خدمات أكاديمية
كفاءات وطنية
معايير عالمية
دراسة
للاستشارات والدراسات والترجمة
UNIVERSITY
drasah1
00966555026526
00966560972772
www.drasah.com
info@drasah.com


شركة دراسة
للاستشارات و الدراسات والترجمة
خدماتنا
توفير المراجع العربية والأجنبية
التحليل الاحصائي وتفسير النتائج
الاستشارات الأكاديمية
جمع المادة العلمية
الترجمة المعتمدة
drasah1
info@drasah.com
00966555026526
00966560972772
drasah.com


دراسة
للاستشارات والدراسات والترجمة
تواصل معنا
00966555026526
00966560972772
متواجدون على مدار الســـاعة
drasah.com

الجامعة الإسلامية/ غزة

كلية الآداب

قسم الجغرافيا ونظم المعلومات الجغرافية

الاختبارات الإحصائية

الاختبارات المعلمية واللامعلمية

الجزء الرابع

من محاضرات الاحصاء الجغرافي بالحاسوب

د. كامل أبو ضاهر

2020

الاختبارات الإحصائية

1) تعريفها وشروط استخدامها

2) الاختبارات المعلمية

أ- اختبار t لعينة واحدة

ب- اختبار t لعينتين مستقلتين

ج- اختبار t لعينتين مرتبطتين

د- اختبار تحليل التباين الأحادي

3)الاختبارات اللامعلمية

أ- الاختبارات اللامعلمية للبيانات الكمية التي ليست لها توزيع طبيعي والبيانات الرتبية

1) اختبار الإشارة لعينة واحدة
2) اختبار مان وتني لعينتين مستقلتين
3) اختبار ويلكوكسن لعينتين مرتبطتين
4) اختبار كروسكال ولاس لثلاث عينات فأكثر

ب- الاختبارات اللامعلمية للبيانات التصنيفية والنوعية

1) اختبار مربع كاي
2) اختبار فاي

الاختبارات الاحصائية كثيرة ومتنوعة ومنها اختبار الفرق بين معدل المجتمع الاحصائي والمتوسط الحسابي لعينة واحدة، والفروق بين المتوسطات الحسابية لعينات مأخوذة من مجتمعات إحصائية مختلفة. وتنقسم الاختبارات الاحصائية إلى اختبارات معلمية للبيانات الكمية ذات التوزيع الطبيعي واختبارات لامعلمية للبيانات الرتبية والبيانات الكمية التي ليس لها توزيع طبيعي واختبار مربع كاي للبيانات التصنيفية.

1) تعريفها وشروط استخدامها :

الفرضية العلمية: عبارة عن حدس علمي أو تفسير أولي لأي ظاهرة نقوم بدراستها، أو أنها تمثل العلاقة القائمة بين عدد من الظواهر وتقدم تفسيراً مبدئياً لها. وليس كل التصورات الأولية بالضرورة فرضية علمية.

شروط الفرضية العلمية :

- أن تكون جزءا من بحث علمي متكامل.
- صياغة الفرضية بحيث تكون قابلة للاختبار.

فإذا اختبرت الفرضية وثبت صدقها تصبح نموذج علمي

الفرضيات الإحصائية : تشكل الفرضيات ركناً هاماً في البحث العلمي ووسيلة للوصول إلى تعميمات تفسر الظاهرة الإعلامية، فالباحث يبحث في مجالات مختلفة، ويضع مجموعة من التساؤلات التي يجب أن يضع لها فرضيات تجيب عن هذه التساؤلات.

مصادر الفرضيات

أ- الدراسة الميدانية. ب- الدراسات السابقة. ج- الملاحظات الشخصية والخبرة العملية.

الشروط التي يجب توافرها في الفرضيات

1) صياغة الفرضيات بعبارات سهلة وبسيطة وواضحة.
2) أن تكون جزءاً من خطة متكاملة للبحث العلمي.
3) إمكانية اختبارها والتأكد من صدقها وثباتها.
4) أن لا تتعارض مع الحقائق العلمية.

5) أن يكون لها قدرة تفسيرية.

6) أن يكون لها نتيجة واحدة واضحة ومحددة.

الإطار العام لاختبار الفرضيات الإحصائية:

الفرضيات الإحصائية: هي فرضيات يضعها الباحثون عندما يستخدمون خصائص العينات، لتقدير معالم المجتمعات الإحصائية التي أخذت منها، أو عندما يوازنون بين المعالم الإحصائية لمجتمعات متعددة مستخدمين عينات مختارة من تلك التجمعات مستخدمين عينات مختارة من تلك المجتمعات.

ترتبط الفرضيات بـ:

- مفاهيم إحصائية خاصة بفئات الثقة ومستويات المعنوية
- يتم اختبارها باختبارات إحصائية مناسبة .
- اختيار أسلوب التحليل الإحصائي

خطوات اختبار الفرضيات :

1- تحديد فرضية العدم أو الفرضية المبدئية للبحث(H_0) وهي الفرضية التي سيتم اختبارها إحصائياً، وهي تمثل عكس ما يتوقعه الباحث.

مثال على صياغة فرضية العدم :

- لا توجد علاقة بين متغير الجنس ومتغير الرضا عن الخدمات البلدية.
- لا يوجد اختلاف بين متوسط دخل الأسرة الريفية والأسرة الحضرية في محافظة خانيونس.
- لا توجد علاقة بين متوسط دخل الفرد ومشاركته في الانتخابات النيابية.

2- **الفرضية البديلة(H_1)** : و هي الفرضية التي تمثل توقعات الباحث لمشكلة البحث وهي تتناقض تماماً مع الفرضية المبدئية .

تتضمن الفرضية البديلة توقعاتنا بشأن النتائج المحتملة للبحث والتي تشمل الموازنة بين أي خاصيتين إحصائيتين (أ) و (ب) أو بين خاصية إحصائية لعينة ومعلم من معالم المجتمع الإحصائي، الذي أخذت منه تلك العينة النتائج التالية :

- (أ) و (ب) **متساويتان**
- **(أ) أكبر من (ب)**
- **(أ) أصغر من (ب)**

مثال 1 :

أ – **الفرضية البديلة** :على أن النساء يتقاضين في بداية عملهن، بعد تخرجهن من الجامعة رواتب تقل عن رواتب الرجال

ب– **الفرضية المبدئية** : تتضمن كل الاحتمالات الأخرى :

- أن النساء تتقاضى رواتب مساوية لرواتب الرجال
- أن النساء تتقاضى رواتب أعلى من رواتب الرجال

مثال 2 :

الفرضية البديلة: أن المشاركة في الانتخابات النيابية في المدن تختلف عن المشاركة في الريف

الفرضية المبدئية : تتضمن:

- أن المشاركة في الانتخابات النيابية في المدن أقل من المشاركة في الريف.
- أن المشاركة في الانتخابات النيابية لا تتأثر بمكان السكن.

3– **تحديد مستوى الدلالة (α)** : وهو يمثل احتمال خطأ من النوع الأول ، وهو يمثل مستوى عدم الثقة في التقدير الذي نحصل عليه، أو يمثل احتمالية أن نكون مخطئين عند رفضنا لفرضية العدم وقبولنا للفرضية البديلة، وتستخدم أغلب الدراسات الإحصائية مستوى الدلالة **0.05 ، 0.01**

4–اختيار توزيع المعاينة المناسب لإجراء الاختبار وتحديد منطقة الرفض، وهي قيمة محددة تستخرج من جداول خاصة وتعرف باسم القيمة الحرجة .

5– **حساب الخاصية الاختبارية** Test statistic: وهي تمثل قيمة محددة تختلف من اختبار لآخر.

6- موازنة الخاصية الإحصائية بالقيمة الحرجة التي تم تحديدها، وإصدار قرار بشأن رفض فرضية العدم، أو عدم التمكن من ذلك.

ملاحظة : إذا تمكنا من رفض فرضية العدم حسب نتائج الاختبار الإحصائي، فإننا نقبل الفرضية البديلة، ونكون قد توصلنا إلى النتائج التي كنا نتوقعها، والتي تتفق مع الأساس النظري للموضوع قيد الدراسة .

أما إذا لم نتمكن من رفض فرضية العدم، فإننا لا نستطيع أن نقبل الفرضية البديلة، أو بعبارة أصح لا نستطيع أن نصدر حكماً بشأنها، ولا يكون لا يجوز في جميع الأحوال، مهما كانت نتائج الاختبار الإحصائي أن يكون قرارنا هو قبول الفرضية المبدئية.

أسباب عدم رفض فرضية العدم:

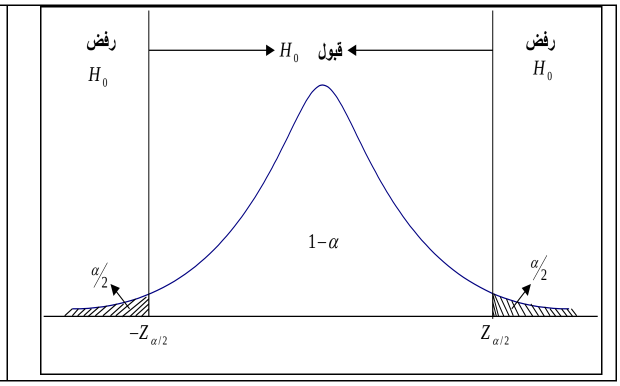
1. **خطأ في تصميم العينة :** فالعينات الإحصائية متعددة فمنها العينات العمدية والعينات العشوائية، والعينات العشوائية أنواع: فمنها العينات البسيطة والمنتظمة والطبقية وغيرها، فكل مجتمع إحصائي له نوع مناسب من العينات يمثله تمثيلاً صحيحاً.
2. **أسلوب اختيار العينة :** هل العينة عشوائية أم عينة عمدية
3. **حجم العينة:** تختلف أحجام العينات حسب نوع الدراسة ونوع المجتمع الإحصائي وحجمه، وبالتالي يؤثر حجم العينة في نتائج الاختبار الإحصائي.
4. عدم ملائمة **الاختبار الإحصائي** المستخدم لموضوع الدراسة.

أخطاء الاختبارات الإحصائية :

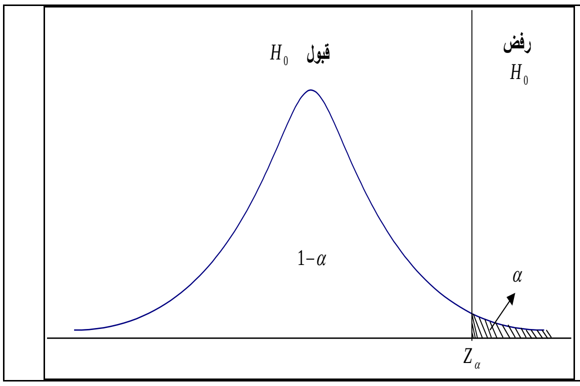
1- خطأ من النوع الأول (α): يحدث عندما يرفض الباحث فرضية العدم بالرغم من كونها فرضية صحيحة .

2- خطأ من النوع الثاني(B): الفشل في رفض فرضية العدم بالرغم من كونها فرضية خاطئة

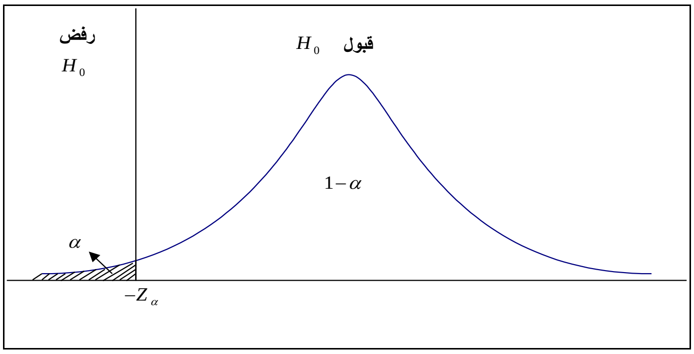
3- قوة الاختبار: القدرة على رفض الفرضية المبدئية عندما تكون تلك الفرضية خاطئة، وهي تساوي $(1 - B)$.



$H_1: \mu \neq \mu_0$



$H_1: \mu \succ \mu_0$



$H_1: \mu < \mu_0$

مناطق القبول والرفض

العوامل التي تؤثر في اختيار الأسلوب الإحصائي:

1- مشكلة البحث وأهداف الدراسة والمنهجية المستخدمة تؤثر في تحديد الأسلوب الإحصائي.

2- خصائص البيانات (هل تتبع التوزيع الطبيعي أم لا، حجم العينة).

3- خصائص الأساليب الإحصائية.   4- خلفية الباحث وفلسفته.

فحص البيانات قبل تطبيق التحليلات الإحصائية المتقدمة:

المخاطر الناشئة عن استخدام أساليب التحليل الإحصائي المتقدمة:

أ- فشل الباحث في الفهم الصحيح لبيانات البحث.

ب- استخدام أساليب غير مناسبة للتحليل تؤدي إلى نتائج خاطئة.

من الأساليب المستخدمة في فحص البيانات:

أولاً / فحص شكل التوزيع : تستخدم هذه الطرق لمعرفة هل البيانات تتبع التوزيع الطبيعي.

1- استخدام التكرار للحصول على المدرج التكراري لمتغير واحد.

2- استخدام مقاييس النزعة المركزية (المتوسط والوسيط والمنوال)، حيث في التوزيع الطبيعي تتساوى القيم الثلاث.

3- اختبارات متاحة في البرامج الاحصائية مثل: K-S, Shapiro-Wilks.

يفضل استخدام كلاهما في نفس الوقت. السبب:

1-اختبار المعنوية غير مفيد عندما تكون العينة أقل من 30.

2-اختبار حساس للعينة التي تزيد عن 1000 مفردة.

شروط تطبيق الاختبارات الإحصائية :

يجب أن نأخذ بعين الاعتبار عند تطبيق أي اختبار إحصائي بأن لكل اختبار إحصائي شروط يجب توافرها، حتى يتم تطبيقه بشكل سليم، ويؤدي إلى نتائج صحيحة، ومن شروط تطبيق الاختبارات الإحصائية:

1- نوعية البيانات المستخدمة وطبيعتها: فالبيانات الكمية يصلح لها اختبارات معلمية Parametric tests، بينما البيانات النوعية والتصنيفية يصلح لها اختبارات غير معلمية Nonparametric tests .

2- طبيعة توزيع المعاينة: الاختبارات المعلمية تشترط بأن تكون العينة المستخدمة عينة عشوائية

الاختبارات الإحصائية

أنواع البيانات	**البيانات الكمية/ توزيع طبيعي**	**البيانات الرتبية + البيانات الكمية /ليست توزيع طبيعي**	**البيانات التصنيفية**
العينة واحدة	اختبار t لعينة واحدة	اختبار الإشارة	اختبار مربع كاي اختبار فاي
عينتان مستقلتان	اختبار t لعينتين مستقلتين	اختبار مان وتني	
عينتان مرتبطتان	اختبار t لعينتين مرتبطتين	اختبار ويلكوكسن	
ثلاث عينات وأكثر	اختبار تحليل التباين	اختبار كروسكال ولاس	

الاختبارات المعلمية

تستخدم الاختبارات المعلمية للبيانات الكمية ذات التوزيع الطبيعي، وحتى نستطيع استخدام الاختبارات المعلمية لا بد أن نتأكد أن البيانات تتبع التوزيع الطبيعي، باستخدام اختبار كولمجروف أو اختبار شابيرو.

تعتبر الاختبارات المعلمية Parametric Test واحدة من أهم التطبيقات التي قدمها علم الإحصاء بشتى العلوم، حيث أن الإحصاء المعلمي هو أحد أنواع الأساليب الإحصائية الاستدلالية التي تهتم بالكشف والاستدلال على معالم المجتمع اعتماداً على ما توافر من بيانات لدى الباحث خاصة بالعينة المأخوذة من هذا المجتمع، كما تتعادل أساليب اتخاذ القرارات الاحصائية. تستخدم الاختبارات المعلمية في حالة العينات الكبيرة التي يشترط توافر المعلومات عن مجتمعاتها مثل:

1) أن يكون توزيع البيانات توزيعاً طبيعياً، وأن يتم اختبارها للتأكد من توزيعها الطبيعي.
2) اختبار مدى تجانس التباين بين العينات، واختبار إذا ما كان يوجد تجانس أم لا.
3) يشترط في الاختبارات المعلمية أن تكون العينات العشوائية.
4) أن تكون العلاقة بين المتغيرات علاقة خطية.
5) يشترط في اختبار t أن تكون العينتين مستقلتين أو مرتبطتين
6) يستخدم فقط مع العينات التي تكون عددية حقيقية.

مميزات الأساليب الاحصائية المعلمية:

1) تستخدم الاختبارات المعلمية في حالة العينات الكبيرة.
2) أن الإحصاء المعلمي يكون أدق وأكثر كفاءة من الإحصاء اللامعلمي.
3) يستخدم في حالة توفر الشروط الاحصائية الخاصة بتحليل التباين.
4) تتناسب الاختبارات المعلمية مع مستويات القياس العليا الفترية والنسبية.

عيوب الأساليب الإحصائية المعلمية:

1) الاحصاء المعلمي يعتبر أكثر صعوبة من الاختبارات الإحصائية اللامعلمية.
2) محدودية نوع البيانات التي يمكن تطبيقها في الاحصاء المعلمي.
3) الاختبارات المعلمية تعتمد على فروض كثيرة تحتاج لتفهمها واستيفائها إلى إحصائي متخصص.
4) تستخدم في البيانات الحقيقية.

أولاً : اختبار t ‏(T Test)

يعد اختبار t من أكثر اختبارات الدلالة شيوعاً في أبحاث العلوم الاجتماعية ومنها بحوث الرأي العام والاعلام، ومن أهم المجالات التي يستخدم فيها هذ الاختبار الكشف عن الفروق بين متغيرين لظاهرة ما، وذلك عن طريق حساب دلالة فرق المتوسطين بين المتغيرين. ويستخدم لقياس دلالة فروق المتوسطات المستقلة والمرتبطة للعينات المتساوية وغير المتساوية.

أ- شروط استخدام اختبار t لدلالة الفروق بين المتوسطات:

1) **حجم كل عينة:** يجب أن يزيد حجم كل من العينتين عن "5" حالات، ويفضل أن يزيد عن "30" حالة .
2) **الفرق بين حجم عينتي البحث:** تقارب بين حجمي العينتين، فالتفاوت الكبير في حجم العينتين أو المجموعتين يؤدي إلى نتائج مضللة، لأن القيمة الحرجة المستخرجة من جدول t تتأثر بحجم العينة لأن درجات الحرية تعتمد على عدد مفردات العينة.
3) **مدى تجانس العينتين:** يقصد بتجانس العينات مدى انتسابها إلى أصل واحد أو أصول متعددة.
4) **مدى اعتدالية التوزيع التكراري لكل من عينتي البحث:** يقصد بها أن البيانات خالية من القيم المتطرفة أو العشوائية وأن منحنى البيانات معتدل ويشبه شكل الجرس.

ب- استخدامات اختبار t :

1) يستخدم في حالة عينة واحدة وذلك بأخذ الفروق بين متوسط العينة ومعدل المجتمع الذي أخذت منه العينة. ويشترط معرفة الانحراف المعياري للمجتمع وأن تكون العينة ذات توزيع طبيعي.
2) يستخدم في حالة عينتين مستقلتين وذلك بقياس الفروق بين متوسطي عينتين مأخوذتين من مجتمعين مستقلين، وأن يكون حجم العينتين متقاربين، ومعرفة التباين للعينتين.
3) يستعمل في حالة أخذ عينتين مرتبطتين وذلك بقياس الفرق بين متوسطي العينتين قبل وبعد التأثير من الظاهرة المدروسة، وهو من أهم الاختبارات المعلمية التي تقيس التغيرات في الرأي العام، وهو يقيس التغيرات التي حدثت في آراء المبحوثين قبل التغيرات التي ننوي إجراؤها على العينة.

2- اختبار t لعينة واحدة

يستخدم اختبار t لاختبار إذا ما كان المتوسط الحسابي للعينة مختلفاً عن المتوسط الحسابي للمجتمع الذي أخذت منه العينة، ويشترط في حالة استعماله أن يكون حجم العينة معروفاً، والمتوسط الحسابي والانحراف المعياري للعينة معروفاً، كما يشترط معرفة معدل المجتمع الاحصائي، وأن يكون التوزيع لبيانات العينة توزيعاً معتدلاً.

مثال 1: عينة مأخوذة من مربي مزارع الدواجن في محافظات غزة يوضح عدد الوفيات في الدجاج/ألف دجاجة لشهر يناير

رقم المزرعة	1	2	3	4	5	6	7	8	9	10
أعداد الوفيات/ ألف دجاجة	**20**	**50**	**30**	**100**	**10**	**70**	**80**	**40**	**20**	**15**

فإذا كان حسب إحصاءات وزارة الزراعة متوسط الوفيات في الدجاج في شهر يناير هو 70 دجاجة لكل ألف دجاجة.

المطلوب : هل يختلف متوسط عدد الوفيات لهذه العينة عن 70 ؟ وهو المعدل العام لوفيات الدجاج في شهر يناير حسب بيانات وزارة الزراعة.

أولاً / الحل بالطريقة الحسابية اليدوية :

رقم المزرعة	1	2	3	4	5	6	7	8	9	10
أعداد الوفيات/ ألف دجاجة	**20**	**50**	**30**	**100**	**10**	**70**	**80**	**40**	**20**	**15**
X^2	**400**	**2500**	**900**	**10000**	**100**	**4900**	**6400**	**1600**	**400**	**225**

مجموع x = 435 مجموع x^2 = 27425

المتوسط الحسابي =

$$\bar{\mathbf{X}} = \frac{\sum x}{N} = \frac{435}{10} = 43.5$$

الانحراف المعياري	
	$\sigma = \sqrt{\frac{\sum X^2 - N\bar{X}^2}{N-1}} = \sqrt{\frac{27425 - 10(43.5)^2}{10-1}} = \sqrt{\frac{27425 - 10(1892.25)}{9}}$ $= \sqrt{\frac{27425 - 18922.5}{9}} = \sqrt{\frac{8502.5}{9}} = \sqrt{944.72} = 30.7$

1- **فرضية العدم:** لا يوجد اختلاف بين المتوسط الحسابي للعينة والمعدل العام لوفيات الدجاج في شهر يناير

2- **الفرضية البديلة:** يوجد اختلاف بين المتوسط الحسابي للعينة والمعدل العام لوفيات الدجاج في شهر يناير

بما أن المتوسط الحسابي للعينة معروف (43.5)، والانحراف المعياري للعينة معروف(30.7)، والمعدل العام لوفيات الدجاج (المجتمع) معروف (70) ، وحجم العينة أقل من 10 مفردة ، فلذلك يمكننا استخدام اختبار t

3- ايجاد قيمة t المحسوبة :

$$t = \frac{(\bar{x} - \mu)}{S}\sqrt{N} = \frac{(43.5 - 70)}{30.7}\sqrt{10} = \frac{(-26.5)}{30.7} * 3.16 = 0.86 * 3.16 = 2.73$$

4- ايجاد قيمة اختبار t الجدولية، فإذا كان مستوى المعنوية 0.05

أ- درجات الحرية = N – 1 = 10 – 1 = 9

ب- مستوى الدلالة = 0.05

ج- فإن قيمة t الجدولية= **1.83**

0.01	0.03	0.05	0.1	0.15	0.2	0.25	0.5	مستوى الدلالة
								Df
31.82	12.71	6.31	3.08	1.96	1.38	1	0	1
6.97	4.3	2.92	1.89	1.39	1.06	0.82	0	2
4.54	3.18	2.35	1.64	1.25	0.98	0.77	0	3
3.75	2.78	2.13	1.53	1.19	0.94	0.74	0	4
3.37	2.57	2.02	1.48	1.16	0.92	0.73	0	5
3.14	2.45	1.94	1.44	1.13	0.91	0.72	0	6
3	2.37	1.9	1.42	1.12	0.9	0.71	0	7
2.9	2.31	1.86	1.4	1.11	0.89	0.71	0	8
2.82	2.26	1.83	1.38	1.1	0.88	0.7	0	9
2.76	2.23	1.81	1.37	1.09	0.88	0.7	0	10
2.72	2.2	1.8	1.36	1.09	0.88	0.7	0	11

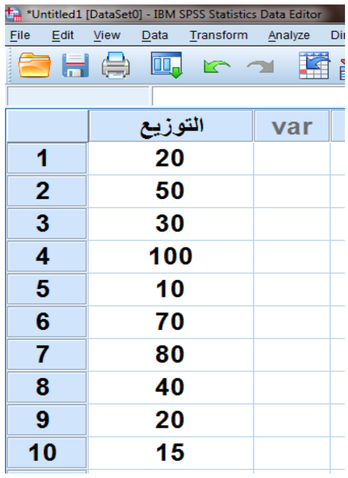
5-**المقارنة:** نقارن بين قيمة اختبار t المحسوبة وقيمة اختبار t الجدولية ، فنجد أن قيمة اختبار t المحسوبة(2.73) أكبر من قيمة اختبار t الجدولية (1.83)

6- **القرار:** بما أن قيمة اختبار t المحسوبة أكبر من قيمة اختبار t الجدولية ، فإننا نرفض فرضية العدم ونقبل الفرضية البديلة

7- **النتيجة:** المتوسط الحسابي للعينة يزيد عن المعدل العام للمجتمع الاحصائي زيادة ذات دلالة إحصائية عند مستوى دلالة 0.05 ، ونستنتج أن الفرق بين المتوسط الحسابي للعينة والمعدل للمجتمع الاحصائي فرق ذو دلالة إحصائية عند مستوى دلالة 0.05.

الحل :بعد إدخال البيانات في الحاسوب

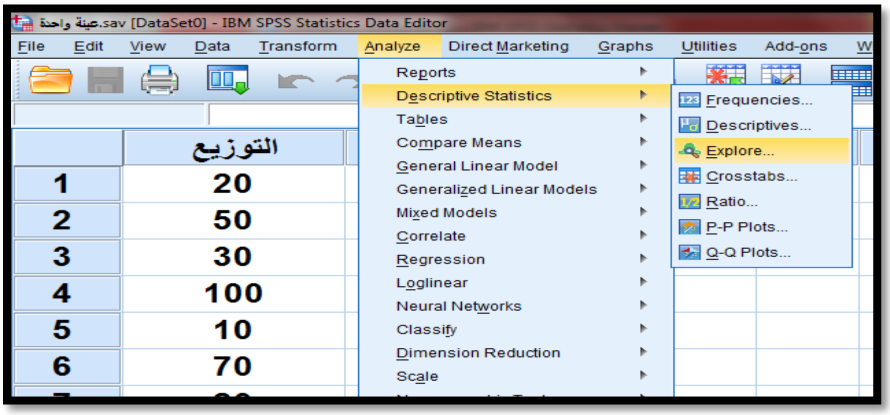
بعد إدخال البيانات في الحاسوب



<u>**أولاً نختبر إذا ما كانت البيانات تتبع التوزيع الطبيعي**</u>

فرضية العدم : البيانات تتبع التوزيع الطبيعي

الفرضية البديلة : البيانات لا تتبع التوزيع الطبيعي



Tests of Normality

	Kolmogorov–Smirnov[a]			**Shapiro–Wilk**		
	Statistic	Df	Sig.	Statistic	**Df**	**Sig.**
الوفيات	0.178	10	0.200*	0.907	**10**	**0.261**

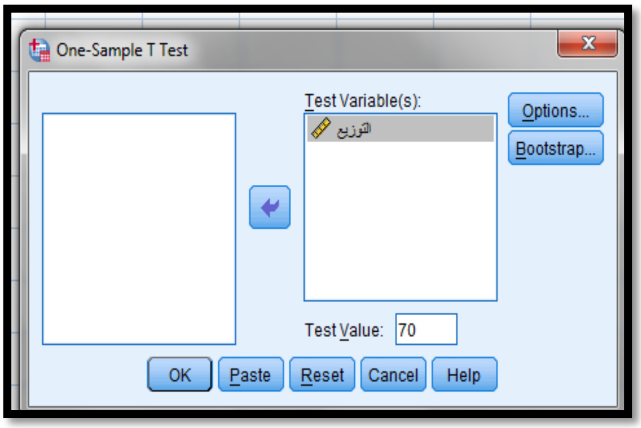
بما أن قيمة مستوى الدلالة = **0.261** حسب اختبار Shapiro–Wilk وهي أكبر من مستوى الدلالة **0.05** لذلك نقبل فرضية العدم القائلة **بأن البيانات تتبع التوزيع الطبيعي.**

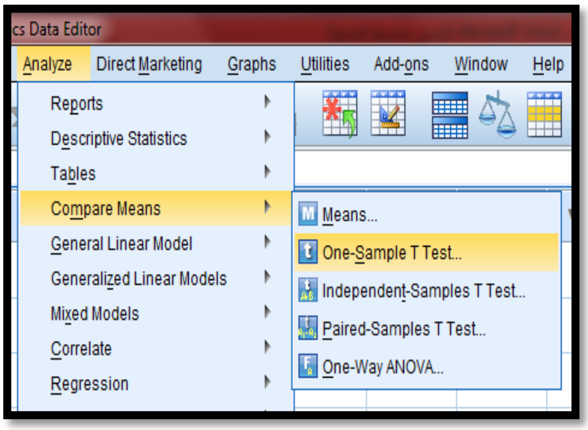
القرار : بما أن البيانات للعينة تتبع التوزيع الطبيعي، لذلك نستخدم اختبار t لعينة واحدة.

ثانياً/ يمكن استخدام (اختبار t) لعينة واحدة لمعرفة اختلاف متوسط عدد التوزيع في العينة عن 70 (المعدل العام)

فرضية العدم : لا يختلف متوسط عدد الوفيات في العينة عن المتوسط العام 70

الفرضية البديلة : يختلف متوسط الوفيات في العينة عن المتوسط العام 70 بدلالة إحصائية .





One-Sample Test						
	Test Value = 70					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
التوزيع	-2.726	9	0.023	-26.500	-48.4874-	-4.5126-

بما أن قيمة اختبار t = ‎- 2.726‎ ومستوى الدلالة لاختبار t = ‎0.023‎ و هو أقل من 0.05 لذلك نرفض فرضية العدم ونقبل الفرضية البديلة القائلة **بأن متوسط أعداد توزيع الوفيات في العينة يختلف عن المتوسط العام لأعداد الوفيات حسب معدل وزارة الزراعة، والفرق بينهما فرق حقيقي وهو ذو دلالة إحصائية عند مستوى دلالة 0.05.**

مثال 2 :

في دراسة حول إنتاج البندورة في قطاع غزة ، أخذت عينة من البيوت البلاستيكية بالشكل التالي:

رقم المشاهدة	1	2	3	4	5	6	7	8	9	10	11
الانتاجية/ طن	**20**	**15**	**20**	**17**	**15**	**25**	**16**	**20**	**13**	**12**	**10**

المطلوب : إذا كان معدل متوسط إنتاجية الدونم في قطاع غزة 19 طن للدونم ، فهل يوجد اختلاف حقيقي بين متوسط إنتاجية البندورة في هذه العينة عن المعدل العام .

الحل :

أولاً / الحل بالطريقة الحسابية اليدوية :

رقم المزرعة	1	2	3	4	5	6	7	8	9	10	11	12
الإنتاجية x	**20**	**15**	**20**	**17**	**15**	**25**	**16**	**20**	**13**	**12**	**10**	**30**
X^2	**400**	**225**	**400**	**289**	**225**	**625**	**256**	**400**	**169**	**144**	**100**	**900**

مجموع x = 213   مجموع x^2 = 4133

المتوسط الحسابي =

$$\bar{X} = \frac{\sum x}{N} = \frac{213}{12} = 17.75$$

الانحراف المعياري	
	$\sigma = \sqrt{\frac{\sum X^2 - N\bar{X}^2}{N-1}} = \sqrt{\frac{4133 - 12(17.75)^2}{12-1}} = \sqrt{\frac{4133 - 12(315.1)}{11}}$ $= \sqrt{\frac{4133 - 3780.75}{11}} = \sqrt{\frac{352.25}{11}} = \sqrt{32.02} = 5.659$

فرضية العدم: لا يوجد اختلاف بين المتوسط الحسابي للعينة والمعدل العام لإنتاجية البندورة في قطاع غزة.

الفرضية البديلة: يوجد اختلاف بين المتوسط الحسابي للعينة والمعدل العام لإنتاجية البندورة في قطاع غزة.

بما أن المتوسط الحسابي للعينة معروف (17.75)، والانحراف المعياري للعينة معروف(5.659)، والمعدل العام لوفيات الدجاج (المجتمع) معروف (19) ، وحجم العينة 12 مفردة ، فلذلك يمكننا استخدام اختبار t

8- ايجاد قيمة t المحسوبة :

$$\boldsymbol{t = \frac{(\bar{x} - \mu)}{S}\sqrt{N} =}$$

$$\boldsymbol{\frac{(17.75 - 19)}{5.659}\sqrt{12} = \frac{(-1.25)}{5.659} * 3.46 = 0.22 * 3.16 = 0.76}$$

9- ايجاد قيمة اختبار t الجدولية، فإذا كان مستوى المعنوية 0.05

د- درجات الحرية = N – 1 = 12 – 1 = 11

هـ- مستوى الدلالة = 0.05 نرجع لجدول اختبار t

و- فإن قيمة t الجدولية= **1.8**

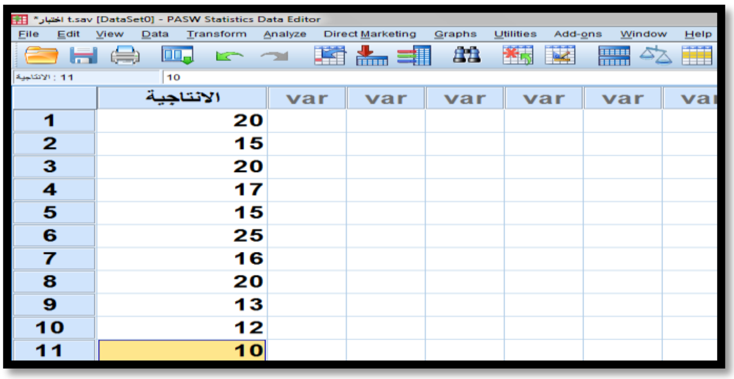
0.01	0.03	0.05	0.1	0.15	0.2	0.25	0.5	مستوى الدلالة
								Df
31.82	12.71	6.31	3.08	1.96	1.38	1	0	1
6.97	4.3	2.92	1.89	1.39	1.06	0.82	0	2
4.54	3.18	2.35	1.64	1.25	0.98	0.77	0	3
3.75	2.78	2.13	1.53	1.19	0.94	0.74	0	4
3.37	2.57	2.02	1.48	1.16	0.92	0.73	0	5
3.14	2.45	1.94	1.44	1.13	0.91	0.72	0	6
3.00	2.37	1.90	1.42	1.12	0.9	0.71	0	7
2.90	2.31	1.86	1.40	1.11	0.89	0.71	0	8
2.82	2.26	1.83	1.38	1.10	0.88	0.7	0	9
2.76	2.23	1.81	1.37	1.09	0.88	0.7	0	10
2.72	2.2	1.80	1.36	1.09	0.88	0.7	0	11

المقارنة: نقارن بين قيمة اختبار t المحسوبة وقيمة اختبار t الجدولية ، فنجد أن قيمة اختبار t المحسوبة(0.76) أقل من قيمة اختبار t الجدولية (1.83)

القرار: بما أن قيمة اختبار t المحسوبة أقل من قيمة اختبار t الجدولية ، فإننا لا نستطيع أن نرفض فرضية العدم ولا نستطيع أن نقبل الفرضية البديلة

النتيجة: المتوسط الحسابي للعينة يزيد عن المعدل العام للمجتمع الاحصائي زيادة ليست لها دلالة إحصائية عند مستوى دلالة 0.05 ، ونستنتج أن الفرق بين المتوسط الحسابي للعينة والمعدل للمجتمع الاحصائي فرق ليس له دلالة إحصائية عند مستوى دلالة 0.05.

ثانياً / الحل بالحاسوب :



أولاً / نختبر إذا ما كانت العينة تتبع التوزيع الطبيعي

فرضية العدم: البيانات تتبع التوزيع الطبيعي **الفرضية البديلة:** البيانات لا تتبع التوزيع الطبيعي

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	Df	Sig.
التكرار	.179	12	0.200*	.940	12	0.504

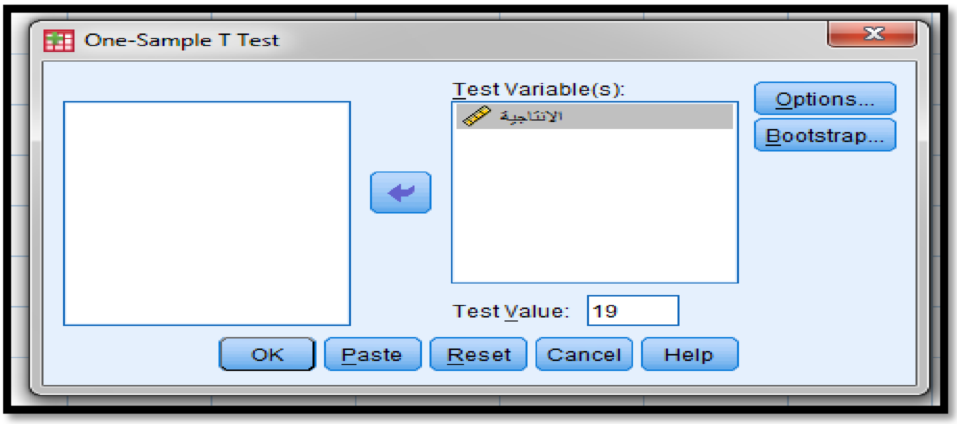
*. This is a lower bound of the true significance.

بما أن مستوى الدلالة **لاختبار شابيرو**= 0.835 و هي أكبر من 0.05 ، لذلك نقبل فرضية العدم القائلة بأن البيانات تتبع التوزيع الطبيعي

القرار: عليه نستخدم اختبار t لعينة واحدة

فرضية العدم : لا يوجد اختلاف بين متوسط العينة ومتوسط المجتمع = 19 طن

الفرضية البديلة : يوجد اختلاف بين متوسط العينة ومتوسط المجتمع = 19 طن



One-Sample Test

	Test Value = 19					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
التكرار	-.765-	11	0.460	-1.250-	-4.85-	2.35

بما أن قيمة اختبار $t = -0.765$ وقيمة مستوى الدلالة $= 0.460$ وهو أكبر من 0.05

القرار: لذلك لا نستطيع أن نرفض فرضية العدم ، ولا نستطيع أن نقبل الفرضية البديلة

النتيجة : لذلك يمكن أن نقول أن الفرق بين متوسط العينة ومتوسط المجتمع فرق ليس له دلالة احصائية .

ثانياً / اختبار t لعينتين مستقلتين

يستخدم اختبار t لاختبار الفرق بين متوسطي عينتين مستقلتين، إذا ما كان الفرق بينهما فرق حقيقي ذو دلالة إحصائية أم أنه فرق جاء عن طريق الصدفة وليس له دلالة إحصائية، وهو من الاختبارات المعلمية الشائعة الاستخدام، ويستعمل للبيانات الكمية ذات التوزيع الطبيعي ويستخدم في حالة تجانس أو عدم تجانس البيانات .

شروط استخدام اختبار t:

لابد على الباحث قبل استخدامه لاختبار (ت) أن يدرس خصائص متغيرات بحثه من النواحي التالية :

أ- **حجم كل عينة :** إن الأصل في هذا الاختبار أنه من مقياس دلالة العينات الصغيرة ، ولكن هذا لا يمنع استخدامه للعينات الكبيرة، واستخدامه للعينات الصغيرة جداً (التي يقل عدد أفرادها عن 30فرداً) أمر مشكوك فيه إذ يميل فيها التوزيع إلى أن يكون مدبباً، أما العينات الكبيرة فهي التي يزيد عدد أفرادها عن 30فرداً وفيها يميل التوزيع إلى أن يكون اعتدالياً طبيعياً، في حين أن العينات الصغيرة جداً يستخدم معها أحد الاختبارات اللابارامترية.

ب- **الفرق بين حجم العينتين** : يُفضل أن يكون حجم عينتي الدراسة متقارباً ، فلا يكون مثلاً حجم أحد العينتين 600 فرداً والأخرى 70فرداً ، لأن درجات الحرية (وهي المدخل المباشر للكشف عن مستوى الدلالة) تعتمد على عدد أفراد كل عينة، كما أن لحجم العينة تأثيراً على المؤشرات الإحصائية المستخدمة في حساب اختبار (t) وهي المتوسط والتباين.

ج- **تجانس العينتين** يُقاس مدى تجانس العينتين بالفرق بين تباين العينتين ، وذلك باستخدام اختبار Levene ، وهو مترافق مع اختبار t للعينتين المستقلتين.

د- **اعتدالية التوزيع التكراري لكل من عينتي البحث** والمقصود بالاعتدالية هي مدى تحرر التوزيع التكراري من الالتواء ، والالتواء قد يكون سالباً أو موجباً، في حين أن التوزيع الاعتدالي لا التواء فيه، ويمتد معامل الالتواء من -3 إلى +3 ، وكلما اقترب معامل الالتواء من الصفر كان التوزيع اعتدالياً، ففي التوزيع الاعتدالي يكون المتوسط الحسابي = الوسيط. ويمكن استخدام اختبار كولمجروف للعينات الكبيرة، واختبار شابيرو للعينات الصغيرة (أقل من 30 مفردة).

مثال : في دراسة حول مدى ملائمة التربة لزراعة محصول البطاطا في قطاع غزة أخذت عينتان بالشكل التالي:

رقم المشاهدة	أ	ب	ج	د	هـ	و	ز	ح	ط	ي
العينة الأولى"رملية "	500	600	700	800	750	650	450	520	630	120
العينة الثانية "طينية "	1000	1200	1300	1400	1500	2000	1300	1400	1450	1200

المطلوب : هل يوجد اختلاف حقيقي بين متوسط العينتين ؟

العينة	العينة الأولى X1	$(X1)^2$
أ	500	250000
ب	600	360000
ج	700	490000
د	800	640000
ه	750	562500
و	650	422500
ز	450	202500
ح	520	270400
ط	630	396900
ي	120	14400
المجموع	**5720**	**3609200**

العينة	العينة الثانية X2	$(X2)^2$
أ	1000	1000000
ب	1200	1440000
ج	1300	1690000
د	1400	1960000
ه	1500	2250000
و	2000	4000000
ز	1300	1690000
ح	1400	1960000
ط	1450	2102500
ي	1200	1440000
المجموع	**13750**	**19532500**

المتوسط الحسابي العينة الأولى :

$$\bar{X} = \frac{\sum x}{N} = \frac{5720}{10} = 572$$

المتوسط الحسابي للعينة الثانية :

$$\bar{X} = \frac{\sum x}{N} = \frac{13750}{10} = 1375$$

الانحراف المعياري للعينة الأولى :

$$\sigma = \sqrt{\frac{\sum X^2 - N\bar{X}^2}{N-1}} = \sqrt{\frac{3609200 - 10(572)^2}{10-1}} = \sqrt{\frac{3609200 - 10(327184)}{9}}$$

$$= \sqrt{\frac{3609200 - 3271840}{9}} = \sqrt{\frac{337360}{9}} = \sqrt{37484.44} = 193.6$$

الانحراف المعياري للعينة الثانية =

$$\sigma = \sqrt{\frac{\sum X^2 - N\bar{X}^2}{N-1}} = \sqrt{\frac{19532500 - 10(1375)^2}{10-1}}$$

$$= \sqrt{\frac{19532500 - 10(1890625)}{9}} = \sqrt{\frac{19532500 - 18906250}{9}} =$$

$$\sqrt{\frac{626250}{9}} = \sqrt{69583.3} = 263.78$$

1- **فرضية العدم :** لا يوجد اختلاف بين المتوسط الحسابي إنتاجية المحصول في التربة الطينية والمتوسط الحسابي لإنتاجية المحصول في التربة الرملية

2- **الفرضية البديلة :** يوجد اختلاف بين المتوسط الحسابي إنتاجية المحصول في التربة الطينية والمتوسط الحسابي لإنتاجية المحصول في التربة الرملية

3- قيمة اختبار t

$$t = \frac{|\overline{X_1} - \overline{X_2}|}{\sqrt[2]{\frac{(S_1)^2}{N_1} + \frac{(S_2)^2}{N_2}}} = \frac{|572 - 1375|}{\sqrt[2]{\frac{(193.6)^2}{10} + \frac{(263.78)^2}{10}}}$$

$$= \frac{|803|}{\sqrt[2]{\frac{37480.96}{10} + \frac{69579.9}{10}}} = \frac{803}{\sqrt[2]{3748.1 + 6957.99}}$$

$$= \frac{803}{\sqrt[2]{10706.1}} = \frac{803}{103.47} = 7.76$$

إذن قيمة اختبار t المحسوبة = 7.76

4- القيمة الحرجة ، قيمة اختبار t الجدولية :

درجات الحرية = $N_1 + N_1 - 2 = 10 + 10 - 2 = 18$

مستوى الدلالة = 0.05

إذن قيمة اختبار t الجدولية = 1.734

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01
df									
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845

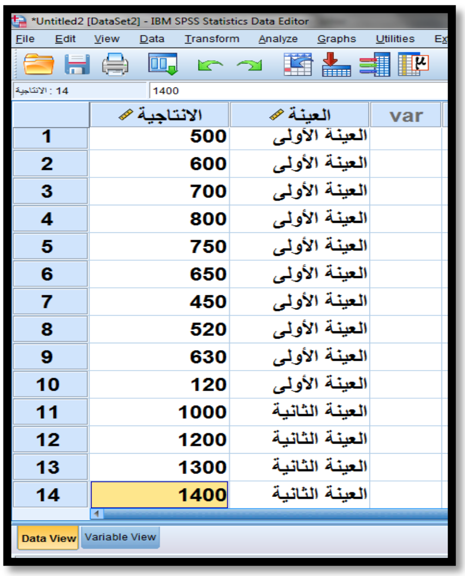
5- **المقارنة** : نقارن بين قيمة اختبار t المحسوبة وقيمة اختبار t الجدولية ، فنجد أن قيمة اختبار t المحسوبة أكبر من قيمة اختبار t الجدولية

6- **القرار** : بما أن قيمة اختبار t المحسوبة أكبر من قيمة اختبار t الجدولية ، فإننا نرفض فرضية العدم ونقبل الفرضية البديلة

7- **النتيجة** : أن الفرق بين متوسطي العينتين فرق حقيقي ذو دلالة إحصائية عند مستوى دلالة 0.05 ، ومن ذلك نستنتج أن زراعة محصول البطاطا في الأرض الطينية قد أعطى زيادة في الانتاجية زيادة حقيقية .

ثانياً / الحل بالحاسوب

أولاً / نقوم بإدخال البيانات كما في الشكل :



	الانتاجية	العينة	var
1	500	العينة الأولى	
2	600	العينة الأولى	
3	700	العينة الأولى	
4	800	العينة الأولى	
5	750	العينة الأولى	
6	650	العينة الأولى	
7	450	العينة الأولى	
8	520	العينة الأولى	
9	630	العينة الأولى	
10	120	العينة الأولى	
11	1000	العينة الثانية	
12	1200	العينة الثانية	
13	1300	العينة الثانية	
14	1400	العينة الثانية	

Tests of Normality

	العينة	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
الانتاجية	العينة الأولى	.164	10	.200*	.893	10	.183
	العينة الثانية	.218	10	.196	.880	10	.130

*. This is a lower bound of the true significance.

a Lilliefors Significance Correction

ثانياً/ نختبر إذا ما كانت البيانات تتبع التوزيع الطبيعي

فرضية العدم: البيانات تتبع التوزيع الطبيعي

الفرضية البديلة: البيانات لا تتبع التوزيع الطبيعي

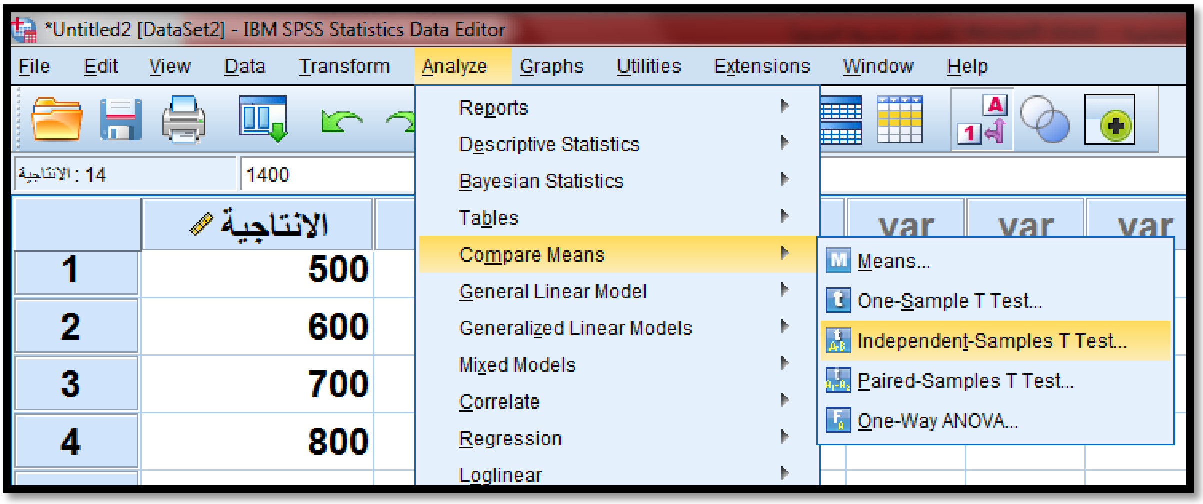
بما أن مستوى الدلالة لاختبار شابيرو للعينة الأولى = 0.183 ، ويساوي 0.130 للعينة الثانية ، وهما أكبر من 0.05 ، لذلك نقبل فرضية العدم القائلة بأن البيانات تتبع توزيع طبيعي

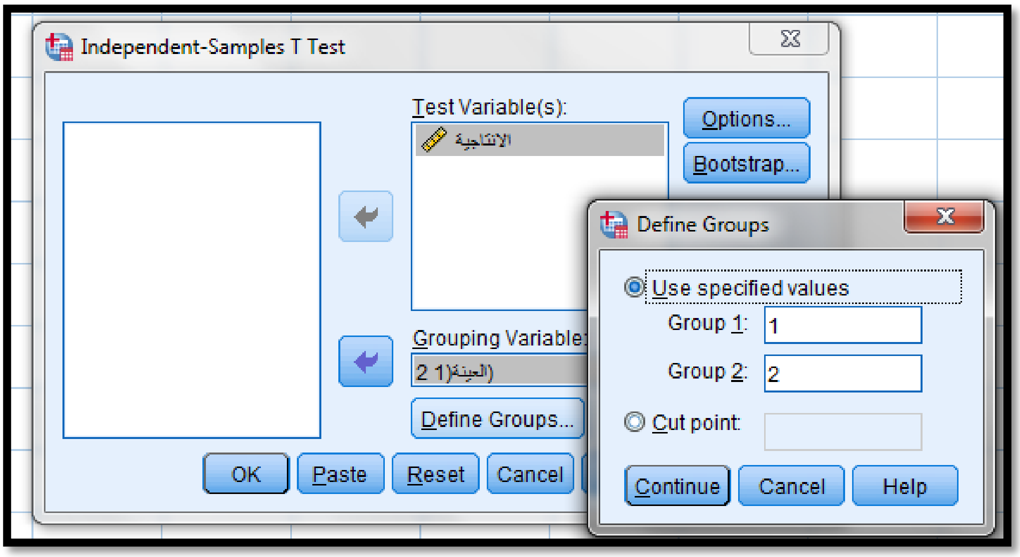
القرار : نستخدم اختبار t لعينتين مستقلتين

ثالثاً / نستخدم اختبار t لعينتين مستقلتين :

فرضية العدم : لا يوجد اختلاف بين متوسط العينة الأولى ومتوسط العينة الثانية

الفرضية البديلة : يوجد اختلاف بين متوسطي العينتين





Group Statistics

	العينة	N	Mean	Std. Deviation	Std. Error Mean
الانتاجية	العينة الأولى	10	572.00	193.609	61.225
	العينة الثانية	10	1375.00	263.787	83.417

Independent Samples Test

	Levne test		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
Equal variances assumed	**0.244**	**0.627**	-7.760-	18	0.000	-803.00000-	103.47356	-1020.38988-	-585.61012-
Equal variances not assumed			-7.760-	16.51	0.000	-803.00000-	103.47356	-1021.79893-	-584.20107-

رابعاً / نختبر تجانس البيانات

فرضية العدم: يوجد تجانس التباين للبيانات   **الفرضية البديلة:** لا يوجد تجانس التباين للبيانات

النتيجة : بما أن قيمة اختبار Levne = 0.244 ومستوى الدلالة 0.627 ، لأنه أكبر من 0.05 ، لذلك نقبل فرضية العدم القائلة بتجانس التباين بين العينتين

وبما أن البيانات بها تجانس للتباين فإننا نأخذ القيمة في الصف الأول من جدول t (بفرض تجانس التباين) (**Equal variances assumed**)

بما أن قيمة اختبار t = 7.7− ، ومستوى الدلالة أقل من 0.000 وهو أقل من 0.05 ، لذلك نرفض فرضية العدم ونقبل الفرضية البديلة

أي أنه يوجد اختلاف بين متوسط العينة الأولى ومتوسط العينة الثانية والفرق بينهما فرق له دلالة إحصائية عند مستوى دلالة أقل من 0.01 ، أي أن طبيعة ونوع التربة له تأثير في انتاجية المحصول

3- اختبار t لعينتين مزدوجتين

يستخدم اختبار t لاختبار الفرق بين متوسطي عينتين مرتبطتين، أي أن الحالة تأخذ قراءتين، القراءة الأولى في العينة الأولى وتكون **قبل** احداث التغيير في الظاهرة المدروسة، ثم القراءة الثانية تكون في العينة الثانية وتكون **بعد** حدوث التغيير.

مثال : يبين الجدول التالي عملية تقويم لطلبة مساق الخرائط، حيث اجري امتحان للطلبة العشرة في بداية المساق ثم بعد دراستهم المساق أجري امتحان للطلبة لمعرفة التحسن في مهارتهم :

الطالب	الوقت قبل دراسة مادة الخرائط	الوقت بعد دراسة مادة الخرائط	الفرق d	d^2
1-	16	12	4	16
2-	23	19	4	16
3-	17	13	4	16
4-	14	11	3	9
5-	16	16	0	0
6-	21	23	-2	4
7-	19	16	3	9
8-	24	20	4	16
9-	26	21	5	25
10-	19	20	1-	1
المجموع	195	171	24	112
المتوسط	19.5	17.1	$\bar{d}$ =2.4	

المطلوب : هل يوجد اختلاف في الوقت الذي يحتاج إليه الطلبة لاستخراج المعلومات من الخريطة ؟

خطوات الحل :

1- **فرضية العدم** : لا يوجد اختلاف في الوقت الذي يحتاج إليه الطلبة لاستخراج المعلومات من الخريطة ، أي أنهم لم يكتسبوا مهارات جديدة من دراستهم لمساق الخرائط .

2- **الفرضية البديلة** : يوجد فرق ذو دلالة إحصائية بين متوسط الوقت الذي كان الطلبة يحتاجون إليه لاستخراج المعلومات من الخريطة قبل دراستهم لمادة الخرائط وبعد دراستهم لها .

3- قيمة اختبار t الحسابية :

أولاً / ايجاد قيمة الانحراف المعياري :

$$SD = \sqrt{\frac{\sum d^2 - n\bar{d}^2}{n-1}} = \sqrt{\frac{112 - 10(2.4)^2}{10-1}} = \sqrt{\frac{54.4}{9}} = \sqrt{6.04}$$

$$SD = 2.46$$

ثانياً / إيجاد قيمة اختبار t الحسابية:

$$\mathrm{T} = \frac{\bar{d}}{SD}(\sqrt{N}) = \frac{2.4}{2.46}(\sqrt{10}) = 0.976 \times (3.162) = 3.08$$

T =3.08

إذن قيمة t المحسوبة = 3.08

4- قيمة t الجدولية (القيمة الحرجة)

أ- درجات الحرية = N − 1 = 10 − 1 = 9

ب- مستوى الدلالة 0.05

ج-إذن قيم t الجدولية عند مستوى دلالة 0.05 ، ودرجات حرية 9 = 1.833

5- **المقارنة :** نقارن بين قيمة t المحسوبة وقيمة t الجدولية، فنجد أن قيمة t المحسوبة أكبر من قيمة t الجدولية

6- **القرار :** بما أن قيمة t المحسوبة أكبر من قيمة t الجدولية، إذن نرفض فرضية العدم ونقبل الفرضية البديلة.

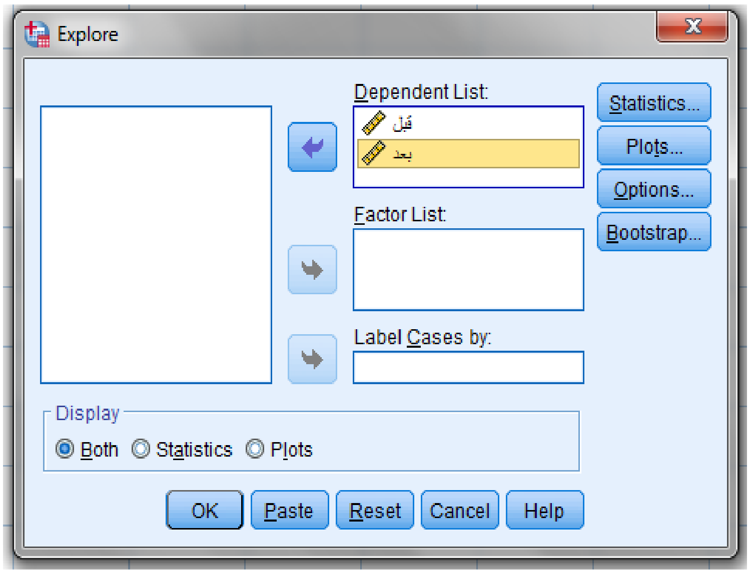
7- **النتيجة :** إن الفرق بين متوسط الوقت الذي كان الطلبة يحتاجون إليه لاستخراج المعلومات من الخريطة قبل دراستهم لمادة الخرائط وبعد دراستهم لها فرق له دلالة إحصائية عند مستوى دلالة 0.05.

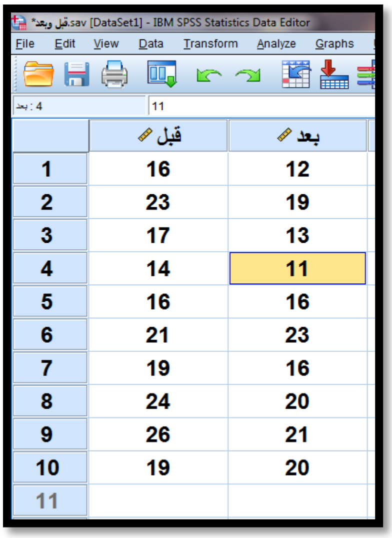
الحل بالحاسوب :

خطوات الحل :

أولاً / نختبر هل البيانات تتبع توزيع طبيعي أم لا

فرضية العدم: البيانات تتبع توزيع طبيعي **الفرضية البديلة:** البيانات لا تتبع توزيع طبيعي





Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
قبل	0.151	10	.200*	0.956	10	0.742
بعد	0.178	10	.200*	0.935	10	0.500

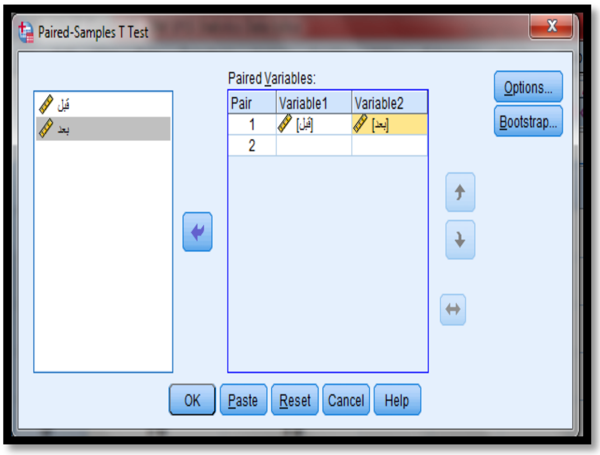
*. This is a lower bound of the true significance.

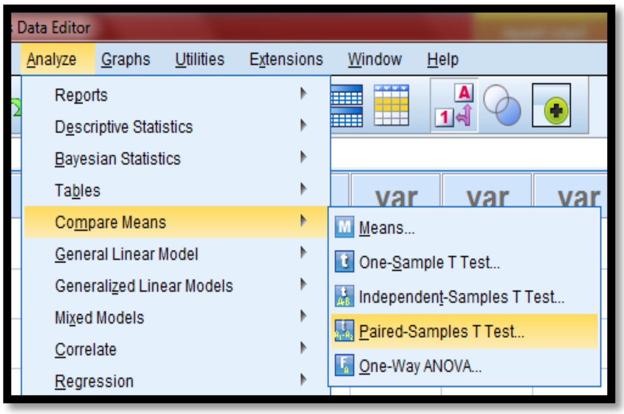
a. Lilliefors Significance Correction

بما أن مستوى الدلالة للعينتين (0.742 ، 0.50) , هي أكبر من 0.05

لذلك نقبل فرضية العدم القائلة بأن البيانات تتبع التوزيع الطبيعي

إذن نستخدم اختبار t المعلمي لعينتين مرتبطتين (مزدوجتين)





ثانياً / اختبار ارتباط العينتين

نجد أن متوسط الزمن قبل الدورة = 16.5 و متوسط الزمن بعد الدورة = 18.9 ، أي أنه يوجد فرق = 2.4

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	قبل	19.50	10	3.923	1.241
	بعد	17.10	10	4.122	1.303

Paired Samples Correlations

	N	Correlation	Sig.
Pair 1 بعد & قبل	10	0.814	0.004

نجد أن معامل الارتباط بين العينتين كبير وهو يساوي 0.814 و هو ذو دلالة إحصائية 0.004 وهي أقل من 0.05 .

لذلك يمكن أن نستخدم اختبار t لعينتين مرتبطتين

1- **فرضية العدم** : لا يوجد اختلاف في الوقت الذي يحتاج إليه الطلبة لاستخراج المعلومات من الخريطة ، أي أنهم لم يكتسبوا مهارات جديدة من دراستهم لمساق الخرائط .

2-**الفرضية البديلة** : يوجد فرق ذو دلالة إحصائية بين متوسط الوقت الذي كان الطلبة يحتاجون إليه لاستخراج المعلومات من الخريطة قبل دراستهم لمادة الخرائط وبعد دراستهم لها .

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 قبل - بعد	2.4	2.459	0.777	0.641	4.159	3.087	19	0.013

نلاحظ أن قيمة اختبار t = 3.087 ، و مستوى الدلالة = 0.013

3-**القرار:** إذن نجد أن قيمة مستوى الدلالة أقل من 0.05 لذلك نرفض فرضية العدم ونقبل الفرضية البديلة

4-النتيجة: يوجد اختلاف بين متوسط الوقت للعينة الأولى قبل التدريب ومتوسط الوقت للعينة الثانية بعد التدريب والفرق بينهما فرق حقيقي عند مستوى دلالة 0.05

بصيغة أخرى: إن الفرق بين متوسط الوقت الذي كان الطلبة يحتاجون إليه لاستخراج المعلومات من الخريطة قبل دراستهم لمادة الخرائط وبعد دراستهم لها فرق له دلالة إحصائية عند مستوى دلالة 0.05.

مثال 2: في دراسة لإحدى الجمعيات الزراعية حول تطوير إنتاجية القمح في قطاع غزة ، استخدمت نوعاً من البذور في 11 موقعاً ، ثم في العام اللاحق استخدمت بذوراً محسنة في نفس المواقع السابقة بالشكل التالي :

الموقع	الإنتاجية 2012(قبل)	الانتاجية 2013 (بعد)	الفرقd	d^2
أ	120	130	10	100
ب	130	150	20	400
ج	140	160	20	400
د	160	180	20	400
هـ	150	190	40	1600
و	110	160	50	2500
ز	100	130	30	900
ح	115	190	75	5625
ط	130	200	70	4900
ك	170	210	40	1600
ل	180	220	40	1600
المجموع	**1505**	**1920**	415	**20025**
المتوسط	**136.8**	**174.5**	37.7	

المطلوب: **هل استخدام البذور المحسنة قد زاد من إنتاجية القمح في المواقع السابقة؟ مستوى الدلالة 0.05**

خطوات الحل :

1- **فرضية العدم** : لا يوجد اختلاف بين انتاجية القمح قبل وبعد استخدام البذور المحسنة .

2- **الفرضية البديلة** : يوجد فرق ذو دلالة إحصائية بين إنتاجية القمح قبل وبعد استخدام البذور المحسنة .

3- **قيمة اختبار t الحسابية :**

$$\mathbf{T} = \frac{\bar{d}}{SD}\sqrt{N}$$

أولاً / ايجاد قيمة الانحراف المعياري :

$$SD = \sqrt{\frac{\sum d^2 - n\bar{d}^2}{n-1}} = \sqrt{\frac{20025 - 11(37.7)^2}{11-1}} = \sqrt{\frac{4390.81}{10}}$$

$$= \sqrt{439.81} = 20.95$$

ثانياً / إيجاد قيمة اختبار t :

$$T = \frac{\bar{d}}{SD}\sqrt{N} = \frac{37.7}{20.95}\sqrt{11} = 1.799 \times 3.316 = 5.968$$

إذن قيمة t المحسوبة = 5.968

4-قيمة t الجدولية (القيمة الحرجة)

أ- درجات الحرية = N − 1 = 11 − 1 = 10

ب- مستوى الدلالة 0.05

ج- إذن قيم t الجدولية عند مستوى دلالة 0.05 ، ودرجات حرية 10 = 1.812

5- **المقارنة :** نقارن بين قيمة t المحسوبة وقيمة t الجدولية ، فنجد أن قيمة t المحسوبة أكبر من قيمة t الجدولية

6- **القرار :** بما أن قيمة t المحسوبة أكبر من قيمة t الجدولية، إذن نرفض فرضية العدم ونقبل الفرضية البديلة

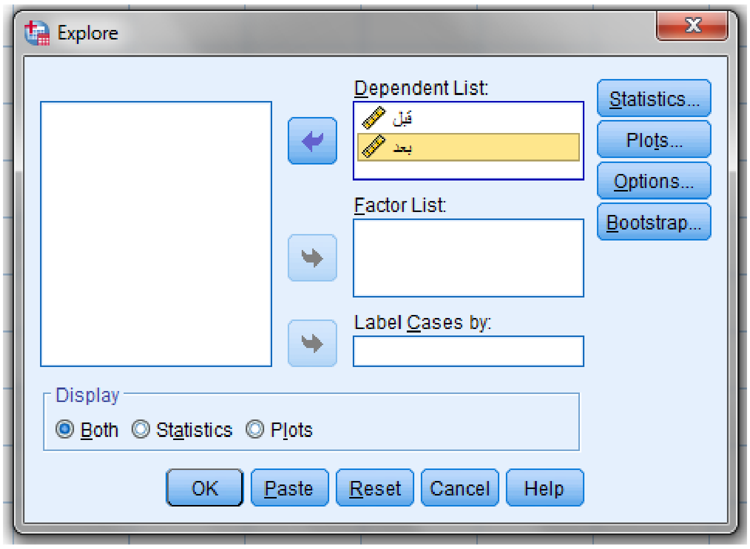
7- **النتيجة :** الفرق بين متوسط إنتاجية القمح قبل استخدام البذور المحسنة وبين إنتاجية البذور المحسنة فرق له دلالة إحصائية عند مستوى دلالة 0.05.

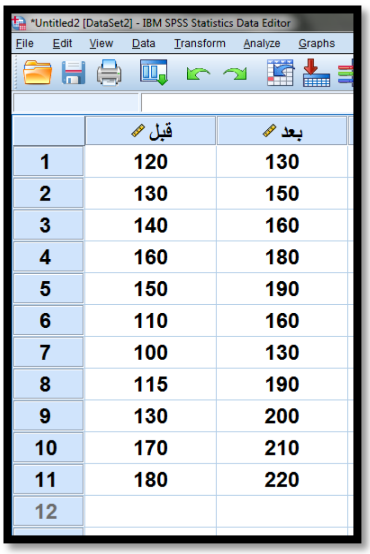
ثانياً / الحل بالحاسوب :

خطوات الحل :

أولاً / نختبر هل البيانات تتبع توزيع طبيعي أم لا

فرضية العدم: البيانات تتبع توزيع طبيعي   **الفرضية البديلة:** البيانات لا تتبع توزيع طبيعي





	قبل	بعد
1	120	130
2	130	150
3	140	160
4	160	180
5	150	190
6	110	160
7	100	130
8	115	190
9	130	200
10	170	210
11	180	220
12		

Tests of Normality

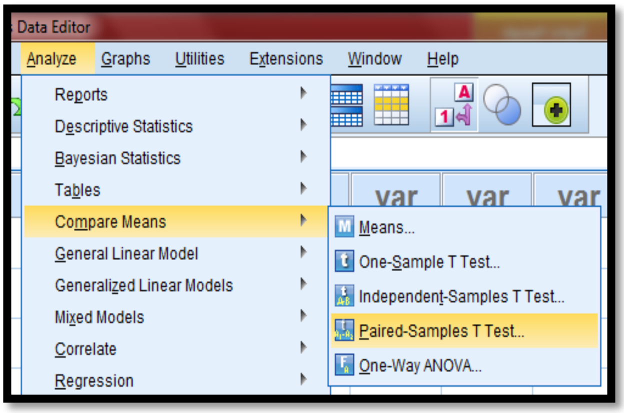
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
قبل	0.15	11	.200*	0.964	11	0.824
بعد	0.147	11	.200*	0.945	11	0.581

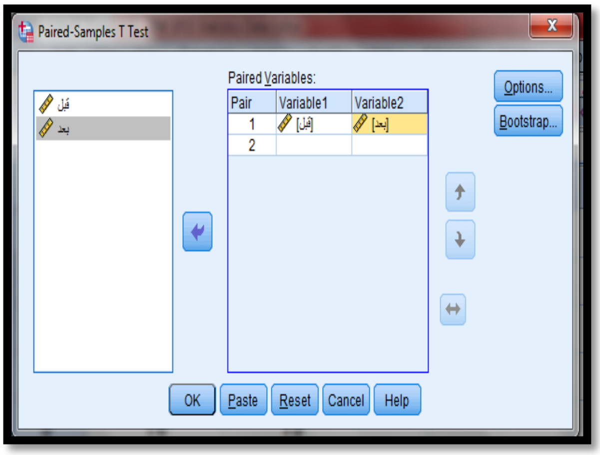
*. This is a lower bound of the true significance.

بما أن مستوى الدلالة للعينتين (0.824 ، 0.581) , هي أكبر من 0.05

لذلك نقبل فرضية العدم القائلة بأن البيانات تتبع التوزيع الطبيعي

إذن نستخدم اختبار t المعلمي لعينتين مرتبطتين (مزدوجتين)





ثانياً / اختبار ارتباط العينتين

نجد أن متوسط الزمن قبل الدورة = 16.5 و متوسط الزمن بعد الدورة = 18.9 ، أي أنه يوجد فرق = 2.4

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	قبل	136.82	11	25.717	7.754
	بعد	174.55	11	30.778	9.280

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	بعد & قبل	11	0.740	0.009

نجد أن معامل الارتباط بين العينتين كبير وهو يساوي 0.74 و هو ذو دلالة إحصائية 0.009 وهي أقل من 0.05 .

لذلك يمكن أن نستخدم اختبار t لعينتين مرتبطتين

1- **فرضية العدم** : لا يوجد اختلاف بين انتاجية القمح قبل وبعد استخدام البذور المحسنة .

2- **الفرضية البديلة** : يوجد فرق ذو دلالة إحصائية بين إنتاجية القمح قبل وبعد استخدام البذور المحسنة .

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 قبل - بعد	-37.727	20.9	6.302	-51.768	-23.686	-5.987	10	0.000

نلاحظ أن قيمة اختبار t =− 5.987 ، ومستوى الدلالة = 0.000

5−**القرار:** إذن نجد أن قيمة مستوى الدلالة أقل من 0.05 لذلك نرفض فرضية العدم ونقبل الفرضية البديلة

6− **النتيجة:** الفرق بين متوسط إنتاجية القمح قبل استخدام البذور المحسنة وبين إنتاجية البذور المحسنة فرق له دلالة إحصائية عند مستوى دلالة 0.05.

بصيغة أخرى: **يوجد اختلاف بين متوسط إنتاجية القمح للعينة الأولى قبل استخدام البذور المحسنة ومتوسط إنتاجية القمح للعينة الثانية بعد** استخدام البذور المحسنة **والفرق بينهما فرق حقيقي عند مستوى دلالة 0.05**

بصيغة أخرى: إن الفرق بين متوسط الوقت الذي كان الطلبة يحتاجون إليه لاستخراج المعلومات من الخريطة قبل دراستهم لمادة الخرائط وبعد دراستهم لها فرق له دلالة إحصائية عند مستوى دلالة 0.05.

اختبار تحليل التباين

يستخدم اختبار تحليل التباين لاختبار الفروق بين متوسطات ثلاث عينات فأكثر، ويشترط لاستخدامه بأن تكون البيانات تتبع التوزيع الطبيعي. يستخدم في حالة وجود متغير تابع وله متغير مستقل ولكن بمستويات متعددة.

جدول (1) المتغيرات المستقلة والتابعة في تحليل التباين الأحادي One – Way ANOVA

<table dir="rtl">
<tr><td rowspan="2">المتغير التابع
درجة التحضر</td><td colspan="5">المتغير المستقل: المجموعات الاقتصادية
(ثلاثة مستويات) Factor</td></tr>
<tr><td>الدول الصناعية</td><td colspan="2">الشرق الوسط</td><td colspan="2">إفريقيا</td></tr>
<tr><td rowspan="2">المتغير التابع
الدخل الشهري</td><td colspan="5">المتغير المستقل: المستوى التعليمي
أربعة مستويات Factor</td></tr>
<tr><td>ابتدائي</td><td>إعدادي</td><td>ثانوي</td><td colspan="2">جامعي</td></tr>
<tr><td rowspan="2">المتغير التابع
عدد العمال</td><td colspan="5">المتغير المستقل: (تصنيف المصانع)
خمسة مستويات Factor</td></tr>
<tr><td>مصانع صغيرة جداً</td><td>مصانع صغيرة</td><td>مصانع متوسطة</td><td>مصانع كبيرة</td><td>مصانع كبيرة جدا</td></tr>
</table>

نلاحظ في تحليل التباين الأحادي وجود متغير مستقل واحد (بثلاثة مستويات فأكثر) ومتغير تابع واحد. وفيه نختبر فرضية اختلاف الأوساط الحسابية للمتغير التابع بين المستويات الموجودة في المتغير المستقل.

منطق اختبار تحليل التباين: التباين هو مقياس للتشتت والاختلاف، فالاختبار يقيس التباين داخل كل مجموعة(عينة) With Variance ويقيس التباين بين المجموعات Between Variance

ومن شروط استخدام اختبار تحليل التباين :

1-أن تكون مفردات العينات مستقلة

2-أن يكون المتغير التابع مقاس على الأقل على المستوى الفئوي.

3-المتغير التابع موزع توزيعاً قريباً من التوزيع الطبيعي في كل مجموعة.

4-تجانس التباين بين المجموعات

مثال 1 :

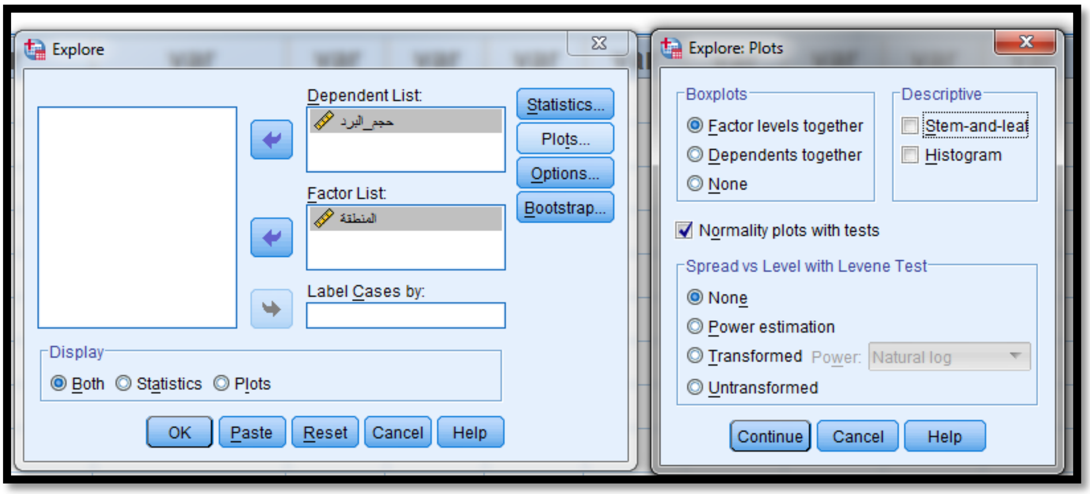
في دراسة حول حجم حبيبات البرد الساقطة على إحدى المدن، قام أحد الباحثين بتقسيم المدينة إلى أربعة أقسام حسب بعدها عن مركز المدينة، لمعرفة أثر التلوث الهوائي في حجم حبيبات البرد ، فحصل على البيانات التالية:

المطلوب: هل يوجد اختلاف حقيقي بين حجم حبيبات البرد في المناطق الأربع؟ وهل يوجد اختلاف بيت المنطقة A والمنطقة B ؟

المركز A	الثانية B	الثالث C	الرابع D
10	8	8	11
11	10	9	12
18	15	12	10
17	12	10	12
12	10	6	8
13	11	12	9
15	14	11	8
8	7	6	6
6	6	5	3
8	5	5	4
13	12	5	12

تختبر إذا ما كانت البيانات تتبع التوزيع الطبيعي

فرضية العدم: البيانات تتبع التوزيع الطبيعي **الفرضية البديلة:** البيانات لا تتبع التوزيع الطبيعي



Tests of Normality							
	المنطقة	Kolmogorov-Smirnov[a]			**Shapiro-Wilk**		
		Statistic	df	Sig.	Statistic	Df	Sig.
الحجم	A	.109	11	.200*	.967	11	0.886
	B	.136	11	.200*	.966	11	0.848
	C	.223	11	.131	.868	11	0.073
	D	.149	11	.200*	.902	11	0.196

بما أن قيمة مستوى الدلالة في العينات الأربع أكبر من 0.05 ، لذا نقبل فرضية العدم القائلة **بأن البيانات تتبع التوزيع الطبيعي**، لذلك سنستخدم **اختبار تحليل التباين**

2- نختبر إذا ما كانت العينات متجانسة التباين :

فرضية العدم : يوجد تجانس للتباين بين العينات الأربع

الفرضية البديلة : لا يوجد تجانس للتباين بين العينات الأربع

Test of Homogeneity of Variance					
		Levene Statistic	df1	df2	Sig.
حجم_البرد	Based on Mean	**0.216**	**3**	**40**	**0.885**

النتيجة : بما أن قيمة اختبار LEVNE = 0.216 ومستوى الدلالة= 0.885، وهي أكبر من 0.05 ، لذلك نقبل فرضية العدم القائلة بوجود تجانس للتباين.

لذلك نستخدم اختبار تحليل التباين ، واختبار LS'D

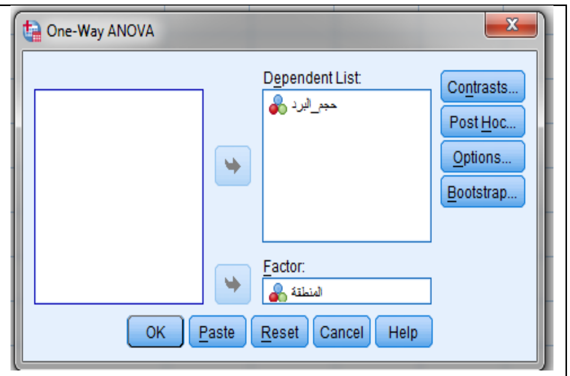
	N	Mean	Std. Deviation
المركز	11	11.91	3.807
الثاني	11	10.00	3.225
الثالث	11	8.09	2.844
الرابع	11	8.64	3.202
Total	44	9.64	3.518

من الجدول المقابل : نجد أن الفرق بين متوسطات أحجام الحبيبات مختلف ، فهل الاختلاف بين هذه المتوسطات اختلاف حقيقي ؟؟

فرضية العدم : لا يوجد اختلاف حقيقي بين متوسطات العينات الأربع، أي أن المتوسطات الأربعة متساوية

الفرضية البديلة : يوجد اختلاف حقيقي بين متوسطات العينات الأربع ، المتوسطات الأربعة غير متساوية

NOVA					
	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	95.523	3	31.841	2.946	0.044
Within Groups	432.364	40	10.809		
Total	527.886	43			



نجد أن قيمة اختبار تحليل التباين= 2.946 و مستوى الدلالة = 0.044 و هي أقل من 0.05

القرار: لذلك نرفض فرضية العدم نقبل الفرضية البديلة

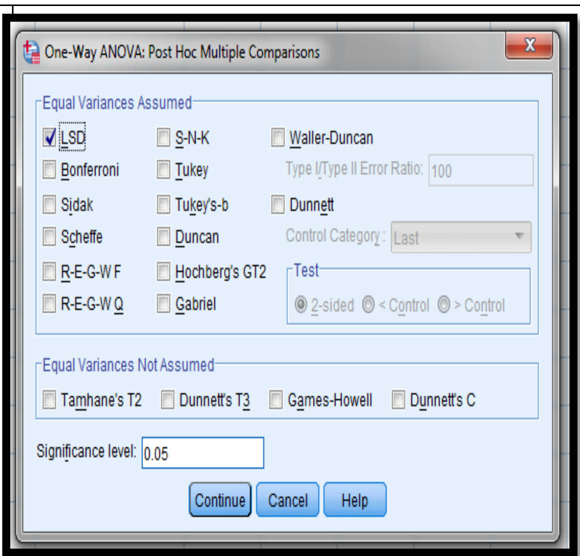
النتيجة : الفرق بين متوسط العينات الأربعة فرق له دلالة احصائية. أي أنه يوجد اختلاف حقيقي بين متوسطات العينات الأربع، والمتوسطات الأربعة غير متساوية.

أظهرت نتيجة اختبار التباين الأحادي أن الفروق بين المتوسطات الحسابية بين المجموعات الأربعة، فهل كل المجموعات الفروق بينها غير دالة إحصائياً أم أن هناك مجموعات دالة إحصائياً :

نجري اختبار Post Hoc لمعرفة الفرق بين العينات الأربع بالشكل التالي :

Multiple Comparisons LSD						
Dependent Variable: حجم_البرد						
					95% Confidence Interval	
(I) المنطقة	(J) المنطقة	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
المركز	الثاني	1.909	1.402	0.181	-1.04-	4.68
	الثالث	3.818*	1.402	0.01	0.87	6.59
	الرابع	3.273*	1.402	0.025	0.32	6.04
الثاني	المركز	-1.909-	1.402	0.181	-4.68-	1.04
	الثالث	1.909	1.402	0.181	-0.95-	4.77
	الرابع	1.364	1.402	0.337	-1.50-	4.23
الثالث	المركز	-3.818-*	1.402	0.01	-6.59-	-.87-
	الثاني	-1.909-	1.402	0.182	-4.77-	.95
	الرابع	-.545-	1.402	0.699	-3.41-	2.32
الرابع	المركز	-3.273-*	1.402	0.025	-6.04-	-.32-
	الثاني	-1.364-	1.402	0.337	-4.23-	1.50
	الثالث	.545	1.402	0.699	-2.32-	3.41

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



من الجدول: الفرق بين متوسط حبيبات البرد :

* **دال إحصائياً** بين مركز المدينة والمستوى الثالث (0.012) والمستوى الرابع (0.03)

* **غير دال إحصائياً** للفروق بين باقي المستويات

السؤال الثاني: أخذت ثلاث عينات من مزارعي الفراولة حول كمية الأسمدة التي تضاف للتربة في محافظة شمال غزة

رقم المنطقة التي أخذت منها العينة	1	2	3	4	5	6	7	8	9	10	11
العينة " 1 "(تربة رملية)	1000	1200	1500	700	600	800	500	900	1500	1800	2000
العينة " 2 "(تربة رملية طينية)	800	752	1533	1200	1000	900	800	750	850	654	123
العينة "3 " (تربة طينية)	700	800	500	300	200	150	100	400	300	450	600

المطلوب: **هل يوجد اختلاف حقيقي بين متوسطات العينات الثلاث ؟ وهل يوجد اختلاف حقيقي بين العينة الأولى والثانية؟**

خطوات الحل :

أولاً / نختبر إذا ما كانت البيانات تتبع توزيع طبيعي أم لا .

فرضية العدم: البيانات تتبع التوزيع الطبيعي **الفرضية البديلة:** البيانات لا تتبع التوزيع الطبيعي

بعد إدخال البيانات لبرنامج SPSS نحصل على النتيجة التالية :

Tests of Normality							
	العينة	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
كميةالسماد	العينة1	.152	11	.200*	.939	11	0.508
	العينة2	.204	11	.200*	.918	11	0.300
	العينة3	.139	11	.200*	.964	11	0.817

بما أن مستوى الدلالة في اختبار شابيرو في العينات الثلاث (0.508 ، 0.3 ، 0.817) أكبر من 0.05 ، فإننا نقبل الفرضية المبدئية القائلة **بأن البيانات تتبع التوزيع الطبيعي .**

النتيجة: يمكن استخدام اختبار تحليل التباين

ثانياً / اختبار تحليل التباين

فرضية العدم: لا يوجد فرق بين متوسطات العينات الثلاث

الفرضية البديلة: يوجد فرق حقيقي بين متوسطات العينات الثلاث

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2954123.879	2	1477061.939	10.438	0.0000
Within Groups	4245070.364	30	141502.345		
Total	7199194.242	32			

و بما أن قيمة اختبار تحليل التباين = 10.438 و مستوى الدلالة = 0.000
لذلك نرفض فرضية العدم ونقبل الفرضية البديلة

النتيجة : أي أن يوجد اختلاف حقيقي بين متوسطات العينات الثلاث عند مستوى دلالة 0.01

ثالثاً / اختبار Bost Hoc : لمعرفة الاختلاف بين متوسطات كل عينتين

أ- معرفة تجانس التباين للعينات الثلاث:

Test of Homogeneity of Variances			
كمية الانتاج من الفراولة			
Levene Statistic	df1	df2	Sig.
3.918	2	30	0.031

فرضية العدم: يوجد تجانس للتباين بين العينات الثلاث

الفرضية البديلة:لا يوجد تجانس للتباين بين العينات الثلاث

نجد أن مستوى الدلالة 0.031 وهو أقل من 0.05، لذلك نرفض فرضية العدم ونقبل الفرضية البديلة

أي أن البيانات للعينات الثلاث ليس بها تجانس للتباين.

القرار : بما أن بيانات العينات الثلاث غير متجانسة التباين، فإننا نستخدم من خيارات Bost Hoc ، الخيار عدم تجانس التباين ونختار الخيار: Tamhane :

Multiple Comparisons						
Dependent Variable: الكمية لانتاجية الفراولة						
Tamhane						
					95% Confidence Interval	
(I) التربة_نوع	(J) التربة_نوع	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
رملية	طينية رملية	285.273	184.168	0.362	-199.91-	770.46
	طينية	727.273*	166.229	0.002	276.55	1177.99
طينية رملية	رملية	-285.273-	184.168	0.362	-770.46-	199.91
	طينية	442.000*	125.031	0.007	111.51	772.49
طينية	رملية	-727.273-*	166.229	0.002	-1177.99-	-276.55-
	طينية رملية	-442.000-*	125.031	0.007	-772.49-	-111.51-

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

بتطبيق اختبار Tamhane نجد أن:

* دالة إحصائياً الفروق للإنتاجية بين التربة الرملية والتربة الطينية ، وبين التربة الطينية الرملية والتربة الطينية .

* غير دالة إحصائياً: الفرق بين عينة التربة الرملية والتربة الطينية الرملية

مثال 3 : أخذت ثلاث عينات لثلاث تصنيفات من دول العالم حسب نسبة التحضر في هذه الدول، فكانت النتائج التالية :

الدولة	التصنيف	التحضر
Australia	1	85
Austria	1	58
Belgium	1	96
Canada	1	77
Denmark	1	85
Finland	1	60
France	1	73
Germany	1	85
Greece	1	63
Iceland	1	91
Ireland	1	57
Italy	1	69
Netherlands	1	89
New Zealand	1	84
Norway	1	75
Portugal	1	34
Spain	1	78
Sweden	1	84
Switzerland	1	62
UK	1	89
USA	1	75

الدولة	التصنيف	التحضر
Armenia	2	68
Azerbaijan	2	54
Bahrain	2	83
Egypt	2	44
Iran	2	57
Iraq	2	72
Jordan	2	68
Kuwait	2	96
Lebanon	2	84
Libya	2	82
Oman	2	11
Saudi Arabia	2	77
Syria	2	50
Turkey	2	61
U.Arab Em.	2	81
Uzbekistan	2	41

الدولة	التصنيف	التحضر
Argentina	3	86
Barbados	3	45
Bolivia	3	51
Brazil	3	75
Chile	3	85
Colombia	3	70
Costa Rica	3	47
Cuba	3	74
Domincan R.	3	60
Ecuador	3	56
El Salvador	3	44
Guatemala	3	39
Haiti	3	29
Honduras	3	44
Mexico	3	73
Nicaragua	3	60
Panama	3	53
Paraguay	3	48
Peru	3	70
Uruguay	3	89
Venezuela	3	91

(1) دول صناعية (2) دول الشرق الأوسط (3) دول أمريكا الجنوبية

المطلوب: هل يوجد اختلاف حقيقي بين متوسطات العينات الثلاث ؟ وهل يوجد اختلاف حقيقي بين العينة الأولى والثانية؟

خطوات الحل :

أولاً / نختبر إذا ما كانت البيانات تتبع توزيع طبيعي أم لا .

فرضية العدم: البيانات تتبع التوزيع الطبيعي **الفرضية البديلة:** البيانات لا تتبع التوزيع الطبيعي

بعد إدخال البيانات لبرنامج SPSS نحصل على النتيجة التالية :

Tests of Normality

التحضر		Kolmogorov-Smirnov[a]			**Statistic Shapiro-Wilk**		
	المجموعة	Statistic	df	Sig.		**df**	**Sig.**
	الدول الصناعية	.162	21	.155	**.924**	**21**	**0.103**
	الشرق الأوسط	.131	16	.200*	**.941**	**16**	**0.361**
	أمريكا اللاتينية	.113	21	.200*	**.953**	**21**	**0.392**

بما أن مستوى الدلالة في اختبار شابيرو في العينات الثلاث (0.508 ، 0.3 ، 0.817) أكبر من 0.05 ، فإننا نقبل الفرضية المبدئية القائلة **بأن البيانات تتبع التوزيع الطبيعي** .

النتيجة : يجب استخدام اختبار تحليل التباين

ثانياً / اختبار تجانس التباين

فرضية العدم : يوجد تجانس للتباين بين العينات الثلاث

الفرضية البديلة: لا يوجد تجانس للتباين بين العينات الثلاث

Test of Homogeneity of Variances			
التحضر			
Levene Statistic	df1	df2	Sig.
1.260	2	55	0.292

النتيجة : بما أن قيمة اختبار LEVNE = 1.260 ومستوى الدلالة= 0.292، وهي أكبر من 0.05 ، لذلك نقبل فرضية العدم القائلة بوجود تجانس للتباين.

لذلك نستخدم اختبار تحليل التباين ، واختبار LS'D

ثالثاً / اختبار تحليل التباين

فرضية العدم : لا يوجد فرق بين متوسطات العينات الثلاث

الفرضية البديلة : يوجد فرق حقيقي بين متوسطات العينات الثلاث

ANOVA					
التحضر					
	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	2028.307	2	1014.154	3.162	0.05
Within Groups	17638.676	55	320.703		
Total	19666.983	57			

و بما أن قيمة اختبار تحليل التباين = 3.162 ومستوى الدلالة = 0.05 ، لذلك نرفض فرضية العدم ونقبل الفرضية البديلة

أي أن يوجد اختلاف حقيقي بين متوسطات العينات الثلاث عند مستوى دلالة 0.05

ثالثاً / لمعرفة الفروق بين متوسطات العينات الثلاث ذات الدلالة الإحصائية نستعمل نجري اختبار Post Hoc لمعرفة الفرق بين العينات الثلاث بالشكل التالي :

Multiple Comparisons

Dependent Variable: التحضر

LSD

المجموعة (I)	المجموعة (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
الدول المتقدمة	الشرق الأوسط	10.402	5.943	**0.086**	-1.51-	22.31
	أمريكا اللاتينية	13.333*	5.527	**0.019**	2.26	24.41
الشرق الأوسط	الدولة المتقدمة	-10.402-	5.943	**0.086**	-22.31-	1.51
	أمريكا اللاتينية	2.932	5.943	**0.624**	-8.98-	14.84
أمريكا اللاتينية	الدول المتقدمة	-13.333-*	5.527	**0.019**	-24.41-	-2.26-
	الشرق الأوسط	-2.932-	5.943	**0.624**	-14.84-	8.98

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

من الجدول السابق نستنتج ما يلي :

1) بما أن مستوى الدلالة بين عينتي الدول المتقدمة والشرق الأوسط تساوي 0.086 وهو أكبر من 0.05 ، لذلك فإن الفرق بين متوسطي التحضر للدول المتقدمة و الشرق الأوسط ليس له دلالة إحصائية.

2) بما أن مستوى الدلالة بين الدول المتقدمة وأمريكا اللاتينية = 0.019 وهو أقل من 0.05 لذلك الفرق بين متوسطي التحضر بين العينتين ذو دلالة إحصائية.

الاختبارات اللا معلمية Non-Parametric Tests

تستخدم الاختبارات اللامعلمية للبيانات الرتبية والبيانات الكمية التي لا تتبع التوزيع الطبيعي، وبالرغم من استخدامها للبيانات الرتبية إلا أنه يمكن استعمالها للبيانات الكمية التي ليس لها توزيع طبيعي سواء كان الالتواء سالباً أو موجباً، ويستعمل اختبار الإشارة لاختبار الفرق بين المتوسط الحسابي لعينة مع مجتمعها الإحصائي وهي تقابل اختبار t لعينة واحدة في الاختبارات المعلمية ، أما في حالة عينتين مستقلتين فنستعمل اختبار مان وتني لاختبار الفرق بين رتب المتوسطين الحسابيين للعينتين، وإذا كانت العينتان مرتبطتين فإننا نستخدم اختبار ويلكوكسن، أما في حالة ثلاث عينات فأكثر فإننا نستعمل اختبار كروسكال ولاس وهو يقابل اختبار تحليل التباين المعلمي.

أولاً / اختبار متوسط عينة واحدة مع المتوسط الحسابي لمجتمع إحصائي مأخوذة منه

1: **اختبار الإشارة** : يستخدم لاختبار متوسط عينة واحدة مأخوذة من مجتمع إحصائي، مع المتوسط الحسابي للمجتمع الإحصائي

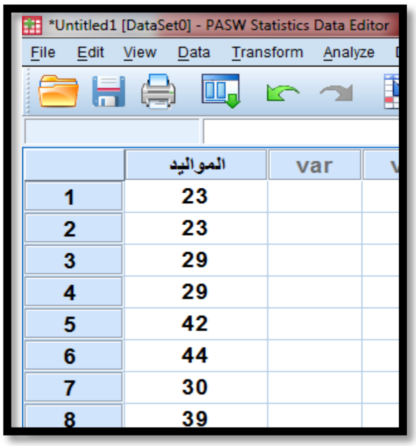
السؤال الأول : يبين الجدول التالي معدل المواليد في مجموعة من دول الشرق الأوسط

الدولة	أرمينيا	أذربيجان	البحرين	مصر	إيران	العراق	فلسطين	الأردن	الكويت
نسبة المواليد /ألف نسمة	23	23	29	29	42	44	30	39	28
الدولة	لبنان	ليبيا	عمان	السعودية	سوريا	تركيا	الأمارات	أوزبكستان	
نسبة المواليد /ألف نسمة	27	45	40	38	44	26	28	30	

إذا كان معدل المواليد في العالم 26 في الألف ، فهل يوجد اختلاف بين المتوسط الحسابي لدول الشرق الأوسط عن معدل العالم ؟

خطوات الحل :

أولاً : نقوم بادخال البيانات كما هو في الشكل التالي :



ثانياً / عمل اختبار التوزيع الطبيعي للبيانات

فرضية العدم : البيانات تتبع التوزيع الطبيعي

الفرضية البديلة : البيانات لا تتبع التوزيع الطبيعي

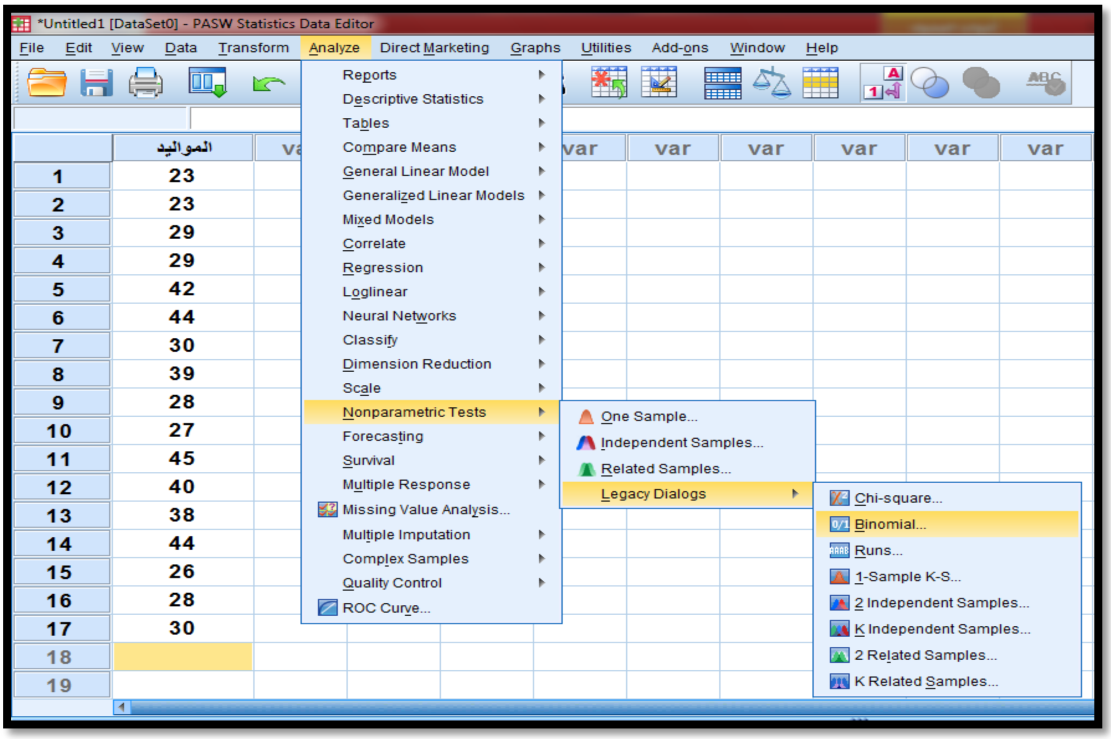
	Kolmogorov-Smirnov(a)			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
المواليد	0.250	17	.006	.881	17	0.033

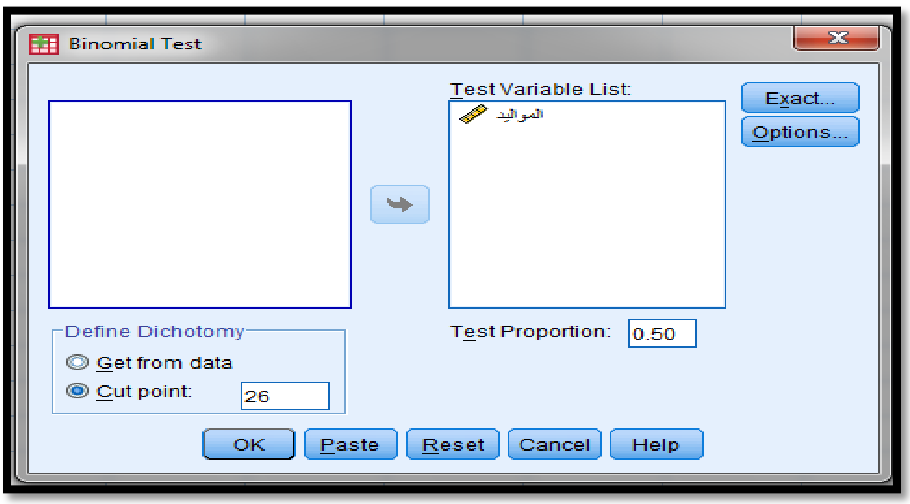
النتيجة : بما أن مستوى الدلالة في اختبار التوزيع الطبيعي (اختبار Shapiro) يساوي 0.033 وهي أقل من 0.05 لذلك نرفض فرضية العدم و نقبل الفرضية البديلة القائلة بأن البيانات لا تتبع التوزيع الطبيعي، إذن **البيانات لا تتبع التوزيع الطبيعي**

في ضوء ذلك يجب استعمال اختبار غير معلمي و هو <u>اختبار الإشارة</u>

فرضية العدم: لا يختلف المتوسط الحسابي للمواليد في مجموعة الشرق الأوسط عن المعدل العام للمواليد في دول العالم

الفرضية البديلة: يختلف المتوسط الحسابي للمواليد في مجموعة الشرق الأوسط عن المعدل العام للمواليد في دول العالم





اختبار الإشارة

		Category	N	Observed Prop.	Test Prop.	Exact Sig. (2-tailed)
العينة	Group 1	<= 26	3	0.18	0.50	0.013
	Group 2	> 26	14	0.82		
	Total		17	1.00		

بما أن قيمة مستوى الدلالة = 0.013 و هي أقل من 0.05 لذلك نرفض فرضية العدم ونقبل الفرضية البديلة

يوجد اختلاف حقيقي بين المتوسط الحسابي للمواليد في مجموعة الشرق الأوسط والمعدل العام في العالم (الفرق بين متوسط العينة والمتوسط العام فرق حقيقي ودال إحصائياً)

مثال2 :

يبين الجدول التالي عينة مأخوذة من عشر خزانات لمياه الشرب في محافظات قطاع غزة ووجد نسبة الكلوريد بها كالتالي:

رقم الخزان	1	2	3	4	5	6	7	8	9	10
نسبة الكلوريد	40	60	50	180	180	170	190	60	220	230

هل يختلف المتوسط الحسابي للعينة السابقة عن المتوسط العام للمياه في خزانات مياه الشرب التابعة لمصلحة بلديات مياه الساحل 150 مليجرام /لتر.

الاجابة :

الحل : <u>أولاً نختبر إذا ما كانت البيانات تتبع التوزيع الطبيعي</u>

فرضية العدم : البيانات تتبع التوزيع الطبيعي

الفرضية البديلة : البيانات لا تتبع التوزيع الطبيعي

Tests of Normality						
	Kolmogorov-Smirnova			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
الكلوريد	.263	10	.048	.832	10	0.035

بما أن قيمة مستوى الدلالة = 0.035 حسب اختبار Shapiro-Wilk و هي أقل من مستوى الدلالة 0.05 لذلك نرفض فرضية العدم ونقبل الفرضية البديلة القائلة بأن البيانات لا تتبع التوزيع الطبيعي .

<u>لذلك نستخدم اختبار الاشارة .</u>

فرضية العدم: لا يختلف متوسط كمية الكلوريد في العينة عن المتوسط العام 150

الفرضية البديلة: يختلف متوسط كمية الكلوريد في العينة عن المتوسط العام 150 بدلالة إحصائية.

Binomial Test						
		Category	N	Observed Pro	Test Prop.	Exact Sig. (2-tailed)
الكلوريد	Group 1	<= 150	4	.40	.50	0.7540
	Group 2	> 150	6	.60		
	Total		10	1.00		

بما أن مستوى الدلالة لاختبار الاشارة = 0.754 وهو أكبر من 0.05 لذلك لا نستطيع أن نرفض فرضية العدم، ولا نستطيع قبول الفرضية البديلة .

لذلك فإن الاختلاف أو الفرق بين متوسط العينة والمتوسط العام فرق ليس له دلالة إحصائية.

مثال3:

أخذت عينة من 10 محطات للأمطار فوجد بها نسبة كلوريد كالتالي :

رقم المشاهدة	1	2	3	4	5	6	7	8	9	10	11
نسبة الكلوريد	20	18	20	30	17	25	40	20	30	19	21

المطلوب : إذا كانت نسبة الكلوريد في الأمطار التي تسقط على هذه المحطات تساوي 30 فهل يوجد اختلاف بين متوسط نسبة الكلوريد في هذه المحطات و المعدل العام ؟

الحل :

أولاً / اختبار التوزيع الطبيعي: نختبر إذا ما كانت البيانات تتبع التوزيع الطبيعي أم لا

فرضية العدم: البيانات تتبع التوزيع الطبيعي **الفرضية البديلة:** البيانات لا تتبع التوزيع الطبيعي

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
الكلوريد	.282	11	.014	0.819	11	0.0170

بما أن قيمة اختبار شابيرو = 0.017، وهي أقل من 0.05، لذلك نرفض فرضية العدم ونقبل الفرضية البديلة أي أن البيانات لا تتبع التوزيع الطبيعي، وعليه يجب استخدام اختبار الاشارة لمعرفة الاختلاف بين متوسط العينات والمتوسط العام.

ثانياً / اختبار الإشارة :

فرضية العدم: لا يوجد اختلاف حقيقي بين متوسط نسبة الكلوريد في العينات العشر والمتوسط العام .

الفرضية البديلة: يوجد اختلاف حقيقي بين متوسط نسبة الكلوريد في العينات العشر والمتوسط العام .

Binomial Test						
		Category	N	Observed Prop.	Test Prop.	Exact Sig. (2-tailed)
الكلوريد	Group 1	<= 30	10	.910	0.50	0.012
	Group 2	> 30	1	.090		

Binomial Test						
		Category	N	Observed Prop.	Test Prop.	Exact Sig. (2-tailed)
الكلوريد	Group 1	<= 30	10	.910	0.50	0.012
	Group 2	> 30	1	.090		
	Total		11	1.00		

بما أن مستوى الدلالة لاختبار الإشارة = 0.012 ، وهي أقل من 0.05

لذلك نرفض فرضية العدم ونقبل الفرضية البديلة

أي أنه يوجد اختلاف حقيقي بين متوسط نسبة الكلوريد في العينات العشر والمتوسط العام .

مثال2 :

في دراسة حول ضرورة توجه الجغرافيين نحو الدراسات البيئية، أخذت عينة من مجموعة من الجغرافيين العاملين في مجال البيئي في قطاع غزة:

حسب مقياس ليكرت

الموافقة	معارض جداً	معارض	محايد	موافق	موافق جداً
الدرجة	1	2	3	4	5

جدول يبين رأي الصحفيين في التوجه نحو الصحافة الإلكترونية

رقم الجغرافي	1	2	3	4	5	6	7	8	9
الرأي	موافق	موافق	موافق بشدة	موافق	موافق بشدة	موافق	موافق بشدة	موافق	موافق بشدة
رقم الجغرافي	10	11	12	13	14	15	16	17	18
	محايد	موافق	موافق بشدة	معارض	معارض بشدة	موافق	موافق بشدة	موافق	محايد

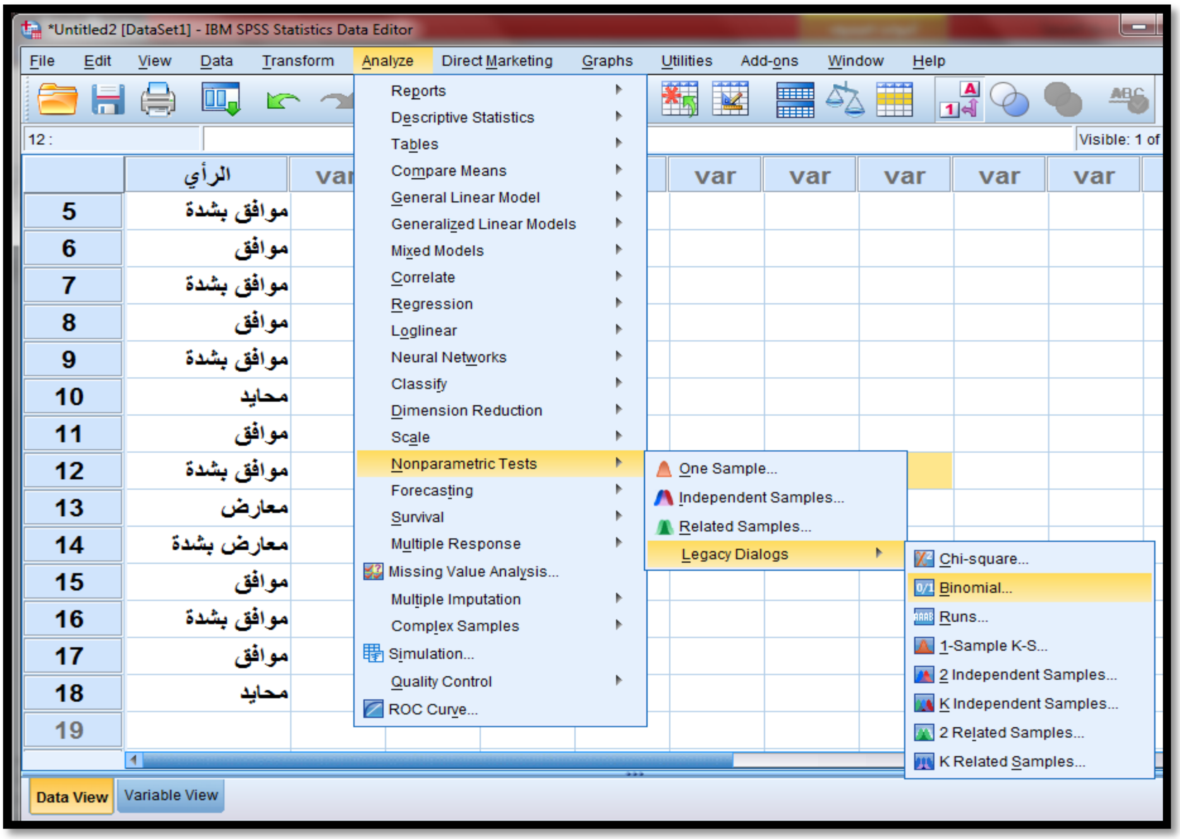
هل يختلف المتوسط الحسابي للعينة السابقة عن المتوسط العام لآراء الجغرافيين في التوجه نحو العمل في المجال البيئي .

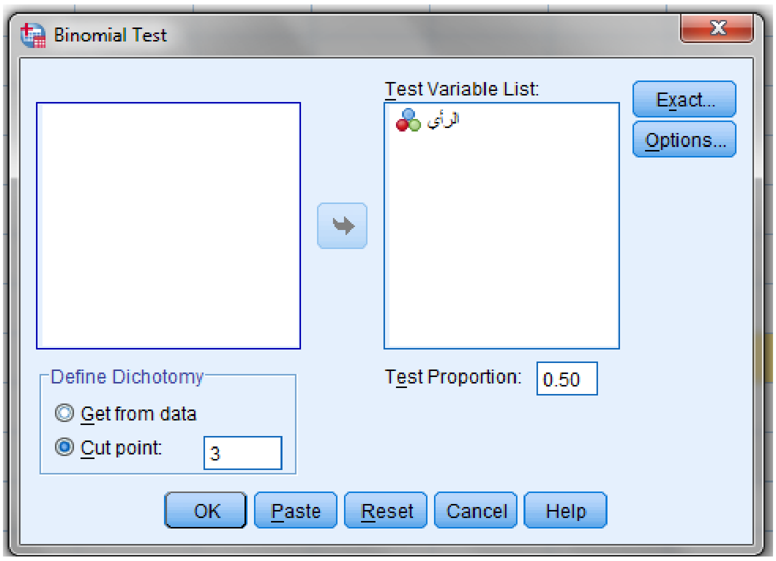
الاجابة :

بما أن البيانات رتبية ، لذلك نستخدم اختبار الاشارة .

فرضية العدم : لا يختلف متوسط رأي العينة عن المتوسط العام

الفرضية البديلة : يختلف متوسط رأي العينة عن المتوسط العام بدلالة إحصائية أقل من 0.05





Binomial Test						
		Category	N	Observed Prop.	Test Prop.	Exact Sig. (2-tailed)
الرأي	Group 1	<= 3	4	.22	0.50	0.031
	Group 2	> 3	14	.78		
	Total		18	1.00		

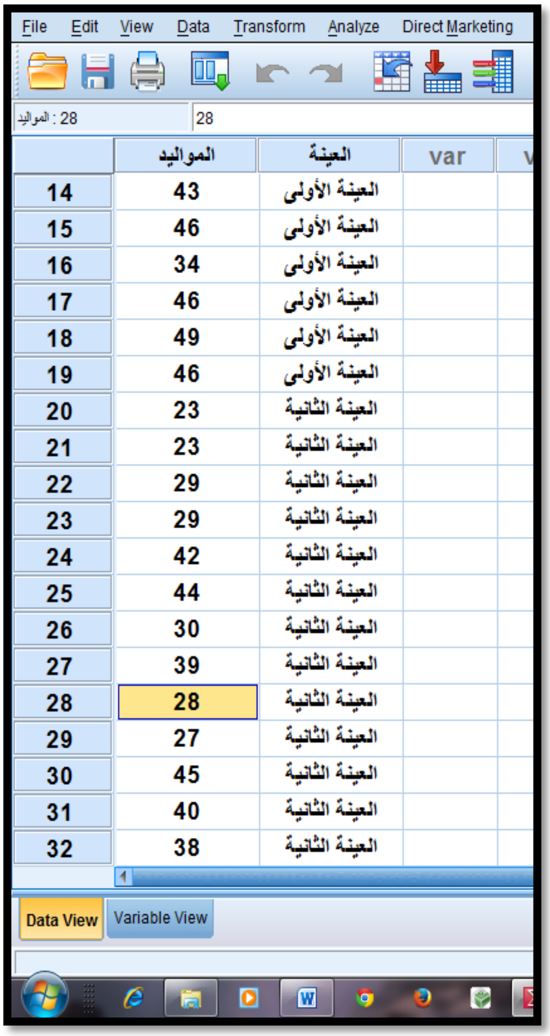
بما أن مستوى الدلالة لاختبار الاشارة = 0.031 و هو أقل من 0.05 لذلك نرفض فرضية العدم ونقبل الفرضية البديلة

لذلك فإن الاختلاف أو الفرق بين متوسط العينة والمتوسط العام فرق له دلالة إحصائية .

ثانياً / اختبار مان وتني لعينتين مستقلتين:

يستخدم لاختبار متوسط الحسابي لعينتين مستقلتين، للبيانات الرتبية بشكل أساسي، كما يستخدم للبيانات الكمية لا تتبع التوزيع الطبيعي.

مثال 1 :



العينة الأولى (افريقيا)		العينة الثانية (الشرق الأوسط)	
الدولة	عدد المواليد في الألف	الدولة	عدد المواليد في الألف
بتسوانا	32	أرمينيا	23
بوركينا فاسو	47	أذربيجان	23
بورندي	44	البحرين	29
الكاميرون	41	مصر	29
افريقيا الوسطى	44	إيران	42
إثيوبيا	45	العراق	44
الجابون	28	فلسطين	30
جامبيا	46	الأردن	39
كينيا	42	الكويت	28
ليبيريا	43	لبنان	27
المغرب	29	ليبيا	45
نيجيريا	44	عمان	40
رواندا	49	السعودية	38
السنغال	43	سوريا	44
الصومال	46	تركيا	26
جنوب افريقيا	34	الإمارات	28
تنزانيا	46	أوزبكستان	30
أوغندا	49		
زامبيا	46		

المطلوب : هل يوجد اختلاف بين متوسط العينتين بدلالة احصائية ؟

خطوات الحل :

أولاً / عمل الاختبار التوزيع الطبيعي للبيانات

فرضية العدم: البيانات تتبع التوزيع الطبيعي **الفرضية البديلة:** البيانات لا تتبع التوزيع الطبيعي

Tests of Normality							
	المجموعة	Kolmogorov-Smirnova			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
المواليد	افريقيا	.246	19	.004	0.818	19	.002
	الشرق الاوسط	.250	17	.006	0.881	17	.033

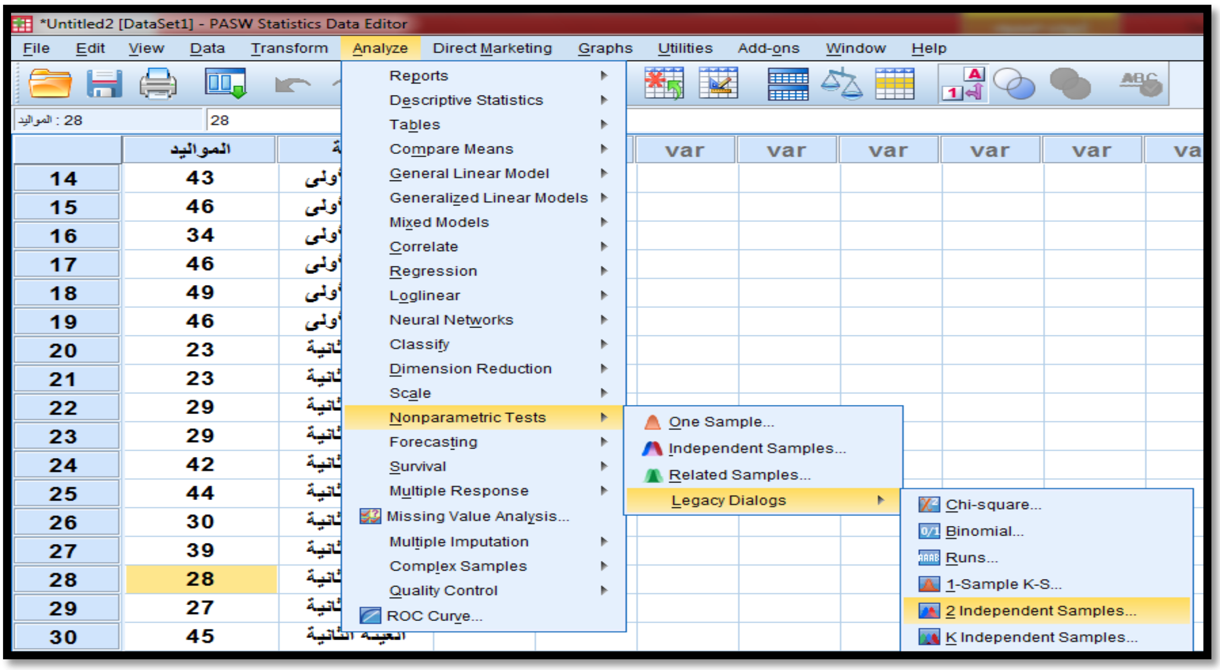
بما أن مستوى الدلالة في اختبار التوزيع الطبيعي اختبار Shapiro يساوي (0.033 و 0.002) هي أقل من 0.05 ، لذلك نرفض فرضية العدم ونقبل الفرضية البديلة

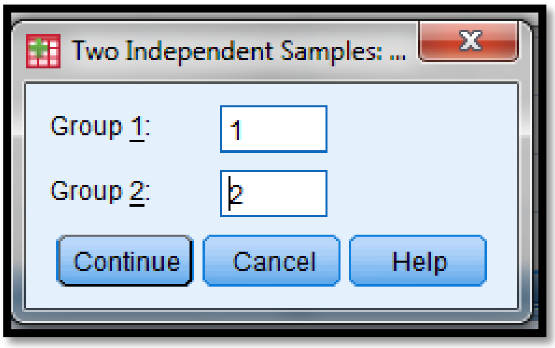
النتيجة البيانات لا تتبع التوزيع الطبيعي

في ضوء ذلك يجب استعمال اختبار غير معلمي وهو اختبار مان وتني

فرضية العدم: لا يختلف المتوسط الحسابي للمواليد في مجموعة الشرق الأوسط عن المعدل مواليد في افريقيا

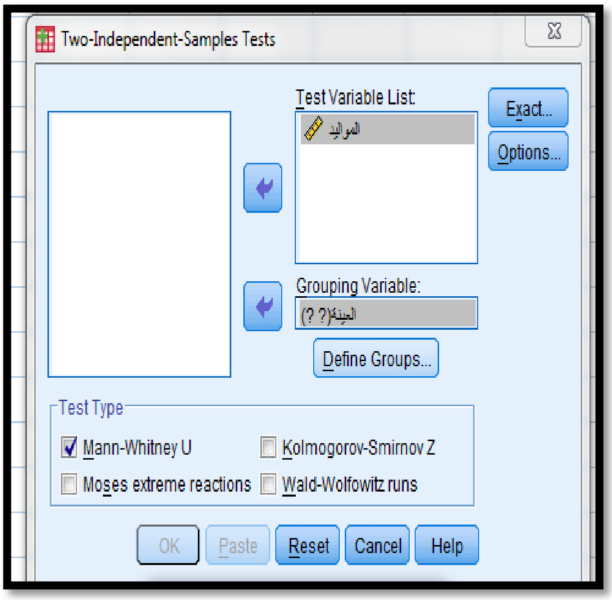
الفرضية البديلة: يختلف المتوسط الحسابي للمواليد في مجموعة الشرق الأوسط عن المعدل للمواليد في افريقيا





Test Statistics[b]

	المواليد
Mann-Whitney U	56.000
Wilcoxon W	209.000
Z	-3.353-
Asymp. Sig. (2-tailed)	0.001
Exact Sig. [2*(1-tailed Sig.)]	0.001[a]



بما أن قيمة مستوى الدلالة = (0.001) وهي أقل من 0.05 لذلك نرفض فرضية العدم ونقبل الفرضية البديلة .

يوجد اختلاف حقيقي بين المتوسط الحسابي للمواليد في مجموعة الشرق الأوسط والمتوسط الحسابي لإفريقيا (الفرق بينهم حقيقي ودال إحصائياً) عند مستوى دلالة 0.05

مثال2 :

في دراسة جغرافية حول ظاهرة المساحات الفراغ داخل الكتل العمرانية، أجرى أحد الباحثين مقارنة بين مدينتين بالشكل التالي :

مساحة القطعة /دونم	أقل من دونم	1	2	3	4	5	6	7	8	9	أكثر من 9
عدد القطع/ المدينة "أ"	400	300	180	120	100	90	70	60	50	10	2
عدد القطع/ المدينة"ب"	600	500	140	80	70	50	30	20	10	5	1

المطلوب : هل يوجد اختلاف بين متوسط عدد القطع في المدينتين ؟

الحل :

أولاً / نختبر إذا ما كانت البيانات تتبع التوزيع الطبيعي أم لا

فرضية العدم: البيانات تتبع التوزيع الطبيعي **الفرضية البديلة:** البيانات لا تتبع التوزيع الطبيعي

Tests of Normality							
	النوع	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
العدد	المدينة أ	.245	11	.063	.846	11	.0380
	المدينة ب	.334	11	.001	.666	11	.0000

بما أن عدد البيانات اقل من 30 مفردة ، لذا سنستعمل اختبار شابيرو Shapiro–Wilk

بما أن قيمة مستوى الدلالة في اختبار شابيرو أقل من 0.05 لذلك نرفض فرضية العدم ونقبل الفرضية البديلة **(البيانات لا تتبع التوزيع الطبيعي)**

لذلك نستخدم اختبار لا معلمي (مان وتني) لعينتين مستقلتين

ثانياً / اختبار مان وتني لعينتين مستقلتين :

فرضية العدم : لا يوجد فرق حقيقي بين متوسطي العينتين

الفرضية البديلة: يوجد فرق حقيقي بين متوسطي العينتين عند مستوى دلالة 0.01

Test Statistics[b]	
Mann-Whitney U	48.500
Wilcoxon W	114.500
Z	-.789-
Asymp. Sig. (2-tailed)	0.4300
Exact Sig. [2*(1-tailed Sig.)]	0.438[a]0

Ranks

	النوع	N	Mean Rank	Sum of Ranks
العدد	المدينة أ	11	12.59	138.50
	المدينة ب	11	10.41	114.50
	Total	22		

نلاحظ وجود فرق واضح في متوسط الرتب

(10.4 ، 12.6)

كما أن قيمة اختبار مان وتني = 48.5 ، وبمستوى دلالة 0.438 ، و هي أكبر من 0.05 لذلك لا نستطيع أن نرفض فرضية العدم، ولا نستطيع قبول الفرضية البديلة

وعليه فالفرق بين المتوسطين فرق ليس له دلالة إحصائية

مثال 3:

هل توجد فروق ذات دلالة إحصائية عند مستوى دلالة أقل من 0.05 بين متوسطات استجابات المبحوثين من طلبة الدراسات العليا حول استخدام برامج GIS في الدراسات الجغرافية تعزى لمتغير الجنس " النوع" :

جدول يبين رأي الطلبة في استخدام برامج GIS في الدراسات الجغرافيا

رقم الطالب	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
الرأي	موافق	موافق	موافق بشدة	موافق	موافق بشدة	موافق	موافق بشدة	موافق	موافق بشدة	موافق
رقم الطالب	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
الرأي	محايد	موافق	موافق بشدة	معارض	معارض بشدة	موافق	موافق بشدة	موافق	معارض بشدة	محايد

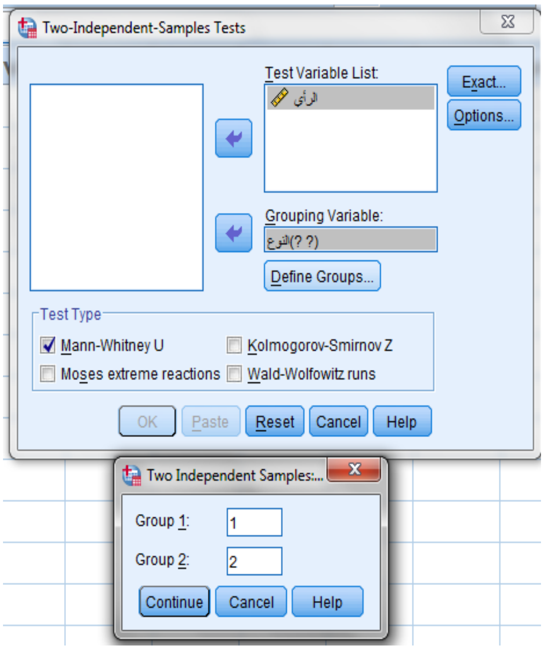
الحل:

بما أن البيانات للعينتين بيانات رتبية :

لذلك نستخدم اختبار لا معلمي (مان وتني) لعينتين مستقلتين

فرضية العدم: لا يوجد فرق حقيقي بين متوسطي آراء العينتين في استخدام برامج GIS في الدراسات الجغرافية

الفرضية البديلة: يوجد فرق حقيقي بين متوسطي آراء العينتين في ا ستخدام برامج GIS في الدراسات الجغرافية.



Ranks

	النوع	N	Mean Rank	Sum of Ranks
الرأي	male	10	13.00	130.00
	female	10	8.00	80.00
	Total	20		

Test Statistics[a]

	الرأي
Mann-Whitney U	**25.000**
Wilcoxon W	**80.000**
Z	**-2.012-**
Asymp. Sig. (2-tailed)	**0.044**
Exact Sig. [2*(1- ailed Sig.)]	**0.063[b]**

النتيجة: نلاحظ وجود فرق واضح في متوسط الرتب (13 ، 8)، كما أن قيمة اختبار مان وتني= − 2.01، وبمستوى دلالة 0.044 ، و هي أقل من 0.05

لذلك نرفض فرضية العدم، ونقبل الفرضية البديلة

وعليه فالفرق بين المتوسطين فرق له دلالة إحصائية عند مستوى دلالة أقل من 0.05 ، أي أنه يوجد اختلاف بين الذكور والإناث في آرائهم حول استخدام برامج GIS في الدراسات الجغرافية.

ثالثاً / اختبار ويلكوكسن لعينتين مرتبطتين :

يستخدم في حالة عينتين مزدوجتين، أي لها قيمة **قبل** التعديل، وقيمة **بعد** التعديل

مثال1 :

في دراسة هيدرولوجية حول معالجة المياه الجوفية، قامت سلطة المياه بأخذ عينة من 10 آبار في محافظة شمال غزة وتم قياس نسبة النترات في هذه الآبار، وتم تطبيق نظام معالجة لخفض نسبة النترات، فأخذت عينة جديدة من نفس الآبار بعد سنة من تطبيق هذا النظام، فتم الحصول على النتائج التالية :

رقم البئر	A	B	C	D	E	F	H	J	I	K
العينة الأولى (قبل)	80	90	70	100	150	250	150	80	90	100
العينة الثانية (بعد)	60	75	70	80	100	150	100	55	53	60

المطلوب :

هل تستطيع أن نحكم على ضوء تلك النتائج ، و بمستوى معنوية 0.05 ، أن برنامج خفض نسبة النترات في المياه الجوفية بمحافظة شمال غزة كان ناجحاً ؟

الحل:

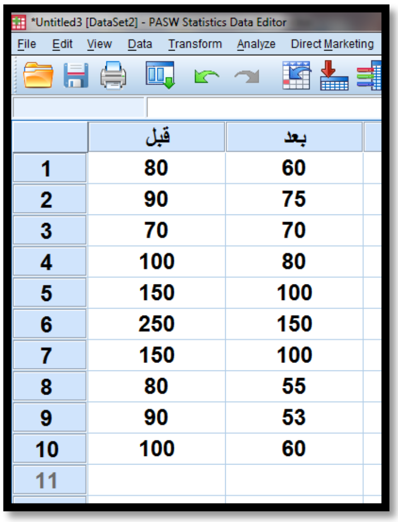
أولاً / اختبار التوزيع الطبيعي: تختبر إذا ما كانت البيانات تتبع التوزيع الطبيعي :

فرضية العدم: البيانات تتبع التوزيع الطبيعي **الفرضية البديلة:** البيانات لا تتبع التوزيع الطبيعي

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
قبل	0.315	10	0.006	0.761	10	0.005
بعد	0.204	10	0.200*	0.836	10	0.040

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.



أن مستوى الدلالة لاختبار شابيرو = (0.005 ، 0.040) أقل من 0.05 ، لذلك نرفض فرضية العدم ونقبل الفرضية البديلة القائلة بأن البيانات لا تتبع توزيع طبيعي، وعليه نستخدم **اختبار ويلكوكسن**

ثانياً/ اختبار ويلكوكسن اللامعلمي:

فرضية العدم: لا يوجد اختلاف بين متوسط العينتين

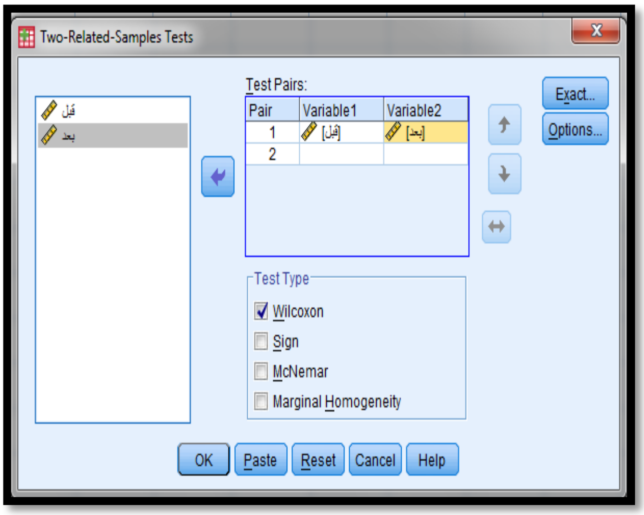
الفرضية البديلة: يوجد اختلاف حقيقي بين متوسطي العينتين

Test Statistics[b]

	بعد - قبل
Z	-2.670-[a]
Asymp. Sig. (2-tailed)	0.008

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test



نجد أن قيمة اختبار ويلكوكسن= −2.67 و مستوى الدلالة = 0.008 وهو أقل من 0.05

لذلك نرفض فرضية العدم ونقبل الفرضية البديلة.

النتيجة : أي أنه يوجد اختلاف حقيقي بين متوسطي العينتين وهذا الفرق دال احصائياً

مثال 2:

في دراسة حول جدوى الدراسة الميدانية في فهم الظاهرات الجيومورفولوجية ، قام مدرس الجيومورفولوجيا في نهاية تدريس المساق بعمل اختبار للطلبة به أشكال جيومورفولوجية مختلفة، ثم خرج بالطلبة للعمل الميداني، وأجرى نفس الاختبار مرة ثانية، فكانت النتائج كالتالي :

رقم العينة	1	2	3	4	5	6	7	8	9	10
قبل الرحلة الميدانية	59	60	50	60	60	55	40	60	88	90
بعد الرحلة الميدانية	88	64	85	65	90	95	80	85	90	90

المطلوب : هل استفاد الطلبة من الرحلة الميدانية في تحسين فهمهم للظواهر الجغرافية حسب العينة، وهل يمكن الاعتماد على هذه النتيجة و تعميمها على باقي الطلبة

أولاً / اختبار التوزيع الطبيعي : نختبر هل البيانات تتبع توزيع طبيعي أم لا

فرضية العدم: البيانات تتبع توزيع طبيعي

الفرضية البديلة: البيانات لا تتبع توزيع طبيعي

Tests of Normality

Shapiro-Wilk			Kolmogorov-Smirnov(a)			
Sig.	df	Statistic	Sig.	df	Statistic	
0.035	10	0.832	.200(*)	10	.198	قبل
0.043	10	0.839	.200(*)	10	.214	بعد

بما أن مستوى الدلالة للعينتين في اختبار شابيرو (**0.035**، **0.043**)، وهي أقل من 0.05

لذلك نرفض فرضية العدم و نقبل الفرضية البديلة.

أي أن البيانات لا تتبع التوزيع الطبيعي

إذن نستخدم اختبار ويلكوكسن غير المعلمي لعينتين مرتبطتين (مزدوجتين)

ثانياً / اختبار ويلكوكسن :

فرضية العدم : لا يوجد اختلاف بين متوسط درجات الطلبة في مساق الجيومورفولوجيا قبل الرحلة الميدانية و متوسط درجات الطلبة بعد الرحلة الميدانية

الفرضية البديلة : يوجد اختلاف بين متوسط درجات الطلبة في مساق الجيومورفولوجيا قبل الرحلة الميدانية و متوسط درجات الطلبة بعد الرحلة الميدانية

تطبيق اختبار ويلكوكسن

Test Statistics[b]

	بعد - قبل
Z	-2.499-[a]
Asymp. Sig. (2-tailed)	0.012

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

بما أن قيمة اختبار ويلكوكسن – 2.499 ، ومستوى الدلالة للاختبار يساوي **0.012** و هي أقل من 0.05

نرفض فرضية العدم و نقبل الفرضية البديلة

النتيجة : يوجد اختلاف بين متوسط درجات الطلبة في مساق الجيومورفولوجيا قبل الرحلة الميدانية ومتوسط درجات الطلبة بعد الرحلة الميدانية أن الفرق بين المتوسطين فرق حقيقي ذو دلالة إحصائية عند مستوى 0.05

و من ذلك نستنتج أن الطلبة قد استفادوا من الرحلة الميدانية .

مثال 3:

في دراسة حول أهمية مساق الإحصاء التطبيقي في البحوث الجغرافية أخذت عينة من عشرة طلبة وأخذت آراءهم في مدى قناعتهم في أهمية الإحصاء في البحوث الجغرافية ثم بعد دراسة المساق أخذت آراؤهم مرة ثانية فكانت النتائج التالية:

جدول يبين رأي الجغرافيين في أهمية الإحصاء في البحوث الجغرافية

رقم الباحث	1	2	3	4	5	6	7	8	9	10
الاختبار القبلي	موافق	موافق	موافق	معارض	معارض بشدة	موافق	موافق	موافق	معارض بشدة	موافق
الاختبار البعدي	موافق	موافق	موافق بشدة	موافق	موافق بشدة	موافق بشدة	محايد	موافق بشدة	موافق	محايد

بما أن البيانات للعينتين بيانات رتبية :

إذن نستخدم اختبار ويلكوكسن غير المعلمي لعينتين مرتبطتين (مزدوجتين)

فرضية العدم : لا يوجد اختلاف بين متوسط آراء الجغرافيين في أهمية الإحصاء في بحوث الجغرافيا

الفرضية البديلة : يوجد اختلاف بين متوسط آراء الجغرافيين في أهمية الإحصاء في بحوث الجغرافيا

Ranks

		N	Mean Rank	Sum of Ranks
الأول - الثاني	Negative Ranks	1[a]	3.50	3.50
	Positive Raks	7[b]	4.64	32.50
	Ties	2[c]		
	Total	10		

Test Statistics[a]

	الأول - الثاني
Z	-2.124-[b]
Asymp. Sig. (2-tailed)	0.034

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

	الأول	الثاني
1	موافق	موافق
2	موافق	موافق
3	موافق	موافق بشدة
4	محايد	موافق
5	معارض بشدة	موافق بشدة
6	موافق	موافق بشدة
7	موافق	محايد
8	موافق	موافق بشدة
9	معارض بشدة	موافق
10	محايد	موافق

بما أن مستوى الدلالة لاختبار ويلكوكسن يساوي **0.034** ، وهي أقل من 0.05

لذلك نرفض فرضية العدم ونقبل الفرضية البديلة

يوجد اختلاف بين متوسط بين متوسط آراء الجغرافيين في أهمية الإحصاء في بحوث الجغرافيا

أن الفرق بين المتوسطين فرق حقيقي ذو دلالة إحصائية عند مستوى 0.05

ومن ذلك نستنتج أن الطلبة قد استفادوا من دراسة مساق الإحصاء .

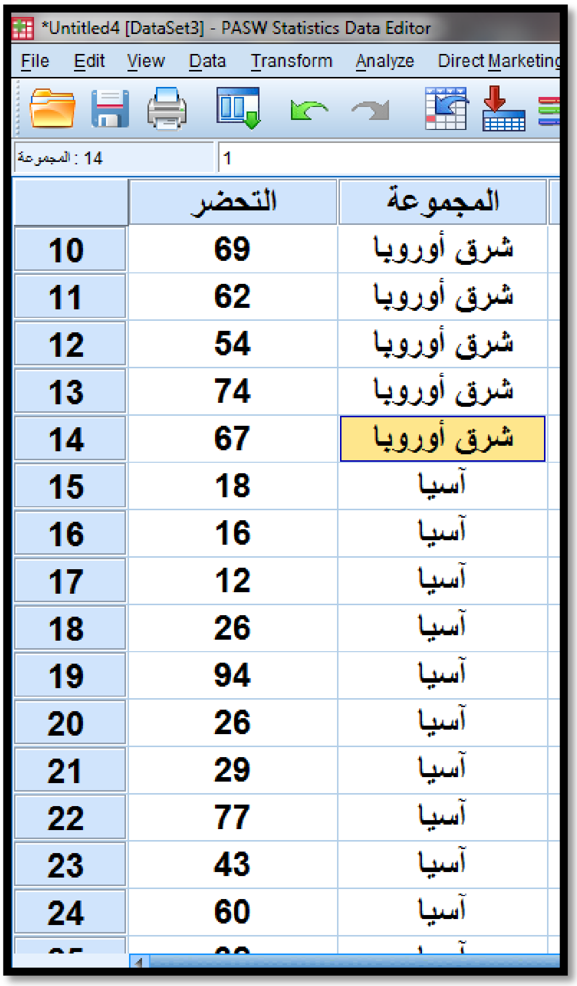
رابعاً / اختبار كروسكال ولاس:

يستخدم لثلاث عينات فأكثر لبيانات رتبية ، أو بيانات لا تتبع التوزيع الطبيعي

مثال 1 :

يبين الجدول التالي نسبة التحضر في الدول التالية :

شرق أوروبا	آسيا	إفريقيا
65	18	25
36	16	15
68	12	5
51	26	40
30	94	47
72	26	12
56	29	46
64	77	23
71	43	24
69	60	45
62	32	46
54	43	35
74	72	6
67	100	40
	71	24
	22	49
	20	21
		8
		8



*Untitled4 [DataSet3] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing

14 : المجموعة 1

	التحضر	المجموعة
10	69	شرق أوروبا
11	62	شرق أوروبا
12	54	شرق أوروبا
13	74	شرق أوروبا
14	67	شرق أوروبا
15	18	آسيا
16	16	آسيا
17	12	آسيا
18	26	آسيا
19	94	آسيا
20	26	آسيا
21	29	آسيا
22	77	آسيا
23	43	آسيا
24	60	آسيا

المطلوب : هل يوجد اختلاف في نسبة التحضر بين متوسطات المجموعات الثلاثة؟

خطوات الحل :

أولاً/ اختبار التوزيع الطبيعي: نختبر إذا ما كانت البيانات تتبع التوزيع الطبيعي

فرضية العدم: البيانات تتبع التوزيع الطبيعي **الفرضية البديلة:** البيانات لا تتبع التوزيع الطبيعي

Tests of Normality

المجموعة	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
التحضر شرق أوروبا	.205	14	.116	.860	14	0.030
آسيا	.202	17	.064	.886	17	0.041
إفريقيا	.159	19	.200*	.900	19	0.049

a. Lilliefors Significance Correction

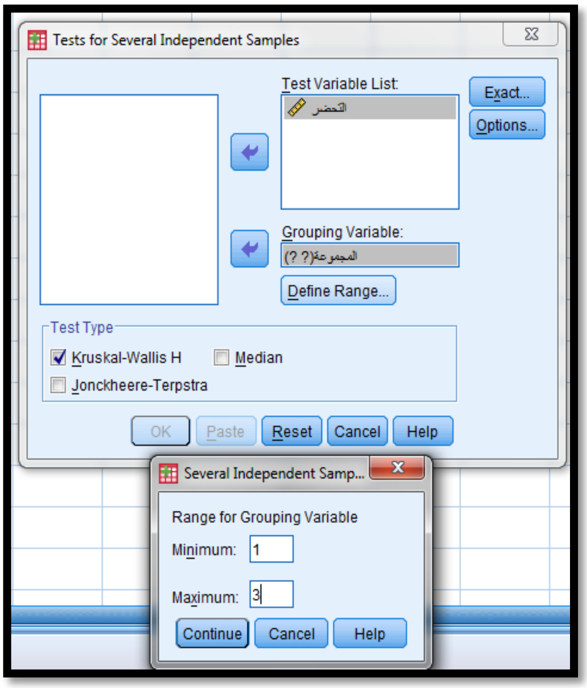
*. This is a lower bound of the true significance.

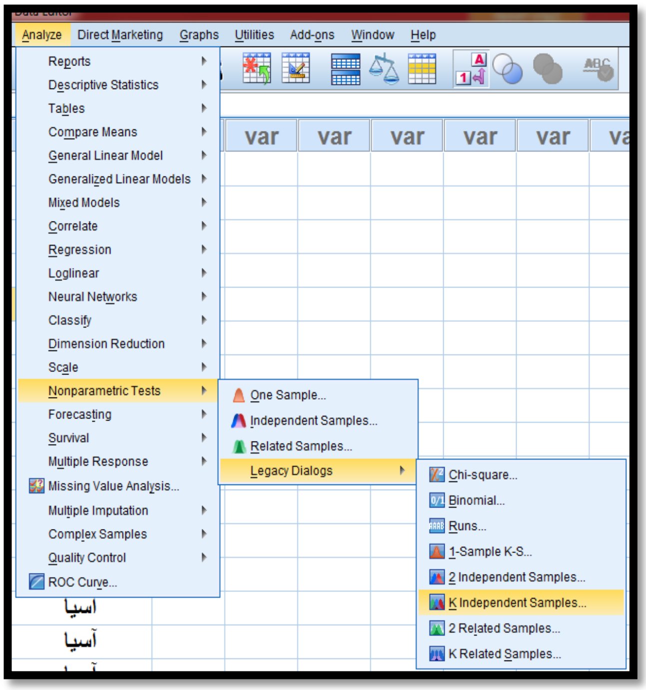
النتيجة: بما أن العينات الثلاث أقل من 30 مفردة، لذلك نستخدم اختبار شابيرو، ونجد أن مستوى الدلالة للعينات الثلاث (0.03 ، 0.041 ، 0.049) وهي أقل من 0.05، **لذلك نرفض فرضية العدم ونقبل الفرضية البديلة، أي أن البيانات لا تتبع التوزيع الطبيعي، لذلك يجب أن نستخدم اختبار لا معلمي (كروسكال ولاس) لوجود ثلاث عينات لا تتبع التوزيع الطبيعي.**

ثانياً / اختبار كروسكال ولاس:

فرضية العدم: لا يوجد اختلاف بين متوسطات التحضر للمجموعات الثلاثة

الفرضية البديلة: يوجد اختلاف (فرق) بين متوسطات التحضر للمجموعات الثلاثة





Test Statistics[a,b]

	التحضر
Chi-square	16.081
Df	2
Asymp. Sig.	0.000

a. Kruskal Wallis Test

b. Grouping Variable: المجموعة

Ranks

	المجموعة	N	Mean Rank
التحضر	شرق أوروبا	14	37.00
	آسيا	17	26.15
	إفريقيا	19	16.45
	Total	50	

نلاحظ وجود اختلاف بين متوسط الرتب للمجموعات الثلاث

نجد أن مستوى الدلالة لاختبار كروسكال ولاس 0.000 و هو أقل من 0.05 لذلك نرفض فرضية العدم و نقبل الفرضية البديلة

أي أنه يوجد اختلاف حقيقي بين متوسطات التحضر للمجموعات الثلاثة و الفرق بينهما فرق حقيقي و دال إحصائياً .

مثال 2 :

أخذت ثلاث عينات من طلبة الجامعات من ثلاث محافظات في قطاع غزة حول رأيهم في المشاركة في الانتخابات الفلسطينية ، فكانت النتائج التالية :

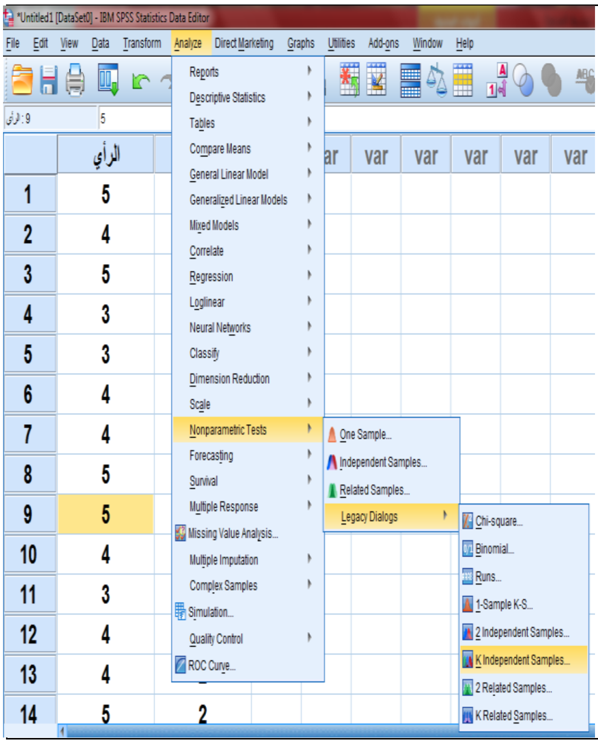
المطلوب : هل يوجد اختلاف في آراء طلبة الجامعات حول المشاركة في الانتخابات تعزى لمتغير مكان الإقامة ؟

الحل :

بعد إدخال البيانات في برنامج SPSS ، فإننا نستخدم اختبار كروسكال ولاس الرتبي ، لأن البيانات بيانات رتبية.

حسب مقياس ليكرت

الموافقة	معارض جداً	معارض	محايد	موافق	موافق جداً
الدرجة	1	2	3	4	5



رقم الحالة	محافظة شمال غزة	رقم الحالة	محافظة غزة	رقم الحالة	محافظة رفح
-1	موافق جداً	-1	محايد	-1	معارض
-2	موافق	-2	موافق	-2	محايد
-3	موافق جداً	-3	موافق	-3	موافق
-4	محايد	-4	موافق جداً	-4	موافق
-5	محايد	-5	محايد	-5	موافق
-6	موافق	-6	محايد	-6	موافق جداً
-7	موافق	-7	معارض جداً	-7	محايد
-8	موافق جداً	-8	معارض	-8	محايد
-9	موافق جداً	-9	محايد	-9	معارض جداً
-10	موافق	-10	موافق جداً	-10	معارض
		-11	موافق	-11	معارض
				-12	معارض جداً
				-13	معارض جداً

فرضية العدم : لا يوجد اختلاف بين متوسطات الرتب بين ثلاث عينات ، أي لا يوجد اختلاف في آراء الطلبة في المشاركة في الانتخابات يعزى لمتغير مكان الإقامة .

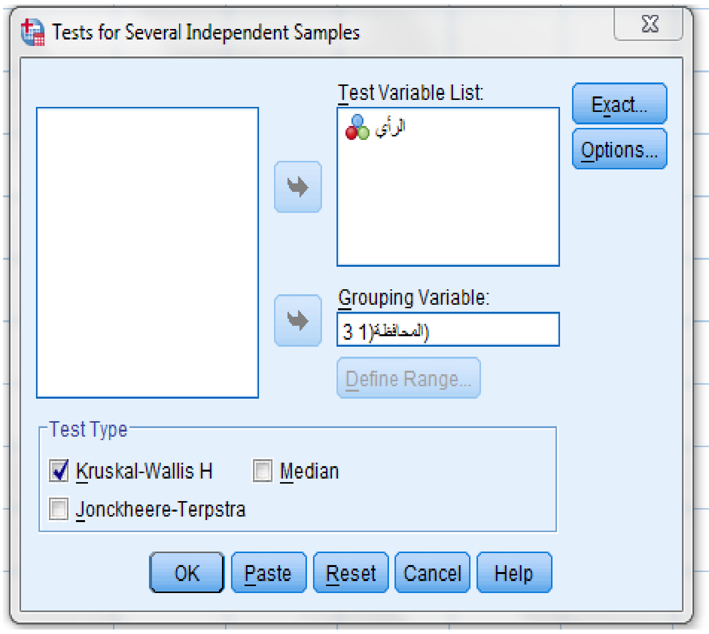
الفرضية البديلة : يوجد اختلاف بين متوسطات الرتب بين ثلاث عينات ، أي لا يوجد اختلاف في آراء الطلبة في مسيرات العودة يعزى لمتغير مكان الإقامة .

Ranks

	المحافظة	N	Mean Rank
الرأي	غزة شمال	10	24.00
	غزة	11	17.32
	رفح	13	12.65
	Total	34	

Test Statistics[a,b]	الرأي
Chi-Square	7.772
df	2
Asymp. Sg.	0.021

a. Kruskal Wallis Test



نلاحظ وجود فرق بين رتب العينات الثلاث (24 ، 17.32 ، 12.65) ، كما أن قيمة اختبار كروسكال ولاس هو (7.772) ومستوى الدلالة 0.02 ، وهي أقل من 0.05

لذلك نرفض فرضية العدم ونقبل الفرضية البديلة

النتيجة: يوجد اختلاف بين متوسطات الرتب بين ثلاث عينات، أي لا يوجد اختلاف في آراء الطلبة في المشاركة في الانتخابات يعزى لمتغير مكان الإقامة عند مستوى دلالة أقل من 0.05

ثانياً/ الاختبارات اللامعلمية للبيانات التصنيفية:

تستعمل الاختبارات اللامعلمية للبيانات التصنيفية، حيث نختبر العلاقة بين المتغيرات، أو نختبر مدى موافقة التوزيع الفعلي للبيانات مع التوزيع النظري المتوقع، ومن الاختبارات المشهورة للبيانات التصنيفية، اختبار مربع كاي واختبار فاي، واختبار ماكنمار وغيرهما.

اختبار مربع كاي

يستخدم اختبار مربع كاي في تحليل البيانات الاسمية، فالمتغيرات يجب أن تكون مصنفة ومقاسة بمقياس إسمي، وهو اختبار يستخدم للموازنة بين التوزيعات التكرارية للمتغيرات، وهو يصلح لمعالجة البيانات النوعية التي تكون على شكل تكرارات لمجموعات أو أصناف معينة. ويستخدم مربع كاي لدراسة الارتباط بين المتغيرات الاسمية أو على الأقل متغير واحد اسمي والآخر قد يكون ترتيبي أو رقمي منفصل.

1) شروط استخدام اختبار مربع كاي:

يجب أن يكون التوزيع الفعلي للتكرارت كما يلي :

1- أن تكون البيانات على شكل تكرارات و ليس نسباً مئوية أو كسوراً

2- ألا يقل مجموع التكرارات الفعلية عن 20 تكراراً و يفضل أن يزيد عددها عن 40 تكراراً

3- ألا يقل مجموع التكرارات المتوقعة في أي فئة من فئات التصنيف عن خمسة تكرارت و **إذا كان عدد الفئات خمس فئات أو أكثر، فينبغي** :

أ- ألا تقل التكرارات المتوقعة عن خمسة في 20% من تلك الفئات.

ب- وألا يزيد عدد الفئات التي يكون تكرارها واحداً على فئة واحدة.

4- الافتراض بأن جزءاً من تباين المجموعات يرجع إلى عامل الصدفة.

2) استخدامات اختبار مربع كاي: يستخدم اختبار مربع كاي في الحالات التالية :

أ- تحديد وجود علاقة (ارتباط) بين متغيرين مصنفين (ولكنه لا يقيس هذه العلاقة)

ب-لاختبار مدى تطابق (Goodness – of – fit) التوزيع المتوقع مع التوزيع الحقيقي ويستخدم في دراسة متغير مصنف واحد

3) توزين الحالات حسب تكرارات حدوثها :

عملية توزين الحالات هي عملية تستخدم عندما تكون البيانات المتوفرة ملخصة بشكل جدول تقاطعي أو بشكل جدول تكراري، ونريد إعلام SPSS بأن يتعامل معها خلال عمليات التحليل الإحصائي كأنها بيانات خام مرتبة بشكل مصفوفة مكونة من أعمدة (متغيرات) وصفوف (حالات).

أولاً / اختبار مربع كاي لمقارنة توزيع نظري مع توزيع فعلي :

مثال :

الجدول التالي يبين توزيع 200 مزرعة دواجن في منطقة ما، حسب تصنيف الأراضي للمنطقة:

فئة الأراضي	النسبة إلى مساحة المنطقة	التوزيع الفعلي O	التوزيع النظري E	(O – E)	$(O-E)^2$	$\frac{(O-E)^2}{E}$
فيضية	%10	10	20	10−	100	5
معتدلة الانحدار	%35	100	70	30	900	12.85
شديدة الانحدار	%10	2	20	18−	324	16.2
جيرية منبسطة	%25	38	50	12−	144	2.88
رملية	%20	50	40	10	100	2.5
المجموع	%100	200	200			39.43

المطلوب: هل يوجد اختلاف بين التوزيع النظري والتوزيع الفعلي (هل طبيعة الأرض أثر على اختيار مواقع المزارع ؟

الحل بالطريقة الحسابية العادية:

أولاً / فرضية العدم : لا يوجد اختلاف بين التوزيع الفعلي للمزارع والتوزيع النظري ، أي أن طبيعة الأرض لا تؤثر في اختيار موقع المزارع ، وأن الاختلاف بين التوزيعين الفعلي والنظري ناتج عن عامل الصدفة .

ثانياً / فرضية البديلة : يوجد اختلاف بين التوزيع الفعلي للمزارع والتوزيع النظري ، أي أن طبيعة الأرض تؤثر في اختيار موقع المزارع ، وأن الاختلاف بين التوزيعين الفعلي والنظري اختلاف حقيقي مرتبط بالاختلاف في نوع الأرض.

ثالثاً / قيمة اختبار كاي المحسوبة = 39.43

$$X = \sum \frac{(O-E)^2}{E} = 39.43$$

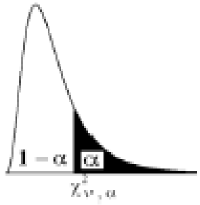
رابعاً / قيمة اختبار مربع كاي الحرجة (الجدولية) :

درجات الحرية = N – 1 = 5 – 1 = 4

مستوى الدلالة : 0.05

Percentage Points of the χ^2 Distribution; $\chi^2_{v,\alpha}$

$P(\chi^2 > \chi^2_{v,\alpha}) = \alpha$



	α														
v	0.001	0.005	0.010	0.025	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.975	0.990	0.995	0.999
1	10.83	7.88	6.63	5.02	3.84	2.71	1.32	0.45	0.10	0.02					
2	13.82	10.60	9.21	7.38	5.99	4.61	2.77	1.39	0.58	0.21	0.10	0.05	0.02	0.01	
3	16.27	12.84	11.34	9.35	7.81	6.25	4.11	2.37	1.21	0.58	0.35	0.22	0.11	0.07	0.02
4	18.47	14.86	13.28	11.14	9.49	7.78	5.39	3.36	1.92	1.06	0.71	0.48	0.30	0.21	0.09
5	20.52	16.75	15.09	12.83	11.07	9.24	6.63	4.35	2.67	1.61	1.15	0.83	0.55	0.41	0.21
6	22.46	18.55	16.81	14.45	12.59	10.64	7.84	5.35	3.45	2.20	1.64	1.24	0.87	0.68	0.38
7	24.32	20.28	18.48	16.01	14.07	12.02	9.04	6.35	4.25	2.83	2.17	1.69	1.24	0.99	0.60
8	26.12	21.95	20.09	17.53	15.51	13.36	10.22	7.34	5.07	3.49	2.73	2.18	1.65	1.34	0.86
9	27.88	23.59	21.67	19.02	16.92	14.68	11.39	8.34	5.90	4.17	3.33	2.70	2.09	1.73	1.15

القيمة الحرجة (الجدولية) = 9.49

خامساً / المقارنة: نقارن بين قيمة اختبار مربع كاي المحسوبة وقيمة مربع كاي الجدولية ، فنجد أن قيمة مربع كاي المحسوبة أكبر من قيمة مربع كاي الجدولية

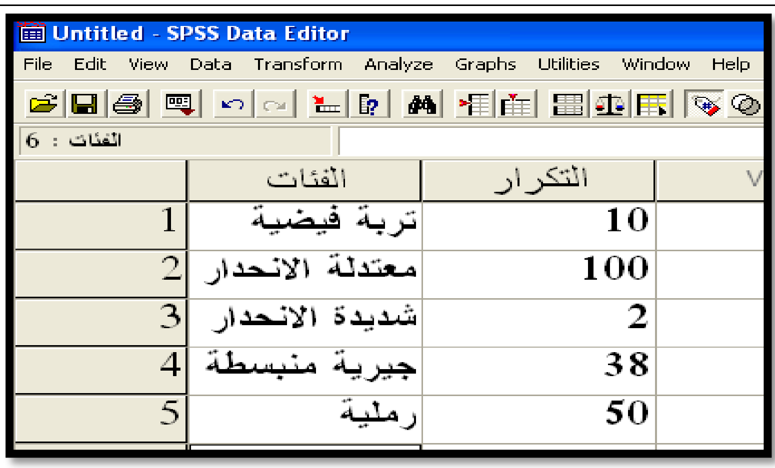
سادساً / القرار: بما أن قيمة مربع كاي المحسوبة أكبر من قيمة مربع كاي الجدولية، نرفض فرضية العدم ونقبل الفرضية البديلة .

سابعاً/ النتيجة : يوجد اختلاف بين التوزيع الفعلي للمزارع والتوزيع النظري، أي أن طبيعة الأرض تؤثر في اختيار موقع المزارع، وأن الاختلاف بين التوزيعين الفعلي والنظري اختلاف حقيقي مرتبط بالاختلاف في نوع الأرض، عند مستوى دلالة 0.05

ثانياً / الحل بالحاسوب :

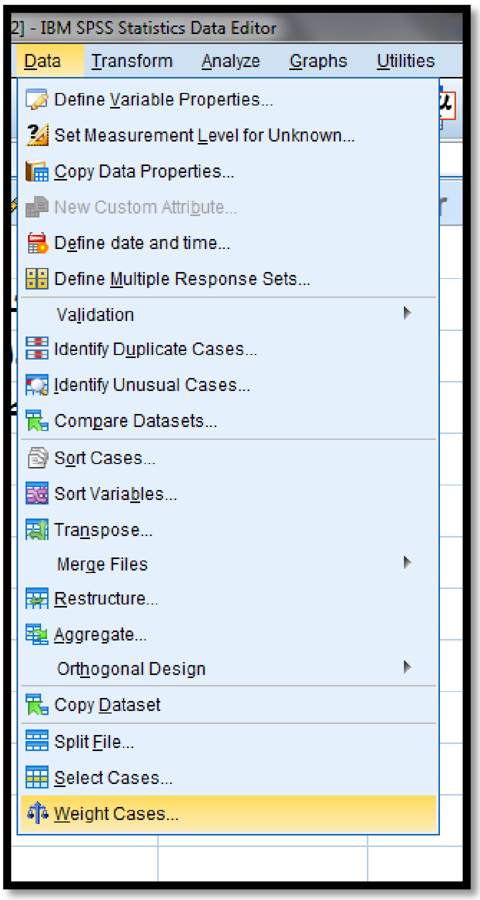
الحل : إدخال البيانات في برنامج SPSS بالشكل التالي

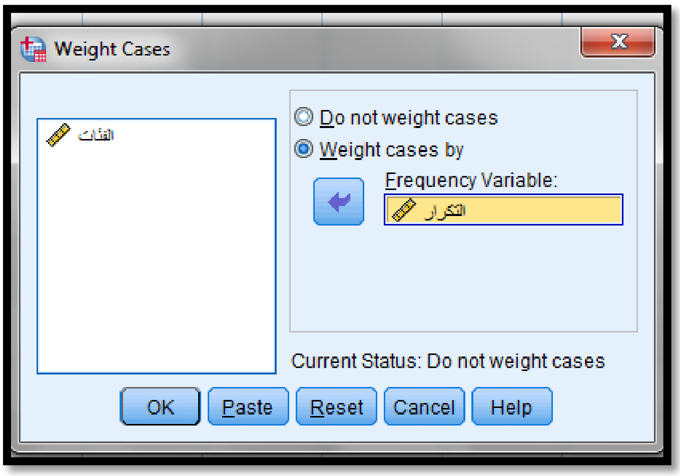
فئة الأراضي		التكرار
فيضية	1	10
معتدلة الانحدار	2	100
شديدة الانحدار	3	2
جيرية منبسطة	4	38
رملية	5	50



	الفئات	التكرار
1	تربة فيضية	10
2	معتدلة الانحدار	100
3	شديدة الانحدار	2
4	جيرية منبسطة	38
5	رملية	50

1–ثم نقوم بتوزين البيانات :





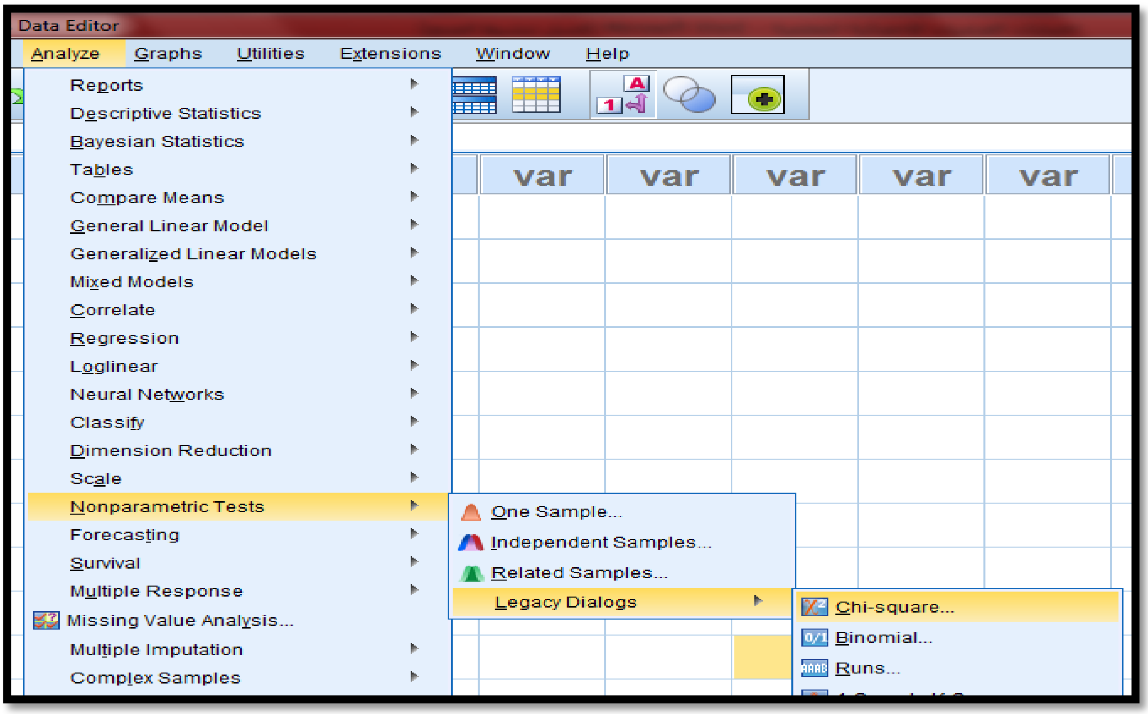
نفرض فرضيات العدم و البديلة :

فرضية العدم : لا يوجد اختلاف بين التوزيع الفعلي والتوزيع النظري

لا تؤثر طبيعة الأرض على اختيار موقع المزرعة

الفرضية البديلة : يوجد اختلاف بين التوزيع الفعلي والتوزيع النظري

يوجد تأثير لطبيعة الأرض على اختيار موقع المزرعة



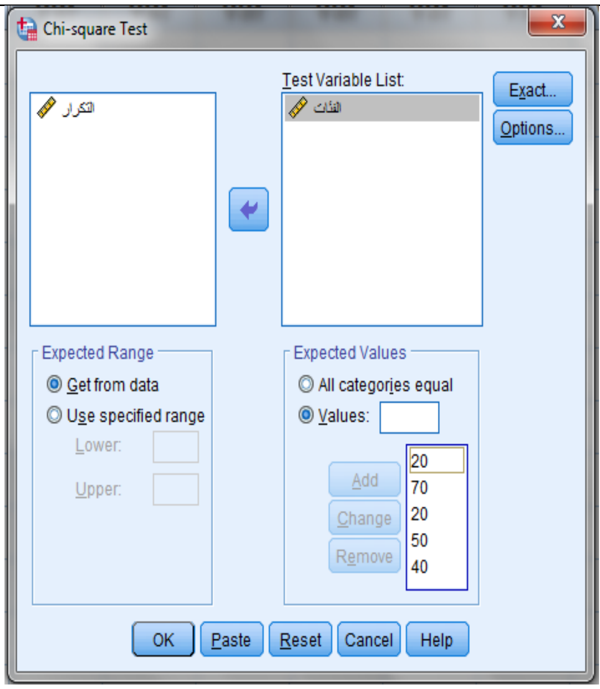
الفئات

	Observed N	Expected N	Residual
تربة فيضية	10	20.0	-10.0-
معتدلة الانحدار	100	70.0	30.0
شديدة الانحدار	2	20.0	-18.0-
جيرية منبسطة	38	50.0	-12.0-
رملية	50	40.0	10.0
Total	200		

Test Statistics

	الفئات
Chi-Square	**39.437[a]**
df	**4**
Asymp. Sig.	**0.000**

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 20.0.



النتيجة :

بما أن قيمة مربع كاي المحسوبة = 39.5 ومستوى الدلالة 0.000 و هو أقل من 0.05

النتيجة : نرفض فرضية العدم و نقبل الفرضية البديلة

أي أن طبيعة الأرض لها تأثير على إختيار موقع المزرعة

مثال 2 : في دراسة حول أثر مستوى الدخل في اصابة السكان بمرض السكر

فئة الدخل	**أقل من 500**	**500 - 1000**	**1000 – 1500**	**أكثر من 1500**
عدد حالات الاصابة	50	100	150	200

المطلوب : **هل توجد علاقة بين عدد حالات الاصابة بمرض السكر ومستوى الدخل ؟**

الفئات	**التوزيع الفعلي O**	**التوزيع النظري E**	**(O – E)**	$(O-E)^2$	$\frac{(O-E)^2}{E}$
أقل من 500	50	125	75 –	5625	45
500 – 1000	100	125	25 –	625	5
1000 – 1500	150	125	25	625	5
أكثر من 1500	200	125	75	5625	45
	500	500			100

1- فرضية العدم : لا توجد علاقة بين عدد حالات الإصابة بمرض السكر ومستوى الدخل

2- الفرضية البديلة : توجد علاقة بين عدد حالات الإصابة بمرض السكر ومستوى الدخل.

3- **قيمة اختبار مربع كاي:** كما في الجدول = 100

4- **قيمة اختبار مربع كاي الجدولية** (القيمة الحرجة) :

أ- درجات الحرية = N -1 = 4 -1 = 3

ب- مستوى الدلالة = 0.05

ت- قيمة اختبار مربع كاي الجدولية عند درجة حرية 3 ومستوى دلالة 0.05 = **7.815**

5- **المقارنة** : نقارن بين قيمة اختبار مربع كاي المحسوبة وقيمة اختبار مربع كاي الجدولية، فنجد أن قيمة اختبار مربع كاي المحسوبة أكبر من قيمة اختبار مربع كاي الجدولية.

6- **القرار:** بما أن قيمة اختبار مربع كاي المحسوبة أكبر من قيمة اختبار مربع كاي الجدولية، فإننا نرفض فرضية العدم ونقبل الفرضية البديلة

7- **النتيجة:** توجد علاقة ذات دلالة إحصائية بين الإصابة بمرض السكر ومستوى الدخل

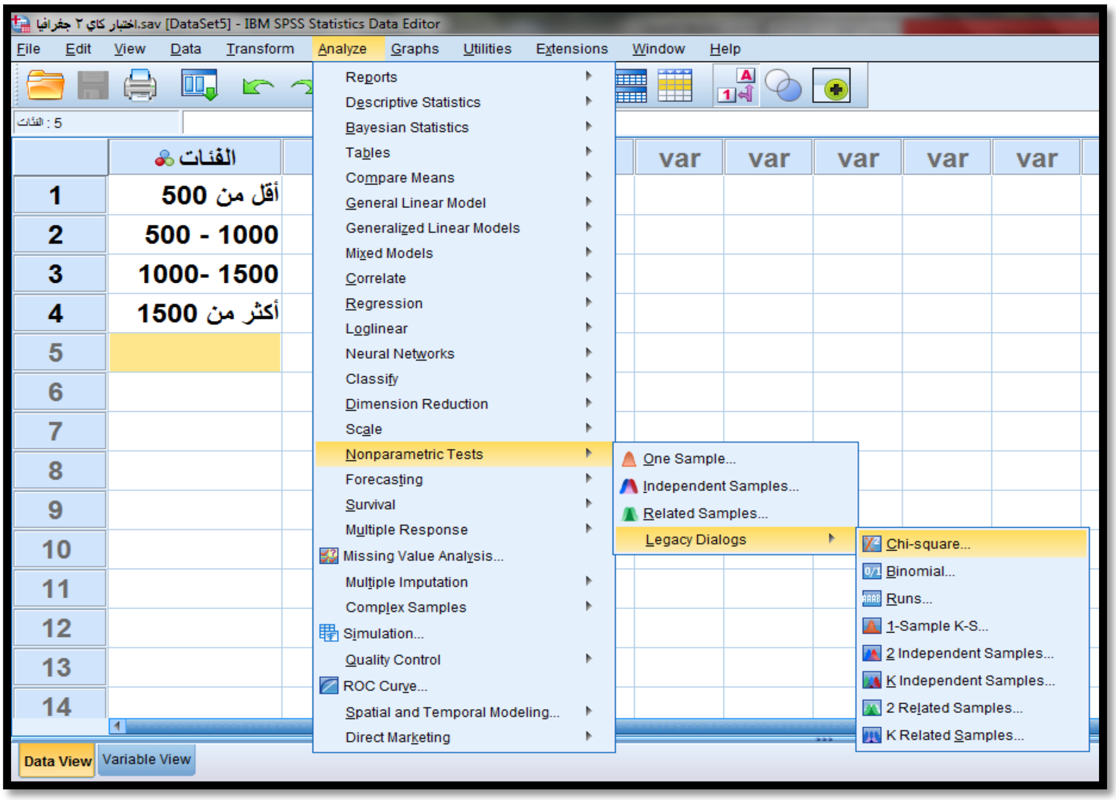
الحل بالحاسوب :

1) الفرضيات :

فرضية العدم : لا توجد علاقة بين عدد حالات الإصابة بمرض السكر ومستوى الدخل

الفرضية البديلة : توجد علاقة بين عدد حالات الإصابة بمرض السكر ومستوى الدخل

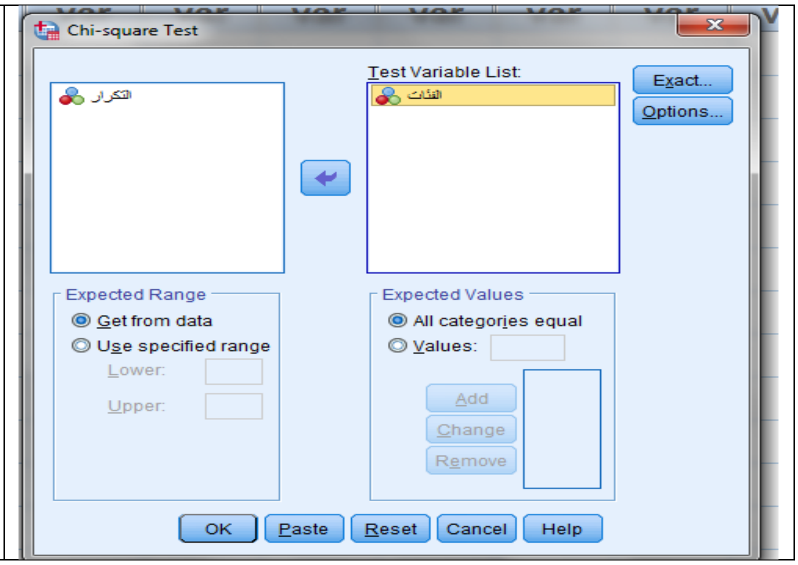
2) نقوم بإدخال البيانات ثم توزينها weight cases



نختار في Expected Value

الاختيار All categories equal

وهنا يكون توزيع القيم المتوقعة بالتساوي وهي = 125 بعد قسمة المجموع على عددها 4



النتيجة :

Test Statistics

	الفئات
Chi-Square	100.000[a]
df	3
Asymp. Sig.	.000

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 125.0.

الفئات	Observed N	Expected N	Residual
500 من أقل	50	125.0	-75.0-
500 - 1000	100	125.0	-25.0-
1000- 1500	150	125.0	25.0
1500 من أكثر	200	125.0	75.0
Total	500		

نجد أن قيمة اختبار مربع كاي = 100 ومستوى الدلالة = 0.000 وهو أقل من 0.01 لذلك نرفض فرضية العدم و نقبل الفرضية البديلة

النتيجة : (توجد علاقة بين عدد حالات الإصابة بمرض السكر و مستوى الدخل)

ثانياً / اختبار مربع كاي للجداول المتقاطعة :

مثال1 : يوضح الجدول التالي استخدام المزارعين لنوع جديد من الأشتال في الزراعة من خلال استبيان وزع عليهم بالشكل التالي :

مستوى التعليم	استعمال الاشتال المحسنة	
	يستعمل (نعم)	لا يستعمل (لا)
أمي	20	30
متعلم	40	10
المجموع	60	40

المطلوب : هل المستوي التعليمي له علاقة باستعمال المزارعين للأشتال المحسنة ؟

صيغة أخرى : هل استعمال المزارعين للأشتال المحسنة له علاقة بمتغير التعليم ؟

خطوات الحل :

1) نعمل جدول التوزيع النظري بالشكل التالي :

التوزيع النظري المتوقع E

مستوى التعليم	استعمال الاشتال المحسنة		المجموع
	يستعمل (نعم)	لا يستعمل (لا)	
أمي	30 = 100/(50×60)	20 = 100/(50×40)	50
متعلم	30 = 100/(50×60)	20 = 100/(50×40)	50
المجموع	60	40	100

2) جدول الحل :

مستوى التعليم	استعمال الاشتال المحسنة	العدد(التكرار) O	التوزيع النظري E	(O – E)	$(O-E)^2$	$\frac{(O-E)^2}{E}$
أمي	نعم	**20**	**30**	**10–**	**100**	**3.33**
أمي	لا	**30**	**20**	**10**	**100**	**5**
متعلم	نعم	**40**	**30**	**10**	**100**	**3.33**
متعلم	لا	**10**	**20**	**10–**	**100**	**5**
المجموع		**100**	**100**			**16.66**

3) الفرضيات

أ- **فرضية العدم**: لا توجد علاقة ذات دلالة احصائية بين متغير التعليم ومتغير استعمال الأشتال المحسنة.

ب- **الفرضية البديلة** : توجد علاقة ذات دلالة إحصائية بين متغير التعليم ومتغير استعمال الأشتال المحسنة (أي أن المستوى التعليمي له علاقة باستعمال المزارعين للأشتال المحسنة).

<u>بما أن البيانات بيانات اسمية فإننا سنستخدم اختبار مربع كاي لإثبات العلاقة السابقة</u>

4- قيمة اختبار مربع كاي المحسوبة = 16.66

5- قيمة اختبار مربع كاي الجدولية (القيمة الحرجة) :

أ- درجات الحرية = N (عدد الفئات) – 1 = 4 – 1 = 3

ب-مستوى الدلالة = 0.05

ج- قيمة اختبار مربع كاي الجدولية عند درجة حرية 3 ومستوى دلالة 0.05 = **7.815**

6- **المقارنة** : نقارن بين قيمة اختبار مربع كاي المحسوبة وقيمة اختبار مربع كاي الجدولية، فنجد أن قيمة اختبار مربع كاي المحسوبة أكبر من قيمة اختبار مربع كاي الجدولية

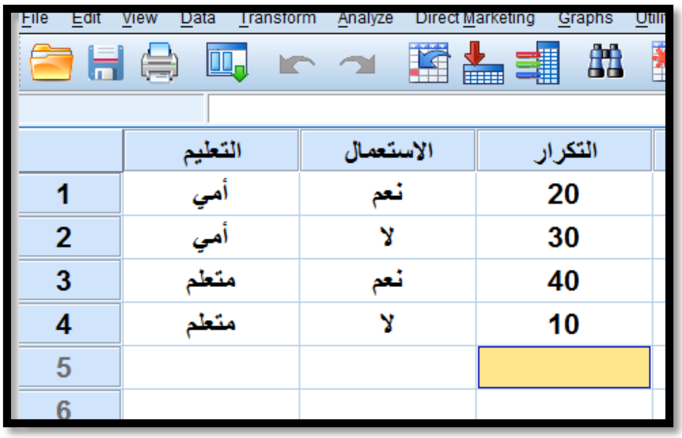
7- **القرار**: بما أن قيمة اختبار مربع كاي المحسوبة أكبر من قيمة اختبار مربع كاي الجدولية، فإننا نرفض فرضية العدم ونقبل الفرضية البديلة

8- النتيجة: توجد علاقة ذات دلالة إحصائية بين المستوى التعليمي و استعمال الأشتال المحسنة في الزراعة و له دلالة إحصائية عند مستوى دلالة 0.05

ثانياً / المعالجة الإحصائية بالحاسوب :

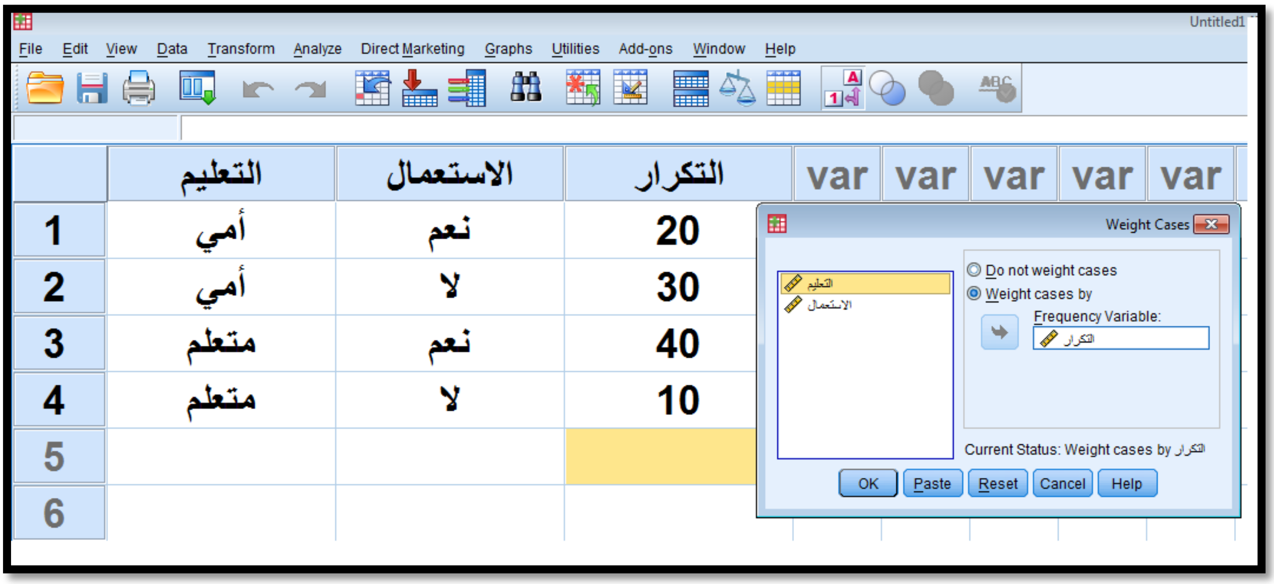
حتى تتم معالجة البيانات السابقة في برنامج SPSS يجب ادخالها بالشكل التالي :

مستوى التعليم	استعمال الاشتال المحسنة	العدد (التكرار)
أمي (1)	نعم (1)	20
أمي (1)	لا (2)	30
متعلم (2)	نعم (1)	40
متعلم (2)	لا (2)	10



يعطى أمي = 1 متعلم = 2 نعم = 1 لا = 2

يجب توزين البيانات Weight Cases :

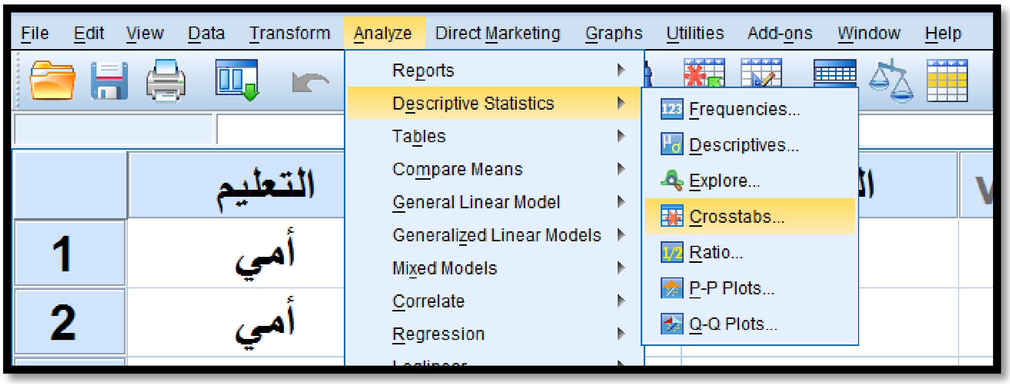


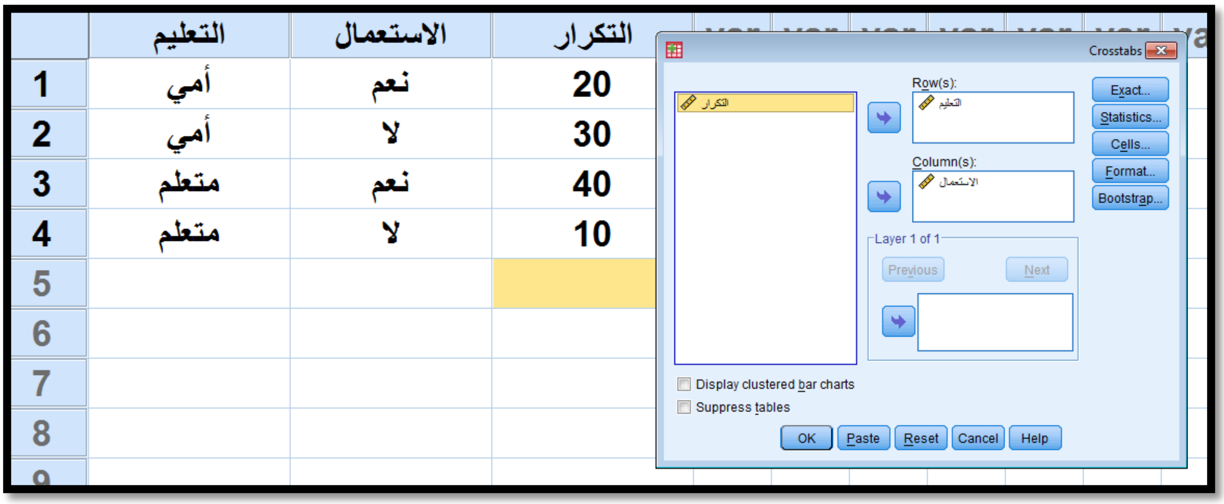
بعد توزين الحالات (التكرار)

فرضية العدم: لا توجد علاقة ذات دلالة احصائية بين متغير التعليم ومتغير استعمال الأشتال المحسنة

الفرضية البديلة : توجد علاقة ذات دلالة إحصائية بين متغير التعليم ومتغير استعمال الأشتال المحسنة (أي أن المستوى التعليمي له علاقة باستعمال المزارعين للأشتال المحسنة)

<u>استخدام اختبار مربع كاي لاثبات العلاقة السابقة</u>





النتيجة :

الأشتال * التعليم Crosstabulation

			الأشتال		
			نعم	لا	Total
التعليم	أمي	Count	20	30	50
		Expected Count	30.0	20.0	50.0
	متعلم	Count	40	10	50
		Expected Count	30.0	20.0	50.0
Total		Count	60	40	100
		Expected Count	60.0	40.0	100.0

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	**16.667**[a]	1	0.000		
Continuity Correction[b]	15.042	1	.000		
Likelihood Ratio	17.261	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	16.500	1	.000		
N of Valid Cases	100				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 20.
b. Computed only for a 2x2 table

النتيجة : نجد أن قيمة اختبار مربع كاي = 16.7 و مستوى الدلالة = 0.000 و هي أقل من 0.05 **لذلك نرفض فرضية العدم و نقبل الفرضية البديلة أي توجد علاقة بين المستوى التعليمي و استعمال الأشتال المحسنة في الزراعة و له دلالة إحصائية.**

مثال 2 :

في دراسة في الجغرافيا السياسية حول الانتخابات الفلسطينية، جاءت إجابات العينة المأخوذة من المجتمع الفلسطيني في قطاع غزة، حول مشاركتهم في الانتخابات المقبلة كالتالي :

مكان السكن	مدن	ريف	مخيم	المجموع
أجاب (نعم)	100	30	150	**280**
أجاب (لا)	60	20	155	**235**
أجاب (متردد)	40	50	95	**185**
المجموع	**200**	**100**	**400**	**700**

المطلوب :هل يوجد اختلاف في أراء الناخبين حول مشاركتهم في الانتخابات الفلسطينية القادمة تعزى لمكان السكن.

أولاً / الحل بالطريقة اليدوية

خطوات الحل :

1) جدول التوزيع النظري

مكان السكن	مدن	ريف	مخيم	المجموع
أجاب (نعم)	80	40	160	**280**
أجاب (لا)	67.1	33.6	134.3	**235**
أجاب (متردد)	52.9	26.4	105.7	**185**
المجموع	**200**	**100**	**400**	**700**

2) عمل جدول اختبار مربع كاي

مكان السكن	**المشاركة في الانتخابات**	**التوزيع الفعلي "O"**	**التوزيع النظري "E "**	**(O – E)**	$(O-E)^2$	$\frac{(O-E)^2}{E}$
مدن (1)	نعم (1)	100	80	20	400	5
مدن (1)	لا (2)	60	67.1	7.1 −	50.41	0.75
مدن (1)	متردد (3)	40	52.9	12.9 −	166.41	3.15
ريف (2)	نعم (1)	30	40	10 −	100	2.5
ريف (2)	لا (2)	20	33.6	13.6 −	184.96	5.5
ريف (2)	متردد (3)	50	26.4	23.6	556.96	21.1
مخيمات (3)	نعم (1)	150	160	10 −	1000	0.62
مخيمات (3)	لا (2)	155	134.3	20.4	428.5	3.19
مخيمات (3)	متردد (3)	95	105.7	10.7 −	114.5	1.1
المجموع		700	700	/////////////////		**42.91**

3) الفرضيات

أ- **فرضية العدم**:لا توجد علاقة ذات دلالة احصائية بين متغير السكن ومتغير المشاركة في الانتخابات

ب- **الفرضية البديلة:** توجد علاقة ذات دلالة إحصائية بين متغير السكن ومتغير المشاركة في الانتخابات(أي أن المسكن له علاقة بالمشاركة في الانتخابات) أي أنه يوجد اختلاف في قرار الأفراد في المشاركة في الانتخابات حسب نوع المسكن .

<u>بما أن البيانات بيانات اسمية فإننا سنستخدم اختبار مربع كاي لإثبات العلاقة السابقة</u>

4- قيمة اختبار مربع كاي المحسوبة = **42.91**

5- قيمة اختبار مربع كاي الجدولية (القيمة الحرجة) :

أ- درجات الحرية = N (عدد الفئات) – 1 = 9 – 1 = 8

ب- مستوى الدلالة = 0.05

ج- قيمة اختبار مربع كاي الجدولية عند درجة حرية 8 ومستوى دلالة 0.05 = **15.51**

6- **المقارنة** : نقارن بين قيمة اختبار مربع كاي المحسوبة وقيمة اختبار مربع كاي الجدولية، فنجد أن قيمة اختبار مربع كاي المحسوبة أكبر من قيمة اختبار مربع كاي الجدولية

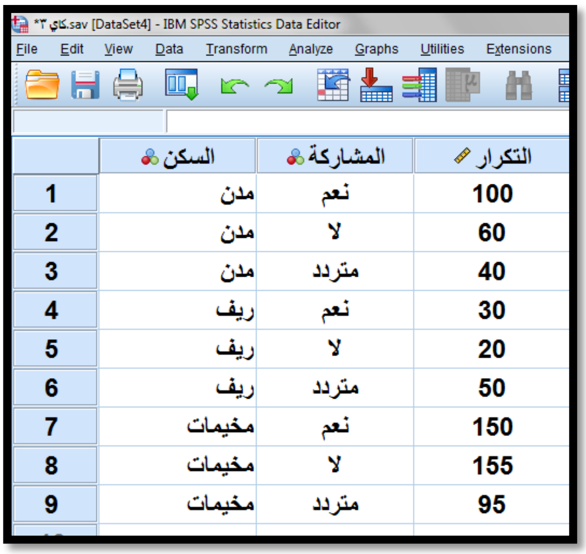
7- **القرار**: بما أن قيمة اختبار مربع كاي المحسوبة أكبر من قيمة اختبار مربع كاي الجدولية، فإننا نرفض فرضية العدم ونقبل الفرضية البديلة

8- **النتيجة:** توجد علاقة ذات دلالة إحصائية بين متغير السكن ومتغير المشاركة في الانتخابات(أي أن المسكن له علاقة بالمشاركة في الانتخابات) أي أنه يوجد اختلاف في قرار الأفراد في المشاركة في الانتخابات حسب نوع المسكن .

ثانياً / المعالجة الإحصائية بالحاسوب :

حتى تتم معالجة البيانات السابقة في برنامج SPSS يجب ادخالها بالشكل التالي :

مكان السكن	المشاركة في الانتخابات	العدد (التكرار)
مدن (1)	نعم (1)	100
مدن (1)	لا (2)	60
مدن (1)	متردد (3)	40
ريف (2)	نعم (1)	30
ريف (2)	لا (2)	20
ريف (2)	متردد (3)	50
مخيمات (3)	نعم (1)	150
مخيمات (3)	لا (2)	155
مخيمات (3)	متردد (3)	95



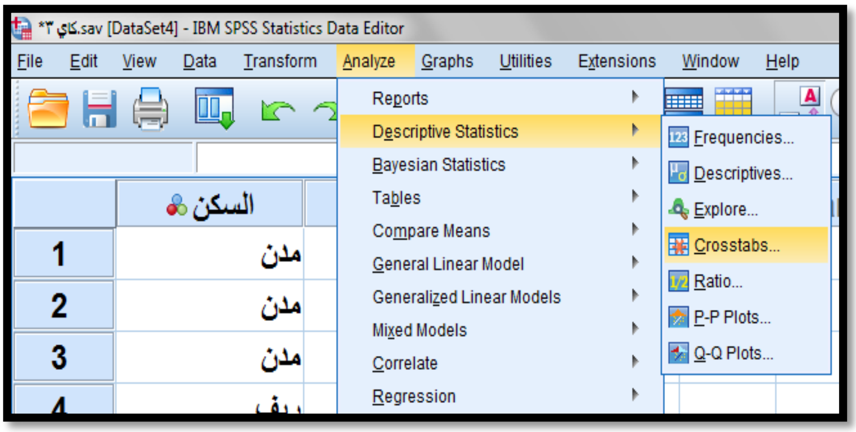
	السكن	المشاركة	التكرار
1	مدن	نعم	100
2	مدن	لا	60
3	مدن	متردد	40
4	ريف	نعم	30
5	ريف	لا	20
6	ريف	متردد	50
7	مخيمات	نعم	150
8	مخيمات	لا	155
9	مخيمات	متردد	95

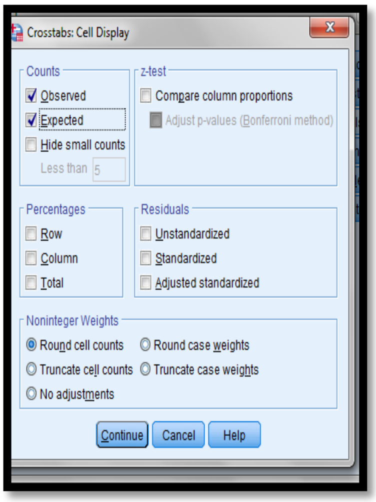
فرضية العدم : لا توجد علاقة ذات دلالة احصائية بين متغير السكن ومتغير المشاركة في الانتخابات

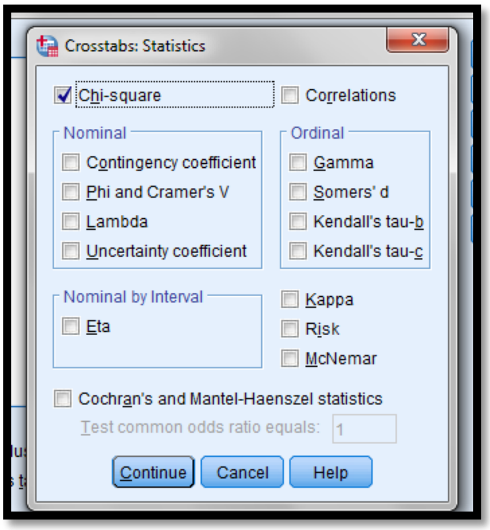
الفرضية البديلة: توجد علاقة ذات دلالة إحصائية بين متغير السكن ومتغير المشاركة في الانتخابات(أي أن المسكن له علاقة بالمشاركة في الانتخابات) أي أنه يوجد اختلاف في قرار الأفراد في المشاركة في الانتخابات حسب نوع المسكن .

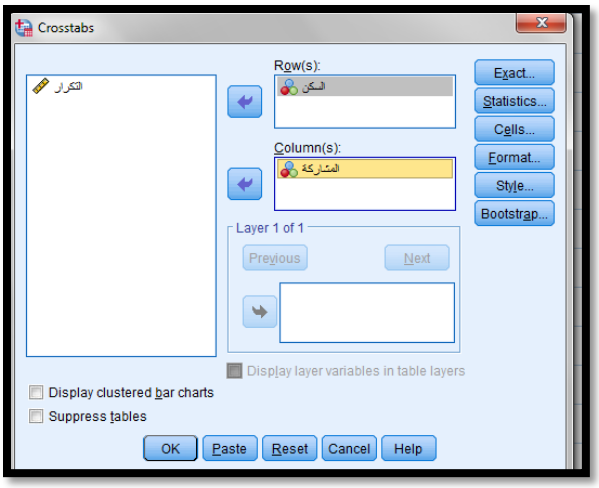
الاختبار المستخدم : استخدام اختبار مربع كاي لإثبات العلاقة السابقة :

بعد إدخال البيانات ، نقوم بتوزينها من DATA ثم Weight









المشاركة * السكن Crosstabulation

			المشاركة			
			نعم	لا	مترددة	Total
السكن	مدن	Count	100	60	40	200
		Expected Count	80.0	67.1	52.9	200.0
	ريف	Count	30	20	50	100
		Expected Count	40.0	33.6	26.4	100.0
	مخيمات	Count	150	155	95	400
		Expected Count	160.0	134.3	105.7	400.0
Total		Count	280	235	185	700
		Expected Count	280.0	235.0	185.0	700.0

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	**42.803**[a]	4	**0.000**
Likelihood Ratio	39.419	4	.000
Linear-by-Linear Association	2.943	1	.086
N of Valid Cases	700		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 26.43.

قيمة اختبار مربع كاي = 42.8 و مستوى الدلالة = 0.000

بما أن قيمة اختبار كاي أقل من 0.01 لذا نرفض فرضية العدم و نقبل الفرضية البديلة

النتيجة : توجد علاقة ذات دلالة إحصائية بين متغير السكن و متغير المشاركة في الانتخابات (أي أن المسكن له علاقة بالمشاركة في الانتخابات) أي أنه يوجد اختلاف في قرار الأفراد في المشاركة في الانتخابات حسب نوع المسكن.

انتهت المحاضرات لاختبارات الفرضيات