

King County House Sale Price Prediction

Recommendation for Nicole Johnson(Buyer)

Recommended by Tanjina Afroj (Junior Data Scientist)

15 th May 2023



Agenda

- ❖ Introduction
- ❖ General data overview
- ❖ Hypothesis
- ❖ The initial results of the analysis
- ❖ Insights and recommendations for client
- ❖ Conclusion

Introduction

Tanjina Afroj

- Junior Data Scientist at XYZ Company
- XYZ is a renowned company with sensible employees
- Recommend clients with the best choice based on Data Science prediction
- The ultimate goal is a happy client.

I'm working on The King County House Prices dataset which includes data from May 2014 to May 2015 from the state of Washington DC. The King County House Sales dataset contains 21,597 rows and 21 columns.

Introduction

Nicole Johnson

My client Nicole wants to buy a house in King County.

Requirements: The house should be

1. lively
2. Located in central neighborhood
3. middle price range
4. right timing(within a year)

General data overview

Column name	Description
id	Unique identifier for a house
view	An index from 0 to 4 of how good the view of the property was
date	Date of the home sale
price	Price of each home sold
bedrooms	Number of bedrooms
bathrooms	Number of bathrooms
sqft_living	Square footage of the apartments interior living space
sqft_lot	Square footage of the land space
floors	Number of floors
waterfront	The apartment was overlooking the waterfront or not
condition	An index from 1 to 5 on the condition of the apartment,
grade	An index from 1 to 13, where 1-3 falls short of building construction and design, 7 has an average level of construction and design, and 11-13 have a high quality level of construction and design.
sqft_above	The square footage of the interior housing space that is above ground level
sqft_basement	The square footage of the interior housing space that is below ground level
yr_built	The year the house was initially built
yr_renovated	The year of the house's last renovation
zipcode	What zipcode area the house is in
lat	latitude
long	longitude
sqft_living15	The square footage of interior housing living space for the nearest 15 neighbors
sqft_lot15	The square footage of the land lots of the nearest 15 neighbors

The King County Housing Dataset contains information about the price, size, location, condition and other features of the houses in the Washington DC's King County area.

The size of the Dataset:

- 21,597 house sales between May 2014 –May 2015
- 23 variables

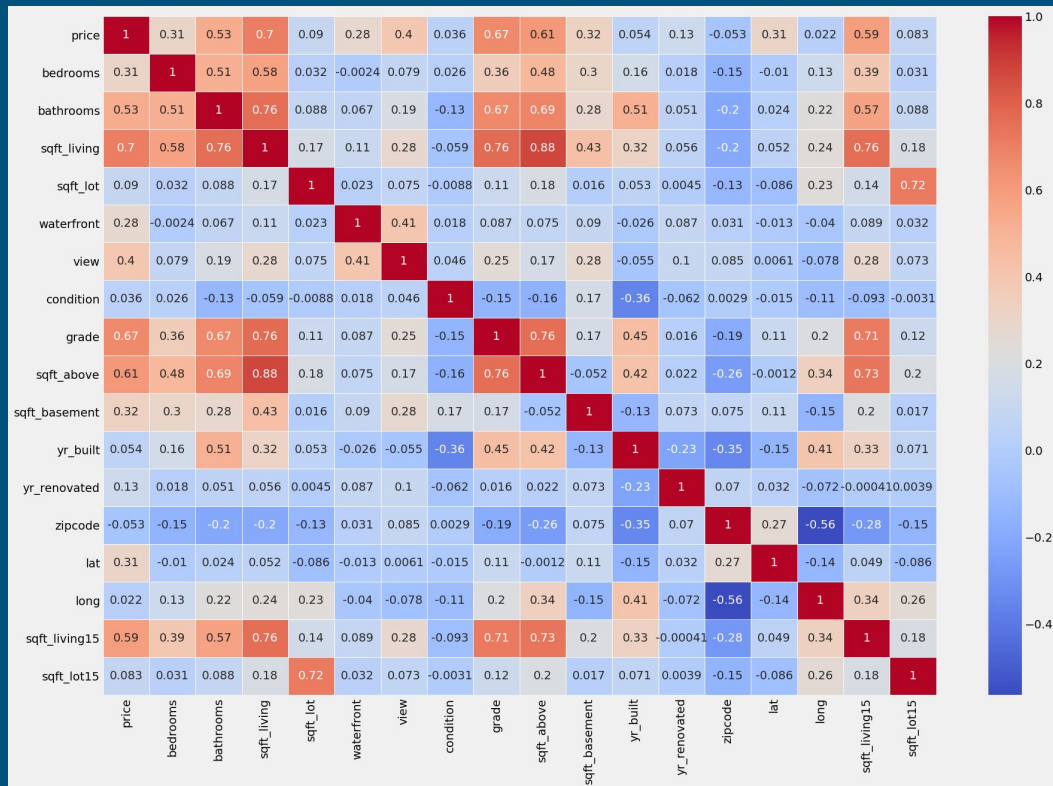
Missing values present in the column:

- waterfront
- view
- yr_renovated

Business Question & Hypothesis

1. How is location affecting the prices in King County and how should you price your house based on the location?
2. What is the factor that is not affecting the price of a house?
3. Is there any pick time when the price drop within the year? when and how could it benefits my client?
4. Does the age of a house affect its sale price in central and lively neighborhoods, for houses in the mid-price range?
5. What is the main factor outside the square foot that's driving the price

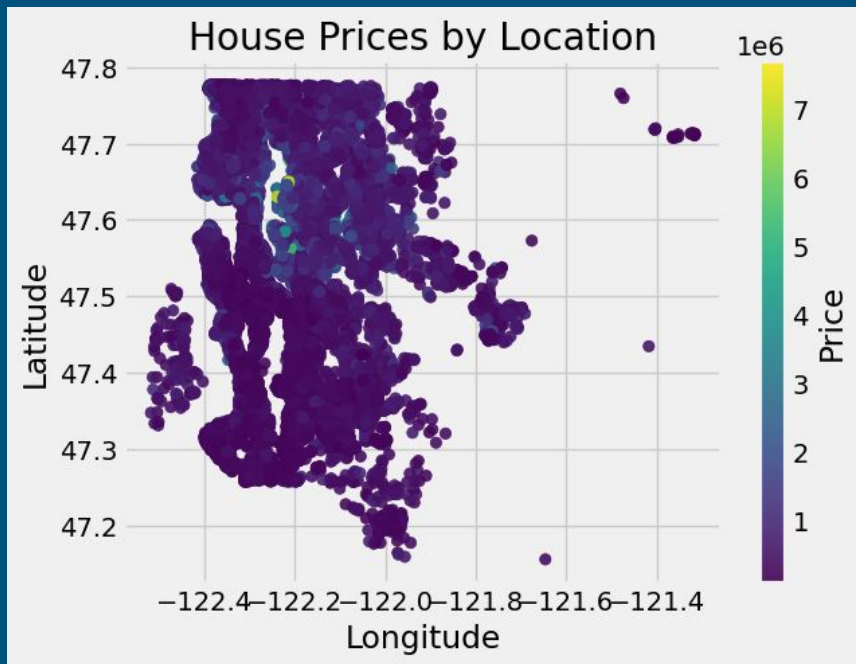
Correlation Heat Map



➤ Correlation between price and the square footage of the house

➤ Correlation between the grade and price

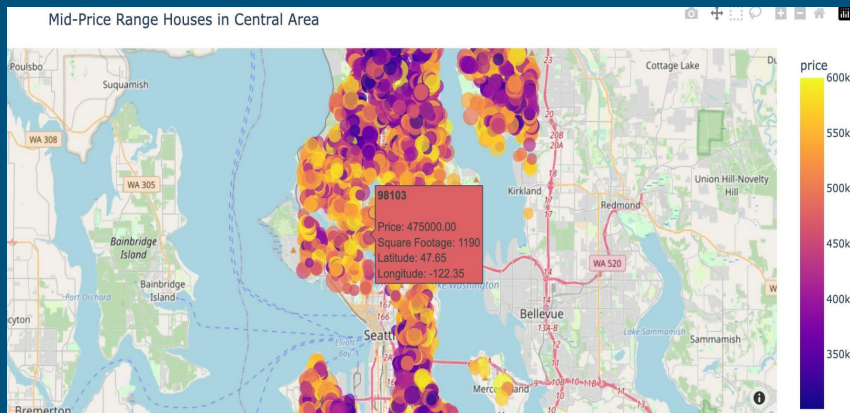
Checking Hypothesis #1



Q: How is location affecting the prices in King County and how should you price your house based on the location?

- From the scatter plot, we can observe that there is a positive correlation between the price of the house and its location. This means that as we move closer to the central area, the prices tend to increase. However, we can also see that there are certain locations that have high prices despite being farther away from the central area. This could be due to factors such as the neighborhood, amenities, or view.

Geographical Factor

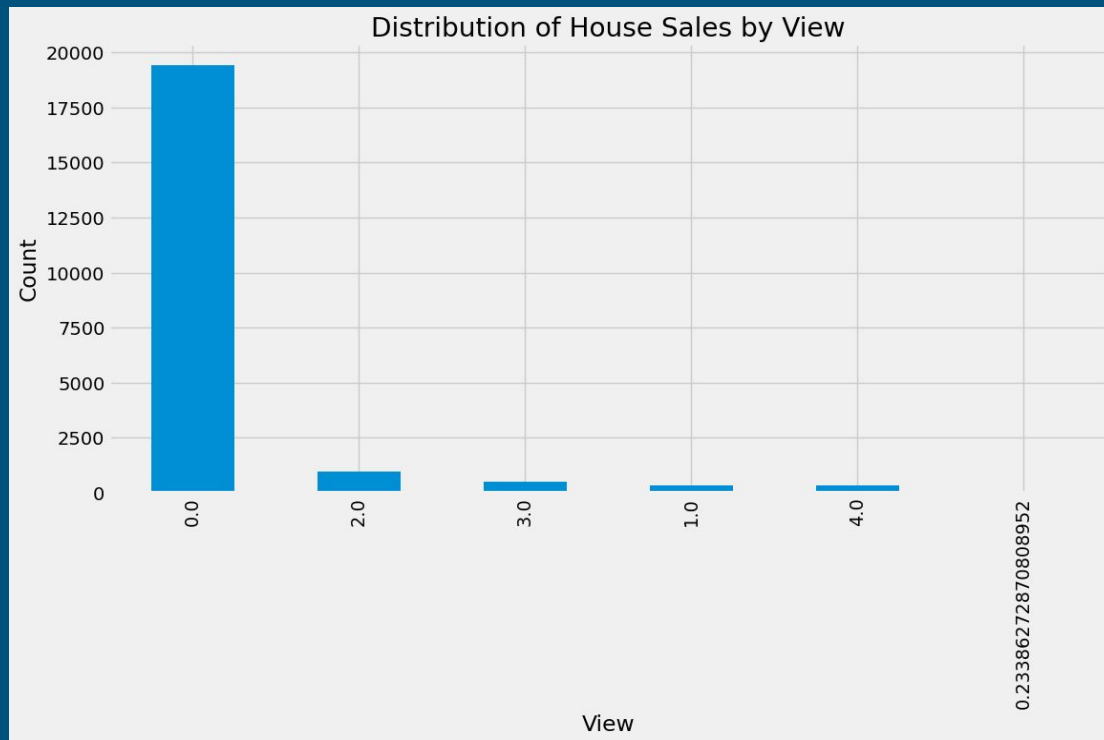


This scatter mapbox plot showing the locations of houses in the mid-price range in the central area of King County. The color of each dot represents the sale price, and the size represents the square footage of the living area

Location Matters!

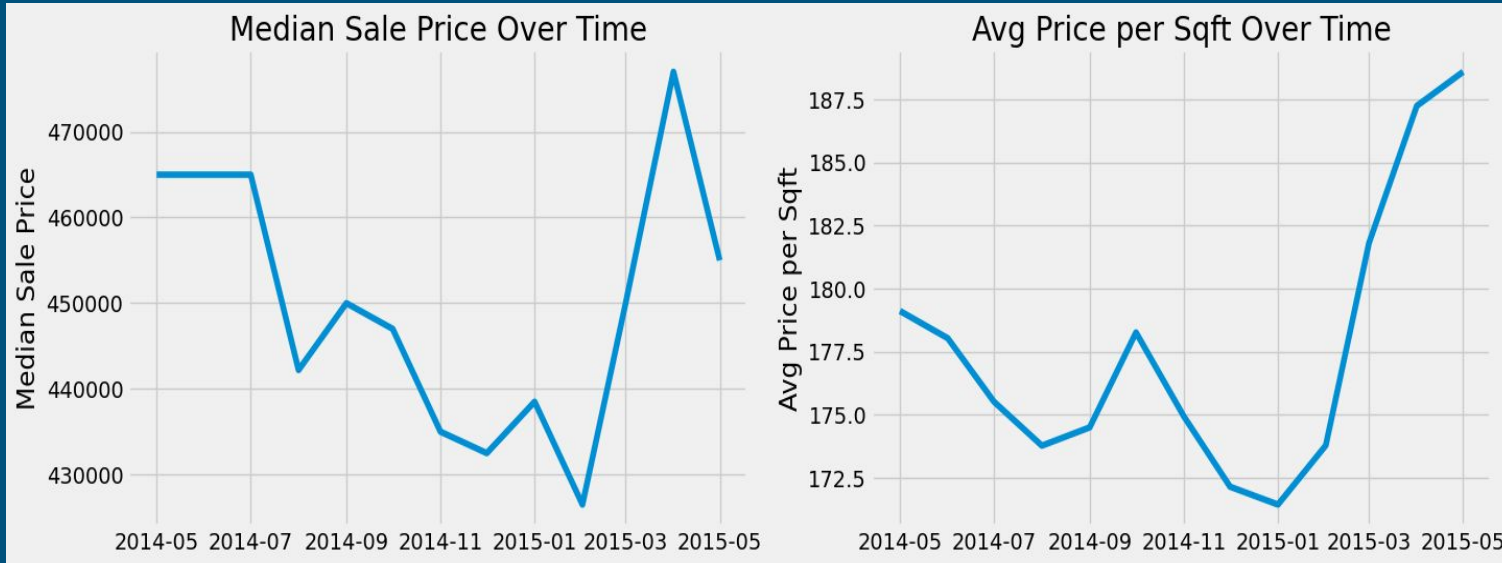
- ★ Higher Prices:
 - Houses between downtown Seattle and downtown Redmond tend to have higher prices
 - Tend to be clustered
 - Near water area
- ★ Lower Prices:
 - Tend to be spread out
 - Houses in the southern part of King County have lower prices

2. What is the factor that is not affecting the price of a house?



this shows that view does not impact the sales of a house much

3. Is there any pick time when the price drop within the year? when and how could it benefits my client?



There is a pattern to both graphs. Prices dip some time around January and February 2015. So, presumably the best timing of the year to buy a house can be January or February.

Conclusion with recommendation

1. Before Buying a house know your average price for your zipcode first. For example:

Zip code 98103 corresponds to the area around the neighborhood of Fremont, which is located in the north-central part of Seattle.

Zip code 98115 corresponds to the area around the neighborhoods of Wedgwood and View Ridge, which are located in the northeast part of Seattle.

Zip code 98117 corresponds to the area around the neighborhood of Ballard, which is located in the northwest part of Seattle.

2. House prices depends on square footage not the view.
3. Best time to buy a house can be January or February.

Thank You
