

CS 5010 Final Project

Authors

- André Zazzera (alz9cb)
- Annie Williams (maw3as)
- Hannah Frederick (hbf3k)
- Sean Grace (smg2mx)

plots/

This folder contains each graph saved as a pdf file. This folder is for personal reference and bookkeeping.

scrapers/

This folder contains the following three files, each containing a different webscraper. This folder also contains the corresponding csv output for each web scraper. These csv files are read in by data_cleaning.ipynb, and later condensed into complete_cleaned.csv.

Each python file can be run individually to perform a fresh scrape and produce corresponding csv outputs.

- voter_registration.py
- voter_registration.csv
- governor_results.py
- governor_results.csv
- presidential_results.py
- presidential_results.csv

Team NULL CS 5010 Final Project Report.ipynb

This jupyter notebook contains our final report, combining all of our graphs and analysis. This notebook was used to produce our final report pdf.

TestWebScraper.py

This file contains nine unit tests. Running this python file is expected to show output from nine passing tests. This file is imported and run in the Final Report notebook, where the test results are also displayed.

complete_cleaned.csv

Output file from data_cleaning.ipynb. This csv file was used for our final analysis, and is the result of cleaning and combining the csv files from the scrapers folder.

data_cleaning.ipynb

Data cleaning and processing notebook that takes the csv results from the scrapers folder, and produces one final dataset to complete_cleaned.csv

data_graphing.ipynb

Jupyter notebook that creates all of the visuals included in the plots folder.

data_querying.ipynb

Jupyter notebook that contains all of the queries in one consolidated place. All of the queries that appear in this file also appear in the Final Report notebook.