

Tema VI: Representación gráfica

Maicel Monzón

Introducción

- **Sesión anterior:** Trabajamos con tablas para organizar y sintetizar información estructurada
- **Ahora:** Transformamos esos valores en **representaciones visuales**

Gráficos > tablas para mostrar tendencias en series temporales

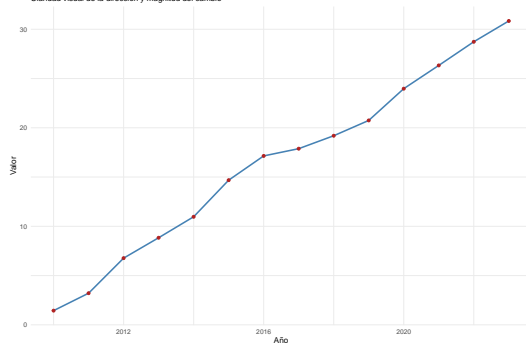
Table 1: Valores anuales (2010-2023)

Año	Valor
2010	1.439524
2011	3.209347
2012	6.768055

Gráfico de tendencia

Tendencia Temporal (2010-2023)

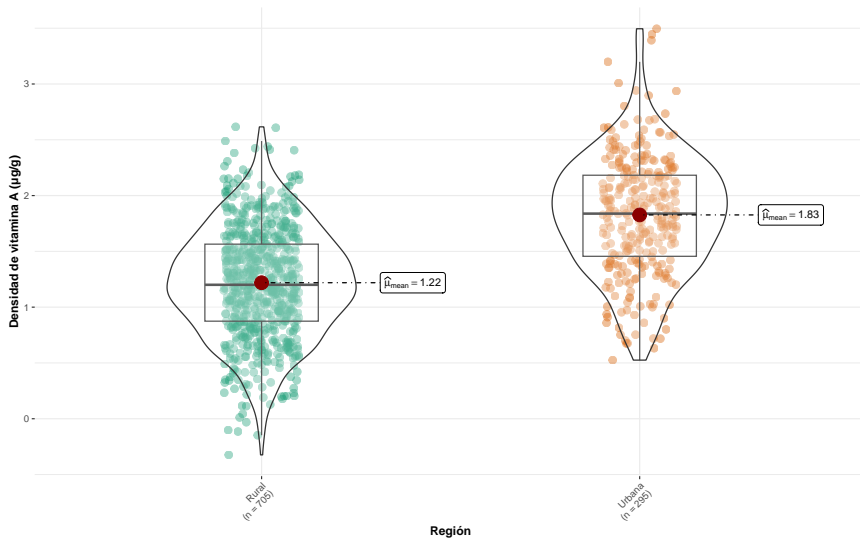
Claridad visual de la dirección y magnitud del cambio



Gráficos > Comparaciones directas entre variables o categorías

Comparación de densidad de vitamina A entre regiones y escenarios de fortificación

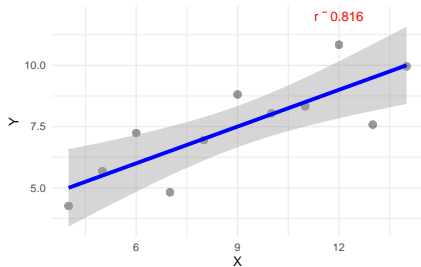
$t_{\text{Welch}}(517.32) = -16.85$, $p = 4.36\text{e-}51$, $\hat{g}_{\text{Hedges}} = -1.18$, $\text{CI}_{95\%} [-1.34, -1.03]$, $n_{\text{obs}} = 1,000$



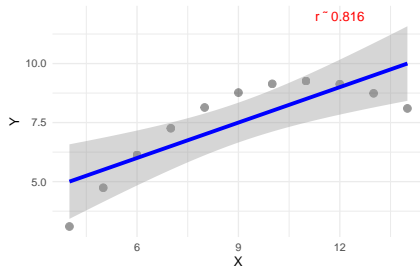
$\log_{10}(\text{BF}_{01}) = -127.78$, $\hat{\delta}_{\text{difference}}^{\text{posterior}} = -0.60$, $\text{CI}_{\text{ETI}}^{\text{ETI}} [-0.67, -0.54]$, $r_{\text{Cohen}}^{\text{JZS}}$ = 0.71

Gráficos > Relaciones y patrones ocultos

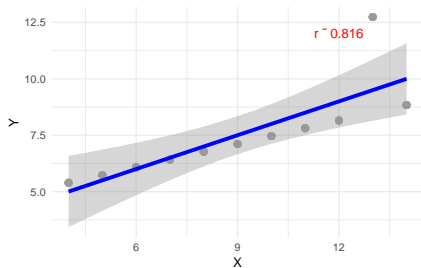
Conjunto 1



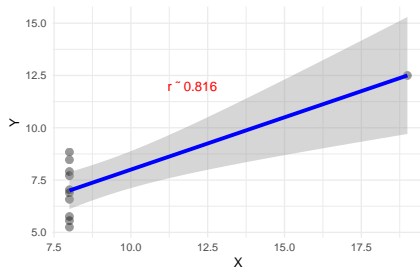
Conjunto 2



Conjunto 3



Conjunto 4



Tablas estadísticas > gráficos

“Aunque las **tablas** estadísticas son el **principal recurso** para presentar datos detallados en publicaciones científicas, los **gráficos** cumplen un papel esencial al revelar patrones, tendencias y relaciones que **complementan el análisis**. Su uso estratégico, como en el caso del Cuarteto de Anscombe, permite identificar información oculta que las tablas no muestran”

Sumario

- Los gráficos estadísticos y sus partes.
- Tipos de gráficos estadísticos en función del número y tipo de variables.
- Errores más comunes en la confección de un gráfico.
- Biblioteca **ggstatsplot** y sus principales funciones.
- Casos de uso de visualización de datos de fortificación de alimentos a gran escala.

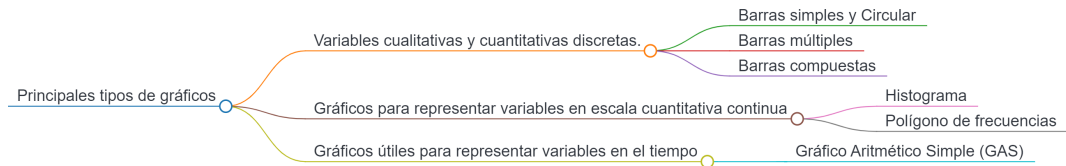
Gráficos estadísticos: Representación Visual de Datos

- **Definición:** Un gráfico es una herramienta visual que representa datos numéricos en un plano.
 - **Objetivo :** Facilitar la comprensión rápida y evidenciar patrones o tendencias.
 - **Ventaja :** Muestra de forma más clara y objetiva la evolución de fenómenos en comparación con tablas.

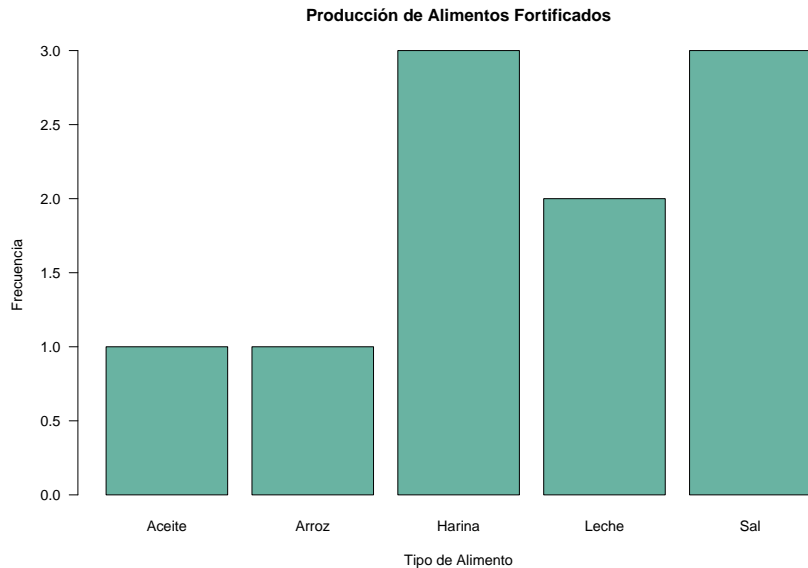
Partes del Gráfico

1. Presentación (Identificación y Título)
2. Gráfico propiamente dicho
3. Fuente
4. Notas explicativas
5. Leyenda (Identificación de sus componentes)

Principales tipos de gráficos

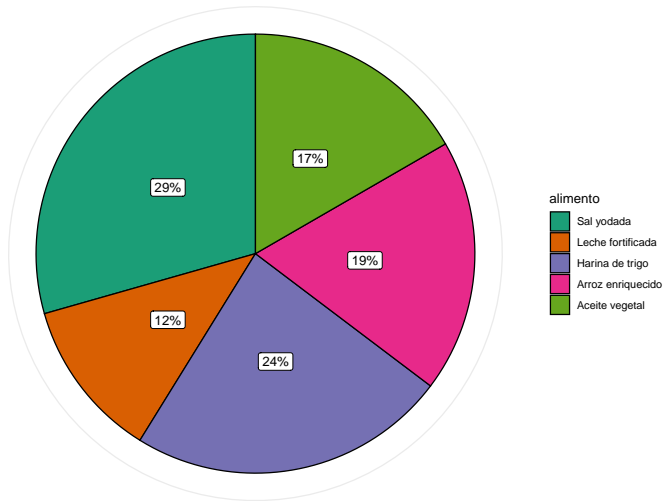


Barras simples: una sola variable

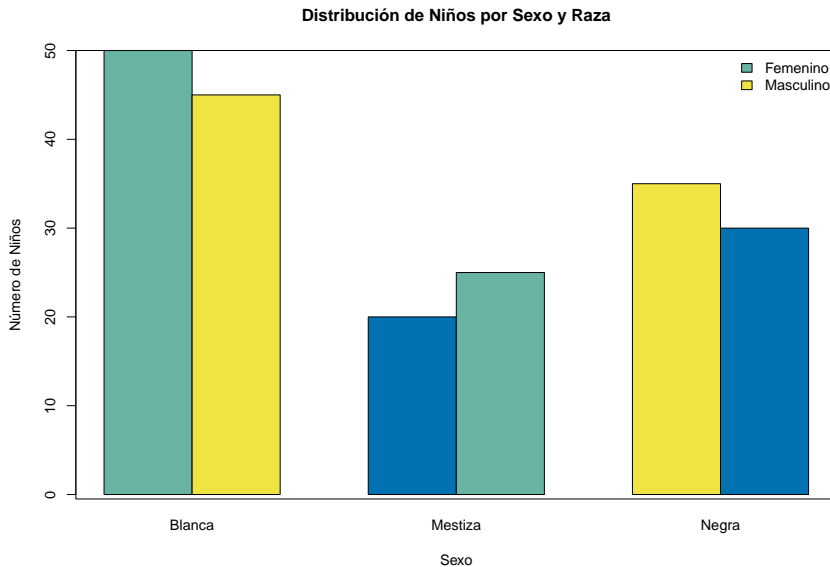


Pastel: una sola variable, frecuencias relativas

Distribución de la Producción de Alimentos Fortificados (2023)



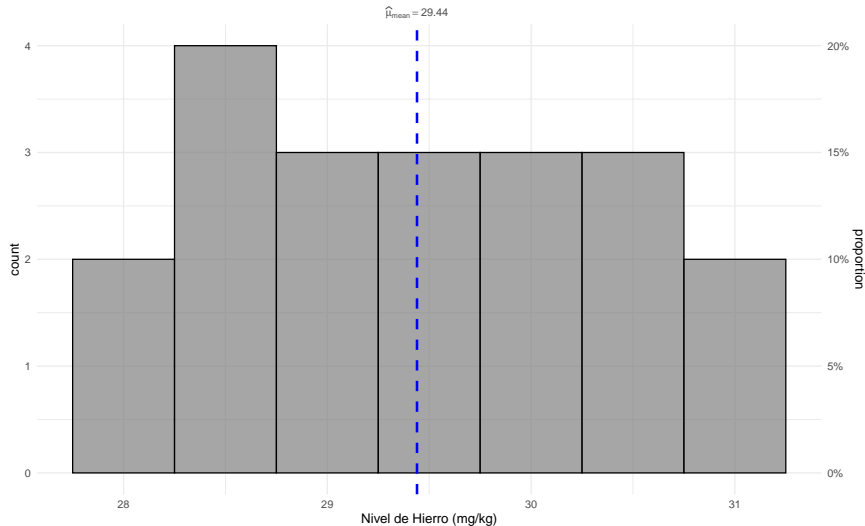
Barras múltiples: dos o más variables cualitativas



Histograma: variable cuantitativa

Distribución de Hierro en Harina Fortificada

$t_{\text{Student}}(19) = 6.67$, $p = 2.21\text{e-}06$, $\hat{g}_{\text{Hedges}} = 1.43$, $\text{CI}_{95\%} [0.81, 2.04]$, $n_{\text{obs}} = 20$

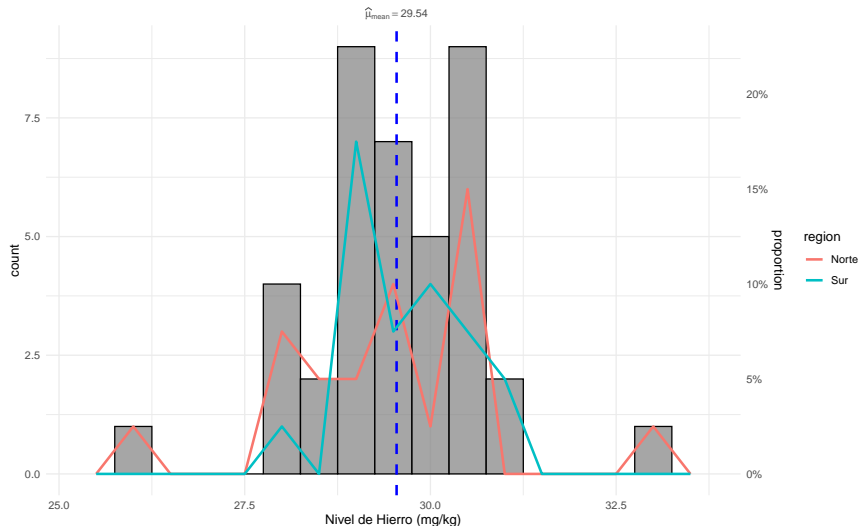


$\log_0(\text{BF}_{01}) = -9.07$, $\hat{\delta}_{\text{posterior}} = 1.44$, $\text{CI}_{95\%}^{\text{ETI}} [0.97, 1.89]$, $r_{\text{Cauchy}}^{\text{JZS}} = 0.71$

Polígono de frecuencias: variables cuantitativas

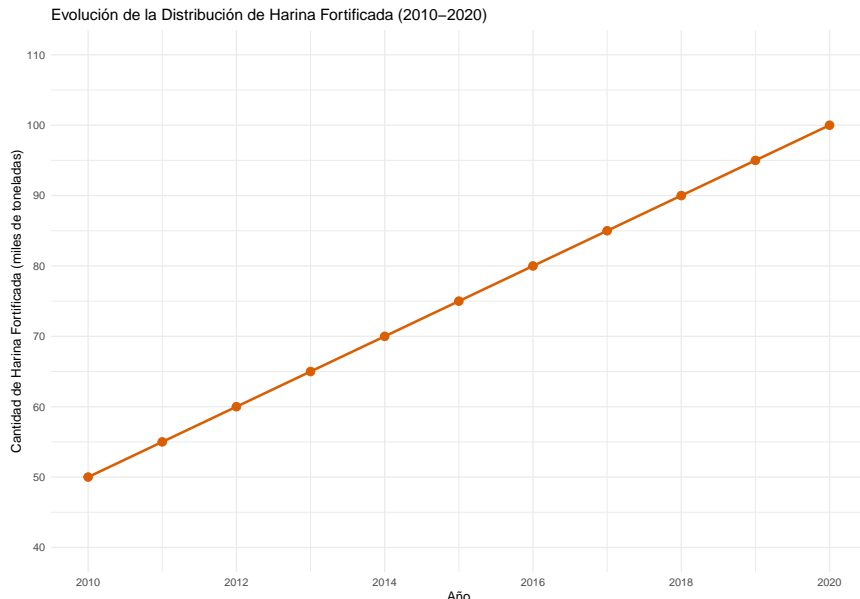
Distribución de Hierro en Harina Fortificada por Región

$t_{\text{Student}}(39) = 165.54, p = 3.80\text{e-}57, \hat{\theta}_{\text{Hedges}} = 25.67, \text{CI}_{95\%} [19.80, 31.21], n_{\text{obs}} = 40$



$\log_6(\text{BF}_{01}) = -122.21, \hat{\theta}_{\text{posterior}} = 29.55, \text{CI}_{95\%}^{\text{ETI}} [29.18, 29.91], t_{\text{Cauchy}}^{\text{JZS}} = 0.71$

Gráfico Arimético Simple: variables en el tiempo



Errores más comunes

Errores Comunes en Gráficos Estadísticos



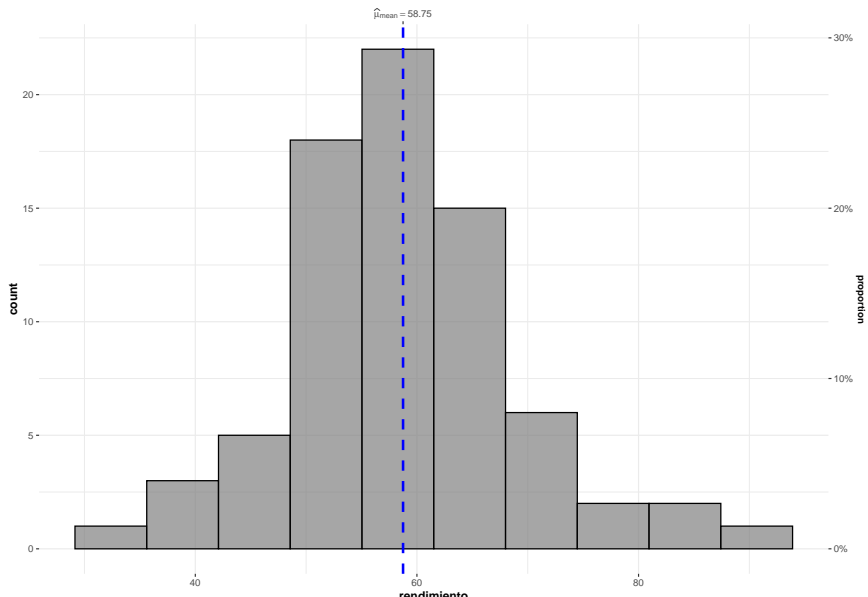
Biblioteca ggstatsplot

- Crea **gráficos elegantes** con detalles estadísticos incluidos.
- Permite incorporar **diferentes enfoques estadísticos**:
- pruebas **paramétricas**
- pruebas **no paramétricas**
- inferencia **bayesiana**
- pruebas **robustas**
- calcular **tamaños del efecto**
- identificar *outliers*

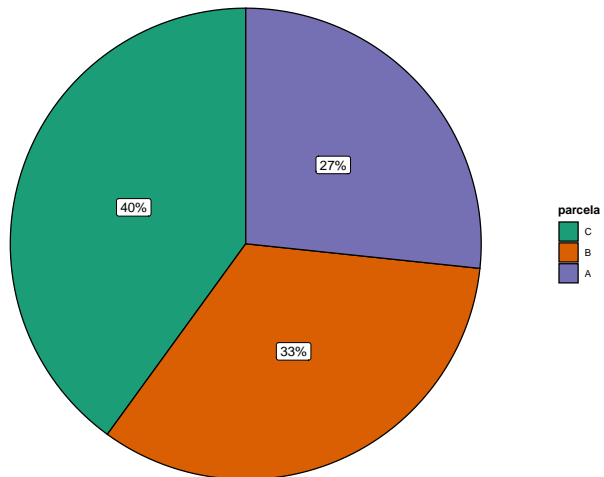
Funciones de ggstatsplot por tipo de gráfico

Tipo de gráfico	Función	Argumentos clave
Gráfico de pastel	<code>ggpiestats</code>	<code>data, x</code>
Gráfico de barras simples	<code>ggbarstats</code>	<code>data, x</code>
Gráfico de barras múltiples	<code>ggbarstats</code>	<code>data, x, fill = grouping.var</code>
Gráfico de barras compuestas	<code>ggbarstats</code>	<code>data, x, fill = grouping.var, position = "fill"</code>
Histograma	<code>gghistostats</code>	<code>data, x</code>
Polígono de frecuencias	<code>gghistostats</code>	<code>data, x, grouping.var</code>

Descripción de una variable cuantitativa



Descripción de una variable cualitativa



Comparación de Grupos Independientes: variable cuantitativa

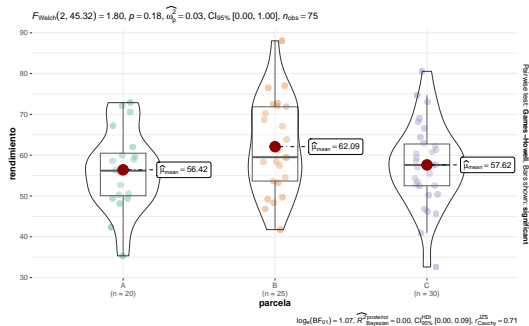
- **Ejemplo:** Imaginemos que queremos comparar el rendimiento de un cultivo (variable continua) en tres parcelas diferentes (grupos independientes).

	parcela	rendimiento	fortificado
1	A	49.39524	No
2	A	52.69823	No
3	A	70.58708	No

- Se emplea la función `ggbetweenstats` para visualizar un **gráfico de dispersión por grupo**

Comparación de Grupos Independientes: variable cuantitativa

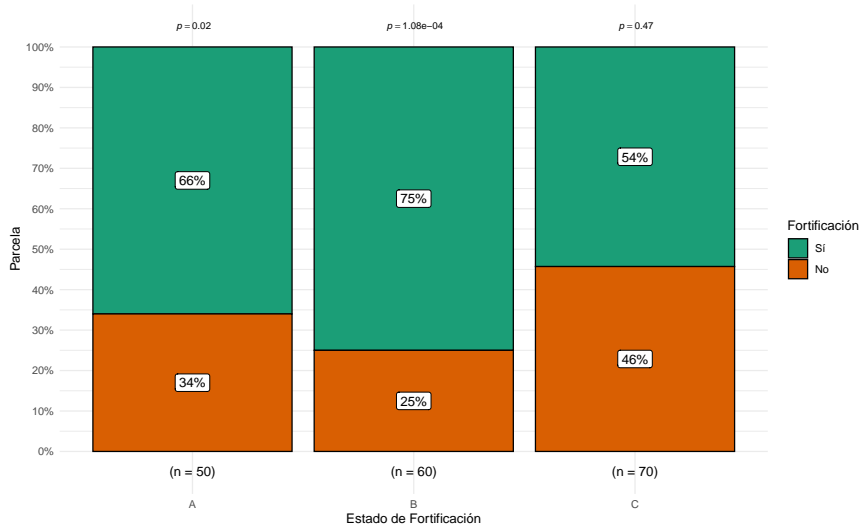
```
# Comparación paramétrica  
grupos_Independientes <-  
  ggbetweenstats(  
    data = data, # datos  
    x = parcela, # v cual  
    y = rendimiento, # v cuant  
    type = "parametric" # met param  
  )
```



Comparación de Grupos Independientes: variable cualitativa

Comparación de Parcelas por Estado de Fortificación

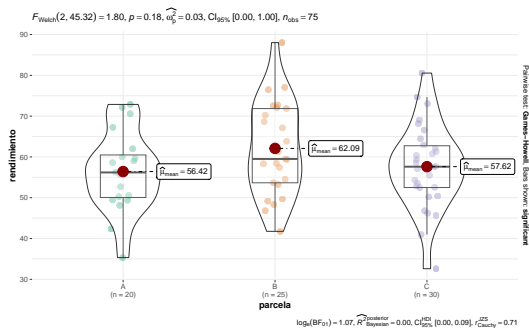
$\chi^2_{\text{Pearson}}(2) = 6.12, p = 0.05, \hat{V}_{\text{Cramer}} = 0.15, \text{CI}_{95\%} [0.00, 0.30], n_{\text{obs}} = 180$



$\log_2(\text{BF}_{01}) = -0.13, \hat{V}_{\text{posterior}} = 0.16, \text{CI}_{\text{ETI}}^{95\%} [0.00, 0.30], a_{\text{Bayes-Dickey}} = 1.00$

En la visualización de la comparación se muestra

- La **distribución de los datos** para cada grupo y medidas resúmenes.
- Los **resultados de la prueba estadística** (estadístico de prueba ANOVA, grados de libertad, valor p).
- El **tamaño del efecto** con su intervalo de confianza .



Otras Funciones Importantes (5 minutos)

- **ggwithinstats()**: Para comparar **medidas repetidas** dentro de los mismos sujetos o grupos a lo largo del tiempo .
- **ggscatterstats()**: Para visualizar la **relación entre dos variables continuas**, incluyendo correlación y regresión .
- **ggbarstats()**: Para analizar la **asociación entre variables categóricas** .
- **Funciones con prefijo group_**: Para realizar análisis **estratificados** por una tercera variable .

Conclusión y Recursos (3 minutos)

- **ggstatsplot** es una herramienta poderosa y fácil de usar para crear gráficos estadísticos directamente en R .
- **Integra la visualización y el análisis estadístico**, facilitando la exploración de datos y la comunicación de resultados .
- Es ideal para **informes científicos y académicos** gracias a la riqueza de detalles estadísticos que incorpora .

Bibliografía

- Recursos para seguir aprendiendo:
 - Documentación oficial del paquete `ggstatsplot`.
 - Sitio web del proyecto `easystats`.
 - Tutoriales y ejemplos en línea.
 - Comunidad activa en foros y plataformas como GitHub.