(I might roll this document into Vignettes, but for now it is the easiest starting place for glptools.)

**Purpose**

glptools has two main purposes. First, it houses functions that streamline many common tasks when working with Greater Louisville Project data and graphs. Second, it contains basic reference data, crosswalks, and maps that are clean and ready to merge with data. Hopefully, using the glptools package reduces the amount of time required sourcing functions and reading in reference data. The package is designed to integrate with both the glpdata package (which houses the bulk of common GLP data) and other projects that involve similar tasks.

This document contains an overview of the contents of the glptools package and will be flushed out with information on maintaining and updating glptools.

**Structure**

The **data-raw** folder contains raw reference data and scripts used to clean and process the data.

* **basic\_info.R** contains FIPS codes, MSA codes, and state abbreviations for GLP peer cities.
* **MSA\_FIPS.R** creates a crosswalk from MSA codes to FIPS codes.
* **MSA\_PUMA.R** and **FIPS\_PUMA.R** create crosswalks from MSA and FIPS codes to PUMAs.
* **zip\_codes.R** creates crosswalks from MSA and FIPS codes to zip codes.
* **tract\_crosswalk.R** creates a crosswalks from 2000 to 2010 census tracts.
* **inflation\_and\_COLA.R** creates a data frame containing cost-of-living and inflation data.
* **population.R** creates data frames of populations for peer cities. (Should this be supplanted by population data from glpdata?)
* **maps.R** creates other crosswalks between geographies and creates SpatialPolygonDataFrames for GLP maps.

The **R** folder houses functions and documentation for the reference data.

Reading data

* **readers.R** contains functions to help read folders of files from the ACS, BRFSS, CBP, SAIHIE, CPS unemployment, or generic .csv files

Processing data

* **general.R** contains a large number of functions that take care of tasks not aligned with other scripts. It should probably be broken up into multiple files…and generally re-examined.
* **microdata.R** contains functions to process ACS and CPS microdata, as well as summarize microdata by peer geography and demographic group.
* **wonder.R** containsfunctionsspecifically to process CDC wonder data.

Graphing data

* **ranking.R** contains a function to produce GLP ranking graphs.
* **trendline.R** contains wrapper functions used to produce GLP-style trendline graphs (or output the data behind them.)
* **trendline\_helpers.R** contains a common trendline function that is called by the trendline wrapper functions. The main trendline function calls upon several intermediate functions designed to make the function more digestible.
* **map.R** contains a function to produce leaflet graphs of Louisville data.

Other

* **data.R** contains documentation for reference data.

**Using and updating glptools**

Some of scripts rely on reference data produced in other scripts (such as a list of GLP peer FIPS codes or an MSA to FIPS crosswalk). The file data-raw/generate\_all\_data.R runs each script sequentially so that the global environment is populated with any necessary objects.

After editing a script, run **generate\_all\_data.R** to update all of the reference information.