تطبيقات حاسوبية
في الأسواق السياحية

وقل رب زدني علما
جامعة دمشق
Damascus University

منشورات جامعة دمشق

كلية السياحة

الدكتورة ريم محمود رمضان
أستاذ مساعد
قسم الإدارة – كلية الاقتصاد

أسامة أنور محسن
قائم بالأعمال
المعهد العالي للدراسات والبحوث السكانية

جامعة دمشق

1438-1439هـ
2016-2017م

وقل رب زدني علما
جامعة دمشق
Damascus University

فهرس المحتويات

المحتويات	رقم الصفحة
المقدمة	9
الفصل الأول: مدخل إلى البحث العلمي	13
1. البحث ومفهومه	14
2. خصائص البحث العلمي	15
3. مراحل البحث العلمي	15
4. دور الإحصاء في البحث العلمي	19
5. أقسام علم الإحصاء	20
6. المتغيرات وأنواعها في البحث العلمي	21
7. القياس ومستوياته في البحث العلمي	25
أسئلة الفصل الأول	30
الفصل الثاني: جمع البيانات الإحصائية	31
1. مصادر جمع البيانات	32
2. أساليب جمع البيانات	34
3. أدوات جمع البيانات	35
أسئلة الفصل الثاني	43
الفصل الثالث: المجتمع والعينة	45
1. المجتمع الإحصائي	46
2. إطار المجتمع الإحصائي	48
3. العينة	49
4. أنواع العينات	50
5. تحديد حجم العينة	65
أسئلة وتمارين الفصل الثالث	73

الفصل الرابع: التعرّف على بيئة البرنامج الإحصائي SPSS	77
1. تشغيل البرنامج الإحصائي (SPSS)	78
2. شاشات البرنامج الإحصائي (SPSS)	78
3. ملفات البرنامج الإحصائي (SPSS)	81
4. القوائم الرئيسية في البرنامج الإحصائي (SPSS)	81
5. شريط الأدوات (SPSS Toolbar)	84
6. استيراد وتصدير البيانات	87
7. فتح ملفات بيانات	90
8. الخروج من البرنامج الإحصائي	91
أسئلة الفصل الرابع	92
الفصل الخامس: تعريف المتغيرات وإدخال البيانات في البرنامج الإحصائي SPSS	93
1. تعريف المتغيرات باستخدام البرنامج الإحصائي SPSS	94
2. حفظ البيانات	105
3. حفظ ملف المخرجات	106
4. عرض خصائص المتغيرات	107
5. قائمة التحرير (EDIT)	109
6. إدخال البيانات	111
7. القائمة View	114
أسئلة الفصل الخامس	118
الفصل السادس: الإجراءات التنظيمية في البرنامج الإحصائي SPSS	121
1. الإجراءات المتعلقة بتعريف البيانات	123
2. الإجراءات التنظيمية للبيانات	141
3. الإجراءات التنظيمية غير المباشرة للبيانات	157

أسئلة الفصل السادس	166
الفصل السابع: معالجة وتحويل البيانات باستخدام البرنامج الإحصائي SPSS	169
1. إجراء العمليات الحسابية Compute	171
2. عدّ القيم Count Value Within Cases	179
3. ترميز قيم المتغير Recode	182
4. ترتيب الحالات Rank Cases	197
5. استبدال الحالات المفقودة Replace Missing Value	201
أسئلة الفصل السابع	205
الفصل الثامن: وصف المتغيرات النوعية	207
1. الجدول التكراري لمتغير نوعي واحد	208
2. الجداول التقاطعية Cross Table	217
أسئلة الفصل الثامن	225
الفصل التاسع: الوصف الإحصائي للمتغيرات الكمية	229
1. مقاييس النزعة المركزية	230
2. مقاييس التشتت	232
3. معامل الالتواء	235
4. معامل التفلطح	236
5. التطبيق باستخدام البرنامج الإحصائي SPSS.	237
أسئلة وتمارين الفصل التاسع	255
الفصل العاشر: الإحصاء الاستدلالي	259
1. أساليب الاستدلال الإحصائي (الإحصاء الاستدلالي)	260
2. تعريفات أولية	262
3. أساليب التقدير الإحصائي	263

4. اختبار الفرضيات	266
5. اختبار التوزيع الطبيعي	272
6. اختبار الفرضية المتعلقة بوسط حسابي واحد	283
7. مقارنة وسطي عينتين مستقلتين	290
8. مقارنة وسطي عينتين مرتبطتين	299
9. مقارنة أكثر من متوسطي عينتين	307
أسئلة وتمارين الفصل العاشر	317
الفصل الحادي عشر: الأساليب الإحصائية لدراسة العلاقات	321
1. الارتباط	323
2. الارتباط الجزئي	340
3. الانحدار	345
المراجع	359

Damascus University

المقدمة

تلعب السياحة دوراً هاماً في اقتصاد الدول، وتحتل مكاناً مرموقاً واهتماماً بالغاً من طرف الحكومات التي أخذت في تطوير وتنمية القطاع السياحي، ويظهر الأثر الاقتصادي للسياحة في زيادة الإيرادات السياحية من النقد الأجنبي مما يساهم بتوفير قدر أكبر من العملات الأجنبية التي تسهم بدفع عملية التنمية نحو الأمام، ونتيجةً لذلك اهتمت الدول بتطوير القطاع السياحي من خلال تهيئة الجو المناسب لذلك من بنية تحتية ودراسات علمية متخصصة وأكاديمية، ولهذا يأتي كتاب تطبيقات حاسوبية في الأسواق السياحية لتزويد طلاب كلية السياحة المقبلين على التخرج بالمهارات والمبادئ الأساسية في التطبيقات الحاسوبية لإجراء البحوث التي تسهم بتطوير القطاع السياحي الذي يسهم بتنشيط الناتج القومي المحلي، لا سيما أنّ البحث العلمي آخذ بالتطور السريع على صعيد البحوث النظرية والتطبيقية، وجاء هذا الكتاب ليواكب التطور الحاصل في إعداد البحوث باستخدام برامج حاسوبية ذات مواصفات وخصائص علمية، ليعكس التطورات الحاصلة في البرنامج الاحصائي SPSS، ويقع الكتاب في أحد عشر فصل ليبين الجوانب النظرية والتحليلية على النحو التالي:

يأتي **الفصل الأول** كمدخل إلى البحث العلمي ليعرّف الطالب على البحث العلمي وخصائصه ومراحله، وكذلك يعرّف الطالب على أهمية الإحصاء في البحث العلمي، ويناقش التصنيفات المختلفة للمتغيرات ومستويات القياس في البحث العلمي.

أما **الفصل الثاني** فيهدف إلى التعرّف على مصادر وأساليب وأدوات وخطوات جمع البيانات.

ويناقش **الفصل الثالث** المفاهيم الأساسية المستخدمة في اختيار العينة وأنواع العينات الاحتمالية (العشوائية) وغير الاحتمالية (غير العشوائية) وطرق تحديد حجم العينة.

أما **الفصل الرابع** فيركز على بيئة البرنامج الإحصائي SPSS، وكيفية تشغيله، والتعرف على شاشات وملفات وشريط أدوات البرنامج الإحصائي (SPSS) وقوائمه الرئيسية.

والفصل الخامس يناقش تعريف المتغيرات ضمن البرنامج الإحصائي SPSS، وحفظ البيانات والمتغيرات في البرنامج الإحصائي SPSS، وعرض خصائص المتغيرات.

أما **الفصل الساس** فيركز على الإجراءات التنظيمية في البرنامج الإحصائي SPSS التي تتعلق بالبيانات.

ويهدف **الفصل السابع** للتعرف على كيفية معالجة وتحويل البيانات باستخدام البرنامج الإحصائي SPSS.

ويركز **الفصل الثامن** على وصف المتغيرات النوعية باستخدام أوامر التكرار والوصف والجداول التقاطعية والأشكال البيانية.

وكذلك **الفصل التاسع** يركز على الوصف الإحصائي للمتغيرات الكمية باستخدام مقاييس النزعة المركزية ومقاييس التشتت، باستخدام التكرار والوصف والاستكشاف.

ويأتي **الفصل العاشر** الإحصاء الاستدلالي للتعرف طرق اختبار التوزيع الطبيعي، وطرق اختبار الفرضيات باستخدام البرنامج الإحصائي SPSS.

وأخيراً فقد جاء **الفصل الحادي عشر** لتوضيح معاملات الارتباط، ومعامل الارتباط الجزئي، ومعامل الانحدار وكيفية التطبيق باستخدام البرنامج الإحصائي SPSS.

ويحتوي هذا الكتاب على المفاهيم الأساسية والأمثلة التطبيقية التي تتعلق بالتطبيقات الحاسوبية في الأسواق السياحية التي سيتم عرضها لاحقاً، أما المعادلات

المقدمة

الإحصائية والقوانين الرياضية ستعرض في نطاق ضيق جداً حيث أن البرامج الإحصائية تتكفل بهذه المهمة.

وأخيراً حاولنا في هذا الكتاب تبسيط المادة للطالب حتى تكون مستساغة لديه ليتقبلها بتفهم ويستخدمها بإتقان مما يسمح له بمتابعة البحث العلمي الأكاديمي والموضوعي.

كما بذلنا كل الجهد لتقديم كل ما هو ضروري في التطبيقات الحاسوبية لطلبة كلية السياحة ولجميع الباحثين والمهتمين في هذا المجال كون البحوث التسويقية أصبحت في هذه الأيام بمثابة مرآة تعكس الواقع والحقائق، ويعالج الكتاب كافة الجوانب التطبيقية والإحصائية التي تهم طلبة كلية السياحة والفئات الأخرى التي لها علاقة وطيدة بالأبحاث السياحية ليعتبر مرجعاً كاملاً لطلبة كلية السياحة من جهة، وللمختصين والمهتمين والعاملين في مجال البحوث السياحية من جهة أخرى.

وأخيراً نقدم شكرنا وتقديرنا لكل من ساهم بإنجاز هذا الكتاب، والله ولي التوفيق

دمشق: 2016/12/1

المؤلفان

الدكتورة: ريم محمود رمضان **أسامة أنور محسن**

وقل رب زدني علما
جامعة دمشق
Damascus University

الفصل الأول CHAPTER 1

مدخل إلى البحث العلمي

أهداف الفصل الأول:

1. التعرف على البحث العلمي وخصائصه.
2. التعرف على مراحل البحث العلمي.
3. التعرف على أهمية الإحصاء في البحث العلمي.
4. التعرف على التصنيفات المختلفة للمتغيرات في البحث العلمي.
5. التعرف على مستويات القياس في البحث العلمي.
6. التعرف على العلاقة بين مستويات القياس والأساليب الإحصائية الملائمة لها.

لقد أصبح الحاسوب أداة لا غنى عنها في البحث العلمي، فبرامج الحاسوب الإحصائية التي يُعتنى بتصميمها وإعدادها تيسر عملية التبادل بين أفكار الباحث والنتائج المستمدة من بيانات دراسته، والعكس بالعكس، مما يمكنه من الفحص المستنير لهذه النتائج وإعادة النظر في أفكاره بيسر وسهولة، كما أن هذه البرامج أسهمت إسهاماً كبيراً في توسيع نطاق الأساليب الإحصائية التي تستخدم في تحليل البيانات بمختلف أنواعها، متعددة المتغيرات والعينات خاصة.

فأغلب البحوث العلمية تتطلب تحليل البيانات المستمدة منها أساليب إحصائية متقدمة تتضمن عمليات رياضية معقدة ومتكررة يصعب تنفيذها بسرعة ودقة دون استخدام برامج إحصائية حاسوبية، فتوفر مثل هذه البرامج أسهم في انتشار استخدام متغيرات متعددة في البحوث التي كانت تقتصر عادة على متغير واحد أو متغيرين.

ولعل أهم البرامج الإحصائية الشائعة الاستخدام هو البرنامج الإحصائي SPSS فهذا البرنامج يتجدد باستمرار ويتضمن التطورات المعاصرة في الإحصاء وتقنيات الحاسوب والبرمجيات، ويشمل على عدد كبير من الأساليب الإحصائية الاستدلالية البارامترية واللابارامترية.

1. البحث ومفهومه:

يمثل البحث طريقة منظمة أو فحص استفساري منظم لاكتشاف حقائق جديدة، أو التثبت والتحقق من حقائق قديمة والعلاقات التي تربط بينها أو القوانين التي تحكمها وبما يسهم في نمو المعرفة الإنسانية. إنه شيء يباشره الأفراد بنظام لإيجاد مخرجات منطقية لزيادة المعرفة. إنه طريقة منظمة لإيجاد تفسيرات لظواهر اجتماعية أو توضيح حقائق لم تفهم بصورة واقعية.

ونلاحظ تعبيرين هامين من هذا التعريف هما "البحث العلمي"، و "لإيجاد مخرجات منطقية" وهذا يدلل على أن البحث يركز على علاقات منطقية وليس على

معتقدات، ولذلك لابد أن يتضمن البحث عدداً من الأهداف قد تتضمن وصفاً، إيضاحاً، فهماً وتحليلاً، ولا بد أن يمتلك البحث هدفاً واضحاً أو مجموعة أشياء تمثل الإجابة على سؤال أو عدة أسئلة، وكذلك إيضاح للطرق المستخدمة في جمع البيانات، وأن يناقش الباحث منطقية النتائج التي حصل عليها بحيث تكون نتائج ذات معنى، مع شرح المحددات التي تواجه الباحث في البحث، أما البحث العلمي فهو المحاولة الدقيقة للتوصل إلى حلول للمشكلات.

فالبحث وسيلة وليس غاية لأن الباحث يحاول من خلال بحثه إشباع حاجته من المعرفة وتوسيعها، أو دراسة ظاهرة معينة أو مشكلة ما؛ للتعرف على العوامل التي أدت إلى وقوعها ثم الخروج بنتيجة أو حل وعلاج للمشكلة.

2. خصائص البحث العلمي:

يتصف البحث العلمي بمجموعة مترابطة من الخصائص التي لا بد من توافرها لتتحقق أهدافه، وهي كما يلي:

- الموضوعية.
- الاختبارية والدقة.
- إمكانية تكرار النتائج.
- التبسيط والاختصار.
- أن يكون للبحث العلمي غاية وأهداف.
- استخدام نتائج البحث لاحقاً في التنبؤ بحالات ومواقف مشابهة.

3. مراحل البحث العلمي:

يرى بعض الباحثين أن مراحل البحث العلمي تمر بعدة خطوات أهمها تحديد المشكلة، ووضع الفروض، وتحديد أهداف البحث وأهميته، ووضع منهجية للبحث، وسنناقش هذه المراحل بإيجاز كما يلي:

1.3. مشكلة البحث:

1.1.3. الشعور والإحساس بالمشكلة:

هو التساؤل الذي يدور في ذهن الباحث حول موقف معين يكتنفه الغموض ويحتاج إلى تفسير وتوضيح، فهي خطة ذهنية تعد نقطة البداية لتحديد المشكلة الفعلية في صورة تساؤلات واستفسارات يمكن ترجمتها فيما بعد إلى أسئلة نحاول الوصول إلى إجابات عليها، وعليه فإن الشعور بمشكلة ما تدفع الباحث إلى البحث عن حلول لتلك المشكلة مثل الشعور بنقص عدد السياح في أحد المنتجعات السياحية مما يدفع إلى البحث عن الأسباب التي أدت إلى ذلك النقص لمحاولة علاجه، وما يساعد الباحث في زيادة الإحساس والشعور بالمشكلة والوصول إلى منابعها:

- الخبرة الشخصية واهتمامات الباحث بالموضوع.
- القراءة الناقدة التحليلية للإطار النظري المبحوث.
- الدراسات السابقة التي أجريت في مجال التخصص من قبل باحثين آخرين، سواء في نفس البيئة أو بيئة مشابهة.

2.1.3. تحديد مشكلة البحث:

تبدأ عملية البحث بتحديد مشكلة البحث، والمشكلة هي تساؤل أو عبارة تستفسر عن نوع العلاقة بين متغيرين أو أكثر، فالمشكلة تبدأ بفكرة عامة عما ينوي الباحث دراسته، وبعدها يقوم بتهذيب الهدف العام وتحسينه ليصبح جملة بليغة ودقيقة بشكل أكثر تحديداً، وهناك طريقتان لصياغة أية مشكلة، فمن الممكن صياغة المشكلة بصيغة إخبارية، أو صيغة استفهامية، ويمكن صياغة المشكلة بعبارة واحدة أو عدة عبارات، ويتوقف ذلك على طبيعة المشكلة ودرجة تعقيدها، ولابد أن تتحدد في صياغة المشكلة العلاقات بين المتغيرات والمجتمع الخاص الذي تهتم به الدراسة، ويجب أن تكون المشكلة قابلة للبحث بعد تعريف متغيراتها.

ومن الاعتبارات الرئيسية التي تؤخذ عند تحديد مشكلة البحث:

- أن تكون المشكلة قابلة للبحث وليست خيالية أو مستحيلة، ويمكن أن ينبثق عنها أسئلة وفرضيات يمكن إثباتها أو نفيها.
- أن تتمتع المشكلة بالأصالة بحيث تؤدي دراستها إلى إضافات علمية جديدة ولا تكون تكراراً لأبحاث سابقة، وإن تناولت موضوعاً سبق تناوله من باحثين سابقين فلا بد أن تكون معالجته من زوايا جديدة تضيف شيئاً ما في توضيح المشكلة وطرق معالجتها.
- أن تكون المشكلة ضمن اهتمامات الباحث العلمية، لأن الباحث الذي يخوض في معالجة مشاكل ليست من اختصاصه واهتماماته قد لا تكون نتائجه دقيقة ويمكن البناء عليها.
- أن تكون المشكلة في حدود إمكانيات الباحث المادية فدراسة أي مشكلة يترتب عليها تكاليف عديدة في مراحل البحث اللاحقة سواء في تحديد العينة ومتابعتها والتنقل بين المناطق الجغرافية التي تقطنها العينة لجمع المعلومات.
- أن تكون المشكلة ملائمة للبيئة التي يجري البحث بها سواء من الناحية السياسية أو الاجتماعية أو الثقافية أو الاقتصادية بحيث لا تتضارب مع منظومة القيم السائدة في المجتمع.

2.3. وضع الفروض:

بعد أن يقوم الباحث بتحديد المشكلة، يبدأ بالبحث عن حل للمشكلة أو تفسير للظاهرة التي ملاحظتها لذلك يستخدم الباحث الفرض لكي يحل المشكلة.

ويعرف الفرض بأنه تخمين أو استنتاج ذكي يصوغه ويتبناه الباحث مؤقتاً لشرح بعض ما يلاحظه من الحقائق والظواهر... وليكون هذا الفرض كمرشد له في البحث والدراسة التي يقوم بها، وتتبع الفروض من مصادر متعددة، فيمكن استنتاجها من

النظريات المعروفة في مجال علمي معين، أو من الثقافة المتخصصة للباحث في مجاله الموضوعي أو المدرسة الفكرية التي ينتمي إليها.

3.3. تحديد أهداف البحث:

يعد تحديد الأهداف المتوخى الوصول إليها من أساسيات البحث العلمي، لذا لا بد من أن يتسم الهدف (أو الأهداف) بالشفافية والوضوح وعلى درجة معقولة من التفصيل لنكون على علم كاف بالمعطيات اللازم تغطيتها، فالهدف من البحث إضافة علمية، أو تشخيصي لظاهرة ما للتعرف على عوامل معينة في تلك الظاهرة، أو البحث في علاقة السبب والأثر لمشكلة ما، وتحديد الباحث لأهداف دراسته بدقة سيساعده في التحديد الدقيق للمجتمع والعينة التي سيتعامل معها، وسيكون مقنعاً للقارئ في الأخذ بالنتائج التي توصل إليها. فمثلاً إذا كان هدف البحث هو دراسة "مستوى الخدمات السياحية"، عندها يجب أن نوضح وبالتفصيل إذا كان الأمر سيقتصر على الفنادق والمنتجعات فقط أم أن ذلك يتضمن شبكات النقل البري والجوي والبحري ...الخ، وعادة ما يراعى في تصميم الأهداف طبيعة العلاقات الإحصائية المستهدفة بين المتغيرات ذاتها أو بين المتغيرات والمشاهدات، ليتم في ضوئها تصميم الاستبانة التي سيتم التطرق إليها لاحقاً.

4.3. تحديد أهمية البحث:

لابد أن يقوم الباحث من توضيح أهمية بحثه الذي ينوي الشروع به وما يمكن أن يؤدي ذلك من إضافات علمية تنفع الباحثين اللاحقين، وما يمكن أن يصل له البحث من نتائج عملية على أرض الواقع تفيد في حال المشكلة المعروضة وتعميمها على المشاكل المشابهة، لأن ذلك سيعطي قوة وقناعة بالبحث.

5.3. الإطار النظري للدراسة:

يقوم الباحث بعد تحديد مشكلة الدراسة بالمراجعة النظرية لما تم تناوله ممن سبقوه فيما يتعلق بالموضوع المبحوث ويشمل الإطار النظري للدراسة:

- مراجعة الأدب السابق سواء من الكتب والوثائق والانترنيت والتي تناولت الموضوع لإثراء الموضوع المبحوث.
- مراجعة الدراسات السابقة المتعلقة بموضوع البحث.

إن الاهتمام بالإطار النظري للدراسة سيوفر للباحث نظرة سريعة للتطورات البحثية عن الموضوع، ويساعد الباحث في تحديد معنى المشكلة وأهميتها وربط نتائج الدراسة مع المعرفة السابقة وكذلك تقديم ملخصات للدراسات السابقة التي تناولت الموضوع ومدى الاتفاق أو الاختلاف مع الدراسات السابقة؟ واقترح مزيد من الدراسات مستقبلاً وتجنب التكرار في الدراسات المشابهة.

6.3. تحديد منهج البحث:

هو الطريقة التي سيسلكها الباحث في الإجابة عن الأسئلة، إنها خطة تبين وتحدد طرق وإجراءات جمع البيانات وتحليلها، حيث يقوم الباحث من خلال منهج البحث بتحديد تصميم البحث ويختلف تصميم البحث باختلاف الهدف منه فقد يكون استكشاف عوامل معينة لظاهرة ما، أو توصيفها، أو إيجاد العلاقة أو السبب والأثر بين مجموعة من العوامل.

4. دور الإحصاء في البحث العلمي:

علم الإحصاء دوراً أساسياً في مجال البحث العلمي كأداة في تحليل واستخراج النتائج لمختلف البحوث والدراسات، وحيث أن الإحصاء علم له قواعده وقوانينه، فضلاً عن كونه طريقة علمية تستخدم القيم والأرقام في تحليل الصفات والظواهر المراد بحثها وصولاً إلى نتائج موثوقة يستدل منها الباحثون في عمليات التحليل

والتفسير لتلك الظواهر، فتعريف علم الإحصاء يركز إلى ما يهدف إليه الإحصاء، ومن هذه التعاريف:

1) علم يبحث في جمع البيانات وتنظيمها وعرضها في جداول وتحليلها واستنتاج النتائج، ومن ثم اتخاذ القرارات المناسبة.
2) هو العلم الذي يبحث في الطرق والأساليب المختلفة لجمع البيانات وعرضها وتبويبها وتحليلها، ثم استخدام هذه البيانات في التنبؤ أو التحقق من بعض الظواهر، وبالتالي قبول فرضيات الأبحاث أو رفضها أو الإجابة عن أسئلتها الأساسية.
3) هو فرع من العلوم الذي يختص بالطرق العلمية لجمع البيانات وتنظيمها وتلخيصها وعرضها وتحليلها وذلك للوصول إلى نتائج مقبولة وقرارات سليمة على ضوء هذا التحليل.
4) هو علم تعميم الجزء على الكل.

ومن هذه التعاريف نجد أن الإحصاء هو طريقة منظمة تسير في خطواتٍ متسلسلة بدءاً من جمع البيانات عن الظاهرة ثم وصف هذه الظاهرة، وتحليل البيانات المتجمعة عن الظاهرة وفق قواعد وقوانين إحصائية خاصة، واتخاذ القرارات المناسبة بناءً على البيانات، وغالباً ما تكون هذه القرارات على شكل تعميمات، أو تقديرات وذلك من أجل التنبؤ أو لرفض الفرضيات الإحصائية أو عدم رفضها، فالإحصاء هو الوسيلة التي يُستَدِلُ من خلالها الباحثون على الكيفية التي ينجز فيها البحث العلمي بأفضل الطرق وأيسرها، وبأقل كلفة وجهد، مع اختصار المدة المعينة لذلك الانجاز، ومثل هذه الصفات الحميدة للإحصاء جعلت عملية الإقبال عليه من قبل الباحثين واستخدامه في تزايد مستمر.

5. أقسام علم الإحصاء:

بناءً على تعريف علم الإحصاء يمكننا تقسيم علم الإحصاء إلى ما يلي:

1) **الإحصاء الوصفي** (Descriptive Statistics): وهو نوع من أنواع الإحصاء الذي يهتم بأساليب جمع البيانات وتبويبها وعرضها في جداول ورسوم وأشكال بيانية، كما يهتم بإظهار الخصائص الأساسية للظاهرة المدروسة وذلك بحساب مؤشرات مختلفة كالتكرارات، والنسب المئوية، مقاييس النزعة المركزية، مقاييس التشتت وغيرهما من المقاييس.

2) **الإحصاء الاستدلالي** (Inferential Statistics): وهو نوع من أنواع الإحصاء الذي يهتم بالطرق والأساليب التي تكشف عن مؤشرات المجتمع من خلال حسابها في العينة التي سحبت منه، وهناك أسلوبان للإحصاء الاستدلالي الأول **التقدير** (Estimate) وهو تقدير مؤشرات المجتمع(المتوسط الحسابي – الانحراف المعياري) بناءً على مؤشرات العينة ويتم ذلك باستخدام **مجالات الثقة** (Confidence Interval) وهو مجال حول أحد تلك المؤشرات، والثاني **اختبار الفرضيات** (Tests of Hypotheses) وفيه يتم استخدام بيانات العينة للوصول إلى قرار علمي سليم بخصوص الفروض المحددة حول مؤشرات المجتمع.

3) **التنبؤ** (Forecasting): وفيه يتم استخدام نتائج الاستدلال الإحصائي، والتي تدلنا على سلوك الظاهرة في الماضي في معرفة ما يمكن أن يحدث لها في الحاضر والمستقبل. هناك العديد من الأساليب الإحصائية المعروفة التي تستخدم في التنبؤ، ومن أبسطها **أسلوب الانحدار الخطي**، وهي معادلة رياضية يتم تقدير معاملاتها باستخدام بيانات العينة، ثم بعد ذلك استخدام المعادلة المقدرة في التنبؤ بما يمكن أن يحدث للظاهرة في المستقبل.

6. المتغيرات وأنواعها في البحث العلمي:

إنّ مجموعة المشاهدات أو الملاحظات أو القياسات الأولية المأخوذة من أفراد دراسة ما أو اختبار معين وتتعلق بصفة أو خاصة معينة تدعى **بالبيانات** (Data) وهذه البيانات إما أن تكون رقمية (كمية) أو بيانات غير رقمية (نوعية) والبيانات تمثل

خصائص أفراد الدراسة وهي لا تأخذ نفس القيمة بالنسبة لجميع الأفراد وهذه الخصائص تسمى **بالمتغير** (Variable)، فمثلاً مجموعة من طلبة الجامعة قد يختلفون في الجنس أو الكلية أو السنة الدراسية أو مستوى الذكاء أو درجة التحصيل، وبالتالي يمكننا تعريف **المتغير** بأنه التعبير عن الظاهرة من خلال مقدار له خصائص رقمية (كمية) أو صفات غير رقمية (وصفية) تتغير من عنصر (مفردة أو مشاهدة) إلى آخر من العناصر محل الظاهرة المدروسة.

يمكن تصنيف المتغيرات بطرق متعددة وهذه التصنيفات لها فوائدها في البحوث المختلفة وخاصة عند جمع البيانات.

1.6. التصنيف العام:

وتنقسم المتغيرات إلى:

1.1.6. المتغيرات العشوائية (Random Variable): وهي عبارة عن البيانات الكمية المعبرة عن نتائج التجارب العشوائية والتي تنتج من خلال نظام احتمالي معين وتختلف هذه القيم من ظاهرة إلى أخرى.

2.1.6. المتغيرات الإحصائية (Statistical Variable): وهي عبارة عن البيانات التي تمثل قيماً فعلياً لظاهرة ما، وتنقسم هذه البيانات إلى نوعين، فمنها الكمية والأخرى النوعية.

ومن جهة ثانية إذا كانت هذه الخصائص أو السمات نفسها بالنسبة إلى كل فرد من أفراد المجموعة فإنّ هذه السمة تدعى بالثابت (Constant)، وكمثال على ذلك درجات طلاب كلية السياحة في مقرر الإحصاء فإنه من الطبيعي أن يأخذ الطلبة درجات متفاوتة في مقرر الإحصاء، هنا تسمى درجة الطالب المتغير، أما الثابت هو أنهم من طلاب كلية السياحة.

2.6. تصنيف المتغيرات حسب مستوى القياس [1]:

1.2.6. المتغيرات الكمية (Quantative Variable): وهي المتغيرات الناتجة عن عملية قياس أو عد والتي يعبّر عنها بمقادير أو قيم معينة مثل الدخل، مستوى الذكاء، درجة التحصيل الدراسي والوزن والطول وغير ذلك، **وتقسم إلى نوعين:**

أ- متغيرات مستمرة (Variable Continuous): وتكون بياناته ناتجة عن عملية استخدام جهاز أو أداة للقياس مثل متغير طول، الوزن، درجة الحرارة أو درجة التحصيل الدراسي أو مستوى الذكاء...

ب- متغيرات منقطعة (Discrete Variable): وتكون بياناته ناتجة عن عملية عد أو تعداد وفيها يأخذ المتغير قيماً صحيحة ولا يوجد فيها أرقام كسرية مثل عدد أفراد الأسرة، أو العدد السنوي لحالات الولادة أو الطلاق أو الزواج فجميع الحالات تأخذ أعداداً صحيحة ولا يمكن التعبير عنها بشكل أعداد مستمرة.

2.2.6. المتغيرات التصنيفية أو النوعية (Variable Qualitative): إن المتغيرات التصنيفية تستخدم لتصنيف قيم المتغير في فئات متعددة وتقسم لقسمين:

أ- متغيرات اسمية (Nominal Variable): عبارة عن اسم أو وصف لأي عنصر أو مفردة في المجتمع مثل الجنس يصنف إلى ذكر أو أنثى.

ب- متغيرات ترتيبية (Ordinal Variable): عبارة عن اسم أو وصف يعبّر عن التفضيل أو الترتيب لأي عنصر في المجتمع.

---

(1) William Mendenhall; Robert J.Beaver; Barbara M.Beaver. **Introduction To Probability and Statistics**, Duxbury Tenth Edition, 1999, P10.

3.6. تصنيف المتغيرات في البحوث الارتباطية أو السببية:

1.3.6. المتغير المستقل (Independent Variable): وهو المتغير الذي يتحكم بالتجربة، كما يعتبر في بعض الدراسات متغيراً تصنيفياً يتم من خلاله تصنيف المشاهدات (الحالات) المدروسة إلى مستويات، فعلى سبيل المثال إذا كان الباحث مهتماً بدراسة تأثير درجة الفندق (نجمتان – ثلاث نجوم) في مدينة ما على عدد السياح فإن درجة الفندق هي المتغير المستقل ونجمتان وثلاث نجوم تعبر عن مستويات المتغير المستقل.

2.3.6. المتغير التابع (Dependent Variable): هو المتغير المقابل للمتغير المستقل وهو يتأثر بتغيرات المتغير المستقل، مثال على ذلك: دراسة العلاقة بين الحركة السياحية وإيرادات الدولة، فالحركة السياحية هي المتغير المستقل وإيرادات الدولة هي المتغير التابع كونها تتأثر بالحركة السياحية.

وفي الأبحاث والدراسات يمكن أن يكون هناك أكثر من متغير مستقل واحد ومتغير تابع واحد فقط، كما تدعى المتغيرات المستقلة بالعوامل (Factors).

والطريقة الأفضل لتحديد أي المتغيرين المستقل وأيهما المتغير التابع هو طرح السؤال التالي:

من هو المتغير المسبب لحدوث الظاهرة، أو من هو المتغير المثير للظاهرة، أو من هو المتغير المتنبئ، أو من هو المتغير الأسبق زمنياً؟ والجواب هو المتغير المستقل.

3.3.6. المتغير المعدل (Moderator Variable): وهو نوع خاص من المتغيرات المستقلة إذ يعدّ متغيراً مستقلاً ثانوياً يتم اختياره من قبل الباحث لمعرفة أثره على العلاقة بين المتغير المستقل والمتغير التابع ويختاره عادة الباحث ويقيسه لمعرفة فيما إذا كان هذا المتغير يعدل العلاقة بين المتغير المستقل والمتغير التابع وهذا النوع من المتغيرات يمكن التعرف عليه عند دراسة التصاميم العاملية، فعلى سبيل المثال إذا كان الباحث

مهتم بدراسة العلاقة بين عدد السياح وإيرادات الدولة وهناك متغير آخر يؤثر على هذه العلاقة مثل عدد المعالم السياحية، فإن عدد المعالم السياحية يعدّ متغيراً معدلاً.

4.3.6. المتغيرات الضابطة (Control Variable): إن المتغيرات التي تؤثر على المتغير التابع من الصعب دراستها في الوقت نفسه، فلذلك قد يلجأ الباحث في بعض الأحيان إلى تحييد أثرها أو ضبطها حتى يضمن أن هذه المتغيرات ليس لها تأثير على المتغير التابع وهذه المتغيرات تدعى بالمتغيرات الضابطة والضبط يكون إما عن طريق العزل أو خلق التكافؤ بين المجموعات من خلال التعيين العشوائي للأفراد إلى المجموعات. فعندما نقوم بالمقارنة بين مجموعتين من الأفراد من صف معين فإن الصف يعتبر متغيراً ضابطاً لأننا لا نريد أن ندرس أثره، وبالتالي تم سحب العينة من نفس المستوى الصفي.

7. القياس ومستوياته في البحث العلمي:

1.7. تعريف القياس:

لكلمة قياس استعمالاتها الواسعة في العلوم كافة وفي مجالات الحياة الإنسانية المختلفة، فيقال إن درجة الحرارة اليوم (25) درجة مئوية، ويقال طول زيد 169 سم ....وهكذا ولعل هذا الانتشار الواسع لكلمة قياس في العلوم المختلفة والحياة الإنسانية كان قد أغنى مضامينها وشعب دلالاتها وأكسبها معان عديدة. حيث قام العالم لورج بمحاولة حصر المعاني الكثيرة التي ينطوي عليها مصطلح القياس باللغة الإنكليزية فكان هناك ما يقارب من أربعين معنى.

هناك العديد من التعريفات لعملية القياس نذكر منها:

- جليفورد (Gilford, 1950): وصف للبيانات أو المعطيات بالأرقام.
- ستيفنز (Stevens, 1951): العملية التي يتم بوساطتها التعبير عن الأشياء والحوادث برموز وأعداد استناداً إلى قواعد.

- ❖ كامبل (Campbell, 1952): تمثيل للخصائص أو السمات المقيسة بالأرقام.
- ❖ مهرنز (Mehrens, 1969): العملية التي تمكن الإحصائي من الحصول على معلومات كمية عن ظاهرة ما.
- ❖ ننالي (Nunnally, 1970): القياس يتكون من قواعد استخدام الأعداد بحيث تدل على الأشياء بطريقة تشير إلى كميات من الخاصية.

ومن الواضح أن هذه المجموعة من التعريفات تلتقي على ضرورة توافر قواعد محددة تخضع لها عملية القياس ويتم من خلالها تمثيل الخصائص أو السمات المقاسة بأرقام.

فقد تنحصر مهمة الباحث في القياس مجرد تصنيف أو ترتيب الأفراد تنازلياً أو تصاعدياً وفقاً لدرجة امتلاكهم للسمة المقاسة أو المتغير وقد يكون الترتيب معتمداً على وحدات متساوية تدل على مقادير متساوية من الخاصية أو السمة التي يقيسها وقد يستخدم شكلاً أرقى من الأشكال السابقة جميعاً فيلجأ إلى الصفر المطلق ليشير به إلى انعدام السمة المقيسة.

بناءً على ما سبق إن المقاييس يمكن تصنيفها في أربعة مستويات بمجموعها ذات صفة تراكمية بمعنى أن مستويات المقاييس متضمنة فيما بينها، وهي تختلف بإمكانية معالجتها رياضياً.

2.7. مستويات القياس: وهذه المستويات هي:

1.2.7. المقياس الاسمي (Nominal Scale):

وهو أدنى مستويات القياس وأبسطها، وهو مقياس تصنف فيه حالات المتغير إلى مجموعات أو فئات استناداً إلى سمة معينة تعتمد أساساً للتصنيف، فأفراد الجامعة يصنفون حسب الجنس إلى ذكور أو إناث، أو حسب النتيجة إلى ناجحين أو راسبين....، ويناسب هذا المقياس المتغيرات النوعية (الكيفية) ويمكن الاستعاضة عن

الأسماء بأرقام فالأفراد يمكن أن يصنفوا إلى ذكور وإناث بحيث يعطى الرقم (1) للإناث والرقم (2) للذكور ولكن هذه الأرقام لا تحمل أي معنى كمي وإنما تشير إلى فئات أو مجموعات معينة في الخاصية، ولا يمكن في هذا المستوى استخدام العمليات الحسابية باستثناء الجمع على مستوى الفئة الواحدة، فالأرقام هنا مجرد رموز للفئات أو المجموعات ويستحيل استخراج أي مقياس من المقاييس الإحصائية سوى المنوال.

2.2.7. المقياس الترتيبي (Ordinal Scale):

وجدنا أن المقياس السابق يشير إلى التصنيف أي الانتماء لفئة استناداً إلى خاصة معينة اتخذت أساساً للتصنيف، وبمتابعة التصنيف بين أفراد المتغير حسب خاصة معينة تصاعدياً أو تنازلياً فتصبح عملية التصنيف تشمل الترتيب وهو ما يستخدم لأجله المقياس الرتبي أو الترتيبي، حيث يقوم المقياس الترتيبي أو الرتبي بترتيب الحالات وفق تدرج معين يشير إلى الاختلاف بينهم دون تحديد لمقدار هذا الاختلاف حيث إن هذا الترتيب لا يشير إلى انتظام في الفروق، والأرقام المستخدمة هنا ليس لها مدلول كمي، فالترتيب نسبي وليس مطلقاً فقد يؤثر انضمام فرد إلى فئة أو خروج فرد منها إلى تعديل الموقع النسبي للفئة وبالتالي لا يمكننا هنا القيام بالعمليات الحسابية الأربع ولا نستطيع حساب المتوسط أو الانحراف المعياري ولكن يمكننا استخراج الوسيط كما يمكننا تحويل الرتب إلى رتب مئينيه واستخدام معامل ارتباط الرتب.

3.2.7. مقياس المسافات أو الفترات (Interval Scale):

ويدعى أيضاً المقياس الفئوي، وجدنا أن المقياس الاسمي يهتم بالتصنيف والمقياس الرتبي يهتم بالتصنيف والترتيب وفي كلا المقياسين الأعداد لا تحمل معناً كمياً أما في المقياس الفئوي فهو يهتم بالتصنيف والترتيب كما أن الأعداد تحمل معناً كمياً فالمقياس الفئوي أرقى وأدق من المقياسين السابقين كما أنه يعتمد على وحدات قياس متساوية للتعبير عن الفرق في السمة أو الخاصة المقاسة مما يمكننا من تقدير المسافة

التي تفصل بين فردين أو شيئين بدقة وسهولة، فالفرق بين الدرجتين 20 و40 هو نفس الفرق بين الدرجتين 40 و60 لذلك يسمح لنا المقياس الفئوي بتنظيم الفروق والتعبير عنها بوحدات متساوية وبالتالي نكون هنا قادرين على إجراء عمليتي الجمع والطرح ضمن المستوى الفئوي، ولكن هذا المقياس لا يتضمن صفراً حقيقياً فالصفر هنا نسبي (افتراضي) وليس مطلقاً (رياضياً) فعندما نقول إن درجة الحرارة صفراً فإن ذلك لا يعني انعدام الحرارة، وبما أن هذا الصفر ليس رياضياً في هذا المستوى فتكون عملية القسمة أو إيجاد النسبة غير ممكنة. إنّ هذا المستوى من القياس يستخدم كثيراً في القياس النفسي والتربوي فنحن لا نقيس الذكاء أو الميول أو الاتجاهات وإنما نقيس الفرق الحقيقي بين ذكاء شخصين طبق عليهما نفس مقياس الذكاء، ومن الأمثلة على هذا النوع من المقاييس أيضاً التحصيل، والقلق، والاتجاهات وغير ذلك.

4.2.7. مقياس النسبة (Ratio Scale):

يعد هذا المقياس أرقى وأدق المقاييس السابقة لأنه يحمل جميع صفات المقاييس السابقة بالإضافة إلى وجود صفر حقيقي يشير إلى انعدام السمة أو الخاصة كما يمكننا هذا المستوى استخدام جميع العمليات الحسابية من جمع وطرح وضرب وقسمة ومقارنة فالفرد الذي وزنه (70 كغم) هو ضعف الفرد الذي وزنه (35 كغم) ومن الأمثلة على هذا النوع من المقاييس الطول والوزن والعمر وغير ذلك. وتستخدم مقاييس النسبة على نطاق واسع في العلوم الدقيقة ويندر استعمالها في التربية وعلم النفس نظراً لأن المقاييس النفسية والتربوية تفتقر إلى وجود نقطة صفر حقيقية تدل على غياب أو انعدام السمة.

إنّ معرفة نوع المقياس المستخدم يحدد الاختبار الإحصائي الذي سنستخدمه في التحليل الإحصائي فعلى سبيل المثال إذا كان المتغير الذي نتعامل معه اسمي فإننا نستطيع حساب المنوال، أما في حالة تصنيف المتغير ضمن مقياس رتبي فإننا نستطيع أن نحسب المنوال والوسيط، أما في حالة تصنيف المتغير ضمن مقياس فترات أو مقياس نسبة فإننا نستطيع أن نحسب المنوال والوسيط والمتوسط الحسابي، إنّ هذا الوضع يكون

في حالة الإحصاء الوصفي وكذلك الحال عندما نستخدم أساليب الاستدلال الإحصائي، حتى الإحصاء المتقدم فلا نستطيع حساب تحليل التباين إلا إذا كان المتغير التابع يصنف على الأقل ضمن مقياس المسافات أو الفترات وهكذا.

3.7. علاقة القياس بالإحصاء:

يختلف القياس عن الإحصاء فهما مفهومان مختلفان، ولكل منهما معنى وإجراءات مختلفة، فيقصد بالقياس تعيين أرقام أو مستويات مختلفة للصفة المقاسة باختلاف الأفراد، أما الإحصاء فهو يستخدم هذه الأرقام أو المستويات ويتعامل معها بأساليب معينة تناسب مشكلة الدراسة أو تساؤلاتها.

وبذلك يكون القياس هو عملية التوصل إلى الأرقام التي نستخدمها في التحليلات الإحصائية، فيقوم الباحث بقياس المتغيرات مثل الذكاء والتحصيل والمستوى الاجتماعي وغيرها، ثم يستخدم الأرقام التي يحصل عليها في إجراء التحليلات الإحصائية لوصف الظاهرة أو تفسيرها، وفي اختبار صحة الفروض المتعلقة بالظاهرة.

ومن هنا فإن الأرقام المستخدمة تؤثر في التحليلات الإحصائية، فإذا كانت الأرقام دقيقة وتدل على الصفة المقاسة دون تحيز، فإن التحليلات الإحصائية تتعامل مع أرقام دقيقة وجيدة، وبذلك لا نستطيع أن نقرر أن فهم النتائج وتفسيرها مستقل عن أدوات القياس المستخدمة، وإذا استخدم الباحث أدوات غير جيدة، وكانت الأرقام لا معنى لها فإن تحليل تلك الأرقام يؤدي إلى نتائج لا معنى لها أيضاً.

أسئلة الفصل الأول:

1. ماذا تعني المصطلحات التالية:

   بيانات، متغير، ثابت، إحصاء.

2. ما الفرق بين الإحصاء الوصفي والإحصاء الاستدلالي.
3. من خلال دراستك أعطِ مثالاً على المتغير الاسمي والمتغير الرتبي والمتغير المنفصل والمتغير المستمر.
4. عند دراسة العلاقة بين عدد السياح والإنفاق على السياحة في بلدٍ ما، المطلوب حدد ما يلي:

   أ. المتغير التابع. ب. المتغير المستقل.

   ج. المتغير المنفصل. د. المتغير المستمر.

5. ماذا يقصد بالمتغير المعدل، أعطِ مثالاً يوضح ذلك.
6. أعطي مثالاً على كل نوع من أنواع المقاييس.

الفصل الثاني CHAPTER 2

جمع البيانات الإحصائية

أهداف الفصل الثاني:

1. التعرف على مصادر جمع البيانات.
2. التعرف على أساليب جمع البيانات.
3. التعرف على أدوات جمع البيانات.
4. التعرف على خطوات جمع البيانات.

بعد التعُّرف على البحث العلمي وخصائصه ومراحل البحث العلمي وأنواع المتغيرات والمقاييس المختلفة يتضح أن البيانات هي القاعدة الأساسية لأي بحث علمي، فكل المراحل الإحصائية (جمع وتبويب وعرض بياني وتحليل) تتم على البيانات، لذلك فعملية جمع البيانات هي العملية الأولى التي ستبنى عليها كل المراحل التالية، وبالتالي كلما كانت البيانات المجموعة عن موضوع الدراسة صحيحة ودقيقة ومفصلة وشاملة لكل جوانب الموضوع كانت القاعدة سليمة وبالتالي تكون القرارات التي تتخذ بشأنها مبنية على أساس سليم.

1. مصادر جمع البيانات:

يمكن تحديد مصادر جمع البيانات بطريقتين هما:

1.1. حسب طبيعة البيانات:

هناك مصدران لجمع البيانات حسب طبيعة البيانات:

1.1.1. المصادر الأولية لجمع البيانات:

يطلق عليها أيضاً المصادر المباشرة أو المصادر الميدانية، حيث يتم الحصول على البيانات أو المعلومات من مصادرها الأصلية عن طريق الاتصال بوحدات أو مفردات الدراسة مباشرة، ويلجأ الباحث إلى هذا المصدر في حالة عدم توافر البيانات المطلوبة في أي مصدر من المصادر التاريخية، أو إذا كانت البيانات المنشورة قديمة ولا تعبر عن الظاهرة موضوع الدراسة في الوقت الحاضر، أو إذا كانت البيانات المنشورة لا تغطي كل جوانب الدراسة الحالية، ومن مزايا المصادر المباشرة في الحصول على البيانات أن درجة الدقة وحدود الثقة في البيانات تكون مرتفعة، مما يساعد على استخلاص نتائج موثوق فيها بدرجة كبيرة، إلا أنّ أهم المشاكل التي تواجه الاعتماد على هذا النوع من المصادر هي الحاجة إلى الوقت والتكلفة المادية اللازمين لإنجاز مهمة الحصول على البيانات ونتيجة لذلك فإنّ الباحث يجد نفسه مضطراً إلى بذل قصارى

جهده في جمع البيانات التي يحتاج إليها بالطريقة المباشرة في وقت قصير بأقل تكلفة مادية ممكنة.

2.1.1. المصادر الثانوية لجمع البيانات:

يطلق عليها أيضاً المصادر غير المباشرة أو المصادر التاريخية، حيث يتم الحصول على البيانات من الإحصاءات أو النشرات الإحصائية التي تنشرها الوزارات والمؤسسات الحكومية وشبه الحكومية والخاصة كل في مجال نشاطها، فبالإضافة إلى الهيئات المتخصصة في جمع البيانات تقوم أغلب الوزارات والهيئات بإنشاء نظام معلومات للإدارة يحوي إحصاءات عديدة عن طبيعة نشاطها في فترات زمنية مختلفة، وقد تشمل هذه النشرات الإحصائية بعض المعدلات والمقارنات بين العام الحالي للدراسة والأعوام السابقة، وهذه البيانات الموجودة في النشرات الإحصائية هي التي تمثل المصادر التاريخية للبيانات.

وحين يستخدم الباحث أو الدارس هذه البيانات عليه أن يتحرى عن مصدر هذه البيانات المنشورة، ويفضل أن يعود إلى مصدرها الأول الذي تولى جمعها ونشرها لأنه لا يخفى علينا أن تعاقب البيانات من ناشر لآخر قد يصيبها بشيء من التغيير سواءً عن قصد أو غير قصد، كما أنه يحسن أن نحيط بالظروف التي جمعت به حتى نستطيع أن نحيط بنواحي قصورها أو تحيزها عند استخدامها أو مقارنتها ببيانات أخرى، كما أنه يجب أن يشار إلى المصدر الذي استقيت منه هذه البيانات حماية للمستخدم من أي تحيز فيها، هذا فضلاً على الالتزام بمبادئ الخلق العلمي القويم في رد الفضل لذويه، وأهم ما يميز المصادر غير المباشرة أنها توفر الوقت والجهد والمال، ومن عيوبها أن البيانات قد تكون غير دقيقة في بعض الأحيان.

2.1. حسب طبيعة البحث:

1.2.1. بيانات المجتمع (Population Data):

وهي البيانات التي تجمع من مجموعة كلية أو كينونة كاملة (Complete Group) يتشاركون في قواعد عامة من الخصائص، وقد تكون مجموعة من الأفراد أو الظواهر أو المنشآت.

2.2.1. بيانات العينة (Sample Data):

وهي البيانات التي تجمع من مجموعة جزئية تمثل جزءاً من المجتمع أو عدداً من الحالات التي تؤخذ من المجتمع الأصلي، أي استخدام بيانات تمثل عدداً من الوحدات أو جزءاً من المجتمع (A Subset) الكبير للوصول إلى استنتاجات عن المجتمع الكبير معتمدين على التحليل الإحصائي.

2. أساليب جمع البيانات:

يتحدد الأسلوب المستخدم في جمع البيانات، حسب الهدف من البحث، وحجم المجتمع محل البحث، وهناك أسلوبان لجمع البيانات هما:

1.2. أسلوب الحصر الشامل:

يسمى في بعض الأحيان بأسلوب المسوحات الشاملة، وفيه يتم جمع البيانات من جميع عناصر المجتمع الإحصائي دون استثناء، وبالتالي فإن النتائج تكون على مستوى عالٍ من الدقة والتفصيل والوضوح، ولكن يعاب عليها ارتفاع تكاليفها وحاجتها للكثير من الوقت والجهد وحاجاتها لعدد كبير من الباحثين وأحياناً يصعب أو يستحيل اتباعها.

ويستخدم هذا الأسلوب في الحالات التالية:

- إذا كان المطلوب الحصول على بيانات على مستوى عالٍ من الدقة كما هو الحال في التعدادات العامة للسكان أو الصحة أو الزراعة أو الصناعة وغيرها...

- إذا كان المجتمع قيد الدراسة صغيراً.
- إذا تعذر الحصول على الحجم والتوزع الفعلي للمجتمع المدروس.

2.2. أسلوب العينة:

يسمى في بعض الأحيان بمسوحات العينة، وفيه يتم جمع البيانات من جزء معين من مجتمع الظاهرة قيد الدراسة وفق قواعد علمية معينة تناسب الدراسة لكي تمثل المجتمع المدروس أصدق وأفضل، وبالتالي الحصول على عينة جيدة نظرياً أي يجب أن تكون عشوائية والعشوائية تعني توافر طرق تحديد عدد عناصرها وتحديد فرصة اختيار أي عنصر من عناصرها كما أنه بواسطة العينة الجيدة يمكننا تحديد الثقة في تقدير مؤشرات المجتمع الذي سحبت منه.

والأسباب التي تدعو إلى استخدام العينات:

- قد يؤدي المسح الشامل أحياناً إلى فساد عناصر المجتمع قبل فحص البيانات.
- توفير الوقت والجهد والنفقات.
- تعذّر الوصول إلى جميع أفراد المجتمع.

3. أدوات جمع البيانات:

يقصد بأداة جمع البيانات الوسيلة التي تتم بواسطتها عملية جمع البيانات بهدف اختبار فرضيات البحث أو الإجابة عن تساؤلات البحث.

وقد يستخدم الباحث طريقة أو أكثر لجمع البيانات لذلك يجب على الباحث اختيار الطريقة أو الوسيلة التي ستستخدم في عملية الحصول على البيانات، ويجب على الباحث أن يقرر مسبقاً الطريقة المناسبة لبحثه أو دراسته وأن يكون ملماً بالأدوات والأساليب المختلفة لجمع البيانات لأغراض البحث العلمي، وعملية اختيار أداة البحث لا تخضع لرغبة الباحث أو ذوقه في استخدام أداة دون أخرى وإنما حسب أغراض البحث، لأن لأدوات وأساليب جمع البيانات ميزاتها الإيجابية والسلبية ويعتمد اختيار الباحث

لطريقة جمع البيانات على عدة عوامل منها طبيعة البحث ومدى ملائمة طريقة جمع البيانات، طبيعة مجتمع البحث وعينة الدراسة، وظروف الباحث وقدراته المالية والوقت المتاح له، مدى معرفة الباحث بالطريقة أو الأداة.

ومن طرق جمع البيانات:

- الملاحظة (Observation).
- المقابلة (Interview).
- الإستبانة (Questionnaire).
- الاختبارات (Tests).
- الوثائق (Documents).

وسوف نعرض هنا مفهوم كل أداة من الأدوات جمع البيانات، وسنتناول بشيء من التفصيل الاستبانة.

1.3. الملاحظة (Observation):

تعد الملاحظة أقدم وسيلة من وسائل جمع البيانات والمعلومات الخاصة بظاهرة ما. تعني الملاحظة بمعناها البسيط: الانتباه العفوي إلى حادثة أو ظاهرة أو أمر ما، أما الملاحظة العلمية فهمي: انتباه مقصود ومنظم ومضبوط للظاهرات أو الحوادث أو الأمور بغية اكتشاف أسبابها وقوانينها.

وتستخدم الملاحظة كأداة في جمع البيانات والمعلومات وخاصة في دراسة الظاهرات الاجتماعية ودراسة تحليل المضمون والوثائق كما أنها تستخدم في حالة عدم تمكن الباحث من استخدام المقابلة والاستبيان لجمع المعلومات كملاحظة السلوك العدواني عند مجموعة من الأطفال، وتتميز الملاحظة بدقة المعلومات بسبب ملاحظة الظاهرات في ظروفها الطبيعية وفي أثناء حدوثها، ويمكن إجراء الملاحظة على قليل من المفحوصين، لكن الملاحظة تتطلب وجود المُلاحظ نفسه في موقع العمل لفترة طويلة من

الزمن كما أن الملاحظين قد يغيرون سلوكهم إذا شعروا بإجراء الملاحظة، وقد يحدث تحيز من الباحث إما بسبب تأثره بالأفراد أو بسبب عدم نجاحه في تفسير ظاهرة ما.

وهناك نوعان من الملاحظة:

1.1.3. الملاحظة المُقننة: ويكون لدى المُلاحظ مجموعة محددة مسبقاً من الأنشطة والظواهر التي يخطط لملاحظتها ومراقبتها.

2.1.3. الملاحظة غير المُقننة: ولا يكون لدى المُلاحظ أدنى فكرة عن الجوانب التي يرغب التركيز عليها بالملاحظة، لذلك يسجل ما يلاحظه.

2.3. المقابلة:

تعتبر المقابلة أداة من أدوات جمع البيانات وهي عبارة عن استبيان منطوق، والفرق الرئيسي بين المقابلة والاستبيان أنها تتضمن تفاعلاً مباشراً بين الباحث والمستجيب، فالمقابلة مرنة ويمكن تعديلها حسب الموقف، وتستخدم خاصة مع الأشخاص الذين هم على درجة كبيرة من الأمية، أو في الحالات التي تتصف بالسرية كعدم اعتماد على الاستبانة في جمع البيانات، وأهم عيوب المقابلة هي أنها معرضة للذاتية والتحيز من جانب الباحث، كما أن تكاليفها عالية وتستغرق وقتاً طويلاً، ويتوقف نجاح المقابلة على مهارة وتدريب الباحث ومساعديه، وللتغلب على عيوب المقابلة يجب أن يكون الباحث وسيطاً محايداً لتبادل المعلومات.

ويمكن تقسيم المقابلات من حيث درجة المرونة إلى:

1.2.3. المقابلات المقننة: وفيها تكون أسئلة المقابلة محددة ويتبع كل سؤال مجموعة من الاختبارات أو الإجابات يختار من بينها المستجيب الإجابة التي تتفق مع رأيه، وتمتاز المقابلة المقننة بأن الاختلافات تكون محدودة، مما يرفع من ثبات المقابلة، كما أنه لا مجال أمام الباحث لتفسير أو تعديل الاستجابات مما يزيد من صدق المقابلة.

2.2.3. المقابلات شبه المقننة: وهنا نجد أن الأسئلة لا يتبعها اختيارات محددة ولكن تصاغ الأسئلة بحيث تسمح بالإجابات الفردية، فالسؤال مفتوح ولكنه محدد للغاية في محتواه، وتمتاز بأنها تسمح بوجود علاقة تفاعلية بين الباحث والمستجيب، مما يجعله يقبل على المقابلة ويبذل جهده في إعطاء المعلومات الضرورية، ولكن يعيبها أنها أقل موضوعية من المقابلة المقننة.

3.2.3. المقابلات غير المقننة: في هذا النوع من المقابلات يقوم الباحث بتوجيه أسئلة عريضة في أي ترتيب يراه مناسباً، وهنا نجد أن المقابلة تركز على المستجيب، وعلى الرغم من أن هذا الأسلوب يتميز بوجود علاقة ألفة كبيرة بين الباحث والمستجيب، إلا أن درجة صدقها وثباتها محدودة مقارنة بالنوعين السابقين.

وأفضل طريقة لتسجيل الإجابات هي تدوينها حرفياً كما أعطاها المستجيب، ويجب توجيه الأسئلة بالنص كما هي مكتوبة في دليل المقابلة، كما يجب على الباحث أن يتجنب إعادة صياغة السؤال بعبارات من عنده لأنه قد يؤثر على معنى السؤال.

3.3. الاستبيان:

يعدّ الاستبيان أداة مفيدة من أدوات البحث العلمي للحصول على الحقائق، والتوصل إلى الوقائع والتعرف على الظروف والأحوال وجمع البيانات عن الظواهر التي لا يمكن ملاحظتها بشكل مباشر (كالميول والآراء والمعتقدات.....) أو التي يمكن ملاحظتها بشكل مباشر، وهو في بعض الأحيان الوسيلة العملية الوحيدة للقيام بالدراسة العملية.

والاستبيان عبارة عن وثائق مكتوبة توجه نفس الأسئلة إلى جميع الأفراد في العينة، أي تمكننا الاستبيانات من جمع بيانات عن الفرد من الفرد ذاته، حيث يسجل المستجيبون إجابات مكتوبة عن كل سؤال (عبارة) من أسئلة (عبارات) الاستبيان، ولذلك

نجد أن المستجيبين عملياً يتحكمون في جمع البيانات لأنهم يملؤون الاستبيان بالطريقة التي يرونها مناسبة.

هناك تعريفات عديدة للاستبيان ومن الصعب حصرها، لكننا سوف نقوم باستعراض أهمها:

- أداة تتضمن مجموعة من الأسئلة أو الجمل الخبرية التي يطلب من المفحوص الإجابة عنها بطريقة يحددها الباحث حسب أغراض البحث.
- مجموعة من الأسئلة المكتوبة والتي تعد بقصد الحصول على معلومات أو آراء المبحوثين حول ظاهرة أو موقف معين.
- وسيلة لجمع المعلومات المتعلقة بموضوع البحث عن طريق إعداد استمارة يتم تعبئتها من قبل عينة ممثلة من الأفراد ويسمى الشخص الذي يقوم بإملاء الاستمارة بالمستجيب.

وبعد هذا العرض لبعض تعريفات الاستبيان يمكننا صياغة التعريف التالي:

هي أداة تتضمن مجموعة من الأسئلة والعبارات يقوم الباحث بوضعها بشكل مرتب بأسلوب منطقي مناسب يتم الإجابة عليها من قبل عينة ممثلة من الأفراد بطريقة يحددها الباحث حسب أغراض البحث وتستخدم لجمع البيانات عن الظواهر التي يمكن ملاحظتها بشكل مباشر وغير مباشر.

تتطلب عملية تصميم الاستبيان المرور بالخطوات التالية:

1) تحديد موضوع البحث أو الدراسة بشكل عام.
2) تعريف هدف الدراسة.
3) تحديد الأفراد الذين سيطلب منهم الإجابة على الاستبيان.
4) وضع الأسئلة حول كل موضوع من موضوعات الدراسة بحيث تكون هذه الأسئلة مناسبة وتراعي عدم التكرار.

5) الاطلاع على الدراسات السابقة التي تناولت نفس موضوع الاستبيان.
6) عرض الاستبيان على مجموعة من المحكمين المتخصصين في مجال البحث العلمي للحكم عليها وملائمتها لموضوع البحث، كما يجب عرض الاستبيان على عدد محدد من أفراد مجتمع الدراسة والطلب منهم التعليق عليها وبيان الأسئلة الغامضة أو غير المفهومة ومدى تغطيتها لموضوع الدراسة.
7) إخضاع الاستبانة للتجريب بعد إعدادها بصورتها الأولى وعرضها على المحكمين.
8) التأكيد على طباعة الاستبيان بشكل واضح من حيث: الوضوح، الترتيب، المسافات الثابتة بين كل سؤال وآخر.

هناك بعض الأمور التي يجب مراعاتها عند صياغة أسئلة وعبارات الاستبيان:

1) يجب صياغة عبارات وأسئلة بشكل واضح ومحدد بحيث يفسرها جميع المستجيبين بنفس الطريقة.
2) يجب استخدام لغة سهلة واضحة تتناسب مع مستوى المبحوثين حتى يمكن إعطاء معلومات صحيحة موثوق بها.
3) تجنب استخدام المصطلحات أو التعابير غير المفهومة وعند ورود مثل هذه المصطلحات أو التعابير فيجب على الباحث تعريفها.
4) تجنب كتابة الأسئلة العامة حتى لا تكون تفسيراته كثيرة.
5) تجنب كتابة الأسئلة أو العبارات المزدوجة لأنها تحتوي أكثر من فكرة وتسبب إرباكاً للمستجيب فقد يوافق على جزء من السؤال ويرفض الجزء الآخر، أي كل سؤال يعالج فكرة واحدة فقط.
6) عند صياغة السؤال أو العبارة يجب الابتعاد عن الصياغة المتحيزة التي توحي بإجابة معينة.
7) عند كتابة أسئلة أو عبارات منفية يجب توضيح أداة النفي بشكل جيد أو وضع تحتها خط لأنه كثيراً ما يغفل المستجيبون أداة النفي.

8) يجب أن تكون الأسئلة متعلقة بالموضوع ومهمة للمستجيب فلو سألنا أسئلة تافهة وغير مهمة بالنسبة له فلن يعطي الاهتمام الكافي للاستبيان وبالتالي الحصول على إجابات غير موثوق بها.
9) يفضّل البدء بالأسئلة السهلة التي لا تحتاج إلى تفكير عميق من المبحوث ثم التدرج إلى الأسئلة الأكثر صعوبة.
10) يجب أن يكون طول السؤال مناسباً ويجب تجنب الأسئلة الطويلة التي قد تؤدي لتملل المبحوث.
11) تجنب الأسئلة الموجبة (هل تتقيد بالأنظمة والقوانين؟) والأسئلة الشخصية.

1.3.3. أنواع أسئلة الاستبيان:

يمكن للباحث استخدام أنواع مختلفة من الأسئلة في الاستبانة، ويعتمد ذلك على طبيعة الدراسة وإمكانيات الباحث ومهاراته في تحليل البيانات وبشكل عام تقسم الأسئلة على النحو التالي:

1.1.3.3. الأسئلة المغلقة أو محدودة الإجابات:

وفي هذا النوع من الأسئلة يحدد الباحث الإجابات الممكنة لكل سؤال وتكون مهمة المستجيب اختيار أحدها، ومن مزايا هذا النوع من الأسئلة أنه يسهّل التفاعل مع الأسئلة وتوضيح معاني الأسئلة، كما أنّ الميزة الأهم سهولة عملية تصنيف الإجابات وتبويبها وتحليلها، **أنواع الأسئلة المغلقة:**

- **الأسئلة الثنائية:** وتكون الإجابات محصورة بخيارين وعلى المستجيب اختيار أحدها، مثل: هل أنت مدخن (نعم- لا) أو الإجابة على سؤال الجنس (ذكر- أنثى).
- **أسئلة الاختيار من متعدد:** يوجد هناك العديد من الإجابات المطلوب من المستجيب اختيار أحدها حسب رأيه، مثل الإجابة على سؤال الدخل الشهري

يوجد عدة إجابات للسؤال كل جواب يمثّل فئة محددة من الدخل ويختار المستجيب الجواب الذي يمثّل دخله الشهري.

- **الأسئلة المدرجة:** وهي الأسئلة التي تقدّم عدة اختيارات أو بدائل ويجب على المستجيب أن يرتّب هذه البدائل وفق تدرّج يحدده السؤال، مثلاً ترتيب الإجابات المحددة في السؤال وفق الأهمية بالنسبة للمستجيب من الأهم إلى الأقل أهمية.

2.1.3.3. الأسئلة المفتوحة:

يستخدم هذا النوع من الأسئلة عندما لا يكون لدى الباحث معلومات كافية عن كافة موضوعات الدراسة أو الإجابات المحتملة ويرغب في الحصول على معلومات تفصيلية ومحققة حول الظاهرة أو المشكلة، ويمتاز هذا النوع من الأسئلة بأنه لا يقيد المبحوث بإجابة محددة وإنما يعطيه حرية في الإجابة على السؤال المطروح بطريقته ولغته وأسلوبه الخاص الذي يراه مناسباً.

3.1.3.3. الأسئلة المغلقة المفتوحة:

في مثل هذا النوع من الأسئلة يطرح الباحث في البداية سؤالاً مغلقاً ثم يتبعه بسؤال مفتوح يطلب فيه من المبحوثين توضيح أسباب اختياره للإجابة المعينة أو إضافة إجابة أخرى.

أسئلة الفصل الثاني

1. هناك مصدران لجمع البيانات، اذكرهما مع تحديد الفرق بينهما.
2. عرف الملاحظة، عدد أنواعها مع ذكر مثال مناسب لكل نوع من أنواع الملاحظة.
3. عرف المقابلة، عدد أنواعها مع ذكر مثال مناسب لكل نوع من أنواع المقابلة.
4. حدد الطريقة المناسبة في جمع البيانات لإجراء بحث للظواهر الآتية مع ذكر الأسباب:
   - لدراسة عدد وجنس دخول السياح إلى أحد المطارات في وقت محدد.
   - لدراسة حالة الأميين الاجتماعية والاقتصادية.
   - لاستطلاع آراء أعضاء الهيئة التدريسية في كلية السياحة بشأن تطوير العملية التدريسية.
   - لمراقبة سلوك طلاب من الصف الأول خلال زيارتهم لحديقة الحيوانات.
5. صمم نموذجاً لاستبيان لاستطلاع آراء عينة من طلاب كلية السياحة في جامعة دمشق بهدف تحديد العوامل المؤثرة في العمل التطوعي عند طلاب كلية السياحة، بحيث يحوي الاستبيان الأقسام التالية:
   - القسم الأول: خصائص الطالب ويحوي الجنس، العمر، السنة الدراسية، الاختصاص، دخل الأسرة الشهري.
   - القسم الثاني: ممارسة العمل التطوعي.
   - القسم الثالث: أسباب ممارسة العمل التطوعي.
   - القسم الرابع: معوقات العمل التطوعي.

وقل رب زدني علما
جامعة دمشق
Damascus University

الفصل الثالث CHAPTER 3

المجتمع والعينة

أهداف الفصل الثالث:

1. التعرف على المفاهيم الأساسية المستخدمة في اختيار العينة (المجتمع الإحصائي – إطار المجتمع).
2. التعرف على العينات الاحتمالية (العشوائية) وأنواعها.
3. التعرف على العينات غير الاحتمالية (غير العشوائية) وأنواعها.
4. التعرف على طرق تحديد حجم العينة.

تتطلب دراسة مشكلة أو ظاهرة ما توافر بيانات ومعلومات ضرورية عن هذه الظاهرة أو المشكلة لتساعد الباحث في اتخاذ القرار السليم حيالها، وإذا استطاع الباحث إجراء دراسته على جميع أفراد المجتمع، فإن دراسته تكون ذات نتائج أقرب للواقع وأكثر دقة، ولكن الباحث قد يجد صعوبة في التعامل مع كل مشاهدة من مشاهدات المجتمع لعدة أسباب، مما سيضطره لإجراء الدراسة على مجموعة جزئية من مجتمع البحث، وهذه المجموعة نسميها عينة الدراسة، فالتحديد الدقيق والواضح لمجتمع البحث أمر ضروري جداً لأنه سيساعد في تحديد الأسلوب العلمي الأمثل لدراسة هذا المجتمع واشتقاق العينة الإحصائية التي تمثل مجتمع البحث.

1. المجتمع الإحصائي:

عند القيام بدراسة ظاهرة معينة أو مشكلة ما فيجب علينا أن نحدد الوحدات الإحصائية التي تتوافر فيها الظاهرة المراد دراستها، ومن هنا تعرف **الوحدة الإحصائية** على أنها كائن أو شيء يتصف بالظاهرة أو الخاصة المدروسة ويمكن أن تأخذ أشكالاً متعددة:

- الوحدة الأولية.
- الوحدة السردية أو وحدة القائمة.
- الوحدة المبدئية.

وكمثال عن أشكال الوحدة الإحصائية: تم أخذ عينة من السائحين في مدينة اللاذقية لدراسة نفقات الأسرة على السياحة، ومن أجل ذلك تقسم المدينة إلى قطاعات متقاربة في المساحة، ومن هذه القطاعات تختار عشوائياً عدة قطاعات كل قطاع يشكل **وحدة أولية** ثم يُعد لكل وحدة أولية منتقاه قائمة بالمنشآت السياحية (فنادق – منتجعات – شاليهات – منازل) الموجودة في كل وحدة أولية منتقاه حيث يشكل كل مبنى وحدة من

وحدات القائمة أو وحدة سردية، ثم في كل مبنى نحصر عدد الأسر السائحة التي تقطن المبنى حيث تشكل كل أسرة **وحدة مبدئية**.

وفي بعض الأحيان قد تنطبق وحدتان مع بعضهما، ويطلق على أول وحدة مسحوبة اسم وحدة المعاينة Sampling Unit لهذا قد تكون وحدة المعاينة وحدة أولية أو وحدة سردية أو قد تنطبق على الوحدة المبدئية نفسها.

ويطلق على مجموع الوحدات الإحصائية التي تتصف بخاصة معينة اسم المجتمع الإحصائي (Population) ويكون تعريف المجتمع الإحصائي كما يلي: المجتمع الإحصائي هو كل أو مجموع الوحدات الإحصائية التي تتوافر فيها الظاهرة أو الخاصة المدروسة، فلو أراد الباحث دراسة متوسط العمر الحالي للفرد السوري فإن المجتمع هو كل السوريين صغاراً وكباراً ذكوراً وإناثاً.

ونذكر أيضاً أن هناك نوعين من المجتمعات الإحصائية:

- مجتمع الهدف (Target Population): وهو المجتمع المستهدف بالدراسة.
- مجتمع العينة (Sample Population): وهو المجتمع الذي سيتم اختيار العينة منه، ويعمم عليه النتائج.

ويوجد مجتمع محدود ومجتمع غير محدود (لانهائي)، فالمجتمع المحدود هو ذلك المجتمع الذي يمكن حصر عدد مفرداته، مثل مجتمع طلاب كلية السياحة في جامعة دمشق، وفي كثير من الحالات التي يهتم بها الباحثون تهدف مجتمع محدود ولكنه كبير جداً والذي يُنظر إليه على أنه مجتمع غير محدود مثل مجتمع يحتوي على عدة ملايين من الأفراد. أما المجتمع الذي يصعب حصر مفرداته مثل مجتمع الأسماك في البحر، فيقال عنه مجتمع غير محدود.

ويتصف المجتمع الإحصائي بصفتين رئيسيتين:

- المرونة: ويقصد به إمكانية تحديد الإطار الزماني والمكاني.

- التجانس: ويقصد به أن جميع الأفراد متجانسون من حيث الخاصية المدروسة.

2. إطار المجتمع الإحصائي:

الإطار لغةً هو الشيء الذي يحيط بشيء آخر ويحصره حصراً تاماً كإطار الصورة مثلاً، فإن إطار المجتمع الإحصائي أو اختصاراً الإطار هو قائمة تشمل جميع وحدات (مفردات) المجتمع الإحصائي والتي سوف يؤخذ منها العينة، ويسمى في بعض الأحيان إطار المعاينة (Sample Frame)، فالإطار هو الأداة أو الوسيلة التي تمثل المجتمع ويجب أن نشير هنا إلى أن الإطار قد لا يمثل المجتمع الإحصائي تماماً مئة بالمئة وذلك نتيجة لطبيعة المجتمع المدروس وبوجه عام هناك بعض الشروط الواجب توافرها في إطار المجتمع (المعاينة) حتى يتجنب الباحث الوقوع فيما يسمى بأخطاء المعاينة، ومن هذه الشروط:

- أن يكون الإطار شاملاً لكافة وحدات المجتمع المدروس.
- حديثاً حتى يعبر عن الظاهرة موضوع الدراسة في الوقت الحاضر.
- منتظماً ليسهل عملية اختيار الوحدات فتسلسل الأرقام يساعد على دقة اختيار وحدات العينة.

ويأخذ الإطار واحداً من عدة أشكال لعل من أهمها:

- القائمة: كأن تتوافر قائمة بأسماء المنتجعات السياحية في منطقة جغرافية محددة حيث تشكل هذه القائمة إطاراً للمنتجعات السياحية في تلك المنطقة الجغرافية.
- السجل: هناك العديد من السجلات التي تشكل إطاراً للظواهر المدروسة، كعدد العمال وعدد الغرف في المنتجعات السياحية.
- الدليل: في العديد من الدول تتوافر أنواع من الأدلة التي تحتوي على كل أو معظم الفعاليات التي تجمعهم صفة معينة، فمثلاً الدليل الأصفر الذي يحوي أسماء وعناوين وهواتف المكاتب السياحية المنتشرة في مدينة ما.

ومن أهم استخدامات الإطار وفق أي شكل من الأشكال هو سحب وحدات العينة التي يراد دراستها، ومن الجدير بالذكر أن الإطار يحتاج إلى عمليات تحديث بين الحين والحين الآخر لكثرة التغيرات التي تطرأ عليه.

3. العينة:

بعد أن تعرفنا على مفهوم الوحدات الإحصائية والمجتمع الذي تتألف منه، يمكن للباحث أن يدرس الظاهرة أو الخاصة المطلوبة من خلال أسلوب الحصر الشامل أو التعداد العام حيث يدرس الظاهرة من خلال جميع وحدات المجتمع، وقد يلجأ الباحث لأسلوب العينة حيث ينتقي مجموعة من الوحدات أو الأفراد من المجتمع ويجمع المعلومات منها أو يخضعها لإجراء معين وذلك ليحصل على صورة ممثلة للمجتمع الذي انتقيت منه، ووفقاً لذلك يمكن تعريف العينة كما يلي:

هي مجموعة جزئية منتهية أو محدودة من مجتمع الدراسة يتم اختيارها بطريقة علمية ومناسبة بهدف دراسة خصائص المجتمع الذي سحبت منه العينة واستخلاص نتائج حوله، بحيث إن عناصرها تماثل عناصر المجتمع من حيث المؤشرات الإحصائية (المتوسط – الانحراف المعياري .....)، أي هي مجموعة جزئية من مجتمع لهما خصائص مشتركة.

وعند اختيار العينة يراعي الباحث الخطوات التالية:

- تحديد مجتمع الدراسة.
- تحديد خصائص المجتمع.
- تحديد حجم العينة بدقة وببساطة لتنفيذها في الواقع.
- الاختيار العشوائي للعينة.

4. أنواع العينات:

من خلال تعريف العينة نلاحظ أن هناك عدة أنواع للعينات تبعاً لإجراءات اختيارها بحيث يمكن للباحث أن يختار العينة حسب ظروف الدراسة التي يقوم بها، وتصنف العينات إلى:

1.4. العينات الاحتمالية (العشوائية):

هي العينات التي يتم اختيار مفرداتها وفقا لقواعد الاحتمالات، بمعنى آخر هي التي تعتمد على العشوائية في اختيار مفردات العينة من مجتمع الدراسة، بحيث يكون لكل فرد من أفراد المجتمع فرصة متساوية لاختياره فيها، وأيضاً له نفس الفرصة لاختياره ضمن الفئة الواحدة المتجانسة المكونة للمجتمع الإحصائي، بهدف تجنب التحيز الناتج عن اختيار المفردات، ويتم السحب العشوائي بترقيم جميع عناصر المجتمع بأرقام متسلسلة دون تميز ، ثم نستخدم واحدة من الطرق التالية لاختيار مفردات العينة حسب حالة المجتمع الذي ندرسه، وهذه الطرق هي:

- طريقة جداول الأرقام العشوائية: جداول الأرقام العشوائية هي جداول معدة خصيصاً لاختيار العينات وهي عبارة عن تسلسل أرقام حيث نقوم بوضع إصبعنا على رقم ما عندها يتم اختيار العنصر من المجتمع الذي يحمل الرقم نفسه وهكذا يتم تحديد المفردات التي ستدخل بعينة الدراسة، وتستخدم هذه الطريقة في المجتمعات الإحصائية الكبيرة.
- طريقة دواليب الحظ: حيث نقوم بوضع عدد الدواليب حسب مرتبة حجم المجتمع ثم نقوم بتحريك الدواليب كما في السحب على جوائز يانصيب، وتستخدم هذه الطريقة من أجل المجتمعات الإحصائية الكبيرة والصغيرة.

- طريقة القرعة: نضع أوراق ضمن وعاء أو كيس بحيث كل ورقة تحمل رقماً لعنصر من عناصر المجتمع الإحصائي ثم نقوم بسحب العدد المطلوب من الوعاء وبشكل عشوائي، وتستخدم هذه الطريقة في المجتمعات الإحصائية الصغيرة.
- طريقة الحاسوب: حيث نقوم بإدخال حجم المجتمع الإحصائي وحجم العينة ثم يقوم الحاسوب بتحديد المفردات التي تنتمي للعينة.

ومن أهم أنواع العينات الاحتمالية، ما يلي:

1.1.4. العينة العشوائية البسيطة (Simple Random Sample):

وهي أبسط أنواع العينات، ويتم اختيارها على أساس أن يكون هناك فرص متساوية أمام جميع عناصر المجتمع للظهور في العينة أي لا يتم التحيز لأي عنصر من عناصر المجتمع على حساب عنصر آخر، ويتم اختيار عناصر هذه العينة حسب طرق السحب العشوائي.

تتميز العينة العشوائية البسيطة ببساطة تطبيقها واستعمالها، وهي الأساس الذي يرتكز عليه لاختيار العينات العشوائية الأخرى، كما تتميز بأن نتائجها تكون قابلة للتعميم على مجتمع الدراسة الأصلي إذا كان حجم العينة كبيراً نسبياً، ولكن يؤخذ عليها صعوبة تمثيلها لكل عناصر المجتمع وخاصة عندما يكون عناصر مجتمع الدراسة منتشرين في مناطق جغرافية متباعدة، أو عندما يكون المجتمع غير متجانس مثال إذا كان المجتمع الذي ندرسه يحوي 1000 سائح، 600 ذكر و400أنثى ونريد اختيار 100 سائح وحصلنا على 80 ذكر و20 أنثى نكون متحيزين للذكور أو الإناث وهكذا.

2.1.4. العينة العشوائية المنتظمة (Systematic Random Sample):

وهي الشكل الثاني من أشكال العينات الاحتمالية وتعني اختيار عناصر العينة من بين كل عناصر المجتمع بطريقة منتظمة حيث نقوم بتقسيم المجتمع الأصلي إلى فئات متساوية في الحجم (عدد الفئات = حجم المجتمع الأصلي ÷ حجم العينة) بحيث نختار العنصر الأول بشكل عشوائي من الفئة الأولى، ثم بعد ذلك يتم اختيار بقية العناصر بطريقة منتظمة من كل الفئات ومن هنا جاءت تسميتها بالعينة العشوائية المنتظمة.

مثال 1:

بفرض أننا نريد اختيار عينة عشوائية منتظمة حجمها (25) طالب من طلاب السنة الرابعة في كلية السياحة المكون من (125) طالب.

في الخطوة الأولى نقسم المجتمع إلى فئات متساوية بحيث عدد الفئات يساوي حجم العينة فيكون لدينا 25 فئة.

وفي الخطوة الثانية نحدد عدد العناصر في كل فئة من الفئات وذلك بتقسيم حجم المجتمع على حجم العينة أي (125 ÷ 25= 5) فيكون ضمن كل فئة من فئات المجتمع 5 عناصر، وسيكون إطار المجتمع على الصورة التالية:

الفئة	العناصر				
الأولى	1	2	3	4	5
الثانية	6	7	8	9	10
⋮	⋮	⋮	⋮	⋮	⋮
الرابعة والعشرون	116	117	118	119	120
الخامسة والعشرون	121	122	123	124	125

الخطوة الثالثة: نختار العنصر الأول من الفئة الأولى بطريقة عشوائية (حسب طرق الاختيار العشوائي) وليكن اختيارنا للعنصر الذي يحمل الرقم 4 وبالتالي العنصر الأول في العينة الذي اخترناه هو الطالب الذي يحمل الرقم 4، نختار الطالب الثاني وذلك بجمع 4 مع عدد عناصر الفئة أي مع 5 فنحصل على 9 فيكون العنصر الثاني الذي اخترناه في العينة هو الطالب الذي رقمه 9 وهكذا نستمر حتى نحصل على آخر مفردة والتي يكون رقمها 124 أي العنصر الأخير في العينة هو الطالب الذي رقمه 124 ونكون قد حصلنا على عينة حجمها 25 مفردة.

الفئة	العناصر				
الأولى	1	2	3	(4)	5
الثانية	6	7	8	(9)	10
⋮	⋮	⋮	⋮	⋮	⋮
الرابعة والعشرون	116	117	118	(119)	120
الخامسة والعشرون	121	122	123	(124)	125

ويتميز هذا النوع من العينات بقلة تكاليفه وسهولة إجرائه ولكن يؤخذ على هذا النوع من العينات العشوائية أنه إذا كان العنصر الأول المختار في العينة متحيزاً تصبح العينة كلها متحيزة، كما أنها لا تصلح في الاختيار إذا كان أفراد المجتمع مرتبين في قوائم بنظام خاص يزيد من فرص ظهور أفراد بذاتهم يختلفون عن غيرهم في سمة من السمات المتصلة بموضوع البحث، وإذا كان حجم المجتمع كبيراً فإنه يصعب سحب مفردات العينة بهذه الطريقة.

3.1.4. العينة العشوائية الطبقية (Stratified Random Sample):

تعد العينة العشوائية الطبقية من أفضل أنواع العينات العشوائية التي يتم سحبها عندما يشمل المجتمع المدروس على عناصر غير متجانسة من حيث الخصائص التي يقوم الباحث بدراستها، حيث نقوم بتقسيم المجتمع إلى مجموعات مؤلفة من عناصر متجانسة على أساس متغير واحد أو أكثر (السن، الفئات العمرية، المستوى التعليمي..) وتدعى كل مجموعة طبقة ومجموع هذه الطبقات يكون المجتمع الأصلي، والسبب بتقسيم المجتمع إلى طبقات للتأكد من تمثيل كل طبقة لتكون العينة ممثلة بقدر الإمكان للمجتمع المدروس ومن جهة ثانية لتكون الطبقات على درجة كبيرة من التجانس.

ثم نختار عينة عشوائية (بسيطة – منتظمة) من كل طبقة من هذه الطبقات ويتوقف تحديد حجم العينة المسحوبة من كل طبقة على العاملين التاليين:

- حجم الطبقة: فحجم العينة المسحوبة من كل طبقة يتناسب طرداً مع حجم هذه الطبقة في المجتمع، وهذه الطريقة تسمى التخصيص النسبي (Proportional) وأحياناً يكون حجم العينة الفرعية متساوياً لجميع الطبقات.
- مدى التجانس داخل الطبقة: فكلما زادت درجة التجانس بين مفردات الطبقة قل حجم العينة المسحوبة من هذه الطبقة.

وهناك عدة أساليب لتوزيع العينة على الطبقات منها:

i. أسلوب التوزيع المتساوي (Equal Allocation): وهو أدنى مستويات الدقة في اختيار العينة العشوائية الطبقية وفيه نقسم عدد عناصر العينة الكلية على طبقات المجتمع بالتساوي حتى لو اختلف عدد الأفراد ضمن كل طبقة عن عدد أفراد الطبقة الأخرى في هذا المجتمع.

مثال 2:

حسب التعداد السكاني العام في الجمهورية العربية السورية عام 2004 كان عدد السكان في الفئة العمرية (20 – 24) سنة (1864000) نسمة يتوزعون (920000) إناث بنسبة (49.35%) والذكور (944000) بنسبة (50.65%)، وإذا أردنا سحب عينة عشوائية حجمها (1864) نسمة وبطريقة العينة الطبقية بأسلوب التوزيع المتساوي.

بما أن نوع العينة طبقية بأسلوب التوزيع المتساوي فيجب أخذ 50% من حجم العينة الذكور و50% من حجم العينة إناث أي 932 ذكر و932 أنثى وليس حسب توزعهم في المجتمع.

ii. أسلوب التوزيع المتناسب (Proportional Allocation): يقصد به تمثيل الطبقة في العينة حسب وزنها النسبي في المجتمع وذلك حسب المعادلة التالية:

$$n_r = (N_r / N) * n$$

حيث: n_r: حجم العينة في الطبقة r.

N_r: عدد عناصر الطبقة r ضمن المجتمع.

N: عدد عناصر المجتمع الكلي.

n: حجم العينة الكلي.

مثال 3:

باحث مهتم بدراسة المنشآت السياحية في العاملة في محافظات الجمهورية العربية السورية من مختلف درجات التصنيف السياحي وقد تم تحديد حجم العينة بـ 50 مفردة وكانت المنشآت تتوزع كما يلي (25 خمس نجوم – 47 أربع نجوم – 66 ثلاث نجوم – 196 نجمتين، و**المطلوب:** سحب عينة عشوائية حجمها (50) مفردة وبطريقة العينة العشوائية الطبقية بأسلوب التوزيع المتناسب.

نوجد الحجم النسبي للعينة وفق كل طبقة حيث:

$$N = 334$$
$$N_1 = 25$$
$$N_2 = 47$$
$$N_3 = 66$$
$$N_4 = 196$$
$$n = 50$$

حجم العينة في طبقة المنشآت ذات الخمس نجوم:

$$n_1 = (N_1/N) * n = (25/334) * 50 = 3.74 \simeq 4$$

أي سيكون في العينة المنشآت ذات الخمس نجوم حجمها 4 مفردات.

حجم العينة في طبقة المنشآت ذات الأربع نجوم:

$$n_2 = (N_2/N) * n = (47/334) * 50 = 7.03 \simeq 7$$

أي سيكون في العينة المنشآت ذات الأربع نجوم حجمها 7 مفردات.

حجم العينة في طبقة المنشآت ذات ثلاث نجوم:

$$n_3 = (N_3/N) * n = (66/334) * 50 = 9.88 \simeq 10$$

أي سيكون في العينة المنشآت ذات ثلاث نجوم حجمها 10 مفردات.

حجم العينة في طبقة المنشآت ذات النجمتين:

$$n_4 = (N_4/N) * n = (196/334) * 50 = 29.34 \simeq 29$$

أي سيكون في العينة المنشآت ذات النجمتين حجمها 29.

وفي النهاية نجد: $n = n_1 + n_2 + n_3 + n_4 = 4 + 7 + 10 + 29 = 50$

من خلال المثال السابق نجد أن الوزن النسبي لتمثيل طبقات العينة نفس الوزن

النسبي للمجتمع.

iii. أسلوب التوزيع الأمثل (Optimum Allocation): يقصد به تمثيل الطبقة في العينة حسب وزنها النسبي في المجتمع وذلك حسب المعادلة التالية:

$$n_r = \left(\frac{N_r * S_r}{N_1 * S_1 + N_2 * S_2 + \ldots + N_k * S_k}\right) * n$$

حيث: n_r: حجم العينة في الطبقة r.

N_r: عدد عناصر الطبقة r ضمن المجتمع.

S_r: الانحراف المعياري للطبقة r.

n: حجم العينة الكلي.

مثال 4:

باحث مهتم بدراسة رضا السياح عن الخدمات السياحية المقدمة لهم، من مجتمع يحوي 1200 مفردة، ويريد سحب عينة طبقية حجمها 150 مفردة بطريقة أسلوب التوزيع الأمثل من هذا المجتمع علماً أن حجم كل طبقة والانحراف المعياري بين مفردات كل طبقة، يبينه الجدول التالي:

الجدول رقم (3-1) توزيع أعداد السياح والانحراف المعياري في المثال حسب الجنسية

الجنسية	حجم الطبقة	الانحراف المعياري للمفردات
سوري ومغترب	500	2
عربي	400	3
أجنبي	300	2
المجموع	**1200**	

المطلوب: سحب عينة عشوائية حجمها (150) مفردة وبطريقة العينة العشوائية الطبقية بأسلوب التوزيع الأمثل.

نوجد الحجم النسبي الأمثل للعينة وفق كل طبقة حيث:

$$N = 1200$$
$$N_1 = 500 \quad S_1 = 2$$
$$N_2 = 400 \quad S_2 = 3$$
$$N_3 = 300 \quad S_3 = 2$$
$$n = 150$$

حجم العينة في طبقة السائح السوري والمغترب:

$$n_1 = \left(\frac{N_1 * S_1}{N_1 * S_1 + N_2 * S_2 + N_3 * S_3}\right) * n$$

$$= \left(\frac{500 * 2}{500 * 2 + 400 * 3 + 300 * 2}\right) * 150 = 53.57 \simeq 54$$

أي سيكون في العينة السوري والمغترب حجمها 54 مفردة.

حجم العينة في طبقة السائح العربي:

$$n_2 = \left(\frac{N_2 * S_2}{N_1 * S_1 + N_2 * S_2 + N_3 * S_3}\right) * n$$

$$= \left(\frac{400 * 3}{500 * 2 + 400 * 3 + 300 * 2}\right) * 150 = 64.28 \simeq 64$$

أي سيكون في العينة عربي حجمها 64 مفردة.

حجم العينة في طبقة السائح الأجنبي:

$$n_3 = \left(\frac{N_3 * S_3}{N_1 * S_1 + N_2 * S_2 + N_3 * S_3}\right) * n$$

$$= \left(\frac{300 * 2}{500 * 2 + 400 * 3 + 300 * 2}\right) * 150 = 32.14 \simeq 32$$

أي سيكون في العينة أجنبي حجمها 32 مفردة.

وفي النهاية نجد: $n = n_1 + n_2 + n_3 = 54 + 64 + 32 = 150$

ومن **مزايا** العينة المسحوبة بطريقة العينات الطبقية الدقة العالية في النتائج لأن المقاييس

المحسوبة من خلالها تقترب كثيراً من مقاييس المجتمع المدروس، كما أن العينات المسحوبة بهذا الأسلوب تعد أكثر تمثيلاً لمجتمعاتها بسبب التجانس الداخلي بين طبقاتها وحجم الخطأ فيها أقل بكثير من حجم الخطأ لعينة عشوائية بسيطة من نفس الحجم.

ومن أهم **عيوب** العينة العشوائية الطبقية أنه من الصعب الحصول عليها في حالة المجتمعات الإحصائية كبيرة الحجم وخاصة إذا كانت مترامية الأطراف لأنها تتطلب من الباحث معرفة بخصائص مجتمع البحث قبل اختيار العينة، كما أن التصنيف إلى عدة طبقات قد يؤدي إلى حدوث خطأ في التصنيف.

4.1.4. العينة العشوائية المتعددة المراحل (Multi- Stage Sample):

وجدنا من خلال دراستنا للعينات العشوائية البسيطة والمنتظمة والطبقية أنه من عيوب استخدامها إذا كان حجم المجتمع المدروس كبيراً وخاصة إذا كان أفراد المجتمع منتشرين على مساحة جغرافية واسعة، لذلك نستخدم العينة العشوائية المتعددة المراحل حيث يلاءم هذا النوع من العينات العشوائية دراسة الحالات التي يكون المجتمع المدروس كبيراً، كما تدعى بالعينة المساحية (Area Sample) وذات أهمية كبيرة لإمكانية الحصول على عينات تمثل المناطق الجغرافية المختلفة، ويتم اختيار عناصر العينة العشوائية كما يلي: نقسم المجتمع الأصلي إلى وحداته الأولية، في المرحلة الأولى نختار عشوائياً عدداً من الوحدات الأولية، وفي المرحلة الثانية نختار من الوحدات السابقة المختارة بشكل عشوائي عدداً من الوحدات الفرعية أيضاً بشكل عشوائي وهكذا تتعاقب مراحل المعاينة حتى تتشكل عناصر العينة، لذلك سميت بهذا الاسم لأن اختيار عناصر العينة يتم على عدة مراحل.

مثال 5:

في بحث دراسة مستوى المعيشة للأسر التي تعيش في القرى ضمن الجمهورية العربية السورية **والمطلوب** سحب عينة عشوائية متعددة المراحل.

نعلم أن جميع المحافظات السورية تحوي قرى ماعدا محافظة دمشق، لذلك في المرحلة الأولى تم اختيار عشوائي لأربع محافظات من أصل 12 محافظة وهي (السويداء – حمص – اللاذقية – دير الزور) حيث عدد القرى في هذه المحافظات على الترتيب (132 – 501 – 438 – 86)، في المرحلة الثانية نختار بشكل عشوائي عدداً من هذه القرى ولنفرض أننا اخترنا من هذه القرى في السويداء (13) في حمص (50) ومن اللاذقية (44) وفي دير الزور (9)قرى، في المرحلة الثالثة يتم اختيار أسر من هذه القرى بشكل عشوائي، وهكذا تتكون عناصر العينة بطريقة العينة العشوائية متعددة المراحل، كما يبينه المخطط التالي:

المخطط رقم (3-1) يبين مراحل اختيار العينة العشوائية المتعددة المراحل

وتمتاز هذه العينات بتوفير كثير من المال والجهد، ويعاب عليها الوقت الطويل في تحديد عناصر العينة، وزيادة أخطاء المعاينة بسبب المراحل المتعددة في اختيار عناصر العينة.

5.1.4. العينة العشوائية العنقودية (Cluster Random Sample):

وجدنا في العينات العشوائية البسيطة والمنتظمة والطبقية أن وحدة المعاينة هي أفراد ولكن في المعاينة العنقودية تكون وحدة المعاينة هي مجموعات وندعو هذه المجموعات بالعناقيد، ولا تختلف خطوات المعاينة العنقودية كثيراً عن خطوات المعاينة العشوائية البسيطة أو المنتظمة أو المتعددة المراحل سوى بوحدة المعاينة وهذه الطريقة في المعاينة تستخدم من أجل المجتمعات الكبيرة. حيث نقوم بتقسيم مجتمع الدراسة إلى عناقيد تتماثل من حيث الخصائص ثم نقوم باختيار عشوائي لعدد من العناقيد وتكون عناصر العينة هم الأفراد ضمن العنقود المختار، كما يبينه الشكل التالي:

المخطط رقم (3-2) يبين مراحل اختيار العينة العشوائية العنقودية

2.4. العينات غير الاحتمالية:

غالباً ما يتطلب اختيار عناصر عينة عشوائية إمكانيات مادية وفنية أو يجد الباحث صعوبة في وضع إطار للمجتمع المدروس في هذه الحالة يضطر الباحث إلى اتباع أسلوب الاختيار الشخصي لعناصر العينة، وهذا الأسلوب

يدعى المعاينة غير الاحتمالية وتعرف بأنها المعاينة التي يتم اختيار مفرداتها بطريقة غير عشوائية فهي طريقة شخصية وغير موضوعية، حيث يقوم الباحث باختيار مفردات العينة بالصورة التي تحقق الهدف من المعاينة، مثل اختيار عينة من السياح الأجانب حاملي الجنسية الفرنسية.

وتستخدم المعاينة غير الاحتمالية في الأبحاث الاستطلاعية ومن أجل محاولة تصور تقريبي عن خصائص ظاهرة ما لا يمكن تعميم نتائجها على المجتمع لأنها فقدت شرط العشوائية، كما يستخدمها الباحثون في الاختبار القبلي للاستبيان قبل البدء في المعاينة.

وأهم أنواع العينات غير الاحتمالية:

1.2.4. العينة الميسرة (المتيسرة) (Convenience Sample):

كما وتدعى عينة المصادفة، حيث يتم اختيار عناصر هذه العينة كما يفهم من الاسم أي جمع معلومات من أفراد المجتمع الإحصائي المتيسرين للدراسة كأن يقوم الباحث بدراسة عينة من الطلاب في المدرسة المجاورة لمنزله، أو يقوم بسؤال أشخاص يقابلهم بالطريق قبل غيرهم، وهذا ما يحدث مع الصحفيين الذين يقابلون أفراداً بالشارع ويتحدثون معهم ويعتبرون أنهم اختاروا عينة عشوائية من المواطنين، ومن ميزات هذه الطريقة أنها سريعة التنفيذ وقليلة التكلفة، أما عيوبها أنها تؤدي إلى الحصول على عينة غير سليمة وغير ممثلة للمجتمع وبالتالي لا يمكن تعميم نتائجها، وغالباً ما تخدم هذه العينة كدراسة أولية أكثر من كونها دراسة هيكلية متكاملة.

2.2.4. العينة التحكمية (Judgmental Sample):

وتدعى العينة الغرضية أو العمدية أو الهادفة (Purposive Sample) وتستخدم للحصول على معلومات من شريحة محددة من الأشخاص القادرين على توفير المعلومات المطلوبة بسبب موقعهم أو لأن المعايير التي يضعها الباحث تتوافر فيهم، كما وتستخدم هذا النوع من العينات للحصول على عينة من مجتمع يصعب الوصول إليه، وللتعرف على أنواع معينة من الحالات نريد منها معلومات خاصة لدراستها دراسة معمقة.

ومن مزايا هذه العينة أنها تضمن وحدات معاينة مناسبة للدراسة ضمن العينة المختارة، كما أنها غير مكلفة وسهلة الحصول عليها ومن عيوبها أنها غير ممثلة لمجتمع الدراسة بناءً على نظرية الاحتمالات.

ومن الأمثلة التي نستخدم فيها مثل هذه العينات: عند دراسة فعالية برنامج سياحي فإننا نتوجه للأشخاص المختصين بهذا المجال إذ يكون لديهم معرفة متخصصة في ذلك الموضوع نتيجة الخبرة.

إذا أردنا دراسة بعض الأطفال المتخلفين عقلياً فإننا نلجأ إلى دور الرعاية المتخصصة للحصول على عينة من الأطفال المتخلفين عقلياً.

3.2.4. العينة الحصصية (Quota Sample):

تستخدم المعاينة بطريقة العينة الحصصية للحصول على عينة مماثلة في خصائصها لخصائص المجتمع المدروس حيث تمر هذه المعاينة بعدة مراحل: ففي المرحلة الأولى نقوم بدراسة المجتمع الأصلي وتصنيفه حسب متغيرات الدراسة، تليها المرحلة الثانية بتحديد نسب هذه المتغيرات في المجتمع الأصلي، أما في المرحلة الأخيرة نحدد عدد (حصة) الأفراد الداخلين في الدراسة بناءً على حجم العينة، ومن خلال مراحل المعاينة بطريقة العينة الحصصية نلاحظ تشابهاً

بينها وبين العينات الطبقية ولكن الاختيار هنا لمفردات العينة يتم بشكل غير عشوائي.

وتتميز العينة الحصصية بأن اختيار مفرداتها يتم بكلفة أقل وجهد أقل كما يمكننا من الحصول على بيانات أولية وسريعة تمثل المجتمع المدروس ولكن يؤخذ على هذا النوع من العينات أنها غالباً ما تكون متحيزة لأن الباحث يقوم بتصنيفها حسب رغبته وخبرته الشخصية وعند استخدامها لا يمكننا التحكم بعدد المفردات التي تمثل حصة كل متغيرات الدراسة بالإضافة لكونها عينة غير عشوائية فلا يمكننا تعميم نتائجها على مجتمع الدراسة ككل.

4.2.4. عينة كرة الثلج (Snowball Sample):

تستخدم عينة كرة الثلج عندما نواجه صعوبة في تحديد أعضاء المجتمع المرغوب دراسته، حيث يبدأ الباحث بعينة صغيرة ميسرة، ثم تبدأ العينة بالكبر شيئاً فشيئاً مع سير الدراسة.

وفي هذه الحالة نحتاج إلى الخطوات التالية:

- الاتصال بواحد أو اثنين من حالات المجتمع المرغوب دراسته.
- سؤال هؤلاء لتحديد حالات أخرى يمكن الرجوع إليها لتوافر المعلومات لديهم.
- سؤال الحالات الجديدة لتحديد حالات أخرى جديدة، وهكذا.
- التوقف عندما لا نستطيع الوصول إلى حالات جديدة، أو الوصول إلى حجم عينة مقبول.

مثال 6:

إذا أراد مجموعة من الطلاب أن يدرسوا تأسيس كلية السياحة في جامعة دمشق عن طريق المقابلات مع الأساتذة الذين عايشوا الحدث، ولذلك يقومون بتحديد واحد أو اثنين من هؤلاء الأساتذة والاتصال بهم، ثم يقوم بالاستدلال منهم على أفراد آخرين وهكذا حتى لا يستطيعون الوصول إلى أفراد جدد، أو يكونوا قد استوفوا البيانات المرغوب جمعها.

5. تحديد حجم العينة:

كيفية تحديد حجم العينة سؤال يشغل تفكير الباحثين في كل مكان، حيث إن تحديد حجم العينة المناسب من أهم قرارات الباحث للحصول على بيانات تزوده بمعلومات يمكن الاعتماد عليها لتعميم النتائج، وتتفق آراء كثير من الباحثين على أن حجم العينة يتوقف على مجموعة من العوامل منها:

- طبيعة المجتمع المدروس: ويقصد به مدى التجانس والتباين بين أفراد مجتمع الدراسة، فإذا كان هناك تباين كبير بين أفراد المجتمع تطلب ذلك زيادة حجم العينة.
- نوع العينة المستخدم والهدف من البيانات.
- درجة الثقة المطلوبة: فكلما زادت درجة الثقة المطلوبة، زاد حجم العينة اللازم.
- حدود الخطأ الذي يمكن أن يتوقعه الباحث أو يحدده عن طريق خصائص الظاهرة في المجتمع المدروس.
- الإمكانيات المادية والزمن اللازم لإنجاز الدراسة.
- مدى التفصيل المطلوب في نتائج العينة كتقديرات لخصائص المجتمع فكلما زادت درجة التفصيل المطلوبة زاد حجم العينة المسحوبة.
- يجب أن يتناسب حجم العينة مع حجم المجتمع الأصلي.

وهناك ثلاثة أساليب لتقدير حجم العينة:

1.5. الخبرة السابقة:

يعتمد على خبرات الباحث وخبرات الباحثين الآخرين في هذا المجال، ووفقاً لهذا الأسلوب فإننا نجد أن الإحصائيين ينصحون بأن يكون الحد الأدنى المقبول لعدد الأفراد كما يلي:

- في الدراسات الوصفية الحد الأدنى المقبول 100 فرد.
- في الدراسات الارتباطية على الأقل 50 فرداً.
- وفي الدراسات التجريبية 30 فرداً لكل مجموعة، وفي بعض الأحيان يمكن أن تتألف كل مجموعة من 15 فرداً.
- أما في الدراسات المسحية فالعدد المناسب 20% من أفراد المجتمع الكلي إذا كان حجم المجتمع من500 إلى 1000فرد وتنخفض النسبة إلى 5% في المجتمعات الكبيرة جداً.

ويتميز هذا الأسلوب بالسهولة كما أنه يفيد بعض الباحثين قليلي الخبرة في مجال العمل الإحصائي، ويفضل استخدامه عند الاعتماد على العينات غير الاحتمالية.

2.5. الطرق الإحصائية:

يتطلب هذا الأسلوب من الباحث الإلمام بقدر وافر من المعلومات الإحصائية حتى يستطيع استخدام الأساليب الإحصائية في تقدير حجم العينة الأمثل.

وتعتمد الطرق الإحصائية على تحديد العوامل التي يتوقف عليها حجم العينة واعتبارها دلائل رئيسية أو مؤشرات أساسية لهذا الغرض، وهو أمر لا تأخذ فيه الطريقة الأولى، كما تعتمد هذه الطرق على توافر بعض المعلومات عن معالم المجتمع الأصلي عن طريق العينة الاستطلاعية (التي يكون حجمها عادة 30 مفردة)، وتتمثل العوامل الرئيسية المحددة لحجم العينة في:

- أكبر خطأ للتقدير يسمح به.
- درجة الثقة المطلوبة.
- نوع العينة المستخدمة.

ومن هذه الطرق:

1.2.5. تحديد حجم العينة عند تقدير نسبة ظاهرة ما في المجتمع:

هناك العديد من المعادلات الإحصائية التي نستطيع من خلالها تحديد حجم العينة نذكر منها:

a) معادلة ستيفن ثامبسون:

$$n = \frac{N \times p(1-p)}{\left[\left[N-1 \times \left(d^2 \div z^2\right)\right] + p(1-p)\right]}$$

b) معادلة روبرت ماسون:

$$n = \frac{N}{\left[\left(S^2 \times (N-1)\right) \div pq\right] + 1}$$

c) معادلة ريتشارد جيجر:

$$n = \frac{\left(\frac{z}{d}\right)^2 \times P^2}{1 + \frac{1}{N} \times \left[\left(\frac{z}{d}\right)^2 \times P^2 - 1\right]}$$

d) معادلة هيربرت أركن:

$$n = \frac{p(1-p)}{(d \div t) + \left[p(1-p) \div N\right]}$$

والصيغة الأعم:

$$n = \frac{Z^2 * p(1-p)}{(SE)^2}$$

حيث:

N: حجم المجتمع المدروس.

t: الدرجة المعيارية.

Z: القيمة الجدولية المستخرجة من جداول التوزيع الطبيعي بمستوى ثقة معين وتقدر كما يلي:

$\alpha = 0.05$	$Z = 1.96$
$\alpha = 0.001$	$Z = 2.58$

P: نسبة حدوث الظاهرة التي نهتم بها في المجتمع المدروس ونقوم بتقديرها من العينة الاستطلاعية أو نستعيض عنها بالقيمة 0.5 التي تعطي أكبر حجم ممكن للعينة.

q: نسبة عدم توافر الخاصة في المجتمع المدروس ($q = 1 - p$).

d: نسبة الخطأ وتساوي قيمة α.

S: حاصل قسمة الدرجة المعيارية المقابلة لمستوى الدلالة $1-\alpha$ على α أي $S = \frac{Z}{\alpha}$.

SE: أكبر خطأ يسمح به عند تقدير نسبة حدوث الظاهرة وتقدر عادة بقيمة ما بين (0.01 – 0.05)

وبعد حساب قيمة n تحسب قيمة كسر المعاينة التالي $\frac{n}{N}$ وهنا نميز حالتين:

- إذا كانت قيمة كسر المعاينة أصغر من 0.05 يكون حجم العينة n .

- إذا كانت قيمة كسر المعاينة أكبر من 0.05 يصبح حجم العينة n المبدئي ويرمز له بالرمز n_0 ويكون حجم العينة النهائي: $n = \frac{n_0}{1+(\frac{n_0}{N})}$.

مثال 7:

في دراسة عن استخدام وسائل تنظيم الأسرة بين السيدات في مدينة ما، وجد من خلال عينة استطلاعية أن نسبة السيدات اللاتي يستخدمن وسائل تنظيم الأسرة تساوي (0.83) وبفرض أن أكبر خطأ للتقدير مسموح به هو 0.05 وبدرجة ثقة 0.95 عين حجم العينة **المطلوب** إذا علمت أن عدد السيدات في المدينة ككل بلغ N = 4079 .

الحل:

لدينا المعطيات التالية:

$$\left.\begin{array}{l} P = 0.83 \\ 1-P = 0.17 \end{array}\right\} \Rightarrow P.(1-P) = 0.1411$$
$$SE = 0.05 \quad \Rightarrow (SE)^2 = 0.0025$$
$$1-\alpha = 0.95 \quad \Rightarrow Z = 1.96$$

نطبق المعادلة التالية:

$$n = \frac{Z^2 * p(1-p)}{(SE)^2} = \frac{(1.96)^2 * 0.1411}{0.0025} = 216.82$$

n يقرب دائماً للأعلى فيصبح n = 217 .

نحسب كسر المعاينة:

$$\frac{n}{N} = \frac{217}{4079} = 0.053$$

أكبر من 0.05 فيصبح حجم العينة مبدئي(n_0 = 217) ويكون حجم العينة:

$$n = \frac{n_0}{1+(\frac{n_0}{N})} = \frac{217}{1+(\frac{217}{4079})} = 206.078 \simeq 206$$

وعند توزيع الاستبيانات علينا أن نوزع 227 استبيان بزيادة عن حجم العينة بمقدار (10%) لتلافي عدم الاستجابة.

مثال 8:

من عينة استطلاعية لدراسة مدى تأثير البرامج التلفزيونية في زيادة الوعي بالقضايا السياحية بمدينة ما. وجد أن النسبة المئوية لمتابعي هذه البرامج 25%، **المطلوب** تحديد الحجم الأمثل للعينة التي يمكن عن طريقها دراسة هذا التأثير عند مستوى ثقة 95% وبحيث لا يزيد الخطأ عن 0.05.

الحل:

لدينا المعطيات التالية:

$$\left.\begin{array}{l} P = 0.25 \\ 1-P = 0.75 \end{array}\right\} \Rightarrow P.(1-P) = 0.1875$$
$$SE = 0.05 \quad \Rightarrow (SE)^2 = 0.0025$$
$$1-\alpha = 0.95 \Rightarrow Z = 1.96$$

نطبق المعادلة التالية:

$$n = \frac{Z^2 * p(1-p)}{(SE)^2} = \frac{(1.96)^2 * 0.1875}{0.0025} = 288.12 \simeq 289$$

بما أن حجم المجتمع غير معلوم فنعتبر n الحجم النهائي.

وعند توزيع الاستبيانات علينا أن نوزع 317 استبياناً بزيادة عن حجم العينة بمقدار (10%) لتلافي عدم الاستجابة.

2.2.5. تحديد حجم العينة عند تقدير متوسط ظاهرة ما في المجتمع:

لتحديد حجم العينة عند تقدير المتوسط، يتطلب منا معرفة كل من تباين الظاهرة المدروسة في المجتمع ودرجة الثقة المرغوبة بالبيانات وأكبر خطأ للتقدير الذي يسمح به عند حدوث الظاهرة، ويحسب حجم العينة من القانون التالي:

$$n = \frac{\sigma_x^2 * Z^2}{(SE)^2}$$

حيث:

Z: القيمة الجدولية المستخرجة من جداول التوزيع الطبيعي بمستوى ثقة معين وتقدر كما يلي:

$\alpha = 0.05$	$Z = 1.96$
$\alpha = 0.001$	$Z = 2.58$

SE: أكبر خطأ يسمح به عند تقدير نسبة حدوث الظاهرة وتقدر عادة بقيمة ما بين (0.01 – 0.05).

σ_x : الانحراف المعياري للعينة.

3.5. الطريقة الجدولية:

وجدنا أن المعادلات السابقة كانت تناقش عندما يكون هناك متغير تابع واحد فقط، ولكن غالباً ما تكون مشكلة البحث أو فرضيات البحث تحوي على عدة متغيرات وقد قدم كل من العالمين كريجيسي و مورجان (Kerjecie and Morgan 1970) جدولاً لتحديد حجم العينة حيث يتضمن الجدول عمودين الأول لعدد عناصر المجتمع والعمود الثاني يمثل العدد المقابل لعدد عناصر العينة.

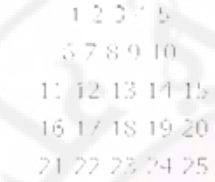
كما يبينه الجدول التالي:

N	n	N	n	N	n
10	10	220	140	1200	291
15	14	230	144	1300	297
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	169	2000	322
55	48	320	175	2200	327
60	52	340	181	2400	331
65	56	360	186	2600	335
70	59	380	191	2800	338
75	63	400	196	3000	341
80	66	420	201	3500	346
85	70	440	205	4000	351
90	73	460	210	4500	354
95	76	480	214	5000	357
100	80	500	217	6000	361
110	86	550	226	7000	364
120	92	600	234	8000	367
130	97	650	242	9000	368
140	103	700	248	10000	370
150	108	750	254	15000	375
160	113	800	260	20000	377
170	118	850	265	30000	379
180	123	900	269	40000	380
190	127	950	274	50000	381
200	132	1000	278	75000	382
210	136	1100	285	100000	384

حيث (N) تدل على حجم المجتمع، و(n) تدل على حجم العينة.

أسئلة وتمارين الفصل الثالث

1. حدد طريقة اختيار العينة العشوائية حسب الأشكال التالية:



2. اختر الإجابة الصحيحة:

- عندما يبدأ الباحث بعينة صغيرة ثم تبدأ بالكبر مع سير الدراسة، تكون العينة:
  أ. عينة كرة الثلج.
  ب. عينة عشوائية منتظمة.
  ج. عينة ميسرة.
  د. عينة عنقودية.
- العينة التي يتساوى أفرادها بنفس الفرصة والاختيار ليكون أحد أفرادها هي:
  أ. العينة العشوائية البسيطة.
  ب. العينة المنتظمة.
  ج. العينة الطبقية.
  د. العينة العنقودية.

3. قارن بين العينة العشوائية والعينة غير العشوائية من حيث استخدام كل منها؟

4. اذكر الخطوات اللازم اتّباعها لسحب عينة عشوائية عنقودية من طلاب قسم الإدارة السياحية في السنة الرابعة نريد سحب عينة عشوائية منتظمة حجمها 10 طلاب من طلاب القسم البالغ عددهم 50 طالباً، حدد العناصر التي يجب سحبها من المجتمع والتي تشكل العينة المطلوبة، إذا كان الرقم العشوائي الأول المسحوب 3.
5. إذا أردنا سحب عشوائية طبقية حجمها 60 طالباً من طلاب السنة الرابعة في كلية السياحة في جامعة دمشق البالغ عددهم 300 طالب ومكون من 3 طبقات، ما هو عدد الطلاب في كل طبقة في الحالتين التاليتين:
   - إذا كانت عينة عشوائية طبقية بطريقة التوزيع المتساوي.
   - إذا كانت عينة عشوائية طبقية بطريقة التوزيع المتناسب.
6. من عينة استطلاعية تتكون من 30 ناخباً وجد أن 12 ناخباً سينتخبون مرشح الحزب A، عند مستوى ثقة 95% وبحيث لا يزيد الخطأ عن 0.01 عين حجم العينة المطلوبة من الناخبين.

وقل رب زدني علما
جامعة دمشق
Damascus University

الفصل الرابع CHAPTER 4

التعرّف على بيئة البرنامج الإحصائي SPSS

أهداف الفصل الرابع:

1. التعرف على كيفية تشغيل البرنامج الإحصائي (SPSS).
2. التعرف على شاشات البرنامج الإحصائي (SPSS).
3. التعرف على ملفات البرنامج الإحصائي(SPSS).
4. التعرف على القوائم الرئيسية في البرنامج الإحصائي(SPSS).
5. التعرف على شريط الأدوات وكيفية تعديله في البرنامج الإحصائي(SPSS).
6. التعرف على كيفية استيراد وتصدير البيانات في البرنامج الإحصائي(SPSS).
7. التعرف على كيفية فتح ملف بيانات والخروج من البرنامج الإحصائي(SPSS).

يُعتبر البرنامج الإحصائي SPSS أحد البرامج الإحصائية التي لاقت شيوعاً في استخدامها من قبل الباحثين للقيام بالتحليلات الإحصائية، ويُستخدم الكثير من الباحثين البرنامج الإحصائي في ميادين العلوم الاقتصادية والتربوية والاجتماعية وغيرها من العلوم للقيام بإجراء التحليلات الإحصائية لبياناتهم المختلفة بهدف رسم المخططات البيانية وإيجاد المؤشرات الإحصائية المطلوبة لأبحاثهم كحساب مقاييس النزعة المركزية ومقاييس التشتت وحساب معاملات الارتباط ومقارنة المتوسطات....الخ، والقيام بمثل هذه التحليلات الإحصائية بالطرائق اليدوية ليس سهلاً، وخاصة إذا كان حجم البيانات كبيراً، وعلى كل حال لم تَعد هناك مشكلة مع تطور أجهزة الحاسوب وتصميم أنظمة خاصة، مثل (SPSS) الذي يمثل الأحرف الأولى من العبارة (**S**tatistical **P**ackage **F**or **S**ocial **S**ciences).

1. تشغيل البرنامج الإحصائي (SPSS):

هناك عدة طرق لتشغيل البرنامج الإحصائي(SPSS):

- من خلال النقر المزدوج على أيقونة البرنامج التي تظهر على سطح المكتب.
- أو من خلال قائمة ابدأ (Start) نتبع الخطوات التالية:

1. انقر فوق زر ابدأ (Start)، اختر كافة البرامج (PROGRAMSALL)
2. ثم اختر (SPSS Inc)، ثم (PASW Statistics 18.0).
3. انقر فوق أيقونة (PASW SPSS Statistics 18.0).

2. شاشات البرنامج الإحصائي (SPSS):

يحتوي البرنامج الإحصائي SPSS على ثلاث شاشات رئيسية هي:

2.1. شاشة محرر البيانات (Data Editor Windows):

وهي الشاشة التي تحوي البيانات الإحصائية المراد تحليلها، وهذه الشاشة يتم فتحها تلقائياً عند تشغيل البرنامج الإحصائي (SPSS)، وتتكون هذه الشاشة من ورقتين

تشبهان ورقة العمل في برنامج الجداول الإلكترونية (Microsoft Excel) حيث تتكون كل ورقة من أعمدة وصفوف، فالورقة الأولى ورقة عرض البيانات (Data View) ولهذه الورقة مهمة في إدخال وتعديل وعرض البيانات، وتمثل الأعمدة في هذه الورقة المتغيرات في حين تمثل الصفوف الحالات التي تدرس، وبذلك تمثل كل خلية مشاهدة المتغير للحالة المقابلة، أما الورقة الثانية فهي ورقة عرض المتغيرات (Variable View) التي تستخدم لعرض وتعريف المتغيرات، ويمكن التنقل بين هاتين الصفحتين بالنقر على اسم الصفحة المراد الانتقال إليها في أسفل الزاوية اليسرى من شاشة محرر البيانات أو باستخدام الاختصار (Ctrl+T).

الشكل رقم (4-1) يبين شاشة محرر البيانات في البرنامج الإحصائي SPSS

2.2. شاشة المخرجات (Output Viewer):

وهي الشاشة التي تظهر فيها نتائج الإجراءات الإحصائية كالرسومات البيانية والجداول التكرارية ونتائج الاختبارات الإحصائية.

الشكل رقم (4-2) يبين شاشة المخرجات في البرنامج الإحصائي SPSS

3.2. شاشة محرر التعليمات (Syntax Editor):

وهي الشاشة التي تتم من خلالها كتابة التعليمات (البرنامج) للعمليات المختلفة، وهذه التعليمات يمكن تخزينها وتعديلها وتنفيذها في أي وقت، كما أنه في النسخة 17 من البرنامج وما بعدها أصبحت تظهر هذه التعليمات ضمن شاشة المخرجات.

الشكل رقم (4-3) يبين شاشة محرر التعليمات في البرنامج الإحصائي SPSS

3. ملفات البرنامج الإحصائي(SPSS):

يتعامل البرنامج الإحصائي(SPSS) مع مجموعة الملفات التي يخرجها البرنامج حسب المعلومات الموجودة فيها، وهناك ثلاثة أنواع مهمة من هذه الملفات تستخدم دائماً، وهي:

1.3. ملفات البيانات(Data Files):

وهي الملفات التي تحوي على البيانات التي تظهر في شاشة محرر البيانات (Data Editor) ويميز هذه الملفات اسمها الذي ينتهي دائماً باللاحقة (.sav) فأي ملف له لاحقة (.sav) يحتوي على بيانات خام، ويتم فتح هذا النوع من الملفات من خلال شاشة محرر البيانات (Data Editor).

2.3. ملفات المخرجات الإحصائية (نتائج الإجراءات الإحصائيةOutput Viewer):

وهو الملف الذي يحتوي على نتائج الإجراءات الإحصائية التي تظهر في شاشة المخرجات ويميزه اسمه الذي ينتهي دائماً باللاحقة(.SPO) فأي ملف له لاحقة(.SPO) يحتوي على نتائج إجراءات إحصائية معينة ويتم فتح هذا النوع من الملفات من خلال شاشة عرض المخرجات(Output Viewer).

3.3. ملفات التعليمات (Syntax):

وهو الملف الذي يحوي على التعليمات المراد إجراؤها كالإجراءات الإحصائية مثلاً ويميز هذا الملف لاحقة (.SPS) فأي ملف له لاحقة(.SPS) هو ملف تعليمات، ويتم فتح هذا النوع من الملفات من خلال شاشة محرر التعليمات (Syntax Editor).

4. القوائم الرئيسية في البرنامج الإحصائي(SPSS):

تمثل القوائم (Menus) المفاتيح الأساسية للقيام بأي عملية في أنظمة النوافذ، ويزودنا البرنامج الإحصائي(SPSS) بإحدى عشرة قائمة رئيسية (تتخللها قوائم فرعية)

تستطيع من خلالها القيام بجميع العمليات التي يوفرها البرنامج الإحصائي وهذه القوائم هي:

الشكل رقم (4-4) يبين القوائم الرئيسية في البرنامج الإحصائي SPSS

1.4. قائمة ملف (File):

إن الهدف الرئيسي لهذه القائمة التعامل مع الملفات من حيث إنشاء ملف جديد، أو فتح ملف مخزن مسبقاً، أو حفظ ملف، أو طباعة ملف، أو التعرف على المتغيرات الموجودة في ملف البيانات، كذلك فإن قائمة الملف تعرض قائمة بآخر الملفات التي تم استخدامها وكذلك الخروج من البرنامج الإحصائي(SPSS).

2.4. قائمة تحرير(Edit):

تحتوي هذه القائمة على كثير من الأدوات المهمة فهي تستخدم لعمليات النسخ واللصق والقص ونقل البيانات من مكان إلى آخر، والبحث عن حالات مهمة.

3.4. قائمة عرض (View):

يمكن باستخدام قائمة العرض إظهار وإخفاء شريط الأدوات (الأيقونات المختصرة المناسبة) Toolbar التي يمكن استخدامها بدل البحث عن الأوامر ضمن القوائم (سنتحدث عن هذه الأيقونات لاحقاً) وكذلك نستطيع من خلال هذه القائمة إظهار أو إخفاء خطوط الشبكة وتغيير نوع الخط المستخدم، وإظهار أو إخفاء عناوين القيم.

4.4. قائمة بيانات (Data):

تحتوي قائمة البيانات على العديد من الأدوات المهمة التي تسمح بتعريف المتغيرات وتغيير أسمائها، وكذلك القيام بالعمليات المختلفة على البيانات من إدراج وفرز

وتحويل ودمج مع بيانات أخرى وتجميع الحالات، وغير ذلك من عمليات تقسيم الملفات وتحديد ووزن الحالات.

5.4. قائمة التحويلات(Transform):

نستطيع من خلال هذه القائمة القيام بالعمليات الحسابية المختلفة مثل استخدام الدوال الإحصائية التي يزودنا بها البرنامج الإحصائي (SPSS) وإنشاء متغيرات جديدة وإعادة ترميز البيانات وتحديد الرتب واستبدال القيم المفقودة وعمليات إنشاء قيم عشوائية وغيرها.

6.4. قائمة الإجراءات الإحصائية (Analyze):

تهتم هذه القائمة بالتحليلات الإحصائية الكثيرة، وتحتوي على جميع أدوات التحليلات الإحصائية العادية والمتقدمة كإجراءات الإحصاء الوصفي والإحصاء الاستدلالي والتنبؤ لذلك تعد هذه القائمة من أهم قوائم البرنامج الإحصائي (SPSS).

7.4. قائمة الرسومات (Graphs):

نستطيع عن طريق هذه القائمة تمثيل البيانات بيانياً بأشكال مختلفة لتلاءم التحليل والنتائج المطلوبة.

8.4. قائمة الأدوات (Utilities):

وهنا نستطيع إيجاد معلومات مفصلة عن الملف المستخدم والمتغيرات المدخلة ضمن الملف، وتعريف واستخدام المجموعات Sets للمتغيرات المختلفة.

9.4. قائمة إطار (Window):

نستطيع عن طريق هذه القائمة التنقل بين النوافذ المختلفة والتحكم بحجم هذه النوافذ.

10.4. قائمة المساعدة (Help):

تزودنا هذه القائمة ضمن البرنامج الإحصائي بمساعدة تفاعلية، نستطيع من خلالها الحصول على إجابات وإيضاحات كثيرة للتساؤلات التي تثار عند مواجهة مشكلة ما مع البرنامج الإحصائي (SPSS).

5. شريط الأدوات (SPSS Toolbar):

الشكل رقم (4-5) يبين شريط الأدوات الرئيسية في البرنامج الإحصائي SPSS

يزودنا البرنامج الإحصائي بالإضافة للقوائم الرئيسية بشريط الأدوات الذي يحتوي على أيقونات رسومية تمثل أوامر موجودة ضمن القوائم الرئيسية، تغنينا عن استخدام القوائم الرئيسية.

كما يمكننا إضافة شريط أدوات جديد باسم يحدده المستخدم أو إجراء تعديلات على هذا الشريط سواء بإضافة أو حذف بعض الأوامر، ويتم ذلك من خلال القائمة (View)، كما تبينه الخطوات التالية:

- من قائمة View نختار الأمر Toolbars ثم نختار الأمر الفرعي Customize.
- يظهر صندوق حوار بعنوان Show Toolbars الذي له الشكل التالي:

الشكل رقم (4-6) يبين صندوق حوار Show Toolbars

والذي يحوي عدة خيارات:

- Window: لتحديد مكان تفعيل شريط الأدوات في الشاشة المرغوبة وهذه الشاشات هي: (Data Editor – Syntax – Output).
- Show ToolTips: بحذف التحديد لا يتم عرض تلميحات الأوامر (وظيفة الأمر الموجود ضمن شريط الأدوات عند تمرير مؤشر الفأرة عليه).
- Large Buttons: عند تحديده تظهر الأيقونات الرسومية بحجم كبير، وعند إزالة التحديد تظهر بحجم صغير.
- Edit: لتحرير شريط الأدوات وذلك بإضافة أو حذف إحدى الأوامر، وعند النقر عليه يظهر صندوق حوار جديد بعنوان Edit Toolbar له الشكل التالي:

الشكل رقم (4-7) يبين صندوق حوار لتحرير شريط الأدوات

ولإضافة أي أمر نحدد القائمة التي تحوي الأمر من القسم المعنون بـ Categories ثم نحدد الأمر الذي نريد إضافته من القسم المعنون بـ Tools وننقر عليه مرتين وسوف ينتقل مباشرة إلى شريط الأدوات الموجود في أسفل صندوق الحوار، وبعد الانتهاء من اختيار الأوامر المرغوب إضافتها إلى شريط الأدوات ننقر على الزر Continue فنعود إلى صندوق المعنون بـ Show Toolbars

ونضغط على OK.

- New: لإضافة شريط أدوات جديد باسم جديد، وعند النقر عليه يظهر صندوق حوار جديد بعنوان Toolbar Properties له الشكل التالي:

الشكل رقم (4-8) يبين صندوق حوار لإضافة شريط أدوات جديد

فيكتب اسم شريط الأدوات الجديد الذي نرغب بإضافته في المربع الخاص المعنون بـ Toolbar Name، أما الشاشات التي نرغب بإضافة شريط الأدوات الجديد لها نقوم بتحديدها في القسم المعنون بـ Display on the following windows.

6. استيراد وتصدير البيانات:

تعتبر عملية الحصول على البيانات من الأولويات التي تشغل بال الباحثين، ولكن ليس بالضروري أن تكون هذه البيانات مخزنة في ملفات (SPSS)، إذ قد تكون ضمن بيئة النوافذ (Windows) مثل ملفات (Excel) أو (Access) وغيرها.

كذلك، فإننا تحتاج إلى نقل البيانات التي قمنا بإدخالها إلى برنامج (SPSS) إلى تطبيقات أخرى مثل (Excel) لذلك يوفّر البرنامج الإحصائي(SPSS) إمكانية تصدير البيانات التي يتعامل بها إلى أنظمة أخرى (Exporting Data) وكذلك استيراد بيانات من أنظمة أخرى (Importing Data) ويوضح الجدول التالي أنواع الملفات التي يتعامل معها البرنامج الإحصائي(SPSS):

الجدول رقم (4-1) يبين أنواع الملفات التي يتعامل معها البرنامج الإحصائي SPSS

Application	Extension
Spss/ pc+	*.Sys
Systat	*.Syd
Systat	*.Sys
Spss portable	*.Por
Excel	*.Xls
Lotus	*.W
Slyk	*.Slk
D Base	*.Dbf
Sas long life name	*.Sas7dbat
Sas short life name	*.Sd7
Sas v6 for windows	*.Sd2
Sas v6 for unix	*.Ssd1
Sas transport	*.Xpt
Text	*.Txt
Data	*.Dat

1.6. تصدير البيانات (Exporting Data):

إذا أردنا تصدير ملف من تطبيق (SPSS) إلى الملف بتطبيق (Excel) اتبع الخطوات التالية:

1. بعد فتح الملف وإدخال البيانات ننقر على قائمة ملف (file) ثم نختار الأمر الفرعي (Save as) ليظهر مربع حوار بعنوان (save data as).

2. من المربع نحدد نوع الملف الذي نرغب بتصدير البيانات له (xls لملفات Excel)
3. انقر فوق موافق OK سوف ينشأ ملف جديد اسمه اسم الملف الذي وضعته الذي يستطيع تطبيق Excel التعرف عليه.
4. افتح تطبيق Excel.
5. انقر فوق Open من قائمة File في Excel وافتح الملف.

وهذا ما يبنه الشكل التالي:

الشكل رقم (4-9) يبين تصدير ملف من SPSS إلى Excel

2.6. استيراد البيانات (Importing Data):

نستطيع استيراد البيانات من بعض التطبيقات التي يشبه تنظيمها تنظيم برنامج (SPSS) كالموجودة في الجدول السابق إلى (SPSS) باتباع الخطوات التالية:

1. ننقر على قائمة (File) ثم نختار الأمر الفرعي(Open) ثم (Data).

2. ننقر السهم إلى يمين قائمة (File of type) ستظهر قائمة بأنواع الملفات التي يمكن لبرنامج (SPSS) التعامل معها، نختار بالنقر على نوع الملف الذي نريد فتحه.
3. نحدد المكان الموجود عليه الملف الذي نريد فتحه، وذلك باختيار المكان من قائمة Look in.
4. ننقر فوق الملف الذي نريد فتحه، ثم ننقر فوق (OK). سنجد الملف في شاشة (SPSS).

كما في الشكل التالي:

الشكل رقم (4-10) يبين استيراد ملف Excel إلى SPSS

7. فتح ملفات بيانات:

يمكننا فتح ملف بيانات الذي تم تخزينه مسبقاً لإجراء التعديلات عليه أو الاطلاع عليه أو لإجراء عمليات إحصائية جديدة، وذلك باتباع الخطوات التالية:

1) نفتح البرنامج الإحصائي (SPSS).
2) ننقر فوقOpen من قائمة File.

3) من مربع الحوار Open Data File نحدد الملف الذي نريد فتحه من قائمة الملفات فيظهر ذلك في مربع File Name، ثم نختار الأمر Open.

أو بالنقر مرتين على الملف الذي نريد فتحه.

8. الخروج من البرنامج الإحصائي:

تستطيع الخروج من البرنامج الإحصائي (SPSS) بالنقر على Exit من قائمة File.

أسئلة الفصل الرابع

1- أكتب خطوات فتح البرنامج الإحصائي SPSS من قائمة ابدأ.

2- عدد الشاشات التي يظهرها البرنامج الإحصائي SPSS، وماهي شاشة محرر التعليمات Syntax Editor.

3- اذكر أهم الملفات التي ينتجها البرنامج الإحصائي SPSS، وماهي لاحقة كل منها.

4- ما فائدة قائمة التحويلات Transform.

5- عرف شريط الأدوات Toolbar وأكتب خطوات إدراج شريط أدوات جديد.

6- ماذا تعني عملية استيراد وتصدير البيانات.

الفصل الخامس CHAPTER 5

تعريف المتغيرات وإدخال البيانات في البرنامج الإحصائي SPSS

أهداف الفصل الخامس:

1. تعريف المتغيرات ضمن البرنامج الإحصائي SPSS.
2. حفظ البيانات في البرنامج الإحصائي SPSS.
3. حفظ ملف المخرجات البرنامج الإحصائي SPSS.
4. عرض خصائص المتغيرات الموجودة في ملف البيانات.
5. التعرف على قائمة التحرير Edit.
6. إدخال البيانات في البرنامج الإحصائي SPSS.
7. التعرف على قائمة العرض View

سبق لنا أن درسنا المتغيرات وأنواعها، وسنتعلم في هذا الفصل كيفية تعريف المتغيرات ضمن البرنامج الإحصائي (SPSS) تمهيداً لإدخال البيانات، فتعريف المتغيرات تعد الخطوة الأولى في إدخال البيانات وذلك من خلال التبويب (Variable View) ضمن شاشة محرر البيانات (Data Editor).

فعند النقر على التبويب (Variable View) نلاحظ أنّ عدد الأسطر مرقمة بشكل متسلسل (1-2-3 .....) فكل سطر يعبّر عن متغير واحد في هذا التبويب، ويعبر كل عمود في التبويب (Data View) على قيم متغير واحد (بيانات ذلك المتغير)، أما الأعمدة في التبويب (Variable View) فإنها تعرّف خاصة واحدة من خصائص المتغيرات المعرفة، ويعبر كل سطر في التبويب (Data View) على حالة واحدة (استبيان واحد).

1. تعريف المتغيرات باستخدام البرنامج الإحصائي SPSS:

قبل البدء بتفريغ البيانات في البرنامج الإحصائي (SPSS) نقوم بتعريف متغيرات الدراسة وخصائصها من خلال التبويب (Variable View) ويشمل تعريف المتغيرات ذكر خصائص هذه المتغيرات، وهي:

1.1. اسم المتغير (Name):

وهو العمود الأول، حيث يتم في هذه الموقع تحديد اسم قصير لكل متغير من متغيرات الدراسة أو للأسئلة المدخلة، حيث ندخل الاسم ضمن التبويب (Variable View) ليظهر في رأس العمود في التبويب (Data View)، وعند كتابة اسم المتغير يجب مراعاة عدد من الشروط تتمثل في:

- يجب أن يبدأ اسم المتغير بحرف أبجدي، أما بقية خانات الاسم فمن الممكن أن تكون حرفاً أو رقماً أو رمزاً مثل @- $- #- أو (.) أو (-)، فمثلاً إذا تمت كتابة اسم المتغير بالشكل التالي (1education) فهذا الاسم خاطئ لكونه بدأ برقم

وبالتالي يظهر لدينا رسالة تحذير تبين أن اسم المتغير يشمل على أحد الأحرف غير المسموح بها في بداية اسم المتغير، كما يلي:

الشكل رقم (5-1) يبين رسالة الخطأ عند بدء اسم المتغير بحرف غير مسموح به

- ❖ يجب ألا ينتهي اسم المتغير بنقطة (.) أو شحطة سفلية (_) أو أي رمز خاص (؟ ...).
- ❖ يجب ألا يتجاوز طول اسم المتغير أكثر من ثماني خانات قبل الإصدار 12 أما بعده يجب ألا يتجاوز أكثر من 64 خانة.
- ❖ يجب ألا يحتوي اسم المتغير على فراغات، ويستبدل الفراغ بالشحطة السفلية(_)، فمثلاً إن كان اسم المتغير يحوي فراغاً تظهر رسالة تحذير تبين أن اسم المتغير يشمل على فراغ، كما يلي:

الشكل رقم (5-2) يبين رسالة الخطأ عند يحوي اسم المتغير على فراغ

- ❖ يجب أن يكون اسم المتغير وحيداً لا يتكرر، كما ينصح ألا يكون هناك تشابه في أسماء المتغيرات حتى لا يحدث خلط بين المتغيرات في أثناء عملية التحليل أو قراءة النتائج.
- ❖ يمكن الكتابة في البرنامج SPSS 10.0 وما بعده باللغتين العربية والأجنبية على خلاف الإصدارات السابقة للبرنامج، والتي لا تقبل الكتابة باللغة العربية.
- ❖ اسم المتغير هنا غير حساس فهو لا يميز بين أسماء المتغيرات ذات الأحرف الكبيرة والأسماء ذات الأحرف الصغيرة فكلاهما سواء بالنسبة للبرنامج ولكن تختلف فقط في الإظهار.

2.1. نوع المتغير (Type):

وهو العمود الثاني، بعد أن أدخلنا اسم المتغير يجب أن نحدد نوعه، حيث أننا ذكرنا أنواع المتغيرات سابقاً (راجع الفصل الأول)، إنّ العمود Type يتيح لنا تعريف نوع المتغير الذي نقوم بإدخال قيمه، حيث يفترض البرنامج الإحصائي (SPSS) أنّ جميع المتغيرات رقمية (Numeric) فإنه يجب علينا في بعض الأحيان تغيير نوع المتغير، وذلك بالنقر على الرمز [...] الموجود بجانب كلمة Numeric فيظهر صندوق حوار باسم (Variable Type) والذي يحتوي على أنواع المتغيرات الممكنة، كما يبينه الشكل التالي:

الشكل رقم (5-3) صندوق حوار تحديد نوع المتغير (Variable Type)

- **Numeric:** وهو المتغير الرقمي وهو النوع الافتراضي للمتغيرات في ورقة (Data View)، ويستخدم هذا الخيار لعرض البيانات على شكل أرقام، مع العلم أنه من الممكن أن تحتوي هذه البيانات على إشارة + /- أو على جزء عشري.
- **Comma:** الفاصلة، ويستخدم هذا الخيار لعرض القيم الرقمية للمتغير (كما في خيار Numeric) حيث تظهر فاصلة (,) بعد كل ثلاث خانات صحيحة ويفصل بين الخانات الصحيحة والعشرية بنقطة (.)، فمثلاً إذ قمت بإدخال الرقم (3765.76) سيظهر بالشكل (3,765.76).
- **Dot:** النقطة، ويفيد هذا الخيار لعرض القيم الرقمية للمتغير (كما في خيار Numeric) حيث تظهر نقطة (.) بعد كل ثلاث خانات صحيحة ويفصل بين الخانات الصحيحة والعشرية فاصلة (.) ، فمثلاً إذ قمت بإدخال الرقم (3765.76) سيظهر بالشكل (3.765,76)
- **Scientific Notation:** الرمز العلمي، يستخدم هذا الخيار لعرض البيانات بصيغة علمية، حيث يستخدم الحرف E بدلاً من الأساس (10) وذلك كما يلي:

$$0.0025 \rightarrow 2.5\times10^{-3} \rightarrow 2.5E-003$$

$$3500 \rightarrow 3.5\times10^{+3} \rightarrow 2.5E+003$$

❖ **Data:** التاريخ، ويعرض بيانات المتغير على شكل تاريخ و/ أو وقت، مع ضرورة الالتزام بخيارات العرض الزمني والتاريخي على يمين صندوق الحوار، كما في الشكل التالي:

الشكل رقم (5-4) صندوق حوار تحديد خيارات عرض متغير التاريخ

يمكنا أن نختار أياً منها، فإذا اخترنا الخيار (dd.mm.yyyy) فإن تاريخ العشرين من نيسان لعام ألفين وعشرة سيُعرض بالشكل 20.04.2010 ، حيث يعبّر الرمز mm عن ترتيب الشهر (month) في السنة، والرمز dd عن تاريخ اليوم (day)، والرمز yyyy عن العام (year) الذي سيتم عرضه في أربع خانات.

❖ **Dollar :** وهو متغير يستعمل كرمز للدولار الأمريكي، حيث تشمل بيانات المتغير على إشارة S وخانات عشرية وفاصلة لكل ثلاث خانات، كما يمكننا اختيار شكل عرض الرقم الذي نرغب فيه من خلال الخيارات الموجودة على يمين صندوق الحوار، كما في الشكل المجاور:

الشكل رقم (5-5) صندوق حوار تحديد شكل عرض الرقم

❖ **Custom Currency:** وهو متغير يستعمل للدلالة على العملة المحلية، حيث نلاحظ هناك خمسة خيارات يمكن تخزينها ضمن الملف الواحد وهي: (CCA, CCB ,CCC, CCD, CCE) ويتم ضبط هذا النوع من المتغيرات كما يلي:

- نحدد أحد الخيارات التي يتيحها لنا صندوق الحوار المعنون (Variable Type).
- من قائمة Edit نختار الأمر Option.
- من صندوق الحوار المعنون Option نختار التبويب Currency.
- نحدد اسم العملة التي تم اختيارها في أثناء تعريف المتغير، ثم نكتب اسم العملة التي نرغب أن تظهر وذلك قبل الرقم Prefix أو بعد الرقم Suffix ضمن مستطيل All Values، وإذا رغبنا بظهور الإشارة السالبة نضعها في أحد خيارات Negative Value ثم نضغط على OK، كما يبينه الشكل التالي:

الشكل رقم (5-6) كيفية تعريف العملة المحلية

❖ **String:** ويستعمل هذا الخيار لتعريف متغير من النوع المحرفي أو النوعي أو أرقام أو رموز، وهذا البديل يدل على أن قيم المتغير هي رموز وليست أرقاماً، والتعامل مع هذا النوع في البرنامج الإحصائي SPSS يقتصر على عدد محدود من الأوامر.

نلاحظ في جميع الخيارات السابقة تكرر معنا خياران هما Width و Decimal Places اللذان يظهران في العامودين التاليين.

3.1. العرض (Width):

وهو العمود الثالث، ويحدد عدد الخانات المدخلة للبيانات وليس للمتغيرات، بحيث تؤخذ الفواصل العشرية والإشارات والرموز بالحسبان، مع العلم أن أقصى عدد للخانات المدخلة هو 40 خانة.

4.1. الخانات العشرية (Decimal Places):

وهو العمود الرابع، ويمثل عدد المراتب العشرية للمتغيرات العددية فقط، مع العلم أن العدد الأقصى لعدد الخانات العشرية هو (16) خانة.

5.1. عنوان المتغير (Variable Label):

وهو العمود الخامس، وجدنا أن اسم المتغير يفرض على المستخدم مجموعة من القيود وفي كثير من الأحيان لا يتم كتابة اسم المتغير بشكل كامل بل يكتب بشكل مختصر وهذا الاسم يشكل لبساً للقارئ لعدم فهم الاسم القصير، لذلك يلجأ المستخدم لكتابة اسم المتغير بشكل مختصر وهذا الاسم يشكل لبساً للقارئ لعدم فهمه هذه الرموز ولذلك جاء عنوان المتغير ليعطي المتغير عنواناً يصل عدد محارفه إلى 256 محرف ولا يفرض على المستخدم القيود القاسية التي يفرضها اسم المتغير، وهو يقابل إدراج تعليق في برامج (Office) فمثلاً إذا كان اسم المتغير (Edu_level) يمكننا أن نكتب عنوان المتغير (Education Level) كما يمكن أن يظهر العنوان بدلاً من اسم المتغير في شاشة المخرجات للبرنامج الإحصائي حسب رغبة المستخدم.

6.1. عناوين قيم المتغير (Value Labels):

وهو العمود السادس، في بعض الأحيان نحدد عناوين للبيانات كون المتغير يستعمل قيماً عددية للتعبير عن قيم غير عددية مثلاً متغير الجنس (Gender) يستعمل الرقم 1 للتعبير عن الذكور (Male)، والرقم 2 للتعبير عن الإناث (Female)، ولإدخال عناوين لقيم المتغير نتبع الخطوات التالية:

- ننقر على الخلية التي تقع تحت عامود Value في سطر المتغير مثلاً Genderوذلك بالنقر على الرمز [...].
- يظهر صندوق حوار بعنوان Value Labels كما في الشكل المجاور:

الشكل رقم (5-7) صندوق حوار Value Labels

- ننقر بداخل المستطيل المعنون بـ(Value) ونضع الرقم 1 ثم نكتب في المستطيل المعنون بـ(Label) وصف للقيمة 1 وهو: (male) ثم ننقر على زر (Add) لإضافة القيمة الأولى، ومن أجل تعريف القيمة الثانية ندخل القيمة 2 في المستطيل المعنون بـ (Value) وفي المستطيل المعنون بـ(Label) نكتب (female) ثم ننقر على زر (Add)، وبعد الانتهاء من تعريف قيم المتغير ننقر الزر (OK)، كما في الشكل التالي:

الشكل رقم (5-8) كيفية تعريف قيم المتغير باستخدام صندوق حوار Value Labels

- كما يمكننا إجراء تعديلات على القيم المدخلة باستخدام الزر (Change)، بعد تحديد القيمة التي نريد تعديلها.
- أما لحذف قيمة مدخلة بأكملها نحدد القيم التي نريد حذفها ثم نضغط على زر (Remove).

7.1. القيم المفقودة (Missing):

ويمثل العمود السابع، ويستخدم هذا العمود لتعريف القيم المفقودة حيث إن التحليل الإحصائي يتطلب في بعض الأحيان تحديد بعض قيم المتغير على أنها قيم مفقودة، أي أن هذه القيم موجودة أصلاً ولكننا ليس بحاجة إلى إدخالها في التحليل الإحصائي كونها قيماً شاذة.

إن البرنامج الإحصائي يتيح لنا خيارات تعريف القيم المفقودة لمتغير معين، ويتم ذلك من خلال نقر على الرمز [...] في عمود (Missing) للمتغير، فيظهر صندوق حوار فيه ثلاثة خيارات، وهي كما يلي:

- الخيار الأول (No Missing Values): وهو الخيار الافتراضي، ويفترض أن القيم المفقودة هي البيانات التي لا يتم إدخالها في صفحة (Data View) والتي تظهر نقطة (.) في صفحة (Data View).

الشكل رقم (5-9) صندوق حوار لتحديد القيم المفقودة

- الخيار الثاني (Discrete Missing Values): يمكن هذا الخيار تحديد ثلاث قيم على أنها قيم مفقودة، مثال: القيم 90000، 120000، 150000 تمثل دخل أفراد في عينة. سوف تعدّ هذه القيم قيماً مفقودة في حالة إدخالها.

- الخيار الثالث (Discrete Missing Values Range Plus Optional): يمكن هذا الخيار تحديد مجال من القيم على أنها قيم مفقودة بالإضافة إلى إمكانية تحديد قيمة مفقودة واحدة اختيارياً، مثال: لنحدد القيم التالية في المجال 10.3 – 7.5 بالإضافة إلى تحديد قيمة مفقودة واحدة اختيارياً خارج المجال المحدد مسبقاً كالقيمة 12 مثلاً.

8.1. عرض العمود (Columns):

وهو العمود الثامن، نحدد من خلالها عرض العمود في صفحة (Data View) كما أنه يمكننا تغيير عرض العمود من خلال وضع مؤشر الماوس على الحد الفاصل بين اسمي متغيرين في صفحة (Data View) وتحريكه مع استمرار الضغط على الحد الفاصل إلى اليمين واليسار، أو من خلال زيادة العرض من نفس الخلية، ويجب ملاحظة أن عدد المحارف في هذه الخاصية يجب أن يكون أكبر أو يساوي Width بالنسبة لجميع المتغيرات وذلك لعرض جميع خانات المتغير.

9.1. المحاذاة (Align):

ويستخدم لضبط موقع البيانات داخل الخلايا التابعة لمتغير معين في صفحة (Data View)، ولها ثلاثة خيارات:

- Left: للمحاذاة على يسار الخلية.
- Right: للمحاذاة على يمين الخلية.
- Center: للمحاذاة في وسط الخلية.

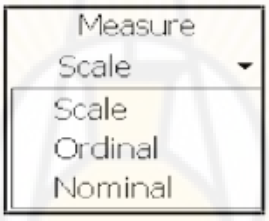
الشكل رقم (5-9) القائمة المنسدلة التي تظهر في عمود المحاذاة

10.1. مقياس المتغير (Measure):

لقد تم تصنيف المتغيرات في القسم الأول إلى عدة مقاييس، تستخدم هذه الخلية لضبط تعريف مقياس لمتغير معين وتصنيفه وفق التصنيف المتاح في البرنامج الإحصائي وهي كما يلي:

- Scale: ويستخدم للبيانات الكمية المستمرة.
- Ordinal: ويستخدم للبيانات الترتيبية.
- Nominal: ويستخدم للبيانات الاسمية.

الشكل رقم (5-9) القائمة المنسدلة التي تظهر في عمود مقياس المتغير



2. حفظ البيانات:

لحفظ البيانات المدخلة نتبع الخطوات التالية:

1) نختر الأمر (Save As) من قائمة ملف (File) لحفظ البيانات التي تخزن للمرة الأولى، أما البيانات التي تم تعديلها فيستخدم لحفظها الأمر (Save) على نفس الملف، وعندما نختار الأمر (Save As) يظهر صندوق حوار بعنوان (Save Data AS).
2) نحدد المكان الذي نريد حفظ البيانات فيه في مربع (Look In) وندخل اسم الملف في مربع (File Name).
3) نحدد النوع الذي نريد الحفظ به من مستطيل (Save as type).
4) المستطيل (Variables) يحدد لنا عدد المتغيرات التي نريد حفظها في قاعدة البيانات.

وهذا ما يبينه الشكل التالي:

الشكل رقم (5-10) خطوات حفظ ملف البيانات

ملاحظات:

- يمكن استخدام الأيقونة – من شريط الأدوات بدلاً من الأمر (Save).
- يمكن استخدام الأمر (Save As) لإنشاء نسخة إضافية من الملف العامل ولكن باسم جديد أو مكان جديد أو نوع جديد.

3. حفظ ملف المخرجات:

عند التعامل مع البرنامج الإحصائي SPSS نكون بحاجة لحفظ ملف المخرجات الذي يحوي على المخططات والجداول الإحصائية التي حصلنا عليها من الإجراءات الإحصائية، فهناك مفهوم خاطئ عند بعض المستخدمين أن حفظ ملف البيانات سيؤدي إلى حفظ ملف المخرجات وهذا خاطئ، لذلك نتبع الخطوات التالية:

1) من الأمر (Save As) من قائمة ملف (File) في شاشة المخرجات يظهر صندوق حوار بعنوان (Save Output AS).

2) نحدد المكان الذي نريد حفظ البيانات فيه في مربع (Look In) وندخل اسم الملف في مربع (File Name).
3) في المستطيل (Save as type) فلا يوجد سوى خيار واحد وهو (Viewer Files ° .spv).
4) ثم نضغط علىSave.

4. عرض خصائص المتغيرات:

يتيح البرنامج الإحصائي SPSS التعرف على خصائص المتغيرات الموجودة في قاعدة البيانات المفتوحة، وذلك من خلال الطريقتين التاليتين:

1.4. الطريقة الأولى: طريقة قائمة File:

من قائمة File ثم نختار Display Data File information ويمكننا هذا الخيار من عرض المتغيرات ضمن الملف الذي نعمل به أو التعرف على بيانات من ملف خارجي، فيظهر جدولان في شاشة المخرجات، فالجدول الأول يعرض معلومات عن جميع المتغيرات المدخلة إلى قاعدة البيانات المفتوحة، أما الجدول الثاني يعرض قيم المتغيرات مع عناوينها، كما في الأشكال التالية:

الشكل رقم (5-11) جدول يبين معلومات عن جميع المتغيرات المدخلة في قاعدة البيانات المفتوحة

الشكل رقم (5-11) جدول يبين قيم المتغيرات مع عناوينها في قاعدة البيانات المفتوحة

2.4. طريقة قائمة Utilities:

من قائمة Utilities ثم نختار الأمر الفرعي Variableفيظهر صندوق حوار معنون Variables له الشكل التالي:

الشكل رقم (5-12) صندوق حوار لعرض خصائص كل متغير من متغيرات الدراسة

وبالنقر على المتغير الذي نرغب بعرض خصائصه التي تظهر في المستطيل المعنون بـ Variable Information.

5. قائمة التحرير (EDIT):

الشكل رقم (5-13) قائمة تحرير في البرنامج الإحصائي SPSS

نستطيع من خلال هذه القائمة القيام بالكثير من المهام، وهي:

- **Undo**: للتراجع عن آخر عملية قمنا بها.
- **Redo**: وذلك لاستعادة آخر عملية قمنا بها.
- **Cut**: للقيام بالقص سواء للمتغير أم للحالة.
- **Copy**: للقيام بالنسخ سواء للمتغير أم للحالة.
- **Paste**: للقيام باللصق للحالات.
- **Paste Variables**: وذلك لنسخ متغير.

- **Clear**: لمسح محتويات خلية.
- **Insert Variable**: لإدراج متغير بين متغيرين.
- **Insert Cases** : لإدراج حالة بين حالتين.
- **Find**: للقيام بعملية البحث
- **Go to--**: للذهاب إلى حالة أو متغير.
- **Options**: وهو زر للخيارات المتاحة.

وسنتكلم هنا عن أهم هذه العمليات:

1.5. حذف المتغيرات (الأعمدة):

لحذف عمود أو أكثر بما يحتويه من بيانات، علينا تحديد العمود بالنقر على اسم المتغير في أعلى العمود ثم ننقر فوق زر Delete في لوحة المفاتيح أو بالنقر على خيار Clear ضمن القائمة Edit.

2.5. حذف الحالات (الصفوف):

لحذف صف أو أكثر بما يحتويه من بيانات، علينا تحديد الصف بالنقر على رقم الصف في الجانب الأيسر ثم ننقر فوق زر Delete في لوحة المفاتيح أو بالنقر على خيار Clear ضمن القائمة Edit.

3.5. نسخ البيانات ونقلها:

يمكننا نسخ بيانات متغير أو حالة إلى مكان جديد وبالتالي الحصول على نسختين متطابقتين من البيانات أو نقل بيانات خلية إلى موقع آخر، وذلك باتباع الخطوات الآتية:

أ- نحدد المتغير أو الحالة التي نريد نسخها أو نقلها.

ب- نختر الأمر Copy للنسخ أو Cut للنقل من قائمة Edit.

ج- ننتقل للخلية التي نريد النقل أو نسخ البيانات إليها.

د- ننقر فوق Paste من قائمة Edit.

4.5. البحث عن القيم:

إذا رغبنا بالبحث عن قيمة لمتغير معين نتبع الخطوات الآتية:

- ننقر فوق الأمر Find من قائمة Edit فيظهر صندوق حوار معنون بـ Find and Replace- Data View.
- في مربع Find ندخل القيمة التي نبحث عنها.

الشكل رقم (5-14) صندوق الحوار الخاص بالبحث عن القيم

6. إدخال البيانات:

بعد القيام بتعريف متغيرات الدراسة يتم إدخال البيانات في صفحة محرر البيانات (Data View)، وذلك باتباع إحدى الطريقتين:

الطريقة الأولى: إدخال الأرقام لكل متغير من الحالة الأولى إلى آخر حالة وذلك بالضغط على الزر Enter في لوحة المفاتيح.

الطريقة الثانية: إدخال الأرقام لكل حالة من اليسار إلى اليمين وذلك بالضغط على الزر Tab في لوحة المفاتيح أو استخدام الأسهم من لوحة المفاتيح.

حالة عملية:

يرغب باحث في إجراء دراسة حول خصائص العاملين في مكاتب السياحة، وقد قام بإعداد الاستمارة التالية:

1. رقم الموظف:
2. الجنس: () ذكر () أنثى
3. تاريخ الولادة: --/--/----.
4. عدد سنوات الدراسة: ---.
5. الوظيفة: () إداري () دليل سياحي
6. الخبرة بالسنوات:_____.
7. الراتب الشهري:_____.

وبفرض أن استجابات 10 موظفين كانت على النحو التالي:

الحالة	رقم الموظف	الجنس	تاريخ الولادة	عدد سنوات الدراسة	الموقع الوظيفي	الخبرة بالسنوات	الدخل الشهري
1	14523A	ذكر	07/11/1980	12	دليل سياحي	5	460
2	11367A	ذكر	18/07/1962	15	إداري	25	950
3	14569D	ذكر	19/02/1982	12	دليل سياحي	4	440
4	17548G	ذكر	23/01/1991	12	دليل سياحي	1	380
5	12365A	أنثى	19/03/1988	14	إداري	2	460
6	14523D	أنثى	20/04/1985	15	دليل سياحي	6	570
7	16579G	ذكر	08/04/1981	12	دليل سياحي	8	520
8	19874D	أنثى	26/06/1987	14	إداري	4	500
9	14693A	أنثى	14/09/1979	15	إداري	11	670
10	12965H	أنثى	01/01/1970	15	دليل سياحي	20	850

الخطوة الأولى: نقوم بتعريف المتغيرات ومن أجل ذلك ننشئ جدولاً بسيطاً يحوي بعض خصائص تعريف المتغيرات:

Name	Type	Label	Values	Measure
رقم_الموظف	String	رقم الموظف		Nominal
الجنس	Numeric	الجنس	1 = ذكر 2 = أنثى	Nominal
ت_الولادة	Date	تاريخ الولادة		Nominal
سنوات_الدراسة	Numeric	عدد سنوات الدراسة		Scale
م_الوظيفي	Numeric	الموقع الوظيفي	1= دليل سياحي 2 = إداري	Nominal
الخبرة	Numeric	الخبرة بالسنوات		Scale
الدخل	Numeric	الدخل الشهري		Scale

بعد وضع هذا الجدول كدليل نقوم بفتح البرنامج الإحصائي SPSS ومن شاشة محرر البيانات Data Editor نعرف المتغيرات في صفحة (Variable View) فتظهر الصفحة كما في الشكل التالي:

الشكل رقم (5-15) صفحة تعريف المتغيرات

الخطوة الثانية: تفريغ البيانات في صفحة (Data View) ضمن شاشة محرر البيانات كما في الشكل التالي:

الشكل رقم (5-16) صفحة إدخال البيانات

7. القائمة View:

تعد هذه القائمة أيضاً من القوائم البسيطة والتي تندرج ضمن قوائم المعالجة الفنية للبيانات، وتحوي القائمة Viewعدداً من الوظائف التي تشابه مثيلتها في البرامج الأخرى بصفة عامة مع بعض الاختلاف تبعاً لطبيعة البرنامج، فيمكن من خلالها إلغاء خطوط الشبكة المحددة للخلايا في ملف البيانات، كما يمكن تغيير حجم الخط ونمطه كما تمكننا من الانتقال إلى صفحة تعريف المتغيرات أو إلى صفحة إدخال البيانات، وهذه الخيارات موضحة كما في الشكل التالي:

الشكل رقم (5-17) قائمة View والخيارات التي تحويها

1.7. الأمر Status Bar:

ويستخدم هذا الأمر لإظهار شريط الحالة الموجود في أسفل شاشة البرنامج SPSS أو إخفائه وهو شريط يظهر وضع معالج نظام SPSS ويظهر ضمنه حالات تقسيم الملف Split file، أو تحديد الحالات Filter on، أو تثقيل الحالات Weight on.

2.7. الأمر Toolbars:

وهذا الأمر تم شرحه سابقاً.

3.7. الأمر Fonts:

لتغيير حجم الخط ونوعه المستعمل عند الكتابة في صفحة محرر البيانات وعند اختيار الأمر Fonts يظهر صندوق حوار بعنوان Font كما يبينه الشكل التالي:

الشكل رقم (5-18) صندوق حوار Font لتعديل الخط المستخدم

وكما هو مُلاحظ فصندوق الحوار يحوي ثلاثة خيارات:

7.3.1. Font:

يستخدم لتحديد نوع الخط المستخدم، وتتوافر أنواع عديدة من الخطوط التي يمكن استخدامها حسب رغبة المستخدم.

7.3.2. Font Style:

ويستخدم لتحديد نمط الخط المستخدم حيث تتوافر الخيارات التالية:

- الخط العادي Regular.
- الخط المائل Italic.
- الخط الغامق Bold.
- الخط المائل الغامق: Bold Italic.

7.3.3. Size:

ويستخدم لتحديد حجم الخط.

أما الجزء المعنون بـ Sample فهو لإظهار شكل الخط بعد التعديل.

4.7. الأمر Grid Lines:

ويستخدم لعرض خطوط الشبكة للجداول الموجودة ضمن شاشة Data Editor أو إخفائها فعند تنشيطه تظهر الخطوط الفاصلة بين الخلايا الخاصة بالبيانات، أما في حالة عدم تنشيطه فتختفي تلك الخطوط.

5.7. الأمر Value Labels:

يتضمن هذا الأمر خاصية عرض وصف المتغيرات الاسمية والرتبية، فبدلاً من أن تظهر على شكل أسماء تظهر على شكل قيم رقمية مع ملاحظة أن ليس لهذه الأرقام أي دلالة حسابية.

6.7. الأمر Variable:

للانتقال من صفحة عرض البيانات Data View إلى صفحة تعريف المتغيرات Variable View وبالعكس أي الرجوع إلى صفحة عرض البيانات.

أسئلة الفصل الخامس

السؤال الأول: اختر الإجابة الصحيحة:

1- لتعريف متغير **رقم الطالب** في البرنامج SPSS:

فيكون اسم المتغير هو: a. رقم الطالب b. رقم_الطالب c. رقم-الطالب d. رقم+الطالب

2- لمتابعة تعريف متغير **رقم الطالب** نختار نوع المتغير من عمود **Type**:

النوع (Type) هو: a. Data b. Comma c. Dot d. String

3-لمتابعة تعريف متغير **رقم الطالب** نختار المقياس **Measure**:

مقياس المتغير (Measure) هو: a. Numeric b. Nominal c. Ordinal d. Scale

4-لتحديد نوع متغير **تاريخ الميلاد** نختار نوع المتغير من عمود **Type**:

النوع (Type) هو: a. Data b. Comma c. Dot d. Date

5-لإظهار التنسيق التالي للبيانات 23,512.89 نختار نوع المتغير من عمود **Type**:

نحدد نوع المتغير: a. Data b. Comma c. Dot d. Date

6-لإظهار التنسيق التالي للبيانات 23.512,89 نختار نوع المتغير من عمود **Type**:

نحدد نوع المتغير: e. Data f. Comma g. Dot h. Date

السؤال الثاني: يرغب باحث بتعريف متغيرات استمارة بحثه على البرنامج الإحصائي SPSS حيث كانت تشمل على الفقرات التالية:

1- العمر بالسنوات:

2- الجنسية: سوري عربي أجنبي

3- عدد أفراد الأسرة:

4- الرجاء الإجابة على الفقرات التالية بوضع اشارة (X) عند الإجابة المختارة:

	أوافق بشدة	أوافق	محايد	لا أوافق	لا أوافق بشدة
1. يتأثّر القرار السياحي لاختيار المقصد السياحي بتوافر الخدمة على مدار الوقت.					
2. يؤثر التنوع في الخدمات السياحية المقدمة في القرار السياحي.					
3. تؤثر جودة الخدمة السّياحيّة في قرارك السّياحيّ عند اختيار المقصد.					
4. التلاؤم والتناسب بين السّعر والخدمة المقدمة عامل جذب فعال.					
5. التّرويج السياحي يدفعك لاتخاذ قرارك في اختيار المقصد.					
6. يؤدي الترويج لموقع المنشأة السياحية دوراً هاماً في قرار اختيارك المقصد السياحي.					
7. يعتبر التّرويج لمقصد سياحي من أهم عوامل التّعريف به.					
8. يؤثّر المظهر العام للمنشأة في فعالية التّرويج السّياحيّ لها.					
9. يؤثر المظهر الداخلي للمنشأة في فعالية التّرويج السّياحيّ لها.					
10. إنّ توافر عنصر الأمن والأمان في سوريّة عموماً يحفّزك لاختيارها مقصداً سياحياً					

والمطلوب تعريف متغيرات الاستمارة وإدخالها على البرنامج الإحصائي SPSS.

السؤال الثالث: بفرض لدينا النتائج كما في الجدول التالي، والمطلوب تعريف المتغيرات وترميزها وإدخالها على البرنامج الإحصائي SPSS.

الرقم	المجموعة	السؤال 1	السؤال 2	السؤال 3
1	ضابطة	7	8	5
2	تجريبية	5	5	9
3	ضابطة	9	10	7
4	ضابطة	10	10	7
5	ضابطة	9	10	7
6	تجريبية	6	8	9
7	تجريبية	7	8	8
8	تجريبية	6	9	10
9	تجريبية	6	9	10
10	ضابطة	9	9	4

الفصل السادس CHAPTER 6

الإجراءات التنظيمية
في البرنامج الإحصائي SPSS

أهداف الفصل السادس:

1. الإجراءات التي تتعلق بتعريف البيانات.
2. الإجراءات التنظيمية للبيانات.
3. الإجراءات التنظيمية غير المباشرة للبيانات

في كثير من الأحيان لا تكون بيانات الدراسة مدخلة بالطريقة التي نرغب بها، لذلك نحتاج للاستفادة من بيانات موجودة بملف آخر، أو ترتيب البيانات وفقاً لأحد المتغيرات تصاعدياً أو تنازلياً، أو تحديد بعض الحالات المحددة من ملف البيانات، أو تقسيم ملف البيانات بناءً على متغير ما للمقارنة بين مجموعات هذا المتغير .... فالبرنامج الإحصائي SPSS يتيح العديد من الإجراءات التنظيمية على ملف البيانات لتسهل عملية المعالجة والتعامل مع بيانات لذلك لا بد لنا من التعرف على قائمة Data والتي تنقسم إلى عدة أجزاء، حيث:

الشكل رقم (6-1) يبين قائمة Data

إن هناك جزء يتعلق بتعريف البيانات الخام قبل استخدامها في التحليل الإحصائي، وجزء يحوي على إجراءات تنظيمية تظهر نتائجها مباشرة على الملف كترتيب بيانات الملف حسب قيم متغير ما تصاعدياً أو تنازلياً، والعمليات الخاصة بهذا الجزء تستخدم بعد إتمام عملية إدخال البيانات، أما الجزء الأخير فيحتوي على إجراءات تنظيمية لا تظهر نتائجها مباشرة على الملف وإنما عند استخدام الإجراءات الإحصائية أي في شاشة المخرجات، كما يمكننا ملاحظة أن هناك دلالة على استخدام هذه الإجراءات من خلال شريط الحالة (Status Bar) كما ذكر سابقاً.

1. الإجراءات المتعلقة بتعريف البيانات:

1.1. تعريف متغير بواسطة متغير آخر Define Variable Properties:

ويستخدم هذا الأمر لنسخ خصائص متغير إلى متغير آخر أو لنسخ خصائص متغير إلى مجموعة متغيرات من نفس الملف أو لتعريف متغير أو لمجموعة متغيرات، ولكن قبل البدء بتنفيذ هذا الإجراء علينا تحديد أسماء المتغيرات في صفحة تعريف المتغيرات Variable View.

ومن أجل تطبيق هذا الأمر لنأخذ المثال التالي: بفرض أننا قمنا بسؤال عينة من القادمين إلى القطر من أحد المطارات عن سبب القدوم إلى القطر، فكانت الأسئلة التالية:

سبب القدوم إلى القطر:	نعم	لا
1. متابعة التعليم	1	2
2. سياحة دينية	1	2
3. سياحة رياضية	1	2
4. سياحة تاريخية	1	2
5. سياحة علاجية	1	2
6. زيارة الأهل والأصدقاء	1	2
7. الاستثمار	1	2

1- نقوم بإدخال أسماء المتغيرات في صفحة تعريف المتغيرات (Variable View).

كما يلي: حيث نعبر عن السؤال الأول (q1) ونعبر عن السؤال الثاني بـ (q2) وهكذا حتى آخر سؤال....

2- من قائمة Data نختار الأمر Define Variable Proprieties فيظهر صندوق حوار بعنوان Define Variable Proprieties كما يلي:

الشكل رقم (6-2) يبين صندوق حوار Define Variable Proprieties

3- ننقل المتغيرات المراد تعريف خصائصها إلى مستطيل المعنون بـ Variable To scan وسننقل في هذا المثال جميع المتغيرات، ثم نضغط على Continue.

4- فيظهر صندوق حوار لتحديد خصائص المتغيرات المعرفة كما في الشكل التالي:

الشكل رقم (6-3) يبين صندوق حوار Define Variable Proprieties

نحدد المتغير الذي نرغب بتعريف خصائصه من خلال النقر عليه، حيث يمكننا أن نضيف له عنواناً من خلال الكتابة في مستطيل (Label)، ويمكننا تغير المقياس من قائمة (Measurement Value) ونوع المتغير من قائمة (Type) لأنه يأخذ بشكل افتراضي المقياس النسبي (Scale) والنوع الرقمي (Numeric)، كما يمكننا أن نضيف للمتغير فئات في مستطيل (Value Label grid)، فمثلاً المتغير (q1) يعبر عن السؤال التالي: سبب القدوم إلى القطر للسياحة الدينية؟ وهنا يوجد إجابتان فقط (نعم – لا) يتم إدخالها في المستطيل (Value Label grid)، كما يمكننا أن نستخدم الأمر (Suggest) في صندوق الحوار السابق بعد تحديد الفئات ليقدم لنا المقياس المقترح للمتغير.

فالخطوات السابقة تمكننا من تعريف خصائص متغير باستخدام الأمر Define Variable Proprieties، وإذا أردنا أن ننقل صفات المتغير (المقياس – الفئات – القيم المفقودة المعرفة – التنسيق) إلى بقية المتغيرات ننقر على الخيار (To Other Variables..) أما لنقل خصائص متغير إلى متغير آخر ننقر على (From Another Variable..) حيث يظهر صندوق حوار بعنوان (Apply Labels and Level From) فنحدد المتغير الذي نرغب بنسخ خصائصه، ثم نضغط على Copy.

2.1. نسخ خصائص متغير ..Copy Data Properties:

يقوم هذا الأمر بنسخ خصائص متغير (ماعدا الاسم والنوع) إلى متغير آخر من نفس الملف أو من أجل نسخ متغير أو عدة متغيرات من ملف خارجي إلى الملف المفتوح، ولهذا الخيار نستخدم الأمر ..Copy Data Properties.

3.1. تعريف التاريخ للسلسلة الزمنية Define Dates:

يقوم هذا الأمر بتوليد متغيرات التاريخ Date Variables التي تستعمل كتاريخ لقيم السلسلة الزمنية فقط، أي أن متغير التاريخ لا يستعمل في العمليات على السلسلة

الزمنية التي تنجز باستخدام الأمر Create time series ضمن قائمة Transform التي سنتعرف عليها لاحقاً.

مثال 1:

بفرض لدينا البيانات التالية التي تعبر عن أعداد السياح الأجانب لقلعة دمشق ابتداءً من شهر أيلول 2009.

والمطلوب: تعريف متغير تاريخ لمتغير أعداد السياح حسب السنة والفصل والشهر خلال المدة المذكورة.

لتنفيذ المطلوب نعرف متغير أعداد السياح في صفحة (Variable View) ثم ندخل البيانات في صفحة (Data View) كما يبينه الشكل التالي:

الشكل رقم (6-4) تعريف متغير أعداد السياح وإدخال البيانات

ثم نتبع الخطوات التالية:

1) من القائمة Data نختار الأمر Define Dates.
2) يظهر صندوق حوار بعنوان Define Dates كما يبينه الشكل التالي:

الشكل رقم (6-5) صندوق حوار تعريف متغير التاريخ

والذي يحوي الأقسام التالية:

القسم الأول: **Cases Are**: وفيه نختار الصيغة التي نريدها حيث تتوافر خيارات عديدة يمكن استخدامها حسب الحاجة، وفي مثالنا هنا نختار Years, quarters, months من قائمة Cases Are، ومنها على سبيل المثال لا الحصر:

- Years: تستخدم لعرض قيم السلسلة الزمنية حسب السنوات فقط.
- Years, quarters: تستخدم لعرض قيم السلسلة الزمنية حسب السنوات وأرباع السنوات.
- Years, months: تستخدم لعرض قيم السلسلة الزمنية حسب السنوات والأشهر.
- No date: لإزالة متغير التاريخ الذي سبق وتم تعريفه.

- Custom: لعرض قيم السلسلة الزمنية حسب تاريخ معين يحدد من قبل المستفيد (غير متوافر ضمن خيارات قائمة Cases Are)، مثلاً: أسبوع يتكون من أربعة أيام عمل، وهذا يتم عادةً باستخدام الأمر Syntax.

القسم الثاني: First Cases IS: يستخدم لتعريف التاريخ للحالة الأولى في السلسلة الزمنية، وفي مثالنا نقوم بما يلي:

- في مستطيل السنة (Year) ندخل سنة البداية 2009.
- في مستطيل الفصل(Quarter) ندخل فصل البداية 3 (شهر أيلول يقع في الفصل الثالث من السنة).
- في مستطيل الشهر (Month) ندخل شهر البداية، 9.

ملاحظة: لا يمكن تنفيذ هذا الأمر في حالة وجود تعارض بين الفصول والأشهر، مثلاً، استعمال الفصل 3 بدلاً من الفصل 2 لشهر حزيران.

- أما في الجزء Periodicity at higher level فإنه يبين دورية التاريخ أو أنه يعطي أكبر قيمة يمكن تنفيذها للبرنامج، مثلاً أعلى دورية للفصول هي 4 وأعلى دورية للأشهر هي 12 أي أنها وسيلة ضبط إدخال البيانات في الجزء الثاني من شاشة Define Dates.

عند نقر OK في شاشة حوار Define Dates تضاف متغيرات التاريخ إلى (Data View)، كما يوضحه الشكل التالي:

الشكل رقم (6-6) صفحة ادخال البيانات (Data View) بعد تعريف متغير التاريخ

4.1. التحقق من عملية الإدخال Validation:

في أثناء عملية إدخال البيانات تتم بعض الأخطاء البشرية، وقبل إجراء أي تحليل لهذه البيانات يجب علينا التأكد من خلو هذه البيانات من الأخطاء لأن الخطأ في إدخال بعض الأرقام يغير من مسار النتائج بشكل كبير، يتيح لنا البرنامج الإحصائي SPSS مجموعة من القواعد التي تمكننا من التأكد من صحة الإدخال، كما يمكننا بناء قواعد خاصة بالمستخدم للتحقق من البيانات، وذلك من خلال بناء:

- قواعد تحقق لبيانات متغير واحد (Single- Variable Rules).
- قواعد تحقق من بيانات عدة متغيرات وتتم عملية التحقق بناءً على صيغ منطقية أو علاقات يحددها المستخدم بين المتغيرات (Cross – Variable Rules).

كما يتم تحميل قواعد التحقق الموجودة مسبقاً ضمن البرنامج من خلال اتباع الخطوات التالية:

1) من القائمة Data نختار الأمر Validation.
2) من الخيار Validation نختار الأمر الفرعي Load Predefines Rules.
3) يظهر صندوق حوار بعنوان Load Predefined Validation Rules.
4) نضغط على (OK) ليتم تحميل قواعد التحقق الموجودة مسبقاً.

1.4.1. الحالة الأولى: بناء قواعد التحقق من أجل متغير واحد فقط:

يتيح لنا البرنامج الإحصائي (SPSS) تعريف قواعد للتحقق من صحة الإدخال إضافة إلى قواعد التحقق المعرفة مسبقاً، وذلك من خلال اتباع الخطوات التالية:

1) من قائمة Data نختار الأمر Validation.
2) من الخيار Validation نختار الأمر الفرعي Define Rules.
3) يظهر صندوق حوار بعنوان Defined Validation Rules نختار التبويب (Single-Variable Rules) المفتوح بشكل افتراضي كما في الشكل التالي:

الشكل رقم (6-7) صندوق حوار للتحقق من إدخال البيانات لمتغير واحد

من الشكل السابق نجد أنه يحوي على القواعد المعدة مسبقاً الموجودة ضمن جزء Rules، حيث يمكننا التعديل على هذه القواعد وذلك بتحديد القاعدة التي نرغب كما

يمكننا حذف القاعدة التي نريد من خلال الضغط على الزر (Delete) ومن أجل إضافة قواعد جديدة نتبع ما يلي:

I. نضغط على زر New ضمن التبويب Single-Variable Rules لبناء قاعدة تحقق صحة جديدة، حيث يظهر اسم القاعدة الجديدة (SingleVarRule1) بشكل افتراضي في قسم Rule Definition، يمكننا تعديل الاسم من خلال المستطيل (Name) مع مراعاة قواعد تسمية المتغيرات، ويمكننا أيضاً تعديل نوع القاعدة الجديدة والتي تأخذ النوع (Numeric) بشكل افتراضي، وهناك خياران آخران String و Date.

II. من قسم Valid Values نحدد نوع قيم التحقق والتي تقسم إلى قسمين:

- List In: وتستخدم للتحقق حسب قيم متقطعة أي يستخدم من أجل التحقق بالنسبة للمتغيرات المنفصلة أو الفئوية ونكتب القيم المحددة في مستطيل Values .
- Within a Range: وتستخدم للتحقق حسب قيم تأخذ مجالاً من القيم أي يستخدم هذا الخيار من أجل المتغيرات المستمرة، وفيها نحدد أصغر قيمة وأكبر قيمة.

مثال 2:

بفرض لدينا البيانات الواردة في الشكل رقم (6-6) لمجموعة من الأشخاص والتي تحوي متغير الجنس (1- ذكر ، 2- أنثى)، والمطلوب إضافة قاعدة للتحقق من عملية إدخال متغير الجنس.

الشكل رقم (6- 8) متغير الجنس لمجموعة أشخاص

لتنفيذ المطلوب نتبع الخطوات التالية:

1) من القائمة Data نختار الأمر Validation .
2) نختار الأمر الفرعي Define Rules.

يظهر صندوق حوار بعنوان Defined Validation Rules

3) نضغط على الزر New لتعريف قاعدة تحقق صحة فتأخذ اسم SingleVarRule 1.
4) بما أن المتغير الذي نريد تعريف قاعدة تحقق صحة من أجله من النوع المنقطع فإننا نختار من Valid Values نختار In a list فيظهر مستطيل معنون بـ Values نضع القيم التي يجب أن يأخذها المتغير وفي حالتنا هنا يأخذ القيمتين (1 ، 2).
5) ثم نضغط على زر OK حتى تتعرف قاعدة تحقق جديدة، كما في الشكل التالي:

الشكل رقم (6-9) خطوات بناء قاعدة للتحقق من إدخال البيانات لمتغير واحد

6) لإجراء قاعدة التحقق من الصحة من قائمة Data نختار Validation ثم نختار Validate data فيظهر صندوق الحوار بعنوان Validate data.
7) ننقل متغير الرقم على القسم المعنون بـ Case Identifier Variables.
8) ننقل المتغير الذي نريد وضع قاعدة للتحقق من صحته ضمن المستطيل المعنون: Analysis Variables.
9) ننقر على التبويب Rules Single-Variables لتفعيل القاعدة التي نريدها للتحقق، كما نجد من مستطيل Analysis Variables أنه يحوي على المتغير الذي نريد التحقق من صحته وعندما نفعل القاعدة نجد ضمن عمود Rules أصبحت القيمة 1 بدل 0 كما في الشكل التالي:

الشكل رقم (6-10) خطوات التحقق من إدخال البيانات لمتغير واحد

10) ثم نضغط على الزر OK سيظهر في شاشة المخرجات الجداول التالية:

- الجدول الأول المعنون (Rule Descriptions) يتم كتابة القاعدة التي تم على أساسها عملية التحقق.
- الجدول الثاني المعنون (Variable Summary) يتم بيان عدد الحالات التي شذت عن القاعدة التي تم تعريفها.
- الجدول الثالث المعنون (Case Report) يتم كتابة كافة أرقام الحالات التي شذت عن القاعدة المعرفة وذلك في عمود Case.

والشكل التالي يبين الجداول الثلاثة السابقة:

الشكل رقم (6-11) مخرجات برنامج SPSS للتحقق من إدخال البيانات لمتغير واحد

2.4.1. الحالة الثانية: بناء قواعد التحقق من أجل المتغيرات التقاطعية:

يتيح لنا البرنامج الإحصائي (SPSS) تعريف قواعد للتحقق من صحة الإدخال بناءً على إجراء علاقة بين متغيرين أو أكثر من ضمن المتغيرات المتاحة، وذلك من خلال اتباع الخطوات التالية:

1) من القائمة Data نختار الأمر Validation.
2) نختار الأمر الفرعي Define Rules.

3) يظهر صندوق حوار بعنوان Defined Validation Rules من التبويب Cross – Variable Rules وضمن المستطيل Rule Definition نكتب القاعدة التي تربط بين المتغيرين أو عدة متغيرات، كما يوجد طريقة ثانية يتم مناقشتها كما في المثال التالي:

مثال 3:

لتكن البيانات الواردة في الجدول التالي والتي تحوي على المتغيرات التالية:

رقم الموظف – الجنس (1=ذكر، 2= أنثى) – الحالة الاجتماعية (1= عازب، 2= متزوج) – عدد الأولاد، **والمطلوب:** التحقق من أن إدخال البيانات كان صحيحاً.

	رقم_الموظف	الجنس	الحالة_الاجتماعية	عدد_الأولاد
1	125	ذكر	عازب	0
2	168	أنثى	عازب	0
3	189	ذكر	عازب	2
4	197	أنثى	متزوج	2
5	213	أنثى	عازب	3
6	228	ذكر	عازب	1
7	236	أنثى	متزوج	1
8	267	أنثى	عازب	3
9	279	ذكر	متزوج	2
10	296	أنثى	عازب	0
11	299	أنثى	عازب	0
12	310	أنثى	متزوج	3
13	320	أنثى	متزوج	3
14	327	ذكر	عازب	2
15	356	ذكر	متزوج	1

لتنفيذ المطلوب نتبع الخطوات التالية:

1) من القائمة Data نختار الأمر Validate Data.
2) يظهر صندوق حوار بعنوان Validate Data.
3) ننقل متغيري الحالة الاجتماعية وعدد الأولاد (لأن الشخص الأعزب يجب ألا يكون لديه أولاد (عدد الأولاد =0) لذلك يجب أن نضع الشرط التالي (الشخص العازب عدد أولاده =0)) إلى المستطيل المعنون: Analysis Variables ومتغير رقم الموظف إلى Case Identifier Variables.
4) من التبويب Cross – Variable Rules ومنه نختار Define Rules فيظهر صندوق حوار بعنوان Define Validation Rules نكتب قاعدة التحقق من الصحة والتي تنص أن الأفراد العازبين ليس لديهم أولاد أي إذا كان عدد الأولاد أكبر من 0 للشخص العازب فهي حالة خاطئة، كما في الشكل التالي:

الشكل رقم (6-12) خطوات بناء قاعدة للتحقق من إدخال بيانات لمتغيرين

ثم نضغط على Continue وإذا أردنا أن يدرج متغير يعبر عن الحالات التي تشذ عن القاعدة المدخلة، نضغط على التبويب Save ونحدد الخيار Save Indicator Variables That Record All Validation Rule Violations ثم نضغط على OK.

فيظهر في شاشة المخرجات الجدولان التاليان:

- الجدول الأول المعنون (Cross Variable Rules) يتم كتابة القاعدة التي تم على أساسها عملية التحقق.
- الجدول الثاني المعنون (Case Report) يظهر لنا أرقام الحالات التي شذت عن القاعدة وكذلك يتم كتابة الرقم التعريفي الذي أدخلناه في المستطيل Case Identifier Variables وفي مثالنا هنا نستخدم رقم الموظف.

وهذا ما يبينه الشكل التالي:

الشكل رقم (6-13) مخرجات برنامج SPSS للتحقق من إدخال البيانات لمتغيرين

كما نجد أنه أضاف متغيراً على قاعدة البيانات لتصبح كما في الشكل التالي:

الشكل رقم (6-14) شكل قاعدة البيانات بعد التحقق من صحة الإدخال

5.1. تحديد الحالات المتكررة Identify Duplicate Cases:

وجدنا في الفقرة السابقة أنه يتم التحقق من خلو البيانات من الأخطاء، ولكن ماذا لو كان لدينا تكرار لبعض الحالات حسب متغير فريد يعتبره المستخدم مفتاحاً أساسياً أي لا تتكرر قيمه إلا مرة واحدة وكمثال عن هذا المتغير هو الرقم الوطني في الهوية فلكل شخص رقم وطني واحد ومميز، ففي المتغير الفريد يكون لكل سطر (حالة) رقم مميز وفريد.

فالبرنامج الإحصائي SPSS يتيح لنا هذه الإمكانية من خلال الأمر Identify Duplicate Cases والذي يستفاد منه بتحديد الحالات المتكررة التي تحدث نتيجة الخطأ في الإدخال.

2. الإجراءات التنظيمية للبيانات:

يوفر البرنامج الإحصائيSPSS مجموعة من الإجراءات الخاصة بتنظيم البيانات، وذلك من خلال قائمة Data، وتبدأ التعليمات الخاصة بالإجراءات التنظيمية من الأمر Sort Cases وتنتهي بالأمر Aggregate فتمكننا هذه الأوامر بترتيب البيانات ودمج الملفات وخلق بيانات جديدة، وهذا ما سنتناوله في الفقرات التالية:

1.2. ترتيب البيانات Sort Cases:

يستخدم الأمر Sort Cases لترتيب البيانات (الحالات) المدخلة في صفحة إدخال البيانات (Variable View) حسب قيم متغير معين أو عدة متغيرات في ملف ما، ترتيباً تصاعدياً أو تنازلياً، وعند اختياره يظهر صندوق حوار بعنوان Sort Cases كما في الشكل التالي:

الشكل رقم (6-15) صندوق الحوار الخاص بترتيب البيانات

يظهر في صندوق الحوار ما يلي:

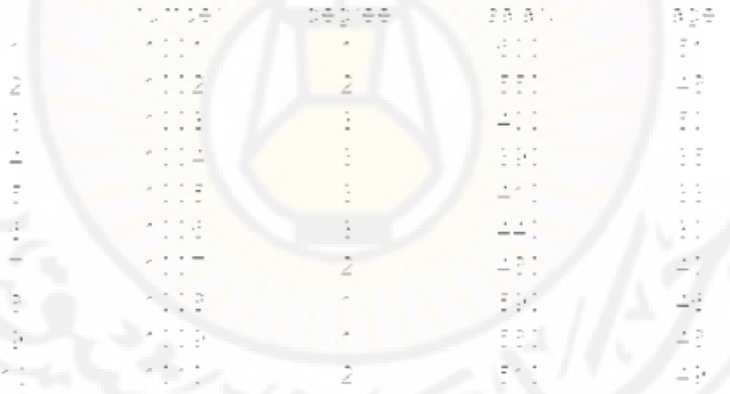
- Sort By: نضع ضمنه المتغير الذي نرغب بترتيب البيانات بناءً عليه.
- sort order: ضمن هذا الحقل نحدد نوع الترتيب الذي نريده فنختار أحد الخيارين التاليين:

1) Ascending: لغرض الترتيب التصاعدي للبيانات ويشار إليها بالحرف A إلى جانب المتغير الذي نريد الترتيب بالنسبة له.
2) Descending: من أجل الترتيب التنازلي للبيانات ويشار إليها بالحرف D إلى جانب المتغير الذي نريد الترتيب بالنسبة له.

ملاحظة: يمكننا ترتيب البيانات حسب متغيرين أو أكثر .

مثال 4:

اعتماداً على البيانات الواردة في الجدول التالي، المطلوب ترتيب البيانات تصاعدياً حسب المتغير Salary ضمن فئات الدرجة الوظيفية Degree، (أي الترتيب حسب متغير Salary أولاً ثم حسب المغير Degree).



خطوات التنفيذ:

1) من القائمة Data نختار الأمر Sort Cases.
2) يظهر صندوق حوار بعنوان Sort Cases.
3) ننقل متغير Salary إلى حقل Sort by.
4) ننقل متغير Degree إلى حقل Sort by.
5) ننقر على الزر Ok، نحصل على ترتيب البيانات حسب المتغيرين المشار إليهما وكما هو واضح في الشكل التالي:

الشكل رقم (6-16) خطوات ترتيب البيانات ونتائج الترتيب

2.2. ترتيب المتغيرات Sort Variables:

يمكن ترتيب المتغيرات المعرفة في صفحة تعريف المتغيرات حسب خصائص المتغير كالاسم أو النوع .....، ويمكننا ترتيبها تصاعدياً أو تنازلياً، ويتم ذلك من القائمة Data باختيار الأمر Sort Variables فيظهر صندوق حوار بعنوان Sort Variables، له الشكل التالي:

الشكل رقم (6-17) صندوق الحوار الخاص بترتيب المتغيرات

ومنه نلاحظ:

- من قسم Variable View Columns نختار الصفة التي نريد ترتيب المتغير بالنسبة لها، والتي يظهر ضمن هذا القسم جميع الصفات التي ندرجها في أثناء تعريف أي متغير.
- Sort Order: ضمن هذا الحقل نحدد نوع الترتيب الذي نريده فنختار أحد الخيارين التاليين:

I. Ascending: لغرض الترتيب التصاعدي للمتغيرات حسب الصفة أو الخاصة التي رتبنا المتغيرات بناءً عليها.

II. Descending: من أجل الترتيب التنازلي للمتغيرات حسب الصفة أو الخاصة التي رتبنا المتغيرات بناءً عليها.

- الخيار (Save the current (pre-sorted) variable order in a new variable order in a new attribute) وعند التأشير عليه يقوم بإضافة صفة إلى صفحة تعريف المتغيرات (Variable View) ويكون ضمن هذه الصفة رقم ترتيب المتغير في صفحة إدخال البيانات (Data View) قبل ترتيب المتغيرات، حيث يطلب منا وضع اسم لهذه الصفة.

3.2. المبادلة بين المتغيرات والحالات Transpose:

يستعمل هذا الأمر لتبديل الحالات (الصفوف) إلى متغيرات (أعمدة) وبالعكس، أي تحويل المتغيرات Variables إلى حالات Cases وبالعكس.

مثال 5:

بفرض لدينا مصفوفة x معرفة أعمدتها، متغيراتها، x_1, x_2, x_3 إضافة متغير تسميه Y.

	X1	X2	X3	Y
1	3.00	6.00	9.00	y1
2	4.00	7.00	10.00	y2
3	5.00	2.00	8.00	y3

وقد أدخلت البيانات في شاشة Data Editor وكما هو موضح في الشكل.

المطلوب: إيجاد تبديل المتغيرات إلى حالات وتسمية أعمدتها بقيم المتغير y.

لتنفيذ ذلك نتبع الخطوات التالية:

1) من قائمة Data نختار الأمر Transpose.
2) يظهر صندوق حوار بعنوان Transpose.
3) نختار المتغيرات (الأعمدة) x_1, x_2, x_3 ونقلها إلى مستطيل Variable (s).
4) ننقل متغير التسمية بعد التبديل Y وننقله إلى مستطيل Name Variable.
5) نضغط على OK، فنحصل على تبديل المتغيرات إلى حالات والتي تظهر في شاشة Data Editor.

الشكل رقم (6-18) خطوات المبادلة بين الحالات والمتغيرات ونتيجة التبديل

4.2. دمج الملفات Merge Files:

دمج الملفات عبارة عن عملية تجميع أكثر من ملف ويستعمل الأمر Merge Files لذلك، وهذه العملية مهمة جداً في حالة استخدام البرنامج الإحصائيSPSS كقاعدة بيانات، وتتم عملية دمج الملفات بإحدى الطريقتين التاليتين:

أ- دمج ملفين يحتويان المتغيرات نفسها وبحالات مختلفة Add Cases.

ب- دمج ملفين يحتويان متغيرات مختلفة وحالات نفسها Add Variables.

1.4.2. Add Cases:

تتيح هذه الطريقة دمج ملفين يحتويان على نفس المتغيرات بحالات مختلفة، وكمثال على ذلك قيام فردين بتفريغ بيانات دراسة معينة حيث يقوم كل منهما بتفريغ عدد من الاستمارات في ملف معين ونرغب الآن بدمج الملفين معاً لتكون جميع البيانات في ملف واحد، فهنا نلاحظ ضرورة أن يتوافر في الملفين المدمجين تماثل المتغيرات (بجميع الخصائص).

مثال 6:

بفرض لدينا البيانات في الملف Merge Data1 الذي يحوي المتغيرات التالية:

وبفرض لدينا البيانات في الملف الثاني Data2 Merge الذي يحتوي على المتغيرات التالية:

المطلوب: قم بدمج الملفين Data1 Merge و Data2 Merge لنحصل على ملف بيانات جديد.

لتنفيذ المطلوب نتبع الخطوات التالية:

1) تأكد أن أحد الملفين مفتوح أمامك وليكن Data1 Merge ويعرف باسم الملف العامل Working Data File.
2) من القائمة Data نختار Merge Files.
3) من Merge Files نختار Add Cases.
4) فيظهر صندوق حوار بعنوان Add Cases To Merge Data1 ثم نضغط هنا على Browse فتظهر نافذة لاختيار الملف الذي نرغب بدمجه مع الملف المفتوح.
5) اختر الملف Data2 Merge المراد دمجه مع الملف Data1 Merge والذي يعرف بملف البيانات الخارجي، External Data File.
6) فيظهر صندوق حوار بعنوان Add Cases Form.

الشكل رقم (6-19) صندوق حوار دمج الحالات

تتضمن هذه الشاشة الخيارات والمكونات التالية:

- Variables in New Active Dataset: تتضمن هذا القائمة المتغيرات التي سوف تتضمن في الملف الناتج من عملية الدمج والتي تتطابق من ناحية الاسم ونوع المتغير (عددي أم رمزي) ويمكن حذف أي متغير من هذه القائمة وعدم تضمينه في الملف الجديد (المدمج).
- Unpaired Variable: تتضمن هذه القائمة أسماء المتغيرات التي لن تتضمن في الملف الناتج عن عملية الدمج.
  - ❖ ويرمز للمتغيرات من الملف العامل بالرمز *.
  - ❖ ويرمز للمتغيرات من الملف الخارجي بالرمز + .
  - ❖ وتشمل القائمة Unpaired Variables المتغيرات التالية:
    - المتغيرات من الملفين المدمجين التي لا تتطابق من ناحية الاسم.

- المتغيرات التي تم تعريفها كمتغيرات عددية في أحد الملفين وكمتغيرات رمزية في الملف الآخر حيث لا يمكن دمج المتغيرات العددية بالمتغيرات الرمزية.
- المتغيرات الرمزية التي لها أطوال غير متساوية، Unequal Width في كلا الملفين.

▪ Indicate Case Source Variable: عند تأشير هذا الخيار يتم إضافة متغير جديد باسم Source1 في الملف المدمج وإعطاء القيمة (0) لحالات الملف العامل و القيمة(1) لحالات الملف الخارجي.

7) عند ضغط OK في شاشة حوار Add Cases Form يتم دمج الملفين ويظهر الملف الجديد في شاشة Data Editor كما في الشكل التالي والذي يمكن تسميته من خلال الأمر Save as:

الشكل رقم (6-20) شكل قاعدة البيانات بعد عملية دمج الحالات

2.4.2. الحالة الثانية إضافة المتغيرات Add Variables:

تستخدم هذه الحالة لدمج الملف العامل الحالي مع الملف الخارجي واللذين يحتويان نفس الحالات ولكن المتغيرات مختلفة، وتكون عملية الدمج صحيحة إحصائياً ومنطقياً إذا كان الملفان يحتويان على العينة نفسها، أي أنّ المتغيرات الموجودة في الملفين هي متغيرات متعلقة بالمجموعة نفسها من الأفراد فليس منطقياً ولا صحيحاً أن تضاف معلومات عن زيد وخلدون إلى المعلومات المتعلقة بحالة رهف.

ولذلك فإنّ الأصل في عملية إضافة المتغيرات، أن تكون الملفات مرتبة بطريقة واحدة يراعى فيها أن يكون الشخص الأول في الملف الأول هو الشخص الأول نفسه في الملف الثاني، وإذا لم نكن متأكدين من هذا الوضع، يفضل أن تتم عملية الدمج بناءً على متغير مشترك بين الملفين يسمى المتغير المفتاح Key Variable ويجب أن ترتب البيانات حسب هذا المتغير في الملفين المراد دمجهما قبل إجراء عملية الدمج (أنظر عملية ترتيب البيانات Sort Cases) حيث تتم مطابقة البيانات الموجودة في الملفين حسب تطابق قيم هذا المتغير، وسيقوم الحاسوب بإظهار خطأ في حالة أن الملفين لم يكونا مرتبين حسب هذا المتغير.

مثال 7:

لنفترض لدينا ملف Merge Variable 1 والذي يحوي المتغيرات التالية:

الرقم – الاسم– الدرجة في مقرر الرياضيات – الدرجة في مقرر الكيمياء – الدرجة في مقرر العربي – الدرجة في مقرر الانكليزي، والبيانات كما يلي:

والملف الثاني Merge Variable 2 والذي يحوي المتغيرات التالية:

الرقم – الاسم– الدرجة في مقرر المعلوماتية – الدرجة في الاحصاء، والبيانات كما يلي:

المطلوب: دمج الملفين السابقين بملف واحد.

لتنفيذ المطلوب نتبع الخطوات التالية:

1) افتح الملف الأول Merge Variable 1.
2) من القائمة Data نختار Merge Files.
3) من Merge Files نختار Add Variables.
4) يظهر صندوق حوار بعنوان Add Variables: Read File ثم نضغط هنا على Browse فتظهر نافذة لاختيار الملف الذي نرغب بدمجه مع الملف المفتوح، وهنا نختار الملف الثاني الذي نرغب بدمجه Merge Variable 2.

5) وعند نقر الزر Open في شاشة حوار Add Variables: Read File ثم Continue يظهر صندوق حوار Add Variables Form إن هذا الصندوق يتضمن المكونات والخيارات التالية:

الشكل رقم (6-21) صندوق حوار دمج متغيرات

a) New Active Dataset: وهي قائمة بأسماء المتغيرات التي سوف تضمن في الملف الجديد المدمج، وهي متغيرات لا تتشابه من ناحية الاسم في كلا الملفين المدمجين.

b) Excluded Variables: وهي قائمة بأسماء المتغيرات التي تستبعد من الملف المدمج الجديد، وعادة تكون هذه المتغيرات من الملف الثاني الخارجي، ونلاحظ أن هذه القائمة تتضمن متغير الاسم والرقم من ملف 2 Merge Variable (رمزه+) وذلك لتكراره في كلا الملفين.

c) Key Variables: تستعمل هذه المتغيرات في حالة عدم تطابق بعض الحالات في كلا الملفين المراد دمجهما، مثلاً احتواء الملفين على بعض الحالات غير المتشابهة أو أن هناك حالات مفقودة في المتغير المشترك.

علماً أنه يتوجب توافر الشروط التالية في Key Variables:

- يجب أن يضمن Key Variables في كلا الملفين المراد دمجهما وبنفس الاسم.
- يجب ترتيب الملفين تصاعدياً، Sorting Ascending وبموجب Key Variables.

6) عند نقر الزر ok في صندوق حوار Add Variables Form، نحصل على الملف الناتج عن عملية الدمج، كما في الشكل التالي:

الشكل رقم (6-22) شكل قاعدة البيانات بعد عملية دمج المتغيرات

5.2. تجميع البيانات Aggregate:

يستعمل هذا الأمر لحساب بعض القياسات والإجراءات الإحصائية لمجموعة من الحالات وفق معيار (متغير) محدد ووضع النتائج في ملف جديد تمهيداً للتحليل فمثلاً يمكن أن نعرض الوسط الحسابي لمعدلات الطلبة في كل مقرر أو كلية وكذلك الانحراف المعياري، في ملف تجميعي جديد بدلاً من عرض معدلات الطلبة لكافة الكليات.

مثال 8:

ليكن البيانات التي حصلنا عليها من المثال رقم 6 بعد عملية الدمج.

المطلوب: إيجاد الانحراف المعياري لدخل موظفي الشركة حسب متغير الشهادة.

لتنفيذ المطلوب نتبع الخطوات التالية:

1) من القائمة Data نختار Aggregate.
2) يظهر صندوق حوار Aggregate Data والذي يحوي:

a. Break Variables: وهو متغير (متغيرات) تجزئة (التقسيم) يستعمل لتعريف مجاميع (من الحالات) ويجب أن يكون المتغير نوعياً، وفي تطبيقنا ننقل متغير الشهادة.

b. Aggregate variables: وهو المتغير (المتغيرات) الذي نرغب في تجميع حالاتها والتي يتكون منها حسب متغير التجزئة، وفي تطبيقنا فإن المتغير التجميعي هو الدخل وعند نقله إلى مربع Aggregate Variables فإنه يظهر باسم افتراضي هو الدخل_Mean أي متغير متوسط الدخل لأن دالة المتوسط يضيفها بشكل تلقائي، ولكن المطلوب في المثال هو الانحراف المعياري ويمثل الانحراف المعياري للرواتب لكل شهادة، ولتغير الدالة الافتراضية (المتوسط

الحسابي Mean) ننقر على الزر Function بعد تحديد اسم المتغير في مربع Aggregate variables، فيظهر صندوق حوار بعنوان :Aggregate Data Aggregate Function كما في الشكل التالي:

الشكل رقم (6-23) صندوق حوار لتحديد دالة تجميع البيانات

من صندوق الحوار يمكن اختيار الدالة المطلوبة والمناسبة وفي مثالنا هنا نختار Standard deviation، بعد اختيار الانحراف المعياري يصبح اسم متغير التجميع الدخل_SD ويمكننا تعديل هذا الاسم باستعمال الزر Name & Label بعد تحديد اسم المتغير في مربع Aggregate variables.

c. Save number of cases in break group as variable: عند تأشير المربع المجاور له يتم تكوين متغير جديد باسم افتراضي هو N- Break يبين عدد الحالات لكل مجموعة.

d. Add Aggregate Variables To Active Dataset: عند تفعيل المربع المجاور له سيتيح تكوين متغير جديد هو متغير التجميع على ملف قاعدة البيانات.

e. Create new data file: عند تفعيل مربع المجاور له سيتيح تكوين ملف جديد يحتوي المعلومات التجميعية باسم افتراضي Aggr في نفس الدليل Directory، الذي يقع فيه الملف الأصلي وقد تم تفعيل هذا الخيار ويمكن تغيير اسم الملف التجميعي وموقعه من خلال ضغط الزر File.

f. Replace working data file: يستخدم هذا الخيار لإحلال الملف التجميعي محل الملف الحالي، Salary علماً أن ذلك لا يلغي خزن الملف الأصلي، وأن الملف التجميعي الناتج لا يخزن ما لم يتم خزنه باسم محدد (معين)، فيصبح شكل صندوق الحوار كما يلي:

الشكل رقم (6-24) صندوق حوار التجميع

g. عند نقر زر OK يتم تكوين وتخزين الملف التجميعي Aggr (لقراءة محتويات الملفAggr يتطلب فتحه من قائمة File)، فتصبح قاعدة البيانات كما يلي:

الشكل رقم (6-25) شكل قاعدة البيانات بعد عملية التجميع

h. ومن قاعدة البيانات أعلاه نلاحظ مقدار الانحراف المعياري بين دخل العاملين وفقاً لكل شهادة، حيث نرى أن الانحراف بين رواتب الموظفين ضمن الشهادة الثانوية أعلى بكثير من الانحراف في شهادتي المعهد المتوسط والجامعية وهو مؤشر قد يستفاد منه في اتخاذ القرارات الإدارية.

ملاحظة: يمكن إدخال المتغير الدخل إلى مربع Aggregate variables عدة مرات في نفس الوقت حيث يأخذ برقم تسلسلي حيث الدخل-1، الدخل-2، ...الخ، مثلاً المتغير الدخل-1 يمثل تجميع الحالات حسب المتوسط الحسابي، والمتغير الدخل-2 يمثّل التجميع حسب الانحراف المعياري، ...وهكذا.

3. الإجراءات التنظيمية غير المباشرة للبيانات:

يتضمن هذا الجزء الأوامر من Split file إلى Weight Cases، التالية وحسب تسلسلها في الحزمة:

1.3. تجزئة الملفات: Split file:

يستعمل هذا الأمر لغرض تجزئة (فصل) ملف البيانات لأغراض إحصائية أو تنظيمية وفق معيار محدد، وعند اختيار الأمر Split file من قائمة Data يظهر صندوق حوار بعنوان Split File كما في الشكل التالي:

الشكل رقم (6-26) صندوق حوار تجزئة الملفات

- Analyze all cases, do not create groups: ويعني عدم تجزئة الملف.
- Compare groups: ويتم تجزئة الملف حسب فئات متغير معين (يحدد هذا المتغير أو مجموعة من المتغيرات في خانة Group Based on) ويتم عرض النتائج لأي عملية إحصائية على شكل مقارنة النتائج وفقاً لفئات متغير التجزئة.
- Organize output by groups: هذا الخيار لا يختلف عن الخيار السابق Compare groups، ولكن يتم عرض النتائج بصورة مستقلة لكل فئة من فئات متغير التجزئة، أي الاختلاف بين هذا الخيار والخيار السابق يكمن في طريق عرض النتائج العمليات الإحصائية.
- Sort the file by grouping variable(s): عند تفعيل هذا الخيار يتم ترتيب الملف حسب فئات متغير (متغيرات) التجزئة، وهذا الخيار مفعل بشكل افتراضي.
- File is already sorted: وهذا الخيار يعبر عن أن الملف لا يحتاج إلى ترتيب والبيانات مرتبة بصورة صحيحة حسب متغير التجزئة.

- عند ضغط زر OK في صندوق حوار Split file يتم تجزئة الملف إلى عدد فئات متغير التجزئة ولن نلاحظ أي تغيير في شاشة محرر البيانات وسنجد على شريط الحالة مكتوب Split by.

2.3. اختيار الحالات: Select Cases:

إن تطبيق الأمر اختيار الحالات Select Cases يمكننا من اختيار مجموعة من مفردات العينة بطرق ومعايير مختلفة تتراوح من مدى تطابق تعبير رياضي أو منطقي معقد إلى أن تكون المفردات مختارة بطريقة عشوائية، كما أن المفردات التي لا تنطبق عليها شروط الاختيار يمكن استبعادها أو حذفها نهائياً، فالمفردات التي يتم استبعادها سوف تبقى في ملف البيانات ولكن لن تدخل في التحليل وسوف تظهر في صفحة محرر البيانات (Data View) كأنها مشطوبة ويظهر الشطب على رقم الحالة (السطر) كما يقوم البرنامج الإحصائي SPSS بتكوين متغير جديد باسم $_FILTER ليكون مؤشراً على حالة المفردة وما إذا تم اختيارها أم لا، فإذا تم اختيار المفردة سيأخذ هذا المتغير القيمة 1 وإذا لم يتم اختيارها سيأخذ القيمة 0، ويمكن إلغاء هذا التجميد للمفردات في أي وقت عن طريق تطبيق الأمر "اختيار جميع المفردات" ولكن خيار حذف المفردات سيلغيها نهائياً، لذا يجب الحذر عند حفظ ملف البيانات، فحفظه بنفس الاسم القديم يعني أنه لا يمكن العودة إلى الملف الأصلي وستفقد جميع البيانات التي تم حذفها.

وعند اختيار الأمر Select Cases من قائمة Data يظهر صندوق حوار بعنوان Select Cases كما في الشكل التالي:

الشكل رقم (6-27) صندوق حوار اختيار الحالات

من صندوق الحوار المعنون بـ Select Cases نجد الخيارات التالية:

1) All cases: يعني اختيار كافة الحالات ضمن العينة.
2) If condition is satisfied: وهذا الخيار يتيح اختيار مفردات على أن يتحقق شرط معين، لذا فإنه بالضغط على هذا الخيار If Condition is Satisfied ثم الضغط على الشرط ..IF أسفل هذا الخيار سيفتح صندوق حوار فرعي بعنوان Select Cases: If يمكننا من اختيار مفردات بشرط تحقق شرط معين، وهذا الشرط يمكن أن يكون على شكل صيغة رياضية أو صيغة منطقية، ويتم اختيار المفردة في المجموعة فقط إذا كانت محققة للشرط المكتوب وستستبعد المفردة في

حالة عدم تحقق الشرط، ويمكن أن تحتوي الصيغة الرياضية على أكثر من دالة في نفس الوقت ويمكنها أيضاً أن تحتوي على جميع الإشارات الجبرية، وكذلك يمكنها أن تحتوي على أسماء متغيرات، والشكل التالي يبين صندوق حوار الاختيار الشرطي للمفردات:

الشكل رقم (6-28) صندوق حوار الاختيار الشرطي للمفردات

3) Random Sample of Cases: ويعتبر أحد الخيارات الهامة لاختيار عينات عشوائية صغيرة من جميع الحالات المدخلة في صفحة إدخال البيانات (Data View)، وهذا الخيار يتيح للمستخدم اختيار عينة عشوائية بسيطة من البيانات المدخلة بالحجم المرغوب فيه، لذا فإنه عند اختيار" عينة عشوائية من الحالات" Random Sample of Cases ثم الضغط على عينة Sample أسفله سيفتح

صندوق حوار بعنوان Select Cases: Random Sample ليتم تحديد حجم العينة المرغوب (عدد مفرداتها) كما يبينه الشكل التالي:

الشكل رقم (6-29) صندوق حوار الاختيار الشرطي للحالات

وتتيح هذه النافذة الفرصة لتحديد حجم العينة إما كنسبة مئوية من عدد الحالات الكلي للبيانات أو كعدد محدد من هذه البيانات، كما يلي:

- **Approximately**: يتم اختيار عينة مكونة من النسبة المئوية التي يتم تحديدها من الحالات في المربع الملاصق تقريباً، وحيث إن طريقة الاختيار العشوائي تعتمد على أسس الاحتمالات وعلى أن النسبة المعطاة تمثل احتمال ظهور أي مفردة في العينة فإن حجم العينة المختارة لن يكون بالضرورة مساوياً لتلك النسبة تماماً خاصة عندما يكون عدد مفردات العينة صغيراً.
- **Exactly**: يحدد عدد الحالات التي يجب اختيارها عشوائياً بدقة ولكن يجب هنا تحديد عدد الحالات بالإضافة إلى عدد آخر يمثل بداية الحالات التي يجب الاختيار منها، وهذا العدد يجب ألا يزيد عن عدد الحالات الكلي.

4) Based on time or case rang: وهذا الخيار يتيح للمستخدم فرصة اختيار الحالات استناداً إلى مدى معين من داخل مدى البيانات يتم تحديده، فعند تحديد "اختيار مفردات معتمدًا على المدى" Based on Time or Case Range ثم

الضغط على المدى " Range " سيفتح صندوق حوار بعنوان :Select Cases Range لتحديد القيمة الأولى والقيمة الأخيرة في هذه المفردات التي يتطلب اختيارها ويمكن أن يكون هذا المدى معتمدًا على أرقام المفردات أو تاريخ أو أوقات محددة ولكنه في النهاية سيعتمد على أرقام الحالات (الأسطر) التي تقع بها تلك الحالات وذلك كما تظهر في صفحة إدخال البيانات (Data View)، لكن الاعتماد على التاريخ أو الوقت لن يتم إلا إذا كانت هناك بيانات سلسلة زمنية بمتغيرات تاريخ Date Variables ومعرفة مسبقًا، وهذا ما يبينه الشكل التالي:

الشكل رقم (6-30) صندوق حوار تحديد المدى للحالات المختارة

5) Use Filter Variable: هذا الخيار يتيح لنا إمكانية اختيار الحالات من خلال إضافة Filter Variable إلى الملف حيث نعطي القيمة 1 للحالات التي نرغب في اختيارها والقيمة0 للحالات التي لا نرغب في اختيارها.
6) Filter Unselected cases: في هذه الحالة يتم إبقاء الملف الحالي، ووضع إشارة خط مائل (/) للحالات التي تم استبعادها ويظهر متغير جديد للتعبير عن الحالات التي تم اختيارها بإعطائها الرقم (1) والرقم (0) للحالات التي تم استبعادها.
7) Deleted Unselected cases: في هذه الحالة يتم حذف الحالات التي تم استبعادها من الملف.

3.3. وزن الحالات، Weight Cases:

يقصد بهذه العملية ترجيح المفردات وإعطاء أهمية نسبية مختلفة للحالات المختلفة بإعطائها أوزان مختلفة ولها العديد من الفوائد، فعندما لا يكون لدينا قيم جميع الحالات بل تكرارات حدوث تلك القيم أو فئات الظاهرة ، ففي هذه الحالة تصبح أهمية المفردات النسبية مقاسة بالتكرارات وتكون الأخيرة بمثابة الأوزان، والحالة الثانية عند تحليل العينات الطبقية حيث تكون في داخل البيانات عينات جزئية مختارة من طبقات مختلفة، وتقاس في هذه الحالة الأهمية النسبية للقيم بحجم كل طبقة أو نسبة تمثيلها في المجتمع، وهناك حالات أخرى كثيرة قد نحتاج إلى ترجيح البيانات بأوزان مختلفة مثل قناعتنا بضعف الثقة في بعض القيم بحيث نحتاج إلى جعل تأثير هذه القيم على نتائج التحليل محدود.

وعند اختيار الأمر Weight Cases من قائمة Data يظهر صندوق حوار بعنوان Weight Cases كما في الشكل التالي:

الشكل رقم (6-31) صندوق حوار وزن الحالات

من صندوق الحوار السابق نجد:

- Weight Cases By: ويستخدم لترجيح الحالات باستخدام المتغير الذي يحتوي على أوزان المفردات.
- Do not weight cases: عدم ترجيح الحالات (أو إلغاء عملية ترجيح المفردات).

عند الضغط على زر OK في صندوق حوار Weight Cases سيتم ترجيح الحالات ولن نلاحظ أي تغيير في شاشة محرر البيانات Data Editor أو في شاشة محرر النتائج Output Viewer ولكن هذه الأوزان ستدخل في جميع العمليات الحسابية في التحليل الإحصائي اللاحق لحين إلغاء عملية ترجيح المفردات وسنجد على شريط الحالة مكتوب Weight On .

أسئلة الفصل السادس

1- بفرض لدينا إيرادات إحدى الشركات السياحية (كما في الجدول التالي) من شهر أيلول لعام 2009 ولغاية تشرين الثاني من العام 2010، المطلوب أكتب خطوات إضافة متغير تاريخ حسب السنة والفصل والشهر خلال المدة السابقة.

إيرادات الشركة	112	110	99	95	105	107	111
	120	123	126	130	135	122	113

2- يبين الجدول التالي البيانات الخاصة بطلاب كلية السياحة الجدد في السنة الأولى:

الجنس	عدد سنوات الدراسة	استخدام الحاسوب	عدد ساعات الاستخدام
ذكر	12	نعم	3
ذكر	13	نعم	1
أنثى	12	لا	0
ذكر	12	لا	1
أنثى	15	نعم	2
أنثى	12	لا	2
ذكر	12	لا	0
أنثى	16	نعم	2

المطلوب:

أ. تعريف المتغيرات وإدخال البيانات في البرنامج الاحصائي SPSS وحفظها باسم Example2.

ب. التحقق من عملية إدخال متغير عدد سنوات الدراسة باعتبار أن عدد سنوات الدراسة يجب ألا يزيد عن 13 سنة.

ج. التحقق من عملية إدخال عدد ساعات استخدام الحاسوب باعتبار أن ساعات استخدام الحاسوب للطلاب غير المستخدمين للحاسوب يساوي صفراً.

3- بفرض أن البيانات في الجدول التالي تمثل تكملة بيانات الطلاب في السؤال الثاني:

مكان الاقامة	الاختصاص
ريف	أدبي
ريف	أدبي
مدينة	أدبي
مدينة	علمي
ريف	علمي
مدينة	أدبي
مدينة	أدبي
مدينة	علمي

أ. تعريف المتغيرات وإدخال البيانات في البرنامج الاحصائي SPSS وحفظها باسم Example3.

ب. قم بدمج حالات الملف Example2 و Example3، وحفظه باسم Merge1.

4- قم بتجميع الحالات في الملف Merge1 حسب متغير الجنس وحساب الوسط الحسابي لعدد سنوات الدراسة.

5- قم باختيار حالات الاختصاص الأدبي في الملف Merge1 ثم قم بتجميع الحالات حسب متغير مكان الإقامة وحساب الوسط الحسابي لعدد ساعات استخدام الحاسوب.

وقل رب زدني علما
جامعة دمشق
Damascus University

الفصل السابع CHAPTER 7

معالجة وتحويل البيانات
باستخدام البرنامج الإحصائي SPSS

أهداف الفصل السابع:

1. إجراء العمليات الحسابية.
2. عدّ القيم.
3. ترميز قيم المتغير.
4. ترتيب الحالات.
5. استبدال القيم المفقودة.

غالباً ما تكون بيانات الدراسة مناسبة للتحليل ولا تحتاج لتدخل الباحث، وقد يحتاج الباحث أحياناً لإجراء بعض التحويلات على هذه البيانات من خلال إنشاء متغير جديد اعتماداً على المتغيرات الموجودة في الملف الأساسي وهذه التحويلات غالباً ما تكون حسب دالة أو عملية رياضية أو مؤشر معين، أو قد تكون إعادة ترميز للبيانات، هذه التحويلات وغيرها يتيحها لنا البرنامج الإحصائي SPSS ولهذه التحويلات أهمية في التحليل والتطبيق الإحصائي وخاصة في الملفات ذات البيانات الكبيرة.

الشكل رقم (7-1) يبين قائمة Transform

فقائمة Transform توفر العديد من الأدوات لتحويل البيانات ومعالجتها حسب متطلبات العمل وتشمل على عدة أوامر تتصل بالعمليات الحسابية وإعادة الترميز لبيانات متغير سواء على نفس المتغير أو متغير آخر، كما تحوي القائمة على إعادة تصنيف المتغيرات الكمية المستمرة إلى فئات يقترحها الحاسب وإعطاء رتب للبيانات، وإنشاء سلاسل زمنية والتعامل مع القيم المفقودة، واختيار بعض البيانات عشوائياً، هذه الأوامر سيتم التطرق إليها حسب تسلسل ورودها في القائمة مع توضيح وشرح عن كيفية استخدامها.

1. إجراء العمليات الحسابية Compute:

يتيح لنا الأمر Compute إمكانية إنشاء متغير (متغيرات) جديد يتم حساب قيمه بالاعتماد على قيم متغيرات الدراسة الموجودة في ملف البيانات الأساسي، وذلك بناءً على معادلة رياضية أو مقياس إحصائي أو صيغة منطقية تصاغ وفقاً لحاجة الباحث، كذلك باستطاعة الباحث استخدام الدوال الرياضية Functions والتي يتوافر منها أكثر من 70 دالة تتضمن (دوال حسابية، إحصائية، توزيعات احتمالية)، وعند اختيار الأمر Compute من القائمة Transform يظهر صندوق الحوار التالي:

الشكل رقم (7-2) صندوق حوار إجراء العمليات الحسابية

ويتضمن صندوق الحوار السابق ما يلي:

- **Target Variable**: لتحديد اسم المتغير الهدف وهو المتغير الجديد الذي نرغب بإنشائه، وعند كتابة اسم المتغير يجب مراعاة قواعد تسمية المتغيرات.

- **Type & Label**: من خلال الضغط على زر Type & Label يظهر صندوق حوار فرعي بعنوان Compute Variable: Type and Label كما في الشكل التالي:

الشكل رقم (7-3) صندوق حوار تحديد نوع ووصف اسم المتغير

فعند تحديد اسم متغير الهدف يعرف تلقائياً على أنه متغير كمي ولكن إذا كان التعامل مع عمليات ينتج عنها متغير وصفي فلا بد من تعريف المتغير على أنه متغير وصفي بطول مناسب، كما يمكننا تحديد وصف للمتغير من خلال المستطيل Label، أو يكون التعبير الرياضي Numeric Expression الذي سنقوم بكتابته هو عنوان متغير الهدف عند نقر الدائرة المجاورة لـ Use expression as label.

- **الحقل Type**: وهو يحتوي على خيارين أيضاً:
  - إما أن يكون نوع المتغير الناتج عددي Numeric وهو الخيار الافتراضي.
  - أو أن يكون المتغير الناتج حرفياً String وهنا يجب تحديد عدد خانات المتغير Width.
- **Numeric Expression**: ويستخدم هذا المستطيل لكتابة الصيغة الرياضية المطلوب حسابها ويمكن اختيار أي من المفاتيح الموجودة في صندوق الحوار والتي تمثل الأرقام وإشارات العمليات الحسابية والعلاقات وكذلك أي من الدوال الموجودة ضمن المستطيل Function group، وعند كتابة الصيغة الرياضية لا بد من

مراعاة كتابة العمليات الحسابية داخل الكمبيوتر حيث تبدأ بحساب الدوال ثم حساب الأسس ثم عمليات الضرب والقسمة ثم عمليات الجمع والطرح ثم فك الأقواس والعمليات المنطقية.

- **If:** ويستخدم عندما يرغب الباحث بإجراء العمليات الحسابية على جزء محدد من الحالات التي تحقق شرطاً معيناً، وعند النقر على زر If يفتح صندوق حوار جديد بعنوان Compute Variable: If Cases له الشكل التالي:

الشكل رقم (7-4) صندوق حوار تحديد الدالة الشرطية

مثال 1:

لبيانات الواردة في الجدول التالي تبين دخل مجموعة من موظفي إحدى الشركات السياحية، **والمطلوب:** حساب صافي دخل الموظفين إذا علمت أنّ ضريبة الدخل 3%.

	ID	Salary
1	1	800.00
2	2	450.00
3	3	390.00
4	4	455.00
5	5	420.00
6	6	422.00
7	7	489.00
8	8	390.00
9	9	444.00
10	10	519.00

لتنفيذ ذلك نتبع الخطوات التالية:

1- من قائمة Transform نختار Compute .

2- يظهر صندوق حوار بعنوان Compute Variable.

3- نحدد اسم متغير الهدف Target Variable والذي هو عبارة عن صافي الدخل وليكن netSalary.

4- نكتب معادلة إيجاد صافي الدخل في مربع Numeric Expression والتي هي عبارة عن الراتب مطروحاً منه الضريبة، كما في الشكل التالي:

الشكل رقم (7-5) صندوق حوار حساب دالة صافي الدخل

salary-(0.03*salary)

5- عند النقر على زر OK في صندوق حوار Compute Variable نحصل على نتائج الحل في شاشة محرر البيانات Data Editor والتي تتضمن عموداً جديداً (متغيراً) معنوناً بـ netSalary وهو عبارة عن صافي الراتب لكل موظف، كما موضح في الشكل التالي:

الشكل رقم (7-6) شاشة محرر البيانات بعد حساب صافي الدخل

	ID	Salary	netSalary
1	1	800.00	776.00
2	2	450.00	436.50
3	3	390.00	378.30
4	4	455.00	441.35
5	5	420.00	407.40
6	6	422.00	409.34
7	7	489.00	474.33
8	8	390.00	378.30
9	9	444.00	430.68
10	10	519.00	503.43

مثال 2:

بفرض لدينا البيانات الخاصة بموظفي إحدى الشركات السياحية، **والمطلوب:** حساب صافي رواتب الموظفين إذا رغبت إدارة الشركة زيادة رواتب الموظفين الذين تقل رواتبهم عن 500 وحدة نقدية بمقدار 5%.

	ID	Salary	Age
1	1	800.00	51.00
2	2	450.00	46.00
3	3	390.00	24.00
4	4	455.00	47.00
5	5	420.00	33.00
6	6	422.00	32.00
7	7	489.00	49.00
8	8	390.00	25.00
9	9	444.00	42.00
10	10	519.00	48.00

لتنفيذ ذلك، اتبع الخطوات التالية:

1. من قائمة Transform نختار Compute.
2. يظهر صندوق الحوار بعنوان Compute Variable.
3. نحدد اسم متغير الهدف Target Variable والذي هو عبارة عن صافي الدخل وليكن netSalary.
4. ننقر على الزر if في صندوق حوار Compute Variable فيظهر صندوق حوار فرعي بعنوان Compute Variable :if cases وبما أننا نريد زيادة رواتب الموظفين الذين تقل رواتبهم عن 500 وحدة نقدية أي جزء من الحالات، نختار Include if case satisfied condition ونكتب الشرط التالي (Salary<=500) ثم ننقر على زر Continue للعودة إلى صندوق الحوار السابق Compute Variable وكذلك نكتب المعادلة الحسابية في مربع Numeric Expression كما يلي: (Salary + 0.05 * Salary)
5. ثم نضغط على المفتاح ok، ستظهر شاشة محرر البيانات Data Editor التي تحوي على المتغير netSalary.

الشكل رقم (7-7) يوضح خطوات المثال مع قاعدة البيانات الجديدة

مثال3:

لنفترض لدينا ملف البيانات والذي يحوي المتغيرات التالية: اسم الطالب (name) – الدرجة في مقرر الرياضيات (mathematics) – الدرجة في مقرر الإحصاء (statistics) – الدرجة في مقرر الإدارة (management) – الدرجة في مقرر القانون (law)، والمطلوب: إيجاد معدل الطلاب في هذه المقررات.

	name	mathematics	statistics	management	law
1	Ahmad	88.00	87.00	70.00	84.00
2	Mosa	77.00	80.00	77.00	85.00
3	Rahf	94.00	90.00	88.00	89.00
4	Zeid	80.00	79.00	89.00	79.00

لتنفيذ ذلك نتبع الخطوات التالية:

1- من قائمة Transform نختار Compute.

2- يظهر صندوق حوار بعنوان Compute Variable.

3- ندخل اسم المتغير الجديد وليكن Average والذي هو عبارة عن معدل الطلاب في خانة Target Variable.

4- في مستطيل Functions Group نختار دالة Statistical.

5- من مستطيل Functions and Special Variables نختار الدالة الفرعية Mean والتي تعبر عن المتوسط.

6- ننقل الدالة بعد تحديدها بواسطة السهم على يمين Functions ستظهر الدالة في خانة Numeric Expression.

7- ندخل أسماء المتغيرات إلى داخل قوسين بحيث نستبدل ؟ بأسماء المتغيرات بحيث يكون الفصل بين هذه المتغيرات باستخدام الفاصلة، كما يلي:

(mathematics, statistics, management ,law) .

8- ثم نضغط على المفتاح OK، ستظهر شاشة محرر البيانات Data Editor التي تحوي على المتغير الجديد Average.

الشكل رقم (7-8) يوضح الخطوات المثال مع قاعدة البيانات الجديدة

2. عدّ القيم Count Value Within Cases:

يقوم هذا الأمر بإيجاد عدد المرات التي يتكرر فيها رقم أو رمز معين في عدد من المتغيرات، ولهذا الأمر أهمية تطبيقية خاصة في الاستبيانات الإحصائية لحساب عدد القيم المتشابهة لمجموعة من المتغيرات ولكل حالة من حالات الاستبانة، وعند اختيار الأمر Count Value within Cases من القائمة Transform يظهر صندوق الحوار التالي:

الشكل رقم (7-9) صندوق حوار عد القيم داخل الحالات

من صندوق حوار عد القيم داخل الحالات (Count Occurrences of Value within Cases) نجد:

1) في قسم Target Variable: نضع اسم المتغير الجديد الذي يقوم بعدّ الحالات الذي نريد إنشاءه.
2) في قسم Target Label: نضع عنوان المتغير الذي نريد إنشاءه إذا تطلب الأمر.
3) في قسم Numeric Variable: ننقل المتغير أو المتغيرات التي نرغب بعد حالاتها.
4) عند النقر على زر Define Variable يظهر صندوق حوار بعنوان Count values within cases: values to count ، والذي له الشكل التالي:

الشكل رقم (7-10) صندوق حوار تعريف الحالات

يهدف صندوق الحوار السابق لتعريف الحالات التي نرغب بعدها والذي يقسم إلى قسمين:

القسم الأول: المعنون بـ Value ويستخدم لتعريف الحالات التي سنعدّها والتي تكون كما يلي:

- Value: لعدّ قيم منفصلة.
- System-missing: لعدّ القيم المفقودة.
- System-or user-missing: لعدّ القيم المفقودة أو القيم المفقودة التي يعرفها المستخدم.
- Range: لعدّ القيم التي تقع ضمن مجال من القيم بحيث نحدد الحد الأدنى والحد الأعلى لهذا المجال.
- Range, LOWEST through Value: لعدّ القيم التي تقع من قيمة محددة وأصغر منها.
- Range, Value through HIGHEST: لعدّ القيم التي تقع أكبر من قيمة محددة.

القسم الثاني: المعنون بـ Values to Count: وهو قسم خاص بالقيم التي تم تعريفها في القسم الأول.

5) **If**: ويستخدم عندما يرغب الباحث بعدّ الحالات التي تحقق شرطاً معيناً، وعند النقر على زر if يظهر صندوق حوار بعنوان: Count Occurrence: if Cases. له الشكل التالي:

الشكل رقم (7-11) صندوق حوار تعريف الحالات

3. ترميز قيم المتغير Recode:

تتطلب بعض إجراءات التحليل الإحصائي أن تكون قيم المتغير على شكل مجموعات قليلة العدد تحوي عدداً من الحالات، فليس من المفيد أن تكون قيم المتغير كثيرة مثلاً، عند تكوين جدول تكراري يمثل تكرارات قيم ظاهرة معينة يصعب الاستفادة من الجدول إذا احتوى على 50 قيمة وبالتالي لا بد من معالجة هذه المشكلة عن طريق تجميع القيم المتقاربة والاستعاضة عنها بقيم محددة أي أن الباحث لا بد له في بعض الحالات من إعطاء عدد محدود من القيم أو الرموز لأحد المتغيرات ذات القيم الكثيرة ليسهل فهم البيانات والتعامل معها إحصائياً.

إن هذه الخيارات والإمكانيات يوفرها البرنامج الإحصائي SPSS عن طريق الخيار Recode ويتضمن هذا الأمر عدة أنواع من الترميز:

النوع الأول: الترميز في المتغير نفسه Recode Into Same Variable.

النوع الثاني: الترميز في المتغير جديد Recode Into different Variable.

النوع الثالث: الترميز في متغير جديد تلقائياً Automatic Recode.

النوع الرابع: تحويل المتغيرات الرقمية إلى مجموعات Visual Binning.

1.3. ..Recode into Same Variable:

يتيح البرنامج الإحصائي SPSS إعادة ترميز القيم للبيانات الرقمية Numeric والحرفية String مع مراعاة ما يلي:

- ✓ في حالة إعادة ترميز قيم أكثر من متغير فإنه من الضروري أن تكون المتغيرات من نفس النوع، لأنه لا يمكن إعادة ترميز قيم متغيرات رقمية وحرفية معاً.
- ✓ عند استخدام أمر إعادة ترميز القيم على نفس المتغيرات في البيانات الأساسية فإن البرنامج الإحصائي SPSS سيقوم بالكتابة على القيم القديمة للمتغير واستبدالها بالقيم الجديدة ولا يمكن التراجع عن هذا الأمر.

عند اختيار الأمر Recode into Same Variable من قائمة Transform، يظهر صندوق حوار بعنوان Recode into Same Variable.

الشكل رقم (7-12) صندوق حوار إعادة الترميز على نفس المتغير

يتضمن صندوق الحوار السابق ما يلي:

1) في قسم Numeric Variable نضع المتغير الذي نريد إعادة ترميز قيمه مع العلم أنه يمكننا اختيار أكثر من متغير واحد بشرط أن تكون جميعها من نفس النوع.

2) عند النقر على زر Old and New Values يظهر صندوق حوار بعنوانRecode into---Variable: Old and New Values كما يبينه الشكل التالي:

الشكل رقم (7-31) صندوق حوار ترميز القيم الجديدة والقديمة

القسم الأول والمعنون Old values: وهي القيم التي نريد إعادة ترميزها وتحوي الخيارات التالية:

- Value: لإعادة ترميز قيم منفصلة.
- System-missing: لإعادة ترميز القيم المفقودة.
- System-or user-missing: لإعادة ترميز القيم المفقودة أو القيم المفقودة التي يعرفها المستخدم.
- Range: لإعادة ترميز القيم التي تقع ضمن مجال من القيم بحيث نحدد الحد الأدنى والحد الأعلى لهذا المجال.
- Range, LOWEST through Value: لإعادة ترميز القيم التي تقع من قيمة محددة وأصغر منها.

- Range, Value through HIGHEST: لإعادة ترميز القيم التي تقع أكبر من قيمة محددة.
- All other values: لإعادة ترميز القيم غير المحددة سابقاً.

القسم الثاني والمعنون New Values: وهي القيم التي تعطى للقيم التي تم اختيارها في القسم الأول.

3) عند النقر على زر If يظهر صندوق حوار بعنوان Recode into-- Variable: if Cases وهو لكتابة شرط أو عبارة شرطية على الحالات التي نريد إعادة ترميزها.

4) ثم ننقر على Continue ثم Change ثم OK.

مثال 4:

بفرض لدينا بيانات لدخل مجموعة من الموظفين مع رقم كل موظف كما في ملف البيانات المجاور:

والمطلوب: أعد ترميز المتغير Salary حسب الفئات التالية:

الفئة	الرمز
أقل من 249	1
250 –499	2
500 –750	3
أكثر من 750	4

	number	salary
1	1.00	850.00
2	2.00	600.00
3	3.00	510.00
4	4.00	280.00
5	5.00	300.00
6	6.00	220.00
7	7.00	900.00
8	8.00	460.00
9	9.00	350.00
10	10.00	530.00
11	11.00	650.00
12	12.00	880.00
13	13.00	950.00
14	14.00	160.00
15	15.00	200.00

لتنفيذ المطلوب نتبع الخطوات التالية:

1- من قائمة Transform اختر الأمر Recode into Same Variable.

2- يظهر صندوق حوار بعنوان Recode into Same Variable.

3- نضع المتغير المراد إعادة ترميزه وهو Salary في قسم :Numeric Variables .

4- ننقر على خيار Old and New Values يظهر صندوق حوار بعنوان: Recode into Same Variable: Old and New Values ونضع في قسم Range, LOWEST through Value القيمة 249 وفي قسم New Value نضع القيمة 1 ثم نضغط على Add، ومن أجل إدخال الفئة من 250 إلى 499 نفعل الخيار Range ثم ندخل ضمن الحد الأدنى للفئة 250 وفي الحقل تحت الخيار Through ندخل الحد الأعلى وهو 499 وفي قسم New Value نضع القيمة 2 ثم نضغط على Add ، ثم نعيد الخطوة السابقة من أجل إدخال الفئة 500 إلى 749 وفي قسم New Value نضع القيمة 3 ثم نضغط على Add ، ولإدخال الفئة الأخيرة نفعل الخيار Range value Through Highest ثم ندخل 750 أو نفعل الخيار All Other Values وفي قسم New Value نضع القيمة 4 ثم نضغط على Add، حتى نحصل على الترتيب التالي كما في صندوق الحوار:

الشكل رقم (7-14) صندوق حوار ترميز القيم الجديدة والقديمة في المثال رقم 4

5- نختار Continue ثم Change ثم OK فنحصل على قاعدة البيانات كما في الشكل:

2.3. ..Recode into Different Variable:

يتيح البرنامج الإحصائي SPSS إعادة ترميز قيم البيانات الرقمية Numeric والحرفية String على متغير جديد، وهو وما يميز هذا الخيار عن الخيار السابق (Recode into same Variables) بحيث يمكننا إعادة الترميز على متغير جديد باسم جديد بحيث لا يتم إحلال الرموز مكان القيم الأساسية مع الإبقاء على قيم المتغير السابق مع ملاحظة أنه لا يمكن إعادة ترميز أكثر من متغير واحد في هذه الحالة.

عند اختيار الأمر Recode into Different Variable من قائمة Transform، يظهر صندوق حوار بعنوان Recode into Different Variable، كما يبينه الشكل التالي:

الشكل رقم (7-15) صندوق حوار إعادة الترميز على متغير جديد

نلاحظ أن صندوق الحوار لا يختلف إلا قليلاً عن سابقه وهذا الاختلاف يكمن في قسم Output Variable لوضع اسم المتغير الجديد وفي قسم العنوان (Label) نضع عنوان المتغير الجديد ثم نتبع الخطوات التي تمت في الأمر Recode into same Variable ثم نضغط على Change ثم OK.

مثال5:

بفرض لدينا بيانات لدخل مجموعة من الموظفين مع رقم كل موظف كما في ملف البيانات المجاور:

والمطلوب: أعد ترميز المتغير Salary بمتغير آخر جديد حسب الفئات التالية:

الفئة	الرمز
أقل من 249	1
250 – 499	2
500 – 750	3
أكثر من 750	4

	number	salary
1	1.00	850.00
2	2.00	600.00
3	3.00	510.00
4	4.00	280.00
5	5.00	300.00
6	6.00	220.00
7	7.00	900.00
8	8.00	460.00
9	9.00	350.00
10	10.00	530.00
11	11.00	650.00
12	12.00	880.00
13	13.00	950.00
14	14.00	160.00
15	15.00	200.00

لتنفيذ المطلوب نتبع الخطوات التالية:

1- من قائمة Transform اختر الأمر Recode Into Different Variable.

2- يظهر صندوق حوار بعنوان Recode Into Different Variable.

3- نضع المتغير المراد إعادة ترميزه وهو Salary في قسم :Numeric Variables .

4- في قسم Output Variable نضع اسم المتغير الجديد وهو Num_Salary وفي قسم العنوان نضع عنوان المتغير الجديد Number Salary.

5- ننقر على خيار Old and New Values يظهر صندوق حوار بعنوان: Recode into Different Variable: Old and New Values ونضع في قسم Range, LOWEST through Value القيمة 249 وفي قسم New Value نضع القيمة 1 ثم نضغط على Add، ومن أجل إدخال الفئة من 250 إلى 499 نفعل الخيار Range ثم ندخل ضمن الحد الأدنى للفئة 250 وفي الحقل تحت الخيار Through ندخل الحد الأعلى وهو 499 وفي قسم New Value نضع القيمة 2 ثم نضغط على Add ، ثم نعيد الخطوة السابقة من أجل إدخال الفئة 500 إلى 749 وفي قسم New Value نضع القيمة 3 ثم نضغط على Add ، ولإدخال الفئة الأخيرة نفعل الخيار Range value Through Highest ثم ندخل 750 أو نفعل الخيار All Other Values وفي قسم New Value نضع القيمة 4 ثم نضغط على Add، حتى نحصل على الترتيب التالي كما في صندوق الحوار:

6- ثم ننقر على Continue ثم Change في صندوق الحوار المعنون Recode into Different Variable ثم OK، فنحصل على قاعدة البيانات كما في الشكل التالي:

3.3. إعادة ترميز القيم تلقائياً ..Automatic Recode:

يتيح الخيار إعادة ترميز القيم تلقائياً إمكانية ترميز المتغيرات الرقمية والحرفية إلى أعداد منطقية صحيحة، وتكمن الفائدة من هذا الخيار أنه في أساليب التحليل الإحصائي المتقدم لا يتم تفعيل خياراتها إذا كانت القيم حرفية أو قيم رقمية مستمرة، لذا يجب علينا إعادة ترميز مثل هذه القيم لتصبح على شكل مجموعات، وعند استخدام هذا الخيار يجب ملاحظة ما يلي:

- ✓ المتغير الجديد الذي سنحصل عليه سيكون من النوع الرقمي والمقياس الرتبي وتأخذ قيمه عنوانيها قيم البيانات الأساسية.
- ✓ إذا وجد قيم مفقودة لن يتم إعادة ترميزها.
- ✓ بالنسبة للمتغيرات الحرفية التي يتم إعادة ترميزها تلقائياً سيتم ترتيبها أبجدياً تصاعدياً أو تنازلياً حسب رغبة المستخدم ولكن مع ملاحظة أن القيم الحرفية التي تتشابه لكنها تختلف بطريقة بداية الكتابة بحرف صغير أو كبير سيتم التعامل على أن القيم ذات الحروف الأبجدية الصغيرة تأتي قبل القيم ذات الحروف الأبجدية الكبيرة في الترتيب.

ولكن قبل البدء بتنفيذ مثال عن حالة إعادة ترميز القيم تلقائياً لنناقش صندوق الحوار الذي يظهر:

1) من قائمة Transform اختر الأمر ..Automatic Recode.
2) يظهر صندوق حوار بعنوان: Automatic Recode.
3) في قسم Variable-> New Name نضع المتغير الذي نريد إعادة ترميزه تلقائياً.
4) في قسم New Name نضع اسم المتغير الجديد الذي نرغب بوضع قيم البيانات المعاد ترميزها تلقائياً.
5) في قسم Recode Starting From نختار نوع إعادة الترميز بشكل تنازلي (Lowest Value) أو تصاعدي (Highest Value).
6) ثم نضغط على Add New Name ليظهر الاسم الجديد في Variable-> New Name ثم على OK.

الشكل رقم (7-16) صندوق حوار إعادة الترميز التلقائي

مثال 6:

بفرض لدينا بيانات الدخل لمجموعة من الموظفين كما في المثال 5، والمطلوب: أعد ترميز المتغير Salary باستخدام خيار إعادة الترميز التلقائي.

لتنفيذ المطلوب نتبع الخطوات التالية:

أ. من قائمة Transform اختر الأمر ..Automatic Recode.

ب. يظهر صندوق حوار بعنوان: Automatic Recode.

ج. ننقل متغير الدخل Salary قسم Variable -> New Name.

د. في قسم New Name نضع اسم المتغير الجديد وهو Auto_Salary.

هـ. نختار نوع الترميز Lowest Value من قسم Recode Starting From.

7) ثم نضغط على Add New Name ليظهر الاسم الجديد في Variable-> New Name ثم على OK.

8) ثم على OK فيظهر ملف البيانات وملف المخرجات على الشكل التالي:

4.3. تحويل البيانات الرقمية إلى مجموعات ..Visual Binning:

يتيح لنا البرنامج الإحصائي SPSS إمكانية تحويل البيانات الرقمية وخاصة البيانات المستمرة إلى مجموعات تصنيفية حيث يتم تصنيف البيانات إلى فئات مناسبة بناءً على مدى البيانات وعدد الفئات، وفق الخطوات التالية:

1- من قائمة Transform نختار ..Visual Binning .

2- يظهر صندوق حوار بعنوان: Visual Binning.

3- نقوم بإدخال المتغيرات المرغوب إعادة تصنيفها إلى مستطيل Variables to Bin، ثم نضغط على الزر Continue فيظهر صندوق حوار جديد بعنوان Visual Binning.

4- من قسم Scanned Variable List نحدد المتغير المرغوب بإعادة تصنيفه فيظهر اسم المتغير في المستطيل Current Variable.

5- نقوم بكتابة اسم المتغير الجديد في المستطيل Binned Variable.

6- نضغط على Make Cutpoints فيظهر صندوق حوار بعنوان Make Cutpoints ويفيد صندوق الحوار الذي يظهر لتحديد عدد الفئات التي تظهر ومدى كل فئة وموقع كل فئة.

7- ثم نضغط على زر Apply .

8- نضغط على مستطيل Make Label ليظهر لدينا وصف تلقائي للفئات.

مثال 7:

بفرض لدينا بيانات لدخل مجموعة من الموظفين مع رقم كل موظف كما في المثال 5، والمطلوب: أعد ترميز المتغير Salary باستخدام الخيار Visual Binning.

لتنفيذ المطلوب نتبع الخطوات التالية:

1) من شريط القوائم اختر Transform→ Visual Binning.. .
2) يظهر صندوق حوار بعنوان: Visual Binning.
3) نقوم بإدخال متغيرات الدخل Salary إلى مستطيل Variables to Bin.
4) ثم نضغط على الزر Continue فيظهر صندوق حوار جديد بعنوان Visual Binning.
5) من قسم Scanned Variable List نحدد متغير الدخل Salary فيظهر اسم المتغير في المستطيل Current Variable.
6) نقوم بكتابة اسم المتغير الجديد في المستطيل Binned Variable وليكن Bin_Salay كما في صندوق الحوار التالي:

الشكل رقم (7-17) صندوق حوار Visual Binning

حيث يظهر لنا صندوق الحوار السابق أصغر قيمة ضمن المستطيل Minimum حيث ظهرت في مثالنا القيمة 160 ويظهر أيضاً أكبر ضمن المستطيل Maximum حيث تبلغ هنا 950 ، وأيضاً يظهر لنا إذا كان لدينا قيم مفقودة أم لا.

7) نضغط على Make Cutpoints فيظهر صندوق حوار بعنوان Make Cutpoints ويظهر فيه:

- ✓ نقوم بتحديد موقع نقطة الفاصل الأولى في مستطيل First Cutpoint Location، ويجب أن تكون أصغر من أقل قيمة حتى تتضمن أصغر قيمة ضمن التصنيف وليكن في مثالنا 150.
- ✓ نحدد عدد الفئات المرغوب إظهارها في مستطيل Number Of Cut Points، وليكن عدد الفئات هنا 5 سنلاحظ أن عدد الفئات التي ستظهر 6 فئات.
- ✓ نضغط على Apply، فنعود إلى صندوق الحوار السابق Visual Binning.

8) نضغط على مستطيل Make Label ليظهر لدينا وصف تلقائي للفئات التي قمنا بإنشائها.

9) نضغط على زر OK فنحصل على ملف البيانات التالي:

	number	salary	Bin_Salary
1	1.00	850.00	790.00+
2	2.00	600.00	470.00 - 629.00
3	3.00	510.00	470.00 - 629.00
4	4.00	280.00	150.00 - 309.00
5	5.00	300.00	150.00 - 309.00
6	6.00	220.00	150.00 - 309.00
7	7.00	900.00	790.00+
8	8.00	460.00	310.00 - 469.00
9	9.00	350.00	310.00 - 469.00
10	10.00	530.00	470.00 - 629.00
11	11.00	650.00	630.00 - 789.00
12	12.00	880.00	790.00+
13	13.00	950.00	790.00+
14	14.00	160.00	150.00 - 309.00
15	15.00	200.00	150.00 - 309.00

4. ترتيب الحالات Rank Cases:

يتيح البرنامج الإحصائي SPSS إمكانية استحداث متغير يحوي رتباً لحالات متغير كمي موجود في قاعدة البيانات وذلك من خلال استخدام الأمر Rank Cases مع ملاحظة أن البرنامج سوف يقوم تلقائياً بإعداد خصائص المتغير الجديد (الاسم – الوصف – المقياس) بناءً على خصائص المتغير الأساسي كما توضح لنا هذه الخصائص في جدول ضمن شاشة المخرجات وإحدى فوائد هذا الخيار أنه في كثير من طرق العرض والتحليل الإحصائي قد نرغب في التعامل مع الرتب الخاصة بالمتغير وليست مع قيم المتغير، مع ملاحظة أن برنامج الـ SPSS يتيح الآتي:

- ✓ ترتيب الحالات تصاعدياً أو تنازلياً.
- ✓ تنسيق الترتيب أو التصنيف على هيئة مجموعات فرعية حسب متغير آخر.

وعند اختيار الأمر ..Rank Cases من القائمة Transform يظهر صندوق الحوار بعنوان Rank Cases الذي له الشكل التالي:

الشكل رقم (7-18) صندوق حوار ترتيب الحالات

1- المتغير المرغوب بترتيب حالته ننقله إلى مستطيل Variables.

2- في قسم BY نحدد المتغير الذي نرغب على أساسه تنسيق الترتيب أو التصنيف على هيئة مجموعات فرعية (وهذه الخيار اختياري) وذلك من خلال اختيار متغير أو أكثر لذلك.

3- في قسم Assign Rank 1 to نحدد أسلوب الترتيب إما تصاعدياً (Smallest Value) أو تنازلياً (Largest Value).

4- عند التأشير على خيار Display summary tables سيظهر جدول في شاشة المخرجات يحوي خصائص المتغير الجديد الذي يحوي ترتيب الحالات.

5- عند النقر على الزر Rank Types.. يظهر صندوق حوار بعنوان: Rank Types :Cases. ويمكننا هنا تحديد نوع الرتب وكذلك تحديد طريقة الترتيب المطلوبة فهنا يمكن إنتاج رتب بسيطة أو قيم معيارية أو نسب مئوية أو معدلات أو أن تكون الرتب الناتجة بناء على أحد الصيغ المتاحة لإنتاج الرتب، كما يمكننا حساب القيم المعيارية وكذلك تقدير المعدل فالعديد من الصيغ التي يمكن اتباعها عادة في تقدير المعدلات المتاحة للاستخدام هنا وهي: Blom, Tukey, Rankit, Van der Waerden .

الشكل رقم (7-19) صندوق حوار لتحديد نوع الرتب وطريقة ترتيب الحالات

6- عند النقر على خيار Ties يظهر صندوق حوار بعنوان: Rank Cases: Ties يمكننا التحكم في كيفية التعامل مع ترتيب الحالات المتساوية بالقيمة فأغلب الطرق الإحصائية تعطي قيمة المتوسط للرتب المتساوية كرتبة موحدة لهذه القيم، إلا أن البرنامج الإحصائي SPSS يمكننا من تغير هذه الطريقة لإعطاء قيماً مختلفة للرتب المتساوية فالطرق المتاحة في النظام للتعامل مع القيم المتساوية هي إما بإعطائها المتوسط الحسابي (Mean) أو أصغر قيمة لجميع الرتب المتساوية (Low) أو أكبر قيمة لجميع الرتب المتساوية (High) أو إعطائها قيماً متسلسلة حسب ترتيب ورودها في قاعدة البيانات (Sequential rank To unique values).

الشكل رقم (7-20) صندوق حوار لترتيب الحالات المتساوية بالقيمة

مثال 8:

بفرض لدينا درجات مجموعة من الطلاب مع رقم كل طالب كما في ملف البيانات المجاور:

	Number	Degree	Gender
1	1.00	65.00	male
2	3.00	88.00	male
3	5.00	47.00	male
4	7.00	66.00	male
5	9.00	77.00	male
6	11.00	98.00	male
7	13.00	62.00	male
8	2.00	79.00	female
9	4.00	48.00	female
10	6.00	65.00	female
11	8.00	78.00	female
12	10.00	92.00	female
13	12.00	64.00	female

والمطلوب:

- إعطاء رتب لدرجات الطلاب.
- ثم إعطاء رتب لدرجات الطلاب حسب الجنس.

لتنفيذ المطلوب الأول نتبع الخطوات التالية:

1) من شريط القوائم اختر Transform→ Rank Cases.. .
2) يظهر صندوق حوار بعنوان: Rank Cases.
3) نقوم بإدخال متغيرات الدخل Degree إلى مستطيل Variables.
4) في قسم Assign Rank 1 to نحدد أسلوب الترتيب التصاعدي Smallest value.
5) ثم نضغط على OK، فيظهر متغير جديد باسم RDegree حيث يظهر رتبة حالات متغير Degree.

لتنفيذ المطلوب الثاني نتبع الخطوات التالية:

1) من شريط القوائم اختر Transform→ Rank Cases.. .
2) يظهر صندوق حوار بعنوان: Rank Cases.
3) نقوم بإدخال متغيرات الدخل Degree إلى مستطيل Variables.
4) نقوم بإدخال متغير الجنس Gender إلى المستطيل المعنون بـ BY.
5) في قسم Assign Rank 1 to نحدد أسلوب الترتيب التصاعدي Smallest value.
6) ثم نضغط على OK، فيظهر متغير جديد باسم RAN001 حيث يظهر رتبة حالات متغير Degree وفقاً لمتغير الجنس Gender. فتظهر قاعدة البيانات كما يلي:

5. استبدال الحالات المفقودة Replace Missing Value:

في كثير من الأحيان يواجه المستخدم بعض القيم المفقودة في البيانات نتيجة لعدم استجابة بعض أفراد العينة عن بعض الأسئلة، وهو ما يتسبب بحدوث مشاكل في التحليل الإحصائي للبيانات وإن عملية استبدال القيم المفقودة تتم بإحدى الطرق

الإحصائية، ومن أجل هذه الحالة يتيح لنا البرنامج الإحصائي SPSS أمر التعامل مع بيانات كهذه البيانات فنستبدل القيم المفقودة بعدد من الإجراءات الإحصائية عن طريق:

- ✓ وسط العينة أو وسط القيم المجاورة.
- ✓ وسيط العينة أو وسيط القيم المجاورة.
- ✓ الاستكمال الخطي حيث تستخدم القيم القبلية والبعدية للقيم المفقودة في تقدير القيم المفقودة.
- ✓ استخدام معادلة الخط للعينة في تقدير القيم المفقودة.

وعند اختيار الأمر Replace Missing Value .. من القائمة Transform يظهر صندوق الحوار بعنوان ..Replace Missing Value الذي له الشكل التالي:

الشكل رقم (7-21) صندوق حوار استبدال الحالات المفقودة

المتغير المراد استبدال الحالات المفقودة فيه نضعه في قسم New Variable.

1- في المستطيل المعنون بـ Name and Method نحدد اسم المتغير الجديد الذي تم إنشاؤه دون حالات مفقودة حيث نلاحظ أن اسم المتغير الجديد يحمل اسم المتغير الأساسي متبوعة بـ شحطة سفلية Underscore ثم بعد ذلك رقم متسلسل.

2- من قسم Method نحدد أسلوب التعامل مع القيم المفقودة المتاحة في البرنامج الإحصائي:

Series mean: يتم إضافة متوسط كافة القيم الموجودة في المتغير للقيم المفقودة حيث يكون المتوسط كالتالي:

- ✓ Mean nearby point: وهنا يتم تحديد عدد القيم التي سيتم حساب المتوسط لها قبل القيم المفقودة وبعدها ثم يتم استبدال القيم المفقودة.
- ✓ Median nearby point: وهنا يتم تحديد عدد القيم التي سيتم حساب الوسط لها قبل القيم المفقودة وبعدها ثم يتم استبدال القيم المفقودة.
- ✓ Linear interpolation: وهنا يتم حساب المتوسط للقيمة الأعلى والأسفل من القيمة المفقودة ثم يتم استبدال القيم المفقودة.
- ✓ Linear Tread at point: يتم إجراء علاقة تربط كافة القيم لمتغير مع سلسلة زمنية يتم إعطاء أول رقم فيها 1 ثم الرقم الثاني 2 وهكذا يتم تحديد القيم المفقودة على خط العلاقة.

مثال 9

بفرض لدينا المتغير الذي تظهر بياناته كما في الجدول المجاور:

حيث نلاحظ أن الحالات رقم 4 و 8 هي حالات مفقودة والمطلوب تعويض عن هذه القيم.

لتنفيذ المطلوب نتبع الخطوات التالية:

1) من شريط القوائم اختر Transform→ Replace Missing Value.. .
2) يظهر صندوق حوار بعنوان: Rank Cases.
3) في قسم New Variable نضع Variable.
4) في المستطيل المعنون بـ Name and Method نلاحظ أن اسم المتغير الجديد أصبح نفس الاسم السابق مع إضافة الرقم 1_ فيكون Variable_1.
5) نحدد طريقة Mean nearby point من Series mean .
6) ثم نضغط على Change ثم OK كما يبينه ملف البيانات التالي:

	Variable	Variable_1
1	2.00	2.00
2	3.00	3.00
3	2.00	2.00
4	.	4.00
5	6.00	6.00
6	3.00	3.00
7	6.00	6.00
8	.	4.00
9	3.00	3.00
10	7.00	7.00

أسئلة الفصل السابع

1. بفرض لدينا الدرجات التالية التي تمثل علامات طلاب كلية السياحة في مقرر الاحصاء:

60	59	58	57	57	56	52	50	48	47
67	67	66	65	64	63	63	60	61	60
75	74	73	73	72	70	71	70	69	69
81	81	80	78	79	79	77	77	76	75
88	87	88	86	86	85	84	83	83	82

والمطلوب:

أ. عرف متغير الدرجة، وأدخل البيانات السابقة في البرنامج الإحصائي SPSS.

ب. إعادة ترميز القيم باستخدام الأمر Visual Binning.

ج. إعادة ترميز القيم باستخدام الأمر Recode into Different Variable كما يلي:

القيم	الرمز
من 49 وأقل	1
من 50 إلى 59	2
من 60 إلى69	3
من 70 إلى 79	4
من 80 وأكبر	5

د. بفرض أن مدرس المادة قرر زيادة 3 درجات للطلاب الذين تقل درجتهم عن 60، استخدم الأمر Compute لتنفيذ الأمر.

وقل رب زدني علما
جامعة دمشق
Damascus University

الفصل الثامن CHAPTER 8

وصف المتغيرات النوعية

أهداف الفصل الثامن:

1. التعرف على كيفية العرض الجدولي لمتغير نوعي واحد.
2. التعرف على كيفية العرض الجدولي لمتغيرين نوعيين.

بعد أن يجمع الباحث بياناته ينتقل إلى عملية عرضها عرضاً مبسطاً يسهل على القارئ قراءتها، لذلك يقوم في المرحلة الأولى بتصنيف هذه البيانات إلى فئات ثم يظهر عدد الحالات التي تدخل في كل فئة من هذه الفئات، وهذا ما يعرف بالتكرار (التكرار المطلق) حيث يكون مجموع هذه التكرارات مساوياً لحجم العينة (مع الأخذ بالحسبان القيم المفقودة)، ويضاف عمود آخر لعرض التكرارات النسبية التي تعبر عن نسبة أفراد الفئة إلى مجموع أفراد العينة بحيث يكون مجموع نسب التكرارات مساوياً للمئة.

1. الجدول التكراري لمتغير نوعي واحد:

وهي الجداول التي تتوزع فيها البيانات حسب متغير واحد (صفة واحدة) ويتألف عادة من عمودين الأول يمثل فئات المتغير المدروس، والعمود الثاني يمثل عدد المفردات التابعة لكل فئة وتدعى التكرارات، ومن هذا المنطلق يمكن تسمية الجدول بجدول التوزيع التكراري Frequency Distribution Table، كما يضاف لهذا الجدول عمود ثالث يمثل النسبة المئوية لفئات المتغير ،

من خلال الأمر Frequencies في البرنامج الإحصائي SPSS يمكننا إنشاء جدول تكراري لوصف توزيع مفردات العينة حسب متغير نوعي واحد (اسمي أو نوعي).

من قائمة Analyze نختار الأمر Descriptive Statistics ثم نختار الأمر الفرعي Frequencies فيظهر صندوق حوار له الشكل التالي:

الشكل رقم (8-1) صندوق حوار الأمر Frequencies

ويتضمن Frequencies ما يلي:

1.1. الجداول التكرارية:

وتعرض بالنقر على المربع الفارغ بجانب Display Frequency tables، ويتضمن الجدول التكراري ما يلي:

a) العمود الأول المسمى Frequency: والذي يعني تكرار أفراد العينة في كل فئة من فئات المتغير.

b) العمود الثاني المسمى Percent: والذي يتضمن النسب المئوية لأفراد العينة في كل فئة من فئات المتغير.

c) العمود الثالث المسمى Valid Percent: والذي يتضمن النسب المئوية لأفراد العينة في كل فئة من فئات المتغير بعد استبعاد البيانات المفقودة Missing.

d) العمود الرابع المسمى Cumulative Percent: يمثل النسب المئوية التراكمية لأفراد العينة في كل فئة من فئات المتغير.

2.1. مقاييس الإحصاء الوصفي:

وتعرض من خلال النقر على الزر Statistics فيظهر صندوق حوار فرعي بعنوان Frequencies Statistics له الشكل التالي:

الشكل رقم (8-2) صندوق حوار الأمر الفرعي Statistics

إنّ معظم هذه القياسات الإحصائية غالباً ما تستخدم مع متغيرات من النوع الترتيبي Ordinal، أو الكمي، ولا تستخدم مع متغيرات نوعية، ونظراً لأنّ الأمر Frequencies يمكن استخدامه مع متغيرات من النوع الاسمي الترتيبي أو الكمي في بعض الحالات فقد وضعت هذه الخيارات ضمن الإجراء المذكور، وسنتحدث عن هذا الأمر في الفصل التالي.

3.1. الأشكال البيانية:

وتعرض من خلال النقر على الزر Charts فيظهر صندوق حوار له الشكل التالي:

الشكل رقم (8-3) صندوق حوار الرسوم البيانية (Charts)

نلاحظ أن صندوق الحوار السابق يقسم إلى قسمين:

a) القسم الأول المعنون بـ (Chart Type): يستخدم لتحديد نوع الشكل والأشكال البيانية التي يتم عرضها:

- مخطط القرص Pie Chart: والذي يستخدم لتمثيل المتغيرات النوعية أو الترتيبية على شكل قرص دائري.
- مخطط الأعمدة Bar Chart: والذي يستخدم لتمثيل المتغيرات النوعية أو الترتيبية على شكل أعمدة.
- المدرج التكراري Histogram: والذي يستخدم فقط في حالة المتغيرات الكمية لتمثيل البيانات على شكل أعمدة متلاصقة ويمكننا من عرض المنحني الطبيعي Normal Curve متداخلاً مع المدرج التكراري وذلك بمجرد التأشير والنقر على المربع.

b) القسم الثاني (Chart Values): ويستخدم لتحديد طريقة التمثيل البياني للفئات المتغير ويتم ذلك من خلال:

- Frequencies: لتمثيل التكرارات.
- Percentages: لتمثيل النسب المئوية.

4.1. الترتيب والتنظيم:

إضافة إلى إمكانات الأمر Frequencies فإنه يتيح إمكانية تنظيم النتائج وعرضها وترتيبها بطرق متعددة بحيث يسهل تحليلها وتفسيرها ومقارنتها وذلك من خلال النقر على الزر Format فيظهر صندوق الحوار التالي:

الشكل رقم (8-4) صندوق حوار الترتيب والتنظيم (Format)

ويتضمن ما يلي:

a) Ordered By: لترتيب الفئات في الجدول التكراري تصاعدياً أو تنازلياً إما حسب قيمها Values أو حسب تكراراتها Counts.

b) Multiple Variables : يستخدم هذا الأمر في حالة وجود أكثر من متغير في قائمة المتغيرات، وتشمل ما يلي:

- Compare variables : لعرض المؤشرات الإحصائية للمتغيرات كافة في جدول واحد.
- Organize output by variables : لعرض مؤشرات كل متغير في جدول مستقل.

c) Suppress Tables With More Than n Categories: لإخفاء الجدول التكراري للمتغيرات التي يزيد عدد فئاتها عن العدد المحدد من قبل المستفيد.

ملاحظة: في حالة وجود متغير واحد ليس ضرورياً أي من الخيارين.

مثال 1:

تم تسجيل نوع المهنة لـ (30) سائح عند قدومهم إلى مطار دمشق الدولي كما يلي:

طالب - تاجر - أعمال حرة - دبلوماسي - تاجر - طالب - تاجر - أعمال حرة - دبلوماسي - طالب - تاجر - أعمال حرة - تاجر - دبلوماسي - طالب - تاجر - أعمال حرة - طالب - تاجر - طالب - تاجر - أعمال حرة - طالب - تاجر - أعمال حرة - طالب - تاجر - طالب - تاجر.

والمطلوب:

1- أدخل البيانات السابقة في البرنامج الإحصائي SPSS.

2- أنشئ جدول تكراري ثم اقرأه.

3- ارسم مخطط الدائرة للبيانات السابقة.

4- ما هو المؤشر الإحصائي الذي يمكن استخدامه هنا.

الحل:

1) إدخال البيانات:

نعرف المتغير في صفحة (Variable View) كما في الجدول التالي:

Name	Type	Label	Values		Measurement
			Value	Label	
نوع_المهنة	Numeric	نوع المهنة	1	طالب	Nominal
			2	تاجر	
			3	أعمال حرة	
			4	دبلوماسي	

(الأعمدة التي لم ترد في الجدول السابق تترك بشكل افتراضي)

ندخل البيانات في صفحة (Data View).

2) إيجاد الجدول التكراري:

من قائمة Analyze نختار الأمر Descriptive Statistics ثم نختار الأمر الفرعي Frequencies فيظهر صندوق حوار بعنوان Frequencies.

نحدد المتغير المراد إيجاد الجدول التكراري له، وهو في المثال أعلاه (نوع_المهنة). ثم ننقر على السهم لنقل المتغير إلى مربع Variables، نلاحظ أنّ الخيار Display Frequency tables محدد بشكل افتراضي، وبعدها ننقر على الزر OK فتظهر في شاشة المخرجات الجداول التالية وهي:

Statistics		
نوع المهنة		
N	Valid	30
	Missing	0

نوع المهنة					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	طالب	9	30.0	30.0	30.0
	تاجر	12	40.0	40.0	70.0
	أعمال حرة	6	20.0	20.0	90.0
	دبلوماسي	3	10.0	10.0	100.0
	Total	30	100.0	100.0	

فالجدول الأول المسمى (Statistics) يعرض لنا عدد الحالات المدخلة (Valid) وهي (30) حالة وعدد الحالات المفقودة (Missing) وهي (0).

والجدول الثاني والذي يحمل اسم المتغير المدروس (نوع المهنة) يمثل الجدول التكراري حيث يحوي على الأعمدة التالية:

a) العمود الأول المسمى Frequency: والذي يعني عدد أفراد العينة في كل فئة من فئات المتغير.

b) العمود الثاني المسمى Percent: والذي يتضمن النسب المئوية لأفراد العينة في كل فئة من فئات المتغير.

c) العمود الثالث المسمى Valid Percent: والذي يتضمن النسب المئوية لأفراد العينة في كل فئة من فئات المتغير بعد استبعاد البيانات المفقودة Missing.

d) العمود الرابع المسمى Cumulative Percent: يمثل النسب المئوية التراكمية لأفراد العينة في كل فئة من فئات المتغير.

ويقرأ كما يلي: السياح الطلاب 9 أشخاص بنسبة (30%) بينما 12 سائح تاجر بنسبة (40%) و 6 أشخاص أعمال حرة بنسبة (20%) و دبلوماسي عددهم 3 بنسبة (10%).

3) رسم مخطط الدائرة البيانية:

لرسم المخطط الدائري من صندوق الحوار المعنون بـ Frequencies السابق، وفق الخطوات التالية:

- من خلال صندوق حوار Frequencies نضغط على الزر Chart .
- نختار الدائرة البيانية Pie Chart .
- نختار طريقة تمثيل القطاعات بطريقة النسب المئوية Percentages.
- ثم نضغط على Continue ثم OK، فيظهر في شاشة المخرجات الشكل البياني التالي:

4) المؤشر الإحصائي:

المؤشر الإحصائي الذي يمكن استخدامه في حالة المتغيرات النوعية هو المنوال والذي يبين القيمة الأكثر تكراراً، ولإدراج المؤشر الإحصائي نستخدم الأمر الإحصائي Frequencies السابق، وفق الخطوات التالية:

- من خلال صندوق حوار Frequencies نضغط على الزر Statistics، فيظهر صندوق حوار فرعي بعنوان Frequencies Statistics.
- من القسم المعنون بـ Central Tendency نختار الأمر Mode.
- ثم نضغط على Continue ثم OK، فيظهر في شاشة المخرجات الجدول التالي:

Statistics		
نوع المهنة		
N	Valid	30
	Missing	0
Mode		2.00

نلاحظ أن قيمة المنوال كما يبينه الجدول السابق في صف Mode هي القيمة 2 والتي تدل على أن الفئة التي تحمل رقم 2 أي التاجر هو الفئة الأكثر تكراراً.

مما سبق نجد:

- يستفاد من الجداول التكرارية في دراسة الأهمية النسبية للفئات المختلفة بالإضافة إلى استخدامه كأساس لإجراء المقارنات بين عدد التوزيعات.
- ويضاف التكرار النسبي في الجداول لأن النسبة المئوية تعطي دلالة أكبر من العدد المطلق.
- أما النسبة المئوية التصاعدية يستفاد منها في بيان نسبة الأفراد الذين يقعون تحت فئة أو أقل من فئة معينة، كما له فائدة في حساب بعض المؤشرات الإحصائية البسيطة كالوسيط والترتيب المئوي.

2. الجداول التقاطعية Cross Table:

وهي تلك الجداول التي تتوزع فيها البيانات حسب متغيرين أو أكثر في نفس الوقت، فهو يتكون من صفوف وأعمدة، وتمثل الصفوف فئات أو مجاميع إحدى الظاهرتين والأعمدة تمثل فئات المتغير الآخر والمربعات الناتجة من الصفوف والأعمدة فتحتوي على التكرارات المشتركة بين الظاهرتين، كما يستخدم لعرض عدد الحالات(التكرارات) التي لها مجاميع مختلفة من قيم متغيرين مصنفين أو أكثر (Categorical Variables)، ويعتبر وضع البيانات في جداول تقاطعية الخطوة الأولى في وصف العلاقة ودراستها بين الظواهر المختلفة، ويمكن أن يرافق الجدول التقاطعي حساب مؤشرات واختبارات إحصائية.

ويسمى الجدول التقاطعي حسب المتغيرات التي يتم عرضها في الجدول فالجدول التقاطعي لمتغيرين باسم (Two-Way Crosstabulation). ويسمى الجدول التقاطعي لأكثر من متغيرين باسم (Multi-Way Crosstabulation).

وعند اختيار الأمر Crosstabs يظهر صندوق حوار له الشكل التالي:

الشكل رقم (8-5) صندوق حوار الجداول التقاطعية (Crosstabs)

ويتطلب صندوق الحوار السابق تحديد متغير الصف Row وتحديد متغير العمود Column.

ويحوي صندوق حوار Crosstabs الأزرار الرئيسية التالية:

1.2. Statistics:

عند اختياره يظهر صندوق حوار يتيح إمكانية حساب معاملات الارتباط بين المتغيرين المحددين سابقاً في العمود والصف:

الشكل رقم (8-6) صندوق الحوار الفرعي الخاص بحساب الاختبارات الإحصائية (Statistics)

2.2. Cells:

عند اختياره يظهر صندوق حوار يتيح إمكانية التحكم بعرض محتويات الجدول التقاطعي، عند الضغط على هذا الخيار يظهر صندوق الحوار التالي:

الشكل رقم (8-7) صندوق الحوار الفرعي الخاص بحساب النسب المئوية والبواقي في الجدول التقاطعي

يضم صندوق الحوار الخيارات التالية:

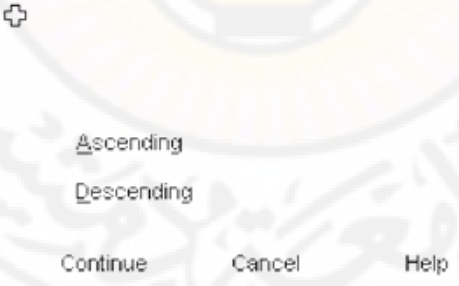
- حقل Counts: وهو لتحديد تكرار الإجابة على خيار معين، ويحوي الخيارات التالية:
  - Observed: وهو محدد بشكل افتراضي حيث تملأ الخلايا بالتكرارات المشاهدة والتي يرمز لها بـ O_i.
  - Expected: فهو لحساب التكرارات المتوقعة والتي يرمز لها بالرمز E_i.
- حقل Percentages: وهو لحساب النسب المئوية ويحوي على الخيارات التالية:
  - Rows: تملأ خلايا الجدول بالنسب المئوية من مجموع الصف (أي يتم قسمة تكرارات أي صف على مجموع الصف وسنحصل على نسبة الخلية للصف).
  - Columns: تملأ خلايا الجدول بالنسب المئوية من مجموع العمود (أي يتم قسمة تكرارات أي عمود على مجموع العمود وسنحصل على نسبة الخلية للعمود).
  - Total: تملأ خلايا الجدول بالنسب المئوية من المجموع الكلي (أي يتم قسمة جميع أرقام الجدول على مجموع التكرارات وسنحصل على نسبة كل خلية للمجموع الكلي).
- حقل Residuals: وهذا الحقل مخصص لحساب الفروق بين القيم المشاهدة والقيم المتوقعة وذلك حسب كل خيار، ويحوي على الخيارات التالية:
  - Unstandardized: لملئ الخلايا بالفرق بين التكرار المشاهد والتكرار المتوقع.
  - Standardized: نفس الخيار السابق مقسوماً على الخطأ المعياري.
  - Adj. standardized: نفس الخيار السابق معبراً عنه بوحدات الانحراف المعياري عن المتوسط.

- حقل Noninteger Weights: وهذا الحقل مخصص لتدوير التكرارات إذا كان هناك أوزان للمتغيرات معطاة بشكل رقمي مستمر (مر معنا مسبقاً الأمر Weight Cases) ويحوي على الخيارات التالية:
- Round cell counts ، Round case weights: يتم التقريب (التدوير) لأقرب عدد صحيح بعد العدد المعطى.
- Truncate cell counts ، Truncate case weights: يتم اقتطاع الجزء العشري من الأوزان ويبقى العدد الصحيح.
- No adjustments: تترك الأوزان على حالها دون أي تغيير.

3.2. Format:

يتيح لنا إظهار ترتيب صفوف الجدول تصاعدياً (Ascending) أو تنازلياً (Descending) حسب فئات متغير الصف، عند الضغط على هذا الخيار يظهر صندوق الحوار التالي:

الشكل رقم (8-8) صندوق الحوار الفرعي الخاص بالترتيب في الجدول التقاطعي



4.2. Display clustered bar charts:

يستخدم لإظهار مخطط الأعمدة.

5.2. Suppress Tables:

يستخدم لإخفاء الجداول التقاطعية.

مثال 2:

تم تسجيل (30) سائح عند قدومهم إلى مطار دمشق الدولي، حول مهنة كل منهم وفقاً لجنس المجيب فكانت إجابتهم كما يلي:

نوع المهنة	طالب	طالب	طالب	طالب	طالب	طالب	طالب	طالب	طالب	تاجر
الجنس	ذكر	ذكر	ذكر	أنثى	أنثى	أنثى	أنثى	ذكر	أنثى	أنثى
نوع المهنة	تاجر	تاجر	تاجر	تاجر	تاجر	تاجر	تاجر	تاجر	تاجر	تاجر
الجنس	ذكر	أنثى	أنثى	أنثى	أنثى	ذكر	ذكر	ذكر	ذكر	ذكر
نوع المهنة	تاجر	أعمال حرة	أعمال حرة	أعمال حرة	أعمال حرة	أعمال حرة	أعمال حرة	دبلوماسي	دبلوماسي	دبلوماسي
الجنس	ذكر	ذكر	أنثى	ذكر	أنثى	ذكر	أنثى	ذكر	أنثى	أنثى

والمطلوب:

1- أدخل البيانات السابقة في البرنامج الإحصائي SPSS.

2- تكوين جدول تقاطعي بين متغير الجنس ومتغير نوع المهنة، وارسم مخطط يمثل ذلك.

الحل:

1) إدخال البيانات: نعرف المتغيرات في صفحة (Variable View) كما في الجدول التالي:

Name	Type	Label	Values		Measurement
			Value	Label	
نوع_المهنة	Numeric	نوع المهنة	1	طالب	Nominal
			2	تاجر	
			3	أعمال حرة	
			4	دبلوماسي	
الجنس	Numeric	الجنس	1	ذكر	Nominal
			2	أنثى	

(الأعمدة التي لم ترد في الجدول السابق تترك بشكل افتراضي)

ندخل البيانات في صفحة (Data View).

2) تكوين جدول تقاطعي بين متغير الجنس ومتغير نوع المهنة:

من قائمة Analyze نختار الأمر Descriptive Statistics ثم نختار الأمر الفرعي Crosstabs فيظهر صندوق حوار بعنوان Crosstabs، نحدد المتغير المستقل ونضعه في مربع (Row) وهو في مثالنا الجنس، ونحدد المتغير التابع ونضعه في مربع (Column) وهو في مثالنا نوع المهنة.

نضغط على زر Cell فيظهر صندوق حوار بعنوان Crosstabs: Cell Display ونحدد من مستطيل Percentages الخيارات Row, Column, Total، ثم نضغط على Continue، ولإظهار مخطط أعمدة نفعل الخيار Display clustered bar charts في صندوق حوار Crosstabs، ثم نضغط على الزر OK.

فيظهر في شاشة المخرجات الجدول المعنون بـ (**الجنس * نوع المهنة Crosstabulation**) كما يلي:

الجنس * نوع المهنة Crosstabulation							
			نوع المهنة				Total
			طالب	تاجر	أعمال حرة	دبلوماسي	
الجنس	ذكر	Count	5	7	3	1	16
		% within الجنس	31.3%	43.8%	18.8%	6.3%	100.0%
		% within نوع_المهنة	55.6%	58.3%	50.0%	33.3%	53.3%
		% of Total	16.7%	23.3%	10.0%	3.3%	53.3%
	أنثى	Count	4	5	3	2	14
		% within الجنس	28.6%	35.7%	21.4%	14.3%	100.0%
		% within نوع_المهنة	44.4%	41.7%	50.0%	66.7%	46.7%
		% of Total	13.3%	16.7%	10.0%	6.7%	46.7%
Total		Count	9	12	6	3	30
		% within الجنس	30.0%	40.0%	20.0%	10.0%	100.0%
		% within نوع_المهنة	100.0%	100.0%	100.0%	100.0%	100.0%
		% of Total	30.0%	40.0%	20.0%	10.0%	100.0%

من الجدول السابق نجد أن عدد الذكور الطلاب 5 (من سطر Count تقاطع سطر ذكر مع طالب) بنسبة 16.70% (من سطر of Total% تقاطع سطر ذكر مع طالب)، وعدد الذكور التجار 7 بنسبة 23.30%، وعدد الذكور الأعمال الحرة 3 بنسبة 10%، وعدد الذكور الدبلوماسيين 1 بنسبة 3.30%، أما عدد الإناث الطلاب 4 (من سطر Count تقاطع سطر ذكر مع طالب) بنسبة 13.30% (من سطر of Total% تقاطع سطر ذكر مع طالب)، وعدد الإناث التجار 5 بنسبة 16.70%، وعدد الإناث الأعمال الحرة 3 بنسبة 10%، وعدد الإناث الدبلوماسيين 2 بنسبة 6.70%.

أسئلة الفصل الثامن

1. ليكن لدينا الجدول التالي الذي يصف متغير الشهادة لعينة تم استطلاع آرائهم:

الشهادة					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ثانوية	20	11.8	12.5	12.5
	معهد متوسط	60	35.3	37.5	50.0
	إجازة جامعية	80	47.1	50.0	100.0
	Total	160	94.1	100.0	
Missing	System	10	5.9		
Total		170	100.0		

والمطلوب:

أ. ما حجم العينة التي تم استطلاع رأيها.

ب. ما نسبة حملة شهادة المعهد المتوسط في العينة المستطلع آرائهم.

ج. قم بوصف متغير الشهادة.

2. ليكن لدينا الجدول التقاطعي للجنس (ذكر – أنثى) والشهادة (معهد متوسط – إجازة جامعية)، والمطلوب:

الجنس * الشهادة Crosstabulation					
			الشهادة		Total
			معهد متوسط	إجازة جامعية	
الجنس	ذكر	Count	20	25	45
		الجنس% within	44.4%	55.6%	100.0%
		الشهادة%within	55.6%	51.0%	52.9%
		% of Total	23.5%	29.4%	52.9%
	أنثى	Count	16	24	40
		الجنس% within	40.0%	60.0%	100.0%
		الشهادة%within	44.4%	49.0%	47.1%
		% of Total	18.8%	28.2%	47.1%
Total		Count	36	49	85
		الجنس% within	42.4%	57.6%	100.0%
		الشهادة%within	100.0%	100.0%	100.0%
		% of Total	42.4%	57.6%	100.0%

1- عرف كلاً من متغيري الجنس والشهادة في البرنامج SPSS.

2- ما نسبة الذكور في العينة؟

3- ما نسبة الإناث في العينة؟

4- ما نسبة حملة شهادة المعهد المتوسط في العينة؟

5- ما نسبة حملة الإجازة الجامعية في العينة؟

6- ما نسبة الذكور حملة شهادة المعهد المتوسط في العينة؟

7- ما نسبة الذكور حملة الإجازة الجامعية في العينة؟

8- ما نسبة الإناث حملة شهادة المعهد المتوسط في العينة؟

9- ما نسبة الإناث حملة الإجازة الجامعية في العينة؟

10-ما نسبة حملة شهادة المعهد المتوسط من الذكور في العينة السابقة؟

11-ما نسبة حملة الإجازة الجامعية من الذكور في العينة السابقة؟

12-ما نسبة حملة شهادة المعهد المتوسط من الإناث في العينة السابقة؟

13-ما نسبة حملة الإجازة الجامعية من الإناث في العينة السابقة؟

14-ما نسبة الذكور من حملة شهادة المعهد المتوسط في العينة السابقة؟

15-ما نسبة الإناث من حملة شهادة المعهد المتوسط في العينة السابقة؟

16-ما نسبة الذكور من حملة الإجازة الجامعية في العينة السابقة؟

17-ما نسبة الإناث من حملة الإجازة الجامعية في العينة السابقة؟

Damascus University

وقل رب زدني علما
جامعة دمشق
Damascus University

الفصل التاسع CHAPTER 9

الوصف الإحصائي للمتغيرات الكمية

أهداف الفصل التاسع:

1. التعرف على مقاييس النزعة المركزية.
2. التعرف على مقاييس التشتت.
3. التعرف على الأمر Frequencies.
4. التعرف على الأمر Descriptive.
5. التعرف على الأمر Explore.

الهدف الأساسي من وصف البيانات الكمية هو تلخيص البيانات للتعرف على مركز البيانات ومقدار تشتت (تبعثر) البيانات حول هذا المركز (درجة تجانس البيانات) ومن خلال هذين المؤشرين يتمكن الباحث من فهم أبعاد الظاهرة المدروسة، وبالتالي يتم الوصف الإحصائي للبيانات الكمية من خلال مقاييس النزعة المركزية والتشتت لأن مقاييس النزعة المركزية لا تعتبر كافية لوصف مجموعة من البيانات وصفاً كاملاً فقد تتساوى بعض العينات في أحد مقاييس النزعة المركزية على الرغم من اختلاف توزيع بياناتها حول مركزها (درجة تجانس البيانات)، كما أن مقاييس النزعة المركزية تمثل مركز البيانات لكنها لا تبين مدى التفاف أو بعثرة البيانات حول هذا المقياس، ولهذا لا بد من وجود مقياس آخر مع مقاييس النزعة المركزية لقياس درجة التشتت أو التبعثر في داخل هذه البيانات، لذا وجب وضع مقاييس النزعة المركزية والتشتت معاً.

1. مقاييس النزعة المركزية:

نعرف مقياس النزعة المركزية بأنه القيمة (الرقم) الذي تتمركز حوله البيانات، ومن أهم مقاييس النزعة المركزية التي سنتعرض إليها بالدراسة الوسط الحسابي والوسيط والمنوال، كما سنتعرض بالدراسة لحساب كل من البيانات الآتية:

1.1. الوسط الحسابي (المتوسط الحسابي) Mean:

يعرف الوسط الحسابي (المتوسط الحسابي): بأنه مجموع قيم المشاهدات مقسوماً على عددها ويرمز له بالرمز ($\overline{X}$)، ويحسب من العلاقة التالية:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

حيث: X_i المشاهدة i ، n هو حجم العينة.

ويفسر الوسط الحسابي لمجموعة من المشاهدات أو (القيم) بأنه القيمة التي لو أعطيت لكل مشاهدة أو (قيمة) في العينة المدروسة لكانت جميع المشاهدات (القيم) الجديدة مساوية للوسط الحسابي (للمتوسط الحسابي).

حيث يعتبر الوسط الحسابي من أشهر مقاييس النزعة المركزية ويصلح للبيانات الكمية وهو وحيد، لكنه يتأثر بالقيم الشاذة والمتطرفة وعادةً ما يتم اللجوء إلى حساب المتوسط المشذب (Trimmed Mean) للتخلص من تأثير المشاهدات الشاذة في المتوسط ويحسب المتوسط المشذب بعد استثناء P% من أدنى المشاهدات و P% وأعلى المشاهدات فإذا كان P=5 في هذه الحالة يتم استثناء 5% من أدنى المشاهدات و 5% وأعلى المشاهدات ويحسب المتوسط لـ 90% من المشاهدات المتبقية.

2.1. الوسيط Median :

يعرف الوسيط بأنه القيمة الواقعة في وسط المشاهدات بعد ترتيب المشاهدات ترتيباً تصاعدياً أو تنازلياً، ولحساب الوسيط نقوم أولاً بترتيب المشاهدات ترتيباً تصاعدياً أو تنازلياً ومن ثم نأخذ القيمة الواقعة في المنتصف ويعتمد هذا على عدد تلك المشاهدات فهناك حالتان إذا كان عدد المشاهدات فردياً أما إذا كان عدد البيانات زوجياً فإن الوسط هو متوسط القيمتين الواقعتين في وسط البيانات، فإذا كان لدينا مجموعة من المشاهدات $x_1، x_2، x_3، \ldots، x_n$ المرتبة تصاعدياً عندئذٍ يكون:

- إذا كان عدد المشاهدات فردياً فإن الوسيط يكون:

$$Median = X_{(\frac{n+1}{2})} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

- إذا كان عدد المشاهدات فردياً فإن الوسيط يكون:

$$Median = \frac{X_{(\frac{n}{2})} + X_{(\frac{n+2}{2})}}{2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

ويفسر الوسيط لمجموعة من المشاهدات أو (القيم) بأنه القيمة التي تكون أكبر منها 50% من المشاهدات وأصغر منها 50% من المشاهدات.

كما أن الوسيط لا يتأثر بالقيم الشاذة أو المتطرفة لأن المتوسط لا يتعلق بقيمة المشاهدة وإنما بموقع المشاهدة لذلك يفضل استخدام الوسيط مع البيانات الترتيبية، وبما أن الوسيط لا يتعلق بعدد المشاهدات وقيمتها لذا يعد أقل أهمية من المتوسط الحسابي.

3.1. المنوال Mode:

يعرف المنوال بأنه قيمة المشاهدة التي تتكرر أكثر من غيرها، أو هو المفردة ذات القيمة الأكثر شيوعاً أو تكراراً، كما يستخدم مع المتغيرات النوعية والكمية.

وحسب تعريف المنوال يمكن أن يكون للعينة منوال واحد، أو منوالان إذا تكررت مشاهدتان بنفس عدد مرات التكرار، وفي حالة عدم تكرار أي قيمة لا يكون لدينا أي منوال.

ويفسر بأنه القيمة الأكثر رغبة (تكراراً) من أفراد العينة.

2. مقاييس التشتت:

يقصد بالتشتت دراسة مدى تقارب أو تباعد المشاهدات عن بعضها بعضا أي عن أي مقياس من مقاييس النزعة المركزية، فكلما كانت المشاهدات قريبة من بعضها بعضا أي تكون قريبة من مقاييس النزعة المركزية تكون البيانات متجانسة والعكس كلما كانت المشاهدات بعيدة عن بعضها بعضا أي بعيدة عن مقاييس النزعة المركزية تكون البيانات متباعدة أو مشتتة، وتصنف مقاييس التشتت إلى صنفين:

الصنف الأول: هي المقاييس التي توضح مدى وكيفية اختلاف القيم فيما بينها.

الصنف الثاني: هي المقاييس التي توضح مدى تشتت أو تجمع القيم حول أحد المتوسطات.

ومن أهم مقاييس التشتت التي سنتعرض إليها بالدراسة المدى والتباين والانحراف المعياري والانحراف عن المتوسط.

ولكن يجب التمييز بين نوعين من مقاييس التشتت:

مقاييس التشتت المطلقة: والتي تكون وحداتها وحدات القيم الأصلية نفسها.

مقاييس التشتت النسبية: والتي تكون خالية من وحدات القياس.

1.2. المدى Range:

وهو أبسط مقياس من مقاييس التشتت ويعبر عن الفرق بين أكبر مشاهدة وأصغر مشاهدة، وفي معظم الأحيان يكون المدى مقياساً مضللاً لاعتماده على أكبر قيمة أو أصغر قيمة اللتين غالباً ما تكونان شاذتين وخاصةً إذا كان حجم العينة أو المجتمع كبيراً نسبياً.

2.2. الانحراف الربيعي Semi – interquaritile Range:

يعتبر الاهتمام بأكبر قيمة وأصغر قيمة من البيانات في حساب المدى من أهم عيوبه لإهماله بقية القيم الواقعة بين هاتين القيمتين، وللتغلب على هذه العيوب ولتجنب القيم المتطرفة يتم حساب المدى الربيعي الذي يأخذ بالحسبان قيمتي الربيع الأول والربيع الثالث.

ويعرف الانحراف الربيعي بأنه الفرق بين الربيع الثالث والربيع الأول ويستخدم الانحراف الربيعي كمقياس للتشتت عندما يكون الوسيط هو مقياس النزعة المركزية للبيانات نفسها.

ويحسب من خلال العلاقة التالية:

$$Q = \frac{Q_3 - Q_1}{2} \quad \text{.....(4)}$$

حيث Q_3 الربيع الثالث، و Q_1 الربيع الأول.

3.2. الانحراف المتوسط Mean Deviation

الهدف من مقاييس التشتت معرفة تشتت القيم عن القيمة المركزية لذلك من الممكن حساب التشتت بحساب انحراف كل قيمة عن الوسط الحسابي فإذا كانت هذه الانحرافات كبيرة يعني أن تشتت البيانات أكبر من الحالة التي يكون فيها الانحرافات صغيرة.

لذا يعتبر الانحراف المتوسط من أحد مقاييس التشتت ويعبر عنه بمتوسط الانحرافات المطلقة للقيم عن وسطها الحسابي ويحسب من العلاقة:

$$MD = \frac{\sum_{i=1}^{n} |X_i - \overline{X}|}{n} \quad \text{.....(5)}$$

حيث X_i المشاهدة i ، $\overline{X}$ المتوسط الحسابي، n حجم العينة.

4.2. الانحراف المعياري Standard Deviation

يعد الانحراف المعياري من أكثر مقاييس التشتت شيوعاً وأهمية واستخداماً في التطبيقات العملية لاعتماده في العديد من العمليات الإحصائية المتعلقة بإجراء المقارنات واختبار الفرضيات.

ويعرف الانحراف المعياري بأنه الجذر التربيعي لمربعات انحرافات القيم عن وسطها الحسابي على عددها، ويرمز له في حالة المجتمع بـ σويرمز للعينة S ويحسب من العلاقة:

للمجتمع:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{N}} \quad \text{.....(6)}$$

للعينة:

$$S = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}} \quad \text{.....(7)}$$

ويفيد الانحراف المعياري في إيجاد درجة الاختلاف بين القيم ووسطها الحسابي وعندما يكون الانحراف المعياري قيمته صغيرة فهذا يدل على أن التوزيع متقارب وتتجمع بياناته قرب متوسطها، ويعرف التباين بأنه مربع الانحراف المعياري.

3. معامل الالتواء:

يعرف معامل الالتواء (**Skewness**) بأنه عبارة عن بعد المنحني عن التماثل، ويقصد بالتماثل أنه إذا أسقطنا عموداً من قمة المنحني التكراري وقسمه إلى قسمين منطبقين يكون التوزيع متماثلاً والعكس يكون التوزيع غير متماثل أي ملتوٍ وهنا نميز حالتين:

الالتواء الموجب (نحو اليمين): إذا كان التوزيع له ذيل طويل ناحية اليمين (القيم كبيرة) أو اعتماداً على مقاييس النزعة المركزية كما في العلاقة التالية:

$$\text{Mode} < \text{Median} < \text{Mean}$$

الالتواء السالب (نحو اليسار): إذا كان التوزيع له ذيل طويل ناحية اليسار (القيم صغيرة) أو اعتماداً على مقاييس النزعة المركزية كما في العلاقة التالية:

$$\text{Mode} > \text{Median} > \text{Mean}$$

ويحسب معامل الالتواء من خلال إحدى العلاقتين:

$$S.K = \frac{\overline{X} - Mode}{S} \quad \text{.....(8)}$$

$$S.K = \frac{3(\overline{X} - Median)}{S} \quad \text{.....(9)}$$

تنحصر قيمة معامل الالتواء بين 3- و 3+ ، وحسب قيمة معامل الالتواء نميز الحالات التالية:

S.K=0 يكون التوزيع متماثلاً.

S.K>0 يكون التوزيع موجباً (ملتوٍ نحو اليمين).

S.K<0 يكون التوزيع سالباً (ملتوٍ نحو اليسار).

وهذا ما يبينه الشكل التالي:

الشكل رقم (9-1) أشكال الالتواء للتوزع

ويفسر معامل الالتواء بأنه عبارة عن بعد المنحني عن التماثل وهو بذلك يقيس اتجاه تركز القيم كما يحدد مناطق وجود القيم المتطرفة.

ملاحظة: عند المقارنة بين درجة الالتواء لتوزيعين أو أكثر يجب أن نستخدم نفس الصيغة.

4. معامل التفلطح:

هو مقياس نسبي يصف ارتفاع قمة منحني التوزيع، يشير التفلطح إلى درجة تدبب (Peakness) أو درجة تطاول (Flatness) التوزيع في المنطقة المحيطة بالمنوال، وتقاس درجة تفلطح أي توزيع بالنسبة إلى تفلطح التوزيع الطبيعي، فإذا كانت قاعدة التوزيع ضيقة وطرفاه مرتفعين فإنه يكون متطاولاً (مدبباً)، أما إذا كانت قاعدة التوزيع واسعة وطرفاه منخفضين فإنه يكون متفلطحاً، وهذا ما يبينه الشكل التالي:

الشكل رقم (9-2) أشكال التفلطح للتوزع

ويقاس معامل التفلطح باستخدام مقاييس الموقع على الربيعين الأول والثالث والمئينين العاشر والتسعين، ويحسب باستخدام الصيغة التالية:

$$K = \frac{\frac{1}{2}(Q_3 - Q_1)}{D_9 - D_1} \quad \text{.....(10)}$$

ويأخذ معامل التفلطح قيمه من [3+ , 3-] وحسب قيمة معامل التفلطح نميز الحالات التالية:

K=3 يكون المنحني معتدلاً.

K >3 يكون المنحني متطاولاً (مرتفعاً).

K <3 يكون المنحني متفلطحاً (منخفضاً).

5. التطبيق باستخدام البرنامج الإحصائي SPSS

سيتم التركيز على وصف البيانات من خلال الأمر (Descriptive statistic) الذي يمكن استخدامه لوصف كافة أنواع المتغيرات سواء الاسمية أو الكمية.

نلاحظ وجود مجموعة من الخيارات المتاحة، وأيها نختار فيعتمد ذلك على نوعية المتغير الذي نتعامل معه، ونميز هنا الحالات التالية:

- **Frequencies**: ومن أجل المتغيرات المنفصلة أو بشكل عام من أجل المتغيرات الوصفية مثل (المستوى التعليمي) كما يمكننا استخدامها من أجل المتغيرات الكمية أو المستمرة.
- **Descriptive**: من أجل المتغيرات الكمية المستمرة كالعمر .
- **Explore**: يتعامل مع جميع المتغيرات وخصوصاً عندما نريد تلخيص متحول مستمر بالنسبة للمتحولات.

وسندرس هذه التعليمات بنوع من التفصيل:

1.5. التكرارات Frequencies:

يستخدم الأمر Frequencies لوصف توزيع مفردات العينة حسب أحد أنواع المتغيرات كالاسمي أو النوعي، وعند اختيار الأمر Frequencies يظهر صندوق حوار كما في الشكل رقم (8-1)، وتُعرض نتائج استخدام الأمر Frequencies للمتغيرات الكمية من خلال النقر على الزر وتعرض من خلال النقر على الزر Statistics فيظهر صندوق حوار فرعي بعنوان Frequencies Statistics له الشكل التالي:

الشكل رقم (9-3) صندوق حوار الأمر الاحصاءات الوصفية للمتغيرات الكمية

ويتضمن صندوق الحوار السابق المؤشرات الإحصائية التالية:

أولاً: مقاييس النزعة المركزية، Central Tendency، والتي تتضمن:

- الوسط الحسابي Mean.
- المنوال Mode.
- الوسيط Median.
- المجموع Sum.

ثانياً: مقاييس التشتت، Dispersion، ومنها:

- الانحراف المعياريStd deviation.
- التباين، Variance.
- المدى، Range.
- الخطأ المعياري، S.E. Mean.
- أدنى قيمة، Minimum
- أعلى قيمة، Maximum.

ثالثاً: الإحصاءات الخاصة بالموقع:

- المئينيات، Percentiles.
- الربيعيات، Quartiles.

رابعاً: إحصاءات شكل التوزيعات التكرارية وهي:

- توزيع الالتواء، Distribution Skewness.
- توزيع التفلطح، أو التفرطح، Kurtosis Distribution.

إنّ معظم هذه القياسات الإحصائية غالباً ما تستخدم مع متغيرات من النوع الترتيبي Ordinal، أو الكمي، ولا تستخدم مع متغيرات نوعية، ونظراً لأنّ الأمر Frequencies يمكن استخدامه مع متغيرات من النوع الترتيبي أو الكمي في بعض الحالات فقد وضعت هذه الخيارات ضمن الإجراء المذكور.

مثال 1:

يرغب باحث بدراسة الإنفاق على السياحة لعينة من 10 أفراد، كما في الجدول التالي:

535	140	387	340	426	340	240	262	235	425

والمطلوب: عرض نتائج الإحصاءات الوصفية للمتغير في الإنفاق.

لتنفيذ المطلوب نتبع الخطوات التالية:

- ندخل المتغير Salary والبيانات في شاشة محرر البيانات Data Editor.

Measurement	Values		Label	Type	Name
Scale	Label	Value	Salary	Numeric	Spending

- من خلال شريط القوائم، Menu Bar، اختر

Analyze→ Descriptive Statistic→ Frequencies

- بالنقر على زر Statistics يظهر صندوق حوار فرعي بعنوان Frequencies Statistics ونختار من صندوق الحوار:

1) مقاييس الموقع Percentiles Value:
   - الربيعيات: Quartiles.
   - Cut points for equal groups: العشيرات.
   - Percentiles: المئينيات
2) مقاييس التشتت Dispersion .
3) مقاييس النزعة المركزية Central Tendency.
4) مقاييس شكل التوزيع Distribution (الالتواء والتفلطح).

- عند نقر الزر Continue نعود إلى صندوق حوار Frequencies.
- عند النقر على المفتاح ok في صندوق حوار Frequencies، نحصل على الوصف الإحصائي للمتغيرات في شاشة المخرجات كما في الجدول التالي:

Statistics

Salary

N	Valid	10	الفعلية	العدد
	Missing	0	المفقودة	
Mean		333.0000	المتوسط الحسابي	
Std. Error of Mean		36.73841	الخطأ المعياري	
Median		340.0000	الوسط	
Mode		340.00	المدى	
Std. Deviation		116.17707	الانحراف المعياري	
Variance		13497.111	التباين	
Skewness		.061	الالتواء	
Std. Error of Skewness		.687	الخطأ في الالتواء	
Kurtosis		-.266	التفلطح	
Std. Error of Kurtosis		1.334	الخطأ في التفلطح	
Range		395.00	المدى	
Minimum		140.00	أصغر قيمة	
Maximum		535.00	أكبر قيمة	
Sum		3330.00	المجموع	
Percentiles	10	149.5000	10	النسب المئوية
	20	236.0000	20	
	25	238.7500	25	
	30	246.6000	30	
	40	293.2000	40	
	50	340.0000	50	
	60	368.2000	60	
	70	413.6000	70	
	75	425.2500	75	
	80	425.8000	80	
	90	524.1000	90	

2.5. الأمر Descriptive:

يهتم هذا الأمر بالمتغيرات الكمية Quantitative Variables والتي عادة تكون ذات قيم عديدة ولذلك يعتبر استخدام التكرارات لوصف مثل هذه المتغيرات غير مناسب وبدلاً من ذلك غالباً ما تستخدم طرائق إحصائية أخرى لوصف العينة.

يفيد هذا الأمر في عرض مقاييس الإحصاء الوصفي لمجموعة من المتغيرات الكمية في جدول واحد مع عرض القيم المعيارية لبيانات المتغيرات Z scores ويتم حساب القيم المعيارية وفق العلاقة التالية:

$$Z = \frac{X - \mu}{S} \quad \text{.....(11)}$$

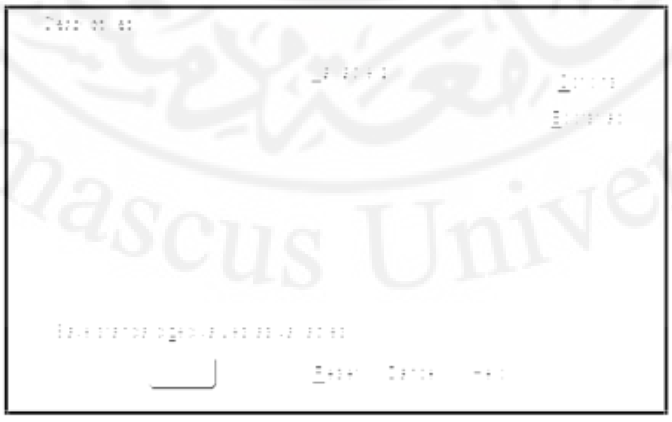
حيث:

X هو القيم المدخلة، μ هو المتوسط الحسابي للمتغير، S وهو الانحراف المعياري للمتغير، حيث تظهر القيم المعيارية في نهاية ملف البيانات.

عند اختيار الأمر الفرعي Descriptive من الأمر Descriptive Statistic من قائمة Analyze

يظهر صندوق حوار بعنوان Descriptive له الشكل التالي:

الشكل رقم (9-4) صندوق حوار الأمر Descriptives



حيث إنّ Save standardized values as variables لإضافة القيم المعيارية Z scores، كمتغيرات إلى شاشة البيانات Data Editor، عند تأشير (نقر) المربع المجاور علماً انّ:

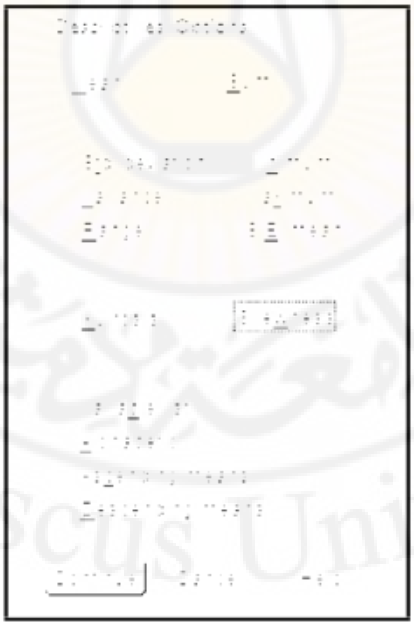
$$ZScores = \frac{X - \bar{X}}{S.D} \quad \text{.....(12)}$$

$\bar{X}$ تمثل المتوسط الحسابي للمتغير X.

X تمثل قيم المتغير.

S.D تمثل الانحراف المعياري لقيم المتغير.

وعند الضغط على زر Options يظهر صندوق حوار فرعي بعنوان Descriptives: Options له الشكل التالي:

الشكل رقم (9-5) صندوق حوار الأمر Options

حيث يتضمن المؤشرات الإحصائية التالية:

أولاً: مقاييس النزعة المركزية Central Tendency، وتشمل ما يلي:

1- الوسط الحسابي mean.

2- المجموع sum.

ثانياً: مقاييس التشتت Dispersion وتشمل:

1- الانحراف المعياري STD. Deviation مقدار تشتت القيم عن وسطها الحسابي مقاساً بوحدات المتغير نفسها S.D = $\sqrt{VAR}$ ويعتبر أهم القياسات الإحصائية لاعتماد معظم القياسات الإحصائية عليه.

2- التباين Variance: مقدار تشتت القيم عن وسطها الحسابي، وهو مربع الانحراف المعياري VAR = $(s.d)^2$.

3- المدى Range: الفرق بين أكبر قيمة وأقل قيمة R= Max – Min.

4- Minimum: أقل قيمة في المتغير.

5- Maximum: أكبر قيمة للمتغير.

6- S.E.mean: الخطأ المعياري مقدار الخطأ الموجود في الوسط الحسابي ويمكن التعبير عنه رياضياً بالصيغة التالية: $SE = \frac{S.D}{\sqrt{n}}$ وهو يستخدم للدلالة على دقة الوسط الحسابي كتقدير لوسط المجتمع.

مثال 2:

اعتماداً على بيانات المثال رقم (1) وباستخدام الأمر Descriptive أوجد ما يلي:

1. مقاييس النزعة المركزية للمتغيرات.
2. مقاييس التشتت للمتغيرات.
3. تحديد شكل توزيع البيانات المتغيرات.

خطوات الحل: لغرض إيجاد المقاييس الوصفية باستخدام الأمر Descriptive نتبع الخطوات التالية:

من قائمة Analyze نختار عند اختيار الأمر Descriptive Statistic ثم الأمر الفرعي Descriptive فيظهر صندوق حوار بعنوان Descriptive ننقل المتغير المراد وصفه إلى مربع Variable(s)، ننقر على الزر Options فيظهر صندوق حوار Descriptive: Options نختار المقاييس التالية:

1. بعض مقاييس النزعة المركزية (Mean, Sum)
2. مقاييس التشتت Dispersion.
3. شكل التوزيع Distribution حيث يمكن اختيار المؤشرات الوصفية المرغوبة، المطلوبة، في المثال أعلاه.

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Spending	10	395.00	140.00	535.00	333.00	36.738	116.17707	13497.111	.061	.687	-.266-	1.33
Valid N (listwise)	10											

كما تم إضافة الدرجات المعيارية للمتغيرات إلى شاشة البيانات Data Editor وكما هو مبين في الجدول:

وشاشة البيانات Data Editor يتضمن قيم المتغير ودرجاته المعيارية مما يتيح إمكانية الاستفادة منها لأغراض متعددة.

3.5. الأمر Explore:

يستخدم الأمر Explore للتحقق من دقة البيانات ونوعيتها وطبيعتها وتجانسها وشكلها قبل إجراء التحليلات الإحصائية، إن الأمر Explore يتضمن القيام بما يلي:

1.3.5. فحص البيانات Screening، ومحاولة تصحيح الأخطاء إن وجدت ومن أهم هذه الاخطاء:

- تتضمن أرقاماً غير منطقية، كوجود فترات انقطاع في البيانات.
- إذا كانت جميع البيانات زوجية مثلاً.
- تحتوي على قيم شاذة.

2.3.5. التحقق من بعض الشروط التي يجب توافرها قبل استخدام الاختبارات الإحصائية، مثل تحليل الانحدار، وتحليل التباين وكما سيتم ذلك في الأقسام القادمة، إذا

يستخدم الأمر Explore للتحقق من الشروط التي تتطلبها هذه الاختبارات الإحصائية ومن أهمها:

- التحقق من كون التوزيع طبيعياً للمتغير.
- التحقق من شرط تجانس التباين.

3.3.5. مقارنة توزيع متغير ما لمجموعتين من الأفراد (ذكور وإناث مثلاً)، ويمكن مقارنة توزيع متغيرين للمجموعة الواحدة من الأشخاص.

عند اختيار الأمر الفرعي Explore من الأمر Descriptive Statistic من قائمة Analyze

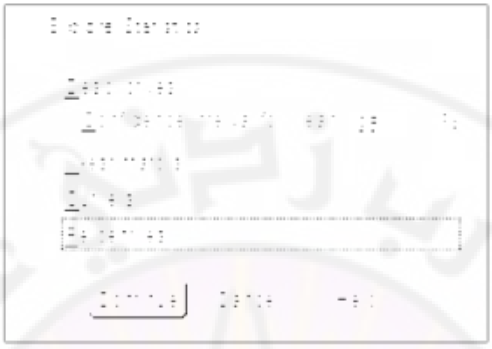
يظهر صندوق حوار بعنوان Explore له الشكل التالي:

الشكل رقم (9-6) صندوق حوار الأمر Explore

يتضمن صندوق الحوار السابق ما يلي:

- Dependent List: للمتغير الكمي.
- Factor List: نضع فيه المتغير الاسمي (كالجنس مثلا).
- Statistics: ويستخدم لإظهار أهم المؤشرات الإحصائية وبالنقر على Statistics يظهر صندوق حوار بعنوان بـ Explore: statistics له الشكل التالي:

الشكل رقم (9-7) صندوق حوار الإحصاءات ضمن الأمر Explore



ويحوي ما يلي:

- Descriptive: لإيجاد مؤشرات الإحصاءات الوصفية مثل: مقاييس النزعة المركزية Central بالإضافة إلى الوسط المقطوع (المشذب) Trimmed Mean 5% وهو الوسط الحسابي بعد حذف أعلى 5% وأقل 5% من البيانات وذلك لإلغاء أثر القيم الشاذة إن وجدت في البيانات، بالإضافة لمجال الثقة حول المتوسط وهو بشكل افتراضي 95% ويمكن أن نغيره إلى القيمة التي نرغب بها، كما يظهر لنا خيار Descriptive أيضاً مقاييس التشتت والتي تحوي على الخطأ المعياري، والانحراف المعياري، والتباين وأقل قيمة وأكبر قيمة والمدى الربيعي، مقاييس التوزيع الالتواء والتفلطح.
- M-Estimators: وهي تقديرات للمتوسط الحسابي حيث يتم إعطاء القيم البعيدة عن المركز أوزان أقل من القيم القريبة منه، وهناك أربعة أنواع من المقدرات، وهي Huber,Tukey,Hampel,Andrew وبذلك يصبح أثرها على النتائج أضعف مما لو بقيت كما هي.
- Outliers: لإظهار أكبر خمس قيم وأقل خمس قيم سواء كانت شاذة أم غير شاذة، ويشار إليها باسم Extreme Values في نتائج SPSS فإذا كانت قيماً شاذة ينصح بحذفها من البيانات حتى لا تؤثر على الاختبارات الإحصائية الأخرى.

- Percentiles: المئينيات وهي القيم التي يقل عنها نسبة معينة من البيانات فالمئين العاشر 10^{th} percentile، هو قيم المتغير المعتمد التي يسبقها 10% من الحالات ويليها 90% من الحالات بعد ترتيب قيم المتغير تصاعدياً عند الرغبة في استخراج أحد هذه المؤشرات نقوم بتأشير المربع المجاور له، حيث يقوم بعرض القيم المئينية التالية:95-90-75-50-25-10-5، عند نقر المفتاح Continue نعود إلى شاشة حوار Explore.

– plots: ويستخدم لإظهار بعض الرسومات البيانية لاستكشاف البيانات وبالنقر على Plots يظهر صندوق حوار بعنوان Explore: Plots.

الشكل رقم (9-8) صندوق حوار الرسومات البيانية ضمن الأمر Explore

- Boxplots: لإظهار مخطط الصندوق ذو القرنين من الجزء المعنون.
- Descriptive: لإظهار مخطط الساق والأوراق Stem – and– leaf، كما يظهر مخطط الأعمدة المتلاصقة عند اختيار Histogram.
- Normality Plots with tests: للتحقق من أن توزيع المتغير المطلوب متماثل (للمتغير التوزيع الطبيعي) أم لا ، وذلك عن طريق:
  - ✓ اختبار Kolmogorov–Smirnov واختبار Shapiro–Wilk .

✓ بعض الرسوم البيانية مثل:

❖ Normal probability plots: بحيث يتم رسم كل قيمة من هذا المتغير مع نظيرها من التوزيع الطبيعي، فإذا ما وقعت جميع هذه النقاط على خط مستقيم فإن هذا المتغير يكون متماثل التوزيع وإذا تشتتت النقاط عن الخط المستقيم فإن هذا المتغير يكون غير متماثل التوزيع.

❖ هناك بعض الرسوم الأخرى التي تعطي فكرة عن شكل التوزيع لمتغير ما، ومن خلالها يمكن تقدير ما إذا كان توزيع هذا المتغير قريباً من التوزيع الطبيعي أم لا، ومن أهم هذه الرسومات Histogram, steam and Leaf plot.

- Power Estimation : تقدير الأس المناسب ، لإجراء تحويل على البيانات لجعل التباين أكثر تماثلاً إذا كان غير ذلك ويكون الأس المناسب للتحويل أحد مضاعفات القيمة 2/1 الأقرب للقيمة المقدرة:

✓ إذا كانت القيمة المقدرة 1.95 فإن قيمة الأس المناسبة هي القيمة 2، وبذلك يكون التحويل المناسب هو مربع القيمة.

✓ وإذا كانت القيمة المقدرة هي 0.1 فإن قيمة الأس المناسبة للتحويل ستكون لوغاريتم القيم.

✓ اختبار تجانس التباين (Homogeneity of variance) عن طريق اختبار Levene test الضروري لإجراء اختبار تحليل الانحدار وتحليل التباين.

مثال 3:

يرغب باحث بدراسة الدخل اليومي لعينة من 10 أفراد، كما في الجدول التالي:

53	14	38	34	42	34	24	26	23	42

والمطلوب: أوجد المؤشرات الإحصائية لمتغير الدخل وفقاً لطريقة Explore.

لتنفيذ المطلوب نتبع الخطوات التالية:

1- ندخل المتغير Salary والبيانات في شاشة محرر البيانات Data Editor كما سبق.

2- من خلال شريط القوائم Menu Bar اختر

Analyze→ Descriptive Statistic→ Explore.

يظهر صندوق حوار بعنوان Explore ننقل متغير Salary إلى قسم: Dependent list.

Options يظهر صندوق حوار Descriptive.

3- ننقر على Statistics يظهر صندوق حوار بعنوان Explore: Statistics.

4- ننقر على المفتاح Plots في شاشة حوار Explore يظهر صندوق حوار Explore: plots.

5- ننقر على OK في صندوق حوار Explore لتظهر النتائج كما في الجدول التالي:

Descriptives

					Statistic	Std. Error
المتوسط الحسابي		Salary	Mean		333.0000	36.73841
مجال الثقة حول المتوسط بمقدار %95	الحد الأدنى		95% Confidence Interval for Mean	Lower Bound	249.8919	
	الحد الأعلى			Upper Bound	416.1081	
المتوسط المشذب 5%			5% Trimmed Mean		332.5000	
الوسيط			Median		340.0000	
التباين			Variance		13497.111	
الانحراف المعياري			Std. Deviation		116.17707	
أكبر قيمة			Minimum		140.00	
أصغر قيمة			Maximum		535.00	
المدى			Range		395.00	
المدى الربيعي			Interquartile Range		186.50	
معامل الالتواء			Skewness		.061	.687
معامل التفلطح			Kurtosis		-.266-	1.334

من الجدول السابق نجد أن من المؤشرات التي يعرضها الأمر Explore مجال الثقة حول المتوسط (95% Confidence Interval for Mean) والذي يقرأ أننا واثقون بمقدار 95% أن المتوسط لن يقل عن (249.8919) (قيمة الحد الأدنى) ولن يزيد عن (416.1081) (قيمة الحد الأعلى).

كما أنه يعرض المتوسط المشذب (5% Trimmed Mean).

والمدى الربيعي (Interquartile Range).

أما مخطط الساق والأوراق فهو كما يلي:

Salary Stem-and-Leaf Plot

Frequency	Stem & Leaf
1.00	1 . 4
3.00	2 . 346
3.00	3 . 448
2.00	4 . 22
1.00	5 . 3

Stem width: 10.00

Each leaf: 1 case(s)

ومخطط الصندوق ذو القرنين كما يلي:

أسئلة وتمارين الفصل التاسع

1- لتكن لدينا البيانات التالية التي تمثل عدد ساعات تدريب طلاب السنة الرابعة في كلية السياحة:

32	36	39	40	34	33	35	30	10
35	46	45	33	30	41	45	45	50

والمطلوب: باستخدام البرنامج الإحصائي (SPSS):

- أحسب مقاييس النزعة المركزية.
- أحسب مقاييس التشتت.
- أحسب مقياسي الالتواء والتفلطح.

2- بفرض لدينا الجدول التالي الذي يمثل مؤشرات إحصائية لمتغير الدخل:

Damascus University

الدخل

123.4091	Lower Bound	95% Confidence Interval for Mean
125.5909	Upper Bound	
124.5000	Mean	
124.5	5% Trimmed Mean	
124.000	Median	
122.00	Mode	
8.534	Variance	
0.000	Skewness	
0.427	Std. Error of Skewness	
0.226	Kurtosis	
0.633	Std. Error of Kurtosis	
9.00	Range	
120.00	Minimum	
129.00	Maximum	
120.1000	10	Percentiles
121.2000	20	
122.0000	25	
122.3000	30	
123.4000	40	
124.5000	50	
125.6000	60	
126.7000	70	
127.0000	75	
127.8000	80	
128.9000	90	

والمطلوب:

- أكتب مجال الثقة حول المتوسط وفسره.
- أوجد المقاييس التالية: المتوسط الحسابي، المنوال، الوسيط مع تفسيرهما.

- أوجد مقياسين من مقاييس التشتت.
- أوجد معامل الالتواء ومعامل التفلطح من الجدول السابق.
- ارسم مخطط (Boxplot) وضع عليه التسميات مع قيمها.

3- ليكن لدينا مخطط الساق والأوراق التالي:

Stem-and-Leaf Plot

Frequency	Stem & Leaf
3.00	22 . 013
6.00	22 . 556669
1.00	23 . 0
4.00	23 . 5569
4.00	24 . 0024
1.00	25 . 0

Stem width: 10.00

Each leaf: 1 case(s)

المطلوب:

1- ما حجم العينة المدروسة.

2- ما مدى البيانات.

3- ما هو المنوال للبيانات السابقة.

4- من المخطط السابق ارسم المخطط الصندوقي (Boxplot) وحدد عليه التسميات والقيم التي يمكن الحصول عليها من مخطط الساق والأوراق السابق.

وقل رب زدني علما
جامعة دمشق
Damascus University

الفصل العاشر CHAPTER 10

الإحصاء الاستدلالي

أهداف الفصل العاشر:

1. التعرف على أساليب الاستدلال الإحصائي.
2. التعرف على طرق اختبار التوزيع الطبيعي باستخدام SPSS.
3. التعرف على طرق اختبار الفرضية المتعلقة بوسط حسابي واحد باستخدام SPSS.
4. التعرف على طرق اختبار الفرضية المتعلقة بمتوسطي عينتين مستقلتين باستخدام SPSS.
5. التعرف على طرق اختبار الفرضية المتعلقة بمتوسطي عينتين مرتبطتين باستخدام SPSS.
6. التعرف على طرق اختبار الفرضية المتعلقة بمتوسطي أكثر من عينتين باستخدام SPSS.

يستخدم الباحث العديد من المقاييس الإحصائية التي تلخص بيانات بحثه وتنظمها وتمثلها بطرق مختلفة، وتم عرض العديد من المقاييس الإحصائية ومن أهمها: مقاييس النزعة المركزية، ومقاييس التشتت، والمئينيات، والعشيريات، ومعامل الالتواء، ومعامل التفلطح. وقد لاحظنا أن هذه المقاييس تُمكن الباحث من اختزال البيانات، وتكشف عن بنيتها وتستخلص معلومات مفيدة منها.

وعلى الرغم من أهمية هذه المعلومات الوصفية في إلقاء الضوء على الظاهرة التي يقوم الباحث بدراستها، إلا أنها لا تنطبق إلا على العينة التي استمدت منها هذه البيانات، ولكن الباحث لا يهدف عادة من جمع بياناته من عينة منتقاة الاقتصار على وصف خصائص الظاهرة موضع الدراسة في إطار هذه العينة فحسب، وإنما تعميم نتائج بحثه من العينة المدروسة إلى المجتمع المستهدف. وهذا الاستنتاج بالتعميم من العينة إلى المجتمع هو ما يسمى بالاستدلال الإحصائي.

والاستدلال الإحصائي هو استخدام الأساليب الإحصائية التي تساعد الباحث في التوصل إلى استنتاجات وتعميمات صادقة تمكنه من تقدير احتمال الوقوع في أخطاء الاستدلال استناداً إلى المعلومات التي توفرها له العينة المدروسة، والأساليب الكمية المستخدمة في الاستدلال الإحصائي لا تؤدي إلى نتائج حتمية مؤكدة وإنما تكون نتائجها احتمالية، فالاستنتاج من العينة إلى المجتمع يخضع لبعض الخطأ ويمكن تقدير هذا الخطأ، وإذا لم يتم تقديره فلا يكون لهذا التعميم أية فائدة.

1. أساليب الاستدلال الإحصائي (الإحصاء الاستدلالي):

يمكن تصنيف أساليب الاستدلال الإحصائي تبعاً للعديد من العوامل منها:

1.1. التصنيف حسب الهدف من الاستدلال الإحصائي:

أ) **أساليب التقدير:**

تستخدم هذه الأساليب في البحوث الاستكشافية بهدف تقدير بعض خصائص المجتمع المدروس، كتقدير نسبة العاملين في القطاع السياحي في إحدى المدن، أو تقدير

متوسط دخل الأسرة في أحد البلدان، أو تقدير متوسط عدد ساعات مشاهدة التلفزيون يومياً، أو تقدير معدل الجريمة في إحدى المناطق، أو تقدير معدل البطالة في إحدى الدول، أو تقدير الارتباط بين متوسط عدد ساعات مشاهدة التلفزيون ومستوى الثقافة العامة... الخ.

ب) **اختبارات الفروض**

تستخدم هذه الاختبارات غالباً في البحوث التوكيدية، بهدف اختبار الفروض حول خصائص المجتمع مثل: هل نسبة الذكور في المجتمع هي (40%)؟ هل متوسط دخل الأسرة في المجتمع لا يقل عن (850) دولاراً شهرياً؟ هل يوجد ارتباط طردي قوي بين التدخين والإصابة بمرض سرطان الرئة؟....الخ.

2.1. التصنيف حسب الهدف من البحث:

تختلف أساليب الاستدلال الإحصائي بحسب الهدف من البحث، فالهدف قد يكون:

أ) **دراسة الفروق (الاختلافات) بين المجموعات.**

ب) **دراسة العلاقة (الارتباط) بين متغيرات الدراسة.**

ج) **دراسة التنبؤ، والكشف عن الأثر.**

3.1. التصنيف حسب مستوى القياس للمتغيرات:

يمكن أن يتم تقسيم أساليب الاستدلال الإحصائي حسب مستويات القياس للمتغيرات، فتكون الأساليب مرتبة تنازلياً حسب مستوى القياس كما يلي:

أ) **أساليب القياس الكمي**

1. المستوى النسبي.

2.مستوى الفترة.

ب) أساليب القياس الكيفي

1. المستوى الترتيبي.

2. المستوى الاسمي.

وفي هذا السياق تجدر الإشارة إلى الملاحظات التالية:

- كلما زاد مستوى القياس للمتغيرات أمكن استخدام أساليب إحصائية على مستوى أفضل.
- المتغيرات بمستوى قياس معين يمكن التعامل معها بالأساليب الإحصائية المخصصة لهذا المستوى، وكذا الأساليب الإحصائية المخصصة لمستوى القياس الأقل.
- إن استخدام أسلوب إحصائي مستواه أعلى من مستوى قياس المتغير، يعد خطأ منطقياً، كما أن استخدام أسلوب إحصائي مستواه أقل من مستوى قياس المتغير يعد إهداراً وتضحية ببعض المعلومات المتاحة.

4.1. التصنيف إلى أساليب معلميه وغير معلميه:

يوجد تقسيم آخر شائع الاستخدام لأساليب الاستدلال الإحصائي، حيث يتم تقسيمها إلى أساليب معلميه وأخرى لا معلميه، وأساس هذا التقسيم هو مدى توافر بعض الشروط.

2. تعريفات أولية:

1- **وسيط (معلمة) المجتمع:** هو كل مقياس إحصائي يرتبط بالمجتمع، ويميز مفرادته.

2- **التابع الإحصائي (إحصائية العينة):** هو كل مقياس إحصائي يرتبط بالعينة، ويميز مفرادتها.

3- **مستوى المعنوية (مستوى الأهمية – مستوى الدلالة الإحصائية):** هو الحد الأقصى للوقوع في الخطأ من النوع الأول ونرمز له بـ (α) (احتمال رفض H_0 علماً أنها صحيحة)، فمثلاً إذا كان $\alpha = 0.05$ فهذا يعني أنه إذا تكررت

التجربة مئة مرة متماثلة فإنه يمكن رفض فرضية العدم (وهي في الواقع صحيحة) خمس مرات فقط مقابل خمس وتسعين حالة يكون فيها رفض فرضية العدم غير صحيح، أي نسبة الثقة بالقرار الذي تم أخذه يكون سليماً وصحيحاً بنسبة 95%، مقابل نسبة خطأ 5% بحد أقصى، فدرجة الثقة تساوي $(1-\alpha)$.

3. أساليب التقدير الإحصائي:

وجدنا أن هدف الاستدلال الإحصائي هو القيام بقرارات واستقراءات حول المجتمع المستهدف من خلال المعلومات التي توفرها العينة، على أن تكون هذه الاستقراءات جيدة حول معالم (وسطاء) المجتمع (مثل: المتوسط الحسابي، التباين، الانحراف المعياري، النسبة المئوية ...)، وبالتالي يمكن تقدير وسيط المجتمع أو التنبؤ بقيمته من خلال طرق التقدير النقطي أو طرق التقدير المجالي.

1.3. التقدير النقطي:

يهدف هذا النوع من التقدير إلى تقدير معلمة المجتمع الإحصائي بقيمة واحدة فقط يتم حسابها من بيانات العينة، مثل تقدير المتوسط الحسابي للمجتمع من خلال المتوسط الحسابي للعينة، أو تقدير النسبة للمجتمع من خلال النسبة المتوفرة في العينة، أو تقدير الانحراف المعياري للمجتمع من خلال الانحراف المعياري للعينة، وفي هذا النوع من التقدير يتم استخدام قيمة إحصائية واحدة للاستدلال على معلمة المجتمع.

ويعتبر التقدير (في كثير من الأحيان) تقديراً جيداً إذا احتوى على الصفات المرغوب فيها بالتقدير، مثل: عدم التحيز، الاتساق، الكفاءة النسبية، الكفاية. ويوجد العديد من طرق التقدير منها: طريقة العزوم، وطريقة الإمكان الأكبر، وطريقة المربعات الصغرى.

2.3. التقدير المجالي:

التقدير المجالي هو عبارة عن تقدير معلمة المجتمع الإحصائي ضمن مجال محدد عند احتمال معين بدلالة التابع الإحصائي المقابل له في العينة، آخذين بالحسبان الخطأ

المعياري للتابع الإحصائي المراد التقدير بدلالته، فنحصل على قيمة معلمة المجتمع الإحصائي التي تقع عند ذلك الاحتمال بين حدين أعلى وأدنى، ويطلق على هذين الحدين اسم حدي الثقة أو مجال الثقة أو فترة الثقة، وذلك بافتراض أن العينة تتبع التوزيع الطبيعي، ووقوع التقدير بين هذين الحدين يدعى احتمال الثقة أو درجة الدقة، وإن المتمم الحسابي لاحتمال الثقة يسمى بمستوى الدلالة، فإذا رمزنا بـ (α) لمستوى الدلالة فإن درجة الثقة تساوي ($1-\alpha$) التي تعبر وتدل على المساحة المحصورة تحت المنحني الطبيعي، وبناءً على جدول المساحات تحت المنحني الطبيعي نورد القيم المختلفة لاحتمالات الثقة المستخدمة في الحياة العملية في الجدول التالي:

الجدول رقم (10-1) قيم التوزيع لاحتمالات الثقة

احتمال الثقة	90%	95%	99%
مستوى الدلالة (α)	10%	5%	1%
قيمة التوزيع $z_{\alpha/2}$	1.645	1.96	2.58

ولإيجاد حدي مجال الثقة حول المتوسط نطبق الصيغة الرياضية التالية:

$$CI = \bar{X} \pm z_{\frac{\alpha}{2}} . \sigma_{\bar{X}} \qquad (1-10)$$

حيث:

$\bar{X}$: المتوسط الحسابي للعينة.

$\sigma_{\bar{X}}$: الخطأ المعياري للتقدير ويحسب من القيمة التالية:

$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$ حيث σ الانحراف المعياري للمجتمع المدروس.

وتستخدم الصيغة رقم (10-1) عندما يكون تباين المجتمع معلوماً، وفي حال كان تباين المجتمع مجهولاً وحجم العينة أصغر من 30، تستخدم الصيغة التي تتبع توزيع ستيودنت بدلاً من التوزيع الطبيعي، كما يلي:

$$CI = \bar{X} \pm t_{(\frac{\alpha}{2},df)} S_{\bar{X}} \qquad (2\text{-}10)$$

S_X: الخطأ المعياري للتقدير ويحسب من الصيغة التالية:

$$S_{\bar{X}} = \frac{S}{\sqrt{n}}$$ حيث S الانحراف المعياري للعينة.

ولإيجاد قيمة t الحرجة تستخدم الجداول الخاصة بالتوزيع التائي والتي تتطلب معرفة عدد درجات الحرية (df) أي عدد المشاهدات في تقدير أحد معالم المجتمع الإحصائي. ومن حيث القيمة العددية يكون عدد درجات الحرية في حالة العينة الواحدة: $(df = n-1)$.

وفي حالة العينتين: $(df = n_1 + n_2 - 2)$.

وتتأثر جودة التقدير بصفة عامة بعدة عوامل نذكر منها:

- حجم العينة، فكلما كان حجم العينة كبيراً كان التقدير أكثر كفاءة.
- التباين داخل العينة، فكلما كان التباين صغيراً كان التقدير أكثر كفاءة.
- نوع العينة، فالتقديرات المحسوبة من عينة عشوائية بسيطة أكفأ من التقديرات المحسوبة من الأنواع الأخرى من العينات.
- درجة الثقة المطلوبة (التقدير بفترة)، فكلما كانت درجة الثقة أكبر كانت فترة الثقة أكبر، إلا أن هذا لا يعد ميزة، فكبر فترة الثقة قد لا يفيد كثيراً في النواحي العلمية، وقد يعتبر "تحصيل حاصل".

4. اختبار الفرضيات:

اختبار الفرضيات هو الأسلوب الثاني من أساليب الاستدلال الإحصائي حول أية معلمة من معالم المجتمع. وتعرف الفرضية بشكل عام بأنها تخمين ذكي وتفسير محتمل يتم بواسطته ربط الأسباب بالمسببات، كتفسير مؤقت للمشكلة أو الظاهرة المدروسة، وبالتالي فالفرضية عبارة عن حدس أو تكهن يضعه الباحث كحل ممكن ومحتمل لمشكلة الدراسة.

وتأخذ الفروض غالباً صيغة التعميمات أو المقترحات التي تصاغ بأسلوب منسق ومنظم يظهر العلاقات التي يحاول الباحث من خلالها حل المشكلة. وتشتمل الفرضيات عادة على بعض العلاقات المعروفة كحقائق علمية والتي يقوم الباحث بربطها ببعض الأفكار المتصورة التي ينسجها من خياله أو خبرته البحثية ليعطي بذلك تفسيرات وحلول أولية مقبولة لأوضاع الظاهرة أو المشكلة التي مازالت مجهولة.

واختبار الفروض عبارة عن تحديد فيما إذا كانت بيانات العينة تؤيد اعتقاداً معيناً في المجتمع أم لا، فعلى سبيل المثال يدعي أحد الباحثين الاجتماعيين أن نسبة استخدام وسائل تنظيم الأسرة في مدينته هو 90%، ومن الطبيعي ألا يتم التسليم بادعائه مالم يتم اختبار هذا الادعاء، الذي يعطينا حكماً على قبول أو رفض هذا الادعاء.

واختبار الفرضيات يستخدم في معظم البحوث التجريبية وشبه التجريبية، إذ يلجأ الباحثون إلى صياغة افتراض معين حول ظاهرة ما ويعملون على جمع البيانات الإحصائية المناسبة من أجل الحكم على مدى صحة هذا الافتراض أو عدم صحته.

ويُعتمد عادة على مصادر مختلفة لاشتقاق فرضيات البحث، فبعض الباحثين يستندون إلى الدراسات السابقة، أو النظريات العلمية المختلفة، أو التفسيرات العلمية لحقائق معينة. وبعضهم الآخر يعتمد على أدوات عقلية معينة مثل الحدس، أو الإلهام، أو التخيل أو الاستبصار، كما يمكن اشتقاق فرضيات البحث بالاعتماد على الخبرات

والتجارب الشخصية. وتجدر الإشارة إلى أنه يمكن للباحث أن يقتصر على مصدر واحد لاشتقاق فرضيات بحثه.

وتنقسم فرضيات البحث إلى عدة أنواع، ويمكن التركيز هنا على نوعين رئيسين هما:

1.4. الفرض البحثي:

يشتق الفرض البحثي عادة اشتقاقاً مباشراً من إطار نظري معين، فمثلاً إذا كان الباحث بصدد دراسة تأثير البيئة الأسرية على جُناح الأحداث، ويود على وجه التحديد التعرف إلى أثر الطلاق في التفكك الأسري، أو تأثير عدم التوافق الزوجي على ظاهرة جناح الأحداث، فهو يقوم بالاستناد إلى إطار نظري معين بتوقع زيادة حالات جُناح الأحداث في الأسر المفككة بدرجة أكبر من الأسر المستقرة، ولكي يتأكد أو يتحقق من صحة العلاقة بين التفكك الأسري وجناح الأحداث، يقوم بصياغة توقعه المستمد أساساً من الاطار النظري على شكل فرض بحثي على النحو الآتي: تزداد معدلات جناح الأحداث في الأسر المفككة عنها في الأسر المستقرة.

2.4. الفرض الإحصائي:

تعرف الفرضية الإحصائية على أنها صياغة مبدئية حول معلمة أو أكثر من معالم المجتمع الإحصائي، ويمكن أن تكون صحيحة أو خاطئة بحيث نلجأ إليها لتقدير معالم المجتمع الإحصائي عندما لا تتوافر لدينا معلومات كافية عن المجتمع الإحصائي، فالفرضية الإحصائية تعد بمثابة أفضل تخمين أو تنبؤ لمعالم المجتمع من خلال إحصائيات العينة، وهي تساعد الباحثين على إصدار بعض الأحكام أو اتخاذ القرارات المناسبة حول المعالم الإحصائية في ظل البيانات المتوافرة.

وعند اللجوء إلى استخدام إجراء اختبار الفرضيات الإحصائية، يتطلب الأمر دائماً صياغة نوعين من الفرضيات هي:

1.2.4. الفرضية الصفرية:

تشير الفرضية الصفرية إلى عدم وجود فرق أو عدم وجود علاقة أو عدم وجود أثر بين المتغيرات، وتعود دائماً إلى المجتمع، وتسمى الفرضية الصفرية بالفرضية العدمية أو فرضية النفي، وتصاغ عادة بصيغة النفي، ويرمز لها بالرمز (H_0) وفيما يلي أمثلة على الفرضية الصفرية:

- لا يوجد فروق ذات دلالة إحصائية بين النساء والرجال تجاه الرأي بالعمل التطوعي.
- لا يوجد علاقة ذات دلالة إحصائية بين المستوى التعليمي للأم ومعدل خصوبتها.
- لا يوجد أثر ذو دلالة إحصائية للتدخين على السرطان.

ومن الملاحظ أن الفرضية الصفرية هي التي يتم اختبارها إحصائياً، كما أنه إذا لم يتم قبول الفرضية الصفرية (النفي / العدم)، فإنه يمكن قبول الفرضية البديلة. مع التأكيد بأننا لا نستطيع رفض الفرضية الصفرية ولكن لا نستطيع قبولها، وعندما لا نستطيع قبولها نقبل الفرضية البديلة.

2.2.4. الفرضية البديلة:

تشير الفرضية البديلة إلى وجود فرق أو علاقة أو أثر بين المتغيرات، وتعود دائماً إلى العينة، وتسمى بفرضية البحث أو فرضية الإثبات، وتصاغ عادة بصيغة الإثبات، ويرمز لها بالرمز (H_1) وفيما يلي أمثلة على الفرضية الصفرية:

- يوجد فروق ذات دلالة إحصائية بين النساء والرجال تجاه الرأي بالعمل التطوعي.
- يوجد علاقة ذات دلالة إحصائية بين المستوى التعليمي للأم ومعدل خصوبة الأم.
- يوجد أثر ذو دلالة إحصائية للتدخين على السرطان.

ويمكن للفرضية البديلة أن تأخذ الأشكال التالية:

1) الفرض البديل ذو الاتجاهين (عدم المساواة):

تعكس هذه الفرضية الاختلاف بين المجموعات (وجود فروق بين المجموعات) ولكن لا تحدد اتجاه الاختلاف، أي أنها تحدد بأن المتوسطات غير متساوية فقط دون أن تحدد أيهما أكبر أو أصغر، أو لا تحدد اتجاه التأثير بين المتغيرات، وإنما تهتم فقط بوجوده أو عدم وجوده، ويرمز للفرض البديل ذو الاتجاهين في حالة مقارنة المتوسطات بما يلي:

$$H_1 : \bar{X}_1 \neq \bar{X}_2$$

وفي حالة دراسة الارتباط (الأثر) يمكن التعبير عن الفرض البديل بالرموز على النحو التالي:

$$H_1 : R \neq 0$$

2) الفرض البديل المُتجهة (من اتجاه واحد):

تعكس الفرضية المُتجهة الاختلاف بين المجموعات، مع تحديد اتجاه الاختلاف إن كان أكبر أو أصغر، ففي حالة مقارنة المتوسطات، يمكن التعبير عنها بالرموز كما يلي:

$$H_1 : \bar{X}_1 < \bar{X}_2$$
$$H_1 : \bar{X}_1 > \bar{X}_2$$

أما في حالة الارتباط، فإنه يحدد نوعية التأثير (إيجابي أو سلبي) كأن نقول إن هناك علاقة (طردية أو عكسية) بين المتغيرين، يمكن التعبير عنها بالرموز كما يلي:

$$H_1 : R < 0$$
$$H_1 : R > 0$$

وتكمن أهمية الفرض البديل في كونه يحدد قيمة الدرجة الحرجة التي تستخدم للتحقق من الفرضية إحصائياً، فإذا كان الفرض البديل عديم الاتجاه فإن القيم المحسوبة

كنتائج للبحث تتم مقارنتها مع التوزيع النظري بما يسمى اختبار من الطرفين، أما إذا كان الفرض البديل ذا اتجاه مُحدد فنقارن النتائج مع التوزيع النظري بما يسمى اختبار من طرف واحد (اتجاه واحد).

3.4. الأخطاء المحتملة والمتعلقة باختبار الفرضيات:

لا يوجد قرار إحصائي منزه عن الخطأ، فالقرارات الإحصائية هي دائماً قرارات احتمالية، بمعنى أنه لا مفر من وجود احتمال للخطأ في أي قرار يصدره الباحث حول المجتمع عن طريق عينة عشوائية مأخوذة منه. ولما كانت هذه القرارات مؤسسة على ما يجريه الباحث من اختبارات للفرضيات وتزيد الثقة بزيادة دقة هذه الاختبارات، وجب على الباحث أن يدرس كيف يزيد من هذه الدقة، أي من قدرة الاختبارات على اتخاذ قرار سليم الذي لا يشوبه إلا قدر ضئيل من الخطأ، ويتأتى ذلك عن طريق التحكم ما أمكن في احتمالات الأخطاء التي تنجم حتماً عند استخدام الاختبارات.

وهناك أربعة احتمالات تنبثق عن قبول أو رفض الفرضية الصفرية مبينة في الجدول التالي:

الجدول رقم (10-2)

فرضية العدم H0	القرار	
	رفض H0	قبول H0
H0 صحيحة	خاطئ	صحيح
H0 خاطئة	صحيح	خاطئ

من الجدول السابق نجد أن القرار خاطئ في حالتين:

1- رفض فرضية العدم H_0 بينما هي في الواقع صحيحة، ويدعى مثل هذا الخطأ بالخطأ من النوع الأول Type I error، ويرمز له بالرمز α ويدعى عادة بحجم الاختبار، أو هو مستوى الأهمية أو مستوى الدلالة الإحصائية أو مستوى المعنوية.

2- قبول فرضية العدم H_0 بينما هي في الواقع خاطئة، ويدعى مثل هذا الخطأ بالخطأ من النوع الثاني Type II error، ويرمز له بالرمز β .

فمثلاً شخص ما متّهم بالقتل قُدّم للمحاكمة، وكما هو معلوم في القضاء أن المتهم بريء حتى تثبت إدانته، بمعنى أن فرضية العدم H_0 هي الشخص المتهم بريء.

يعتمد قرار المحكمة في هذه القضية على المعلومات المتوفرة بهذا الخصوص، ويكون لدينا الحالات الأربع المبينة في الجدول التالي:

جدول رقم (10-3)

الشخص المتهم بريء H0	قرار المحكمة	
	الشخص المتهم قاتل (رفض H0)	الشخص المتهم بريء (قبول H0)
H0 صحيحة	خاطئ	صحيح
H0 خاطئة	صحيح	خاطئ

نلاحظ أن الخطأ من النوع الأول يتمثل في قرار المحكمة بأن الشخص المتهم قاتل بينما هو في الحقيقة بريء، أما الخطأ من النوع الثاني فيتمثل بقرار المحكمة بأن المتهم بريء في حين أنه قاتل فعلاً.

وهكذا عند اختيار فرضية H_0 يمكن الوصول إلى قرار صحيح أو الوقوع في خطأ من النوع الأول أو الثاني.

4.4. مستوى المعنوية:

في اختبار فرضية معينة يكون مستوى معنوية الاختبار هو أقصى احتمال يمكن أن تتحمل به خطأ من النوع الأول ونرمز له بـ α ويحدد بشكل عام قبل سحب أي عينة، وعادة نستخدم $\alpha = 0.05$ أو $\alpha = 0.01$ علماً أن هناك قيماً أخرى يمكن استخدامها أيضاً، فمثلاً إذا استخدمنا $\alpha = 0.05$ كمستوى معنوية لتصميم اختبار

للفرضية، فهذا يعني أن هناك حوالي 5 فرص من ضمن كل 100 عينة من الحجم n نفسه سوف نرفض من خلالها فرضية العدم H_0 في حين هي صحيحة.

5.4. قوة الاختبار الإحصائي:

تعني احتمال رفض فرضية العدم H_0 وهي خاطئة، أي تتحدد قوة الاختبار بالفرق بين الواحد الصحيح وقيمة احتمال الخطأ من النوع الثاني (β)، أي أن قوة الاختبار ($1-\beta$)، وتعني قوة الاختبار المفاضلة بين اختبارات لها نفس مستوى المعنوية (α) وتعطي أصغر قيمة لـ (β).

وعند إجراء أي اختبار نطمح بأن يكون (β و α) أصغر ما يمكن، وهذا الأمر صعب في الاختبار الاحصائي، وبالتالي نطمح إلى إيجاد قاعدة أو اختبار إحصائي لقبول أو رفض فرضية العدم بحث تكون احتمالات الوقوع في خطأ من النوع الأول والثاني أقل ما يمكن، وكما هو معروف فإنه إذا صغرت قيمة α فإن قيمة β ستكبر (والعكس صحيح فإذا كبرت قيمة α فإن قيمة β ستصغر) لذلك لجأ الإحصائيون إلى قاعدة مفادها تثبيت قيمة α والبحث عن اختبار يجعل قيمة β أصغر ما يمكن.

5. اختبار التوزيع الطبيعي:

التوزيع الطبيعي Normal Distribution هو أشهر التوزيعات الاحتمالية ويعود ذلك لأن الكثير من الظواهر تتبع التوزيع الطبيعي، وهناك اختبارات تحتاج فيها لمعرفة ما إذا كان المجتمع الذي سحبت منه عينة الدراسة تتبع توزيعاً معيناً كالتوزيع الطبيعي أم لا، فاختبارات ت ستيودنت (T-Test) تتطلب أن يكون توزيع المجتمع طبيعياً. لذلك فإن التوزيع الطبيعي هو شيء محوري في علم الإحصاء، ومنحني التوزيع الطبيعي يشبه الجرس (الناقوس) ويتميز بوجود تماثل بين جانبيه الأيمن والأيسر حول المتوسط، ومن سمات منحني التوزيع الطبيعي أن المتوسط يساوي الوسيط ويساوي المنوال. وفي الشكل

التالي لمنحني التوزيع الطبيعي يلاحظ تماثل المنحني يعني أن 50% من القيم هي أقل من المتوسط و 50% من القيم هي أكبر من المتوسط وهذا يعني أن الوسيط يساوي المتوسط ويساوي المنوال.

الشكل رقم (10-1) منحنى التوزيع الطبيعي

وسنتعرف على عدد من الطرق التي تستخدم في اختبار التوزيع الطبيعي للبيانات، ومنها:

1.5. اعتماداً على العلاقة بين مقاييس النزعة المركزية:

يمكن معرفة شكل التوزيع للعينة من خلال الاعتماد على العلاقة بين الوسط الحسابي والوسيط والمدى والتي تسمى علاقة بيرسون، عندها نميز ثلاث حالات:

أ- إذا كان $Mode < Median < Mean$ فإن التوزيع ملتو نحو اليمين ولا يتبع التوزيع الطبيعي.

ب- إذا كان $Mode > Median > Mean$ فإن التوزيع ملتو نحو اليسار ولا يتبع التوزيع الطبيعي.

ج- إذا كان $Mode = Median = Mean$ فإن التوزيع متماثل ويتبع التوزيع الطبيعي.

2.5. اعتماداً على معامل الالتواء والتفلطح:

يمكن معرفة شكل التوزيع للعينة من خلال الاعتماد على معامل الالتواء (S.K) والتفلطح (Ku)، عندها نميز ثلاث حالات:

أ- S.K>0 و Ku >3 يكون التوزيع موجباً (ملتوٍ نحو اليمين) ولا يتبع التوزيع الطبيعي.

ب- S.K<0 و Ku <3 يكون التوزيع سالباً (ملتوٍ نحو اليسار) ولا يتبع التوزيع الطبيعي.

ج- S.K=0 و Ku=3 يكون التوزيع متماثلاً ويتبع التوزيع الطبيعي.

3.5. اعتماداً على الأشكال البيانية:

من خصائص منحني التوزيع الطبيعي أن يكون متماثلاً حول الوسط الحسابي فدراسة التوزيع الطبيعي باستخدام الأشكال البيانية يعتمد على مفهوم التماثل عن طريق إسقاط عمود من قمة المنحنى بحيث يكون الجزيئان المقسمان متماثلين (متماثلين حول العمود) فتتبع البيانات التوزيع الطبيعي أما إذا كانا غير متساويين فإن البيانات عندها ليس لها توزيع الطبيعي، وسوف نستخدم الأشكال التالية:

1.3.5. المدرج التكراري Histogram:

المدرج التكراري هو أحد الرسومات البيانية التي تستخدم فقط مع المتغيرات الكمية المتصلة، ويعطينا المدرج التكراري معلومات غزيرة في شكل بسيط، فهو يمكننا من فهم البيانات وتوزيعها، ويكون للبيانات توزيع طبيعي إذا كانت الأعمدة التي تقع في الوسط هي الأعلى ارتفاعاً كما هو موضح في الشكل التالي:

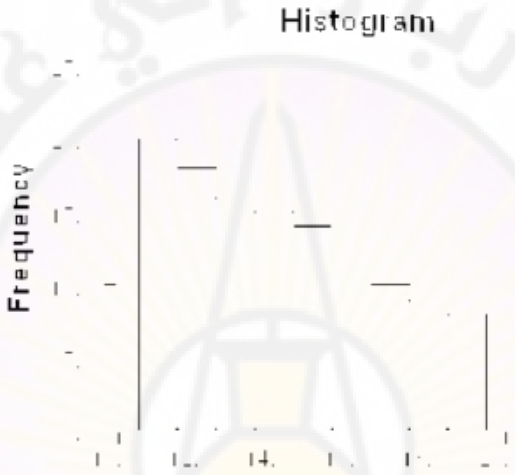
الشكل رقم (10-2) المدرج التكراري مع منحنى التوزيع الطبيعي

حيث نلاحظ أن الأعمدة الأعلى ارتفاعاً تقع في الوسط، كما أن منحني التوزيع الطبيعي متناظر (متماثل) على طرفي المتوسط، أما في الشكل رقم (10-3) نلاحظ أن الأعمدة الأكثر ارتفاعاً في الجهة اليمنى من التوزيع:

الشكل رقم (10-3) المدرج التكراري مع منحنى التوزيع الطبيعي

فالمنحني غير متناظر بالنسبة للمتوسط، وبالتالي ليس هناك توزيع طبيعي للبيانات، كما أنه في الشكل رقم (10-4) نلاحظ أن الأعمدة الأكثر ارتفاعاً في الجهة اليسرى من التوزيع:

الشكل رقم (10-4) المدرج التكراري مع منحنى التوزيع الطبيعي



فالمنحني غير متناظر بالنسبة للمتوسط، وبالتالي لا يوجد توزيع طبيعي للبيانات.

2.3.5. مخطط الصندوق ذو القرنين:

هو طريقة للتمثيل البياني لمجموعة من القيم العددية لعينة احصائية من خلال تمثيل المؤشرات الإحصائية التالية من الأدنى إلى الأعلى: القيمة الصغرى Min، الربيع الأول Q1، الوسيط Q2، الربيع الثالث Q3، والقيمة العظمى Max. ويظهر مخطط الصندوق القيم الشاذة والمتطرفة.

وتتبع البيانات التوزيع الطبيعي إذا كان البعد بين الربيع الأدنى والوسيط يساوى البعد بين الربيع الأعلى والوسيط، كما في الشكل التالي:

الشكل رقم (10-5) الصندوق ذو القرنين في حالة التوزيع الطبيعي

أما إذا كان الوسيط أقرب إلى الربيع الثالث فيكون التوزيع ذو التواء نحو اليسار (التواء سلبي) وبالتالي لا يوجد توزيع طبيعي للبيانات. كما في الشكل التالي:

الشكل رقم (10-6) الصندوق ذو القرنين في حالة الالتواء السلبي

أما إذا كان الوسيط أقرب إلى الربيع الأول فيكون التوزيع ذو التواء نحو اليمين (التواء إيجابي) وبالتالي لا يوجد توزيع طبيعي للبيانات. كما في الشكل التالي:

الشكل رقم (10-6) الصندوق ذو القرنين في حالة الالتواء السلبي

4.5. اختبار كولموغروف- سميرنوف أو شابيرو- ويلك:

يستخدم هذان الاختباران لمعرفة طبيعة توزيع بيانات ظاهرة معينة في كونها تتبع التوزيع الطبيعي (الاعتدالي) من عدمه، وهذان الاختباران ضروريان في اختبار الفرضيات لأن معظم الاختبارات المعلمية تشترط أن يكون توزيع البيانات طبيعيًا.

مع ملاحظة أنه يستخدم اختبار كولمجروف - سميرنوف (Kolmogorov-Smirnov) لمعرفة توزيع البيانات إذا كان حجم العينة أكبر من أو يساوي 50، بينما يستخدم اختبار شابيرو-ويلك (Shapiro-Wilk) إذا كان حجم العينة أقل من 50.

وفي هذا الاختبار نضع الفرضيات التالية:

الفرضية الابتدائية (H_0): البيانات تتبع التوزيع الطبيعي عند مستوى دلالة إحصائية α.

الفرضية البديلة (H_1): البيانات لا تتبع التوزيع الطبيعي عند مستوى دلالة إحصائية α.

ولمناقشة الفرضية نعتمد على اختبار كولمجروف - سميرنوف أو شابيرو-ويلك حسب قيمة المعنوية (Sig) إذا كانت قيمة المعنوية Sig أكبر من α نقبل الفرضية الابتدائية (H_0)، أما إذا كانت قيمة المعنوية Sig أصغر من α نرفض الفرضية الابتدائية (H_0).

مثال 1:

بفرض لدينا البيانات التالية:

4.5 ، 5.2 ، 4.9 ، 4.3 ، 4.6 ، 4.8 ، 4.6 ، 4.9 ،

4.5 ، 5.0 ، 4.8 ، 4.6، 4.6 ، 4.7 ، 4.5 ، . 4.7

والمطلوب: هل البيانات تتبع التوزيع الطبيعي وذلك عند مستوى دلالة إحصائية 0.05.

الحل:

بعد أن ندخل المتغير ونعرفه في شاشة (Variable View ضمن Data Editor)، ثم ندخل البيانات في شاشة (Data View ضمن)، نجري اختيار التوزيع الطبيعي كما يلي:

فرضيات التوزيع الطبيعي:

- الفرضية الصفرية: بيانات العينة تتبع التوزيع الطبيعي عند مستوى دلالة إحصائية 0.05.
- الفرضية البديلة: بيانات العينة لا تتبع التوزيع الطبيعي عند مستوى دلالة إحصائية 0.05.

نجري الاختبار باستخدام برنامج (SPSS) نتبع ما يلي:

Analyze Descriptive statistics Explore

فيظهر صندوق حوار بعنوان (Explore) نضع المتغير الكمي في المستطيل Dependent List كما يلي:

فتظهر شاشة المخرجات نهتم فقط بجدول التوزيع الطبيعي والمعنون بـ (Tests of Normality) والذي يكون له الشكل التالي:

من الجدول المعنون بـ Test of Normality نختار اختبار شابيرو- ويلك (Shapiro-Wilk) لأن حجم العينة أصغر من 50. فنجد أن قيمة المعنوية أكبر من قيمة الدلالة الإحصائية 0.05، وبالتالي البيانات تتوزع وفق التوزيع الطبيعي، ونقبل الفرضية الابتدائية.

مثال 2:

لتكن لدينا درجات طلاب في مقرر الإحصاء في كلية الاقتصاد في مقرر الإحصاء وذلك لطلاب شعبتي إدارة الأعمال (management) والمحاسبة (accounting) فكانت النتائج كما يلي:

management	75	74	83	75	70	73	76	80	85	79
accounting	60	79	79	76	85	64	90	91	95	69

والمطلوب: هل درجات الطلاب تتبع التوزيع الطبيعي وذلك عند مستوى دلالة إحصائية 0.05 في الشعبتين.

الحل:

ندخل متغير الدرجات ونعرفه في شاشة (Variable View ضمن Data Editor)، ثم ندخل متغيراً ثانياً هو متغير التقسيم وهو متغير نوع الشعبة ويحوي مجموعتين (إدارة الأعمال- المحاسبة)، كما في الجدول التالي:

اسم المتغير	Name	Type	Values	Measurement
درجات الطلاب	درجات_الطلاب	Numeric	×	Scale
نوع الشعبة	نوع_الشعبة	Numeric	1= إدارة الأعمال 2= محاسبة	Nominal

لإجراء الاختبار باستخدام برنامج (SPSS) نتبع ما يلي:

نضع فرضيات التوزيع الطبيعي:

الفرضية الأولى: الفرضية الصفرية: درجات طلاب شعبة المحاسبة تتبع التوزيع الطبيعي

الفرضية البديلة: درجات طلاب شعبة المحاسبة لا تتبع التوزيع الطبيعي.

الفرضية الثانية: الفرضية الصفرية: درجات طلاب شعبة إدارة الأعمال تتبع التوزيع الطبيعي.

الفرضية البديلة: درجات طلاب شعبة إدارة الأعمال لا تتبع التوزيع الطبيعي.

ومن أجل إجراء هذا الاختبار حاسوبياً باستخدام برنامج (SPSS) نتبع ما يلي:

Analyze Descriptive statistics Explore

فيظهر لدينا صندوق حوار بعنوان Explore نضع المتغير الكمي وهو الدرجات في المستطيل Dependent List ونضع متغير التقسيم (المجموعات) وهو متغير نوع الشعبة في مستطيل Factor List كما يلي:

نهتم فقط بجدول التوزيع الطبيعي والمعنون بـ (Tests of Normality) والذي يكون له الشكل التالي:

نختار اختبار شابيرو – ويلك (Shapiro-Wilk) لأن حجم العينة أصغر من 50.

من خلال الجدول السابق نجد أن الجدول يحوي على سطرين من خلال كل سطر يمكننا أن نناقش كل فرضية على حدة كما يلي:

الفرضية الأولى: من السطر الأول الذي يحوي على السطر المعنون المحاسبة (accounting) نجد قيمة المعنوية أكبر من قيمة الدلالة الإحصائية 0.05 وبالتالي يمكننا قبول الفرضية الصفرية الأولى أي: درجات طلاب شعبة المحاسبة تتوزع وفق التوزيع الطبيعي.

الفرضية الثانية: من السطر الثاني الذي يحوي على السطر المعنون إدارة الأعمال (management) نجد قيمة المعنوية أكبر من قيمة الدلالة الإحصائية 0.05 وبالتالي يمكننا قبول الفرضية الصفرية الثانية أي: درجات طلاب شعبة إدارة الأعمال تتوزع وفق التوزيع الطبيعي.

6. اختبار الفرضية المتعلقة بوسط حسابي واحد:

يستخدم هذا الاختبار لمعرفة فيما إذا كان وسط العينة يختلف اختلافاً معنوياً عن وسط المجتمع الذي سحبت منه العينة، ولتطبيقه يجب أن تكون المفردات مستقلة

ويتحقق ذلك بالاختيار العشوائي للمفردات، وعند اختبار الفرضية المتعلقة به نكون أمام حالتين للاختبار الإحصائي:

الحالة الأولى: استخدام الاختبار الإحصائي الطبيعي:

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \qquad (3\text{-}10)$$

وذلك عندما يكون تباين المجتمع (أو الانحراف المعياري للمجتمع) معلوم أو حجم العينة أكبر من 30 مفردة حتى ولو لم يكن تباين المجتمع معلوماً.

الحالة الثانية: استخدام الاختبار الإحصائي ستيودنت:

$$t = \frac{\bar{X} - \mu}{s / \sqrt{n}} \qquad (4\text{-}10)$$

وذلك عندما يكون تباين المجتمع (أو الانحراف المعياري للمجتمع) مجهولاً أو حجم العينة صغيراً جداً أصغر من 30 مفردة.

تطبيق مقارنة متوسط العينة مع متوسط المجتمع الذي سحبت منه العينة باستخدام البرنامج الإحصائي SPSS:

وقبل البدء بدراسة الأساليب المتاحة، هناك بعض الأمور العامة التي يجب التحقق منها قبل البدء بتنفيذ الاختبار، وهذه الافتراضات هي:

1) مستوى القياس: يجب أن يكون المتغير الذي نجري عليه الاختبار كمياً (مستمراً – منفصلاً) ويكون مقياسه (فترة – نسبة).
2) المعاينة العشوائية: يجب اختيار العينة بشكل عشوائي ودون تحيز.
3) استقلالية المشاهدات: لا بد أن تكون كل المشاهدات التي تتكون منها البيانات مستقلة عن بعضها بعضا أي أنه يجب ألا تكون القياسات أو عملية قياس المشاهدات متأثرة بأية مشاهدة أو عملية قياس مشاهدة أخرى، وهذا الافتراض مهم

جداً وتتم عملية التشكيك في موقف تتم عندما نجمع المشاهدات أو القياسات في موقف جماعي أو يكون فيه شكل ما من التفاعل بين أفراد التجربة.

4) التوزيع الطبيعي: يجب أن تكون المشاهدات أو القياسات تتبع التوزيع الطبيعي، فإذا لم تكن تتبع التوزيع الطبيعي يجب أن يكون حجم العينة أكبر من 30 مشاهدة، وسنقوم باستعراض آلية إجراء اختبار التوزيع الطبيعي ضمن كل حالة.

مثال 3:

بفرض لدينا النتائج التالية:

4.5 ، 5.2 ، 4.9 ، 4.3 ، 4.6 ، 4.8 ، 4.6 ،4.9 ،

4.7 ، 4.5 ، 4.7 ، 4.6 ، 4.6 ، 4.8 ، 5.0 ، 4.5

والمطلوب: اختبر الفرض القائل أن متوسط البيانات يساوي 5، وذلك عند مستوى المعنوية 0.05.

الحل:

سيتم اختبار الفرضيات التالية:

الفرضية الصفرية: لا يوجد فروق ذات دلالة معنوية (إحصائية) بين متوسط البيانات والقيمة 5 عند مستوى دلالة إحصائية 0.05.

الفرضية البديلة: يوجد فروق ذات دلالة معنوية (إحصائية) بين متوسط البيانات والقيمة 5 عند مستوى دلالة إحصائية 0.05.

خطوات تنفيذ الاختبار في برنامج (SPSS):

1- ندخل البيانات ونعرفها في شاشة (Variable View ضمن Data Editor)، ثم ندخل البيانات في شاشة (Data View ضمن Data Editor).

2- نجري اختيار التوزيع الطبيعي، وذلك لمعرفة هل البيانات تتبع التوزيع الطبيعي أم لا، والهدف من هذا الاختبار هو تحديد نوع الاختبار المستخدم من أجل مقارنة

متوسط العينة مع متوسط المجتمع الذي سحبت منه العينة (مر معنا سابقاً هذا الاختبار).

نضع أولاً فرضيات التوزيع الطبيعي:

- الفرضية الصفرية: بيانات العينة تتبع التوزيع الطبيعي عند مستوى دلالة إحصائية 0.05.
- الفرضية البديلة: بيانات العينة لا تتبع التوزيع الطبيعي عند مستوى دلالة إحصائية 0.05.

في شاشة المخرجات نهتم فقط بجدول التوزيع الطبيعي والمعنون بـ (Tests of Normality) والذي يكون له الشكل التالي:

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
weight	.172	16	.200*	.962	16	.691

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

نلاحظ أن البرنامج يقوم بإجراء اختبار التوزيع الطبيعي للبيانات باستخدام اختبارين هما كولموكجروف – سميرنوف (Kolmogorov-Smirnov) واختبار شابيرو – ويلك (Shapiro-Wilk)، فنجد أن قيمة المعنوية في كلا الاختبارين أكبر من 0.05 وبالتالي نقبل التوزيع الطبيعي للبيانات.

3- من أجل اختبار مقارنة المتوسط (الفرضية الرئيسية) من شريط الأوامر:

Analyze ⟶ Compare Means ⟶ One Samples T Test

4- نحصل على صندوق حوار يحوي ثلاثة مستطيلات موزعة كما يلي:

- المستطيل الأول: يحوي جميع المتغيرات التي تحويها قاعدة البيانات.

- المستطيل الثاني والمعنون بـ ((Test Variable(s): ونضع فيها المتغير (المتغيرات) التي نريد اختبار متوسطها.
- المستطيل الثالث والمعنون بـ (Test Value): نضع فيه قيمة المتوسط الذي نريد مقارنته (متوسط المجتمع) وهي في حالتنا 5.

وبالضغط على زر (Options..) يفيدنا بتغير قيمة المعنوية وبالتالي تغير حدود مجال الثقة وأيضاً في التعامل مع القيم المفقودة.

5- نضغط على زر (OK) فتظهر شاشة المخرجات، والتي تحوي الجدولين الآتيين:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
[illegible]	16	4.7000	.22509	.05627

One-Sample Test

	Test Value = 5					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
weight	-5.331	15	.000	-.30000	-.4199	-.1801

ولنفسر الجداول التي حصلنا عليها من شاشة المخرجات:

✓ **الجدول رقم (1):**

ويحوي بعض المؤشرات الإحصائية للمتغير (للعينة) الذي نقارن متوسطه مع المتوسط المعلوم حيث يحوي:

One-Sample Statistics

	عدد الحالات	المتوسط الحسابي	الإنحراف المعياري	الخطأ المعياري
weight	16	4.7000	.22509	.05627

✓ **الجدول رقم (2):**

ويحوي معلومات عن دالة الاختبار وقيمتها وقيمة المعنوية كما يبينه الجدول التالي:

One-Sample Test

	القيمة(المتوسط) المختبرة (المقارن)=5					
	ت المحسوبة	عدد درجات الحرية	قيمة المعنوية	الفرق بين المتوسطين	95%مجال الثقة حول الفرق بين المتوسطين	
					الحد الأدنى	الحد الأعلى
weight	-5.331-	15	.000	-.30000-	-.4199-	-.1801-

من خلال هذا الجدول يمكننا أن نتخذ القرار بقبول أو رفض الفرضية، وذلك بعدة طرق:

الطريقة الأولى: نعتمد على قيمة المعنوية (significance value) والمبينة في الجدول الثاني ((Sig. (2-tailed) فإذا كانت قيمة المعنوية أكبر من مستوى المعنوية يكون هناك دليل قوي لقبول الفرضية الابتدائية وإذا كانت قيمة المعنوية أصغر من مستوى المعنوية يكون هناك دليل قوي لرفض الفرضية الابتدائية، وفي حالتنا هنا نجد:

$$\text{Sig. (2-tailed)}=0.00<0.05$$

وبالتالي لدينا إثبات قوي لرفض الفرضية الابتدائية.

الطريقة الثانية: الطريقة الثالثة المستخدمة في الاختبار تعتمد على مجال الثقة حول الفرق بين المتوسطين (Confidence Interval of the Difference) والمبينة في الجدول الثاني فإذا كانت حدود مجال الثقة من إشارتين مختلفتين يكون هناك دليل قوي لقبول الفرضية الابتدائية وإذا كانت حدود مجال الثقة من نفس الإشارة يكون هناك دليل قوي لرفض الفرضية الابتدائية، وفي حالتنا هنا نجد أن الحد الأدنى والحد الأعلى لها إشارة سالبة، وبالتالي لدينا إثبات قوي لرفض الفرضية الابتدائية.

كما يحتوي الجدول رقم (2) على حدود مجال الثقة حول الفرق بين المتوسطين (95 % Confidence Interval of the Difference) حيث نجد أن:

$[-0.4199\ ,\ -0.1801]$ ويقرأ مجال الثقة: أننا واثقون بمقدار 95% على أن الفرق بين المتوسطين لن يزيد عن -0.1801 ولن يقل -0.4199.

كتابة النتائج:

استخدم اختبار ستيودنت t لاختبار معنوية الفرق بين متوسط البيانات والمتوسط المعلوم 5. فقد وجد أنه يوجد فروق معنوية بين متوسط البيانات حيث بلغ (4.7)

بانحراف معياري (0.225) والمتوسط المعطى (5) حيث بلغت قيمة t المحسوبة (-5.331) وهي غير دالة إحصائياً عند مستوى المعنوية (0.05).

7. مقارنة وسطي عينتين مستقلتين:

تعرف العينات المستقلة بأنها العينات التي لا يؤثر فيها اختيار الفرد في العينة الأولى على اختيار الفرد في العينة الثانية، والبيانات المأخوذة منهما تكون مستقلة بمعنى أنه لا يوجد بينهما ارتباط، ويمكن تحقيق ذلك بالاختيار العشوائي.

وتكون دالة الاختبار الإحصائي في حالة مقارنة وسطي عينتين مستقلتين على النحو الآتي:

الحالة الأولى: الاختبار الإحصائي الطبيعي كما يلي:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \qquad (4-10)$$

حيث:

$\bar{X}_1$: وسطي العينة الأولى	σ_1 : الانحراف المعياري للعينة الأولى	n_1 : حجم العينة الأولى
$\bar{X}_2$: وسط العينة الثانية	σ_2 : الانحراف المعياري للعينة الثانية	n_2 : حجم العينة الثانية

وتستخدم عندما يكون تباين المجتمع الأول معلوم، وتباين المجتمع الثاني معلوم أو كلاً من حجم العينة للمجتمع الأول وحجم العينة للمجتمع الثاني أكبر من 30 والتباينات للمجتمع مجهولة عندها نقدر σ_2, σ_1 بـ S_2, S_1 تباين العينة الأولى والثانية على الترتيب وتصبح قيمة دالة الاختبار الاحصائي كما يلي:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \qquad (5-10)$$

الحالة الثانية: الاختبار الاحصائي ستيودنت كما يلي:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_C \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad (6\text{-}10)$$

حيث:

$$S_C = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} \qquad (7\text{-}10)$$

وتستخدم في حالة العينات الصغيرة الحجم (حجم العينة أقل من 30) والتباينات للمجتمع مجهولة.

مقارنة متوسطي عينتين مستقلتين باستخدام البرنامج الإحصائي SPSS:

قبل البدء بدراسة الأساليب المتاحة، هناك بعض الأمور العامة المشتركة التي يجب التحقق منها قبل البدء بتنفيذ الاختبار، وهذه الافتراضات هي:

1) مستوى القياس: يجب أن يكون المتغير الذي نجري عليه الاختبار كمياً (مستمراً – منفصلاً) ويكون مقياسه (فترة – نسبة).
2) المعاينة العشوائية: يجب اختيار العينة بشكل عشوائي ودون تحيز.
3) استقلالية المشاهدات: لا بد أن تكون كل المشاهدات التي تتكون منها البيانات مستقلة عن بعضها بعضا أي أنه يجب ألا تكون القياسات أو عملية قياس المشاهدات متأثرة بأية مشاهدة أو عملية قياس مشاهدة أخرى، وهذا الافتراض مهم جداً وتتم عملية التشكيك في موقف عندما نجمع المشاهدات أو القياسات في موقف جماعي أو يكون فيه شكل ما من التفاعل بين أفراد التجربة.

4) التوزيع الطبيعي: يجب أن تكون المشاهدات أو القياسات تتبع التوزيع الطبيعي، فإذا لم تكن تتبع التوزيع الطبيعي يجب أن يكون حجم العينة أكبر من 30 مشاهدة، وسنقوم باستعراض آلية إجراء اختبار التوزيع الطبيعي ضمن كل حالة.
5) تجانس البيانات: ويقصد بهذا الشرط أن البيانات مأخوذة من مجتمعات ذات تباينات متساوية أي تماثل تغيرية الدرجات لكل مجموعة من المجموعات، وسنقوم باستعراض آلية إجراء اختبار التجانس ضمن الاختبارات التي تتطلب إجراء اختبار التجانس.

مثال 4:

لتكن لدينا درجات طلاب في مقرر الإحصاء وذلك لطلاب شعبتي الإدارة (management) والمحاسبة (accounting) فكانت النتائج كما يلي:

management	75	74	83	75	70	73	76	80	85	79
accounting	60	79	79	76	85	64	90	91	95	69

والمطلوب: هل هناك فروق ذات دلالة معنوية عند مستوى المعنوية (0.05) بين متوسطي درجات الطلاب في الشعبتين.

الحل:

سيتم اختبار الفرضيات التالية:

الفرضية الصفرية: لا يوجد فروق ذات دلالة معنوية (إحصائية) بين متوسطي درجات الطلاب في الشعبتين عند مستوى دلالة إحصائية 0.05.

الفرضية البديلة: يوجد فروق ذات دلالة معنوية (إحصائية) بين متوسطي درجات الطلاب في الشعبتين عند مستوى دلالة إحصائية 0.05.

خطوات تنفيذ الاختبار في برنامج (SPSS):

1- ندخل متغير الدرجات ونعرفه في شاشة (Variable View ضمن Data Editor)، ثم ندخل البيانات في شاشة (Data View)، ثم ندخل متغير ثاني هو متغير التقسيم وهو متغير نوع الشعبة ويحوي مجموعتين (إدارة الأعمال- المحاسبة) كما مر معنا سابقاً.

2- نجري اختبار التوزيع الطبيعي:

نضع أولاً فرضيات التوزيع الطبيعي:

الفرضية الأولى: الفرضية الصفرية: درجات طلاب شعبة المحاسبة تتبع التوزيع الطبيعي.

الفرضية البديلة: درجات طلاب شعبة المحاسبة لا تتبع التوزيع الطبيعي.

الفرضية الثانية: الفرضية الصفرية: درجات طلاب شعبة إدارة الأعمال تتبع التوزيع الطبيعي.

الفرضية البديلة: درجات طلاب شعبة إدارة الأعمال لا تتبع التوزيع الطبيعي.

فتظهر شاشة المخرجات نهتم فقط بجدول التوزيع الطبيعي والمعنون بـ (Tests of Normality) والذي يكون له الشكل التالي:

Tests of Normality

	DEPARTMENT	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
DEGREE	accounting	.185	10	.200*	.959	10	.779
	management	.129	10	.200*	.955	10	.725

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

نختار اختبار شابيرو – ويلك (Shapiro-Wilk) لأن حجم العينة أصغر من 50.

من خلال الجدول السابق نجد أن الجدول يحوي على سطرين من خلال كل سطر يمكننا أن نناقش كل فرضية على حدة كما يلي:

الفرضية الأولى: من السطر الأول الذي يحوي على السطر المعنون المحاسبة (accounting) نجد قيمة المعنوية أكبر من قيمة الدلالة الإحصائية 0.05 وبالتالي يمكننا قبول الفرضية الصفرية الأولى أي: درجات طلاب شعبة المحاسبة تتوزع وفق التوزيع الطبيعي.

الفرضية الثانية: من السطر الثاني الذي يحوي على السطر المعنون إدارة الأعمال (management) نجد قيمة المعنوية أكبر من قيمة الدلالة الإحصائية 0.05 وبالتالي يمكننا قبول الفرضية الصفرية الثانية أي: درجات طلاب شعبة إدارة الأعمال تتوزع وفق التوزيع الطبيعي.

3- من أجل اختبار مقارنة المتوسطين (الفرضية الرئيسية) من شريط الأوامر :

4- نحصل على صندوق حوار يحوي ثلاثة مستطيلات موزعة كما يلي:

المستطيل الأول: يحوي جميع المتغيرات التي تحويها قاعدة البيانات.

المستطيل الثاني والمعنون بـ ((Test Variable(s): ونضع فيه المتغير (المتغيرات) التي نريد اختبار متوسطها وهي المتغيرات الكمية وفي مثالنا نضع متغير الدرجة Degree.

المستطيل الثالث والمعنون بـ (Grouping Variables): ونعرف فيها المجموعات التي تقسم البيانات وفي مثالنا نضع Department، ثم نضغط على (Define Groups..) لتحديد المجموعات فنضغط الرقم المعطى للمجموعة الأولى 1 والرقم المعطى للمجموعة 2.

ثم نضغط على (continue) كما في الشكل التالي:

وبالضغط على زر (Options..) يفيدنا بتغير قيمة المعنوية وبالتالي تغير حدود مجال الثقة وأيضاً في التعامل مع القيم المفقودة.

5- نضغط على زر (OK) فتظهر شاشة المخرجات، والتي تحوي على الجدولين التاليين:

ولنفسر الجداول التي حصلنا عليها:

الجدول الأول: ويحوي بعض المؤشرات الإحصائية عن كل مجموعة من مجموعات العينة

الجدول الثاني: ويبين مخرجات اختبار العينات المستقلة حيث نلاحظ أن الجدول مقسوم إلى قسمين الأول يشمل الجزء المتعلق باختبار ليفين (Levene's Test) لتساوي التباينات والقسم الثاني يشمل اختبار (ت) لمقارنة المتوسطات (t-test for Equality of Means):

مناقشة المعلومات التي حصلنا عليها من الجدول رقم (2):

القسم الأول من جدول (Independent Samples Test): ويهتم بتساوي التباينات وفي البداية علينا كتابة الفرضية المتعلقة بهذا الاختبار:

الفرضية الصفرية: يوجد تجانس في درجات الطلاب حسب الشعبة.

الفرضية البديلة: لا يوجد تجانس في درجات الطلاب حسب الشعبة.

ومن اختبار ليفين (Levene's Test) نجد قيمة المعنوية أصغر من 0.05 وبالتالي نرفض الفرضية الصفرية (الابتدائية) ونقبل البديلة أي لا يوجد تجانس في الدرجات وهذا يعني أن تباين درجات الطلاب ليس متطابق بالنسبة للمجموعتين، ولا داعي للقلق بهذا الشأن لأن برنامج (SPSS) يوفر لنا اختبار T دون تحقق شرط تساوي التباينات.

القسم الثاني من جدول (Independent Samples Test): ويهتم هذا القسم بكشف هل هناك فروق ذات دلالة إحصائية بين المجموعتين، ويلاحظ أن هذا الجزء مقسوم إلى سطرين فالسطر الأول يناقش قيمة T عندما يكون هناك تجانس في البيانات، أما السطر الثاني فهو يناقش حالة عدم تجانس في البيانات. من خلال هذا القسم في جدول (Independent Samples Test) يمكننا أن نتخذ القرار بقبول أو رفض الفرضية الرئيسية (التي وضعناها بداية الحل وهي:

الفرضية الصفرية: لا يوجد فروق ذات دلالة إحصائية بين متوسط درجات الطلبة في الشعبتين عند مستوى دلالة إحصائية 0.05.

الفرضية البديلة: يوجد فروق ذات دلالة إحصائية بين متوسط درجات الطلبة في الشعبتين عند مستوى دلالة إحصائية 0.05)، وذلك بعدة طرق:

<u>الطريقة الأولى</u>: الطريقة الأولى المستخدمة في الاختبار تعتمد على قيمة المعنوية (significance value) والمبينة في الجدول الثاني ((Sig. (2-tailed) فإذا كانت قيمة المعنوية أكبر من مستوى المعنوية يكون هناك دليل قوي لقبول

الفرضية الابتدائية وإذا كانت قيمة المعنوية أصغر من مستوى المعنوية يكون هناك دليل قوي لرفض الفرضية الابتدائية، وفي حالتنا هنا نجد:

$$Sig.\ (2-tailed)=0.662 > 0.05$$

وبالتالي لدينا إثبات قوي لقبول الفرضية الابتدائية.

الطريقة الثانية: الطريقة الثانية المستخدمة في الاختبار تعتمد على مجال الثقة حول الفرق بين المتوسطين (Confidence Interval of the Difference) والمبينة في الجدول الثاني فإذا كانت حدود مجال الثقة من إشارتين مختلفتين يكون هناك دليل قوي لقبول الفرضية الابتدائية وإذا كانت حدود مجال الثقة من نفس الإشارة يكون هناك دليل قوي لرفض الفرضية الابتدائية، وفي حالتنا هنا نجد أن الحد الأدنى والحد الأعلى من إشارتين مختلفتين، وبالتالي لدينا إثبات قوي لقبول الفرضية الابتدائية.

كما يحتوي الجدول رقم (2) على حدود مجال الثقة حول الفرق بين المتوسطين (95% Confidence Interval of the Difference) حيث نجد أن:

[-10.5587 , 6.95877] ويقرأ مجال الثقة: أننا واثقون بمقدار 95% على أن الفرق بين المتوسطين لن يزيد عن 6.95877 ولن يقل -10.5587

كتابة النتائج:

استخدم اختبار ستيودنت t لمقارنة متوسطي عينتين مستقلتين، وذلك لمقارنة متوسطي درجات الطلاب في شعبتي المحاسبة وإدارة الأعمال، حيث دل الاختبار على عدم وجود فروق معنوية بين درجات طلاب شعبة المحاسبة و درجات طلاب شعبة إدارة الأعمال. حيث كانت درجات الطلبة في شعبة المحاسبة بمتوسط (77) وانحراف معياري (4.66) وخطأ معياري (1.475)، أما بالنسبة لدرجات الطلبة في شعبة إدارة الأعمال بمتوسط (78.8) وانحراف معياري (11.79) وخطأ معياري (3.72)، ومن خلال اختبار ستيودنت t بلغت قيمة t المحسوبة (-0.449) وهي غير دالة إحصائياً عند مستوى المعنوية (0.05).

8. مقارنة وسطي عينتين مرتبطتين:

كثيراً من الدراسات الاجتماعية تكون بياناتها على شكل مشاهدات مزدوجة متقابلة، فمثلاً عند مقارنة وسطي عدد سنوات الدراسة للإناث حسب رأي الآباء والأمهات فإننا نحصل من كل أسرة على إجابات مزدوجة (عدد سنوات الدراسة حسب رأي الأب، وعدد سنوات الدراسة حسب رأي الأم)، وكمثال آخر فإنه لمعرفة فيما إذا كان لبرامج التوعية الأسرية أثر في تخفيض عدد المواليد المرغوب إنجابهم لعينة من الفتيات المقبلات على الزواج فستكون بيانات الدراسة على شكل زوج متقابل (عدد الأولاد المرغوب إنجابهم قبل اتباع الدورة، وعدد الأولاد المرغوب إنجابهم بعد اتباع الدورة)... في مثل هذه الحالات نكون أمام عينتين زوجتين (عينتين غير مستقلتين) أي أن أحدهما مرتبطة بالأخرى (صورة أزواج من القيم) كما لو أن هناك عينة واحدة أخذ لها القياس مرتين مرة قبل التجربة ومرة بعدها.

والاختبار الإحصائي المستخدم لاختبار مثل هذه الفرضية هو اختبار T للبيانات المرتبطة، وبعدد درجات حرية ($df = n - 1$) ويعطى بالعلاقة التالية:

$$t = \frac{\bar{D}}{S_D / \sqrt{n}} \qquad (8-10)$$

ولإيجاد قيمة t نتبع الخطوات التالية:

أ- نوجد الفروق ما بين القياسات $D_i = X_i - Y_i$.

ب- نحسب المتوسط الحسابي للفرق $\bar{D}$ حيث $\bar{D} = \frac{\sum_{i=1}^{i=n} D_i}{n}$

ج- نحسب الانحراف المعياري لهذه الفروق ونرمز له بـ S_D، ويعطى بالصيغة التالية:

$$S_D = \sqrt{\frac{n\sum_{i=1}^{i=n} D_i^2 - (\sum_{i=1}^{i=n} D_i)^2}{n(n-1)}} \qquad (9-10)$$

د- نعوض في الصيغة (10-8).

تطبيق مقارنة متوسطي عينتين مرتبطتين باستخدام البرنامج الإحصائي SPSS:

قبل البدء بدراسة الأساليب المتاحة، هناك بعض الأمور العامة المشتركة التي يجب التحقق منها قبل البدء بتنفيذ الاختبار، وهذه الافتراضات هي:

1) مستوى القياس: يجب أن يكون المتغير الذي نجري عليه الاختبار كمياً (مستمراً – منفصلاً) ويكون مقياسه (فترة – نسبة).
2) المعاينة العشوائية: يجب اختيار العينة بشكل عشوائي ودون تحيز.
3) استقلالية المشاهدات: لا بد أن تكون كل المشاهدات التي تتكون منها البيانات مستقلة عن بعضها بعضا أي أنه يجب ألا تكون القياسات أو عملية قياس المشاهدات متأثرة بأية مشاهدة أو عملية قياس مشاهدة أخرى، وهذا الافتراض مهم

جداً وتتم عملية التشكيك في موقف عندما نجمع المشاهدات أو القياسات في موقف جماعي أو يكون فيه شكل ما من التفاعل بين أفراد التجربة.

4) التوزيع الطبيعي: يجب أن تكون المشاهدات أو القياسات تتبع التوزيع الطبيعي، فإذا لم تكن تتبع التوزيع الطبيعي يجب أن يكون حجم العينة أكبر من 30 مشاهدة، وسنقوم باستعراض آلية إجراء اختبار التوزيع الطبيعي ضمن كل حالة.

مثال 5:

قام أحد الخبراء التربويين بقياس درجة الذكاء لعدد من الطلبة فحصل على الدرجات الموضحة في الجدول التالي ضمن سطر قبل (before)، ثم قام بإخضاعهم لدورة تنمية المهارات التفكير العليا فحصل على البيانات ضمن سطر بعد (after)، فكانت الدرجات موضحة في الجدول التالي:

before	134	103	116	113	124	120	128	122	123	108	134	108	111	125	134
after	134	106	110	115	122	126	130	118	125	110	138	111	115	125	130

والمطلوب: بيان فيما إذا كانت الدورة قد حسنت مهارات التفكير عند الطلبة.

الحل:

سيتم اختبار الفرضيات التالية:

الفرضية الصفرية: لا يوجد فروق ذات دلالة معنوية (إحصائية) بين متوسطي درجات الطلاب قبل بدء الدورة وبعد الدورة عند مستوى دلالة إحصائية 0.05.

الفرضية البديلة: يوجد فروق ذات دلالة معنوية (إحصائية) بين متوسطي درجات الطلاب قبل بدء الدورة وبعد الدورة عند مستوى دلالة إحصائية 0.05.

خطوات تنفيذ الاختبار في برنامج (SPSS):

1- ندخل متغير الدرجات قبل البدء بالدورة ونعرفه في شاشة (Variable View ضمن Data Editor)، ثم ندخل البيانات في شاشة (Data View)، ثم ندخل متغيراً

ثانٍ هو متغير الدرجات بعد البدء بالدورة ونعرفه في شاشة (Variable View ضمن Data Editor)، ثم ندخل البيانات في شاشة (Data View).

2- نجري اختبار التوزيع الطبيعي:

نضع أولاً فرضيات التوزيع الطبيعي:

الفرضية الأولى: الفرضية الصفرية: درجات ذكاء الطلاب قبل الدورة تتبع التوزيع الطبيعي.

الفرضية البديلة: درجات ذكاء الطلاب قبل الدورة لا تتبع التوزيع الطبيعي.

الفرضية الثانية: الفرضية الصفرية: درجات ذكاء الطلاب بعد الدورة تتبع التوزيع الطبيعي.

الفرضية البديلة: درجات ذكاء الطلاب بعد الدورة لا تتبع التوزيع الطبيعي.

فتظهر شاشة المخرجات نهتم فقط بجدول التوزيع الطبيعي والمعنون بـ (Tests of Normality) والذي يكون له الشكل التالي:

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
before	.114	15	.200*	.943	15	.418
after	.131	15	.200*	.958	15	.662

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

من خلال الجدول السابق نجد أن الجدول يحوي على سطرين من خلال كل سطر يمكننا أن نناقش فرضية على حدة كما يلي:

الفرضية الأولى: من السطر الأول الذي يحوي على السطر المعنون قبل (before) نجد قيمة المعنوية أكبر من قيمة الدلالة الإحصائية 0.05 وبالتالي يمكننا قبول الفرضية الصفرية للتوزيع الطبيعي للفرضية الأولى أي: درجات ذكاء الطلاب قبل الدورة تتبع التوزيع الطبيعي.

الفرضية الثانية: من السطر الثاني الذي يحوي على السطر المعنون بعد (after) نجد قيمة المعنوية أكبر من قيمة الدلالة الإحصائية 0.05 وبالتالي يمكننا قبول الفرضية الصفرية للتوزيع الطبيعي للفرضية الثانية أي: درجات ذكاء الطلاب قبل الدورة تتبع التوزيع الطبيعي.

3- من أجل اختبار مقارنة المتوسطين (الفرضية الرئيسية) من شريط الأوامر :

4- نحصل على صندوق حوار يحوي مستطيلين يحويان كما يلي:

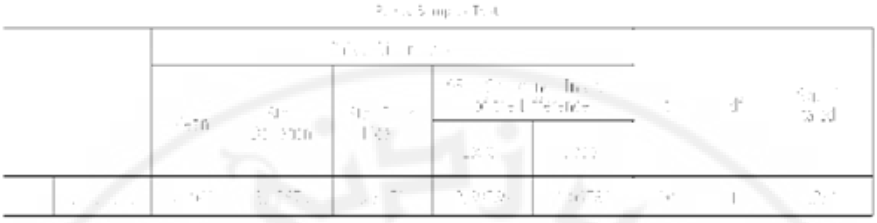
المستطيل الأول: يحوي جميع المتغيرات التي تحويها قاعدة البيانات.

المستطيل الثاني والمعنون بـ ((Paired Variable(s)): ونضع هنا المتغيرين المرتبطين اللذين نريد مقارنة متوسطيهما وهما القبل والبعد.

ثم نضغط على (OK)، فتظهر شاشة المخرجات، والتي تحوي على الجداول التالية:

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	[illegible]	15	.94_	.000



ولنفسر الجداول التي حصلنا عليها في شاشة المخرجات:

- ✓ **الجدول رقم (1):**

ويحوي بعض المؤشرات الإحصائية للمتغيرين قبل وبعد، كما في الجدول التالي:

- ✓ **الجدول رقم (2):**

ويحوي معلومات عن معامل الارتباط بين المتغيرين ومعنويته على اعتبار أننا نقارن بين عينتين مرتبطتين وهذا الجدول لا داعي أن نناقشه لأننا منطقياً نعلم أن العينتين مرتبطتان لأننا أجرينا الدراسة على فرض أن العينتين مرتبطتان وكانت البيانات لنفس الأشخاص قبل وبعد إجراء الدورة.

- ✓ **الجدول رقم (3):**

ويحوي معلومات دالة الاختبار قيمها ومعنويتها، حيث يحوي:

من خلال هذا الجدول يمكننا أن نتخذ القرار بقبول أو رفض الفرضية الصفرية، وذلك بعدة طرق:

الطريقة الأولى: الطريقة الأولى المستخدمة في الاختبار تعتمد على قيمة المعنوية (significance value) والمبينة في الجدول الثالث ((Sig. (2-tailed) فإذا كانت قيمة المعنوية أكبر من مستوى المعنوية يكون هناك دليل قوي لقبول الفرضية الابتدائية وإذا كانت قيمة المعنوية أصغر من مستوى المعنوية يكون هناك دليل قوي لرفض الفرضية الابتدائية، وفي حالتنا هنا نجد:

$$\text{Sig. } (2-\text{tailed})=0.381 > 0.05$$

وبالتالي لدينا إثبات قوي لقبول الفرضية الابتدائية.

الطريقة الثانية: الطريقة الثانية المستخدمة في الاختبار تعتمد على مجال الثقة حول الفرق بين المتوسطين (Confidence Interval of the Difference) والمبينة في الجدول الثالث فإذا كانت حدود مجال الثقة من إشارتين مختلفتين يكون هناك دليل قوي لقبول الفرضية الابتدائية وإذا كانت حدود مجال الثقة من نفس الإشارة يكون هناك دليل قوي لرفض الفرضية الابتدائية، وفي حالتنا هنا نجد أن الحد الأدنى والحد الأعلى من إشارتين مختلفتين، وبالتالي لدينا إثبات قوي لقبول الفرضية الابتدائية.

كما يحتوي الجدول رقم (2) على حدود مجال الثقة حول الفرق بين المتوسطين (95 % Confidence Interval of the Difference) حيث نجد أن:

[-2.6976 , 1.0976] ويقرأ مجال الثقة: أننا واثقون بمقدار 95% على أن الفرق بين المتوسطين لن يزيد عن 1.0976 ولن يقل-2.6976

كتابة النتائج:

استخدم اختبار ستيودنت t لمقارنة متوسطي عينتين مرتبطتين، وذلك لتقييم تأثير الدورة على الطلبة، حيث دل الاختبار على عدم وجود فروق معنوية بين متوسط درجات الطلبة قبل البدء بالدورة وبعد الدورة. حيث كانت درجات الطلبة قبل البدء بالدورة بمتوسط (120.2) وانحراف معياري (10.10) وخطأ معياري (2.60)، أما بالنسبة لدرجات الطلبة بعد الدورة بمتوسط (121) وانحراف معياري (9.73) وخطأ معياري (2.51)، ومن خلال اختبار ستيودنت t بلغت قيمة t المحسوبة (-0.904) وهي غير دالة إحصائياً عند مستوى المعنوية (0.05).

9. مقارنة أكثر من متوسطي عينتين:

وجدنا أنه لدراسة الفرق بين متوسطي مجتمعين (عينتين) يستخدم اختبار T-Test الذي يتسم ببساطة التحليل وسهولة التطبيق، غير أن أغلب الدراسات تهتم بدراسة أكثر من مجتمعين (عينتين) بهدف التعرف على الفروق بين متوسطات تلك المجتمعات (العينات). وباستخدام اختبار T-Test يتطلب أن تتم المقارنة بين متوسطات هذه المجتمعات (العينات) مثنى مثنى، وهذا ينطوي عليه كثير من الصعوبات للأسباب التالية:

1) الجهد المبذول في إجراء المقارنات الثنائية: نعلم أن عدد الاختبارات المطلوبة تزيد مع زيادة عدد المجتمعات المقارنة، فإذا كان عدد متوسطات المجتمعات المطلوب مقارنتها k فيكون عدد المقارنات الثنائية:

$$\frac{k.(k-1)}{2} \qquad (10-10)$$

2) إضعاف عملية المقارنة: إن إجراء الاختبار بين مجتمعين معاً وترك بقية المجتمعات الأخرى ولو مؤقتاً يعني ترك معلومات إضافية متاحة عن المجتمع وضياع فرص الحصول على أفضل تقدير لتباين المجتمع.

3) زيادة مقدار الخطأ: إن استخدام اختبار T Test لمقارنة متوسطي مجتمعين (عينتين) عند مستوى معنوية ($\alpha = .05$)، فإن احتمال اتخاذ قرار خاطئ هو ($.05$) وإذا كان لدينا 3 مجتمعات ونريد المقارنة عند مستوى معنوية ($\alpha = .05$) فإن احتمال اتخاذ قرار خاطئ هو $(.05)^3$، إذ كلما زادت المجتمعات (العينات) المدروسة زاد احتمال اتخاذ قرار خاطئ.

في ضوء ما تقدم من أسباب فإن الاختبار الإحصائي المناسب في هذه الحالات هو اختبار فيشر F-Test والذي نحصل عليه من تحليل التباين (Analysis of Variance) وبالرموز (ANOVA) والذي يمكننا من المقارنة بين أكثر من متوسطي مجتمعين (عينتين) بآن واحد، مما يقلل من احتمال الوقوع في الخطأ من النوع الأول، كما يقلل الجهد المبذول، كما أنه يعد أكثر قوة من اختبار مقارنة متوسطي عينتين، لأنه عند استخدام اختبار مقارنة متوسطي عينتين لا نستفيد من كل المعلومات المتوافرة بالمجتمعات، بالإضافة لميزة تحليل التباين والتي تتلخص بإمكانية اختبار أثر التفاعل المشترك بين المجتمعات المدروسة، لذلك يعد تحليل التباين من أهم الأساليب المستخدمة في البحوث في مختلف مجالات العلوم النفسية والتربوية والاجتماعية وغيرها.

فكلمة تباين تعني اختلاف الأشياء عن بعضها بعضاً، هذا الاختلاف هو الذي يجعلنا نميز بين هذه الأشياء، أي أن أي مجموعة من الأشياء مختلفة عن بعضها نسميها متباينة، والمعنى النفسي للتباين يتشابه مع معنى الفروق الفردية، أي اختلاف الأفراد عن بعضهم بعضاً، وأحياناً يكون الاختلاف داخل الأفراد، أي اختلاف مجموعة

من الظواهر الاجتماعية أو النفسية، في حين أن المعنى الإحصائي للتباين هو مربع الانحراف المعياري.

أما تحليل التباين فهو البحث عن مكونات هذا الاختلاف (أو التباين)، فالبحث عن المكونات وحساب كل مكون على حدة يسمى تحليلاً، أي دراسة مكونات الاختلاف بين مجموعة من الأفراد في ظاهرة معينة وحساب نصيب كل مكون بواسطة معادلات إحصائية معينة.

فمثلاً: إذا كان لدينا ثلاث مجموعات من الأفراد مصنفة حسب الحالة الاجتماعية (عازب، متزوج، مطلق) فالاختلاف بين العازبين والمتزوجين أو العازبين والمطلقين أو المتزوجين والمطلقين اختلاف (تباين) بين المجموعات، أما الاختلاف بين العازبين وبعضهم بعضا أو الاختلاف بين المتزوجين وبعضهم بعضا أو الاختلاف بين المطلقين وبعضهم بعضا يكون اختلافًا أو تباينًا داخل المجموعات، فمهمة تحليل التباين هي البحث عن مقدار الاختلاف بين المجموعات.

1.9. فرضيات تحليل التباين أحادي الاتجاه:

- الفرضية الابتدائية: لا يوجد فروق ذات دلالة إحصائية بين متوسطات المجتمعات وفقاً لمجموعات المتغير النوعي ($H_0 : \mu_1 = \mu_2 = ... = \mu_k$)
- الفرضية البديلة: يوجد فروق ذات دلالة إحصائية بين متوسطي مجتمعين على الأقل بين هذه المجتمعات المدروسة.

تطبيق تحليل التباين باستخدام البرنامج الإحصائي SPSS:

وجدنا مما سبق أن اختبار T Test يستخدم لمقارنة متوسط عينة مع متوسط معلوم أو من أجل مقارنة تساوي متوسطين .. ولكن السؤال المطروح الآن: ماذا لو أردنا مقارنة أكثر من متوسطين ..؟ نستخدم هنا تحليل التباين من أجل مقارنة أكثر من متوسطين.

يسمى تحليل التباين بتحليل التباين الأحادي لأنه يحوي على متغير مستقل واحد مع متغير تابع واحد، المتغير المستقل من نوع اسمي Nominal أو ترتيبي Ordinal ولكننا ندخله في البرنامج من النوع الرقمي، أما المتغير التابع فيكون من النوع الكمي Scale.

الشروط الواجب توافرها قبل إجراء تحليل التباين:

أ- يجب أن يكون توزيع المتغير التابع توزيعاً طبيعياً لكل مجموعة من مجموعات المتغير العاملي (يمكن تجاوز هذا الشرط إذا كان عدد أفراد العينة أكبر من 15 لكل مجموعة من المجموعات).

ب- يجب أن يكون التباين متساوٍ لكل مجموعة من المجموعات.

ج- يجب أن تكون العينات من كل مجتمع من مجتمعات المتغير العاملي عشوائية وأن تكون قيم المتغير التابع مستقلة عن بعضها لكل فرد من أفراد العينات.

مثال 6:

رغب مدير إحدى الشركات بمعرفة عدد الأيام المناسب لتدريب موظفي الشركة الجدد أفضل عدد لأيام تدريب الموظفين الجدد الشركة، فقام قسم التدريب بإجراء تدريب للموظفين الجدد في ثلاث مجموعات (المجموعة الأولى مدتها شهر؛ والمجموعة الثانية مدتها شهرين؛ والثالثة مدتها ثلاثة أشهر)، وبعد الانتهاء من التدريب أجري اختبار لقياس مستوى المعارف التي حصل عليها الموظفون وكانت النتائج كما في الجدول التالي:

Group1	8	60	32	69	57	75	52	86	68	63
Group2	75	82	81	81	47	81	71	80	88	72
Group3	81	76	81	74	75	74	85	81	89	82

والمطلوب: مساعدة مدير الشركة باختيار أفضل مدة للتدريب.

لتنفيذ المطلوب نتبع الخطوات التالية:

- ندخل متحولين الأول عدد النقاط (Points) ويحوي النقاط المسجلة.
- والثاني المجموعات (Groups) ويحوي المجموعات الثلاثة.

	Points	Groups
1	63	Group 1
2	68	Group 1
3	86	Group 1
4	52	Group 1
5	75	Group 1
6	57	Group 1
7	60	Group 1
8	32	Group 1
9	60	Group 1
10	58	Group 1
11	72	Group 2
12	88	Group 2
13	80	Group 2
14	71	Group 2
15	81	Group 2
16	47	Group 2
17	81	Group 2
18	81	Group 2
19	82	Group 2
20	75	Group 2
21	82	Group 3
22	80	Group 3
23	81	Group 3
24	85	Group 3
25	71	Group 3
26	75	Group 3
27	74	Group 3
28	61	Group 3
29	76	Group 3
30	61	Group 3

أولاً: نجري اختبار التوزيع الطبيعي:

- ✓ الفرضية الابتدائية: متغير المجموعة 1 يتبع التوزيع الطبيعي.
- ✓ الفرضية البديلة: متغير المجموعة 1 لا يتبع التوزيع الطبيعي.
- ✓ الفرضية الابتدائية: متغير المجموعة 2 يتبع التوزيع الطبيعي.
- ✓ الفرضية البديلة: متغير المجموعة 2 لا يتبع التوزيع الطبيعي.
- ✓ الفرضية الابتدائية: متغير المجموعة 3 يتبع التوزيع الطبيعي.
- ✓ الفرضية البديلة: متغير المجموعة 3 لا يتبع التوزيع الطبيعي.

وكما مر معنا سابقاً لاختبار التوزيع الطبيعي نجد أن البيانات بالنسبة للمجموعات الثلاث تتوزع وفق التوزيع الطبيعي.

ثانياً: نجري اختبار التباين:

- ✓ الفرضية الابتدائية: يوجد تجانس في البيانات.
- ✓ الفرضية البديلة: لا يوجد تجانس في البيانات.

وكما مر معنا سابقاً لاختبار تجانس البيانات، كما في الشكل التالي:

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
Points	Based on Median	1.500	2	27	.241
	Based on Median and with adjusted df	1.500	2	19.866	.247
	Based on trimmed mean	1.690	2	27	.204

من الجدول المعنون بـ Test of Homogeneity of Variance ومن السطر الأول نجد أن قيمة المعنوية 0.191 = Sig وهي أكبر من 0.05 وبالتالي نقبل الفرضية الابتدائية وبالتالي يوجد تجانس بين البيانات.

ثالثاً: نجري اختبار تحليل التباين:

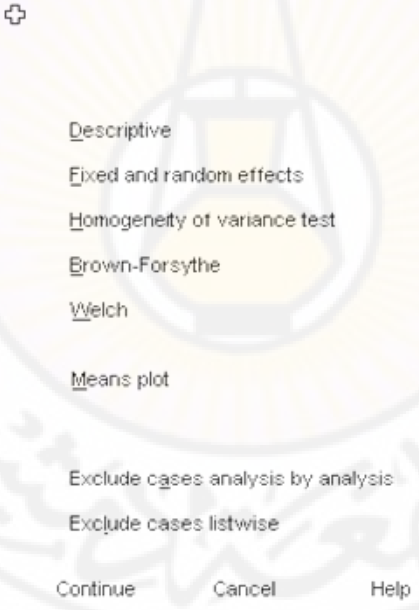
- ✓ الفرضية الابتدائية: لا يوجد فروق معنوية (ذات دلالة إحصائية) بين متوسطات المجموعات الثلاث.
- ✓ الفرضية البديلة: يوجد فروق معنوية (ذات دلالة إحصائية) بين متوسطين على الأقل.

1- من قائمة Analyze نختار Compare Means ثم One-Way ANOVA..:

- ننقل المتغير Points المستمر إلى القسم Dependent List .
- ننقل متغير التقسيم Groups إلى القسم Factor.

2- ثم ننقر على الأيقونة ..Options فيظهر صندوق الحوار التالي:

- Descriptive : وذلك لحساب الإحصاءات الوصفية (N ;Mean ;Std. Deviation ;Std. Error)
- Homogeneity Of variance test: لاختبار تجانس البيانات.
- Brown-Forsythe\ Welch: اختباران يستخدمان كبديلين عن اختبار F في حالة عدم تساوي التباينات .
- Means plot : لرسم بياني يمثل الفروقات بين المتوسطات حسب فئات المتغير Factor.



- وبالنقر على Continue ثم OK، تظهر شاشة المخرجات كما يلي:

الجدول المعنون بـ Descriptive يبين لنا الإحصاء الوصفي لكل فئة من الفئات.

والجدول التالي المعنون بـ ANOVA

ونلاحظ أن قيمة Sig=.003 أصغر من 0.05 وبالتالي نرفض الفرضية الابتدائية ونقبل البديلة وبالتالي يوجد هناك اختلاف بين متوسطين على الأقل.

ولنبحث عن هذا الاختلاف بين أي متوسطين يكمن، لذلك نتبع ما يلي:

من قائمة Analyze نختار الأمر Compare Means ثم نختار الأمر الفرعي One-Way ANOVA.. فيظهر صندوق الحوار السابق نختار منه الأيقونة Post Hoc فيظهر صندوق الحوار التالي:

يُظهر لنا هذا صندوق الحوار خياران: الأول عندما تكون التباينات متساوية، و الثاني عندما تكون التباينات غير متساوية .

نختار من هذا الصندوق اختبار Tukey لأن التباينات متساوية وبالنقر على Continue ثم OK يظهر ما يلي:

وهذا الجدول يظهر لنا ما يلي:

العمود الأول المعنون ((I–J) Mean Difference) الفرق بين متوسط المجموعة j–i ، والعمود الثاني المعنون (Std. Error) يبين الخطأ ،وقيمة Sig ، ومجال الثقة (95% Confidence Interval).

ولمناقشة الفروق بين المجموعات نتبع أحد الخيارات التالية:

أ- إذا ظهرت إشارة (*) فوق الفرق بين المتوسطات يدل على وجود فروق في المتوسطات.

ب- حسب قيمة المعنوية (Sig) إذا كانت أصغر من 0.05 فهذا يدل على وجود فروق معنوية.

ج- إذا كان مجال الثقة من إشارة واحدة فهذا يدل على وجود فروق معنوية.

فنلاحظ من الجدول المعنون (Multiple Comparisons):

- يوجد فروق معنوية بين متوسط المجموعة الأولى ومتوسط المجموعة الثانية.
- يوجد فروق معنوية بين متوسط المجموعة الأولى ومتوسط المجموعة الثالثة.
- عدم وجود فروق معنوية بين متوسط المجموعة الثانية والثالثة.

كما نجد التالي، والذي يبين لنا:

Points

Tukey HSD[a]

Groups	N	Subset for alpha = 0.05	
		1	2
Group 1	10	62.00	
Group 2	10		75.80
Group 3	10		79.80
Sig.		1.000	.697

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 10.000.

N : عدد مفردات كل عينة من العينات .

ويحسب لنا متوسط كل عينة من العينات ويضع المتوسطات الذين لا يوجد بينهم فروق معنوية عند مستوى الأهمية Subset for alpha=.05 (نحن نحدده) ضمن عامود واحد.

أسئلة وتمارين الفصل العاشر

1. يرغب مدير مؤسسة نقل سياحي أن يتأكد أن متوسط عدد الركاب في الحافلة التي تتسع لـ 45 راكب يساوي 41 راكب، وللتأكد من ذلك تم أخذ عينة مؤلفة من 10 حافلات وكانت النتيجة على النحو التالي:

رقم الحافلة	1	2	3	4	5	6	7	8	9	10
عدد الركاب	40	41	39	38	42	44	41	41	39	40

والمطلوب:

1) اختبر إذا كانت البيانات تتبع التوزيع الطبيعي عند مستوى دلالة إحصائية $\alpha=0.05$.

2) هل متوسط عدد الركاب في الحافلة يساوي 41 راكب عند مستوى دلالة إحصائية $\alpha=0.05$.

2. يرغب مدير أحد الفنادق باختبار ملائمة مذاق الطعام الذي يقدمه المطعم لزبائنه بناءً على آراء 15 زبوناً تم اختيارهم عشوائياً، وذلك بإعطائه درجة على مقياس يتراوح من 1 (غير ملائم) إلى 20 (ممتاز) وقد كانت نتيجة آراء الزبائن على النحو التالي:

7	18	15	18	13	17	15	16
17	14	19	10	6	14	12	

والمطلوب:

1) اختبر إذا كانت البيانات تتبع التوزيع الطبيعي عند مستوى دلالة إحصائية $\alpha=0.05$.

2) حتى يوجد فروق ذات دلالة إحصائية بين متوسط آراء الزبائن والقيمة 15 عند مستوى دلالة إحصائية $\alpha=0.05$.

3. يرغب باحث بدراسة اتجاهات مجموعة من السائحين نحو فاعلية برنامج سياحي جديد، ومن أجل ذلك قام باحث باختيار 10 أشخاص وطلب منهم تقييم الرحلة السياحية قبل وبعد تطبيق البرنامج السياحي الجديد بناءً على مقياس من 1 إلى 10 وكانت النتيجة على النحو التالي:

	قبل	بعد
1	5	9
2	5	8
3	2	10
4	7	7
5	8	8
6	8	7
7	3	9
8	5	10
9	4	10
10	5	8

والمطلوب:

1) اختبر إذا كانت البيانات تتبع التوزيع الطبيعي عند مستوى دلالة إحصائية $\alpha=0.05$.
2) اختبر إذا كان هناك فروق في اتجاهات الأشخاص نحو الرحلة السياحية قبل وبعد البرنامج السياحي، عند مستوى دلالة إحصائية $\alpha=0.05$.

4. شارك 11 شخصاً في برنامج لتخفيف الوزن، والجدول التالي يبين أوزانهم قبل البدء بالبرنامج وبعده.

	قبل	بعد
1	101	100
2	121	115
3	131	120
4	120	116
5	100	99
6	105	102
7	115	110
8	110	105
9	102	100
10	112	107
11	122	120

والمطلوب:

1) اختبر إذا كانت البيانات تتبع التوزيع الطبيعي عند مستوى دلالة إحصائية $\alpha=0.01$.

2) هل ساهم البرنامج في تخفيف الوزن، عند مستوى دلالة إحصائية $\alpha=0.01$.

5. الجدول التالي يبين درجات الذكور والإناث في أحد المقررات:

ذكور	65	66	71	75	80	77	69	60	83
إناث	63	72	74	76	83	84	86	69	70

والمطلوب: هل يوجد فروق ذات دلالة إحصائية بين درجات الذكور والإناث عند مستوى دلالة إحصائية $\alpha=0.01$.

6. الجدول التالي يبين نتائج مجموعة من الطلاب في ثلاث طرق تدريس:

طريقة 1	65	66	63	72	74	76	69	60	69
طريقة 2	83	84	86	81	70	71	75	80	77
طريقة 3	81	91	90	93	86	80	79	77	85

والمطلوب: أي طريقة من الطرق السابقة أفضل في التدريس، عند مستوى دلالة إحصائية $\alpha=0.01$.

الفصل الحادي عشر CHAPTER 11

الأساليب الإحصائية لدراسة العلاقات

أهداف الفصل الحادي عشر:

1. التعرف على معامل الارتباط وكيفية التطبيق باستخدام SPSS.
2. التعرف على معامل الارتباط الجزئي وكيفية التطبيق باستخدام SPSS.
3. التعرف على معامل الانحدار وكيفية التطبيق باستخدام SPSS.

Damascus University

تهتم البحوث بدراسة العلاقة بين متغيرين أو أكثر، فينصب اهتمام الباحث لمعرفة هل يوجد علاقة بين المتغيرين أو لا؟ وما هو شكل العلاقة بين المتغيرين إن وجدت؟، وما هي شدة العلاقة بين المتغيرين أو المتغيرات؟

ولا بد من التمييز بوضوح بين وجود ارتباط مرتفع بين ظاهرتين وبين وجود علاقة سببية بينهما، فوجود ارتباط مرتفع لا يعني بالضرورة أن إحدى الظاهرتين هي سبب للأخرى. إذ قد يكون الارتباط المرتفع بينهما نتيجة لتأثر كل منهما بظاهرة ثالثة لم تدخل في الحساب. فمثلاً من المعروف أن هناك ارتباطاً مرتفعاً بين التدخين وسرطان الرئة، وهناك ارتباط بين التدخين وتلون الأسنان، ولو حصل أن أخذنا بيانات إحصائية تتضمن درجة تلون الأسنان ونسبة الإصابة بسرطان الرئة، وكانت هذه البيانات في غالبيتها من أفراد تلونت أسنانهم بفعل التدخين فنجد ارتباطاً مرتفعاً بين ظاهرة تلون الأسنان وظاهرة سرطان الرئة، وهذا لا يعني بالطبع أن تلون الأسنان يؤدي إلى الإصابة بسرطان الرئة أو العكس، فالارتباط المرتفع كان نتيجة لوجود عامل ثالث خفي يرتبط معه المتغيران معاً هو عادة التدخين.

وتتم دراسة العلاقات وفقاً للأسلوب الذي يتم بموجبه تحليل البيانات ونوع البيانات والمتغيرات، ونميز ثلاثة أنواع لدراسة العلاقات هي:

- **الارتباط** (Correlation):

وهو أسلوب إحصائي يستخدم لتحليل العلاقة من حيث النوع والاتجاه بين المتغيرات أو الظواهر التي يمكن قياس المشاهدات المأخوذة منها والتعبير عنها بشكل كمي مثل علاقة مستوى الذكاء بالتحصيل الدراسي، أو علاقة الدخل بالإنفاق الشهري على السلع الغذائية والاستهلاكية.

- **الاقتران** (Association):

وهو أسلوب إحصائي يستخدم لتحديد نوع العلاقة بين المتغيرات أو الظواهر التي لا يمكن قياس المشاهدات المأخوذة منها والتعبير عنها بصورة رقمية (كمية) وإنما بصورة

نوعية (وصفية)، مثل علاقة المهنة (موظف، عامل، طالب) بالاتجاه نحو التغيير الاجتماعي (اتجاه إيجابي، اتجاه محايد، اتجاه سلبي).

- **التوافق** (Contingency):

وهو أسلوب إحصائي يستخدم لتحديد نوع العلاقة بين المتغيرات أو الظواهر التي يمكن قياس المشاهدات المأخوذة منها والتعبير عنها بصورة رقمية (كمية) والمتغيرات أو الظواهر الأخرى بصورة نوعية (وصفية) مثل العلاقة بين جنس المدرس الجامعي (ذكر، أنثى) والدخل الشهري.

1. الارتباط (Correlation):

يقاس الارتباط بواسطة معامل الارتباط، وقيمة معامل الارتباط لا تقل عن -1 ولا تزيد عن $+1$، وإذا كان مساوياً للصفر فهذا يعني أنه لا توجد أية علاقة تربط بين قيم الظاهرتين، وأن أي قيمة للمتغير المستقل X لا تقابلها أي قيمة للمتغير التابع Y، وفي هذه الحالة فإن شكل البيانات في مخطط الانتشار تظهر متبعثرة على شكل سحابة من البيانات أو مساحة دائرية مليئة بالبيانات.

أما إذا كان معامل الارتباط $+1$ أو -1 فهذا يعني أن الارتباط بين الظاهرتين شديد جداً أو مثالي والنقط تظهر على الرسم متبعة لخط مستقيم وإن أي قيمة للمتغير المستقل X تقابلها قيمة للمتغير التابع Y دون زيادة أو نقصان.

والحالة الأكثر شيوعاً هي عندما تكون قيمة معامل الارتباط لا تساوي $+1$ أو -1 ففي هذه الحالة تزداد شدة العلاقة الارتباطية بين الظاهرتين كلما اقترب المعامل من $+1$ أو -1 وتنقص شدة العلاقة كلما اقترب المعامل من الصفر، وبصفة عامة فإن قيمة معامل الارتباط أو مستوياته يمكن أن تتحدد كما هو موضح في الجدول التالي:

جدول رقم (11-1) وصف معاملات الارتباط حسب قيمتها

قوة معامل الارتباط	مدى معامل الارتباط
تام	+1 أو -1
عالٍ جداً	من 0.8 إلى 0.99
عالٍ	من 0.6 إلى 0.79
متوسط	من 0.4 إلى 0.59
ضعيف	من 0.2 إلى 0.39
ضعيف جداً	من 0.01 إلى 0.19
لا يوجد علاقة خطية	صفر

1.1. معامل ارتباط بيرسون (Person Correlation Cofficents):

يستخدم معامل ارتباط بيرسون لدراسة العلاقة بين متغيرين كميين. ولحساب هذا المعامل لا بد من توافر قيمة لكل متغير عند نفس الفرد، بمعنى أن هذه القيم عبارة عن قيم مزدوجة، وهذا يعني أيضاً تساوي عدد هذه القيم في المتغيرين ويحسب هذا المعامل باستخدام العلاقة التالية:

$$R = \frac{1}{n-1}\sum_{i=1}^{n}\frac{(x_i-\bar{x})}{S_x}\bullet\frac{(y_i-\bar{y})}{S_y} \qquad (1-11)$$

كما يمكن استنتاج الصيغة التالية من الصيغة السابقة:

$$R = \frac{\sum_{i=1}^{n}(x_i-\bar{x})\bullet(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\bullet\sum_{i=1}^{n}(y_i-\bar{y})^2}} \qquad (2-11)$$

وبإصلاح الصيغة السابقة يمكن كتابتها بأشكال أخرى أكثر سهولة لحساب معامل الارتباط:

$$R = \frac{n(\sum_{i=1}^{n} x_i \cdot y_i) - (\sum_{i=1}^{n} x_i) \cdot (\sum_{i=1}^{n} y_i)}{\sqrt{\left[n\sum_{i=1}^{n} x_i^{\,2} - (\sum_{i=1}^{n} x_i)^2\right]\left[n\sum_{i=1}^{n} y_i^{\,2} - (\sum_{i=1}^{n} y_i)^2\right]}} \qquad (3\text{-}11)$$

أو :

$$R = \frac{\sum_{i=1}^{n} x_i \cdot y_i - n \cdot \overline{X} \cdot \overline{Y}}{\sqrt{\left[n\sum_{i=1}^{n} x_i^{\,2} - n \cdot \overline{X}^{\,2}\right]\left[n\sum_{i=1}^{n} y_i^{\,2} - n \cdot \overline{Y}^{\,2}\right]}} \qquad (4\text{-}11)$$

وللتأكد من الأهمية الإحصائية لمعامل الارتباط نحسب الخطأ المعياري لمعامل الارتباط S_R . فإذا كان $R > 3 \cdot S_R$ يكون لمعامل الارتباط أهمية إحصائية أو مدلول إحصائي، حيث تحسب قيمة الخطأ المعياري لمعامل الارتباط (S_R) وفق الصيغة التالية:

$$S_R = \sqrt{\frac{1-R^2}{n-2}} \qquad (5\text{-}11)$$

وبالاستناد إلى هذه الصيغة يتم معرفة فيما إذا كان معامل الارتباط يمثل المجتمع الذي سحبت منه العينة، فإذا كان لمعامل الارتباط أهمية إحصائية أو مدلول إحصائي، فإن العينة تمثل مجتمعها الإحصائي، أما إذا كان معامل الارتباط ليس له أهمية إحصائية أو مدلول إحصائي، فإن العينة لا تمثل مجتمعها الإحصائي، أو أن يكون الارتباط زائفاً، أي لا وجود لعلاقة سببية بين المتغيرين.

مثال 1:

بفرض لدينا البيانات التالية التي تعبر عن درجات عشرة طلبة في مادتي اللغة الأجنبية (Lang) والرياضيات (Math)، كما هي موضحة في الشكل المجاور:

	Lang	Math
1	60.00	56.00
2	68.00	60.00
3	60.00	64.00
4	74.00	82.00
5	80.00	76.00
6	84.00	72.00
7	80.00	74.00
8	72.00	60.00
9	62.00	64.00
10	82.00	86.00

والمطلوب:

1) اختبر معنوية معامل الارتباط بمستوى دلالة 5%.
2) 2. أوجد قيمة معامل الارتباط.

لتنفيذ المطلوب نتبع الخطوات التالية:

سيتم اختبار الفرضيات التالية:

الفرضية الصفرية: لا يوجد علاقة ذات دلالة معنوية (إحصائية) بين درجات الطلبة في مقرر اللغة الأجنبية ومقرر الرياضيات.

الفرضية البديلة: يوجد علاقة ذات دلالة معنوية (إحصائية) بين درجات الطلبة في مقرر اللغة الأجنبية ومقرر الرياضيات.

ولكن قبل البدء في اختبار الفرضية علينا معرفة إذا ما كان هناك ارتباط عن اختبار النقاط باستخدام شكل الانتشار، نتبع الخطوات التالية:

ودائماً ننقل المتغير المستقل إلى Axis (X) وننقل المتغير التابع إلى Axis (Y) ثم نضغط على OK فيظهر شكل الانتشار في شاشة المخرجات:

خطوات تنفيذ الاختبار في برنامج (SPSS):

1- ندخل متغيري Math, Lang ونعرفهما في شاشة (Variable View ضمن Data Editor)، ثم ندخل البيانات في شاشة (Data View ضمن Data Editor).

اسم المتغير	Name	Type	Values	Measurement
الدرجة في اللغة	Lang	Numeric	×	Scale
الدرجة في الرياضيات	Math	Numeric	×	Scale

2- نجري اختبار التوزيع الطبيعي، وذلك لمعرفة هل البيانات تتبع التوزيع الطبيعي أم لا، والهدف من هذا الاختبار معرفة معامل الارتباط الذي نريد استخدامه:

نضع أولاً فرضيات التوزيع الطبيعي:

الفرضية الأولى: الفرضية الصفرية: درجات الطلبة في مقرر اللغة الأجنبية تتبع التوزيع الطبيعي.

الفرضية البديلة: درجات الطلبة في مقرر اللغة الأجنبية لا تتبع التوزيع الطبيعي.

الفرضية الثانية: الفرضية الصفرية: درجات الطلبة في مقرر الرياضيات تتبع التوزيع الطبيعي.

الفرضية البديلة: درجات الطلبة في مقرر الرياضيات لا تتبع التوزيع الطبيعي.

فتظهر شاشة المخرجات نهتم فقط بجدول التوزيع الطبيعي والمعنون بـ (Tests of Normality) ويكون كما يلي:

من خلال الجدول السابق نجد أن الجدول يحوي على سطرين من خلال كل سطر يمكننا أن نناقش كل فرضية على حدة كما يلي:

- **الفرضية الأولى:** من السطر الأول الذي يحوي على السطر المعنون اللغة الأجنبية (Lang) نجد قيمة المعنوية أكبر من 0.05 وبالتالي يمكننا قبول الفرضية الصفرية للتوزيع الطبيعي للفرضية الأولى أي: درجات الطلبة في مقرر اللغة الأجنبية تتوزع وفق التوزيع الطبيعي.
- **الفرضية الثانية:** من السطر الثاني الذي يحوي على السطر المعنون الرياضيات (Math) نجد قيمة المعنوية أكبر من 0.05 وبالتالي يمكننا قبول الفرضية الصفرية للتوزيع الطبيعي للفرضية الثانية أي: درجات الطلبة في مقرر الرياضيات تتوزع وفق التوزيع الطبيعي.

3- من أجل اختبار الفرضية الرئيسية وهي دراسة العلاقة بين المتغيرين من شريط الأوامر:

Analyze Correlate Bivariate

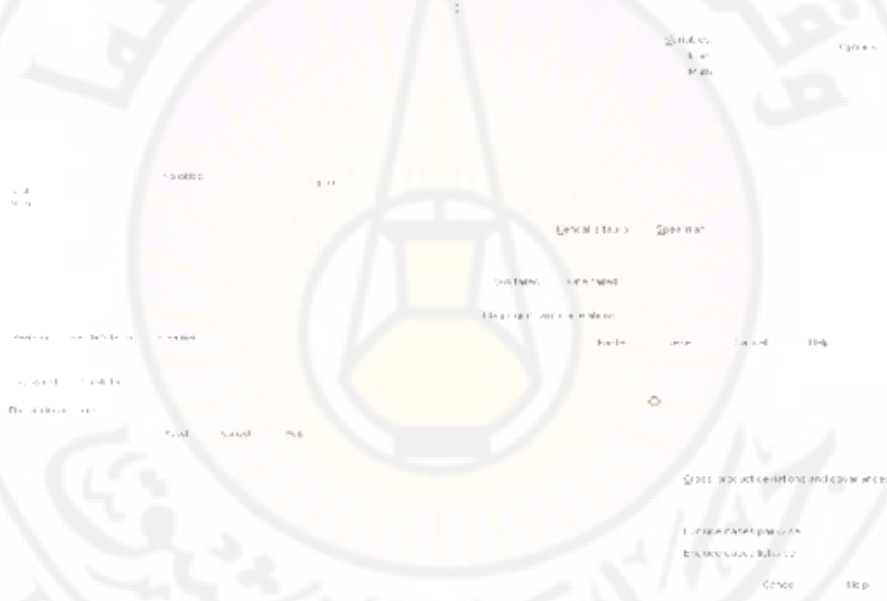
سيظهر صندوق الحوار كما في الشكل المجاور:

1) نختار المتغيرات الكمية التي نريد حساب معامل الارتباط بينها وننقلها إلى مربع Variables
2) من الحقل المعنون بـ Correlation Coefficient نختار معامل الارتباط الذي نريده حسب قواعد اختيار معاملات الارتباط وهنا نختار معامل ارتباط بيرسون (Person) لأن المتغيرين كميان ويتبعان التوزيع الطبيعي.
3) من الحقل المعنون بـ Test Significance: وهنا نختار حسب نوع الفرضية التي نريد اختبارها إذا كانت من طرف واحد نختار One Tailed أما في حالة اختبار الفرضية من الطرفين نختار Two Tailed وهو الخيار الافتراضي.
4) Flag Significant Correlation عند تأشير هذا الخيار فإنه سوف يظهر علامة (*) فوق معامل الارتباط عندما يكون معامل الارتباط معنوياً عند مستوى المعنوية 5%، وسوف يظهر (**) فوق معامل الارتباط عندما يكون معامل الارتباط معنوياً عند مستوى المعنوية 1%.

5) بالنقر على زر Option ستظهر شاشة حوار Bivariate Correlations: Options

ونختار منها Means and Standard Deviation وذلك لعرض المتوسطات الحسابية والانحرافات المعيارية لكل متغير من المتغيرات المراد حساب معامل الارتباط لها.

6) ثم ننقر على Continue ثم OK.



فتظهر شاشة المخرجات التي تحوي الجداول التالية:

الجدول الأول: والذي يبين لنا المتوسط الحسابي والانحراف المعياري وعدد الحالات لكل متغير من المتغيرات.

Descriptive Statistics

	Mean	Std. Deviation	N
lang	72.2000	9.30711	10
Math	70.0000	9.66092	10

Correlations

		lang	Math
lang	Pearson Correlation	1	.776**
	Sig. (2-tailed)		.008
	N	10	10
Math	Pearson Correlation	.776**	1
	Sig. (2-tailed)	.008	
	N	10	10

**. Correlation is significant at the 0.01 level (2-tailed).

من الجدول السابق أول شيء نناقشه وهي قيمة الـ Sig فنلاحظ أن قيمة الـ Sig أصغر من 0.01 وأصغر من 0.05 أي أنه يوجد علاقة ارتباطية ذات دلالة إحصائية بين متغيري Math, Lang أي أننا رفضنا الفرضية الصفرية وقبلنا الفرضية البديلة والتي تنص على وجود علاقة ارتباطية بين المتغيرين، وقيمة معامل الارتباط بيرسون بين المتغيرين Math , Lang هو $r = 0.776$.

2.1. معامل الارتباط سبيرمان للرتب (Spearman Correlation Cofficents):

يعتبر معامل الارتباط سبيرمان من الطرائق الإحصائية الهامة في قياس العلاقة بين متغيرين رتبيين، وخاصة عندما يكون حجم العينة أقل أو يساوي 30 ($n \leq 30$) أو يكون المتغيران من النوع الرتبي، ويحسب معامل ارتباط سبيرمان كما يلي:

بفرض أن عينة من قيم x تتضمن n قيمة، نضع قيم المتغير x في عمود وفي العمود المجاور نضع أمام كل قيمة لـ x رتبة يبدأ الترتيب من الأصغر إلى الأكبر، ولكن إذا تكررت إحدى قيم x أكثر من مرة فإننا نعطي رتبة لكل قيمة مساوية للمتوسط الحسابي للرتب القيم، وبالطريقة نفسها نرتب قيم المتغير y ، ليصبح لكل قيمة من المتغير x رتبة تقابلها رتبة لقيمة من المتغير y .

وللحصول على قيمة معامل ارتباط سبيرمان للرتب يتم تطبيق الصيغة التالية:

$$r_s = 1 - \frac{6 \cdot \sum_{i=1}^{n} d^2}{n(n^2 - 1)} \qquad (6-11)$$

مثال 2:

بفرض لدينا البيانات التالية التي تعبر عن درجات عشرة طلبة في مادتي اللغة الأجنبية (Lang) والرياضيات (Math)، كما هي موضحة في الشكل المجاور:

	lang	Math
1	50.00	56.00
2	58.00	60.00
3	50.00	64.00
4	64.00	82.00
5	60.00	76.00
6	94.00	72.00
7	90.00	74.00
8	62.00	66.00
9	62.00	64.00
10	95.00	86.00

والمطلوب:

1) اختبر معنوية معامل الارتباط بمستوى دلالة 5%.
2) ما قيمة معامل الارتباط.

لتنفيذ المطلوب نتبع الخطوات التالية:

سيتم اختبار الفرضيات التالية:

الفرضية الصفرية: لا يوجد علاقة ذات دلالة معنوية (إحصائية) بين درجات الطلبة في مقرر اللغة الأجنبية ومقرر الرياضيات.

الفرضية البديلة: يوجد علاقة ذات دلالة معنوية (إحصائية) بين درجات الطلبة في مقرر اللغة الأجنبية ومقرر الرياضيات.

خطوات تنفيذ الاختبار في برنامج (SPSS):

1- ندخل متغيري Lang , Math ونعرفهما في شاشة (Variable View ضمن Data Editor)، ثم ندخل البيانات في شاشة (Data View ضمن Data Editor).

Measurement	Values	Type	Name	اسم المتغير
Scale	×	Numeric	Lang	الدرجة في اللغة
Scale	×	Numeric	Math	الدرجة في الرياضيات

2- نجري اختيار التوزيع الطبيعي، وذلك لمعرفة هل البيانات تتبع التوزيع الطبيعي أم لا، والهدف من هذا الاختبار معرفة معامل الارتباط الذي نريد استخدامه:

نضع أولاً فرضيات التوزيع الطبيعي:

الفرضية الأولى: الفرضية الصفرية: درجات الطلبة في مقرر اللغة الأجنبية تتبع التوزيع الطبيعي.

الفرضية البديلة: درجات الطلبة في مقرر اللغة الأجنبية لا تتبع التوزيع الطبيعي.

الفرضية الثانية: الفرضية الصفرية: درجات الطلبة في مقرر الرياضيات تتبع التوزيع الطبيعي.

الفرضية البديلة: درجات الطلبة في مقرر الرياضيات لا تتبع التوزيع الطبيعي.

فتظهر شاشة المخرجات نهتم فقط بجدول التوزيع الطبيعي والمعنون بـ (Tests of Normality) والذي يكون له الشكل التالي:

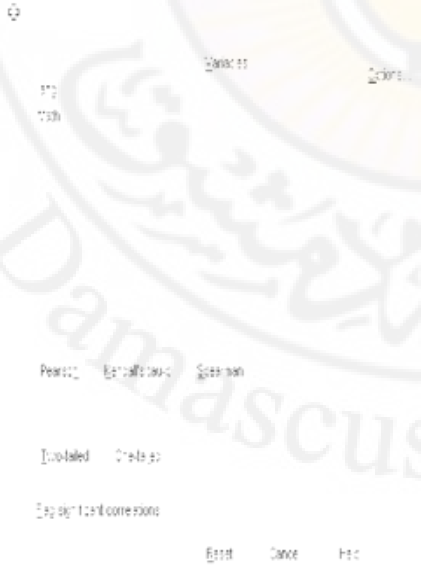
من خلال الجدول السابق نجد أن الجدول يحوي على سطرين من خلال كل سطر يمكننا أن نناقش كل فرضية على حدة كما يلي:

- **الفرضية الأولى:** من السطر الأول الذي يحوي على السطر المعنون اللغة الأجنبية (Lang) نجد قيمة المعنوية أصغر من 0.05 وبالتالي نرفض الفرضية الصفرية للتوزيع الطبيعي للفرضية الأولى أي: درجات الطلبة في مقرر اللغة الأجنبية لا تتوزع وفق التوزيع الطبيعي.
- **الفرضية الثانية:** من **السطر** الثاني الذي يحوي على السطر المعنون الرياضيات (Math) نجد قيمة المعنوية أكبر من 0.05 وبالتالي يمكننا قبول الفرضية الصفرية للتوزيع الطبيعي للفرضية الثانية أي: درجات الطلبة في مقرر الرياضيات تتوزع وفق التوزيع الطبيعي.

3- من أجل اختبار الفرضية الرئيسية وهي دراسة العلاقة بين المتغيرين من شريط الأوامر:

Analyze Correlate Bivariate

سيظهر صندوق الحوار كما في الشكل المجاور:



أ- نختار المتغيرات الكمية التي نريد حساب معامل الارتباط بينها وننقلها إلى مربع Variables

ب- من الحقل المعنون بـ Correlation Coefficient نختار معامل الارتباط الذي نريده حسب قواعد اختيار معاملات الارتباط وهو معامل ارتباط سبيرمان (Spearman) لأن المتغيرين كميان ولا يتبعان التوزيع الطبيعي.

1) ثم ننقر على Continue ثم OK.

فتظهر شاشة المخرجات التي تحوي الجداول التالية:

من الجدول السابق أول شيء نناقشه وهي قيمة الـ Sig فنلاحظ أن قيمة الـ Sig أصغر من 0.05 أي أنه يوجد علاقة ارتباطيه ذات دلالة إحصائية بين متغيري Math Lang , أي أننا رفضنا الفرضية الصفرية وقبلنا الفرضية البديلة والتي تنص على وجود علاقة ارتباطية بين المتغيرين، وقيمة معامل ارتباط سبيرمان بين المتغيرين Lang , Math هو $r = 0.758$.

مثال 3:

	الرضا	إنتاجية
1	1.00	2.00
2	2.00	3.00
3	3.00	1.00
4	4.00	4.00
5	5.00	8.00
6	6.00	7.00
7	7.00	5.00
8	8.00	6.00

تمت دراسة الرضا الوظيفي وأثره على إنتاجية العامل لدى عينة من العاملين في إحدى الشركات تم اختيارهم عشوائياً، وكانت رتبهم كما في الجدول المجاور، والمطلوب هل يوجد علاقة بين متغيري الرضا والإنتاجية.

لتنفيذ المطلوب نتبع الخطوات التالية:

سيتم اختبار الفرضيات التالية:

الفرضية الصفرية: لا يوجد علاقة ذات دلالة معنوية (إحصائية) بين درجات الرضا الوظيفي ورتب الإنتاجية.

الفرضية البديلة: يوجد علاقة ذات دلالة معنوية (إحصائية) بين درجات الرضا الوظيفي ورتب الإنتاجية.

خطوات تنفيذ الاختبار في برنامج (SPSS):

1- ندخل متغيري الرضا والإنتاجية ونعرفهما في شاشة (Variable View ضمن Data Editor)، ثم ندخل البيانات في شاشة (Data View ضمن Data Editor).

2- من أجل اختبار الفرضية الرئيسية وهي دراسة العلاقة بين المتغيرين من شريط الأوامر:

Analyze Correlate Bivariate

سيظهر صندوق الحوار كما في الشكل المجاور:

1) نختار المتغيرات التي نريد حساب معامل الارتباط بينها وننقلها إلى مربع Variables
2) من الحقل المعنون بـ Correlation Coefficient نختار معامل الارتباط الذي نريده حسب قواعد اختيار معاملات الارتباط وهو معامل ارتباط سبيرمان (Spearman) أو معامل كاندل تاو ب (Kendall tau-B).
3) ثم ننقر على Continue ثم OK.

فتظهر شاشة المخرجات التي تحوي الجداول التالية:

بالنسبة لمعامل ارتباط كاندال تاو ب: نجد أن قيمة المعنوية Sig أكبر من 0.05 وهذا يدل على عدم وجود علاقة ارتباطيه ذات دلالة إحصائية بين متغيري رتب الرضا الوظيفي ومتغير رتب الإنتاجية، وهذا يعني أننا قبلنا الفرضية الابتدائية ورفضنا الفرضية البديلة، وفي هذه الحالة لا نناقش قيم معامل الارتباط، لأنه بالأساس لا يوجد ارتباط بين المتغيرين.

أما بالنسبة لمعامل ارتباط سبيرمان: نجد أن قيمة المعنوية Sig أصغر من 0.05 وهذا يدل على وجود علاقة ارتباطية ذات دلالة إحصائية بين متغيري رتب الرضا الوظيفي ومتغير رتب الإنتاجية، وهذا يعني أننا رفضنا الفرضية الابتدائية وقبلنا الفرضية البديلة، وفي هذه تكون قيمة معامل ارتباط سبيرمان $r = 0.714$.

2. الارتباط الجزئي Partial Correlation:

درسنا الارتباط بين متغيرين (Y,X) من حيث كيفية تحديد درجة واتجاه العلاقة، ولكن هذه العلاقة هي في الحقيقة علاقة إجمالية، ذلك لأن معامل الارتباط بين المتغيرين لا يبين درجات الارتباط بين المتغيرين فقط وإنما قد يتأثر بمتغير ثالث (أو بمتغيرات أخرى) ترتبط بالمتغيرات المستقلة أو المتغير التابع.

X Y

C

مثلاً: قد نحصل على قيمة عالية لمعامل الارتباط لبيرسون للعلاقة بين أسعار اللحوم البيضاء واللحوم الحمراء فقد لا توجد علاقة فعلية بين المتغيرين ولكن كلا المتغيرين يتأثر بعامل ثالث هو المستوى العام للأسعار فإذا استبعدنا المستوى العام للأسعار (أو تثبيته) عند قياس العلاقة بين أسعار اللحوم البيضاء والحمراء، فسيتم الحصول على قيمة أقل لمعامل الارتباط، وهذا ما يعرف بالارتباط الجزئي علماً أنه يمكن استبعاد أي عدد من المتغيرات عند قياس العلاقة بين ظاهرتين.

الشروط التي يجب توافرها لإيجاد معاملات الارتباط الجزئية: لضمان معامل الارتباط الجزئي، دقيقاً وموثوقاً، يجب توافر الشروط التالية:

1) **التوزيع الطبيعي:** يجب أن يكون توزيع كل متغير من المتغيرات الداخلة في إيجاد معامل الارتباط الجزئي طبيعياً، فإذا تحقق هذا الشرط فإن العلاقة الوحيدة الموجودة بين المتغيرات هي العلاقة الخطية، وإذا لم يتحقق هذا الشرط فإن العلاقة ربما تكون غير خطية ومن الممكن فحص نوع العلاقة من خلال الرسوم البيانية.
2) **المعاينة العشوائية:** يجب اختيار العينة بشكل عشوائي ومن دون تحيز.

3) **استقلالية المشاهدات:** لا بد أن تكون كل المشاهدات التي تتكون منها البيانات مستقلة عن بعضها بعضا أي أنه يجب ألا تكون القياسات أو عملية قياس المشاهدات متأثرة بأية مشاهدة أو عملية قياس مشاهدة أخرى، وهذا الافتراض مهم جداً وتتم عملية التشكيك في موقف تتم عندما نجمع المشاهدات أو القياسات في موقف جماعي أو يكون فيه شكل ما من التفاعل بين أفراد التجربة.

مثال 4:

لغرض تطبيق الارتباط الجزئي فقد تم استخدام البيانات الخاصة لمجموعة من العمال في أحد المعامل حيث تم تعريف المتغيرات التالية: إنتاجية العامل وعدد ساعات العمل اليومية وعدد سنوات الخبرة بالعمل على الآلة.

	الإنتاجية	ساعات	الخبرة
1	43.00	5.00	5.00
2	44.00	6.00	5.00
3	58.00	8.00	6.00
4	61.00	9.00	7.00
5	63.00	9.00	7.00
6	64.00	10.00	8.00
7	67.00	11.00	8.00
8	68.00	12.00	9.00
9	69.00	12.00	10.00
10	71.00	13.00	12.00
11	71.00	13.00	12.00
12	72.00	10.00	15.00
13	81.00	11.00	15.00

المطلوب:

1. هل يوجد ارتباط بين إنتاجية العامل وعدد سنوات الخبرة، بعد استبعاد أثر عدد سنوات العمل.
2. اختبر معنوية معاملات الارتباط من طرفين بمستوى دلالة 5%.

لتنفيذ المطلوب نتبع الخطوات التالية:

علينا التأكد أولاً من أن المتغيرات تتبع التوزيع الطبيعي.

إجراءات حساب معامل الارتباط الجزئي Partial correlation:

1- من شريط القوائم اختر

Analyze → correlate → partial

سيظهر صندوق حوار بعنوان partial correlation.

في مربع Variables يتم إدخال المتغيرات التي يراد حساب معامل الارتباط الجزئي لها.

في مربع Controlling for يتم إدخال المتغير الذي يراد استبعاد أثره من الحقل المعنون بـ Test Significance: ومن هنا نختار حسب نوع الفرضية التي نريد اختبارها إذا كانت من طرف واحد نختار One Tailed أما في حالة اختبار الفرضية من الطرفين نختار Two Tailed وهو الخيار الافتراضي.

2- انقر Options سيظهر صندوق حوار Partial Correlation: Option، ومن هذه الشاشة، يمكن إجراء ما يلي:

a. عند تأشير الخيار Means and Standard Deviation سنحصل على وصف إحصائي للمتغيرات الداخلة في التحليل، وهي Standard Deviation و N أي عدد الحالات والوسط الحسابي Mean.

b. أما في حالة تأشير الخيار Zero-order Correlation سنحصل على معاملات الارتباطات لجميع المتغيرات بما فيها المتغير الذي تم تثبيته وعلى شكل مصفوفة.

3- اضغط continue، ستعود إلى صندوق الحوار المعنون بـ: Partial Correlation.

4- عند التأشير على Display actual significance level سيتم إظهار قيمة المعنوية في جداول الارتباط.

5- اضغط Ok، ستظهر نتائج الحل في شاشة المخرجات Output.

الجـــدول الأول: والمعنـــون بـــ Descriptive Statistics وهو جـــدول يحـــوي المتوســـطات الحســابية والانحرافـات المعياريـة وحجـم العينـة لكـل متغيـر مـن المتغيرات التي تم اختيارها.

Descriptive Statistics

	Mean	Std. Deviation	N
y	64.0000	10.75484	13
x2	6.0000	2.08167	13
x1	9.0000	2.61406	13

نلاحظ من الجدول المعنون بـ Correlations أنه قسم إلى قسمين حيث:

- القسم الأول والمعنون بـ non: ويمثل قيم معاملات الارتباط وقيم المعنوية لجميع المتغيرات بما فيها المتغير الذي نريد استبعاد أثره، حيث نلاحظ أنه يوجد ارتباط بين جميع المتغيرات.
- القسم الثاني والمعنون بـ عدد ساعات العمل تم حساب معاملات الارتباط الجزئية كما يظهر في الجدول السابق والذي يتضمن معاملات الارتباط الجزئية، وعدة مفردات العينة n وقيم المعنوية Sig فإذا كانت قيمة المعنوية أقل من مستوى المعنوية 0.05 فإن قيمة معامل الارتباط الجزئي مقبولة إحصائياً ما بين المتغير التابع Y والمتغير المستقل المراد فحص علاقته معه، أما إذا كانت هذه القيمة أكبر من المستوى المقبول، فإن معامل الارتباط غير مقبول إحصائياً، وعندئذ لا يمكن الإقرار بوجود علاقة بين المتغيرين. في مثالنا هنا نجد أن قيمة المعنوية هنا (Sig = 0.005) أصغر من مستوى الدلالة 0.05 وهذا يدل على وجود ارتباط بين إنتاجية العامل وعدد سنوات الخبرة بعد استبعاد أثر متغير ساعات العمل كما نلاحظ أن قيمة معامل الارتباط قد نقصت عما كانت عليه قبل استبعاد أثر متغير عدد ساعات العمل.

3. الانحدار:

كان جالتون (Galton) أول من استخدم فكرة الانحدار في مجال علم الوراثة، إذ لفت نظره وراثة صفة طول القامة، فالأطفال الذين يكون آباؤهم طوال القامة يميلون لأن يكونوا أقصر قامة من آبائهم، والعكس من ذلك الأطفال الذين يكون آباؤهم قصار القامة يميلون لأن يكونوا أطول قامة من آبائهم، واستنتج أن طول الأبناء يميل إلى التراجع أو الانحدار نحو المتوسط العام للطول في المجتمع الأصلي، وفي ضوء هذه النتائج أطلق جالتون على ذلك مفهوم الانحدار ولكن هذا المفهوم توسع وأخذ أبعاد متعددة في الوقت الحاضر ليعني الانحدار نحو المتوسط والتنبؤ، فيعرف الانحدار بأنه معادلة رياضية تعبر عن العلاقة بين متغير/متغيرات مستقلة والمتغير التابع فإذا تمكن الباحث من تحديد

هذه العلاقة فيمكنه أن يحدد المعادلة الرياضية بين المتغيرين/المتغيرات التي تستخدم فيما بعد للتنبؤ بقيم المتغير التابع من خلال قيم المتغير/المتغيرات المستقلة، فعملية التنبؤ هدف من أهداف العلوم الاجتماعية والتي يسعى الباحث الاجتماعي من خلالها الحصول على بيانات غير معروفة بناءً على بيانات معروفة وذات صلة بالظاهرة المدروسة، ويمكن الوصول إلى قيمة مستقبلية في المتغير التابع اعتماداً على قيمة معينة معروفة في المتغير/ المتغيرات المستقلة عن طريق أساليب الانحدار التالية:

- الانحدار الخطي البسيط (Simple Linear Regression): يستخدم الانحدار الخطي البسيط عندما نفرض وجود متغير مستقل واحد يؤثر على متغير تابع، أي المتغير المستقل يفسر مقدار التغير الموجود في المتغير التابع.
- الانحدار الخطي المتعدد (Multiple Linear Regression): يستخدم الانحدار الخطي المتعدد عندما نفرض وجود أكثر من متغير مستقل تؤثر جميعها في نفس الوقت على المتغير التابع، أي المتغيرات المستقلة تفسر مقدار التغير الموجود في المتغير التابع.

وسيتم في هذا الفصل مناقشة **الانحدار الخطي البسيط فقط:**

إن دراسة الانحدار تهدف بشكل أساسي إلى تحديد العلاقة بين المتغيرين بطريقة رياضية للاستفادة منها في عملية التنبؤ فكلما كانت العلاقة قوية بين المتغيرين كلما كان التقدير (التنبؤ) أفضل، فإذا كانت العلاقة بين المتغيرين خطية فالانحدار خطي والعلاقة الرياضية بين المتغيرين تكون على شكل معادلة خط مستقيم يسمى خط الانحدار وهو خط الملائمة الأفضل Line of fit والذي يمثل جميع النقاط في شكل الانتشار أفضل تمثيل والذي تكون فيه مجموع مربعات انحراف القيم المقدرة عن القيم الحقيقية (الفعلية) أقل ما يمكن، أو مجموع انحرافات القيم عن خط الانحدار تساوي الصفر، وهذا ما يبينه الشكل التالي:

ويأخذ نموذج الانحدار الخطي البسيط الصيغة التالية:

$$\hat{y}_i = a + b.x_i + e \qquad (7-11)$$

حيث إن: $\hat{y}_i$ يمثل القيمة المقدرة أو المحسوبة للمتغير التابع.

x_i يمثل المتغير المستقل.

a ثابت الانحدار (الحد الثابت).

b معامل الانحدار (يمثل الميل).

e الخطأ العشوائي.

ويتم إيجاد هذه المعادلة بطريقتين:

أ- **طريقة الرسم الحر (طريقة الرسم البياني):**

تعتمد هذه الطريقة على رسم شكل الانتشار للظاهرة المدروسة باعتبار المحور (X) يمثل المتغير المستقل والمحور الأفقي (Y) يمثل المتغير التابع ويتم رسم خط الانحدار بحيث يمر بين منتصف نقاط الانتشار ويتم تحديد معادلة خط الانحدار على النحو التالي:

$$\hat{y} = a + b.x$$

a عبارة عن نقطة تقاطع خط الانحدار مع المحور (Y) عندما تكون قيمة المتغير المستقل (X) تساوي الصفر.

b عبارة عن ميل خط الانحدار (ظل الزاوية) مع المحور (x).

ويتم حساب ظل الزاوية من خلال أخذ نقطتين على خط الانحدار وإجراء عملية إسقاط منها على المحور (x) والمحور (y) كما هو موضح في المثال التالي:

وتتميز هذه الطريقة بالسهولة والسرعة ولكن من سيئاتها أنها قد تعطي نتائج غير دقيقة ومختلفة تبعاً للخبرة الشخصية، واختلاف أسلوب الرسم، حيث يمكن أن نحصل على عدد من الخطوط المستقيمة تساوي عدد القائمين على عملية الرسم.

ب- الطريقة الرياضية:

وتتلخص هذه الطريقة بإيجاد قيمة الثوابت a و b التي تجعل مجموع مربعات انحرافات القيم عن خط الانحدار أصغر ما يمكن، أي إيجاد a و b التي تجعل مجموع فروقات القيم ومقدراتها أصغر ما يمكن أي $(\sum_{i=1}^{i=n}(y_i-\hat{y}_i)^2 \to \min)$ ويتم ذلك بطريقتين:

الطريقة الأولى: وتسمى طريقة المربعات الصغرى وهي طريقة رياضية تعتمد على اشتقاق معادلة الانحدار $((y-a-bx)^2=0)$ بالنسبة لكل من (a) و (b)، سنحصل على المعادلتين التاليتين:

$$\sum_{i=1}^{n} y = na + b\sum_{i=1}^{n} x \qquad (8-11)$$

$$\sum_{i=1}^{n} x.y = a\sum_{i=1}^{n} x + b\sum_{i=1}^{n} x^2 \qquad (9-11)$$

وبحل جملة المعادلتين نحصل على قيم الثوابت (a) و (b) وبذلك على معادلة خط الانحدار.

الطريقة الثانية: وتعتمد على استخدام الأوساط الحسابية لكل متغير (التابع والمستقل) لنحصل على الثوابت:

$$b = \frac{n\sum_{i=1}^{n} x_i . y_i - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \qquad (10-11)$$

$$a = \bar{y} - b\bar{x} \qquad (11-11)$$

ويمكن إيجاد قيمة b بصيغة أخرى:

$$b = \frac{\sum_{i=1}^{n} x_i . y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^{n} x_i^2 - n\bar{X}^2} \qquad (12-11)$$

$$b = \frac{\sum_{i=1}^{n} (x_i - \bar{X}).(y_i - \bar{Y})}{\sum_{i=1}^{n} (x_i - \bar{X})^2} \qquad (11-13)$$

مثال 5:

البيانات التالية تمثل العمر X، ضغط الدم Y لعينة مكونة من 10 أشخاص وقد تم إدخال البيانات في ورقة Data Editor.

	Age	Pressure
1	35	112
2	40	128
3	38	130
4	44	138
5	67	158
6	64	162
7	59	140
8	69	175
9	25	125
10	50	142

المطلوب:

1- اكتب نموذج التنبؤ التقديري مفترضاً العلاقة الخطية بين X_i و Y_i واختبر معنوية معالم النموذج (أكتب معادلة الانحدار بالشكل $\hat{y}_i = b_0 + b_1 x_i$)

2- استخرج فترة الثقة 95% لكل من معلمتي الانحدار B_1 , B_0.

3- اختبر التوزيع الطبيعي للأخطاء العشوائية بيانياً.

لتنفيذ المطالب أعلاه نتبع الخطوات التالية:

أ- من شريط القوائم:

a. Analyze→Regression→Linear

فيظهر صندوق حوار بعنوان Linear Regression

b. في مستطيل Dependent تضع المتغير التابع.

c. في مستطيل Independent(s) يمكن إدخال متغير مستقل واحد هنا أو مجموعة من المتغيرات المستقلة، كما نلاحظ أن هذا القسم معنون بـ Block له فائدة بحيث إذا كان لدينا نموذجان لأحدهما متغير مستقل X والنموذج الآخر له متغير مستقل آخر Z مع نفس المتغير التابع Y ، في هذه الحالة يتم إدخال المتغير X في Block1 ثم نضغط على next ثم يتم إدخال المتغير Z في Block.

ب- Method: تمثل نوع الطريقة المستخدمة في الانحدار، وتكون لها فائدة عندما نكون أمام حالة الانحدار المتعدد، وهي التي سوف نناقشها فيما بعد، وبشكل افتراضي يكون خيار Enter هو الفعال.

ج- اضغط الزر **Statistics** فيظهر صندوق حوار بعنوان Linear Regression Statistic وكما هو مبين في الشكل:

- Estimate: لتقدير معالم نموذج الانحدار.
- Confidence Interval: لتقدير فترة ثقة لمعالم الانحدار.
- Model Fit: لعرض R^2 و ANOVA.
- Descriptives: لعرض المتوسط الحسابي والانحراف المعياري لكلا المتغيرين التابع والمستقل.
- Part and partial correlations: لعرض معامل ارتباط بيرسون بين المتغيرين التابع والمستقل ومعنوية الارتباط بين المتغيرين.
- ثم نضغط على Continue للعودة إلى صندوق حوار Linear Regression.

د- انقر الخيار **Plots**، فيظهر صندوق حوار بعنوان Linear Regression: plots.

ومنها يتم:

- Normal probability plot: لاختبار التوزيع الطبيعي للأخطاء العشوائية باستخدام مخطط Normal P–P.
- Histogram: أيضاً لاختبار التوزيع الطبيعي للأخطاء العشوائية.
- ثم نضغط على Continue لعودة إلى صندوق حوار Linear Regression.

ه- لغرض تضمين النموذج الحد الثابت أو حذفه، فيتم ذلك من الخيار **options**.

و- عند اختيار الزر OK من صندوق حوار Linear Regression تظهر في شاشة المخرجات الجداول التالية:

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	Age[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: Pressure

وهذا الجدول لتحديد الطريقة المعتمدة في التحليل واسم المتغير المستقل.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.900[a]	.810	.787	8.824

a. Predictors: (Constant), Age

b. Dependent Variable: Pressure

وهذا الجدول يبين قيمة معامل الارتباط بالقيمة المطلقة في عامود R وفي مثالنا هنا نجد أن قيمته هي 0.90 كما أن قيمة R Square وهي مربع معامل الارتباط هي 0.810 والذي يدعى

معامل التحديد.

الجدول التالي المعنون بـ ANOVA

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2661.050	1	2661.050	34.174	.000[a]
	Residual	622.950	8	77.869		
	Total	3284.000	9			

a. Predictors: (Constant), Age
b. Dependent Variable: Pressure

هنا نضع فرضيات الانحدار كما يلي:

H_0: لا يوجد انحدار.

H_1: يوجد انحدار.

نلاحظ أن قيمة Sig=0.00 أصغر من 0.05 وبالتالي نرفض فرضية العدم وبالتالي يوجد انحدار.

الجدول التالي:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations		
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part
1	(Constant)	85.044	9.970		8.530	.000	62.052	108.036			
	Age	1.140	.195	.900	5.846	.000	.690	1.589	.900	.900	.900

a. Dependent Variable: Pressure

الجدول المعنون بـ Coefficients يظهر لنا معاملات الارتباط ولكن قبل وضع المعاملات في معادلة الانحدار يجب أن نجري اختبار للفرضيات حول هذه المعاملات:

$$H_0 : b_0 = 0$$
$$H_1 : b_0 \neq 0$$

لمناقشة هذه الفرضية من السطر المعنون بـ Constant نجد أن قيمة Sig =0.0 أصغر من 0.05 وبالتالي نرفض الفرضية الابتدائية ونقبل البديلة أي أن $b_0 = 85.044$

$$H_0 : b_1 = 0$$
$$H_1 : b_1 \neq 0$$

لمناقشة هذه الفرضية من السطر المعنون باسم المتغير المستقل Age نجد أن قيمة Sig =0.0 أصغر من 0.05 وبالتالي نرفض الفرضية الابتدائية ونقبل البديلة أي أن $b_1 = 1.140$ وبعد أن تم قبول معاملات الارتباط يمكننا كتابة معادلة الانحدار، فيكون لها الشكل التالي: $\hat{y}_i = 85.044 + 1.140 * x_1$.

وتشير المعادلة إلى أن زيادة العمر سنة واحدة تؤدي إلى زيادة مستوى الضغط بمقدار 1.140 وحدة الأرقام داخل الأقواس تشير إلى الخطأ المعياري للمعلمة المقابلة.

كما نجد من الجدول السابق أن معاملات الانحدار لها مجال ثقة بمقدار 95%.

يبين الجدول المعنون بـ Residuals Statistics الإحصاء الوصفي للقيم المتنبأ ضمن مجال قيم المتغير المستقل في سطر Predicted Value.

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	113.53	163.68	141.00	17.195	10
Residual	-12.931-	11.465	.000	8.320	10
Std. Predicted Value	-1.597-	1.319	.000	1.000	10
Std. Residual	-1.465-	1.299	.000	.943	10

a. Dependent Variable: Pressure

أما السطر الثاني المعنون Residual فيبين الإحصاء الوصفي للفرق بين القيم الحقيقية للمتغير التابع والقيم المتنبأ بها.

كما نجد من شاشة المخرجات رسماً مخططياً Histogram ومخطط Normal P-P للفروق بين القيم الحقيقية للمتغير التابع والقيم المتنبأ بها وذلك لاختبار التوزيع الطبيعي للفروق.

المراجع

1. المراجع العربية

- البلداوي، عبد الحميد عبد المجيد. **أساليب البحث العلمي والتحليل الإحصائي**، دار الشروق، عمان، 2007.
- الرفاعي، أحمد حسين. **مناهج البحث العلمي: تطبيقات إدارية واقتصادية**، دار وائل، عمان الأردن.
- العاقل، محمد عادل. **محاضرات في المفاهيم الأساسية للعينات والتحيز والأخطاء العينية وغير العينية**، المكتب المركزي للإحصاء، دمشق، 1996.
- القاضي، دلال؛ العبدالله، سهيلة؛ البياتي، محمود، **الإحصاء للإداريين والاقتصاديين**، دار مكتبة الحامد، عمان الأردن، 2003.
- الكيلاني، عبدالله زيد؛ الشريفين، نضال كمال. **مدخل إلى البحث في العلوم التربوية والاجتماعية أساسياته مناهجه تصاميمه أساليبه الإحصائية**، دار المسيرة للنشر والتوزيع، 2004.
- المحمد، محمد جاسم، **المعلوماتية تحليل البيانات**، منشورات جامعة دمشق، ط1، 2012.
- النجار، عبدالله بن عمر. **استخدام حزمة البرامج الإحصائية SPSS في تحليل البيانات، جامعة الملك فيصل الإحساء، ط1.**
- النجار، فايز جمعه؛ النجار، نبيل جمعه؛ الزعبي ماجد راضي. **أساليب البحث العلمي منظور تطبيقي**، دار الحامد للنشر والتوزيع، عمان الأردن، 2009.
- أبو زيد، محمد خير سليم. **التحليل الاحصائي للبيانات باستخدام برمجية SPSS**، دار جرير، عمان الأردن، 2010.
- أبو علام، رجاء محمود، **مناهج البحث في العلوم النفسية والتربوية**، دار النشر للجامعات، القاهرة مصر، 2004.

المراجع

- بدر، أحمد. **أصول البحث العلمي ومناهجه**، المكتبة الأكاديمية، ط9، القاهرة، 1996.
- دويدري، رجاء وحيد. **البحث العلمي أساسياته النظرية وممارسته العلمية**، دار الفكر، دمشق، سورية، 2000.
- عبيدات، محمد؛ أبو نصار، محمد؛ أبو نصار، عقلة، **منهجية البحث العلمي**، دار وائل، عمان الأردن، ط1، 1997.
- عليان، ربحي مصطفى؛ غنيم، عثمان محمد. **مناهج وأساليب البحث العلمي: النظرية والتطبيق**، دار صفاء للنشر والتوزيع، عمان الأردن، 2000.
- عمورة، عدنان، **الإحصاء (2) الطرق الوسيطية في الاستدلال الإحصائي**، منشورات جامعة دمشق، 2011.
- عمورة، عدنان؛ صبح، محمد؛ قاسم، عزات. **نظرية العينات**، منشورات جامعة دمشق، 1999.
- عودة، أحمد سليمان؛ خليل، خليل يوسف، **أساسيات البحث العلمي في التربية والعلوم الإنسانية**، مكتبة الكتاني، إربد، ط2، 1992.
- غانم، عدنان؛ علي، نبيل، **الإحصاء التطبيقي**، منشورات جامعة دمشق، 2006.
- فهمي، محمد شامل بهاء الدين. **الإحصاء بلا معاناة: المفاهيم مع التطبيقات باستخدام برنامج SPSS**، الرياض 1426 هـ.
- كنجو، أنيس، الإحصاء والاحتمال، مكتبة العبيكان، ط1، الرياض السعودية، 2000.
- محمد، أماني موسى. **التحليل الإحصائي للبيانات**، مركز تطوير الدراسات العليا والبحوث في العلوم الهندسية، 2007، ط1.
- محمد، جمال محمد شاكر، **المرشد في تحليل البيانات باستخدام SPSS**، **الدار الجامعية**، ط1، القاهرة، 2005.
- مخائيل، امطانيوس. **القياس والتقويم في التربية الحديثة**، منشورات جامعة دمشق.

المراجع

- مخـول، مطـانيوس؛ غـانم، عـدنان. **مبـادئ الإحصـاء، منشـورات جامعـة دمشق**،2011.
- مصـطفى، عبـد الحفيظ محمـد فـوزي، **الاسـتدلال الإحصـائي (2): نظريـة اختبـار الفرضيات**، مجموعة النيل العربية، القاهرة مصر، ط1، 2002.
- منيـزل، عبـدالله؛ غرايبـة، عـايش. **الإحصـاء التربـوي- تطبيقـات باسـتخدام الـرزم الإحصائية للعلوم الاجتماعية**، دار المسيرة للنشر والتوزيع، عمان، 2005.
- نجيـب، حسـين علـي؛ الرفـاعي، غالـب عـوض صـالح، **تحليـل ونمذجـة البيانـات باستخدام الحاسوب**، الأهلية للنشر والتوزيع، عمان الأردن، 2006.

2. المراجع الأجنبية

- Arthur Griffith, **SPSS For Dummies**, Wiley Publishing, Inc, 2007.
- Mark Vernoy; Diana J.Kyle, **Behavioral Statistics in Action**, Mc Graw Hill third Edition, 2002.
- Serkan, Uma. **Research methods for business**, 1992.
- Sheridan J Coakes, **SPSS: analysis without anguish**, John Wiley & Sons Australia, LTD, 2005.
- SPSS Statistics 17.0 Brief guide 2014.
- William Mendenhall; Robert J.Beaver; Barbara M.Beaver, **Introduction To Probability and Statistics**, Duxbury Tenth Edition, 1999.
- Zar,Jerrold H. **Biostatistical Analysis**, Prentice Hall, 5th ed, 2010.

اللجنة العلمية:

د. عدنان غانم

د. حيدر عباس

د. ياسر الجندي

المدقق اللغوي

د. سكينة موعد

حقوق الطبع والترجمة والنشر محفوظة لمديرية الكتب والمطبوعات الجامعية

وقل رب زدني علما
جامعة دمشق
Damascus University