Data 227 - Autumn 2023 Term project

Data Visualization & Communication - Trimble

Due: Friday, Dec 1, 2023, 11:59pm (Friday 9th week)

Visualization project - tell a story with graphs (and text)

For this assignment, you will create a series of four related visualizations that tell a story. Related visualizations could show an overview of a dataset, breakdowns and aggregations at different levels, or could compare or combine datasets on related topics to expand the strength of inferences (you might want to look up ancillary data like population, consumer price index, migration rates, etc.). Find a truth in the data, find a way to share this truth, and share some context about the dataset in which it was found in using the visualizations.

- You should investigate and briefly explain the origin and the meaning of the dataset. When, where is the data from and why might we care about it?
- At least two visualizations which encode more than 9 data elements. Data-poor (less than 10 numbers) graphics are permitted, but no more than two graphics should be considered data poor.
- A figure caption for each graphic. This can be short. You will probably need to number the graphics, and you have to split your text between body and caption.
- Describe the content of your graphs in the text. Include summary statistics when necessary, and you are welcome to include tables that may help your audience interpret your graph if appropriate. Try to avoid low-effectiveness visualizations. If a visualization is just as effective as a table, a table should be preferred.
- A citation for the source/provenance of your dataset. This must be in the report somewhere, but does not need to be in every caption. Include footnotes for other facts / data as needed.
- A paragraph of critique or commentary on your design (or a draft of your design) from another student in the class.
- A footnote or an entry in the bibliography must indicate the origin of the dataset in sufficient detail to permit it to be found, with dataset name, author, revision number in addition to the URL where it can be found today. Footnotes for background and other work must be adequate.
- Turn in a project proposal (or project draft), and we will look at it and give you the same kind of feedback.
- Visualizations should be free of correctable flaws; everything that needs a label must be labeled, fonts must be no smaller than half the font size of the text in the report; included images must not be grainy or illegible. Best practice is to write a caption for each figure that succinctly explains the figure.
- Most of your grade will be on the quality of your data reporting. Explaining what is in a dataset is difficult, and doing it in a way that is easy to read is even more so. Does the dataset have obvious utility or immediately raise questions? Can you answer at least some of those for someone unfamiliar with the dataset?
- Page limit: 5 pages including visualizations, not including bibliography or submitted code. Some people could do a good job in 3.5 pages.
- You do not have to use any specific tools to produce the visualization, but you must submit the substantial code you used to perform modeling and generate visualizations.
- This is an individual project. You are welcome to collaborate, look at the same data, and share code (with credit) but I expect one report per student, with a different angle on the data if two students are working together on a dataset.