

Diagnostico del Modelo de Regresión Lineal

Sergio Nava

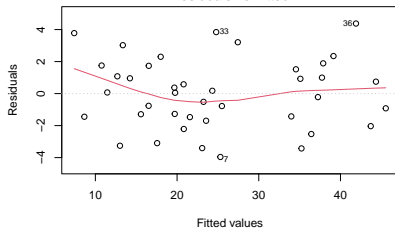
2023-03-23

Residuals vs Fitted.

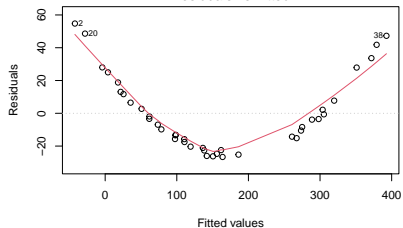
El gráfico Residuals vs Fitted es una herramienta útil para diagnosticar relaciones no lineales en los datos.

- ▶ Si los residuos se distribuyen por igual alrededor de una línea horizontal, esto indica que no existe una relación no lineal subyacente presente y que el modelo ha capturado todas las relaciones lineales con precisión.
- ▶ Si ve patrones o grupos distintos en su gráfico residual, podría indicar una relación no lineal subyacente entre las variables predictoras y la variable de resultado que su modelo no ha tenido en cuenta.
- ▶ Es importante investigar dichos patrones más a fondo antes de sacar conclusiones sobre lo que significan para su análisis.

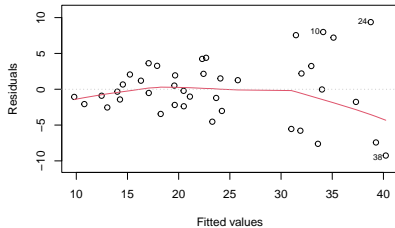
Modelo 1
Residuals vs Fitted



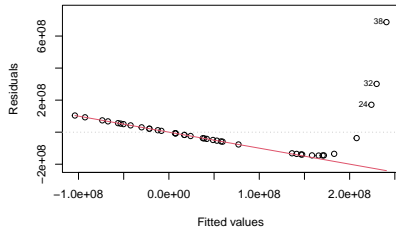
Modelo 2
Residuals vs Fitted



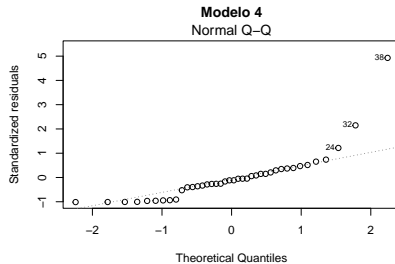
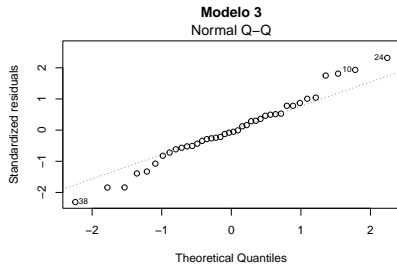
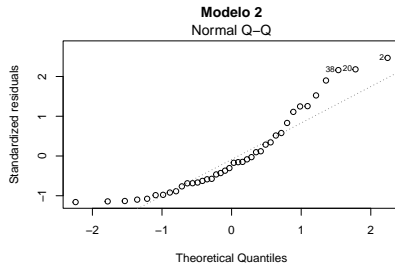
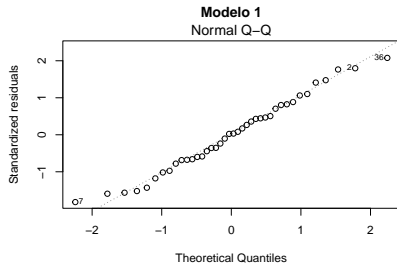
Modelo 3
Residuals vs Fitted



Modelo 4
Residuals vs Fitted



Normal Q-Q

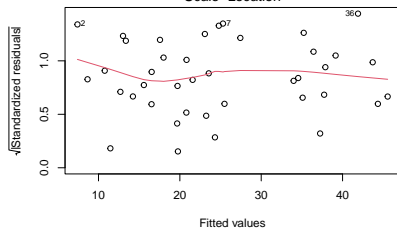


Scale-Location

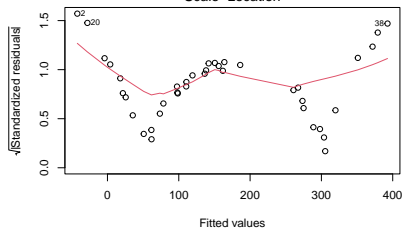
Los gráficos de *Scale-Location* se utilizan para evaluar el supuesto de igualdad de varianza en un modelo de regresión lineal.

- ▶ Este gráfico muestra si los residuos tienen una distribución uniforme en diferentes rangos de predictores, lo que se conoce como homocedasticidad.
- ▶ Al observar este gráfico, debe mostrar una línea horizontal con puntos distribuidos por igual (aleatoriamente) para obtener los mejores resultados.
- ▶ Si hay patrones o grupos visibles en el gráfico, entonces puede indicar que hay una variación desigual y se debe realizar una mayor investigación sobre por qué esto podría estar ocurriendo.

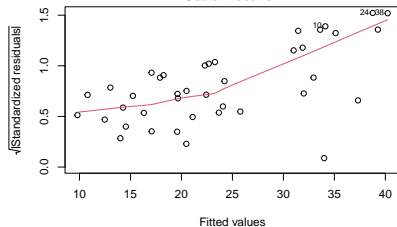
Modelo 1
Scale-Location



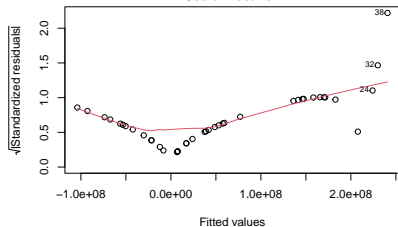
Modelo 2
Scale-Location



Modelo 3
Scale-Location



Modelo 4
Scale-Location



Residuals vs Leverage

Esta gráfica nos ayuda a encontrar casos influyentes (es decir, sujetos) si los hay.

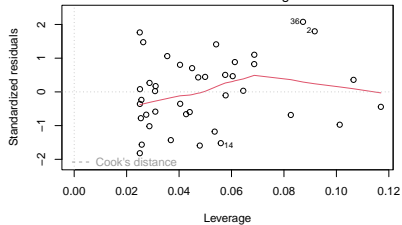
- ▶ No todos los valores atípicos influyen en el análisis de regresión lineal (cualquiera que sea el significado de los valores atípicos).
- ▶ Aunque los datos tienen valores extremos, es posible que no sean influyentes para determinar una línea de regresión. Eso significa que los resultados no serían muy diferentes si los incluyéramos o excluyéramos del análisis.
- ▶ Siguen la tendencia en la mayoría de los casos y realmente no importan; no son influyentes.
- ▶ Por otro lado, algunos casos podrían ser muy influyentes incluso si parecen estar dentro de un rango razonable de valores. Pueden ser casos extremos contra una línea de regresión y pueden alterar los resultados si los excluimos del análisis.

A diferencia de las otras tramas, esta vez los patrones no son relevantes.

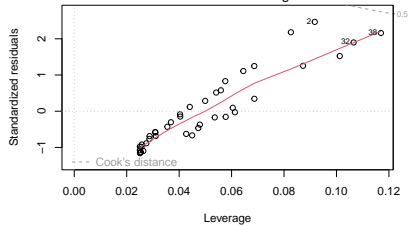
- ▶ Estamos atentos a los valores atípicos en la esquina superior derecha o en la esquina inferior derecha.
- ▶ Busque casos fuera de una línea discontinua, la distancia de Cook.
- ▶ Cuando los casos están fuera de la distancia de Cook (lo que significa que tienen puntajes altos de distancia de Cook), los casos influyen en los resultados de la regresión.
- ▶ Los resultados de la regresión se verán alterados si excluimos esos casos.

Modelo 1

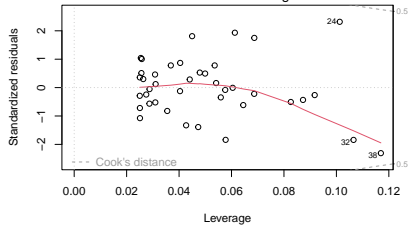
Residuals vs Leverage

**Modelo 2**

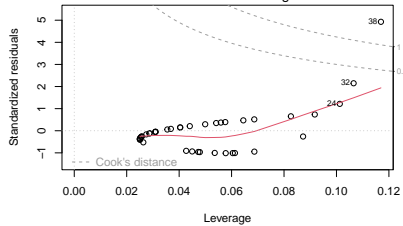
Residuals vs Leverage

**Modelo 3**

Residuals vs Leverage

**Modelo 4**

Residuals vs Leverage



8e+08

08