King County Housing Analysis

By Vijaya Gunukula

Goal

To develop a model that will predict house prices in King County as accurate as possible.

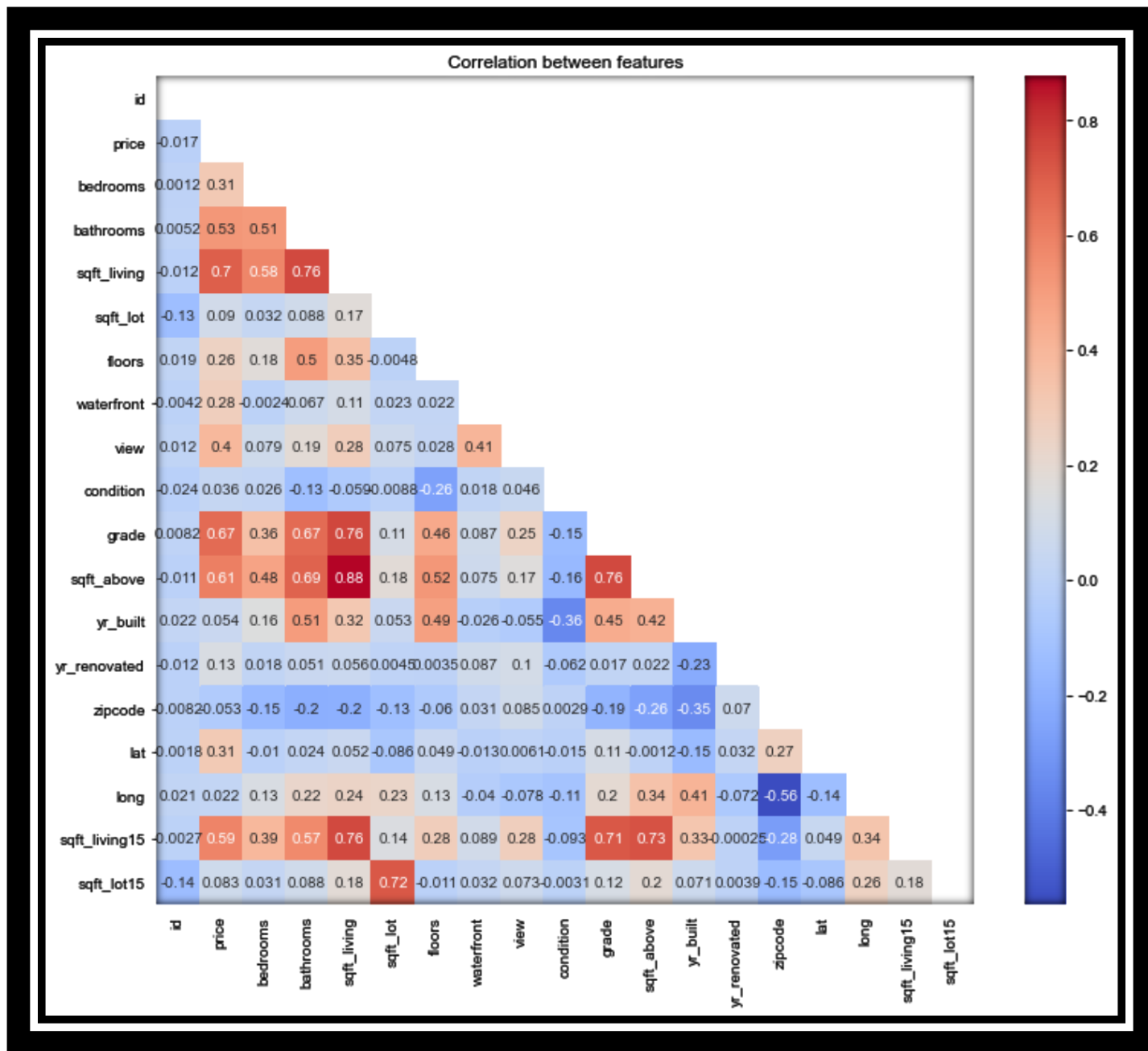# Data

❖King County's House Sales data of over 20,000 homes

❖Number of bedrooms, bathrooms and floors

❖Size of the houses

❖Location by zip code as well as longitude and latitude

❖Condition and grade of the houses based on county grading system
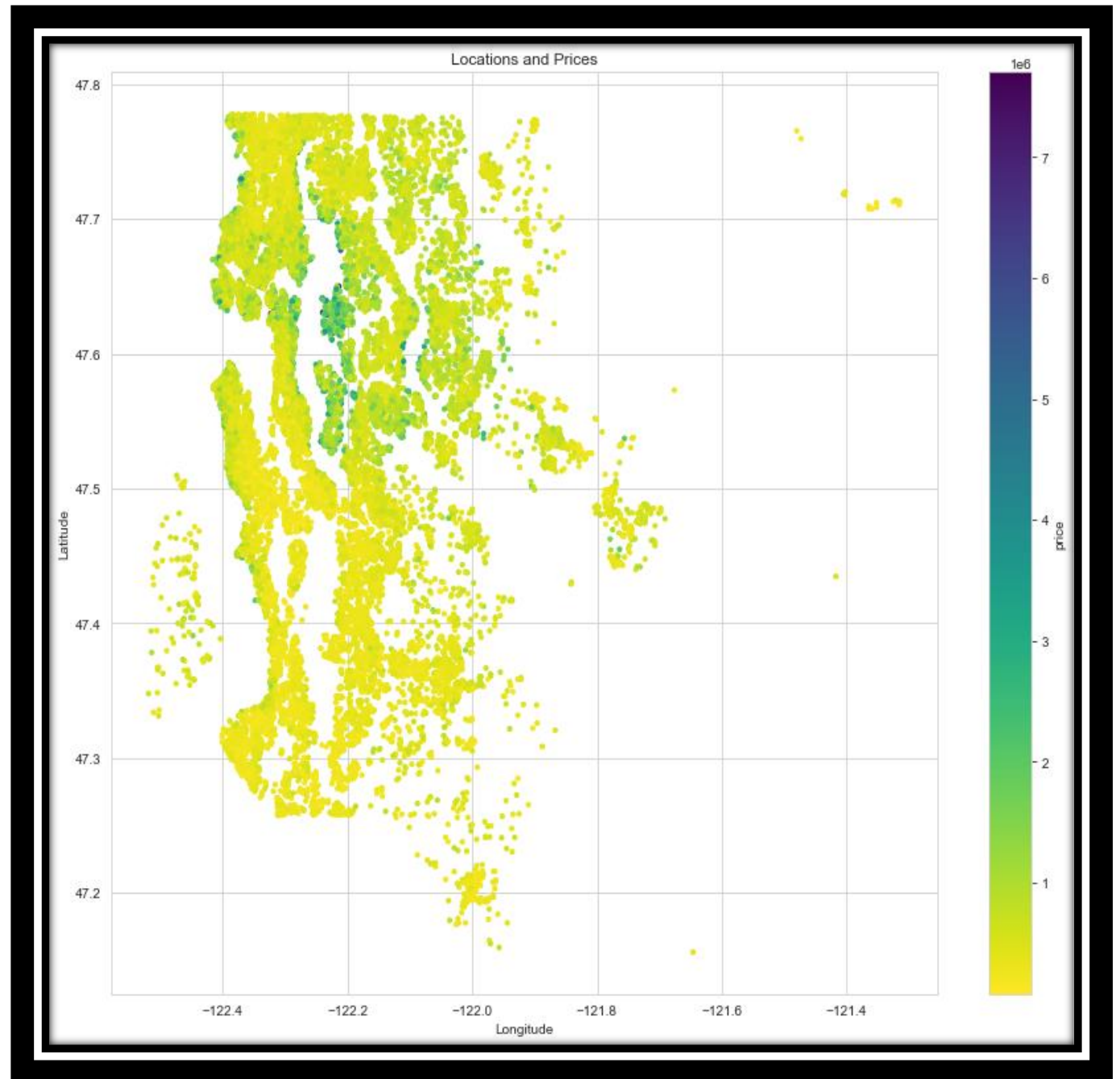
❖View of the houses

❖Date of the transaction

Data Exploration

•The darkest red outputs illustrate the strongest positive linear relationships as far as we can strictly from correlation coefficients and the blue values are the negative linear relationships.

•The values that describe each house's sqft_living and grade, sqft_living and price are strongly correlated with each other.
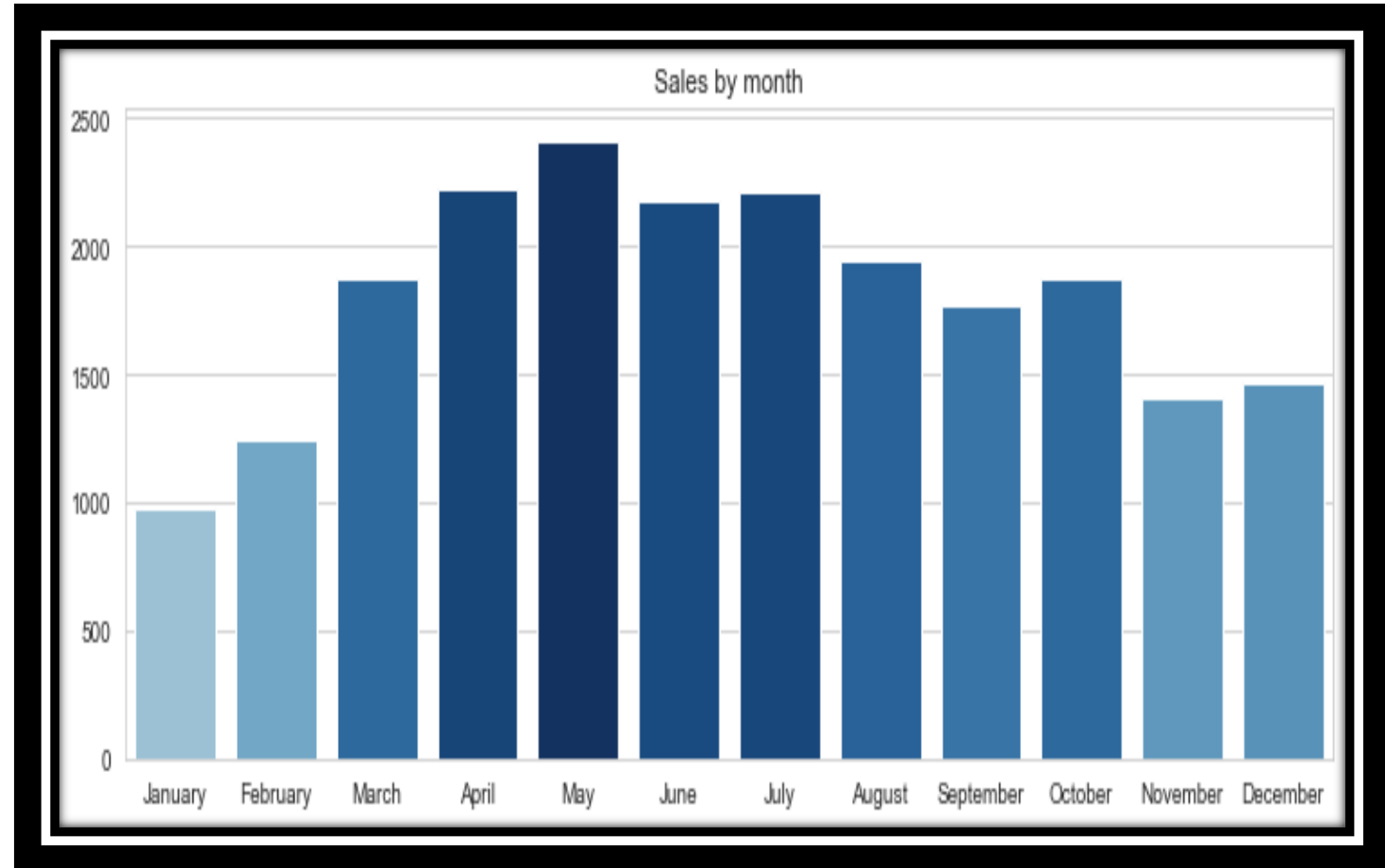

Correlation between features

Best Location

Prices for square feet ranges from $87 to $810. Darker spots on the map identify premium locations with properties over $5 Million.

Locations and Prices

Most House Sales in a Month

❖ The most houses were sold in the months of May, April and July.

Conclusion

Model performs with 75% accuracy across all the data.

Selected important features and rejected the ones that can negatively impact result of the prediction.

Condition and size of the house have biggest impact on house price.

Future Work

Analyze more on **neighborhoods, which vary greatly in average house price.**

Thank You