Ka Shing Wai, Weicheng Huang, Zhuoying Cai
Mr. Long
CSC 59970
03/17/2019

Project Proposal – NYC Housing Prices Analysis

GitHub Repo: https://github.com/zhuoyingcai/NYC-Housing-Prices-Analysis

New York City is known as the financial and cultural center. In this fast-growing city, it becomes one of the most prosperous real estate markets in the United States. New York City's real estate market has been grown non-stop for years. New York City has five boroughs, and each area has different housing prices. Among these five boroughs, Manhattan has always been the first choice for investors. However, in the past years, other boroughs have also increased popularity. In this project, we will be focusing on how the housing prices were changing in the five boroughs of NYC for the past years.

Potential Questions
1. How were the housing prices changed among the five boroughs over the past years?
2. What are some major causes that affect housing prices?
3. Which borough of New York City has the highest housing turnover rate?
4. What is the price of a house based on given criteria (location, gross area, year built, etc)?

Independent variables: Sale Price
Dependent variables: Borough, Neighborhood, Building Class Category, Tax Class at Present, Zip Code, Gross Square Feet, Year Built, Sale Date, Population, Households with children under 18 years old, Population aged 65+, Median household income, Unemployment rate, Serious crime rate, Students performing at grade level in math, Mean travel time to work.

Datasets
- *Annualized Sales Update* – yearly sales information of properties sold in New York City.
  https://www1.nyc.gov/site/finance/taxes/property-annualized-sales-update.page
- *Rolling Sales Data* – lists properties that sold in the last twelve-month period in New York City for all tax classes.
  https://www1.nyc.gov/site/finance/taxes/property-rolling-sales-data.page
- *New York City Neighborhood Data Profiles* – critical for understanding local housing and demographic trends, identifying community needs, and informing policy conversations.
  http://furmancenter.org/neighborhoods

The major theme of the project is about how the housing prices were changing in the five boroughs of NYC for the past years. The first question targets this theme directly; we are going to analyze the datasets, Annualized Sales Update, and Rolling Sales Data, and find out how the housing prices varied and changed in each borough. Moreover, it is significant to discover what are the factors that contribute the most at affecting the housing prices. The second and third questions are related to the causes of change in housing prices. There are many major factors, such as population, crime rate, schools, etc. that influence the housing price of an area. In addition to the first two datasets, New York City Neighborhood Data Profiles provide detailed information about all neighborhoods in NYC. This dataset is helpful for us to have further analysis and a deeper understanding of the relationship between major factors and housing prices. Furthermore, for the last question, we are going to build a model that can offer people some insights about the housing price based on given criteria. The model can indicate and provide suggestions about whether if a property is worth investment for the price.