For this project we started by gathering our data from various different sites credible on the web.

* Pennsylvania school data (csv)
  + <http://www.edna.pa.gov/Screens/wfHome.aspx>
* Real estate data (csv)
  + <https://www.kaggle.com/harry007/philly-real-estate-data-set-sample>
* Ice cream shop data (JSON)
  + Yelp API

Once we had our sources, we loaded our CSVs into the notebook and connected to the Yelp API. Before we were able to work with our data we needed to clean up any unnecessary columns from the data set. This would help with readability seeing as useless columns would cause clutter and slow down our program. We utilized the Pandas module in Python, which allowed us to manipulate the data as we saw fit. As we examined the data during the cleaning process, we were able to understand the unique relationships between our chosen data sets. Subsequently we were able to map out the structure of our server by creating an ERD and designing our database and how our many tables would relate back to each other. After cleaning the data set we then performed an inner join on the data on a primary key, in this case we used zipcode. This allowed us to only keep the data that related to each other. Finally after having all the data cleaned the data we used sqlalchemy to import the data onto a postgres server