

Group 1

Organization of the repo

Group 1's repo is very organized. The folders make sense and there don't seem to be any excess irrelevant documents floating around the repo. However, the "images" folder is unnecessary since it doesn't contain any actual images - it just includes a readme that says that the folder contains images.

Documentation in the readmes

The readmes are very clear, included in every folder, and don't contain unnecessary information. However, the content of the readme for the output folder still has the generic content saying that it contains DTA and DO files to demonstrate how to export data. However, it contains data visualization images from the group's analysis and the readme should explain the content of the images.

Code readability

Pretty much every cell contains hashtags explaining the code. All the steps were easy to understand and there were also some markdown explanations of the hypothesis testing and even a conclusion write up at the end. The binder works!

Output

There are two data visualizations included in the output file.

Improvement

Perhaps include a significance test about that explores old vs young or conservative vs liberal news consumption and examine the patterns based on just one variable - instead of many variables like conservative leaning, order, and highly engaged with political news so you can go beyond visual data analysis.

Group 2

Organization of the repo

The folders are logically and nicely organized, however there are still checkpoints in the repo. The images notebook is still full of files copied from the original structure such as the binder image that were originally put there for reference. The final copy shouldn't still include them.

Documentation of the readmes

Mostly, the documentation is fine. The documentation for the notebooks folder describes the ipynb files well. However, some of the readmes still contain the generic reference information and were not updated to reflect the content of the folders.

Code readability

I really like how the analysis code contains a table of contents and an extremely clear markdown explanation of each step. The code is very well organized and easily understandable - the markdown explanations draw conclusions along the way which is also very nice. However, it can't be opened in binder.

Output files

The output files contain many useful data visualizations, but the readme doesn't explain what's what and because there are so many, the images are not as easy to use and process through.

Improvements

The analysis is great I'd just make the repo more polished.