الاختبارات الاحصائية لعينة وعينتين

- تعتبر اختبارات الفروض الاحصائية Statistical Testing Hypotheses واحدة من أهم التطبيقات التى قدمها علم الاحصاء كحل للمشاكل العلمية المختلفة بشتى فروع العلم. فباستخدام نظرية الاحتمالات وخصائص التوزيعات العينية امكن التعرف على ما يسمى باختبارات الفروض الاحصائية ومن خلالها يمكن لأى شخص ان يتخذ قرار برفض أو قبول فرض معين أو مجموعة من الفروض المتعلقة بمشكلة معينة موجوده فى الحياه العامه.
- وبصفه عامه فان قبول او رفض اى قرار يجب ان يمر بعدة مراحل وهى:

1. سحب عينة من المجتمع بحيث تكون ممثله احسن تمثيل للمجتمع
2. تجميع البيانات المتعلقة بالمشكلة من العينة
3. تطبيق قواعد معينه لاختبار الفروض الموضوعه عن طريق الباحث وهى مشكله تحتاج لخبره احصائية
4. اتخاذ القرار بناء على ما توصل اليه الباحث من نتائج.

خطوات اجراء الاختبار الاحصائى

- الاختبار الاحصائى قد يكون متعلقا بعينة واحدة او عينتين او اكثر وقد يكون اختبارا معلميا او غير معلميا ويجب ان يمر الاختبار اى كان نوعه بعدة خطوات

1. يبدأ الاختبار بتفهم اهداف البحث ثم اعادة صياغه هذه الأهداف فى فرضين احداهما يسمى فرض العدم Null Hypothesis والأخر يسمى الفرض البديل Alternative Hypothesis
2. يحدد احتمال الخطأ (الخطر) ويمثل الخطأ من النوع الأول (احتمال رفض فرض العدم وهو صحيح) Type I Error ويرمز له بالرمز α وغالبا يساوى 0.05 أو 0.01
3. يتم اجراء التجربه وتجميع البيانات من العينه
4. تراجع الفروض Assumptions اللازمه للاختبار وهى تختلف عن الفروض الاحصائية(فرض العدم والفرض البديل).
5. استخدام نظرية الاحتمالات نبحث عن متغير عشوائى يسمى احصائى الاختبار Test Statistic وعن التوزيع العينى Sampling Distribution الملائم لهذا المتغير
6. تستخدم المعلومات المتاحه من العينة والمجتمع والتوزيع العينى لاتخاذ قرار معين اما بقبول او رفض فرض العدم.

اختبار ت T Test

- احصاء الاختبار فى هذه الحاله له التوزيع الاحتمالى والمسمى توزيع T (T Distribution) لذا يسمى هذا الاختبار باختبار ت T ويستخدم اختبار ت عندما يكون المجتمع (او المجتمعات) لها التوزيع الطبيعى والتباين (التباينات) غير معلوم وايضا حجم العينه (العينات) أقل من 30 لكن اذا زاد حجم العينة عن 30 سوف يكون لأحصاء الاختبار التوزيع الطبيعى المعيارى وهو اختبار معلمى.
- واختبار T يتوقف على عدة شروط يجب توافرها قبل اجراء الاختبار وهى

1. يجب ان تكون وحدة القياس المقاسه بها البيانات بفتره على الأقل وهذا يعنى ان البيانات يجب ان تكون لمتغيرات مستمره. فاذا كانت وحدة القياس اسميه او ترتيبيه فلا يطبق الاختبار
2. العينة يجب ان تكون مختارة عشوائيا من المجتمع
3. بيانات العينة او العينات الداخله فى الاختبار يجب ان تتوزع حسب التوزيع الطبيعى

الشرط الأول والثانى يجب ان يتحقق الباحث منهما لكن الشرط الثالث يمكن التحقق منه عن طريق الأمر Frequencies أو الأمر Explore كما تم سابقا

- وسوف نهتم هنا بالشرح والتوضيح لاستخدامات اختبارات T والذى يستخدم فى عدد من الاختبارات منها:

1. هل متوسط عينه واحده يساوى متوسط مجتمع.
2. اختبار متوسط عينتين مستقلتين Independent Samples ويكون المطلوب اختبار هل العينتين مسحوبتان من نفس المجتمع ام لا؟
3. اختبار الفرق بين متوسطى عينتين غير مستقلتين.

الاختبارات الاحصائية لعينة واحدة
One Sample Test

يستخدم هذا الاختبار في مقارنة متوسط عينة بقيمة مفترضة للمجتمع.

على سبيل المثال:

مقارنة متوسط تحصيل الطلاب بإحدى المدارس في مادة الجغرافيا بمتوسط تحصيل الطلاب العام في مدارس المحافظة ككل.

ويكون إجابة على السؤال: هل يختلف متوسط العينة بالمدرسة عن المتوسط العام بالمحافظة؟

ويمكن إجراء الاختبار من خلال الخطوات التالية:

- **صياغة الفرض الصفري والفرض البديل.**
- **تحديد أعلى نسبة خطأ يسمح الباحث بها (مستوى الدلالة @) .**
- **جمع المعلومات والبيانات من خلال أدوات القياس المختلفة.**
- **إجراء الاختبار على البيانات.**
- **اتخاذ القرار في ضوء النتائج.**

وهناك مجموعة من الشروط لاستخدام اختبار ت لعينة واحدة وهي:

- **أن يكون المتغير التابع مقاسا على المستوى الكمي .**
- **أن يتبع المتغير التابع التوزيع الاعتدالي.**
- **استقلالية المشاهدات.**
- **العينة مختارة عشوائيا.**

- بفرض ان لدينا عينة من 20 شخص وتم قياس وزن كل شخص بالكليوجرام وكانت لدينا البيانات التالية:

65 75 80 120 110 90 60 100 90 70 165 100 90 80 70 75 85 55 130 100

بفرض اننا نريد اختبار الفرض العدمى: متوسط الوزن فى المجتمع المسحوب منه العينه يساوى 100كجم

الفرض البديل: متوسط المجتمع لا يساوى 100كجم

الحل:

اولا: بوضع هذه البيانات ى ملف وليكن weight.sav

ثانيا: يجب اختبار هل هذه البيانات تتبع التوزيع الطبيعى ام لا وذلك باستخدام الأمر Frequencies أو Explore كما سبق

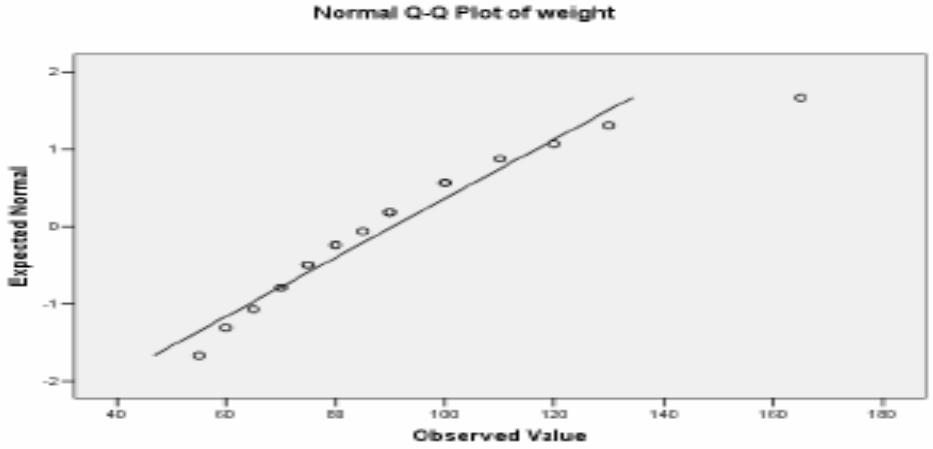
Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
weight	.158	20	.200*	.913	20	.072

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

a. Lilliefors Significance Correction

من اختبار Test of Normality نجد ان قيمة Sig. اكبر من 0.05 لذا سوف نقبل فرض العدم وهو ان البيانات لها التوزيع الطبيعى



من الرسم البيانى نجد ان البيانات تتركز حول الخط المستقيم لذا فانها تتوزع تبعا للتوزيع الطبيعى

ثالثا: بما ان حجم العينه 20 اى اقل من 30 لذا فان احصاء الاختبار هو T بمعنى اننا سوف نستخدم اختبار T وسنجرى الاختبار كالتالى

1- الفروض الاحصائية

فرض العدم Null Hypothesis سيكون له الشكل التالى:

$$H_0 : \mu = 100$$

الفرض البديل Alternative Hypothesis سيكون له الشكل التالى:

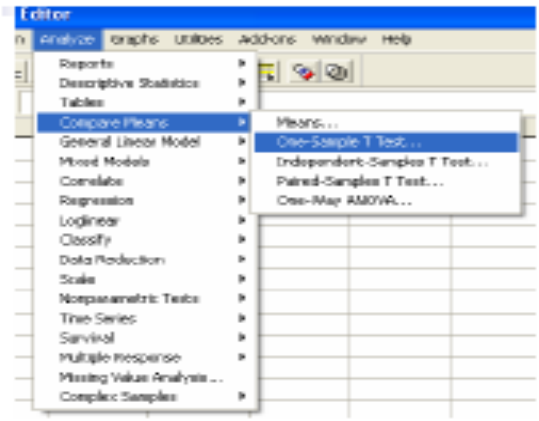
$$H_0 : \mu \neq 100$$

احصاء الاختبار هو

$$T = \frac{\overline{X} - \mu_0}{S / \sqrt{n}}$$

وبتعين الوسط الحسابى والانحراف المعيارى للعينه يمكن حساب قيمة احصاء الاختبار T

ومن جدول توزيع T يمكن حساب القيمة الحرجة $t_{\alpha/2} = t(n-1;\alpha/2)$ فاذا كانت |T| أكبر من القيمة الحرجة نرفض فرض العدم ونقبل الفرض البديل لكن اذا كانت أقل سوف نقبل فرض العدم ونرفض الفرض البديل



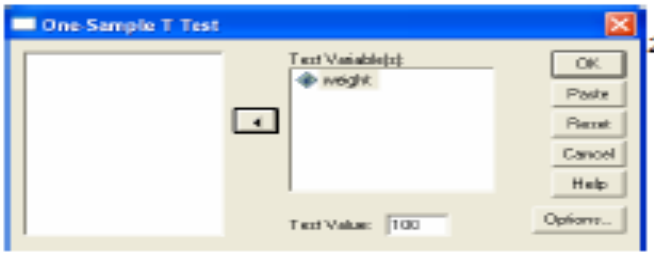
ويمكن اجراء هذا الاختبار عن طريق حزمة SPSS بسهوله كما يلى:

1- بتخزين البيانات فى ملف Weight.sav

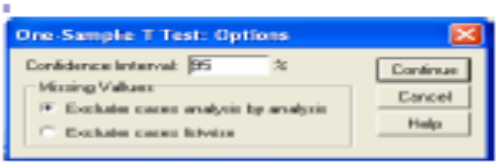
2- من قائمة Analyze نختار Compare Means

3- من القائمة الفرعيه نختار One Sample T Test

4- تظهر شاشة جديده بعنوان One-Sample T test ننقل المتغير Weight لقائمة Test Variable(s)



5- فى المستطيل Test Value: نكتب قيمة فرض العدم وهى 100



6- نختار الأمر Options تظهر شاشة جديده بعنوان One- Sample T Test: Options لاحظ انه فى خانة Confidence interval القيمة 95% بمعنى ان مستوى المعنويه α=0.05 ويمكن تغيرها ثم نختار Continue نعود للشاشة السابقة ثم نضغط Ok

7- تظهر النتائج التالية

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
weight	20	90.5000	26.20295	5.85909

One-Sample Test

	Test Value = 100					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower	Upper
weight	-1.621	19	.121	-9.50000	-21.7632	2.7632

الجدول الأول:

بعنوان One-Sample Statistics وعرض عدد حالات المتغير والوسط الحسابي والانحراف المعيار وخطأ التقدير للوسط الحسابي.

الجدول الثاني:

بعنوان One-Sample Test وعرض نتيجه الاختبار حيث يحتوى الجدول على

1. فرض العدم اعلى الجدول Test Value = 100
2. واسم المتغير
3. قيمه احصاء الاختبار T= -1.621
4. درجة الحرية n-1 = 19
5. القيمة الاحتمالية P-value وهى Sig. = 0.121
6. الفرق بين وسط العينة وقيمة الفرض العدمى Mean Difference
7. 95% فترة ثقة لوسط المجتمع المسحوب منه العينة

- من الجدول الثانى يمكن اتخاذ قرار بناء على قيمة Sig.=0.121 حيث ان الاختبار ذو طرفين فاننا سوف نقارنها بالقيمة α/2 =0.025 وهى اكبر لذا سوف نقبل فرض العدم.
- يمكن للمستخدم ان يستخدم القيمة الموجبه لأحصاء الاختبار |T| ومقارنتها بالقيمة الحرجة والتى نحصل عليها من جدول T عند مستوى معنويه α/2 =0.025 ودرجة حرية n-1 = 19
- ومن الملاحظ ان

$$t_{\alpha/2} = t(0.025,19) = 2.262, \qquad |T| = 1.621$$

- وبالتالى فان

$$|T| < t_{\alpha/2}$$

- لذا سوف نقبل فرض العدم وهو ان وسط المجتمع المسحوب منه العينة يساوى 100كجم
- ويفضل دائما عند التعامل مع الحزمة ان نستخدم Sig. لاتخاذ القراروليست احصاء الاختبار

اختبارعينتين مستقلتين

يستخدم هذا الاختبار في مقارنة متوسط عينتين مستقلتين (أي أن الأشخاص في المجموعة 1 ليسوا نفس الأشخاص في المجموعة 2).

- بفرض ان لدينا عينتين مستقلتين ونهتم بمتغير معين في كلا العينتين ونرغب في اختبار ان متوسطي المجتمعين المسحوب منهما العينتين لهما نفس الوسط الحسابي ام لا لذا سوف تصاغ الفروض الاحصائية كالتالي:

1. اختبار ذو طرفين

$$H_0 : \mu_1 = \mu_2 \ , \quad H_1 : \mu_1 \neq \mu_2$$

2. اختبار ذو طرف واحد

$$H_0 : \mu_1 = \mu_2 \ , \quad H_1 : \mu_1 <(>) \ \mu_2$$
$$\text{or} \quad H_0 : \mu_1 < \mu_2 \ , \quad H_1 : \mu_1 \geq \mu_2$$
$$\text{or} \quad H_0 : \mu_1 > \mu_2 \ , \quad H_1 : \mu_1 \leq \mu_2$$

ولاجراء هذا الاختبار يجب توافر بعض الشروط

الشروط اللازم توافرها

1. حجم العينات المسحوبه اقل من 30 لامكانيه استخدام اختبار T لكن اذا كانت اكبر من 30 سوف نستخدم Z
2. يجب ان تكون العينات مستقله
3. يجب ان تكون المجتمعات المسحوب منها العينات متجانسه
4. يجب ان تكون المجتمعات لها التوزيع الطبيعي

- الشرطين الأول والثاني يتأكد منهما الباحث لكن الشرطين الثالث والرابع يمكن التاكد منهما باستخدام حزمة SPSS

- اختبار طبيعية البيانات تم اجرائه سابقا باستخدام الأمر Explore والأمر Frequencies
- اختبار التجانس ويعني ان تباين المجتمعين متساوى وستكون الفروض الاحصائية لها الشكل التالي

$$H_0 : \sigma_1^2 = \sigma_2^2 \ , \quad H_1 : \sigma_1^2 \neq \sigma_2^2$$

- اذا تم قبول فرض العدم فهذا يعني ان هناك تجانس لذا سوف نستمر في الاختبار اما اذا تم رفض فرض العدم فيعني عدم وجود تجانس لذا لا يجوز اجراء الاختبار.

تمرين:

بفرض ان لدينا عينتين من الطلاب وتم تسجيل درجاتهم فى مقرر الاحصاء كالتالى

Sample 1	20	17	10	13	15	14	14		
Sample 2	19	15	17	10	3	8	19	10	16

والمطلوب معرفة هل هناك فرق بين مستوى التحصيل للمجموعتين ام لا؟

الحل:

1- من الواضح ان حجم العينات اقل من 30

2- العينات مستقله

3- يجب اختبار هل المجتمعات لها التوزيع الطبيعى ام لا كما يلى

- سنقوم بادخال البيانات فى متغيرين احداهما يسمى Factor والأخر Data حيث تحتوى data على القراءات فى العينتين والمتغير factor يحتوى الرقم 1 اذا كانت القيمة فى المتغير data من العينة الاولى والرقم 2 اذا كانت القيمة من العينة الثانية، كالتالى

	Factor	Data
1	1.00	20.00
2	1.00	17.00
3	1.00	10.00
4	1.00	13.00
5	1.00	15.00
6	1.00	14.00
7	1.00	14.00
8	2.00	19.00
9	2.00	15.00
10	2.00	17.00
11	2.00	10.00
12	2.00	3.00
13	2.00	8.00
14	2.00	19.00
15	2.00	10.00
16	2.00	16.00
17		

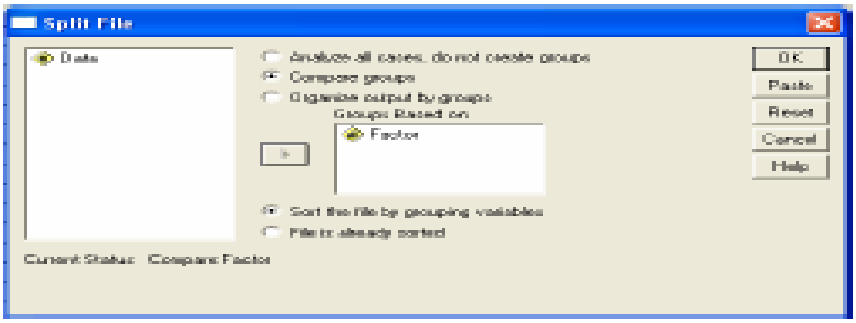
لاجراء اختبار الطبيعية على البيانات data سوف نقوم بالخطوتين التاليتين:

الخطوه الاولى:

1- من قائمة Data نختار Split File

2- تظهر الشاشة التالية نختار Compare groups وننقل المتغير Factor لخانة Groups Based on

3- نختار Ok سيتم تقسيم البيانات لمجموعتين تبعا للمتغير Factor وهو تقسيم غير ظاهرى

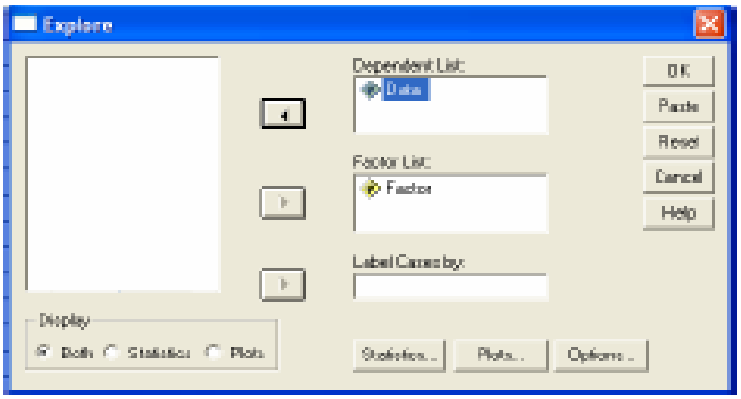


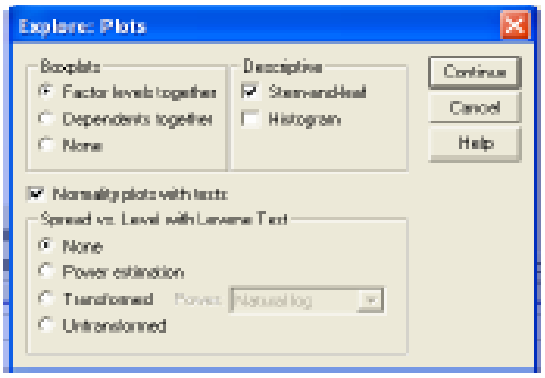
الخطوه الثانية:

1- من قائمة Analyze نختار Descriptive Statistics ثم نضغط Explore

2- تظهر الشاشة التالية ننقل المتغير Data لخانة Dependent List وننقل المتغير Factor لخانة Factor List

3- نختار Plots ثم نحدد الاختبار Normality plots with tests سو نختار Continue نعود للشاشة السابقة ثم نختار Ok تظهر النتائج التالية





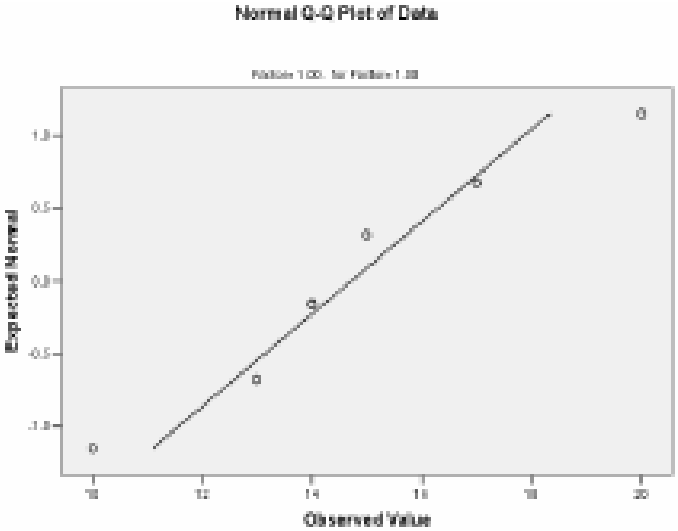
Tests of Normality

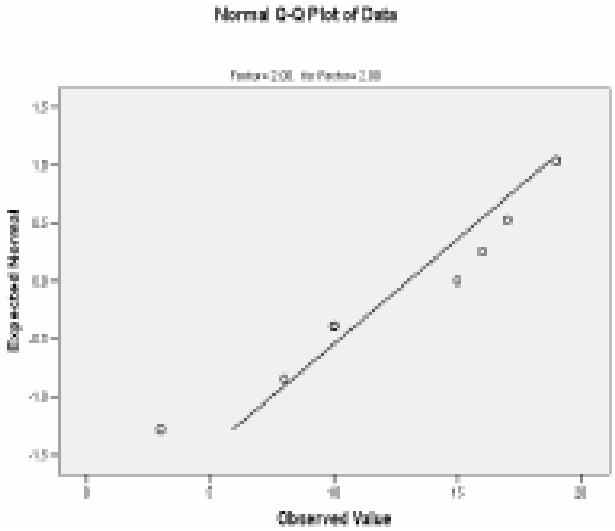
			Factor	
			1.00	2.00
Statistic	Kolmogorov-Smirnov[a]	Data	.178	.197
	Shapiro-Wilk	Data	.988	.915
df	Kolmogorov-Smirnov[a]	Data	7	9
	Shapiro-Wilk	Data	7	9
Sig.	Kolmogorov-Smirnov[a]	Data	.200*	.200*
	Shapiro-Wilk	Data	.883	.354

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

a. Lilliefors Significance Correction

- نجد من جدول Tests of Normality أن قيمة Sig. في جميع الحالات اكبر من 0.05 لذا سوف نقبل فرض العدم القائل ان البيانات لها التوزيع الطبيعي
- ويتضح ذلك ايضا من الرسم البياني التالي





يجب الغاء التقسيم لبيانات الملف indp_samples قبل اجراء اى عمليه اخرى على البيانات.

4- هل المجتمعات المختار منها العينات متجانسه؟

1- من قائمة Analyze نختار Compare Means

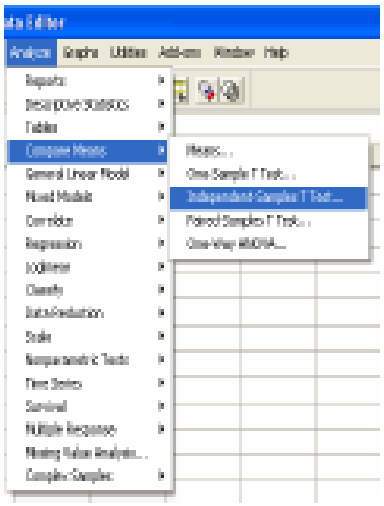
2- من القائمة المنسدله نختار Independent Sample T Test

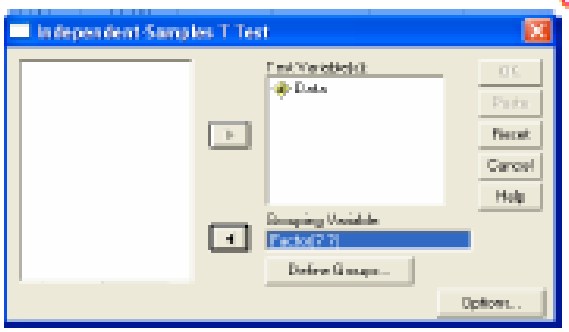
3- ننقل المتغير Data لخانة Test variable(s)

4- ننقل المتغير Factor لخانة Grouping Variable:

5- نضغط على Define Groups لتحديد المجموعات

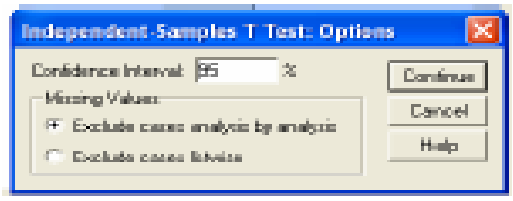
6- تظهر شاشة جديدة نكتب الرقم 1 المميز للمجموعة الأولى والرقم 2 المميز للمجموعة الثانية ثم نختار Continue





7- بالضغط على Option سوف تظهر شاشة جديدة نحدد فيها قيمة فترة الثقة (1- α)100% وستكون باستمرار 95% اى ان مستوى المعنويه α =0.05 وايضا كيفية التعامل مع القيم المفقودة.

8- بالضغط على Continue تم Ok تظهر النتائج التاليه



الجدول الأول:

يحتوى على حجم العينات والوسط والانحراف المعيارى والخطأ المعيارى لكل عينه

Group Statistics

	Factor	N	Mean	Std. Deviation	Std. Error Mean
Data	1.00	7	14.7143	3.14718	1.18952
	2.00	9	13.0000	5.52268	1.84089

الجدول الثاني: يحتوى على اختبارى التجانس واختبار T

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Data	Equal variances assumed	4.481	.053	.731	14	.477	1.71429	2.34614	-3.31788	6.74626
	Equal variances not assumed			.782	13.043	.448	1.71429	2.19177	-3.01916	6.44772

1. العمود الأول يحتوى اسم المتغير Data
2. العمود الثانى والثالث يسارا لاجراء اختبار التجانس وحيث ان قيمة Sig. = 0.053 فهى اكبر من 0.05 لذا سوف نقبل فرض العدم وهوتجانس المجتمعين
3. العمود الرابع والخامس والسادس لاجراء اختبار T وحيث ان المجتمعات متجانسه سوف نهتم بالصف الأول ومن العمود السادس Sig. = 0.477 وهى اكبر من 0.025 لذا سوف نقبل فرض العدم وهو ان وسطى المجتمعين متساوى أى لا يوجد فرق بين مستوى الطلاب فى المجموعتين.
4. الاعمده الاخيرة تقدم فترة الثقة للفرق بين متوسطى المجتمعين.

اختبار عينتين غير مستقلتين

- يستخدم ذلك الاختبار عندما يكون لدينا عينتين غير مستقلتين
- بمعنى ان لدينا عينة واحده ولكل مفرده من مفردات العينه قرائتين، القراءه الاولى تمثل العينة الأولى والقراءه الثانية تمثل العينة الثانيه.
- ويجب قبل اجراء الاختبار التحقق من الشروط التالية

1- العينات غير مستقله

2- عدد المفردات اقل من 30

3- العينات لها التوزيع الطبيعى

- اذا كان عدد المفردات اكبر من 30 فاننا لن نهتم بالشرط الثالث.
- **تمرين:**

بفرض ان لدينا عينه مكونه من 10 اشخاص تم قياس ضغط الدم لكل شخص فى العينه ثم بعد فترة زمنيه تم قياس ضغط الدم مره ثانية لنفس الأشخاص وكانت البيانات كالتالى

Id	1	2	3	4	5	6	7	8	9	10
Before	130	140	150	130	140	145	135	110	120	150
After	120	140	130	150	130	110	110	120	110	140

*

هل يوجد فرق معنوى بين ضغط الدم قبل وبعد تلك الفترة الزمنيه.

الحل: للاجابه على هذا السؤال يجب اختبار الفروض الاحصائية التاليه

$$H_0 : \mu_d = 0,\ H_1 : \mu_d \neq 0,\quad \mu_d = \mu_{Before} - \mu_{After}$$

	Before	After
1	130.00	120.00
2	140.00	140.00
3	150.00	130.00
4	130.00	150.00
5	140.00	130.00
6	145.00	110.00
7	135.00	110.00
8	110.00	120.00
9	120.00	110.00
10	150.00	140.00
11		

وبإدخال البيانات فى متغيرين Before للقراءات قبل الفتره الزمنيه والمتغير After للقراءات بعد الفتره الزمنيه. ثم التحقق من بعض الشروط

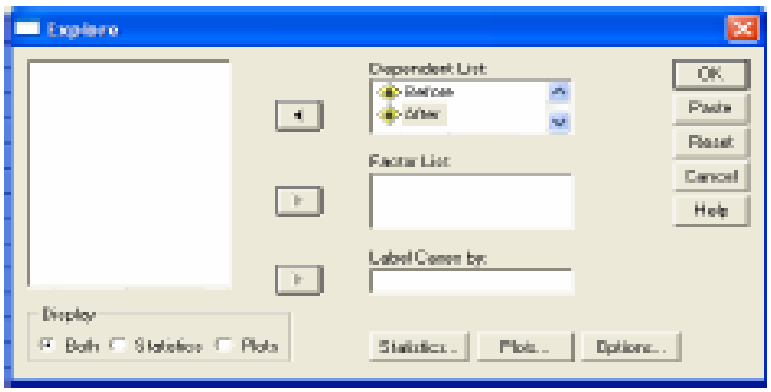
1- القراءات قبل الفتره الزمنيه تختبر عينه من مجتمع القراءات قبل الفترة والقراءات بعد تلك الفتره هى عينه اخرى من مجتمع القراءات بعد الفترة الزمنية وتلك العينات غير مستقله لأنها لنفس الأشخاص.

2- حجم العينات 10 أقل من 30

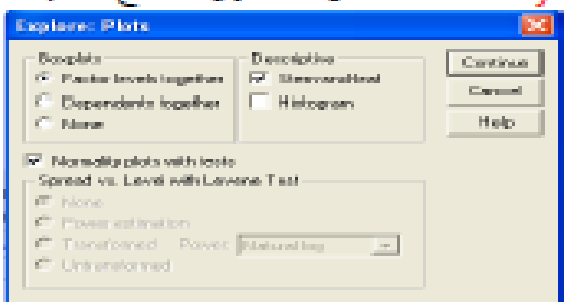
3- يمكن اختبار هل العينات لها التوزيع الطبيعى ام لا؟

من قائمة Analyze نختار Descriptive Statistics ثم نختار Explore

تنقل المتغيرات Before وايضا After لخانه Dependent List



وبالضغط على Plots وبتحديد الاختبار Normality plots with tests سوف تظهر النتائج التاليه



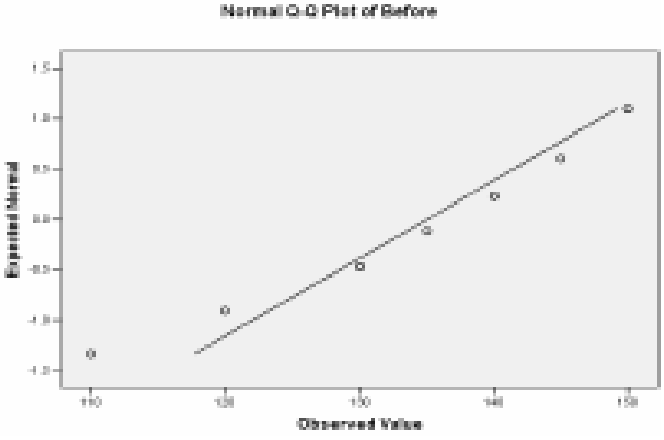
Tests of Normality

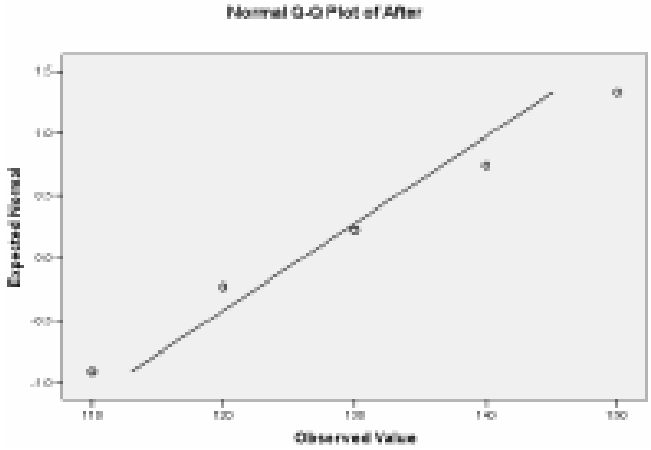
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Before	.151	10	.200*	.936	10	.505
After	.188	10	.200*	.908	10	.268

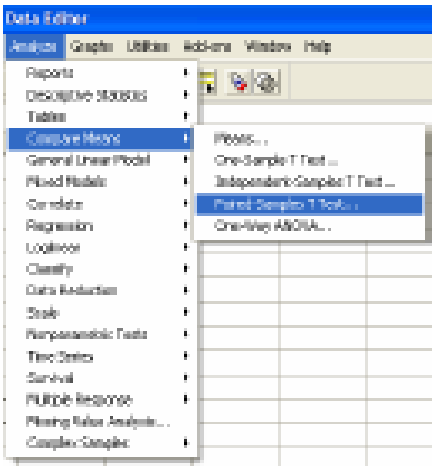
*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

جدول بعنوان Tests of Normality ومنه نجد ان البيانات لها التوزيع الطبيعي ونحصل على نفس النتيجه من الرسم البيانى التالى

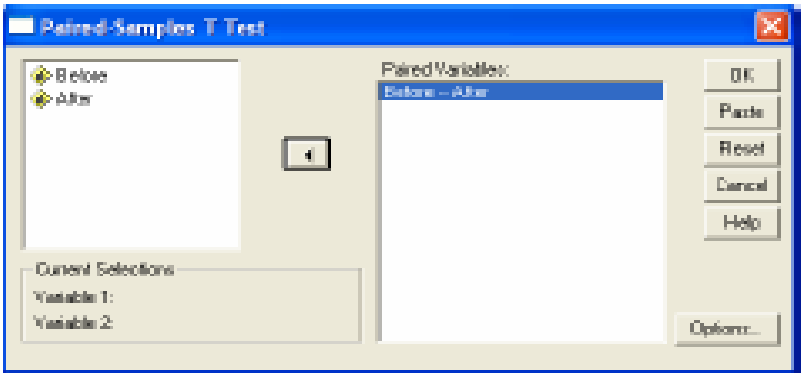


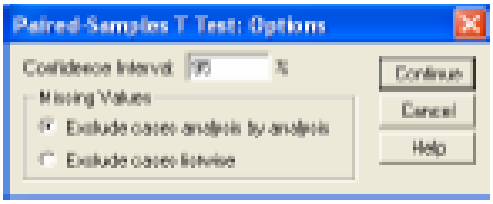




مما سبق نستطيع اجراء اختبار T كما يلي:

1. من قائمة Analyze نختار Compare Means
2. من القائمة المنسدله نختار Paired Samples t Test
3. تظهر شاشة جديده بعنوان Paired Samples T Test ننقل المتغيرين After, Before معا لقائمة Paired Variables
4. نضغط على الاختيار Option تظهر شاشة نحدد فيها مستوى المعنويه وكيفية التعامل مع القيم المفقوده.
5. بالضغط على Continue نعود للشاشة السابقه نختار Ok تظهر النتائج





النتائج Output

الجدول الأول: بعنوان Paired Samples Statistics

يحتوى على الوسط الحسابي وعدد القيم والانحراف المعياري والخطأ المعياري لكل عينه

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Before	136.0000	10	12.90994	4.08248
	After	126.0000	10	14.28841	4.52155

الجدول الثاني: بعنوان Paired Samples Correlations

يحتوى على عدد القيم ومعامل الارتباط بين المتغيرين وايضا قيمة Sig. =0.398 لاختبار معنويه معامل الارتباط وقيمتها اكبر من 0.05 لذا سوف نقبل فرض العدم وهو عدم وجود ارتباط بين عناصر المجتمعين (لا يوجد معنوية للارتباط).

Paired Samples Correlations

	N	Correlation	Sig.
Pair 1 Before & Afte	10	.301	.398

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	Before - After	9.00000	16.12452	5.09902	-2.53478	20.53478	1.765	9	.111

الجدول الثالث: بعنوان Paired Samples Test

1. يحتوى العمود الأول على اسم المتغير الجديد وهو الفرق بين القراءات قبل وبعد الفترة الزمنيه (d=Before-After)
2. يحتوى العمود الثانى على قيمة الوسط الحسابى للفروق بين القراءات
3. العمود الثالث يحتوى على الانحراف المعيار للفروق بين القراءات قبل وبعد الفتره الزمنيه
4. العمود الرابع يحتوى الخطا المعيارى للفروق
5. العمود الخامس يحتوى على فترة 95% ثقه للفرق بين متوسطى المجتمعين
6. العمود السادس يحتوى على قيمة احصاء الاختبار T ودرجة الحرية فى العمود السابع (n-1=9)
7. العمود الاخير يحتوى على Sig. = 0.111 لاختبار Tوهى اكبر من α/2=0.025 لذا سوف نقبل فرض العدم : انه متوسطى المجتمعين متساوى ونرفض الفرض البديل: متوسطى المجتمعين مختلف.