Problem Set Instructions

Lauren Nichols

Problem Set Instructions

For Problem Sets 2 and 3, your task will be to create various visualizations and answer specific questions related to these visualizations.

When you are *explicitly* asked to make an **exploratory** figure, you should assume your audience is another data scientist or analyst on your internal team. In other words, you do not need to worry about data ink, highlighting an insight, or removing legends, **but the data should still be represented clearly and accurately**.

For **all other figures**, make sure that you:

1. Chose appropriate **chart types** and **metrics** that will reduce cognitive load and support accurate interpretation
2. Clearly state the intended take-away in the **title** and ensure it matches the data you are showing
3. Include and format **labels** (axis, variables, legends) **and annotations** in a way that promotes quick understanding and interpretation
4. Choose **scales** (ordering, orientation, axis limits, etc.) that support accurate interpretation
5. Use **color** effectively according to best practices discussed in class

Tools

You have the option to complete this assignment in R or Python, and can use any libraries for data processing (pandas, NumPy, dplyr, etc). For visualizations, you must only use the following libraries/packages:

- matplotlib
- seaborn
- plotnine

- altair
- plotly
- ggplot2 (and extension libraries)

Figure Post Processing

If you are struggling to implement a particular change using code, you may export the figure and make the necessary changes in Power Point, Illustrator, or other post-processing tool.

After post-processing, save the figure as an image and add the image to your Jupyter notebook/Quarto document.

Make sure to include:

1. the original code,
2. the original figure created by the code *before* post-processing, and
3. your final figure.

Submission

Format the assignment as a Jupyter notebook, Markdown, or Quarto document.

Export the final document as a PDF, ensuring that all code and outputs are displayed, and upload your submission to Canvas.