

# Ejercitación Práctica

Unidad: Visualización de Datos con R

Prof. Nicolás Sidicaro

2025-11-01

## Índice

<b>1</b>	<b>Introducción</b>	<b>2</b>
1.1	Objetivos generales . . . . .	2
1.2	Consideraciones para la ejercitación . . . . .	2
<b>2</b>	<b>Encuentro 5: Series de Tiempo</b>	<b>3</b>
2.1	Objetivos del encuentro . . . . .	3
2.2	Ejercitación . . . . .	3
<b>3</b>	<b>Encuentro 6: Variables Numéricas</b>	<b>5</b>
3.1	Parte 1: Relaciones entre Variables Cuantitativas . . . . .	5
3.2	Parte 2: Distribuciones de Variables Numéricas . . . . .	6
<b>4</b>	<b>Encuentro 7: Variables Cualitativas vs Cuantitativas</b>	<b>8</b>
4.1	Objetivos del encuentro . . . . .	8
4.2	Ejercitación . . . . .	8
<b>5</b>	<b>Encuentro 8: Modificaciones a los Gráficos</b>	<b>10</b>
5.1	Objetivos del encuentro . . . . .	10
5.2	Ejercitación . . . . .	10
5.3	Consignas específicas . . . . .	11
5.4	Criterios de evaluación . . . . .	11
<b>6</b>	<b>Anexo: Recursos y Referencias</b>	<b>12</b>
6.1	Sitios web recomendados . . . . .	12
6.2	Paquetes de R utilizados . . . . .	12
6.3	Estructura recomendada para scripts . . . . .	12
6.4	Consejos finales . . . . .	13

# 1 Introducción

El presente documento contiene la ejercitación práctica correspondiente a las clases de visualización de datos con R. Los ejercicios propuestos tienen como objetivo principal que los estudiantes desarrollen habilidades para replicar y adaptar visualizaciones profesionales utilizando datos reales de Argentina.

Todos los ejercicios se basan en gráficos disponibles en el sitio **Argendata** (<https://argendata.fund.ar/>), una plataforma de acceso abierto que presenta datos sobre Argentina de forma visual y accesible.

## 1.1 Objetivos generales

- Aplicar los conceptos teóricos de visualización de datos en casos prácticos
- Desarrollar habilidad para replicar gráficos profesionales
- Comprender la relación entre tipo de datos y tipo de visualización
- Dominar las técnicas de personalización en `ggplot2`
- Practicar la importación, limpieza y transformación de datos

## 1.2 Consideraciones para la ejercitación

Para todos los ejercicios presentados a continuación, se deberá:

1. **Replicar los componentes esenciales** de los gráficos presentados
2. **Incluir título, subtítulo y fuente** de los datos
3. **Seleccionar una paleta de colores** apropiada
4. **Mostrar los países o categorías** señalados en los gráficos originales
5. **Formatear correctamente** los datos (especialmente porcentajes cuando corresponda)

## 2 Encuentro 5: Series de Tiempo

### 2.1 Objetivos del encuentro

- Dominar la visualización de datos temporales
- Comprender el manejo de fechas en R
- Aplicar técnicas de series múltiples y faceting
- Trabajar con escalas temporales

### 2.2 Ejercitación

A continuación se presentan seis ejercicios de complejidad creciente para trabajar con series de tiempo. Los estudiantes deberán reconstruir los gráficos indicados, prestando especial atención al formato de las fechas, las etiquetas de los ejes y la presentación de múltiples series.

#### 2.2.1 Ejercicio 1: Gasto público en relación al PIB

**Fuente:** <https://argendata.fund.ar/graficos/gasto-publico-en-relacion-al-pib/>

**Descripción:** Reconstruir el gráfico de serie temporal que muestra la evolución del gasto público como porcentaje del PIB.

**Elementos a considerar:**

- Eje temporal correctamente formateado
- Escala del eje vertical en porcentajes
- Título y subtítulo descriptivos
- Fuente de los datos

#### 2.2.2 Ejercicio 2: Gasto público social

**Fuente:** <https://argendata.fund.ar/graficos/gasto-publico-social/>

**Descripción:** Replicar el gráfico de evolución del gasto público social.

**Elementos a considerar:**

- Manejo de series temporales largas
- Formateo apropiado de valores
- Claridad en las etiquetas

#### 2.2.3 Ejercicio 3: Gasto público por país

**Fuente:** <https://argendata.fund.ar/graficos/gasto-publico-de-paises-seleccionados/>

**Descripción:** Reconstruir el gráfico comparativo del gasto público de países seleccionados a lo largo del tiempo.

**Elementos a considerar:**

- Múltiples series en un mismo gráfico
- Uso efectivo del color para distinguir países
- Leyenda clara y bien posicionada
- Manejo de superposición de líneas

### 2.2.4 Ejercicio 4: Crecimiento del PIB

**Fuente:** <https://argendata.fund.ar/graficos/cambio-en-el-pib-per-capita-en-la-historia-reciente/>

**Descripción:** Replicar el gráfico que muestra el cambio en el PIB per cápita en la historia reciente.

**Elementos a considerar:**

- Visualización de cambios porcentuales
- Manejo de valores positivos y negativos
- Líneas de referencia (si corresponde)

### 2.2.5 Ejercicio 5: PIB de Latinoamérica

**Fuente:** <https://argendata.fund.ar/graficos/participacion-de-argentina-brasil-y-mexico-en-el-pib-de-america-latina/>

**Descripción:** Reconstruir el gráfico que muestra la participación de Argentina, Brasil y México en el PIB de América Latina.

**Elementos a considerar:**

- Gráfico de áreas o líneas apiladas
- Representación de porcentajes que suman 100%
- Diferenciación clara entre países

### 2.2.6 Ejercicio 6: VAB per cápita por provincia

**Fuente:** <https://argendata.fund.ar/graficos/vab-de-cada-provincia/>

**Descripción:** Replicar el gráfico del Valor Agregado Bruto per cápita de cada provincia.

**Elementos a considerar:**

- Múltiples series (una por provincia)
- Estrategias para evitar saturación visual
- Posible uso de faceting o selección de provincias destacadas
- Manejo de la leyenda con muchos elementos

## 3 Encuentro 6: Variables Numéricas

### 3.1 Parte 1: Relaciones entre Variables Cuantitativas

#### 3.1.1 Objetivos

- Visualizar correlaciones y relaciones numéricas
- Dominar el uso de scatter plots
- Incorporar múltiples dimensiones en gráficos de dispersión
- Aplicar técnicas de regresión en visualizaciones

#### 3.1.2 Ejercitación

Los siguientes ejercicios progresivamente aumentan en complejidad, incorporando elementos como color, múltiples escalas y líneas de tendencia.

**3.1.2.1 Ejercicio 1 (Fácil): Cambio en los destinos de exportación** Fuente: <https://argendata.fund.ar/graficos/cambio-en-los-destinos-de-exportacion/>

**Descripción:** Reconstruir el gráfico de dispersión básico.

**Elementos a considerar:**

- Scatter plot simple
- Etiquetado de puntos relevantes
- Ejes claramente identificados

**3.1.2.2 Ejercicio 2 (Fácil): Participación de los sectores en el empleo vs PIB** Fuente: <https://argendata.fund.ar/graficos/participacion-de-los-sectores-en-el-empleo-versus-su-participacion-en-el-pib/>

**Descripción:** Replicar el gráfico de dispersión incorporando color por categoría.

**Elementos a considerar:**

- Uso de color para distinguir sectores
- Leyenda apropiada
- Posible inclusión de línea de 45 grados como referencia

**3.1.2.3 Ejercicio 3 (Medio): IDH y penalidad por presiones planetarias** Fuente: <https://argendata.fund.ar/graficos/idh-y-penalidad-por-presiones-planetarias/>

**Descripción:** Reconstruir gráfico con mayor complejidad, incorporando dos escalas simultáneamente.

**Elementos a considerar:**

- Múltiples variables en un mismo gráfico
- Uso de tamaño o forma de puntos
- Manejo de dos escalas diferentes
- Claridad en la comunicación de múltiples dimensiones

**3.1.2.4 Ejercicio 4 (Medio): PIB per cápita y formalidad en América Latina** Fuente: <https://argendata.fund.ar/graficos/pib-per-capita-y-formalidad-en-america-latina/>

**Descripción:** Replicar gráfico de dispersión con línea de regresión.

**Elementos a considerar:**

- Scatter plot con línea de tendencia
- Método de regresión (`geom_smooth()`)
- Intervalo de confianza (opcional)
- Etiquetado de países relevantes

### 3.1.2.5 Ejercicio 5 (Difícil): Exceso de desigualdad en Argentina y América Latina

**Fuente:** <https://argendata.fund.ar/graficos/exceso-de-desigualdad-en-argentina-y-america-latina/>

**Descripción:** Reconstruir visualización compleja con múltiples elementos.

**Elementos a considerar:**

- Gráfico de dispersión avanzado
- Múltiples grupos o categorías
- Anotaciones y referencias visuales
- Posible faceting por región o período

## 3.2 Parte 2: Distribuciones de Variables Numéricas

### 3.2.1 Objetivos

- Comprender y aplicar gráficos de distribución
- Dominar boxplots, histogramas y violin plots
- Comparar distribuciones entre grupos

### 3.2.2 Ejercitación

Los siguientes ejercicios requieren adaptar los datos de gráficos existentes para crear visualizaciones de distribución.

**3.2.2.1 Ejercicio 1: Boxplot del IDH según región** **Fuente:** <https://argendata.fund.ar/graficos/idh-y-penalidad-por-presiones-planetarias/>

**Consigna:** Realizar un boxplot que muestre el Índice de Desarrollo Humano (IDH) según región geográfica.

**Elementos a considerar:**

- Agrupación por región
- Identificación de outliers
- Ordenamiento de regiones (por mediana, alfabéticamente, etc.)
- Interpretación de la distribución

**3.2.2.2 Ejercicio 2: Histograma de relevancia del agro y la pesca** **Fuente:** <https://argendata.fund.ar/graficos/importancia-del-agro-y-la-pesca-en-argentina-y-otros-paises/>

**Consigna:** Realizar un histograma que muestre la relevancia del agro y la pesca como porcentaje del PIB total en el año 2021.

**Elementos a considerar:**

- Selección apropiada del número de bins
- Escala del eje horizontal (porcentajes)
- Posible inclusión de línea de densidad
- Interpretación de la forma de la distribución

**3.2.2.3 Ejercicio 3: Violin plot por país** **Fuente:** <https://argendata.fund.ar/graficos/importancia-del-agro-y-la-pesca-en-argentina-y-otros-paises/>

**Consigna:** Realizar un gráfico de violín que muestre la relevancia del agro y la pesca como porcentaje del PIB total en todo el período disponible, separado por país (países seleccionados: Argentina, Uruguay, Colombia).

**Elementos a considerar:**

- Violin plot con tres categorías
- Comparación visual de distribuciones
- Posible inclusión de boxplot interno
- Uso efectivo del color
- Interpretación de las diferencias entre países

## 4 Encuentro 7: Variables Cualitativas vs Cuantitativas

### 4.1 Objetivos del encuentro

- Dominar técnicas de comparación entre grupos
- Aplicar gráficos de barras en sus diferentes variantes
- Utilizar Cleveland dot plots para comparaciones efectivas
- Comprender cuándo usar cada tipo de visualización

### 4.2 Ejercitación

Los ejercicios propuestos cubren diferentes técnicas para visualizar la relación entre variables categóricas y numéricas, desde gráficos de barras simples hasta visualizaciones más complejas de brechas y composiciones.

#### 4.2.1 Ejercicio 1 (Fácil): Principales países productores de carne aviar

**Fuente:** <https://argendata.fund.ar/graficos/principales-paises-productores-de-carne-aviar/>

**Descripción:** Replicar gráfico de barras simple.

**Elementos a considerar:**

- Gráfico de barras horizontales o verticales
- Ordenamiento de categorías (usualmente de mayor a menor)
- Formateo de números en el eje
- Selección de color apropiado

#### 4.2.2 Ejercicio 2 (Medio): Desigualdad pre y post intervención fiscal

**Fuente:** <https://argendata.fund.ar/graficos/desigualdad-pre-y-post-intervencion-fiscal-en-paises-seleccionados/>

**Descripción:** Replicar gráfico de brechas utilizando Cleveland dot plot.

**Elementos a considerar:**

- Dot plot con dos puntos por categoría (pre y post)
- Líneas conectando los puntos para visualizar la brecha
- Uso de color para diferenciar pre/post
- Ordenamiento efectivo de países
- Leyenda clara

#### 4.2.3 Ejercicio 3 (Medio): Egresos del gobierno en relación al PIB

**Fuente:** <https://argendata.fund.ar/graficos/egresos-del-gobierno-en-relacion-al-pib/>

**Descripción:** Replicar gráfico de barras con categorías anidadas e incorporar leyenda de tipo de gasto. Adicionalmente, ordenar las barras en orden descendente.

**Elementos a considerar:**

- Barras agrupadas por tipo de gasto
- Ordenamiento descendente total
- Leyenda identificando tipos de gasto
- Posible uso de `fct_reorder()` para ordenamiento
- Paleta de colores diferenciando categorías



#### 4.2.4 Ejercicio 4 (Medio): Gasto público por nivel de gobierno

**Fuente:** <https://argendata.fund.ar/graficos/gasto-publico-por-nivel-de-gobierno/>

**Descripción:** Replicar gráfico de barras apiladas.

**Elementos a considerar:**

- Barras apiladas al 100% o en valores absolutos
- Diferenciación clara entre niveles de gobierno
- Ordenamiento lógico de categorías
- Etiquetas dentro o fuera de las barras
- Interpretación de proporciones

#### 4.2.5 Ejercicio 5 (Difícil): Brechas de participación en la tasa de empleo

**Fuente:** <https://argendata.fund.ar/graficos/empleo-por-genero-y-provincia/>

**Descripción:** Armar un gráfico que muestre las brechas de participación de la tasa de empleo según sexo por provincia.

**Elementos a considerar:**

- Visualización de brechas de género
- Comparación entre múltiples provincias
- Ordenamiento efectivo (por tamaño de brecha, alfabético, etc.)
- Uso de color para distinguir géneros
- Posible uso de dot plot con líneas
- Claridad en la comunicación de la magnitud de las brechas

## 5 Encuentro 8: Modificaciones a los Gráficos

### 5.1 Objetivos del encuentro

- Aplicar estilos profesionales a visualizaciones
- Incorporar anotaciones y elementos narrativos
- Dominar la personalización avanzada en `ggplot2`
- Comprender el uso de temas predefinidos

### 5.2 Ejercitación

Esta práctica tiene por objetivo que los estudiantes tomen los gráficos replicados de Argendata en los encuentros anteriores y realicen cambios sustantivos en la presentación de **al menos 5 gráficos**.

#### 5.2.1 Elementos a modificar

Los estudiantes deberán incorporar mejoras significativas en los siguientes aspectos:

##### 5.2.1.1 1. Anotaciones

- Agregar texto explicativo sobre los datos
- Destacar puntos o períodos relevantes
- Incluir interpretaciones breves
- Utilizar `annotate()` o `geom_text()/geom_label()`

##### 5.2.1.2 2. Estilos predefinidos Aplicar temas profesionales como:

- `theme_economist()` (estilo The Economist)
- `theme_wsj()` (estilo Wall Street Journal)
- `theme_fivethirtyeight()` (estilo FiveThirtyEight)
- `theme_minimal()` con personalizaciones
- `theme_bw()` con ajustes personalizados

**Nota:** Estos temas requieren el paquete `ggthemes`

##### 5.2.1.3 3. Manejo de colores

- Aplicar paletas profesionales (`viridis`, `ColorBrewer`, etc.)
- Considerar accesibilidad (daltonismo)
- Usar colores con propósito comunicativo
- Mantener consistencia visual
- Explorar `scale_color_*` y `scale_fill_*`

##### 5.2.1.4 4. Otros elementos vistos en el encuentro

- Ajustes de tipografía y tamaños
- Mejoras en leyendas (posición, formato)
- Optimización de etiquetas de ejes
- Líneas de referencia estratégicas
- Mejoras en la relación señal-ruido
- Aplicación de principios de buena visualización

### 5.3 Consignas específicas

1. **Seleccionar 5 gráficos** de los elaborados en encuentros anteriores
2. **Documentar los cambios** realizados en cada gráfico
3. **Justificar las decisiones** de diseño tomadas
4. **Comparar** versión original vs versión mejorada
5. **Presentar** ambas versiones en el documento final

### 5.4 Criterios de evaluación

- Aplicación apropiada de principios de diseño
- Coherencia visual entre elementos
- Mejora efectiva de la comunicación de datos
- Creatividad en las soluciones aplicadas
- Documentación clara de las modificaciones

## 6 Anexo: Recursos y Referencias

### 6.1 Sitios web recomendados

- **Argendata:** <https://argendata.fund.ar/>
- **From Data to Viz:** <https://www.data-to-viz.com/>
- **R Graph Gallery:** <https://r-graph-gallery.com/>
- **ggplot2 extensions:** <https://exts.ggplot2.tidyverse.org/>

### 6.2 Paquetes de R utilizados

```
# Instalación de paquetes necesarios
install.packages(c(
  "ggplot2",      # Visualización
  "dplyr",        # Manipulación de datos
  "readr",        # Lectura de datos
  "readxl",       # Lectura de Excel
  "tidyr",        # Limpieza de datos
  "lubridate",    # Manejo de fechas
  "scales",       # Formateo de escalas
  "ggthemes",     # Temas adicionales
  "viridis",      # Paletas de colores
  "patchwork"     # Composición de gráficos
))
```

### 6.3 Estructura recomendada para scripts

```
# 1. Cargar librerías
library(ggplot2)
library(dplyr)
library(readr)

# 2. Importar datos
datos <- read_csv("ruta/archivo.csv")

# 3. Explorar datos
glimpse(datos)
summary(datos)

# 4. Limpiar y transformar
datos_limpios <- datos %>%
  filter(...) %>%
  mutate(...) %>%
  select(...)

# 5. Crear visualización
grafico <- ggplot(datos_limpios, aes(x = ..., y = ...)) +
  geom_...(...) +
  labs(...) +
  theme_...(...)

# 6. Guardar gráfico
```

```
ggsave("grafico.png", grafico, width = 10, height = 6, dpi = 300)
```

## 6.4 Consejos finales

- **Iterar:** Los buenos gráficos requieren múltiples versiones
- **Simplicidad:** Menos suele ser más en visualización
- **Audiencia:** Considerar siempre a quién va dirigido el gráfico
- **Contexto:** Proporcionar información necesaria para interpretación
- **Reproducibilidad:** Documentar el código y las fuentes de datos

---

**Contacto:** Nicolás Sidicaro - UBA

**Fecha de última actualización:** 2025-11-01