الجمهـوريــــة الجزائـريــــة الديمقراطيــــة الشـعـبيــة

وزارة التعليم العالي والبحث العلمي

جامعة قاصدي مرباح ورقلة

كلية الرياضيات و علوم المادة

Department of Mathematics

قسم الرياضيات

تخصص: احصاء واحتمالات

مُذكّرَة تخَرج لنيَل شهَادة الماستر

تحت عنوان :

الإنحدار الخطي المتعدد

من إعداد الطالب :

★ نايلي محمد الأخضر

تحت إشراف الأستاذ :

★ عقون رشيد

نوقشت يوم 2022/06/15 ، من طرف لجنة المناقشة :

♦ عباسي حسين ......أستاد بجامعة قاصدي مرباح ورقلة ...... رئيسا.

♦ عقون رشيد ......أستاد بجامعة قاصدي مرباح ورقلة ...... مشرفا.

♦ زيبار سعيد ......أستاد بجامعة قاصدي مرباح ورقلة ...... مناقشا.

السنة الجامعية : 2022/2021

بسم الله الرحمن الرحيم

❖ إهداء

بدأت بأكثر من يد وقاسيت أكثر من هم وعانيت الكثير من الصعوبات وها أنا اليوم والحمد الله أطوي سهر الليالي وخلاصة مشواري بين دفتي هذا العمل المتواضع الذي أهديه: الى من بلغ الرسالة وأدى الأمانة ونصح الأمة الى نبي الرحمة رسولنا صلى الله عليه وسلم.

الى من قال الله عز وجل في حقهما: (وقضى ربك ألا تعبدوا الا اياه وبالوالدين احسانا).

الى بهجة القلب وهبة الرب وكمال الود وصفاء الحب الى التي حملت الحياة بين يديها بريقا وشعاعا لدربي.

الى التي اهدتني رضاها ولم تبخل عني بدعواتها (أمي الغالية حفضها الله)

الى مصدر فخري أبي

الى أخواتي واخوتي وجميع أفراد عائلتي

الى كل من اجتمعت بهم على محبة الله والتقينا على طاعته وتوحدنا على دعوته

الى كل من شاطرني ثمرة جهدي

الى من يسعهم قلبي ولم تسعهم هذه الورقة.

نايلي محمد الأخضر

الفهرس

مقدمـــــــــة

يلعب تحليل الانحدار دورا كبيرا في دارسة ظواهر مهمة في مختلف مجالات الحياة وذلك من خلال تحليل هذه الظواهر والتنبؤ بنتائجها المستقبلية.
كما أن نموذج الانحدار يعبر عن علاقة إحصائية بين متغير الظاهرة المدروسة والذي يسمى بالمتغير التابع مع متغيرات أخرى تسمى بالمتغيرات المستقلة. يهدف تحليل الانحدار الى تقدير دالة الانحدار باستخدام طرق وسيطية وغير وسيطية، من بين الطرق الوسيطية المعتمدة في هدا المدكرة طريقة المربعات الصغرى وهي من الطرق الأكثر شيوعاً ، تهدف إلى تقدير خط الانحدار البسيط كما تسمح بتقدير المعلمات في الانحدار الخطي المتعدد الذي يؤدي إلى تقليل مجموع الانحرافات الرئيسية أو الأخطاء الواردة في النقاط التي تمت ملاحظتها في خط الانحدار ، حيث تتميز بأنها طريقة واضحة بسيطة وتسمح ببناء نماذج من بيانات إحصائية تسمح لنا بإجراء تنبؤات تحت افتراضات معينة.
يوسيب بوسكوفيتش [4] [12] هو أول رياضياتي يقوم بحساب معاملات نموذج انحدار خطي خلال أعماله في مجال الجغرافيا 1755 .
أول ظهور لمصطلح انحدار بالإنجليزية:(Regression) سنة 1866 في اعمال فرانسيس غالتون، حيث طبق نموذج الانحدار الخطي البسيط في نمذجة قامات الأطفال المولودين لآباء ذوي قامات غير اعتيادية (فارعة أو قصيرة الطول) مبينا بأن قامات الأطفال تؤول نحو المتوسط الحسابي لقامة الساكنة. تتألف هده المدكرة من ثلاث فصول:

- الفصل الأول بعنوان: الوسائل القاعدية المستعملة في كل من الاحتمالات والجبر الخطي
- الفصل الثاني بعنوان: الانحدار الخطي المتعدد
- الفصل الثالث بعنوان: دراسة مثال باستخدام برنامج SPSS لمنودج في الانحدار الخطي المتعدد.

الفصل الأول

قوانين في الجبر الخطي والاحتمالات

1.1 مقدمة

يلعب الجبر الخطي والمصفوفات دورا حيويا وهاما في التعبير عن العلاقات الرياضية متعددة المتغيرات بشكل بسيط يسهل فهمها وبالتالي إيجاد الحلول المناسبة لهذه العلاقات. فضلا عن ذلك فإن المصفوفات لها تطبيقات فى مجالات عديدة، فى الاقتصاد، والإحصاء وبحوث العمليات والعمليات الإدارية وغيرها من المجالات.

2.1 وسائل الجبر الخطي المستخدمة

تعريف 2.1.1 (المصفوفة). ليكن حقل حيث $(K = \mathbb{R} V \mathbb{C})$ و ليكن: $a_{ij} \in K; i = 1,\dots,n; j = 1,\dots,m; n,m \in \mathbb{N}$ و ليكن الشكل التالي:

$$A = \begin{pmatrix} a_{1.1} & \cdots & a_{1,\ \mathrm{m}} \\ \vdots & \ddots & \vdots \\ a_{\mathrm{n},1} & \cdots & a_{\mathrm{n,m}} \end{pmatrix} = a_{\mathrm{i,j}}$$

هذا الجدول يسمى مصفوفة .

- نسمي العناصر التي لها نفس الدليل الاول سطر
- نسمي العناصر التي لها نفس الدليل الثاني عمود

و نقول عندئذ أن المصفوفة هي ذات n سطر و m عمود.

أنواع المصفوفات

توجد المصفوفات على أنواع كثيرة ومختلفة وكل نوع له ما يميزه عن النوع الآخر بما يحتويه من صفوف وأعمدة.

المصفوفة المربعة: هى مصفوفة عدد صفوفها يساوى عدد أعمدتها ولذلك فهى تكون من الرتبة $n \times n$

المصفوفة المستطيلة: هى مصفوفة عدد صفوفها لا يساوى عدد اعمدتها ولذلك فهى تكون من الرتبة $n \times m$

مصفوفة الوحدة: وهى مصفوفة مربعة كل عنصر من عناصر قطرها الرئيسي يساوي الواحد الصحيح، وباقى عناصر المصفوفة أصفار. ويرمز لها بالرمز I .

المصفوفة القياسية: هى مصفوفة مربعة عناصر قطرها الرئيسي متساوية القيمة وباقي عناصرها أصفار.

المصفوفة القطرية: هي مصفوفة مربعة عناصر قطرها الرئيسي مقادير حقيقية ليست بالضرورة متساوية. وعلى ذلك فإن كلا من مصفوفة الوحدة والمصفوفة القياسية هى حالة خاصة من المصفوفة القطرية.

المصفوفة الصفرية: هي مصفوفة جميع عناصرها أصفار وقد تكون المصفوفة الصفرية مربعة أو مستطيلة.

مصفوفة الصف الواحد: وهي مصفوفة مستطيلة تحتوي على صف واحد وأي عدد من الأعمدة وصورتها العامة هي: $\left(a_{1.1}\ a_{1.2} \ldots a_{1.n} \right)$
متجه صف من الدرجة .$(1 \times n)$

مصفوفة العمود الواحد: وهي مصفوفة مستطيلة تحتوي على عدة صفوف وعمود واحد فقط وصورتها العامة هي: $a_{1.1}\ a_{2.1} \vdots a_{n.1}$ متجه صف من الدرجة .$(1 \times n)$

منقول المصفوفة:

نحصل على منقول المصفوفة بتبديل صفوفها بأعمدتها أي بجعل صفها الأول مكان العمود الأول وصفها الثاني مكان العمود الثاني وهكذا. فإذا كان لدينا المصفوفة :

$$A = \begin{bmatrix} a_{1.1} & a_{1.2} & a_{1.3} \\ a_{2.1} & a_{2.2} & a_{2.3} \\ a_{3.1} & a_{3.2} & a_{3.3} \end{bmatrix}$$

فإن مبدول هذه المصفوفة A ويرمز لها بالرمز A' يكون على الصورة الآتية:

$$A' = \begin{bmatrix} a_{1.1} & a_{2.1} & a_{1.3} \\ a_{1.2} & a_{2.2} & a_{2.3} \\ a_{1.3} & a_{3.2} & a_{3.3} \end{bmatrix}$$

خصائص منقول المصفوفة :

- $(A')' = A$
- $(A' + B' + C') = (A + B + C)'$
- $(A' \times B' \times C') = (A \times B \times C)'$
- $(A')^{-1} = \left(A^{-1}\right)'$

ملاحظة 2.1.1. إذا كان $A' = A$ فإن المصفوفة تسمى مصفوفة متماثلة ، ونلاحظ أنه عند إيجاد منقول المصفوفة A نجد أن عناصر القطر الرئيسي لا تتغير وبالتالي فإنه يمكننا القول أن المصفوفة القطرية والمصفوفة القياسية ومصفوفة الوحدة والمصفوفة الصفرية المربعة جميعها مصفوفات متماثلة. كما أن المصفوفات المتماثلة تعطى معكوسات متماثلة أيضا.

درجة المصفوفة: المصفوفة التي لها n سطر و m عمود نقول عنها أنها من الدرجة $n \times m$ و إذا $n = m$ فإننا نقول أن المصفوفة من الدراجة n
نرمز لمجموعة المصفوفات من الدرجة $n \times m$ حيث عناصرها موجودة في الحقل $(K = \mathbb{R} \vee \mathbb{C})$ بالرمز $\mathscr{M}_{n\times m}(K)$
نرمز لمجموعة المصفوفات من الدراجة n حيث عناصرها موجودة في الحقل $(K = \mathbb{R} \vee \mathbb{C})$ بالرمز $\mathscr{M}_n(K)$

ضرب مصفوفة في مقدار سلمي

لتكن $A = (a_(i.j))$ حيث $A \in \mathscr{M} n \times m(\mathbb{K})$ فإننا نعرف المصفوفة $\lambda \in \mathbb{K}$: λA بالشكل التالي : $\lambda A = (\lambda a_(i.j))$ أي نقوم بضرب العدد السلمي λ في جميع عناصر المصفوفة.

ضرب المصفوفات:

شرط الجمع)لازم و كافي (: حتى يمكن ضرب مصفوفتان , يجب أن يتحقق الشرط التالي: عدد أعمدة المصفوفة الاولى يساوي عدد أسطر المصفوفة الثانية, أي أنه إذا كانت : $A = (a_{i.j}) \in \mathscr{M}\mathrm{n} \times \mathrm{m(K)}$ فيجب أن تكون المصفوفة الثانية $B = (b_{i.j}) \in \mathscr{M}\mathrm{m} \times \mathrm{p(K)}$ و في هذه

الحالة يكون الجداء $A\times B$ ممكن لأن عدد أعمدة المصفوفة A يساوي عدد أسطر المصفوفة B و في هذه الحالة لدينا:

$$A\times B=\sum_{e=1}^{m}a_{i.e}b_{e.k}$$

حيث : $n,m,p\in\mathbb{N}\ ;K=1,,p\ ;\ i=1,,n$

نتيجة 2.1.1. في الحالة العامة فإن ضرب المصفوفات ليس تبديلي أي:

$$A\times B\neq B\times A$$

ملاحظات 2.1.1. • الضرب توزيعي على الجمع في المصفوفات أي:

$$\mathrm{A(B+C)=A\cdot B+A\cdot C;\ \forall A\in\mathscr{M}_{n,m}(K),\ \forall C,B\in\mathscr{M}_{m,p}(K)}$$

- الجداءات الشهيرة ليست محققة في الحالة العامة في المصفوفات و هذا لأن الضرب غير تبديلي أي أن على سبيل المثال .
لدينا:

$$(A+B)^2\neq A^2+B^2+2A\cdot B$$

$$(A-B)^2\neq A^2+B^2-2A\cdot B$$

$$(A+B)(A-B)\neq A^2-B^2$$

- مربع المصفوفات لا يكون إلا في المصفوفة المربعة لأن في المصفوفة المربعة فقط يتساوى عدد الاعمدة مع عدد الاسطر.

مقلوب مصفوفة: لتكن $A\in\mathscr{M}_n(K)$ أي مصفوفة مربعة) لحساب مقلوب مصفوفة يجب ان تكون المصفوفة مربعة (إذا وجدت مصفوفة B من نفس دراجة المصفوفة A حيث يكون لدينا:

$$A\times B=B\times A=I_n$$

فإننا نقول في هذه الحالة أن المصفوفة A قابلة للقلب و مقلوبها هو المصفوفة B و نرمز لها بالرمز $A^(-1)$

رتبة مصفوفة: لتكن $A\in\mathscr{M}_{n,m}(K)$ أي أن المصفوفة ليست بالضرورة مربعة, نسمي رتبة مصفوفة درجة أكبر مصفوفة مربعة مستخرجة من المصفوفة A تكون قابلة للقلب, أو هي عدد أشعة الاعمدة أو الاسطر التي تكون مستقلة خطيا.

ترميز: نرمز لرتبة المصفوفة A بالرمز $\mathrm{rg}(A)$.

نتيجة 2.1.2. من الواضح من خلال التعريف على ان رتبة المصفوفة $A \in \mathscr{M}_{n,m}(K)$ تكون أقل أو تساوي من أصغر العددين n و m أي: $\mathrm{rg}(A) \leq \min(n,m)$

أثر المصفوفة: في المصفوفات المربعة يكون لدينا القطر الرئيسي و مجموع عناصر القطر الرئيسي يسمى أثر المصفوفة أي إذا كان لدينا : $A = (a_{i,j}) \in \mathscr{M}_n(K)$ فإن عناصر القطر الرئيسي هي : $a_{1.1}, a_{2,2},, a_(n,n)$ و بالتالي فإن أثر المصفوفة هو مجموع هذه العناصر و نرمز له بالرمز $\mathrm{Tr}(A)$ أي:

$$\mathrm{Tr}(A) = a_{1.1}, a_{2,2},, a_(n,n)$$

بعض الخواص: مما سبق يمكن استنتاج الخواص التالية:

$$\forall A, B \in \mathscr{M} n \times m(K); (A+B)^T = A^T + B^T$$

$$\forall A \in \mathscr{M} n \times m(K), \forall B \in \mathscr{M} m \times p(K); (A.B)^T = B^T \cdot A^T$$

$$\forall \lambda \in K, \forall A \in \mathscr{M} n \times m(K); (\lambda A)^T = \lambda A^T$$

$$\forall A, B \in \mathscr{M} n(K); \mathrm{Tr}(A+B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$$

$$\forall \lambda \in K, \forall A \in \mathscr{M}_n(K); \mathrm{Tr}(\lambda A) = \mathrm{Tr}(\lambda A)$$

التوقع الرياضي و التباين المشترك

التوقع الرياضي : للتوقع الرياضي أهمية كبيرة في التحليل الإحصائي و يستخدم للحصول على عدد من المقاييس الإحصائية، التي تصف الخواص الهامة للتوزيعات الاحتمالية، من هذه المقاييس على سبيل المثال التباين .
ليكن المتغير العشوائي X حيث

$$X = (x_1, x_2,, x_n)$$

فإن التوقع الرياضي ل X سيكون بالصيغة التالية :

$$\begin{aligned}\mathbb{E}(\boldsymbol{X}) &= x_1 p\,(X = x_1), \ldots\ldots\ldots, x_n p\,(X = x_n)\\ &= \sum_{i=1}^{n} x_i p\,(X = x_i)\end{aligned}$$

عندما تكون الاحتمالات متساوية $\mathbb{E}(X)$ تكتب من الشكل :

$$\mathbb{E}(X) = \frac{x_1 + x_2 \cdots\cdots\cdots + x_n}{n}$$

التباين المشترك (التغاير):

يعبر عن العلاقة بين متغيرين X و Y وهو القيمة المتوقعة لانحرافهما عن الوسط. التباين المشترك بين متغيرين عشوائيين يقيس درجة العلاقة الخطية بين هذين المتغيرين، يرمز للتباين المشترك ب $\operatorname{cov}(x;y)$ أو $\sigma_(x.y)$ و صيغته الرياضية هي من الشكل :

$$\operatorname{cov}(x;y) = \mathbb{E}[(\mathrm{X}-\mu_X)(Y-\mu_Y)]$$

حيث: μ_X هو متوسط المتغير X.
μ_Y هو متوسط المتغير Y.
أي أن التبيان المشترك هو القيمة المتوقعة لحاصل ضرب $(Y-\mu_Y)$ في $(\mathrm{X}-\mu_X)$ فإذا كان موجبا، أي $cov(x;y)>0$ فهذا يدل على أن هناك علاقة طردية بين المتغيرين. إذا كانت قيمته سالبة، أي $cov(x;y)<0$ فهذا يدل على أن هناك علاقة عكسية بين المتغيرين. إذا كانت قيمته مساوية للصفر، أي $cov(x;y)=0$ فهذا يعني أن : المتغيرين X و Y مستقلين و بالتالي لا توجد علاقة بين هاذين المتغيرين.

المتغيرات العشوائية و التوزيعات الاحتمالية:

نقول عن المتغير X أنه متغير عشوائي إذا كانت القيمة العددية لهذا المتغير غير مؤكدة فالقيم المختلفة للمتغير العشوائي ترتبط باحتمالات معينة، حيث أنها تشكل معا ما يسمى بالتوزيع الاحتمالي. و تنقسم المتغيرات العشوائية إلى متغيرات عشوائية منفصلة، و متغيرات عشوائية متصلة. إن القانون الذي يعطي الاحتمالات الخاصة بمختلف القيم التي يأخذها المتغير العشوائي المستمر X يسمى بالتوزيع الاحتمالي حيث أن:

$$p(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x)dx$$

التوزيع الطبيعي -لابلاس غوس-

هو توزيع احتمالي لمتغير عشوائي متصل و هو أكثر التوزيعات استخداما في التحليل الإحصائي. التوزيع الطبيعي جرسي الشكل و متماثل حول الوسط الحسابي.
إذا كان X متغير عشوائي متصل، نقول أن X يتبع قانون طبيعي بوسط μ و انحراف معياري σ

$$X \longrightarrow N(\sigma,\mu)$$

إذا كانت دالة كثافته الاحتمالية من الشكل:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2\right)\cdot\forall\mathrm{X}\in\mathrm{R}$$

توزيع فيشر (F):

قانون فيشر هو قانون احتمالي مكون من متغيرين يتبعان قانون كاي تربيع ومستقلين عن بعضهما و مقسومين على درجات حريتيهما.
الصياغة الرياضية لتوزيع فيشر هي كمايلي:

$$F=\frac{X/n_1}{Z/n_2}$$

حيث أن: n_1 : عدد درجات الحرية للمتغير X.
n_2 : عدد درجات الحرية للمتغير Z.
لقانون فيشر دور مهم في المقارنة بين تباينات المتغيرات العشوائية للعينات المستخرجة من مجتمعين وهذا النوع من الاختبار مفيد في تحليل التباين وفي تحليل الانحدار .

التقدير:

التقدير هو عملية استنتاج أو تقدير أحد معالم المجتمع مثل المتوسط الحسابي أو الانحراف المعياري... من المعلومات التي تقدمها العينات و كي يكون سليما ينبغي أن نعتمد على عينة ممثلة للمجتمع و يمكن تحقيق ذلك بالمعاينة العشوائية، حيث يكون لكل مفردة في المجتمع فرصة متكافئة للدخول في العينة.
إن تقدير أحد معالم المجتمع يمكن أن يكون :

تقدير نقطي: هو تقدير معلمة المجتمع بنقطة تحسب من بيانات العينة.

التقدير بفترة: لا يمكن الحصول على معالم المجتمع بدون خطأ مهما كان التقدير، لذلك يجب إعطاء فترة معينة لتوقع وقوع معالم المجتمع داخلها، هذه الفترة تسمى بفترة الثقة، علما بأنه تزداد فترة الثقة بزيادة حجم العينة و يمكن تحديد الفترة $(a;b)$ التي تضم معالم المجتمع باحتمال قدره $(1-a)$، كمايلي:

$$\boldsymbol{P}(\boldsymbol{a}<\mu<b)=1-\boldsymbol{a}$$

إن أهم الطرق المعروفة في التقديرهي :

- طريقة المربعات الصغرى .Méthode des moindres carrés
- طريقة العزوم .Méthode des moments
- طريقة المعقولية العظمى .Maximum de vraisemblance

الفصل الثاني

الانحدار الخطي المتعدد

1.2 تقديم النموذج

نموذج الانحدار الخطي المتعدد ويسمى أحيانا النموذج الخطي العام هو امتداد للنموذج البسيط حيث انه يتضمن أكثر من متغير مستقل واحد، في حالة النموذج البسيط كان الأمر يعتمد على متغيرين متغير تابع والآخر متغير مستقل، لكن في حالة النموذج العام قد يتضمن عدد من المتغيرات من بينها قد يكون هناك تابع واحد والعديد من المتغيرات المستقلة. [5]

2.2 الصياغة العامة للنموذج

يستند النموذج الخطي العام على افتراض وجود علاقة خطية ما بين المتغير التابع وعدد من المتغيرات المستقلة، يمكن كتابة نموذج انحدار خطي عام أو متعدد على الشكل [3] [5] التالي:

$$Y_i = \beta_0 x_{i,0} + \beta_1 x_{i,1} + \ldots + \beta_{p-1} x_{i,p-1} + \varepsilon_i, \quad \forall i = 1 \ldots n$$

حيث:

- $p \leq n$،
- $x_{i,0}, x_{i,1}, \ldots, x_{i,p-1}$ متغيرات مستقلة
- $\beta_0, \beta_1, \ldots, \beta_{p-1}$ تسمى معالم النموذج وعددها: p
- ε_i متغير عشوائي يعبر عن الأخطاء الموجودة في النموذج أو أيضا البواقي لأنه يمكن كتابتها عند كل مشاهدة (i) على الشكل التالي:

$$\varepsilon_i = Y_i - \beta_0 x_{i,0} - \beta_0 x_{i,1} - \ldots - \beta_0 x_{i,p-1}$$

يمكن كتابة هذا النظام على الشكل المصفوفي التالي [1] :

$$Y = X\beta + \varepsilon$$

$$Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \quad X = \begin{pmatrix} x_{1,0} & \ldots & x_{1,p-1} \\ \vdots & & \vdots \\ x_{n,,0} & \ldots & x_{n,p-1} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_{p-1} \end{pmatrix} \quad \text{et} \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

3.2 أنواع الانحدار الخطي المتعدد

هناك ثلاثة نماذج انحدار متعدد رئيسية هي:

- الانحدار العياري أو القياسي (standard) .
- 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. [10]

4.2 فرضيات النموذج

ان بناء نموذج الانحدار الخطي يجب أن يكون مستوفيا لعدد من الفرضيات التي يمكن اجمالها فيما يلي:

1- $E[\varepsilon_i]=0$ من أجل كل i=1...n

2- $cov(\varepsilon_i, \varepsilon_j) = 0$ من أجل كل $i \neq j$

3- $Var(\varepsilon_i) = \sigma^2$ من أجل كل i=1...n

5.2 شروط النموذج

$E(\varepsilon) = 0$ و $E(Y) = X\beta$,

$Var(\varepsilon) = Var(Y) = \sigma^2 I_n$

المصفوفة $(X'X)^{-1}$ موجودة

جميع المتغيرات التفسيرية غير عشوائية.

6.2 مثال على النموذج الانحدار المتعدد

ان العائد من المحاصيل يعتمد على كمية النيتروجين والأسمدة الفوسفاتية المستخدمة. يتم السيطرة على هذه المتغيرات من قبل الباحث، ولكن العائد من المحاصيل يمكن أن يعتمد أيضاً على متغيرات لا يمكن السيطرة عليها مثل تلك المرتبطة بالطقس.
قد يتم وضع النموذج المفترض بشكل مثالي لتبسيط حالة العالم الحقيقي المعقدة ولكن في كثير من الحالات توفر نماذج تقريبية مفيدة للعلاقات بين المتغيرات. ربما تكون هذه العلاقات إما ترابطية أو سببية.

7.2 طريقة المربعات الصغرى وتطبيقها على النموذج العام

تهدف هذه الطريقة إلى إيجاد تقدير للشعاع β الذي يصغر مجموع مربعات الانحراف $\hat{\epsilon}$ بين القيمة المقدرة $\hat{Y}$ والقيمة الحقيقية [8]

$$Min \sum_{i=1}^{n} \widehat{\epsilon_i}^2 = Min \sum_{i=1}^{n} (Y_i - \widehat{Y_i})^2$$
$$\widehat{\epsilon_i} = Y_i - \widehat{Y_i} \ ; \quad i = 1 \ldots n$$

$$Min\sum_{i=1}^{n}(Y_i-\widehat{Y_i})^2=Min(Y_i-\widehat{Y_i})(Y-\widehat{Y_i})=Min\widehat{\epsilon}\widehat{\epsilon}$$

$$\Gamma(Y,X,\widehat{\beta})=(Y-\widehat{Y})(Y-\widehat{Y})=\widehat{Y}\widehat{Y}-2\widehat{Y}Y+YY=\widehat{\beta}\widehat{X}XY-2\widehat{\beta}XY+YY$$

حيث $\widehat{Y}=X\widehat{\beta}$ومنه الهدف هو

$min_{\widehat{\beta}}\Gamma(Y.X.\widehat{\beta})$ اذا كان$\widehat{\beta}$ موجود فيجب ان يحقق الشرط الضروري:

$$\frac{\partial(Y,X,\widehat{\beta})}{\partial\widehat{\beta'}}=2(X'X)\widehat{\beta}-2X'Y=0$$

بما ان رتبة X هي p فان:$(X'X)$ مصفوفة مربعة $(p)*(p)$رتبتها p وتقبل معكوس $(X'X)^{-1}$

ومنه :

$$2(X'X)\widehat{\beta}-2X'Y=0\Rightarrow(X'X)\widehat{\beta}-X'Y=0$$

نضرب المعادلة ب $(X'X)^{-1}$لنحصل على $\widehat{\beta}=(X'X)^{-1}X'Y$ وهو تقدير ل β

وللتأكد من أن$\widehat{\beta}$ المتحصل عليه هو قيمة دنيا ل $\Gamma(Y.X.\widehat{\beta})$ يجب تحقيق الشرط من الدرجة الثانية:

$$\frac{\sigma^2\Gamma(Y,X,\widehat{\beta})}{\sigma\widehat{\beta}'\sigma\widehat{\beta}}=X'X0$$

وهي مصفوفة معرفة موجبة ومنه فإن$\widehat{\beta}$هو نهاية صغرى

والآن لنرمز ب A للمصفوفة $(X'X)^{-1}X'$حيث :

ومنه نرى أن مختلف المقدرات $(\widehat{\beta}_1,\widehat{\beta}_2,.....,\widehat{\beta}_k)$ هي على شكل خطي مع المتغير Y

كذلك لدينا: $\widehat{\beta}=(X^{'}X)^{-1}X^{'}Y$

و أيضا: $Y=X\beta+\varepsilon$

اذن:

$$\widehat{\beta}=(X^{'}X)^{-1}X^{'}[X\beta+\varepsilon]=(X^{'}X)^{-1}X^{'}\varepsilon\Rightarrow\widehat{\beta}=\beta+(X^{'}X)^{-1}X^{'}\varepsilon$$

بإدخال التوقع الرياضي:

$$E(\widehat{\beta})=\beta+(X^{'}X)^{-1}X^{'}E(\varepsilon)$$

حيث : $E(\varepsilon)=0$ نحصل في الأخير : $E(\widehat{\beta})=\beta$

نستنتج أن التقدير $\widehat{\beta}$ ل β المحصل عليه بطريقة المربعات الصغرى غير متحيز.

بالإضافة إلى ذلك فإن $\widehat{\beta}$ هو التقدير الأفضل من ضمن كل التقديرات الخطية غير المتحيزة ل β

- التقدير MCO $\widehat{\beta}$ مقدر خطي غير متحيز ل β

$$Var(\widehat{\beta})=\sigma^2(X'X)^{-1}$$

برهان.
لدينا: $\widehat{\beta} = (X'X)^{-1}X'Y$ إنه بالفعل مقدر خطي (في Y). وبالإضافة $E(\widehat{\beta}) = E((X'X)^{-1}X'Y) = (X'X)^{-1}X'X\beta = \beta$ ومنه $\widehat{\beta}$ غير متحيز. ومنه:

$$\begin{aligned} Var(\widehat{\beta}) &= Var((X'X)^{-1}X'Y) \\ &= (X'X)^{-1}X'Var(Y)X(X'X)^{-1} \\ &= (X'X)^{-1}X'(\sigma^2 I_n)X(X'X)^{-1} \\ &= \sigma^2(X'X)^{-1} \end{aligned}$$

□

نظرية 7.2.1 (gauss markov).
المقدر $\widehat{\beta}$ للمربعات الصغرى العادية هو المقدر الوحيد خطية دون الحد الأدنى من التحيز التبايني بين المقدرات الخطية دون تحيز β [11]
البرهان :
$\widetilde{\beta}$ مقدر خطي غير متحيز ل β . ومنه $\widetilde{\beta}$ تكتب من الشكل :

$\widetilde{\beta} = AY$ حيث $AX\beta = \beta$ من أجل كل β يعني أن : $AX = I_p$.

$$\begin{aligned} Var(AY) &= Var(A(I_n - \Pi_X + \Pi_X)Y) \\ &= Var(A(I_n - \Pi_X)Y + A\Pi_X Y) \\ &= A(I_n - \Pi_X)\sigma^2 I_n(I_n - \Pi_X)A' + 2A(I_n - \Pi_X)\sigma^2 I_n \Pi_X A' + Var(AX(X'X)^{-1}X'Y) \\ &= \sigma^2 A(I_n - \Pi_X)A' + 0 + Var((X'X)^{-1}X'Y) \\ &= \sigma^2 A(I_n - \Pi_X)A' + Var(\widehat{\beta}) \end{aligned}$$

8.2 القيم المقدرة، البواقي

تعريف 8.2.1.
المتجه العشوائي $\widehat{Y} = \Pi_X Y = X(X'X)^{-1}X'Y$ يسمى متجه القيم المقدرة .
المتجه $\widehat{\varepsilon} = Y - \widehat{Y} = (I_n - \Pi_X)Y$ يسمى متجه البواقي.[10]

ملاحظة 8.2.1.
$\widehat{\varepsilon}$ متعامد مع المتجه $\widehat{Y}$ يتوافق مع الإسقاط المتعامد لـ Y على $\epsilon(X)^{\perp}$.

نظرية 8،2،1.

$E(\widehat{\varepsilon})=0, X'\widehat{\varepsilon}=0$ و $Var(\widehat{\varepsilon})=\sigma^2(I_n-\Pi_X)$.

9،2 تقدير تباين الأخطاء σ^2

إحدى فرضيات النموذج هي

$$E(\varepsilon\varepsilon')=\Omega_\varepsilon=\sigma^2 I_n$$

وبما أن σ^2 غير معروف ينبغي تقديره:

$$\begin{aligned}\widehat{\varepsilon}&=Y-X\widehat{\beta}\\&=X\beta+\varepsilon-X\widehat{\beta}\\&=\varepsilon-X(\widehat{\beta}-\beta)\\&=\varepsilon-X(X'X)^{-1}X'\varepsilon\\&=(I_n-X(X'X)^{-1}X')\varepsilon\end{aligned}$$

نضع :

$$M_x=(I_n-X(X'X)^{-1}X')$$

حيث M_x تسمى المصفوفة الدورية أي:

$$M_x=M_x'M_x=M_x^2=M'$$

بالإضافة إلى ذلك:

$$M_xX=0$$

ومنه

$$\widehat{\varepsilon}'\widehat{\varepsilon}=\varepsilon'M_x'M_x\varepsilon$$

اي :

$$\widehat{\varepsilon}'\widehat{\varepsilon}=\varepsilon'M_x\varepsilon$$

ندخل التوقع الرياضي على الطرفين

$$E(\widehat{\varepsilon}'\widehat{\varepsilon})=E(\varepsilon'M_x\varepsilon)$$

ويجب الملاحظة أن أثر $\widehat{\varepsilon}'\widehat{\varepsilon}$ يساوي أثر $\varepsilon'M_x\varepsilon$, ونعلم أيضا أن أثر $(AB)=(BA)$.

يكون لدينا إذن: أثر $\varepsilon'M\varepsilon$ يساوي $\varepsilon\varepsilon'M$

$$E(\widehat{\varepsilon}'\widehat{\varepsilon})=E(\varepsilon'\varepsilon)Tr(M_x)$$

نعلم أن:

$$E(\widehat{\varepsilon}'\widehat{\varepsilon}) = \sigma^2$$

وعليه:

$$E(\widehat{\varepsilon}'\widehat{\varepsilon}) = \sigma^2 \left\{ Tr(I_n) - Tr(X(X'X)^{-1}X') \right\}$$

ومنه

$$E(\widehat{\varepsilon}'\widehat{\varepsilon}) = \sigma^2(n-k-1)$$

حيث

$$Tr(X(X'X)^{-1}X') = k+1$$

$$Tr(I_n) = n$$

لكي نحصل على تقدير غير متحيز لσ^2 يكفي قسمة العبارة على $(n-k-1)$

$$E(\frac{\widehat{\varepsilon}'\widehat{\varepsilon}}{n-k-1}) = \sigma^2$$

في حالة الانحدار المتعدد حيث هناك $k+1$ معلم للتقدير و n عدد المشاهدات، وهذ يعطي عدد درجات الحرية $n-k-1$، ادن

$$\widehat{\sigma}_\varepsilon^2 = \frac{\widehat{\varepsilon}'\widehat{\varepsilon}}{n-k-1} = \frac{\sum_{i=1}^{n} \widehat{\varepsilon}_i^2}{n-k-1}$$

لقد برهنا أن $\widehat{\beta}-\beta$ تساوي $(X'X)^{-1}X'\sigma$ وإحدى فرضيات النموذج هي :

$$E(\varepsilon\varepsilon') = \Omega_\varepsilon = \sigma^2 I_n$$

لدينا:

$$\widehat{\beta} - \beta = (X'X)^{-1}X'\varepsilon$$

نقوم بحساب :

$$(\widehat{\beta}-\beta)(\widehat{\beta}-\beta)'$$

حيث :

$$(\widehat{\beta}-\beta)(\widehat{\beta}-\beta)' = (X'X)^{-1}X'\varepsilon\varepsilon'X(X'X)^{-1}$$

بإدخال التوقع الرياضي على الطرفين، نحصل على مصفوفة التباين المشترك للمقدرات:

$$\begin{aligned}\Omega_{\widehat{\beta}} &= E((\widehat{\beta}-\beta)(\widehat{\beta}-\beta)') \\ &= (X'X)^{-1}X'E(\varepsilon\varepsilon')X(X'X)^{-1} \\ &= X(X'X)^{-1}X\Omega_\varepsilon X(X'X)^{-1}\end{aligned}$$

اذن :

$$\Omega_{\widehat{\beta}} = \sigma_{\varepsilon}^{2}(X'X)^{-1}$$

بما أن σ^2 غير معروف، فانه يمكن استبداله بمقدر تباين الأخطاء $\widehat{\sigma}_{\varepsilon}^{2}$ وعليه:

$$\widehat{\Omega}_{\widehat{\beta}} = \widehat{\sigma}_{\varepsilon}^{2}(X'X)^{-1}$$

مجموع المربعات المقدرة ونقطة تقدير التباين

كما هو الحال في نموذج الانحدار الخطي البسيط ، يمكن استخدام متجه البواقي في تقدير σ^2.
نقدم مجموع مربع البواقي :

$$SCR = \sum_{i=1}^{n} \widehat{\varepsilon}_i^{\,2} = \|\widehat{\varepsilon}\|.$$

لدينا :

$$E[\|\widehat{\varepsilon}\|^2] = E[\widehat{\varepsilon}'\widehat{\varepsilon}] = E[tr(\widehat{\varepsilon}'\widehat{\varepsilon})]$$

ليكن A,B,C ثلاث مصفوفات :

$$tr(ABC) = tr(CBA) = tr(BCA)$$

يعني أن :

$$\begin{aligned} E[\|\widehat{\varepsilon}\|^2] &= E[tr(\widehat{\varepsilon}\widehat{\varepsilon}')] \\ &= trE[\widehat{\varepsilon}\widehat{\varepsilon}'] \\ &= tr(Var(\widehat{\varepsilon})) \\ &= \sigma^2 tr(I_n - \Pi_X) \\ &= \sigma^2(n - tr(X(X'X)^{-1}X')). \end{aligned}$$

ومنه :

$$\begin{aligned} E[\|\widehat{\varepsilon}\|^2] &= \sigma^2(n - tr(X'X(X'X)^{-1})) = \sigma^2 tr(n - tr(I_p)) \\ &= \sigma^2(n-p). \end{aligned}$$

التنبؤ

القيمة التفسيرية الجديدة

$$x_{n+1} = (x_{n+1,0}, \ldots, x_{n+1,p-1})$$

نريد أن نتوقع ملاحظة جديدة للمتغير

$$Y_{n+1} = \beta_0 x_{n+1,0} + ... + \beta_{p-1} x_{n+1,p-1} + \varepsilon_{n+1} = x_{n+1}\beta + \varepsilon_{n+1}$$

حيث :

$$E(\varepsilon_{n+1}) = 0,$$

$$Var(\varepsilon_{n+1}) = \sigma^2$$

و

$$Cov(\varepsilon_{n+1}, \varepsilon_n) = 0$$

من أجل كل $i = 1...n$

Y_{n+1} و Y_n غير مترابطان $i = 1..n,$ تستخدم لحساب $\widehat{\beta}$. لهذا، نقدم :

$$\widehat{Y}^p_{n+1} = x_{n+1}\widehat{\beta}.$$

يتم تعريف خطأ التنبؤ كما يلي :

$$\widehat{\varepsilon}^p_{n+1} = Y_{n+1} - \widehat{Y}^p_{n+1}$$

إنه مركزي، التباين يساوي :

$$\begin{aligned} Var(\widehat{\varepsilon}^p_{n+1}) &= Var(x_{n+1}\beta + \varepsilon_{n+1}\widehat{x}_{n+1}\beta) \\ &= Var(\varepsilon_{n+1}) + x_{n+1}Var(\widehat{\beta})x'_{n+1} \\ &= \sigma^2(1 + x_{n+1}(X'X)^{-1}x'_{n+1}). \\ Var(\widehat{\varepsilon}^p_{n+1}) &= E[(Y_{n+1} - \widehat{Y}^p_{n+1})^2] \end{aligned}$$

يسمى أيضا الخطأ التربيعي متوسط التنبؤ (EQMP) التي ستستخدم لاختيار المتغيرات أو نموذج.

1.9.2 تقدير فترة الثقة واختبار فرضيات المتقاربة

لبناء فترات الثقة أو اختبارات الفرضيات على β نحن في حاجة الى تعرف قانون $(\widehat{\beta} - \beta)$.
في الحالة العامة، لا توجد افتراضات بشأن قانون εلذلك يمكننا أن نسعى للحصول على معرفة تقريبية بها عندما n كبيرة جدا.

نظريات الحد

لكل عينة حجمها n. سيتم تحديد الاعتماد في n باستخدام$^{(n)}$ وبالتالي: Y، X، $\widehat{\beta}$، $\widehat{\varepsilon}$، $\widehat{\sigma^2}$
تصبح على التوالي : $Y^{(n)}$، $X^{(n)}$،$\widehat{\beta}^{(n)}$، $\widehat{\varepsilon}^{(n)}$، $\widehat{\sigma^2}^{(n)}$

نظرية 9.2.1.

إذا كان: A: $X^{(n)'}X^{(n)}/n \longrightarrow_{n\longrightarrow+\infty}$ معرفة موجبة .

إذن:

$$\sqrt{n}(\widehat{\beta}^n - \beta) \rightarrow N(0, \sigma^2 A^{-1})$$

ملاحظة 9.2.1.

σ^2 كونها غير معروفة نقدرها ب$\widehat{\sigma^2}^n$.

2.9.2 اختبار المتغيرات المضافة:

إذا كان هناك نظرية اقتصادية تعتمد على نموذج معين، النموذج يقول أن المتغير التابع المراد تفسيره يتأثر بعدد من المتغيرات المستقلة K

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon_i$$

ونظرية أخرى تستخدم نموذج آخر يقول أن هذا النموذج ناقص وهناك متغيرات إضافية تؤثر على المتغير التابع كما يلي:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 \varepsilon_i$$

أي أن النظرية الثانية تضيف مجموعه أخرى إلى النموذج السابق.
نظريتين متعارضتين مثلا هناك نظرية الاستهلاك محددة بالدخل هذه نظرية الدخل الدائم .
لكن هناك عوامل أخرى تؤثر على الاستهلاك مثل السعر والثروة والعادات، يمكن اختبار النموذجين كما يلي:
فرضية العدم:

$$H_0 = \beta_3 = \beta_4 = 0$$

الفرضية البديلة: فرضية العدم غير صحيحه

المطلوب هو اختبار النظرية الثانية إذا رفضنا فرضية العدم معناه أن النموذج الثاني هو أفضل من النموذج الأول. النموذج الأول هو النموذج المقيد أي النموذج الذي يساوي فيه المعاملات $H_0 = \beta_3 = \beta_4 = 0$ [2] فإذا قبلت فرضية العدم معناه أننا قبلنا النموذج الأول.

$$F = \frac{SSE_R - SSE_\varepsilon}{SSE_\varepsilon / n - k}$$

حيث تمثل SSE_{ε} مجموع مربعات البواقي الغير مقيدة.
SSE_R مجموع مربعات البواقي المقيدة و r عدد القيود المفروضة على فرضية العدم.

جدول تحليل التباين لمعادلة الانحدار ANOVA

ان تحليل مجموع المربعات الصغرى الى مجموع مربعات البواقي ومجموع مربعات الانحدار. الغرض من هدا التحليل اختبار معنوية مجموع مربعات الانحدار وهدا يدخل ايضا في اختبار معنوية العامل β. ونمثل هدا التحليل في جدول تحليل التباين[1]:

متوسط المربعات	مجموع المربعات	درجة الحرية	التباين
مجموع المربعات الانحدار	$\mathrm{SSR} = \sum \hat{\mathrm{y}}^2 = \hat{\beta} \sum \mathrm{xy} = \hat{\beta}^2 \sum \mathrm{x}^2$	$k-1=2-1$	$\mathrm{SSR}/1$
مجموع المربعات البواقي	$\mathrm{SSE} = \sum \mathrm{u}^2 = \sum \mathrm{y}^2 - \sum \hat{\mathrm{y}}^2$	$n-k=n-2$	$\mathrm{SSE}/(\mathrm{n}-2)$
مجموع المربعات الاجمالي	$\mathrm{SST} = \sum \mathrm{y}^2 = \sum \hat{\mathrm{y}}^2 + \sum \mathrm{u}^2$	$n-2=3$	$\mathrm{F} = \frac{\mathrm{SSR}}{\mathrm{SSE}/\mathrm{n}-\mathrm{k}}$

جدول 2.1: ANOVA، جدول تحليل التباين

انتهى

الفصل الثالث

تحليل الانحدار باستخدام الحزمة الإحصائية SPSS

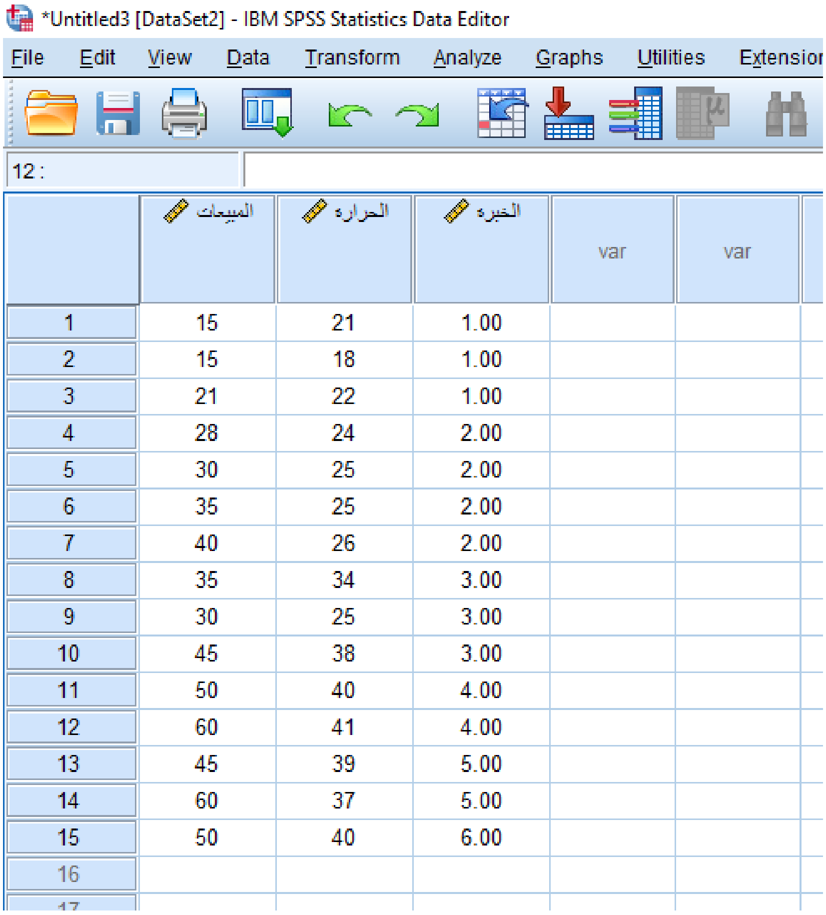
مثال 0.3.1. لدراسة العلاقة بين المبيعات اليومية لعدد 02 محل تجاري في يوم معين من سلعة معينة ودرجات الحرارة المسجلة في منتصف اليوم وسنوات الخبرة للبائع الذي يقوم بالخدمة. تم تجميع النتائج في الجدول التالي:

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
1	40	50

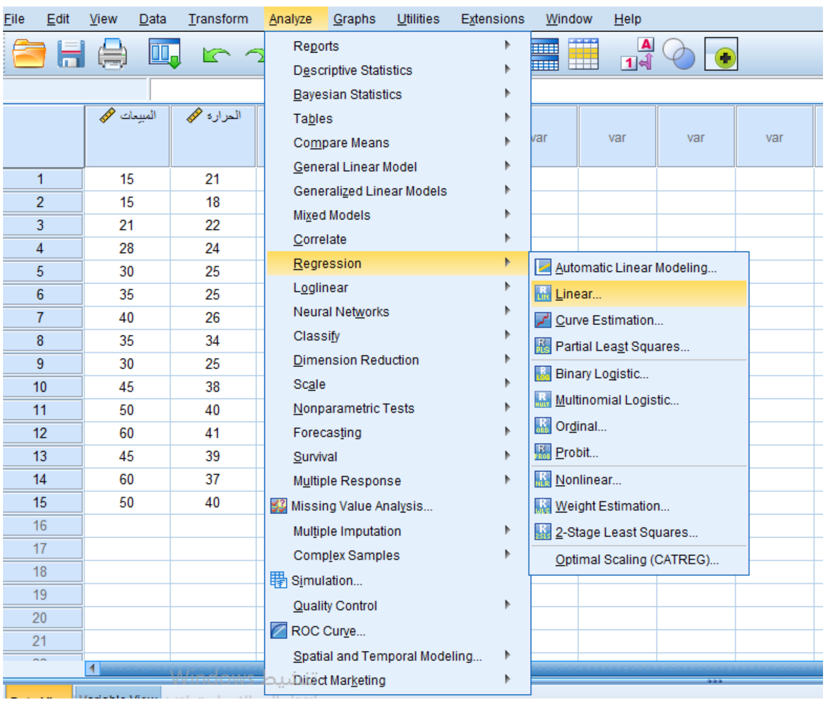
سنفترض أن المبيعات هي المتغير التابع وأن سنوات الخبرة ودرجه الحرارة هي المتغيرات المستقلة.

كيفية استخدام الحزمة spss في الحصول على:
الانحدار العياري أو القياسي Standard

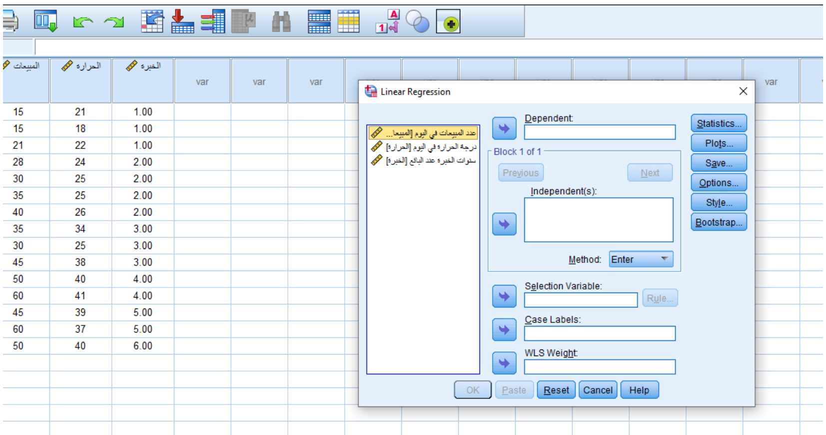
أو الآني (المتزامن) . Simultaneous
باستخدام بيانات الملف " regression.sav " للحصول على الانحدار العياري نتبع الخطوات التالية:



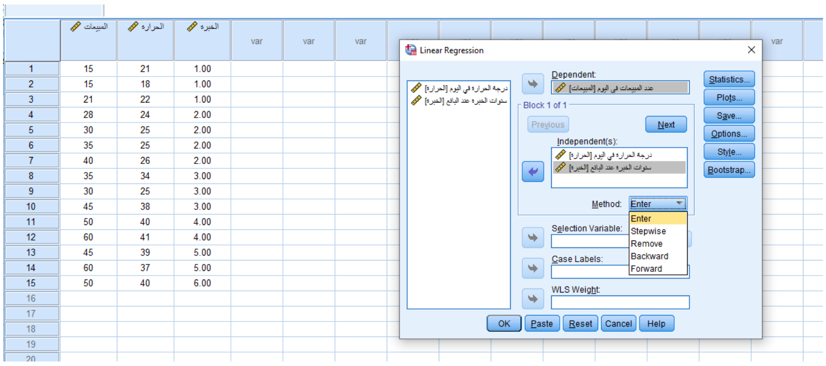
- من قائمة Analyze نختار regression.
- تظهر قائمة منسدلة تحتوي على عدد من أنواع الانحدار سوف نختار هنا منها Linear .



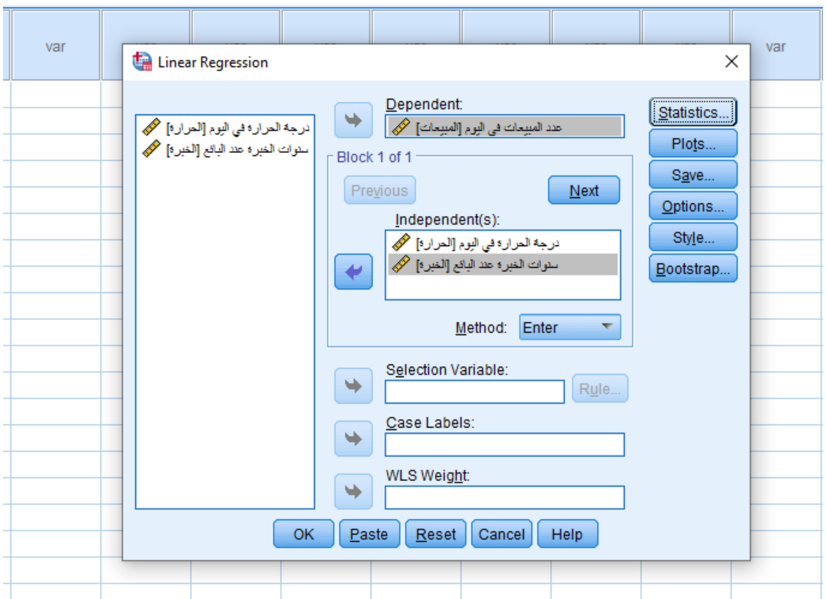
- تظهر شاشة جديدة بعنوان Linear regression.



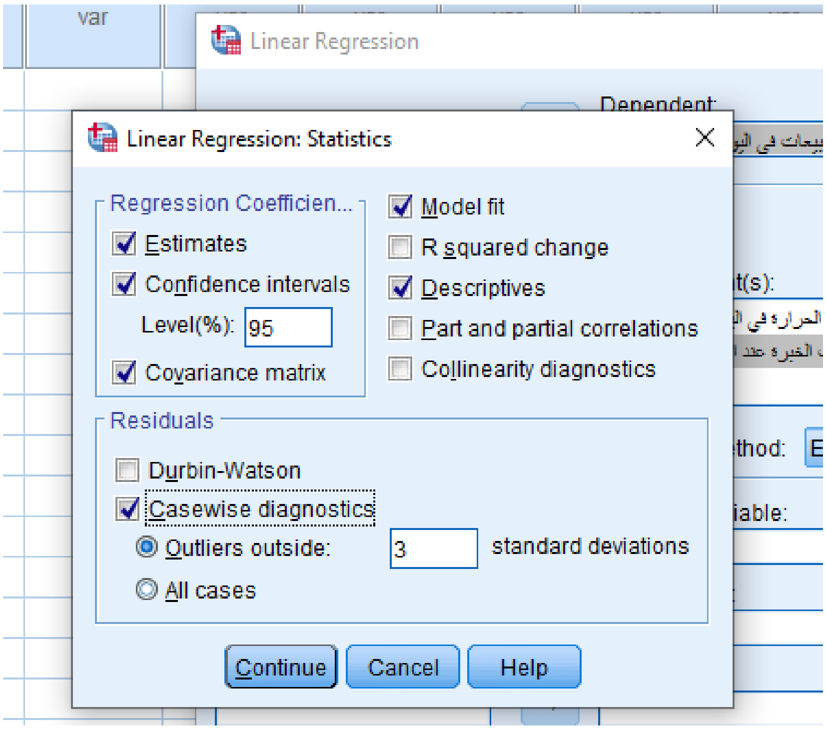
- ننقل المتغير التابع المبيعات لخانة Dependent .
- ننقل المتغيرات المستقلة الحرارة و سنوات الخبرة لخانة Independent .



- نختار نوع الانحدار من خانة Method الطريقة العيارية وهي Enter.



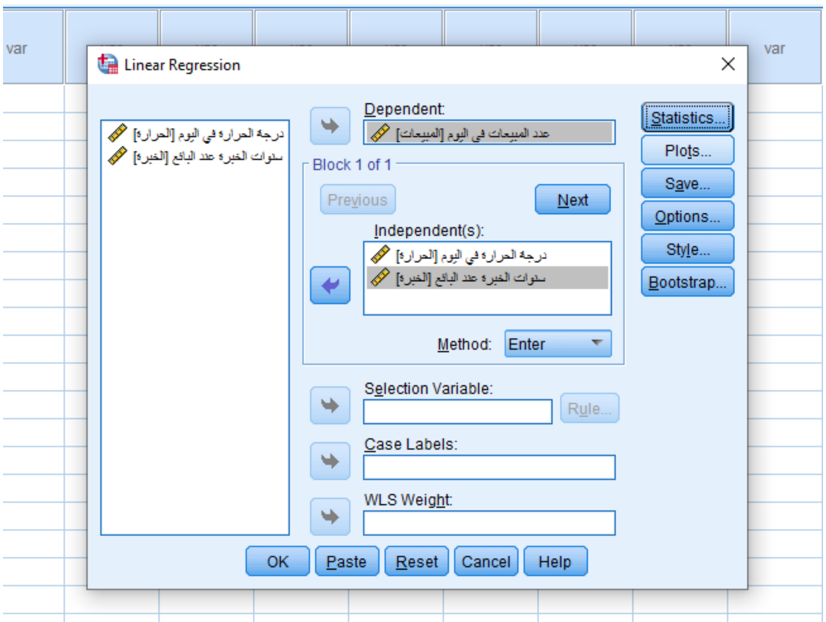
- نضغط على الأمر Statistics تظهر شاشة جديدة بعنوان : Linear regression statistics.



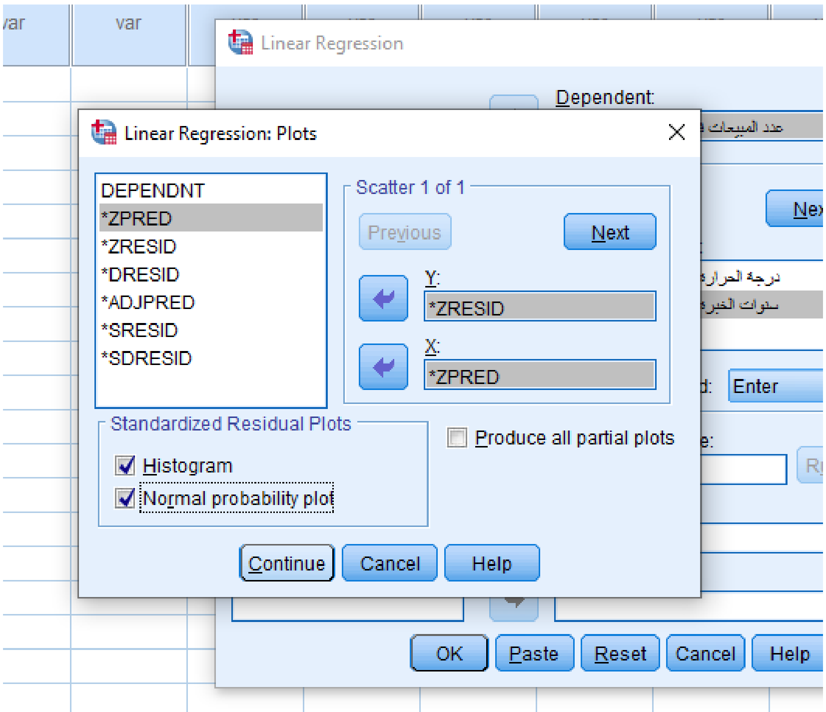
تتأكد أن Casewise diagnostics Model fit, Estimates مختارة ويمكن أيضا إضافة اختيارات أخرى.

- نختار Continue لنعود للشاشة السابقة.

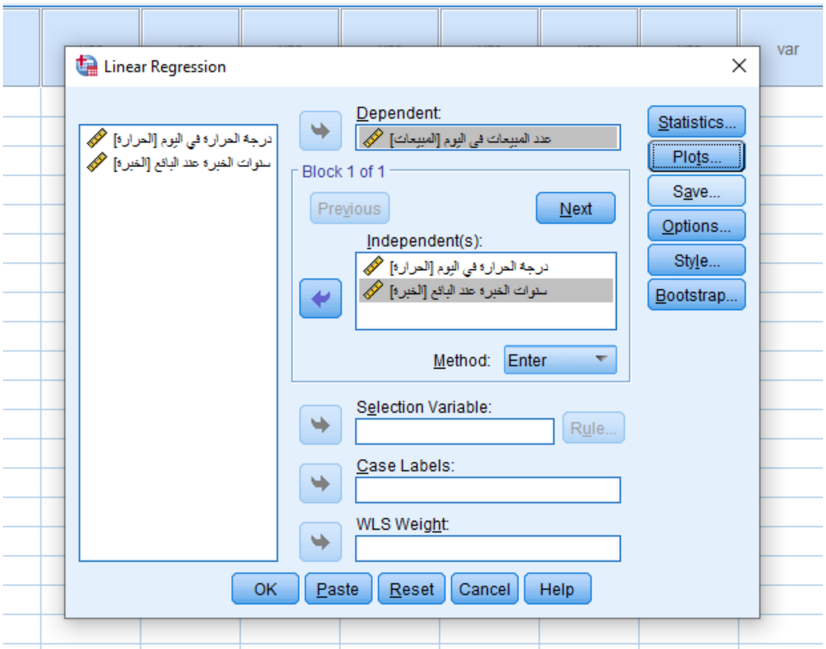
- نضغط على الأمر Plot فتظهر شاشة جديدة بعنوان .Linear regression: Plots



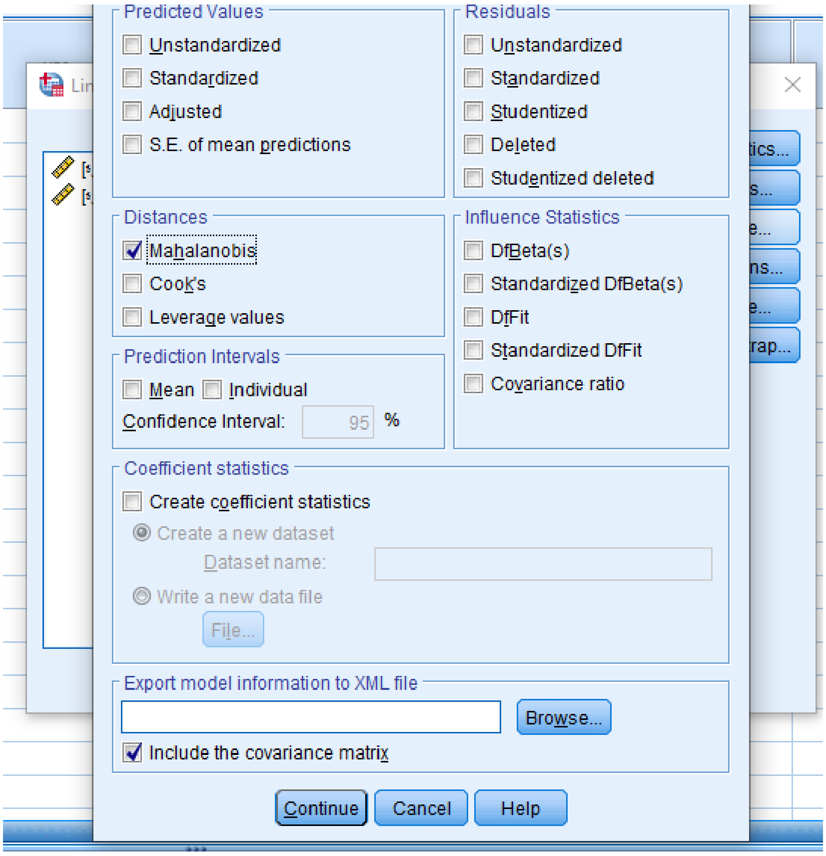
- ننقل ZRESID * للمستطيل المقابل ل: Y وايضا ZPRED * للمستطيل المقابل ل: X.
- من قائمة Standardized Residual Plots نختار كلا من Histogram و Normal probability plot.



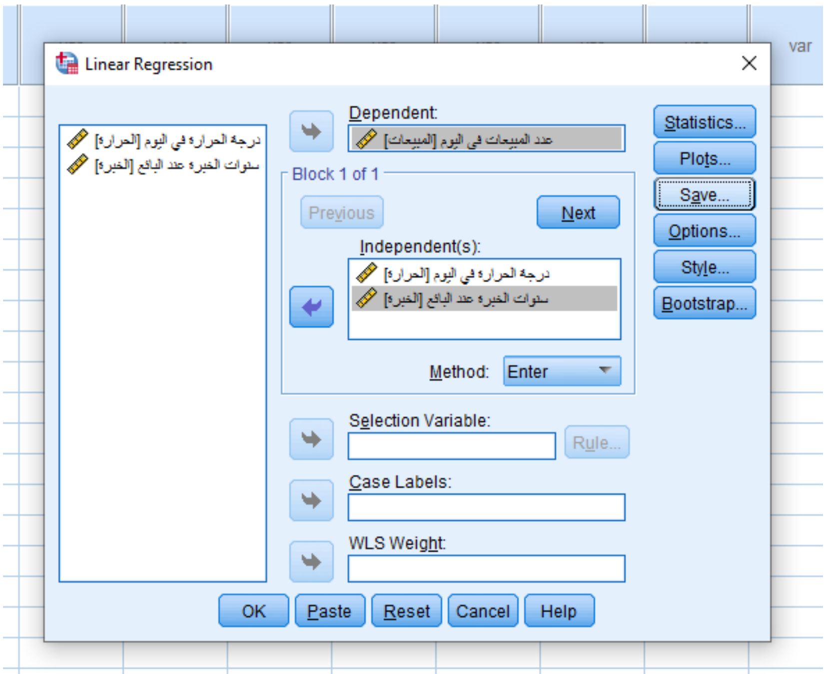
- نضغط على Continue فنعود للشاشة السابقة.



- نضغط على الأمر Save تظهر 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.

[a]Entered/Removed Variables

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 Removed/Entered Variables ويحتوي على أسماء المتغيرات التي دخلت في معادلة الانحدار وهما متغيرين (الحرارة و سنوات الخبرة) والمتغيرات التي استبعدت من الدخول في المعادلة 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), البائع عند الخبرة سنوات, اليوم في الحرارة درجة

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 .
- تعيين خطأ التقدير Std. Error of the estimate وهو هنا 6.45606 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.

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B 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
	عدد الخبرة سنوات البائع	1.907	2.362	.208	.807	.435	-3.240-	7.054

a. Dependent Variable: اليوم في المبيعات عدد

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:

سنوات الخبرة 1.907 + الحرارة 1.275 + 6.993- = المبيعات

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 $sig = 0.016$ 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 عدد الخبرة سنوات	اليوم في الحرارة درجة
1	Correlations	البائع عدد الخبرة سنوات	1.000	-.888-
		اليوم في الحرارة درجة	-.888-	1.000
	Covariances	البائع عدد الخبرة سنوات	5.581	-.952-
		اليوم في الحرارة درجة	-.952-	.206

a. Dependent Variable: اليوم في المبيعات عدد

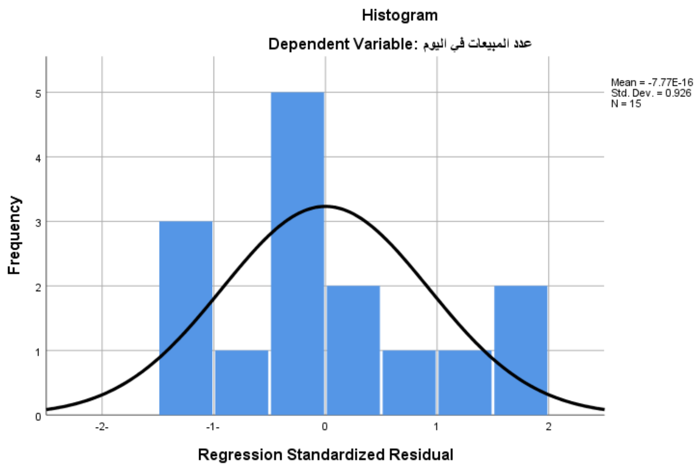
الجدول السابع: بعنوان Correlations Coefficient ويعطى مصفوفة الارتباط بين المتغيرات المستقلة ونلاحظ كما سبق انه لا يوجد ارتباط تام بين المتغيرات وبعضها الآخر.

Residuals Statistics[a]

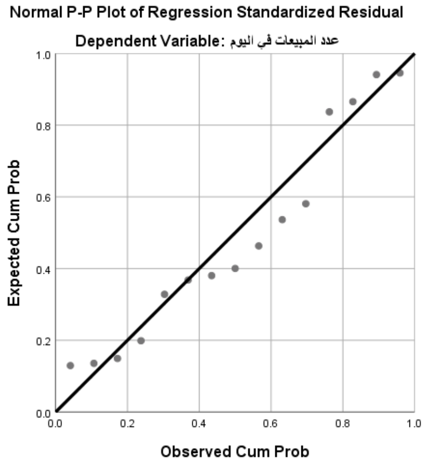
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	17.86	55.44	37.27	13.222	15
Std. Predicted Value	-1.468-	1.374	.000	1.000	15
Standard Error of Predicted Value	1.945	4.247	2.801	.667	15
Adjusted Predicted Value	18.72	59.65	37.52	13.420	15
Residual	-7.255-	10.294	.000	5.949	15
Std. Residual	-1.129-	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: عدد المبيعات في اليوم

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

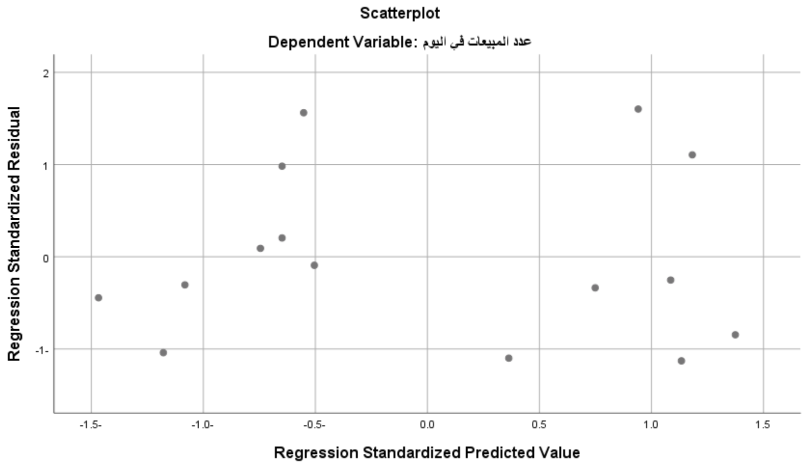


شكل 3.1: هو المدرج التكراري ويستخدم للتعرف هل البيانات تتوزع حسب التوزيع الطبيعي أم لا؟

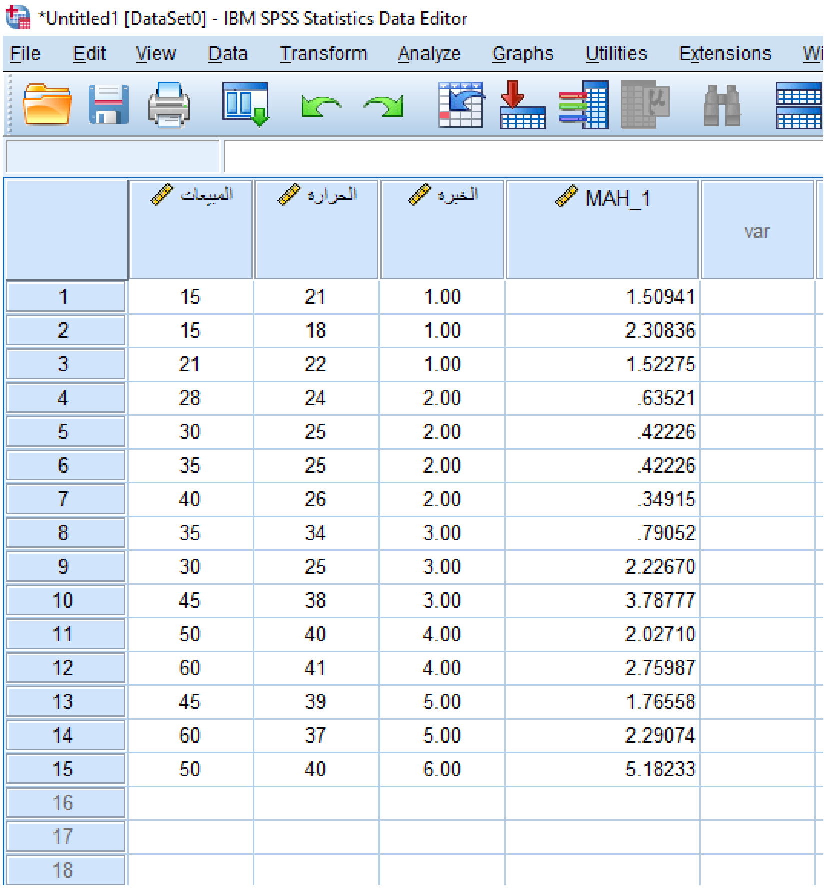


شكل 3.2: يختبر هل البواقي تتبع التوزيع الطبيعي ام لا؟

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



شكل 3.3: يمثل شكل الانتشار للبواقي مع القيم المتوقعة ومنه يتضح عدم وجود نمط معين للنقاط في الشكل وهذا يتسق مع شرط الخطية.



	المبيعات	الحراره	الخبرة	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.

- 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 اج راء 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.

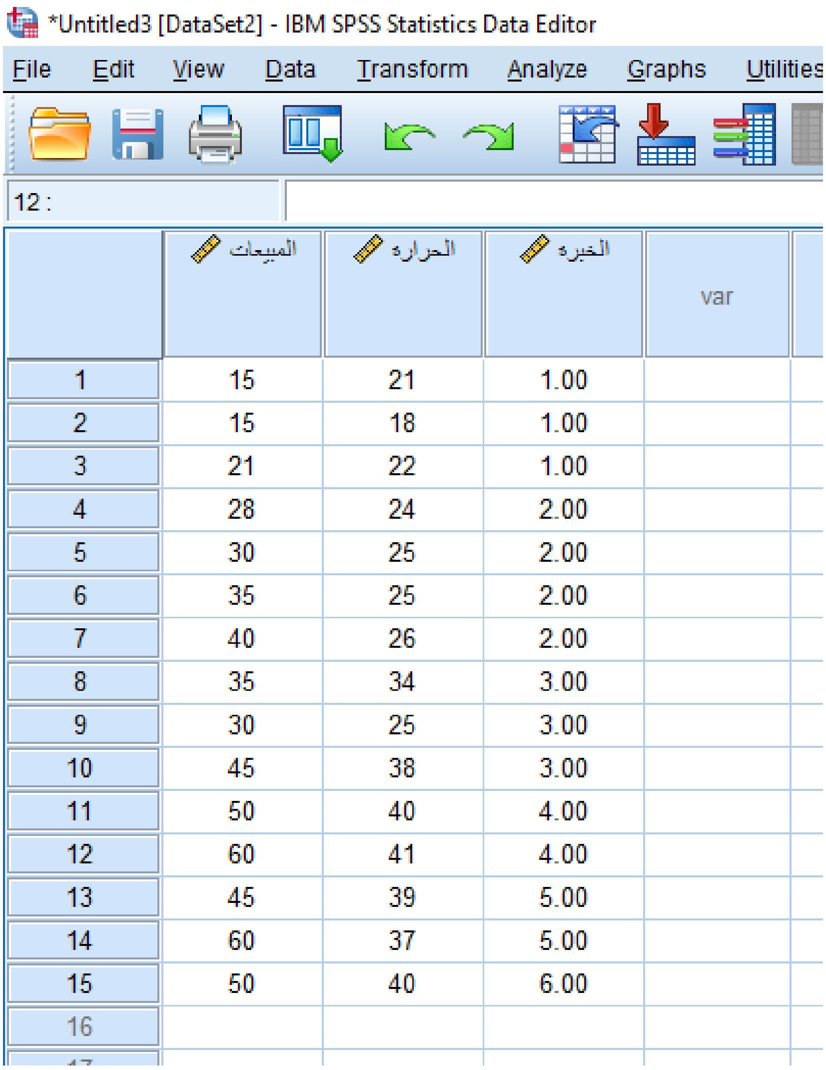
مما سبق يمكن القول وبصفه عامه فان الانحدار التدريجي يحقق بعض المزايا منها:

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،43-40)

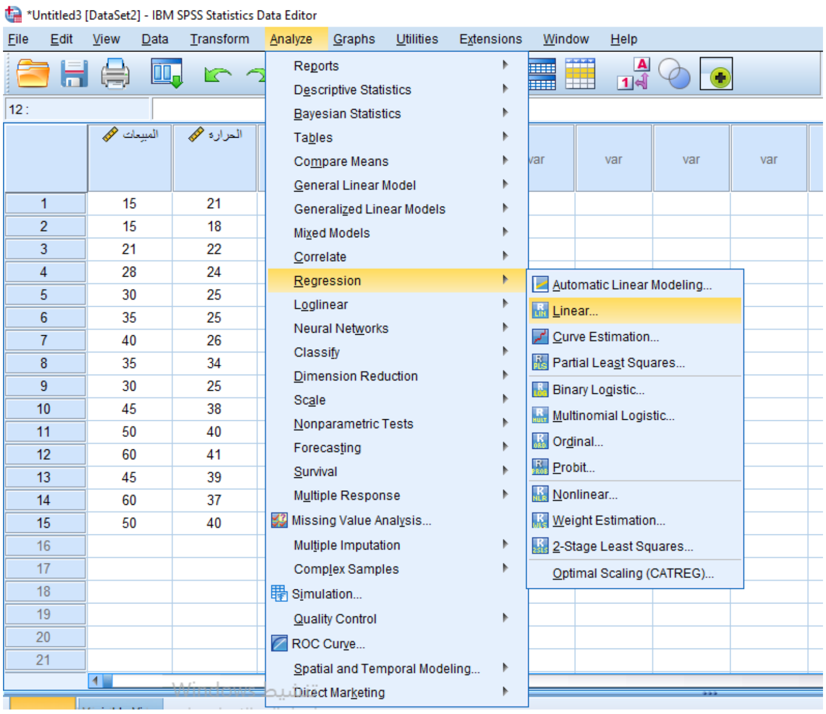
ولإجراء خطوات الانحدار التدريجي باستخدام SPSS نمر بالخطوات التالية:

- نعرف بالمتغيرات وندخل البيانات على صفحة Spss

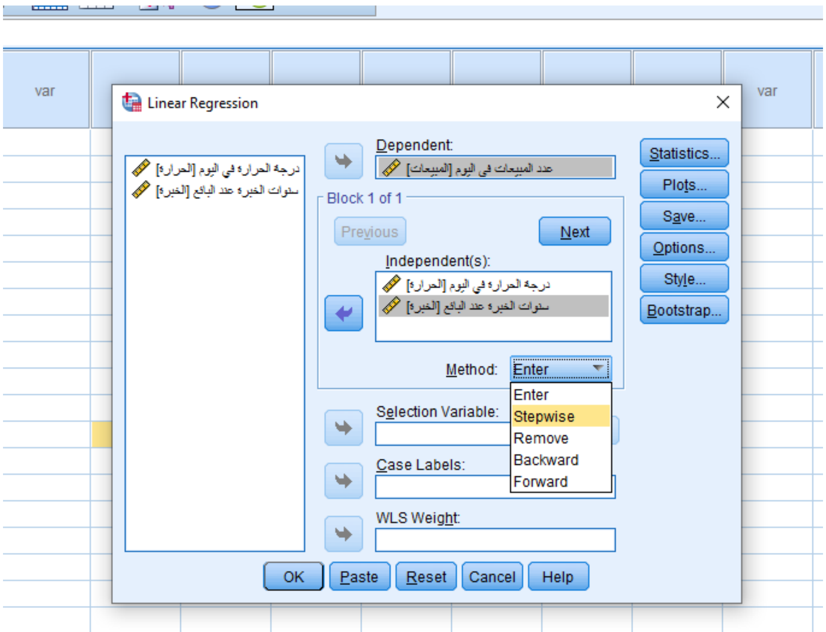


	المبيعات	الحرارة	الخبرة	var
1	15	21	1.00	
2	15	18	1.00	
3	21	22	1.00	
4	28	24	2.00	
5	30	25	2.00	
6	35	25	2.00	
7	40	26	2.00	
8	35	34	3.00	
9	30	25	3.00	
10	45	38	3.00	
11	50	40	4.00	
12	60	41	4.00	
13	45	39	5.00	
14	60	37	5.00	
15	50	40	6.00	
16				

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

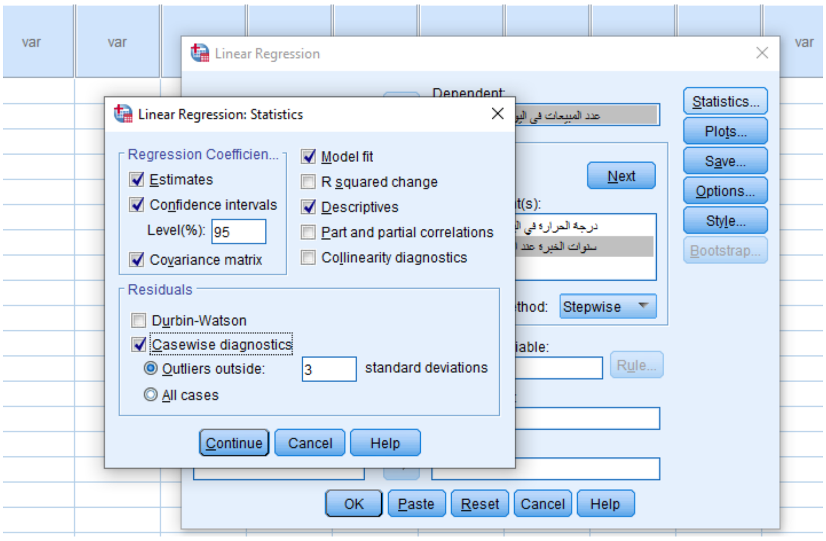


- تظهر شاشه جديده بعنوان Linear regression.

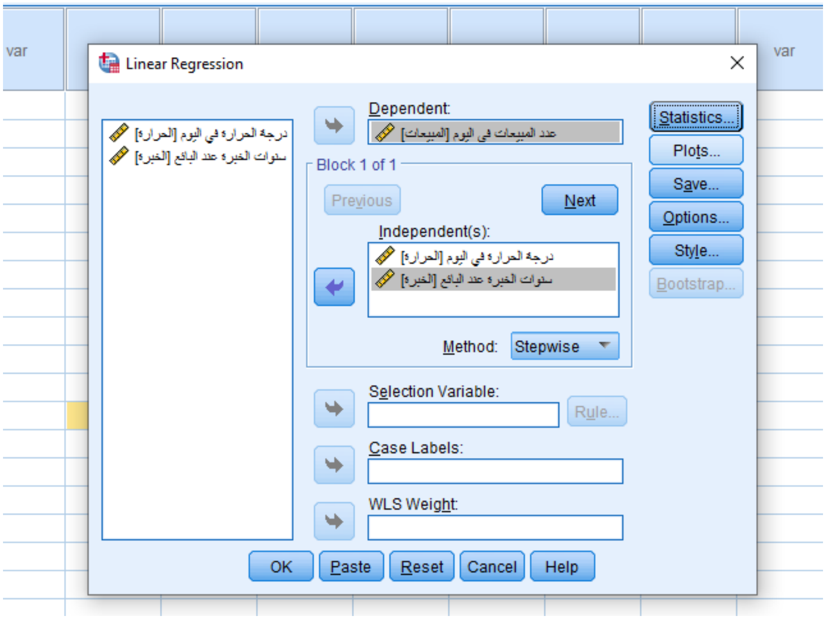


- ننقل المتغير المبيعات لخانة Dependent وننقل المتغيرين المستقلين الحرارة، سنوات الخبرة لخانة Independent .
- من قائمة Method نختار Stepwise.

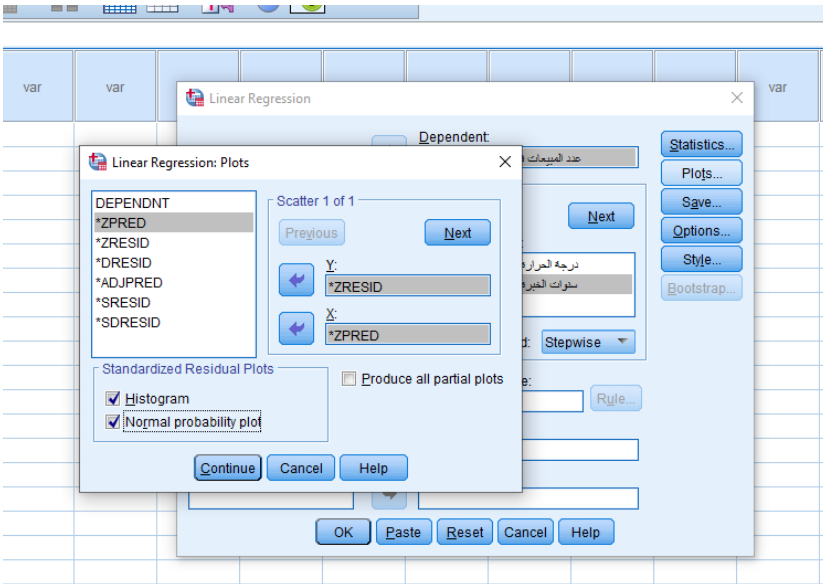
- نضغط على الامر Statistics تظهر شاشه جديده بعنوان Linear Regression: Statistics.



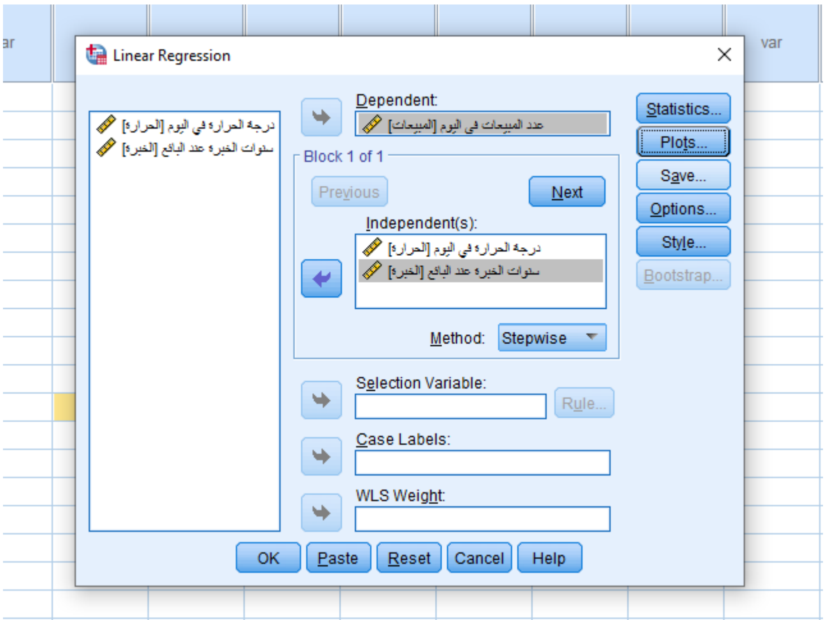
- تتأكد أن Casewise diagnostics, Estimates, Model fit ، مختارة ويمكن أيضا إضافة اختيارات أخرى.
- نختار Continue لنعود للشاشة السابقة.



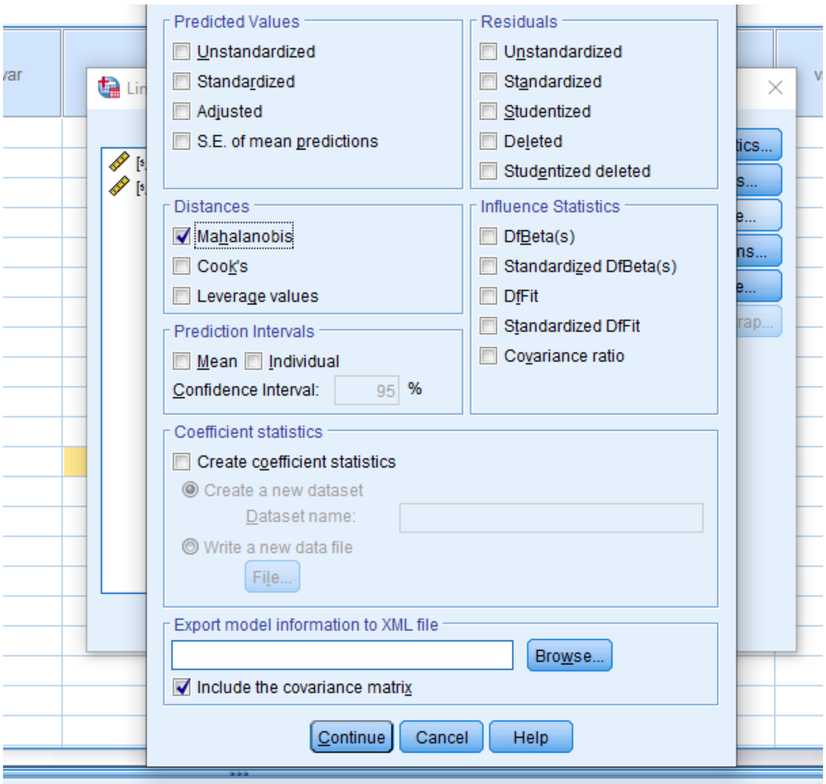
- نضغط على الامر Plot فتظهر شاشه جديده بعنوان Linear Regression: Plots



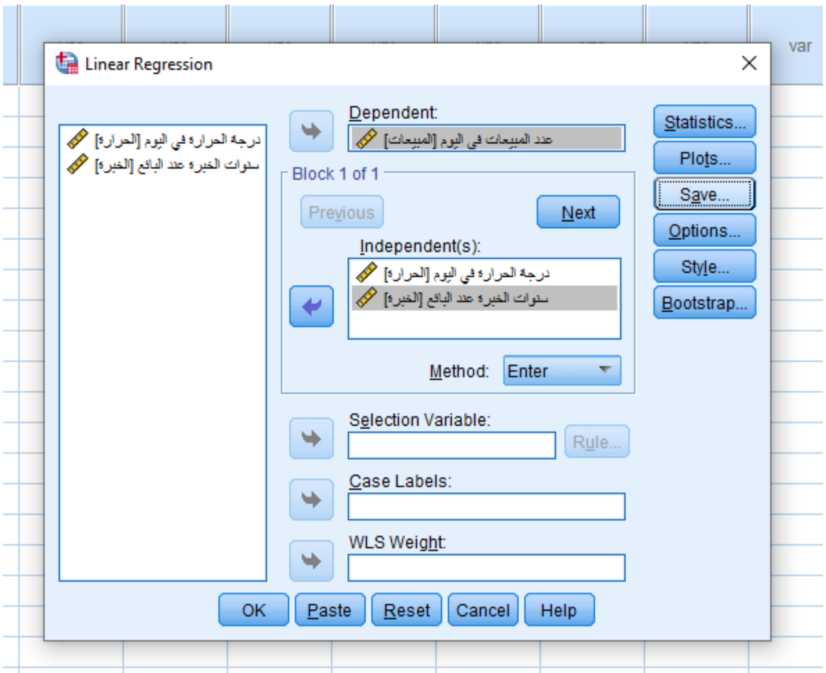
- ننقل *ZRESID للمستطيل المقابل ل: Y وأيضا *ZPRED للمستطيل المقابل : X
- من قائمة Standardized Residual Plots نختار كلا من Histogram, Normal probability plot
- نضغط على Continue فنعود للشاشة السابقة.



- 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			

الجدول الثالث: بعنوان ANOVA وهو تحليل التباين للانحدار ويتضح أن الانحدار معنوي حيث $sig = 0.000$ 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
	في الحرارة درجة اليوم	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.

$$\text{درجة الحرارة}*1.600 + -11.271 = \text{المبيعات}$$

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

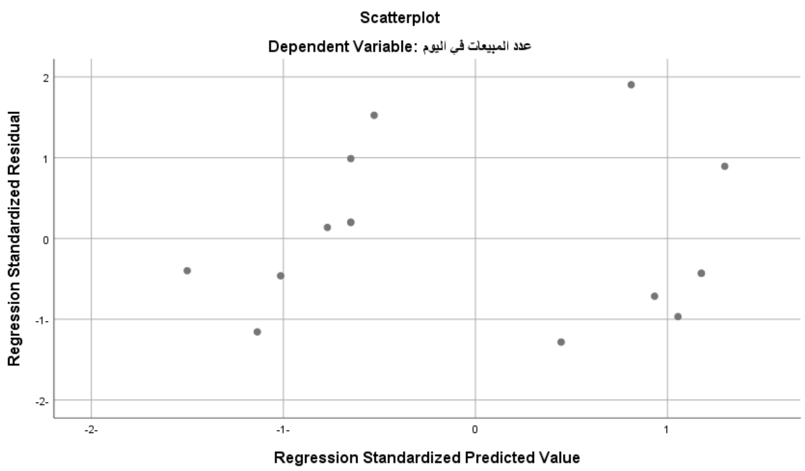
الجدول الخامس: بعنوان Excluded Variables ويعرض بيانات تخص المتغير الذي استبعد وهو سنوات الخبرة.

Residuals Statistics[a]

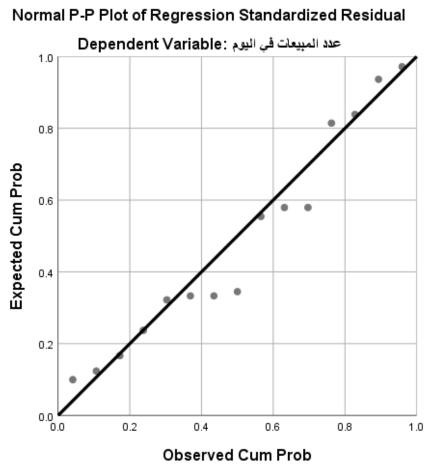
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	17.53	54.33	37.27	13.149	15
Std. Predicted Value	-1.501-	1.298	.000	1.000	15
Standard Error of Predicted Value	1.803	3.024	2.289	.356	15
Adjusted Predicted Value	18.28	53.28	37.35	13.062	15
Residual	-8.134-	12.066	.000	6.109	15
Std. Residual	-1.283-	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

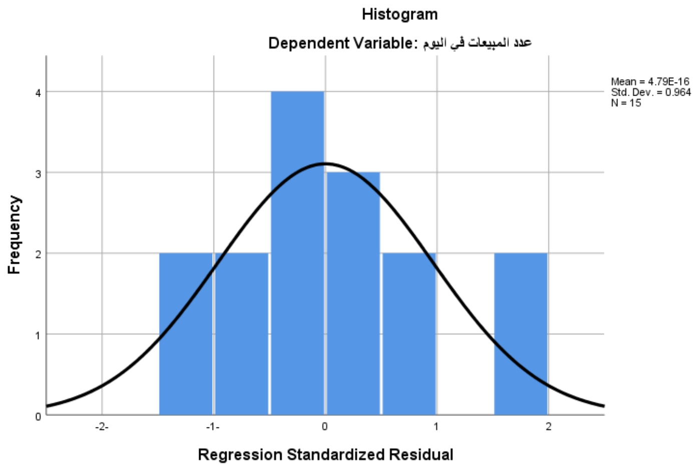
الجدول السادس: بعنوان Statistics Residuals ويعرض بيانات خاصه بتحليل البواقي.



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؟

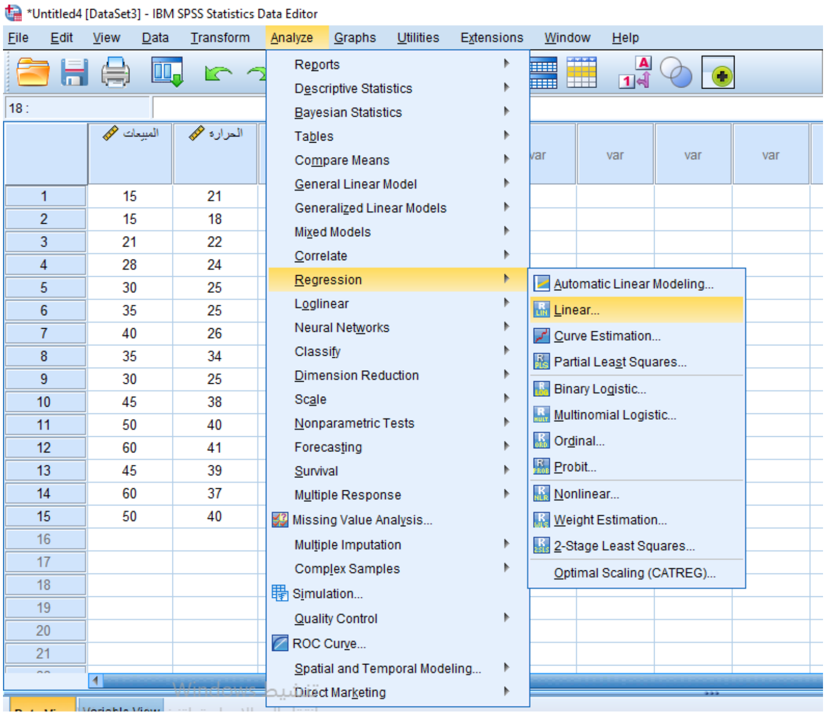
الانحدار الهرمي: Hierachical Regression

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:

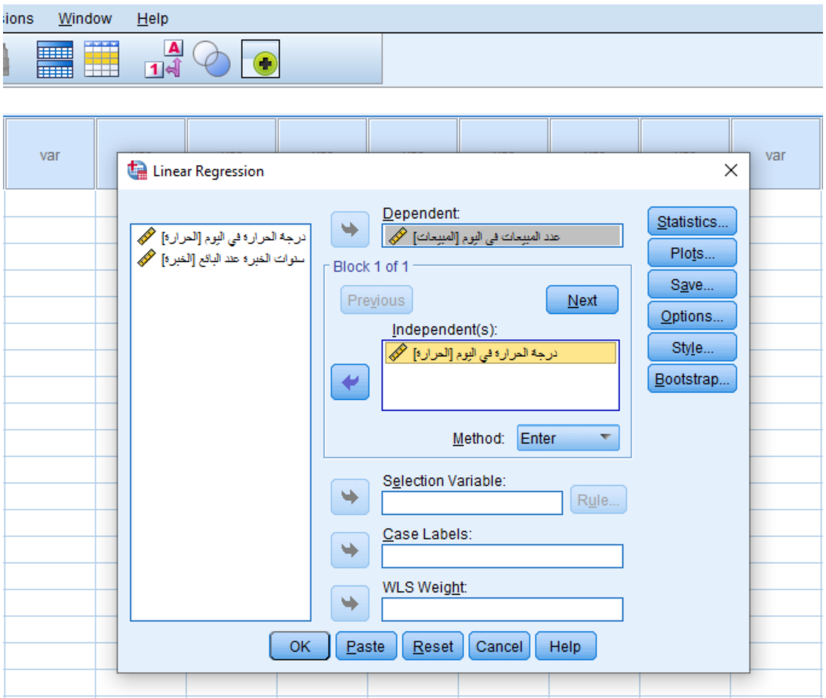
- نعرف بالمتغيرات وندخل البيانات على صفحة Spss

	المبيعات	الحرارة	الخبرة	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				

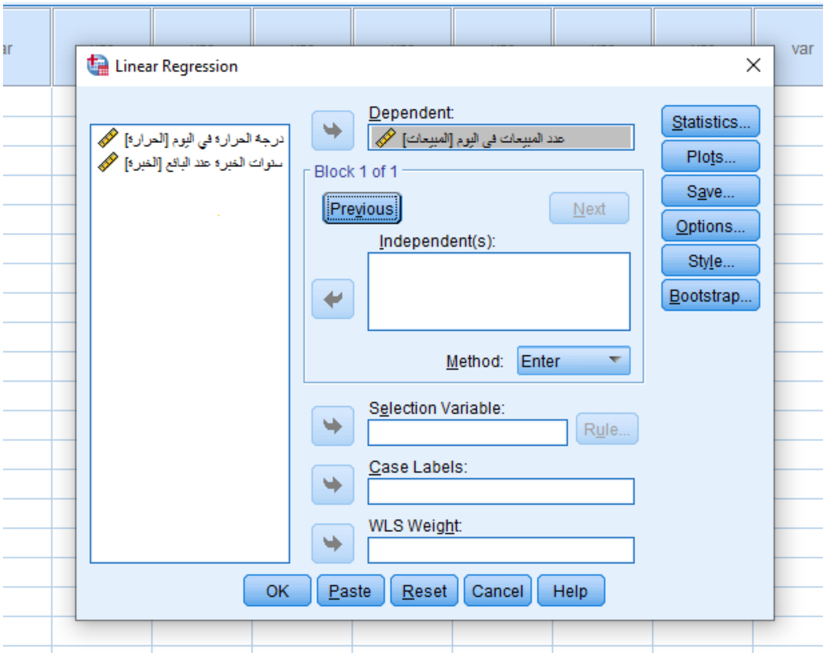
- من قائمة Analyze نختار Regression .
- من القائمة المنسدلة نختار Linear .



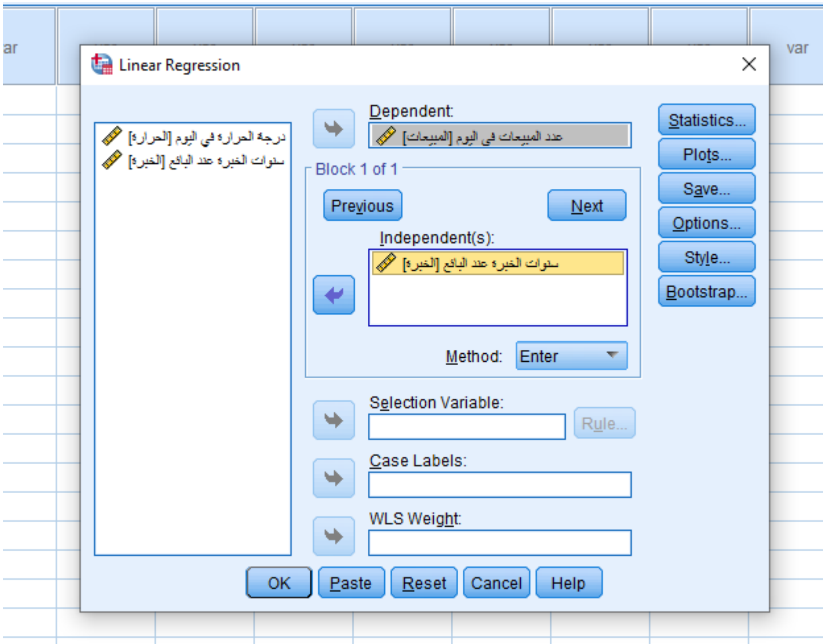
- تظهر شاشة بعنوان Linear Regression.



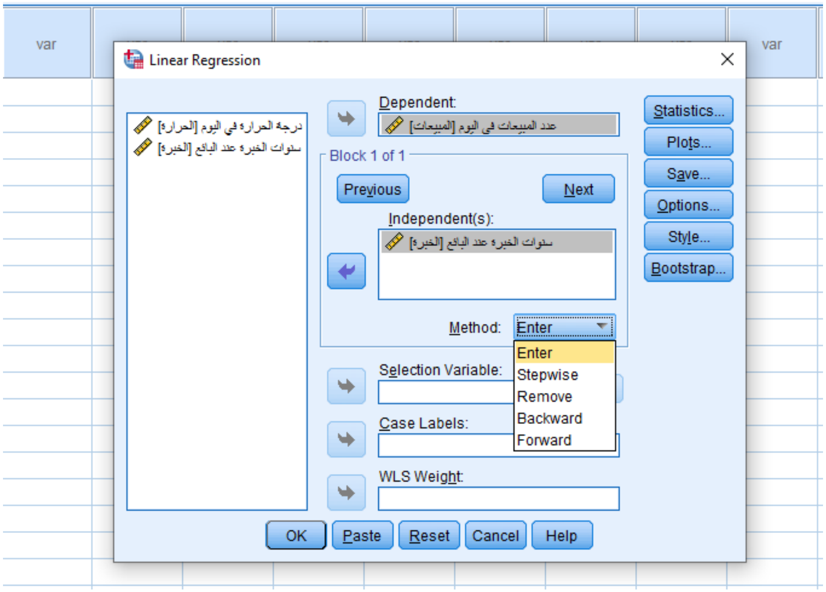
- ننقل المتغير التابع المبيعات لخانة Dependent وننقل أول المتغيرات المستقلة درجة الحرارة لخانة Independent(s) لاحظ فوق كلمة previous وجود Block 2 of 2
- نضغط على الامر Next.



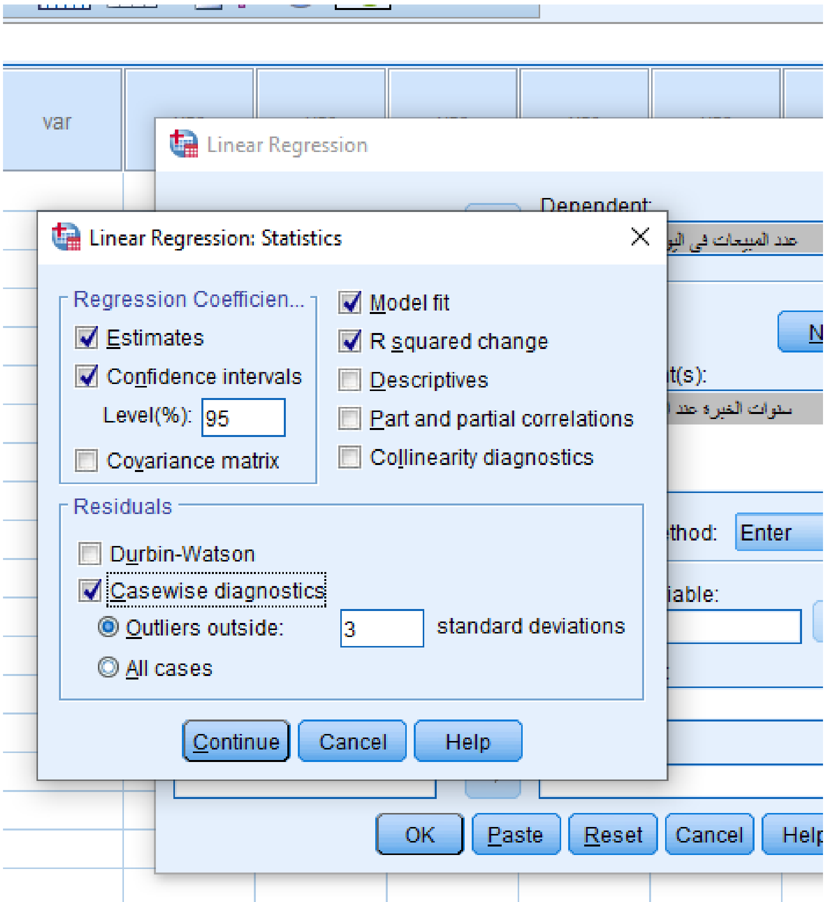
- ننقل المتغير المستقل الثاني السنوات لخانة Independent.



- سوف نختار الطريقة العيارية Method Enter.

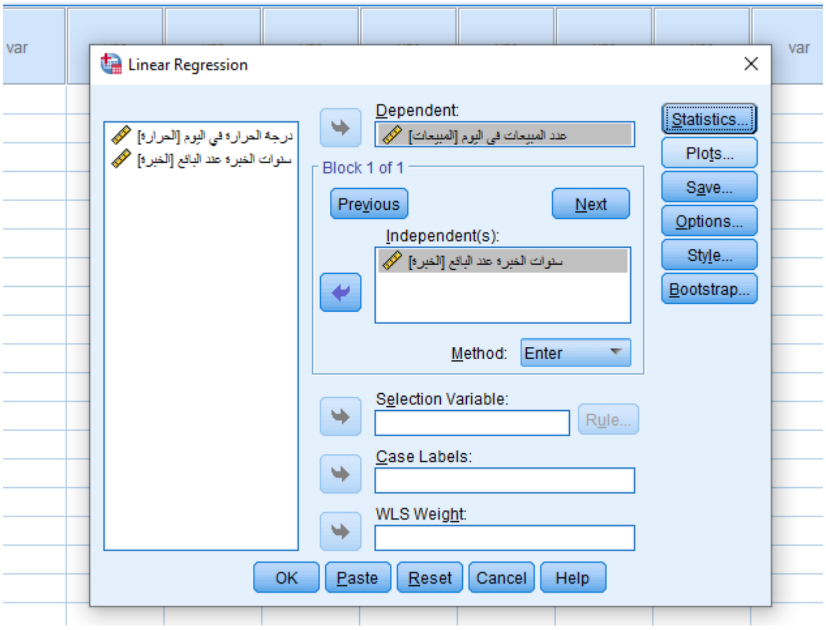


- نضغط على الأمر Statistics تظهر شاشة جديدة بعنوان Linear Regression: Statistics.

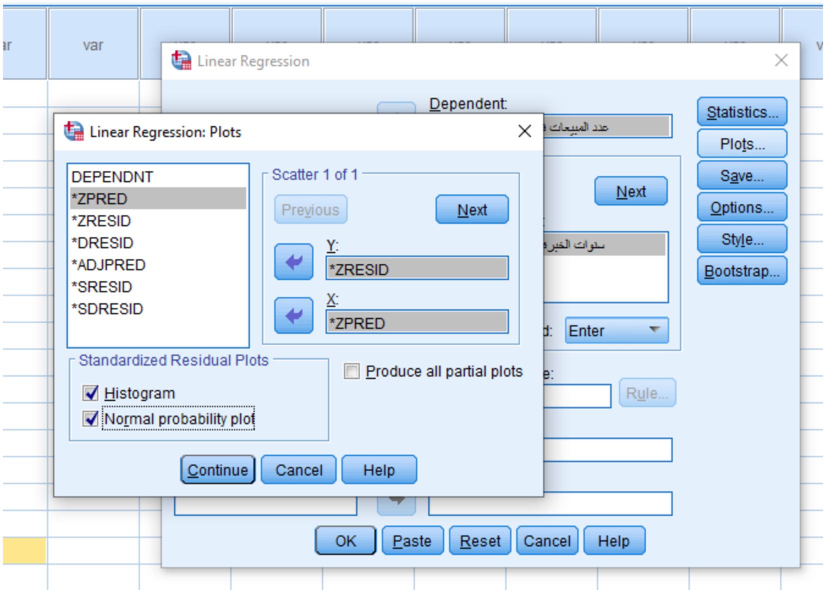


- Model fit, Estimates diagnostics, R Casewise Squared Change، مختاره ويمكن أيضا إضافة اختيارات أخرى.

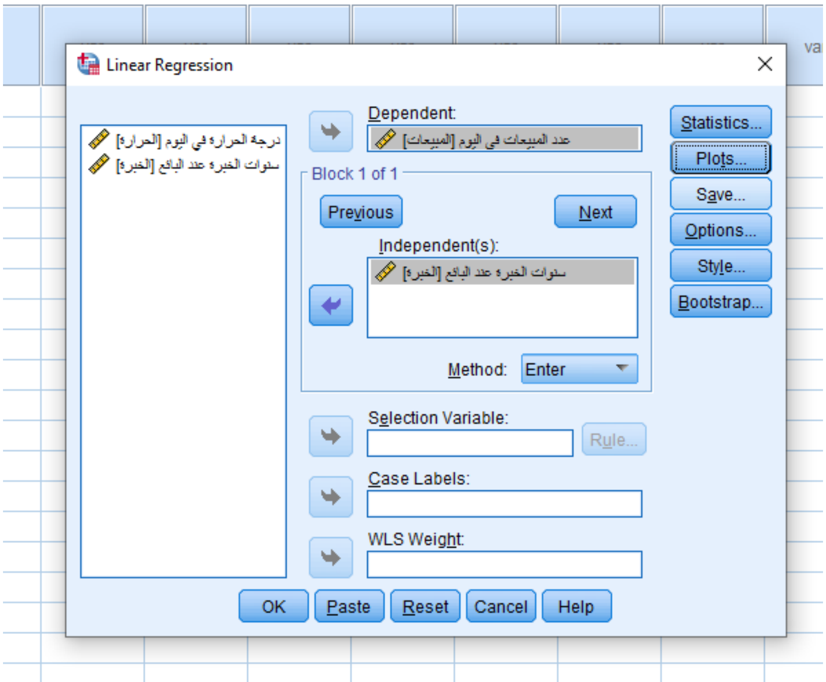
- نختار Continue لنعود للشاشة السابقة.



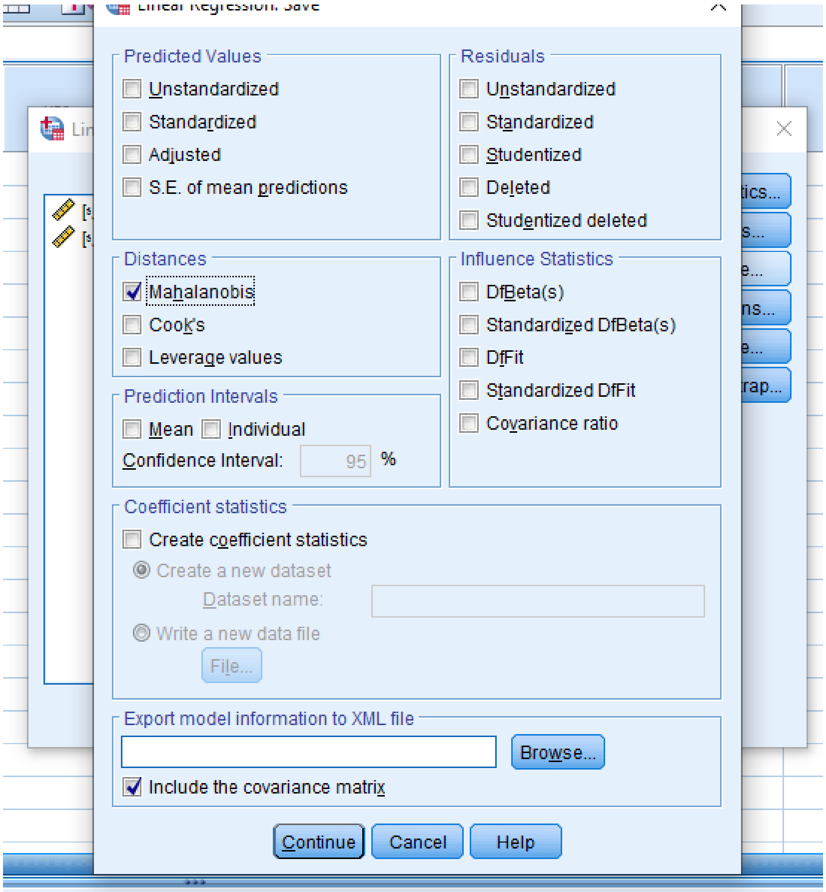
- نضغط على الامر Plot فتظهر شاشه جديده بعنوان Linear Regression: Pltos



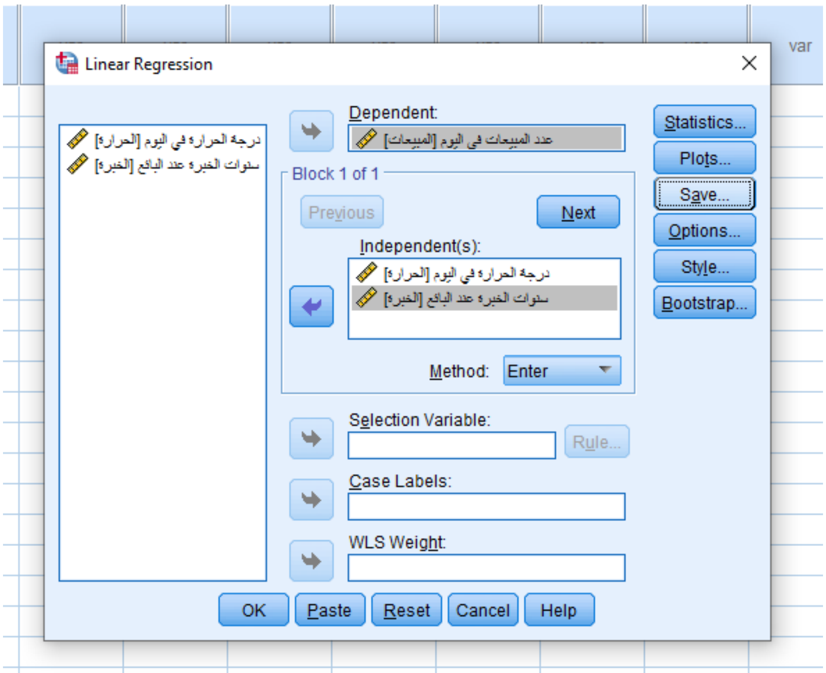
- ننقل *ZRESID للمستطيل المقابل ل : Y وأيضا *ZPRED للمستطيل المقابل : X.
- من قائمة Standardized Residual Plots نختار كلا من، Normal probability plot، Histogram
- نضغط على Continue فنعود للشاشة السابقة.



- 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:

Entered/Removeda Variables

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

a. Dependent Variable: عدد المبيعات في اليوم

b. All requested variables entered.

الجدول الاول: بعنوان Entered/Removed Variables 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), درجة الحرارة في اليوم

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

c. Dependent Variable: عدد المبيعات في اليوم

الجدول الثاني: بعنوان Model Summary ويعطي ملخص عن النموذج لاحظ كلا من R ، Sig, F, 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), اليوم في الحرارة درجة, البائع عند الخبرة سنوات

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

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			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 معن وي.
- الخطوة الثانية: تم ادخال سنوات الخبرة مع درجات الحرارة فكانت غير معنوية.

وعلى ذلك يكون أفضل خط مقدر هو:

درجة الحرارة*1،600 + 11،271- = المبيعات

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

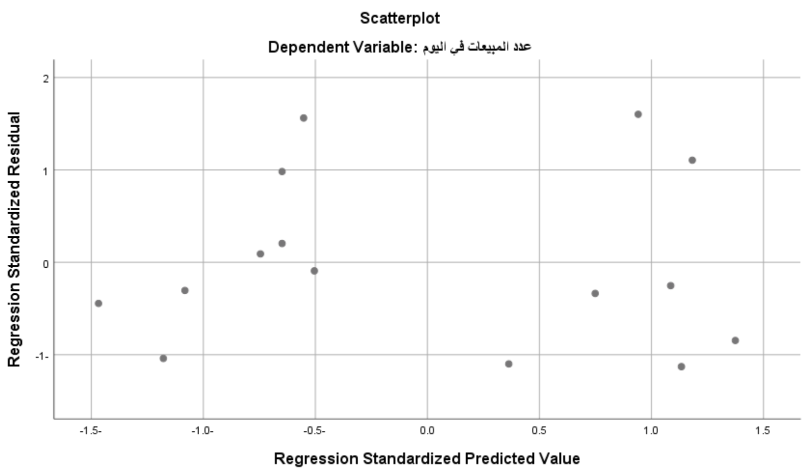
الجدول الخامس: بعنوان Excluded Variables وهو خاص بالمتغيرات المستبعدة.

Residuals Statistics[a]

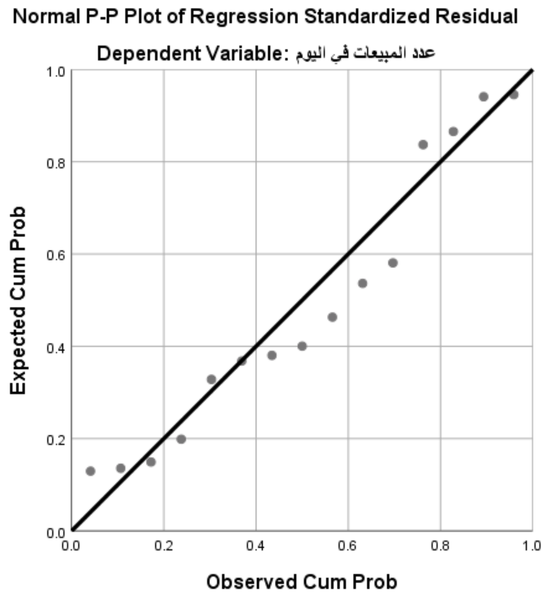
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	17.86	55.44	37.27	13.222	15
Std. Predicted Value	-1.468-	1.374	.000	1.000	15
Standard Error of Predicted Value	1.945	4.247	2.801	.667	15
Adjusted Predicted Value	18.72	59.65	37.52	13.420	15
Residual	-7.255-	10.294	.000	5.949	15
Std. Residual	-1.129-	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

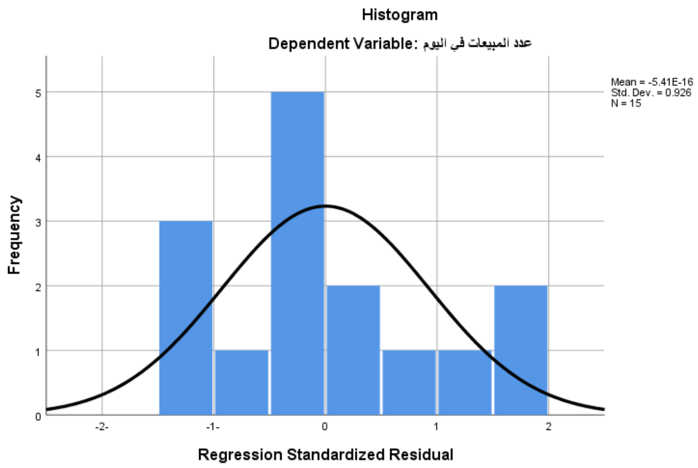
الجدول السادس: بعنوان Statistics Residual ويستخدم لتحليل البواقي.



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؟

خاتمة

ختاما، تجدر الاشارة الى أننا عبر هذا البحث قد قمنا بجمع وتفسير وتحليل أهم المعلومات الخاصة بمضوع الانحدار الخطي المتعدد، وتقدير معالمه باستخدام طريقة المربعات الصغرى. وفي الأخير قمنا بإعطاء مثال وتحليله باستخدام البرنامج الإحصائي spss . من أجل معرفة العلاقة بين المبيعات و درجة الحرارة و سنوات الخبرة تم استخدام نموذج الانحدار الخطي المتعدد .والذي اعتبرت فيه متغيرات الحرارة و سنوات الخبرة كمتغيرات تفسيرية ومتغير المبيعات كمتغير تابع , أظهرت نتائج نموذج الانحدار انه معنوي وذلك من خلال قيمة sig=0.000 وهي أصغر من مستوي المعنوية (0.05), وتفسر النتائج أن هناك علاقة بين المتغيرات المستقلة والمتغير التابع .

أتمنى من الله أن تكون نالت كلماتي إعجابكم، بتوفيق منه تعالى تمكنت من اجتياز جملة من العراقيل والصعوبات لإتمام هذا العمل والذي لا يخلو من النقائص فلله الكمال والتمام.

المراجع العلمية

[1] تومي صالح، مدخل لنظرية القياس الاقتصادي، الجزائر: ديوان المطبوعات الجامعية، 1999.

[2] السعيد بومنجل، الدليل الإحصائي للطالب، الجزائر: ديوان المطبوعات الجامعية، 2000

[3] عبد الرحمان بن محمد سليمان أبو عمه، أنور أحمد محمد عبد الله، الإحصاء التطبيقي: مطابع جامعة الملك سعود، 1995

[4] مقدمة في الاستدلال الإحصائي، الإسكندرية: دار الطباعة والنشر والتوزيع، 1993

[5] محمد شيخي، طرق الاقتصاد القياسي محاضرات وتطبيقات، دار الحامد للنشر والتوزيع، الطبعة الأولى، الاردن، 2012

[6] أسامة ربيع أمين 2008 ، التحليل الاحصائي للمتغيرات المتعددة باستخدام برنامج spss .

[7] مركز الدراسات الاحصائية والسياسية العامة. www.sabr-ssp.com

[8] عبد الفتاح مصطفى محمد . الانحدار المتعدد ، كلية العلوم ، جامعة مصر.

المراجع باللغة الأجنبية

[9] Mekkaoui Zineb, Analyse Des déterminants De la Production Du Blé en Algérie Cas Des wilayas Tiaret, Sétif et Médéa L'échantillon 1990-2009, mémoire de Master, spécialité Économie appliquée et ingénierie financière, Faculté des sciences économiques et de gestion et Sciences commerciales, Université Abdel Rahman Mera Bejaia, 2013.

[10] Aza¨is, J.-M., Bardet, J.-M. (2006). Le modèle linéaire par l'exemple. Régression analyse de la variance et plans d'expérience, Dunod, Paris.

[11] Dodge, Y., Rousson, V. Analyse de régression appliquée, Dunod.

[12] Kleinbaum, D. et al. Applied regression analysis and other multivariate methods.

ملخص

الانحدار الخطي هو اجراء احصائي راسخ، أصبحت نماذجه وسيلة مثبتة للتنبؤ بالمستقبل بشكل علمي وموثوق فيه ودلك لأن خصائص هده النمادج مفهومة جيدا ويمكن تدريبها بسرعة كبيرة بواسطة وسائل الاعلام الالي الحديثة مثل برنامج spss، ويعود ذلك لبساطتها وتوفرها على صيغة رياضية سهلة التفسير نعبر عنها بمعادلة خطية تتضمن متغير واحد أو أكثر من المتغيرات المستقلة التي تقدر قيمة المتغير التابع. يناسب الانحدار الخطي طرق احصائية كثيرة حسب فرضيات دراسة الحالات. تستخدم عادة طريقة "المربعات الصغرى" لتقدير المعلمات سواء في الانحدار الخطي البسيط او المتعدد كما أن هناك حاسبات عن طريق برامجها تمكننا من إكتشاف أفضل خط ملائم لمجموعة من البيانات المقترنة. ثم تقوم بتقدير قيمة المتغير التابع Y بدالة المتغير المستقل X، مع دراسة جوانب أخرى مهمة في تحليل النمودج كالإرتباط الخطي المتعدد (multicollinearity) الدي يعبر عنه بمصفوفة التغاير(matrix covariance) .

الكلمات المفتاحية : الإنحدار الخطي، التنبؤ، تقدير المعلمات ، الإرتباط المتعدد ، مصفوفة التغاير

Abstract

Linear regression is a well-established statistical procedure, and its models have become a proven way to predict the future scientifically, reliably and massagely because the characteristics of this model are well understood and can be trained very quickly by modern automated media such as spss This is due to its simplicity and availability of an easy-to-interpret mathematical formula expressed in a linear formula that includes one or more independent variables that value the dependent variable. Linear regression fits many statistical methods according to case study hypotheses. The "micro squares" method is commonly used to estimate parameters in both simple and multiple linear regression and there are computers through its software that enable us to detect the best suitable line for a set of associated data. Then you estimate the value of the Y variable of the independent variable X, while studying other important aspects of growth analysis such as multiple linear correlation, expressed in the variability matrix..

Keywords: linear regression, prediction, parameter estimation, multiple correlation, variability matrix

Résumé

La régression linéaire est une procédure statistique bien établie, et ses modèles sont devenus un moyen éprouvé de prédire l'avenir scientifiquement, fiable et massagely parce que les caractéristiques de ce modèle sont bien compris et peut être formé très rapidement par des médias automatisés modernes tels que spss Cela est dû à sa simplicité et à la disponibilité d'une formule mathématique facile à interpréter exprimée dans une formule linéaire qui comprend une ou plusieurs variables indépendantes qui valorisent la variable dépendante. La régression linéaire correspond à de nombreuses méthodes statistiques selon des hypothèses d'études de cas. La méthode des "microcarrés" est couramment utilisée pour estimer les paramètres en régression linéaire simple et multiple et il y a des ordinateurs grâce à son logiciel qui nous permettent de détecter la ligne la plus appropriée pour un ensemble de données associées. Ensuite, vous estimez la valeur de la variable Y de la variable indépendante X, tout en étudiant d'autres aspects importants de l'analyse de la croissance tels que la corrélation linéaire multiple, exprimée dans la matrice de variabilité.

Mots-clés : régression linéaire, prédiction, estimation des paramètres, corrélation multiple, matrice de variabilité