

Generación de gráficas con ggplot2

J. Eduardo Gamboa U.

2025-10-12

Contents

1	Introducción	2
2	Gráficos para variables cualitativas	3
2.1	Gráfico de barras	3
2.2	Gráfico circular	7
2.3	Gráfico de mosaico	10
2.4	Diagrama de treemap	11
3	Gráficos para variables cuantitativas	12
3.1	Histograma	12
3.2	Gráfico de densidad	13
3.3	Ojiva (Función de distribución acumulada empírica)	15
3.4	Boxplot	16
3.5	Violín	17

1 Introducción

Una representación visual permite detectar patrones, tendencias y valores atípicos que serían difíciles de identificar en tablas. En R, el paquete ggplot2 implementa la Gramática de los Gráficos (*The Grammar of Graphics*), propuesta por Wilkinson (2005), donde toda visualización se construye combinando datos, mapeos estéticos y geometrías.

Cargar los paquetes ggplot2 y dplyr

```
library(ggplot2)
library(dplyr)
```

La función principal es ggplot(), que combina datos (data), mapeos estéticos (aes()) y capas geométricas (geom_*()).

```
df |>
  ggplot(aes(x = variable_x, y = variable_y)) +
  geom_<tipo>()
```

Para la aplicación utilizaremos el archivo Programacion_Academica_2025_II_UNALM.csv, que contiene información sobre las secciones programadas para el ciclo 2025-II en la UNALM. Cabe señalar que un mismo curso puede aparecer más de una vez si cuenta con varias secciones programadas.

```
datos <- read.csv2('Programacion_Academica_2025_II_UNALM.csv')
datos|> head(5)
```

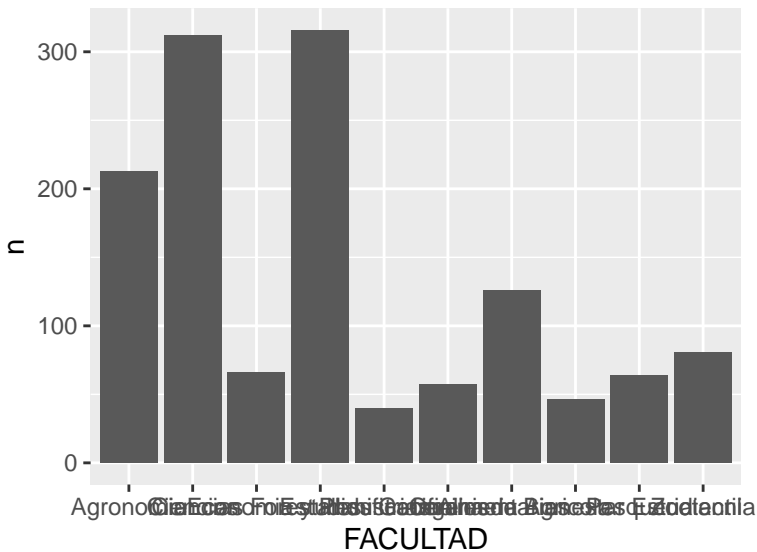
##	FACULTAD	DEPARTAMENTO	PERIODO	CURSO	SECCION
## 1	Agronomia	Entomologia	2025-II	Apicultura General	1490
## 2	Agronomia	Entomologia	2025-II	Control Biologico	1500
## 3	Agronomia	Entomologia	2025-II	Crianza y Evaluacion de Insectos	1510
## 4	Agronomia	Entomologia	2025-II	Entomologia Agricola	1520
## 5	Agronomia	Entomologia	2025-II	Entomologia Agricola	1530
##	MATRICULADOS				
## 1					30
## 2					17
## 3					12
## 4					41
## 5					30

2 Gráficos para variables cualitativas

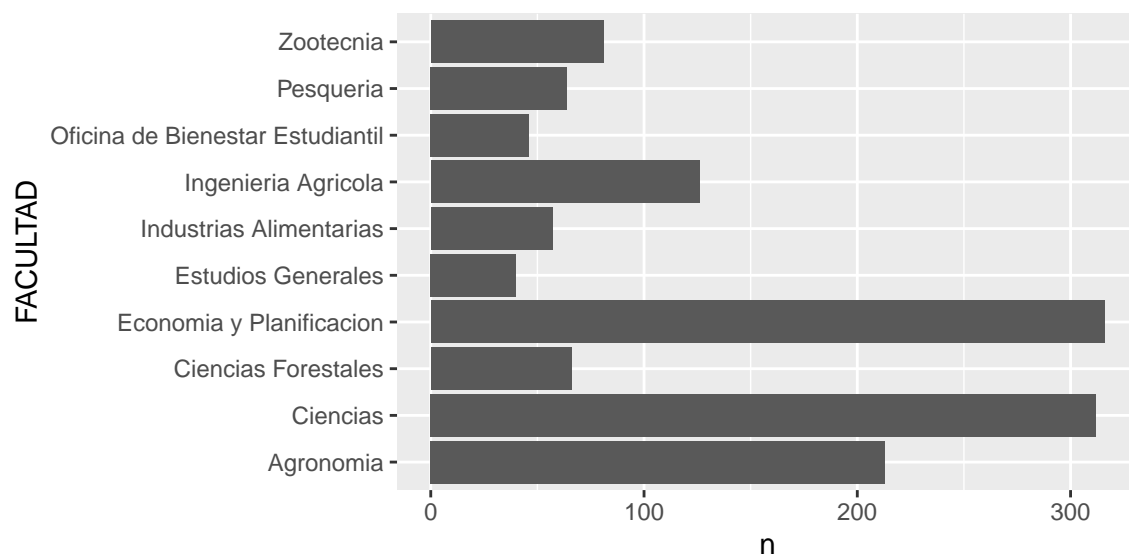
2.1 Gráfico de barras

Representa frecuencias o totales por categoría.

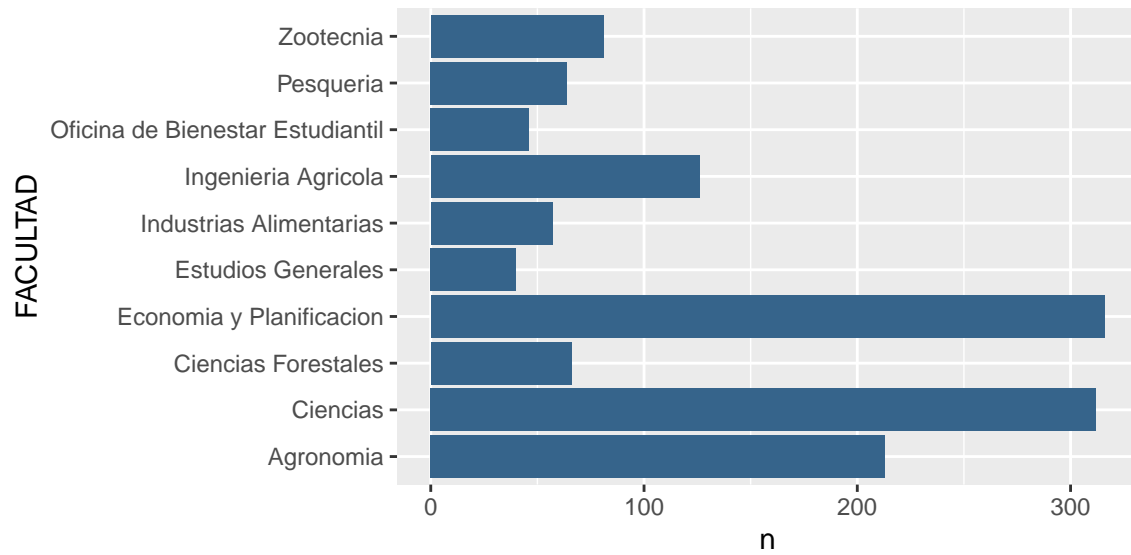
```
datos |>  
  count(FACULTAD) |>  
  ggplot(aes(x = FACULTAD, y = n)) +  
  geom_col()
```



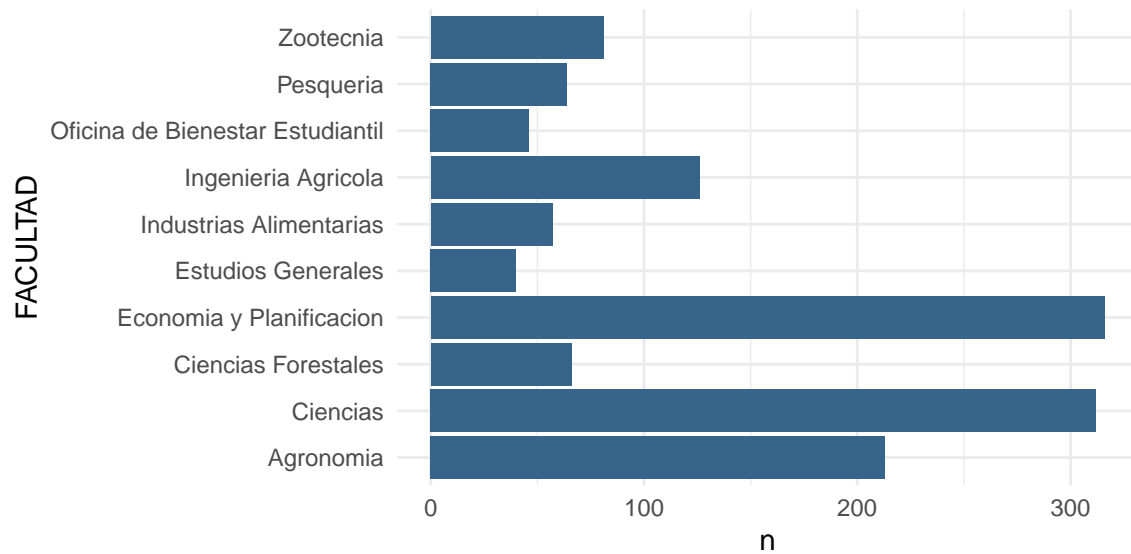
```
datos |>  
  count(FACULTAD) |>  
  ggplot(aes(x = FACULTAD, y = n)) +  
  geom_col() +  
  coord_flip()
```



```
datos |>
  count(FACULTAD) |>
  ggplot(aes(x = FACULTAD, y = n)) +
  geom_col(fill = "steelblue4") +
  coord_flip()
```



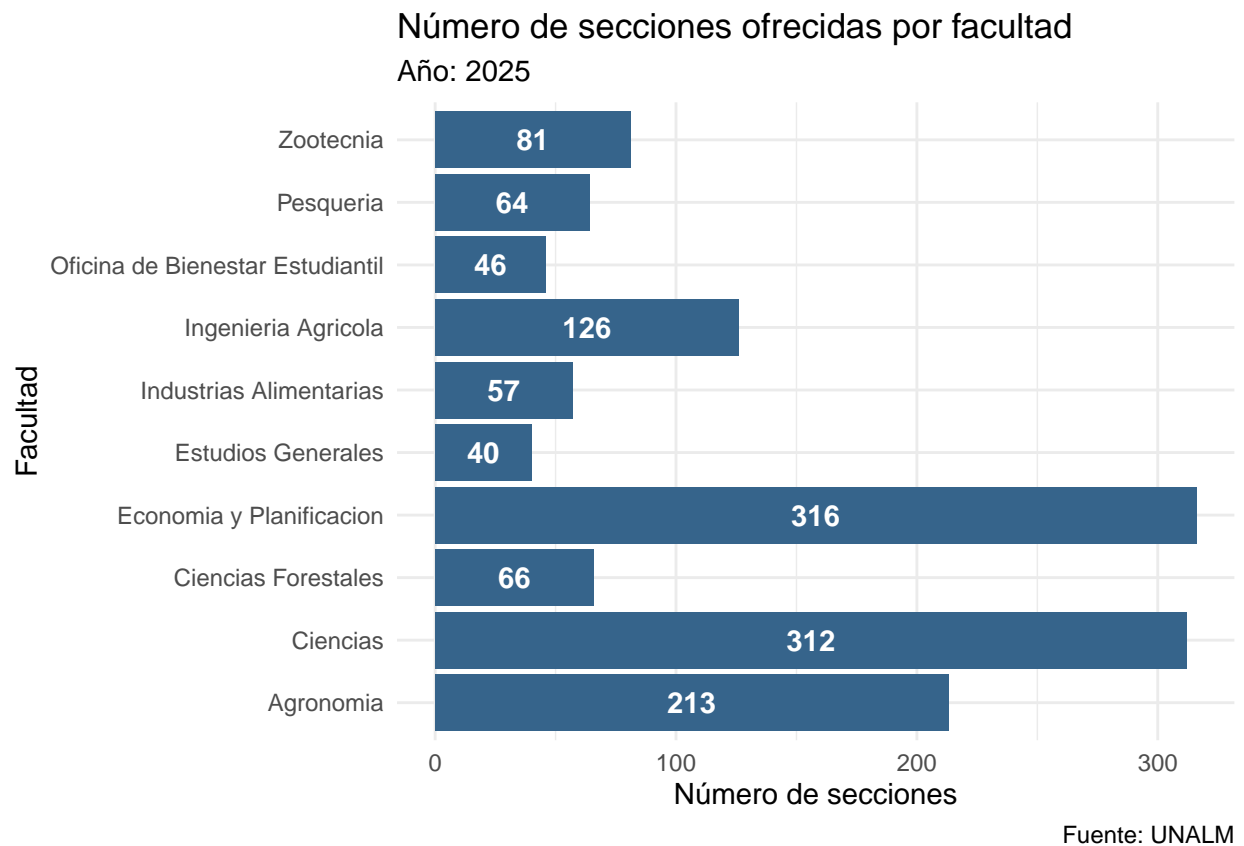
```
datos |>
  count(FACULTAD) |>
  ggplot(aes(x = FACULTAD, y = n)) +
  geom_col(fill = "steelblue4") +
  coord_flip() +
  theme_minimal()
```



```

datos |>
  count(FACULTAD) |>
  ggplot(aes(x = FACULTAD, y = n)) +
  geom_col(fill = "steelblue4") +
  geom_text(aes(label = n,
                position = position_stack(vjust = 0.5),
                color = "white",
                size = 4,
                fontface = "bold")) +
  coord_flip() +
  labs(title = "Número de secciones ofrecidas por facultad",
       subtitle = "Año: 2025",
       x = "Facultad",
       y = "Número de secciones",
       caption = "Fuente: UNALM") +
  theme_minimal()

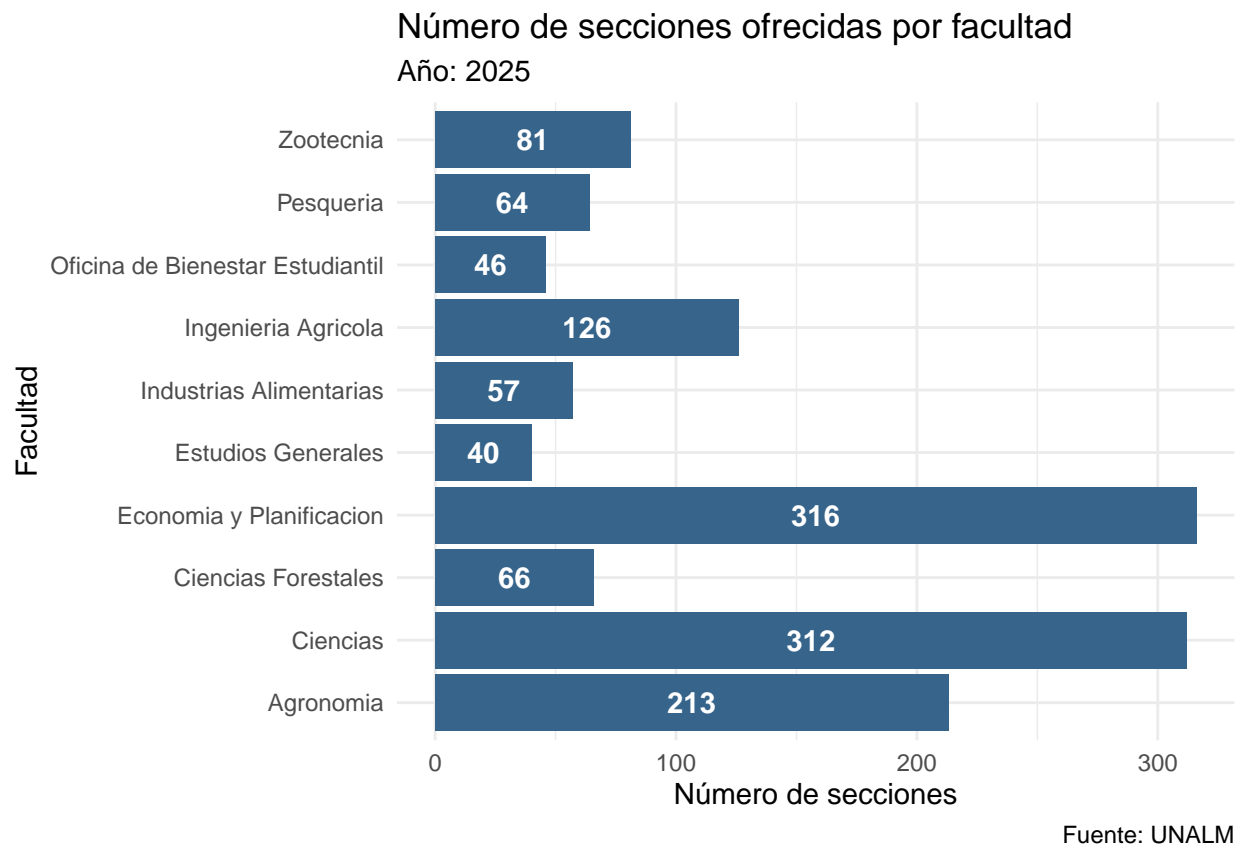
```



```

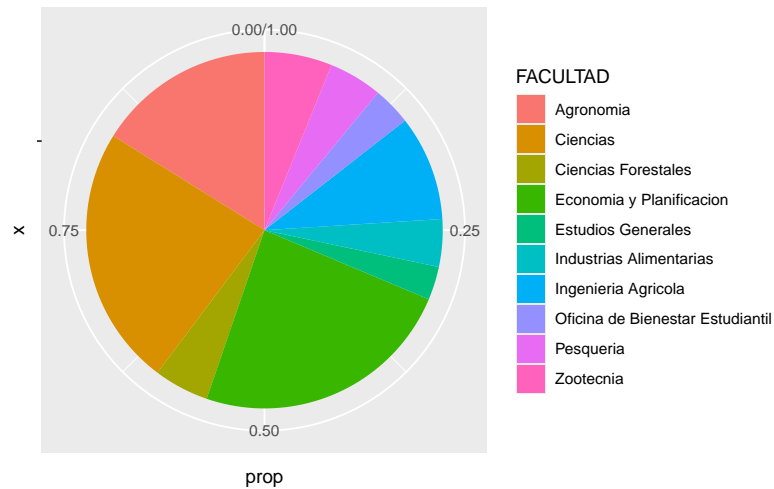
datos |>
  ggplot(aes(x = FACULTAD)) +
  geom_bar(fill = "steelblue4") +
  geom_text(aes(label = after_stat(count)),
            stat = "count",
            position = position_stack(vjust = 0.5),
            color = "white",
            size = 4,
            fontface = "bold") +
  coord_flip() +
  labs(title = "Número de secciones ofrecidas por facultad",
       subtitle = "Año: 2025",
       x = "Facultad",
       y = "Número de secciones",
       caption = "Fuente: UNALM") +
  theme_minimal()

```

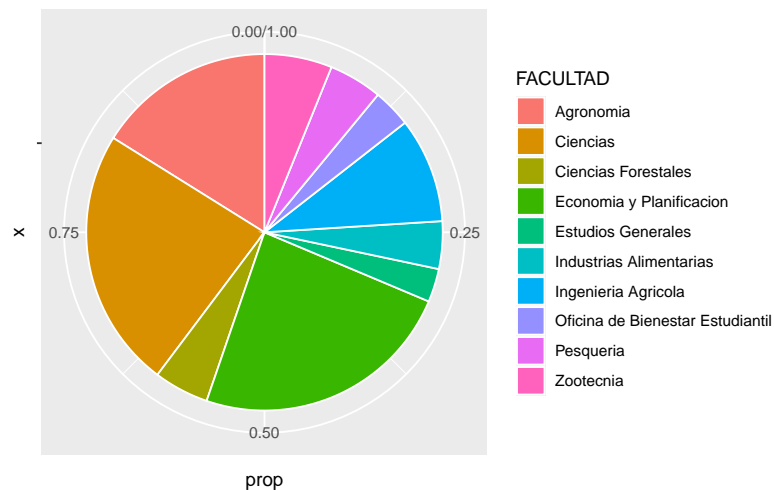


2.2 Gráfico circular

```
datos |>
  count(FACULTAD) |>
  mutate(prop = n / sum(n)) |>
  ggplot(aes(x = "", y = prop, fill = FACULTAD)) +
  geom_col() +
  coord_polar(theta = "y")
```



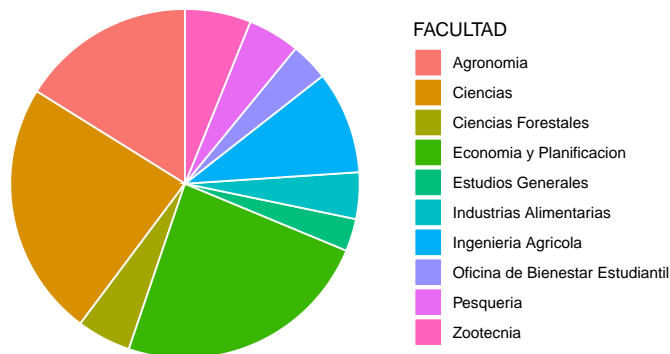
```
datos |>
  count(FACULTAD) |>
  mutate(prop = n / sum(n)) |>
  ggplot(aes(x = "", y = prop, fill = FACULTAD)) +
  geom_col(width = 1, color = "white") +
  coord_polar(theta = "y")
```



```

datos |>
  count(FACULTAD) |>
  mutate(prop = round(n / sum(n) * 100,1)) |>
  ggplot(aes(x = "", y = prop, fill = FACULTAD)) +
  geom_col(width = 1, color = "white") +
  coord_polar(theta = "y") +
  theme_void()

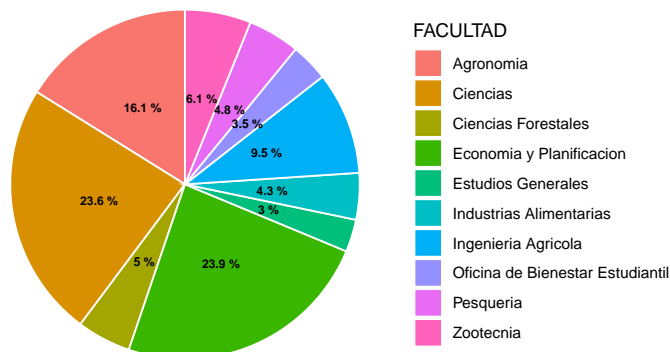
```



```

datos |>
  count(FACULTAD) |>
  mutate(prop = round(n / sum(n) * 100,1)) |>
  ggplot(aes(x = 1, y = prop, fill = FACULTAD)) +
  geom_col(width = 1, color = "white") +
  coord_polar(theta = "y") +
  geom_text(aes(label = paste(prop,"%"),
    position = position_stack(vjust = 0.5),
    color = "black",
    size = 2.5,
    fontface = "bold")) +
  theme_void()

```



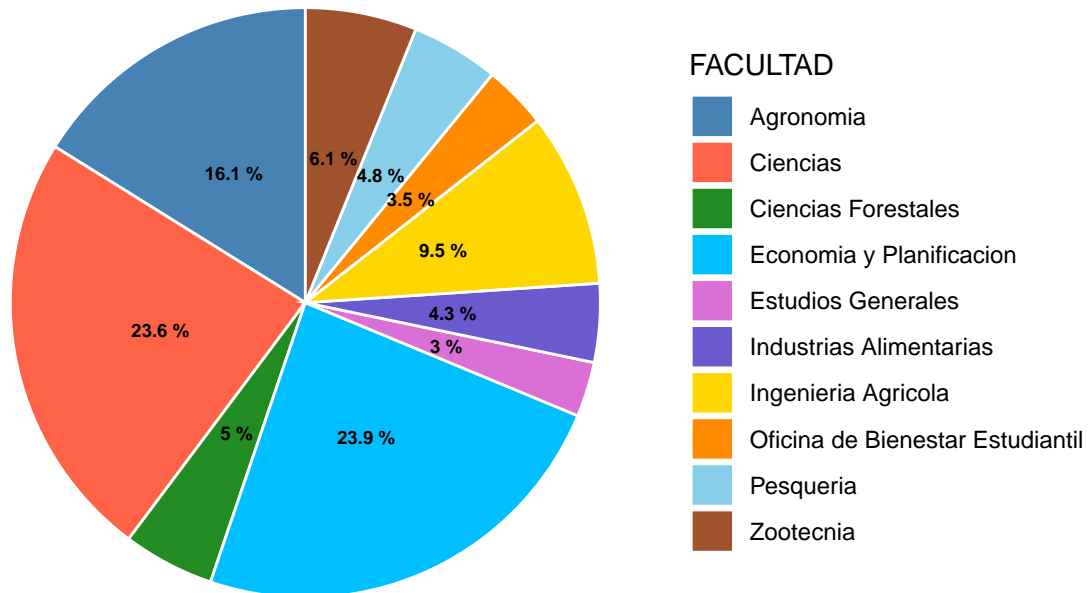

```

datos |>
  count(FACULTAD) |>
  mutate(prop = round(n / sum(n) * 100,1)) |>
  ggplot(aes(x = 1, y = prop, fill = FACULTAD)) +
  geom_col(width = 1, color = "white") +
  coord_polar(theta = "y") +
  scale_fill_manual(values=c("steelblue","tomato","forestgreen","deepskyblue","orchid",
                             "slateblue","gold","darkorange","skyblue","sienna"))+
  geom_text(aes(label = paste(prop,"%"),
                    position = position_stack(vjust = 0.5),
                    color = "black",
                    size = 2.5,
                    fontface = "bold")) +
  labs(title = "Porcentaje de secciones ofrecidas por facultad",
        subtitle = "Año: 2025",
        caption = "Fuente: UNALM") +
  theme_void()

```

Porcentaje de secciones ofrecidas por facultad

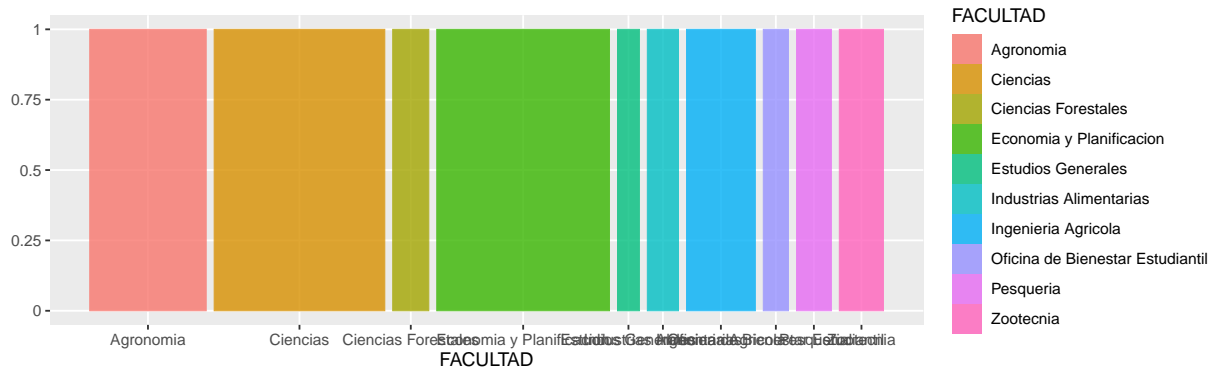
Año: 2025



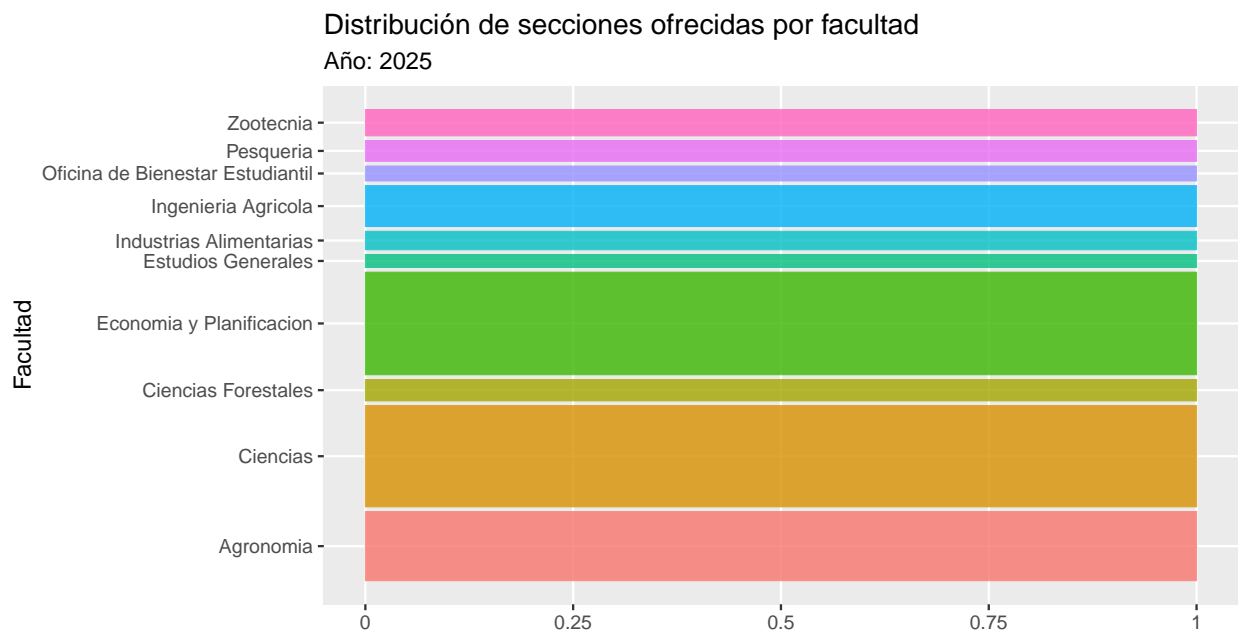
Fuente: UNALM

2.3 Gráfico de mosaico

```
library(ggmosaic)
datos |>
  ggplot() +
  geom_mosaic(aes(x = product(FACULTAD), fill = FACULTAD))
```



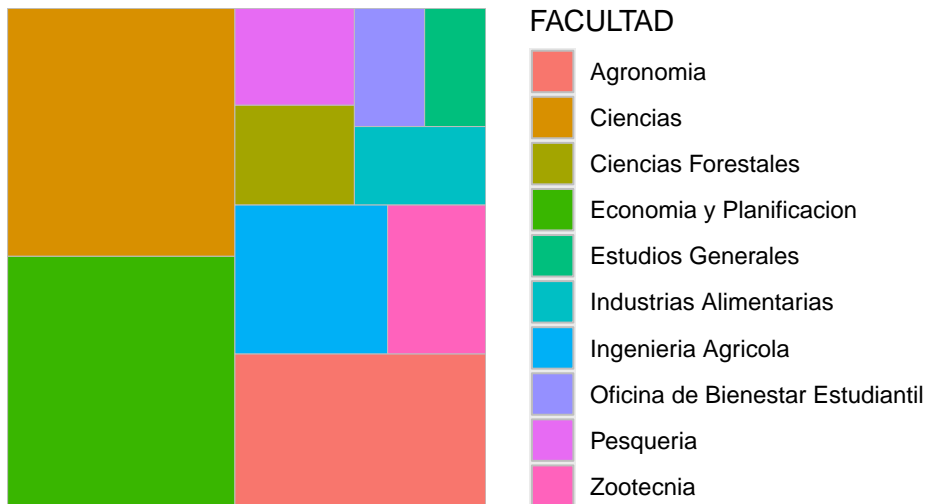
```
datos |>
  ggplot() +
  geom_mosaic(aes(x = product(FACULTAD), fill = FACULTAD)) +
  labs(title = "Distribución de secciones ofrecidas por facultad",
       subtitle = "Año: 2025",
       x = "Facultad",
       y = "",
       caption = "Fuente: UNALM") +
  coord_flip() +
  theme(legend.position = "none")
```



Fuente: UNALM

2.4 Diagrama de treemap

```
library(treemapify)
datos |> count(FACULTAD) |>
  ggplot(aes(area = n, fill = FACULTAD, label = FACULTAD)) +
  geom_treemap()
```



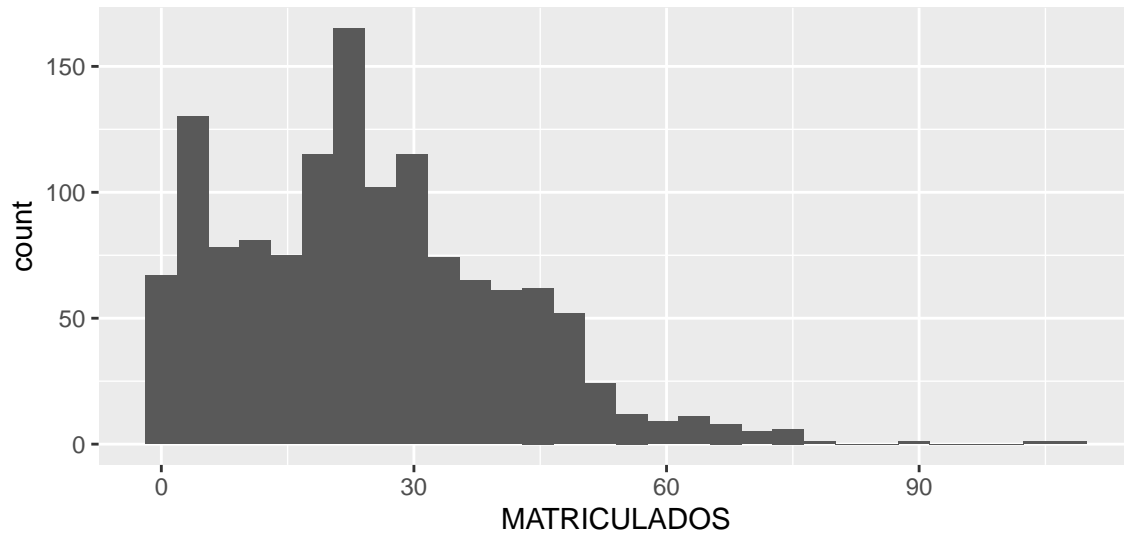
```
datos |> count(FACULTAD) |>
  ggplot(aes(area = n, fill = FACULTAD, label = FACULTAD)) +
  geom_treemap() +
  geom_treemap_text(colour = "white",
                    place = "centre",
                    grow = FALSE,
                    reflow = TRUE) +
  scale_fill_manual(values=c("steelblue","tomato","forestgreen","deepskyblue","orchid",
                             "slateblue","gold","darkorange","skyblue","sienna")) +
  theme(legend.position = "none")
```



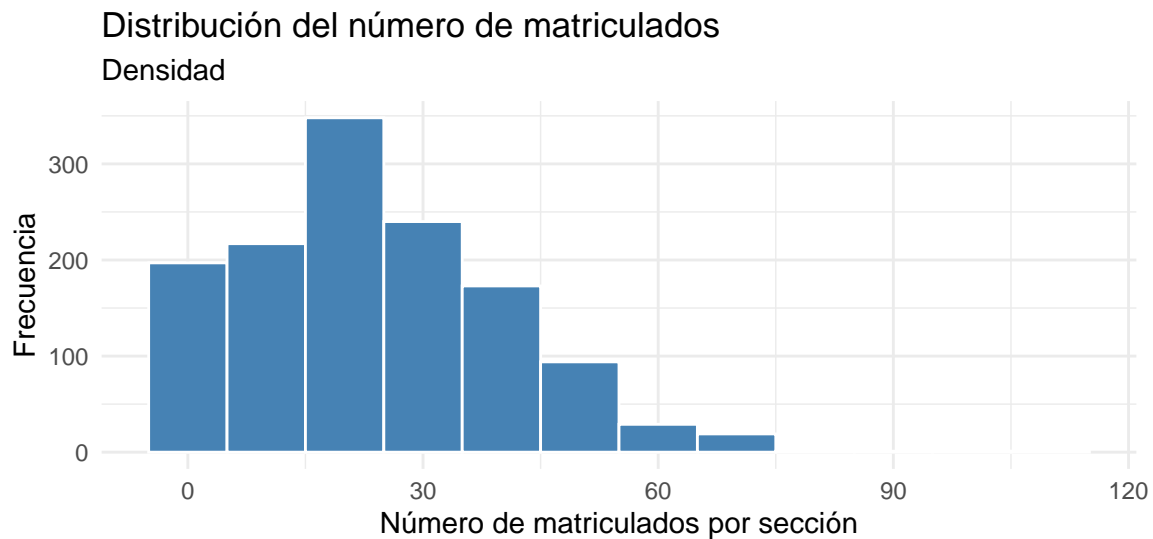
3 Gráficos para variables cuantitativas

3.1 Histograma

```
datos |>  
  ggplot(aes(x = MATRICULADOS)) +  
  geom_histogram()
```

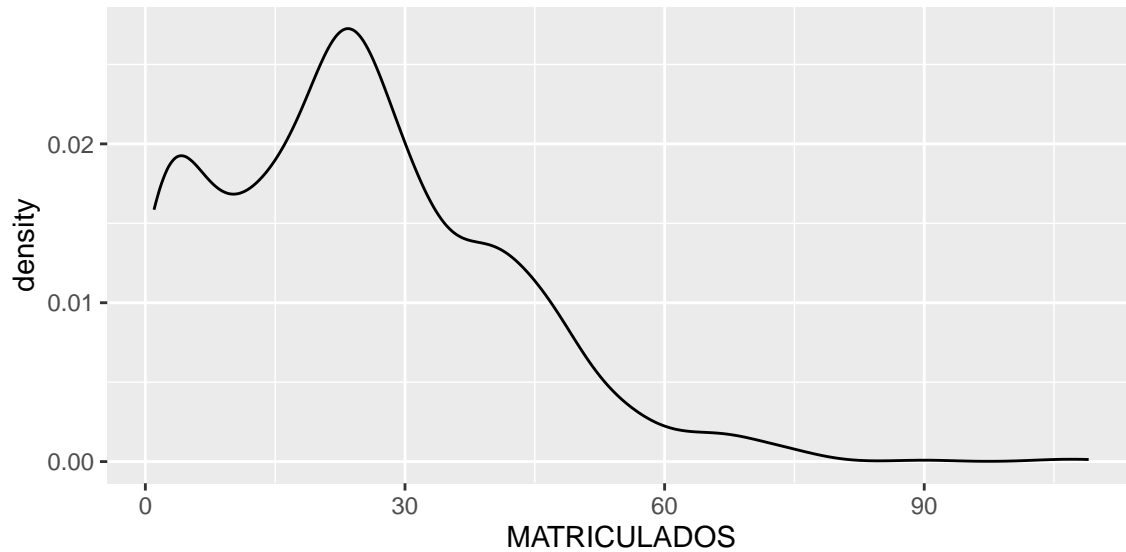


```
datos |>  
  ggplot(aes(x = MATRICULADOS)) +  
  geom_histogram(binwidth = 10, fill = "steelblue", color = "white") +  
  labs(title = "Distribución del número de matriculados",  
       subtitle = "Densidad",  
       x = "Número de matriculados por sección",  
       y = "Frecuencia") +  
  theme_minimal()
```

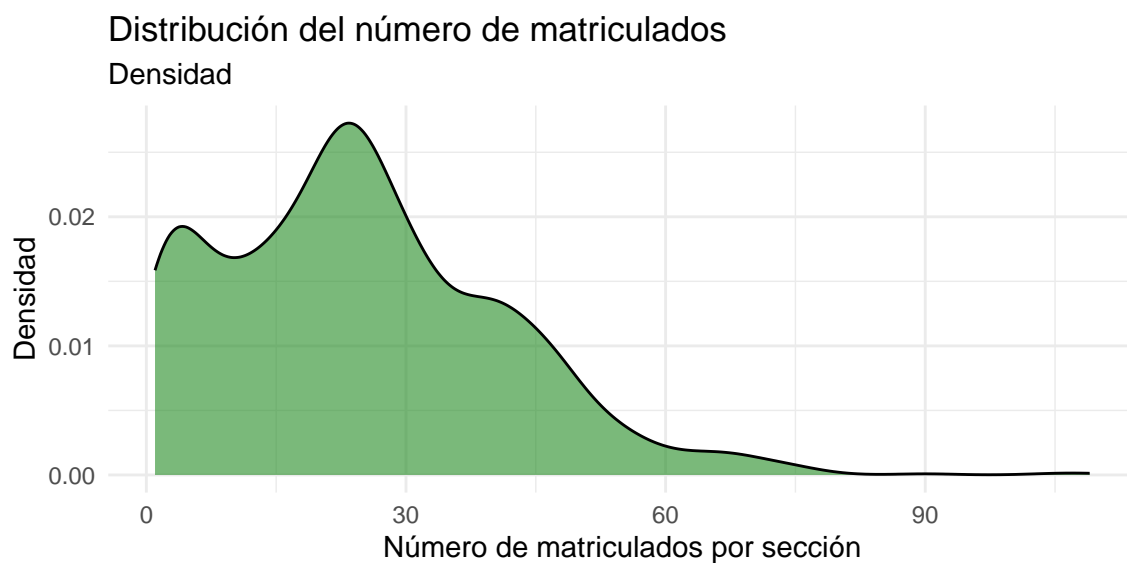


3.2 Gráfico de densidad

```
datos |>  
  ggplot(aes(x = MATRICULADOS)) +  
  geom_density()
```

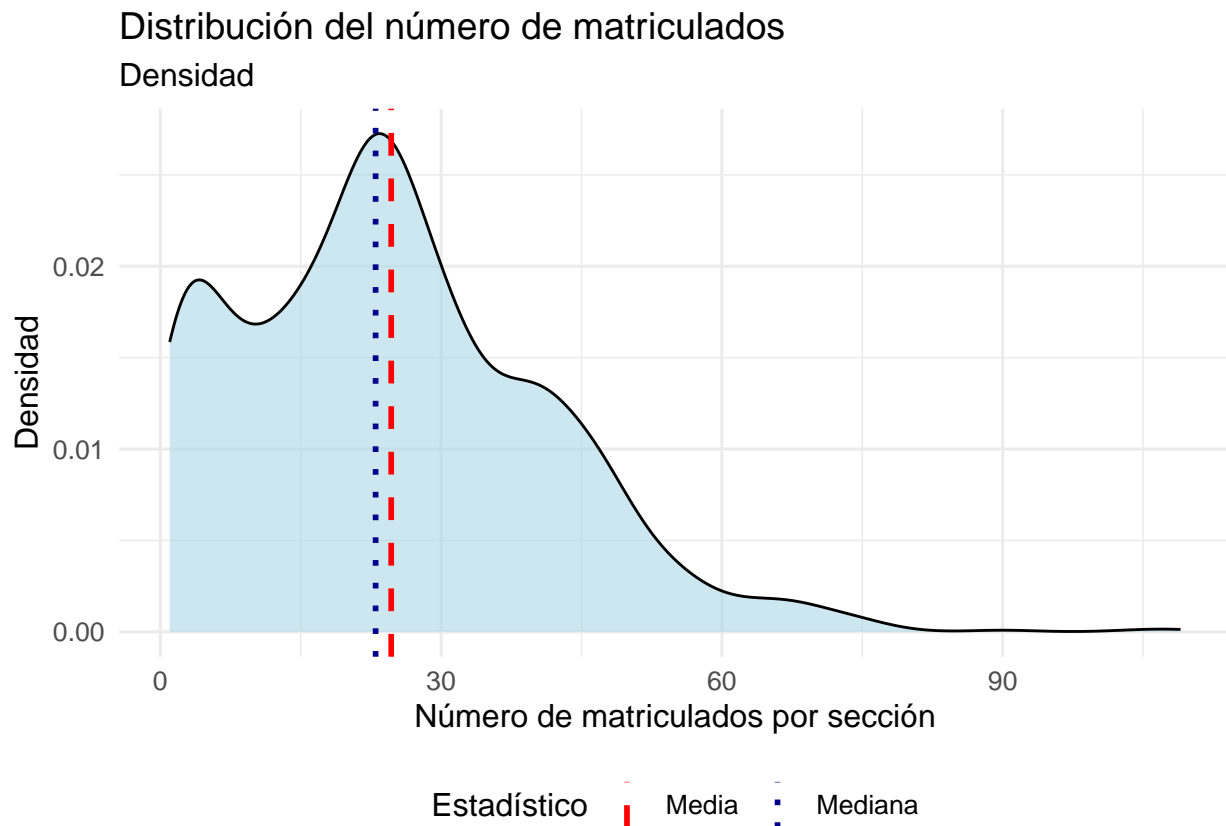


```
datos |>  
  ggplot(aes(x = MATRICULADOS)) +  
  geom_density(fill = "forestgreen", alpha = 0.6, color = "black") +  
  labs(title = "Distribución del número de matriculados",  
       subtitle = "Densidad",  
       x = "Número de matriculados por sección",  
       y = "Densidad") +  
  theme_minimal()
```



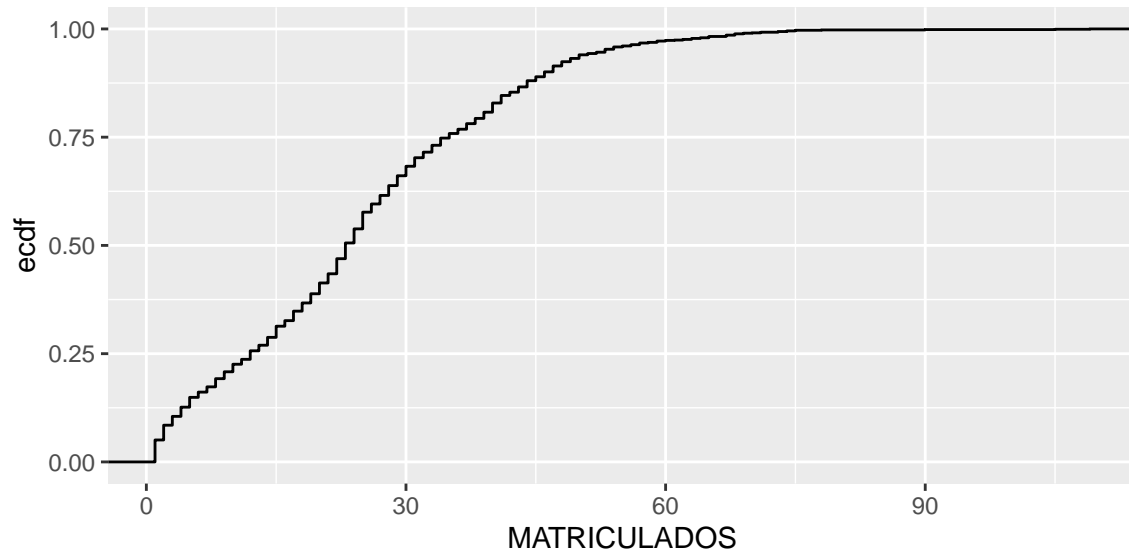
```
library(ggplot2)

ggplot(datos, aes(x = MATRICULADOS)) +
  geom_density(fill = "lightblue", alpha = 0.6) +
  geom_vline(data = data.frame(valor = mean(datos$MATRICULADOS), tipo = "Media"),
    aes(xintercept = valor, color = tipo),
    linetype = "dashed", linewidth = 1) +
  geom_vline(data = data.frame(valor = median(datos$MATRICULADOS), tipo = "Mediana"),
    aes(xintercept = valor, color = tipo),
    linetype = "dotted", linewidth = 1) +
  scale_color_manual(name = "Estadístico",
    values = c("Media" = "red", "Mediana" = "darkblue")) +
  labs(title = "Distribución del número de matriculados",
    subtitle = "Densidad",
    x = "Número de matriculados por sección",
    y = "Densidad") +
  theme_minimal(base_size = 12) +
  theme(legend.position = "bottom")
```

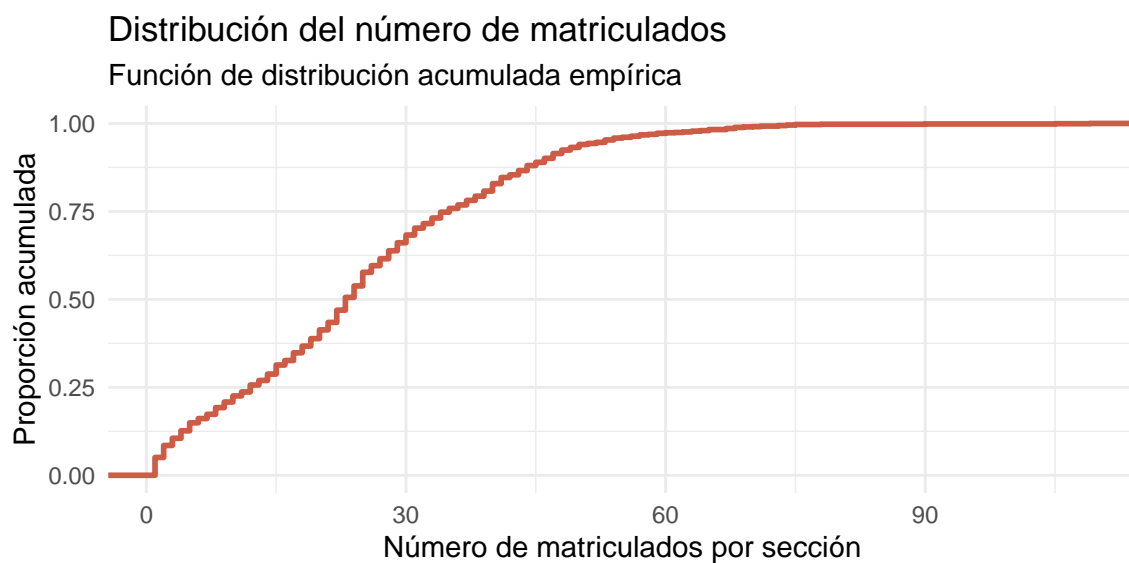


3.3 Ojiva (Función de distribución acumulada empírica)

```
datos |>  
  ggplot(aes(x = MATRICULADOS)) +  
  stat_ecdf()
```

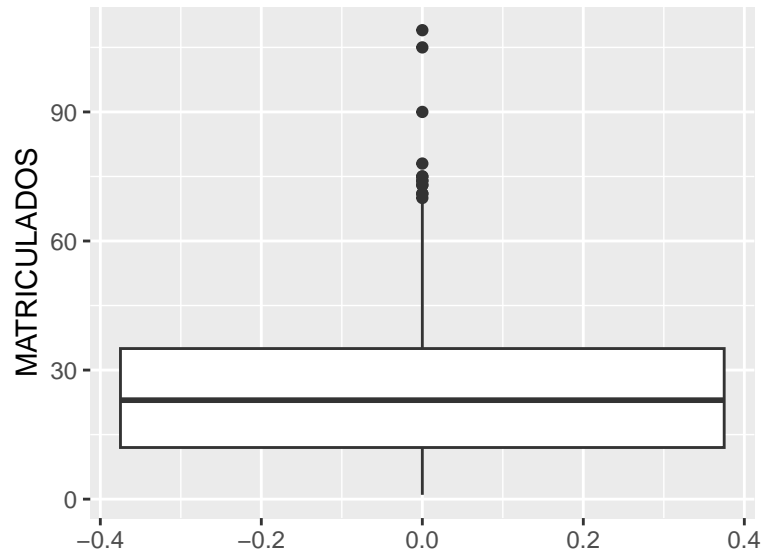


```
datos |>  
  ggplot(aes(x = MATRICULADOS)) +  
  stat_ecdf(geom = "step", color = "coral3", linewidth = 1) +  
  labs(title = "Distribución del número de matriculados",  
       subtitle = "Función de distribución acumulada empírica",  
       x = "Número de matriculados por sección",  
       y = "Proporción acumulada") +  
  theme_minimal()
```

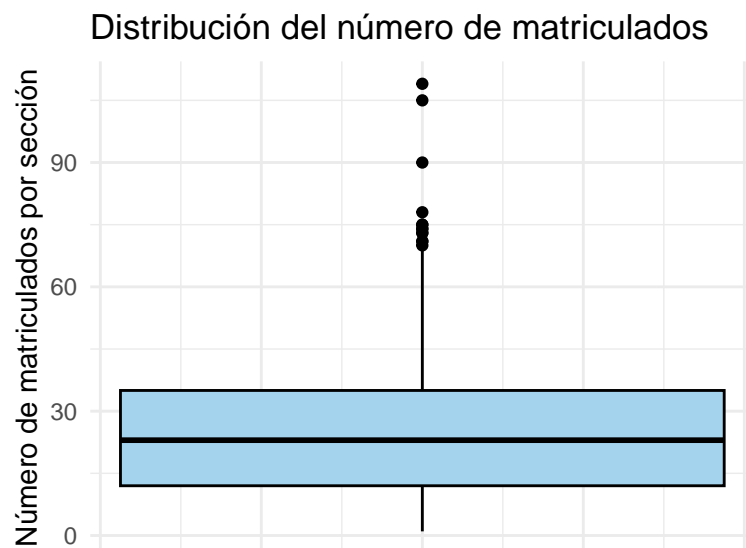


3.4 Boxplot

```
datos |>  
  ggplot(aes(y = MATRICULADOS)) +  
  geom_boxplot()
```

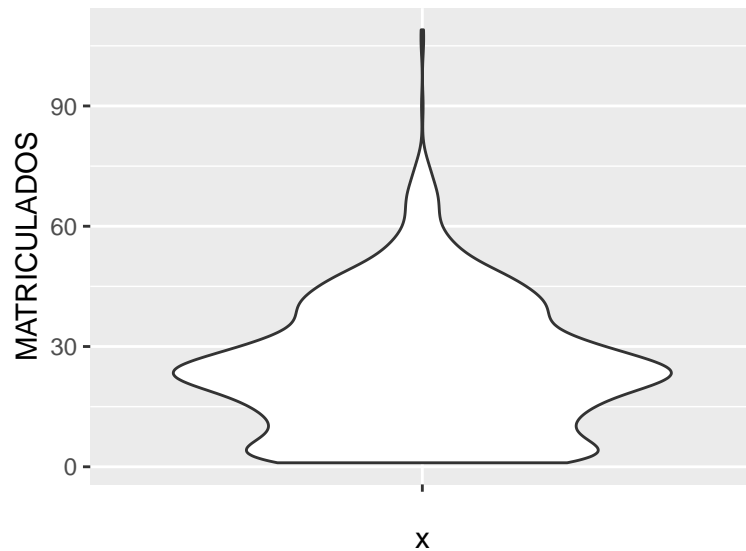


```
datos |>  
  ggplot(aes(y = MATRICULADOS)) +  
  geom_boxplot(fill = "lightskyblue2", color = "black") +  
  labs(title = "Distribución del número de matriculados",  
       y      = "Número de matriculados por sección") +  
  theme_minimal() +  
  theme(axis.text.x = element_blank(),  
        axis.ticks.x = element_blank())
```



3.5 Violín

```
datos |> ggplot(aes(x = "", y = MATRICULADOS)) + geom_violin()
```



```
datos |>
  ggplot(aes(x = "", y = MATRICULADOS)) +
  geom_violin(trim = TRUE, alpha = 0.7, fill = "gold", color = "black") +
  geom_boxplot(width = 0.1, alpha = 0.7, fill = "red", color = "black",
    outlier.color = "black", outlier.shape = 15) +
  coord_flip() +
  labs(title = "Distribución del número de matriculados",
    x = "",
    y = "Número de matriculados por sección") +
  theme_minimal()
```

Distribución del número de matriculados

