

Tarea 9 - ANOVA

Héctor Hibran Tapia Fernández - A01661114

2024-08-28

Problema 1.

Resuelve las dos partes del problema “El rendimiento”. Se encuentra en los apoyos de clase de “ANOVA”. Para ello se te recomienda que sigas los siguientes pasos:

El rendimiento

En un instituto se han matriculado 36 estudiantes. Se desea explicar el rendimiento de ciencias naturales en función de dos variables: género y metodología de enseñanza. La metodología de enseñanza se analiza en tres niveles: explicación oral y realización del experimento (1er nivel) explicación oral e imágenes (2º nivel) y explicación oral (tercer nivel).

En los alumnos matriculados había el mismo número de chicos que de chicas, por lo que formamos dos grupos de 18 sujetos; en cada uno de ellos, el mismo profesor aplicará a grupos aleatorios de 6 estudiantes las 3 metodologías de estudio. A fin de curso los alumnos son sometidos a la misma prueba de rendimiento. Los resultados son los siguientes:

Chicos			Chicas		
Método 1	Método 2	Método 3	Método 1	Método 2	Método 3
10	5	2	9	8	2
7	7	6	7	3	6
9	6	3	8	5	2
9	6	5	8	6	1
9	8	5	10	7	4
10	4	3	6	7	3

¿Existe alguna influencia de la metodología de enseñanza y el género de los estudiantes en el rendimiento de los estudiantes?

1. Análisis Exploratorio.

- Calcula la media para el rendimiento por método de enseñanza, puedes consultar el código en R en los apoyos de clase de “ANOVA”.

```
rendimiento = c(10, 7, 9, 9, 9, 10, 5, 7, 6, 6, 8, 4, 2, 6, 3, 5, 5, 3, 9, 7, 8, 8, 10,
6, 8, 3, 5, 6, 7, 7, 2, 6, 2, 1, 4, 3)
metodos = factor(rep(c("Método 1", "Método 2", "Método 3"), each = 6, times = 2))
medias = tapply(rendimiento, metodos, mean)
print(medias)
```

```
## Método 1 Método 2 Método 3
##      8.5      6.0      3.5
```

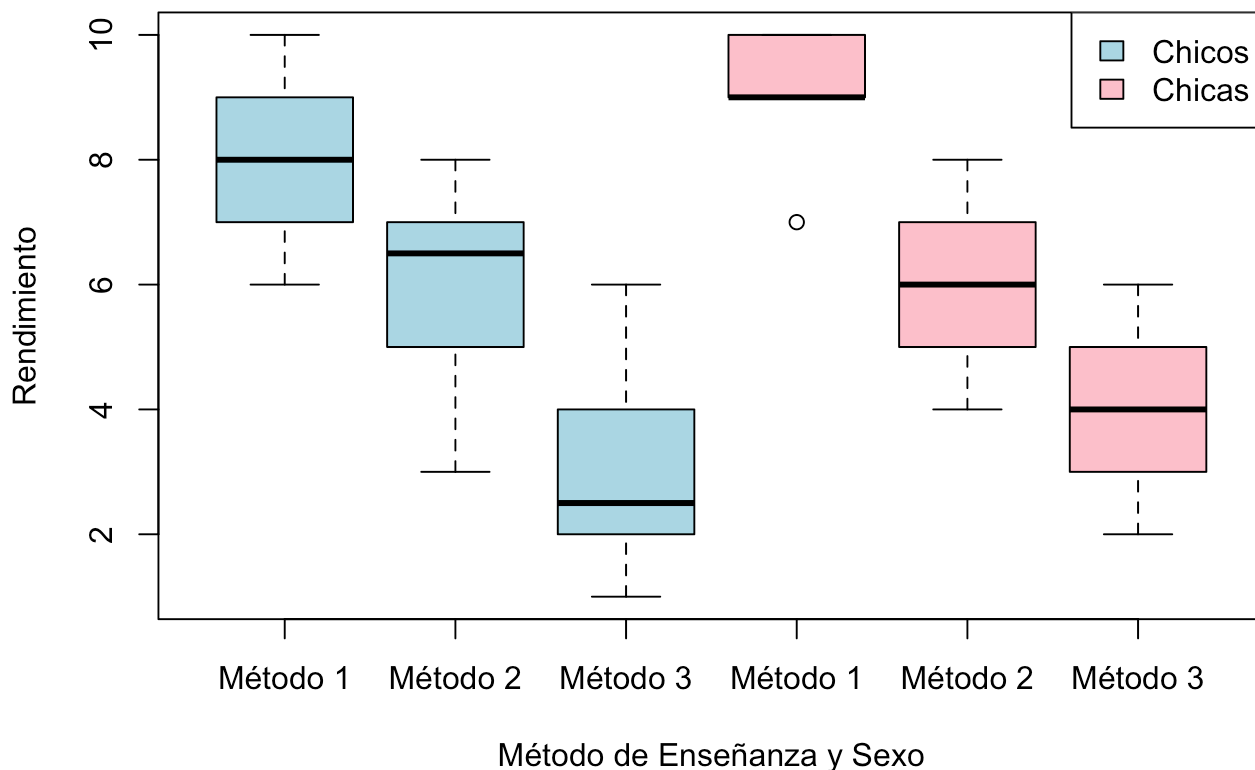
- Haz el boxplot de la evaluación de los estudiantes por método de enseñanza y sexo.

```
sexo = factor(rep(c("Chicos", "Chicas"), each = 18))

boxplot(rendimiento ~ metodos + sexo, main = "Evaluación de los Estudiantes", xlab = "Método de Enseñanza y Sexo", ylab = "Rendimiento", col = c("lightblue", "lightblue", "lightblue", "pink", "pink", "pink"), names = c("Método 1", "Método 2", "Método 3", "Método 1", "Método 2", "Método 3"))

legend("topright", legend = c("Chicos", "Chicas"), fill = c("lightblue", "pink"))
```

Evaluación de los Estudiantes



- Interpreta el resultado desde la perspectiva estadística y en el contexto del problema. Escribe tus conclusiones parciales.

Parece que tanto chicos como chicas tienen un rendimiento mejor y más consistente con los Métodos 1 y 2 en comparación con el Método 3.

Ahora hablando del Método 1 se muestra que es más consistente el rendimiento de los chicos y una variabilidad moderada para las chicas.

El Método 3 podría ser el menos efectivo para ambos sexos, ya que muestra menores rendimientos y alta variabilidad.

2. Las Hipótesis.

- Establece las hipótesis estadísticas (tienen que ser 3).

1. Hipótesis I:

- **Hipótesis Nula (H_{01}):** No hay diferencias significativas en el rendimiento promedio de los estudiantes entre los diferentes métodos de enseñanza ($H_{01} : \mu_{\text{Método 1}} = \mu_{\text{Método 2}} = \mu_{\text{Método 3}}$).
- **Hipótesis Alternativa (H_{11}):** Hay al menos una diferencia significativa en el rendimiento promedio de los estudiantes entre los métodos de enseñanza ($H_{11} : \text{Al menos una de las medias } (\mu_{\text{Método 1}}, \mu_{\text{Método 2}}, \mu_{\text{Método 3}}) \text{ es diferente}$).

2. Hipótesis II:

- **Hipótesis Nula (H_{02}):** No hay diferencias significativas en el rendimiento promedio entre chicos y chicas ($H_{02} : \mu_{\text{Chicos}} = \mu_{\text{Chicas}}$).
- **Hipótesis Alternativa (H_{12}):** Hay una diferencia significativa en el rendimiento promedio entre chicos y chicas ($H_{12} : \mu_{\text{Chicos}} \neq \mu_{\text{Chicas}}$).

3. Hipótesis III:

- **Hipótesis Nula (H_{03}):** $H_{03} : \text{No hay interacción entre método de enseñanza y género.}$
- **Hipótesis Alternativa (H_{13}):** $H_{13} : \text{Existe una interacción entre método de enseñanza y género.}$

3. Realiza el ANOVA para dos niveles con interacción:

- Consulta el código en R en los apoyos de clase de “ANOVA”.

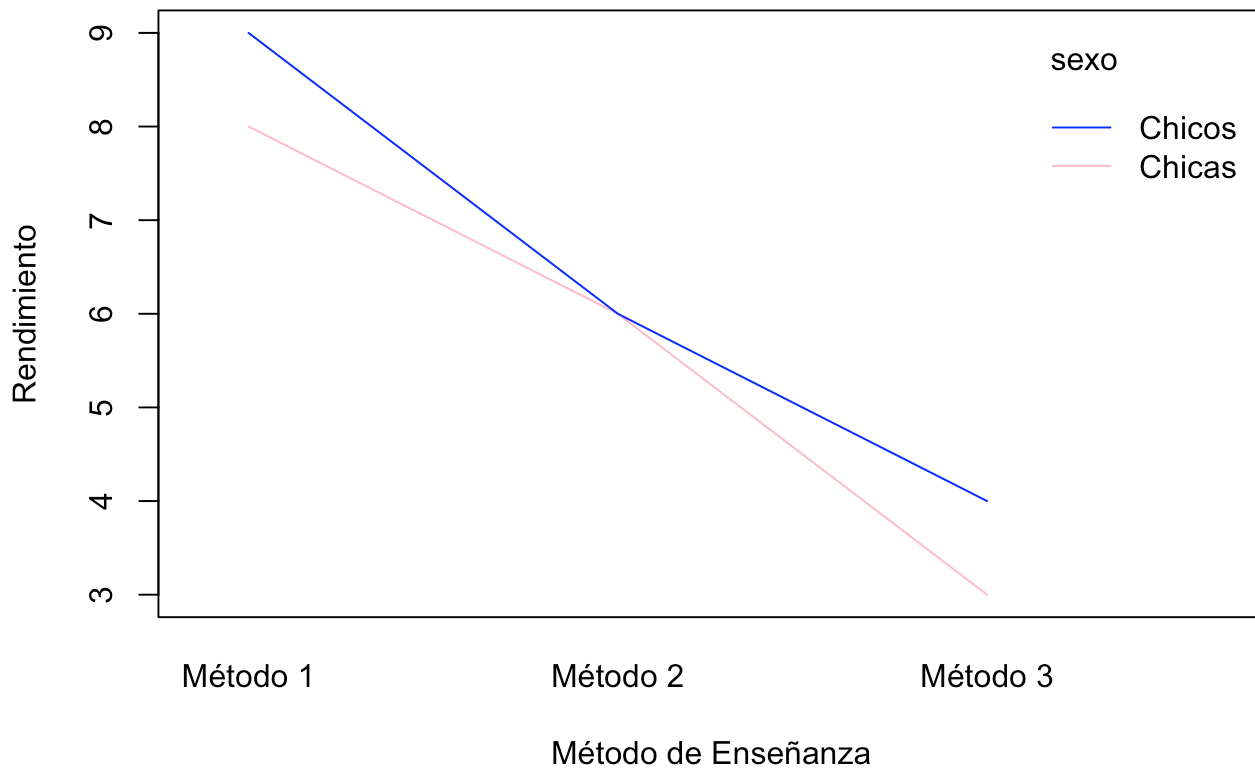
```
A = aov(rendimiento ~ metodos * sexo)
summary(A)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## metodos      2     150    75.00   32.143 3.47e-08 ***
## sexo         1       4     4.00    1.714   0.200
## metodos:sexo  2       2     1.00    0.429   0.655
## Residuals    30      70     2.33
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Haz la gráfica de interacción de dos factores en ANOVA.

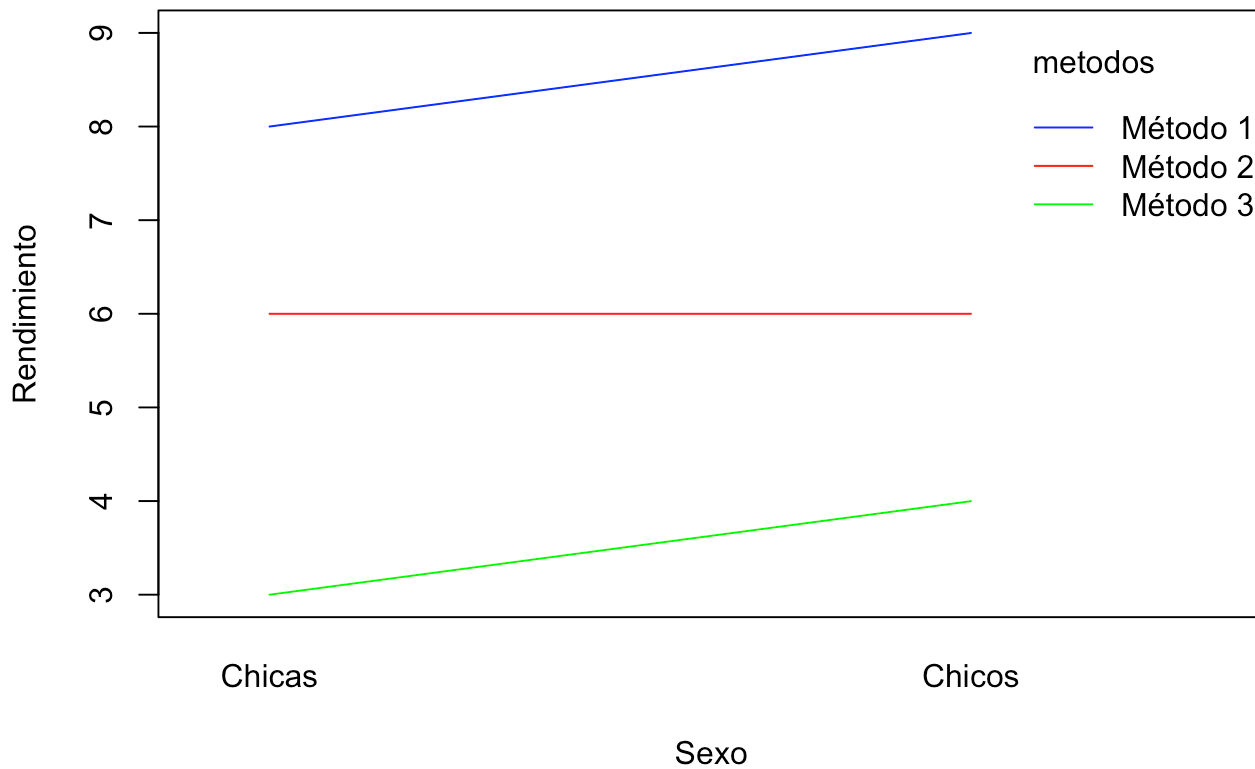
```
interaction.plot(metodos, sexo, rendimiento, main = "Interacción entre Método de Enseñanza y Sexo", xlab = "Método de Enseñanza", ylab = "Rendimiento", col = c("pink", "blue"), lty = 1, legend = TRUE)
```

Interacción entre Método de Enseñanza y Sexo



```
interaction.plot(sexo, metodos, rendimiento, main = "Interacción entre Sexo y Método de Enseñanza", xlab = "Sexo", ylab = "Rendimiento", col = c("blue", "red", "green"), lty = 1, legend = TRUE)
```

Interacción entre Sexo y Método de Enseñanza

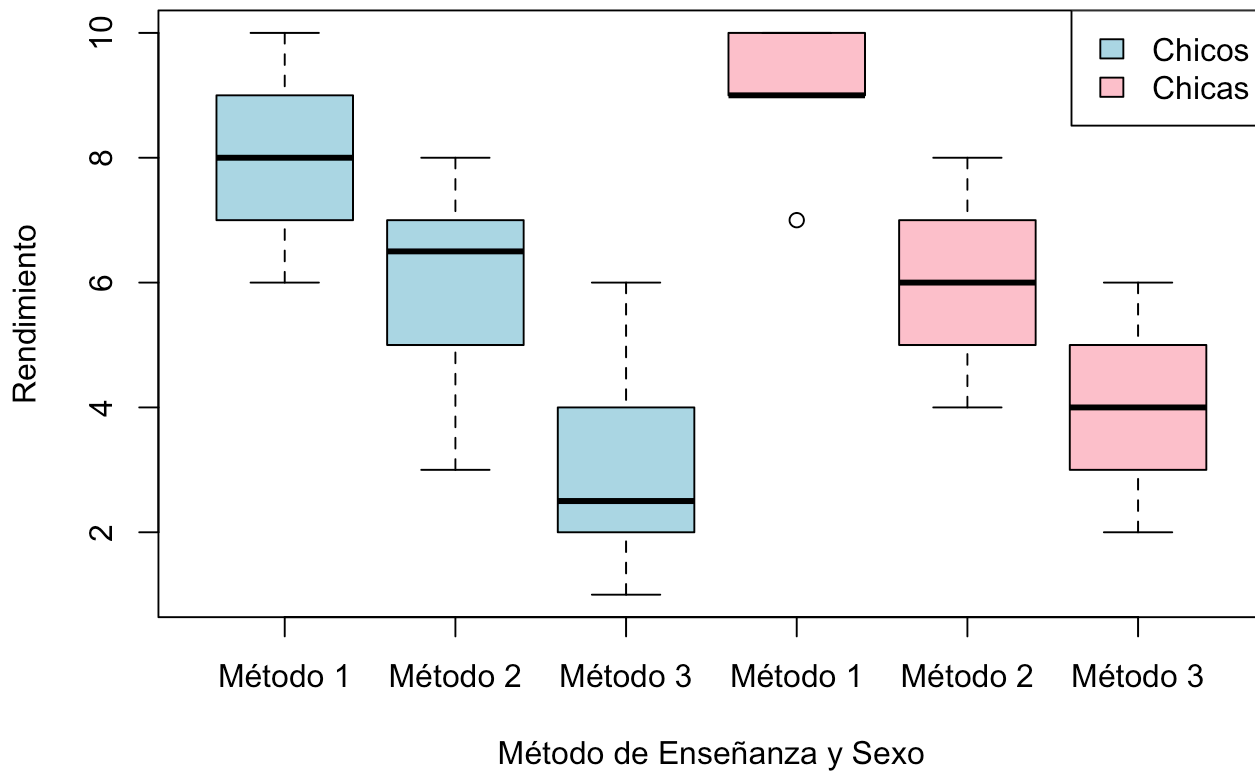


- Haz el boxplot para visualizar la interacción de los factores, por ejemplo, peso por dieta interacción ejercicio:

```
boxplot(rendimiento ~ metodos * sexo, main = "Evaluación de los Estudiantes", xlab = "Método de Enseñanza y Sexo", ylab = "Rendimiento", col = c("lightblue", "lightblue", "lightblue", "pink", "pink", "pink"), names = c("Método 1", "Método 2", "Método 3", "Método 1", "Método 2", "Método 3"))
```

```
legend("topright", legend = c("Chicos", "Chicas"), fill = c("lightblue", "pink"))
```

Evaluación de los Estudiantes



- Interpreta el resultado desde la perspectiva estadística y en el contexto del problema.

Gráfica “Interacción entre Método de Enseñanza y Sexo”:

- Método 1: Chicos como chicas obtuvieron los rendimientos más altos, los chicos ligeramente superando a las chicas.
- Método 2: El rendimiento disminuye para ambos sexos en comparación con el Método 1, pero más para los chicos.
- Método 3: El rendimiento disminuye aún más para ambos sexos, la diferencia es más entre chicos y chicas. Los chicos obtienen mejores rendimientos que chicas en este método.

Gráfica “Interacción entre Sexo y Método de Enseñanza”:

- Método 1: Ambos sexos obtienen buenos resultados, pero los chicos tienen un rendimiento superior a las chicas.
- Método 2: Hay una reducción en el rendimiento, es más notable en los chicos, que tienen un rendimiento menor comparado con las chicas.
- Método 3: El rendimiento cae para ambos sexos, pero más para las chicas. Los chicos tienen un mejor rendimiento comparado con las chicas.

- Escribe tus conclusiones parciales.

El rendimiento promedio de los estudiantes varía significativamente según el método de enseñanza utilizado. El Método 1 puede ser el más efectivo, y el Método 3 es el menos efectivo para ambos sexos. También es claro que existen diferencias en el rendimiento promedio entre chicos y chicas, en general, los chicos tienden a tener un rendimiento ligeramente mejor que las chicas.

Y por último hay una interacción significativa entre el sexo y el método de enseñanza. Lo que significa que el efecto del método de enseñanza en el rendimiento de los estudiantes no es el mismo para chicos y chicas.

4. Realiza el ANOVA para dos niveles sin interacción:

- Consulta el código de R en los apoyos de clase de “ANOVA”

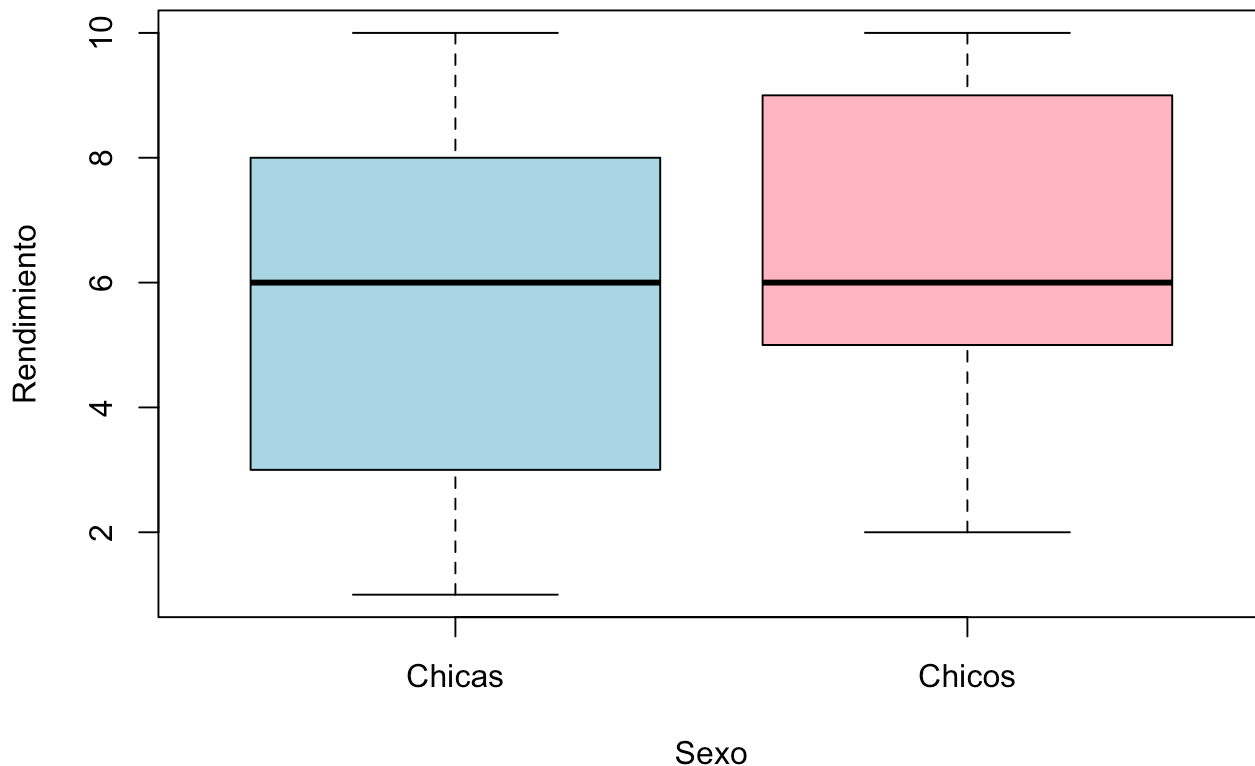
```
B = aov(rendimiento ~ metodos + sexo)
summary(B)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## metodos      2    150   75.00  33.333 1.5e-08 ***
## sexo         1     4    4.00   1.778  0.192
## Residuals   32     72    2.25
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Haz el boxplot de rendimiento por sexo. Calcula la media para el rendimiento por sexo y método.

```
boxplot(rendimiento ~ sexo, main = "Boxplot de Rendimiento por Sexo", xlab = "Sexo", ylab = "Rendimiento", col = c("lightblue", "lightpink"))
```

Boxplot de Rendimiento por Sexo



```
media_por_sexo = tapply(rendimiento, sexo, mean)
print("Media del rendimiento por sexo:")
```

```
## [1] "Media del rendimiento por sexo:"
```

```
print(media_por_sexo)
```

```
##      Chicas      Chicos
## 5.666667 6.333333
```

```
media_por_sexo_metodo = tapply(rendimiento, list(sexo, metodos), mean)
print("Media del rendimiento por sexo y método:")
```

```
## [1] "Media del rendimiento por sexo y método:"
```

```
print(media_por_sexo_metodo)
```

```
##           Método 1 Método 2 Método 3
## Chicas           8         6         3
## Chicos           9         6         4
```

- Haz los intervalos de confianza de rendimiento por sexo. Grafícalos.

```
ic_chicos = t.test(rendimiento[sexo == "Chicos"])$conf.int
ic_chicas = t.test(rendimiento[sexo == "Chicas"])$conf.int

print("Intervalo de confianza para Chicos:")
```

```
## [1] "Intervalo de confianza para Chicos:"
```

```
print(ic_chicos)
```

```
## [1] 5.103347 7.563320
## attr("conf.level")
## [1] 0.95
```

```
print("Intervalo de confianza para Chicas:")
```

```
## [1] "Intervalo de confianza para Chicas:"
```

```
print(ic_chicas)
```



```
## [1] 4.356505 6.976828
## attr("conf.level")
## [1] 0.95
```

```
media_chicos = mean(rendimiento[sexo == "Chicos"])
media_chicas = mean(rendimiento[sexo == "Chicas"])

media = c(media_chicos, media_chicas)
ic_lower = c(ic_chicos[1], ic_chicas[1])
ic_upper = c(ic_chicos[2], ic_chicas[2])
sexos = c("Chicos", "Chicas")

plot(1:2, media, ylim = c(min(ic_lower), max(ic_upper)), pch = 19, col = c("blue", "red"),
     xlab = "Sexo", ylab = "Rendimiento", main = "Intervalos de Confianza de Rendimiento por Sexo",
     xaxt = "n")
axis(1, at = 1:2, labels = sexos)
arrows(1, ic_lower[1], 1, ic_upper[1], angle = 90, code = 3, length = 0.1, col = "blue")
arrows(2, ic_lower[2], 2, ic_upper[2], angle = 90, code = 3, length = 0.1, col = "red")

points(1:2, media, pch = 19, col = c("blue", "red"))
legend("topright", legend = c("Chicos", "Chicas"), col = c("blue", "red"), pch = 19)
```

Intervalos de Confianza de Rendimiento por Sexo



- Interpreta el resultado desde la perspectiva estadística y en el contexto del problema.

El intervalo de confianza para los chicos es de 5.10 a 7.56 con un nivel de confianza del 95%. Esto significa que, tenemos un 95% de confianza, la media del rendimiento de los chicos se encuentra entre 5.10 y 7.56. Lo mismo para las chicas.

- Escribe tus conclusiones parciales.

Visualmente, el punto que representa el rendimiento promedio de los chicos está más alto que el de las Chicas, pero debido a la superposición de los intervalos de confianza, no se puede concluir que esta diferencia sea significativa.

5. Realiza el ANOVA para un efecto principal.

- Consulta el código de R en los apoyos de clase de “ANOVA”.

```
levels(metodos)
```

```
## [1] "Método 1" "Método 2" "Método 3"
```

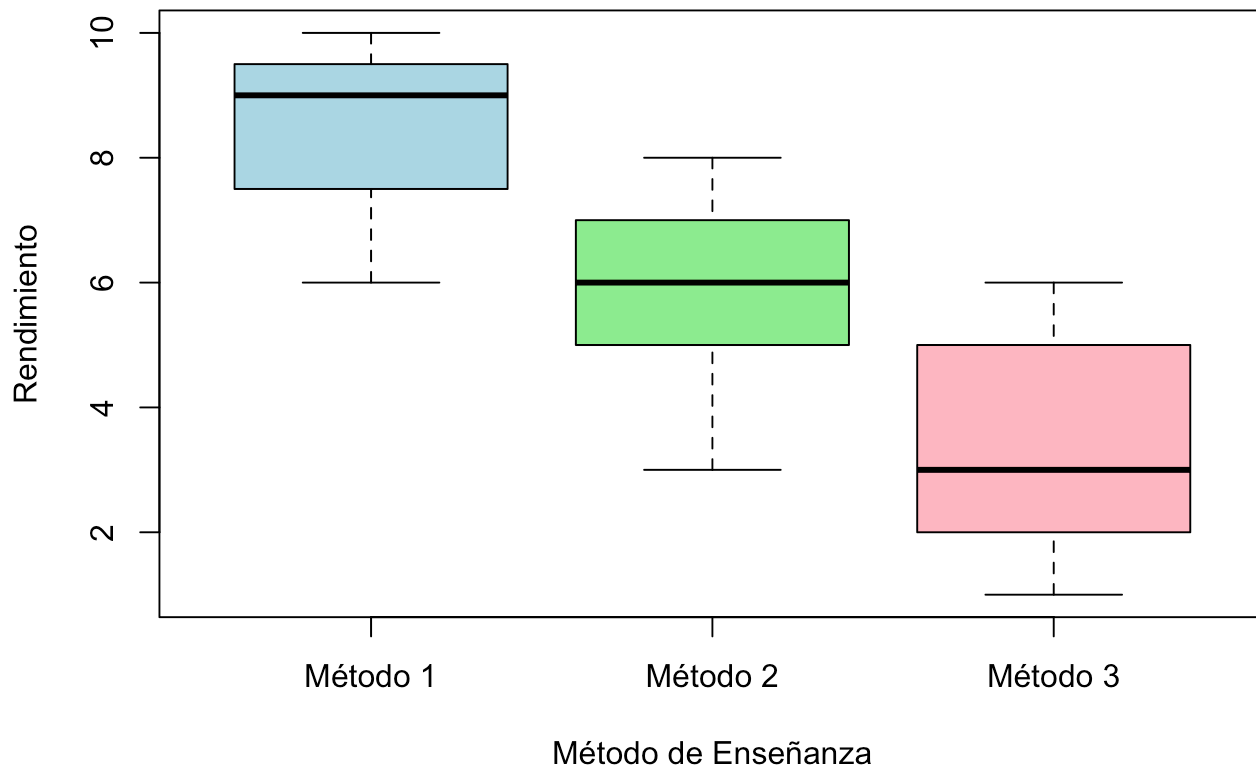
```
C = aov(rendimiento ~ metodos)
summary(C)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## metodos        2     150    75.0    32.57 1.55e-08 ***
## Residuals     33       76     2.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Haz el boxplot de rendimiento por método de enseñanza. Calcula la media.

```
boxplot(rendimiento ~ metodos, main = "Rendimiento por Método de Enseñanza", xlab = "Método de Enseñanza", ylab = "Rendimiento", col = c("lightblue", "lightgreen", "lightpink"))
```

Rendimiento por Método de Enseñanza



```
media_por_metodo = tapply(rendimiento, metodos, mean)
media_por_metodo
```

```
## Método 1 Método 2 Método 3
##      8.5      6.0      3.5
```

```
media_general = mean(rendimiento)
media_general
```

```
## [1] 6
```

- Haz los intervalos de confianza de rendimiento por método. Gráficalos.

```

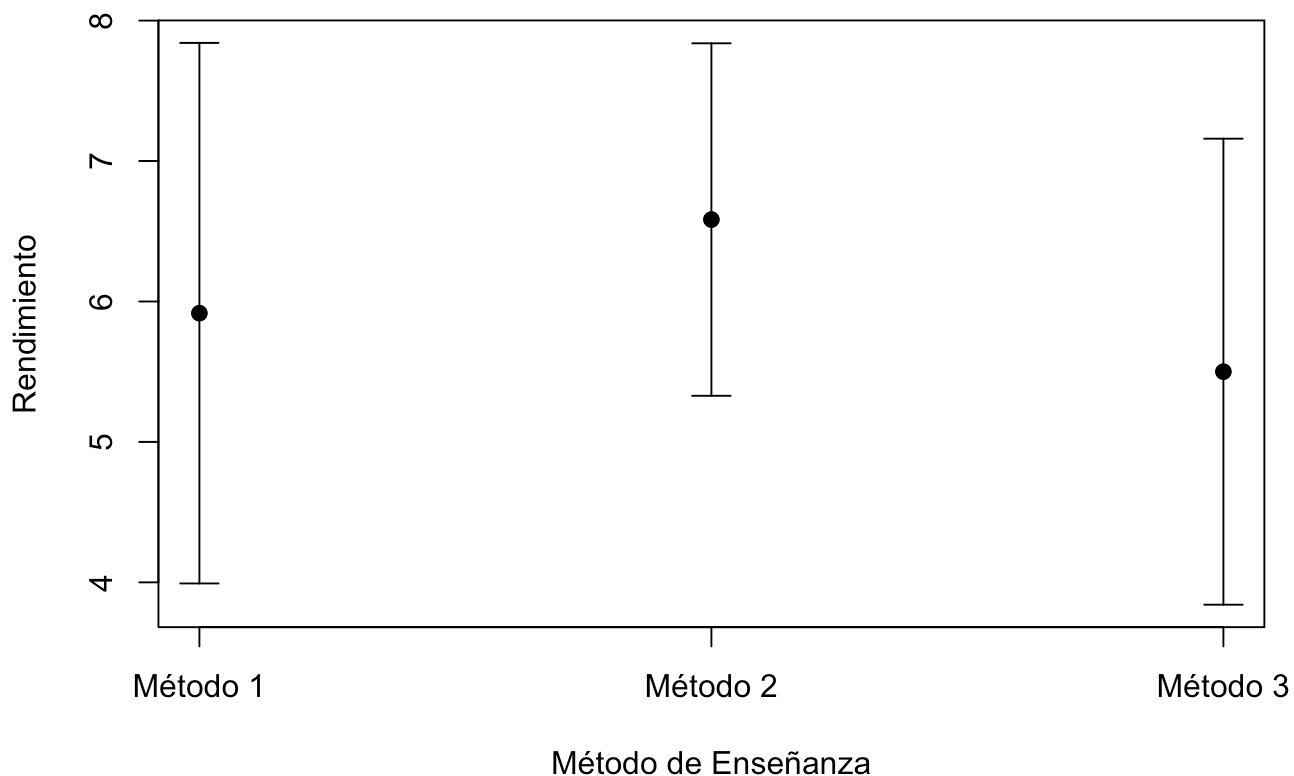
metodos = unique(metodos)
medias = numeric(length(metodos))
ic_lower = numeric(length(metodos))
ic_upper = numeric(length(metodos))

for (i in 1:length(metodos)) {
  datos = rendimiento[metodos == metodos[i]]
  t_test = t.test(datos)
  medias[i] = mean(datos)
  ic_lower[i] = t_test$conf.int[1]
  ic_upper[i] = t_test$conf.int[2]
}

plot(1:length(metodos), medias, ylim = c(min(ic_lower), max(ic_upper)), pch = 19, xlab =
"Método de Enseñanza", ylab = "Rendimiento", main = "Intervalos de Confianza de Rendimiento por Método de Enseñanza", xaxt = "n")
axis(1, at = 1:length(metodos), labels = metodos)
arrows(1:length(metodos), ic_lower, 1:length(metodos), ic_upper, angle = 90, code = 3, length = 0.1)
points(1:length(metodos), medias, pch = 19)

```

Intervalos de Confianza de Rendimiento por Método de Enseñanza



- Realiza la prueba de comparaciones múltiples de Tukey. Grafica los intervalos de confianza de Tukey.

```

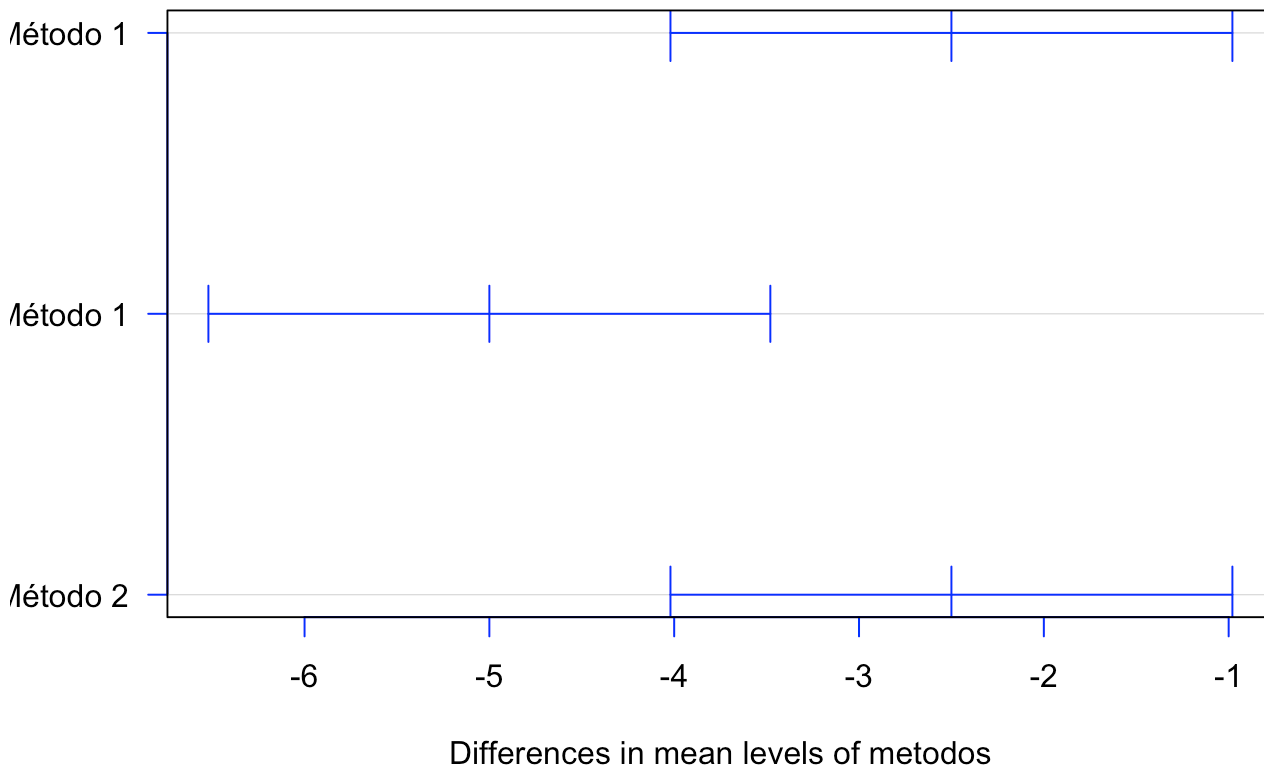
tukey_result <- TukeyHSD(C)
print(tukey_result)

```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = rendimiento ~ metodos)
##
## $metodos
##              diff      lwr      upr    p adj
## Método 2-Método 1 -2.5 -4.020241 -0.9797592 0.0008674
## Método 3-Método 1 -5.0 -6.520241 -3.4797592 0.0000000
## Método 3-Método 2 -2.5 -4.020241 -0.9797592 0.0008674
```

```
plot(tukey_result, las = 1, col = "blue")
```

95% family-wise confidence level



- Interpreta el resultado desde la perspectiva estadística y en el contexto del problema.

Tukey multiple comparisons of means: se usa para la prueba de Tukey donde realiza comparaciones múltiples de medias. **95% family-wise confidence level:** Nos dice que los intervalos de confianza para las diferencias de medias entre los métodos se calculan con un nivel de confianza del 95%. Y “Family-wise” se refiere a que este nivel de confianza se aplica a todas las comparaciones simultáneamente.

Fit: aov(formula = rendimiento ~ metodos): Nos dice que se ajustó un modelo ANOVA antes de realizar la prueba de Tukey.

Método 2 - Método 1:

- **diff = -2.5:** La media del rendimiento para el Método 2 es 2.5 unidades menor que la del Método 1.
- **lwr = -4.020241, upr = -0.9797592:** El intervalo de confianza del 95% para esta diferencia es de -4.02 a -0.98. Dado que este intervalo no incluye el cero, hay una diferencia significativa en el rendimiento entre Método 1 y Método 2.
- **p adj = 0.0008674:** El valor p ajustado es mucho menor que 0.05, lo que indica que la diferencia es estadísticamente significativa.

Y así podemos irnos con cada comparación.

- Escribe tus conclusiones parciales

El Método 1 tiene un rendimiento significativamente mayor que el Método 2 y Método 3.

6. Comprueba la validez del modelo. Comprueba:

- Normalidad
- Homocedasticidad
- Independencia
- Relación lineal entre las variables (coeficiente de determinación).

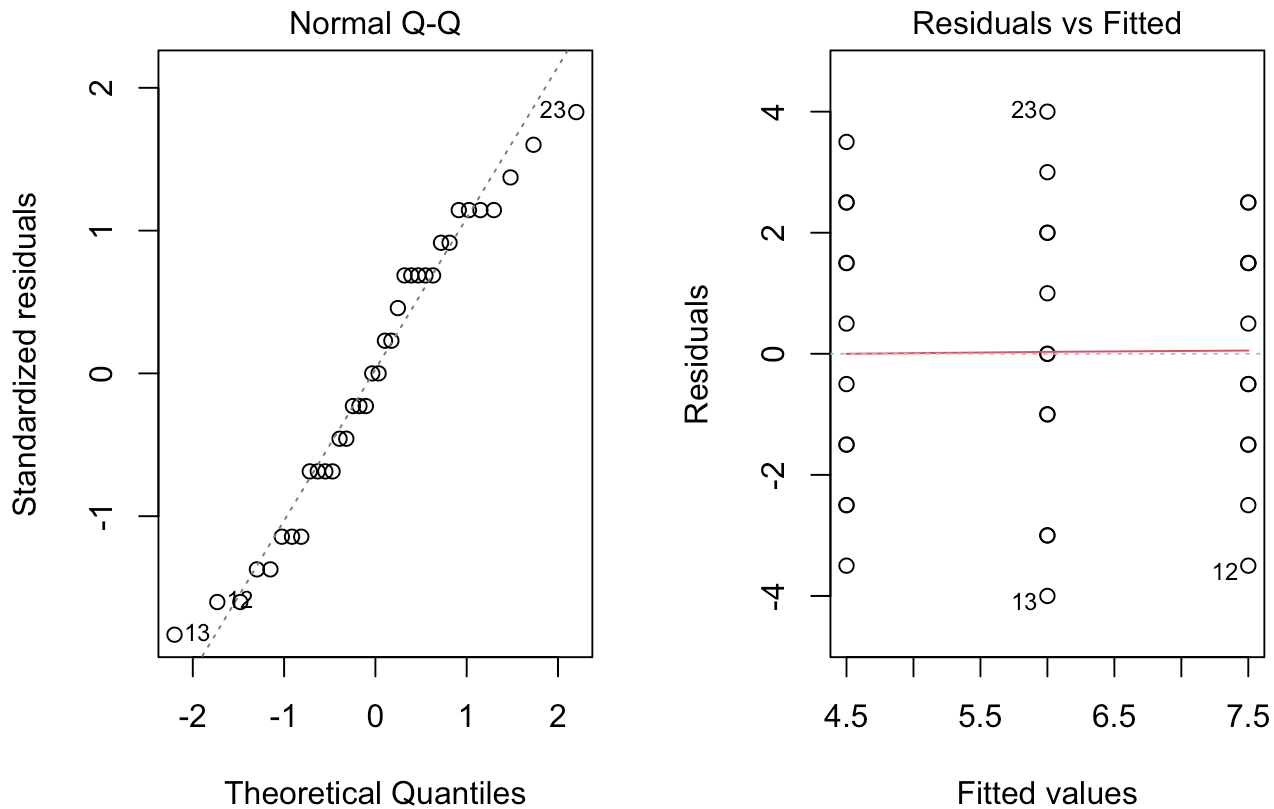
```
rendimiento = c(10, 7, 9, 9, 9, 10, 5, 7, 6, 6, 8, 4, 2, 6, 3, 5, 5, 3, 9, 7, 8, 8, 10,
6, 8, 3, 5, 6, 7, 7, 2, 6, 2, 1, 4, 3)
```

```
metodos <- factor(rep(c("Método 1", "Método 2", "Método 3"), each = 12))
C <- aov(rendimiento ~ metodos)
```

```
par(mfrow = c(1, 2))
plot(C, which = 2)
shapiro.test(resid(C)) # Prueba de Shapiro-Wilk para normalidad
```

```
##
## Shapiro-Wilk normality test
##
## data:  resid(C)
## W = 0.95995, p-value = 0.2142
```

```
plot(C, which = 1) # Gráfico de residuos vs. valores ajustados
```



```
bartlett.test(rendimiento ~ metodos) # Prueba de Bartlett para homocedasticidad
```

```
##
## Bartlett test of homogeneity of variances
##
## data:  rendimiento by metodos
## Bartlett's K-squared = 0.62817, df = 2, p-value = 0.7305
```

```
summary(C)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## metodos     2      54   27.000    5.18 0.0111 *
## Residuals   33     172    5.212
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

7. Concluye en el contexto del problema.

- Los residuos del modelo están distribuidos normalmente.
- Las varianzas entre los grupos son homogéneas o iguales.
- Hay diferencias significativas en el rendimiento entre los métodos de enseñanza.

Problema 2.

Vibración de motores

Un ingeniero de procesos ha identificado dos causas potenciales de vibración de los motores eléctricos, el material utilizado para la carcasa del motor (factor A) y el proveedor de cojinetes utilizados en el motor (Factor B). Los siguientes datos sobre la cantidad de vibración (micrones) se obtuvieron mediante un experimento en el cual se construyeron motores con carcasas de acero, aluminio y plástico y cojinetes suministrados por cinco proveedores seleccionados al azar.

	Proveedor				
Material	1	2	3	4	5
Acero	13.1	16.3	13.7	15.7	13.5
	13.2	15.8	14.3	15.8	12.5
Aluminio	15.0	15.7	13.9	13.7	13.4
	14.8	16.4	14.3	14.2	13.8
Plástico	14.0	17.2	12.4	14.4	13.2
	14.3	16.7	12.3	13.9	13.1

¿Existe evidencia de que el proveedor y el material son causas de la vibración de los motores eléctricos?

1. Análisis Exploratorio.

- Calcula la media para el rendimiento por método de enseñanza, puedes consultar el código en R en los apoyos de clase de “ANOVA”.

```
vibracion = c(13.1, 13.2, 16.3, 15.8, 13.7, 14.3, 15.7, 15.8, 13.5, 12.5, 15.0, 14.8, 15.7, 16.4, 13.9, 14.3, 13.7, 14.2, 13.4, 13.8, 14.0, 14.3, 17.2, 16.7, 12.4, 12.3, 14.4, 13.9, 13.2, 13.1)
```

```
material = factor(rep(c("Acero", "Aluminio", "Plástico"), each = 10))
medias = tapply(vibracion, material, mean)
print(medias)
```

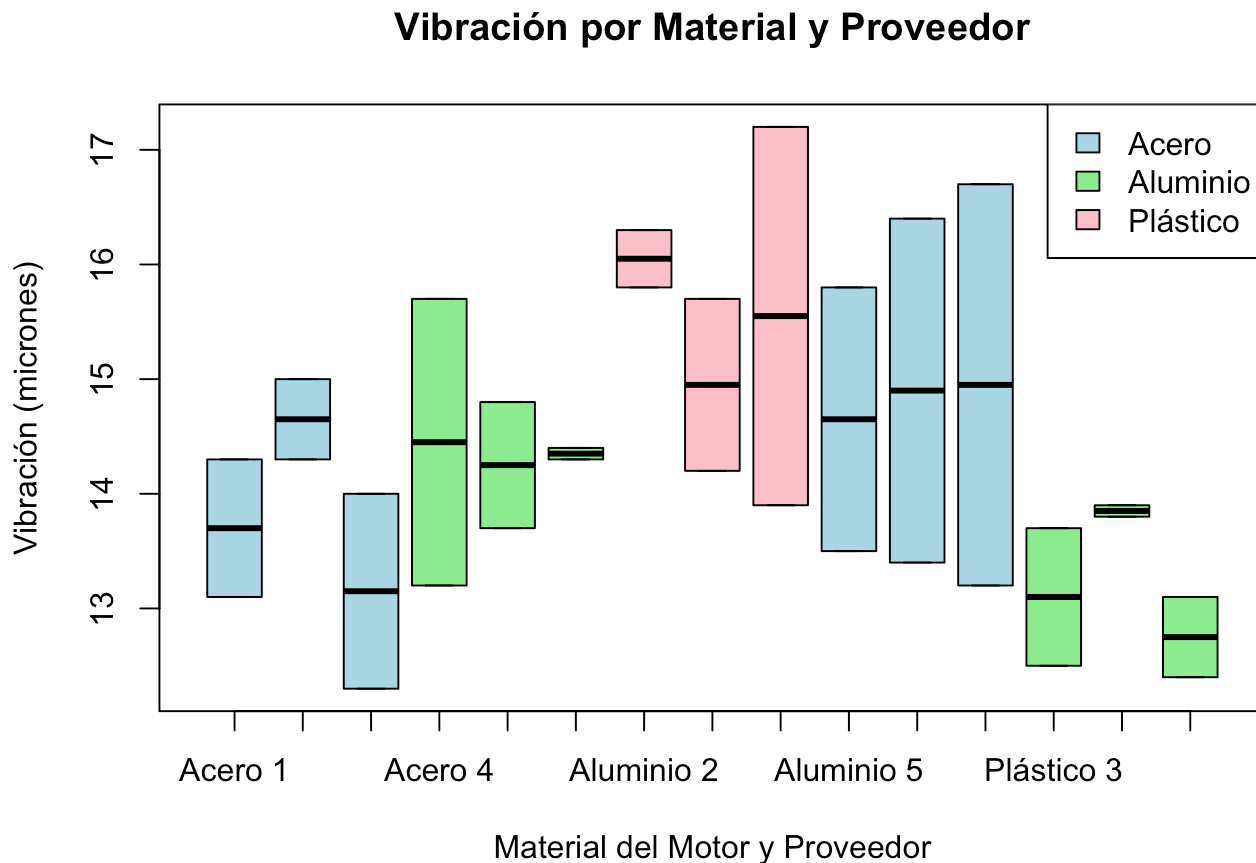
```
##      Acero Aluminio Plástico
##      14.39    14.52    14.15
```

- Haz el boxplot de la evaluación de los estudiantes por método de enseñanza y sexo.


```
proveedor = factor(rep(1:5, times = 6))
```

```
boxplot(vibracion ~ material + proveedor, main = "Vibración por Material y Proveedor", xlab = "Material del Motor y Proveedor", ylab = "Vibración (micrones)", col = c("lightblue", "lightblue", "lightblue", "lightgreen", "lightgreen", "lightgreen", "pink", "pink", "pink"), names = c("Acero 1", "Acero 2", "Acero 3", "Acero 4", "Acero 5", "Aluminio 1", "Aluminio 2", "Aluminio 3", "Aluminio 4", "Aluminio 5", "Plástico 1", "Plástico 2", "Plástico 3", "Plástico 4", "Plástico 5"))
```

```
legend("topright", legend = c("Acero", "Aluminio", "Plástico"), fill = c("lightblue", "lightgreen", "pink"))
```



2. Las Hipótesis.

- Establece las hipótesis estadísticas (tienen que ser 3).

1. Hipótesis I:

- **Hipótesis Nula (H_{01}):** El material del motor no tiene un efecto significativo en la vibración del motor. Es decir, las medias de vibración son iguales para todos los materiales.
- **Hipótesis Alternativa (H_{11}):** Al menos uno de los materiales del motor tiene un efecto significativo diferente en la vibración del motor.

2. Hipótesis II:

- **Hipótesis Nula (H_{02}):** El proveedor de cojinetes no tiene un efecto significativo en la vibración del motor. Es decir, las medias de vibración son iguales para todos los proveedores.

- **Hipótesis Alternativa (H_{12}):** Al menos un proveedor de cojinetes tiene un efecto significativo diferente en la vibración del motor.

3. Hipótesis III:

- **Hipótesis Nula (H_{03}):** H_{03} : No existe interacción significativa entre el material del motor y el proveedor de cojinetes en términos de su efecto sobre la vibración del motor. Es decir, el efecto de un material del motor es el mismo independientemente del proveedor de cojinetes.
- **Hipótesis Alternativa (H_{13}):** H_{13} : Existe una interacción significativa entre el material del motor y el proveedor de cojinetes, lo que significa que el efecto de un material del motor depende del proveedor de cojinetes.

3. Realiza el ANOVA para dos niveles con interacción:

- Consulta el código en R en los apoyos de clase de “ANOVA”.

```
anova = aov(vibracion ~ material * proveedor)
summary(anova)
```

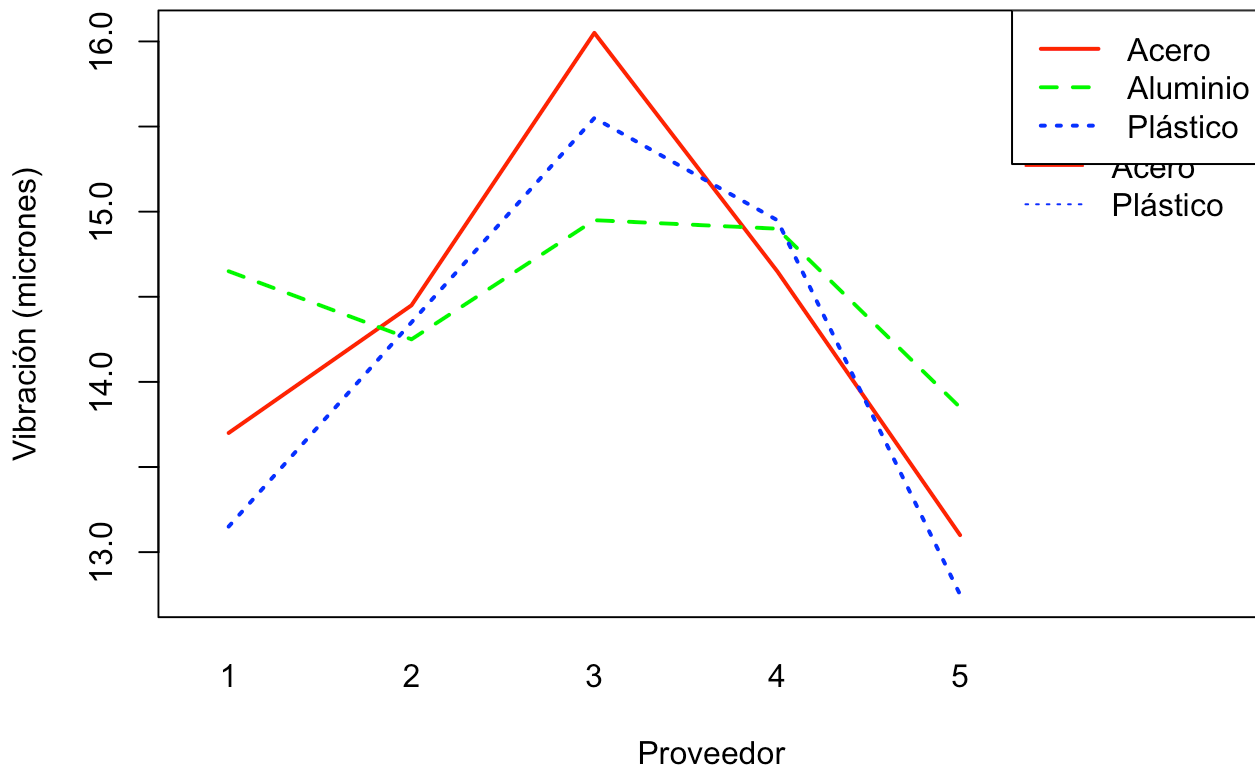
```
##                Df Sum Sq Mean Sq F value Pr(>F)
## material         2  0.705    0.352    0.195 0.8248
## proveedor        4 18.651    4.663    2.583 0.0797 .
## material:proveedor 8  4.219    0.527    0.292 0.9579
## Residuals       15 27.080    1.805
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Haz la gráfica de interacción de dos factores en ANOVA.

```
interaction.plot(proveedor, material, vibracion, col = c("red", "green", "blue"), lty =
1:3, lwd = 2, main = "Gráfica de Interacción entre Material y Proveedor", xlab = "Provee
dor", ylab = "Vibración (micrones)")

legend("topright", legend = levels(material), col = c("red", "green", "blue"), lty = 1:
3, lwd = 2)
```

Gráfica de Interacción entre Material y Proveedor

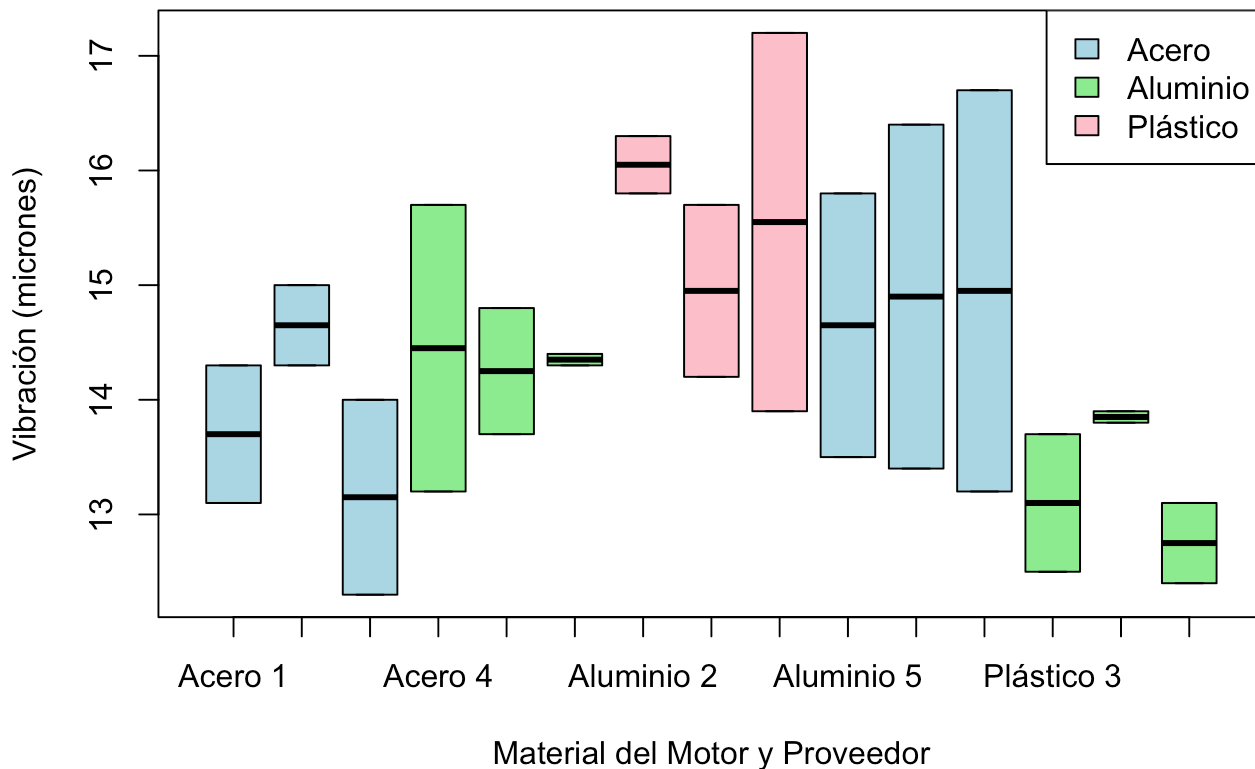


- Haz el boxplot para visualizar la interacción de los factores, por ejemplo, peso por dieta interacción ejercicio:

```
boxplot(vibracion ~ material * proveedor, main = "Vibración por Material y Proveedor", xlab = "Material del Motor y Proveedor", ylab = "Vibración (micrones)", col = c("lightblue", "lightblue", "lightblue", "lightgreen", "lightgreen", "lightgreen", "pink", "pink", "pink"), names = c("Acero 1", "Acero 2", "Acero 3", "Acero 4", "Acero 5", "Aluminio 1", "Aluminio 2", "Aluminio 3", "Aluminio 4", "Aluminio 5", "Plástico 1", "Plástico 2", "Plástico 3", "Plástico 4", "Plástico 5"))
```

```
legend("topright", legend = c("Acero", "Aluminio", "Plástico"), fill = c("lightblue", "lightgreen", "pink"))
```

Vibración por Material y Proveedor



4. Realiza el ANOVA para dos niveles sin interacción:

- Consulta el código de R en los apoyos de clase de “ANOVA”

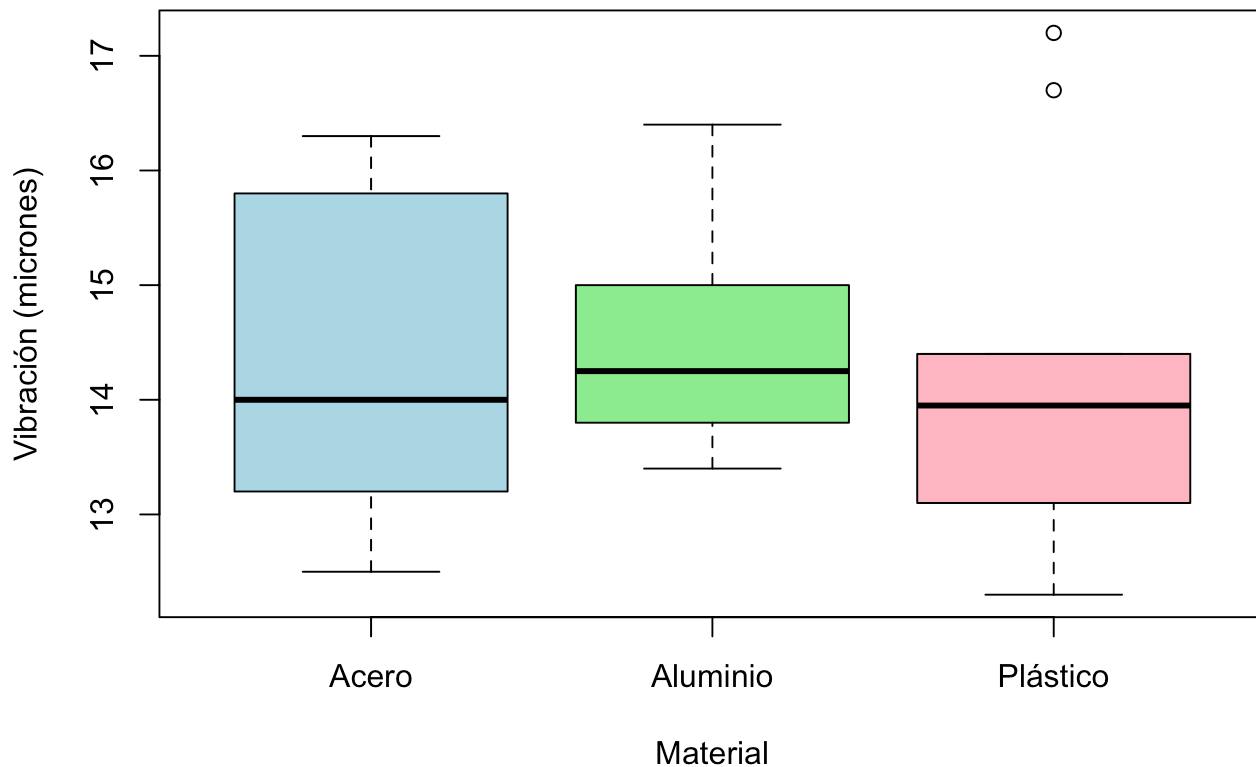
```
anova_sin_interaccion = aov(vibracion ~ material + proveedor)
summary(anova_sin_interaccion)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## material    2  0.705    0.352   0.259 0.7741
## proveedor    4 18.651    4.663   3.427 0.0245 *
## Residuals   23 31.299    1.361
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Haz el boxplot de rendimiento por sexo. Calcula la media para el rendimiento por sexo y método.

```
boxplot(vibracion ~ material, main = "Boxplot de Vibración por Material", xlab = "Material", ylab = "Vibración (micrones)", col = c("lightblue", "lightgreen", "lightpink"))
```

Boxplot de Vibración por Material



```
media_por_material = tapply(vibracion, material, mean)
print("Media de la vibración por material:")
```

```
## [1] "Media de la vibración por material:"
```

```
print(media_por_material)
```

```
##      Acero Aluminio Plástico
##      14.39   14.52   14.15
```

```
media_por_material_proveedor = tapply(vibracion, list(material, proveedor), mean)
print("Media de la vibración por material y proveedor:")
```

```
## [1] "Media de la vibración por material y proveedor:"
```

```
print(media_por_material_proveedor)
```

```
##           1      2      3      4      5
## Acero    13.70 14.45 16.05 14.65 13.10
## Aluminio 14.65 14.25 14.95 14.90 13.85
## Plástico 13.15 14.35 15.55 14.95 12.75
```

- Haz los intervalos de confianza de rendimiento por sexo. Grafícalos.

```
ic_acero = t.test(vibracion[material == "Acero"])$conf.int
ic_aluminio = t.test(vibracion[material == "Aluminio"])$conf.int
ic_plastico = t.test(vibracion[material == "Plástico"])$conf.int

print("Intervalo de confianza para Acero:")
```

```
## [1] "Intervalo de confianza para Acero:"
```

```
print(ic_acero)
```

```
## [1] 13.39909 15.38091
## attr("conf.level")
## [1] 0.95
```

```
print("Intervalo de confianza para Aluminio:")
```

```
## [1] "Intervalo de confianza para Aluminio:"
```

```
print(ic_aluminio)
```

```
## [1] 13.83651 15.20349
## attr("conf.level")
## [1] 0.95
```

```
print("Intervalo de confianza para Plástico:")
```

```
## [1] "Intervalo de confianza para Plástico:"
```

```
print(ic_plastico)
```

```
## [1] 12.97056 15.32944
## attr("conf.level")
## [1] 0.95
```

```

media_acero = mean(vibracion[material == "Acero"])
media_aluminio = mean(vibracion[material == "Aluminio"])
media_plastico = mean(vibracion[material == "Plástico"])

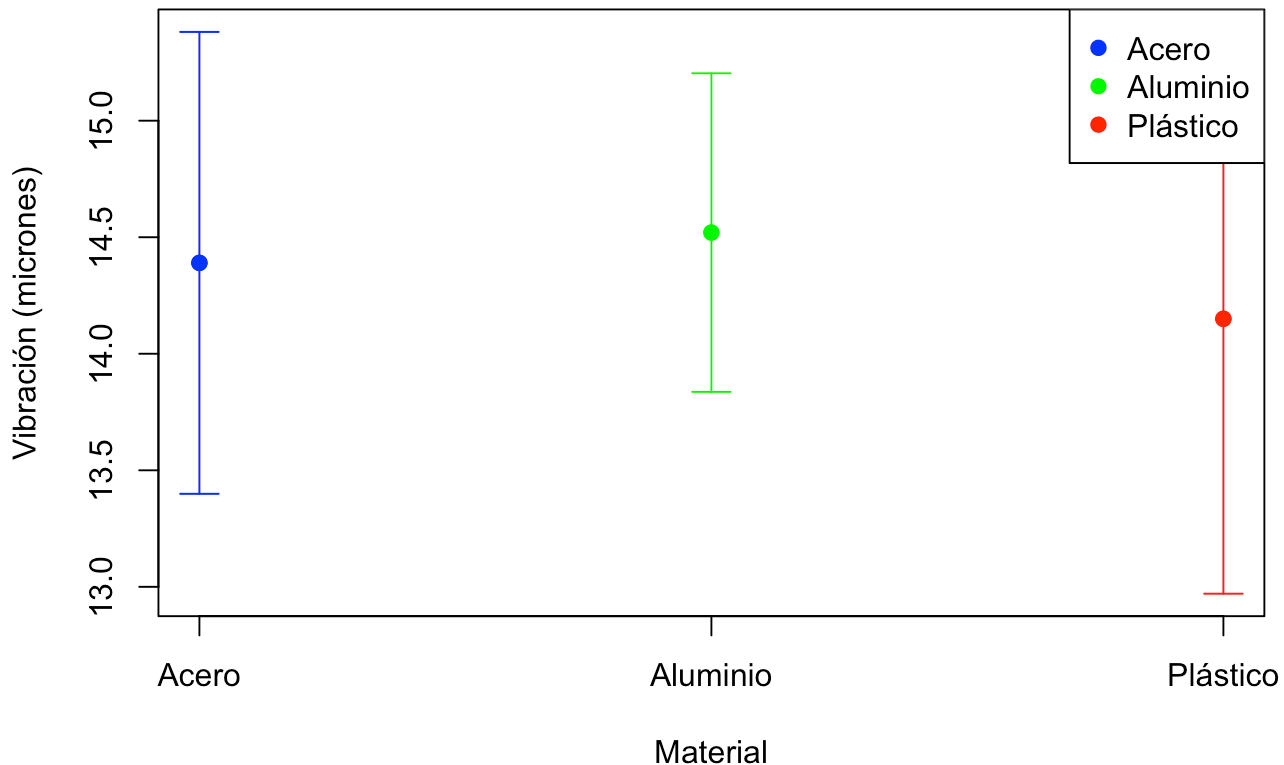
media = c(media_acero, media_aluminio, media_plastico)
ic_lower = c(ic_acero[1], ic_aluminio[1], ic_plastico[1])
ic_upper = c(ic_acero[2], ic_aluminio[2], ic_plastico[2])
materiales = c("Acero", "Aluminio", "Plástico")

plot(1:3, media, ylim = c(min(ic_lower), max(ic_upper)), pch = 19, col = c("blue", "green", "red"), xlab = "Material", ylab = "Vibración (micrones)", main = "Intervalos de Confianza de Vibración por Material", xaxt = "n")
axis(1, at = 1:3, labels = materiales)
arrows(1, ic_lower[1], 1, ic_upper[1], angle = 90, code = 3, length = 0.1, col = "blue")
arrows(2, ic_lower[2], 2, ic_upper[2], angle = 90, code = 3, length = 0.1, col = "green")
arrows(3, ic_lower[3], 3, ic_upper[3], angle = 90, code = 3, length = 0.1, col = "red")

points(1:3, media, pch = 19, col = c("blue", "green", "red"))
legend("topright", legend = c("Acero", "Aluminio", "Plástico"), col = c("blue", "green", "red"), pch = 19)

```

Intervalos de Confianza de Vibración por Material



5. Realiza el ANOVA para un efecto principal.

- Consulta el código de R en los apoyos de clase de “ANOVA”.

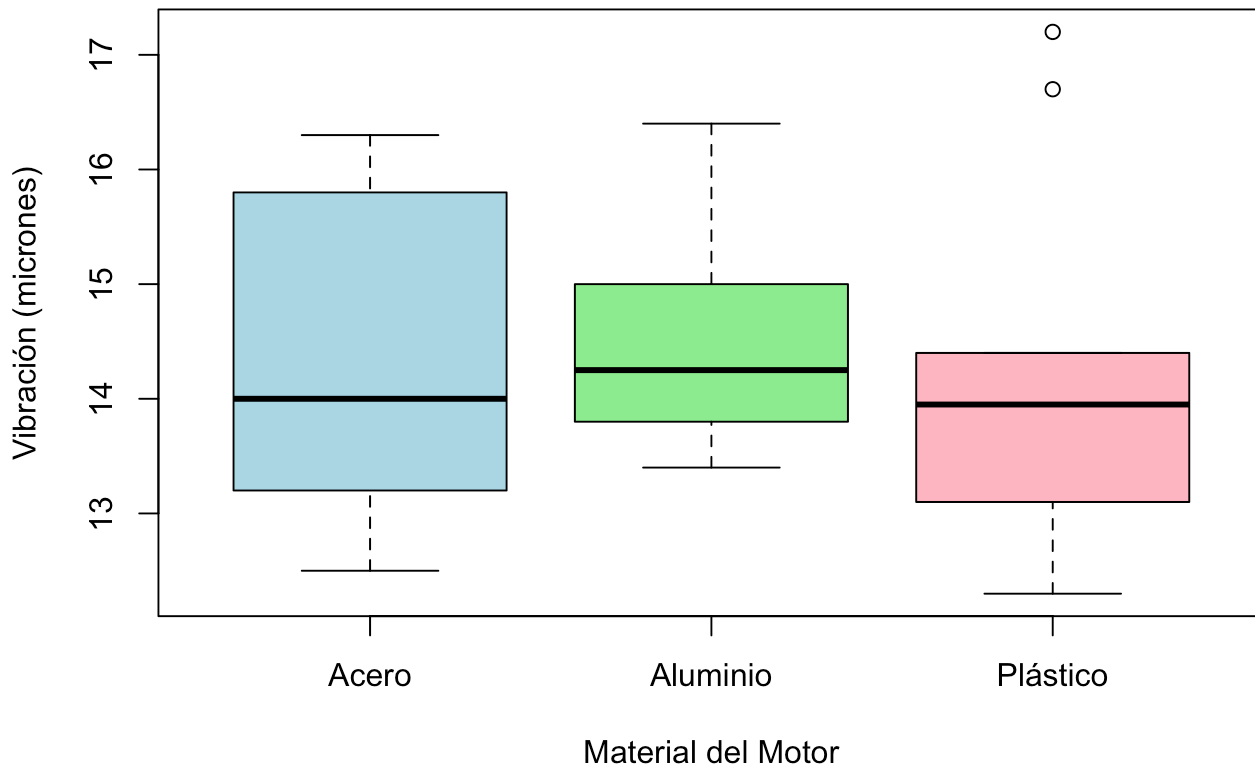
```
anova_material = aov(vibracion ~ material)
summary(anova_material)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## material    2   0.70   0.3523    0.19  0.828
## Residuals  27  49.95   1.8500
```

- Haz el boxplot de rendimiento por método de enseñanza. Calcula la media.

```
boxplot(vibracion ~ material, main = "Vibración por Material del Motor", xlab = "Material del Motor", ylab = "Vibración (micrones)", col = c("lightblue", "lightgreen", "lightpink"))
```

Vibración por Material del Motor



```
media_por_material = tapply(vibracion, material, mean)
media_por_material
```

```
##   Acero Aluminio Plástico
##  14.39   14.52   14.15
```



```
media_general = mean(vibracion)
media_general
```

```
## [1] 14.35333
```

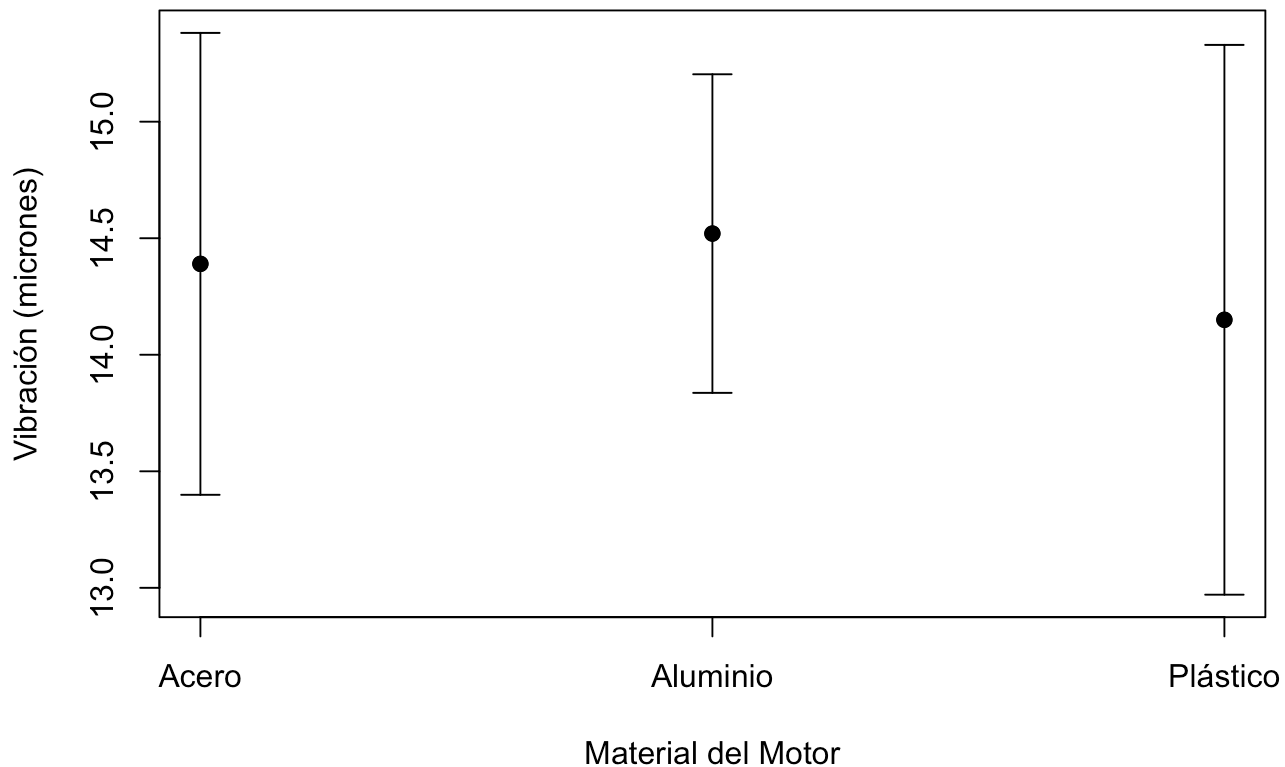
- Haz los intervalos de confianza de rendimiento por método. Grafícalos.

```
materiales = unique(material)
medias = numeric(length(materiales))
ic_lower = numeric(length(materiales))
ic_upper = numeric(length(materiales))

for (i in 1:length(materiales)) {
  datos = vibracion[material == materiales[i]]
  t_test = t.test(datos)
  medias[i] = mean(datos)
  ic_lower[i] = t_test$conf.int[1]
  ic_upper[i] = t_test$conf.int[2]
}

plot(1:length(materiales), medias, ylim = c(min(ic_lower), max(ic_upper)), pch = 19, xlab = "Material del Motor", ylab = "Vibración (micrones)", main = "Intervalos de Confianza de Vibración por Material", xaxt = "n")
axis(1, at = 1:length(materiales), labels = materiales)
arrows(1:length(materiales), ic_lower, 1:length(materiales), ic_upper, angle = 90, code = 3, length = 0.1)
points(1:length(materiales), medias, pch = 19)
```

Intervalos de Confianza de Vibración por Material



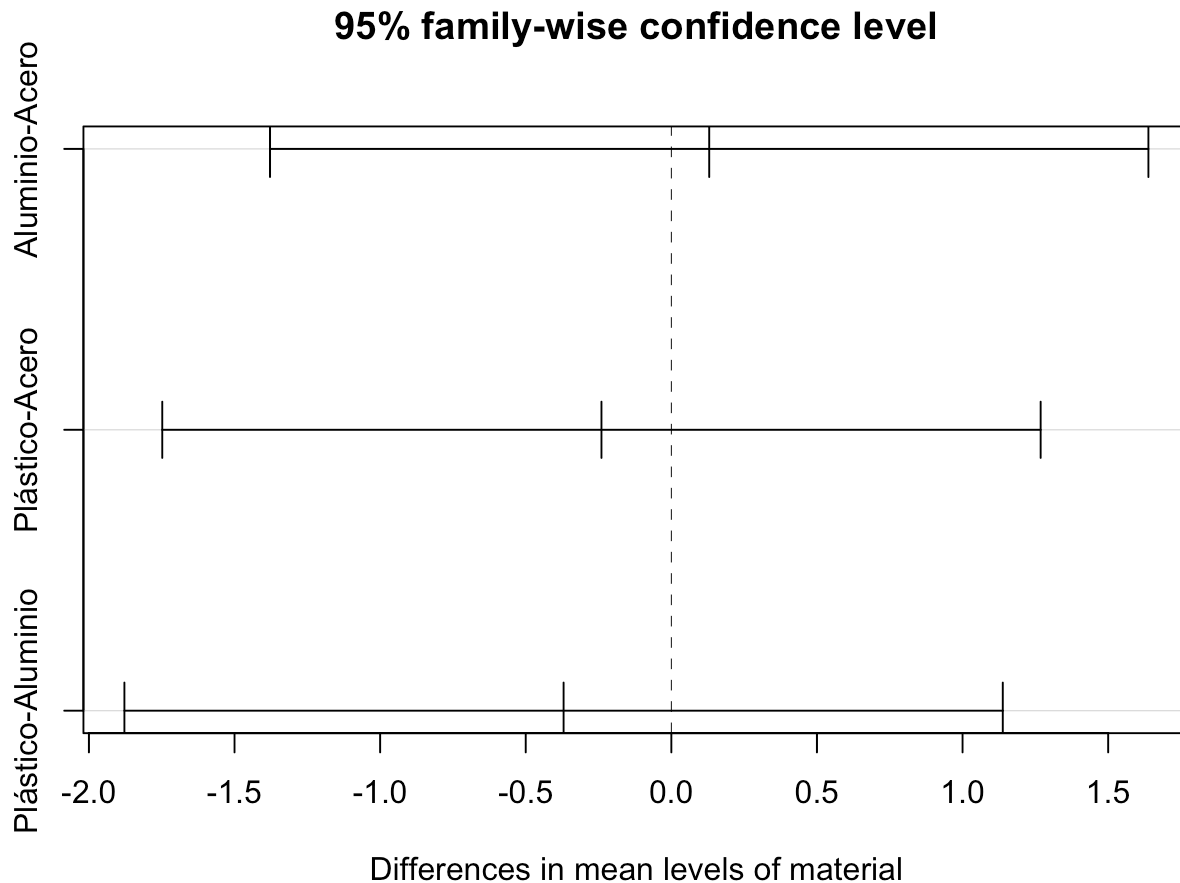
- Realiza la prueba de comparaciones múltiples de Tukey. Grafica los intervalos de confianza de Tukey.

```
anova_material = aov(vibracion ~ material)
tukey_resultados = TukeyHSD(anova_material)
print(tukey_resultados)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = vibracion ~ material)
##
## $material
##
```

	diff	lwr	upr	p adj
Aluminio-Acero	0.13	-1.378171	1.638171	0.9751575
Plástico-Acero	-0.24	-1.748171	1.268171	0.9180284
Plástico-Aluminio	-0.37	-1.878171	1.138171	0.8168495

```
plot(tukey_resultados)
```



6. Comprueba la validez del modelo. Comprueba:

- Normalidad
- Homocedasticidad
- Independencia
- Relación lineal entre las variables (coeficiente de determinación).

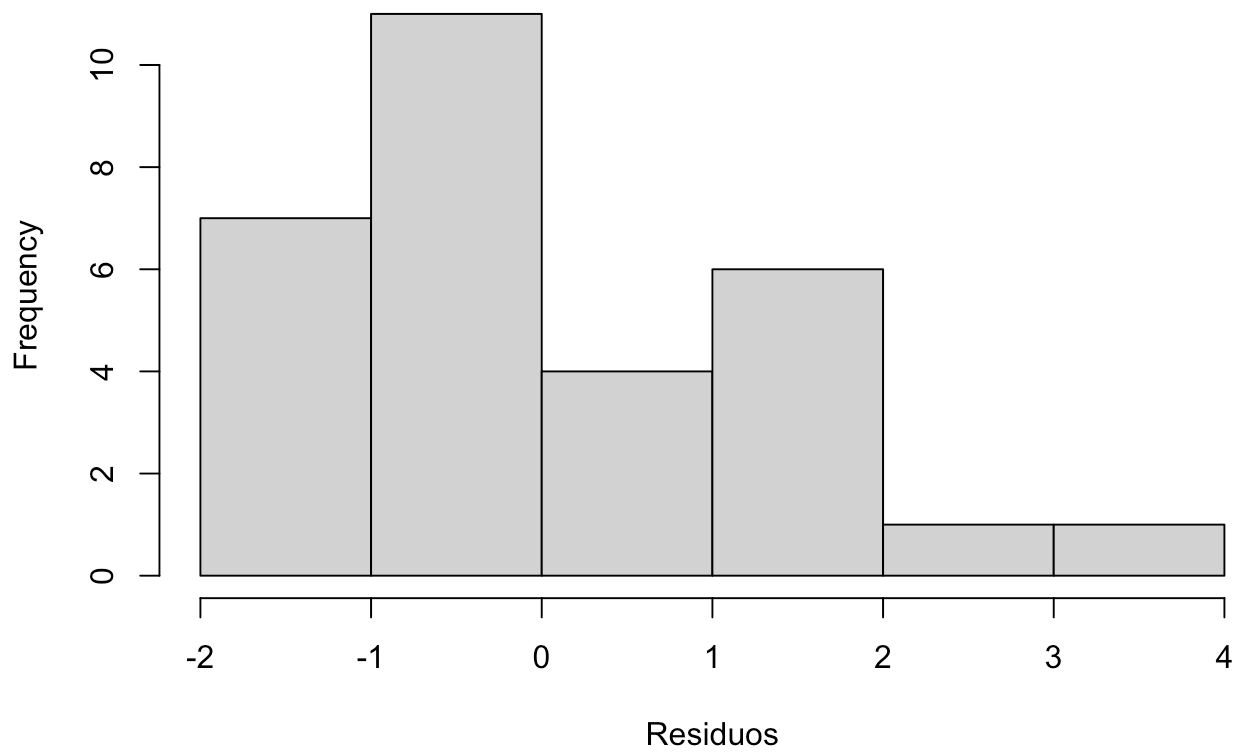
```
anova_material = aov(vibracion ~ material)
```

```
# 1. Comprobamos la normalidad
```

```
residuos = residuals(anova_material)
```

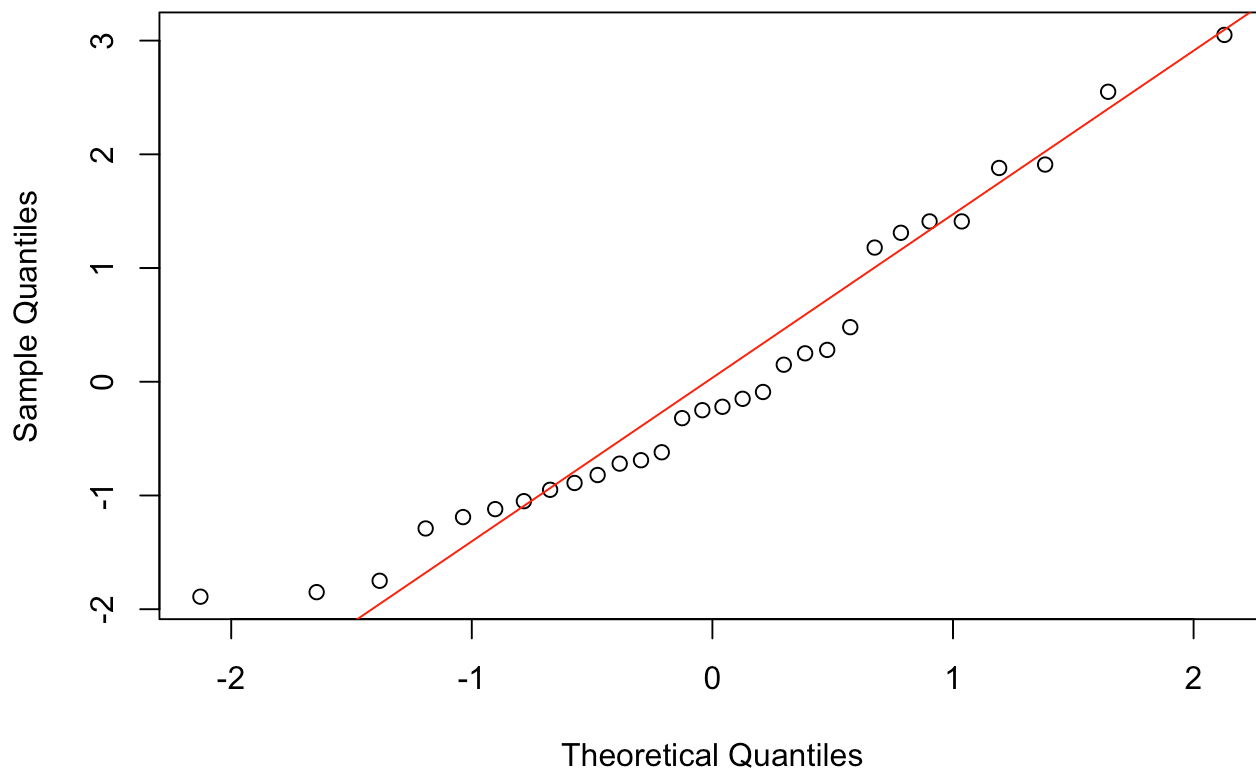
```
hist(residuos, main = "Histograma de los Residuos", xlab = "Residuos")
```

Histograma de los Residuos



```
qqnorm(residuos)  
qqline(residuos, col = "red")
```

Normal Q-Q Plot

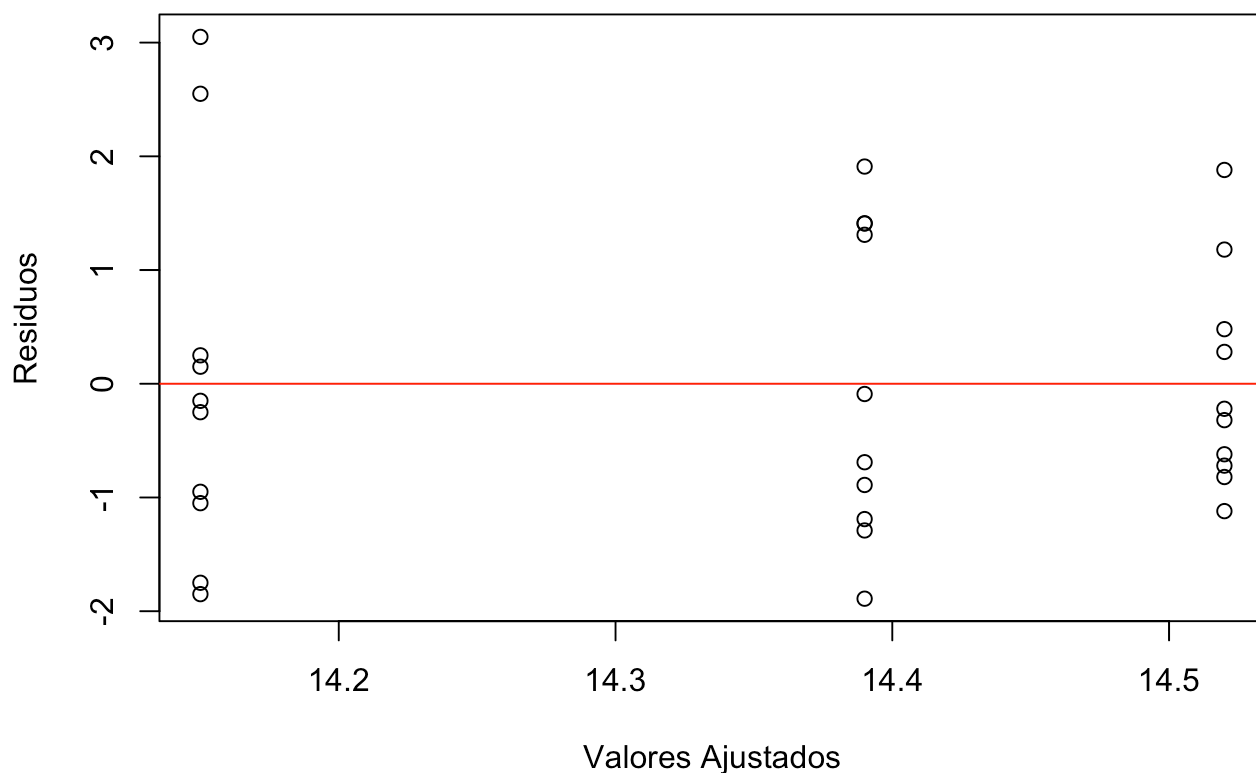


```
shapiro_test = shapiro.test(residuos)
print(shapiro_test)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuos
## W = 0.94362, p-value = 0.1139
```

```
# 2. Comprobamos la homocedasticidad
plot(fitted(anova_material), residuos, main = "Residuos vs Valores Ajustados", xlab = "V
alores Ajustados", ylab = "Residuos")
abline(h = 0, col = "red")
```

Residuos vs Valores Ajustados



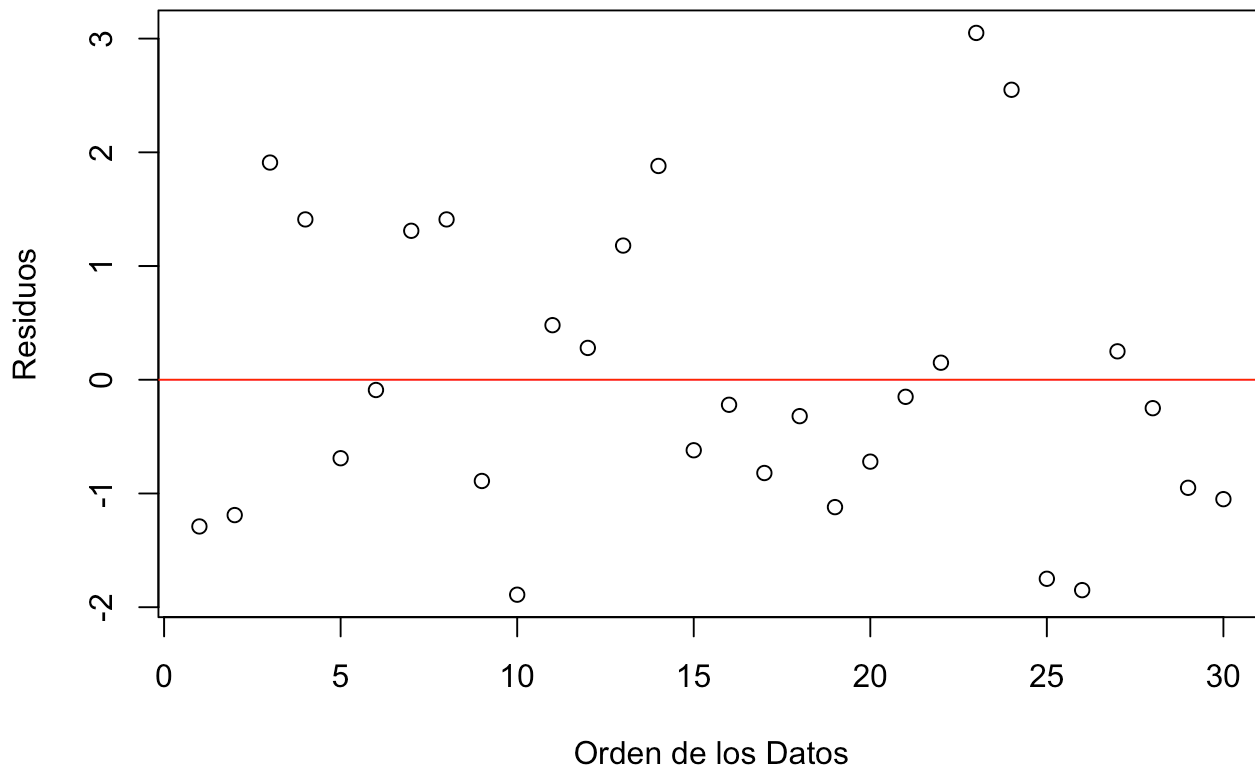
```
# Prueba de homocedasticidad de Bartlett
bartlett_test = bartlett.test(vibracion ~ material)
print(bartlett_test)
```

```
##
## Bartlett test of homogeneity of variances
##
## data:  vibracion by material
## Bartlett's K-squared = 2.4442, df = 2, p-value = 0.2946
```

3. Comprobamos la independencia vamos a observar los residuos en función del orden de los datos para ver si hay alguna tendencia.

```
plot(residuos, main = "Residuos en función del Orden de los Datos", xlab = "Orden de los Datos", ylab = "Residuos")
abline(h = 0, col = "red")
```

Residuos en función del Orden de los Datos



```
# 4. Relación lineal y Coeficiente de Determinación (R^2)
summary_anova = summary(anova_material)
suma_cuadrados_total = sum((vibracion - mean(vibracion))^2)
suma_cuadrados_residuales = sum(residuos^2)
R2 = 1 - (suma_cuadrados_residuales / suma_cuadrados_total)
print(paste("Coeficiente de Determinación (R^2):", R2))
```

```
## [1] "Coeficiente de Determinación (R^2): 0.0139111894922459"
```