

# Proba\_Anova\_fer

Fernanda Pérez

2024-08-27

## Anova

```
#install.packages("tidyverse")
```

```
library(tidyverse)
```

```
## — Attaching core tidyverse packages —————
tidyverse 2.0.0 —
## ✓ dplyr      1.1.4      ✓ readr      2.1.5
## ✓ forcats    1.0.0      ✓ stringr    1.5.1
## ✓ ggplot2    3.5.1      ✓ tibble     3.2.1
## ✓ lubridate  1.9.3      ✓ tidyr      1.3.1
## ✓ purrr      1.0.2
## — Conflicts —————
tidyverse_conflicts() —
## ✗ dplyr::filter() masks stats::filter()
## ✗ dplyr::lag()     masks stats::lag()
## ⓘ Use the conflicted package (<http://conflicted.r-lib.org/>) to
force all conflicts to become errors
```

## Problema1

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

Análisis exploratorio. Calcula la media para el rendimiento por método de enseñanza. Consulta el código en R en los apoyos de clase de “ANOVA” Haz el boxplot de resistencia a la tensión por concentración de madera dura. Interpreta el resultado desde la perspectiva estadística y en el contexto del problema. Escribe tus conclusiones parciales

## Análisis exploratorio. Calcula la media para el rendimiento por método de enseñanza.

```
calificacion <- c(10, 7, 9, 9, 9, 10, 5, 7, 6, 6, 8, 4, 2, 6, 3, 5, 3, 9,
7, 8, 8, 10, 6, 8, 8, 3, 5, 6, 7, 7, 2, 6, 2, 1, 4, 3)
metodo <- factor(c(rep("M1", 6), rep("M2", 6), rep("M3", 6), rep("M1",
6), rep("M2", 6), rep("M3", 6)))
sexo <- factor(c(rep("h", 18), rep("m", 18)))
datos <- data.frame(calificacion, metodo, sexo)
```

```
media_por_metodo <- aggregate(calificacion ~ metodo, FUN = mean)
print(media_por_metodo)
```

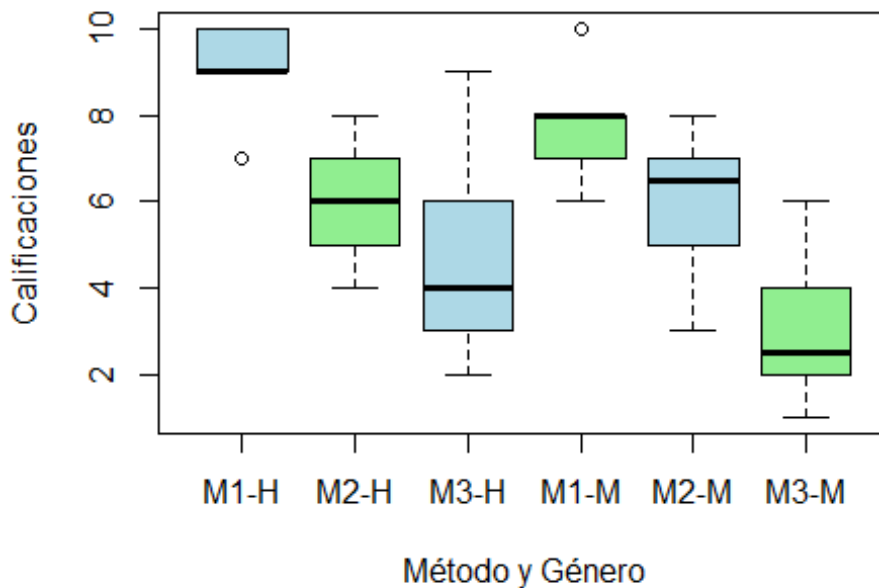
```
##  metodo calificacion
## 1      M1      8.416667
## 2      M2      6.000000
## 3      M3      3.833333
```

Con este primer análisis nos podemos dar una idea de los datos e información que estamos manejando. vemos que el método 1 (M1) tiene una mayor media de calificaciones, mientras que el método 3 (M3) tiene la media de calificaciones más baja.

### Boxplot

```
boxplot(calificacion ~ metodo * sexo,
        main = "Boxplot de Calificaciones por Método de Enseñanza y Género",
        xlab = "Método y Género",
        ylab = "Calificaciones",
        col = c("lightblue", "lightgreen"),
        names = c("M1-H", "M2-H", "M3-H", "M1-M", "M2-M", "M3-M"))
```

### Boxplot de Calificaciones por Método de Enseñanza y Género



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

En el boxplot vemos la distribución de las calificaciones de los estudiantes de acuerdo al método de enseñanza (M1, M2 y M3) y por género Hombres (H) y Mujeres (M).

Vemos que tanto en hombres como en mujeres el método 1 tiene una media de calificaciones más alta . Y el método 3 tanto en hombres como en mujeres tiene la media de calificaciones más baja.

Usaremos anova para hacer una comparación entre los diferentes métodos de enseñanza y poder decidir si es que existen diferencias significativas entre ellas.

Hasta el momento con la información que hemos obtenido nos damos cuenta que no afecta el genero sino el metodo que se este empleando.

#2. Las hipótesis. Establece las hipótesis estadísticas (tienen que ser 3).

Hipótesis 1:  $H_0$ :No hay diferencias significativas en el rendimiento entre los tres métodos de enseñanza.  $H_1$ :Por lo menos una de las medias será diferente

Hipótesis 2:  $H_0$ :El género no tiene un efecto significativo en el rendimiento.  $H_{1_1}$  El género si tiene un efecto significativo en el rendimiento.

Hipótesis 3:  $H_{0_0}$  No hay interacción entre el método de enseñanza y el género.  $H_1$ :Si existe interacción entre el método de enseñanza y el género.

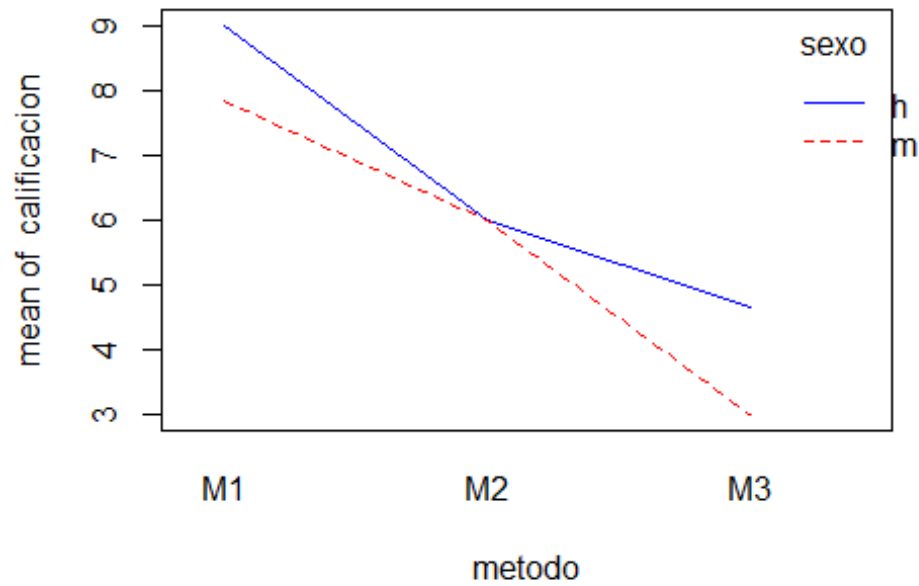
### 3. ANOVA para dos niveles con interacción

```
A <- aov(calificacion ~ metodo * sexo)
summary(A)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## metodo         2 126.17   63.08   20.989 1.99e-06 ***
## sexo           1   8.03    8.03    2.671    0.113
## metodo:sexo     2   4.39    2.19    0.730    0.490
## Residuals     30  90.17    3.01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

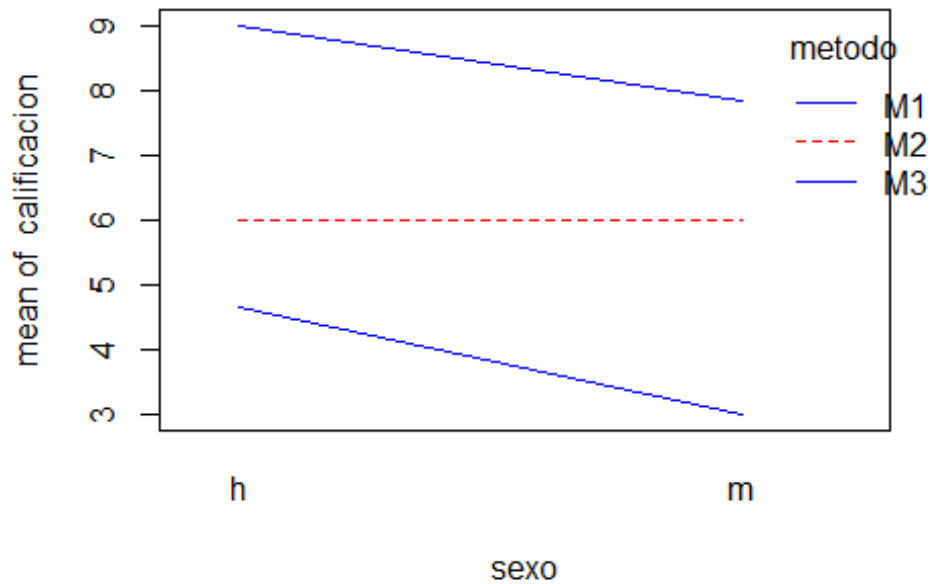
```
interaction.plot(metodo, sexo, calificacion,
                 col = c("blue", "red"),
                 lty = 1:2,
                 main = "Interacción entre Método y Género")
```

## Interacción entre Método y Género



```
interaction.plot(sexo, metodo, calificacion,  
                 col = c("blue", "red"),  
                 lty = 1:2,  
                 main = "Interacción entre Género y Método")
```

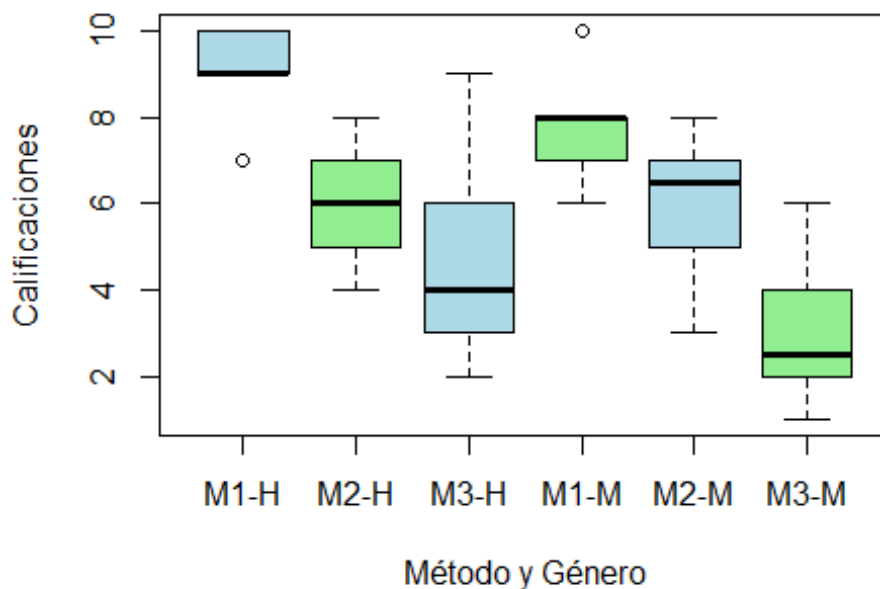
## Interacción entre Género y Método



##Haz el boxplot para visualizar la interacción de los factores

```
boxplot(calificacion ~ metodo * sexo,  
        main = "Boxplot de Calificaciones por Método de Enseñanza y  
Género",  
        xlab = "Método y Género",  
        ylab = "Calificaciones",  
        col = c("lightblue", "lightgreen"),  
        names = c("M1-H", "M2-H", "M3-H", "M1-M", "M2-M", "M3-M"))
```

## xplot de Calificaciones por Método de Enseñanza y (



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

```
anova_resultado <- aov(calificacion ~ metodo * sexo, data = datos)
```

```
cat("Interpretaciones:\n")
```

## Interpretaciones:

```
cat("1. Método de enseñanza (metodo):\n")
```

## 1. Método de enseñanza (metodo):

```
cat("F-value:", summary(anova_resultado)[[1]]$F.value[1], "\n")
```

## F-value: 20.98891

```
cat("p-value:", summary(anova_resultado)[[1]]$Pr(>F)[1], "\n")
```

## p-value: 1.989934e-06

```
cat("Interpretación: Existe una diferencia significativa en el  
rendimiento según el método de enseñanza. Por lo que el método de  
enseñanza que se elija si tiene un impacto directo en el rendimiento de  
los estudiantes.\n\n")
```

## Interpretación: Existe una diferencia significativa en el rendimiento según el método de enseñanza. Por lo que el método de enseñanza que se elija si tiene un impacto directo en el rendimiento de los estudiantes.

```

cat("2. Género (sexo):\n")
## 2. Género (sexo):
cat("F-value:", summary(anova_resultado)[[1]]$`F value`[2], "\n")
## F-value: 2.67098
cat("p-value:", summary(anova_resultado)[[1]]$`Pr(>F)`[2], "\n")
## p-value: 0.1126451

cat("Interpretación: El género no tiene un efecto significativo en el
rendimiento. \n\n")
## Interpretación: El género no tiene un efecto significativo en el
rendimiento.

cat("3. Interacción entre Método y Género (metodo:sexo):\n")
## 3. Interacción entre Método y Género (metodo:sexo):
cat("F-value:", summary(anova_resultado)[[1]]$`F value`[3], "\n")
## F-value: 0.7301294
cat("p-value:", summary(anova_resultado)[[1]]$`Pr(>F)`[3], "\n")
## p-value: 0.4902126

cat("Interpretación: No hay una interacción significativa entre el método
de enseñanza y el género.\n")
## Interpretación: No hay una interacción significativa entre el método
de enseñanza y el género.

```

### Escribe tus conclusiones parciales

Por los resultados obtenidos puedo deducir que el metodo de enseñanza si es un factor crucial , pero por otro lado el género y la interacción entre género y método no parece tener un impacto significativo.

## 4.Realiza el ANOVA para dos niveles sin interacción

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

```

calificacion <- c(10, 7, 9, 9, 9, 10, 5, 7, 6, 6, 8, 4, 2, 6, 3, 5, 3, 9,
7, 8, 8, 10, 6, 8, 8, 3, 5, 6, 7, 7, 2, 6, 2, 1, 4, 3)
metodo <- factor(c(rep("M1", 6), rep("M2", 6), rep("M3", 6), rep("M1",
6), rep("M2", 6), rep("M3", 6)))
sexo <- factor(c(rep("h", 18), rep("m", 18)))

```

```

anova_sin_interaccion <- aov(calificacion ~ metodo + sexo)
summary(anova_sin_interaccion)

##              Df Sum Sq Mean Sq F value    Pr(>F)    
## metodo         2 126.17   63.08   21.349 1.29e-06 ***
## sexo           1   8.03    8.03    2.717   0.109    
## Residuals     32  94.56    2.95                 
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

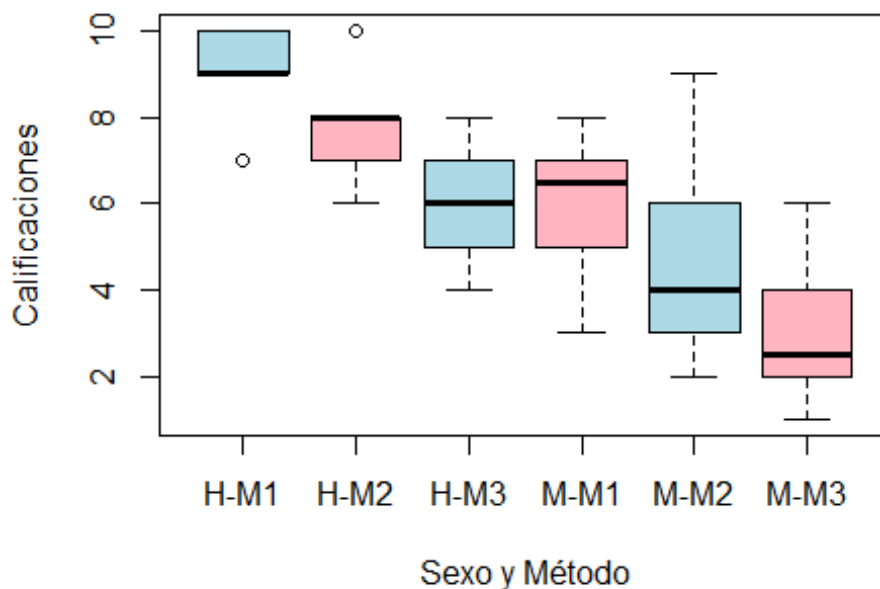
##Haz el boxplot de rendimiento por sexo. Calcula la media para el rendimiento por sexo y método.

```

boxplot(calificacion ~ sexo + metodo,
        main = "Boxplot de Calificaciones por Sexo y Método de
Enseñanza",
        xlab = "Sexo y Método",
        ylab = "Calificaciones",
        col = c("lightblue", "lightpink"),
        names = c("H-M1", "H-M2", "H-M3", "M-M1", "M-M2", "M-M3"))

```

## Boxplot de Calificaciones por Sexo y Método de Enseñanza



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

```

media_sd_sexo <- aggregate(calificacion ~ sexo, FUN = function(x) c(mean
= mean(x), sd = sd(x), n = length(x)))

medias <- media_sd_sexo$calificacion[, "mean"]
desviaciones <- media_sd_sexo$calificacion[, "sd"]

```



```

n <- media_sd_sexo$calificacion[, "n"]

error <- qt(0.975, df = n - 1) * (desviaciones / sqrt(n))

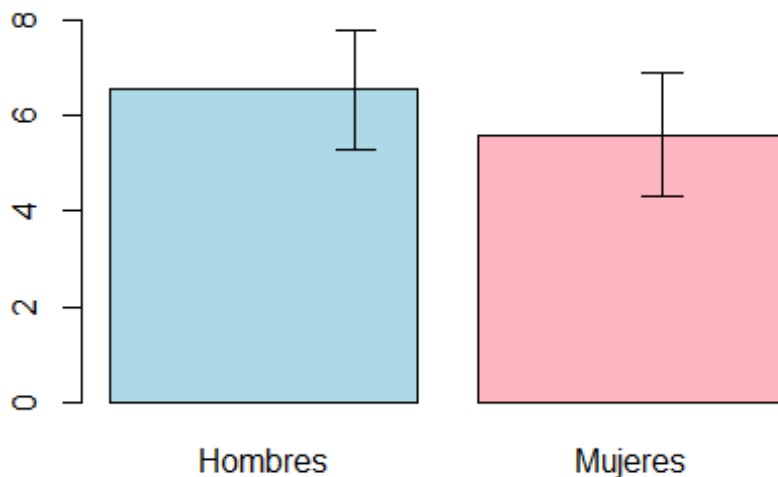
lower_ci <- medias - error
upper_ci <- medias + error

barplot(medias, beside = TRUE,
        ylim = c(0, max(upper_ci) + 1),
        names.arg = c("Hombres", "Mujeres"),
        col = c("lightblue", "lightpink"),
        main = "Medias de Rendimiento por Sexo con Intervalos de
Confianza")

arrows(1:length(medias), lower_ci,
       1:length(medias), upper_ci,
       angle = 90, code = 3, length = 0.1)

```

## Medias de Rendimiento por Sexo con Intervalos de Confianza



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

```

cat("Interpretaciones:\n")

## Interpretaciones:

cat("1. Método de enseñanza (metodo):\n")

```

```
## 1. Método de enseñanza (metodo):

cat("F-value:", summary(anova_sin_interaccion)[[1]]$`F value`[1], "\n")
## F-value: 21.349

cat("p-value:", summary(anova_sin_interaccion)[[1]]$`Pr(>F)`[1], "\n")
## p-value: 1.286636e-06

cat("Interpretación: El método de enseñanza si tiene un efecto
significativo en el rendimiento de los estudisntes \n\n")

## Interpretación: El método de enseñanza si tiene un efecto
significativo en el rendimiento de los estudisntes

cat("2. Género (sexo):\n")

## 2. Género (sexo):

cat("F-value:", summary(anova_sin_interaccion)[[1]]$`F value`[2], "\n")
## F-value: 2.716804

cat("p-value:", summary(anova_sin_interaccion)[[1]]$`Pr(>F)`[2], "\n")
## p-value: 0.1090821

cat("Interpretación: Por los resultados deducimos que le género no tiene
un impacto relevante en las calificaciones de los estudiantes \n\n")

## Interpretación: Por los resultados deducimos que le género no tiene un
impacto relevante en las calificaciones de los estudiantes

cat("Conclusiones parciales:\n")

## Conclusiones parciales:

cat("De acuerdo a los p-valores obtenidos, vemos que no es tan
importante centrarse en el género sino en los métodos de enseñanza.\n")

## De acuerdo a los p-valores obtenidos, vemos que no es tan importante
centrarse en el género sino en los métodos de enseñanza.
```

#5. Realiza el ANOVA para un efecto principal

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

```
calificacion <- c(10, 7, 9, 9, 9, 10, 5, 7, 6, 6, 8, 4, 2, 6, 3, 5, 3, 9,
7, 8, 8, 10, 6, 8, 8, 3, 5, 6, 7, 7, 2, 6, 2, 1, 4, 3)
metodo <- factor(c(rep("M1", 12), rep("M2", 12), rep("M3", 12)))

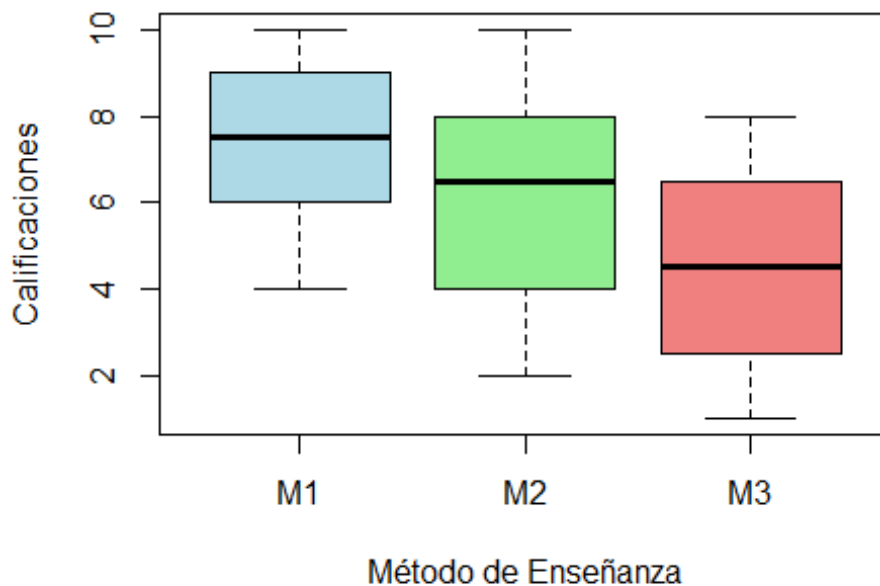
anova_efecto_principal <- aov(calificacion ~ metodo)
summary(anova_efecto_principal)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## metodo      2   54.5   27.25   5.161 0.0112 *
## Residuals  33  174.2    5.28
## ---
## 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(calificacion ~ metodo,
        main = "Boxplot de Calificaciones por Método de Enseñanza",
        xlab = "Método de Enseñanza",
        ylab = "Calificaciones",
        col = c("lightblue", "lightgreen", "lightcoral"))
```

### Boxplot de Calificaciones por Método de Enseñanza



```
media_por_metodo <- aggregate(calificacion ~ metodo, FUN = mean)
print(media_por_metodo)
```

```
##  metodo calificacion
## 1     M1           7.50
## 2     M2           6.25
## 3     M3           4.50
```

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

```
media_sd_metodo <- aggregate(calificacion ~ metodo, FUN = function(x)
c(mean = mean(x), sd = sd(x), n = length(x)))
```

```
medias <- media_sd_metodo$calificacion[, "mean"]
```

```

desviaciones <- media_sd_metodo$calificacion[, "sd"]
n <- media_sd_metodo$calificacion[, "n"]

error <- qt(0.975, df = n - 1) * (desviaciones / sqrt(n))

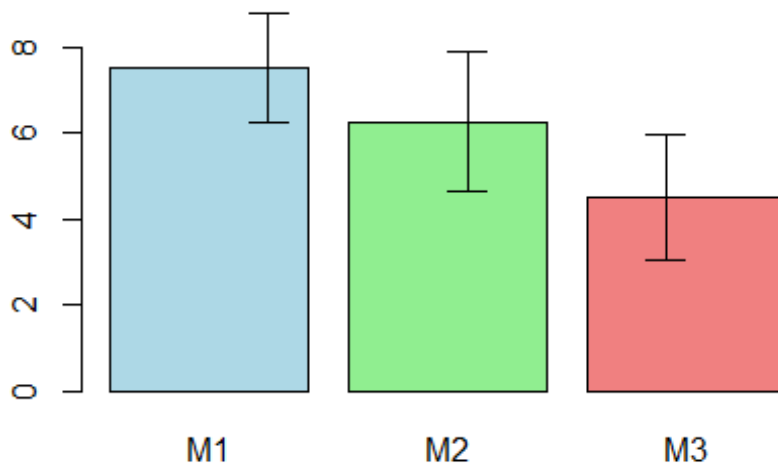
lower_ci <- medias - error
upper_ci <- medias + error

barplot(medias, beside = TRUE,
        ylim = c(0, max(upper_ci) + 1),
        names.arg = c("M1", "M2", "M3"),
        col = c("lightblue", "lightgreen", "lightcoral"),
        main = "Medias de Rendimiento por Método con Intervalos de
Confianza")

arrows(1:length(medias), lower_ci,
       1:length(medias), upper_ci,
       angle = 90, code = 3, length = 0.1)

```

## Medias de Rendimiento por Método con Intervalos de Confianza



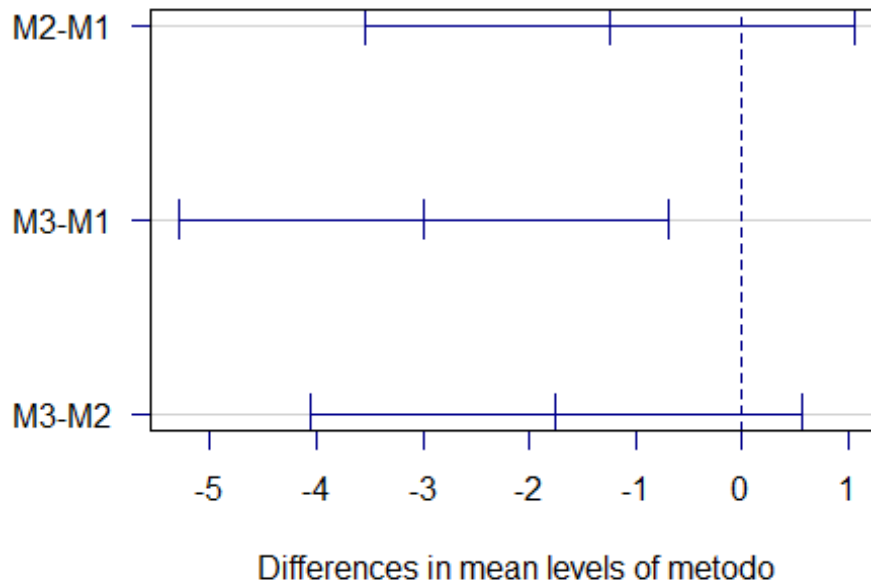
##Realiza la prueba de comparaciones múltiples de Tukey. Grafica los intervalos de confianza de Tukey.

```

tukey_result <- TukeyHSD(anova_efecto_principal)
plot(tukey_result, las = 1, col = "darkblue")

```

### 95% family-wise confidence level



```
print(tukey_result)

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = calificacion ~ metodo)
##
## $metodo
##      diff      lwr      upr    p adj
## M2-M1 -1.25 -3.551929  1.0519294 0.3876940
## M3-M1 -3.00 -5.301929 -0.6980706 0.0083475
## M3-M2 -1.75 -4.051929  0.5519294 0.1647216
```

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

```
cat("Interpretaciones:\n")
```

```
## Interpretaciones:
```

```
cat("Interpretación: Existe una diferencia significativa en el
rendimiento académico según el método de enseñanza utilizado.\n\n")
```

```
## Interpretación: Existe una diferencia significativa en el rendimiento
académico según el método de enseñanza utilizado.
```

```
cat("Conclusiones parciales:\n")
```

```
## Conclusiones parciales:
```

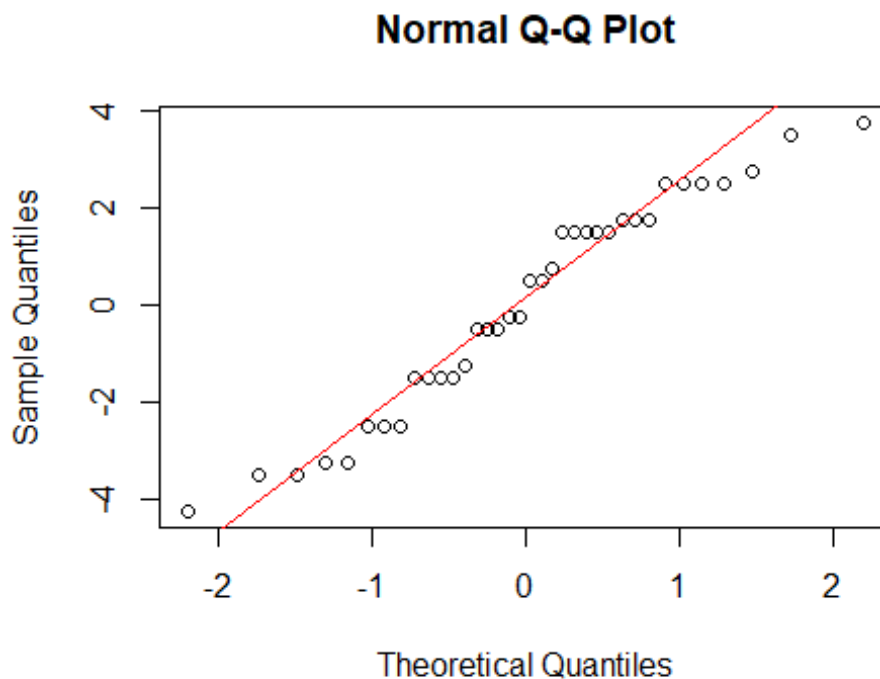
```
cat("por el analisis anova nos damos cuenta que el método de enseñanza  
tiene un impacto significativo en las calificaciones de los estudiantes.  
pareciera que el método 1 y el método 2 son más efectivos que el metodo 3  
\n")
```

```
## por el analisis anova nos damos cuenta que el método de enseñanza  
tiene un impacto significativo en las calificaciones de los estudiantes.  
pareciera que el método 1 y el método 2 son más efectivos que el metodo 3
```

## Comprueba la validez del modelo. Comprueba:

### Normalidad

```
qqnorm(residuals(anova_efecto_principal))  
qqline(residuals(anova_efecto_principal), col = "red")
```



```
shapiro_test <- shapiro.test(residuals(anova_efecto_principal))  
print(shapiro_test)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: residuals(anova_efecto_principal)  
## W = 0.95058, p-value = 0.1091
```

### Homocedasticidad

```
chooseCRANmirror(graphics = FALSE, ind = 1)
```

```

install.packages("lmtest")

## Installing package into 'C:/Users/Usuario/AppData/Local/R/win-
library/4.4'
## (as 'lib' is unspecified)

## package 'lmtest' successfully unpacked and MD5 sums checked

## Warning: cannot remove prior installation of package 'lmtest'

## Warning in file.copy(savedcopy, lib, recursive = TRUE): problema al
copiar
## C:\Users\Usuario\AppData\Local\R\win-
library\4.4\00LOCK\lmtest\libs\x64\lmtest.dll
## a C:\Users\Usuario\AppData\Local\R\win-
library\4.4\lmtest\libs\x64\lmtest.dll:
## Permission denied

## Warning: restored 'lmtest'

##
## The downloaded binary packages are in
## C:\Users\Usuario\AppData\Local\Temp\RtmpEpX3RN\downloaded_packages

library(lmtest)

## Cargando paquete requerido: zoo

##
## Adjuntando el paquete: 'zoo'

## The following objects are masked from 'package:base':
##
##   as.Date, as.Date.numeric

bp_test <- bptest(anova_efecto_principal)
print(bp_test)

##
## studentized Breusch-Pagan test
##
## data: anova_efecto_principal
## BP = 1.6907, df = 2, p-value = 0.4294

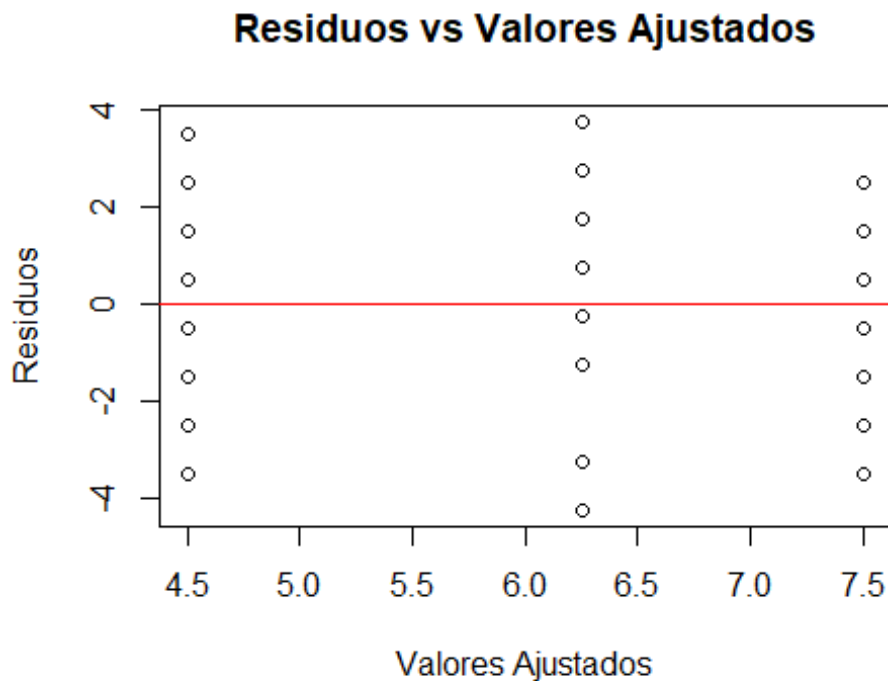
#install.packages("lmtest")

#library(lmtest)
#bp_test <- bptest(anova_efecto_principal)
#print(bp_test)

plot(fitted(anova_efecto_principal), residuals(anova_efecto_principal),
     main = "Residuos vs Valores Ajustados",
     xlab = "Valores Ajustados",

```

```
ylab = "Residuos")
abline(h = 0, col = "red")
```



```
library(lmtest)
bp_test <- bptest(anova_efecto_principal)
print(bp_test)

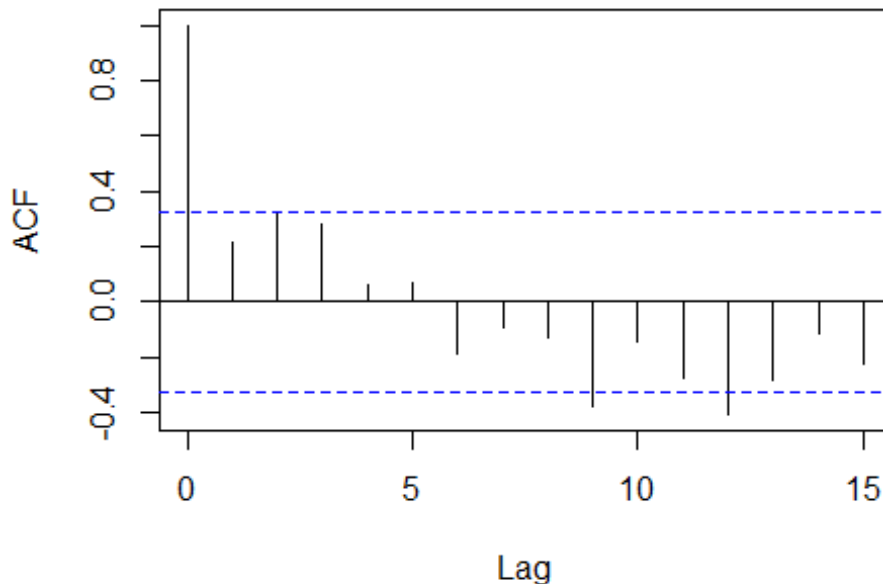
##
## studentized Breusch-Pagan test
##
## data: anova_efecto_principal
## BP = 1.6907, df = 2, p-value = 0.4294
```

#### Independencia:

```
acf(residuals(anova_efecto_principal), main = "Autocorrelación de los Residuos")
```



## Autocorrelación de los Residuos



### Relación lineal entre las variables (coeficiente de determinación).

```
SST <- sum((calificacion - mean(calificacion))^2)
SSM <- sum((fitted(anova_efecto_principal) - mean(calificacion))^2)
r_squared <- SSM / SST
cat("R²: ", r_squared, "\n")

## R²:  0.2382514
```

### Concluye en el contexto del problema.

Normalidad: con el grafico Q-Q vemos que los residuos siguen una distribución normal con desviaciones en los extremos.

Homocedasticidad: por los resultados obtenidos vemos en el grafico que no hay un patrón claro.

Independencia: vemos que no hay una autocorrelación significativa en los residuos.

$R^2$ : el 23.8% de la variabilidad en las calificaciones es explicada por el modelo.

Conclusión general:

Se verificó que el método de enseñanza si influye significativamente en el rendimiento academico, el género no afecta significativamente las calificaciones así que se puede decir que el género es independiente del rendimiento. Así que como conclusión se deben centrar en mejorar los métodos de enseñanza sin tener necesidad de adaptarse de acuerdo al género.

## Problema 2

### Análisis exploratorio

```
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 = 5))
proveedor <- factor(rep(1:5, times = 3))

datos <- data.frame(vibracion, material, proveedor)

media_por_material <- aggregate(vibracion ~ material, data = datos, FUN =
mean)
print(media_por_material)

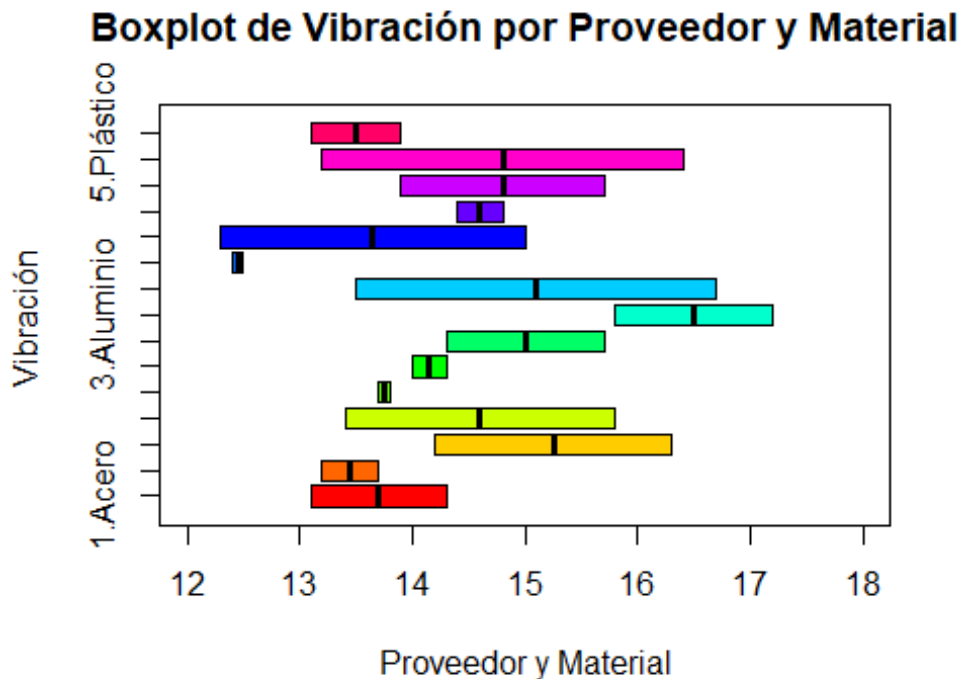
##   material vibracion
## 1   Acero      14.15
## 2 Aluminio      14.64
## 3 Plástico      14.27

media_por_proveedor <- aggregate(vibracion ~ proveedor, data = datos, FUN
= mean)
print(media_por_proveedor)

##   proveedor vibracion
## 1          1  13.83333
## 2          2  14.35000
## 3          3  15.51667
## 4          4  14.83333
## 5          5  13.23333
```

### Boxplot por material y proveedor

```
boxplot(vibracion ~ interaction(proveedor, material), data = datos,
        col = rainbow(length(levels(interaction(proveedor, material)))),
        main = "Boxplot de Vibración por Proveedor y Material",
        xlab = "Proveedor y Material",
        ylab = "Vibración ",
        ylim = c(12, 18),
        horizontal = TRUE)
```



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

La variación que hay en la media de de vibraciones entre los materiales es muy pequeña , el aluminio es el que presenta mayor medida de vibración.

Por otro lado hay una mayor variabilidad entre la media de los proveedores, el proveedor 3 muestra mayor media de vibración.

Vemos que el aluminio causa más vibración que el acero y el plástico, mientras que el proveedor 3 causa más vibraciones. Tanto el material como el proveedor influyen significativamente en la vibración de los motores eléctricos.

##Las hipótesis. Establece las hipótesis estadísticas (tienen que ser 3).

Hipotesis 1: H0: No hay diferencias significativas en la vibración media entre los diferentes materiales utilizados para la carcasa del motor (Acero, Aluminio, Plástico).  
H1: Hay por lo menos un material que genera una vibración media significativamente diferente.

Hipotesis2: H0: No hay una diferencia significativa en la vibración media entre los diferentes proveedores de cojinetes. H1: Hay por lo menos un proveedor que genera una vibración media significativamente diferente.

Hipotesis3: H0: No hay una interacción significativa entre el material y el proveedor en términos de la vibración media. H1: Si existe una interacción significativa entre el material y el proveedor

### 3: Realiza el ANOVA para dos niveles con interacción:

```
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 = 5))
proveedor <- factor(rep(1:5, times = 3))

anova_resultado <- aov(vibracion ~ material * proveedor, data = datos)
summary(anova_resultado)

##              Df Sum Sq Mean Sq F value Pr(>F)
## material      2  1.305   0.652   0.410 0.6707
## proveedor     4 18.651   4.663   2.933 0.0563 .
## material:proveedor 8  6.849   0.856   0.538 0.8102
## Residuals    15 23.850   1.590
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

### Haz la gráfica de interacción de dos factores en 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))
proveedor <- factor(rep(1:5, each = 6))

length(vibracion)

## [1] 30

length(material)

## [1] 30

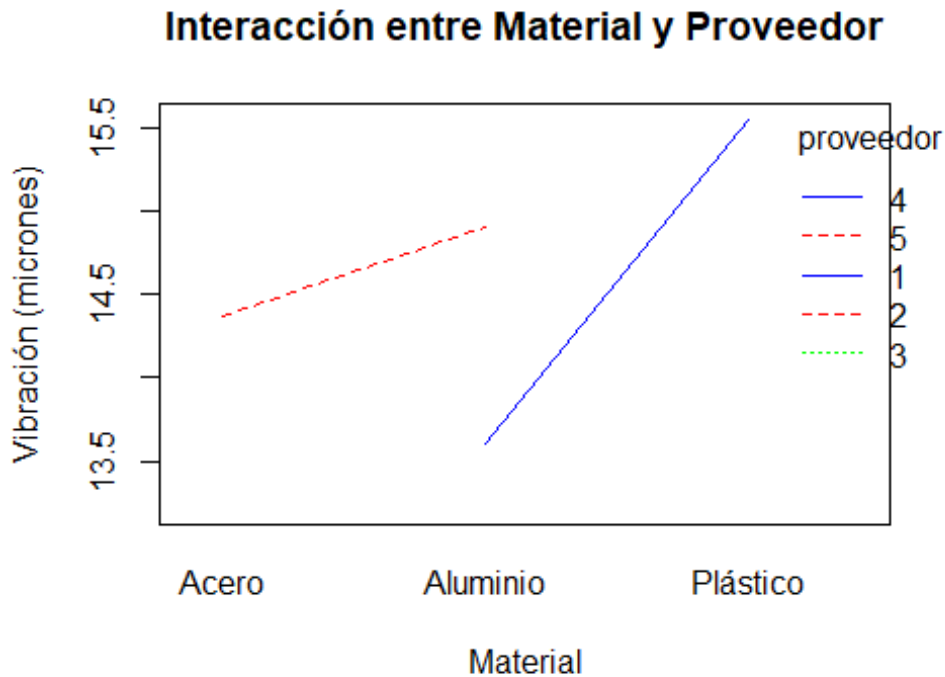
length(proveedor)

## [1] 30

datos <- data.frame(vibracion, material, proveedor)
print(head(datos))
```

```
##   vibracion material proveedor
## 1    13.1     Acero         1
## 2    13.2     Acero         1
## 3    16.3     Acero         1
## 4    15.8     Acero         1
## 5    13.7     Acero         1
## 6    14.3     Acero         1

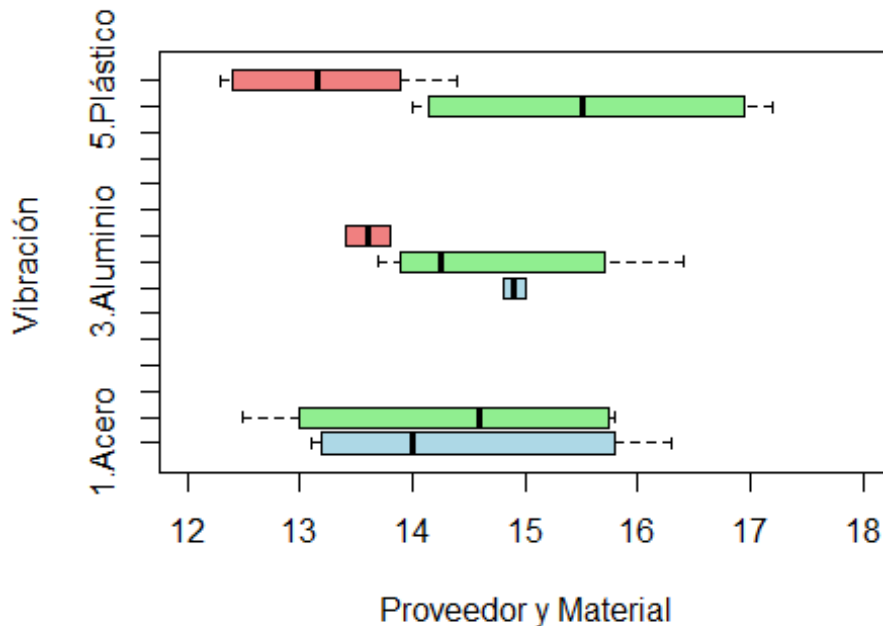
interaction.plot(material, proveedor, vibracion,
  col = c("blue", "red", "green"),
  lty = 1:3,
  main = "Interacción entre Material y Proveedor",
  xlab = "Material",
  ylab = "Vibración (micrones)")
```



##Haz el boxplot

```
boxplot(vibracion ~ interaction(proveedor, material), data = datos,
  col = c("lightblue", "lightgreen", "lightcoral"),
  main = "Boxplot de Vibración por Proveedor y Material",
  xlab = "Proveedor y Material",
  ylab = "Vibración ",
  ylim = c(12, 18),
  horizontal = TRUE)
```

## Boxplot de Vibración por Proveedor y Material



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

Vemos que el material no afecta de manera significativa la vibración. Vemo un poco que el proveedor muestra una pequeña tendencia a influir. NO se ve una interacción significativa entre el material y proveedor.

## 4. Realiza el ANOVA para dos niveles sin interacción

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

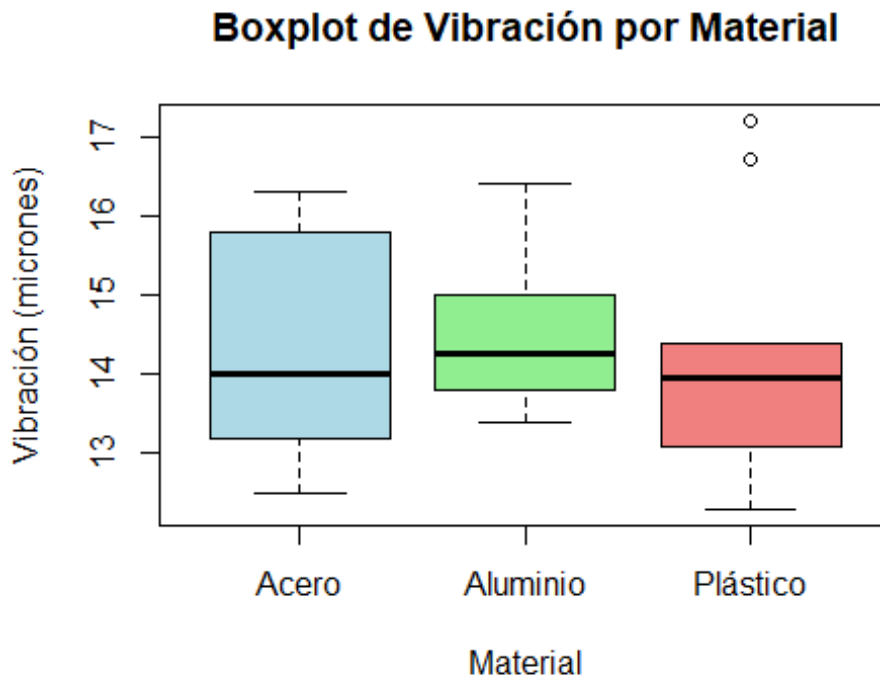
```
anova_sin_interaccion <- aov(vibracion ~ material + proveedor, data =
datos)
summary(anova_sin_interaccion)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
## material	2	0.70	0.352	0.233	0.7936
## proveedor	4	15.24	3.811	2.526	0.0685 .
## Residuals	23	34.71	1.509		
## ---					

## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

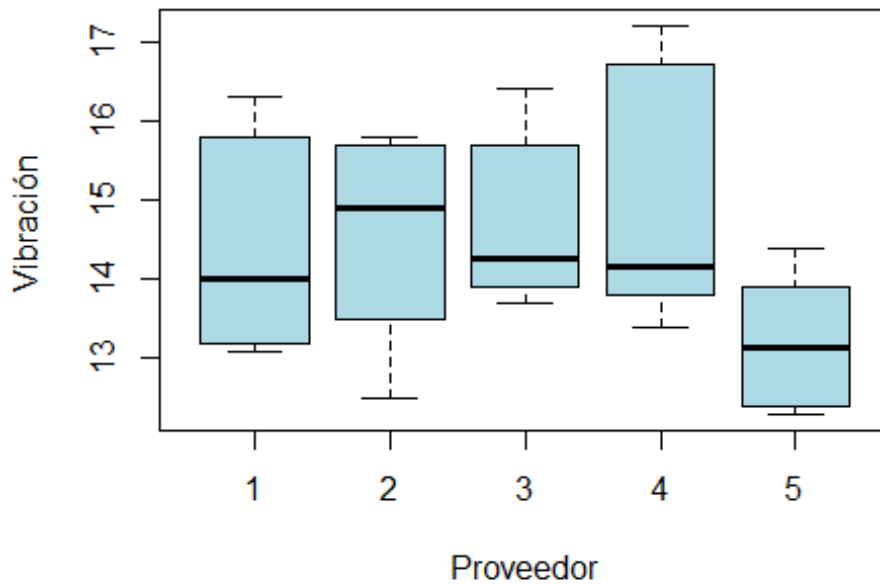
Haz el boxplot por Material y Proveedor. Calcula la Media para la Vibración por Material y Proveedor.

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



```
boxplot(vibracion ~ proveedor, data = datos,  
        main = "Boxplot de Vibración por Proveedor",  
        xlab = "Proveedor",  
        ylab = "Vibración ",  
        col = "lightblue")
```

## Boxplot de Vibración por Proveedor



```
media_por_material <- aggregate(vibracion ~ material, data = datos, FUN =
mean)
print(media_por_material)

##   material vibracion
## 1   Acero      14.39
## 2 Aluminio     14.52
## 3 Plástico     14.15

media_por_proveedor <- aggregate(vibracion ~ proveedor, data = datos, FUN
= mean)
print(media_por_proveedor)

##   proveedor vibracion
## 1         1  14.40000
## 2         2  14.55000
## 3         3  14.70000
## 4         4  14.90000
## 5         5  13.21667
```

### Haz los intervalos de confianza de rendimiento

```
media_sd_material <- aggregate(vibracion ~ material, data = datos, FUN =
function(x) c(mean = mean(x), sd = sd(x), n = length(x)))
medias <- media_sd_material$vibracion[, "mean"]
desviaciones <- media_sd_material$vibracion[, "sd"]
n <- media_sd_material$vibracion[, "n"]
```



```

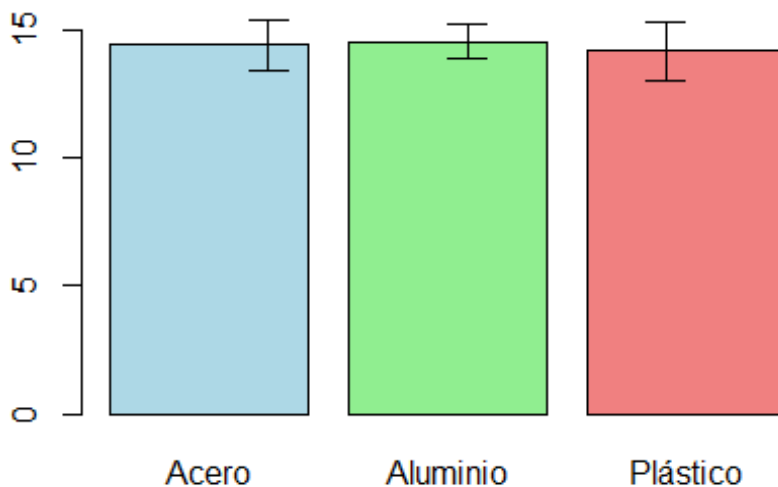
error <- qt(0.975, df = n - 1) * desviaciones / sqrt(n)

lower_ci <- medias - error
upper_ci <- medias + error

barplot(medias, beside = TRUE,
        ylim = c(0, max(upper_ci) + 1),
        names.arg = c("Acero", "Aluminio", "Plástico"),
        col = c("lightblue", "lightgreen", "lightcoral"),
        main = "Medias de Vibración por Material con Intervalos de
Confianza")
arrows(1:length(medias), lower_ci, 1:length(medias), upper_ci,
       angle = 90, code = 3, length = 0.1)

```

### Medias de Vibración por Material con Intervalos de Confianza



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

El material no se ve que afecte significativamente la media de la vibración, hay variaciones entre los proveedores y vemos que posiblemente el proveedor tendrá un mayor impacto en la vibración que el material.

## 5. Realiza el ANOVA para un efecto principal

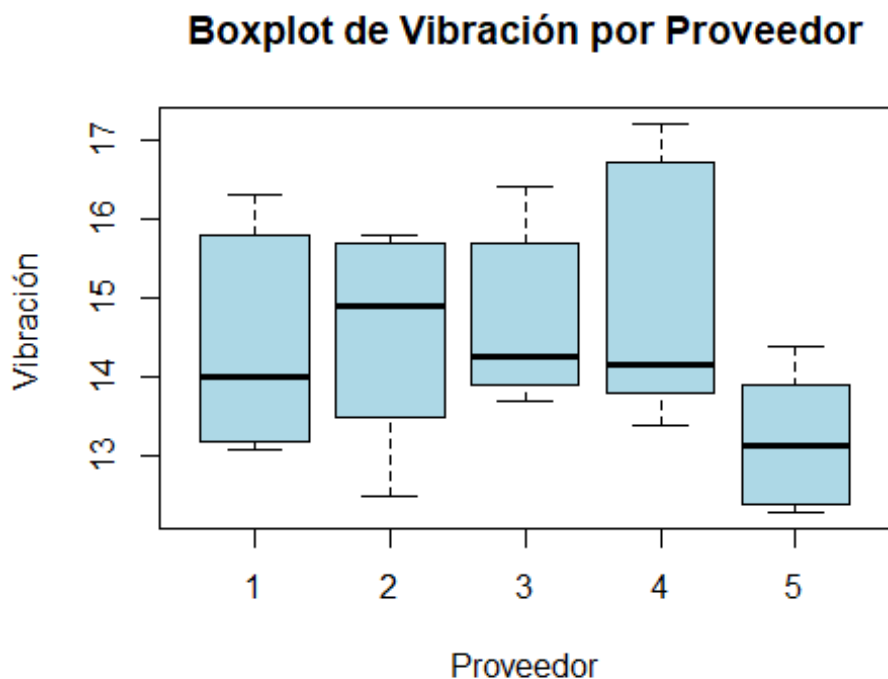
Consulta el código de R en los apoyos de clase de “ANOVA”

```
anova_efecto_principal <- aov(vibracion ~ proveedor, data = datos)
summary(anova_efecto_principal)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## proveedor   4  10.51   2.628   1.637  0.196
## Residuals  25  40.14   1.606
```

**Boxplot de Vibración por Proveedor. Calcula la Media.**

```
boxplot(vibracion ~ proveedor, data = datos,
        main = "Boxplot de Vibración por Proveedor",
        xlab = "Proveedor",
        ylab = "Vibración ",
        col = "lightblue")
```



```
media_por_proveedor <- aggregate(vibracion ~ proveedor, data = datos, FUN
= mean)
print(media_por_proveedor)
```

```
## proveedor vibracion
## 1         1  14.40000
## 2         2  14.55000
## 3         3  14.70000
## 4         4  14.90000
## 5         5  13.21667
```

### Haz los Intervalos de Confianza de Vibración por Proveedor.

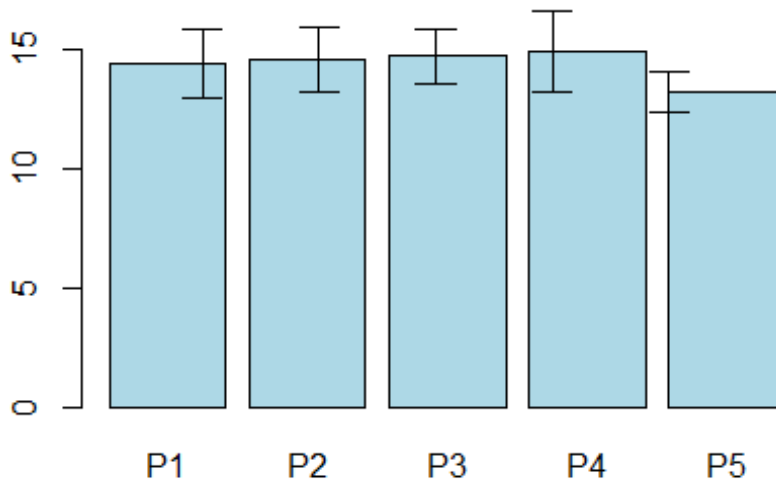
```
media_sd_proveedor <- aggregate(vibracion ~ proveedor, data = datos, FUN
= function(x) c(mean = mean(x), sd = sd(x), n = length(x)))
medias <- media_sd_proveedor$vibracion[, "mean"]
desviaciones <- media_sd_proveedor$vibracion[, "sd"]
n <- media_sd_proveedor$vibracion[, "n"]

error <- qt(0.975, df = n - 1) * desviaciones / sqrt(n)

lower_ci <- medias - error
upper_ci <- medias + error

barplot(medias, beside = TRUE,
        ylim = c(0, max(upper_ci) + 1),
        names.arg = c("P1", "P2", "P3", "P4", "P5"),
        col = "lightblue",
        main = "Medias de Vibración por Proveedor con Intervalos de
Confianza")
arrows(1:length(medias), lower_ci, 1:length(medias), upper_ci,
       angle = 90, code = 3, length = 0.1)
```

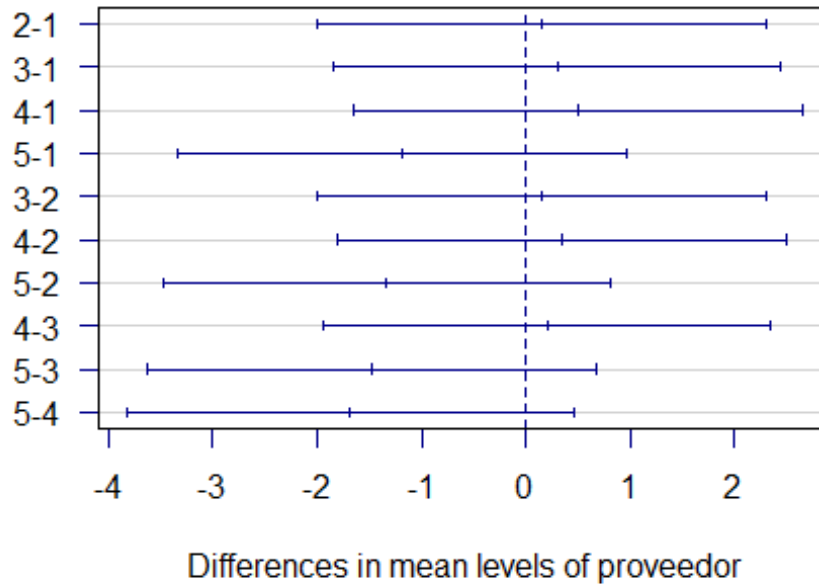
### Medias de Vibración por Proveedor con Intervalos de Confianza



### Realiza la prueba de comparaciones múltiples de Tukey

```
tukey_result <- TukeyHSD(anova_efecto_principal)
plot(tukey_result, las = 1, col = "darkblue")
```

### 95% family-wise confidence level



```
print(tukey_result)

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = vibracion ~ proveedor, data = datos)
##
## $proveedor
##      diff      lwr      upr    p adj
## 2-1  0.150000 -1.998627 2.2986268 0.9995773
## 3-1  0.300000 -1.848627 2.4486268 0.9936776
## 4-1  0.500000 -1.648627 2.6486268 0.9582384
## 5-1 -1.183333 -3.331960 0.9652935 0.5005896
## 3-2  0.150000 -1.998627 2.2986268 0.9995773
## 4-2  0.350000 -1.798627 2.4986268 0.9886564
## 5-2 -1.333333 -3.481960 0.8152935 0.3835892
## 4-3  0.200000 -1.948627 2.3486268 0.9986874
## 5-3 -1.483333 -3.631960 0.6652935 0.2825187
## 5-4 -1.683333 -3.831960 0.4652935 0.1778945
```

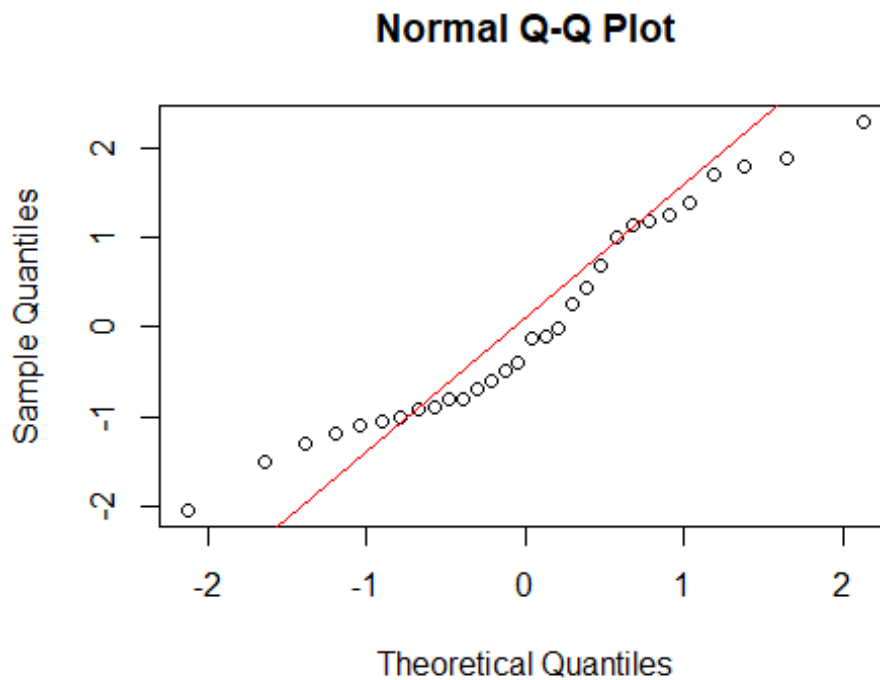
**Interpreta el resultado desde la perspectiva estadística y en el contexto del problema. Escribe tus conclusiones parciales**

Vemos que no hay diferencias significativas entre los proveedores, lo cual nos sugiere que el proveedor no genera un impacto considerable en la vibración.

## 6. Comprueba la validez del modelo. Comprueba:

### Normalidad

```
qqnorm(residuals(anova_efecto_principal))  
qqline(residuals(anova_efecto_principal), col = "red")
```

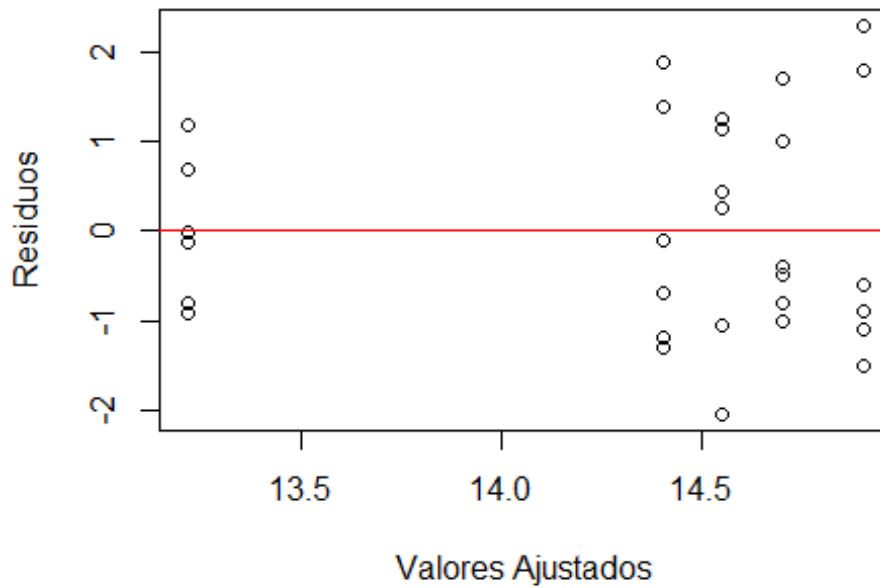


```
shapiro_test <- shapiro.test(residuals(anova_efecto_principal))  
print(shapiro_test)  
  
##  
##  Shapiro-Wilk normality test  
##  
## data:  residuals(anova_efecto_principal)  
## W = 0.94401, p-value = 0.1167
```

### Normalidad

```
plot(fitted(anova_efecto_principal), residuals(anova_efecto_principal),  
     main = "Residuos vs Valores Ajustados",  
     xlab = "Valores Ajustados",  
     ylab = "Residuos")  
abline(h = 0, col = "red")
```

## Residuos vs Valores Ajustados



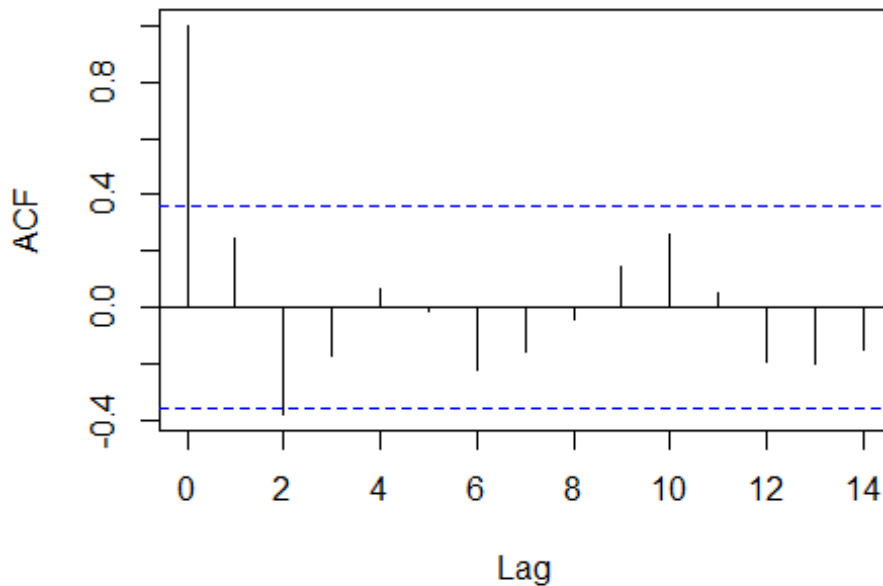
```
library(lmtest)
bp_test <- bptest(anova_efecto_principal)
print(bp_test)

##
## studentized Breusch-Pagan test
##
## data: anova_efecto_principal
## BP = 5.2302, df = 4, p-value = 0.2645
```

## Independencia

```
acf(residuals(anova_efecto_principal), main = "Autocorrelación de los  
Residuos")
```

## Autocorrelación de los Residuos



### Relación lineal entre las variables (coeficiente de determinación).

```
SST <- sum((datos$vibracion - mean(datos$vibracion))^2)
SSM <- sum((fitted(anova_efecto_principal) - mean(datos$vibracion))^2)
r_squared <- SSM / SST
cat("R²: ", r_squared, "\n")

## R²: 0.2075097
```

### Concluye en el contexto del problema.

Normalidad: en el grafico de Q-Q vemos que los residuos se desvían de la línea diagonal o sea que no siguen de manera certera una distribución normal.

Homocedasticidad: No vemos un patrón claro , o sea que la varianza de los residuos es constante

Independencia: Vemos que no hay autocorrelación significativa.

R<sup>2</sup>: solo el 20.8% de la variabilidad en la vibración se explica por el modelo

En un resumen general vemos que ni el material ni el proveedor tienen un impacto significativo en la vibración de los motores eléctricos en este experimento.