|  |  |
| --- | --- |
| |  | | --- | | **ADTA 5130 DATA ANALYTICS**  **Group 3**  123.PNG    **Real Estate Home Pricing**  **Report by**  Sneha Maddi - 11533404  Sneha Sahithi Pamarthi –11514108  Sushma Sree Vadla -- 11514108  Midhi Naveen Kommalapati -- 11522489 | |

**Business Understanding :**

**Data Understanding :**

For the final project “Real Estate Home Pricing” that includes seven variables and 550 observations. Below are the following variables.

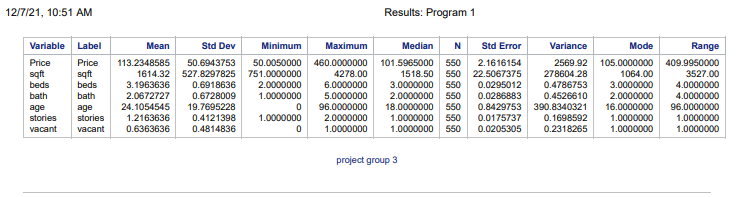
|  |
| --- |
| Variables  price  sqft  beds  baths  age  stories  vacant |

**Dependent/Target Variable :**

On this project we are evaluating the variable **“Price”.**

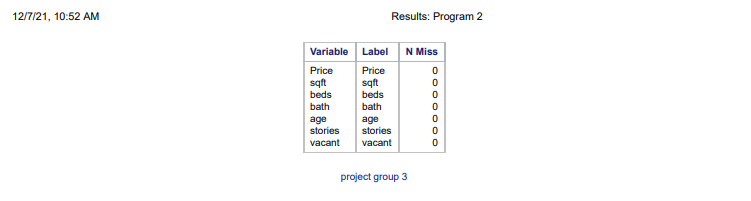
**Problem Statement :**

**Summary Statistics :**



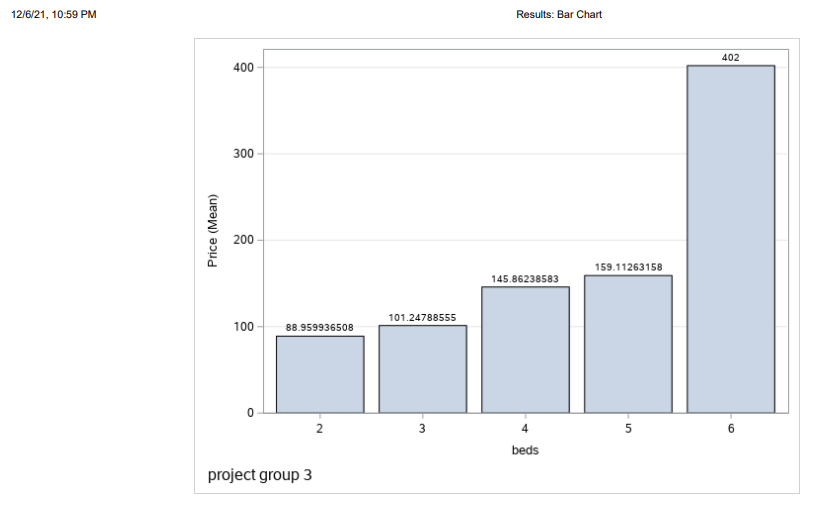
**Misssing Values :**

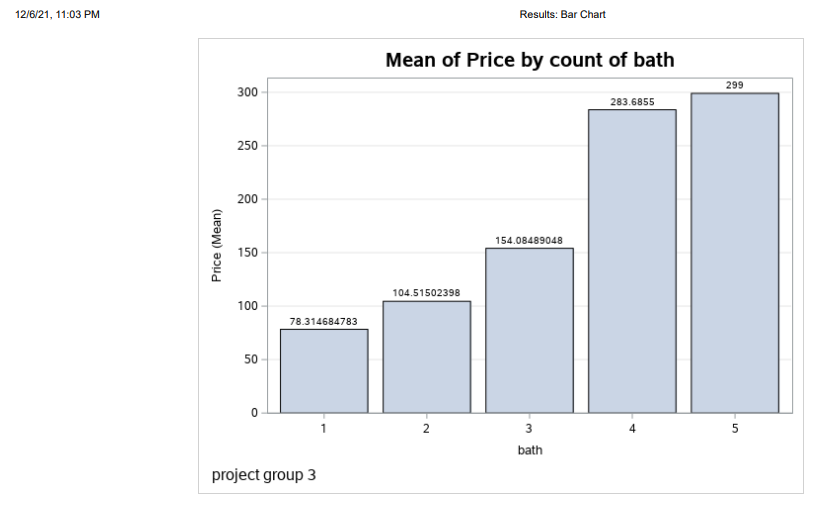
There are no missing values for this dataset.



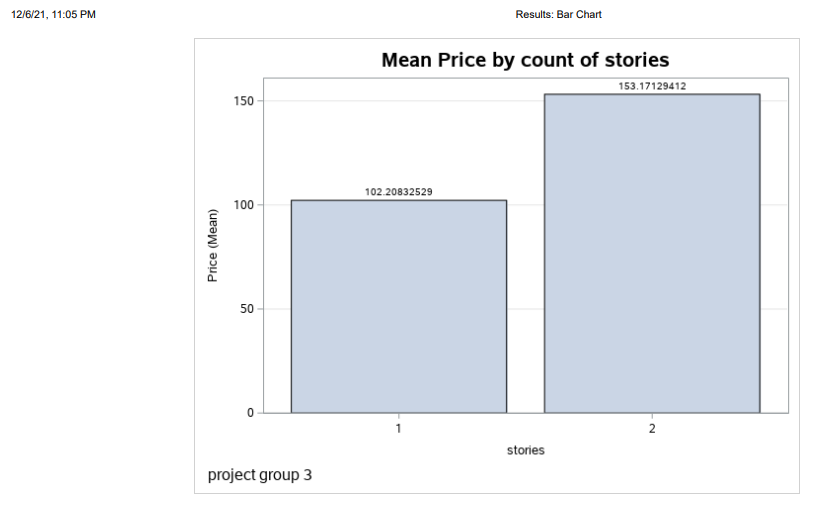
**Data Analysis :**

On the below we can see number of **beds** in a house increases.

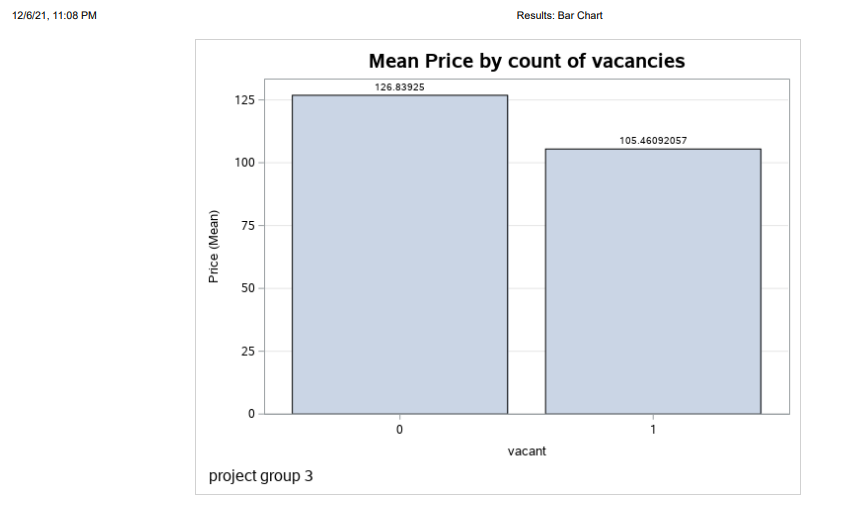


On the below we can see number of **baths** in a house increases.

On the below we can see number of **stories** in a house increases.

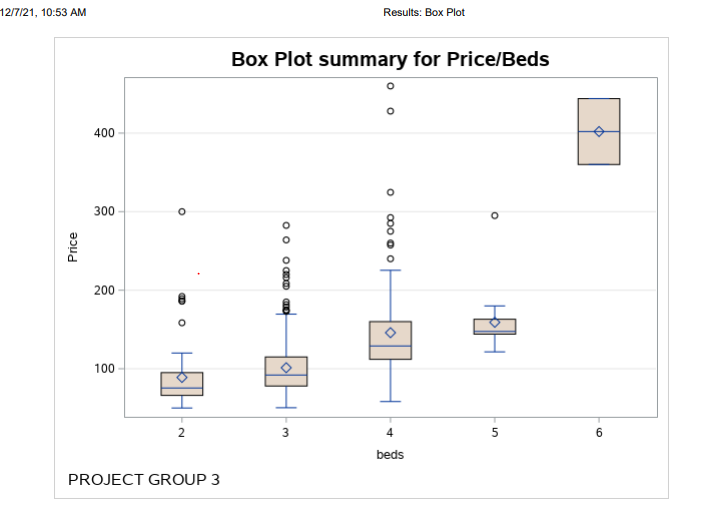


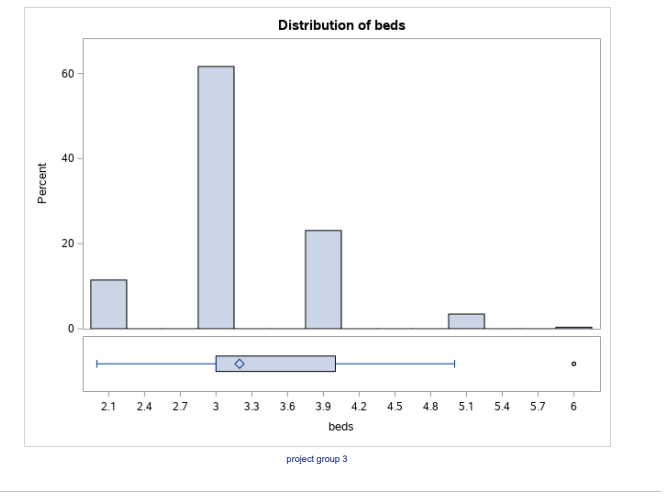
On the below we can see number of **vacancies** in a house increases.

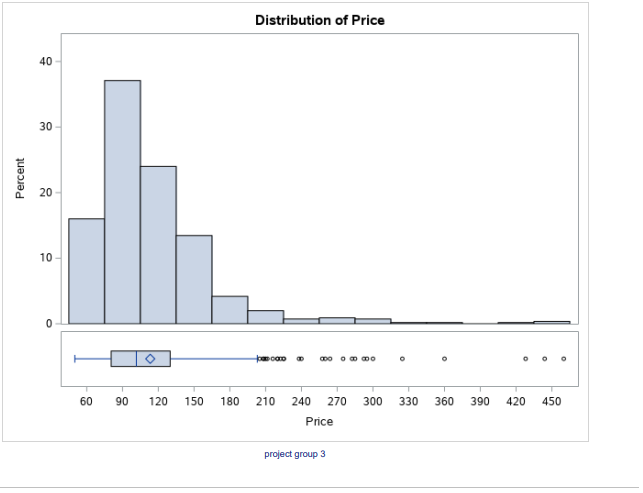


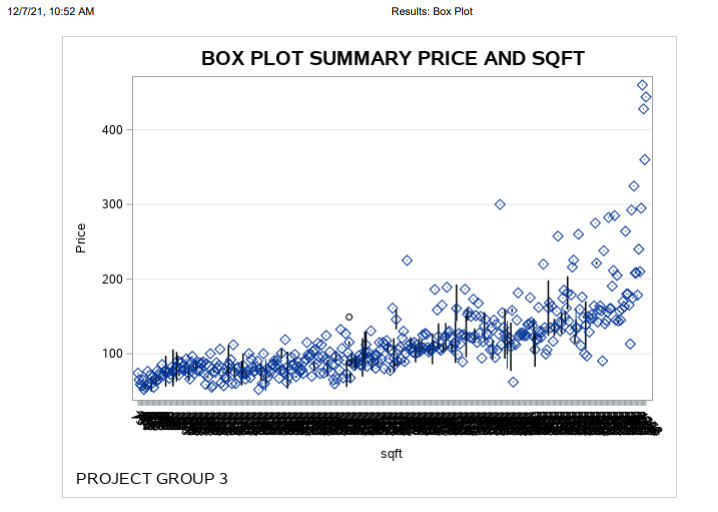
**Distribution/Outliers :**

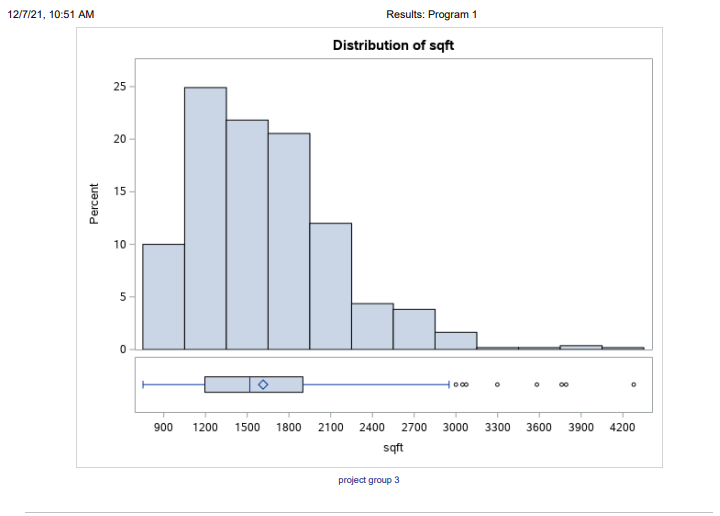
In the same way, we are using to check the outliers using box-whisker plot and box plot for the fields visual representation. Check below for the visual representation.

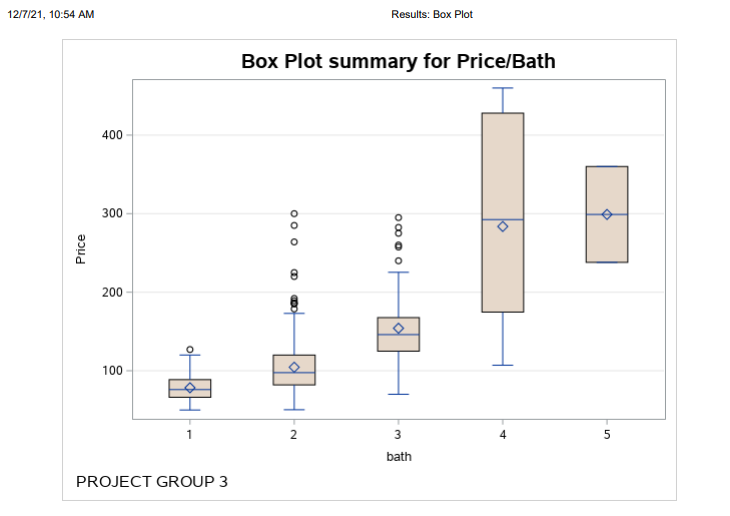


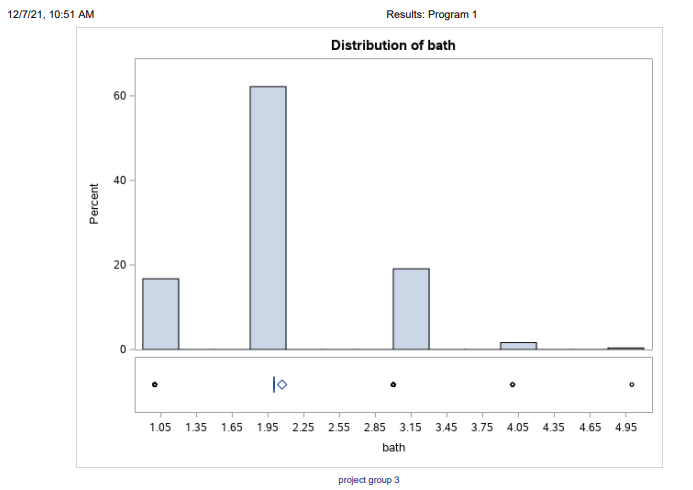
****

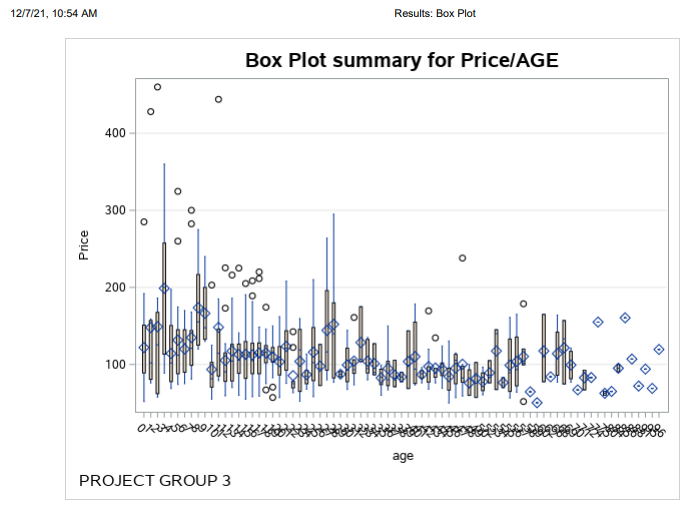


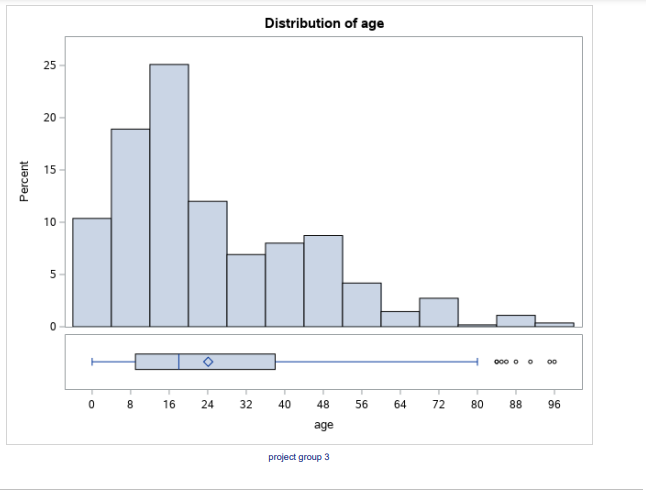


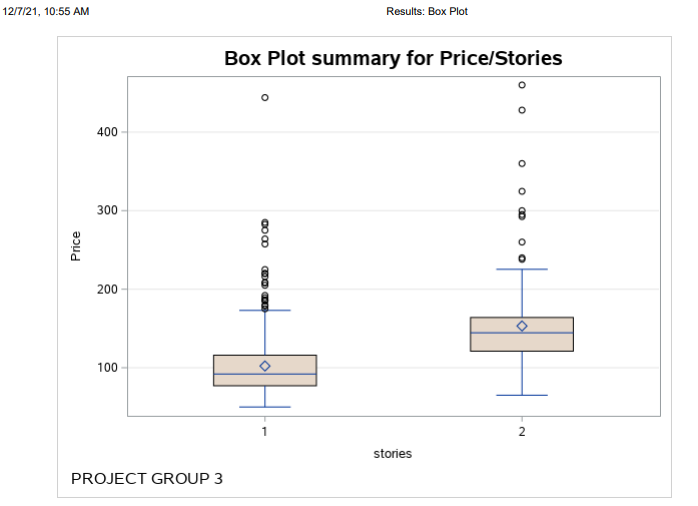


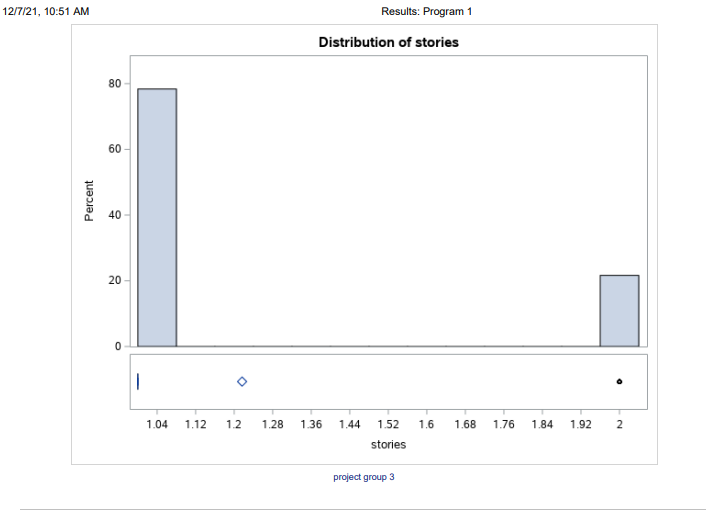


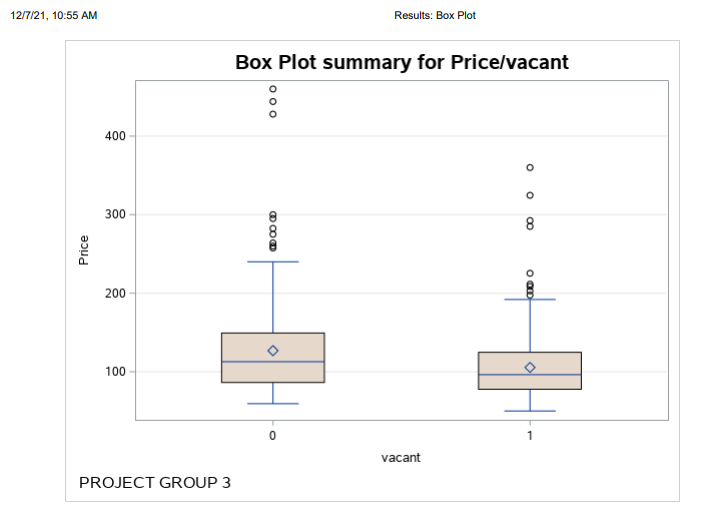


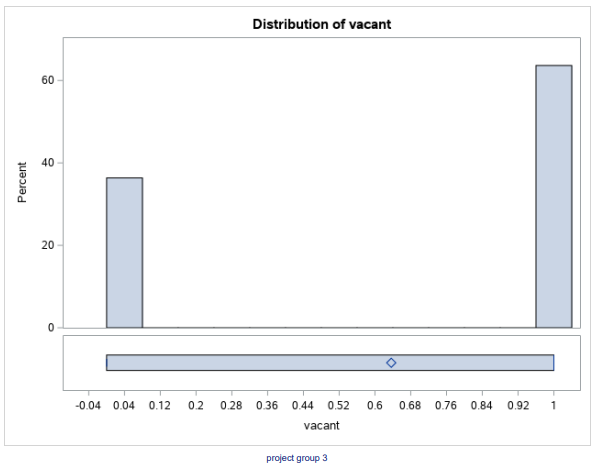












**Linear Regression :**

For the Linear regression we performed in SAAS.

Linear Regression was used to address our research hypothesis, and SAS was used to compute the results.

Linear regression analysis is a statistical approach for predicting the value of one variable based on another's value. The variable you want to forecast is known as the dependent variable. The independent variable is the one you use to predict the value of the dependent variable.

Business and organizational leaders may be able to make better decisions by employing linear regression techniques. Organizations amass enormous amounts of data, which linear regression helps them manage more successfully.

Graphical user interface, text, application, Word

Description automatically generated

Table

Description automatically generated

Graphical user interface, application, table

Description automatically generated

Chart, scatter chart

Description automatically generated

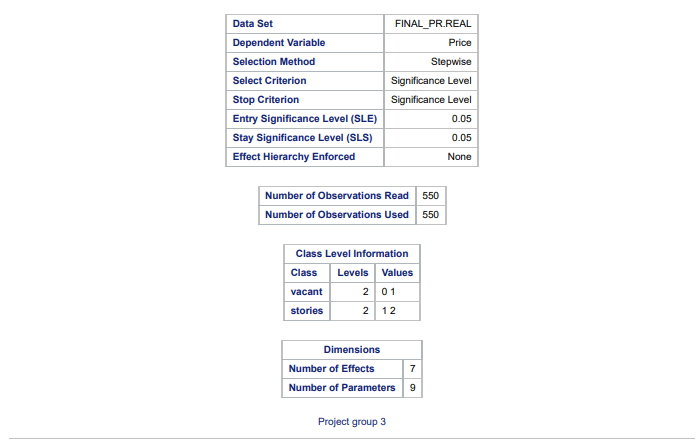
Diagram, engineering drawing

Description automatically generated

A picture containing graphical user interface

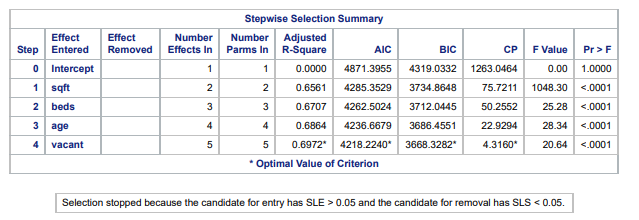
Description automatically generated

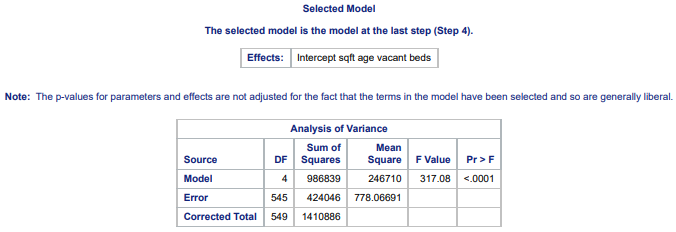
Second:

**** In evaluation of our linear regression outputs, we receive the following data.

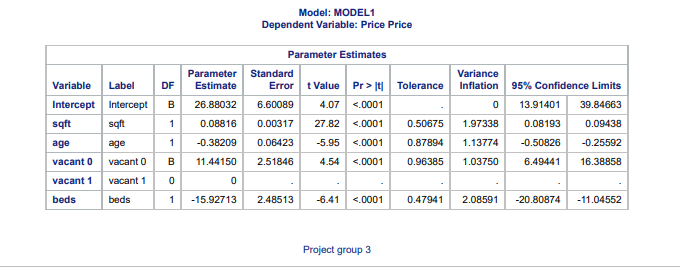
|  |
| --- |
| SqFt = 0.088  Beds = -15.927  Age = -0.382  Vacant = -11.441  House Price = .88  And we note that our p-values are <.05  Thus the regression equation is **[House Price = 088 – 15.927 -0.382 – 11.44].**  **R square** value is : |

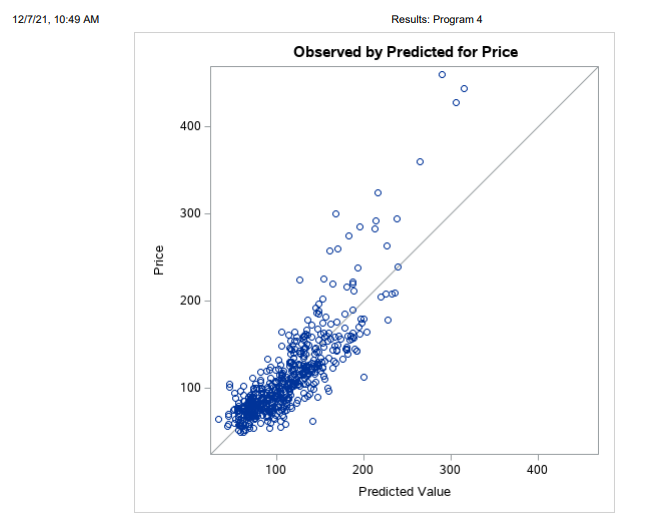
Stepwise Selection Summary which shows an effect being entered and removed in each step along with each model’s Adj R-square, AIC, BIC, Mallow’s Cp, F statistic and P-value.

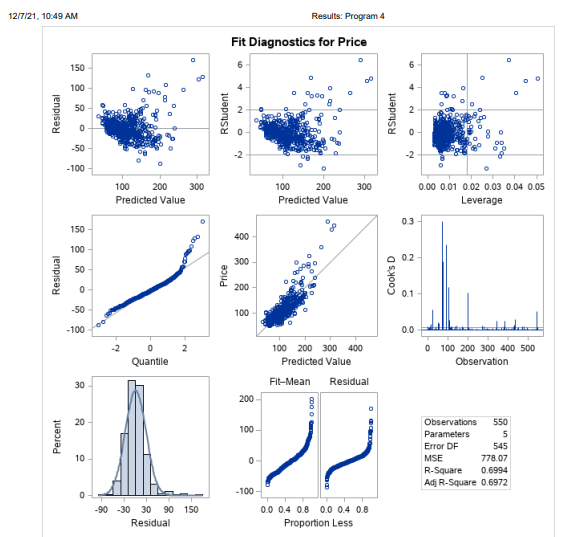
****

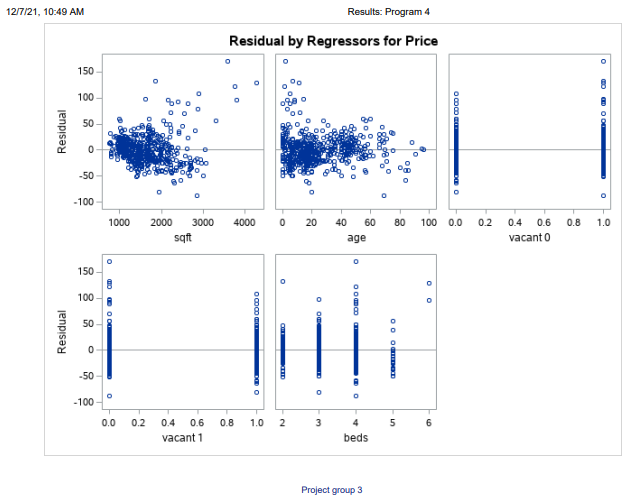
****

****

****

We might conclude that the normality assumption is untrue since the distribution is not normal or symmetric. The Q-Q graphic backs this up. We investigate the following in addition to our Box-Chart charting for outlier detection:****

****

****

**Pearson correlation:**

**Table

Description automatically generated**

**A picture containing table

Description automatically generated**

**Conclusion :**

House prices are not influenced by elements such as the number of bedrooms, bathrooms, square footage, or whether or not the house is vacant, we can reject the null hypothesis based on our project research in the above-mentioned scenario.

We may deduce from our data that the price of a home can be predicted by considering characteristics such as the number of bedrooms, baths, square footage, number of stories, and the age of the property. This information will assist our customer in making informed strategic decisions and sensible investments.

* **Project -3**