Module 3: R

Final Summative Assessment

Project steps:

1. Select two related datasets (i.e. that share a common column and could be joined).

Suggested sources include:

a. <https://open.canada.ca/en/open-data>

b. <https://data.ontario.ca/en/>

c. <https://open.toronto.ca/>

d. <https://stats.oecd.org/>

e. <https://datasetsearch.research.google.com/>

f. <https://archive.ics.uci.edu/ml/datasets.php>

g. <https://ec.europa.eu/eurostat/home>

h. <https://www.kaggle.com/datasets>

i. <https://data.nasdaq.com/>

j. <https://dataverse-harvard-edu.myaccess.library.utoronto.ca/>

k. <https://data.worldbank.org/>

l. <https://www150.statcan.gc.ca/n1/en/type/data>

m. <http://www.icpsr.umich.edu.myaccess.library.utoronto.ca/web/pages/ICPSR/index.html>

n. Student’s own data

2. Create an RMarkdown document in an RProject.

3. Import dataset.

4. Join 2 datasets.

5. Create 2 and present summary tables.

6. Create 2 data visualizations.

7. Write about findings.

8. Knit the report to PDF.

9. Create a shiny application out of one summary table or data visualization.

Guidelines

A. Write code that is reproducible and human-readable.

B. Customize your tables and plots.

C. Create a PDF report that does not include code and presents your results in a professional manner.

Notes for instructor:

- To prevent all the project work falling at the end of the week, the project steps above ideally would be be completed:

1. After Module 0 and before Module 2

2. After Module 2 and before Module 5

3. After Module 5 and before Module 7

4. After Module 5 and before Module 7

5. After Module 5 and before Module 7

6. After Module 7 and before Module 8

7. After Module 7

8. After Module 7

9. After Module 8

- Some stages should be checked and approved before the student has progressed too far. Specifically, Step 1 should be checked before the students begin Step 4, and the document from Step 2 should be test-knitted before the student begins Step 3.