

Real Estate Price Prediction and Purchase Recommendation

APMA E4990 Sec 002 Introduction to Data Science in Industry

By Jingya Yan and Yuxuan Liu

Tools

Python, Jupyter Notebook, Javascript

Data source

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

Project Description

When new real estate buyers decide to purchase their first homes, they usually rely on brokers for pricing information and property recommendation. This may result in an informational gap that puts clients at a disadvantaged position. They run the risk of perceiving the real estate market from a limited, and potentially biased perspective. Through collecting and processing a large dataset of property sales by appropriate machine learning techniques, we aim to supply an analytical and objective view to interested real estate buyers of New York City. Clients will be able to find out the price of a property given a series of features and receive recommendations on the best valued properties based on a price range. They can compare our results with information obtained from real estate brokers. Our goal is to make real estate buyers more informed about the market and make the best decisions for themselves.

Audience

Real estate buyers who already obtained details of the real estate they are interested in. With the details they can know whether the price provided by brokers is underestimated or overestimated.

Our project is also useful for brokers who want to build a better trading strategy with clients.

Algorithms

Regression:

Based on the data provided(i.e. neighborhood, building class category, zip code and etc.) combined with the real sales price in the past, we are going to give a prediction for the sales price under the given variables by using regression models.

Recommendation Engine:

We will utilize a prescriptive machine learning model to produce a list of properties whose market values lie in a specific price range. The market values of these properties are obtained from our previous regression model.

User Interaction

We are going to build an interactive website for users to input their preferences. A Python kernel will be running behind the scene to process user inputs with machine learning algorithms. When the algorithms are completed, price predictions and property recommendations will be shown to users.

Reference

Link for property prices in NYC:

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