Intro_to_data_vis_lab_5

woessner.21

2025-10-08

```
# Load needed packages
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(readr)

# 1. Read in both datasets
wdi_data <- read.csv("C:/Users/Caleb Woessner/Downloads/WDICSV.csv")
wdi_country <- read.csv("C:/Users/Caleb Woessner/Downloads/WDICountry.csv")
```

```
names(wdi_data)
```

```
##  [1] "Country.Name"   "Country.Code"   "Indicator.Name" "Indicator.Code"
##  [5] "X1960"          "X1961"          "X1962"          "X1963"
##  [9] "X1964"          "X1965"          "X1966"          "X1967"
## [13] "X1968"          "X1969"          "X1970"          "X1971"
## [17] "X1972"          "X1973"          "X1974"          "X1975"
## [21] "X1976"          "X1977"          "X1978"          "X1979"
## [25] "X1980"          "X1981"          "X1982"          "X1983"
## [29] "X1984"          "X1985"          "X1986"          "X1987"
## [33] "X1988"          "X1989"          "X1990"          "X1991"
## [37] "X1992"          "X1993"          "X1994"          "X1995"
## [41] "X1996"          "X1997"          "X1998"          "X1999"
## [45] "X2000"          "X2001"          "X2002"          "X2003"
## [49] "X2004"          "X2005"          "X2006"          "X2007"
## [53] "X2008"          "X2009"          "X2010"          "X2011"
## [57] "X2012"          "X2013"          "X2014"          "X2015"
## [61] "X2016"          "X2017"          "X2018"          "X2019"
## [65] "X2020"          "X2021"          "X2022"          "X2023"
## [69] "X2024"
```

```r
# 2. Filter only for GDP per capita (current US$)
gdp_data <- wdi_data %>%
  filter(`Indicator.Name` == "GDP per capita (current US$)")
```

```r
# 3. Join with region info using Country Code
merged_data <- gdp_data %>%
  left_join(wdi_country %>% select(`Country.Code`, Region), by = "Country.Code")
```

```r
# 4. Randomly sample 10 and 400 countries
set.seed(42)  # for reproducibility
small_world <- merged_data %>% sample_n(10)
large_world <- merged_data %>% sample_n(266)
```

```r
# 5. Write to CSVs (same folder as your HTML file)
write_csv(small_world, "C:/Users/Caleb Woessner/OneDrive/Documents/small_world.csv")
write_csv(large_world, "C:/Users/Caleb Woessner/OneDrive/Documents/large_world.csv")
```