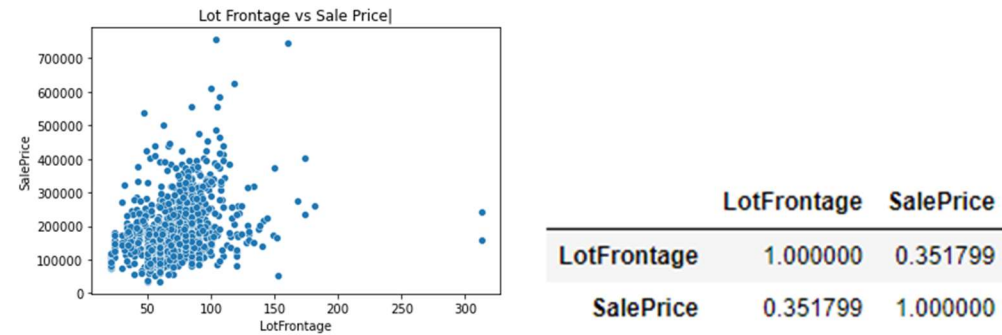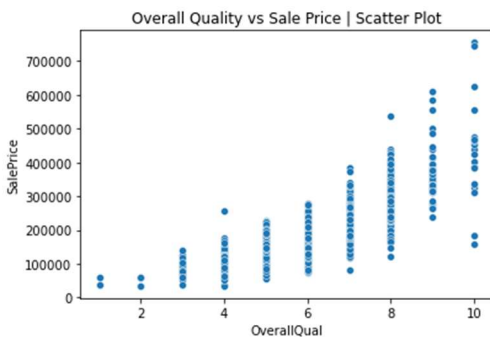Assignment 3 Part 3 CPSC 4800
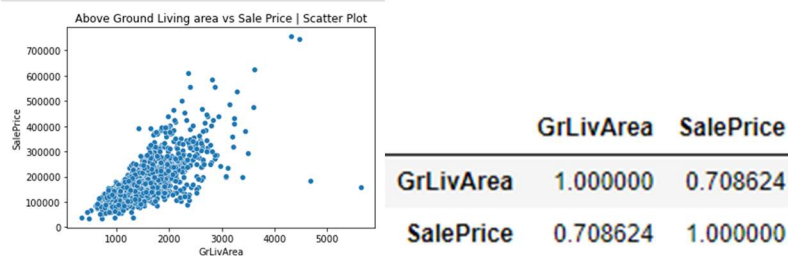
I conducted an analysis into the sale price of houses from the train dataset from
https://www.kaggle.com/competitions/house-prices-advanced-regression-techniques/overview .
Through some exploratory data analysis, I created 3 hypotheses to test: Was housing sale price
associated with the lot frontage of the home, the overall quality of the home, and the above ground
living area of the home.

	LotFrontage	SalePrice
LotFrontage	1.000000	0.351799
SalePrice	0.351799	1.000000

Lot Frontage showed weak, positive correlation. It's difficult to say if there is any association between
Lot Frontage and sale price. The Lot Frontage was specifically defined as linear feet of street connected
to property. There are also significant outliers in both directions of the scatter plot, and deeper
understanding and perhaps removal could improve analysis.

We see that the minimum, maximum, and mean sale price generally increase with the "overall quality"
of the house parameter. Although I did not conduct any statistical tests to see if it was significant, there
seems to be some preliminary evidence that sale price will increase as overall quality increases.

	GrLivArea	SalePrice
GrLivArea	1.000000	0.708624
SalePrice	0.708624	1.000000

Finally, we saw that there was significant, positive correlation between above ground living area and
sale price.