

Correlación Estadística

- 1) (10%) Una empresa comercial tiene establecimientos en varias ciudades. La gerente general de ventas piensa lanzar un anuncio comercial por televisión, varias veces antes de una venta gigante que empezará el sábado y terminará el domingo. La gerente planea conseguir las cifras de las ventas del sábado y domingo en sus diferentes locales y emparejarlas con el número de veces que apareció el comercial en televisión. El objetivo principal de la investigación es determinar si existe relación entre el número de veces que se transmitió el anuncio y las ventas de sus productos. Los pares de datos son:

Establecimiento de Ventas	Número de Anuncios Transmitidos	Ventas de Sábado y Domingo (miles de Bs)
Caracas	4	150
Maracay	2	80
Valencia	5	210
Barquisimeto	6	240
Maracaibo	3	170

Se pide:

- ¿Cuál es la variable dependiente y por qué?
 - Dibuje el diagrama de dispersión. Indique el título, las leyendas y unidades correspondientes.
 - Viendo el diagrama de dispersión, ¿parece haber alguna relación entre la variable x y la y ? Explique su respuesta.
 - Calcule el coeficiente de correlación.
 - Interprete el resultado anterior.
-

Solución

Tabla de datos

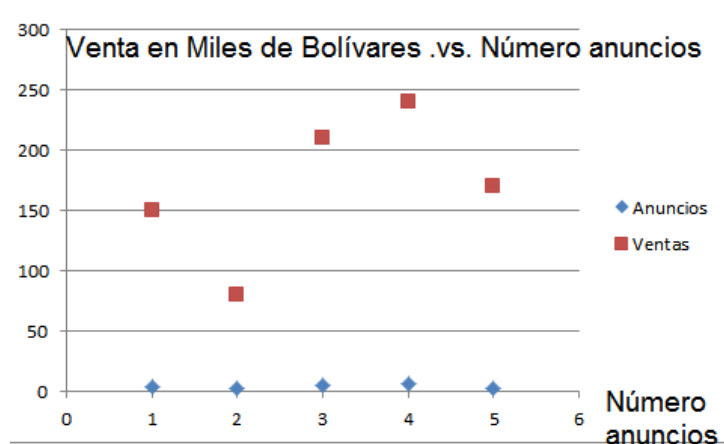
Ciudad	Anuncios	Ventas
Caracas	4	150
Maracay	2	80
Valencia	5	210
Barquisimeto	6	240
Maracaibo	3	170

Las ventas están en miles de Bolívars

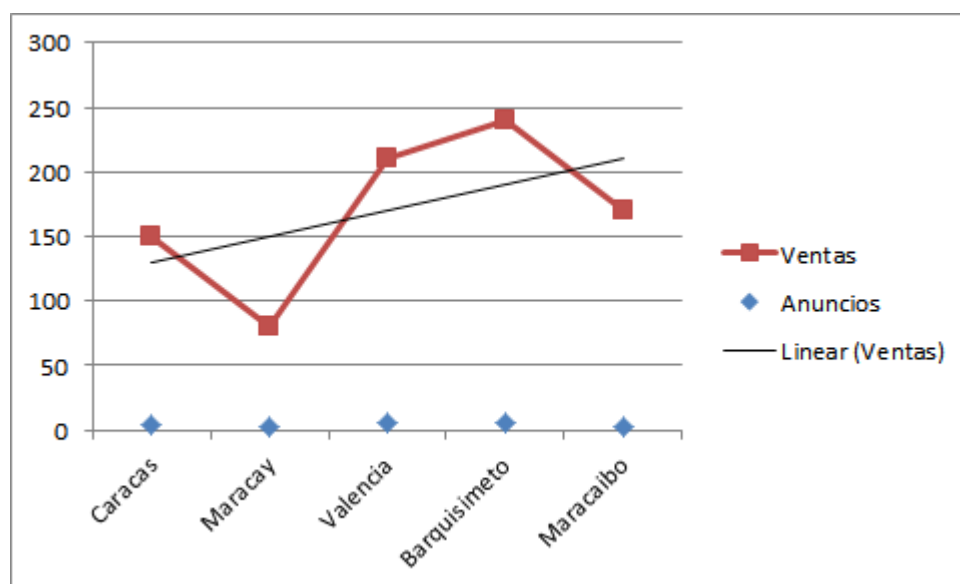
- a)Cuál es la variable dependiente y por qué

La escogencia de la variable dependiente depende de nuestro interés y otros factores. Dado que podemos controlar el número de anuncios que aparecen en televisión, escogemos este número como la variable independiente x : *número de anuncios*. Asumimos que las ventas en miles de Bolívars dependerán del número de anuncios y por lo tanto nuestra variable dependiente será: y : *ventas (en miles de Bolívars)*

b) Diagrama de dispersión



c)



Observando el gráfico superior se ve que los puntos no reflejan una situación de dependencia lineal perfecta, y que en algunos casos el número de anuncios ha aumentado y las ventas disminuido como es el caso de Caracas y Maracay y Barquisimeto y Maracaibo y también ha sucedido que se aumentaron los avisos entre Maracay y Valencia y las ventas aumentaron y lo mismo sucede con Valencia y Barquisimeto. Esta relación no corresponde a un comportamiento lineal perfecto. Esta relación la mide el coeficiente de correlación que se calculará más adelante. Este coeficiente da alto y cercano a 1 debido a que la nube de puntos es simétrica respecto a la recta de regresión que se presenta en el gráfico.

d) Cálculo del coeficiente de correlación

$$r = \frac{n(\sum_{i=1}^n x_i y_i) - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{\sqrt{[n(\sum_{i=1}^n x_i^2) - (\sum_{i=1}^n x_i)^2][n(\sum_{i=1}^n y_i^2) - (\sum_{i=1}^n y_i)^2]}}$$

Haremos los cálculos parciales en EXCEL

Ciudad	x_i	y_i	$x_i y_i$	x_i^2	y_i^2
Caracas	4	150	600	16	22500
Maracay	2	80	160	4	6400
Valencia	5	210	1050	25	44100
Barquisim	6	240	1440	36	57600
Maracaibo	3	170	510	9	28900
Sumas=	20	850	3760	90	159500

Luego:

$$r = \frac{5(3760) - (20)(850)}{\sqrt{[5(90) - (20)^2][5(159.500) - (850)^2]}} = 0,9295$$

Este coeficiente de correlación señala una correlación positiva fuerte. Esta medida da alta por que la recta de regresión que se muestra en el gráfico junto a la nube de puntos está promediándolos “adecuadamente” aun cuando como vemos tiene pendiente positiva indicando una relación creciente para la variable “y” respecto a la variable “x” aun cuando como vimos en las fluctuaciones que tiene a nube respecto a la recta esta relación no es siempre creciente ya que hubo decrecimiento entre Caracas y Maracay y entre Barquisimeto y Maracaibo.