

House recommendations

For Larry Sanders



Alexander Engelhardt, 15.09.2022

The stakeholder

Larry Sanders

- 42 years old, 2 kids
- Wants to buy a house for his family in King County, Washington
- **What Larry asked for:**
 - Waterfront view
 - Isolated but central w/o kids because of germ phobia
 - Enough rooms to accommodate family
- **Limiting factors:**
 - Low budget (limit at 1.5M \$)



The data

Housing data from Kings County, Washington

- Information on 21597 houses described in 21 columns
- House records between 02.05.2014 and 27.05.2015

Overview table of data for selected features

	id	date	price	bedrooms	bathrooms	sqft_living	waterfront	condition	yr_built	zipcode	lat	long	sqft_lot15
0	7129300520	2014-10-13	221900	3	1	1180	NaN	3	1955	98178	47.511	-122.257	5650
1	6414100192	2014-12-09	538000	3	2	2570	0.000	3	1951	98125	47.721	-122.319	7639
2	5631500400	2015-02-25	180000	2	1	770	0.000	3	1933	98028	47.738	-122.233	8062
3	2487200875	2014-12-09	604000	4	3	1960	0.000	5	1965	98136	47.521	-122.393	5000
4	1954400510	2015-02-18	510000	3	2	1680	0.000	3	1987	98074	47.617	-122.045	7503

Based on his criteria, which houses can be recommended to Larry?

Research questions and hypotheses

1) What is the availability and price of houses with a waterfront view?

- Houses with waterfront are less common than those without.
- Houses with waterfront are more expensive than those without.

2) What is the price of houses in central, isolated neighborhoods?

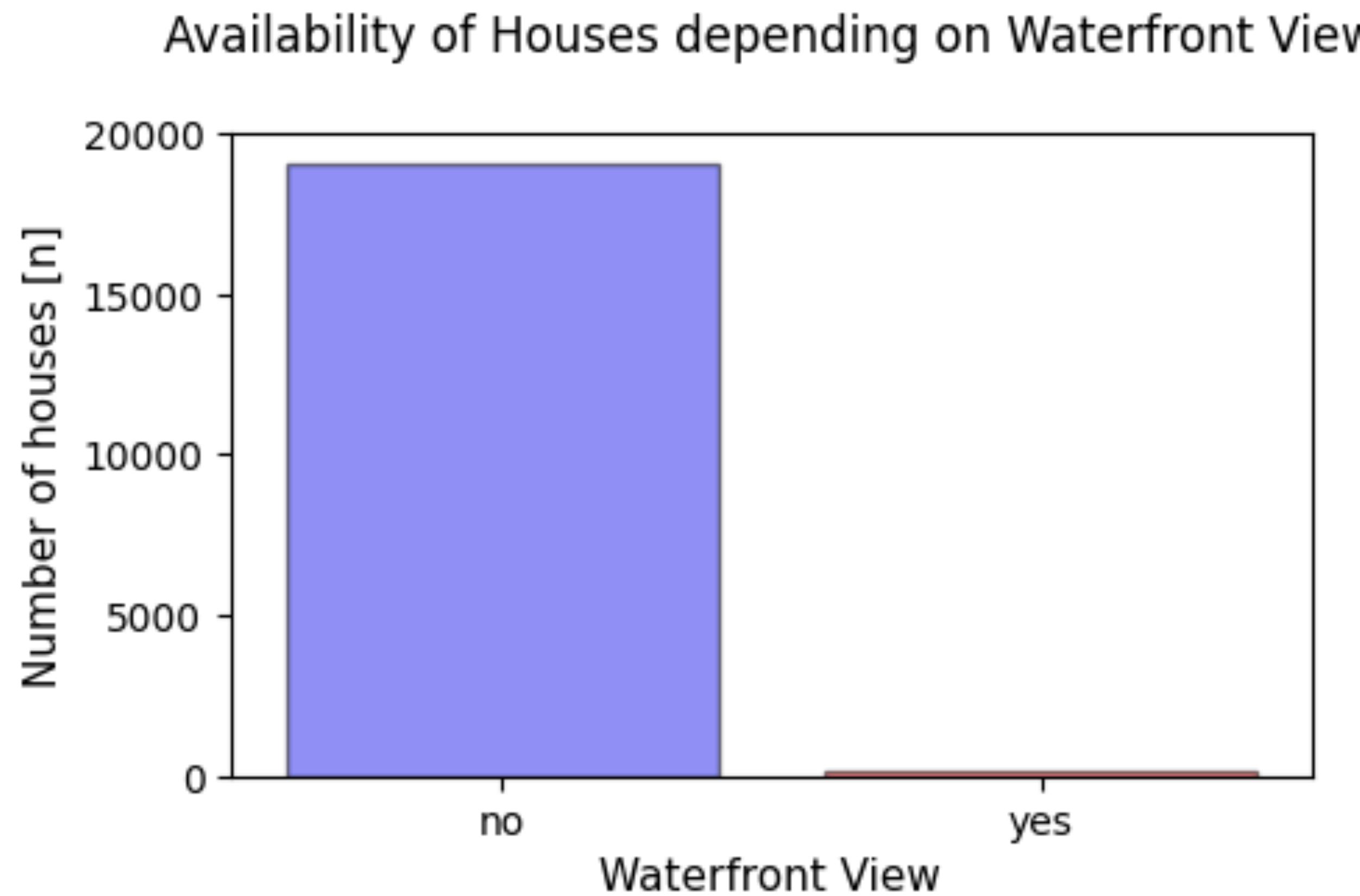
- Houses at central, but isolated neighborhoods are more expensive than at other sites.

3) How does the number of bed- and bathrooms affect housing price?

- Housing prices increase with increasing number of bed- and bathrooms.

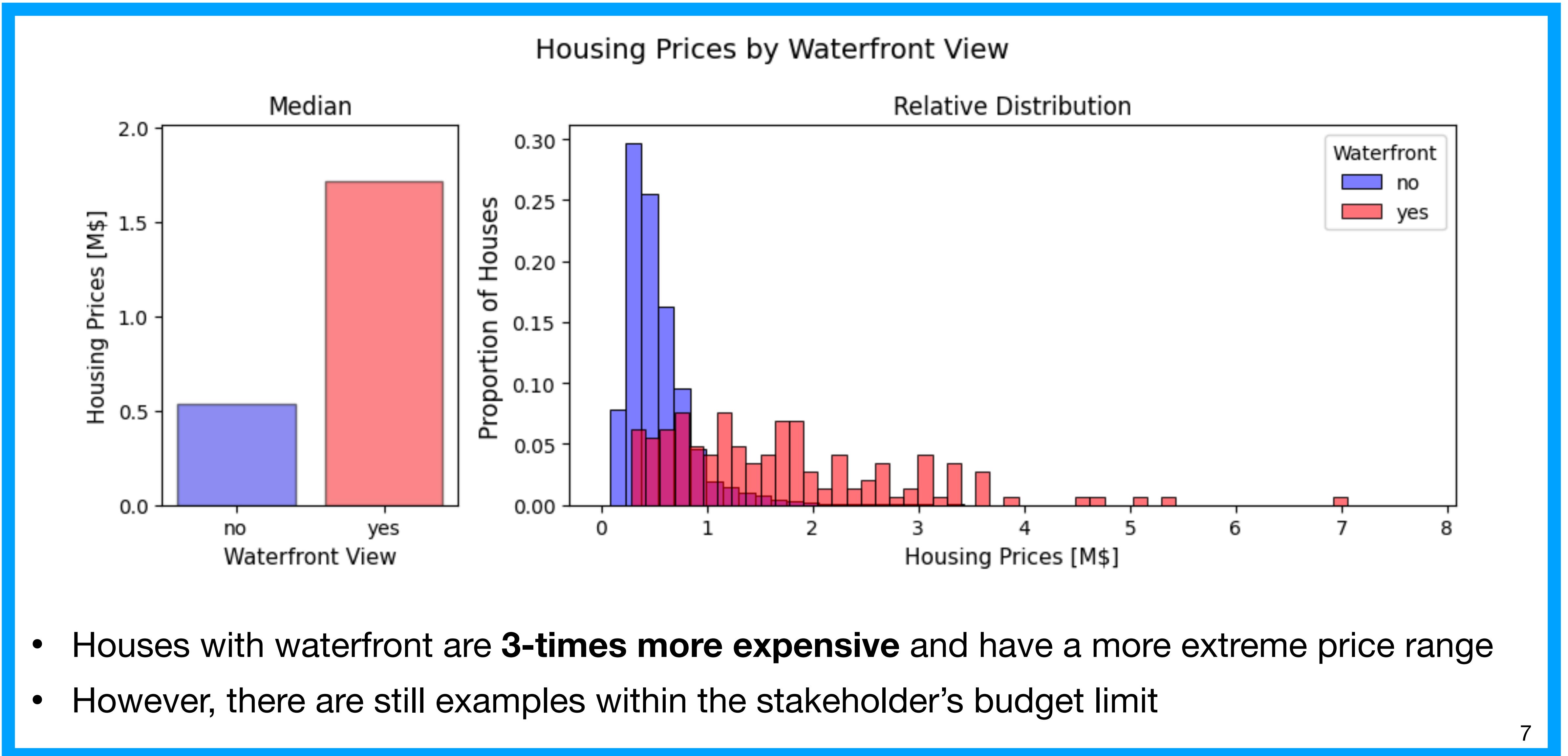
Q1: What is the availability and price of houses with a waterfront view?

Houses with waterfront are less common than those without



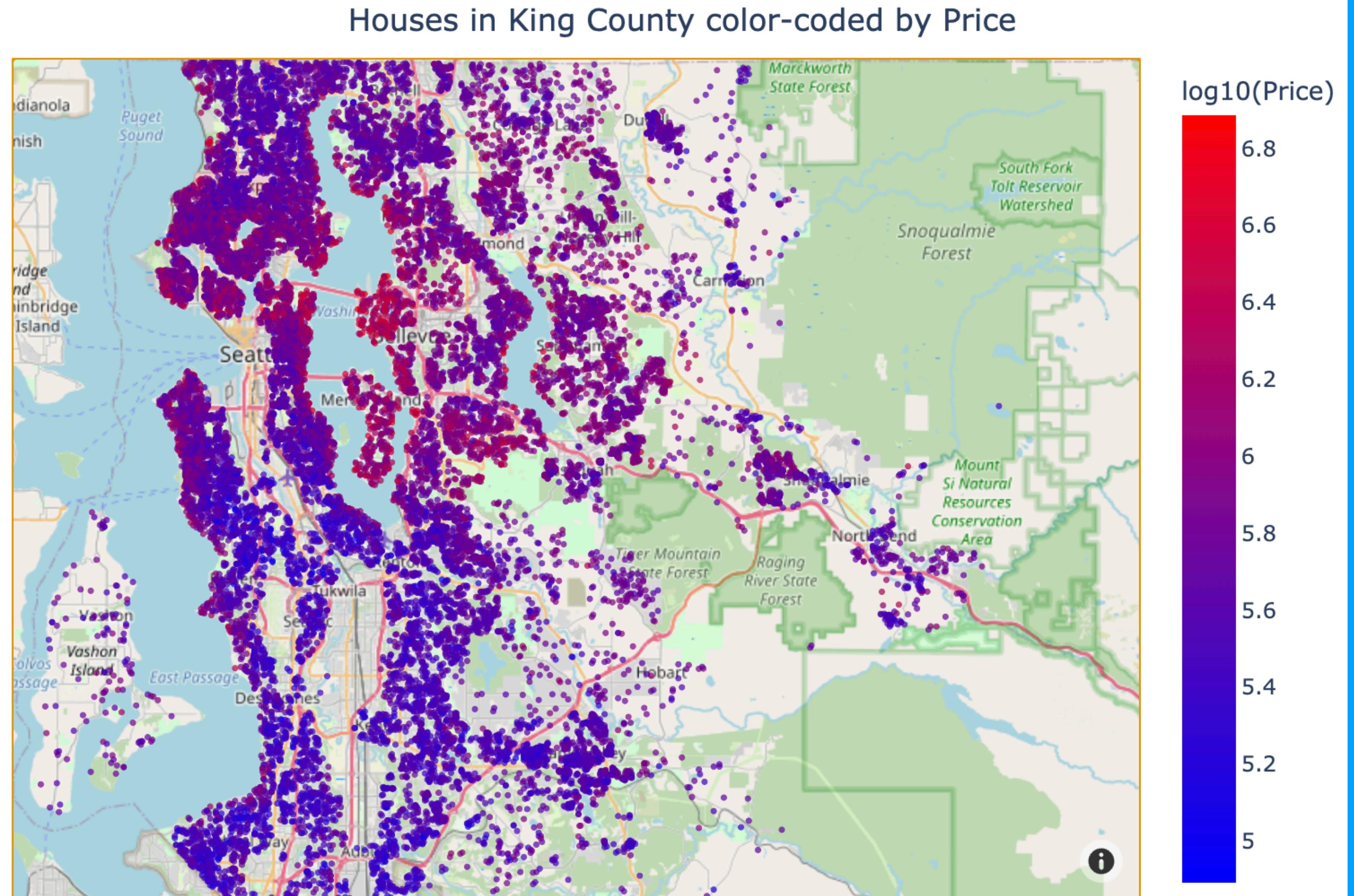
- The majority of houses does not have a waterfront
- In total, there are only **146** houses fulfilling the stakeholder's preference

Houses with waterfront are more expensive than those without



Houses with waterfront are more expensive than those without

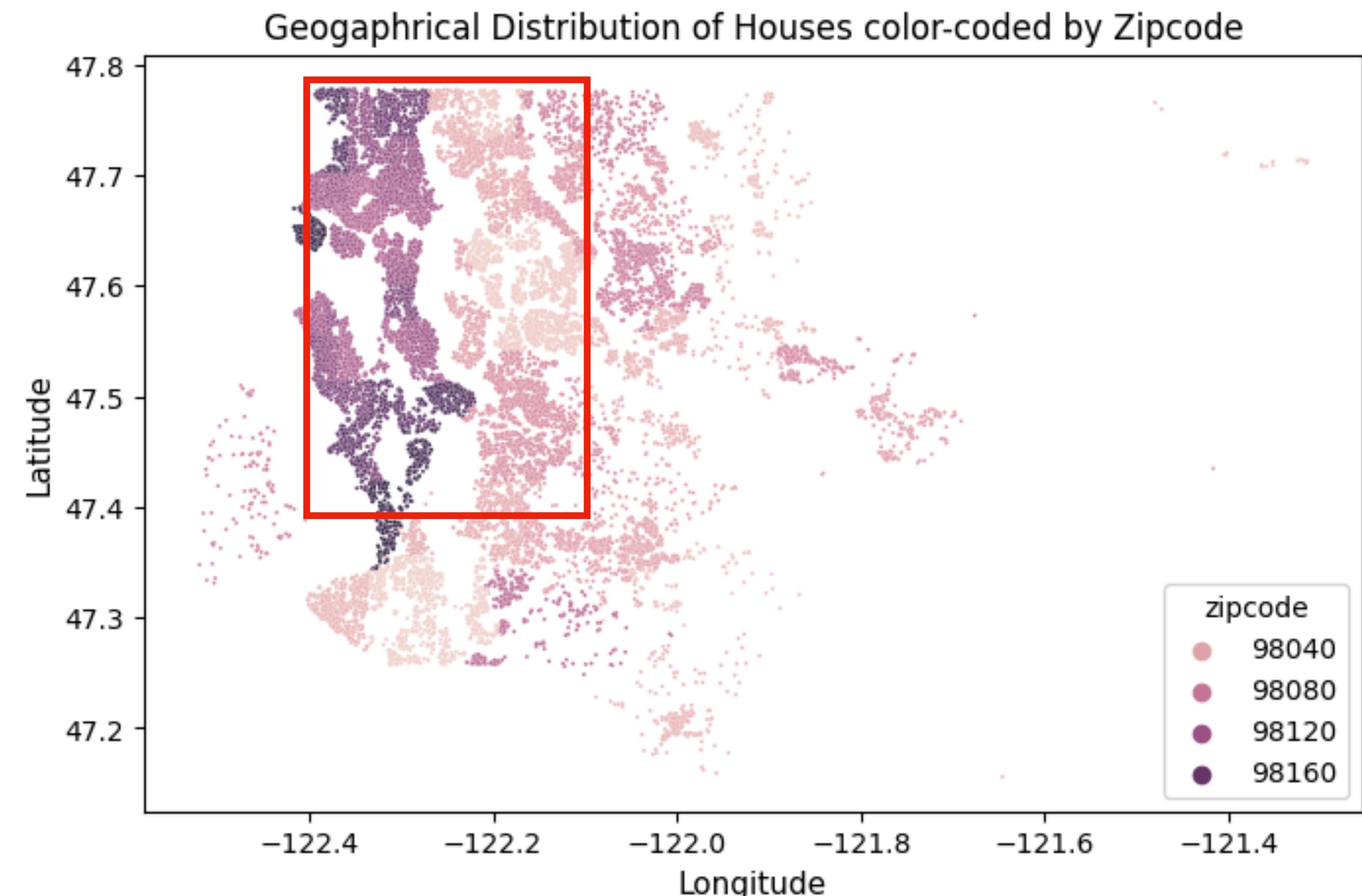
- Consistent with the previous data, a geospatial map shows a clustering of the most expensive houses near the coastline



Q2: What is the price of houses in central, isolated neighborhoods?

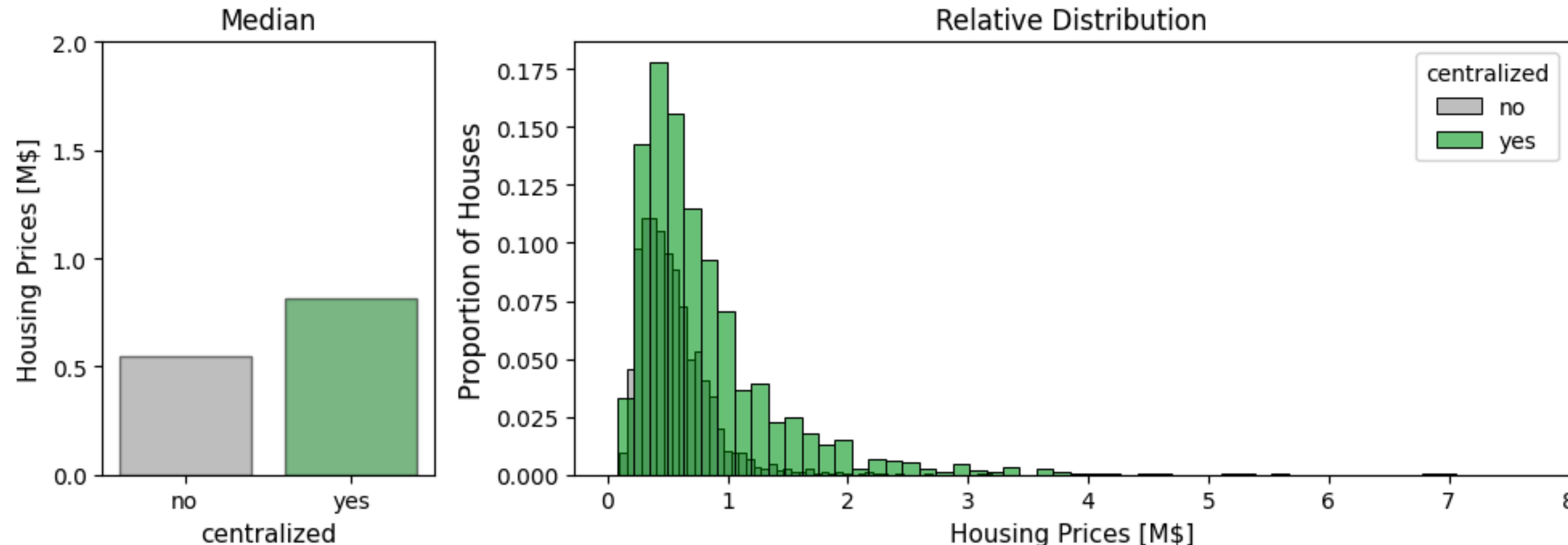
Extracting features from data to define central, isolated neighbourhoods

- **Square footage of land lots from 15 nearest neighbours**
 - used to define feature ‘isolated’
 - threshold: 75% quantile
- **Latitude and longitude**
 - used define feature ‘centralized’
- In total, **2637** houses meeting both criteria



Houses at central, isolated neighborhoods are more expensive than at other sites.

Prices of Houses in Isolated Neighborhood depending on Location



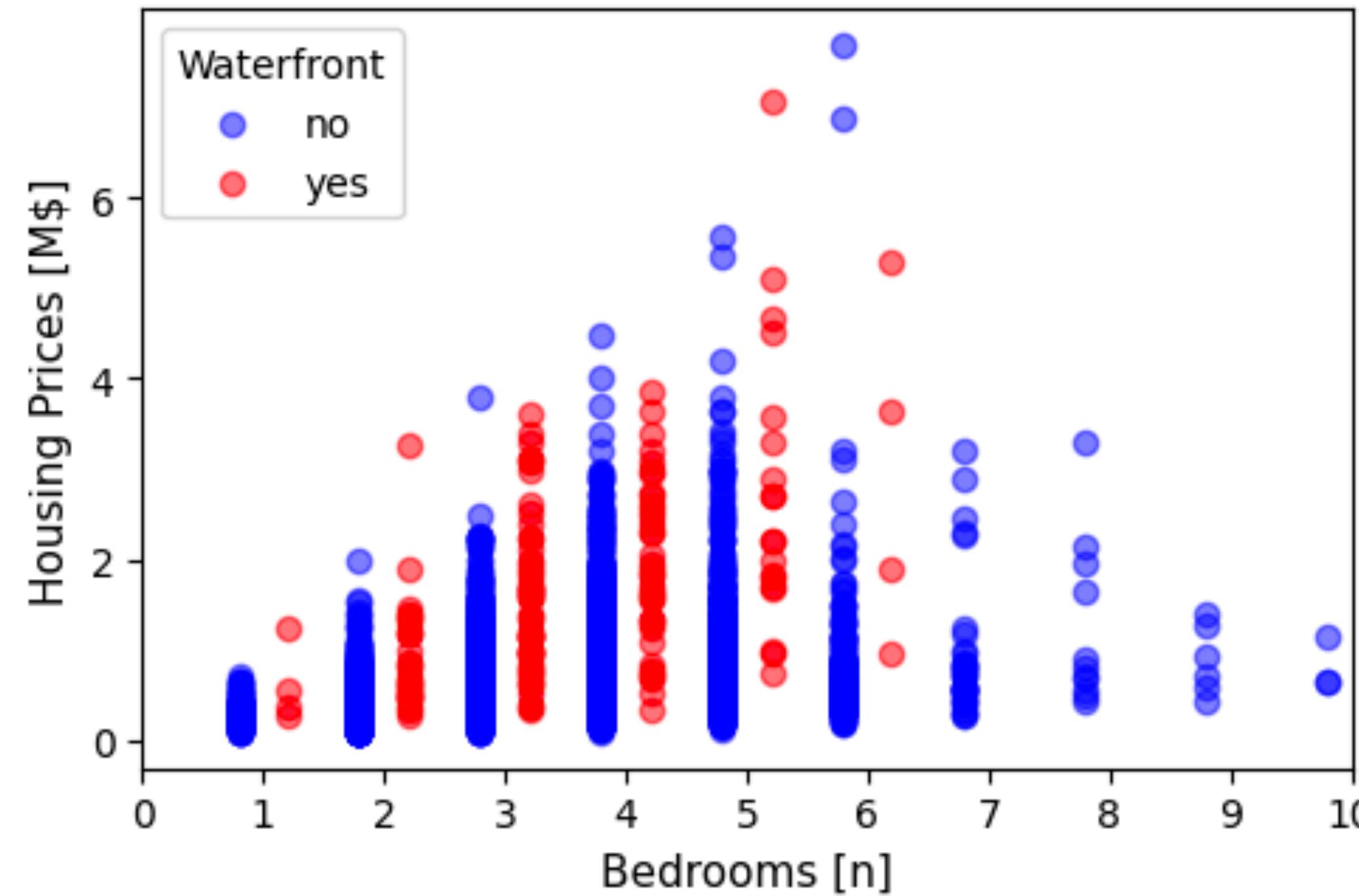
- Houses at central, isolated neighborhoods are only slightly more expensive
- Thus, prices are within the stakeholder's budget limit for the majority of those houses

Q3: How does the number of bed- and bathrooms affect housing price?

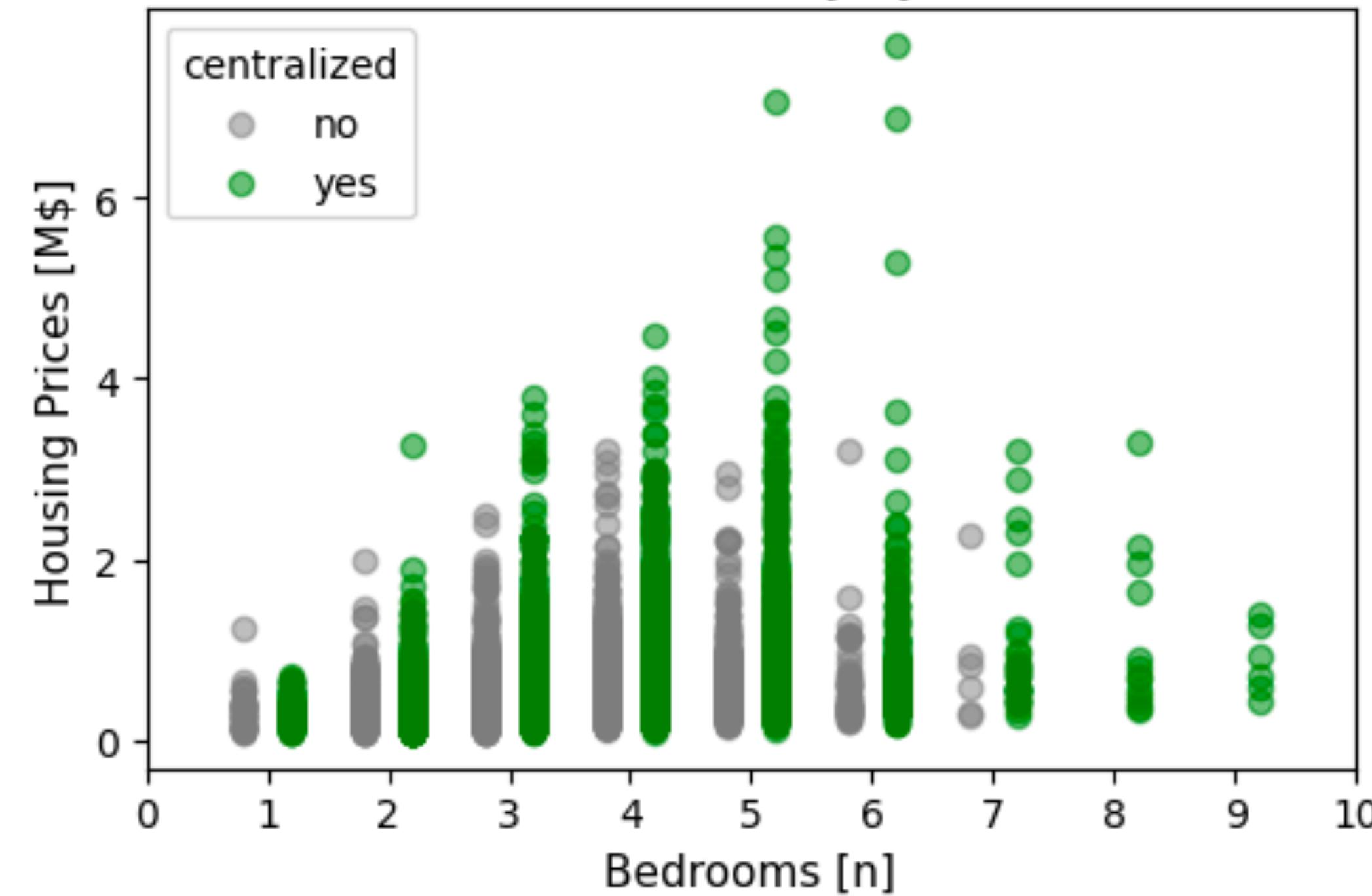
Housing prices increase with increasing number of bed- and bathrooms

Change in Price Depending on Bedroom Number

Houses with or without Waterfront



Isolated Houses only by Location



- Prices show curvilinear relationship with number of bedrooms (similar for bathrooms)
- Importantly, many houses have enough bedrooms and fit the stakeholder's price range and other criteria

What houses can be recommended based on those criteria?

Selecting houses based on stakeholder criteria

- **Filtering based on stakeholder's housing criteria and budget limit:**
 - Maximum price: 1.5 M\$
 - Waterfront view
 - Central, isolated neighbourhood
 - At least 3 bedrooms
 - At least 1 bathroom

Recommendations

	id	price	condition	bedrooms	bathrooms	sqft_living	yr_built	sqft_lot15
18833	7631800110	380000	3	3	2	1980	1984	17313
17460	6329000185	540000	3	3	2	2600	1912	14700
13225	2623039018	685000	3	4	1	1550	1930	22047
15811	8073000491	700000	3	4	1	1950	1957	13122
11155	7567600030	750000	4	5	1	2640	1954	11942
11832	2623039082	770000	4	3	3	2050	1930	12200
14174	7631200085	947500	5	3	2	2980	1917	12514
17881	7110000068	975000	3	6	2	2520	1954	26809
19955	2623039019	988500	3	3	2	2015	2007	12310
1259	7636800041	995000	3	3	4	4380	1968	18512
4291	7781600025	1160000	4	3	2	2490	1961	24121
11103	7936500172	1180000	3	3	2	1970	1937	34689
2073	3623500260	1200000	4	3	1	1560	1928	16710
16173	6117500460	1310000	3	4	2	2680	1956	19964
6471	7936500190	1340000	3	4	3	2130	1955	28598
15826	7351200050	1340000	3	4	1	2300	1934	13342
18899	9253900408	1400000	3	3	2	3130	1947	18782

- Based in Larry's criteria, we can recommend a set of 17 houses
- All recommended houses are in average condition or better (values ≥ 3)

Contact Information

- If you liked my work, we can stay connected!
- Just reach out at:

Alexander Engelhardt
Data Science Death Road 666
60066 Data Science Hell
Mobile: +666/66666606

