

قررت وزارة التعليم تدريس
هذا الكتاب وطبعه على نفقتها



المملكة العربية السعودية

الإحصاء

التعليم الثانوي - نظام المسارات
السنة الثالثة

قام بالتأليف والمراجعة
فريق من المتخصصين

يُوزع مجاناً للإيحاء

٣) وزارة التعليم، ١٤٤٤هـ

فهرسة مكتبة الملك فهد الوطنية أثناء النشر

وزارة التعليم

الإحصاء - التعليم الثانوي - نظام المسارات - السنة الثالثة . /

وزارة التعليم . - ط ١٤٤٥ . - الرياض ، ١٤٤٤ هـ

١٨٤ ص ٢١ : ٢٥,٥ سم

ردمك ٩ - ٤٥٠ - ٥١١ - ٦٠٣ - ٩٧٨

١ - الرياضيات - تعليم ٢ - الحقائق التعليمية ٣ - التعليم الثانوي -

السعودية أ. العنوان

١٤٤٤ / ٩٣٠٢

ديوي ٥١٠,٧١٢

رقم الإيداع: ١٤٤٤ / ٩٣٠٢

ردمك: ٩ - ٤٥٠ - ٥١١ - ٦٠٣ - ٩٧٨

حقوق الطبع والنشر محفوظة لوزارة التعليم

www.moe.gov.sa

مواد إثرائية وداعمة على "منصة عين الإثرائية"



ien.edu.sa

أعزاءنا المعلمين والمعلمات، والطلاب والطالبات، وأولياء الأمور، وكل مهتم بالتربية والتعليم؛
يسعدنا تواصلكم؛ لتطوير الكتاب المدرسي، ومقترحاتكم محل اهتمامنا.



fb.iien.edu.sa

أخي المعلم/أختي المعلمة، أخي المشرف التربوي/أختي المشرفة التربوية؛
نقدر لك مشاركتك التي ستسهم في تطوير الكتب المدرسية الجديدة، وسيكون لها الأثر الملموس في دعم
العملية التعليمية، وتجويد ما يقدم لأبنائنا وبناتنا الطلبة.



fb.iien.edu.sa/BE

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

الفصل الثالث

الارتباط والانحدار الخطي

Correlation and Linear Regression





المتغيرات العشوائية

Random Variables

الدرس الأول

• أتعرف المتغيرات العشوائية المنفصلة والمتصلة، وأميزها، وأوجد قيمها في تجارب عشوائية.

في هذا
الدرس



◀ ما الرقم الذي يمكن أن يظهر على الوجه العلوي لحجر نرد عند رميه؟
في تجربة رمي حجر النرد يظهر على الوجه العلوي رقم من مجموعة الأرقام {1, 2, 3, 4, 5, 6}. تعد هذه التجربة تجربة عشوائية، وتمثل تلك الأرقام قيم المتغير العشوائي.
◀ ما التجربة العشوائية؟

التجربة العشوائية (Random Experiments):

أي عملية يتم من خلالها الحصول على نتائج سواء كانت أرقاماً أم قياسات أم استجابات، بحيث تكون النتائج الممكنة لهذه العملية معلومة مسبقاً، ولا يمكن تحديد أيها سيتحقق فعلاً قبل إجراء التجربة.

المتغير العشوائي (Random Variable):

هو متغير في تجربة أو حدث معين، يأخذ مجموعة من القيم لكل منها احتمال محدد.
وقد يكون المتغير العشوائي كمياً أو نوعياً، مثلاً: عدد خريجي كلية العلوم متغير كمي، ومن سيفوز في سباق الجري متغير نوعي. ومن الممكن دائماً التعبير عن البيانات النوعية باستخدام البيانات الكمية؛ عن طريق تخصيص رقم لكل قيمة نوعية وفقاً لنظام محدد مسبقاً. على سبيل المثال: يتم تسجيل الرقم (1) إذا فاز أحمد بالسباق والرقم (0) إذا لم يفز أحمد بالسباق.

إذا كان عدد الخريجين يرمز له بالرمز X وكانت نتيجة أحمد في سباق الجري يرمز لها بـ Y ، فإنه في نهاية كل سنة يمكن الحصول على قيمة للمتغير X ، وفي ختام كل سباق يشارك فيه أحمد، يمكن الحصول على قيمة للمتغير Y .
من الطبيعي أن يقال عن متغير مثل X أو Y إنه متغير عشوائي؛ لأن القيم التي يتخذها مرتبطة بتجارب عشوائية.
تسمى مجموعة جميع النتائج الممكنة لتجربة عشوائية فضاء العينة ويرمز لها بالرمز S .

الحادثة (Event): مجموعة جزئية من فضاء العينة، وقد تساويه.

مثل: فضاء العينة في تجربة رمي حجر النرد مرة واحدة: $S = \{1, 2, 3, 4, 5, 6\}$

وتتضمن هذه التجربة عدداً من الحوادث مثل: $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$

كما أن ظهور عدد زوجي $\{2, 4, 6\}$ يمثل حادثة أيضاً، وحيث إن عدد النتائج الممكنة لهذه الحادثة هو ثلاث نتائج، فإن هذا يعني أن المتغير العشوائي $X = 3$

مثال 1

في تجربة رمي حجر النرد مرة واحدة، عيّن الحادثة A التي تمثل ظهور عدد أكبر من أو يساوي 2، وحدد عدد نتائجها.

الحل:

الحادثة A هي $\{2, 3, 4, 5, 6\}$ ، وعدد نتائجها $= 5$.

مثال 2

اكتب فضاء العينة لتجربة رمي قطعة نقود مرة واحدة.

الحل:

نفرض أن H يعني ظهور الصورة و T يعني ظهور الكتابة فيكون:

فضاء العينة لرمي قطعة نقود مرة واحدة: $S = \{H, T\}$

في تجربة رمي قطعة النقود مرة واحدة، تسمى المجموعات $\{H\}, \{T\}$ وكذلك $\{H, T\}$ حوادثاً.



1. في تجربة رمي حجر النرد مرة واحدة، عَيّن الحادثة B التي تمثل ظهور عدد أصغر من أو يساوي 4، وحدد عدد نتائجها. **الحادثة B هي: $B = \{1, 2, 3, 4\}$ ، عدد النتائج يساوي 4.**

2. في التجربة السابقة؛ هل يُعدّ ظهور الأعداد الأكبر من 7 حادثة؟ فسّر إجابتك.

ظهور الأعداد الأكبر من 7 حادثة مستحيلة الظهور، لأن عدد أوجه حجر النرد يساوي 6

تنقسم المتغيرات العشوائية إلى نوعين:

المتغير العشوائي المنفصل (Discrete Random Variable):

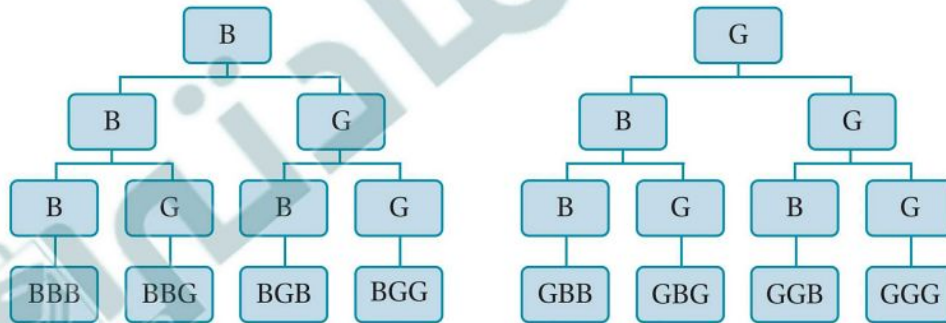
هو المتغير الذي يأخذ قيمًا تنتمي إلى مجموعة منتهية، أو غير منتهية، ولكنها قابلة للعد؛ أي أنه يأخذ قيمًا منفصلة عن بعضها.

مثلاً: أسرة لديها ثلاثة أبناء، وليكن X المتغير العشوائي الذي يدل على عدد الذكور في هذه الأسرة.

وحتى نتمكن من إيجاد قيم المتغير العشوائي X ، نوجد أولاً فضاء العينة لحادثة وجود ثلاثة أبناء في الأسرة.

نفرض أن G يرمز للذكور، و B يرمز للإناث؛ فتكون مجموعة النتائج الممكنة لفضاء العينة ممثلة في الرسم

الشجري الآتي:



$$S = \{GGG, GGB, GBG, GBB, BGG, BGB, BBG, BBB\}$$

وبالعودة للمتغير العشوائي الذي يمثل عدد الذكور في الأسرة، فإن:

$$X(GGG)=3, \quad X(GGB)=2, \quad X(GBG)=2, \quad X(GBB)=1, \quad X(BGG)=2, \quad X(BGB)=1, \quad X(BBG)=1, \\ X(BBB)=0$$

من الواضح أن القيم التي يمكن أن يأخذها المتغير X هي: 0, 1, 2, 3

يلاحظ أن هذا المتغير X يأخذ قيمًا منتهية وقابلة للعد. فهو متغير عشوائي منفصل.

المتغير العشوائي المتصل (Continuous Random Variable):

يأخذ قيماً تنتمي إلى مجموعة غير منتهية وغير قابلة للعد؛ بمعنى أن المتغير يأخذ جميع القيم التي تقع في نطاق تغيره.

مثلاً: تجربة اختيار نقطة من دائرة مركزها نقطة الأصل $(0, 0)$ ونصف قطرها r .

وليكن X المتغير الذي يدل على المسافة بين النقطة المختارة ومركز الدائرة.

فتكون مجموعة القيم الممكنة التي يمكن أن يأخذها المتغير X هي:

$\{X: 0 \leq X \leq r\}$ وهي مجموعة غير منتهية وغير قابلة للعد.

أوجد نوع المتغير في تجربة اختيار نقطة على دائرة مركزها نقطة الأصل $(0, 0)$. ونصف

قطرها r ، حيث يشير X إلى المسافة بين النقطة المختارة ومركز الدائرة.

ليس متغير؛ ذلك أن المسافة بين النقطة المختارة والمركز تساوي قيمة ثابتة



تفكير ناقد

3

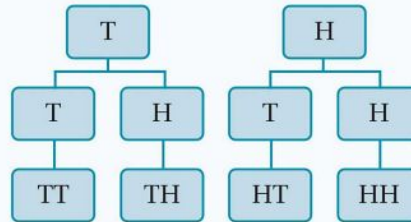
مثال



في تجربة رمي قطعة نقود مرتين على التوالي، لنفترض أن المتغير العشوائي X هو عدد مرات ظهور الصورة H في الرميّتين، ما قيم المتغير العشوائي X ؟ حدد ما إذا كان X منفصلاً أو متصلاً.

الحل:

في تجربة رمي قطعة النقود مرتين، تكون مجموعة النتائج الممكنة لفضاء العينة ممثلة في الرسم الشجري الآتي:



يكون فضاء العينة كالآتي:

$$S = \{HH, HT, TH, TT\}$$

المتغير العشوائي X هو عدد مرات ظهور الصورة في الرميتين، وبالتالي:

النتاج HH يعطي قيمة للمتغير العشوائي $X = 2$.

النتاج HT يعطي قيمة للمتغير العشوائي $X = 1$.

النتاج TH يعطي قيمة للمتغير العشوائي $X = 1$.

النتاج TT يعطي قيمة للمتغير العشوائي $X = 0$.

من الواضح أن القيم التي يمكن أن يأخذها هذا المتغير هي 0, 1, 2.

يلاحظ أن المتغير يأخذ قيمًا منتهية وقابلة للعد. فهو متغير عشوائي منفصل.

4

مثال

إذا كان المتغير العشوائي X يمثل أوزان الطلاب في مدرستك، فهل X متغير منفصل أم متصل؟

الحل:

يمكن أن تتراوح أوزان الطلاب في المدرسة بين 40 كيلوجرام إلى 90 كيلوجرام.

من الواضح أن هناك أعدادًا غير منتهية وغير قابلة للعد من القيم الممكنة التي يمكن أن يأخذها هذا المتغير العشوائي، إذن X متغير عشوائي متصل.

2

تحقق من فهمك

1. هل الجملة صحيحة أم غير صحيحة فيما يأتي؟

a. المتغير الذي يمثل أسعار الأسهم في البورصات العالمية، يعد متغيرًا عشوائيًا منفصلًا. **خطأ**

b. المتغير الذي يمثل عدد المصابين بمرض السكري في المملكة العربية السعودية، يعد متغيرًا عشوائيًا

متصلًا. **خطأ**

2. أعط أمثلة لمتغيرات عشوائية منفصلة، وأخرى متصلة، ثم قدم مقترحات حول استعمالات كل منها في

مسارك التخصصي **بالنسبة لمسار الصحة والحياة**

نسبة الإصابات بالحمى القلاعية بين الأغنام (متغير عشوائي متصل)

عدد العظام في جسم الإنسان (متغير عشوائي منفصل)

مستوى السكر في الدم (متغير عشوائي متصل)

بالنسبة لمسار إدارة الأعمال:

عدد الموظفين في قسم الموارد البشرية (متغير عشوائي منفصل)



الارتباط Correlation

الدرس الثاني



في هذا
الدرس

- أتعرّف الارتباط في البيانات الكمية والنوعية، وأميز المتغيرات المستقلة والتابعة.
- أحسب معامل الارتباط بين متغيرين، وأستخدمه في تحديد نوع الارتباط وقوته، وأفسره لاتخاذ القرارات المناسبة.



كيف يمكن وصف العلاقة بين متغيرين؟

يريد باحث معرفة ما إذا كان هناك علاقة بين كمية الكهرباء المولدة (Y) عن مروحة هوائية وسرعة الرياح (X)، بمعنى هل التغير في سرعة الرياح يؤثر في كمية توليد الكهرباء؟

لتوليد الكهرباء باستخدام مروحة هوائية؛ إذا زادت سرعة الرياح، سيزداد ناتج التوليد وفقًا لذلك. وبالتالي، فإن السرعة المتغيرة X وكمية الكهرباء Y بينهما علاقة.

ما مفهوم الارتباط بين متغيرين؟

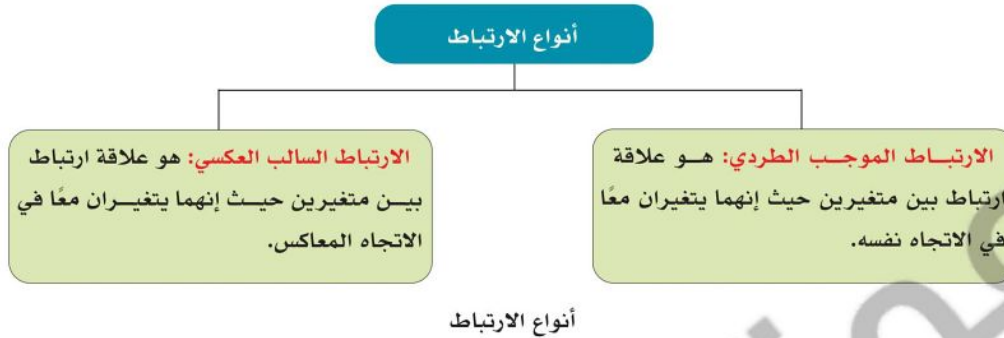
الارتباط (Correlation)

وصف علاقة بين متغيرين من حيث قوتها، كل من المتغيرين يأخذ قيمًا ويمثلان مجموعة من البيانات عند عرضهما معًا بواسطة زوج من المتغيرات العشوائية (X,Y)، حيث X هو المتغير المستقل (independent variable) وY متغير تابع لـ X (dependent variable)، يقوم الباحث بتحديد قيم المتغير المستقل X الذي تعتمد قيم المتغير التابع Y عليه.

فإذا كان تغير أحد المتغيرين يعتمد كليًا على تغير الآخر فنقول إن الارتباط بينهما تام، مثل: العلاقة بين مساحة الدائرة ونصف قطرها. أما إذا كان تغير أحدهما لا يعتمد على تغير الآخر فنقول إن الارتباط بينهما غير تام، مثل: العلاقة بين وزن الفرد وطوله، أو بين التحصيل وعدد ساعات الدراسة، أو بين الدخل والمصروفات.

أنواع الارتباط:

للاارتباط نوعان: ارتباط موجب (طردي)، وارتباط سالب (عكسي).



عند النظر في الارتباط بين متغيرين يؤخذ بعين الاعتبار نوع البيانات التي يمثلانها؛ فالارتباط في البيانات الكمية يختلف عنه في البيانات النوعية.

الارتباط في البيانات الكمية

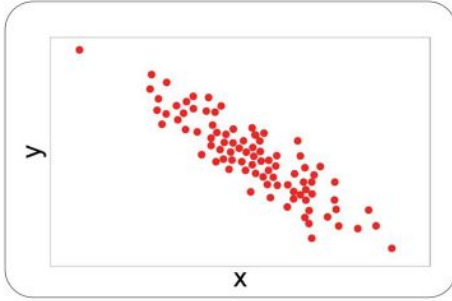
يمكن التعبير عن الارتباط في البيانات الكمية ووصفه بيانياً من خلال رسم شكل الانتشار، أو جبرياً بحساب معامل الارتباط.

شكل الانتشار (Scatter Plot):

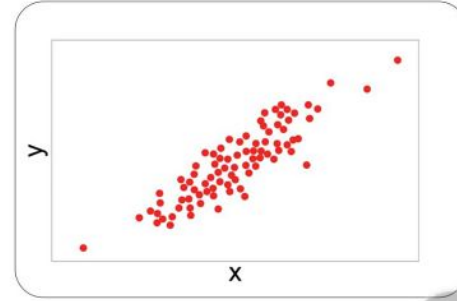
وسيلة أولية يعرف الباحث من خلالها نوع الارتباطات بين المتغيرات الكمية (طردي، عكسي)، ومستوى قوتها (قوي، ضعيف، لا يوجد ارتباط).

يحدد شكل الانتشار طبيعة العلاقة بين المتغيرين X و Y ، حيث يمثل أحد المتغيرين X على المحور الأفقي، والمتغير الآخر Y على المحور الرأسي، وتكون لكل قيمة للمتغير X قيمة مناظرة للمتغير Y في الرسم البياني، ويتم تحديد الأزواج المرتبة (X, Y) بنقاط تمثل شكل الانتشار.

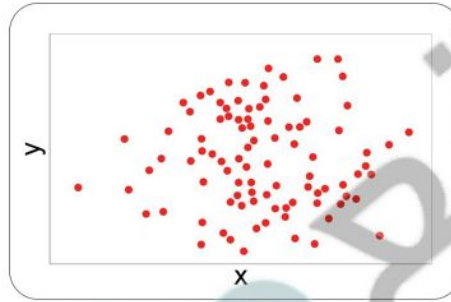
أي أن رسم شكل الانتشار وتحليله يساعد في الإجابة عن الأسئلة الآتية: هل هناك علاقة بين المتغيرين؟ وإن وجدت فما مدى قوتها واتجاهها؟



الارتباط العكسي (السالب)



الارتباط الطردي (الموجب)



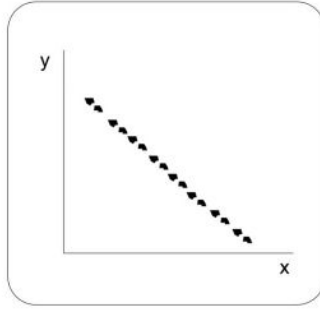
لا يوجد ارتباط بين المتغيرين

الارتباط وفق شكل الانتشار

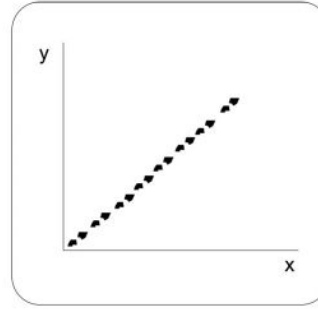
حالات شكل الانتشار:

1. إذا وقعت جميع النقاط على خط مستقيم، فإن الارتباط يُعبر عنه بعلاقة خطية، تمثل أقوى نوع من الارتباط بين المتغيرين.

وفي هذه الحالة؛ إذا كانت العلاقة طردية (كلما زادت قيمة المتغير X ، زادت قيمة المتغير Y)، يكون الارتباط طرديًا تامًا. وإذا كانت العلاقة عكسية (كلما زادت قيمة المتغير X ، انخفضت قيمة المتغير Y) يكون الارتباط عكسيًا تامًا، كما يوضح الشكل الآتي:



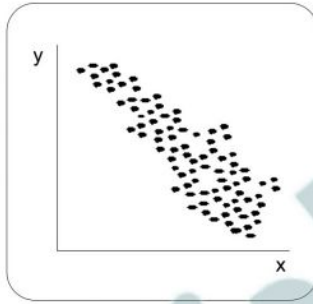
ارتباط عكسي تام



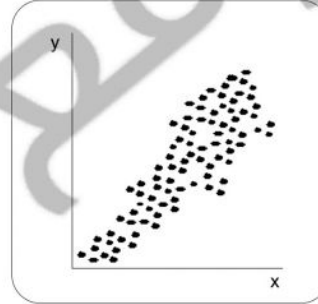
ارتباط طردي تام

أنواع الارتباط الخطي التام

2. إذا كانت النقاط تتخذ شكل خط مستقيم، ولكن لا تقع جميعها على هذا الخط، فإن العلاقة خطية غير تامة (طرديّة أو عكسيّة). كما يوضح الشكل الآتي:



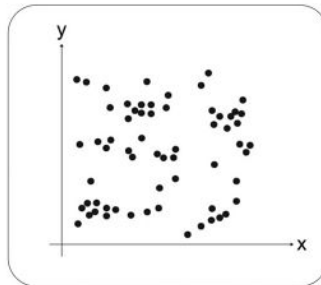
ارتباط خطي عكسي



ارتباط خطي طردي

أنواع الارتباط الخطي غير التام

3. إذا كانت النقاط مشتتة بدون نظام معين فهذا يدل على عدم وجود علاقة بين المتغيرين، أو أن العلاقة بينهما ضعيفة جداً، كما يوضح الشكل الآتي:



لا يوجد ارتباط

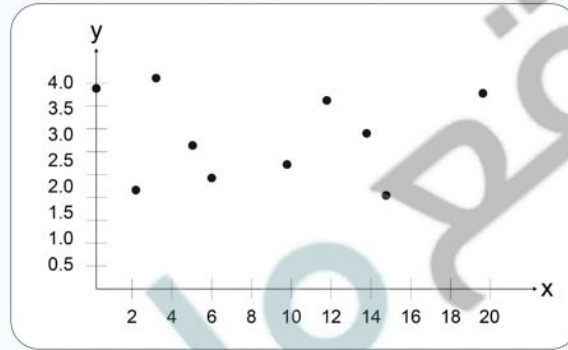


تعرض البيانات الآتية عدد ساعات العمل الإضافية التي يقضيها أحد الموظفين أسبوعياً، ونسبة إنجازه.

ارسم شكل الانتشار، وهل يوجد هناك ارتباط بين المتغيرين؟ حدد نوع الارتباط.

5	15	14	20	2	10	6	0	3	13	عدد الساعات (X)
3.1	1.8	3	3.7	2.2	2.4	2.5	3.9	4	3.6	نسبة الإنجاز (Y)

الحل:

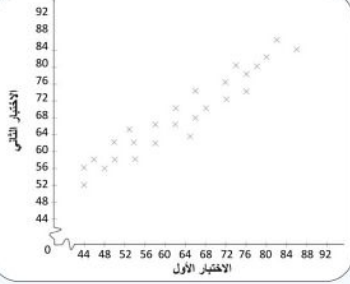


في رسم شكل الانتشار يظهر عدم وجود ارتباط خطي بين المتغيرين، مما يعني أن عدد الساعات الإضافية التي يقضيها الموظف في الأسبوع قد لا تكون ذات علاقة بمعدل إنجاز العمل.



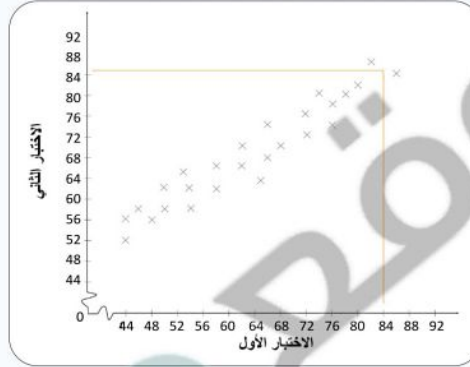
كيف يستفيد صاحب العمل من هذا النوع من الدراسات؟
تفكير ناقد يمكن لصاحب العمل أن يوقف أو يقتن ساعات العمل الإضافي عندما يلاحظ عدم وجود ارتباط بين عدد ساعات العمل ونسبة الإنجاز

مثال 2



يعرض شكل الانتشار نتائج اختبارين لفصل من 26 طالبًا، إذا حصل طالب على 84 في الاختبار الأول، فكيف يمكن أن تكون نتيجته في الاختبار الثاني؟

الحل:



من شكل الانتشار يتضح أنه عندما تكون درجة الطالب في الاختبار الأول 84، يمكن أن تكون 85 تقريبًا.

تحقق من فهمك 1

1. تعرض البيانات الآتية أطوال عدد من نباتات الظل ومعدلات الإصابة بالتبقرش

الأطوال بالسنتيمتر (X)	68	72	65	70	62	75
معدل التبقرش (Y)	90	85	88	85	93	98

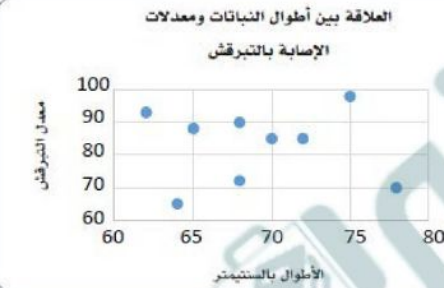
2.

ارسم شكل الانتشار لهذه البيانات، وهل هناك ارتباط بين المتغيرين؟ حدد نوع

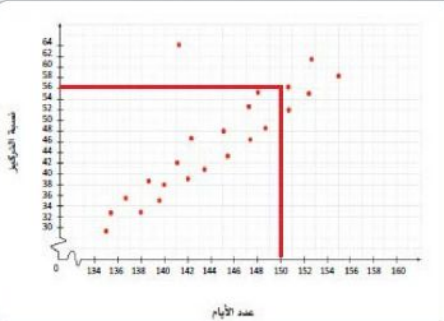
2. يوضح شكل الانتشار الآتي العلاقة بين عدد الأيام التي خضع فيها عدد من

فرط الحركة لتلقي العلاج ونسبة التركيز لديهم أثناء الحصة الدراسية، استخدم

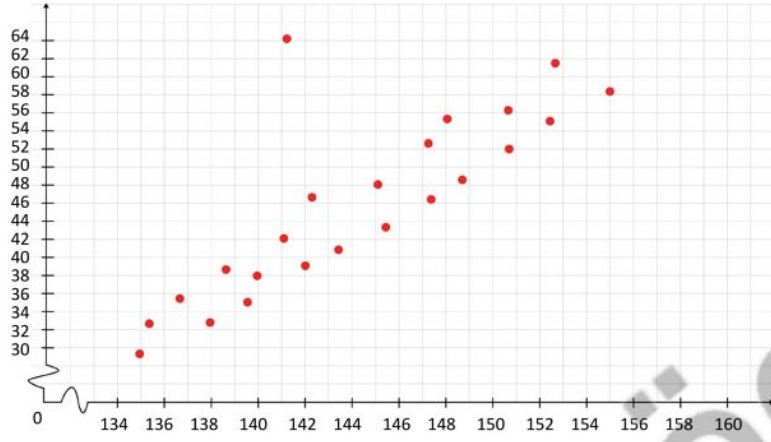
نسبة تركيز يمكن أن يحققها طفل تلقى العلاج على مدى 150 يومًا.



رسم شكل الانتشار يوضح أنه لا يوجد ارتباط بين متوسط أطوال نباتات الظل ومعدلات الإصابة بالتبقرش.



يوضح الرسم أن أفضل نسبة تركيز يمكن أن يحققها طفل تلقى العلاج على مدى 150 يومًا هي 56%.



اضطراب نقص الانتباه

اضطراب نقص الانتباه وفرط الحركة هو مرض يصيب الأطفال قبل سن 12 ، ويتسبب في خلق مشكلات لدى الأطفال في المنزل والمدرسة بشكل مستمر، مثل: مشاكل التعلم أو اللغة، الأمراض المزاجية: مثل الاكتئاب أو القلق، واضطرابات النوم، ويخضع المصابون بالمرض لتقييم دقيق للحالة، مع تقديم أدوية متخصصة ومعالجات سلوكية متنوعة.

قوة الارتباط:

يستخدم معامل الارتباط لحساب قوة الارتباط بين متغيرين.

معامل الارتباط (Correlation Coefficient):

قيمة تحدد الارتباط بين متغيرين، ويرمز له بالرمز r .

تتراوح قيمة معامل الارتباط من -1 إلى +1، أي أن قيمته

لا يمكن أن تتجاوز هذا النطاق.

تعني قيمة معامل الارتباط -1 أن هناك ارتباطاً عكسياً تاماً، بينما تعني قيمة معامل الارتباط +1 أن هناك ارتباطاً

طردياً تاماً، وعندما تكون قيمة معامل الارتباط 0 فتعني أنه لا توجد علاقة بين المتغيرين.



معامل الارتباط r يسمى أيضاً
معامل بيرسون Pearson. نسبة
للإحصائي الإنجليزي كارل
بيرسون (1857-1936).

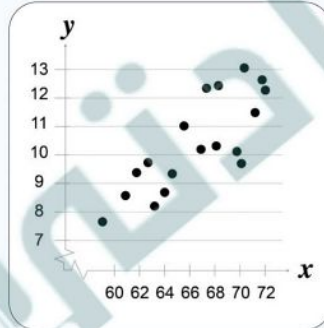
ويوضح الجدول الآتي قوة الارتباط حسب قيمة معامل:

قوة الارتباط	قيمة معامل الارتباط
ارتباط طردي تام	+1
ارتباط طردي قوي	من 0.70 إلى 0.99
ارتباط طردي متوسط	من 0.50 إلى 0.69
ارتباط طردي ضعيف	من 0.01 إلى 0.49
لا يوجد ارتباط	0
ارتباط عكسي ضعيف	من -0.01 إلى -0.49
ارتباط عكسي متوسط	من -0.50 إلى -0.69
ارتباط عكسي قوي	من -0.70 إلى -0.99
ارتباط عكسي تام	-1

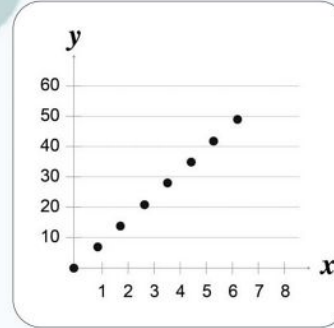
3

مثال

حدد نوع الارتباط ودرجته في الأشكال الآتية:



معامل الارتباط $r = 0.81$



معامل الارتباط $r = 1$

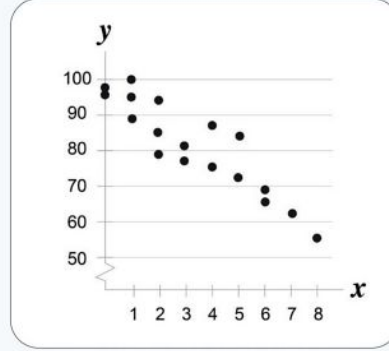
الحل:

a. ارتباط طردي تام.

b. ارتباط طردي قوي.

تحقق من فهمك 2

حدد نوع الارتباط ودرجته في الشكل الآتي:



ارتباط عكسي قوي

معامل الارتباط $r = -0.8$

تذكر:

الرمز الرياضي $(\sum X)^2$ يعني مربع مجموع القيم، ولا يعني مجموع مربعات القيم.

$$(\sum X)^2 \neq \sum X^2$$

يمكن حساب قيمة معامل الارتباط لبيانات المتغيرين X, Y باستخدام الصيغة الآتية:

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2} \sqrt{n\sum Y^2 - (\sum Y)^2}}$$

حيث n عدد أزواج البيانات، ويكون نطاق معامل الارتباط في الفترة $[-1, 1]$.

مثال 4

توضح البيانات الآتية درجات خمسة طلاب في مادتي الرياضيات والفيزياء، احسب معامل الارتباط بين درجات الطلاب في المادتين. ماذا تستنتج؟

الرياضيات (X)	الفيزياء (Y)
4	4
8	7
5	5
3	4

الحل:

الرياضيات (X)	الفيزياء (Y)	XY	X ²	Y ²
4	4	16	16	16
8	7	56	64	49
5	5	25	25	25
3	4	12	9	16
$\sum x = 20$	$\sum Y = 20$	$\sum XY = 109$	$\sum x^2 = 114$	$\sum Y^2 = 106$

بالتعويض عن هذه المجاميع، وعن $n=4$ ، في قانون معامل الارتباط:

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2} \sqrt{n\sum Y^2 - (\sum Y)^2}}$$

$$= \frac{4(109) - (20)(20)}{\sqrt{4(114) - (20)^2} \sqrt{4(106) - (20)^2}}$$

$$= \frac{36}{\sqrt{56}\sqrt{24}} = 0.98$$

إذن يوجد ارتباط طردي قوي بين درجات الطلاب في مادتي الرياضيات والفيزياء.

Y ²	X ²	XY	Y	X
4489	4225	4355	67	65
4624	4624	4624	68	68
2601	2500	2550	51	50
1225	1521	1365	35	39
1521	2304	1872	39	48
2704	3364	3016	52	58
$\sum Y^2 = 17164$	$\sum X^2 = 18538$	$\sum XY = 17782$	$\sum Y = 312$	$\sum X = 328$

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2} \sqrt{n\sum Y^2 - (\sum Y)^2}}$$

$$= \frac{6(17782) - (328)(312)}{\sqrt{6(18538) - (328)^2} \sqrt{6(17164) - (312)^2}}$$

$$= \frac{4356}{\sqrt{3644}\sqrt{5640}} = 0.96$$

إذن يوجد ارتباط طردي قوي بين دخل الأسرة والإنفاق الشهري.

تحقق من فهمك

توضيح البيانات الآتية:

X
Y

احسب معامل الارتباط

بالتعويض عن هذه المجاميع وعن $n=6$ ، في قانون معامل الارتباط

الارتباط في البيانات النوعية

بفرض أن الظاهرة قيد الدراسة تحتوي على متغيرين نوعيين ترتيبيين، على سبيل المثال: قياس العلاقة بين تقديرات الطلاب في مادتين، أو قياس العلاقة بين درجة تفضيل المستهلك لسلعة معينة ومستوى دخله.

في هذا النوع من البيانات يستخدم معامل سبيرمان (Spearman) لقياس مستوى الارتباط بين المتغيرين X و Y.

معامل ارتباط سبيرمان (Spearman Correlation Coefficient):

مقياس لمستوى الارتباط بين متغيرين نوعيين ترتيبيين، بناءً على رتب البيانات المرصودة.

يعتمد حساب معامل سبيرمان على استبدال البيانات بترتيبها داخل العينة، بفرض أن العينة مكونة من n زوج مرتب وفق متغيرين X و Y. وأن المتغير X له رتبة R_X ، والمتغير Y له رتبة R_Y ، و d تعني الفرق بين الرتبتين، أي أن: $d = R_X - R_Y$ ، فإن معامل سبيرمان لارتباط الرتبة ويرمز له بالرمز R يعطى بالصيغة الآتية:

$$R = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

ويأخذ معامل سبيرمان قيمة من -1 إلى +1، تعبر عن قوة ارتباط المتغيرين.

5

مثال

لدراسة العلاقة بين تقديرات الطلاب في مادة الإحصاء وتقديراتهم في مادة الرياضيات، اختير ثمانية طلاب وكانت تقديراتهم على النحو الآتي:

الإحصاء (X)	ضعيف	ممتاز	جيد	مقبول	جيد جداً
الرياضيات (Y)	مقبول	جيد	جيد جداً	ضعيف	ممتاز

هل هناك ارتباط بين المتغيرين X, Y؟ ما نوعه؟ وما مدى قوته؟

ثانياً: حساب قيمة معامل سبيرمان كالتالي:

d^2	d	R_y	R_x	Y	X
1	-1	2	1	مقبول	ضعيف
1	-1	4	3	جيد جداً	جيد
1	-1	3	2	جيد	مقبول
1	-1	5	4	ممتاز	جيد جداً
1	1	1	2	ضعيف	مقبول
1	-1	3	2	جيد	مقبول
1	-1	4	3	جيد جداً	جيد
0	0	3	3	جيد	جيد
1	-1	5	4	ممتاز	جيد جداً
4	2	1	3	ضعيف	جيد
$\sum d^2 = 12$	$\sum d = -4$				

$$R = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6(12)}{10(100 - 1)} = 0.928$$

أي أن هناك علاقة ارتباط طردية قوية بين تقديرات

الطلاب في اللغة الانجليزية وتقديراتهم في اللغة العربية.

4 تحقق من فهمك

فيما يأتي تقديرات عشرة طلاب في اختبارين، أحدهما للغة الإنجليزية والآخر للغة العربية:

اللغة الإنجليزية	ضعيف	جيد	مقبول	جيد جداً	مقبول	جيد	جيد جداً	مقبول	جيد	جيد جداً	جيد
اللغة العربية	مقبول	جيد جداً	جيد	ممتاز	ضعيف	جيد	جيد جداً	جيد	ممتاز	ضعيف	ضعيف

هل هناك ارتباط بين المتغيرين؟ ما نوعه؟ وما مدى قوته؟

أولاً: حساب رتب التقديرات في المتغيرين، نفرض أن التقدير (ضعيف) تناظره رتبة منخفضة تساوي 1، يليها تقدير مقبول وتناظره الرتبة 2، ثم تقدير جيد وتناظره الرتبة 3 وهكذا

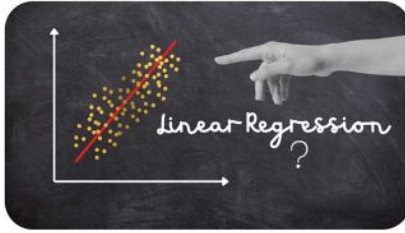


الانحدار الخطي البسيط

Simple Linear Regression

الدرس الثالث

- أكتب معادلة الانحدار الخطي البسيط، وأستخدمها في تمييز العلاقات الخطية بين المتغيرات والتنبؤ بقيمها.



« **فكر** كيف يمكن الاستفادة من شكل الانتشار في مجالات الحياة مختلفة؟

يوظف العلماء شكل الانتشار في التنبؤ في العديد من المجالات، وذلك باستخدام نموذج رياضي يُعرف بـ "خط الانحدار"، فعلى سبيل المثال: يستخدم علماء الأحياء والعلوم السلوكية خط الانحدار للتنبؤ بسلوك الكائنات والبشر، ويستخدمه علماء التقنيات الرقمية وعلم البيانات والذكاء الاصطناعي في حل مشكلات حاسوبية معقدة.

وهو عبارة عن ارتباط خطي بين متغيرين، حيث يمثل المتغير المستقل X متغير التنبؤ، فيما يمثل المتغير Y الاستجابة أو التوقع الناتج عن هذا التنبؤ، ومن المهم الإشارة إلى أن خط الانحدار يمثل معادلة خطية من الدرجة الأولى.

« **ما** الغرض من استخدام الانحدار الخطي البسيط؟

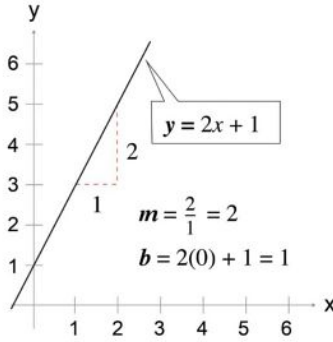
يستخدم الانحدار الخطي البسيط لدراسة تأثير متغير كمي (المتغير المستقل) على متغير كمي آخر (المتغير التابع)، على سبيل المثال:

- دراسة تأثير الإنتاج على التكلفة.
- دراسة تأثير كمية السماد على إنتاجية المزرعة.
- دراسة تأثير كمية البروتين التي يتناولها الشخص على وزنه.

الانحدار الخطي البسيط (Simple Linear Regression)

هو معادلة خطية من الدرجة الأولى، تعبر عن المتغير التابع كدالة في المتغير المستقل.

معادلة الخط المستقيم (Straight-line Equation):



يمكن كتابة معادلة الخط المستقيم بمعلومية ميله m والمقطع b على المحور Y ، بالصيغة الآتية:

$$Y = mx + b$$

حيث يمثل ميل الخط المستقيم m نسبة تغيره الرأسى إلى الأفقى، ويقطع المستقيم المحور Y عند القيمة b . ووفقاً للتمثيل المجاور فإن:

$$m = \frac{2}{1} = 2, \quad b = 1$$

وبالتالى فإن معادلة هذا المستقيم هي $Y = 2x + 1$

ويمكن إيجاد عدد لانتهائي من النقاط على هذا المستقيم، فعلى سبيل المثال تكون قيمة $Y = 5$ عندما $x = 2$.

معادلة خط الانحدار البسيط (Simple Regression Line Equation):

معادلة خط الانحدار البسيط للمتغير المستقل X والمتغير

التابع Y هي:

$$\hat{y} = mx + b$$

حيث $\hat{y}$ هي القيمة المتوقعة لـ y عند قيمة معينة لـ x .

ويحسب ميل خط الانحدار البسيط m من المعادلة التالية:

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

حيث n عدد البيانات.

ويعتمد نوع الارتباط بين المتغيرين على قيمة ميل خط الانحدار m ، حيث تكون المتغيرات مرتبطة ارتباطاً طردياً

إذا كانت قيمة الميل موجبة، ومرتبطة عكسياً إذا كانت قيمة الميل سالبة.

وتعطى نقطة التقاطع b مع المحور Y من المعادلة:

$$b = \bar{y} - m\bar{x}$$

حيث $\bar{y}$ هو المتوسط الحسابي لقيم المتغير y ، و $\bar{x}$ هو المتوسط الحسابي لقيم المتغير x . ويمر خط الانحدار

البسيط دائماً بالنقطة $(\bar{x}, \bar{y})$.

لاحظ:

في الرياضيات تتحدد معادلة الخط المستقيم بنقطتين عليه. وفي الإحصاء تستخدم كل نقطة في البيانات لتحديد معادلة خط الانحدار البسيط.



في الجدول أدناه بيانات عن كمية البروتين اليومية بالجرام التي يحتاجها العجل الرضيع، ومقدار الزيادة في وزن العجل بالكيلوجرام، لعينة من n عجل رضيع.

10	11	14	15	20	25	46	50	59	70	كمية البروتين x
10	10	12	12	13	13	19	15	16	20	زيادة الوزن y

اكتب معادلة خط الانحدار البسيط.

الحل:

من الجدول: عدد العجول الرضيعة: $n = 10$

x^2	xy	زيادة الوزن y	كمية البروتين x
4900	1400	20	70
3481	944	16	59
2500	750	15	50
2116	874	19	46
625	325	13	25
400	260	13	20
225	180	12	15
196	168	12	14
121	110	10	11
100	100	10	10
$\sum x^2 = 14664$	$\sum xy = 5111$	$\sum y = 140$	$\sum x = 320$

$$\bar{x} = \frac{\sum x}{n} = \frac{320}{10} = 32$$

$$\bar{y} = \frac{\sum y}{n} = \frac{140}{10} = 14$$

بالتعويض في معادلة ميل خط الانحدار البسيط:

$$\begin{aligned}
 m &= \frac{n \sum xy - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2} \\
 &= \frac{10(5111) - 320(140)}{10(14664) - (320)^2} \\
 &= 0.14
 \end{aligned}$$

بالتعويض في معادلة التقاطع مع محور Y:

$$\begin{aligned} b &= \bar{y} - m\bar{x} \\ &= 14 - 0,14(32) \\ &= 9,4 \end{aligned}$$

بالتعويض عن قيمتي m و b في معادلة خط الانحدار البسيط تكون معادلة خط الانحدار البسيط لكمية البروتين (x) على زيادة الوزن (y):

$$\hat{y} = 0,14x + 9,4$$

بما أن قيمة الميل $m = 0,14$ موجبة فإن الارتباط بين المتغيرين ارتباط طردي، أي أن وزن العجل الرضيع يرتبط طرديًا مع كمية البروتين.

1 تحقق من فهمك

من الجدول: $n = 5$

x^2	xy	التكلفة بالريالات y	كمية الزجاج المسطح بالطن X
64	144	18	8
16	44	11	4
225	435	29	15
49	126	18	7
36	84	14	6
$\sum x^2 = 390$	$\sum xy = 833$	$\sum y = 90$	$\sum x = 40$

إذا كانت "x" هي أسعار استيراد كمية الزجاج المسطح بالطن بالريالات.

4	8	كمية الزجاج المسطح بالطن X
11	18	التكلفة بالريالات y

اكتب معادلة خط الانحدار البسيط لهذه البيانات، وفسر العلاقة

بالتعويض في معادلة التقاطع مع محور Y:

$$\begin{aligned} b &= \bar{y} - m\bar{x} \\ &= 18 - 1.61(8) \\ &= 5.12 \end{aligned}$$

بالتعويض عن قيمتي m و b في معادلة خط الانحدار البسيط تكون معادلة خط الانحدار البسيط لكمية البروتين (x) على زيادة الوزن (y):

$$\hat{y} = 1.61x + 5.12$$

بما أن قيمة الميل $m = 1.61$ موجبة فإن الارتباط بين المتغيرين ارتباط طردي.

$$\bar{x} = \frac{\sum x}{n} = \frac{40}{5} = 8$$

$$\bar{y} = \frac{\sum y}{n} = \frac{90}{5} = 18$$

بالتعويض في معادلة ميل خط الانحدار البسيط:

$$\begin{aligned} m &= \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2} \\ &= \frac{5(833) - 40(90)}{5(390) - (40)^2} \\ &= 1.61 \end{aligned}$$

إذا كانت معادلة خط الانحدار البسيط لكمية الإنتاج (x) بتكلفة (y) بآلاف الريالات هي

$$\hat{y} = 0.134x + 9.44$$

استخدم هذه المعادلة للتنبؤ بمقدار التكلفة المتوقعة في الحالات الآتية، وفسر تأثير قيمة الميل في المعادلة:

a. عدم وجود إنتاج.

b. كمية الإنتاج $x = 50$.

الحل:

a. عدم وجود إنتاج يعني أن $x = 0$ فتكون التكلفة 9,44 ألف ريال.

b. إذا كانت كمية الإنتاج $x = 50$ وحدة فإن:

$$\hat{y} = 0.134(50) + 9.44$$

$$= 16.59$$

وهذا يعني أنه عندما تصل كمية الإنتاج إلى 50 وحدة فإن التكلفة تبلغ نحو 16,59 ألف ريال.

وتدل قيمة الميل $m = 0,134$ على أنه كلما زادت كمية الإنتاج وحدة واحدة، حدثت زيادة في التكلفة بمقدار 134 ريالاً.

إذا كانت معادلة خط الانحدار البسيط لإنتاج مادة الإسفلت (بملايين البراميل) x والاستهلاك المحلي لها

y هي:

$$\hat{y} = 0.36x + 3.26$$

استخدم هذه المعادلة للتنبؤ بقيمة الاستهلاك المحلي عندما يصل الإنتاج إلى 16 مليون برميل.

$$X = 16$$

وبالتعويض في المعادلة:

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

إذن عندما يكون الإنتاج قد وصل إلى 16 مليون برميل فإن كمية الاستهلاك المحلي تساوي 9.02 مليون برميل اسمنت

1: عيّن نوع المتغير العشوائي (متصل، منفصل) في العبارات الآتية:

- عدد الأطفال في الأسرة **منفصل**
- معدل هطول الأمطار على مدار عام. **متصل**
- نسبة الزيوت الطيارة في البخور. **متصل**
- متوسط عدد ساعات الجرد السنوي في المخازن. **متصل**
- عدد السكان في جمهورية التشيك. **منفصل**
- معدل النمو السكاني في دول الخليج العربي. **متصل**
- عدد الأهداف التي سجلها أحد اللاعبين **منفصل**
- أوزان الطلاب في مدارس مدينة صبيا. **متصل**
- سرعة المركبات والسيارات التي تعبر طريق الملك عبد الله. **متصل**
- أطوال الطلاب الملتحقين بكلية الملك فهد للعلوم الأمنية. **متصل**

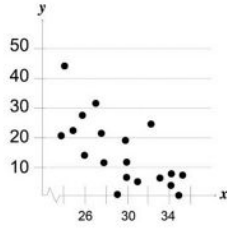
2: اكمل الجدول بما يناسب من العبارات الآتية:

- التدخين سبب رئيس للإصابة بأمراض القلب والشرابين.
- انخفاض الانبعاث الكربوني مرهون بخفض أدخنة المصانع.
- بطء سلاسل الإمداد أحد أسباب ارتفاع الأسعار.
- أحد أسباب استعمال البوليمرات في صنع أواني المطبخ هو تمتعها بخاصية العزل الحراري.

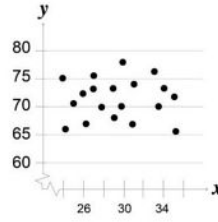
العبارة	المتغير المستقل	المتغير التابع
a	التدخين	الإصابة بأمراض القلب والشرابين
b	الانبعاث الكربوني	أدخنة المصانع
c	بطء سلاسل الإمداد	ارتفاع الأسعار
d	بخاصية العزل الحراري	استعمال البوليمرات في صنع أواني المطبخ

نعم؛ يمكن التنبؤ بقوة العلاقة من شكل الانتشار: فإذا (أو سالب) فتكون العلاقة طردية قوية (أو عكسية قوية) وإذا كانت جميع النقاط على شكل الانتشار منتشرة بشكل قريب من خط مستقيم ميله موجب (أو سالب) فتكون العلاقة طردية (أو عكسية) وإذا كانت جميع النقاط على شكل الانتشار منتشرة بشكل بعيد عن خط المستقيم فلا توجد هناك علاقة

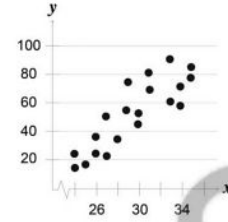
3: a. اكتب نوع الارتباط المناسب (ارتباط عكسي، ارتباط طردي، لا يوجد ارتباط) لكل شكل انتشار فيما يأتي:



ارتباط عكسي



لا يوجد ارتباط



ارتباط طردي

b. هل يمكن التنبؤ بقوة العلاقة من شكل الانتشار؟ وضع إجابتك.

4: اختر الإجابة الصحيحة فيما يأتي:

1. يمثل معامل ارتباط قيمته 0.8:

a. ارتباط خطي طردي قوي.

b. ارتباط خطي طردي متوسط.

c. ارتباط خطي عكسي قوي.

d. ارتباط خطي عكسي متوسط.

2. القيمة التي تمثل

a. 0.7

b. -0.6

c. -0.2

d. 1

3. الارتباط الذي تعبر

a. ارتباط طردي متوسط.

b. ارتباط طردي تام.

c. ارتباط عكسي ضعيف.

d. ارتباط عكسي تام.

ج 5- C

Y ²	X ²	XY	Y	X
16	2500	200	4	50
64	4225	520	8	65
25	2304	240	5	48
25	1764	210	5	42
100	3481	590	10	59
49	5184	504	7	72
36	3600	360	6	60
$\sum Y^2 = 315$	$\sum X^2 = 23058$	$\sum XY = 2624$	$\sum Y = 45$	$\sum X = 396$

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2} \sqrt{n\sum Y^2 - (\sum Y)^2}}$$

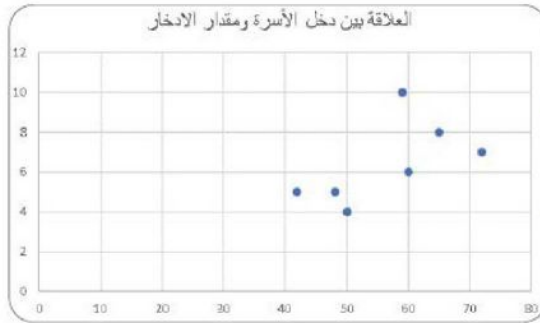
$$= \frac{7(2624) - (396)(45)}{\sqrt{7(23058) - (396)^2} \sqrt{7(315) - (45)^2}}$$

$$= \frac{548}{\sqrt{4590} \sqrt{180}} = 0.60$$

إذن يوجد ارتباط طردي متوسط بين دخل الأسرة ومقدار

الادخار.

(b) شكل الانتشار:



5: يريد باحث تحديد ما إذا كانت هناك علاقة خطية بين دخل الأسرة X ومقدار الادخار Y، لتوضيح البيانات الآتية للمتغيرين X, Y لسبع أسر:

42	48	65	50	دخل الأسرة X
5	5	8	4	مقدار الادخار Y

a. حدد المتغير المستقل والمتغير التابع.

b. ارسم شكل الانتشار للبيانات.

c. احسب معامل الارتباط، هل يوجد ارتباط بين المتغيرين والتعويض عن هذه المجاميع، وعن $n=7$ ، في قانون معامل الارتباط:

d. باعتقادك؛ ما العوامل المؤثرة على ادخار الأسرة؟ وضع إجابتك.

عدد أفراد الأسرة، المستوى التعليمي للأبوين، الوضع الصحي لأفراد الأسرة

س6: صل العبارات في (A) بما يناسبها من الرموز في (B).

(B)	(A)
b	قيمة Y المتوقعة عند قيمة X المقابلة لها
$\hat{y}$	ميل معادلة خط الانحدار البسيط
M	مقطع المحور y
$\bar{y}$	النقطة التي يمر بها خط الانحدار دائماً
$(\bar{x}, \bar{y})$	المتوسط الحسابي للمتغير Y

عدد الساعات التي يقضيها الطالب في ممارسة الرياضة

7: أعط أمثلة من مسارك التخصصي على علاقات ارتباطية بين متغيرات كل أسبوع ومؤشر كتلة الجسم لديه

طول الشخص والتاريخ الوراثي للطول

8: العبارة "يكشف معامل ارتباط بيرسون عن قوة واتجاه العلاقة بين رتب متغيرين" هل هي صحيحة دائماً، أم

صحيحة أحياناً، أم غير صحيحة أبداً؟ برر إجابتك.

غير صحيحة أبداً، يقيس معامل ارتباط بيرسون العلاقة بين متغيرين كل منهما متصل: وعند حساب هذا المعامل يفترض أن العلاقة بين المتغيرين علاقة خطية كما يفترض تجانس التباين؛ بينما معامل ارتباط سبيرمان يقيس العلاقة بين متغيرين ترتيبيين، لكن لا يشترط وجود علاقة خطية بين المتغيرين؛ وعلى الرغم من أن معامل ارتباط بيرسون وسبيرمان قد يتساوى بين متغيرين عندما تكون العلاقة متعدمة أو عشوائية. فإنه من المنطقي عدم التعويل على معامل ارتباط بيرسون لإيجاد العلاقة بين رتب متغيرين

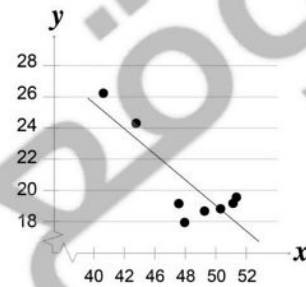
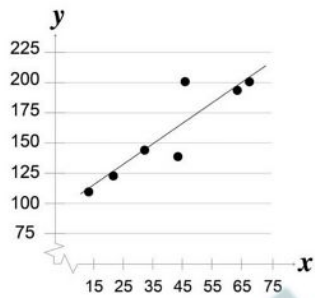
9: اكتب معادلة الانحدار الخطي البسيط من المجموعة (A) تحت الرسم البياني المناسب لها من المجموعة (B):

المجموعة (A)

a. $\hat{y} = -x + 50$

b. $\hat{y} = 1.6x + 83$

المجموعة (B)



$\hat{y} = 1.6x + 83$

$\hat{y} = -x + 50$

تطبيقات - مسار الصحة والحياة

- أجري تحليلًا إحصائيًا استدلاليًا للكشف عن الارتباط والعلاقات الخطية بين المتغيرات المرتبطة بالصحة والحياة، باستخدام برنامج الجداول الإلكترونية، بشكل احترافي، فرديًا وضمن فريق.

تطبيق:

في الجدول الآتي قائمتان من البيانات تمثلان متغيرين: "X" الذي يعبر عن العمر لعشرة أشخاص، و"Y" الذي يعبر عن قياس ضغط الدم لديهم. باستخدام برنامج الجداول الإلكترونية (إكسل Excel) احسب معامل الارتباط، وارسم شكل الانتشار، وخط الانحدار البسيط للمتغيرين.

العمر X	16	25	39	45	49	64	70	29	57	22
قياس ضغط الدم Y	109	122	143	132	199	185	199	130	175	118

اتبع الخطوات الآتية:

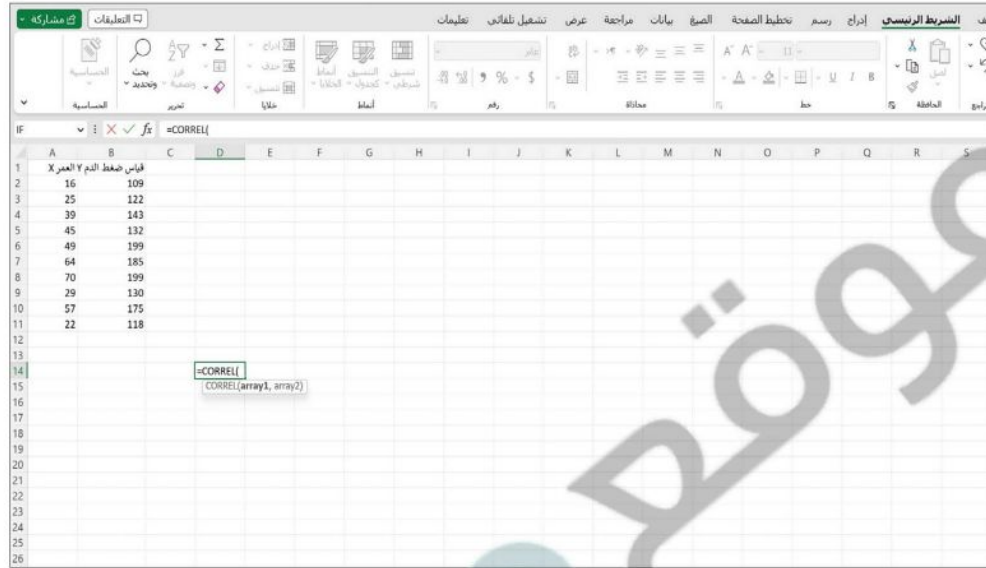
- أولاً: حساب قيمة معامل الارتباط باستخدام وظيفة CORREL :

إدخال البيانات في برنامج الجداول الإلكترونية كالآتي:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1		قياس ضغط الدم Y																	
2		16	109																
3		25	122																
4		39	143																
5		45	132																
6		49	199																
7		64	185																
8		70	199																
9		29	130																
10		57	175																
11		22	118																

لحساب معامل الارتباط بين هذين المتغيرين حدد خلية فارغة ثم أدخل الصيغة =CORREL، كما يوضح

الشكل الآتي:

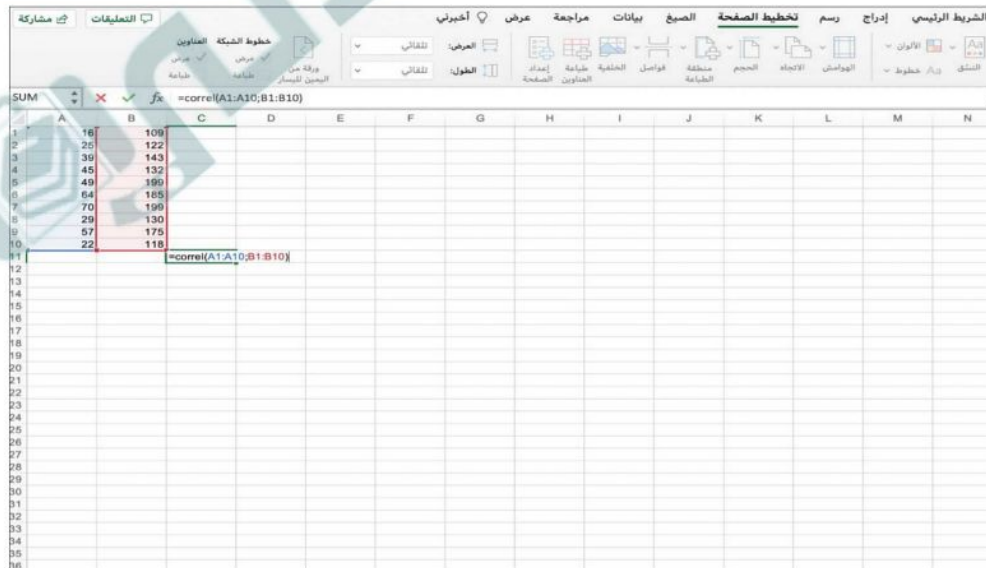


لاحظ:

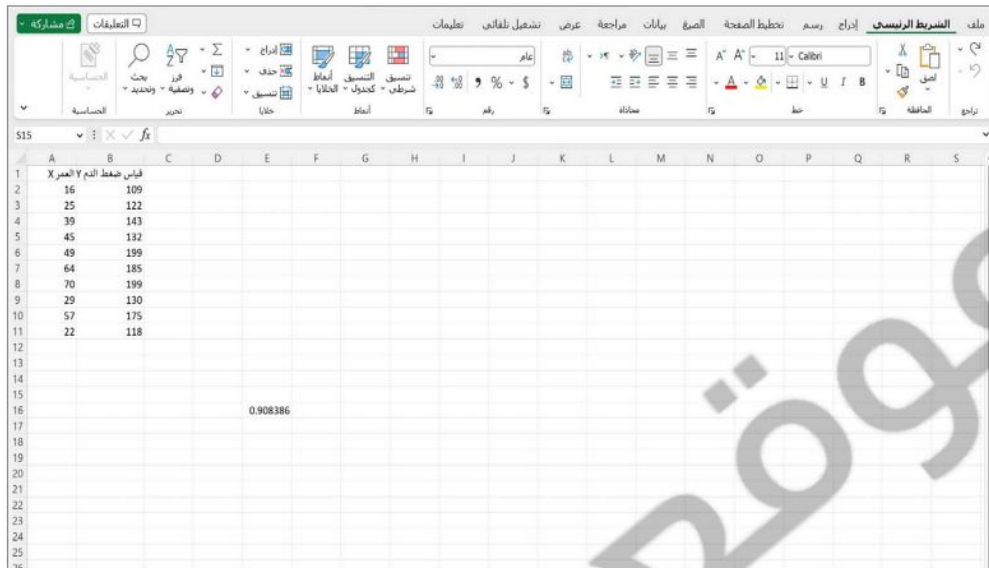
في بعض الإصدارات والإصدارات من Excel،
توضع فاصلة منقوطة بدلاً من فاصلة.

بالضغط على array1، ثم حدد العمود A كاملاً.

ضع فاصلة، ثم اضغط على array2 وحدد العمود B كاملاً.



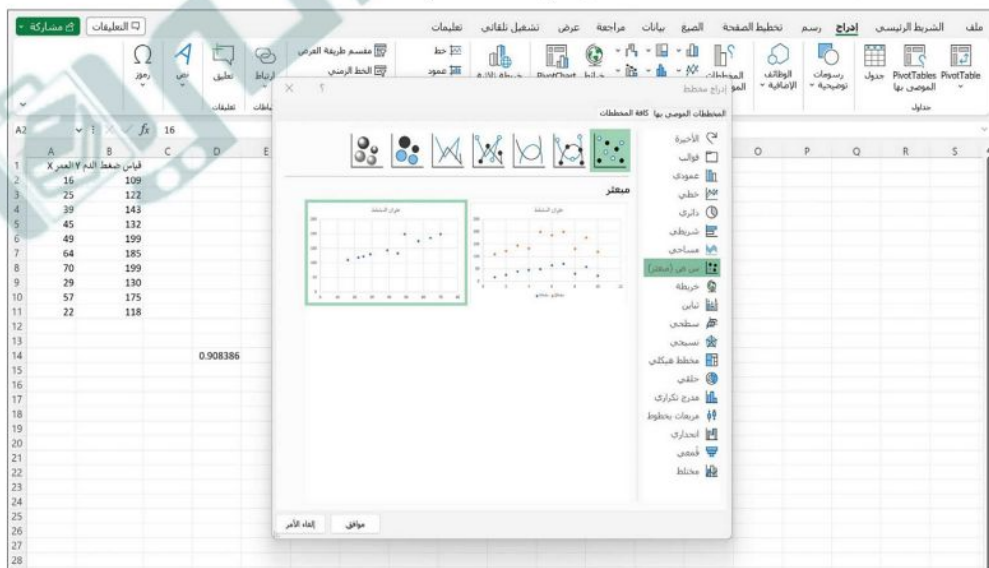
اضغط على زر إدخال فتظهر نتيجة الارتباط كما في الشكل أدناه:



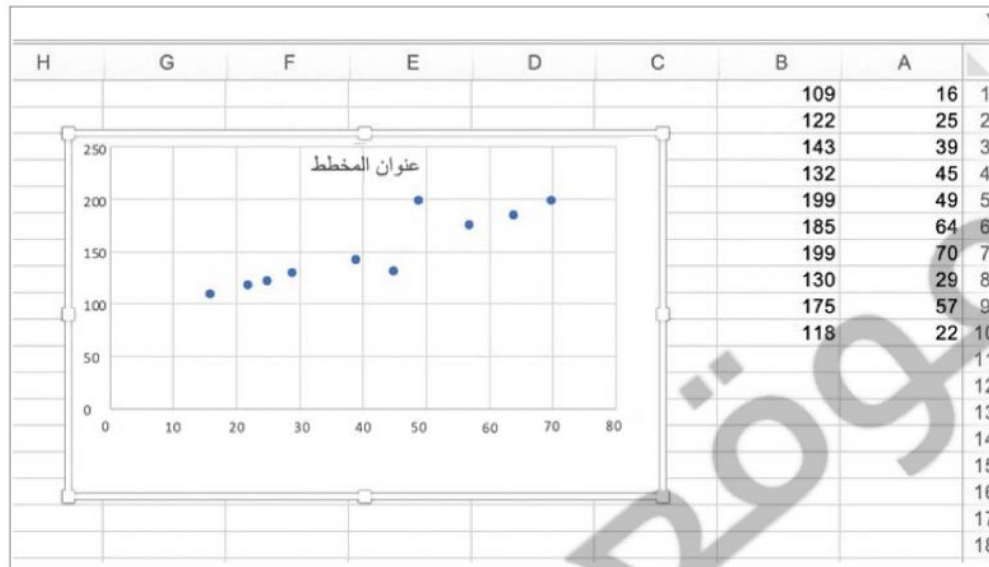
النتيجة: معامل الارتباط يساوي تقريباً 0.9، وهذا يدل على وجود ارتباط قوي بين المتغيرين، وهو ما يعني وجود علاقة بين التقدم في العمر وبين الإصابة بارتفاع ضغط الدم.

• ثانياً: رسم شكل الانتشار:

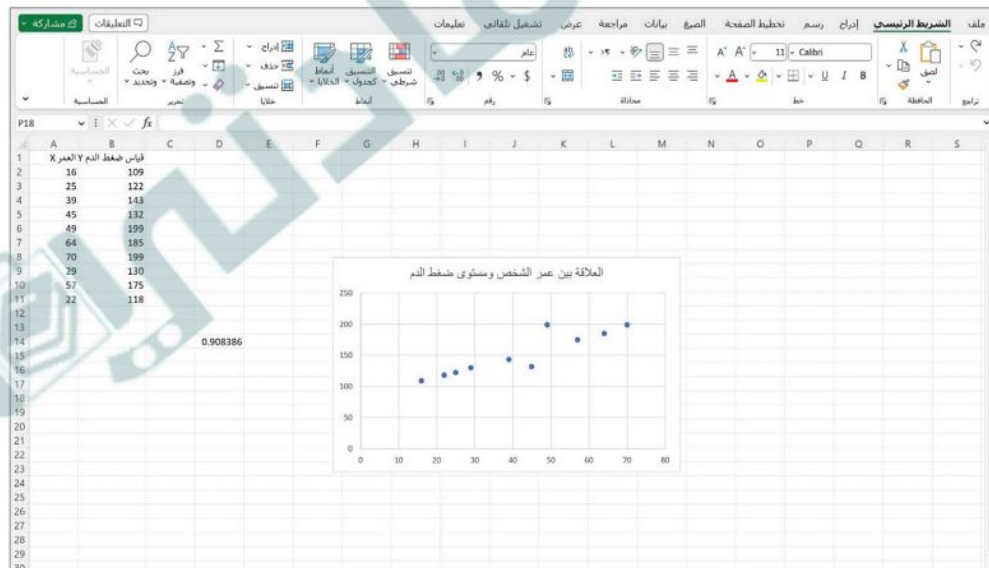
لرسم شكل الانتشار: حدد البيانات في العمودين، ثم انقر فوق علامة التبويب إدراج على الشريط، وحدد موقع قسم المخططات، ثم اختر مبعثر، كما هو موضح في الشكل الآتي:



فيظهر شكل الانتشار للبيانات كما يأتي:



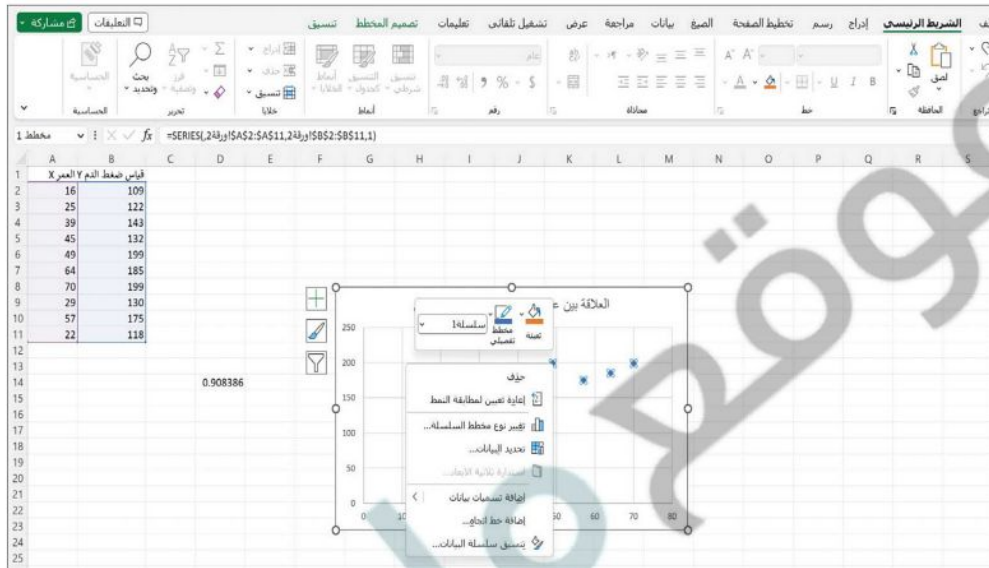
وبالضغط على عنوان المخطط يمكن كتابة عنوان المخطط حسب مسمى التمثيل.



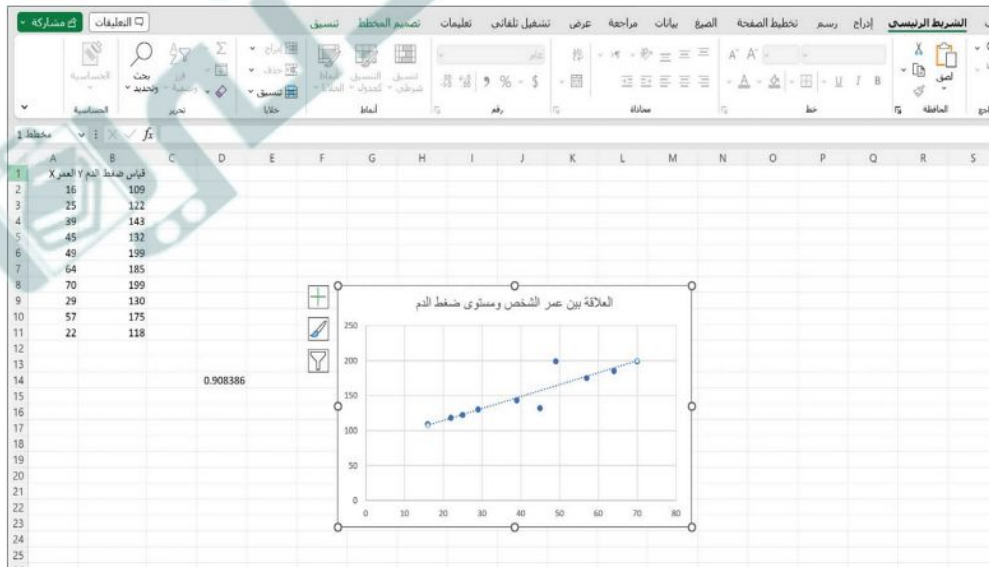
ويوضح الشكل أنه يوجد ارتباط طردي (موجب) بين المتغيرين.

• ثالثاً: رسم خط الانحدار البسيط:

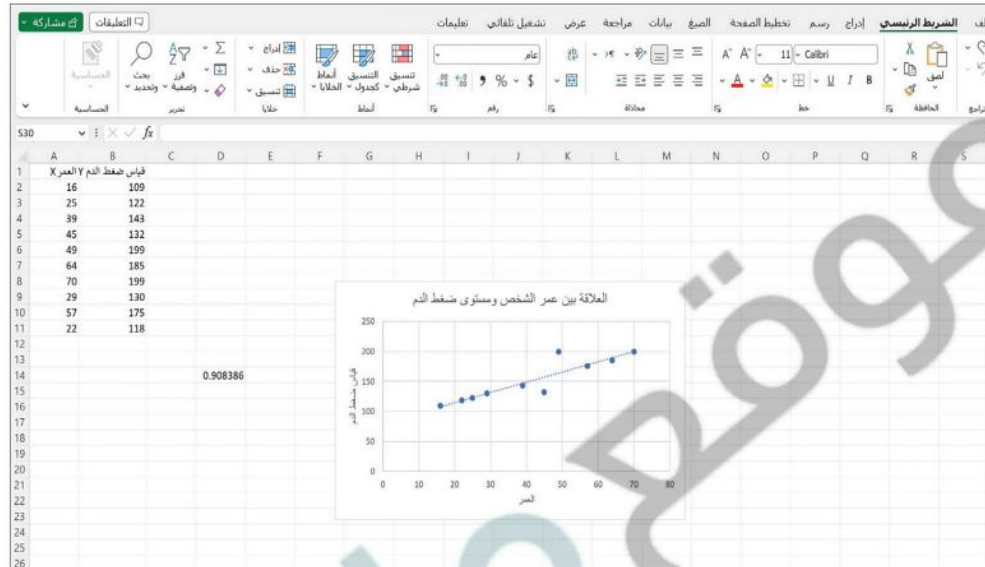
ابداً بالنقر مرة واحدة على أي نقطة بيانات في مخطط الانتشار، ثم اضغط مرة واحدة على زر الفأرة الأيمن، فتظهر قائمة كما في الشكل أدناه.



اختر إضافة خط اتجاه من القائمة، فتحصل على خط الانحدار كما هو موضح في الشكل الآتي:



لاحظ أنه عند النقر على إشارة (+) في أعلى يسار الرسم تظهر قائمة مقترحة بإضافات مهمة للتمثيل البياني، مثل: كتابة عناوين للمحاور، وإضافة تنسيقات أخرى.



تمارين:

1: يرى الاختصاصي النفسي أن خضوع المصابين في حوادث سير عنيفة لدعم نفسي يمكن أن يحسن حالتهم بشكل أسرع نسبياً، ويستدل بتجربة قام فيها بإخضاع مجموعة من المصابين لدعم نفسي، حيث عرض البيانات من خلال الجدول الآتي:

D	C	B	A
		2	3
		2	4
		2	2
		1	2
		1	2
		1	2
		0.65465367	7
			8
			9
			10
			11
			12

3	عدد المصابين X
2	عدد الذين تحسنت حالتهم بشكل أسرع نتيجة الخضوع لدعم نفسي Y

احسب معامل الارتباط بين المتغيرين، باستخدام برنامج الج

محققاً. فسّر إجابتك.

معامل الارتباط = 0.65 إذن: نعم يوجد ارتباط طردي متوسط بين عدد المصابين وعدد الذين تحسنت حالتهم بشكل أسرع نتيجة للخضوع للدعم النفسي

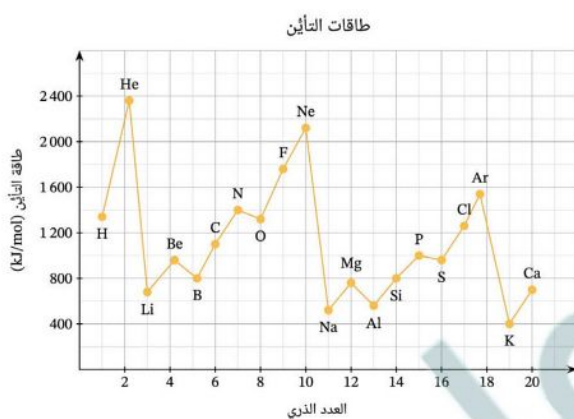
أولاً: حساب رتب التقديرات في المتغيرين، نفرض أن التقدير (مقبولة) تناظره رتبة (منخفضة) تساوي 1، يليها تقدير (جيدة) وتناظره الرتبة 2، ثم تقدير (جيدة جداً) وتناظره الرتبة 3 وهكذا

2: توضح البيانات الآتية إجابات عينة من سبعة أشخاص؛ عن سؤالين الأول حول برامج العلاج الطبيعي، والثاني عن مدى استجابتهم للعلاج:

السؤال الأول X	جيدة	مقبولة	ممتازة	جيدة	جيدة جداً	مقبولة
السؤال الثاني Y	جيدة جداً	مقبولة	جيدة جداً	جيدة	جيدة	ممتازة

احسب معامل سبيرمان لارتباط الرتب. فسر إجابتك.

3: يعرض التمثيل البياني المجاور طاقة التأين لعدد من العناصر.



a. دون إجراء حسابات؛ توقع قوة واتجاه الارتباط الخطي للعناصر:

Li, Be, B, C, N, O, F, Ne •

Ar, K, Ca •

b. وفقاً للتمثيل؛ هل العبارة "كلما زاد العدد الذري زادت طاقة التأين" صحيحة؟ فسر إجابتك.

4: توضح البيانات الآتية عدد ساعات ممارسة الرياضة في الأسبوع X ومعدل ساعات النوم اليومية Y كما يأتي:

9	6	7	4	2	1	X
7	4	5	8	8	6	Y

احسب معامل الارتباط الخطي، ما مدى قوة العلاقة الخطية بين المتغيرين؟

5: توضح البيانات الآتية العلاقة بين كثافة الهواء X وبين درجة الحرارة المثوية Y:

1.40	1.30	1.20	1.05	1	0.9	0.8	X (kg/m³) الكثافة
80	100	140	200				Y (C°) الحرارة

باستخدام برنامج الجداول الإلكترونية:

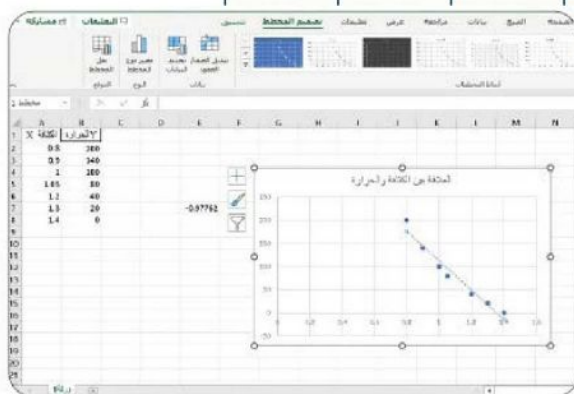
a. احسب معامل الارتباط. 0.98 تقريباً

b. ارسم شكل الانتشار.

c. ارسم خط الانحدار.

d. اكتب تقريراً عن النتيجة.

من خلال قيمة معامل الارتباط وشكل الانتشار نلاحظ أن الارتباط بين كثافة الهواء ودرجة الحرارة يمثل ارتباطاً عكسياً شبه تام، وذلك يعني أنه كلما زادت كثافة الهواء قلت درجة الحرارة والعكس صحيح



ثانياً: حساب قيمة معامل سبيرمان كما يأتي:

d^2	d	R_y	R_x	Y	X
1	-1	3	2	جيدة جداً	جيدة
0	0	1	1	مقبولة	مقبولة
1	1	3	4	جيدة جداً	ممتازة
0	0	2	2	جيدة	جيدة
1	1	2	3	جيدة	جيدة جداً
1	-1	2	1	جيدة	مقبولة
4	-2	4	2	ممتازة	جيدة
$\sum d^2 = 8$	$\sum d = 2$				

$$R = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6(8)}{7(49 - 1)} = 0.86$$

أي أن هناك علاقة ارتباط طردية قوية بين السؤالين X,Y.

ج3: 3:

(a) - Li, Be, B, C, N, O, F, Ne (ارتباط طردي قوي).

Ar, K, Ca (لا يوجد ارتباط).

(b) غير صحيحة، مثال للعدد الذري H طاقة التأين تساوي

1240 تقريباً؛ بينما العدد الذري Ca طاقة التأين تساوي

470 تقريباً.

4:

ج4:

Y^2	X^2	XY	Y	X
36	1	6	6	1
64	4	16	8	2
64	16	32	8	4
25	49	35	5	7
16	36	24	4	6
49	5184	63	7	9
$\sum Y^2 = 254$	$\sum X^2 = 187$	$\sum XY = 176$	$\sum Y = 38$	$\sum X = 29$

بالتعويض عن هذه المجاميع، وعن $n = 6$ ، في قانون معامل الارتباط.

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2} \sqrt{n\sum Y^2 - (\sum Y)^2}}$$

$$= \frac{6(176) - (29)(38)}{\sqrt{6(187) - (29)^2} \sqrt{6(254) - (38)^2}}$$

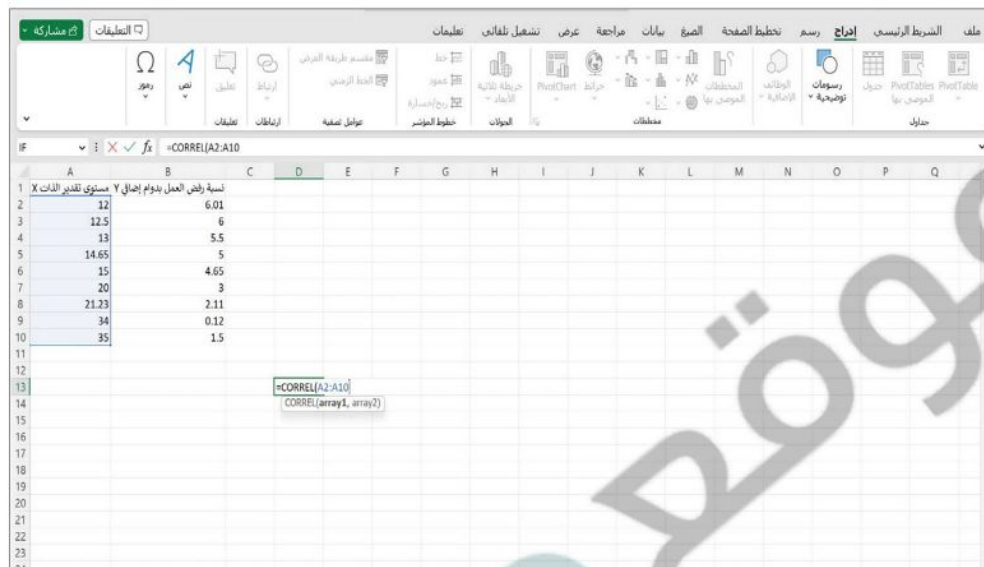
$$= \frac{-46}{\sqrt{281} \sqrt{80}} = -0.31$$

إذن يوجد ارتباط عكسي ضعيف بين عدد ساعات ممارسة الرياضة في الأسبوع وعدد

ساعات النوم.

لحساب معامل الارتباط بين هذين المتغيرين؛ حدد خلية فارغة ثم أدخل الصيغة =CORREL، كما يوضح

الشكل الآتي:

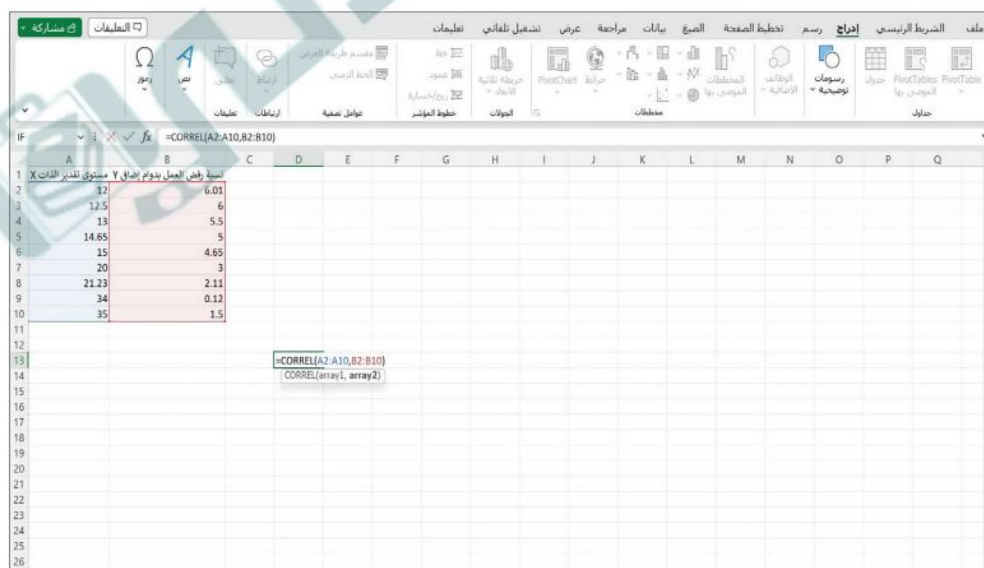


لاحظ:

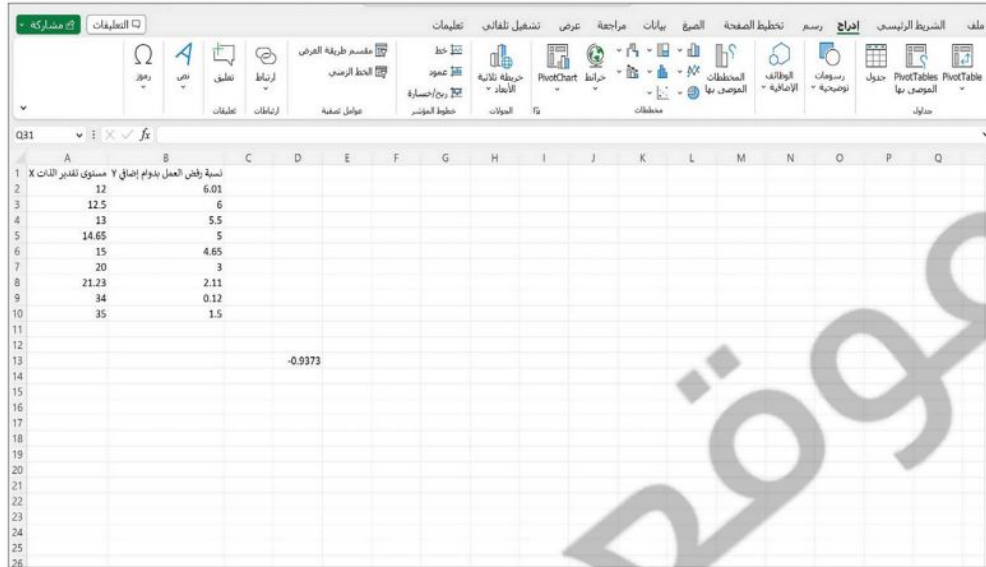
في بعض الإصدارات والإصدارات من Excel،
توضع فاصلة منقوطة بدلاً من فاصلة.

بالضغط على array1 ثم حدد العمود A كاملاً.

ضع فاصلة، ثم اضغط على array2 وحدد العمود B كاملاً.



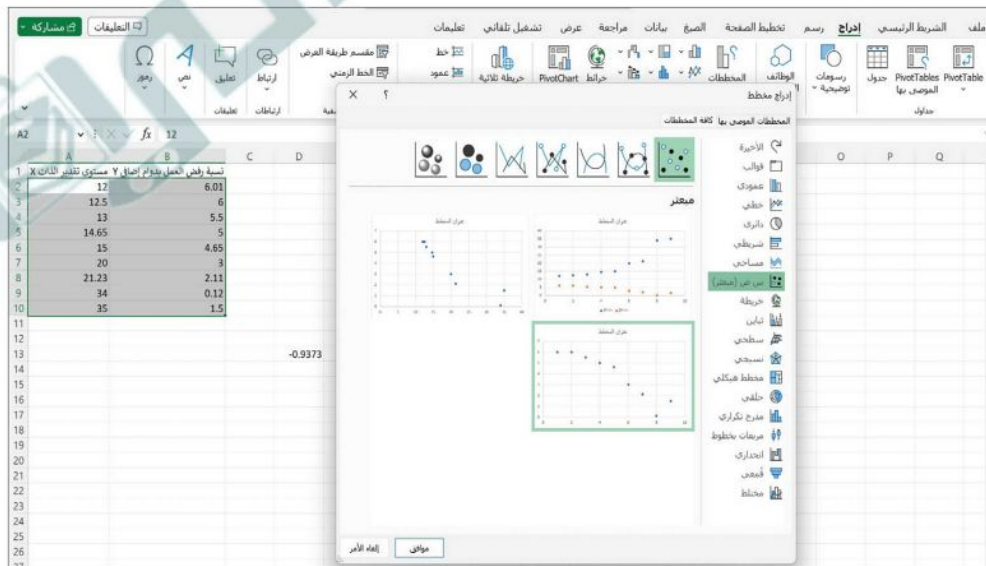
اضغط على زر إدخال فتظهر نتيجة الارتباط كما في الشكل أدناه:



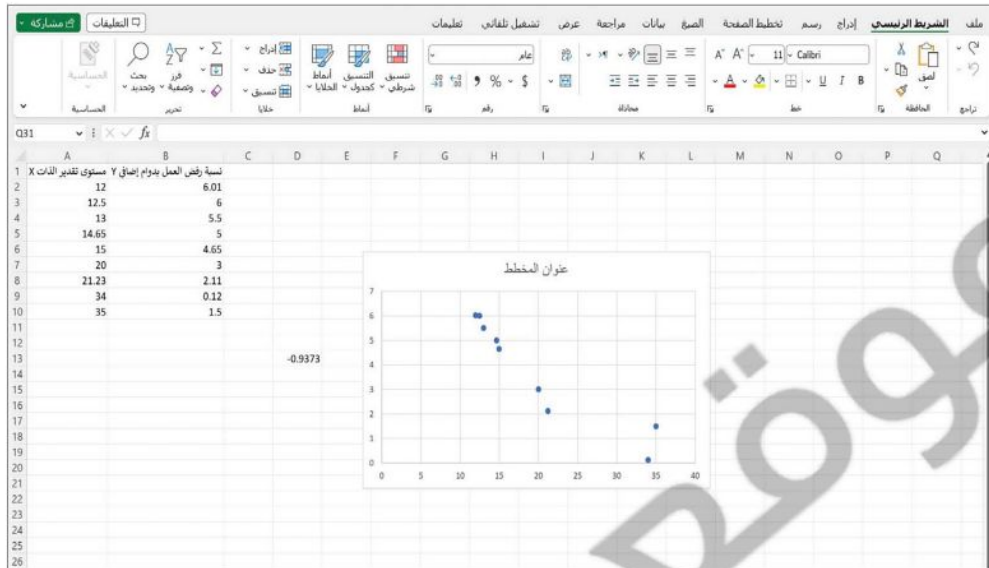
النتيجة: معامل الارتباط يساوي 0.9- وهذا يعني أنه يوجد ارتباط عكسي قوي بين المتغيرين؛ أي أنه كلما زاد مستوى تقدير الذات لدى الموظف قلت نسبة رفضه العمل بدوام إضافي.

• ثانيًا: رسم شكل الانتشار:

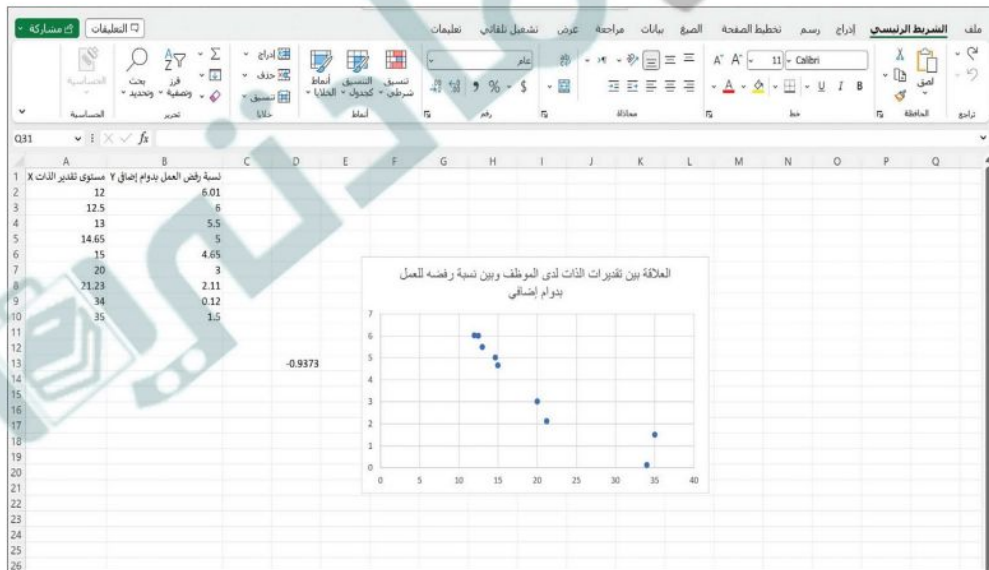
لرسم شكل الانتشار حدد البيانات في العمودين، بعد ذلك انقر فوق علامة التبويب إدراج على الشريط وحدد موقع قسم المخططات. اختر مبعثر، كما هو موضح في الشكل الآتي:



فيظهر شكل الانتشار للبيانات كما يأتي:



وبالضغط على عنوان المخطط يمكن كتابة عنوان المخطط حسب مسمى التمثيل:



• ثالثاً: رسم خط الانحدار البسيط:

ابدأ بالنقر مرة واحدة على أي نقطة بيانات في مخطط الانتشار، ثم اضغط مرة واحدة على زر الفأرة الأيمن،

فتظهر لك قائمة كما في الشكل أدناه:



ثانياً: حساب قيمة معامل سبيرمان كما يأتي:

d^2	d	R_y	R_x	Y	X
1	-1	3	2	جيدة جداً	جيدة
0	0	1	1	مقبولة	مقبولة
1	1	3	4	جيدة جداً	ممتازة
0	0	2	2	جيدة	جيدة
1	1	2	3	جيدة	جيدة جداً
1	-1	2	1	جيدة	مقبولة
4	-2	4	2	ممتازة	جيدة
$\sum d^2 = 8$	$\sum d = 2$				

D	C	B	A	
		2	3	1
		2	4	2
		2	2	3
		1	2	4
		1	2	5
		1	2	6
	0.65465367			7
				8
				9
				10
				11
				12
				13

$$R = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6(8)}{7(49 - 1)} = 0.86$$

أي أن هناك علاقة ارتباط طردية قوية بين السؤالين X, Y.

تمارين:

1: توضح البيانات أدناه حجم الإنتاج X وحجم الصادرات Y للنقط الخام، في إحدى الدول (بالمليار برميل) خلال عدة سنوات كما يأتي:

حجم الإنتاج X	3	4	2	2	2
حجم الصادرات Y	2	2	2	1	1

احسب معامل الارتباط بين المتغيرين، باستخدام برنامج الجداول الإلكترونية، وفسر إجابتك.

معامل الارتباط يساوي 0.65 إذن يوجد ارتباط طردي متوسط بين حجم الإنتاج وحجم الصادرات

2: توضح البيانات الآتية إجابات عينة من سبعة أشخاص؛ عن سؤالين الأول حول برامج الضمان الاجتماعي، والثاني

أولاً: حساب رتب التنديرات في المتغيرين: نفرض أن التقدير (مقبولة) تتناظره رتبة (منخفضة) عن مدى ملاءمتها لاحتياجات المستفيدين: تساوي 1؛ يليها تقدير (جيدة) وتناظره الرتبة 2 ثم تقدير (جيدة جداً) وتناظره الرتبة 3، وهكذا

السؤال الأول X	جيدة	مقبولة	ممتازة	جيدة	جيدة جداً	مقبولة
السؤال الثاني Y	جيدة جداً	مقبولة	جيدة جداً	جيدة	جيدة	ممتازة

احسب معامل سبيرمان لارتباط الرتب. فسر إجابتك.

3: يعرض التمثيل البياني المجاور مقارنة

بين كمية الإنتاج الصناعي في المملكة العربية

السعودية خلال عدد من السنوات:

a. بدون إجراء حسابات؛ حدد قوة واتجاه

خلال الفترات:

2022 (ارتباط طردي قوي)

الأشهر الثلاثة الأولى من العام 2020 (لا

يوجد ارتباط)

2019 (لا يوجد ارتباط)

الربع الثاني من 2020 (ارتباط عكسي قوي)

م.2020.

في الربع الأول من العام (أي خلال الأشهر الثلاثة الأولى)؛ كان هناك ثبات في مستوى الإنتاج الصناعي ثم بدأ تزايد الإنتاج الصناعي بشكل كبير خلال نهاية الربع الأول (الشهر الثالث) في الربع الثاني (خلال الأشهر الرابع والخامس والسادس) كان هناك تناقص ملحوظ في الإنتاج الصناعي، ولكن مع بداية الربع الثالث بدأ الإنتاج الصناعي في التحسن: ثم ثبت مستوى الإنتاج واستمر حتى نهاية الربع الرابع

b. قدم وصفاً للإنتاج الصناعي للمملكة خلال العام 2021م على أساس ربع سنوي.

4: يريد مدير شركة معرفة ما إذا كانت هناك علاقة خطية بين عدد سنوات الخدمة للموظفين X ومساهماتهم في

الأنشطة السنوية المجتمعية التابعة للشركة (بآلاف الريالات) Y. ويوضح الجدول الآتي بيانات هذه الدراسة:

30	24	3	15	5	10	1	X
2	3	9	5	14	8	12	Y

a. احسب معامل الارتباط الخطي، وهل توجد علاقة بين المتغيرين؟

b. باعتقادك؛ لماذا ظهرت هذه النتيجة؟ وضع إجابتك.

5: توضح البيانات الآتية المساحة بالمتر المربع X، وأسعار المنزل بآلاف الريالات Y؛ لسبعة منازل في أحد أحياء مدينة

الرياض:

1278	1312	1552	2332	2413	1592	1924	المساحة X
1450	990.9	1200	2190.9	2750	1360.9	1740.9	السعر Y

باستخدام برنامج الجداول الإلكترونية:

a. احسب معامل الارتباط.

b. ارسم شكل الانتشار.

c. ارسم خط الانحدار.

d. اكتب تقريراً عن النتيجة.

