

# **METODOLOGIA DEL ANALISIS DE CORRELACIÓN Y REGRESIÓN LINEAL SIMPLE APLICADOS AL CASO ACCIÓN DE ECOPETROL Y LOS PRECIOS INTERNACIONALES DEL PETRÓLEO**

**Jairo Luis Pimienta Quiroz. Programa de Administración de empresas de Universidad de Cartagena**

## **Resumen**

El propósito de este artículo es explicar cómo se realizan los análisis de correlación y regresión lineal por mínimos cuadrados simple utilizando el programa Excel para determinar y describir la relación lineal entre dos variables cuantitativas continuas e independientes. De esta manera los estudiantes y personas en general tendrán acceso a una herramienta útil que le permitirá establecer y cuantificar relaciones en un campo amplio de las ciencias en general y en especial en la administración. En particular esto se hará mediante la aplicación de las herramientas al caso precios internacionales del petróleo-precio acción Ecopetrol, con indicaciones paso a paso de cómo realizarlo en el programa Microsoft Excel.

## **Abstract**

This purpose of this article is to explain how to make a correlation and linear simple regression (using least squares) analysis using Excel to determinate and describe the linear relation between two independent quantitative continuous variables. This way the students and people in general will have access to a useful tool that will allow them to stablish and quantify relations on a wide range of the science in general and particularly in the administration. This will come to be through the application of the tools to the case of the crude international prices-Ecopetrol stock prices, with step by step indications about how to do it on the Microsoft Excel program.

**Palabras clave:** Coeficiente de correlación, Microsoft Excel, Precio de la acción de Ecopetrol, Precios del petróleo, Relación, Regresión lineal.

## INTRODUCCIÓN

Para los profesionales que se desenvuelven en áreas administrativas es de vital importancia la recopilación de datos acerca de fenómenos o variables, y de métodos o herramientas que le permitan el estudio y análisis de estos para soportar su proceso de toma de decisiones y desempeñarse mejor. Por eso los métodos estadísticos y matemáticos siempre están ligados a los procesos gerenciales y en especial a ésta toma de decisiones. Entre estos tenemos el análisis de correlación y regresión lineal, los cuales son comunes en diferentes áreas del conocimiento y son usualmente utilizados por las áreas relacionadas con la medicina y las ciencias económicas en general.

En Administración, estos métodos son a menudo utilizados para pronosticar la tendencia de ciertas variables (como las ventas) a medida que pasa el tiempo, pero, la aplicación podría ser más amplia dada su sencillez y utilidad. Y es por esto que el enfoque de este trabajo es mostrar cómo se realiza e interpretan los análisis de regresión lineal y correlación.

Estos dos métodos permiten establecer y describir la relación lineal entre dos variables, relación que si bien puede ser vista y representada mediante el trazo de una línea por parte del espectador, éste lo haría “al ojo”. El uso de métodos que permitan la cuantificación de la relación y que provean precisión para el trazo de la nueva línea de tendencia es de notable importancia para el área de estudio pues este resultado puede resultar útil para describir la tendencia de una variable a medida que la otra se comporta de cierta manera. En el documento se realiza la aplicación de ambos análisis en el caso precios internacionales del petróleo-precio de la acción de Ecopetrol, los cuales serán considerados como variables cuantitativas continuas e independientes. La fuerza y la dirección, en este caso proporcionan una idea de la dirección de la relación, así como de la proporcionalidad de la relación, es decir que tan cercano es su comportamiento. Esto se hace mediante el coeficiente de correlación de Pearson y seguidamente con el método de los mínimos cuadrados con el que se llega a la recta que mejor muestre la relación lineal entre estos, no solo aportando dirección si no también una pendiente y un intercepto que proporcionan información de la relación lineal. La herramienta que se utiliza para realizar los análisis será Excel dado que posee las herramientas integradas (Pearson y regresión por mínimos cuadrados) y es un software de fácil acceso, lo cual facilitaría su utilización por parte de los estudiantes y otras personas en general.

## Análisis de correlación

El coeficiente de correlación Pearson es una medida de la relación lineal entre dos variables. El coeficiente de Pearson muestra no solo dirección de la relación lineal, sino, también fuerza de esta, la cual se clasifica en rangos.

Puede tomar valores entre -1 y 1. Si el coeficiente es positivo se dice que si una de las variables aumenta la otra tiene tendencia a aumentar, y viceversa. Mientras que cuando el valor es negativo la relación sería inversa, pues al aumentar una de las variables la otra disminuiría. Cuando el coeficiente toma el valor de 0 se dice que no existe relación lineal entre ambas. Además, también se dice que mientras más cercanos sean los valores a -1 y 1 la relación es débil o fuerte respectivamente, siendo perfecta cuando toma alguno de los dos valores. Esto implica que tendrían una relación perfectamente lineal. Los rangos estarán expresados de manera rápida en la tabla 1:

TABLA 1			
INTERPRETACION DEL COEFICIENTE DE CORRELACIÓN			
Valor	Interpretation		
	Direccion	Rango	Fuerza
$0 \geq r \geq 1$	Positiva	1	Perfecta
		$1 < r \leq 0.8$	Muy fuerte
		$0,6 < r \leq 0.4$	Fuerte
		$0,2 < r \leq 0,4$	Moderada
		$0.2 < r \leq 1$	Debil
		$0 < r \leq 0.2$	Muy débil
$r=0$	No hay relación		
$0 \geq r \geq -1$	Negativa	$-0 > r \geq -0.2$	Muy débil
		$-0.2 > r \geq -0.4$	Débil
		$-0.4 > r \geq -0.6$	Moderada
		$-0.6 r \geq -0.8$	Fuerte
		$-0.8 > r > -1$	Muy fuerte
		$r = 1$	Perfecta

En el caso de los rangos que denotan la fuerza de la relación, los autores no usan los mismos rangos, pero concuerdan que cuando el coeficiente se acerca a 0.5 y a -0.5 la fuerza de la relación es moderada. En este caso se usaran los expresados en la tabla.

El coeficiente esta expresado matemáticamente en la ecuación 1 de la siguiente manera:

$$\text{Ecuacion 1: } r = \text{Corr}(X, Y) = \frac{\text{Cov}(x, y)}{Sx \cdot Sy} = \frac{\sum(x_i - \bar{x}) \cdot (y_i - \bar{y})}{(n - 1) \left( \left( \sqrt{\frac{\sum(x_i - \bar{x})^2}{n - 1}} \right) \cdot \left( \sqrt{\frac{\sum(y_i - \bar{y})^2}{n - 1}} \right) \right)}$$

Siendo  $r$  el coeficiente de correlación,  $\text{Cov}(x, y)$  la covarianza de  $x$  y  $y$ .

$(x_i - \bar{x})$  la diferencia entre cada valor que tomo  $x$  ( $x_i$ ) y la media aritmética de  $x$ .

$(y_i - \bar{y})$  la diferencia entre cada valor que tomo  $y$  ( $y_i$ ) y la media aritmética de  $y$ .

$Sx$  y  $Sy$  la desviación estándar de  $x$  y  $y$ .

$n$  la muestra.

El coeficiente de correlación es entonces el resultado de dividir la covarianza de X y Y entre el producto de las desviaciones de X y Y. Siendo la covarianza la suma del producto de todos los residuos para todos los  $x_i$  y  $y_i$  dividida entre  $n-1$ , es ésta en si una medida de variación.

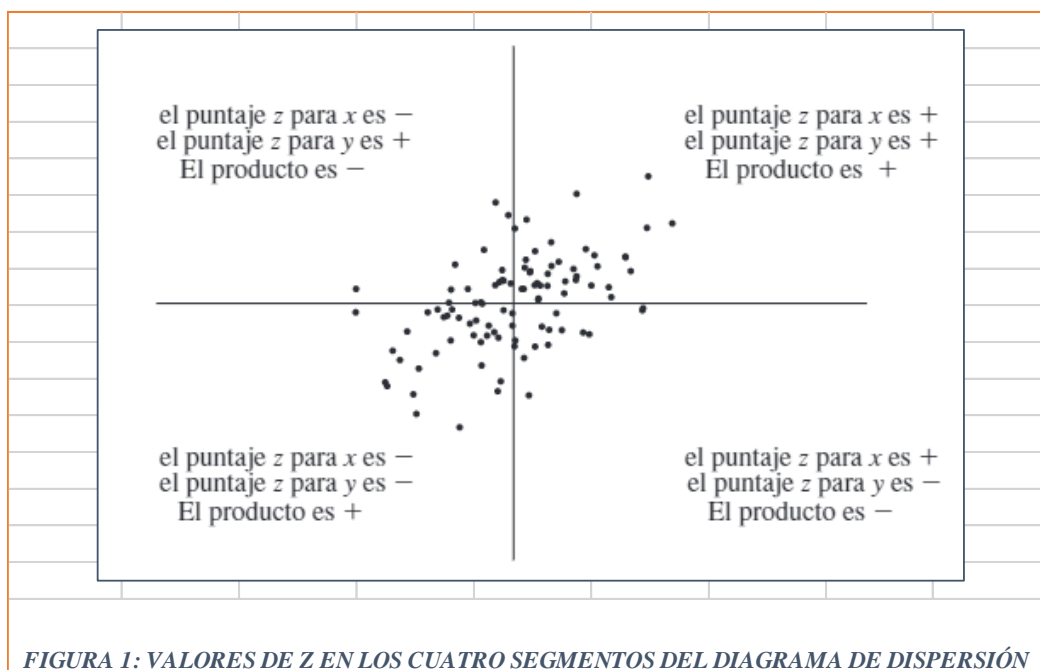
Usualmente el análisis de correlación se empieza con la observación de los datos en un diagrama de dispersión en el cual se muestran puntos, los cuales están compuestos por un valor de X y uno de Y proporcionados por las variables, teniendo así una idea preliminar de cómo se comportan o relacionan.

En este diagrama se trazan dos rectas que cortan (una totalmente horizontal y otra totalmente vertical) el diagrama en 4 cuartiles o segmentos, y pasan a través de un punto compuesto por las medias de ambas variables. Dependiendo de cómo se distribuyan los puntos en los 4 segmentos se puede conocer la relación lineal sin cuantificarla. Esto es por la manera en que se calcula el coeficiente. Dado que las varianzas siempre se elevan al cuadrado estas no cambiaran el signo del termino (coeficiente) por lo tanto, en este caso, el signo depende del dividendo, el cual está

determinado por el producto de los residuos (valor de la observación menos la media de esa variable) de ambas variables.

En la figura 1 (Navidi, 2006, pág. 510) se observa como  $z$  ( $z = \{x_i - \bar{x}\} / S_x$ ) toma valores que pueden ser positivos o negativos, los cuales “arrastran” al coeficiente hacia en cierta dirección. En los casos de los residuos, al multiplicarse, si ambos son negativos o positivos tomaran un valor positivo y contribuirán a una relación positiva, mientras que si tienen distinto signo, el producto será negativo y la relación será negativa. Si los puntos se concentran en mayoría en los segmentos inferior izquierdo y superior derecho se sabe que a medida que uno aumenta el otro también aumenta, mientras que si se concentran en los segmentos opuestos (superior izquierdo e inferior derecho) la relación sería negativa. Cuando los puntos se esparcen por todo el diagrama o se aglomeran en forma circular alrededor del centro la relación, tienden a ser 0 por lo cual tiende a no tener relación lineal y según la interpretación del coeficiente de Pearson, a no tener relación o a relaciones muy débiles.

Luego de la vista preliminar en el diagrama de dispersión se procede a calcular el coeficiente. Este como se expresó matemáticamente anteriormente es igual a la covarianza de las dos variables dividida entre las desviaciones estándar de ambas variables, siendo la covarianza a su vez la sumatoria de los productos de cada pareja de residuos para cada  $x_i$  y  $y_i$  divididos entre  $n-1$ . Luego de hallar el coeficiente se procede a interpretarlo según fue dispuesto en la tabla 1.



## **Análisis de regresión por mínimos cuadrados**

*“Determina una ecuación de regresión al minimizar la suma de los cuadrados de las distancias verticales entre los valores reales de Y y los que se pronosticaron” (Lind A, 2015).*

El método de los mínimos cuadrados es una técnica que utilizando datos de dos variables relacionadas y estudiadas en parejas de datos, expresa la relación lineal entre éstas, tomando ambas variables (una como independiente y la otra como dependiente) para trazar la línea o recta de tendencia, que mejor represente la relación lineal de ambas variables y que minimice la suma total de los residuos.

La línea de tendencia puede trazarse a criterio de la persona que desea conocer la relación lineal, pero el uso de un método que represente la relación lineal de la manera más aproximada hace ideal su utilización. Esto se logra a través de la ecuación final de los mínimos cuadrados para la regresión, que es una función lineal simple que traza una línea recta. Esta consta de un intercepto, una pendiente y la variable de la siguiente forma:

*Ecuacion 2:* 
$$\check{Y} = a + bx$$

*Donde:*

$\check{Y}$  (Y prima) es el valor estimado de Y cuando X toma cierto valor.

$a$  es el intercepto

$b$  es la Pendiente

$x$  la variable independiente

Para hallar el valor del intercepto y la pendiente se utilizan las siguientes ecuaciones:

$$b = r * \left( \frac{S_y}{S_x} \right)$$

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

Los valores de a y b que resulten serán respectivamente el intercepto, el valor que toma  $\check{Y}$  cuando x es 0, y, b el valor de la pendiente según el cual  $\check{Y}$  aumentará por cada punto en la variable x. Es

por esto que se considera que la regresión lineal ofrece más información que el análisis de correlación, pues esta relación se plasma en datos claros de como aumenta o disminuye la variable independiente al tomar la dependiente cierto valor. Con los valores de  $a$  y  $b$  se puede proceder a remplazar para obtener el valor “pronosticado” o al que tendería la variable dependiente al tomar la independiente cierto valor.

### **Delimitación de los datos**

En este caso de estudio las variables a analizar serán los precios SPOT del petróleo en sus referencias WTI, Brent y la canasta OPEP, y, los precios de negociación de la acción de Ecopetrol. El periodo sobre el cual se realizara la correlación y la regresión será del 2014 al 2015 en periodos mensuales.

Se utilizaran los precios SPOT en cuanto al crudo para el WTI y el BRENT, pues estos reflejan mejor la realidad de la demanda actual (o de ese momento) de éste, ya que los contratos a futuro que se negocian cambian de portador usualmente varias veces dado que son renegociados antes de la entrega del producto. En cuanto a los precios de las acciones, se promediaran los datos históricos de los precios de cierre de las acciones de manera mensual, creando así nuestra serie de tiempo con datos mensuales. Se debe tener en cuenta que para promediar se debe hacer en base al número de observaciones mas no a la cantidad de días del mes o a 30 días, pues no todos los días son tranzadas.

La información acerca de los datos históricos que conforman las series de tiempo de cada variable están disponibles y libres al público en las siguientes páginas: En la página oficial de la Independent Statistics and Analysis U.S. Energy Information Administration (Administración de información de estadísticas y análisis de energía de los Estados Unidos) se encuentran los precios históricos del barril de petróleo WTI y Brent, y en la página oficial de la OPEP el precio de la canasta OPEP (que es un promedio ponderado de la producción del petróleo producido por los países pertenecientes a la OPEP).

En cuanto los datos de la acción de Ecopetrol, están disponibles en la página de la Bolsa de Valores de Colombia, y, éstos y los precios del petróleo están dispuestos en tablas que se encuentran en los anexos del documento.

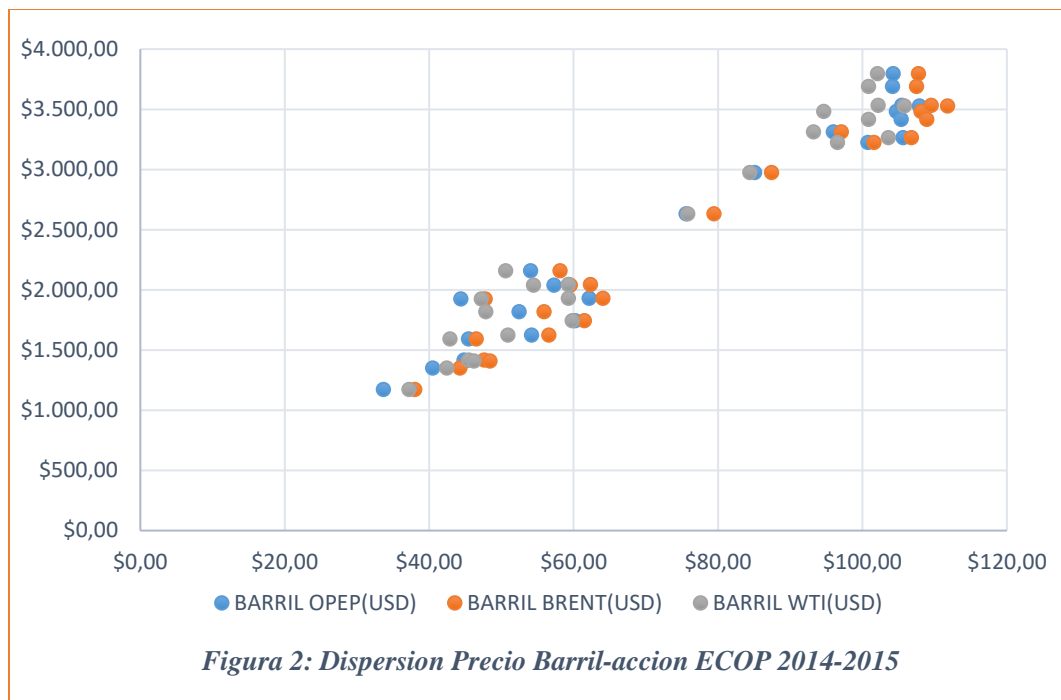
Los precios de la acción de Ecopetrol y los precios internacionales del petróleo se tabularan para facilitar su análisis y posteriormente se llevara a cabo la síntesis de la correlación de las dos variables, entendiendo como correlación la correlación lineal medida mediante el coeficiente de correlación de Pearson, para luego utilizar la regresión lineal por medio de mínimos cuadrados para hallar el valor esperado de la acción (siendo esta nuestra variable dependiente) cuando el precio del petróleo toma cierto valor.

Posteriormente se realizara la interpretación de los resultados, y, las conclusiones y observaciones. El tratamiento de los datos se realizara con Microsoft Excel dado que este es un software de alta difusión y por su simplicidad, y que haría más sencillo el proceso de tratamiento y tabulación de los datos, así como las operaciones (las cuales, dado que serán numerosas puede hacer tedioso el procedimiento).

### **Aplicación de los modelos**

Como se mencionó los modelos serán aplicados con Excel dada la facilidad que proporciona, ya que tiene las formulas y ayuda incorporadas para calcular de manera instantánea, tanto el coeficiente de correlación, como la recta de tendencia por mínimos cuadrados. En primera instancia ya teniendo tabulados los datos en parejas para el estudio de caso, se procede a realizar el diagrama de dispersión, el cual muestra las puntos dados por cada pareja de datos mostrando la relación de estos a medida que aumentan y disminuyen. En la figura 2 se encuentra los datos de las tres referencias de petróleo en el eje X y el precio de la acción de Ecopetrol en el eje Y.





En éste los valores de Y serán siempre los mismos para cada pareja de datos por lo cual los puntos de las tres parejas tendrán la misma altura y variaran en el eje X (el análisis del diagrama de dispersión realizara en los resultados, de manera separada pues las medias, referencia el análisis son distintas en el caso de la variable X, y son necesarias para el análisis). Para el análisis preliminar se toman por separado los datos en tres diagramas distintos, trazando dos rectas en cada uno que corten vertical y horizontalmente el diagrama a través de las medias de x y y formando los segmentos para observar cómo se distribuyen los puntos en estos segmentos.

Luego se procede a hallar los términos necesarios para resolver la ecuación y calcular el coeficiente. Para esto necesita encontrar los valores de las medias para las dos variables. Esto dividiendo el total de los valores de cada observación (para cada variable) entre el total de las observaciones ( $\sum x_i/n$  en el caso de la variable x). En el caso de la referencia WTI la media sería 70,97 y para la acción de Ecopetrol 2461,92. Posteriormente se buscan los residuos para cada valor de  $x_i$  y  $y_i$  siendo para  $x_1$  23,65 y para  $y_1$  1.020,93. Los datos y las tablas con los cálculos realizados se encuentran en los anexos para realizar la consulta.

Después encontrar los valores para todos los  $x_i$  y  $y_i$  se procede a multiplicar cada pareja de residuos ( $\{x_i - \bar{x}\} \cdot \{y_i - \bar{y}\}$ ) siendo el resultado de los residuos para la primera pareja de datos 24.140,81. Seguidamente después de hallar todos los productos de los residuos se realiza la

sumatoria de estos y se divide entre  $n-1$ . Esto resultara en un valor de 21.852,04 siendo este la covarianza.

Teniendo la covarianza se procede a hallar las desviaciones estándar para X y Y. Esto elevando al cuadrado cada residuo de una variable y posteriormente haciendo la sumatoria de todos los resultados dividiéndolo entre  $n-1$ . El primer residuo al cuadrado para X sería 559,125434 y la sumatoria de todos los residuos de x al cuadrado sería 14507,22698. De esta manera al dividir entre  $n-1$  se tendría una varianza de  $630,7489993 \approx 630,75$  y al aplicar la raíz cuadrada se obtendría la desviación, que toma un valor de 25,11. Conociendo el valor de la covarianza de X y Y, y las desviaciones de ambas variables se puede hallar el coeficiente de correlación.

El procedimiento descrito actualmente dadas las numerosas operaciones puede resultar tedioso por lo cual es recomendable utilizar las herramientas incorporadas en el programa Excel.

El coeficiente de correlación de Pearson puede ser accesado en Excel 2013 desde el menú predeterminado de fórmulas por medio el menú secundario **INSERTAR FUNCIÓN**.

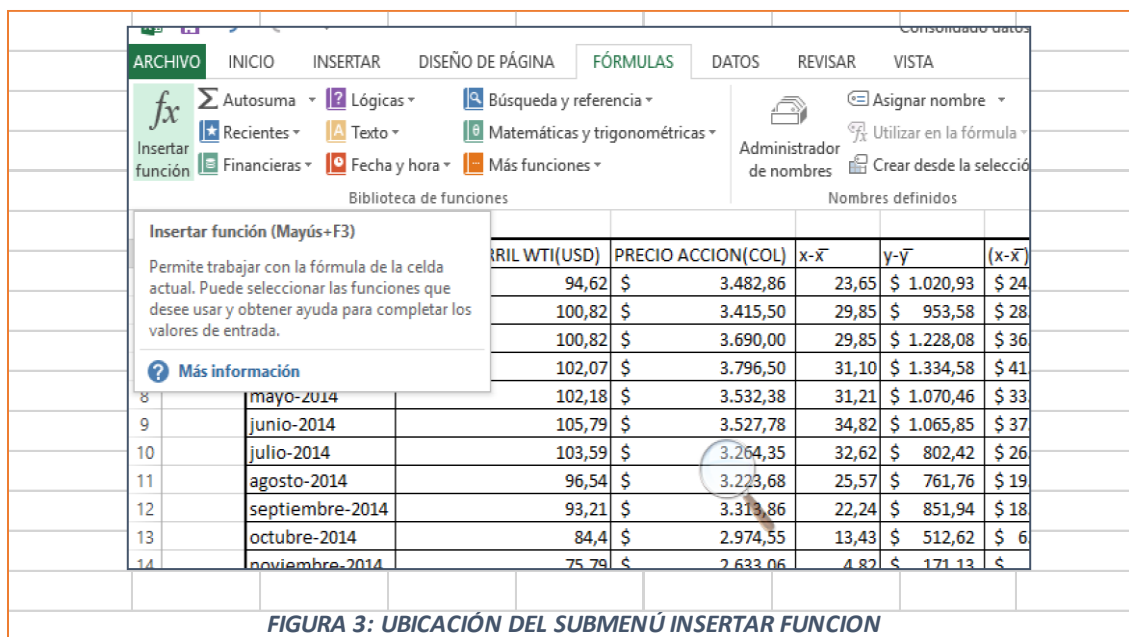


FIGURA 3: UBICACIÓN DEL SUBMENÚ INSERTAR FUNCIÓN

Luego en la ventana emergente se realizara la búsqueda de la formula escribiendo Pearson o correlación. La fórmula aparecerá como **PEARSON**. Se arrastra el cursor hasta allí, se presiona clic izquierdo y se presiona aceptar seguidamente. Después se seleccionan los valores, los cuales son todos los valores de ambas variables (figura 5).

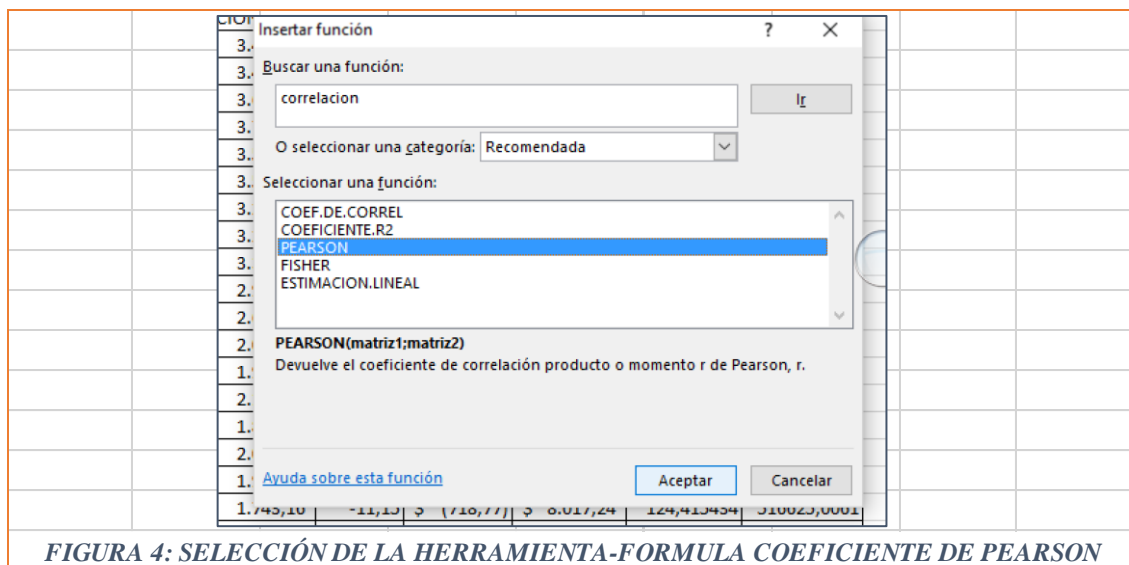


FIGURA 4: SELECCIÓN DE LA HERRAMIENTA-FORMULA COEFICIENTE DE PEARSON

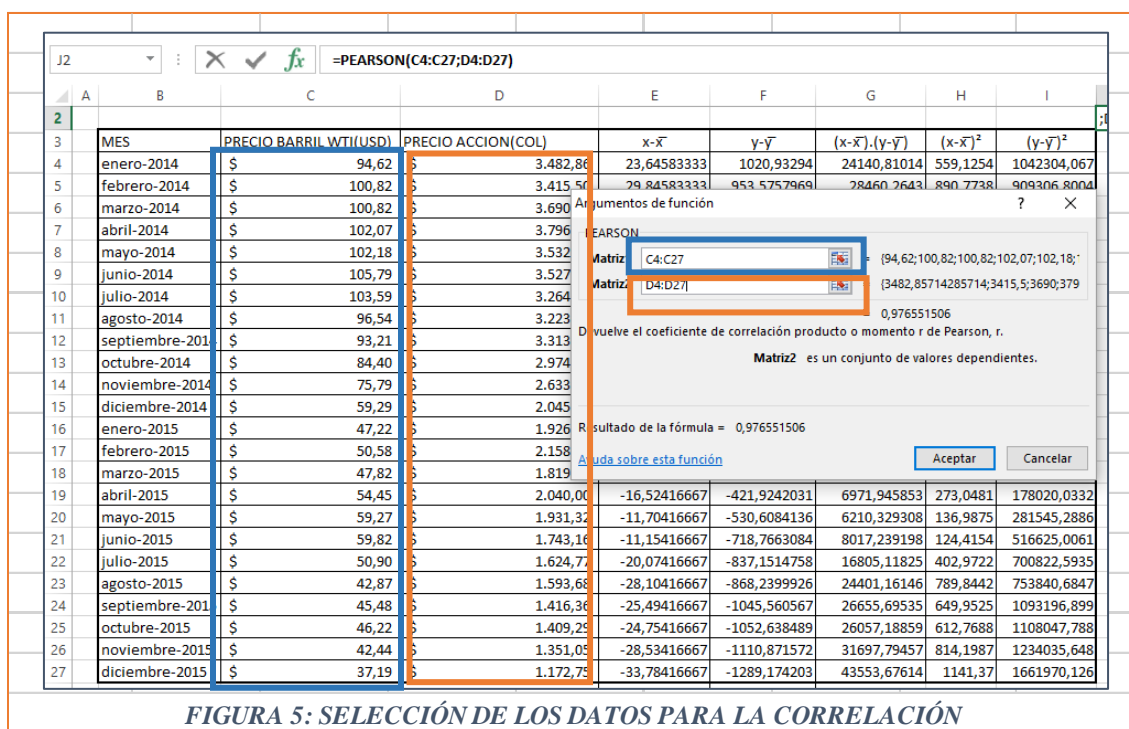
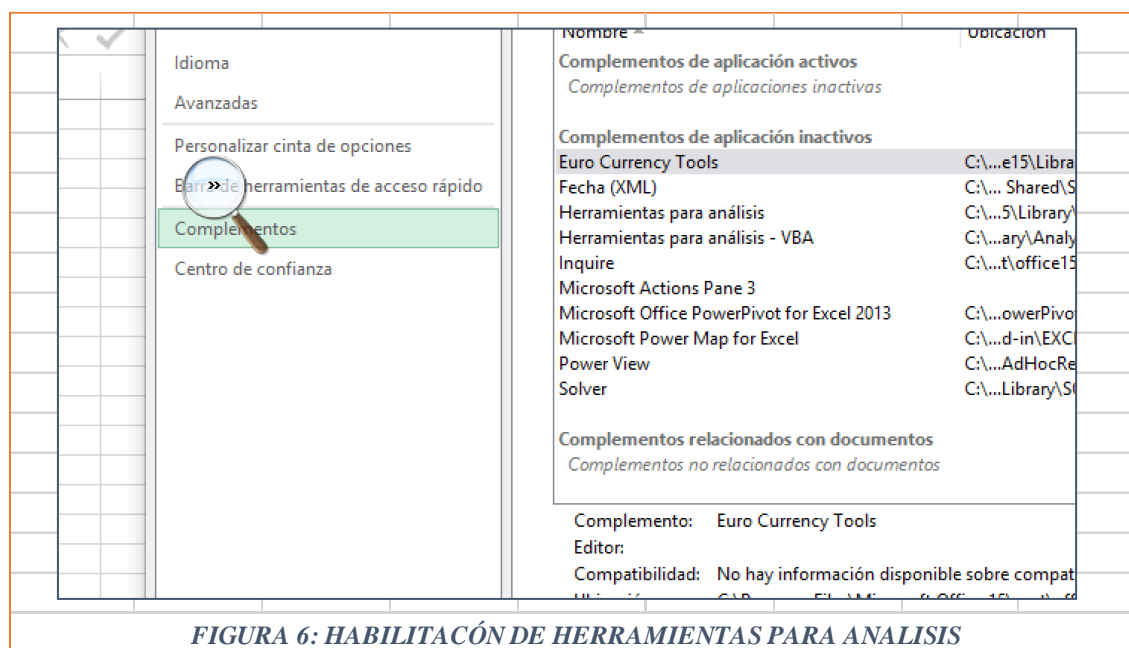


FIGURA 5: SELECCIÓN DE LOS DATOS PARA LA CORRELACIÓN

En el caso de los mínimos cuadrados, se puede utilizar la información obtenida en el análisis de correlación para encontrar el intercepto y la pendiente, pues ya se tienen los valores de las desviaciones y de la covarianza para hallar la pendiente, y, con la pendiente y las medias necesarias para hallar el intercepto (utilizando las formulas).

En cuanto a la utilización de la herramienta en Excel para el análisis de regresión, se tiene que tener en cuenta que en Excel no siempre está dispuesta en el menú principal. Para acceder a esta se inicia yendo al menú **datos** en el cual se dará clic derecho y luego en personalizar cinta de opciones. En la ventana emergente se presiona primero **complementos** (figura 6) y luego se buscara en el lado izquierdo **Herramientas para análisis** para señalarla. Posteriormente se presiona en la casilla **ir** que está en la parte baja de la ventana y en la ventana emergente se selecciona **herramientas para análisis** y se le da aceptar.



**FIGURA 6: HABILITACIÓN DE HERRAMIENTAS PARA ANALISIS**

Después de esto, el submenú **Análisis de datos** debería aparecer en el menú **Datos**. Al ingresar a análisis de datos, en la ventana emergente se selecciona la opción **Regresión** (figura 7) y se presiona aceptar. En la nueva ventana se seleccionan en las casillas de rangos de Y y de X (figura 8), para Y los datos de la acción de Ecopetrol, que será la variable dependiente, y para x el precio del petróleo, y se señala la casilla de curva de regresión ajustada. Por último se elige la celda de preferencia para la salida de los datos.

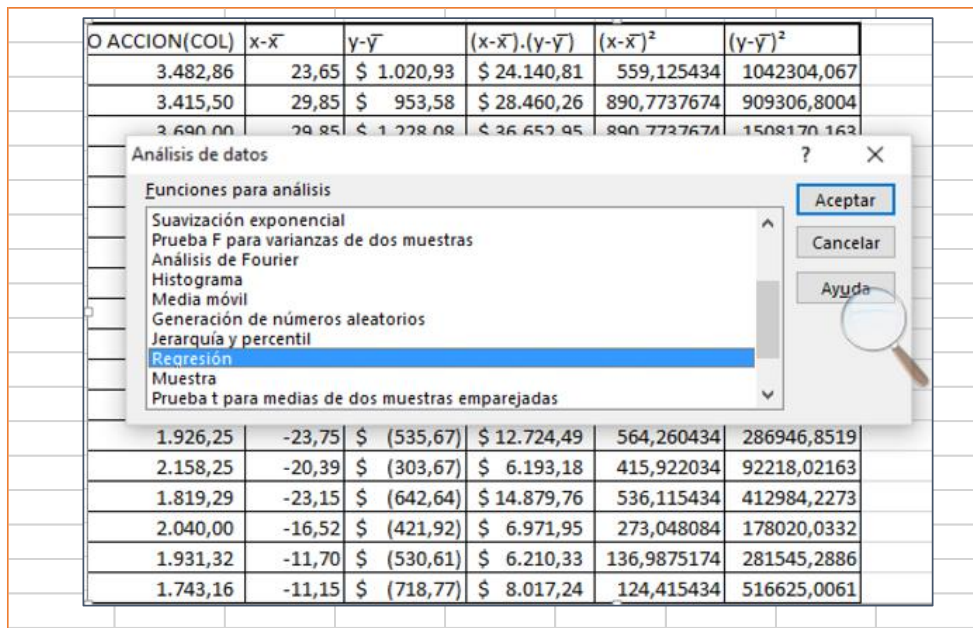


FIGURA 7: SELECCIÓN DE LA HERRAMIENTA DE ANALISIS DE REGRESION

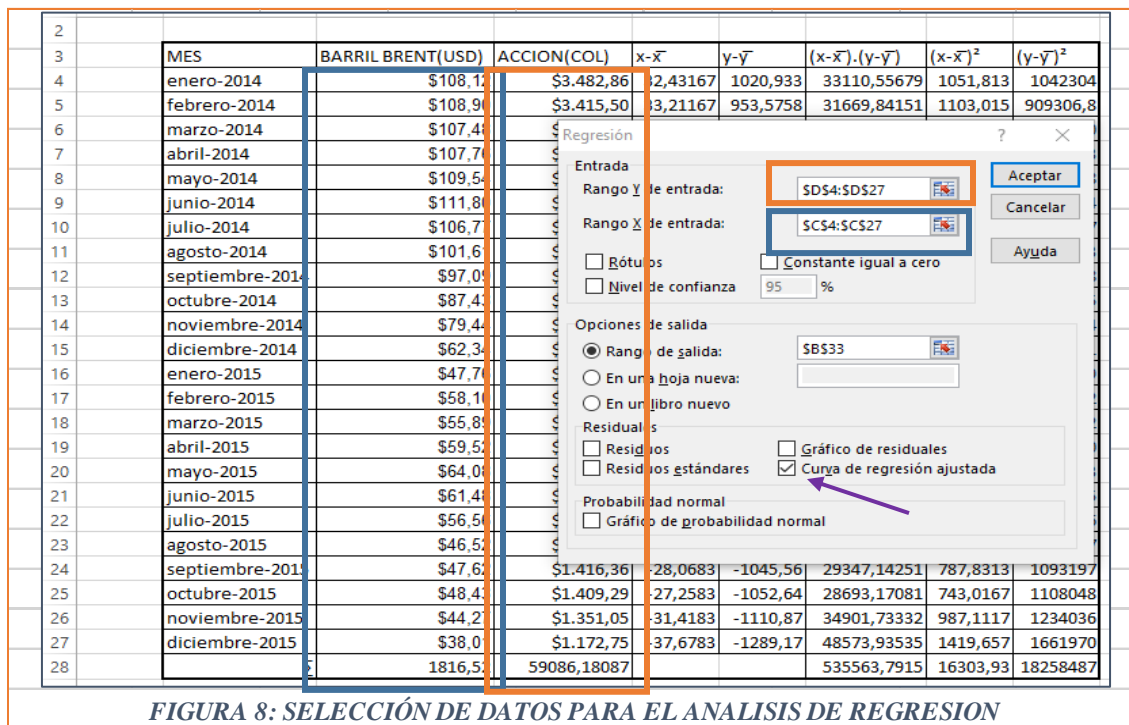
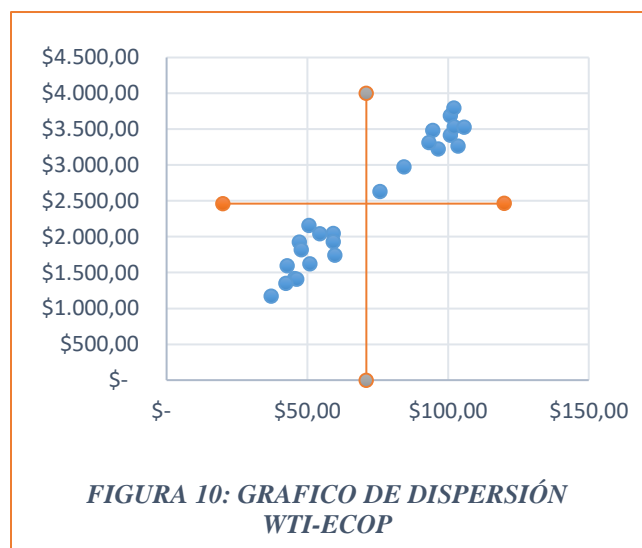
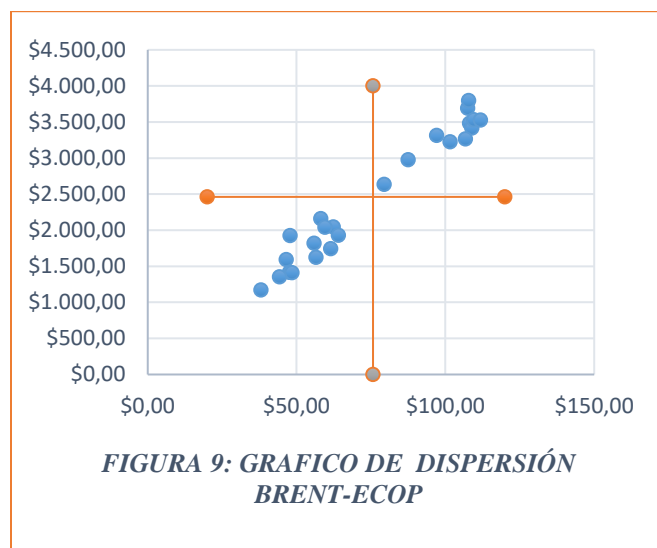
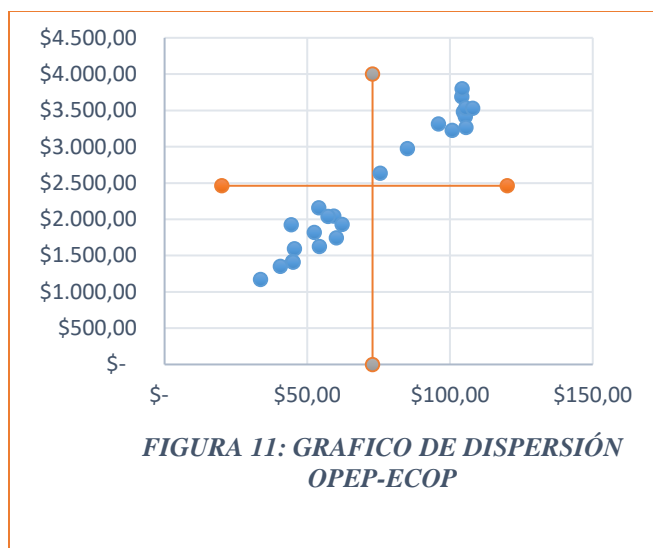


FIGURA 8: SELECCIÓN DE DATOS PARA EL ANALISIS DE REGRESION

## Resultados

En los diagramas de dispersión se tuvieron resultados similares. En el caso del diagrama de dispersión WTI-ECOP en la figura 10 se puede observar una relación lineal positiva, pues los puntos se encuentran concentrados en el segmento inferior izquierdo y superior derecho. Esto a la vez indica que en las observaciones realizadas, los valores de una variable tienden a aumentar a medida que los de la otra aumentan. En cuanto la fuerza se prevé una relación fuerte o muy fuerte dada la cercanía de los puntos. Para el WTI-ECOP, BRENT-ECOP y el OPEP-ECOP (figuras 10, 9 y 11 respectivamente) la interpretación es la misma tanto para la dirección lineal y fuerza de la relación. De esto también se deduce que los coeficientes de correlación son positivos (los tres), pues los productos de los residuos fueron todos positivos al encontrarse todos los datos en los segmentos inferior izquierdo y superior derecho (los productos de residuos con signo negativo estarían en los cuadrantes superior izquierdo e inferior derecho).





El comportamiento de las variables fue bastante parecido lo cual puede ser observado en las líneas de tendencia y los diagramas de dispersión por lo cual se esperaba preliminarmente resultados altos en cuanto a la fuerza y una relación positiva.

Los coeficientes fueron para WTI-ECOP, BRENT-ECOP y OPEP-ECOP, 0.9766, 0.9816 y 0.979 en ese orden. Todos estos demuestran una relación positiva y están en el rango de la relación muy fuerte (0.8-1). Esto implica que a medida de que los precios del petróleo en sus tres referencias aumentan, la tendencia de la acción de Ecopetrol es a aumentar y el hecho de que las relaciones lineales sean muy fuertes implica que las variables tendrán comportamientos lineales parecidos y variaciones muy similares (proporcionalmente). No solo esto, sino que también por su cercanía a 1 se puede decir que su relación lineal es casi perfecta según el coeficiente de Pearson.

En cuanto a la aplicación del análisis de regresión lineal por mínimos cuadrados, este arrojó la siguiente función lineal de Y (acción de Ecopetrol) y X (Precio Brent, dado que este tuvo el coeficiente de correlación más fuerte). En este caso Excel arroja los siguientes datos:

Resumen		
<i>Estadísticas de la regresión</i>		
Coeficiente de correlación múltiple	0,981595651	
Coeficiente de determinación R^2	0,963530023	
R^2 ajustado	0,961872297	
Error típico	173,9757805	
Observaciones	24	
<b>ANÁLISIS DE VARIANZA</b>		
	<i>Grados de libertad</i>	<i>Suma de cuadrados</i>
Regresión	1	1759
Residuos	22	66588
Total	23	18258
	<i>Coefficientes</i>	<i>Error típico</i>
Intercepción	-24,34278785	109,07
Variable X 1	32,84874804	1,8625

FIGURA 12: SALIDA DE DATOS DE ANALISIS DE REGRESION

Esto implica que la función lineal de la recta de regresión que muestra la tendencia de Y según los valores de X es la siguiente:

$$\check{Y} = a + bx = -24.34 + 32.85x$$

De esto se entiende que a, el intercepto, denota el valor que toma  $\check{Y}$  cuando x es 0. En cuanto a b, la pendiente, esta indica en cuanto aumentara  $\check{Y}$  por cada punto que aumente x. De esta manera dando valores a x se obtiene el valor esperado, o al que tiende la acción de Ecopetrol. Siendo, por ejemplo para un x = \$40,  $\check{Y} = \$1289.6$ . La siguiente es la recta de regresión y los datos de  $\check{Y}$ , es decir, los de Y después del ajuste de la regresión junto con una tabla que contiene los nuevos pares con los valores ajustados de y para los valores originales de x.



Observación	Pronóstico para Y	Residuos
1	3527,26385	-44,406707
2	3552,885873	-137,385873
3	3506,240651	183,759349
4	3515,438301	281,061699
5	3573,909072	-41,5281197
6	3648,147243	-120,369465
7	3482,91804	-218,570214
8	3313,4185	-89,7342896
9	3164,942159	148,921477
10	2847,623253	126,922202
11	2585,161756	47,8937994
12	2023,448165	22,0518353
13	1544,513418	381,736582
14	1884,169473	274,080527
15	1811,57374	7,71197438
16	1930,814695	109,185305
17	2080,604986	-149,289197
18	1995,198241	-252,040347
19	1833,582401	-208,809674
20	1503,780971	89,9032397
21	1539,914594	-123,550957
22	1566,52208	-157,236365
23	1429,871288	-78,8186561
24	1224,238125	-51,488125

**FIGURA 13: VALORES DE PRONOSTICADOS DE Y PARA CADA OBSERVACIÓN**



En la recta están representados los puntos con los valores ajustados de Y, los cuales pueden ser observados en la tabla. La nueva recta es la recta que mejor se ajusta a para representar la tendencia lineal de ambas variables.

## Discusión

Los análisis de correlación y regresión por mínimos cuadrados ordinarios son herramientas sencillas y representan una forma rápida de examinar la relación entre dos variables, obteniendo así una descripción de la relación entre estas, y así, de la dirección y la fuerza entre ambas. Esto hace que sean útiles en más de uno los ámbitos del administrador, ya que le permite aplicarlos (aunque con limitaciones) en una amplia gama de situaciones, desde pronósticos de ventas hasta la correlación vista en este trabajo.

En el caso particular de los precios de la acción de Ecopetrol y los precios internacionales del petróleo, los análisis aplicados revelan una relación positiva fuerte entre las variables, siendo el precio de la referencia Brent el que guarda la relación más estrecha con el valor de la acción de Ecopetrol, por lo cual la tendencia de una será parecida en proporción en cuanto a variaciones de la otra. La ecuación de mínimos cuadrados revela que por cada punto que  $x$  habrá una variación de 32.85 en la acción de Ecopetrol.

Esta información es relevante para describir o predecir tendencias en cuanto a variaciones futuras, entendiendo que el modelo tiene limitaciones. En el caso de las series de tiempo, para estimar comportamientos futuros utilizando el tiempo en sí como variable independiente, es a menudo erróneo asumir que siempre habrá aumentos, por ejemplo, en el caso de la producción o de las ventas. Las series de tiempo a menudo siguen tendencias cíclicas y siempre estarán presentes variables y condiciones exógenas que no son tenidas en cuenta por el método de los mínimos cuadrados. Entonces es de gran importancia entender que las limitaciones en el modelo pueden resultar en predicciones erróneas, por ejemplo en casos que las variables muestren relaciones no lineales (exponenciales, logarítmicas, etc.) o en caso de extender los modelos a datos que no se encuentren en el rango de los datos estudiados, pues no es seguro que las variables se comporten de la misma manera fuera del rango de estudio. También hay que entender que así como los valores de  $\hat{Y}$  de línea de tendencia difieren de los valores reales que tomo  $Y$  para los valores de  $X$  observados, la realidad es que no deja de ser una estimación y que existe un margen de error. En este caso se hace una aproximación a este margen de error mediante el error estándar o “error típico” como está dispuesto en la salida de datos de Excel para el análisis de regresión (figura 12).

En cuanto al caso de estudio la aplicación del trabajo es clara, por ejemplo, para inversionistas potenciales, corredores de bolsa o accionistas de Ecopetrol, pues ante la expectativa de una subida

en los precios del petróleo, el aumento de precio de la acción de podría ser lo suficientemente relevante como para ser tenido en cuenta, ya sea modificando su portafolio de inversión, obteniendo acciones para esperar el “revote” de una acción a precios bajos aguardando a que el precio aumente para vender generando mayor rentabilidad o simplemente esperando mayores dividendos si estaba interesado en vender. Los resultados de la correlación y regresión serian significativos para su toma de decisiones y su desempeño financiero. Todo esto sin olvidar las limitaciones ya mencionadas.

## **Bibliografía**

Independent Statistics and analysis us energy information administration, página oficial portal Petroleum and other liquids. Ver: [http://www.eia.gov/dnav/pet/pet\\_pri\\_spt\\_s1\\_d.htm](http://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm)

Bolsa de Valores de Colombia, página oficial. Ver <https://www.bvc.com.co/pps/tibco/portalbvc>

Gujarati, D. (2009). *Econometria*.

Lind A, D. (2015). *Estadística aplicada a los negocios y la economía*.

Navidi, W. (2006). *Estadística para ingenieros y científicos*.

Organization of the petroleum exporting countries, portal OPEP BASKET PRICE. Ver: [http://www.opec.org/opec\\_web/en/data\\_graphs/40.htm](http://www.opec.org/opec_web/en/data_graphs/40.htm)

## Anexos

MES	PRECIO BARRIL WTI(USD)	PRECIO ECOP(COL)	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x}) \cdot (y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$
enero-2014	\$94,62	\$3.482,86	23,6458333	1020,93294	24140,81014	559,125434	1042304,07
febrero-2014	\$100,82	\$3.415,50	29,8458333	953,575797	28460,2643	890,7737674	909306,8
marzo-2014	\$100,82	\$3.690,00	29,8458333	1228,0758	36652,94555	890,7737674	1508170,16
abril-2014	\$102,07	\$3.796,50	31,0958333	1334,5758	41499,74655	966,9508507	1781092,56
mayo-2014	\$102,18	\$3.532,38	31,2058333	1070,45675	33404,49491	973,804034	1145877,65
junio-2014	\$105,79	\$3.527,78	34,8158333	1065,85357	37108,58041	1212,142251	1136043,84
julio-2014	\$103,59	\$3.264,35	32,6158333	802,423623	26171,71515	1063,792584	643883,671
agosto-2014	\$96,54	\$3.223,68	25,5658333	761,760007	19475,02939	653,611834	580278,309
septiembre-2014	\$93,21	\$3.313,86	22,2358333	851,939433	18943,58325	494,432284	725800,798
octubre-2014	\$84,40	\$2.974,55	13,4258333	512,621251	6882,367485	180,2530007	262780,547
noviembre-2014	\$75,79	\$2.633,06	4,81583333	171,131352	824,1400715	23,19225069	29285,9398
diciembre-2014	\$59,29	\$2.045,50	-11,6841667	-416,424203	4865,569793	136,5197507	173409,117
enero-2015	\$47,22	\$1.926,25	-23,7541667	-535,674203	12724,4943	564,260434	286946,852
febrero-2015	\$50,58	\$2.158,25	-20,3941667	-303,674203	6193,182311	415,922034	92218,0216
marzo-2015	\$47,82	\$1.819,29	-23,1541667	-642,638489	14879,75868	536,115434	412984,227
abril-2015	\$54,45	\$2.040,00	-16,5241667	-421,924203	6971,945853	273,048084	178020,033
mayo-2015	\$59,27	\$1.931,32	-11,7041667	-530,608414	6210,329308	136,9875174	281545,289
junio-2015	\$59,82	\$1.743,16	-11,1541667	-718,766308	8017,239198	124,415434	516625,006
julio-2015	\$50,90	\$1.624,77	-20,0741667	-837,151476	16805,11825	402,9721674	700822,593
agosto-2015	\$42,87	\$1.593,68	-28,1041667	-868,239993	24401,16146	789,844184	753840,685
septiembre-2015	\$45,48	\$1.416,36	-25,4941667	-1045,56057	26655,69535	649,952534	1093196,9
octubre-2015	\$46,22	\$1.409,29	-24,7541667	-1052,63849	26057,18859	612,7687674	1108047,79
noviembre-2015	\$42,44	\$1.351,05	-28,5341667	-1110,87157	31697,79457	814,1986674	1234035,65
diciembre-2015	\$37,19	\$1.172,75	-33,7841667	-1289,1742	43553,67614	1141,369917	1661970,13
$\Sigma$	1703,38	59086,18087			\$ 502.596,83	14507,22698	18258486,6
Medias	70,97	2461,92		Covarianza	\$ 21.852,04	630,7489993	793847,245
					Desviaciones	25,11471679	890,981058
					Correlación	0,9766	

En la tabla están dispuestos los datos de los precios históricos de la acción de Ecopetrol provenientes del portal acciones de la página de la Bolsa de valores de Colombia en el link <https://www.bvc.com.co/pps/tibco/portalbvc/Home/Mercados/enlinea/acciones?action=dummy> y de los precios SPOT del barril de petróleo referencia WTI encontrados en el portal Precios SPOT en la página de la Administración de información de energía de los Estados unidos en el link [http://www.eia.gov/dnav/pet/pet\\_pri\\_spt\\_s1\\_d.htm](http://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm), así como los datos necesarios para el cálculo del coeficiente de correlación.

MES	BARRIL BRENT(USD)	PRECIO ECOP(COL)	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x}) \cdot (y - \bar{y})$	$(x - \bar{x})^2$	$(y - \bar{y})^2$
enero-2014	\$108,12	\$3.482,86	32,43166667	1020,93294	33110,55679	1051,813	1042304,1
febrero-2014	\$108,90	\$3.415,50	33,21166667	953,5757969	31669,84151	1103,015	909306,8
marzo-2014	\$107,48	\$3.690,00	31,79166667	1228,075797	39042,57638	1010,71	1508170,2
abril-2014	\$107,76	\$3.796,50	32,07166667	1334,575797	42802,0701	1028,592	1781092,6
mayo-2014	\$109,54	\$3.532,38	33,85166667	1070,456749	36236,74506	1145,935	1145877,7
junio-2014	\$111,80	\$3.527,78	36,11166667	1065,853575	38489,749	1304,052	1136043,8
julio-2014	\$106,77	\$3.264,35	31,08166667	802,423623	24940,66357	966,07	643883,67
agosto-2014	\$101,61	\$3.223,68	25,92166667	761,7600074	19746,08899	671,9328	580278,31
septiembre-2014	\$97,09	\$3.313,86	21,40166667	851,9394333	18232,92377	458,0313	725800,8
octubre-2014	\$87,43	\$2.974,55	11,74166667	512,6212514	6019,027861	137,8667	262780,55
noviembre-2014	\$79,44	\$2.633,06	3,751666667	171,1313524	642,0277906	14,075	29285,94
diciembre-2014	\$62,34	\$2.045,50	-13,34833333	-416,4242031	5558,569071	178,178	173409,12
enero-2015	\$47,76	\$1.926,25	-27,92833333	-535,6742031	14960,4877	779,9918	286946,85
febrero-2015	\$58,10	\$2.158,25	-17,58833333	-303,6742031	5341,123109	309,3495	92218,022
marzo-2015	\$55,89	\$1.819,29	-19,79833333	-642,6384888	12723,17101	391,974	412984,23
abril-2015	\$59,52	\$2.040,00	-16,16833333	-421,9242031	6821,811157	261,415	178020,03
mayo-2015	\$64,08	\$1.931,32	-11,60833333	-530,6084136	6159,479335	134,7534	281545,29
junio-2015	\$61,48	\$1.743,16	-14,20833333	-718,7663084	10212,4713	201,8767	516625,01
julio-2015	\$56,56	\$1.624,77	-19,12833333	-837,1514758	16013,31248	365,8931	700822,59
agosto-2015	\$46,52	\$1.593,68	-29,16833333	-868,2399926	25325,11352	850,7917	753840,68
septiembre-2015	\$47,62	\$1.416,36	-28,06833333	-1045,560567	29347,14251	787,8313	1093196,9
octubre-2015	\$48,43	\$1.409,29	-27,25833333	-1052,638489	28693,17081	743,0167	1108047,8
noviembre-2015	\$44,27	\$1.351,05	-31,41833333	-1110,871572	34901,73332	987,1117	1234035,6
diciembre-2015	\$38,01	\$1.172,75	-37,67833333	-1289,174203	48573,93535	1419,657	1661970,1
$\Sigma$	1816,52	59086,18087			535563,7915	16303,93	18258487
Medias	75,69	2461,92		Covarianza	23285,38224	708,8667	793847,24
					Desviación	26,62455	890,98106
					Correlación		0,981596

En la tabla están dispuestos los datos de los precios históricos de la acción de Ecopetrol provenientes del portal acciones de la página de la Bolsa de valores de Colombia en el link <https://www.bvc.com.co/pps/tibco/portalebvc/Home/Mercados/enlinea/acciones?action=dummy> y de los precios SPOT del barril de petróleo referencia BRENT encontrados en el portal Precios SPOT en la página de la Administración de información de energía de los Estados unidos en el link [http://www.eia.gov/dnav/pet/pet\\_pri\\_spt\\_s1\\_d.htm](http://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm), así como los datos necesarios para el cálculo del coeficiente de correlación y de la función de regresión lineal.

MES	PRECIO OPEP(USD)	PRECIO ECOP(COL)	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x}) \cdot (y - \bar{y})$	$(x - \bar{x})^2$	$(y - \bar{y})^2$
enero-2014	\$105	\$3.483	31,861198	1020,9329	32528,14692	1015,13596	1042304,07
febrero-2014	\$105	\$3.416	32,534517	953,5758	31024,12755	1058,49477	909306,8
marzo-2014	\$104	\$3.690	31,300159	1228,0758	38438,96822	979,699979	1508170,16
abril-2014	\$104	\$3.797	31,421112	1334,5758	41933,85532	987,286267	1781092,56
mayo-2014	\$105	\$3.532	32,592107	1070,4567	34888,44141	1062,24547	1145877,65
junio-2014	\$108	\$3.528	35,045398	1065,8536	37353,26221	1228,17989	1136043,84
julio-2014	\$106	\$3.264	32,761277	802,42362	26288,42293	1073,3013	643883,671
agosto-2014	\$101	\$3.224	27,907778	761,76001	21259,02953	778,844099	580278,309
septiembre-2014	\$96	\$3.314	23,13438	851,93943	19709,09075	535,199547	725800,798
octubre-2014	\$85	\$2.975	12,213451	512,62125	6260,874711	149,168394	262780,547
noviembre-2014	\$76	\$2.633	2,7190166	171,13135	465,3089811	7,39305105	29285,9398
diciembre-2014	\$59	\$2.046	-13,535165	-416,4242	5636,370407	183,200699	173409,117
enero-2015	\$44	\$1.926	-28,468412	-535,6742	15249,79392	810,450482	286946,852
febrero-2015	\$54	\$2.158	-18,786983	-303,6742	5705,122225	352,950747	92218,0216
marzo-2015	\$52	\$1.819	-20,386983	-	13101,46023	415,629094	412984,227
abril-2015	\$57	\$2.040	-15,546983	-421,9242	642,63849	6559,648599	241,708694
mayo-2015	\$62	\$1.931	-10,686983	-	5670,60333	114,211615	281545,289
junio-2015	\$60	\$1.743	-12,636983	-	530,60841	9083,037937	159,69335
julio-2015	\$54	\$1.625	-18,656983	-	718,76631	15618,72122	348,083031
agosto-2015	\$45	\$1.594	-27,386983	-	837,15148	23778,4743	750,046862
septiembre-2015	\$45	\$1.416	-28,016983	-	868,23999	29293,45308	784,951361
octubre-2015	\$45	\$1.409	-27,826983	-	1045,5606	29291,7538	774,341007
noviembre-2015	\$41	\$1.351	-32,346983	-	1052,6385	35933,34433	1046,32734
diciembre-2015	\$34	\$1.173	-39,206983	-	1110,8716	50544,63163	1537,18755
				-	1289,1742	535615,9436	16393,7306
Σ	1748,327603	59086,18087					18258486,6
medias	72,85	2461,92		Covarianza	23287,64972	712,770894	793847,245
					Desviaciones	26,6977694	890,981058
					Correlación	0,9790	

En la tabla están dispuestos los datos de los precios históricos de la acción de Ecopetrol provenientes del portal acciones de la página de la Bolsa de valores de Colombia en el link <https://www.bvc.com.co/pps/tibco/portalebvc/Home/Mercados/enlinea/acciones?action=dummy> y de los precios de la canasta OPEP encontrados en el portal OPEP BASKET PRICE en la página de la Organización de países exportadores de petróleo (OPEP) en el link [http://www.opec.org/opec\\_web/en/data\\_graphs/40.htm](http://www.opec.org/opec_web/en/data_graphs/40.htm), así como los datos necesarios para el cálculo del coeficiente de correlación.