

## ANEXO 8

# Regresión múltiple para los indicadores de las variables de entrada (tres estados)

María Camila Cardona<sup>a,c</sup>, María José Martínez<sup>a,c</sup>, Paula Andrea Bravo<sup>a,c</sup>, Carlos Andrés Hernández<sup>a,c</sup>

Isabel Cristina García<sup>b,c</sup>

<sup>a</sup>Estudiante de Ingeniería Industrial

<sup>b</sup>Profesora, Directora del Proyecto de Grado, Departamento de Ciencias Naturales y Matemáticas

<sup>c</sup>Pontificia Universidad Javeriana, Cali, Colombia

Fecha de elaboración: 06/08/2021

Teniendo en cuenta un intervalo de tiempo desde el 25 de marzo del 2019 hasta el 25 de marzo del 2021, se realiza la recolección de información por medio de bases de datos encontradas en la web para obtener los valores de la tasa representativa del mercado (TRM), el depósito a término fijo (DTF), el precio del petróleo Brent, la tasa de desempleo en Estados Unidos, la tasa prime y el estado de confinamiento en tres estados (pre-confinamiento, confinamiento y post-confinamiento). El propósito de esto es estudiar un modelo de regresión que relacione la variable del índice COLCAP con las restantes.

En la TABLA XXXI se presenta un resumen de las variables y las unidades en las cuales están dadas:

TABLA XXXI.  
 IDENTIFICACIÓN DE VARIABLES CON UNIDADES

Variable	Nombre	Unidades
Dependiente	Y: Índice COLCAP	Pesos (\$)
Independiente	X1: Tasa Representativa del Mercado (TRM)	Pesos (\$)
	X2: Depósito a Término Fijo (DTF)	Porcentaje (%)
	X3: Precio del Petróleo Brent	Dólares/Barril (\$)
	X4: Tasa de Desempleo USA	Porcentaje (%)
	X5: Tasa Prime	Porcentaje (%)
	X6: Pre-Confinamiento	Binaria
	X7: Confinamiento	Binaria

A partir de esto se realiza un análisis exploratorio por medio de diagramas de dispersión o nube de puntos para identificar puntos extraños.

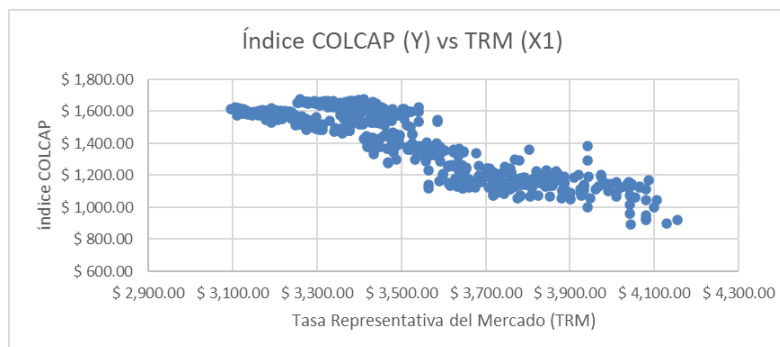


Fig. 80. Diagrama de dispersión del índice COLCAP y TRM

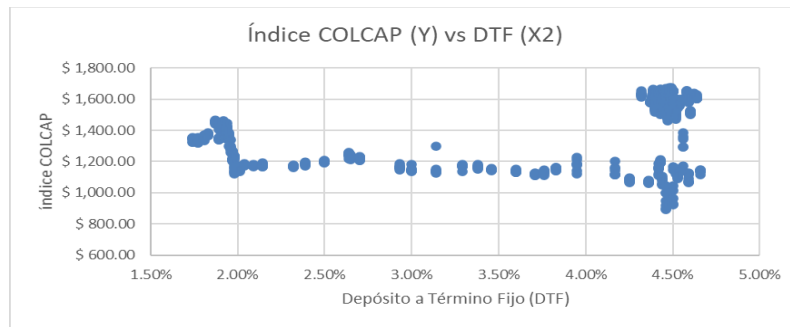


Fig. 81. Diagrama de dispersión del índice COLCAP y DTF

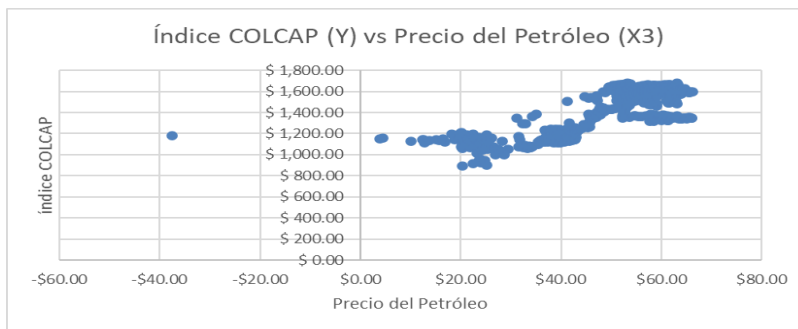


Fig. 82. Diagrama de dispersión del índice COLCAP y Precio del Petróleo

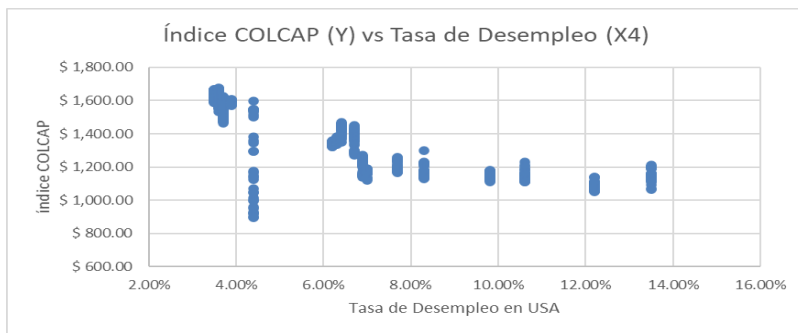


Fig. 83. Diagrama de dispersión del índice COLCAP y Tasa de Desempleo

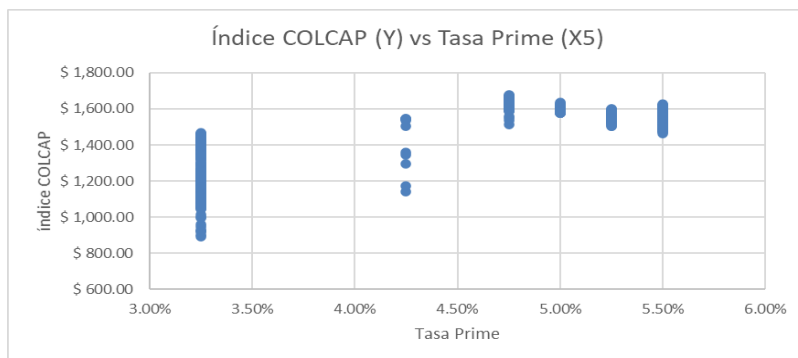


Fig. 84. Diagrama de dispersión del índice COLCAP y Tasa Prime

A partir de estos diagramas se puede identificar que **el único en el que se observa un dato atípico** es en la comparación del índice COLCAP con el precio del petróleo, ya que se observa como uno de sus puntos es negativo (-\$37.63) y al realizar una investigación de lo ocurrido el 20 de abril del 2020 que fue el día en el cual el valor del precio del petróleo se registra negativo, se identifica que hubo un colapso de la demanda por la crisis del COVID-19 y se tuvieron dificultades para almacenar el exceso de producción [1]. Por lo cual, se toma la decisión de eliminar ese día de la base de datos para la realización del estudio, al realizar la corrección del dato se genera un nuevo diagrama, en donde se observa un mejor comportamiento.

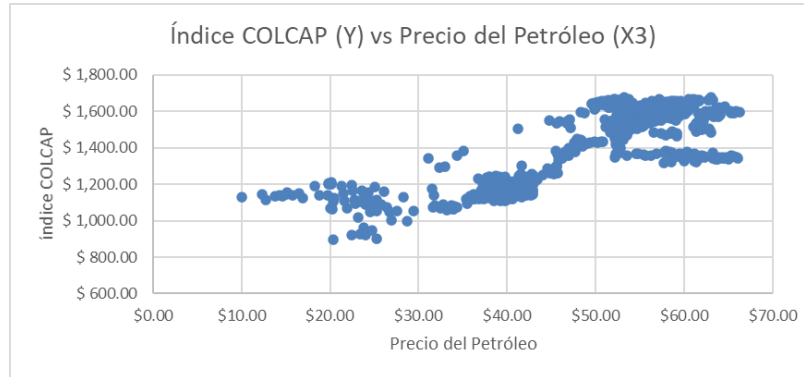


Fig. 85. Diagrama de dispersión del índice COLCAP y Precio del Petróleo corregido

Cabe resaltar que, al realizar la eliminación de ese día en la base de datos, el número total de datos en cada variable es de 731.

#### 1. Transformaciones curvilíneas para cada variable independiente.

Al realizar transformaciones curvilíneas para cada variable independiente (X) se busca seleccionar la transformación que mejor se ajuste con la variable de respuesta (Y: Índice COLCAP), en la cual se obtienen los siguientes resultados:

	Y	X1/1000	RAIZ(X1)	X1 <sup>2</sup>	1/X1	LN(X1)	Exp(x1)
Y	1						
X1/1000	-0.887490374	1					
RAIZ(X1)	-0.886803797	0.99983549	1				
X1 <sup>2</sup>	<b>-0.887648412</b>	0.99934246	0.998520895	1			
1/X1	0.882325566	-0.99737706	-0.998524861	-0.994104742	1		
LN(X1)	-0.885712601	0.99934241	0.999835634	0.997372368	-0.99934482	1	
Exp(x1)	-0.879792718	0.99148781	0.988983857	0.99553585	-0.97960087	0.98616476	1

Fig. 86. Correlación de las transformaciones de la TRM

	Y	X2	RAIZ(X2)	X2 <sup>2</sup>	1/X2	LN(X2)	Exp(x2)
Y	1						
X2	0.425402131	1					
RAIZ(X2)	0.407879087	0.99918948	1				
X2 <sup>2</sup>	<b>0.456385007</b>	0.99717116	0.99335526	1			
1/X2	-0.350557514	-0.98637054	-0.99215107	-0.9715175	1		
LN(X2)	0.389326741	0.99666916	0.99914175	0.98778582	-0.99646389	1	
Exp(x2)	0.426447752	0.99999701	0.99908868	0.99735196	-0.98597539	0.99646899	1

Fig. 87. Correlación de las transformaciones del DTF

	Y	X3	RAIZ(X3)	X3 <sup>2</sup>	1/X3	LN(X3)	Exp(x3)
Y	1						
X3	0.81781233	1					
RAIZ(X3)	0.79927413	0.99512753	1				
X3 <sup>2</sup>	0.8229258	0.98706972	0.96676003	1			
1/X3	-0.65077166	-0.89304457	-0.93155979	-0.81650157	1		
LN(X3)	0.76619486	0.97792414	0.99369677	0.93303497	-0.96579903	1	
Exp(x3)	0.08285161	0.25091115	0.22479985	0.29863185	-0.14027281	0.197247933	1

Fig. 88. Correlación de las transformaciones del Precio del Petróleo

	Y	X4	RAIZ(X4)	X4 <sup>2</sup>	1/X4	LN(X4)	Exp(x4)
Y	1						
X4	-0.82949537	1					
RAIZ(X4)	-0.85623939	0.99532814	1				
X4 <sup>2</sup>	-0.75584877	0.98199519	0.95932615	1			
1/X4	0.8901917	-0.94129662	-0.96926744	-0.86251232	1		
LN(X4)	-0.87463198	0.98228592	0.99577385	0.9297249	-0.98769502	1	
Exp(x4)	-0.82485083	0.99989238	0.9938182	0.98466078	-0.93640164	0.97947522	1

Fig. 89. Correlación de las transformaciones de la tasa de desempleo en USA

	Y	X5	RAIZ(X5)	X5 <sup>2</sup>	1/X5	LN(X5)	Exp(x5)
Y	1						
X5	0.8347627	1					
RAIZ(X5)	0.84167063	0.99971981	1				
X5 <sup>2</sup>	0.81846846	0.99873225	0.99726134	1			
1/X5	-0.8577525	-0.99620017	-0.99798188	-0.99055959	1		
LN(X5)	0.84777736	0.99893476	0.9997471	0.99534709	-0.99915726	1	
Exp(x5)	0.83410806	0.99999768	0.99966646	0.99883847	-0.99601034	0.998833	1

Fig. 90. Correlación de las transformaciones de la tasa prime

## 2. Matriz con las mejores transformaciones.

Al terminar de identificar cuáles son las mejores transformaciones que tienen relación con el índice COLCAP por medio de la correlación, se procede a realizar una matriz que no solo incluye las variables seleccionadas, sino que también se calculan las interacciones entre pares de X's transformadas.

Esto se hace con el fin de realizar un análisis de multicolinealidad calculando el factor inflacionario de varianza (VIF) por medio del software estadístico MINITAB, en donde se busca que el valor del VIF sea menor a 5 para descartar la presencia de multicolinealidad entre las variables.

Después de hacer una eliminación de variables para identificar quienes quedan en el modelo se llega al resultado mostrado por la Fig. 91.

Predictor	Coef	SE Coef	T	P	VIF
Constant	1949.78	42.28	46.11	0.000	
X1^2	-77.385	2.280	-33.94	0.000	2.380
X1*X3 T	0.0028694	0.0004250	6.75	0.000	3.231
X2*X5 T	2412.0	186.5	12.93	0.000	1.318
X4*X5 T	0.53519	0.02938	18.22	0.000	1.616

S = 68.4078 R-Sq = 88.9% R-Sq(adj) = 88.8%

Fig. 91. Análisis final del factor inflacionario de varianza (VIF)

En la cual se identifica que el VIF de las 4 variables restantes son menores a 5 y no se continúa excluyendo ninguna variable por problemas de multicolinealidad.

### 3. Método de selección de variables por eliminación hacia atrás.

En este punto se siguen escogiendo variables, por lo cual se hace uso del método de eliminación hacia atrás para observar las variables que se van a quedar porque aportan significativamente el comportamiento del índice COLCAP.

Al realizar regresión en estos datos, se puede identificar por medio de la Fig. 92 cuáles son los coeficientes de cada variable y la probabilidad (P-Valor) que ayudan a identificar quienes estarán en el modelo final por medio del planteamiento de hipótesis para cada  $\beta$  mostrado en la Fig. 93.

#### Resumen

Estadísticas de la regresión	
Coefficiente de correlación múltiple	0.948092369
Coefficiente de determinación R <sup>2</sup>	0.89887914
R <sup>2</sup> ajustado	0.898041122
Error típico	65.25241518
Observaciones	731

#### ANÁLISIS DE VARIANZA

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	6	27402632.98	4567105.497	1072.624869	0
Residuos	724	3082703.445	4257.877686		
Total	730	30485336.43			

	Coefficientes	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%	Inferior 95.0%	Superior 95.0%
Intercepción	1903.085572	42.02417655	45.28549346	2.0617E-213	1820.581776	1985.589369	1820.581776	1985.589369
X1^2	-64.53797978	3.009596712	-21.44406243	2.33657E-79	-70.44655847	-58.6294011	-70.44655847	-58.6294011
X1*X3 T	0.003186443	0.000415552	7.667972727	5.63888E-14	0.002370612	0.004002274	0.002370612	0.004002274
X2*X5 T	2504.854729	429.0576555	5.838037608	7.97225E-09	1662.509006	3347.200452	1662.509006	3347.200452
X4*X5 T	0.251136272	0.043656185	5.752593188	1.29684E-08	0.165428442	0.336844102	0.165428442	0.336844102
X6	46.12974422	16.95405757	2.720867498	0.006667426	12.84475861	79.41472983	12.84475861	79.41472983
X7	-78.94595262	14.90254258	-5.297482104	1.5595E-07	-108.2033096	-49.6885956	-108.2033096	-49.6885956

Fig. 92. Regresión para las variables del modelo

Como se mencionó anteriormente por medio de la siguiente figura se plantea la hipótesis para evaluar cada uno de los  $\beta$ 's a un nivel de confianza del 95%, donde se ve de forma clara que el P-Valor de cada una de las variables es menor a el alfa de 0.05, lo que indica que todas las variables presentadas hasta este punto aportan significativamente al comportamiento del índice COLCAP.

#### Planteamiento de hipótesis para cada $\beta$

Planteamiento de hipótesis para cada $\beta$		Regla de decisión: Rechazo $H_0$ si $p\text{-valor} < \alpha$		nivel de confianza del 95%		Con un nivel de confianza de 95% podemos afirmar que	
	Planteamiento de hipótesis	p-valor		$\alpha$	Decisión	Conclusión	
X1^2	$\beta_1$	Ho: $\beta_1 = 0$ vs Ha: $\beta_1 \neq 0$	2.33657E-79	<	0.05	Rechazo $H_0$	$\beta_1 \neq 0$
X1*X3 T	$\beta_2$	Ho: $\beta_2 = 0$ vs Ha: $\beta_2 \neq 0$	5.63888E-14	<	0.05	Rechazo $H_0$	$\beta_2 \neq 0$
X2*X5 T	$\beta_3$	Ho: $\beta_3 = 0$ vs Ha: $\beta_3 \neq 0$	7.97225E-09	<	0.05	Rechazo $H_0$	$\beta_3 \neq 0$
X4*X5 T	$\beta_4$	Ho: $\beta_4 = 0$ vs Ha: $\beta_4 \neq 0$	1.29684E-08	<	0.05	Rechazo $H_0$	$\beta_4 \neq 0$
X6	$\beta_5$	Ho: $\beta_5 = 0$ vs Ha: $\beta_5 \neq 0$	0.006667426	<	0.05	Rechazo $H_0$	$\beta_5 \neq 0$
X7	$\beta_6$	Ho: $\beta_6 = 0$ vs Ha: $\beta_6 \neq 1$	1.5595E-07	<	0.05	Rechazo $H_0$	$\beta_6 \neq 0$

Fig. 93. Planteamiento de hipótesis para cada  $\beta$

El modelo final tendrá entonces las siguientes variables:  $X_1^2$ ,  $X_1 * X_3$  T,  $X_2 * X_5$  T,  $X_4 * X_5$  T,  $X_6$ ,  $X_7$ .

### 4. Estimación con el modelo de regresión $Y^\wedge$ .

Al tener las variables que aportan significativamente, se obtiene un modelo completo de regresión lineal con la fórmula general

$$Y = \beta_0 + \beta_1 * X_1^2 + \beta_2 * X_1 X_3 T + \beta_3 * X_2 X_5 T + \beta_4 * X_4 X_5 T + \beta_5 * X_6 + \beta_6 * X_7$$

Esto permite realizar estimaciones teniendo en cuenta los coeficientes obtenidos en la Fig. 92 lo que finalmente se resume a una fórmula más precisa

$$Y = 1903.085 - 64.538 * X_{12} + 0.00319 * X_{13} T + 2504.855 * X_{25} T + 0.2511 X_{45} T + 46.13 * X_6 - 78.94 * X_7$$

Cuando se reemplaza en el modelo las observaciones de cada una de las  $X$ 's se obtiene una estimación del valor del índice COLCAP. Sin embargo, al tener una base de datos tan grande en la Fig. 94 solo se muestran las estimaciones del índice COLCAP, para los primeros 10 datos encontrados en la base de datos.

	TRM	DTF	Petróleo	Tasa Desempleo	Tasa Prime	Pre-Confinamiento	Confinamiento	COLCAP								Pronóstico
#	X1	X2	X3	X4	X5	X6	X7	Y	X1*X2	X1*X3 T	X2*X5 T	X4*X5 T	X6	X7		Y <sup>h</sup>
1	\$ 3,635.12	1.77%	\$58.56	6.20%	3.25%	0	0	\$ 1,320.11	\$ 13.21	\$ 45,314.76	0.964%	496.2779156	0	0		\$ 1,343.45
2	\$ 3,589.82	1.77%	\$61.18	6.20%	3.25%	0	0	\$ 1,325.33	\$ 12.89	\$ 48,235.22	0.964%	496.2779156	0	0		\$ 1,373.88
3	\$ 3,553.34	1.77%	\$57.76	6.20%	3.25%	0	0	\$ 1,319.82	\$ 12.63	\$ 42,123.83	0.964%	496.2779156	0	0		\$ 1,371.22
4	\$ 3,553.34	1.77%	\$61.55	6.20%	3.25%	0	0	\$ 1,335.96	\$ 12.63	\$ 47,833.22	0.964%	496.2779156	0	0		\$ 1,389.41
5	\$ 3,553.34	1.74%	\$63.24	6.20%	3.25%	0	0	\$ 1,344.07	\$ 12.63	\$ 50,499.23	0.932%	496.2779156	0	0		\$ 1,397.10
6	\$ 3,553.34	1.74%	\$63.96	6.20%	3.25%	0	0	\$ 1,348.70	\$ 12.63	\$ 51,657.56	0.932%	496.2779156	0	0		\$ 1,400.79
7	\$ 3,569.45	1.74%	\$61.42	6.20%	3.25%	0	0	\$ 1,324.54	\$ 12.74	\$ 48,064.26	0.932%	496.2779156	0	0		\$ 1,381.93
8	\$ 3,578.02	1.74%	\$60.00	6.20%	3.25%	0	0	\$ 1,325.59	\$ 12.80	\$ 46,088.02	0.932%	496.2779156	0	0		\$ 1,371.68
9	\$ 3,553.51	1.74%	\$64.60	6.20%	3.25%	0	0	\$ 1,343.56	\$ 12.63	\$ 52,696.30	0.932%	496.2779156	0	0		\$ 1,404.02
10	\$ 3,575.63	1.74%	\$64.80	6.20%	3.25%	0	0	\$ 1,350.14	\$ 12.79	\$ 53,685.27	0.932%	496.2779156	0	0		\$ 1,396.99

Fig. 94. Pronostico del índice COLCAP de los primeros diez datos

Para realizar una interpretación más precisa de lo que se muestra en la Fig. 94 se selecciona la fila cuatro. Esto nos indica que cuando la tasa representativa del mercado (TRM) es de \$3,553.34 pesos, el depósito a término fijo (DTF) es de 1.77%, el precio del petróleo es de \$61.55 dólares, la tasa de desempleo en USA es de 6.20%, la tasa prime es de 3.25% y se encuentre en post-confinamiento, se espera que el Índice COLCAP sea de \$1,389.41 pesos.

X1: Tasa Representativa del Mercado (TRM)	X2: Depósito a Término Fijo (DTF)	X3: Precio del Petróleo	X4: Tasa de Desempleo USA	X5: Tasa Prime	X6: Pre- Confinamiento	X7: Confinamiento	Y <sub>Observado</sub> : Índice COLCAP	Y <sub>Estimado</sub> : Índice COLCAP	Residuo o error
\$ 3,553.34	1.77%	\$ 61.55	6.20%	3.25%	0	0	\$ 1,335.96	\$ 1,389.41	\$ 53.45

Fig. 53. Comparación entre el Y observado y el Y estimado

Al realizar la comparación entre los dos valores del índice COLCAP, se observa que el  $Y_{\text{Estimado}}$  con los coeficientes arrojados por el ANOVA ( $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5$  y  $\beta_6$ ) es cercano al  $Y_{\text{Observado}}$ , en este caso con una diferencia de \$53.45 pesos. A partir de esto, se podría decir que el modelo permite pronosticar el valor aproximado del índice COLCAP con ciertos rangos de valores para la variable de la TRM ( $X_1$ ), DTF ( $X_2$ ), Precio del Petróleo ( $X_3$ ), Tasa de Desempleo en USA ( $X_4$ ), la Tasa Prime ( $X_5$ ), el estado de pre-confinamiento ( $X_6$ ) y el estado de confinamiento ( $X_7$ ).

##### 5. Interpretación del coeficiente de determinación y planteamiento de hipótesis para validar los coeficientes ( $\beta$ 's) del modelo.

###### Resumen

Estadísticas de la regresión	
Coefficiente de correlación múltiple	0.948092369
Coefficiente de determinación $R^2$	0.89887914
$R^2$ ajustado	0.898041122
Error típico	65.25241518
Observaciones	731

Fig. 96. Coeficiente de determinación ( $R^2$  ajustado)

En este caso se tienen más de 2 variables independientes, por lo cual, el coeficiente de determinación  $R^2$  ajustado es quien indica el porcentaje de variación que tiene variable dependiente ( $Y$ ) que está siendo explicada por las variables independientes ( $X$ 's). En este caso el  $R^2 = 0.8980$  lo que indica que el 89.80% de la variación en el índice COLCAP, está siendo explicada por la tasa representativa del mercado (TRM), el depósito a término fijo (DTF), el precio del petróleo, la tasa de desempleo en USA, la tasa prime y el estado de confinamiento (pre-confinamiento, confinamiento y post-confinamiento) y el otro 10.20% es la variación que no está siendo explicada por el modelo es debido a otros factores que no han sido incluidos en el modelo de regresión.

### ANÁLISIS DE VARIANZA

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	P-Valor
Regresión	6	27402632.98	4567105.497	1072.624869	0.000
Residuos	724	3082703.445	4257.877686		
Total	730	30485336.43			

Fig. 97. Planteamiento de hipótesis correspondiente a la tabla ANOVA

### Planteamiento de hipótesis de la tabla ANOVA

**Ho:**  $\beta_1=\beta_2=\beta_3=\beta_4=0$  vs **Ha:** los coeficientes del modelo  $\beta$ 's  $\neq 0$

**Regla de decisión:** Rechazo Ho si p-valor  $< \alpha$  (95% de confianza)

**Decisión:** Como el P-valor = 0.0000  $< \alpha = 0.05$  **Rechazo Ho**

**Conclusión:** Con un nivel de confianza del 95%, podemos afirmar que el modelo si explica significativamente el comportamiento del índice COLCAP.

### 6. Intervalos de confianza y de predicción del modelo final

Al tener una base de datos tan grande en la Fig. 98 sólo se mostrarán los intervalos para los primeros 10 datos encontrados en la base de datos y así poder realizar la interpretación de forma clara.

	TRM	DTF	Petróleo	Tasa Desempleo	Tasa Prime	Pre-Confinamiento	Confinamiento	COLCAP
#	X1	X2	X3	X4	X5	X6	X7	Y
1	\$ 3,635.12	1.77%	\$58.56	6.20%	3.25%	0	0	\$ 1,320.11
2	\$ 3,589.82	1.77%	\$61.18	6.20%	3.25%	0	0	\$ 1,325.33
3	\$ 3,553.34	1.77%	\$57.76	6.20%	3.25%	0	0	\$ 1,319.82
4	\$ 3,553.34	1.77%	\$61.55	6.20%	3.25%	0	0	\$ 1,335.96
5	\$ 3,553.34	1.74%	\$63.24	6.20%	3.25%	0	0	\$ 1,344.07
6	\$ 3,553.34	1.74%	\$63.96	6.20%	3.25%	0	0	\$ 1,348.70
7	\$ 3,569.45	1.74%	\$61.42	6.20%	3.25%	0	0	\$ 1,324.54
8	\$ 3,578.02	1.74%	\$60.00	6.20%	3.25%	0	0	\$ 1,325.59
9	\$ 3,553.51	1.74%	\$64.60	6.20%	3.25%	0	0	\$ 1,343.56
10	\$ 3,575.63	1.74%	\$64.80	6.20%	3.25%	0	0	\$ 1,350.14

Fig. 98. Base de datos de las variables originales

Con el fin de tener una interpretación más clara de los intervalos mostrados en la Fig. 99 se hace uso de los datos que se muestran en la Fig. 96 en la cual se hace la elección de la fila diez.

#	Pronosticos Y^A	Intervalo de confianza del 95% (Corto plazo)		Intervalo de predicción del 95% (Largo plazo)	
		Limite inferior	Limite superior	Limite inferior	Limite superior
	PFIT1	CLIM1	CLIM2	PLIM1	PLIM2
1	\$ 1,343.45	\$ 1,329.77	\$ 1,357.13	\$ 1,214.61	\$ 1,472.28
2	\$ 1,373.88	\$ 1,358.99	\$ 1,388.76	\$ 1,244.91	\$ 1,502.84
3	\$ 1,371.22	\$ 1,359.76	\$ 1,382.68	\$ 1,242.60	\$ 1,499.84
4	\$ 1,389.41	\$ 1,375.06	\$ 1,403.76	\$ 1,260.50	\$ 1,518.32
5	\$ 1,397.10	\$ 1,381.18	\$ 1,413.01	\$ 1,268.00	\$ 1,526.19
6	\$ 1,400.79	\$ 1,384.12	\$ 1,417.45	\$ 1,271.60	\$ 1,529.97
7	\$ 1,381.93	\$ 1,367.40	\$ 1,396.46	\$ 1,253.00	\$ 1,510.86
8	\$ 1,371.68	\$ 1,358.23	\$ 1,385.13	\$ 1,242.87	\$ 1,500.49
9	\$ 1,404.02	\$ 1,386.67	\$ 1,421.37	\$ 1,274.74	\$ 1,533.29
10	\$ 1,396.99	\$ 1,378.75	\$ 1,415.24	\$ 1,267.59	\$ 1,526.39

Fig. 99. Intervalos de confianza y de predicción al 95%



A partir de esto se puede concluir que con un nivel de confianza del 95% se espera a **CORTO PLAZO** que el índice COLCAP oscile entre \$1,378.75 pesos y \$1,415.24 pesos, cuando la tasa representativa del mercado (TRM) sea de \$3,575.63 pesos, el depósito a término fijo (DTF) sea de 1.74%, el precio del petróleo sea de \$64.80 dólares, la tasa de desempleo en USA sea de 6.20%, la tasa prime sea de 3.25% y se encuentre en estado de post-confinamiento.

Por otra parte, con un nivel de confianza del 95% se espera a **LARGO PLAZO** que el índice COLCAP oscile entre \$1,267.59 pesos y \$1,526.39 pesos, cuando la tasa representativa del mercado (TRM) sea de \$3,575.63 pesos, el depósito a término fijo (DTF) sea de 1.74%, el precio del petróleo sea de \$64.80 dólares, la tasa de desempleo en USA sea de 6.20%, la tasa prime sea de 3.25% y se encuentre en estado de post-confinamiento.

X1: Tasa Representativa del Mercado (TRM)	X2: Depósito a Término Fijo (DTF)	X3: Precio del Petróleo	X4: Tasa de Desempleo USA	X5: Tasa Prime	X6: Pre-Confinamiento	X7: Confinamiento	Y <sub>Observado</sub> : Índice COLCAP	Y <sub>Estimado</sub> : índice COLCAP	
								Limite inferior	Limite Superior
\$ 3,575.63	1.74%	\$ 64.80	6.20%	3.25%	0	0	\$ 1,350.14	\$ 1,378.75	\$ 1,415.24

Fig. 100. Comparación entre el Y observado y el intervalo del Y estimado a corto plazo

Cuando la variable X1= \$3,575.63, X2= 1.74%, X3= \$64.80, X4= 6.20%, X5= 3.25%, X6= 0 y X7= 0. El Y<sub>Observado</sub> no se encuentra dentro del intervalo de predicción de corto plazo, lo que indica que la predicción a corto plazo no es asertiva y el intervalo tiene poca amplitud.

X1: Tasa Representativa del Mercado (TRM)	X2: Depósito a Término Fijo (DTF)	X3: Precio del Petróleo	X4: Tasa de Desempleo USA	X5: Tasa Prime	X6: Pre-Confinamiento	X7: Confinamiento	Y <sub>Observado</sub> : Índice COLCAP	Y <sub>Estimado</sub> : índice COLCAP	
								Limite inferior	Limite Superior
\$ 3,575.63	1.74%	\$ 64.80	6.20%	3.25%	0	0	\$ 1,350.14	\$ 1,267.59	\$ 1,526.39

Fig. 101. Comparación entre el Y observado y el intervalo del Y estimado a largo plazo

Cuando la variable X1= \$3,575.63, X2= 1.74%, X3= \$64.80, X4= 6.20%, X5= 3.25%, X6= 0 y X7= 0. El Y<sub>Observado</sub> se encuentra dentro del intervalo de predicción de largo plazo, esto nos indica que la predicción es asertiva, ya que, el intervalo al 95% tiene una amplitud entre \$1,267.59 pesos y \$1,526.39 pesos. Sin embargo, el intervalo es muy amplio, por lo cual el rango de error del Y<sub>Estimado</sub> puede ser mayor.

## VALIDACIÓN DE SUPUESTOS DEL ANOVA

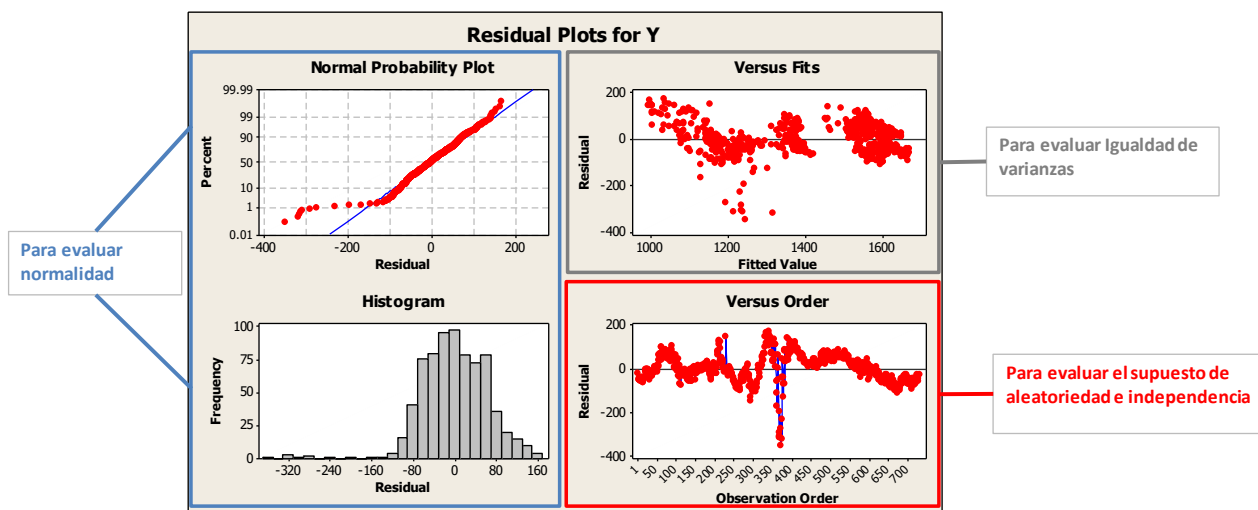


Fig. 102. Gráfico 4 en 1 para evaluación de supuestos con 3 estados

El 4 en 1 nos muestra clara cuál es el comportamiento de los datos y ayuda a evaluar e identificar de forma gráfica los supuestos del ANOVA.



7. Planteamiento de hipótesis para verificar el cumplimiento del supuesto de normalidad de los residuos.

**Planteamiento de hipótesis de prueba de bondad de ajuste a la distribución normal**

**H<sub>0</sub>:** Los residuos se ajustan a un comportamiento normal

**H<sub>a</sub>:** Los residuos no se ajustan a un comportamiento normal

**Regla de decisión:** Rechazo H<sub>0</sub> si p-valor <  $\alpha$  (para un nivel de confianza del 95%)

**Decisión:** Como el P-valor = 0.005 <  $\alpha = 0.05$  **Rechazo H<sub>0</sub>**

**Conclusión:** Con un nivel de confianza del 95%, se puede afirmar que los residuos NO se ajustan a un comportamiento normal. Lo que quiere decir, que hubo factores de ruido que afectaron la información tomada para el experimento y que dañaron el comportamiento normal en las variables.

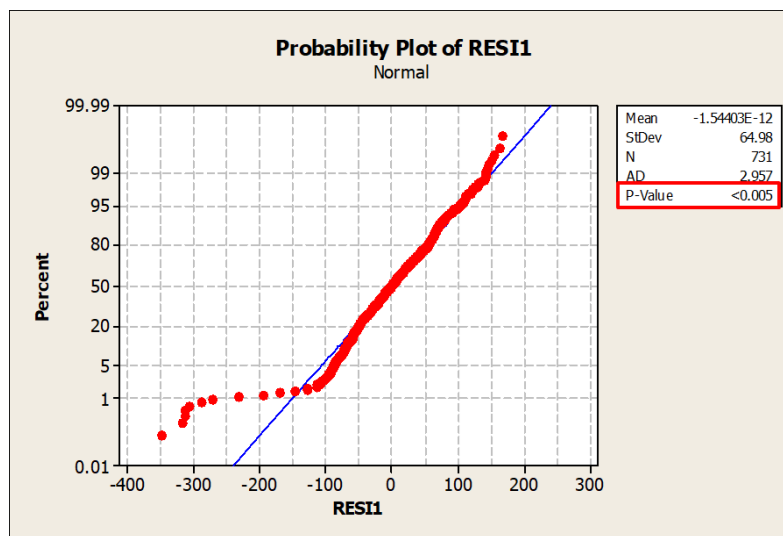


Fig. 103. Prueba de normalidad para los residuos

8. Planteamiento de hipótesis para verificar el cumplimiento del supuesto de igualdad de varianzas.

**Planteamiento de hipótesis de prueba de igualdad de varianzas "homocedasticidad"**

**H<sub>0</sub>:** Los residuos se tienen varianzas estadísticamente iguales

**H<sub>a</sub>:** Los residuos no tienen varianzas estadísticamente iguales

**Regla de decisión:** Rechazo H<sub>0</sub> si p-valor <  $\alpha$  (para un nivel de confianza del 95%)

**Decisión:** Como el P-valor = 0.002539 <  $\alpha = 0.05$  **Rechazo H<sub>0</sub>**

**Conclusión:** Con un nivel de confianza del 95%, se puede afirmar que los residuos NO tienen varianzas estadísticamente iguales. Esto significa que las parejas de datos no oscilan aleatoriamente por encima y por debajo de cero con una variación estadísticamente igual, lo que permite decir que los datos no son comparables entre ellos y hay variables que tienen variación significativamente diferente.

#### ANÁLISIS DE VARIANZA

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	P-Valor
Regresión	1	37725993.67	37725993.67	9.24019614	0.002539358
Residuos	363	1482061149	4082813.083		
Total	364	1519787143			

Fig. 104. Prueba de igualdad de varianza

9. Planteamiento de hipótesis para verificar el cumplimiento del supuesto de aleatoriedad e independencia de los residuos.

#### Planteamiento de hipótesis de prueba de aleatoriedad e independencia

**H<sub>0</sub>:** Los residuos son aleatorios e independientes

**H<sub>a</sub>:** Los residuos no son aleatorios ni independientes

#### Regla de decisión:

Si  $D.W. < dl$  Rechazo  $H_0$

Si  $dl < D.W. < du$  Cumple con reserva

Si  $D.W. > du$  No rechazo  $H_0$

**Decisión:** Como  $D.W. = 0.207284 < dl = 1.707$  **Rechazo  $H_0$**

**Conclusión:** Con un nivel de confianza del 95%, podemos afirmar que los residuos NO son aleatorios e independientes, lo que quiere decir que si existe autocorrelación y sesgos entre las variables analizadas.

Datos	
n =	731
k =	6
dl =	1.707
du =	1.831

Fig. 105. Datos para encontrar el valor dl y du en tablas

Dato arrojado en Minitab
Durbin-Watson statistic = 0.207284

Fig. 106. Prueba de aleatoriedad e independencia

#### 10. Conclusión

Después de plantear y validar las respectivas hipótesis para el supuesto de normalidad, igualdad de varianzas y aleatoriedad e independencia a un nivel de confianza del 95%, se puede decir que el modelo no es confiable para realizar estimaciones del índice COLCAP a corto y largo plazo, ya que los 3 supuestos del ANOVA no se cumplen.

Por otra parte, con base al coeficiente de determinación ( $R^2$  ajustado) el modelo mostró que el 89.80% de la variación del índice COLCAP está siendo explicado por él, y el 10.20% restante está explicado por otros factores que no han sido tomados en cuenta en el modelo.

### *11. Referencias*

[1]C. Tiempo, "Petróleo WTI en terreno negativo por primera vez en la historia", *Portafolio.co*, 2020. [En línea]. Disponible: <https://www.portafolio.co/internacional/noticias-hoy-precios-del-petroleo-20-de-abril-de-2020-crudo-hoy-540051>. [Accedido: 07- May- 2021].