From Kaggle there is a bunch of data from the city of San Francisco. I took 2 datasets, one with building permit data and the other describing the patrons of the libraries. The goal was to take the well-structured CSV files and create a database in 3NF that allows mixed querying between the dataset. Additionally, there is enough geographic and demographic data in the dataset to make some exploratory judgements about some of the neighborhoods.

WE downloaded the data from Kaggle and imported it into a jupyter notebook for exploration.