وزارة التعليم العالي والبحث العلمي
جامعة الفرات الاوسط التقنية
كلية البوليتكنك / القادسية
قسم تقنيات ادارة الاعمال

محاضرات مادة مبادئ الإحصاء

لطلبة قسم تقنيات ادارة الاعمال

السنة الدراسية الأولى – كلية البوليتكنك القادسية

م. فاطمة عبد الرزاق عبود
م.م.ضحى سالم وحيد

2025- 2026

الاسبوع الاول

تعريف علم الأحصاء

تعريف علم الأحصاء(statistics Science)

عرف علم الأحصاء بأنه (هو العلم الذي يهتم بالطرق العلمية لجمع وتنظيم وتلخيص البيانات وعرضها وتحليلها بأساليب علمية للتوصل الى الاستنتاجات لاتخاذ القرارات المناسبة) .

ازدادت اهمية علم الاحصاء وخاصة في السنوات الاخيرة ، حيث اخذ يحتل المكانة البارزة بين العلوم الاخرى ، ولاهمية التعرف على الطرائق العلمية او الادوات الفنية التي يمكن الاستعانة بها من اجل جمع البيانات الاحصائية ومن ثم تبويبها وعرضها وتحليلها وبالتالي تفسير النتائج المتعلقة بها لابد من دراسة مبسطة للمباديء الاساسية للاحصاء الوصفي والحياتي.

علاقة علم الاحصاء بالعلوم الاخرى

ماكان علم الاحصاء ليأخذ اهميته الكبيرة دون الحاجة اليه كأداة تحليلية في مختلف العلوم، بالاضافة لارتباطه الوثيق بالرياضيات كون أسسه وأصول قوانينه رياضية ، فأن الحاجة اليه في العلوم الاقتصادية هي حاجة حيوية، اذ من المستحيل ان نتعرف على سلوك ظاهرة اقتصادية ما في الانتاج او الاستهلاك او ..... ا وان نؤثر في سلوك هذه الظاهرة او تلك بما يلائم ا وان نقوم بأي عملية تخطيط ا وان نتخذ اي قرار اقتصادي دون ان يساعدنا علم الاحصاء في شرح وبيان الواقع الاقتصادي الحالي والتنبؤ بما ستكون عليه الظاهرة الاقتصادية في المستقبل ، كذلك العلاقة بين علم الاحصاء وعلم السكان هي علاقة وطيدة ايضا حيث بواسطته يمكن التعبير عن مختلف المؤشرات السكانية كنسب ومعدلات الولادات والوفيات والزواج والطلاق والهجرة وكذلك التوزيعات السكانية والتعدادات السكانية ، وهناك علاقة وثيقة لهذا العلم مع العلوم الانسانية حيث ان اغلب دراسات السلوك الانساني وفي مختلف النواحي الاجتماعية والنفسية تحتاج للبحث الميداني .

اما علاقة علم الاحصاء بعلوم الطب والصيدلة والهندسة وعلم الحياة والوراثة والزراعة وغيرها فهي علاقة وثيقة .

وبأختصار طالما ان هناك حاجة لدراسات ميدانية فهناك حاجة لامناص منها للاحصاء.

الطريقة الاحصائية

هي الطريقة العلمية الخاصة بمعالجة النواحي الخاضعة للتحليل الكمي (الارقام) وتتألف من المراحل الرئيسية التالية:-

1. جمع البيانات Collecting of data
2. مراجعة البيانات Verifying data
3. تصنيف وجدولة البيانات Classification and Tabulation
4. التمثيل البياني للبيانات Craph presentation
5. حساب المؤشرات الاحصائية
6. التحليل والتفسير والتنبؤ

اولا : جمع البيانات(Collection of Data)

وهي مرحلة مهمة من مراحل الطريقة الاحصائية حيث ان دقة النتائج تتوقف على دقة البيانات او المعلومات التي يتم جمعها عن الظاهرة موضوعة البحث ،وقبل البدء بجمع البيانات الاحصائية المطلوبة لدراسة المشكلة المبحوثة ، يجب دراسة موضوع البحث لمعرفة نوعية البيانات المراد جمعها وعلى نوعية الاسئلة التي ستوجه وحصر المصادر التي يمكن استقصاء البيانات منها ومدى توفر الامكانات للقيام بمثل هذه الدراسة ، وبعد توفر هذه العناصر يمكن تخمين مدى نجاح هذه الدراسة والبحث بشرط اجراء دراسة تمهيدية وعلى نطاق ضيق لدراسة مدى نجاح هذه العملية قبل البدء بالدراسة .

المصادر التي يمكن الحصول على البيانات المطلوبة منها ، والتي تقسم عموما الى نوعين من المصادر هما:

أ. **المصادر التاريخية او المكتبية :**وهي البيانات المدونة او المحفوظة والتي جمعت النشرات والاصدارات التي تصدرها الدوائر والمنظمات المحلية او العالمية ، حيث يدخل ضمن هذه المصادر المؤلفات والبحوث التي قام بانجازها المختصون ، فمثلا اذا اردنا حساب معدل الوفيات الخام السنوي فاننا نرجع الى بيانات التسجيل الحيوي او لنتائج التعداد السكاني في حالة وقوع عملية التعداد خلال تلك السنة.

ب.المصادر الميدانية :ويتم فيها جمع البيانات المطلوبة مباشرة من مجتمع الدراسة او البحث ، وتتعدد الطرائق والاساليب التي يتم جمع البيانات بموجب هذا النوع من المصادر ، فقد تستخدم طريقة المقابلة الشخصية او احد وسائل البريد المتوفرة كالرسائل او الهاتف او بوساطة وسائل الاعلام المتاحة والتي يمكن ان تؤمن الحصول على البيانات المطلوبة، ويوجد ايضا مايعرف بطريقة او اسلوب المشاهدة وبالاخص في حالة التجارب العلمية ذات الطابع التطبيقي ، وعموما فأن الحصول على البيانات في اغلب الاحيان يكون من خلال استمارة تعد لهذا الغرض تعرف باستمارة او صحيفة الاستبيان .

اساليب جمع البيانات من المصادر الميدانية: هناك اسلوبان

1. **اسلوب الحصر الشامل** :-وهو ان يتم جمع البيانات من جميع وحدات المجتمع الاحصائي مثل التعداد العام للسكان ويعتبر من افضل اساليب جمع البيانات.
2. **اسلوب العينة**:- وهو اختيار جزء من المجتمع الاحصائي بطريقة معينة تضمن تمثيل المجتمع الاصلي بجميع وحداته تمثيلا صادقا.

.

وبعد تحديد الاسلوب الذي سيعتمد في عملية جمع البيانات المطلوبة عن الظاهرة المبحوثة ، تأتي مرحلة تجهيز البيانات ومراجعة صحائف الاستبيان مكتبيا والفرز والتبويب حيث تحول البيانات الخام التي تم جمعها الى جداول احصائية بحيث تصبح جاهزة لاغراض الوصف والاستدلال الاحصائي .

الاستبيانات ، شروطها ، اجرائها

استمارة (صحيفة)الاستبيان :-هي عبارة عن ورقة مصممة لاستيعاب مجموعة من البيانات المراد جمعها وذلك باحتوائها على مجموعة من الاسئلة.

- صياغة مقدمة تعنون بعبارة احترام توجه للمجيب مع الاشارة الى اهمية المعلومات المطلوبة وعن كيفية الاجابة عن الاسئلة الموضوعة مع مراعاة جانب الدقة والامانة وبما يعبر ذلك عن واقع الحال فعلا ، واخيرا توجيه كلمة شكر وأمل بافادة المجيب .

مكونات الاستمارة الاحصائية:

1.الهدف من الاستمارة.
2.اسئلة ثانوية تتعلق بصفات المبحوث.
3.اسئلة رئيسية تتعلق بموضوع البحث.
4.تعليمات على كيفية الاجابةعلى الاسئلة الغامضة في الاستمارة.

النقاط الاساسية الواجب ملاحظتها عند صياغة استمارة الاستبيان

1.صياغة مقدمة تعنون بعبارة احترام توجه للمجيب مع الاشارة الى اهمية المعلومات المطلوبة وعن كيفية الاجابة عن الاسئلة الموضوعة مع مراعاة جانب الدقة والامانة وبما يعبر ذلك عن واقع الحال فعلا، واخيرا توجيه كلمة شكر وامل بافادة المجيب.
2.تكون الاسئلة واضحة وبسيطة وشاملة لموضوع البحث وقليلة.
3.تكون الاجابة محددة قدر الامكان لاتقبل التأويل (كالاجابة بنعم او لا)

4.الترتيب المنطقي للاسئلة حتى لايتشتت ذهن المبحوث.
5.تكرار الاسئلة المهمة باكثر من صيغة.
6.الابتعاد عن كل ما من شأنه ان يخلق حالة الاحراج او الارباك او الكذب عند المعنيين بالاجابة كذلك خلق الثقة لديهم بان البيانات ليس لها اي علاقة الا لاغراض البحث والدراسة مع الحفاظ على سرية المعلومات المدونة.

انواع اسئلة الاستبيان:-

أ.الاسئلة الثنائية:وهي الاسئلة التي تحتمل احد امرين فقط مثل اسئلة الصواب والخطأ وتمتاز هذه الاسئلة بأنها سهلة الاعداد والاجابة والتقييم لكن يعاب عليها بأنها تسهل الامور اكثر مما يجب.

ب.اسئلة الاختيار المتعدد : وهذه افضل من الاسئلة الثنائية اذ انها تعطي عدة بدائل ممكنة لكنه يجب ان يراعى فيها ان تغطي جميع الامكانات.

ج. الاسئلة المفتوحة : وهي الاسئلة التي تترك للمجيب حرية التعبير عن رأيه دون قيد لكن من مساويء هذه الاسئلة يصعب التقييم اثناء عملية تحليلها(تحليل الاجابات).

ومن الاجراءات المتبعة لتطوير استمارة الاستبيان او زيادة فاعليتها بأحد (او اكثر) الوسائل الاتية:-

- اجراء تجربة تقويمية للاستمارة من خلال توزيعها على عينة صغيرة من افراد المجتمع وملاحظة اجاباتهم وبالتالي الاستفادة في اعادة صياغة الجوانب التي قد تكون اقل فاعلية قبل اقرارها بصورة نهائية وتوزيعها على المجيبين .
- عرض الاستمارة على خبير او مجموعة خبراء لمراجعتها او الاستفادة من ارائهم بشأن تطويرها.
- تقديم استمارة الاستبيان من خلال الاطلاع على بعض الاستمارات المعدة لنفس الغرض سابقا.

مراجعة البيانات: تأتي هذه الخطوة مباشرة بعد عملية جمع البيانات وتعتبر من العمليات المهمة لنجاح البحث لان البيانات حينما تجمع تعتبر اولية ولايمكن تبويبها قبل مراجعة الاجوبة المدرجة فيها ، فعملية المراجعة تكشف لنا الاخطاء او النقصان في تلك الاستمارات وعندها يحاول القائمون بالعملية الاحصائية من اكمالها او تصحيحها وان تعذر عليهم ذلك يجب ترك تلك الاستمارات واستبعادها عن البحث كي لاتظهر النتائج بعيدة عن الصحة او تظهر الوضع على غير حقيقته

عرض البيانات الاحصائية (Presentation & Tabulation of Statistical Data)

بعد اجراء عملية المراجعة تبدأ عملية تصنيف البيانات فتقسم الى مجموعات متجانسة ، أي ان مفردات كل مجموعة تشترك بصفة معينة ، فقد يكون التصنيف على اساس الجنس او العمر او الطول ....الخ
حيث ان هذه الخطوة من شأنها ان تساعد او تمكن من الاستفادة منها الى اقصى حد ممكن، فقد يصعب على الجهة المستفيدة من استيعاب البيانات خاصة اذا كانت كثيرة العدد .
بعد اتمام عملية تصنيف البيانات تبدأ عملية التبويب ، وهي عملية تفريغ البيانات ووضعها في جداول بقصد ابرازها بشكل اوضح وابسط ومختصر، ويختلف اسلوب التبويب وفقا لاختلاف طبيعة البيانات وهناك عدة انواع للتبويب منها :-

تبويب البيانات (Tabulation of Data)

1.**التبويب الوصفي (النوعي)** :- ونعني به عملية تقسيم البيانات الى مجموعات على اساس ان مفردات كل مجموعة تشترك بصفة معينة لها علاقة بموضوع البحث فمثلا يمكن تقسيم الاشخاص الى مرضى واصحاء.

2.التبويب الكمي :-هو توزيع البيانات الى مجموعات معينة عندما يكون بالامكان قياس بعض الصفات وتضم هذه المجموعات مدى معين من قيم الظاهرة، فمثلا تبويب مجموعة من المرضى حسب الاعمار.

3.التبويب الزمني :هو عبارة عن تقسيم البيانات الى مجموعات تخص وحدة زمنية معينة (يوم ،اسبوع ،شهر ، سنة) مثل كميات الانتاج

4.التبويب الجغرافي :-هو توزيع البيانات الى مجموعات موزعة توزيعا جغرافيا.مثلا توزيع تصدير النفط الخام حسب الدول ويستخدم للمقارنة بينها.

طرق التبويب: هناك طريقتان للتبويب السابق:-

1. **الطريقة اليدوية**:تستخدم هذه الطريقة اذا كان المجتمع الاحصائي(البيانات) محدودا (صغيرا)
2. **الطريقة الالية** : تستخدم هذه الطريقة اذاكان المجتمع الاحصائي كبيرا وواسعا فان تبويبها يتطلب جهدا ووقتا طويلا لذا يتم اللجوء الى انظمة الحاسوب الجاهزة التي تتميز بالسرعة والدقة.

عرض البيانات الاحصائية (Presentation & Tabulation of Statistical Data)

ان كبر حجم البيانات وتعدد ا رقامها يشكل صعوبة كبيرة تحول دون فهمها والتوصل الى نتائج مهمة عنها لذلك تصنف البيانات بتبويبها وفق نظام معين في مجموعات متجانسة بهدف تلخيصها ووضعها في حجم مناسب من اجل فهمها وتحليلها .
ان عملية تصنيف البيانات يجب ان يعتمدعلى نظام يتم بموجبه تصنيف المعلومات وتبويبها وحتى يكون هذا النظام المعتمد فى عملية التصنيف نظام ناجح يجب ان يتمتع بخواص هى:-
أ.**عدم التداخل:** يجب ان لاتتداخل المجموعات التي تقسم اليها البيانات مع بعضها البعض بمعنى انه يجب ان لايوجد الا مكان واحد للمفردة الواحدة فى النظام.
ب.**الشمولية:** بمعنى يجب ان تجد كل مفردة من المفردات مكانا لها ضمن احدى مجموعات النظام.
ج.**الاستمرارية في تطبيق الاساس المستخدم في النظام:** بمنى انه اذا اتبع اساس معين كالاساس الزمني فيجب الاستمرار في تطبيق هذا الاساس في في تصنيف كل مفردات المجتمع الاحصائي.

1.العرض الجدولي للبيانات (التبويب): حيث توضع البيانات في جداول تتحدد درجتها بعدد المتغيرات الداخلة في بنائها بعد الانتهاء من مرحلة جمع البيانات التي نرغب في دراسة الظاهرة من خلالها حيث يكون من الصعوبة او من المستحيل في بعض الاحيان استيعابها خاصة عندما يكون عددها كبيرا جدا، وتعتبر هذه الطريقة من العرض احدى الخطوات الاساسية المهمة واللازمة لتلخيص البيانات المبحوثة ، كما تصنف الجداول من جانب اخر الى نوعين :

وذلك لعرض البيانات بطريقة مشوقة وسهلة وواضحة وذلك بتبويبها او تفريغها حسب تصنيف معين يتوقف على نوع المتغير او المتغيرات وعددها والغرض من الدراسة ،وتقسم الجداول الى نوعين :الوصفية والتي يكون متغيراتها نوعية ،والكمية والتي يكون متغيراتها كمي .

البيانات غير المبوبة :- هي البيانات الاولية الخام التي جمعت ولم تبوب.

1.عرض البيانات غير المبوبة :

1.العرض الجدولي(الجداول الوصفية) تقسيم البيانات الى مجموعات تشترك بصفة معينة ، **و يمكن عرض بياناته في شكل جدول تكراري بسيط، وهو جدول يتكون من عمودين، أحدهما به مستويات (مجموعات) المتغير، والثاني به عدد المفردات (التكرارات) لكل مستوى (مجموعة).**

فمثلا يقسم المتغير النوعي الخاص بالحالة التعليمية الى ثلاث مستويات نوعية هي (امي ، يقرأ ويكتب ، متعلم) او شدة الاصابة بمرض معين (بسيطة ، متوسطة ، شديدة ، مزمنة).

ملاحظات على الجدول عند تكوين جدول ما لعرض البيانات، يجب مراعاة الآتي:

1- كتابة رقم للجدول.
2- كتابة عنوان للجدول.
3- لكل عمود من أعمدة الجدول عنوان يدل على محتواه.

مثال(1) :البيانات التالية تمثل دراسة عن اسباب اختيار الطلبة لقسم تقنيات ادارة المواد في المعهد التقني الديوانية فاذا كانت اسباب الاختيار هي :-

1.الرغبة 2.لم تسمح معدلاتهم في الاعدادية بأفضل من ذلك 3.اسباب اخرى وكانت الاجابات كالتالي :-
(الرغبة ، الرغبة ، المعدل ، اخرى ، المعدل ، اخرى ، المعدل ، المعدل ، الرغبة ، المعدل)

الحل :

جدول رقم (1)
توزيع عينة حجمها 10 حسب اسباب اختيار القسم العلمي

السبب	عدد الطلبة	عدد الطلبة
الرغبة	///	3
معدل الاعدادية	////	5
اخرى	//	2
المجموع		10

مثال (2):البيانات التالية تمثل المستوى التعليمي لمجموعة من العاملين في احدى المصانع وكانت كالاتي:(اعدادية ، اعدادية ، يقرأ ويكتب ، امي ، ابتدائية ،يقرأ ويكتب ، اعدادية ، جامعي ، عليا ، عليا، ابتدائية ، جامعي ، جامعي ،امي ، يقرأ ويكتب)

2.العرض البياني للبيانات *Graphic of Data Presentation*:بالرغم من ان طبيعة المسالة الاحصائية هي التي تحدد في اغلب الاحيان استخدام احد الاشكال البيانية لغرض تمثيل بياناتها ، فان اختيار شكل بياني او اكثر ليكون اكثر ملائمة لنوع البيانات المبحوثة يتحدد في ضوء بعض المحددات او المؤشرات ، وتهيء طريقة العرض البياني مجالا مبسطا تمكن المستفيد من خلالها استيعاب محتويات تلك الجداول وبأقل جهد ممكن.

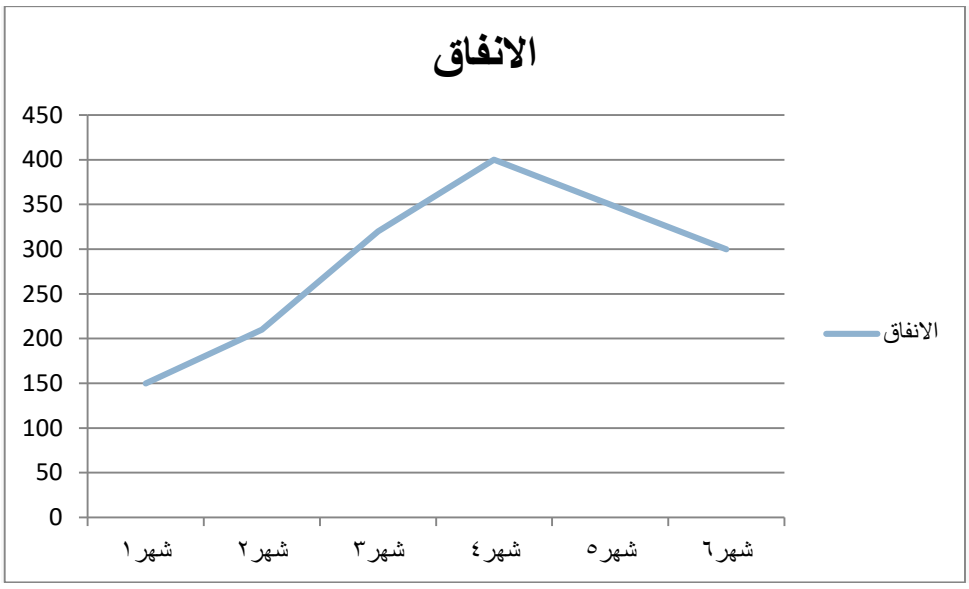
كما انها من الطرائق الاحصائية الوصفية التي تمكن بعض المستفيدين من الذين يجدون صعوبة في ادراك واستيعاب اعداد قيم المتغيرات التي تتضمنها الجداول الاحصائية لذا توضح هذه الجداول بصيغة بيانية او رسوم تصويرية ،واهم هذه الاشكال:

1.الخط البيانيLine Chart

تستخدم هذه الطريقة لتوضيح التغيرات التي تحدث نتيجة لسير ظاهرة ما خلال فترة زمنية معينة فمثلا تغير درجات الحرارة خلال ايام الاسبوع ، حيث يكون المحور الافقي ممثلا لعامل الزمن وهو المتغير المستقل والمحور العمودي ممثلا لقيم الظاهرة (درجات الحرارة)وهو المتغير التابع.

مثال(1) :الجدول التالي يتضمن الانفاق الاستهلاكي لاحدى العوائل خلال ستة اشهر.والمطلوب عرضها بيانيا.

الانفاق	150	210	320	400	350	300
الشهر	1	2	3	4	5	6



مثال(2) : البيانات التالية توضح اعداد المرضى الخارجين من مستشفى الاطفال والولادة خلال اسبوع واحد ، المطلوب عرضها بيانيا حسب الطريقة التي تناسبها.

ايام الاسبوع	السبت	الاحد	الاثنين	الثلاثاء	الاربعاء	الخميس	الجمعة
عدد المرضى	20	15	48	10	17	28	10

2.طريقة المستطيلات او الاعمدة البيانية Bar Charts

تستخدم هذه الطريقة في حالة دراسة صفات ظاهرتين (او اكثر) ايضا.وهي عبارة عن مستطيلات ترسم بعرض واحد ولكن بأطوال مختلفة بحيث تدل على قيم الظاهرة.

مثال(1) : من بيانات الجدول الخاص باعداد طلبة المعهد التقني موزعين حسب الاقسام العلمية، المطلوب وصف بيانات الجدول بيانيا.

الاقسام العلمية	اعداد الطلبة
تقنيات ادارة المواد	65
تقنيات التأمين	60
تقنيات المحاسبة	84
تقنيات الادارة الصحية	155
تقنيات الادلة الجنائية	110
تقنيات الميكانيك	35
تقنيات الاشعة والسونار	160
تقنيات التمريض	218
تقنيات التأهيل الطبي	120
تقنيات صناعة الاطراف	97
تقنيات صحة المجتمع	190

مثال (2): البيانات التالية تمثل اعداد المرضى الراقدين في مستشفى الديوانية التعليمي مصنفين حسب نوع المرض

نوع المرض	A	B	C	D	E
اعداد المرضى	15	20	8	19	31

المطلوب عرضها بيانيا بالطريقة التي تناسبها.

العرض الجدولي للبيانات المبوبة (التوزيعات التكرارية)

***الجداول الكمية Quantitative Tables* :-** اذا كانت البيانات المبحوثة تعبر عن ظاهرة لمتغير او اكثر يمكن قياسها كميا ، فأن الجدول الذي يبنى على هذا النوع من البيانات يسمى بجدول التوزيع التكراري (Frequency Table) وهو من الجداول الاحصائية المهمة حيث نتمكن بوساطتها من تنظيم البيانات الكثيرة العدد وذلك بتوبيبها في مجموعات متساوية او غير متساوية وذلك بناءا على بعض المؤشرات بحيث لاتخسر البيانات المبوبة من اهميتها الا الشيء اليسير او ربما لاتخسر شيئا.

ويعرف جدول التوزيع التكراري(Frequency Distribution) :بانه عبارة عن تلخيص وترتيب البيانات الاحصائية كبيرة العدد في مجموعات تسمى الفئات وحصر عدد المفردات التي تنتمي الى كل فئة وذلك لتسهيل قراءتها ومعالجتها رياضيا .

مصطلحات وقوانين

الفئة (Class) : عبارة عن مدى محدود من قيم الظاهرة ، لكل فئة بداية يسمى الحد الادنى للفئة (Lower class limit) ولها نهاية تسمى الحد الاعلى للفئة (Upper class limit) ويرمز لها.

طول الفئة (CLASS SIZE (WIDTH : هو المدى المحصور بين حدي الفئة ، اي
طول الفئة = الحد الاعلى للفئة – الحد الادنى للفئة

$$w = \text{Upper Class Limit} - \text{Lower Class Limit}$$

مركز الفئة Class Mid Point :القيمة الواقعة عند منتصف الفئة ، اي القيمة التي تتوسط المسافة بين الحد الادنى والحد الاعلى للفئة

$$mi = \frac{u.c.l + l.c.l}{2}$$

تكرار الفئة Frequency Of Class : عدد المفردات التي تقع ضمن مدى الفئة ودائما يكون مجموع التكرارات مساويا لحجم العينة اي مساويا لعدد قيم الظاهرة.

المدى Range : الفرق بين اكبر قيمة big value واصغر قيمة small value من قيم الظاهرة المدروسة .

$$R = B.V - S.V$$

عدد الفئات Number Of Classes : يمكن تحديد عدد الفئات بالصيغة التالية:

$$n.c = 2.5\sqrt[4]{n}$$

يجب ان يكون عدد الفئات يلائم ظروف البيانات بحيث لانحصل على فئات بعضها خال من التكرارات او ان نحصل على توزيع تتركز التكرارات في فئة دون اخرى او يكون عددها قليل فتضيع بذلك معالم التوزيع وان لايكون عددها كثيرا فيضيع القصد من عملية التلخيص او الاختصار.

الجدول التكراري المغلق Close frequency Table :هو الجدول الذي يكون فيه الحد الادنى للفئة الاولى والحد الاعلى للفئة الاخيرة معلومين.

الجدول التكراري المفتوح Open Frequency Table : هو الجدول الذي تكون في احدى الحالات التالية:-

1. جدول مفتوح من الاسفل اذا كان الحد الادنى للفئة الاولى غير معلوم.
2. جدول مفتوح من الاعلى اذا كان الحد الاعلى للفئة الاخيرة غير معلوم.
3. جدول مفتوح من الطرفين اذا كان الحد الادنى للفئة الاولى والحد الاعلى للفئة الاخيرة غير معلومين.

جدول توزيع تكراري منتظم Regular frequency distribution table: هو الجدول الذي تكون فئاته متساوية الاطوال وذلك تسهيلا للعمليات الحسابية.

جدول توزيع تكراري غير منتظم Irregular frequency distribution table : هو الجدول الذي تكون فئاته غير متساوية في الطول فقد تكون البيانات المبحوثة مفصلة في جزء ومجملة في جزء اخر او قد نكون راغبين في دراسة احدى او اكثر الفئات ذات طول معين وذلك للاهمية او لاي سبب اخر وبغض النظر عن اطوال الفئات الاخرى .

*والجداول التكرارية تكون بنوعين:

1.جدول تكراري بسيط
2.جدول تكراري مزدوج او مركب

1. ***جدول التوزيع التكراري البسيط Simple frequency distribution table*** :هو تبوب البيانات على شكل فئات تكرارية مع تحديد عدد المشاهدات لكل من هذه الفئات ويعرف عدد المشاهدات هنا بالتكرار اذا كان هناك متغير واحد.

مثال (1) :البيانات الاتية تمثل اوزان مجموعة من التلاميذ في احدى المدارس

42	28	27	25	26	27	34	31	31	48
38	45	30	32	39	49	38	44	43	35

المطلوب عرض هذه البيانات في جدول توزيع تكراري.

الحل :

$R = B.V - S.V$

$= 49 - 25$

$= 24$

$n.c = 2.5\sqrt[4]{n}$

$= 2.5\sqrt[4]{20}$

$= (2.5)(2.1)$

$= 5.25 \cong 6$

$w = \frac{R+1}{n.c}$

$= \frac{24+1}{6}$

$= 4.1 \cong 4$

الدرجة			العلامات الإحصائية	fi (التكرارات)
©فئات	©فئات	©فئات		
25 to les than 29	25 – 29	25-	////	5
29 to les than 33	29 – 33	29-	////	4
33 to les than 37	33 – 37	33-	//	2
37 to les than 41	37 – 41	37-	///	3
41 to les than 45	41 – 45	41-	///	3
45 to les than 49	45 – 49	45-49	///	3
Total			**20**	**20**

مثال (2) : البيانات التالية تمثل درجات مجموعة من الطلبة في مادة الاحصاء.

92	73	62	66	97	65	65	87	98	83	57	63	75	51
66	79	75	90	79	88	75	93	74	88	61	69	65	61

المطلوب عرض البيانات في جدول توزيع تكراري.

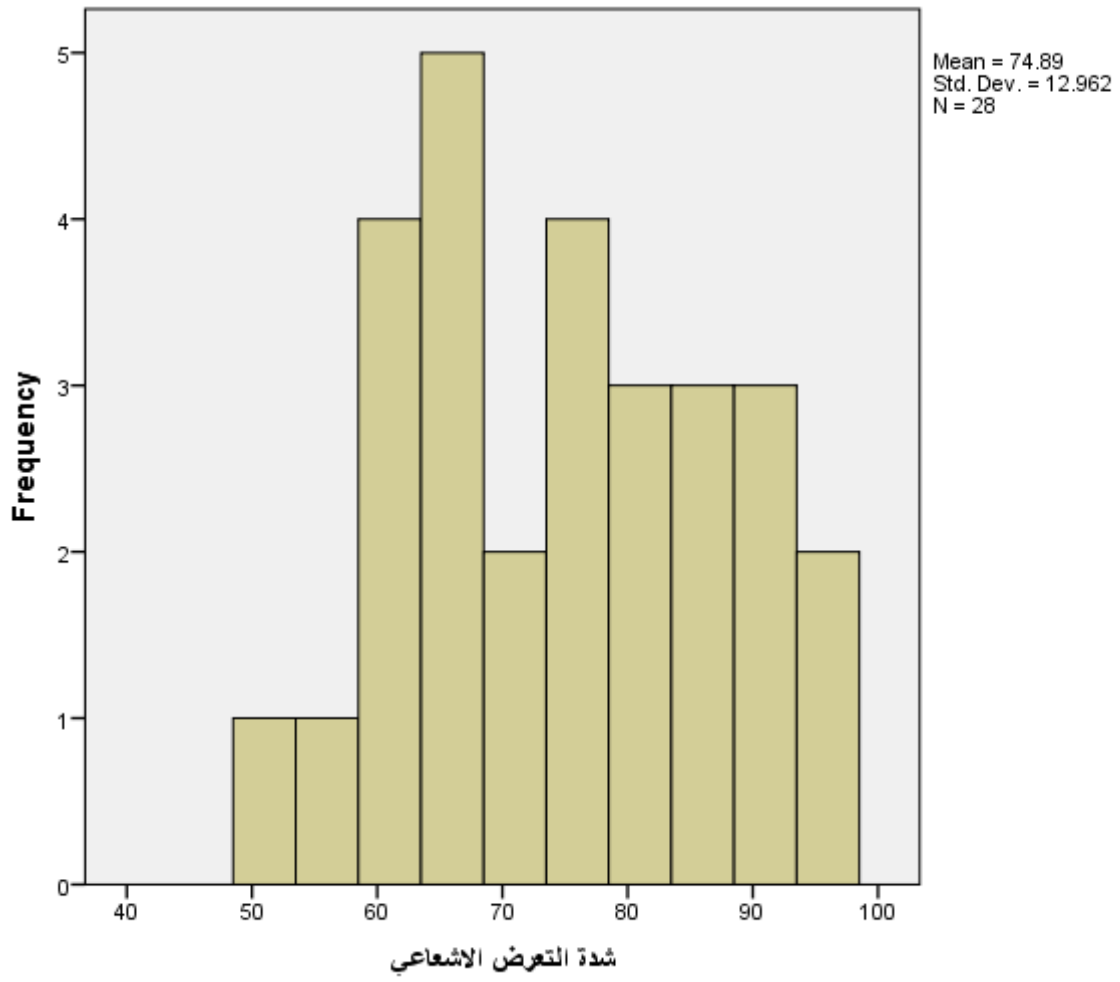
العرض البياني للتوزيعات التكرارية(البيانات المبوبة)

1. المدرج التكراري Frequency Histogram

هو عبارة عن شكل بياني يتألف من أعمدة رأسية (مستطيلات) تتناسب مساحتها مع تكرارات الفئات إذا كانت البيانات التي لدينا تمثل متغيراً كمياً متصلاً ، وفي حالة الفئات المتساوية بالطول فأن الفئات التي تتمثل على المحور الافقي تتساوى فيها قواعد المستطيلات المتجاورة المرسومة على ذلك المحور في حين تمثل ارتفاعها على المحور العمودي يتناسب وتكرار كل منها .

مثال : عرض التوزيع التكراري لقراءات شدة التعرض الاشعاعي لمجموعة من العاملين نتيجة لتعاملهم ع نوع معين من الاشعاع بطريقة المدرج التكراري

C	Fi
50 -	1
55 -	2
60 -	11
65 -	10
70 -	12
75 -	21
80 -	6
85 -	9
90 -	4
95 - 100	4
Total	80

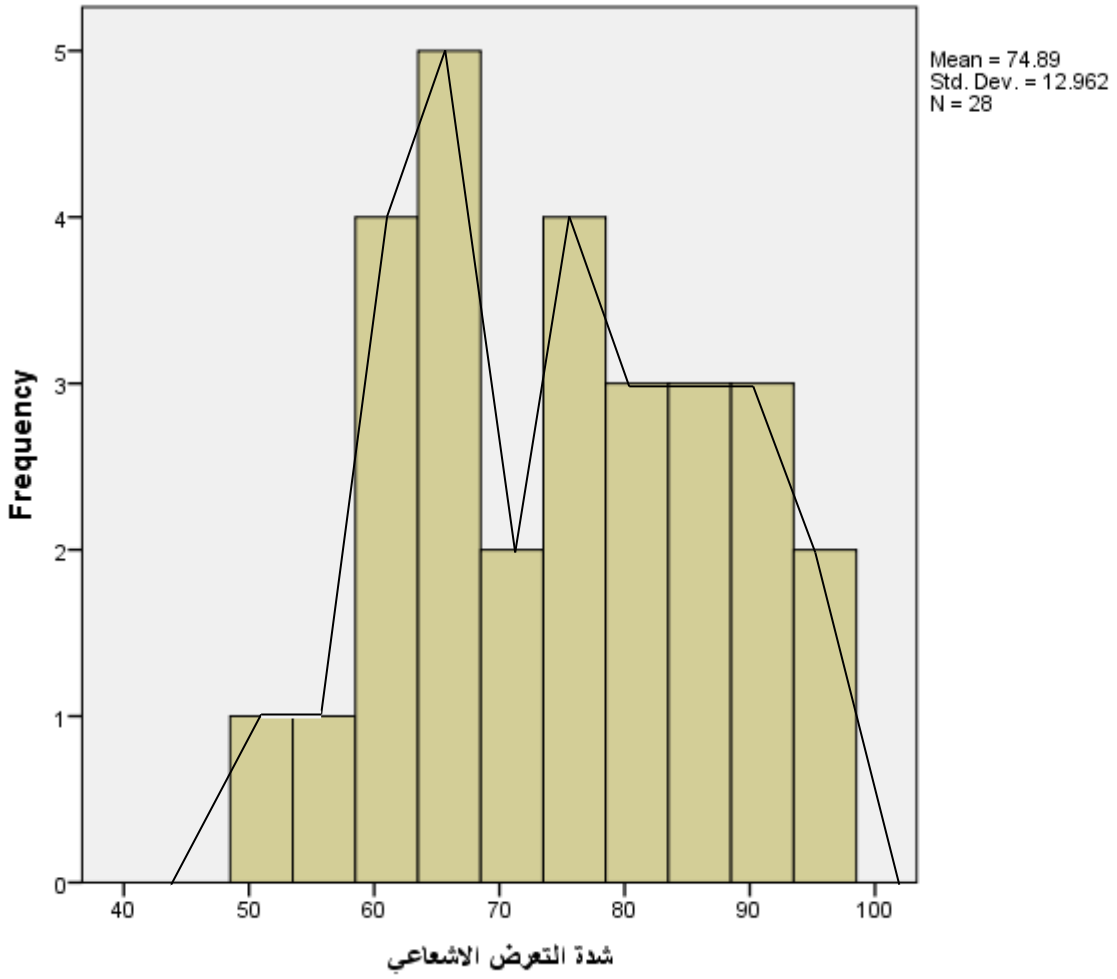


م.فاطمة عبد ا

2. المضلع التكراري Frequency – Polygon

تستخدم هذه الطريقة بشكل خاص عند اجراء المقارنة بين توزيعين تكراريين (او اكثر) حيث يصعب اجراء الطريقة السابقة وذلك برسم المدرجين التكراريين على نفس المحور وذلك نتيجة لتداخل المستطيلات او التمييز بين توزيعين.

ولعرض البيانات المبوبة بموجب هذه الطريقة فأننا نقوم بتقسيم المحورين المتعامدين بنفس الاسلوب المتبع عند تقسيم المحورين باستخدام المدرج التكراري وبعد تحديد مراكز الفئات على المحور الافقي والتكرارات على المحور العمودي وبعد تحديد النقاط يتم ايصالها بمستقيمات تبدأ من نقطة على المحور الافقي تبتعد بنصف طول الفئة الاولى عن الحد الادنى لها وتنتهي بنصف طول الفئة الاخيرة عن الحد الاعلى لهاوبذلك نحصل على مضلع تكراري مقفل مع المحور الافقي ، كما يمكن رسم المضلع في المدرج وذلك باخذ منتصفات القواعد العليا للمستطيلات في المدرج التكراري ومن ثم نصلها بمستقيمات فنحصل على المضلع التكراري.

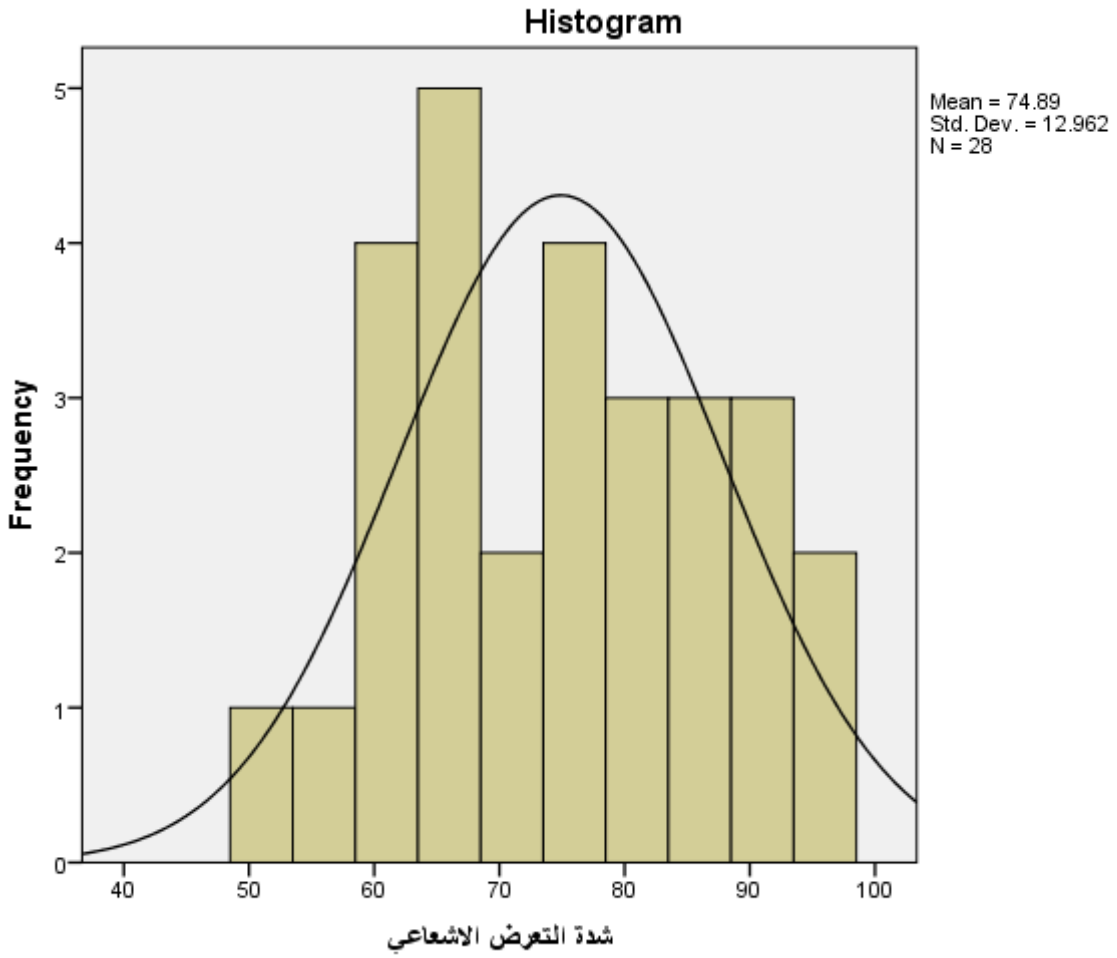


3. المنحنى التكراري Frequency Curve

تعتبر هذه الطريقة من الطرق الشائعة الاستخدام قياسا بالطرق الاخرى ويفضل استخدامها في حالة اجراء المقارنات بين ظاهرتين (او اكثر) ولتنفيذ هذه الطريقة يلزم تحديد نقط المضلع التكراري ومن ثم نمهد منحنى غير منكسر لايشترط ان يمر بجميع النقط.

مثال:بيانات الجدول التالي تمثل اوزان 300 شخص ، المطلوب عرضها بيانيا.

الفئات c	التكرارfi
25 -	3
35 -	12
45 -	72
55 -	47
65 -	97
75 -	52
85 - 95	27



الاسبوع الثاني

المنهج الاحصائى فى البحث العلمى ومنهج تصميم البحوث :

مفهوم المنهج الإحصائي

يُعرَّف المنهج الإحصائي على أنه مجموعة من المبادئ والأساليب المستخدمة لجمع البيانات الإحصائية وتحليلها، بهدف استخراج استدلالات علمية قد تكون غير واضحة. كما يعتمد هذا المنهج على توظيف الطرق الرقمية والرياضية، إلى جانب مجموعة من الأدوات الإحصائية، من أجل جمع البيانات بدقة وتحليلها وتفسيرها بشكل منهجي.

أنواع المنهج الإحصائي

يشمل المنهج الإحصائي في البحث العلمي عدة أنواع، أبرزها ما يلي:-

1.المنهج الإحصائي الوصفي

يحتل المنهج الإحصائي الوصفي مكانة بارزة بين المناهج الإحصائية نظرًا لأهميته الكبيرة، حيث يُستخدم في جمع البيانات الرقمية المتعلقة بموضوع أو ظاهرة معينة ، ثم وصفها وتحليلها بدقة. ويهدف هذا المنهج إلى تقديم تفسير منطقي وصحيح يعتمد على النتائج المستخلصة، مما يسهم في فهم الظواهر المختلفة بشكل أكثر دقة ووضوح.

2.المنهج الإحصائي الاستدلالي

يرتكز المنهج الإحصائي الاستدلالي على تحليل عينة ممثلة لمجتمع واسع، حيث يتم دراسة البيانات الرقمية المستخلصة منها وتفسيرها، بهدف التوصل إلى تعميمات واستنتاجات قابلة للتطبيق على نطاق أوسع. ويهدف هذا المنهج إلى الوصول إلى نتائج دقيقة ومنطقية تعكس خصائص المجتمع المدروس

مجالات استخدامات المنهج الإحصائي

يستخدم المنهج الإحصائي في البحث العلمي في العديد من المجالات ومن اهم استخداماته ما يلي

1. تحليل البيانات في البحوث العلمية لاستخلاص النتائج الدقيقة
2. دراسة الظواهر الاجتماعية وفهم تأثير العوامل المختلفة عليها
3. دعم القرارات الاقتصادية والتجارية بناءً على البيانات والإحصائيات
4. التنبؤ بالاتجاهات المستقبلية في مختلف المجالات مثل السوق والطقس
5. تقييم فعالية السياسات والبرامج في المؤسسات والمنظمات
6. تحسين جودة العملية الإنتاجية في المجالات الصناعية والتجارية

الأهداف الرئيسية للمنهج الإحصائي

1. جمع البيانات عبر خطوات منظمة وبطريقة منهجية
2. تحليل البيانات واستخلاص استنتاجات دقيقة
3. تفسير الظواهر العلمية والاجتماعية والاقتصادية
4. دعم عملية اتخاذ القرار بناءً على الأدلة الإحصائية
5. التنبؤ بالاتجاهات المستقبلية من خلال تحليل البيانات
6. قياس العلاقة بين المتغيرات المختلفة وفهم تأثيرها
7. تحسين جودة الأبحاث العلمية والدراسات الميدانية

مراحل المنهج الإحصائي في البحث العلمي

يمر المنهج الإحصائي في البحث العلمي بعدة مراحل أساسية تضمن دقة نتائجه، مما يساعد الباحثين في تحليل الظواهر المختلفة بكفاءة، ويوفر الوقت والجهد المبذول في عملية البحث. وفيما يلي عرض لهذه المراحل :-

1. **تحديد المشكلة أو الظاهرة محل الدراسة:** يجب أن تكون الظاهرة المختارة مرتبطة بموضوع البحث، حيث تُعد الخطوة الأساسية التي ينطلق منها الباحث في دراسته.

2. **صياغة الفروض:** تُستخدم الفروض لتحديد العلاقات بين الظواهر المختلفة، سواء بإثبات وجود ارتباط بينها أو نفيه، مما يسهم في توجيه البحث.

3. **التعاريف الإجرائية:** يتم وضع مؤشرات كمية واضحة لكل ظاهرة، مما يسهل عملية القياس والتحليل لاحقًا.

4. **جمع البيانات الإحصائية:** تُجمع البيانات بطرق مختلفة مثل المقابلات، الملاحظة، الاستبيانات، أو من خلال المصادر المتعددة، مع مراعاة تحديد نطاق البحث من حيث الزمن، الفئة المستهدفة، والموقع الجغرافي، إضافة إلى تحديد المفاهيم البحثية لضمان دقة البيانات.

5. **مراجعة البيانات وتنقيحها:** يتم فحص البيانات المُجمعة والتأكد من خلوها من الأخطاء أو التضارب أو المعلومات المضللة لضمان صحة التحليل.

6. **فرز البيانات وتبويبها:** تُصنف البيانات وتُجدول بطريقة منظمة، مع تمثيلها بيانيًا لتوفير رؤية واضحة للباحث حول الظاهرة، مما يسهل عملية التحليل والاستنتاج.

7. **تحليل البيانات:** تُعد هذه المرحلة من أهم خطوات البحث وتعتمد على دقة الفرز والتبويب السابق، ويأخذ التحليل شكلين

- تحليل كيفي: يتم من خلاله مقارنة البيانات، وضع الفروض المفسرة للظاهرة، والاعتماد على القوانين والنظريات العلمية في صياغة استنتاجات واضحة.
- تحليل كمي: يعتمد على الأساليب الرياضية والإحصائية، مثل حساب المتوسطات العامة، دراسة الارتباطات، واستخدام الرسوم البيانية لتوضيح النتائج.

8. **تفسير البيانات والنتائج:** بعد تحليل البيانات، يبدأ الباحث في تفسيرها وربطها بالنتائج المستخلصة، مما يساعد في التأكد من صحتها ودقتها، ثم تقديم تفسيرات منطقية تسهم في فهم الظاهرة بوضوح.

9. **عرض النتائج والتوصيات:** في المرحلة الأخيرة، يقوم الباحث بعرض النتائج التي توصل إليها بشكل موجز، إلى جانب تقديم التوصيات والتوقعات المستقبلية المتعلقة بالظاهرة المدروسة.

مزايا المنهج الإحصائي في البحث العلمي

1. يوفر أدوات دقيقة لتحليل البيانات والتوصل إلى استنتاجات موثوقة
2. يساعد في تفسير الظواهر المختلفة استنادًا إلى أسس علمية
3. يسهم في دعم عملية اتخاذ القرارات بناءً على بيانات رقمية واضحة
4. يتيح التنبؤ بالاتجاهات المستقبلية من خلال تحليل البيانات المتوفرة
5. يمكن تطبيقه في مختلف المجالات مثل العلوم الاجتماعية، الاقتصاد، والصحة
6. يسهل مقارنة البيانات ودراسة العلاقات بين المتغيرات المختلفة
7. يوفر طرقًا إحصائية متنوعة تساهم في زيادة دقة النتائج
8. يقلل من نسبة الأخطاء البحثية من خلال استخدام أساليب منظمة في جمع البيانات وتحليلها

عيوب المنهج الإحصائي في البحث العلمي

1.يعتمد بشكل كبير على صحة البيانات المُجمعة، مما قد يؤدي إلى نتائج غير دقيقة إذا كانت البيانات غير موثوقة
2.قد يكون تفسير البيانات الإحصائية معقدًا ويتطلب خبرة في التحليل الرياضي والإحصائي
لا يمكنه دائمًا تفسير الأسباب العميقة وراء الظواهر، حيث يركز على العلاقة بين المتغيرات دون التعمق في
3.العوامل المؤثرة
4.يحتاج إلى أدوات وبرامج متخصصة في التحليل الإحصائي، مما قد يشكل تحديًا لبعض الباحثين
5.قد يؤدي سوء اختيار العينة إلى نتائج غير دقيقة وغير قابلة للتعميم
6.يقتصر على تحليل البيانات الكمية، مما يجعله غير كافٍ بمفرده لدراسة الظواهر التي تحتاج إلى تفسيرات نوعية
7.قد يكون مكلفًا من حيث الوقت والموارد، خاصة عند جمع كميات كبيرة من البيانات ومعالجتها

وبذلك يُعتبر المنهج الإحصائي في البحث العلمي أداة فعالة تسهم في تنظيم البيانات وتحليلها بطريقة علمية دقيقة، مما يساعد في فهم الظواهر المختلفة واتخاذ قرارات مبنية على أسس رقمية واضحة. وعلى الرغم من التحديات المرتبطة بتطبيقه، إلا أن مزاياه العديدة تجعله جزءًا لا غنى عنه في الأبحاث العلمية بمختلف مجالاتها. لذا، فإن تطوير المهارات الإحصائية لدى الباحثين واستخدام الأدوات الحديثة في التحليل يسهم في تحقيق نتائج أكثر دقة وفعالية، مما يعزز من .جودة البحث العلمي ويدعم مسيرة التقدم في شتى المجالات

منهجية تصميم البحث :- تم تعريف تصميم البحث على أنه الخطة أو الهيكل الشامل الذي يوجه عملية إجراء البحث. وهو عنصر حاسم في عملية البحث ويعمل كمخطط لكيفية إجراء الدراسة، بما في ذلك الأساليب والتقنيات التي سيتم استخدامها لجمع البيانات وتحليلها. تعد الدراسة البحثية المصممة جيدًا أمرًا ضروريًا لضمان تحقيق أهداف البحث وأن النتائج صحيحة وموثوقة.

تشمل العناصر الرئيسية لتصميم البحث ما يلي :-

أهداف البحث : تحديد أهداف وغايات الدراسة البحثية بوضوح. ما الذي يحاول البحث تحقيقه أو التحقيق فيه؟

أسئلة أو فرضيات البحث : صياغة أسئلة أو فرضيات بحثية محددة تتناول أهداف الدراسة. هذه الأسئلة توجه عملية البحث.

طرق جمع البيانات : تحديد كيفية جمع البيانات، سواء من خلال المسوحات، أو التجارب، أو الملاحظات، أو المقابلات، أو البحوث الأرشيفية، أو مزيج من هذه الأساليب.

أخذ العينات : تحديد السكان المستهدفين واختيار العينة التي تمثل تلك السكان. يمكن أن تختلف طرق أخذ العينات، مثل أخذ العينات العشوائية، أو أخذ العينات الطبقية، أو أخذ العينات الملائمة.

أدوات جمع البيانات : تطوير أو اختيار الأدوات والأدوات اللازمة لجمع البيانات، مثل الاستبيانات أو المسوحات أو المعدات التجريبية.

تحليل البيانات : تحديد الأساليب الإحصائية أو التحليلية التي سيتم استخدامها لتحليل البيانات المجمعة. وقد **البحث** أو كمية، اعتمادا على أهداف نوعية أساليب يشمل ذلك.

الإطار الزمني : وضع جدول زمني لمشروع البحث، بما في ذلك متى سيتم جمع البيانات وتحليلها والإبلاغ . عنها

الاعتبارات الأخلاقية : معالجة القضايا الأخلاقية، بما في ذلك الحصول على موافقة مستنيرة من المشاركين، وضمان خصوصية وسرية البيانات، والالتزام بالمبادئ التوجيهية الأخلاقية.

الموارد: بما في ذلك التمويل والموظفين والمعدات والوصول إلى مصادر للبحث تحديد الموارد اللازمة للبيانات.

عرض البيانات وإعداد التقارير : تخطيط كيفية عرض نتائج البحث والإبلاغ عنها، سواء من خلال التقارير المكتوبة أو العروض التقديمية أو غيرها من التنسيقات.

هناك العديد من التصاميم البحثية، مثل التجريبية، والرصدية، والمسحية، ودراسة الحالة، والتصميمات الطولية، وكل منها يناسب أسئلة وأهداف البحث المختلفة. يعتمد اختيار تصميم البحث على طبيعة البحث وأهداف الدراسة يعد تصميم البحث الجيد أمرًا بالغ الأهمية لأنه يساعد على ضمان صحة نتائج البحث وموثوقيتها وقابليتها للتعميم، .مما يسمح للباحثين باستخلاص استنتاجات ذات معنى والمساهمة في مجموعة المعرفة في مجال تخصصهم

انواع تصاميم البحوث

هناك عدة أنواع من تصاميم البحث، كل منها مصمم للإجابة على أسئلة بحثية محددة وتحقيق أهداف معينة. يعتمد اختيار تصميم البحث على طبيعة مشكلة البحث وأهداف الدراسة. فيما يلي عدة أنواع نموذجية من تصاميم البحث

1.تصميم البحث التجريبي : يتم تعيين الأفراد بشكل عشوائي

2.تصميم البحوث شبه التجريبية : تتم مقارنة مجموعتين أو أكثر، ولكن لا يتم تعيين المشاركين بشكل عشوائي.

3.تصميم البحوث الرصدية : يتم جمع البيانات من نقطة زمنية واحدة لفحص العلاقات أو الاختلافات بين المتغيرات. أنها لا تنطوي على المتابعة مع مرور الوقت

4.تصميم البحث الوصفي : تتضمن دراسات الحالة فحصًا متعمقًا لفرد أو مجموعة أو ظاهرة واحدة. غالبًا ما يتم استخدامها للحصول على فهم عميق لحالة فريدة

5.تصميم البحوث الارتباطية يفحص هذا التصميم العلاقات بين متغيرين أو أكثر لتحديد ما إذا كانت مرتبطة أم لا.

6.تصميم البحث بأثر رجعي : يقوم الباحثون بفحص الظروف أو السلوكيات الحالية والبحث عن الأسباب المحتملة بأثر رجعي

7.تصميم البحث الاستكشافي : الدراسة التجريبية هي تحقيق أولي صغير النطاق يتم إجراؤه قبل مشروع بحثي واسع النطاق لاختبار إجراءات البحث وجمع البيانات الأولية

8.دراسة الأتراب : تتبع الدراسات الأترابية مجموعة من الأفراد (الأتراب) على مدى فترة من الزمن لتقييم تطور نتائج أو ظروف محددة. وهي شائعة في علم الأوبئة

9.البحث العملي : غالبًا ما يستخدم البحث الإجرائي في البيئات التعليمية أو التنظيمية. يعمل الباحثون بشكل تعاوني مع الممارسين لمعالجة المشاكل العملية وإجراء التحسينات

10.التحليل التلوي: يتضمن التحليل التلوي التوليف الإحصائي للبيانات من دراسات متعددة حول نفس الموضوع لتقديم نظرة عامة أكثر شمولاً لنتائج البحث

11. تصميم متسلسل : يجمع هذا التصميم بين عناصر البحث المقطعي والطولي لفحص التغيرات المرتبطة بالعمر أثناء مقارنة الأفواج المختلفة

12.النظرية المؤرضة : النظرية المرتكزة هي منهج بحث نوعي يركز على تطوير النظريات أو التفسيرات المرتكزة على البيانات التي تم جمعها أثناء عملية البحث

.

ولكل من هذه التصاميم البحثية نقاط قوة ونقاط ضعف، ويعتمد اختيار التصميم على سؤال البحث، والموارد المتاحة، والاعتبارات الأخلاقية، وطبيعة البيانات اللازمة لتحقيق أهداف البحث. غالبًا ما يختار الباحثون التصميم الذي يتوافق بشكل أفضل مع أهدافهم وقيودهم البحثية المحددة

مقاييس النزعة المركزية Measures of Central Tendency : هي المقاييس الاحصائية التي تحقق وصف حسابي او عددي للبيانات التي جمعت ورتبت سابقا ى أن الغرض من حساب المتوسطات هو الحصول على قيمة مركزية لمجموعة بيانات واستخدامها لتوفير معلومات قيّمة .

1.الوسط الحسابي Arithmetic Mean : اهم مقاييس النزعة المركزية واكثرها استخداما ،فمن مميزاته انه معرف جبريا ويمكن فهمه واستيعابه ويستخدم في حسابه جميع المعلومات المتاحة

البيانات غير المبوبة

الوسط الحسابي لمجموعة مؤلفة من (n) من القيم التي يأخذها متغير ما مثل (x) حيث ان $x_i : x_1 , x_2 , x_3 , \ldots\ldots, x_n$

فأن الوسط الحسابي لهذه المجموعة من القيم والذي يرمز اليه بالرمز $(\bar{x})$ ويقرأ (x bar) ويستخرج كما يلي :

1.الطريقة المطولة

$$\bar{x} = \frac{x1 + x2 + \cdots + xn}{n}$$

وللاختصار في الكتابة نستعمل الحرف اليوناني ($\sum$) ونعني به هنا جمع الحدود التي بداخله اي ان :

$$\sum_{i=1}^{n} xi = x1 + x2 + \cdots . + xn$$

وبذلك يصبح الوسط الحسابي

$$\bar{x} = \frac{\sum_{i=1}^{n} xi}{n}$$

2. ألطريقه المختصرة

3.

خطوات هذه الطريقه :-

1- أختيار احدى القيم كوسط فرضي a من بين قيم المتغير xi ويستحسن ان تكون قريبه من الوسط.
2- نجد انحرافات القيم عن الوسط الفرضي ونرمزله بالرمز di حيث ان $di = xi - a$.
3- ثم نطبق الصيغة التاليه لأيجاد الوسط الحسابي

$$\bar{x} = a + \frac{\sum_{i=1}^{n} di}{n}$$

حيث ان

a = تمثل وسط فرضي يتم اختياره من بين قيم المتغير Xi
∑ di تمثل مجموع انحرافات قيم المتغير xi عن الوسط الفرضي .
n = عدد قيم الظاهره.

مثال (1):البيانات التالية تمثل اليوريا في الدم ل(15) شخصا , 44 , 37 , 43 , 48 , 38 , 46 , 42 , 43 , 37)
38 , 39 , 37 , 42 , 38 , 45)
المطلوب استخراج الوسط الحسابي بالطريقتين.

مثال(2) :جد الوسط الحسابي للبيانات التالية بالطريقتين المطولة والمختصرة.

xi : 5,18,5,15,7,12,9,11

البيانات المبوبة

1.*الطريقة المطولة* :اذا كانت القيم $(x_1,x_2,\ldots.x_n)$ تحدث بتكرارات $(f_1,f_2,\ldots.,f_n)$ مرة على الترتيب ، فان الوسط الحسابي سيكون:

$$\bar{x} = \frac{f1m1 + f2m2 + \cdots. + fnmn}{f1 + f2 + \cdots. + fn}$$

اي

$$\bar{x} = \frac{\sum_{i=1}^{n} fimi}{\sum_{i=1}^{n} fi}$$

حيث (mi) مركز الفئة ويساوي $mi = \frac{u.c.l+l.c.l}{2}$

مثال:لبيانات الجدول التالي جد الوسط الحسابي باستخدام الطريقة المطولة .

c	Fi
5 -	18
10 -	25
15 -	10
20 - 25	7

مزايا الوسط الحسابي

1.سهولة الحساب والفهم : يتميز بأنه سهل الحساب والفهم مقارنة بالعديد من المقاييس الإحصائية الأخرى .

2.استخدام جميع القيم : يأخذ في الاعتبار كل قيمة من القيم في مجموعة البيانات عند حسابه .

3.مرونة في العمليات الجبرية :يمكن استخدامه بسهولة في العمليات الجبرية، مما يجعله مفيدًا في التحليلات الإحصائية
4.مقياس شائع : هو أكثر المقاييس الإحصائية استخدامًا وشيوعًا وفهمًا.اللاحقة

عيوب الوسط الحسابي:-

1.يتأثر بالقيم المتطرفة او الشاذة
2.يصعب استخدامه في حالة الجداول التكرارية المفتوحة

المتوسط التوافقي Harmonic Mean

المتوسط التوافقي هو نوع من المتوسط يكون مفيدًا بشكل خاص في المواقف التي تكون فيها المعدلات المتوسطة مرغوبة. على عكس الوسط الحسابي، الذي يجمع القيم ويقسمها على العدد، يتم حساب المتوسط التوافقي عن طريق أخذ مقلوب متوسط مقلوب مجموعة من الأرقام. تجعل طريقة الحساب هذه المتوسط التوافقي فعالاً بشكل خاص في السيناريوهات التي تتضمن معدلات، مثل السرعة أو الكثافة، حيث ترتبط القيم التي يتم حساب متوسطها عكسيًا بالكمية التي يتم قياسها،المتوسط التوافقي لقائمة أرقام يميل بشدة نحو العناصر الأصغر فيها، فإنه يميل (مقارنةً بالمتوسط الحسابي (إلى تخفيف تأثير القيم المتطرفة الكبيرة وتفاقم تأثير القيم الصغيرة . وكثيرًا ما يُستخدم المتوسط الحسابي خطأً في الحالات التي تستدعي استخدام المتوسط التوافقي.

- **للبيانات غير المبوبة**

الوسط التوافقي (H) لمجموعة من القيم $(x_1, x_2, x_3, .., x_n)$ هو مقلوب الوسط الحسابي لمقلوبات هذه القيم ، ويمكن حسابه بهذه الصيغة :-

$$H = \frac{1}{\frac{1}{n}\sum\frac{1}{x_i}} = \frac{n}{\sum\frac{1}{x_i}}$$

حيث ان :-
Xi : مفردات القيم ، n : عدد القيم

مثال : جد الوسط التوافقي للقيم التالية : 2 4 8 :xi

- **للبيانات المبوبة**

$$H = \frac{\sum fi}{\sum\frac{fi}{mi}}$$

حيث ان :-
Fi : التكرارات
Mi : مراكز الفئات

مثال : بيانات الجدول التالي يبين ذوبان 100 مركب في محلول معين (مقاسة بالثواني)

فئات سرع الذوبان c	2	7	12	17-22	∑
عدد المركبات المذابة fi	20	50	20	10	100

مزايا الوسط التوافقي

1. يستخدم جميع البيانات : يتم حساب المتوسط التوافقي بناءً على جميع البيانات في السلسلة، دون تجاهل أي عنصر منها، مما يجعله أكثر شمولية في بعض الحالات.

2.موثوقية عالية للبيانات المتطابقة : يوفر نتيجة موثوقة عندما تكون النتائج المطلوبة متطابقة لوسائل مختلفة.

3.لا يتأثر بأخطاء المعاينة : يعتبر الوسط التوافقي أكثر مقاومة لأخطاء المعاينة، مما يزيد من موثوقيته في بعض التحليلات.

4.مناسب لحسابات المعدلات : يستخدم بشكل خاص عند حساب متوسط المعدلات أو السرعات، حيث يكون القيمة المقلوبة هي المهم.

عيوب الوسط التوافقي

1.صعوبة الحساب : قد يكون حساب المتوسط التوافقي أكثر تعقيدًا وصعوبة في بعض الأحيان مقارنة بالمتوسط الحسابي

2.عدم امكانية استخدامه مع قيم صفرية او سالبة : لايمكن حسابه في حال وجود قيم صفرية او سالبة في مجموعة البيانات ، لان ذلك يؤدي الى نتائج غير محددة.

3.غير مناسب للبيانات غير المعدلات : إذا كانت البيانات لا تمثل معدلات أو نسب، فقد لا يوفر المتوسط التوافقي نتيجة ذات معنى.

4.صعوبة حسابه للتوزيعات المفتوحة **:** لا يمكن حساب المتوسط التوافقي للتوزيعات التكرارية المفتوحة.

المتوسط التربيعي Squares Mean

الوسط التربيعي $(x_1, x_2, \ldots\ldots, x_n)$ هو الجذر التربيعي لمتوسط مربعات تلك القيم ويعطى بالصيغة الاتية :- لمجموعة من القيم

للبيانات غير المبوبة

$$S.M = \sqrt{\frac{\sum xi^2}{n}}$$

مثال :- جد المتوسط التربيعي للبيانات التالية :-

Xi : 12 7 5 2 -3

للبيانات المبوبة

يتم ايجاد المتوسط التربيعي للبيانات المبوبة بالصيغة التالية:-

$$S.M = \sqrt{\frac{\sum fi\ mi^2}{\sum fi}}$$

حيث ان :-

Fi: التكرارات

mi^2 : مربع مراكز الفئات

ويستخدم هذا النوع من المتوسطات بكثرة في التطبيقات الطبيعية

مزايا الوسط التربيعي

1. بسبب تربيع الفروقات، فإن الوسط التربيعي يعطي وزناً أكبر للأخطاء الكبيرة، مما يجعله حساساً للقيم المتطرفة
2. تعريف الوسط التربيعي بأنه الجذر التربيعي الموجب لمتوسط مربعات القيم يعني أن نتيجته تكون دائمًا موجبة
3. يستخدم في تقييم أداء النماذج الإحصائية حيث يقيس مدى تشتت الخطأ بين القيم المتوقعة والفعلية

عيوب الوسط التربيعي

1. يتأثر الوسط التربيعي بشكل كبير بالقيم المتطرفة، حيث يؤدي تربيع الأخطاء إلى تضخيم تأثير هذه القيم
2. قيمة الوسط التربيعي تعتمد على مقياس المتغير المستهدف، مما يجعل من الصعب مقارنة القيم بين متغيرات ذات مقاييس مختلفة.
3. يختلف عن المتوسط الحسابي في أنه ليس بسيطًا وسهل الحساب، بل يتطلب عمليات حسابية أكثر تعقيد

الوسط الهندسي Geometric Mean

. الوسط الهندسي لمجموعة من القيم يكون دائما اقل من الوسط الحسابي

للبيانات غير المبوبة

الوسط الهندسي لمجموعة من القيم $(x_1, x_2, \dots\dots, x_n)$ هو الجذر النوني (n) لحاصل ضرب تلك القيم ويعطى بالصيغة التالية :-

$$G = \sqrt[n]{(x_1)(x_2)\dots(x_n)}$$

حيث ان :-
Xi : مفردات القيم ، n : عدد القيم

مثال : اوجد الوسط الهندسي للقيم الموجبة التالية التي تتعلق بتركيز السكر في الدم لخمسة اشخاص (140 138 123 130 160)

للبيانات المبوبة

اذا كان لدينا توزيع تكراري ذو فئات وكانت مراكز فئاته قيما موجبة وهي $(m_1, m_2, \dots\dots, m_n)$ وكانت التكرارات المقابلة لها $(f_1, f_2, \dots, f_n)$ على الترتيب فأن قيمة الوسط الهندسي يتم ايجادها بالصيغة التالية :-

$$\boldsymbol{G = \sqrt[\Sigma f_i]{(f_i m_i)(f_i m_i)\dots(f_i m_i)}}$$

وباستخدام اللوغاريتمات فان الصيغة اعلاه تكون بالشكل الاتي :-

$$\boldsymbol{Log\ G = \frac{1}{\sum f_i}\sum f_i\, Log m_i}$$

مثال :- بيانات الجدول التالي تتضمن نسب تلوث الهواء ل (30) مدينة كبيرة بموجب مقياس (المايكرو غرام لكل متر مكعب) جد قيمة الوسط الهندسي.

c	10 -	20 -	30 -	40 -	50 -	60 -	70 - 80	T
fi	2	3	5	10	6	2	2	30

مزايا الوسط الهندسي

1. يعتبر انسب المتوسطات في قياس النزعة المركزية لمعدلات التغير. فمثلا في حالة تقدير عدد السكان بين سني التعداد فهنا يكون التغيير في السكان متناسبا مع عدد السكان نفسه.
2. يصلح لتمثيل متوسط مجموعة من النسب وذلك لما يتوافر فيه من مزايا لا تتوفر في غيره من المتوسطات .
3. الوسط الهندسي هو انسب الاوساط في حالة كون البيانات معبرا عنها بالنسب او بالمعادلات.

عيوب الوسط الهندسي

1. لايصح استخدامه لمجموعة من القيم اذا كانت احداها صفر فأن الوسط الهندسي بالتالي يساوي صفر.
2. لايصح استخدامه لمجموعة من القيم اذا كانت احداها سالبة فأن الوسط الهندسي يكون قيمة خيالية.

الوسيط Median

هو القيمة التي يسبقها نصف البيانات المرتبة ويليها النصف الآخر. أو يمكن تعريفه بأنه القيمة التي تقسم القيم إلى جزئيين بحيث يكون عدد القيم التي أقل منها يساوي عدد القيم التي أكبر منها. فالوسيط إذن هو متوسط ترتيب القيم.

البيانات غير المبوبة :

اذا كان عددها فرديا اي انه القيمة ذات الترتيب $\frac{n+1}{2}$

$$Me=x_{\frac{n+1}{2}}$$

اما اذاكان عدد القيم زوجي فان الوسيط هو الوسط الحسابي للقيمتين بالمنتصف اي ذات الترتيبين $\frac{n}{2}+(\frac{n}{2}+1)$

$$Me=\frac{x_{\frac{n}{2}}+x_{\frac{n}{2}+1}}{2}$$

مثال :فيما يلي مجموعتان من تفاعل لقياس المناعة لمرض التدرن (بالملم) جد قيمة الوسيط لكل مجموعة.

xi : 8,3,8,4,8,4,6,5,10

yi : 5,18,5,15,7,12,9,11

البيانات المبوبة

يمكن ايجاد الوسيط للبيانات المبوبة بطريقتين:

1.الطريقة الحسابية وفق الصيغة التالية

$$Me = l_1 + (\frac{\frac{\sum fi}{2} - (\sum f_1)}{f_m})w_m$$

بعد تكوين جدول تكراري متجمع صاعد

حيث ان :

l_1 :الحد الادنى للفئة الوسيطية (أي الفئة التي يقع فيها الوسيط)

$\frac{\sum fi}{2}$:ترتيب الفئة الوسيطية

$(\sum f_1)$:مجموع التكرارات لجميع الفئات قبل الفئة الوسيطية

f_m :تكرار الفئة الوسيطية

w_m :طول الفئة الوسيطية

ويتم تطبيق الصيغة اعلاه بعد تكوين جدول تكراري متجمع صاعد.

مثال تمثل البيانات التالية اطوال حياة خمسين حيوانا من حيوانات التجارب (الاشهر)
المطلوب :جد قيمة الوسيط بالطريقتين الحسابية والبيانية

c	25 -	30 -	35 -	40 -	45 - 50
fi	7	19	14	7	3

المنوال Mode

البيانات غير المبوبة :
هو القيمة او الصفة التي تتكرر اكثر من غيرها او هو القيمة الاكثر شيوعا وقد لايكون للقيم او الصفات منوال وقد يوجد منوال واحد او اكثر.

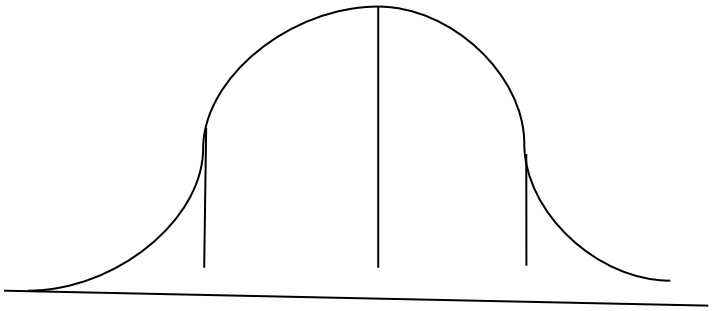
البيانات المبوبة : المنوال هو القيمة التي يقابلها اكبر تكرار
1.الطريقة الحسابية : يمكن ايجاد المنوال بالصيغة:-

$$Mo = L_1 + \left(\frac{\Delta 1}{\Delta 1+\Delta 2}\right) w$$

حيث ان :
L_1 الحد الادنى للفئة المنوالية
$\Delta 1$:الفرق الاول بين تكرار الفئة المنوالية والتكرار السابق لها$(\Delta_1=f-f_1)$
$\Delta 2$:الفرق الثاني بين تكرار الفئة المنوالية والتكرار اللاحق لها$(\Delta_2=f-f_2)$
w:طول الفئة المنوالية

مقاييس الربع Quarter measures

لقد تبين سابقا ان قيمة الوسيط هي النقطة التي تقع على المحور الافقي بحيث تكون المساحة تحت المضلع التكراري متساوية على طرفيها ، وبتعميم هذه الفكرة يمكن ان نضع نقاطا على المحور الافقي تقسم المجموعة (المساحة تحت المضلع التكراري) الى اجزاء متساوية من التقسيمات تسمى بقيم التقسيمات الجزئية وان اهم التقسيمات التي نحن بصددها هي تقسيم المجموعة الى اربعة اجزاء متساوية والتي يرمز لها بالرموز (Q1,Q2,Q3) وهي الربيع الاول والربيع الثاني والربيع الثالث على الترتيب كما في الشكل:



مثال : التالي درجات الاحصاء في صفٍّ من 19 طالبًا

98، 95، 90، 87، 84، 82، 81، 77، 76، 75، 72، 75، 70، 69، 68، 65، 65، 60، 59

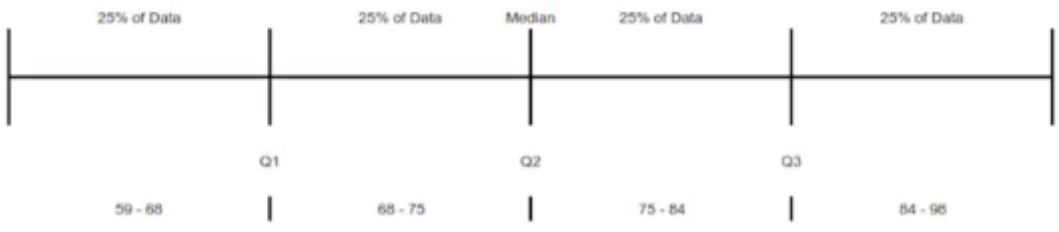
تتم حساب الارباع كما يلي

- الربع الأول (Q1) = (n + 1) × 1/4
- الربع الثاني (Q2)الوسيط أو ، = (n + 1) × 2/4
- الربع الثالث (Q3) = (n + 1) × 3/4

- الربع الأول (Q1) = 20 × 1/4 = 5
- الربع الثاني (Q2) = 20 × 2/4 = 10

- الربع الثالث 20 × 3/4 = 15

الرسم البياني كالتالي



العلاقة بين المتوسطات (الوسط الحسابي ، الوسيط ، المنوال)

1. في المنحنيات التكرارية المتماثلة فيتطابق كل من الوسط الحسابي والوسيط والنوال اي :-

$$\overline{x = me = mo}$$

2.(في التوزيعات غير المتماثلة (المنحرفة
.الوسط الحسابي >الوسيط >المنوال

3. لتوزيع المنحرف إلى اليسار (سالب الانحراف)
.المنوال >الوسيط >الوسط الحسابي

4. توجد صيغة تقريبية تربط بين القيم الثلاث في حالة الانحراف

المتوسط الحسابي = 3* الوسط الحسابي – 2* المنوال

يمكن استخدام هذه الصيغة لتقدير قيمة واحدة إذا كانت القيمتان الأخريان معروفتين

5.- عند حساب الوسط الحسابي نأخذ في الحسبان كل البيانات، بينما الوسيط يهتم فقط بالأماكن المركزية، وعند المنوال نركز على القيم الأكثر تكراراً.

6. عند وجود البيانات غير المبوبة ، الوسط الحسابي هو الأكثر ملائمة عندما تكون كل البيانات قريبة من بعضها، لكن إذا حصل العكس (وجود احد البيانات أو أكثر بعيداً نسبياً عن الباقي) فإن الوسيط هو الأكثر ملائمة.
7.- عند وجود بيانات مبوبة ، فالوسط الحسابي هو الملائم إذا توفرت 3 شروط مجتمعة:
- التوزيع مغلق (نعرف كل حدود الفئات).
- كل الفئات تحتوي على نفس الطول.
- التوزيعات متماثلة.

إذا لم يتوفر احد هذه الشروط أو أكثر فالوسيط هو الملائم.
8.- إذا كان التوزيع مفتوح نستطيع فقط حساب الوسيط. لكن إذا وقع مكان الوسيط في فئة مفتوحة لا يمكن أن نحسب الوسيط.
9.- إذا كل الفئات تحتوي على نفس الطول يمكننا حساب الوسط الحسابي بالطريقة المطولة والطريقة المختصرة، لكن إذا وجدنا إحدى الفئات بطول مختلف يمكن حساب الوسط الحسابي فقط بواسطة الطريقة المطولة.
10.- لا يمكن أن نحسب المنوال المعدل إذا كانت إحدى فئات البيانات المبوبة مفتوحة. لكن يمكن أن نحسب المنوال غير المعدل.

الاسبوع الثامن والتاسع

مقاييس التشتت Measures of Dispersion

مفهوم التشتت :هي احدى المقاييس التي تلقي الضوء على مدى اختلاف البيانات فيما بينها وتقيس مدى ذلك التفاوت او التغير بين مفرداتها اي ان التشتت(درجة التغير التي تتجه بها البيانات الرقمية عن بعضها البعض او حول قيمة وسطى)

1.المدى Range :

للبيانات غير المبوبة

يعرف المدى لمجموعة من القيم غير المبوبة بأنه الفرق بين اكبر قيمة واصغر قيمة في المجموعة.
فاذا تم ترتيب مجموعة من القيم ($x_1,x_2,\ldots..,x_n$) ترتيبا تصاعديا فان المدى يعطى بالصيغة الاتية:-

$$R = x_n - x_1$$

مثال:اذا كان لدينا القيم (30 , 33 , 35 , 38 , 40) فان المدى

$$R= 40 - 30 =10$$

للبيانات المبوبة

بانه الفرق بين الحد الاعلى للفئة العليا (الاخيرة) والحد الادنى للفئة الدنيا(الاولى)

$$R = \frac{u.c.l - l.c.l}{2}$$

وبذلك فان المدى لايعتمد على جميع البيانات ، بل على اكبر واصغر قيمة فقط، الامر الذي يقلل من اهمية استخدامه في حالة كون احدى او كلا القيمتين المتطرفتين (اكبر واصغر قيمة) شاذة ففي هذه الحالة يكون المدى كبيرا بينما نجد ان باقي مفردات البيانات ليست متباعدة عن بعضها البعض ، كما ان قيمة المدى تتأثر وفقا لحجم العينة الماخوذة فيزداد بازدياد مفرداتهاولذلك لايعتمد عليه لتقدير مقدار التغير في المجتمع ، اذ ان قيمته تتوقف على حجم العينة المختارة

ويمكن التخلص من العيب الذي يشوب المدى وهو تأثره بالقيم الشاذة وذلك باستخدام مقاييس اخرى تعبر عن تحسين لقيمته تدعى بشبيهات المدى.

2.المدى الربيعي Quartile Deviation

يعرف المدى الربيعي لمجموعة من القيم (المبوبة وغير المبوبة)بالفرق بين الربيعين الاعلى والادنى ويعطى بالصيغة التالية :-

$$Q.D = Q_3 - Q_1$$

للبيانات غير المبوبة

كيفية حساب المدى الربيعي

1.ترتيب جميع القيم في مجموعة البيانات ترتيبًا تصاعديًا من الأصغر إلى الأكبر

2.ايجاد قيمة الوسيط بعد تحديد القيمة الوسطى لمجموعة البيانات بالكامل

3.تحديد القيمة الوسطى للنصف السفلي من مجموعة البيانات قبل الوسيط (Q1) الربيع الأول

4.تحديد القيمة الوسطى للنصف العلوي من مجموعة البيانات بعد الوسيط (Q3) الربيع الثالث

5:حساب المدى الربيعي بطرح الربيع الأول من الربيع الثالث.

مثال احسب المدى الربيعي للبيانات التالية :- 1، 2، 4، 4، 6، 7، 9، 12

للبيانات المبوبة

حيث ان :-

Q1 : الربيع الاول او الادنى وهو القيمة التي يسبقها (1/4)البيانات ويتبعها (¾) من البيانات ، وعى فرض ان ان القيم مرتبة ترتيبا تصاعديا ، هذه في حالة البيانات غير المبوبة ، اما في حالة البيانات المبوبة فتستخرج تلك القيمة بموجب الصيغة التالية :-

$$\mathbf{Q_1 = q_1 + (\frac{\frac{n}{4} - (\sum f)q_1}{f_{q1}})w_{q1}}$$

حيث ان :-

Q1 :الحد الادنى لفئة الربيع الادنى

$\sum Fq_1$:مجموع التكرارات لجميع الفئات قبل فئة الربيع الادنى

Fq_1 : تكرار فئة الربيع الادنى

Wq_1: طول فئة الربيع الادنى

اما Q3 تشير الى الربيع الثالث (الاعلى) وهو القيمة التي يسبقها (3/4) من البيانات ويبعها (1/4) والبيانات المرتبة ترتيبا تصاعديا او تنازليا وتستخرج تلك القيمة في حالة البيات المبوبة بموجب الصيغة التالية :-

$$\boldsymbol{Q_3 = q_3 + (\frac{\frac{3n}{4} - (\sum f_3)q}{f_{q3}} w_{q3}}$$

حيث ان :-

q_3 :الحد الادنى لفئة الربيع الاعلى

$(\sum f3)q$:مجموع التكرارات لجميع الفئات قبل فئة الربيع الاعلى

Fq_3 : تكرار فئة الربيع الاعلى

Wq_3 طول فئة الربيع الاعلى

$$\boldsymbol{Q = Q_3 - Q_1}$$

الانحراف المتوسط The Mean Deviation

الانحراف المتوسط او متوسط الانحرافات المطلقة لمجموعة من القيم عن وسطها الحسابي وهو احد مقاييس التشتت.

للبيانات غير المبوبة

$$\mathbf{M.D} = \frac{\sum|x_i - \bar{x}|}{n}$$

مثال :اوجد متوسط الانحراف لمجموعة القيم (40 37 35 33 30)

للبيانات المبوبة

$$M.D = \frac{\sum f_i|m_i - \bar{x}|}{\sum f_i}$$

مثال : اوجد متوسط الانحرافات للتوزيع التكراري التالي:

c	10 -	20 -	30 -	40 -	50 - 60
fi	9	27	44	15	5

مميزات الانحراف المتوسط

1. سهولة حسابه
2. اكثر شمولا من المدى

عيوب الانحراف المتوسط

1. يعتمد على جميع المفردات في حسابه
2. اهماله للاشرات الجبرية والذي يجعل منه مفهوما غير جبري وبالتالي فأنه قليل الاهمية.

الاسبوع العاشر

التباين Variance

مقياس التباين هو مقياس إحصائي يُستخدم لقياس مدى تشتت أو انتشار مجموعة من البيانات عن متوسطها الحسابي . يتم حسابه عن طريق جمع مربعات الفروقات بين كل نقطة بيانات والمتوسط، ثم قسمة الناتج على عدد نقاط البيانات (أو عددها ناقص واحد إذا كانت عينة). كلما ارتفعت قيمة التباين، زاد انتشار البيانات، والعكس صحيح

للبيانات غير المبوبة

متوسط مربعات الانحراف لمجموعة من القيم(x_1 , x_2 , …. ,x_n) عن وسطها الحسابي.

$$s^2 = \frac{\sum_i^n (x_i - \bar{x}}{n-1}$$

حيث ان :-

S^2 : **التباين**

Xi : قيم الظاهرة

$\bar{x}$: الوسط الحسابي

n : عدد قيم الظاهرة

للبيانات المبوبة

يعطى بالصيغة التالية:

$$\boldsymbol{s^2 = \frac{\sum(m_i - \bar{x})^2 f_i}{\sum f_i - 1}}$$

حيث ان:-

S^2 :الانحراف المعياري

Mi :مراكز الفئات

$\bar{x}$: الوسط الحسابي

fi:التكرارات

مما سبق تبين ان مقياس التباين قد تضمن مربعات الانحرافات لذلك فان قيمته يعبر عنها بوحدات مربعة ولاعتبارات معينة تخص الكثير من المسائل يكون المرغوب فيه اعتماد نفس الوحدات الاحصائية المستخدمة للبيانات الاصلية وصولا الى تحقيق هذا الهدف لابد من اخذ الجذر التربيعي الموجب للتباين ةبذلك نحصل على الانحراف المعياري.

الانحراف المعياري The Standard Deviation

هو الجذرالتربيعي متوسط مربعات الانحراف لمجموعة من القيم(x_1 , x_2 , …. ,x_n) عن وسطها الحسابي.

البيانات غير المبوبة :

يعطى بالصيغة التالية:

$$s = \sqrt[2]{\frac{\sum_{i-1}^{n}(xi - \overline{x})^2}{n-1}}$$

حيث ان :-

S :الانحراف المعياري

Xi : قيم الظاهرة

$\bar{x}$: الوسط الحسابي

n : عدد قيم الظاهرة.

البيانات المبوبة :

يعطى بالصيغة التالية:

$$s = \sqrt[2]{\frac{\sum_{i-1}^{n}(mi - \overline{x})^2 fi}{\sum_{i-1}^{n} fi - 1}}$$

حيث ان:-

S :الانحراف المعياري

Mi :مراكز الفئات

$\bar{x}$: الوسط الحسابي

fi:التكرارات

تكون قيمة الانحراف المعياري اقل مايمكن اذا وفقط اذا كانت الانحرافات مقاسة عن الوسط الحسابي لذا يتم اعتماد قيمة الوسط الحسابي من بين مقاييس النزعة المركزية لايجاد تلك الانحرافات.

مقاييس التشتت النسبيMeasures of relative dispersion

يطلق على مقاييس التشتت السابقة بمقاييس التشتت المطلق حيث تعتمد جميعها على وحدة قياس مفردات المجموعة ، ففي حالة اجراء مقارنة تشتت مجموعتين من القيم تختلفان في وحدة القياس او عند اجراء مقارنة تشتت بين صفتين متصلتين بنفس المجموعة فأن من الخطأ ان نقارن دون ازالة اثر الاختلاف في وحدات القياس لذلك نلجأ الى مايعرف بمقاييس التشتت النسبي والتي تعرف:

مقاييس التشتت النسبي لمجموعة من القيم $(x_1, x_2, \dots, x_n)$يعطى بالصيغة التالية:

$$\text{مقياس التشتت النسبي} = \frac{\text{مقياس التشتت المطلق}}{\text{مقياس المتوسط}}$$

من مقاييس التشتت النسبي

1.معامل التباين او الاختلاف Coeffcient Of Variation

وهو معامل التشتت النسبي اذا كان مقياس التشتت المطلق الانحراف المعياري والمتوسط هو الوسط الحسابي ويعتبر من اكثر معاملات التغير انتشارا ويستخدم للمقارنة بين التغير في عدة مجموعات او توزيعات تكرارية تختلف بعضها او جميعها في وحدة القياس المستعملة.

وتعطى صيغة هذا المعامل كالاتي:-

$$C.V = \left(\frac{S}{\bar{x}}\right) 100\%$$

حيث ان :

$C.V$: معامل التغير او الاختلاف

s :الانحراف المعياري

$\bar{x}$:الوسط الحسابي

2.المتغير المعياري والدرجة المعيارية Standardized Variable and Standard Units-(Scores)

المتغير المعياري(Z) هو مقدار لاحجم له (بمعنى انه مستقل عن وحدات القياس المستخدمة للبيانات) والذي يقيس انحراف القيم عن الوسط الحسابي بوحدات من الانحراف المعياري ، ويستخدم في حال الاحتياج الى اجراء مقارنة بين قيم محددة لمجموعات مختلفة ويعطى بالصيغة التالية:

$$Z_i = \frac{\mathrm{x_i} - \bar{\mathrm{x}}}{\mathrm{s}}$$

حيث ان :

Z_i :المتغير المعياري

I=1, 2,.....,n

s :الانحراف المعياري

$\bar{x}$:الوسط الحسابي

مثال: في احدى الدورات التي تنظمها مؤسسة صحية معينة لمجموعة من منتسبيها حصل احد المشاركين في اختبار معين على درجة (84) حيث كان متوسط الدرجات لهذا الاختبار (76) وانحرافها المعياري (10) ، وفي اختبار اخر حصل على درجة (90) حيث كان متوسط الدرجات (82) وانحرافها المعياري (16) ، ففي اي الاختبارين كانت درجة استيعابه اعلى؟

الارتباط الخطي البسيط (بيرسون) Simple linear correlation

الارتباط : هو مقياس شدة العلاقة بين ظاهرتين او متغيرين (x , y) ان وجدت ، كالعلاقة بين الانفاق والدخل او العلاقة بين زيادة الانتاج ومهارة العاملين من خلال علاقة رياضية معينة. لما كانت العلاقة بين المتغيرين أقوى، اقتربت قيمة معامل بيرسون من +1 أو -1. الوصول إلى القيم +1 أو -1 يعني أن جميع نقاط البيانات تقع على خط مستقيم يمثل "أفضل خط ملاءمة"، أي أن التغير في أي متغير لا يضعف الارتباط مع المتغير الآخر

ملاحظات

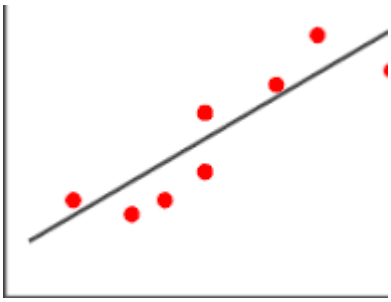
-على الرسم البياني يمكن ملاحظة العلاقة بين المتغيرين وتكوين افتراضات حتى قبل حسابها

-اذا كانت المخططات المبعثرة قريبة من الخط فهذا يشير الى علاقة قوية بين المتغيرات.

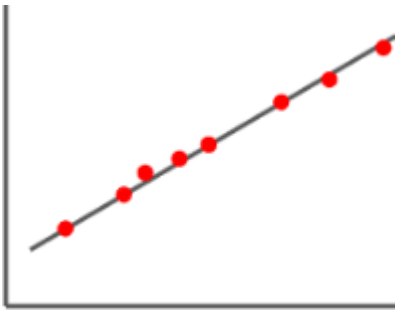
-كلما اقتربت النقاط من الخط، زادت قوة العلاقة، وكلما ابتعدت عنه، أصبحت العلاقة أضعف

أما إذا كان الخط قريبًا من أن يكون موازيًا لمحور**x** وكانت النقاط مبعثرة عشوائيًا فيمكن القول إن

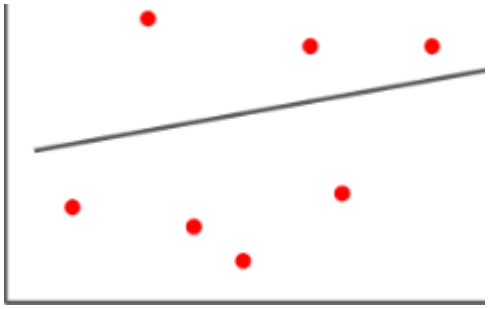
الارتباط شبه منعدم بين المتغيرين



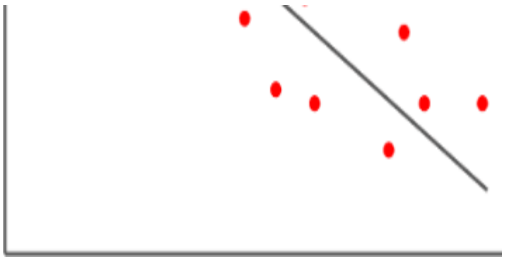
ارتباط موجب متوسط



ارتباط موجب قوي



عدم وجود ارتباط



ارتباط عكسي ضعيف

ويحسب بالصيغة التالية :-

$$r = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}$$

حيث ان :-

x_i: المتغير الاول

y_i: المتغير الثاني

ان قيمة (r) تقع بين (1- ،1+) فاذا كانت نتيجة المقياس (1+) فهذا يدل على ارتباط تام طردي واذا كان الاختلاف كاملا وعكسيا اصبحت قيمة المقياس(1-) واذا كانت النتيجة (0) فهذا يدل على انعدام الارتباط.

مثال:- اخذت البيانات التالية عم مصاريف الجيب الاسبوعية (x) ودرجات احدى المواد الدراسية (y) لمجتمع احصائي وكانت كالآتي

رقم الطالب	1	2	3	4	5	6	7	8
مصروف الطالب	2	4	10	8	2	13	5	12
الدرجات	97	90	60	55	92	44	82	40

المطلوب :

1. هل يمكن ان يكون التغير في احدى الظاهرتين يسبب تغير في الظاهرة الاخرى.
2. حساب معامل ارتباط بيرسون
3. تفسير شدة الارتباط وطبيعته؟

ارتباط الرتب Rank correlation

في كثير من الاحيان لايمكننا تطبيق معامل ارتباط بيرسون وذلك نظرا لطبيعة البيانات المتوفرة ، لذا نستخدم معاملات ارتباط الرتب ومنها:-

معامل ارتباط الرتب سبيرمان Spearman's rank correlation coefficient

اذا كانت احدى الظاهرتين او كلاهما لاتأخذان قيما عددية تقيس كل مشاهدة في هذه الحالة لانستطيع استخدام معامل ارتباط بيرسون،لذا نلجأ الى استخدام معامل ارتباط سبيرمان
ويمكن استخراج قيمة معامل ارتباط الرتب سبيرمان بالصيغة التالية:-

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

حيث ان :-
$-1 \le r_s \le 1$
di : الفرق بين كل ترتيبين متقابلين

مثال :- البيانات التالية تمثل سن المرأة عند الزواج والمستوى العلمي لمجموعة من النساء المطلوب :-

رقم المشاهدة i	1	2	3	4	5	6
المستوى التعليمي x	ثانوية	ابتدائية	امية	اعدادية	جامعية	عليا
سن الزواج y	22	17	16	17	26	32

خطوات الحل:-
1.نرتب القيم تصاعديا او تنازليا لكل من x , y من خلال اعطاء ارقام تسلسلية تسمى الرتب
2.ايجاد الفرق بين رتب الظاهرة الاولى ورتب الظاهرة الثانية ثم تربيع القيم المستخرجة.
3.تطبيق صيغة القانون.

المشاهدة	المستوى التعليمي xi	سن الزواج yi	r_{x_i}	r_{y_i}	$r_{x_i} - r_{y_i}$	d_i^2
1	ثانوية	22	4	4	0	0
2	ابتدائية	17	2	2.5	-0.5	0.25
3	امية	16	1	1	0	0
4	اعدادية	17	3	2.5	0.5	0.25
5	جامعية	26	5	5	0	0
6	دراسات عليا	32	6	6	0	0
∑						0.5

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

$$=1-\frac{6(0.5)}{6(36-1)} \qquad = 0.99$$

ان العلاقة بين مستوى تعليم الاناث والسن عند الزواج هي علاقة طردية (اي يتأخر زواج الفتاة مع استمرارها بالتعليم) وشديدة جدا (شبه تامة).

الانحدار الخطي البسيط : هو خط مستقيم يمر بمجموعة من النقاط بطريقة تجعل مجموع مربع النقط المتبقية من النموذج (أي، المسافات الرأسية بين النقطة المتبقية والخط) أقل ما يمكن. هذا يشير الي حقيقة أن الانحدار هو واحد من أبسط الأساليب المستخدمة في مجال الإحصاء حيث أن ميل الخط يساوي العلاقة بين y و x مصححة بنسبة الانحرافات المعيارية لهذه المتغيرات. نقطة تقاطع الخط مع محور الصادات هي مركز كتلة نقاط البيانات .(x, y) توجد طرق انحدار أخرى بجانب المربعات الصغرى البسيطة (انظر الانحدار الخطي). علي وجه الخصوص، عندما يريد شخص أن يقوم بفعل الانحدار عن طريق العين فانه يميل عادة الي رسم خط حاد قليلا ويكون قريبا من ذلك الذي ينتج من طريقة أقل مربعات كليه. يحدث هذا لأنه طبيعي أكثر لعقل الإنسان ملاحظة المسافات المتعامدة علي خط الانحدار بدلا من تلك الراسية كما يحدث في طريقة المربعات الصغرى

إذا كانت الظاهرة تزيد أو (تنقص) بمقدار ثابت كل فترة زمنية فإن معادلة الاتجاه العام تكون على صورة خط مستقيم أي أن معادلته هي :-

$$\widehat{Y} = \widehat{B_0} + \widehat{B_1} X$$

حيث -:

$\widehat{B_0}$ هو الجزء المقطوع من المحور الرأسي

$\widehat{B_1}$ هو ميل خط الاتجاه

$\widehat{Y}$ قيمة الظاهرة الاتجاهية

X : دليل الزمن (تبدأ بالواحد لأول فترة ثم اثنين للفترة الثانية وهكذا .....)

مثال :-البيانات التالية تمثل عدد العاملين (بالألف) في احدى الشركات العالمية

السنة	2016 (0)	2017 (1)	2018 (2)	2019 (3)	2020 (4)
عدد العاملين	7	8	10	11	13

والمطلوب:

1- ايجاد معادلة الانحدار الخطي

2- تقدير عدد العاملين عام 2026

الحل

ولايجاد معادلة الانحدار الخطي نحسب المجاميع التالية-:

خطوات الحل :-

1. ايجاد قيمة $\hat{\beta}_1$

السنوات XI	عدد العاملينYI	XiYi	Xi^2
0	**7**	**0**	**0**
1	**8**	**8**	**1**
2	**10**	**20**	**4**
3	**11**	**33**	**9**
4	**13**	**52**	**16**
10	**49**	**113**	**30**

$$\bar{X} = \frac{\sum XI}{N} = \frac{10}{5} = 2$$

$$\bar{Y} = \frac{\sum YI}{N} = \frac{49}{5} = 9.8$$

- ويكون حساب $\widehat{B_1}$ كما يلي:

$$\hat{\beta}_1 = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} = \frac{(5)(113) - (10)(49}{(5)(30) - (10)^2}$$

$$= \frac{75}{50} = 1.5$$

.2.ايجاد قيمة $\widehat{B_0}$ **كما يلي:**

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 9.8 - (1.5)(2) = 6.8$$

3. ايجاد معادلة الانحدار الخطي هي:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

$$\hat{y} = 6.8 + 1.5x$$

لتقدير عدد العاملين 2026 أي أن X= 10 اذن

$$\hat{y} = 6.8 + 1.5 * (10) = 21.8$$

أي أن عدد العاملين المتنبأ به عام 2026 هو 21800 عامل.

مثال (2)

البيانات التالية توضح ساعات العمل اليومية لـ 10 عمال في أحد المصانع وكذلك الإنتاج اليومي لكل منهم.

ساعات العمل xi	الإنتاج Yi
10	11
7	10
10	12
5	6
8	10
8	7
6	9
7	10
9	11
10	10

1- **بين أثر ساعات العمل علي الانتاج (قم بتقدير نموذج الانحدار الخطي).**

2- **قدر انتاج عامل عدد ساعات عمله 9 ساعات.**

العلاقة بين الارتباط والانحدار الخطي البسيط : من المقاييس الوصفية مقاييس النزعة المركزية، والتشتت، وغيرها من المقاييس ألاخرى والتي يمكن من خلالها وصف شكل توزيع البيانات التي تم جمعها عن متغير واحد، وننتقل من التعامل مع متغير واحد إلى التعامل مع متغيرين أو أكثر، لدراسة وتحليل العلاقة بين متغيرين، وذلك باستخدام بعض طرق

التحليل الاحصائي مثل تحليل الارتباط، والانحدار الخطي البسيط، فإذا كانت دراسة العلاقة بين متغيرين استخدم لذلك أسلوب تحليل الارتباط، وإذا كان المطلوب دراسة أثر أحد المتغيرين على الاخر استخدم لذلك أسلوب تحليل الانحدار، ومن الامثلة على ذلك- 1 :الانفاق، والدخل العائلي- 2 .سعر السلعة، والكمية المطلوبة منها- 3 .الفترة الزمنية لتخزين الخبز، وعمق طراوة الخبز- 4 .تقديرات الطالب في مادة الاحصاء، وتقديراتهم في مادة الرياضيات- 5 .كميات السماد المستخدمة، وكمية اإلنتاج من محصول معين تم تسميده بهذا النوع من السماد- 6 .عدد مرات ممارسة نوع معين من الرياضة البدنية، ومستوى الكلسترول في الدم- 7 .وزن الجسم، وضغط الدم .، والامثلة على ذلك في المجال التطبيقي كثيرة.

تكمن العلاقة بين معامل الارتباط ومعامل الانحدار في ان معامل الارتباط يقيس قوة واتجاه العلاقة الخطية بين متغيرين بينما معامل الانحدار يصف هذه العلاقة بمعادلة رياضية ، يظهر التشابه في ان اشارة كلا المعاملين تكون متطابقة ، فاذا كان معامل الارتباط موجبا فان معامل الانحدار يكون موجبا ايضا والعكس صحيح ، ويمكن استخدام معامل الارتباط مربعا (r^2) لتقييم دقة نموذج الانحدار.

العلاقة الرياضية

1. يمكن التعبيرعن معامل الارتباط(r) من خلال معاملي الانحدارباستخدام العلاقة التالية :

$$r = \sqrt{b_{xy}.\, b_{yx}}$$

2. يعرف معامل الانحدار(β_1) بأنه ميل خط الانحدار ويعبر عن مقدار التغير في المتغير التابع (Y) مقابل كل وحدة تغيرفي المتغير المستقل (x) .
3. يعرف معامل الارتباط (r) بانه مقياس لقوة واتجاه العلاقة الخطية بين المتغيرين وتتراوح قيمته بين +1 و -1 .