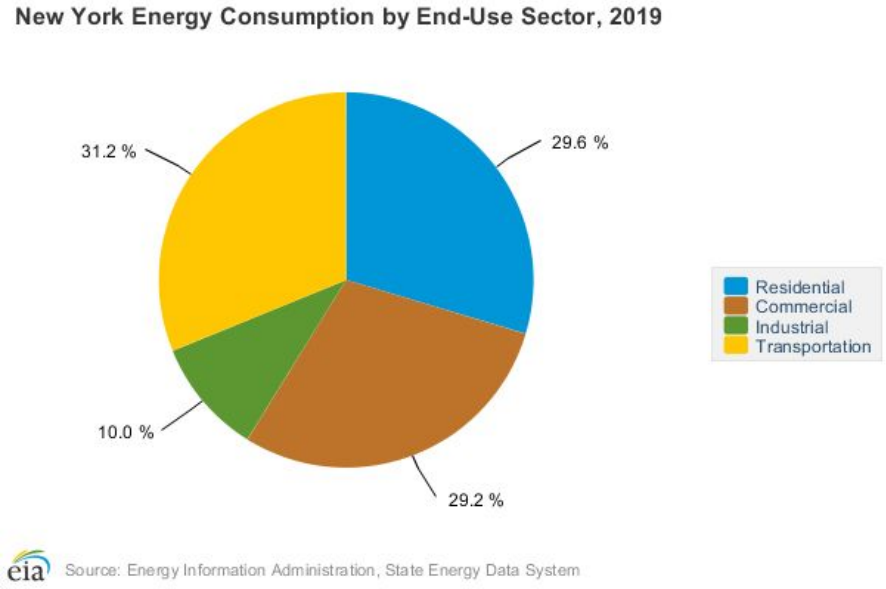
NYC Building Energy Efficiency

Steven Litvack-Winkler (Data Scientist)

Problem Statement

Buildings are a major source of energy consumption in New York. Given New York's goals to greatly reduce greenhouse gas emissions, what aspects are associated with high or low energy efficiency?



Data Source

- ENERGY STAR Portfolio Manager
  - Private buildings > 25,000 square feet
- NYC Open Data aggregates and publishes this information (here)

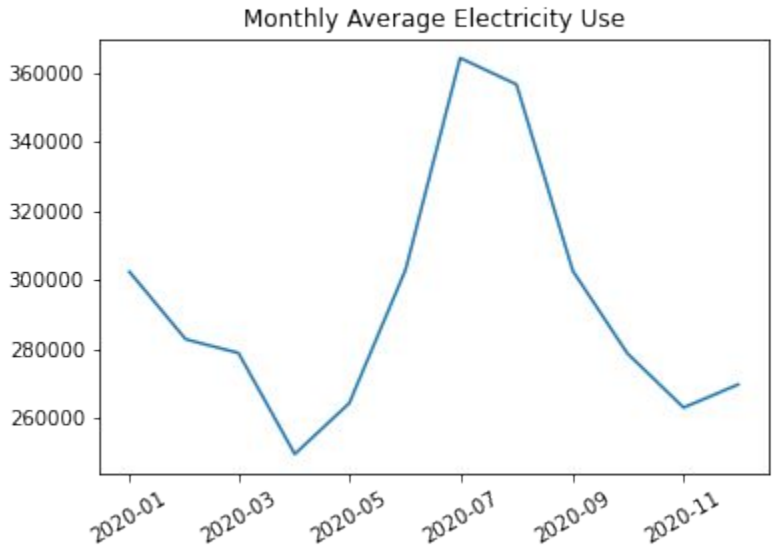
Data Overview

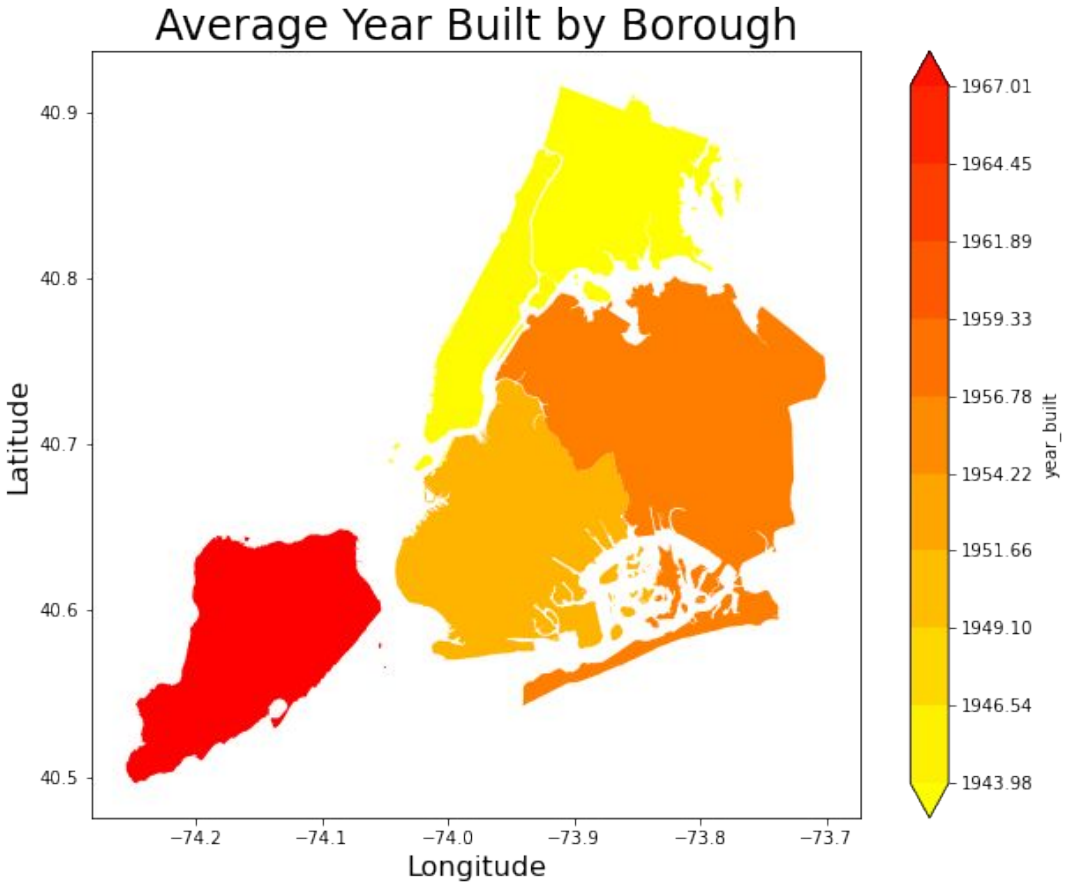
- 28,067 total properties with 16,894 in multifamily housing category
- Identification columns
  - Name, id
- Property information
  - Type, square footage, year built
- Location columns
  - Address, borough, neighborhood
- Energy usage columns
  - Energy use intensity, total emissions
- Messy data with missing values, duplicates, inconsistent field values / names and outliers
  - One rehab / nursing center used 6.4e10 kBtu of steam power in 2020, about equal to an atomic bomb or calories consumed by 10 million people in a year

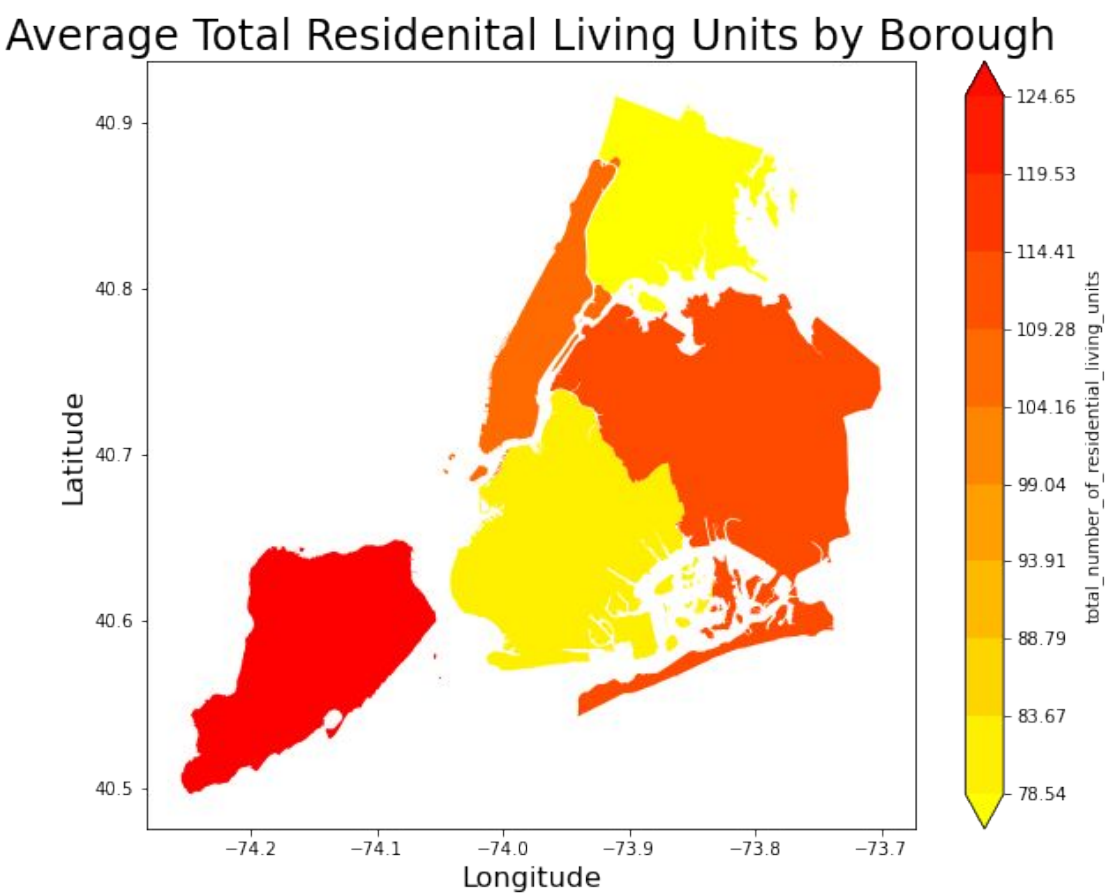
Property Type Comparison

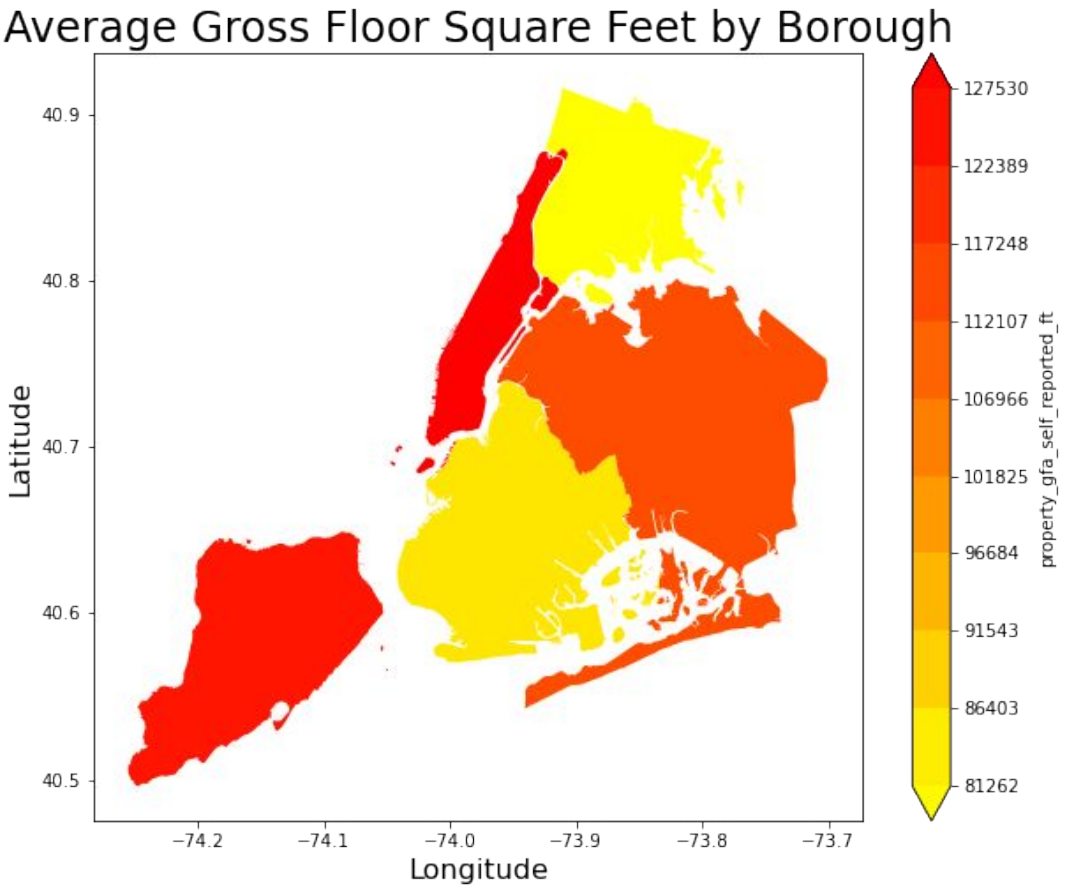
Property Type	Total GHG Emissions	Number of Properties
Multifamily Housing	9616037.2	18702
Office	2909336.4	2338
Hospital (General Medical & Surgical)	859730.0	82
K-12 School	719067.5	1614
College/University	650218.8	301

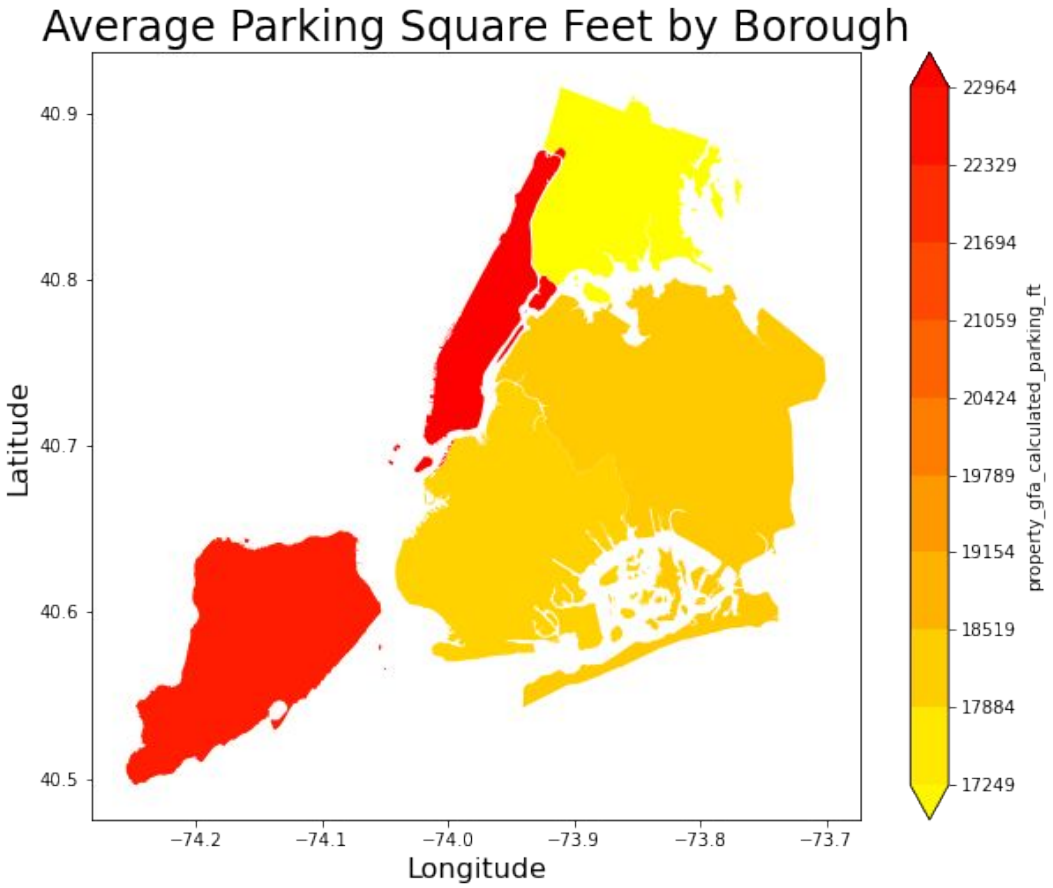
Monthly Electricity Consumption

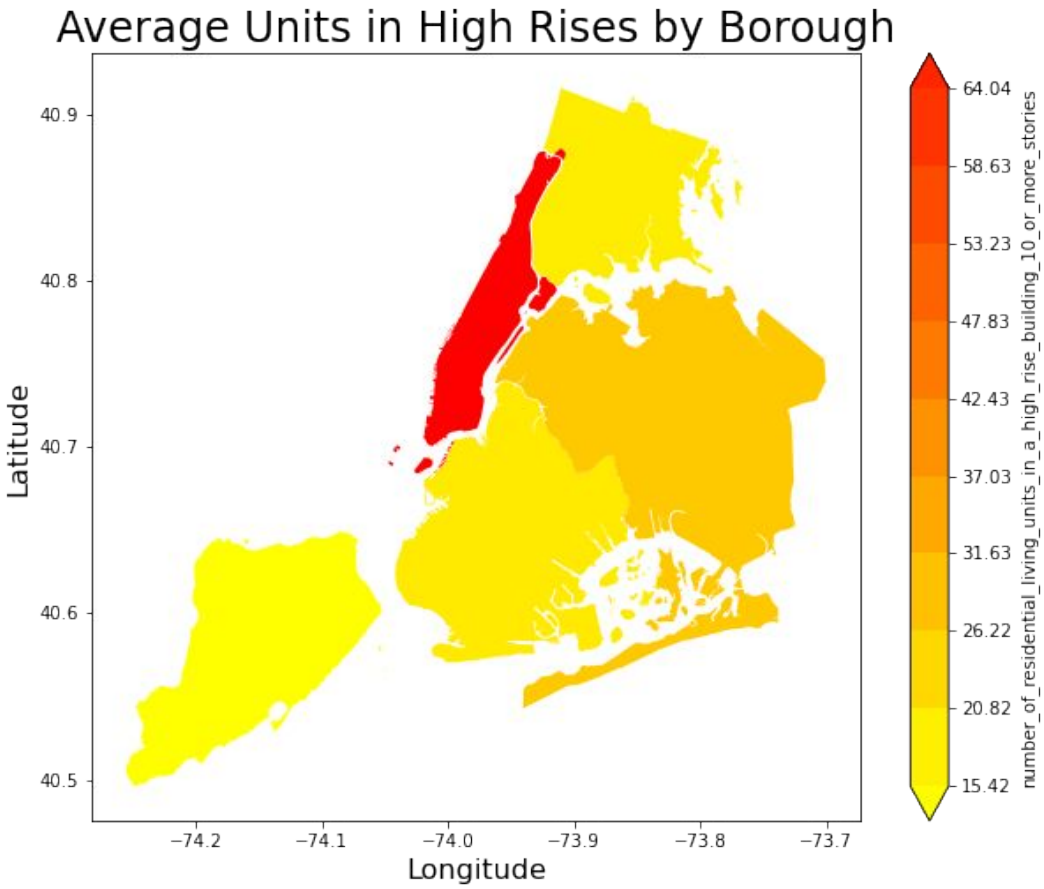



Average Year Built by Borough
Latitude
40.9
40.8
40.7
40.6
40.5
−74.2
−74.1
−74.0
−73.9
−73.8
−73.7
Longitude
1967.01
1964.45
1961.89
1959.33
1956.78
1954.22
1951.66
1949.10
1946.54
1943.98
year_built

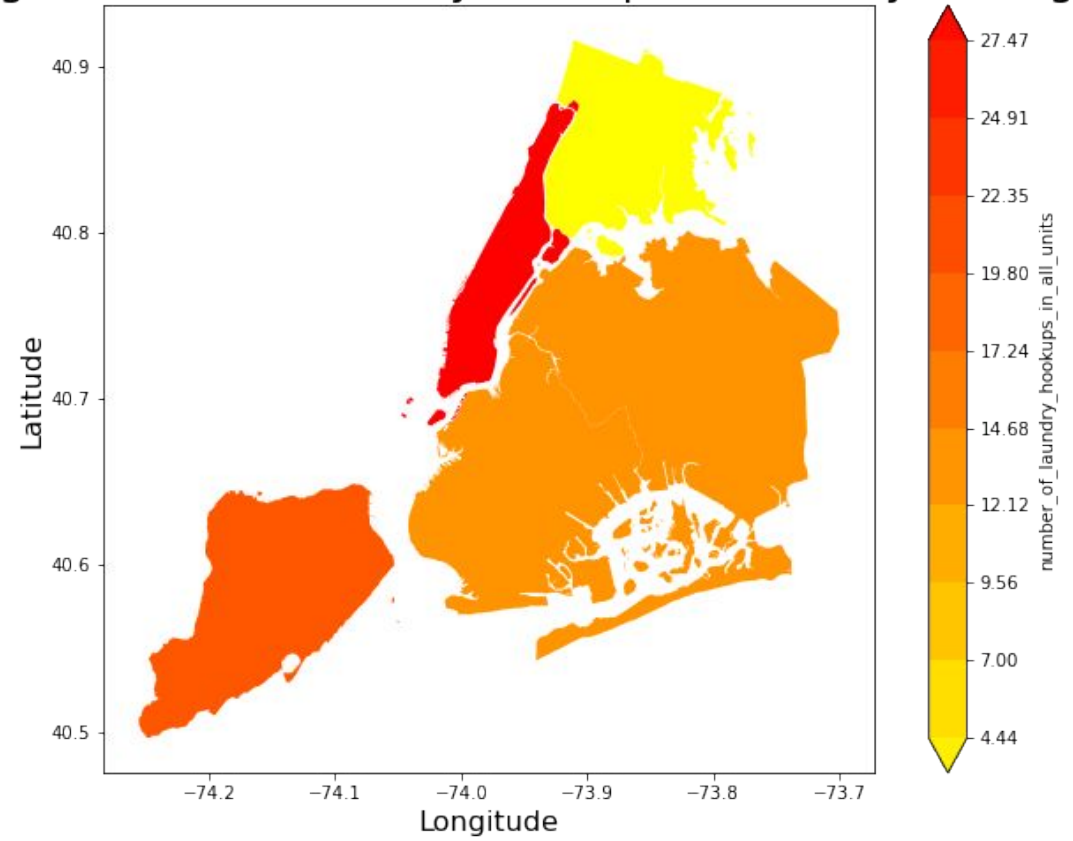

Average Total Residenital Living Units by Borough
Latitude
40.9
40.8
40.7
40.6
40.5
−74.2
−74.1
−74.0
−73.9
−73.8
−73.7
Longitude
124.65
119.53
114.41
109.28
104.16
99.04
93.91
88.79
83.67
78.54
total_number_of_residential_living_units


Average Gross Floor Square Feet by Borough
Latitude
40.9
40.8
40.7
40.6
40.5
−74.2
−74.1
−74.0
−73.9
−73.8
−73.7
Longitude
127530
122389
117248
112107
106966
101825
96684
91543
86403
81262
property_gfa_self_reported_ft

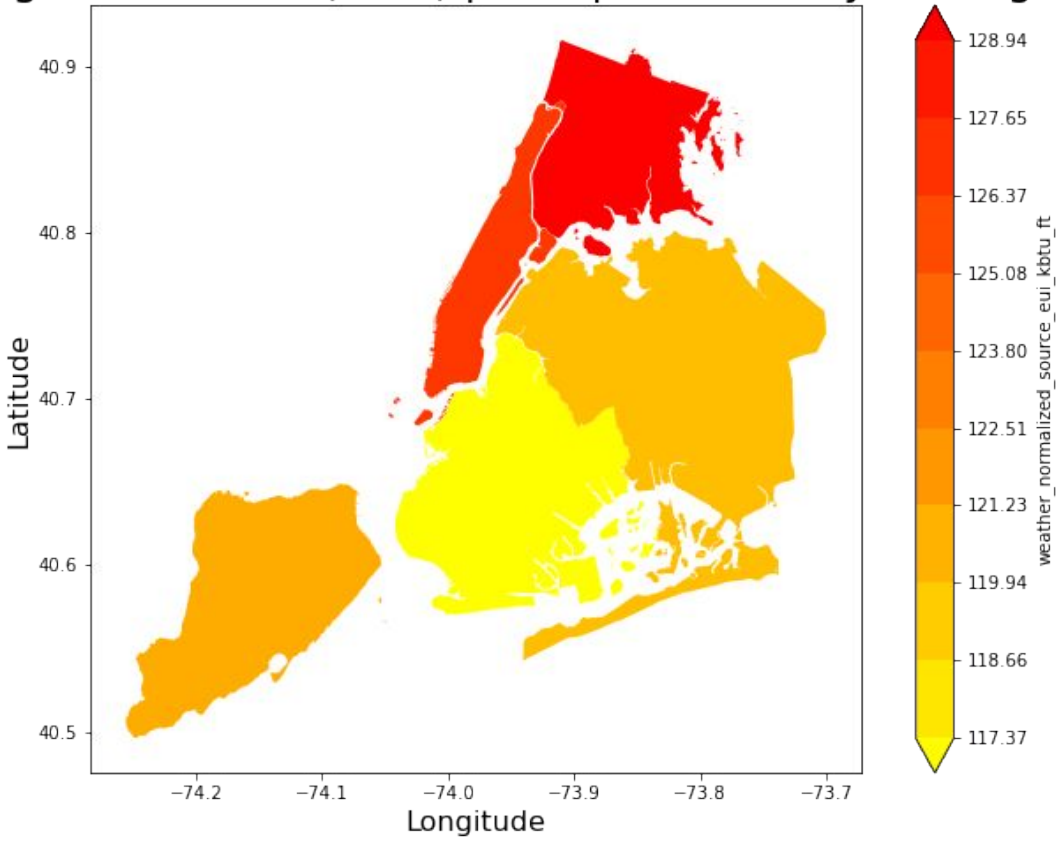

Average Parking Square Feet by Borough
Latitude
40.9
40.8
40.7
40.6
40.5
−74.2
−74.1
−74.0
−73.9
−73.8
−73.7
Longitude
22964
22329
21694
21059
20424
19789
19154
18519
17884
17249
property_gfa_calculated_parking_ft


Average Units in High Rises by Borough
Latitude
40.9
40.8
40.7
40.6
40.5
−74.2
−74.1
−74.0
−73.9
−73.8
−73.7
Longitude
64.04
58.63
53.23
47.83
42.43
37.03
31.63
26.22
20.82
15.42
number_of_residential_living_units_in_a_high_rise_building_10_or_more_stories

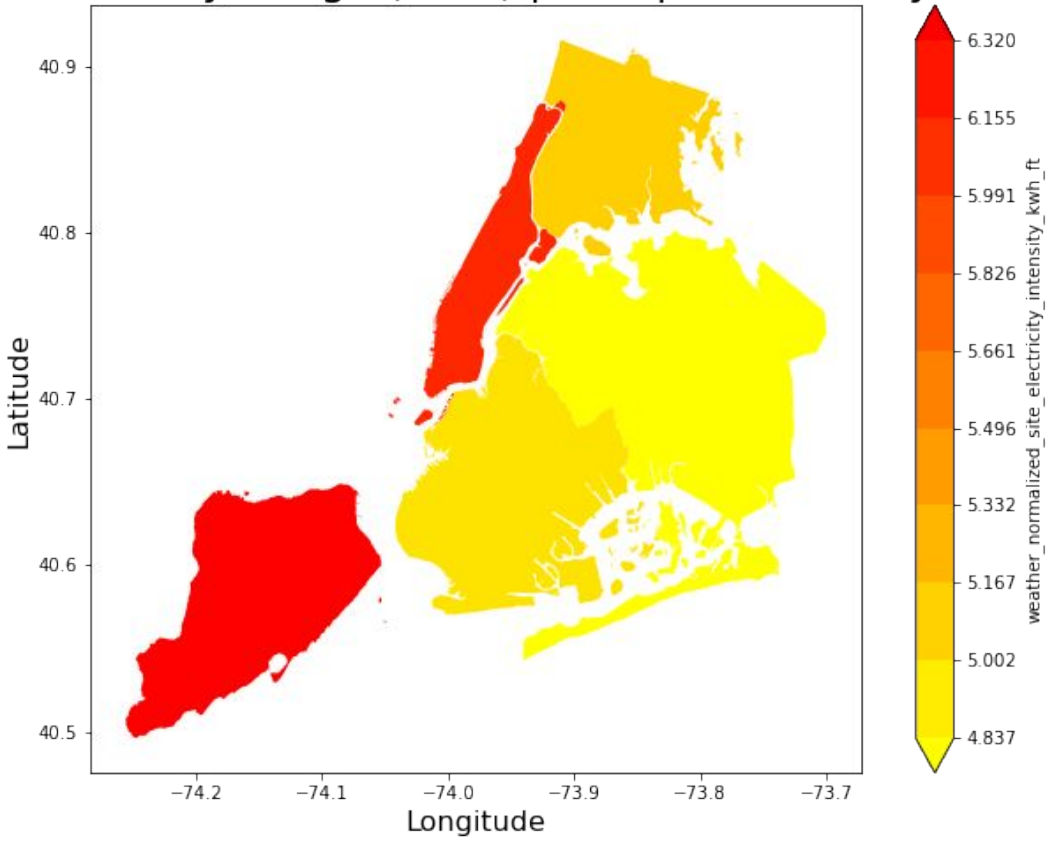
Average Number of Laundry Hookups in Units by Borough



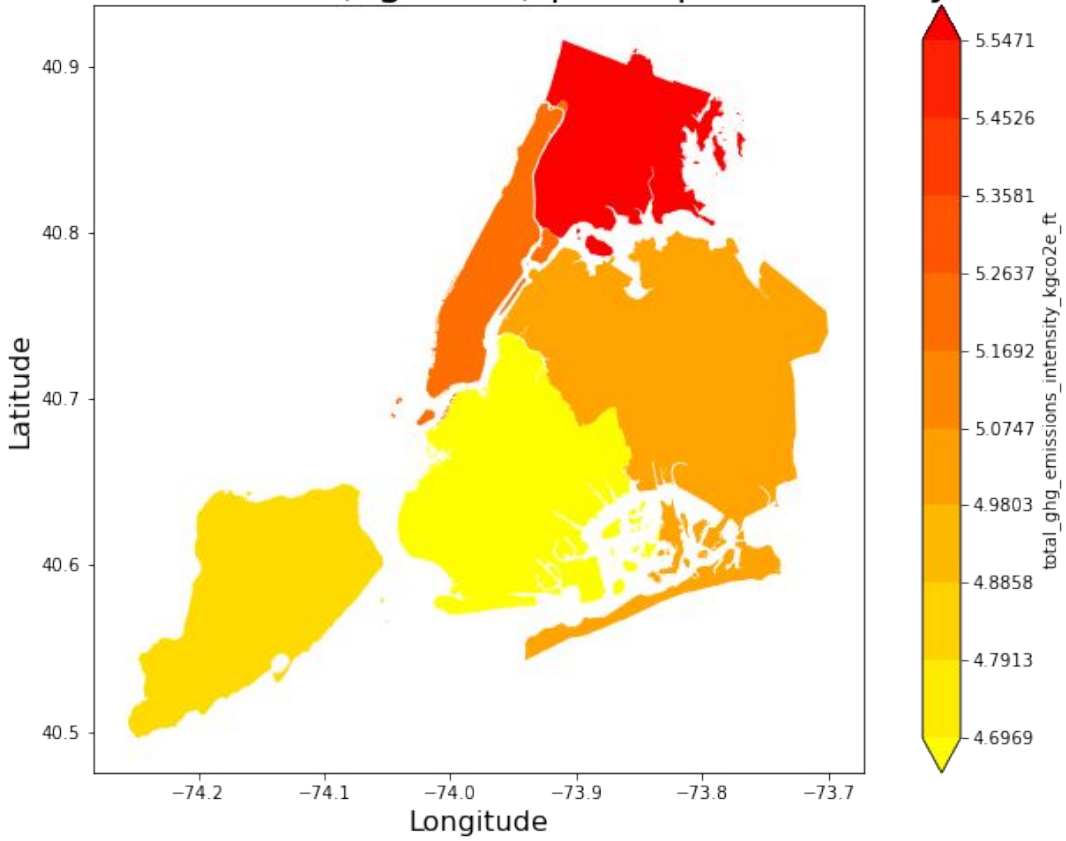
Average Source EUI (kBtu) per Square Foot by Borough



Average Electricity Usage (kBtu) per Square Foot by Borough



Average Total Emissions (kgCO2e) per Square Foot by Borough



Model to predict energy consumption

Target - total greenhouse gas emissions intensity

Features
* borough (with a one hot encoding)
* gross floor square footage
* ratio of square footage for parking
* year built
* occupancy ratio
* government_subsidized_housing flag
* high rise and mid rise flags

Model to predict energy consumption

- Target - total greenhouse gas emissions intensity
- Linear regression
  - Almost all features are statistically significant but R2 is only 0.05

Conclusions / Takeaways

- Greenhouse gas emissions are not easily explained by factors like demographics, location, and size.
- Real world data is messy and collecting / cleaning it is very time consuming.
- Choropleth maps are powerful ways to visualize geographic data.

Future Steps

- Compare to prior years of data, study if buildings become more / less efficient over time.
- Maps at more localized levels.
- Use more powerful models to predict ghg emissions.
- Explore other publicly available data that explains more of the variance in ghg emissions.
- Understand current official websites and reports on NYC environmental building data.