USA Real-Estate Market Analysis

Proposal

The USA has one of the hottest real estate markets right now. We needed a massive amount of data with various factors to know and analyze the real-estate data. Forecasting and evaluating the real estate market helps assess the stability of the US real estate market. Predicting home prices may also give a solid foundation for real estate investors to develop investment strategies and minimize losses.

The issue arises when numerous variables, such as location and property demand, may influence the house price. Thus most stakeholders, including buyers and developers, homebuilders, and the real estate sector, would like to know the exact attributes or the specific factors that influence house prices to help customers and investors make decisions.

To get all this data, we got a dataset which was scraped from the website https://www.realtor.com (found on Kaggle). This dataset contains features like no. of beds, baths, streets, city, and zip code and has 923,160 observations. To get meaningful insights into this massive pile of data, we have to perform statistical analysis, Exploratory Data Analysis on the dataset, and visualize some of the observations.

SMART questions

1. Can we predict real-world housing prices based on this data?

2. What is the correlation between house prices and other attributes?

3. Which location contains the house with the highest and lowest prices?

4. What could be the trend behind housing prices?

Dataset link

https://www.kaggle.com/datasets/ahmedshahriarsakib/usa-real-estate-dataset

Github repo

Link: https://github.com/JayZhangApril/22FA_DATS6101_10W_GROUP6.git