معهد الدراسات والبحوث الإحصائية
قسم الإحصاء الحيوي والسكاني

جامعة القاهرة

الأسلوب الإحصائي: التحليل والتفسير

باستخدام الحاسب وبرنامج SPSS

إعداد دكتور / عبدالحميد محمد العباسي

أستاذ ورئيس قسم الإحصاء الحيوي والسكاني

2013

بسم الله الرحمن الرحيم

تقديم:

الحمد لله الذى أودع كتابه الكريم لطائف بره ورحمته ، وجلى بآياته السطور من بديع علمه وحكمته فسرت روحه فى الكائنات تحيى دارس العلم وأشرقت أنواره فى الوجود تبدد ظلمة العدم فأزالت عن القلوب رانها وعن النفوس أدرانها وأطلقت العقل من عقاله وأعطت الروح انطلاقها، وصلوات الله وسلامه على سيدنا محمد الذى كان خلقه القرآن صلى الله عليه وعلى آلة وصحبه والمرتشفين من علمه وحكمته الوارثين لكريم شمائله وأدبه و الناهلين من معين محبته ومتابعته .

يهدف هذا الكتاب الى مساعدة القارئ فى تعلم الفكرة الأساسية للأساليب الإحصائية وسوف نركز هنا على المفاهيم وشروط التطبيق وطريقة التفسير لأساليب التحليل الإحصائى البسيط والمتقدم، ومتى تستخدم وليس على المعادلات و الاشتقاقات الرياضية وطرق الحساب. فإعطاء معنى للبيانات هى المهمة الصعبة للباحث وليس تجميع البيانات، حيث يرى كثير من الباحثين أن جمع البيانات هو المهمة الصعبة، وما تبقى هو مهمة سهلة هدفها ربط البيانات التى تم جمعها بأهداف الدراسة ومما لاشك فيه أن الباحث الذى لديه هذا الاتجاه لا يقدر أهمية وصعوبة التحليل والتفسير.

ومما تجدر الإشارة إليه أن هذا الكتاب قد صيغت بشكل يجعلها مفيدة للباحثين فى فروع المعرفة المختلفة سواء كانوا متخصصين فى العلوم الاجتماعية أو الإنسانية أو الطبيعية. لهذا فهى غير موجهه الى فئة معينة من الأفراد بقدر ما هى موجهه الى جميع الباحثين المستخدمين للإحصاء لمساعدتهم فى تحليل بياناتهم والإجابة عن الأسئلة التى يسعون الى الإجابة عنها من خلاله. وليس هذا الكتاب كتابا إحصائيا. ولذلك كان على من يستخدمونها ممن ليس لديهم معرفة إحصائية كافية الرجوع الى بعض الكتب الإحصائية التى قد تكون مفيدة لهم فى الحصول على تفسير إحصائى للنتائج بشكل أعمق مما هو وارد هنا. و سنتناول بالتفصيل كل طرق العرض والتحليل الشائعة فى صورة مبسطة خاصة بعد التقدم الهائل فى الحاسبات والبرمجيات خاصة الإحصائية منها.

وبعد فقد حاولت قدر جهدى أن أخرج عملاً يرضى به الله عنى وأرضى به عن نفسى فإن كنت قد أصبت فهذا توفيق من الله وإن كنت قد أخطأت فهذا من عندى فاللهم إنى أسألك تكتب لهذا العمل ما حاولت أن أحققه فيه وأن تجعلنى من عبادك العلماء .

ربنا عليك توكلنا وإليك أنبنا وإليك المصير ...

أ.د/عبد الحميد العباسى

الفهـــــرس

الفصل الأول	المقدمة	20-1
	مقدمة وأهمية الإحصاء وأهدافه	2
	بعض المفاهيم الأساسية	6
	الظواهر والمتغيرات الاجتماعية وقياسها	8
	مستويات القياس	9
	خطوات البحث العلمى	10
	الطريقة العلميــة	10
	الطريقة الإحصائية	12
	خطوات الطريقة الإحصائية	12
	المناهج المستخدمة فى البحوث الاجتماعية	14
	مصادر و وسائل جمع البيانات	15
	تحليل البيانات	18
	الخطوات التى تمر بها عملية تحليل البيانات	19
	المراحل التى تمر بها عملية التحليل الإحصائى على الحاسب	20
الفصل الثانى	عرض وتبويب البيانات الإحصائية	31-21
	العرض الجدولى للبيانات	22
	العرض البياني والهندسي للبيانات	24
	التعاريف	31
الفصل الثالث	المقاييس الإحصائية الأساسية	39-32
	مقاييس للبيانات غير الرقمية (الوصفية)	33
	مقاييس للبيانات الرقمية (الكمية)	36
الفصل الرابع	اختبارات الفروض	45-40
	مقدمة	41
	اختبارات الفروض باستخدام الحاسب	43
الفصل الخامس	الارتباط والانحدار	53-46
	الارتباط	47
	الانحدار	51
الفصل السادس	اختيار الأسلوب الإحصائى متعدد المتغيرات	62-54
	الأساليب المختلفة للتحليل الإحصائى للمتغيرات المتعددة	56
	جدول الاستشارات الإحصائية	60
	المراجع	63

الفصل الأول
المقدمة

المقدمة

أولا مقدمة وأهمية الإحصاء وأهدافه:

تطورت الأساليب الكمية التى تستخدمها الإدارة فى سبيل إتخاذ القرارات المثلى اللازمة لتنفيذ عملية إنتاجية معينة، وقد زادت أهمية الأساليب الكمية فى جميع مجالات العلوم بتطور العلوم الرياضية وتطبيقاتها وكذلك بالاستخدام المتزايد للحاسبات الآلية عامة والشخصية منها خاصة مما سهل حل كثير من المشاكل المتعلقة بإتخاذ القرارات المختلفة.

وتستخدم الأساليب الإحصائية فى العديد من المؤسسات بما فى ذلك المؤسسات الأمنية والصناعية والإدارية ومؤسسات الإدارة العامة والتى يتطلب عملها إتخاذ القرارات بناء على البيانات المتاحة لديها، فالاستخدام السليم للإحصاء يمكن متخذ القرار من إستخلاص النتائج المفيدة من هذه البيانات.

وقد أتسع علم الإحصاء وتشعب وتعددت مجالاته وتطبيقاته كما أصبح هناك متخصصون فى الإحصاء داخل كل فرع من فروع العلم وفى كل مجالاته، ولما كان هدف العلم هو البحث عن الحقيقة فإن العلماء يحتاجون الى عدد من الطرق والوسائل التى تمكنهم من إشتقاق المعلومات والمعرفة وتحليل تلك المعلومات ومحاولة التنبؤ بالمستقبل وتحقيق وتعميم نتائجهم.

ولكل فرع من فروع العلم مجالاته ومشكلاته الخاصة به لذلك فالإحصائى يقدم الأدوات والوسائل بصورة عامة، وعلى المتخصصين فى فرع معين من فروع العلم إستخدام الصالح منها لأغراضهم والإلمام بها وبشرط تطبيق كل أداة او وسيلة. ولا يختص الإحصاء فقط بوسائل التحليل الكمى للظواهر والمتغيرات التى يمكن قياسها بل يمتد للظواهر النظرية غير المقيسة كمياً. وعلى ذلك فإن هذا النوع من التحليل يكون مكملا للتحليل الكمى وليس منافساً له .

والإحصاء لا يعوض سواء القياس، بل العكس فالإحصاء يهتم بدقة القياس، ومصير التحليل الإحصائى الفشل إذا كنا نشك فى دقة القياس عموماً. ولا تقف مهمة الإحصاء على التحليل فحسب حيث أن الإحصاء يدخل فى كل مراحل البحث العلمى. فلا يمكن للباحث أن يخطط وينفذ كل بحثه دون خبرة ومعلومات الإحصائى. فإذا لم يكن فى الإمكان ضبط المراحل الأولى من البحث وفق المواصفات الإحصائية السليمة فإن النتائج و التحليلات النهائية ستصبح غير ذات فائدة. فمهما كانت التحليلات دقيقة ومتعمقة فإن اعتمادها على تصميم قاصر أو قياس خاطئ للمتغيرات والظواهر لا يؤدى الى تعويض هذه العيوب أو العمل على الإقلال منها. هذا ويجب جمع المعلومات بطرق غير متحيزة حتى نصبح قادرين على وضع تفسيرات وتعميمات صادقة عن الظاهرة أو الظواهر محل البحث.

ماهية الإحصاء:

لقد كان علم الإحصاء في بدايته يهتم فقط بعملية العد والحصر للأشياء ، ومن هنا جاءت تسميته العربية " إحصاء" فهي مشتقة من كلمة أحصى ، وتعني استخدام الحصى أو الحجارة الصغيرة كوسيلة بدائية لعد الأشياء الكثيرة ، فقد كان الإنسان قديماً يستعين بالحصى في عملية العد. وكان علم الإحصاء مقصوراً على تعدادات السكان وثرواتهم ، وعدد المواليد والوفيات لمعرفة القوى البشرية المتوفرة في الدولة ، وذلك للأهتداء بها في تصريف أمور الدولة ورسم سياستها ، وحيث إن الإحصاء كان مقصوراً على الحقائق الخاصة بالدولة فمن هنا جاءت التسمية باللغة الأجنبية " Statistics" فهي مشتقة من كلمة " State " أي الدولة. وهناك تعريفات عدة للإحصاء تتراوح بين ما كان مألوفاً وشائعاً في الماضي إلى ما هو حديث وجامع. فقديماً عرف الإحصاء على أنه مجرد جمع المعلومات وترتيبها في جداول أو إبرازها في رسوم بيانية . وقد أخذ هذا المعنى للإحصاء يتلاشى من الأذهان فاسحاً المجال للمعنى الحديث الذي يعرف علم الإحصاء بأنه العلم الذي يبحث في جمع البيانات ، وتنظيمها ، وعرضها ، وتحليلها، واستقراء النتائج ، واتخاذ القرارات بناء عليها .

نبذة تاريخية عن تطور علم الإحصاء:

إن الإحصاء بمعنى الحصر والعد ، يرجع إلى قرون عديدة . فقد قام المصريون القدامى بعمل تعدادات للأيدي العاملة والثروات الموجودة ، واستفادوا من النتائج المستخلصة في بناء الأهرامات ، حتى يعينهم ذلك في التخطيط لعملية البناء وتنفيذها . وتبين البحوث التاريخية أن هناك دراسات مماثلة لما قام به الفراعنة أجريت من قبل الصينيين والرومان والإغريق للحصول على المعلومات التي تفيد في جمع الضرائب، والحروب وغيرها. كما تجدر الإشارة إلى قيام الدولة الإسلامية في عهد الخليفة العباسي المأمون قامت الدولة بعمل تعداد للسكان والثروات بهدف تحديد الإمكانيات المادية والفكرية . وتجدر الإشارة إلى ما قام به أفلاطون من معالجة لقضايا السكان في كتابه (الجمهورية) ، وكذلك أر سطو في كتابه (السياسة) ، وكما أن مقدمة ابن خلدون تشير إلى بعض الأساليب الإحصائية. غير أن هذه المحاولات لم تكن مبنية على أساس علمي. ويمكن اعتبار القرن السابع عشر بداية لظهور أنواع مختلفة من الأساليب الإحصائية . فقد جمع العالم كاسبر نيومان عام 1601م بيانات عن بعض الوفيات وأعمارهم ، وتوصل إلى استنباط دقيق لمتوسط العمر لبعض المجموعات ، أما العالم جون جرانت في الفترة (1624-1674) فقد ربط السكان بالإحصاء ، في حين أعد العالم ادموند هيلس أول جدول حياة. ولعل ما حدث خلال القرن الثامن عشر من تطورات في العلوم المختلفة، أدى إلى الاهتمام بالأساليب الإحصائية وتطورها . ومن العلماء الذين لعبوا دورا كبيراً في تطور علم الإحصاء الألماني فر يدريك جاوس (1777 - 1858) ، والفرنسي لابلاس (1749 – 1827) . والعالم الإنجليزي كارل بيرسون (1857 – 1936) إسهامات كثيرة في علم الإحصاء منها تعريف معامل الارتباط ومعامل الارتباط الجزئي وتقديره واستخدام اختبار مربع كاي لاختبارات جودة التوفيق .ويعتبر العالم الإحصائي رونالد فيشر (1890 – 1962) من الذين أضافوا الكثير لعلم الإحصاء ، وهو الذي وضع أساسيات علم تصميم التجارب وتحليل التباين وغيرها من الإسهامات في علم الإحصاء.

وصار علم الإحصاء يدرس في الكثير من الجامعات الأوروبية كأحد العلوم الرئيسية وخصوصا بعد منتصف القرن التاسع عشر، وبدأت الحكومات تهتم بإنشاء وتطوير الأجهزة الإحصائية. وفي بداية القرن العشرين تعقدت الأمور الإدارية لأجهزة الدولة المختلفة وذلك نتيجة حتمية للثورة الصناعية . وكان لابد من تطوير الدراسات والنظريات الإحصائية لمواجهة مشاكل تلك الإدارة . وبالفعل ففي العشرينيات من القرن الماضي تطور تطبيق النظريات الإحصائية ، ويرجع الفضل في هذا التطور إلى مجموعة من الأخصائيين أمثال شوهارت ، ورومج ، ودمنج وغيرهم.

أهمية علم الإحصاء:

توجد حالة من عدم التأكد في حياتنا اليومية وفي حياتنا العامة بشكل لم يسبق له مثيل، فنحن نناقش كل ما توارثناه عن آبائنا وأجدادنا من أفكار وتعليم ،ونطلب إجابة علمية على كل ما نوجهه من أسئلة مثل : هل توجد علاقة بين التدخين والسرطان ؟ وهل فعلا يتناقص الميل الحدي لاستهلاك الفرد كلما زاد دخله ؟ إننا نقرأ إجابات لمثل هذه الأسئلة ولكنها في الغالب تكون إجابات مماثلة لآراء أصحابها واعتقاداتهم الأمر الذي يفقدها الصفة العلمية . وحتى نستطيع الإجابة عن هذه الأسئلة وغيرها إجابة علمية صحيحة يجب أن نتبع المنهج الإحصائي .

ويعتبر الإحصاء أداة للتخطيط بالغة الأهمية في تخليص الاقتصاد من التطور العشوائي وعدم التوازن في نمو قطاعا ته المختلفة وفي تسريع عجلة التقدم . وقد تأثر نمو اقتصاديات البلدان المتقدمة والنامية بتطور جمع ودراسة البيانات الإحصائية وتحليل نتائجها ، وأصبحت الإحصاءات القاعدة المتينة التي تبني عليها سياسة الدول في كافة المجالات . ومن هنا فقد اهتمت الدول في عالمنا المعاصر بخلق كادر ذى خلفية علمية متينة متفهم للأساليب الإحصائية المختلفة في التحليل بشكل يؤهله للمساهمة الفاعلة في عملية تخطيط وتنفيذ البرامج الإنمائية . وتتجلى أهمية الإحصاء في التعرف على المجالات الحيوية التي تعتمد على الأساليب الإحصائية في البحث والتحليل . وكذلك تبرز أهمية علم الاحصاء في البحوث حيث إنه يساعد على تقديم أدق نوع ممكن من الوصف للمعطيات ، إذ الوصف والموضوعية من سمات العلم الحديث .

دور المنظمات الدولية في تطوير الإحصاء:

بدأ الاهتمام بالإحصاء من خلال المؤتمرات الدولية، وهي تنظيمات عالمية شبه رسمية انضم إليها خبراء الإحصاء وممثلو الأجهزة الإحصائية الحكومية، وكان الغرض من عقد تلك المؤتمرات توحيد الأساليب الإحصائية الحكومية في عمل التعدادات. ويعود الاهتمام بالإحصاء السكاني على النطاق الدولي إلى أواسط القرن التاسع عشر عندما عقد أول مؤتمر دولي للإحصاء عام 1853 م في بروكسل وصار يعقد كل سنتين أو ثلاث سنوات حتى عام 1878م . ولقد لعبت تلك المؤتمرات دورا كبيرا في تطوير الإحصاء من الناحيتين النظرية والتطبيقية . وساهمت منظمة الأمم المتحدة في هذا التطور ، واستمر الاهتمام الدولي بتأسيس المعهد الإحصائي الدولي ، وقد بلغ هذا الاهتمام ذروته في عام 1974م حيث اعتبرت المنظمة الدولية هذه السنة سنة دولية للسكان لجذب الاهتمام العالمي بالمشكلة السكانية .

ولعله من نافلة القول التأكيد على أن وجود جهاز إحصائي كفؤ في أية دولة يعتبر أحد مؤشرات تقدمها الاقتصادي والاجتماعي ، ويعبر عن مستوى إلمام ومعرفة تلك الدولة بالواقع الذي يعيشه مجتمعها في كافة النواحي الحياتية.

أهداف علم الإحصاء:

لعلم الإحصاء أهداف متعددة يمكن حصرها في الآتي:

1. تبسيط البيانات الإحصائية بعرضها في جداول أو رسومات بيانية ، وذلك لتسهيل فهمها وتحليلها.
2. التعبير عن الحقائق بصورة عددية واضحة ودقيقة بدلاً من عرضها ، والتعبير عنها بطريقة إنشائية.
3. مقارنة المجموعات المختلفة وإيجاد العلاقات القائمة بينها .
4. التنبؤ ببيانات مستقبلية مما يساعد عملية التخطيط .
5. استخلاص النتائج واتخاد القرارات المناسبة بقدر كبير من الصحة ، وذلك بعد قيام الباحث في أي فرع من فروع العلوم المختلفة بتحليل البيانات المتوفرة لديه.

علاقة علم الإحصاء بالعلوم الأخرى:

اكتسب علم الاحصاء أهميته من إمكانية تطبيق نظرياته ، ومبادئه وأساليبه في كل المجالات، يمكن التعبير عن ظواهره ببيانات يمكن تجميعها . فقد أصبح بالإمكان استخدام الأساليب الإحصائية وتطبيقها في مختلف العلوم. ونعرض فيما يلي علاقة الإحصاء ببعض العلوم:

- ففي علم الاقتصاد ، استخدم علم الإحصاء لتفسير الظواهر الاقتصادية المختلفة كنظريات الطلب والعرض ، والعلاقة بين مستويات الدخل والإنفاق الاستهلاكي، ونوع العلاقات الاقتصادية المختلفة وكيفية قياسها .. الخ.
- وفي مراقبة الإنتاج في الشركات الصناعية من حيث كمية ودرجة وجودة ومدى ملاءمة كل ذلك لاحتياجات السوق وأذواق المستهلكين . وغيرها من الدراسات الاقتصادية . ولايمكن للدراسات الاقتصادية أن تطور دون استخدام النظريات والطرق الإحصائية . فالإحصاء يعتبر أداة مهمة في تلك الدراسات . ويمكن القول بأن المجال الاقتصادي هو الإطار الذي نشأ وتطور في كنفه علم الإحصاء.
- وفي علم النفس ، استخدمت الطرق الإحصائية في قياس درجة ذكاء الأشخاص، وفي دراسة العلاقة بين ذكاء الأشخاص ومهارتهم .. الخ.
- وفي علم الفلك استخدمت الطرق الإحصائية في دراسات خاصة بتحديد مدارات الكواكب والنجوم وغيرها من الأبراج السماوية ..الخ.
- وفي علم الجغرافيا) بشقيها الطبيعي والبشري) استخدمت الأساليب الإحصائية في دراسة أشكال سطح الأرض ، والجغرافية المناخية وجغرافية البحار والمحيطات . فضلا عن تطبيق الطرق الإحصائية في جغرافية المدن وعلم الخرائط .. الخ.
- وفي العلوم الطبية ، يطبق علم الإحصاء في أغلبية الدراسات الطبية لمقارنة الأمراض المختلفة وسبل علاجها وتحديد العلاقة بين بعض الأمراض ومسبباتها، ولقياس كفاءة الأدوية .. الخ.

وفي الواقع لا يخلو علم من العلوم في عصرنا هذا من استخدام علم الإحصاء.

بعض المفاهيم الأساسية:

الهدف من أى بحث هو اختبار لافتراضات موضوعة فى إطار نظرى. ولنأخذ على سبيل المثال الافتراض النظرى للعلاقة القائلة بأنه كلما تزايدت درجة عدم الرضا كلما زاد التغيب من العمل. فالباحث الذى يرغب فى القيام بهذه الدراسة لا يستطيع القفز رأسا الى جمع بياناته لاختبار هذا الافتراض Proposition فلا شك أن عليه مهمة صعبة وهى تحديد المقصود بالظاهرتين محل الدارسة وهما " عدم الرضا " و " التغيب من العمل ". ولعل كل منهما كمفهوم نظرى قد يكون واضحاً. وتبدأ مشكلة الباحث من هذه المرحلة فعليه أن يقترح طريقة لقياس هاتين الظاهرتين.

و أهم الصعوبات التى تواجه البحث المنهجى أو العلمى وإحدى المعضلات التى تقف حائلا دون إيجاد إجابات مقنعة للمشكلات المختلفة، هى أن كثير من الظواهر من النوع النظرى ومن هنا فإن مهمة الباحث الأساسية فى هذه المرحلة هى تحديد الوسائل والعمليات اللازمة لعملية القياس ويمكن أن نسمى ذلك بمرحلة وضع التعاريف المنهجية Operational definitions .

والصعوبة الثانية ، هى أن معظم الظواهر الاجتماعية لا توجد فى فراغ أو تجريد وإنما توجد متشابكة متداخلة مع غيرها من الظواهر الأخرى، ومن ثم فهى اشد تعقيداً من الظاهرة فى العلوم الطبيعية فمن الاستحالة بمكان عزل الظاهرة الاجتماعية معملياً لكشف القوانين المفترضة عن طريق التجريب. لذلك نجد من الضرورى قبل الدخول فى التفاصيل أن نتعرض لبعض المصطلحات الإحصائية والتميز بين بعض المفاهيم .

نبدأ بتعريف **المجتمع الإحصائى Population** وهو ببساطة مجموعة من الأفراد أو الأشياء أو الأحداث (سنطلق عليها مفردات) محددة ومعرفة مسبقاً بحدود مكانية وزمانية. فعلى سبيل المثال ، إذا كان الهدف من الدراسة هو التعرف على أعمار مجموعة المتدربين فإن المجتمع الإحصائى هو جميع المتدربين المسجلون فى دورة معينة منعقدة فى فترة معينة.

وقد شاعت شهرة استخدام الإحصاء للعينات بل لقد عرفه البعض بأنه علم العينات ويمكن تعريف **العينة Sample** ببساطة أنها جزء من المجتمع فدراسة العينة هى دراسة لمجموعة من مفردات المجتمع بدلاً من إجراء حصر شامل للمجتمع ككل .

ويفترض فى العينة ان تكون ممثلة Representation للمجتمع، وعشوائية Random, وذات تكلفة معقولة. وتقسم العينات الى نوعين رئيسين: الأول **عينات احتمالية** ومنها العينة العشوائية البسيطة، والطبقية، والمنتظمة، ومتعددة المراحل، اما الثانى **عينات غير احتمالية** ومنها العينة القصدية، و العرضية.

- **الاعتبارات التى تدعو إلى سحب عينة:**

هناك عديد من الاعتبارات التى تدعو الى سحب عينة بدلا من دراسة جميع مفردات المجتمع الإحصائى كله. والواقع أن دراسة جزء من مفردات المجتمع مثلاً 5% أو 10% يوفر الكثير من الوقت والجهد والمال. فدراسة مائة مفردة بدلا من ألف لاشك يؤدى الى توفير المجهود المبذول ونفقات جامعى البيانات والطباعة والإعداد ... الخ. وبلا شك فإن الوقت الذى يتطلبه الأمر سوف يتضاءل كثيراً.

وبطبيعة الحال فإن البيانات يجب أن تجمع ويستفاد منها بسرعة وإلا أصبحت قديمة مما يقلل من قيمتها. كما يتطلب الأمر بسرعة الحصول على النتائج لأسباب اقتصادية أو سياسية أو غيرها. ولا يقف الأمر عند هذا الحد، فإن استخدام أسلوب العينة قد يكون حتمياً أو لا يمكن إجراء حصر شامل، وهنا فقد تتعدد الأسباب. فقد يكون المجتمع لا نهائى أى لا يمكن حصره أو عده مثل مجتمع النمل أو الأسماك. كما قد يتطلب الأمر إستخدام العينة إذا كانت دراسة المفردة يؤدى إلى فنائها مثل إنتاج بيض مزرعة أو تحليل دم مريض .

والاعتبارات السابقة ليست وحدها التى تبرر إستخدام العينات فالإحصائى يلجأ الى العينة حتى ولو كان الحصر الشامل ممكناً. فهو يثق بنتائج العينات ويرى أن العينة اقل تعرضا للخطأ عن الحصر الشامل. فدراسة جزء صغير من المجتمع كثيراً ما يؤدى الى الدقة فى القياس والتركيز فى جمع البيانات وتوحيد المفاهيم. كما تؤدى الى تقليل أخطاء الأعداد الكبيرة أثناء عمليات تفريغ وجدولة البيانات. والواقع أن هذا صحيح إلى حد كبير فإنه كلما صغر حجم المفردات تحت الدراسة كلما استطاع الإحصائى أحكام الإشراف والرقابة على بياناته فى جميع مراحل إعدادها.

- **الاعتبارات التى تدعو الى دراسة المجتمع كاملاً:**

وبالمثل فهناك اعتبارات تبرر إستخدام أسلوب الحصر الشامل دون أسلوب العينة مقدمة هذه الاعتبارات، طبيعة المجتمع المدروس، فإذا كانت المفردات معروفة بعدم التجانس أو ضآلة حجمها فلا مناص من حصر شامل وقد يلجأ الباحث للحصر الشامل لمجرد عدم توافر إمكانيات سحب العينة تمثل المجتمع محل الدراسة .

والاعتبار الأخير الذى قد يبرر الأخذ بأسلوب الحصر الشامل هو اعتبارات الدقة المتناهية فعلى سبيل المثال، فإن مصنعاً لإنتاج الساعات يحتم ضرورة التأكد من كل ساعة ينتجها ولا يلجأ الى أساليب الجودة التى تعتمد على أسلوب العينة. كذلك اختبار القدرات لجميع المتقدمين لشغل وظيفة معينة وليس لمجرد عينة منهم.

- **الوصف والاستنباط الإحصائى:**

بعد أن يحصل الباحث على بياناته وفقاً للأصول والقواعد المراعية فإنه عادة ما يتبع خطوتين أساسيتين، أولا يلجأ الى تلخيص مجموعة القرارات بطريقة مناسبة وهذا ما نطلق عليه عادة الوصف الإحصائى .

أما إذا بدأ الباحث استخراج تعميمات واستنتاجات تعتمد على أسس نظرية الاحتمالات مما يعطى فكرة عن مقدار الخطأ الذى يمكن أن يقع فيه الباحث بعملية التعميم على المجتمع ككل، فإن الباحث بذلك قد دخل فى الخطوة التالية وهو ما نطلق عليه الاستنباط (الاستنتاج) الإحصائى Statistical Inferences .

الظواهر والمتغيرات الاجتماعية وقياسها :

إن قياس السلوك الإنسانى وسيلة لفهم الإنسان كفرد وكعضو فى المجتمع فهو الوسيلة التى يمكن عن طريقها معرفة خصائص الأشخاص الذين نتعامل معهم من حيث قدراتهم وميولهم وإنجازاتهم وما يمكن لهم أن ينجزوه الى غير ذلك من الصفات. ويقصد بالقياس إعطاء قيمة رقمية لصفة من الصفات طبقاً لبعض القواعد أو الأسس. وهذه القيمة الرقمية هى التى تعطى القياس طبيعته الخاصة التى يختلف بها عن الوسائل الأخرى لوصف السلوك الإنسانى .

وتمكننا البيانات الرقمية التى نحصل عليها عن طريق القياس من :

1- تقدير السلوك .

2- الحكم على حالة الفرد أو الجماعة .

3- التنبؤ بسلوك الفرد أو الجماعة فى المستقبل .

ويلعب القياس كذلك دورا متزايد الأهمية فى البحوث الإحصائية خاصة وفى علم النفس والاجتماع والسكان والتربية عامة، إذا تمكننا المقاييس التى نستخدمها فى البحوث والدراسات من الحصول على البيانات بطريقة تجعلها قابلة للمعالجة والتفسير والتحليل الإحصائى، ومعنى هذا أن القياس هو الوسيلة التى يمكن عن طريقها تكميم الظواهر الاجتماعية، ويمكننا أن ندرك أهمية القياس إذا تذكرنا أنه بدون التكميم فإن العلم لن يكون له وجود حقيقي ولسنا نبالغ إذا قلنا أن العلم وتطبيقاته تصبح أمور مستحيلة بدون القياس، وغياب القياس عن العلوم الاجتماعية يعود بها إلى القرون الماضية عندما كانت خاضعة للتفسيرات الذاتية والفلسفية.

كما ذكرنا من قبل فالظواهر الاجتماعية تختلف كثيرا عن الظواهر الطبيعية خصوصاً فيما يتعلق بالقياس فمعظم الظواهر الاجتماعية توجد فى صيغة نظرية مما يتطلب أساليب دقيقة فى القياس تعتمد على نوع الظاهرة. فبعض الظواهر الاجتماعية توجد فى صورة يسهل قياسها رقميا أو كمياً والكثير من الظواهر يصعب بل يستحيل قياسها رقميا مثل الحالة الزواجية. فكيف يمكن قياس الحالة الزواجية رقمياً ؟ وهنا فالباحث يقوم بعملية تصنيفية بحتة، لا يستطيع الجزم بأى ترتيب كان لهذه الأوجه المعينة للظاهرة محل القياس. وأحياناً يستطيع الباحث وضع ترتيب معين لأوجه الظاهرة محل القياس، فعلى سبيل المثال إذا أردنا قياس مقدرة الطالب (المتدرب) على تحصيل مادة معينة فهنا نستطيع ترتيب الطلبة بحسب مقدراتهم : ممتاز – جيد جداً – جيد – مقبول – ضعيف، والواقع أن الظاهرة نفسها تحكم وتحدد نوع ومستوى القياس الذى يلائمها.

أن القياس فى أى مجال يتضمن ثلاث خطوات عامة :

1. تعريف أو تحديد الخاصية أو الصفة التى نريد قياسها.
2. تحديد مجموعة إجراءات يمكن عن طريقها إظهار هذه الصفة وعزلها وإدارتها وملاحظتها.
3. وضع مجموعة من الطرق أو الوسائل لترجمة هذه الملاحظات فى عبارات كمية سواء أكانت درجة أو كمية.

مستويات القياس

توجد عدة مستويات للقياس ويهمنا أن نفرق بينها، وعلى الباحث أن يكون ملما بها وبالوسائل والطرق الإحصائية المناسبة والمرتبطة بكل من هذه المستويات.

وتوجد أربع مستويات للقياس هى :

أ. الاسمية Nominal

هى اقل مستويات القياس بالنسبة لإجراء العمليات الحسابية (جمع وطرح وضرب وقسمة) ويسمى هذا المستوى أسميا لأن الأعداد التى تستخدم فيه تقوم مقام الأسماء او فئات ينتمى إليها الأشخاص أو الأشياء كأفراد. وذلك مثل أرقام التليفونات التى لا يتطلب استخدامها أى تحليل إحصائى، وكذلك أرقام لاعبي كرة القدم وغير كذلك...، فالأرقام فى مثل هذه الحالات لا تشير الى كميات من خصائص فعندما يستخدم الباحث رقم واحد ليشير للذكور ورقم أثنين ليشير للإناث أو يستخدم أرقاما معينة لتصنيف المهن أو المدارس فهو لا يهدف من استخدامها الى أى مضمون كمى. والعمليات الحسابية التى يمكن إجراؤها على مثل هذا المستوى من المقاييس هو العد أو التعداد للحالات الفردية، ولكن لا يمكن إستخدام أى من العمليات الحسابية الأربعة معها .

ب. الترتيبية Ordinal

هذا النوع من القياسات هو المستوى الثانى ويستخدم هذا المستوى عندما يمكن تنظيم البيانات فى سلسلة تمتد من الأدنى الى الأعلى (أو بالعكس) فى الخاصية التى نقيسها إلا أننا لا نستطيع أن نحدد الفرق بدقة بين أى رتبتين. وهذا النوع من المقاييس لا يتضمن أى معلومات قليلة عما نقيسه مثل ترتيب المتقدمين لعمل معين وفى هذا النوع من القياسات لا يمكن إستخدام العمليات الحسابية المعروفة ولا يمكن استخراج متوسط للرتب وإنما تصلح لهذه القياسات أساليب إحصائية أكثر ملائمة وهى الإحصاءات اللامعلمية مثل معامل ارتباط الرتب لسبيرمان أو كا[2].

ج. المسافة Interval

هذا النوع من القياسات هو المستوى الثالث للقياس ويميزه عن موازين الرتبة انه يسمح لنا بتحديد بعد شيئين أو شخصين عن بعضهما فى الخاصية موضوع القياس. إلا أن أهم نواحى قصور هذا النوع من القياسات أنها ليس لها صفر مطلق. فمثلاً قد يحصل طالب على صفر فى اختبار معين ولكن الصفر لا يمكن أن يعنى انعدام القدرة التحصيلية. ويمكن أن تستخدم عمليات الجمع والطرح مع هذا النوع إلا أن عملية القسمة بالذات لا يجوز استخدامها إطلاقا فلا نستطيع ان نقسم الدرجة التى حصل عليها الطالب (أ) فى اختبار ما على الدرجة التى حصل عليها الطالب (ب) ، وبالتالى لا يمكن الحصول على النسبة بين الدرجتين مثل الضعف أو النصف .

د. النسبة Ratio

هذا النوع من القياسات من اعلى مستويات القياس حيث يمكن إجراء جميع العمليات الحسابية فيه، وله صفر مطلق يفيد انعدام الصفة التى نقيسها وتتوافر فيه جميع خصائص المقاييس بالإضافة الى الصفر المطلق. وهذا النوع من القياسات هو المألوف اكثر من غيره. فأبعاد الجسم الطول والوزن والحجم يمكن قياسها بهذه الطريقة ولهذا نستطيع القول أن الشخص الذى يبلغ طوله 180 سم ضعف طول الشخص الذى طوله 90 سم. وتسمية هذا النوع باسم مقاييس النسبة جاءت من قابليته لاستخراج النسبة بين الأعداد والتعبير عن القياس فى صورة نسبة .

- **خطوات البحث العلمى:**

الهدف من البحث العلمى هو التوصل الى معرفة الحقائق بطريقة أكثر ما تكون صدقا ودقة فى ذلك بقصد التوصل الى إجابات لعديد من التساؤلات فيما يتعلق بأنماط وعناصر وتغيرات ظواهر الحياة المختلفة، وكذلك العناصر التى تؤدى الى هذه التغيرات وأيضا أشكال وأنواع العلاقات بين المتغيرات المختلفة بقصد معرفة علاقات الأسباب والآثار التى تزيد بين هذه المتغيرات المختلفة. وفيما يلى نبذة مختصرة عن أهم خطوات كل من :

أ - الطريقة العلميــة .

ب - الطريقة الإحصائية .

أ- <u>الطريقة العلمية</u>:

تتلخص الطريقة العلمية فى جمع حقائق عن جزئيات أو مفردات وربطها فى نظام منسق وإيجاد ما بينها من علاقات على صورة تعميمات أو قواعد عامة أو قوانين علمية. وتختلف الطرق العلمية لكل علم من العلوم تبعا لاختلاف طرق جمع البيانات واختلاف أدوات الملاحظة والقياس والتقدير الكمى، ووسائل التحليل والإثبات بالنسبة لكل علم من العلوم، ويهمنا هنا إستخدام الطريقة العلمية فى العلوم الإنسانية والاجتماعية وخاصة السكانية وهناك خطوات أساسية مشتركة للطريقة العلمية هى :

أولاً : <u>الشعور بوجود مشكلة</u> :

هو بداية عملية للتفكير وينشأ من ملاحظة أو ظهور بيانات تصادف الباحث ويرى فيها ما يخالف المألوف فينشط العقل مع الحواس لجمع بيانات أكثر تكون متعلقة بالمشكلة ولها صلة بها مع عزل واستبعاد البيانات التى لا صلة لها بالمشكلة وذلك للوصول الى الخطوة التالية.

ثانياً : <u>تحديد المشكلة</u> :

ويكــون ذلــك بتكــرار الملاحظــة لتحديــد المــشكلة وأبعــاد العناصــر التــى لا تتعلــق أو تتصل بها بالمشكلة.

ثالثاً : مرحلة فرض الفروض :

بعد تحديد المشكلة بدقة، تنشا الرغبة فى تفسير هذه المشكلة، ويلجأ الباحث الى فرض فروض علمية لتفسير المشكلة والفرض هو اقتراح مبدئى لتفسير المشكلة ويصل إليه المفكر عن طريقة معرفته السابقة بحالات مشابهة للظاهرة الحالية التى يريد تفسيرها.

رابعاً : جمع البيانات اللازمة لإثبات أو تغيير الفرض:

ويكون ذلك بإجراء التجارب الممكنة أو تكرار الملاحظة المنظمة وبعد جمع البيانات ترتب وتصنف وتعرض بطريقة مبسطة موجزة ومفصلة، لتسهل الإفادة منها.

خامساً : مرحلة التفسير:

بعد جمع البيانات بدقة وحياد وتحليلها رياضياً توضع معا فى نظام منسق مرتب يرى الباحث هل تؤيد الفرض العلمى أو تنفيه فإذا أيدت البيانات الفرض العلمى ليكون هذا الفرض تفسيرا للمشكلة أو حكماً عليها وإذا تكرار هذا التفسير مراراً وتكراراً فقد يسمى تعميماً .

أما إذا أوضحت البيانات التى جمعت خطأ الفرض فيجب تعديله أو البحث عن فرض جديد وجمع بيانات جديدة لتفنيد الفرض او تأكيد صحته.

سادساً : تحقيق النموذج:

بتحقيق النموذج الاحتمالى المقترح أو تفسيره أى اختبار صحته، ويكون ذلك بتطبيق هذا الحكم على الحالات والظواهر المماثلة الماضية والحاضرة والمتجددة، فإذا وجد من الظواهر ما يخالف هذا النموذج الاحتمالى المقترح بدأ فى الشك فى صحته.

ب - الطريقة الإحصائية:

هناك طرق متعددة للوصول الى الحقائق المختلفة المتعلقة بأى فرع من فروع العلوم. ولعل أهم وأدق طريقة هى الطريقة العلمية والواقع أن الطريقة الإحصائية تختلف عن غيرها من طرق الإثبات الأخرى فى أنها تعتمد اعتمادا كليا على التعبير الرقمى عن الظواهر.

لذلك فالظواهر غير القابلة للقياس العددى أو التى لا يمكن التعبير عنها بالأرقام لا تكون موضوعاً للدراسة الإحصائية. فالطريقة الإحصائية تستخدم الأرقام للوصول الى حقائق موضوعية دقيقة ومحددة وهى تعتبر من أقوى وسائل الإقناع والإثبات إذا أن منطق الأرقام موضوعى ومحايد بعيد عن الميل الشخصى للباحث.

ويلاحظ أنه لابد من الحرص عند إستخدام البيانات الإحصائية وذلك لاحتمال الوقوع فى الخطأ. وأهم الأخطاء التى تنجم عن عدم الحرص فى إستخدام البيانات الإحصائية.

1- خطأ التحيز :

كما لو توصل الباحث الى النتيجة أولاً بعد ذلك يحاول إيجاد الأرقام التى تثبت صحة هذه النتيجة .

2 - الأخطاء الناتجة من استعمال بيانات غير كافية :

مثل محاولة وصف الظاهرة عن طريق تسجيل عدد صغير من القيم .

3 - أخطاء ناتجة عن عمل مقارنة بين بيانات لا يصح المقارنة بينها :

مثل مقارنة الدخل النقدي فى مصر بالدخل النقدى فى الولايات المتحدة الأمريكية.

خطوات الطريقة الإحصائية:

1- تحديد المشكلة موضوع البحث :

أن الشعور بوجود مشكلة معينة أو الشعور بالحاجة الى أيجاد تعليل أو تفسير لحالة يراد علاجها. ولابد من تحديد موضوع البحث تحديدا واضحا دقيقا وكذلك تحديد النقاط التى تُكون الإطار العام للمشكلة.

2 - وضع الفروض :

أى فرض من الفروض العلمية التى يمكن أن تؤدى الى حل المشكلة يجب أن يصاغ صياغة سهلة واضحة وأن تكون فروضاً موضوعية بعيدة عن التحيز والميل الشخصى للباحث ويجب أن ترتب الفروض ترتيبا تفصيليا حسب أهميتها.

3 - جمع البيانات وتصنيفها وتبويبها :

الواقع أنه يلزم قبل البدء فى بحث المشكلة الإطلاع على جميع الأبحاث السابقة التى أجريت فى الموضوع وتقديرها من الناحية العلمية أى الإطلاع على المصادر التاريخية ومعرفة كل ما نشر عن المشكلة أو المشكلات المشابهة فربما نجد فيها الكفاية لتفسير المشكلة، وربما كان هناك بعض البيانات المتوفرة

والبعض الآخر غير متوفر فى المصادر التاريخية وبذلك يقتصر العمل على جمع البيانات الناقصة من الميدان سواء بالمشاهدة أو المقابلة أو البريد أو التليفون ... الخ. وبعد جمع البيانات لا بد من تبويبها وتصنيفها وتلخيصها تم عرضها بطريقة أو بأخرى باستخدام الجداول أو الرسوم البيانية.

4 - <u>تحليل البيانات</u> :

بعد تبويب البيانات وعرضها نجد أنه لا بد من معالجة هذه البيانات رياضيا لاستخلاص أرقاماً أكثر تلخيصا وأكثر إيضاحاً. فيمكن معرفة أين تتركز القيم والأهمية النسبية للأجزاء المكونة للمجموعة ومدى التقارب أو التباعد بين المفردات أى مدى تشتتها كما يمكن الوصول الى العلاقة التى قد تكون موجودة بين ظاهرتين أو أكثر.

5 - <u>التفسـير</u> :

هو اختبار صحة الفروض ويفضل أن يتخذ التفسير صورة معادلة رياضية. ولا يغيب عن بالنا أن القوانين العلمية التى نحصل عليها هى قوانين احتمالية أى ليست يقينية . كما يجب الإلمام بالقوانين العلمية الحديثة فى مجال البحث فتفسير المشكلات السكانية يستلزم أن يكون الباحث ملماً بالعلوم السكانية الحديثة. وكذلك تفسير، المشكلات التجارية يستلزم أن يكون الباحث ملماً بمواضيع الاقتصاد والتجارة والمستحدث فيها وللباحث الإحصائى عند قيامه بالبحث أن يستعين بالعالم الديموجرافى و الإجتماعى او النفسى او الطبى أو الاقتصادى حسب ميدان البحث وأن لا يستأثر وحدة بالتفسير إذا كان غير ملماً بمواضيع علم البحث.

6 - <u>تحقيق التفسير</u> :

يكون علاج المشكلة موضوع البحث حسب الفرض الصحيح الذى تم الوصول إليه فإذا انتهت المشكلة ثبت احتمالياً صحة الفرض أو ما يحل محله من حكم أو تعميم أو قانون أو نظرية. أما إذا لم تنته المشكلة وجب إعادة فرض فروض جديدة أو تكملة الدورة السابقة.

و الآن وبعد تعرفنا على الطريقة العلمية و الإحصائية ننتقل الى المناهج المستخدمة فى البحوث الاجتماعية ثم مصادر و وسائل جمع البيانات.

المناهج المستخدمة فى البحوث الاجتماعية:

يعد موضوع الدراسة وأهدافها احد الأسس التى يختار بناء عليها المنهج أو المناهج المستخدمة فى الدراسة ، فإذا كان موضوع الدراسة قد تحدد فى الآثار أو الأبعاد الاجتماعية - الاقتصادية المؤثرة على انتشار ظاهرة الإدمان فى المجتمع الكويتي ، أو فى دراسة عن دور المرأة فى تنمية المجتمع ، أو فى دراسة عن ظاهرة التضخم وآثاره الاجتماعيـة والاقتصادية . فهنا لابد ان يستخدم المنهج الوصفى Descriptive Methodلان استخدام هذا المنهج يساعد فى وصف وتحليل المتغيرات المرتبطة بالظاهرة المدروسة ، والأبعاد المؤثرة على تلك الظاهرة ، وكذلك الأبعاد الناجمة عنها.

وإذا كان موضوع الدراسة قد تحدد فى تأثير الخدمات التأهيلية على إعادة التكيف الاجتماعى للمصابين فى حرب تحرير الكويت 1991 ، أو فى تأثير الطلاق على التفكك الأسري فى المجتمع الكويتي . فان أى من هاتين الدراستين يستطيع أن يستخدم المنهج التجريبي Experimental Method ، ففي الدراسة الأولي الخاصة بتأثير الخدمات التأهيلية التى تقدم للمصابين فى حرب تحرير الكويت على إعادة التكيف الاجتماعى ، يمكن أن نكّون مجموعتين من المصابين فى هذه الحرب ، إحداهما تلقت هذه الخدمات التأهيلية (تسمى المجموعة التجريبية Experimental Group) والأخرى لم تتلق هذه الخدمات التأهيليـة (وتسمى المجموعة الضابطة Control Group). بشرط أن تتشابه المجموعتين كلما أمكن ذلك بين المتغيرات أو الخصائص مثل العدد ، والسن ، والحالة التعليمية ، والدخل الشهرى ، بكل من المجموعتين بينما الاختلاف الوحيد بين المجموعة التجريبية والمجموعة الضابطة ، أن المجموعة التجريبية قد تلقى أفرادها قدر منتظم ومستمر من الخدمات التأهيلية بينما لم يتوافر ذلك لأفراد المجموعة الضابطة ، ويهدف من خلال استخدام هذا المنهج الى التعرف على تأثير هذه الخدمات التأهيلية على التكيف الاجتماعي للمجموعه التجريبية ، وعدم حدوث هذا التكيف الاجتماعي لأفراد المجموعة الضابطة نتيجة لعدم تلقيها لهذه الخدمات التأهيلية .

وبالإضافـة الـى المنهجين السابقين يمكن استخـدام المنـهج الاستطلاعى الكشـفى (Explonatory) للبحث عن الجديد بالمجـالات التـى لـم يسبق إليهـا احد من قبل أو مستوى المعرفة بهـا قليل ، وتوجد ملحوظـة منهجية أساسية فيما يتعلق بالمناهج المستخدمة فى الدراسة وتتمثل فى أننا نستطيع بناءاً علـى موضـوع الدراسـة وأهدافها اختيار اكثر من منهج لتحقيق أهداف الدراسة .

مصادر و وسائل جمع البيانات:

تعتبر عملية جمع البيانات إحدى الخطوات الهامة فى البحث الإحصائى والواقع ان تحديد المشكلة ومتغيراتها تحديداً دقيقاً يؤدى إلى تحديد البيانات المطلوبة ونوعها، وبصفة عامة يمكن القول أنه يوجد ثلاث مصادر رئيسية للبيانات هى :-

أ. المصادر التاريخية .

ب. المصادر الداخلية .

ج. المصادر الميدانية .

وعموماً على الباحث أن يجمع كل ما يستطيع من بيانات يمكن أن تلقى ضوءاً على المشكلة محل البحث من المصدرين الأولين ويلجأ بعد ذلك الى استكمال بياناته من المصدر الثالث أى الدراسة الميدانية.

أ- المصادر التاريخية :

هى عبارة عن بيانات تم جمعها مسبقا بواسطة جهات أخرى أو لبحوث أخرى وأهم الأمثلة على ذلك بيانات تعداد السكان والإحصاءات المختلفة التى تقوم بها الدولة سواء جمعها أو نشرها كذلك أبحاث العينة والبحوث الشاملة التى قامت بها بعض جهات وجمعت على أساسها بيانات يمكن أن تفيد الباحث فى دراسته.

ب - المصادر الداخلية :

يقصد بها مجموعة البيانات التى سبق لنفس الباحث أو الجهة القائمة بالبحث تجميعها. فعلى سبيل المثال قد يلجأ الباحث الى البيانات الموجودة بالمؤسسة التى يعمل بها من خلال الإطلاع على الدفاتر والسجلات والتقارير .

ج - المصدر الميدانى :

هنا يقوم الباحث بنفسه بجمع البيانات من مفردات العينة التى قام بتصميمها وسحبها من المجتمع.

- **جمع البيانات الميدانية:**

ليست عملية تصميم وتجميع البيانات الميدانية بعملية سهلة أو هينة، و إنما تخضع للشروط وللقيود التى يضعها الباحث و إلا حصل على بيانات مضللة غير ممثلة للمجتمع المستهدف كذلك فهى عملية طويلة تتضمن الخطوات الآتية :-

1 - تحديد نوع المجتمع المستهدف :

هل هو مجتمع بشرى أم غير بشرى هل هو مجتمع ريفى أم حضرى ؟ ما هى درجة التعليم، متوسط السن مدى صعوبة المواصلات ... وغير ذلك من الاعتبارات التى يجب ان تؤخذ فى الاعتبار.

2 - تحديد وحدة المعاينة :

هذه الخطوة وأن بدت سهلة إلا أنها تحتاج الى تحديد دقيق، فمثلاً الباحث الذى يريد جمع بيانات عن السكان مثلا نجد أن وحدة العد هى الفرد، ولو كنا بصدد دراسة استهلاك السلع والخدمات فوحدة العد المناسبة

هى الأسرة، وهذا التحديد بلا شك يؤدى الى الوضوح وإن كان يثير مشكلة أخرى هى الإطار Frame فإيجاد الإطار السليم يكون عادة إحدى المشكلات التى تجابه الباحثين فى ميدان العلوم الاجتماعية.

3 - <u>تحديد ما إذا كان سيتم جمع البيانات من المجتمع ككل أو من عينة ممثلة</u> :

عادة ما يلجأ الباحث الى أسلوب العينة . ولكن هناك عديد من الاعتبارات التى تدفع الباحث الى إجراء حصر شامل. إذا ما قرر الباحث جمع عينة فعليه أن يحدد بقدر الإمكان حجم العينة التى يراد سحبها وأن تكون ممثلة للمجتمع الذى يرغب فى دراسته.

4 - <u>تحديد وسيلة جمع البيانات وتصميم صحيفة البحث</u> :

بعد أن قام الباحث بتحديد النقاط التى سيشملها بحثه وقام بدراسة مصادر البيانات التاريخية خصوصاً دراسة البيئة التى سيجرى فيها الباحث ثم قام بتصميم العينة الممثلة فأنه عليه أن يحدد أولا الوسيلة التى سوف يتبعها فى جمع بياناته هل هى عن طريق المقابلة الشخصية مع المبحوثين أم عن طريق اتصال التليفونى او البريد ، او سيتم ذلك عن طريق الملاحظة... أو غيرها من وسائل جمع البيانات.

ثانيا تصميم ما نسميه بصحيفة البحث او الاستبيان Questionnaire وصحيفة البحث هى الاستمارة التى تحتوى على مجموعة الأسئلة التى يجب أن نعرف إجابتها الدقيقة من المبحوثين .

5 - <u>اختيار وتدريب جامعى البيانات</u> :

على الباحث أن يبدأ فى عملية اختيار جامعى البيانات وتحديد أهم صفاتهم وهذا يعتمد الى حد كبير على طبيعة البحث ذاته خصوصا إذا كان حجم العينة كبير نسبيا أو ذا أتساع جغرافي كبير. فالباحث فى مجال الدراسات الطبية (الصحية) يحتاج الى جامعى بيانات ذو خلفية طبية .

6 - <u>تجربة الاستمارة</u> :

من أهم الخطوات الواجب إتباعها هى اختبار صحيفة البحث للتأكد من سلامتها. وتسمى هذه الدراسة الاستكشافية Pilot Study. وتجرى على عينة مماثلة للعينة الأصلية والهدف من هذه الدراسة المسبقة :-

أ. التعرف على الصعوبات التى قد تواجه البحث الأصلى فى الميدان .

ب. المساعدة على التعرف على مدى فهم جامعى البيانات للاصطلاحات .

ج. تعرف مفردات العينة للأسئلة والألفاظ المستخدمة و وضوحها .

فالدراسة الاستكشافية ضرورية للاطمئنان على صحيفة الاستبيان من ناحية الشكل والصياغة والمضمون .

وسائل جمع البيانات:

تتعدد وسائل جمع البيانات الميدانية ، و اختيار الوسيلة يعتمد الى حد كبير على طبيعة البحث ذاته ودرجة الدقة المطلوبة ، وبطبيعة الحال على ميزانية البحث المالية والزمنية كذلك فإن طبيعة المجتمع المستهدف قد تحتم أحيانا اختيار وسيلة معينة دون أخرى.

أهم وسائل جمع البيانات المباشرة هى:

- المشاهدة أو الملاحظة Observation .
- البريـــــد Mail.
- الاتصال التليفونى Telephone .
- المقابلة الشخصية Interview .
- الصحف والمجلات والإذاعة .

- **تصميم استمارة البحث :**

الهدف من استمارة البحث هو ترجمة الأهداف العامة للبحث الى لغة أو صياغة يستطيع المبحوث فهمها لجمع الحقائق التى تلقى الضوء على مشكلة البحث. ولذلك فتصميم استمارة البحث جزء أساسى وهام قد يساعد على نجاح البحث أو يدفع الى فشله ويجب أن تتوافر فى الاستمارة الجيدة الشروط الآتية :-

1. يجب تحديد التعاريف التى تستخدم فى الاستمارة تحديداً واضحاً.
2. يجب استخدام الألفاظ البسيطة.
3. فى الأسئلة التى تتطلب إجابتها وحدات رقمية يجب تحديدها تحديداً دقيقاً.
4. يتحسن أن تكون إجابة السؤال بكلمة " نعم" أو " لا " قدر الإمكان.
5. إذا كان السؤال يحتمل إجابات متعددة يفضل كتابة كل الحالات الممكنة.
6. إذا كان السؤال كبيراً يستحسن تقسيمه الى عدد من الأسئلة البسيطة الواضحة.
7. يجب ترك مسافات كافية عند الطبع للإجابة على كل سؤال.
8. يجب ألا تتطلب الأسئلة عمليات حسابية معقدة.
9. يجب البعد عن الأسئلة التى يختلف فيها الحكم الشخصى.
10. يجب ألا توحى الأسئلة بإجابات معينة للمجيب.
11. يجب أن يكون عدد الأسئلة قليلاً بقدر الإمكان.
12. يجب أن تغطى الأسئلة جميع نواحى البحث وعناصره الأساسية الهامة.
13. يجب ألا تحتوى على أسئلة محرجة أو منافية للعرف.
14. يجب تكرار بعض الأسئلة بصيغ مختلفة لإمكان مراجعة الأسئلة.
15. يجب أن يكون هناك تسلسل منطقى بين الأسئلة.

- **شكل الاستمارة:**

يجب أن تكون الاستمارة ذات حجم معقول ومطبوعة جيداً و من السهل قراءتها دون مجهود وتكون على شكل كراسة صغيرة وليست أوراق متفرقة. وعلى الصفحة الأولى يجب أن نكتب عنوان البحث بوضوح والجهة القائمة به وعادة ما يكتب فى مكان واضح بصدر الاستمارة ما يفيد أن البيانات سرية ولن تستخدم لغير الأغراض الإحصائية .

- **تنسيق الاستمارة:**

بطبيعة الحال لا يجب أن تضع الأسئلة فى الاستمارة دون خطة معينة وترتيب منطقى فأولاً يجب أن تقسم الاستمارة الى موضوعات أساسية، يحتوى كل منهما على عدد من الأسئلة يغطى الموضوع. وهذه الأسئلة يجب أن ترتب بدورها بطريقة منطقية فيجب أن تعطى الأسئلة أرقاماً مسلسلة حتى لا يحدث أى لبس وحتى يمكن إحالة البيانات أحياناً الى أسئلة سابقة أو أن نطلب منه القفز مباشرة الى سؤال معين قادم فى حالة عدم انطباق الأسئلة الأخرى مثلاً.

ثانيا تحليل البيانات:

البيانات هى قيم المتغيرات التى يتم جمعها من المفردات سواء فى العينة او المجتمع محل الدراسة، وبصفة عامة هى مجموعة من المواد التى تستخدم كقاعدة للتوصل الى قرارات. والتوصل الى قرارات أمر مهم يمارسه جميع الأفراد على اختلاف تخصصاتهم. وتوجد عدة مصادر يمكن أن نحصل من خلالها على هذه البيانات فمنها ما نحصل عليه من المسوح Surveys او التجارب او مصادر أخرى كالمصادر التاريخية و السجلات.

والحصول على البيانات لا يعد كافيا بحد ذاته بل يجب أن نتوصل الى الاستنتاج المبنى عليها وهذا الأمر يعرف بتحليل البيانات لذلك فمن المهم جدا أن يعرف المرء البيانات التى يحتاجها وكيفية تحليلها. ويمكن تحليل البيانات بطرق مختلفة ومتعددة ففى بعض الأحيان لا يلزمنا سوى عملية وصف البيانات، وتختلف المعالجة الإحصائية او تحليل البيانات تبعا للهدف الذى نسعى إليه، وشكل البيانات التى نتعامل معها. ويساعدنا الحاسب فى إجراء التحليلات التى نريدها لأنه يقلل من احتمالية الخطأ والوقت المطلوب للقيام بعملية التحليل وسهولة ويسر الأداء، و الحاسب لا يختار الأسلوب الملائم لعملية التحليل ولا يقوم بتفسير النتائج ولكن الباحث هو الذى يحدد ذلك طبقا لهدفه وطبيعة البيانات المتاحة.

والخطوات التى تمر بها عملية تحليل البيانات هى:

1- تصميم الدراسة:

تعد الخطوة الأولى فى إجراء أى بحث هو تقرير طبيعة البيانات التى نحتاجها و من أين نجمعها وكيف تجمع.

2- استخدام الحاسب:

يتضمن إدخال البيانات فى الحاسب بحيث تصبح جاهزة للاستخدام. ويجب أن نتأكد من دقة الإدخال وسلامة الترميز حتى نضمن نتائج صحيحة ودقيقة.

3- توصيف البيانات:

بعد إدخال البيانات فى ملف للحاسب تبدأ عملية التحليل والخطوة الأولى فى عملية التحليل هى توصيف البيانات التى تم إدخالها.

اختيار الأسلوب الإحصائى المناسب:

أن اختيار أسلوب مناسب لتحليل البيانات تحكمه أمور عديدة منها:

أ. هدف الدراسة.

ب. نوع الدراسة (وصفية ـ تحليلية ـ ...الخ) .

ج. تحقق الشروط الخاصة بالأسلوب الذى يستخدم (حجم العينة ـ ...الخ).

وبصفة عامة يمكن أن **تصنف الأساليب (المعالجات) الإحصائية** فى ثلاث مجموعات:

1- وصف البيانات:

تستخدم لوصف الظاهرة (الظواهر) التى ندرسها، ومن هذه الأساليب الرسوم و الجداول التكرارية والمتوسطات ومقاييس التشتت والالتواء و التفرطح.

2- اختبارات الفروض:

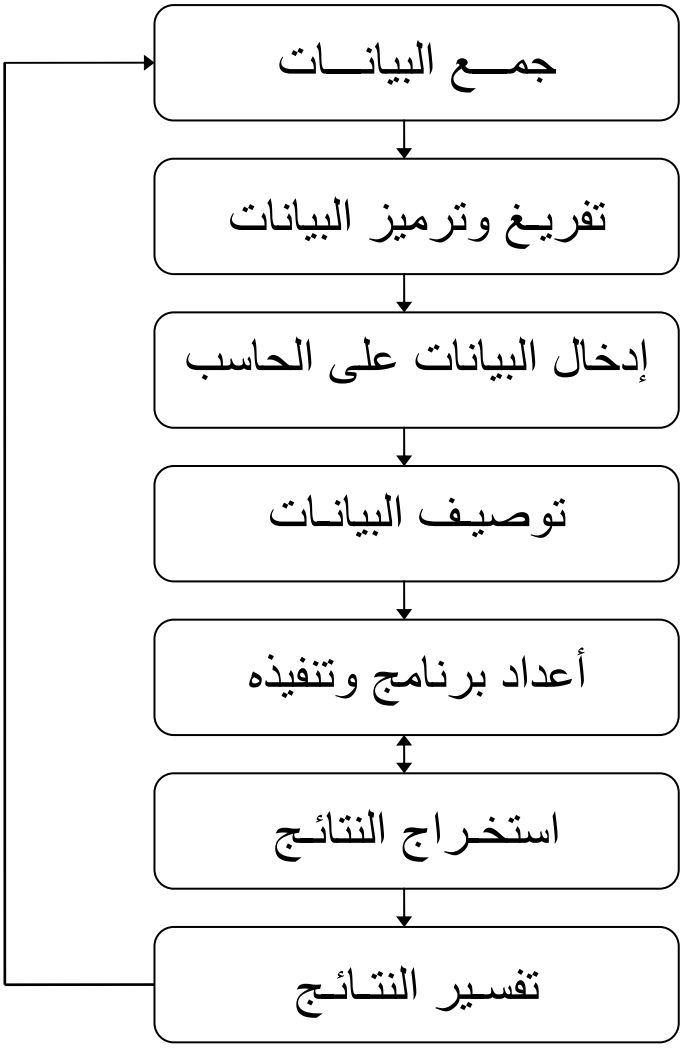
تستخدم للتأكد من فرضية معينة او عدم تحقق معقوليتها او لاختبار أن الفرق ذات دلالة إحصائية أم لا ومن أمثلتها اختبار Z, T, χ^2.

3- وصف العلاقات:

تستخدم لتقرير وجود علاقة بين متغيرين او أكثر ومنها تحليل الانحدار والتباين والتحليل العاملى و تحليل التمايز و السلاسل الزمنية ...الخ.

ويوضح الشكل التالى المراحل التى تمر بها عملية التحليل الإحصائى على الحاسب.

المراحل التى تمر بها عملية التحليل الإحصائى على الحاسب



الفصل الثانى

عرض وتبويب البيانات الإحصائية

عرض وتبويب البيانات الإحصائية

يعتبر عرض وتبويب البيانات الإحصائية الخطوة الثانية (بعد تجميع هذه البيانات) في مفهوم العملية الإحصائية . ففي مجال الدراسات والأبحاث الاقتصادية والاجتماعية والإدارية والأمنية وغيرها ، يحصل الباحث على معلومات ضخمة يصعب عليه وعلى القاريء وصفها أو تفسيرها بسهولة لاستخلاص بعض الحقائق أو المدلولات الإحصائية منها. لذلك يلجأ الباحث الى ترتيب البيانات وعرضها بطريقة مختصرة تساعد على قراءتها ومعالجتها إحصائيا للتعرف عليها ووصفها ومقارنتها بغيرها من الظواهر، والخروج ببعض الإستنتاجات الإحصائية عن مجتمع الدراسة .

وقد ساعد ظهور وتطور الحاسب الآلي بكل أشكاله على تسهيل عملية تبويب وعرض البيانات التي يمكن الحصول عليها سواء من المصادر الأصلية (أي من خلال الإستبيان أو قاعدة البيانات) أو من المصادر المطبوعة أو المنشورة، حيث اصبح هناك برامج متخصصة فى العرض البيانى مثل Harvard Graphics, K-Plot ، كما ادرجت داخل العديد من البرامج امكانية العرض البيانى مثل برامج مجموعة Offeice. ولكن لا بد من استعراض الأساس الذي تسند إليه عملية عرض البيانات وتبويبها. وتتوقف طريقة عرض البيانات على نوع هذه البيانات وعلى الحقائق المطلوب إبرازها. وهناك طريقتان أساسيتان لعرض وتبويب البيانات الإحصائية وهما:

أولاً : العرض الجدولي للبيانات (Statistical Tables)

بعد عملية تبويب وتعيين الصفات التي تميز المشاهدات (المفردات)، ترصد النتائج في جداول مناسبة توضح الشكل النهائي للمجموعات المميزة. وتسمى هذه العملية التي يتم بها تجميع البيانات في مجموعات مميزة ومتجانسة بعملية التصنيف. وتصنف البيانات الإحصائية بوجه عام وفقًا لإحدى القواعد التالية :

1 – **تصنيف جغرافي** : حيث تجمع المشاهدات (المفردات) التي تشترك في صفة مكانية واحدة في مجموعة مستقلة، مثل توزيع الجرائم حسب المحافظات ، ومن الأمثلة الأخرى للتصنيف الجغرافي توزيع سكان الدولة حسب المحافظات .

2 - **تصنيف تاريخي أو زمني** : حيث تجمع المشاهدات (المفردات) المرتبطة بزمن معين في مجموعة واحدة. أي أن توزيع البيانات حسب زمن حدوثها وتسمى بيانات هذا النوع من التصنيف بالسلاسل الزمنية .

3 – **تصنيف نوعي أو وصفي** : حيث تجمع المشاهدات (المفردات) التي تشترك في صفة واحدة بمجموعة واحدة ، مثل تصنيف المتهمين حسب الجنس أو الجنسية أو كلاهما ...الخ .

4 - **تصنيف كمي** : حيث يتم تجميع المشاهدات (المفردات) المشتركة في صفة معينه بأخذ شكلا رقميًا في مجموعة واحدة ، مثال ذلك توزيع المتهمين حسب فئات العمر ، أو توزيع السكان حسب فئات الإنفاق الشهري . ويسمى الجدول المصمم على أساس التصنيف الكمي بالجدول التكراري بفئات .

وقد يكون التصنيف خليطًا من الأنواع الآنفة الذكر بحيث تظهر البيانات تبعًا لنوعين أو أكثر من هذه الفئات ، ويترتب على ذلك ضياع المعالم الشخصية للمعلومات بحيث تصبح مجرد أرقام معينة في تصنيفات معينة يراها الباحث ، ومن أجل أن يكون الجدول الإحصائي واضحًا يضع الإحصائيون قواعد يجب مراعاتها عند بناء وتصميم الجداول الإحصائية. ومجمل هذه القواعد أن يكون الجدول واضحًا ومفسرًا لنفسه بنفسه دون الحاجة الى الرجوع الى مصممه أو واضعه.

وعليه فإن إعداد الجدول يتطلب مراعاة ما يلي :

1 – وجود عنوان واضح له .

2 – ذكر المصدر الذي أخذت منه البيانات .

3 – تسجيل الملاحظات الخاصة في أسفل الجدول وتحديدها بعلامات خاصة .

4 – توضيح عناوين الأعمدة والصفوف .

5 – ضرورة ذكر وحدات القياس الكمية أما في عناوين الأعمدة والصفوف أو عنوان الجدول .

6 – يفضل فصل الأعمدة بخطوط رأسية عندما تكون متعددة .

ويمكن التمييز بين شكلين رئيسيين من الجداول الإحصائية :

أ – **الجدول البسيط** : يشمل بياناً واحدا فقط ، حيث يحتوي على عمود واحد للبيان موضع الدراسة يقابله عمود آخر بعدد التكرارات مثل الجدول الذي يبين توزيع الجرائم حسب نوعها او حسب المحافظات.

توزيع الجرائم خلال النصف الأول عام 2001 بالكويت

المحافظة	عدد الجرائم
العاصمة	2026
حولي	1367
الأحمدي	526
الفروانية	1149
الجهراء	614
مبارك الكبير	330
المجموع	6012

ب – الجدول المتعدد : حيث يصنف الجدول حسب ظاهرتين او اكثر، مثل الجدول الذي يبين توزيع الجرائم حسب المحافظات ونوع الجريمة معا أو المحافظات ونوع الجريمة والمعلومية.

توزيع الجرائم حسب نوع القضية خلال النصف الأول عام 2001 بالكويت

نوع القضية	جنح	جنايات	اجمالى
مصلحة عامة	190	48	238
نفـــس	1397	732	2129
عرض وسمعة	601	195	796
مـــال	2349	4105	6454
المجموع	4537	5080	9617

توزيع اجمالى الجرائم حسب نوع القضية خلال النصف الأول لعامى 2000 و2001 بالكويت

نوع القضية	النصف الأول 2000		النصف الأول 2001	
	جنح	جنايات	جنح	جنايات
مصلحة عامة	342	64	190	48
نفـــس	1911	794	1397	732
عرض وسمعة	814	252	601	195
مـــال	2922	4690	2349	4105
المجموع	5989	5800	4537	5080

ثانيًا : العرض البياني والهندسي للبيانات (Graphs)

حيث تعرض البيانات وتمثل في رسومات وأشكال بيانية خاصة. ويكون الشكل البياني تعبيرًا تصويريًا للعلاقة بين المتغيرات التي يحاول الباحث دراستها. وتستخدم في الإحصاء أنواع مختلفة من الأشكال البيانية والتصويرية التي تعتمد على طبيعة البيانات الإحصائية والهدف المرجو ابرازه من الشكل البياني .

وسنقوم باستعراض أهم الأشكال وأكثرها استعمالا وشيوعًا في الحياة العملية.

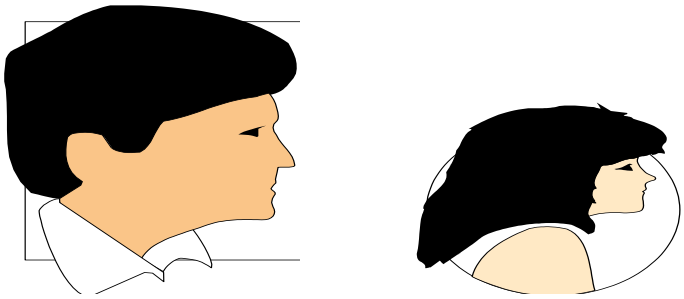
1 – الأشكال أو الرسومات المصورة (Pictograms) :

حيث تستخدم الصور والأشكال للتعبير عن القيم المختلفة للظواهر المراد عرضها ويستخدم ذلك لعرض القيم الإجمالية للظواهر موضع الدراسة لغرض المقارنة . وقد يتبع في ذلك تكرار الصور للدلالة على حجم الظاهرة في فترات زمنية مختلفة . كما قد يتبع اسلوب آخر باستعمال صور صغيرة يتم تكبيرها للدلالة على زيادة حجم هذه الظاهرة .

وعلى سبيل المثال سنأخذ البيانات التالية عن مرتكبي الجرائم بالكويت عام 2000 .

ذكور	24774
إناث	3686

ويمكن تمثيل هذه المجموعة من البيانات برسم مصور كالرسم الموضح بالشكل أدناه .



وهذه الطريقة شائعة الإستخدام ، وتعطي انطباعًا بصريًا جيدا عن مجموعة البيانات المبحوثة ، ولكنها ليست دقيقة إذا انجزت يدويًا ، لذلك يفضل عدم استخدامها الا إذا استخرجت باستخدام الحاسب الآلي .

2 – الدوائر المجزأة (Pie Charts)

حيث تمثل الدائرة مجموع القيم الكلية للظاهرة ، ويتم تقسيمها الى قطاعات جزئية تتناسب مع قيم المجموعات الجزئية التي تتكون منها الظاهرة ، وتميز تلك القطاعات عن بعضها أما بألوان مختلفة أو بظلال مختلفة من أجل ضمان الإيضاح . ويستخدم هذا النوع من الرسوم البيانية في الحالات التالية :

أ – عندما يكون الهدف منها مقارنة الأجزاء المختلفة بالنسبة للمجموع الكلي .

ب – عندما بكون الأجزاء المقارنة قليلة العدد نسبيًا .

ج – كما يمكن استخدامها أيضًا لتوضيح التطور النسبي لإجزاء الظاهرة في فترات زمنية مختلفة .

ولتوضيح ذلك ، يمكن استخدام الدائرة لبيان توزيع جرائم الجنايات بالكويت خلال النصف الأول عام 2001 حسب نوع الجريمة كما يلي :

نوع القضية	جنح	جنايات	اجمالى
مصلحة عامة	190	48	238
نفـــس	1397	732	2129
عرض وسمعة	601	195	796
مـــال	2349	4105	6454
المجموع	4537	5080	9617

بالطريقة التقليدية نتبع الخطوات التالية :

1 – نرسم دائرة بمقياس رسم مناسب .

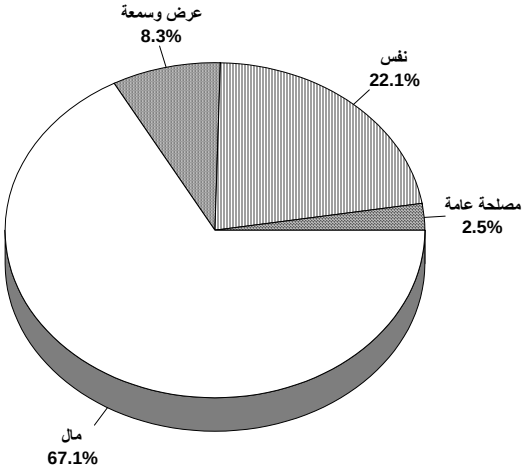
2 – نحسب نسبة كل مجموعة الى المجموع الكلي .

3 – تقسيم 360 درجة على المجموعات حسب نسبة كل مجموعة .

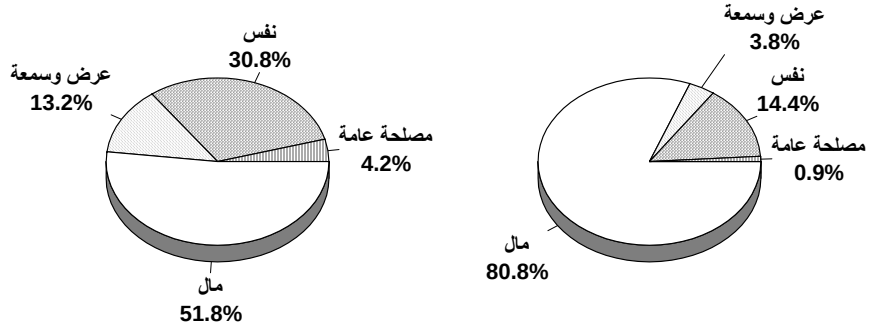
وبالتالي فإننا نحصل على تقسيم للدائرة الى قطاعات تتناسب كل منها مع حجمها من المجموع الكلي .

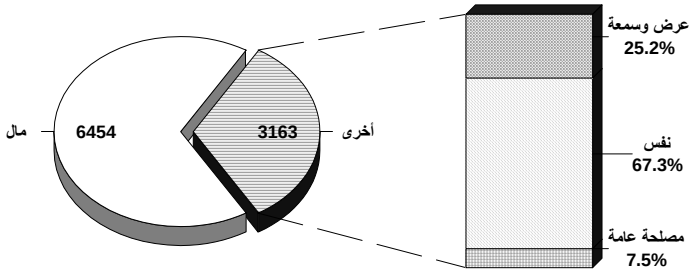
	المصلحة العامة	النفس	العرض والسمعة	المال
النسبة المئوية %	2.5	22.1	8.3	67.1
الزاوية بالدرجات	9	80	30	241

ونقوم بتقسيم مساحة الدائرة الى القطاعات المذكورة حسب قيمة زوايا القطاع كما هو مبين في الشكل أدناه . ويتم تمييز كل قطاع من هذه القطاعات بلون معين أو بتظليل معين ، مع ملاحظة كتابة اسم الجزء الذي يخص كل قطاع من هذه القطاعات الى جانبه (أو داخله) .



جنايات جنح





3 – الأعمدة والمستطيلات (Bar Charts)

وتبني على أساس أعمدة ومستطيلات متساوية القاعدة ، يتناسب ارتفاعها مع قيم الظاهرة المدروسة مع استخدام مقياس رسم مناسب يمكن من عرض أكبر القيم للظاهرة . واذا ما وجدت أحد القيم الشاذة في الكبر فعندها يمكن استخدام طريق العمود المقطوع . وتستخدم الأعمدة البيانية في الحالات التالية :

1 – عندما تكون أجزاء الظاهرة المقارنة كثيرة العدد نسبيًا , حيث يصعب تمثيلها بالدائرة المجزأة لأنها تبدو مكتظة لدرجة يصعب معها مقارنة التوزيع النسبي للظاهرة المدروسة .

ب – عندما تكون الأجزاء المقارنة في فترات زمنية مختلفة .

ج – عندما نرغب في توضيح قيم الأجزاء المقارنة المختلفة للظاهرة موضع الدراسة ، وذلك من أجل ابراز المقارنة بين هذه الأجزاء وتوضيح التغير أو التطور عبر الزمن سواء لظاهرة واحدة أو لعدة ظواهر بين فترات زمنية مختلفة .

ويمكن أن يكون التمثيل بالأعمدة والمستطيلات بأشكال مختلفة :

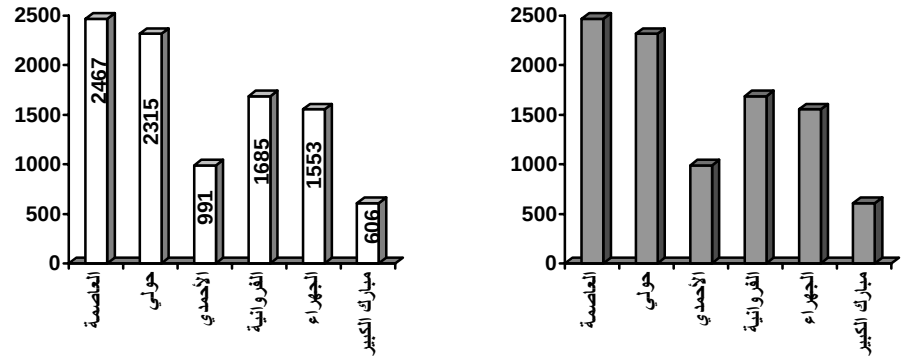
أولا : الأعمدة المبسطة (Simple Bar Charts)

وتستخدم هذه الطريقة لعرض تقسيمات ظاهرة واحدة فى نفس الفترة الزمنية، وكذلك لعرض الظاهرة في عدة فترات زمنية حيث يخصص لكل تقسيم او فترة زمنية عمود مستقل مع ترك مسافات متساوية بين كل عمود وآخر وبالتالي نحصل على شكل عبارة عن مجموعة أعمدة متجاورة .

ولتوضيح فكرة الأعمدة المبسطة ، سندرس البيانات التالية بخصوص توزيع جرائم الجنح حسب المحافظات بالكويت خلال النصف الأول عام 2001 .

المحافظة	عدد الجرائم
العاصمة	2467
حولي	2315
الأحمدي	991
الفروانية	1685
الجهراء	1553
مبارك الكبير	606
المجموع	9717

ويوضح الشكل التالى تمثيلا بيانيا مبسطًا بالأعمدة بحيث يمثل هذه البيانات .



ويعطي هذا الشكل انطباعًا بصريًا جيدًا عن الأهمية النسبية في كل محافظة . ولكن عندما يكون لدينا فترات زمنية مختلفة لتوزيع الجرائم فإننا سنلجأ الى طريقة أخرى للتمثيل بالأعمدة البيانية .

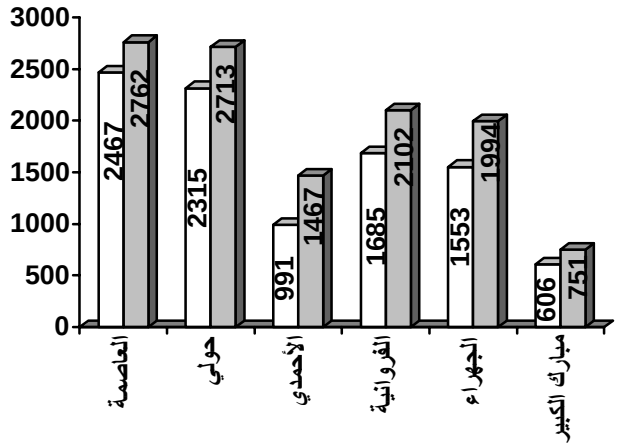
ثانيًا : الأعمدة المتلاصقة (Multiple Bar Charts)

وتستخدم عندما يكون المطلوب عرض ظاهرة معينة في عدة فترات زمنية أو عدة أماكن جغرافية أو التقسيم خلال فترة زمنية واحدة . ففي هذه الحالة تمثل كل فترة بعدد من الأعمدة المتلاصقة مساوية لعدد أوجه الظاهرة ، مع تمييز كل عمود بلون أو تظليل معين .

ولتوضيح فكرة الأعمدة المتلاصقة ، لو كان لدينا البيانات التالية للجرائم بالكويت فى النصف الأول لعامى 2000 و 2001.

المحافظة	جرائم 2000	جرائم 2001
العاصمة	2762	2467
حولي	2713	2315
الأحمدي	1467	991
الفروانية	2102	1685
الجهراء	1994	1553
مبارك الكبير	751	606
المجموع	11789	9617

فبدلا من رسم شكلين بيانيين منفصلين لا تسهل مقارنتها ، فإنه يفضل عمل رسم بياني موحد تكون فيه قيمة الأعمدة أزواجا متجاورة يمثل كل زوج نفس الخاصية في مجموعتي البيانات. ويبين الشكل التالى أدناه هذا النوع من الرسم البياني بالأعمدة المتلاصقة أو المتجاورة والذي استخدمت فيه المواد الخام الواردة في المثال .

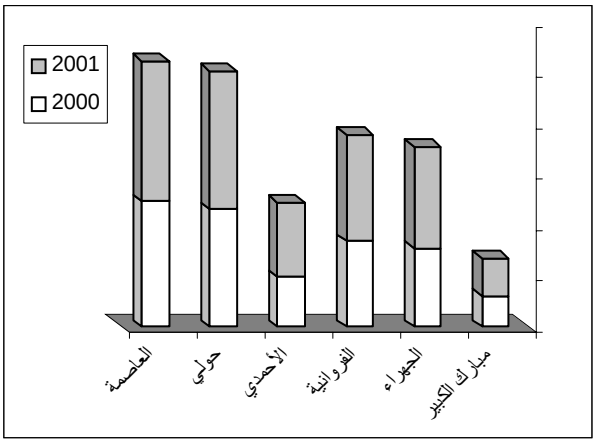


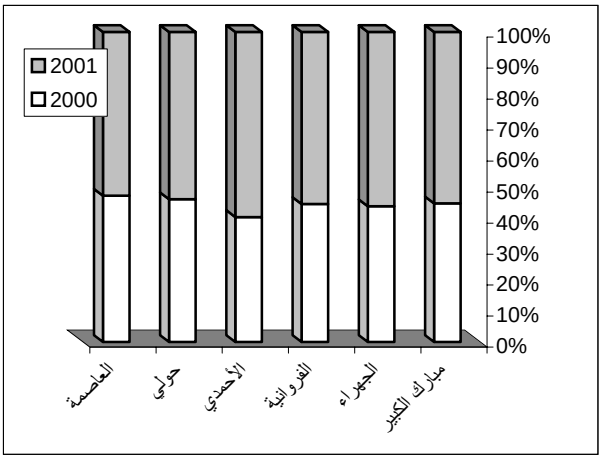
ثالثًا : الأعمدة المجزأة (Component Bar Charts)

ويتم اللجوء لهذا النوع من الرسوم عندما تكون القيمة الإجمالية للظاهرة موزعة على مجموعات فرعية مميزة ، ويشمل البيان عدة فترات زمنية أو تقسيمات ومن أجل تبديد بعض الغموض التي يكتنف تفسير أو مقارنة البيانات نتيجة اختلاف الأطوال الكلية للأعمدة. لذلك يتم اللجوء الى رسم عمود يمثل قيمة الظاهرة الكلية لكل فترة زمنية ، ويقطع هذا العمود الى أجزاء تتناسب مع قيم المجموعات الجزئية ويميز كل جزء من الأعمدة بألوان معينة أو بأشكال وتظليلات مختلفة لكل مجموعة فرعية على حدة .

إلا أن الإختلاف في اجمالي القيم يعوق المقارنة الى حد ما وخصوصًا عندما يكون المقصود هو المقارنة النسبية . ومن أجل تسهيل هذه المقارنة يمكن اللجوء الى استخدام النسب المئوية . ويبين الشكلين التاليين أعمدة بيانية ذات مركبات تمثل القيم والنسب المئوية لتوزيع الجرائم حسب نوعها في المحافظات

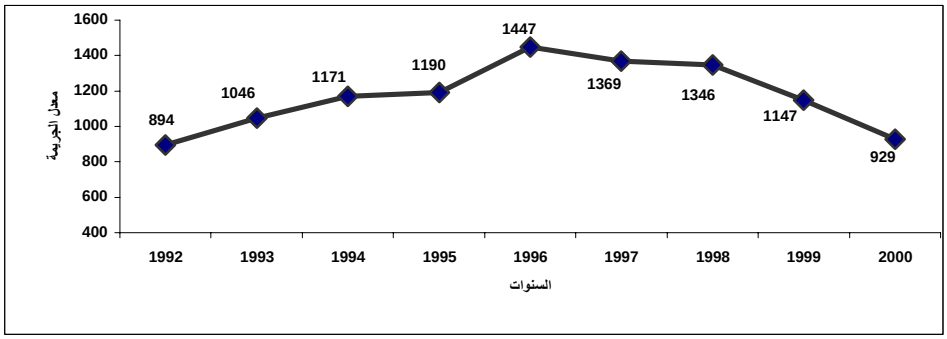
بالكويت خلال النصف الأول عامى 2000 و 2001 وهذا النوع من الأعمدة والمستطيلات البيانية هو أفضلها . فكون المستطيلات أو الأعمدة لها نفس الطول الكلي يجعل مقارنة البيانات الفرعية المناظرة أسهل .



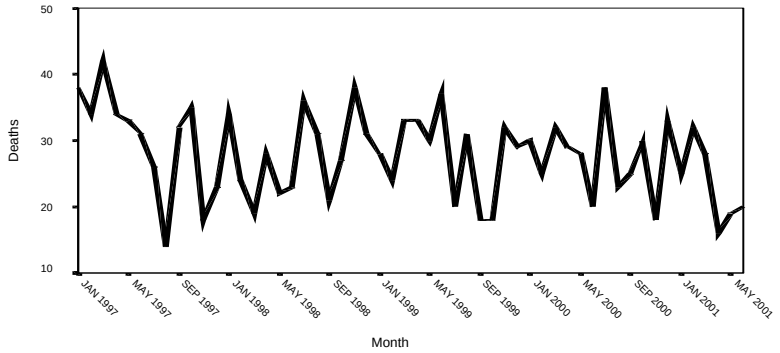


رابعا: الخطوط البيانية Scatter

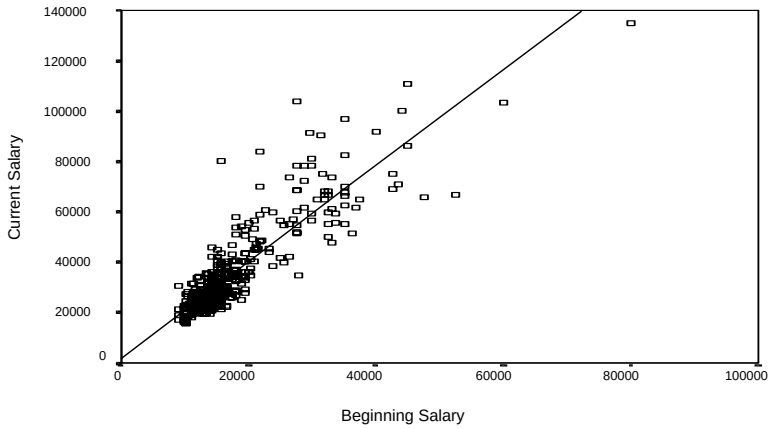
للتعبير عن العلاقة بين ظاهرتين بخط بياني ، يتم رسم محورين متقاطعين في نقطة الأصل، يمثل المحور الأفقي أحد الظاهرتين والمحور الرأسي الظاهرة الأخرى. وقد يكون الموضوع متعلقًا بتمثيل قيم ظاهرة ما خلال فترة زمنية معينة، وفي هذه الحالة يكون المحور الأفقي ممثلا للفترات الزمنية المتعاقبة، ويمثل المحور الرأسي قيم الظاهرة في تلك الفترة. وتمثل أزواج القيم لكلا الظاهرتين على الرسم ويوصل بين تلك النقاط على الخط البياني . ويبين الشكل التالى معدل الجرائم بالكويت خلال الفترة 1996 – 2000 .



ويبين الشكل التالى أعداد الوفيات الشهرية الناتجة من حوادث المرور خلال الفترة من يناير 1997 حتى يونيو 2001 .



ويبين الشكل التالى اخرمرتب (معاش) ومرتب اول تعيين لعينة من المتهمين فى قضايا الرشوة.



Definitions

Indicator: Unlimited number of a specific event in specific area in particular time.

Proportion: It is the relation between a category and the total, which means dividing one category by another.

Ratio: Arithmetic relation between two categories belong ton the same group. Which means dividing one category on another.

Rates: Consecution of events among population in a particular time, Rates clarify degree of repetition of events and spreading degree. Row rates are accounted according to the total population where as private rates are accounted according to limited branch group.

Constant: It is, arbitrary number for example (100, 100.000) we can multiply it with rates, proportion quantities to express this measurement in order to simplify it to the readers.

Period Measures: It is the statistics, which measures the events that happen for whole population on part of them during a certain period of time.

Change Ratio: change amount (the difference) in comparing between two years and the original year.

Mean: A measure of central tendency. The arithmetic average; the sum divided by the number of cases.

Standard Deviation: A measure of dispersion around the mean.

Coefficient of Variation: Data is considered ill conditioned if has a small coefficient of variation (the standard deviation divided by the mean), to compare the measurement not the same .

التعاريف

المؤشر: الرقم المطلق لحدث معين في منطقة محددة وفي وقت محدد.

التناسب: هو العلاقة بين مجموعة فرعية وبين الإجمالي, وهذا يعني قسمة مجموعة فرعية على أجمالي المجموعات.

النسبة: هي العلاقة الحسابية بين مجموعة فرعية ومجموعة فرعية أخرى من نفس المجموعات ويعني هذا قسمة مجموعة فرعية على أخرى.

المعدلات: هي تواتر الأحداث بين السكان في فترة زمنية محددة, وتبين المعدلات مدى تكرار الحدث ومدى شيوعه , والمعدلات الخام هي التي تحسب بالنسبة لمجموع السكان . أما المعدلات الخاصة فهي تلك التي تحسب النسبة لمجموعة فرعية محدودة.

الرقم الثابت: هو رقم اعتباطي لا يتغير (مثلا 100 , 100.000) يمكن أن نضرب فيه المعدلات أو النسب أو المقادير للتعبير عن هذه المقاييس بطريقة تجعلها أكثر فهما للقارئ .

مقاييس الفترة الزمنية هى الإحصاءات التي تقيس أحداث تقع لكل السكان أو لجزء منهم خلال فترة من الزمن.

نسبة التغير: مقدار التغير(الفرق) بين سنتي المقارنة والأساس منسوبا لسنة الأساس.

المتوسط (الوسط): مقياس للنزعة المركزية ويحسب بقسمة مجموع القيم على عددها.

الانحراف المعيارى: مقياس للتشتت النسبى للقيم حول متوسطها ويستخدم للمقارنة بين بيانات لها نفس القياس.

معامل الاختلاف : مقياس للتشتت المطلق ويفضل عندما تكون قيمتـه صغيـرة، ويحسـب بقسمـة الانحراف المعيارى على المتوسط، ويستخدم للمقارنة بين بيانات ليس لها نفس القياس.

الفصل الثالث
المقاييس الإحصائية الأساسية

المقاييس الإحصائية الأساسية

يوجد العديد من المقاييس الأساسية التى يمكن استخدامها لتلخيص وعرض البيانات الإحصائية وأهم هذه المقاييس يمكن تقسيمها حسب طبيعة قياس البيانات الى:

أ - مقاييس للبيانات غير الرقمية (الوصفية) Categorical .

ب- مقاييس للبيانات الرقمية (الكمية) Numerical ، وتنقسم الى:

1. مقاييس النزعة المركزية.
2. مقاييس التشتت.
3. مقاييس الالتواء والتفرطح.

أولا: مقاييس للبيانات غير الرقمية (الوصفية) Categorical

يوجد العديد من المقاييس الأساسية التى يمكن استخدامها لتلخيص وعرض البيانات الإحصائية غير الرقمية وأهم هذه المقاييس:

التناسب (الحصة) والنسبة المئوية ، والنسبة ، والمعدل

Proportion , Percentages , Ratio & Rates

يمكن تلخيص البيانات المقاسة على المستوى الاسمي مثلما هي الحالة في باقي أنواع المقاييس . وتتركز العمليات الحسابية في عملية العد (Counting) للحالات الواقعة في كل فئة (Category) وبيان حجم كل فئة مقارنة بالفئات الاخرى . فلنفرض ان لدينا توزيع مجموعة من السجناء وفق متغير النوع (متغير اسمي) فيمكننا ان نعد عدد الذكور ولنفرض ان عددهم (20) والاناث وعددهن (40) فرد .

التناسب (الحصة) Proportions

مقارنة الجزء بالكل . وترمز (P) الى النسبة ، إما (F) الى عدد التكرارات اما (N) فترمز الى اجمالي عدد الحالات . ويلاحظ ان قيمة النسبة لا تتعدى الواحد الصحيح (1) . وهذا يعني ان :

التناسب (الحصة) = التكرار مقسوم على العدد الكلي للحالات $P = \frac{F}{N}$

وهذه الخاصية يمكن تعميمها الى أي نوع من الفئات . وهنا نتمكن من المقارنة بين الفئات المختلفة (مثلا المحافظات في نسب الجرائم المرتكبة في كل محافظة) . ويمكن القول أن الهدف من استخدام الحصة هو توحيد مقياس المقارنة للبيانات الخام .

لنفرض أن لدينا 20 مذنب منهم 12 انثى ، فإن حصة الإناث في العينة $= 12 / 20 = 0.6$

النسبة المئوية : Percentage %

النسبة المئوية = مجموع التكرارات x 100 مقسومًا على العدد الكلي

في المثال السابق : عدد الإناث المذنبات = 12

عدد المذنبين الكلي = 20

النسبة المئوية للإناث في السجن = 12 x 100 مقسومًا على 20 = 60 %

$$12 \times 100 / 20 = 60\%$$

المعدل : Ratio

المعدل = قسمة مجموعة على مجموعة أخرى مثل معدل الإناث الى الذكور $12/8 = 1.5$

وهذا يعني ان حصة الإناث هي 1.5 لكل ذكر وتكتب هكذا : 1.5 : 1

والجدول التالي يبين صعوبة مقارنة الأرقام الخام فقط :

جدول رقم (1) عدد جرائم التعدي على الإنسان والتعدي على الممتلكات في ثلاثة محافظات كويتية*

المحافظة	العاصمة	الجهراء	الفروانية
التعدي على الممتلكات	147	182	191
التعدي على الإنسان	59	75	108
المجموع	206	257	299

* البيانات افتراضية.

ويمكن ان تكون المقارنة ذات معنى أكثر عند مقارنة هذه المحافظات.

جدول رقم (2) عدد و نسبة وتناسب (الحصة) لجرائم التعدي على الممتلكات والإنسان في المحافظات.

المحافظة	العاصمة			الجهراء			الفروانية		
التعدى	N	%	P	N	%	P	N	%	P
التعدي على الممتلكات	147	71.4	0.714	182	70.8	0.708	191	63.9	0.639
التعدي على الإنسان	59	28.6	0.286	75	29.2	0.292	108	33.1	0.331
المجموع	206	100	1.000	257	100	1.000	299	100	1.000
نسبة التعدي على الممتلكات الى التعدي على الإنسان	1 : 2.49			1 : 2.43			1 : 1.77		

ويمكن ملاحظة أن النسبة المئوية والحصة لجرائم التعدي على الإنسان في محافظتى العاصمة والجهراء متقاربة جدا و متشابهة .

المعدلات Rates

هل تبدو محافظة الفروانية أكثر المحافظات في التعديات ضد الإنسان لا ، انها تشير فقط الى أن عدد كبير من جرائم التعدي على الممتلكات مقارنة بجرائم التعدي على الإنسان . ولكي نحدد أي من هذه المحافظات أكثر في جرائم التعدي ضد الممتلكات أو الأفراد فإننا بحاجة الى تحديد معلومة أخرى وهي عدد سكان كل محافظة ، وذلك لتحديد التعديات مقارنة بالسكان .

والمعدلات نوع خاص من النسب التي تقارن عدد من التكرارات لحادثة معينة (مثل الجريمة) لإجمالي حدوثها في فترة زمنية محددة منسوبا لعدد السكان ، فمثلا جميع السكان في أي محافظة يحتمل ان يرتكبوا جرائم ، ولو فرضنا في المثال السابق أن :

سكان محافظة العاصمة = 115,000 الف نسمة

سكان المحافظة الجهراء = 175,000 الف نسمة

سكان المحافظة الفروانية = 540,000 الف نسمة

أما معادلة احتساب المعدلات فهي : عدد الجرائم / عدد السكان x 100,000

معدل التعدي على الإنسان لكل 100,000 من السكان في كل من :

محافظة العاصمة = 59\115,000x100,000= 51,3

محافظة الجهراء = 75\175,000x100,000=42,9

محافظة الفروانية = 108\540,000x100,000=20

ولاحظ أن معدلات التعدي على الإنسان في محافظة العاصمة هي الأعلى .

ثانيا: مقاييس للبيانات الرقمية (الكمية) Numerical

يوجد العديد من المقاييس الأساسية التى يمكن استخدامها لتلخيص وعرض البيانات الإحصائية الرقمية وأهم هذه المقاييس:

أ. مقاييس النزعة المركزية Central Tendency

تهدف هذه المجموعة من المقاييس الى بيان القيمة التي تنزع مجموعة من البيانات أن تتمركز حولها. وتمثل القيمة التى تتمركز حولها مجموعة من البيانات وصفا لطبيعة تلك البيانات .

الوسط Mean

يمثل الوسط (المتوسط) Average مجموع القيم مقسوم على عددها .

مثال : البيانات التالية تمثل عدد تكرار الحدث للانحراف : 1, 3 , 5 , 6 , 8 , 9

لحساب الوسط $= \sum x / n$ =1+3+5+6+8+9 \ 6 = 36\6 = 6

الميزات: 1 – يدخل في احتساب الوسط جميع القيم

2 – المجموع الجبري لانحرافات القيم عن وسطها يساوي صفر .

3 – مجموع مربع الانحرافات عن الوسط يقل عن مجموع مربعات انحرافات القيم عن أي قيمة أخرى .

4 – يتأثر الوسط بالقيم الشاذة الكبيرة والصغيرة .

الوسيط Median

وهو القيمة التي تقسم مجموعة من البيانات الى 50% أعلى منها و 50% أدنى منها .

مثال البيانات التالية تمثل اعمار متهمين بقضية ما 18, 17 , 16 , 15 , 14 , 13 , 12 أحسب وسيط العمر ومتوسطه.

1 . ترتب البيانات تنازليا أو تصاعديا 12، 13، 14 ،15، 16، 17، 18

ترتيب قيمة الوسيط = (عدد القيم + 1) مقسوم على 2

= (7 + 1) \ 2 = 4

ترتيب قيمة الوسيط = القيمة التي ترتيبها 4 وهي 15 حيث يقع دونها 50% مـن البيانات (3 قيم) وأعلى منها 50 % من البيانات (3) قيم .

المتوسط = $\sum x / n = 105 / 7 = 15$

مثال آخر لحساب الوسيط في حالة العدد الزوجي للبيانات .

البيانات التالية تمثل اعمار متهمين بقضية ما 17 , 16 , 15 , 14 , 13 , 12 أحسب وسيط العمر ومتوسطه.

1 . ترتب البينات تنازليا أو تصاعديا 17، 16، 15، 14، 13، 12

ترتيب قيمة الوسيط = (عدد القيم + 1) مقسوم على 2

=(6+1)/2=3.5

ترتيب قيمة الوسيط = القيمة التي ترتيبها 3.5 وهي القيمة الواقعة بين الترتيب 4 و 3 وهما القيمتان 15 ، 14 وبحساب متوسطهما = (14+15)\2=14.5 .

المتوسط = $\sum x / n = 88/6 = 14.7$

أما في حالة وجود قيم متشابهة فيكون الوسيط أيا منها .

مثال 17 , 16 , 15 , 14 , 14 , 14 , 14 , 13 , 12 فيكون الوسيط 14

المميزات:

1 – أكثر مقاييس النزعة الركزية ثباتًا

2 – لا يتأثر بالقيم المتطرفة والشاذة

3 – يعتمد عليه بدلا من الوسط في حالة وجود قيم شاذة

يوجد مقاييس أخرى للنزعة المركزية منها على سبيل المثال:

1. المنوال Mode
2. الوسط الهندسى Geometric
3. الوسط التوافقى Harmonic

ب. مقاييس التشتت Dsipartion

تهدف هذه المجموعة من المقاييس الى بيان مقدار اختلاف مجموعة من البيانات عن وسطها (قربها وبعدها عن المتوسط). وتمثل مقاييس التشتت القيمة التى تبين مقدار الاختلاف لمجموعة البيانات .

المدى Range

يمثل المدى ابسط أنواع مقاييس التشتت ، ولا يتطلب عمليات رياضية في احتسابة ويساوي الفرق بين أعلى قيمة وأدنى قيمة. $Range = Max - Min$

فقد يتساوى المتوسطان في مجموعة بيانات ولا يتساوى توزيعهم حول كل متوسط. مثال :

أ. 120، 119 ، 118 ، 117 ، 116 ، 114 ، 113 ، 112 ، 111 ، 110

ب. 190 ، 175 ، 160 ، 145 ، 130 ، 115 ، 100 ، 85 ، 70 ، 55 ، 40

يلاحظ من هذا المثال أن الوسط = 115 ولكن الإحتلاف بين التوزيعين واضح .

ففي التوزيع أ البيانات تدور حول الوسط وقريبة منه ، أما توزيع ب فهناك اختلاف كبير بين البيانات .

المدى أ = 120-110 = 10

ب = 190-40 = 50

يلاحظ مقدار الإختلاف بين المجموعتين . وبالتالى فأن التوزيع أ اقل تشتتا من التوزيع ب.

المميزات :

1 – أقل دقة

2 – مؤشر سريع للتباين

3 – سهل الإحتساب

4 – تقل كفاءة مع حجم العينة

التباين والانحراف المعياري Variance & Standard Deviation

يعرف التباين بأنه متوسط مجموع مربعات الإنحراف حول الوسط .

$$S^2 = \frac{\sum (x-\bar{x})^2}{n-1} \quad Or \quad \frac{n\sum x^2 - (\sum x)^2}{n-1}$$

والانحراف المعيارى هو الجذر التربيعى للتباين $S = \sqrt{S^2}$

مثال لاحتساب التباين والانحراف المعيارى :

عدد المخالفات	الإنحرافات عن الوسط	مربع الإنحرافات	مربع عدد المخالفات
10	8-	64	100
12	6-	36	144
14	4-	16	196
15	3-	9	225
17	1-	1	289
19	1+	1	361
20	2+	4	400
23	5+	25	529
24	6+	36	576
26	8+	64	676
180	صفر	256	3496

1 – احتساب الوسط = 180\10 = 18 مخالفة

2- بيان مقدار الانحرافات عن الوسط = كل قيمة – الوسط

3 – احتساب مجموع مربع جميع الإنحرافات عن الوسط

التباين = 256/9 = 28.4 او = $\frac{3496-(180\times180)/10}{9} = 256/9$ ، **والانحراف المعياري =** $\sqrt{28.4} = 5.3$

يوجد مقاييس أخرى للتشتت منها على سبيل المثال:

1. معامل الأختلاف C. V
2. القيم المعيارية Standardized Value
3. نصف المدى الربيعى Mid Range

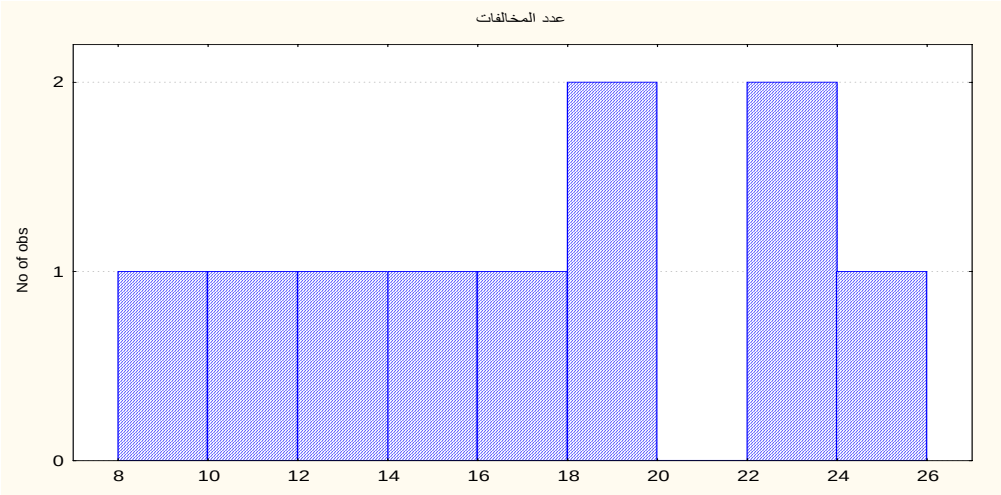
مستخدما بيانات المثال السابق احسب المقاييس الأساسية لها مع عرضها بيانيا.

Statistics

X

N	Valid	10
	Missing	0
Mean		18.0000
Median		18.0000
Mode		10.00[a]
Std. Deviation		5.3333
Variance		28.4444
Range		16.00
Minimum		10.00
Maximum		26.00
Percentiles	25	13.5000
	50	18.0000
	75	23.2500

a. Multiple modes exist. The smallest value is shown



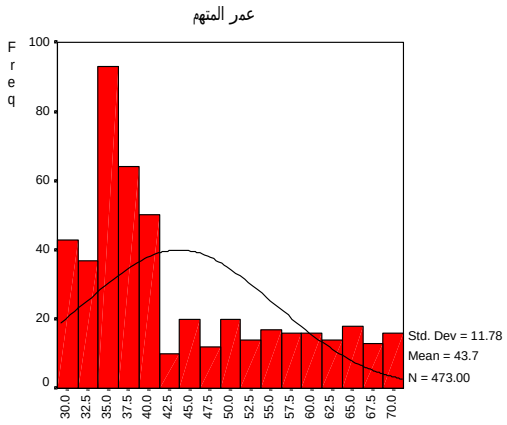
مثال يوضح المقاييس الأساسية والمدرج التكرارى لأعمار المتهمين فى أحد الجرائم

Statistics

عمر المتهم

N	Valid	473
	Missing	1
Mean		43.67
Median		38.00
Mode		36[a]
Std. Deviation		11.78
Variance		138.86
Range		42
Minimum		29
Maximum		71
Percentiles	25	35.00
	50	38.00
	75	52.50

a. Multiple modes exist. The smallest value is shov



الفصل الرابع
اختبارات الفروض

اختبارات الفـروض
TEST OF HYPOTHESES

مقدمة:

قد نرغب فى معرفة ما إذا كانت العينة المختارة من مجتمع ما ممثلة له تماماً وأن أى فروق تظهر فى التحليل لبيانات العينة ما هى إلا فروق عشوائية ناتجة عن طريقة اختيار العينة أو بعض الاختلافات غير الجوهرية بين مفرداتها أم لا.

ونستخدم لذلك اختبارات الفروض (اختبارات المعنوية).

- يختص الاستنتاج الإحصائى بعمل التقديرات للمعالم أو اختبارات الفروض.

الفروض

- أصلية (فرض الصفرى) (فرض العدم) Null hypotheses
- بديلة (فرض بديل) Alternative hypotheses

- الفرض الإحصائى هو الذى يختص بالقيم الحقيقية أو التوزيع الاحتمالى لمجتمع معين أو أنه عبارة عن تقدير قد يكون صحيح أو غير صحيح.
- الأنواع المختلفة من الفروض الممكنة:
  - ... أكبر من قيمة معينة
  - ... أصغر من قيمة معينة
  - ... يساوى قيمة معينة
- أى أن الفرض الذى يجرى اختباره هو الفرض الأصلى ، أما الفرض المقابل هو البديل.
- الفرق المعنوى هو الفرق الذى يعتبر كبيراً جداً لأنه ينسب فقط الى تقديرات المعاينة .
- مستوى المعنوية هو احتمال أن العينة سينتج عنها رفض الفرض الأصلى .

وعادة ما يعرض مستوى المعنوية فى صورة نسبة مئوية 5% أو 1% ولإجراء اختبار معنوية فإنه من الضرورى تكوين فرضاً ومستوى معنوية مناسب .

* **اختبارات الجانب الواحد والجانبين**

- فى حالة ما إذا كانت المشكلة تحت الدراسة تختص بقيم التساوى عن الفرض الأصلى فإنه يطلق على الاختبار الإحصائى فى هذه الحالة أنه اختبار طرفين . أو بمعنى آخر إذا كانت منطقة الرفض تشتمل على مساحات عند طرفى التوزيع .

- أما إذا كانت المشكلة تختص بقيم أصغر أو أكبر من الفرض الأصلى فإن الاختبار الإحصائى فى هذه الحالة يعرف بأنه اختبار ذو جانب واحد . أو بمعنى آخر إذا كانت منطقة الرفض موجودة كلها عند أحد طرفى التوزيع .

ويوضح الجدول التالى الفروض والقرارات الممكنة للفروض الإحصائية:

الفرض الصحيح	القــرار	
	قبول الفرض الأصلى	قبول الفرض البديل
الفرض الأصلى صحيح	لا يوجد خطأ	خطأ النوع الأول I
الفرض البديل صحيح	خطا النوع الثانى II	لا يوجد خطأ

احتمال الخطأ من النوع الأول α ينتج إذا رفض فرض إحصائى وكان يجب قبوله.

$\alpha = P$ (Type I Error). (رفض فرض صحيح).

احتمال الخطأ من النوع الثانى β ينتج إذا تم قبول فرض إحصائى كان يجب رفضه.

$\beta = P$ (Type II Error). (قبــول فرض خطأ).

* أن احتمال وقوع الخطأ من النوع الأول يساوى مستوى المعنوية .

أى أن كلما نقص مستوى المعنوية كلما قل احتمال ارتكاب الخطأ من النوع الأول وقد يثار سؤال إذا كانت هذه هى الحقيقة فلماذا لا نضع α الى أقل حد ممكن ؟.

الإجابة لأنه لحجم عينه ثابت كلما صغر احتمال الخطأ من النوع الأول كلما كبر احتمال الخطأ من النوع الثانى ويجب الموازنة بين هذين النوعين من الأخطاء .

* ويمكن تحديد حجم العينة المختارة وطريقة اختيارها إذا علمنا قيمة α ، β .

$$n = \frac{Z^2 \sigma^2}{D^2}$$

حيث Z, σ, D قيمة الإحصاء Z بمعلومية α ، β والانحراف المعيارى والفرق بين القيمة الفعلية والمقدرة (الخطأ المسموح به).

أما فى حالة النسبة فأن حجم العينة يكون:

$$n = \frac{Z^2 p\, q}{D^2}$$

حيث P النسبة فى المجتمع، $q = 1-P$.

وإذا لم يكن معلوم قيمة P او تقديرها فأننا نختار أقرب تقدير الى 0.5 .

اختبارات الفروض باستخدام الحاسب:

عرفنا احتمال خطأ النوع الأول α واحتمال خطأ النوع الثانى β ويمكن عن طريقها تحديد حجم العينة المناسب وذلك بتثبيت احتمال خطأ النوع الأول عند 5% وتحدد حجم العينة الذى يجعل احتمال خطأ النوع الثانى 10 % تقريبا .

وكذلك فإن الفرق المعنوى الموجود فعلاً فى المجتمع والفرق غير المعنوى لا يوجد فى المجتمع .

ويعطى النظام قيمة P - Value (Significance) فى كل اختبار من الاختبارات .

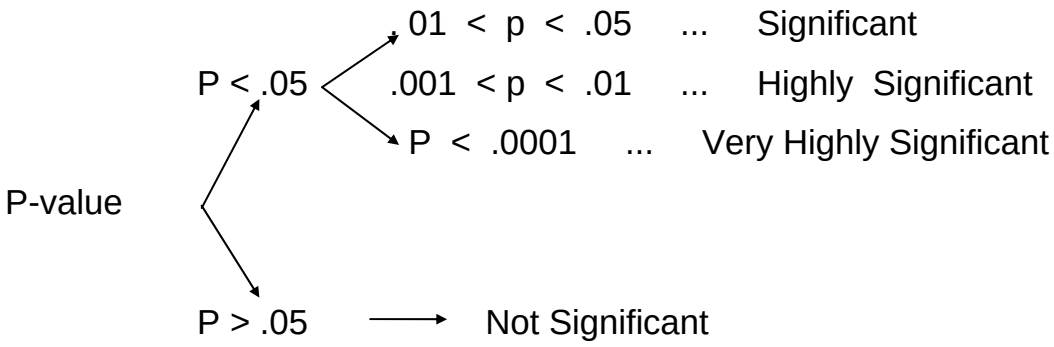
* خطوات أى اختبار من اختبارات الفروض كما يلى :

1. H_0 : (Null hypotheses) تحديد فرض العدم .1
2. H_1 : (Alternative hypotheses) تحديد الفرض البديل .2
3. Calculate Statistic استخدام الإحصاء المناسب .3
4. Compute P - value حساب القيمة الاحتمالية .4
5. Compare p - value مقارنة القيمة الاحتمالية .5

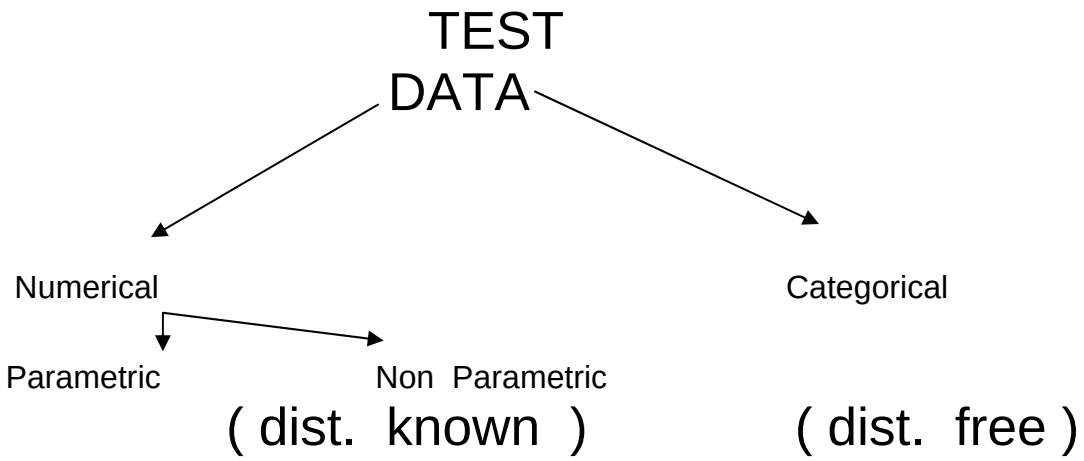
P - value < .05 ... Reject H_0 or Accept H_1

P - value >= .05 ... Don't Reject H_0

P - value ... Where
Observed Significant Level
.05 or .01 ... Nominal Significant Level



يجب ملاحظة أن القيم الناتجة لـ P – Value لاختبار طرفين فإذا أردنا اختبار طرف واحد نقوم بقسمتها على 2 .



One Sample problem		* T - test pairs	* Sign - test * Wilcoxon - test	* Binomial - test
Two Samples	Related (paired)	* T - test pairs	* Sign - test * Wilcoxon - test	* Mcnemar - test
	Independent	* T- test groups	* Mamn - Whitney test	* χ^2 - test
Multi samples problem		* F-test	* Kruskal - Wallis test	* χ^2 - test

ملاحظات :

1. لاختيار الاختبار المناسب يتم ذلك من خلال رسم المدرج التكرارى وبيان الشكل البيانى له :

   أ - إذا كان قريب من التوزيع الطبيعى (او توزيع معلوم أخر) نختار اختبار معلمى.

   ب- إذا كان غير ذلك نختار اختبار لا معلمى .

2. أن الحاسب لا يختار التحليل الملائم للمشكلة المراد تحليلها ولا يتدخل فى ذلك بل يترك للمستخدم اختيار التحليل الملائم وتنفيذه .

3. أن الوصول الى قرار بعدم المعنوية يجب إبرازه فهذه نتيجة يجب على الباحث إظهارها ولا يحاول أن يبدلها او يغيرها او يستبدل البيانات المستخدمة.

مستخدماً البيانات التالية أختبر الفرض القائل أن متوسط X = 25 عند مستوى معنوية 5 % .

Cases	X
1	26.00
2	29.12
3	22.88
4	24.63
5	33.11
6	28.21
7	29.20
8	23.12
9	27.15
10	28.90
11	27.40
12	29.16
13	28.41
14	30.20
15	36.17
16	28.62

Statistics

		X
Mean	Statistic	28.2675
	Std. Error	.8378
Std. Deviation	Statistic	3.3513
Variance	Statistic	11.2314
Skewness	Statistic	.531
	Std. Error	.564
Kurtosis	Statistic	1.199
	Std. Error	1.091

أ - <u>باستخدام الطريقة العادية</u>

1- H_0 : μ = 25
2- H_1 : μ # 25

3- Calculated Statistic $T = \dfrac{\overline{X} - \mu}{\dfrac{s}{\sqrt{n}}} = \dfrac{28.27 - 25}{\dfrac{3.35}{\sqrt{16}}} = 3.9$

4- Tabulated Statistic T_{15} = 2.131 at α = 5%
5- Compare Tabulated Statistic < Calculated Statistic
We accept H_1

ب - <u>باستخدام P - Value</u>

1- H_0 : μ = 25
2- H_1 : μ # 25
3- Calculated Statistic = 3.9
4- Calculate P- Value P (T < 3.9)
= 1 - P ($-\infty$ < T < 3.9) = 0.0007 (< 0.005) Note $T_{15,.005}$ =2.947
5- Compare P Value < α (0.05)
We accept H_1

الفصل الخامس
الارتباط والانحدار

الارتباط Correlation

لقياس ودراسة العلاقة بين متغيرين أو اكثر يكون أمامنا اختيارين أولهما هو قياس درجة العلاقة بين هذه المتغيرات أى هل ترتبط هذه المتغيرات بعلاقة خطية أم غير خطية بصورة أو بأخرى ، وكذلك تعيين اتجاه تلك العلاقة بمعنى هل هى عكسية أم طردية .

أما الهدف الأخر - هو غير مستقل عن الأول - فهو استخدام تلك العلاقة بعد تحديدها فى التنبؤ بقيم متغير ما بدلالة المتغير أو المتغيرات الأخرى .

والهدف الأول هو مجال دراسة الارتباط Correlation أما الثانى وهو تعيين العلاقة واستخدامها فى التنبؤ فهو مجال دراسة الانحدار Regression .

ولبيان صورة مبدئية عن شكل العلاقة بين متغيرين ، يتم ذلك من خلال الشكل الانتشارى Scatter Diagram وهو الشكل البيانى الناتج عن انتشار النقط بين المحورين المتعامدين وأهم مقاييس الارتباط يمكن تقسيمها حسب قياس المتغيرات الى ما يلى :

أولاً : الارتباط لمتغيرات مقاسة بفترة او نسبة Ratio or Interval

1. معامل الارتباط الخطى البسيط Simple Linear Correlation Coefficient .
2. معامل إتيا Eta Coefficient .

ثانياً : الارتباط لمتغيرات مقاسة بمقياس ترتيبى Ordinal

3. مقياس جاما لجودمان وكيرسكال Goodman & Kurskall Gama .
4. مقياس كندل تاو B Kendell's Taw - B .
5. مقياس كندل تاو C Kendell's Taw - C .
6. مقياس سومرز Somer's D .
7. معامل ارتباط الرتب لسبيرمان Spearman's rho .

ثالثاً : الارتباط لمتغيرات مقاسة بمقياس أسمى Nominal

8. معامل فاى The Phi Coefficient .
9. معامل الاقتران The Coefficient of Contingency.
10. معامل كرامير Cramers Coefficient .
11. معامل لامدا Lambda Coefficient .
12. معامل عدم التأكد Uncertainty Coefficient .

وسنعطى فكرة مبسطة عن كل مقياس .

أولاً : الارتباط لمتغيرات مقاسة بفترة أو نسبة

1- معامل الارتباط الخطى البسيط Simple Linear Correlation Coefficient

يستخدم لقياس العلاقة بين المتغيرات الكمية وأهم الملاحظات عليه :

1. يستخدم لقياس درجة قوة العلاقة الخطية بين متغيرين فكلما زادت قيمته كلما زاد تركز المفردات حول خط الانحدار .
2. أن قيمته تتراوح بين 1 ، -1 فإذا كانت قيمته 1 كانت العلاقة طردية تامة ، وإذا كانت صفر فتعنى عدم وجود علاقة خطية ولكن لا يعنى وجود علاقة غير خطية وإذا كانت قيمته -1 فإن العلاقة عكسية تامة .
3. إن إشارة معامل الارتباط تبين اتجاه العلاقة فعندما تكون موجبة تكون العلاقة طردية وسالبة حين تكون العلاقة عكسية .
4. معامل التحديد R^2 هو مربع معامل الارتباط الخطى وقيمته محصورة ما بين الصفر والواحد وهو بين النسبة التى يفسرها خط الانحدار .

ويحسب هذا المقياس طبقا للعلاقة الآتية :

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

حيث :

σ_x Standard deviation for variable X

σ_y ,, ,, ,, ,, Y

σ_{xy} Covariance between X, Y .

2- معامل إتيا

يحسب بالعلاقة التالية :

$$\sqrt{\frac{\text{Explaind Variance}}{\text{Total Variance}}}$$

وتتراوح قيمته ما بين الصفر والواحد. ويحسب لمتغيرين أحدهما اسمى والأخر فترى و لابد من توافر شروط مثل الخطية والتبعية للتوزيع الطبيعى.

ثانيا : الارتباط لمتغيرات مقاسه بمقياس ترتيبى

3- مقياس جاما لجودمان وكيرسكال

يحسب هذا المقياس بالعلاقة التالية :

$$\text{Gama} = \sqrt{\frac{Q - P}{Q + P}}$$

حيث : P عدد الأزواج المتفقة و Q عدد الأزواج المختلفة.

4- مقياس كندال تاو B

يأخذ هذا المقياس فى الاعتبار التنقلات فى العينة الأولى T_X والثانية T_Y ويحسب بالعلاقة :

$$\tau_b = \frac{P - Q}{\sqrt{(P + Q + T_X) \times (P + Q + T_Y)}}$$

ويحسب هذا المقياس للجداول المربعة فقط .

5- مقياس كندال تاو C

يحسب لجميع الجداول المربعة وغير المربعة بالعلاقة :

$$\tau_c = \frac{2m(P - Q)}{N^2 (m - 1)}$$

حيث : m عدد الصفوف أو الأعمدة أيهما أقل .

6- مقياس سومزر D

يحسب كالاختبارات السابقة ولكنه يختلف عنها فى أنه يحسب فى حالة التماثل ثم فى حالة المتغير الأول مستقل ثم فى حالة المتغير الثانى مستقل .

7- معامل الارتباط الرتب لسبيرمان

يستخدم هذا المقياس لقياس العلاقة بين متغيرين إذا ما كانت القياسات ترتيبية حيث لا يصلح معامل الارتباط الخطى البسيط لبيرسون لقياس تلك العلاقة ، ويحسب بالعلاقة التالية:

$$\text{معامل ارتباط سيبرمان} = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

حيث $\sum d^2$ مجموع مربع الفروق بين رتب المتغير الأول ورتب المتغير الثانى ،

n عدد أ زواج القيم .

ثالثاً : الارتباط لمتغيرات مقاسه بمقياس أسمى

8- معامل فاى Φ

يحسب هذا المقياس معتمداً على χ^2 وحجم العينة N بالعلاقة التالية :

$$\Phi = \sqrt{\frac{\chi^2}{N}}$$

9- معامل الاقتران

يحسب للجدوال أكبر من 2 X 2 معتمداً χ^2 وحجم العينة N بالعلاقة التالية :

$$CC = \sqrt{\frac{\chi^2}{\chi^2 + N}}$$

10- معامل كرامير

يعتمد هذا المقياس على χ^2 وحجم العينة N وعدد الصفوف أو الأعمدة أيهما أقل (K) ويحسب بالعلاقة التالية :

$$V = \sqrt{\frac{\chi^2}{N(K-1)}}$$

11- معامل لامدا

يحسب كالاختبارات السابقة ولكنه يعطى ثلاث قيم للارتباط الأولى فى حالة التماثل والثانية فى حالة كون المتغير الأول مستقل والثالثة فى حالة كون المتغير الثانى مستقل.

$$\lambda = \frac{\text{Misclassified in situation 1} \ - \ \text{Misclassified in situation 2}}{\text{Misclassified in situation 1}}$$

12- معامل عدم التأكد

ويحسب هذا المقياس وتتراوح قيمته ما بين الصفر والواحد. و لابد من توافر شروط مثل الخطية والتبعية للتوزيع الطبيعى .

من الجدير بالذكر انه توجد مقاييس اخرى مثل:

Odds Ratio, Relative Risk, Kappa

الانحدار Regression

اولا: الانحدار البسيط

أن العلاقة الخطية بين متغيرين Y , X هى :

$$Y = A + BX$$

حيث : Y Dependent variable المتغير التابع .

X Independent variable المتغير المستقل .

ويتم تقدير معالم النموذج بطريقة المربعات الصغرى (OLS) .

ويظهر بمخرجات SPSS , Model Summary وبه $R, R^2, adj.R^2, S.E$

حيث : R هى معامل الارتباط ، و R^2 معامل التحديد ،

$adj.R^2$ معامل التحديد المصحح ، و S.E الخطأ المعيارى للتقدير.

2- Anova

ولبيان معنوية الانحدار (النموذج) يعطى جدول تحليل تباين الانحدار. ويقبل كلما كانت قيمة معنوية F أقل من 5% .

3- Coefficients

ويظهر بها Value in Equation and Testing

ويظهر قيمتين :

أ- تحت عمود B وهى القيمة غير المعيارية .

ب- تحت عمود Beta قيم معيارية .

ونحن نأخذ القيمة غير المعيارية منها نحدد A (Constant) ، B (Slope)

و يلى ذلك عمود SE(B) وهو عمود الخطأ المعيارى للتقدير B

ثم عمود لقيمة T test و الذى ينتج من قسمة B / SE (B)

ثم عمود معنوية Sig. T test ويقبل كلما كانت القيمة أقل من 5 % .

ثانيا: الانحدار المتعدد

تكون فيه العلاقة بين المتغيرات كما يلى :

$$Y = b0 + b1\ X1 + b2\ X2 + b3\ X3 + \ldots + bn\ Xn$$

حيث : Y المتغير التابع Dependent variable .

X1 to Xn المتغيرات المستقلة Independent variables .

يوجد ثلاث طرق للانحدار المتعدد كل منها يمثل طريقة لمعالجة مشكلة الانحدار وهى:

1- الانحدار العيارى Standard Regression ونقوم بإدخال جميع المتغيرات المستقلة بنموذج الانحدار مرة واحدة بغض النظر عن أية معايير إحصائية للدخول بالنموذج.

2- الانحدار الهرمى Hierarchical Regression ونقوم بإدخال جميع المتغيرات المستقلة بنموذج الانحدار تباعا ونحدد ترتيب دخول هذه المتغيرات بالنموذج على اساس إحصائى نظرى.

3- الانحدار التدريجى Stepwise Regression ونقوم بإدخال عدد من المتغيرات المستقلة بنموذج الانحدار وليس كلها، ونحدد ترتيب دخول اى متغير مستقل بالنموذج طبقا لمعيار إحصائى Criteria ، ويكون الإدخال بالنموذج امامى Forward او خلفى Backward.

والمفاضلة بين الطرق الثلاثة تختلف من باحث لأخر وتتوقف على خبرتة الإحصائية والنظرية بموضوع دراستة.

وتظهر نفس النتائج السابقة كما بالانحدار البسيط .

إذا حدث مخالفة لفرض من فروض المربعات الصغرى فإنه يؤدى بالطبع الى مشاكل فى التقدير تبرز مبدئيا من **تحليل البواقى**، وأهم هذه المشاكل وطرق اكتشافها وعلاجها:

(اكتشاف المشكلة - علاج المشكلة)

1- الارتباط الذاتى. (اختبار دوربن واطسون - متغيرات إبطاء مع r)

2- تعدد العلاقات الخطية . (ارتباط المستقلين - الانحدار التدريجى)

3- عدم ثبات تباين الحد العشوائى . (شكل البواقى - تحويلة رياضية)

4- وجود ارتباط بين المتغير العشوائى ومتغير مستقل. (البواقى مع المستقل - تحويلة رياضية)

5- المتغير التابع و العشوائى لا يتبع التوزيع الطبيعى . (اختبار التبعية - تحويلة رياضية)

6- أخطاء فى قياس المتغيرات عند التجميع . (الدالة العكسية - تحويلة رياضية)

نموذج الانحدار اللوجستى:

يعتبر نموذج الانحدار الخطى اللوجستى من أهم النماذج المستخدمة لصياغة دالة التمييز، ويتميز هذا النموذج بملاءمته للعديد من الاستخدامات ، ويستخدم الانحدار اللوجستى بصفة عامة لتحليل العلاقة بين ما حدث فى الماضى وما يمكن أن يحدث فى المستقبل .

وينشأ الانحدار اللوجستى عن استخدم تحويلة اللوجيت السابقة عندما يكون المتغير التابع (Y) يعبر عن احتمال حدوث حدث معين (θ) ، وهناك مجموعة من المتغيرات التفسيرية { X_1 , X_2, ... , X_p } = X عددها P متغير ، وهناك علاقة بين الاحتمال (θ) والمتغيرات التفسيرية على الصورة الخطية التالية :

$$\theta_k = \beta_0 + \beta_1 x_{1k} + \beta_2 x_{2k} + \ldots + \beta_p x_{pk}$$

حيث أن الطرف الأيسر محدود ($0 \leq \theta_k \leq 1$) بينما الطرف الأيمن من المعادلة غير محدود .

وباستخدام تحويلة اللوجيت نصل إلى أن احتمال الحدوث :

$$P_r (Y_k = 1/X) = \frac{\exp(\beta' X)}{1+\exp(\beta' X)} = P_r(X)$$

الفصل السادس
اختيار الأسلوب الإحصائى
للتحليل متعدد المتغيرات

Rough guide to multivariable methods
Classification of Variables

Method	Dependent	Independent
Multiple regression analysis	Continuous	Classically all continuous, but in practice any type(s) can be used. (To describe the extent, direction, and strength of the relationship between several independent variables and a continuous dependent variable.)
Analysis of variance	Continuous	All nominal. (To describe the relationship between a continuous dependent variable and one or more nominal independent variables.)
Analysis of covariance	Continuous	Mixture of nominal variables and continuous variables. (the latter used as control variables)* (To describe the relationship between a continuous dependent variable and one or more nominal independent variables, controlling for the effect of one or more continuous independent variables.)
Discriminant (Logistic)	Nominal	Classically all continuous, but in practice a mixture of various types analysis can be used as long as some are continuous. (To determine how one or more independent. variables can be used to discriminate among different categories of a nominal dependent variable.) logistic if dependent variable two category only.
Factor	(The variables used in a factor analysis are classically continuous, but in practice analysis may be of any type. These variables are not clearly identifiable as either dependent or independent, although the resulting factors may be used as dependent or independent variables in a later analysis.) (To define one or more new composite variables called factors from other, specially constructed or reduced variables.)	
Categorical data analysis	Nominal	Mostly nominal, but sometimes ordinal using linear models. (To describe the relationship between a nominal dependent variable and several nominal or ordinal independent variables, although applications to situations involving only dependent variables are possible.)

* Generally speaking, a control variable is a variable that has to be considered before any relationships of interest can be quantified; this is because a control variable may be related to the variables of primary interest and must be taken into account in studying the relationships among the 'primary variables. For example, in describing the relationship between blood pressure and physical activity, we would probably consider "age" and "sex" as control variables because they are-related to blood pressure and physical activity and so, unless taken into account, could confound any conclusions regarding the primary relationship of interest.

الأساليب المختلفة للتحليل الإحصائى للمتغيرات المتعددة

يعتمد التحليل الإحصائى للمتغيرات المتعددة من مجموعة من الأساليب بعضها موجود ومألوف للباحثين مثل تحليل الارتباط والانحدار وتحليل التباين والتحليل العاملى ، والبعض الأخر اكثر حداثة وغير مألوف للباحثين مثل تحليل المسارات وارتباط كانونكل ونماذج المعادلات الخطية ... ، وفيما يلى تعريف مبسط لأهم تلك الأساليب.

تحليل المكونات والتحليل العاملى Principle Components and Factor Analysis

أسلوب إحصائى يستخدم لتحليل العلاقات المتداخلة بين عدد كبير من المتغيرات، وشرح تلك المتغيرات بواسطة الأبعاد (العوامل)، والهدف من ذلك التحليل تقليل المتغيرات الأصلية والاستعاضة عنها بعدد أقل من المتغيرات بدون فقدان الكثير من المعلومات، ويعتبر هذا الأسلوب الأساس الذى يستعمل فى تصميم المقاييس المجمعة.

تحليل الانحدار المتعدد Multiple Regression Analysis

أسلوب إحصائى يستخدم عندما يكون هناك متغير تابع كمى يفترض وجود علاقة بينه وبين مجموعة من المتغيرات المستقلة الكمية. والهدف منه هو بيان اثر ومسئولية كل متغير مستقل فى تفسير التابع ، بالإضافة الى التنبؤ بالتغير فى التابع بمعرفة التغيرات فى المتغيرات المستقلة.

تحليل الانحدار اللوجيستى المتعدد Multiple logistic Regression Analysis

أسلوب إحصائى يستخدم عندما يكون هناك متغير تابع وصفى (نوعى اسمى او ترتيبى) يفترض وجود علاقة بينه وبين مجموعة من المتغيرات المستقلة الكمية والوصفية. والهدف منه هو بيان اثر ومسئولية كل متغير مستقل فى تفسير التابع ، بالإضافة الى التنبؤ بالتغير فى التابع بمعرفة التغيرات فى المتغيرات المستقلة.

تحليل التباين والتغاير المتعدد Multivariate Analysis of Variance & Covariance

أسلوب إحصائى يستخدم عندما يكون هناك مجموعة متغيرات تابعة (اثنين على الأقل) كمية يفترض وجود علاقة بينه وبين مجموعة من المتغيرات المستقلة الوصفية (يسمى تحليل تباين) ، وإذا أضيف مجموعة من المتغيرات المستقلة الكمية الى الوصفية السابقة (يسمى تحليل تغاير). والهدف منه هو بيان اثر ومسئولية كل متغير مستقل فى تفسير التابع ، بالإضافة الى بيان الاختلاف المعنوى لتقسيمات المتغيرات المستقلة الوصفية.

تحليل التمايز المتعدد Multiple Discriminant Analysis

أسلوب إحصائى يستخدم عندما يكون هناك متغير تابع وصفى له وجهان أو أكثر يفترض وجود علاقة بينه وبين مجموعة من المتغيرات المستقلة الكمية أو الوصفية أو كلاهما. والغرض منه هو فهم الفروق بين اوجه المتغير التابع والتنبؤ بها بمعلومية المتغيرات المستقلة.

التحليل المتصل Conjoint Analysis

أسلوب إحصائى يستخدم لتقييم الأشياء (مثل المنتجات والإصدارات الجديدة والخدمات) مع الحفاظ على سهولة هذا التقسيم ، حيث يتم تحديد أهمية الخصائص المختلفة ومستوى كل خاصية فى الوقت الذى يقوم المستقصى منه بتقييم خصائص محدودة للمنتج فى صورة مجموعة من التوليفات.

تحليل (الارتباط التوافقى) ارتباط كانونكل Canonical Correlation

أسلوب إحصائى يستخدم عندما يكون هناك مجموعة من المتغيرات التابعة الكمية و مجموعة من المتغيرات المستقلة الكمية وهو يختلف عن الانحدار فى أن التابع متغير واحد كمى بينما هنا مجموعة من المتغيرات التابعة الكمية. والهدف منه هو تعظيم الارتباط بين مجموعة من المتغيرات المستقلة ومجموعة من المتغيرات التابعة عن طريق الحصول على توليفة خطية لمجموعة المتغيرات المستقلة وأخرى لمجموعة المتغيرات التابعة تعظم الارتباط بين المجموعتين ، بالتالى فهذا الأسلوب يحدد مجموعة من الأوزان لكل من المتغيرات المستقلة و المتغيرات التابعة تمكن من الحصول على أعلى ارتباط بينهما.

تحليل المجموعات (العناقيد) Cluster Analysis

أسلوب إحصائى يستخدم لتقسيم الأشياء أو الأشخاص أو المفردات (حالات أو متغيرات) الى عدد صغير من المجموعات الفرعية غير المتداخلة والشاملة بناء على مدى التشابه والتوافق بينهم مع الحفاظ على سهولة هذا التقسيم ، حيث يتم تحديد أهمية الخصائص المختلفة لكل مجموعة.

القياس متعدد الأبعاد Multidimensional Scaling

أسلوب إحصائى يستخدم لتقسيم الأشياء أو المفردات الى مسافات فى فراغ متعدد الأبعاد لقياس التشابه أو التفضيل بين الأشياء ووضع المتشابهين بالقرب من بعضهما ، وذلك حسب طبيعة القياس.

تحليل التوافق Correspondence Analysis

أسلوب إحصائى يتعامل مع المتغيرات المرتبطة داخليا والتى تساعد على تقليل الأبعاد التى تستخدم لترتيب الأشياء ، بالإضافة الى تحديد الخريطة التى تربط هذه الأشياء بالخصائص. حيث يستخدم المتغيرات الوصفية سواء أكانت اسمية أو ترتيبية بالاعتماد على الجداول التكرارية المزدوجة أو المتعددة.

التحليل اللوغاريتم خطى Log Linear Analysis

يعتبر من الأساليب شائعة الاستخدام فى العلوم الاجتماعية والسكانية عامة والطبية منها خاصة. ويستخدم عندما تكون جميع المتغيرات وصفية أو مبوبة فى جداول تكرارية ثنائية أو اكثر، وذلك لدراسة الآثار المشتركة لمتغيرات الدراسة ، وبناء نموذج (تحويل المتغيرات الوصفية لنماذج كمية باستخدام تحويلة اللوجيت).

تحليل المسارات (نموذج معادلات البناء) Path Analysis (Structural Equ. Model)

أسلوب إحصائى يستخدم عندما نريد ربط مجموعة من المتغيرات والتى تكون تابعة تارة ومستقلة تارة أخرى حيث تحل معادلات الانحدار المتعدد فى نفس الوقت وذلك لبيان الآثار المباشرة وغير المباشرة للمتغيرات الداخلة فى النموذج.

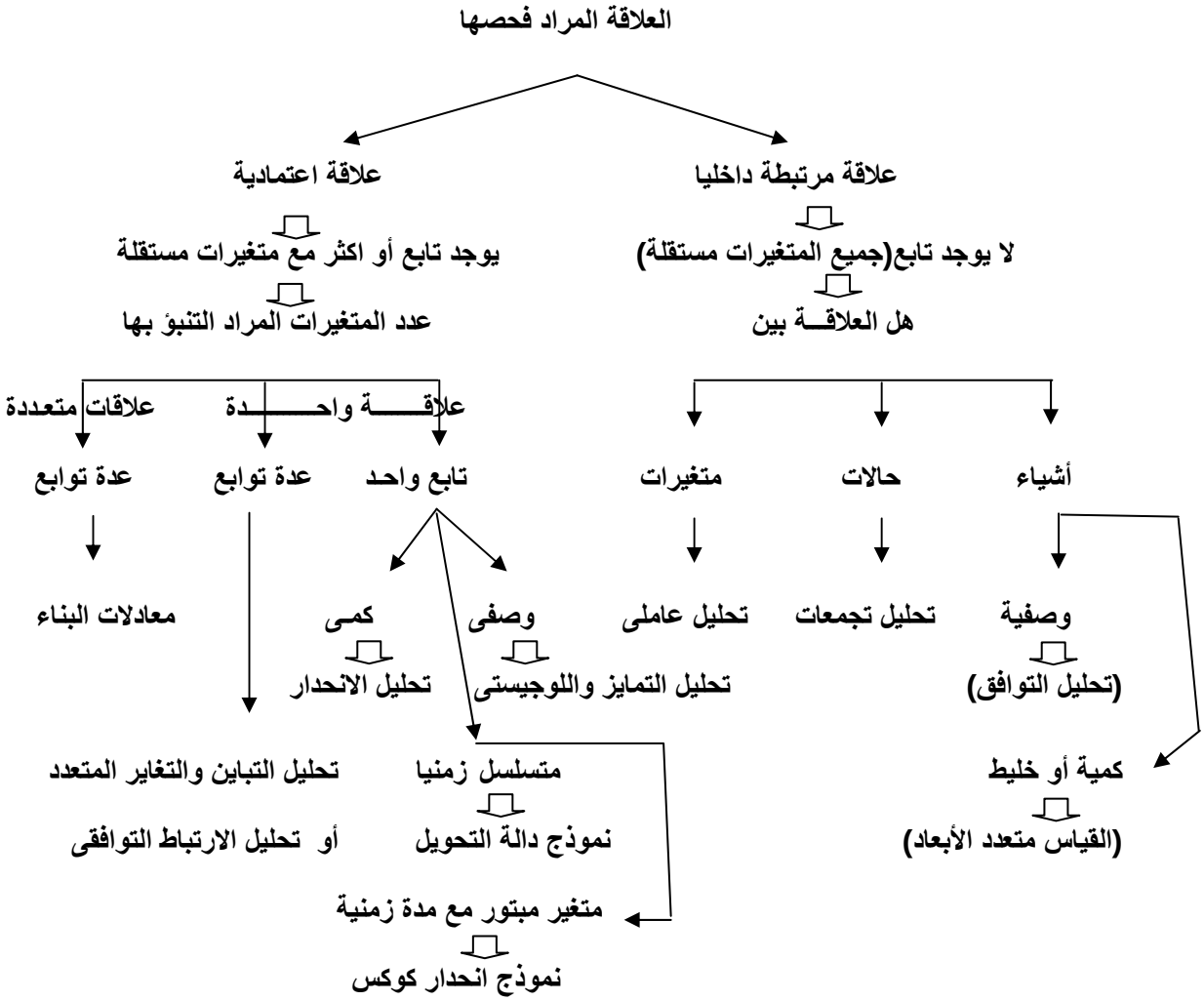
تحليل البقاء (انحدار كوكس) Survival Analysis

يعتبر من الأساليب شائعة الاستخدام فى العلوم الطبية خاصة والاجتماعية عامة وذلك لعدة أسباب منها أن التقديرات المحسوبة باستخدام الانحدار مقبولة النتائج خاصة بعد التغلب على بعض المشاكل التى تنشأ نتيجة غياب بعض فروض النموذج . حيث يقوم النموذج على فرض أساسى أن المتغير التابع وصفى مبتور Censored مع مدة البقاء حتى يمكن تقدير معدل المخاطرة أو الفشل.

تحليل السلاسل الزمنية (نموذج دالة التحويل) Time Series Analysis

أسلوب إحصائى يستخدم عندما نرغب فى بناء نموذج يتضمن اكثر من سلسلة زمنية واحدة، ويسمى مثل هذا النموذج بنموذج السلسلة الزمنية المتعدد أو نموذج دالة التحويل. ومن الأمثلة الشائعة لهذا النوع السلسلة الخاصة بالتنبؤ بالمبيعات باستخدام المنفق على الإعلان وحجم المشتريات و أسعارها.

ويوضح الشكل التالى الأسلوب الإحصائى المناسب لنوع العلاقة التى يراد فحصها حسب الاعتمادية وعدد المتغيرات التابعة ونوع القياس.



الإطار العام للتحليل الإحصائى للمتغيرات المتعددة:

أن هناك العديد من الأمور التى يجب أخذها فى الاعتبار للتطبيق السليم لأدوات التحليل الإحصائى المتقدم أكثر من مجرد اختيار الأسلوب المناسب، والخطوات التالية يمكن أن ترشد الباحث مهما اختلف أسلوب التحليل المختار:

1. تحديد مشكلة البحث وأهدافه والأسلوب المستخدم.
2. تصميم خطة التحليل.
3. تقييم مدى توافر شروط تطبيق أسلوب التحليل.
4. تقييم الملائمة الكلية لنموذج التحليل.
5. تفسير التوليفة الخطية.
6. تقييم صلاحية نموذج التحليل وتعميم النتائج.

جدول الاستشارات الإحصائية (نوع النموذج والمؤشرات والأساليب والاختبارات الإحصائية المناسبة)										
نوع النموذج الإحصائي المترجم للظاهرة	حالات النموذج	المتغيرات المستقلة				المتغيرات التابعة				المؤشرات والأساليب والاختبارات الإحصائية المناسبة
		منفصلة وعددها	متقطعة		يوجد تفاعل أو علاقة		منفصلة وعددها	متقطعة وعددها	يوجد تفاعل أو علاقة	
			عددها	مستوياتها	نعم	لا			نعم	لا
1- النموذج صفر: حيث يهتم الباحث بالدراسة الإحصائية لمتغير واحد فقط ، أو متغير واحد فقط كل مرة لظاهرة متعددة المتغيرات، والمتغير هنا مجرد متغير واحد حيث لا معنى لصفة مستقل وتابع .	أ	1	-	-	-	-	-	-	-	-
	ب	-	1	2≤	-	-	-	-	-	-
2- النموذج البسيط حيث يهتم الباحث بمتغيرين فقط أو متغيرين كل مرة من متغيرات ظاهرة متعددة المتغيرات. غالبا ما يستطيع الباحث التمييز بين ما هو تابع وما هو مستقل في مثل هذا النموذج وإذا لم يستطع فالمتغيران يعتبران في مجرد علاقة مطلوب.	أ	-	1	2	-	-	1	-	-	-
	ب	-	1	2<	-	-	1	-	-	-
	ج	1	-	-	-	-	1	-	-	-
	د	1	-	-	-	-	1	-	-	-
	هـ	-	1	-	-	-	-	1	-	-
	و	1	-	-	-	-	-	1	-	-

المؤشرات الإحصائية مثل: المتوسط – الوسيط – التباين- الانحراف المعياري – الالتـواء –التفرطح.	-
الاختبارات الإحصائية: مثل اختبار(ت) T-test واختبار(ذ)Z لاختبار الفروض حول متوسط المجتمع μ ، (ذ) في حالة إذا كان حجم العينة كبيرا .	-
المؤشرات الإحصائية مثل : النسبة المئوية	-
الاختبارات الإحصائية مثل : اختبار كا^2 لجودة المطابقة - كا^2 لدلالة الفروق بين نسبتين – اختبار (ذ) لدلالة الفرق بين نسبتين في حالة إذا كان حجم العينتين كبيرا.	-
اختبار(ت) T-test لدلالة الفرق بين متوسطي مجتمعين أو اختبار(ذ) Z لدلالة الفرق بـين مجتمعين في حالة كبر حجم العينتين.	-
أسلوب تحليل التباين الأحادي ANOVA مع اسـتخدام اختبـار(ف)F-test المـصحوب باختبارات المقارنات الزوجية المتعددة مثل شيفية.	-
أسلوب تحليل الارتباط والانحدار البسيطين المصحوبين باختبارات(ت) و(ف) F المناسبين.	-
إذا كانت المتغيرات فى تسلسل زمنى يستخدم أسلوب تحليل السلاسل الزمنية.	-
اختبار(كا^2) لاستقلالية متغيرين المصحوب بحـساب Lambda (λ) أو Uncertainty Coefficient (U)	-
أسلوب التحليل التمييزي البسيط Simple Discriminate Analysis	-

نوع النموذج الإحصائي المترجم للظاهرة	حالات النموذج	المتغيرات المستقلة				المتغيرات التابعة				المؤشرات والأساليب والاختبارات الإحصائية المناسبة
		منفصلة وعددها	متقطعة		يوجد تفاعل أو علاقة	منفصلة وعددها	متقطعة وعددها	يوجد تفاعل أو علاقة		
			عددها	مستوياتها	نعم	لا		نعم	لا	
3- النموذج المتعدد وهنا نترجم الظاهرة إحصائيا من خلال العديد من المتغيرات واحد منها فقط متغير تابع والأخريات متغيرات مستقلة. المتغيرات المستقلة قد تكون كلها أو بعضها في حالة ترابط أو تفاعل	أ	-	≤2	≤2	√	√	1	-	-	تحليل التباين ذو البعدين أو اكثر ANOVA 2- Way Or more
	ب	≤1	≤1	≤2	-	√	1	-	-	تحليل التغاير ANCOVA
	ج	≤1	≤1	≤2	-	√	1	-	-	تحليل التفاعل(Trait-Treatment Interaction(TTI
	د	≤1	-	-	√	√	1	-	-	تحليل الانحدار المتعددMultiple Regression
	هـ	≤1	≤1	≤2	√	√	-	1	-	التحليل التمييزي المتعدد Multiple Discriminate Analysis
	و	≤1	-	-	√	√	1	-	-	تحليل الارتباط المتعدد، حيث يحسب معامل الارتباط المتعدد بين متغير وبقية المتغيرات مجتمعة مقدرة كل مرة (R)
4- النموذج المتعدد المتدرج هو نفسه النموذج المتعدد ولكن عدد المتغيرات التابعة والتي قد تكون كلها أو بعضها مرتبطة – عددها أكثر من متغير تابع واحد. أما الحالة (ز) من هذا النموذج فيختفي الفارق بين ما هو تابع وما هو مستقل فالمتغيرات تصنف إلى مجرد مجموعتين من المتغيرات	أ	-	1	2	-	-	1<	-	√	اختبار هوتلنج Hotelling's
	ب	-	1	2<	-	-	1<	-	√	تحليل التباين المتعدد الاحادي MANOVA
	ج	-	≤2	≤2	√	√	1<	-	√	تحليل التباين المتعدد المتدرج MMANOVA
	د	≤1	≤1	≤2	-	√	1<	-	√	تحليل التغاير المتعدد المتدرج MANCOVA
	هـ	≤1	≤2	≤2	√	-	1<	-	√	تحليل التفاعل المتعدد المتدرج MMTTI
	و	≤1	-	-	√	√	1<	-	√	تحليل الانحدار المتعدد المتدرج MMRA
	ز	≤1	-	-	√	√	1<	-	√	تحليل الارتباط التوافقي المتعدد Canonical Correlation

نوع النموذج الإحصائي المترجم للظاهرة	حالات النموذج	المتغيرات المستقلة				المتغيرات التابعة				المؤشرات والأساليب والاختبارات الإحصائية المناسبة
		منفصلة وعددها	متقطعة		يوجد تفاعل أو علاقة		منفصلة وعددها	متقطعة وعددها	يوجد تفاعل أو علاقة	
			عددها	مستوياتها	نعم	لا			نعم	لا
5- النموذج المتدرج : حيث يختفي الفارق بين ماهو تابع وماهو مستقل ولكن المتغيرات المترجمة للظاهرة تتصف بالتعدد هذه المتغيرات قد تكون كلها أو بعضها في حالة ارتباط	أ	متغيرات عددها n في مجموعة واحدة حيث يختفي الفارق بين ما هو تابع وما هو مستقل بهدف الحصول على تحويلات خطية متعامدة عددها m حيث $m \leq n$								تحليل المحاور المتعامدة Principle Component Analysis
	ب	نفس الحالة (أ) ولكن الهدف اكتشاف البناء العاملي لهذه المتغيرات والذى يتكون من m العوامل حيث $m < n$								التحليل العاملى الاستكشافى Expletory Factor Analysis
	ج	نفس الحالة (أ) حيث مجموعة المتغيرات أمكن اكتشاف بناؤها العاملي أو أمكن صياغة فروض حوله والمطلوب توكيد هذا البناء								التحليل العاملى التوكيدى Confirmatory Factor Analysis
	د	الحالتين (أ،ب) السابقتين تستلزمان حساب مصفوفة ارتباط بيرسون كمدخل للتحليل الإحصائي أما في حالة (د) المصفوفة المدخلة محسوبة من تشابهات similarities أو عدم تشابهاتdissimilarities بين أزواج المتغيرات أو بين أزواج الحالات كما انها تعمل في حالة ان تكون المصفوفة مكونة من معاملات ارتباط بيرسون								يمكن استخدام الأسلوبين: تحليل المجموعات Cluster Analysis التحليل متعدد الأبعاد Multidimensional Scaling
6- نموذج تحليل المسارات	أ	متغيرات فى مجموعة واحدة حيث يكون التابع مستقلا أو العكس فى نفس الوقت، ويكون الهدف الحصول على النموذج السببى بين المتغيرات								تحليل المسارات Path Analysis
	ب	نفس الحالة (أ) ولكن بهدف العثور على النموذج السببى لعوامل البناء العاملى لمجموعة المتغيرات								نماذج البناء الخطى LISREL
7- النموذج الشامل		هو ذلك النموذج الإحصائى المترجم للظاهرة تحت الدراسة والذى يجمع ما بين كل أو بعض حالات النموذجين رقم (5) و (6) مع كل أو بعض حالات النماذج السابقة عليهما								حسب الحالة

المراجـع

أحمد عبادة سرحان (1965): "طرق التحليل الإحصائى" – دار المعارف – القاهرة – مصر.

جلال مصطفى الصياد (1993): الاستدلال الإحصائى" – دار المريخ – الرياض – السعودية.

عبد الجبار توفيق (1985):"التحليل الاحصائى فى البحوث التربوية والنفسية والاجتماعية والنفسية : الطرق اللامعلمية" الطبعة الثانية – مؤسسة الكويت للتقدم العلمى – الكويت.

عبد اللطيف أبو العلا (1994): " الأسلوب الإحصائى – الطرق والتحليل " – مكتبة الجلاء – المنصورة - مصر.

عبدالحميد العباسى (2002): "التحليل الاحصائى البسيط والمتقدم باستخدام SPSS" – معهد الدراسات والبحوث الاحصائية جامعة القاهرة – الجيزة – مصر.

عبدالحميد العباسى (1999): "التحليل الاحصائى باستخدام SPSS" – معهد الدراسات والبحوث الاحصائية جامعة القاهرة – الجيزة – مصر.

فؤاد ابو حطب و آمال صادق (1996): "مناهج البحث وطرق التحليل الاحصائى: فى العلوم النفسية والتربوية والاجتماعية – مكتبة الانجلو المصرية – القاهرة – مصر.

Aitkin, M., Anderson, D. A., Francis, B. and Hinde, J. P. (1989). Statistical Modeling in GLIM, Clarendon Press, Oxford.

Anderson, T.W, (1984), An Introduction to Multivariate Statistical Analysis (2nd) New York John Wiley.

Bagozzi, R. P., & Yi, Y. (1989). On the use of structural equation models in experimental design. Journal of Marketing Research , 26 , 271-284.

Cronbach, L. J. (2004). My current thoughts on Coefficient Alpha and successor procedures. *Educational and Psychological Measurement, 64*, 391-418.

Dobson, A. J. (1990). An Introduction to Generalized Linear Models, Chapman and Hall, London.

Draper, N.R. and H. Smith (1981). Applied Regression Analysis, Second Edition, John Wiley & Sons, Inc.

Johnson, Richard A. and W. Wichern, (1996). Applied Multivariate Statistical Analysis (4th), New York Partice-Hel international.

Ostle, B., & Malone, L. C. (1988). Statistics in research: Basic concepts and techniques for research workers (4th ed.). Ames, IA: Iowa State Press.

Stevens, J. (1986). Applied multivariate statistics for the social sciences. Hillsdale, NJ: Erlbaum.

الأسلوب الإحصاء: التحليل والتفسير

باستخدام الحاسب وبرنامج SPSS

- أول مرجع من نوعه يتناول التحليلات الإحصائية البسيطة والمتقدمة باستخدام SPSS.
- يحتوى المرجع على ترجمة لكافة المصطلحات المرتبطة بالأساليب الإحصائية المتقدمة.
- يمكن استخدام هذا المرجع بواسطة كل من المبتدئين وأصحاب الخبرة فى مجال البحث والتحليل الإحصائى.
- المرجع يربط بين النظرية والتطبيق ويحتوى على العديد من الأمثلة التى تساعد الأكاديميين والممارسين للتحليل الإحصائى.
- المرجع يوضح كيفية التعليق على المخرجات، وكتابة التقرير النهائى باحتوائه على العديد من الأمثلة التى تساعد كل من الأكاديميين والممارسين للتحليل الإحصائى.
- يبدأ المرجع معك بالتدريج بالتحليلات البسيطة لمتغير واحد ولمتغيرين ثم التحليل المتعدد المتغيرات.
- يبدأ المرجع معك تدريجيا بالتحليلات البسيطة الأساسية من توزيع تكرارى ونسبى وجداول مزدوجة واختبارات فروض ثم ينتقل معك بأسلوب مبسط ويسير للأساليب الأكثر تعقيداً.
- يساعدك على تفسير نتائج التحليل الإحصائى لـ SPSS واختيار المهم والمناسب منها لوضعه بالتقرير النهائى.
- يحدد لك شروط تطبيق أساليب التحليل الإحصائى المتعدد المتغيرات وكيفية التأكد من توافرها.
- المرجع يستخدم بيانات فعلية فى التحليل والتعليق على المخرجات.
- يساعدك فى اختيار الأسلوب الإحصائى المناسب لتحليل البيانات .

المؤلف: عبدالحميد محمد العباسي (دكتوراه الفلسفة في الإحصاء الحيوي والسكاني، جامعة القاهرة، 1999) أستاذ الإحصاء رئيس قسم الإحصاء الحيوي والسكانى – معهد الدراسات والبحوث الإحصائية – جامعة القاهرة.

البريد الإلكتروني: . drabdo1@gmail.com / drabdo1@yahoo.com / dabdo1@hotmail.com

الاهتمامات البحثية في مجالات:

1- الإحصاء التطبيقي،
2- الإحصاء الحيوي والسكاني،
3- الإحصاء الجنائي،
4- الإحصاء السياحي والإحصاء التربوي والإحصاء الاجتماعي،
5- نمذجة وتقييم البيانات،
6- الاستشراف المستقبلي (التنبؤ)،
7- دعم اتخاذ القرارات،
8- النماذج السكانية،
9- تقييم الأداء المؤسسي،
10- مؤشرات سوق العمل،
11- قياس وتقييم الأداء والجودة،
12- التخطيط الأمني ودعم اتخاذ القرار،
13- النماذج الاقتصادية والاقتصاد القياسي،
14- بحوث العمليات والرقابة على الجودة،
15- التنقيب في البيانات Data Mining،
16- تقييم جودة البيانات الإحصائية وصدقها وثباتها.

الحزم الإحصائية: SPSS- AMOS- STATGRAPHICS -TSM -MINITAB – NSCC- EVIEWS- LISREL.