

# Tema VI - Representación Gráfica con ggstatsplot

Desafío 6:

Dr. Maicel Monzón Pérez

2025-03-24

## Índice

<b>Tareas 1. Gráfico de pastel</b>	<b>1</b>
<b>Tareas 2. Gráfico de barras para comparar dos variable cualitativas</b>	<b>2</b>
<b>Tareas 3. Histograma para observar la distribución del consumo de hierro</b>	<b>3</b>
<b>Tareas 5. Graficos de dispersión por variables</b>	<b>4</b>
<b>Tareas 6. Usar ggscatterstats() para mostrar correlación.</b>	<b>5</b>
<b>Tareas 7. Comparación entre grupos (Países)**</b>	<b>6</b>
<b>Conclusión</b>	<b>7</b>

```
library(readr)
datos <- read_csv("static/csv/datos_fortificacion.csv",na = c("NA", "-", ""))
# View(datos_fortificacion)
```

## Tareas 1. Gráfico de pastel

- Crear un gráfico de barras con `ggpiestats()` para describir los principales grupos de edades de la muestra (variable `grupo_objetivo`)

Solución:

```
library(ggstatsplot)
```

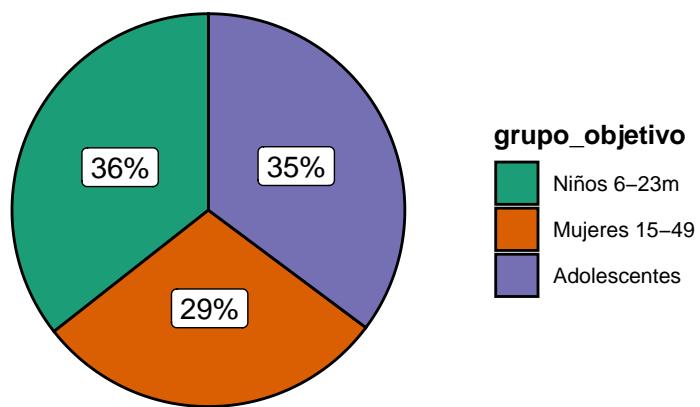
You can cite this package as:

Patil, I. (2021). Visualizations with statistical details: The 'ggstatsplot' approach. Journal of Open Source Software, 6(61), 3167, doi:10.21105/joss.03167

```
ggpiestats(  
  data = datos,  
  x = grupo_objetivo,  
  title = "Distribución de grupos objetivo en el estudio",  
  caption = "Fuente: datos_fortificacion.csv",  
  bf.message = FALSE,  
  proportion.test = FALSE # Omitir prueba de proporciones para simplificar  
)
```

### Distribución de grupos objetivo en el estudio

$\chi^2_{\text{gof}}(2) = 3.86, p = 0.15, \widehat{C}_{\text{Pearson}} = 0.09, \text{CI}_{95\%} [0.00, 0.16], n_{\text{obs}} = 500$



Fuente: datos\_fortificacion.csv

### Tareas 2. Gráfico de barras para comparar dos variable cualitativas

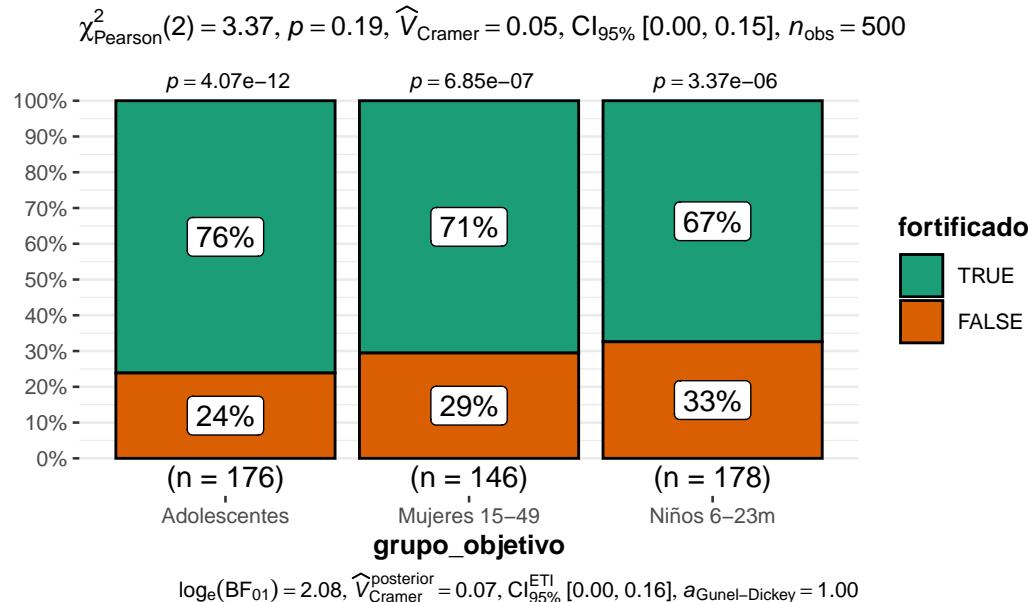
```
ggbarsstats(  
  data = datos,
```

```

x = fortificado, # TRUE/FALSE
y = grupo_objetivo,
title = "Alimentos no fortificados por grupo objetivo"
)

```

### Alimentos no fortificados por grupo objetivo



### Tareas 3. Histograma para observar la distribución del consumo de hierro

```

library(ggstatsplot)
library(ggplot2)

# Histograma con análisis de normalidad
gghistogram(
  data = datos,
  x = hierro_mg,           # Variable numérica a analizar
  title = "Distribución de hierro (mg) en alimentos",
  xlab = "Hierro (mcg)",    # Etiqueta del eje X
  ylab = "Frecuencia",      # Etiqueta del eje Y
  binwidth = 1,              # Ancho de los intervalos (ajustable)
)

```

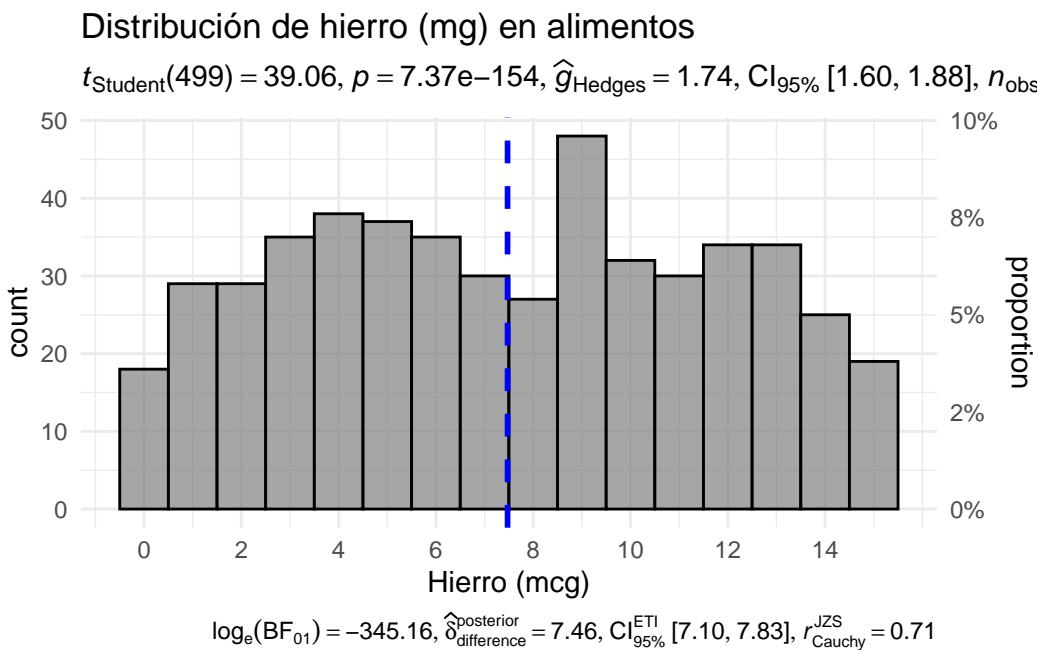
```

fill = "#4DBBD5",           # Color de las barras
test.normal = TRUE,          # Incluir prueba de normalidad (Shapiro-Wilk)
centrality.para = "mean",    # Línea para la media
messages = FALSE             # Omitir mensajes técnicos
) +
theme_minimal() +
scale_x_continuous(breaks = seq(0, 15, by = 2)) # Personalizar ejes

```

Scale for x is already present.

Adding another scale for x, which will replace the existing scale.



## Tareas 5. Graficos de dispersión por variables

- Visualice las diferencias de consumo de hierro por alimentos

```

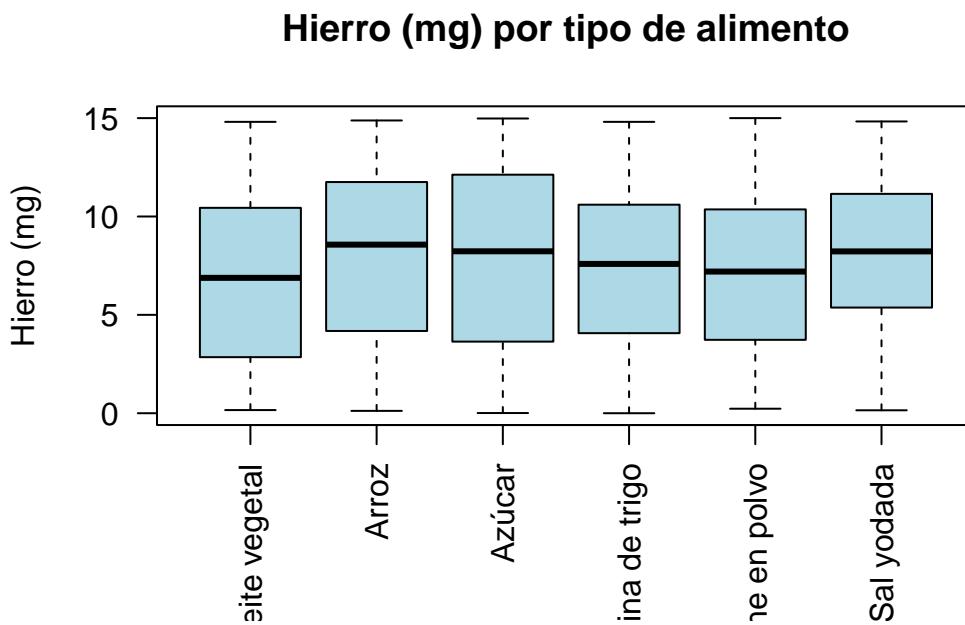
boxplot(
  hierro_mg ~ alimento,
  data = datos,
  main = "Hierro (mg) por tipo de alimento",
  xlab = ""
)

```

```

    ylab = "Hierro (mg)",
    col = "lightblue",
    las = 2 # Rotar etiquetas eje X
)

```



### Tareas 6. Usar ggscatterstats() para mostrar correlación.

Solución:

```

grafico_dispersion <- ggscatterstats(
  data = datos,
  x = "ingreso_mensual",
  y = "hierro_mg",
  title = "Relación ingreso vs. hierro en alimentos",
  xlab = "Ingreso mensual (USD)",
  ylab = "Hierro (mg)"
)

```

```

Registered S3 method overwritten by 'ggside':
  method from
  +.gg   ggplot2

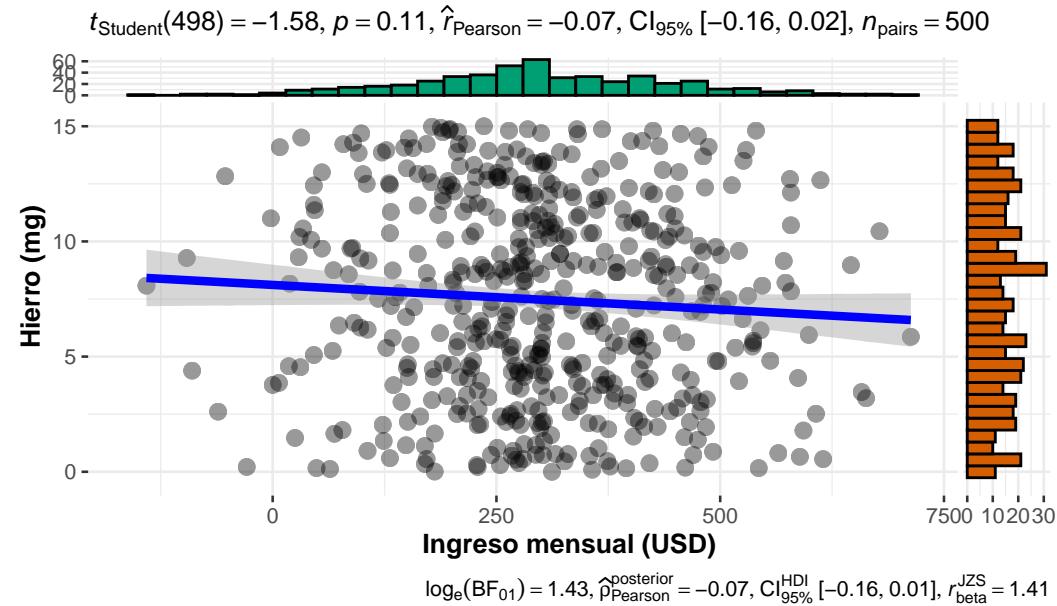
```

## grafico\_dispersión

```
`stat_xsidebin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
`stat_ysidebin()` using `bins = 30`. Pick better value with `binwidth`.
```

### Relación ingreso vs. hierro en alimentos



---

## Tareas 7. Comparación entre grupos (Países)\*\*

### Problema:

Comparar zinc\_mg entre Nepal y Nigeria.

### Instrucciones:

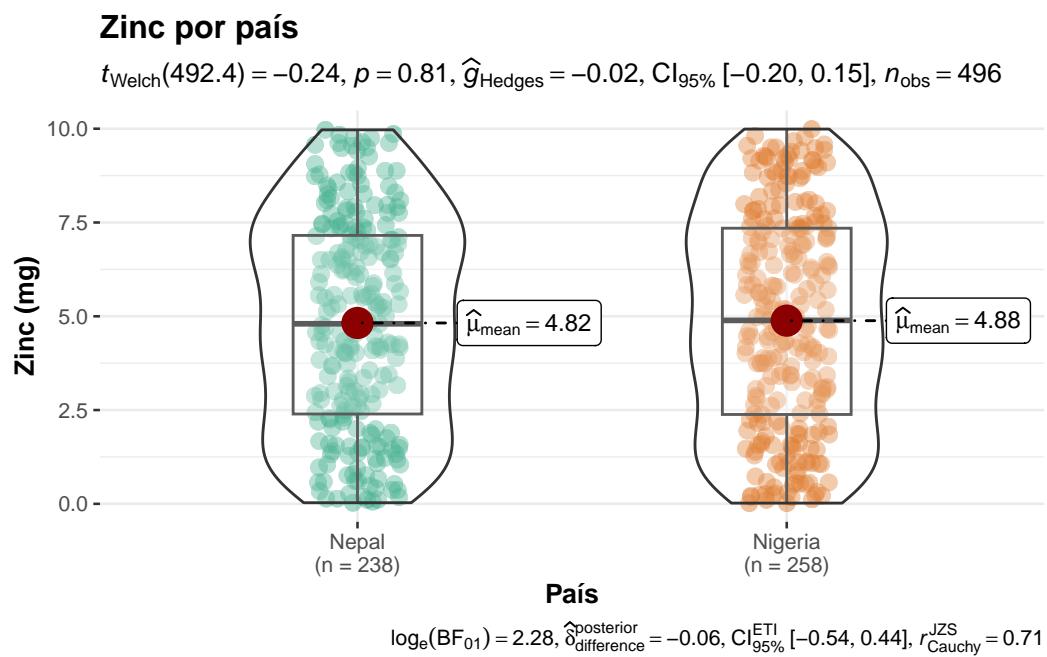
1. Filtrar datos por pais.
2. Usar ggbetweenstats() para comparar grupos.

### Solución:

```

grafico_grupos <- ggbetweenstats(
  data = datos,
  x = "pais",
  y = "zinc_mg",
  title = "Zinc por país",
  xlab = "País",
  ylab = "Zinc (mg)"
)
grafico_grupos

```



## Conclusión

- Crear gráficos comparativos y de dispersión con `ggstatsplot`.
- Corregir errores en presentación y escalas.
- Usar `group_` para análisis estratificados.