Analysis of the Housing Market in Two US Cities from 2020-2022

Real Estate Investment Group

**Data Science Capstone Project   
Launch Report**

Date:

09/27/2022

Team Members:

Name: Nikhil Muthuvenkatesh

Name: Seyi Oyesiku

Name: Luqing Qi

Name: Rohan Ukkalam

**The System/Product**

**System/Product Name:** East Coast Housing Price Predictor

**Introduction:**

Since the COVID-19 pandemic in 2020, real estate pricing has been quite volatile. The initial lockdowns led to a large drop in property values across the country, but shortly thereafter there was a huge spike that has continued up until now. This project will implement a model that is trained with real estate data over the last two years in order to predict the future prices for housing. The goal of this capstone project is to correctly predict when someone that is interested in purchasing real estate should make a purchase of a home in a city on the East Coast, such as Philadelphia, Boston, or New York.

The motivation behind this project comes from the team members. We are all graduate students who are interested in purchasing real estate in the future, and we would like to know what to expect from this market in the future. It could also potentially be useful for banks that give out home loans, mortgage companies, and for escrow companies. In addition, anyone that is interested in purchasing a home in one of the cities that we use for analysis in the next couple years would find this information useful.

**Highlighted Features:**

* Model that is able to predict house pricing in these two cities within the next couple months
* The training set will include real estate data from both cities from the time period of 2020-2022
* End product will include visualizations for laymen to better understand the trends for the past few years
* End product will include several areas/location of interest

**Sponsor or Proxy User:**

Potential users include mortgage companies, banks, investors, and people searching for homes.

**Issues:**

Potential issues in data acquisition include finding credible and recent data on real estate pricings. Team has intermediate understanding of the subject, and thus might run into issues in the interpretation of insights should enough data be gathered.

**The Team**

**Team Name:** Not sure about our group number, but we will be the Real Estate Investment Group

**Team Members and their specialties:**

Nikhil Muthuvenkatesh- Intermediate level of programming in Python, background in sales/engineering to help market the final product, some knowledge of the real estate market

* Would like to work on the data acquisition, pre-processing, and the exploratory data analytics

Rohan Ukkalam - Intermediate level of programming in Python, some experience with programming in Java and C.

* Would like to work on data acquisition, pre-processing and cleaning, exploratory data analytics, and report writing.

Seyi Oyesiku - Intermediate level python programmer, my background is networking and supply chain logistics. Other familiar programming languages include C and Python.

* Would like to work on data acquisition, pre-processing and cleaning, analytics, presentation and report writing.

Luqing Qi - Intermediate level Python programmer, my working experience coming from the healthcare industry.

* Would like to work on data acquisition, pre-processing and cleaning.

**Team Communication:**

Team communication will happen with a combination of in-person and Zoom meetings. Since we all have different schedules as to when we are on campus, we must have some online meetings. We also all live in different parts of the city, so it may not always be feasible to meet in person. We also have a text group chat, and a Google Drive folder of all files that we can collaborate together on.

**Team Issues:**

Potential issues include scheduling conflicts, as half of the team are online students, and one member lives far away enough from the Drexel Campus that it is not entirely feasible for in person meetings.

Plan

The initial draft of the plan is to first locate credible sources for the information we seek. We will first look at census information and, if they exist, public information from the larger realtors in the areas we are focusing on. Most likely for the later sources data harvesting via APIs will be required, so we will research that as well. After finding several credible sources, we will then begin acquisition, and afterwards processing.

Table of Contributions

The table below identifies contributors to various sections of this document.

|  | **Section** | **Writing** | **Editing** |
| --- | --- | --- | --- |
| **1** | **Project** | Nikhil, Rohan | Seyi |
| **2** | **Team** | Seyi | Nikhil |
| **3** | **Plan** | Luqing | Rohan |

**Grading**

The grade is given on the basis of quality, clarity, presentation, completeness, and writing of each section in the report. This is the grade of the group. Individual grades will be assigned at the end of the term when peer reviews are collected.