EDA PROJECT

KING COUNTY, SEATTLE (USA)
HOUSING DATA

LUKAS KAUFMANN

22.09.2023

I. introduction

II. data overview

III. hypotheses

IV. analysis & visualisation

V. conclusions

I. CLIENT INTRODUCTION

| William Rodriguez
| Buyer
| Two people (want two houses)
| country house (best timing & non-renovated)
| & city house (fast & central location)

** added information after talk with clients:*

| *budget of 400.000 $ per house*
| *looking for two single apartments (limited size)*
| *no multi-story houses / appartments (live accessibly in the future)*
| *want to renovate their country house themselves*
| *city house should be in good condition*

II. DATA OVERVIEW

- House & Sales Details
- Data collected between 2014/05 and 2015/05
- >21.000 Rows and >20 Columns

III. HYPOTHESES

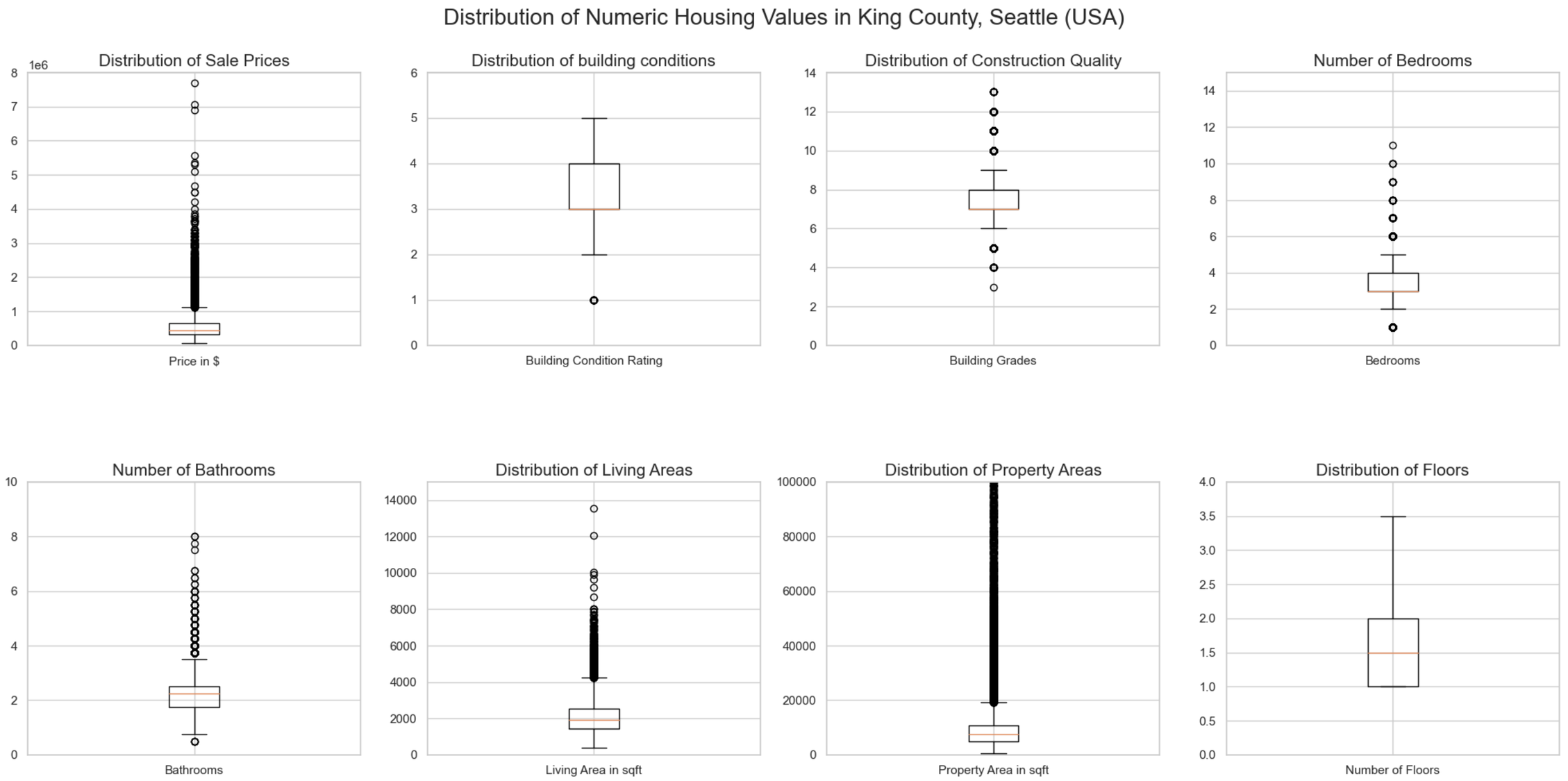
A. Questions:

- When is the best time to buy a (country) house?
- Are the houses on the country side bigger?
- Are the houses in the city more expensive?
- What effects the price of a house?

B. Hypotheses:

- There is no time of the year where it is “better” to buy a house. (price drops within a year or month, etc.)
- House prices (in King County) generally increase over time.
- The country houses are bigger compared to the city houses.
- The city houses will be more expensive than the country houses.
- The newer the house, the better building grade it has. (construction quality / meets code).
- The better rated a house, the more expensive it is.

IV. ANALYSIS & VISUALISATION



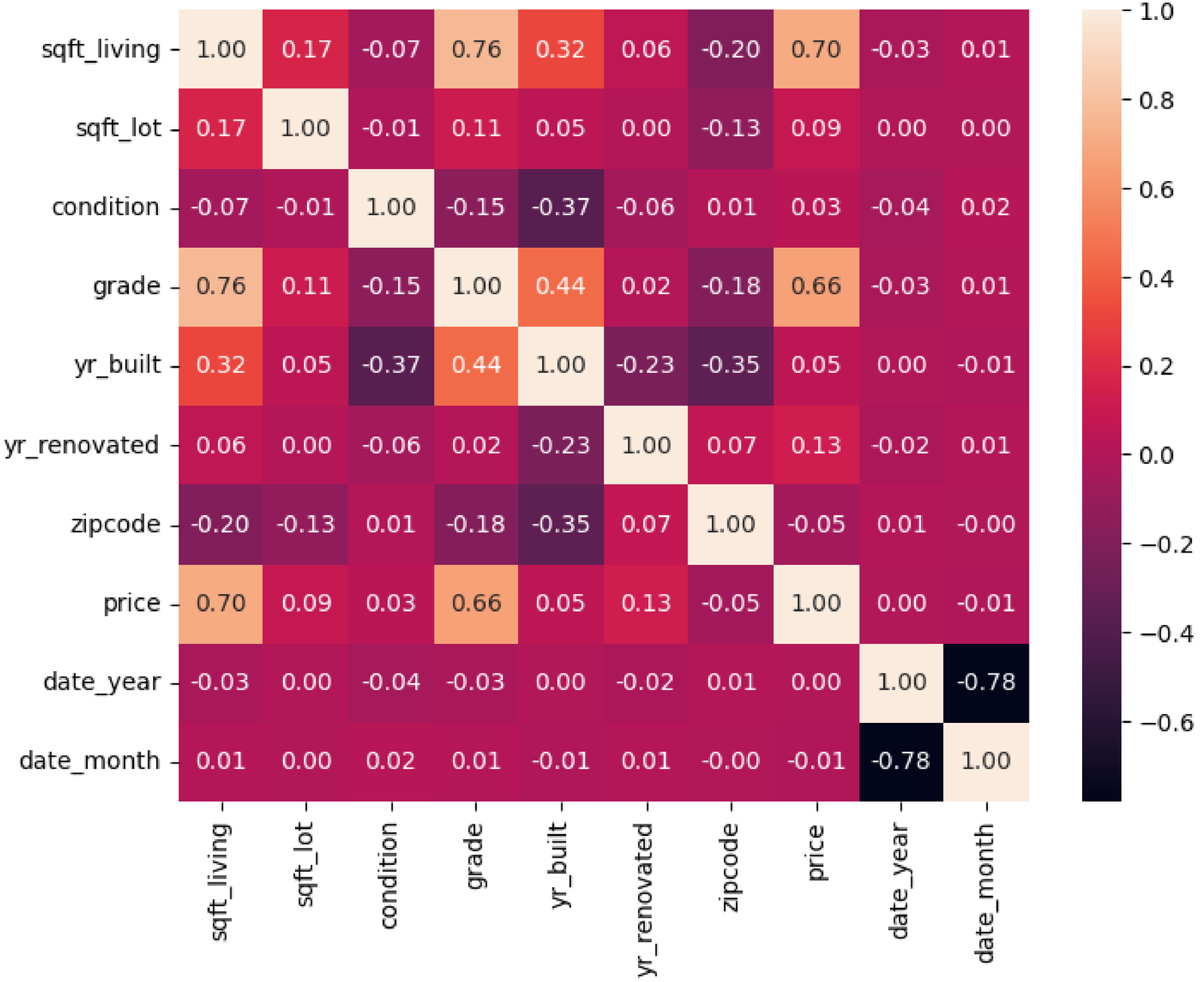
boxplot - average distributions of numeric housing values

IV. ANALYSIS & VISUALISATION

- year_built has an effect on the grade and on sqft_living

- grade and sqft_living have a high correlation to the price of the house

- grade and sqft_living are related to each other as well



heatmap - feature correlation values

IV. ANALYSIS & VISUALISATION

- older houses are smaller in size

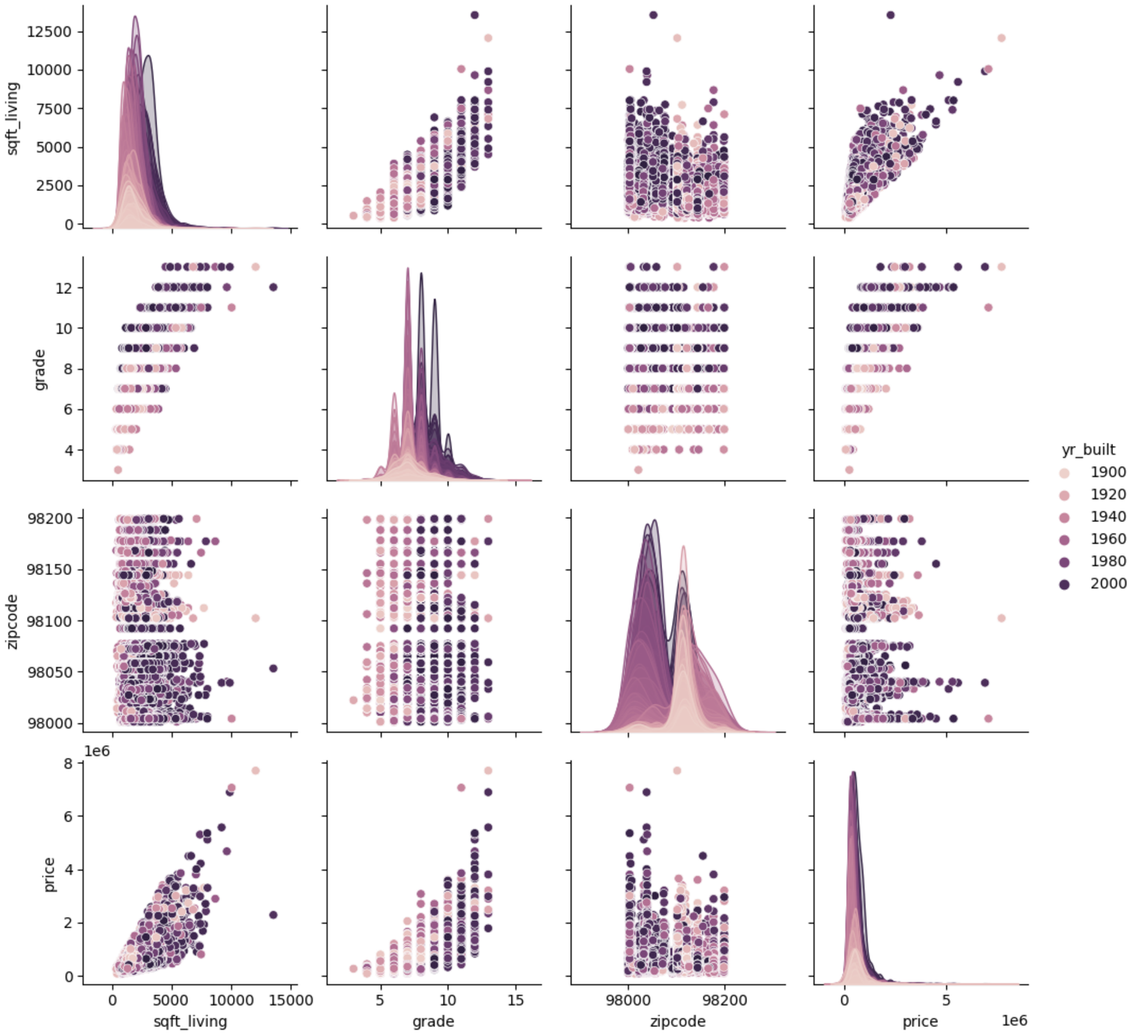
- newer houses have better grades

- The better rated and the bigger a house, the more expensive it is

- The price range is rather evenly distributed throughout the zipcodes

- Is the time of the sale related to the price?

pairplot - correlation visualisations

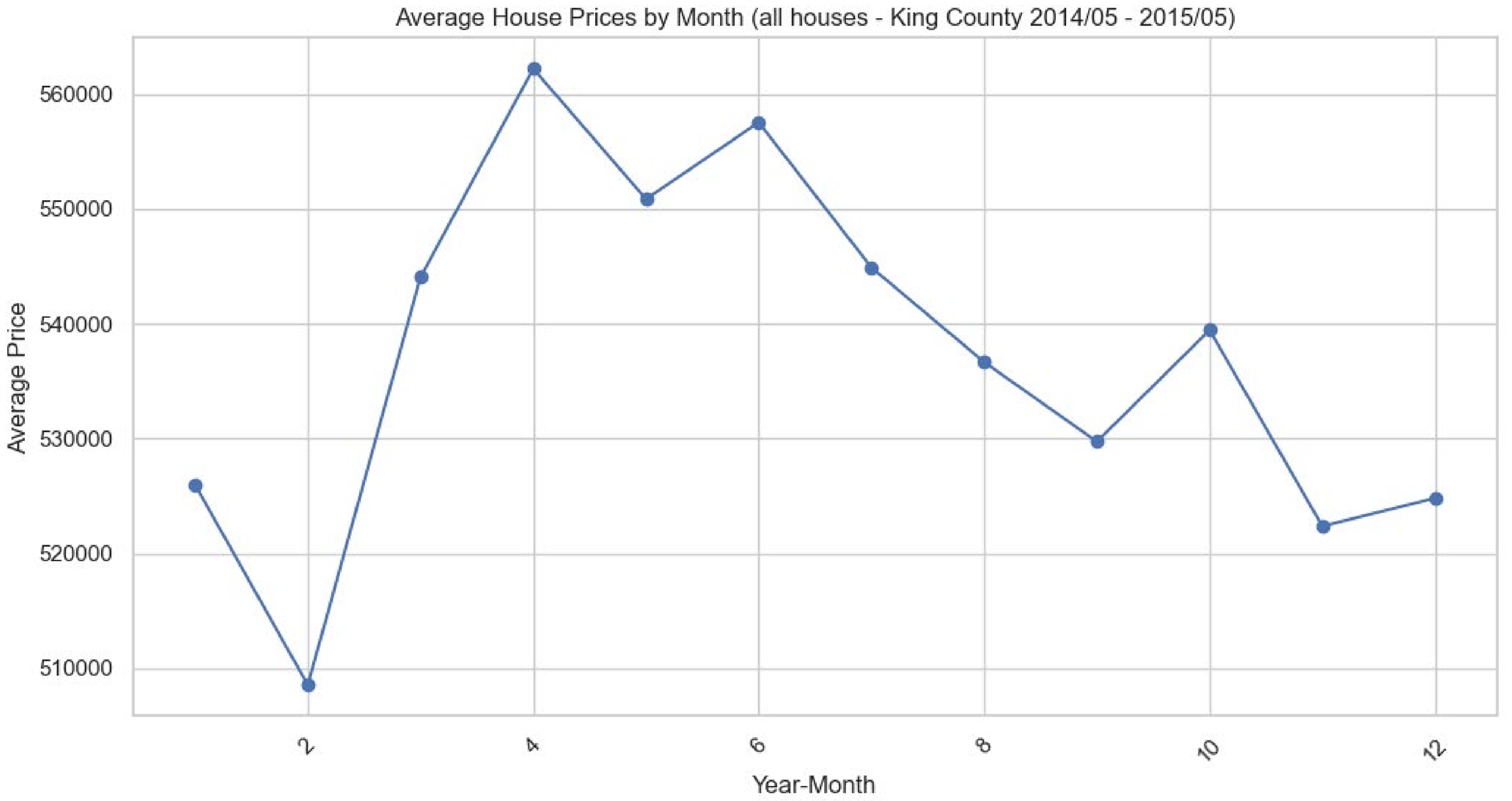


IV. ANALYSIS & VISUALISATION

- no long term data

- too many factors

- random distribution



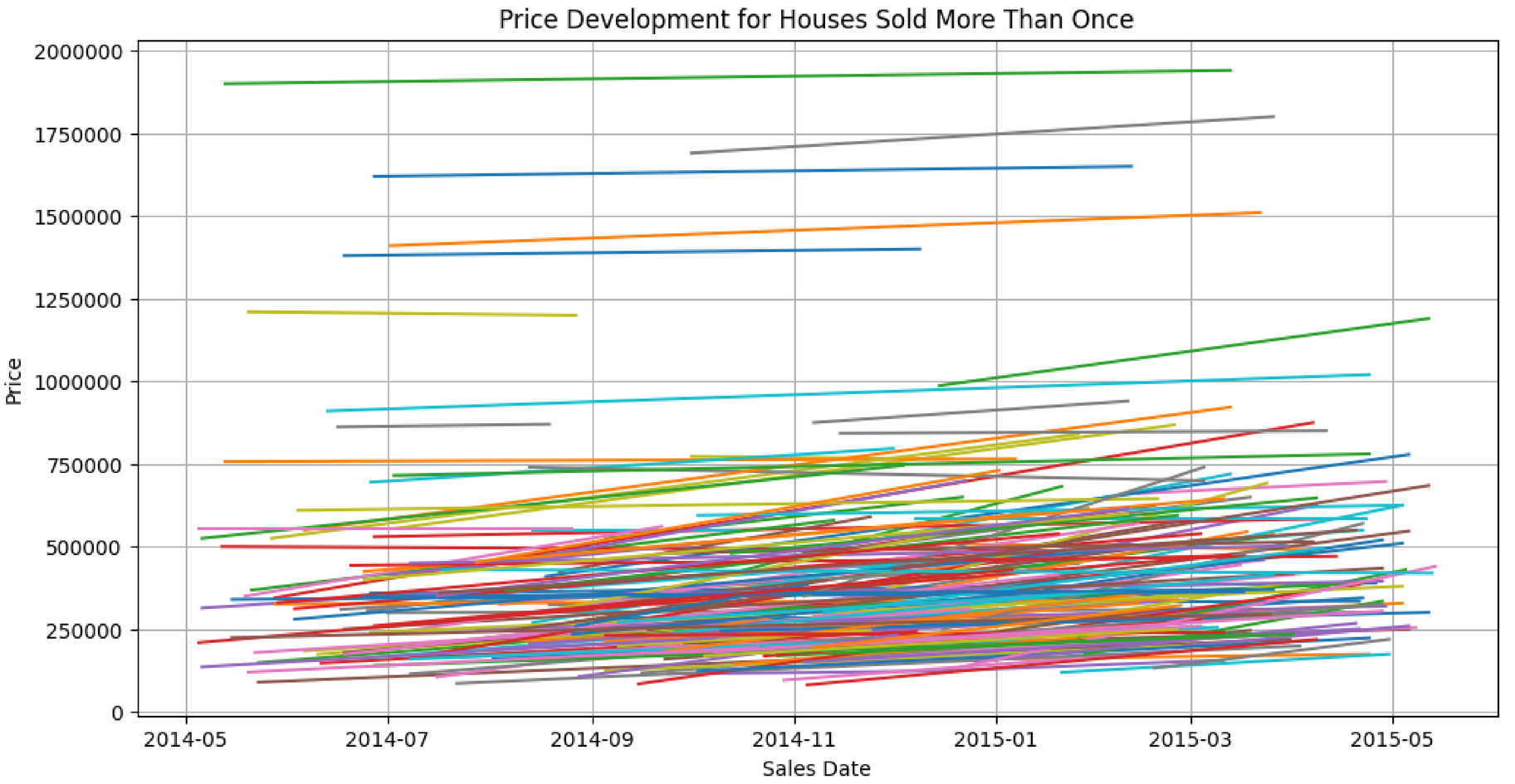
lineplots - average house prices by month of sale

IV. ANALYSIS & VISUALISATION

- as expected:
  upward trend

- price decrease:
  only 3.4%
  (6 of 176 houses)

= outliers/exceptions



lineplots - price development for houses sold more than once

IV. ANALYSIS & VISUALISATION

Conclusion:

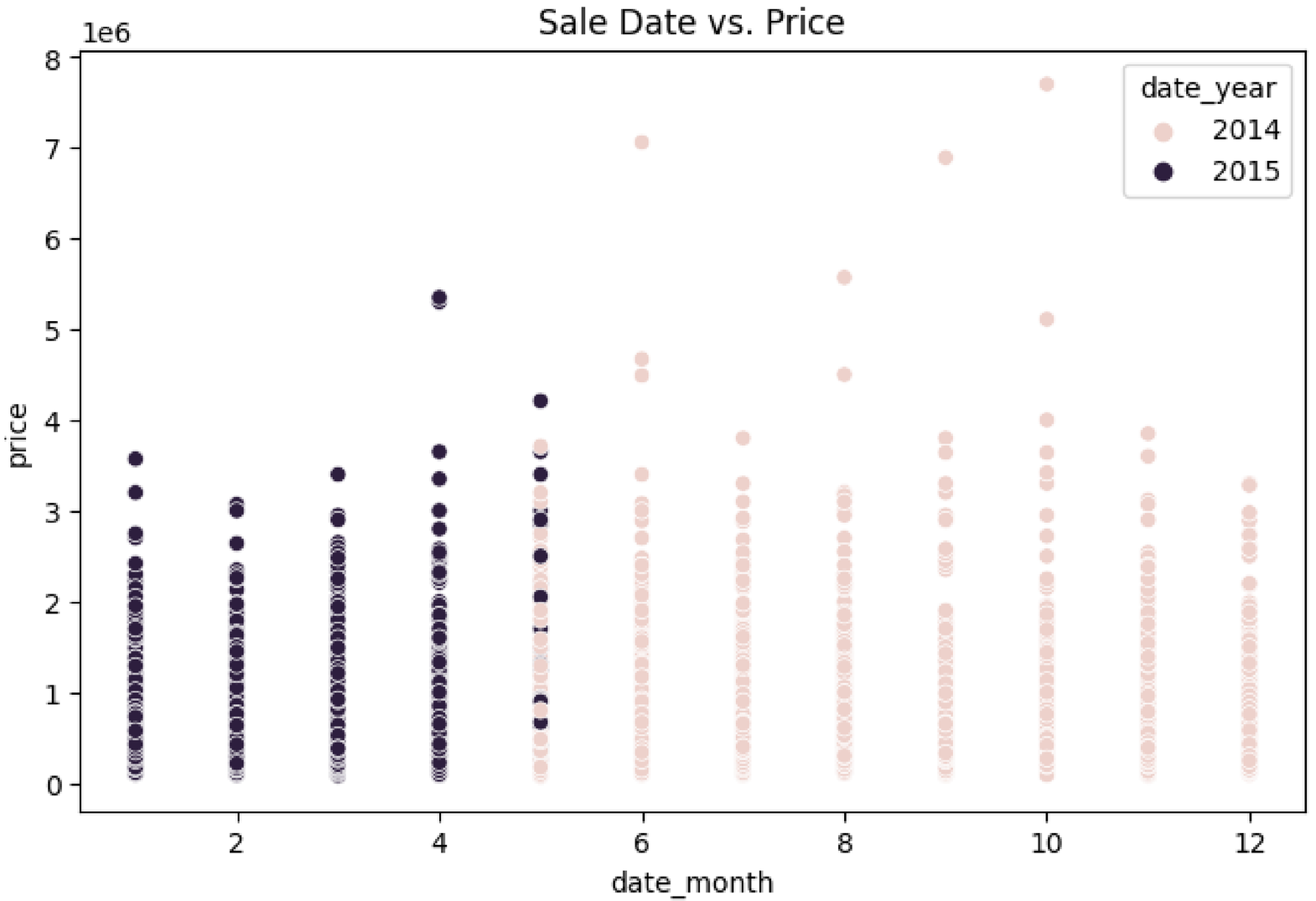
Analyzing the "best time" to buy a house based solely on the month is challenging for several reasons:

- real estate prices are influenced by a multitude of factors (market conditions, economic factors, location, individual property attributes, etc.)

- No long-term trends available. (Real estate trends often follow long-term cycles.)

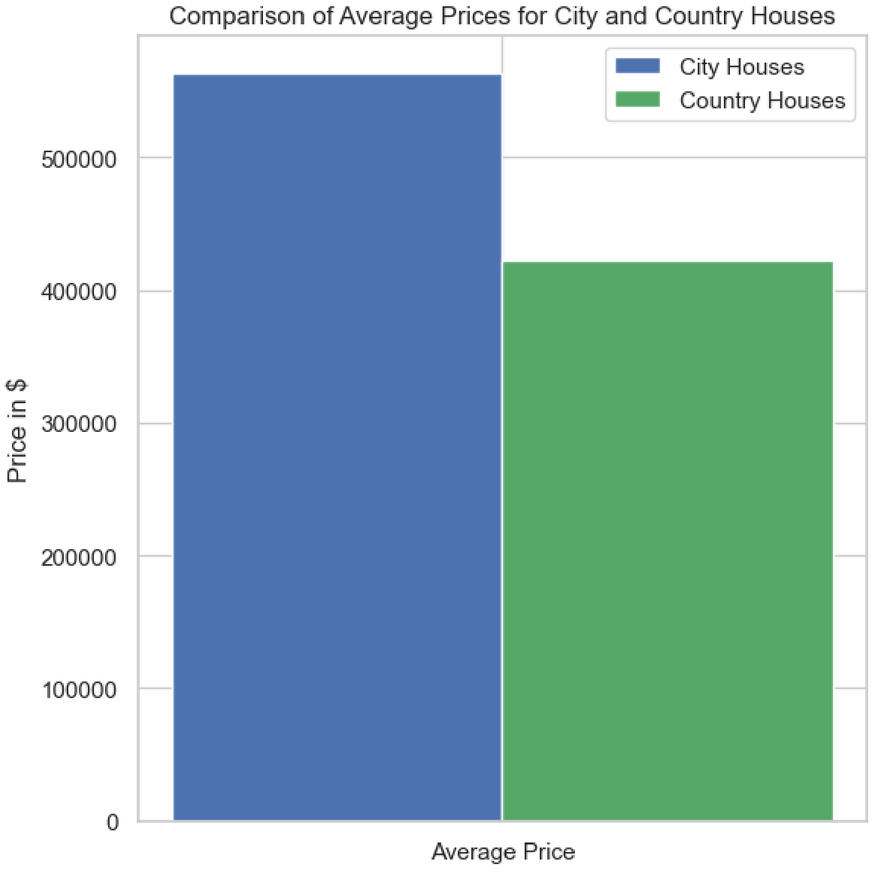
-> the best time to buy a house depends on individual circumstances and goals.

- prices tend to increase over time.



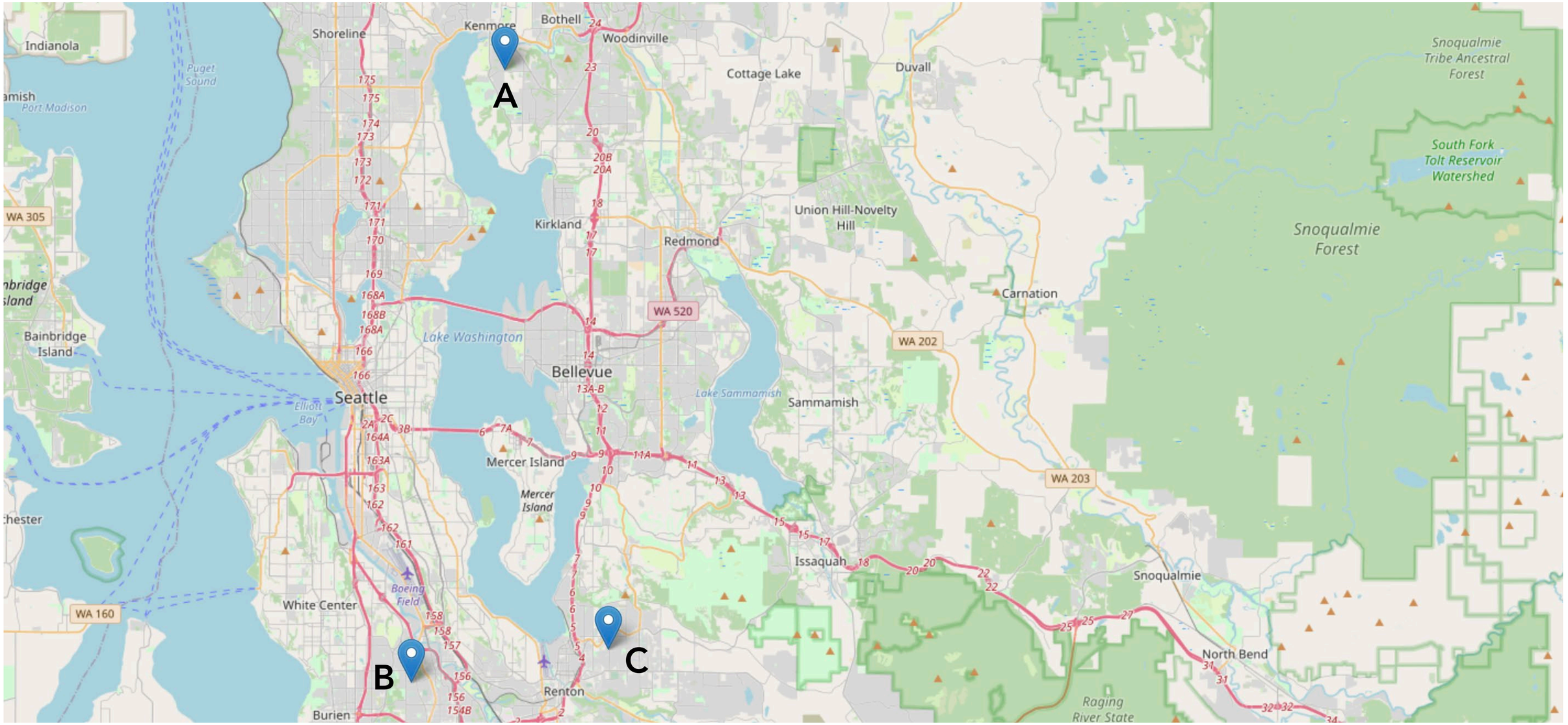
scatterplot - distribution and relation between month of sale and price

IV. ANALYSIS & VISUALISATION



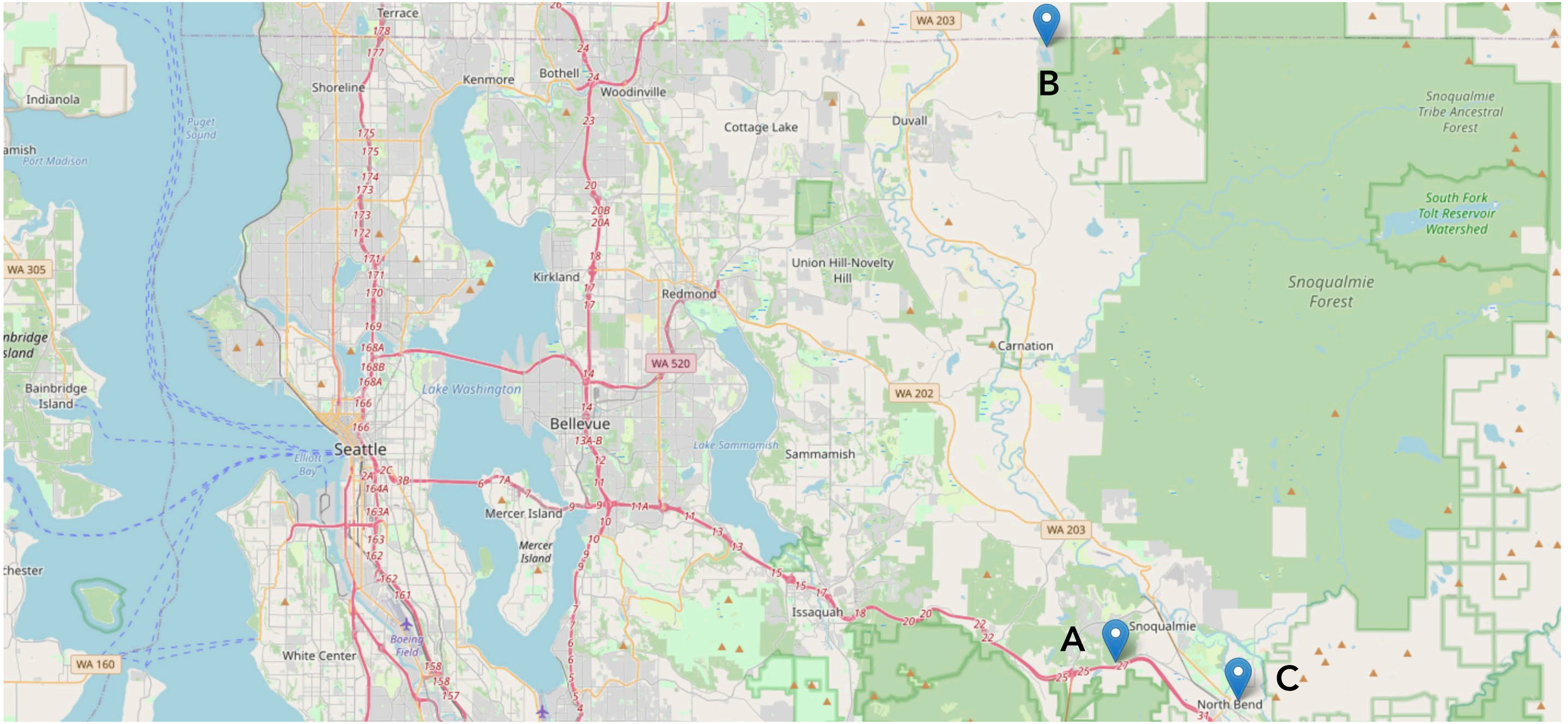
bar chart - comparison of average size and price between city and country houses

IV. ANALYSIS & VISUALISATION



map 01 - top three city house locations

IV. ANALYSIS & VISUALISATION



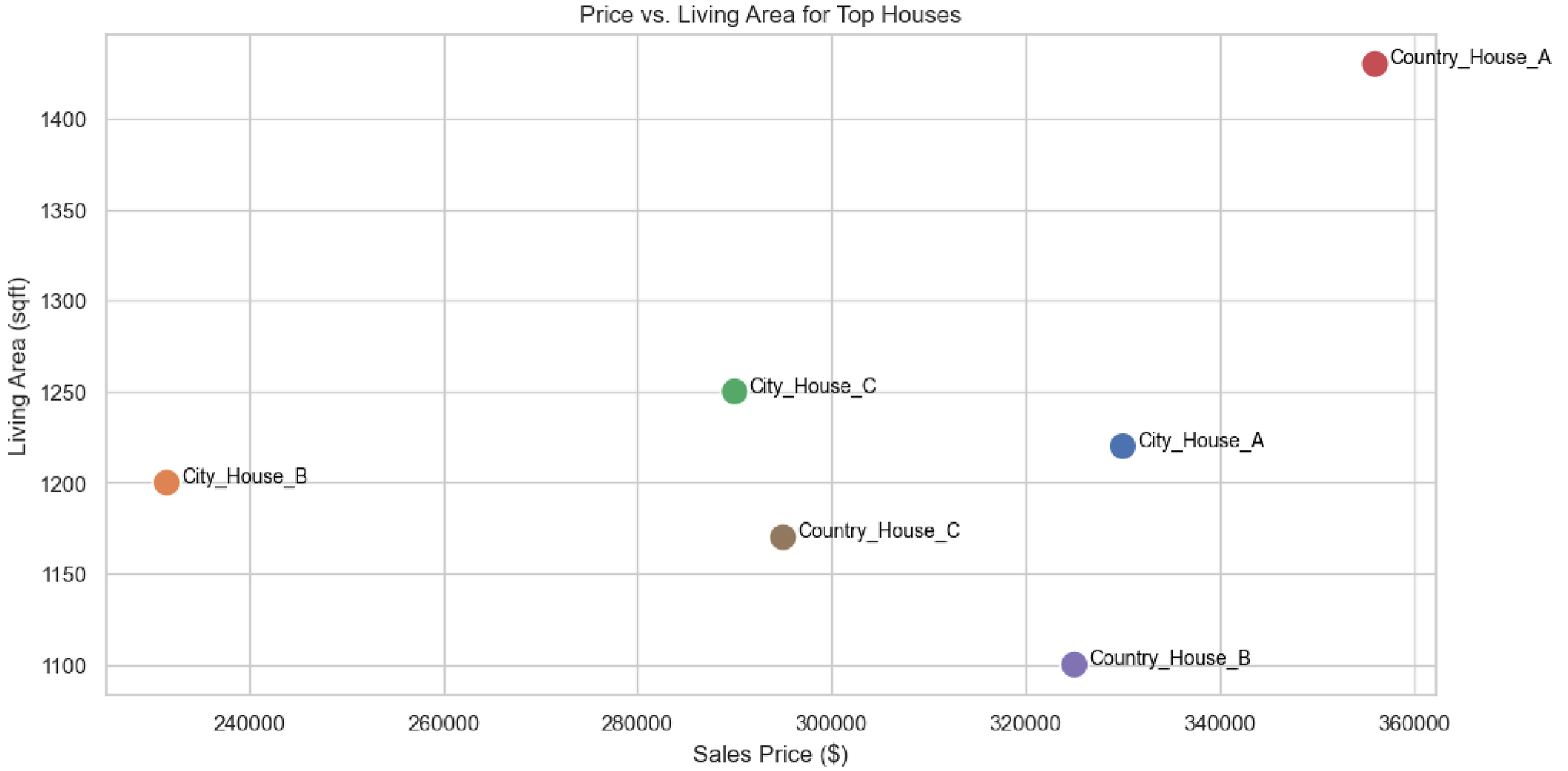
map 02 - top three country house locations

IV. ANALYSIS & VISUALISATION

	Top_3	price	bedrooms	bathrooms	sqft_living	sqft_lot	condition	grade	yr_built	zipcode
0	City_House_A	330000.000	2.000	1.000	1220.000	10000.000	5	7	1950	98028
1	City_House_B	231500.000	2.000	1.000	1200.000	9488.000	5	7	1941	98168
2	City_House_C	290000.000	2.000	1.000	1250.000	12507.000	5	7	1958	98056
3	Country_House_A	356000.000	2.000	1.000	1430.000	365904.000	3	7	1991	98065
4	Country_House_B	325000.000	2.000	1.000	1100.000	17817.000	3	7	1980	98019
5	Country_House_C	295000.000	2.000	1.000	1170.000	10621.000	3	7	1963	98045

data - top houses

IV. ANALYSIS & VISUALISATION



scatterplot - top houses | price vs. living area

IV. ANALYSIS & VISUALISATION

scatterplot - top houses | price vs. property area

IV. ANALYSIS & VISUALISATION

house images - country house a

“largest property!”

IV. ANALYSIS & VISUALISATION

house images - country house b

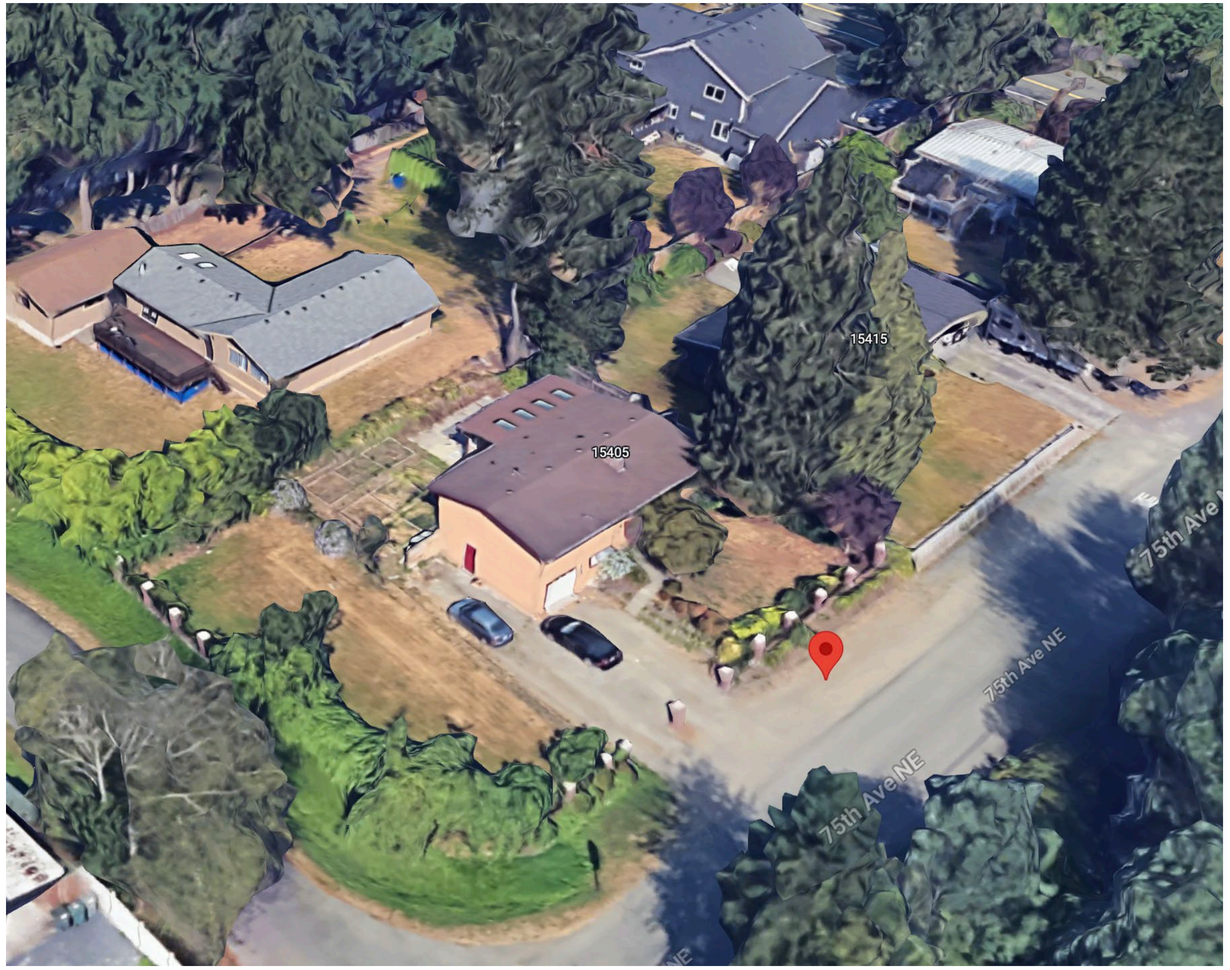
“It’s on a lake!”

IV. ANALYSIS & VISUALISATION

house images - country house c

“why not?”

IV. ANALYSIS & VISUALISATION



house images - city house a

“decent.”

IV. ANALYSIS & VISUALISATION

house images - city house b

"cheapest!"

IV. ANALYSIS & VISUALISATION

house images - city house c

“best size-to-cost ratio.”

V. CONCLUSIONS

A. Insights:

- Older houses tend to be smaller. Newer Houses tend to have better building grades.
- The better rated and the bigger a house, the more expensive it usually is.
- The price range is relatively evenly distributed throughout the zipcodes.
- Country houses tend to be bigger, city houses tend be more more expensive compared to each other.

B. Recommendations:

- Country House A has the biggest plot, but **Country House B** is on a lake...
- **City House C** has the best size-to-cost ratio and is also the newest of the Top 3 city houses.
- The best time to buy a house depends on your individual circumstances and goals. The sooner the better though, since prices tend to increase over time.