**Predicting NYC Housing Prices**

**Introduction:**

New York City is notoriously known for its expensive real estate market. Across the different boroughs you see a lot of diversity in in NYC is very diverse across the different boroughs and very dynamic from year to year. Our data is originally from the NYC Department of Finance’s record of properties sold in New York City from September 2016 to September 2017 and December 2017 to December 2018.

**Datasets**: <https://www.kaggle.com/new-york-city/nyc-property-sales>

<https://www1.nyc.gov/site/finance/taxes/property-rolling-sales-data.page>

**Goal**:

1. Perform various regression analyses to predict housing prices
2. Identify factors that are most important to determining price
3. Learn Python packages and modeling techniques

**Timeline**:

March 8, 15: Exploratory Analysis

Deliverables: Tableau Visualizations, Correlations, Variable Distributions

Troubleshooting: nested regression model to fill in missing square footage values, get longitude and latitude of addresses

March 22, 29: Classification Building

Deliverables: Linear Regression, Random Forest Regression, Analyze Findings

April 5: Explore More Complex Models

Deliverables: Ridge Regression, ElasticNet?, Neural Net, etc

April 12, 19: Model refinement and interpretation

Deliverables: Gridsearch, Visualizations of Learning Curves, ROC, Accuracy

April 26: Finalize

Deliverables: Slide Deck, Final Predictive Ability

May 3: Practice Presentation

Deliverables: Practice

Meeting Times:

Tuesdays - 6PM