Lab 6

Objective

In this report, a recommendation is provided to the Syracuse Real Estate Investment Trust (SREIT) for the best zip codes in the United States to invest in housing. Using data from Zillow, the zip codes with the highest increases in home value homes were predicted by using a time series forecasting procedure called Prophet.

The data was first examined to identify market changes in cities in Arkansas and their forecasted growth until 2025. It was then used to forecast growth across the U.S. to identify cities with the highest gross and percentage increases. The forecasting procedure resulted in a variety of results and are used to outline 3 investment recommendations.

Load and clean Datasets

The Zillow Home Value Index (ZHVI) was loaded to a Pandas data frame for analysis. This initial data set contains a list of over 30,000 zip codes and includes what Zillow defines as the "typical home value" from 1996 to 2020. Because Prophet can handle missing data, no initial cleansing was needed.

Source: https://www.zillow.com/research/data/

Analysis

The following graph shows the mean housing prices per month from 1996 until 2020 across the United States. Prices dropped significantly during the 2008 housing crisis and only recently recovered to pre-2008 values in recent years.
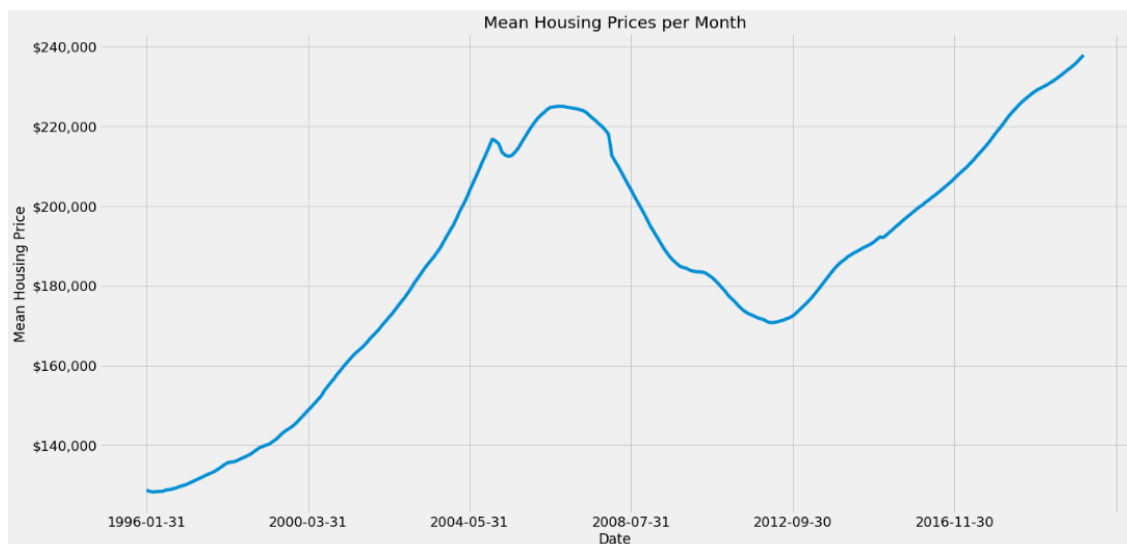
The data can additionally be analysed at a granular level including specific states, counties, and metro areas. Here, the metro areas of Fayetteville-Springdale-Rogers, Jonesboro, Memphis, and Searcy in the state of Arkansas are shown in a time-series graph.
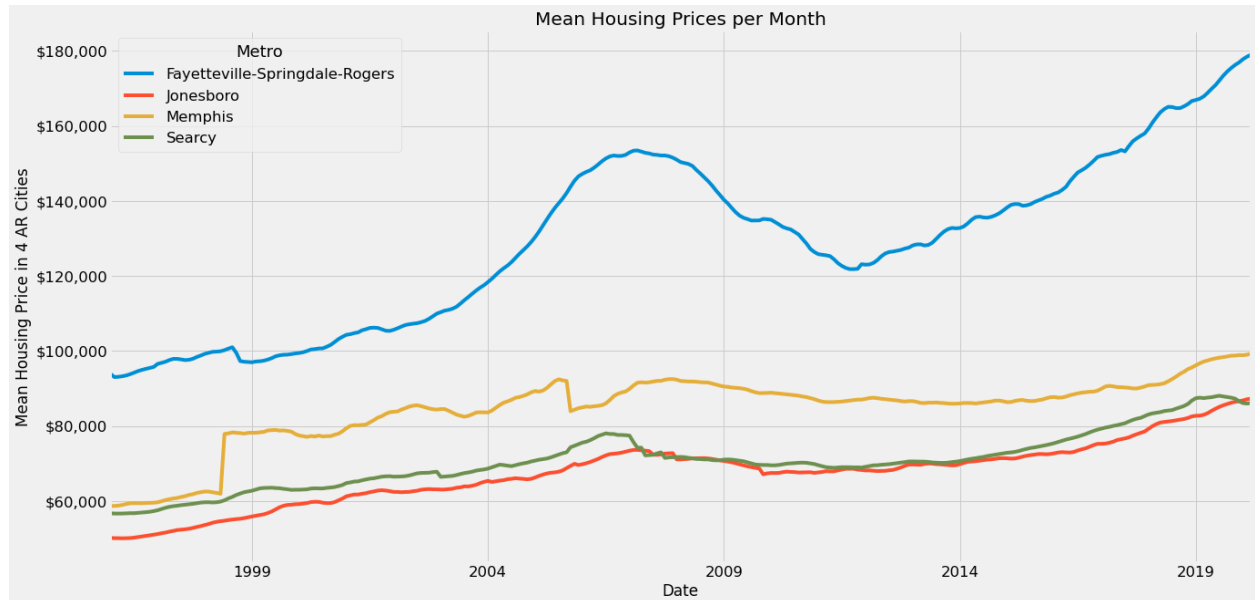
Fayetteville clearly showed the greatest downswing in housing prices during the 2008 crisis and is likely to be the most volatile of the 4 metro areas. To take a closer look at how these past trends can predict future trends, the mean prices for each metro area are forecasted for the next 60 months using Prophet.

	Fayetteville-Springdale-Rogers	Jonesboro	Memphis	Searcy
Predicted 2025 Value	$155,634	$102,888	$110,394	$105,404
Total Growth 2020-2025	$16,256	$16,290	$12,270	$16,124
Percent Growth 2020-2025	11.66%	18.81%	12.51%	18.06%

Jonesboro and Searcy are similarly forecasted to yield the highest percentage growth between 2020 and 2025. Despite having a similar total growth increase as Fayetteville, Jonesboro and Searcy have a lower housing cost in 2020 and would be wiser choices for investment based on this forecasting data.

Modelling

The data was then used to predict housing prices across the U.S. Due to limitations in time and computing efficiency, the initial 30,000+ zip codes were cut down to the 3,000 zip codes with the highest percentage growths in the last 5 years. These areas are more likely to produce high forecasted value increases between 2020 and 2025.

The prophet code used in this report was leveraged from a Medium article *Forecasting National Home Values through Time Series Analysis* by Dr. Wyatt Sharber.

Source: https://medium.com/@wvsharber/forecasting-national-home-values-through-time-series-analysis-924ac911c5a4

```
1.  def make_dict(df, zipcode):
2.      """This function returns a summary dictionary from the combined dataframe for a giv
    en zipcode. This dictionary
3.      can be appended to a list that will be used to make a final dataframe."""
4.      pct_change_1year = df.loc[(df['ds']>'2021-03-31')
5.                                  & (df['ds']<'2021-05-
    01')]['pct_change'].values[0]
6.      gross_profit_1year = df.loc[(df['ds']>'2021-03-31')
7.                                  & (df['ds']<'2021-05-
    01')]['gross_profit'].values[0]
8.      pct_change_5year = df.loc[(df['ds']>='2025-03-31')]['pct_change'].values[0]
9.      gross_profit_5year = df.loc[(df['ds']>='2025-03-31')]['gross_profit'].values[0]
10.     estimate_1year = df.loc[(df['ds']>'2021-03-31')
11.                                 & (df['ds']<'2021-05-01')]['y'].values[0]
12.     upper_1year = df.loc[(df['ds']>'2021-03-31')
13.                                 & (df['ds']<'2021-05-
    01')]['yhat_upper'].values[0]
14.     lower_1year = df.loc[(df['ds']>'2021-03-31')
15.                                 & (df['ds']<'2021-05-
    01')]['yhat_lower'].values[0]
16.     estimate_5year = df.loc[(df['ds']>='2025-03-31')]['y'].values[0]
17.     upper_5year = df.loc[(df['ds']>='2025-03-31')]['yhat_upper'].values[0]
18.     lower_5year = df.loc[(df['ds']>='2025-03-31')]['yhat_lower'].values[0]
19.     zipcode_dict = {'Zipcode': zipcode,
20.                     'estimate_1year': estimate_1year,
21.                     'upper_1year': upper_1year,
22.                     'lower_1year': lower_1year,
23.                     'pct_change_1year': pct_change_1year,
24.                     'estimate_5year': estimate_5year,
25.                     'upper_5year': upper_5year,
26.                     'lower_5year': lower_5year,
27.                     'gross_profit_1year': gross_profit_1year,
28.                     'pct_change_5year': pct_change_5year,
29.                     'gross_profit_5year': gross_profit_5year}
30.     return zipcode_dict
```

```
31.
32. def prophet_master(df, start, stop):
33.     """Puts together all the previous functions into one function that will return a li
     st of dictionaries with the
34.     results of the prophet predictions. Can control with start/stop."""
35.     preds = []
36.     for i in range(start, stop):
37.         zipcode = get_zipcode(df, i)
38.         forecast = Prophet_analysis(one_zipcode(df, i))
39.         combined = combine_dataframes(one_zipcode(df, i), forecast)
40.         preds.append(make_dict(combined, zipcode))
41.     return preds
```

Model Results

The prophet results have been summarized to show the top zip codes by total gross increase and percentage increase from 2020-2025.

Highest Gross Profit

Zipcode	State	CountyName	gross_profit_5year
90210	CA	Los Angeles County	$ 2,562,837
94027	CA	San Mateo County	$ 2,367,465
90020	CA	Los Angeles County	$ 2,099,482
90402	CA	Los Angeles County	$ 1,961,700
90212	CA	Los Angeles County	$ 1,656,484

These five zip codes in California provided the highest gross profit after 60 months. However, it is often suggested to diversify investments. Should a situation happen that causes housing prices in California specifically to suddenly fall, an investment in other areas of the country would help curb this loss.

Highest Gross Profit outside of California

Zipcode	State	CountyName	gross_profit_5year
98039	WA	Seattle-Tacoma-Bellevue	$ 1,258,495.00
98004	WA	Seattle-Tacoma-Bellevue	$ 950,437.60
98004	WA	Seattle-Tacoma-Bellevue	$ 700,572.10
98005	WA	Seattle-Tacoma-Bellevue	$ 644,921.40
59716	MT	Bozeman	$ 619,625.80

Outside of California, four zip codes in the Seattle area of Washington state are forecasted to provide the highest gross profit in 60 months. The last is in Montana and is the only zip code so far seen outside of the West Coast.

Highest Percentage Increase

Zipcode	State	CountyName	pct_change_5year	gross_profit_5year
48238	MI	Wayne County	102.38%	$ 26,851.48
25160	WV	Clay County	100.81%	$ 52,518.96
41743	KY	Perry County	96.30%	$ 40,284.01
75247	TX	Dallas County	85.85%	$ 130,720.89
30314	GA	Fulton County	85.567%	$ 157,909.11

Fortuitously, the zip codes that provided the highest percentage change are all outside of the western half of the contiguous United States. Despite having lower gross profits, it is important to note that housing costs in these zip codes have much lower costs and thus require a smaller initial investment.

Comparing the results graphically between all these zip codes do not provide prominent differences that are not already clear in the table summaries. An example of the graphed results is shown below for Beverly Hills, CA:

Beverly Hills, CA 90210 Projected Growth 2020-2025
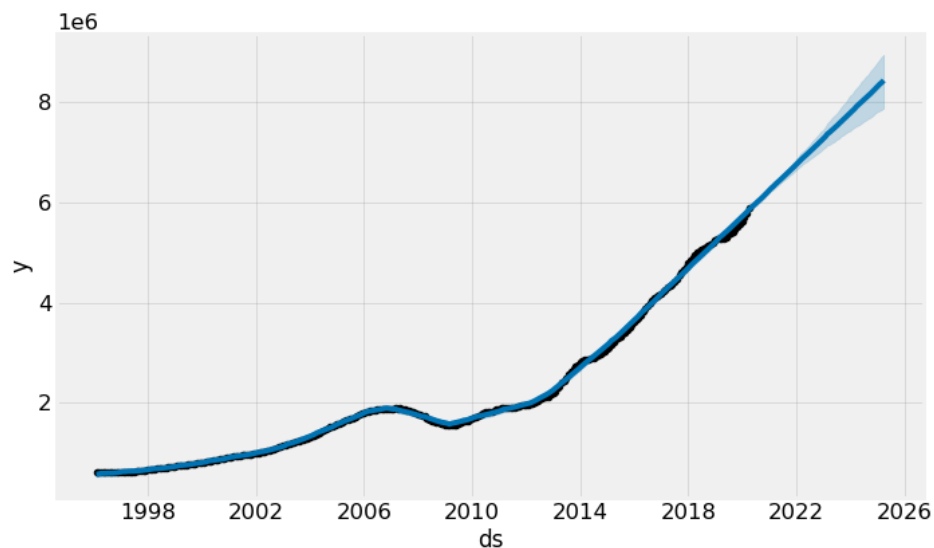
In five years, an investment in Beverly Hills is forecasted to yield a profit of $ 2,562,837.The forecast includes an upper and lower limit of $1.4 to $3.6 million respectively.

Based on the forecasting results, the following recommendations have been made for investment:

Recommendation 1: Huge Investment, Maximum Gross Profit

City and Zip Code	Cost in 2020	Gross Profit in 2025	Percentage Increase
Beverly Hills, CA 90210	$5,902,641.00	$ 2,562,837	43.41%
Atherton, CA 94027	$6,661,460.00	$ 2,367,465	35.54%
Los Angeles, CA 90020	$4,013,833.00	$2,099,482	52.31%

Investing in these California zip codes will provide the maximum projected gross profits in 2025. However, this investment plan requires a massive initial investment of several million dollars per house and includes a great deal of risk as they are all in the same state.

Recommendation 2: Diversified Investment, Not-as-big Gross Profit

City and Zip Code	Cost in 2020	Gross Profit in 2025	Percentage Increase
Beverly Hills, CA 90210	$5,902,641.00	$ 2,562,837	43.41%
Medina, WA 98039	$2,999,237.00	$ 950,437.60	42.96%
Bozeman, MT 59716	$1,312,365.00	$ 619,625.80	47.21%

This recommendation provides lower gross profits than recommendation 1 but is less risky as the funds are invested in different states and regions. Additionally, these zip codes require a smaller initial investment on average.

Recommendation 3: Small Investment, Big Percentage Profit

City and Zip Code	Cost in 2020	Gross Profit in 2025	Percentage Increase
Detroit, MI 48238	$26,228.00	$ 26,851.48	102.38%
Pond Gap, WV 25160	$52,097.00	$ 52,518.96	100.81%
Atlanta, GA 30314	$184,547.00	$ 40,284.01	96.30%

The last recommendation requires the lowest investment of the 3 and the highest percentage increase. While having much lower gross profits, investing in these zip codes is projected to provide nearly dollar for dollar returns after 5 years.

In summary, each investment recommendation has its strengths and weaknesses and will depend on additional information such as SREIT's initial funds, budget, and timeline for expected returns. It is even possible that a mix of three zip codes from different recommendations is the best plan.

Without knowing this additional information and for the purposes of providing a recommendation of three zip codes in this report, **Recommendation 2** is likely the most robust investment plan. It provides a mix of regions for a diversified investment while also returning high gross profits in five years. Overall, this is the safest and best bet.

Conclusion

Three investment plans for housing are recommended to the Syracuse Real Estate Investment Trust (SREIT). Each plan varies by initial investment costs, projected percentage increase, total gross profits, and diversity in regions.

The recommendations are based on forecasting results from Zillow data. The initial set of 30,000 zip codes was initially cut down to the top 3,000 zip codes that showed the highest housing value increases over the last five years. A forecasting procedure called Prophet was then used to predict the housing prices in these zip codes for the next 60 months. The results were summarised to show the zip codes with the highest gross profit and percentage increases.

Overall, there are strengths and weaknesses associated with each investment plan and picking the right one will depend on more information such as SREIT's budget and timeline for expected returns in investment. The zip codes 90210, 98039, and 59716 are suggested for investment as it provides a variety of locations and high gross profits.