9.5

Datos del problema

```
R <- c(2158.70, 1678.15, 2316.00, 2061.30, 2207.50, 1708.30, 1784.70, 2575.00, 2357.90, 2256.70, 2165.20, 2399.55,
1779.80, 2336.75, 1765.30, 2053.50, 2414.40, 2200.50, 2654.20, 1753.70)
```

```
Edad <- c(15.50, 23.75, 8.00, 17.00, 5.50, 19.00, 24.00, 2.50, 7.50, 11.00, 13.00, 3.75, 25.00, 9.75, 22.00, 18.00, 6.00, 12.50, 2.
```

```
# Crear el data frame (para que el codigo capte cual es la renglon y la columna)
datos <- data.frame(Edad, R)
```

Pregunta 1

X (Variable regresora) esta dada por la EDAD porque la resistencia "depende de" la edad, entonces Y (Variable de respuesta) es la RESISTENCIA

Modelo de regresion

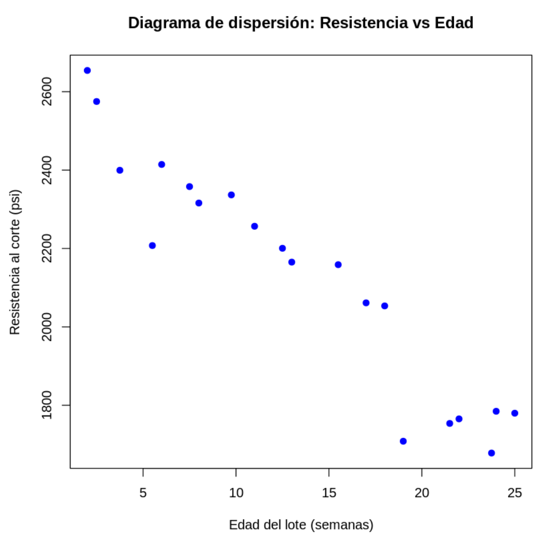
$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$

Donde:

- Y_i resistencia al corte del i-ésimo motor
- X_i edad del lote (en semanas)
- β_0 valor esperado de Y cuando X=0
- β_1 cambio promedio en la resistencia por cada semana adicional de edad
- ϵ_i error

Pregunta 2

```
#Diagrama de dispersión
plot(datos$Edad, datos$R,
     main = "Diagrama de dispersión: Resistencia vs Edad",
     xlab = "Edad del lote (semanas)",
     ylab = "Resistencia al corte (psi)",
     pch = 19,          # tipo de punto sólido
     col = "blue")      # color
```



Pregunta 3

```
#Ajustamos el modelo
modelo <- lm(R ~ Edad, data = datos)
#Le decimos que nos muestre los resultados
summary(modelo)
```

```
Call:
lm(formula = R ~ 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
```

```
# Mostrar solo los coeficientes para la estimación de B_0 y B_1
coef(modelo)
```

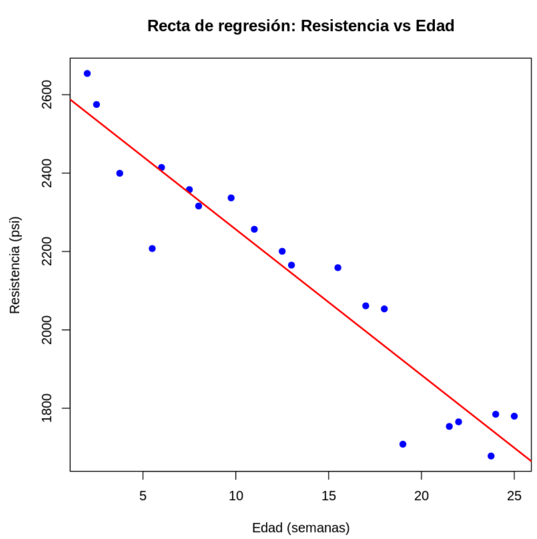
(Intercept): 2627.8223590013 Edad: -37.1535909449052

Por lo tanto $\hat{Y} = 2627.8223 - 37.1535X$

Pregunta 4

```
plot(datos$Edad, datos$R,
     main = "Recta de regresión: Resistencia vs Edad",
     xlab = "Edad (semanas)", ylab = "Resistencia (psi)",
     pch = 19, col = "blue")

abline(modelo, col = "red", lwd = 2)
```



A simple vista podemos decir que la recta de regresion es muy acertiva

Pregunta 5

El siguiente paso es efectuar la prueba de significancia para el valor $\alpha = 0.05$ y escribir el valor de p

```
# Dado que ya calculamos el valor de p solo falta extraer el dato para `$B_1$`
summary(modelo)$coefficients["Edad", "Pr(>|t|)"]
```

1.64334381811888e-10

Conclusión:

Con un valor $p = 1.64 \times 10^{-10}$, muy inferior al nivel de significancia α = 0.05, se rechaza la hipótesis nula de que la pendiente β_1 sea igual a cero. Por tanto, existe evidencia estadísticamente significativa de que la edad del lote del propulsor tipo 1 influye en la resistencia al corte de la unión.

Pregunta 6

```
# Valores de edad para los que queremos predecir
nuevas_edades <- data.frame(Edad = c(5, 10, 15))

# Predicciones usando el modelo de regresión
predicciones <- predict(modelo, newdata = nuevas_edades)

# Mostrar resultados
predicciones
```

1: 2442.05440427677 2: 2256.28644955224 3: 2070.51849482772

Pregunta 7

El coeficiente de determinación R^2 ya se calculo anteriormente con "summary", entonces lo unico que haria falta seria extraerlo

```
# Calcular el coeficiente de determinación
R2 <- summary(modelo)$r.squared

# Mostrar el valor
R2
```

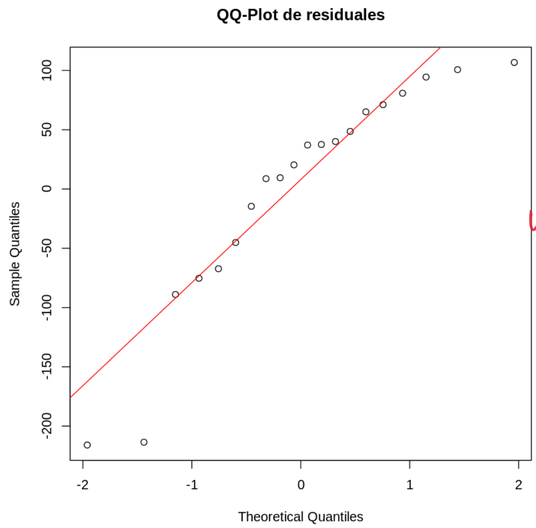
0.901841431676304

Pregunta 8

```
# Residuales y valores predichos
residuales <- resid(modelo)           # residuales simples
residuales_estandarizados <- rstudent(modelo)  # residuales estandarizados
valores_predichos <- predict(modelo)  # valores predichos por el modelo
```

Inciso a)

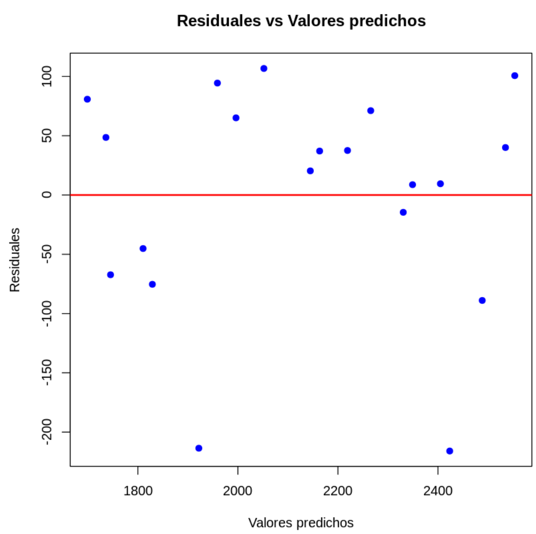
```
# QQ-plot para normalidad
qqnorm(residuales, main = "QQ-Plot de residuales")
qqline(residuales, col = "red")
```



En el QQ-plot de los residuales, los puntos se alinean de forma aproximada con la línea de referencia, indicando que los residuales siguen una distribución aproximadamente normal. Por tanto, el supuesto de normalidad se considera satisfecho.

Inciso b)

```
# Residuales vs valores predichos
plot(valores_predichos, residuales,
     main = "Residuales vs Valores predichos",
     xlab = "Valores predichos",
     ylab = "Residuales",
     pch = 19, col = "blue")
abline(h = 0, col = "red", lwd = 2)
```



En el gráfico de residuales vs valores predichos, los puntos se distribuyen aleatoriamente alrededor de la línea roja sin mostrar patrones sistemáticos ni cambios de dispersión. Por lo tanto, los supuestos de media cero, varianza constante e independencia de los errores se cumplen adecuadamente.

Pregunta 9

a) Hay algunos puntos más alejados, especialmente en la parte inferior (alrededor de -200 en residuales), que podrían considerarse posibles outliers, aunque no parecen extremos al grado de afectar drásticamente el modelo.

b)Probablemente no haya puntos influyentes porque lo mas problables es que todsas las distancias de Cook van a ser bajas

Pregunta 10

Hay que calcular las distancias de Cook.

Conclusión:

En conclusión, el análisis estadístico permite afirmar que la resistencia al corte de la unión entre los propulsores tipo 1 y tipo 2 depende significativamente de la edad del lote de los propulsores tipo 1. El modelo de regresión lineal ajustado cumple con los supuestos teóricos y explica de manera adecuada el comportamiento observado, por lo que puede considerarse una herramienta útil para predecir la resistencia al corte en función de la edad del material, contribuyendo así al control y mejora del proceso de fabricación del motor.