جامعة محمد بوضياف بالمسيلة — السنة الأولى ماستر علوم التربية

كلية العلوم الإنسانية والاجتماعية — تخصص إرشاد وتوجيه

قسم علم النفس — السداسي الثاني

مقياس المعالجة الإحصائية للبيانات التربوية (2)

تحليل الانحدار الخطي المتعدد

المحاضرة الخامسة	المحاضرة السادسة
-أن يميز بين الانحدار الخطي المتعدد والبسيط. - أن يعدد أنواع الانحدار الخطي المتعدد. - أن يحدد شروط استخدام الانحدار الخطي المتعدد. - أن يوجد معادلة الانحدار الخطي المتعدد باستخدام برنامج Excel. - أن يقرأ ويفسر نتائج تحليل الانحدار الخطي المتعدد باستخدام برنامج Excel.	- أن يميز بين أنواع الانحدار الخطي المتعدد من حيث إدخال البيانات. - أن يمتلك مهارة استخدام تحليل الانحدار باستخدام حزمة Spss. - أن يوجد معادلة الانحدار الخطي المتعدد باستخدام حزمة Spss. - أن يقرأ ويفسر نتائج تحليل الانحدار الخطي المتعدد باستخدام حزمة Spss.

العام الدراسي: 2020/2019

الدكتور أحمد سعودي

المحاضرة الخامسة

الانحدار الخطي المتعدد

Multiple Linear Regression

مقدمة:

الانحدار المتعدد (متغيرين مستقلين أو أكثر)، تحليل الانحدار ملائم لمعرفة كم من التباين يستطيع المتغيران معا أن يفسرانه من التباين في المتغير التابع. و لمعرفة تباين المقدار الذي يشرحه المتغير الأول ومعرفة تباين المقدار الذي يشرحه المتغير الثاني نستطيع أن نكشف عن الأهمية لكل من المتغيرين (النجار، 2015، 38).

- من المعروف أن الارتباط الخطي بين متغيرين يقيس العلاقة غير السببية بينهما ليتعرف على درجتها واتجاهها.
- كلمه العلاقة غير السببية نعني بها أن الارتباط لا يهتم بتصنيف المتغيرات إلى متغيرات مستقله ومتغيرات تابعه ولا يهتم بمعرفة أي من المتغيرات يؤثر ومن منها يتأثر.
- نفرض ان لدينا متغيرا (أو مجموعة متغيرات) مستقلا وآخر تابعا، يهدف الانحدار إلى الحصول على الصورة الرياضية للعلاقة التي تربط المتغير (المتغيرات) المستقل بالمتغير التابع. لو أمكن الحصول على هذه الصورة لأطلق عليها اسم معادلة الانحدار Regression Equation .
- معادلة الانحدار قد تكون مضبوطة Exact وقد تكون تقريبيه Approximate .
- وتستخدم معادلة الانحدار في التنبؤ بقيمة المتغير التابع عند قيمة معينه للمتغير (المتغيرات) المستقل.
- إذا كانت المعادلة مضبوطة يتم التنبؤ بدون خطأ لكن إذا كانت تقريبية يكون التنبؤ بخطأ يمكن تقديره والتحكم فيه.
- بصفة عامة يمكن أن يكون الانحدار خطيا Linear، وقد يكون غير خطي Non Linear فإذا كان لدينا متغيران وكانت العلاقة بينما خطية فيمكن تمثيل العلاقة بينهما بمعادلة خط مستقيم، أما إذا كانت العلاقة غير خطية فإنه يمكن تمثيل العلاقة بينهما بمعادلة غير خطية ويمكن معرفة نوع العلاقة في حالة متغيرين عن طريق شكل الانتشار.(عبد الفتاح مصطفى محمد، 2-4)
- ونلاحظ أن الانحدار الخطي البسيط بين متغيرين هي حالة خاصة من الانحدار الخطي المتعدد وسوف نتعرض هنا لكيفية استخدام برنامج Excel والحزمة الإحصائية SPSS للتعامل مع الانحدار وأنواعه المختلفة وخواصه والمقاييس والاحصاءات المرتبطة به.

معادلة الانحدار الخطي المتعدد:

في الانحدار المتعدد نفترض أن لدينا متغيرا تابعا واحدا نرمز له بالرمز Y ومجموعة من المتغيرات المستقلة عددها m متغيرا نرمز لها بالرمز Xm……,X2,X1 كحالة خاصة في الانحدار البسيط يكون: m=1.

نريد هنا الحصول على أفضل معادله انحدار تمثل العلاقة بين المتغير التابع والمتغيرات المستقلة وذلك إذا توافرت لدينا بيانات من عينة حجمها n فيكون النموذج الرياضي الخطي هو:

$$Yj = \beta 0 + \beta 1\ X1j + \beta 2\ X2j + \ldots + \beta\ m{-}1\ X\ mj + \varepsilon\, j\ ,\ j = 1,2,\ldots,n$$

(عبد الفتاح، 17-19)

وإجمالا فإن معادلة الانحدار الخطي المتعدد في العينات:

$$bmxm\ldots\ldots + b3x3 + \hat{y} = a + b1x1 + b2x2$$

a: المقطع الصادي، وهو بعد النقطة التي يقطع عندها خط الانحدار.

b1: معامل الانحدار للمتغير الأول x1 (وزن الانحدار)، تعني التغير في y الذي يقابل تغير x1 وحدة واحدة مع بقاء الأشياء الأخرى ثابتة.

b2: معامل الانحدار للمتغير الثاني x2 (وزن الانحدار)، تعني التغير في y الذي يقابل تغير x2 وحدة واحدة مع بقاء الأشياء الأخرى ثابتة. (النجار، 2015، 38 - 39)

أنواع الانحدار الخطي المتعدد:

الانحدار المتعدد هو امتداد للارتباط بين متغيرين، نتيجة الانحدار هي معادلة تمثل أفضل تقدير للمتغير التابع من عدة متغيرات مستقلة، يستخدم تحليل الانحدار عندما تكون المتغيرات المستقلة مرتبطة ببعضها البعض وبالمتغير التابع.(نجيب وآخرون، 2009، 168- 169)

هناك ثلاثة نماذج انحدار متعدد رئيسية هي:

- **الانحدار العياري أو القياسي (Standard) أو الآني (المتزامن) (Simultaneous).**
- **الانحدار الهرمي (Hierarchical).**

- **الانحدار المتدرج (stepwise).**

هذه النماذج تختلف من جهتين:

- الأولى في معالجة الاختلافات المتداخلة بسبب ارتباط المتغيرات المستقلة.
- والثانية في ترتيب إدخال المتغيرات المستقلة في المعادلة.

على ضوء هذين الاختلافين فإن نماذج الانحدار يمكن تقسيمه الى الثلاثة أقسام التالية:

1) **الانحدار المعياري:** في هذه الطريقة ندخل المتغيرات المستقلة في معادلة الانحدار دفعة واحدة لنحصل على المعادلة التي تصف العلاقة بين كل المتغيرات المستقلة والمتغير التابع مرة واحدة دون مناقشة هل كل المتغيرات المستقلة يجب أن تدخل في المعادلة أم لا؟ ولا نتعرض لمناقشة هل المتغيرات المستقلة مرتبطة بعضها بالبعض أم مستقلة.

2) **الانحدار الهرمي:** في الانحدار الهرمي تدخل المتغيرات المستقلة في المعادلة المقترحة تباعا ونحدد ترتيب دخول هذه المتغيرات في المعادلة المقترحة على أساس إحصائي نظري.

3) **الانحدار التدريجي:**

في نموذج الانحدار المتدرج عدد المتغيرات المستقلة المدخلة في النموذج وكذلك ترتيب إدخالها يحدد من خلال معيار إحصائي يتم الوصول إليه عن طريق إجراء الانحدار المتدرج:

- **طريقة الإدخال الأمامية Forward**: وتتضمن إدخال المتغيرات المستقلة كل واحد على حدة، ويعتمد ترتيب الإدخال ومدى قبول المتغير المستقل على أساس ما إذا كانت قيمة F حرجة محددة (FIN) وما إذا كان المستوى الحرجة لألفا (PIN) قد تحقق.
- **طريقة الخيار العكسي Backward**: يبدأ الخيار العكسي بكل المتغيرات المستقلة في المعادلة ويقوم بالتدرج بحذف أسوء المتغيرات على أساس أن قيمة F الجزئية تكون أقل من قيمة حرجة (FOUT) وكما أن المعيار المقبول (POUT) يجب أن يتحقق أيضا.
- **طريقة الخيار المتدرج (التدريجي)**: خلط بين الطريقتين (Stepwise)، هو مزيج بين الطريقتين الأمامية والعكسية وهي تسمح بالإزاحة المتأخرة للمتغيرات التي سبق أن دخلت في النموذج.
- **إن اختيار الطريقة يعتمد بصورة كبيرة على أهداف الباحث** (نجيب وآخرون، 2009، 169- 170)
- **المفاضلة بين الطرق الثلاثة الموضحة تختلف من باحث الى أخر وتتوقف كثيرا على خبرة الباحث في الانحدار أو خبرة الباحث الإحصائية.**

فروض الاختبار Assumption Testing:

هناك عدد من الفروض تحكم استخدام الانحدار:

1. **نسبة الحالات إلى المتغيرات المستقلة:** عدد الحالات المطلوبة يعتمد على نوع نموذج الانحدار المستخدم، بالنسبة إلى الانحدار القياسي أو الهرمي فيجب أن يكون العدد المثالي للحالات أو المفردات عشرين ضعف عدد المتغيرات المستقلة، بينما الحالات التي نحتاجها في الانحدار المتدرج يتعدى ذلك. والحد الأدنى المطلوب هو خمس أضعاف عدد المتغيرات المستقلة على الأقل.
2. **القيم المتطرفة:** الحالات المتطرفة أو الشاذة لها تأثير قوي على حل معادلة الانحدار ولذا يجب حذفها أو تعديلها لتقليل تأثيرها. يمكن اكتشاف القيم المتطرفة في المتغير الواحد خلال عرض البيانات. أما القيم المتطرفة في المتغيرات المتعددة فإنه يمكننا اكتشافها باستخدام الطرق الإحصائية مثل: مسافة Mahalanobis وكذلك بطريقة العرض البياني مثل: شكل الانتشار للأخطاء. يجب أن يستخدم قرار حذف القيم المتطرفة من البيانات بحذر لأن حذفها سينتج عنه عادة حالات أكثر تطرفا.
3. **الارتباط القوي بين المتغيرات المستقلة:** الازدواج الخطي Multicollinearity تشير للارتباط القوي بين المتغيرات المستقلة، في حين أن singularity تحدث عند حدوث ارتباط تام بين المتغيرات المستقلة. هذه المشاكل تؤثر على طريقة تفسير أي علاقة بين المتغيرات المستقلة والمتغير التابع، ويمكن اكتشاف ذلك من خلال فحص مصفوفة الارتباط المتعدد. معظم برامج الحاسب لها قيمة محددة في الازدواج الخطي ولا تسمح بالمتغيرات التي تسبب مشكلة.
4. **التوزيع الطبيعي والعلاقة الخطية والتجانس واستقلالية البواقي:** يسمح فحص شكل الانتشار للبواقي في اختبار الفروض السابقة. من المفترض أن الفرق بين القيم الحقيقية والمقدرة للمتغير التابع له توزيع طبيعي. وأيضا من المفروض أن البواقي لها علاقة خطية مع القيمة المقدرة للمغير التابع، وأن التباين للبواقي متساو لكل القيم المقدرة. والانحرافات البسيطة في العلاقة الخطية لا تكون جدية، أما الانحرافات المتوسطة إلى المرتفعة قد تؤدي إلى قيم تقديرية أقل للعلاقة.

الفرض الأول خاص بتصميم البحث، أما الفرض 2، 3، 4 فنتأكد منها من خلال تحليل الانحدار. (نجيب وآخرون، 2009، 170- 171)

تحليل الانحدار باستخدام برنامج Excel

مثال: لتكن البيانات التالية لعينة عشوائية مسحوبة من موظفي مؤسسة ما:

الرقم	الراتب الشهري	سنوات الخبرة Ex	عدد سنوات التعليم الجامعي Edu	الجنس
1	56900	5.5	4	F
2	60500	9	4	M
3	58900	4	5	F
4	59000	8	4	M
5	57500	9.5	5	M
6	55500	3	4	F
7	56000	7	3	F
8	52700	1.5	4.5	F
9	65000	8.5	5	M
10	60000	7.5	6	F
11	56000	9.5	2	M
12	53600	6	2	F
13	55000	2.5	4	M
14	52500	1.5	4.5	M

المطلوب:

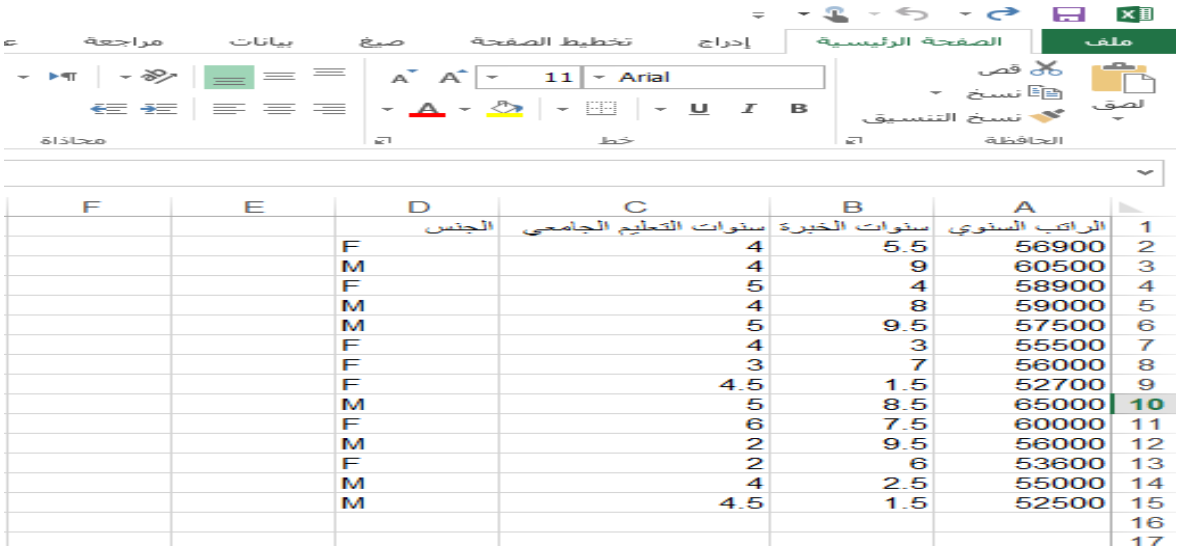
1) أوجد معادلة الانحدار الخطي المتعدد للمتغيرات المستقلة الثلاثة على المتغير التابع (الراتب السنوي).

2) أعد تقدير معادلة الانحدار المعنوية.

3) بفرض أنه يوجد مهندس موظف بالمؤسسة، عدد سنوات خبرته عشر سنوات، ودرس بالجامعة 5 سنوات .كم تتوقع راتبه السنوي؟

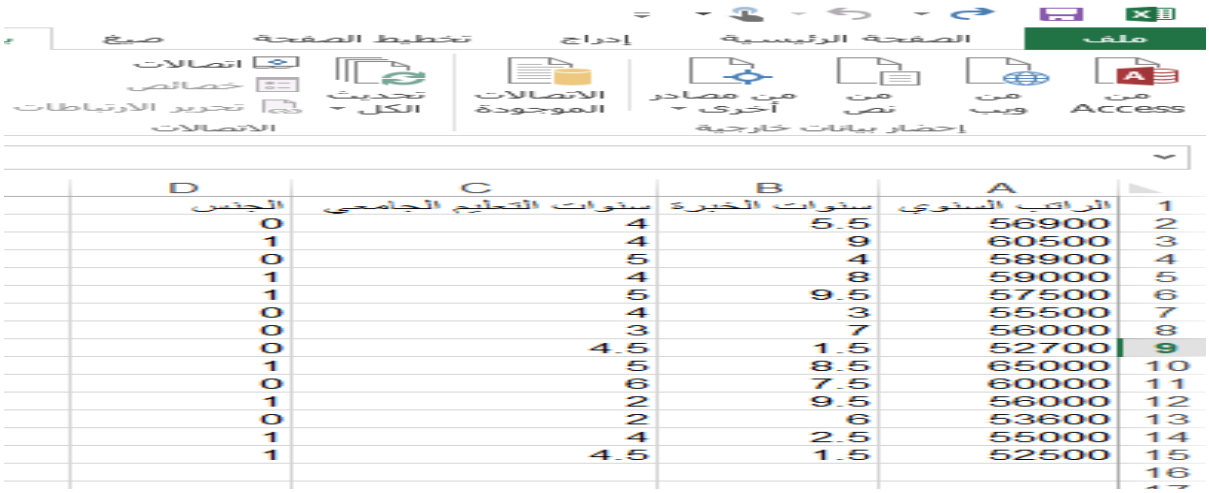
لحساب تحليل الانحدار باستخدام برنامج Excel نتبع الخطوات التالية:

- **نصب البيانات على صفحة Excel**



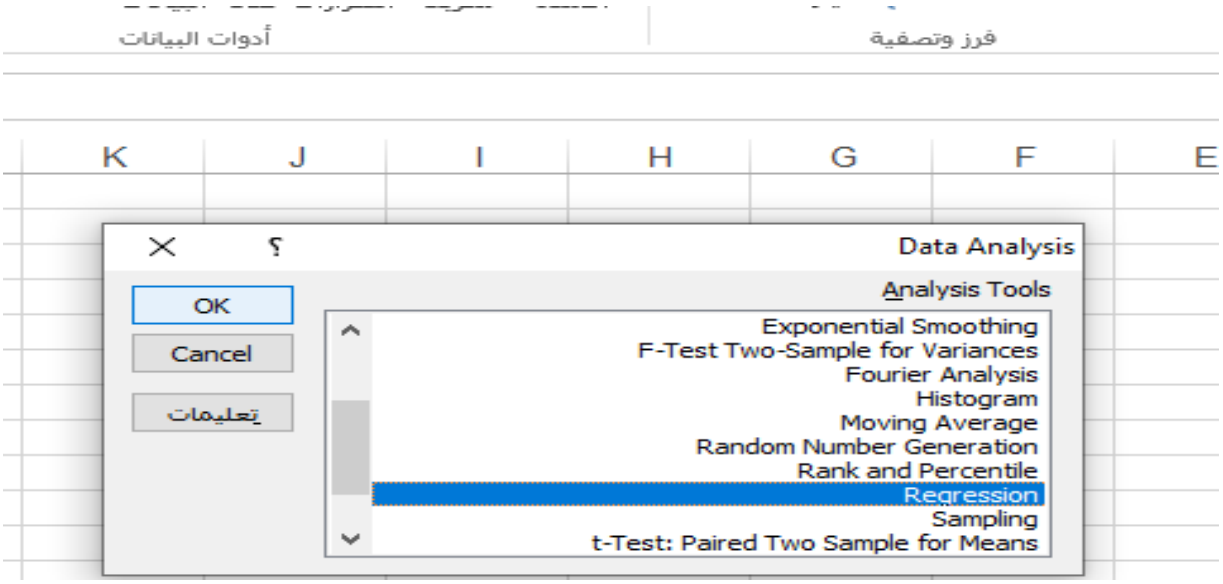
	A	B	C	D	E	F
1	الراتب السنوي	سنوات الخبرة	سنوات التعليم الجامعي	الجنس		
2	56900	5.5	4	F		
3	60500	9	4	M		
4	58900	4	5	F		
5	59000	8	4	M		
6	57500	9.5	5	M		
7	55500	3	4	F		
8	56000	7	3	F		
9	52700	1.5	4.5	F		
10	65000	8.5	5	M		
11	60000	7.5	6	F		
12	56000	9.5	2	M		
13	53600	6	2	F		
14	55000	2.5	4	M		
15	52500	1.5	4.5	M		
16						
17						

نستبدل رمز الإناث **F** برقم **0** ورمز الذكور **M** بالرقم **1** (تشفير البيانات الإسمية) لأن تحليل الانحدار لا يتعامل إلا مع الأرقام.

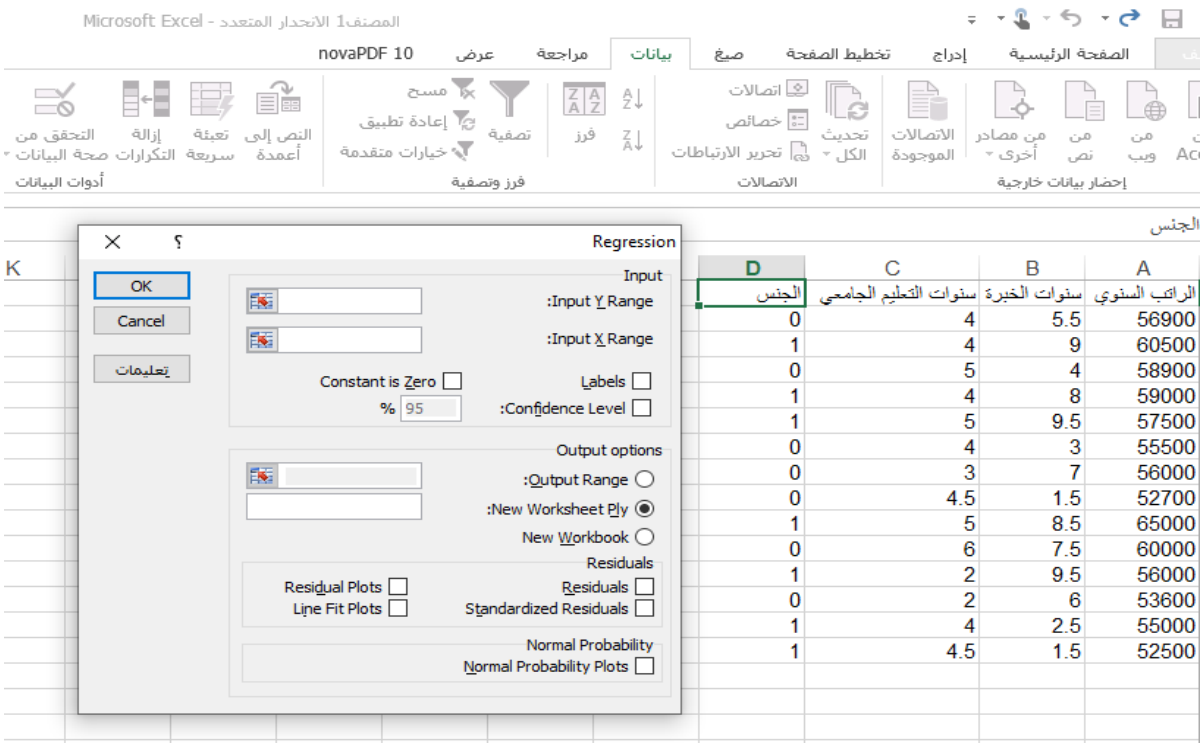


	A	B	C	D
1	الراتب السنوي	سنوات الخبرة	سنوات التعليم الجامعي	الجنس
2	56900	5.5	4	0
3	60500	9	4	1
4	58900	4	5	0
5	59000	8	4	1
6	57500	9.5	5	1
7	55500	3	4	0
8	56000	7	3	0
9	52700	1.5	4.5	0
10	65000	8.5	5	1
11	60000	7.5	6	0
12	56000	9.5	2	1
13	53600	6	2	0
14	55000	2.5	4	1
15	52500	1.5	4.5	1
16				

من شريط المهام العلوي نختار أيقونة بيانات **Data** ثم نضغط على **Data Analysis**

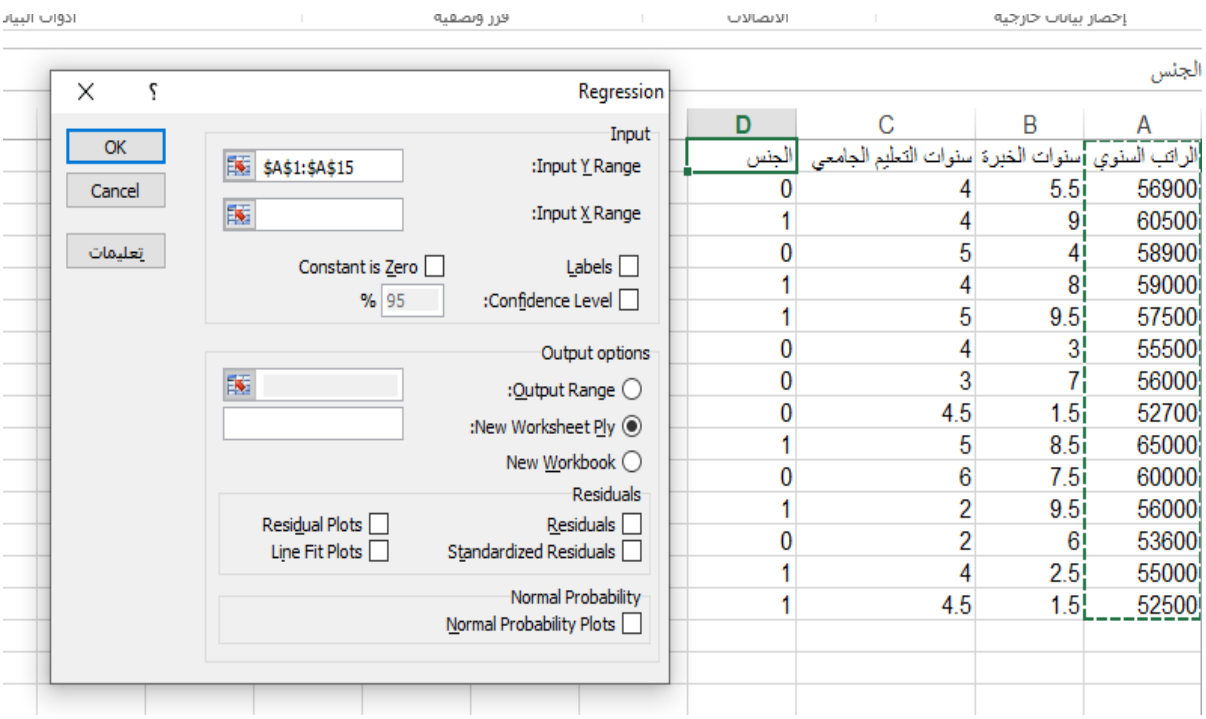


نحصل على مربع الحوار الموضح في الشكل والذي يتضمن جملة الأساليب الإحصائية المتاحة، نختار منها **Regression** ثم نضغط على الزر موافق **OK**.



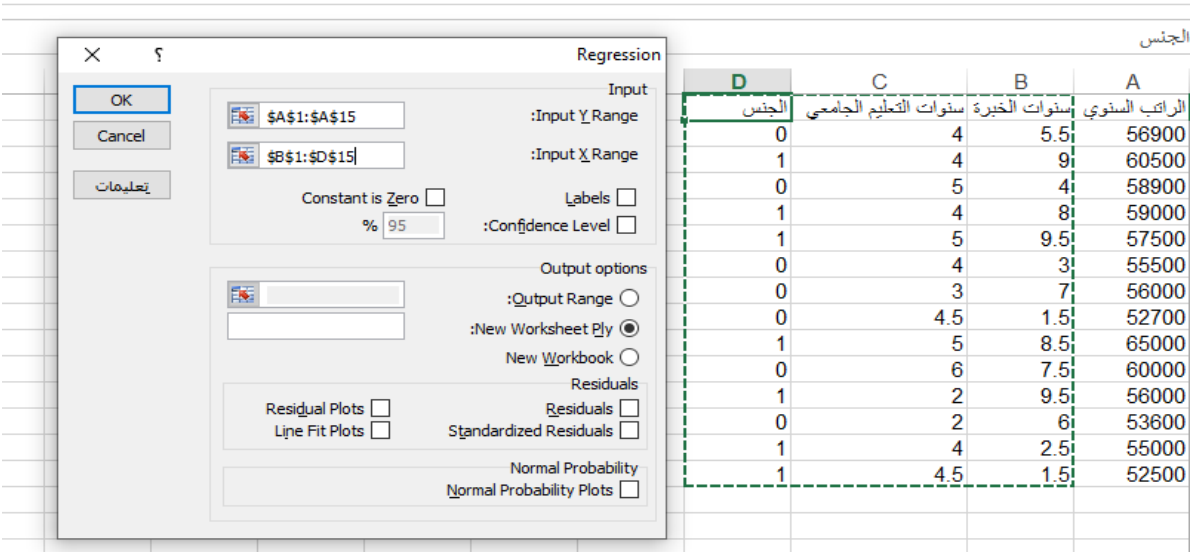
الجنس	سنوات التعليم الجامعي	سنوات الخبرة	الراتب السنوي
0	4	5.5	56900
1	4	9	60500
0	5	4	58900
1	4	8	59000
1	5	9.5	57500
0	4	3	55500
0	3	7	56000
0	4.5	1.5	52700
1	5	8.5	65000
0	6	7.5	60000
1	2	9.5	56000
0	2	6	53600
1	4	2.5	55000
1	4.5	1.5	52500

يظهر مربع الحوار في الأعلى، ونشرع في إدخال بيانات المتغيرات على النحو التالي:

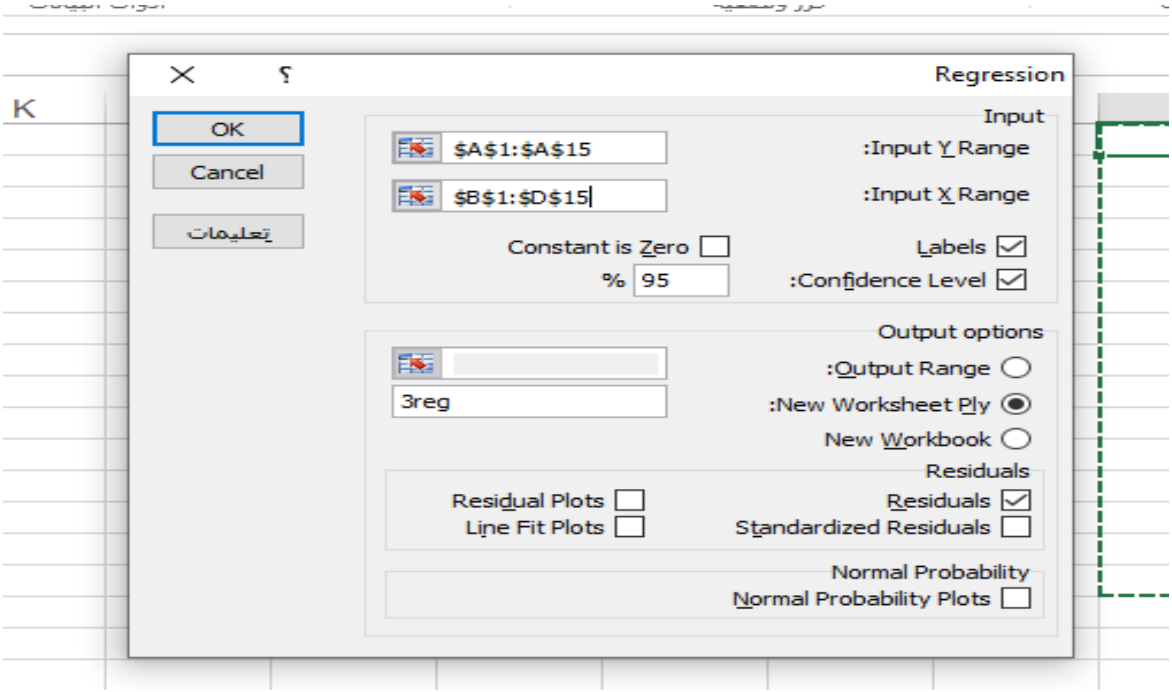


الجنس	سنوات التعليم الجامعي	سنوات الخبرة	الراتب السنوي
0	4	5.5	56900
1	4	9	60500
0	5	4	58900
1	4	8	59000
1	5	9.5	57500
0	4	3	55500
0	3	7	56000
0	4.5	1.5	52700
1	5	8.5	65000
0	6	7.5	60000
1	2	9.5	56000
0	2	6	53600
1	4	2.5	55000
1	4.5	1.5	52500

ندخل بيانات المتغير التابع (الراتب الشهري) في الخانة Input Y Range



ثم ندخل بيانات المتغيرات المستقلة (الخبرة، سنوات التعليم، الجنس) في الخانة Input X Range



نتأكد من التأشير على الخانات المبينة في الشكل، ثم نضغط على زر موافق OK.

فنصحل على المخرجات التالية:

SUMMARY OUTPUT

Regression Statistics	
0.82159	Multiple R
0.675011	R Square
0.577514	Adjusted R Square
2251.716	Standard Error

14 Observations

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	1.05E+08	35103300	6.923422	0.008376
Residual	10	50702242	5070224		
Total	13	1.56E+08			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	45495.33	2809.771	16.19183	1.67E-08	39234.77	51755.89	39234.77	51755.89
سنوات الخبرة	801.5711	228.4684	3.508455	0.005646	292.5117	1310.63	292.5117	1310.63
سنوات التعليم الجامعي	1595.737	560.644	2.846256	0.017361	346.5439	2844.929	346.5439	2844.929
الجنس	382.5721	1287.411	0.297164	0.772423	-2485.96	3251.102	-2485.96	3251.102

تفسير المخرجات:

الجدول الأول:

- معامل الارتباط المتعدد (Multiple R): يساوي 0.82 وبالتالي العلاقة بين المتغير التابع من جهة (المرتب السنوي)، وبين المتغيرات المستقلة من جهة أخرى علاقة قوية طردية، بمعنى أنه مع ازدياد قيم المتغيرات المستقلة المعنوية يزداد المرتب السنوي.
- معامل التحديد (R Square): يساوي 0.68 وبالتالي فإن النموذج يفسر 68% من البيانات. بمعنى أننا إذا تنبأنا باستخدام هذا النموذج سنحصل على نتائج دقيقة باحتمال 68%.
- معامل التحديد المعدل (Adjusted R Square): يساوي 0.58، وهو يفيد في اختيار النموذج الأكثر معنوية، إذ أن هذا المقياس يزيد في إحدى الحالتين: إضافة متغير مستقل معنوي إلى المعادلة، أو حذف متغير مستقل غير معنوي من المعادلة كما سنرى لاحقا عندما سنحذف متغير الجنس غير المعنوي.
- الخطأ المعياري للتقدير (Standard Error): يساوي 2,251.7.

جدول تحيل التباين:

يتضح من قيمة المعنوية (Significant F = 0.0084) أن النموذج معنوي، بمعنى أن النموذج يصلح للتنبؤ.

جدول المعاملات:

- معامل التقاطع: يساوي 45495.329 وهو معنوي (**8-10×1.67=08-E1.67**).
- معامل متغير عدد سنوات الخبرة: يساوي 801.57 وهو معنوي (0.0056)، ويلاحظ بأن قيمته الموجبة تدل على العلاقة الطردية بين عدد سنوات الخبرة و الراتب السنوي. بمعنى أنه مع زيادة عدد سنوات الخبرة يزداد المرتب السنوي.
- معامل متغير عدد سنوات التعليم الجامعي: يساوي 1595.74 وهو معنوي (0.017)، ويلاحظ أيضا من قيمته الموجبة أن علاقته طردية مع الراتب السنوي. ونلاحظ أن معنويته أكبر من معنوية معامل عدد سنوات الخبرة، الذي يدل على أن عدد سنوات الخبرة أكثر تأثير في الراتب السنوي من عدد سنوات التعليم الجامعي.
- معامل متغير الجنس: يساوي 382.57 لكنه غير معنوي (0.772) بمعنى أنه لا يوجد تمييز بين مرتبات الذكور والإناث في هذه المؤسسة.

RESIDUAL OUTPUT

Residuals	Predicted الراتب السنوي	Observa
-1386.92	56286.92	1
1025.012	59[illegible]74.99	2
2219.704	56680.3	3
326.5836	58673.42	[illegible]
-3971.51	61471.51	[illegible]
1217.011	54282.99	6
106.4633	55893.54	7
-1178.5	53878.5	8
4330.061	60669.94	9
-1081.53	61081.53	10
103.7709	53496.23	12
735.2246	54264.78	13
-1761.07	54261.07	14

تؤخذ المعلومات من السطر الأول في الشكل وتحسب القيمة المتوقعة لهذا الراتب مثلا بالشكل:

56,286.9= 45,495 +801.6(5.5)+ 1,595.7(4)+ 382.6(0)

الفرق بين الراتب المتنبأ به باستخدام المعادلة (56,286.9) والراتب الفعلي للمشاهدة الأولى (السطر الأول في

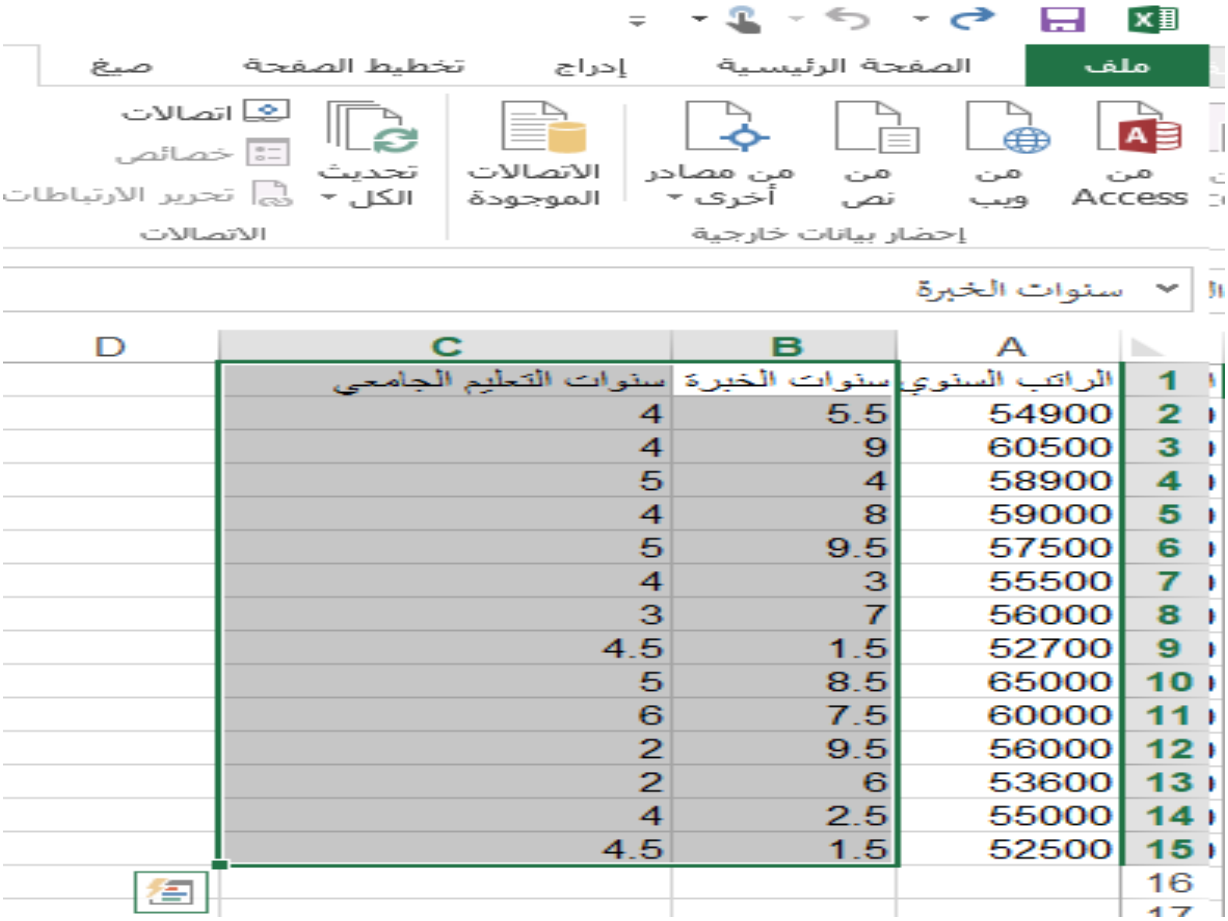
بالطبع كلما صغرت قيم الفروقات دل ذلك على أن النموذج أفضل للتنبؤ، والنتائج أدق.

الخلاصة: وجد أن أكثر عامل مؤثر في الراتب السنوي هو الخبرة، ثم يأتي عدد سنوات التعليم، لكن لم يكن هنالك تأثير للجنس في الراتب السنوي. والمعادلة الناتجة هي:

الجنس* 382.6 + سنوات الدراسة* 1595.7 + الخبرة* 801.6 + 45495.3 = الراتب

Salary = 45495.3 + 801.6 Experience + 1,595.7 Education + 382.6 Gender

لكن يوجد في هذه المعادلة متغير الجنس الذي يؤثر وجوده على نتائج التنبؤ، لكنه بالمقابل ليس معنويا. لذا لا بد من حذفه من المعادلة للحصول على نتائج تنبؤ أدق. أي لا بد من إعادة إيجاد معادلة دالة الانحدار لكن هذه المرة بدون إدخال متغير الجنس في المعادلة؛ بمعنى أن نطاق المتغيرات المستقلة الجديد الخلية Input X (Range) في نافذة مدخلات الانحدار (Regression) يوضحه الشكل الآتي:



	A	B	C	D
1	الراتب السنوي	سنوات الخبرة	سنوات التعليم الجامعي	
2	54900	5.5	4	
3	60500	9	4	
4	58900	4	5	
5	59000	8	4	
6	57500	9.5	5	
7	55500	3	4	
8	56000	7	3	
9	52700	1.5	4.5	
10	65000	8.5	5	
11	60000	7.5	6	
12	56000	9.5	2	
13	53600	6	2	
14	55000	2.5	4	
15	52500	1.5	4.5	
16				

أما نطاق المتغير التابع فيبقى على حاله، فتظهر لدينا المخرجات الموضحة في الجداول التالية والتي تقارن أيضا بن المخرجات قبل وبعد حذف متغير الجنس:

قبل حذف متغير الجنس	بعد حذف متغير الجنس
SUMMARY OUTPUT	SUMMARY OUTPUT

Regression Statistics	
0.82159	Multiple R
0.675011	R Square
0.577514	Adjusted R Square
2251.716	Standard Error
14	Observations

Regression Statistics	
0.819842	Multiple R
0.672141	R Square
0.61253	Adjusted R Square
2156.385	Standard Error
14	Observations

نلاحظ أن قيم معاملات الارتباط المتعدد لم تتغير تقريبا، وكذلك معامل التحديد. والسبب في ذلك أن متحول الجنس لم يكن يساهم في الأصل في النموذج معامل معنوي، بمعنى أنه لا توجد أية علاقة بين متحول الجنس ومتحول الراتب التابع، لذا فإن حذفه أو إضافته لا تؤثر في علاقة الارتباط المتعدد، لكن بالنظر إلى قيمة معامل التحديد المعدل (Adjusted R Square) نلاحظ أن قيمته قد ازدادت بعد حذف متحول الجنس المستقل من المعادلة وهذا دليل على أن متحول الجنس غير معنوي، إذ أن قيمة معامل التحديد المعدل تزداد عند حذف متحول مستقل غير معنوي، أو عند إضافة متحول مستقل معنوي.

	ANOVA					
	df	*SS*	*MS*	*F*	*F*	*Significance F*
قبل حذف متغير الجنس	3	1.05E+08	35103300	6.923422	0.008376	
	10	50702242	5070224			
	13	1.56E+08				
	ANOVA					
	df	*SS*	*MS*	*F*	*F*	*Significance F*
بعد حذف متغير الجنس	2	1.05E+08	52431084	11.27551	0.002169	
	11	51149975	4649998			
	13	1.56E+08				

نلاحظ أن معنوية النموذج قد ازدادت بعد حذف متغير الجنس من 0.008 إلى 0.002 بمعنى أن النموذج الجديد أفضل للتنبؤ من النموذج السابق.

	Upper 95.0%	Lower 95.0%	Upper 95%	Lower 95%	P-value	t Stat	Standard Error	Coefficients	
قبل حذف متغير الجنس	51755.89	39234.77	51755.89	39234.77	1.67E-08	16.19183	2809.771	45495.33	Intercept
	1310.63	292.5117	1310.63	292.5117	0.005646	3.508455	228.4684	801.5711	سنوات الخبرة
	2844.929	346.5439	2844.929	346.5439	0.017361	2.846256	560.644	1595.737	سنوات التعليم الجامعي
	3251.102	-2485.96	3251.102	-2485.96	0.772423	0.297164	1287.411	382.5721	الجنس
	Upper 95.0%	*Lower 95.0%*	*Upper 95%*	*Lower 95%*	*P-value*	*t Stat*	*Standard Error*	*Coefficients*	
بعد حذف متغير الجنس	51432.69	39589.98	51432.69	39589.98	3.19E-09	16.9167	2690.32	45511.34	Intercept
	1275.881	375.4545	1275.881	375.4545	0.00196	4.036492	204.5509	825.668	سنوات الخبرة
	2784.075	423.3226	2784.075	423.3226	0.01229	2.990332	536.2946	1603.699	سنوات التعليم الجامعي

من الواضح أن المتغيرات المستقلة المعنوية قد حافظت على معنويتها.

قبل حذف متغير الجنس			بعد حذف متغير الجنس		
RESIDUAL OUTPUT			RESIDUAL OUTPUT		
Observation	*Predicted الراتب السنوي*	*Residuals*	*Observation*	*Predicted الراتب السنوي*	*Residuals*
1	56286.92	-1386.92	1	56467.31	-1567.31
2	59474.99	1025.012	2	59357.14	1142.855
3	56680.3	2219.704	3	56832.5	2067.496
4	58673.42	326.5836	4	58531.48	468.5233
5	61471.51	-3971.51	5	61373.68	-3873.68
6	54282.99	1217.011	6	54403.14	1096.863
7	55893.54	106.4633	7	56102.11	-102.11
8	53878.5	-1178.5	8	53966.48	-1266.48
9	60669.94	4330.061	9	60548.01	4451.99
10	61081.53	-1081.53	10	61326.04	-1326.04
11	56684.3	-684.3	11	56562.58	-562.581
12	53496.23	103.7709	12	53672.74	-72.7427
13	54264.78	735.2246	13	53990.3	1009.697
14	54261.07	-1761.07	14	53966.48	-1466.48

من الواضح أن قيم الفروقات قد صغرت بعد حذف متحول الجنس، الأمر الذي يدل على حصولنا على نموذج أفضل وأدق للتنبؤ.

في المحصلة، إن النموذج المعنوي النهائي:

الراتب = 45511.3 + 825.6 الخبرة* + 1603.7 سنوات الدراسة*

Salary = 45,511.3 + 825.6 Experience + 1,603.7 Education

المهندس الموظف في المؤسسة صاحب 10 سنوات خبرة ، و عدد سنوات تعليمه الجامعي 5 سنوات يكون دخله السنوي من خلال المعادلة السابقة:

Salary = 45,511.3+ 825.6 (10) + 1,603.7 (5) = 61,785.8

حصة الأعمال الموجهة

تطبيق: للتنبؤ بالمتغير y انطلاقا من المتغيرين X1 و X2 وباستخدام برنامج Excel

فتحصل على النتائج التالية:

SUMMARY OUTPUT

Regression Statistics	
0.982008	Multiple R
0.964339	R Square
0.958395	Adjusted R Square
8.139439	Standard Error
15	Observations

ANOVA

Significance F	*F*	*MS*	*SS*	*df*	
0.0	162.3	10749.2	21498.3	2	Regression
		66.3	795.0	12	Residual
			22293.3	14	Total

Upper 95.0%	*Lower 95.0%*	*Upper 95%*	*Lower 95%*	*P-value*	*t Stat*	*Standard Error*	*Coefficients*	
148.9825	-15.691	148.9825	-15.691	0.10322	1.763595	37.78971	66.64574	Intercept
0.08472	0.024312	0.08472	0.024312	0.001989	3.932596	0.013863	0.054516	X1
-0.39718	-6.03377	-0.39718	-6.03377	0.028648	-2.48587	1.2935	-3.21548	X2

RESIDUAL OUTPUT

Residuals	*Predicted Y*	*Observation*
-3.28324	63.28324	1
-5.16581	75.16581	2
5.813535	74.18646	3
3.93097	86.06903	4
2.048405	97.95159	5
6.950362	113.0496	6
-5.28577	115.2858	7
2.12453	107.8755	8
-2.61995	132.6199	9
-11.287	141.287	10
-3.52317	143.5232	11
18.16331	161.8367	12
-10.5038	170.5038	13
-2.73991	172.7399	14
5.377529	184.6225	15

المحاضرة السادسة

تحليل الانحدار باستخدام الحزمة الإحصائية SPSS

مثال:

لدراسة العلاقة بين المبيعات اليومية لعدد 15 محل تجاري في يوم معين من سلعة معينة ودرجات الحرارة المسجلة في منتصف اليوم وسنوات الخبرة للبائع الذي يقوم بالخدمة. تم تجميع النتائج في الجدول التالي:

سنوات الخبرة	درجة الحرارة	المبيعات
1	21	15
1	18	15
1	22	21
2	24	28
2	25	30
2	25	35
2	26	40
3	34	35
3	25	30
3	38	45
4	40	50
4	41	60
5	39	45
5	37	60
6	40	50

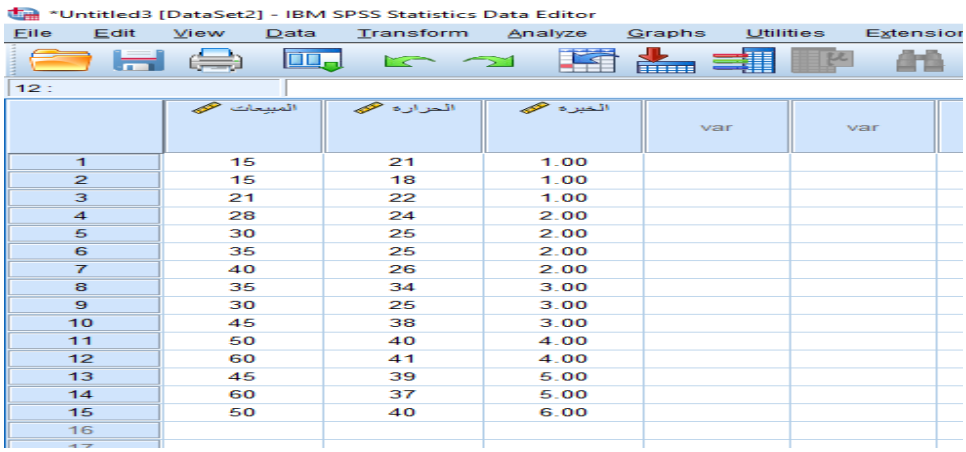
- سنفترض أن المبيعات هي المتغير التابع وأن سنوات الخبرة ودرجه الحرارة هي المتغيرات المستقلة.

كيفية استخدام الحزمة spss في الحصول على:

الانحدار العياري أو القياسي Standard

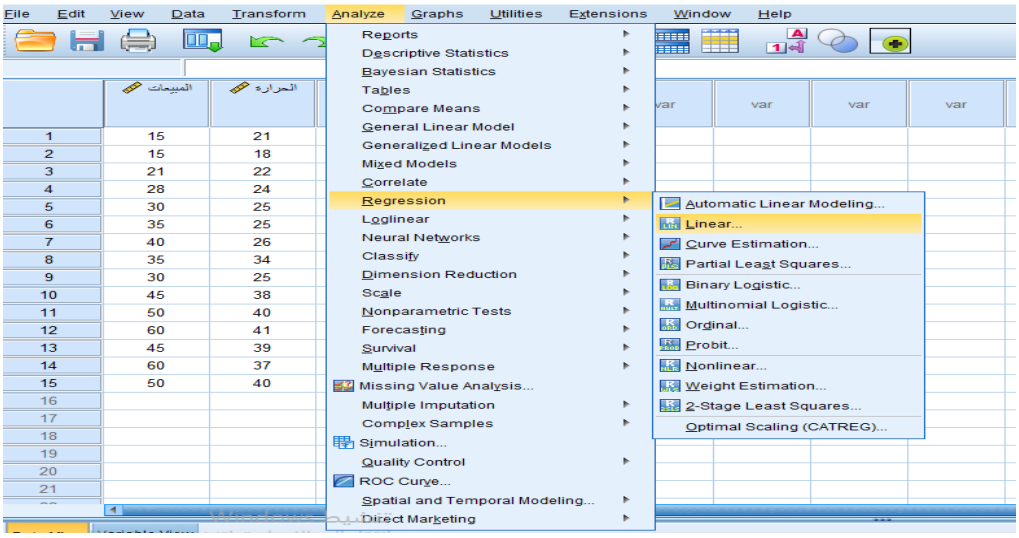
أو الآني (المتزامن) Simultaneous.

باستخدام بيانات الملف "regression.sav" للحصول على الانحدار العياري نتبع الخطوات التالية:

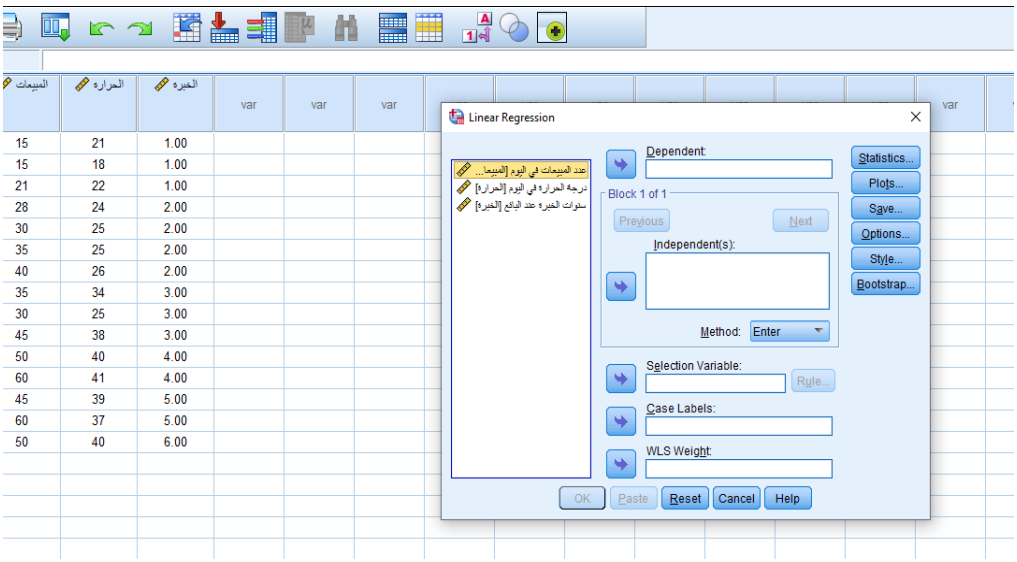


	المبيعات	الحراره	الخبره	var	var
1	15	21	1.00		
2	15	18	1.00		
3	21	22	1.00		
4	28	24	2.00		
5	30	25	2.00		
6	35	25	2.00		
7	40	26	2.00		
8	35	34	3.00		
9	30	25	3.00		
10	45	38	3.00		
11	50	40	4.00		
12	60	41	4.00		
13	45	39	5.00		
14	60	37	5.00		
15	50	40	6.00		
16					

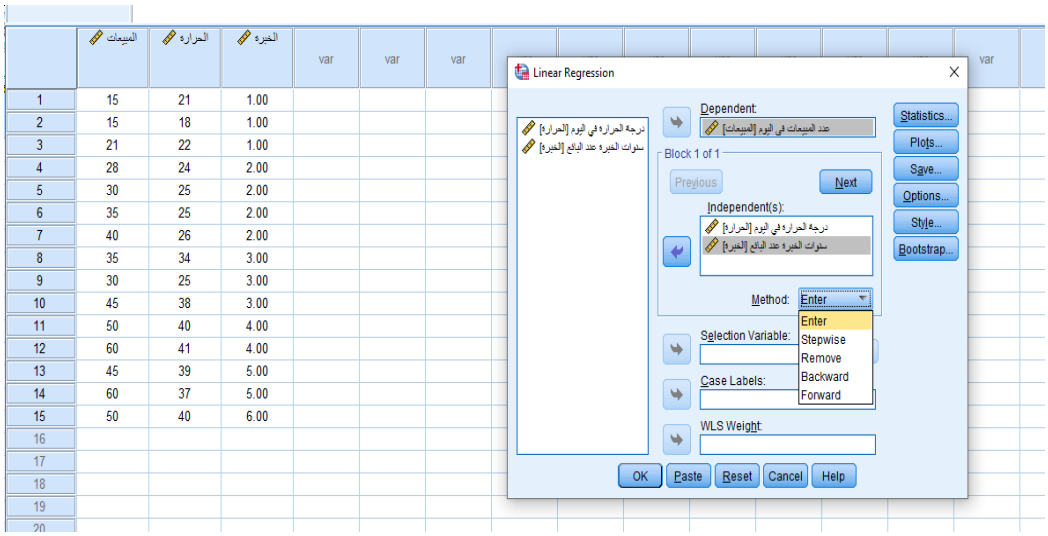
- من قائمة Analyze نختار regression.
- تظهر قائمة منسدلة تحتوي على عدد من أنواع الانحدار سوف نختار هنا منها Linear.



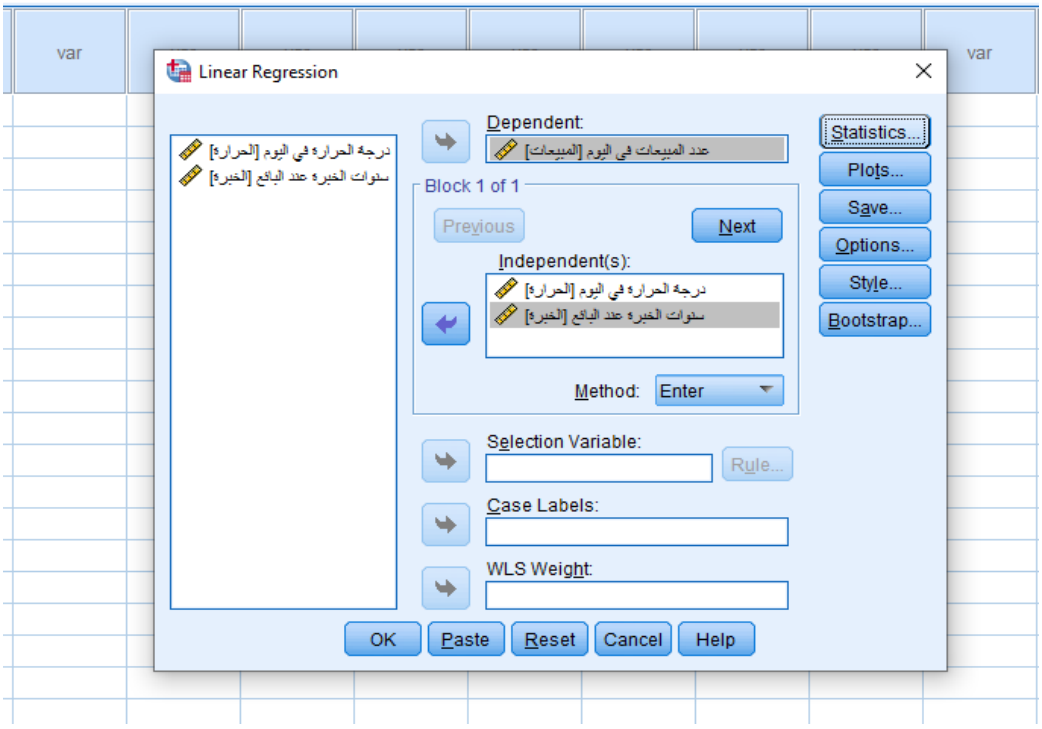
- تظهر شاشة جديدة بعنوان Linear regression.



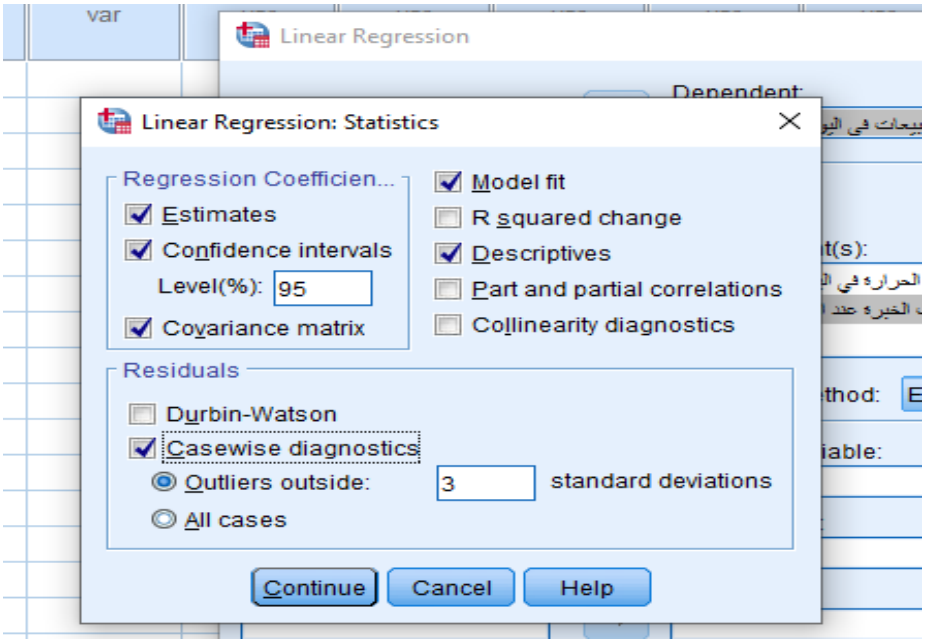
- ننقل المتغير التابع **المبيعات** لخانة Dependent .
- ننقل المتغيرات المستقلة **الحرارة** و **سنوات الخبرة** لخانة Independent .



- نختار نوع الانحدار من خانة Method ونجد أن الطريقة المختارة من قبل البرنامج هي الطريقة العيارية وهي Enter .

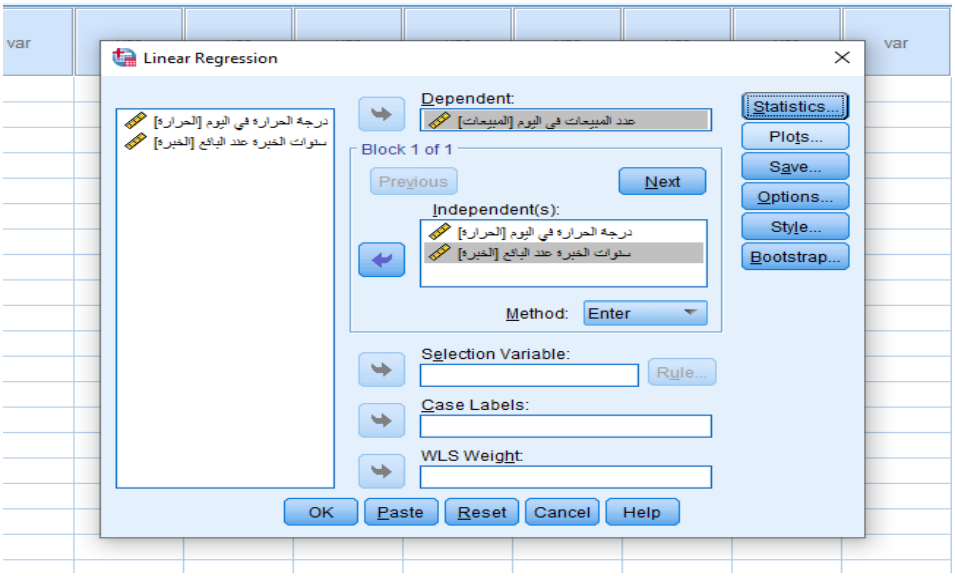


- نضغط على الأمر Statistics تظهر شاشة جديدة بعنوان Linear regression: Statistics.

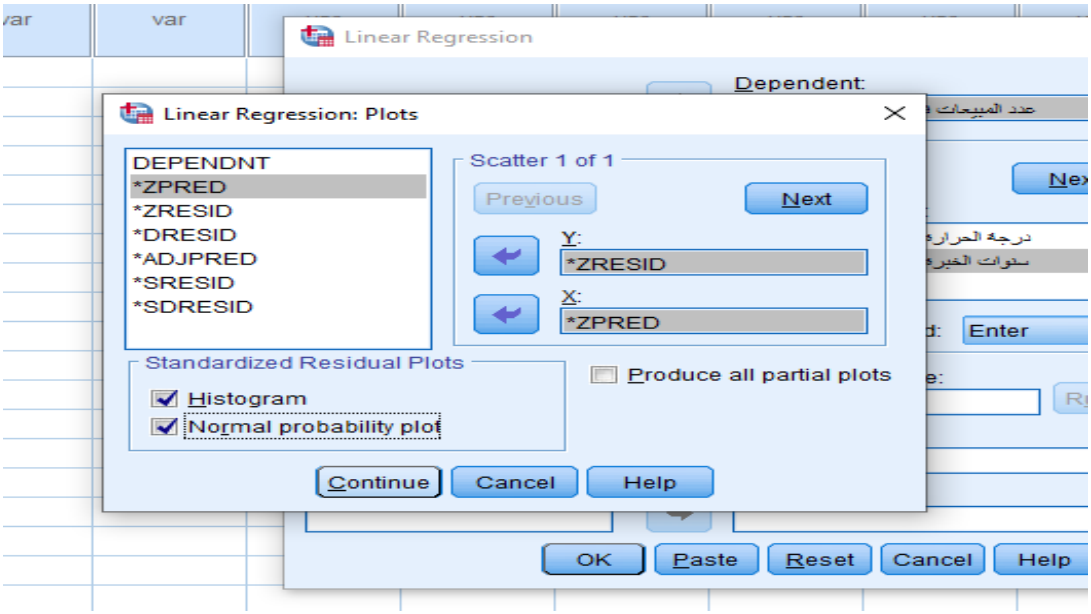


نتأكد أن Model fit, Estimates, Casewise diagnostics مختارة ويمكن أيضا إضافة اختيارات أخرى.

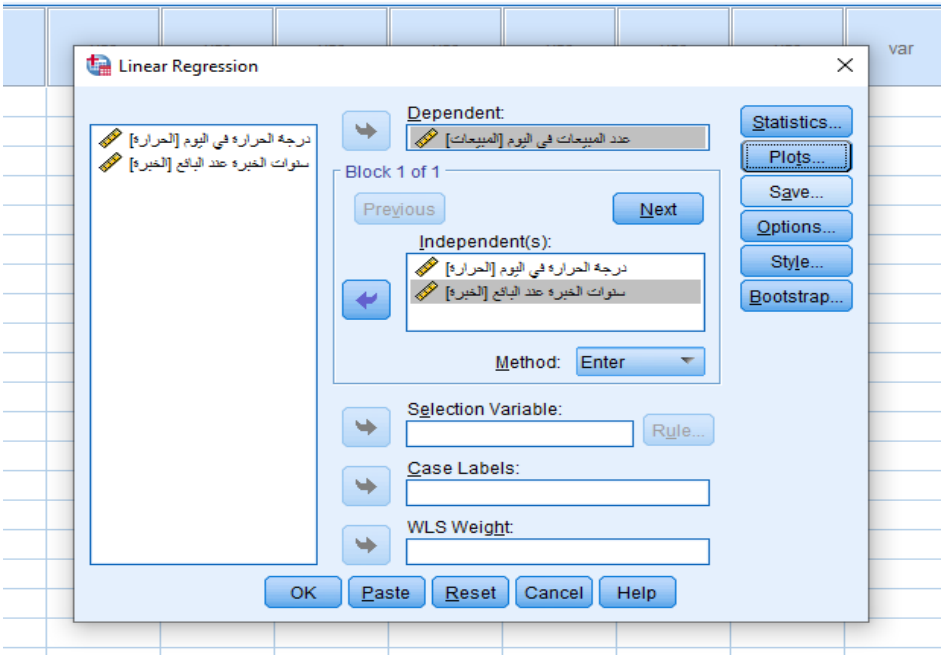
- نختار Continue لنعود للشاشة السابقة.
- نضغط على الأمر Plot فتظهر شاشة جديدة بعنوان Linear regression: Plots.



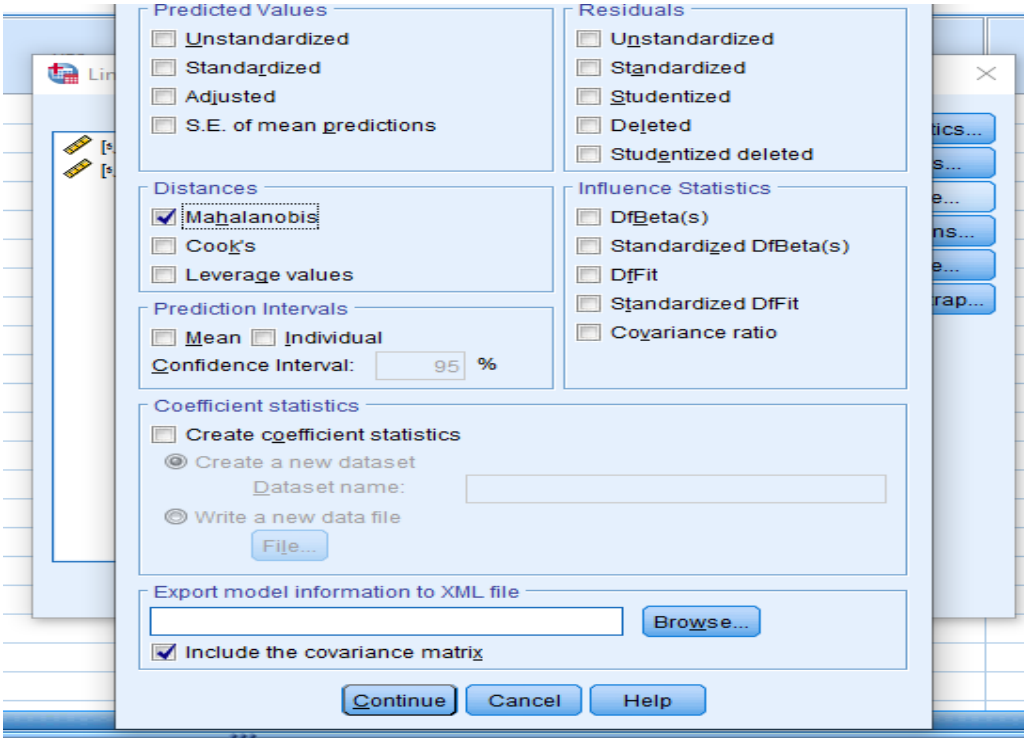
- ننقل * ZRESID للمستطيل المقابل لـ: Y . وايضا* ZPRED للمستطيل المقابل :X.
- من قائمة Plots Residual Standardized نختار كلا من Histogram و Normal probability plot.



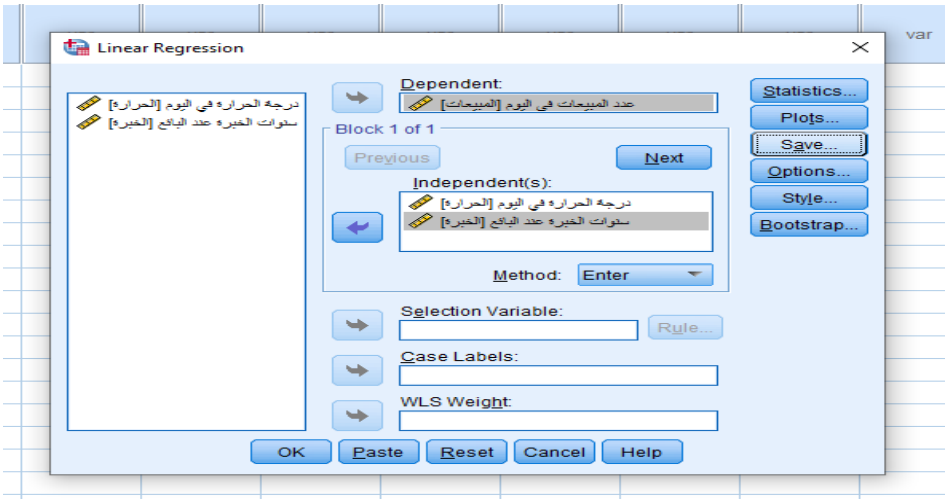
- نضغط على Continue فنعود للشاشة السابقة.



- نضغط على الأمر Save تطهر شاشة جديدة بعنوان Linear regression: Save.



- من الأمر Distances نختار Mahalanobis.
- نضغط على Continue فنعود للشاشة السابقة.



- نضغط على OK فنحصل على النتائج التالية.

Descriptive Statistics

	Mean	Std. Deviation	N
اليوم في المبيعات عدد	37.27	14.499	15
اليوم في الحرارة درجة	30.33	8.217	15
البائع عند الخبرة سنوات	2.9333	1.57963	15

الجدول الأول: بعنوان Descriptive Statistics ويعطي لنا المتوسط والانحراف المعياري وعدد الحالات لكل متغير على حده.

Correlations

		اليوم في المبيعات عدد	اليوم في الحرارة درجة	البائع عند الخبرة سنوات
Pearson Correlation	اليوم في المبيعات عدد	1.000	.907	.849
	اليوم في الحرارة درجة	.907	1.000	.888
	البائع عند الخبرة سنوات	.849	.888	1.000
Sig. (1-tailed)	اليوم في المبيعات عدد	.	.000	.000
	اليوم في الحرارة درجة	.000	.	.000
	البائع عند الخبرة سنوات	.000	.000	.
N	اليوم في المبيعات عدد	15	15	15
	اليوم في الحرارة درجة	15	15	15
	البائع عند الخبرة سنوات	15	15	15

الجدول الثاني: بعنوان Correlations وهي مصفوفة الارتباط بين جميع المتغيرات وأيضا معنوية الارتباط ونلاحظ أنه لا يوجد ارتباط تام بين المتغيرات المستقلة وبعضها الآخر.

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	درجة ,البائع عند الخبرة سنوات اليوم في الحرارة[b]	.	Enter

a. Dependent Variable: اليوم في المبيعات عدد

b. All requested variables entered.

الجدول الثالث: بعنوان Variables Removed/Entered ويحتوي على أسماء المتغيرات التي دخلت في معادلة الانحدار وهما متغيرين (**الحرارة** و **سنوات الخبرة**) والمتغيرات التي استبعدت من الدخول في المعادلة وهنا في الطريقة العيارية لا تستبعد متغيرات.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.912[a]	.832	.804	6.426

a. Predictors: (Constant), اليوم في الحرارة درجة ,البائع عند الخبرة سنوات

b. Dependent Variable: اليوم في المبيعات عدد

الجدول الرابع: بعنوان Summary Model يحتوي على بعض المقاييس التي تم حسابها للنموذج المقدر وهي:

- قيمة معامل الارتباط R=0.912 وهو عالي جدا.
- مربع معامل الارتباط ويستخدم تعين مدى البيانات المستخدمة من المتغيرات المستقلة في تقدير المتغير التابع ونلاحظ أن النموذج المقدر يعبر عن (المتغيرين المستقلين معا) 83 % من البيانات وزيادة قيمة هذا المقياس يفسر أن النموذج المقترح ملائم .
- تعيين مربع معامل الارتباط المعدل Adjusted R Square ويستخدم لنفس الغرض السابق ولكنه أدق.
- تعيين خطأ التقدير Std. Error of the estimate وهو هنا 6.42616 كلما قل دل على خطأ أقل للنموذج.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2447.388	2	1223.694	29.633	.000[b]
	Residual	495.546	12	41.295		
	Total	2942.933	14			

a. Dependent Variable: اليوم في المبيعات عدد

b. Predictors: (Constant), اليوم في الحرارة درجة ,البائع عند الخبرة سنوات

الجدول الخامس: يحتوي على نتائج تحليل التباين ANOVA لاختبار معنويه الانحدار

الفرض الصفري: الانحدار غير معنوي (لا يختلف عن الصفر)

الفرض البديل: الانحدار معنوي (يختلف عن الصفر)

ومن جدول ANOVA نجد ان Sig = .000 وهي أقل من مستوى المعنوية 0.05 لذا سوف نرفض فرض العدم ونقبل الفرض البديل وهي أن الانحدار معنوي وبالتالي توجد علاقة بين المتغيرات المستقلة والمتغير التابع.

ونتيجة قبولنا الفرض البديل والذي ينص على معنويه الانحدار هو معامل واحد على الاقل من المعاملات الثلاثة يمكن أن يكون معنوي. ويمكن تحديد العامل السبب في المعنوية.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound
1	(Constant)	-6.993-	8.428		-.830-	.423	-25.356-	11.371
	في الحرارة درجة اليوم	1.275	.454	.722	2.807	.016	.285	2.264
	عند الخبرة سنوات البائع	1.907	2.362	.208	.807	.435	-3.240-	7.054

a. Dependent Variable: عدد المبيعات في اليوم

الجدول السادس: بعنوان Coefficients ويساعد هذا الجدول في الحصول على ما يلي:

1. معادلة خط الانحدار المقدرة والخطأ في التقدير لكل معامل وذلك من العمود Coefficients Unstandarized حيث:

المبيعات = – 6.993 + 1.275 الحرارة + 1.907 سنوات الخبرة

2. لتعيين أي من المعاملات يكون معنويا وسببا في معنوية تحليل التباين للانحدار ننظر إلى العمود الثالث من اليمين الذى يعطى قيمة Sig. لاختبار معنوية كل معامل على حده نجد انه في حاله سنوات الخبرة والثابت Constant

قيمة sig اكبر من 0.05 لذا فان الثابت ومعامل السنوات غير معنوي لكن في حاله درجات الحرارة sig = .016.0 أقل من 0.05 لذا فان معامل درجات الحرارة معنوي وهو سبب معنوية تحليل التباين للانحدار.

3. العمود الأخير يقدم التقدير بفترة لمعاملات خط الانحدار والثابت كلا على حده.

Coefficient Correlations[a]

Model			البائع عند الخبرة سنوات	اليوم في الحرارة درجة
1	Correlations	البائع عند الخبرة سنوات	1.000	-.888-
		اليوم في الحرارة درجة	-.888-	1.000
	Covariances	البائع عند الخبرة سنوات	5.581	-.952-
		اليوم في الحرارة درجة	-.952-	.206

a. Dependent Variable: اليوم في المبيعات عدد

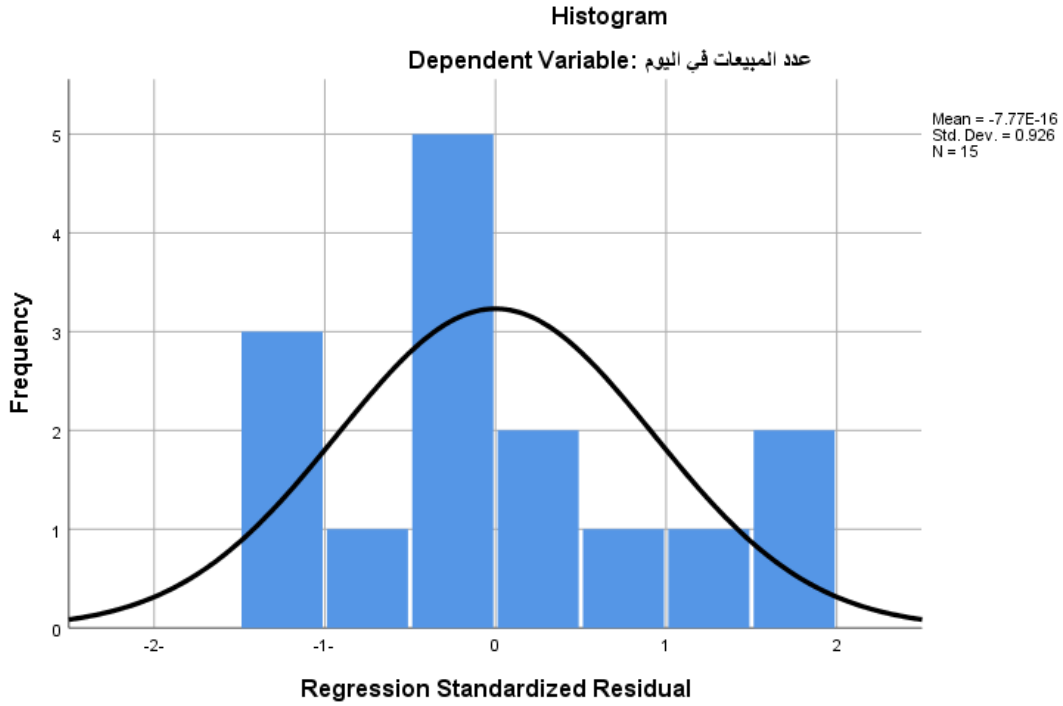
الجدول السابع: بعنوان Correlations Coefficient ويعطى مصفوفة الارتباط بين المتغيرات المستقلة ونلاحظ كما سبق انه لا يوجد ارتباط تام بين المتغيرات وبعضها الآخر.

Residuals Statistics[a]

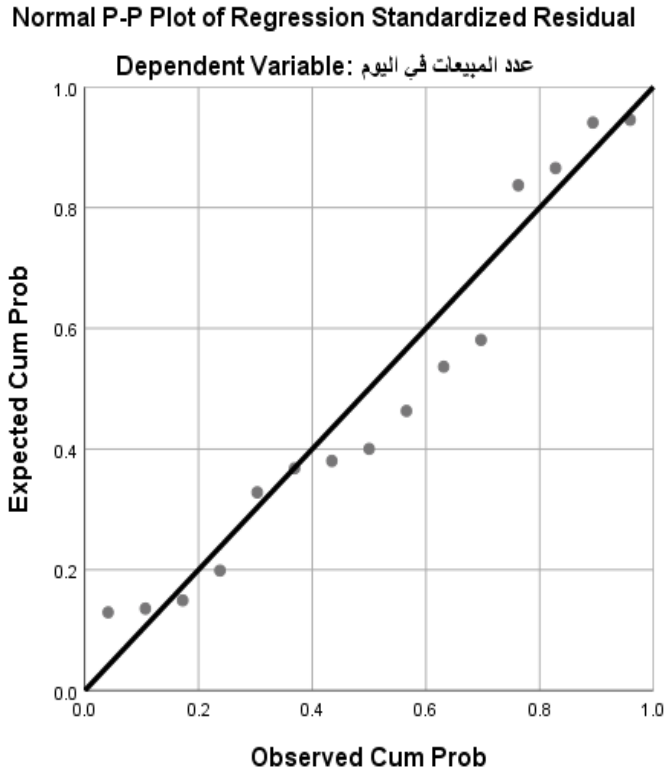
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	17.86	55.44	37.27	13.222	15
Std. Predicted Value	-1.468-	1.374	.000	1.000	15
Standard Error of Predicted Value	1.945	4.247	2.801	.667	15
Adjusted Predicted Value	18.72	59.65	37.52	13.420	15
Residual	-7.255-	10.294	.000	5.949	15
Std. Residual	-1.129-	1.602	.000	.926	15
Stud. Residual	-1.257-	1.826	-.017-	1.034	15
Deleted Residual	-9.654-	13.374	-.254-	7.483	15
Stud. Deleted Residual	-1.291-	2.057	.008	1.090	15
Mahal. Distance	.349	5.182	1.867	1.347	15
Cook's Distance	.000	.333	.090	.113	15
Centered Leverage Value	.025	.370	.133	.096	15

a. Dependent Variable: اليوم في المبيعات عدد

الجدول الثامن: بعنوان Statistics Residuals يستخدم لمعرفه بعض المقاييس الخاصة بالبواقي.

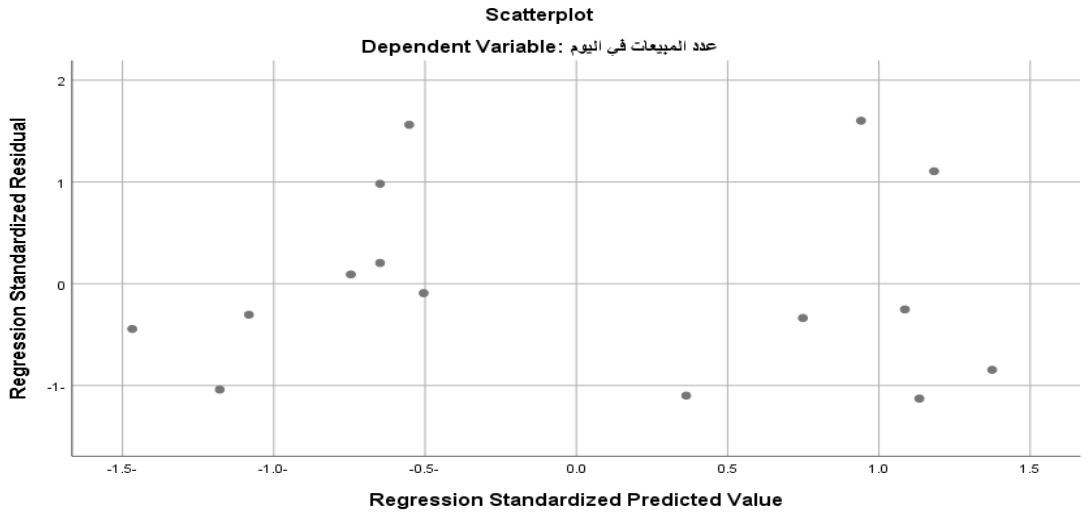


الشكل البياني 1: هو المدرج التكراري ويستخدم للتعرف هل البيانات تتوزع حسب التوزيع الطبيعي أم لا؟

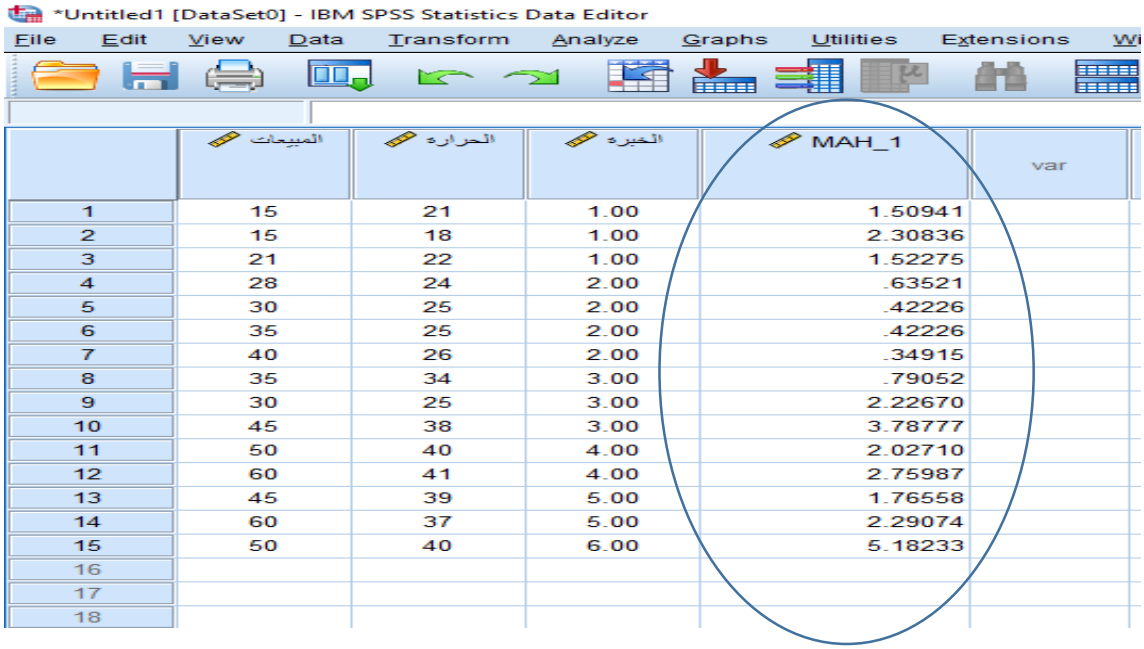


الشكل البياني 2: يختبر هل البواقي تتبع التوزيع الطبيعي ام لا؟

ومن الشكل نجد أن النقاط تتجمع حول الخط وبالتالي فان البيانات (البواقي) تتوزع حسب التوزيع الطبيعي.



الشكل البياني 3: يمثل شكل الانتشار للبواقي مع القيم المتوقعة ومنه يتضح عدم وجود نمط معين للنقاط في الشكل وهذا يتسق مع شرط الخطية.



	المبيعات	الحراره	الخبره	MAH_1	var
1	15	21	1.00	1.50941	
2	15	18	1.00	2.30836	
3	21	22	1.00	1.52275	
4	28	24	2.00	.63521	
5	30	25	2.00	.42226	
6	35	25	2.00	.42226	
7	40	26	2.00	.34915	
8	35	34	3.00	.79052	
9	30	25	3.00	2.22670	
10	45	38	3.00	3.78777	
11	50	40	4.00	2.02710	
12	60	41	4.00	2.75987	
13	45	39	5.00	1.76558	
14	60	37	5.00	2.29074	
15	50	40	6.00	5.18233	
16					
17					
18					

بالعودة لملف البيانات نجد انه قد أضيف متغير جديد mah_1 وذلك لأننا طلبنا اختبار Mahalanobis فنقوم بمقارنة قيم هذا المتغير بقيمة Chi- Square عند درجة حريه n-1 = 2

ومستوى معنويه مثلا 0.001 فنجد أن Chi- Square =13.8 وجميع قيم المتغير أقل من هذه القيمة لذا فانه لا يوجد قيم متطرفة متعددة.

الانحدار التدريجي

Stepwise Regression

مما سبق في الانحدار العياري يتضح أنه إذا زاد عدد المتغيرات المستقلة عن عدد معين فان هذا يؤدى الى ظهور العديد من المشاكل عند معالجة مشكله الانحدار.

- هناك شرط يحدد العلاقة بين عدد الحالات وعدد المتغيرات المستقلة فاذا لم نحقق هذا الشرط فان النتائج والتقديرات ستكون غير سليمة.

- زيادة عدد المتغيرات المستقلة عن عدد الحالات يقلل أيضا من درجات حرية الخطأ في اختبار تحليل التباين ومعه قد تصل درجات الحرية إلى الصفر ويستحيل معه بعد ذلك اجراء اي اختبار لمعنوية الانحدار.

- ادخال عدد كبير من المتغيرات المستقلة يؤدى ايضا الى فقدان القدرة على تحقيق شروط تطبيق الانحدار (الارتباط الذاتي والخطية والتجانس،).
- لمعالجة هذه المشكلة نستخدم الانحدار التدريجي والذي يسميه البعض (خطوه خطوه) وذلك للتحكم في عدد المتغيرات التي تدخل في معادلة الانحدار.
- يهدف الانحدار التدريجي أساسا الى ايجاد علاقة بين المتغير التابع والمتغيرات المستقلة الاكثر ارتباطا به ويتم ذلك تدريجيا.
- يبدأ الانحدار التدريجي بالخطوات التالية:

1- نحسب مصفوفة الانحدار لجميع المتغيرات

2- نختار المتغير المستقل الذي له أكبر ارتباط بالمتغير التابع وندخله في معادله الانحدار.

3- نختار المتغير المستقل الثاني الذي له أكبر ارتباط بعد المتغير الذي دخل المعادلة فنبحث اولا هل هناك ارتباط كبير بينه وبين المتغير الذي اختير إذا كانت الإجابة بنعم يستبعد ذلك المتغير وإذا كانت بلا ندخله في الاختيار.

4- نكرر هذه العملية مع بقية المتغيرات المستقلة مع استبعاد المتغيرات التي لها ارتباط كبير مع المتغيرات المختارة

5- تكون عملية الإضافة مجديه إذا كان هناك تأثير على معامل التحديد Determination of Coefficient (مربع معامل الارتباط) وكذلك قيمه F من جدول تحليل التباين

6- نتوقف عن الإضافة إذا لم يكن للإضافة تأثير على معامل التحديد وقيمة F المحسوبة (أو له تأثير ضعيف).

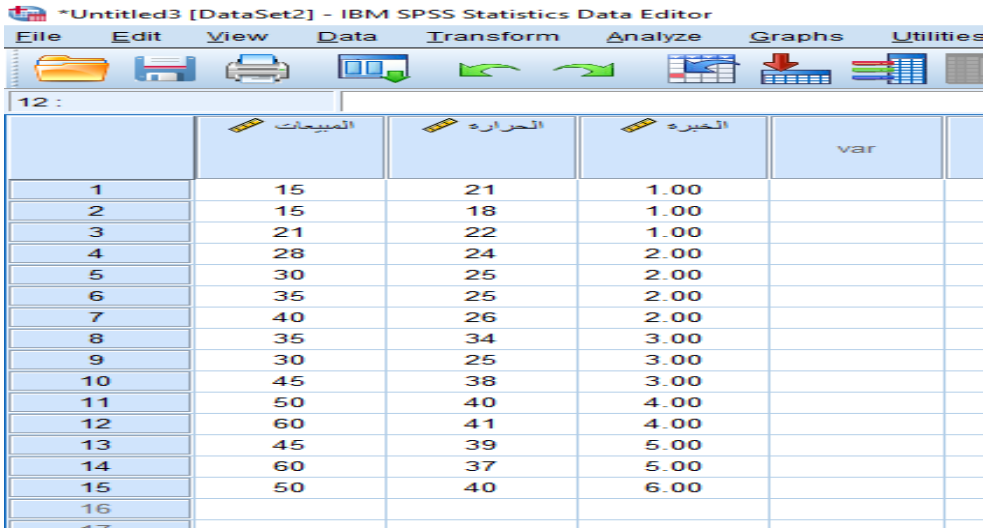
مما سبق يمكن القول وبصفه عامه فان الانحدار التدريجي يحقق بعض المزايا منها:

1- تقليل عدد المتغيرات المستقلة الداخلة في النموذج عندما لا يتلاءم عدد الحالات مع عدد المتغيرات المستقلة.

2- التخلص من الازدواج الخطى بين المتغيرات المستقلة في النموذج المقدر. (عبد الفتاح، 40-43)

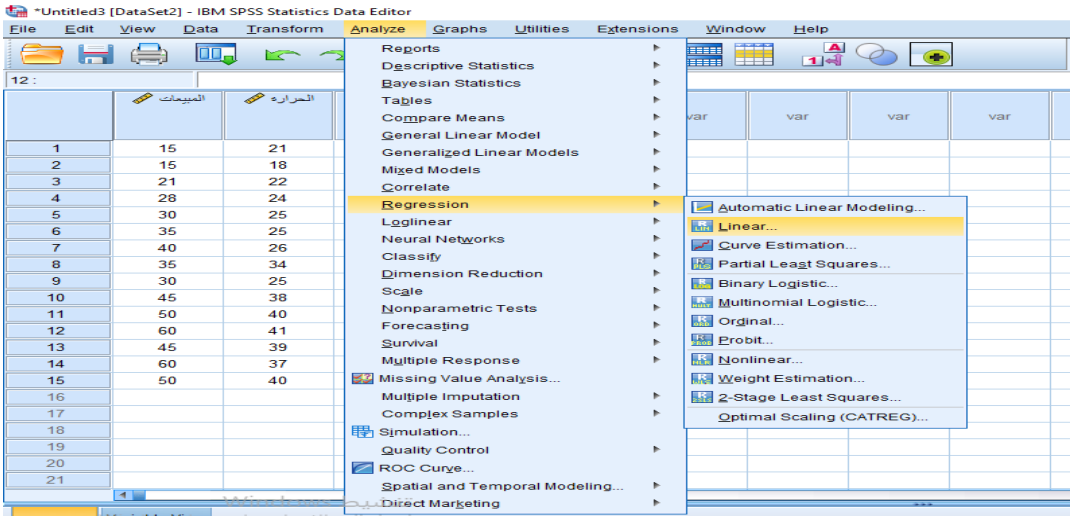
ولإجراء خطوات الانحدار التدريجي باستخدام SPSS نمر بالخطوات التالية:

- نعرف بالمتغيرات وندخل البيانات على صفحة Spss

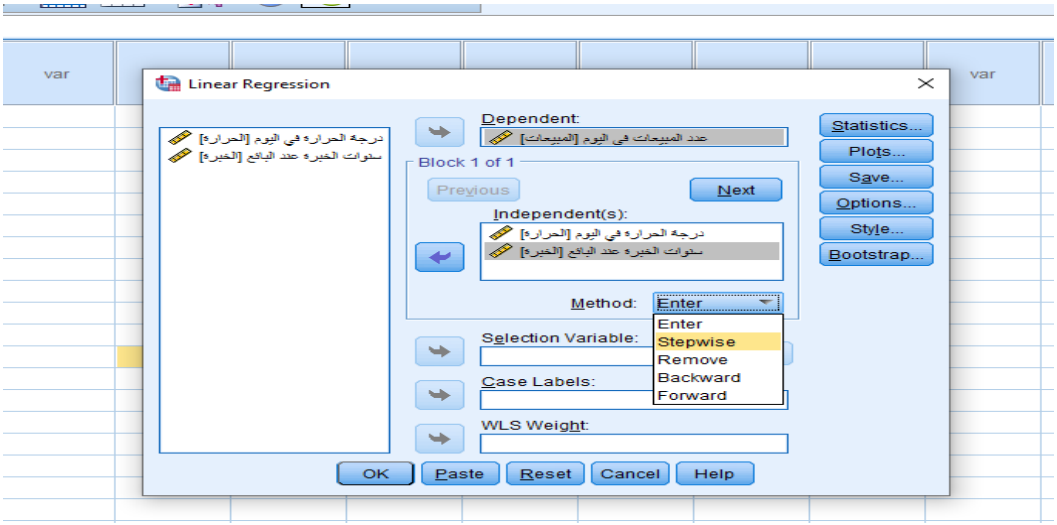


	المبيعات	الحراره	الخبره	var
1	15	21	1.00	
2	15	18	1.00	
3	21	22	1.00	
4	28	24	2.00	
5	30	25	2.00	
6	35	25	2.00	
7	40	26	2.00	
8	35	34	3.00	
9	30	25	3.00	
10	45	38	3.00	
11	50	40	4.00	
12	60	41	4.00	
13	45	39	5.00	
14	60	37	5.00	
15	50	40	6.00	
16				

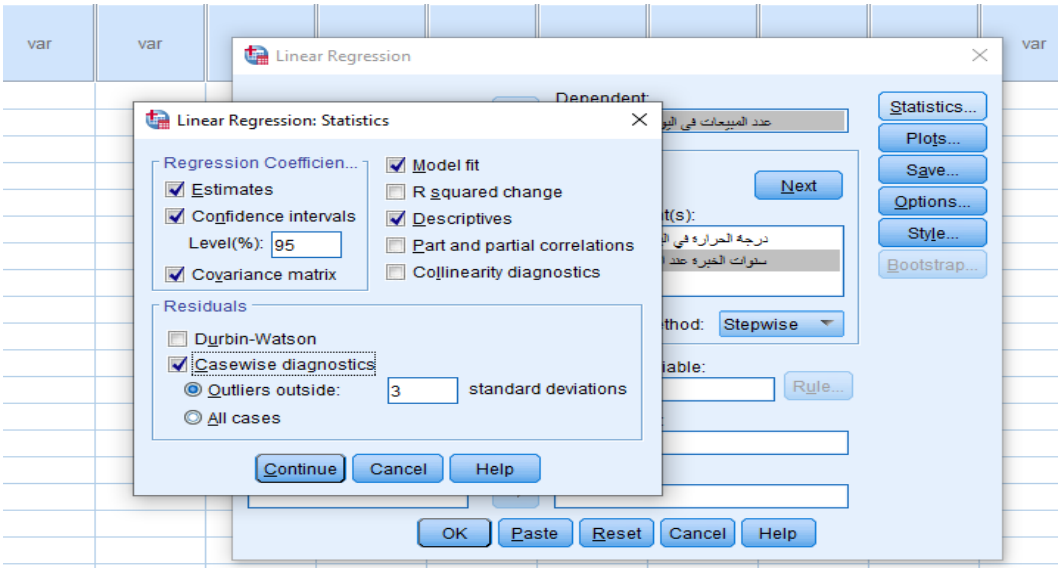
- من قائمة Analyze نختار الامر Regression.
- من القائمة المنسدلة نختار Linear.



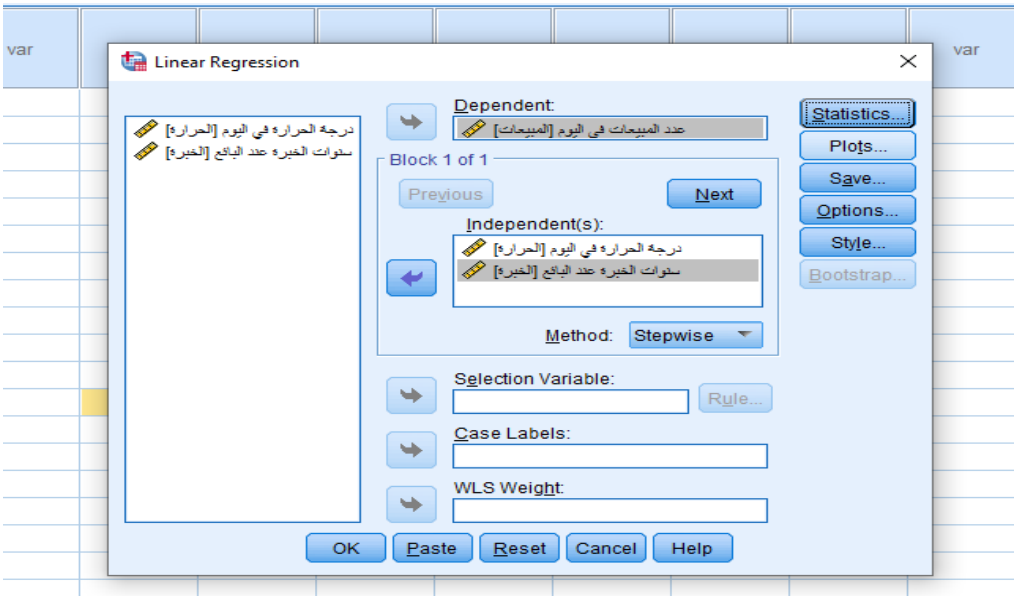
- تظهر شاشه جديده بعنوان Linear regression.



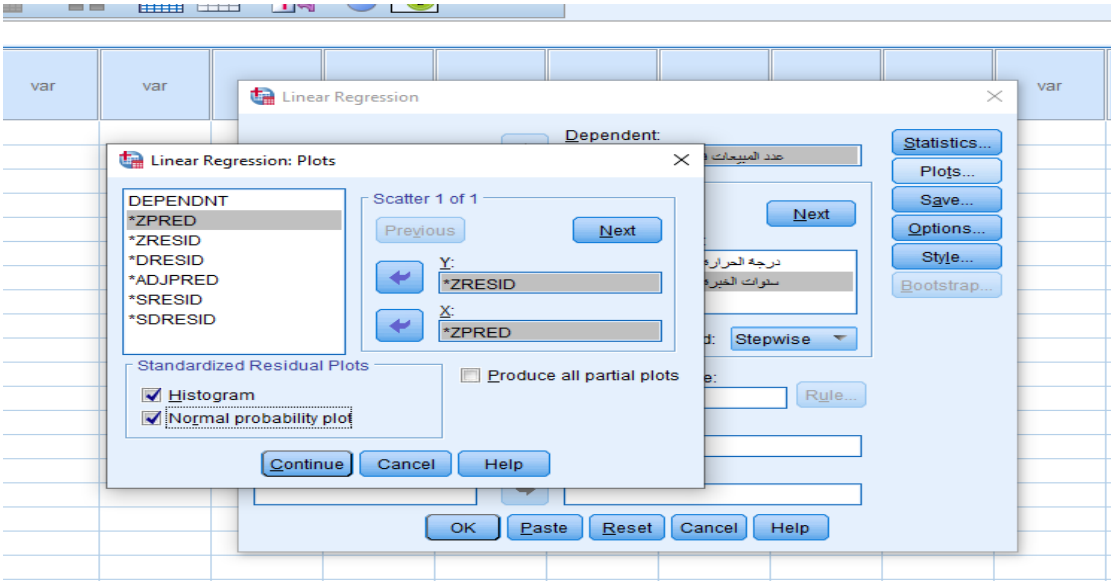
- ننقل المتغير المبيعات لخانة Dependent وننقل المتغيرين المستقلين الحرارة، سنوات الخبرة لخانة Independent.
- من قائمة Method نختار Stepwise .
- نضغط على الامر Statistics تظهر شاشه جديده بعنوان Linear Regression: Statistics.



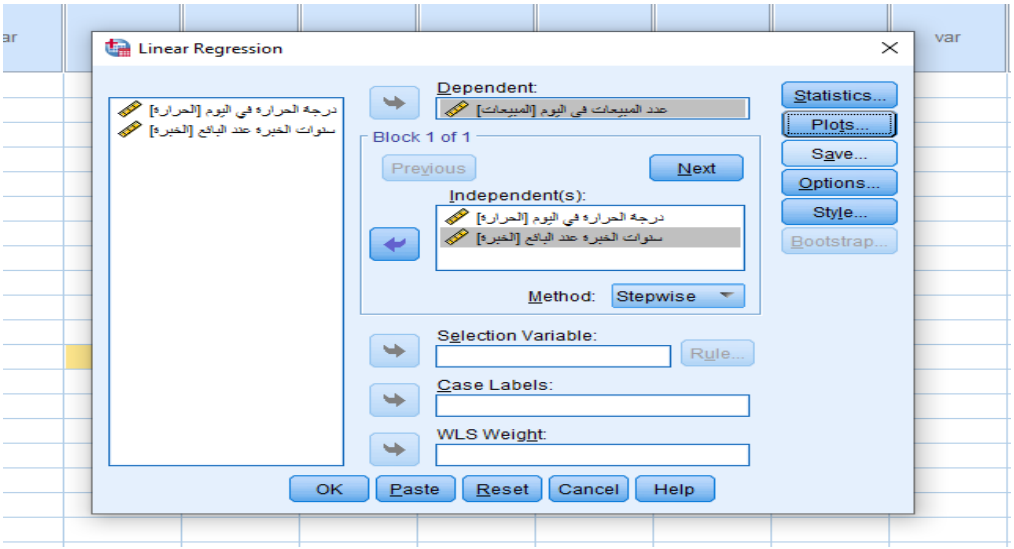
- أن نتأكد Model fit, Estimates, diagnostics Casewise ، مختارة ويمكن أيضا إضافة اختيارات أخرى.
- نختار Continue لنعود للشاشة السابقة.



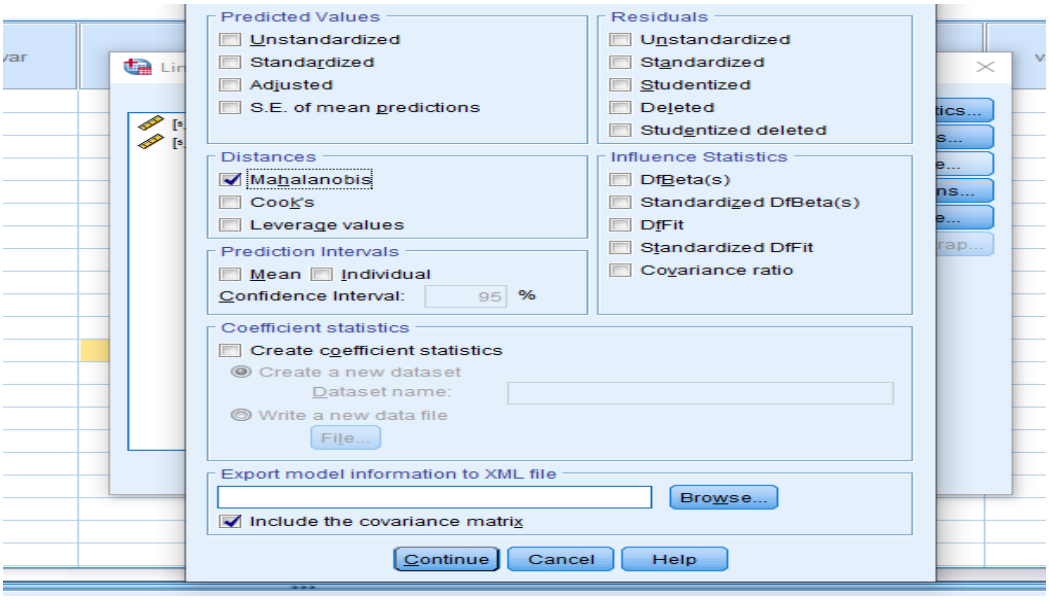
- نضغط على الامر Plot فتظهر شاشه جديده بعنوان Linear Regression: Plots.



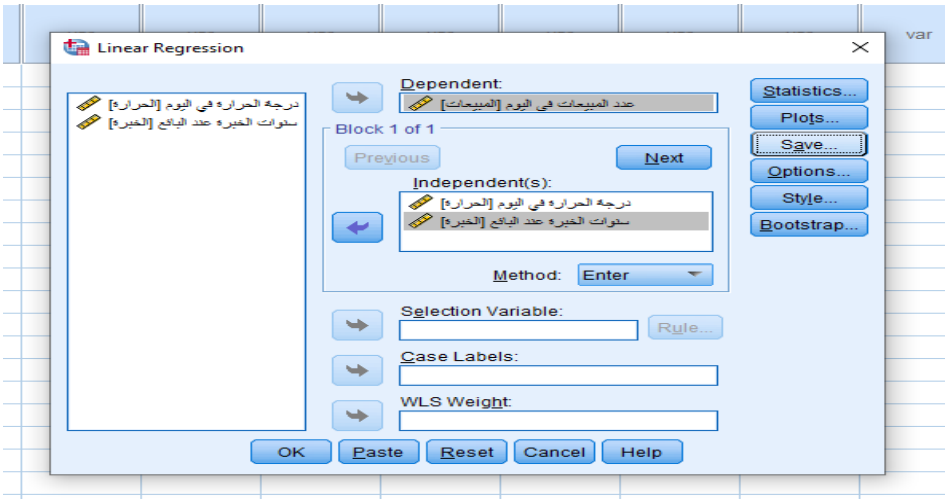
- ننقل ZRESID * للمستطيل المقابل لـ Y: وأيضا ZPRED * للمستطيل المقابل X:
- من قائمة Standardized Residual Plots نختار كلا من Histogram, Normal probability plot
- نضغط على Continue فنعود للشاشة السابقة.



- نضغط على الأمر Save تظهر شاشة جديدة بعنوان Linear Regression: Save.



- من الامر Distances نختار Mahalanobis.
- نضغط على Continue فنعود للشاشة السابقة.



- نضغط على Ok فنحصل على النتائج التالية:

Descriptive Statistics

	Mean	Std. Deviation	N
اليوم في المبيعات عدد	37.27	14.499	15
اليوم في الحرارة درجة	30.33	8.217	15
البائع عند الخبرة سنوات	2.9333	1.57963	15

الجدول الاول: بعنوان Variables Entered/removed ويوضح المتغيرات الداخلة في المعادلة وطريقة المعالجة ويتضح أن المتغير درجة الحرارة هو المتغير الوحيد الذي تم إدخاله في معادلة الانحدار.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.907[a]	.822	.809	6.339

a. Predictors: (Constant), اليوم في الحرارة درجة

b. Dependent Variable: اليوم في المبيعات عدد

الجدول الثاني: بعنوان Model Summary ويعطي بعض المقاييس الهامة والمحسوبة من البيانات وأهمها معامل التوافق (مربع معامل الارتباط) ويستخدم للحكم على عملية التوفيق ومنه نجد أن معادلة الانحدار تمثل 80% من البيانات.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2420.480	1	2420.480	60.228	.000[b]
	Residual	522.453	13	40.189		
	Total	2942.933	14			

a. Dependent Variable: اليوم في المبيعات عدد

b. Predictors: (Constant), اليوم في الحرارة درجة

الجدول الثالث: بعنوان ANOVA وهو تحليل التباين للانحدار ويتضح أن الانحدار معنوي حيث 0 .000 = Sig وهي أقل من 0.05.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound
1	(Constant)	-11.271-	6.465		-1.743-	.105	-25.238-	2.696
	اليوم في الحرارة درجة	1.600	.206	.907	7.761	.000	1.155	2.046

a. Dependent Variable: اليوم في المبيعات عدد

الجدول الرابع: بعنوان Coefficients ومنه يمكن إيجاد معادلة الانحدار بين المبيعات ودرجة الحرارة فقط ومقدار الخطأ في التقدير واختبار معنوية المعاملات والتقدير بفترة للمعاملات.

المبيعات = – 11.271 + 1.600*درجة الحرارة

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics Tolerance
1	البائع عند الخبرة سنوات	.208[b]	.807	.435	.227	.212

a. Dependent Variable: اليوم في المبيعات عدد

b. Predictors in the Model: (Constant), اليوم في الحرارة درجة

الجدول الخامس: بعنوان Excluded Variables ويعرض بيانات تخص المتغير الذي استبعد وهو سنوات الخبرة.

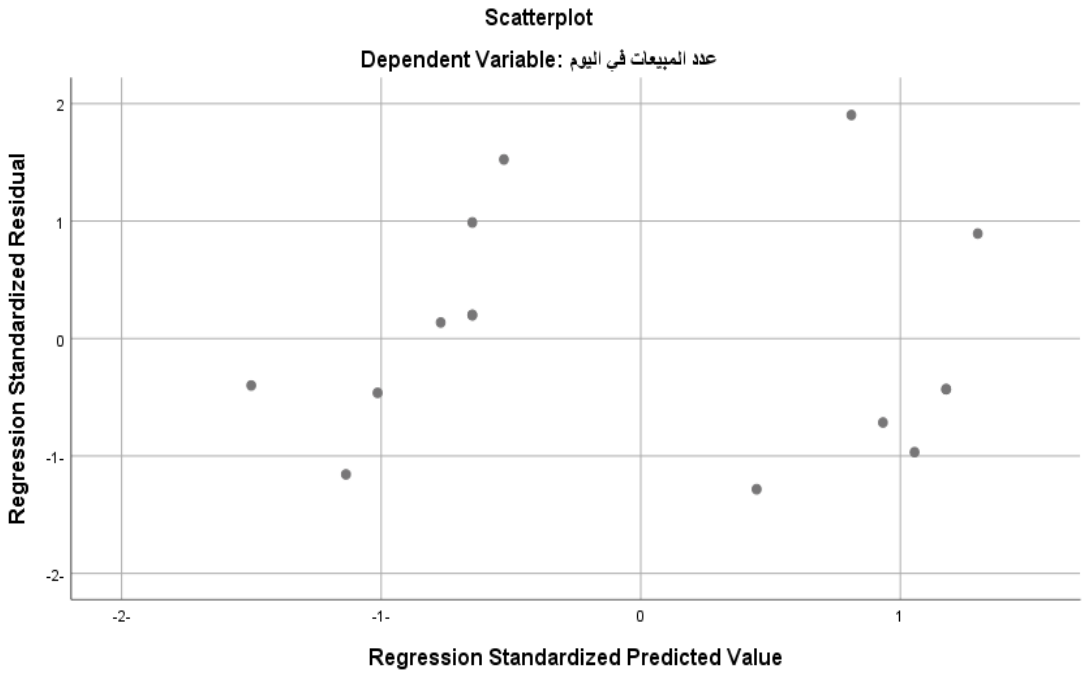
Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	17.53	54.33	37.27	13.149	15
Std. Predicted Value	-1.501-	1.298	.000	1.000	15
Standard Error of Predicted Value	1.803	3.024	2.289	.356	15
Adjusted Predicted Value	18.28	53.28	37.35	13.062	15

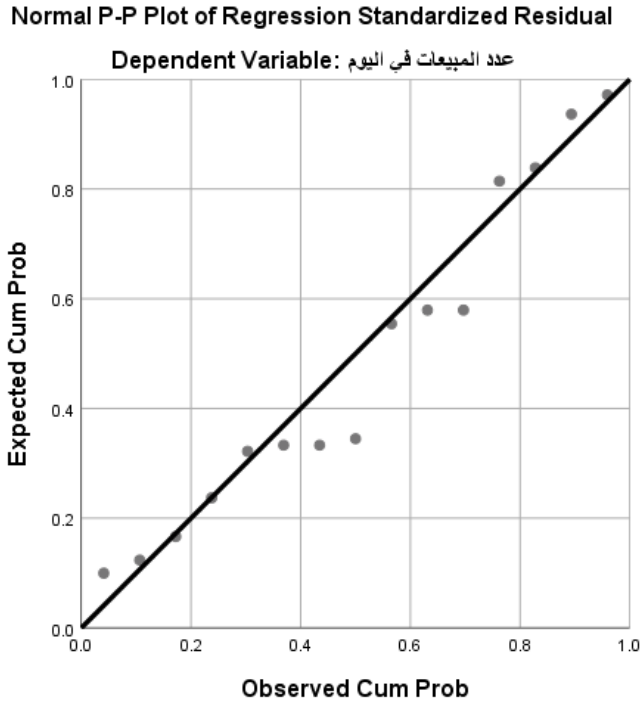
Residual	-8.134-	12.066	.000	6.109	15
Std. Residual	-1.283-	1.903	.000	.964	15
Stud. Residual	-1.338-	2.022	-.006-	1.027	15
Deleted Residual	-8.850-	13.613	-.088-	6.947	15
Stud. Deleted Residual	-1.385-	2.346	.021	1.092	15
Mahal. Distance	.199	2.253	.933	.587	15
Cook's Distance	.001	.262	.068	.072	15
Centered Leverage Value	.014	.161	.067	.042	15

a. Dependent Variable: اليوم في المبيعات عدد

الجدول السادس: بعنوان Residuals Statistics ويعرض بيانات خاصه بتحليل البواقي.

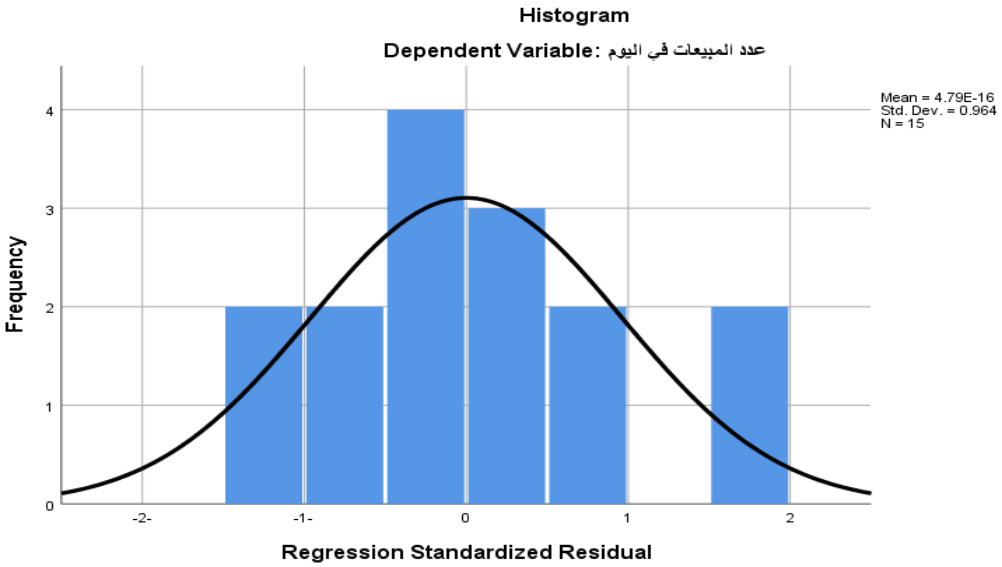


الشكل البياني الاول: يمثل شكل الانتشار للبواقي مع القيم المتوقعة ومنه يتضح عدم وجود نمط معين للنقاط في الشكل وهذا يتسق مع شرط الخطية.



الشكل البياني الثاني: يختبر هل البواقي تتبع التوزيع الطبيعي ام لا؟

ومن الشكل نجد أن النقاط تتجمع حول الخط وبالتالي فان البيانات (البواقي) تتوزع حسب التوزيع الطبيعي.



الشكل البياني الثالث: هو المدرج التكراري ويستخدم للتعرف هل البيانات تتوزع حسب التوزيع الطبيعي أم لا؟

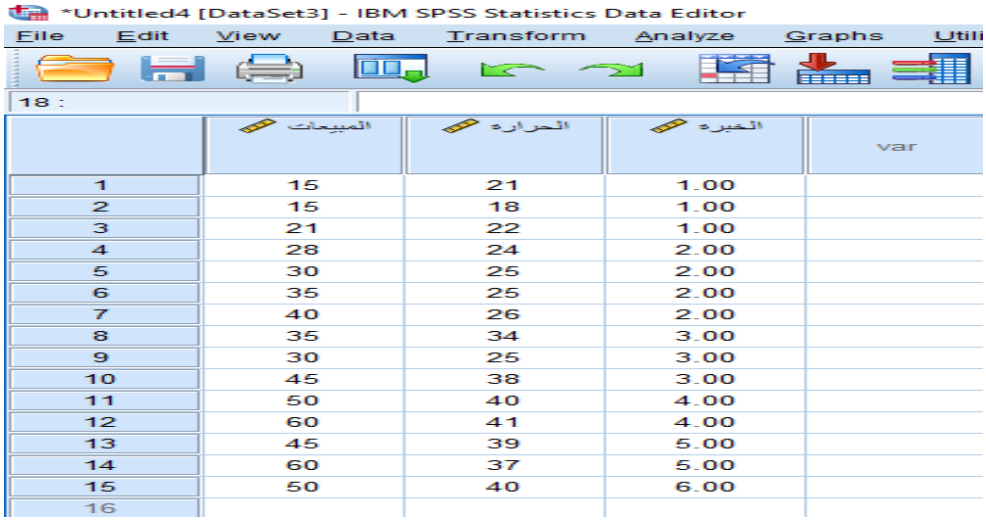
الانحدار الهرمي:

Hierachical Regression

يتم في هذه الحالة استخدام الطريقة العيارية في تحديد معادلة الانحدار ولكن ليس لكل المتغيرات بل ندخل المتغيرات تباعا فندخل أول متغير ثم يليه المتغير وهكذا.

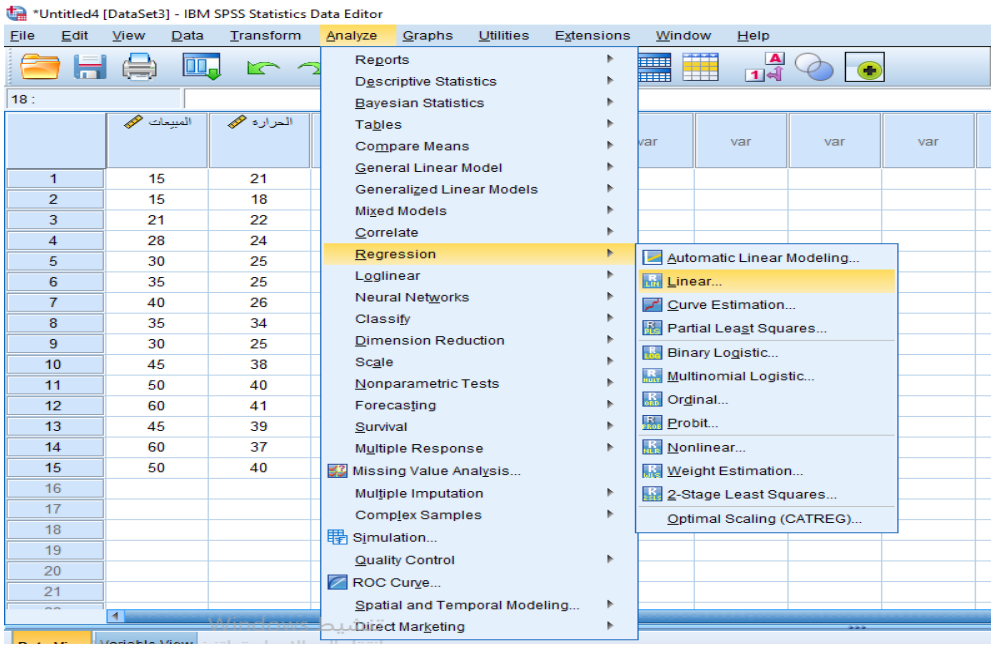
ويمكن تنفيذ ذلك على بيانات الملف تبعا للخطوات التالية:

- نعرف بالمتغيرات وندخل البيانات على صفحة Spss

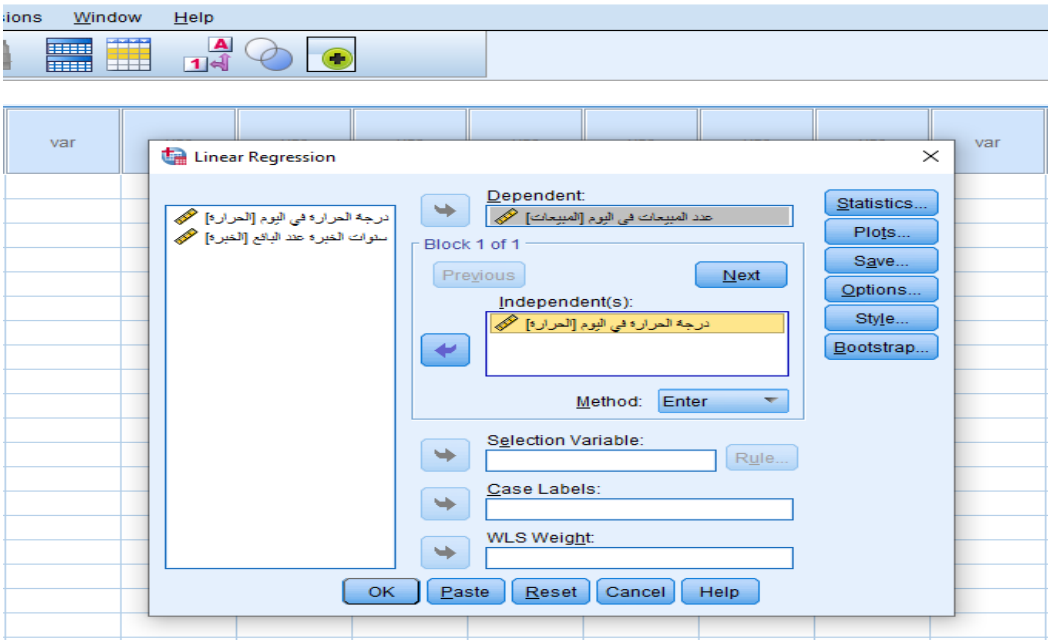


	المبيعات	الحراره	الخبره	var
1	15	21	1.00	
2	15	18	1.00	
3	21	22	1.00	
4	28	24	2.00	
5	30	25	2.00	
6	35	25	2.00	
7	40	26	2.00	
8	35	34	3.00	
9	30	25	3.00	
10	45	38	3.00	
11	50	40	4.00	
12	60	41	4.00	
13	45	39	5.00	
14	60	37	5.00	
15	50	40	6.00	
16				

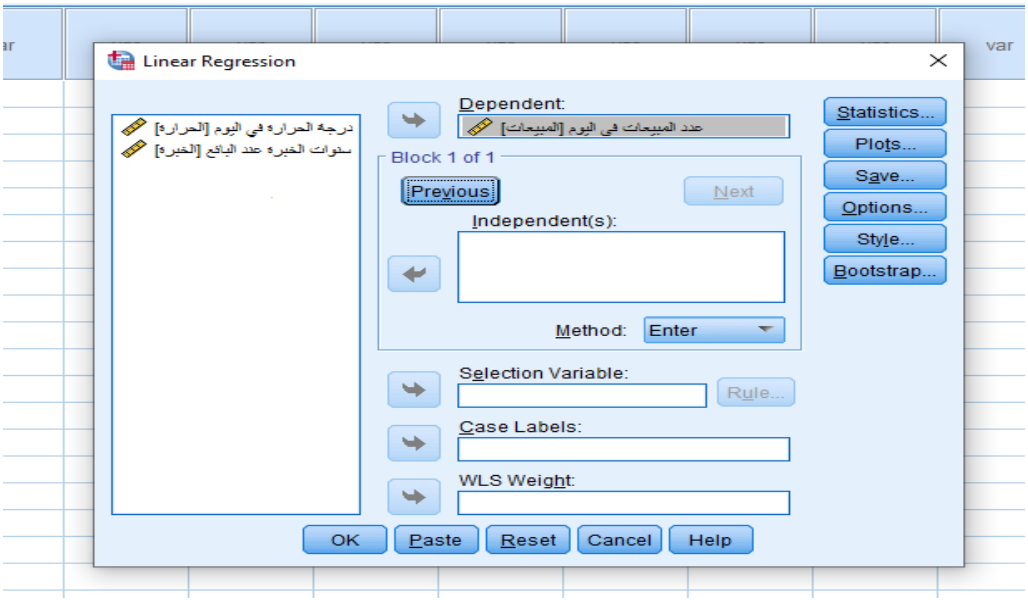
- من قائمة Analyze نختار Regression.
- من القائمة المنسدلة نختار Linear.



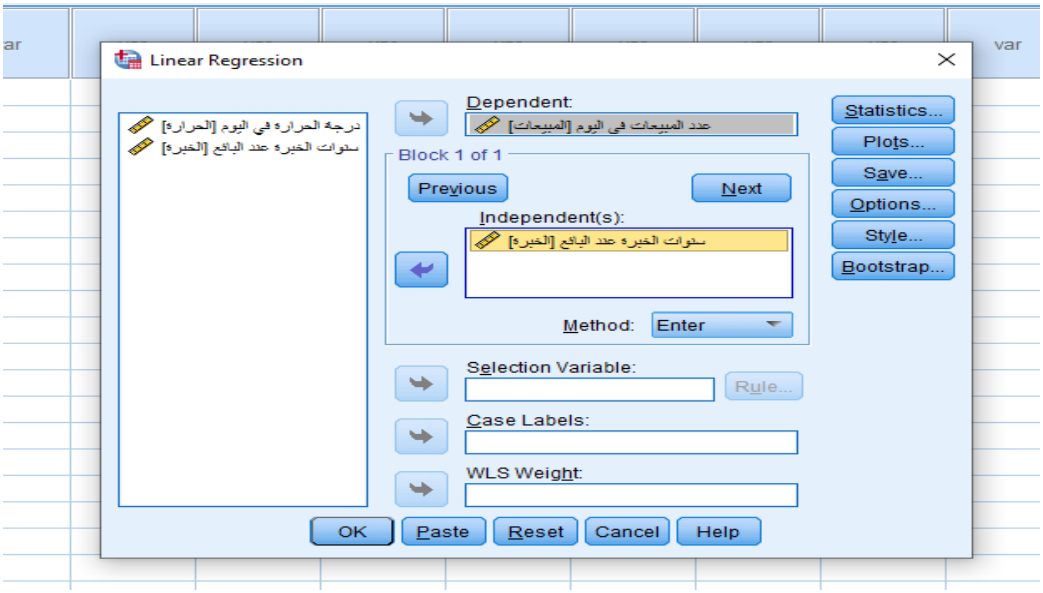
- تظهر شاشة بعنوان Linear Regression.



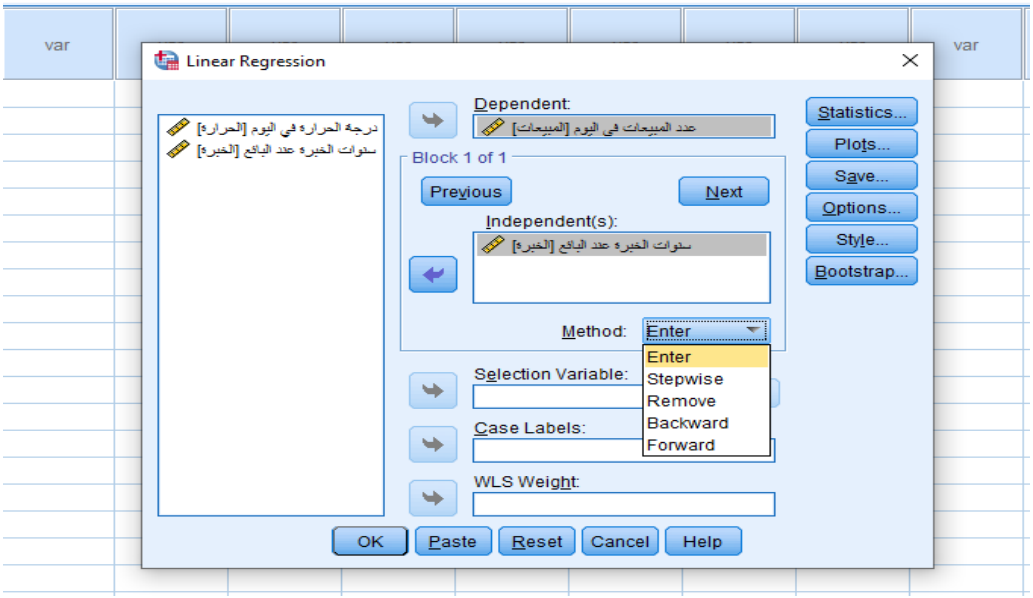
- ننقل المتغير التابع المبيعات لخانة Dependent وننقل أول المتغيرات المستقلة درجة الحرارة لخانة Independent (s) لاحظ فوق كلمة previous وجود Block 2 of 2.
- نضغط على الامر Next.



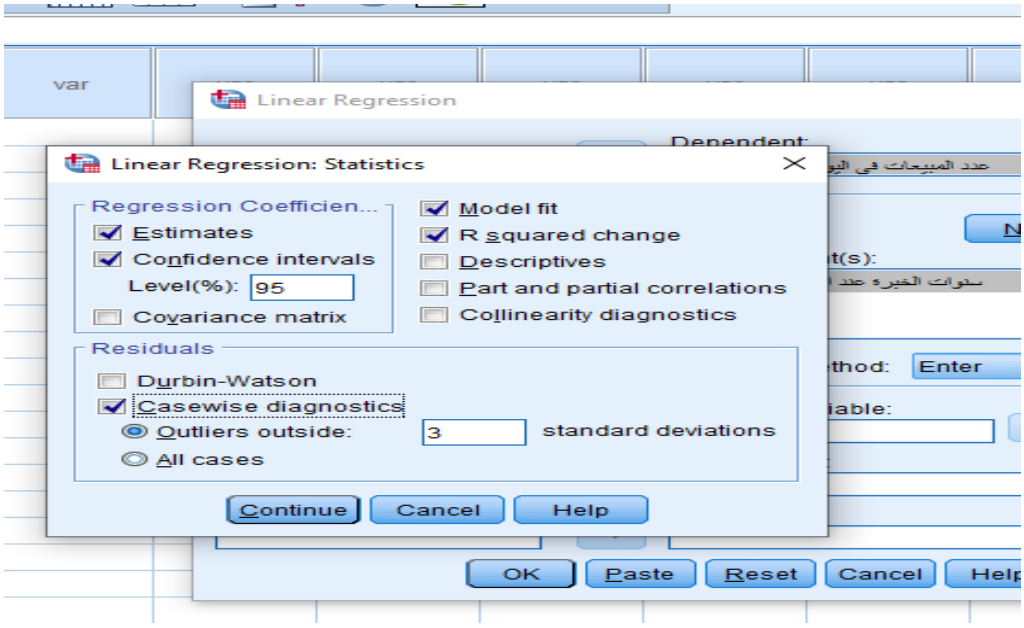
- ننقل المتغير المستقل الثاني السنوات لخانة Independent.



- لاحظ الصندوق الحواري فوق Previous لم تغير لأننا لم نضغط على Next
- سوف نختار الطريقة Method العيارية Enter.



- نضغط على الأمر Statistics تظهر شاشة جديدة بعنوان Linear Regression: Statistics.

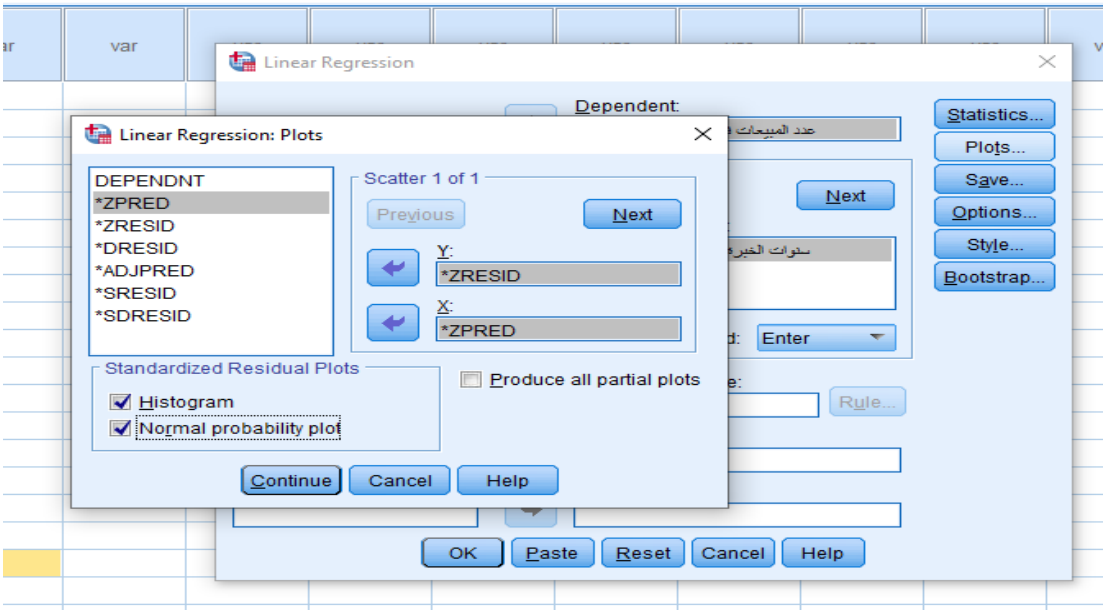


- Model fit, Estimates diagnostics, R Casewise Squared Change، مختاره

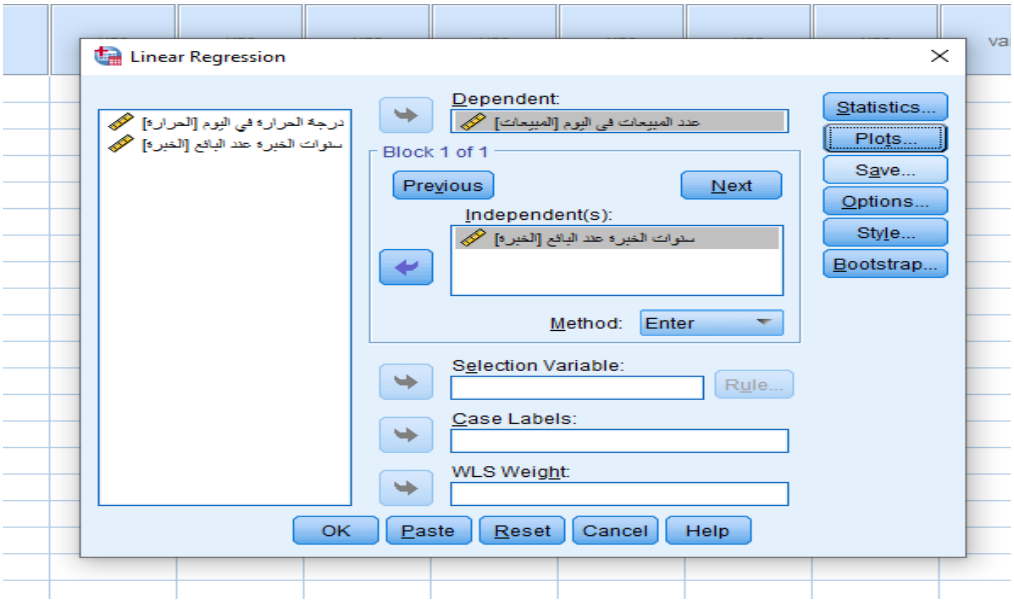
ويمكن أيضا إضافة اختيارات أخرى.

- نختار Continue لنعود للشاشة السابقة.

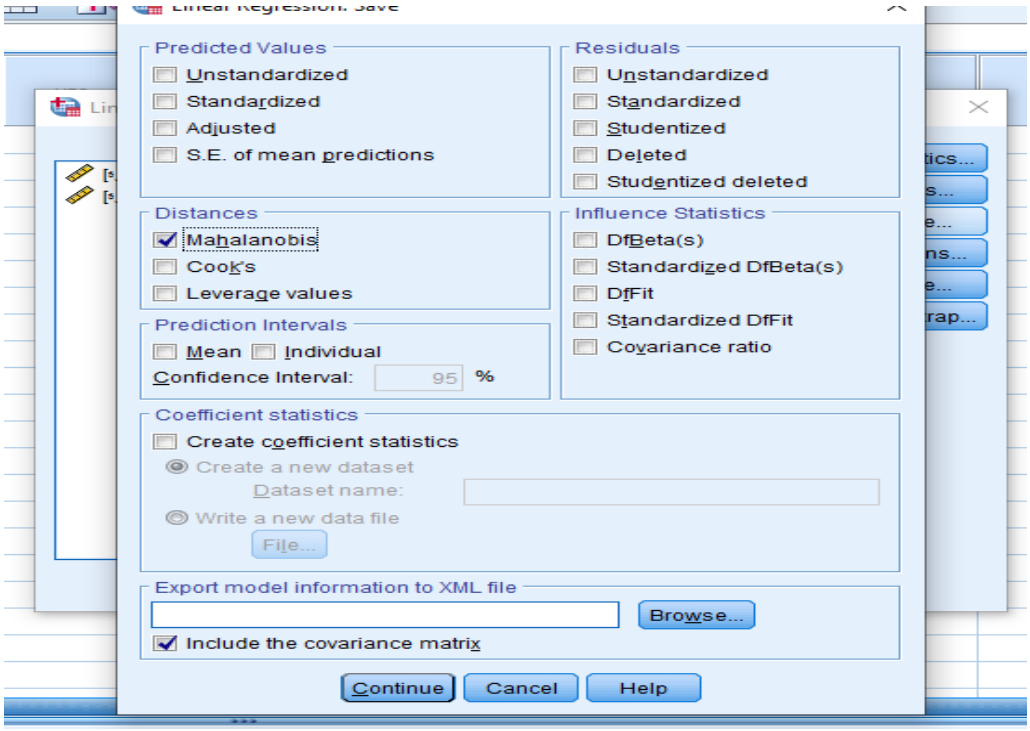
- نضغط على الامر Plot فتظهر شاشه جديده بعنوان Linear Regression: Pltos



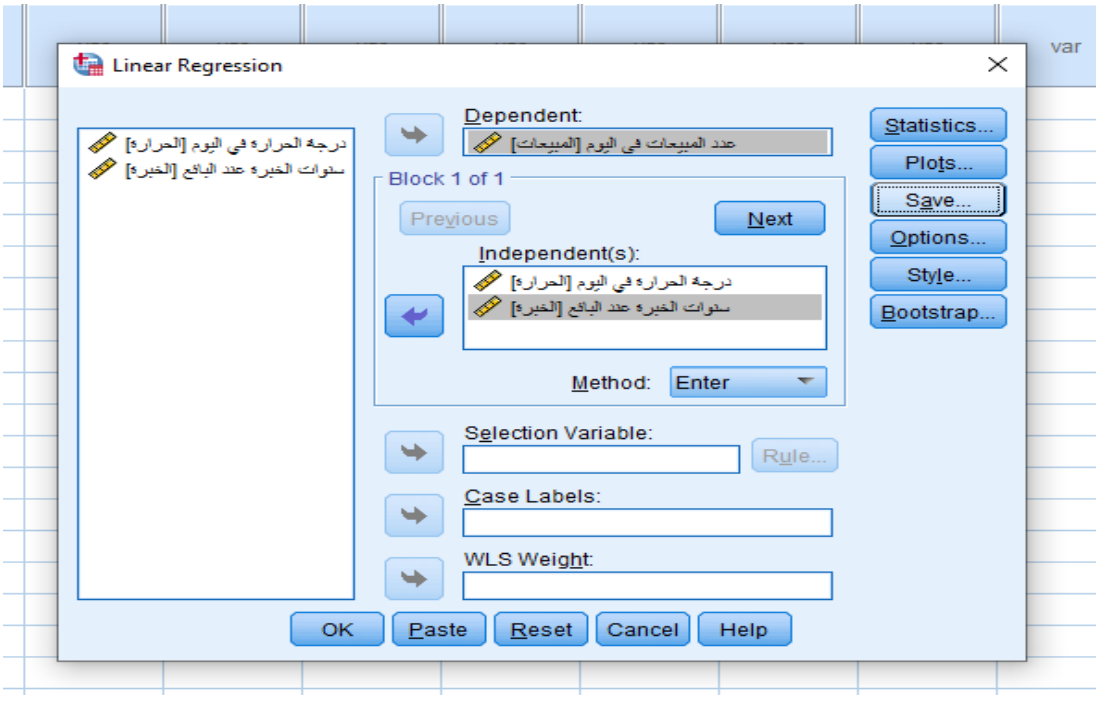
- ننقل ZRESID* للمستطيل المقابل لـ :Y وأيضا ZPRED* للمستطيل المقابل :X.
- من قائمة Standardized Residual Plots نختار كلا من ، Histogram .Normal probability plot
- نضغط على Continue فنعود للشاشة السابقة.



- نضغط على الامر Save تظهر شاشة جديده بعنوان Linear Regression: Save.



- من الامر Distances نختار Mahalanobis.
- نضغط على Continue فنعود للشاشة السابقة.



- نضغط على Ok فنحصل على النتائج التالية:

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	اليوم في الحرارة درجة[b]	.	Enter
2	البائع عند الخبرة سنوات[b]	.	Enter

a. Dependent Variable: اليوم في المبيعات عدد

b. All requested variables entered.

الجدول الاول: بعنوان Variables Entered/Removed يوضح المتغيرات التي أدخلت للنموذج وطريقة الاختيار.

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.907[a]	.822	.809	6.339	.822	60.228	1	13	.000
2	.912[b]	.832	.804	6.426	.009	.652	1	12	.435

a. Predictors: (Constant), اليوم في الحرارة درجة

b. Predictors: (Constant), البائع عند الخبرة سنوات, اليوم في الحرارة درجة

c. Dependent Variable: اليوم في المبيعات عدد

الجدول الثاني: بعنوان Model Summary ويعطي ملخص عن النموذج لاحظ كلا من R. ,F, Sig Change square نجد أن اضافة المتغير الأول درجة الحرارة معنوي وله جدوى بخلاف المتغير الثاني السنوات.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2420.480	1	2420.480	60.228	.000[b]
	Residual	522.453	13	40.189		
	Total	2942.933	14			
2	Regression	2447.388	2	1223.694	29.633	.000[c]
	Residual	495.546	12	41.295		
	Total	2942.933	14			

a. Dependent Variable: اليوم في المبيعات عدد

b. Predictors: (Constant), اليوم في الحرارة درجة

c. Predictors: (Constant), البائع عند الخبرة سنوات, اليوم في الحرارة درجة

الجدول الثالث: بعنوان ANOVA ويعطي تحليل التباين لكل خطوة إدخال ومنه يتضح معنوية الانحدار في كل خطوة.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound
1	(Constant)	-11.271-	6.465		-1.743-	.105	-25.238-	2.696
	اليوم في الحرارة درجة	1.600	.206	.907	7.761	.000	1.155	2.046
2	(Constant)	-6.993-	8.428		-.830-	.423	-25.356-	11.371
	اليوم في الحرارة درجة	1.275	.454	.722	2.807	.016	.285	2.264
	البائع عند الخبرة سنوات	1.907	2.362	.208	.807	.435	-3.240-	7.054

a. Dependent Variable: اليوم في المبيعات عدد

الجدول الرابع: بعنوان Coefficients ومنه يحسب خط الانحدار المقترح (المقدر) واختبار معنوية المعاملات وخطأ التقدير. ومن الجدول نجد أن:

- الخطوة الأولى: تم ادخال درجات الحرارة وكان الانحدار معنوي.
- الخطوة الثانية: تم ادخال سنوات الخبرة مع درجات الحرارة فكانت غير معنوية.

وعلى ذلك يكون أفضل خط مقدر هو:

المبيعات = – 11.271 + 1.600*درجة الحرارة

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics Tolerance
1	البائع عند الخبرة سنوات	.208[b]	.807	.435	.227	.212

a. Dependent Variable: اليوم في المبيعات عدد

b. Predictors in the Model: (Constant), اليوم في الحرارة درجة

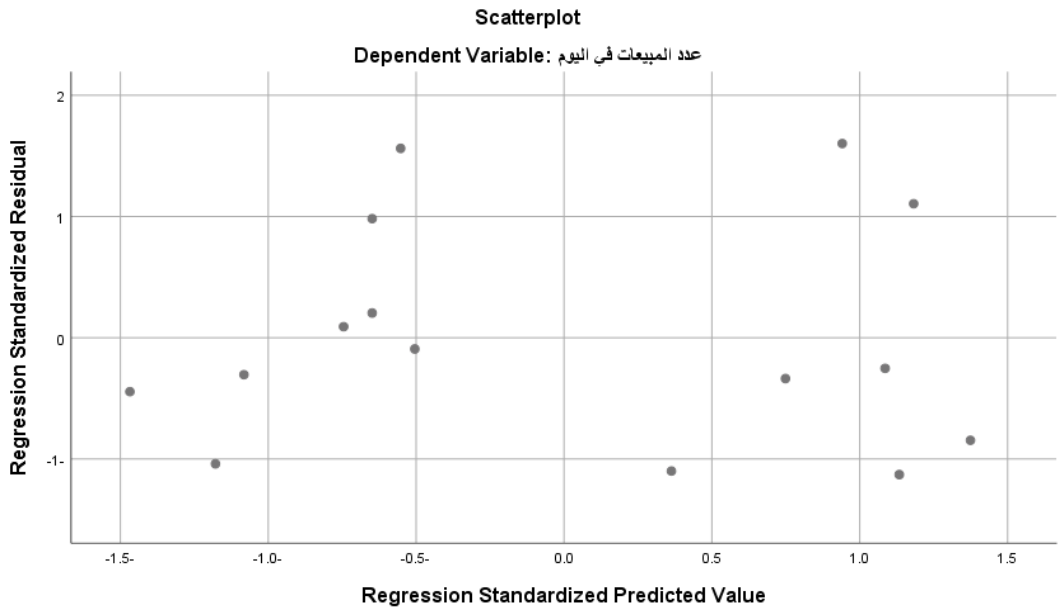
الجدول الخامس: بعنوان Excluded Variables وهو خاص بالمتغيرات المستبعدة.

Residuals Statistics[a]

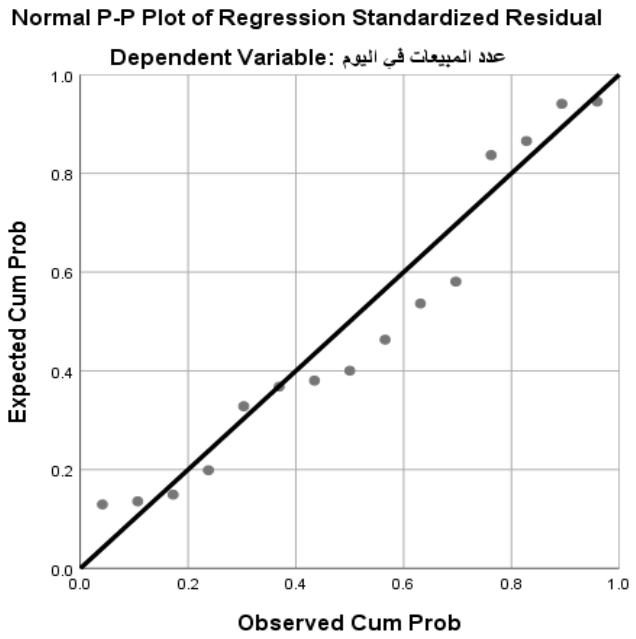
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	17.86	55.44	37.27	13.222	15
Std. Predicted Value	-1.468-	1.374	.000	1.000	15
Standard Error of Predicted Value	1.945	4.247	2.801	.667	15
Adjusted Predicted Value	18.72	59.65	37.52	13.420	15
Residual	-7.255-	10.294	.000	5.949	15
Std. Residual	-1.129-	1.602	.000	.926	15
Stud. Residual	-1.257-	1.826	-.017-	1.034	15
Deleted Residual	-9.654-	13.374	-.254-	7.483	15
Stud. Deleted Residual	-1.291-	2.057	.008	1.090	15
Mahal. Distance	.349	5.182	1.867	1.347	15
Cook's Distance	.000	.333	.090	.113	15
Centered Leverage Value	.025	.370	.133	.096	15

a. Dependent Variable: اليوم في المبيعات عدد

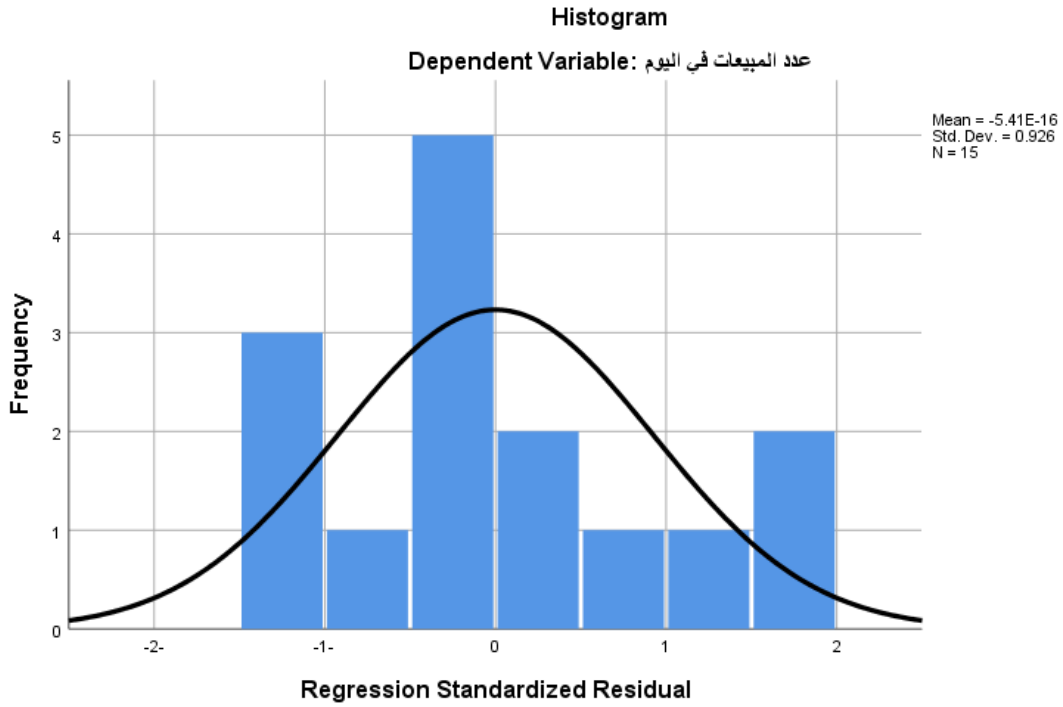
الجدول السادس: بعنوان Residual Statistics ويستخدم لتحليل البواقي.



الشكل البياني: يمثل شكل الانتشار للبواقي مع القيم المتوقعة ومنه يتضح عدم وجود نمط معين للنقاط في الشكل وهذا يتسق مع شرط الخطية.



الشكل البياني: يختبر هل البواقي تتبع التوزيع الطبيعي ام لا؟ ومن الشكل نجد أن النقاط تتجمع حول الخط وبالتالي فان البيانات (البواقي) تتوزع حسب التوزيع الطبيعي.



الشكل البياني: هو المدرج التكراري ويستخدم للتعرف هل البيانات تتوزع حسب التوزيع الطبيعي أم لا؟

حصة الأعمال الموجهة

تطبيق: لتكن البيانات التالية لعينة عشوائية مسحوبة من موظفي مؤسسة ما:

الرقم	الراتب الشهري	سنوات الخبرة Ex	عدد سنوات التعليم الجامعي Edu	الجنس
1	56900	5.5	4	F
2	60500	9	4	M
3	58900	4	5	F
4	59000	8	4	M
5	57500	9.5	5	M
6	55500	3	4	F
7	56000	7	3	F
8	52700	1.5	4.5	F
9	65000	8.5	5	M
10	60000	7.5	6	F
11	56000	9.5	2	M
12	53600	6	2	F
13	55000	2.5	4	M
14	52500	1.5	4.5	M

المطلوب:

1) أوجد معادلة الانحدار الخطي المتعدد للمتغيرات المستقلة الثلاثة على المتغير التابع (الراتب السنوي) باستخدام حزمة Spss.

2) أعد تقدير معادلة الانحدار المعنوية.

المراجع:

- أسامة، ربيع أمين (2008). التحليل الإحصائي للمتغيرات المتعددة باستخدام برنامج spss.
- التنجي معن (2016). الإحصاء المهني باستخدام إكسل. الإصدار الثالث: مركز سبر للدراسات الإحصائية والسياسات العامة. www.sabr-sp.com.
- دجلة، إبراهيم مهدي (2010). حول أسلوب تحليل التغاير المتعدد باستخدام تصميم قطع منشقة. مجلة كلية الإدارة والاقتصاد. المجلد 16. العدد 60: جامعة بغداد. العراق.
- الطريري، عبد الرحمان بن سليمان (2013). القياس النفسي والتربوي: الرياض. السعودية.
- النجار، نبيل جمعة صالح (2015). الإحصاء التحليلي مع تطبيقات برمجية spss. ط1: دار الحام للنشر والتوزيع. عمان. الأردن.
- عبد الفتاح، محمد نجيب وآخرون (2009). التحليل المعمق للبيانات باستخدام حزمة البرامج الجاهزة SPSS. دليل منهجي للمستخدم: جامعة الدول العربية.
- عبد الفتاح، مصطفى محمد . الانحدار المتعدد، كلية العلوم، جامعة المنصورة: مصر.
- علام، صلاح الدين محمود (2016). مقدمة لحزمة البرامج الإحصائية spss في علم النفس. ط1: دار الفكر. عمان. المملكة الهاشمية الأردنية.