☰ Vee2002 / phase_4_project

Search: Type / to search

<> Code   Issues   Pull requests   Actions   Projects   Wiki   Security   Insights   Settings

phase_4_project Public

master   1 Branch   0 Tags

Go to file   Go to file   Add file   Code   ...

VivianMosomi432   Merge b...   f1efbdf · 1 minute ago   5 Commits

Presentation .pdf	Time Series Analysis	1 minute ago
README.md	Update README.md	19 minutes ago
final_ipynb.ipynb	Time Series Analysis	19 minutes ago
notebook.pdf	Time Series Analysis	1 minute ago
sarimax_model.pkl	Time Series Analysis	1 hour ago
starter_notebook.i...	Time Series Analysis	1 hour ago
zillow_data.csv	Time Series Analysis	1 hour ago

README

No description, website, or topics provided.

📖 Readme
⎓ Activity
☆ 0 stars
👁 1 watching
⑂ 0 forks

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Contributors 2

VivianMosomi432

Vee2002 Vivian Mosomi

Languages

● Jupyter Notebook 100.0%

Zillow Time Series Analysis

Introduction

In this project, we utilize data sourced from Zillow Research, a reputable provider of real estate insights to predict future property prices. By employing advanced time series technologies, we aim to assist Grand Realty Investment firm to pinpoint the most promising zip codes in the USA for investment.

Our role as a consultancy agency revolves around unearthing actionable insights that enable us to furnish our stakeholder to make informed decisions amidst the evolving landscape of real estate investment.

Objective

This project aims to develop a time series model that predicts the future prices of hoses for our client, Grand Realty Investment firm to finance.

Notebook Structure

The Python notebook is structured as follows:

1. Business Understanding
2. Data Understanding
3. Data Preparation
4. Modelling
5. Conclusions and Recommendations

Data Preparation

The process incorporated to prepare our data for analysis involved several steps including: feature engineering and column renaming, dealing with missing values, conversion of data from wide to long format, checking for stationarity, as well as detrending our dataset to make it suitable for analysis.

Modelling

The baseline model for our analysis was an ARIMA (2,1,2) model. To improve on the performance of the prediction, we incorporated a FB Prophet model which recorded a higher performance than the baseline model. Our final model, chosen as a SARIMAX model, allowed for seasonality of the data and consideration of exogenous variability of our data, showed great prediction success and was used to make recommendations for our stakeholders.

Recommendations

Based on the time series analysis, our project revealed the top five zipcodes for our investor firm to invest in and recommended next steps as well as potential for future research to be considered.