README Challenge 5

Table of Contents

[**1. Description 1**](#_dgtkdb6n8g1r)

[**2. How to use the files 1**](#_66lx6kt2g0f4)

## Description

This repository contains Jupyter notebooks (.ipynb) for Challenge 5, as well as how to use the file and prerequisites for use.

Both notebooks contain Markdown notes within the files themselves, as well as comments, please you those for references as you’re running the files.

## How to use the files

**Challenge 5 (Geospatial)**:

Filename 1: ‘ShapeFile Generation’.ipynb

This file will generate a shapefile which maps US state legislative districts to their representative, along with the representatives party, and information about business sectors in their district.

This can be used within mapping software to display the data or used to understand different what business sectors may be influential in a particular area given the level of employment or payroll in an area.

**PREREQ/REFERENCES:**

1. Get a Census API Key: <https://api.census.gov/data/key_signup.html>
   1. One is provided within the code for now, feel free to use that one
2. Get OpenStates API Key: <https://openstates.org/accounts/login/?next=/accounts/profile/>
   1. One is provided within the code for now, feel free to use that one
   2. Once you create a profile, there is an area to request an API key, it is free to use this API
3. Install non-standard Python packages if you need to:
   1. geopandas, spacy
   2. To do this run pip install <package\_name> from your terminal
4. To get more information about NAICS and business sectors, visit: <https://www.census.gov/naics/?58967?yearbck=2022>

Filename 2: ‘Business Addresses’.ipynb

This file will generate addresses for businesses using the SEC Edgar API. There are no prerequisites, it only uses standard Python packages.

**REFERENCES**

1. **SEC EDGAR** Datastore:<https://www.sec.gov/edgar/search-and-access>
   1. Contains general information about SEC Filings
2. **CIK Information**: <https://www.sec.gov/edgar/searchedgar/cik>
   1. Explains CIK and how it is used as a unique identifier within the SEC ecosystem
3. **Wikipedia Source:** <https://en.wikipedia.org/wiki/List_of_S%26P_500_companies>