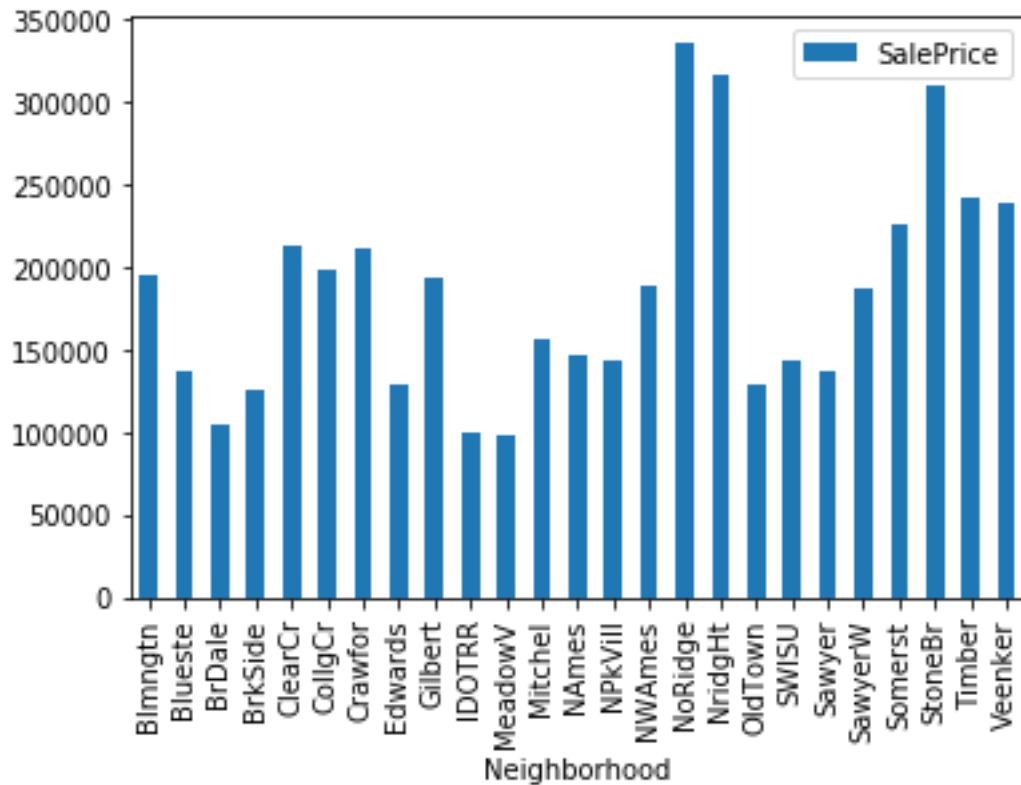


Abstract

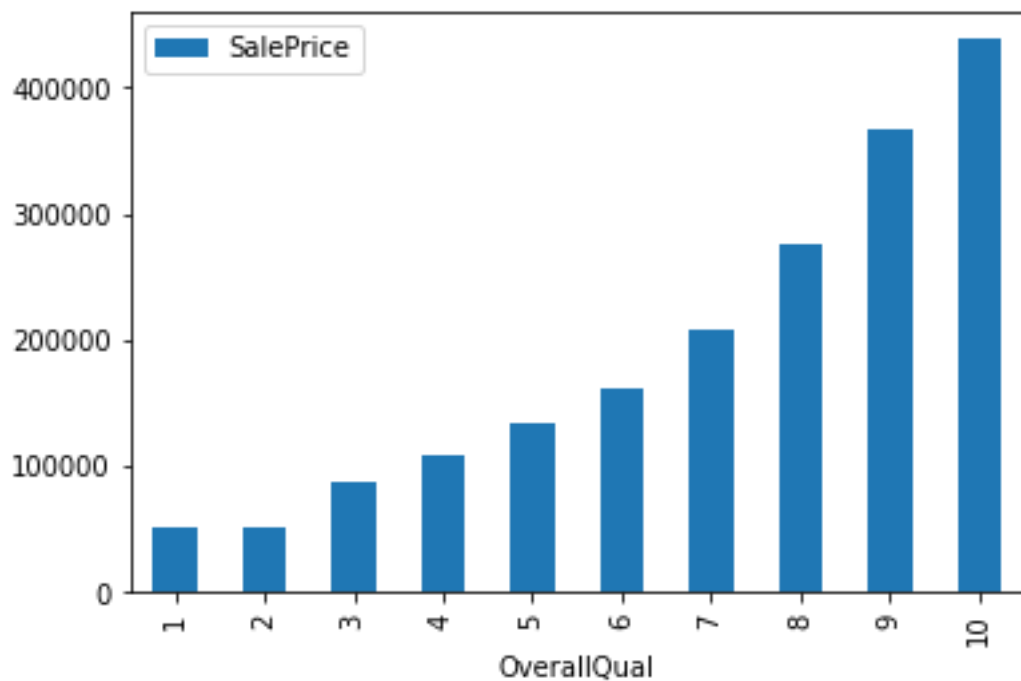
Analysis of the factors affecting the sales prices of houses across different neighborhoods in the Ames United States, based on the Ames Housing Dataset. The dataset contains records of 1460 house sales.

Mean Sales Price By Neighborhood



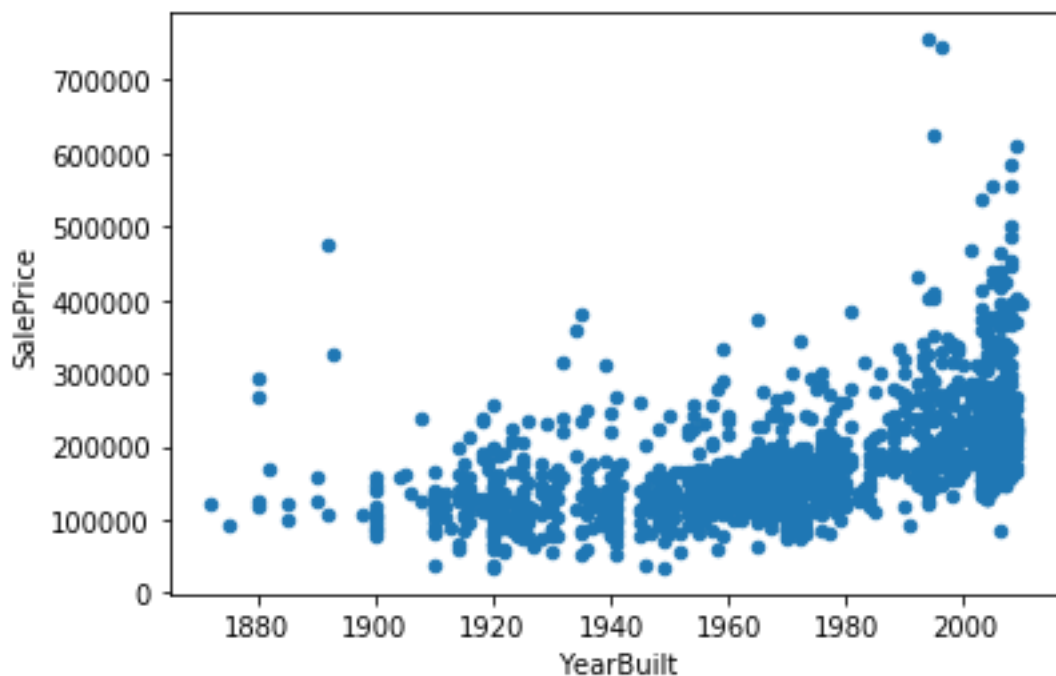
The above chart shows the mean sales price of the houses as it spreads across different neighborhoods. To compute the above, a pivot table was created, containing the aggregate mean of house prices, indexed by the neighborhood. The bar chart shows that **StoneBr**, **Nright**, and **NoRidge** have the highest house prices by a far margin. On the lower side, **Blueste**, **BrkSide**, **ClearCr**, **Idotrr** and **MeadowV** represents the least priced neighborhoods, with less than \$150,000 in mean price sales of houses. The majority of the neighborhoods are similar in house sale prices, falling in the range of \$150,000 to \$250,000.

Mean Sales Price By Overall Quality

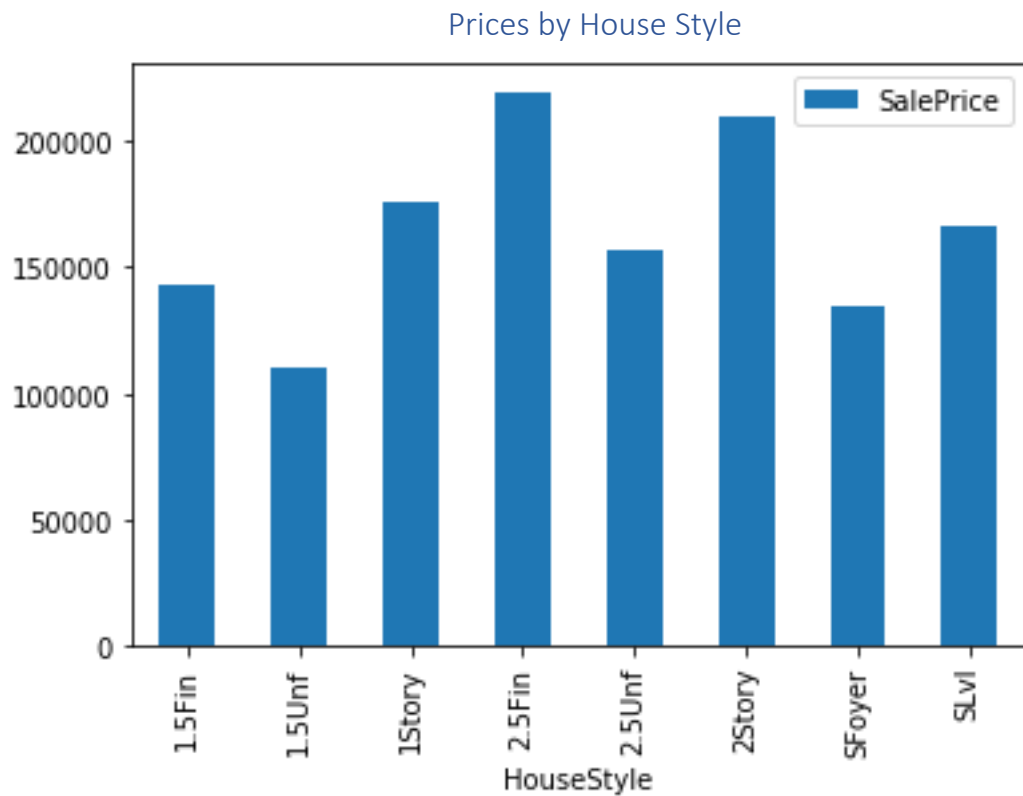


The sales price directly correlates with the overall quality rating of the houses.

Mean Sales Price by Year Built



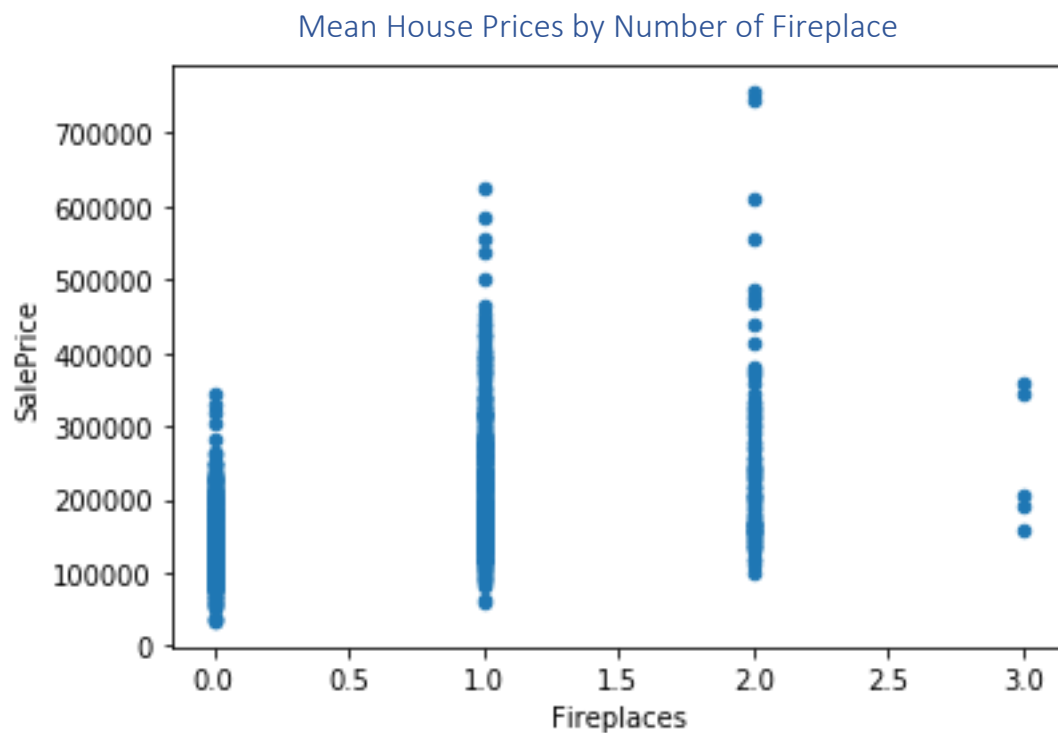
The scatter plot above shows the aggregated mean prices of house sales grouped by the year the houses were built. Houses built between 1910 and 1980 show similar trend in price, however, houses built about the 2000 indicate a significant increase in the prices they were sold.



The aggregate mean price by house style above indicates a high buying preference for two house styles, **2.5Fin** and **2 Story**. The least preferred style is **1.5Unf** above.



The scatter plot above represents the sales by the year each house was sold. Surprisingly, between 2006 to 2010, the trend seems to indicate an almost constant price with no significant changes.



To understand how the number of fireplaces influences the selling price of the houses, the sales price is plotted against the number of fire places. The scatter plot indicates a high preference for either 1 or 2 fire places. Houses with 3 fire places are the least preferred.



The box plot above is for the Sales Price. The high number of whiskers here indicates a large number of outliers in the data, with prices between 350000 to over 700000.