وزارة التعليم العالي والبحث العلمي

جامعة الفرات الاوسط التقنية

المعهد التقني بابل

قسم تقنيات ادارة المواد

المرحلة الاولى

الإحصـاء

اسم الأستاذ

حيدر المعمار

اللقب العلمي

مدرس مساعد

السنة الدراسية

2026-2025

الاحصاء

مفاهيم اساسية

علم الاحصاء

هو الطريقة العلمية التي تختص بجمع البيانات والحقائق عن ظاهرة معينة او مجموعة ظواهر معينة، وتنظيم، وتبويب هذه البيانات، والحقائق بطريقة يسهل تحليلها، وتفسيرها، ومن ثم استخلاص النتائج، واتخاذ القرار على ضوء ذلك.

يتضمن علم الاحصاء فرعين رئيسين هما:

1- الاحصاء الوصفي

يتضمن هذا الفرع الطرق والاساليب المستخدمة في جمع البيانات، والمعلومات عن ظاهرة معينة او مجموعة ظواهر، وكيفية تنظيم، وتصنيف، وتبويب هذه البيانات، وعرضها في جداول، ورسوم بيانية، وحساب بعض المؤشرات الاحصائية.

2- الاحصاء الاستدلالي

وهذا الفرع يهتم بتقدير مؤشرات المجتمع الاحصائي، واختبار الفرضيات.

مجال تطبيق علم الاحصاء

يستخدم في كل مجال يتضمن بحثاً علمياً، مثل البحوث الزراعية، والصناعية، والنفسية، والاجتماعية، والرياضة والشباب، والطبية، والاقتصادية، والادارية، والهندسية، والخ.

مراحل الطريقة الاحصائية

1- تحديد مشكلة، او فرضية البحث.

2- جمع البيانات، والمعلومات عن الظاهرة او الظواهر ذات العلاقة بالبحث.

3- تصنيف البيانات، وتبويبها، وعرضها.

4- حساب المؤشرات الاحصائية مثل تقدير معلمات مجتمع البحث.

5- تحليل معطيات الدراسة، والتوصل الى النتائج.

6- تفسير النتائج، واتخاذ القرار بشأن فرضيات البحث.

المجتمع الاحصائي

هو جميع مفردات الظاهرة موضوع البحث، او الدراسة. وتلك المفردات يفترض ان تشترك بصفة او صفات معينة، وقد تكون تلك المفردات كائناً حياً، او اي شيء اخر. وقد يكون المجتمع الاحصائي محدد، او غير محدد.

المفردة الاحصائية

هي اصغر وحدة في المجتمع الاحصائي.

جمع وتصنيف وتبويب البيانات

اساليب جمع البيانات

يمكن الحصول على بيات البحث ثيد الدراسة من مصدرين هما:

1- المصادر التأريخية

وهي البيانات والمعلومات المحفوظة والمتجمعة لدى اجهزة ومؤسسات الدولة المختلفة نتيجة لمسوحات قامت بها هذه الجهات او الهيئات لاغراض خاصة بها. مثال ذلك: بيانات تعداد السكان، واحصاءات الانتاج الزراعي، والصناعي، واحصاءات التجارة الخارجية، والداخلية، واحصاءات الطلبة المتخرجين من الجامعات العراقية، وغيرها.

2- المصادر الميدانية

في حالة تعذر الحصول على البيانات من المصادر التأريخية يتم اللجوء الى الميدان للحصول على تلك البيانات، اي الحصول على البيانات من من مصادرها الاصلية.

اساليب جمع البيانات من الميدان

1- اسلوب الحصر الشامل

بموجب هذا الاسلوب يتم جمع البيانات عن كافة مفردات المجتمع الاحصائي، وفي هذه الحالة يجب ان يكون المجتمع الاحصائي محددا، اي يمكن ملاحظة كل مفردة فيه. مثال ذلك التعداد العام للسكان، او

حصر المنشآت والوحدات الصناعية في العراق. ويعتبر اسلوب التسجيل الشامل افضل اسلوب في جمع البيانات كونه يجهز الباحث ببيانات كاملة عن كافة مفردات مجتمع الدراسة، الا انه يحتاج الى وقت وجهد وموارد مادية وبشرية كبيرة في انجازمهمة جمع البيانات، بالاضافة الى احتمال الوقوع في اخطاء نتيجة التعامل مع مفردات المجتمع الاحصائي بشكل كامل.

2- اسلوب العينات

ويقصد به جمع البيانات والمعلومات عن مجموعة معينة من مفردات المجتمع الاحصائي، تدعى هذه المجموعة من المفردات بالعينة، بحيث يتم اختيارها بطريقة تضمن تمثيلها للمجتمع الاحصائي بشكل دقيق. ومن ميزات اسلوب العينات ما يلي:

- لا تحتاج الى جهد ووقت وموارد مادية وبشرية كبيرة.
- يمكن استخدامه في حالة المجتمعات غير المحددة.
- امكانية الحصول على بيانات لها صفات اكثر مما لو استخدمنا اسلوب الحصر الشامل.
- امكانية اختبار دقة النتائج.

انواع العينات

أ- العينة العشوائية

يقصد بالعينة العشوائية بانها تلك المجموعة من المفردات المختارة من المجتمع الاحصائي بحيث ليس للباحث اي تدخل في عملية الاختيار، بمعنى ان لكل مفردة نفس فرصة الظهور في العينة. ومن انواع العينات العشوائية ما يلي:

1- العينة العشوائية البسيطة

وهي تلك العينة المختارة بشكل عشوائي يضمن لان تمتلك اية مفردة من مفردات المجتمع الاحصائي نفس الفرصة في الظهورضمن مفردات العينة. ويراعى عند استعمال هذا النوع من العينات ان يكون المجتمع الاحصائي متجانس من حيث الصفة او الصفات ذات العلاقة بالبحث. وفيما يلي وصفا لاسلوب المعاينة العشوائية البسيطة:

بافتراض ان المجتمع الاحصائي متجانس ومحدود وعدد مفرداته يساوي N . فعند اختيار عينة عشوائة بسيطة ذات حجم n نجد ان كل مفردة في هذا المجتمع الاحصائي لها نفس احتمال الظهور في العينة، وهذا الاحتمال يساوي $\frac{1}{N}$. اما عدد العينات العشوائية البسيطة التي يمكن اختيارها من هذا المجتمع يحسب باستخدام قانون التوافيق كما يأتي:

$$\text{عدد العينات العشوائية البسيطة} = C_n^N = \frac{N!}{n!(N-n)!}$$

مثال: مجتمع احصائي متجانس عدد مفرداته (4) وهي A, B, C, D يراد اختيار عينة عشوائية بسيطة قوامها (3) مفردات. ما هو عدد العينات العشوائة البسيطة التي يمكن اختيارها من هذا المجتمع؟ وما هو احتمال اختيار اي مفردة منه؟

$$\text{عدد العينات العشوائية البسيطة} = C_3^4 = \frac{4!}{3!(4-3)!} = \frac{24}{6} = 4$$

$$\text{احتمال سحب اي مفردة} = \frac{1}{N} = \frac{1}{4}$$

$$4! = 4 * 3 * 2 * 1$$

العينات العشوائية البسيطة هي:

A, B, C
A, B, D
A, C, D
B, C, D

2- العينة الطبقية العشوائية

تعتبر العينة الطبقية العشوائية افضل انواع العينات واكثرها دقة في تمثيل المجتمع الاحصائي غير المتجانس، حيث ان في كثير من الاحوال تكون مفردات المجتمع الاحصائي غير متجانسة من حيث الصفة او الصفات المدروسة. ففي حالة كون المجتمع الاحصائي ذو مفردات غير متجانسة لايجوز سحب عينة عشوائية بسيطة لتمثيل هذا المجتمع. فعلى سبيل المثال اذا اريد دراسة دخل الاسرة فاننا نجد ان هناك اسرا ذات دخول عالية، واخرى ذات دخول متوسطة، واخرى ذات دخول منخفضة، اذن مفردات المجتمع الاحصائي هنا غير متجانسة من حيث الصفة المدروسة، ولا يجوز سحب عينة

عشوائية بسيطة لاننا سنحصل على تقدير لمتوسط الدخل قد يكون منحازا لاحدى الفئات الثلاث. وعليه يجب تقسيم المجتمع الاحصائي الى ثلاث فئات : الاولى تضم الاسر ذات الدخول المرتفعة، والثانية تضم الاسر ذات الدخول المتوسطة، والثالثة تضم الاسر ذات الدخول المنخفضة. وبعد ذلك يتم سحب عينة عشوائية بسيطة من كل فئة (تسمى الفئة او المجموعة بالطبقة) يتناسب حجمها وحجم الطبقة في المجتمع، ومجموع حجوم العينات العشوائية الثلاث تؤلف العينة العشوائية الطبقية.

وفيما يلي **مثال** يوضح اسلوب المعاينة العشوائية الطبقية:-

بافتراض ان مجتمعا احصائيا مؤلفا من (N) من المفردات والذي يمكن تجزئته الى (L) من الطبقات حجومها $(N_1, N_2 ,\ldots,N_L)$ علما ان $(N_1 + N_2 +\ldots+ N_L = N)$

وان تقسيم المجتمع الاحصائي الى (L) من الطبقات يتم على اساس اشتراك مفردات كل طبقة بصفة او عدة صفات مشتركة. فاذا يراد اختيار عينة طبقية عشوائية ذات حجم (n) فان هنالك عدة طرق لاختيار هذه العينة منها ما تسمى **بطريقة التوزيع المتناسب** وبموجب هذه الطريقة يتم سحب عينة عشوائية بسيطة من كل طبقة يتناسب حجمها وحجم الطبقة في المجتمع. فاذا كانت نسبة كل طبقة في المجتمع (وزن الطبقة) يساوي

$W_h=N_h / N$, $h= 1, 2 ,3 ,\ldots,L$, where $W_1+W_2+\ldots+W_L= 1$

فان حجم العينة العشوائية البسيطة المسحوبة من الطبقة الاولى يساوي $(n_1= W_1 . n)$

وان حجم العينة العشوائية البسيطة المسحوبة من الطبقة الثانية يساوي $(n_2= W_2 . n)$

وهكذا لاخر طبقة . وبذلك فان حجم العينة الطبقية العشوائية $(n= n_1+ n_2 +\ldots+ n_L)$

مما سبق نلاحظ ان $(n_h / n = N_h /N)$ اي ان مساهمة الطبقة (h) في العينة تساوي نسبة مساهمة الطبقة (h) في المجتمع.

مثال: مجتمع مؤلف من 2200 اسرة . يراد دراسة متوسط الدخل للاسرة، علما ان هذا المجتمع يضم اسر ذات دخول مرتفعة عددها (700) اسرة، واسر ذات دخول متوسطة عددها (900) اسرة، واسر

ذات دخول واطئة عددها (600) اسرة. المطلوب: سحب عينة عشوائية طبقية ذات حجم (110) اسرة. فما هو حجوم عينات الطبقات باعتماد طريقة التوزيع المتناسب؟

$N= 2200$, $N_1= 700$, $N_2= 900$, $N_3= 600$

$W_1= 700/2200= 7/22$, $W_2= 900/2200= 9/22$, $W_3= 600/2200= 6/22$

$n_1= 7/22*110= 35$, $n_2= 9/22*110=45$, $n_3= 6/22*110=30$

نلاحظ ان : $W_1= 7/22=35/110$, $W_2= 9/22=45/110$, $W_3= 6/22=30/110$

3- العينة العشوائية المنتظمة

بموجب هذا النوع من المعاينة يتم تقسيم مفردات المجتمع البالغ عددها (N) من المفردات الى (n) من المجاميع وكل مجموعة تحتوي على (k) من الفردات. حيث ان : (k=N/n) علما بان مفردات المجتمع مقسمة وفق نظام معين، كأن يكون ترتيبا تصاعديا او تنازليا او ترتيب الدور السكنية حسب تسلسلها في شارع معين. وعند اختيار مفردات العينة المنتظمة يتم اولا اختيار مفردة واحدة بصورة عشوائية من المجموعة الاولى، وعلى ضوء تسلسل المفردة المختارة من المجموعة الاولى يتم اختيار مفردة من المجموعة الثانية بعد اضافة العدد (k) الى تسلسل المفردة الاولى، وهكذا يضاف العدد (k) الى تسلسل المفردة الثانية لنحصل على المفردة الثالثة، وهكذا لاخر مفردة. بمعنى اخر يتم اختيار مفردة من المجموعة الاولى بطريقة عشوائية اما بقية المفردات يتم اختيارها على ابعاد متساوية من المفردة الاولى.

مثال: تم ترتيب (24) طالبا حسب تسلسل درجاتهم تنازليا، ويراد اختيار عينة عشوائية بحجم (6) طلاب للتعرف على اسباب انخفاض مستواهم في الامتحان. المطلوب تحديد تسلسل هؤلاء الطلبة.

1- يتم تقسيم الطلبة الى (6) مجاميع كل منها يحتوي k من الطلبة حيث (k=N/n=24/6=4)

2- يتم اختيار طالب في المجموعة الاولى بصورة عشوائية وليكن تسلسل رقم (3) مثلا.

3- بقية الطلبة فيتم الحصول على تسلسلاتهم باضافة 4 (والتي تمثل k) الى تسلسل (3) فيكون الطالب الثاني بتسلسل 7 والطالب الثالث بتسلسل 11 والطالب الرابع بتسلسل 15 وهكذا الى تسلسل اخر طالب في العينة سيكون 23.

المجموعة6	المجموعة5	المجموعة4	المجموعة3	المجموعة2	المجموعة1
21 22 **23** 24	17 18 **19** 20	13 14 **15** 16	9 10 **11** 12	5 6 **7** 8	1 2 **3** 4

اذن العينة العشوائية المنتظمة تكون على اساس ان الطالب الذي تم اختياره عشوائيا يحمل التسلسل الثالث وتكون تسلسلات الطلبة الذين يمثلون العينة هي: (23 ، 19 ، 15 ، 11 ، 7 ، 3)

ملاحظة: ان عدد العينات العشوائية المنتظمة التي يمكن اختيارها من المجتمع يساوي عدد مفردات كل مجموعة (k).

4- العينة متعددة المراحل

بموجب اسلوب المعاينة المتعددة المراحل يتم تقسيم المجتمع الاحصائي الى وحدات اولية، ثم يتم سحب عينة عشوائية بسيطة من هذه الوحدات الاولية وهي المرحلة الاولى. ثم يتم تقسيم الوحدات الاولية المختارة الى وحدات اصغر تدعى بالوحدات الثانوية، ويتم اختيار عينة عشوائية بسيطة من الوحدات الثانوية لكل وحدة اولية وهي المرحلة الثانية. ثم تقسم الوحدات الثانوية المختارة الى وحدات اصغر يتم اخيار عينة عشوائية بسيطة منها وهي المرحلة الثالثة. وتستمر عملية التقسيم والاختيار لحين الوصول الى المفردات التي تجمع منها البيانات. فمثلا عند اجراء داسة لتقدير متوسط استهلاك العائلة العراقية لمادة السكر فان الوحدة الاحصائية التي يمكن الحصول منها على البيانات هي العائلة العراقية، فعند اختيار عينة متعددة المراحل يتم تقسيم العراق الى محافظات كونها الوحدات الاولية نختار منها عشوائيا عينة من المحافظات في المرحلة الاولى، ثم تقسم المحافظات المختارة في المرحلة الاولى الى اقضية وهي الوحدات الثانوية يتم اختيار عينة عشوائية منها في المرحلة الثانية. ثم تقسم الوحدات

الثانوية المختارة في المرحلة الثانية الى نواحي يتم اختيار عينة عشوائية منها وهي المرحلة الثالثة، ثم تقسم النواحي المختارة في المرحلة الثالثة الى محلات سكنية يتم اختيار عينة عشوائية منها وهي المرحلة الرابعة، وتقسم المحلات المختارة الى ازقة ومنها يتم اختيار عينة عشوائية منها وهي المرحلة الخامسة. وتقسم الازقة المختارة الى الدور السكنية التي يتم اختيار عينة عشوائية منها وهذه المرحلة السادسة. وبهذا يتم الوصول الى العائلة التي يتم جمع البيانات منها.

ب- العينات غير العشوائية

وهي تلك المجموعة من المفردات المختارة من المجتمع الاحصائي بطريقة يكون للباحث دخلاً في اختيار تلك المفردات، وذلك لاعتبارات تتعلق بطبيعة البحث او الدراسة. وهذه العينات نوعين هما:-

1- العينة العمدية

وهي العينة المختارة بشكل متعمد، ونعتقد مسبقا ان مفردات هذه العينة هي خير من يمثل مجتمع الدراسة. فمثلا عند دراسة السبل الكفيلة للارتقاء برياضة كرة القدم فمن الافضل اختيار عينة من المختصين برياضة كرة القدم وبشكل عمدي كون ان هذه العينة هي ذات خبرة بشؤون هذه الرياضة.

2- العينة الحصصية

بموجب هذا النوع من المعاينة يتم تقسيم المجتمع الاحصائي الى عدة طبقات بالاستناد الى معايير تتعلق بطبيعة الدراسة ثم يتم اختيار عينة عمدية من كل طبقة (بشكل غير عشوائي) يتناسب حجمها وحجم الطبقة في المجتمع، ومجموع حجوم هذه العينات العمدية يمثل حجم العينة الحصصية.

وسائل جمع البيانات

بعد تحديد حجم العينة واسلوب المعاينة الملائم يتطلب الامر اختيار الوسيلة الملائمة في جمع البيانات والمعلومات عن الظاهرة او الظواهر المتعلقة بالدراسة. واهم وسائل جمع البيانات ما يلي:

1- اسوب الجمع المباشر

بموجب هذا الاسلوب يتم جمع البيانات والمعلومات المتوفرة لدى اجهزة معينة ذات علاقة بالدراسة. فمثلا عند دراسة تطور الدخل القومي فيمكن الرجوع الى الجهاز المركزي للاحصاء لاختيار سلسلة زمنية من بيانات الدخل القومي في العراق. وكذلك وفق هذا الاسلوب يمكن توفير البيانات ذات الطابع المختبري او الحقلي التي يتم عملها من قبل الباحث فمثلا عند اجراء دراسة حول اختبار فاعلية سماد معين على زيادة كمية محصول الحنطة فان الباحث يقوم بنفسه بعمل تصميم لهذه التجربة وتسجيل نتائجها التي تمثل البيانات اللازمة لها البحث.

2- الاستبيان

وهو استمارة يتم من خلالها جمع البيانات والمعلومات من مفردات مجتمع الدراسة او عينة البحث عن طريق مواجهة الباحث الشخصية للمفردة الاحصائية او عن طريق المراسلة. ويجب مراعاة النقاط التالية عند تصميم الاستمارة :

أ- اعداد مقدمة ايضاحية تكتب في بداية الاستمارة.

ب- ان تكون فقرات الاستمارة متسلسلة وغير مبعثرة بحيث ان كل صنف منها يحقق غرض معين.

جـ - ان تكون الاسئلة متوسطة العدد.

د- ان تكون الاسئلة واضحة المعنى.

هـ - ان تهيئ الاسئلة بالشكل الذي تكون اجابة الفرد عليها محددة وقصيرة.

و- ان يؤخذ بنظر الاعتبار ظروف تفريغ وتصنيف وتبويب وترميز الاجابات.

الاخطاء الشائعة في جمع البيانات

يحدث في بعض الاحيان ان يقع الباحث في بعض الاخطاء لدى جمعه للبيانات. وهذه الاخطاء تحدث نتيجة سوء استخدام الطريقة الاحصائية وهي:

1- خطأ التحيز

عند جمع البيانات يفترض ان تجمع من مصادرها الاصلية، الا انه في بعض الاحيان تجمع البيانات من مصادر اخرى غير مصادرها الاصلية، فمثلا عند دراسة رغبات الاطفال على اساس عينة من اطفال مجموعة من الاسر، فان المصدر الصحيح لجمع البيانات حول هذه الظاهرة هو الام كونها تمتلك معلومات كافية عن طفلها بسبب معايشتها له، ولكن في حالة اخذ المعلومات من غير هذا المصدر فمن المحتمل جدا ان نقع في خطأ تسجيل رغبات الطفل، ومن ثم تأثر نتائج الدراسة بمجمل هذه الاخطاء.

2- خطأ الصدفة

يحصل هذا النوع من الاخطاء عندما يقوم الباحث باستيفاء بعض البيانات والمعلومات بالاعتماد على معلوماته الشخصية او التعمد بجمع البيانات من بعض المفردات دون الاخرى، او يقوم بجمع بيانات ناقصة لسبب او لاخر.

تصنيف وتبويب البيانات

ان البيانات المستحصل عليها بخصوص ظاهرة معينة تسمى بالبيانات الخام او الاولية، وتلك البيانات تكون غير منظمة ويصعب على الباحث تكوين فكرة عن الظاهرة المدروسة وكذلك يصعب الاعتماد عليها لاغراض التحليل الاحصائي. لذا فان الخطوات الهامة التي تلي عملية جمع البيانات هي عملية مراجعة وتصنيف البيانات وتبويبها.

1- مراجعة البيانات

بعد اتمام عملية جمع البيانات يجب مراجعة تلك البيانات وتدقيقها لغرض التأكد من مطابقتها وتكاملها لمتطلبات الدراسة.

2- تصنيف البيانات

بعد التأكد من تكامل ووضوح ودقة البيانات التي تم الحصول عليها نبدأ بعملية تصنيف على اساس الظواهر التي جمعت عنها البيانات حيث يتم فرز البيانات على هيئة مجموعة، فقد يكون التصنيف على اساس ظاهرة العمر او الجنس او المهنة او الوزن او الطول او محل الاقامة وغيرها.

3- تبويب البيانات

بعد اتمام تصنيف البيانات نبدأ لعملية تبويب البيانات ويقصد بتبويب البيانات تفريغ البيانات المصنفة في جداول خاصة بحيث ان كل جزء من البيانات المصنفة عن الظاهرة المعنية يعود الى مستوى معين لتلك الظاهرة. فالهدف من عملية التبويب هو ابراز البيانات وتوضيحها في اضيق حيز ممكن كي نتمكن من تكوين فكرة عنها. وهناك اربعة اشكال للتبويب هي:

أ- التبويب الزمني

وهو تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل مجموعة منها يعود لوحدة زمنية معينة، مثل اليوم او الاسبوع او الشهر او السنة.

ب- التبويب الجغرافي

هو تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل مجموعة منها تخص وحدة جغرافية معينة او تقسيم اداري معين.

جـ - التبويب الكمي

هو تجميع البيانات المصنفة وترتيبها في جداول على اساس ان يكون كل مجموعة منها تخص وحدة كمية معينة مثل وحدات الوزن او الطول او العمر او المسافة والى اخره.

د – التبويب على اساس صفة معينة

هو تجميع البيانات المصنفة وترتيبها في جداول على اساس ان يكون كل مجموعة منها تشترك بصفة خاصة معينة مثل الحالة الاجتماعية او عنوان الوظيفة او النوع او القومية.

التوزيعات التكرارية واساليب عرض البيانات

المتغير العشوائي

هو صفة لشيء ما، او الشيء نفسه ويعرف ايضا بانه دالة ذات قيمة حقيقية معرفة على فضاء يدعى فضاء العينة. ويرمز للمتغير العشوائي بالحروف الكبيرة، ويرمز لقيمه عند تنفيذ التجربة بالحروف الصغيرة.

مثال: عند رمي زهر النرد وملاحظة العدد الذي سوف يظهر على وجهه بعد رميه، هنا المتغير X هو العدد الذي سوف يظهر على وجه الزهر بعد رميه. ان الحالات الممكنة الوقوع اي ظهور العدد على وجه زهر النرد هي مجموعة الاعداد الحقيقية (6, 5, 4, 3, 2 , 1) ولا يمكن اطلاقا ظهور غيرها . هذه المجموعة يطلق عليها فضاء العينة للمتغير X، وعناصرها تمثل القيم الممكنة للمتغير X، فذلك يعني ان المتغير X دالة معرفة على هذا الفضاء. وحيث ان تجربة رمي الزهر هي تجربة عشوائية اي ان عملية رمي الزهر تتم دون تحيز من قبل صاحب التجربة لهذا الوجه من الزهر او ذاك، فعليه يكون المتغير X متغيرا عشوائيا. وغالبا ما يتم الرمز الى مجموعة القيم الممكنة للمتغير (فضاء العينة) بالرمز Ω (ويلفظ اوميكا) اي ان

$$\Omega = \{X : x=1, 2, 3, 3, 4, 5, 6\}$$

انواع المتغيرات العشوائية

تقسم المتغيرات العشوائية الى قسمين رئيسين هما:

أ- المتغيرات النوعية (الوصفية)

وهي المتغيرات التي لايمكن قياسها بوسائل القياس المألوفة، وانما تشكل صفات مثل متغير لون العين ويمكن ان يكون اسود او ازرق او عسلي او اخضر .

ب- المتغيرات الكمية

وهي المتغيرات التي يمكن قياسها بوسائل القياس المألوفة ، مثل عدد الطلبة في صف معين او عدد اشجار البرتقال في بستان ما، او وزن حمولة من السمنت بالطن. وتكون المتغيرات الكمية على نوعين هما:

1- المتغيرات المتقطعة

اذا كانت مجموعة القيم الممكنة للمتغير العشوائي هي مجموعة قابلة للعد، سواء كانت مجموعة محدودة اوغير محدودة، عندئذ يقال ان المتغير العشوائي متقطع. (الارقام صحيحة بدون كسور)

2- المتغيرات المستمرة

اذا كانت مجموعة القيم الممكنة للمتغير العشوائي هي مجموعة غير قابلة للعد، سواء كانت مجموعة محدودة اوغير محدودة، عندئذ يقال ان المتغير العشوائي مستمر. (الارقام حقيقية تتقبل الكسور)

اساليب عرض البيانات:

وهي نوعين :

اولاً: العرض الجدولي للبيانات

أ- التوزيع التكراري:

وهو تلخيص وترتيب البيانات الخاصة بالمتغير العشوائي والتي سبق وان جمعت وصنفت في جداول مقسمة الى مجاميع كل منها تسمى الفئات، وهذه الفئات قد تكون مرتبة تصاعديا او تنازليا حسب طبيعة البيانات.

ولشرح مكونات التوزيع التكراري نفرض ان:

المتغير العشوائي X يأخذ القيم $(x_1 , x_2 , x_3 , \ldots x_n)$ من عينة عشوائية ذات حجم (n)

وان x_S تعني اصغر قيمة، و x_L تمثل اكبر قيمة في تلك البيانات.

ويراد عمل جدول توزيع تكراري عدد فئاته m.

خطوات عمل جدول التوزيع التكراري:

1- حساب المدى الكلي للتوزيع: وهو الفرق بين اكبر قيمة واقل قيمة مضافا اليها واحد.

$$\mathbf{T.R = x_L - x_S + 1}$$

2- حساب عدد فئات التوزيع: لحساب عدد فئات التوزيع التكراري يمكن اعتماد احدى الصيفتين التقريبتين الاتيتين:

صيغة يول (Yule)

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

صيغة سترجس (Sturgess)

$$\mathbf{m = 1 + 3.322\log_{10}(n)}$$

وعند التطبيق يتم تقريب الناتج لاقرب عدد صحيح.

3- حساب طول الفئة (المدى الفئوي) : يمثل طول الفئة مقدار سعة الفئة اي مقدار المسافة بين الحد الادنى والحد الاعلى للفئة، ويرمز له (L).

$$\mathbf{L = \frac{T.R}{m}}$$

4- تحديد الحد الاعلى والحد الادنى للفئة: لكل فئة حد ادنى وحد اعلى. الحد الادنى يمثل بداية الفئة والحد الاعلى يمثل نهاية الفئة. ويمكن تحديد الحد الادنى والحد الاعلى للفئات اذا كانت اطوالها متساوية كالاتي:

تسلسل الفئة	الحد الادنى	الحد الاعلى
1	x_S	x_S+L
2	x_S+L	x_S+2L
3	x_S+2L	x_S+3L
. .	. .	. .
m	$x_S+(m-1)L$	x_S+mL

وهناك طرق مختلفة لكتابة حدود الفئات للتوزيع التكراري استنادا الى نوع المتغير العشوائي، وكالاتي:

أ- **في حالة المتغيرات المتقطعة**: تكتب الفئات كما يلي:

تسلسل الفئة	الحد الادنى	الحد الاعلى
1	x_S	x_S+L-1
2	x_S+L	x_S+2L-1
3	x_S+2L	x_S+3L-1
. . .	. . .	. . .
m	$x_S+(m-1)L$	x_S+mL-1

مع ملاحظة انه يجب ضمان ان كل قيمة من قيم البيانات تكون في فئة واحدة فقط دون ان تتكرر في الفئات الاخرى.

ان طول الفئة عندما يكون المتغير العشوائي متقطعا يحسب بالصيغة ادناه:

$$L=U.L - L.L + 1$$

U.L : الحد الاعلى للفئة

L.L : الحد الادنى للفئة

ب- في حالة المتغيرات المستمرة: تكتب الفئات كما يلي:

الحد الاعلى		الحد الادنى	تسلسل الفئة
x_S+L	Less than	x_S	1
x_S+2L	Less than	x_S+L	2
x_S+3L	Less than	x_S+2L	3
.		.	.
.		.	.
.		.	.
x_S+mL	Less than	$x_S+(m-1)L$	m

او تكتب اختصارا بذكر الحد الادنى للفئات ما عدا الاخيرة وكما يلي:

الفئات	تسلسل الفئة
x_S —	1
x_S+L —	2

3 x_S+2L —

. .

. .

m $x_S+(m-1)L$ — x_S+mL

ان طول الفئة عندما يكون المتغير العشوائي مستمرا يحسب بالصيغة ادناه:

L=U.L – L.L

بافتراض ان الحد الاعلى لكل فئة هو الحد الادنى للفئة السابقة لها.

5- حساب تكرار كل فئة: يمثل تكرار الفئة ذلك الجزء من مفردات العينة التي تقع ضمن تلك الفئة. بحيث ان مجموع هذه الاجزاء (تكرارات الفئات) يشكل عدد مفردات العينة، ويرمز لتكرار الفئة بالرمز (f_i).

مركز الفئة: هي قيمة من قيم المتغير العشوائي X التي تتوسط المسافة بين الحد الادنى والحد الاعلى للفئة، ويرمز لمركز الفئة بالرمز X ويحسب كالاتي:

$$\mathbf{X = \frac{L.L + U.L}{2}} \quad , \quad \textbf{or} \ ,$$

$\mathbf{X = L.L + 0.5\ (L - 1)}$ في حالة المتغير العشوائي المتقطع

$\mathbf{X = L.L + 0.5\ L}$ في حالة المتغير العشوائي المستمر

ويمثل مركز الفئة كل قيمة تقع ضمن تلك الفئة، اي ان كل قيمة واقعة ضمن الفئة ينظر لها وكأنها تساوي مركز تلك الفئة من حيث القيمة العددية.

انواع التوزيعات التكرارية

1- توزيع تكراري مفتوح:

أ- مفتوح من احد طرفيه: ويكون فيه الحد الادنى للفئة الاولى غير معلوم، او يكون فيه الحد الاعلى للفئة الاخيرة غير معلوم.

ب- مفتوح من الطرفين: ويكون فيه الحد الادنى للفئة الاولى غير معلوم، وكذلك فيه الحد الاعلى للفئة الاخيرة غير معلوم.

<u>2- توزيع تكراري مغلق</u>: وفيه يذكر الحد الادنى للفئة الاولى ، وكذلك يذكر فيه الحد الاعلى للفئة الاخيرة.

كما يمكن تقسيم التوزيعات التكرارية من حيث **طول الفئة** الى نوعين :

<u>1- توزيع تكراري منتظم</u>: وفيه تكون اطوال الفئات متساوية.

<u>2- توزيع تكراري غيرمنتظم</u>: وفيه تكون اطوال الفئات غيرمتساوية.

مثال: البيانات التالية تمثل عدد الطلبة المرحلة الاولى في 60 مدرسة ابتدائية. **المطلوب** تبويب البيانات في توزيع تكراري.

60	76	80	120	132	82	90	65	68	72	150	142
157	164	88	90	98	101	103	110	119	116	120	126
109	114	120	122	111	116	90	78	93	95	98	104
120	113	121	119	125	126	130	131	136	118	120	142
150	154	122	123	139	125	156	154	136	137	110	136

الحل:

ان المتغير العشوائي الذي يمثل اعداد الطلبة هو متغير متقطع (ارقامه صحية وليست كسرية). لذا نستعمل توزيعاً تكراياً خاصاً بالمتغيرات المتقطعة.

القيمة الكبرى $X_L=164$, القيمة الصغرى $X_S=60$, عدد المشاهدات $n= 60$

المدى الكلي $T.R = 164 - 60 + 1 = 105$

عدد الفئات $m = 2.5\sqrt[4]{60} = 2.5(2.783) = 6.958 \approx 7$

طول الفئة $\mathbf{L} = \frac{\mathbf{T.R}}{\mathbf{m}} = \frac{105}{7} = 15$

نكتب حدود الفئات

الفئات	الحد الادنى	الحد الاعلى
60 – 74	60	60 + 15 – 1
75 – 89	60+15	60 + 30 – 1
90 – 104	60+30	60 + 45 – 1
105 – 119	60 + 45	60 + 60 – 1
120 – 134	60 + 60	60 + 75 – 1
135 – 149	60 + 75	60 + 90 – 1
150 – 164	60 + 90	60 + 105 – 1

توزيع البيانات على الفئات كما في الجدول ادناه:

الفئات	الحزم	التكرار fi	مركز الفئات X

60 – 74	IIII	4	67
75 – 89	𝍸	5	82
90 – 104	𝍸𝍸	10	97
105 – 119	𝍸𝍸I	11	112
120 – 134	𝍸𝍸𝍸I	16	127
135 – 149	𝍸II	7	142
150 – 164	𝍸II	7	157
المجموع		60	

ملاحظة: الرمز 𝍸 يشير الى الرقم 5.

والجدول التكراري اعلاه يشير الى ان:

4 مدارس لديها طلبة المرحلة الاولى بين 60 الى 74 طالب

10 مدارس لديها طلبة المرحلة الاولى بين 75 الى 89 طالب

11 مدرسة لديها طلبة المرحلة الاولى بين 90 الى 104 طالب. وهكذا بقية الجدول.

وتم حساب مراكز الفئات كما يلي:

مركز الفئة الاولى $X_1 = \frac{60+74}{2} = 67$

مركز الفئة الثانية $X_2 = \frac{75+89}{2} = 82$

مركز الفئة الثالثة $X_3 = \frac{90+104}{2} = 97$

وهكذا لبقية الفئات.

واجب بيتي:

اختر 20 رقما صحيحا مختلفا، وكون جدول التوزيع التكراري لها.

مثال2: البيانات التالية تمثل اوزان عينة من طلبة احدى الكليات قوامها 20 طالب. **المطلوب** تفريغ البيانات في جدول توزيع تكراري، ومن ثم حساب مراكز الفئات.

65.3	70.5	83	94	55	46	47.8	62.3	77.2	61.3
66.2	95	68.3	80.2	78.3	80.1	76.3	81.5	51.8	74.1

الحل:

ان المتغير العشوائي الذي يمثل اوزان الطلبة هو متغير مستمر (ارقامه تتقبل الكسور). لذا نستعمل توزيعاً تكراراياً خاصاً بالمتغيرات المستمرة.

$n = 20 \qquad , X_S = 46 \qquad , X_L = 95$

$T.R = 95 - 46 + 1 = 50$ المدى الكلي

$m = 1 + 3.322\log(20)$

$m = 1 + 3.322(1.30103)$

$m = 1 + 4.32202 = 5.322 \approx 5$ عدد الفئات

$\boldsymbol{L} = \frac{\textbf{T.R}}{\boldsymbol{m}} = \frac{50}{5} = 10$ طول الفئة

نكتب حدود الفئات

الحد الادنى للفئات	الحد الاعلى للفئات
46 —	Less than 46 + 10 = 56
56 —	Less than 46 + 20 =66
66 —	Less than 46 + 30 = 76
76 —	Less than 46 + 40 = 86
86 —	Less than 46 + 50 = 96

توزيع البيانات على الفئات كما في الجدول ادناه:

الفئات	الحزم	التكرار fi	مركز الفئات X
46 —	IIII	4	51

56 –	III	3	61
66 –	IIII	4	71
76 –	~~IIII~~ II	7	81
86 – 96	II	2	91
المجموع		20	

ملاحظة: الرمز ~~IIII~~ يشير الى الرقم 5.

وتم حساب مراكزالفئات كما يلي:

$X_1 = \frac{46+56}{2} = 51$ مركز الفئة الاولى

$X_2 = \frac{56+66}{2} = 61$ مركز الفئة الثانية

$X_3 = \frac{66+76}{2} = 71$ مركز الفئة الثالثة

وهكذا لبقية الفئات.

ب- الجدول التكراري النسبي

هو توزيع تكراري تكون تكرات فئاته معبرا عنها بنسب مئوية، ويرمز للتكرار النسبي بالرمز (f_i^*) ويحسب كالاتي:

$$f_i^* = \frac{f_i}{n} \times 100$$

حيث ان f_i يمثل تكرار الفئة i

وان n تمثل مجموع تكرارات الفئات

مثال 3: كون توزيع تكراري نسبي للبيانات الواردة في التوزيع التكراري للمثال 1.

الفئات	f_i	f_i^*
60 – 74	4	4/60*100=6.67%

75 – 89	5	5/60*100=8.33%
90 – 104	10	10/60*100=16.67%
105 – 119	11	11/60*100=18.33%
120 – 134	16	16/60*100=26.67%
135 – 149	7	7/60*100=11.67%
150 – 164	7	7/60*100=11.67%
المجموع	60	%100

تمرين : كون توزيع تكراري نسبي للبيانات الواردة في التوزيع التكراري للمثال 2.

جـ - التوزيع التكراري المتجمع

هو التوزيع التكراري الذي تتجمع فيه التكرارات تصاعديا او تنازليا. وهو نوعين صاعد ونازل.

1 - التوزيع التكراري المتجمع الصاعد

وهو التوزيع الذي يبين تراكم التكرارات ابتداءا من الفئة الاولى للتوزيع وانتهاءا بالفئة الاخيرة. ويتم احتساب التكرارات المتجمعة الصاعدة على اساس الحدود العليا للفئات.

ويتم عمله باتباع مايلي:

1- عمل جدول توزيع تكراري كما ذكر سابقا.

2- يتم اعادة كتابة الحدود العليا للفئات مسبوقة بعبارة اقل من *(less than)* في حالة المتغيرات المستمرة، او عبارة اقل من او يساوي *(less than or equal)* في حالة المتغيرات المتقطعة.

3- يتم تجميع التكرارات تصاعديا حيث يكون تكرار الفئة الاولى مساوياً لتكرارها الاصلي. وتكرار الفئة الثانية مساوياً لمجموع تكراري الفئتين الاولى والثانية، وهكذا لغاية الفئة الاخيرة التي يكون تكرارها المتجمع الصاعد يساوي مجموع تكرارات كل الفئات (اي عدد بيانات العينة n).

مثال 4: كون توزيعا تكراريا متجمعا صاعدا يعرض توزيع اوزان عينة من 20 طالبا وبالاعتماد على بيانات المثال 2.

الفئات	التكرار fi	الحدود العليا للفئات	التكرار المتجمع الصاعد Fi
46 —	4	Less than 56	4
56 —	3	Less than 66	7 =4+3
66 —	4	Less than 76	11=4+3+4
76 —	7	Less than 86	18=4+3+4+7
86 — 96	2	Less than 96	20=4+3+4+7+2
المجموع	20		

تمرين : كون توزيع تكراري متجمع صاعد بالاعتماد على بيانات المثال 1.

2 - التوزيع التكراري المتجمع النازل

وهو التوزيع الذي يبين تناقص التكرارات ابتداءا من الفئة الاولى للتوزيع وانتهاءا بالفئة الاخيرة. ويتم احتساب التكرارات المتجمعة النازلة على اساس الحدود الدنيا للفئات.

ويتم عمله باتباع مايلي:

1- عمل جدول توزيع تكراري كما ذكر سابقا.

2- يعاد كتابة الحدود الدنيا للفئات مسبوقة بعبارة اكبر من او يساوي *(more than or equal)* في حالة المتغيرات المستمرة، او المتقطعة.

3- يكون التكرار المتجمع النازل للفئة الاولى مساوياً للمجموع الكلي للتكرارات، اما التكرارالمتجمع النازل للفئة الثانية فيكون مساويا للتكرار المتجمع النازل للفئة الاولى مطروحا منه تكرار الفئة الاولى الاصلي. والتكرارالمتجمع النازل للفئة الثالثة فيكون مساويا للتكرار المتجمع النازل للفئة الثانية مطروحا منه تكرار الفئة الثانية الاصلي، وهكذا لغاية الفئة الاخيرة التي يكون تكرارها المتجمع النازل يساوي تكرارها الاصلي.

مثال 5: كون توزيعا تكراريا متجمعا نازلا بعرض اعداد طلبة الصف الاول ابتدائي الواردة في المثال 1.

الفئات	fi	الحدود الدنيا للفئات	$\grave{F}_i$
60 – 74	4	More than or equal 60	60
75 – 89	5	More than or equal 75	56=60-4
90 – 104	10	More than or equal 90	51=56-5
105 – 119	11	More than or equal 105	41=51-10
120 – 134	16	More than or equal 120	30=41-11
135 – 149	7	More than or equal 135	14=30-16
150 – 164	7	More than or equal 150	7=14-7
المجموع	60		

ثانيا : العرض الهندسي للبيانات

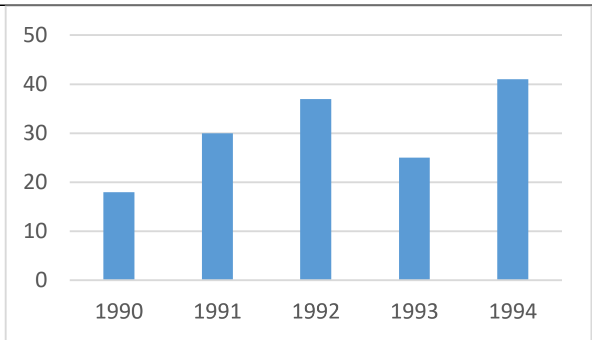
يمكن عرض البيانات بهيئة رسوم بيانية واشكال هندسية معينة لغرض اعطاء فكرة واضحة وسريعة عن البيانات. ومن وسائل عرض البيانات ما يلي:-

أ- الاشرطة البيانية:

هي مستطيلات رأسية او افقية قواعدها متساوية وتمثل الصفة التي على اساسها تم التوبيب (سنة، شهر ، محافظة، قطر ، صنف دم،....)، اما ارتفاعاتها تمثل البيانات المقابلة لتلك الصفة.

مثال : استخدم الاشرطة البيانية لعرض عدد الدورات المنفذة من قبل الجامعات العراقية للمدة 1990-1994.

السنة	عدد الدورات
1990	18
1991	30
1992	37
1993	25
1994	41



السنوات

<u>ب- المستطيل البياني</u>

هو مستطيل (نفترض له طولا معينا) مجزء الى عدة مستطيلات حيث ان طول قاعدة المستطيل الكبير تمثل البيانات الكلية اما قواعد المستطيلات الجزئية فتتناسب وحجم البيانات الجزئية. فاذا كان:

طول قاعدة المستطيل الجزئي (L_i) ، **عدد البيانات في الصنف (n_i)** ،

مجموع البيانات الكلية (n) ، **طول قاعدة المستطيل (L)**

فان:

$$L_i = \frac{n_i}{n} * \mathrm{L}$$

مثال: بلغ عدد طلاب احدى الكليات 2000 طالبا، في الصف الاول 800 طالب، وفي الصف الثاني 500 طالب، وفي الصف الثالث 400 طالب، وفي الصف الرابع 300 طالب.
استخدم المستطيل البياني لتمثيل هذه البيانات.

نفرض ان طول المستطيل الكبير = 10سم

طول قاعدة المستطيل للصف الاول $L_1 = \frac{800}{2000} * 10 = \frac{8}{2} = 4 \text{ cm}$

طول قاعدة المستطيل للصف الثاني $L_2 = \frac{500}{2000} * 10 = 2.5 \text{ cm}$

طول قاعدة المستطيل للصف الثالث $L_3 = \frac{400}{2000} * 10 = 2 \text{ cm}$

طول قاعدة المستطيل للصف الرابع $L_4 = \frac{300}{2000} * 10 = 1.5 \text{ cm}$

الصف الاول	الصف الثاني	الصف الثالث	الصف الرابع

جـ - الدائرة البيانية

هي دائرة مقسمة الى عدد من القطاعات التي تمثل البيانات الجزئية في حين تمثل الدائرة البيانات الكلية، وتحسب زاوية كل قطاع كما يلي:

$$a_i = \frac{n_i}{n} * 360^0$$

حيث ان: زاوية القطاع (a_i) ، عدد بيانات الصنف (n_i) ، مجموع البيانات الاصلية (n)

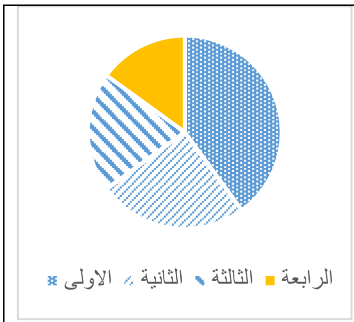
ولبيانات المثال السابق يكون:

زاوية القطاع للصف الاول = $a_1 = \frac{800}{2000} * 360^0 = 144^0$

زاوية القطاع للصف الثاني= $a_2 = \frac{500}{2000} * 360^0 = 90^0$

زاوية القطاع للصف الثالث= $a_3 = \frac{400}{2000} * 360^0 = 72^0$

زاوية القطاع للصف الرابع= $a_4 = \frac{300}{2000} * 360^0 = 54^0$



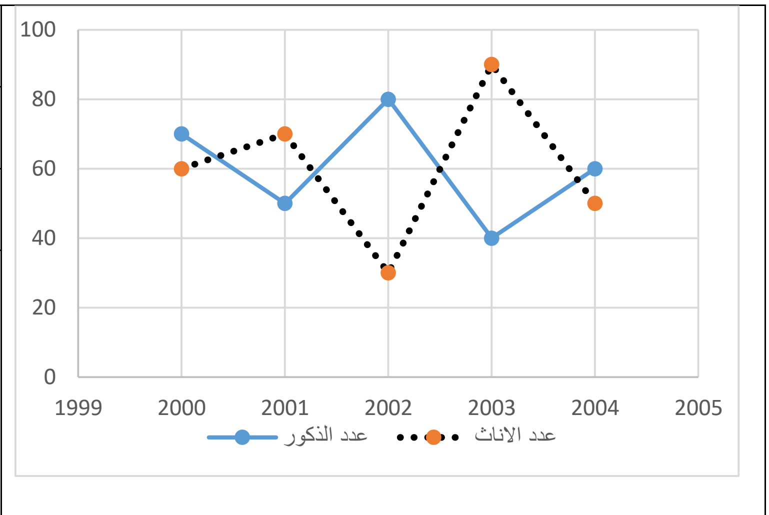
د – الخط البياني

هو شكل بياني يوضح التغيرات الحاصلة في ظاهرة معينة عبر الزمن. وهو شكل نافع في حالة اجراء مقارنة بين ظاهرتين اواكثر مقاسة بنفس وحدات القياس. مثل مقارنة التغيرات الحاصلة بين تكاليف انتاج سلعة معينة والارباح المتحققة من مبيعات تلك السلعة خلال فترة زمنية معينة.

مثال: قارن بين عدد الذكور والاناث للمرحلة الاولى خلال السنوات 2000 -2004 باستخدام الخط البياني.

السنوات	عدد الذكور	عدد الاناث	
2000	70	60	

2001	50	70
2002	80	30
2003	40	90
2004	60	50



المجموع (Summation)

نحتاج في التحليل الاحصائي الى عملية جمع سلسلة من الاعداد والكميات. فمثلا اذا كانت لدينا عينة مؤلفة من خمسة طلاب سجلت درجاتهم في امتحان الاحصاء، ولو فرضنا ان (X) هو المتغير العشوائي الذي يمثل درجة الطالب في امتحان الاحصاء، فان $(x_1, x_2, x_3 , x_4, x_5)$ تمثل درجات الطلاب من الاول الى الخامس على التوالي. وعندئذ فان المجموع الكلي للدرجات هو

$(x_1 + x_2 + x_3 + x_4 + x_5)$

وبهدف تسهيل عملية كتابة المجموع بطريقة اكثر اختصارا، فيعبر عنه بالشكل $\sum_{i=1}^{5} x_i$ حيث ان الرمز $\sum$ يشير الى عملية الجمع، وهو حرف اغريقي يلفظ (Sigma) وان i يمثل دليل او تسلسل العدد في عملية الجمع فاذا كان (i=1) فان فهذا يعني تسلسل العدد الاول، وعندما (i=4) فان تسلسل العدد هو الرابع وهكذا. وتقرأ مجموعة سلسلة الاعداد ابتداءا بالعدد من التسلسل الاول وانتهاءا بالعدد في التسلسل الخامس.

وبشكل عام اذا كان لدينا n من القيم وكل قيمة هي x_i فان المجموع الكلي للقيم تكتب بالشكل:

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n$$

خواص المجموع

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

2- $\sum_{i=1}^{n} a x_i = a \sum_{i=1}^{n} x_i$, a عدد ثابت

3- $\sum_{i=1}^{n}(x_i \mp a) = \sum_{i=1}^{n} x_i \mp na$, a عدد ثابت

4- $\sum_{i=1}^{n}(x_i y_i) \neq \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i$

5- $\sum_{i=1}^{n} x_i^2 \neq (\sum_{i=1}^{n} x_i)^2$

مثال: اذا كان لدينا المتغيرين X,Y بالقيم الاتية:

x_i	2	4	5	8	3
y_i	1	2	3	2	4

احسب

$\sum_{i=1}^{n} x_i$, $\sum_{i=1}^{n} y_i$, $\sum_{i=1}^{n} x_i + \boldsymbol{y_i}$, $\sum_{i=1}^{n} x_i - \boldsymbol{y_i}$, $\sum_{i=1}^{n} 3x_i$

$\sum_{i=1}^{n} x_i - \mathbf{4}$, $\sum_{i=1}^{n} x_i\, \boldsymbol{y_i}$, $\sum_{i=1}^{n} x_i^2$, $\sum_{i=1}^{n} y_i^2$

فان:

x	y	x+y	x-y	3x	x-4	xy	x²	y²
2	1	3	1	6	-2	2	4	1
4	2	6	2	12	0	8	16	4
5	3	8	2	15	1	15	25	9
8	2	10	6	24	4	16	64	4
3	4	7	-1	9	-1	12	9	16
$\sum_{i=1}^{n} x_i = 22$	$\sum_{i=1}^{n} y_i = 12$	$\sum_{i=1}^{n} x_i$ 34	$\sum_{i=1}^{n} x_i$ 10	$\sum_{i=1}^{n} 3x_i = 66$	$\sum_{i=1}^{n} x_i$ 2	$\sum_{i=1}^{n} x_i y_i = 53$	$\sum_{i=1}^{n} x_i^2$ 118	$\sum_{i=1}^{n} y_i^2$ 34

لاحظ الخواص:

$$\left[(\mathbf{22})^2 = \mathbf{484} = \left(\sum_{i=1}^{n} x_i\right)^2\right] \neq \left[\sum_{i=1}^{n} x_i^2 = \mathbf{118}\right]$$

$$\left[\mathbf{22} * \mathbf{12} = \mathbf{264} = \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i\right] \neq \left[\sum_{i=1}^{n} x_i\, y_i = \mathbf{53}\right]$$

$$\sum_{i=1}^{n} 3x_i = 66 = 3\sum_{i=1}^{n} x_i = 3 * 22 = 66$$

تمرين: اذا كان لدينا المتغيرين X , Y بالقيم الاتية:

x_i	2	3	4	2	1	5	4	6	5
y_i	3	6	2	5	2	1	4	4	3

احسب

$\sum_{i=1}^{n} x_i$, $\sum_{i=1}^{n} y_i$, $\sum_{i=1}^{n} x_i + \boldsymbol{y_i}$, $\sum_{i=1}^{n} |x_i - \boldsymbol{y_i}|$, $\sum_{i=1}^{n} 2y_i$

$\sum_{i=1}^{n} y_i + \mathbf{4}$, $\sum_{i=1}^{n} (x_i / \boldsymbol{y_i})$, $\sum_{i=1}^{n} x_i^2$, $\sum_{i=1}^{n} y_i^2$

مقاييس النزعة المركزية (المتوسطات)

وهي مقاييس احصائية تستخدم لايجاد قيمة واحدة تمثل المجتمع الاحصائي، وسميت بهذا الاسم لانها تميل الى التمركز في وسط قيم المجتمع الاحصائي، ومنها ما يلي:

اولاً : الوسط الحسابي (Mean or Average)

وهو اهم مقاييس النزعة المركزية لما يمتاز به من خصائص جيدة، وسهولة حسابه، ويحسب لاي مجموعة قيم، بقسمة مجموع القيم على عددها.

أ- حساب الوسط الحسابي للبيانات غير المبوبة

توجد اكثر من طريقة لحسابه ومنها:

1- الطريقة المباشرة (باستخدام القيم الاصلية)

لتكن $x_1, x_2, \ldots , x_n$ عينة عشوائية من البيانات عددها n التي تمثل المتغير العشوائي X .
فالوسط الحسابي والذي يرمزله بالرمز $\overline{X}$ يحسب حسب الصيغة الاتية:

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

مثال 1: البيانات التالية تمثل اعداد العاملين في 15 منشأة حكومية، المطلوب تقدير الوسط الحسابي لعدد العاملين من بيانات هذه العينة.

50 60 68 59 58 62 65 52 61 63 59 69 64 65 55

الحل:

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{50+60+\cdots+55}{15} = \frac{910}{15} = 60.667 \qquad \text{عامل}$$

2- الطريقة المختصرة (طريقة الانحرافات)

تستخدم في حال كون قياسات العينة اعددها كبيرة يصعب التعامل معها، لذا يفضل اختزال هذه الاعداد الى اعداد اصغر يسهل التعامل معها، ولنفرض ان $x_1, x_2, \ldots , x_n$ تمثل قياسات العينة

التي عددها n . وليكن a اي قيمة ثابتة (من البيانات او من خارجها)، فان الوسط الحسابي يحسب بالصيغة:

$$\overline{X} = a + \frac{\sum_{i=1}^{n} d_i}{n} \quad , \qquad d_i = x_i - a$$

ويسمى الثابت a وسطا فرضيا، ويفضل ان يكون احد قيم البيانات القريبة من مركزها.

مثال 2: استخدم الطريقة المختصرة لحساب متوسط عدد العاملين للمثال 1.

الحل : ليكن (a = 55) فيتم حساب d_i كما ياتي:

x_i	x_i-55	=	d_i
59	59- 55	=	4
58	58- 55	=	3
62	62- 55	=	7
65	65- 55	=	10
52	52- 55	=	-3
61	61- 55	=	6
63	63- 55	=	8
59	59- 55	=	4
69	69- 55	=	14
64	64- 55	=	9
65	65- 55	=	10
55	55- 55	=	0
910			85

$$\overline{X} = a + \frac{\sum_{i=1}^{n} d_i}{n} = 55 + \frac{85}{15} = 55 + 6.667 = 60.667$$ عامل

تمرين: احسب الوسط الحسابي بالطريقتين المباشرة والانحرافات للبيانات التالية والتي تمثل وزن ثمانية اشخاص.

67 90 76 59 63 102 84 91

ب- حساب الوسط الحسابي للبيانات المبوبة

توجد اكثر من طريقة لحسابه ومنها:

1- الطريقة المباشرة (باستخدام القيم الاصلية)

لتكن $x_1, x_2, \dots , x_m$ تمثل مراكز الفئات لتوزيع تكراري عدد فئاته m. وان $f_1, f_2, \dots , f_m$ تمثل التكرارات المقابلة لهذه الفئات. فان الوسط الحسابي $\overline{X}$ يحسب (سواء كانت الفئات متساوية بالطول او غير متساوية. وسواء كان المتغير العشوائي متغيرا متقطعا او مستمرا) حسب الصيغة الاتية:

$$\overline{X} = \frac{\sum_{i=1}^{m} x_i f_i}{\sum_{i=1}^{m} f_i}$$

مثال 1: الاتي هو جدول توزيع تكراري لعينة قوامها 75 اسرة، مبوبة حسب عدد افراد الاسرة ، المطلوب حساب الوسط الحسابي لعدد افراد الاسرة في العينة.

Classes	f_i
2 – 4	8
5 – 7	12
8 – 10	20
11 – 13	13
14 – 16	10
17 – 19	8
20 – 22	4

الحل:

Classes	f_i	x_i	x_if_i
2 – 4	8	3	24
5 – 7	12	6	72
8 – 10	20	9	180
11 – 13	13	12	156
14 – 16	10	15	150
17 – 19	8	18	114
20 – 22	4	21	84
	75		810

$\overline{X} = \frac{\sum_{i=1}^{m} x_i f_i}{\sum_{i=1}^{m} f_i} = \frac{810}{75} = 10.8$ فرد

مثال 2: الاتي هو جدول توزيع تكراري لرواتب مجموعة من الموظفين حسب فئات الراتب الاسمي بالدينار، المطلوب حساب الوسط الحسابي لرواتب الموظفين.

Classes	f_i
69.5 –	6
72.5 –	10
76.5 –	18
88.5 –	25
97.5 –	12
112.5 – 130.5	7

الحل: **ملاحظة: ان اطوال الفئات غير متساوية لكن ذلك لا يؤثر على الحل**

Classes	f_i	x_i	x_if_i
69.5 –	6	71	426
72.5 –	10	74.5	745
76.5 –	18	82.5	1485
88.5 –	25	93	2325
97.5 –	12	105	1260
112.5 – 130.5	7	121.5	850.5
	78		7091.5

$$\overline{X} = \frac{\sum_{i=1}^{m} x_i f_i}{\sum_{i=1}^{m} f_i}$$

$$= \frac{7091.5}{78} = 90.917$$

دينار متوسط الراتب

2- الطريقة المختصرة للبيانات المبوبة

بالامكان اختصار قيم مراكز الفئات بطرح قيمة ثابتة تساوي a (وسط فرضي) من كل منها، وهذا الوسط الفرضي قد يكون مساوياً لاحد مراكز الفئات او لايساويه، ويفضل ان يكون الوسط الفرضي مساويا لمركز الفئة ذات التكرار الاكبر، ويحسب الوسط الحسابي كما يأتي:

$$\bar{X} = a + \frac{\sum_{i=1}^{m} d_i f_i}{\sum_{i=1}^{m} f_i} \quad , \quad d_i = x_i - a$$

مثال 3: الاتي هو جدول توزيع تكراري لاطوال عينة حجمها 70 شخص مقاسة بالسنتيمتر، المطلوب حساب الوسط الحسابي للطول بالطريقة المختصرة.

Classes	f_i
140 –	4
148 –	6
156 –	15
164 –	20
172 –	17
180 –	7
188 – 196	1

الحل: ليكن $(a = 168)$ وهي قيمة مركز الفئة الرابعة التي تقابل اكبر تكرار.

Classes	f_i	x_i	$d_i=x_i -168$	d_if_i
140 –	4	144	-24	-96
148 –	6	152	-16	-96
156 –	15	160	-8	-120
164 –	20	168	0	0
172 –	17	176	8	136
180 –	7	184	16	112
188 – 196	1	192	24	24
	70			-40

متوسط الطول $\bar{X} = a + \frac{\sum_{i=1}^{m} d_i f_i}{\sum_{i=1}^{m} f_i} = 168 + \frac{-40}{70} = 168 - 0.571 = 167.43cm$

تمرين : احسب متوسط الطول للمثال 3 باستعمال الطريقة المباشرة

الوسط الحسابي المرجح (الموزون)

بعض الظواهر تختلف من حيث الاهمية النسبية، وتسمى هذه الاهمية النسبية بالوزن او الترجيح لكل مفردة، وعند حساب الوسط الحسابي لتلك القيم يؤخذ الوزن لكل قيمة، لذا يسمى بالوسط الحسابي الموزون او المرجح ويرمز له بالرمز $\bar{X}_w$ ويحسب كالاتي:

$$\bar{X}_w = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i}$$

حيث ان:

$x_1, x_2, \dots, x_n$: تمثل قيم الظاهرة

$w_1, w_2, \dots, w_n$: تمثل الاوزان

مثال 4: الجدول الاتي يمثل درجات احد الطلبة في اربع مواد دراسية وعدد الساعات الاسبوعية لتلك المواد. المطلوب حساب معدل الطالب في تلك المواد مرجحا بعدد الساعات.

الدرجة	عدد الساعات
60	3
65	3
50	6
70	2

الحل:

الدرجة x_i	عدد الساعات w_i	$x_i\, w_i$
60	3	180
65	3	195
50	6	300
70	2	140
	14	815

$$\bar{X}_w = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i} = \frac{815}{14} = 58.2$$

مزايا الوسط الحسابي

1- سهولة حسابة وبساطة فكرته.

2- خضوعه للعمليات الحسابية.

عيوب الوسط الحسابي

1- يتاثر بالقيم الشاذة (الكبيرة جدا والصغيرة جدا)

2- لايمكن حسابه للتوزيعات التكرارية المفتوحة.

3- لايمكن حسابة في حالة فقدان قيمة اواكثر.

4- لايمكن حسابه للمتغيرات النوعية، الا اذا امكن تحويلها الى متغيرات كمية.

5- لا يفضل استخدامه في حالة التوزيعات غير المتماثلة.

خواص الوسط الحسابي

1- مجموع انحرافات القيم عن وسطها الحسابي يساوي صفرا. اي ان:

<u>أ- في حالة البيانات غير المبوبة</u>

$$\sum_{i=1}^{n}(x_i-\bar{x})=0$$

والبيانات الاتية توضح ذلك.

x_i		$(x_i-\bar{x})$
3	$\bar{X}=\frac{\sum_{i=1}^{n}x_i}{n}=\frac{20}{4}=5$	$3-5=-2$
8		$8-5=\ \ 3$
2		$2-5=-3$
7		$7-5=\ \ 2$
20		0

$$\sum_{i=1}^{n}(x_i-\bar{x})=(3-5)+(8-5)+(2-3)+(7-2)$$
$$=-2+3-3+2=0$$

ب- في حالة البيانات المبوبة

$$\sum_{i=1}^{n}(x_i-\bar{x})f_i=0$$

للبيانات ادناه نحسب الوسط الحسابي ونجده يساوي 38

Classes	f_i	x_i	$x_i f_i$	x_i-38	**(x_i-38)f_i**
10 –	9	15	135	-23	**-207**
20 –	8	25	200	-13	**-104**
30 –	18	35	630	-3	**-54**
40 –	23	45	1035	7	**161**
50 – 60	12	55	660	17	**204**
	70		2660		**0**

$$\bar{X}=\frac{\sum_{i=1}^{m}x_i f_i}{\sum_{i=1}^{m}f_i}=\frac{2660}{70}=38$$

2- مجموع مربعات انحرافات القيم عن وسطها الحسابي يكون اقل ما يمكن.اي ان:

$$\sum_{i=1}^{n}(x_i-\bar{x})^2\le\sum_{i=1}^{n}(x_i-a)^2$$

عند اي قيمة لـ a ، ويتساوى الطرفان فقط عندما $a=\bar{x}$.

فلمثالنا السابق (في الخاصية 1- أ) فان الوسط الحسابي يساوي 5 $\bar{x}=5$ ، ولنفرض مرة a=3 ومرة اخرى a=6 ونحسب مجاميع المربعات فنرى ان اقلها هو مجموع مربعات انحرافات القيم عن الوسط الحسابي الذي هو 5.

x_i	$(x_i-\bar{x})$	$(x_i-\bar{x})^2$	(x_i - 3)	**(x_i - 3)^2**	(x_i - 6)	**(x_i - 6)^2**
3	-2	**4**	0	**0**	-3	**9**
8	3	**9**	5	**25**	2	**4**
2	-3	**9**	-1	**1**	-4	**16**
7	2	**4**	4	**16**	1	**1**
20		**26 الاقل**		**42**		**30**

3- اذا كان لدينا مجتمع احصائي يتكون من عدد (k) من المجاميع الجزئية (وكل مجموعة حجمها n_i) فبامكاننا حساب الوسط الحسابي للمجتمع الاحصائي الرئيس بالاعتماد على الاوساط الحسابية للمجاميع الجزئية وكالاتي:

$$\bar{X} = \frac{\sum_{i=1}^{k} \bar{x}_i n_i}{\sum_{i=1}^{k} n_i}$$

حيث ان :

$$\sum_{i=1}^{k} n_i = n_1 + n_2 + \cdots + n_k$$

مثال: الجدول التالي يبين الاوساط الحسابية لدرجات ثلاث شعب واعداد الطلبة لكل شعبة لاحد الصفوف الدراسية. المطلوب حساب الوسط الحسابي لذلك الصف.

الشعبة	عدد الطلبة n_i	الوسط الحسابي $\bar{x}_i$	$n_i\, \bar{x}_i$
a	50	60	3000
b	55	50	2750
c	60	65	3900
	165		9650

$$\bar{X} = \frac{\sum_{i=1}^{k} \bar{x}_i n_i}{\sum_{i=1}^{k} n_i} = \frac{9650}{165} = 58.48$$

تمارين الوسط الحسابي

1- الاتي توزيع تكراري للدخل الشهري لعينة من الاسر. المطلوب حساب متوسط الدخل الشهري للاسرة في هذه العينة

classes	100 –	120 –	140 –	160 –	180 –	200 –220
f_i	4	5	6	7	10	8

2- الاتي توزيع تكراري لعدد النداءات الهاتفية العاجلة التي استقبلتها مستشفى معينة خلال ثلاثون يوما. المطلوب حساب متوسط عدد النداءات اليومية.

classes	10 – 19	20 – 29	30 –39	40 –49	50 –59	60 –69
f_i	10	7	4	4	3	2

3- جد قيمة w التي تجعل الوسط الحسابي لقيم المجموعة ادناه يساوي 30.

37 w 31 28 25 20

4- اذا علمت بان الوسط الحسابي للتوزيع التالي يساوي 39 فجد قيمة b.

classes	10 –	20 –	30 –	40 –	50 –60
f_i	2	4	6	8	b

5- مصنع مؤلف من ثلاثة اقسام، عدد العاملين في كل منها ومتوسط الاجر الشهري موضح ادناه. المطلوب حساب الوسط الحسابي لاجر العامل الشهري في المصنع.

القسم	A	B	C
عدد العاملين	95	80	106
متوسط الاجر	110	170	99

6- جد الوسط الحسابي المرجح للبيانات الاتية:

A-

x_i	10	18	15	11	19	22	20
w_i	1	4	3	2	5	7	6

B-

x_i	-1	0	2	-2	-3	1	-4
w_i	1.5	1.3	1.4	1.6	1.5	1.3	1.9

ثانياً الوسيط The Median:

هو احد مقاييس النزعة المركزية ، وهو القيمة التي تتوسط مجموعة القيم بعد ترتيبها تصاعديا او تنازليا، ويرمز له بالرمز (Me) .

أ - حساب الوسيط للبيانات غير المبوبة

- ترتيب القيم تصاعديا او تنازليا.
- اذا كان عدد القيم:-

فرديا فان الوسيط هو القيمة التي تقابل الترتيب $\frac{n+1}{2}$.

زوجيا فان الوسيط هو الوسط الحسابي للقيمتين المقابلتين للترتيبين $\frac{n}{2}+1$ ، $\frac{n}{2}$.

(عدد القيم التي هي اقل من الوسيط = عدد القيم التي هي اكبر من الوسيط)

مثال 1: الاتي درجات تسعة طلاب . المطلوب حساب الوسيط.

65	45	50	52	61	62	70	75	80

الحل : نرتب القيم تصاعديا

1	2	3	4	5	6	7	8	9
45	50	52	61	**62**	65	70	75	80

ان عدد القيم فردي، لذلك فان ترتيب الوسيط يكون:

$$\text{ترتيب الوسيط} = \frac{n+1}{2} = \frac{9+1}{2} = 5$$

الوسيط هو القيمة التي تقابل الترتيب الخامس وعليه

Me= 62

مثال 2: الاتي اوزان عشرة طلاب . المطلوب حساب الوسيط.

30	48	50	31	38	35	40	45	37	39

الحل : نرتب القيم تصاعديا

1	2	3	4	5	6	7	8	9	10
30	31	35	37	**38**	**39**	40	45	48	50

ان عدد القيم زوجي ، لذلك فان ترتيب الوسيط يكون:

ترتيبي الوسيط $= \left[\frac{n}{2} = \frac{10}{2} = 5 \quad , \quad \frac{n}{2} + 1 = \frac{10}{2} + 1 = 6\right]$

الوسيط هو الوسط الحسابي للقيمتين الخامسة والسادسة بعد الترتيب

$$\text{Me} = \frac{38+39}{2} = 38.5$$

أ - حساب الوسيط للبيانات المبوبة

يحسب الوسيط بالاعتماد على التوزيع التكراري المتجمع الصاعد او النازل

1- حساب الوسيط للتوزيع التكراري المتقطع

- عمل توزيع تكراري متجمع صاعد.
- حساب ترتيب الوسيط الذي يكون $F_k = \frac{\sum_{i=1}^{m} f_i}{2}$.
- تثبيت ترتيب الوسيط بين التكرارات المتجمعة الصاعدة. ثم تحديد التكرار المتجمع الصاعد اللاحق لترتيب الوسيط (F_{k+1})، والتكرار المتجمع الصاعد السابق لترتيب الوسيط (F_{k-1}).
- تحديد الفئة الوسيطية باعتبارها الفئة التي تقابل التكرار المتجمع الصاعد اللاحق لترتيب الوسيط (F_{k+1}).
- يحسب الوسيط باعتباره مركز الفئة الوسيطية.

مثال 3: التوزيع التكراري الاتي يبين توزيع 80 اسرة حسب عدد الافراد. المطلوب حساب الوسيط لعدد افراد الاسرة.

Classes	f_i
2 – 4	6
5 – 7	9
8 – 10	12
11 – 13	20
14 – 16	14
17 – 19	11
20 – 22	8

الحل:

Classes	f_i	الحدود العليا للفئات	F_i		
2 – 4	6	Less than or equal 4	6		
5 – 7	9	Less than or equal 7	15		
8 – 10	12	Less than or equal 10	27	← F_{k-1}	←F_k=40
11 – 13	**20**	**Less than or eq. 13**	**47**	← $\mathbf{F_{k+1}}$	
14 – 16	14	Less than or equal 16	61		
17 – 19	11	Less than or equal 19	72		
20 – 22	8	Less than or equal 22	80		
	80				

ترتيب الوسيط $F_k = \frac{\sum_{i=1}^{m} f_i}{2} = \frac{80}{2} = 40$

الفئة الوسيطية هي **(11–13)**

الوسيط $Me = \frac{11+13}{2} = 12$

2- حساب الوسيط للتوزيع التكراري المستمر

- عمل توزيع تكراري متجمع صاعد.
- حساب ترتيب الوسيط الذي يكون $F_k = \frac{\sum_{i=1}^{m} f_i}{2}$.
- تثبيت ترتيب الوسيط بين التكرارات المتجمعة الصاعدة. ثم تحديد التكرار المتجمع الصاعد اللاحق لترتيب الوسيط (F_{k+1})، والتكرار المتجمع الصاعد السابق لترتيب الوسيط (F_{k-1}).
- تحديد الفئة الوسيطية باعتبارها الفئة التي تقابل التكرار المتجمع الصاعد اللاحق لترتيب الوسيط (F_{k+1}).
- يحسب الوسيط وفق الصيغة الاتية: $Me = h_k + \frac{F_k - F_{k-1}}{f_k} \times L_k$

سواء كانت الفئات متساوية او غير متساوية.

حيث ان:

h_k : الحد الادنى للفئة الوسيطية

F_{k-1} : التكرار المتجمع الصاعد السابق لترتيب الوسيط

f_k : تكرار الفئة الوسيطية

L_k : طول الفئة الوسيطية

مثال 4: التوزيع التكراري التالي يمثل اعمار عينة من 90 طالب. المطلوب تقدير متوسط العمر باستخدام الوسيط.

Classes	f_i
Less than 6	3
6 –	6
7 –	9
8 –	12
9 –	**20**
10 –	18
11 –	17
12 or more	5

الحل:

Classes	f_i	الحدود العليا للفئات	F_i		
Less than 6	3	Less than 6	3		
6 –	6	Less than 7	9		
7 –	9	Less than8	18		
8 –	12	Less than9	30	← F_{k-1}	← F_k=45
9 –	**20**	**Less than 10**	**50**	← $\mathbf{F_{k+1}}$	
10 –	18	Less than11	68		
11 –	17	Less than12	85		
12 or more	5	Less than ∞	90		
	90				

ترتيب الوسيط $\mathrm{F}_k = \frac{\sum_{i=1}^{m} f_i}{2} = \frac{90}{2} = 45$

الفئة الوسيطية هي(9 ــ) وحدها الادنى يساوي 9 وطولها يساوي 1.

$$Me = h_k + \frac{F_k - F_{k-1}}{f_k} \times L_k$$

$$Me = 9 + \frac{45 - 30}{20} * 1 = 9 + \frac{15}{20} = 9 + 0.75 = 9.75$$

ويمكن حساب الوسيط من التوزيع المتجمع النازل ايضا

- عمل توزيع تكراري متجمع نازل.
- حساب ترتيب الوسيط الذي يكون $\mathrm{F}_k = \frac{\sum_{i=1}^{m} f_i}{2}$.
- تثبيت ترتيب الوسيط بين التكرارات المتجمعة النازلة. ثم تحديد التكرار المتجمع النازل السابق لترتيب الوسيط (F_{k-1}^*)، والتكرار المتجمع النازل اللاحق لترتيب الوسيط (F_{k+1}^*).
- تحديد الفئة الوسيطية باعتبارها الفئة التي تقابل التكرار المتجمع النازل السابق لترتيب الوسيط (F_{k-1}^*).
- يحسب الوسيط وفق الصيغة الاتية: $Me = h_k + \frac{F_{k-1}^* - F_k}{f_k} \times L_k$

حيث ان:

h_k : الحد الادنى للفئة الوسيطية

F_{k-1}^* : التكرار المتجمع النازل السابق لترتيب الوسيط

f_k : تكرار الفئة الوسيطية

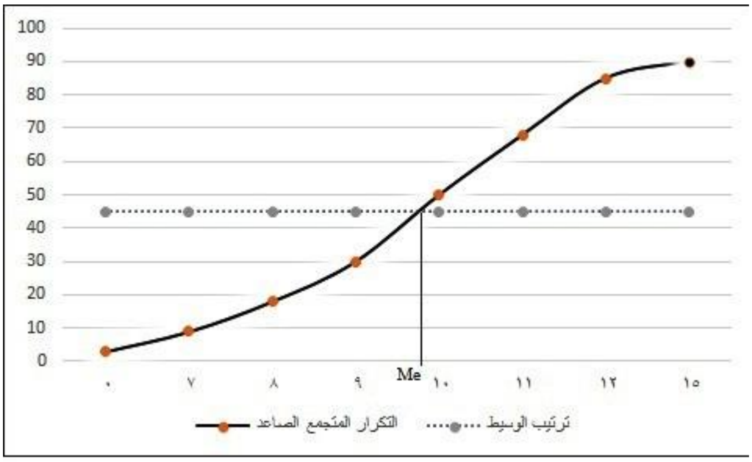
L_k : طول الفئة الوسيطية

Classes	f_i	الحدود الدنيا للفئات	F_i^*		
Less than 6	3	more than or = -∞	90		
6 —	6	more than or = 6	87		
7 —	9	more than or = 7	81		
8 —	12	more than or = 8	72		
9 —	**20**	more than or = **9**	**60**	← $\boldsymbol{F_{k-1}^*}$	← F_k=45
10 —	18	more than or = 10	40	← F_{k+1}^*	
11 —	17	more than or = 11	22		
12 or more	5	more than or = 12	5		
	90				

$$Me = h_k + \frac{F_{k-1}^* - F_k}{f_k} \times L_k = 9 + \frac{60 - 45}{20} \times 1 = 9.75$$

3- حساب الوسيط بيانياً للتوزيع التكراري

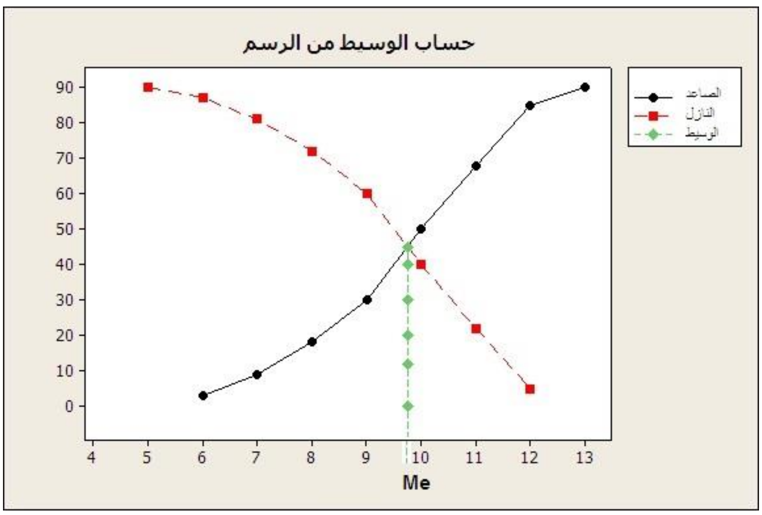
- يرسم منحنى التكرار المتجمع الصاعد على المحور العمودي، والحدود العليا للفئات ترسم على المحور الافقي.
- يحسب ترتيب الوسيط F_k، ويثبت على المحور العمودي، ومنه يرسم خط افقي موازٍ للمحور الافقي سيتقاطع مع منحنى التكرار المتجمع الصاعد.
- من نقطة التقاطع ينزل عمود على المحور الافقي وسيتقاطع معه في نقطة هي الوسيط.



ويمكن ايجاد الوسيط ايضا

- يرسم منحنى التكرار المتجمع الصاعد على المحور العمودي، وحدود الفئات ترسم على المحور الافقي.
- يرسم منحنى التكرار المتجمع النازل على المحور العمودي، وحدود الفئات ترسم على المحور الافقي.
- من نقطة تقاطع الرسمين ينزل عمود على المحور الافقي وسيتقاطع معه في نقطة هي الوسيط.

للمثال 4 يكون:



مزايا الوسيط	عيوب الوسيط
1- سهولة حسابة وبساطة فكرته. 2- يمكن تقديره عن طريق التخمين. 3- لايتاثر بالقيم الشاذة (الكبيرة جدا والصغيرة جدا) 4- يمكن حسابه للتوزيعات التكرارية المفتوحة. 5- يمكن تعيينه هندسياً. 6- يمكن حسابه للبيانات الوصفية.	1- لا يخضع للعمليات الجبرية. 2- يتأثر باخطاء المعاينة. 3- لا يعتمد في حسابه على جميع البيانات.

تمارين الوسيط (كل طالب يغير الفئات فقط)

1- التوزيع التكراري الاتي يمثل عدد النداءات الهاتفية العاجلة التي استقبلتها مستشفى خلال مدة شهر واحد. جد الوسيط لعدد النداءات

Classes	f_i
0 – 4	12
5 – 9	36
10 – 14	56
15 – 19	51
20 – 24	44

2- التوزيع التكراري ادناه لكميات الكهرباء المستهلكة (كيلو واط) من قبل 40 دار سكنية خلال شهر. جد الوسيط لكمية الكهرباء المستهلكة بطريقتين.

Classes	f_i
800 –	7
900 –	8
1000 –	15
1100 –	5
1200 –	3
1300 or more	2

3- جد الوسيط للتوزيع التكراري التالي حسابياً وبيانياً بثلاث طرق.

Classes	f_i
10 –	3
30 –	8
65 –	10
70 –	15
80 –	22
85 –	20
90 – 98	14

المنوال Mode

هو احد مقاييس النزعة المركزية، وهو القيمة التي تتكرر اكثر من غيرها، ويمكن ان يكون للمجتمع الاحصائي اكثر من منوال واحد اذا تكررت اكثر من قيمة واحدة بنفس العدد من المرات بحيث يكون هذا التكرار هو الاكبر، ويحدث ذلك في المجتمعات غير المتجانسة. كما ان هناك مجتمعات احصائية ليس لها منوالاً. ويرمز له بالرمز (Mo).

1- حساب المنوال للبيانات غير المبوبة

مثال 1: احسب المنوال لمجموعات البيانات الاتية:

A- 19 **18** 17 20 **18** 21 19 **18** 22

الحل: ان العدد 18 تكرر 3 مرات، اي اكثر من غيره لذا فانه المنوال (Mo = 18)

B- **10** 11 12 9 **10** 4 3 **5** 6 2 **5**

الحل: ان العددين 10و5 تكررا مرتين اي اكثر من غيرهما لذا فهنا **يوجد منوالين**

(Mo = 10) and (Mo = 5)

C- 12 15 16 11 10 9

الحل: لايوجد عدد تكرر اكثر من غيره لذا **لايوجد منوال** لهذه البيانات

2- حساب المنوال للبيانات المبوبة

أ- حساب المنوال للتوزيعات التكرارية المتقطعة

- اذا كان التوزيع التكراري منتظماً (فئاته متساوية الطول) فان الفئة المنوالية هي التي تقابل اكبر تكرار.
- اذا كان التوزيع التكراري غير منتظم (فئاته ليست متساوية الطول) فان الفئة المنوالية هي التي تقابل اكبر تكرار معدل (التكرار المعدل يساوي التكرار الاصلي مقسوما على طول الفئة المقابلة له اي ان $f_i^* = f_i/L_i$)
- يحسب المنوال باعتباره مركزا للفئة المنوالية

مثال 2: للتوزيع التكراري الاتي احسب المنوال.

Classes	f_i
60 – 74	2
75 – 89	6
90 – 104	14
105 – 119	10
120 – 134	8

الفئة المنوالية (90 – 104)، اكبر تكرار ← (14)

الحل : نلاحظ ان التوزيع منتظم لذا فان المنوال يكون:

$$Mo = \frac{90+104}{2} = \frac{194}{2} = 97$$

مثال 2: للتوزيع التكراري الاتي احسب المنوال.

Classes	f_i
5 – 9	2
10 – 14	6
15 – 24	10
25 – 34	22
35 – 49	27
50 – 59	11

الحل: نلاحظ ان الفئات غير متساوية بالطول لذا يجب تعديل التكرارات قبل حساب المنوال

Classes	f_i	L_i طول الفئة	f_i^* التكرار المعدل
5 – 9	2	5	2/5= 0.4
10 – 14	6	5	6/5= 1.2
15 – 24	10	10	10/10=1
25 – 34	22	10	22/10= 2.2
35 – 49	27	15	27/15= 1.8
50 – 59	11	10	11/10= 1.1

الفئة المنوالية (25 – 34)، اكبر تكرارمعدل ← (22/10= 2.2)

$$Mo = \frac{25+34}{2} = \frac{59}{2} = 29.5=30$$

تمرين: للتوزيعين التكرارين الاتيين احسب المنوال.

A

Classes	f_i
0 – 4	12
5 – 9	36
10 – 14	56
15 – 19	51
20 – 24	44

B

Classes	f_i
0 – 8	10
9 – 21	11
22 – 34	13
35 – 39	8
40 – 50	9

ب- حساب المنوال للتوزيعات التكرارية المستمرة

- اذا كان التوزيع التكراري منتظم (فئاته متساوية الطول) فان الفئة المنوالية هي التي تقابل اكبر تكرار.
- اذا كان التوزيع التكراري غير منتظم (فئاته ليست متساوية الطول) فان الفئة المنوالية هي التي تقابل اكبر تكرار معدل (التكرار المعدل يساوي التكرار الاصلي مقسوما على طول الفئة المقابلة له اي ان $f_i^* = f_i/L_i$)
- يحسب المنوال وفق الصيغة الاتية:

$$Mo = h_k + \frac{(f_k - f_{k-1})}{(f_k - f_{k-1}) + (f_k - f_{k+1})} \times L_k = h_k + \frac{d_1}{d_1 + d_2} \times L_k$$

حيث ان:

h_k : الحد الادنى للفئة المنوالية

L_k : طول الفئة المنوالية

f_k : تكرار الفئة المنوالية او التكرار المعدل للفئة المنوالية (اعلى تكرار او اعلى تكرار معدل)

f_{k-1} : التكرار السابق للفئة المنوالية او التكرار المعدل السابق للفئة المنوالية

f_{k+1} : التكرار اللاحق للفئة المنوالية او التكرار المعدل اللاحق للفئة المنوالية

مثال 4: التوزيع التكراري الاتي يبين توزيع اعمار عدد من المرضى الراقدين في احدى المستشفيات . المطلوب ايجاد العمر الشائع للمريض في هذه المجموعة.

Classes	f_i
Less than 20	2
20 –	8
30 –	16
40 –	17
50 –	23
60 – 70	14

الحل :ان الجدول التكراري اعلاه منتظم

	Classes	f_i	
	Less than 20	2	
	20 –	8	
	30 –	16	
	40 –	17	f_{k-1}
الفئة المنوالية	50 –	23	f_k اكبر تكرار
	60 – 70	14	f_{k+1}

$$Mo = h_k + \frac{d_1}{d_1 + d_2} \times L_k$$

$$d_1 = (f_k - f_{k-1}) = 23 - 17 = 6$$

$$d_2 = (f_k - f_{k+1}) = 23 - 14 = 9$$

$$Mo = 50 + \frac{6}{6+9} \times 10 = 50 + \frac{60}{15} = 54$$

مثال 5: للتوزيع التكراري ادناه. احسب المنوال.

Classes	f_i
5 –	2
10 –	6
15 –	10
25 –	22
35 –	27
50 – 60	11

الحل :ان الجدول التكراري اعلاه غير منتظم لذا سنستخدم التكرار المعدل

	Classes	f_i	L_i	f_i^*	
	5 –	2	5	2/5= 0.4	
	10 –	6	5	6/5= 1.2	
	15 –	10	10	10/10= 1	f_{k-1}
الفئة المنوالية	25 –	22	10	22/10= 2.2	f_k اكبر تكرار معدل
	35 –	27	15	27/15= 1.8	f_{k+1}
	50 – 60	11	10	11/10= 1.1	

$$Mo = h_k + \frac{d_1}{d_1 + d_2} \times L_k$$

$$d_1 = (f_k - f_{k-1}) = 2.2 - 1 = 1.2$$

$$d_2 = (f_k - f_{k+1}) = 2.2 - 1.8 = 0.4$$

$$Mo = 25 + \frac{1.2}{1.2 + 0.4} \times 10 = 25 + \frac{12}{1.6} = 32.5$$

المدرج التكراري:

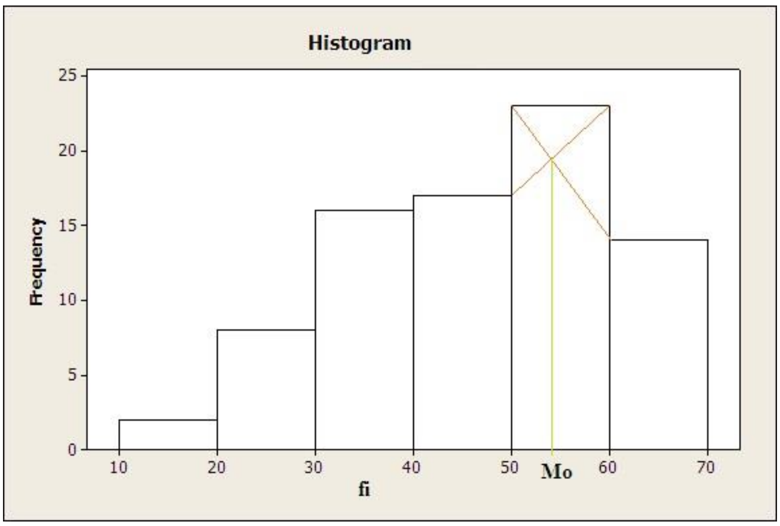
هو احد اساليب العرض البياني (الهندسي) لجداول التوزيع التكرارية، ويرسم بشكل مجموعة من المستطيلات، قاعدة كل منها تمثل طول الفئة في جدول التوزيع التكراري (ترسم على المحور الافقي)، وارتفاع كل منها هو قيمة التكرار المقابل لتلك الفئة (ترسم على المحور العمودي). وفي حالة عدم تساوي اطوال الفئات يجب تعديل التكرارات بقسمة كل تكرار f_i على طول الفئة المقابلة له L_i لنحصل على التكرار المعدل f_i^* اي ان $(f_i^* = f_i / L_i)$.

ايجاد المنوال من رسم المدرج التكراري

1- رسم المدرج التكراري وتحديد الفئة المنوالية التي تقابل اكبر تكرار.
2- رسم قطعة مستقيم من الحد الاعلى للفئة السابقة للفئة المنوالية الى الحد الاعلى للفئة المنوالية.
3- رسم قطعة مستقيم من الحد الادنى للفئة اللاحقة للفئة المنوالية الى الحد الادنى للفئة المنوالية.
4- رسم عمود من نقطة تقاطع قطعتي المستقيمين السابقين الى المحور الافقي. ونقطة تقاطع العمود مع المحور الافقي تمثل المنوال.

للمثال 4: يمكن ايجاد المنوال من الرسم كالاتي:

	Classes	f_i	
	Less than 20	2	
	20 –	8	
	30 –	16	
	40 –	17	f_{k-1}
الفئة المنوالية	50 –	23	اكبر تكرار f_k
	60 – 70	14	f_{k+1}



Mo = 54

ملاحظة: في حالة الجداول ذات الفئات غير المتساوية يتم رسم التكرار المعدل.

مزايا المنوال	عيوب المنوال
1- سهولة حسابة وفهمه. 2- يمكن تقديره عن طريق التخمين. 3- لايتاثر بالقيم الشاذة (الكبيرة جدا والصغيرة جدا) 4- يمكن حسابه للتوزيعات التكرارية المفتوحة. 5- يمكن تعيينه هندسياً. 6- يمكن حسابه للبيانات الوصفية.	1- لا يخضع للعمليات الجبرية. 2- يتأثر باخطاء المعاينة. 3- لا يعتمد في حسابه على جميع البيانات.

تمارين المنوال

1- الاتي توزيع الدخل الشهري لمجموعة من الاسر في منطقة معينة. جد الدخل الشائع في هذه المنطقة.

Classes	f_i
Less than 100	6
100 –	13
150 –	28
200 –	52
250 –	24
300 or more	12

2- كان توزيع اعمار المصابين بمرض البلهارزيا في احدى المناطق الريفية كما مبين في ادناه. جد العمر الشائع للمصابين بهذا المرض.

Classes	f_i
5 –	20
10 –	28
15 –	26
20 –	43
25 –	32
30 –	18
35 – 40	41

3- لجدول التوزيع التكراري التالي. جد قيمة a , b اذا علمت ان a اكبر من b بوحدة واحدة وان Mo=24 .

Classes	f_i
0 –	14
10 –	a
20 –	27
30 –	b
40 – 50	15

4- الاتي توزيع تكراري لعدد الموظفين في دائرة ما حسب فئات الراتب، المطلوب ايجاد متوسط الراتب باستخدام المنوال.

Classes	f_i
18 –	14
28 –	20
36 –	36
50 –	45
70 –	26
100 – 130	12

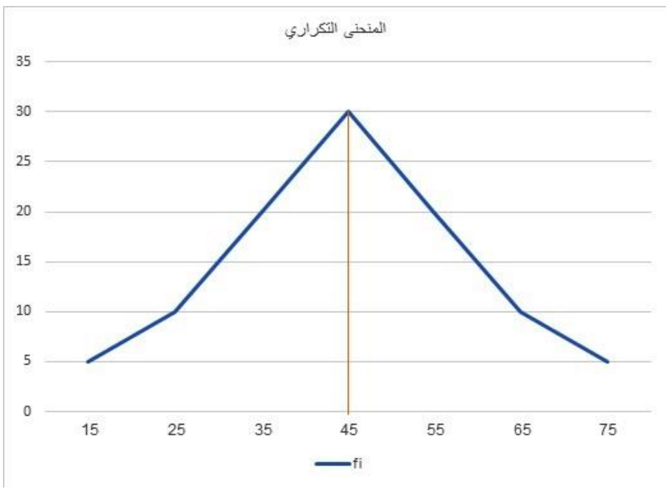
العلاقة بين المتوسطات

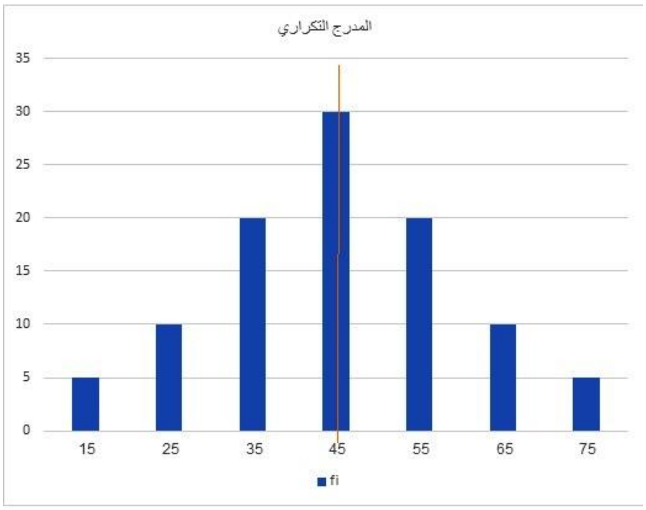
التوزيع التكراري المتماثل:

هو التوزيع التكراري الذي تكون بياناته حول نقطة معينة متشابهة، مثل الجدول الاتي:

Classes	f_i
10 –	5
20 –	10
30 –	20
40 –	30
50 –	20
60 –	10
70 – 80	5

وعند رسم المنحنى التكراري والمدرج التكراري له يظهر بالشكلين الاتيين





وكما يظهر فان الجزء الايمن من التوزيع يماثل الجزء الايسر من التوزيع (يقسم التوزيع بالخط الاحمر الى جزئين متماثلين). وفي هذا التوزيع المتماثل يتساوى الوسط الحسابي مع المنوال والوسيط.

العلاقة بين الوسط الحسابي والوسيط والمنوال

لوحظ ان العلاقة بين الوسط الحسابي والوسيط والمنوال، اذا كان التوزيع التكراري قريب من حالة التماثل تكون بالصيغة الاتية:

$$\bar{x} - Me = \frac{1}{3}(\bar{x} - Mo)$$

ان هذه العلاقة مهمة جدا اثناء التطبيق حيث يتعذر مثلا حساب الوسط الحسابي للتوزيعات التكرارية المفتوحة، في حين يمكن حساب الوسيط والمنوال لها. لذا يمكن استخدام هذه العلاقة لايجاد قيمة الوسط الحسابي، بعد حساب الوسيط والمنوال. كما يمكن حساب اي متوسط من المتوسطات الثلاثة اذا علم المتوسطين الاخرين.

ملاحظة: اذا كان التوزيع التكراري متماثلا فان المتوسطات الثلاثة تكون متساوية. اي ان:

$$\bar{x} = Me = Mo$$

وكلما ابتعدت المتوسطات عن بعضها البعض كلما كان التوزيع ابعد عن التماثل.

مثال 5: تعذر في احد التوزيعات التكرارية القريبة من التماثل الحصول على قيمة الوسط الحسابي.، في حين امكن الحصول على قيمتي الوسيط والمنوال وكانتا:

$Mo = 53 \quad , \quad Me = 52$. جد قيمة الوسط الحسابي.

الحل: العلاقة بين المتوسطات الثلاثة هي

$$\bar{x} - Me = \frac{1}{3}(\bar{x} - Mo)$$

$$\bar{x} - 52 = \frac{1}{3}(\bar{x} - 53)$$

$$3(\bar{x} - 52) = \bar{x} - 53$$

$$3\bar{x} - 156 = \bar{x} - 53$$

$$3\bar{x} - \bar{x} = 156 - 53$$

$$2\bar{x} = 103$$

$$\bar{x} = \frac{103}{2} = 51.5$$

تمارين العلاقة بين المتوسطات

تمرين1: اذا كان الوسط الحسابي يساوي 73 ، والمنوال يساوي 69 ، احسب الوسيط لهذا التوزيع التكراري مفترضا قربه من التماثل.

تمرين2: اذا كان الوسط الحسابي يساوي 60 ، والوسيط يساوي 50 ، احسب المنوال لهذا التوزيع التكراري. وهل ان التوزيع التكراري متماثل ؟

تمرين 3: اذا كان التوزيع التكراري التالي يقترب من التماثل، فاحسب الوسط الحسابي له.

Classes	f_i
60 –	6
80 –	14
100 –	20
120 –	30
140 –	22
160 –	12
180 or more	8