

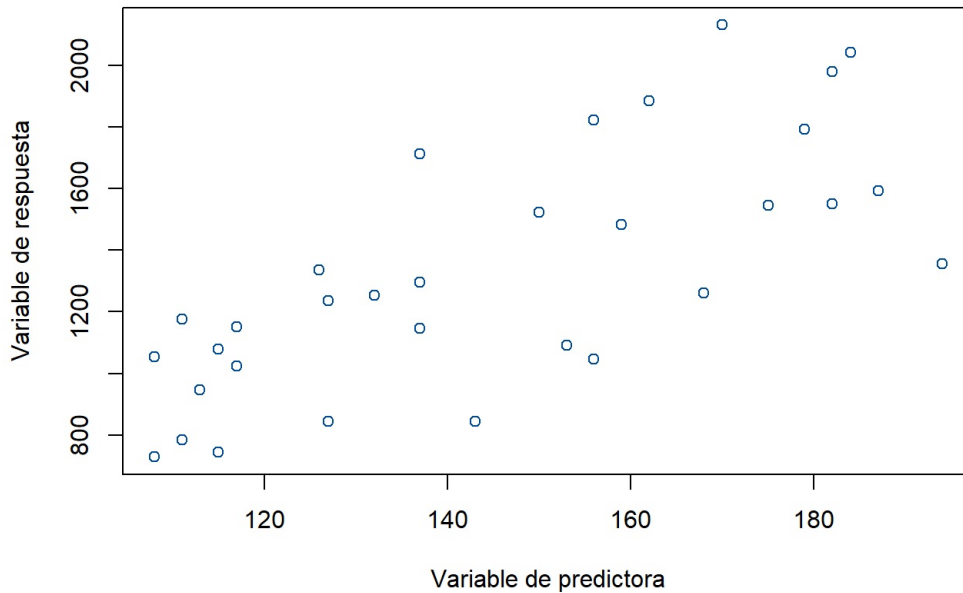
# Práctica 3

Agustin Riquelme y Heriberto Espino

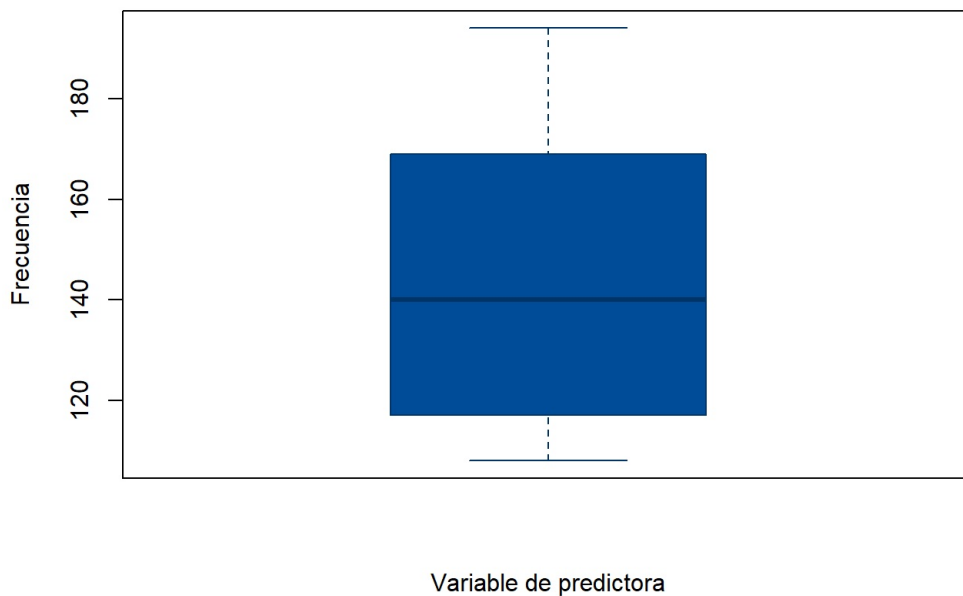
Con el objetivo de sintetizar las líneas de código lo más posible, se han realizado una funciones con los argumentos de la variable de predictora, la variable de respuesta y el nivel de significancia  $\alpha$ . Las funciones se usarán a lo largo del documento con el objetivo de visualizar los datos de manera más amigable, el código de la función se ocultará para el reporte pero permanecerá en el formato markdown para su correcta ejecución.

A continuación se presentará el análisis de regresión lineal sobre la predicción del precio del reloj con sus años de antigüedad. Los datos presentados incluyen valores relacionados con un modelo estadístico y las pruebas realizadas en relación con ese modelo.

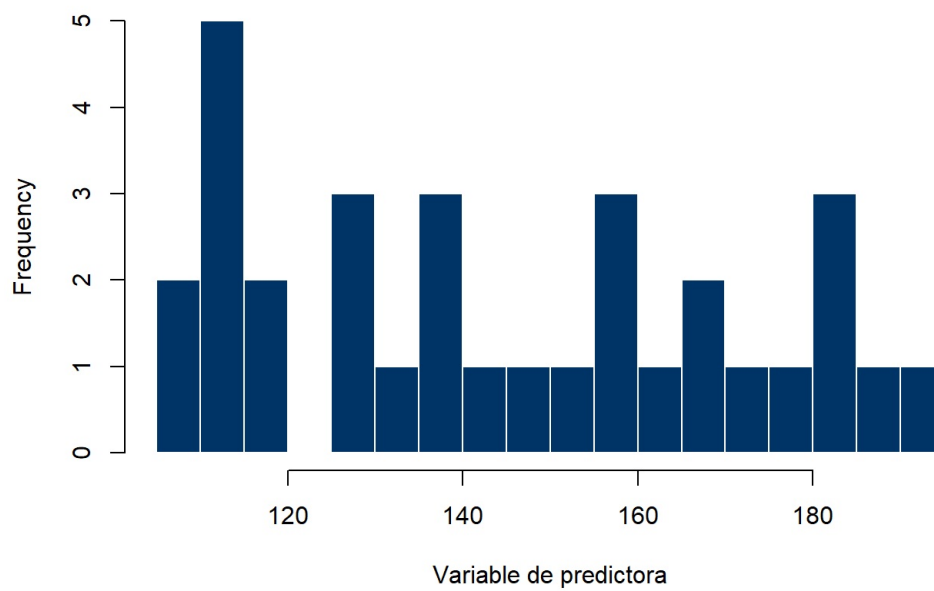
Gráfico dispersión



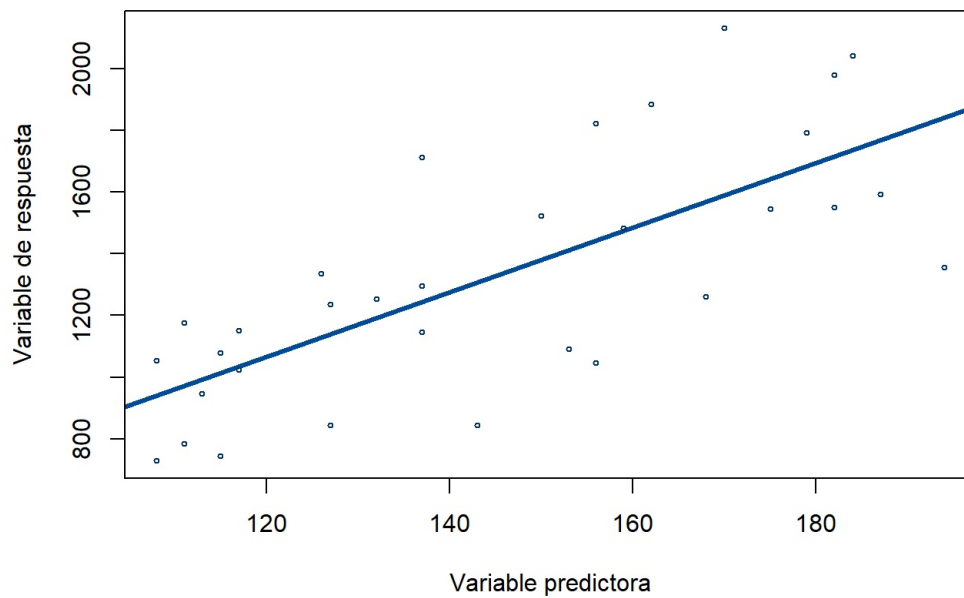
Boxplot x



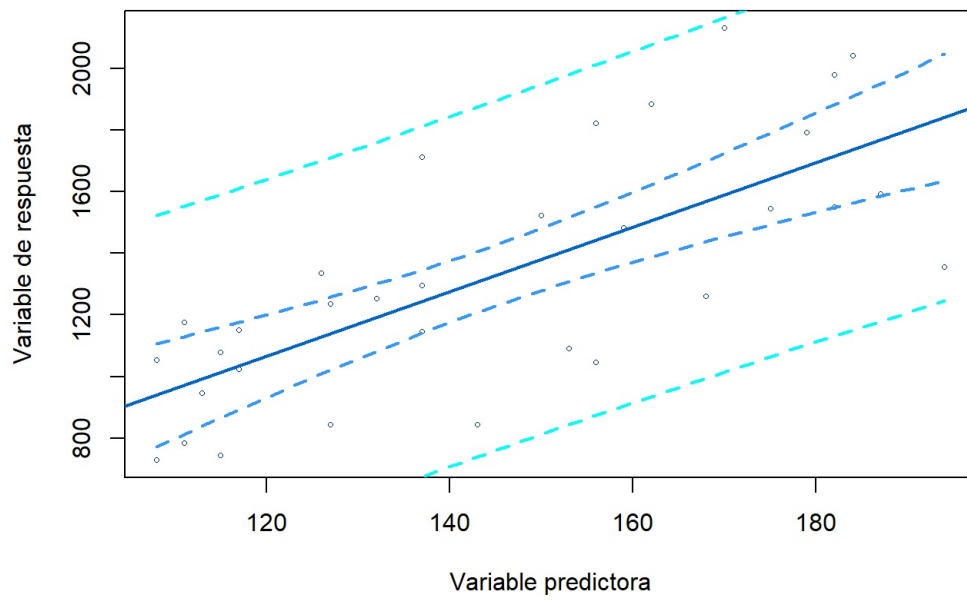
**Histograma x**



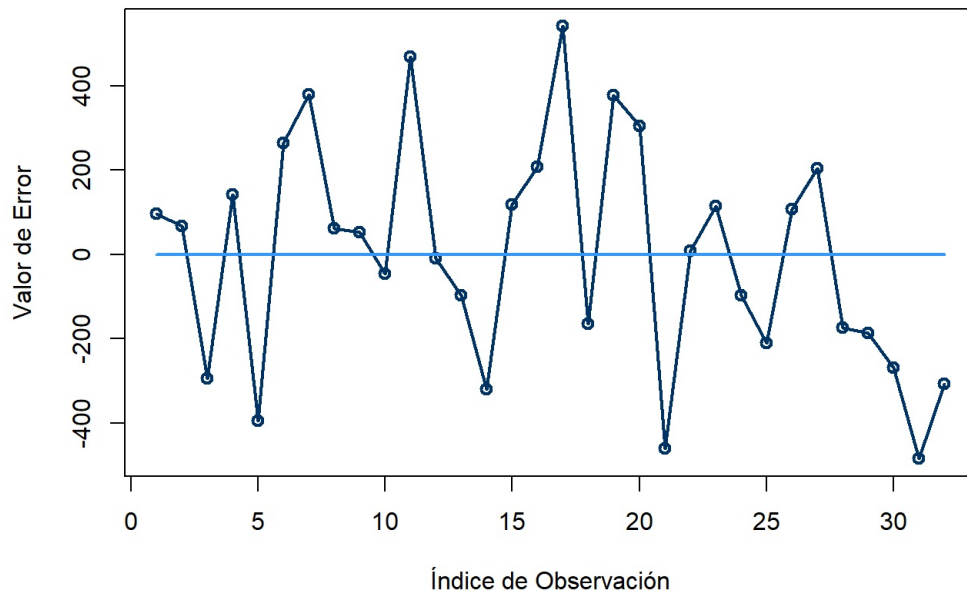
**Recta de regresión**



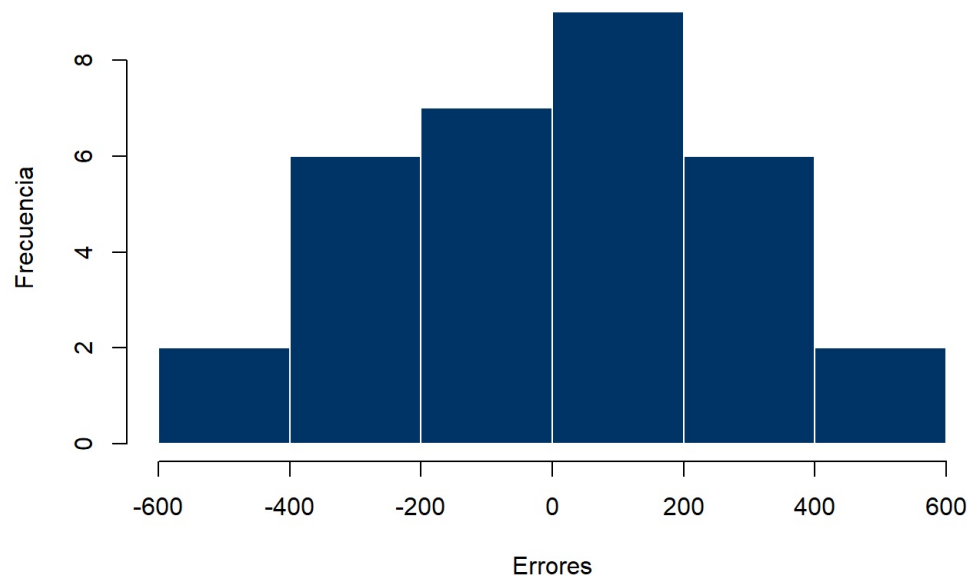
**Intervalos de confianza**



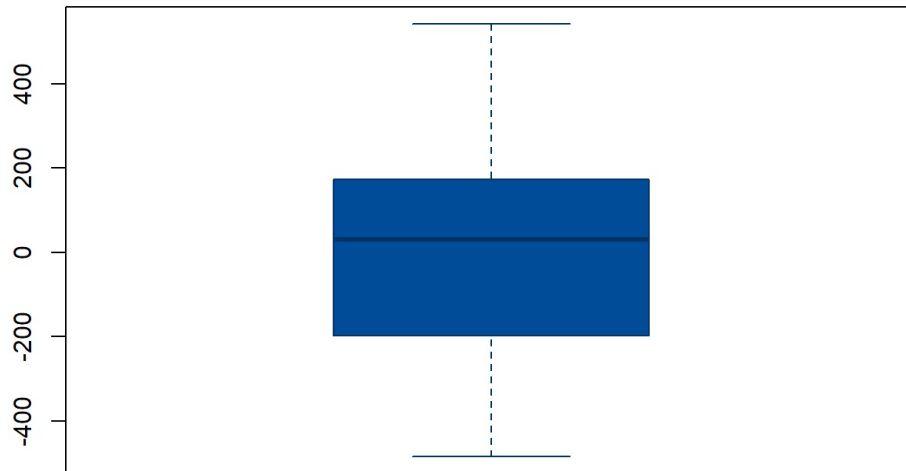
**Gráfico de Errores**



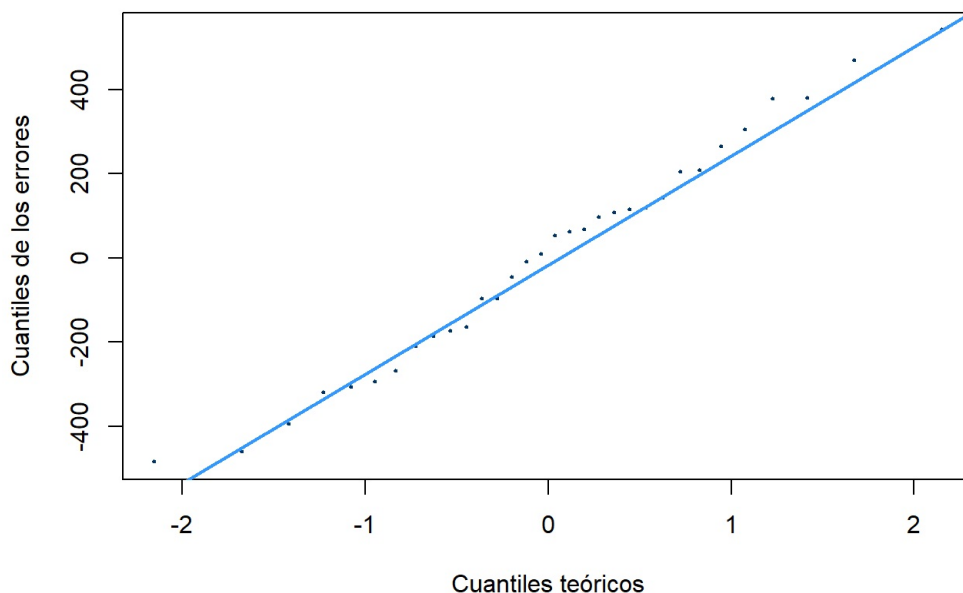
**Histograma de Errores**



**Boxplot de errores**



**Gráfico QQ de Errores**



**Informacion del modelo**

Datos	Valores
$\alpha$	0.05
$\beta_0$	-192.047448784969
$\beta_1$	10.4798444073133
S	273.530316237936
$S^2$	74818.8339012252
CV	20.6146258116202
df	30
qti	-2.04227245630124
qtd	2.04227245630124
tc	5.84398459480107
P valor	2.15808351730652e-06
Mín. $\beta_1$	6.81749772720417
Máx. $\beta_1$	14.1421910874224
Prom. err.	-5.55111512312578e-16
Var. error	72405.3231302179
Lillie P valor	0.88429697708903
Anderson P valor	0.924115811061527
Rho Pearson	0.729631263195947
Rho Spearman	0.759725090364974
dw	1.80461504164592

**Coefficiente de correlación**

Rho	Conclusión
Pearson	Hay una correlación fuerte
Spearman	Hay una correlación fuerte

**Inferencia para  $\beta_1$   $H_0: \beta_1 = 0$  vs  $H_1: \beta_1 \neq 0$**

Prueba.de.hipótesis	Conclusión
tc	Rechazamos $H_0$ . $\beta_1 \neq 0$
p valor	Rechazamos $H_0$ . $\beta_1 \neq 0$

Test de normalidad  $H_0$ : Los datos son normales vs  $H_1$ : Los datos no son normales

Test.de.normalidad	Valor
Normalidad Lillie	pvalor $\geq \alpha$ , no rechazamos $H_0$
Normalidad Anderson	pvalor $\geq \alpha$ , no rechazamos $H_0$

**Test Durbin-Watson**

Test	Conclusión
dw	No hay una correlación en los residuos

El valor  $\alpha$ , que es 0.05 en este caso, representa el nivel de significancia utilizado en las pruebas estadísticas. Este valor es crucial para determinar si los resultados obtenidos son estadísticamente significativos.

Los valores  $\beta_0$  y  $\beta_1$  son coeficientes de regresión.  $\beta_0$ , que es igual a -192.05, es el coeficiente de intercepción, mientras que  $\beta_1$ , con un valor de 10.48, es el coeficiente de pendiente en un modelo de regresión lineal. Estos coeficientes son esenciales para comprender cómo se relacionan las variables en el modelo y cómo una variable (en este caso,  $\beta_1$ ) afecta a la variable dependiente.

S y  $S^2$  representan la desviación estándar y la varianza, respectivamente. En este caso, la desviación estándar es de aproximadamente 273.53 y la varianza es de alrededor de 74818.83. Estos valores indican la dispersión de los datos y la variabilidad en torno a la media.

CV (coeficiente de variación) es una medida de la variabilidad relativa en comparación con la media y es igual a aproximadamente 20.61%. Cuanto mayor sea el valor del coeficiente de variación, mayor será la variabilidad en relación con la media.

El valor df (grados de libertad) es 30, lo que sugiere que se utilizaron 32 observaciones en el análisis estadístico, a las cuales se le restaron 2 grados de libertad por ser una prueba de dos colas.

qti y qtd son valores relacionados con una prueba de t-student bilateral. El valor tc representa el valor crítico de la prueba t, y el p-valor es extremadamente bajo ( $2.16e-06$ ), lo que sugiere que hay evidencia suficiente para rechazar la hipótesis nula en esta prueba. Esto indica que los coeficientes en el modelo de regresión son significativos.

Los valores Mín.  $\beta_1$  y Máx.  $\beta_1$  indican los valores mínimos y máximos que puede tomar el coeficiente  $\beta_1$  con un nivel de confianza determinado. Esto es útil para comprender la variabilidad en el coeficiente de pendiente.

Prom. err. (promedio de error) es prácticamente cero, lo que indica que el modelo parece estar bien ajustado a los datos.

Var. error (varianza del error) es de aproximadamente 72405.32, lo que sugiere la cantidad de variabilidad en los datos que no se explica mediante el modelo de regresión.

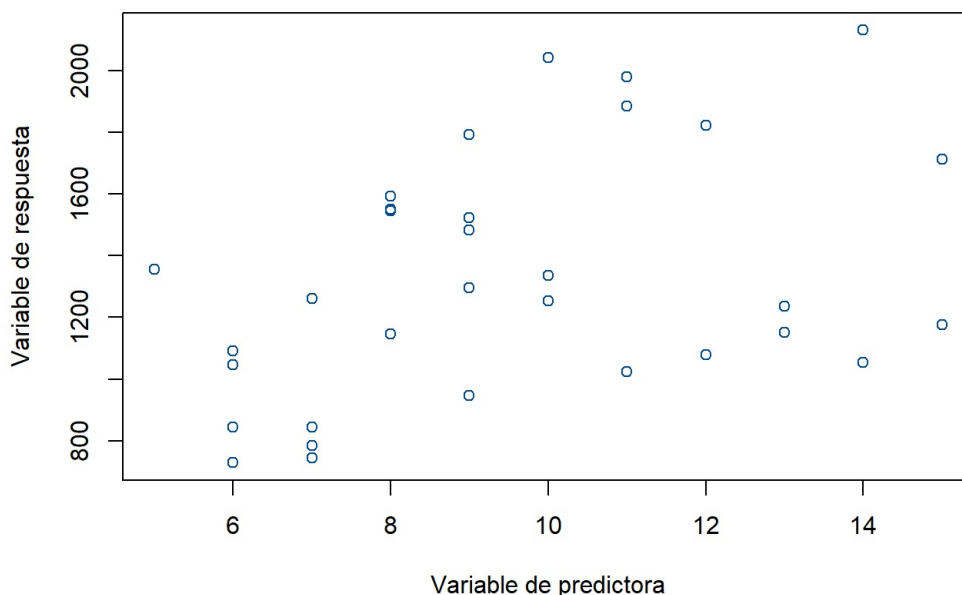
Los valores de Lillie y Anderson P son p-valores asociados con pruebas de normalidad. Un p-valor alto (0.88 y 0.92, respectivamente) sugiere que los datos se ajustan bien a una distribución normal.

Rho Pearson y Rho Spearman son coeficientes de correlación. Rho Pearson mide la correlación lineal y es 0.73, lo que indica una correlación positiva moderada. Rho Spearman mide la correlación de rango y es 0.76, lo que también sugiere una correlación positiva.

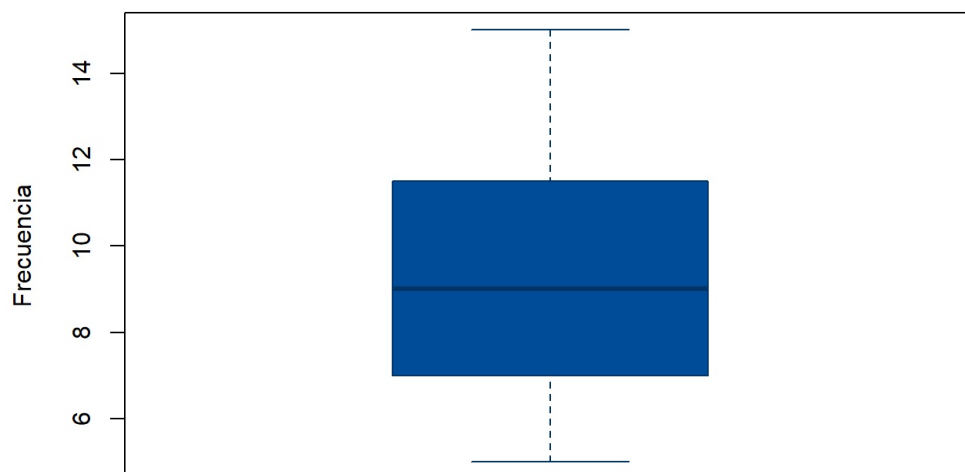
Finalmente, dw es el estadístico de Durbin-Watson, que se utiliza para detectar la presencia de autocorrelación en los residuos de un modelo de regresión. Un valor de 1.80 indica que es probable que no haya autocorrelación en los residuos.

A continuación se presentarán el análisis de regresión lineal sobre la predicción del precio del reloj con el número de postores. Los datos presentados incluyen valores relacionados con un modelo estadístico y las pruebas realizadas en relación con ese modelo.

### Gráfico dispersión

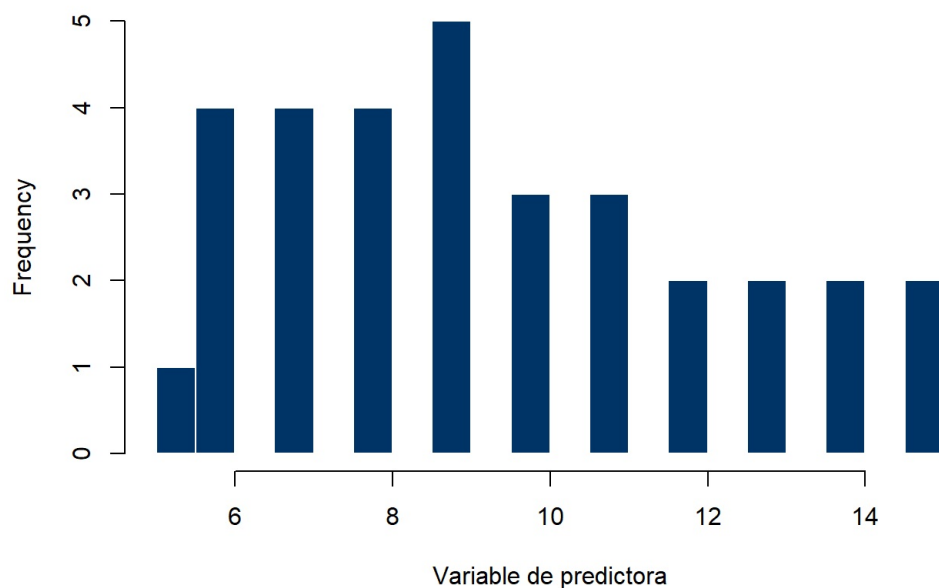


Boxplot x

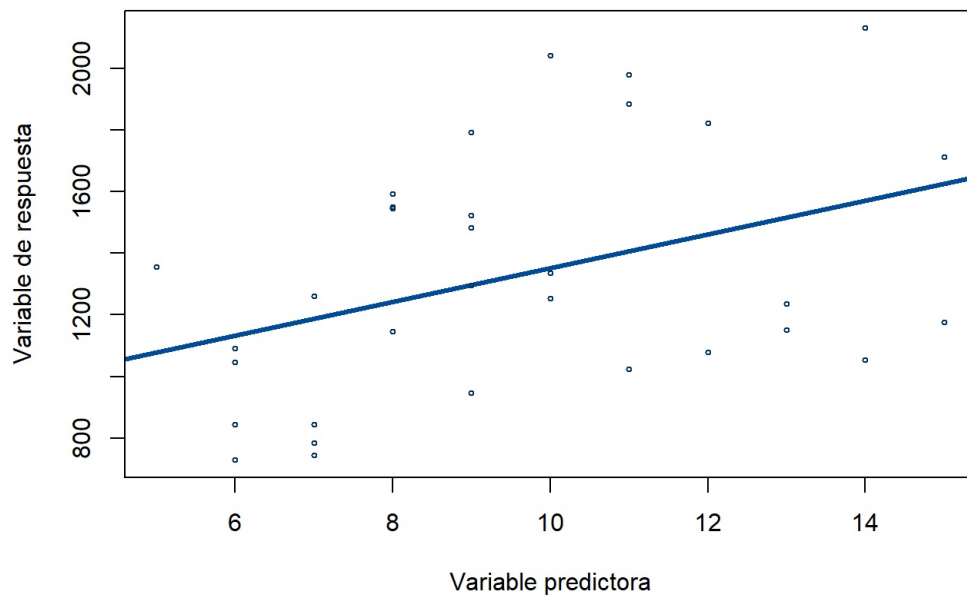


Variable de predictora

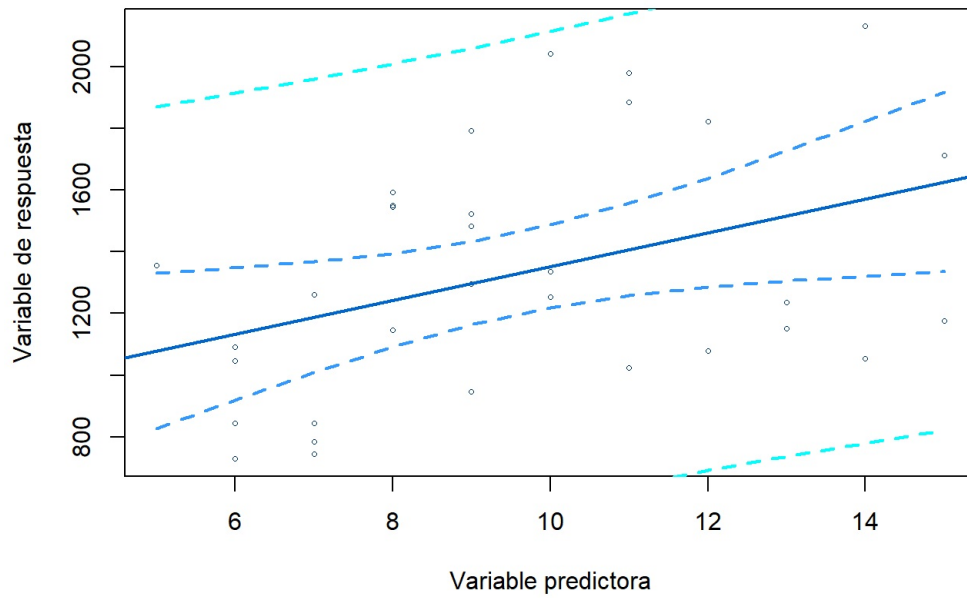
Histograma x



**Recta de regresión**

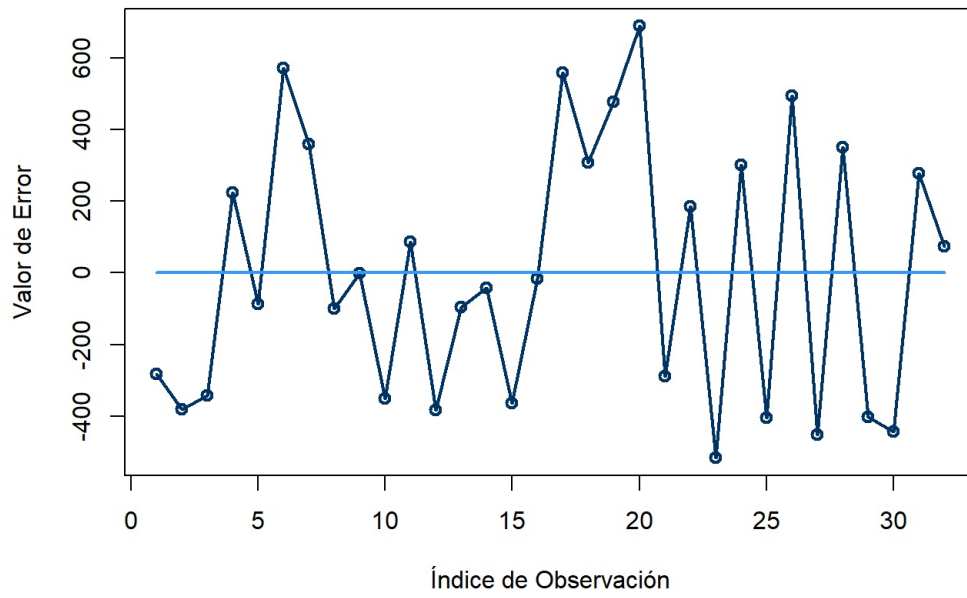


**Intervalos de confianza**

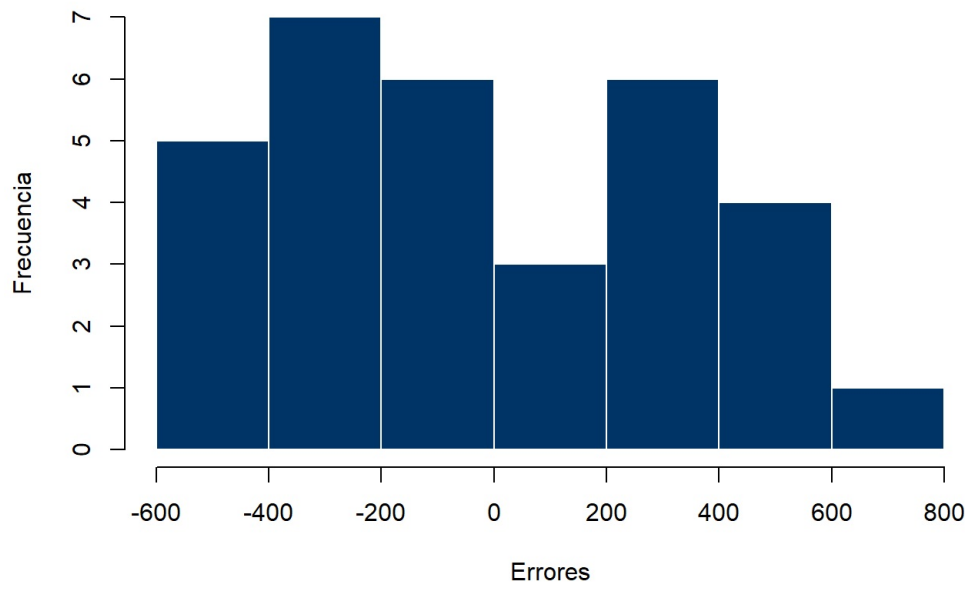




**Gráfico de Errores**



**Histograma de Errores**



Boxplot de errores

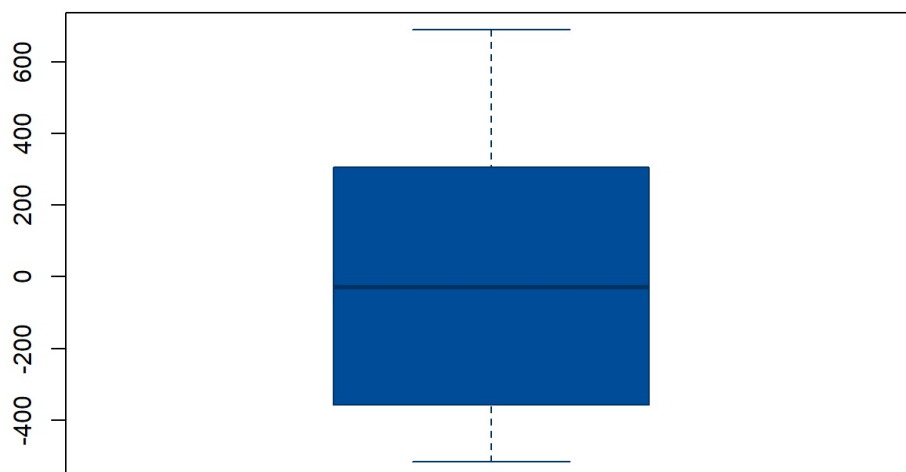
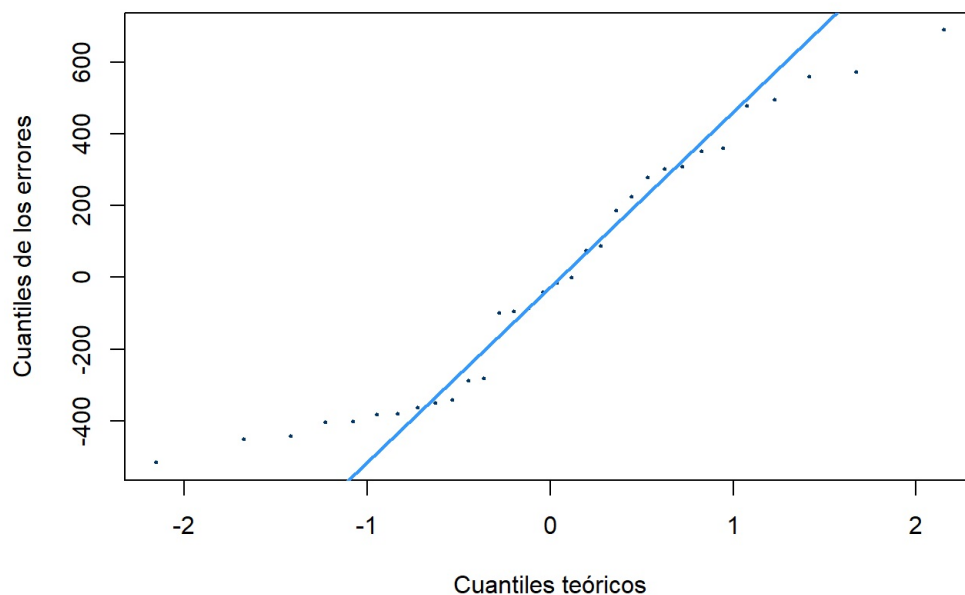


Gráfico QQ de Errores



Informacion del modelo

Datos	Valores
$\alpha$	0.05
$\beta_0$	804.911863982997
$\beta_1$	54.7633454181773
S	367.429232150101
$S^2$	135004.240638413
CV	27.6913222534226
df	30
qti	-2.04227245630124
qtd	2.04227245630124
tc	2.35645485940133
P valor	0.0251762709603326
Mín. $\beta_1$	7.30151031607473
Máx. $\beta_1$	102.22518052028
Prom. err.	-0.00000000000000555111512312578
Var. error	130649.265133948
Lillie P valor	0.0429165726900164
Anderson P valor	0.0447925109080511
Rho Pearson	0.395204366463315
Rho Spearman	0.408892074997163

## Coeficiente de correlación

Rho	Conclusión
Pearson	Hay una correlación baja
Spearman	Hay una correlación media

Inferencia para  $\beta_1$   $H_0: \beta_1 = 0$  vs  $H_1: \beta_1 \neq 0$

Prueba.de.hipótesis	Conclusión
tc	Rechazamos $H_0$ . $\beta_1 \neq 0$
p valor	No hay evidencia suficiente para rechazar $H_0$

Test de normalidad  $H_0$ : Los datos son normales vs  $H_1$ : Los datos no son normales

Test.de.normalidad	Valor
Normalidad Lillie	pvalor < $\alpha$ , rechazamos $H_0$ . No son normales
Normalidad Anderson	pvalor < $\alpha$ , rechazamos $H_0$ . No son normales

## Test Durbin-Watson

Test	Conclusión
dw	No hay una correlación en los residuos

En primer lugar, el valor  $\alpha$ , que es 0.05 en este caso, representa el nivel de significancia utilizado en las pruebas estadísticas. Este valor es fundamental para determinar si los resultados son estadísticamente significativos.

Los valores  $\beta_0$  y  $\beta_1$  son coeficientes de regresión.  $\beta_0$ , que tiene un valor de 804.91, es el coeficiente de intercepción, mientras que  $\beta_1$ , con un valor de 54.76, es el coeficiente de pendiente en un modelo de regresión lineal. Estos coeficientes son esenciales para comprender cómo se relacionan las variables en el modelo y cómo una variable (en este caso,  $\beta_1$ ) afecta a la variable dependiente.

S y  $S^2$  representan la desviación estándar y la varianza, respectivamente. En este caso, la desviación estándar es de aproximadamente 367.43 y la varianza es de alrededor de 135004.24. Estos valores indican la dispersión de los datos y la variabilidad en torno a la media.

CV (coeficiente de variación) es una medida de la variabilidad relativa en comparación con la media y es igual a aproximadamente 27.69%. Cuanto mayor sea el valor del coeficiente de variación, mayor será la variabilidad en relación con la media.

El valor df (grados de libertad) es 30, lo que sugiere que se utilizaron 32 observaciones en el análisis estadístico, a las cuales se le restaron 2 grados de libertad por ser una prueba de dos colas.

qti y qtd son valores relacionados con una prueba de t-student bilateral. El valor tc representa el valor crítico de la prueba t, y el p-valor es 0.0252, que es menor que el nivel de significancia  $\alpha$  (0.05). Esto sugiere que hay evidencia suficiente para rechazar la hipótesis nula en esta prueba. En otras palabras, los coeficientes en el modelo de regresión son significativos.

Los valores Mín.  $\beta_1$  y Máx.  $\beta_1$  indican los valores mínimos y máximos que puede tomar el coeficiente  $\beta_1$  con un nivel de confianza determinado. Esto es útil para comprender la variabilidad en el coeficiente de pendiente.

Prom. err. (promedio de error) es prácticamente cero, lo que indica que el modelo parece estar bien ajustado a los datos.

Var. error (varianza del error) es de aproximadamente 130649.27, lo que sugiere la cantidad de variabilidad en los datos que no se explica mediante el modelo de regresión.

Los valores de Lillie y Anderson P son p-valores asociados con pruebas de normalidad. Un p-valor bajo (0.0429 y 0.0448, respectivamente) sugiere que los datos no se ajustan bien a una distribución normal. Esto puede indicar la necesidad de considerar transformaciones o modelos diferentes.

Rho Pearson y Rho Spearman son coeficientes de correlación. Rho Pearson mide la correlación lineal y es 0.395, lo que indica una correlación positiva débil. Rho Spearman mide la correlación de rango y es 0.409, lo que también sugiere una correlación positiva media.

Finalmente, dw es el estadístico de Durbin-Watson, que se utiliza para detectar la presencia de autocorrelación en los residuos de un modelo de regresión. Un valor de 2.163 indica que puede haber autocorrelación en los residuos.

Si bien la prueba t indica que los coeficientes en el modelo son significativos, es importante tener en cuenta que los datos no se ajustan bien a una distribución normal, y la correlación entre las variables es débil.