Capstone Project: Project Proposal
Jeffrey Ebert

New York City Property Sales

As the United State's most populous city, every year New York City undergoes thousands of property transactions. These types of sales range from classes of buildings such as homes, apartment buildings, hotels, warehouses and more. As either a buyer or a seller involved in New York City real estate, it would be important to know the rates at which properties are sold and how these evaluations are determined based on a building's features.

This project analyzes a dataset from September 1st 2016 to August 31st 2017 that includes all of the property transactions throughout the five boroughs of New York City in this timeframe along with the property's accompanying location, and characteristics about these properties. The data used was developed by the City of New York and hosted on Kaggle.com as a public dataset.

The dataset contains 84,548 data entries of property transactions, all providing information regarding the type of building it is, when it was built, where it is located, the amount of units it has, its square footage, when it was sold, and the price that it was sold for. The main purposes of the project are to investigate the varying features of the transactions and discover the most important factors in determining a property's price as well as comparing and contrasting the various types of transactions made across New York's five boroughs. The project will analyze the important qualitative characteristics of a property sale as well as finding the underlying correlations between its numerical features and its price while building predictive models based on all of these variables to determine a price.

A potential buyer or seller of New York City properties in the immediate future could use this project to gain a more nuanced understanding of how certain types of buildings are evaluated. This could help someone list a more competitive price if they are selling their property or save money for someone looking to buy an unfairly priced listing.