

Coeficiente de correlación

Por:

Carmen J. Velázquez Rivera

CORRELACION LINEAL ENTRE DOS VARIABLES

ESTADISTICA INFERENCIAL

Materia: Matemáticas

NIVEL: Secundario

Grado: 10

Concepto Principal: Coeficiente de correlación lineal r de Pearson

Conceptos Secundarios:

- Desviación de la media

- Media aritmética de la muestra

- Regresión lineal (recta de regresión)

- Correlación positiva y negativa

- Fuerza de la correlación

- Índices de correlación

- Diferencia entre correlación y causalidad

CORRELACION LINEAL ENTRE DOS VARIABLES

Conocimiento Previo:

Determinar las medidas de tendencia central y de dispersión (media, mediana, moda, rango)

Determinar la variable dependiente e independiente de los datos

Construir tabla de datos de una ecuación lineal en dos variables

Localizar coordenadas en el plano cartesiano

Representar datos de forma gráfica

Objetivos específicos

Mediante la experiencia de aprendizaje, los estudiantes:

Determinarán la media de la muestra.

Representarán los datos de forma gráfica.

Identificará si la correlación es positiva, negativa o cero.

Determinarán la fuerza de la correlación.

Calcularán el coeficiente de correlación lineal r de Pearson.

Interpretarán la información de los datos obtenidos.

Determinarán la diferencia entre correlación y causalidad.

Glosario

1. Estadística – es la ciencia de obtener, clasificar, organizar e interpretar datos que se presentan a una valoración numérica.
2. Estadística descriptiva – nos indica cual tal es una situación, describe e informa lo que hay de tal modo que permite describir y resumir las observaciones que se hagan sobre un asunto, fenómeno o problema de investigación.
3. Estadística inferencial – cuando se infiere o se deduce una observación de unos datos estadísticos obtenidos de una muestra, la cual se generaliza sobre la población en total.

Glosario

- 4. Causalidad – una relación causa-efecto entre variables.
- 5. Coeficiente de correlación lineal r de Pearson – mide la fuerza de la relación lineal entre los valores cuantitativos apareados x y y en una muestra. También se conoce como coeficiente de correlación producto momento de Pearson, en honor a Karl Pearson (1857 – 1936), quien lo desarrollo originalmente.
- 6. Correlación – la relación entre dos conjuntos de datos.
- 7. Correlación positiva y negativa – (dirección de la relación), dos conjuntos de datos pueden tener correlación positiva si aumenta o disminuyen juntos y correlación negativa si uno de los conjuntos aumenta y el otro conjunto disminuye. Si uno de los conjuntos aumenta o disminuye y el otro se queda igual, la correlación es cero.

Glosario

8. Desviación media – estadística que depende o toma como referencia la media o puntuación promedio de un grupo o distribución para indicar el nivel de dispersión de los datos en una distribución.
9. Índices de correlación – los valores dados a la r de Pearson que van desde -1.0 hasta +1.0, pasando por cero. Estos determinan el nivel de la relación entre las variables.
10. Niveles de correlación – niveles que clasifican la relación entre las variables desde cero, baja, moderada, alta o perfecta, ya sean positiva o negativas.

Glosario

- 11. Media de la muestra – la puntuación promedio de la distribución de los datos de la muestra.
- 12. Recta de regresión – recta que representa la relación entre dos variables, se encuentra lo más próxima posible a todos los puntos de datos.

Materiales

1. libreta de papelotes
2. marcadores de colores
3. cinta adhesiva de papel (masking tape)
4. reglas de 12 pulgadas
5. lápices
6. papel cuadriculado
7. pega
8. tijeras
9. calculadora gráfica
10. hoja de actividades
11. cronómetro

Pre y pos prueba

La pre y pos prueba consistirá de varios ítems de selección múltiple los cuales consistirán en:

- a) Modelos de gráficas con diferentes tipos de correlación en el cual el participante debe identificar la dirección de la correlación.
- b) Modelos de gráficas con diferentes tipos de correlación en el cual el participante debe identificar la fuerza de la correlación.
- c) Indicar cuál es la media de la muestra en un ejercicio.
- d) Mostrar la diferencia entre lo que es correlación y lo que es causalidad.

Assessment

A través de la lección se realizarán actividades de exploración, desarrollo y cierre dirigidas al desarrollo de las destrezas y conceptos establecidos, asegurando el entendimiento y el aprendizaje de parte de los participantes.

Entre las actividades se incluyen hojas de trabajo individual, trabajo en grupo y uso de la calculadora gráfica.

Inicio: Exploración

Se entrega a cada estudiante la copia de una prueba de inteligencia lógico matemática, cada una identificada con un número para mantener la confidencialidad de los participantes. Esta prueba contiene varias preguntas que el estudiante debe contestar haciendo uso de una escala de distribución Likert. Se le indica que escriban en la parte inferior de la hoja el total de la puntuación obtenida en la prueba, así como el porcentaje de aprovechamiento académico obtenido en mayo 2014 en la clase de matemáticas.

Inicio: Exploración

Mediante un dialogo socializado se discuten las siguientes preguntas:

¿Qué podemos esperar de un estudiante que obtiene una puntuación alta en la prueba de inteligencia lógica matemática comparado con el porciento de aprovechamiento académico en el curso de matemáticas?

¿Existirá una relación entre los resultados de la prueba y el aprovechamiento académico de los estudiantes en el curso de matemáticas?

Inicio: Exploración

¿Cómo podemos establecer la relación entre la prueba y el porcentaje de la clase?

Luego de la discusión de la situación se procede a realizar las preguntas de la actividad 1, ya que los resultados de la prueba de inteligencia emocional se trabajará en la actividad de cierre.

Desarrollo

Se presentan varias actividades

Actividad 1: ¿Qué conoces sobre estadística?

Instrucciones:

Dividir el grupo en sub grupos de 5 estudiantes. Se entrega a cada grupo la hoja de trabajo #1. Contestar las preguntas, luego escribirán en un papelote su contestación, el cual pegarán a la pared y se discutirán al finalizar la capacitación.

Hoja de trabajo actividad 1:

¿Qué conocimiento tienes?

Instrucciones:

Contesta las siguientes preguntas. Escríbelas en el papelote y pégalas a la pared.

1. ¿Qué es estadística?
2. ¿La estadística, se divide en varias categorías? Explica tu contestación.
3. Menciona algunos conceptos estadísticos que recuerdes.
4. ¿Para qué se utiliza la estadística?
5. Menciona algunas profesiones que utilizan estadística.

Actividad 2: Marca el punto X

En esta actividad se recolectarán datos.

1. Se entrega a los estudiantes una hoja en blanco y una con una tabla y un plano cartesiano.
2. Se indica que escriban en la tabla la predicción de cuantas X pueden escribir en 5 segundos.
3. En la hoja en blanco, anotarán las X que pueden escribir en 5 segundos. El maestro toma el tiempo indicando a los estudiantes cuando comienzan y cuando terminan de escribir las X según el tiempo.
4. El estudiante contará la cantidad de X que escribieron en 5 segundos y las registrarán en la columna de 5 segundos.

Actividad 2: Marca el punto X

5. Luego harán la predicción de cuántas X pueden escribir en 10 segundos y 20 segundos y las anotará en la tabla.
6. Ahora se tomará el tiempo real de 10 y 20 segundos de cuántas X logran escribir en esos tiempos y las anotarán en la tabla.
7. Los estudiantes graficarán los datos en el plano cartesiano y observarán su gráfica.

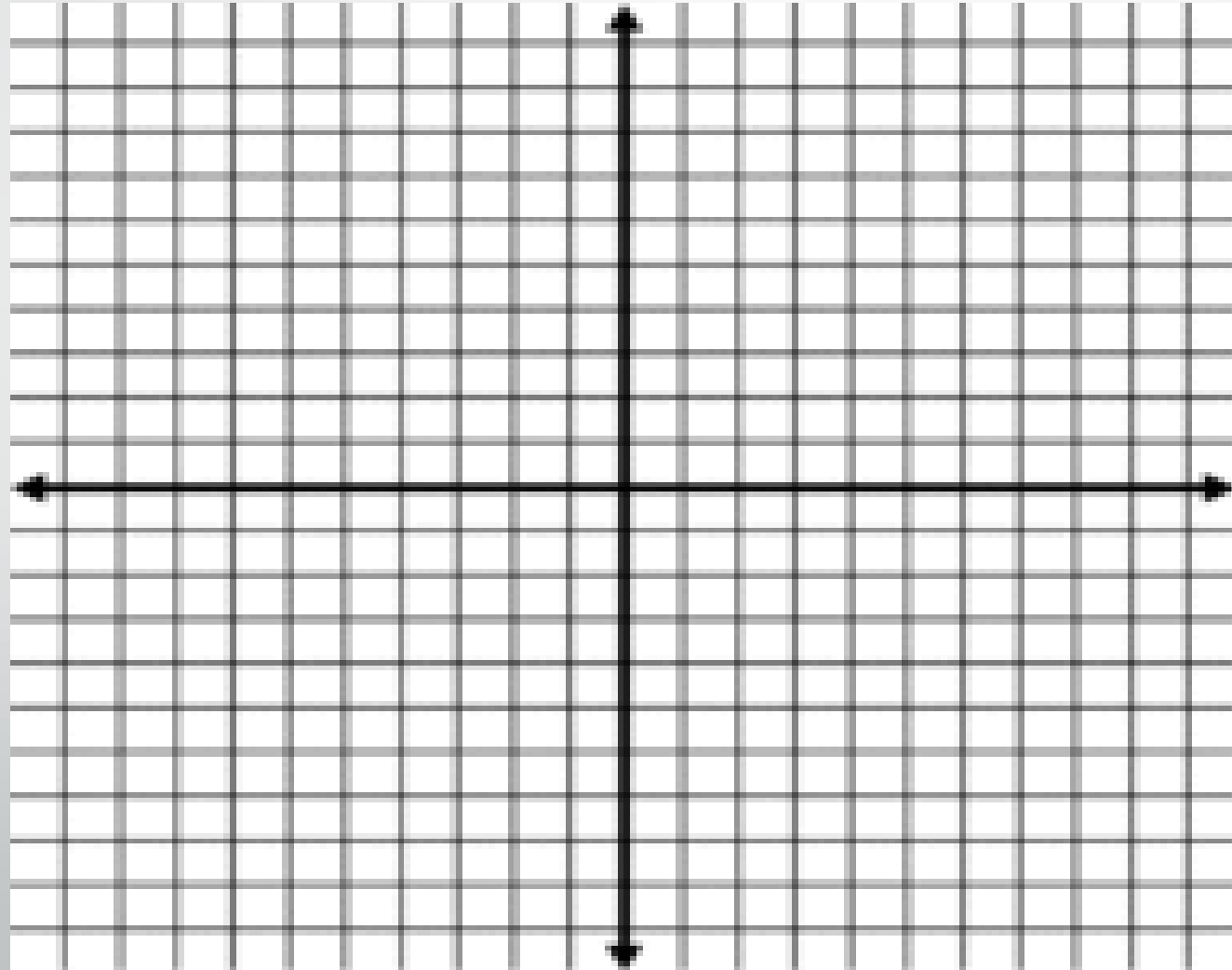
Se realizarán las siguientes preguntas basadas en la observación de las gráficas.

1. ¿Todas las gráficas son iguales? ¿Son todas lineales?
2. ¿Qué factores podrían haber contribuido a que los datos no sean perfectamente lineales?

Hoja de actividad 2

Tiempo	Predicción	Tiempo Real
5 segundos		
10 segundos		
20 segundos		

Hoja de actividad 2



Coeficiente de Correlación lineal r de Pearson

Definición: Una correlación existe entre dos variables cuando una de ellas está relacionada con la otra de alguna manera.

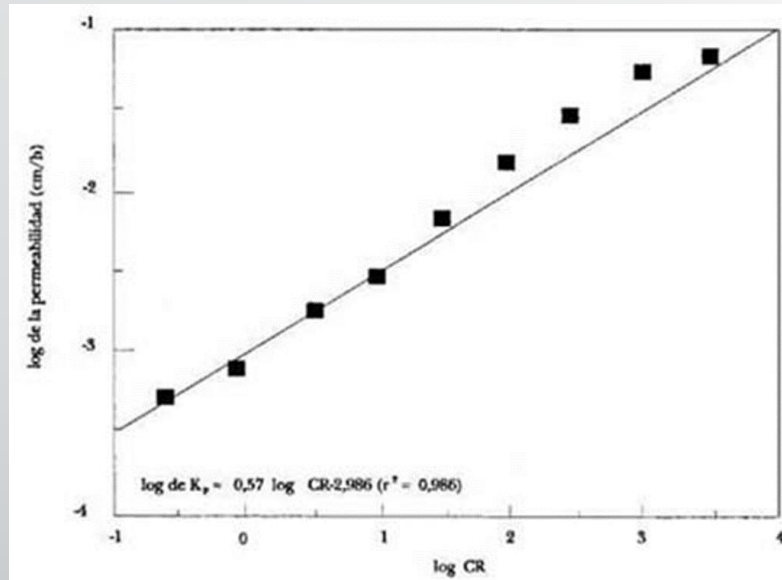
El coeficiente de correlación r , mide la fuerza de la relación lineal entre los valores cuantitativos apareados x y y en una muestra.

En estos diagramas se observa la relación que existe entre dos variables. Cuanto más se acercan los datos de las coordenadas, mayor es la relación entre la variable X y la variable Y .

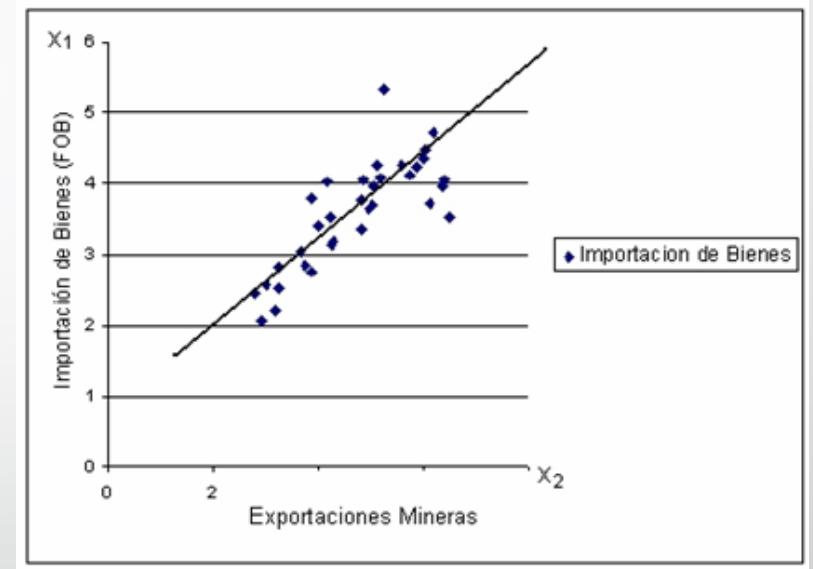
En la gráfica de dispersión podemos observar la dirección y la fuerza o nivel de la relación entre las variables.

Tipos de correlación: (algunos ejemplos)

Correlación positiva alta



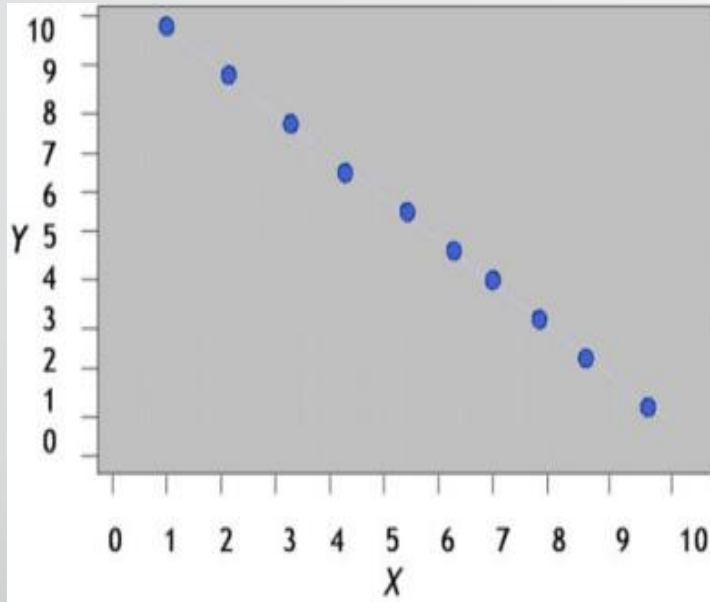
Correlación moderada alta



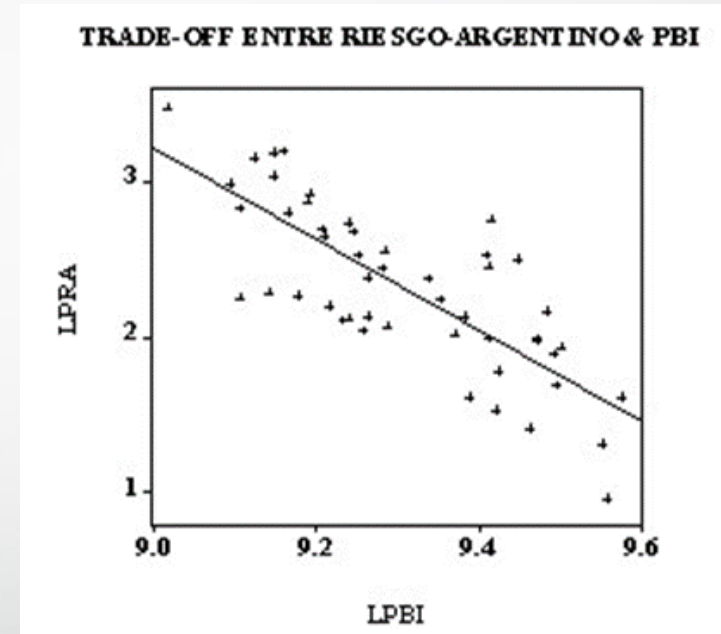
Cuando la dirección es positiva, los datos van en forma ascendente tanto en X como en Y. Si los datos se agrupan en una línea ascendente, esto es indicativo que la relación entre variables es una positiva perfecta. Cuando X aumenta, Y también aumenta.

Tipos de correlación: (algunos ejemplos)

Negativa perfecta



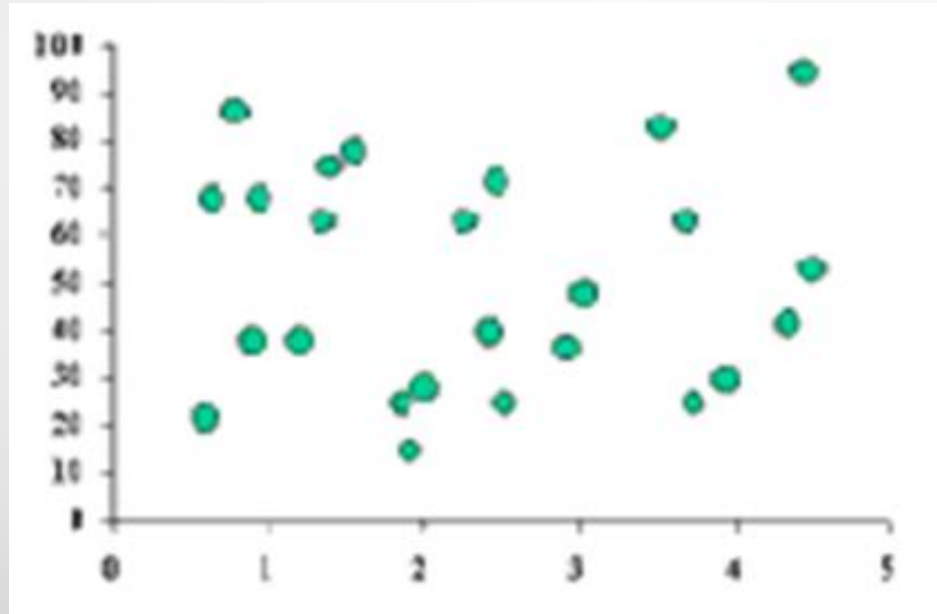
Negativa moderada



Cuando los datos van en forma descendente la correlación es negativa, X aumenta, Y disminuye. Si los datos se agrupan en una línea descendente, esto es indicativo que la relación entre variables es una negativa perfecta.

Tipos de correlación: (algunos ejemplos)

Correlación nula



Cuando las puntuaciones de una variable no se relacionan con las puntuaciones de la otra variable, es una correlación cero.

Interpretación de los índices de correlación

Los índices de correlación fluctúan entre:

- Negativo uno (-1) hasta positivo uno (+1) pasando por el cero (0).
- Los signos determinan la dirección.
- Cuando $r = (+1)$, la correlación entre las variables es positiva perfecta y el diagrama de dispersión presenta una línea perfecta.
- Cuando $r = (-1)$, la correlación entre las variables es negativa perfecta y el diagrama de dispersión presenta una línea recta descendente.

Interpretación de los índices de correlación

Para saber si la correlación existe, necesitas un número, una información cuantitativa que indique la dirección de la correlación y cuán fuerte es esta relación.

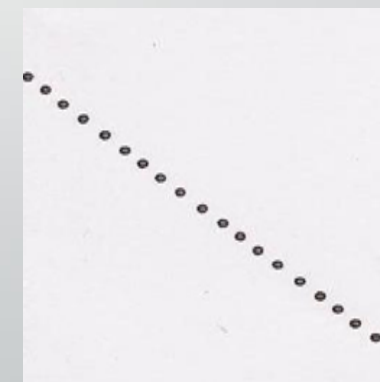
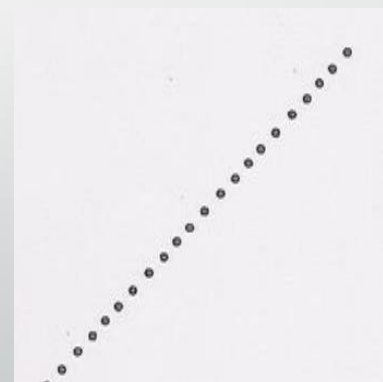
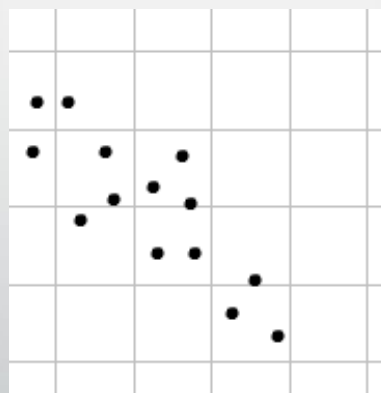
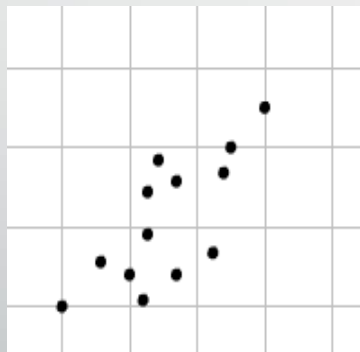
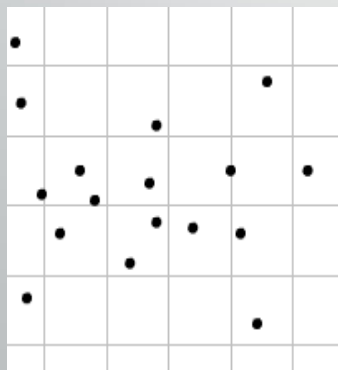
Correlación positiva	Correlación negativa
+ 1.00 - correlación perfecta	- 1.00 - correlación perfecta
+ 0.76 – 0.99 - correlación alta	- 0.76 – -0.99 - correlación alta
+ 0.51 – 0.75 - correlación moderada alta	- 0.51 – -0.75 - correlación moderada alta
+ 0.26 – 0.50 - correlación moderada baja	- 0.26 – -0.50 - correlación moderada baja
+ 0.01 – 0.25 - correlación débil baja	- 0.01 – -0.25 - correlación débil baja
0.00 cero correlación	0.00 cero correlación

No existen coeficientes de correlación sobre 1.00, ni bajo -1.00. Cuando se informa de un coeficiente sobre 1.00 o bajo -1.00, se concluye que se realizaron los cálculos erróneamente.

Actividad 3: ¿Qué tipo de correlación es?

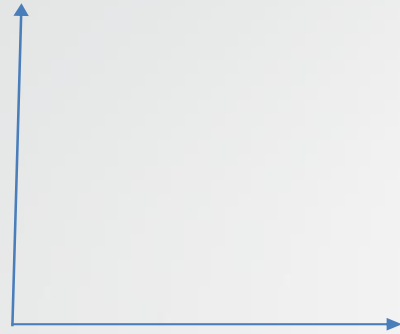
Construye tu gráfica de correlación positiva, negativa, positiva perfecta, negativa perfecta y correlación cero.

Instrucciones: Recorta los diagramas de dispersión y pégalos en el papel de la gráfica que corresponde según la dirección de la correlación.

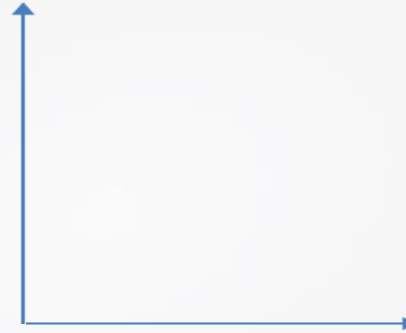


Pega aquí los diagramas de dispersión según su correlación.

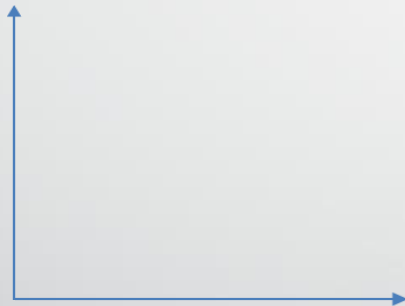
Correlación cero



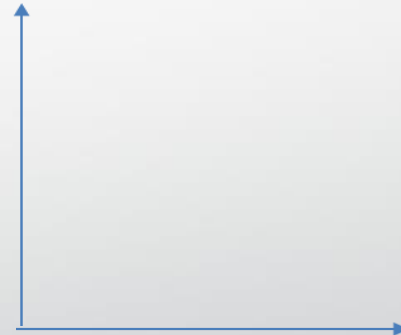
Correlación positiva perfecta



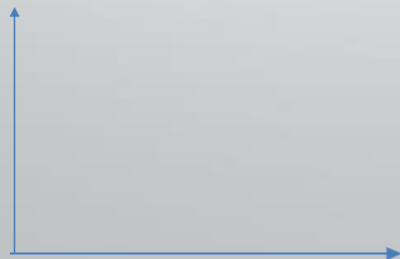
Correlación negativa



Correlación negativa perfecta



Correlación positiva



Correlación o causalidad

Una correlación no es lo mismo que causalidad. Cuando se habla de causalidad se habla de causa-efecto. Para entender este concepto veremos el siguiente ejemplo.

No siempre el hecho de que haya una alta correlación implica causalidad.

Por ejemplo: Un estudio de escuela elemental con niños de edades de 6 a 11 años, encuentra una alta correlación positiva de la talla de zapatos y las puntuaciones en el IQ. ¿Tiene sentido decir que el que tenga la talla de zapato más grande tiene IQ mayor? O, ¿un IQ alto causa el tener una talla grande? ¿Qué explicaría esta correlación? Los estudiantes deben justificar su razonamiento como clase.

Actividad 4: Correlación o causalidad

Los estudiantes determinarán las diferencias entre correlación y causalidad al explicar estos pares de variables fuertemente asociados.

- a. Usar un audífono.
- b. La cantidad de leche que toma una persona.
- c. La cantidad de dinero que una gana.
- d. La capacidad de gimnasio de una escuela de pueblo.

- a. Morir en los siguientes diez años.
- b. La fuerza en los huesos de una persona.
- c. El número de años que una persona asistió a la escuela.
- d. Número de iglesias en el mismo pueblo.

Fórmula: Coeficiente de correlación lineal r de Pearson

Se presentan dos fórmulas distintas para determinar el coeficiente de correlación de Pearson

Primera fórmula

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

Para esta actividad se escoge de forma aleatoria 10 estudiantes de la clase. A estos estudiantes la maestra le proveerá una hoja de papel con un numero asignado, esto para mantener la confidencialidad de los estudiantes. En la misma el estudiante escribirá el resultado de la prueba de inteligencia emocional y el porcentaje obtenido del curso de matemáticas de mayo 2014. Entregará la hoja con los datos a la maestra. Estos datos se utilizarán para determinar el coeficiente de correlación lineal r de Pearson.

Prueba de inteligencia lógica

➔ 3. INTELIGENCIA LÓGICO-MATEMÁTICA

Ítem	S	s	n	N
1. Aprecia enigmas, contraseñas, problemas de lógica.				
2. Realiza cálculos mentales.				
3. Aprecia proponer problemas de cálculos u otras operaciones.				
4. Analiza datos con facilidad.				
5. Trabaja bien con promedios, proporciones y otros esquemas.				
6. Trabaja bien con medidas, números, nociones de estadística.				
7. Aprecia realizar experiencias con palitos, agua, arena, etc.				
8. Percibe la geometría en los objetos y paisajes que ve.				
9. Busca la secuencia lógica en las ideas.				
10. La falta de patrones de regularidad en las cosas lo molesta.				

Explicación de las variables de la fórmula

r = coeficiente de correlación lineal Pearson

X = una variable independiente de intervalo razón (puntuaciones del test de inteligencia lógica-matemática)

Y = una variable dependiente de intervalo razón (aprovechamiento académico del curso de matemáticas en porcentajes)

n = tamaño de la muestra

Σ = el resultado de la suma total de la columna

ejemplo: ΣY = total de la columna Y

Σx^2 = total de la columna x^2

Pasos a seguir para determinar el coeficiente de correlación lineal usando la primera fórmula. Esta práctica es dirigida.

1. Organizar los datos a comparar, según el valor de X , en forma descendente.

El valor de Y depende de X , según la relación de las puntuaciones.

2. Añade una columna de x^2 , cuadrando los valores de x , y una columna de y^2 , cuadrando los valores de y .
3. Multiplica cada valor de X por cada valor de Y , para crear la columna XY .
4. Suma todos los valores de cada columna.
5. Sustituye los valores de las variables en la fórmula.
6. Resuelve según el orden de las operaciones.

Determina el coeficiente de correlación lineal usando la primera fórmula

Estudiantes	X	Y	x^2	y^2	xy
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

Segunda fórmula

$$r = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

Pasos a seguir para determinar el coeficiente de correlación lineal usando la segunda fórmula. Esta práctica es dirigida.

1. Organizar los datos de las columnas X y Y
2. Determinar la media de las columnas X y Y.
3. Para la columna $(x = X - \bar{x})$ se resta la media \bar{x} a cada valor de X, estos valores son las desviaciones de la media de x.

Pasos para determinar el coeficiente de correlación lineal utilizando la segunda fórmula.

4. Para la columna $(y = Y - \bar{y})$ se resta la media \bar{y} a cada valor de Y , estos valores son las desviaciones de la media de y .
5. Crea una nueva columna elevando al cuadrado cada valor de x y otra columna elevando al cuadrado cada valor de y .
6. Multiplicar cada valor de x por y para crear la columna xy .
7. Sumar cada columna y anotar el total en la parte inferior de la misma.
8. Sustituir los valores en la fórmula.
9. Resuelve según el orden de las operaciones.

Determina el coeficiente de correlación lineal utilizando la segunda fórmula.

Estudiante	X	Y	$x = X - \bar{x}$	$y = Y - \bar{y}$	x^2	y^2	xy
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							

Cierre

- En la actividad de cierre los participantes entrarán los datos de la situación presentada al principio, en la calculadora para crear una gráfica. Los datos que se utilizarán son de 10 participantes que se seleccionaron al azar previamente. Luego se analizará la gráfica para indicar el tipo de correlación de la situación. En el análisis los participantes indicarán si el caso es de correlación o causalidad.

Actividad utilizando la calculadora gráfica TI Nspire CX CAS.



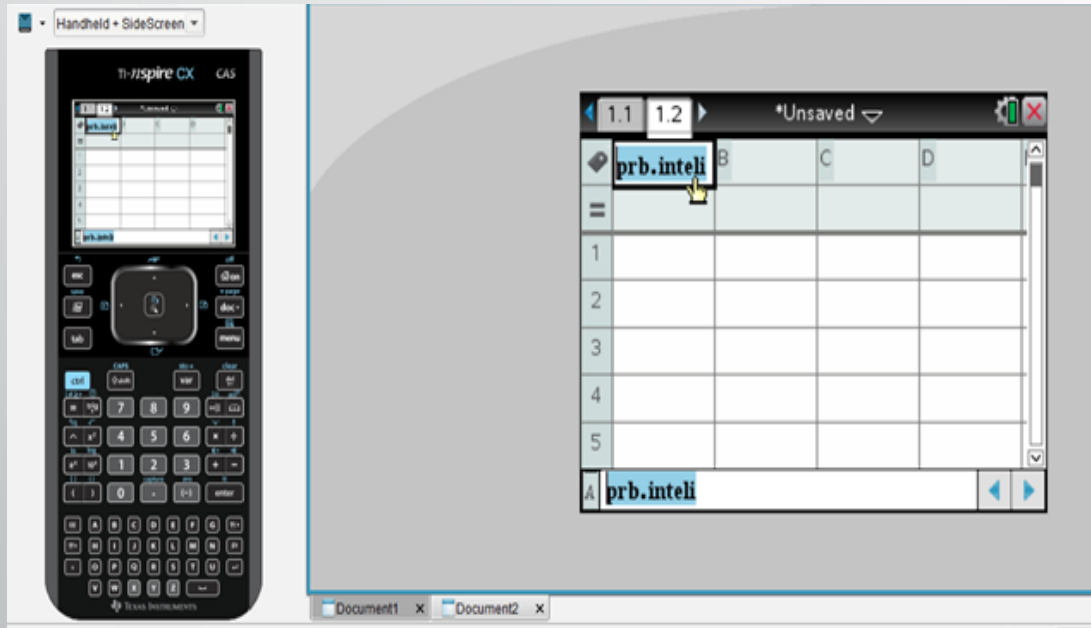
Home – este botón te lleva donde están todos los iconos.

Actividad utilizando la calculadora gráfica TI Nspire CX CAS



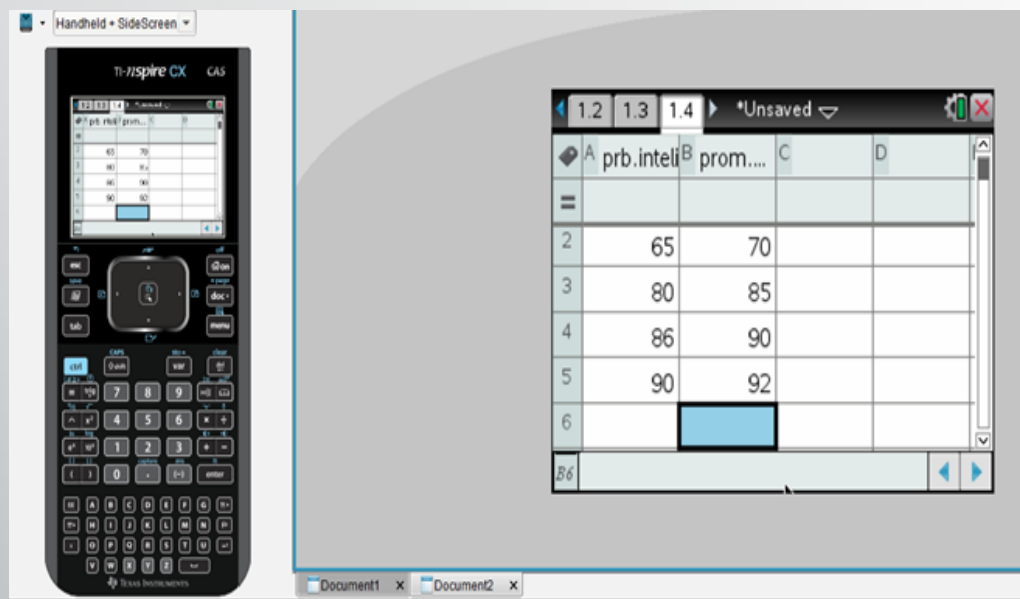
2. Colocar el cursor sobre el icono de las tablas y oprimir enter.

Actividad utilizando la calculadora gráfica TI Nspire CX CAS



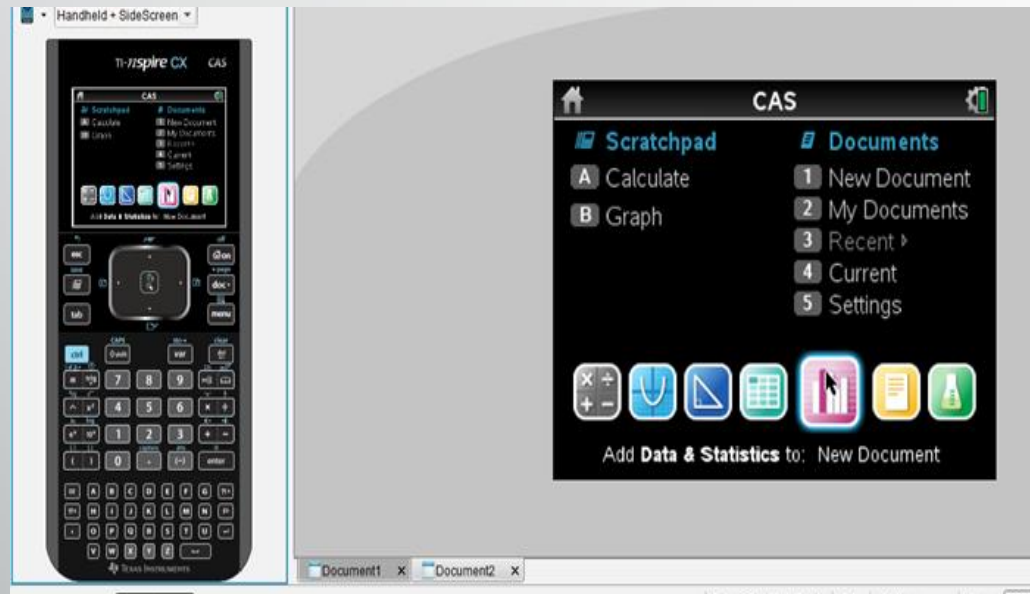
3. Ubicarse en la primera columna, primera fila donde esta la A.
4. Escribir el nombre de la columna (prb.inteli), esto es prueba de inteligencia.
5. Entrar los datos de X (prueba de inteligencia) comenzando en la celda uno, de la tabla realizada anteriormente en la columna A, un dato por celda. Bajar a la próxima celda utilizando el cursor.

Actividad utilizando la calculadora gráfica TI Nspire CX CAS



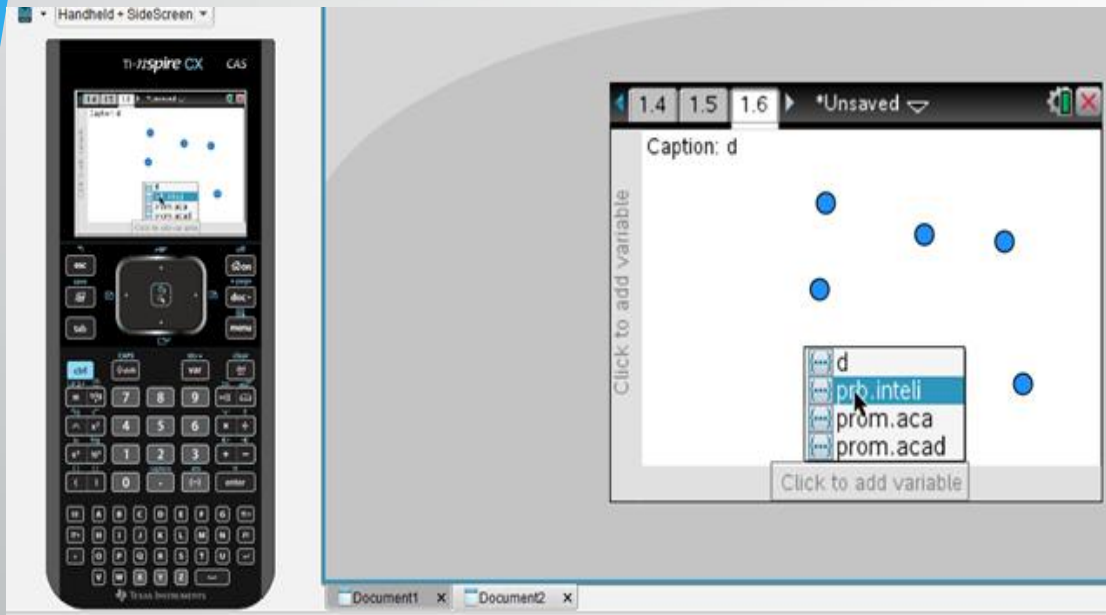
6. Ubicarse en la segunda columna (B) y escribir el nombre de la columna (prom.acad).
7. Entrar los datos de Y (promedio académico) de la tabla realizada anteriormente en la segunda columna (B).

Actividad utilizando la calculadora gráfica TI Nspire CX CAS



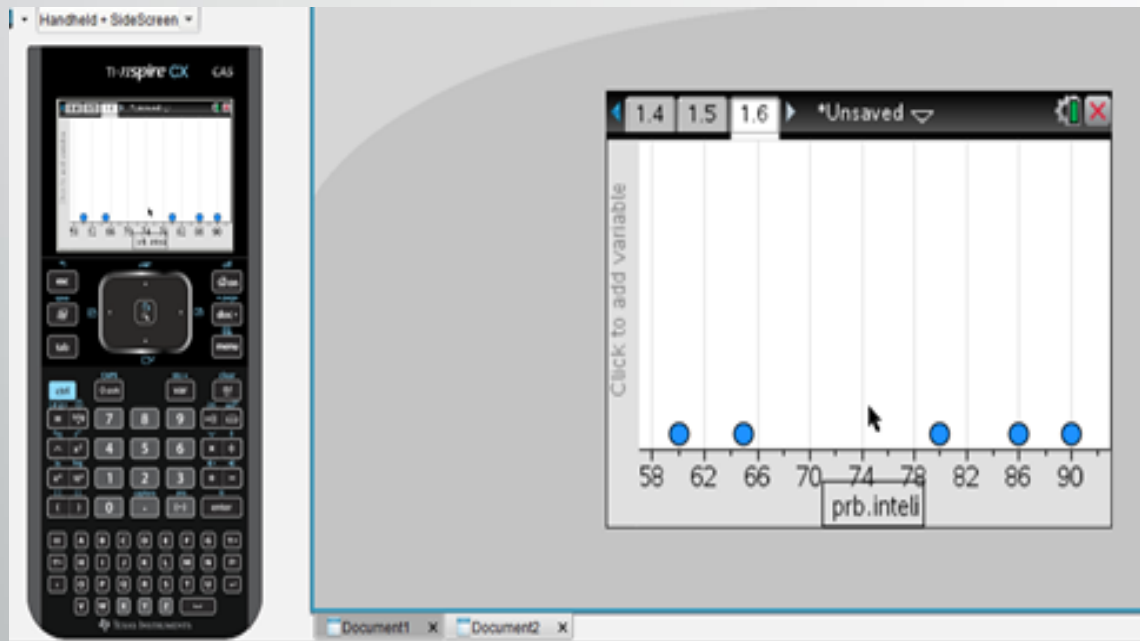
8. Oprimir el botón de Home
(botón que tiene el dibujo de la casita).
9. Seleccionar el icono de las gráficas.

Actividad utilizando la calculadora gráfica TI Nspire CX CAS



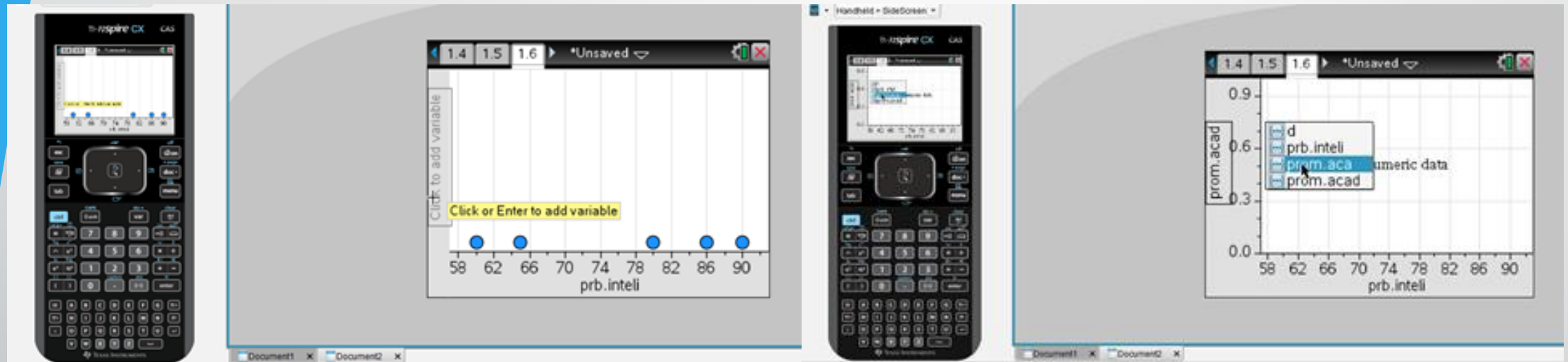
10. Se mostrarán los datos de forma desordenada.

Actividad utilizando la calculadora gráfica TI Nspire CX CAS



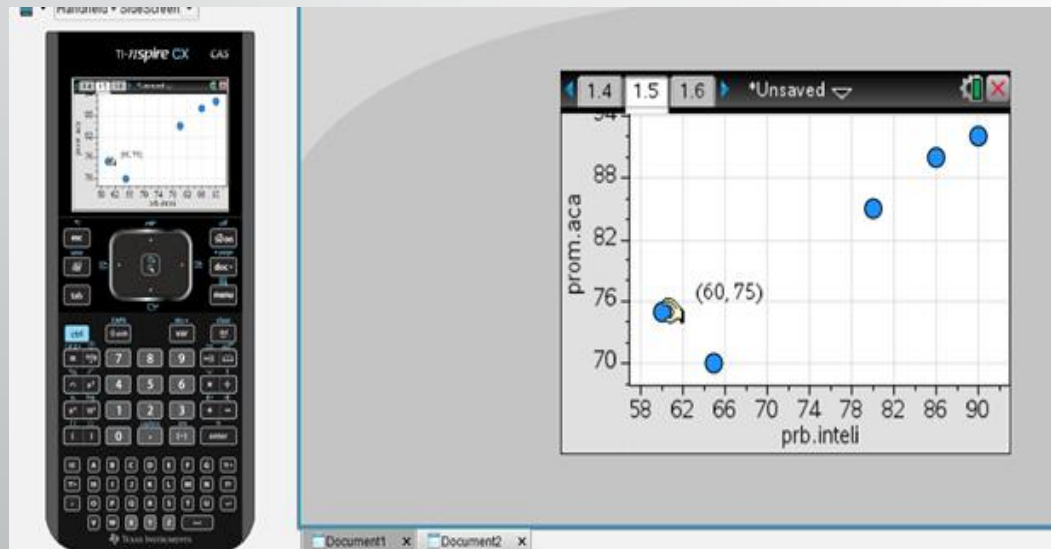
11. Frota el mouse pad , saldrá un mensaje para escoger una variable (click or enter to add variable).
12. Selecciona prb.inteli. ya que este es el nombre de la columna A, que es nuestro eje de X.
13. Los datos se acomodan en el eje X.

Actividad utilizando la calculadora gráfica TI Nspire CX CAS



14. Frota el mouse pad , saldrá un mensaje para escoger una variable (click or enter to add variable).
15. Selecciona prom.acad. ya que este es el nombre de la columna B, que es nuestro eje de Y.
16. Los datos se acomodan en el eje Y mostrando la gráfica finalizada.

Actividad utilizando la calculadora gráfica TI Nspire CX CAS

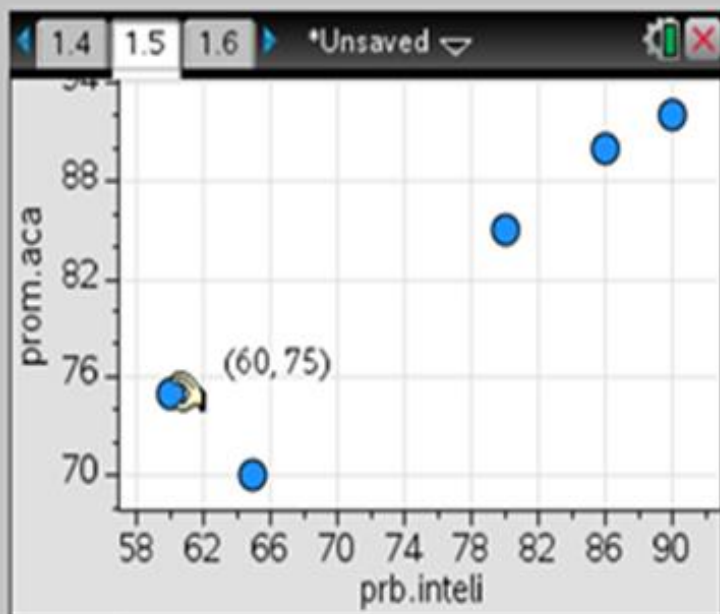


17. Si ubica el cursor sobre los puntos, se muestran las coordenadas de los mismos.

Observa la gráfica creada e indica qué tipo de correlación es. Explica tu respuesta.

La situación presentada es de correlación o de causalidad. Explica tu respuesta.

Actividad utilizando la calculadora gráfica TI Nspire CX CAS



Observa la gráfica creada e indica qué tipo de correlación es.

Explica tu respuesta.

La situación presentada es de correlación o de causalidad.

Explica tu respuesta.



Pos prueba

Referencias

- Antúnez, C. (2006) ¿Cómo identificar en usted y en sus alumnos las inteligencias múltiples? Recuperado el 11 de septiembre 2014 en: [ile:///C:/Users/User/Desktop/21ANTUNES-Celso-Cap2-Descubriendo-las-IM-en-sus-alumnos.pdf](file:///C:/Users/User/Desktop/21ANTUNES-Celso-Cap2-Descubriendo-las-IM-en-sus-alumnos.pdf)
- Aron, A., Aron, E., Coups, E. (2006). Statistics for Psychology. Sexta Edición. New Jersey. Pearson Prentice Hall.
- Departamento de Educación de Puerto Rico, (2014). Estándares de Contenido y Expectativas de Grado. Programa de Matemáticas. San Juan, PR.
- Departamento de Educación de Puerto Rico, (2014). Mapas Curriculares Grado 10 (2014). Programa de Matemáticas. San Juan, PR.
- Larousse, diccionario esencial Matemáticas, (s.f.), (2008). Editorial Larousse. México.