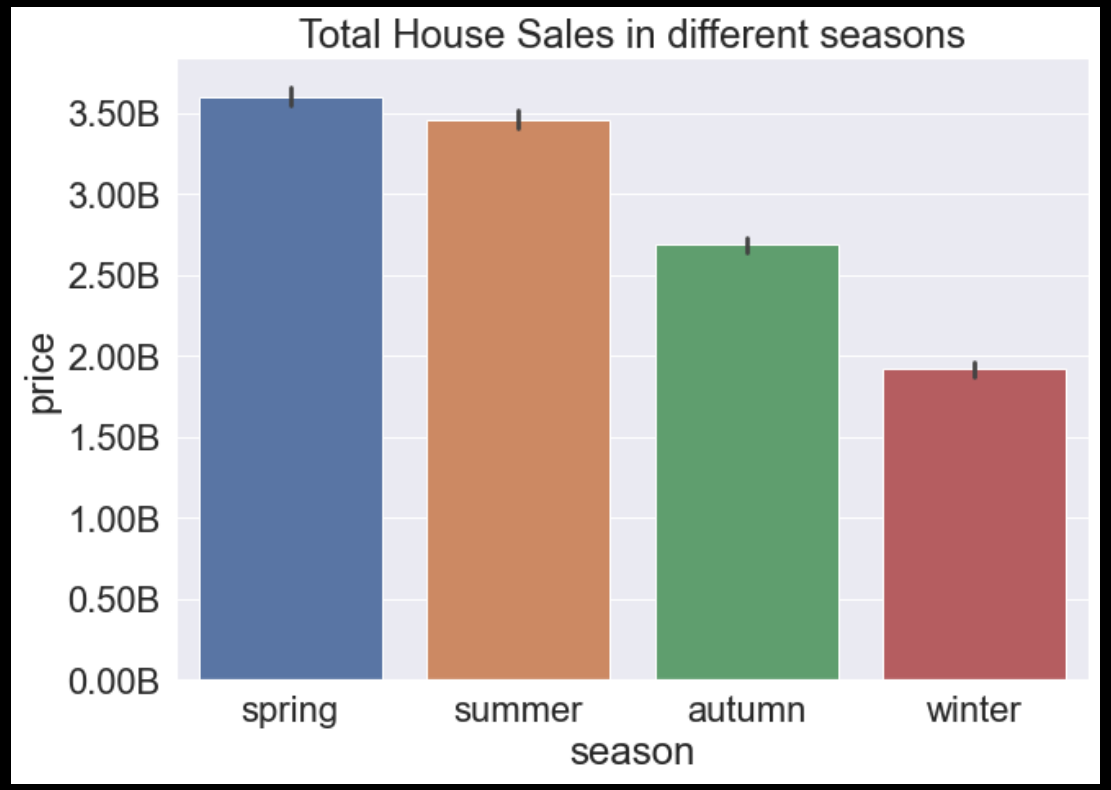
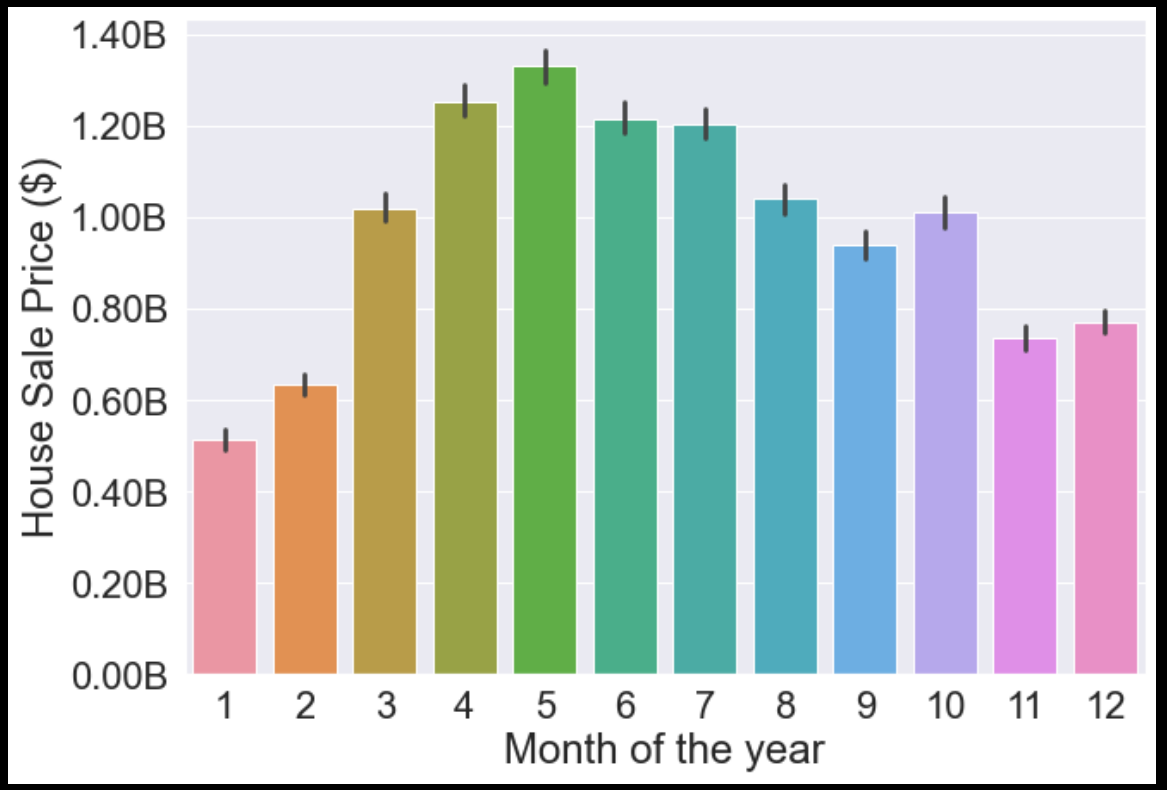
How to hack house prices

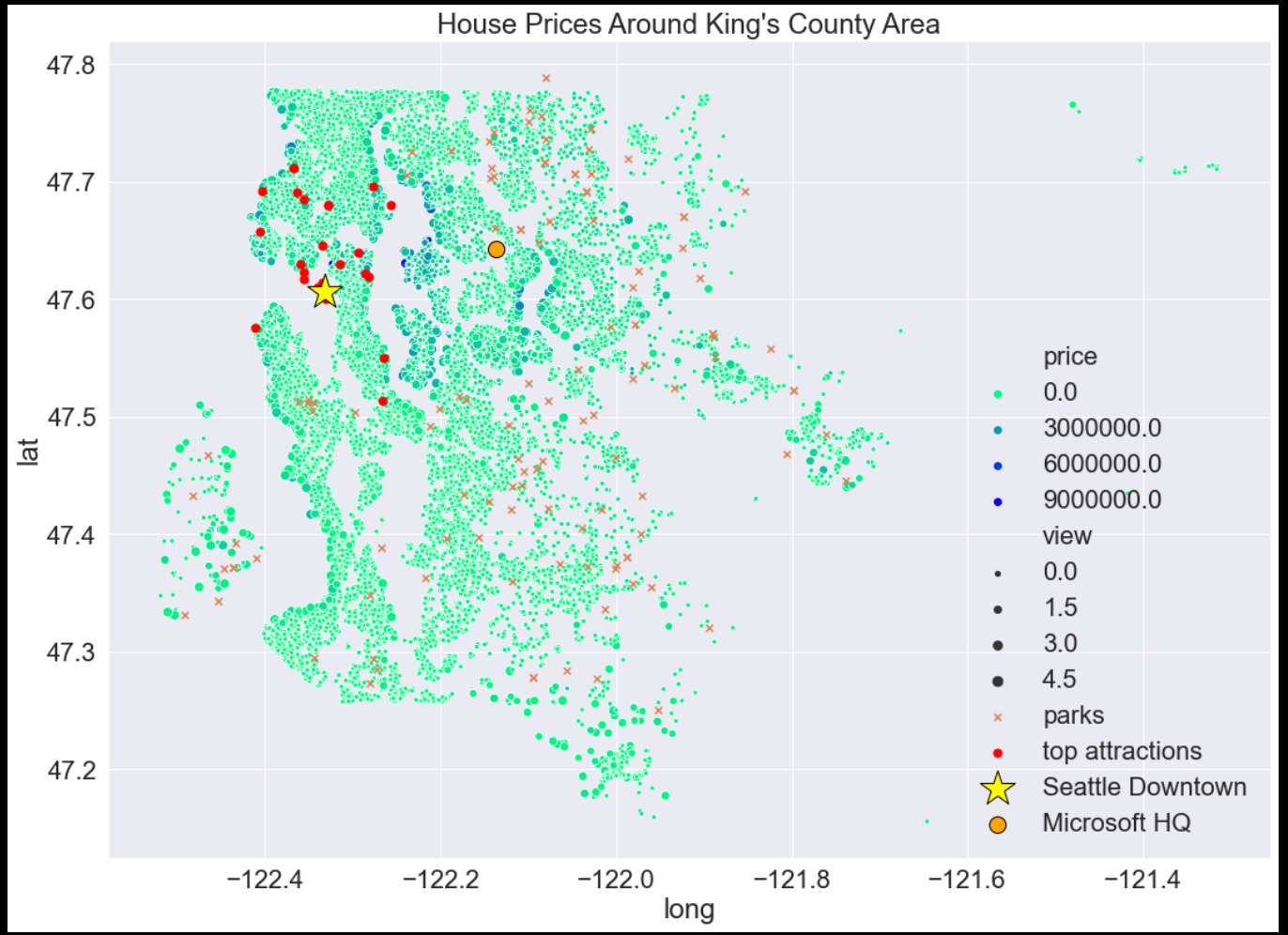
Sung Bae

Introduction





- House sales market is huge – around billions in each season
- There are different factors that could affect house price
- Example. Time/seasonal factor



- Geographical factor

Higher Prices
- Usually clustered
- Near water area

Lower Prices
- Spread out
- Away from the main city

Question 1

What factors affect the house prices the most in most interpretable and reasonable way?

Model Goals

1. Prioritize interpretability
2. Have reasonable predictors that can be altered by the owners
3. May not result in highest r^2 value

Model Result – Positive Effectors

Major Positive Effectors:

1. waterfront = water scenery matters
2. average room space = more space is better
3. grade x bathroom = having great bathrooms count!
4. Seattle = Further you live from Seattle is better

Minor Positive Effectors:

1. Effective age
2. Selling in Spring
3. Selling in Summer

		coef	std err	t	P>|t|	[0.025	0.975]
	Intercept	10.2369	0.053	194.138	0.000	10.134	10.340
Predictors	living_div_total_rooms	0.5292	0.009	61.506	0.000	0.512	0.546
	bathrooms_x_grade	0.4683	0.007	67.582	0.000	0.455	0.482
	eff_age	0.0402	0.003	14.284	0.000	0.035	0.046
	rate	-0.0239	0.003	-7.661	0.000	-0.030	-0.018
	mean_attractions	-0.5824	0.011	-53.090	0.000	-0.604	-0.561
	microsoft	-0.1694	0.003	-59.956	0.000	-0.175	-0.164
	seattle	0.2056	0.007	31.496	0.000	0.193	0.218
	season_spring	0.0302	0.007	4.627	0.000	0.017	0.043
	season_summer	0.0219	0.006	3.864	0.000	0.011	0.033
	season_winter	-0.0344	0.008	-4.247	0.000	-0.050	-0.019
	waterfront_10	0.7012	0.033	21.164	0.000	0.636	0.766

Model Result – Negative Effectors

Major Negative Effectors:

1. mean distance from main attractions = further you live, the value goes down generally
2. distance from Microsoft HQ = further you live, the value goes down generally

Minor Negative Effectors:

1. Mortgage rate – higher rate lowers the house value
2. Selling in winter

		coef	std err	t	P>|t|	[0.025	0.975]
	Intercept	10.2369	0.053	194.138	0.000	10.134	10.340
Predictors	living_div_total_rooms	0.5292	0.009	61.506	0.000	0.512	0.546
	bathrooms_x_grade	0.4683	0.007	67.582	0.000	0.455	0.482
	eff_age	0.0402	0.003	14.284	0.000	0.035	0.046
	rate	-0.0239	0.003	-7.661	0.000	-0.030	-0.018
	mean_attractions	-0.5824	0.011	-53.090	0.000	-0.604	-0.561
	microsoft	-0.1694	0.003	-59.956	0.000	-0.175	-0.164
	seattle	0.2056	0.007	31.496	0.000	0.193	0.218
	season_spring	0.0302	0.007	4.627	0.000	0.017	0.043
	season_summer	0.0219	0.006	3.864	0.000	0.011	0.033
	season_winter	-0.0344	0.008	-4.247	0.000	-0.050	-0.019
	waterfront_10	0.7012	0.033	21.164	0.000	0.636	0.766

Model Result - Accuracy

Dep. Variable:	price	**R-squared:**	0.725
Model:	OLS	**Adj. R-squared:**	0.725
Method:	Least Squares	**F-statistic:**	3694.
Date:	Mon, 21 Sep 2020	**Prob (F-statistic):**	0.00
Time:	11:45:16	**Log-Likelihood:**	-418.50
No. Observations:	15409	**AIC:**	861.0
Df Residuals:	15397	**BIC:**	952.7
Df Model:	11		
Covariance Type:	nonrobust		

R-squared: 0.725

- 72.5 % of the price can be described by adding in the factors
- This model can be used to intuitively understand how house prices are effected by different factors

Question 2

So what can be done to maximize house price by the homeowners?

Actions that can be done:

Increase Living Area	*grade x bathrooms*	*Make some waterfront*
- second strongest factor - Shows that people love having bigger independent open spaces	- Shows how important it is to have great quality bathrooms - Renovating bathrooms can be key to increasing your house value!	- The strongest factor - Shows people love nature - construct or install some sort of water fountains or pond that can be seen from the house

Actions that can be done:

Conclusion

- Model is limited in accuracy (R^2 = 0.725)
- However it can give us a good idea of what factors we can control and change to change the house price
  - Actionable changes
    - Renovate to
      - Increase open space
      - Waterfront presence
      - Better quality bathrooms
  - What NOT to do
    - increase number of bathrooms and bedrooms in total
    - sell in the winter season
    - sell during high mortgage rate

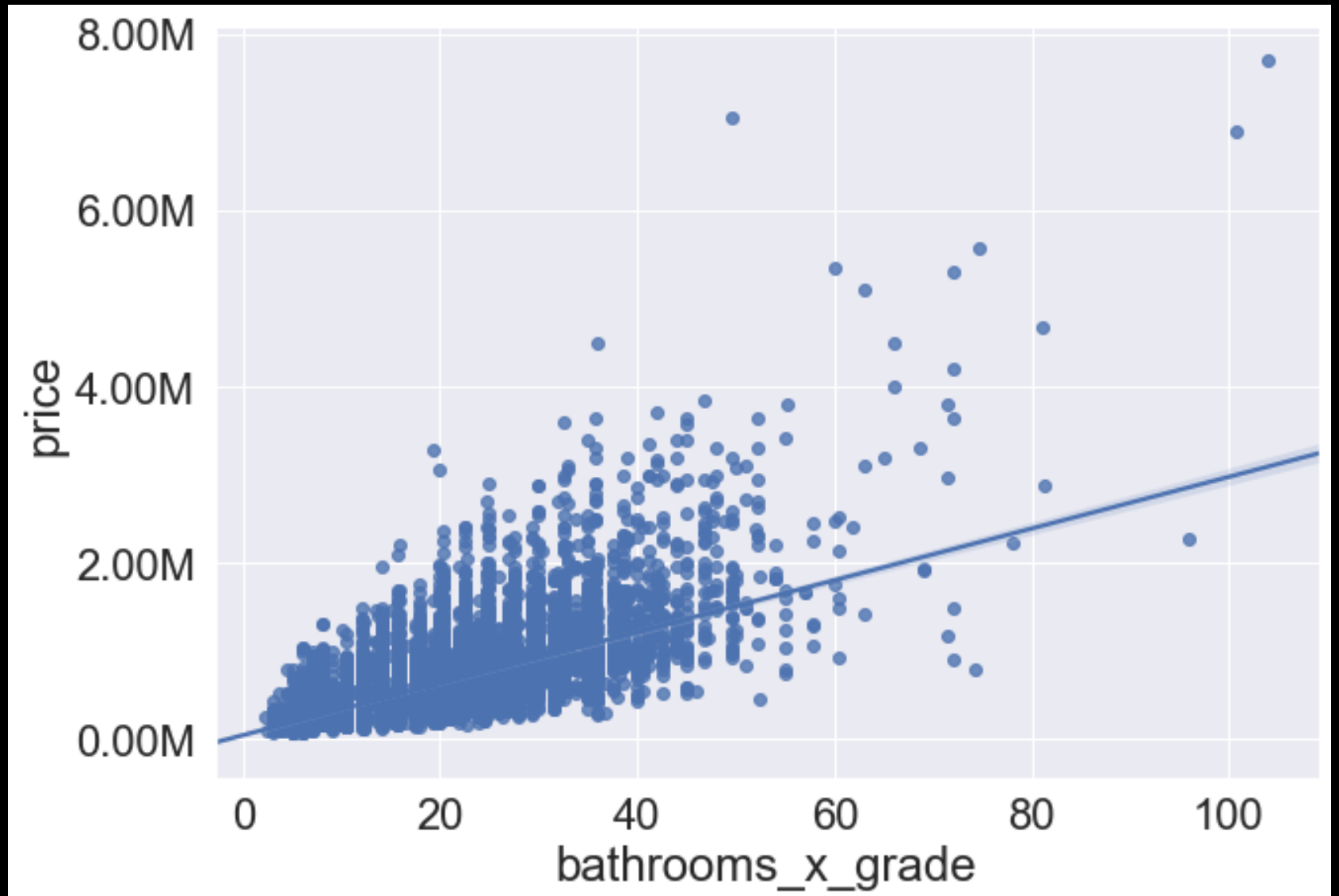
Future Studies

- The following features could be added to increase the accuracy of the model:
  - school district for elementary, middle and high school rating
  - average traffic time to major attractions
  - noise level
  - flood factors

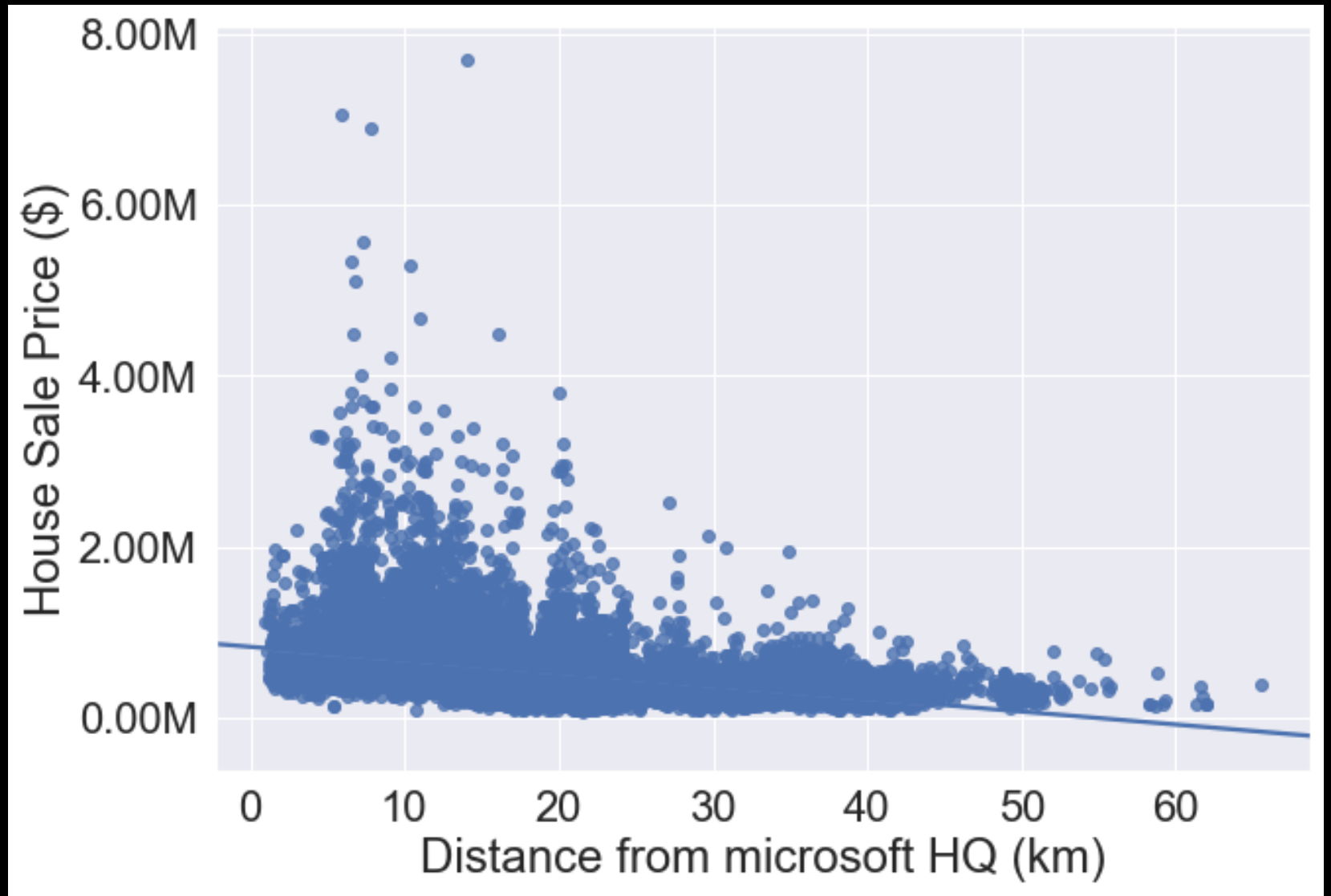
Thank you for listening

Appendix

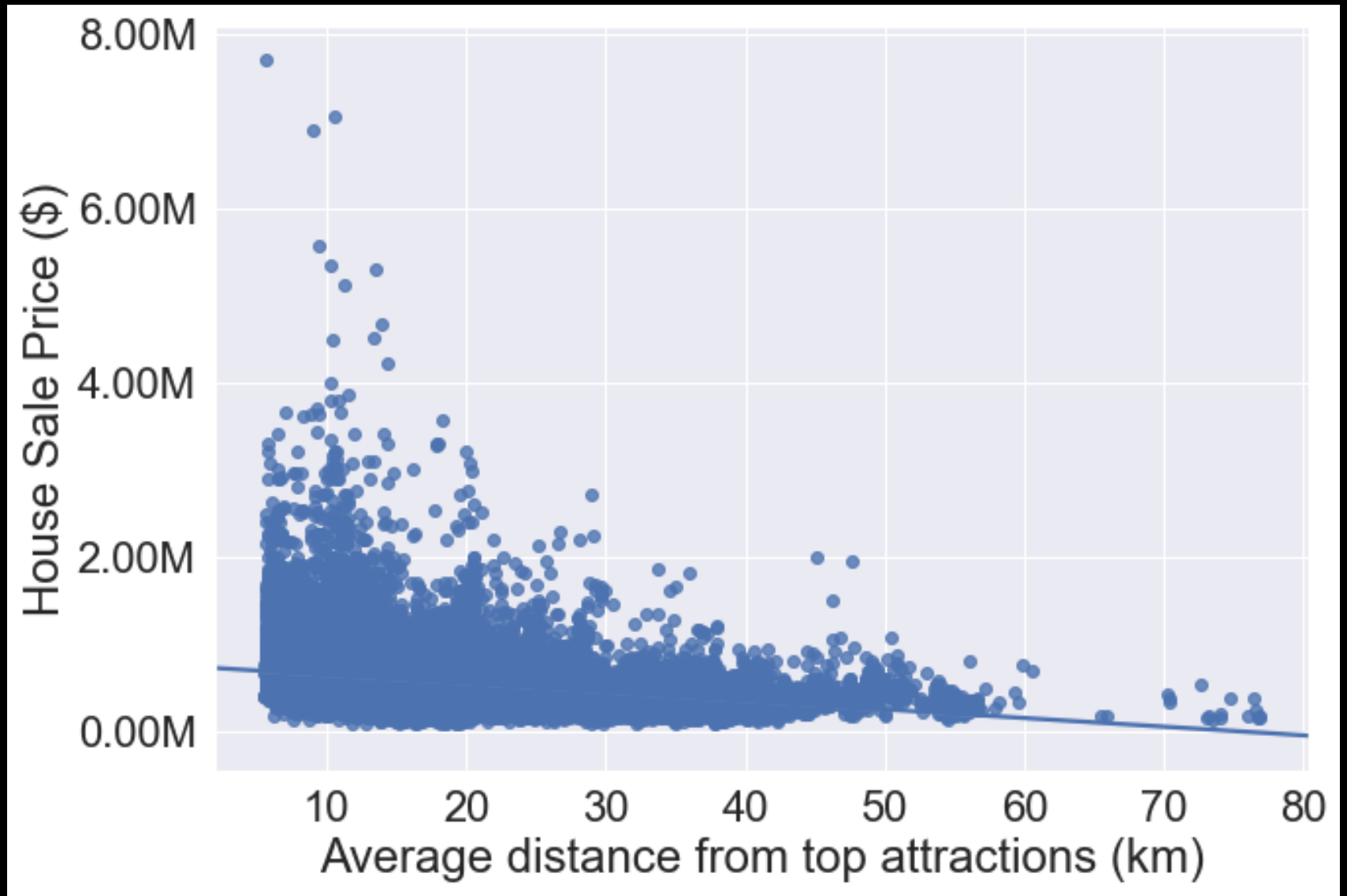
Price vs. Bathroom x grade

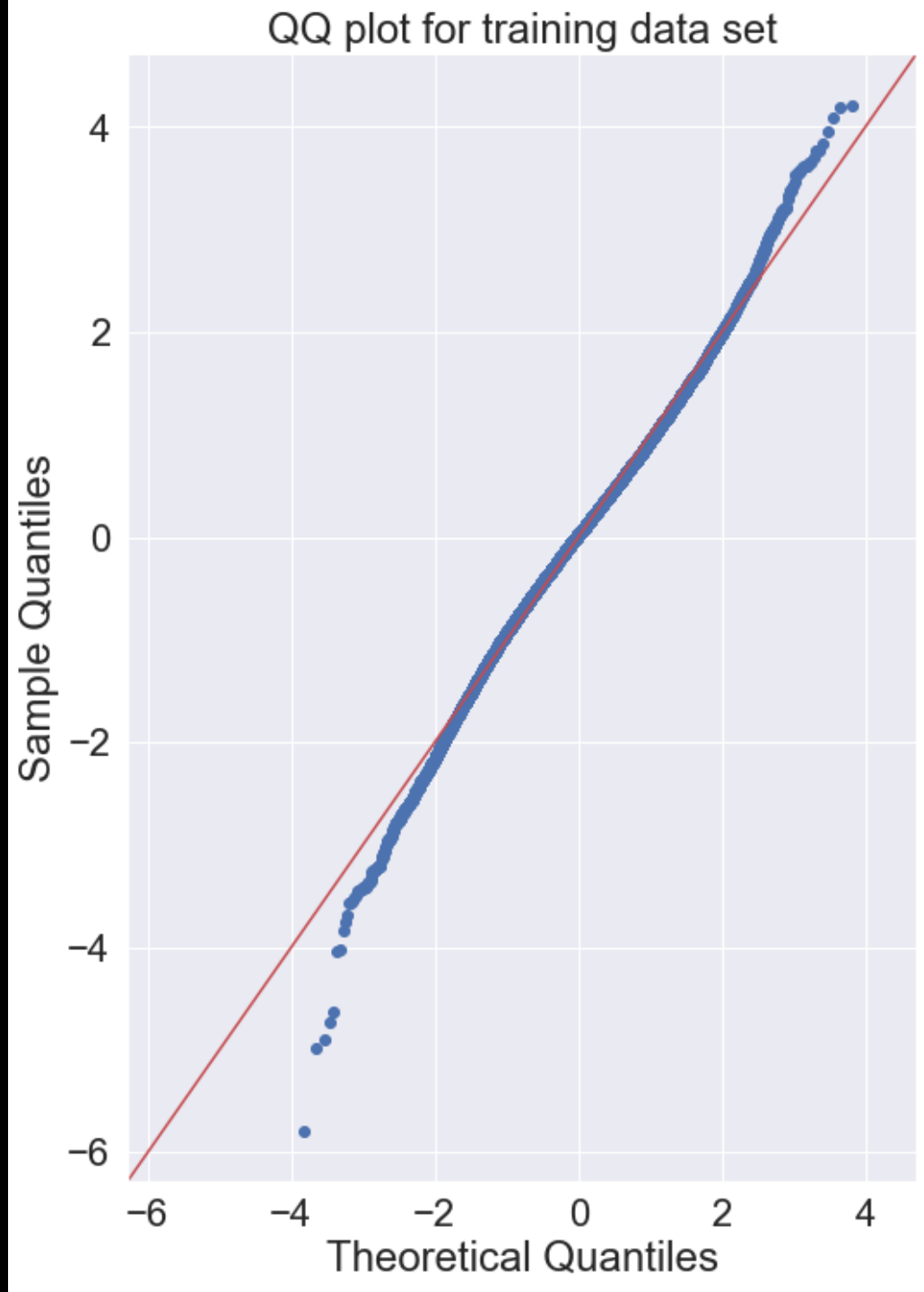


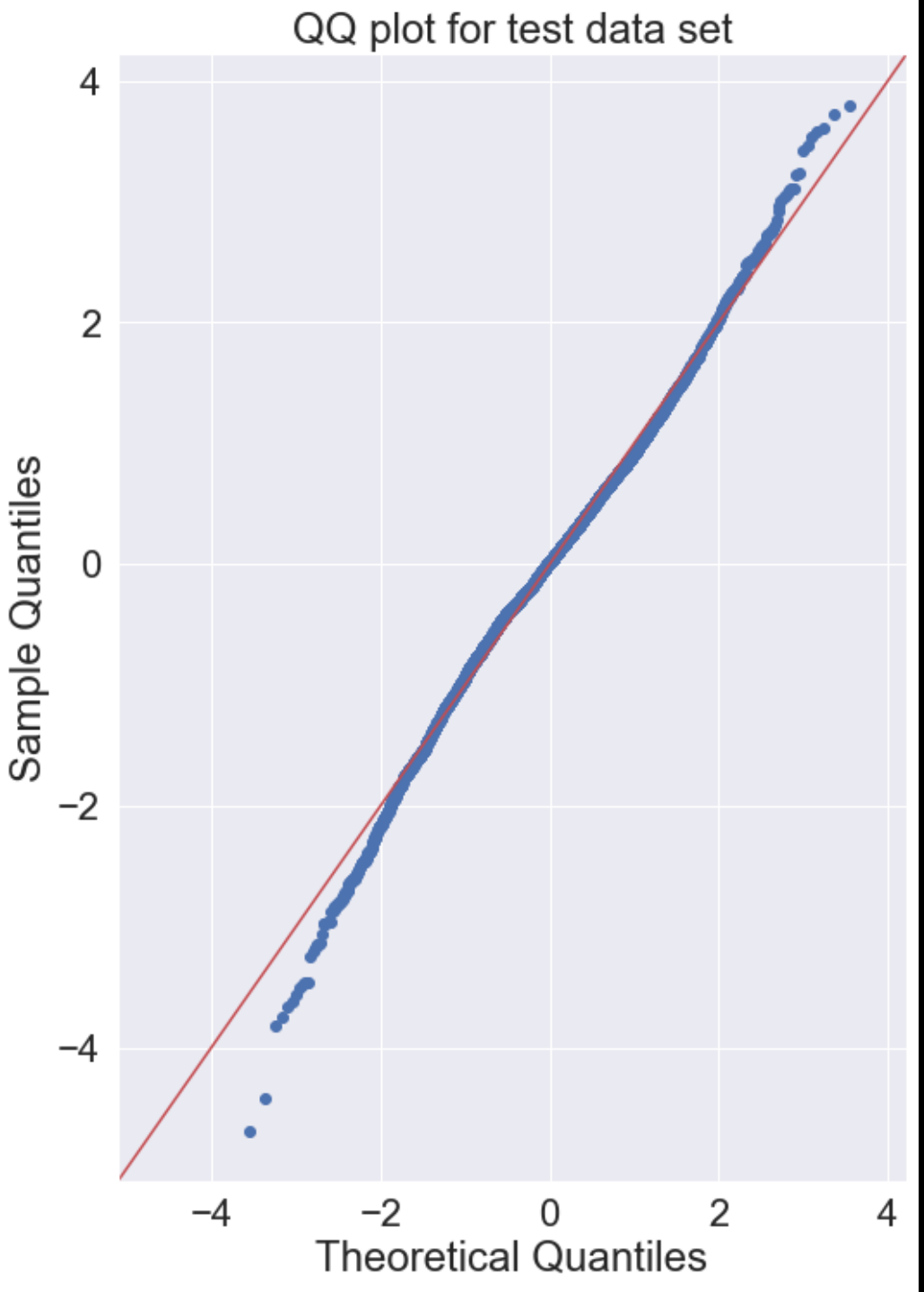
Price vs. Distance from MS HQ



Price vs. Average distance from main attractions




QQ plot for training data set
Sample Quantiles
Theoretical Quantiles
4
2
0
−2
−4
−6
−6
−4
−2
0
2
4


QQ plot for test data set
Sample Quantiles
Theoretical Quantiles
4
2
0
−2
−4
−4
−2
0
2
4


Training Residual Graph
Training Residuals
Residuals
y_train
1.0
0.5
0.0
−0.5
−1.0
−1.5
12
13
14
Test Residual Graph
Testing Residuals
Residuals
y_test
1.0
0.5
0.0
−0.5
−1.0
12
13
14