

# Predicting Home Sale Prices: Training the Trend

-Manjila Moktan

A decorative light blue triangle is located in the bottom right corner of the slide.

## Problem Statement :

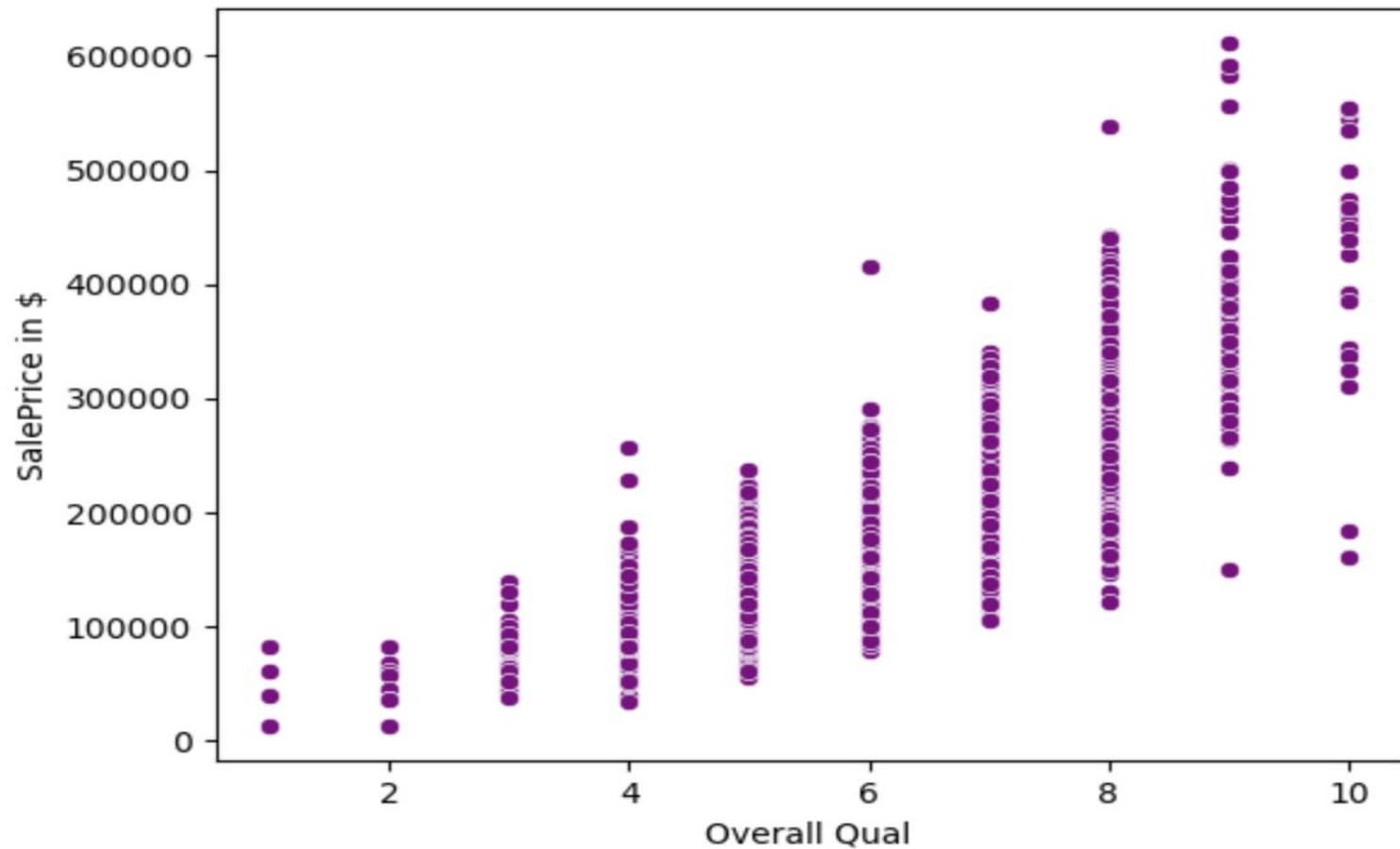
Accurate Pricing of a home for sale is essential to any real estate agency. When listed price is too high, it can deter potential buyers whereas the opposite will result in loss of income for sellers as well as reduced commission for the realtors. There needs to be the right balance of listing price for your listing to be sold quickly in the amount it is listed for.

# Data Cleaning

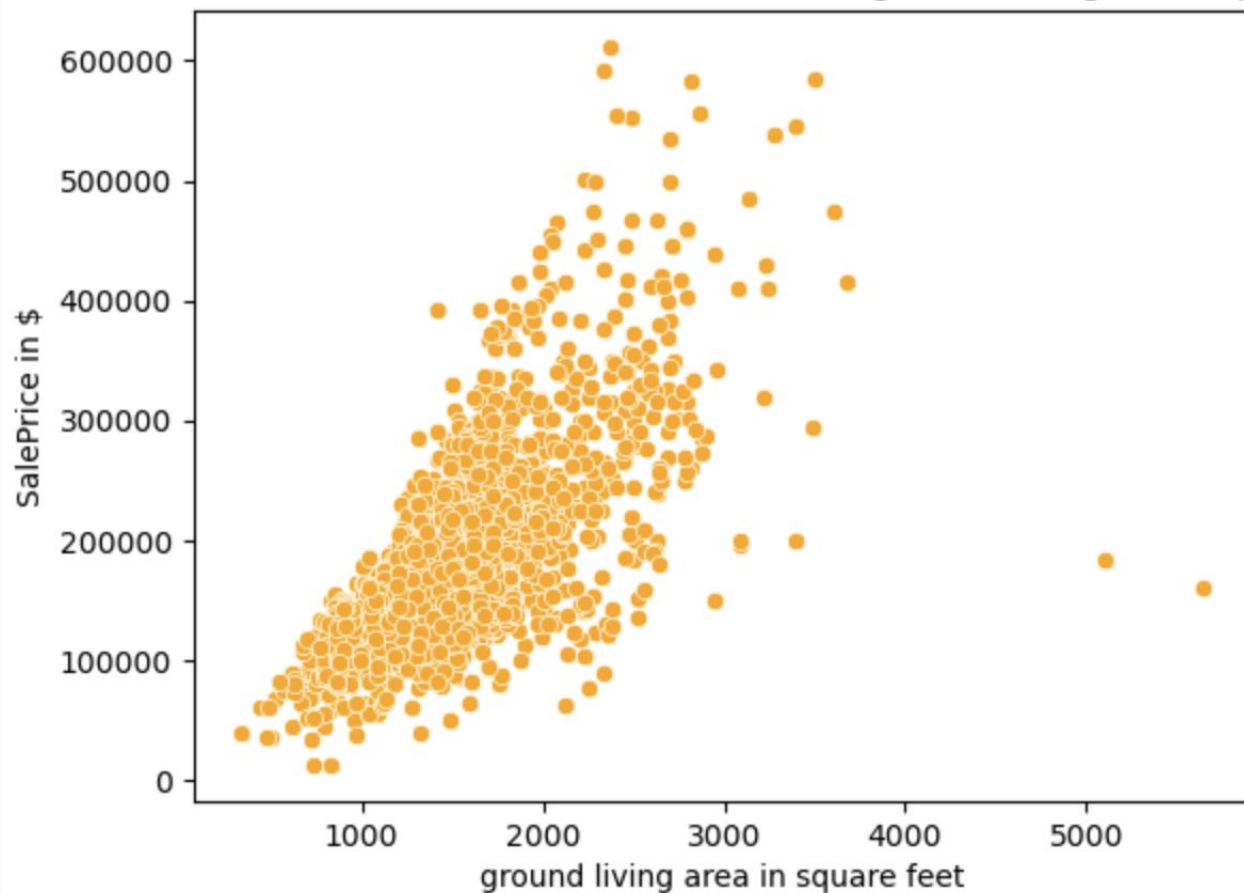
Empty Values: (data dictionary, median, mean)

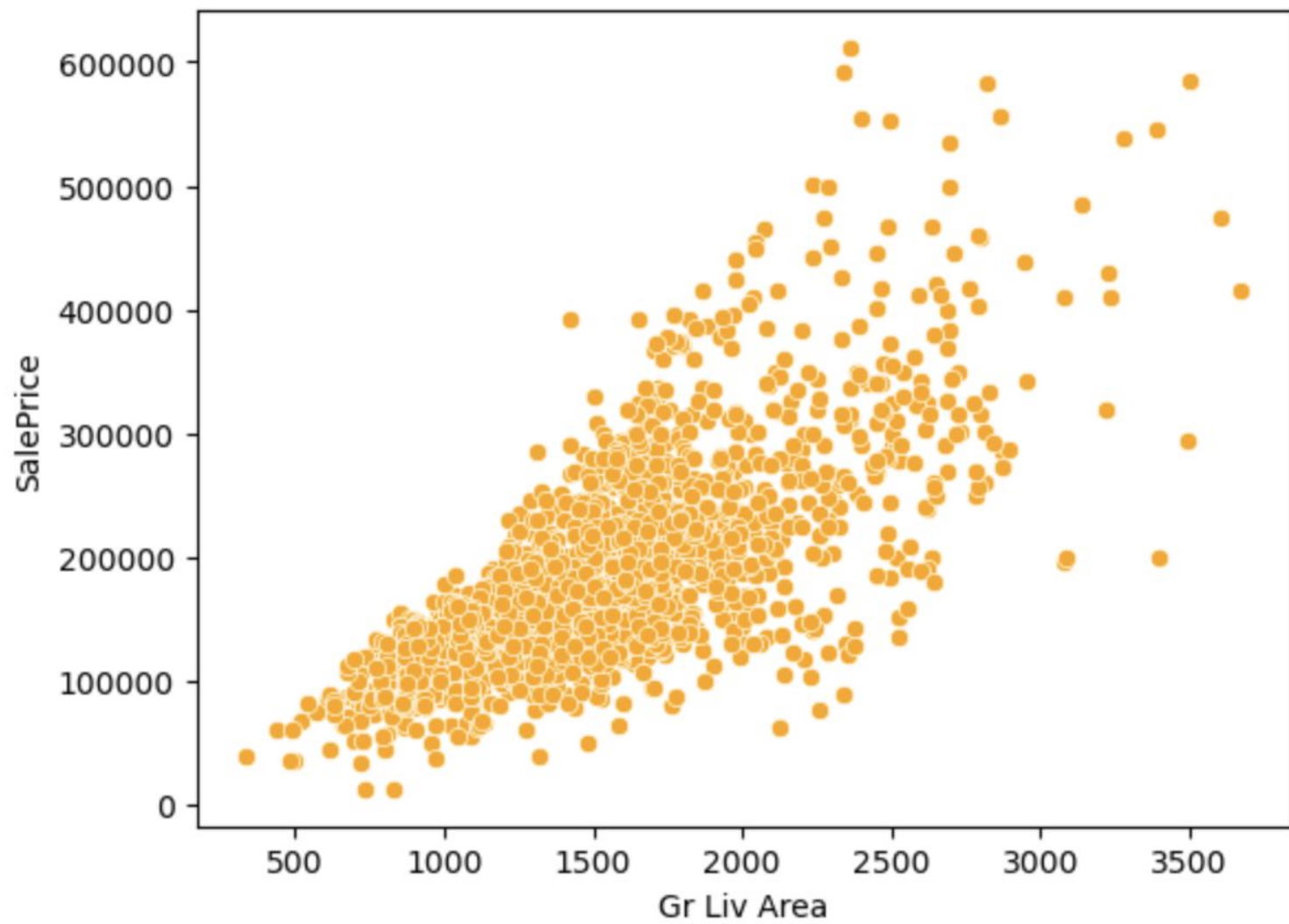
- Data dictionary : garage types,: no garage for empty one, similar for fence and pool
- Median for discrete variables eg. number of car in garage
- And mean continuous variables e.g. for most of the areas

Does SalePrice increase with increase in Overall Quality?

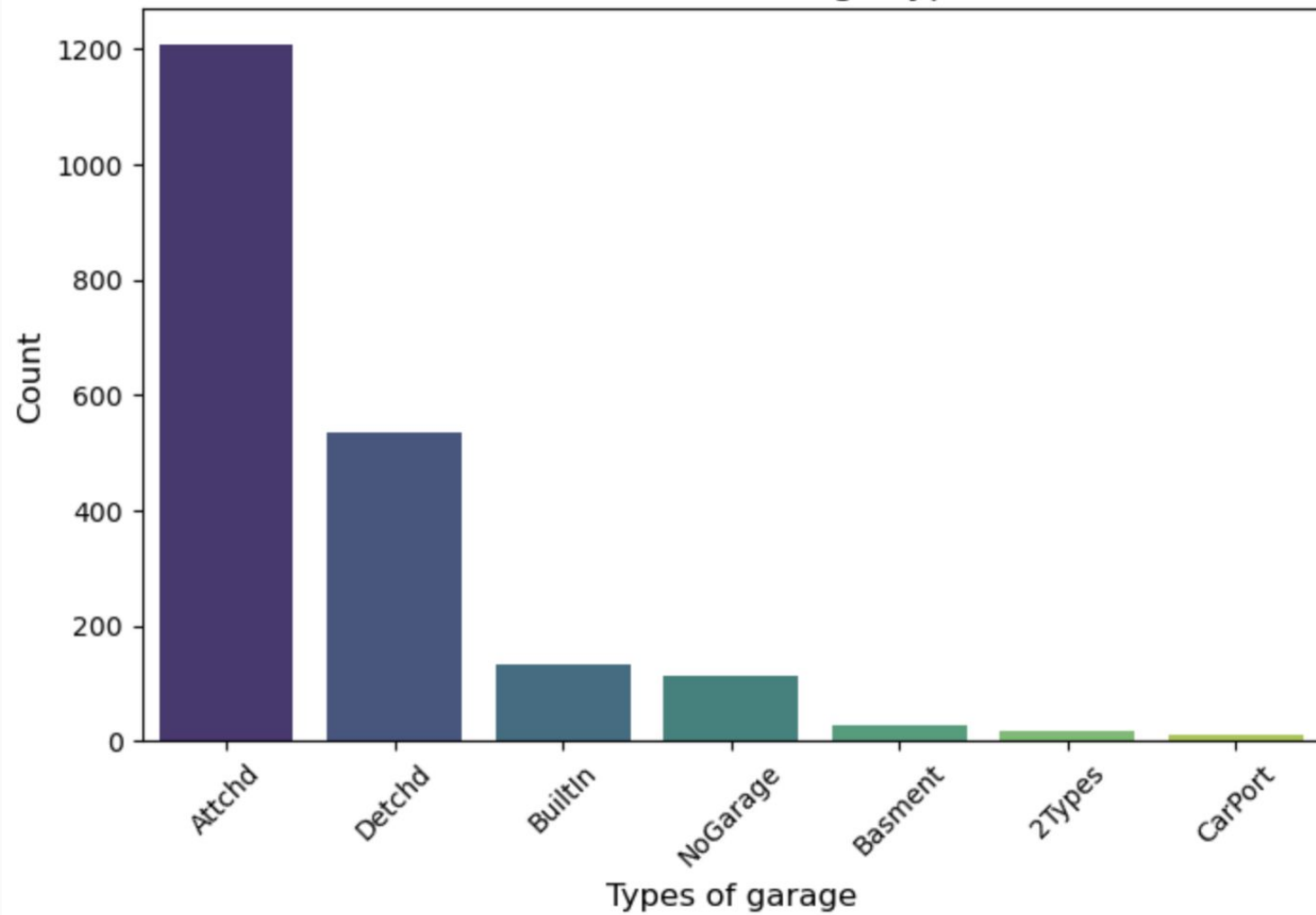


Does SalePrice increase with increase in Above ground living area square feet?

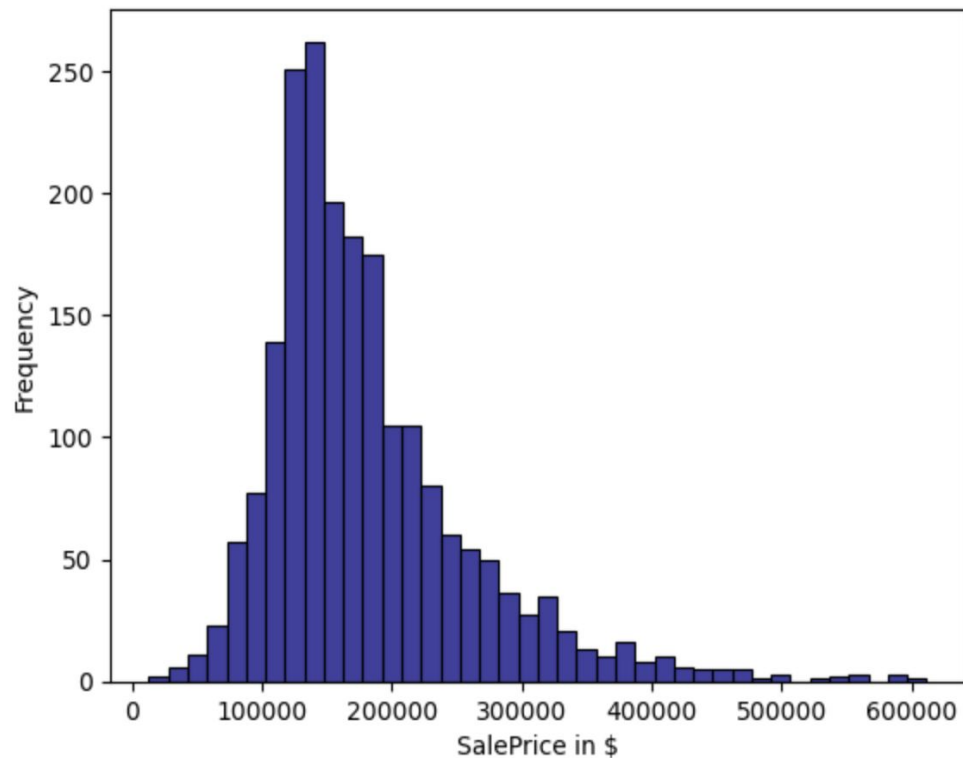




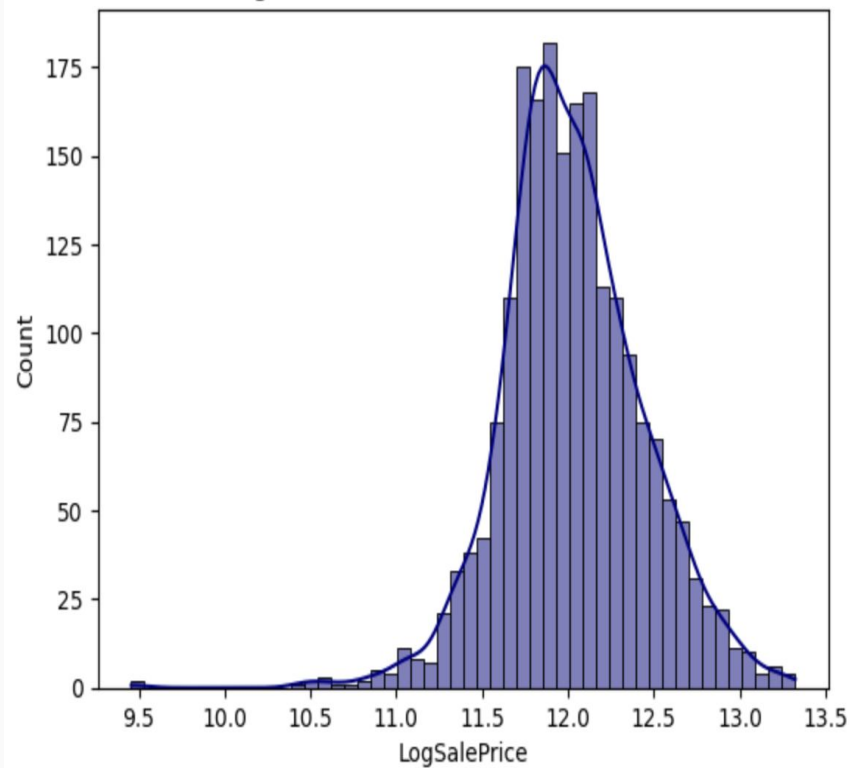
Distribution of Garage Types



Distribution of SalePrice



Log-Transformed SalePrice Distribution





## Best Model:

Root Mean Squared Error (RMSE)	$R^2$ (coefficient of determination)
\$24334.03	0.9057

Feature	Coefficient
Overall Qual	0.084117

For every step increase in Overall condition, the price of the house increases by 8.41%, holding all else constant.

I.e. \$200,000 home increases to \$217,540

### Recommendation:

- When pricing a home, the overall quality of the home must be taken into great consideration.

### Benefits of modeling:

- Accuracy
- Cost effective
- Time saving

