Statistical Report for Iowa Housing Dataset

I. Overview of Dataset

Objective: To find what factors affect housing prices.

Sample size: 1460

Features Used: Since there 79 explanatory variables, selective ones chosen for setting up hypothesis

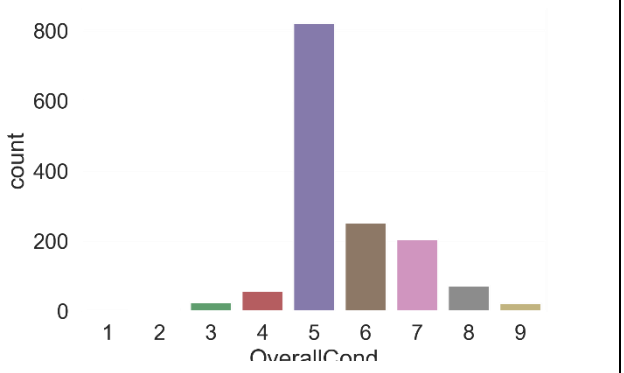
Categorical	Numerical
SaleType: Type of sale	SalePrice: Property's sale price (in USD)
	OverallQual: Overall material and finish quality
	GarageArea: Size of garage in square feet
	GrLivArea: Above grade (ground) living area square feet
	YearBuilt: Original construction date
	FullBath: Full bathrooms above grade
	TotalBsmtSF: Total square feet of basement area
	MSSubClass: The building class

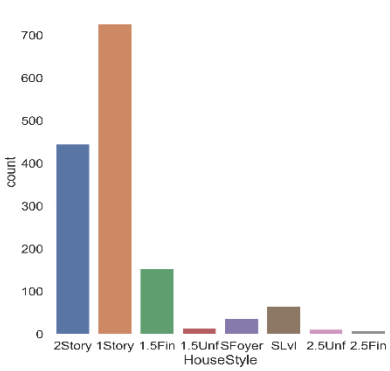
Target Feature: SalePrice

II. **Data Preparation**: After inspecting the dataset using .head() and .shape, dropped the ID column since it would not have any impact on the analysis. Then checked for missing values (identified: LotFrontage). Dropped LotFrontage since it had the majority of values missing. Examined distribution of target variable, was right-skewed since mean was greater than median. Segregated the numerical and categorical variables to ease Univariate and Bivariate Analysis.

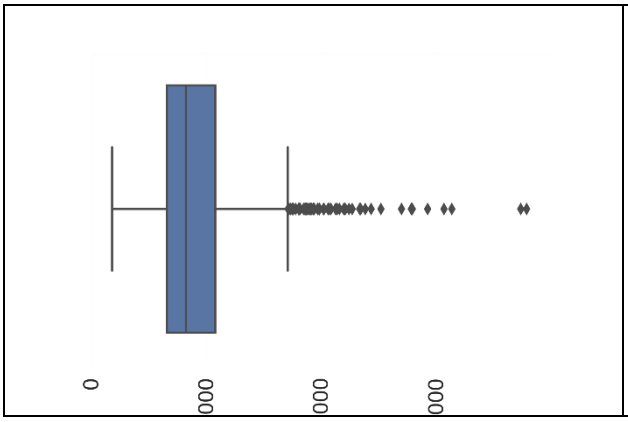
III. Univariate and Bivariate Analysis

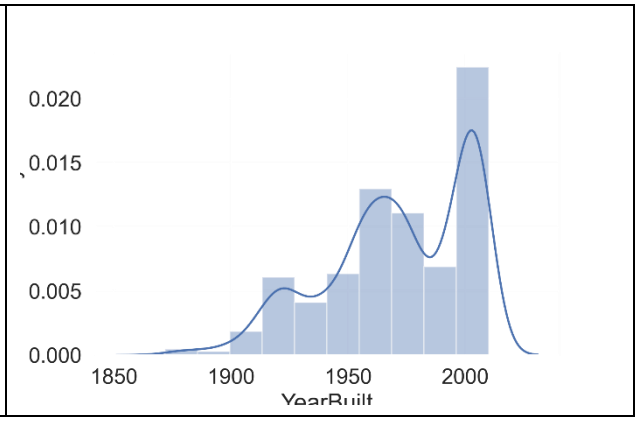
Univariate for categorical variables (for overall condition and house style)



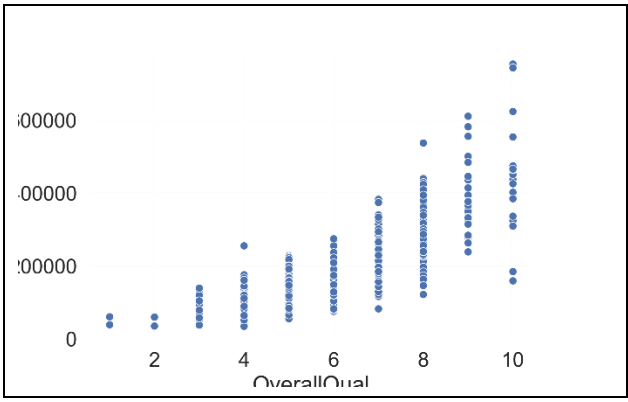


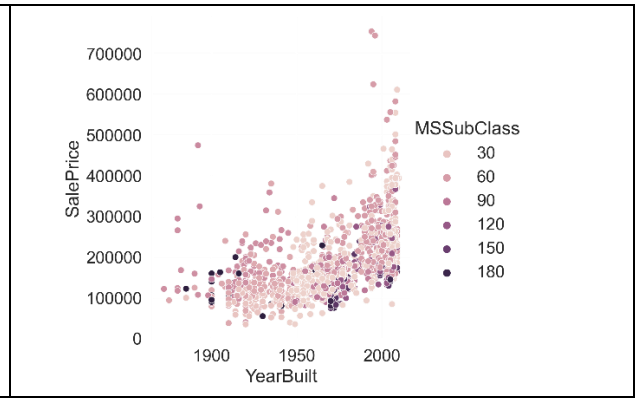
Univariate for numerical variables (Target Variable: Sale Price and Yearbuilt)





Bivariate and Multivariate Analysis (OverallQuality and)

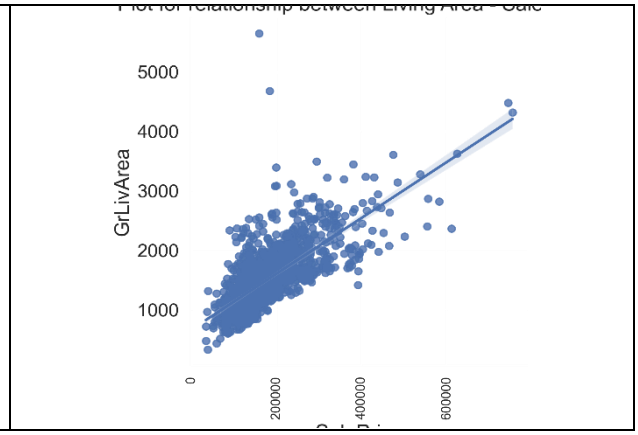




IV. Hypothesis Testing

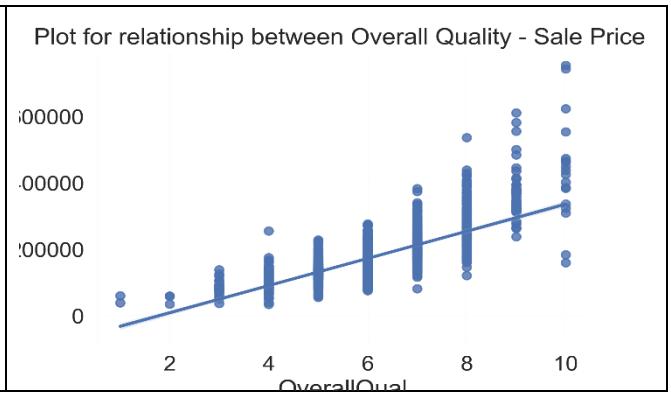
Determine if living area affect housing price

- There is a strong positive correlation between living area and housing price.
- As the living area increases, the price of the house tends to increase as well. Therefore, living area is a significant factor in determining the housing price.



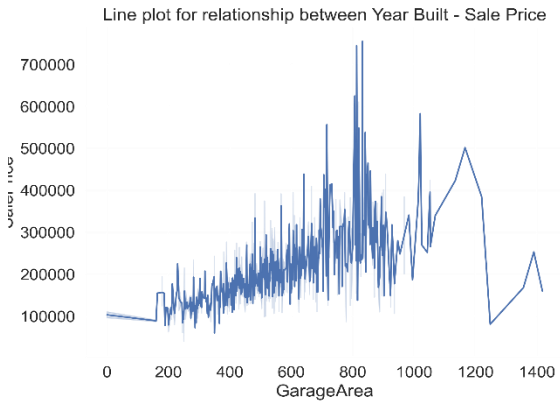
Determine if there is any correlation between overall quality of the house and price.

- There is a strong positive correlation between overall quality and price.
- As the overall quality of the house increases, the price of the house tends to increase as well.



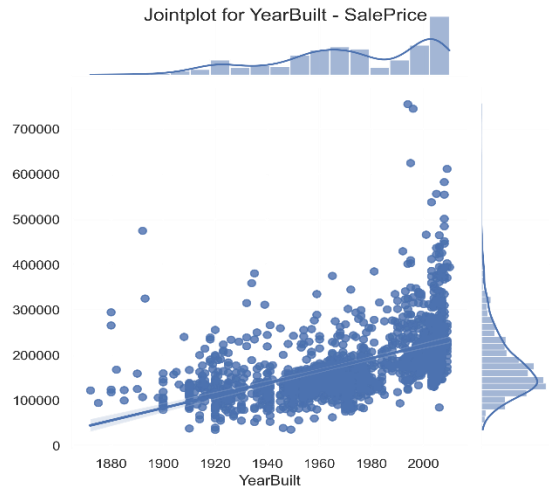
Does Garage area influence the housing price?

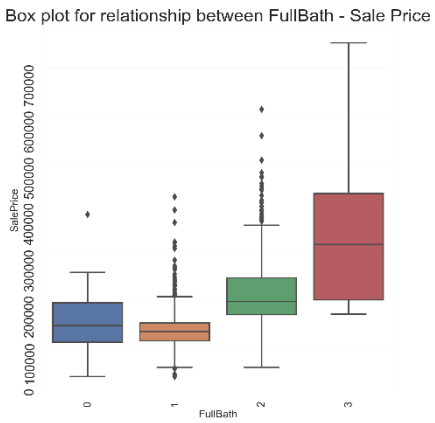
- Positive correlation between garage area and housing price.
- As the garage area increases, the price of the house tends to increase as well. However, there is a significant amount of variability in the data, and some houses with smaller garage areas also have high prices.
- while garage area is a factor in determining housing price, it is not the only significant factor.



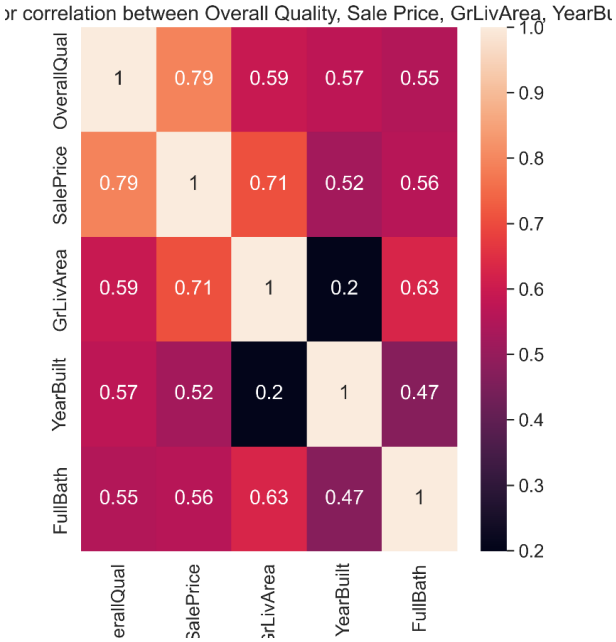
Determine if there is correlation between the year house is built and cost

- Weak positive correlation between the year the house is built and the housing price
- As the year the house was built increases, the price of the house tends to increase slightly.
- A lot of variability in the data, and some older houses also have high prices. Year the house is built is a factor in determining housing price, it is not a strong predictor.



Is there any association between FullBath and SalePrice? - weak positive correlation between FullBath and SalePrice. - As the number of FullBaths increases, the price of the house tends to increase slightly - A lot of variability in the data, and some houses with fewer FullBaths also have high prices. - while FullBath is a factor in determining housing price, it is not a strong predictor.	

V. Conclusion



Overall, analysis of the housing price dataset reveals that various factors impact the price of a house. Among these, overall quality, living area and garage area are the strongest predictors of housing price. A larger living area tends to correlate with a higher sale price, while houses with better quality also generally have higher prices. On the other hand, the impact of year the house was built, and FullBath on housing prices is comparatively weaker. While houses with larger garages or more FullBath tend to have slightly higher prices, the relationship is not as strong as that of living area or overall quality.

From the 5-hypothesis stated, living area and overall quality are strong predictors of housing price, while garage area, the year the house was built, and FullBath are weaker predictors.