

Assignment 3

Data visualization for economics

Due date: December 5th, 2025

Instructions

This assignment focuses on global taxation data and visualization techniques. It should be done using RMarkdown and the final output should be a pdf file (2 pts). The answers should show the R code used to produce the results as well as the results themselves. Please also provide the dataset you have produced or downloaded as a CSV file.

First part

1. Scrape Tax Data from Wikipedia

Use R to scrape the table containing top marginal income tax rates by country from the following Wikipedia page: [List of countries by tax rates](#). (3 pts)

2. Clean the Table by Changing the Column Names

Rename the columns in the scraped table to standard names (e.g., Country, Top Marginal Income Tax Rate, etc.). (1 pt)

3. Clean the Table by Removing Non-Numeric Elements

Clean the content of the cells by removing footnotes and non-numeric characters. Ensure the tax rates are numeric values. (1 pts)

Second Part

If you did not manage to complete the first part of the assignment, you can work on the second part using [this csv file](#)

4. Add a Continent Column

Create a new column in your dataset indicating the continent of each country (*Hint: you may want to use the `countrycode` package*). (1 pt)

5. Analyze Average Rates by Continent

Answer the following questions: (1 pts)

- Which continent has the highest average top marginal income tax rate?

- Which continent comes second?
- Which continent has the highest average VAT rate?

6. Graph: Average Top Income Tax Rate by Continent

Create a bar chart showing the average top marginal income tax rate for each continent. (1 pt)

7. Graph: Compare VAT and Income Tax Rates

Create a single grouped bar chart comparing the average VAT rates and top marginal income tax rates for each continent. (1.5 pts)

8. Map: Top Marginal Income Tax Rate by Country

Build a map visualizing the top marginal income tax rate for each country. (2 pts)
Include:

- A clear title and legend.
- Gray shading for countries with missing data.

9. Handle Missing Data:

When the top marginal income tax rate is missing, replace it with the corporate income tax rate. Update your dataset accordingly. (1 pt)

10. Graph: New Income Tax Rate Variable

Create a graph to visualize the new variable that includes the substituted values for missing income tax rates. (1 pt)

11. Enhanced Map: Custom Colors

Use a `scale` function to change the colors on the map created in Question 10. Choose a color palette that enhances the storytelling of your visualization. (1 pt)

Third Part

12. Free visualization

From the Insee website ([here](#)) or another data producer (Banque de France, IMF, ILO etc.) pick a theme you are interested in (demography, wealth, labor market, poverty, firms, education etc.). Then, choose one / a combination of datasets and make a visualization. The visualization should tell us a story, and the visualization choices should be consistent with this story. Comment the visualization (3.5 pts).