**Project Title:** ETL\_School\_Rating&Home\_Price

**Contributors:**

Sreedevi Agasthya

Nirguna Joshi

**Scope:** Providing school information and home prices for prospective home buyers by zip code for school calendar year 2017-2018 in state of Georgia.

**Extract (Data set Source):** we pulled the \*csv data set from below source and read them into Python

<https://www.zillow.com/research/data/>

<https://catalog.data.gov/dataset/public-school-locations-2017-18>

**Transform** (Data Cleaning)**:**

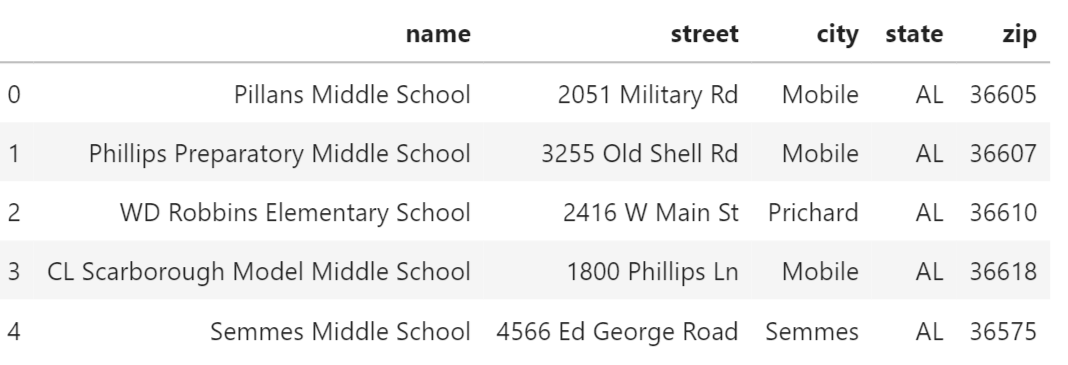
Dropped the columns and duplicates, which are not required from both Schools and Homes data frames

Here is the example before and after:

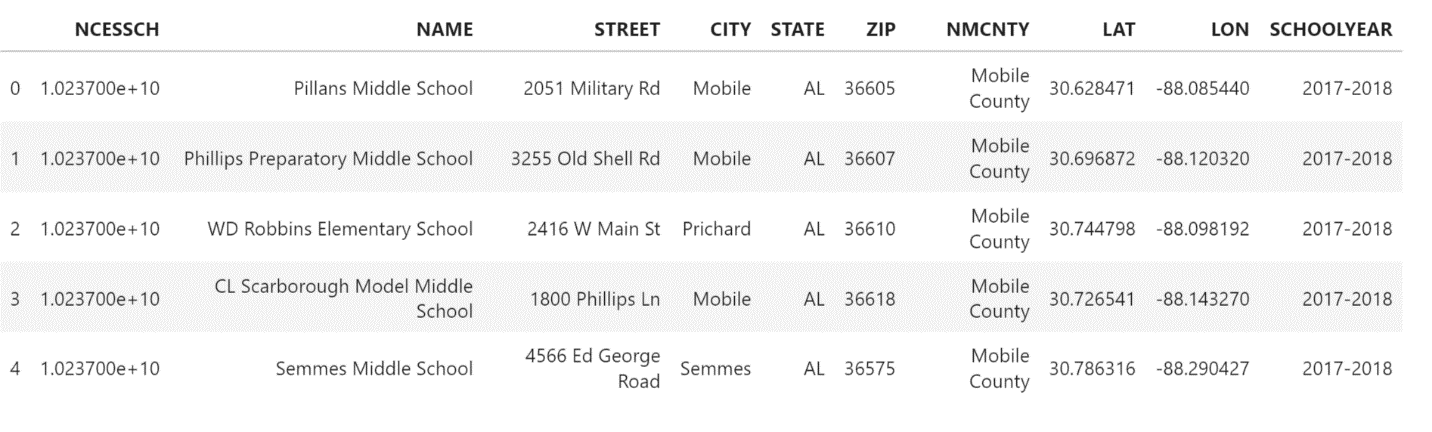
Schools before:



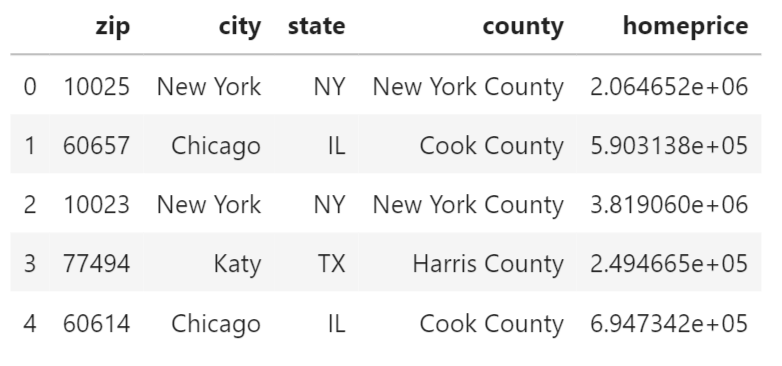
Schools after:



Homes before:



Homes after:



**Load** (Loaded the data to PostgreSQL):

Below is the result set after querying both the tables:

