USA Real Estate Market Analysis by Lisa Greer

Overview:

- "Location, location, location" – We all know the age-old adage that location is the biggest factor driving real estate prices. But what if there is more to it? This analysis explores a variety of housing features to find if there is an equally important factor in predicting housing prices.

Purpose:

- This project was completed out of my own curiosity and may be relevant to realtors and home buyers/sellers in the USA.

Objective:

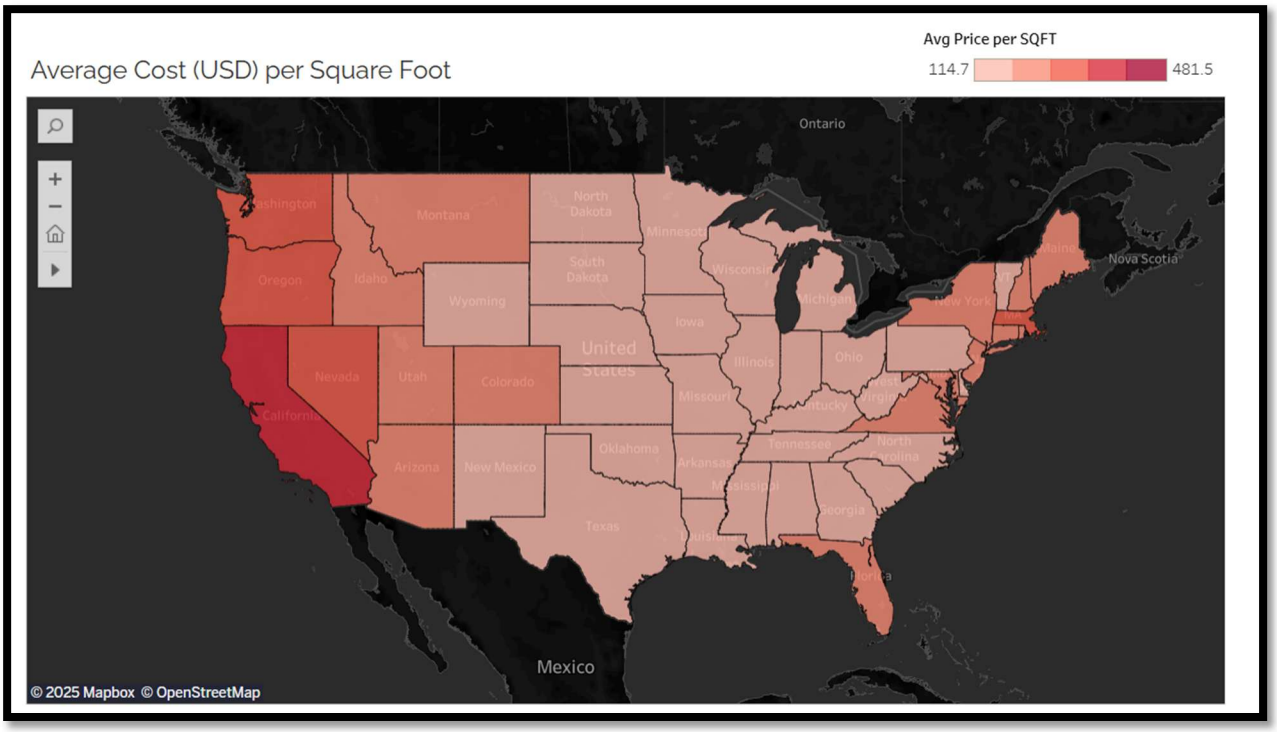
- The goal of this project was to find out what factors truly influence housing prices in the USA. This is especially relevant now as many people conjecture about housing costs due to the lack of affordable housing in the USA, but looking at the data is the best way to find the truth behind housing costs.

Tools, Skills, Methodologies:

- This analysis was completed using data from realtor.com; a real estate listing website that is the second most visited of its kind in the United States as of 2024, with over 100 million monthly active users.

- Python was used to clean, manage, and wrangle the data so that it could be analyzed using statical techniques and visualized using Python libraries and Tableau.

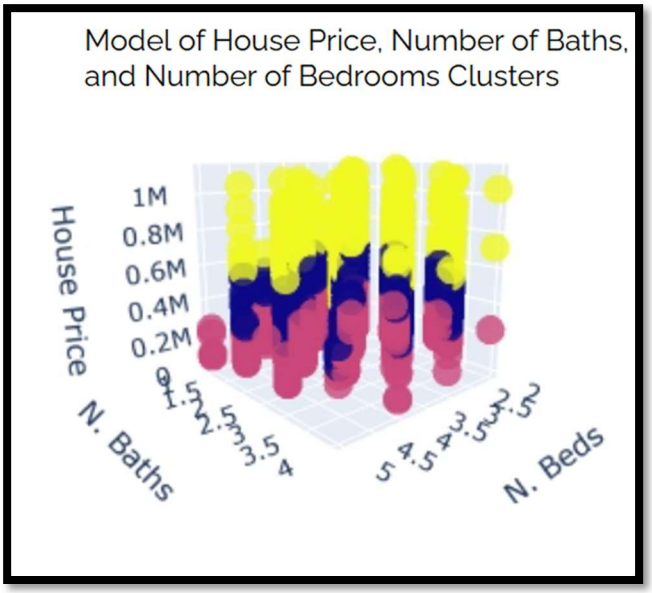
The Findings:



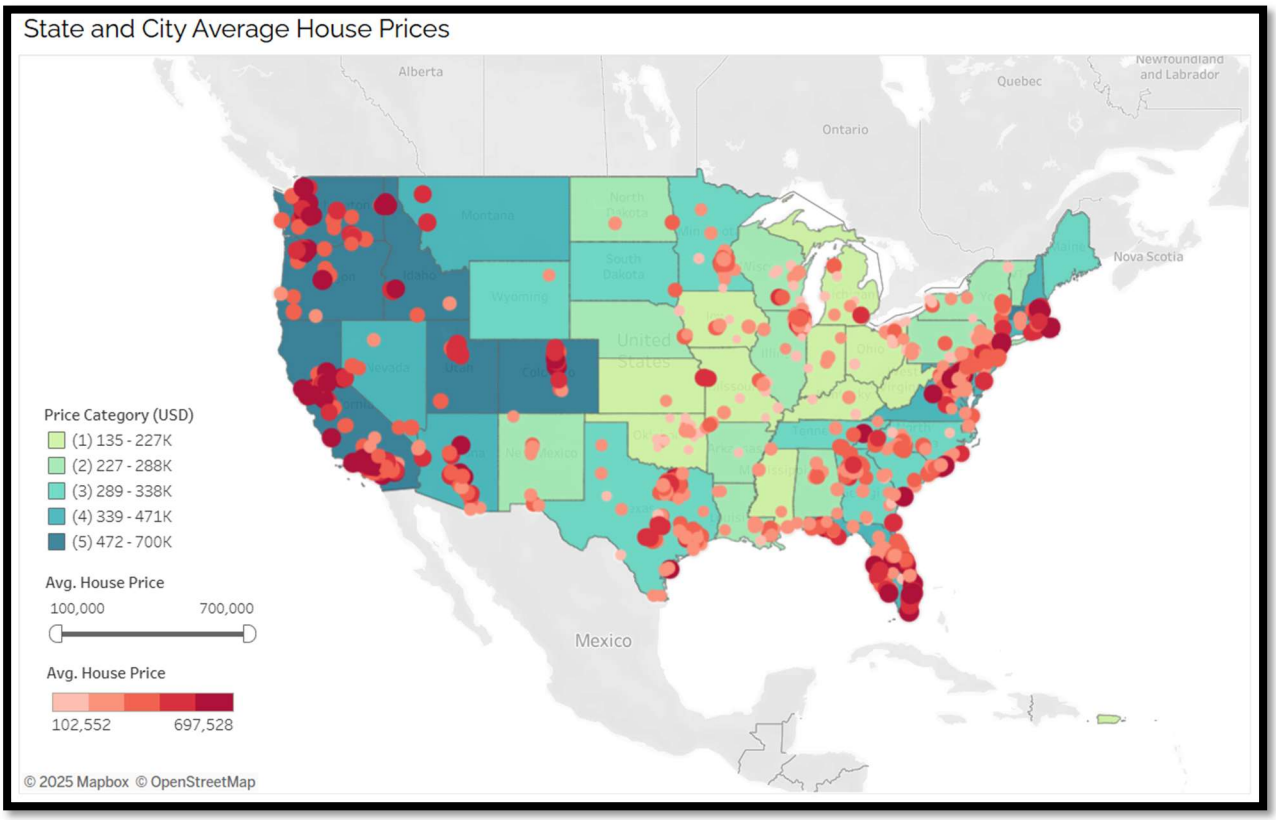
To begin the analysis a map of price by square foot for housing across the USA was created. The darker the state, the more expensive the housing. This map confirmed typical expectations of high demand areas being higher priced (ie. vacation areas and coasts).

Next a comparison of median house prices was conducted with housing features including average number of bathes, bedrooms, acres, and house square footage. This chart was created to look for linear dependence between cost and the four house features. No linear dependence was found meaning none of these may be significant factors affecting national housing costs.

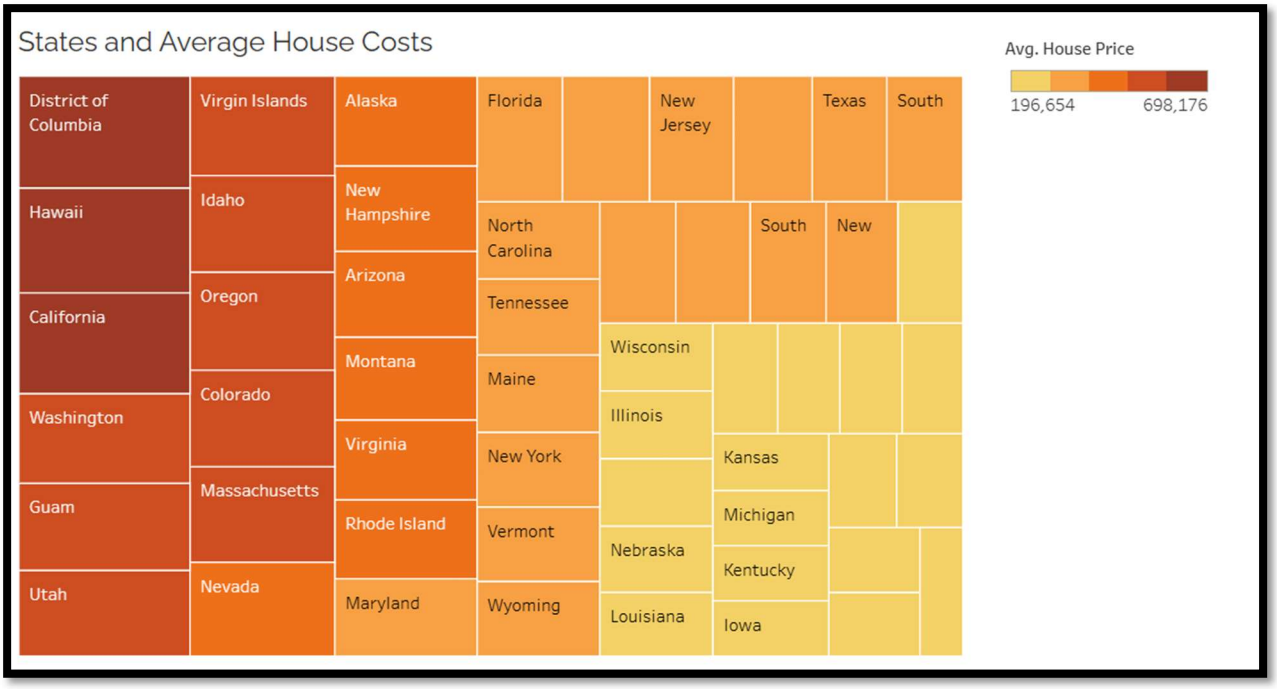
To confirm the hypothesis that house features (house square footage being the example here) did not significantly impact house price a linear regression was conducted. The results indicated that only 19% of house costs could be attributed to square footage. Therefore a linear regression was not enough to explain house costs.



Since a linear regression was not enough, a cluster analysis was conducted. This type of analysis groups data points in into “clusters” that can be compared to one another to find patterns. The analysis revealed three distinct groups. However, as seen by the three different colors in the chart, these groups were divided by price. This reveals that the largest pattern in the data is house pricing over house features. Meaning it is time to look at other factors like geographical location.



The data revealed that location is in fact the largest factor influencing housing prices. As seen in the map; large, dense cities and popular places to live (coast, islands, and metropolitan areas) are still the most expensive parts of the USA real estate market.



Conclusion:

- When buying or selling a house in the USA it is important to understand the market and how it works.
- Location is ultimately the most influential factor on pricing with desirable locations and metropolitan areas accounting for the highest prices. To get more space for the money, the Midwest is the most cost-effective part of the USA. If location is the most important factor for a buyer they will be paying premium prices.

Limitations:

- The data needed significant cleaning with a lot of outliers being removed to identify trends (or lack thereof). There are some insights that might have been missed if those outliers were to be further analyzed.

Next Steps:

- Narrowing the analysis to specific states or cities could yield far more specific insights into how the market is operating in certain parts of the USA.
- It would be interesting to cross-reference this with more historical data to find trends over the lifespan of the USA as a country.

- Cross-referencing this data with weather patterns (particularly high-risk insurance areas) would provide further insight into house costs.
- Adding house age as a factor would also be an interesting avenue to follow.

Follow – Up:

- For more details about the project and data here are links to the interactive charts and GitHub repository:
    - Tableau Link
    - GitHub Link