

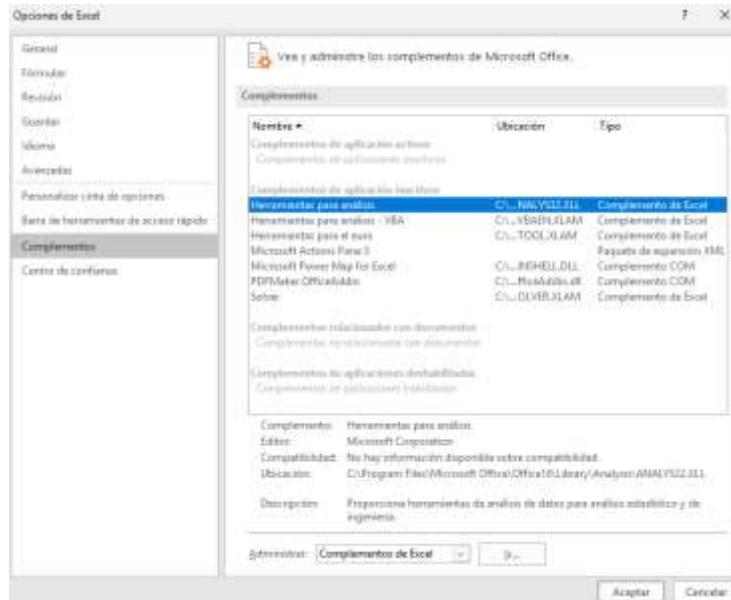
EJERCICIO 1

Paquete de herramientas de análisis en Excel

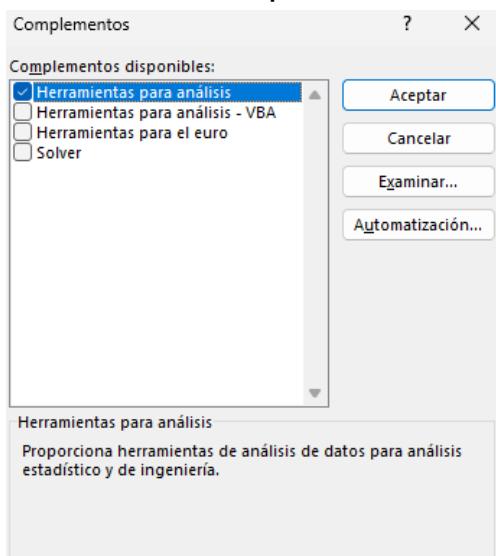
El paquete de herramientas de análisis es un programa complementario de Excel que proporciona herramientas de análisis de datos para el análisis de datos financieros, estadísticos y de ingeniería.

Para cargar el complemento Analysis ToolPak, siga los siguientes pasos.

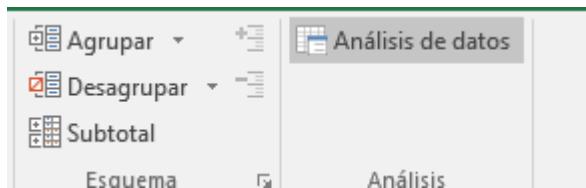
1. En la pestaña Archivo, haga clic en Opciones.
2. En Complementos, seleccione Analysis ToolPak y haga clic en el botón Ir.



3. Seleccione la opción Herramientas de análisis y haga clic en Aceptar.

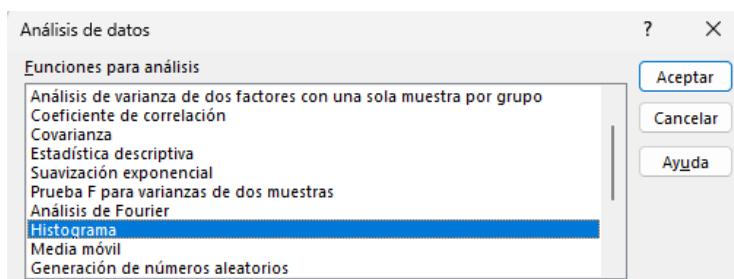


4. En la pestaña Datos, en el grupo Análisis, ahora puede hacer clic en Análisis de datos.



Aparece el siguiente cuadro de diálogo.

5. Por ejemplo, seleccione Histograma y haga clic en Aceptar para crear un histograma en Excel.



EJERCICIO 2

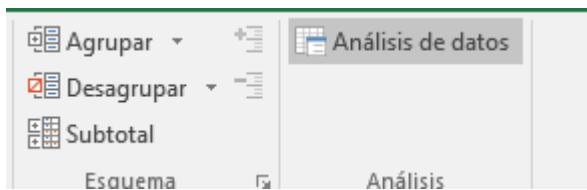
Histograma en Excel

Este ejemplo te enseña cómo crear un histograma en Excel.

1. Primero, ingrese los números de bin (niveles superiores) en el rango C4:C8.

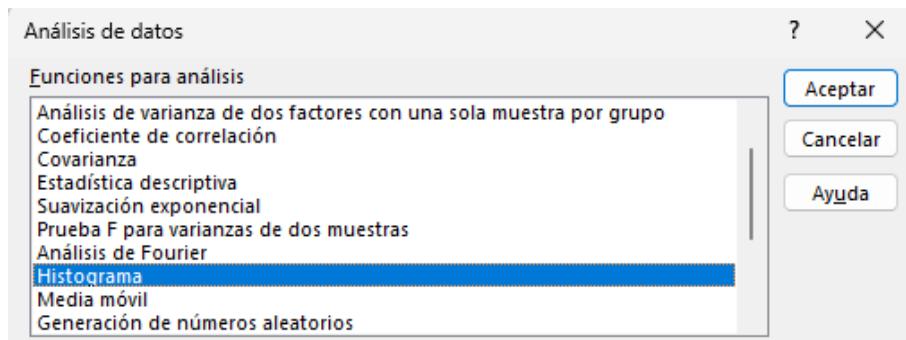
A1	B	C	D
Estudiantes			Estudiantes
1	22		
2	29		
3	40	20	
4	30	25	
5	45	30	
6	24	35	
7	21	40	
8	19		
9			
10	24		
11	22		
12	25		
13	52		
14	35		
15	40		
16	31		
17	37		
18	21		
19	23		

2. En la pestaña Datos, en el grupo Análisis, haga clic en Análisis de datos.



Nota: ¿No encuentra el botón "Análisis de datos"? Haga clic aquí para cargar el complemento Analysis ToolPak.

3. Seleccione Histograma y haga clic en Aceptar.

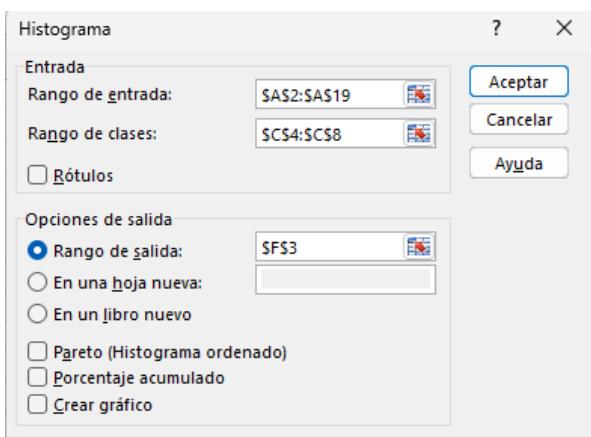


4. Seleccione el rango A2:A19.

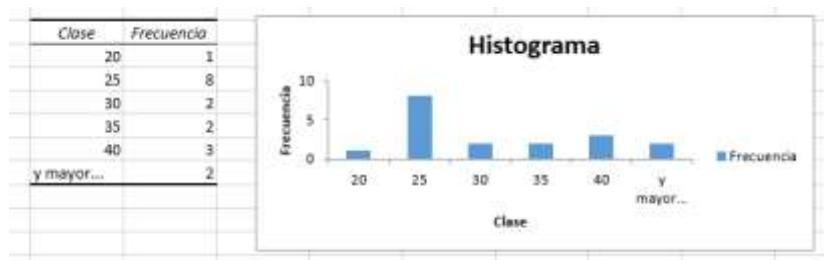
5. Haga clic en el cuadro Rango de contenedores y seleccione el rango C4:C8.

6. Haga clic en el botón de opción Rango de salida, haga clic en el cuadro Rango de salida y seleccione la celda F3.

7. Compruebe la salida del gráfico.



8. Haga clic en Aceptar.



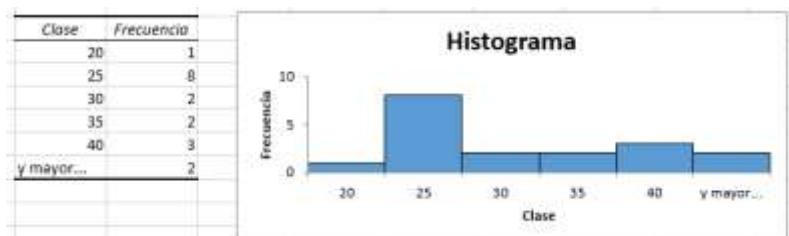
9. Haga clic en la leyenda del lado derecho y presione Eliminar.

10. Etiquete correctamente sus contenedores.

11. Para eliminar el espacio entre las barras, haga clic con el botón derecho en una barra, haga clic en Formato de serie de datos y cambie el Ancho del espacio a 0%.

12. Para agregar bordes, haga clic con el botón derecho en una barra, haga clic en Formato de serie de datos, haga clic en el ícono Relleno y Línea, haga clic en Borde y seleccione un color.

Resultado:



Si tiene Excel 2016 o posterior, simplemente utilice el tipo de gráfico Histograma.

13. Seleccione el rango A1:A19.

A
1 Estudiantes
2 22
3 29
4 40
5 30
6 48
7 24
8 21
9 19
10 24
11 22
12 25
13 52
14 35
15 40
16 31
17 37
18 21
19 23
20

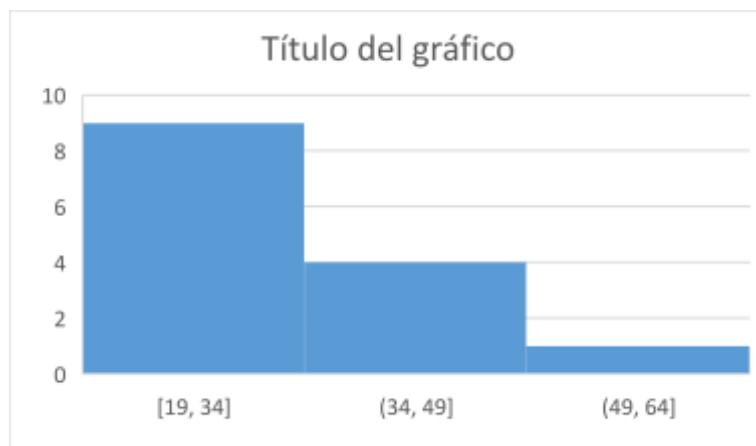
14. En la pestaña Insertar, en el grupo Gráficos, haga clic en el símbolo Histograma.



15. Haga clic en Histograma.

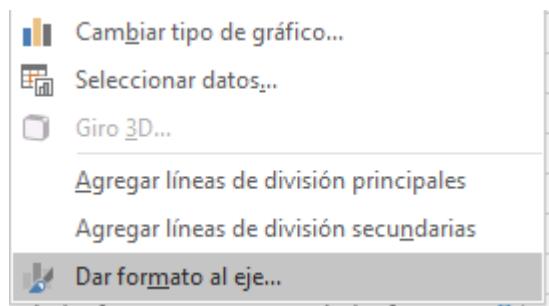


Resultado: un histograma con 3 intervalos.



Nota: Excel utiliza la regla de referencia normal de Scott para calcular el número de intervalos y el ancho de los intervalos.

16. Haga clic con el botón derecho en el eje horizontal y, a continuación, haga clic en Formato de eje.

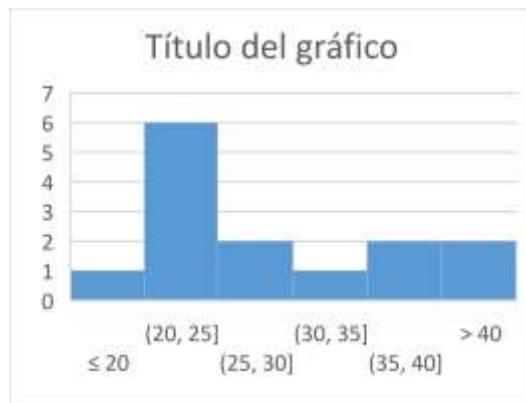


Aparece el panel Formato del eje.

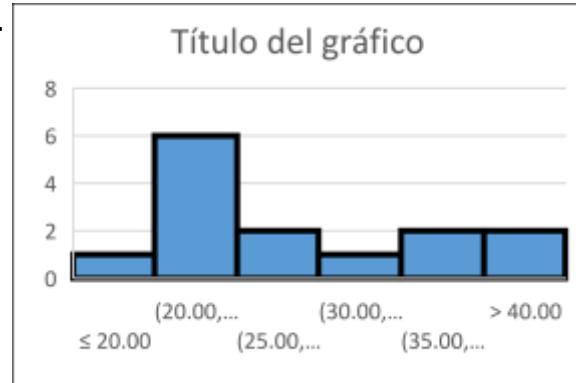
17. Defina los intervalos del histograma. Usaremos los mismos números de intervalo que antes (vea la primera imagen de esta página). Ancho del intervalo: 5. Número de intervalos: 6. Intervalo de desbordamiento: 40. Intervalo de subdesbordamiento: 20.



Resultado:



Recuerde que hicimos el siguiente histograma usando el paquete de herramientas de análisis (pasos 1-12).



Conclusión: las etiquetas de los intervalos parecen diferentes, pero los histogramas son iguales. ≤ 20 es lo mismo que 0-20, (20, 25] es lo mismo que 21-25, etc.

Consejo: también puedes usar tablas dinámicas para crear fácilmente una distribución de frecuencias en Excel.

EJERCICIO 3

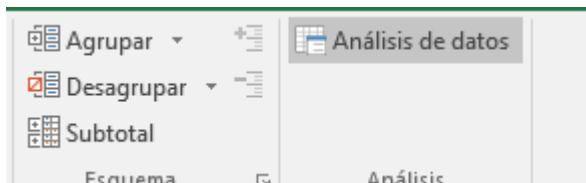
Estadísticas descriptivas en Excel

Puede utilizar el complemento Analysis Toolpak para generar estadísticas descriptivas. Por ejemplo, puede tener las puntuaciones de 14 participantes en una prueba.

	A
1	PUNTUACION
2	22
3	29
4	40
5	30
6	48
7	24
8	21
9	19
10	24
11	22
12	25
13	52
14	35
15	40

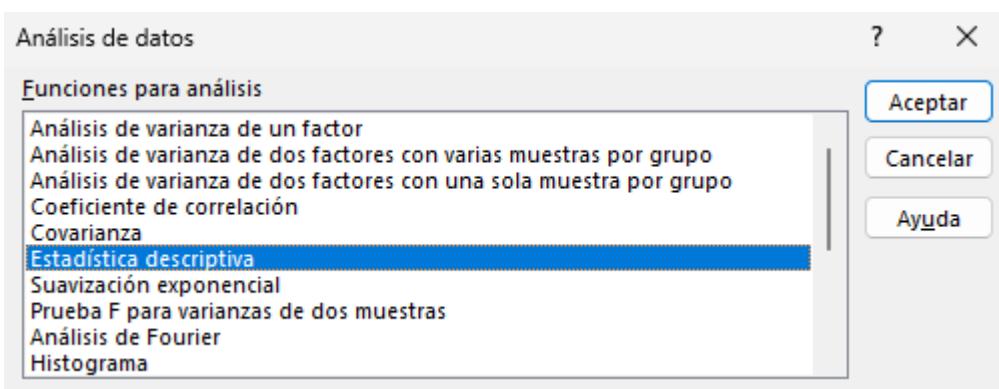
Para generar estadísticas descriptivas para estas puntuaciones, siga los siguientes pasos.

1. En la pestaña Datos, en el grupo Análisis, haga clic en Análisis de datos.

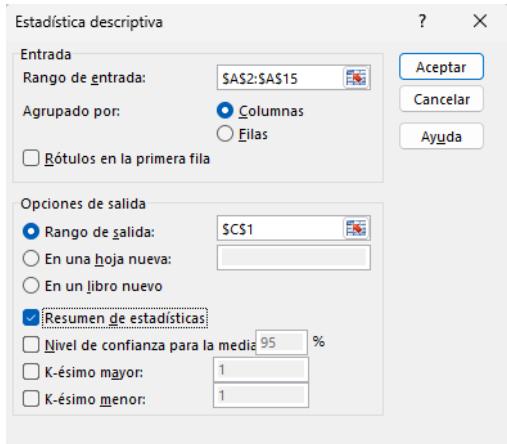


Nota: ¿No encuentra el botón "Análisis de datos"? Haga clic aquí para cargar el complemento Analysis ToolPak.

2. Seleccione Estadísticas descriptivas y haga clic en Aceptar.



3. Seleccione el rango A2:A15 como rango de entrada.
4. Seleccione la celda C1 como Rango de salida.
5. Asegúrese de que la casilla de Estadísticas resumidas esté marcada.



6. Haga clic en Aceptar.

Resultado:

	A	B	C	D
1	PUNTUACION		Columna1	
2		22		
3		29	Media	30.7857143
4		40	Error típico	2.8248866
5		30	Mediana	27
6		48	Moda	22
7		24	Desviación e	10.5697578
8		21	Varianza de	111.71978
9		19	Curtosis	-0.40098707
10		24	Coeficiente	0.87648187
11		22	Rango	33
12		25	Mínimo	19
13		52	Máximo	52
14		35	Suma	431
15		40	Cuenta	14

Consejo: visita nuestra página sobre funciones estadísticas para obtener más información sobre este tema.

EJERCICIO 4

ANOVA en Excel

Este ejemplo le enseña cómo realizar un ANOVA (análisis de varianza) de un factor en Excel. Un ANOVA de un factor o de una vía se utiliza para probar la hipótesis nula de que las medias de varias poblaciones son todas iguales.

A continuación, puedes encontrar los salarios de personas que tienen un título en economía, medicina o historia.

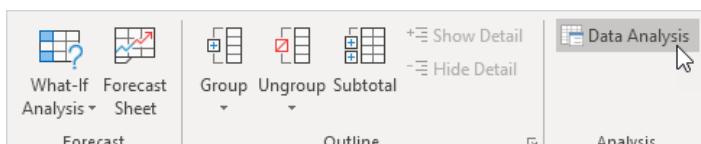
$$H_0 : \mu_1 = \mu_2 = \mu_3$$

H_1 : al menos una de las medias es diferente.

	A	B	C	D
1	economics	medicine	history	
2	42	69	35	
3	53	54	40	
4	49	58	53	
5	53	64	42	
6	43	64	50	
7	44	55	39	
8	45	56	55	
9	52		39	
10	54		40	
11				

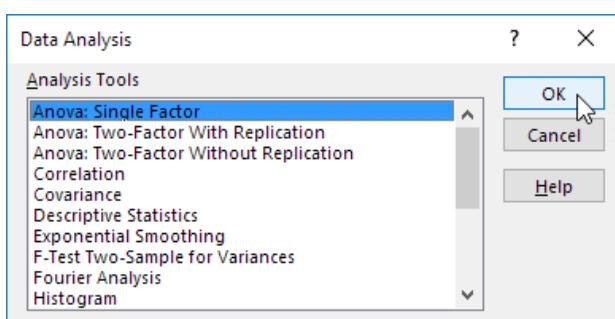
Para realizar un ANOVA de un solo factor, siga los siguientes pasos.

1. En la pestaña Datos, en el grupo Análisis, haga clic en Análisis de datos.



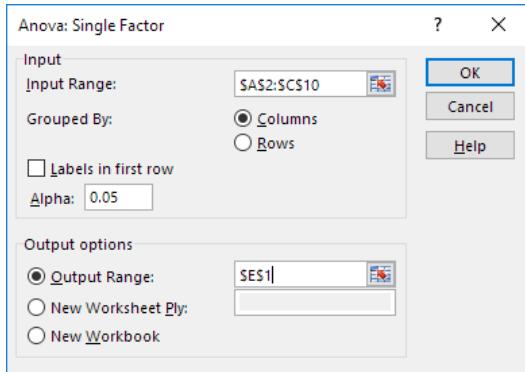
Nota: ¿No encuentra el botón "Análisis de datos"? Haga clic aquí para cargar el complemento Analysis ToolPak.

2. Seleccione Anova: Factor único y haga clic en Aceptar.



3. Haga clic en el cuadro Rango de entrada y seleccione el rango A2:C10.

4. Haga clic en el cuadro Rango de salida y seleccione la celda E1.



5. Haga clic en Aceptar.

Resultado:

E	F	G	H	I	J	K
Anova: Single Factor						
SUMMARY						
Groups	Count	Sum	Average	Variance		
Column 1	9	435	48.33333	23.5		
Column 2	7	420	60	32.33333		
Column 3	9	393	43.66667	50.5		
ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	1085.84	2	542.92	15.19623	7.16E-05	3.443357
Within Groups	786	22	35.72727			
Total	1871.84	24				

Conclusión: si $F > F$ crítico, rechazamos la hipótesis nula. Este es el caso, $15.196 > 3.443$. Por lo tanto, rechazamos la hipótesis nula. Las medias de las tres poblaciones no son iguales. Al menos una de las medias es diferente. Sin embargo, el ANOVA no indica dónde radica la diferencia. Se necesita una prueba t para comparar cada par de medias.

EJERCICIO 5

Prueba F en Excel

Este ejemplo le enseña a realizar una prueba F en Excel. La prueba F se utiliza para comprobar la hipótesis nula de que las varianzas de dos poblaciones son iguales.

A continuación, se muestran las horas de estudio de 6 alumnas y 5 alumnos.

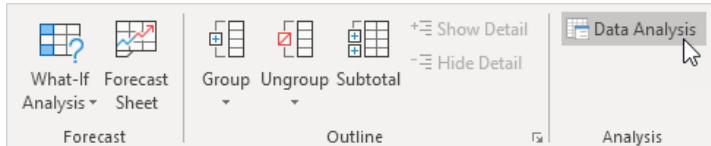
$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

	A	B	C
1	Female	Male	
2	26	23	
3	25	30	
4	43	18	
5	34	25	
6	18	28	
7	52		
8			

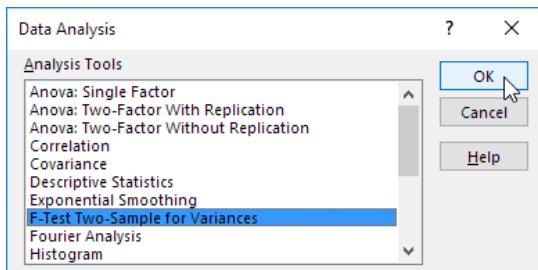
Para realizar una prueba F, ejecute los siguientes pasos.

1. En la pestaña Datos, en el grupo Análisis, haga clic en Análisis de datos.



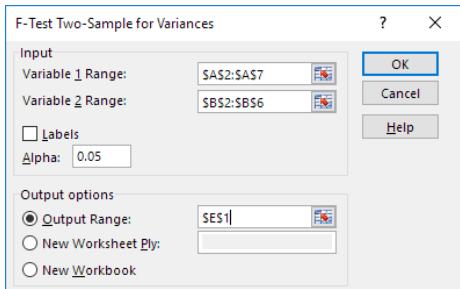
Nota: ¿No encuentra el botón "Análisis de datos"? Haga clic aquí para cargar el complemento Analysis ToolPak.

2. Seleccione Prueba F de dos muestras para varianzas y haga clic en Aceptar.



3. Haga clic en el cuadro Rango de variable 1 y seleccione el rango A2:A7.

4. Haga clic en el cuadro Rango de variable 2 y seleccione el rango B2:B6.
5. Haga clic en el cuadro Rango de salida y seleccione la celda E1.



6. Haga clic en Aceptar.

Resultado:

E	F	G
F-Test Two-Sample for Variances		
	Variable 1	Variable 2
Mean	33	24.8
Variance	160	21.7
Observations	6	5
df	5	4
F	7.373271889	
P(F<=f) one-tail	0.037888376	
F Critical one-tail	6.256056502	

Importante: asegúrese de que la varianza de la Variable 1 sea mayor que la de la Variable 2. En este caso, $160 > 21,7$. De lo contrario, intercambie los datos. Como resultado, Excel calcula el valor F correcto, que es el cociente entre la varianza 1 y la varianza 2 ($F = 160 / 21,7 = 7,373$).

Conclusión: si $F > F$ crítica unilateral, rechazamos la hipótesis nula. En este caso, $7,373 > 6,256$. Por lo tanto, rechazamos la hipótesis nula. Las varianzas de las dos poblaciones son desiguales.

EJERCICIO 6

Prueba t en Excel

Este ejemplo le enseña a realizar una prueba t en Excel. La prueba t se utiliza para comprobar la hipótesis nula de que las medias de dos poblaciones son iguales.

A continuación, se muestran las horas de estudio de 6 alumnas y 5 alumnos.

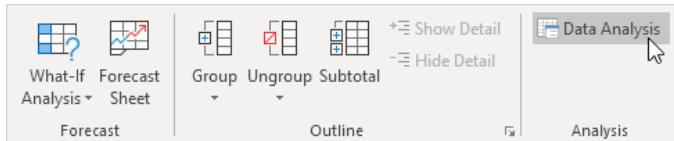
$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

	A	B	C
1	Female	Male	
2	26	23	
3	25	30	
4	43	18	
5	34	25	
6	18	28	
7	52		
8			

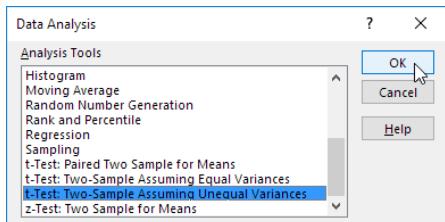
Para realizar una prueba t, ejecute los siguientes pasos.

1. Primero, realice una prueba F para determinar si las varianzas de ambas poblaciones son iguales. Esto no es así.
2. En la pestaña Datos, en el grupo Análisis, haga clic en Análisis de datos.

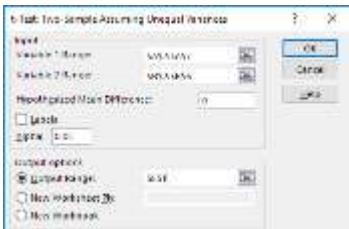


Nota: ¿No encuentra el botón "Análisis de datos"? Haga clic aquí para cargar el complemento Analysis ToolPak.

3. Seleccione Prueba t: Dos muestras asumiendo varianzas desiguales y haga clic en Aceptar.



4. Haga clic en el cuadro Rango de variable 1 y seleccione el rango A2:A7.
5. Haga clic en el cuadro Rango de variable 2 y seleccione el rango B2:B6.
6. Haga clic en el cuadro Diferencia de medias hipotética y escriba 0 ($H_0: \mu_1 - \mu_2 = 0$).
7. Haga clic en el cuadro Rango de salida y seleccione la celda E1.



8. Haga clic en Aceptar.

Resultado:

E	F	G
t-Test: Two-Sample Assuming Unequal Variances		
	Variable 1	Variable 2
Mean	33	24.8
Variance	160	21.7
Observations	6	5
Hypothesized Mean Difference	0	
df	7	
t Stat	1.47260514	
P(T<=t) one-tail	0.092170202	
t Critical one-tail	1.894578605	
P(T<=t) two-tail	0.184340405	
t Critical two-tail	2.364624252	

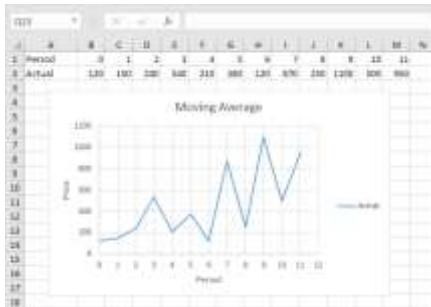
Conclusión: Realizamos una prueba de dos colas (desigualdad). Si t Stat < -t Critical two-tail o t Stat > t Critical two-tail, rechazamos la hipótesis nula. Este no es el caso: $-2,365 < 1,473 < 2,365$. Por lo tanto, no rechazamos la hipótesis nula. La diferencia observada entre las medias muestrales (33 - 24,8) no es lo suficientemente convincente como para afirmar que el promedio de horas de estudio entre estudiantes mujeres y hombres difiere significativamente.

EJERCICIO 7

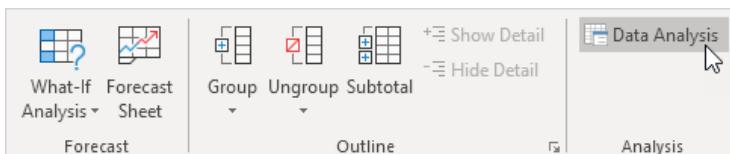
Promedio móvil en Excel

Este ejemplo le enseña a calcular la media móvil de una serie temporal en Excel. La media móvil se utiliza para suavizar irregularidades (picos y valles) y así identificar tendencias fácilmente.

1. Primero, echemos un vistazo a nuestra serie temporal.

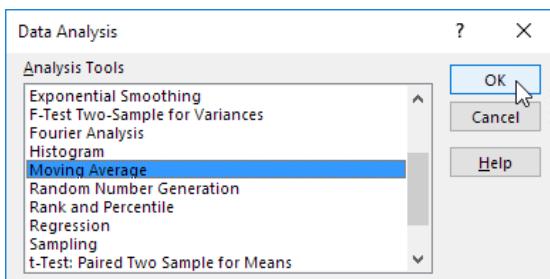


2. En la pestaña Datos, en el grupo Análisis, haga clic en Análisis de datos.



Nota: ¿No encuentra el botón "Análisis de datos"? Haga clic aquí para cargar el complemento Analysis ToolPak.

3. Seleccione Promedio móvil y haga clic en Aceptar.

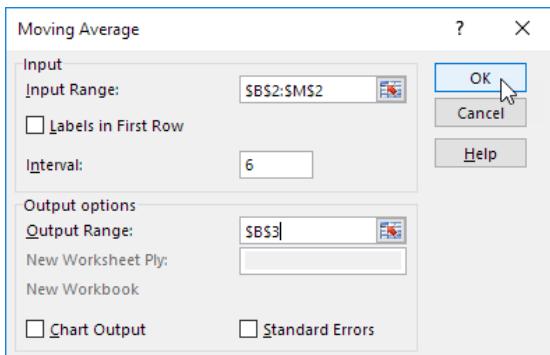


4. Haga clic en el cuadro Rango de entrada y seleccione el rango B2:M2.

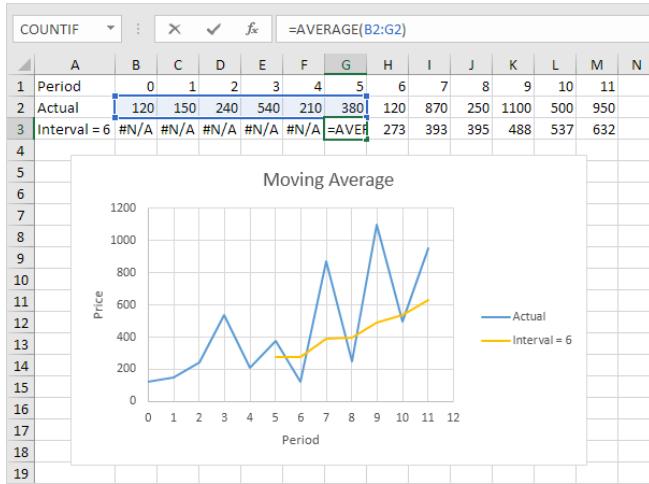
5. Haga clic en el cuadro Intervalo y escriba 6.

6. Haga clic en el cuadro Rango de salida y seleccione la celda B3.

7. Haga clic en Aceptar.



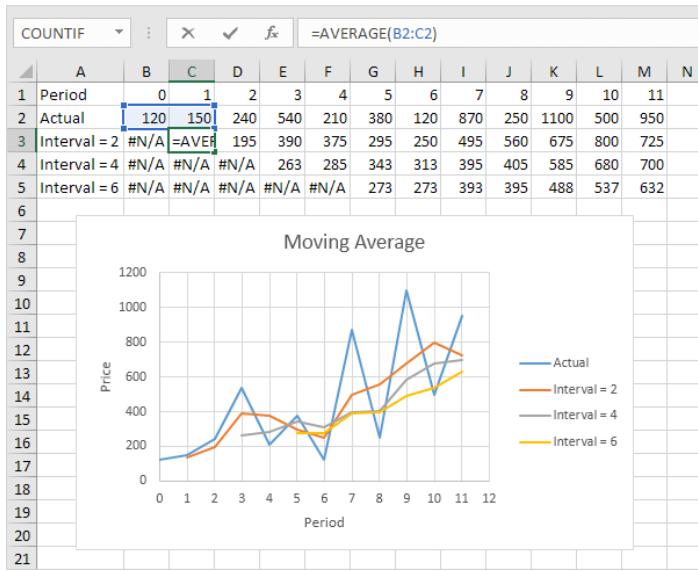
8. Dibuje un gráfico de estos valores.



Explicación: Como el intervalo es 6, la media móvil es el promedio de los 5 puntos de datos anteriores y el punto de datos actual. Como resultado, los picos y valles se suavizan. El gráfico muestra una tendencia creciente.

Excel no puede calcular la media móvil de los primeros 5 puntos de datos porque no hay suficientes puntos de datos anteriores.

9. Repita los pasos 2 a 8 para el intervalo = 2 y el intervalo = 4.



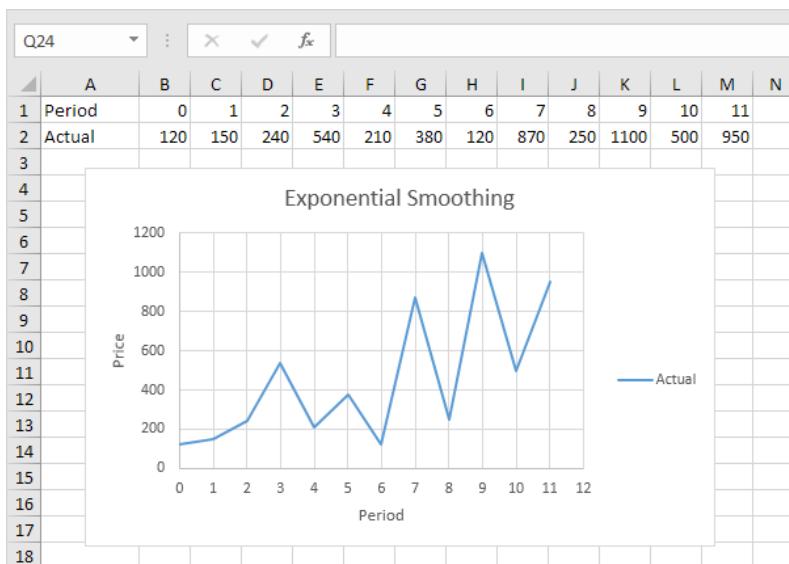
Conclusión: Cuanto mayor sea el intervalo, más se suavizan los picos y valles. Cuanto menor sea el intervalo, más cerca estarán las medias móviles de los puntos de datos reales.

EJERCICIO 8

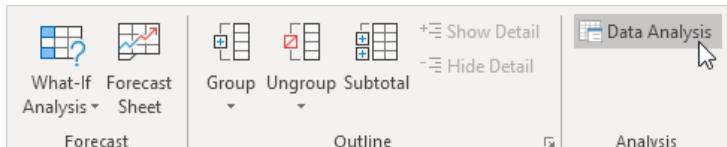
Suavizado exponencial en Excel

Este ejemplo le enseña a aplicar el suavizado exponencial a una serie temporal en Excel. El suavizado exponencial se utiliza para suavizar irregularidades (picos y valles) y así identificar tendencias fácilmente.

1. Primero, echemos un vistazo a nuestra serie temporal.

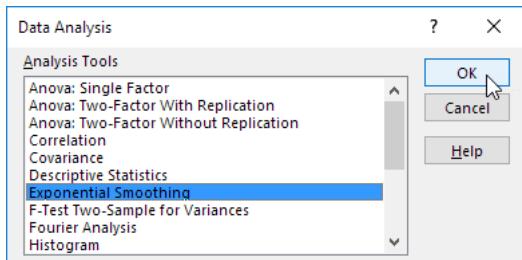


2. En la pestaña Datos, en el grupo Análisis, haga clic en Análisis de datos.

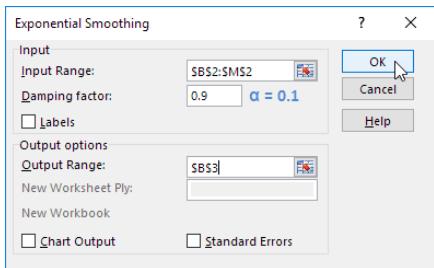


Nota: ¿No encuentra el botón "Análisis de datos"? Haga clic aquí para cargar el complemento Analysis ToolPak.

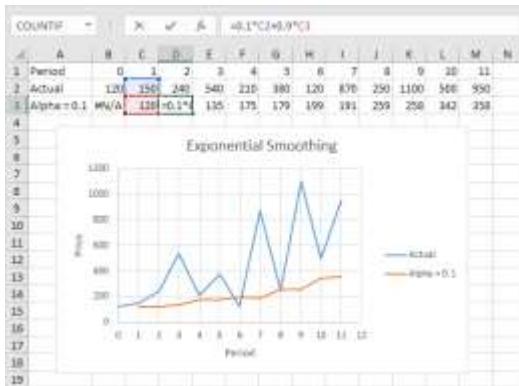
3. Seleccione Suavizado exponencial y haga clic en Aceptar.



4. Haga clic en el cuadro Rango de entrada y seleccione el rango B2:M2.
5. Haga clic en el cuadro "Factor de amortiguamiento" e introduzca 0,9. La literatura suele referirse a la constante de suavizado α (alfa). El valor $(1-\alpha)$ se denomina factor de amortiguamiento.
6. Haga clic en el cuadro Rango de salida y seleccione la celda B3.
7. Haga clic en Aceptar.

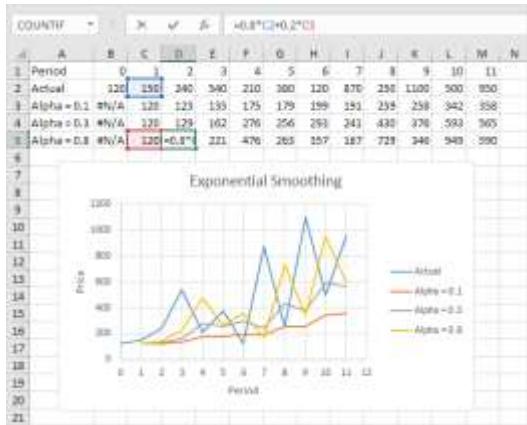


8. Dibuje un gráfico de estos valores.



Explicación: Dado que establecemos alfa en 0,1, el punto de datos anterior tiene una ponderación relativamente baja, mientras que el valor suavizado anterior tiene una ponderación alta (es decir, 0,9). Como resultado, se suavizan los picos y valles. El gráfico muestra una tendencia creciente. Excel no puede calcular el valor suavizado del primer punto de datos porque no existe ningún punto de datos anterior. El valor suavizado del segundo punto de datos es igual al anterior.

9. Repita los pasos 2 a 8 para alfa = 0,3 y alfa = 0,8.



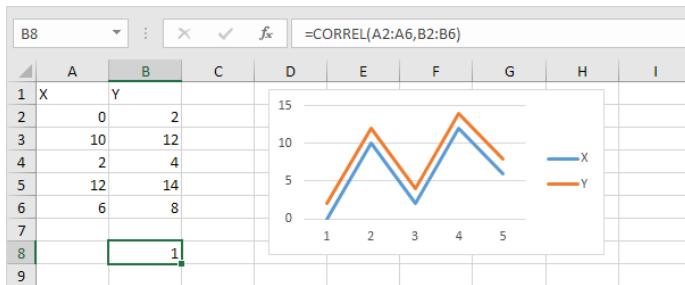
Conclusión: Cuanto menor sea el alfa (y mayor el factor de amortiguamiento), más se suavizan los picos y valles. Un alfa mayor (y un factor de amortiguamiento menor) hace que los valores suavizados se acerquen más a los puntos de datos reales.

EJERCICIO 9

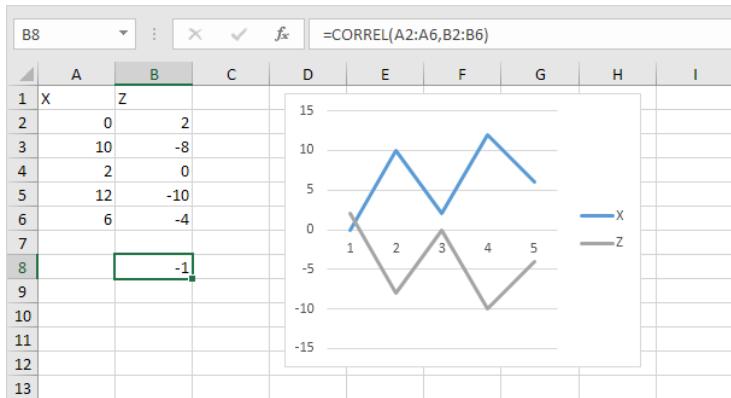
Análisis de correlación en Excel

El coeficiente de correlación (un valor entre -1 y +1) indica la intensidad de la relación entre dos variables. Podemos usar la función CORREL o el complemento Analysis Toolpak de Excel para calcular el coeficiente de correlación entre dos variables.

Un coeficiente de correlación de +1 indica una correlación positiva perfecta. A medida que la variable X aumenta, la variable Y aumenta. A medida que la variable X disminuye, la variable Y disminuye.



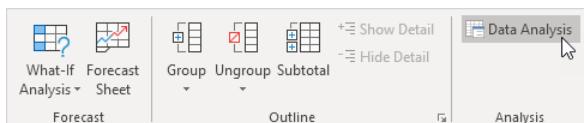
Un coeficiente de correlación de -1 indica una correlación negativa perfecta. A medida que la variable X aumenta, la variable Z disminuye. A medida que la variable X disminuye, la variable Z aumenta.



- Un coeficiente de correlación cercano a 0 indica que no hay correlación.

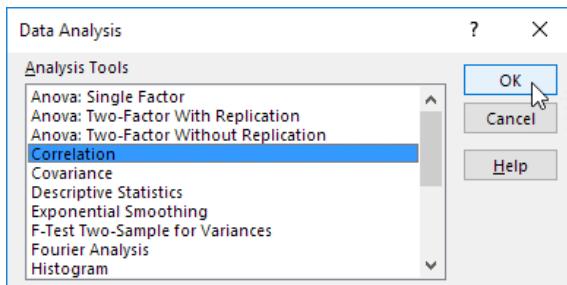
Para utilizar el complemento Analysis Toolpak en Excel para generar rápidamente coeficientes de correlación entre múltiples variables, ejecute los siguientes pasos.

1. En la pestaña Datos, en el grupo Análisis, haga clic en Análisis de datos.

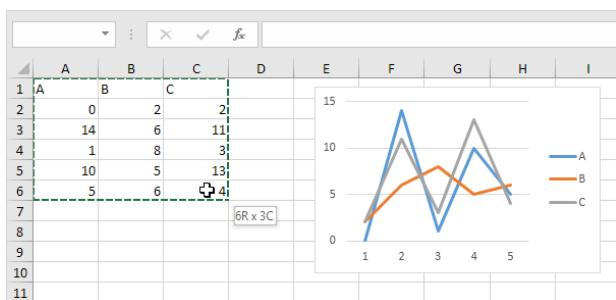


Nota: ¿No encuentra el botón "Análisis de datos"? Haga clic aquí para cargar el complemento Analysis ToolPak.

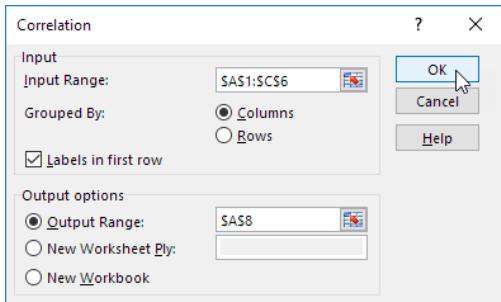
2. Seleccione Correlación y haga clic en Aceptar.



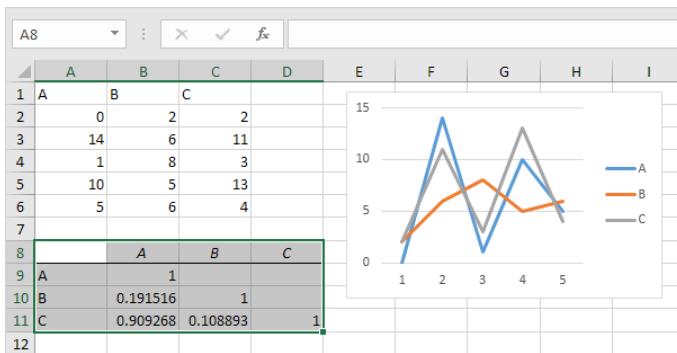
3. Por ejemplo, seleccione el rango A1:C6 como rango de entrada.



4. Verifique las etiquetas en la primera fila.
5. Seleccione la celda A8 como Rango de salida.
6. Haga clic en Aceptar.



Resultado:



Conclusión: Las variables A y C presentan una correlación positiva (0,91). Las variables A y B no presentan correlación (0,19). Las variables B y C tampoco presentan correlación (0,11). Puede comprobar estas conclusiones observando el gráfico.

EJERCICIO 10

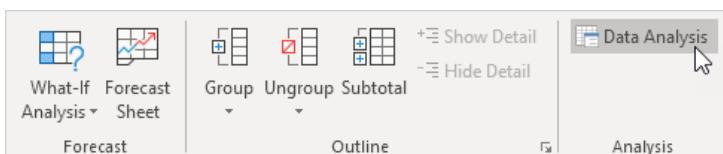
Análisis de regresión en Excel

Este ejemplo le enseña cómo ejecutar un análisis de regresión lineal en Excel y cómo interpretar el resultado del resumen.

A continuación, encontrará nuestros datos. La pregunta clave es: ¿existe una relación entre la cantidad vendida (producción) y el precio y la publicidad (insumo)? En otras palabras: ¿podemos predecir la cantidad vendida si conocemos el precio y la publicidad?

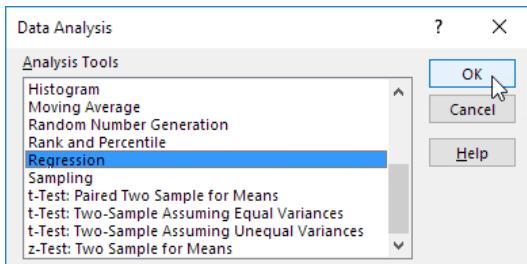
	A	B	C	D
1	Quantity Sold	Price	Advertising	
2	8500	\$2	\$2,800	
3	4700	\$5	\$200	
4	5800	\$3	\$400	
5	7400	\$2	\$500	
6	6200	\$5	\$3,200	
7	7300	\$3	\$1,800	
8	5600	\$4	\$900	
9				

1. En la pestaña Datos, en el grupo Análisis, haga clic en Análisis de datos.



Nota: ¿No encuentra el botón "Análisis de datos"? Haga clic aquí para cargar el complemento Analysis ToolPak.

2. Seleccione Regresión y haga clic en Aceptar.



3. Seleccione el rango Y (A1:A8). Esta es la variable predictora (también llamada variable dependiente).

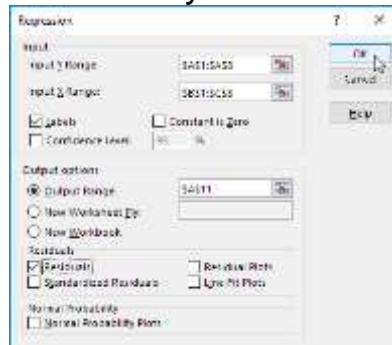
4. Seleccione el rango X (B1:C8). Estas son las variables explicativas (también llamadas variables independientes). Estas columnas deben ser adyacentes.

5. Revise las etiquetas.

6. Haga clic en el cuadro Rango de salida y seleccione la celda A11.

7. Verificar residuos.

8. Haga clic en Aceptar.



Excel produce la siguiente salida de resumen (redondeada a 3 decimales).

R cuadrado

El R cuadrado es igual a 0,962, lo cual es un ajuste muy bueno. El 96 % de la variación en la cantidad vendida se explica por las variables independientes precio y publicidad. Cuanto más cercano a 1, mejor se ajusta la línea de regresión (seguir leyendo) a los datos.

11	SUMMARY OUTPUT
12	
13	Regression Statistics
14	Multiple R 0.981
15	R Square 0.962
16	Adjusted R Square 0.943
17	Standard Error 310.524
18	Observations 7
19	

Significancia de los valores F y P

Para comprobar si sus resultados son fiables (estadísticamente significativos), consulte la Significancia F (0,001). Si este valor es inferior a 0,05, no hay problema. Si la Significancia F es superior a 0,05, probablemente sea mejor dejar de usar este conjunto de variables independientes. Elimine una variable con un valor p alto (superior a 0,05) y vuelva a ejecutar la regresión hasta que la Significancia F sea inferior a 0,05.

La mayoría o todos los valores p deben ser inferiores a 0,05. En nuestro ejemplo, este es el caso (0,000 , 0,001 y 0,005).

20	ANOVA						
21		df	SS	MS	F	Significance F	
22	Regression	2	9694299.568	4847149.784	50.269	0.001	
23	Residual	4	385700.432	96425.108			
24	Total	6	10080000.000				
25							
26		Coefficients	Std Error	t Stat	P-values	Lower 95%	Upper 95%
27	Intercept	8536.214	386.912	22.062	0.000	7461.975	9610.453
28	Price	-835.722	99.653	-8.386	0.001	-1112.404	-559.041
29	Advertising	0.592	0.104	5.676	0.005	0.303	0.882
30							

Coeficientes

La línea de regresión es: $y = \text{Cantidad vendida} = 8536,214 - 835,722^* \text{Precio} + 0,592^* \text{Publicidad}$. En otras palabras, por cada aumento de precio unitario, la cantidad vendida disminuye en 835,722 unidades. Por

cada aumento de publicidad unitaria, la cantidad vendida aumenta en 0,592 unidades. Esta información es valiosa.

También puede usar estos coeficientes para hacer un pronóstico. Por ejemplo, si el precio es de \$4 y la publicidad es de \$3000, podría alcanzar una cantidad vendida de $8536.214 - 835.722 * 4 + 0.592 * 3000 = 6970$.

Derechos residuales de autor

Los residuos muestran la distancia entre los puntos de datos reales y los puntos de datos predichos (usando la ecuación). Por ejemplo, el primer punto de datos es 8500. Usando la ecuación, el punto de datos predicho es $8536.214 - 835.722 * 2 + 0.592 * 2800 = 8523.009$, lo que da un residuo de $8500 - 8523.009 = -23.009$.

33	RESIDUAL OUTPUT		
34			
35	Observation	Predicted Quantity Sold	Residuals
36	1	8523.009	-23.009
37	2	4476.048	223.952
38	3	6265.938	-465.938
39	4	7160.883	239.117
40	5	6252.733	-52.733
41	6	7095.058	204.942
42	7	5726.330	-126.330
43			

También puede crear un diagrama de dispersión de estos residuos.

