

**Universidad de Costa Rica**  
**Facultad de Ingeniería**  
**Escuela de Ciencias de la Computación e Informática**

**CI-0131 Diseño de Experimentos**

Dr. Ignacio Díaz Oreiro

Grupo 01

**Laboratorio 4: ANOVA de una vía**

Nathalie Alfaro Quesada, B90221

David Meléndez Aguilar, C04726

I ciclo 2025

## 1. Cargado de la base de datos

Se establece la columna motor como factor con el siguiente código:

```
db$motor <- as.factor(db$motor)
```

Al utilizar el código:

```
summary(db)
```

Aparece lo siguiente:

```
> summary(db)
      motor      respuesta
MongoDB  :12   Min.    : 8200
MySQL    :12   1st Qu.: 9840
PostgreSQL:12   Median :10622
Redis    :12   Mean    :10870
          :      3rd Qu.:12048
          :      Max.    :14698
          :
```

## 2. Revise las primeras 6 líneas del dataframe

Al ejecutar:

```
head(db)
```

Nos da el siguiente resultado:

```
> head(db)
      motor respuesta
1   MongoDB    13592
2     Redis    11851
3     MySQL    10334
4     Redis     9594
5 PostgreSQL   10926
6 PostgreSQL     8200
```

## 3. Calculamos la media y la desviación estándar para cada motor

Podemos obtener la media y la desviación estándar de cada motor al ejecutar el siguiente código:

```
db %>%
```

```
  group_by(motor) %>%
```

```
  summarise(mean = mean(respuesta), sd = sd(respuesta))
```

Lo cuál nos dará el siguiente resultado:

```

+           sd = sd(respuesta))
# A tibble: 4 × 3
  motor      mean    sd
  <fct>    <dbl> <dbl>
1 MongoDB  12484.  1373.
2 MySQL    10706.  1534.
3 PostgreSQL 10498.  1551.
4 Redis     9792.  1134.

```

Donde podemos ver que el motor MongoDB tiene la mayor media y Redis la menor, mientras que la desviación estándar más grande la tiene PostgreSQL, mientras que Redis vuelve a ser el valor más bajo, esta vez con la desviación estándar.

#### 4. Diagrama de cajas para cada uno de los cuatro motores para visualizar la distribución del tiempo de respuesta para cada motor

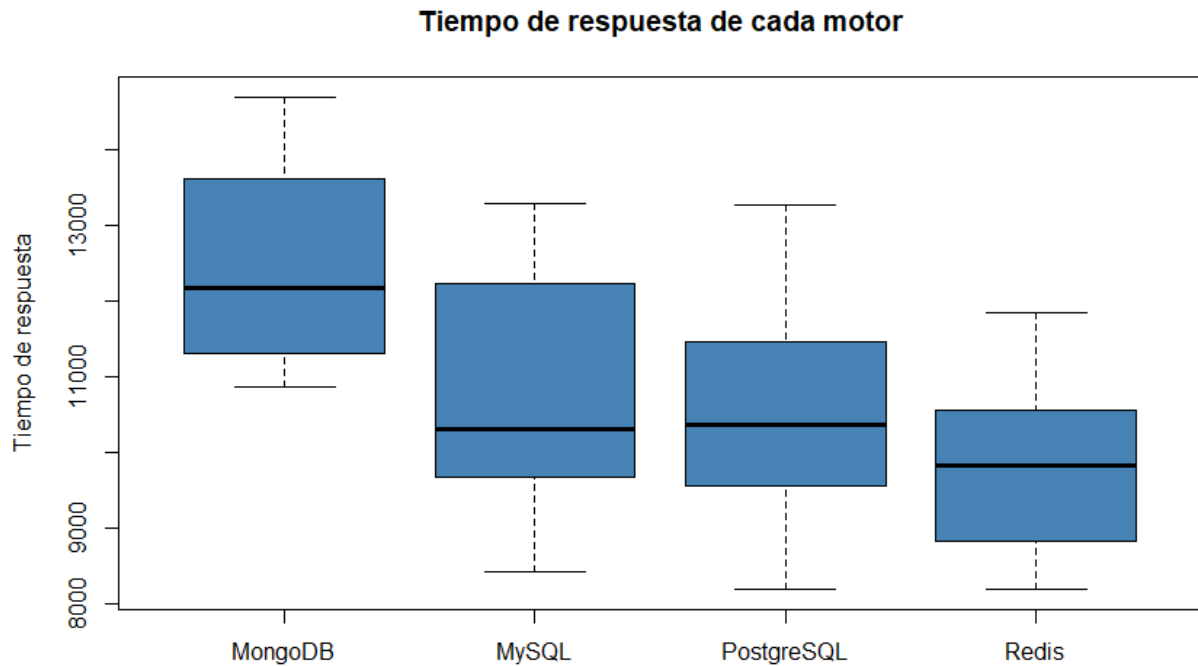
Podemos ejecutar este código:

```

boxplot(respuesta ~ motor,
        data = db, # data2,
        main = "Tiempo de respuesta de cada motor",
        xlab = "Motor",
        ylab = "Tiempo de respuesta",
        col = "steelblue",
        border = "black")

```

Para obtener el siguiente diagrama de cajas:



Donde vemos que la mediana del MongoDB tiene la mayor mediana y Redis la menor, y que el rango intercuartil de Redis es el más pequeño y el de MySQL el más grande.

## 5. Modelo de datos e hipótesis nula y alternativa

Se utilizará el siguiente modelo de datos Anova:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad \left\{ \begin{array}{l} i = 1, 2, 3 \\ j = 1, 2, \dots, 30 \end{array} \right\}$$

Donde  $\mu$  es un parámetro común a todos los motores (la media global) y  $\tau_i$  es la eficacia  $i$ -ésima, donde  $i = 1$  corresponde a MongoDB,  $i = 2$  corresponde a MySQL,  $i = 3$  corresponde a PostgreSQL e  $i = 4$  corresponde a Redis. Además,  $j = 1, 2, \dots, 30$  corresponde a cada una de las 48 repeticiones que se realizan. Adicionalmente, los errores del modelo ( $\varepsilon_{ij}$ ) representan una variable aleatoria que sigue una distribución normal e independiente con media 0 y varianza  $\sigma^2$ .

Seguidamente definimos la hipótesis alternativa y la hipótesis nula:

$$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$$

$$H_1 : \tau_i \neq 0 \text{ para algún } i \in \{1, 2, 3, 4\}$$

Es decir, la hipótesis nula indica que la eficacia es de 0 para los motores, y la hipótesis alternativa indica que alguno de los 4 motores sí tiene un efecto distinto de 0 en la variable de respuesta. Dicho de otra manera (con el modelo de las medias),

sería que las medias de los 4 motores son iguales, mientras que la hipótesis alternativa indica que al menos dos medias son distintas.

## 6. Ajuste del modelo ANOVA unidireccional

Con el siguiente código:

```
model <- aov(respuesta ~ motor, data = db)

summary(model)
```

Ajustamos el modelo ANOVA unidireccional usando el tiempo de respuesta y motor, y luego lo visualizamos.

```
> summary(model)
          Df    Sum Sq Mean Sq F value    Pr(>F)
motor      3 47158833 15719611   7.929 0.000246 ***
Residuals 44 87233088  1982570
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

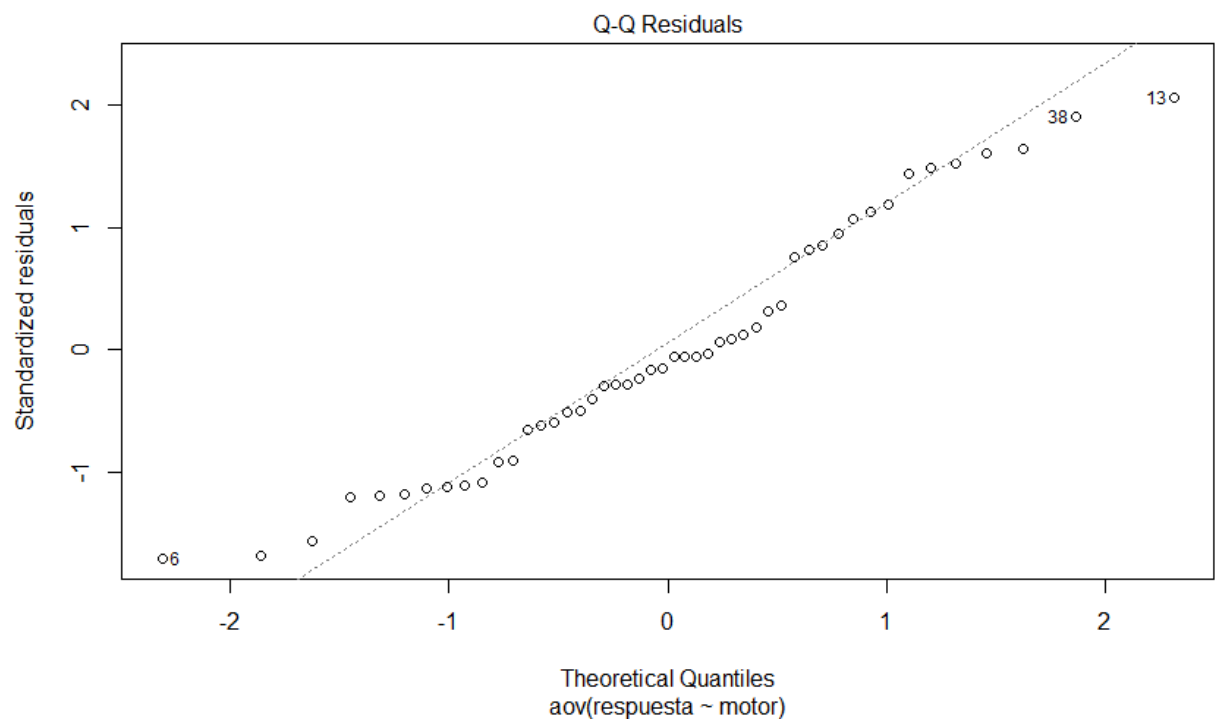
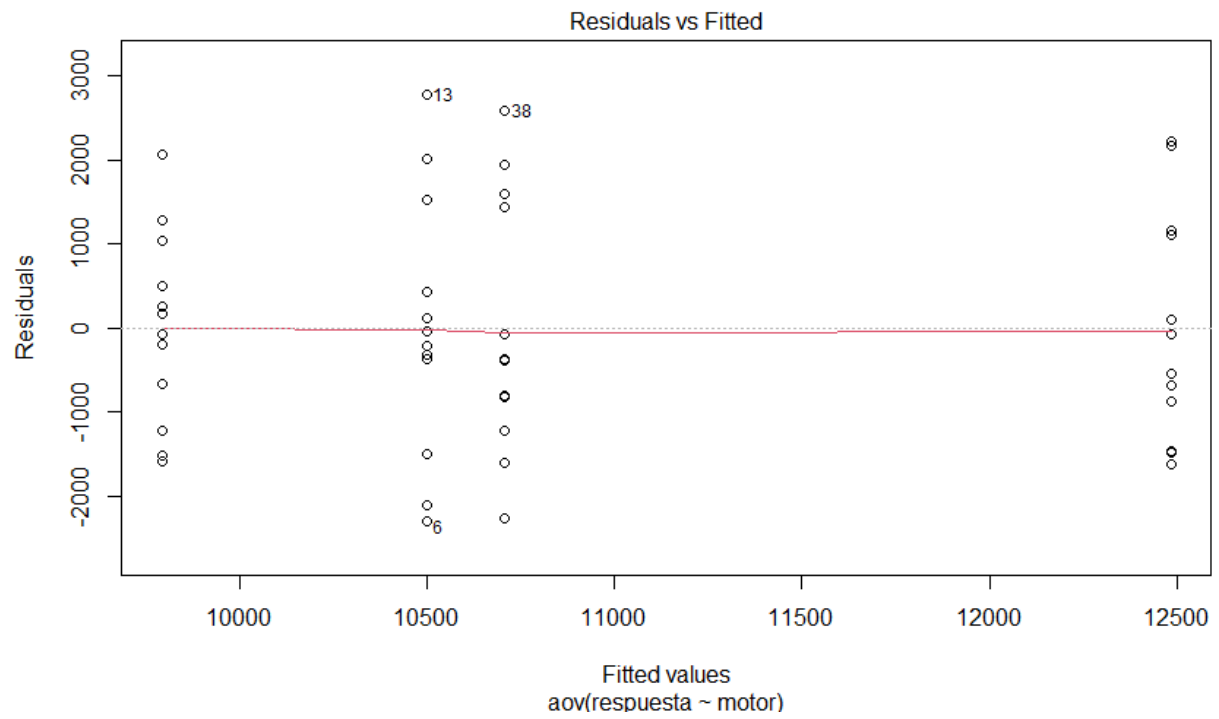
Teniendo el nivel de significancia de 0.05, 0.000246 es estadísticamente significativo al ser menor, por lo que se rechaza la hipótesis nula.

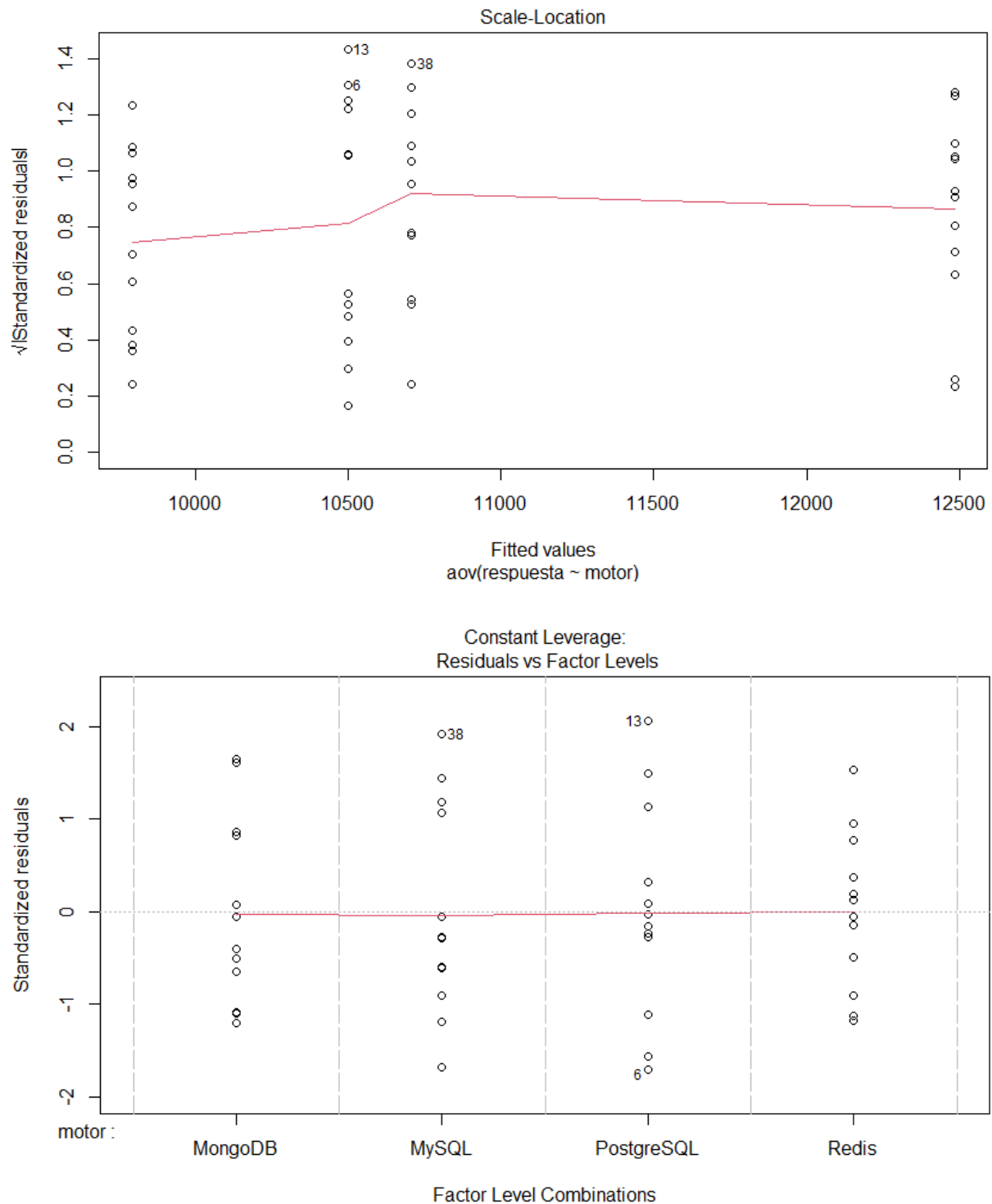
## 7. Comprobación de los supuestos del modelo

Debemos verificar que se cumplan los supuestos de nuestro modelo para que los resultados del modelo sean confiables. Para esto utilizaremos el código:

```
plot(model)
```

Para obtener los siguientes gráficos:





En Q-Q Residuals podemos ver que los residuos se desvían bastante de la línea en todo momento, por lo que podríamos pensar que la suposición de normalidad no se cumple.

En Residuals vs fitted tenemos una línea casi recta horizontal y cercana a 0, por lo que no hay indicios de problemas graves. De igual forma, podemos ver que los valores están más dispersos para los valores de MySQL y PostgreSQL que de los otros, por lo que la igualdad de varianzas podría no darse.

## 8. Prueba Shapiro Wilks sobre los residuales

Para probar formalmente el supuesto de normalidad debemos ejecutar la prueba Shapiro Wilks sobre los residuales.

### Código:

```
shapiro.test(model$residuals)
```

### Resultado:

```
Shapiro-Wilk normality test

data:  model$residuals
W = 0.96064, p-value = 0.1072
```

### Análisis:

No se puede rechazar la hipótesis nula de que los residuales siguen una distribución normal, para un nivel de significancia de 0.05, porque p es mayor.

## Pruebas de Levene o de Bartlett

También debemos probar formalmente la igualdad de varianzas, con las pruebas de Levene o de Bartlett.

### Código:

```
bartlett.test(respuesta ~ motor, data = bd_server)
```

### Resultado:

```
Bartlett test of homogeneity of variances

data:  respuesta by motor
Bartlett's K-squared = 1.2584, df = 3, p-value = 0.739
```

### Código:

```
library(car)

leveneTest(respuesta ~ motor, data = bd_server)
```

### Resultado:

```
Levene's Test for Homogeneity of Variance (center = median)

  Df F value Pr(>F)
group 3  0.2826 0.8377

44
```



### Análisis:

En ambos casos, el valor de p es mayor a 0.05, por lo cual no se puede rechazar la hipótesis nula de igualdad de varianzas.

## 9. Análisis de las diferencias de tratamiento

Para determinar cuáles medias son distintas realizamos una prueba post hoc para determinar exactamente qué grupos de tratamiento difieren entre sí. Para nuestra prueba post hoc, usaremos Tukey para comparaciones múltiples:

```
TukeyHSD(model, conf.level=.95)
```

### Resultado:

```
Tukey multiple comparisons of means
```

```
95% family-wise confidence level
```

```
Fit: aov(formula = respuesta ~ motor, data = bd_server)
```

```
$motor
```

	diff	lwr	upr	p adj
MySQL-MongoDB	-1777.6667	-3312.463	-242.8700	0.0174469
PostgreSQL-MongoDB	-1985.5000	-3520.297	-450.7034	0.0065296
Redis-MongoDB	-2691.0833	-4225.880	-1156.2867	0.0001565
PostgreSQL-MySQL	-207.8333	-1742.630	1326.9633	0.9836113
Redis-MySQL	-913.4167	-2448.213	621.3800	0.3952445
Redis-PostgreSQL	-705.5833	-2240.380	829.2133	0.6130353

### Análisis:

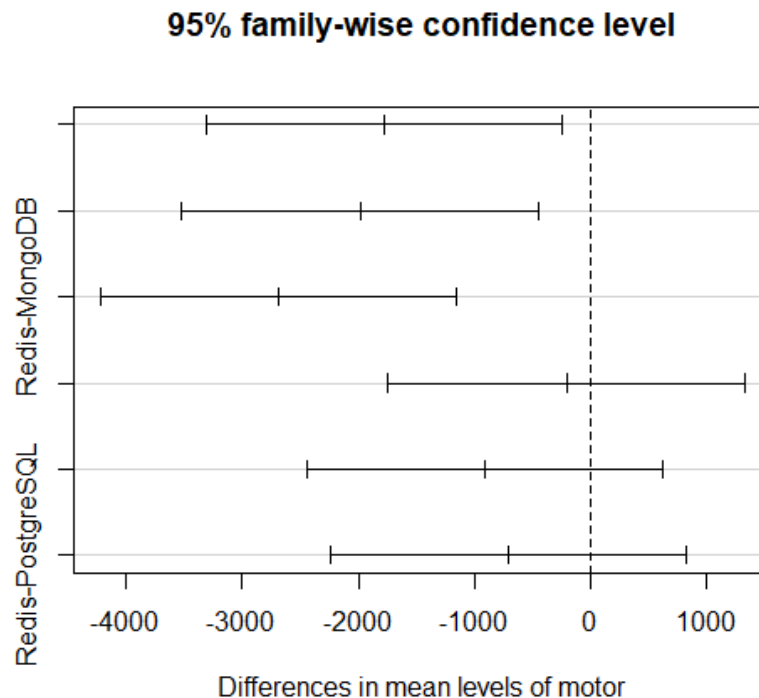
Se observa que las comparaciones entre MongoDB y los otros motores muestran diferencias estadísticamente significativas, ya que los valores p ajustados son menores a 0.05. Esto indica que MongoDB presenta una media significativamente diferente.

En cambio, las comparaciones entre MySQL, PostgreSQL y Redis entre sí no presentan diferencias significativas. Estos tres motores tienen medias similares en cuanto a la variable de respuesta.

## 10. Visualizar los intervalos de confianza del 95% que resultan de la prueba de Tukey usando la función plot(TukeyHSD()) en R

```
plot(TukeyHSD(model, conf.level=.95))
```

## Resultado:



El gráfico muestra los intervalos de confianza del 95% para las diferencias de medias entre los distintos niveles del factor motor.

El eje horizontal representa la diferencia entre medias de los pares de grupos comparados, cada línea es un par de motor, si el intervalo toca la vertical de 0 es que no hay diferencia significativa y sino entonces la diferencia es estadísticamente significativa (MongoDB es significativamente diferente).

### 11. Reportando los resultados de una ANOVA de una vía

Se realizó un ANOVA de una vía para ver la eficiencia de cuatro motores diferentes de bases de datos (PostgreSQL, MySQL, MongoDB y Redis) donde hubo una diferencia estadísticamente significativa entre la mitad de grupos ( $F(3, 44) = 7.929$ ,  $p = 0.000246$ ). Se realizaron pruebas post hoc HSD de Tukey.

Las comparaciones entre MongoDB y otros motores muestran una media con diferencias estadísticamente significativas, ya que los valores  $p$  ajustados son menores a 0.05.

En cambio, las comparaciones entre MySQL, PostgreSQL y Redis no presentan diferencias significativas. Estos tres motores tienen medias similares en cuanto a la variable de respuesta ya que sus valores para  $p$  son mayores a 0.05.

**12. También, en R puede obtenerse mediante la función etaSquared() de paquete lsr:**

```
#install.packages(lsr)

library(lsr)

etaSquared(model, anova=TRUE)
```

**Resultado:**

```
> etaSquared(model, anova=TRUE)
      eta.sq eta.sq.part      SS df      MS      F
motor    0.3509053    0.3509053 47158833  3 15719611 7.928905
Residuals 0.6490947           NA 87233088 44 1982570      NA

      p
motor 0.0002459618
Residuals      NA
> |
```

En este caso, puede verse que el efecto de la variable motor es grande, reflejado en el eta.sq = 0.35.

El 35% de la variabilidad total en la variable de respuesta se debe al tipo de motor.

El valor de p indica que la diferencia entre al menos dos tipos de motores es estadísticamente significativa.

### 13. Para determinar la potencia

```
power.anova.test(groups = 4, n = 30, between.var = 98.92, within.var = 139.56,
sig.level = 0.05)
```

**Resultado:**

```
Balanced one-way analysis of variance power calculation
```

```
groups = 4
```

```
n = 30
```

```
between.var = 98.92
```

```
within.var = 139.56
```

```
sig.level = 0.05
```

```
power = 0.9999999
```

NOTE: n is number in each group

### Análisis:

En este caso, con  $n = 30$  para cada grupo, y las varianzas identificadas entre grupos y dentro del grupo, la potencia de la prueba es 0.99, lo que es muy bueno.

**14. Para determinar el tamaño de  $n$  necesario (o suficiente) para una potencia de prueba específica, no se incluye el  $n$  y se agrega  $\text{power} = \text{al valor deseado}$ . Por ejemplo, si quiero encontrar el tamaño de  $n$  para una potencia de 0.85:**

```
power.anova.test(groups = 4, between.var = 98.92, within.var = 139.56, sig.level = 0.05, power = 0.85)
```

### Resultado:

```
Balanced one-way analysis of variance power calculation
```

```
groups = 4
```

```
n = 6.848857
```

```
between.var = 98.92
```

```
within.var = 139.56
```

```
sig.level = 0.05
```

```
power = 0.85
```

NOTE:  $n$  is number in each group

### Análisis:

Se necesitan 6.85 (o 7 redondeado) de muestra por grupo para alcanzar la potencia de 0.85, una alta probabilidad de detectar diferencias verdaderas con las varianzas identificadas entre grupos y dentro del grupo.

### En resumen

Se realizó un ANOVA de una vía para ver la eficiencia de cuatro motores diferentes de bases de datos (PostgreSQL, MySQL, MongoDB y Redis).

Un ANOVA de una vía a través de la prueba TukeyHSD con un gráfico, reveló que había una diferencia estadísticamente significativa de la media en el grupo de MongoDB, pero en el grupo de PostgreSQL, MySQL y Redis no se presenta diferencia significativa, donde  $(F(3, 44) = 7.929, p = 0.000246)$ .

La prueba HSD de Tukey para comparaciones múltiples encontró que el valor medio fue significativamente diferente entre Redis-MongoDB y MySQL-MongoDB ( $p = 0.0001565$ , IC del 95 % =  $([-4225.880, -1156.2867])$ ).

Además, en la prueba HSD de Tukey se encontró que no hubo una diferencia estadísticamente significativa entre PostgreSQL-MySQL y Redis-PostgreSQL ( $p=0.9836113$ ), IC del 95 % = [-1742.630, 1326.9633]).

También, para encontrar su potencia, se prueba con  $n = 30$  para cada grupo, y las varianzas identificadas entre grupos y dentro del grupo, la potencia de la prueba es 0.99, lo que es muy bueno, porque hay una alta probabilidad de detectar una diferencia entre grupos, lo que la hace más confiable.