Las Vegas, NV
Single Family Houses
Best Investments

HOMEBRAIN INNOVATION

Presented by
Joy Mickle
Binh Do
Nasr Salahuddin
Beenish Mehbob

Project Background - Las Vegas Best Investments

Las Vegas economic dashboard: Real Estate market: population (2.3million, population increasing but housing permits are down from last year, tight sellers market due to lack of inventory but sales taxes are increasing year to year. There is a gap between capacity and availability of housing.

Buyer Profile: (avg income increasing year to year ($133k to $58k for per capita income increasing year to year)

Popular zip codes colored map: Zip Codes 89121, 89131, 89121 are the top housing sales zip codes (most housing sales occurred)

Clark County Tax Assessor : Dataset includes all real property sale in the category of single family house or vacant land zoned for single family usage from 2018 to 2023 in Las Vegas city limits.

Redfin: Dataset includes scraped data from years (to present), including listings of single family homes for sale in Las Vegas city limits (2900 current listings -1800 sold)

Data Cleaning - Beenish

In the provided code, data cleaning and preprocessing steps were performed on a dataset before training a Linear Regression model. Here's an overview of how the data was cleaned:

1. Handling Missing Values:

- Rows with missing values (NaN) were dropped from the dataset using `data.dropna(inplace=True)` to ensure that the model is trained on complete data.

2. Cleaning the 'Lot Size' Column

- The 'Lot Size' column contained values with commas (',') as thousands separators. These commas were removed using `data['Lot Size'] = data['Lot Size'].str.replace(',', '')`. - Additionally, specific values like '1 Acre' in the 'Lot Size' column were replaced with the equivalent numeric value (4840) using `data.loc[data['Lot Size'] == '1 Acre', 'Lot Size'] = 4840`
- The 'Lot Size' column was converted to a float data type to ensure it contains numerical values using `data['Lot Size'] = data['Lot Size'].astype(float)`.

List of Columns in the Data Set:
URL
Walk Score (out of 100)
Transit Score (out of 100)
Bike Score (out of 100)
Elementary School Score (out of 10)
Middle School Score (out of 10)
High School Score (out of 10)
Address
Price
Beds
Baths
Sq Ft
Year Built
Property Type
HOA Dues
Lot Size
Garage Spaces
Has Pool
Has Patio
Flood Factor
Fire Factor
Heat Factor
Wind Factor
Full Address
Zipcode

Heat Map- Beenish

Findings from the Correlation Heatmap:

The correlation heatmap visually represents the relationships between different features in the dataset.

1. Positive correlations:

- 'Price' has a strong positive correlation with 'Baths' and 'Sq Ft', indicating that as these values increase, the price tends to increase as well.
- 'Beds' also show a positive correlation with 'Price', although slightly weaker than 'Baths' and 'Sq Ft'.
- Garage Spaces' have a moderate positive correlation with 'Price'.

2. Negative correlations:

- 'Walk Score (out of 100)' and 'Transit Score (out of 100)' have negative correlations with 'Price', suggesting that higher walk and transit scores are associated with lower prices.
- 'Year Built' has a negative correlation with 'Price', indicating that older properties tend to have lower prices.

Front End- Back End Working-Nasr

With our data storage processes we were able to successfully assemble all of the data and convert it into a user accessible HTML format that created our website using Javascript and Machine Learning algorithms from our models.

- We created six HTML scripts for our website and used models for our back end to generate our website.
- Dropping, binning, and relocating classifications in order to create a comprehensive back end that could lead to a well portrayed front end.
- Created four different price ranges measuring real estate value in order to categorize our data,
- Dropping unnecessary variables
- Compiled a set of models that would help us predict the best investment in the Las Vegas area according to predicted resale value.

Conclusion- Nasr

Our project compiled a set of predictive models to assist users in identifying the best real estate investment opportunities in the Las Vegas area, specifically focusing on predicting resale values. This approach aimed to provide users with data-driven recommendations for their investment decisions.

- Our findings conclude that Las Vegas area has opportunities for real estate growth but we feel that there are other ways we could optimize our model in order to predict a more accurate result. Our findings led us to believe that there are some good investments across different price ranges according to a few zip code hot spots. Our models show there are significant correlations between house price, zip code, and walk score. However, we feel as a group there are still some ways we could probably achieve a more accurate result.

Question ?