US Housing Market Trends

Team Dive In Data:

Aswini Pusuluri - 015310269
Heer Parekh - 015270320
Sowmya Ravichandran - 015337400
Vamshi Krushna Lakavath - 015351310

Github:

https://github.com/Sowmyadiya/Data228-Project

Objective:

The United States Housing Market is one of the fastest-growing sectors and safest places to invest for many people. Nowadays, we can see a lot of volatility in the housing market. It can be due to the COVID-19 pandemic, work-from-home, rising incoming in few sectors, shortage of inventory, high lumber prices, and record-low mortgage rates. Housing prices have touched an all-time high in most US cities while the affordability issue is worsening and increasing job losses combine to keep homeownership within reach for many potential buyers. The sudden growth of house prices looks like a bubble because in the past we have seen some housing market crashes like in the 1980s and 2008. The latest housing bubble was observed in 2008. Housing prices touched the peak in 2006 and started falling from early 2007 and reached a record low in 2012. Our main objective in this project is to find answers to the below analytical questions:

1) How much was the price fluctuation from 2016 to 2021?
2) Which were the best years and months to buy the house?
3) Maximum percentage change in the market trend in the past five years.
4) The cheapest and costliest cities to buy a house in the US according to price and area?
5) Analyzing the housing market trend during the COVID-19 pandemic.
6) This dataset can also help us in predicting the price of houses in the future.

Source data: https://www.kaggle.com/maedemaftouni/real-estate-market-trends

For this project, we are using the dataset from realtor.com. This is a large dataset that contains house listings details from 2016 to 2021, having a total of 40 fields and each field has 900 thousand plus records of batch data. The dataset contains many measures for diverse levels of geology and different state zip codes. The “Current Month Data” column displays the information for the most recent month, while the historic data contains monthly records dating to previous months. The dataset also contains some other important fields like prices of the houses, mean and median house prices within the time period and geolocation, an active number of listing houses, price increased or reduced month-month and year-year, and area of the house. Some of these fields have null values where data preparation is required. Out of the 40 fields, few are not required for our analysis so we will remove those fields under the data cleaning part.

Milestone and different phases of the project with tentative end dates:

Phases	Tentative Start Date	Tentative End Date
Data Analysis	08/30/2021	09/15/2021
Designing	09/16/2021	09/29/2021
Coding	09/30/2021	10/21/2021
Testing	10/22/2021	10/31/2021
Visualization	11/01/2021	11/10/2021
Documentation	11/11/2021	11/15/2021