



Dieter Janzen

June 3, 2022

EDA Project King County Housing

Stakeholder

Jacob Phillips

- Buyer with unlimited budget

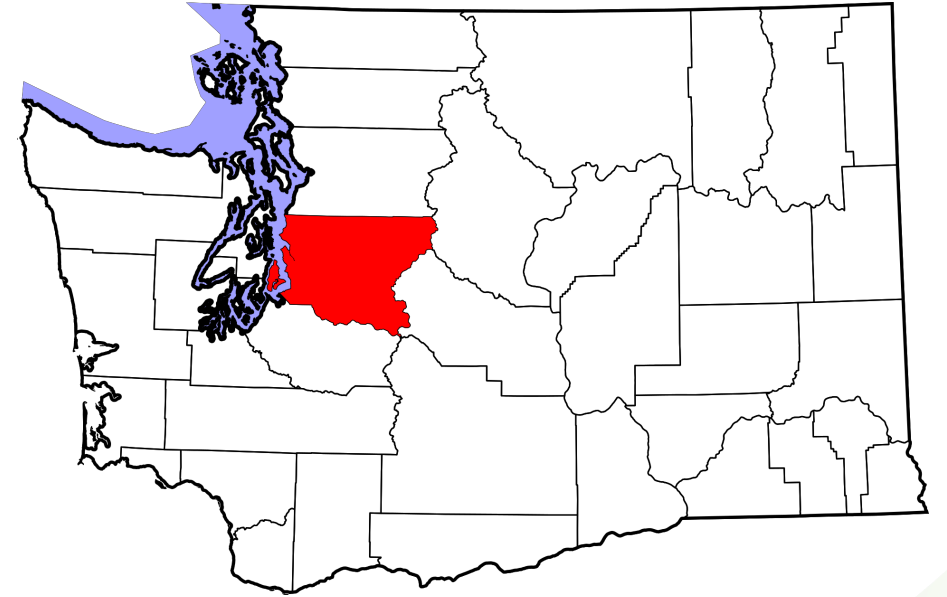
He wants a historic house with:

- 4+ bathrooms or a second smaller house nearby
- A big lot with room for tennis court and pool
- A golf course nearby
- No waterfront location

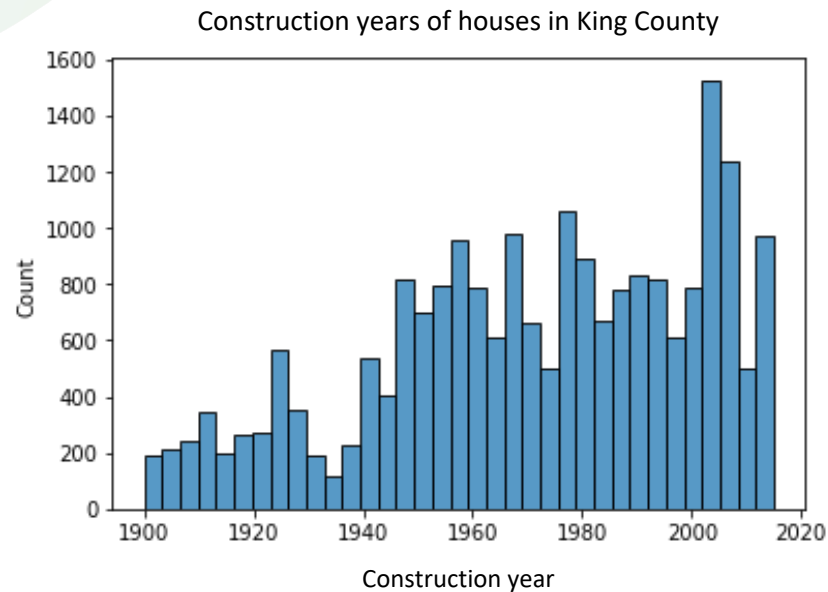
Dataset

The dataset contains the data of 21597 houses in King County, Washington, USA.

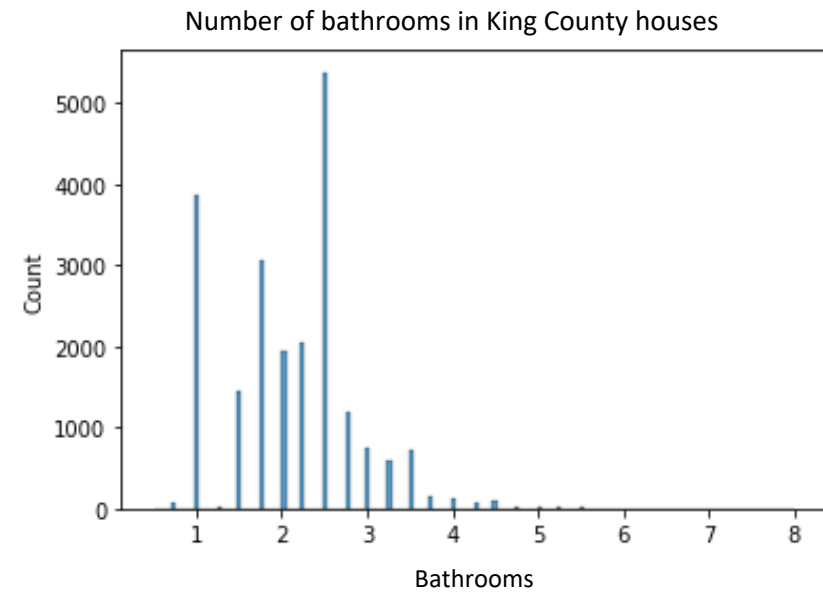
The Houses are described by 19 features.



Data overview

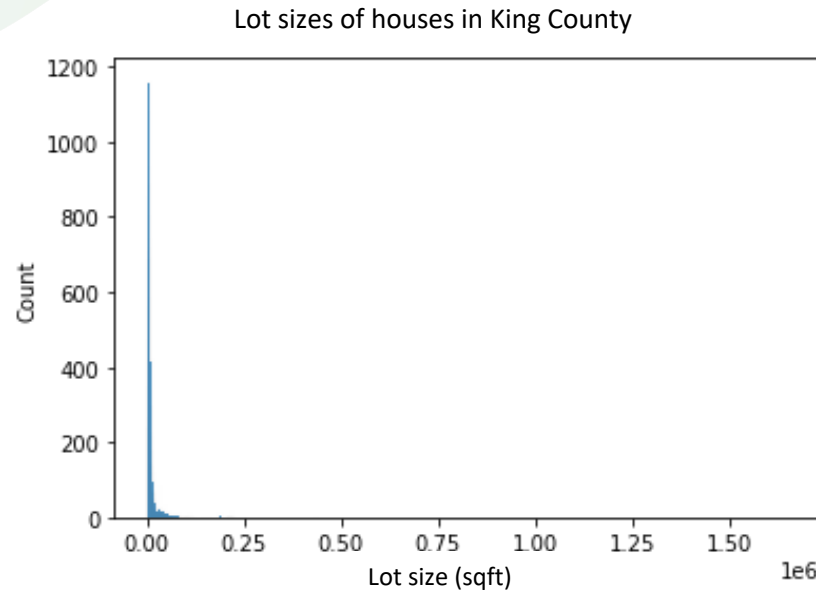


- More houses built in the last 50 years
- Still a lot of old houses available

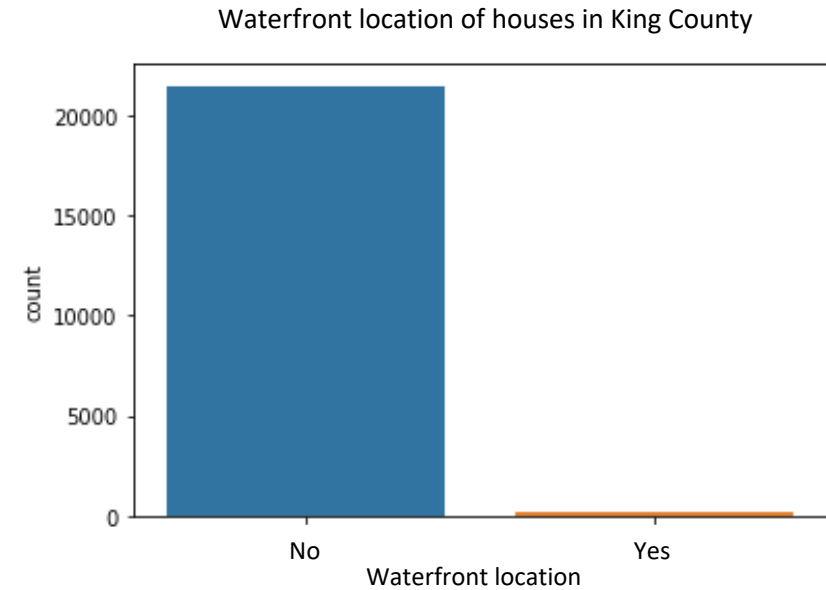


- Most houses have 2.5 or less bathrooms
- This will be a limiting factor.

Data overview



- There are some giant lots
- Not outliers, because stakeholder is interested in big lots



- Only few houses at a waterfront location



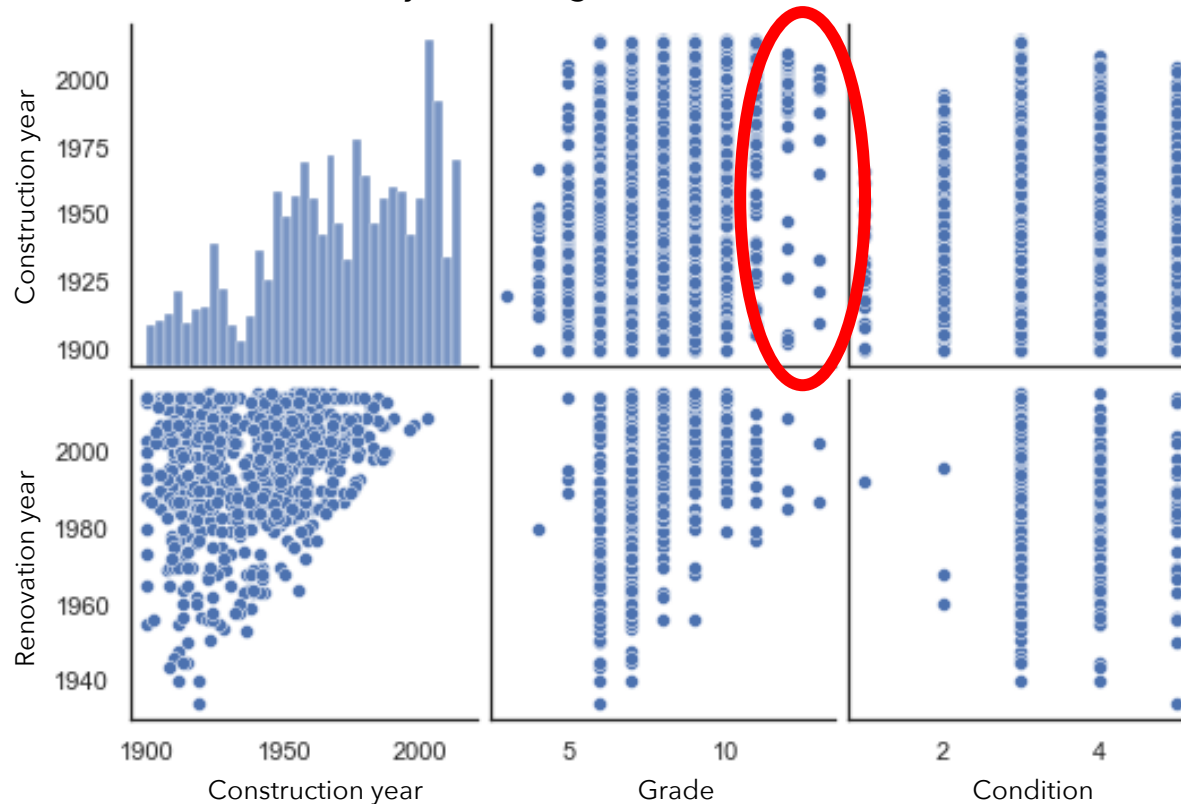
Questions

1. How does the age of the house correlate with renovation, condition and grade?
 - If a house is historic (> 50 years old), it either has low condition/grade, or was recently renovated and has a good condition/grade
2. How does a nearby golf course correlate with lot size and other features?
 - If a golf course is close to the property, lot size and also other features are higher compared to properties further away

Question 1

How does the age of the house correlate with renovation, condition and grade?

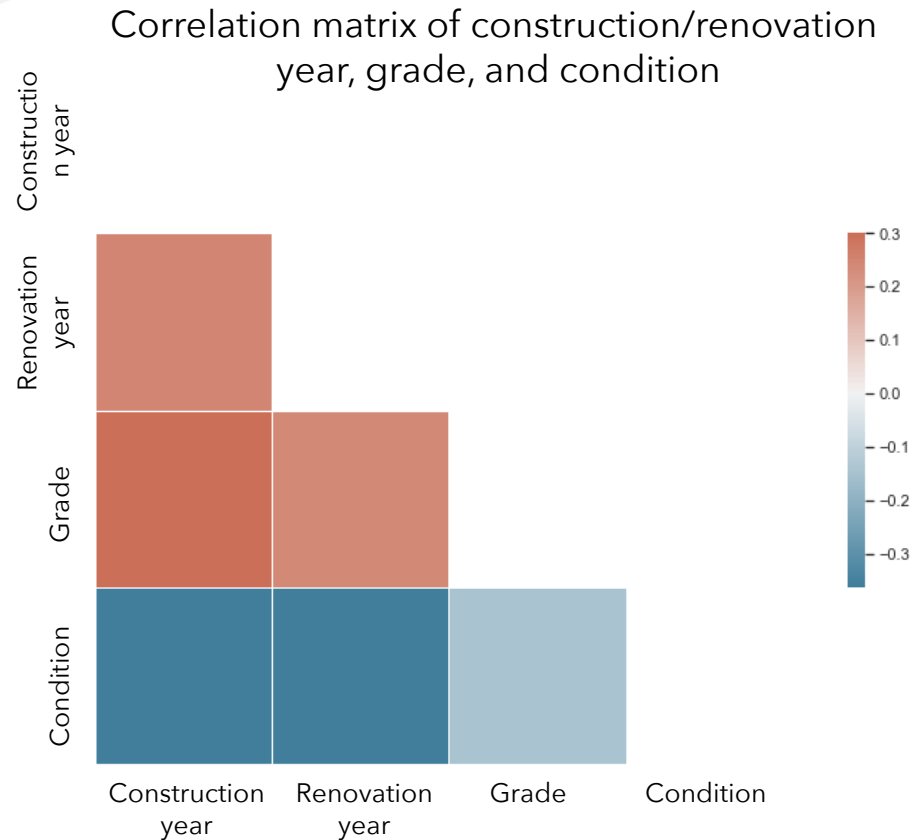
Relationship of construction/renovation year with grade and condition



- The older the house, the more likely it was renovated
- Highest grade houses tend to be either new or very old
- Low grade houses with bad condition were mostly built before 1960
- Renovated houses are in better condition
- High grades only for houses renovated in the last 40 years

Question 1

How does the age of the house correlate with renovation, condition and grade?



- Negative correlation of condition with other features?

Houses (construction < 1972)	Grade (mean)	Condition (mean)
Renovated	7.68314	3.225291
Not renovated	7.060890	3.665311

- Non-renovated old houses seem to have lower grade, but higher condition.



Question 1

How does the age of the house correlate with renovation, condition and grade?

Hypothesis:

If a house is historic (> 50 years old), it either has low condition/grade, or was recently renovated and has a good condition/grade.

This hypothesis seems to be false, because:

- Renovated historic houses on average seem to have actually worse condition compared to non-renovated houses
- However, the average grade seems to be higher for renovated historic houses.

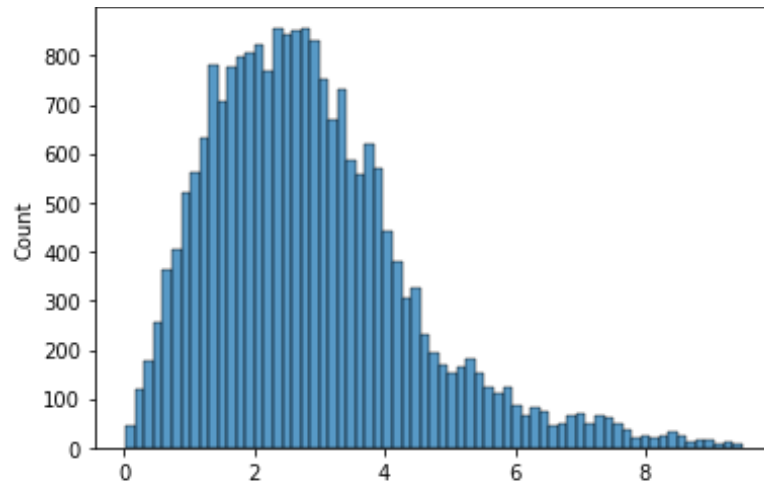
Takeaway: it doesn't matter if a historic house is renovated or not.

Question 2

How does a nearby golf course correlate with lot size and other features?

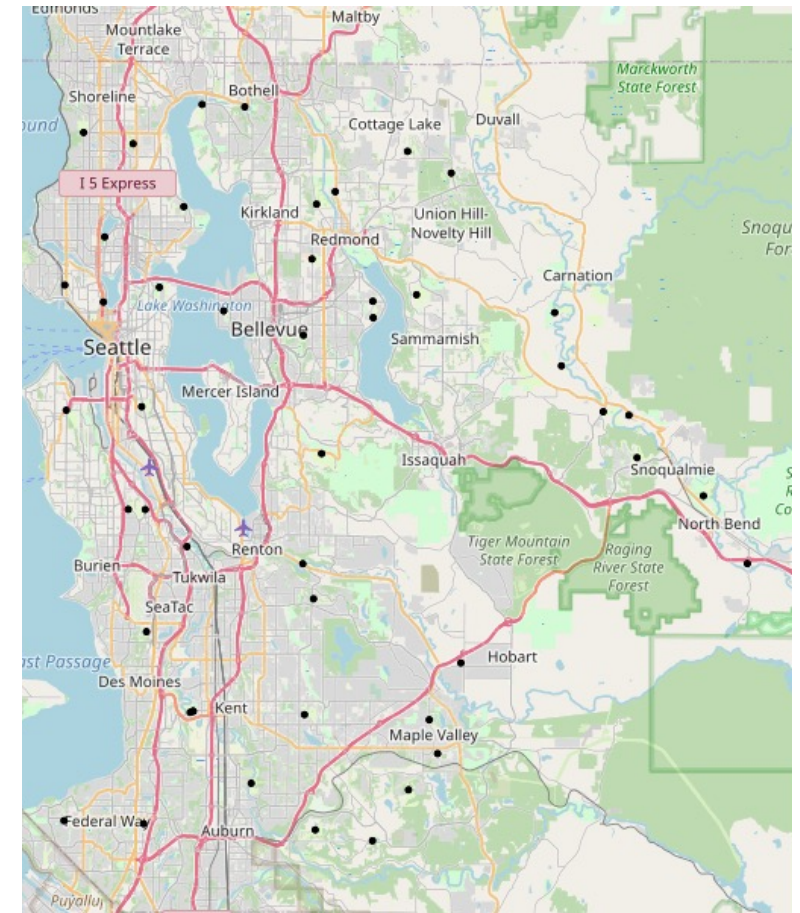
- Gathered the locations of golf courses in King County and looked at the nearest distance for each house

Distance to nearest golf course of houses in King County



- Most houses seem to be within 4 km of a golf course

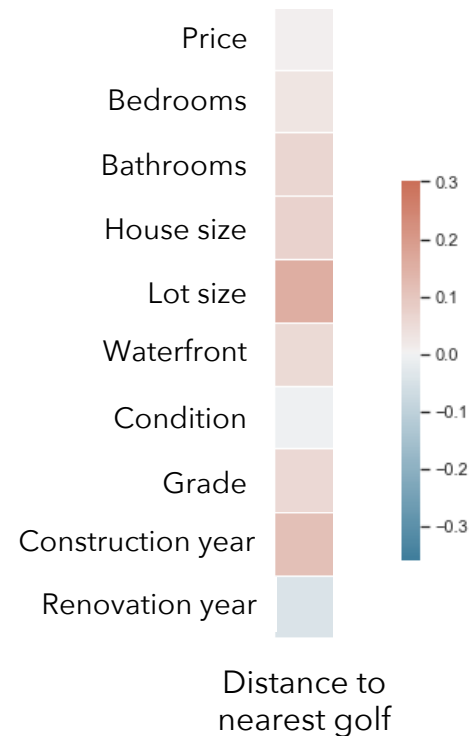
Golf course locations in King County



Question 2

How does a nearby golf course correlate with lot size and other features?

Correlation matrix of distance to golf course with other features

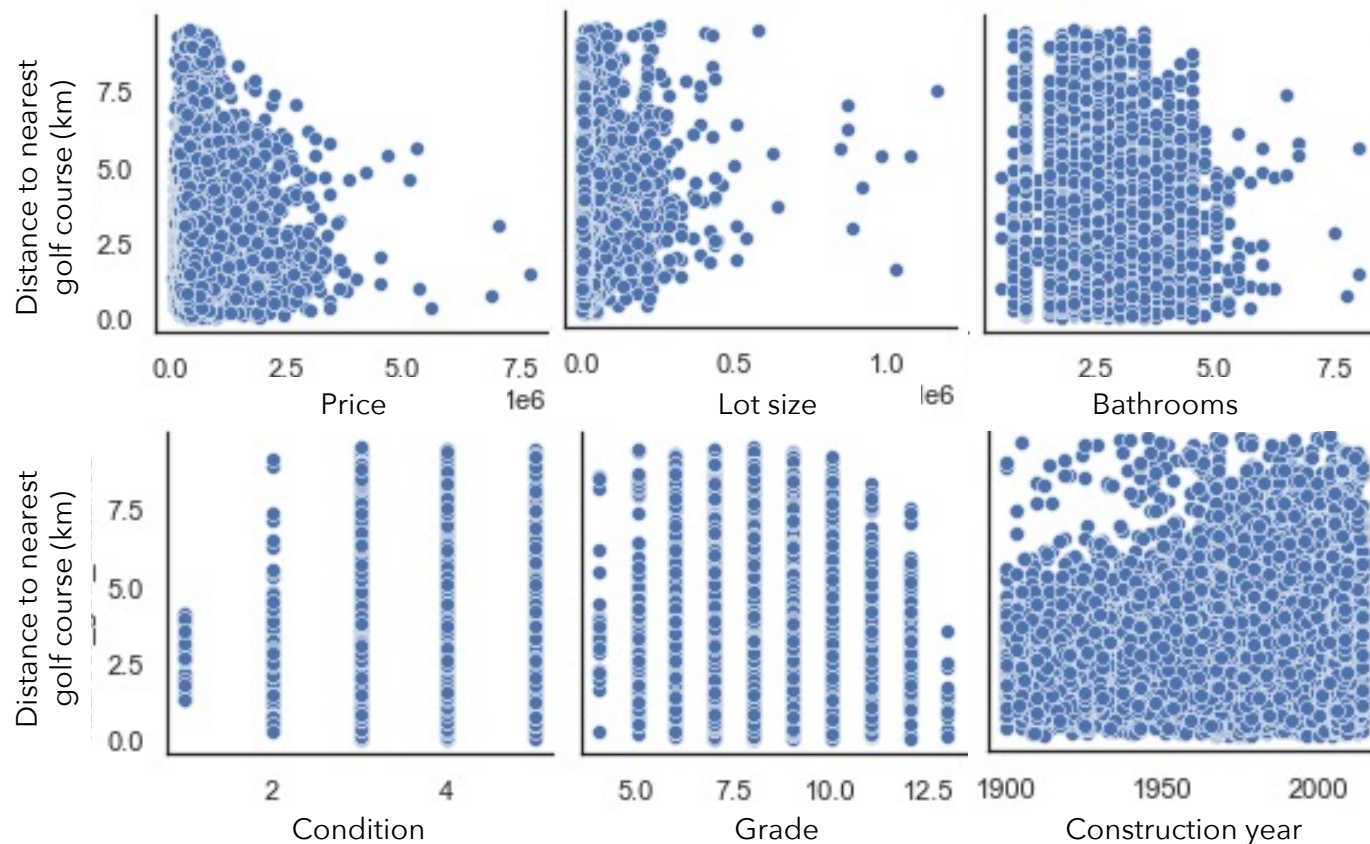


- There seems to be slight positive correlation with most values
- Except:
 - Price
 - Condition
 - Renovation year

Question 2

How does a nearby golf course correlate with lot size and other features?

Relationship of golf course distance with other features



- Highest grade and very expensive houses seem to be close to golf courses
- Houses with bigger lots are not very close to golf courses
- Very old houses and houses with very bad condition are located within 5 km of a golf course



Question 2

How does a nearby golf course correlate with lot size and other features?

Hypothesis:

If a golf course is close to the property, lot size and also other features are higher compared to properties further away.

This hypothesis seems to be true for:

- Price, grade, and bathrooms
- However, it doesn't seem to be true for house condition, as all bad condition houses are within 5 km of a golf course.

Takeaway: there is likely a suitable historic house close to a golf course

Finding the right house

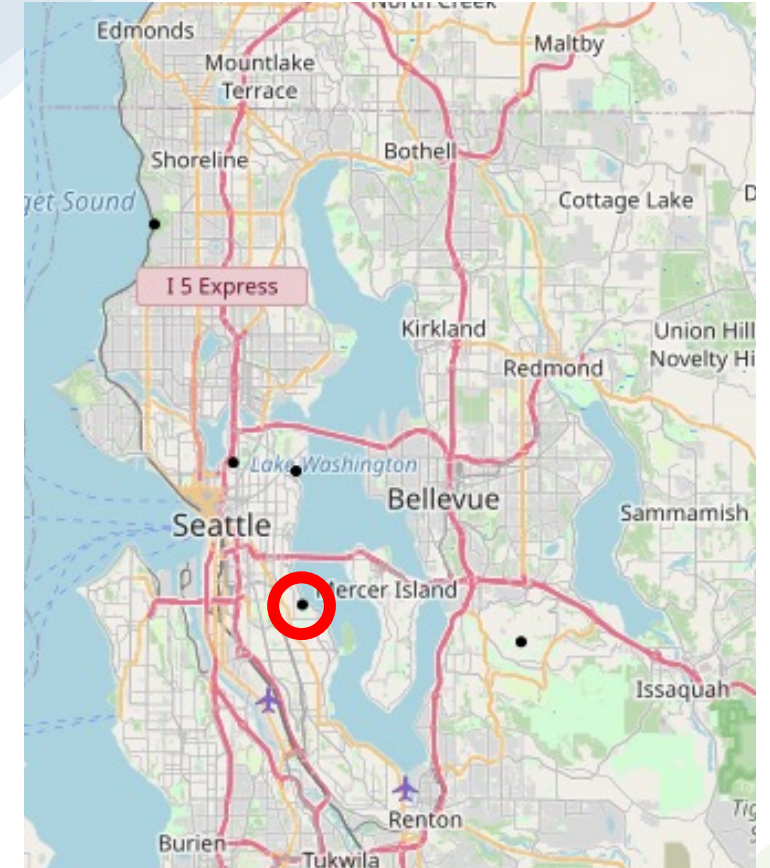
Based on the EDA, I tried to find houses meeting the following conditions:

- older than 50 years
- with 4+ bathrooms
- with a lot bigger than 20000 sqft
- with a golf course within 2.5 km
- with no waterfront
- good grade (9) and condition (3)

Five houses meet these requirements.

However, one seems to be too close to the waterfront.

Suitable houses



Recommendations for the stakeholder

The stakeholder should look at the following houses:

- 3377900195
- 6762700020
- 2303900035
- 5317100750

If these houses are unsuitable, houses with < 4 bathrooms, combined with a nearby smaller house, can also be taken into consideration.