State Processing Confluence Map

Notebooks:

* main\_label\_pipe
* main\_label\_geo
* main\_HASHING
* main\_concat
* file\_update

Selenium Notebooks: (23/56)

AK, AL, AR, AZ, CA, CT, DC, DE, GA, IL, KY, MA, MD, ME, NC, NH, NJ, OH, PA, TN, VT, WV, VA

State Aggregation Notebooks: (3/56)

IL, MD, VA

Files:

* ppp\_pipe
* ppp\_geo
* ppp\_hash
* ppp\_hashSort
* ppp\_hpgf

State Files:

* pipe\_states (56/56)
* geocoded (24/56)
* geo\_fips (23/56)
* hash\_states (23/56)
* state\_agg (3/56)

Raw Files:

* public\_150k\_plus\_211121
* public\_150k\_plus\_220102
* TIGERLINE\_Shapefiles (State, County, Block Group, Block)
* NAICS codes

External Resources:

* Google Maps Geocoder
* U.S. Census Geocoder
* TIGERLINE\_Shapefiles
* U.S. SBA Loan Program Data
* MapShaper

Raw File Overview:

* public\_150k\_plus\_211121 – U.S. SBA PPP Loan program data as of November 21, 2021
* public\_150k\_plus\_220102 – U.S. SBA PPP Loan program data as of January 02, 2022
* TIGERLINE\_Shapefiles (State, County, Block Group, Block) –
  + State and County Data – 1 file each.
    - Reduced (2/2) \*
  + Block Group Data (3/56)
    - Reduced (3/56) \*
  + Block Data (3/56)
    - Reduced (3/56) \*
* NAICS – 2-6 Digit Industry Code Pairs

External Resources:

* Google Maps Geocoder – pairing coordinates with addresses of Loan Recipients
* U.S. Census Geocoder – pairing FIPS codes with coordinates of Loan Recipients
* TIGERLINE\_Shapefiles – for pairing State, County, Block Group, and Block Codes with geospatial map bounds.
* U.S. SBA Loan Program Data – main data resource
* Mapshaper – for reducing file size of TIGERLINE\_Shapefiles

Notebook Process: