**Scott Snow**

**IST 718 – Lab 2**

**May 10, 2019**

**Data Outline and Sources**

* Housing Data, from Zillow, April 1996 – March 2019
* Each row represents a Zip Code in the US, contains ID, State, City, Metropolitan Area, County and a Size Ranking
* Primary is Median Housing Prices for each zip code by month for the given range
* Initially 15481 rows by 283 columns
* 13266 rows after rows with missing values removed

**Initial Analysis**

**A screenshot of a cell phone

Description generated with very high confidenceA screenshot of a cell phone

Description generated with very high confidence**

**A screenshot of a social media post

Description generated with very high confidence**

The left graph is the overall distribution of current values. Most zip codes have prices less than $600,000 but some have prices around $1,000,000 and some even higher.

Above the current prices are compared to both a zip codes peak, and to its value from the start of 1997.

To gain a feel of the pricing change over time, some metro areas in Arkansas are visualized below.

A close up of a map

Description generated with high confidence

A picture containing text, map

Description generated with very high confidence

Most temporal data collections often show 1 or 2 zones of increased change. Late 2001, and 2006-2008. The former due to the September 11 attacks and the latter due to the financial crisis. Prior to 2007, the prices were increasing but there was a significant drop due to 2008. Mostly, this does not need to be accounted for since such effects were so widespread, but occasionally data does need to be normalized.

**Method**

Using, the Prophet library, forecasts were made for 2018 based on data from 1997 to 2017. Due to time constraints, a sample approach was taken. First, the approach to aggregate the metro areas from above was used to aggregate all metro areas. There were 712 unique metro areas included in the data. Based off an average processing time of 4 seconds per row, the top 80 metro areas for projected housing price growth for 2018 was used to sample the original set of 13266 zip codes into a more manageable 2311 zip codes. From the forecast obtained from each of these zip codes, a percentage of price gain for December 2018 over December 2017 was the final metric used to compare and sort the zip codes.

From there, a list of the top nine zip codes was obtained. To be more risk adverse, additional data was incorporated to confirm or refute some of the selected zip codes. Below are visuals of these zip codes for the data in 1997-2017.

A picture containing text, map

Description generated with very high confidence

A picture containing map, text

Description generated with very high confidence

**Additional Data**

To inform the analysis, data was also incorporated from data.gov’s dataset on Individual Tax Filings broken down by zip code. Using the adjusted gross income category included in the data, a visual for the top 9 zip codes was created. Concerning zip codes are those with either low population, or a high percentage of tax filings with adjusted gross income of 25,000 or less which could indicate that the price growth is not sustainable comparative to the model.

**A screenshot of a cell phone

Description generated with very high confidence**

Based on some of the criteria, zip code 93620 can be eliminated for having a low population. 33125 and 33135 can be eliminated for having a high low-income percentage. Despite the lower population of 37210, the previous visual shows much higher growth compared to pre 2008 values suggesting true growth versus recovery growth. Outside of those, the remaining zip codes are similar. The most likeliest candidate to be the third zip code is 34207 which had a quicker rebound in prices compared to some of the other zips, and has a decent population size with a low percentage of adjusted gross below 25,000.

**What are the three recommended zip codes**

* **37207, Nashville TN**
* **37210, Nashville TN**
* **34207, Bradenton, FL**

**Conclusion**

It seems that Nashville is a good place to invest in property. It is a growing large city with a thriving music and sports scene having most recently hosted the 2019 NFL Draft. Bradenton is also about 50 miles from Tampa, FL which is a large city on the coast. Overall, the entire country was affected by the financial crisis. Some areas are still recovering from where they were in 2007. Some other areas, however, have experienced true growth.