Exploratory Data Analysis of King County
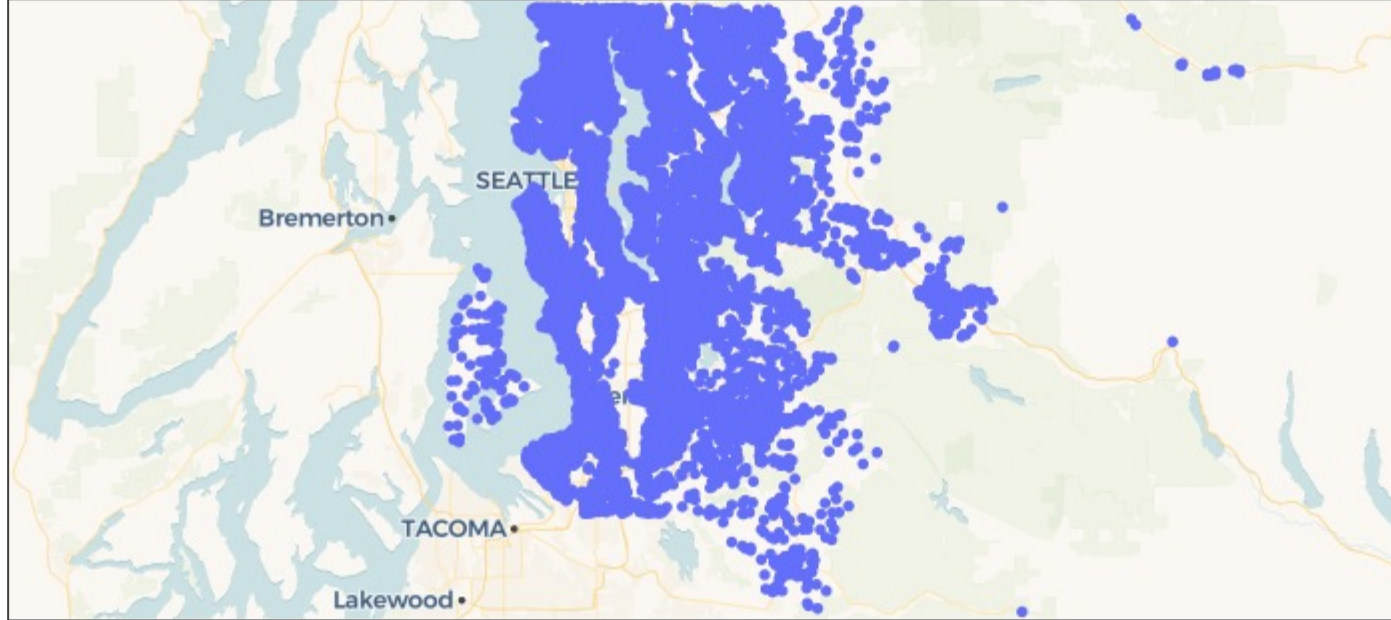
02/20/2026

by Isabell Durieux

🏠 Buying requirements

Houses of King County

Erin Robinson

- Invest in poor neighborhood

- Buying & selling

- Costs back + little profit

- Socially responsible

Goal: Recommendations for affordable houses in King County

Summary of Hypotheses

Hypothesis 1: The lower the grade and condition, the cheaper the house

Hypothesis 2: Cheap houses are accumulated in a specific council district

Hypothesis 3: Houses at the waterfront are more expensive compared to houses without a waterfront

Data analysis

Data overview

Target variable: Price

Main features: Bedrooms, bathrooms, sqft living, sqft lot, floors, sqft above, sqft basement, sqft living 15, sqft lot15, date of selling

Quality and location factors: Waterfront, view, condition, grade, zipcode, lat, long
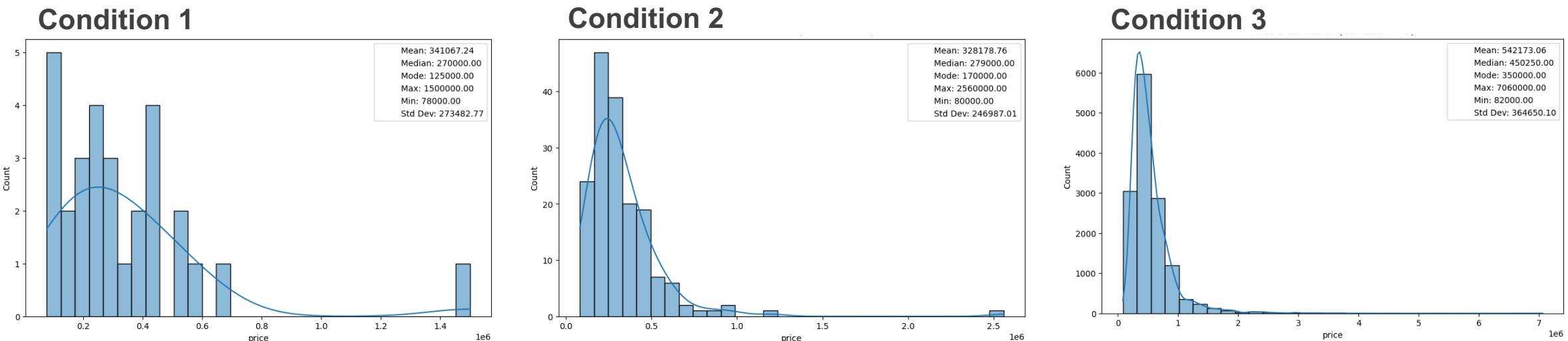
Age, Modernization: Year built, year renovated

Quality and location factors

Parameter	Information
Condition	How good is the overall condition of the house
Grade	Overall grade given to the housing unit, based on King County grading system
Lat	Latitude coordinate
Long	Longitude coordinate
Price	Is prediction target
Waterfront	House which has a view to a waterfront
Id	Unique identified for a house

» neue fische
School and Pool for Digital Talent

SPICED

Hypothesis 1: The lower the grade and condition, the cheaper the house

Price distribution of conditions 1-5

Condition 1

Mean: 341067.24
Median: 270000.00
Mode: 125000.00
Max: 1500000.00
Min: 78000.00
Std Dev: 273482.77

Condition 2

Mean: 328178.76
Median: 279000.00
Mode: 170000.00
Max: 2560000.00
Min: 80000.00
Std Dev: 246987.01

Condition 3

Mean: 542173.06
Median: 450250.00
Mode: 350000.00
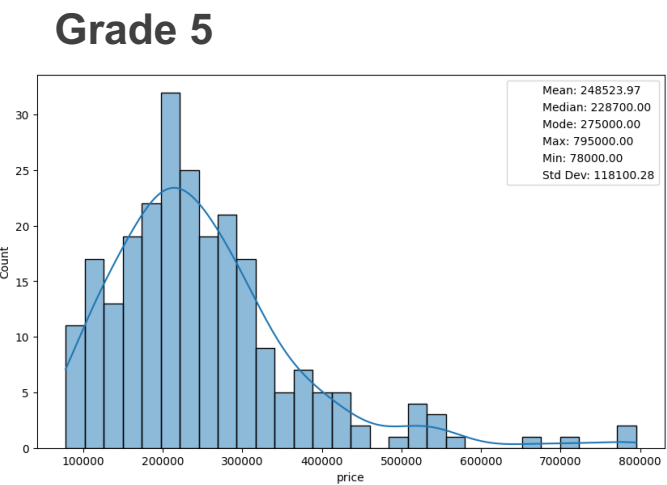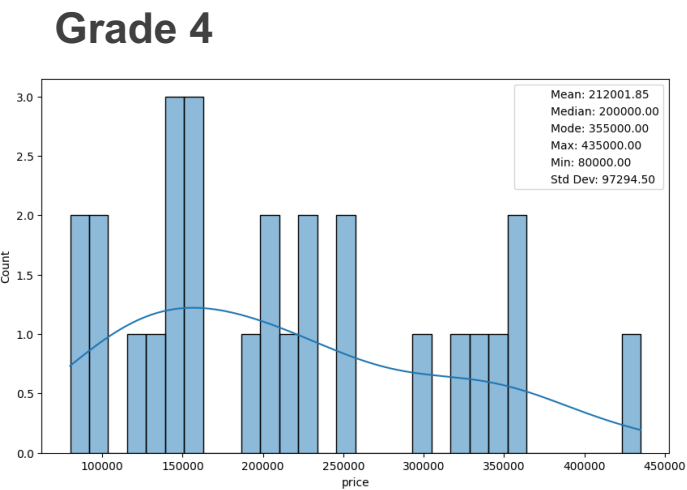Max: 7060000.00
Min: 82000.00
Std Dev: 364650.10

Condition 4 and 5 were plotted in the same way as the other conditions.

Condition	Mean price in mio dollar	Median price in mio dollar	Mode price in mio dollar	Max price in mio dollar	Min price in mio dollar	Std. dev. in mio dollar
1	0,341	0,270	0,125	1,500	0,078	0,273
2	0,328	0,279	0,170	2,560	0,080	0,247
3	0,542	0,450	0,350	7,060	0,082	0,365
4	0,521	0,440	0,450	7,700	0,089	0,359
5	0,613	0,526	0,550	3,650	0,110	0,411

- House prices generally increase with better condition

- The standard deviation increases with condition, suggesting greater price variability among homes in better condition

Mode = Price with the highest frequency
Median = Mean value, resistant to extreme prices

Hypothesis 1: The lower the grade and condition, the cheaper the house

Price distribution of grade 3-13

Grade 4


```
Mean: 212001.85
Median: 200000.00
Mode: 355000.00
Max: 435000.00
Min: 80000.00
Std Dev: 97294.50
```

Grade 5


```
Mean: 248523.97
Median: 228700.00
Mode: 275000.00
Max: 795000.00
Min: 78000.00
Std Dev: 118100.28
```
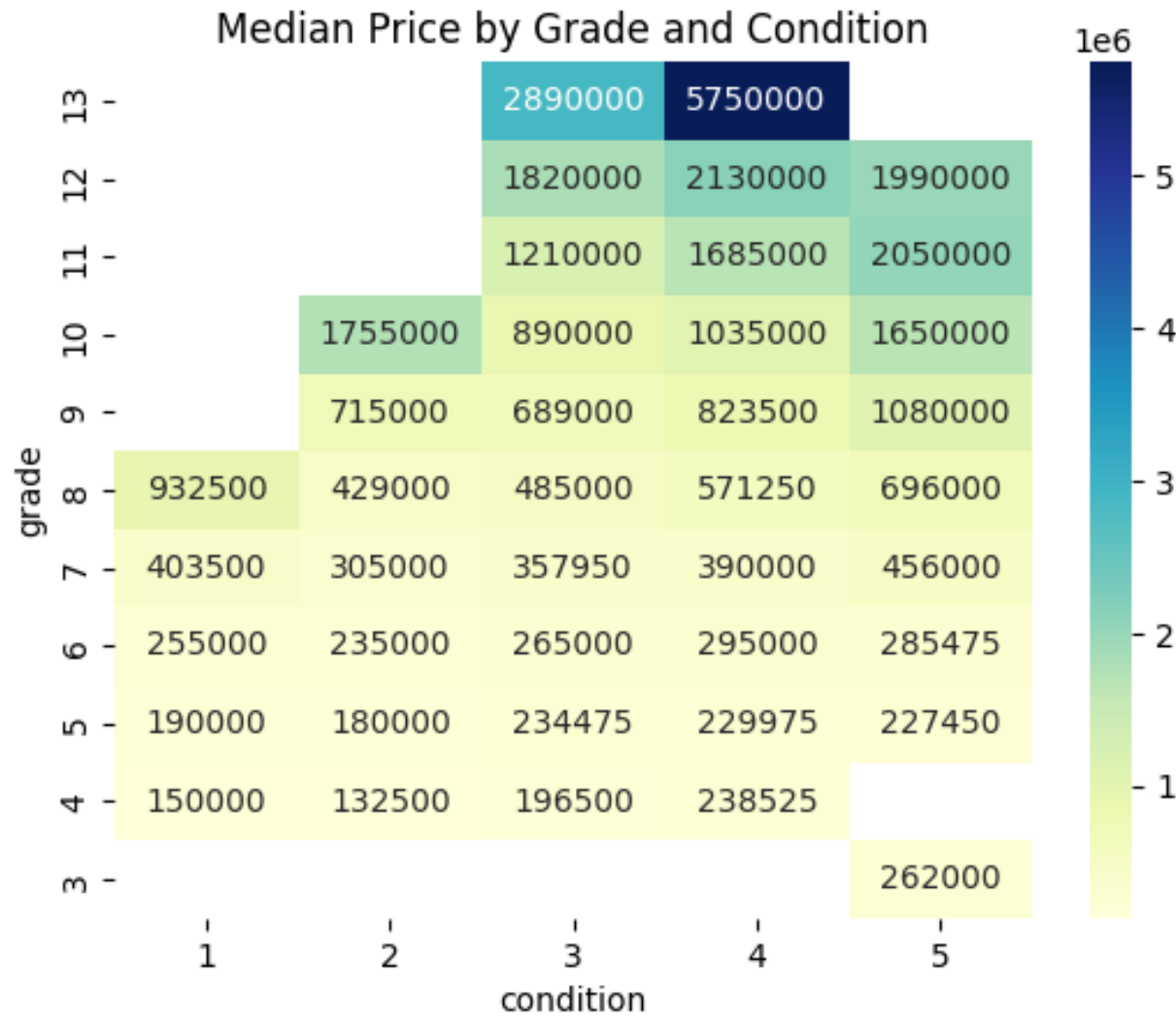
Grade 3, and 7-13 were plotted in the same way as the other conditions.

Grade	Mean price in mio dollar	Median price in mio dollar	Mode price in mio dollar	Max price in mio dollar	Min price in mio dollar	Std. dev. in mio dollar
3	0,262	0,262	0,262	0,262	0,262	nan
4	0,212	0,200	0,355	0,435	0,080	0,097
5	0,249	0,229	0,275	0,795	0,078	0,118
6	0,302	0,275	0,250	1,200	0,082	0,123
7	0,403	0,375	0,350	2,050	0,090	0,156
8	0,543	0,510	0,550	3,070	0,140	0,218
9	0,774	0,720	0,650	2,300	0,230	0,317
10	1,072	0,914	1,050	3,600	0,316	0,484
11	1,498	1,280	1,150	7,060	0,420	0,705
12	2,203	1,820	1,490	5,350	0,835	1,029
13	3,711	2,980	3,800	7,700	0,178	1,859

The table shows that house prices increase strongly with higher grade levels.

Lower grades and conditions are associated with lower median prices of homes
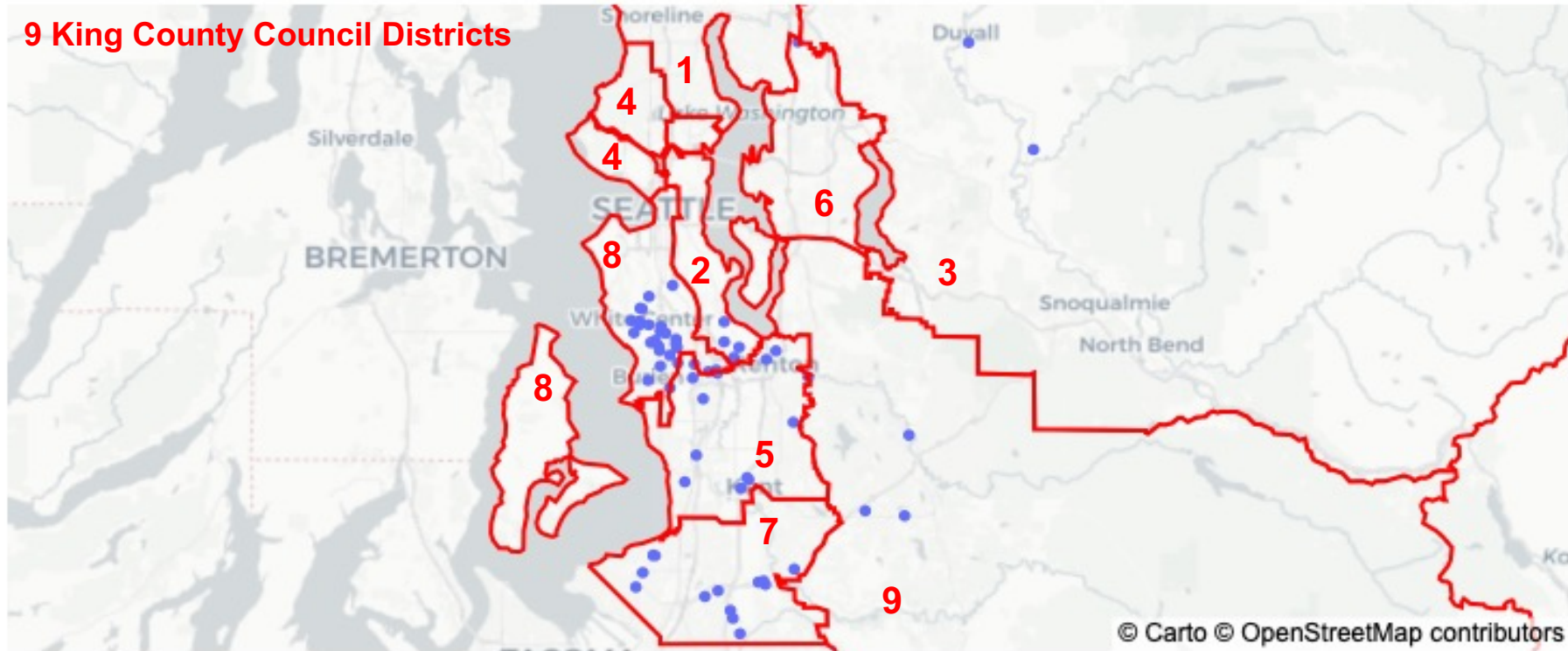

Median Price by Grade and Condition

- The heatmap shows the median house price by grade (y-axis) and condition (x-axis)

- Higher grades (especially 10–13) are associated with substantially higher median prices

- Condition has a positive but smaller impact compared to grade on median prices

→ **Hypothesis 1 is confirmed**
→ **Focus on lower grades (grade 4-7) and conditions (condition 1-3) to find houses**

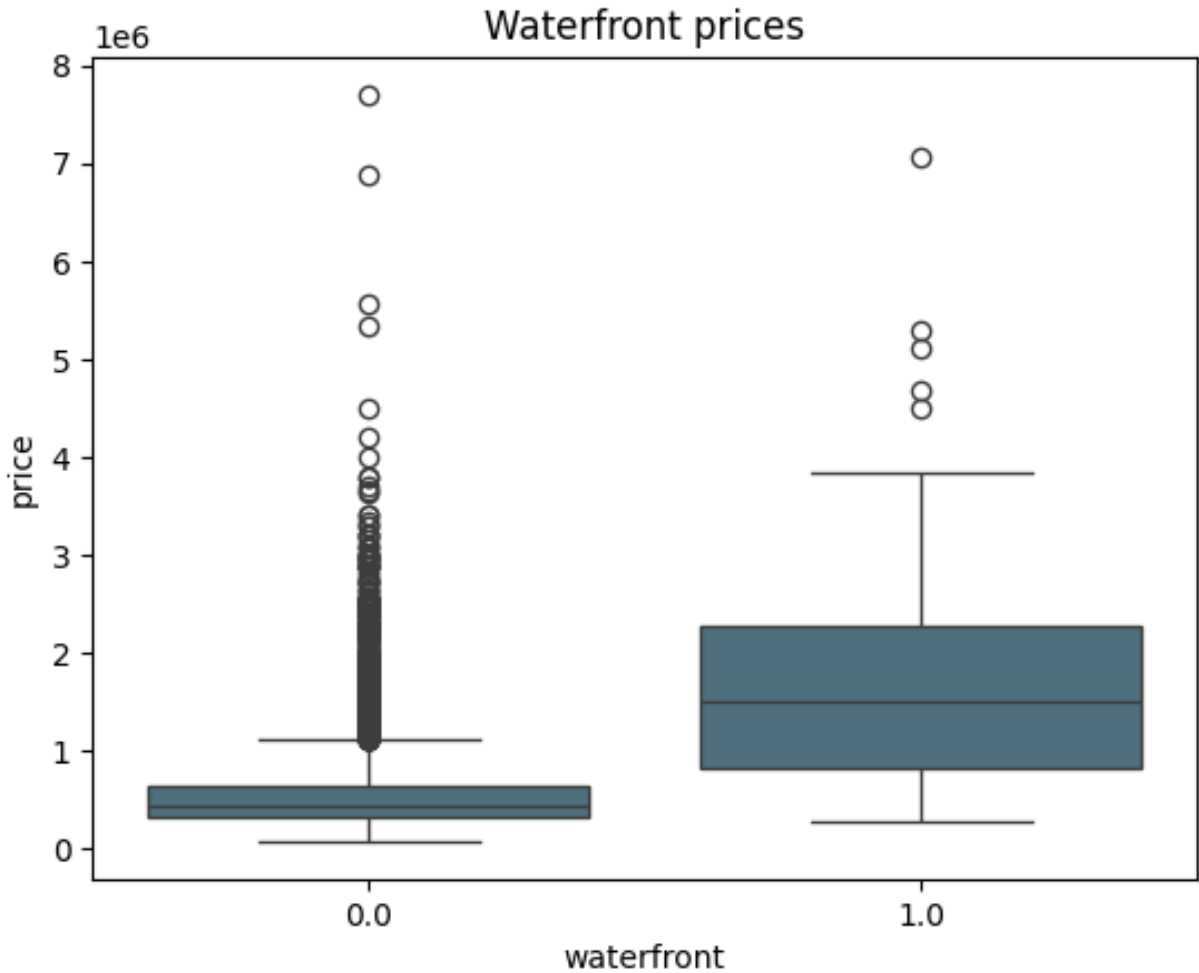# Hypothesis 2: Cheap houses are accumulated in a specific council district

Definition cheap houses: Up to 120,000 dollars

→ Most accumulation of cheap houses in district 5, 7, and 8

→ Hypothesis 2 is confirmed

Hypothesis 3: Houses at the waterfront are more expensive compared to houses wo a waterfront

Waterfront prices

Waterfront	Mean price	Median price	Max price	Min price	Std. dev. price
Yes	0.533	0.450	7.700	0.078	0.345
No	1.717	1.510	7.060	0.285	1.145

- Prices of houses with a waterfront are 3x higher compared to houses without a waterfront

- Focus on houses without a waterfront

- Hypothesis 3 is confirmed

Recommendations for King County

Summary for affordable houses in King County

→ Most accumulation of affordable houses in district 5, 7, 8

→ Use the King Counting grading system, focus on low

   grades (grade 4-7) and conditions (condition 1-3)

→ Houses without waterfront are more likely to be cheaper

→ Lower prices on the King County real estate market

   between November and February

»» neue fische
School and Pool for Digital Talent   **SPICED**