|  |
| --- |
| **Executive Summary - What Is Needed** |

* Currently, I have a Python program that pulls United States Census data from the API censusdata
  + The issue is US Census API is being retired and the data is no longer being refreshed
* I need to change my program to pull data from the Census API of <https://api.census.gov/data>
* All programs are run in Python and in Jupyter Notebooks specifically

|  |
| --- |
| **Current Python Program** |

Below is my current program - - it pulls data at the block group level from the American Community Survey – 5 yr edition. For ease of reading, comments are in bold yellow, *program code is in italics*:

**# Install required packages**

*!pip install censusdata*

*!pip install numpy*

*import numpy as np*

*import pandas as pd*

**# get the 2019 data**

*cars\_data = censusdata.download('acs5', 2019,*

*censusdata.censusgeo([('state', '30'), ('county', '111'), ('block group', '\*')]), \*

*['B25077\_001E', 'C24050\_001E', 'C24050\_011E', 'C24050\_014E', 'B25013\_009E', \*

*'B25013\_008E', 'B25013\_007E', 'B25002\_001E', 'B25002\_002E', 'B01003\_001E', \*

*'B25031\_001E', 'B25103\_001E', 'B25118\_014E', 'B25118\_015E', 'B25118\_016E', \*

*'B25118\_017E', 'B25118\_018E', 'B25118\_019E', 'B25118\_020E', 'B25118\_021E', \*

*'B25118\_022E', 'B25118\_023E', 'B25118\_024E', 'B25118\_025E', 'B25118\_001E', \*

*'B25026\_010E', 'B25026\_011E', 'B25026\_012E', 'B25026\_013E', 'B25026\_014E', \*

*'B25026\_015E', 'B25046\_003E', 'B08013\_001E', 'B25003I\_003E', 'B25003A\_003E', \*

*'B25003B\_003E', 'B25003C\_003E', 'B25003D\_003E', 'B25003E\_003E', \*

*'B25003F\_003E', 'B25003G\_003E', 'B25003H\_003E', 'B19013\_001E', 'B25119\_003E', \*

*'B25010\_002E', 'B25010\_003E', 'B25009\_011E', 'B25009\_012E', 'B25009\_013E', \*

*'B25009\_014E', 'B25009\_015E', 'B25009\_016E', 'B25009\_017E', \*

*'B25032\_013E', 'B25032\_014E', 'B25032\_015E', 'B25032\_016E', 'B25032\_017E', \*

*'B25032\_018E', 'B19057\_001E', 'B19057\_002E', 'B19057\_003E', 'B25042\_001E', 'B25042\_003E', \*

*'B25042\_004E', 'B25042\_005E', 'B25042\_006E', 'B25042\_007E', 'B25042\_008E', \*

*'B25042\_010E', 'B25042\_011E', 'B25042\_012E', 'B25042\_013E', 'B25042\_014E', 'B25042\_015E'])*

**# show the top few rows of the data using the head function**

*cars\_data.head()*

|  |
| --- |
| **What Specifically Is Needed for this Project:** |

This project will be successful when:

1. I have a Python program that runs in Jupyter notebook that can pull data from <https://api.census.gov/data>
2. I’m able to pull this data at the block group level using the ACS-5 data source

1. The data that needs to be pulled is the same as listed above - - it is the list of approximately 75 different variables from *'B25077\_001E'* to *'B25042\_015E'.* I’m looking to pull all these variables in one run of the program, and not 75 different programs pulling one variable each
2. Please note, you may need to register for a Census API key, and these are available for free at : [Key Signup (census.gov)](https://api.census.gov/data/key_signup.html)
3. Here’s the additional Census developer info - - [link](https://www.census.gov/data/developers/data-sets/acs-5year.html)