-

yi= 85 , 67 , 80 , 75 , 60 , 55

اوجد الوسط الحسابي؟

الحل :نختار وسط فرضي وليكن = 75

yi	di =yi - a
85	10
67	-8
80	5
75	0
60	15
55	$\frac{-20}{-28}$

$$\bar{y} = a + \sum di$$

$$\bar{y} = 75 - \frac{28}{6}$$

$$\bar{y} = 75 - 4.67 = 70.33$$

وحسب الطريقة الاعتيادية:

$$\bar{y} = \frac{\sum yi}{n} = \frac{422}{6} = 70.33$$

ملاحظة// لا يتغير الوسط الحسابي بتغير الوسط الفرضي

ب – الوسط الحسابي في حالة البيانات المبوبة

$$\bar{y} = \frac{\sum f_i y_i}{\sum f_i}$$

مثال // اوجد الوسط الحسابي للبيانات التالية التي تبين توزيع (100) طالب من طلبة كلية الادارة والاقتصاد حسب صفة الوزن ، اوجد الوسط الحسابي لوزن طلبة الكلية .

الحل:

الفئات(الوزن كغم)	التكرار fi(عدد الطلبة)	مركز الفئات yi	fiyi
60 – 62	5	61	305
63- 65	15	64	960
66 – 68	45	67	3015
69 – 71	27	70	1890
74 - 72	8	73	584
	100		6754

$$\bar{y}=\frac{\sum fiyi}{\sum fi}$$

$$\bar{y}=\frac{6754}{100}=67.54$$

جـ - الوسط الحسابي المرجح أو الموزون Weighted Mean

من الناحية العملية هناك الكثير من الحالات تكون بعض المفردات اكثر اهمية من الاخرى مما يتوجب الامر اخذ ذلك بنظر الاعتبار لدى حساب الوسط الحسابي ، فمثلا عند حساب معدل درجات الطالب فأن الامر يستوجب الاخذ بنظر الاعتبار عدد الساعات الاسبوعية المخصصة لكل درس وهذا يعني ترجيح المفردات بأوزان معينة تمثل اهمية كل منها وعنده ادخال اهمية المفردات في حساب الوسط الحسابي فأن عندئذ يسمى الوسط الحسابي المرجح وبتعبير آخر لكل قيمة من المشاهدات (yi) وزن خاص يتناسب مع اهميتها (wi) فالوسط الحسابي لهذه القيم يحسب كما يلي:-

$$\bar{y}_w=\frac{\sum wiyi}{\sum wi}$$

حيث ان $\bar{y}_w$ الوسط الحسابي الموزون

wi اوزان وأهمية (المفردة)

yi قيمة المشاهدة

مثال // اذا كانت درجات احد الطلبة في الصف الاول في قسم الادارة في الدروس المقررة في هذه المرحلة حسب الساعات الاسبوعية المحددة لكل درس وكما يلي:

الدرجة	عدد الساعات
62	2
80	2
75	2
88	3
84	3
84	3
86	3

المطلوب حساب معدل الطالب ؟

الحل //

الدرجة yi	الاهمية wi	wiyi
62	2	124
80	2	160
75	2	150
88	3	264
84	3	252
84	3	258
86	3	270
	18	1478

$$\frac{\sum wiyi}{\sum wi} = \frac{1478}{18} = 80.714$$

اما الوسط الحسابي الموزون في حالة البيانات المبوبة :- $\bar{y}_w = \frac{\sum wifiyi}{\sum wifi}$ اذ ان :

yi= مركز الفئة
Fi= التكرار
wi= الاهمية

مثال // اوجد الوسط الحسابي الموزون للبيانات التالية التي تمثل انتاج معمل الادوية في سامراء من الادوية بالطن وعدد المكائن العاملة وعدد ساعات العمل ؟

فئات الانتاج بالطن	عدد المكائن العاملة fi	عدد ساعات العمل wi	yi	wifi	wifiyi
2 – 4	4	6	3	24	72
4 – 6	5	5	5	25	125
6 – 8	6	6	7	36	252
8 – 10	3	4	9	12	108
10 - 12	2	4	11	8	88
	20			105	645

$$\bar{y}_w = \frac{\sum wifiyi}{\sum wifi} = \frac{645}{105} = 6.134$$ طن

خصائص الوسط الحسابي :-

1- مجموع انحرافات القيم عن وسطها الحسابي = صفر

$\sum_{i=1}^{n}(y_i - \bar{y}) = 0$ حالة البيانات الغير المبوبة

$\sum_{i=1}^{n} f_i(y_i - \bar{y}) = 0$ حالة البيانات المبوبة

2- مجموع مربعات انحرافات القيم عن وسطها الحسابي = اقل ما يمكن

$\sum_{i=1}^{n}(y_i - \bar{y}) =$ اقل ما يمكن حالة البيانات الغير المبوبة

$\sum_{i=1}^{n} f_i(y_i - \bar{y}) =$ اقل ما يمكن حالة البيانات المبوبة

3- يأخذ الوسط الحسابي بعين الاعتبار جميع القيم في حسابه

4- يتأثر الوسط الحسابي بالقيم الشاذة او المتطرفة لان الوسط الحسابي يأخذ بنظر الاعتبار جميع القيم .

5- هناك صعوبة في حساب الوسط الحسابي في حالة الفئات المفتوحة لأنه من الصعب تحديد مراكز الفئات وهذه المشكلة تحل بتحديد مراكز الفئات بصورة تقريبية.

الوسيط Median:-

يعرف الوسيط بأنه القيمة التي تمثل المرتبة الوسطى عندما ترتب القيم قيد الدرس تصاعدياً او تنازلياً وهذا يعني ان نصف القيم تقل عن قيمة الوسيط والنصف الاخر يزيد عنها.

أ- ايجاد الوسيط لبيانات غير مبوبة

1- يتم ترتيب القيم تصاعدياً او تنازلياً

2- اذا كان عدد القيم فردي (n) فالوسيط يكون القيمة التي ترتيبها $\frac{n+1}{2}$ واذا كان عدد القيم زوجي (n) فالوسيط هو الوسط الحسابي للقيمتان اللتان ترتيبهما $\frac{n}{2}$ و $\frac{n}{2}$ +1

مثال // اوجد الوسيط للبيانات التالية:

yi=11 , 12 , 13 , 12 , 13 , 11 , 14

الحل :- 1- ترتب البيانات ترتيب تصاعدي 11 , 11 , 12 , 12 ,13 , 13 ,14

3- ايجاد رتبة الوسيط

بما ان عدد القيم (n) = فردي

رتبة الوسيط T = $\frac{n+1}{2}$ = $\frac{7+1}{2}$ = $\frac{8}{2}$ = 4

Me = 12

مثال // اوجد الوسيط للبيانات التالية: yi= 11 , 12, 13 ,12 , 13 , 11 , 14 , 10

الحل :-

1- ترتيب البيانات تصاعدياً 10 , 11 , 11, 12 , 12 , 13 , 13 , 14

2- ايجاد رتبة الوسيط

بما ان عدد القيم زوجي = 8

فالوسيط هو الوسط الحسابي للقيمتان التي ترتيبهما $\frac{n}{2}$ ، $\frac{n}{2}$ +1

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

$$5 = \frac{8}{2} + 1 = \frac{n}{2} + 1$$

$$Me = \frac{12+12}{2} = \frac{24}{2} = 12$$

ب – ايجاد الوسيط لبيانات مبوبة :

1 -الوسيط لبيانات مبوبة (متغير متقطع):

ويمكننا ايجاد الوسيط من الجداول التكرارية البسيطة بتحويلها الى جداول تكرارية صاعدة او نازلة.

الوسيط في حالة التكرار المتجمع الصاعد

خطوات ايجاد الوسيط لبيانات مبوبة لمتغير متقطع :

1- نجد التكرار المتجمع الصاعد

2- نجد ترتيب الوسيط والذي يساوي $T = \frac{\sum f_i}{2}$

ملاحظة : ***fi*** **هنا التكرارات الاصلية وليس التكرار المتجمع الصاعد.**

3- نحدد قيمة الوسيط وهي التي تقع بين التكرارين يعني ترتيب الوسيط بين التكرارين

4-نحدد فئة الوسيط وان مركز هذه الفئة يمثل الوسيط.

مثال/ الاتي توزيع لعينة من الاسر حسب عدد افراد الاسر ة ، المطلوب حساب الوسيط لعدد افراد الاسرة.

الفئات	2-4	5-7	8-10	11-13	14-16	17-19	20-22	المجموع
التكرارات	6	9	12	20	14	11	8	80

الحل/

الفئات	التكرارات f_i	الحدود العليا للفئات	التكرار المتجمع الصاعد F_i
2-4	6	4 فاقل	6
5-7	9	7 فاقل	15
8-10	12	10 فاقل	27
11-13	20	13 فاقل	47
14-16	14	16 فاقل	61
17-19	11	19 فاقل	72
20-22	8	22 فاقل	80
المجموع	80		

ترتيب الوسيط :

$$T = \frac{\sum f_i}{2} = \frac{80}{2} = 40$$

اي الوسيط يقع بين التكرارين الثالث والرابع أي بين 27 و47

وعليه فان فئة الوسيط هي الفئة (11-13) لانها الاقرب الى 47 وان الوسيط يمثل مركز هذه الفئة

وعليه فالوسيط يساوي 12 وكالاتي:

$$Me = \frac{11 + 13}{2} = 12$$

2 -الوسيط لبيانات مبوبة (متغير مستمر):

نطبق القانون التالي:

$$Me = L_k + [\frac{\frac{\sum fi}{2} - F_{k-1}}{F_{k+1} - F_{k-1}}] \times h_k$$

L_k = هي الحد الادنى لفئة الوسيط

$\frac{\sum fi}{2}$ = رتبة الوسيط في حالة مجموع التكرارات عدد زوجي

$\frac{\sum fi+1}{2}$ = رتبة الوسيط في حالة مجموع التكرارات عدد فردي

F_{K-1}= التكرار المتجمع الصاعد السابق لفئة الوسيط

F_{K+1}= التكرار المتجمع الصاعد اللاحق لفئة الوسيط

h_k= طول الفئة (طول فئة الوسيط)

مثال // اوجد الوسيط للبيانات التالية التي تبين توزيع 100 طالب من طلبة كلية الصيدلة حسب صفة الوزن.

الحل 1- : نجد التكرار المتجمع الصاعد وكالتالي:

الفئات	التكرارات f_i	الحدود العليا للفئات	التكرار المتجمع الصاعد F_i
60 – 62	5	اقل او يساوي62	5
63 – 65	15	= = 65	20
66 – 68	45	= = 68	65
69 – 71	27	= = 71	92
72 - 74	8	= = 74	100
المجموع	100		

2-نجد ترتيب الوسيط من الصيغة :

$$T = \frac{\sum f_i}{2} = \frac{100}{2} = 50$$

3-نحدد فئة الوسيط والتي تقع بين التكرارين الثاني والثالث وهي (**68 – 66**) لان 50 اقرب الى 65

اي الوسيط يقع بين التكرارين الثالث والرابع أي بين **20** و **65**

4-نطبق صيغة القانون $Me = L_k + [\frac{\frac{\sum fi}{2} - F_{k-1}}{F_{k+1} - F_{k-1}}] \times h_k$

$$66 + Me = L_k + \left[\frac{\frac{\sum fi}{2} - F_{k-1}}{F_{k+1} - F_{k-1}}\right] * h_k = 66 + \left[\frac{50-20}{92-20}\right] * 3 = 67.25$$

ملاحظات عن الوسيط :-

1- يستخدم الوسيط كمقياس للنزعة المركزية بدلاً عن الوسط الحسابي عندما تكون هناك قيمة شاذة في التوزيع .

2- يستخدم في حالة الفئات المفتوحة

3- الوسيط قليل الحساسية للمتغيرات التي تحدث في قيم البيانات الاصلية لانه يهتم بالقيم الواقعة في المنتصف ويهمل الاطراف على عكس الوسط الحسابي الذي يعتبر شديد الحساسية لانه يأخذ بعين الاعتبار جميع القيم في حسابه .

4- يمكن استخدامه في حالة المتغيرات الوصفية التي لا تعبر عنها في الارقام كما هو الحال في ترتيب الاشخاص وفقاً لخصائصهم .

3- المنوال Mode:-

هي القيمة الاكثر شيوعاً او تكراراً في التوزيع وهو ابسط مقاييس النزعة المركزية

أ- المنوال في حالة البيانات غير المبوبة

مثال 1// اوجد المنوال للبيانات التالية : 7 , 4 , 8 , 6 , 4

Mo = 4

مثال 2// اوجد المنوال للبيانات التالية: 7 , 6 , 3 , 8 , 5 ,6 , 3

Mo = 3

Mo = 6 التوزيع ثنائي المنوال

مثال3// اوجد المنوال للبيانات التالية:, 7 , 9 , 3 , 8 , 5 ,6

لايوجد منوال

ب – المنوال في حالة البيانات المبوبة

$$Mo = Li + [\frac{d1}{d1+d2}] \times C$$

Li = هي الحد الادنى الحقيقي للفئة المنوالية

$d1$ = الفرق بين فئة المنوال والفئة السابقة لها في التكرار

$d2$ = الفرق بين فئة المنوال والفئة اللاحقة لها في التكرار

C= طول الفئة

مثال // اوجد المنوال للبيانات التالية التي تمثل توزيع طلبة كلية الصيدلة حسب صفة الوزن

الفئات	التكرار fi
60 – 62	5
63 – 65	15
66 – 68	45 (الفئة المنوالية)
69 – 71	27
72 - 74	8

الحل :-

$d1 = 45 - 15 = 30$

$d2 = 45 - 27 = 18$

$$Mo = Li + \left[\frac{di}{d1+d2}\right] \times C$$

$$Mo = 65.5 + \left[\frac{30}{30+18}\right] \times 3$$

$$Mo = 65.5 + \frac{30}{48} \times 3$$

$$Mo = 65.5 + 0.625 \times 3$$

$$Mo = 65.5 + 1.88 = 67.38$$

احصاء ١ - د.بريسان