

## Phase 2 Project: Business Case Study using Data Science

Dennis Mwanzia  
Ian Macharia  
Samuel Igecha  
Pamela Otieno  
Pauline Ndung'u



# BACKGROUND

- ▶ King County is a region in Washington state that includes Seattle and is known for its strong economy, diverse population, and natural beauty. Major employers include technology companies, healthcare providers, financial institutions, and government agencies. The real estate market in King County is highly competitive due to strong demand, limited supply, and high levels of income and wealth among buyers, with median home prices above the national average and increasing steadily in recent years.





THE OBJECTIVE OF THIS PROJECT IS TO APPLY OUR NEWLY ACQUIRED DATA SCIENCE SKILLS TO FORMULATE AND SOLVE A REAL BUSINESS PROBLEM.



WE HAVE BEEN PROVIDED WITH A RAW DATASET CONSISTING OF HOUSE SALES PRICES IN NORTHWESTERN COUNTY.



TASK IS TO FORMULATE A REAL BUSINESS CASE STUDY, ANALYZE THE DATA, AND PROVIDE VALUABLE INSIGHTS TO AID IN MAKING INFORMED BUSINESS DECISIONS.

## Business Understanding



The aim of this project is to provide useful business insights to a hypothetical client, Kings Wajuzi Developers who is interested in investing in the real estate market of Northwestern County.



Our goal is to analyze the provided dataset and formulate a business problem related to this domain, which can be solved using data science techniques.



By leveraging our expertise, we aim to provide valuable recommendations that can help our client in making informed decisions regarding their investment in the local real estate market.

# OBJECTIVES



To determine the best season that favors sale of prices of homes.

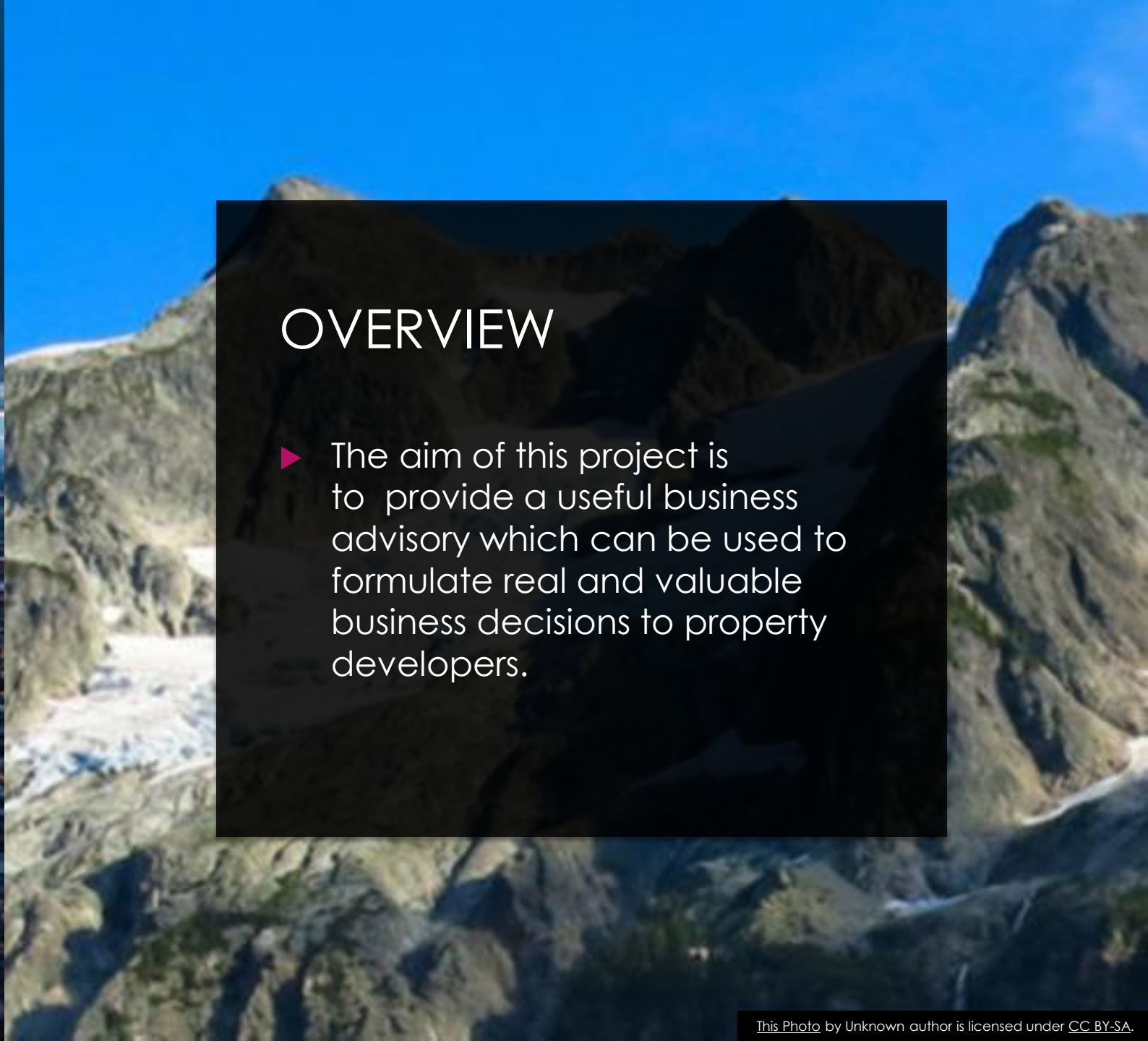


To determine which locations are the most popular to home owners.



To determine which house features home owners look out for the most.





## OVERVIEW

- ▶ The aim of this project is to provide a useful business advisory which can be used to formulate real and valuable business decisions to property developers.

# TOOLS

- Python
- Python Libraries

## DATA SOURCE:

- [Kaggle.com](https://www.kaggle.com)

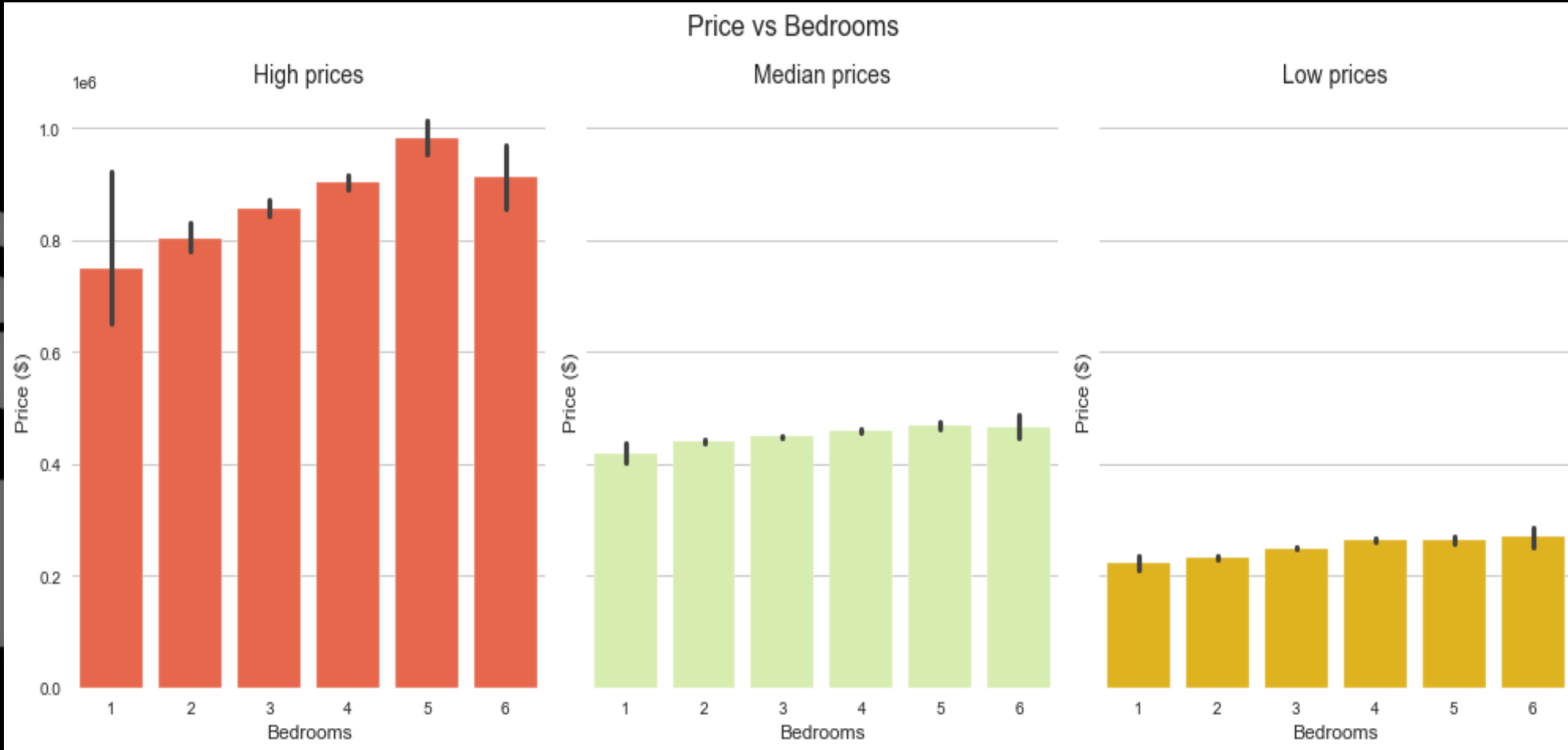




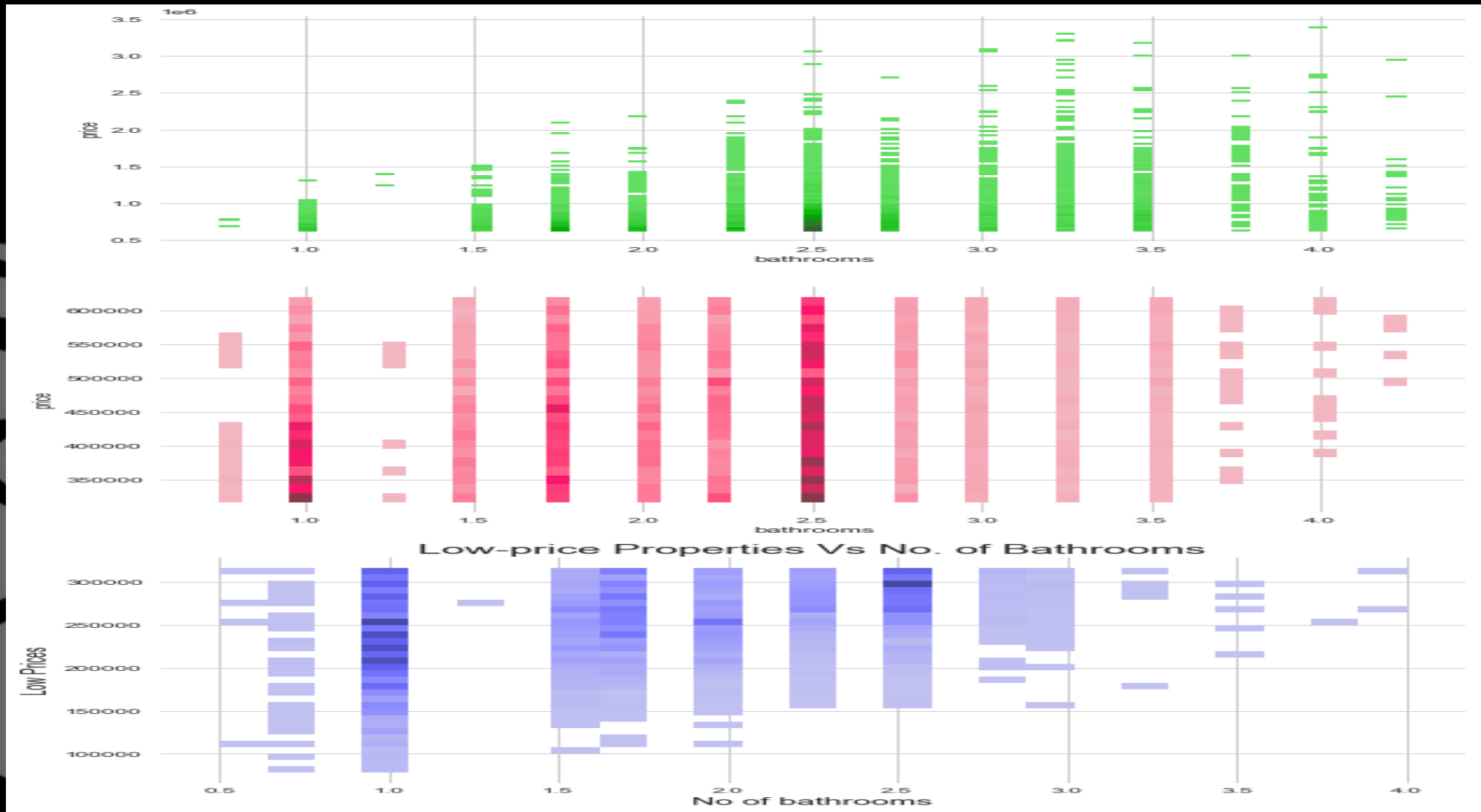
# DATA VISUALIZATION



# 1. How do different priced houses relate with the number of bedrooms?

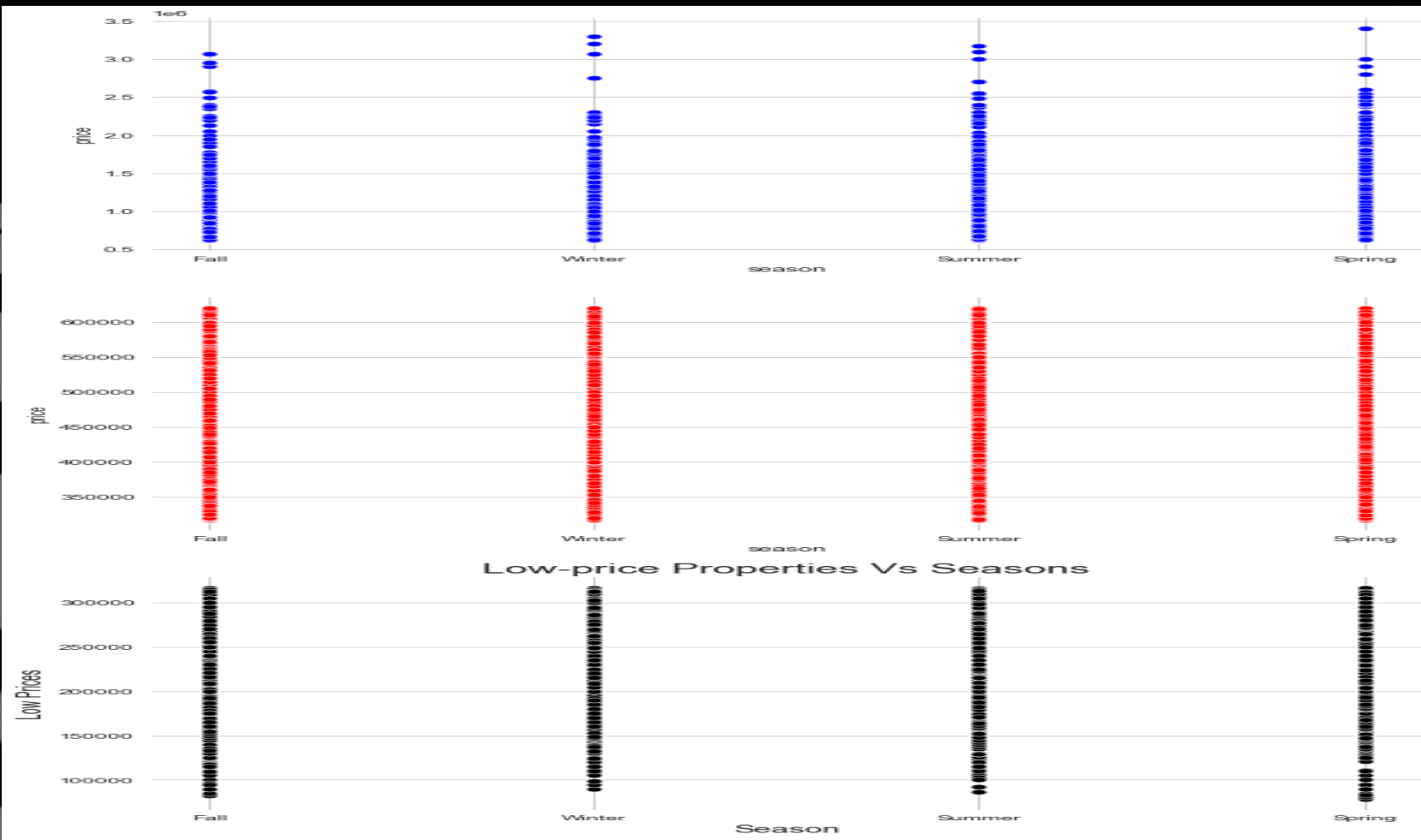


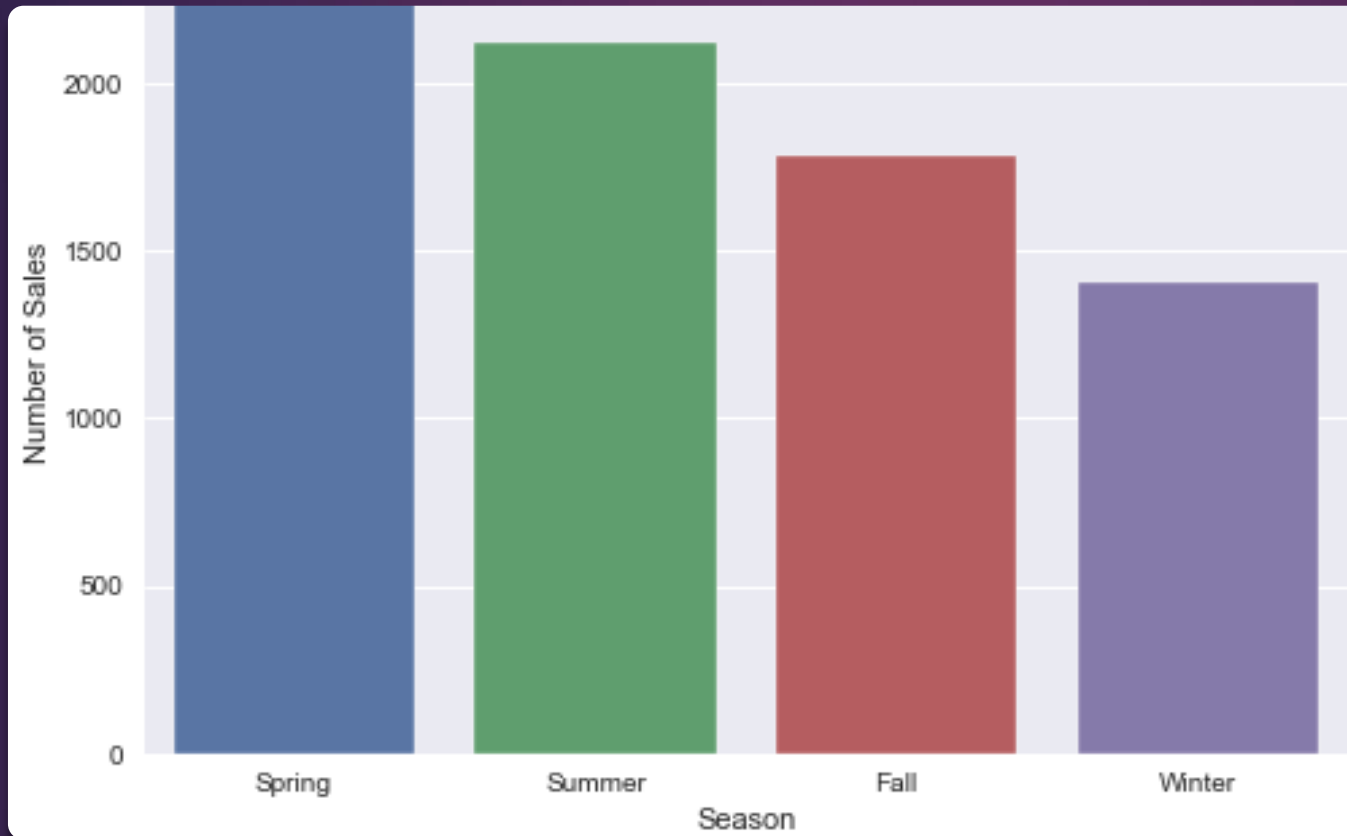
## 2. How do different priced houses relate with the number of bathrooms?





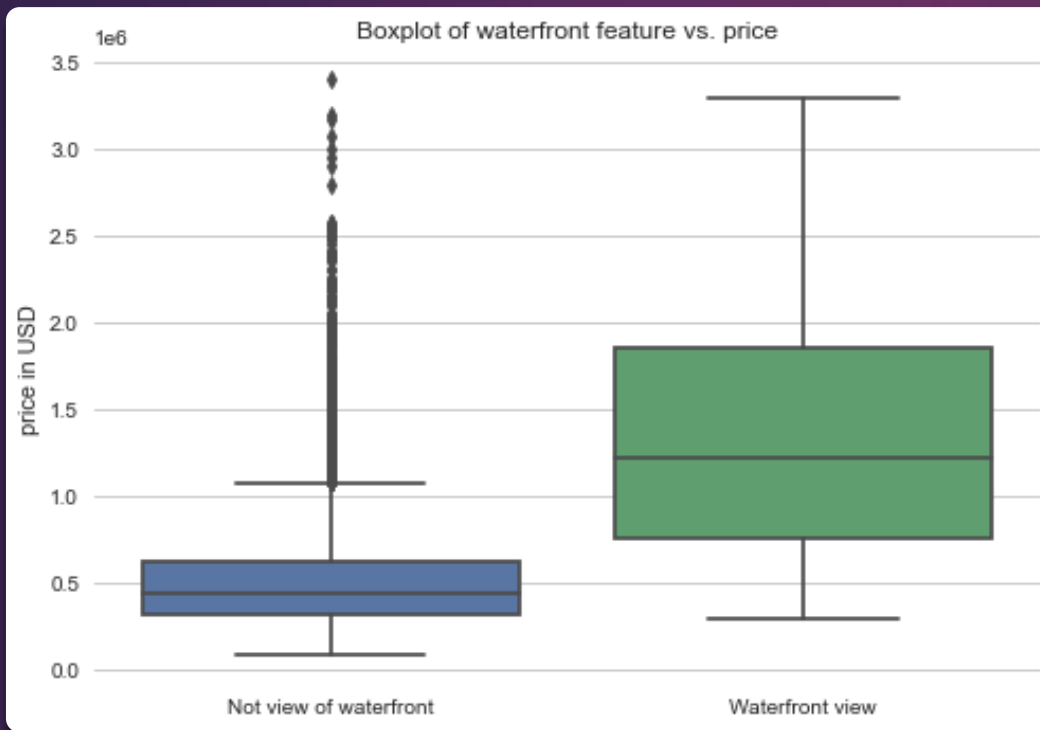
### 3. How do different-priced houses relate with the seasons



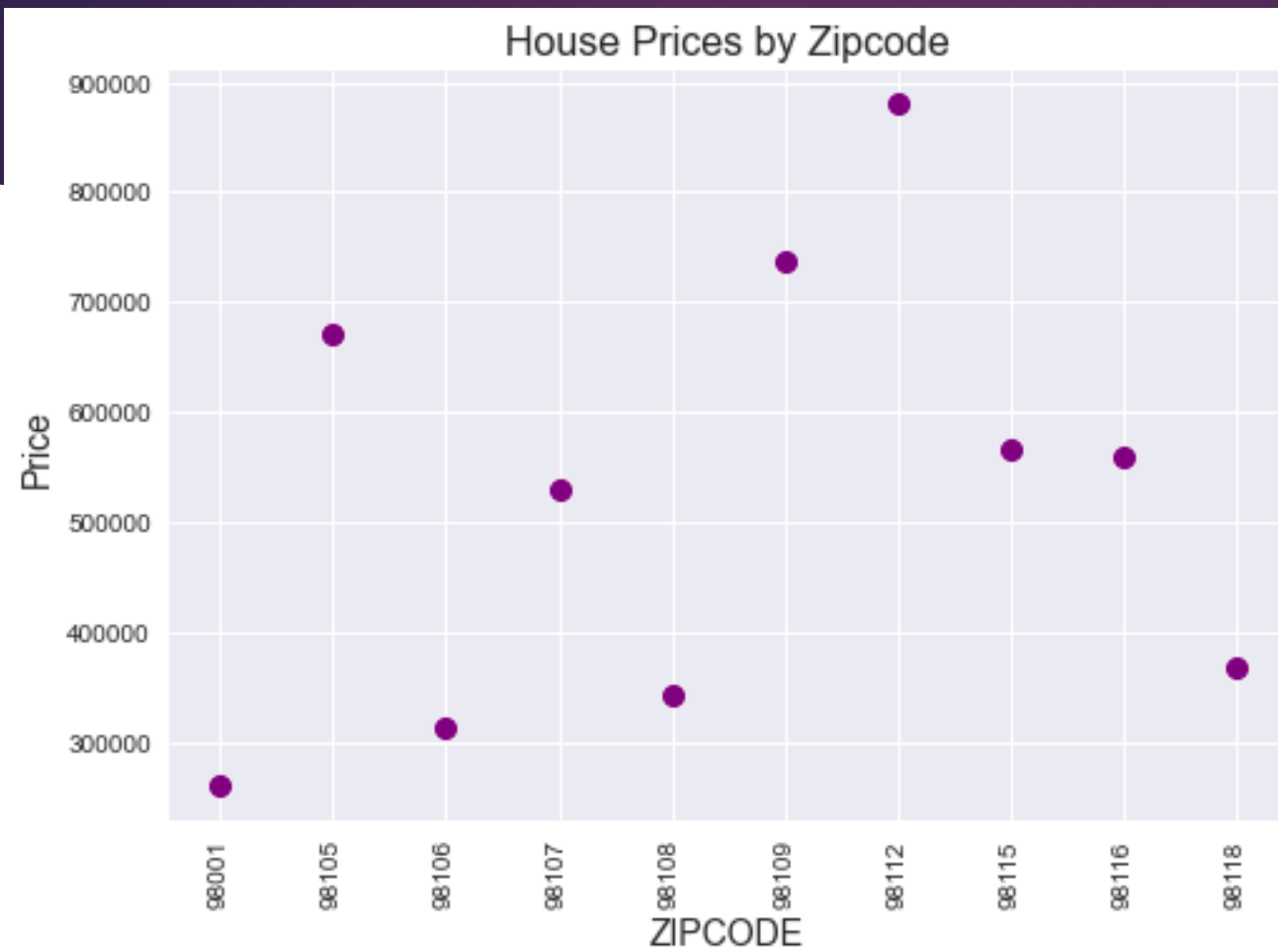


How do  
Seasons  
Affect  
Sales of  
Properties



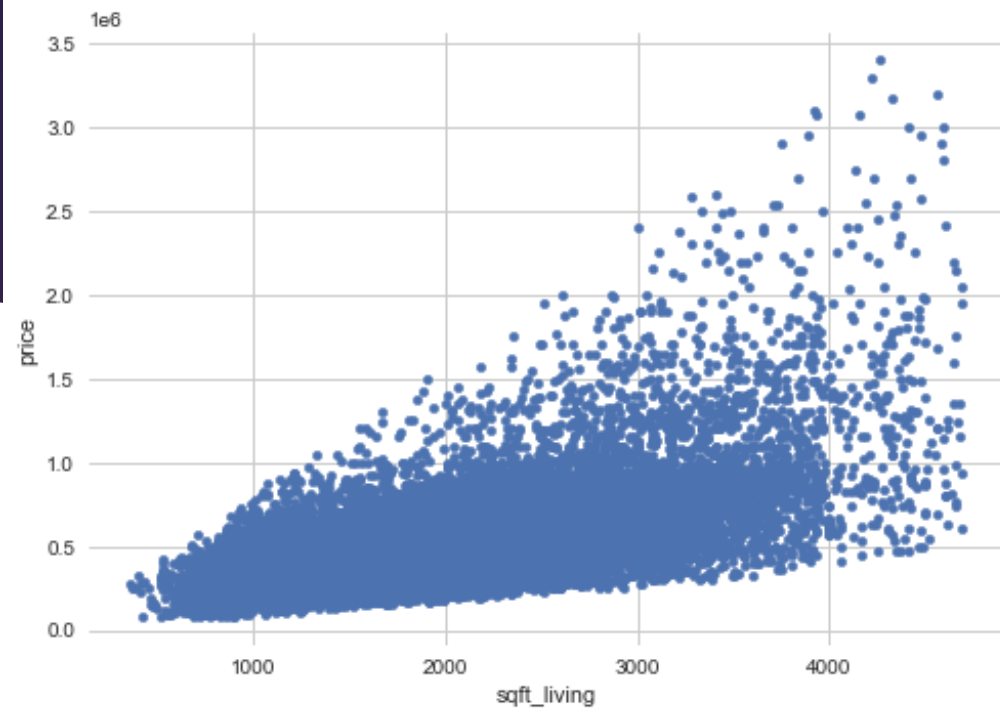


Does the Presence of a waterfront feature affect its price?



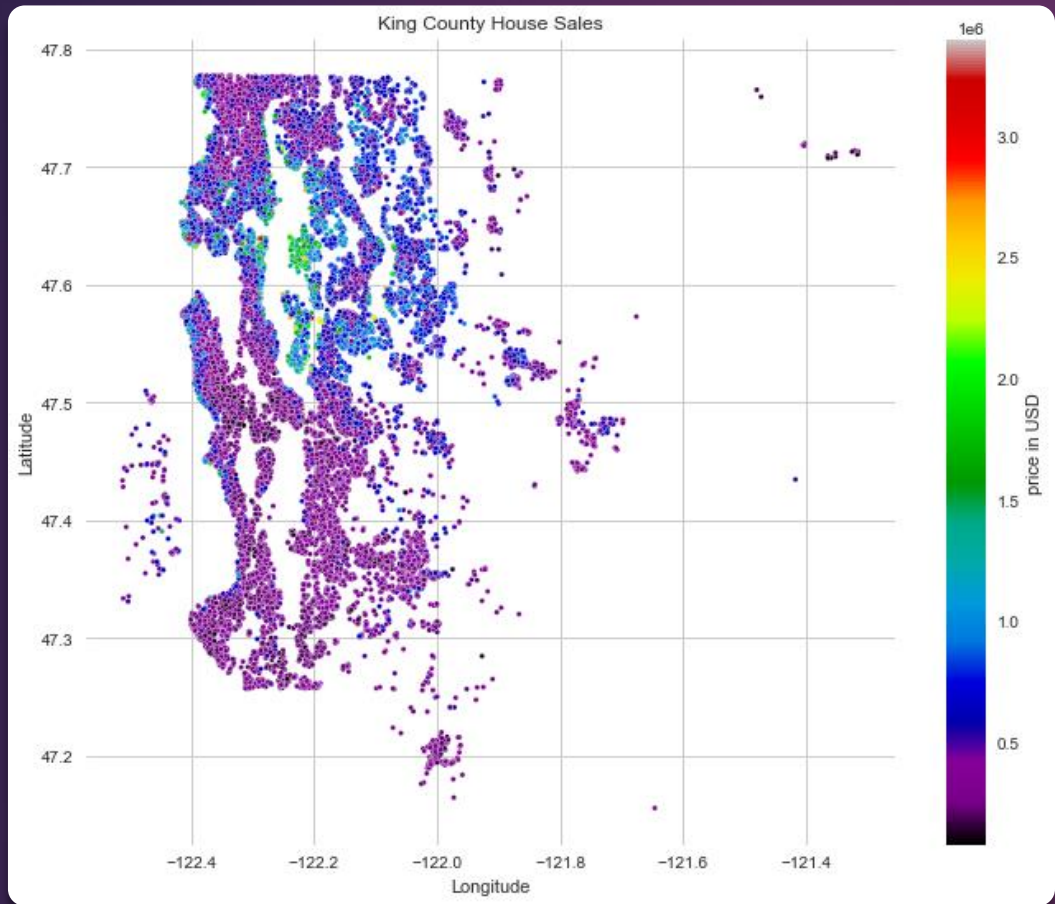
Zip  
Code





price	1.000000
bedrooms	0.287461
bathrooms	0.440446
sqft_living	0.633164
sqft_lot	0.084227
floors	0.242786
sqft_above	0.513823
sqft_basement	0.300752
yr_renovated	0.132620
lat	0.365129
long	-0.006183
sqft_living15	0.569679
sqft_lot15	0.045560
month_sold	-0.018926
year_sold	0.013148
age	-0.009455
renovated	0.105561

Baseline Model Generation



How does  
Location affect  
the Sale Price of  
a house?





# THANK YOU