

Instructions for using census tract rescaling code. See final report for description of running GIS code. The GIS code should not need to be rerun anytime soon since the census tract information it relies on and generates corresponds to the 2020 census tracts. The output file `Combinations_Union.csv` file is all that is needed and is in the `gisscripts` directory in this repository.

Base Requirements: Python 3 or higher. Numpy and NETCDF4 modules. Any additional requirements by these modules (e.g. a particular iteration of Python 3) also need to be observed.

First let's assume a directory structure where the source code is in the current directory. Our mapping directory will be `../map_data`, our output directory will be `../output`, and our input files (netcdf gridded nclimgrid-d data) is in `../ncei_data`. (We are using / assuming Unix. Use \ instead for windows).

- 1.) Put the ncei data in the `ncei_data` directory. For this, setup subdirectories (e.g. `ncei_data/1981`) for each year. Put each monthly `.nc` file in it's proper subdirectory.
- 2.) Setup directories in the output directory (e.g. `../output/1981`), for the output `.txt` files containing the estimates.
- 3.) Edit the `mapproc*` scripts to utilize these directories.
- 4.) Put the `Combinations_Union.csv` file in the `map` directory.
- 5.) Run `mapproc.g2c.py` and `mapproc.c2g.py` scripts. These will generate pickle files for the dictionaries that are used hereafter which will be put into the output directory.
- 6.) Edit `forproc.py` for directories and dates. Remember to ensure the correct directories exist as described at the beginning of this document.

Estimates have now been produced and are in the text files in `../output/yyyy`.

- 7.) Create subdirectory `../output/err` and `../output/err/stat`.
- 8.) Edit `devcomp.py` and run it for the desired dates and field (one field at a time). This produces deviation arrays that will go into subdirectory `err`. There will be one for each day and each field.
- 9.) Edit `errstat.py` and run it for the desired dates and field. Output will go to `err/stat` subdirectory and the filenames contain enough info to know what is in them. These are numpy arrays and should be easily to load with `np.load`.

Last modified 5/22/2021 by C. Travis Ashby