الانحدار Regression والانحدار هو أسلوب يمكن بواسطته تقدير قيمة أحد المتغيرين بمعلومية قيمة المتغير الآخر عن طريق معادلة الانحدار، وله أنواع :

الانحدار الخطي البسيط : فكلمة " بسيط " تعني أن المتغير التابع Y يعتمد على متغير مستقل واحد وهو X وكلمة " خطي" تعني أن العلاقة بين المتغيرين (X , Y) علاقة خطية .

الانحدار المتعدد : إذا كان المتغير Y يعتمد على أكثر من متغير مستقل

الانحدار غير الخطي : إذا كانت العلاقة بين المتغير Yوالمتغيرات المستقلة غير خطية كأن تكون من الدرجة الثانية أو آسية.

الانحدار الخطي البسيط

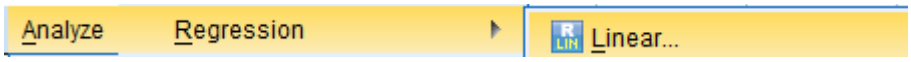
بعد تمثيل الأزواج المرتبة بالمستوى نحصل على شكل الانتشار فإذا أظهر الشكل الانتشاري للبيانات أن هناك علاقة خطية بين المتغيرين نقوم بتقدير خط الانحدار y على X بواسطة العلاقة :

$$\hat{y} = a + bx$$

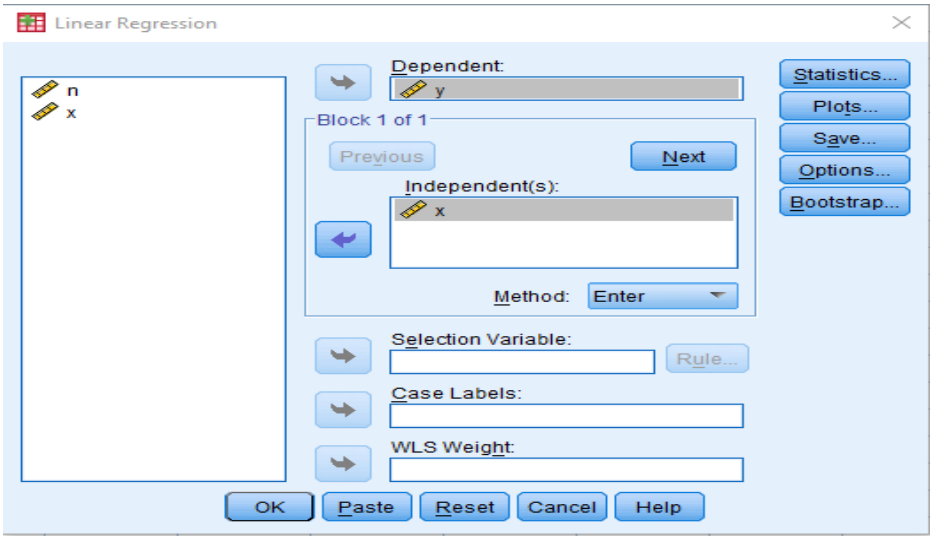
حيث a: ثابت الانحدار أو الجزء المقطوع من محور y

b: ميل الخط المستقيم أو معامل انحدارy على X او Y/X

تحليل الانحدار في لبرنامج spss



يظهر صندوق حوار باسم Linear Regressionوالذي يكون ترتيبه كما يأتي



Dependent: يمثل المتغير المعتمد.

Independent: المتغير او المتغيرات المستقلة يمكن ادخال اكثر من مجموعة من المتغيرات المستقلة كل مجموعة تدخل ضمن **Block** له رقم تسلسلي.

Method: الطريقة المستخدمة في الانحدار والطريقة الافتراضية هي طريقة الانتر (Enter).

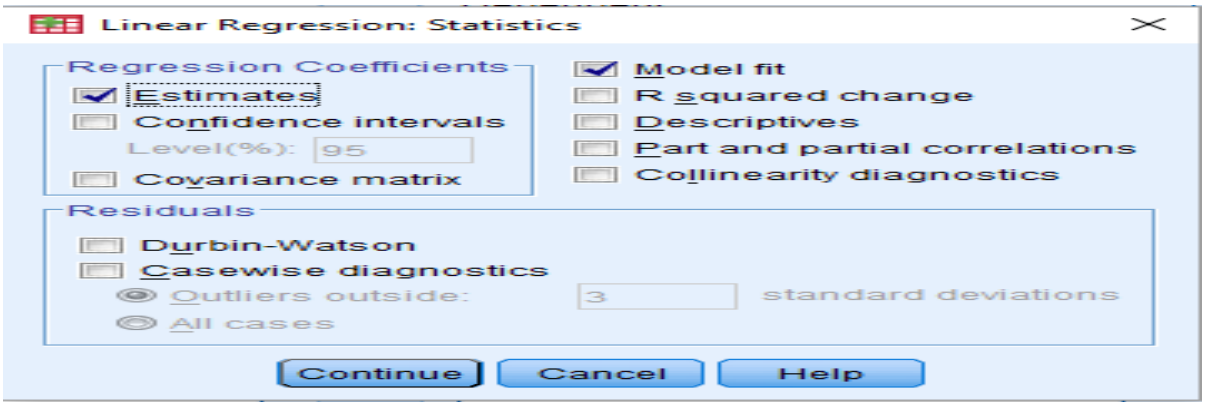
ملاحظة: تختلف الطريقة الافتراضية (Enter) عن الطرق الاخرى وذلك من خلال ان الطريقة الافتراضية يتم بموجبها ايجاد النتائج للنموذج المحدد في صندوق الحوار بينما الطرق الاخرى يتم من خلالها ايجاد افضل نموذج انحدار.

Selection Variable: يستعمل في تحديد التحليل لمجموعة من البيانات لمشاهدات المتغير الاختبار في المثال المتغير (n).

ملاحظة: ان المتغير في قائمة (**Selection Variable**) يجب ان لايكون من المتغيرات المستقلة او المتغير الرئيسي ويجب ان بتصف بمجموعة من الصفاة لكي يقوم البرنامج بتحديد الحالات التي سيتم ادخالها في تحليل الانحدار

Case Labels: متغير تستخدم قيمة كعناوين لنقاط شكل الانتشار **scatterplots**.

عند النقر على الزر **Statistic** يظهر صندوق حوار **Statistics** الذي يكون ترتيبه كالاتي



Estimates: لتقدير معالم نموذج الانحدار واختبارات T المرافقة.

Confidence Interval: لتقدير فترات الثقة 95% لكل من معلمتي الانحدار.

Model Fit: لعرض معامل التصحيح المصحح وغير المصحح وجدول تحليل الانحدار (ANOVA)

Descriptive: لحساب الوسط الحسابي والانحراف المعياري للمتغير المعتمد والمتغير او المتغيرات المستقلة

Part and partial correlations: لاحتساب الارتباط بين المتغير المستقل والمتغير المعتمد باستبعاد المتغيرات المستقلة الاخرى والارتباط الجزئي بين المتغيرات المستقلة والمتغير المعتمد. اضافة الى احتساب معامل ارتباط person r Value.

Collinearity diagnostics: يستخدم هذا الاختيار لاختبار مشكلة الارتباط المتعدد بين المتغيرات المستقلة في حالة الانحدار المتعدد. حيث يتم في البداية حساب المعامل التسامح (Tolerance) لكل المتغيرات المستقلة حيث ان

$$Tolerance = 1 - R^2_{x_{i,others}}$$

$R^2_{x_{i,others}}$ تمثل مربع معامل الارتباط المتعدد بين المتغير المستقل (x_i) وبقية المتغيرات المستقلة. فاذا كانت معامل التسامح (Tolerance) أصغر من (0.1) دل ذلك على وجود مشكلة التعدد الخطي اما إذا كان معامل التسامح أصغر من (0.2) دل ذلك باحتمال وجود مشكلة التعدد الخطي. بعدها يتم حساب معامل التضخم التباين (VIF) وهو اختصار لـ (Variance Inflation Factor) حيث ان

$$VIF = \frac{1}{Tolerance}$$

ان الحصول على قيمة لمعامل تضخم التباين تزيد عن (5) او (10) يشير الى ان تأثير المعلمة المرافقة بمشكلة التعدد الخطي.

الاختبار الثاني لمشكلة التعدد الخطي يتمثل باستخدام الجذور المميزة لمصفوفة (xx) ففي حالة وجود عدة جذور قريبة من الصفر فهذا دليل على وجود مشكلة التعدد الخطي . اما الاختبار الثالث لجود مشكلة التعدد الخطي بين المتغيرات المستقلة يستعمل دليل الحالة (condition index) وهو الجذر التربيعي لحاصل قسمة اكبر جذر مميز على بقية الجذور المميزة.

ملاحظة: إذا زادة قيمة دليل الحالة عن (15) دل ذلك على امكانية وجود مشكلة الارتباط الذاتي اما في حالة كانت قيمة دليل الحالة أكبر من (30) فهذا مؤشر على خطورة المشكلة.

المقياس الرابع لمشكلة التعدد الخطي هو (Variance Proportion) وهو يساوي نسبة تباين المفسر بواسطة المكون الاساسي (Principle Component) المرافقة لكل جذر مميز حيث تعتبر مشكلة التعدد الخطي مؤثرة اذا كان المكون الاساسي المرافق لدليل الحالة (Condition Index) مرتفع نسبيا.

وللكشف عن وجود المشكلة بواسطة هذا الاختبار يتم بالبحث عن النسب الكبيرة للتباين والتي تقابل الجذور المميزه الصغيرة والبحث عن المتغيرات التي تملك نسب عالية للتباين على نفس الجذر المميز ذات القيمة الصغيرة

ملاحظـة: تتميـز مشـكلة التعـدد الخطـي بارتفـاع تبـاين معـالم النمـوذج وبالتـالي عـدم ظهـور المعلمـة معنوية نتيجة انخفاض احصائية (t) بالرغم من ان المتغير قد يكون مهم في النموذج.

Residuals:

فـي هـذه القائمـة يمكـن تفعيـل اختبـار Durbin – Watson لاختبـار الارتبـاط الـذاتي (اسـتقلالية الخطـأ) ولكـن برنـامج spss لا يعطـي قيمـة الدلاليـة لاختبـار دربـن واتسـون لـذلك عليـك ان تقـرر بنفسك إذا كانت القيمة مختلفة بشكل كافي عن (2)

اختيـار casewise diagnostics اذا تـم تفعيـل هـذا الاختيـار فـان البرنـامج سـوف يعطـي جـدول

Casewise Diagnostics[a]

Case Number	Std. Residual	Record Sales (thousands)	Predicted Value	Residual
1	2.125	330.00	229.9203	100.07975
2	-2.314-	120.00	228.9490	-108.94899-
10	2.114	300.00	200.4662	99.53375
47	-2.442-	40.00	154.9698	-114.96982-
52	2.069	190.00	92.5973	97.40266
55	-2.424-	190.00	304.1231	-114.12308-
61	2.098	300.00	201.1897	98.81030
68	-2.345-	70.00	180.4156	-110.41564-
100	2.066	250.00	152.7133	97.28666
164	-2.577-	120.00	241.3240	-121.32405-
169	3.061	360.00	215.8675	144.13246
200	-2.064-	110.00	207.2061	-97.20606-

a. Dependent Variable: Record Sales (thousands)

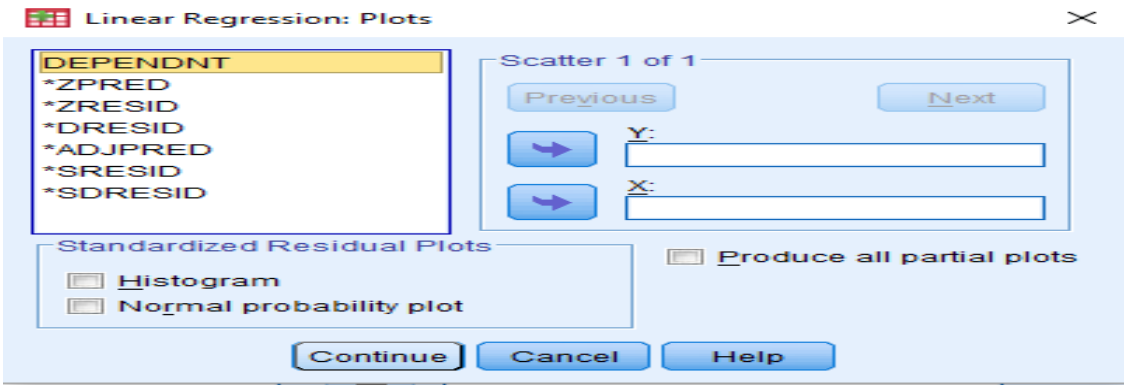
بـالقيم المشـاهدة للمتغيـر المعتمـد والقـيم المتوقعـة والفـرق بـين القيمـة المشـاهدة والقيمـة المتوقعـة والقيمـة المعياريـة للبـواقي ويمكـن ان يعطـي هـذه القـيم لجميـع الحـالات او الحـالات التـي يكـون فيهـا الانحـراف المعيـاري للبـاقي اكبـر مـن (3) (بغـض النظـر عـن الاشـارة) ويمكـن تغيـر المعيـار الـى (3) وسـوف تسـتخدم القيمـة (2) كمـا سـوف يعطـي جـدول يلخـص احصـائيات البـواقي والتـي تتضـمن الاشـارة الـى القيمـة الـدنيا والقيمـة والعظمـى والمتوسـط والانحـراف المعيـاري لكـل مـن القيمـة المتوقعـة وللبواقي.

طرق تشخيص الحالات

يعطي برنامج spss جدولا بإحصائيات الباقي والتي يجب تدقيقها للبحث عن الحالات الشاذة ويعرض الجدول ادناه الحالات التي تملك باقيا معياريا اصغر من (2−) واكبر من (2+) .

ملاحظة: يجب ان تكون هنالك 95% من الحالات تمتلك باقي معياري اصغر من 2− او اكبر من 2+ فاذا كانت لدينا عينة مؤلفة من 200 مشاهدة فننا نتوقع ان عدد الحالات التي باقي معياري خارج الحدود هي 10 حالات وفي المثال نلاحظ ان عدد الحالات التي تكن خارج الحدود هي 12 حالة اي 6% اضافة الى ذلك يجب ان تكون نسبة عدد الحالات التي تكون ضمن الحدود (2±) 99% وبالتالي فان 1% فقط من هذه الحالات تقع خارج الحدود .

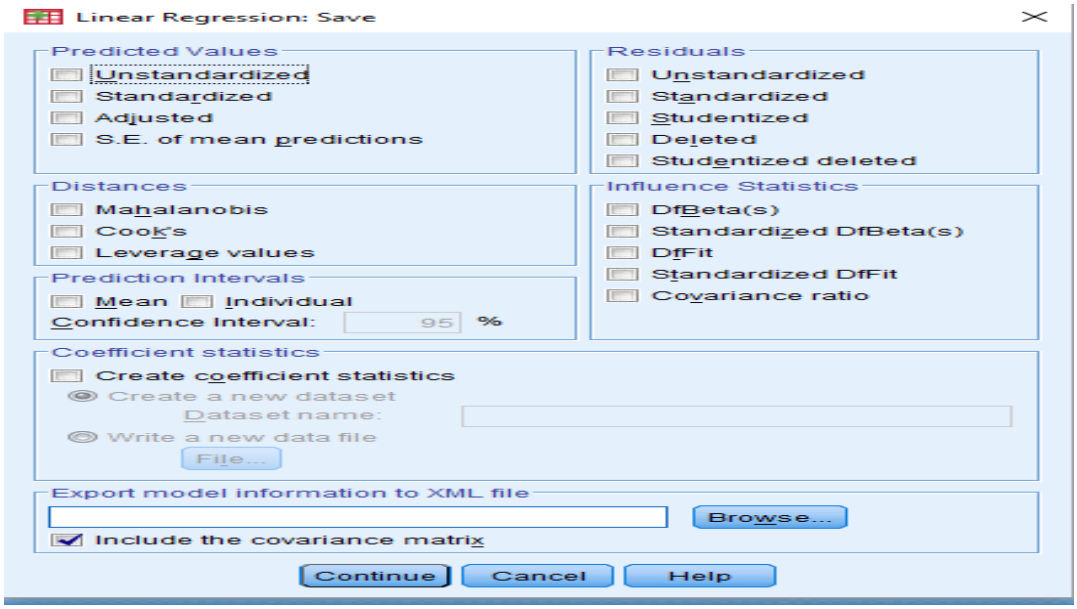
عند النقر على Plot يظهر صندوق حوار باسم Plots والذي يكون ترتيبه كما يلي:



يستخدم صندوق حوار Plots لا يجاد بعض الاشكال البيانية والتي يمكن استعمالها في التحقق من فرضيات تحليل الانحدار مثل (Normality ,Linearity and equality of variance) ، ويمكن استخدام برنامج الـ (Spss) لإيجاد رسم الانتشار لمشاهدات المتغير المعتمد Dependent Variable (DEPENDNT) مع اي من البواقي او القيمة التنبئية الاتية .

عند تفعيل الاختيار (Produce all partial plots) فان البرنامج يقوم بإيجاد رسم الانتشار للبواقي للمتغير المستقل مع المتغير المعتمد في حالة الانحدار البسيط وفي حالة الانحدار المتعدد يقوم البرنامج بإيجاد رسم الانتشار لكل متغير مستقل مع المتغير المعتمد وبصورة مستقلة

Standardized Residual Plots: يمكن من خلاله الحصول على المدرج التكراري (Histogram) للبواقي القياسية كما يمكن من خلاله ايجاد ((Normal p p) (Normal probability plot ومقارنة توزيع البواقي مع منحنى التوزيع الطبيعي.عند النقر على الزر Save يظهر صندوق حوار باسم (Save) والذي يكون كما يلي:



يستفاد من صندوق حوار save للحصول على القيم التنبئية والبواقي وبعض المتغيرات الاحصائية التي يمكن استخدامها في التشخيص حيث تظهر هذه المتغيرات في نافذة البيانات.

Predicted Values : القيم التنبؤية التي يمكن حسابها من نموذج الانحدار.

Unstandardized : القيمة التنبؤية غير المعيارية للمتغير المعتمد $(\hat{y})$.

Standardized : القيمة التنبؤية المعيارية للمتغير المعتمد.

Adjusted : القيمة المتوقعة للحالة عندما يتم استبعادها من نموذج الانحدار.

S.E. of mean predictions: الخطاء القياسي للقيمة التنبؤية.

Distances : تستخدم هذه القائمة لتحديد الحالات الشاذة للمتغيرات المستقلة التي تؤثر في النموذج عن طريق ايجاد متغيرات لثلاثة من المقاييس وهي

1- Mahalanobis

2- Cook's.

3- Leverage values

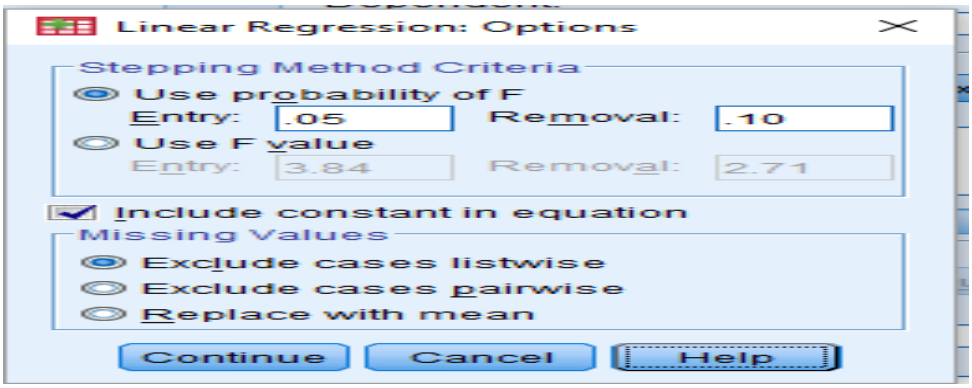
Residuals : من خلال هذا الاختيار يمكن ايجاد صيغ مختلفة لمتغير البواقي وهو المتغير الناتج من طرح القيمة التنبوئيه من القيمة المشاهدة للمتغير المعتمد. وكما يأتي:

Influence Statistics: تستخدم هذه القائمة لعرض التأثير في معلمات الانحدار والقيمة التنبؤية الناتج من حذف الحالات المقابلة.

Coefficient Statistics : تستخدم لحفظ معلمات الانحدار كملفات نافذة محرر بيانات مستقلة .

Prediction Intervals : الحدود العليا والدنيا للقيم للوسط الحسابي والانحراف المعياري للقيمة التنبؤية

عند اختيار **Option** تظهر نافذة باسم **Option** والتي يمكن توصيفها كما يلي :



Stepping Method Criteria: يستخدم في حالة الانحدار المتعدد لتحديد القيمة المعنوية للإحصاء (f) (القيمة الاحتمالية او القيمة الجدولية) لقبول او رفض المتغيرات لنموذج الانحدار في حالة اختيار احدى الطرق في النافذة الرئيسية حيث يكن

Include constant in equation: اختيار افتراضي يستخدم لتضمين الحد الثابت في نموذج الانحدار اي الانحدار من نقطة الاصل.