**Extract - Data**

* winemag-data\_first150k.csv – winemag.com data obtained from Kaggle.com
* acs2015\_census\_tract\_data.csv – Census data obtained from Kaggle.com

**Transformation**

* Columns were renamed to be clear and uniformed for both datasetsfr.
* The winemag dataset was filtered to only include data from the US.
* NaaN values were dropped from the winemag dataset.
* Brought in selected columns for the Census dataset.
* Grouped the Census dataset by State.
* Merged the winemag and census dataset on the State column.
* The merged dataset had non-ASCII characters so these were removed.
* A dataframe was created to only include State, Vineyard, Winery, Cost Per Bottle, Category and TotalPop.
* A dataframe was create to Total Population, number of vineyards, number of vineyards, number of wineries, number of categories, and average cost per bottle by state.

**Load**

* Our data was store in MySql using a database called wine\_db
* There were 3 tables created within the wine\_db