Stephanie Cruz

Lab Assignment 6

Contents

[Introduction: 2](#_Toc55593280)

[About the Data: 2](#_Toc55593281)

[Data Cleaning: 2](#_Toc55593282)

[Exploratory Data Analysis: 2](#_Toc55593283)

[Investment Strategy Analysis: 5](#_Toc55593284)

# Introduction:

# About the Data:

The provided base data set was from Zillow and included housing values. The data includes information about housing prices for over 30,000 zip codes. The pricing houses are for each month starting January 1996 through March 2020. Additional data was used from the census bureau, including information on the number of employees in each zip code as well as the annual payroll in each zip code.

# Data Cleaning:

Unneeded columns were dropped from the dataset, however no other data cleaning was performed.

# Exploratory Data Analysis:

An exploratory data analysis was completed for the following four Arkansas Metro areas: Hot Springs, Little Rock, Fayetteville and Searcy. For each area, the average housing value was plotted over the time from January 1996 through March 2020. Below is the graph for Hot Springs:

Chart, line chart

Description automatically generated

In this graph we can see a dip in housing prices from shortly after 2014 through 2012, and then housing prices sharply increased.

Chart, line chart

Description automatically generated

In Fayetteville, there was also a dip in housing prices, however it started later than the Hot Springs dip and didn’t last as long before housing prices increased again.

Chart, line chart

Description automatically generated

Little Rock also experienced a dip in housing prices; however the dip was both shallower and shorter before increases in home prices resumed.

Chart, line chart

Description automatically generated

In Search, there was a sharp dip as in Hot Springs and then recovery. However, Searcy shows the beginning of a dip in housing prices in the current market, the only one of the four metro areas to show that trend.

Chart, line chart, histogram

Description automatically generated

The four metro areas are shown together above. Solely going on the information in the graph, it is advised to invest in the Little Rock area. Both Hot Springs and Fayetteville showed larger dips, and while Searcy shows a more stable line, there is a concerning dip in the current market. Due to its stability and steady increase in value, Little Rock is the area that should be invested in.

# Investment Strategy Analysis:

In order to determine which three zip codes Syracuse should invest in, the following strategy was used: It was determined that the REIT would invest in properties with the intent to sell the shares in those properties in five years’ time. Additionally, the REIT wants to invest in areas where there are a large number of employees as well as high payroll numbers to ensure there are buyers for the properties the trust is purchasing. The following process was used to determine which zip codes should get the investment:

1. For each zip code in the Zillow base data set, each zip code was run through prophet to project the housing value in five years.
2. The base price was pulled from the Zillow base data set (with the base price defined as the price on 3/31/2020, the price the REIT would buy the property at)
3. The mean absolute error was subtracted from the projected value (value five years from now)
4. The base value was subtracted from the future value calculated in [3]
5. To ensure the REIT is investing in locations with lots of employees and high salaries, zip codes were graphed and the top five zip codes that exist in areas where the annual payroll of the zip code was over 1 million dollars where pulled.

Chart, scatter chart

Description automatically generated

Above is the graph showing a scatterplot of each zip code plotted with the increase in home value on the y-axis and the annual payroll of the zip code on the x-axis. The zip codes the REIT is targeting are in the middle section, where annual payroll hits 1.0 and higher. Below are the five zip codes that the REIT is targeting:



It can be seen above that the zip codes we are targeting are on the west coast: three in California and two in Washington. To narrow down our list from five to three, we will look for the zip codes with the highest R-values, which are 98004, 94107, and 98052.

Zip code 98004 is Bellevue, WA, and the projected increase in home value in that zip-code is almost 1 million dollars. Zip code 94107 is in San Francisco, CA and the projected increase in home value is three quarters of a million dollars, and the third zip code is 98052, Redmond, WA and the projected increase in home value in that zip code is around four hundred thousand dollars.