

9

Tarea 2: Estadística y Diseño de Experimentos

Canales Torres Jose Carlos 2203070368

1. Identifique quién es la variable de respuesta Y y quién es la variable regresora X y escriba el modelo de regresión.

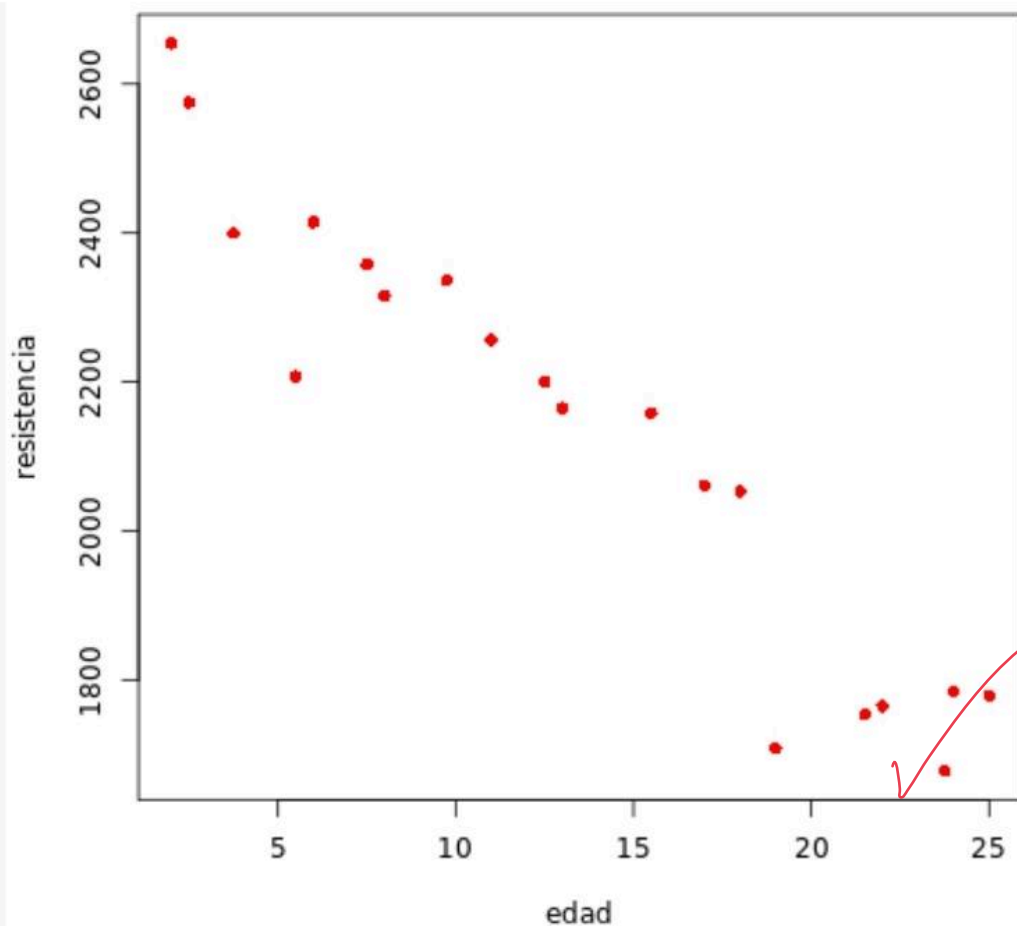
R= Variable respuesta Y: resistencia

Variable regresora X: edad.

Modelo de regresión lineal simple: $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$

2. Grafique el diagrama de dispersión de los datos.

Salida del programa



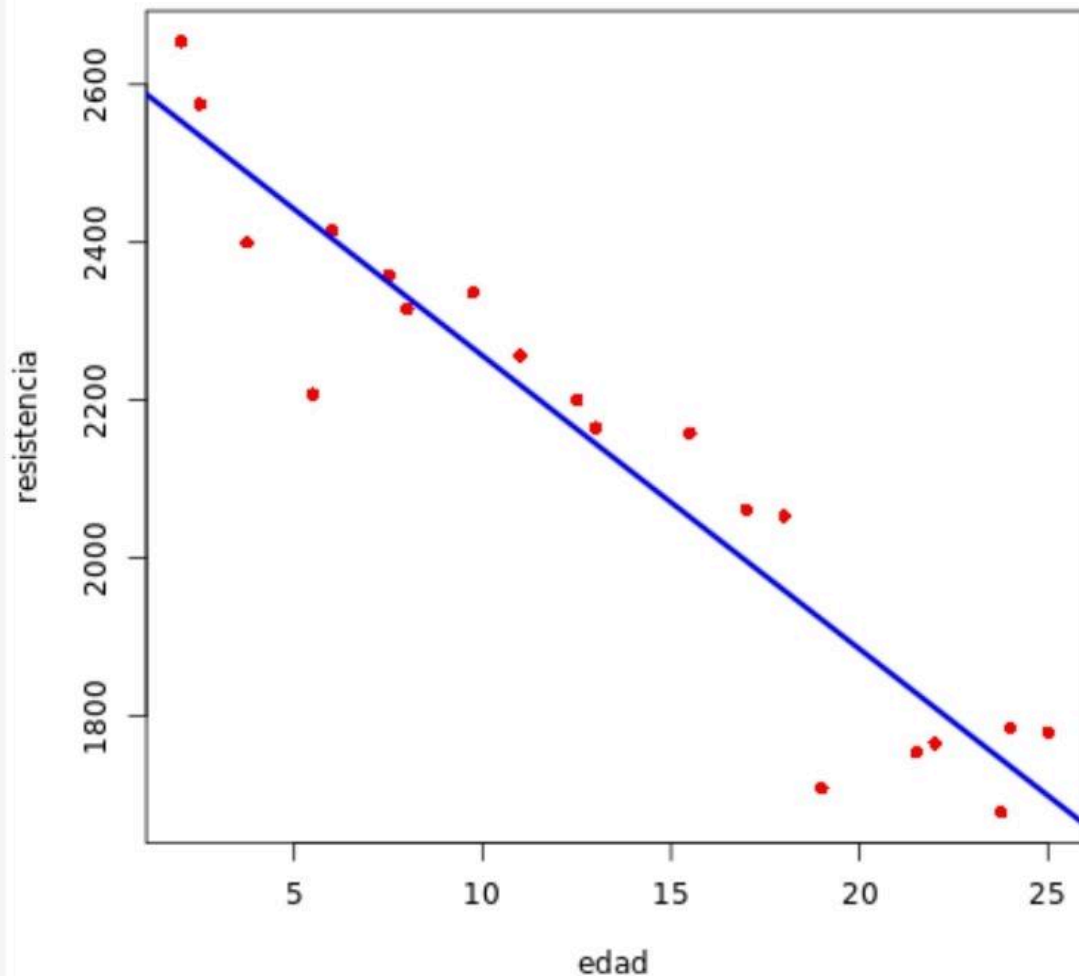
3. Obtenga los estimadores para β_0 y β_1 y escriba la ecuación de la recta ajustada.

R= $\beta_0 = 2672.822$, $\beta_1 = -37.154$

La ecuación de la recta ajustada es: Resistencia = $2672.822 - 37.154 (X)$

4. Grafique la recta de la regresión junto con los datos. ¿Qué tan bueno cree que es el ajuste?

Salida del programa



La recta parece ajustarse razonablemente los datos. La tendencia decreciente es clara y los puntos no se alejan demasiado de la recta, lo que sugiere un buen ajuste lineal.

5. Efectúe la prueba de significancia de la regresión para un nivel $\alpha = 0.05$. Escriba el valor del p-valor. ¿Qué conclusiones puede hacer sobre β_1 ?

R= Hipotesis nula: $H_0: \beta_1 = 0$
Hipotesis alternativa $H_1: \beta_1 \neq 0$

$\beta_1 = -37.154$

p-valor a β_1 : 1.64×10^{-10} , como p-valor es mucho menor que 0.05, se rechaza hipotesis nula

¿ Conclusión ?

Salida del programa

```
(Intercept)      edad
2627.82236    -37.15359

Call:
lm(formula = resistencia ~ edad)

Residuals:
    Min       1Q   Median       3Q      Max
-215.98  -50.68   28.74   66.61  106.76

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2627.822     44.184   59.48  < 2e-16 ***
edad        -37.154       2.889  -12.86 1.64e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 96.11 on 18 degrees of freedom
Multiple R-squared:  0.9018,    Adjusted R-squared:  0.8964
F-statistic: 165.4 on 1 and 18 DF,  p-value: 1.643e-10
```

6. Suponga que se tienen tres lotes del propulsor tipo 1, con 5, 10 y 15 semanas de edad respectivamente. ¿Cuál es la estimación para la resistencia según el modelo de regresión (para cada lote)?

R= Ecuacion estimada: $RESISTENCIA = 2672.822 - 37.154 * (X)$; Donde X es valores de semanas de edad

X=5

$$2672.822 - 37.154 * (5) = 2442.05$$

X=10

$$2672.822 - 37.154 * (10) = 2256.28$$

X=15

$$2672.822 - 37.154 * (15) = 2070.51$$

7. Calcule el valor del coeficiente de determinación (Multiple R Squared). Según este coeficiente, ¿qué tan bueno es el ajuste de la regresión?

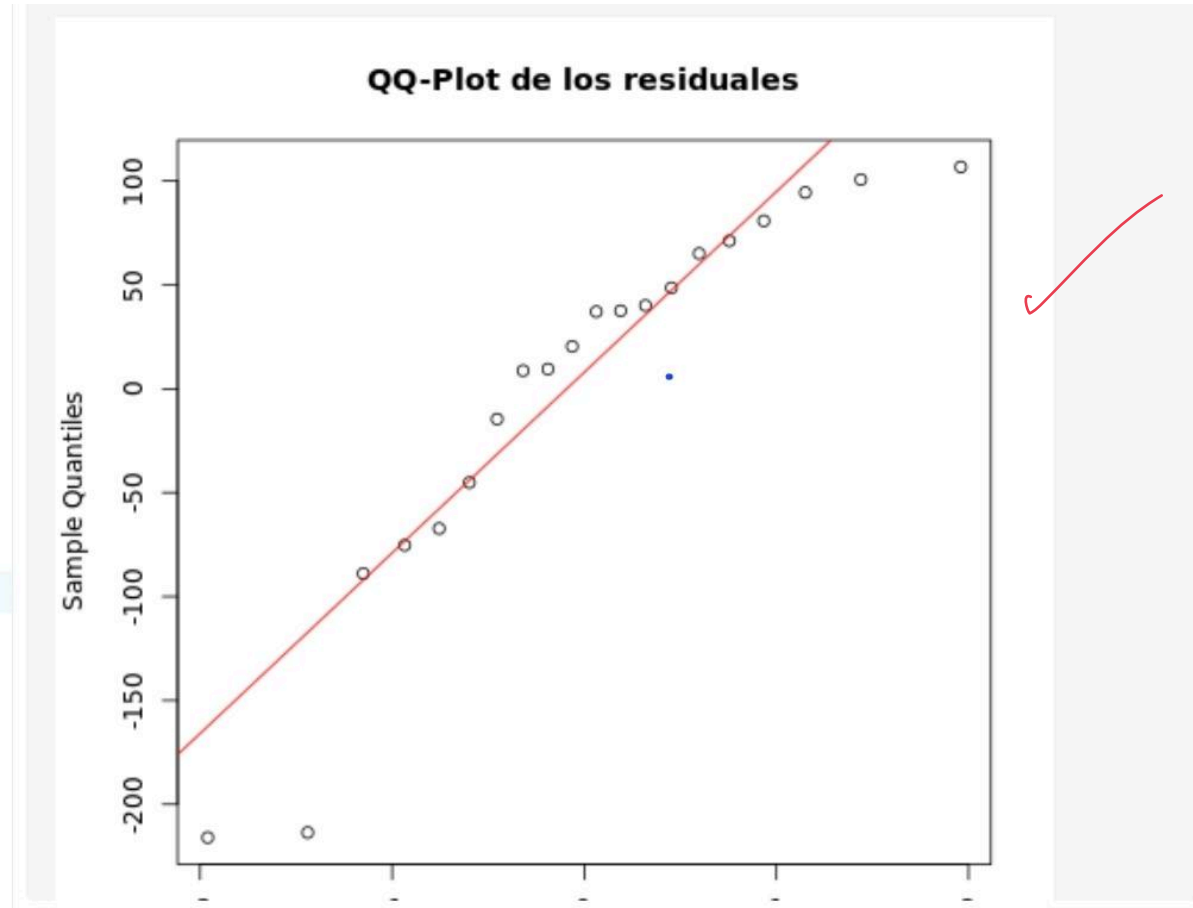
$R = R^2 = 0.9018$: El 90.18 % de la variación en la resistencia se explica por la edad.

Esto indica un excelente ajuste lineal.

8. Verificación de supuestos del modelo. (Obtenga primero los residuales)

a) Normalidad. Grafique los residuales contra los cuantiles de una normal (`qqnorm``, `qqline``). ¿Se satisface este supuesto?

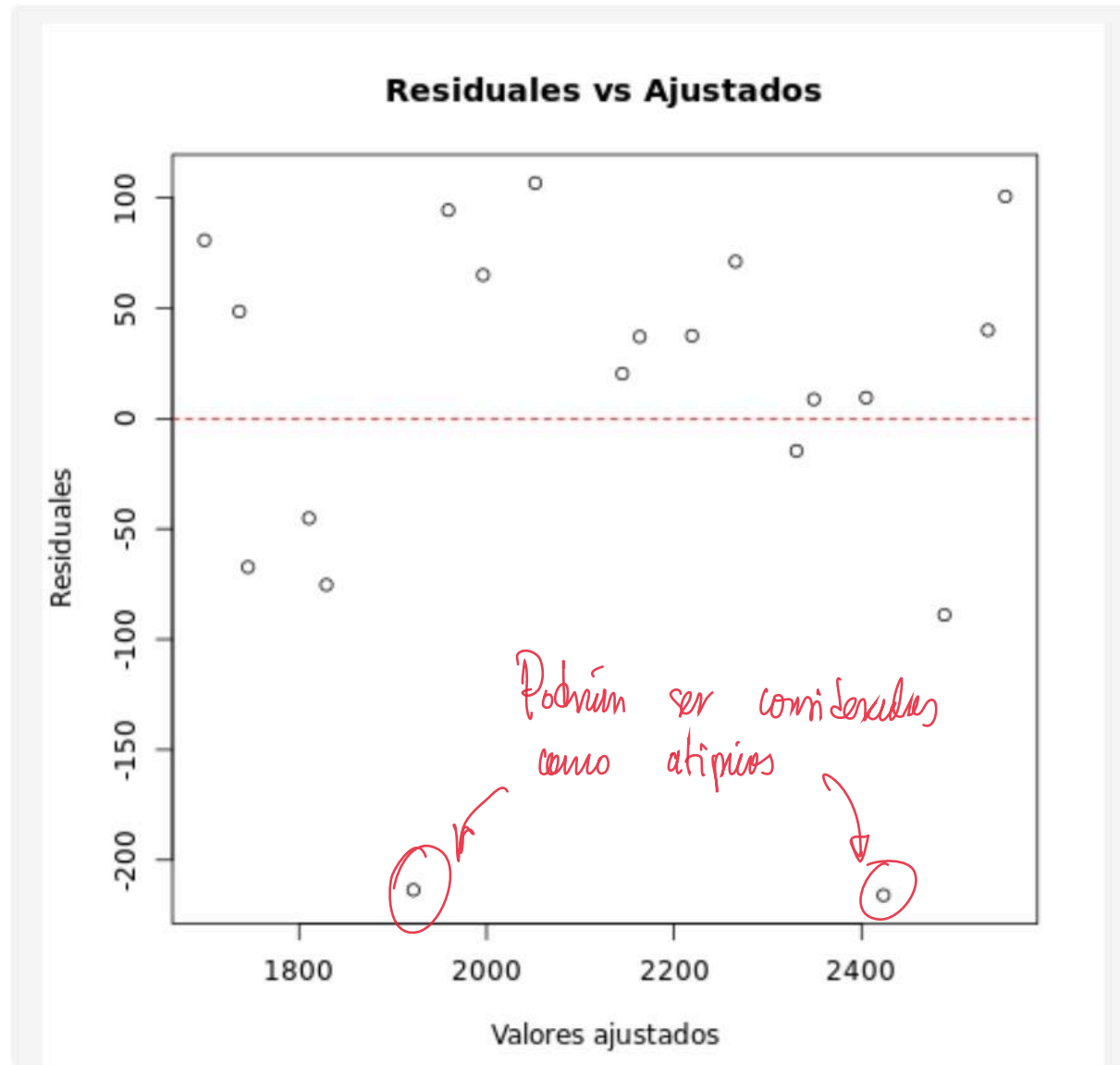
R= Los puntos deberían alinearse con la línea, se cumple la normalidad aproximadamente.



b) Media cero, varianza constante e independencia. Grafique los residuales contra los predichos para verificar los tres supuestos (`predict`, `rstudent`). ¿Observa alguna anomalía?**

R= Los residuales se distribuyen aleatoriamente no hay patrones visibles.
Se cumplen los supuestos de media cero, homocedasticidad e independencia.

Salida del programa

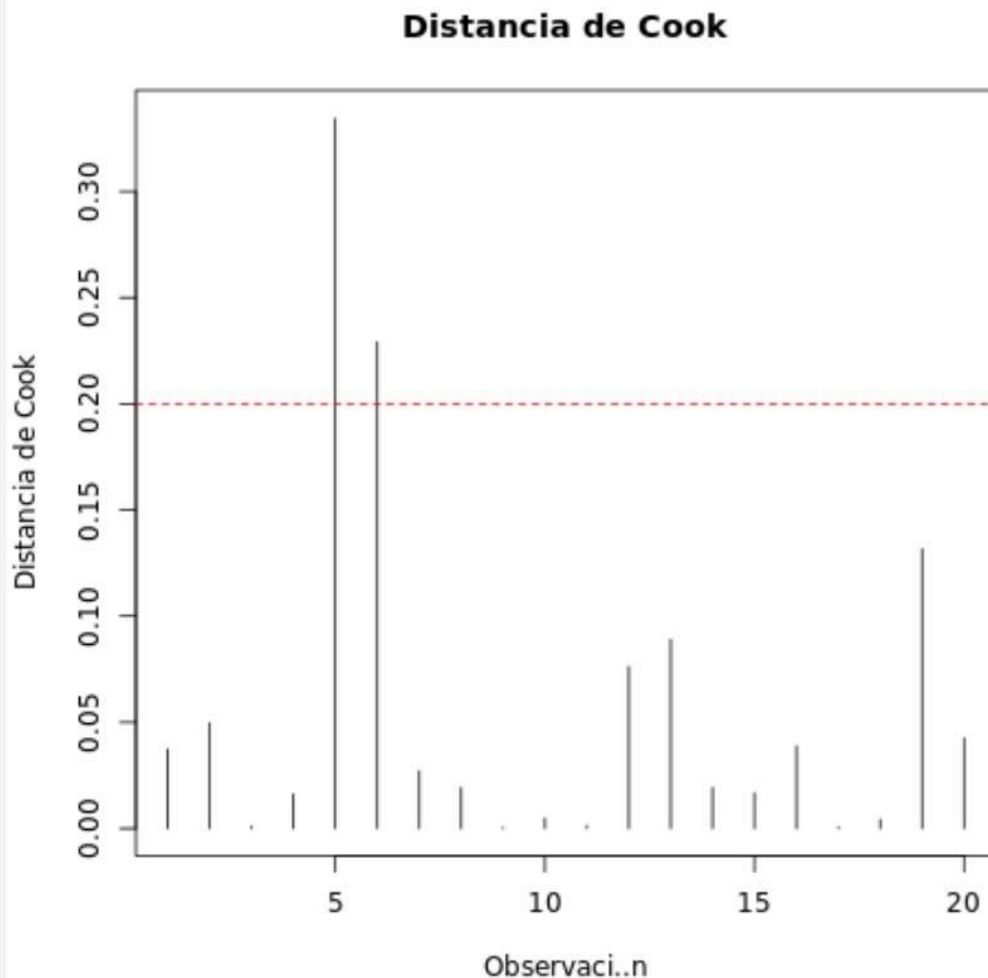


9. Puntos atípicos e influyentes.

a) Utilizando la gráfica anterior, ¿se observan puntos que puedan considerarse como atípicos (*outliers*)? R= Con la gráfica de residuales no se observan valores extremos → sin outliers evidentes

b) Utilizando la distancia de Cook, verifique si hay puntos influyentes. R= Ningún valor excede 1 no hay puntos influyentes.

Salida del programa



10. Escriba una conclusión general para este problema.

Existe una relación lineal negativa y significativa entre la edad del propulsor y la resistencia al corte.

Por cada semana adicional de edad, la resistencia disminuye en promedio 37.15 psi.

El modelo explica aproximadamente el 90% de la variación en la resistencia, lo cual indica un ajuste excelente.

Además, los supuestos del modelo se cumplen y no se identifican observaciones influyentes.