وزارة التعليم العالي والبحث العلمي

جامعة محمد خيضر بسكرة

كلية العلوم الإنسانية والاجتماعية

قسم العلوم الاجتماعية

شعبة علم الاجتماع

مطبوعة في مقياس:

تحليل ومعالجة البيانات الاجتماعية

موجهة لطلبة السنة الثالثة علم الاجتماع

من إعداد الأستاذة:

د/ شاوش اخوان جهيدة

السنة الجامعية: 2021-2022

المحتوى الرسمي المقرر للمقياس

السداسي الخامس:

وحدة التعليم المنهجية

المادة :تحليل و معالجة المعطيات الاجتماعية

الرصيد: 03

المعامل: 02

الحجم الساعي الأسبوعي: 1سا و30د (أعمال موجهة)

طريقة التقييم: متواصل + امتحان

أهداف التعليم

التعرف على كيفيات عرض وتحليل البيانات الكمية والكيفية

المعارف المسبقة المطلوبة

التحكم في تقنيات البحث السوسيولوجي و المقاييس الاحصائية

محتوى المادة:

1- تفريغ و عرض المعطيات

- تبويب و تفريغ المعطيات (التبويب ، الترميز، التفريغ المسطح .....)
- العرض الجدولي للمعطيات الجداول البسيطة ، الجداول المركبة .....)
- الرسومات البيانية

2- طرق تحليل البيانات

- التحليل الوصفي للمعطيات
- أساليب الإحصاء الاجتماعي
- المقاييس و المعاملات الإحصائية

3- المعالجة الآلية للمعطيات

- تطبيقات لبرامج الإعلام الآلي في عرض و تحليل المعطيات

طريقة التقييم :

المراجع: (*كتب ومطبوعات ، مواقع انترنت*، إلخ)

1- **Guibert Joël.**Méthodologie des pratiques de terrain en sciences humaines et sociales, Paris : A. Colin, 1997.

2- **Stéphane Beaud**. Guide de l'enquête de terrain**,** 4e édition augmentée. la découverte, paris,2010

3- **Frédéric Lebaron**L'enquête quantitative en sciences sociales : Recueil et analyse des données, Dunod, coll. « psycho sup », 2006

4- **Philippe Cibois**Les méthodes d'analyse d'enquêtes,.2007

5- **Renaud et G. Pini**. Introduction à l'analyse exploratoire des données avec SPSSSS

6- **Rafael Costa et G. MasuyStroobant**Pratique de l'analyse des donnéesSPSS appliqué à l'enquête « Identités et Capital social en WallonieUCL/IACCHOS/DEMO Louvain la Neuve 2013

7- **فضيل دليو**، تقنيات تحليل البيانات في العلوم الاجتماعية والاعلامية دار الثقافة للنشر والتوزيع، **2010**.

8- محمد خير أبو زيد ، أساليب التحليل الاحصائي باستخدام برمجيةspss ،دار جرير للنشر و التوزيع، 2005.

9- رضا عبدالله أبو سريع ،تحليل البيانات باستخدام برنامجspss،دار الفكر للنشر والتوزيع ،**2004**.

10- صلاح الدين حسين الهيثي ،الاساليب الاحصائية في العلوم الادارية/ تطبيقات باستخدامspss،دار وائللطباعة و النشر والتوزيع،2004.

11- ابراهيم الحكيم ، المرجع في تحليل البياناتspss، شعاع للنشر و العلوم، 2004

فهرس المحتويات:

عنوان المحور	عنوان الدرس	الصفحة
مقدمة	مقدمة	06
مدخل عام	الإحصاء وتحليل البيانات في الدراسات الاجتماعية	10
	مفاهيم أساسية	14
تفريغ وعرض البيانات	ترميز البيانات	23
	العرض الجدولي للبيانات	28
	العرض البياني	34
طرق تحليل البيانات	مقاييس النزعة المركزية	45
	مقاييس التشتت	58
	معاملات الارتباط	65
	اختبار الفرضيات	77
المعالجة الآلية للبيانات	المعالجة الآلية للبيانات باستخدام برنامج الحزمة الإحصائية spss	92
خاتمة	خاتمة	107
	قائمة لأهم المراجع	111

مقدمة

مقدمة:

تعد مرحلة تحليل البيانات من أهم مراحل البحث العلمي، وعليها تتوقف تفسيراته ونتائجه، وعادة ما يخصص في البحث فصل لتحليل البيانات؛ يبدأ هذا الفصل بتحديد نوعية البيانات التي يجب جمعها، وذلك بناءا على أهداف الباحث وتقديراته وطبيعة المجال الذي يعمل فيه، وفي هذه المرحلة يقوم الباحث باستعراض البيانات التي تم جمعها باستخدام أحد طرق جمع البيانات الوصفية أو الكمية .

وليست عملية جمع البيانات باليسيرة، فهي عملية تتطلب من الباحث العمل لوقت طويل، فعليه أن يقوم بجمع البيانات من مصادر مختلفة، ومن ثم ترتيبها بطريقة تسهل عليه استخراج المعلومات منها، بعد ذلك تأتي المرحلة الثالثة والتي بموجبها يعالج الباحث هذه البيانات، وبعد الانتهاء من معالجة البيانات يقوم الباحث بتنقيتها من الأخطاء، وذلك من خلال استخدام عدد من الوسائل الإحصائية المتاحة، وبعد ذلك تكون البيانات قد أصبحت جاهزة للعرض .

يقدم علم الإحصاء للباحثين في مجال العلوم الاجتماعية العديد من الطرق والأساليب اللازمة والضرورية للقيام بالدراسات والبحوث علي أساس من القياس لحركة العديد من المتغيرات المحددة للظواهر موضوع الدراسة. ولقد أصبح الاتجاه العام في البحوث والدراسات الاجتماعية هو استخدام طرق القياس الكمية ووسائل الإحصائية وذلك لتحديد الخصائص وإبراز الاتجاهات العامة في الظواهر الاجتماعية، وتحليل العلاقات المتشابكة والمتبادلة بين الظواهر. حيث أن الإحصاء لا غنى عنه لأي باحث في شتى المجالات المختلفة إذ اعتمد في بحثه على الأسلوب العلمي. أي أن الإحصاء هو عصا الباحث التي تقوده إلى الطريق الصحيح، وهى الأداة التي تساعده على تفسير الظواهر التي يدرسها وتوضيح النتائج التي يحصل عليها ودلالات البيانات والأرقام التي يحصل عليها.

وقد أدى التقدم المتسارع في تكنولوجيا المعلومات واستخدام الحاسبات الآلية إلى مساعدة الباحثين ومتخذي القرارات في الوصول إلى درجات عالية ومستويات متقدمة من

التحليل ووصف الواقع ومتابعته ثم إلى التنبؤ بالمستقبل . ومع التطور التكنولوجي الكبير في السنوات الماضية أصبحت مسألة جمع البيانات وتحليلها مسألة سهلة وسريعة جدا وقليلة الأخطاء، وذلك من خلال الاستعانة بالمعالجة الآلية للمعطيات من خلال مختلف تطبيقات برامج الإعلام الآلي المتنوعة المتاحة.

مدخل عام

الدرس 1: الإحصاء وتحليل البيانات في الدراسات الاجتماعية

تمهيد:

يهتم علم الإحصاء بالبحث في طرق جمع الحقائق الخاصة بالظواهر العلمية، التي تتمثل في حالات أو مشاهدات متعددة، وفى كيفية تسجيل هذه الحقائق في صورة قياسية رقمية، وتلخيصها بطريقة يسهل بها معرفة اتجاهات الظواهر وعلاقات بعضها ببعض، ويبحث أيضاً في دراسة هذه العلاقات والاتجاهات واستخدامها في فهم حقيقة الظواهر ومعرفة القوانين التي تسير تبعاً لها. وتستخدم كلمة الإحصاء لتشير إلي عملية جمع البيانات الكمية والأساليب المستعملة في معالجة تلك البيانات، كما نعني بهذه الكلمة أيضا عملية استخلاص بعض الاستنتاجات من دراسة عينة صغيرة ومن ثم صياغة تعميمات يمكن تطبيقها علي مجتمعات اكبر حجما.

أهمية علم الإحصاء

يعتبر علم الإحصاء احد الوسائل المهمة في البحث العلمي من خلال استخدام قواعده وقوانينه وطرقه في جمع البيانات و المعلومات اللازمة في البحث العلمي وتحليل هذه المعلومات لغرض الوصول إلى النتائج التي يهدف لها البحث كما وان للإحصاء دورا بارزا في وضع الخطط المستقبلية عن طريق التنبؤ بالنتائج لكافة القطاعات حيث يمكن تطبيق علم الإحصاء في مجالات العلوم الصرفة أو العلوم الإنسانية (مثل الزراعة ,الصناعة ,الطب, الهندسة, الإدارة والاقتصاد...الخ). ولهذا أصبحت البحوث والدراسات الاجتماعية الحديثة تتجه أكثر فأكثر نحو استخدام طرق القياس الكمية والوسائل الإحصائية، وذلك لتجنب الذاتية في تحديد خصائص الظواهر الاجتماعية وإبراز الاتجاهات العامة لها، وإضفاء الموضوعية في تحليل العلاقات المتشابكة والمتبادلة بين الظواهر، وذلك من أجل تحقيق الخاصية العلمية لعلم الاجتماع.

الطريقة الإحصائية في البحث العلمي

1- تحديد المشكلة ووضع الفرضيات.

2- جمع البيانات والمعلومات ذات العلاقة بالبحث.

3- تصنيف وتبويب وعرض البيانات.

4- حساب المؤشرات الإحصائية.

5- تحليل وتفسير النتائج على ضوء فرضية البحث .

6- تفسير النتائج واتخاذ القرار.

مصادر جمع البيانات:

1- مصادر أولية (أصلية): وهي البيانات التي يقوم الباحث بجمعها بنفسه من مصادر ميدانية وتشمل البيانات والمعلومات التي يمكن الحصول عليها من مصادرها الأساسية عن طريق المواجهة أو المراسلة أو بأي طريقة أخرى وتسمى البيانات المجموعة بهذه الطريقة بالبيانات الأولية.

2- مصادر ثانوية: وهي بيانات تم إعدادها مسبقا، أي يتم جمعها من دراسات سابقة أو كتب أو مجلات، وتشمل الوثائق والمطبوعات والنشرات الإحصائية التي تصدرها الهيئات الحكومية مثل تعداد السكان والصادرات والواردات...ومن عيوب هذه الطريقة عدم معرفة طريقة تجميعها والهدف الذي جمعت من أجله.

وهناك عدة طرق للقيام بجمع البيانات من مصادرها الأولية منها:

- الملاحظة (المشاهدة): كمعرفة حركة المرور في منطقة معينة وتسجيل البيانات منها.
- الاستبيان: بطرح أسئلة يتم الإجابة عليها على أن تكون تلك الأسئلة تتناول موضوع معين كمجانية التعليم أو طبيعة السكن.
- المقابلة: أواللقاء المباشر بين الباحث والمبحوثين شخصياً للحصول على البيانات المطلوبة مع ضرورة شرح المطلوب للمبحوث للحصول على أفضل الإجابات.

- الاتصال الهاتفي: في حال توفر الهاتف عند الفئة المستهدفة .

- واليوم تستخدم طرق أخرى كالبريد الإلكتروني أو نشر المطلوب عبر شبكة الإنترنت وطلب الإجابة عليه من قبل عينة من المجتمع أو الفئة المستهدفة موضوع البحث[1].

أساليب جمع البيانات

1- **أسلوب الحصر الشامل:** يتم جمع البيانات هنا عن طريق تسجيل جميع مفردات المجتمع، لذلك فان هذا الأسلوب يتطلب وقتا وجهدا، وينبغي عند استخدام هذا الأسلوب أن يكون المجتمع محدودا مثل تعداد السكان في منطقة ما أو حصر عدد العاملين في مؤسسة معينة...، ويمتاز هذا الأسلوب بالحصول على نتائج دقيقة لكونه يتعامل مع جميع بيانات مجتمع البحث.

2- **أسلوب العينات**: وهي عملية جمع البيانات والمعلومات عن مجموعة معينة من مفردات المجتمع مثل استفتاء بعض الأشخاص عن موضوع معين، وهذا الأسلوب يحتاج إلى وقت وجهد وموارد اقل مقارنة بالأسلوب الأول، بالإضافة إلى أنه مفيد في حالة دراسة المجتمعات الغير محدودة، ولكن دقة النتائج هنا تكون أقل كونه لا يتعامل مع جميع المفردات.

الأخطاء الشائعة في جمع البيانات

1- خطا التحيز: يحدث هذا الخطأ نتيجة جمع البيانات من مصادرها غير الرئيسية أو التحيز لمفردة معينة في المجتمع قيد الدراسة دون الأخرى لسبب أو لأخر.

2- خطا الصدفة: يحدث هذا الخطأ نتيجة اعتماد الباحث على معلوماته الشخصية أو جمع البيانات ناقصة .

[1]: وليد عبد الرحمان خالد الفرا: تحليل بيانات الاستبيان باستخدام البرنامج الإحصائي Spss، الندوة العالمية للشباب الإسلامي، 1430ه، ص5.

تصنيف البيانات

هي عملية تقسيم البيانات إلى مجاميع صغيرة أو أصناف على اساس قاعدة معينة كاشتراكها في بعض الصفات أو الخصائص كالمهنة أو الحالة الاجتماعية وحسب متطلبات البحث.

تبويب البيانات

وهي عملية إدخال البيانات المصنفة في جداول خاصة بهدف إبراز البيانات وتوضيحها في أضيق نطاق ممكن لتكوين فكرة عنها، ويختلف أسلوب تبويب البيانات تبعا لطبيعتها كما يأتي:

1- **التبويب الزمني:** ومعنى ذلك فرز البيانات إلى مجموعات على أساس وحدة زمنية معينة كالسنة أو الشهر.

2- **التبويب الجغرافي:** ومعنى ذلك فرز البيانات إلى مجموعات على أساس وحدة جغرافية معينة مثل محافظة معينة أو بلد معين.

3- **التبويب الكمي:** ومعنى ذلك فرز البيانات إلى مجموعات على أساس وحدات كمية أي أرقام كالطول والوزن....الخ.

4- **التبويب الوصفي:** ومعنى ذلك فرز البيانات إلى مجموعات على أساس صفة معينة مشتركة مثل الحالة الاجتماعية أو لون العينين ....الخ.

الدرس 02: مفاهيم أساسية

1- تعريف علم الإحصاء:

هو العلم الذي يختص بجمع وتصنيف وتبويب البيانات وعرض وتحليل البيانات واستخلاص النتائج منها. ويقسم الإحصاء بصورة عامة إلى قسمين رئيسين هما

- **الإحصاء الوصفي Descriptive Statistic:** يتضمن طرق جمع البيانات وتصنيفها وعرضها في جداول ورسومات بيانية وحساب بعض المؤشرات الإحصائية.
- **الإحصاء الاستدلالي Inference Statistic:** يتضمن طرق اختبار الفرضيات وتقدير المعالم

2- المجتمع الاحصائي والعينة الاحصائية

- **المجتمع الاحصائي**: هو مجموعة من العناصر او المفردات التي تخص ظاهرة معينة، ويمكن ان تكون عناصر المجتمع الاحصائي عبارة أفراد أو عائلات أو جماعات أو منظمات...الخ . ويتمثل المجتمع الاحصائي بعدد العناصر او المفردات التي يتضمنها ويرمز له بالرمز N.
- **العينة الاحصائية**: تعرف العينة بانها جزء من مفردات المجتمع الاحصائي يتم اختيارها بطريقة علمية دقيقة، وتحمل نفس خصائص المجتمع الاحصائي، وتستخدم لأغراض الدراسة في حال تعذر اجراء الدراسة على المجتمع عندما يكون حجمه كبيرا .

3- البارامترات (المَعلمات) والإحصاءات :

للمجتمع خصائص متعددة مثل المتوسط والوسيط والانحراف المعياري وكذلك لكل عينة تسحب من هذا المجتمع خصائصها أيضاً وما يتعلق بخصائص المجتمع يسما مَعلَّم أو بارامتر Parameter بينما كل ما يتعلق بخصائص العينات يسمى إحصاءه Statistic ويمكن الاستفادة من إحصاءات العينة تقدير معلمات المجتمع.

4- الإحصاء البارامتري والإحصاء اللابارامتري :

تصنف الأساليب الإحصائية الاستدلالية إلى أساليب بارامترية وأساليب لا بارامبرية:

- **والأساليب البارامترية:** ويطلق عليه البعض الطرق المعلمية هي الأساليب التي تتطلب استيفاء افتراضات معينة حول المجتمع الذي تسحب منه عينة البحث ومن هذه الافتراضات أن يكون التوزيع طبيعياً وأن يكون هناك تجانس في التباين.
- **أما الأساليب اللابارامترية:** والتي يطلق عليه البعض الطرق اللامعلمية فهي الأساليب التي تستخدم في الحالات التي لا يكون فيها نوع التوزيع الاحتمالي للأصل الذي سحبت منه العينة معروفاً، أو في حالة عدم استيفاء شرط التوزيع الاعتدالي للمجتمع.

وإذا كانت الأساليب البارامترية تصلح للبيانات في المستوى الفتري والمستوى النسبي فإن الأساليب الإحصائية اللابارامترية تصلح في حالة البيانات الرتبية والاسمية.

5- المتغيرات : Variables

يقصد بالمتغير أي خاصية يمكن قياسها وتتباين قيمها من فرد إلى أخر أو من مجموعة إلى أخرى، فالبيانات الإحصائية التي يتعامل معها الباحث أو يقوم بجمعها ما هي إلا درجات

أو مؤشرات لمقدار الشيء أو الصفة أو الخاصية موضوع القياس لدى الفرد، وعليه عندما نهتم بتحديد نوع الفرد (ذكر أو أنثى) فإننا نكون بصدد متغير النوع أو الجنس وعندما نهتم بتحديد درجة ذكاء الفرد نكون بصدد متغير الذكاء وعندما نهتم بتحديد درجة التحصيل الدراسي عند الفرد نكون بصدد متغير التحصيل الدراسي.

يرمز للمتغيرات عادة بالأحرف الكبيرة مثل X,Y,Z وتقسم الى قسمين هما:

1- **المتغيرات الوصفية:** وهي المتغيرات التي تكون مفرداتها غير قابلة للقياس بالأرقام بل تكون على شكل صفات مثل لون العينين أو تقديرات الطلاب .

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

ا) **المتغيرات المتقطعة :** وهي المفردات التي تأخذ قيما متميزة عن بعضها وتكون قابلة للعد والحساب ولا تتقبل الكسور مثل عدد الطلاب في كلية معينة أو عدد العائلات في منطقة معينة.

ب) **المتغيرات المستمرة:** وهي المفردات التي تأخذ مدى معين أو مجال معين وتكون غير قابلة للعد والحساب و تتقبل الكسور مثل أعمار الأفراد أو أطوالهم أو درجاتهم.

6- <u>مستويات القياس:</u>

أ- القياس الاسمي (التصنيفي):

يعد هذا المستوى أدنى مستويات القياس وأضعفها بل أن تسميته مقياساً تعد أحياناً تسمية مجازية وهنا تستخدم الأعداد بغرض التصنيف فقط لا غير مثال : تحديد العدد 1 ليدل على أن جنس المفحوص ذكر والعدد 2 ليدل على أن الجنس أنثى، فهنا لا يعني ذلك أن 2 أكبر من 1 وإنما الغرض تصنيفي بحت، وهكذا فإن تحديد بعض الأعداد لتدل على التخصص ولون العينين والجنسية كلها من نوع المقاييس الاسمية أو التصنيفية، وبالتالي **فالقياس الاسمي** أو **التصنيفي** يعنى بتصنيف الأفراد في الظاهرة ولكن لا يوضح ترتيبهم في هذه الظاهرة.

ب- **القياس الترتيبي (الرتبي) :**

وهنا الهدف من الأعداد هو ترتيب الأفراد في ظاهرة معينة، ولكن يجب ملاحظة أن الفروق بين الرتب ليست متساوية، أي أن المسافات البينية بين الأفراد غير متساوية، فمثلاً إذا حددنا الأعداد من واحد إلى خمسة لتدل على ترتيب الأفراد من حيث المستوى الدراسي فلا يكون الفرق بين الشخص الأول (الأقل مستوى دراسيا) والشخص الموالي له هو نفس الفرق بين الشخص الخامس (الأكثر مستوى دراسيا) والسابق له. ويجب ملاحظة أن القياس الترتيبي قد يحمل معنى التساوي أحياناً ، **فالقياس الترتيبي** يعطي فكرة عن ترتيب الأفراد في ظاهرة معينة ولكن لا يعطى فكرة عن الفروق في الظاهرة بين الأفراد.

ج- **القياس الفتري أو الفئوي :**

وهنا تتساوى الفروق أو المسافات بين المستويات المتتالية مثل درجات الأطفال في اختبار الذكاء أو درجاتهم في التحصيل، فهنا يكون هناك وحدة قياس ثابتة متفق عليها يقاس بها الفرق بين كل درجة والتالية لها بحيث يصبح الفرق مثلاً بين 4 ، 5 مساوي للفرق بين 10 ، 11 أو المسافة بين 10 ، 15 مساوي للفرق بين 20 ، 25 ولا يعنى ذلك وجود صفر مطلق يعنى غياب الصفة وإنما البداية أو الصفر هنا صفر اختياري أو نسبي وليس صفراً مطلقاً ويجوز إجراء العمليات الحسابية لتقليدية كالجمع والطرح.

د- **المستوى النسبي :**

وهنا للدرجات صفر مطلق يعنى غياب الصفة مثل مقياس الوزن ودرجة الحرارة ويسمى هذا المستوى بالمستوى النسبي لأن النسبة بين أي درجتين لا تتأثر بوحدة القياس فمثلاً النسبة بين واحد كيلو جرام و10 كيلو جرام هي نفسها النسبة بين 1000 جرام و 10.0000جرام وهنا الأعداد المستخدمة أعداد حقيقة لها صفر مطلق.

ويمكن اختصار ذلك في الجدول التالي:

مستوى القياس	الخصائص	مثال
الاسمي	**1) ابسط وأدنى مستويات القياس** **2) يصنف المبحوثين إلى مجموعات متمايزة طبقاً لخصائص نوعية** **3) لا يسعى لتسجيل أفضلية، لأن هذه الأفضلية منتفية أصلا.**	**- متغير الجنس(ذكر، أنثى)** **- متغير مكان الإقامة (ريف،حضر، بدو)**
الترتيبي	**1) أعلى مرتبة من القياس الاسمي حيث يمكن باستخدامه ترتيب المفردات ترتيب تصاعدياً أو تنازلياً حسب درجة امتلاكها لخاصية معينة.** **2) ترتيب المفردات يتضمن الأفضلية.** **3) لا توجد وحدة قياس في هذا المستوى من مستويات القياس.**	**- متغير المستوى المعيشي (مرتفع، متوسط، متدني)** **- متغير المستوى التعليمي (أمي، ابتدائي، متوسط، ثانوي، جامعي، تعليم عالي)**
الفتري أو الفئوي	**1) أرقى من القياس الترتيبي حيث تحمل الأرقام هنا كميا ويكون الحصول على وحدة قياس بالتالي متاحا هذا بالإضافة إلى سمتي التصنيف والترتيب** **2) الصفر في هذا القياس لا يعني انعدام الصفة وإنما صفر نسبي وليس مطلقاً.**	**- علامات الطلاب في مادة الإحصاء (0، 5، 15، 20) الطالب الحاصل على صفر لا يعني عدم امتلاكه أي معلومة في المادة.**
النسبي	**1) أرقى مستويات القياس الثلاثة السابقة، حيث يتفوق على مستوى القياس الفتري بأنه يمتلك سمة "الصفر المطلق" الذي يدل على انعدام الخاصية أو السمة، بالإضافة إلى اشتماله لجميع سمات القياس**	**- عدد التلاميذ في الصف** **- عدد الأبناء في الأسرة**

7- الترميز Coding:

- **الترميز العددي:** ويقصد به استخدام الأرقام بصورة متتالية لتمييز مفردات البيانات؛ فمثلا يستخدم الرقم 1 للذكور والرقم 2 للإناث.
- **الترميز الحرفي:** وذلك باستخدام الحروف كأن نرمز مثلا بالحرف M للذكور والحرف F للإناث.
- **الترميز الرقمي والحرفي:** مثلا الأرقام للجنس (1=ذكر، 2=أنثى) و الحروف للمستوى الدراسي (L=ليسانس ، M= ماستر، D= دكتوراه)، فيكون: D1= ذكر ذو مستوى دكتوراه، M2= أنثى ذات مستوى ماستر، L2= أنثى ذات مستوى ليسانس وهكذا.

8- عرض البيانات:

ويمكن ذلك بالطرق التالية[2]:

- **جدولة البيانات Tabulation:**

ويتم في هذه الطريقة وضع البيانات في جداول إحصائية يختلف شكلها حسب نوع البيانات (وصفية أو كمية متقطعة أو مستمرة).

- **تمثيل البيانات بيانيا Graphical Representation:**

وهو التعبير عن البيانات برسوم بيانية تتلائم مع طبيعتها ، مثل:

- الأعمد البيانية (Bar Chart) والرسوم الدائرية (Pie Chart): تستخدم في حالة البيانات التي لها وحدة قياس اسمية أو ترتيبية ، بشرط أن تكون تقسيمات المتغير ليست كبيرة (أقل من عشر تقسيمات).
- المدرج التكراري (Histogram) والمضلع التكراري (Polygram) والمنحنى التكراري(Frequency Curve) تستخدم في حالة البيانات المستمرة (وحدة قياسها

[2]: وليد عبد الرحمن الفرا: مرجع سبق ذكره، ص 15.

فترة أو نسبة) الموضوعة في الجداول التكرارية بعد الترميز، ويستخدم المدرج التكراري أيضًا في حالة البيانات الكمية المتقطعة.

- رسم الصندوق (Box Plot) : ويستخد للبيانات المستمرة التي تعتمد على الوسيط والربيعين.
- رسم الساق والأوراق (Stem and Leaf) :يستخدم لتمثيل البيانات الكمية (متقطعة أو مستمرة)

9- حساب المقاييس الإحصائية (Statisticals Measures) :

توجد عدة مقاييس إحصائية منها

أ- مقاييس النزعة المركزية Measures of Central Tendency:

ويعرف مقياس النزعة المركزية للبيانات بأنه العدد الذي تتمركز حوله البيانات . وتوجد عدة مقاييس للنزعة المركزية وهي:

- **المتوسط الحسابي (Average or Mean) :** ويصلح للبيانات الكمية فقط وهو وحيد ويتأثر بالقيم الشاذة.
- **والوسيط (Median) :** وهو القيمة التي تقع في منتصف البيانات بعد الترتيب التنازلي أوالتصاعدي وهو وحيد ولا يتأثر بالقيم الشاذة.
- **المنوال (Mode) :** وهو القيمة الأكثر شيوعا، ولا يتأثر بالقيم الشاذة، لكنه قد لا يكون وحيدا وقد لا يكون موجودا.
- وتعتبر مقاييس النزعة المركزية كلها مقاييس مطلقة؛ أي لها نفس تمييز البيانات الأصلية، وبذلك لا تصلح للمقارنة بين مجموعتين أو أكثر إذا اختلفت وحدة المقياس.

ب- مقاييس التشتت (Variance measurement):

وهي تقيس مدى البعد أو التشتت بين مفردات المتغير عن مقياس النزعة المركزية الخاص بها، وتوجد مقاييس مطلقة ومقاييس نسبية منها:

- **المدى (Range):** وهو البعد بين أكبر وأصغر قيمة، وهو يتأثر بالقيم الشاذة.
- **التباين (Variance):** يقيس تشتت البيانات عن الوسيط
- **الانحراف المعياري (Standard deviation):** هو الجذر التربيعي الموجب للتباين.
- **معامل الاختلاف (Coefficient of variation):** وهو من أحسن المقاييس النسبية للتشتت، ويستخدم أحيانا للتعرف على القيم الشاذة للبيانات.
- **مقياس الالتواء (Measure of skewness):** وهو مقياس تشتت نسبي يحدد هل البيانات متماثلة أم ملتوية، ويكون ناحية اليمين إذا كان الإلتواء موجبا، ويكون ناحية اليسار إذا كان الإلتواء سالبا، ويكون متماثلا إذا كانت قيمة الالتواء صفرا.
- **مقياس التفرطح (Measure of kurtosis):** وهو مقياس نسبي يقيس قمة المنحنى ويأخذ الشكل المدبب عندما تكون معظم القيم بالقرب من الوسط الحسابي والذيلين، ويأخذ الشكل المفرطح عندما تكون معظم القيم بعيدة عن الوسط والذيلين، ويكون متوسط التفرطح (معتدل) عندما يكون معامل التفرطح مساويا للصفر.

المحور الأول:

تفريغ وعرض المعطيات

الدرس 3: ترميز البيانات

تمهيد:

قبل الشروع في التعامل مع محرر البيانات في برنامج Spss ينبغي القيام بترميز البيانات وإعدادها للإدخال في هذا البرنامج، وهو خطوة جد هامة من خطوات تحليل البيانات، ويتم من خلالها تهيئة البيانات سواء كانت استمارة أو مقابلة أو ملاحظة... أوغيرها، وذلك لكي يتمكن البرنامج من فهمها والتعامل معها.

ترميز البيانات Coding data:

ويقصد به عملية تحويل بيانات كل سؤال أو فقرة من أسئلة أو فقرات أداة جمع البيانات إلى أرقام أو أحرف ليسهل إدخالها إلى الحاسب الآلي، وغالبا ما نستخدم الأرقام لتمثيل البيانات، وذلك بأن يعطى كل متغير ترميزا معينا يعني مؤشرا معينا للبرنامج. ويجب التفريق هنا بين أنواع المتغيرات المختلفة. ويسري الترميز على كل الأداة بحيث تصحح جميع الاستمارات المراد إدخالها مثلا وترقم حسب أفراد العينة. وحسب برنامج Spss فإن الأشخاص الذين قاموا بالإجابة على أسئلة الاستمارة يطلق عليهم اسم الحالات Cases (الأفراد) وكل سؤال في الاستمارة هو بمثابة متغيرة Variable، وتشكل الإجابات قيم المتغيرة Valeur de variable.

ويفضل أولا استخدام أوراق الترميز لإدخال البيانات من الاستمارات، ثم القيام بعد ذلك بإدخال البيانات إلى الحاسب لإزالة الإرباك خاصة إذا كان عدد الأسئلة كبيرا.

دليل الترميز:

هو البيان العام الذي يتضمن اسئلة استمارة المعلومات وفقراتها ومتغيراتها في صورة اختصارات حرفية أو رقمية ترمز إلى وصف المتغيرات أو الأسئلة في صياغتها التامة

المصاغة بها. وظيفته مساعدة الباحث على تذكر السؤال أو المتغيرات بعد اختصارها حتى لا يقع في أخطاء التحليل، ويرجع إليه الباحث لتفسير الاختصارات التي ترمز إلى البيانات بعد إدخالها إلى الحاسب الآلي.

طريقة ترميز الاستمارة:

تختلف طريقة الترميز حسب نوع السؤال، وفيما يلي بعض أنواع الأسئلة الأكثر استخداما في البحوث الاجتماعية، وكيفية التعامل معها لترميزها يدويا أوعلى برنامج Spss[3]:

1- **سؤال مغلق يسمح باختيار إجابة واحدة:**

مثل:- الجنس: ذكر ☐ أنثى ☐

عندما يكون لدينا سؤال يسمح فيه باختيار إجابة واحدة فقط فإنه يكفي صياغته في متغيرة واحدة، وفي هذا المثال لدينا متغير الجنس مكون من ذكور وإناث، يمكن صياغته على برنامج Spss بإضافة متغيرة جديدة باسم الجنس ونقوم بترميز الذكور بالرقم 1 والإناث بالرقم 2.

2- **سؤال مفتوح عن متغير كمي:**

مثل:- السن:.....

تتطلب الإجابة على هذا السؤال تحديد عمر الشخص وفي هذه الحالة تعتبر المتغيرة كمية متصلة ولا تحتاج إلى أي ترميز. وفي بعض الأحيان يمكن تحويلها إلى فئات وهنا ترمز كل فئة بمركزها؛ مثلا الفئة [0- 10[ترمز بمركز هذه الفئة وهو 5.

3- **سؤال يتضمن أكثر من متغير كمي:**

مثل: عدد الإخوة: ..... الذكور: ..... الإناث: .....

في هذا السؤال يجب معرفة عدد الإخوة الذكور وعدد الإخوة الإناث، ويشكلان معا متغيرتين كميتين، وفي نفس السؤال مطلوب معرفة المجموع الذي يشكل متغيرا

---

[3]: جمال شعوان: **مدخل لدراسة وتحليل البيانات الإحصائية تطبيقات على برنامج Spss**، (pdf)، 2014، ص. ص12 .16 .

ثالثا ضمن هذا السؤال. ولربح الوقت يمكن حسابه من خلال جمع عدد الإخوة الذكور والإناث وذلك للاختصار ولربح الوقت في إدخال البيانات.

4- **سؤال مفتوح لمتغير اسمي**:

مثل:- المهنة: .....

يصاغ السؤال في متغيرة واحدة اسمية، ولترميز هذا السؤال يمكن تحديد بعض المهن وترميزها بأرقام، ولاختيار أي المهن سيتم الاعتماد عليها فإنه يجب الأخذ بعين الاعتبار عدة عوامل منها:

المهن الأكثر تكرارا في الاستمارة

الهدف من الاستمارة

الدقة المطلوبة

5- **سؤال منطقي يسمح باختيار إجابة واحدة فقط:**

مثل: هل ترغب بمتابعة الدراسة بالجامعة: نعم ☐ لا ☐

يعتبر هذا النوع من الأسلة شبيها بالنوع الأول إلا أن الأمر يختلف في طبيعة المتغيرة، حيث أن الأسلة المنطقية التي تتطلب الإجابة بنعم أو لا غالبا ما يتم دراستها من الناحية الكمية وبالتالي فهي تختلف عن متغيرة الجنس من حيث طبيعة القياس، ونستخدم قيمة 1 و0 في ترميزها حيث 1= نعم و0= لا، وتأخذ وحدة القياس صيغة كمية Scale.

6- **سؤال رتبي على شكل جدول:**

مثال: أماكن الاتصال بالأنترنت

مكان الاتصال	صعب	نادر	دائم
المنزل			
نادي الأنترنت			
مكان العمل			

المطلوب في هذه السؤال تحديد درجة الاتصال بالأنترنت من الأماكن الثلاثة المذكورة، وبالتالي سيتم التعامل مع هذا السؤال على أساس أنه ثلاث أسئلة، وستحدد ثلاث متغيرات هي: (الاتصال من المنزل- الاتصال من نادي الأنترنت- الاتصال من مكان العمل).

وكل متغيرة ستأخذ ثلاث قيم رتبية يجب ترميزها بأرقام يراعى فيها احترام التراتبية، مثلا (1=صعب، 2= نادر، 3= دائم).

ويجب تحديد المقياس الرتبي للمتغيرات.

7- **سؤال رتبي:**

مثل: - هل توافق أن يكون تسجيل الطالب في الجامعة عن طريق الأنترنت فقط:

موافق بشدة ☐ موافق ☐ محايد ☐ معارض ☐ معارض بشدة ☐

هذا السؤال يشبه السؤال السابق؛ فهو عبارة عن سؤال واحد يقبل جوابا واحدا، ويمكن ترميزه على شكل متغيرة واحدة ذات أجوبة اختيارية رتبية، مثلا (1= معارض بشدة، 2= معارض، 3= محايد، 4= موافق، 5= موافق بشدة)

8- **سؤال مغلق يسمح بأكثر من إجابة واحدة:**

مثل:- ماهي أهم الهوايات التي تمارسها:

القراءة ☐ الرياضة ☐ السفر ☐ أخرى ☐

في هذا السؤال نلاحظ أن الشخص يمكن أن يجيب على أكثر من اختيار واحد، لذلك فإن متغيرا واحدا لا يكفي لتمثيل السؤال. في هذه الحالة يفضل إنشاء أربعة متغيرات، كل متغير له احتمال إجابتين (نعم / لا)، حيث (1= نعم ، 0 =لا)

9- **سؤال كمي مغلق عبارة عن فئات يقبل إجابة واحدة:**

مثل: -الدخل الشهري:

أقل من 50.000 ☐ 100.000-50.000 ☐ 200.000-100.000 ☐

أكثر من 200.000 ☐

المتغيرة المدروسة في هذه الحالة هي متغيرة الدخل الشهري وهي متغيرة كمية صيغت على شكل فئات، ولترميز الفئات نستخدم مركز الفئة.

مركز الفئة هو المتوسط الحسابي لطرفي الفئة، بمعنى مجموع الطرفين مقسوما على 2، فمثلا مركز الفئة 100.000-200.000 هو 15.000.

بالنسبة للفئة التي تتضمن عبارة "أقل من" فإن مركز الفئة يساوي الحد الأعلى للفئة ناقص نصف طول الفئة التي بعدها، أما بالنسبة للفئة "أكثر من" فإن مركزها يساوي الحد الأدنى للفئة زائد نصف طول الفئة التي قبلها.

10- **أسئلة تحديد الأهمية:**

ويطلب فيها من المبحوث وضع ترتيب للإجابات حسب أهميتها مثل:

-حدد أي رابط من الروابط التالية تفضل الانتماء إليها قبل غيرها (رتبها حسب الأهمية):

الروابط	الأول	الثاني	الثالث	الرابع	الخامس
الأسرة					
القبيلة					
القرية					
المدينة					
الوطن					
الوطن العربي					

يتم تحويل كل فقرة من فقرات السؤال إلى متغير مستقل ثم يعرف كل متغير على حدة، وتحدد قيمته ويتم وصفها وترميزها وتصبح على النحو التالي:

الأسرة: س1 س2 س3 س4 س5 س6

القبيلة: ق1 ق2 ق3 ق4 ق5 ق6

وهكذا

الدرس 4: العرض الجدولي للبيانات

تمهيد:

بعد التأصيل النظري للدراسة الاجتماعية يقوم الباحث بالنزول إلى الميدان وجمع البيانات حول موضوع بحثه من مختلف المصادر الممكنة والمتاحة، هذه البيانات تحتاج إلى تصنيف وتبويب وتنظيم من أجل تسهيل دراستها وتحليلها والتمكن من تفسيرها للوصول إلى نتائج دقيقة وموضوعية.

وهناك عدة طرق لعرض البيانات وتنظيمها يمكن تقسيمها إلى قسمين أساسيين هما: الجداول الإحصائية والرسوم البيانية، ولكل من القسمين أشكال عديدة منها ما سيتم توضيحه في هذا الدرس.

أنواع الجداول الإحصائية:

هناك العديد من الصيغ التي يمكن أن تكون عليها الجداول، ويمكن عموما تقسيمها إلى نمطين أساسيين:

أولا: الجداول البسيطة:

وهي جداول تمثل متغيرا واحدا، وهي تتكون عادة من ثلاثة أعمدة تمثل:

الوحدات i - البيانات Xi - والتكرارات Fi .

مثال: المعلومات الأساسية لجدول البيانات البسيطة

جدول رقم(01): البيانات التالية تمثل الحالة العائلية لعينة مكونة من 50 مفردة (n=50)

الوحدات i	البياناتXi	التكراراتFi
1	أعزب	19
2	متزوج	27
3	مطلق	1
4	أرمل	3
-	المجموع Σ	50

مجموع التكرارات يكون دائما مساويا لحجم العينة؛ أي أن: $\Sigma Fi=n$

1- جداول البيانات المبوبة في فئات:

تستخدم في حالة العينات الكبيرة ($n>30$) بحيث تقسم البيانات إلى فئات متساوية في الطول (بقدر الإمكان) تكون كافية لاحتواء جميع البيانات.

مثال: الأعمدة الأساسية لجدول البيانات المبوبة في فئات:

جدول رقم (02): البيانات التالية تمثل الفئات العمرية لعينة مكونة من 50 مفردة

الوحدات i	الفئات	التكراراتFi
1	]35-30]	6
2	]40-35]	4
3	]45-40]	17
4	]50-45]	12
5	]55-50]	6
6	[60-55]	5
-	المجموع Σ	50

كيفية تحديد عدد الفئات في جداول التوزيع التكراري:

يمكن حساب عدد الفئات المناسب لبيانات ما وفق القاعدة التالية:

$$K=1+(3,3 \times \log n)$$

حيث: k: عدد الفئات.

1 و 3,3: ثوابت.

log n : اللوغاريتم العشري لحجم العينة.

كيفية تحديد طول الفئات:

أما طول الفئات فيحسب بقسمة المدى العام على عدد الفئات، وذلك وفقا للمعادلة التالية:

$$\Delta = \frac{H - L}{K}$$

حيث أن:

Δ: طول الفئات

H: الحد الأعلى للتوزيع

L: الحد الأدنى للتوزيع

K: عدد الفئات

كيفية حساب مركز الفئة (Ci):

مركز الفئة هو مجموع حدي الفئة مقسوما على اثنين، ويحسب كما يلي:

$$Ci = \frac{Hi + Li}{2}$$

أنواع التوزيعات التكرارية:

- **التكرار النسبي (Pi):**

التكرار النسبي لأية فئة هو حاصل قسمة تكرار هذه الفئة على مجموع التكرارات، و يحسب كما يلي:

$$Pi = \frac{Fi}{\Sigma Fi}$$

مجموع التكرارات النسبية دائما يساوي الواحد (01)

- **التكرار النسبي المئوي (Pi%):**

التكرار النسبي المئوي لأية فئة هو حاصل قسمة تكرار هذه الفئة على مجموع التكرارات والناتج مضروبا في مئة، و يحسب كما يلي:

$$Pi\% = \frac{Fi \times 100}{\Sigma Fi}$$

مجموع التكرارات النسبية المئوية دائما يساوي مئة (100)

- **التكرار المتجمع الصاعد:**

التكرار المتجمع الصاعد لأي فئة هو مجموع تكرارات الفئات السابقة لها مع تكرارها.

- **التكرار المتجمع النازل:**

التكرار المتجمع النازل لأي فئة هو حاصل طرح مجموع تكرارات الفئات السابقة لهذه الفئة من المجموع الكلي للتكرارات.

ثانيا: الجداول المركبة:

وهي عبارة عن جداول تتضمن متغيرين أو أكثر وتستخدم عادة لمقارنة البيانات أو لحساب معاملات الارتباط أو للاختصار

مثال:

جدول رقم (03): الجدول التالي يبين توزيع مفردات عينة بحجم 50 مفردة وذلك حسب متغيري الجنس وممارسة الرياضة:

X / y	دائما	أحيانا	أبدا	المجموع
ذكور	15	5	2	22
إناث	06	12	10	28
المجموع	21	17	12	50

مثال2: الجدول رقم (04) تقييمات أستاذين لعينة من البحوث

الوحدات	المنتج 1	المنتج 2
1	جيد	متوسط
2	جيد	جيد
3	متوسط	متوسط
4	ممتاز	جيد جدا
5	جيد	جيد
6	متوسط	ضعيف
7	ممتاز	جيد جدا
8	ضعيف	ضعيف
9	متوسط	ضعيف
10	جيد جدا	جيد

الجداول المغلقة والجداول المفتوحة:

- الجداول المغلقة: تكون نهاياتها محدودة من الجهتين
- الجداول المفتوحة: تكون إحدى نهاياتها أو كلاهما غير محدودة، فيكون لدينا:
  - الجداول المفتوحة من الأعلى
  - الجداول المفتوحة من الأسفل
  - الجداول المفتوحة من الأعلى والأسفل

الجداول المنتظمة والجداول غير المنتظمة:

- **الجداول المنتظمة:** تكون أطوال جميع الفئات متساوية
- **الجداول غير المنتظمة:** تكون أطوال الفئات متغيرة وغير متساوية. وهنا نلجأ إلى تعديل الفئات كما يلي:

التكرار المعدل= تكرار الفئة/طول الفئة

الدرس 5: الرسومات البيانية

تمهيد:

يعتبر العرض البياني أفضل من العرض الجدولي في عدة نقاط ؛ حيث يبرز بوضوح المعالم الرئيسية للسلاسل الإحصائية من اتجاه تغير، المقارنة، إيجاد أرقام أو معدلات بيانية، تطور الظاهرة، التنبؤ بقيم غير موجودة في الجداول، براز أسباب التغيرات بصفة عامة للظاهرة المدروسة...

وهناك العديد من التمثيلات البيانية التي يمكن استخدامها في عرض البيانات ويستحسن إرفاقها بالجداول البيانية، منها المنحنيات الرسومات البيانية التالية:

1- الاشرطة البيانية Bar- Charts:

وهي عبارة عن مجموعة من المستطيلات الافقية او العمودية قواعدها متساوية وتمثل الصفة التي تم على اساسها التبويب (سنة ,محافظة , وهكذا) وارتفاعاتها تمثل البيانات المقابلة لتلك الصفة، إذن يمكن اعتبار عدة أشرطة على محور الفواصل، لها نفس العرض وطولها متناسب مع التكرارات على محور التراتيب. المسافة بين الأشرطة ليس لها أي دلالة.والاشرطة البيانية على نوعين هما:

أ- الاشرطة البيانية المفردة

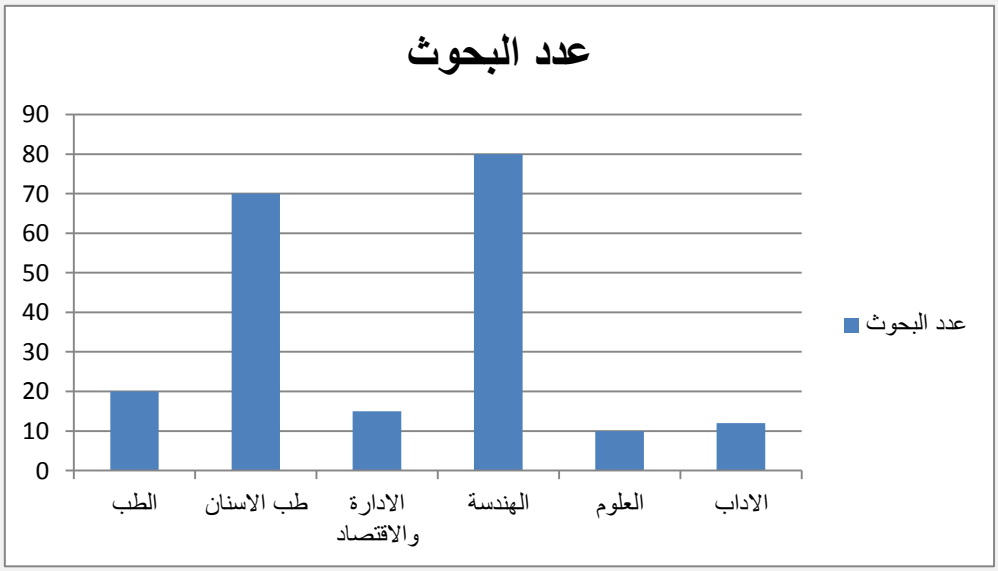
وهي اشرطة بيانية تخص صنف واحد للبيانات مثل عدد الطلبة المقبولين او تطور عدد سكان بلد ما حسب التعداد السكاني..

مثال:

البيانات الواردة في الجدول التالي تمثل عدد البحوث العلمية المنجزة من قبل أساتذة في جامعة معينة موزعين حسب كلياتهم خلال سنة 2020:

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

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



ب- **الاشرطة البيانية المركبة**:

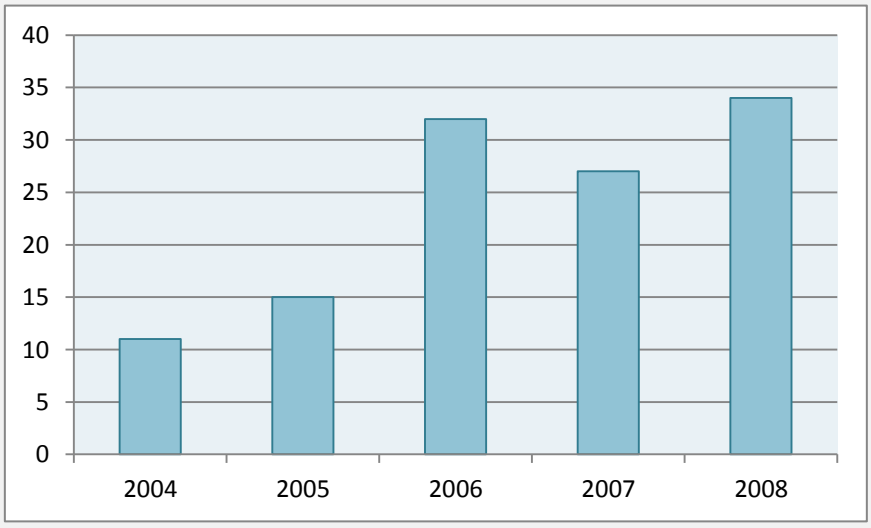
وهي أشرطة بيانية تخص صنفين أو اكثر للبيانات مثل عدد الكتب الموجودة في إحدى المكتبات مصنفة حسب أنواعها الى كتب علمية وأدبية وتاريخية وغيرها أو عدد الموظفين في إحدى الدوائر مصنفين حسب درجاتهم الوظيفية.

مثال :

البيانات التالية تمثل عدد الندوات والمؤتمرات العلمية التي عقدتها كليات جامعة ما:

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

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



2- الشريط الأفقي:

يمكن اعتبار شريط مستطيل أفقي واحد، طوله مقسم تقسيما متناسبا مع التكرارات. عرض الشريط ليس له أي أهمية أو دلالة.

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

مثال:

بلغت تكاليف إنتاج سلعة معينة (300) دينار كما هي موضحة بالجدول التالي:

مستلزمات الإنتاج	التكاليف (دينار)
أجور العمال	120
المواد الخام	60
المصاريف المباشرة	90
المصاريف غير المباشرة	30
المجموع	300

نفرض أن طول قاعدة المستطيل الكلي = 10

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

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

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

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

أجور العمال	المواد الخام	المصاريف المباشرة	المصاريف غير المباشرة

3- الدائرة البيانية Pie-chart :

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

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

أو **زاوية القطاع = (البيانات الجزئية* °180) / البيانات الكلية**

يمكن اعتبار دائرة أو نصف دائرة، يتم تقسيمها إلى زوايا متناسبة مع التكرارات. يتم حساب زاوية القطاع (درجة القطاع) حسب القاعدة الثلاثية كمايلي:

في حالة الدائرة: $Xi = \frac{Fi}{\Sigma Fi} \times 360$

في حالة نصف الدائرة: $Xi = \frac{Fi}{\Sigma Fi} \times 180$

مثال:

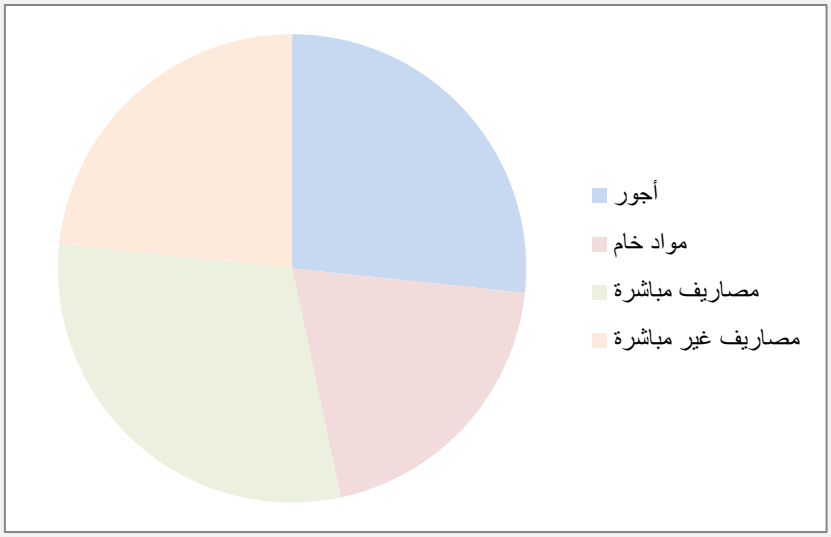
بالعودة الى بيانات مثال السابق ارسم الدائرة البيانية

زاوية قطاع الاجور = (80* °360) /300 = °96

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

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

زاوية قطاع المصاريف غير المباشرة = (70* °360) /300 = °84

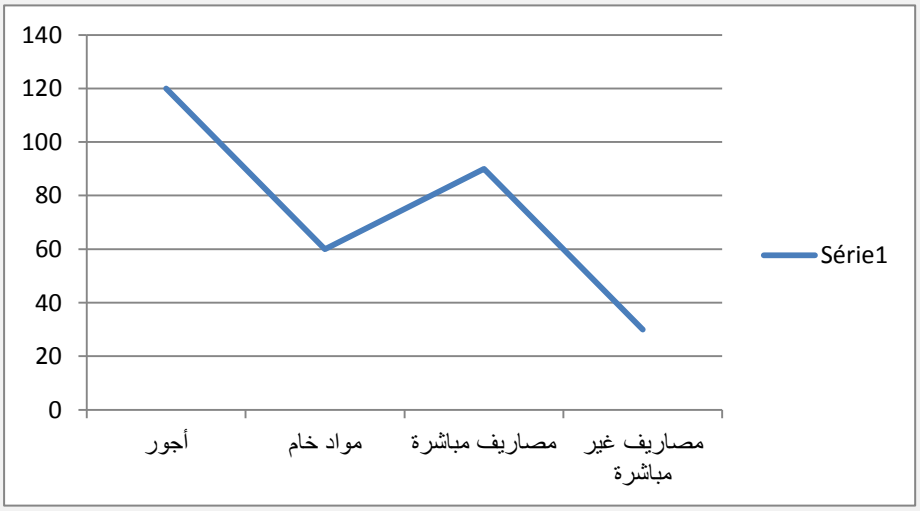


4- الخط البياني Line-chart

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

مثال: بالعودة الى بيانات المثال السابق

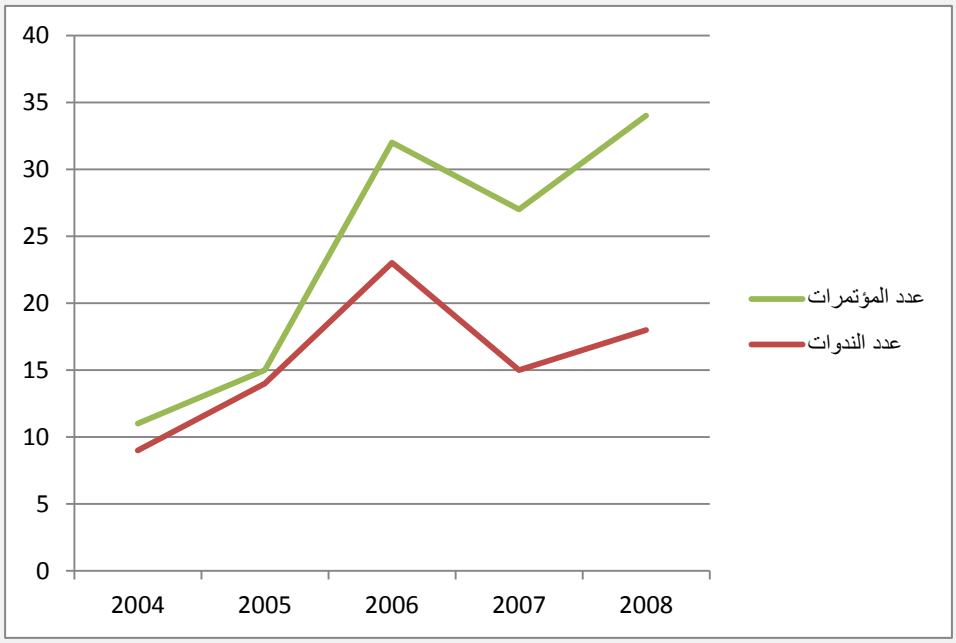
يتم رسم الخط البياني بعد وضع المحور الافقي X والذي يمثل مستلزمات الإنتاج في هذا المثال والمحور العمودي Y والذي يمثل التكاليف بعد ان يؤخذ تقسيم مناسب للمحور (Y)



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

بالعودة لبيانات المثال الأسبق

يكون الخط البياني كالتالي:



ثانيا: العرض الهندسي للبيانات المبوبة

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

1- المدرج التكراري Histogram

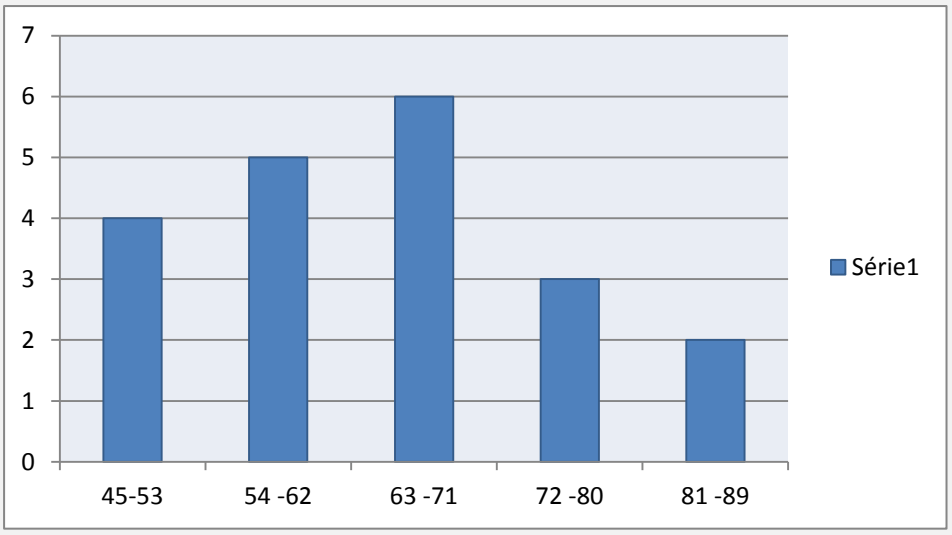
وهو عبارة عن مجموعة من المستطيلات قاعدة كل منها تمثل طول الفئة في التوزيع التكراري وارتفاعها يمثل التكرار المقابل لتلك الفئة أي ان المحور السيني (X) تستقر فيه الفئات والمحور الصادي (Y) تستقر فيه التكرارات هذه المستطيلات تكون منفصلة عن بعضها في حالة المتغير المتقطع ومتصلة مع بعضها في حالة المتغير المستمر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

مثال:

للتوزيع التكراري التالي ارسم المدرج التكراري

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

المحور الأفقي (X) تستقر فيه الفئات (عدد اشجار النخيل) والمحور العمودي (Y) تستقر فيه التكرارات (عدد العائلات التي تملكها).



2- المضلع التكراري Polygon

وهو عبارة عن عدد من المستقيمات التي تتصل ببعضها بواسطة نقاط هذه النقاط تمثل مراكز الفئات أي ان المحور السيني (X) تستقر فيه مراكز الفئات والمحور الصادي (Y)

تستقر فيه التكرارات مع مراعاة غلق المضلع بمركزي فئة وهميين (مركز الفئة الاولى – طول الفئة ومركز الفئة الاخيرة + طول الفئة) وبتكرارين مساويين للصفر.

مثال : ارسم المضلع التكراري للتوزيع التكراري في مثال السابق

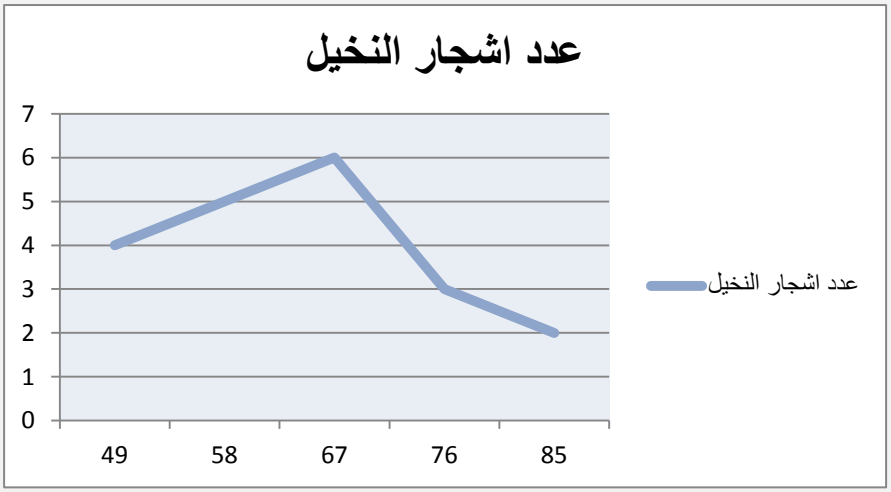
عدد الأشجار	التكرار	مراكز الفئات
53 –45	4	49
62 –54	5	58
71 – 63	6	67
80 –72	3	76
89– 81	2	85
المجموع	20	/

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

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

M6=85+9=94 Mo=49-9=40

فيكون المضلع بالشكل التالي:



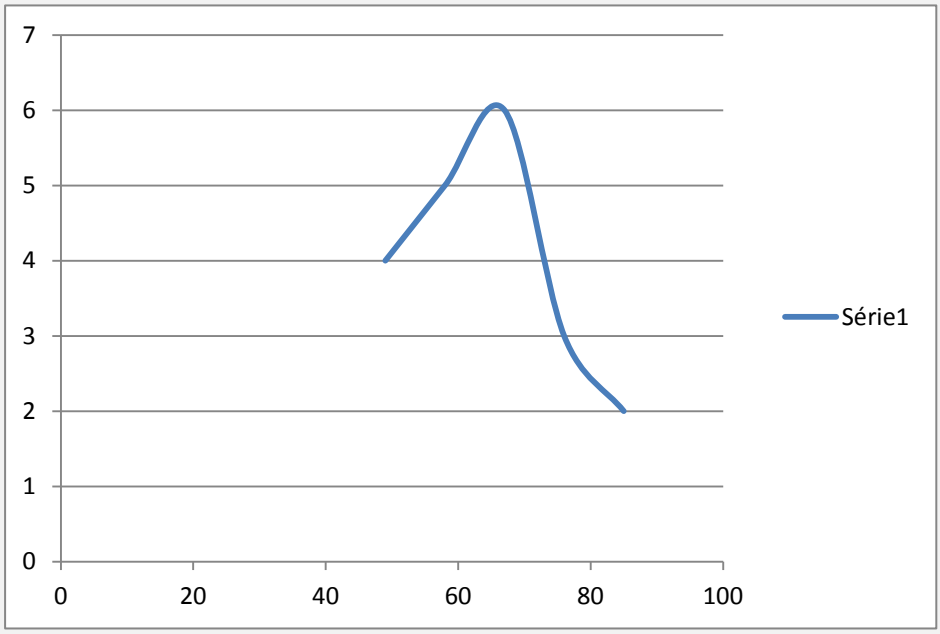
3- المنحنى التكراري Curve

ان فكرة رسم المنحنى التكراري لا تختلف عن رسم المضلع التكراري الا انه يتم ايصال النقاط بخط منحنى بدلا من الخطوط المستقيمة و لا يوجد داعي لغلقه.

مثال: نرسم المنحنى التكراري لنفس المثال السابق (13) .

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

فيكون المنحنى بالشكل التالي:



نلاحظ أنه نفس الشكل السابق ولكنه ممهد أي أنه لا يحتوي على زوايا.

المحور الثاني:

طرق تحليل البيانات

الدرس 6: مقاييس النزعة المركزية

تمهيد:

يساهم تبويب البيانات وعرضها بيانيا في إيضاح معالمها وسهولة وصفها للظاهرة محل الدراسة، غير أن ذلك لا يكفي للوصول إلى نتائج وافية، وإنما يشكل خطوة أولى نحو عملية التحليل الاحصائي للبيانات وتفسيرها، وذلك بحساب قيم خاصة (مقاييس) توضح السمات الأساسية لهذه البيانات؛ من هذه المقاييس ما يوضح نزعة البيانات المدروسة نحو قيمة مركزية ما بعينها وتسمى مقاييس النزعة المركزية، ومنها ما يوضح نزعة هذه البيانات نحو التشتت أو التباعد أو التباين وتسمى مقاييس التشتت.

ماهية مقاييس النزعة المركزية:

مقاييس النزعة المركزية (measures of central tendency) هي المقاييس التي تحاول أن تصف نقطة تجمع المشاهدات. وتعود فكرتها إلى الباحث الإنجليزي فرانسيس جالتون. إن الطرق الإحصائية التي تقوم بحساب القيمة التي تتمركز حولها معظم المشاهدات تسمى مقاييس النزعة المركزية، ومن أهمها وأكثرها استخداما ثلاثة مقاييس هي: الوسط الحسابي،الوسيط، المنوال. وسنتعلم حساب كل منها حسب طبيعة تبويب البيانات: غير المبوبة والمبوبة (الجداول التكرارية).

1- الوسط الحسابي (المتوسط):

من أكثر المتوسطات استخداما حتى لدى العامة (متوسط دخل الفرد، متوسط عمر السكان، متوسط درجة الحرارة...)، ويعرف عامة بالمعدل. لكل توزيع متوسط حسابي واحد و يرمز له بالرمز $\overline{x}$.

يقوم المتوسط الحسابي لمجوعة من القيم لمتغير ما على مبدأ بسيط هو: مجموع هذه القيم مقسوما على عددها؛ وبصيغة رياضية :

$$\overline{x} = \frac{\Sigma \times i}{n} = \frac{x1 + x2 + x3 + \cdots + xn}{n} = \frac{\text{مجموع القيم الفردية}}{\text{عدد الأفراد}}$$

حيث:

$\overline{x}$: المتوسط الحسابي.

x_i: القيم المتحصل عليها.

n: حجم العينة (عدد القيم).

مثال:

إذا كانت علامات مجموعة من طلاب الصف السابع في مادة الرياضيات هي: 18، 14، 11، 10، 12 أوجد الوسط الحسابي لعلامات الطلاب؟

الحل:

- نجد مجموع القيم: $18 + 14 + 11 + 10 + 12 = 65$

- نقسم المجموع على عدد القيم: $\frac{65}{5} = 13$

- الوسط الحسابي =13

ثانياً: في حالة البيانات المبوبة (الجداول التكرارية)

في حالة البيانات المبوبة في جدول ذو تكرارات فإننا نقوم بضرب كل قيمة فردية في تكرارها تم نجمع النواتج، فتكون المعادلة كما يلي:

$$\overline{x} = \frac{\Sigma(\text{xi. fi})}{\Sigma \text{fi}} = \frac{(x1.f1) + (x2.f2) + (x3.f3) + \cdots + (xn.fn)}{f1 + f2 + f3 + \cdots + fn}$$

حيث:

$\overline{x}$: المتوسط الحسابي.

x_i: القيم المتحصل عليها.

f_i: تكرار القيمة x_i

Σf_i: مجموع التكرارات =n.

مثال:

يبين الجدول التالي أجور 40 عامل في أحد المصانع وكانت كالآتي:

الفئات	التكرار(K)	مراكز الفئات (X)	K×X
9- 3	10	6	60
15- 9	12	12	144
21- 15	8	18	144
27- 21	6	24	144
33- 27	3	30	90
39- 33	1	36	36
المجموع	40		618

$$\text{الوسط الحسابي } (\bar{X}) = \frac{\sum K \times X}{\sum K} = \frac{618}{40} = 15.45 \text{ ديناراً}$$

في حالة البيانات المبوبة في فئات:

في هذه الحالة نحسب أولا مراكز الفئات ثم نجمع حاصل ضرب مركز كل فئة في تكرارها، ونقسم الناتج على مجموع التكرارات، أي أن قيمة مركز الفئة تعوض قيم البيانات التي تندرج ضمنها، فيكون القانون كما يلي:

$$\overline{x} = \frac{\Sigma(\text{ci. fi})}{\Sigma \text{fi}} = \frac{(c1.f1) + (c2.f2) + (c3.f3) + \cdots + (cn.fn)}{f1 + f2 + f3 + \cdots + fn}$$

حيث:

$\overline{x}$: المتوسط الحسابي.

c_i: القيم المتحصل عليها.

f_i: تكرار الفئة c_i

Σf_i: مجموع التكرارات.

من عيوب الوسط الحسابي أنه لا يمكن إيجاده بالرسم ويتأثر بالقيم الشاذة، أما مزاياه فمنها السهولة في الحساب ولذلك فهو أكثر المتوسطات استخداماً كما تدخل جميع قيم المجموعة في حسابه.

خصائص الوسط الحسابي:

1- يعتمد على جميع القيم والمشاهدات

2- هو نقطة اتزان المشاهدتان

3- مربع الانحرافات اقل ما يمكن عن الوسط

4- اقل مقاييس النزعة المركزية تأثرا بالتقلبات العينية

5- الوسط الحسابي يتأثر بالقيم المتطرفة والقيم الشاذة لذا لا يصلح للتوزيعات الملتوية.

6- لا يصلح في حالة الفئات المفتوحة (لعدم وجود مركز فئة)

7- مجموع انحرافات القيم عن المتوسط الحسابي يساو الصفر.

$$\sum (X - \bar{X}) = 0$$

8- مجموع مربعات انحرافات القيم عن الوسط أقل من مجموع مربعات انحرافات القيم عن أي قيمة أخرى.

9- الوسط الحسابي يتأثر بالعمليات الحسابية الأربعة.

ملاحظة:

المتوسط الحسابي إحصاء هام لقياس النزعة المركزية إذا كانت البيانات تعطي توزيعا معتدلا، حيث تكون القيم متماثلة حول المتوسط الحسابي، وكلما كان التوزيع غير متناظر حول المتوسط فإن إحصاء المتوسط الحسابي يصبح غير ملائم لقياس النزعة المركزية، ذلك أن بعض القيم متطرفة إلى الأعلى أو الأدنى تجذب المتوسط نحوها.

2- المنوال MODE:

يعرف المنوال بأنه القيمة الأكثر تكرارا في التوزيع ، ويرمز له بالرمز M_O .

المنوال هو القيمة الأكثر تكراراً في البيانات.

قد يكون للتوزيع أكثر من منوال واحد إذا كانت عدة قيم لديها نفس التكرار الأكبر قيمة، وقد لا يكون لديه أي منوال إذا كانت جميع التكرارات متساوية.

لا يحتاج استخراج المنوال إلى إجراء أية عملية حسابية في حالة البيانات البسيطة والمبوبة؛ إذ يكفي استخراج القيمة التي لديها أكبر تكرار.

أولاً: حساب المنوال في حالة البيانات غير المبوبة:

مثال: 6 ، 5 ، 5، 4، 7، 2، 5، 3، 8

نلاحظ هنا أن القيمة 5 تكررت ثلاث مرات هذا يعني أن قيمة المنوال هنا هي: 5

ثانياً: في حالة البيانات المبوبة (جداول تكرارية):

نقوم بحساب التكرار المتجمع الصاعد أو النازل ، ثم نحسب رتبة الوسيط ونقارنها بقيم التكرارا المتجمع (التراكمي):

إذا وجدنا في قيم التكرار التراكمي قيمة مساوية لرتبة الوسيط، فإن الوسيط هو القيمة التي تقابلها.

وإذا وجدنا رتبة الوسيط محصورة بين قيمتين إحداهما أكبر منها والأخرى أصغر منها. فإن الوسيط هو القيمة التي تقابل التكرارا التراكمي الأكبر منها مباشرة.

ثالثا: في حالة البيانات المبوبة في فئات:

نستخرج الفئة المنوالية وهي التي لديها أكبر تكرار، ثم نحسب قيمة المنوال من خلال المعادلة التالية:

$$M_o = d + \frac{fi+1}{fi+1+fi-1} . L$$

حيث:

- d: الحد الأدنى للفئة المنوالية.
- F_{i+1}: التكرار التالي لتكرار الفئة المنوالية.
- F_{i-1} : التكرار السابق لتكرار الفئة المنوالية.

- L: طول الفئة المنوالية.

الطريقة الثانية:

يمكن حساب المنوال التقريبي وهو مركز الفئة المنوالية.

مثال:

الفئات	3 -9	9 -15	15 -21	21 -27	27 -33	33 -39	المجموع
التكرار	10	12	8	6	3	1	40

من الجدول نلاحظ أن الفئة التي تقابل أكثر تكرار هي الفئة (9 -15) هذا يعني أن المنوال يكون عبارة عن حاصل جمع الحدين مقسوما على 2

خواص المنوال:

1- غير ثابت

2- يتأثر بطول الفئة

3- يفضل عندما يكون المقياس اسمي

4- لا يعتمد عليه في حالة الإحصاءات اللاحقة

3- الوسيط Médiane :

يعرف وسيط أية مجموعة من القيم بأنه القيمة التي تقع في وسط تلك القيم بعد ترتيبها تصاعديا أو تنازليا في حالة وجود رقمين تضع وسيطهما. هو العدد الأوسط في البيانات المرتبة تصاعدياً أو تنازلياً عندما يكون عددها فردياً، أو هو الوسط الحسابي للعددين الأوسطين عندما يكون عدد البيانات زوجياً.

يقسم الوسيط مجموع قيم التوزيع إلى مجموعتين متساويتين . ويرمز له بالرمز Me.

وتختلف طرق حسابه باختلاف طبيعة البيانات:

- **أولاً: حساب الوسيط في حالة البيانات غير المبوبة**

في حالة ما إذا كان عدد البيانات قليلا نرتب القيم تصاعديا أو تنازليا، ثم نقوم بحساب رتبة الوسيط ولنرمز لها بالرمز R كما يلي:

إذا كان عدد مفردات العينة n فرديا فإن رتبة الوسيط هي: $R=\frac{n+1}{2}$

مثال:

لحساب الوسيط لمجموعة القيم: 30 ، 60 ، 20 ، 40 ، 70

الحل:

- نرتب القيم ترتيب تصاعدي أو تنازلي كالتالي: 20 ، 30 ، 40 ، 60 ، 70

-عدد القيم فردي نأخذ القيمة الواقعة في الوسط وهي : 40

إذا كان عدد مفردات العينة n زوجيا فإننا نأخذ كوسيط متوسط العددين الموجودين في الوسط بعد الترتيب التصاعدي أو التنازلي، حيث أن رتبة الوسيط الأول هي: $R_1=\frac{n}{2}$

ورتبة الوسيط الثاني هي: $R_2=\frac{n+2}{2}$

نستخرج الوسيطين الأول والثاني ثم نحسب الوسيط الذي يمثل متوسطهما الحسابي كما يلي:

$$Me=\frac{Me1+Me2}{2}$$

مثال:

نجد الوسيط للقيم 12 ، 4 ، 1 ، 6 ، 11 ، 8

-نرتب القيم ترتيب تصاعدي أو تنازلي: 1 ، 4 ، 6 ، 8 ، 11 ، 12 .

-نلاحظ هنا أن عدد القيم زوجي وبعد ترتبها تصاعدياً نأخذ القيمتان التي بالمنتصف وهما 6 ، 8

-نجمعهما ونقسم الناتج على 2 أي: $\frac{6+8}{2} = \frac{14}{2} = 7$

ثانيا: في حالة البيانات المبوبة في جدول تكراري:

في هذه الحالة نستخرج رتبة الوسيط بالعلاقة: $R = \frac{n}{2}$

ثم نبحث في عمود التكرار المتجمع الصاعد أو النازل؛ الوسيط هو القيمة التي تقابل التكرار المتجمع الصاعد الذي يساوي رتبة الوسيط R أو يكون أكبر منها تماما.

ثالثا: في حالة البيانات المبوبة في فئات:

بنفس الطريقة السابقة نحسب رتبة الوسيط ونستخرج انطلاقا من التكرار المتجمع الصاعد أو النازل الفئة الوسيطية التي تقابل رتبة الوسيط أو تكون أكبر منها مباشرة، ثم نحسب الوسيط بالعلاقة التالية:

$$M_e = d + \frac{R-(Fi-1)}{(fi++1)-(f+i-1)} . L$$

حيث:

d: الحد الأدنى للفئة الوسيطية.

F^{+}_{i+1}: التكرار المجمع الصاعد (أو النازل) التالي لتكرار الفئة الوسيطية.

F^{+}_{i-1} : التكرار المتجمع الصاعد (أو النازل) السابق لتكرار الفئة الوسيطية.

L: طول الفئة الوسيطية.

a. مثال:

الفئات	3- 9	9- 15	15- 21	21- 27	27- 33	33- 39	المجموع
التكرار	10	12	8	6	3	1	40

الحل: نحتاج لتكوين جدول التكرار التراكمي الذي يضم عمودين، العمود الأول يضم الحدود الفعلية العليا والعمود الثاني التكرار التراكمي. لإيجاد التكرار التراكمي نجمع التكرارات، حيث الحد الأعلى الفعلي للفئة الأولى يأخذ أول تكرار ونجمع التكرارات حتى يتم الوصول إلى آخر حد فعلي يأخذ عدد التكرارات جميعها

الحدود الفعلية العليا	التكرار التراكمي
9.5	10
15.5	22
21.5	30
27.5	36
33.5	39
39.5	40

أولاً : نجد رتبة الوسيط وهو عبارة عن مجموع التكرارات مقسومة على 2 ، إذن

$$20=\frac{40}{2}$$

تكون رتبة الوسيط في الجدول التكرار التراكمي بين 10 وَ 22 أي:

9.5	10
س	رتبة الوسيط=20
15.5	22

الآن نستخدم النسبة والتناسب لإيجاد قيمة الوسيط

$$\frac{x-9.5}{15.5-9.5}=\frac{20-10}{22-10}$$

$$\frac{x-9.5}{6}=\frac{10}{12}$$

$$12\times(x-9.5)=10\times 6$$

$$(x-9.5)=\frac{60}{12}$$

$$x-9.5=5\Rightarrow x=14.5$$

إذن، قيمة الوسيط هي 14.5

من مزايا الوسيط أنه لا يتأثر بالقيم الشاذة، ويمكن الحصول عليه بالرسم، ومن عيوبه أنه لا يدخل في حسابه سوى قراءة واحدة أو قراءتين من المجموعة كلها.

خواص الوسيط:

- لا يتأثر بالقيم المتطرفة
- يستخدم في التوزيعات الملتوية
- يفضل استخدامه في حالة الفئات المفتوحة
- يأتي بعد الوسط في تأثره بالتقلبات العينية

الدرس 7: مقاييس التشتت

تمهيد:

لقد سبق لنا وتكلمنا عن مقاييس النزعة المركزية وذلك لوصف البيانات عدديا لهذه التوزيعات المختلفة، ولكن حساب المتوسطات للمجموعات المختلفة من البيانات غير كاف للمقارنة بين هذه المجموعات، فربما تكون المتوسطات (الوسط الحسابي) لأكثر من مجموعة متساوية ولكن هذه المجموعات مختلفة كثيرا.

ماهية التشتت (dispersion):

إذا كانت مجموعة البيانات متباعدة أو متباينة عن بعضها يقال أنها متشتتة، أما إذا كانت البيانات متجانسة وغير متباعدة فيقال أنها غير متشتتة.

مثال:

ولتوضيح ذلك نأتي بمثل بمثال لدراسة ثلاث مجموعات مختلفة من الطلاب X , Y , Z وكانت الدرجات كالأتي:

X= {60، 61 ، 62 ، 58 ، 59 }

Y= {50 ، 60 ، 66 ، 54 ، 70}

Z= {39 ، 65 ، 46 ، 78 ، 72}

بحساب الوسط الحسابي للثلاث مجموعات نجده يساوي 60 درجة لكل منها. ولكن عند النظر لدرجات المجموعة الأولى نجدها متقاربة، ودرجات المجموعة الثانية أقل تقاربا من المجموعة الأولى، ودرجات المجموعة الثالثة أقل تقاربا من درجات المجموعة الثانية. أي أن الثلاث مجموعات مختلفة التجانس رغم أن الوسط الحسابي لهم متساو، وبذلك

تكون مقاييس النزعة المركزية غير كافية للمقارنة بين طبيعة البيانات الإحصائية، لذلك نشأت الحاجة إلى إيجاد مقاييس تقيس درجة تجانس (تقارب) أو تشتت (تباعد) مفردات البيانات عن بعضها البعض، وتعرف هذه المقاييس بمقاييس التشتت.

يستخدم علماء الإحصاء عدة مقاييس لتحديد درجة انحراف البيانات عن القيمة الوسطية أو عن بعضها البعض ويطلقون عليها اسم مقاييس التشتت، ومن أكثرها شيوعاً ما يلي:

1- المدى Range

يعرف المدى (Range) بأنه الفرق بين أكبر مشاهدة وأصغر مشاهدة، وتدل القيمة الكبيرة للمدى على أن البيانات متباعدة، أما القيمة الصغيرة فتدل على أن البيانات قريبة من بعضها البعض.

أي أن : المدى = أكبر مشاهدة – أصغر مشاهدة.

ويرمز له بالرمز E

أ- في حالة توزيع تكراري أو بدون تكرارات:

$$E= H-L$$

حيث:

H: الحد الأعلى للتوزيع.

L: الحد الأدنى للتوزيع.

مثال:

أوجد مدى الدرجات التالية: 25،30،45،29،35،40

نلاحظ أن أكبر قيمة هي:45 وأصغر قيمة هي: 25

إذن المدى E = 25 - 45 =20

ب-في حالة بيانات مبوبة في فئات

وهو حاصل الفرق بين الحد الفعلي الأعلى لآخر فئة و الحد الفعلي الأدنى لأول فئة.

المدى = الحد الفعلي الأعلى للفئة العليا - الحد الفعلي الأدنى للفئة الدنيا .

2- الانحراف المتوسط:

إن الانحراف المتوسط يفيدنا في معرفة في معرفة متوسط انحرافات القيم عن متوسطها الحسابي (X) وهذا بغض النظر عن إشارات الانحراف ويرمز له بالرمز (MD) مع العلم أن قيمة الانحراف المتوسط تزداد كلما تباعدت قيم (XI) عن بعضها البعض وتصغر قيمته كلما تقاربت، ويمكن حساب الانحراف المتوسط بواسطة المعادلة التالية:

$$e = \frac{\sum IXi - \bar{X}I}{n}$$

خطوات حساب الانحراف المتوسط:

- نحسب المتوسط الحسابي.
- نحسب انحراف كل قيمة عن المتوسط.
- نتجاهل إشارات الإنحرافات.
- نجمع هذه الإنحرافات.
- نقسم مجموع الانحرافات على عدد الحالات، فيكون الناتج هو الانحراف المتوسط.

في حالة البيانات المبوبة في جدول تكراري:

$$e = \frac{\sum IXi - \bar{X}I \,.\, Fi}{\Sigma Fi}$$

في حالة البيانات المبوبة في جدول تكراري ذو فئات:

$$e = \frac{\sum ICi - \bar{X}I \,.\, Fi}{\Sigma Fi}$$

3- التباين:

يعرف التباين (Variance) للمشاهدات المفردة أو لتوزيعات البيانات التكرارية بأنه الوسط الحسابي لمربعات انحرافات القيم عن وسطها الحسابي، يرمز إلى الانحراف المعياري بالرمز s^2

ويحسب بالصيغة التالية:

$$S^2 = \frac{\sum(Xi - \bar{X})^2}{n}$$

حيث:

$\bar{x}$: الوسط الحسابي .

n: حجم العينة (أو عدد البيانات رموز رياضية)

والتباين هو مربع الانحراف المعياري.

- **في حالة البيانات المبوبة في جدول تكراري:**

$$S^2 = \frac{\sum(Xi - \bar{X})^2 . Fi}{\Sigma Fi}$$

- **في حالة البيانات المبوبة في جدول تكراري ذو فئات:**

$$S^2 = \frac{\sum(Ci - \bar{X})^2 . Fi}{\Sigma Fi}$$

4- الانحراف المعياري:

الانحراف المعياري هو عبارة عن الجذر التربيعي للتباين. أي أنه الجذر التربيعي لمتوسط مربعات انحرافات القيم عن المتوسط الحسابي. ويرمز له بالرمز S

حساب الانحراف المعياري:

- في حالة بيانات غير مبوبة:

$$S = \sqrt{\frac{\sum(Xi - \bar{X})^2}{n}}$$

- في حالة البيانات المبوبة في جدول تكراري:

$$S = \sqrt{\frac{\sum(Xi - \bar{X})^2 . Fi}{\Sigma Fi}}$$

- في حالة البيانات المبوبة في جدول تكراري ذو فئات:

$$S = \sqrt{\frac{\sum(Ci - \bar{X})^2 . Fi}{\Sigma Fi}}$$

مثال:

أحسب مقاييس النزعة المركزية والتشتت للبيانات التالية:

X_i	F_i
1	6
2	10
3	9
4	4
5	1
المجموع	30

$(X_i-\overline{X})^2.F_i$	$(X_i-\overline{X})^2$	$\lvert X_i-\overline{X}\rvert.F_i$	$\lvert X_i-\overline{X}\rvert$	F+	$X_i.F_i$	F_i	X_i
12.78	2.13	8.76	1.46	6	6	6	1
2.11	0.21	4.6	0.46	16	20	10	2
2.62	0.29	4.86	0.54	25	27	9	3
9.48	2.37	6.16	1.54	29	16	4	4
6.45	6.45	2.54	2.54	30	5	1	5
33.42	/	26.92	/	/	74	30	Σ

1) **الوسط الحسابي:** $\overline{X}=\frac{\sum.Fi}{\sum Fi}=\frac{74}{30}=2.46$

2) **الوسيط :** هو القيمة التي تتوسط التوزيع و رتبة الوسيط هي: $c=\frac{\sum \mathrm{Fi}}{2}=\frac{30}{2}=15$

نبحث في التكرار المتجمع الصاعد (**F+**) فنجدها محصورة بين 6 و16 وبالتالي فإن الوسيط هو قيمة (X_i) المقابلة ل16 وهي 2. **Me=2**

3) **المنوال:** القيمة الأكثر تكرارا هي 2. **Mo=2**

4) **المدى:** **E=H-L=5-1=4**

5) **التباين:** $e=\frac{\sum IXi-X1.Fi}{\sum \mathrm{Fi}}=\frac{26.92}{30}=0.89$

6) **الانحراف المتوسط :** $s^2=\frac{\sum(Xi-X)^2.Fi}{\sum \mathrm{Fi}}=\frac{33.42}{30}=1.11$

7) **الانحراف المعياري:** $S=\sqrt{\frac{\sum(Xi-X)^2.Fi}{\sum Fi}}=\sqrt{s^2}=\sqrt{1.11}=1.05$

الدرس 8: معاملات الارتباط

تمهيد:

يقيس الارتباط مدى العلاقة بين الظواهر المختلفة (ظاهرتين أو أكثر أو متغيرين أو أكثر) لمعرفة ما إذا كان تغير أحدهما أو مجموعة منها مرتبطاً بتغير الاخرى، فقد يريد الباحث معرفة ما إذا كان هناك علاقة بين التدخين والإصابة بمرض في الرئة، أو بين درجة تعليم الشخص ومستوى دخله، أو بين الحالة التعليمية والحالة الاجتماعية للناخب... بل قد يرغب الباحث في دراسة العلاقة بين أكثر من متغيرين في وقت واحد؛ فقد يريد الباحث معرفة تأثير درجة التعليم ومستوى الدخل وحجم الأسرة على درجة الوعي السياسي للشخص. في هذا المثال يريد الباحث معرفة تأثير ثلاثة متغيرات على متغير رابع وهكذا.

يمكن تحديد الارتباط بين متغيرين من خلال استخدام مجموعة من الإحصاءات تعرف بأسم معاملات الارتباط ومعامل الارتباط هو رقم يلخص التحسن في تخمين القيم على متغير واحد لأي حالة على أساس معرفة قيم المتغير الثاني، فكلما ارتفع المعامل قوي الارتباط، ومن ثم تحسنت قدرتنا التنبؤية أو التفسيرية.

معامل الارتباط:

هو تعبير يشير إلى المقياس الإحصائي الذي يدل على مقدار العلاقة بين المتغيرات سلبية كانت أم إيجابية، وتتراوح قيمته بين الارتباط الموجب التام (+1) وبين الارتباط السالب التام (-1) .

إن الارتباط الكامل لا وجود له في الظواهر الاجتماعية، وأن معامل الارتباط الناتج في الأبحاث والدراسات الإنسانية والاجتماعية يكون عادة كسرا موجبا أو سالبا.

والجدول التالي يوضح أنواع العلاقات بين المتغيرات كما يصفها معامل الارتباط:

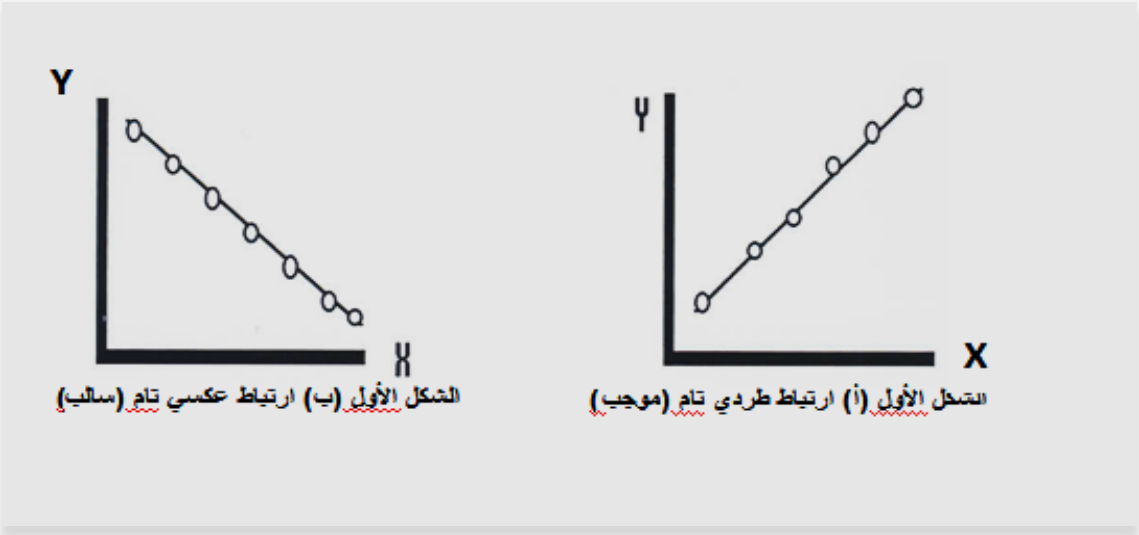
نوع الارتباط	قيمة معامل الارتباط
ارتباط طردي تام	**+1**
ارتباط طردي قوى	**من 0.7 إلى أقل من +1**
ارتباط طردي متوسط	**من 0.4 إلى أقل من 0.7**
ارتباط طردي ضعيف	**من صفر إلى أقل من 0.4**
ارتباط منعدم	**صفر**
ارتباط عكسي قوى	**من -0.7 إلى أقل من -1**
ارتباط عكسي متوسط	**من -0.4 إلى أقل من -0.7**
ارتباط عكسي ضعيف	**من صفر إلى أقل من -.04**

شكل الانتشار Scatter Diagram :

هناك وسيلة مبدئية يعرف الباحث من خلالها نوع الارتباط بين المتغيرين وشدته، وما إذا كانت العلاقة خطية أو غير خطية، موجبة أو سالبة. هذه الوسيلة هي " شكل الانتشار " والتي تصلح إذا كان المتغيران كميين. وجدير بالذكر أن هذه وسيلة مبدئية تساعد فقط في معرفة نوع الارتباط ولا تعتبر بديلاً عن الطرق الإحصائية التي سوف نتناولها بالتفصيل في هذه المحاضرة.

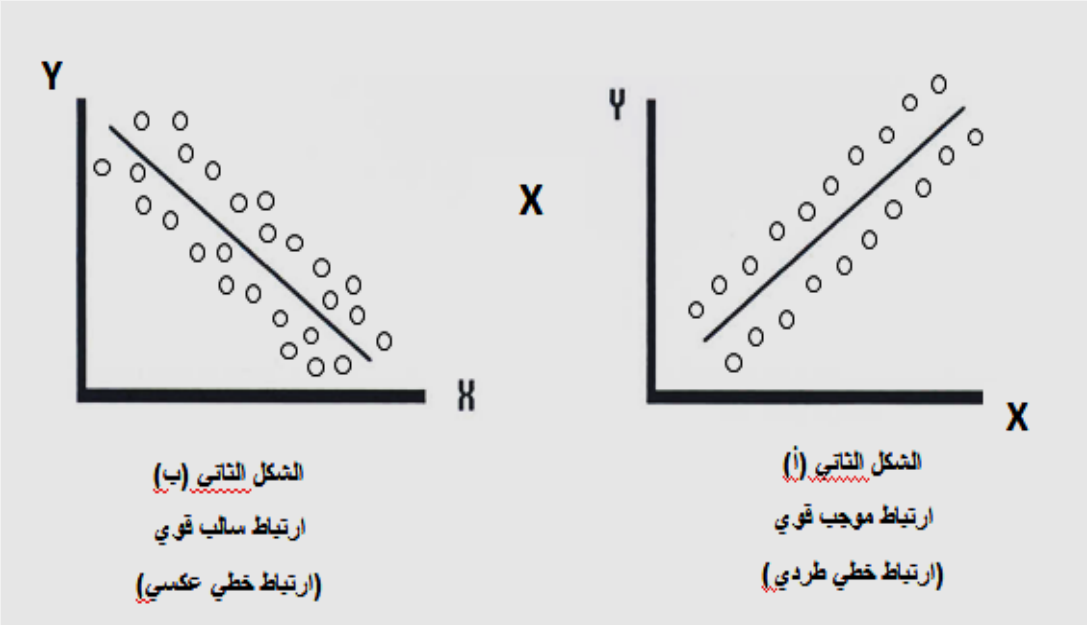
والمقصود بشكل الانتشار هو تمثيل قيم الظاهرتين بيانياً على المحورين، المتغير الأول X على المحور الأفقي، والمتغير الثاني Y على المحور الرأسي، حيث يتم تمثيل كل زوج من القيم بنقطة، فنحصل على شكل يمثل كيفية انتشار القيم على المستوى، وهو الذي يسمى شكل الانتشار.

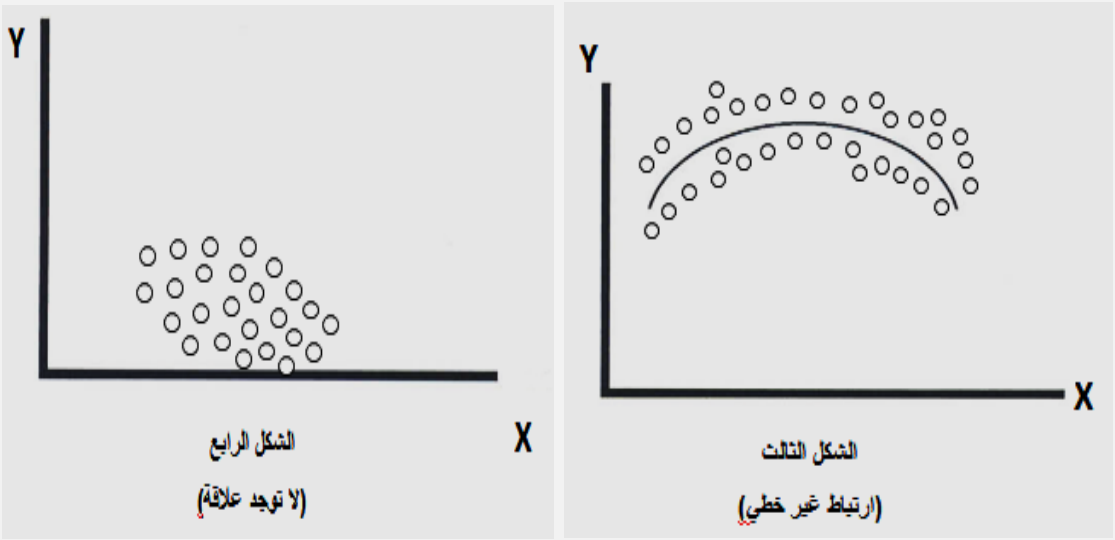
وطريقة انتشار القيم تدل على وجود أو عدم وجود علاقة بين المتغيرين ومدى قوتها ونوعها. فإذا كانت تتوزع بشكل منتظم دل ذلك على وجود علاقة (يمكن استنتاجها)، أما إذا كانت النقط مبعثرة ولا تنتشر حسب نظام معين دل ذلك على عدم وجود علاقة بين المتغيرين أو أن العلاقة بينهما ضعيفة. والأشكال التالية تظهر بعض أشكال الانتشار المعروفة :



الشكل الأول (أ) ارتباط طردي تام (موجب)

الشكل الأول (ب) ارتباط عكسي تام (سالب)





إذا كانت العلاقة تأخذ شكل منحنى فإن الارتباط لا يكون خطياً "ارتباط غير خطي " Non Linear Correlation كما في الشكل الثالث .

من أهم مهاملات الارتباط:

1- معامل بيرسون للارتباط الخطى البسيط Person's Correlation Coefficient

يعتبر معامل الارتباط الخطى البسيط لبيرسون Person's Correlation Coefficient والذى سنرمز له بالرمز r_p يعتبر معامل ارتباط بيرسون من اكثر المعاملات الارتباط شيوعا واستعمالا عندما يكون كلا المتغيرين متغير كميا متصلا مثل الذكاء والتحصيل مثلا او الطول والوزن .

وهناك أكثر من صيغة يمكن الاعتماد عليها فى حساب معامل الارتباط الخطى البسيط لبيرسون منها:

$$r_p = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2}\sqrt{\sum (y-\bar{y})^2}}$$

وكذلك المعادلة الرياضية التالية والتي تعتبر اسهل وابسط:

$$r_p = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}}$$

من أهم خصائص معامل الارتباط الخطى البسيط لبيرسون أنه لا يعتمد على قيم المتغيران نفسها عند حساب قيمته وإنما يعتمد على مقدار التباعد بين هذه القيم بعضها البعض. لذلك لا يتأثر معامل الارتباط الخطى البسيط بأي عمليات جبرية يتم إجراءها على بيانات كل من المتغيرين أو أحدهما من جمع أو طرح أو ضرب أو قسمة.

يشير معامل ارتباط بيرسون إلى درجة العلاقة الخطية بين متغيرين , و من الممكن أن يشير إلى عدم وجود ارتباط بين متغيرين إذا كان لديهم علاقة منحنية، وبالتالي. لذلك من المهم التحقق من خطية الدرجات خلال شكل الانتشار.

مثال:

أوجد معامل الارتباط بين دخل تسعة أسر (X) والإنفاق (Y) اليومي بالدينار والمبينة في الجدول الآتي:

X	**6**	**8**	**7**	**14**	**11**	**12**	**8**	**9**	**10**
Y	**4**	**8**	**6**	**10**	**9**	**11**	**8**	**7**	**8**

نكون جدول البيانات حسب احد القوانين المذكورة أعلاه وليكن الثاني مع العلم بأن الصف الخير يعبر عن المجموع للبيانات في العمود والتي تقع أعلاه.

X	Y	XY	X^2	Y^2
6	4	24	36	16
8	8	64	64	64
7	6	42	49	36
14	10	140	196	100
11	9	99	121	81
12	11	132	144	121
8	8	64	64	64
9	7	63	81	49
10	8	80	100	64
85	71	708	855	595

$$r = \frac{n(\sum X_i Y_i) - (\sum X_i)(\sum Y_i)}{\sqrt{n(\sum X_i^2) - (\sum X_i)^2}\sqrt{n(\sum Y_i^2) - (\sum Y_i)^2}}$$

$$r = \frac{9 \times 708 - 85 \times 71}{\sqrt{9 \times 855 - (85)^2}\sqrt{9 \times 595 - (71)^2}}$$

$$r = \frac{6372 - 6035}{\sqrt{7695 - 7225}\sqrt{5355 - 5041}}$$

$$r = \frac{337}{\sqrt{470}\sqrt{314}}$$

$$r = 0.877$$

قيمة معامل الارتباط تساوي 0.877 يعني أن العلاقة طردية قوية.

معامل التحديد Determination Coefficient

وهو مربع معامل الارتباط لذلك يرمز له بالرمز R^2 أو R-Square و هو يشير إلى نسبة تفسير المتغير أو المتغيرات المستقلة للتغير فى المتغير التابع

فمثلا:

نجد أن الدخل يفسر نسبة (0.877^2) أى76.91 % من التغير فى قيمة الاستهلاك بينما 23.09% من التغير فى الاستهلاك ترجع إلى عوامل أخرى منها الخطاء العشوائى .

2- معامل ارتباط الرتب (اسبيرمان) Spearman

إذا كانت الظاهرة محل الدراسة تحتوي على متغيرين وصفيين ترتيبين، ومثال على ذلك قياس العلاقة بين تقديرات الطلبة في مادتين ، أو العلاقة بين درجة تفضيل المستهلك لسلعة معينة ومستوى الدخل، فإنه يمكن استخدام طريقة "بيرسون" السابقة في حساب معامل ارتباط يعتمد على رتب مستويات المتغيرين كبديل للقيم الأصلية ، ويطلق على هذا المعامل " معامل ارتباط سبيرمان " Spearman ، ويعبر عنه بالمعادلة التالية :

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

حيث أن d هي الفرق بين رتب مستويات المتغير الأول X ، ورتب مستويات المتغير الثاني Y أي أن : $d = R_x - R_y$.

مثـــال :

فيما يلي تقديرات **10** طلاب في مادتي الإحصاء، والاقتصاد:

تقديرات إحصاء	أ	جـ+	د	د+	ب+	جـ+	أ+	ب	ب+	ب+
تقديرات اقتصاد	أ+	د	جـ	جـ	أ	ب	ب+	ب	جـ	ب

المطلوب:

1- احسب معامل الارتباط بين تقديرات الطلبة في المقررين.

2- وما هو مدلوله ؟

الحـــل

1- بفرض أن **x** هي تقديرات الإحصاء، **y** هي تقديرات الاقتصاد، يمكن حساب معامل الارتباط بينهما وذلك بإتباع الآتي:

الرتب	1	2	3	4	5	6	7	8	9	10
تقديرات إحصاء	أ+	أ	ب+	ب+	ب+	ب	جـ+	جـ+	د+	د
رتب X	1	2	(3+4+5)/3=4			6	(7+8)/2=7.5		9	10
تقديرات اقتصاد	أ+	أ	ب+	ب	ب	ب	جـ	جـ	جـ	د
رتب Y	1	2	3	(4+5+6)/3=5			(7+8+9)/3=8			10

- **إذا يمكن حساب المجموع:** $\sum d^2$ **كما يلي:**

x	y	رتب x	رتب y	d	d^2
أ	أ+	2	1	1	1
جـ+	د	7.5	10	-2.5	6.25
د	جـ	10	8	2	4
د+	جـ	9	8	1	1
ب+	أ	4	2	2	1
جـ+	ب	7.5	5	2.5	6.25
أ+	ب+	1	3	-2	4
ب	ب	6	5	1	1
ب+	جـ	4	8	-4	16
ب+	ب	4	5	-1	1
					44.5

$$\sum d^2 = 44.5$$

- **معامل الارتباط هو:**

$$r = 1 - \frac{6\sum d^2}{n(n^2-1)}$$
$$= 1 - \frac{6(44.5)}{10(10^2-1)} = 1 - \frac{267}{990}$$
$$= 1 - 0.2697 = 0.7303$$

2- مدلول معامل الارتباط :

بما أن $r = 0.703$ ، ويدل ذلك على وجود ارتباط طردي قوي بين تقديرات الطالب في مادة الإحصاء ، ومادة الاقتصاد .

ملاحظة:

يمكن استخدام صيغة معامل ارتباط "سبيرمان" في حساب الارتباط بين متغيرين كميين، حيث يتم استخدام رتب القيم التي يأخذها المتغير.

3- معامل الاقتران فاي

تعريف:

معامل اقتران فاي يستخدم لمعرفة العلاقة بين متغيرين اسميين ثنائيي التقسيم كالنوع (ذكر/أنثى) والاصابة بالمرض (مصاب/غير مصاب) أو التدخين (مدخن/غيرمدخن) على أن يكون المتغيرين معرفين في صورة جدول ثنائي مزدوج كما يلي:

	X_1	X_2	المجموع
Y_1	a	b	a+b
Y_2	c	d	c+d
المجموع	a+c	b+d	a+b+c+d

وتستخدم المعادلة التالية لاستخراج قيمة فاي:

$$r_{\phi} = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

ملاحظة:

عند تفسير نتيجة معامل الاقتران ننظر الى قوة العلاقة فقط (ضعيفة/متوسطة/قوية) وليس اتجاه العلاقة (طردية/عكسية).

مثال:

أوجد قيمة معامل الاقتران بين النوع (ذكر/أنثى) والاصابة بمرض الاكتئاب (مصاب/غير مصاب) حسب البيانات التالية:

	مصاب	غير مصاب	المجموع
ذكر	12	8	20
أنثى	4	6	10
المجموع	16	14	**30**

الحل:

$$r\phi = \frac{ad - bc}{\sqrt{(a+b)(c+d)a+c)(b+d)}}$$

$$\frac{(12\times6)-(8\times4)}{\sqrt{20\times10\times16\times14}} = \frac{72-32}{\sqrt{44800}} = \frac{40}{211.66} = 0.19$$

قيمة اقتران فاي = 0.19 وتدل على أن العلاقة بين النوع (ذكور واناث) والاصابة بمرض الاكتئاب ضعيفة، أي أنه ليس هناك تأثير للنوع على الاصابة بالمرض.

الدرس 9: اختبار الفروض

تمهيد:

إذا كان البحث العلمي يتضمن جانبا من الإبداع فإن الموطن الحقيقي للإبداع يكم في الفرض العلمي، فكل تلك الإبداعات العلمية والنظريات والقوانين كانت في البداية مجرد فرضيات علمية. تلعب الفرضية في العلم دور يتعذر علينا وصفه فهي اساس العلم و البوصلة التي توجه البحث، فهي تنقلنا من الجانب التجريدي النظري الى الواقع الملموس، قد تكون لنا تصورات رائعة عن الواقع و لكن هذه الافكار ستبقى من دون ادنى قيمة اذا لم ننجح في صياغتها في فرضية علمية

وتعتبر الفرضيات نقطة جوهرية في البحث تحتاج من الباحث جهدا كبيرا، إذ أنها تتطلب أن يوسع الباحث مجالات اطلاعه ومعارفه، لأن الأمر متعلق بميلاد فكرة وتخمين جديد في مجال البحث العلمي.

تعريف الفرضية: Hypothesis

هي ادعاء حول صحة شيء ما

والفرضية عبارة عن حدس علمي أو تفسير أولي للظاهرة ما، أو أنها تمثل العلاقة القائمة بين عدد من الظواهر و تقدم تفسيرا مبدئيا لها، فهي حل محتمل و مؤقت للمشكلة موضع الدراسة. و يعتمد في صياغة الفرضية على النظريات أو البحوث السابقة أو كليهما، ولكن صحتها تعتمد على مدى تأييد الأدلة والشواهد والبيانات الفعلية للفرضية التي تبحث فيها الدراسة.

وتوجد عدة أنواع من الفروض وهى:

أ- الفرض البحثى : Research Hypothesis يشتق الفرض البحثي عادة اشتقاقا مباشرا من إطار نظري معين ، و يحدد الإجابة المتوقعة للسؤال البحثي، ومن أمثلة الفروض البحثية :

- هناك علاقة بين أساليب المعاملة الوالدية والتوافق النفي
- هناك علاقة سالبة بين المستوى الاقتصادي والتعلم
- هناك علاقة موجبة بين الاكتئاب والأمراض السيكوسوماتية

ب- الفرض الإحصائي : Statistical Hypothesis عندما نعبر عن الفروض البحثية بصيغة رمزية وعددية ، فإنها تسمى عادة الفروض الإحصائية، الفروض الاحصائية نوعان هما:

- **الفرض الصفري** Null Hypothesis و يرمز له بالرمز H_o و يشير إلى النفي، هي الفرضية حول معلمة المجتمع التي نجري اختبار عليها باستخدام بيانات من عينة والتي تشير أن الفرق بين معلمة المجتمع والإحصائي من العينة ناتج عن الصدفة ولا فرق حقيقي بينهما. وهي الفرضية التي ننطلق منها ونرفضها عندما تتوفر دلائل على عدم صحتها، وخلاف ذلك نقبلها وتعني كلمة Nul انه لا يوجد فرق بين معلمة المجتمع والقيمة المدعاة (إحصائية العينة).
- **الفرض البديل** Alternative Hypothesis و يرمز له بالرمز H_1 و يشير الى الاثبات، هي الفرضية التي يضعها الباحث كبديل عن فرضية العدم و نقبلها عندما نرفض فرضية العدم باعتبارها ليست صحيحة بناء على المعلومات المستقاة من العينة.

مثال:

لا توجد فروق ذات دلالة إحصائية بين الذكور و الإناث في مهارة الإملاء(H_o)

توجد فروق ذات دلالة إحصائية بين الذكور و الإناث في مهارة الإملاء(H_1)

لا توجد علاقة ارتباطيه موجبة أو سالبة بين(H_o) .....

توجد علاقة ارتباطيه موجبة أو سالبة بين(H_1) .....

في الفروض الإحصائية نبدأ دائما بالفرض الصفري (لا توجد)، و في حالة عدم تحقق الفرض الصفري H_0 نؤكد الفرض البديل H_1

تعريف اختبار الفرضيات(Hypothesis Testing):

هو أحد أساليب الاحصاء الاستدلالي (inferential statistics) الذي تستخدم فيه بيانات العينة المسحوبة من مجتمع الدراسة لاتخاذ قرارات وإصدار أحكام حول قيمة معلمة (parameter) أو أكثر من معالم المجتمع.

إن الفكرة الأساسية في اختبار الفرضيات هي حساب الفرق بين المعلمة التي نفترضها للمجتمع (في الفرضية الصفرية) والقيمة المقابلة لها في العينة، أي التابع الإحصائي (الإحصاءة) وننسب هذا الفرق إلى الخطأ المعياري للتابع الإحصائي.

المعيار الذي يستطيع من خلاله الباحث الحكم على هذا الفرق تتم من خلال قسمة الفرق على الخطأ المعياري، ثم مقارنة خارج القسمة بالقيمة الجدولية أو ما يسمى بحدود منطقتي القبول أو الرفض.[4]

[4]: حسن محمد حسن: مبادئ الإحصاء الاجتماعي، دار المعرفة الجامعية، 2000.

الفروض الإحصائية التي تخضع للإختبار[5]:

الفرضية الصفرية: يرمز لها بالرمز H_0: تتضمن الهدف المطلوب للاختبار، وقبولها يعني عدم رفض نتائج العينة.

الفرضية البديلة: يرمز لها بالرمزH_1 : وتقبل في حال رفض H_0 .

مثال:

اختبار فرضية استخدام استراتيجية جديدة في التدريس لمادة الإحصاء، متوسط الطلاب وفق الاستراتيجية الحالية تساوي =70

الفرضية الصفرية: $H_0: \mu = 70$ $H_0 : \mu_1 - \mu_2 \leq 0$

الفرضية البديلة: $H_1: \mu \neq 70$ $H_1 : \mu_1 - \mu_2 > 0$

للتأكد من أن الطريقة الجديدة ستلقى بمستوى أداء الطلاب يجب اتخاذ قرار بهذا حتى لا يكون نتائج هذا العمل غير مفيد لاحقا.

وهنا نفرض أن الطريقة الجديدة ليست أفضل من الطريقة الحالية (الفرضية الصفرية H_0)

الطريقة الجديدة أفضل من الطريقة الحالية (الفرضية البديلة H_1)

قرارات اختبار الفرضية الصفرية:

إن اختبار الفرضية بأسلوب إحصائي يؤدي إلى اتخاذ قرار إذا ما كنت الفرضية مقبولة أم مرفوضة.

رفض الفرضية لا يعني بالضرورة أن تكون خاطئة

[5] : عبد الله عبد الحليم وأخرون: الإحصاء المفاهيم الأساسية، 2003،

كما أن قبول الفرضية لا يعني بالضرورة أن تكون صحيحة[6]

القرار	الفرضية الصفرية	
	صحيحة	خاطئة
قبول الفرضية	صحيحة	خطأ من النوع الثاني $(\beta - 1)\ \beta$
رفض الفرضية	خطأ من النوع الأول $(\alpha - 1)\ \alpha$	صحيحة

مع تحديد مستوى الدلالة α(0.05 أو 0.01 ...) أو غير ذلك ، وممكن أن نرفع مستوى الدلالةα إن كنا حريصين على عدم قبول الفرضية الخاطئة وهو احتمال رفض الفرض H_0وهو صحيح، وهو ما يعرف باحتمال الوقوع في الخطأ من النوع الأول (Type1)، أوالعكس بقبول H_0 وهو خاطئ ويعرف باحتمال الوقوع في الخطأ من النوع الثاني (Type2)

أنواع الخطأ وقوة الاختبار[7]:

إن تخفيض الوقوع في الخطأ من النوع الأول (بزيادة مستوى الدلالة) يزيد من فرص الوقوع في الخطأ من النوع الثاني (أي قبول الفرض الصفري بينما هو خاطئ)، والتقليل من الخطأ الثاني يأتي بزيادة حجم العينة بهدف الحصول على قوة اختبار عالية).

[6] : فاروق عبد العظيم وآخرون: مبادئ الإحصاء، دار المعرفة الجامعية،
[7] : فتحي عبد العزيز أبو راضي: مبادئ الإحصاء الاجتماعي، دار المعرفة الجامعية،

تقليل أحد الخطأين يؤدي لزيادة الخطأ الآخر (توجد علاقة عكسية بينهم).

الخطأ الأول أكثر خطورة من الخطأ الثاني.

الموازنة بين الخطأين تستلزم تقدير الباحث للتكلفة المترتبة على الوقوع في أي من الخطأين.

<u>قوة الاختبار:</u>

يعرف بـβ وهو يعتمد على : الابتعاد عن H_0، حجم العينة n، الانحراف المعياري للمجتمع δ، مستوى المعنوية α، نوع الاختبار (جانب أو جانبين)

$$\beta = \sqrt{n}\,(\mu - \mu_0)\delta$$

خطوات اختبار الفرضيات[8]:

تحديد نوع توزيع المجتمع	توزيع طبيعي: اختبارات بارامترية	توزيع غير طبيعي: اختبارات لابارامترية
صياغة الفرضية	الفرض الصفري H_0	الفرض البديل H_1
اختيار مستوى الدلالة الاحصائيةα	0.05	0.01
اختيار دالة الاختبار المناسبة	جمع البيانات من العينة	حساب دالة الاختبار الإحصائية
اتخاذ القرارات	قبول الاحتمال $\leq$ مستوى الدلالة	رفض الاحتمال >مستوى الدلالة

[8] : خليفة عبد السميع خليفة: الإحصاء التربوي، مكتبة الأنجلو المصري،

الدلالة الاحصائية:

دائما ما يستخدم الباحثين مستوى الدلالة في اختبار الفرضيات (hypotheses) حيث لا توجد فرضية تخلو من مستوى الدلالة.

باختصار إن الدلالة الاحصائية هي قيمة احتمالية يمكن أن نوضحها بالمثال التالي:

عندما يريد أي جراح أن يُجري عملية ما، فإنه يشير إلى أهل المريض بقوله إننا نريد أن نعاين الحالة، فإذا كان احتمال الفشل لدينا أكثر من 10%، فإننا لن نُجري العملية، ولكن إذا كان معدل الفشل أقل من 10%، فإننا سوف نجري هذه العملية. أي أن إجراء وعدم إجراء العملية مرتبط بقيمة معنوية هي 10% مثلا، وكتطبيق على هذا المثال استخدم الاحصائيون مستوى الدلالة أو القيمة المعنوية وهي قيمة احتمالية لحدث ما لمحاكمة نتيجة أي دراسة يقوم الباحثون بتطبيقها.

بتفسير بسيط فإن الدلالة الاحصائية تقع في ثلاثة مستويات:

- المستوى الاول: مستوى اقل من 0.05
- المستوى الثاني: مستوى اقل من 0.01
- المستوى الثالث: مستوى اقل من 0.001

والمستوى الثالث اقوى في دلالته من المستوى االاول والثاني

والمستوى الثاني اكثر في قوته من المستوى الاول فقط

بينما المستوى الثالث هو اقل مستواى دلالة احصائية مقبولة.

وفي اوقات كثيرة يُعبر عن مستوى الدلالة بمستوى الخطأ المحتمل.

ماذا يعني ذلك وما معنى هذه الدلالة؟

عندما نقول هذا دال عند مستوى 0.05، فان هذا يعني أننا لو أعدنا هذا البحث أو هذا الإختبار 100 مرة فاننا سوف نحصل على نفس النتيجة 95 مرة، وسوف تكون نسبة الخطأ في أن نحصل على نتيجة مختلفة هو خمسة مرات من اصل المئة (أي 0.05.(

وكذلك الأمر بالنسبة لمستوى الدلالة 0.01، فانه يعني انني سوف احصل على نفس النتيجة 99 مرة من أصل الـ100 مرة التي أعيد فيها التجربة أو الاختبار . بينما نسبة الخطأ تقل بكثير لتكون مرة واحدة فقط في كل 100 مرة.

وبالطبع فان نسبة الدلالة 0.001 هي الأقوى، لأنها تعني أننا لو أعدنا تطبيق هذا البحث أو هذه التجربة ألف مرة (1000 مرة) فاننا سوف نحصل على نفس النتيجة 999 مرة في مقابل مرة واحدة خطأ.

ويمكن تلخيص ما سبق بقولنا أن مستوى الدلالة الإحصائية أو مستوى الاحتمال، هو درجة الإحتمال الذي تُرفض به الفرضية الصفرية H0 عندما تكون هي الصحيحة، أي هو احتمال الوقوع في الخطأ من النوع الأول (ألفا)، وهي ثلاثة أنواع:

-دال عند 0.05 أي مستوى الثقة 95% والشك 5.%

-دال عند 0.01 أي مستوى الثقة 99% والشك 1.%

-دال عند 0.001 أي مستوى الثقة 99.9% والشك 0.1.%

أنواع اختبارات الفروض حسب الاتجاه:

عندما نقبل الفرضية المبدئية فإننا نقبلها بنسبة دقة 90% أو 95% أو 99% أو غير ذلك وتسمى مستويات الثقة Significance Levels أي يوجد نسبة خطأ معين في قبولنا للفرضية المبدئية بمعنى أننا نقبل صحة الفرضية المبدئية وهي خاطئة وهذا الخطأ هو الخطأ α ويسمى مستوى المعنوية، أي إذا كان مستوى الثقة 95% (1- α) فان مستوى المعنوية α تساوي 5% وهي عبارة عن مساحة منطقة تحت منحنى التوزيع تمثل منطقة الرفض وتكون أما على صورة ذيل واحد جهة اليمين أو اليسار أو ذيلين متساويين في المساحة واحد جهة اليمين والثاني جهة اليسار.

- **تعريف اختبار الفروض في جانب واحد:**

هو الاختبار الذي تبين فيه الفروض البديلة أن المعلمة للمجتمع اكبر أو اصغر من إحصائية العينة، فهناك تحديد للاتجاه.

- **تعريف اختبار الفروض في جانبين(ذيلين):**

هو الاختبار الذي لا تبين فيه الفرضية البديلة أن معلمة المجتمع أكبر أو أصغر من إحصائية العينة، بل مجرد أنها تختلف .

أسس اختيار الاختبار المناسب لفروض الدراسة:

هناك مجموعة من الخطوات التي يجب على الباحث القيام بها من أجل إجراء تحليل إحصائي مميز لبحثه العلمي.

-2ففي البداية يجب على الباحث أن يقوم باختيار نوع الاختبار الإحصائي، ويتم هذا الاختيار وفق عدد من الأسس ومن أبرز هذه الأسس:

- نوع البيانات الخاصة والتي تتعلق وترتبط بشكل مباشر بالمتغيرات التابعة.
- نوع العلاقات التي يرغب الباحث في اختبارها، وإجراء التحليل الإحصائي لها.
- تحديد عدد المتغيرات المستقلة التي سيقوم الباحث بدراستها.
- ومن ثم تحديد عدد مستويات المتغيرات المستقلة

وبعد أن يختار الباحث نوع الاختبار الإحصائي يجب عليه أن يمتلك القدرة الكافية على التمييز بين الاختبارات المعلمية والاختبارات غير المعلمية، وسنوضح هذا الفرق فيما يلي :

- للاختبارات المعلمية ملامح معينة ومن أبرز هذه الملامح تحقيق الفرضية التي تنص على أن يكون نوع البيانات التي يقوم الباحث بإجراء التحليل الإحصائي لها في مستوى مقياس الفترة، كما يجب أن يكون توزيع مجتمع الدراسة توزيعا طبيعيا، بالإضافة إلى ذلك فأن مجتمع البحث يجب أن يحتوي على نفس الاختلافات التي توجد في عينة البحث التي سيتم إخضاعها للتحليل الإحصائي
- أما بالنسبة للاختبارات غير المعلمية فإن لها مجموعة من الملامح كأن تحقق الفرضية التي تنص على أن يكون نوع البيانات في مستوى مقياس رتبي وحسب، بالإضافة إلى ذلك فإن مجتمع الدراسة يجب أن يكون موزعا توزيعا حرا.

وبعد أن يطلع الباحث على يطلع الباحث على الاختبارات المعلمية والاختبارات غير المعلمية يجب عليه أن يختار نوع الاختبار المناسب، وتمتلك الاختبارات المعلمية مجموعة من المميزات عن الاختبارات غير المعلمية ومن أبرز هذه المميزات:

- الاختبارات المعلمية أكثر قوة في حال تمت مقارنتها بالاختبارات غير المعلمية
- الاختبارات المعلمية تمتلك قدرة كبيرة على تحديد جميع الدلالات التي تعنى بالاختلافات المهمة.
- الاختبارات المعلمية تقوم باستخدام جميع المعلومات الموجودة في البيانات المجموعة
- لا تهتم الاختبارات غير المعلمية سولا بترتيب البيانات فقط

بعد ذلك يجب على الباحث أن يقوم باختبار فرضياته، ويتم هذا الأمر من خلال قيامه بعدد من الخطوات ومن أبرز هذه الخطوات:

- تحديد فرض العدم(Null Hypothesis)
- تحديد الفرض التجريبي(Experimental Hypothesis)
- اختيار مستوى الدلالة المناسب(Level of Significant)

ومن ثم يجب على الباحث أن يقوم بتحديد مستوى الدلالة الإحصائية، ويوجد للدلالة الإحصائية مستويين وهما:

- اتجاه الاختبار: حيث يجب أن يحدد الباحث اتجاه الاختبار سواء أكان هذا الاختبار في اتجاه واحد أو اتجاهين، ويعتمد اختيار اتجاه الاختبار على المؤثر الذي يريد الباحث دراسته، ويعد اختيار اختبار الاتجاه الواحد مناسبا في حال تم تحديد اتجاه تأثير المتغير

المستقل، سواء أكان هذا الاتجاه تصاعديا أو تنازليا، بينما يكون اختيار الاختبار ذو الاتجاهين هو المناسب في حال لم يتم تحديد اتجاه تأثير المتغير المستقل

- درجة الحرية: ويعني بها عدد القيم التي تقبل إجراء التغييرات عليها في حال تم القيام بحساب خاصية إحصائية معينة، ولتحديد هذه الحرية يجب على الباحث أن يتبع معادلة معينة

متى تستخدم الاختبارات اللابارميترية لقياس الفرق بين عينتين؟

تسخدم الاختبارات اللابراميترية للكشف عن دلالة الفروق بين متوسطي عينتين عندما:

-1لا تتوفر شروط استخدام اختبار " ت " كأن تكون مفردات العينتين صغيرة

-2عندما يكون توزيع أحد العينيتين غير اعتدالي أو ملتوى بدرجة كبيرة

-3عندما يكون تباين العينتين Sample variance مختلف بصورة كبيرة عن بعضها

وهنا يفضل استخدام الرتب فضلا عن القيم الأصلية في حساب دلالة الفروق بين متوسطي عينتين كما يمكن اسنخدام الرتب في حساب معامل الارتباط أو قوة العلاقة بين متغيرين

من أشهر الاختبارات البارامترية:

اختبار T واختبار Z

ومن أشهر الاختبارات اللابراميترية ما يلي:

اختبار كا2: The chi-Square Test

ومن الاختبارات اللابراميترية الأكثر شيوعا هو اختبار كا2 ويستخدم هذا الاختبار عندما تقارن قيمة كا2 المحسوبة (observed) بقيمة كا2 النظرية (الجدولية) بدرجة جرية مقدارها 1 وعند نسبة خطأ = 1 فإذا كانت كا2 المحسوبة تساوي أو أكبر من الجدولية فمعنى ذلك أن هناك ارتباط بين المتغير الأول والثاني ومن ثم يمكن رفض الفرض الصفري أما إذا كانت كا2 المحسوبة أقل من كا2 الجدولية فلا وجود لهذه العلاقة أو أن هذين المتغيرين مستقلان عن بعضهما البعض كما تستخدم كا2 لاختبار مدى اتفاق توزيع القيم مع التوزيع المتوقعGoodness of Fit

اختبار مان وتني يو: The Mann-Whitney U Test

يتشابه هذا الاختبار مع اختبار T-test وهو من المقاييس البراميترية ولكن اختبار مان ونتي يو يعد من الاختبارات اللابراميترية ويستخدم عادة لبيان عما إذا كان وسيط كل من عينيتن مستقلتين يختلفان عن بعضها البعض اختلافا جوهريا

اختبار ويلكوكسون للفروق بين رتب قيم مرتبطة: TheWilcoson runk sum test

وهو اختبار لابراميتري بديل عن اختبار " ت " للقيمة المرتبطة إذا لم تستوفى البيانات التي بين أيدينا الشروط الواجب توافرها لاستخدام اختبار "ت" للقيمة المرتبطة ويمكن استخدام هذا الاختبار عندما تكون البيانات معبرا عنها في شكل رتب لاختبار الفرض بأن عينات الدراسة لها نفس توزيع المجتمع الذي تم سحب العينات منه

معامل ارتباط سبيرمان: The Spearman Rank Order Correlation

احيانا يسمى(Supearman's rho) وهو اختبار لابراميتري يستخدم في الحالات التالية:

·إذا كان المتغيران كل منها ينقسم إلى فئات منفصلة كثيرة

·إذا كان المتغيران ينقسمان إلى فئات ونريد الاستعانة يرتب هذه الفئات عن الفئات نفسها

·إذا كان المتغيران كل منها متغير متصل ولكننا نفضل استخدام الرتب بدلا من القيم الخام لكل متغير

اختبار كيلموجوروف-سمرنوفThe kilmogorow Smirnow :

يقوم بنفس عمل كا2 عند اختبار التوزيع المتوقع كما يقوم بنفس عمل اختبار ويلكوكسون للفروق بين رتب القيم لتحديد عما إذا كانت العينات العشوائية هي من نفس عينة الدراسة

اختبار كروسكال – والاس: The Kruskal Wallis test

ويسمى أحيانا كروسكال – والاس (H) عندما يجري الباحث تجربة ولكن بياناته التي حصل عليها لا تتوزع توزيعا اعتداليا أو أن تكون العينات صغيرة عندئذ يلجأ إلى استخدام كروسكال والاس لدراسة بين العينات

معامل اتفاق كندال:

وهو إجراء يمكن استخدامه عندما يكون هناك رتبا قام بها بعض الحكام المستقلين والمطلوب معرفة مدى اتفاقهم في تحديد الرتب ولقياس دلالة (W) المعنوية يمكن استخدام كا2

المحور الثالث:

المعالجة الآلية للمعطيات

الدرس 10: المعالجة الآلية للبيانات باستخدام برنامج الحزمة الإحصائية spss

تمهيد:

التحليل الإحصائي للبيانات هو عصب العملية البحثية ومؤشر نجاح البحث ووصول الباحث لنتائج دقيقة قائمة على تحليل إحصائي منطقي للبيانات التي تم جمعها، لذا على الباحث إجراء خطوات التحليل الإحصائي للبيانات بشكل صحيح مع الاعتماد على برامج الإحصاء واستخدماها بشكل صحيح.

يجب على كافة الباحثين التأكد من الإلمام بطبيعة برامج التحليل الإحصائي للبيانات بحيث يضمن استخدامها بشكل سليم.

وسائل المعالجة الآلية للبيانات

تتوافر العديد من طرق إجراء التحليل الإحصائي والتي تطورت كثير مع الطفرات التكنولوجية الهائلة التي لحقت بكافة المجالات وخاصة التطورات التقنية في برامج الحاسوب والتي مكنت الباحثين من إجراء التحليل الإحصائي للبيانات بدقة عالية.

على الباحث تجنب الخطأ في استخدام التطبيقات بحيث يتجنب الخطأ في النتائج الخاصة بالتحليل الإحصائي للبيانات ومن أهم البرامج المستخدمة في التحليل الإحصائي للبيانات-:

- برنامج SPSS للتحليل الإحصائي.
- برنامج SAS للتحليل الإحصائي.
- برنامج EVIEWE للتحليل الإحصائي.
- برنامج.MINITAB
- برنامج R

وغيرها.....

لمحة عامة عن برنامج spss:

SPSS هي اختصار للأحرف اللاتينية الأولى من اسم "الحزم الإحصائية للعلوم الاجتماعية"، وهي حزم حاسوبية متكاملة لإدخال البيانات وتحليلها. وتستخدم عادة في جميع البحوث العلمية التي تشمل على العديد من البيانات الرقمية ولا تقتصر على البحوث الاجتماعية فقط بالرغم من أنها أنشأت أصلاً لهذا الغرض ، ولكن اشتمالها على معظم الاختبارات الإحصائية (تقريباً (وقدرتها الفائقة في معالجة البيانات وتوافقها مع معظم البرمجيات المشهورة جعل منها أداة فاعلة لتحليل شتى أنواع البحوث العلمية

وتستطيع SPSS قراءة البيانات من معظم أنواع الملفات لتستخدمه الاستخراج النتائج على هيئة تقارير إحصائية أو أشكال بيانية أو بشكل توزيع اعتدالي أو إحصاءاً وصفياً بسيطاً أو مركباً وتستطيع الحزم جعل التحليل الإحصائي مناسباً للباحث المبتدئ والخبير على حد سواء

تشغيل البرنامج:

يمكن تشغيل برنامج SPSS بإحدى الطريقتين التاليتين:

- عن طريق النقر على قائمة start ثم التأشير على الأمر all programs ثم النقر على الاسم. spss for windows

- أو عن طريق النقر المزدوج على الأيقونة الموجودة غالبا على سطح المكتب.

نوافذ البرنامج:

يتألف برنامج SPSS من عدة نوافذ هي[9]:

1-نافذة محرر البيانات Data Editor : من أهم النوافذ حيث يتم إدخال البيانات فيها قبل إجراء الدراسة الإحصائية

[9]: غيث البحر ومعن التنجي: التحليل الإحصائي للاستبيانات باستخدام برنامج IBM SPSS Statistics، مركز سبر للدراسات الإحصائية والسياسات العامة، تركيا، 2014، ص7.

2-نافذة المخرجات Viewer : تعرض هذه النافذة نتاج الإحصاءات التي نجريها على البيانات من جداول ورسوم بيانية.

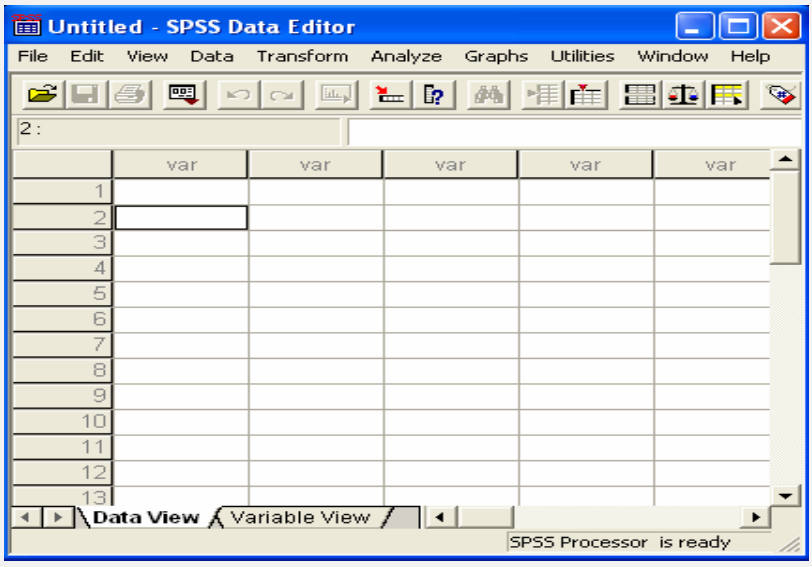
2-محرر الجداول Pivot Table Editor : عند النقر المزدوج على أي جدول في نافذة المخرجات تظهر نافذة جديدة تتيح لنا إمكانية التعديل على الجدول بحيث يصبح بالشكل الذي يناسبنا.

2-محرر الرسوم Chart Editor : للتعديل على شكل أي رسم بياني في نافذة المخرجات ننقر عليه نقرا مزدوجا فيظهر محرر الرسوم البيانية الذي يساعدنا على إجراء تغييرات كبيرة على الرسم البياني.

ملاحظة :يحوي برنامج SPSS على نوافذ أخرى لكننا لن نأتي على ذكرها لأنها لا تخدمنا كثيرا في عملية تحليل البيانات.

الأوامر المهمة التي تحتويها القوائم المختلفة لبرنامج spss

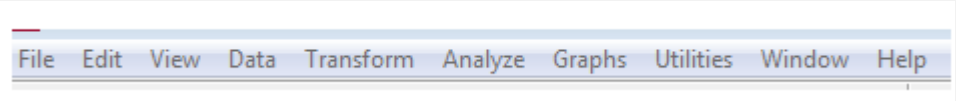
عند فتح البرنامج تظهر لنا شاشة تتضمن جملة من الأجزاء التي يحتاج المستخدم لمعرفتها كالتالي:



1. لائحة الأوامر
2. المفاتيح المساعدة
3. شاشة البيانات
4. شاشة تعريف المتغيرات
5. لائحة التقارير والمخرجات

1- <u>لائحة الأوامر:</u>

تظهر بالشكل التالي وتتضمن:



- **الأمر File(ملف):** لفتح وحفظ الملفات وقراءة بيانات من جداول إلكترونية (مثل اكسل) وطباعة البيانات.
- **الأمر Edit (تحرير):** يقص وينسخ ويلصق القيم، وللحصول على قيم بيانات ولتغير الخيارات.
- **الأمر View (عرض):** للتحكم في شكل القيم وشرحها، كتعديل بعض الأوامر في شريط الأدوات Toolbar والقوائم Menu، وإظهار وإزالة الخطوط الشبكية Gridlines، والتحكم في عرض وصف قيم المتغيرات Value Labels
- **الأمر Data (بيانات):** لعمل تغير شامل على ملف البيانات، كإدخال متغير جديد أو حالة جديدة أو دمج بيانات ملفين.
- **الأمر Transform (تحويل):** لعمل تغير لمتغيرات محددة في ملف البيانات ولحساب متغيرات جديدة بناء على قيم موجودة.
- **الأمر Analyze (تحليل):** لاختيار مجموعة كبيرة ومتباينة من العمليات والاختبارات الإحصائية مثل اختبارات وتحليل التباين والاختبارات اللامعلمية . ويعتبر هذا الخيار بيت القصيد من الحزم كلها ويشمل أكبر كمية من الخيارات الضمنية .
- **الأمر Graphs (الرسوم البيانية):** لإعداد رسوم بيانية بأنواعها : طولي ، دائري ، نقطي .....الخ
- **الأمر Utilities (أدوات):** للحصول على معلومات عن المتغيرات، وللتحكم في ظهور متغيرات معينة في مربع الحوار، وللتحكم في شاشة العرض الرئيسة.

- **الأمر Window (نافذة):** للتنقل بين شاشات البرنامج النشطة وغيرها، أو لتصغير جميع نوافذ SPSS المفتوحة.
- **الأمر Help (مساعدة):** للحصول على الصفحة الأساسية للبرنامج (internet Home Page) أو الدخول على شاشة المساعدة في العديد من أوجه SPSS، ويمكن الحصول على المساعدة أيضا بنقر زر الفأرة الأيمن في المكان الذي تريد الحصول على مساعدة فيه.

2- المفاتيح المساعدة:

فتح ملف

حفظ ملف

طباعة

عرض آخر عملية تم حسابها

تراجع

إعادة

الانتقال إلى حالة محددة

الاستعلام عن خصائص متغير

البحث عن قيمة محددة

إدراج حالة جديدة

إدراج متغير جديد

تقسيم مخرجات الأوامر

الترجيح

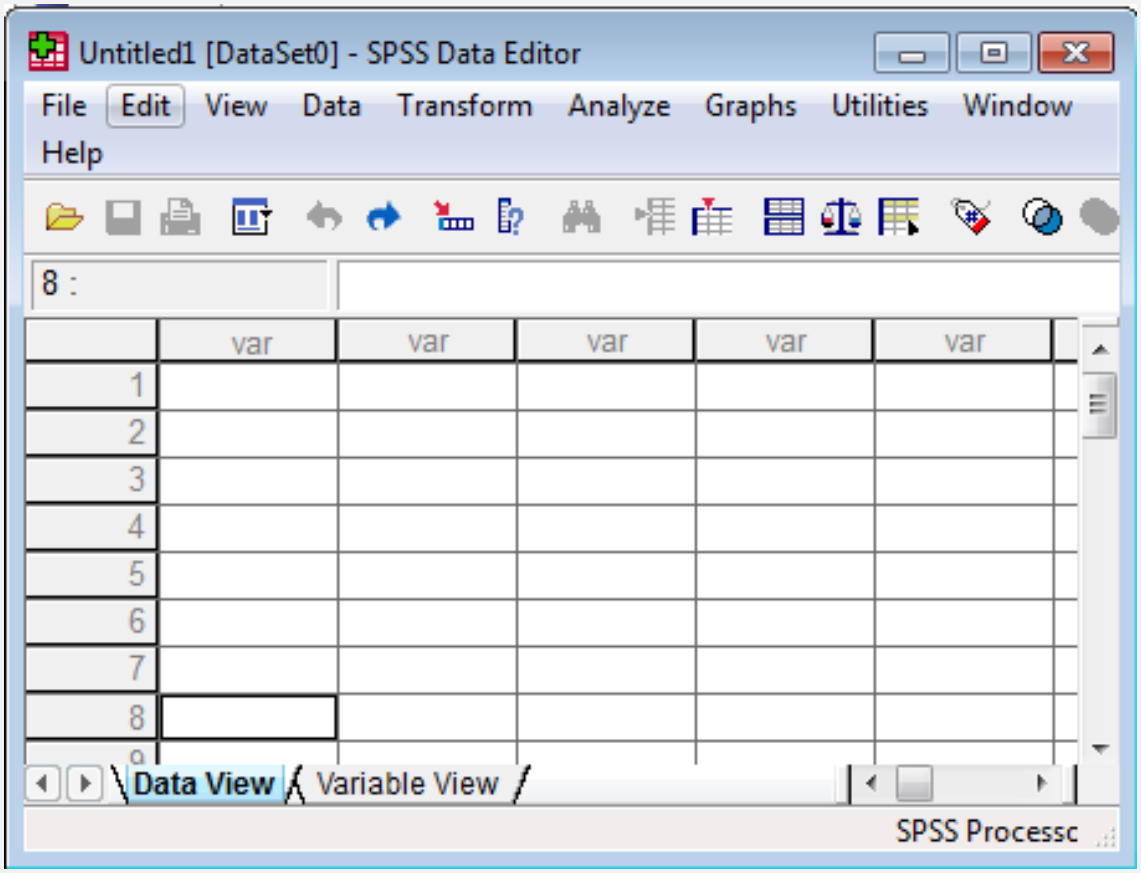
اختيار حالات

عرض قيم المتغيرات (رموز/أرقام)

تحديد فئات محددة للاستعمال

3- شاشة البيانات Data View:

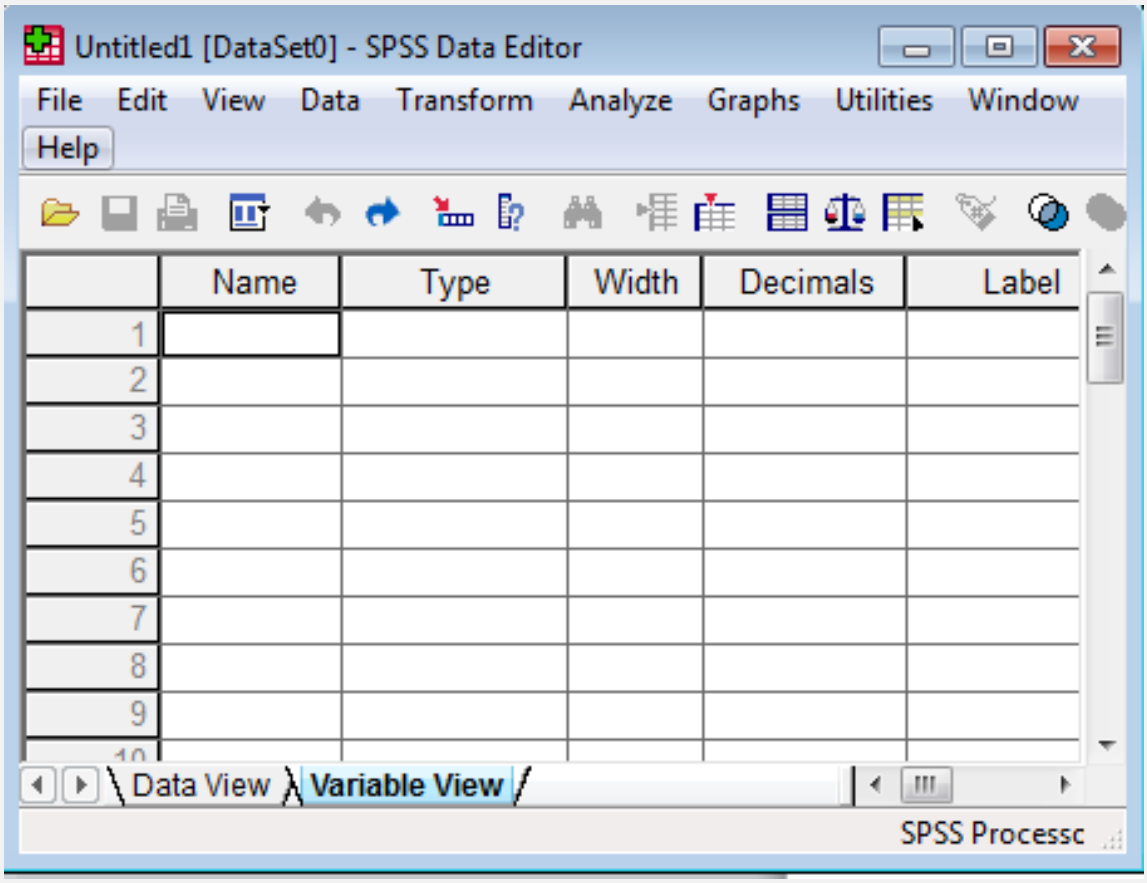
تظهر هذه الشاشة بالشكل التالي:



تستخدم هذه الشاشة لإدخال البيانات (كما سنرى ذلك في درس لاحق)، وهي عبارة عن شبكة من الأعمدة والصفوف، حيث يمثل كل عمود في هذه الشاشة متغيرا معينا من متغيرات البحث، بينما يمثل كل صف حالة أو وحدة من وحدات العينة.

4- شاشة تعريف المتغيرات Variable View:

وتظهر هذه الشاشة بالشكل التالي:



تستخدم هذه الشاشة لتعريف المتغيرات مثل اسم المتغير ونوعه...(كما سنرى ذلك في درس لاحق)، وهي عبارة عن شبكة من الأعمدة والصفوف، حيث يمثل كل صف في هذه الشاشة متغيرا معينا من متغيرات البحث، بينما يمثل كل عمود صفة أو خاصية من خصائص هذا المتغير.

لائحة المخرجات:

وهي عبارة عن شاشة تظهر فيها نتائج العمليات التي تم القيام بها. ويتم فتح النافذة أو الملف تلقائيا بمجرد تنفيذ الأوامر الإحصائية مثل جدول التوزيع التكراري أو رسم بياني.

ومن نافذة المخرجات يمكن الدخول إلى ثلاث شاشات هي[10]:

- محرر الرسوم البيانية Chart Editor.
- محرر النصوص Text Editor .
- محرر الجداول Pivot Editor .

يعتبر محرر بيانات الـ SPSS الواجهة الأولية للحزم ، وهي واجهة تشبه الجداول الإلكترونية وتستخدم لإدخال البيانات الخام لأول مرة . ومن خلال المحرر يمكن قراءة البيانات وتعديلها أو تغيرها التعامل مع المتغيرات وتسميتها أو تغير أسمائها ومن خلال محرر البيانات تحفظ ملفات البيانات وتسمى ملفات بيانات Data files ولا يستطيع هذا الملف استخراج أي نوع من النتائج ، وإنما النتائج ترسل إلى نوع آخر من الملفات وهي ملفات المخرجات.

وملفات المخرجات Output files تحوي على جميع النتائج التي تتم بعد أي عملية إحصائية، وفي كل مرة يطلب البرنامج من المستخدم حفظ الملف أو حذفه ،ويوصى بعدم حفظ جميع ملفات المخرجات إلا ما يحتاجه الباحث أو المستخدم بصفة مستمرة وبعد أن يتأكد من صحة النتائج أما ملفات البيانات فإنه يجب حفظها بأكثر من ملف والحفاظ عليها نظراً لان فقدها يؤدي إلى إعادة الإدخال كاملاً بعكس ملفات المخرجات التي لا يتطلب استرجاعها سوى استرجاع العملية الإحصائية، وطلب النتائج من البرنامج . وفي النسخ الأخيرة من الـ SPSSيمكن التعامل مع المخرجات (بيانات أو رسومات) وتعديلها في نظام شجري جميل وسهل يمكن التحكم فيه بكل يسر وسهولة

.

---

[10]. ؟؟؟؟؟؟؟؟؟؟؟؟؟ : مرجع سبق ذكره، ص 11.

ومن خلال قائمة الأوامر وخيارات البرنامج يستطيع الاختيار بين العديد من عمليات تعديل البيانات وتشكيلها وبين الاختبارات الإحصائية المتعددة وأنواع كثيرة من الرسوم البيانية الجميلة.

وعموما: فإنه يمكن إجمال مراحل تحليل البيانات بالخطوات التالية:

-1ترميز البيانات.

-2إدخال البيانات في الـ.SPSS

-3اختيار الاختبار أو الشكل المناسب.

-4تحديد المتغيرات المراد تحليلها.

يتضمن هذا البرنامج العديد من الإجراءات الاحصائية الشائعة التي يستخدمها الباحثون في ميدان العلوم الاجتماعية عند دراسة الظواهر المطروحة أمامهم وسوف نبدأ بتلك الإجراءات التي يبدأ بها عادة الباحثون كخطوة أولة ثم يليها استخدام إجراءات احصائية أكثر تعقيدا وتميزا فلا يوجد بحث اجتماعي يستخدم كل الإجراءات التي سوف نشرحها مرة واحدة وإنما عادة ما يستخدم الباحث إجراء او طريقة من الطرق الإحصائية منفردة في أي وقت من الأوقات

-1التوزيع التكراري لمتغير واحد وقياس النزعة المركزية والتشتت One- Way Frequency Distribution, Measures of Central Tendency and Dispersion :

في معظم البحوث الاجتماعية فإن أول وظيفة للباحث هو فحص الصفات التوزيعية لكل المتغيرات المستقلة Independent Variable أو المتغيرات التابعة Dependent Variablesتحت الدراسة ولتحقيق هذا الغرض فإن برنامج SPSS يحتوى على إجراء احصائي يسمى: Descriptive Statistics

أ : FREQUENCIES-الذي يقوم بالحساب الوصفي لينتج تقارير مجدولة للتوزيع التكراري البسيط لعدد غير كبير من الحالات المراد دراستها

ب : DESCRIPTIVES-الذي يقوم بحساب عديد من مقاييس النزعة المركزية الشائعة والتشتت للمتغيرات التي قيست باستخدام المقياس الفئوي للوحدات المتساوية والتي تجمع عددا كبيرا من الحالات التي يراد دراستها

ومن الأمثلة التي يتم استخدام النوع الثاني من الإجراءات الإحصائية فيها عددا يقوم الباحث بدراسة الدخل بالجنيه المصري والذي يفترض أنه عبارة عن قيمة متسلسلة a continuous valueوعندما يتم وضع القيم في فئات مجمعة كما هو الحال في المثال التالي : 1001+ ، 1000 -501 ، 500-401 ، 400-301 ، 300-201 ، 200- كما يتيح إجراء FREQUENCIESالحصول على توزيع وصفي لمتغيرات اسمية كما هو الحال في توزيع العينة حسب الديانة ، الجنس ، أو الانتماء الحزبي

ويوفر إجراء Frequencies التحاليل الاحصائية التالية:

-المتوسط الحسابيMein

-الوسيطMedian

-المنوالMode

-الانحراف المعياريStandard Deviation

-تحليل التباينAnalysis of Variance

-تحليل الالتواءSkewness

-تحليل التفرطحKurtosis

-تحليل المدى أو النطاقRange

-الخطأ المعياريStandard Error

كما يوفر إجراء FREQUENCIES الرسوم البيانية مثل الأعمدة البيانية والدوائر البيانية والمضلع التكراري

-2دراسة العلاقات بين متغيرين فئويين أو أكثر Relationship between Two or More Categorical Variables :

بعد أن يقوم الباحث بفهم خصائص كل متغير من متغيرات الدراسة فإن أول عمل يقوم به هو فحص مجموعات من العلاقات وهنا يختار إجراء أو أكثر لدراسة تلك العلاقات ويتوقف هذا بالطبع على خصائص المغيرات وأغراض البحث فالباحث قد يختار معامل الارتباط لدراسة هذه العلاقات أو قد يلجأ إلى عرض نتائج الدراسة في شكل جداول وخاصة إذا كانت المتغيرات اما اسمية أو تعكس ترتيبا معينا

ومن الإجراءات التي تتيح للباحث تحليل العلاقة بين متغيرين أو أكثر استخدام ما يسمى CROSSTABSحيث يمكن تبويب وفقا لمتغيرين فئويين وهذا الإجراء يمكن الباحث من تحليل العلاقة بين متغيرين أو أكثر من متغير بإنتاج تبويب مزدوج Cross robulation يوضح به التوزيع التكراري لمتغيرين ويمكن التعبير عن هذا التوزيع باستخدام النسب المئوية من المجموع الأفقي أو المجموع العمودي أو كنسبة من المجموع الكلي للجدول والتحليل الاحصائي الذي يستخدم عادة لدراسة العلاقة بين متغيرين من توزيع تكراري هو اختيار كا تربيع Chi-square الذي يظهر مدى الارتباط بين المتغيرين

-3تحليل الارتباط أو العلاقة بين متغيرين:

أن تحليل العلاقة بين متغيرين باستخدام إجراء Correlate الذي يمكن الباحث من استخدام طريقة لقياس الارتباط المستقيم أو الخطي Linear Relationship بين متغيرين وينتج عن هذه العملية قيمة احصائية توضح قوة واتجاه العلاقة بين المتغيرين وتعرف هذه القيمة باسم معامل الارتباط Correlation Coefficient ويحتوي برنامج SPSS على طريقتين احصائيتين لحساب هذه العلاقة

وينتج الاجراء الاحصائي الأول ارتباط بيرسون Pearson's الذي يناسب البيانات المستمدة من التوزيع المعتدل Normal Distribution والتي تخضع للمقياس Interval Scale أما الإجراء الاحصائي الثاني NOPAR CORR اللابراميتري فهو يناسب البيانات التي تخضع للمقياس الترتيبي Ordinal Scale حيث تحتوي على فئات كثيرة عن جداول ذات تبويب مزدوج Cross Tabulation وتمكن الباحث من استخدام معامل ارتباط سبيرمان أو معامل الارتباط المعروف باسم كندال Kendall rank-order correlation أو كليهما

ويتيح الإجراءين الاحصاءين في برنامج ال SPSSPEARSON CORR- NONPAR CORRلزوجين مختارين من قائمة المتغيرات بالإضافة إلى مصفوفة متكاملة من معاملات الارتباط والتي من الممكن استخدامها عند تطبيق الطرق الاحصائية على المتغيرات المتعددة Multivariateوبالرغم من أن معامل الارتباط بين متغيرين يعطي ملخص احصائي منفرد لوصف العلاقة بين متغيرين فيوجد كثيرا من المواقف قد يرغب فيها الباحث في دراسة هذه العلاقة بشكل موسع ومفصل ويوجد في برنامج SPSS برنامج فرعي يعرف باسم SCATTEGRAMالذي ينتج رسما بيانيا بالنقاط المبعثرة Scatter plot Diagram بين المتغيرين وبذلك يمكن رؤية نمط العلاقات بشكل واضح وبالإضافة إلى هذا الرسم البياني فيوجد أيضا طرق أخرى كمعامل ارتباط بيرسون ، الخطأ المعياريStandard error

-4معامل الارتباط الجزئي: Partial Correlation

لدراسة العلاقة بين متغيرين لابد من إيجاد مقياس تقيس به مثل هذه العلاقة وهذا المقياس هو ما يسمى بمعامل الارتباط ولحساب معامل الارتباط ينبغى أن يكون لدينا متغيرات لتقيس قوة واتجاه العلاقة بينهما ومعامل الارتباط الجزئي يدرس العلاقة بين متغيرين بينما نتحكم في تأثير متغير آخر أو أكثر وفي هذه الحالة نجد أن معامل الارتباط الجزئي هو في الواقع يشبه التبويب المزدوج للمتغيرات المتصلة والبرنامج المستخدم هو إجراء PARTIAL باستخدام اما بيانات خام أو مصفوفة من معاملات الارتباط البسيط التي تم الحصول عليها باستخدام PEARSON CORRوPARTIAL CORR

-5معامل الارتباط المتعدد وتحليل الانحدار Multiple Correlation and Regression :

معامل الارتباط المتعدد هو بمثابة امتداد لمعامل الارتباط الجزئي إلى التحليل المتعدد ويمكن تحليل الانحدار المتعدد Multiple Correlation لدراسة العلاقة بين مجموعة من المتغيرات المستقلة Independem ومتغير تابع Dependem بينما يؤخذ التحليل في اعتباره العلاقات بين المتغيرات المستقلة والهدف الأساسي هذا هو استنتاج مدى تأثر المتغيرات المستقلة والهدف الأساسي هذا هو استنتاج مدى تأثير المتغيرات المستقلة والهدف الأساسي هذا هو استنتاج مدى تأثير المتغيرات المستقلة مجتمعة على المتغير التابع كما

يمكن استخدام العلاقات بين المتغيرات المستقلة في التنبؤ بقيم المتغير التابع وتحديد مدى أهمية كل متغير من المتغيرات المستقلة في هذا التنبؤ ويمكن الحصول على عدد لا بأس به من معاملات الارتباط المتعددة وتحليل الانحدار باستخدام الإجراء الاحصائي REGRESSIONويمكن القيام بهذه التحليلات الاحصائية باستخدام مادة أولية خام Raw Dataأو استخدام مصفوفة من معاملات الارتباط

كما يمكن هذا الاجراء الباحث من تحليل الانحدار على عدد محدود من المتغيرات أو السماح للمتغيرات بالدخول في التحليل الاحصائي بشكل متتابع حسب قدرتها في تفسير التباين

-6تحليل التباين: ANOVA

تحليل التباين هو إجراء إحصائي لتقدير تأثير مجموعة من المتغيرات المستقلة (عوامل Factors) على متغير تابع تم قياسه بمستوى القياس القائم على الوحدات الفئوية Interval Scaleوتبعا لذلك يمكن تقسيم المتغيرات ووضعها في فئات Categories طبقا لقيم كل من المتغيرات المستقلة وطبقا لدرجة اختلاف المتوسط الحسابي لهذه الفئات عن المتغير التابع وبذلك يمكن حساب تأثير المتغيرات المستقلة على المتغيرات المستقلة على المتغيرات التابعة ودرجة التفاعل بين هاتين المجموعتين

ويوفر SPSS الإجراء الإحصائي ANOVA لتحليل التباين وهو شبيه إلى حد كبير بإجراء حساب تحليل الانحدار

-7التحليل العامليين: Factor Analysis

التحليل العاملي هو إجراء إحصائي عام يستخدم في تحديد الأبعاد الرئيسية المتمثلة في عدد كبير جدا من المتغيرات أو بمعنى آخر التحليل العاملي يكتشف الأبعاد الرئيسية التي تفسر مجموعة كبيرة من المتغيرات ويستخدم في ميادين العلوم الاجتماعية لتقليل المتغيرات الكثيرة إلى عدد صغير من العوامل التي يمكن تفسيرها بها.

خاتمة

خاتمة:

يهدف تحليل البينات إلى عدة أهداف أهمها الحصول على إجابات دقيقة واضحة لأسئلة محددة تمثل أسئلة الانطلاق التي أسست للبحث وقام من أجل الحصول عليها، أو توضيح طبيعة العلاقة بين الأثر والسبب لظاهرة ما وشدة هذه العلاقة واتجاهها. ومن أجل بلوغ هذه الغايات ينبغي أن يولي الباحث عناية كبيرة بالبيانات التي يبحث عنها ويستخدمها للوصول إلى نتائج بحثه، حيث تمر البيانات بثلاثة مراحل قبل أن تصبح جاهزة للعرض وهي :

1- إدخال البيانات: وتلي هذه المرحلة مرحلة جمع البيانات، حيث يقوم الباحث بإدخال البيانات في الحاسوب باستخدام برامج الإدخال .

2- تبويب البيانات .

3- معالجة وتحويل المعلومات إلى نتائج .

ويستخدم الباحث في تحليل البيانات إحدى الطرق التالية العديد من الطرق والأساليب أهمها التحليل الإحصائي الذي يتخذ طرقا وأشكال عديدة منها :

مقاييس التوسط : والتي تعد أكثر الطرق استخداما ، فهي تقيس النزعة المركزية بالنسبة لصفات أو خصائص معينة .

مقاييس التشتت : وتحدد هذه المقاييس درجة اختلاف البيانات عن بعضها البعض أو عن متوسطها .

الانحدار والارتباط : ويقوم هذا النوع من التحليل على دراسة العلاقة بين متغيرين أو أكثر بحيث يمكن التنبؤ بنتيجة أحدهما من خلال معرفة الآخر .

التحليل الوصفي العاملي : ومن خلاله يستطيع الباحث دراسة المتغيرات المتنوعة على ظاهرة معينة من خلال التحليل المنطقي والواقعي .

التحليل النوعي : ويستخدم في حساب المعايير الخاصة ، والتي تعتمد على وصف نوعية وطبيعة البيانات التي يتم جمعها .

قائمة المراجع

قائمة المراجع:

- مجموعة مؤلفين: الإحصاء باستخدام spss، شعاع للنشر والعلوم، حلب، 2007.
- محمد بلال الزعبي وعباس الطلاقحة، النظام الإحصائي spss فهم وتحليل البيانات الإحصائية، دار وائل للنشر، الأردن، ط3، 2006.
- جمال شعوان: **مدخل لدراسة وتحليل البيانات الإحصائية تطبيقات على برنامج Spss**، (pdf)، 2014.
- جوني دانييل: أساسيات اختيار العينة في البحوث العلمية، مركز البحوث الجامعية، الأردن، 2015.
- رعد جعفر حسين: **محاضرات في البرنامج الإحصائي Spss**، لطلبة الأقسام الزراعية السنة الدراسية الثالثة، الكلية التقنية، المسيب، (pdf)، 2012-2013
- غيث البحر ومعن التنجي: **التحليل الإحصائي للاستبيانات باستخدام برنامج IBM SPSS Statistics**، مركز سبر للدراسات الإحصائية والسياسات العامة، تركيا، 2014
- منى قمحية: **البرنامج الإحصائي Spss V10.0**، (pdf)، دت
- مهدي محمد القصاص: مبادئ الإحصاء والقياس الاجتماعي، (pdf)، 2007.
- هشام بركات بشر حسين: **تحليل البيانات باستخدام برنامج الحزمة الإحصائية للعلوم الاجتماعية، (pdf)**، دار النشر الإلكتروني، د ت
- وليد عبد الرحمان خالد الفرا: **تحليل بيانات الاستبيان باستخدام البرنامج الإحصائي Spss**، الندوة العالمية للشباب الإسلامي، 1430ه.
- بدون اسم: **مقدمة في برنامج Spss**، (pdf)، دت.

- غريب محمد سيد محمد وآخرون: الإحصاء والقياس في البحث الاجتماعي، دار المعرفة الجامعية، 1997.
- فتحي عبد العزيز أبو راضي: مبادئ الإحصاء الاجتماعي، دار المعرفة الجامعية،
- خليفة عبد السميع خليفة: الإحصاء التربوي، مكتبة الأنجلو المصري،
- حسن محمد حسن: مبادئ الإحصاء الاجتماعي، دار المعرفة الجامعية، 2000.
- عبد الله عبد الحليم وأخرون: الإحصاء المفاهيم الأساسية، 2003،
- فاروق عبد العظيم وآخرون: مبادئ الإحصاء، دار المعرفة الجامعية،
- أحمد حماد شعبان سعد : الوافي للرياضيات
- عماد توما بني كرش: علم الإحصاء
- أحمد عبدالسميع طبيّه : مبادىء الإحصاء
- ابراهيم عبد ربه : أساسيات الرياضية البحته
- بوسنة محمود، المرجع السابق، 2005، ص ص 164، 165.
- فاضلي إدريس : الوجيز في المنهجية و البحث العلمي، ديوان المطبوعات الجامعية 2008.
- صلاح أحمد مراد: الأساليب الإحصائية في العلوم النفسية والتربوية والاجتماعية، مكتبة الأنجلو المصرية، القاهرة، 2011.
- صلاح الدين شروخ : منهجية البحث العلمي، دار العلوم للنشر و التوزيع .

- فاضلي إدريس : الوجيز في المنهجية و البحث العلمي، ديوان المطبوعات الجامعية 2008.

- رجاء وحيد دويدري: البحث العلمي أساسياته النظرية و ممارساته العلمية . دار الفكر 2002.

- https://e3arabi.com/?p=1090376