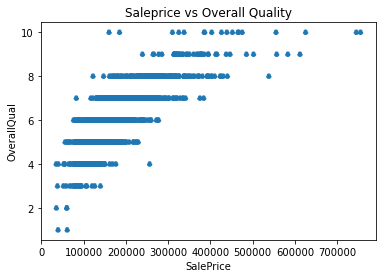
Assignment 3 Part 3

We will be using Housing Price Data from Kaggle to test our hypotheses.

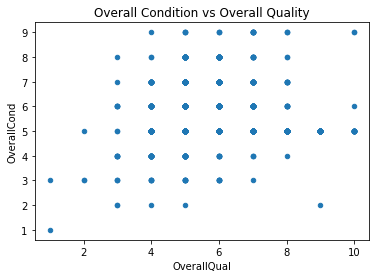
Hypothesis 1:

There is a correlation between sale price vs overall quality (refers to material and finish quality). Strong correlation of 0.79. This is statistically significant, and the scatterplot shows that when the overall quality rating is higher that the price will be higher as well. This is known through linear regression analysis. This scatterplot shows a strong positive linear relationship between price and quality. This could be because when there is higher quality builds, they are worth more as the demand will be higher thus influencing prices.



Hypothesis 2:

There is a correlation overall quality and overall condition(rating). I choose this as you would think the condition of a home would increase when the quality is better. This scatter shows us that there is not enough statistical evidence of a correlation between the two variables. The correlation is -0.091932 which is not large enough to be significant. Thus, we must accept that the overall quality of the house does not affect the overall condition. You can see this in the graph below as there is no linear form.



Hypothesis 3:

Lastly, I want to see if the Sale Price of houses and the Lot Area of the house are correlated. I chose this as larger area usually equates to a higher house price. In our correlation analysis we see a slight positive correlation of 0.29. Our scatterplot shows this below.

Chart, scatter chart

Description automatically generated