

# PHASE 2 PROJECT



## VALUE ANALYSIS FOR HOUSE SALE PRICES IN KING COUNTY

# Business Overview

## Will your renovations pay off?

- This project aims to provide valuable insights for a real estate agency operating in King County, Washington, USA. The agency seeks to provide advice to homeowners on how home renovations can potentially increase the estimated value of their properties and homes, and by how much.



# Business problem

- Is flipping worth it?
- What is the expected return on renovation as an investment?.



# Objectives

- Scale the Impact of renovation
- Construct a model that can be used in predicting house sale prices
- Find out the factors that highly impact property prices
- Have actionable recommendations to stakeholder's satisfaction



# Data Sourcing

- We used publicly available datasets from King County Datasets

## Why King County?

The area is rich in the data resource having information on over 21,000 homes, providing researchers with a rich and diverse set of data to work with.

It houses headquarters for major tech companies which have contributed significantly to the county's economy and have helped to make it one of the wealthiest and most prosperous areas in the United States.

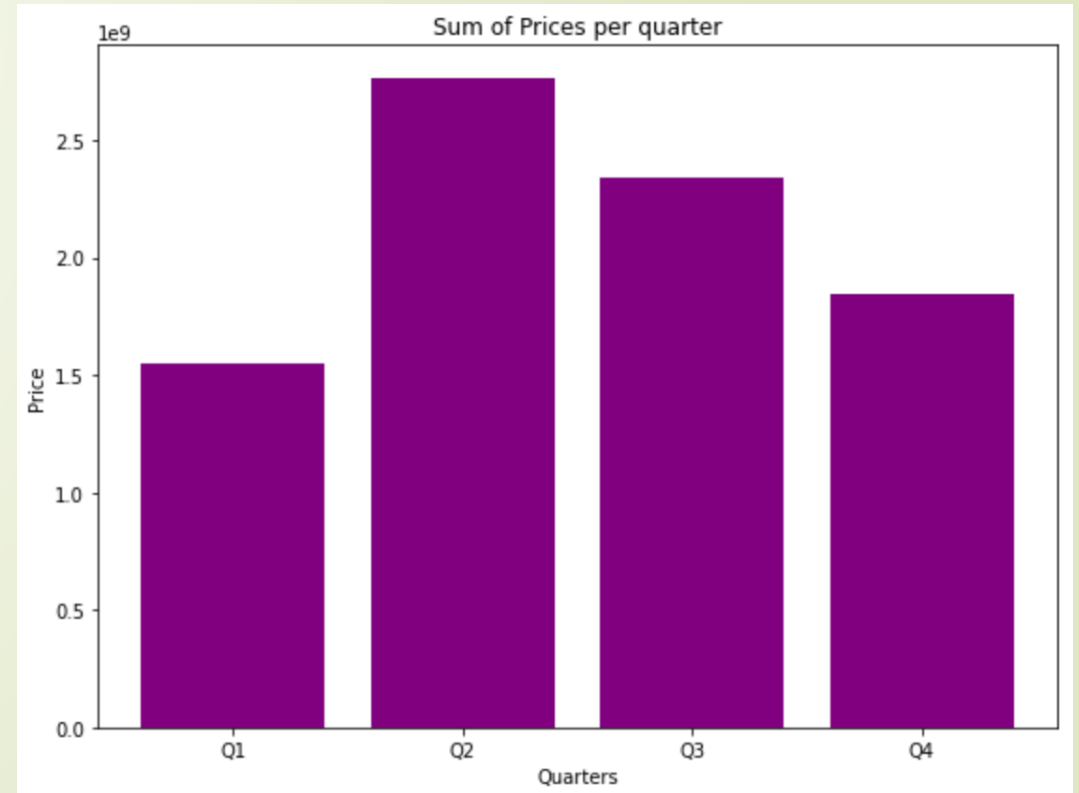


# Exploratory data analysis

The graph shows the sum of houses sold per quarter, with second quarter being the highest.

This is around summertime!

This makes sense because most people are on holiday and crave relaxing spaces, especially those with waterfronts.

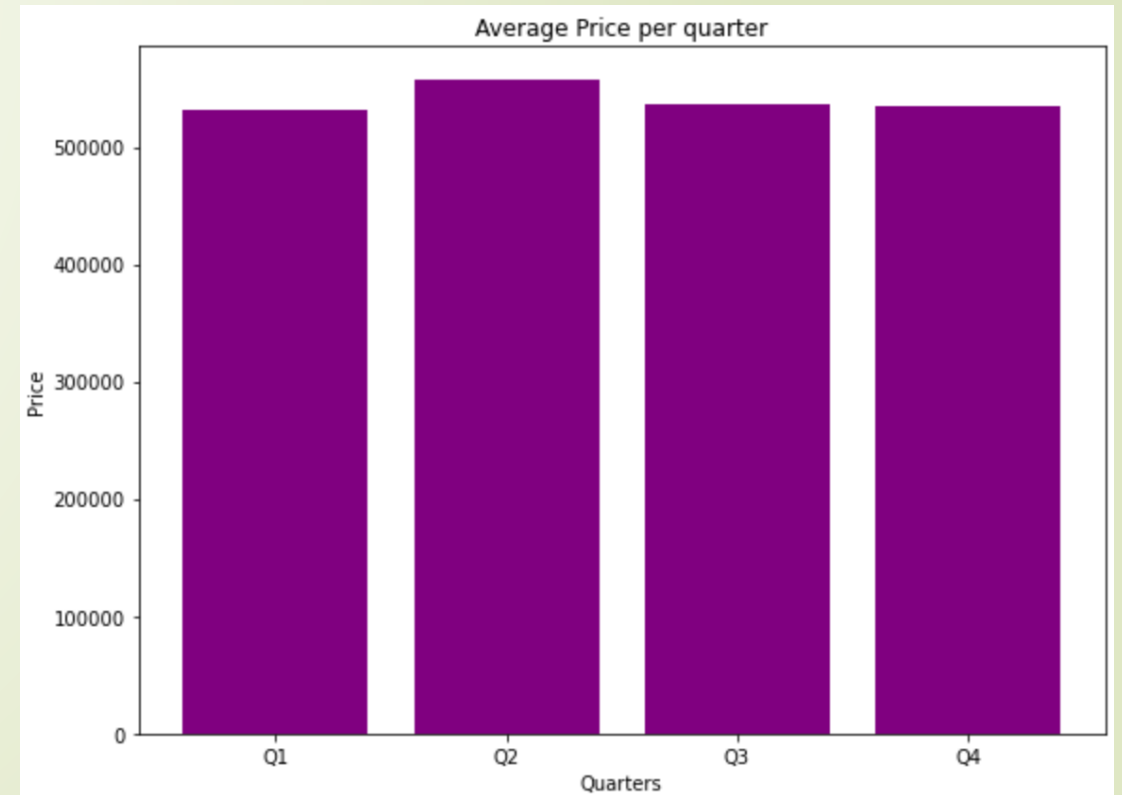




# Exploratory data analysis

The average graph confirms that the sale price is higher in the second quarter.

This could be due to the summertime demand



# Data Modeling

Our first model explained 49% of the variation in sale prices.

We wanted to improve this pilot model. This called for fine-tuning by adding more features to our model, to come up with accurate predictions.

## OLS Regression Results

```
price      R-squared:
      OLS    Adj. R-squared:
Least Squares  F-statistic:
Wed, 19 Apr 2023  Prob (F-statistic):
      23:45:26  Log-Likelihood:
      15684    AIC:
      15682    BIC:
           1
nonrobust
```



# Data Modeling

By adding more predictors, our model accuracy improved by about 10%, to 58% of explained variation in sale prices

Our target metric wasn't met yet!

Some more digging

## OLS Regression Results

```
price    R-squared:
OLS      Adj. R-squared:
Least Squares  F-statistic:
Wed, 19 Apr 2023  Prob (F-statistic):
23:45:27      Log-Likelihood:
15684         AIC:
15673         BIC:
10
nonrobust
```

# Data Modeling

After further research, we added more features that significantly improved our model, surpassing our target.

It meant that the inclusion of location parameters was important for the prediction of sales.

**Eureka!** Our model explains 79% of the variation in sale prices.

## OLS Regression Results

```
price      R-squared:
OLS        Adj. R-squared:
Least Squares  F-statistic:
Wed, 19 Apr 2023  Prob (F-statistic):
23:45:41    Log-Likelihood:
s:          15684    AIC:
           15604    BIC:
              79
           nonrobust
```



# Results

- For an average house that is not renovated, with no waterfront and built since 1900, we would have a sale price of 711,600 dollars
- The most impactful features for inferring and predicting house sale prices were found to be the square footage of the living area, bedrooms, floors, and whether the property had a waterfront view or was renovated.
- The waterfront view proved to be most impactful with houses having a waterfront view having value increase of about 834,000 dollars more than those without a waterfront. This can be inferred for the other variables as well.
- Our model is off by about 100,000 dollars for each prediction of the price



# Recommendations

- Development of a comprehensive database the agency can use to track the renovation projects that would improve property value.
- For customers who would like to sell their properties/houses, it is recommended that they should renovate them first as we see these would see an increase in value by about 48,000 dollars.
- We recommend for customers on a budget should look out for houses situated on higher property floors as we expect a price drop on said houses of about 30,000 dollars for every floor increase



# Conclusion

- The combination of square footage of the living area, bedrooms, floors, and whether the property had a waterfront view or was renovated are the most reliable predictors of a house's price in King County.
- However, there are some limitations to the model. To meet regression assumptions, we had to try out log-transformation on certain variables. Therefore, any new data used with the model would require similar pre-processing. Additionally, since housing prices vary regionally, the model's usefulness for data from other counties may be restricted.
- In summary, if you are seeking affordable housing, it may be advisable to compromise on square footage and have no waterfront view. But, given that many urban residents already do this, it may not be a viable solution for everyone.



# THANK YOU

Presentation by:

Joel Omondi - [joel.omondi@student.moringaschool.com](mailto:joel.omondi@student.moringaschool.com) -TL

James Mbeti - [james.mbeti@student.moringaschool.com](mailto:james.mbeti@student.moringaschool.com)

Kennedy Juma - [kennedy.juma@student.moringaschool.com](mailto:kennedy.juma@student.moringaschool.com)

Aaron Onserio - [aaron.onserio@student.moringaschool.com](mailto:aaron.onserio@student.moringaschool.com)

Victoria Nabea - [victoria.nabea@student.moringaschool.com](mailto:victoria.nabea@student.moringaschool.com)

Nyokabi Waiganjo - [nyokabi.waiganjo@student.moringaschool.com](mailto:nyokabi.waiganjo@student.moringaschool.com)

<https://github.com/Ed-Odhiambo/dsc-phase-2-project-v2-3.git>