جامعة الشهيد حمه لخضر بالوادي
كلية العلوم الاقتصادية والتجارية وعلوم التسيير

دروس عبر الخط في مقياس:

الأدوات الإحصائية لتحليل البيانات

المستوى السنة الثالثة مالية مؤسسة

من إعداد: د/ لطفي مخزومي

السنة الجامعية 2022/2021

الدرس الأول:
مفهوم علم الإحصاء

1. تعريف علم الإحصاء:

يتم تعريف الإحصـــاء على أنه فن وعلم جمع البيانات وتحليلها وتقديمها وتفســـيرها. لا ســـيما في مجال الأعمال والاقتصـاد، فإن المعلومات المقدمة من خلال جمع البيانات وتحليلها وتقديمها وتفسـيرها تمنح المديرين وصـناع القرار فهمًا أفضل لبيئة الأعمال والبيئة الاقتصادية، وبالتالي تمكنهم من اتخاذ قرارات أكثر استنارة وأفضل. في بيئة الأعمال والاقتصاد العالمية اليوم، يمكن لأي شـخص الوصـول إلى كميات هائلة من المعلومات الإحصـائية. يفهم المدراء وصـناع القرار الأكثر نجاحًا المعلومات ويعرفون كيفية استخدامها بفعالية. تاريخيا، ظهر الإحصـاء الوصـفي قبل الإحصـاء الاستدلالي. تم إجراء التعدادات منذ زمن بعيد مثل العصـر الروماني. على مر القرون، أدت سـجلات مثل المواليد والوفيات والزواج والضـرائب بشـكل طبيعي إلى تطوير الإحصـاء الوصـفي. الإحصـاء الاسـتدلالي هو الأحدث، حيث بدأت التطورات الرئيسـية مع البحث الذي أجراه كارل بيرســـون (1857-1936)Karl Pearson ورونالد فيشـــر (1890-1962) Ronald Fisher، الذين نشـروا نتائجهم في السـنوات الأولى من القرن العشـرين. منذ عمل بيرسـون وفيشـر، تطورت الإحصـاء الاستنتاجي بسرعة ويتم تطبيقه الآن في عدد لا يحصى من المجالات (Weiss, 2016, p. 5).

يـتم تطبيـق الأسـاليب الإحصـائية لتفسـير البيـانات مـع وجـود التبـاين. خـلال عمليـة حـل المشـكلات البحثيـة المختلفـة، غالبًـا مـا يواجـه البـاحثون بيـانات تُظهـر التبـاين. يـوفر الإحصـاء أدوات أساسـية للتعامـل مـع التبـاين الملحـوظ. تتضـمن أمثلـة تطبيـق الإحصـاء في عمليـة حـل المشـكلات البحثيـة، علـى سـبيل المثـال لا الحصر، ما يلي (Montgomery & Runger, 2003, p. 12):

- تلخيص وعرض البيانات: ملخص رقمي وتصور في الإحصاء الوصفي؛
- استنتاج الخصائص (المتوسط والوسيط والنسبة والتباين) لمجتمع واحد/إثنين باستخدام الاختبارات الإحصائية المعلمية واللامعلمية؛
- اختبار العلاقة بين المتغيرات: تحليل الارتباط وتحليل البيانات الفئوية؛
- نمذجة العلاقة السببية بين المتغير التابع والمتغيرات المستقلة: تحليل الانحدار، تحليل التباين؛
- تحديد مصادر التباين في متغير الاستجابة: تحليل التباين؛
- تقييم الأهمية النسبية للعوامل لمتغير الاستجابة: تحليل الانحدار، تحليل التباين.

تتمثل الأهداف الرئيسية للإحصاء فيما يلي (Johnso & Bhattacharyya, 1987, p. 8):

- لعمل استنتاجات حول المجتمع من تحليل المعلومات الواردة في عينة البيانات. وهذا يشمل تقييمات لمدى عدم اليقين المتضمن في هذه الاستنتاجات؛
- تصميم العملية ومدى أخذ العينات بحيث تشكل المشاهدات أساسًا لاستخلاص استنتاجات صحيحة.

2. تطبيقات علم الإحصاء في ميدان العلوم الاقتصادية والمالية:

في بيئة الأعمال والاقتصاد العالمية اليوم، يمكن لأي شخص الوصول إلى كميات هائلة من المعلومات الإحصائية. يفهم المدراء وصناع القرار الأكثر نجاحًا المعلومات ويعرفون كيفية استخدامها بفعالية. ويمكن توضيح بعض استخدامات الإحصاء في الأعمال والاقتصاد كما يلي (Anderson, et al., 2019, pp. 3-4):

- تستخدم شركات المحاسبة العامة إجراءات أخذ العينات الإحصائية عند إجراء عمليات تدقيق لعملائها. على سبيل المثال، لنفترض أن شركة محاسبة تريد تحديد ما إذا كان مبلغ الذمم المدينة المعروض في الميزانية العمومية للعميل يمثل إلى حد ما المبلغ الفعلي للحسابات المدينة. عادةً ما يجعل العدد الكبير من الحسابات المستحقة القبض الفردية مراجعة والتحقق من صحة كل حساب مضيعة للوقت ومكلفة للغاية. كممارسة شائعة في مثل هذه الحالات، يختار موظفو التدقيق مجموعة فرعية من الحسابات تسمى عينة. بعد مراجعة دقة الحسابات التي تم أخذ عينات منها، يتوصل المدققون إلى استنتاج حول ما إذا كان مبلغ الذمم المدينة المعروض في الميزانية العمومية للعميل مقبولاً.
- يستخدم المحللون الماليون مجموعة متنوعة من المعلومات الإحصائية لتوجيه توصياتهم الاستثمارية. في حالة الأسهم، يقوم المحللون بمراجعة البيانات المالية مثل نسب السعر / الأرباح وعوائد الأرباح. من خلال مقارنة المعلومات الخاصة بسهم فردي بمعلومات حول متوسطات سوق الأسهم، يمكن للمحلل البدء في استنتاج ما إذا كان السهم استثمارًا جيدًا.
- الماسحات الضوئية الإلكترونية في عدادات البيع بالتجزئة تجمع البيانات لمجموعة متنوعة من تطبيقات أبحاث التسويق. يمكن لمديري العلامات التجارية مراجعة إحصائيات الماسح الضوئي وإحصائيات النشاط الترويجي لاكتساب فهم أفضل للعلاقة بين الأنشطة الترويجية والمبيعات. غالبًا ما تكون مثل هذه التحليلات مفيدة في وضع استراتيجيات تسويق مستقبلية للمنتجات المختلفة.
- يقدم الاقتصاديون في كثير من الأحيان تنبؤات حول مستقبل الاقتصاد أو بعض جوانبه. يستخدمون مجموعة متنوعة من المعلومات الإحصائية في عمل مثل هذه التنبؤات. على سبيل المثال، في التنبؤ بمعدلات التضخم، يستخدم الاقتصاديون المعلومات الإحصائية حول مؤشرات مثل مؤشر أسعار المنتجين، ومعدل البطالة، واستخدام القدرة التصنيعية. غالبًا ما يتم إدخال هذه المؤشرات الإحصائية في نماذج التنبؤ المحوسبة التي تتنبأ بمعدلات التضخم.
- مسؤولو نظم المعلومات مكلفون بالتشغيل اليومي لشبكات الكمبيوتر الخاصة بالمؤسسة. تساعد مجموعة متنوعة من المعلومات الإحصائية المسؤولين على تقييم أداء شبكات الكمبيوتر، بما في ذلك شبكات المنطقة المحلية (LAN)، وشبكات المنطقة الواسعة (WAN)، وقطاعات الشبكة، والشبكات الداخلية، وأنظمة اتصالات البيانات الأخرى. يعد الإحصاء مثل متوسط عدد المستخدمين على النظام، ونسبة الوقت الذي يتم فيه تعطل أي مكون من مكونات النظام، ونسبة النطاق الترددي المستخدم في أوقات مختلفة من اليوم، أمثلة على المعلومات الإحصائية التي تساعد مسؤول النظام على فهم وإدارة شبكة الكمبيوتر بشكل أفضل.

3. أقسام علم الإحصاء:

الإحصاء يعني "الأوصاف العددية" لمعظم الناس. تمثل أرقام البطالة الشهرية، ومعدل فشل الشركات الناشئة، ونسبة المديرات التنفيذيات في صناعة معينة، أوصافًا إحصائية لمجموعات كبيرة من البيانات التي تم جمعها حول بعض الظواهر. غالبًا ما يتم اختيار البيانات من مجموعة أكبر من البيانات التي نرغب في تقدير خصائصها. نسمي عملية الاختيار هذه بأخذ العينات. على سبيل المثال، قد تجمع أعمار عينة من عملاء شركة خدمات بث الفيديو لتقدير متوسط عمر جميع عملاء الشركة. ثم يمكنك استخدام تقديرك لاستهداف إعلانات الشركة للفئة العمرية المناسبة. لاحظ أن الإحصائيات تتضمن عمليتين مختلفتين: (1) وصف مجموعات البيانات و (2) استخلاص النتائج (عمل التقديرات والقرارات والتنبؤات وما إلى ذلك) حول مجموعات البيانات القائمة على أخذ العينات. لذلك، يمكن تقسيم تطبيقات الإحصاء إلى مجالين كبيرين: الإحصاء الوصفي والإحصاء الاستدلالي (McClave, Benson, & Sincich, 2018, p. 28):

- **الإحصاء الوصفي**، والذي يستخدم الأساليب العددية والرسومية لاستكشاف البيانات، أي البحث عن أنماط في مجموعة بيانات، لتلخيص المعلومات التي تم الكشف عنها في مجموعة البيانات، وتقديم المعلومات في شكل مناسب.
- **الإحصاء الاستدلالي**، والذي يستخدم بيانات العينة لإجراء تقديرات أو قرارات أو تنبؤات أو تعميمات أخرى حول مجموعة أكبر من البيانات.

في الإحصاء الوصفي، يحاول الإحصائي وصف الموقف. باعتبار التعداد الوطني الذي تجريه حكومة الولايات المتحدة كل 10 سنوات. تشير نتائج هذا التعداد الى متوسط العمر والدخل والخصائص الأخرى لسكان الولايات المتحدة. للحصول على هذه المعلومات، يجب أن يكون لدى مكتب الإحصاء بعض الوسائل لجمع البيانات ذات الصلة. بمجرد جمع البيانات، يجب على المكتب تنظيمها وتلخيصها. أخيرًا، يحتاج المكتب إلى وسيلة لتقديم البيانات في شكل ذي معنى، مثل المخططات أو الرسوم البيانية أو الجداول. في الإحصاء الاستدلالي، يحاول الإحصائي عمل استنتاجات من عينات إلى المجتمعات. تستخدم الإحصائيات الاستدلالية الاحتمال، أي فرصة وقوع حدث ما. مجال الإحصاء الاستدلالي الذي يسمى اختبار الفرضيات هو عملية صنع القرار لتقييم الادعاءات حول المجتمع، بناءً على المعلومات التي تم الحصول عليها من العينة. على سبيل المثال، قد يرغب الباحث في معرفة ما إذا كان دواء جديد سيقلل من عدد النوبات القلبية لدى الرجال الذين تزيد أعمارهم عن 70 عامًا. لهذه الدراسة، سيتم اختيار مجموعتين من الرجال فوق سن 70. سيتم إعطاء مجموعة واحدة الدواء، وسيتم إعطاء الأخرى دواء وهمي (مادة ليس لها فوائد طبية أو ضرر). في وقت لاحق، سيتم حساب عدد النوبات القلبية التي تحدث في كل مجموعة من الرجال، وسيتم إجراء اختبار إحصائي، وسيتم اتخاذ قرار بشأن فعالية الدواء (Bluman, 2019, pp. 92-93).

4. التفكير الإحصائي والنقدي:

تتكون العملية المتضمنة إجراء دراسة إحصائية من "التحضير والتحليل والاستنتاج". نبدأ بإعداد يتضمن النظر في "السياق"، والنظر في مصدر البيانات، والنظر في طريقة أخذ العينات. بعد ذلك، نقوم ببناء الرسوم البيانية المناسبة، واستكشاف البيانات، وتنفيذ الحسابات المطلوبة للطريقة الإحصائية المستخدمة. أخيرًا، نقوم بتكوين استنتاجات من خلال

تحديد ما إذا كانت النتائج لها دلالة إحصائية ودلالة عملية. يتضمن الشكل رقم (01) العناصر الأساسية في دراسة إحصائية. لاحظ أن الإجراء الموضح في هذا الشكل لا يركز على الحسابات الرياضية. بفضل التطورات الكبيرة في التكنولوجيا، لدينا الآن أدوات تقوم بفعالية بمعالجة الأرقام حتى نتمكن من التركيز على فهم النتائج وتفسيرها (Triola, 2014, p. 05).

الشكل (01): التفكير الاحصائي

التحضير والإعداد:

1. السياق
ماذا تعني البيانات؟
ما هو الهدف من الدراسة؟
2. مصدر البيانات
هل البيانات من مصدر متخصص ومعتمد؟
3. طريقة المعاينة
هل تم جمع البيانات بطريقة غير متحيزة ، أم بطريقة متحيزة؟

التحليل:

1. رسم البيانات
2. استكشاف البيانات
هل هناك أي قيم متطرفة (أرقام بعيدة جدًا عن جميع البيانات الأخرى تقريبًا)؟
ما هي الإحصائيات المهمة التي تلخص البيانات (مثل المتوسط والانحراف المعياري)؟
كيف تتوزع البيانات؟
هل هناك بيانات مفقودة؟
3. تطبيق الأساليب الإحصائية
استخدام التكنولوجيا والبرمجيات للحصول على النتائج

الاستنتاج:

1. الدلالة (المعنوية) الإحصائية
هل النتائج لها دلالة إحصائية؟
هل النتائج لها دلالة عملية؟

المصدر: (Triola, 2014, p. 06)

الدرس الثاني:
البيانات الإحصائية

1. أنواع البيانات الإحصائية:

في كل مرة تجري فيها عملية شراء عبر الإنترنت، يتم التقاط معلومات أكثر من مجرد تفاصيل عملية الشراء نفسها. ما الصفحات التي بحثت عنها للوصول إلى مشترياتك؟ كم من الوقت قضيت في النظر إلى كل منهما؟ تسمى هذه القيم المسجلة، سواء كانت أرقامًا أو تسميات مع سياقها، بالبيانات. يتم تسجيلها وتخزينها إلكترونيًا، في مستودعات رقمية ضخمة تسمى مستودعات البيانات. لطالما اعتمدت الشركات على البيانات لاتخاذ قرارات جيدة، ولكن اليوم، أكثر من أي وقت مضى، تستخدم الشركات البيانات لاتخاذ قرارات بشأن جميع جوانب أعمالها تقريبًا، من المخزون إلى الإعلان إلى تصميم مواقع الويب. تساعد كل تمريرة لبطاقتك الائتمانية وكل نقرة على الماوس في نمو مستودعات البيانات هذه. تندرج تحديات جمع وإدارة وتخزين وتنظيم كل هذه المعلومات بشكل جماعي تحت مصطلح البيانات الضخمة (Big Data). لكن البيانات وحدها لا تستطيع اتخاذ قرارات جيدة. لبدء عملية تحويل البيانات إلى معلومات مفيدة، عليك أولاً معرفة القرارات التي تريد اتخاذها. بدون سؤال، ليس لديك أي فكرة عما قد يكون مثيرًا للاهتمام حول البيانات. هل يجب أن تنظر إلى وقت المعاملات، أم موقعها، أو سعرها، أو المنتجات التي تم شراؤها، أو أي شيء آخر؟ ستساعد معرفتك بقضايا العمل والأسئلة التي تريد الإجابة عليها في توجيه بحثك عن رؤى من البيانات، وتساعدك على تسخير البيانات لاتخاذ قرارات أفضل (Sharpe, De Veaux, & Velleman, 2021, pp. 35-36).

أي مجموعة من البيانات تحتوي على معلومات حول مجموعة من الأفراد [1](Individuals). المعلومات منظمة في متغيرات [2](variables). تتضمن قاعدة بيانات الطلاب بالكلية، على سبيل المثال، بيانات حول كل طالب مسجل حاليًا. الطلاب هم الأفراد الموصوفون في مجموعة البيانات. لكل فرد، تحتوي البيانات على قيم المتغيرات مثل تاريخ الميلاد واختيار التخصص ومتوسط الدرجة (GPA). من الناحية العملية، فإن أي مجموعة من البيانات تكون مصحوبة بمعلومات أساسية تساعدنا على فهم البيانات. عندما تخطط لدراسة إحصائية أو تستكشف بيانات من عمل شخص آخر، ضع في اعتبارك الأسئلة التالية (Moore & Notz, 2021, pp. 94-95):

- من؟ ما الأفراد الذين تصفهم البيانات؟ كم عدد الأفراد الذين يظهرون في البيانات؟
- ماذا؟ كم عدد المتغيرات التي تحتويها البيانات؟ ما هي التعريفات الدقيقة لهذه المتغيرات؟ بأي وحدة قياس يتم تسجيل كل متغير؟ يمكن تسجيل الأوزان، على سبيل المثال، بالجنيه أو بآلاف الجنيهات أو بالكيلوجرامات.
- أين؟ ستختلف درجات الطالب من كلية إلى أخرى، اعتمادًا على العديد من المتغيرات، بما في ذلك "انتقائية" القبول للكلية.
- متى؟ يتغير الطلاب من سنة إلى أخرى، وكذلك الأسعار والرواتب وما إلى ذلك.

---

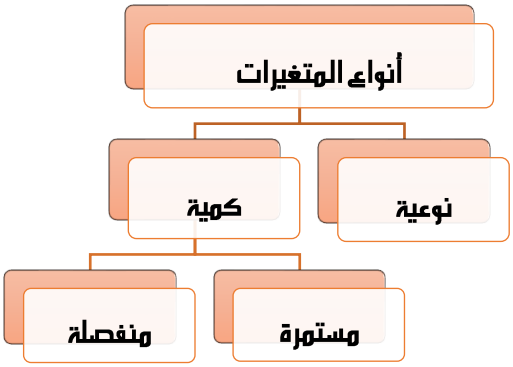
[1] الأفراد هم الكائنات الموصوفة بواسطة مجموعة من البيانات. قد يكون الأفراد بشرًا، لكنهم قد يكونوا أيضًا حيوانات أو أشياء.

[2] المتغير هو أي خاصية للفرد. يمكن أن يأخذ المتغير قيمًا مختلفة لأفراد مختلفين.

- لماذا؟ ما الغرض من البيانات؟ هل نأمل أن نجيب على بعض الأسئلة المحددة؟ هل نريد إجابات لهؤلاء الأفراد فقط أم لمجموعة أكبر من المفترض أن يمثلها هؤلاء الأفراد؟ هل الأفراد والمتغيرات مناسبة للغرض المقصود؟

عندما تكون قيم المتغير هي ببساطة أسماء الفئات، فإننا نسميها متغيرًا فئويًا أو نوعيًا. عندما يتم قياس قيم المتغير بالكميات العددية، فإننا نسميها متغيرًا كميًا. يبدو التمييز بين المتغيرات الفئوية والكمية واضحًا، ولكن هناك أسباب لتوخي الحذر. يمكن اعتبار بعض المتغيرات إما فئوية أو كمية، اعتمادًا على نوع الأسئلة التي نطرحها عنها. على سبيل المثال، سيتم اعتبار العمر المتغير كميًا إذا كانت الردود عددية ولديها وحدات. من ناحية أخرى، قد يهتم البعض بجمع القيم معًا في فئات مثل "طفل (12 عامًا أو أقل)" أو "مراهق (من 13 إلى 19 عامًا)" أو "بالغ (من 20 إلى 64 عامًا)" أو "كبير (65 عامًا أو أكثر). يمكن تلخيص تصنيف المتغيرات على النحو التالي:

الشكل (02): أنواع المتغيرات



المصدر: من اعداد المؤلف

يمكن تصنيف المتغيرات الكمية إلى مجموعتين: منفصلة ومستمرة. يمكن تعيين قيم متغيرات منفصلة مثل 0، 1، 2، 3 ويقال إنها قابلة للعد. من الأمثلة على المتغيرات المنفصلة عدد الأطفال في الأسرة، وعدد الطلاب في الفصل، وعدد المكالمات التي يتلقاها مركز الاتصال كل يوم لمدة شهر. يمكن للمتغيرات المستمرة، من خلال المقارنة، أن تفترض عددًا لا حصر له من القيم في فاصل زمني بين أي قيمتين محددتين. درجة الحرارة، على سبيل المثال، هي متغير مستمر، حيث يمكن للمتغير أن يفترض عددًا لا حصر له من القيم بين أي درجتين من درجات الحرارة.

هناك طريقة شائعة أخرى لتصنيف البيانات وهي استخدام أربعة مستويات للقياس: الاسمي والترتيبي والفاصل الزمني والنسبة. عندما نطبق الإحصائيات على مشاكل حقيقية، يساعدنا مستوى قياس البيانات في تحديد الإجراء الذي يجب استخدامه (Triola, 2014, pp. 18-20):

- المقياس الاسمي (Nominal) يستخدم لتصنيف البيانات ولا تعطي لها أي ترتيب، يتميز المستوى الاسمي للقياس ببيانات تتكون من أسماء، مسميات أو فئات فقط. لا يمكن ترتيب البيانات في مخطط ترتيب (مثل منخفض إلى مرتفع).
- المقياس الترتيبي (Ordinal) يرث خاصية المقياس الاسمي وهي التصنيف بالإضافة الي الترتيب، ولكن الاختلافات (التي يتم الحصول عليها عن طريق الطرح) بين قيم البيانات إما لا يمكن تحديدها أو لا معنى لها.

- مقياس الفترة (Interval) يرث الخاصيتين السابقتين؛ التصنيف والترتيب بالإضافة الي انه يوضح الترتيب والاختلافات الدقيقة بين القيم، ولا قيمه للصفر مع هذا القياس.
- المقياس النسبي (Ratio) يرث المزايا الثلاثة السابقة؛ فهو يصنّف، يرتّب ويوضح المسافات بشكل متساوي، زد على ذلك، فهو يضيف قيمة حقيقة للصفر.

ويمكن تلخيص تصنيف البيانات وفق المستويات الأربعة في الجدول التالي:

الجدول (01):

مستوى القياس	الوصف المختصر	مثال
النسبي	هناك نقطة بداية صفرية طبيعية والنسب لها معنى.	الأطوال، المسافات، الأحجام
الفترة	الاختلافات ذات مغزى، لكن لا توجد نقطة انطلاق طبيعية ولا معنى للنسب.	درجات الحرارة فهرنهايت أو مئوية
الترتيبي	يمكن ترتيب البيانات، لكن الاختلافات لا معنى لها.	مقياس ليكارت الخماسي
الاسمي	الفئات فقط، لا يمكن ترتيب البيانات.	الجنس، لون العينين

المصدر: من اعداد المؤلف

2. أساليب جمع البيانات الإحصائية:

هناك العديد من مصادر البيانات. من السهل جدًا جمع بعض البيانات ولكنها قد لا تكون مفيدة جدًا. تتطلب البيانات الأخرى تخطيطًا دقيقًا وتحتاج إلى موظفين محترفين لجمعها. يمكن أن تكون أكثر فائدة. مهما كان المصدر، سيبدأ التحليل الإحصائي الجيد بدراسة متأنية لمصدر البيانات. يمكن الحصول على البيانات من المصادر الموجودة أو من الدراسات الاستقصائية والدراسات التجريبية المصممة لجمع بيانات جديدة (Anderson, et al., 2019, pp. 10-11):

- في بعض الحالات، تكون البيانات المطلوبة لتطبيق معين موجودة بالفعل. تحتفظ الشركات بمجموعة متنوعة من قواعد البيانات حول موظفيها وعملائها وعملياتها التجارية. يمكن عادةً الحصول على بيانات حول رواتب الموظفين وأعمارهم وسنوات الخبرة من سجلات الموظفين الداخلية. تحتوي السجلات الداخلية الأخرى على بيانات حول المبيعات ونفقات الإعلان وتكاليف التوزيع ومستويات المخزون وكميات الإنتاج. تحتفظ معظم الشركات أيضًا ببيانات مفصلة عن عملائها.
- في بعض الأحيان، لا تتوفر البيانات اللازمة لتطبيق معين من خلال المصادر الموجودة. في مثل هذه الحالات، يمكن الحصول على البيانات غالبًا عن طريق إجراء دراسة إحصائية. يمكن تصنيف الدراسات الإحصائية على أنها إما تجريبية أو قائمة على الملاحظة.

بمجرد تحديد نوع البيانات –كمية أو نوعية – المناسبة للإشكالية المطروحة، سنحتاج إلى جمع البيانات. بشكل عام، يمكن الحصول على البيانات بثلاث طرق مختلفة (McClave, Benson, & Sincich, 2018, pp. 38-39):

- **بيانات من مصدر منشور**، في بعض الأحيان، تكون مجموعة البيانات متوفرة في مصدر منشور، مثل كتاب أو مجلة أو صحيفة أو موقع ويب.
- تتضمن الطريقة الثانية لجمع البيانات إجراء **تجربة مصممة**، حيث يمارس الباحث رقابة صارمة على الوحدات (الأشخاص أو الأشياء أو الأحداث) في الدراسة. التجربة المصممة هي طريقة لجمع البيانات حيث يمارس الباحث سيطرة كاملة على خصائص الوحدات التجريبية التي تم أخذ عينات منها.
- أخيرًا، يمكن استخدام **الدراسات القائمة على الملاحظة** لجمع البيانات. في الدراسة القائمة على الملاحظة، يلاحظ الباحث الوحدات التجريبية في بيئتها الطبيعية ويسجل المتغير (المتغيرات) ذات الأهمية. الدراسة القائمة على الملاحظة هي طريقة لجمع البيانات حيث يتم ملاحظة الوحدات التجريبية التي تم أخذ عينات منها في بيئتها الطبيعية. لم يتم إجراء أي محاولة للتحكم في خصائص الوحدات التجريبية التي تم أخذ عينات منها. (تشمل الأمثلة استطلاعات الرأي والاستطلاعات).

3. المجتمع والعينات:

تبدأ معظم الدراسات الاستقصائية الإحصائية بمجموعة من الأرقام بشكل ما. فعند جمع بيانات لاستطلاع رأي، بيانات لفحص أرباح الشركة، فإن الاحتمالات لا حصر لها. أحد الخيارات هو أن نجمع كل المعلومات ذات الصلة. في الاستطلاع، هذا يعني أننا نسأل الجميع، أو أننا نفحص كل ربح لشركة ما. تتمثل مهمة الإحصائي بعد ذلك في إيجاد طريقة جيدة لتقديم الأرقام لتسهيل تفسير المحتويات للجميع. لكن في كثير من الحالات، قد لا يكون من العملي أو حتى من الممكن جمع كل المعلومات. في مثل هذه الحالات يجب علينا أخذ عينة. يعني هذا في استطلاع للرأي أننا نسأل جزءًا فقط من المجتمع، وفي المحاسبة قد نتحقق فقط من بعض الأرباح المختارة عشوائيًا. هذا يضع الإحصائي في موقف مختلف، يجب عليه فحص النتائج، ولكن بالإضافة إلى ذلك، الحكم على ما إذا كان يمكن تعميم التأثيرات داخل العينة على بقية المجتمع. ما مقدار الثقة التي يمكن أن نتمتع بها في الآثار التي نراها في العينة؟ المشكلة هي أن العناصر في العينة قد تختلف عن بقية المجتمع بطريقة منهجية. تسمى هذه الاختلافات بالتحيز في الاختيار. من المهم أن تضع في اعتبارك أن العينة يجب أن تمثل المجتمع. يجب أن يكون الاختيار عشوائيًا، ويجب أن نسعى لتجنب تأثير مفردات العينة على بعضهم البعض. يجب علينا، على سبيل المثال، ألا نسأل أعضاء مسيرة احتجاجية لأن هؤلاء الأشخاص قد يكون لديهم آراء غير نموذجية للمجتمع (Ubøe, 2017, pp. 1-3).

من المهم في الإحصاء التمييز بين العينة والمجتمع (Bluman, 2019, pp. 91-92):

- يتكون المجتمع من جميع الموضوعات (البشرية أو غير ذلك) التي تتم دراستها. عندما يتم جمع البيانات من كل موضوع في المجتمع، يطلق عليه تعداد. على سبيل المثال، كل 10 سنوات تجري الولايات المتحدة تعداداً سكانيًا. الغرض الأساسي من هذا التعداد هو تحديد توزيع المقاعد في مجلس النواب. في معظم الأحيان، بسبب التكلفة والوقت وحجم

السكان والمخاوف الطبية وما إلى ذلك، لا يمكن استخدام المجتمع بالكامل لإجراء دراسة إحصائية؛ لذلك، يستخدم الباحثون العينات.

- العينة هي مجموعة من الموضوعات المختارة من المجتمع. إذا تم اختيار موضوعات العينة بشكل صحيح، فيجب أن تمتلك في معظم الأحيان نفس الخصائص أو الخصائص المماثلة مثل الموضوعات في المجتمع. ومع ذلك، يقال إن المعلومات التي تم الحصول عليها من عينة إحصائية متحيزة إذا كانت نتائج عينة من المجتمع مختلفة جذريًا عن نتائج المجتمع. أيضًا، يُقال إن العينة متحيزة إذا كانت لا تمثل المجتمع الذي تم اختيارها منه.

يتم اختيار عينة عشوائية بسيطة من n من الموضوعات بحيث يكون لكل عينة ممكنة من نفس الحجم n نفس فرصة الاختيار. (غالبًا ما تسمى العينة العشوائية البسيطة عينة عشوائية). من خلال أخذ العينات العشوائية، نتوقع أن يتم تمثيل جميع مكونات المجتمع (تقريبًا) بشكل متناسب. يتم اختيار العينات العشوائية بعدة طرق مختلفة، بما في ذلك استخدام أجهزة الكمبيوتر لتوليد أرقام عشوائية. بالإضافة إلى العينات العشوائية البسيطة، فيما يلي بعض طرق أخذ العينات الأخرى المستخدمة بشكل شائع في المسوحات (Triola, 2014, pp. 24-26):

- العينة العشوائية المنتظمة (Systematic Random Sample)، يتم استخدامها عند دراسة المجتمعات المتجانسة والتي لا تتباين مفرداتها كثيراً وقد أطلق عليها مصطلح العينة المنتظمة بسبب انتظام المسافات ما بين المفردات المختارة من مجتمع الدراسة.
- العينة العشوائية الطبقية (Stratified Random Sample)، تستخدم في المجتمعات الغير متجانسة والتي يكون فيها التباين ما بين مفرداتها وفقاً لخواص معينة كالمستوى التعليمي والجنس وغيره. حيث يتم تقسيم المجتمع إلى مجموعتين فرعيتين مختلفتين على الأقل (أو طبقات) بحيث تشترك المفردات ضمن نفس المجموعة الفرعية في نفس الخصائص. ثم نسحب عينة من كل مجموعة فرعية (أو طبقة).
- العينة العشوائية العنقودية (Cluster Sample)، نقسم أولاً مناطق المجتمع إلى أقسام (أو مجموعات). ثم نختار بشكل عشوائي بعض هذه المجموعات، ونسحب عينات عشوائية بسيطة من تلك المجموعات المختارة.

تمارين وتطبيقات

1) ما هما النوعان الأساسيان من علم الإحصاء؟ قم بوصفهم بالتفصيل.

2) قم بتحديد بعض الأساليب المستخدمة في الإحصاء الوصفي.

3) اشرح طريقتين يترابط فيهما الإحصاء الوصفي مع الإحصاء الاستنتاجي.

4) ما هو الفرق بين الدلالة الإحصائية والدلالة العملية؟ هل يمكن أن يكون للدراسة الإحصائية دلالة إحصائية، لكن ليس لها دلالة عملية؟

5) يصنف أحد مواقع الويب 50 نوعًا مختلفًا من القهوة. تتضمن البيانات المتغيرات التالية: اسم القهوة، السعر، التصنيف الاجمالي (0 إلى 100)، التحميص (خفيف، أو متوسط، أو داكن)، النكهة، الرائحة، وتصنيفات الشكل (من 0 إلى 10). المطلوب:

- ✓ تحديد المتغيرات وقيمها الممكنة.
- ✓ تصنيف كل متغير الى فئوي (نوعي) أو كمي.
- ✓ تلخيص الخصائص الرئيسية لمجموعة البيانات.

6) يجمع مكتب الإحصاء الأمريكي قدرًا كبيرًا من المعلومات المتعلقة بالتعليم العالي. على سبيل المثال، يوفر المكتب جدولًا يتضمن المتغيرات التالية: الولاية، عدد الطلاب من الولاية الذين يلتحقون بالكلية، عدد الطلاب الذين يلتحقون بالكلية في ولايتهم الأصلية.

- ✓ تصنيف كل متغير الى فئوي (نوعي) أو كمي.
- ✓ اشرح كيف يمكنك استخدام كل متغير كمي لشرح شيء ما عن الولايات.
- ✓ ضع في اعتبارك متغيرًا محسوبًا على أنه عدد الطلاب في كل ولاية الذين يلتحقون بكلية في الولاية مقسومًا على إجمالي عدد الطلاب من نفس الولاية الذين يلتحقون بأي كلية. اشرح كيف يمكنك استخدام هذا المتغير لشرح شيء ما عن الولايات.

7) أفاد استطلاع للرأي أجرته مؤسسة بحثية على 1018 من البالغين أن 39٪ يؤمنون بالتطور.

- ✓ ما هي القيمة الدقيقة لـ 39٪ من 1018؟
- ✓ هل يمكن أن تكون النتيجة من السؤال السابق هي العدد الفعلي للبالغين الذين قالوا إنهم يؤمنون بالتطور؟ لما ولما لا؟
- ✓ ما هو العدد الفعلي للبالغين الذين قالوا إنهم يؤمنون بالتطور؟
- ✓ من بين 1018 مبحوثاً، قال 255 إنهم لا يؤمنون بالتطور. ما هي نسبة المبحوثين الذين قالوا إنهم لا يؤمنون بالتطور؟

8) قدمت وكالة أنباء مقالًا بعنوان رئيسي يفيد بأن حوادث المرور تسببت في مقتل 704 أشخاص في العام الماضي. وأشار المقال إلى أن هذا رقم قياسي جديد ويقارنه بـ 617 حالة وفاة لحوادث المرور في العام السابق له. تم تضمين بيانات أخرى حول تكرار الإصابات. ما هي القيمة المهمة التي لم يتم تضمينها؟ لماذا هي مهمة؟

9) أجرت مؤسسة اعلامية استطلاعًا من خلال مطالبة القراء بالاتصال في ردهم على هذا السؤال: "هل تؤيد تطوير أسلحة ذرية يمكن أن تقتل ملايين الأبرياء؟" حيث ردّ 20 قارئًا، وكانت إجاباتهم أن 87٪ أجابوا بـ "لا" بينما أجاب 13٪ بـ "نعم". حدد العيوب الرئيسية في هذا الاستطلاع.

10) يخطط باحث سوق لإجراء مسح للحاضرين في المؤتمر. يستخدم قائمة أسماء الحضور ويختار كل عشرين اسمًا. هل النتيجة عينة عشوائية بسيطة؟ لماذا ولماذا لا؟ بشكل عام، ما هي العينة العشوائية البسيطة؟

11) حدد نوع العينات (بسيطة، منتظمة، طبقية، عنقودية) المستخدمة عند الحصول على عينة من 1500 إجابة للمسح كما هو موصوف. ثم حدد ما إذا كان من المحتمل أن ينتج عن مخطط أخذ العينات عينة تمثل المجتمع من جميع البالغين:

- يتم تجميع قائمة كاملة بجميع 241472385 بالغًا، ويتم تحديد كل 150000 اسم، حتى الوصول إلى حجم العينة البالغ 1500؛
- تم تجميع قائمة كاملة بجميع 241472385 بالغًا، ويتم اختيار 1500 بالغ عشوائيًا من تلك القائمة؛
- منطقة المسح مقسمة إلى مناطق بها 100 بالغ في كل منطقة. ثم يتم اختيار 15 من هذه المناطق بشكل عشوائي، ويتم مسح جميع الأشخاص المائة في كل منطقة من تلك المناطق؛
- تنقسم منطقة المسح إلى 150 منطقة مع نفس العدد تقريبًا من البالغين في كل منطقة، ثم يتم اختيار 10 أشخاص عشوائيًا من كل منطقة من 150 منطقة؛
- يتم إرسال استطلاع بالبريد إلى 10000 بالغ تم اختيارهم عشوائيًا، ويتم استخدام 1500 إجابة.

الدرس الثالث:
الطرق البيانية والرقمية لوصف البيانات النوعية

1. الوصف الرقمي للبيانات النوعية:

عند وصف المشاهدات النوعية، نحدد الفئات بطريقة يمكن أن تقع فيها كل مشاهدة في فئة (أو قسم) واحدة وواحدة فقط. ثم يتم وصف مجموعة البيانات عدديًا عن طريق إعطاء عدد المشاهدات، أو نسبة العدد الإجمالي للمشاهدات، التي تقع في كل فئة من الفئات.

كما تطرقنا في المحور الأول، تنقسم البيانات النوعية الى بيانات اسمية وبيانات ترتيبية. تتكون البيانات الاسمية عادة من المشاهدات التي تمثل التسميات أو الأسماء؛ المعلومات المتعلقة بالجنس أو العرق على سبيل المثال. تعتبر البيانات الاسمية أقل أشكال البيانات تعقيدًا لأن كل ما يمكننا فعله بالبيانات هو تصنيفها. البيانات الترتيبية أقوى بمعنى أنه يمكننا تصنيف البيانات وترتيبها. تتضمن أمثلة البيانات الترتيبية تقييمات منتج أو أستاذ، حيث يمثل 1 أسوأ ويمثل 4 أفضل. من أجل تنظيم البيانات النوعية، غالبًا ما يكون من المفيد إنشاء توزيع تكراري[1] (Jaggia & Kelly, 2020, p. 20).

لتوضيح بناء توزيع تكراري بالبيانات الاسمية، نستخدم الجدول رقم (02) الذي يبين الطقس لشهر ديسمبر (2020) في الجزائر العاصمة.

الجدول (02) طقس الجزائر العاصمة، ديسمبر 2020

الجمعة	السبت	الاحد	الاثنين	الثلاثاء	الاربعاء	الخميس
				1 غائم	2 ممطر	3 مشمس
4 غائم	5 غائم	6 غائم	7 غائم	8 ممطر	9 غائم	10 غائم
11 غائم	12 غائم	13 غائم	14 غائم	15 غائم	16 مشمس	17 غائم
18 غائم	19 غائم	20 غائم	21 غائم	22 مشمس	23 مشمس	24 مشمس
25 ممطر	26 غائم	27 مشمس	28 مشمس	29 غائم	30 غائم	31 مشمس

المصدر: www.wunderground.com

نلاحظ أولاً أن الطقس في الجزائر العاصمة يصنف على غائم أو ممطر أو مشمس. يسرد العمود الأول في الجدول (03) هذه الفئات. في البداية، نستخدم عمود "العد" لتسجيل عدد الأيام التي تقع ضمن كل فئة. نقوم بتحويل العدد الإجمالي لكل فئة إلى القيمة العددية الخاصة بها في عمود التكرار. نظرًا لظهور 3 علامات عد في الفئة الممطرة، فإننا نسجل القيمة 3 كتكرار لها، وهكذا لبقية الفئات. لاحظ أنه إذا جمعنا عمود التكرار، نحصل على حجم العينة.

[1] يقوم التوزيع التكراري للبيانات النوعية بتجميع البيانات في فئات ويسجل عدد المشاهدات التي تقع في كل فئة.

الجدول (03) التوزيع التكراري لطقس الجزائر العاصمة، ديسمبر 2020

الطقس	العد	التكرار
غائم	𝍸 𝍸 𝍸 𝍸	20
ممطر	𝍸	3
مشمس	𝍸 𝍸	8
		اجمالي = 31 يوم

المصدر: من اعداد المؤلف

من التوزيع التكراري، يمكننا الآن أن نلاحظ بسهولة أن أكثر أنواع الأيام شيوعًا في ديسمبر كانت غائمة لأن هذا النوع من اليوم يحدث بأعلى معدل تكرار.

حينما نرغب في مقارنة الطقس في مارس 2021 بالطقس في ديسمبر 2020، نقوم بتحويل التوزيع التكراري إلى توزيع تكراري نسبي[1]. نحسب التكرار النسبي لكل فئة بقسمة معدل تكرار كل فئة على العدد الإجمالي للمشاهدات. يجب أن يساوي مجموع التكرارات النسبية واحدًا، أو قيمة قريبة جدًا من واحد بسبب التقريب. في الجدول 4، نقوم بتحويل التوزيع التكراري من الجدول 3 إلى توزيع تكراري نسبي. وبالمثل، نحصل على التوزيع التكراري النسبي لشهر مارس 2021؛ حيث لا يتم عرض البيانات الأولية لشهر مارس 2021. كان لشهر مارس 2021، 14 يوماً غائماً و يوم واحد فقط ممطرًا و 16 أيام مشمساً.

الجدول (04) التوزيع التكراري لطقس الجزائر العاصمة، ديسمبر 2020

الطقس	التكرار النسبي مارس 2021	التكرار النسبي ديسمبر 2020
غائم	0.452 = 31/14	0.645 = 31/20
ممطر	0.032 = 31/01	0.097 = 31/03
مشمس	0.516 = 31/16	0.258 = 31/8
	اجمالي = 1	اجمالي = 1

المصدر: www.wunderground.com

يمكننا بسهولة تحويل الترددات النسبية إلى نسب مئوية عن طريق الضرب في 100. على سبيل المثال، النسبة المئوية للأيام الملبدة بالغيوم في ديسمبر 2020 ومارس 2021 تساوي 64.5٪ و45.2٪ على التوالي. من التوزيع التكراري النسبي، يمكننا الآن أن نستنتج أن الطقس في الجزائر العاصمة في شهر ديسمبر 2020 كان في الغالب غائمًا، بينما في شهر مارس 2021 كان في الغالب مشمساً. كما كان الطقس في شهر مارس 2021 أقل تساقطاً، حيث كان ما يقرب من 0.032٪ من الأيام ممطراً، مقابل 0.097٪ من أيام ديسمبر 2020.

[1] التكرار النسبي لكل فئة من المتغيرات النوعية يساوي نسبة (جزء) المشاهدات في كل فئة. يتم حساب التكرار النسبي للفئة بقسمة تكرارها على العدد الإجمالي للمشاهدات. يجب أن يساوي مجموع الترددات النسبية واحدًا. النسبة المئوية للتكرار لكل فئة من المتغيرات النوعية تساوي النسبة المئوية (٪) من المشاهدات في كل فئة؛ إنه يساوي التردد النسبي للفئة مضروبًا في 100.

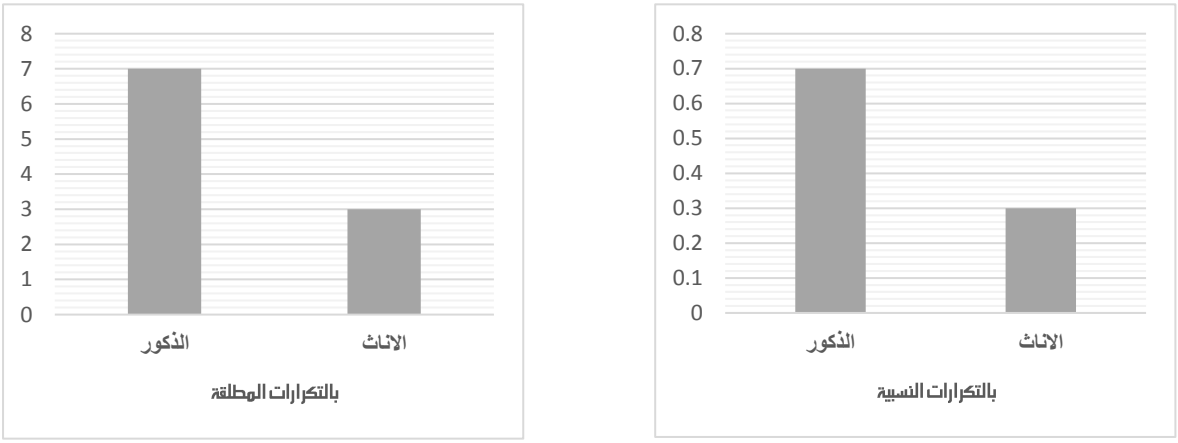
2. الرسوم البيانية الدائرية والمخططات الشريطية:

الرسم البياني الدائري والمخطط الشريطي هما تمثيلان بيانيان مستخدمان على نطاق واسع للبيانات النوعية. يعد المخطط الشريطي أداة بسيطة لتصور التكرارات النسبية أو المطلقة للقيم المشاهدة لمتغير.

يمكن استخدام المخطط الشريطي للمتغيرات الاسمية والترتيبية، طالما أن عدد الفئات ليس كبيرًا جدًا. يتكون من شريط واحد لكل فئة. يتم تحديد ارتفاع كل شريط إما بالتكرار المطلق أو التكرار النسبي للفئة المعنية ويظهر على المحور y. إذا تم قياس المتغير على مستوى ترتيبي، فمن المستحسن ترتيب الأعمدة على المحور x وفقًا لترتيبها أو قيمها. إذا كان عدد الفئات كبيرًا، فسيكون عدد الأشرطة كبيرًا أيضًا وقد لا يظل المخطط الشريطي بدوره مفيدًا (Heumann, Schomaker, & Shalabh, 2016, p. 24).

مثال: عشرة أشخاص يقفون في طابور في سوبر ماركت، إما ذكر (M) أو أنثى (F). التكرارات المطلقة للذكور والإناث هي $n_1 = 7$ و $n_2 = 3$ على التوالي. نظرًا لوجود فئتين، M وF، يلزم وجود شريطين لإنشاء المخطط – أحدهما لفئة الذكور والآخر لفئة الإناث. يتم تحديد ارتفاعات القضبان إما $n_1 = 7$ و $n_2 = 3$ أو $f_1 = 0.7$ و $f_2 = 0.3$. تظهر هذه الرسوم البيانية في الشكل رقم (03).

الشكل (03) المخطط الشريطي

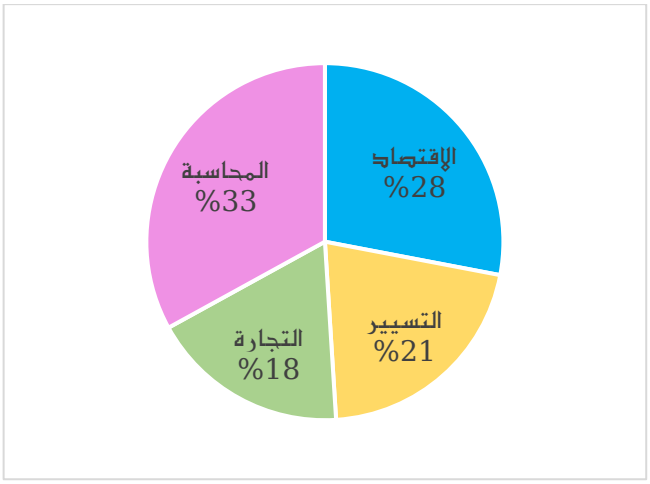


المصدر: من اعداد المؤلف

يوضح الرسم البياني الدائري كيف تنقسم المجموعة بأكملها إلى عدة فئات. تُظهر الرسوم البيانية الدائرية جميع الحالات كدائرة مقسمة إلى أجزاء تتناسب مساحتها مع جزء الحالات في كل فئة. يتم تمثيل فئات (أقسام) المتغير النوعي بشرائح دائرية (دائرة). يتناسب حجم كل شريحة مع التردد النسبي للفئة (Sharpe, De Veaux, & Velleman, 2021, p. 59).

مثال 1: يمكن تمثيل نسب تخصصات الطلبة في كلية العلوم الاقتصادية في إحدى الجامعات الجزائرية، في رسم بياني دائري كما في الشكل رقم (04).

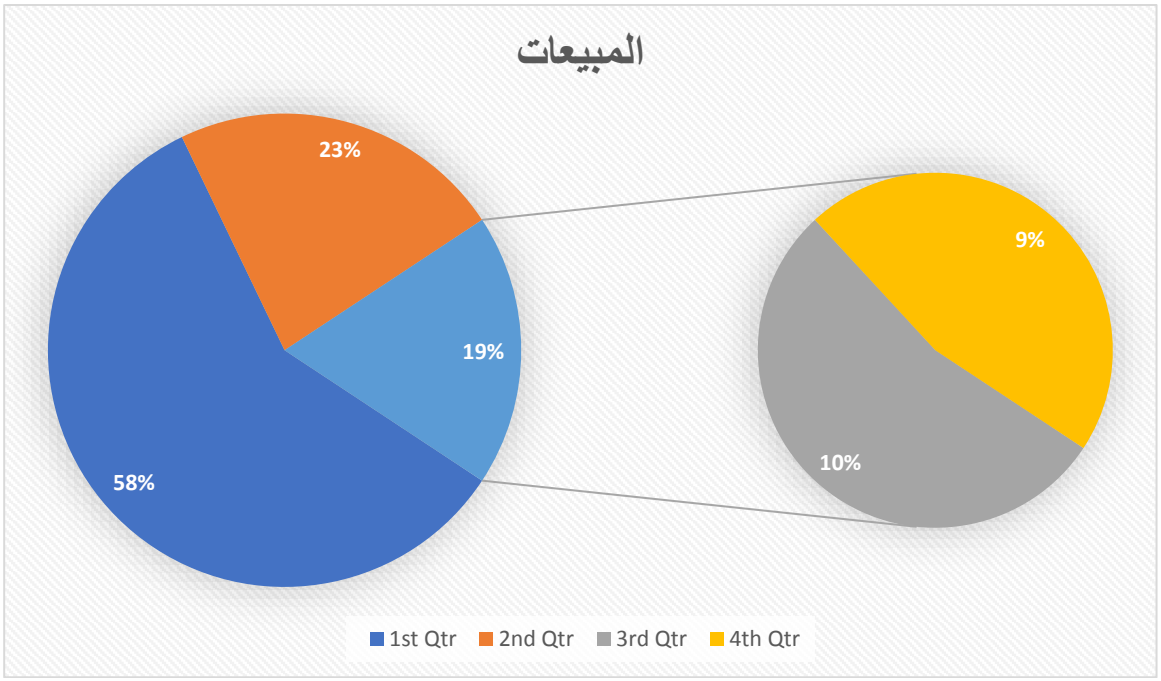
الشكل (04) الرسم البياني الدائري



المصدر: من اعداد المؤلف

مثال 2: يمكن تمثيل المبيعات السنوية لاحدى الشركات مقسمة على الثلاثيات الاربعة، في رسم بياني من دائرتين كما في الشكل رقم (05).

الشكل (05) الرسم البياني بدائرتين



المصدر: من اعداد المؤلف

الدرس الرابع:
الطرق البيانية والرقمية لوصف البيانات الكمية

1. الطرق الرقمية لوصف البيانات الكمية:

مع البيانات الكمية، تمثل كل مشاهدة كمية أو عددًا ذا مغزى. عدد براءات الاختراع التي تمتلكها شركات الأدوية (العدد) ودخل الأسرة (المبلغ) أمثلة على البيانات الكمية. على الرغم من اختلاف طبيعتنا عن البيانات النوعية، فإننا نستخدم التوزيع التكراري لتلخيص البيانات الكمية (Jaggia & Kelly, 2020, p. 27).

لتنظيم البيانات الكمية، نقوم أولاً بتجميع المشاهدات في فئات، ثم نتعامل مع الفئات باعتبارها القيم المميزة لبيانات نوعية. وبالتالي، بمجرد تجميع البيانات الكمية في فئات، يمكننا إنشاء توزيعات التكرار والتكرار النسبي للبيانات بنفس الطريقة تمامًا كما فعلنا مع البيانات النوعية. ثلاثة مبادئ توجيهية مهمة لتجميع البيانات الكمية في فئات هي (Weiss, 2016, p. 52):

- يجب أن يكون عدد الفئات صغيرًا بما يكفي لتقديم ملخص فعال ولكنه كبير بما يكفي لعرض الخصائص ذات الصلة بالبيانات. القاعدة العامة هي أن عدد الفئات يجب أن يكون بين 5 و20.
- يجب أن تنتمي كل مشاهدة إلى فئة واحدة، وواحدة فقط. أي أن كل مشاهدة يجب أن تنتمي إلى فئة معينة ولا يجب أن تنتمي أي مشاهدة إلى أكثر من فئة واحدة.
- كلما كان ذلك ممكنًا، يجب أن يكون لجميع الفئات نفس العرض. بشكل تقريبي، يعني هذا المبدأ التوجيهي، إذا أمكن، أن تغطي جميع الفئات نفس العدد من القيم الممكنة.

2. عرض الساق والورقة:

مخطط الساق والورقة هو نوع من الجداول الأنسب لتمثيل كميات صغيرة من البيانات المنفصلة. يظهر الرقم الأخير من كل قيمة بيانات على هيئة ورقة مرفقة بجميع الأرقام الأخرى، والتي تظهر في الجذع. يتم ترتيب الأرقام الموجودة في الساق عموديًا، ويتم ترتيب الأرقام الموجودة على الأوراق بشكل أفقي، مع وضع أصغر رقم بالقرب من الساق. يشكل كل صف في الجدول فئة من القيم. يجب أن تحتوي الصفوف على فترات عرض متساوية للسماح بإجراء مقارنة مرئية سهلة لمجموعات البيانات. يجب تضمين مفتاح بالوحدة المناسبة لشرح ما تمثله القيم في الرسم التخطيطي (Chalmers, 2018, p. 3).

لإنجاز مخطط الساق والورقة:

- افصل كل مشاهدة في جذع يتكون من جميع الأرقام باستثناء الرقم الأخير (أقصى اليمين) وورقة، الرقم الأخير. قد تحتوي السيقان على عدد من الأرقام حسب الحاجة، لكن كل ورقة تحتوي على رقم واحد فقط.
- اكتب السيقان في عمود رأسي مع أصغرها في الأعلى، وارسم خطًا رأسيًا على يمين هذا العمود.
- اكتب كل ورقة في الصف على يمين ساقها بترتيب تصاعدي من الساق.

مثال: نعتبر النسبة المئوية لعلامات 15 طالبًا في امتحان الاحصاء، الواردة في القائمة التالية: 58، 55، 58، 61، 72، 79، 97، 67، 61، 77، 92، 64، 69، 62 و53.

لتقديم البيانات في مخطط الاق والورقة، نقوم أولاً بتجميع العلامات في فئات مناسبة متساوية العرض. عرض 10 مناسب هنا، كما هو موضح أدناه:

5	8 5 8 3
6	1 7 1 4 9 2
7	2 9 7
8	
9	7 2

بعد ذلك، نرتب العلامات في كل صف بترتيب تصاعدي من اليسار إلى اليمين ونضيف مفتاحًا لإنتاج مخطط الساق والورقة الموضح أدناه:

5	3 5 8 8	Key: **5** | 3
6	1 1 2 4 7 9	represents
7	2 7 9	a score of 53%
8		
9	2 7	

3. المدرج التكراري:

العرض الرسومي الشائع للبيانات الكمية هو المدرج التكراري. يمكن إعداد هذا الملخص الرسومي للبيانات التي تم تلخيصها مسبقًا إما في التكرار أو التكرار النسبي أو النسبة المئوية لتوزيع التكرار. يتم إنشاء المدرج التكراري عن طريق وضع متغير الدراسة المستهدف على المحور الأفقي والتكرار أو التكرار النسبي أو النسبة المئوية للتكرار على المحور العمودي. يتم عرض التكرار أو التكرار النسبي أو النسبة المئوية للتكرار لكل فئة عن طريق رسم مستطيل يتم تحديد قاعدته بواسطة حدود الفئة على المحور الأفقي والذي يكون ارتفاعه هو التكرار المقابل أو التكرار النسبي أو النسبة المئوية للتكرار (Anderson, Sweeney, & Williams, 2008, p. 37). يتطلب انجاز مدرج تكراري يدويًا مزيدًا من العمل مقارنةً بالرسم البياني الساق والورقة. بالإضافة إلى ذلك، لا تعرض المدرجات التكرارية القيم الفعلية المشاهدة. لهذه الأسباب، تفضل مخططات الساق والورقة لمجموعات البيانات الصغيرة (Moore, McCabe, & Craig, 2021, p. 125).

مثال: أدناه التوزيع التكراري للأرباح على مبيعات السيارات لفترة معينة لإحدى الشركات.

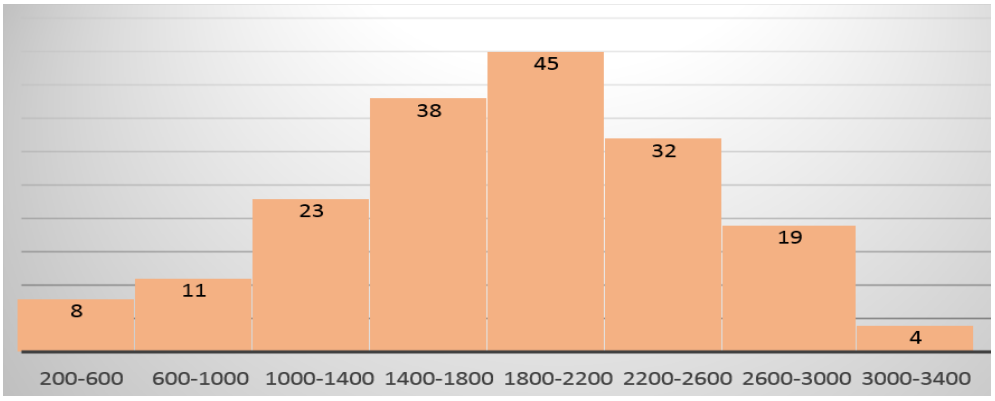
الجدول (05) التوزيع التكراري للأرباح على مبيعات السيارات

الارباح	التكرار
600-200	8
1000-600	11
1400-1000	23
1800-1400	38
2200-1800	45
2600-2200	32
3000-2600	19
3400-3000	4

المصدر: من اعداد المؤلف

نقوم بإنشاء المدرج التكراري، كما في الشكل الموالي:

الشكل (06) المدرج التكراري



المصدر: من اعداد المؤلف

تمارين وتطبيقات

1) اشرح الخطأ في كل مما يلي:

- يمكن استخدام مخطط الساق والورقة لعرض توزيع متغير فئوي.
- يمكن أن ينحرف التوزيع المتماثل جهة اليمين.
- تجاهل دائمًا القيم المتطرفة قبل إجراء تحليل لمجموعة من البيانات.

2) لكل من السيناريوهات التالية، حدد العرض البياني (شكل الدائرة، رسم بياني شريطي، مخطط الساق والورقة، أو مدرج تكراري) الذي ستستخدمه لوصف توزيع المتغير. قدم سببًا لاختيارك، وإذا كان هناك خيار بديل سيكون معقولًا أيضًا، فأشرح لماذا كان اختيارك أفضل من البديل:

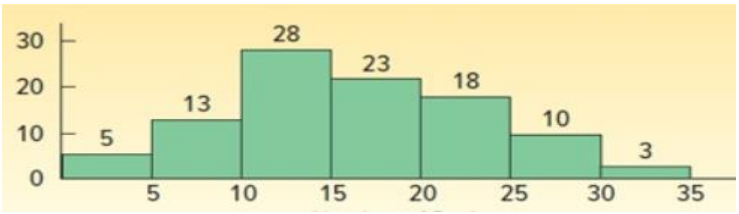
- عدد الدقائق التي قضيتها في النوم في كل يوم من الأيام السبعة في الأسبوع الماضي.
- علامات الامتحان الأول في الإحصاء لعدد 120 طالبًا وطالبة.

- اللون المفضل لكل طالب في تخصص الإحصاء.
- عدد طلاب البكالوريا لكل مدرسة ثانوية في ولاية الوادي.

3) كجزء من مجموعة الموارد البشرية لشركتك، يُطلب منك تلخيص المستويات التعليمية لـ 512 موظفًا في قسمك. من سجلات الشركة، وجدت أن 164 منهم ليس لديهم شهادة جامعية (لا شيء)، و42 منهم حاصلون على درجة تقني سامي، و225 حاصلون على درجة الليسانس، و52 حاصلين على درجة الماجستير، و29 حاصلين على درجة الدكتوراه. بالنسبة للمستوى التعليمي لهؤلاء الموظفين:

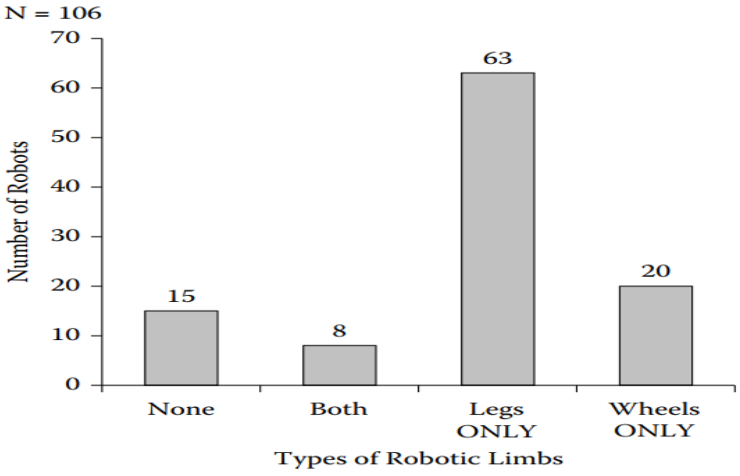
- قم بإعداد جدول تكراري.
- قم بإعداد جدول تكراري نسبي.
- قم برسم مخطط شريطي.
- قم برسم مخطط شريطي للتكرارات النسبية.
- قم برسم شكل الدائرة.

4) يوجد لدى شركة العديد من متاجر البيع بالتجزئة في المناطق المختلفة للوطن. يطلب منها العديد من العملاء شحن مشترياتهم. يوضح الرسم البياني التالي عدد الطرود التي تم شحنها يوميًا لآخر 100 يوم. على سبيل المثال، يُظهر المقطع الأول أنه خلال 5 أيام كان فيها عدد الطرود المشحونة من 0 إلى 5:



- ماذا يسمى هذا المخطط؟
- ما هو العدد الإجمالي للطرود المشحونة؟
- ما هو طول الفئة؟
- ما هو عدد الطرود المشحونة في فئة 10 حتى 15؟
- ما هو التكرار النسبي للطرود المشحونة في فئة 10 حتى 15؟
- في كم يومًا تم شحن 25 طردًا أو أكثر؟

5) في بحث علمي درس المهندسون الاتجاه في تصميم الروبوتات الاجتماعية. باستخدام عينة عشوائية من 106 روبوتًا اجتماعيًا تم الحصول عليها من خلال بحث على الويب، وجد المهندسون أن 63 منها بُنيت بأرجل فقط، و20 بعجلات فقط، و8 بأرجل وعجلات، و15 بلا أرجل ولا عجلات. هذه المعلومات مصورة في الشكل الموالي:



- ما نوع الرسم البياني المستخدم لوصف هذه البيانات؟
- حدد المتغير الذي تم قياسه لكل من تصميمات الروبوت البالغ عددها 106.
- استخدم الرسم البياني لتحديد تصميم الروبوت الاجتماعي الأكثر استخدامًا حاليًا.
- حساب التكرارات النسبية للفئات المختلفة الموضحة في الرسم البياني.

الدرس الخامس:

التحليل الاستكشافي للبيانات باستخدام برنامج SPSS

1. تمهيد.

المرحلة الأولى في أي تحليل للبيانات هي استكشاف البيانات التي تم جمعها. عادة ما نهتم بالبحث في الإحصاء الوصفي مثل المتوسط، المنوال، الوسيط، والتكرار وما إلى ذلك. في كثير من الأحيان، نحن مهتمون بفحص افتراضات البيانات أيضًا (للتذكير أن الاختبارات البارامترية تتطلب بيانات موزعة بشكل طبيعي ولذا فإننا نرغب غالبًا في تقييم الدرجة التي تكون فيها البيانات طبيعية.

2. الإحصاء الوصفي Frequencies and Descriptives:

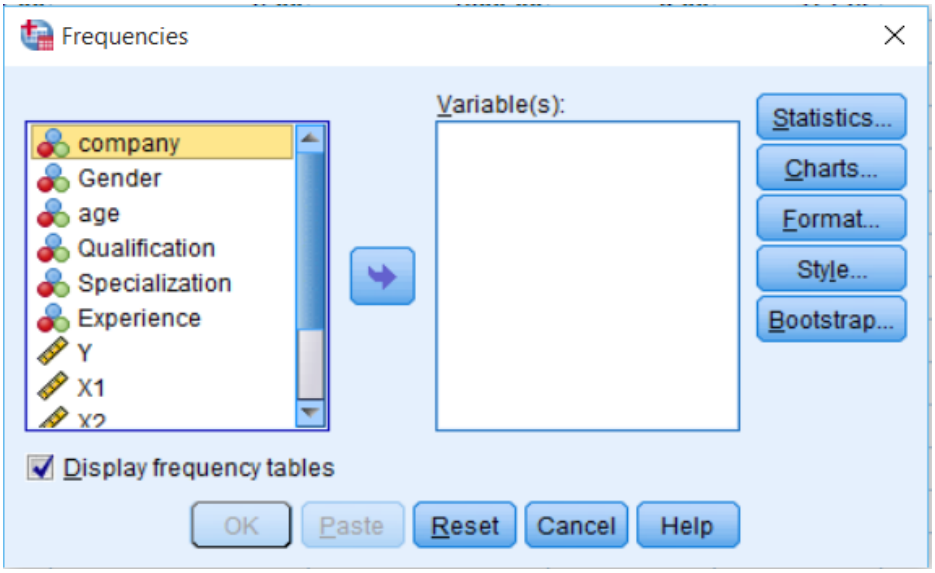
1.2 الأمر Frequencies:

يمكن استخدام أمر Frequencies لتحديد الربيعيات، والمؤينيات، ومقاييس النزعة المركزية (المتوسط، الوسيط، والمنوال)، ومقاييس التشتت (المدى، الانحراف المعياري، التباين، الحد الأدنى والحد الأقصى)، ومقاييس التفرطح والالتواء، وإنشاء الرسوم البيانية.

ولاستخراج التكرارات والنسب المئوية للمتغيرات نتبع الخطوات التالية:

أ- ننقر فوق القائمة Analyze، ثم ننقر على Descriptive Statistics ثم Frequencies، فنتحصل على الشكل التالي:

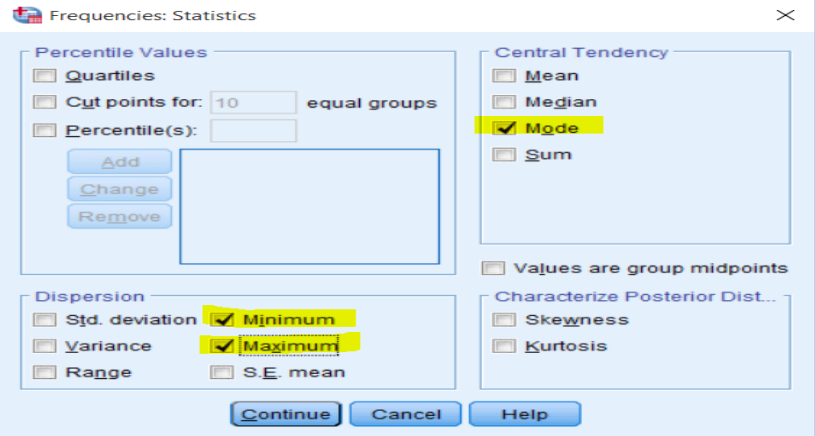
الشكل رقم (59): مربع الحوار Frequencies



المصدر: مخرجات SPSS

ب- نختار المتغيرات المراد استكشاف بياناتها بالنقر عليها مرة واحدة ثم النقر على السهم (او بالنقر المزدوج)، فتتحول الى مربع Variable(s)، ثم ننقر فوق Statistics ليظهر مربع الحوار المبين في الشكل التالي:

الشكل رقم (60): مربع الحوار Frequencies: Statistics



المصدر: مخرجات SPSS

ت- نختار الإحصـــاءات التي نريدها، في هذه الحالة ســـوف نختار المنوال، الحد الأدنى والحد الأقصـــى، ثم ننقر على Continue، ثم ننقر فوق OK، سوف تظهر النتائج في شاشة مستعرض النتائج Output Viewer، كما هو موضح في الجداول الموالية:

الجدول رقم (06): مخرجات نتائج الإحصاءات الوصفية

		الجنس	التخصص
N	Valid	47	47
	Missing	0	0
Mode		1.00	2.00[a]
Minimum		1.00	1.00
Maximum		2.00	7.00

a. Multiple modes exist. The smallest value is shown

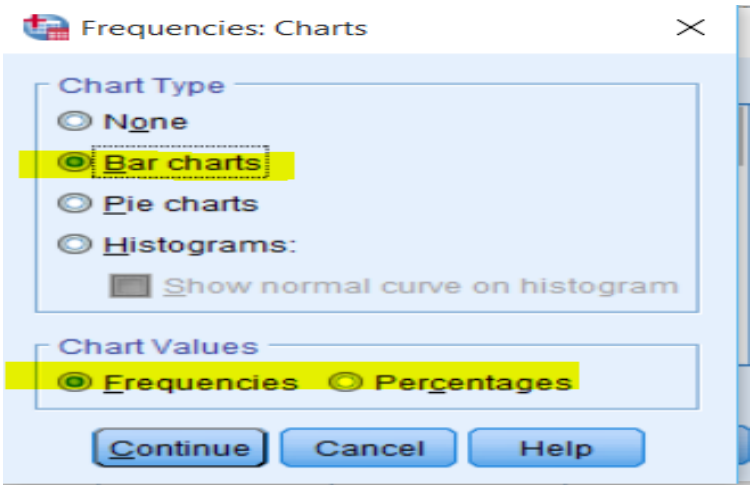
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ذكر	34	72.3	72.3	72.3
	أنثى	13	27.7	27.7	100.0
	Total	47	100.0	100.0	

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	محاسبة	9	19.1	19.1	19.1
	مالية	14	29.8	29.8	48.9
	اقتصاد	14	29.8	29.8	78.7
	موارد بشرية	3	6.4	6.4	85.1
	ادارة اعمال	3	6.4	6.4	91.5
	اخرى	4	8.5	8.5	100.0
	Total	47	100.0	100.0	

المصدر: مخرجات SPSS

ث- لتمثيل النتائج السابقة بيانيا، غالبا ما نستخدم الرسوم البيانية Bar chart وPie chart للمتغيرات النوعية، بينما نستخدم Histograms للمتغيرات الكمية. على سبيل المثال ولإنشاء رسم بياني من نوع Bar chart ننقر فوق القائمة Analyze، ثم ننقر على Descriptive Statistics ثم Frequencies، ونختار المتغير (أو المتغيرات) التي نرغب في انشاء رسوم بيانية لها وننقر على السهم لتحويلها الى مربع Variable(s)، ثم ننقر فوق Charts ليظهر مربع الحوار المبين في الشكل التالي:

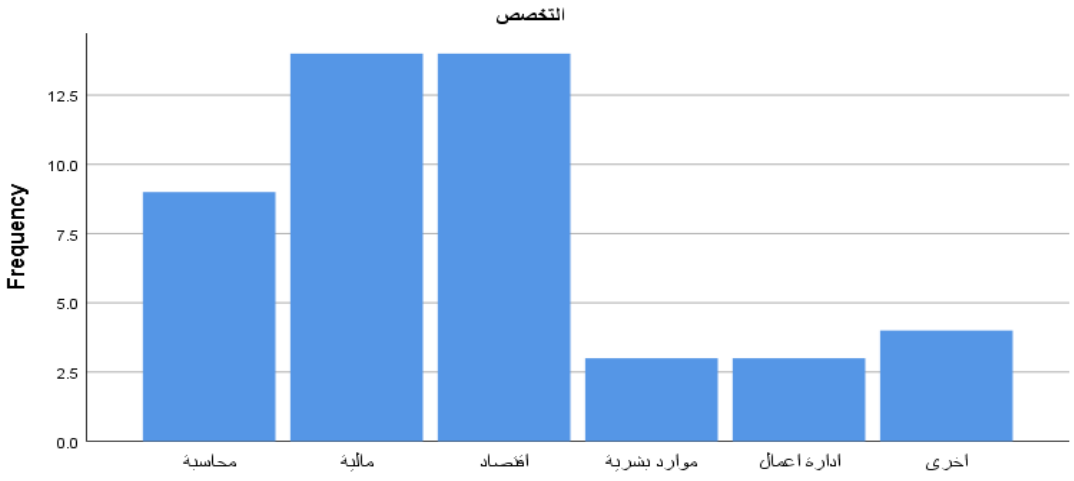
الشكل رقم (61): مربع الحوار Frequencies: Charts1



المصدر: مخرجات SPSS

ج- نختار Bar charts من مربع chart Type، مع إمكانية الاختيار بين التكرار او النسب المئوية من مربع chart Values. ثم ننقر فوق Continue، ثم ننقر فوق OK، فنتحصل الى الرسم البياني التالي:

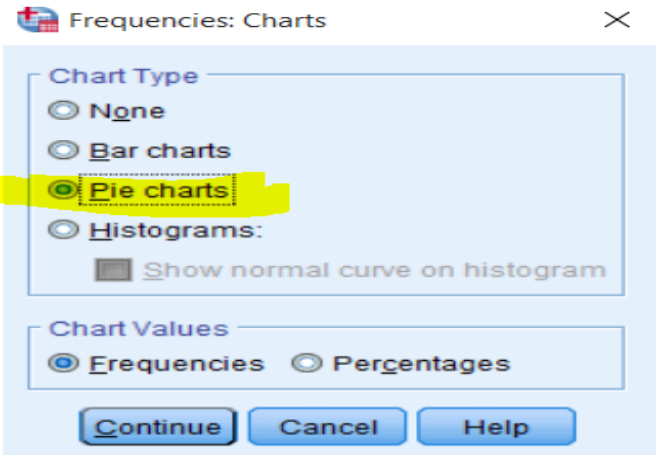
الشكل رقم (62): مخرجات الرسم البياني Bar charts



المصدر: مخرجات SPSS

ح- ولإنشاء رسم بياني من نوع Pie charts ننقر فوق القائمة Analyze، ثم ننقر على Descriptive Statistics ثم Frequencies، ونختار المتغير (أو المتغيرات) التي نرغب في انشاء رسوم بيانية لها وننقر على السهم لتحويلها الى مربع Variable(s)، ثم ننقر فوق Charts ليظهر مربع الحوار المبين في الشكل التالي:

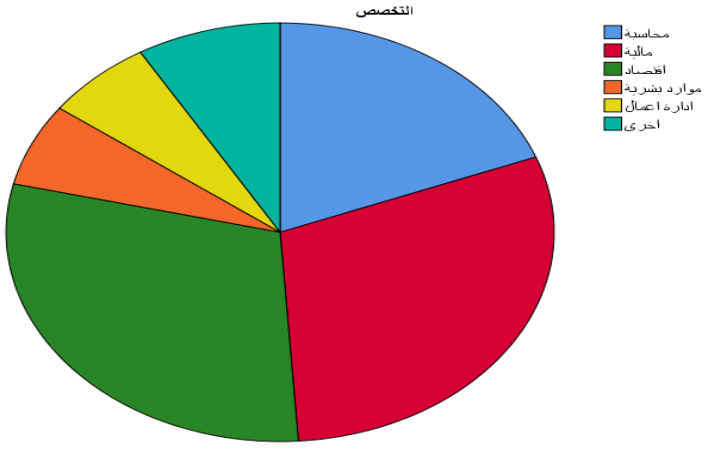
الشكل رقم (63): مربع الحوار Charts2 Frequencies:



المصدر: مخرجات SPSS

خ- نختار Pie charts من مربع chart Type، مع إمكانية الاختيار بين التكرار او النســـب المئوية من مربع chart Values. ثم ننقر فوق Continue، ثم ننقر فوق OK، فنتحصل الى الرسم البياني التالي:

الشكل رقم (64): مخرجات الرسم البياني Pie charts



المصدر: مخرجات SPSS

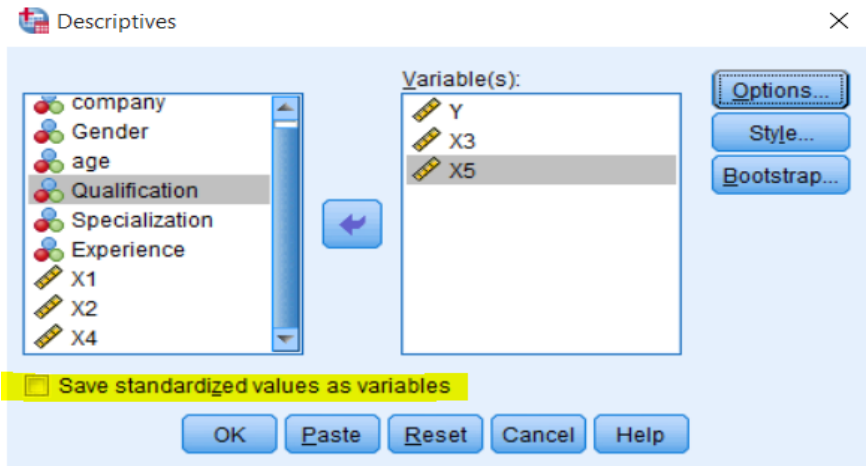
2.2 الأمر Descriptives:

يمكن استخدام أمر Descriptives لتحديد مقاييس النزعة المركزية (المتوسط، الوسيط، والمنوال)، ومقاييس التشتت (المدى، الانحراف المعياري، التباين، الحد الأدنى والحد الأقصى)، ومقاييس التفرطح والالتواء. وتستخدم أكثر مع المتغيرات الكمية.

ولاستخراج بعض مقاييس النزعة المركزية وبعض مقاييس التشتت، لمتغيرات كمية مختارة نتبع الخطوات التالية:

د- ننقر فوق القائمة Analyze، ثم ننقر على Descriptive Statistics ثم Descriptives، كما يمكن اختيار انشـــاء متغير جديد يحتوي على العلامات المعيارية المقابلة لكل مفردة من العينة، وذلك بالنقر على Save standardized values as variables فنتحصل على الشكل التالي:

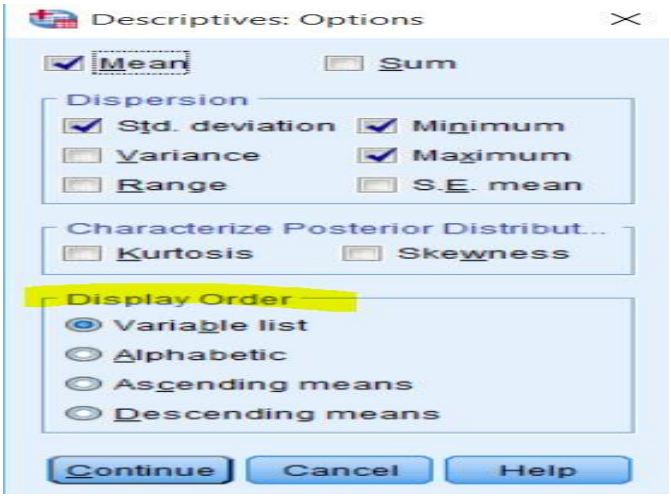
الشكل رقم (65): مربع الحوار Descriptives



المصدر: مخرجات SPSS

ذ- نختار المتغيرات المراد استكشاف بياناتها بالنقر عليها مرة واحدة ثم النقر على السهم (او بالنقر المزدوج)، فتتحول الى مربع Variable(s)، ثم ننقر فوق Options ليظهر مربع الحوار المبين في الشكل التالي:

الشكل رقم (66): مربع الحوار Descriptives: Options



المصدر: مخرجات SPSS

ر- نختار الإحصاءات التي نريدها، ومن مربع Display Order يمكننا ان نختار طريقة ترتيب النتائج من خلال أربعة خيارات: حســب تواجدها في القائمة Variable list، حســب ترتيبها الهجائي Alphabetic، حســب قيم

موسطاتها تصاعديا Ascending means، حسب قيم موسطاتها تنازليا Descending means. بعد ذلك ننقر على Continue، ثم ننقر فوق OK، سوف تظهر النتائج في شاشة مستعرض النتائج Output Viewer، كما هو موضح في الجداول الموالية:

الجدول رقم (07): مخرجات Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Y	32	180.00	5750.00	1154.3594	1001.79975
X3	32	50.34	532.22	137.8147	92.87726
X5	32	418.00	5635.00	1422.9219	1029.88430
Valid N (listwise)	32				

المصدر: مخرجات SPSS

3. استخدام الاجراء الاحصائي Explore:

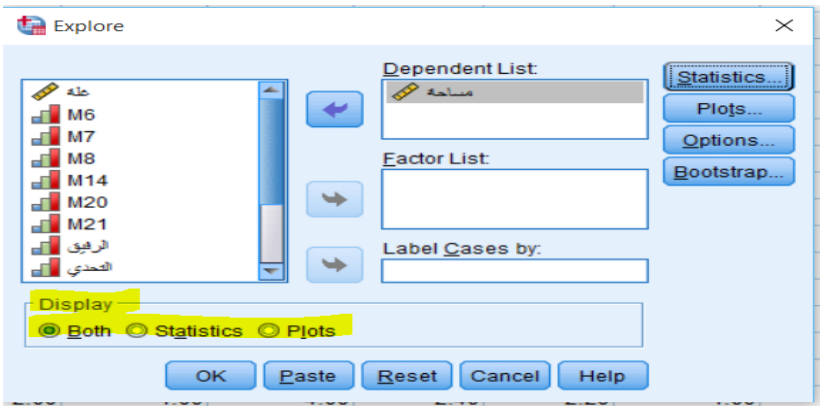
يمكن استخدام أمر Explore لتحديد مقاييس النزعة المركزية (المتوسط والوسيط)، ومقاييس التشتت (المدى، المدى الربيعي، الانحراف المعياري، التباين، والحد الأدنى والحد الأقصى)، ومقاييس التفرطح والالتواء، وإعداد الرسوم البيانية Histograms، وشكل الساق والورقة stem and leaf plots، وشكل الصندوق Tukey box plots. كما يستخدم هذا الاجراء كذلك من التوزيع الطبيعي للمتغير Test of Normality، كذلك اجراء اختبار تجانس التباينات Homogeneity of Variances.

1.3. حساب الإحصاءات الوصفية لمتغير كمي:

لاستخراج الإحصاءات الوصفية لمتغير كمي نتبع الخطوات التالية:

أ- ننقر فوق القائمة Analyze، ثم ننقر على Descriptive Statistics ثم Explore، فنتحصل على شاشة الحوار التالية:

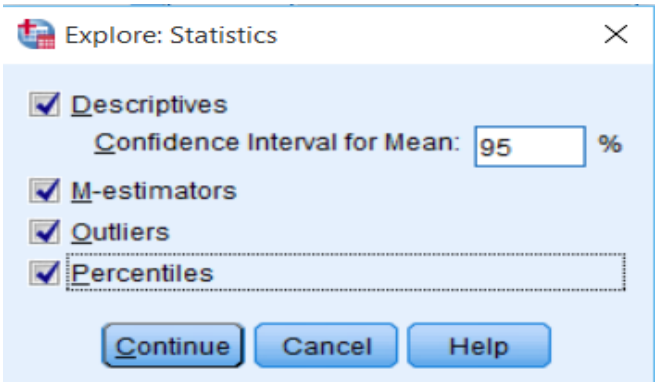
الشكل رقم (67): مربع الحوار Explore1



المصدر: مخرجات SPSS

ب- نختار المتغير المراد حساب الإحصاءات الوصفية له، بالنقر عليه مرة واحدة ثم النقر على السهم (او بالنقر المزدوج)، فيتحول الى مربع Dependent List. من مربع Display يمكننا الاختيار بين عرض الإحصاءات الوصفية فقط دون الرسوم البيانية عبر النقر فوق Statistics، أو عرض الرسوم البيانية فقط دون الإحصاءات الوصفية عبر النقر فوق Plots، أو عرض الإحصاءات الوصفية والرسوم البيانية معا عبر النقر فوق Both وهو ما سنختاره في مثالنا. ثم ننقر فوق Statistics ليظهر مربع الحوار المبين في الشكل التالي:

الشكل رقم (68): مربع الحوار Explore: Statistics1



المصدر: مخرجات SPSS

ت- نختار Descriptives، و M-estimators، و Outliers، وPercentiles. بعد ذلك ننقر على Continue، ثم ننقر فوق OK، سوف تظهر النتائج في شاشة مستعرض النتائج Output Viewer، كما هو موضح في الجداول الموالية:

الجدول رقم (08): مخرجات Explore: Statistics1

Descriptives

			Statistic	Std. Error
مساحة	Mean		4.1389	.41672
	95% Confidence Interval for Mean	Lower Bound	3.2929	
		Upper Bound	4.9849	
	5% Trimmed Mean		4.0432	
	Median		3.0000	
	Variance		6.252	
	Std. Deviation		2.50032	
	Minimum		.50	
	Maximum		9.50	
	Range		9.00	
	Interquartile Range		4.00	
	Skewness		.433	.393
	Kurtosis		-.852-	.768

المصدر: مخرجات SPSS

يبين الجدول نتائج الإحصاءات الوصفية مثل المتوسط الحسابي والوسيط والمتوسط المقطوع (Trimmed Mean)، وهو المتوسط الحسابي بعد حذف أعلى 5% واقل 5% من البيانات وذلك لإلغاء أثر القيم الشاذة ان وجدت. بالإضافة الى مقاييس التشتت التي تحتوي على الخطأ المعياري والانحراف المعياري والتباين واقل قيمة واعلى قيمة والمدى الربيعي، كما يظهر الإحصاءات الخاصة بشكل التوزيع كالالتواء Skewness والتفلطح Kurtosis.

الجدول رقم (09): مخرجات M–Estimators1

M-Estimators

	Huber's M-Estimator[a]	Tukey's Biweight[b]	Hampel's M-Estimator[c]	Andrews' Wave[d]
مساحة	3.9438	3.9519	4.0052	3.9531

a. The weighting constant is 1.339.

b. The weighting constant is 4.685.

c. The weighting constants are 1.700, 3.400, and 8.500

d. The weighting constant is 1.340*pi.

المصدر: مخرجات SPSS

يوضح الجدول تقديرات للمتوسطات التي لا تتأثر بالقيم الشاذة (كما في المتوسط المقطوع).

الجدول رقم (10): مخرجات Percentiles1

Percentiles

		Percentiles						
		5	10	25	50	75	90	95
Weighted Average(Definition 1)	مساحة	.9250	1.0000	2.0000	3.0000	6.0000	8.0000	9.0750
Tukey's Hinges	مساحة			2.0000	3.0000	6.0000		

المصدر: مخرجات SPSS

يوضح الجدول المئينيات، وهي القيم التي يقل عنها نسبة معينة من البيانات، مثلا المئين 20 هو القيمة التي يقل عنها 20% من البيانات.

الجدول رقم (09): مخرجات Extreme Values1

Extreme Values

			Case Number	Value
مساحة	Highest	1	18	9.50
		2	28	9.00
		3	10	8.00
		4	12	8.00
		5	3	7.00[a]
	Lowest	1	23	.50
		2	29	1.00
		3	24	1.00
		4	5	1.00
		5	4	1.00

a. Only a partial list of cases with the value 7.00 are shown in the table of upper extremes.

المصدر: مخرجات SPSS

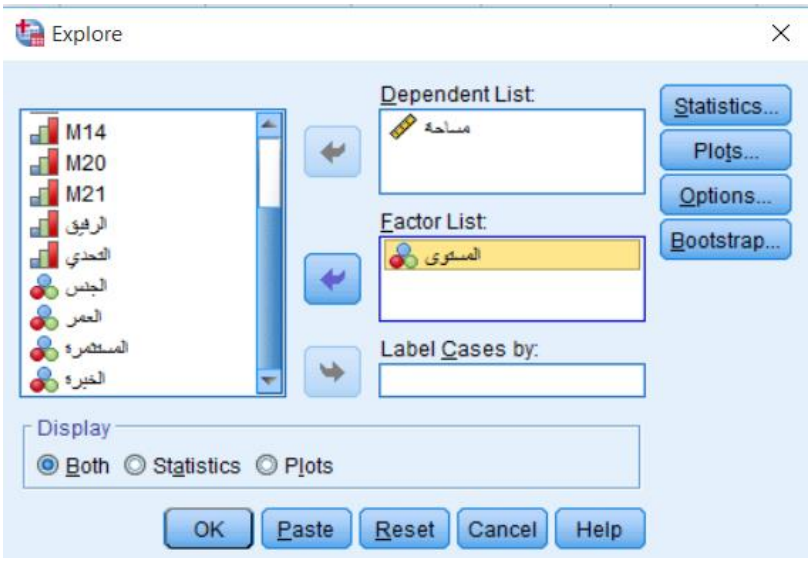
يوضح الجدول ما إذا كانت هناك قيم شاذة. واستخراج أكبر خمسة قيم واقل خمسة قيم شاذة، وذلك تمهيدا لحذفها من البيانات حتى لا تؤثر في الاختبارات الإحصائية الأخرى.

حساب الإحصاءات الوصفية لمتغير كمي حسب فئات متغير نوعي:

بالاعتماد على نفس المثال السابق، لاستخراج الإحصاءات الوصفية للمتغير الكمي "مساحة" حسب فئات المتغير النوعي "المستوى"، نتبع الخطوات التالية:

أ- ننقر فوق القائمة Analyze، ثم ننقر على Descriptive Statistics ثم Explore، فنتحصل على شاشة الحوار التالية:

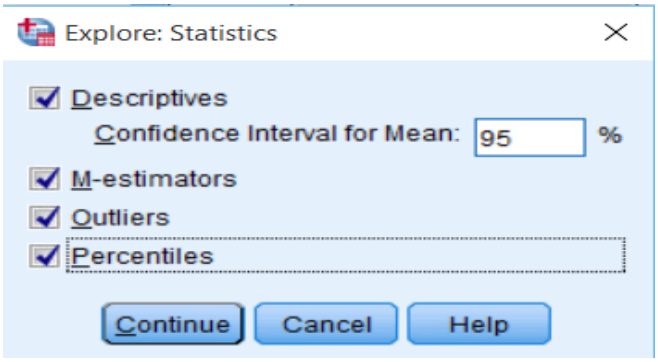
الشكل رقم (69): مربع الحوار Explore2



المصدر: مخرجات SPSS

ب- نختار المتغير الكمي المراد حساب الإحصاءات الوصفية له "مساحة"، بالنقر عليه مرة واحدة ثم النقر على السهم (او بالنقر المزدوج)، فيتحول الى مربع Dependent List. ونختار المتغير النوعي المراد التقسيم حسب فئاته "المستوى"، بالنقر عليه مرة واحدة ثم النقر على السهم (او بالنقر المزدوج)، فيتحول الى مربع Factor List. من مربع Display يمكننا الاختيار بين عرض الإحصاءات الوصفية فقط دون الرسوم البيانية عبر النقر فوق Statistics، أو عرض الرسوم البيانية فقط دون الإحصاءات الوصفية عبر النقر فوق Plots، أو عرض الإحصاءات الوصفية والرسوم البيانية معا عبر النقر فوق Both وهو ما سنختاره في مثالنا. ثم ننقر فوق Statistics ليظهر مربع الحوار المبين في الشكل التالي:

الشكل رقم (70): مربع الحوار Explore: Statistics2



المصدر: مخرجات SPSS

ت– نختار Descriptives، و M-estimators، و Outliers، وPercentiles. بعد ذلك ننقر على Continue، ثم ننقر فوق OK، سوف تظهر النتائج في شاشة مستعرض النتائج Output Viewer، كما هو موضح في الجداول الموالية:

الجدول رقم (11): مخرجات مخرجات Explore: Statistics1

		المستوى		Statistic	Std. Error
مساحة	متوسط	Mean		4.3421	.55284
		95% Confidence Interval for Mean	Lower Bound	3.1806	
			Upper Bound	5.5036	
		5% Trimmed Mean		4.3523	
		Median		5.0000	
		Variance		5.807	
		Std. Deviation		2.40978	
		Minimum		.50	
		Maximum		8.00	
		Range		7.50	
		Interquartile Range		4.00	
		Skewness		-.339-	.524
		Kurtosis		-1.359-	1.014
	ثانوي	Mean		4.0417	.74483
		95% Confidence Interval for Mean	Lower Bound	2.4023	
			Upper Bound	5.6810	
		5% Trimmed Mean		3.9074	
		Median		3.0000	
		Variance		6.657	
		Std. Deviation		2.58015	
		Minimum		1.00	
		Maximum		9.50	
		Range		8.50	
		Interquartile Range		3.50	
		Skewness		1.157	.637
		Kurtosis		.500	1.232
	جامعي	Mean		3.6000	1.39104
		95% Confidence Interval for Mean	Lower Bound	-.2622-	
			Upper Bound	7.4622	
		5% Trimmed Mean		3.4167	
		Median		3.0000	
		Variance		9.675	
		Std. Deviation		3.11047	
		Minimum		1.50	
		Maximum		9.00	
		Range		7.50	
		Interquartile Range		4.50	
		Skewness		1.918	.913
		Kurtosis		3.878	2.000

المصدر: مخرجات SPSS

يبين الجدول نتائج الإحصاءات الوصفية لكل فئة من فئات المتغير النوعي "المستوى" (وهي: متوسط، ثانوي، وجامعي)، مثل المتوسط الحسابي والوسيط والمتوسط المقطوع (Trimmed Mean)، وهو المتوسط الحسابي بعد حذف أعلى 5% واقل 5% من البيانات وذلك لإلغاء أثر القيم الشاذة ان وجدت. بالإضافة الى مقاييس التشتت التي تحتوي على الخطأ المعياري والانحراف المعياري والتباين واقل قيمة واعلى قيمة والمدى الربيعي، كما يظهر الإحصاءات الخاصة بشكل التوزيع كالالتواء Skewness والتفلطح Kurtosis.

الجدول رقم (12): مخرجات M–Estimators2

M-Estimators

	المستوى	Huber's M-Estimator[a]	Tukey's Biweight[b]	Hampel's M-Estimator[c]	Andrews' Wave[d]
مساحة	متوسط	4.5026	4.4713	4.3552	4.4718
	ثانوي	3.2837	2.8740	3.2387	2.8757
	جامعي	2.7522	2.2621	2.7963	2.2503

a. The weighting constant is 1.339.

b. The weighting constant is 4.685.

c. The weighting constants are 1.700, 3.400, and 8.500

d. The weighting constant is 1.340*pi.

المصدر: مخرجات SPSS

يوضح الجدول تقديرات للمتوسطات التي لا تتأثر بالقيم الشاذة (كما في المتوسط المقطوع)، لكل فئة من فئات المتغير النوعي "المستوى" (وهي: متوسط، ثانوي، وجامعي).

الجدول رقم (13): مخرجات Percentiles2

Percentiles

		المستوى	Percentiles 5	10	25	50	75	90	95
Weighted Average(Definition 1)	مساحة	متوسط	.5000	1.0000	2.0000	5.0000	6.0000	7.0000	.
		ثانوي	1.0000	1.3000	2.2500	3.0000	5.7500	9.0500	.
		جامعي	1.5000	1.5000	1.5000	3.0000	6.0000	.	.
Tukey's Hinges	مساحة	متوسط			2.2500	5.0000	6.0000		
		ثانوي			2.5000	3.0000	5.5000		
		جامعي			1.5000	3.0000	3.0000		

المصدر: مخرجات SPSS

يوضح الجدول المئينيات لكل فئة من فئات المتغير النوعي "المستوى" (وهي: متوسط، ثانوي، وجامعي)، وهي القيم التي يقل عنها نسبة معينة من البيانات، مثلا المئين 20 هو القيمة التي يقل عنها 20% من البيانات.

الجدول رقم (14): مخرجات Extreme Values2

Extreme Values[d]

	المستوى			Case Number	Value
مساحة	متوسط	Highest	1	10	8.00
			2	3	7.00
			3	21	7.00
			4	1	6.00
			5	6	6.00[a]
		Lowest	1	23	.50
			2	29	1.00
			3	5	1.00
			4	4	1.00
			5	22	2.00
	ثانوي	Highest	1	18	9.50
			2	12	8.00
			3	19	6.00
			4	32	5.00
			5	7	3.00[b]
		Lowest	1	24	1.00
			2	34	2.00
			3	25	2.00
			4	36	3.00
			5	35	3.00[c]
	جامعي	Highest	1	28	9.00
			2	11	3.00[b]
		Lowest	1	26	1.50
			2	17	1.50

المصدر: مخرجات SPSS

يوضح الجدول ما إذا كانت هناك قيم شاذة لكل فئة من فئات المتغير النوعي "المستوى" (وهي: متوسط، ثانوي، وجامعي). واستخراج أكبر خمسة قيم واقل خمسة قيم شاذة، وذلك تمهيدا لحذفها من البيانات حتى لا تؤثر في الاختبارات الإحصائية الأخرى.

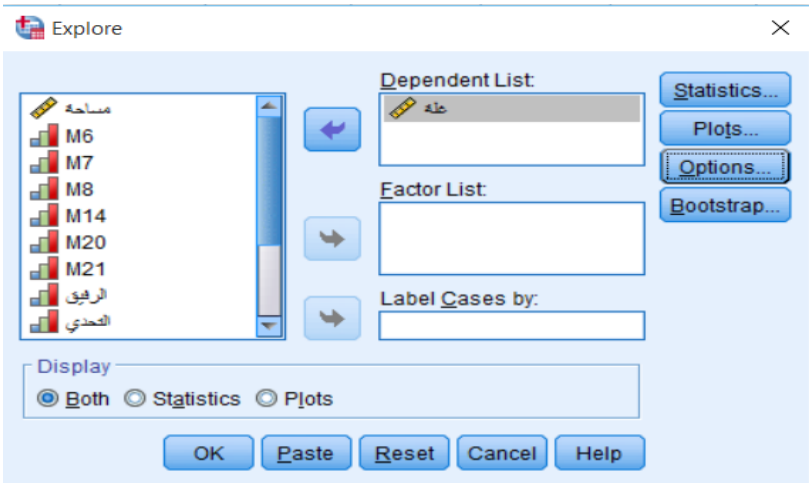
3.3. اختبار التوزيع الطبيعي Testing for Normality:

يعد تقييم التوزيع الطبيعي للبيانات شـرطًا أسـاسـيًا للعديد من الاختبارات الإحصـائية لأن البيانات الموزعة طبيعيا هي افتراض أساسي في الاختبارات البارامترية. هناك طريقتان رئيسيتان لاختبار التوزيع الطبيعي: بيانيًا وعدديًا.

ولإجراء اختبار التوزيع الطبيعي بيانيا، نتبع الخطوات التالية:

أ- ننقر فوق القائمة Analyze، ثم ننقر على Descriptive Statistics ثم Explore، فنتحصل على الشاشة:

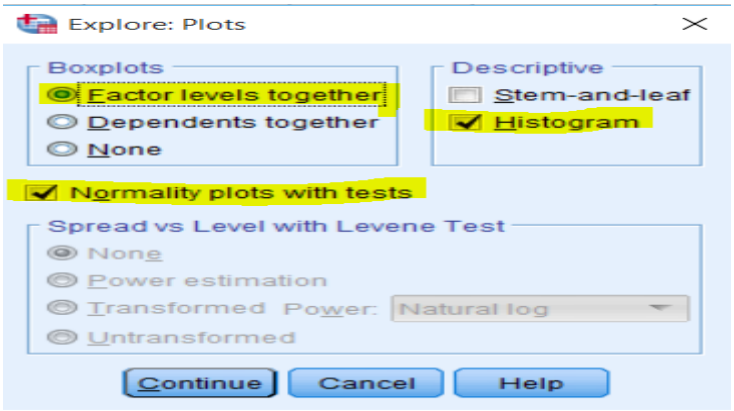
الشكل رقم (71): مربع الحوار Explore3



المصدر: مخرجات SPSS

ب- نختار المتغير الكمي المراد اختبار التوزيع الطبيعي له "غلة"، بالنقر عليه مرة واحدة ثم النقر على السهم (او بالنقر المزدوج)، فيتحول الى مربع Dependent List. من مربع Display نختار عرض الإحصاءات الوصفية والرسوم البيانية معا عبر النقر فوق Both. ثم ننقر فوق Plots ليظهر مربع الحوار المبين في الشكل التالي:

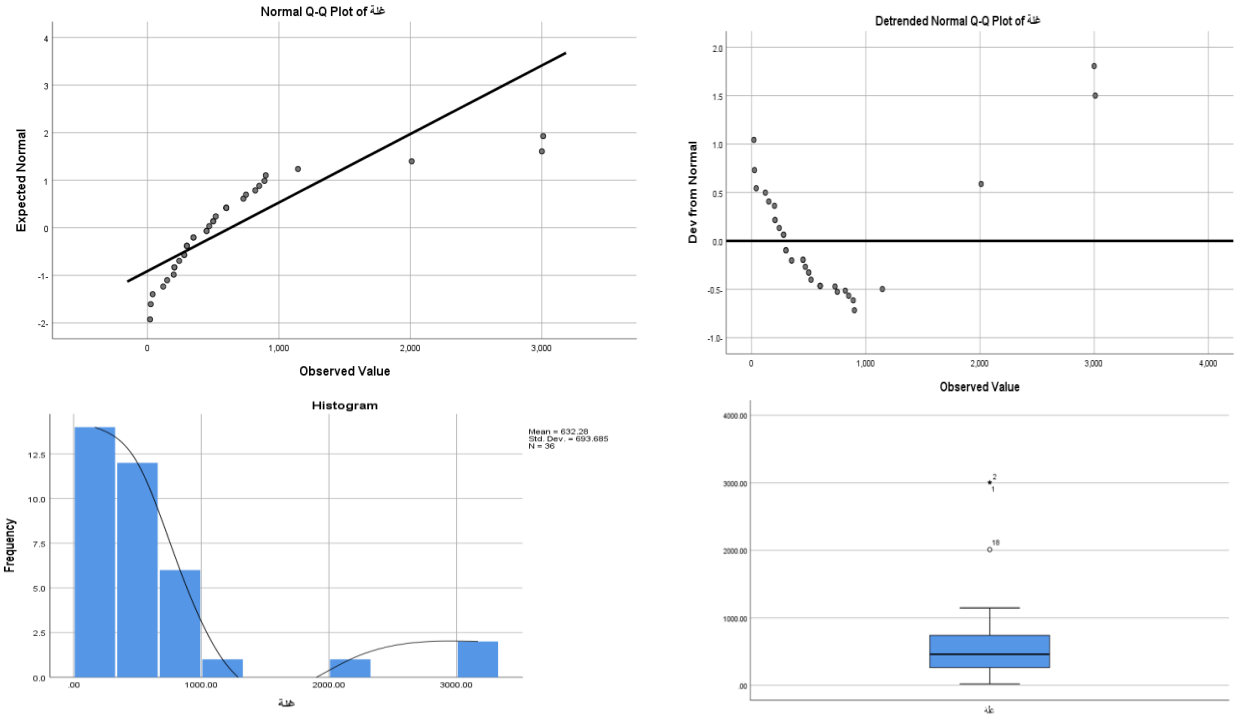
الشكل رقم (72): مربع الحوار Explore: Plots



المصدر: مخرجات SPSS

ت- من مربع Descriptive نختار Histogram، ومن مربع Boxplots نختار Factor levels together او Dependents together، وننقر فوق Continue. ثم ننقر فوق OK، سوف تظهر النتائج في شاشة مستعرض النتائج Output Viewer، كما هو موضح في الجداول الموالية:

الشكل رقم (73): مخرجات Normality Plots



المصدر: مخرجات SPSS

من غير المحتمل جدًا أن ينتج عن الرسم البياني لبيانات العينة منحنى طبيعي سلس تمامًا، خاصةً إذا كان حجم العينة صغيرًا. طالما أن البيانات تكون موزعة بشكل جرسي، مع وجود ذروة في المنتصف ومتناظرة إلى حد ما، فقد يتم استيفاء افتراض الحالة الطبيعية. بالرجوع للرسم البياني السابق لمثالنا فان أيا من هذه المتطلبات متوفرة، وبالتالي يمكننا القول ان بيانات المتغير "غلة" غير موزعة بشكل طبيعي. مخطط Q-Q plot الطبيعي هو طريقة رسومية بديلة لتقييم التوزيع الطبيعي للرسم البياني histogram وهي أسهل في الاستخدام عندما يكون هناك أحجام عينات صغيرة. يجب أن يكون التبعثر قريبًا من الخط قدر الإمكان مع عدم وجود نمط واضح بعيدًا عن الخط حتى يتم اعتبار البيانات موزعة بشكل طبيعي. واضح جدا من خلال نتائج توزيع متغير مثالنا "غلة" انها لا تتوفر على هذه المتطلبات، وان النقاط مبعثرة بعيدا عن الخط في الرسمين، وعليه متغير "غلة" لا يتبع التوزيع الطبيعي.

هناك أيضًا طرق محددة لاختبار التوزيع الطبيعي ولكن يجب استخدامها جنبًا إلى جنب مع الرسم البياني histogram أو مخطط Q-Q plot. يحدد اختبار Kolmogorov-Smirnov واختبار Shapiro-Wilk ما إذا كان التوزيع الأساسي طبيعيًا. كلا الاختبارين حساسان للقيم المتطرفة ويتأثران بحجم العينة:

- بالنسبة للعينات الأصغر، تقل احتمالية اكتشاف الحالة غير الطبيعية ولكن يجب تفضيل اختبار Shapiro-Wilk لأنه أكثر حساسية بشكل عام.
- بالنسبة للعينات الأكبر (أي أكثر من مائة)، تتم اختبارات التوزيع الطبيعي بتحفظ للغاية وقد يتم رفض افتراض الحالة الطبيعية بسهولة. يجب أن يتضمن أي اختبار أيضًا تقييمًا للتوزيع الطبيعي للرسوم البيانية histogram أو مخططات Q-Q plot لأن هذه أكثر ملاءمة لتقييم الحالة الطبيعية في العينات الأكبر.

ولإجراء اختبار التوزيع الطبيعي عددياً، نتبع نفس الخطوات لإجراء اختبار التوزيع الطبيعي بيانيا. سوف تظهر النتائج في شاشة مستعرض النتائج Output Viewer، كما هو موضح في الجداول الموالية:

الجدول رقم (15): مخرجات اختبار التوزيع الطبيعي

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
غلة	.241	36	.000	.680	36	.000

a. Lilliefors Significance Correction

المصدر: مخرجات SPSS

بما ان حجم العينة يساوي 36 فإن الاختبار الأنسب هو اختبارShapiro-Wilk. وبالرجوع الى قيمة والتي تساوي 0.000 وهي أصغر من مستوى المعنوية المفترض 0.05، وعليه فإننا نرفض الفرضية الصفرية ونقبل الفرضية البديلة، أي ان بيانات متغير "غلة" لا تتبع التوزيع الطبيعي.

4. أمثلة وتطبيقات:

التطبيق الأول:

باستخدام عينة عشوائية مكونة من 12 عائلة توصل باحث اقتصادي الى البيانات الاتية فيما يخص الدخل والادخار بآلاف الدنانير:

التسلسل	الدخل	الادخار
1	30.5	2.6
2	26	2.2
3	18	1.5
4	42.5	4
5	30	2.7
6	28	2.9
7	27.5	2.6
8	32.5	3
9	35	3.2
10	26	2.7
11	27.5	2.2
12	39	3.4

(أ) اوجد المتوسط الحسابي للمتغيرات. (ب) اوجد المدى والتباين. (ج) اختبر التوزيع الطبيعي للمتغيرات.

التطبيق الثاني:

إذا توفرت لديك البيانات التالية:

لون العين	لون البشرة	لون الشعر
اخضر	ابيض	اشقر
اخضر	اسمر فاتح	اشقر فاتح
اسود	اسمر	اسود
عسلي	اسمر فاتح	اسود
عسلي	ابيض	اشقر
اخضر	ابيض	اسود
اسود	اسمر فاتح	اشقر فاتح
ازرق	ابيض	اسود
عسلي	اسمر	اسود
اخضر	ابيض	اشقر

(أ) تكوين جدول توزيع تكراري لكل متغير مرتب تصاعديا حسب التكرارات. (ب) الربيعيات، المئين (20)، المئين (60). (ج) الوسط الحسابي، الوسيط، المنوال. (د) الاعمدة البيانية لمتغير لون الشعر بالاعتماد على التكرارات. (هـ) الدائرة البيانية لكل المتغيرات.

الدرس السادس:

إختبار الفرضيات ستيودنت T-test

1. تمهيد:

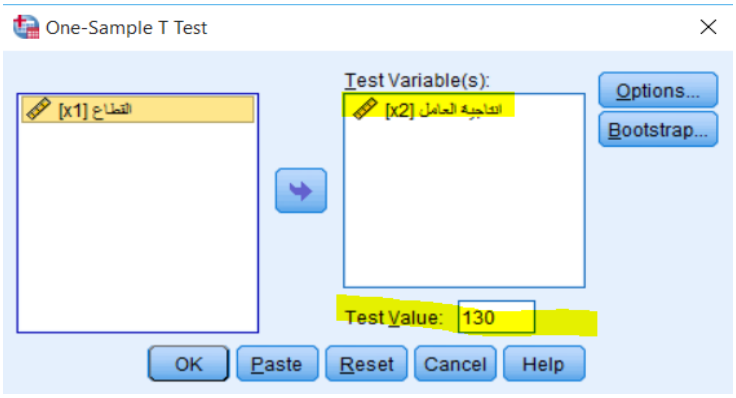
يعد اختبار الفرضيات الإحصائية من أهم الخصائص التي تميز البحوث الميدانية والتجريبية في مجالات علمية عديدة. يعتبر اختبار T-test أحد أشكال اختبار الفرضيات، وواحد من الاختبارات العديدة المستخدمة لهذا الغرض وهو من أكثر اختبارات الدلالة شيوعا في مختلف العلوم الإنسانية والاجتماعية، وترجع نشأته إلى العالم الكيميائي البريطاني وليام سيلي جوست سنة 1908، حيث نسب الاختبار للاسم المستعار "ستيودنت" اي الطالب، وأعطى الحرف الأخير في الكلمة وهو حرف T كاسم لهذا الاختبار.

2. اختبار ستيودنت للعينة الواحدة One Sample T-Test:

يستخدم هذا الاختبار لقياس الفرق المعنوي بين متوسط العينة عن قيمة ثابتة، يطلق عليها في برنامج SPSS القيمة الاختبارية Test Value. ولعمل استدلال حول المتوسط الحسابي لمجتمع في برنامج SPSS، على سبيل المثال لاختبار إنتاجية عامل حول متوسط قدره 135 نتبع الخطوات التالية:

أ- ننقر على قائمة Analyze ثم ننقر على Compare Means ثم One Sample T-Test، ستظهر لنا شاشة الحوار التالية:

الشكل رقم (74): مربع الحوار One Sample T-Test



المصدر: مخرجات SPSS

ب- ننقر على المتغير المراد فحص واختبار متوسطه، وهو في مثالنا متغير "إنتاجية العامل"، ثم ننقر على السهم لتحويله الى مربع Test variable(s). ونكتب رقم 130 في مربع Test Value. ثم ننقر OK، ستظهر لنا نتائج اختبار ستيودنت للعينة الواحدة في شاشة المخرجات كما في الجدولين المواليين:

الجدول رقم (16): مخرجات الإحصاء الوصفي لاختبار One Sample T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
انتاجية العامل	19	138.6316	11.03636	2.53191

المصدر: مخرجات SPSS

قام برنامج SPSS بحساب المتوسط الحسابي (Mean) والانحراف المعياري (Std. Deviation) والخطأ المعياري (Std. Error Mean) للمتغير الذي اختير لفحص متوسطه.

الجدول رقم (17): مخرجات اختبار One Sample T-Test

One-Sample Test

Test Value = 130

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
انتاجية العامل	3.409	18	.003	8.63158	3.3122	13.9509

المصدر: مخرجات SPSS

تشير مخرجات الجدول السابق الى ان متوسط الفرق بين المتغير والقيمة المفترضة (Mean Difference) والتي بلغت في هذا المثال 8.63، والذي يشير الى ان مستوى إنتاجية العمل لدى عينة الدراسة كانت اعلى من المستوى الطبيعي 130، ولكن هل هذا الفرق يعتبر كافيا لكن نقرر ان هناك فرق؟ ام ان الفرق عائد للصدفة نتيجة اختيار العينة؟ نستطيع الإجابة على هذا السؤال من خلال اختبار فرضية ستيودنت. فعند قبول الفرضية الصفرية نقول انه ليس هناك فرق، بينما عند قبول الفرضية البديلة نقر بوجود فرق.

في مثالنا هذا نلاحظ ان قيمة Sig. (2-tailed) والتي تساوي 0.003 وهي أصغر من 0.05 مستوى المعنوية المفترض، وبالتالي فإننا نرفض الفرضية الصفرية ونقبل الفرضية البديلة، أي ان هناك فرق ذو دلالة إحصائية بين متوسط العينة والقيمة المفترضة.

3. اختبار ستيودنت للعينات المزدوجة Paired Samples T-Test:

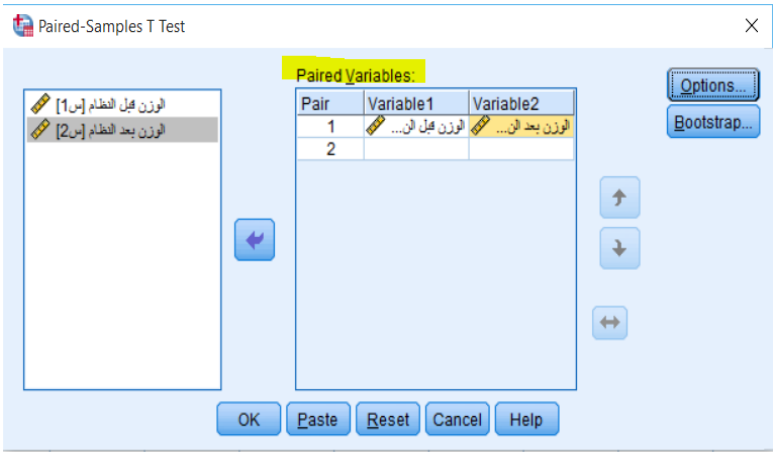
يستخدم هذا الاختبار لقياس الفرق المعنوي (Significant Difference) بين متوسطي متغيرين مرتبطين، او بين متوسطي عينتين مرتبطتين لمتغير واحد. ويفترض في هذا الاختبار تحقق الشرطين التاليين:

- ان يكون توزيع الفرق بين المتغيرين او العينتين طبيعيا.
- ان تكون قيم الفرق بين المتغيرين او العينتين مستقلة عن بعضها البعض.

على سـبيل المثال لإجراء اختبار سـتيودنت للعينات المزدوجة لأوزان مجموعة من الافراد قبل تناول نظام غذائي وبعده، نتبع الخطوات التالية:

ت- ننقر على قائمة Analyze ثم ننقر علىCompare Means ثم Paired Sample T-Test، سـتظهر لنا شاشة الحوار التالية:

الشكل رقم (75): مربع الحوارPaired Sample T-Test



المصدر: مخرجات SPSS

ث- ننقر على المتغيرين المراد فحص متوسـطاتهما، وهما في مثالنا متغير "الوزن قبل النظام" و"الوزن بعد النظام"، ثم ننقر على السـهم لتحويلهما الى مربع Paired variables. ثم ننقر OK، سـتظهر لنا نتائج اختبار سـتيودنت للعينتين المزدوجة في شاشة المخرجات كما في الجداول الموالية:

الجدول رقم (18): مخرجات الإحصاء الوصفي لاختبار Paired Sample T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	الوزن قبل النظام	100.8500	20	12.11035	2.70796
	الوزن بعد النظام	91.7000	20	10.13644	2.26658

المصدر: مخرجات SPSS

قام برنامج SPSS بحساب المتوسط الحسابي (Mean) والانحراف المعياري (Std. Deviation) والخطأ المعياري (Std. Error Mean) للمتغيرين الذين اختيرا لفحص متوسطاتهما.

الجدول رقم (19): مخرجات اختبار العلاقة الارتباطية Paired Sample T-Test

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	الوزن قبل النظام & الوزن بعد النظام	20	.957	.000

المصدر: مخرجات SPSS

يوضح الجدول السابق اختبار العلاقة الارتباطية بين المتغيرين المراد فحص الاختلاف بينهما. حيث بينت النتائج الى وجود علاقة ارتباطية طردية قوية (0.957) وذات معنوية إحصائية.

الجدول رقم (20): مخرجات اختبار Paired Sample T-Test

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	الوزن قبل النظام - الوزن بعد النظام	9.1500	3.78744	.84690	7.37742	10.92258	10.804	19	.000

المصدر: مخرجات SPSS

تشير مخرجات الجدول السابق الى ان متوسط الفرق بين المتغير "الوزن قبل النظام" والمتغير "الوزن بعد النظام" والذي بلغ في هذا المثال 9.15، كما نلاحظ ان قيمة Sig. (2-tailed) والتي تساوي 0.000 وهي أصغر من 0.05 مستوى المعنوية المفترض، وبالتالي فإننا نرفض الفرضية الصفرية ونقبل الفرضية البديلة، أي ان هناك فرق ذو دلالة إحصائية بين اوزان الافراد قبل تناول النظام الغذائي وبعده، بعبارة أخرى ان تناول النظام الغذائي كلن له دور واثر معنوي في انقاص متوسط اوزان الافراد في العينة.

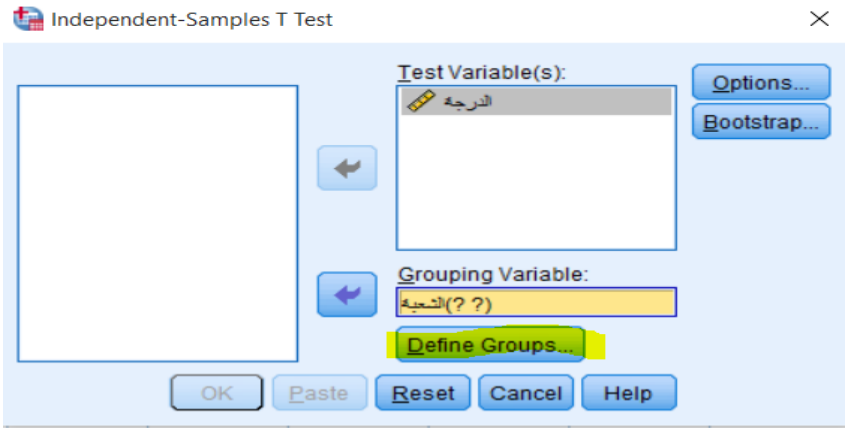
4. اختبار ستيودنت للعينات المستقلة Independent Samples T-Test:

يستخدم هذا الاختبار لقياس الفرق المعنوي (Significant Difference) بين متوسطي عينتين مستقلتين (Independent Samples). حيث يضم هذا الاختبار نوعين من المتغيرات: متغير التجميع (Grouping Variable) ويضم العينتين المستقلتين، ومتغير الاختبار ويضم متغير الدراسة. ويفترض هذا الاختبار ان يكون توزيع متغير الاختبار طبيعيا لكل عينة من عينات متغير التجميع. كما يستخدم هذا الاختبار لحالتين: افتراض ان تباين العينتين متساو، وافتراض ان تباين العينتين غير متساو.

على سبيل المثال لإجراء اختبار ستيودنت لاختبار الفروق في درجات الإحصاء بين عينة من طلاب الاقتصاد وعينة من طلاب التجارة، نتبع الخطوات التالية:

أ- ننقر على قائمة Analyze ثم ننقر على Compare Means ثم Independent Samples T-Test، ستظهر لنا شاشة الحوار التالية:

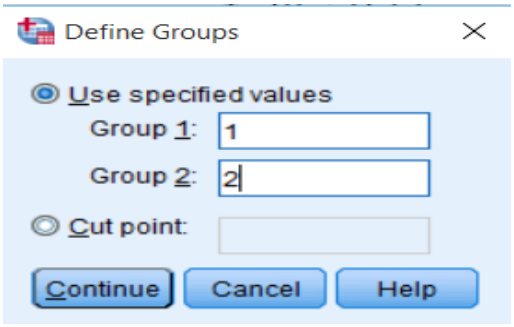
الشكل رقم (76): مربع الحوار Independent Samples T-Test



المصدر: مخرجات SPSS

ب- ننقر على المتغير "الدرجة"، ثم ننقر على السهم لتحويله الى مربع Test Variable(s)، ثم ننقر على المتغير "الشعبة"، ثم ننقر على السهم لتحويله الى مربع Grouping Variable. ثم ننقر فوق زر Define Groups، ستظهر لنا مربع الحوار كما في الشكل الموالي:

الشكل رقم (77): مربع الحوار Define Groups



المصدر: مخرجات SPSS

ت- يتم تحديد متغير التجميع الذين يمثلان المجموعتين المراد اختبار متوسطاتهما، ثم ندخل رقم 1 في Group1 وندخل رقم 2 في Group2. وننقر فوق Continue، ثم ننقر OK، ستظهر لنا نتائج اختبار ستيودنت للعينتين المستقلتين في شاشة المخرجات كما في الجداول الموالية:

الجدول رقم (21): مخرجات الإحصاء الوصفي لاختبار Independent Samples T-Test

Group Statistics

	الشعبة	N	Mean	Std. Deviation	Std. Error Mean
الدرجة	اقتصاد	12	71.0000	15.81714	4.56601
	تجارة	12	69.3333	14.84057	4.28410

المصدر: مخرجات SPSS

قام برنامج SPSS بحساب المتوسط الحسابي (Mean) والانحراف المعياري (Std. Deviation) والخطأ المعياري (Std. Error Mean) لمتغير الدرجة لكل فئة من الفئتين الاقتصاد والتجارة.

الجدول رقم (22): مخرجات اختبار Independent Samples T-Test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
الدرجة	Equal variances assumed	.377	.546	.266	22	.793	1.66667	6.26115	-11.3181	14.65150
	Equal variances not assumed			.266	21.911	.793	1.66667	6.26115	-11.3212	14.65455

المصدر: مخرجات SPSS

يشير اختبار تجانس التباين للفئتين (Homogeneity of Variances) بالاختبار المسمى (Levene's Test)، حيث تشير قيمة Sig. للإحصائية F والتي بلغت 0.546 وهي اكبر من مستوى المعنوية المفترض، وبالتالي نقبل الفرضية الصفرية والتي تنص على وجود تجانس بين تبايني الفئتين (الاقتصاد والتجارة)، وعليه فإننا نعتمد على نتائج صف (Equal variances assumed) في جدول النتائج. كما نلاحظ ان قيمة Sig. (2-tailed) في الصف الأول (Equal variances assumed) والتي تساوي 0.793 وهي أكبر من 0.05 مستوى المعنوية المفترض، وبالتالي فإننا نقبل الفرضية الصفرية ونرفض الفرضية البديلة، أي انه لا يوجد فرق ذو دلالة إحصائية بين درجات الاحصاء لطلبة الاقتصاد ودرجات الإحصاء لطلبة التجارة.

5. أمثلة وتطبيقات:

تطبيق رقم 01: انتجت احدى الشركات غذاء معينا ادعت انه يساعد في تخفيض الوزن بمقدار 2كغ خلال أسبوعين. وقد استهلك 12 اشخاص هذا الغذاء، وكان مقدار النقص في اوزانهم كما في الجدول التالي:

الاشخاص	1	2	3	4	5	6	7	8	9	10	11	12
مقدار النقص	2.4	2	1.8	2.1	1.4	2.3	1	1.2	1.5	1.7	1.1	2.2

المطلوب: اختبار فرضية العدم H_0 القائلة بان المتوسط الحسابي لمقدار النقص في الاوزان تساوي 2كغ.

تطبيق رقم 02: إذا توفرت لديك درجات شعبتين من الطلبة كما في الجدول التالي:

الاولى	78	58	88	90	75	64	81	95	53	46	58	66	67	72
الثانية	55	68	96	45	82	63	77	48	68	78	50	72	58	80

المطلوب: هل يوجد هناك فرق معنوي بين متوسطي المجموعتين عند مستوى دلالة 5%؟

تطبيق03: لدينا عينة تتكون من 11 نبتة، ومعطيات عن اطوال هذه النباتات قبل وبعد تعرضها للضوء كما في الجدول الموالي:

رقم النبتة	1	2	3	4	5	6	7	8	9	10	11
الطول قبل التعرض للضوء	31	33	35	30	36	37	41	35	39	32	30
الطوا بعد التعرض للضوء	33	32	36	29	39	38	41	40	43	34	36

المطلوب: اختبار ان كان هناك فرق جوهري في اطوال النباتات قبل وبعد تعرضها للإضاءة الإضافية عند مستوى معنوية 5%.

الدرس السابع:

تحليل التباين Analysis of variance

1. تمهيد:

هو مجموعة من النماذج الإحصائية (statistical model) مع إجراء مرافقة لهذه النماذج تمكن من مقارنة المتوسطات لمجتمعات إحصائية مختلفة عن طريق تقسيم التباين variance الكلي الملاحظ بينهم إلى أجزاء مختلفة. أول طرق تحليل التباين تم وضعها من قبل الإحصائي رونالد فيشر في العشرينات والثلاثينات من القرن العشرين لذلك تعرف أحيانا بتحليل فيشر للتباين.

2. تحليل التباين الأحادي One Way ANOVA:

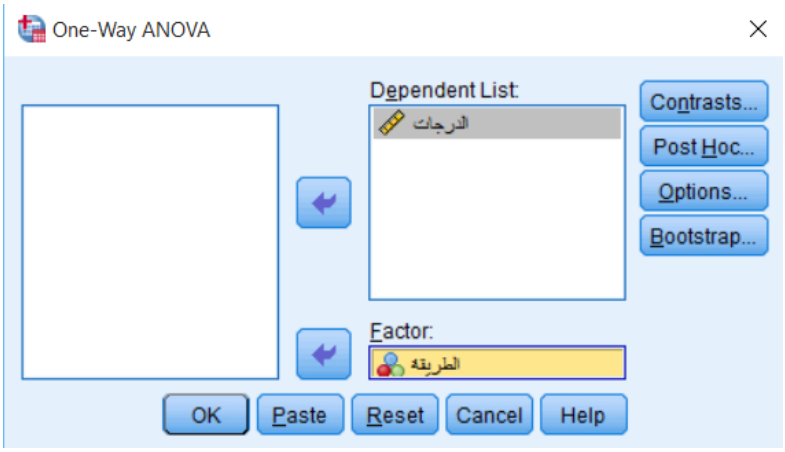
يستخدم هذا الاختبار عند توفر متغير مستقل واحد، الذي يطلق عليه بالمتغير العاملي (Factor)، وهو متغير من النوع الاسمي (Nominal) او الترتيبي (Ordinal)، الذي على أساسه ستقسم العينات المراد اختبار فروقات متوسطاتها. ومتغير تابع (Dependent) واحد وهو متغير من النوع الكمي. ويستخدم هذا الاختبار اذا كان المتغير العاملي مكون من مستويين أو اكثر، واذا كان مكون من مستويين فقط فيمكن استخدام اختبار ستيودنت للعينتين المستقلتين. ويشترط في اختبار تحليل التباين الأحادي تحقق ما يلي:

- ان يكون توزيع المتغير المعتمد طبيعيا لكل عينة من عينات المتغير العاملي.
- ان يكون تباين المتغير المعتمد متساويا لكل عينة من عينات المتغير العاملي.
- ان تكون قيم المتغير المعتمد مستقلة بعضها عن البعض، ولكل عينة من عينات المتغير العاملي.
- ان تكون كل عينة من عينات المتغير العاملي عشوائية.

على سبيل المثال نرغب باختبار الفروق بين ثلاثة طرق تدريسية مختلفة على درجات مجموعة من الطلبة، ومن أجل ذلك سوف نقوم بإجراء تحليل التباين الأحادي وفق الخطوات التالية:

أ- ننقر على قائمة Analyze ثم ننقر علىCompare Means ثم One-Way ANOVA، ستظهر لنا شاشة الحوار التالية:

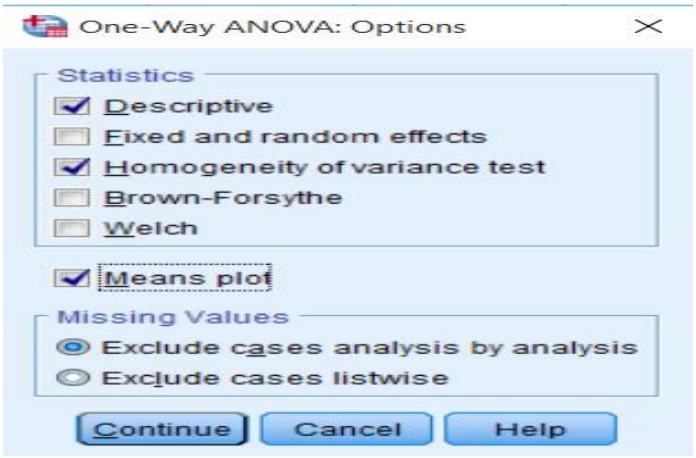
الشكل رقم (78): مربع الحوار One-Way ANOVA



المصدر: مخرجات SPSS

ب- ننقر على المتغير "الدرجات"، ثم ننقر على السهم لتحويله الى مربع Dependent List، ثم ننقر على المتغير "الطريقة"، ثم ننقر على السهم لتحويله الى مربع Factor. ثم ننقر فوق زر Options، ستظهر لنا مربع الحوار كما في الشكل الموالي:

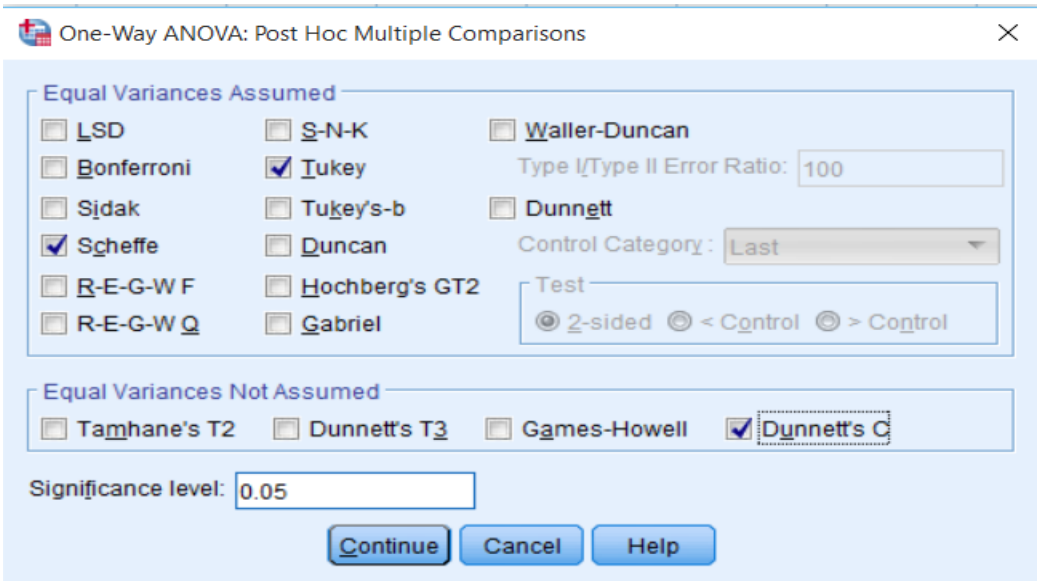
الشكل رقم (79): مربع الحوار One-Way ANOVA: Options



المصدر: مخرجات SPSS

ت- من مربع الحوار نختار Descriptive لعرض الإحصاءات الوصفية، ونختار Homogeneity of variance test لفحص تماثل تباين المجموعات (الشرط الثاني)، كما يمكننا اختيار اختبار Brown-Forsythe او اختبار Welch الذين يستخدمان في حالة عدم تحقق شرط التباين كبديل لاختبار فيشر F. يمكن النقر على لعمل رسم بياني يمثل الفروقات بين متوسطات المتغير المعتمد لكل فئة من فئات المتغير العاملي. ثم ننقر فوق Continue، لنعود الى شاشة الحوار السابقة، وننقر مفتاح الاختبارات البعدية PostHoc، فيظهر لنا مربع الحوار التالي:

الشكل رقم (80): مربع الحوار One-Way ANOVA: PostHoc



المصدر: مخرجات SPSS

ث- هناك مجموعتين من الاختبارات البعدية: الجزء العلوي يشترط تجانس التباين لمجموعات المتغير العاملي Equal Variances Assumed، بينما الجزء السفلي لا يشترط تجانس التباين Equal Variances Not Assumed. وعادة ما يستخدم اختبار شيفيه Scheffe أو توكي Tukey من الجزء الأول، واختبار Dunnett's C من الجزء الثاني. وننقر فوق Continue، ثم ننقر OK، ستظهر لنا نتائج اختبار تحليل التباين الاحادي في شاشة المخرجات كما في الجداول الموالية:

الجدول رقم (23): مخرجات الاحصائيات الوصفية One-Way ANOVA

Descriptives

الدرجات

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
الطريقة الأولى	9	74.4444	3.28295	1.09432	71.9209	76.9679	71.00	81.00
الطريقة الثانية	9	79.1111	2.31541	.77180	77.3313	80.8909	77.00	84.00
الطريقة الثالثة	9	84.7778	2.81859	.93953	82.6112	86.9443	79.00	89.00
Total	27	79.4444	5.09399	.98034	77.4293	81.4596	71.00	89.00

المصدر: مخرجات SPSS

قام برنامج SPSS بحساب المتوسط الحسابي (Mean) والانحراف المعياري (Std. Deviation) والخطأ المعياري (Std. Error Mean) والحد الأدنى (Minimum) والحد العلى (Maximum) للمتغير التابع الدرجات، لكل فئة من الفئات الثلاثة للمتغير العاملي: الطريقة الأولى، الطريقة الثانية، والطريقة الثالثة.

الجدول رقم (24): مخرجات اختبار تجانس التباينOne-Way ANOVA

Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
الدرجات	Based on Mean	.633	2	24	.540
	Based on Median	.235	2	24	.792
	Based on Median and with adjusted df	.235	2	19.612	.793
	Based on trimmed mean	.554	2	24	.582

المصدر: مخرجات SPSS

اختبار تجانس التباين Test of Homogeneity of variance الموضح في الجدول أعلاه، يشير الى تساوي تباينات المجموعات الثلاثة، حيث كانت قيمة .Sig أكبر من مستوى المعنوية المفترض 0.05.

الجدول رقم (25): مخرجات تحليل التباين One-Way ANOVA

الدرجات

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	482.000	2	241.000	30.021	.000
Within Groups	192.667	24	8.028		
Total	674.667	26			

المصدر: مخرجات SPSS

تشير نتيجة تحليل التباين الأحادي الى وجود فروق ذات دلالة إحصائية عند مستوى معنوية 0.05، حيث قيمة .Sig أصغر من مستوى المعنوية المفترض 0.05.

الجدول رقم (26): مخرجات المقارنات البعدية

Multiple Comparisons

Dependent Variable: الدرجات

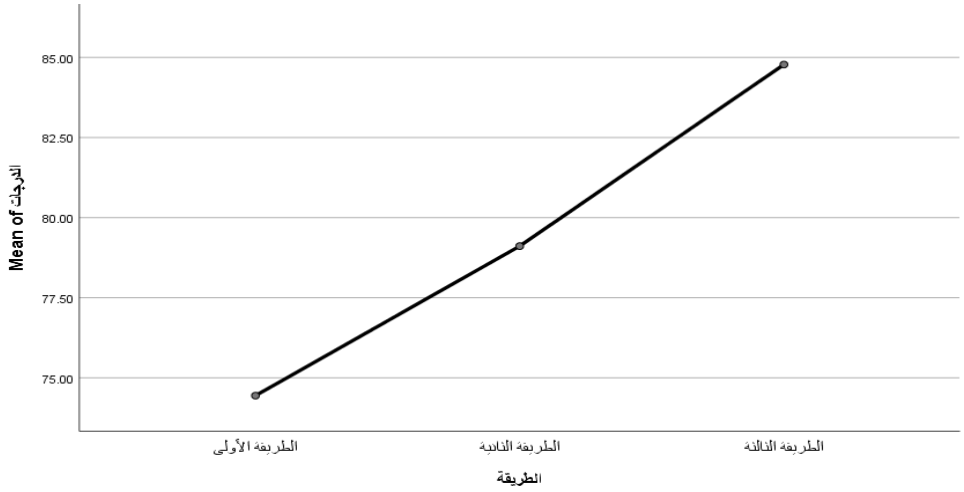
	(I) الطريقة	(J) الطريقة	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Tukey HSD	الطريقة الأولى	الطريقة الثانية	-4.66667-*	1.33565	.005	-8.0022-	-1.3312-
		الطريقة الثالثة	-10.33333-*	1.33565	.000	-13.6688-	-6.9978-
	الطريقة الثانية	الطريقة الأولى	4.66667*	1.33565	.005	1.3312	8.0022
		الطريقة الثالثة	-5.66667-*	1.33565	.001	-9.0022-	-2.3312-
	الطريقة الثالثة	الطريقة الأولى	10.33333*	1.33565	.000	6.9978	13.6688
		الطريقة الثانية	5.66667*	1.33565	.001	2.3312	9.0022
Scheffe	الطريقة الأولى	الطريقة الثانية	-4.66667-*	1.33565	.007	-8.1511-	-1.1823-
		الطريقة الثالثة	-10.33333-*	1.33565	.000	-13.8177-	-6.8489-
	الطريقة الثانية	الطريقة الأولى	4.66667*	1.33565	.007	1.1823	8.1511
		الطريقة الثالثة	-5.66667-*	1.33565	.001	-9.1511-	-2.1823-
	الطريقة الثالثة	الطريقة الأولى	10.33333*	1.33565	.000	6.8489	13.8177
		الطريقة الثانية	5.66667*	1.33565	.001	2.1823	9.1511
Dunnett C	الطريقة الأولى	الطريقة الثانية	-4.66667-*	1.33911		-8.4931-	-.8402-
		الطريقة الثالثة	-10.33333-*	1.44231		-14.4546-	-6.2120-
	الطريقة الثانية	الطريقة الأولى	4.66667*	1.33911		.8402	8.4931
		الطريقة الثالثة	-5.66667-*	1.21589		-9.1410-	-2.1923-
	الطريقة الثالثة	الطريقة الأولى	10.33333*	1.44231		6.2120	14.4546
		الطريقة الثانية	5.66667*	1.21589		2.1923	9.1410

*. The mean difference is significant at the 0.05 level.

المصدر: مخرجات SPSS

من خلال اختبار تجانس التباين Test of Homogeneity of variance تبين ان التباينات متماثلة بين المجموعات الثلاثة، وبالتالي يمكن استخدام نتائج احد الاختبارات البعدية التي تشترط تجانس التباينات اختبار شيفيه Scheffe أو توكيTukey. بالاعتماد على نتائج اختبار شيفيه Scheffe ، وبالاعتماد على عمود Mean Difference (I-J) في جدول النتائج، حيث يشير وجود نجمة * الى ان الفروق ذات دلالة إحصائية عند مستوى 0.05، كما ان الإشارة السالبة للفرق تشير الى ان الفرق لمصلحة الطريقة(J)، بينما الإشارة الموجبة للفرق تشير الى ان الفرق لمصلحة الطريقة(I). يمكن كذلك معرفة معنوية الفروق من خلال عمود .Sig، حيث كلما كانت .Sig أصغر من 0.05 دل ذلك على وجود فروق ذات معنوية إحصائية، بينما إذا كانت .Sig أكبر من 0.05 دل ذلك على عدم وجود فروق ذات معنوية إحصائية.

الشكل رقم (81): رسم متوسطات المجموعات



المصدر: مخرجات SPSS

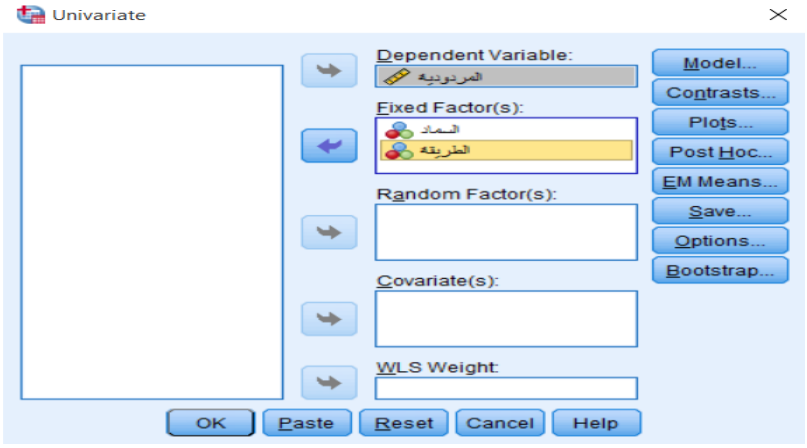
تشير نتيجة رسم متوسطات المجموعات في الشكل أعلاه، والذي يظهر فيه تباعد بين الطرق الثلاثة للتدريس، وهو ما يدعم نتائج الاختبارات السابقة.

3. تحليل التباين ذو المستوى الاعلى Higher-Way ANOVA:

يستخدم تحليل التباين ذو المستوى الأعلى لفحص الفروق لأكثر من متغير عاملي على المتغير التابع. على سبيل المثال نرغب باختبار الفروق بين ثلاثة أنواع سماد مختلفة، وأربعة طرق للري على مردودية محصـــول معين، ومن أجل ذلك سوف نقوم بإجراء تحليل التباين الثنائي وفق الخطوات التالية:

أ– ننقر على قائمة Analyze ثم ننقر على General Linear Model ثم Univariate، ستظهر لنا شاشة الحوار التالية:

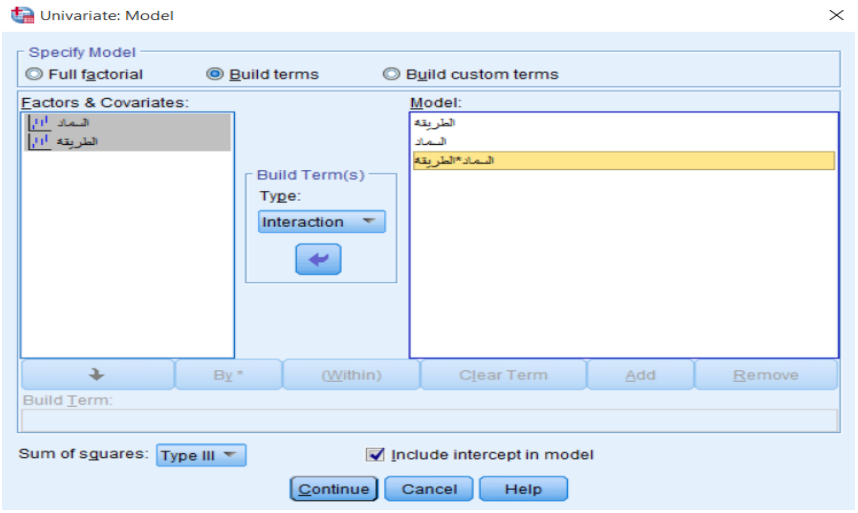
الشكل رقم (82): مربع الحوارUnivariate1



المصدر: مخرجات SPSS

ب- ننقر على المتغير التابع "المردودية"، ثم ننقر على الســـهم لتحويله الى مربع Dependent Variable، ثم ننقر على المتغيرين العامليين الأول "الســـماد" والثاني "الطريقة" (مع ملاحظة انه يمكن ادراج اكثر من متغيرين عامليين)، ثم ننقر على الســـهم لتحويلهما الى مربع Fixed Factor(s). ثم ننقر فوق زر Model، ســـتظهر لنا مربع الحوار كما في الشكل الموالي:

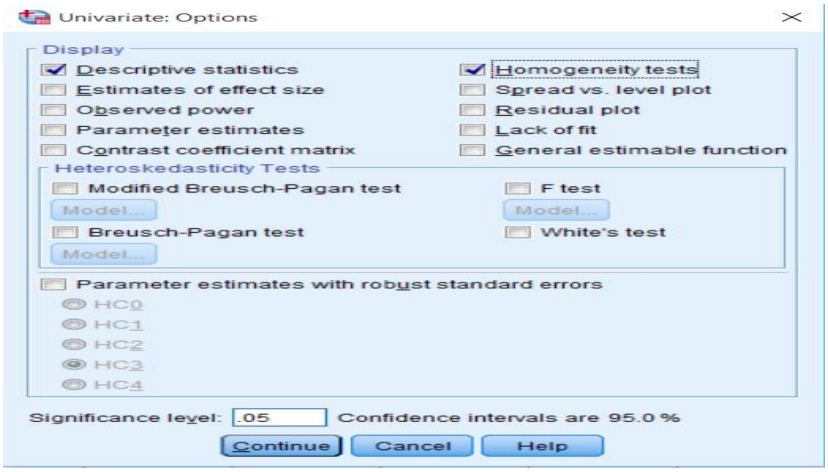
الشكل رقم (83): مربع الحوار Univariate: Model



المصدر: مخرجات SPSS

ت- ننقر على الخيار Build terms في الأعلى بدلا من Full factorial لاختيار المتغيرات العاملية المراد دراســـتها وتفاعلاتها، التابع "المردودية"، ثم من خلال حقل Build term(s) ننقر على الســـهم لاختيار Interaction ونحدد كل متغير عاملي على حدى وننقله الى مربع Model، ثم نحدد كلا المتغيرين سوية وننقلهما الى مربع Model لاختبار تفاعل المتغيرين. ثم ننقر فوق زر Continue، فيتم الرجوع الى مربع الحوار Univariate ومنه ننقر على خيار Options فيظهر لنا مربع الحوار كما في الشكل الموالي:

الشكل رقم (84): مربع الحوار Univariate: Options



المصدر: مخرجات SPSS

ث- ننقر على Descriptive statistics لإيجاد بعض الإحصـاءات الوصـفية، وننقر على Homogeneity tests لاختبار تجانس التباين، ثم ننقر فوق زر Continue، فيتم الرجوع الى مربع الحوار Univariate ومنه ننقر OK، ستظهر لنا نتائج اختبار تحليل التباين الثنائي في شاشة المخرجات كما في الجداول الموالية:

الجدول رقم (27): مخرجات الاحصائيات الوصفية Univariate

Descriptive Statistics

Dependent Variable: الخطأ

مهندس	تقني	Mean	Std. Deviation	N
1.00	1.00	.6667	1.15470	3
	2.00	1.0000	1.00000	3
	3.00	1.6667	1.15470	3
	Total	1.1111	1.05409	9
2.00	1.00	2.6667	1.15470	3
	2.00	1.6667	1.52753	3
	3.00	1.6667	1.15470	3
	Total	2.0000	1.22474	9
3.00	1.00	.3333	.57735	3
	2.00	.3333	.57735	3
	3.00	1.0000	1.00000	3
	Total	.5556	.72648	9
4.00	1.00	1.0000	.00000	3
	2.00	.6667	.57735	3
	3.00	1.0000	1.00000	3
	Total	.8889	.60093	9
Total	1.00	1.1667	1.19342	12
	2.00	.9167	.99620	12
	3.00	1.3333	.98473	12
	Total	1.1389	1.04616	36

المصدر: مخرجات SPSS

قام برنامج SPSS بحساب المتوسط الحسابي (Mean) والانحراف المعياري (Std. Deviation) للمتغير التابع الخطأ، لكل فئة من الفئات الاربعة للمتغير العاملي الأول مهندس: مهندس أول، مهندس ثاني، مهندس ثالث، ومهندس رابع، ولكل فئة من الفئات الثلاثة للمتغير العاملي الثاني: تقني أول، تقني ثاني، وتقني ثالث.

الجدول رقم (28): مخرجات اختبار تجانس التباينات Univariate

Levene's Test of Equality of Error Variances[a,b]

		Levene Statistic	df1	df2	Sig.
الخطأ	Based on Mean	1.479	11	24	.203
	Based on Median	.291	11	24	.982
	Based on Median and with adjusted df	.291	11	15.823	.978
	Based on trimmed mean	1.337	11	24	.265

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Dependent variable: الخطأ

b. Design: Intercept + تقني + مهندس * تقني + مهندس

المصدر: مخرجات SPSS

اختبار تجانس التباين Test of Homogeneity of variance الموضح في الجدول أعلاه، يشير الى تساوي تباينات المجموعات للمتغيرين العاملين، حيث كانت قيمة Sig. أكبر من مستوى المعنوية المفترض 0.05.

الجدول رقم (29): مخرجات اختبار تحليل التباين Univariate

Tests of Between-Subjects Effects

Dependent Variable: الخطأ

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	14.972[a]	11	1.361	1.400	.236
Intercept	46.694	1	46.694	48.029	.000
مهندس	10.306	3	3.435	3.533	.030
تقني	1.056	2	.528	.543	.588
مهندس * تقني	3.611	6	.602	.619	.713
Error	23.333	24	.972		
Total	85.000	36			
Corrected Total	38.306	35			

a. R Squared = .391 (Adjusted R Squared = .112)

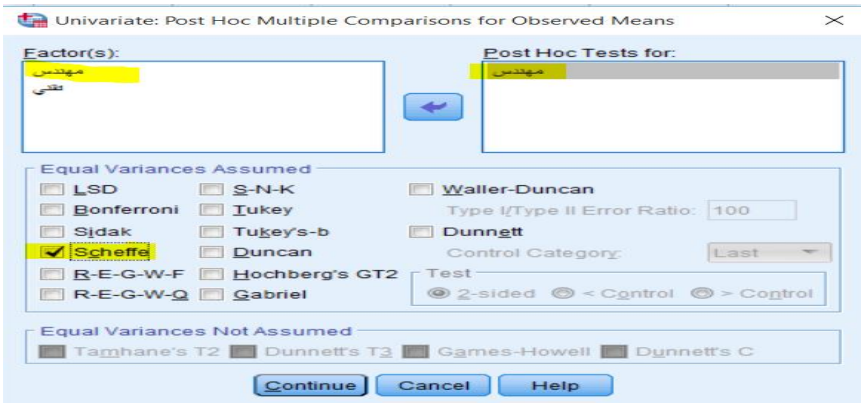
المصدر: مخرجات SPSS

تشير نتيجة تحليل التباين الثنائي الى وجود فروق ذات دلالة إحصائية للمتغير العاملي مهندس عند مستوى معنوية 0.05، حيث قيمة Sig. أصغر من مستوى المعنوية المفترض 0.05، أي انه يوجد فروق معنوية بين المهندسين في الخطأ. بينما تشير نتيجة تحليل التباين الثنائي الى عدم وجود فروق ذات دلالة إحصائية للمتغير العاملي تقني عند مستوى معنوية 0.05، حيث قيمة Sig. أكبر من مستوى المعنوية المفترض 0.05، أي انه لا يوجد فروق معنوية بين التقنيين في الخطأ. كما تشير نتيجة تحليل التباين الثنائي الى عدم وجود فروق ذات دلالة إحصائية لتفاعل المتغيرين العاملين مع بعض (مهندس

* تقني) عند مستوى معنوية 0.05، حيث قيمة .Sig أكبر من مستوى المعنوية المفترض 0.05، أي انه لا يوجد فروق معنوية بين تفاعل المتغيرين العامليين مع بعض (مهندس * تقني) في الخطأ.

ج- ولكون الاختلاف في أداء المهندسين كان له تأثير معنوي فيجب ان نعرف أي مصدر الاختلاف مثنى مثنى. وذلك بتطبيق المقارنات البعدية، حيث ننقر مفتاح الاختبارات البعدية Post Hoc في مربع الحوار Univariate، فيظهر لنا مربع الحوار التالي:

الشكل رقم (85): مربع الحوارPostHoc :Univariate



المصدر: مخرجات SPSS

ح- ننقر على المتغير العاملي مهندس من مربع Factors، ونحوله بواسطة السهم الى مربع Post Hoc Tests for. ثم ننقر على Scheffe في مربع Equal Variances Assumed. وننقر فوق Continue، ثم ننقر OK، ستظهر لنا نتائج اختبار تحليل التباين الثنائي في شاشة المخرجات كما في الجداول الموالية:

الجدول رقم (30): مخرجات اختبار المقارنات المتعددة Univariate

Multiple Comparisons

Dependent Variable: الخطأ

Scheffe

مهندس (I)	مهندس (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1.00	2.00	-.8889-	.46481	.324	-2.2854-	.5076
	3.00	.5556	.46481	.702	-.8409-	1.9520
	4.00	.2222	.46481	.972	-1.1743-	1.6187
2.00	1.00	.8889	.46481	.324	-.5076-	2.2854
	3.00	1.4444*	.46481	.041	.0480	2.8409
	4.00	1.1111	.46481	.156	-.2854-	2.5076
3.00	1.00	-.5556-	.46481	.702	-1.9520-	.8409
	2.00	-1.4444-*	.46481	.041	-2.8409-	-.0480-
	4.00	-.3333-	.46481	.915	-1.7298-	1.0631
4.00	1.00	-.2222-	.46481	.972	-1.6187-	1.1743
	2.00	-1.1111-	.46481	.156	-2.5076-	.2854
	3.00	.3333	.46481	.915	-1.0631-	1.7298

Based on observed means.

The error term is Mean Square(Error) = .972.

*. The mean difference is significant at the .05 level.

المصدر: مخرجات SPSS

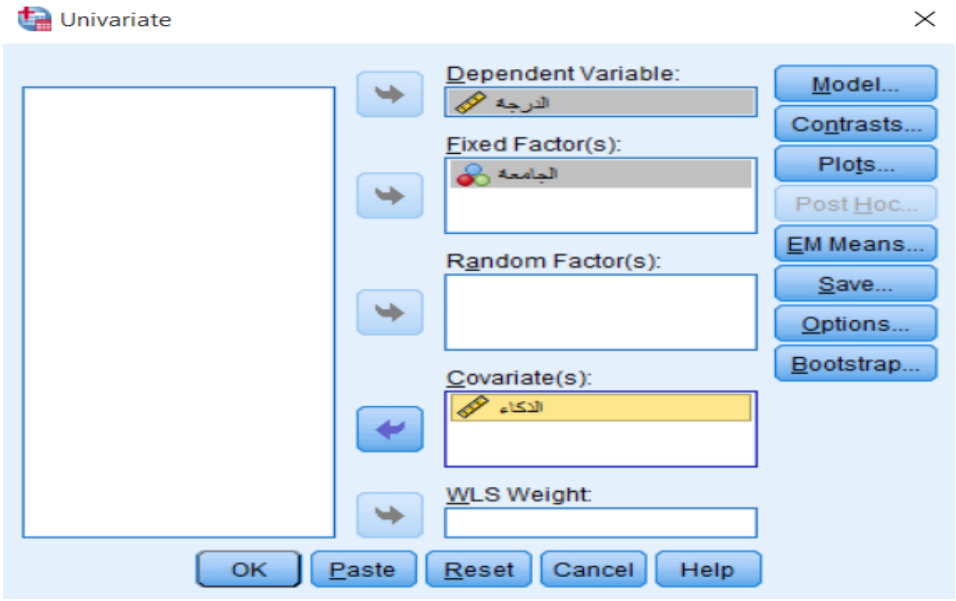
نلاحظ من نتائج اختبار Scheffe ان قيمة .Sig بين المهندس2 والمهندس3 هي الوحيدة معنوية كونها تساوي 0.041 وهي اقل من 0.05 مستوى المعنوية المفترض.

4. تحليل التباين المشترك Analysis of Covariance ANCOVA:

يستخدم تحليل التبيان المشترك ANCOVA عندما نريد مقارنة متوسطات متغير تابع لمجموعتين او أكثر من الافراد بعد ضبط الفروقات بين هذه المجموعات على متغير آخر يسمى المتغير المشترك Covariate. ان الهدف من اجراء هذا التحليل هو محاولة تقليل خطأ التباين. على سبيل المثال نرغب باختبار الفروق بين درجات الإحصاء لطلاب ثلاث جامعات مختلفة، ومع استبعاد أثر الذكاء للطلاب، ومن أجل ذلك سوف نقوم بإجراء تحليل التباين المشترك وفق الخطوات التالية:

أ- ننقر على قائمة Analyze ثم ننقر على General Linear Model ثم Univariate، ستظهر لنا شاشة الحوار التالية:

الشكل رقم (86): مربع الحوار Univariate2



المصدر: مخرجات SPSS

ب- ننقر على المتغير التابع "الدرجة"، ثم ننقر على السهم لتحويله الى مربع Dependent Variable، ثم ننقر على المتغير العاملي "الجامعة"، ثم ننقر على السهم لتحويله الى مربع Fixed Factor(s)، ثم ننقر على المتغير المشترك "الذكاء"، ثم ننقر على السهم لتحويله الى مربع Covariate(s). ثم ننقر OK، ستظهر لنا نتائج اختبار تحليل التباين المشترك في شاشة المخرجات كما في الجداول الموالية:

الجدول رقم (31): مخرجات الاحصائيات الوصفية Univariate–Covariate

Descriptive Statistics

Dependent Variable: الدرجة

الجامعة	Mean	Std. Deviation	N
الوادي	13.8333	3.65605	6
سطيف	12.2500	5.43906	4
خنشلة	12.2000	7.42967	5
Total	12.8667	5.24904	15

المصدر: مخرجات SPSS

قام برنامج SPSS بحساب المتوسط الحسابي (Mean) والانحراف المعياري (Std. Deviation) للمتغير التابع الدرجة، لكل فئة من الفئات الثلاثة للمتغير العاملي: الوادي، سطيف، خنشلة.

الجدول رقم (31): مخرجات اختبار تحليل التباين المشترك ANCOVA

Tests of Between-Subjects Effects

Dependent Variable: الدرجة

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	228.156[a]	3	76.052	5.309	.017
Intercept	2.225	1	2.225	.155	.701
الذكاء	218.806	1	218.806	15.274	.002
الجامعة	14.503	2	7.251	.506	.616
Error	157.578	11	14.325		
Total	2869.000	15			
Corrected Total	385.733	14			

a. R Squared = .591 (Adjusted R Squared = .480)

المصدر: مخرجات SPSS

تشير نتيجة تحليل التباين المشترك الى عدم وجود فروق ذات دلالة إحصائية تعزى للمتغير العاملي الجامعة عند مستوى معنوية 0.05، حيث قيمة Sig. أكبر من مستوى المعنوية المفترض 0.05، أي انه لا يوجد فروق معنوية بين طلاب الجامعات الثلاثة في الاحصاء. بينما تشير نتيجة تحليل التباين المشترك الى وجود فروق ذات دلالة إحصائية للمتغير المشترك الذكاء عند مستوى معنوية 0.05، حيث قيمة Sig. أصغر من مستوى المعنوية المفترض 0.05، أي انه يوجد فروق معنوية في الاحصاء بين الطلاب تعزى للذكاء.

5. أمثلة وتطبيقات:

تطبيق رقم 01: توجد ثلاثة طرق لتدريب الموظفين الجدد، وقد استخدمت كل طريقة لتدريب مجموعة واحدة من ثلاثة مجموعات مستقلة، وبعد انتهاء فترة التدريب اعطي للمجموعات الثلاثة اختبار موحد حيث حصل كل موظف على درجة معينة كما في الجدول التالي:

الطريقة	درجات الموظفين									
الطريقة1	74	78	73	72	72	73	76	81	71	72
الطريقة2	84	77	79	79	81	77	78	80	77	80
الطريقة3	83	85	86	87	89	84	79	86	84	79

المطلوب: – اختبار معنوية الفرق بين متوسطات الطرق التدريبية عند مستوى دلالة 5%.

– في حالة ظهور فروق معنوية بين الطرق التدريبية، اختبر معنوية الفروق بين متوسطي كل طريقتين باستخدام طريقة scheffe عند مستوى دلالة 5%.

تطبيق رقم 02: يرغب باحث اقتصادي في مقارنة الدخل الشهري في أربعة دول عربية، حيث سحب عينة عشوائية حجمها 13 اسرة من كل دولة:

الاولى	2872	2753	2954	2623	2584	2945	2548	2861	3120	2945	2753	2456	2674
الثانية	1753	1648	1523	2241	2154	1582	1654	1945	1783	1658	2135	2040	2145
الثالثة	3125	3521	3471	3654	2970	3578	2846	3214	3346	3965	4100	3159	3095
الرابعة	2571	2468	2123	2587	3753	4200	2546	2863	2368	2469	2541	2943	2864

المطلوب: اجراء تحليل التباين مع المقارنات المتعددة.

تطبيق 03: إذا توفرت لديك البيانات التالية والخاصة بعدد الكيلومترات التي تقطعها كل سيارة في الساعة الواحدة:

	Mercedes	BMW	Fiat	Skoda
عادي	160	155	110	95
بدون رصاص	140	145	140	105
ممتاز	170	165	145	85

المطلوب: هل هناك فرق جوهري في متوسط السرعة:

- ✓ حسب نوع السيارة
- ✓ حسب نوع البنزين

تطبيق رقم 08: لدينا البيانات التالية المتعلقة بنتائج استبيان لمجموعة من الطلبة كما في الجدول التالي:

المشاهدات	الجنس	التخصص	الاعلام الالي	معدل س1	درجة الرضى	الطول	العمر	معدل س2
1	ذ	اقتصاد	متوسط	12.5	راضي	175	21	11.8
2	ذ	حقوق	جيد	11.3	غ راضي	168	22	10.3
3	أ	تكنولوجيا	جيد	9.4	محايد	162	21	11.1
4	أ	النفس	ضعيف	13.5	محايد	170	27	12.5
5	ذ	الاجتماع	ضعيف	14.6	راضي	178	25	15.6
6	أ	تكنولوجيا	متوسط	10.8	راضي محايد	165	24	12.3
7	ذ	حقوق	ضعيف	12.6	غ راضي	180	22	10.2
8	ذ	النفس	متوسط	15.4	محايد	177	21	15.8
9	أ	اقتصاد	جيد	8.6	غ راضي	172	28	10.4
10	أ	الاجتماع	جيد	15.7	راضي	158	23	13.5
11	أ	تكنولوجيا	متوسط	14.2	غ راضي	155	24	12.8
12	ذ	اقتصاد	ضعيف	12.9	غ راضي	168	25	13.6
13	أ	حقوق	متوسط	11.7	محايد	160	21	9.8
14	ذ	حقوق	ضعيف	14.8	راضي	166	22	13.8
15	أ	اقتصاد	ضعيف	14.9	راضي	158	22	14.8
16	أ	النفس	متوسط	16.5	راضي	150	25	15.8
17	أ	الاجتماع	ضعيف	13.3	غ راضي	170	21	12.5
18	ذ	النفس	ضعيف	14.2	غ راضي	182	23	13.8
19	ذ	حقوق	ضعيف	11.8	محايد	175	26	11.5
20	أ	تكنولوجيا	جيد	9.6	محايد	167	24	8.7
21	ذ	اقتصاد	متوسط	10.3	محايد	173	25	9.7

المطلوب: – اختبار الفروق بين معدلي س1 و س2 عند مستوى دلالة 5%.

- اختبار الفروق بين معدلي الذكور والاناث لكل من س1 و س2 عند مستوى دلالة 5%.
- اختبار فرضية ان معدل الطلبة يساوي 12 لكل من س1 و س2 عند مستوى دلالة 5%.

الفصل الثامن:

تحليل الارتباط correlation analysis

1. تمهيد:

يهتم تحليل الارتباط بالطرائق التي يمكن من خلالها إيجاد العلاقة الارتباطية الخطية بين متغيرين أو أكثر. ان قياس نوع ومقدار العلاقة بين المتغيرات يدعى الارتباط correlation، ويقاس الارتباط بين متغيرين بمؤشـــر كمي هو معامل الارتباط.

2. معامل بيرسون للارتباط او معامل الارتباط الخطي البسيط:

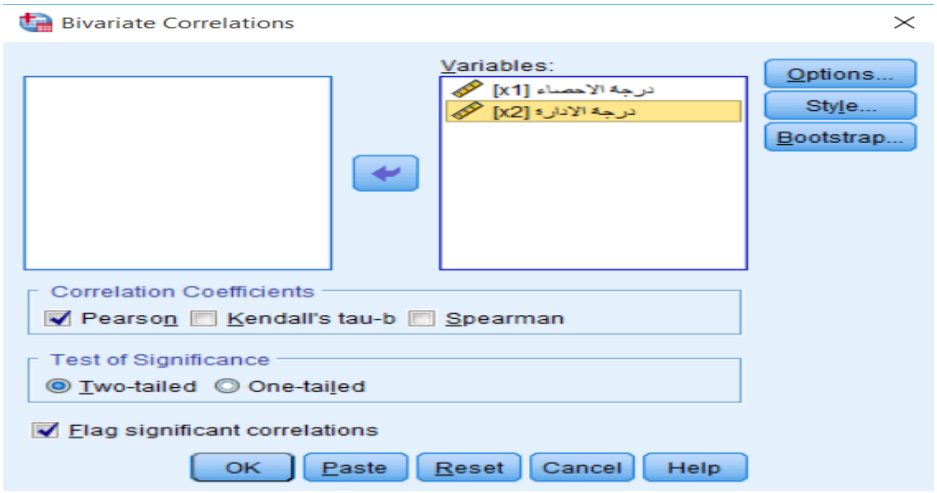
يستخدم معامل بيرسون للارتباط لقياس قوة واتجاه العلاقة الخطية بين متغيرين كميين، ولاعتماد معامل بيرسون يجب توفر الشروط التالية:

- ان يكون كل متغير يتبع التوزيع الطبيعي.
- ان تكون العلاقة خطية بين المتغيرين.
- ان تكون عينة كل متغير مسحوبة بصورة عشوائية.

على ســـبيل المثال لقياس العلاقة الارتباطية بين متغير درجة الإحصـــاء ومتغير درجة الادارة، ويتم ذلك باتباع الخطوات التالية:

أ- نفتح ملف البيانات المطلوب، ثم نختار امر Correlate من قائمة Analyze، ثم نختار امر Bivariate، كما هو موضح في الشكل التالي:

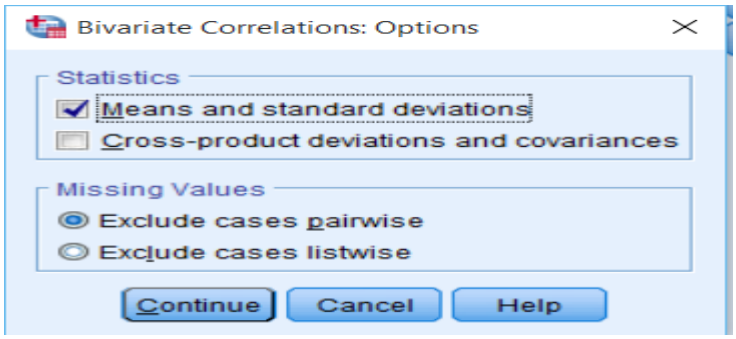
الشكل رقم (87): مربع حوار Bivariate Correlation



المصدر: مخرجات SPSS

ب- نختار المتغيرين من قائمة المتغيرات وبواسطة السهم نقوم بتحويلهما الى مربع Variables، ونختار نوع معامل الارتباط الذي نرغب في استخدامه من مربع Correlation Coefeicients، ثم ننقر على Options فتحصل على مربع الحوار التالي:

الشكل رقم (88): مربع حوار Bivariate Correlation: Options



المصدر: مخرجات SPSS

ت- من مربع Statistics نختار Means and standars deviations وذلك لحساب المتوسطات الحسابية والانحرافات المعيارية لكل متغير. ننقر على Continue لنعود الى مربع الحوار Bivariate Correlation ثم ننقر OK لنتحصل على نتائج هذا الاجراء الاحصائي في شاشة المخرجات كما هو موضح فيما يلي:

الجدول رقم (32): الإحصاءات الوصفية لمتغيري "درجة الإحصاء" و"درجة الإدارة"

Descriptive Statistics

	Mean	Std. Deviation	N
درجة الاحصاء	25.8000	3.44814	30
درجة الادارة	29.5000	2.92138	30

المصدر: مخرجات SPSS

لقد قام البرنامج بحساب الإحصاءات الوصفية للمتغيرين (المتوسط الحسابي والانحراف المعياري).

الجدول رقم (33): اختبار بيرسون للارتباط بين متغيري "درجة الإحصاء" و"درجة الإدارة"

Correlations

		درجة الاحصاء	درجة الادارة
درجة الاحصاء	Pearson Correlation	1	.637**
	Sig. (2-tailed)		.000
	N	30	30
درجة الادارة	Pearson Correlation	.637**	1
	Sig. (2-tailed)	.000	
	N	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

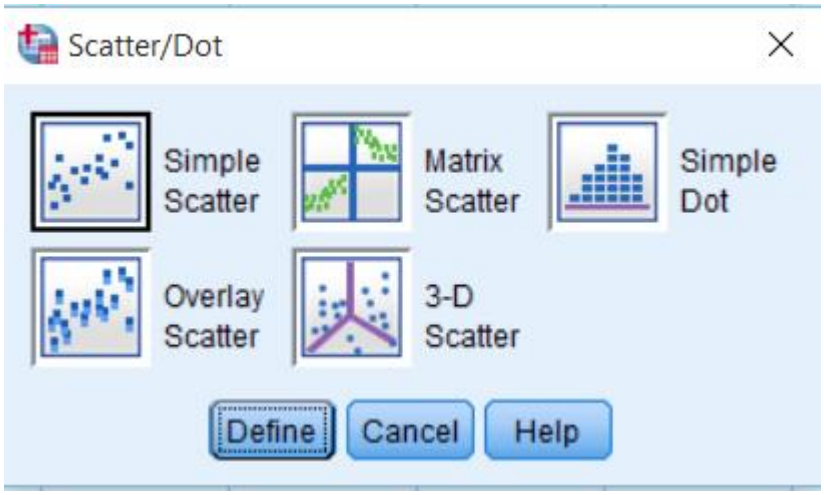
المصدر: مخرجات SPSS

يتضح من نتائج الجدول أعلاه ان معامل الارتباط بيرسون بين متغير درجة الإدارة ومتغير درجة الإحصاء بلغ 0.637، وهو معنوي عند مستوى معنوية 0.01، حسب ما تشير اليه Sig. (2-tailed).

يمكن كذلك تمثيل نتائج الارتباط من خلال الرسوم البيانية، باستخدام لوحة الانتشار Scatter Plot لتمثيل شكل وقوة العلاقة بين متغيرين كميين بيانيا ولإجراء ذلك نتبع الخطوات التالية:

ث- ننقر فوق قائمة Graphs ثم Legacy Dialogs ثم ننقر Scatter/Dot ستظهر لنا شاشة حوار Scatter/Dot كما هي مبينة اسفله:

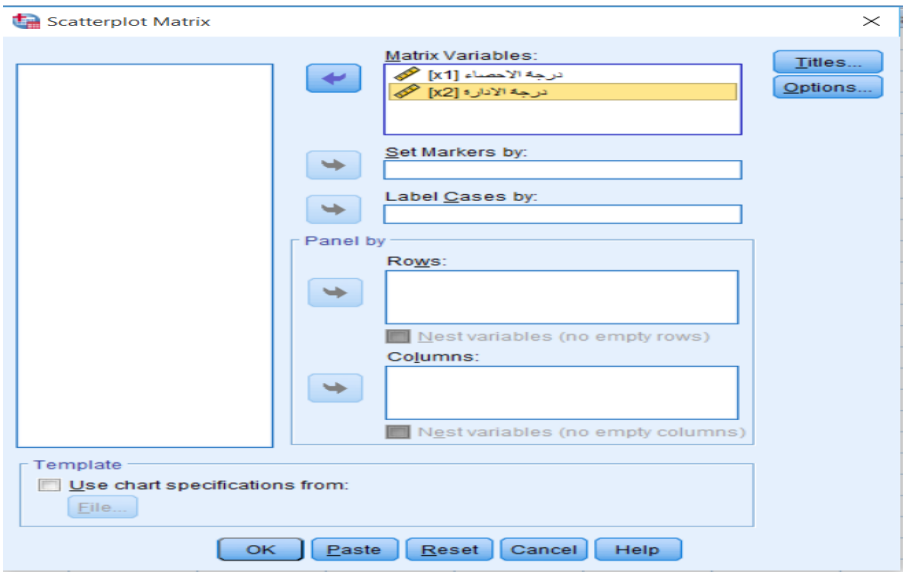
الشكل رقم (89): مربع حوار Scatter/Dot



المصدر: مخرجات SPSS

ج- ننقر فوق Matrix ثم ننقر فوق Define مفتاح سيظهر لنا مربع حوار كما يلي:

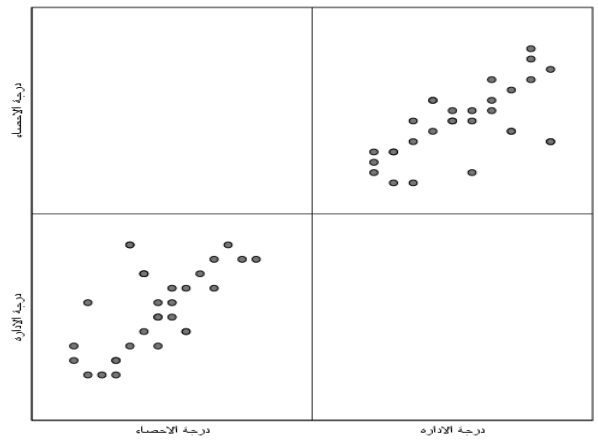
الشكل رقم (90): مربع حوار Scatterplot Matrix



المصدر: مخرجات SPSS

ح- ننقر فوق المتغيرين المراد فحصهما، ونقوم بتحويلهما بواسطة السهم الى مربع Matrix Variables ثم ننقر OK لنتحصل على نتائج هذا الاجراء الاحصائي في شاشة المخرجات كما هو موضح فيما يلي:

الشكل رقم (91): الرسم البياني Matrix Scatter لمتغيري "درجة الإحصاء" و"درجة الإدارة"



المصدر: مخرجات SPSS

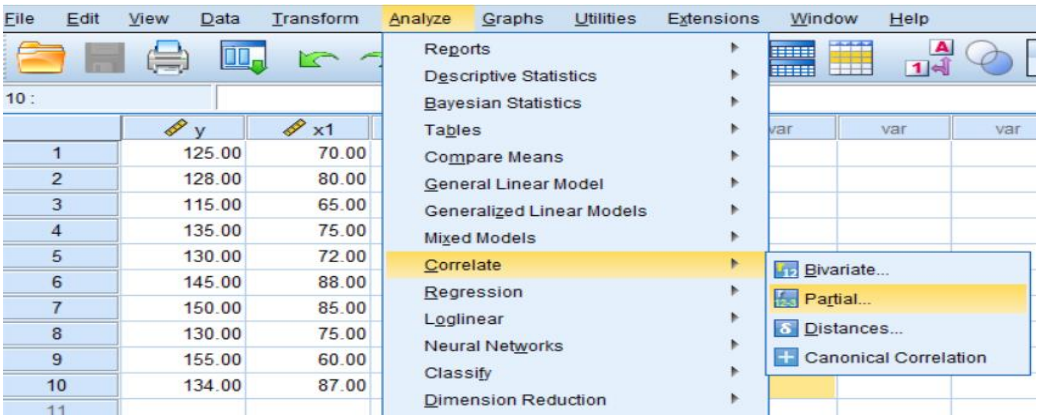
3. الارتباط الجزئي Partial Correlation:

يفضل استخدام معامل الارتباط الجزئي على معامل الارتباط البسيط في كثير من البحوث العلمية، ذلك لأن معامل الارتباط الجزئي يبين نسب تأثر المتغير التابع بمتغير مستقل معين مع ثبات باقي المتغيرات المستقلة (المفسرة) الأخرى على المتغير التابع.

على سبيل المثال لقياس الارتباط الجزئي بين ضغط الدم ووزن الجسم، بعد استبعاد أثر العمر، ومعامل الارتباط الجزئي بين ضغط الدم والعمر، بعد استبعاد أثر وزن الجسم. ويتم ذلك باتباع الخطوات التالية:

أ- نفتح ملف البيانات المطلوب، ثم نختار امر Correlate من قائمة Analyze، ثم نختار امر Partial، كما هو موضح في الشكل التالي:

الشكل رقم (92): قائمة Analyze: Correlate: Partial



المصدر: مخرجات SPSS

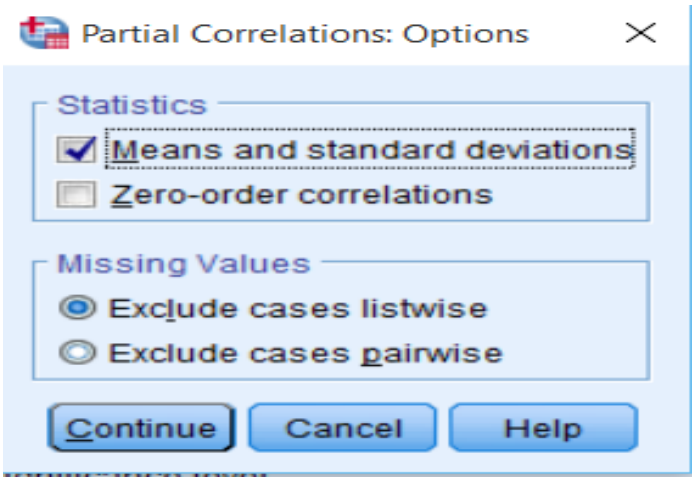
ب- يظهر بعد ذلك مربع الحوار Partial Correlations.

الشكل رقم (93): مربع حوار Partial Correlations

المصدر: مخرجات SPSS

ت- نختار المتغيرين ضغط الدم ووزن الجسم من قائمة المتغيرات وبواسطة السهم نقوم بتحويلهما الى مربع Variables، ونختار متغير العمر المراد استبعاد أثره وبواسطة السهم نقوم بتحويله الى مربع Controlling for، ثم ننقر على Options فتحصل على مربع الحوار التالي:

الشكل رقم (94): مربع حوار Partial Correlations: Options



المصدر: مخرجات SPSS

ث- من مربع Statistics نختار Means and standars deviations وذلك لحساب المتوسطات الحسابية والانحرافات المعيارية لكل متغير. ننقر على Continue لنعود الى مربع الحوار Partial Correlation ثم ننقر OK لنتحصل على نتائج هذا الاجراء الاحصائي في شاشة المخرجات كما هو موضح فيما يلي:

الجدول رقم (34): الإحصاءات الوصفية للمتغيرات "ضغط الدم"، "وزن الجسم" و"العمر"

Descriptive Statistics

	Mean	Std. Deviation	N
ضغط الدم	134.7000	12.12939	10
وزن الجسم	75.7000	9.38142	10
العمر	36.6000	15.12320	10

المصدر: مخرجات SPSS

لقد قام البرنامج بحساب الإحصاءات الوصفية للمتغيرين (المتوسط الحسابي والانحراف المعياري).

الجدول رقم (35): اختبار الارتباط الجزئي بين ضغط الدم ووزن الجسم، بعد استبعاد أثر العمر

Correlations

Control Variables			ضغط الدم	وزن الجسم
العمر	ضغط الدم	Correlation	1.000	.671
		Significance (2-tailed)	.	.048
		df	0	7
	وزن الجسم	Correlation	.671	1.000
		Significance (2-tailed)	.048	.
		df	7	0

المصدر: مخرجات SPSS

يتضح من نتائج الجدول أعلاه ان معامل الارتباط بيرسون بين متغير درجة الإدارة ومتغير درجة الإحصاء بلغ 0.671، وهو معنوي عند مستوى معنوية 0.05، حسب ما تشير اليه Sig. (2-tailed).

4. معامل الارتباط سبيرمان للرتب Spearman's Cofficient of Rank Correlation:

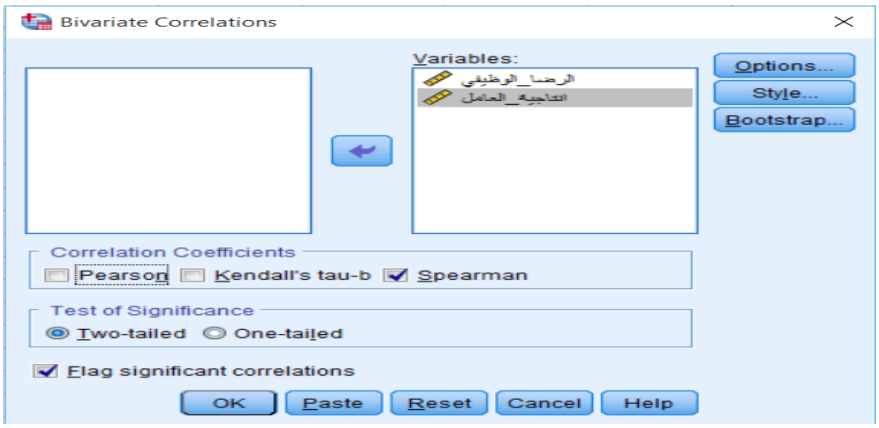
يستخدم لقياس قوة واتجاه العلاقة الخطية بين متغيرين في الحالات التالية:

- إذا كان كلا المتغيرين او أحدهما من النوع الرتبي.
- إذا كان كلا المتغيرين او أحدهما لا يتبع التوزيع الطبيعي، او في حالة البيانات اللامعلمية.

على سبيل المثال لقياس ارتباط سبيرمان للرتب بين الرضا الوظيفي وانتاجية العامل، ويتم ذلك باتباع الخطوات التالية:

أ- نفتح ملف البيانات المطلوب، ثم نختار امر Correlate من قائمة Analyze، ثم نختار امر Bivariate، كما هو موضح في الشكل التالي:

الشكل رقم (95): مربع حوار Bivariate Correlations



المصدر: مخرجات SPSS

ب- نختار المتغيرين من قائمة المتغيرات وبواسطة السهم نقوم بتحويلهما الى مربع Variables، ونختار نوع معامل الارتباط Spearman من مربع Correlation Coefeicients، ننقر على Continue لنعود الى مربع الحوار Bivariate Correlation ثم ننقر OK لنتحصل على نتائج هذا الاجراء الاحصائي في شاشة المخرجات كما هو موضح فيما يلي:

الجدول رقم (36): اختبار Spearman بين متغيري "الرضا_الوظيفي" و"انتاجية_العامل"

Correlations

			الرضا_الوظيفي	انتاجية_العامل
Spearman's rho	الرضا_الوظيفي	Correlation Coefficient	1.000	.714*
		Sig. (2-tailed)	.	.047
		N	8	8
	انتاجية_العامل	Correlation Coefficient	.714*	1.000
		Sig. (2-tailed)	.047	.
		N	8	8

*. Correlation is significant at the 0.05 level (2-tailed).

المصدر: مخرجات SPSS

يتضح من نتائج الجدول أعلاه ان معامل الارتباط سبيرمان بين متغير الرضا الوظيفي ومتغير إنتاجية العامل بلغ 0.714، وهو معنوي عند مستوى معنوية 0.05، حسب ما تشير اليه Sig. (2-tailed).

5. أمثلة وتطبيقات:

التطبيق 01:

إذا توفرت البيانات التالية:

A	10	50	30	20	60	40
B	2	8	5	3	10	4
C	15	25	18	16	30	28
D	15	20	30	50	25	40

اوجد معامل الارتباط الجزئي ومعنويته لكل المتغيرات بعد استبعاد أثر المتغير **D**.

التطبيق 02:

قام أحد الباحثين بدراسة أسباب كفاءة الأشخاص في قيادة المركبات فاعتمد على متغيرين هما عدد سنوات القيادة وشخصية السائق. وقد أجري اختبار لشخصية السائق تضمن (الذوق/ الفن/ الاخلاق) والجدول التالي يبن نتيجة الاختبار:

كفاءة القيادة	عدد سنوات القيادة	شخصية السائق
95	25	4
88	23	4
65	12	2
73	8	3
99	30	5
82	12	3
90	12	5
85	7	2

- اختبار هل ان المتغيرات تتبع التوزيع الطبيعي.
- هل ان العلاقة خطية، بين متغير كفاءة القيادة وكل من عدد سنوات القيادة وشخصية السائق؟
- جد قيمة معامل الارتباط البسيط بين كفاءة القيادة وكل من عدد سنوات القيادة وشخصية السائق.
- جد قيمة معامل الارتباط الجزئي بين كفاءة القيادة وعدد سنوات القيادة بعد استبعاد شخصية السائق.

- جد قيمة معامل الارتباط الجزئي بين كفاءة القيادة وشخصية السائق بعد استبعاد عدد سنوات القيادة.
- تفسير النتائج.

التطبيق 03:

أجرى مركز البحوث العلمية لإحدى الدوائر دراسة لنسب الاعمال المنجزة وعلاقتها بكل من:

- طبيعة قرارات المدير.
- مهارة العاملين.
- عدد الحوافز التشجيعية السنوية.

فكانت النتائج كما يلي:

نسبة الاعمال المنجزة	قرارات المدير	مهارة العاملين	الحوافز التشجيعية
0.8	3	3	2
0.7	2	4	1
0.75	4	4	1
0.6	1	2	0
0.55	2	2	0
0.75	5	2	1
0.5	4	1	0

- جد مصفوفة الارتباط بين المتغيرات أعلاه.
- تحديد نوع ومعنوية الارتباط لبيرسون عند مستوى دلالة 0.05.