

The Final Project

This class will culminate in a final project. The topic of the project is entirely up to you. Vitrally, the project must apply the skills we have learned in class to generate original, high quality data visualizations and analysis in R. The project can be completed individually or in groups of up to four students.

Content

Projects will: address an interesting and well-defined research question, use data that can help answer the question, intelligently employ data visualizations and present the visualizations clearly. In addition to the visualizations, projects should include a written discussion of the research question, data sources, methods and an interpretation of results.

Deadlines

- Wednesday, **March 31**: Send me an email with the following information; (1) the names of all group members (if you're working in a group) (2) a research question or topic. The topic can be general at this point, or if you are deciding between multiple topics, send me a short list of possible topics and I can provide advice.

- Wednesday, **May 5**: the project is due. Failure to meet the deadline will result in a lower grade.

I will set up meeting times with groups to discuss ideas and possible data sources.

Topics

Coming up with an interesting, well defined topic is one of the more challenging aspects of conducting research. The topic is entirely up to you and should be motivated first and foremost by what interests you. Approach me with questions as I can probably provide information regarding what is feasible in terms of obtaining data and conducting analysis.

An important aspect of your topic is that it will lend itself to data analysis and visualization.

Length

Paper lengths should be *roughly* the following:

Group of 1 student: 4 figures + 1500 words

Group of 2 students: 6 figures + 2000 words

Group of 3 students: 8 figures + 2500 words

Group of 4 students: 10 figures + 3000 words

These are just guidelines. Don't be overly concerned about length. Quality is more important than quantity. In addition to the paper, groups should submit their R code as well as any data sets the R code uses. The code should be able to run without errors, producing all the figures in the paper.

Alternatives to a Writing a Paper

If any students would like to suggest a different form for their final project, I am quite open to suggestions. A (virtual) poster or a website seem like potentially good mediums for a data visualization project. Please discuss expectations with me if you would like to pursue an alternative form for your project.

Evaluation

Groups will be given a mark out of 12 according to the following criteria:

(1) Depth of Analysis (5 marks)

- Data analysis should be used to answer the research question.
- There should be an underlying theory or narrative that motivates your topic and can help explain results.
- The data processing, analysis and visualizations should demonstrate the group's proficiency and sophistication with R.

(2) Completeness and Clarity (5 marks)

- Writing should be clear and concise, free of errors, and follow a natural progression.
- Products of data analysis (graphs, maps, tables etc.) should be easily understood, properly labeled/formatted and used to convey specific and relevant points.
- Secondary sources should be included and cited appropriately to give context for the research.

(3) Motivation, Social/Economic Relevance (2 marks)

- The discussion of the topic and exposition of findings should demonstrate that the topic and results are important and interesting. The findings should provide actionable information to some members of society.