**Census Data Automation**

1. U.S. Census Data can be downloaded from <https://data2.nhgis.org/main>
2. Once downloaded, the data is downloaded in the following format of the data in a *csv* file (Figure 1) and a code book in a *txt* file (Figure 2).

Figure 1: nhgis0004\_ds244\_20195\_tract\_E.csv

Graphical user interface, application, table, Excel

Description automatically generated

Figure 2 : nhgis0004\_ds244\_20195\_tract\_E.txt

Graphical user interface, text, application, email

Description automatically generated

1. Download the following Repo <https://github.com/eddiechacha/2022-352CensusTeam/tree/master/census_data_automation>
2. Put the csv file (Figure 1) in the ***data*** folder

In the NHGIS codebook “nhgis0004\_ds244\_20195\_tract\_E.txt” for example, create a new txt file and copy paste all the column names into the file as follows with each column detail in a new line.

Example 3 :

Graphical user interface, text

Description automatically generated

1. Save the following *txt* file and put it in the ***data*** folder.
2. In the terminal, run

*python clean\_data.py*

1. Once the program is running, there will be a series of questions.

Graphical user interface, text

Description automatically generated

* Enter Census File Name : ?
* Enter Field Names .txt File Name : ?
* How many tables?
* Table \* title?

Answer the following questions according to the NHGIS codebook as given in Figure 2.

1. Once all the questions are answered, the program will generate a clean dataset with Chicago specific data points in the ***output*** folder with the data in a *csv* file as well as a codebook in a *txt* file. The new dataset contains all the data points related to Chicago with the column names that have been edited. (Geom coordinates as well as community area number area added).
2. To visualize the dataset using KeplerGL, open up *visualize\_data.ipynb* and enter the directory of the dataset in the following line.

Graphical user interface, text, application

Description automatically generated

1. Run the code and the data can be visualized by clicking census.html in the main directory.

A screenshot of a map

Description automatically generatedA screenshot of a computer

Description automatically generated with medium confidence