**Housing Price Prediction in Colorado**

Regis University

MSDS692 – Data Science Practicum 1

By

Sangita Lamichhane

2022

**Project Overview**

The real estate industry is an important sector including varieties of stakeholders ranging from regulatory bodies, private sectors and investing firms. These stakeholders have a high demand of better understanding the real state’s growth mechanism and its return of investments (ROI). Colorado offers great opportunities for investment in housing market, recently we have seen many real estate investors interested to invest in this state. This project can be considered as a further step towards more evidence-based decision making for the benefit for those investors. The project focused on predicting the most beneficial cities in Colorado to invest and get highest return of investment (ROI) between based on data from Zillow. Also, perform EDA to analyze other features and find out any correlation to housing prices from the redfin data. This project provides a recommendation of which 5 Colorado State metro areas are the most prime for investing in now and selling a year from now.

**Research Problem**

The goal of this research project is to explore the model to predict, with time series analysis, which cities in Colorado state are best for real estate investment. This analysis will appeal investors as it will help them make choices by predicting the estimated Return of Investment (ROI) for their investment in certain cities in the state of Colorado. This study finds out which 5 cities may yield high ROI percentage if invested.

**Data Description**

This project uses the first dataset from Zillow research data to determine the top 5 metro areas for real estate investment in Colorado and forecasts their growth over the next year. The dataset contains the median home prices per metro areas recorded monthly over the period from April 1996 to August 2022. The other dataset used for exploring the relationship to housing price is from Redfin.

**Methodology**

In this project we calculated the Return of Investment (ROI) of metro areas of Colorado through an EDA approach as well as ARIMA model fitting. Since, each of the 17 cities required their own model, which had to be grid searched, fit, and used to generate predictions, we have used a custom function to streamline the process. These functions ran auto Arima, a grid searching function, fit an ARIMA model with the optimal parameters, and made predictions using the fit models. Also, to verify and compare the results from the ARIMA model, we also carried out analysis through Facebook Prophet.

Chart, line chart

Description automatically generated

**Conclusion**

**Recommendations**

There are other factors that I would also like to incorporate additional data about these metro areas to better understand whether median home price can really keep growing as predicted. For instance, the information about wage growth in Colorado, migration data as well as inflation data to determine whether prices in the Colorado metro areas are already close to the maximum the market will allow.

Also, the ROI percentage obtained seems high, it would be better to do more research for the sustainability of the rapid growth of house prices seen in recent years.