Introduction to Web Scraping and Data Management for Social Scientists

ESS 2024—Exam 2

Question 1: Economic Performance and Democracy [20pts]

We want to know whether a higher GDP per capita is indicative of a better quality democracy. To find this out, you should collect the numbers for all countries for different democracy indices and the GDP (nominal) per capita. First, collect the tables containing the relevant information from these websites:

- List of countries by GDP (nominal) per capita
- Democracy indices (V-Dem)
- The Economist Democracy Index
- Bertelsmann Transformation Index

You need to first create a database that contains the information from the democracy indices.

1. Create a local SQLite database.
2. Create a table for each table you collect, making sure the data types are correct.
3. Write a query that combines the data in your analysis dataset you aim to run your analysis on.

Additionally, to what is already there, we want one average measure from the different categories for each index and bring them on the same scale (e.g., from 0 (lowest) to 1 (highest)). We can add the GDP data to a table as is.

Finally, you should produce a plot that:

- Shows the GDP per capita on the x-axis.
- Shows the average democracy score on the y-axis.
- Differentiates between the 3 indices using differently coloured dots.

Make sure to include the code for all steps in the main document and describe what the code does.