Zillow Home Value Forecast (ZHVF)

Python EDA

>>>>>

>>>>>

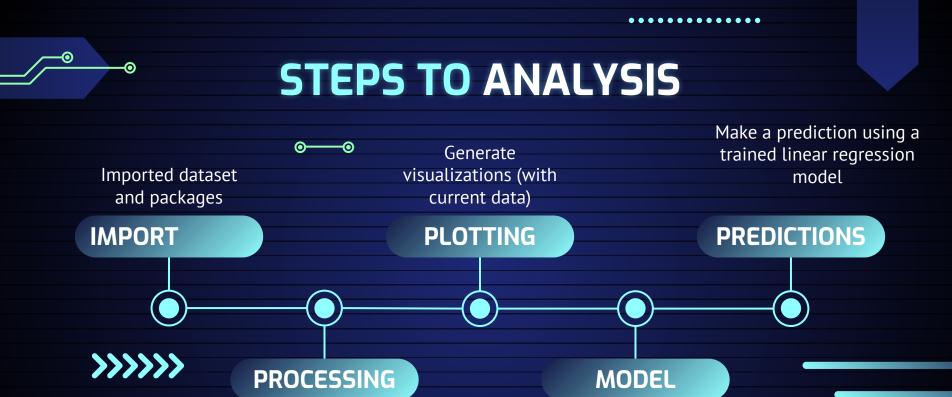
Dalia Ali, Shagufta Ali, Roli Arah, Nikita Thakore

PROBLEM STATEMENT

Providing a more accurate estimate of a home's future value based on current data and market trends to help individuals make informed real estate decisions.

<<<<<





Processing the Data Frame and solve Date column issues (convert Date to a

categorical variable)

Train a linear regression model using the "Date" and "Price" columns



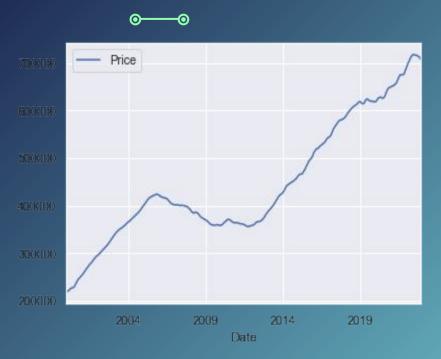
How we used ChatGPT

```
Copy code
      python
      # load your data into a pandas dataframe
      df = pd.read_csv("your_data.csv")
      # convert the date column to a categorical variable
      df['date'] = pd.Categorical(df['date'])
      df['date cat'] = df['date'].cat.codes
      # extract the features and target variable
     X = df.drop(['target_variable', 'date'], axis=1)
     y = df['target_variable']
      # split the data into training and testing sets
      X train, X test, y train, y test = train test split(X, y, test size=0.2, random sta
    3. Train the linear regression mode S Regenerate response
CHATGPT, THANK YOU AND I LOVE YOU .....
```

We mainly used ChatGPT to:

- Troubleshoot when we were having problems
- Better understand certain Python syntax
- Understand different techniques of performing a task in
 Python/Pandas/Scikitlearn

Visualization



Based on current data (left), the prices have been on a general upswing since 2000.

We performed a linear regression to understand what the estimated average house price would be in Boston in June 2023. Our linear regression showed a slope of 1465.21 and intercept of 245064.10.

In other words: with a starting price of \$245,064, each day adds approximately \$1465 to a house's price.*

*Notably, since linear regression accounts for best fit, it will not always accurately predict the housing cost if there are observed dips in the actual data.



Conclusion

We used a linear regression to understand the relationship between Date and Price. Based on the data analysis: As time moves forward (Date "increases"), so does Price. We can expect housing prices to continue to rise.

Our regression model allowed us to predict house values for individual dates 1-1-2023 and beyond. Based on our model, for example, we anticipate the average house price to be approximately \$911,735 on 6-1-2023.

Users who wish to use our model to predict house prices may do so by inputting for how many number of days after 12-23-22 they wish to predict the price.



