



# R para Finanzas

## Guía de Ejercicios 4

**Profesor:** *Víctor Macías E.*

```
# ¡No olvides instalar previamente estos paquetes!  
  
library(readr)      # Pregunta 1  
library(dplyr)      # Preguntas 2 a 9  
library(skimr)       # Pregunta 9  
library(kableExtra) # Pregunta 9
```

### Pregunta 1

Importa los datos descargados de World Development Indicators en la guía anterior.

```
wbData1 <- read_rds("datos/wbData_1970_2023.rds")  
  
wbData1 <- wbData1 |> arrange(country, year) # Ordenando por país y año
```

### Pregunta 2

Renombrar las siguientes variables, usando nombres que faciliten su interpretación:

- NY.GDP.PCAP.PP.KD
- SP.POP.TOTL
- EN.GHG.CO2.MT.CE.AR5
- EN.GHG.ALL.MT.CE.AR5
- EN.GHG.CO2.PC.CE.AR5
- EN.GHG.ALL.PC.CE.AR5

- EN.GHG.ALL.LU.MT.CE.AR5

```
# GDP per capita, PPP (constant 2021 international $)
# NY.GDP.PCAP.PP.KD

# Population
# SP.POP.TOTL

# Carbon dioxide (CO2) emissions (total) excluding LULUCF (Mt CO2e)
# EN.GHG.CO2.MT.CE.AR5

# Total greenhouse gas emissions excluding LULUCF (Mt CO2e)
# EN.GHG.ALL.MT.CE.AR5

# Carbon dioxide (CO2) emissions excluding LULUCF per capita (tCO2e/capita)
# EN.GHG.CO2.PC.CE.AR5

# Total greenhouse gas emissions excluding LULUCF per capita (tCO2e/capita)
# EN.GHG.ALL.PC.CE.AR5

# Total greenhouse gas emissions including LULUCF (Mt CO2e)
# EN.GHG.ALL.LU.MT.CE.AR5

wbData2 <- wbData1 |>
  rename(gdppc = NY.GDP.PCAP.PP.CD,
         population = SP.POP.TOTL,
         co2_total_exc = EN.GHG.CO2.MT.CE.AR5,
         ghg_total_exc = EN.GHG.ALL.MT.CE.AR5,
         co2_pcap_exc = EN.GHG.CO2.PC.CE.AR5,
         ghg_pcap_exc = EN.GHG.ALL.PC.CE.AR5,
         ghg_total_inc = EN.GHG.ALL.LU.MT.CE.AR5)
```

### Pregunta 3

Cambia los nombres de las regiones a español. Por ejemplo, cambia la región “Middle East & North Africa” a “Medio Oriente y Norte de Africa”. ¿Qué países y regiones no tienen asignada una región específica? En el caso de los países sin región asigna la que corresponde.

```
# Número de observaciones por region
wbData2 |> group_by(region) |> count()
```

```
# Observaciones con valores NA en region
wbData2 |> filter(is.na(region)) |>
  group_by(country) |> count()
```

```
wbData2 <- wbData2 |>
  mutate(region2 = case_when(
    country == "Viet Nam" ~ "East Asia & Pacific",
    country == "Czechia" ~ "Europe & Central Asia",
    country == "High income" ~ "Aggregates",
    country == "Latin America & Caribbean" ~ "Aggregates",
    country == "Low income" ~ "Aggregates",
    country == "Lower middle income" ~ "Aggregates",
    country == "Not classified" ~ "Aggregates",
    country == "Sub-Saharan Africa" ~ "Aggregates",
    country == "Upper middle income" ~ "Aggregates",
    TRUE ~ region # keep other values unchanged
  ))
```

```
# Número de observaciones por region2
wbData2 |> group_by(region2) |> count()
```

```
# Cambiando los valores de region2 a español
wbData2 <- wbData2 |>
  mutate(region3 = recode(region2,
    "East Asia & Pacific" = "Este Asiático y Pacífico",
    "Europe & Central Asia" = "Europa y Asia Central",
    "Latin America & Caribbean" = "América Latina y el Caribe",
    "Middle East & North Africa" = "Medio Oriente y Norte de Africa",
    "North America" = "Norteamérica",
    "South Asia" = "Asia del Sur",
    "Sub-Saharan Africa" = "Africa Subsahariana"
  ))
```

```
# Número de observaciones por region3
wbData2 |> group_by(region3) |> count()
```

## Pregunta 4

¿Cuántos países tiene la base de datos?, ¿Cuántos países por región tiene la base de datos?

```
# Descartando observaciones donde region3 es Aggregates
wbData3 <- wbData2 |>
  filter(region3 != "Aggregates")
```

```
# Número de países
wbData3 |> distinct(country) |> count()
```

```
# Número y porcentaje de países por región
wbData3 |> group_by(region3) |> distinct(country) |>
  count() |> ungroup() |>
  mutate(prop_region=round((n/sum(n))*100,1))
```

## Pregunta 5

- (a) Identifica a los 5 países con mayores emisiones de GEI per cápita (excluyendo LULUCF) en 2022

```
wbData3 |> filter(year == 2022) |>
  select(country, ghg_pcap_exc) |>
  slice_max(ghg_pcap_exc, n = 5)
```

- (b) Identifica a los 5 países con mayores emisiones de CO2 per cápita (excluyendo LULUCF) en 2022

```
wbData3 |> filter(year == 2022) |>
  select(country, co2_pcap_exc) |>
  slice_max(co2_pcap_exc, n = 5)
```

## Pregunta 6

- (a) Identifica a los 5 países con mayores emisiones totales de GEI (excluyendo LULUCF) en 2022

```
wbData3 |> filter(year == 2022) |>
  select(country, ghg_total_exc) |>
  slice_max(ghg_total_exc, n = 5)
```

- (b) Identifica a los 5 países con mayores emisiones totales de CO2 (excluyendo LULUCF) en 2022

```
wbData3 |> filter(year == 2022) |>
  select(country, co2_total_exc) |>
  slice_max(co2_total_exc, n = 5)
```

## Pregunta 7

Si ordenas los países de mayor a menor nivel de emisiones de GEI (excluyendo LULUCF) en 2022 ¿Qué lugar ocupa Chile en el ranking? Para mayor información, se sugiere revisar:

<https://data.worldbank.org/indicator/EN.GHG.ALL.MT.CE.AR5>

```
wbData3 |> filter(year == 2022) |>
  arrange(desc(ghg_total_exc)) |>
  mutate(ranking = seq(1, 217, 1)) |>
  filter(country == "Chile") |>
  select(ranking, country, ghg_total_exc)
```

## Pregunta 8

Identifica a los países con los mayores niveles de emisiones totales de GEI y de CO2 en cada una de las regiones.

```
# Países con mayores emisiones totales de GEI (excluyendo LULUCF) por región
wbData3 |> filter(year == 2022) |> group_by(region3) |>
  select(region3, country, ghg_total_exc) %>%
  slice_max(ghg_total_exc, n = 1)
```

```
# Países con mayores emisiones totales de CO2 (excluyendo LULUCF) por región
wbData3 |> filter(year == 2022) |> group_by(region3) |>
  select(region3, country, co2_total_exc) %>%
  slice_max(co2_total_exc, n = 1)
```

## Pregunta 9

- (a) Identifica a los 10 países con un mayor aumento porcentual de sus emisiones totales de GEI entre los años 1990 y 2022.

```
wbData3 |> group_by(country) |>
  filter(year %in% c(1990, 2022)) |>
  mutate(
    variacion =
      round(((ghg_total_exc - lag(ghg_total_exc))/lag(ghg_total_exc))*100,2)
  ) |>
  ungroup() |>
  filter(year == 2022) |>
  select(country, variacion) |>
  slice_max(variacion, n=10)
```

- (b) Identifica a los 10 países con una mayor disminución porcentual de sus emisiones totales de GEI entre los años 1990 y 2022.

```
wbData3 |> group_by(country) |>
  filter(year %in% c(1990, 2022)) |>
  mutate(
    variacion =
      round(((ghg_total_exc - lag(ghg_total_exc))/lag(ghg_total_exc))*100,2)
  ) |>
  ungroup() |>
  filter(year == 2022) |>
  select(country, variacion) |>
  slice_min(variacion, n=10)
```

## Pregunta 10

Construye una tabla con estadística descriptiva de las emisiones de GEI (excluyendo LULUCF) per cápita para el año 2022, separando por regiones.

```
wbData3 |>
  filter(year == 2022) |>
  select(region, ghg_pcap_exc) |>
  group_by(region3) |>
  skim() |>
  yank("numeric") |>
  select(-c("skim_variable", "n_missing", "complete_rate", "hist")) |>
  kbl(booktabs = T, digits = 2, linesep = "",
      col.names = c("Región", "Media", "Desv. Est.", "Min",
                    "p25", "p50", "p75", "Max")) |>
  kable_styling(font_size = 10, latex_options = c("striped", "hold_position"))
```

## Pregunta 11

Guarda los datos después de la última modificación realizada en la pregunta 4.

```
wbData3 |> write_rds("datos/Guia4/wbData3_final.rds")
```