

# Tarea12

Norberto Florez

20/5/2021

```
setwd("C:/Users/ASUS/Documents/GitHub/r-basic/data")
grades = read.table("grades.txt", header = TRUE)
head(grades)
```

```
##      estudio      nota
## 1 Industrial 4.811598
## 2 Telematica 3.825744
## 3 Telematica 3.518318
## 4      Mates 0.000000
## 5 Telematica 5.879675
## 6 Informatica 3.554049
```

```
str(grades)
```

```
## 'data.frame':    200 obs. of  2 variables:
## $ estudio: Factor w/ 4 levels "Industrial","Informatica",...: 1 4 4 3 4 2 3 2 4 4 ...
## $ nota   : num  4.81 3.83 3.52 0 5.88 ...
```

1. Calcula la media, la mediana y la moda redondeada a dos cifras decimales de las notas numéricas de los exámenes tanto de forma conjunta como por estudio universitario. ¿En qué grupo observamos la nota más alta? ¿Qué grupo está por encima de la media de estudiantes?

```
as.numeric(names(which(table(grades$nota)==max(table(grades$nota)))))
```

```
## [1] 0
```

```
round(summary(grades$nota),2)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.00    2.60    3.50    3.87    5.12   10.02
```

```
aggregate(grades$nota~grades$estudio, FUN = summary)
```

```
## grades$estudio grades$nota.Min. grades$nota.1st Qu. grades$nota.Median
## 1 Industrial 0.000000 2.622411 3.435440
## 2 Informatica 1.846116 2.961614 3.477870
## 3 Mates 0.000000 2.533527 4.101367
## 4 Telematica 0.000000 2.442873 3.409993
## grades$nota.Mean grades$nota.3rd Qu. grades$nota.Max.
## 1 3.577426 4.774713 10.024687
## 2 3.954445 4.992463 8.515653
## 3 4.215334 5.932406 8.539979
## 4 3.704336 4.816256 9.290465
```

```
by(grades$nota, grades$estudio, FUN = summary)
```

```
## grades$estudio: Industrial
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.000 2.622 3.435 3.577 4.775 10.025
## -----
## grades$estudio: Informatica
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 1.846 2.962 3.478 3.954 4.992 8.516
## -----
## grades$estudio: Mates
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.000 2.534 4.101 4.215 5.932 8.540
## -----
## grades$estudio: Telematica
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.000 2.443 3.410 3.704 4.816 9.290
```

Respuestas:

La mayor nota la obtuvieron los industriales.

El grupo de mates está por encima de la media.

## 2. ¿En qué grupo observamos más variación de notas? Justifica tu respuesta

```
by(grades$nota, grades$estudio, FUN = sd)
```

```
## grades$estudio: Industrial
## [1] 2.010948
## -----
## grades$estudio: Informatica
## [1] 1.537978
## -----
## grades$estudio: Mates
## [1] 2.195236
## -----
## grades$estudio: Telematica
## [1] 1.909706
```

Respuesta: En el grupo de mates existe la mayor variación de notas pues tiene la mayor desviación estándar.

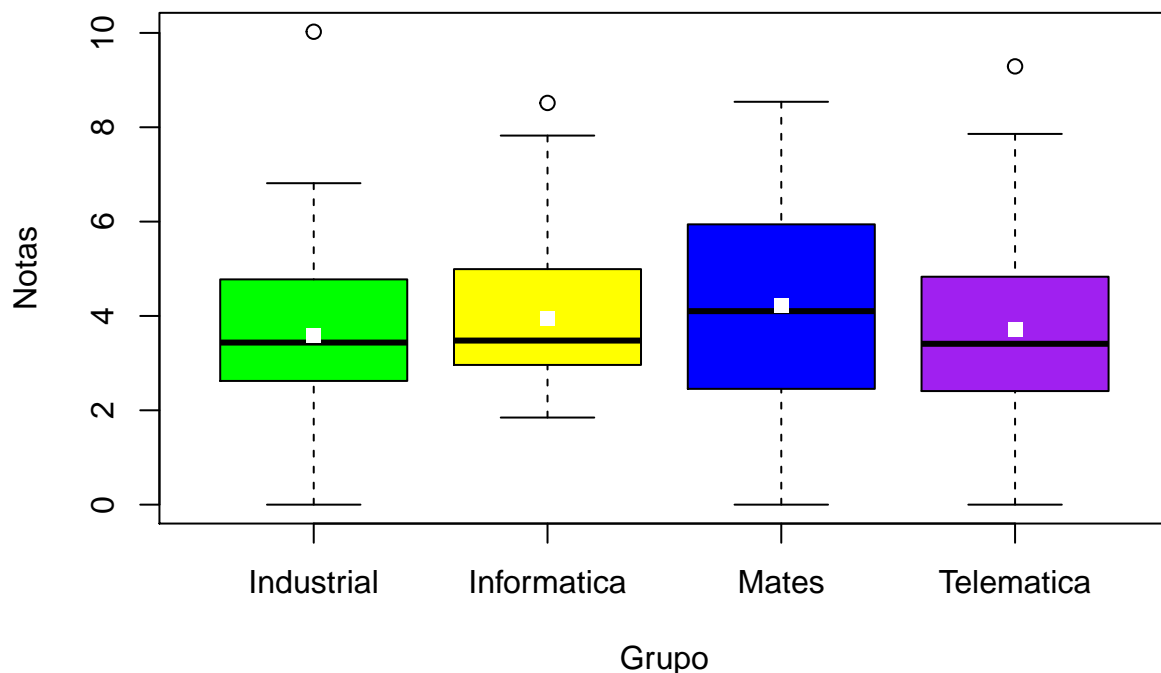
3. ¿Hay mucha diferencia entre el grupo que has respondido en el ejercicio 1 y en el ejercicio 2? Intenta dar una explicación objetiva del suceso

La principal diferencia es que la media de las notas del grupo de mates es considerablemente mayor a la del grupo de industriales. El tercer cuartil del grupo de mates está por encima de 5.93 mientras que el tercer cuartil del grupo de industriales está por encima de 4.77.

4. Dibuja en un único gráfico los cuatro diagramas de caja, uno para cada grupo. Añade así mismo la nota media, pon nombre, título y color al gráfico para documentar el gráfico correctamente.

```
boxplot(grades$nota~grades$estudio, ylab = "Notas", xlab = "Grupo", main = "Gráfico de cajas de notas de estudiantes",  
medias = by(grades$nota, grades$estudio, FUN = mean),  
points(medias, col = "white", pch = 15))
```

**Gráfico de cajas de notas de estudiantes según grupos de estudio**



5. ¿Observas algunos valores atípicos en el boxplot anterior? ¿A qué grupo pertenece?

Se ven outliers en los grupos de industriales, informática y telemática. El mínimo del grupo de informática empieza mas arriba casi en 2, y hay una gran cantidad de alumnos en el rango de 3 y 3.5 (valores aprox del primer cuartil y la media)

**6. Indica a partir del boxplot anterior en qué grupo observamos más variedad de notas.**

Se ve mas variedad de notas en el grupo de Mates. Esto se evidencia en el tamaño del box y la ampliud de los límites respecto de los box de los otros grupos.

**7. En base a todo lo anterior, ¿puedes afirmar de forma estadística que el examen de carrera ha ido mejor en algún estudio que en los demás?**

El examen de carrera ha ido mejor para el grupo de mates.