**This document serves as a guide for the Datasets folder.**

**Datasets Brainstorm**: Legacy document, can be deleted

**cty\_trends.csv**: Original county level dataset of outcome trends

**cty\_covariates.csv**: Original county level dataset of covariates

**ca\_trends.csv**: Processed and extracted California county data of outcome trends for select years

**ca\_covariates.csv**: Processed and extracted California county data of covariates for select years

**ca\_air\_quality.csv**: Processed and extracted California county data of pm 2.5 concentrations for each year, mean

**Codebook\_County\_Trends\_Estimates.pdf**: describes trends data and terminology

**Codebook\_County\_Covariates.pdf**: describes covariates data and terminology

**Data\_Preprocessing.ipynb**: Notebook used to process cty\_trends.csv and cty\_covariates.csv into ca\_trends.csv and ca\_covariates.csv datasets

**Extra\_Preprocessing.ipynb**: Notebook used to process concentrations.csv (found in EDA folder) into ca\_air\_quality.csv. It is provided here for convenience if data needs to be changed.

**Opportunity\_Atlas\_Methods.pdf**: description of data imputation and county level estimates