8.5

Tarea 2.

Lugo Barragán Miriam.

En la fabricación de un motor se deben unir dos tipos de propulsores (tipo 1 y tipo 2). Se sospecha que la resistencia al corte de esta unión está relacionada con la edad (en semanas) del lote de propulsores del tipo 1. En la siguiente tabla se muestra la Resistencia al corte (medida en psi) y la Edad (en semanas) del lote del propulsor tipo 1.

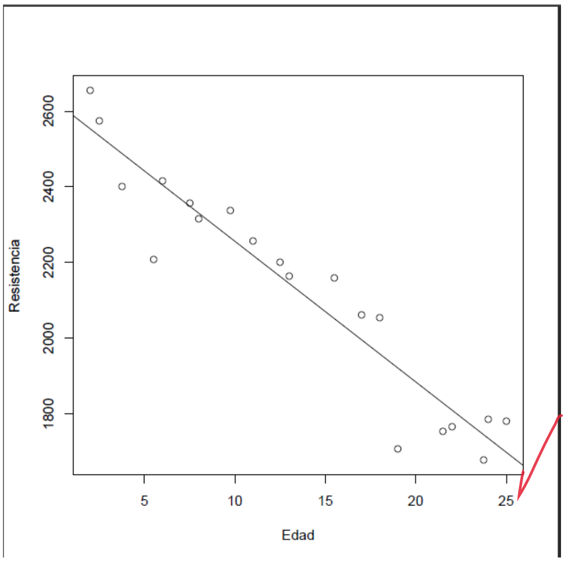
Resistencia	Edad
2158.70	15.50
1678.15	23.75
2316.00	8.00
2061.30	17.00
2207.50	5.50
1708.30	19.00
1784.70	24.00
2575.00	2.50
2357.90	7.50
2256.70	11.00
2165.20	13.00
2399.55	3.75
1779.80	25.00
2336.75	9.75
1765.30	22.00
2053.50	18.00
2414.40	6.00
2200.50	12.50
2654.20	2.00
1753.70	21.50

Cuadro 1: Resistencia obtenida según la edad del propulsor tipo 1.

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

```
> # =====================================================
> # REGRESIÓN LINEAL SIMPLE: Resistencia vs Edad
> # =====================================================
>
> # 1. Identificar variables
> # Variable de respuesta Y: Resistencia
> # Variable regresora X: Edad
> # Modelo: Resistencia = β0 + β1 * Edad
>
> Resistencia <- c(2158.7,1678.15,2316,2061.3,2207.5,1708.3,1784.7,2575,2357.9,2256.7,
+                  2165.2,2399.55,1779.8,2336.75,1765.3,2053.5,2414.4,2200.5,2654.2,1753.7)
> Edad <- c(15.5,23.75,8,17,5.5,19,24,2.5,7.5,11,13,3.75,25,9.75,22,18,6,12.5,2,21.5)
>
> datos <- data.frame(Resistencia, Edad)
>
```

2. Grafique el diagrama de dispersión de los datos.

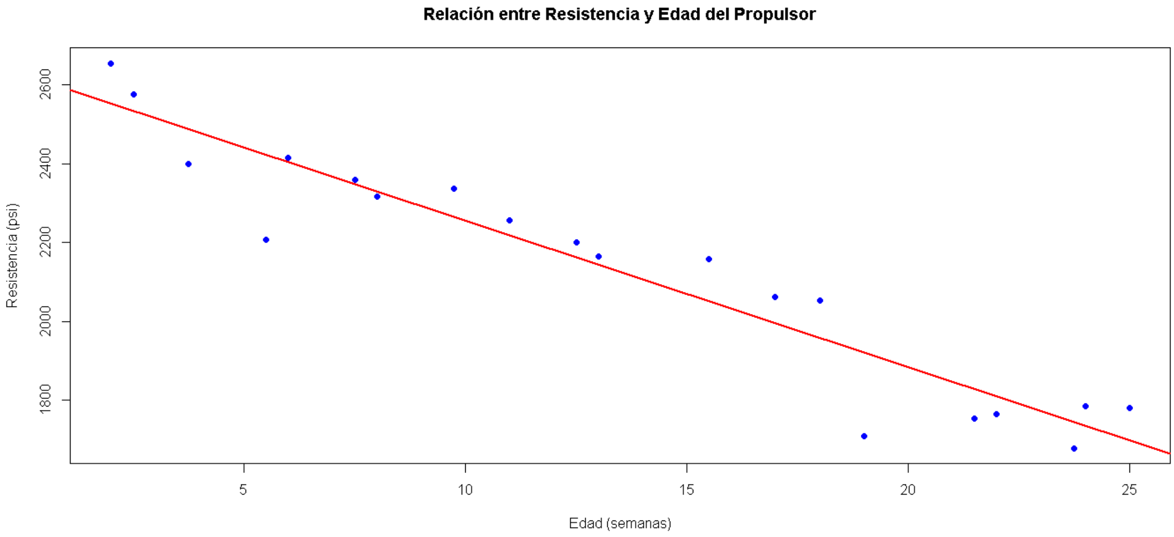


3. Obtenga los estimadores para β_0 y β_1 y escriba la ecuación de la recta ajustada.

```
> summary(RL)

Call:
lm(formula = Resistencia ~ Edad, data = datos)

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

> 
> cat("Ecuación del modelo:\n")
Ecuación del modelo:
> cat("Ŷ =", round(coef(RL)[1],2), "+", round(coef(RL)[2],2), "* Edad\n\n")
Ŷ = 2627.82 + -37.15 * Edad
```

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

Falta responder pregunta



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

¿Cuál es el valor? Faltan conclusiones

```
> # ====================================================
> # 4. Prueba de significancia de β1
> # ====================================================
> cat("Prueba de significancia del modelo:\n")
Prueba de significancia del modelo:
> pvalor <- summary(RL)$coefficients[2,4]
> cat("p-valor =", round(pvalor,5), "\n")
p-valor = 0
> if (pvalor < 0.05) {
+   cat("Conclusión: Se rechaza H0. β1 es significativo. Existe relación lineal$
+ } else {
+   cat("Conclusión: No se rechaza H0. No hay evidencia de relación lineal sign$
+ }
```

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)?

```
>
> # ====================================================
> # 5. Predicciones para edades 5, 10 y 15 semanas
> # ====================================================
> nuevas_edades <- data.frame(Edad = c(5, 10, 15))
> pred <- predict(RL, nuevas_edades, interval = "prediction")
> cat("Estimaciones de Resistencia (con intervalos de predicción):\n")
Estimaciones de Resistencia (con intervalos de predicción):
> print(cbind(nuevas_edades, round(pred,2)))
  Edad     fit     lwr     upr
1    5 2442.05 2229.02 2655.09
2   10 2256.29 2048.38 2464.19
3   15 2070.52 1863.38 2277.65
```

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

```
>
> # =====================================================
> # 6. Coeficiente de determinación R²
> # =====================================================
> R2 <- summary(RL)$r.squared
> cat("\nCoeficiente de determinación (R²):", round(R2,4), "\n")

Coeficiente de determinación (R²): 0.9018
> cat("El modelo explica aproximadamente", round(R2*100,1), "% de la variabilidad en la resistencia.\n\n")
El modelo explica aproximadamente 90.2 % de la variabilidad en la resistencia.

>
```

8. VERIFICACIÓN DE SUPUESTOS DEL MODELO. (Obtenga primero los residuales)

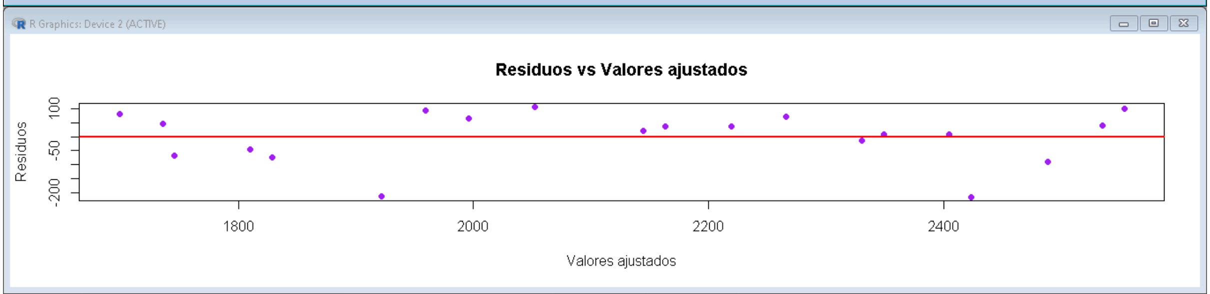
a) NORMALIDAD. Grafique los residuales contra los cuantiles de una normal (qqnorm, qqline). ¿Se satisface este supuesto?

b) MEDIA CERO, VARIANZA CONSTANTE E INDEPENDENCIA. Grafique los residuales contra los predichos para verificar los tres supuestos (predict, rstudent). ¿Observa alguna anomalía?

R Console

```
> # a) Normalidad de los residuales
> par(mfrow = c(1,2))
> qqnorm(residuales, main = "QQ Plot de los residuales")
> qqline(residuales, col = "red")
> hist(residuales, main = "Histograma de residuales", col = "lightblue", xlab = "Residuos")
>
> # b) Media cero, varianza constante e independencia
> par(mfrow = c(1,1))
> plot(ajustados, residuales,
+      main = "Residuos vs Valores ajustados",
+      xlab = "Valores ajustados",
+      ylab = "Residuos",
+      pch = 19, col = "purple")
> abline(h = 0, col = "red", lwd = 2)
> grid()
```

Falta gráfica de supuestos de normalidad

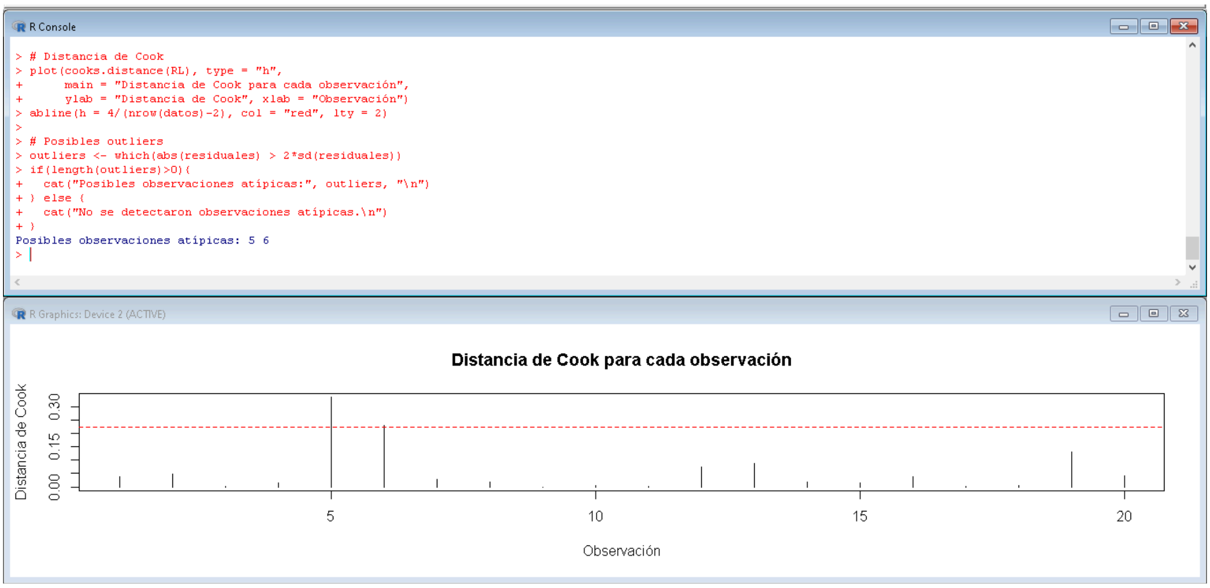


9. PUNTOS ATÍPICOS E INFLUYENTES.

a) Utilizando la gráfica anterior, ¿se observan puntos que puedan considerarse como atípicos (outliers)?

b) Utilizando la distancia de Cook, verifique si hay puntos inlfuyentes.

Hay que responder a las preguntas



10. Escriba una conclusión general para este problema.

Las gráficas presentadas en este trabajo, nos ayudan a estimar de mejor manera los datos y mostrarnos cómo se pueden plasmar en gráficos diferentes dependiendo de lo que el problema se va desarrollando.