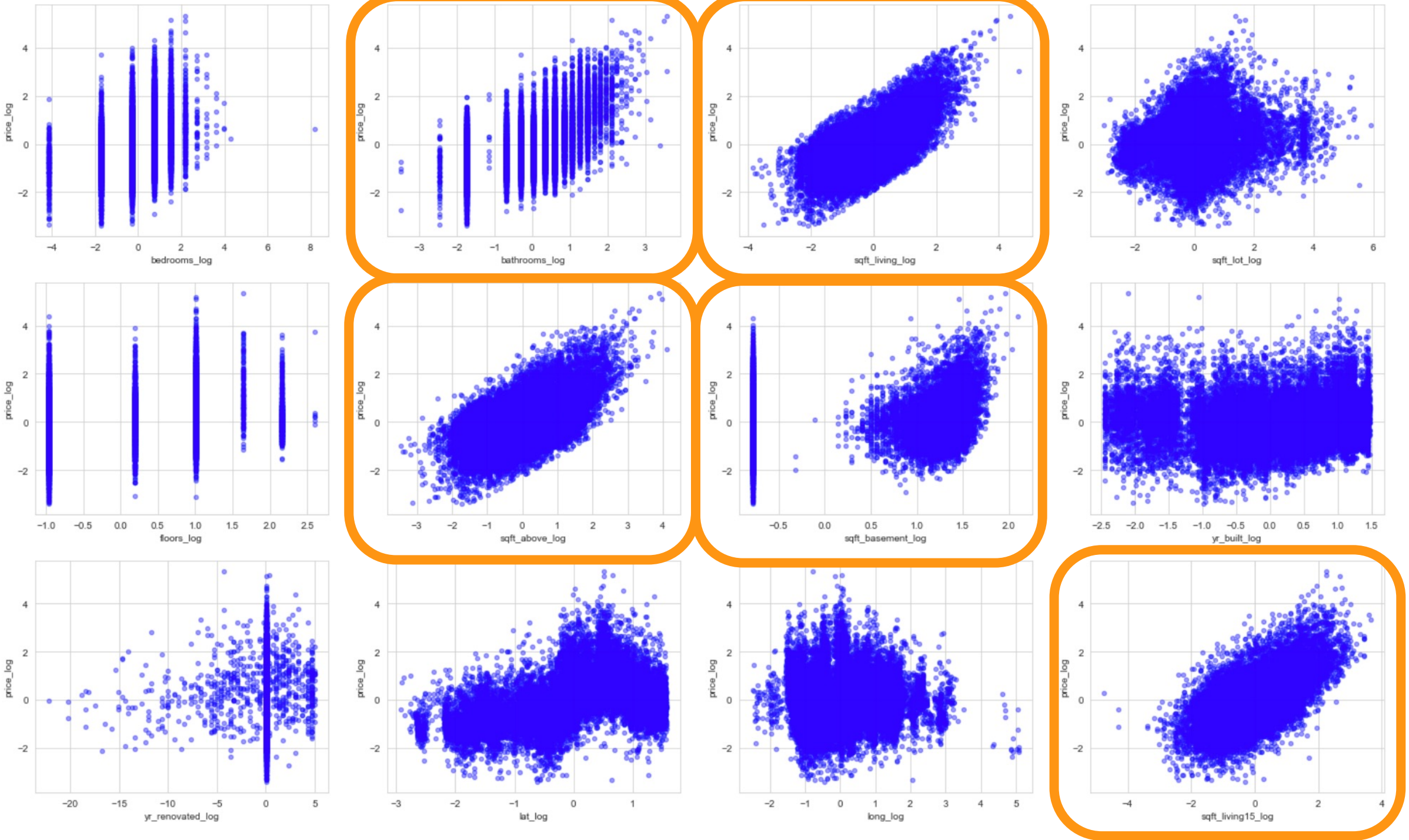
Exploring Housing Sales

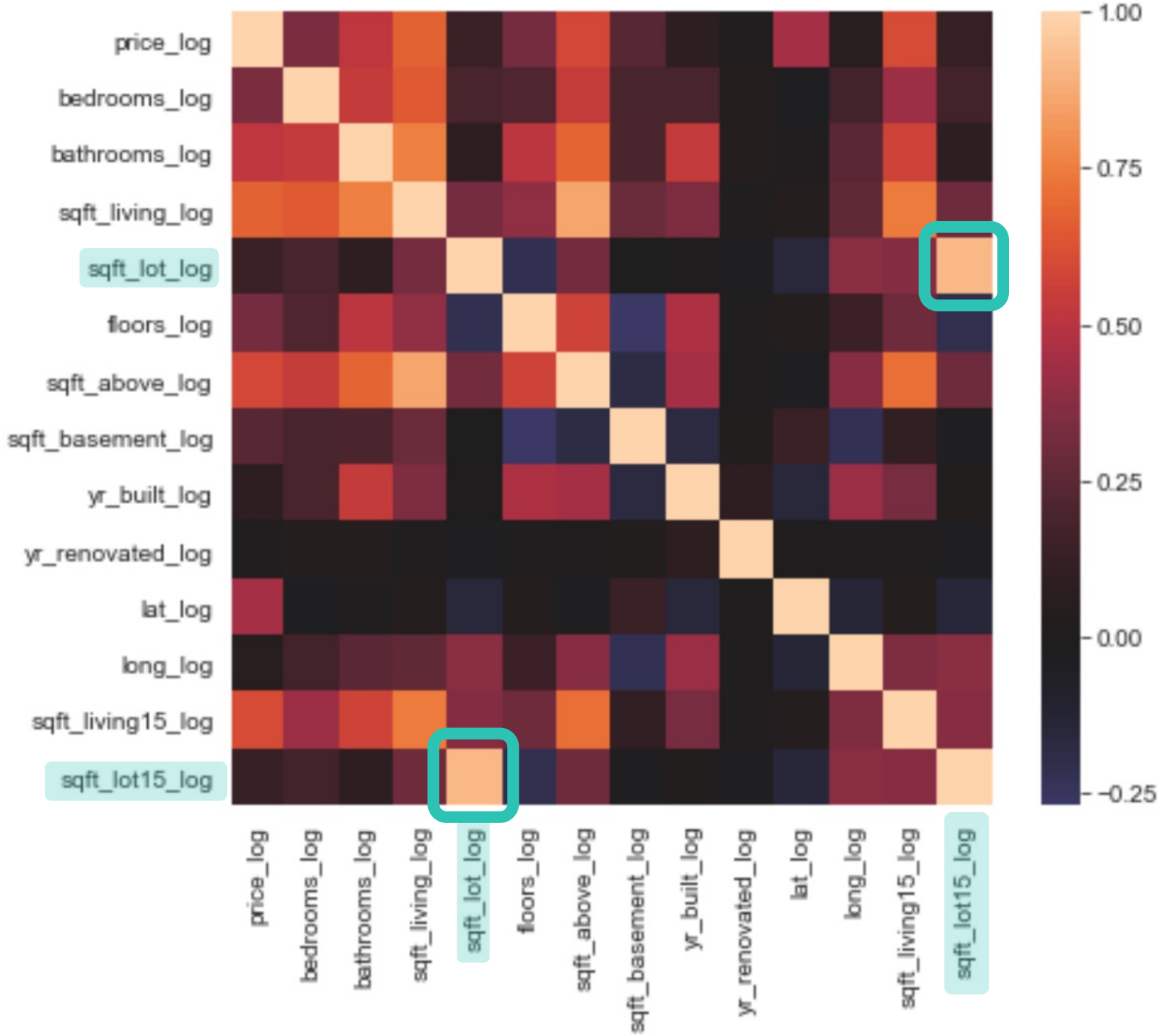
Caitlin Snyder
Flat Iron - Data Science
Module 2

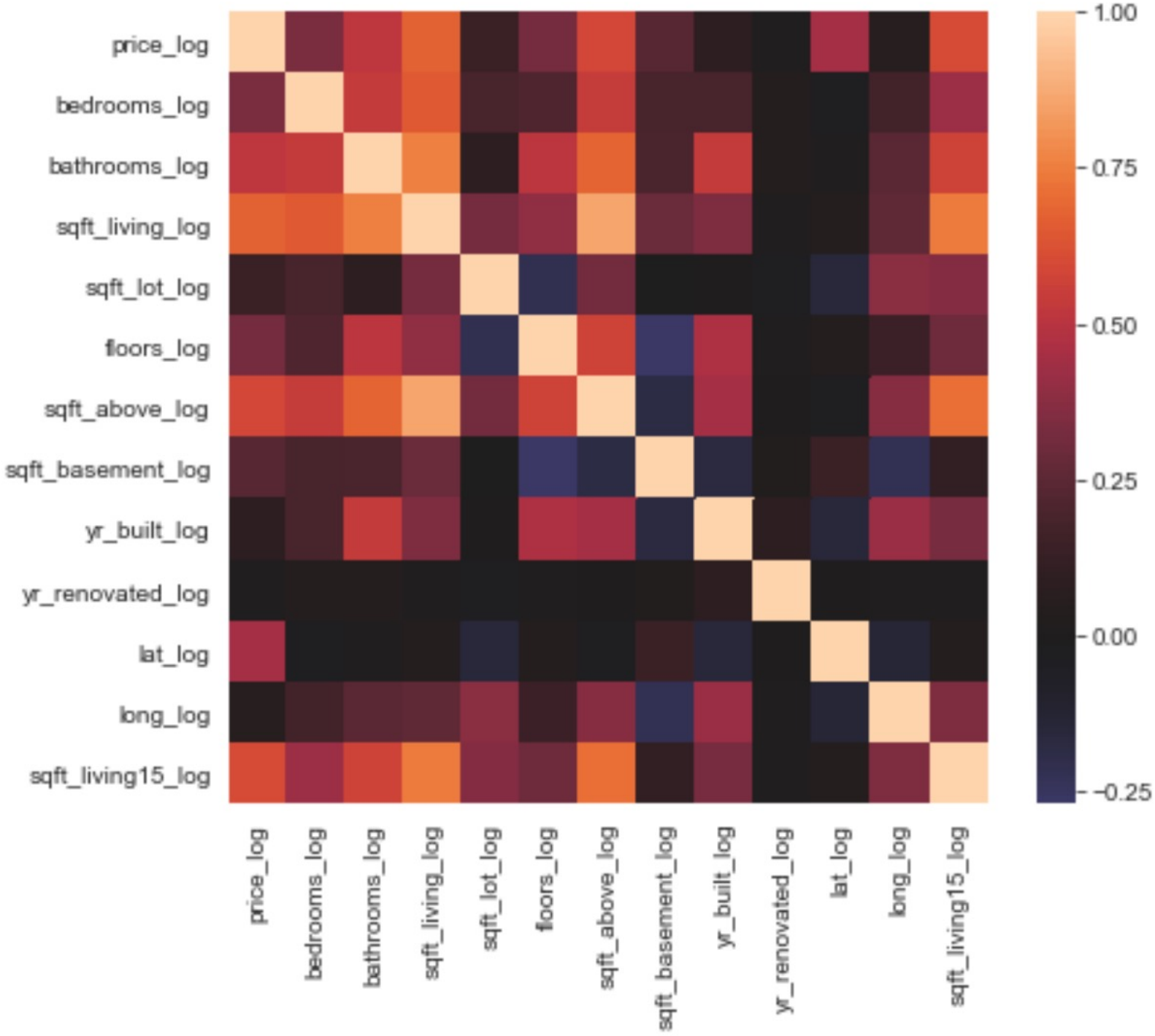
Questions

1. What is the effect on sale price of a low grade* classification?
2. What is the effect on sale price of a high condition* classification?
3. What is the effect on sale price of having a waterfront?

**as defined by the King County Assessor's Office*


price_log
bedrooms_log
bathrooms_log
sqft_living_log
sqft_lot_log
floors_log
sqft_above_log
sqft_basement_log
yr_built_log
yr_renovated_log
lat_log
long_log
sqft_living15_log


price_log
bedrooms_log
bathrooms_log
sqft_living_log
sqft_lot_log
floors_log
sqft_above_log
sqft_basement_log
yr_built_log
yr_renovated_log
lat_log
long_log
sqft_living15_log
sqft_lot15_log
1.00
0.75
0.50
0.25
0.00
−0.25


price_log
bedrooms_log
bathrooms_log
sqft_living_log
sqft_lot_log
floors_log
sqft_above_log
sqft_basement_log
yr_built_log
yr_renovated_log
lat_log
long_log
sqft_living15_log
1.00
0.75
0.50
0.25
0.00
−0.25

Evaluating the model

Training r^2: 0.887

Test r^2: 0.881

Training MSE: 0.113

Test MSE: 0.118

Question 1:
What is the effect on sale price of a low grade classification?

0.7861

decrease in sale price units

Question 2:
What is the effect on sale price of a high condition classification?

0.1305

increase in sale price units

Question 3: What is the effect on sale price of having a waterfront?

0.891

increase in sale price units

Take-aways

- Expect larger, higher quality structures with attractive surroundings to command higher sale prices.
- Looking to increase your property's market value? Invest in the structural integrity and protect against property degradation over time.

Sources

- Outliers: https://towardsdatascience.com/ways-to-detect-and-remove-the-outliers-404d16608dba, https://stackoverflow.com/questions/23199796/detect-and-exclude-outliers-in-pandas-data-frame
- Log transformations: https://discuss.analyticsvidhya.com/t/methods-to-deal-with-zero-values-while-performing-log-transformation-of-variable/2431/2
- OHE: https://towardsdatascience.com/categorical-encoding-using-label-encoding-and-one-hot-encoder-911ef77fb5bd
- Stepwise regression: https://towardsdatascience.com/stepwise-regression-tutorial-in-python-ebf7c782c922
- Interpreting coefficients: https://statisticsbyjim.com/regression/interpret-coefficients-p-values-regression/

Thank you!