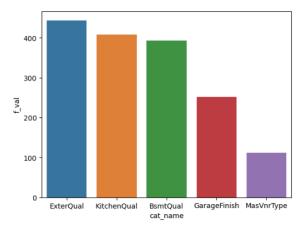
Assignment_3_Part_3 - Conclusion

What are the top 5 categorical variables that most affect house prices?

Figure 1 shows that categorical variables have significant mean differences among groups of each variable corresponding to the sales price by using ANOVA statistical test.

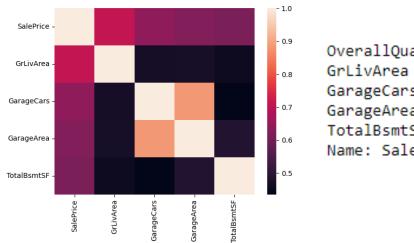


| | cat_name | f_val | p_val | is_statistically_significant |
|----|--------------|------------|---------------|------------------------------|
| 17 | ExterQual | 443.334831 | 1.439551e-204 | True |
| 29 | KitchenQual | 407.806352 | 3.032213e-192 | True |
| 20 | BsmtQual | 392.913506 | 9.610615e-186 | True |
| 32 | GarageFinish | 250.962467 | 1.199117e-93 | True |
| 16 | MasVnrType | 111.672380 | 4.793331e-65 | True |

Figure 1: top 5 categorical variables

What are the top 5 numerical variables that most affect house prices?

Figure 2 shows that numerical variables have a strong correlation coefficient to the sales price by using correlation tests.



OverallQual 0.790982 GrLivArea 0.708624 GarageCars 0.640409 GarageArea 0.623431 TotalBsmtSF 0.613581

Name: SalePrice, dtype: float64

Figure 2: top 5 numerical variables