

4

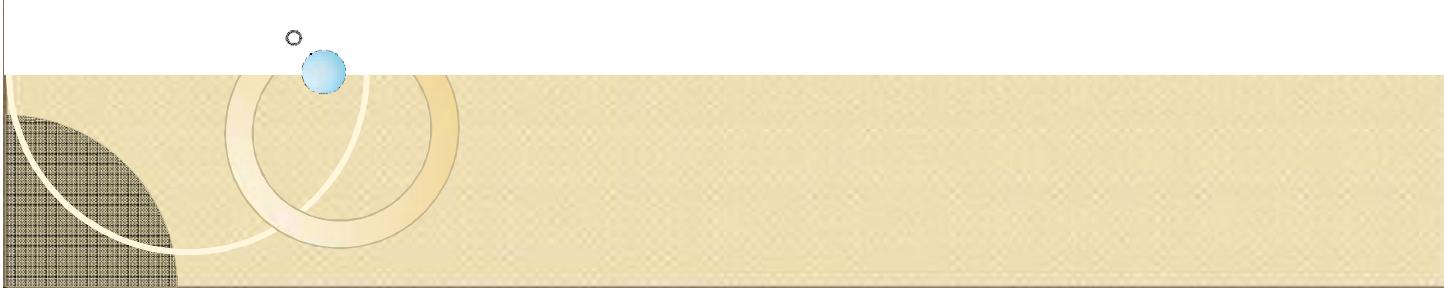
الباب الرابع: الارتباط والانحدار البسيط

Chapter 4: Correlation & Simple Linear Regression

889.7	+5.000
0	+1.500
0	+1.125
0	+1.062

ستتناول في هذا الفصل :

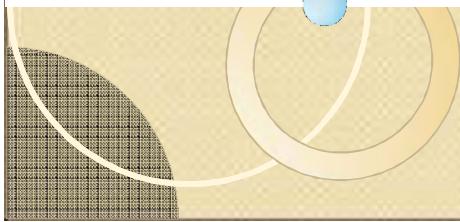
- (١) مفهوم الارتباط وأنواعه.
- (٢) طرق حساب معاملات الارتباط المختلفة.
- (٣) مفهوم الانحدار الخطى البسيط وتطبيقاته .



مقدمة عن الإرتباط

تقابلنا كثيراً في الحياة العملية موافق تضمن متغيرين (ظاهريين) وأكثر ويكون المطلوب معرفة ما إذا كان هناك علاقة بين هذه المتغيرات وما هو شكل هذه العلاقة ؟ وأيضاً كيفية التبرؤ بأحد هذين المتغيرين في حالة معرفتنا بالمتغير الآخر .

فكثيراً ما تجدين في بعض المجالات معادلة الطول مع الوزن فإذا أردت أن تعرفي الوزن المثالي الداخلي طولك في المعادلة ليظهر وزنك المثالي ، وقد توصلوا إلى هذه المعادلة أو إلى هذه الصيغة بدراسة العلاقة ما بين المتغيرين الطول والوزن على مجموعة من الأفراد.



Correlation

- الارتباط: هو تعين طبيعة وقوة العلاقة بين متغيرين أو عدمها
- معامل الارتباط Correlation Coefficient هو مؤشر هذه العلاقة
 - أول خطوه في تحديد طبيعة العلاقة هي رسم شكل الانشار

- إذا كان لدينا متغير ان فقط . المتغير X وهو متغير يتم تحديده من قبل الباحث أو الشخص الذي يقوم بالدراسة وهو يسمى بالمتغير المستقل Independent variable

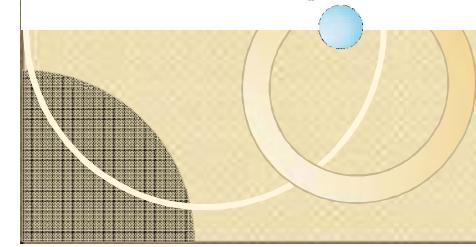
- يرافق المتغير X متغير آخر Y ويسمى بالمتغير التابع dependent variable وهو متغير تابع لأن نتيجته غير محددة وتعتمد على قيم المتغير المسنجل

الارتباط



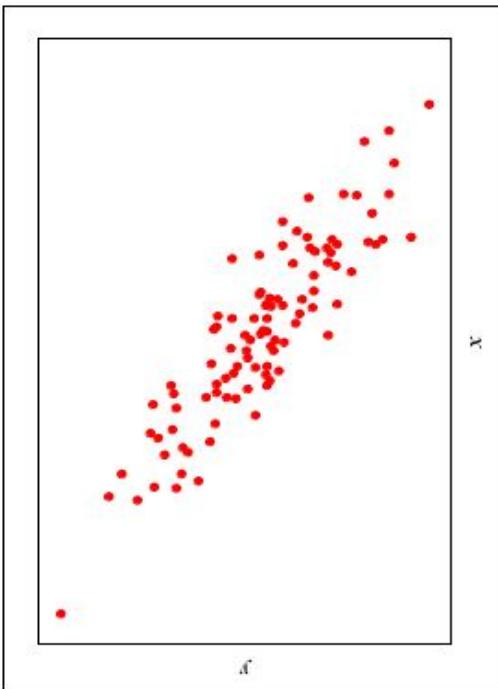
الارتباط السالب (المعكسي) (Correlation Negative) بأنه علاقة بين متغيرين (y, x) بحيث إذا تغير أحد المتغيرين فإن الآخر يبتعد في الاتجاه المضاد.

الارتباط الموجب (الطريدي) (Correlation Positive) بأنه علاقة بين متغيرين (x, y) بحيث إذا تغير أحد المتغيرين فإن الآخر يبتعد في نفس الاتجاه.

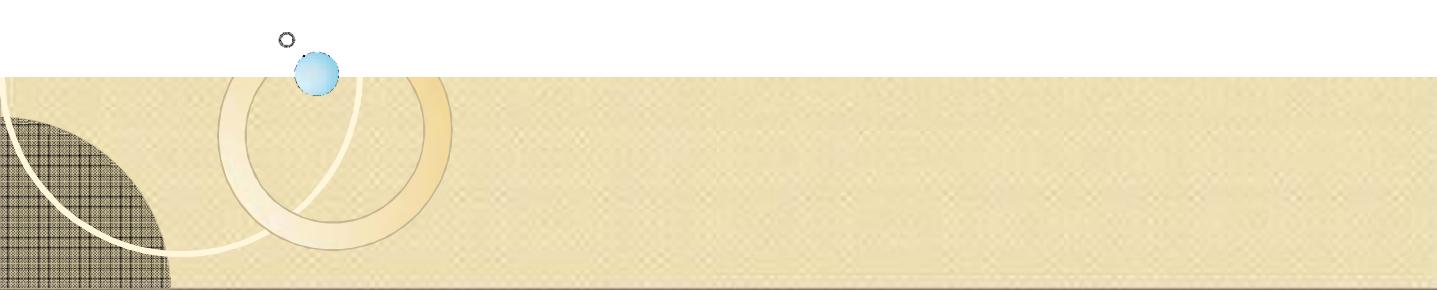
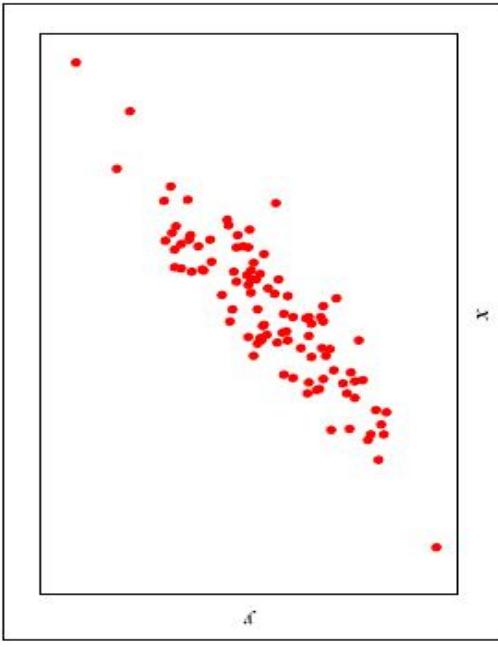


شكل الانتشار الخاص بالارتباط
متغيرين (ظاهريتين)

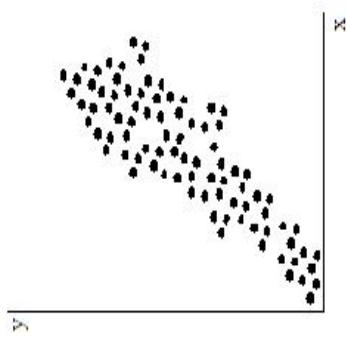
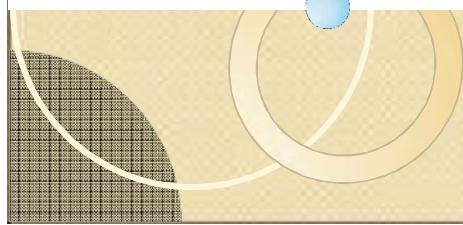
شكل الانتشار الخاص بالارتباط السالب
متغيرين (ظاهريتين)



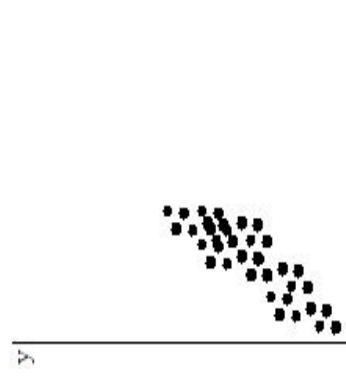
شكل الانتشار الخاص بالارتباط
الموجب (الطريدي)



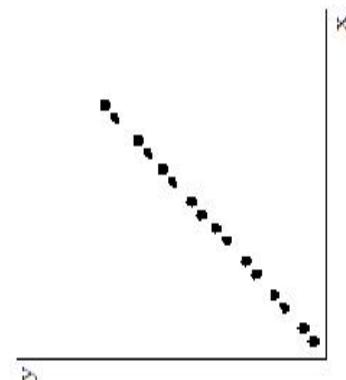
Scatter Plot



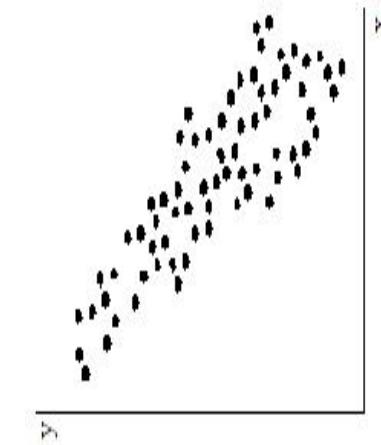
ارتباط طردی



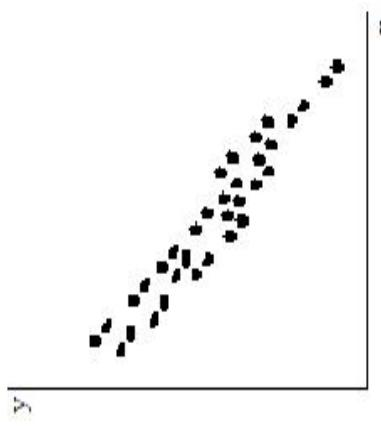
ارتباط طردی قوي



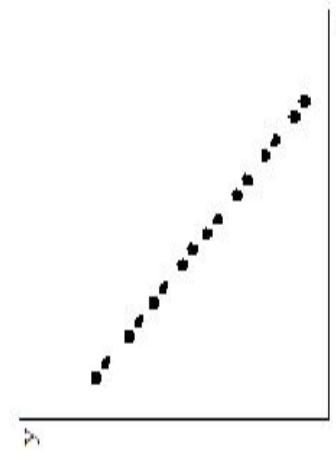
ارتباط طردی تمام



ارتباط عکسي



ارتباط عکسي قوي



ارتباط عکسي تمام

قياس الارتباط

- تستخدم معاملات الارتباط لقياس درجة الارتباط بين متغيرين (ظاهريين) .
- تعريف معامل الارتباط :
يعرف معامل الارتباط الذي يرمز له بالرمز r بأنه عبارة عن مقياس رقمي يقيس قوة الارتباط بين متغيرين ، حيث تترواح قيمته بين $-1 \leq r \leq +1$ و $(+)$ و $(-)$ ، أي أن r ≤ 0 يدل إشارة المعامل الموجبة على العلاقة الطردية ، و $r > 0$ يدل إشارة المعامل السالبة على العلاقة العكسية .
بينما تدل إشارة المعامل السالبة على العلاقة الطردية ، و $r < 0$ يعتمد ذلك على مستوى القياس (سمعي - ترتيبى - فترة - نسبي) للمتغيرات التي تبدو مرتبطة .

ثيماس لا ريزبها

والجدول التالي يوضح أنواع الارتباط واتجاه العلاقة وشكل الانتشار لتحليل نوع :

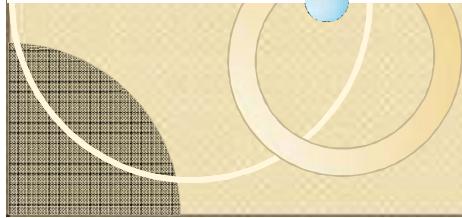
المعنى	قيمة معامل الارتباط
ارتباط طردي تام	+ 1
ارتباط طردي قوي	من 0.99 إلى 0.70
ارتباط طردي متوسط	0.69
ارتباط طردي ضعيف	0.49
لا يوجد ارتباط	0

وما قبل عن الارتباط الطردي ينطبق على الارتباط العكسي (مع وضع إشارة سالبة)

١- معامل بيرسون للارتباط الخطي

Pearson Linear Correlation Coefficient

- معامل بيرسون للارتباط الخطي من أكثر معاملات الارتباط استخداماً خاصة في العلوم الإنسانية و الاجتماعية .
- و مستوى القياس المطلوب عند تطبيق معامل بيرسون للارتباط هو أن يكون كلا المتغيرين مقياس فقرة أو نسبي أو بمعنى آخر أن تكون بيانات كلا المتغيرين (الظاهرتين) بيانات كمية .



١- معامل بيرسون لارتباط الخطي

- حساب معامل بيرسون للارتباط الخطي :

يمكن حساب معامل بيرسون بدلالة القراءات لبيانات المتغيرين X و Y ، باستخدام الصيغة التالية:

$$r_p = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{(n \sum x^2 - (\sum x)^2)(n \sum y^2 - (\sum y)^2)}}$$

حيث :

$\sum_{i=1}^n x_i y_i$: مجموع حاصل ضرب X في Y

$\sum x$: مجموع قيمة المتغير X

$\sum y$: مجموع قيمة المتغير Y

$\sum x^2$: مجموع مربعات قيمة المتغير X

$\sum y^2$: مجموع مربعات قيمة المتغير Y

- الدليل : الدرس وجود علاقة ايجابية بين مقدار المعلم ونواتج التعلم
 - المعلم : يمتلك مهارات تدريس جيدة، ويشعر بالسعادة والرضا في مهنته.

سُجّلت سُتّ فِرَاءَاتٍ تقريريةً لحجم الإنتاج وحجم صادرات النفط الخام بالمملكة العربية السعودية (بالمليار برميل) خلال عدّة سنوات كما يلي:

السنوات	حجم الملايين (أ) ^b
٢٠١٣	٦٧
٢٠١٤	٦٨
٢٠١٥	٦٩
٢٠١٦	٦٩
٢٠١٧	٦٩
٢٠١٨	٦٩

مجموع المدارات (٣)	2	2	2	1	1	1
مجموع الإنتاج (٤)	3	4	2	2	2	2

ادرس و جود علاقه ارتباط خطیه بین حجم الإنتاج و حجم صادرات النفط الخام.

१८

x	y	xy	x^2	y^2
3	2	6	9	4
4	2	8	16	4
2	2	4	4	4
2	1	2	4	1
2	1	2	4	1
2	1	2	4	1
Σ	15	9	24	15

$$r_p = \frac{6(24) - (15)(9)}{\sqrt{((6 \times 41) - 15^2)((6 \times 15) - 9^2)}} = \frac{144 - 135}{\sqrt{(246 - 225)(90 - 81)}} = \frac{9}{\sqrt{189}} = \frac{9}{13.75} = 0.65$$

من الملاحظ أن علاقة الارتباط الخطي بين حجم الإنتاج وحجم صادرات النفط الخام علاقة طردية متوسطة.

- مثال (٤ - ١) :
لدراسة علاقة الصادرات بالميزان التجاري خلال عدة سنوات، أخذنا عشر قراءات تقريبية لقيمة صادرات العربية السعودية (X) وقيمة الميزان التجاري (y) بعشرينات المليارات ريال كما يلي:

X	9	11	17	18	19	16	16	19	23	23
y	1	3	8	7	6	5	7	8	12	12

- احسب معامل الارتباط الخطي، ما مدى قوّة العلاقة الخطية؟
الحل :

x	y	xy	x^2	y^2
9	1	9	81	1
11	3	33	121	9
17	8	136	289	64
18	7	126	324	49
19	6	114	361	36
16	5	80	256	25
16	7	112	256	49
19	8	152	361	64
23	12	276	529	144
23	12	276	529	144
Σ	171	69	1314	3107
$= \sum x = \sum y = \sum xy = \sum x^2 = \sum y^2$				585

من الملاحظ أن علاقة الارتباط الخطي بين قيمة صادرات المملكة العربية السعودية وقيمة الميزان التجاري موجودة وهي علاقة ارتباط طردية قوية.

التجاري

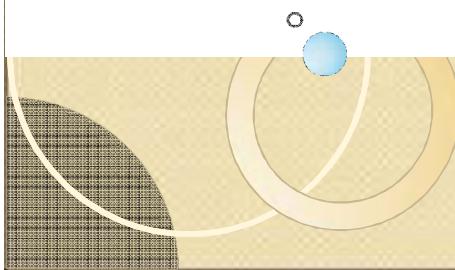
الأندثار

- Prediction
- التنبؤ هو تقدير القيمة المستقبلية لمتغير واحد بناءً على معرفة قيم متغير آخر وبياننا في:
 - تحديد شكل العلاقة بين المتغيرين رياضياً وبيانها (خط الانحدار).
 - توضيح اتجاه العلاقة بين المتغيرين.
 - التنبؤ بقيمة أحد المتغيرين بدلالة المتغير الآخر.



الأنحدار Regression

- والأنحدار هو أسلوب يمكن بواسطته تقدير قيمة أحد المتغيرين بمعلمة قيمة المتغير الآخر عن طريق معادلة الأنحدار، وله أنواع :
 - الانحدار الخطى البسيط : فكلمة "بسيط" تعنى أن المتغير التابع Y يعتمد على متغير مستقل واحد وهو X وكلمة "خطي" تعنى أن العلاقة بين المتغيرين (Y, X) علاقة خطية .
 - الانحدار المتعدد : إذا كان المتغير Y يعتمد على أكثر من متغير مستقل .
 - الانحدار غير الخطى : إذا كانت العلاقة بين المتغير Y والمتغيرات المسئولة غير خطية كأن تكون من الدرجة الثانية أو آسية.



الانحدار الخطى البسيط

- بعد تمثيل الأزواج المرتبة بالمستوى نحصل على شكل الانتشار فإذا أظهر الشكل الانتشاري للبيانات أن هناك علاقة خطية بين المتغيرين تقوم 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 انحدار Y على X (أو Y/X)

- وتحسب القيمتان a و b من العادتين التاليتين :

حيث:

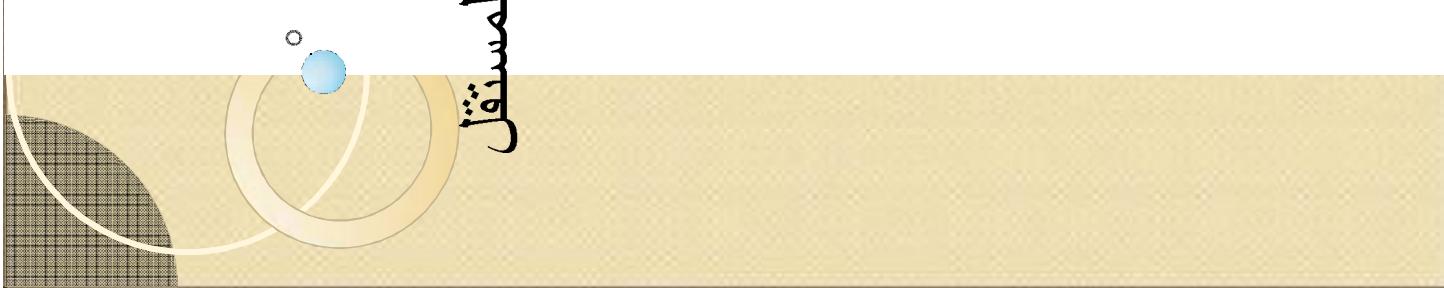
$$a = \frac{\sum x - b \sum X}{n}$$
$$b = \frac{n \sum xy - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2}$$

الأنحدار البسيط

- لإيجاد قيمة مقدرة جديدة \hat{y} نعرض بقيمة معلومة المتغير المستقل x ولتكن y في معادلة تقدير خط الأنحدار Y/X

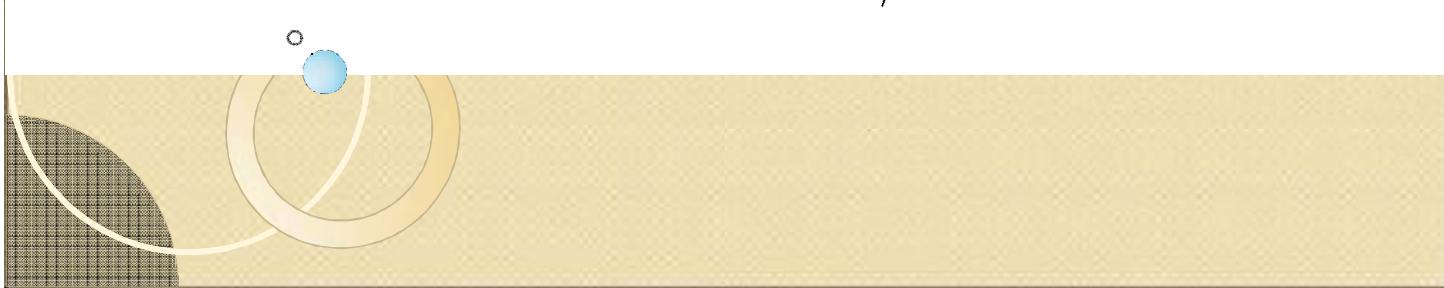
نعرض

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



الانحدار التحليلي البسيط

- ملاحظات هامة :
 - ❖ ميل الخط يمثل كمية التغير في Y المترتبة للتغير في X بمقدار وحدة واحدة .
 - ❖ إشارة معامل الانحدار تدل على نوع الارتباط (طري أو عكسي)
 - ❖ توجد علاقة بين معامل الانحدار ومعامل الارتباط الخططي



- مثال (8 - 4) : لدراسة علاقة الاستهلاك المحلي (y) بإنشاج (x) لمادة الإسفلت (بالمليون برميل) خلال عدة سنوات، أخذنا
- عشر قراءات تقريبية كما يلي :

y	6	8	9	8	7	6	5	6	5
x	10	13	15	14	9	7	6	6	5

أوج معادلة الانحدار الخطى البسيط، وتتوقع قيمة الاستهلاك عندما يصل إنشاج **٦,٠٠٠,٠٠٠** ابرميل .

- الحل :

$$b = \frac{n \sum xy - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2} = \frac{6320 - (90)(65)}{9420 - 90^2} = \frac{6320 - 5850}{1320} = 0.36$$

x	y	xy	x^2
10	6	60	100
13	8	104	169
15	9	135	225
14	8	112	196
9	7	63	81
7	6	42	49
6	5	30	36
5	5	25	25
5	5	25	25
Σ	90	632	942

$$a = \frac{\sum y - b \sum x}{n} = \frac{65 - (0.36 \times 90)}{10} = 3.26$$

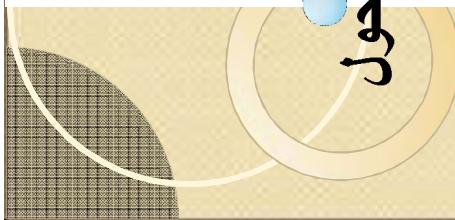
..
معادلة خط الانحدار البسيط في هذه الحالة : $\hat{y} = 3.26 + 0.36x$

تابع مدل مثال (8 - 4)

- ولتوقع قيمة الاستهلاك المحلي عندما يصل الإنتاج **٦٠٠٠٠٠٠٠** برميل، نحوال وحدة هذه القيمة من برميل إلى مليون برميل بالقسمة على مليون أي أن القيمة المستخدمة في توقع الاستهلاك هي
- وبالنطريض في المعادلة السابقة نجد أن:

$$\hat{y}_h = a + bx_h \\ = 3.26 + 0.36(16) = 9.02$$

أي أن الاستهلاك قد يصل إلى **٩.٠٢** مليون برميل، أي ما يعادل **٩٠٢٠٠٠٠** برميل خلال السنة.



٤ - ٣ - تطبيق الانحدار في مجال السلسل الزمنية

أسلوب السلسل الزمنية يتم رصد البيانات التي تعبر عن ظاهرة ما عند نقاط زمنية متتالية.

أمثلة:

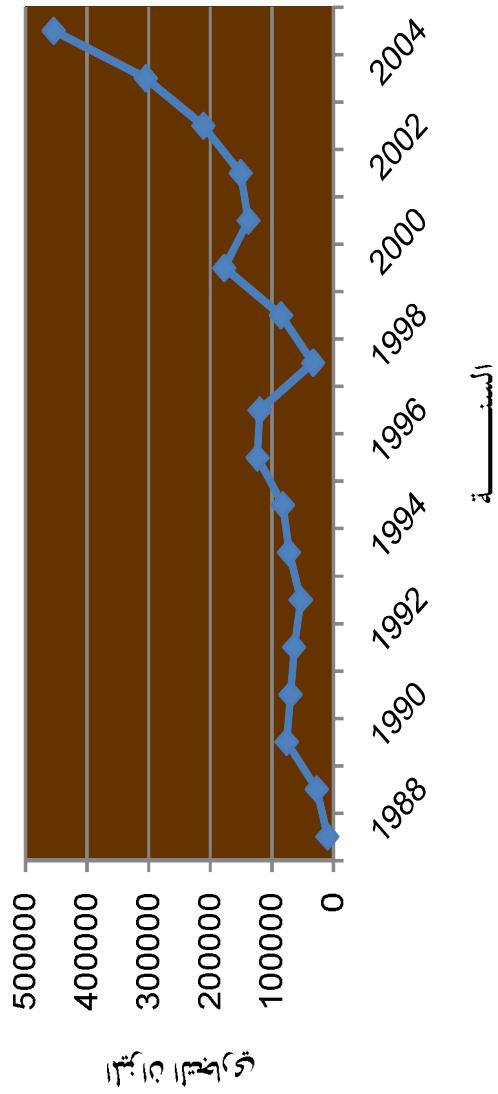
- كمية الصادرات السنوية
- حجم التعاملات الربع السنوي في البورصة
- عدد الجرائم اليومية في إحدى البلاد



3 - 4 - 4 تطبيق الاندثار في مجال السلسلة الزمنية

شكل السلسلة الزمنية

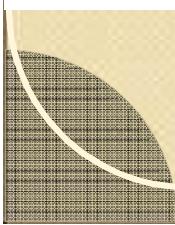
يعتبر شكل السلسلة الزمنية من أبرز الأشكال المستخدمة في كثيর من المجالات العلمية والاقتصادية والاجتماعية، حيث يتميز بعض الظواهر بالتطور خلال الزمن، مثل أسواق البورصة والأسهم، وتطور الأعداد المعتدلين والمحاج في موسمي رمضان والحج وتنمو السكاني وسعر النفط. تطور أعداد المعتدلين والمحاج في موسمي رمضان والحج وتنمو السكاني السنوي. فيما يلي مجموعة من أشكال السلسلة الزمنية لبعض الظواهر من واقع الحياة:



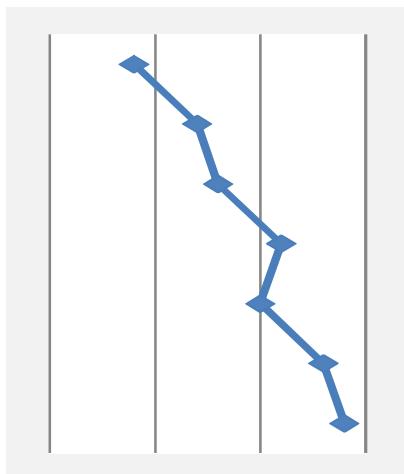
٤ - ٣ - تطبيق الانحدار في مجال السلسلة الزمنية

يستخدم شكل السلسلة الزمنية للتعرف على مكونات السلسلة الزمنية والتي تتمحور في أربع مكونات:

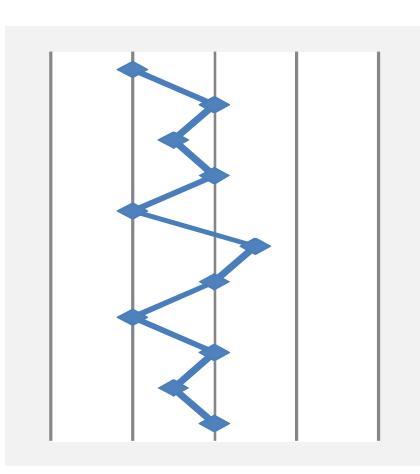
- الاتجاه العام: وهو اتجاه التطور الذي تأخذه السلسلة الزمنية خلال فترة طويلة من الزمن بالرغم من التذبذبات الموجودة بها، ويكون التطور إما بالزيادة أو بالنقصان، وبعض السلسل لا يوجد لها اتجاه.
- التغيرات الموسمية: وهي التغيرات التي تتكرر باانتظام خلال فترة زمنية أقل من السنة، عادة في المواسم.
- التغيرات الدورية: وهي التغيرات التي تحدث في فترات زمنية أكبر من سنة وعادة كل خمس أو عشر سنوات.
- التغيرات المرضية: وهي التغيرات التي تحدث نتيجة حوادث فجائية غير متوقعة مثل الفيضانات والأعاصير والحروب... الخ.



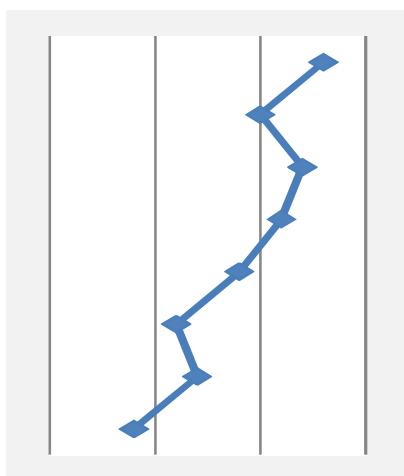
٣ - ٤ - ٤ تطبيق الانحدار في مجال المسلاسل الزمنية



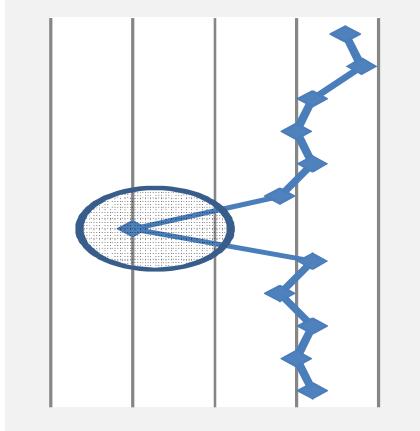
سلسلة ذات اتجاه عم بالزيداد



سلسلة ليس لها اتجاه عم



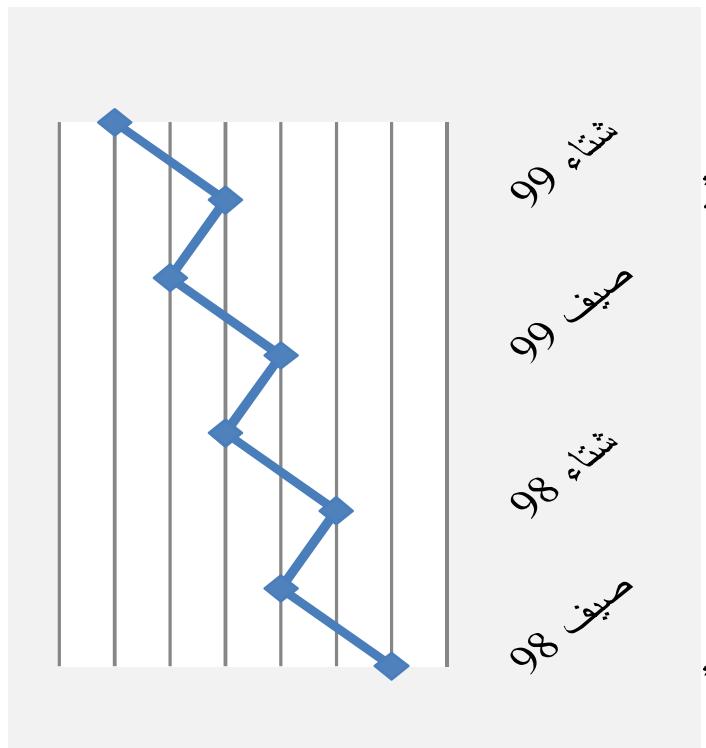
سلسلة ذات اتجاه عم بالتناقض



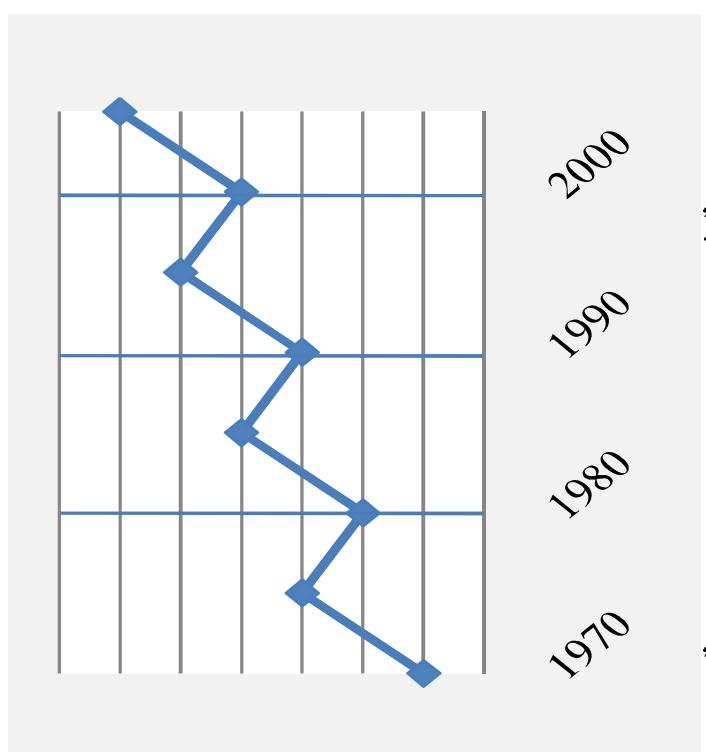
سلسلة ذات عامل عرضي



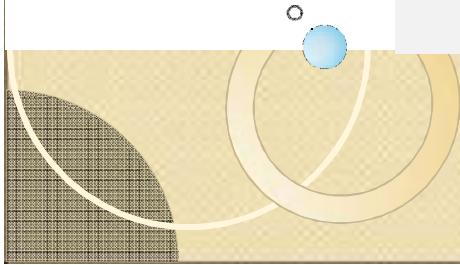
٣ - ٤ - ٤ تطبيق الانحدار في مجال السلسلة الزمنية



سلسلة ذات اتجاه زيادة وتغيرات موسمية



سلسلة ذات اتجاه زيادة تغيرات دورية



٣ - ٤ - ٤ تطبيق الانحدار في مجال السلسل الزمنية

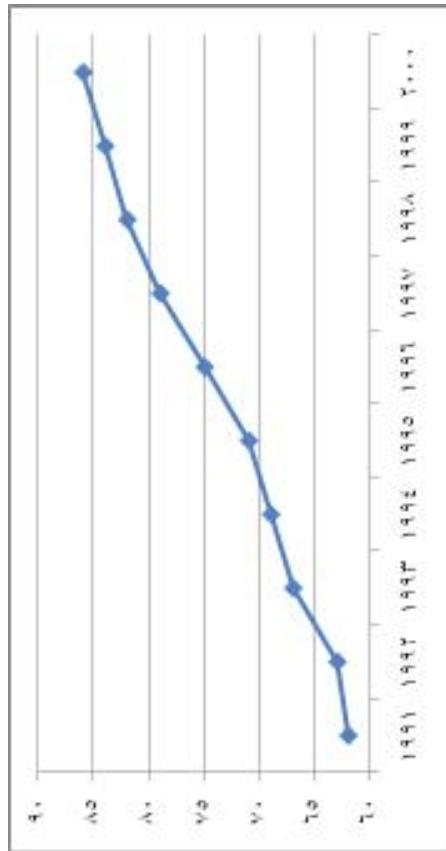
- يتطلب تحليل السلسل الزمنية عادة تحليل المكونات الأربع ولكن قد تختفي السلسل الزمنية من التغيرات الموسمية والدورية والعرضية، ويبقى تعين الاتجاه العام .
- يختلف شكل الاتجاه العام في السلسل الزمنية حسب طبيعة البيانات، وأحد أنواع الاتجاه العام هو الاتجاه الخطي .
- أحد طرق تعين الاتجاه العام الخطي هو استخدام أسلوب الانحدار الخطي البسيط، باعتبار أن الزمن (السنوات، الشهور،...الخ) متغير مستقل **X**، والمتغير التابع **Y** هو الظاهر محل الدراسة.
- ملاحظات:
 - تعين للمتغير المستقل القيم $0, 1, 2, \dots, n = X$ لتمثيل وحدة الزمن.
 - تدل إشارة معامل الانحدار **b** على نوع الاتجاه العام .

- مثال (٩ - ٤) :
البيانات التالية تمثل عدد المحقق المكتشفة (Y) خلال الأعوام ١٩٩١م إلى ٢٠٠٠م :

السنة	١٩٩١	١٩٩٢	١٩٩٣	١٩٩٤	١٩٩٥	١٩٩٦	١٩٩٧	١٩٩٨	١٩٩٩	٢٠٠٠
Y	٦٢	٦٣	٦٧	٦٩	٧٠	٧٥	٧٩	٨٢	٨٤	٨٦

حدد نوع الاتجاه العام، ثم قدر معادلة الاتجاه العام الخطى، ثم توقع عدد المحقق المكتشفة عام ٢٠٠٢م.

- الحل:



يدل الاتجاه العام على الزيادة في قيمة عدد المحقق المكتشفة.

- تابع حل المثال (٤-٩) : تقدير معادلة الاتجاه العام الخطى ، وحساب التوقع

٠

$$b = \frac{n \sum xy - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2} = \frac{35530 - (45)(737)}{2850 - 45^2} = \frac{2365}{825} = 2.87$$

$$a = \frac{\sum y - b \sum x}{n} = \frac{737 - (2.87 \times 45)}{10} = 60.79$$

السنة	x	y	xy	x^2
1991	0	62	0	0
1992	1	63	63	1
1993	2	67	134	4
1994	3	69	207	9
1995	4	70	280	16
1996	5	75	375	25
1997	6	79	474	36
1998	7	82	574	49
1999	8	84	672	64
2000	9	86	774	81
Σ	45	737	3553	285
	$= \sum x$	$= \sum y$	$= \sum xy$	$= \sum x^2$

∴ معادلة الاتجاه العام الخطى في هذه المثال

$$\hat{y} = 60.79 + 2.87x$$

تابع حل المثال (٤-٩) : تقدير معادلة الاتجاه العام الخطي ، ومسار
المتوقع

- ولتوقع عدد الحقول المتوقع اكتشافها عام 2002م نعرض بقيمة
تدل على هذا الزمن؛ حيث أن $2000 \rightarrow 9 = x$
 $x_h = 11 \rightarrow 2002$ إذن
- وبالتعويض في معادلة الاتجاه العام نجد أن:

$$\hat{y}_h = 60.79 + 2.87x_h \\ = 60.79 + 2.87(11) = 92.36 \approx 92$$



بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِيْمِ

