

Final Proposal: School Safety and Home Prices in New York City

Principal Investigators: Emily Bai and Christian Claus

Emails: eb3144@stern.nyu.edu and cjc715@stern.nyu.edu

This project will look at how historical changes in school safety in various districts in New York City have had a correlation in home prices in their respective areas. New York City has 5 boroughs: Manhattan, Queens, Bronx, Brooklyn and Staten Island. We will be looking at data across various schools from NYC Open Data through [this link \(https://data.cityofnewyork.us/Education/2010-2016-School-Safety-Report/qybk-bjjc/data\)](https://data.cityofnewyork.us/Education/2010-2016-School-Safety-Report/qybk-bjjc/data). Specifically, we will be focusing on the crime rate statistics for the years 2013 to 2016.

We will use data from [NYC Open Data \(https://data.cityofnewyork.us/Housing-Development/Annualized-Rolling-Sales-Update/uzf5-f8n2\)](https://data.cityofnewyork.us/Housing-Development/Annualized-Rolling-Sales-Update/uzf5-f8n2) and [Kaggle \(https://www.kaggle.com/new-york-city/nyc-property-sales\)](https://www.kaggle.com/new-york-city/nyc-property-sales) to find information about housing prices in each of the five boroughs in the same time period as the school safety data. The ultimate goal of the project will be to visualize the correlation between school safety statistics and housing prices through a period of 3 years.

While property value is affected by many variables, we want to focus on whether school safety has any correlation with the pricing. This does not imply causation but if there is a correlation, we would like to come up with a few possible explanations for it.

We anticipate that the project will have four main sections:

- Basic statistics about school safety. We will be reading in the excel document with information about schools in all five districts.
- Next, we will aggregate data for house prices in each of the 5 boroughs by reading in the excel document.
- The last step is to merge the two dataframes by matching the zip code and grouping them by boroughs using the `isin` function. - The results will be visualized through several graphs exploring the correlation between crime and housing prices. We want to know 3 things: **1)** What is the overall correlation between housing prices and school safety in NYC? Is there any relationship at all?, **2)** Have there been changes through the three observable years in terms of pricing and crime rate?, and **3)** Do certain boroughs have a higher correlation? Are certain boroughs more or less dynamic in terms of housing prices and do certain boroughs have higher school crime rates, which may affect the pricing?
- Finally, we will try to include possible explanations of what we observed in our visualizations.

Data Report

Overview: Most of the data we are using in this project comes from [NYC Open Link \(https://opendata.cityofnewyork.us/\)](https://opendata.cityofnewyork.us/). From there, we were able to find data on school safety statistics and housing prices for New York City over a 3 year time period from 2013-2016.

Requisite Packages:

```
In [729]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

import geopandas as gpd
from shapely.geometry import Point, Polygon

import os
```

Grabbing Data:

Part 1: School Safety Data

First, we need to read in school safety data:

```
In [730]: crime = pd.read_excel("C:\\me\\data bootcamp\\final project\\2010_-_2016_School_Safety_Report.xlsx")
crime
```

Out[730]:

	School Year	Building Code	DBN	Location Name	Location Code	Address	Borough	Geographic District Co
0	2015-16	K006	17K006	P.S. 006	K006	43 SNYDER AVENUE	K	17
1	2015-16	K005	16K005	P.S. 005 Dr. Ronald Mcnair	K005	820 HANCOCK STREET	K	16
2	2015-16	K003	13K003	P.S. 003 The Bedford Village	K003	50 JEFFERSON AVENUE	K	13
3	2015-16	K002	NaN	655 PARKSIDE AVENUE CONSOLIDATED LOCATION	NaN	655 PARKSIDE AVENUE	K	17
4	2015-16	K002	75K141	P.S. K141	K141	655 PARKSIDE AVENUE	K	17
5	2015-16	K002	84K704	EXPLORE CHARTER SCHOOL(BS)	K704	655 PARKSIDE AVENUE	K	17
6	2015-16	K002	17K002	Parkside Preparatory Academy	K002	655 PARKSIDE AVENUE	K	17
7	2015-16	X953	75XHOS	Hospital Schools - Bronx	X402	3450 EAST TREMONT AVENUE	X	8
8	2015-16	K001	15K001	P.S. 001 The Bergen	K001	309 47 STREET	K	15
9	2015-16	K008	13K008	P.S. 008 Robert Fulton	K008	37 HICKS STREET	K	13
10	2015-16	K095	21K095	P.S. 095 The Gravesend	K095	345 VAN SICKLEN STREET	K	21
11	2015-16	K165	23K165	P.S. 165 Ida Posner	K165	76 LOTT AVENUE	K	23
12	2015-16	Q849	24Q077	I.S. 077	Q077	976 SENECA AVENUE	Q	24
13	2015-16	K857	75K036	P.S. 36	K036	2045 LINDEN BOULEVARD	K	15
14	2015-16	X953	75M401	Hospital Schools	M401	3450 EAST TREMONT AVENUE	X	8

	School Year	Building Code	DBN	Location Name	Location Code	Address	Borough	Geographic District Co
15	2015-16	X953	75KHOS	Hospital Schools - Brooklyn	K403	3450 EAST TREMONT AVENUE	X	8
16	2015-16	X905	10X696	High School of American Studies at Lehman College	X696	2925 GOULDEN AVENUE	X	10
17	2015-16	X886	10X307	Luisa Pineiro Fuentes School of Science and Di...	X307	124 EAMES PLACE	X	10
18	2015-16	X884	NaN	350 GERARD AVENUE CONSOLIDATED LOCATION	NaN	350 GERARD AVENUE	X	7
19	2015-16	X884	07X670	Health Opportunities High School	X670	350 GERARD AVENUE	X	7
20	2015-16	X878	12X682	Fannie Lou Hamer Freedom High School	X682	1021 JENNINGS STREET	X	12
21	2015-16	X876	09X525	Bronx Leadership Academy High School	X525	1710 WEBSTER AVENUE	X	9
22	2015-16	X859	10X315	P.S. 315 Lab School	X315	2246 JEROME AVENUE	X	10
23	2015-16	X852	10X308	Bronx Dance Academy School	X308	3617 BAINBRIDGE AVENUE	X	10
24	2015-16	X843	10X246	P.S. 246 Poe Center	X246	2641 GRAND CONCOURSE	X	10
25	2015-16	X839	11X418	Bronx High School for the Visual Arts	X418	2040 ANTIN PL	X	11
26	2015-16	X834	75X352	The Vida Bogart School for All Children	X352	1330 BRISTOW STREET	X	12
27	2015-16	X826	10X226	P.S. 226	X226	1950 SEDGWICK AVENUE	X	10
28	2015-16	X819	10X207	P.S. 207	X207	3030 GODWIN TERRACE	X	10

	School Year	Building Code	DBN	Location Name	Location Code	Address	Borough	Geographic District Co
29	2015-16	X790	NaN	730 CONCOURSE VILLAGE WEST CONSOLIDATED LOCATION	NaN	730 CONCOURSE VILLAGE WEST	X	7
...
6280	2013-14	K068	18K068	I.S. 068 Isaac Bildersee	K068	956 EAST 82 STREET	K	18
6281	2013-14	K067	NaN	51 SAINT EDWARDS STREET CONSOLIDATED LOCATION	NaN	51 SAINT EDWARDS STREET	K	13
6282	2013-14	K067	84K536	Community Roots Charter School	K536	51 SAINT EDWARDS STREET	K	13
6283	2013-14	K067	13K067	P.S. 067 Charles A. Dorsey	K067	51 SAINT EDWARDS STREET	K	13
6284	2013-14	K066	18K066	P.S. 66	K066	845 EAST 96 STREET	K	18
6285	2013-14	K062	20K062	J.H.S. 062 Ditmas	K062	700 CORTELYOU ROAD	K	20
6286	2013-14	K061	17K061	M.S. 061 Dr. Gladstone H. Atwell	K061	400 EMPIRE BOULEVARD	K	17
6287	2013-14	K059	NaN	211 THROOP AVENUE CONSOLIDATED LOCATION	NaN	211 THROOP AVENUE	K	14
6288	2013-14	K096	21K096	I.S. 096 Seth Low	K096	99 AVENUE P	K	21
6289	2013-14	K059	14K059	P.S. 059 William Floyd	K059	211 THROOP AVENUE	K	14
6290	2013-14	K059	84K125	Success Academy Charter School - Bed- Stuy 2	K125	211 THROOP AVENUE	K	14
6291	2013-14	K058	15K058	P.S. 058 The Carroll	K058	330 SMITH STREET	K	15

	School Year	Building Code	DBN	Location Name	Location Code	Address	Borough	Geographic District Co
6292	2013-14	K057	NaN	125 STUYVESANT AVENUE CONSOLIDATED LOCATION	NaN	125 STUYVESANT AVENUE	K	16
6293	2013-14	K057	16K057	J.H.S. 057 Whitelaw Reid	K057	125 STUYVESANT AVENUE	K	16
6294	2013-14	K104	20K104	P.S./I.S. 104 The Fort Hamilton School	K104	9115 5 AVENUE	K	20
6295	2013-14	K057	16K385	School of Business, Finance and Entrepreneurship	K385	125 STUYVESANT AVENUE	K	16
6296	2013-14	K071	14K071	Juan Morel Campos Secondary School	K071	215 HEYWARD STREET	K	14
6297	2013-14	K057	16K688	The Brooklyn Academy of Global Finance	K688	125 STUYVESANT AVENUE	K	16
6298	2013-14	K056	NaN	170 GATES AVENUE CONSOLIDATED LOCATION	NaN	170 GATES AVENUE	K	13
6299	2013-14	K056	13K056	P.S. 056 Lewis H. Latimer	K056	170 GATES AVENUE	K	13
6300	2013-14	X440	10X353	World View High School	X353	100 WEST MOSHOLU PARKWAY SOUTH	X	10
6301	2013-14	K056	13K103	Satellite Three	K103	170 GATES AVENUE	K	13
6302	2013-14	K056	13K351	The Urban Assembly Unison School	K351	170 GATES AVENUE	K	13
6303	2013-14	K055	NaN	2021 BERGEN STREET CONSOLIDATED LOCATION	NaN	2021 BERGEN STREET	K	23

	School Year	Building Code	DBN	Location Name	Location Code	Address	Borough	Geographic District Co
6304	2013-14	K055	84K626	Achievement First Brownsville Charter School	K626	2021 BERGEN STREET	K	23
6305	2013-14	K055	23K493	Brooklyn Collegiate: A College Board School	K493	2021 BERGEN STREET	K	23
6306	2013-14	K054	13K054	P.S. 054 Samuel C. Barnes	K054	195 SANFORD STREET	K	13
6307	2013-14	K052	22K052	P.S. 052 Sheepshead Bay	K052	2675 EAST 29 STREET	K	22
6308	2013-14	K051	15K051	M.S. 51 William Alexander	K051	350 5 AVENUE	K	15
6309	2013-14	K050	NaN	183 SOUTH 3 STREET CONSOLIDATED LOCATION	NaN	183 SOUTH 3 STREET	K	14

6310 rows × 34 columns

Since there is a lot of information, we have to clean up the data to show only the important information we will be analyzing later. The column "Total Number of Crimes" will include all types of crimes documented in the data file, including


```
In [731]: #crime.set_index("School Year", inplace = True)
crime.rename(columns = {"Borough Name": "Borough_Name", "Postcode": "ZIP_CODE",
"School Year": "Year"}, inplace = True)
crime["Total Number of Crimes"] = crime["Major N"] + crime["Oth N"] + crime["N
oCrim N"] + crime["Prop N"] + crime["Vio N"]
crime["Average Num of Crimes"] = crime["AvgOfMajor N"] + crime["AvgOfOth N"] +
crime["AvgOfNoCrim N"] + crime["AvgOfProp N"] + crime["AvgOfVio N"]

crime = crime[['Year', 'Building Code', 'Location Code', 'Address', 'Borough_Nam
e', 'ZIP_CODE', 'Total Number of Crimes', 'Average Num of Crimes']]
crime.dropna(inplace = True)
crime
```

```
C:\me\Anaconda\lib\site-packages\ipykernel_launcher.py:7: SettingWithCopyWarning:  
ing:
```

A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
import sys
```

Out[731]:

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
0	2015-16	K006	K006	43 SNYDER AVENUE	BROOKLYN	11226.0	4.0	4.41
1	2015-16	K005	K005	820 HANCOCK STREET	BROOKLYN	11233.0	2.0	3.49
2	2015-16	K003	K003	50 JEFFERSON AVENUE	BROOKLYN	11216.0	2.0	3.91
8	2015-16	K001	K001	309 47 STREET	BROOKLYN	11220.0	2.0	12.69
9	2015-16	K008	K008	37 HICKS STREET	BROOKLYN	11201.0	6.0	6.27
10	2015-16	K095	K095	345 VAN SICKLEN STREET	BROOKLYN	11223.0	0.0	6.27
11	2015-16	K165	K165	76 LOTT AVENUE	BROOKLYN	11212.0	13.0	3.91
12	2015-16	Q849	Q077	976 SENECA AVENUE	QUEENS	11385.0	5.0	8.77
13	2015-16	K857	K036	2045 LINDEN BOULEVARD	BROOKLYN	11207.0	1.0	3.91
16	2015-16	X905	X696	2925 GOULDEN AVENUE	BRONX	10468.0	0.0	3.91
17	2015-16	X886	X307	124 EAMES PLACE	BRONX	10468.0	0.0	3.91
20	2015-16	X878	X682	1021 JENNINGS STREET	BRONX	10460.0	12.0	3.91
21	2015-16	X876	X525	1710 WEBSTER AVENUE	BRONX	10457.0	20.0	4.41
22	2015-16	X859	X315	2246 JEROME AVENUE	BRONX	10453.0	0.0	3.91
23	2015-16	X852	X308	3617 BAINBRIDGE AVENUE	BRONX	10467.0	1.0	3.91
24	2015-16	X843	X246	2641 GRAND CONCOURSE	BRONX	10468.0	0.0	4.41
25	2015-16	X839	X418	2040 ANTIN PL	BRONX	10462.0	7.0	3.91
26	2015-16	X834	X352	1330 BRISTOW STREET	BRONX	10459.0	0.0	4.41
27	2015-16	X826	X226	1950 SEDGWICK AVENUE	BRONX	10453.0	1.0	4.41
28	2015-16	X819	X207	3030 GODWIN TERRACE	BRONX	10463.0	0.0	3.91
32	2015-16	X991	X695	560 BROOK AVENUE	BRONX	10455.0	0.0	3.91

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
47	2015-16	X722	X721	2697 WESTCHESTER AVENUE	BRONX	10461.0	2.0	4.41
67	2015-16	X701	X723	3540 BIVONA STREET	BRONX	10475.0	0.0	3.91
77	2015-16	X879	X684	1122 EAST 180 STREET	BRONX	10460.0	26.0	3.91
78	2015-16	X445	X445	75 WEST 205 STREET	BRONX	10468.0	10.0	31.80
150	2015-16	X392	X269	928 SIMPSON STREET	BRONX	10459.0	31.0	4.41
151	2015-16	X368	X368	2975 TIBBETT AVENUE	BRONX	10463.0	8.0	6.27
156	2015-16	X360	X360	2880 KINGSBRIDGE TERRACE	BRONX	10463.0	1.0	3.91
157	2015-16	X340	X340	25 WEST 195 STREET	BRONX	10468.0	0.0	4.41
162	2015-16	X292	X357	800 LYDIG AVENUE	BRONX	10462.0	0.0	3.49
...
6242	2013-14	K112	K112	7115 15 AVENUE	BROOKLYN	11228.0	1.0	4.83
6246	2013-14	K110	K110	124 MONITOR STREET	BROOKLYN	11222.0	2.0	3.73
6247	2013-14	K106	K106	1328 PUTNAM AVENUE	BROOKLYN	11221.0	4.0	4.83
6248	2013-14	K105	K105	1031 59 STREET	BROOKLYN	11219.0	2.0	20.29
6249	2013-14	K101	K101	2360 BENSON AVENUE	BROOKLYN	11214.0	1.0	6.63
6250	2013-14	K100	K100	2951 WEST 3 STREET	BROOKLYN	11224.0	3.0	4.83
6251	2013-14	K099	K099	1120 EAST 10 STREET	BROOKLYN	11230.0	1.0	6.63
6252	2013-14	K095	K095	345 VAN SICKLEN STREET	BROOKLYN	11223.0	3.0	6.63
6253	2013-14	K094	K094	5010 6 AVENUE	BROOKLYN	11220.0	2.0	15.35
6254	2013-14	K093	K093	31 NEW YORK AVENUE	BROOKLYN	11216.0	2.0	3.73
6258	2013-14	K091	K091	532 ALBANY AVENUE	BROOKLYN	11203.0	8.0	3.73
6259	2013-14	K090	K090	2840 WEST 12 STREET	BROOKLYN	11224.0	6.0	4.83

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
6263	2013-14	K086	K086	220 IRVING AVENUE	BROOKLYN	11237.0	0.0	3.73
6264	2013-14	K085	K753	510 CLERMONT AVENUE	BROOKLYN	11238.0	16.0	4.09
6265	2013-14	K084	K084	250 BERRY STREET	BROOKLYN	11249.0	1.0	4.83
6266	2013-14	K081	K081	990 DEKALB AVENUE	BROOKLYN	11221.0	9.0	3.73
6267	2013-14	K078	K078	1420 EAST 68 STREET	BROOKLYN	11234.0	15.0	6.63
6268	2013-14	K075	K075	95 GROVE STREET	BROOKLYN	11221.0	1.0	4.83
6271	2013-14	K097	K097	1855 STILLWELL AVENUE	BROOKLYN	11223.0	3.0	6.63
6279	2013-14	K069	K069	6302 9TH AVENUE	BROOKLYN	11220.0	0.0	6.63
6280	2013-14	K068	K068	956 EAST 82 STREET	BROOKLYN	11236.0	29.0	4.83
6284	2013-14	K066	K066	845 EAST 96 STREET	BROOKLYN	11236.0	0.0	6.63
6285	2013-14	K062	K062	700 CORTELYOU ROAD	BROOKLYN	11218.0	4.0	13.13
6286	2013-14	K061	K061	400 EMPIRE BOULEVARD	BROOKLYN	11225.0	23.0	6.63
6288	2013-14	K096	K096	99 AVENUE P	BROOKLYN	11204.0	13.0	4.83
6291	2013-14	K058	K058	330 SMITH STREET	BROOKLYN	11231.0	0.0	6.63
6294	2013-14	K104	K104	9115 5 AVENUE	BROOKLYN	11209.0	4.0	8.93
6306	2013-14	K054	K054	195 SANFORD STREET	BROOKLYN	11205.0	2.0	3.73
6307	2013-14	K052	K052	2675 EAST 29 STREET	BROOKLYN	11235.0	2.0	6.63
6308	2013-14	K051	K051	350 5 AVENUE	BROOKLYN	11215.0	0.0	8.93

2533 rows × 8 columns

Now, we will be creating separate dataframes for each borough. For each of these dataframes, we want to grab only the rows pertaining to the respective borough.

```
In [732]: bronx = crime[crime['Borough_Name'] == "BRONX"]
```

```
In [733]: bronx.loc[:, 'Borough_Name'] = 'BRONX'  
bronx['Borough_Name'][16]
```

```
Out[733]: 'BRONX'
```

In [734]: `bronx`

Out[734]:

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
16	2015-16	X905	X696	2925 GOULDEN AVENUE	BRONX	10468.0	0.0	3.91
17	2015-16	X886	X307	124 EAMES PLACE	BRONX	10468.0	0.0	3.91
20	2015-16	X878	X682	1021 JENNINGS STREET	BRONX	10460.0	12.0	3.91
21	2015-16	X876	X525	1710 WEBSTER AVENUE	BRONX	10457.0	20.0	4.41
22	2015-16	X859	X315	2246 JEROME AVENUE	BRONX	10453.0	0.0	3.91
23	2015-16	X852	X308	3617 BAINBRIDGE AVENUE	BRONX	10467.0	1.0	3.91
24	2015-16	X843	X246	2641 GRAND CONCOURSE	BRONX	10468.0	0.0	4.41
25	2015-16	X839	X418	2040 ANTIN PL	BRONX	10462.0	7.0	3.91
26	2015-16	X834	X352	1330 BRISTOW STREET	BRONX	10459.0	0.0	4.41
27	2015-16	X826	X226	1950 SEDGWICK AVENUE	BRONX	10453.0	1.0	4.41
28	2015-16	X819	X207	3030 GODWIN TERRACE	BRONX	10463.0	0.0	3.91
32	2015-16	X991	X695	560 BROOK AVENUE	BRONX	10455.0	0.0	3.91
47	2015-16	X722	X721	2697 WESTCHESTER AVENUE	BRONX	10461.0	2.0	4.41
67	2015-16	X701	X723	3540 BIVONA STREET	BRONX	10475.0	0.0	3.91
77	2015-16	X879	X684	1122 EAST 180 STREET	BRONX	10460.0	26.0	3.91
78	2015-16	X445	X445	75 WEST 205 STREET	BRONX	10468.0	10.0	31.80
150	2015-16	X392	X269	928 SIMPSON STREET	BRONX	10459.0	31.0	4.41
151	2015-16	X368	X368	2975 TIBBETT AVENUE	BRONX	10463.0	8.0	6.27
156	2015-16	X360	X360	2880 KINGSBRIDGE TERRACE	BRONX	10463.0	1.0	3.91
157	2015-16	X340	X340	25 WEST 195 STREET	BRONX	10468.0	0.0	4.41

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
162	2015-16	X292	X357	800 LYDIG AVENUE	BRONX	10462.0	0.0	3.49
163	2015-16	X285	X361	200 West 167th Street	BRONX	10452.0	0.0	3.91
164	2015-16	X283	X228	400 EAST FORDHAM ROAD	BRONX	10458.0	0.0	3.91
165	2015-16	X279	X279	2100 WALTON AVENUE	BRONX	10453.0	8.0	8.77
166	2015-16	X254	X254	2452 WASHINGTON AVENUE	BRONX	10458.0	5.0	3.91
170	2015-16	X234	X300	2050 PROSPECT AVENUE	BRONX	10457.0	1.0	4.41
173	2015-16	X465	X231	4143 THIRD AVENUE	BRONX	10457.0	12.0	4.41
176	2015-16	X211	X319	560 EAST 179 STREET	BRONX	10457.0	2.0	3.49
177	2015-16	X209	X209	313 EAST 183RD STREET	BRONX	10458.0	1.0	3.91
178	2015-16	X206	X206	2280 AQUEDUCT AVENUE	BRONX	10468.0	3.0	3.49
...
4737	2013-14	X037	X037	360 WEST 230 STREET	BRONX	10463.0	8.0	4.83
4738	2013-14	X036	X036	1070 CASTLE HILL AVENUE	BRONX	10472.0	0.0	4.83
4739	2013-14	X035	X035	261 EAST 163 STREET	BRONX	10451.0	2.0	4.83
4740	2013-14	X034	X188	770 GROTE STREET	BRONX	10460.0	0.0	3.73
4741	2013-14	X033	X033	2424 JEROME AVENUE	BRONX	10468.0	3.0	8.93
4742	2013-14	X032	X032	690 EAST 183 STREET	BRONX	10458.0	0.0	6.63
4743	2013-14	X030	X030	510 EAST 141 STREET	BRONX	10454.0	2.0	4.83
4744	2013-14	X029	X029	758 COURTLANDT AVENUE	BRONX	10451.0	2.0	4.83
4745	2013-14	X028	X028	1861 ANTHONY AVENUE	BRONX	10457.0	0.0	4.83
4746	2013-14	X027	X277	519 ST ANNS AVENUE	BRONX	10455.0	9.0	3.73

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
4750	2013-14	X025	X025	811 EAST 149 STREET	BRONX	10455.0	0.0	3.73
4751	2013-14	X024	X024	660 WEST 236 STREET	BRONX	10463.0	2.0	6.63
4752	2013-14	X023	X023	2151 WASHINGTON AVENUE	BRONX	10457.0	0.0	4.83
4756	2013-14	X021	X021	715 EAST 225 STREET	BRONX	10466.0	0.0	4.83
4757	2013-14	X020	X020	3050 WEBSTER AVENUE	BRONX	10467.0	2.0	8.93
4758	2013-14	X019	X019	4318 KATONAH AVENUE	BRONX	10470.0	1.0	4.83
4759	2013-14	X018	X018	502 MORRIS AVENUE	BRONX	10451.0	1.0	4.83
4760	2013-14	X016	X016	4550 CARPENTER AVENUE	BRONX	10470.0	3.0	4.83
4764	2013-14	X014	X014	3041 BRUCKNER BOULEVARD	BRONX	10461.0	3.0	4.83
4765	2013-14	X012	X012	2555 TRATMAN AVENUE	BRONX	10461.0	30.0	4.09
4766	2013-14	X011	X011	1257 OGDEN AVENUE	BRONX	10452.0	1.0	6.63
4767	2013-14	X009	X009	230 EAST 183 STREET	BRONX	10458.0	0.0	6.63
4768	2013-14	X008	X008	3010 BRIGGS AVENUE	BRONX	10458.0	6.0	8.93
4769	2013-14	X007	X007	3201 KINGSBRIDGE AVENUE	BRONX	10463.0	1.0	4.83
4770	2013-14	X006	X006	1000 EAST TREMONT AVENUE	BRONX	10460.0	4.0	4.83
4771	2013-14	X005	X005	564 JACKSON AVENUE	BRONX	10455.0	1.0	4.83
4776	2013-14	X003	X003	2100 LAFONTAINE AVENUE	BRONX	10457.0	5.0	3.73
4780	2013-14	X112	X112	1925 SCHIEFFELIN AVENUE	BRONX	10466.0	1.0	3.73
4781	2013-14	X001	X001	335 EAST 152 STREET	BRONX	10451.0	2.0	4.83
6015	2013-14	X107	X107	1695 SEWARD AVENUE	BRONX	10473.0	3.0	4.83

427 rows × 8 columns

We are using a groupby function to organize the data. The "Total Number of Crimes" column is the total number of crimes (these crimes include major crimes, other crimes, non-criminal crimes, property crimes and violent crimes). The "Average Num of Crimes" column represents the average number of crimes, based on historical evidence, that typically occur in the respective area in a given year. We also created a column for standard deviation to compare the number of crimes in either 2013-14, 2014-15 or 2015-16, compared to the average number of crimes typically committed in a year.

```
In [735]: bronxdf = bronx.groupby(["Borough_Name", "Year", "Address"])
bronx_dict = {"ZIP_CODE":"first", "Total Number of Crimes":"first", "Average Num of Crimes":"first"}
bronx_df = bronxdf.agg(bronx_dict)

bronx_df["Standard_deviation"] = bronx_df["Total Number of Crimes"] - bronx_df["Average Num of Crimes"]
# this is the standard deviation of the total number of crimes at the school compared to the average number of crimes in their respective locations
bronx_df
```

Out[735]:

			ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
Borough_Name	Year	Address				
BRONX	2013-14	1000 EAST TREMONT AVENUE	10460.0	4.0	4.83	-0.83
		1007 EVERGREEN AVENUE	10472.0	3.0	6.63	-3.63
		1020 ANDERSON AVENUE	10452.0	2.0	6.63	-4.63
		1021 JENNINGS STREET	10460.0	5.0	3.73	1.27
		1070 CASTLE HILL AVENUE	10472.0	0.0	4.83	-4.83
		1075 PUGSLEY AVENUE	10472.0	5.0	8.93	-3.93
		1122 EAST 180 STREET	10460.0	38.0	4.83	33.17
		1155 CROMWELL AVENUE	10452.0	6.0	6.63	-0.63
		1166 NEILL AVENUE	10461.0	1.0	4.83	-3.83
		1220 GERARD AVENUE	10452.0	0.0	6.63	-6.63
		124 EAMES PLACE	10468.0	0.0	3.73	-3.73
		1245 WASHINGTON AVENUE	10456.0	0.0	3.73	-3.73
		1257 OGDEN AVENUE	10452.0	1.0	6.63	-5.63
		1260 FRANKLIN AVENUE	10456.0	5.0	4.83	0.17
		1290 SPOFFORD AVENUE	10474.0	6.0	6.63	-0.63
		1330 BRISTOW STREET	10459.0	0.0	4.83	-4.83
		1340 SHERIDAN AVENUE	10456.0	0.0	4.09	-4.09
		1375 MACE AVENUE	10469.0	1.0	4.83	-3.83
		1400 NEEDHAM AVENUE	10469.0	6.0	6.63	-0.63
		1434 LONGFELLOW AVENUE	10459.0	4.0	4.83	-0.83

Borough_Name	Year	Address	ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
		1449 SHAKESPEARE AVENUE	10452.0	2.0	6.63	-4.63
		1537 WASHINGTON AVENUE	10457.0	1.0	3.73	-2.73
		1598 TOWNSEND AVENUE	10452.0	0.0	3.73	-3.73
		165 BROWN PLACE	10454.0	5.0	4.83	0.17
		1691 WEEKS AVENUE	10457.0	2.0	13.13	-11.13
		1695 SEWARD AVENUE	10473.0	3.0	4.83	-1.83
		1710 WEBSTER AVENUE	10457.0	25.0	4.83	20.17
		1716 SOUTHERN BOULEVARD	10460.0	20.0	4.83	15.17
		175 WEST 166 STREET	10452.0	3.0	4.83	-1.83
		1771 POPHAM AVENUE	10453.0	5.0	6.63	-1.63

	2015- 16	650 WARING AVENUE	10467.0	1.0	6.27	-5.27
		660 FOX STREET	10455.0	2.0	6.27	-4.27
		660 WEST 236 STREET	10463.0	0.0	8.77	-8.77
		660 WEST 237 STREET	10463.0	14.0	12.69	1.31
		677 EAST 141 STREET	10454.0	4.0	3.91	0.09
		690 EAST 183 STREET	10458.0	1.0	6.27	-5.27
		695 EAST 182ND STREET	10457.0	0.0	3.49	-3.49
		700 BAYCHESTER AVENUE	10475.0	3.0	6.27	-3.27
		700 EAST 179 STREET	10457.0	2.0	3.91	-1.91
		715 EAST 225 STREET	10466.0	1.0	4.41	-3.41
		725 BRADY AVENUE	10462.0	2.0	12.69	-10.69

Borough_Name	Year	Address	ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
		75 WEST 205 STREET	10468.0	10.0	31.80	-21.80
		750 JENNINGS STREET	10459.0	12.0	4.41	7.59
		750 PROSPECT AVENUE	10455.0	0.0	3.91	-3.91
		757 CAULDWELL AVENUE	10456.0	0.0	4.41	-4.41
		758 COURTLANDT AVENUE	10451.0	10.0	6.27	3.73
		770 GROTE STREET	10460.0	3.0	3.91	-0.91
		800 BAYCHESTER AVENUE	10475.0	24.0	6.27	17.73
		800 LYDIG AVENUE	10462.0	0.0	3.49	-3.49
		800 TAYLOR AVENUE	10473.0	5.0	4.41	0.59
		811 EAST 149 STREET	10455.0	2.0	4.41	-2.41
		890 CAULDWELL AVENUE	10456.0	15.0	3.49	11.51
		900 ADEE AVENUE	10469.0	4.0	8.77	-4.77
		916 EAGLE AVENUE	10456.0	6.0	4.41	1.59
		920 EAST 167 STREET	10459.0	8.0	6.27	1.73
		928 SIMPSON STREET	10459.0	31.0	4.41	26.59
		950 RHINELANDER AVENUE	10462.0	16.0	10.22	5.78
		968 CAULDWELL AVENUE	10456.0	2.0	3.91	-1.91
		980 MACE AVENUE	10469.0	5.0	12.69	-7.69
		984 FAILE STREET	10459.0	2.0	4.41	-2.41

424 rows × 4 columns

Now we will repeat this process four more times to create a data frame for each borough like the Bronx one.


```
In [736]: brooklyn = crime[crime['Borough_Name'] == "BROOKLYN "]  
brooklyn
```

Out[736]:

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
0	2015-16	K006	K006	43 SNYDER AVENUE	BROOKLYN	11226.0	4.0	4.41
1	2015-16	K005	K005	820 HANCOCK STREET	BROOKLYN	11233.0	2.0	3.49
2	2015-16	K003	K003	50 JEFFERSON AVENUE	BROOKLYN	11216.0	2.0	3.91
8	2015-16	K001	K001	309 47 STREET	BROOKLYN	11220.0	2.0	12.69
9	2015-16	K008	K008	37 HICKS STREET	BROOKLYN	11201.0	6.0	6.27
10	2015-16	K095	K095	345 VAN SICKLEN STREET	BROOKLYN	11223.0	0.0	6.27
11	2015-16	K165	K165	76 LOTT AVENUE	BROOKLYN	11212.0	13.0	3.91
13	2015-16	K857	K036	2045 LINDEN BOULEVARD	BROOKLYN	11207.0	1.0	3.91
1404	2015-16	KBZT	K746	2186 MILL AVENUE	BROOKLYN	11234.0	0.0	4.41
1405	2015-16	KBXX	K797	475 EAST 57 STREET	BROOKLYN	11203.0	0.0	3.91
1406	2015-16	KBNU	K758	500 19TH STREET	BROOKLYN	11215.0	0.0	3.91
1407	2015-16	KAQX	K731	856 QUINCY STREET	BROOKLYN	11221.0	0.0	4.41
1408	2015-16	KAGI	K707	153 35th St.	BROOKLYN	11232.0	0.0	6.27
1409	2015-16	K994	K595	1119 BEDFORD AVENUE	BROOKLYN	11216.0	5.0	3.91
1410	2015-16	K987	K498	1396 BROADWAY	BROOKLYN	11221.0	15.0	4.41
1411	2015-16	K971	K971	6214 4TH AVENUE	BROOKLYN	11220.0	1.0	3.91
1413	2015-16	K917	K555	350 CONEY ISLAND AVENUE	BROOKLYN	11218.0	2.0	4.41
1414	2015-16	K914	K670	71-77 CLINTON AVENUE	BROOKLYN	11205.0	12.0	6.27
1415	2015-16	K913	K545	1155 DEKALB AVENUE	BROOKLYN	11221.0	11.0	4.41
1416	2015-16	K909	K499	561 GRAND AVENUE	BROOKLYN	11238.0	9.0	3.91
1417	2015-16	K907	K568	1150 EAST NEW YORK AVENUE	BROOKLYN	11212.0	15.0	3.49

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
1421	2015-16	K894	K646	1495 HERKIMER STREET	BROOKLYN	11207.0	27.0	3.49
1422	2015-16	K865	K556	797 BUSHWICK AVENUE	BROOKLYN	11221.0	10.0	3.91
1423	2015-16	K864	K361	3109 NEWKIRK AVENUE	BROOKLYN	11226.0	1.0	4.41
1425	2015-16	K860	K673	9517 KINGS HIGHWAY	BROOKLYN	11212.0	5.0	3.49
1426	2015-16	K849	K134	4001 18 AVENUE	BROOKLYN	11218.0	0.0	4.41
1430	2015-16	K843	K004	530 STANLEY AVENUE	BROOKLYN	11207.0	0.0	3.91
1434	2015-16	K825	K698	173 CONOVER STREET	BROOKLYN	11231.0	5.0	3.49
1435	2015-16	K824	K489	402 EASTERN PARKWAY	BROOKLYN	11225.0	7.0	3.49
1436	2015-16	K819	K409	2057 LINDEN BOULEVARD	BROOKLYN	11207.0	1.0	3.91
...
6242	2013-14	K112	K112	7115 15 AVENUE	BROOKLYN	11228.0	1.0	4.83
6246	2013-14	K110	K110	124 MONITOR STREET	BROOKLYN	11222.0	2.0	3.73
6247	2013-14	K106	K106	1328 PUTNAM AVENUE	BROOKLYN	11221.0	4.0	4.83
6248	2013-14	K105	K105	1031 59 STREET	BROOKLYN	11219.0	2.0	20.29
6249	2013-14	K101	K101	2360 BENSON AVENUE	BROOKLYN	11214.0	1.0	6.63
6250	2013-14	K100	K100	2951 WEST 3 STREET	BROOKLYN	11224.0	3.0	4.83
6251	2013-14	K099	K099	1120 EAST 10 STREET	BROOKLYN	11230.0	1.0	6.63
6252	2013-14	K095	K095	345 VAN SICKLEN STREET	BROOKLYN	11223.0	3.0	6.63
6253	2013-14	K094	K094	5010 6 AVENUE	BROOKLYN	11220.0	2.0	15.35
6254	2013-14	K093	K093	31 NEW YORK AVENUE	BROOKLYN	11216.0	2.0	3.73
6258	2013-14	K091	K091	532 ALBANY AVENUE	BROOKLYN	11203.0	8.0	3.73
6259	2013-14	K090	K090	2840 WEST 12 STREET	BROOKLYN	11224.0	6.0	4.83

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
6263	2013-14	K086	K086	220 IRVING AVENUE	BROOKLYN	11237.0	0.0	3.73
6264	2013-14	K085	K753	510 CLERMONT AVENUE	BROOKLYN	11238.0	16.0	4.09
6265	2013-14	K084	K084	250 BERRY STREET	BROOKLYN	11249.0	1.0	4.83
6266	2013-14	K081	K081	990 DEKALB AVENUE	BROOKLYN	11221.0	9.0	3.73
6267	2013-14	K078	K078	1420 EAST 68 STREET	BROOKLYN	11234.0	15.0	6.63
6268	2013-14	K075	K075	95 GROVE STREET	BROOKLYN	11221.0	1.0	4.83
6271	2013-14	K097	K097	1855 STILLWELL AVENUE	BROOKLYN	11223.0	3.0	6.63
6279	2013-14	K069	K069	6302 9TH AVENUE	BROOKLYN	11220.0	0.0	6.63
6280	2013-14	K068	K068	956 EAST 82 STREET	BROOKLYN	11236.0	29.0	4.83
6284	2013-14	K066	K066	845 EAST 96 STREET	BROOKLYN	11236.0	0.0	6.63
6285	2013-14	K062	K062	700 CORTELYOU ROAD	BROOKLYN	11218.0	4.0	13.13
6286	2013-14	K061	K061	400 EMPIRE BOULEVARD	BROOKLYN	11225.0	23.0	6.63
6288	2013-14	K096	K096	99 AVENUE P	BROOKLYN	11204.0	13.0	4.83
6291	2013-14	K058	K058	330 SMITH STREET	BROOKLYN	11231.0	0.0	6.63
6294	2013-14	K104	K104	9115 5 AVENUE	BROOKLYN	11209.0	4.0	8.93
6306	2013-14	K054	K054	195 SANFORD STREET	BROOKLYN	11205.0	2.0	3.73
6307	2013-14	K052	K052	2675 EAST 29 STREET	BROOKLYN	11235.0	2.0	6.63
6308	2013-14	K051	K051	350 5 AVENUE	BROOKLYN	11215.0	0.0	8.93

763 rows × 8 columns

```
In [737]: brooklyndf = brooklyn.groupby(["Borough_Name", "Year", "Address"])
brooklyn_dict = {"ZIP_CODE": "first", "Total Number of Crimes": "first", "Average Num of Crimes": "first"}
brooklyn_df = brooklyndf.agg(brooklyn_dict)

brooklyn_df["Standard_deviation"] = brooklyn_df["Total Number of Crimes"] - brooklyn_df["Average Num of Crimes"]
brooklyn_df
```

Out[737]:

			ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
Borough_Name	Year	Address				
BROOKLYN	2013-14	1 ALBEMARLE ROAD	11218.0	0.0	13.13	-13.13
		1 WELLS STREET	11208.0	19.0	8.93	10.07
		100 IRVING AVENUE	11237.0	0.0	6.63	-6.63
		100 NOLL STREET	11206.0	8.0	6.63	1.37
		1001 EAST 45 STREET	11203.0	1.0	4.83	-3.83
		1001 HERKIMER STREET	11233.0	0.0	4.09	-4.09
		101 MAUJER STREET	11206.0	7.0	4.09	2.91
		101 WALTON STREET	11206.0	3.0	15.35	-12.35
		101-24 SEAVIEW AVENUE	11236.0	2.0	4.83	-2.83
		1023 NEW YORK AVENUE	11203.0	11.0	8.93	2.07
		1031 59 STREET	11219.0	2.0	20.29	-18.29
		1037 EAST 54 STREET	11234.0	7.0	4.83	2.17
		1053 41ST STREET	11219.0	2.0	4.09	-2.09
		1060 CLARKSON AVENUE	11212.0	0.0	4.83	-4.83
		1070 EAST 83 STREET	11236.0	0.0	6.63	-6.63
		1070 EAST 104 STREET	11236.0	0.0	4.83	-4.83
		108 MONTROSE AVENUE	11206.0	0.0	6.63	-6.63
		110 CHESTER STREET	11212.0	4.0	3.73	0.27
		1100 EAST NEW YORK AVENUE	11212.0	6.0	8.93	-2.93
		1100 ELM AVENUE	11230.0	0.0	3.73	-3.73
		1100 NEWKIRK AVENUE	11230.0	4.0	13.13	-9.13
		111 BERRIMAN STREET	11208.0	0.0	4.83	-4.83
		111 BRISTOL STREET	11212.0	5.0	4.83	0.17

Borough_Name	Year	Address	ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
		1119 BEDFORD AVENUE	11216.0	4.0	3.73	0.27
		112 SCHERMERHORN STREET	11201.0	3.0	4.09	-1.09
		1120 EAST 10 STREET	11230.0	1.0	6.63	-5.63
		114 KOSCIUSKO STREET	11216.0	3.0	3.73	-0.73
		1150 EAST NEW YORK AVENUE	11212.0	14.0	4.09	9.91
		1155 DEKALB AVENUE	11221.0	6.0	4.83	1.17
		1171 65 STREET	11219.0	1.0	6.63	-5.63

	2015- 16	8010 12 AVENUE	11228.0	4.0	12.69	-8.69
		8101 15 AVENUE	11228.0	2.0	8.77	-6.77
		820 HANCOCK STREET	11233.0	2.0	3.49	-1.49
		825 4 AVENUE	11232.0	2.0	4.41	-2.41
		8301 SHORE ROAD	11209.0	29.0	19.33	9.67
		8310 21ST AVENUE	11214.0	6.0	6.27	-0.27
		84 SCHAEFER STREET	11207.0	3.0	6.27	-3.27
		845 EAST 96 STREET	11236.0	10.0	6.27	3.73
		850 KENT AVENUE	11205.0	1.0	3.91	-2.91
		856 QUINCY STREET	11221.0	0.0	4.41	-4.41
		858 JAMAICA AVENUE	11208.0	0.0	6.27	-6.27
		8601 RIDGE BOULEVARD	11209.0	0.0	4.41	-4.41
		87 BAY 49 STREET	11214.0	0.0	4.41	-4.41
		875 WILLIAMS AVENUE	11207.0	0.0	3.91	-3.91
		8787 24 AVENUE	11214.0	7.0	8.77	-1.77
		88 WOODBINE STREET	11221.0	0.0	3.91	-3.91

Borough_Name	Year	Address	ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
		900 ST MARKS AVENUE	11213.0	3.0	3.91	-0.91
		901 CLASSON AVENUE	11225.0	36.0	12.69	23.31
		9115 5 AVENUE	11209.0	4.0	8.77	-4.77
		923 JEROME STREET	11207.0	2.0	3.91	-1.91
		9301 AVENUE B	11236.0	0.0	4.41	-4.41
		942 62ND STREET	11219.0	3.0	3.91	-0.91
		95 GROVE STREET	11221.0	4.0	3.91	0.09
		9517 KINGS HIGHWAY	11212.0	5.0	3.49	1.51
		956 EAST 82 STREET	11236.0	11.0	3.91	7.09
		965 EAST 107TH STREET	11236.0	6.0	3.91	2.09
		970 VERMONT STREET	11207.0	16.0	4.41	11.59
		976 PRESIDENT STREET	11225.0	4.0	4.41	-0.41
		982 HEGEMAN AVENUE	11208.0	3.0	3.91	-0.91
		990 DEKALB AVENUE	11221.0	2.0	3.91	-1.91

763 rows × 4 columns


```
In [738]: manhattan = crime[crime['Borough_Name'] == "MANHATTAN"]  
manhattan
```

Out[738]:

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
990	2015-16	MANN	M353	1818 Amsterdam Ave	MANHATTAN	10031.0	1.0	6.27
991	2015-16	M971	M397	12 SPRUCE STREET	MANHATTAN	10038.0	0.0	3.91
992	2015-16	M970	M670	200-214 WEST 135TH STREET	MANHATTAN	10030.0	10.0	4.41
994	2015-16	M933	M339	425 WEST 33RD STREET	MANHATTAN	10001.0	0.0	3.49
995	2015-16	M920	M570	120 WEST 30 STREET	MANHATTAN	10001.0	1.0	3.91
996	2015-16	M918	M278	421 W 219TH ST	MANHATTAN	10034.0	1.0	4.41
997	2015-16	M916	M517	168 MORNINGSIDE AVENUE	MANHATTAN	10027.0	1.0	3.91
998	2015-16	M911	M285	22 EAST 128 STREET	MANHATTAN	10035.0	4.0	3.91
999	2015-16	M905	M442	890 BROADWAY	MANHATTAN	10003.0	0.0	3.49
1000	2015-16	M898	M018	4124 9 AVENUE	MANHATTAN	10034.0	1.0	3.91
1001	2015-16	M896	M531	123 WEST 43 STREET	MANHATTAN	10036.0	0.0	3.49
1002	2015-16	M895	M610	105 EAST 106TH STREET	MANHATTAN	10029.0	2.0	3.91
1003	2015-16	M894	M425	90 TRINITY PLACE	MANHATTAN	10006.0	16.0	4.41
1007	2015-16	M877	M551	10 SOUTH STREET, SLIP 7	MANHATTAN	10004.0	3.0	3.91
1009	2015-16	M874	M411	55 EAST 25 STREET	MANHATTAN	10010.0	1.0	3.91
1010	2015-16	M873	M439	43 WEST 22 STREET	MANHATTAN	10010.0	0.0	3.91
1011	2015-16	M860	M343	52 CHAMBERS STREET	MANHATTAN	10007.0	1.0	3.91
1012	2015-16	M855	M416	411 EAST 76 STREET	MANHATTAN	10021.0	4.0	4.41
1013	2015-16	M841	M811	466 WEST END AVENUE	MANHATTAN	10024.0	3.0	3.91
1014	2015-16	M837	M479	227-243 WEST 61ST STREET	MANHATTAN	10023.0	0.0	12.69
1019	2015-16	M807	M586	2-10 ASTOR PLACE	MANHATTAN	10003.0	8.0	3.49

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
1020	2015-16	M751	M751	113 EAST 4TH STREET	MANHATTAN	10003.0	3.0	3.49
1021	2015-16	M661	M575	240 2 AVENUE	MANHATTAN	10003.0	11.0	4.41
1022	2015-16	M660	M413	127 EAST 22 STREET	MANHATTAN	10010.0	7.0	4.41
1023	2015-16	M645	M655	320 EAST 96 STREET	MANHATTAN	10128.0	31.0	4.41
1041	2015-16	M600	M600	225 WEST 24 STREET	MANHATTAN	10011.0	37.0	15.65
1042	2015-16	M540	M540	443 WEST 135 STREET	MANHATTAN	10031.0	35.0	12.69
1049	2015-16	M528	M528	180 WADSWORTH AVENUE	MANHATTAN	10033.0	6.0	3.49
1054	2015-16	M824	M418	75 BROAD STREET	MANHATTAN	10004.0	0.0	4.41
1058	2015-16	M840	M150	334 GREENWICH STREET	MANHATTAN	10013.0	0.0	3.49
...
5583	2013-14	M097	M696	525 EAST HOUSTON STREET	MANHATTAN	10002.0	0.0	4.83
5588	2013-14	M079	M079	55 EAST 120 STREET	MANHATTAN	10035.0	10.0	3.73
5589	2013-14	M076	M076	220 WEST 121 STREET	MANHATTAN	10027.0	5.0	3.73
5593	2013-14	M072	M072	131 EAST 104 STREET	MANHATTAN	10029.0	0.0	3.73
5599	2013-14	M137	M184	327 CHERRY STREET	MANHATTAN	10002.0	2.0	4.83
5610	2013-14	M060	M450	420 EAST 12 STREET	MANHATTAN	10009.0	9.0	4.83
5611	2013-14	M058	M035	317 WEST 52ND STREET	MANHATTAN	10019.0	38.0	4.09
5612	2013-14	M057	M057	176 EAST 115 STREET	MANHATTAN	10029.0	5.0	6.63
5618	2013-14	M054	M054	103 WEST 107 STREET	MANHATTAN	10025.0	11.0	6.63
5625	2013-14	M048	M048	4360-78 BROADWAY	MANHATTAN	10033.0	0.0	4.83
5638	2013-14	M645	M655	320 EAST 96 STREET	MANHATTAN	10128.0	12.0	6.63
5643	2013-14	M042	M042	71 HESTER STREET	MANHATTAN	10002.0	0.0	6.63

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
5644	2013-14	M041	M041	116 WEST 11 STREET	MANHATTAN	10011.0	0.0	6.63
5648	2013-14	M036	M036	123 MORNINGSIDE DRIVE	MANHATTAN	10027.0	7.0	4.83
5649	2013-14	M034	M034	730 EAST 12 STREET	MANHATTAN	10009.0	1.0	3.73
5650	2013-14	M033	M033	281 9 AVENUE	MANHATTAN	10001.0	0.0	4.83
5655	2013-14	M028	M028	475 WEST 155 STREET	MANHATTAN	10032.0	2.0	6.63
5660	2013-14	M022	M539	111 COLUMBIA STREET	MANHATTAN	10002.0	10.0	15.35
5661	2013-14	M020	M020	166 ESSEX STREET	MANHATTAN	10002.0	6.0	4.83
5668	2013-14	M015	M015	333 EAST 4 STREET	MANHATTAN	10009.0	7.0	4.09
5675	2013-14	M011	M011	320 WEST 21 STREET	MANHATTAN	10011.0	2.0	6.63
5676	2013-14	M010	M499	2581 7TH AVENUE	MANHATTAN	10039.0	27.0	13.13
5683	2013-14	M006	M006	45 EAST 81 STREET	MANHATTAN	10028.0	0.0	6.63
5684	2013-14	M005	M005	3703 TENTH AVENUE	MANHATTAN	10034.0	1.0	4.83
5685	2013-14	M004	M004	500 WEST 160 STREET	MANHATTAN	10032.0	2.0	4.83
5686	2013-14	M003	M003	490 HUDSON STREET	MANHATTAN	10014.0	4.0	6.63
5687	2013-14	M002	M002	122 HENRY STREET	MANHATTAN	10002.0	3.0	6.63
5688	2013-14	M001	M001	8 HENRY STREET	MANHATTAN	10038.0	2.0	3.73
5835	2013-14	M008	M008	465 WEST 167 STREET	MANHATTAN	10032.0	2.0	4.83
5889	2013-14	M066	M151	421 EAST 88 STREET	MANHATTAN	10128.0	0.0	3.73

377 rows × 8 columns

```
In [739]: manhattandf = manhattan.groupby(["Borough_Name", "Year", "Address"])
manhattan_dict = {"ZIP_CODE":"first", "Total Number of Crimes":"first", "Average Num of Crimes":"first"}
manhattan_df = manhattandf.agg(manhattan_dict)

manhattan_df["Standard_deviation"] = manhattan_df["Total Number of Crimes"] -
manhattan_df["Average Num of Crimes"]
manhattan_df
```

Out[739]:

			ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
Borough_Name	Year	Address				
MANHATTAN	2013-14	10 SOUTH STREET, SLIP 7	10004.0	5.0	3.73	1.27
		100 AMSTERDAM AVENUE	10023.0	13.0	58.39	-45.39
		100 TRINITY PLACE	10006.0	14.0	6.63	7.37
		103 WEST 107 STREET	10025.0	11.0	6.63	4.37
		105 EAST 106 STREET	10029.0	4.0	3.73	0.27
		110 EAST 88 STREET	10128.0	15.0	3.73	11.27
		111 COLUMBIA STREET	10002.0	10.0	15.35	-5.35
		113 EAST 4TH STREET	10003.0	7.0	4.09	2.91
		116 WEST 11 STREET	10011.0	0.0	6.63	-6.63
		12 SPRUCE STREET	10038.0	2.0	3.73	-1.73
		12-18 ELLWOOD STREET	10040.0	2.0	4.09	-2.09
		120 WEST 30 STREET	10001.0	5.0	3.73	1.27
		120 WEST 46 STREET	10036.0	16.0	4.83	11.17
		122 HENRY STREET	10002.0	3.0	6.63	-3.63
		123 MORNINGSIDE DRIVE	10027.0	7.0	4.83	2.17
		123 RIDGE STREET	10002.0	1.0	3.73	-2.73
		123 WEST 43 STREET	10036.0	5.0	4.09	0.91
		127 EAST 22 STREET	10010.0	3.0	4.83	-1.83
		129 WEST 67 STREET	10023.0	1.0	4.09	-3.09
		131 EAST 104 STREET	10029.0	0.0	3.73	-3.73
		132 WEST 89 STREET	10024.0	4.0	4.83	-0.83

Borough_Name	Year	Address	ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
		140 WEST 102 STREET	10025.0	16.0	4.83	11.17
		141 EAST 111 STREET	10029.0	11.0	3.73	7.27
		143 BAXTER STREET	10013.0	0.0	6.63	-6.63
		1458 YORK AVENUE	10075.0	0.0	4.83	-4.83
		160 WEST 78 STREET	10024.0	0.0	6.63	-6.63
		1615 MADISON AVENUE	10029.0	6.0	4.83	1.17
		163 WEST 97 STREET	10025.0	0.0	4.83	-4.83
		166 ESSEX STREET	10002.0	6.0	4.83	1.17
		168 MORNINGSIDE AVENUE	10027.0	0.0	4.09	-4.09

	2015-16	443 WEST 135 STREET	10031.0	35.0	12.69	22.31
		45 EAST 81 STREET	10028.0	3.0	4.41	-1.41
		465 WEST 167 STREET	10032.0	0.0	4.41	-4.41
		466 WEST END AVENUE	10024.0	3.0	3.91	-0.91
		475 WEST 155 STREET	10032.0	6.0	4.41	1.59
		490 HUDSON STREET	10014.0	4.0	6.27	-2.27
		499 WEST 133 STREET	10027.0	6.0	6.27	-0.27
		500 WEST 160 STREET	10032.0	1.0	4.41	-3.41
		501-503 WEST 152 STREET	10031.0	0.0	3.91	-3.91
		504 WEST 158TH STREET	10032.0	3.0	4.41	-1.41
		512 WEST 212 STREET	10034.0	8.0	4.41	3.59
		52 CHAMBERS STREET	10007.0	1.0	3.91	-2.91

Borough_Name	Year	Address	ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
		525 EAST HOUSTON STREET	10002.0	1.0	4.41	-3.41
		525 WEST 44TH STREET	10036.0	0.0	3.91	-3.91
		535 EAST 119 STREET	10035.0	1.0	3.91	-2.91
		55 EAST 25 STREET	10010.0	1.0	3.91	-2.91
		55 EAST 120 STREET	10035.0	9.0	3.91	5.09
		586 WEST 177 STREET	10033.0	3.0	4.41	-1.41
		590 SIXTH AVENUE	10011.0	0.0	3.49	-3.49
		625 WEST 133 STREET	10027.0	6.0	3.49	2.51
		645 MAIN STREET	10044.0	0.0	4.41	-4.41
		71 HESTER STREET	10002.0	0.0	4.41	-4.41
		71-111 CONVENT AVE	10027.0	0.0	3.91	-3.91
		730 EAST 12 STREET	10009.0	4.0	3.91	0.09
		75 BROAD STREET	10004.0	0.0	4.41	-4.41
		8 HENRY STREET	10038.0	2.0	3.91	-1.91
		80 CATHERINE STREET	10038.0	3.0	6.27	-3.27
		890 BROADWAY	10003.0	0.0	3.49	-3.49
		90 TRINITY PLACE	10006.0	16.0	4.41	11.59
		93 NAGLE AVENUE	10040.0	0.0	4.41	-4.41

377 rows × 4 columns


```
In [740]: queens = crime[crime['Borough_Name'] == "QUEENS "]  
queens
```

Out[740]:

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num o Crimes
12	2015-16	Q849	Q077	976 SENECA AVENUE	QUEENS	11385.0	5.0	8.77
596	2015-16	QAZJ	Q404	269-01 76th Avenue	QUEENS	11004.0	1.0	10.22
597	2015-16	QAPM	Q023	74-03 COMMONWEALTH BOULEVARD	QUEENS	11426.0	0.0	3.91
598	2015-16	Q891	Q560	47-07 30TH PLACE	QUEENS	11101.0	4.0	4.41
599	2015-16	Q884	Q065	103-22 99 STREET	QUEENS	11417.0	0.0	3.91
600	2015-16	Q877	Q877	76-05 51 AVENUE	QUEENS	11373.0	1.0	3.91
601	2015-16	Q868	Q078	48-09 CENTER BOULEVARD	QUEENS	11109.0	0.0	4.41
602	2015-16	Q811	Q811	61-25 MARATHON PARKWAY	QUEENS	11362.0	1.0	3.91
603	2015-16	Q799	Q326	188-04 91ST AVENUE	QUEENS	11423.0	5.0	3.91
609	2015-16	Q739	Q286	23-15 NEWTOWN AVENUE	QUEENS	11102.0	0.0	4.41
614	2015-16	Q733	Q950	90-01 SUTPHIN BOULEVARD	QUEENS	11435.0	0.0	31.80
615	2015-16	Q985	Q344	15-15 HAZEN STREET	QUEENS	11370.0	0.0	4.41
616	2015-16	Q725	Q502	21-16 44TH ROAD	QUEENS	11101.0	5.0	6.27
617	2015-16	Q722	Q721	57-12 94 STREET	QUEENS	11373.0	0.0	3.91
618	2015-16	Q721	Q016	41-15 104 STREET	QUEENS	11368.0	0.0	10.22
619	2015-16	Q707	Q670	75-40 PARSONS BOULEVARD	QUEENS	11366.0	13.0	4.41
620	2015-16	Q695	Q680	160-20 GOETHALS AVENUE	QUEENS	11432.0	5.0	4.41
621	2015-16	Q690	Q690	116-25 GUY R BREWER BOULEVARD	QUEENS	11434.0	20.0	4.41
626	2015-16	Q680	Q896	150-91 87 ROAD	QUEENS	11432.0	7.0	4.41
627	2015-16	Q650	Q650	94-06 104TH STREET	QUEENS	11416.0	18.0	6.27
628	2015-16	Q636	Q334	94-25 117TH STREET	QUEENS	11419.0	0.0	3.45

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num o Crimes
629	2015-16	Q633	Q327	171-10 LINDEN BOULEVARD	QUEENS	11434.0	18.0	3.91
630	2015-16	Q620	Q620	165-65 84 AVENUE	QUEENS	11432.0	25.0	33.86
631	2015-16	Q610	Q610	45-30 36 STREET	QUEENS	11101.0	5.0	33.86
632	2015-16	Q600	Q600	37-02 47 AVENUE	QUEENS	11101.0	7.0	10.22
633	2015-16	Q595	Q338	162-02 HILLSIDE AVENUE	QUEENS	11432.0	0.0	3.49
636	2015-16	Q515	Q525	149-11 MELBOURNE AVENUE	QUEENS	11367.0	4.0	8.77
637	2015-16	Q505	Q505	160-05 HIGHLAND AVENUE	QUEENS	11432.0	34.0	31.80
638	2015-16	Q499	Q499	148-20 REEVES AVENUE	QUEENS	11367.0	0.0	3.91
642	2015-16	Q798	Q580	34-12 36 AVENUE	QUEENS	11106.0	0.0	3.91
...
5224	2013-14	Q034	Q034	104-12 SPRINGFIELD BOULEVARD	QUEENS	11429.0	3.0	4.83
5225	2013-14	Q033	Q033	91-37 222 STREET	QUEENS	11428.0	0.0	8.93
5226	2013-14	Q032	Q032	171-11 35 AVENUE	QUEENS	11358.0	2.0	6.63
5227	2013-14	Q031	Q031	211-45 46 ROAD	QUEENS	11361.0	2.0	4.83
5231	2013-14	Q029	Q029	125-10 23 AVENUE	QUEENS	11356.0	0.0	6.63
5232	2013-14	Q028	Q028	109-10 47 AVENUE	QUEENS	11368.0	0.0	4.83
5233	2013-14	Q026	Q026	195-02 69 AVENUE	QUEENS	11365.0	0.0	4.83
5237	2013-14	Q024	Q024	141-11 HOLLY AVENUE	QUEENS	11355.0	0.0	6.63
5238	2013-14	Q023	Q540	138-11 35 AVENUE	QUEENS	11354.0	3.0	3.73
5239	2013-14	Q021	Q021	147-36 26 AVENUE	QUEENS	11354.0	0.0	13.13
5240	2013-14	Q020	Q020	142-30 BARCLAY AVENUE	QUEENS	11355.0	0.0	13.13
5241	2013-14	Q019	Q019	98-02 ROOSEVELT AVENUE	QUEENS	11368.0	1.0	20.29

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num o Crimes
5242	2013-14	Q018	Q018	86-35 235 COURT	QUEENS	11427.0	0.0	4.8%
5243	2013-14	Q017	Q017	28-37 29 STREET	QUEENS	11102.0	0.0	4.8%
5244	2013-14	Q015	Q015	121-15 LUCAS STREET	QUEENS	11413.0	0.0	3.7%
5245	2013-14	Q014	Q014	107-01 OTIS AVENUE	QUEENS	11368.0	0.0	15.3%
5246	2013-14	Q013	Q013	55-01 94 STREET	QUEENS	11373.0	0.0	15.3%
5247	2013-14	Q012	Q012	42-00 72 STREET	QUEENS	11377.0	0.0	13.1%
5248	2013-14	Q011	Q011	54-25 SKILLMAN AVENUE	QUEENS	11377.0	0.0	13.1%
5249	2013-14	Q010	Q010	45-11 31ST AVE	QUEENS	11103.0	5.0	6.6%
5250	2013-14	Q009	Q009	58-74 57 STREET	QUEENS	11378.0	1.0	4.8%
5255	2013-14	Q007	Q007	80-55 CORNISH AVENUE	QUEENS	11373.0	0.0	8.9%
5256	2013-14	Q005	Q005	50-40 JACOBUS STREET	QUEENS	11373.0	6.0	15.3%
5257	2013-14	Q123	Q123	145-01 119 AVENUE	QUEENS	11436.0	1.0	4.8%
5258	2013-14	Q003	Q303	108-55 69 AVENUE	QUEENS	11375.0	2.0	4.0%
5259	2013-14	Q002	Q002	75-10 21 AVENUE	QUEENS	11370.0	3.0	4.8%
5303	2013-14	Q022	Q022	153-33 SANFORD AVENUE	QUEENS	11355.0	0.0	6.6%
5431	2013-14	Q251	Q251	144-51 ARTHUR STREET	QUEENS	11413.0	1.0	3.7%
5972	2013-14	Q188	Q188	218-12 HARTLAND AVENUE	QUEENS	11364.0	0.0	4.8%
6037	2013-14	Q079	Q079	147-27 15 DRIVE	QUEENS	11357.0	2.0	6.6%

778 rows × 8 columns

```
In [741]: queensdf = queens.groupby(["Borough_Name", "Year", "Address"])
queens_dict = {"ZIP_CODE": "first", "Total Number of Crimes": "first", "Average
Num of Crimes": "first"}
queens_df = queensdf.agg(queens_dict)

queens_df["Standard_deviation"] = queens_df["Total Number of Crimes"] - queens
_df["Average Num of Crimes"]
queens_df
```

Out[741]:

			ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
Borough_Name	Year	Address				
QUEENS	2013-14	101-33 124 STREET	11419.0	0.0	4.83	-4.83
		103-22 99 STREET	11417.0	0.0	4.83	-4.83
		104-12 SPRINGFIELD BOULEVARD	11429.0	3.0	4.83	-1.83
		107-01 OTIS AVENUE	11368.0	0.0	15.35	-15.35
		107-25 WREN PLACE	11433.0	10.0	6.63	3.37
		108-10 109 AVENUE	11420.0	0.0	13.13	-13.13
		108-29 155 STREET	11433.0	1.0	4.83	-3.83
		108-55 69 AVENUE	11375.0	2.0	4.09	-2.09
		109-10 47 AVENUE	11368.0	0.0	4.83	-4.83
		109-15 98 STREET	11417.0	20.0	20.29	-0.29
		109-20 UNION HALL STREET	11433.0	3.0	4.83	-1.83
		109-59 INWOOD STREET	11435.0	2.0	4.83	-2.83
		110-08 NORTHERN BOULEVARD	11368.0	0.0	3.73	-3.73
		111-11 118 STREET	11420.0	1.0	6.63	-5.63
		116-25 GUY R BREWER BOULEVARD	11434.0	27.0	4.83	22.17
		120-45 235 STREET	11411.0	0.0	4.83	-4.83
		121-15 LUCAS STREET	11413.0	0.0	3.73	-3.73
		125-10 23 AVENUE	11356.0	0.0	6.63	-6.63
		125-20 SUTPHIN BOULEVARD	11434.0	0.0	4.83	-4.83
		126-10 109 AVENUE	11420.0	0.0	6.63	-6.63

Borough_Name	Year	Address	ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
		126-28 150 STREET	11436.0	4.0	3.73	0.27
		128-02 7 AVENUE	11356.0	1.0	8.93	-7.93
		129-15 150 AVENUE	11420.0	2.0	13.13	-11.13
		130-01 ROCKAWAY BOULEVARD	11420.0	0.0	3.73	-3.73
		130-02 115 AVENUE	11420.0	0.0	4.83	-4.83
		1307 CENTRAL AVENUE	11691.0	0.0	4.83	-4.83
		131-10 97 AVENUE	11419.0	0.0	4.83	-4.83
		132-15 218 STREET	11413.0	0.0	3.73	-3.73
		132-55 RIDGEDALE STREET	11413.0	8.0	4.83	3.17
		133-25 GUY R BREWER BOULEVARD	11434.0	22.0	6.63	15.37

	2015-16	86-37 53RD AVENUE	11373.0	0.0	3.49	-3.49
		86-50 109 STREET	11418.0	1.0	6.27	-5.27
		87-41 PARSONS BOULEVARD	11432.0	1.0	6.27	-5.27
		87-45 117 STREET	11418.0	0.0	3.91	-3.91
		88-02 144 STREET	11435.0	0.0	4.41	-4.41
		88-07 102ND STREET	11418.0	0.0	3.91	-3.91
		88-15 182 STREET	11423.0	27.0	12.69	14.31
		89-02 32 AVENUE	11369.0	0.0	8.77	-8.77
		9 POWER ROAD	11693.0	1.0	3.49	-2.49
		90-01 SUTPHIN BOULEVARD	11435.0	0.0	31.80	-31.80
		90-07 101ST AVENUE	11416.0	0.0	3.91	-3.91

Borough_Name	Year	Address	ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
		90-15 SUTTER AVENUE	11417.0	0.0	12.69	-12.69
		91-02 88 AVENUE	11421.0	0.0	8.77	-8.77
		91-37 222 STREET	11428.0	0.0	8.77	-8.77
		92-07 175 STREET	11433.0	1.0	4.41	-3.41
		93-02 69 AVENUE	11375.0	0.0	6.27	-6.27
		93-06 63 DRIVE	11374.0	0.0	6.27	-6.27
		93-11 101 AVENUE	11416.0	17.0	15.65	1.35
		93-11 34 AVENUE	11372.0	0.0	8.77	-8.77
		94-06 104TH STREET	11416.0	18.0	6.27	11.73
		94-25 117TH STREET	11419.0	0.0	3.49	-3.49
		94-50 159 STREET	11433.0	0.0	3.91	-3.91
		95-16 89TH AVENUE	11421.0	0.0	3.91	-3.91
		97-25 108 STREET	11419.0	0.0	6.27	-6.27
		976 SENECA AVENUE	11385.0	5.0	8.77	-3.77
		98-01 159 AVENUE	11414.0	0.0	4.41	-4.41
		98-01 25 AVENUE	11369.0	1.0	12.69	-11.69
		98-02 ROOSEVELT AVENUE	11368.0	0.0	15.65	-15.65
		98-50 50 AVENUE	11368.0	4.0	33.86	-29.86
		99-01 34 AVENUE	11368.0	0.0	6.27	-6.27

777 rows × 4 columns


```
In [742]: staten = crime[crime['Borough_Name'] == "STATEN IS"]  
staten
```

Out[742]:

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
509	2015-16	R880	R080	715 OCEAN TERRACE	STATEN IS	10301.0	11.0	12.69
510	2015-16	R861	R861	280 REGIS DRIVE	STATEN IS	10314.0	0.0	6.27
511	2015-16	R840	R037	15 FAIRFIELD STREET	STATEN IS	10308.0	0.0	3.91
512	2015-16	R831	R074	211 DANIEL LOW TERRACE	STATEN IS	10301.0	3.0	3.91
513	2015-16	R829	R065	98 GRANT STREET	STATEN IS	10301.0	0.0	3.91
517	2015-16	R722	R721	155 TOMPKINS AVENUE	STATEN IS	10304.0	1.0	3.91
518	2015-16	R600	R600	290 ST MARKS PLACE	STATEN IS	10301.0	15.0	4.41
519	2015-16	R470	R470	109 RHINE AVENUE	STATEN IS	10304.0	6.0	3.49
520	2015-16	R460	R460	1200 MANOR ROAD	STATEN IS	10314.0	46.0	31.80
525	2015-16	R450	R450	105 HAMILTON AVENUE	STATEN IS	10301.0	44.0	33.86
526	2015-16	R445	R445	85 ST JOSEPHS AVENUE	STATEN IS	10302.0	47.0	12.69
527	2015-16	R440	R605	485 CLAWSON STREET	STATEN IS	10306.0	1.0	12.69
531	2015-16	R075	R075	455 HUGUENOT AVENUE	STATEN IS	10312.0	5.0	12.69
532	2015-16	R072	R072	33 FERNDAL AVENUE	STATEN IS	10314.0	6.0	12.69
533	2015-16	R071	R048	1050 Targee Street	STATEN IS	10304.0	0.0	6.27
534	2015-16	R069	R069	144 KEATING PLACE	STATEN IS	10314.0	0.0	6.27
535	2015-16	R061	R061	445 CASTLETON AVENUE	STATEN IS	10301.0	26.0	8.77
536	2015-16	R060	R060	55 MERRILL AVENUE	STATEN IS	10314.0	2.0	6.27
537	2015-16	R059	R059	300 Richmond Terrace	STATEN IS	10301.0	0.0	3.49
538	2015-16	R057	R057	140 PALMA DRIVE	STATEN IS	10304.0	3.0	4.41
539	2015-16	R056	R056	250 KRAMER AVENUE	STATEN IS	10309.0	1.0	4.41

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
540	2015-16	R055	R055	54 OSBORNE STREET	STATEN IS	10312.0	0.0	4.41
541	2015-16	R054	R054	1060 WILLOWBROOK ROAD	STATEN IS	10314.0	1.0	6.27
542	2015-16	R053	R053	330 DURANT AVENUE	STATEN IS	10308.0	2.0	6.27
543	2015-16	R052	R052	450 BUEL AVENUE	STATEN IS	10305.0	2.0	4.41
544	2015-16	R051	R051	20 HOUSTON STREET	STATEN IS	10302.0	35.0	12.69
545	2015-16	R050	R050	200 ADELAIDE AVENUE	STATEN IS	10306.0	2.0	6.27
550	2015-16	R048	R009	1055 TARGEE STREET	STATEN IS	10304.0	0.0	3.49
551	2015-16	R046	R046	41 REID AVENUE	STATEN IS	10305.0	2.0	3.91
552	2015-16	R045	R045	58 LAWRENCE AVENUE	STATEN IS	10310.0	1.0	6.27
...
4829	2013-14	R041	R041	216 CLAWSON STREET	STATEN IS	10306.0	6.0	4.83
4833	2013-14	R039	R039	71 SAND LANE	STATEN IS	10305.0	5.0	4.83
4834	2013-14	R038	R038	421 LINCOLN AVENUE	STATEN IS	10306.0	2.0	3.73
4835	2013-14	R036	R036	255 IONIA AVENUE	STATEN IS	10312.0	0.0	6.63
4836	2013-14	R035	R035	60 FOOTE AVENUE	STATEN IS	10301.0	0.0	3.73
4837	2013-14	R034	R034	528 ACADEMY AVENUE	STATEN IS	10307.0	0.0	8.93
4838	2013-14	R032	R032	32 ELVERTON AVENUE	STATEN IS	10308.0	0.0	6.63
4839	2013-14	R031	R031	55 LAYTON AVENUE	STATEN IS	10301.0	1.0	3.73
4840	2013-14	R030	R030	200 WARDWELL AVENUE	STATEN IS	10314.0	2.0	6.63
4841	2013-14	R029	R029	1581 VICTORY BOULEVARD	STATEN IS	10314.0	0.0	6.63
4842	2013-14	R027	R027	11 CLOVE LAKE PLACE	STATEN IS	10310.0	33.0	8.93
4843	2013-14	R026	R026	4108 VICTORY BOULEVARD	STATEN IS	10314.0	2.0	4.09

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
4844	2013-14	R024	R024	225 CLEVELAND AVENUE	STATEN IS	10308.0	8.0	13.13
4845	2013-14	R023	R023	30 NATICK STREET	STATEN IS	10306.0	0.0	3.73
4846	2013-14	R022	R022	1860 FOREST AVENUE	STATEN IS	10303.0	3.0	8.93
4847	2013-14	R021	R021	168 HOOKER PLACE	STATEN IS	10302.0	0.0	3.73
4848	2013-14	R020	R020	161 PARK AVENUE	STATEN IS	10302.0	4.0	3.73
4849	2013-14	R019	R019	780 POST AVENUE	STATEN IS	10310.0	1.0	4.83
4850	2013-14	R018	R018	221 BROADWAY	STATEN IS	10310.0	5.0	4.83
4851	2013-14	R016	R016	80 MONROE AVENUE	STATEN IS	10301.0	6.0	4.83
4855	2013-14	R013	R013	191 VERMONT AVENUE	STATEN IS	10305.0	2.0	6.63
4856	2013-14	R011	R011	50 JEFFERSON STREET	STATEN IS	10304.0	6.0	3.73
4857	2013-14	R008	R008	112 LINDENWOOD ROAD	STATEN IS	10308.0	1.0	4.83
4858	2013-14	R006	R006	555 PAGE AVENUE	STATEN IS	10307.0	1.0	6.63
4859	2013-14	R004	R004	200 NEDRA LANE	STATEN IS	10312.0	1.0	6.63
4860	2013-14	R003	R003	80 SOUTH GOFF AVENUE	STATEN IS	10309.0	2.0	8.93
4861	2013-14	R002	R002	333 MIDLAND AVENUE	STATEN IS	10306.0	10.0	6.63
4862	2013-14	R001	R001	58 SUMMIT STREET	STATEN IS	10307.0	1.0	4.83
4884	2013-14	R005	R005	348 DEISIUS STREET	STATEN IS	10312.0	5.0	4.09
4920	2013-14	R007	R007	1270 HUGUENOT AVENUE	STATEN IS	10312.0	18.0	8.93

188 rows × 8 columns

```
In [743]: statendf = staten.groupby(["Borough_Name", "Year", "Address"])
staten_dict = {"ZIP_CODE": "first", "Total Number of Crimes": "first", "Average
Num of Crimes": "first"}
staten_df = statendf.agg(staten_dict)

staten_df["Standard_deviation"] = staten_df["Total Number of Crimes"] - staten
_df["Average Num of Crimes"]
staten_df
```

Out[743]:

			ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
Borough_Name	Year	Address				
STATEN IS	2013-14	101 WARREN STREET	10304.0	19.0	6.63	12.37
		105 HAMILTON AVENUE	10301.0	95.0	33.11	61.89
		1050 Targee Street	10304.0	1.0	4.83	-3.83
		1055 TARGEE STREET	10304.0	2.0	4.09	-2.09
		1060 WILLOWBROOK ROAD	10314.0	2.0	6.63	-4.63
		109 RHINE AVENUE	10304.0	2.0	4.09	-2.09
		11 CLOVE LAKE PLACE	10310.0	33.0	8.93	24.07
		112 LINDENWOOD ROAD	10308.0	1.0	4.83	-3.83
		1200 MANOR ROAD	10314.0	62.0	35.31	26.69
		1270 HUGUENOT AVENUE	10312.0	18.0	8.93	9.07
		140 PALMA DRIVE	10304.0	1.0	4.83	-3.83
		144 KEATING PLACE	10314.0	0.0	6.63	-6.63
		15 FAIRFIELD STREET	10308.0	1.0	3.73	-2.73
		155 TOMPKINS AVENUE	10304.0	8.0	3.73	4.27
		1581 VICTORY BOULEVARD	10314.0	0.0	6.63	-6.63
		161 PARK AVENUE	10302.0	4.0	3.73	0.27
		168 HOOKER PLACE	10302.0	0.0	3.73	-3.73
		1860 FOREST AVENUE	10303.0	3.0	8.93	-5.93
		191 VERMONT AVENUE	10305.0	2.0	6.63	-4.63
		20 HOUSTON STREET	10302.0	16.0	8.93	7.07
		200 ADELAIDE AVENUE	10306.0	0.0	6.63	-6.63

Borough_Name	Year	Address	ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
		200 NEDRA LANE	10312.0	1.0	6.63	-5.63
		200 WARDWELL AVENUE	10314.0	2.0	6.63	-4.63
		211 DANIEL LOW TERRACE	10301.0	5.0	3.73	1.27
		216 CLAWSON STREET	10306.0	6.0	4.83	1.17
		221 BROADWAY	10310.0	5.0	4.83	0.17
		225 CLEVELAND AVENUE	10308.0	8.0	13.13	-5.13
		250 KRAMER AVENUE	10309.0	0.0	4.83	-4.83
		255 IONIA AVENUE	10312.0	0.0	6.63	-6.63
		280 REGIS DRIVE	10314.0	1.0	6.63	-5.63

	2015- 16	32 ELVERTON AVENUE	10308.0	5.0	4.41	0.59
		33 FERNDAL AVENUE	10314.0	6.0	12.69	-6.69
		330 DURANT AVENUE	10308.0	2.0	6.27	-4.27
		333 MIDLAND AVENUE	10306.0	13.0	6.27	6.73
		348 DEISIUS STREET	10312.0	0.0	3.91	-3.91
		380 GENESEE AVENUE	10312.0	1.0	6.27	-5.27
		41 REID AVENUE	10305.0	2.0	3.91	-1.91
		4108 VICTORY BOULEVARD	10314.0	0.0	3.49	-3.49
		421 LINCOLN AVENUE	10306.0	3.0	3.91	-0.91
		445 CASTLETON AVENUE	10301.0	26.0	8.77	17.23
		450 BUEL AVENUE	10305.0	2.0	4.41	-2.41
		455 HUGUENOT AVENUE	10312.0	5.0	12.69	-7.69
		485 CLAWSON STREET	10306.0	1.0	12.69	-11.69

Borough_Name	Year	Address	ZIP_CODE	Total Number of Crimes	Average Num of Crimes	Standard_deviation
		50 JEFFERSON STREET	10304.0	2.0	3.91	-1.91
		528 ACADEMY AVENUE	10307.0	3.0	8.77	-5.77
		54 OSBORNE STREET	10312.0	0.0	4.41	-4.41
		55 LAYTON AVENUE	10301.0	0.0	3.91	-3.91
		55 MERRILL AVENUE	10314.0	2.0	6.27	-4.27
		555 PAGE AVENUE	10307.0	1.0	4.41	-3.41
		58 LAWRENCE AVENUE	10310.0	1.0	6.27	-5.27
		58 SUMMIT STREET	10307.0	4.0	4.41	-0.41
		60 FOOTE AVENUE	10301.0	1.0	3.91	-2.91
		71 SAND LANE	10305.0	3.0	4.41	-1.41
		715 OCEAN TERRACE	10301.0	11.0	12.69	-1.69
		77 MARSH AVENUE	10314.0	3.0	6.27	-3.27
		780 POST AVENUE	10310.0	0.0	4.41	-4.41
		80 MAPLE PARKWAY	10303.0	10.0	6.27	3.73
		80 SOUTH GOFF AVENUE	10309.0	2.0	6.27	-4.27
		85 ST JOSEPHS AVENUE	10302.0	47.0	12.69	34.31
		98 GRANT STREET	10301.0	0.0	3.91	-3.91

188 rows × 4 columns

Now we will make a safety dataframe that contains the information we want for all the boroughs. This will be useful later when we merge data.


```
In [744]: safety = bronx.append(brooklyn)
safety = safety.append(manhattan)
safety = safety.append(queens)
safety = safety.append(staten)
safety[['Borough_Name', 'Year', 'ZIP_CODE', 'Address', 'Total Number of Crimes',
'Average Num of Crimes']]
safety["Standard_deviation"] = safety["Total Number of Crimes"] - safety["Average Num of Crimes"]
safety
```

Out[744]:

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
16	2015-16	X905	X696	2925 GOULDEN AVENUE	BRONX	10468.0	0.0	3.91
17	2015-16	X886	X307	124 EAMES PLACE	BRONX	10468.0	0.0	3.91
20	2015-16	X878	X682	1021 JENNINGS STREET	BRONX	10460.0	12.0	3.91
21	2015-16	X876	X525	1710 WEBSTER AVENUE	BRONX	10457.0	20.0	4.41
22	2015-16	X859	X315	2246 JEROME AVENUE	BRONX	10453.0	0.0	3.91
23	2015-16	X852	X308	3617 BAINBRIDGE AVENUE	BRONX	10467.0	1.0	3.91
24	2015-16	X843	X246	2641 GRAND CONCOURSE	BRONX	10468.0	0.0	4.41
25	2015-16	X839	X418	2040 ANTIN PL	BRONX	10462.0	7.0	3.91
26	2015-16	X834	X352	1330 BRISTOW STREET	BRONX	10459.0	0.0	4.41
27	2015-16	X826	X226	1950 SEDGWICK AVENUE	BRONX	10453.0	1.0	4.41
28	2015-16	X819	X207	3030 GODWIN TERRACE	BRONX	10463.0	0.0	3.91
32	2015-16	X991	X695	560 BROOK AVENUE	BRONX	10455.0	0.0	3.91
47	2015-16	X722	X721	2697 WESTCHESTER AVENUE	BRONX	10461.0	2.0	4.41
67	2015-16	X701	X723	3540 BIVONA STREET	BRONX	10475.0	0.0	3.91
77	2015-16	X879	X684	1122 EAST 180 STREET	BRONX	10460.0	26.0	3.91
78	2015-16	X445	X445	75 WEST 205 STREET	BRONX	10468.0	10.0	31.80
150	2015-16	X392	X269	928 SIMPSON STREET	BRONX	10459.0	31.0	4.41
151	2015-16	X368	X368	2975 TIBBETT AVENUE	BRONX	10463.0	8.0	6.27
156	2015-16	X360	X360	2880 KINGSBRIDGE TERRACE	BRONX	10463.0	1.0	3.91
157	2015-16	X340	X340	25 WEST 195 STREET	BRONX	10468.0	0.0	4.41

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
162	2015-16	X292	X357	800 LYDIG AVENUE	BRONX	10462.0	0.0	3.49
163	2015-16	X285	X361	200 West 167th Street	BRONX	10452.0	0.0	3.91
164	2015-16	X283	X228	400 EAST FORDHAM ROAD	BRONX	10458.0	0.0	3.91
165	2015-16	X279	X279	2100 WALTON AVENUE	BRONX	10453.0	8.0	8.77
166	2015-16	X254	X254	2452 WASHINGTON AVENUE	BRONX	10458.0	5.0	3.91
170	2015-16	X234	X300	2050 PROSPECT AVENUE	BRONX	10457.0	1.0	4.41
173	2015-16	X465	X231	4143 THIRD AVENUE	BRONX	10457.0	12.0	4.41
176	2015-16	X211	X319	560 EAST 179 STREET	BRONX	10457.0	2.0	3.49
177	2015-16	X209	X209	313 EAST 183RD STREET	BRONX	10458.0	1.0	3.91
178	2015-16	X206	X206	2280 AQUEDUCT AVENUE	BRONX	10468.0	3.0	3.49
...
4829	2013-14	R041	R041	216 CLAWSON STREET	STATEN IS	10306.0	6.0	4.83
4833	2013-14	R039	R039	71 SAND LANE	STATEN IS	10305.0	5.0	4.83
4834	2013-14	R038	R038	421 LINCOLN AVENUE	STATEN IS	10306.0	2.0	3.73
4835	2013-14	R036	R036	255 IONIA AVENUE	STATEN IS	10312.0	0.0	6.63
4836	2013-14	R035	R035	60 FOOTE AVENUE	STATEN IS	10301.0	0.0	3.73
4837	2013-14	R034	R034	528 ACADEMY AVENUE	STATEN IS	10307.0	0.0	8.93
4838	2013-14	R032	R032	32 ELVERTON AVENUE	STATEN IS	10308.0	0.0	6.63
4839	2013-14	R031	R031	55 LAYTON AVENUE	STATEN IS	10301.0	1.0	3.73
4840	2013-14	R030	R030	200 WARDWELL AVENUE	STATEN IS	10314.0	2.0	6.63
4841	2013-14	R029	R029	1581 VICTORY BOULEVARD	STATEN IS	10314.0	0.0	6.63

	Year	Building Code	Location Code	Address	Borough_Name	ZIP_CODE	Total Number of Crimes	Average Num of Crimes
4842	2013-14	R027	R027	11 CLOVE LAKE PLACE	STATEN IS	10310.0	33.0	8.93
4843	2013-14	R026	R026	4108 VICTORY BOULEVARD	STATEN IS	10314.0	2.0	4.09
4844	2013-14	R024	R024	225 CLEVELAND AVENUE	STATEN IS	10308.0	8.0	13.13
4845	2013-14	R023	R023	30 NATICK STREET	STATEN IS	10306.0	0.0	3.73
4846	2013-14	R022	R022	1860 FOREST AVENUE	STATEN IS	10303.0	3.0	8.93
4847	2013-14	R021	R021	168 HOOKER PLACE	STATEN IS	10302.0	0.0	3.73
4848	2013-14	R020	R020	161 PARK AVENUE	STATEN IS	10302.0	4.0	3.73
4849	2013-14	R019	R019	780 POST AVENUE	STATEN IS	10310.0	1.0	4.83
4850	2013-14	R018	R018	221 BROADWAY	STATEN IS	10310.0	5.0	4.83
4851	2013-14	R016	R016	80 MONROE AVENUE	STATEN IS	10301.0	6.0	4.83
4855	2013-14	R013	R013	191 VERMONT AVENUE	STATEN IS	10305.0	2.0	6.63
4856	2013-14	R011	R011	50 JEFFERSON STREET	STATEN IS	10304.0	6.0	3.73
4857	2013-14	R008	R008	112 LINDENWOOD ROAD	STATEN IS	10308.0	1.0	4.83
4858	2013-14	R006	R006	555 PAGE AVENUE	STATEN IS	10307.0	1.0	6.63
4859	2013-14	R004	R004	200 NEDRA LANE	STATEN IS	10312.0	1.0	6.63
4860	2013-14	R003	R003	80 SOUTH GOFF AVENUE	STATEN IS	10309.0	2.0	8.93
4861	2013-14	R002	R002	333 MIDLAND AVENUE	STATEN IS	10306.0	10.0	6.63
4862	2013-14	R001	R001	58 SUMMIT STREET	STATEN IS	10307.0	1.0	4.83
4884	2013-14	R005	R005	348 DEISIUS STREET	STATEN IS	10312.0	5.0	4.09
4920	2013-14	R007	R007	1270 HUGUENOT AVENUE	STATEN IS	10312.0	18.0	8.93

2533 rows × 9 columns

```
In [745]: safety_2013 = safety[safety["Year"]=="2013-14"]
safety_2014 = safety[safety["Year"]=="2014-15"]
safety_2015 = safety[safety["Year"]=="2015-16"]
```

Part 2: Housing Price Data

Next, we will read in the housing price data:

```
In [746]: bronx_2013 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_15 hous
ing prices\\2013_bronx.xlsx")
brooklyn_2013 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_15 h
ousing prices\\2013_brooklyn.xlsx")
manhattan_2013 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_15
housing prices\\2013_manhattan.xlsx")
queens_2013 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_15 hou
sing prices\\2013_queens.xlsx")
statenisland_2013 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_
15 housing prices\\2013_statenisland.xlsx")
```

```
In [747]: bronx_2013.columns=bronx_2013.iloc[3]
bronx_2013 = bronx_2013[4:] #dont' rerun
```

```
In [748]: bronx_2013["Borough_Name"] = "BRONX"  
bronx_2013
```

Out[748]:

3	BOROUGH	NEIGHBORHOOD	BUILDING CLASS CATEGORY	TAX CLASS AT PRESENT	BLOCK	LOT	EASE- MENT	BUILDING CLASS AT PRESENT
4	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3028	25		A5
5	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3039	28		A1
6	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3039	28		A1
7	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3046	39		A1
8	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3046	52		A1
9	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3048	53		S1
10	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	2927	130		B1
11	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	2929	117		B1
12	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3030	60		B3
13	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3035	27		B1
14	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3036	16		B2
15	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3039	65		B2
16	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3040	5		S2
17	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3046	54		B2
18	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3046	54		B2
19	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3052	37		S2

3	BOROUGH	NEIGHBORHOOD	BUILDING CLASS CATEGORY	TAX CLASS AT PRESENT	BLOCK	LOT	EASE- MENT	BUILDING CLASS AT PRESENT
20	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3036	33		C0
21	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3037	37		C0
22	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3038	125		C0
23	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3041	11		C0
24	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3046	28		C0
25	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3050	1		C0
26	2	BATHGATE	05 TAX CLASS 1 VACANT LAND	1B	3036	36		V0
27	2	BATHGATE	07 RENTALS - WALKUP APARTMENTS	2	2908	15		C1
28	2	BATHGATE	07 RENTALS - WALKUP APARTMENTS	2	2908	17		C1
29	2	BATHGATE	07 RENTALS - WALKUP APARTMENTS	2	2909	12		C7
30	2	BATHGATE	07 RENTALS - WALKUP APARTMENTS	2B	2917	6		C1
31	2	BATHGATE	07 RENTALS - WALKUP APARTMENTS	2B	2924	37		C7
32	2	BATHGATE	07 RENTALS - WALKUP APARTMENTS	2	2927	9		C7
33	2	BATHGATE	07 RENTALS - WALKUP APARTMENTS	2	2928	14		C7
...
5906	2	WOODLAWN	03 THREE FAMILY DWELLINGS	1	3383	76		C0

3	BOROUGH	NEIGHBORHOOD	BUILDING CLASS CATEGORY	TAX CLASS AT PRESENT	BLOCK	LOT	EASE- MENT	BUILDING CLASS AT PRESENT
5907	2	WOODLAWN	03 THREE FAMILY DWELLINGS	1	3385	44		C0
5908	2	WOODLAWN	03 THREE FAMILY DWELLINGS	1	3394	5		C0
5909	2	WOODLAWN	05 TAX CLASS 1 VACANT LAND	1B	3391	45		V0
5910	2	WOODLAWN	06 TAX CLASS 1 - OTHER	1	3381	19		G0
5911	2	WOODLAWN	06 TAX CLASS 1 - OTHER	1	3392	2		G0
5912	2	WOODLAWN	07 RENTALS - WALKUP APARTMENTS	2B	3369	26		C1
5913	2	WOODLAWN	07 RENTALS - WALKUP APARTMENTS	2A	3376	70		C2
5914	2	WOODLAWN	07 RENTALS - WALKUP APARTMENTS	2	3377	80		C1
5915	2	WOODLAWN	07 RENTALS - WALKUP APARTMENTS	2	3388	1		C7
5916	2	WOODLAWN	07 RENTALS - WALKUP APARTMENTS	2	3397	19		D1
5917	2	WOODLAWN	08 RENTALS - ELEVATOR APARTMENTS	2	3379	50		D6
5918	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3363	58		D4
5919	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3363	58		D4
5920	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3363	58		D4
5921	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3368	10		D4
5922	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3398	44		D4

3	BOROUGH	NEIGHBORHOOD	BUILDING CLASS CATEGORY	TAX CLASS AT PRESENT	BLOCK	LOT	EASE- MENT	BUILDING CLASS AT PRESENT
5923	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3398	44		D4
5924	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3398	44		D4
5925	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3398	44		D4
5926	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3398	44		D4
5927	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3398	44		D4
5928	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3398	87		D4
5929	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3398	87		D4
5930	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3398	87		D4
5931	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3398	97		D4
5932	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3398	97		D4
5933	2	WOODLAWN	10 COOPS - ELEVATOR APARTMENTS	2	3398	97		D4
5934	2	WOODLAWN	22 STORE BUILDINGS	4	3377	53		K1
5935	2	WOODLAWN	30 WAREHOUSES	4	3382	75		E1

5932 rows × 22 columns

```
In [749]: brooklyn_2013.columns=brooklyn_2013.iloc[3]
brooklyn_2013 = brooklyn_2013[4:] #dont' rerun
```

```
In [750]: brooklyn_2013["Borough_Name"] = "BROOKLYN"
```

```
In [751]: manhattan_2013.columns=manhattan_2013.iloc[3]
manhattan_2013 = manhattan_2013[4:] #dont' rerun
```

```
In [752]: manhattan_2013["Borough_Name"] = "MANHATTAN"
```

```
In [753]: queens_2013.columns=queens_2013.iloc[3]
queens_2013 = queens_2013[4:] #dont' rerun
```

```
In [754]: queens_2013["Borough_Name"] = "QUEENS"
```

```
In [755]: statenisland_2013.columns=statenisland_2013.iloc[3]
statenisland_2013 = statenisland_2013[4:] #dont' rerun
```

```
In [756]: statenisland_2013["Borough_Name"] = "STATEN IS"
```

Here, we are appending all the data for the year 2013 into one dataframe named housing13. We will make similar dataframes for each year.

```
In [757]: housing2013 = bronx_2013.append(brooklyn_2013)
housing2013 = housing2013.append(manhattan_2013)
housing2013 = housing2013.append(queens_2013)
housing2013 = housing2013.append(statenisland_2013)
housing2013
```

Out[757]:

3	BOROUGH	NEIGHBORHOOD	BUILDING CLASS CATEGORY	TAX CLASS AT PRESENT	BLOCK	LOT	EASE- MENT	BUILDING CLASS AT PRESENT
4	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3028	25		A5
5	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3039	28		A1
6	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3039	28		A1
7	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3046	39		A1
8	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3046	52		A1
9	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3048	53		S1
10	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	2927	130		B1
11	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	2929	117		B1
12	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3030	60		B3
13	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3035	27		B1
14	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3036	16		B2
15	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3039	65		B2
16	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3040	5		S2
17	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3046	54		B2
18	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3046	54		B2
19	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3052	37		S2

3	BOROUGH	NEIGHBORHOOD	BUILDING CLASS CATEGORY	TAX CLASS AT PRESENT	BLOCK	LOT	EASE- MENT	BUILDING CLASS AT PRESENT
20	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3036	33		C0
21	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3037	37		C0
22	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3038	125		C0
23	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3041	11		C0
24	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3046	28		C0
25	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3050	1		C0
26	2	BATHGATE	05 TAX CLASS 1 VACANT LAND	1B	3036	36		V0
27	2	BATHGATE	07 RENTALS - WALKUP APARTMENTS	2	2908	15		C1
28	2	BATHGATE	07 RENTALS - WALKUP APARTMENTS	2	2908	17		C1
29	2	BATHGATE	07 RENTALS - WALKUP APARTMENTS	2	2909	12		C7
30	2	BATHGATE	07 RENTALS - WALKUP APARTMENTS	2B	2917	6		C1
31	2	BATHGATE	07 RENTALS - WALKUP APARTMENTS	2B	2924	37		C7
32	2	BATHGATE	07 RENTALS - WALKUP APARTMENTS	2	2927	9		C7
33	2	BATHGATE	07 RENTALS - WALKUP APARTMENTS	2	2928	14		C7
...
7044	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7091	16		B9

3	BOROUGH	NEIGHBORHOOD	BUILDING CLASS CATEGORY	TAX CLASS AT PRESENT	BLOCK	LOT	EASE- MENT	BUILDING CLASS AT PRESENT
7045	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7091	188		B9
7046	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7092	15		B9
7047	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7092	20		B9
7048	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7092	113		B9
7049	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7092	134		B9
7050	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7092	138		B9
7051	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7105	564		B2
7052	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7105	565		B2
7053	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7267	130		B9
7054	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7267	206		B9
7055	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7267	214		B2
7056	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7267	214		B2
7057	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7316	10		B9
7058	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7316	67		B2
7059	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7316	90		B2
7060	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7317	82		B2

3	BOROUGH	NEIGHBORHOOD	BUILDING CLASS CATEGORY	TAX CLASS AT PRESENT	BLOCK	LOT	EASE- MENT	BUILDING CLASS AT PRESENT
7061	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7317	115		B2
7062	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7317	124		B2
7063	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7317	143		B2
7064	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7317	148		B2
7065	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7346	12		B2
7066	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7349	10		B9
7067	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7349	35		B9
7068	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7355	55		B2
7069	5	WOODROW	05 TAX CLASS 1 VACANT LAND	1B	7092	200		V0
7070	5	WOODROW	05 TAX CLASS 1 VACANT LAND	1B	7105	574		V0
7071	5	WOODROW	22 STORE BUILDINGS	4	7100	16		K6
7072	5	WOODROW	22 STORE BUILDINGS	4	7105	520		K1
7073	5	WOODROW	33 EDUCATIONAL FACILITIES	4	6998	24		W7

89974 rows × 22 columns

Now that housing data for all 5 boroughs have been appended to the dataframe housing2013, we will clean up the data and only keep the columns that we want.


```
In [758]: housing2013.columns = housing2013.columns.to_series().apply(lambda x: x.strip()  
())  
housing2013_df = housing2013[['ADDRESS', 'ZIP CODE', 'SALE PRICE', 'SALE DATE',  
'Borough_Name']]  
housing2013_df.dropna(inplace = True)  
housing2013_df["Year"] = "2013-14"  
housing2013_df
```

```
C:\me\Anaconda\lib\site-packages\ipykernel_launcher.py:3: SettingWithCopyWarning:  
ing:
```

A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

This is separate from the ipykernel package so we can avoid doing imports until

```
C:\me\Anaconda\lib\site-packages\ipykernel_launcher.py:4: SettingWithCopyWarning:  
ing:
```

A value is trying to be set on a copy of a slice from a DataFrame.

Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

after removing the cwd from sys.path.

Out[758]:

3	ADDRESS	ZIP CODE	SALE PRICE	SALE DATE	Borough_Name	Year
4	412 EAST 179TH STREET	10457	355000	2013-07-08 00:00:00	BRONX	2013-14
5	2329 WASHINGTON AVENUE	10458	120000	2013-11-14 00:00:00	BRONX	2013-14
6	2329 WASHINGTON AVENUE	10458	474819	2013-05-20 00:00:00	BRONX	2013-14
7	2075 BATHGATE AVENUE	10457	210000	2013-03-12 00:00:00	BRONX	2013-14
8	2047 BATHGATE AVENUE	10457	343116	2013-07-01 00:00:00	BRONX	2013-14
9	4411 3 AVENUE	10457	225000	2013-08-15 00:00:00	BRONX	2013-14
10	1467 CROTONA PLACE	10456	380000	2013-11-04 00:00:00	BRONX	2013-14
11	3860 3 AVENUE	10457	390000	2013-06-24 00:00:00	BRONX	2013-14
12	4469 PARK AVENUE	10457	207000	2013-06-27 00:00:00	BRONX	2013-14
13	454 EAST 179 STREET	10457	0	2013-07-01 00:00:00	BRONX	2013-14
14	4354 PARK AVENUE	10457	328000	2013-06-26 00:00:00	BRONX	2013-14
15	465 EAST 185 STREET	10458	369000	2013-03-07 00:00:00	BRONX	2013-14
16	4654-4656 PARK AVENUE	10458	308000	2013-02-28 00:00:00	BRONX	2013-14
17	2043 BATHGATE AVENUE	10457	0	2013-08-27 00:00:00	BRONX	2013-14
18	2043 BATHGATE AVENUE	10457	0	2013-01-09 00:00:00	BRONX	2013-14
19	4557 3 AVENUE	10458	0	2013-05-14 00:00:00	BRONX	2013-14
20	2091 WASHINGTON AVENUE	10457	0	2013-06-27 00:00:00	BRONX	2013-14
21	4418 PARK AVENUE	10457	440000	2013-02-15 00:00:00	BRONX	2013-14
22	468 EAST 183 STREET	10458	662115	2013-01-15 00:00:00	BRONX	2013-14
23	452 CYRUS PLACE	10458	0	2013-07-25 00:00:00	BRONX	2013-14
24	2092 WASHINGTON AVENUE	10457	510000	2013-12-19 00:00:00	BRONX	2013-14
25	2186 WASHINGTON AVENUE	10457	460000	2013-05-10 00:00:00	BRONX	2013-14
26	2085 WASHINGTON AVENUE	10457	40730	2013-01-09 00:00:00	BRONX	2013-14

3	ADDRESS	ZIP CODE	SALE PRICE	SALE DATE	Borough_Name	Year
27	442 EAST 176 STREET	10457	0	2013-06-26 00:00:00	BRONX	2013-14
28	446 EAST 176 STREET	10457	0	2013-06-26 00:00:00	BRONX	2013-14
29	4196 PARK AVENUE	10457	0	2013-06-26 00:00:00	BRONX	2013-14
30	1824 WASHINGTON AVENUE	10457	0	2013-06-26 00:00:00	BRONX	2013-14
31	4173 3 AVENUE	10457	50000	2013-08-15 00:00:00	BRONX	2013-14
32	3758 3 AVENUE	10456	0	2013-06-27 00:00:00	BRONX	2013-14
33	538 CLAREMONT PARKWAY	10456	0	2013-06-27 00:00:00	BRONX	2013-14
...
7044	1578 WOODROW ROAD	10309	0	2013-10-18 00:00:00	STATEN IS	2013-14
7045	1576 WOODROW ROAD	10309	0	2013-10-18 00:00:00	STATEN IS	2013-14
7046	96 CRABTREE AVENUE	10309	481549	2013-03-13 00:00:00	STATEN IS	2013-14
7047	78 CRABTREE AVENUE	10309	446059	2013-02-13 00:00:00	STATEN IS	2013-14
7048	98 CRABTREE AVENUE	10309	434501	2013-06-28 00:00:00	STATEN IS	2013-14
7049	30 TRINA LANE	10309	365000	2013-03-05 00:00:00	STATEN IS	2013-14
7050	14 TRINA LANE	10309	260000	2013-09-16 00:00:00	STATEN IS	2013-14
7051	17 PLUMTREE LANE	10309	464990	2013-09-04 00:00:00	STATEN IS	2013-14
7052	15 PLUMTREE LANE	10309	430950	2013-12-12 00:00:00	STATEN IS	2013-14
7053	65 SHARROTT'S LANE	10309	470000	2013-04-09 00:00:00	STATEN IS	2013-14
7054	21 HEMLOCK LANE	10309	485000	2013-07-02 00:00:00	STATEN IS	2013-14
7055	25 WHITE OAK LANE	10309	166666	2013-05-15 00:00:00	STATEN IS	2013-14
7056	25 WHITE OAK LANE	10309	333333	2013-05-15 00:00:00	STATEN IS	2013-14
7057	68 DARNELL LANE	10309	440000	2013-06-26 00:00:00	STATEN IS	2013-14
7058	163 DARNELL LANE	10309	463500	2013-07-03 00:00:00	STATEN IS	2013-14

3	ADDRESS	ZIP CODE	SALE PRICE	SALE DATE	Borough_Name	Year
7059	65 DARNELL LANE	10309	0	2013-09-11 00:00:00	STATEN IS	2013-14
7060	110 SHARROTTTS ROAD	10309	560000	2013-04-18 00:00:00	STATEN IS	2013-14
7061	59 ROBIN COURT	10309	498400	2013-09-27 00:00:00	STATEN IS	2013-14
7062	83 ROBIN COURT	10309	520000	2013-06-18 00:00:00	STATEN IS	2013-14
7063	64 ROBIN COURT	10309	482500	2013-10-03 00:00:00	STATEN IS	2013-14
7064	56 ROBIN COURT	10309	0	2013-12-18 00:00:00	STATEN IS	2013-14
7065	476 HARGOLD AVENUE	10309	816970	2013-12-27 00:00:00	STATEN IS	2013-14
7066	63 PHEASANT LANE	10309	505000	2013-05-15 00:00:00	STATEN IS	2013-14
7067	33 QUAIL LANE	10309	0	2013-04-09 00:00:00	STATEN IS	2013-14
7068	182 BLOOMINGDALE ROAD	10309	677500	2013-12-17 00:00:00	STATEN IS	2013-14
7069	N/A CRABTREE AVENUE	10309	0	2013-09-19 00:00:00	STATEN IS	2013-14
7070	TURNER STREET	10309	0	2013-09-19 00:00:00	STATEN IS	2013-14
7071	639 VETERANS ROAD WEST	10309	1300000	2013-05-29 00:00:00	STATEN IS	2013-14
7072	2790 ARTHUR KILL ROAD	10309	14850914	2013-07-01 00:00:00	STATEN IS	2013-14
7073	1870 DRUMGOOLE ROAD EAST	10309	0	2013-11-07 00:00:00	STATEN IS	2013-14

89974 rows × 6 columns

```
In [759]: bronx_2014 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_15 housing prices\\2014_bronx.xlsx")
brooklyn_2014 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_15 housing prices\\2014_brooklyn.xlsx")
manhattan_2014 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_15 housing prices\\2014_manhattan.xlsx")
queens_2014 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_15 housing prices\\2014_queens.xlsx")
statenisland_2014 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_15 housing prices\\2014_statenisland.xlsx")
```

```
In [760]: bronx_2014.columns=bronx_2014.iloc[3]
bronx_2014 = bronx_2014[4:] #dont' rerun
bronx_2014["Borough_Name"] = "BRONX"

brooklyn_2014.columns=brooklyn_2014.iloc[3]
brooklyn_2014 = brooklyn_2014[4:] #dont' rerun
brooklyn_2014["Borough_Name"] = "BROOKLYN"

manhattan_2014.columns=manhattan_2014.iloc[3]
manhattan_2014 = manhattan_2014[4:] #dont' rerun
manhattan_2014["Borough_Name"] = "MANHATTAN"

queens_2014.columns=queens_2014.iloc[3]
queens_2014 = queens_2014[4:] #dont' rerun
queens_2014["Borough_Name"] = "QUEENS"

statenisland_2014.columns=statenisland_2014.iloc[3]
statenisland_2014 = statenisland_2014[4:] #dont' rerun
statenisland_2014["Borough_Name"] = "STATEN IS"

housing2014 = bronx_2014.append(brooklyn_2014)
housing2014 = housing2014.append(manhattan_2014)
housing2014 = housing2014.append(queens_2014)
housing2014 = housing2014.append(statenisland_2014)
housing2014
```

Out[760]:

3	BOROUGH	NEIGHBORHOOD	BUILDING CLASS CATEGORY	TAX CLASS AT PRESENT	BLOCK	LOT	EASE- MENT	BUILDING CLASS AT PRESENT
4	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3030	59		A1
5	2	BATHGATE	01 ONE FAMILY DWELLINGS	1B	3035	42		V0
6	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3039	64		A1 4
7	2	BATHGATE	01 ONE FAMILY DWELLINGS	1	3046	52		A1 2
8	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	2912	117		B1
9	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	2912	131		B1
10	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	2912	152		B1
11	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3030	60		B3
12	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3035	140		B1
13	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3036	37		B2
14	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3036	37		B2
15	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3036	62		B1
16	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3039	6		B1 4
17	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3039	6		B1 4
18	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3039	30		B1
19	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3039	117		B1

3	BOROUGH	NEIGHBORHOOD	BUILDING CLASS CATEGORY	TAX CLASS AT PRESENT	BLOCK	LOT	EASE- MENT	BUILDING CLASS AT PRESENT	
20	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3039	130		B1	
21	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3040	5		S2	4
22	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3041	7		S2	
23	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3044	167		B1	2
24	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3046	12		B9	
25	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3050	65		B2	
26	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3050	85		B1	2
27	2	BATHGATE	02 TWO FAMILY DWELLINGS	1	3053	22		S2	5
28	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	2908	34		C0	
29	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3030	168		C0	
30	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3036	33		C0	
31	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3037	130		C0	4:
32	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3039	33		C0	
33	2	BATHGATE	03 THREE FAMILY DWELLINGS	1	3039	33		C0	4
...	
7411	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7013	152		B2	

3	BOROUGH	NEIGHBORHOOD	BUILDING CLASS CATEGORY	TAX CLASS AT PRESENT	BLOCK	LOT	EASE- MENT	BUILDING CLASS AT PRESENT
7412	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7020	95		B2
7413	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7020	95		B2
7414	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7020	463		B2
7415	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7022	13		B2
7416	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7022	29		B2
7417	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7022	31		B2
7418	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7091	139		B9
7419	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7091	189		B9 B
7420	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7106	8		B9
7421	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7267	18		B9
7422	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7267	87		B9
7423	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7267	131		B9
7424	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7267	168		B9
7425	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7267	186		B9
7426	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7267	198		B9
7427	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7267	218		B2

3	BOROUGH	NEIGHBORHOOD	BUILDING CLASS CATEGORY	TAX CLASS AT PRESENT	BLOCK	LOT	EASE- MENT	BUILDING CLASS AT PRESENT
7428	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7267	227		B2
7429	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7267	245		B2
7430	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7267	387		B9
7431	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7316	11		B2
7432	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7316	61		B2
7433	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7316	84		B2
7434	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7316	89		B2
7435	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7317	128		B2
7436	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7349	3		B9
7437	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7349	53		B9
7438	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7349	71		B9
7439	5	WOODROW	02 TWO FAMILY DWELLINGS	1	7351	11		B2
7440	5	WOODROW	05 TAX CLASS 1 VACANT LAND	1B	6926	78		V0

88687 rows × 22 columns

```
In [761]: housing2014.columns = housing2014.columns.to_series().apply(lambda x: x.strip()  
())  
housing2014_df = housing2014[['ADDRESS', 'ZIP CODE', 'SALE PRICE', 'SALE DATE',  
'Borough_Name']]  
housing2014_df.dropna(inplace = True)  
housing2014_df["Year"] = "2014-15"  
housing2014_df
```

```
C:\me\Anaconda\lib\site-packages\ipykernel_launcher.py:3: SettingWithCopyWarning:  
  ing:
```

```
A value is trying to be set on a copy of a slice from a DataFrame
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy
```

```
This is separate from the ipykernel package so we can avoid doing imports until
```

```
C:\me\Anaconda\lib\site-packages\ipykernel_launcher.py:4: SettingWithCopyWarning:  
  ing:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.
```

```
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy
```

```
after removing the cwd from sys.path.
```

Out[761]:

3	ADDRESS	ZIP CODE	SALE PRICE	SALE DATE	Borough_Name	Year
4	418 EAST 182 STREET	10457	225000	2014-12-11 00:00:00	BRONX	2014-15
5	1979 WASHINGTON AVENUE	10457	580000	2014-02-07 00:00:00	BRONX	2014-15
6	467 EAST 185TH STREET	10458	225000	2014-09-29 00:00:00	BRONX	2014-15
7	2047 BATHGATE AVE	10457	10	2014-01-22 00:00:00	BRONX	2014-15
8	1538 WASHINGTON AVENUE	10457	297000	2014-06-13 00:00:00	BRONX	2014-15
9	484 CLAREMONT PARKWAY	10457	390000	2014-10-17 00:00:00	BRONX	2014-15
10	517 EAST 171 STREET	10457	0	2014-05-09 00:00:00	BRONX	2014-15
11	4469 PARK AVENUE	10457	300000	2014-03-28 00:00:00	BRONX	2014-15
12	1983 WASHINGTON AVENUE	10457	0	2014-07-15 00:00:00	BRONX	2014-15
13	2083 WASHINGTON AVENUE	10457	525200	2014-12-29 00:00:00	BRONX	2014-15
14	2083 WASHINGTON AVENUE	10457	240000	2014-08-04 00:00:00	BRONX	2014-15
15	459 EAST 179 STREET	10457	380000	2014-01-14 00:00:00	BRONX	2014-15
16	440 EAST 185TH ST	10458	360000	2014-04-21 00:00:00	BRONX	2014-15
17	440 EAST 185TH	10458	360000	2014-01-28 00:00:00	BRONX	2014-15
18	469B EAST 184 STREET	10458	1000	2014-09-08 00:00:00	BRONX	2014-15
19	448 EAST 185 STREET	10458	385000	2014-12-12 00:00:00	BRONX	2014-15
20	469A EAST 184TH STREET	10458	312000	2014-11-05 00:00:00	BRONX	2014-15
21	4654-4656 PARK AVENUE	10458	370000	2014-04-24 00:00:00	BRONX	2014-15
22	444 CYRUS PLACE	10458	0	2014-11-12 00:00:00	BRONX	2014-15
23	2008 BATHGATE AVENUE	10457	0	2014-11-17 00:00:00	BRONX	2014-15
24	2044 WASHINGTON AVENUE	10457	500	2014-09-08 00:00:00	BRONX	2014-15
25	508 EAST 183 STREET	10458	165000	2014-08-13 00:00:00	BRONX	2014-15
26	2241 BATHGATE AVENUE	10457	273000	2014-07-25 00:00:00	BRONX	2014-15

3	ADDRESS	ZIP CODE	SALE PRICE	SALE DATE	Borough_Name	Year
27	511 EAST 183RD STREET	10458	360000	2014-12-08 00:00:00	BRONX	2014-15
28	1835 WASHINGTON AVENUE	10457	0	2014-01-13 00:00:00	BRONX	2014-15
29	4509 PARK AVENUE	10457	0	2014-08-29 00:00:00	BRONX	2014-15
30	2091 WASHINGTON AVE	10457	0	2014-11-06 00:00:00	BRONX	2014-15
31	458 EAST 182ND STREET	10457	327600	2014-09-29 00:00:00	BRONX	2014-15
32	465 EAST 184 STREET	10458	400000	2014-04-10 00:00:00	BRONX	2014-15
33	465 EAST 184TH STREET	10458	0	2014-03-31 00:00:00	BRONX	2014-15
...
7411	141 MARISA CIRCLE	10309	456000	2014-09-04 00:00:00	STATEN IS	2014-15
7412	15 LYNBROOK COURT	10309	0	2014-09-13 00:00:00	STATEN IS	2014-15
7413	15 LYNBROOK COURT	10309	0	2014-04-30 00:00:00	STATEN IS	2014-15
7414	14 LYNBROOK COURT	10309	699000	2014-04-30 00:00:00	STATEN IS	2014-15
7415	34 BROOKSIDE LOOP	10309	0	2014-10-13 00:00:00	STATEN IS	2014-15
7416	61 BROOKSIDE LOOP	10309	763687	2014-10-22 00:00:00	STATEN IS	2014-15
7417	59 BROOKSIDE LOOP	10309	773854	2014-01-07 00:00:00	STATEN IS	2014-15
7418	1624 WOODROW ROAD	10309	0	2014-06-27 00:00:00	STATEN IS	2014-15
7419	630 BLOOMINGDALE ROAD	10309	0	2014-11-25 00:00:00	STATEN IS	2014-15
7420	204 WIRT AVENUE	10309	463000	2014-06-30 00:00:00	STATEN IS	2014-15
7421	30 WHITE OAK LANE	10309	472500	2014-10-28 00:00:00	STATEN IS	2014-15
7422	34 SHARROTT'S LANE	10309	423514	2014-01-27 00:00:00	STATEN IS	2014-15
7423	67 SHARROTT'S LANE	10309	545000	2014-07-23 00:00:00	STATEN IS	2014-15
7424	45 RED CEDAR LANE	10309	475000	2014-04-15 00:00:00	STATEN IS	2014-15
7425	20 RED CEDAR LANE	10309	460425	2014-05-15 00:00:00	STATEN IS	2014-15

3	ADDRESS	ZIP CODE	SALE PRICE	SALE DATE	Borough_Name	Year
7426	36 HEMLOCK LANE	10309	530000	2014-02-21 00:00:00	STATEN IS	2014-15
7427	33 WHITE OAK LANE	10309	445000	2014-08-19 00:00:00	STATEN IS	2014-15
7428	41 WHITE OAK LANE	10309	490000	2014-08-05 00:00:00	STATEN IS	2014-15
7429	17 SPRUCE LANE	10309	495000	2014-12-15 00:00:00	STATEN IS	2014-15
7430	20 SALAMANDER COURT	10309	465000	2014-10-13 00:00:00	STATEN IS	2014-15
7431	64 DARNELL LANE	10309	490000	2014-09-09 00:00:00	STATEN IS	2014-15
7432	178 DARNELL LANE	10309	475000	2014-08-07 00:00:00	STATEN IS	2014-15
7433	57 DARNELL LANE	10309	470000	2014-03-31 00:00:00	STATEN IS	2014-15
7434	131 DARNELL LANE	10309	465000	2014-05-05 00:00:00	STATEN IS	2014-15
7435	106 ROBIN COURT	10309	0	2014-04-09 00:00:00	STATEN IS	2014-15
7436	47 PHEASANT LANE	10309	610000	2014-09-11 00:00:00	STATEN IS	2014-15
7437	15 QUAIL LANE	10309	0	2014-02-04 00:00:00	STATEN IS	2014-15
7438	16 PHEASANT LANE	10309	474000	2014-06-06 00:00:00	STATEN IS	2014-15
7439	40 HERRICK AVENUE	10309	200000	2014-03-21 00:00:00	STATEN IS	2014-15
7440	SHELDON AVENUE	10309	0	2014-07-15 00:00:00	STATEN IS	2014-15

88687 rows × 6 columns

```
In [762]: bronx_2015 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_15 housing prices\\2015_bronx.xlsx")
brooklyn_2015 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_15 housing prices\\2015_brooklyn.xlsx")
manhattan_2015 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_15 housing prices\\2015_manhattan.xlsx")
queens_2015 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_15 housing prices\\2015_queens.xlsx")
statenisland_2015 = pd.read_excel("C:\\me\\data bootcamp\\final project\\2013_15 housing prices\\2015_statenisland.xlsx")
```

```
In [763]: bronx_2015.columns=bronx_2015.iloc[3]
bronx_2015 = bronx_2015[4:] #dont' rerun
bronx_2015["Borough_Name"] = "BRONX"

brooklyn_2015.columns=brooklyn_2015.iloc[3]
brooklyn_2015 = brooklyn_2015[4:] #dont' rerun
brooklyn_2015["Borough_Name"] = "BROOKLYN"

manhattan_2015.columns=manhattan_2015.iloc[3]
manhattan_2015 = manhattan_2015[4:] #dont' rerun
manhattan_2015["Borough_Name"] = "MANHATTAN"

queens_2015.columns=queens_2015.iloc[3]
queens_2015 = queens_2015[4:] #dont' rerun
queens_2015["Borough_Name"] = "QUEENS"

statenisland_2015.columns=statenisland_2015.iloc[3]
statenisland_2015 = statenisland_2015[4:] #dont' rerun
statenisland_2015["Borough_Name"] = "STATEN IS"

housing2015 = bronx_2015.append(brooklyn_2015)
housing2015 = housing2015.append(manhattan_2015)
housing2015 = housing2015.append(queens_2015)
housing2015 = housing2015.append(statenisland_2015)
housing2015

housing2015.columns = housing2015.columns.to_series().apply(lambda x: x.strip())
housing2015_df = housing2015[['ADDRESS', 'ZIP CODE', 'SALE PRICE', 'SALE DATE',
'Borough_Name']]
housing2015_df.dropna(inplace = True)
housing2015_df["Year"] = "2015-16"
housing2015_df
```



```
C:\me\Anaconda\lib\site-packages\ipykernel_launcher.py:29: SettingWithCopyWarning:
```

A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
C:\me\Anaconda\lib\site-packages\ipykernel_launcher.py:30: SettingWithCopyWarning:
```

A value is trying to be set on a copy of a slice from a DataFrame.

Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

Out[763]:

3	ADDRESS	ZIP CODE	SALE PRICE	SALE DATE	Borough_Name	Year
4	1665 WASHINGTON AVENUE	10457	0	2015-08-07 00:00:00	BRONX	2015-16
5	1665 WASHINGTON AVENUE	10457	0	2015-08-07 00:00:00	BRONX	2015-16
6	441 EAST 178 STREET	10457	0	2015-12-18 00:00:00	BRONX	2015-16
7	2329 WASHINGTON AVE	10458	300000	2015-02-20 00:00:00	BRONX	2015-16
8	455 EAST 184 STREET	10458	0	2015-08-18 00:00:00	BRONX	2015-16
9	455 EAST 184 STREET	10458	400000	2015-07-01 00:00:00	BRONX	2015-16
10	2368 WASHINGTON AVENUE	10458	510000	2015-01-06 00:00:00	BRONX	2015-16
11	1540 WASHINGTON AVENUE	10457	0	2015-12-12 00:00:00	BRONX	2015-16
12	3823 3 AVENUE	10457	430000	2015-08-04 00:00:00	BRONX	2015-16
13	517 EAST 171 STREET	10457	10000	2015-06-02 00:00:00	BRONX	2015-16
14	3862 3RD AVENUE	10457	389000	2015-01-29 00:00:00	BRONX	2015-16
15	1557 FULTON AVENUE	10457	505000	2015-08-05 00:00:00	BRONX	2015-16
16	1557 FULTON AVENUE	10457	250000	2015-02-19 00:00:00	BRONX	2015-16
17	2264 WEBSTER AVENUE	10457	490000	2015-09-28 00:00:00	BRONX	2015-16
18	4465 PARK AVE	10457	1000	2015-12-28 00:00:00	BRONX	2015-16
19	4465 PARK AVE	10457	1000	2015-10-20 00:00:00	BRONX	2015-16
20	4461 PARK AVENUE	10457	100000	2015-08-20 00:00:00	BRONX	2015-16
21	1983 WASHINGTON AVENUE	10457	228000	2015-11-16 00:00:00	BRONX	2015-16
22	4354 PARK AVENUE	10457	128778	2015-05-29 00:00:00	BRONX	2015-16
23	465 EAST 179 STREET	10457	420000	2015-04-23 00:00:00	BRONX	2015-16
24	469B EAST 184 STREET	10458	0	2015-10-13 00:00:00	BRONX	2015-16
25	444 EAST 186 STREET	10458	500	2015-02-23 00:00:00	BRONX	2015-16
26	2034 BATHGATE AVENUE	10457	0	2015-09-09 00:00:00	BRONX	2015-16

3	ADDRESS	ZIP CODE	SALE PRICE	SALE DATE	Borough_Name	Year
27	2122 WASHINGTON AVENUE	10457	10	2015-01-19 00:00:00	BRONX	2015- 16
28	508 EAST 183 STREET	10458	449000	2015-06-08 00:00:00	BRONX	2015- 16
29	2238 BASSFORD AVENUE	10457	261000	2015-09-25 00:00:00	BRONX	2015- 16
30	509 EAST 183 STREET	10458	393600	2015-04-01 00:00:00	BRONX	2015- 16
31	2315 BASSFORD AVENUE	10458	232215	2015-08-18 00:00:00	BRONX	2015- 16
32	459 EAST 181 STREET	10457	100	2015-03-16 00:00:00	BRONX	2015- 16
33	468 E 183 STREET	10458	435919	2015-05-29 00:00:00	BRONX	2015- 16
...
7979	41 LENEVAR AVENUE	10309	635000	2015-09-08 00:00:00	STATEN IS	2015- 16
7980	985 RATHBUN AVENUE	10309	468000	2015-12-15 00:00:00	STATEN IS	2015- 16
7981	1013 RATHBUN AVENUE	10309	0	2015-10-13 00:00:00	STATEN IS	2015- 16
7982	449 BLOOMINGDALE ROAD	10309	580000	2015-05-19 00:00:00	STATEN IS	2015- 16
7983	463 BLOOMINGDALE ROAD	10309	0	2015-04-14 00:00:00	STATEN IS	2015- 16
7984	48 MARISA CIRCLE	10309	0	2015-09-09 00:00:00	STATEN IS	2015- 16
7985	48 MARISA CIRCLE	10309	480000	2015-07-29 00:00:00	STATEN IS	2015- 16
7986	72 MARISA CIRCLE	10309	0	2015-09-29 00:00:00	STATEN IS	2015- 16
7987	563 BLOOMINGDALE ROAD	10309	460000	2015-03-25 00:00:00	STATEN IS	2015- 16
7988	63 DEXTER AVENUE	10309	940000	2015-09-16 00:00:00	STATEN IS	2015- 16
7989	133 ANTHONY STREET	10309	0	2015-01-06 00:00:00	STATEN IS	2015- 16
7990	44 BROOKSIDE LOOP	10309	800000	2015-06-01 00:00:00	STATEN IS	2015- 16
7991	44 BROOKSIDE LOOP	10309	0	2015-05-27 00:00:00	STATEN IS	2015- 16
7992	63 BROOKSIDE LOOP	10309	763687	2015-03-30 00:00:00	STATEN IS	2015- 16
7993	49 BROOKSIDE LOOP	10309	0	2015-11-13 00:00:00	STATEN IS	2015- 16

3	ADDRESS	ZIP CODE	SALE PRICE	SALE DATE	Borough_Name	Year
7994	47 CLAY PIT ROAD	10309	485000	2015-02-03 00:00:00	STATEN IS	2015-16
7995	15 PLUMTREE LANE	10309	482500	2015-12-14 00:00:00	STATEN IS	2015-16
7996	200 WIRT AVENUE	10309	465000	2015-11-10 00:00:00	STATEN IS	2015-16
7997	40 SHARROTT'S LANE	10309	515000	2015-11-19 00:00:00	STATEN IS	2015-16
7998	57 SHARROTT'S LANE	10309	0	2015-06-15 00:00:00	STATEN IS	2015-16
7999	56 CLAY PIT ROAD	10309	165000	2015-12-11 00:00:00	STATEN IS	2015-16
8000	26 HEMLOCK LANE	10309	0	2015-08-10 00:00:00	STATEN IS	2015-16
8001	27 WHITE OAK LANE	10309	485000	2015-02-13 00:00:00	STATEN IS	2015-16
8002	54 DARNELL LANE	10309	495000	2015-10-09 00:00:00	STATEN IS	2015-16
8003	125 DARNELL LANE	10309	0	2015-03-16 00:00:00	STATEN IS	2015-16
8004	94 ROBIN COURT	10309	0	2015-02-19 00:00:00	STATEN IS	2015-16
8005	476 HARGOLD AVENUE	10309	825000	2015-09-17 00:00:00	STATEN IS	2015-16
8006	20 QUAIL LANE	10309	490000	2015-01-15 00:00:00	STATEN IS	2015-16
8007	22 PHEASANT LANE	10309	10	2015-11-06 00:00:00	STATEN IS	2015-16
8008	N/A SHELDON AVENUE	10309	0	2015-10-07 00:00:00	STATEN IS	2015-16

91446 rows × 6 columns

Part 3: Merging the data

This part of the project, we will begin to construct a new dataframe by merging the previous two parts of the project. But before that, we will append the housing data frames together and then group them by year to get a comprehensive data frame for housing data.

```
In [768]: housingdf = housing2013_df.append(housing2014_df)
housingdf = housingdf.append(housing2015_df)
housingdf = housingdf.append(housing2016_df)
housingdf
```

Out[768]:

	ADDRESS	Borough_Name	SALE DATE	SALE PRICE	Year	ZIP CODE
4	412 EAST 179TH STREET	BRONX	2013-07-08 00:00:00	355000	2013-14	10457
5	2329 WASHINGTON AVENUE	BRONX	2013-11-14 00:00:00	120000	2013-14	10458
6	2329 WASHINGTON AVENUE	BRONX	2013-05-20 00:00:00	474819	2013-14	10458
7	2075 BATHGATE AVENUE	BRONX	2013-03-12 00:00:00	210000	2013-14	10457
8	2047 BATHGATE AVENUE	BRONX	2013-07-01 00:00:00	343116	2013-14	10457
9	4411 3 AVENUE	BRONX	2013-08-15 00:00:00	225000	2013-14	10457
10	1467 CROTONA PLACE	BRONX	2013-11-04 00:00:00	380000	2013-14	10456
11	3860 3 AVENUE	BRONX	2013-06-24 00:00:00	390000	2013-14	10457
12	4469 PARK AVENUE	BRONX	2013-06-27 00:00:00	207000	2013-14	10457
13	454 EAST 179 STREET	BRONX	2013-07-01 00:00:00	0	2013-14	10457
14	4354 PARK AVENUE	BRONX	2013-06-26 00:00:00	328000	2013-14	10457
15	465 EAST 185 STREET	BRONX	2013-03-07 00:00:00	369000	2013-14	10458
16	4654-4656 PARK AVENUE	BRONX	2013-02-28 00:00:00	308000	2013-14	10458
17	2043 BATHGATE AVENUE	BRONX	2013-08-27 00:00:00	0	2013-14	10457
18	2043 BATHGATE AVENUE	BRONX	2013-01-09 00:00:00	0	2013-14	10457
19	4557 3 AVENUE	BRONX	2013-05-14 00:00:00	0	2013-14	10458
20	2091 WASHINGTON AVENUE	BRONX	2013-06-27 00:00:00	0	2013-14	10457
21	4418 PARK AVENUE	BRONX	2013-02-15 00:00:00	440000	2013-14	10457
22	468 EAST 183 STREET	BRONX	2013-01-15 00:00:00	662115	2013-14	10458
23	452 CYRUS PLACE	BRONX	2013-07-25 00:00:00	0	2013-14	10458
24	2092 WASHINGTON AVENUE	BRONX	2013-12-19 00:00:00	510000	2013-14	10457
25	2186 WASHINGTON AVENUE	BRONX	2013-05-10 00:00:00	460000	2013-14	10457
26	2085 WASHINGTON AVENUE	BRONX	2013-01-09 00:00:00	40730	2013-14	10457

	ADDRESS	Borough_Name	SALE DATE	SALE PRICE	Year	ZIP CODE
27	442 EAST 176 STREET	BRONX	2013-06-26 00:00:00	0	2013-14	10457
28	446 EAST 176 STREET	BRONX	2013-06-26 00:00:00	0	2013-14	10457
29	4196 PARK AVENUE	BRONX	2013-06-26 00:00:00	0	2013-14	10457
30	1824 WASHINGTON AVENUE	BRONX	2013-06-26 00:00:00	0	2013-14	10457
31	4173 3 AVENUE	BRONX	2013-08-15 00:00:00	50000	2013-14	10457
32	3758 3 AVENUE	BRONX	2013-06-27 00:00:00	0	2013-14	10456
33	538 CLAREMONT PARKWAY	BRONX	2013-06-27 00:00:00	0	2013-14	10456
...
8384	1452 WOODROW ROAD	STATEN IS	2016-09-05 00:00:00	875000	2016-17	10309
8385	574 MAGUIRE AVENUE	STATEN IS	2017-05-22 00:00:00	875000	2016-17	10309
8386	72 DEXTER AVENUE	STATEN IS	2016-12-07 00:00:00	840000	2016-17	10309
8387	56 DEXTER AVENUE	STATEN IS	2017-03-30 00:00:00	772500	2016-17	10309
8388	35 LYNBROOK AVENUE	STATEN IS	2017-01-04 00:00:00	740000	2016-17	10309
8389	27 CLAY PIT ROAD	STATEN IS	2017-03-17 00:00:00	-	2016-17	10309
8390	51 CLAY PIT ROAD	STATEN IS	2017-07-06 00:00:00	-	2016-17	10309
8391	1576 WOODROW ROAD	STATEN IS	2016-10-14 00:00:00	-	2016-17	10309
8392	96 CRABTREE AVENUE	STATEN IS	2017-03-01 00:00:00	590000	2016-17	10309
8393	98 CRABTREE AVENUE	STATEN IS	2017-04-26 00:00:00	580000	2016-17	10309
8394	10 TRINA LANE	STATEN IS	2017-06-20 00:00:00	550000	2016-17	10309
8395	118 TURNER STREET	STATEN IS	2016-11-16 00:00:00	545000	2016-17	10309
8396	125 CLAY PIT ROAD	STATEN IS	2017-01-09 00:00:00	800000	2016-17	10309
8397	65 SHARROTT'S LANE	STATEN IS	2017-06-22 00:00:00	642500	2016-17	10309
8398	67 SHARROTT'S LANE	STATEN IS	2016-09-26 00:00:00	408092	2016-17	10309

	ADDRESS	Borough_Name	SALE DATE	SALE PRICE	Year	ZIP CODE
8399	27 RED CEDAR LANE	STATEN IS	2017-02-27 00:00:00	600000	2016-17	10309
8400	30 HEMLOCK LANE	STATEN IS	2016-12-05 00:00:00	610000	2016-17	10309
8401	33 WHITE OAK LANE	STATEN IS	2017-08-04 00:00:00	645000	2016-17	10309
8402	19 SPRUCE LANE	STATEN IS	2017-01-19 00:00:00	168000	2016-17	10309
8403	48 WHITE OAK LANE	STATEN IS	2016-10-18 00:00:00	625000	2016-17	10309
8404	178 DARNELL LANE	STATEN IS	2017-06-30 00:00:00	-	2016-17	10309
8405	137 DARNELL LANE	STATEN IS	2016-12-30 00:00:00	-	2016-17	10309
8406	125 DARNELL LANE	STATEN IS	2016-10-31 00:00:00	509000	2016-17	10309
8407	112 ROBIN COURT	STATEN IS	2016-12-07 00:00:00	648000	2016-17	10309
8408	41 SONIA COURT	STATEN IS	2016-12-01 00:00:00	-	2016-17	10309
8409	37 QUAIL LANE	STATEN IS	2016-11-28 00:00:00	450000	2016-17	10309
8410	32 PHEASANT LANE	STATEN IS	2017-04-21 00:00:00	550000	2016-17	10309
8411	49 PITNEY AVENUE	STATEN IS	2017-07-05 00:00:00	460000	2016-17	10309
8412	2730 ARTHUR KILL ROAD	STATEN IS	2016-12-21 00:00:00	11693337	2016-17	10309
8413	155 CLAY PIT ROAD	STATEN IS	2016-10-27 00:00:00	69300	2016-17	10309

354655 rows × 6 columns

Now we will grab data for year 2013.

```
In [769]: housingdf = housingdf.rename(columns = {'ZIP CODE':'ZIP_CODE', "SALE PRICE":"SALE_PRICE"})
housingdf['Year'] = housingdf['Year'].astype(str)
```



```
In [770]: housing13 = housingdf[housingdf.Year == "2013-14"]
housing13 = housing13.drop(housing13[housing13['SALE_PRICE'].isin([0,1,' ',10
])].index)
housing13 = housing13[~housing13['SALE_PRICE'].isnull()]
housing13
```

Out[770]:

	ADDRESS	Borough_Name	SALE DATE	SALE_PRICE	Year	ZIP_CODE
4	412 EAST 179TH STREET	BRONX	2013-07-08 00:00:00	355000	2013-14	10457
5	2329 WASHINGTON AVENUE	BRONX	2013-11-14 00:00:00	120000	2013-14	10458
39	2252 BASSFORD AVENUE	BRONX	2013-08-05 00:00:00	4429	2013-14	10457
70	490 EAST TREMONT AVENUE	BRONX	2013-05-10 00:00:00	370000	2013-14	10457
73	400 EAST FORDHAM ROAD	BRONX	2013-11-13 00:00:00	47640040	2013-14	10458
74	400 EAST FORDHAM ROAD	BRONX	2013-11-13 00:00:00	86260000	2013-14	10458
77	1156 E. 221 STREET	BRONX	2013-04-25 00:00:00	367000	2013-14	10469
90	3305 FENTON AVENUE	BRONX	2013-09-26 00:00:00	185000	2013-14	10469
94	3301 EASTCHESTER ROAD	BRONX	2013-04-18 00:00:00	375000	2013-14	10469
103	3318 TIEMANN AVENUE	BRONX	2013-12-13 00:00:00	375000	2013-14	10469
109	3133 MICKLE AVENUE	BRONX	2013-11-15 00:00:00	350000	2013-14	10469
117	2837 GUNTHER AVENUE	BRONX	2013-11-22 00:00:00	270500	2013-14	10469
118	2837 GUNTHER AVENUE	BRONX	2013-07-30 00:00:00	205000	2013-14	10469
128	1117 EAST 222ND STREET	BRONX	2013-06-05 00:00:00	200000	2013-14	10469
131	1164 EAST 225TH STREET	BRONX	2013-07-12 00:00:00	222700	2013-14	10466
134	3621 PALMER AVENUE	BRONX	2013-02-13 00:00:00	337000	2013-14	10466
141	2121 LIGHT STREET	BRONX	2013-07-10 00:00:00	390000	2013-14	10466
199	3336 PEARSALL AVENUE	BRONX	2013-09-23 00:00:00	476500	2013-14	10469
212	3229 BRUNER AVENUE	BRONX	2013-08-30 00:00:00	280000	2013-14	10469
221	3048 BRUNER AVENUE	BRONX	2013-07-01 00:00:00	373000	2013-14	10469
222	3149 NEW ENGLAND THRUWAY	BRONX	2013-09-05 00:00:00	345000	2013-14	10469
227	2823 TIEMANN	BRONX	2013-03-11 00:00:00	350000	2013-14	10469
234	2833B GUNTHER AVENUE	BRONX	2013-01-09 00:00:00	300000	2013-14	10469

	ADDRESS	Borough_Name	SALE DATE	SALE_PRICE	Year	ZIP_CODE
236	2901 GRACE AVENUE	BRONX	2013-06-12 00:00:00	430000	2013-14	10469
244	3424 ELY AVENUE	BRONX	2013-12-26 00:00:00	299000	2013-14	10469
248	3530 EDSON AVE	BRONX	2013-11-27 00:00:00	499000	2013-14	10466
257	3648 PALMER AVENUE	BRONX	2013-11-15 00:00:00	199000	2013-14	10466
258	3879 BOSTON ROAD	BRONX	2013-10-07 00:00:00	1750000	2013-14	10466
262	3921 GRACE AVENUE	BRONX	2013-11-25 00:00:00	18000	2013-14	10466
274	3925 HILL AVENUE	BRONX	2013-10-09 00:00:00	459900	2013-14	10466
...
6959	1663 WOODROW ROAD	STATEN IS	2013-04-17 00:00:00	440000	2013-14	10309
6960	41 PLUMTREE LANE	STATEN IS	2013-06-24 00:00:00	405499	2013-14	10309
6962	27 PLUMTREE LANE	STATEN IS	2013-09-19 00:00:00	461370	2013-14	10309
6963	23 PLUMTREE LANE	STATEN IS	2013-08-13 00:00:00	446059	2013-14	10309
6966	55 PLUMTREE LANE	STATEN IS	2013-10-15 00:00:00	476124	2013-14	10309
6973	58 PLUMTREE LANE	STATEN IS	2013-09-30 00:00:00	478507	2013-14	10309
6975	20 PLUMTREE LANE	STATEN IS	2013-09-25 00:00:00	446059	2013-14	10309
6979	844 BLOOMINGDALE ROAD	STATEN IS	2013-08-08 00:00:00	312000	2013-14	10309
6980	146 LORRAINE LOOP	STATEN IS	2013-11-14 00:00:00	343000	2013-14	10309
6988	24 JESSICA LANE	STATEN IS	2013-09-26 00:00:00	459000	2013-14	10309
6993	46 SHARROTT'S ROAD	STATEN IS	2013-07-02 00:00:00	450000	2013-14	10309
6996	157 PEMBROOK LOOP	STATEN IS	2013-02-03 00:00:00	387000	2013-14	10309
6998	30 MALLARD LANE	STATEN IS	2013-08-28 00:00:00	379000	2013-14	10309
6999	30 MALLARD LANE	STATEN IS	2013-05-13 00:00:00	40000	2013-14	10309
7007	42 SONIA COURT	STATEN IS	2013-04-15 00:00:00	450000	2013-14	10309
7013	1002 IONIA AVENUE	STATEN IS	2013-10-03 00:00:00	865512	2013-14	10309

	ADDRESS	Borough_Name	SALE DATE	SALE_PRICE	Year	ZIP_CODE
7020	951 RATHBUN AVENUE	STATEN IS	2013-02-22 00:00:00	470000	2013-14	10309
7022	1047 RATHBUN AVENUE	STATEN IS	2013-09-20 00:00:00	410000	2013-14	10309
7025	996 STAFFORD AVENUE	STATEN IS	2013-03-26 00:00:00	995000	2013-14	10309
7028	59 MARISA CIRCLE	STATEN IS	2013-08-13 00:00:00	450000	2013-14	10309
7035	49 BROOKSIDE LOOP	STATEN IS	2013-11-26 00:00:00	814598	2013-14	10309
7036	47 BROOKSIDE LOOP	STATEN IS	2013-04-15 00:00:00	687318	2013-14	10309
7046	96 CRABTREE AVENUE	STATEN IS	2013-03-13 00:00:00	481549	2013-14	10309
7048	98 CRABTREE AVENUE	STATEN IS	2013-06-28 00:00:00	434501	2013-14	10309
7051	17 PLUMTREE LANE	STATEN IS	2013-09-04 00:00:00	464990	2013-14	10309
7054	21 HEMLOCK LANE	STATEN IS	2013-07-02 00:00:00	485000	2013-14	10309
7055	25 WHITE OAK LANE	STATEN IS	2013-05-15 00:00:00	166666	2013-14	10309
7060	110 SHARROTT'S ROAD	STATEN IS	2013-04-18 00:00:00	560000	2013-14	10309
7061	59 ROBIN COURT	STATEN IS	2013-09-27 00:00:00	498400	2013-14	10309
7063	64 ROBIN COURT	STATEN IS	2013-10-03 00:00:00	482500	2013-14	10309

23282 rows × 6 columns

Here, we tried to merge the data to combine the two data sets for one of the boroughs based on Zip code. We realized that we couldn't do this because the dataframe for school safety had significantly less rows than the dataframe for housing so this resulted in a lot of duplicated rows for the safety data.

```
In [771]: #combo = pd.merge(bronxhousing, bronx,
#                   on = ["Year", "Borough_Name", "ZIP_CODE"],
#                   how = "right",
#                   indicator = True)
#combo.dropna(inplace=True)
#combo.drop_duplicates('ADDRESS_y')
#combo
```

We are going to try and find how many unique zip codes there are for each data frame. Since the housing dataframe has more zip codes than in the school safety data, we will use an `isin` function to only keep the housing data that have the same zip codes as the data in the school safety dataframe. Once we have the same number of unique zip codes, we can try and create new data frames with the average values for Number of Crimes, Average Number of Crimes and Sale Price for each zip code.

```
In [772]: housing13['ZIP_CODE'] = housing13.ZIP_CODE.astype(int)
housing13.ZIP_CODE.value_counts()
```

```
Out[772]: 10019 1591
          10025 719
          11375 517
          11201 503
          10024 482
          10022 460
          11354 423
          10021 389
          10023 368
          11215 356
          11377 327
          11238 311
          11355 310
          11207 307
          11235 306
          11385 301
          10016 300
          10065 283
          11374 282
          10128 279
          10028 275
          11229 254
          11367 254
          10075 254
          11233 252
          11223 239
          11234 235
          11364 231
          11230 229
          11208 216
          ...
          10030 22
          10009 22
          10018 21
          10455 21
          10004 21
          10459 20
          10302 20
          11001 17
          11363 17
          10457 17
          10006 15
          11692 15
          11693 14
          10470 14
          10474 14
          10468 13
          10452 13
          10451 13
          10044 13
          10464 12
          10037 11
          10039 11
          11109 9
          0 8
          10453 7
          11040 6
```

10475	6
10454	6
10129	5
10803	1

Name: ZIP_CODE, Length: 182, dtype: int64


```
In [773]: safety_2013['ZIP_CODE'] = safety_2013['ZIP_CODE'].astype(int)
safety_2013.ZIP_CODE.value_counts()
```

```
C:\me\Anaconda\lib\site-packages\ipykernel_launcher.py:1: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy
```

```
"""Entry point for launching an IPython kernel.
```

```
Out[773]: 10457 15
          11207 14
          11385 13
          10458 12
          10301 10
          10314 10
          10002 10
          11212 10
          11203 10
          11235 9
          11208 9
          10029 9
          11236 9
          11204 9
          11368 9
          11373 9
          11432 9
          11221 9
          11206 8
          10468 8
          11101 8
          10455 8
          11367 8
          10463 8
          11364 8
          11365 8
          10452 8
          11434 8
          11214 8
          11223 8
          ..
          10030 2
          10028 2
          10026 2
          10021 2
          10016 2
          11436 2
          10013 2
          10007 2
          10006 2
          11694 2
          11693 2
          10309 2
          10470 2
          10454 2
          11356 2
          11249 1
          11415 1
          11429 1
          10464 1
          10465 1
          10471 1
          10474 1
          11239 1
          10014 1
          10037 1
          10044 1
```

```
10075    1
11109    1
11213    1
10282    1
```

```
Name: ZIP_CODE, Length: 169, dtype: int64
```

```
In [774]: safetycode = safety_2013['ZIP_CODE'].values
```

```
In [775]: housing13DF = housing13[housing13['ZIP_CODE'].isin(safety_2013["ZIP_CODE"])]  
housing13DF
```

Out[775]:

	ADDRESS	Borough_Name	SALE DATE	SALE_PRICE	Year	ZIP_CODE
4	412 EAST 179TH STREET	BRONX	2013-07-08 00:00:00	355000	2013-14	10457
5	2329 WASHINGTON AVENUE	BRONX	2013-11-14 00:00:00	120000	2013-14	10458
39	2252 BASSFORD AVENUE	BRONX	2013-08-05 00:00:00	4429	2013-14	10457
70	490 EAST TREMONT AVENUE	BRONX	2013-05-10 00:00:00	370000	2013-14	10457
73	400 EAST FORDHAM ROAD	BRONX	2013-11-13 00:00:00	47640040	2013-14	10458
74	400 EAST FORDHAM ROAD	BRONX	2013-11-13 00:00:00	86260000	2013-14	10458
77	1156 E. 221 STREET	BRONX	2013-04-25 00:00:00	367000	2013-14	10469
90	3305 FENTON AVENUE	BRONX	2013-09-26 00:00:00	185000	2013-14	10469
94	3301 EASTCHESTER ROAD	BRONX	2013-04-18 00:00:00	375000	2013-14	10469
103	3318 TIEMANN AVENUE	BRONX	2013-12-13 00:00:00	375000	2013-14	10469
109	3133 MICKLE AVENUE	BRONX	2013-11-15 00:00:00	350000	2013-14	10469
117	2837 GUNTHER AVENUE	BRONX	2013-11-22 00:00:00	270500	2013-14	10469
118	2837 GUNTHER AVENUE	BRONX	2013-07-30 00:00:00	205000	2013-14	10469
128	1117 EAST 222ND STREET	BRONX	2013-06-05 00:00:00	200000	2013-14	10469
131	1164 EAST 225TH STREET	BRONX	2013-07-12 00:00:00	222700	2013-14	10466
134	3621 PALMER AVENUE	BRONX	2013-02-13 00:00:00	337000	2013-14	10466
141	2121 LIGHT STREET	BRONX	2013-07-10 00:00:00	390000	2013-14	10466
199	3336 PEARSALL AVENUE	BRONX	2013-09-23 00:00:00	476500	2013-14	10469
212	3229 BRUNER AVENUE	BRONX	2013-08-30 00:00:00	280000	2013-14	10469
221	3048 BRUNER AVENUE	BRONX	2013-07-01 00:00:00	373000	2013-14	10469
222	3149 NEW ENGLAND THRUWAY	BRONX	2013-09-05 00:00:00	345000	2013-14	10469
227	2823 TIEMANN	BRONX	2013-03-11 00:00:00	350000	2013-14	10469
234	2833B GUNTHER AVENUE	BRONX	2013-01-09 00:00:00	300000	2013-14	10469

	ADDRESS	Borough_Name	SALE DATE	SALE_PRICE	Year	ZIP_CODE
236	2901 GRACE AVENUE	BRONX	2013-06-12 00:00:00	430000	2013-14	10469
244	3424 ELY AVENUE	BRONX	2013-12-26 00:00:00	299000	2013-14	10469
248	3530 EDSON AVE	BRONX	2013-11-27 00:00:00	499000	2013-14	10466
257	3648 PALMER AVENUE	BRONX	2013-11-15 00:00:00	199000	2013-14	10466
258	3879 BOSTON ROAD	BRONX	2013-10-07 00:00:00	1750000	2013-14	10466
262	3921 GRACE AVENUE	BRONX	2013-11-25 00:00:00	18000	2013-14	10466
274	3925 HILL AVENUE	BRONX	2013-10-09 00:00:00	459900	2013-14	10466
...
6959	1663 WOODROW ROAD	STATEN IS	2013-04-17 00:00:00	440000	2013-14	10309
6960	41 PLUMTREE LANE	STATEN IS	2013-06-24 00:00:00	405499	2013-14	10309
6962	27 PLUMTREE LANE	STATEN IS	2013-09-19 00:00:00	461370	2013-14	10309
6963	23 PLUMTREE LANE	STATEN IS	2013-08-13 00:00:00	446059	2013-14	10309
6966	55 PLUMTREE LANE	STATEN IS	2013-10-15 00:00:00	476124	2013-14	10309
6973	58 PLUMTREE LANE	STATEN IS	2013-09-30 00:00:00	478507	2013-14	10309
6975	20 PLUMTREE LANE	STATEN IS	2013-09-25 00:00:00	446059	2013-14	10309
6979	844 BLOOMINGDALE ROAD	STATEN IS	2013-08-08 00:00:00	312000	2013-14	10309
6980	146 LORRAINE LOOP	STATEN IS	2013-11-14 00:00:00	343000	2013-14	10309
6988	24 JESSICA LANE	STATEN IS	2013-09-26 00:00:00	459000	2013-14	10309
6993	46 SHARROTT'S ROAD	STATEN IS	2013-07-02 00:00:00	450000	2013-14	10309
6996	157 PEMBROOK LOOP	STATEN IS	2013-02-03 00:00:00	387000	2013-14	10309
6998	30 MALLARD LANE	STATEN IS	2013-08-28 00:00:00	379000	2013-14	10309
6999	30 MALLARD LANE	STATEN IS	2013-05-13 00:00:00	40000	2013-14	10309
7007	42 SONIA COURT	STATEN IS	2013-04-15 00:00:00	450000	2013-14	10309
7013	1002 IONIA AVENUE	STATEN IS	2013-10-03 00:00:00	865512	2013-14	10309

	ADDRESS	Borough_Name	SALE DATE	SALE_PRICE	Year	ZIP_CODE
7020	951 RATHBUN AVENUE	STATEN IS	2013-02-22 00:00:00	470000	2013-14	10309
7022	1047 RATHBUN AVENUE	STATEN IS	2013-09-20 00:00:00	410000	2013-14	10309
7025	996 STAFFORD AVENUE	STATEN IS	2013-03-26 00:00:00	995000	2013-14	10309
7028	59 MARISA CIRCLE	STATEN IS	2013-08-13 00:00:00	450000	2013-14	10309
7035	49 BROOKSIDE LOOP	STATEN IS	2013-11-26 00:00:00	814598	2013-14	10309
7036	47 BROOKSIDE LOOP	STATEN IS	2013-04-15 00:00:00	687318	2013-14	10309
7046	96 CRABTREE AVENUE	STATEN IS	2013-03-13 00:00:00	481549	2013-14	10309
7048	98 CRABTREE AVENUE	STATEN IS	2013-06-28 00:00:00	434501	2013-14	10309
7051	17 PLUMTREE LANE	STATEN IS	2013-09-04 00:00:00	464990	2013-14	10309
7054	21 HEMLOCK LANE	STATEN IS	2013-07-02 00:00:00	485000	2013-14	10309
7055	25 WHITE OAK LANE	STATEN IS	2013-05-15 00:00:00	166666	2013-14	10309
7060	110 SHARROTT'S ROAD	STATEN IS	2013-04-18 00:00:00	560000	2013-14	10309
7061	59 ROBIN COURT	STATEN IS	2013-09-27 00:00:00	498400	2013-14	10309
7063	64 ROBIN COURT	STATEN IS	2013-10-03 00:00:00	482500	2013-14	10309

22132 rows × 6 columns


```
In [776]: housing13DF["SALE_PRICE"] = housing13DF.SALE_PRICE.astype(float)

#year_dict = {"SALE_PRICE": "first"}
year13 = housing13DF.groupby(["Year", "Borough_Name", "ZIP_CODE"]).mean()
#year13 = year13.agg(year_dict)
year13["SALE_PRICE"] = year13.SALE_PRICE.astype(int)
year13
```

```
C:\me\Anaconda\lib\site-packages\ipykernel_launcher.py:1: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy
```

```
"""Entry point for launching an IPython kernel.
```

Out[776]:

SALE_PRICE			
Year	Borough_Name	ZIP_CODE	
2013-14	BRONX	10451	1365265
		10452	2217611
		10453	349115
		10454	282433
		10455	458322
		10456	2290717
		10457	1416273
		10458	7197350
		10459	1448325
		10460	663833
		10461	749694
		10462	605826
		10463	327278
		10464	549060
		10465	436606
		10466	916320
		10467	739884
		10468	961225
		10469	424092
		10470	447642
		10471	455364
		10472	699696
		10473	289214
		10474	1921813
		10475	755166
	BROOKLYN	11201	2114862
		11203	556248
		11204	900740
		11205	870965
		11206	1009415

	QUEENS	11419	366850
		11420	363263
		11421	468431

SALE_PRICE

Year	Borough_Name	ZIP_CODE	
		11422	625858
		11423	284569
		11426	340599
		11427	373299
		11428	362388
		11429	378807
		11432	590797
		11433	317807
		11434	318328
		11435	390463
		11436	283596
		11691	394026
		11692	421684
		11693	134169
		11694	461443
	STATEN IS	10301	368753
		10302	403625
		10303	840461
		10304	404497
		10305	356664
		10306	567699
		10307	537516
		10308	410269
		10309	504953
		10310	338469
		10312	424609
		10314	382126

169 rows × 1 columns

```
In [777]: year13.reset_index()[year13.reset_index().ZIP_CODE == 11201]
```

```
Out[777]:
```

	Year	Borough_Name	ZIP_CODE	SALE_PRICE
25	2013-14	BROOKLYN	11201	2114862

```
In [778]: housingcode = housing13['ZIP_CODE'].values
safety13df = safety_2013[safety_2013['ZIP_CODE'].isin(housing13["ZIP_CODE"])]
```

```
In [779]: safety13df = safety13df.rename(columns = {"Total Number of Crimes":"Total_Crime", "Average Num of Crimes":"Average_Crime"})

year_dict2 = {"Total_Crime":"first", "Average_Crime":"first", "Standard_deviation":"first"}
year13_2 = safety13df.groupby(["Year", "Borough_Name", "ZIP_CODE"]).mean()
year13_2
```

Out[779]:

			Total_Crime	Average_Crime	Standard_deviation
Year	Borough_Name	ZIP_CODE			
2013-14	BRONX	10451	3.600000	5.190000	-1.590000
		10452	2.000000	5.725000	-3.725000
		10453	1.666667	5.446667	-3.780000
		10454	3.500000	4.830000	-1.330000
		10455	1.750000	4.412500	-2.662500
		10456	4.142857	5.647143	-1.504286
		10457	4.400000	5.338000	-0.938000
		10458	2.833333	6.568333	-3.735000
		10459	4.857143	5.087143	-0.230000
		10460	13.400000	4.390000	9.010000
		10461	8.200000	6.786000	1.414000
		10462	4.500000	9.910000	-5.410000
		10463	4.125000	6.692500	-2.567500
		10464	2.000000	3.730000	-1.730000
		10465	2.000000	6.630000	-4.630000
		10466	2.571429	5.358571	-2.787143
		10467	1.500000	6.673333	-5.173333
		10468	2.625000	10.055000	-7.430000
		10469	5.200000	8.490000	-3.290000
		10470	2.000000	4.830000	-2.830000
		10471	2.000000	4.830000	-2.830000
		10472	3.250000	7.330000	-4.080000
		10473	2.000000	5.550000	-3.550000
		10474	6.000000	6.630000	-0.630000
		10475	4.250000	5.730000	-1.480000
	BROOKLYN	11201	4.000000	5.483333	-1.483333
		11203	3.500000	5.390000	-1.890000
		11204	2.666667	6.685556	-4.018889
		11205	3.250000	5.005000	-1.755000
		11206	4.875000	6.227500	-1.352500

	QUEENS	11419	2.000000	5.430000	-3.430000
		11420	0.500000	8.013333	-7.513333
		11421	0.000000	7.230000	-7.230000

		Total_Crime	Average_Crime	Standard_deviation
Year	Borough_Name	ZIP_CODE		
		11422	2.500000	5.095000
		11423	3.500000	7.185000
		11426	4.000000	5.580000
		11427	17.000000	14.856667
		11428	5.000000	7.563333
		11429	3.000000	4.830000
		11432	7.333333	12.958889
		11433	3.400000	4.970000
		11434	7.500000	4.917500
		11435	4.400000	8.114000
		11436	2.500000	4.280000
		11691	4.500000	5.690000
		11692	3.500000	4.830000
		11693	2.000000	4.460000
		11694	3.000000	8.980000
	STATEN IS	10301	17.200000	8.384000
		10302	22.500000	9.170000
		10303	4.000000	7.780000
		10304	5.571429	4.561429
		10305	2.750000	5.005000
		10306	3.000000	5.746667
		10307	0.666667	6.796667
		10308	2.200000	6.630000
		10309	1.000000	6.880000
		10310	10.500000	6.305000
		10312	4.285714	7.267143
		10314	7.800000	9.894000

168 rows × 3 columns

The dataframe below is what we want. Now, it is easier to analyze school safety data compared to sale prices based on year, borough, and zip code.

```
In [780]: combo = pd.merge(year13.reset_index(), year13_2.reset_index(),
                        on = ["Year", "ZIP_CODE"],
                        how = "left",
                        indicator = True)
combo.dropna(inplace=True)
combo = combo.drop(["Borough_Name_y"], axis = 1)
combo = combo.rename(columns={'Borough_Name_x': 'Borough_Name'})
combo
```


Out[780]:

	Year	Borough_Name	ZIP_CODE	SALE_PRICE	Total_Crime	Average_Crime	Standard_dev
0	2013-14	BRONX	10451	1365265	3.600000	5.190000	-1.5
1	2013-14	BRONX	10452	2217611	2.000000	5.725000	-3.7
2	2013-14	BRONX	10453	349115	1.666667	5.446667	-3.7
3	2013-14	BRONX	10454	282433	3.500000	4.830000	-1.3
4	2013-14	BRONX	10455	458322	1.750000	4.412500	-2.6
5	2013-14	BRONX	10456	2290717	4.142857	5.647143	-1.5
6	2013-14	BRONX	10457	1416273	4.400000	5.338000	-0.9
7	2013-14	BRONX	10458	7197350	2.833333	6.568333	-3.7
8	2013-14	BRONX	10459	1448325	4.857143	5.087143	-0.2
9	2013-14	BRONX	10460	663833	13.400000	4.390000	9.0
10	2013-14	BRONX	10461	749694	8.200000	6.786000	1.4
11	2013-14	BRONX	10462	605826	4.500000	9.910000	-5.4
12	2013-14	BRONX	10463	327278	4.125000	6.692500	-2.5
13	2013-14	BRONX	10464	549060	2.000000	3.730000	-1.7
14	2013-14	BRONX	10465	436606	2.000000	6.630000	-4.6
15	2013-14	BRONX	10466	916320	2.571429	5.358571	-2.7
16	2013-14	BRONX	10467	739884	1.500000	6.673333	-5.1
17	2013-14	BRONX	10468	961225	2.625000	10.055000	-7.4
18	2013-14	BRONX	10469	424092	5.200000	8.490000	-3.2
19	2013-14	BRONX	10470	447642	2.000000	4.830000	-2.8
20	2013-14	BRONX	10471	455364	2.000000	4.830000	-2.8
21	2013-14	BRONX	10472	699696	3.250000	7.330000	-4.0
22	2013-14	BRONX	10473	289214	2.000000	5.550000	-3.5

	Year	Borough_Name	ZIP_CODE	SALE_PRICE	Total_Crime	Average_Crime	Standard_dev
23	2013-14	BRONX	10474	1921813	6.000000	6.630000	-0.6
24	2013-14	BRONX	10475	755166	4.250000	5.730000	-1.4
25	2013-14	BROOKLYN	11201	2114862	4.000000	5.483333	-1.4
26	2013-14	BROOKLYN	11203	556248	3.500000	5.390000	-1.8
27	2013-14	BROOKLYN	11204	900740	2.666667	6.685556	-4.0
28	2013-14	BROOKLYN	11205	870965	3.250000	5.005000	-1.7
29	2013-14	BROOKLYN	11206	1009415	4.875000	6.227500	-1.3
...
139	2013-14	QUEENS	11419	366850	2.000000	5.430000	-3.4
140	2013-14	QUEENS	11420	363263	0.500000	8.013333	-7.5
141	2013-14	QUEENS	11421	468431	0.000000	7.230000	-7.2
142	2013-14	QUEENS	11422	625858	2.500000	5.095000	-2.5
143	2013-14	QUEENS	11423	284569	3.500000	7.185000	-3.6
144	2013-14	QUEENS	11426	340599	4.000000	5.580000	-1.5
145	2013-14	QUEENS	11427	373299	17.000000	14.856667	2.1
146	2013-14	QUEENS	11428	362388	5.000000	7.563333	-2.5
147	2013-14	QUEENS	11429	378807	3.000000	4.830000	-1.8
148	2013-14	QUEENS	11432	590797	7.333333	12.958889	-5.6
149	2013-14	QUEENS	11433	317807	3.400000	4.970000	-1.5
150	2013-14	QUEENS	11434	318328	7.500000	4.917500	2.5
151	2013-14	QUEENS	11435	390463	4.400000	8.114000	-3.7
152	2013-14	QUEENS	11436	283596	2.500000	4.280000	-1.7
153	2013-14	QUEENS	11691	394026	4.500000	5.690000	-1.1
154	2013-14	QUEENS	11692	421684	3.500000	4.830000	-1.3

	Year	Borough_Name	ZIP_CODE	SALE_PRICE	Total_Crime	Average_Crime	Standard_dev
155	2013-14	QUEENS	11693	134169	2.000000	4.460000	-2.4
156	2013-14	QUEENS	11694	461443	3.000000	8.980000	-5.9
157	2013-14	STATEN IS	10301	368753	17.200000	8.384000	8.8
158	2013-14	STATEN IS	10302	403625	22.500000	9.170000	13.3
159	2013-14	STATEN IS	10303	840461	4.000000	7.780000	-3.7
160	2013-14	STATEN IS	10304	404497	5.571429	4.561429	1.0
161	2013-14	STATEN IS	10305	356664	2.750000	5.005000	-2.2
162	2013-14	STATEN IS	10306	567699	3.000000	5.746667	-2.7
163	2013-14	STATEN IS	10307	537516	0.666667	6.796667	-6.1
164	2013-14	STATEN IS	10308	410269	2.200000	6.630000	-4.4
165	2013-14	STATEN IS	10309	504953	1.000000	6.880000	-5.8
166	2013-14	STATEN IS	10310	338469	10.500000	6.305000	4.1
167	2013-14	STATEN IS	10312	424609	4.285714	7.267143	-2.9
168	2013-14	STATEN IS	10314	382126	7.800000	9.894000	-2.0

169 rows × 8 columns

```
In [781]: housing14 = housingdf[housingdf.Year == "2014-15"]
housing14 = housing14.drop(housing14[housing14['SALE_PRICE'].isin([0,1,' ',10
])].index)
housing14 = housing14[~housing14['SALE_PRICE'].isnull()]

housing14['ZIP_CODE'] = housing14.ZIP_CODE.astype(int)
safety_2014['ZIP_CODE'] = safety_2014.ZIP_CODE.astype(int)
housing14DF = housing14[housing14['ZIP_CODE'].isin(safety_2014['ZIP_CODE'])]

housing14DF["SALE_PRICE"] = housing14DF.SALE_PRICE.astype(float)

year_dict3 = {"SALE_PRICE": "first"}
year14 = housing14DF.groupby(["Year", "Borough_Name", "ZIP_CODE"]).mean()
year14["SALE_PRICE"] = year14.SALE_PRICE.astype(int)
year14
```

```
C:\me\Anaconda\lib\site-packages\ipykernel_launcher.py:6: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy
```

```
C:\me\Anaconda\lib\site-packages\ipykernel_launcher.py:9: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy
```

```
if __name__ == '__main__':
```

Out[781]:

SALE_PRICE			
Year	Borough_Name	ZIP_CODE	
2014-15	BRONX	10451	923246
		10452	1811710
		10453	1737000
		10454	1196055
		10455	443431
		10456	1103992
		10457	1497354
		10458	2554219
		10459	415008
		10460	674943
		10461	1286456
		10462	714658
		10463	512521
		10464	437997
		10465	435710
		10466	362956
		10467	616818
		10468	2273016
		10469	416119
		10470	840046
		10471	496135
		10472	354271
		10473	320786
		10474	1160076
		10475	314983
	BROOKLYN	11201	1778961
		11203	418406
		11204	944238
		11205	951021
		11206	1626919

	QUEENS	11419	456323
		11420	429209
		11421	445052

SALE_PRICE		
Year	Borough_Name	ZIP_CODE
		11422
		389787
		11423
		615140
		11426
		376680
		11427
		417036
		11428
		402039
		11429
		360748
		11432
		668534
		11433
		437903
		11434
		346420
		11435
		430110
		11436
		678344
		11691
		835413
		11692
		468563
		11693
		706277
		11694
		534721
	STATEN IS	10301
		469516
		10302
		554404
		10303
		249045
		10304
		408053
		10305
		479529
		10306
		437843
		10307
		528472
		10308
		422041
		10309
		549105
		10310
		457077
		10312
		444672
		10314
		417255

169 rows × 1 columns

```
In [782]: housingcode2 = housing14['ZIP_CODE'].values
safety14df = safety_2014[safety_2014['ZIP_CODE'].isin(housing14["ZIP_CODE"])]

safety14df = safety14df.rename(columns = {"Total Number of Crimes":"Total_Crime", "Average Num of Crimes":"Average_Crime"})

year_dict3 = {"Total_Crime":"first", "Average_Crime":"first", "Standard_deviation":"first"}
year14_2 = safety14df.groupby(["Year", "Borough_Name", "ZIP_CODE"]).mean()
year14_2
```


Out[782]:

			Total_Crime	Average_Crime	Standard_deviation
Year	Borough_Name	ZIP_CODE			
2014-15	BRONX	10451	7.000000	5.000000	2.000000
		10452	1.375000	5.547500	-4.172500
		10453	2.500000	5.438333	-2.938333
		10454	1.800000	4.346000	-2.546000
		10455	2.285714	4.722857	-2.437143
		10456	5.111111	5.097778	0.013333
		10457	7.066667	5.133333	1.933333
		10458	4.500000	6.225833	-1.725833
		10459	6.142857	4.871429	1.271429
		10460	13.600000	4.346000	9.254000
		10461	2.800000	6.104000	-3.304000
		10462	7.000000	9.136667	-2.136667
		10463	6.125000	6.553750	-0.428750
		10464	0.000000	4.040000	-4.040000
		10465	2.000000	6.800000	-4.800000
		10466	1.571429	4.990000	-3.418571
		10467	1.800000	6.988000	-5.188000
		10468	4.000000	9.438750	-5.438750
		10469	5.000000	7.960000	-2.960000
		10470	1.000000	4.550000	-3.550000
		10471	2.000000	4.550000	-2.550000
		10472	3.750000	7.725000	-3.975000
		10473	1.166667	5.130000	-3.963333
		10474	6.000000	6.800000	-0.800000
		10475	6.600000	5.348000	1.252000
	BROOKLYN	11201	4.500000	5.258333	-0.758333
		11203	2.900000	5.590000	-2.690000
		11204	1.875000	7.130000	-5.255000
		11205	9.666667	4.423333	5.243333
		11206	4.700000	5.251000	-0.551000

	QUEENS	11419	1.500000	4.672500	-3.172500
		11420	1.166667	8.031667	-6.865000
		11421	0.333333	5.746667	-5.413333

		Total_Crime	Average_Crime	Standard_deviation
Year	Borough_Name	ZIP_CODE		
		11422	2.500000	4.800000
		11423	5.750000	6.492500
		11426	3.600000	5.166000
		11427	1.500000	6.600000
		11428	3.000000	6.666667
		11429	0.000000	4.550000
		11432	7.333333	12.570000
		11433	1.500000	4.985000
		11434	4.250000	4.486250
		11435	2.666667	11.323333
		11436	6.666667	4.210000
		11691	3.833333	5.091667
		11692	12.000000	4.550000
		11693	4.500000	3.925000
		11694	3.000000	8.775000
	STATEN IS	10301	14.222222	8.927778
		10302	20.250000	7.390000
		10303	4.500000	7.725000
		10304	4.000000	4.338333
		10305	1.250000	4.985000
		10306	2.166667	5.898333
		10307	3.333333	5.916667
		10308	2.400000	6.138000
		10309	1.000000	5.675000
		10310	9.500000	6.137500
		10312	0.857143	7.128571
		10314	7.200000	9.221000

168 rows × 3 columns

```
In [783]: combo2 = pd.merge(year14.reset_index(), year14_2.reset_index(),
                             on = ["Year", "ZIP_CODE"],
                             how = "left",
                             indicator = True)
combo2.dropna(inplace=True)
combo2 = combo2.drop(["Borough_Name_y"], axis = 1)
combo2 = combo2.rename(columns={'Borough_Name_x': 'Borough_Name'})
combo2
```

Out[783]:

	Year	Borough_Name	ZIP_CODE	SALE_PRICE	Total_Crime	Average_Crime	Standard_dev
0	2014-15	BRONX	10451	923246	7.000000	5.000000	2.0
1	2014-15	BRONX	10452	1811710	1.375000	5.547500	-4.1
2	2014-15	BRONX	10453	1737000	2.500000	5.438333	-2.9
3	2014-15	BRONX	10454	1196055	1.800000	4.346000	-2.5
4	2014-15	BRONX	10455	443431	2.285714	4.722857	-2.4
5	2014-15	BRONX	10456	1103992	5.111111	5.097778	0.0
6	2014-15	BRONX	10457	1497354	7.066667	5.133333	1.9
7	2014-15	BRONX	10458	2554219	4.500000	6.225833	-1.7
8	2014-15	BRONX	10459	415008	6.142857	4.871429	1.2
9	2014-15	BRONX	10460	674943	13.600000	4.346000	9.2
10	2014-15	BRONX	10461	1286456	2.800000	6.104000	-3.3
11	2014-15	BRONX	10462	714658	7.000000	9.136667	-2.1
12	2014-15	BRONX	10463	512521	6.125000	6.553750	-0.4
13	2014-15	BRONX	10464	437997	0.000000	4.040000	-4.0
14	2014-15	BRONX	10465	435710	2.000000	6.800000	-4.8
15	2014-15	BRONX	10466	362956	1.571429	4.990000	-3.4
16	2014-15	BRONX	10467	616818	1.800000	6.988000	-5.1
17	2014-15	BRONX	10468	2273016	4.000000	9.438750	-5.4
18	2014-15	BRONX	10469	416119	5.000000	7.960000	-2.9
19	2014-15	BRONX	10470	840046	1.000000	4.550000	-3.5
20	2014-15	BRONX	10471	496135	2.000000	4.550000	-2.5
21	2014-15	BRONX	10472	354271	3.750000	7.725000	-3.9
22	2014-15	BRONX	10473	320786	1.166667	5.130000	-3.9

	Year	Borough_Name	ZIP_CODE	SALE_PRICE	Total_Crime	Average_Crime	Standard_dev
23	2014-15	BRONX	10474	1160076	6.000000	6.800000	-0.8
24	2014-15	BRONX	10475	314983	6.600000	5.348000	1.2
25	2014-15	BROOKLYN	11201	1778961	4.500000	5.258333	-0.7
26	2014-15	BROOKLYN	11203	418406	2.900000	5.590000	-2.6
27	2014-15	BROOKLYN	11204	944238	1.875000	7.130000	-5.2
28	2014-15	BROOKLYN	11205	951021	9.666667	4.423333	5.2
29	2014-15	BROOKLYN	11206	1626919	4.700000	5.251000	-0.5
...
139	2014-15	QUEENS	11419	456323	1.500000	4.672500	-3.1
140	2014-15	QUEENS	11420	429209	1.166667	8.031667	-6.8
141	2014-15	QUEENS	11421	445052	0.333333	5.746667	-5.4
142	2014-15	QUEENS	11422	389787	2.500000	4.800000	-2.3
143	2014-15	QUEENS	11423	615140	5.750000	6.492500	-0.7
144	2014-15	QUEENS	11426	376680	3.600000	5.166000	-1.5
145	2014-15	QUEENS	11427	417036	1.500000	6.600000	-5.1
146	2014-15	QUEENS	11428	402039	3.000000	6.666667	-3.6
147	2014-15	QUEENS	11429	360748	0.000000	4.550000	-4.5
148	2014-15	QUEENS	11432	668534	7.333333	12.570000	-5.2
149	2014-15	QUEENS	11433	437903	1.500000	4.985000	-3.4
150	2014-15	QUEENS	11434	346420	4.250000	4.486250	-0.2
151	2014-15	QUEENS	11435	430110	2.666667	11.323333	-8.6
152	2014-15	QUEENS	11436	678344	6.666667	4.210000	2.4
153	2014-15	QUEENS	11691	835413	3.833333	5.091667	-1.2
154	2014-15	QUEENS	11692	468563	12.000000	4.550000	7.4

	Year	Borough_Name	ZIP_CODE	SALE_PRICE	Total_Crime	Average_Crime	Standard_dev
155	2014-15	QUEENS	11693	706277	4.500000	3.925000	0.5
156	2014-15	QUEENS	11694	534721	3.000000	8.775000	-5.7
157	2014-15	STATEN IS	10301	469516	14.222222	8.927778	5.2
158	2014-15	STATEN IS	10302	554404	20.250000	7.390000	12.8
159	2014-15	STATEN IS	10303	249045	4.500000	7.725000	-3.2
160	2014-15	STATEN IS	10304	408053	4.000000	4.338333	-0.3
161	2014-15	STATEN IS	10305	479529	1.250000	4.985000	-3.7
162	2014-15	STATEN IS	10306	437843	2.166667	5.898333	-3.7
163	2014-15	STATEN IS	10307	528472	3.333333	5.916667	-2.5
164	2014-15	STATEN IS	10308	422041	2.400000	6.138000	-3.7
165	2014-15	STATEN IS	10309	549105	1.000000	5.675000	-4.6
166	2014-15	STATEN IS	10310	457077	9.500000	6.137500	3.3
167	2014-15	STATEN IS	10312	444672	0.857143	7.128571	-6.2
168	2014-15	STATEN IS	10314	417255	7.200000	9.221000	-2.0

169 rows × 8 columns

```
In [784]: housing15 = housingdf[housingdf.Year == "2015-16"]
housing15 = housing15.drop(housing15[housing15['SALE_PRICE'].isin([0,1,' ',10
])].index)
housing15 = housing15[~housing15['SALE_PRICE'].isnull()]

housing15['ZIP_CODE'] = housing15.ZIP_CODE.astype(int)
safety_2015['ZIP_CODE'] = safety_2015.ZIP_CODE.astype(int)
housing15DF = housing15[housing15['ZIP_CODE'].isin(safety_2015['ZIP_CODE'])]

housing15DF["SALE_PRICE"] = housing15DF.SALE_PRICE.astype(float)

year_dict4 = {"SALE_PRICE": "first"}
year15 = housing15DF.groupby(["Year", "Borough_Name", "ZIP_CODE"]).mean()
year15["SALE_PRICE"] = year15.SALE_PRICE.astype(int)
year15
```

```
C:\me\Anaconda\lib\site-packages\ipykernel_launcher.py:6: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy
```

```
C:\me\Anaconda\lib\site-packages\ipykernel_launcher.py:9: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy
```

```
if __name__ == '__main__':
```


Out[784]:

SALE_PRICE			
Year	Borough_Name	ZIP_CODE	
2015-16	BRONX	10451	407360
		10452	1262625
		10453	3141700
		10454	284250
		10455	993614
		10456	305169
		10457	869245
		10458	917753
		10459	333609
		10460	816326
		10461	926303
		10462	275952
		10463	441711
		10464	421864
		10465	444976
		10466	652316
		10467	1449501
		10468	1265253
		10469	476063
		10470	511825
		10471	1288740
		10472	525058
		10473	1695227
		10474	657500
		10475	850000
	BROOKLYN	11201	1727723
		11203	747998
		11204	832474
		11205	1567905
		11206	2042993

	QUEENS	11419	462865
		11420	436256
		11421	522131

SALE_PRICE		
Year	Borough_Name	ZIP_CODE
		11422
		455563
		11423
		559317
		11426
		498640
		11427
		528874
		11428
		489816
		11429
		381153
		11432
		528123
		11433
		530633
		11434
		488843
		11435
		980518
		11436
		348142
		11691
		355604
		11692
		571220
		11693
		555313
		11694
		634530
	STATEN IS	10301
		633260
		10302
		365857
		10303
		277018
		10304
		449887
		10305
		464439
		10306
		647595
		10307
		498018
		10308
		426999
		10309
		492461
		10310
		396361
		10312
		486693
		10314
		443015

169 rows × 1 columns

```
In [785]: housingcode3 = housing15['ZIP_CODE'].values
safety15df = safety_2015[safety_2015['ZIP_CODE'].isin(housing15["ZIP_CODE"])]

safety15df = safety15df.rename(columns = {"Total Number of Crimes":"Total_Crime", "Average Num of Crimes":"Average_Crime"})

year_dict4 = {"Total_Crime":"first", "Average_Crime":"first", "Standard_deviation":"first"}
year15_2 = safety15df.groupby(["Year", "Borough_Name", "ZIP_CODE"]).mean()
year15_2
```

Out[785]:

			Total_Crime	Average_Crime	Standard_deviation
Year	Borough_Name	ZIP_CODE			
2015-16	BRONX	10451	6.200000	5.154000	1.046000
		10452	2.000000	5.527500	-3.527500
		10453	2.000000	5.590000	-3.590000
		10454	4.200000	4.110000	0.090000
		10455	4.714286	4.595714	0.118571
		10456	4.888889	5.014444	-0.125556
		10457	5.800000	5.154000	0.646000
		10458	3.250000	5.971667	-2.721667
		10459	8.142857	4.675714	3.467143
		10460	17.200000	4.110000	13.090000
		10461	2.200000	5.472000	-3.272000
		10462	5.500000	8.628333	-3.128333
		10463	3.500000	6.330000	-2.830000
		10464	0.000000	3.910000	-3.910000
		10465	1.000000	6.270000	-5.270000
		10466	1.571429	5.421429	-3.850000
		10467	3.200000	6.798000	-3.598000
		10468	1.875000	8.865000	-6.990000
		10469	5.000000	7.682000	-2.682000
		10470	1.500000	4.160000	-2.660000
		10471	2.000000	4.410000	-2.410000
		10472	2.750000	7.520000	-4.770000
		10473	3.333333	5.210000	-1.876667
		10474	6.000000	6.270000	-0.270000
		10475	8.200000	4.954000	3.246000
	BROOKLYN	11201	8.833333	5.033333	3.800000
		11203	2.700000	5.174000	-2.474000
		11204	1.750000	6.937500	-5.187500
		11205	7.000000	4.163333	2.836667
		11206	4.900000	4.996000	-0.096000

	QUEENS	11419	0.500000	4.645000	-4.145000
		11420	0.666667	7.706667	-7.040000
		11421	0.000000	5.696667	-5.696667

		Total_Crime	Average_Crime	Standard_deviation
Year	Borough_Name	ZIP_CODE		
		11422	2.000000	4.645000
		11423	8.000000	6.355000
		11426	1.000000	5.082000
		11427	1.000000	6.590000
		11428	6.333333	6.483333
		11429	0.000000	4.410000
		11432	8.333333	11.888889
		11433	1.000000	4.750000
		11434	5.000000	4.347500
		11435	0.833333	10.980000
		11436	6.333333	4.076667
		11691	4.571429	4.810000
		11692	7.500000	4.410000
		11693	4.500000	3.950000
		11694	0.000000	8.550000
	STATEN IS	10301	11.111111	8.762222
		10302	22.000000	8.300000
		10303	6.000000	7.520000
		10304	2.000000	4.246667
		10305	4.750000	4.750000
		10306	3.500000	6.326667
		10307	2.666667	5.863333
		10308	2.200000	6.338000
		10309	1.500000	5.340000
		10310	8.750000	5.965000
		10312	2.142857	6.941429
		10314	6.000000	9.001000

168 rows × 3 columns

```
In [786]: combo3 = pd.merge(year15.reset_index(), year15_2.reset_index(),
                             on = ["Year", "ZIP_CODE"],
                             how = "left",
                             indicator = True)
combo3.dropna(inplace=True)
combo3 = combo3.drop(["Borough_Name_y"], axis = 1)
combo3 = combo3.rename(columns={'Borough_Name_x': 'Borough_Name'})
combo3
```

Out[786]:

	Year	Borough_Name	ZIP_CODE	SALE_PRICE	Total_Crime	Average_Crime	Standard_dev
0	2015-16	BRONX	10451	407360	6.200000	5.154000	1.0
1	2015-16	BRONX	10452	1262625	2.000000	5.527500	-3.5
2	2015-16	BRONX	10453	3141700	2.000000	5.590000	-3.5
3	2015-16	BRONX	10454	284250	4.200000	4.110000	0.0
4	2015-16	BRONX	10455	993614	4.714286	4.595714	0.1
5	2015-16	BRONX	10456	305169	4.888889	5.014444	-0.1
6	2015-16	BRONX	10457	869245	5.800000	5.154000	0.6
7	2015-16	BRONX	10458	917753	3.250000	5.971667	-2.7
8	2015-16	BRONX	10459	333609	8.142857	4.675714	3.4
9	2015-16	BRONX	10460	816326	17.200000	4.110000	13.0
10	2015-16	BRONX	10461	926303	2.200000	5.472000	-3.2
11	2015-16	BRONX	10462	275952	5.500000	8.628333	-3.1
12	2015-16	BRONX	10463	441711	3.500000	6.330000	-2.8
13	2015-16	BRONX	10464	421864	0.000000	3.910000	-3.9
14	2015-16	BRONX	10465	444976	1.000000	6.270000	-5.2
15	2015-16	BRONX	10466	652316	1.571429	5.421429	-3.8
16	2015-16	BRONX	10467	1449501	3.200000	6.798000	-3.5
17	2015-16	BRONX	10468	1265253	1.875000	8.865000	-6.9
18	2015-16	BRONX	10469	476063	5.000000	7.682000	-2.6
19	2015-16	BRONX	10470	511825	1.500000	4.160000	-2.6
20	2015-16	BRONX	10471	1288740	2.000000	4.410000	-2.4
21	2015-16	BRONX	10472	525058	2.750000	7.520000	-4.7
22	2015-16	BRONX	10473	1695227	3.333333	5.210000	-1.8

	Year	Borough_Name	ZIP_CODE	SALE_PRICE	Total_Crime	Average_Crime	Standard_dev
23	2015-16	BRONX	10474	657500	6.000000	6.270000	-0.2
24	2015-16	BRONX	10475	850000	8.200000	4.954000	3.2
25	2015-16	BROOKLYN	11201	1727723	8.833333	5.033333	3.8
26	2015-16	BROOKLYN	11203	747998	2.700000	5.174000	-2.4
27	2015-16	BROOKLYN	11204	832474	1.750000	6.937500	-5.1
28	2015-16	BROOKLYN	11205	1567905	7.000000	4.163333	2.8
29	2015-16	BROOKLYN	11206	2042993	4.900000	4.996000	-0.0
...
139	2015-16	QUEENS	11419	462865	0.500000	4.645000	-4.1
140	2015-16	QUEENS	11420	436256	0.666667	7.706667	-7.0
141	2015-16	QUEENS	11421	522131	0.000000	5.696667	-5.6
142	2015-16	QUEENS	11422	455563	2.000000	4.645000	-2.6
143	2015-16	QUEENS	11423	559317	8.000000	6.355000	1.6
144	2015-16	QUEENS	11426	498640	1.000000	5.082000	-4.0
145	2015-16	QUEENS	11427	528874	1.000000	6.590000	-5.5
146	2015-16	QUEENS	11428	489816	6.333333	6.483333	-0.1
147	2015-16	QUEENS	11429	381153	0.000000	4.410000	-4.4
148	2015-16	QUEENS	11432	528123	8.333333	11.888889	-3.5
149	2015-16	QUEENS	11433	530633	1.000000	4.750000	-3.7
150	2015-16	QUEENS	11434	488843	5.000000	4.347500	0.6
151	2015-16	QUEENS	11435	980518	0.833333	10.980000	-10.1
152	2015-16	QUEENS	11436	348142	6.333333	4.076667	2.2
153	2015-16	QUEENS	11691	355604	4.571429	4.810000	-0.2
154	2015-16	QUEENS	11692	571220	7.500000	4.410000	3.0

	Year	Borough_Name	ZIP_CODE	SALE_PRICE	Total_Crime	Average_Crime	Standard_dev
155	2015-16	QUEENS	11693	555313	4.500000	3.950000	0.5
156	2015-16	QUEENS	11694	634530	0.000000	8.550000	-8.5
157	2015-16	STATEN IS	10301	633260	11.111111	8.762222	2.3
158	2015-16	STATEN IS	10302	365857	22.000000	8.300000	13.7
159	2015-16	STATEN IS	10303	277018	6.000000	7.520000	-1.5
160	2015-16	STATEN IS	10304	449887	2.000000	4.246667	-2.2
161	2015-16	STATEN IS	10305	464439	4.750000	4.750000	0.0
162	2015-16	STATEN IS	10306	647595	3.500000	6.326667	-2.8
163	2015-16	STATEN IS	10307	498018	2.666667	5.863333	-3.1
164	2015-16	STATEN IS	10308	426999	2.200000	6.338000	-4.1
165	2015-16	STATEN IS	10309	492461	1.500000	5.340000	-3.8
166	2015-16	STATEN IS	10310	396361	8.750000	5.965000	2.7
167	2015-16	STATEN IS	10312	486693	2.142857	6.941429	-4.7
168	2015-16	STATEN IS	10314	443015	6.000000	9.001000	-3.0

169 rows × 8 columns

```
In [787]: finaldf = combo.append(combo2)
finaldf = finaldf.append(combo3)
finaldf.head()
```

Out[787]:

	Year	Borough_Name	ZIP_CODE	SALE_PRICE	Total_Crime	Average_Crime	Standard_deviat
0	2013-14	BRONX	10451	1365265	3.600000	5.190000	-1.5
1	2013-14	BRONX	10452	2217611	2.000000	5.725000	-3.7
2	2013-14	BRONX	10453	349115	1.666667	5.446667	-3.7
3	2013-14	BRONX	10454	282433	3.500000	4.830000	-1.3
4	2013-14	BRONX	10455	458322	1.750000	4.412500	-2.6

This finaldf dataframe gives us our final result. In this dataframe, the data is grouped by the individual year and then by the borough after appending the dataframe for each year. From here, we can visualize the data next.

Part 4: Visualizing the data

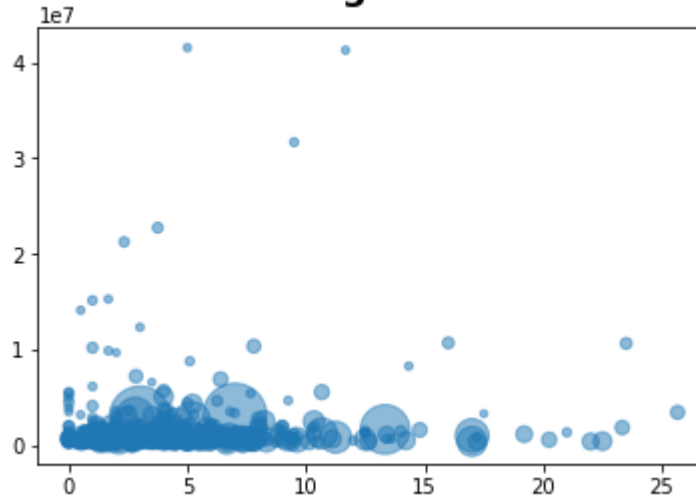
1. Let's visualize the data! For our first graph, we are going to create a scatter plot and plot the crime statistics, based on the average statistics for each unique Zip code, against the Housing Prices for all boroughs of NYC. This will show us in general how whether the **total** and **average** number of crimes for all areas in NYC have any correlation with NYC's housing prices.

```
In [896]: fig, ax = plt.subplots()

fig.suptitle("Total Crime vs. Housing Prices from 2013-2016", fontsize = 18, fontweight = 'bold')
ax.scatter(finaldf["Total_Crime"], finaldf["SALE_PRICE"],
          alpha = .5,
          s = finaldf["Average_Crime"]**2)

plt.show()
```

Total Crime vs. Housing Prices from 2013-2016

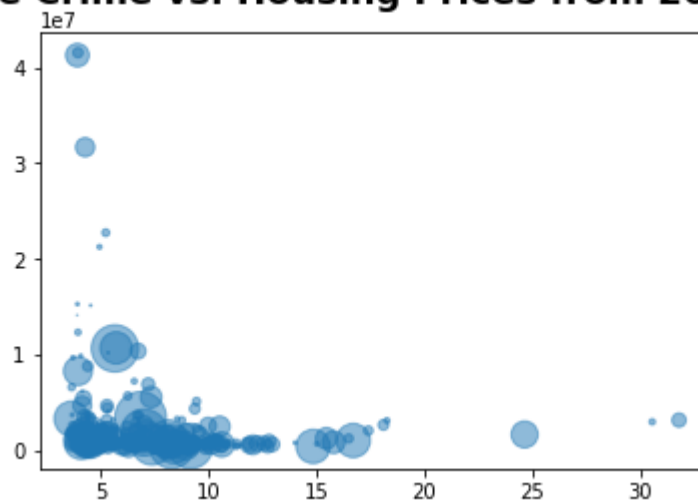


```
In [897]: fig, ax = plt.subplots()

fig.suptitle("Average Crime vs. Housing Prices from 2013-2016", fontsize = 18, fontweight = 'bold')
ax.scatter(finaldf["Average_Crime"], finaldf["SALE_PRICE"],
          alpha = .5,
          s = finaldf["Total_Crime"]**2)

plt.show()
```

Average Crime vs. Housing Prices from 2013-2016



From these two scatter plots, we can observe an overall trend in crime and housing prices. The x-axis is represented by the average number of crimes and the y-axis is represented by the sale price of houses. For the Total Crime vs. Housing Prices graph, the bigger dots represent a higher amount of average crimes in that area. For the Average Crime vs. Housing Prices graph, the bigger dots represent a higher amount of total crimes for a given year. From 2013 - 2016, the graphs show that most of the dots are concentrated around the bottom left, which indicates low crime and low housing prices. However we can see that for the few data points that are on the right hand side of the graph (representing higher crimes), the housing prices tend to be low. While lower crime areas have low housing prices as well, the only data points that represent higher prices are also on the left side of the plot, also the areas with lower crime rates.

The next part of visualizing our data, we are going to break it up into **two parts** to observe two different things: for the **first part**, we will focus on observing the crime rate vs. housing prices over the year time horizon that we have been given data for (2013-2016). This will give us information on whether there are overall changes happening over the years, such as whether crimes are decreasing or

1. Now we will break down the data by each year and visualize the data for the three time frames 2013-14, 2014-15 and 2015-16.

```
In [790]: bronxmean13 = combo[combo.Borough_Name == "BRONX"]
          brooklynmean13 = combo[combo.Borough_Name == "BROOKLYN"]
          manhattanmean13 = combo[combo.Borough_Name == "MANHATTAN"]
          queensmean13 = combo[combo.Borough_Name == "QUEENS"]
          statenmean13 = combo[combo.Borough_Name == "STATEN IS"]
          bronxmean13["SALE_PRICE"].mean()
```

```
Out[790]: 1118724.96
```

```
In [791]: brooklynmean13["SALE_PRICE"].mean()
```

```
Out[791]: 877782.447368421
```

```
In [792]: bronxmean14 = combo2[combo2.Borough_Name == "BRONX"]
          brooklynmean14 = combo2[combo2.Borough_Name == "BROOKLYN"]
          manhattanmean14 = combo2[combo2.Borough_Name == "MANHATTAN"]
          queensmean14 = combo2[combo2.Borough_Name == "QUEENS"]
          statenmean14 = combo2[combo2.Borough_Name == "STATEN IS"]

          bronxmean15 = combo3[combo3.Borough_Name == "BRONX"]
          brooklynmean15 = combo3[combo3.Borough_Name == "BROOKLYN"]
          manhattanmean15 = combo3[combo3.Borough_Name == "MANHATTAN"]
          queensmean15 = combo3[combo3.Borough_Name == "QUEENS"]
          statenmean15 = combo3[combo3.Borough_Name == "STATEN IS"]
```

We will create a new dataframe with information on the mean of crime and housing statistics for 2013-2016.

```

In [803]: data = [['2013-14', 'BRONX', bronxmean13["SALE_PRICE"].mean(), bronxmean13["Total_Crime"].mean()], ['2013-14', 'BROOKLYN', brooklynmean13["SALE_PRICE"].mean(), brooklynmean13["Total_Crime"].mean()],
                ['2013-14', 'MANHATTAN', manhattanmean13["SALE_PRICE"].mean(), manhattanmean13["Total_Crime"].mean()], ['2013-14', 'QUEENS', queensmean13["SALE_PRICE"].mean(), queensmean13["Total_Crime"].mean()],
                ['2013-14', 'STATEN IS', statenmean13["SALE_PRICE"].mean(), statenmean13["Total_Crime"].mean()],
                ['2014-15', 'BRONX', bronxmean14["SALE_PRICE"].mean(), bronxmean14["Total_Crime"].mean()], ['2014-15', 'BROOKLYN', brooklynmean14["SALE_PRICE"].mean(), brooklynmean14["Total_Crime"].mean()],
                ['2014-15', 'MANHATTAN', manhattanmean14["SALE_PRICE"].mean(), manhattanmean14["Total_Crime"].mean()], ['2014-15', 'QUEENS', queensmean14["SALE_PRICE"].mean(), queensmean14["Total_Crime"].mean()],
                ['2014-15', 'STATEN IS', statenmean14["SALE_PRICE"].mean(), statenmean14["Total_Crime"].mean()],
                ['2015-16', 'BRONX', bronxmean15["SALE_PRICE"].mean(), bronxmean15["Total_Crime"].mean()], ['2015-16', 'BROOKLYN', brooklynmean15["SALE_PRICE"].mean(), brooklynmean15["Total_Crime"].mean()],
                ['2015-16', 'MANHATTAN', manhattanmean15["SALE_PRICE"].mean(), manhattanmean15["Total_Crime"].mean()], ['2015-16', 'QUEENS', queensmean15["SALE_PRICE"].mean(), queensmean15["Total_Crime"].mean()],
                ['2015-16', 'STATEN IS', statenmean15["SALE_PRICE"].mean(), statenmean15["Total_Crime"].mean()]]
finalmean = pd.DataFrame(data, columns = ['Year', 'Borough_Name', 'Average_SALE_PRICE', 'Average_Total_Crime'])
finalmean["Average_SALE_PRICE"] = finalmean.Average_SALE_PRICE.astype(int)
finalmean

```

Out[803]:

	Year	Borough_Name	Average_SALE_PRICE	Average_Total_Crime
0	2013-14	BRONX	1118724	3.774857
1	2013-14	BROOKLYN	877782	4.210171
2	2013-14	MANHATTAN	3026404	5.311122
3	2013-14	QUEENS	641648	3.536789
4	2013-14	STATEN IS	461636	6.789484
5	2014-15	BRONX	915980	4.087778
6	2014-15	BROOKLYN	1066377	5.135377
7	2014-15	MANHATTAN	6415353	5.802327
8	2014-15	QUEENS	665295	3.126065
9	2014-15	STATEN IS	451417	5.889947
10	2015-16	BRONX	848557	4.241032
11	2015-16	BROOKLYN	1370560	4.726609
12	2015-16	MANHATTAN	4706927	4.490390
13	2015-16	QUEENS	734162	2.897118
14	2015-16	STATEN IS	465133	6.051720

```
In [866]: finalmean.reset_index()
finalmean
```

Out[866]:

	Borough_Name	Average_SALE_PRICE	Average_Total_Crime
Year			
2013-14	BRONX	1118724	3.774857
2013-14	BROOKLYN	877782	4.210171
2013-14	MANHATTAN	3026404	5.311122
2013-14	QUEENS	641648	3.536789
2013-14	STATEN IS	461636	6.789484
2014-15	BRONX	915980	4.087778
2014-15	BROOKLYN	1066377	5.135377
2014-15	MANHATTAN	6415353	5.802327
2014-15	QUEENS	665295	3.126065
2014-15	STATEN IS	451417	5.889947
2015-16	BRONX	848557	4.241032
2015-16	BROOKLYN	1370560	4.726609
2015-16	MANHATTAN	4706927	4.490390
2015-16	QUEENS	734162	2.897118
2015-16	STATEN IS	465133	6.051720

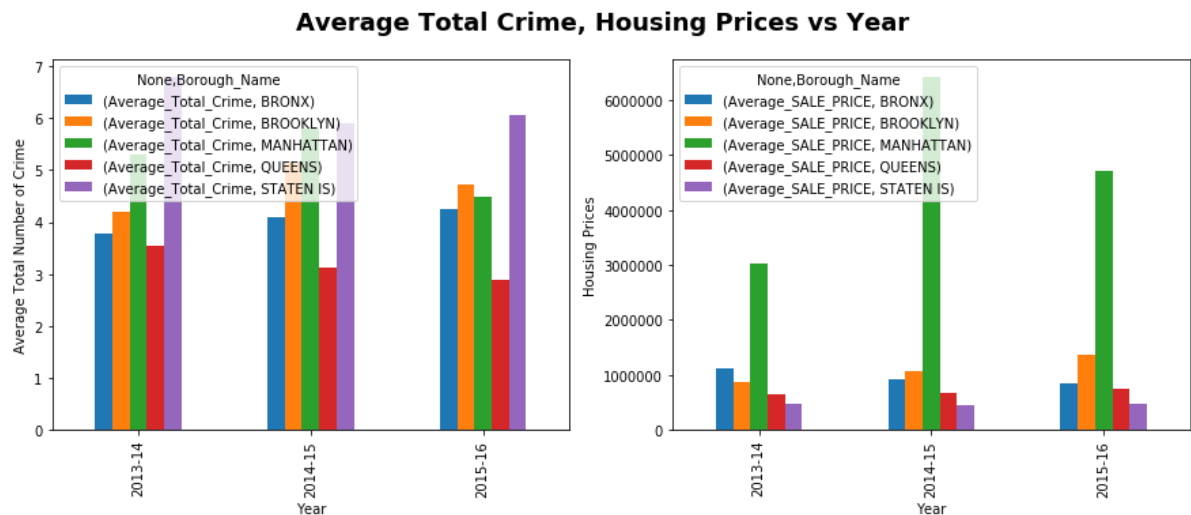
```
In [869]: test10 = finalmean.reset_index().pivot(index = 'Year', columns = 'Borough_Name', values = ['Average_SALE_PRICE', 'Average_Total_Crime'])
```

```
In [906]: fig1, (ax1,ax2) = plt.subplots(1,2, sharey = False,figsize = (15,5))

test10.drop(['Average_SALE_PRICE'], axis = 1).plot(kind = 'bar',ax=ax1, label
= 'Average Total Crime')
test10.drop(['Average_Total_Crime'], axis = 1).plot(kind = 'bar',ax=ax2, label
= 'Housing Prices')

fig1.suptitle("Average Total Crime, Housing Prices vs Year", fontsize = 18, fo
ntweight = 'bold')
ax1.set_ylabel('Average Total Number of Crime')
ax2.set_ylabel('Housing Prices')

plt.show()
```



By visualizing the data for each year, you can see if there are any changes or trends in the past few years regarding the statistics we are focusing on. For this process, we found the average sale price and average crime rate for each borough and then created a histogram to graph the changes per year. From these graphs, we can make the following conclusions:

- The average number of total crime in the Bronx increased slightly in the three year period
- The average number of total crime for both Brooklyn and Manhattan were higher in 2014-15 than in the other two years
- The average number of total crimes in Staten Island is the highest among the 5 boroughs in all three years, with only Manhattan having a similar amount in 2014-15
- Manhattan average housing prices are significantly higher than the other four boroughs, and Staten Island has the lowest average housing prices
- The graphs show that while Staten Island crime rates are the highest in schools, they have the lowest housing prices. This is consistent with the scatter plots above where the dots where there are higher amounts of crime, the housing prices are lower than areas with lower crime.
- Manhattan also has relatively large amounts of crime, but they have significantly higher housing prices than the other boroughs. There are other reasons for this that are unrelated to safety, but from the Total Crime vs. Housing Prices scatter plot, we can see several data points that account for this (higher crime and higher housing prices). However, they are scattered and not concentrated, suggesting that among the entire city, areas with lower crimes tend to have higher property value than areas where school safety is more of a concern.

1. We will visualize the data for each of the boroughs, which can give us information on whether certain boroughs are safer than others or whether the correlation between pricing and safety is stronger in one area than another. We can control for the passage of time by the fact that the time period we had data for was a relatively short period of time and from the graphs in part 2, we can see that there weren't any major disruptions towards the general data trends so we can assume that the trends for each borough through the three years are relatively representative for each borough for the recent years.

```
In [901]: totalmean = finalmean.reset_index().groupby(["Borough_Name"]).mean()  
totalmean["Average_SALE_PRICE"] = totalmean.Average_SALE_PRICE.astype(int)  
totalmean
```

Out[901]:

	Average_SALE_PRICE	Average_Total_Crime
Borough_Name		
BRONX	961087	4.034556
BROOKLYN	1104906	4.690719
MANHATTAN	4716228	5.201280
QUEENS	680368	3.186657
STATEN IS	459395	6.243717


```

In [928]: fig = plt.figure() # Create matplotlib figure

fig.suptitle("Average Housing Price and Average Total Crime vs. Borough [2013
- 2016]", fontsize = 18, fontweight = 'bold')

ax = fig.add_subplot(111) # Create matplotlib axes
ax2 = ax.twinx() # Create another axes that shares the same x-axis as ax.

width = 0.4

totalmean.Average_SALE_PRICE.plot(kind='bar', color='red', ax=ax, alpha = .5,
width=width, position=1, label = 'Housing Price')
totalmean.Average_Total_Crime.plot(kind='bar', color='blue', ax=ax2, alpha = .
5, width=width, position=0, label = 'Total Crime')

ax.set_ylabel('Average Housing Price')
ax2.set_ylabel('Average Total Crime')
ax.legend(loc = 'upper left')
ax2.legend(loc = 'upper right')

plt.show()

```

Average Housing Price and Average Total Crime vs. Borough [2013 - 2016]



Our third graph shows the average housing price and average total crime by borough (by calculating the means of each for the three years we have data for). This side by side comparison allows us to look at the results by sectors of the city. A few conclusions from this graph are:

- Queens has the total average crime around schools and Staten Island has the most. However, both of these boroughs have the lowest housing prices.
- On the other hand, Manhattan has very high housing prices and one of the highest crime rates at schools.
- This graph suggests that the correlation between housing prices and total crime may not be that significant. The rates of crime around schools do not differ that significantly, and housing prices are generally lower in areas with higher crime, but the case of Manhattan goes against that correlation.
- We can conclude that a weak correlation may exist but the difference in housing prices are probably best explained by other factors.

```
In [929]: shape_file = "C:\\me\\data bootcamp\\shape_file\\ZIP_CODE_040114.shx"
```

```
In [930]: shape_file
```

```
Out[930]: 'C:\\me\\data bootcamp\\shape_file\\ZIP_CODE_040114.shx'
```

```
In [931]: nyc_map = gpd.read_file(shape_file)
```

```
In [972]: nyc_map.rename(columns = {"ZIPCODE":"ZIP_CODE"}, inplace = True)  
nyc_map["ZIP_CODE"] = nyc_map["ZIP_CODE"].astype(int)
```

```
In [988]: newmap = pd.merge(nyc_map, finaldf,  
                             on = ["ZIP_CODE"],  
                             how = "inner",  
                             indicator = 'exists')  
  
newmap
```

Out[988]:

	ZIP_CODE	BLDGZIP	PO_NAME	POPULATION	AREA	STATE	COUNTY	ST_FIPS
0	11436	0	Jamaica	18681.0	2.269930e+07	NY	Queens	36
1	11436	0	Jamaica	18681.0	2.269930e+07	NY	Queens	36
2	11436	0	Jamaica	18681.0	2.269930e+07	NY	Queens	36
3	11213	0	Brooklyn	62426.0	2.963100e+07	NY	Kings	36
4	11213	0	Brooklyn	62426.0	2.963100e+07	NY	Kings	36
5	11213	0	Brooklyn	62426.0	2.963100e+07	NY	Kings	36
6	11212	0	Brooklyn	83866.0	4.197210e+07	NY	Kings	36
7	11212	0	Brooklyn	83866.0	4.197210e+07	NY	Kings	36
8	11212	0	Brooklyn	83866.0	4.197210e+07	NY	Kings	36
9	11225	0	Brooklyn	56527.0	2.369863e+07	NY	Kings	36
10	11225	0	Brooklyn	56527.0	2.369863e+07	NY	Kings	36
11	11225	0	Brooklyn	56527.0	2.369863e+07	NY	Kings	36
12	11218	0	Brooklyn	72280.0	3.686880e+07	NY	Kings	36
13	11218	0	Brooklyn	72280.0	3.686880e+07	NY	Kings	36
14	11218	0	Brooklyn	72280.0	3.686880e+07	NY	Kings	36
15	11226	0	Brooklyn	106132.0	3.940860e+07	NY	Kings	36
16	11226	0	Brooklyn	106132.0	3.940860e+07	NY	Kings	36

	ZIP_CODE	BLDGZIP	PO_NAME	POPULATION	AREA	STATE	COUNTY	ST_FIPS
17	11226	0	Brooklyn	106132.0	3.940860e+07	NY	Kings	36
18	11219	0	Brooklyn	92561.0	4.200274e+07	NY	Kings	36
19	11219	0	Brooklyn	92561.0	4.200274e+07	NY	Kings	36
20	11219	0	Brooklyn	92561.0	4.200274e+07	NY	Kings	36
21	11210	0	Brooklyn	67067.0	4.788702e+07	NY	Kings	36
22	11210	0	Brooklyn	67067.0	4.788702e+07	NY	Kings	36
23	11210	0	Brooklyn	67067.0	4.788702e+07	NY	Kings	36
24	11230	0	Brooklyn	80857.0	4.992670e+07	NY	Kings	36
25	11230	0	Brooklyn	80857.0	4.992670e+07	NY	Kings	36
26	11230	0	Brooklyn	80857.0	4.992670e+07	NY	Kings	36
27	11204	0	Brooklyn	77354.0	4.355518e+07	NY	Kings	36
28	11204	0	Brooklyn	77354.0	4.355518e+07	NY	Kings	36
29	11204	0	Brooklyn	77354.0	4.355518e+07	NY	Kings	36
...
516	10019	0	New York	41966.0	1.882838e+07	NY	New York	36
517	10019	0	New York	41966.0	1.882838e+07	NY	New York	36
518	10019	0	New York	41966.0	1.882838e+07	NY	New York	36
519	10065	0	New York	33459.0	1.144258e+07	NY	New York	36

	ZIP_CODE	BLDGZIP	PO_NAME	POPULATION	AREA	STATE	COUNTY	ST_FIPS
520	10065	0	New York	33459.0	1.144258e+07	NY	New York	36
521	10065	0	New York	33459.0	1.144258e+07	NY	New York	36
522	10075	0	New York	25203.0	4.809654e+06	NY	New York	36
523	10282	0	New York	4139.0	1.562845e+06	NY	New York	36
524	10282	0	New York	4139.0	1.562845e+06	NY	New York	36
525	11211	0	Brooklyn	60861.0	4.047088e+07	NY	Kings	36
526	11211	0	Brooklyn	60861.0	4.047088e+07	NY	Kings	36
527	11211	0	Brooklyn	60861.0	4.047088e+07	NY	Kings	36
528	11373	0	Elmhurst	101282.0	4.265486e+07	NY	Queens	36
529	11373	0	Elmhurst	101282.0	4.265486e+07	NY	Queens	36
530	11373	0	Elmhurst	101282.0	4.265486e+07	NY	Queens	36
531	11361	0	Bayside	28496.0	5.016352e+07	NY	Queens	36
532	11361	0	Bayside	28496.0	5.016352e+07	NY	Queens	36
533	11361	0	Bayside	28496.0	5.016352e+07	NY	Queens	36
534	10036	0	New York	23543.0	1.139511e+07	NY	New York	36
535	10036	0	New York	23543.0	1.139511e+07	NY	New York	36
536	10036	0	New York	23543.0	1.139511e+07	NY	New York	36

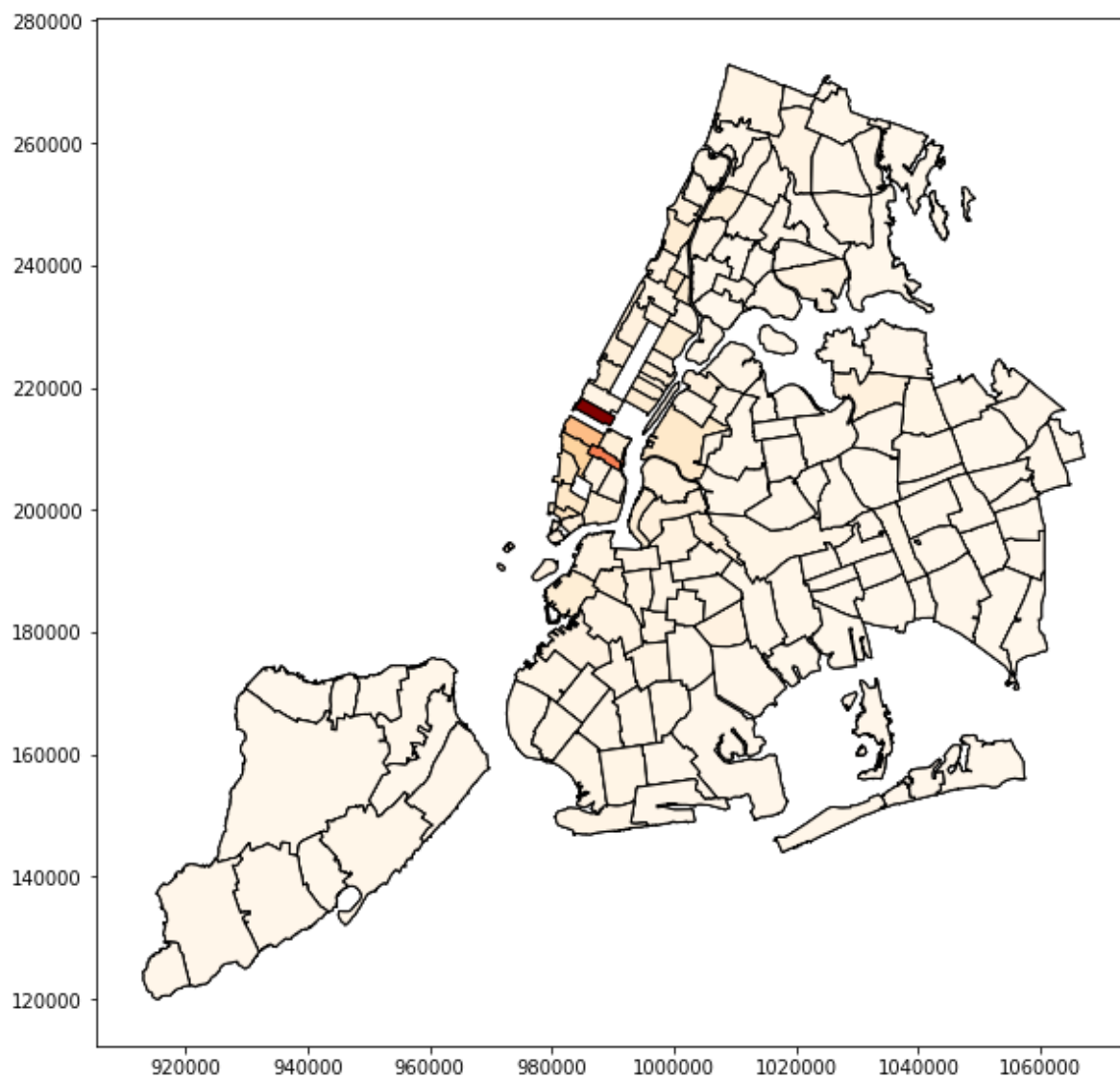
	ZIP_CODE	BLDGZIP	PO_NAME	POPULATION	AREA	STATE	COUNTY	ST_FIPS
537	11414	0	Howard Beach	26148.0	6.392882e+07	NY	Queens	36
538	11414	0	Howard Beach	26148.0	6.392882e+07	NY	Queens	36
539	11414	0	Howard Beach	26148.0	6.392882e+07	NY	Queens	36
540	10310	0	Staten Island	25003.0	5.346328e+07	NY	Richmond	36
541	10310	0	Staten Island	25003.0	5.346328e+07	NY	Richmond	36
542	10310	0	Staten Island	25003.0	5.346328e+07	NY	Richmond	36
543	11249	0	Brooklyn	28481.0	1.777221e+07	NY	Kings	36
544	11249	0	Brooklyn	28481.0	1.777221e+07	NY	Kings	36
545	11249	0	Brooklyn	28481.0	1.777221e+07	NY	Kings	36

546 rows × 21 columns

```
In [989]: fig, ax = plt.subplots(figsize = (10,10))

newmap.plot(ax = ax,
            column = "SALE_PRICE",
            edgecolor = "k",
            cmap = "OrRd"
            )

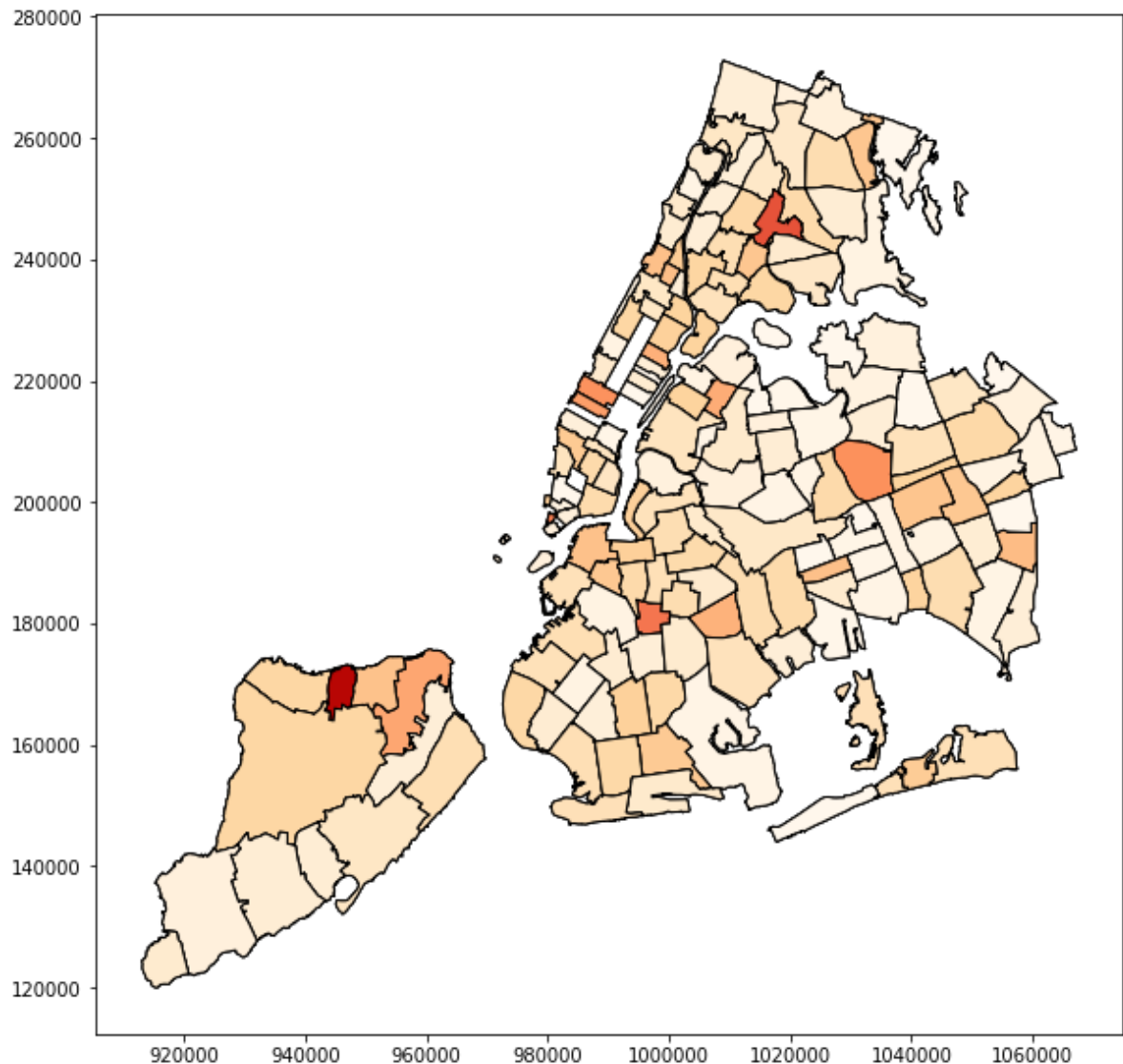
plt.show()
```




```
In [990]: fig, ax = plt.subplots(figsize = (10,10))

newmap.plot(ax = ax,
            column = "Total_Crime",
            edgecolor = "k",
            cmap = "OrRd"
            )

plt.show()
```



Lastly, we imported a shapefile of NYC to better visualize our statistics on a geographical map. From the map showing housing prices, we can see that the Manhattan area is significantly darker, representing higher prices. From the second map showing average total crime rates for 2013-2016, we can see areas where there are more crimes in schools. This map could be useful to visualize regions where school safety is more of an issue and could be addressed.

In conclusion, to recap, we mainly used NYC Open Source as our data source. We read in data for school safety and housing sale prices for three years, 2013-14, 2014-15 and 2015-16. Ideally, we would have analyzed this data over a longer period of time but the accessible data we found was limited so we worked with what we had. We organized the data into various ways with dataframes by year, borough and zip code. The statistics we focused on included Total Crime, Average Total Crime and Sale Prices. We merged our school safety data frame with the housing data to form new dataframes. In the last part, we were able to draw some conclusions from the visualizations we produced.

We did not find a strong correlation with school safety and housing prices for the NYC boroughs. We found that Manhattan had very high housing prices and also higher relative school crimes. However, in general, as shown through the scatter plots which plotted Zip code data for all boroughs and all years, areas with lower crime rates contained higher housing prices. A possible explanation for this could simply be that people want to live in safe school areas for the benefit of their children. A better safety area near schools also could imply higher standard of living since negative activity affects the community less.

In []: