Jeff Registre

Mat 258

Row #2

Lab #5: Multi Reg

We wish to analyze if the price of a house can be predicted by the following variables; number of bedrooms, bathrooms, acres and square footage.

**Correlation:** Below is a correlation matrix for the given data.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Acres | Bedrooms | Bathrooms | Sq.ft |
| Bedrooms | 0.669 | X | X | X |
| Bathrooms | 0.551 | 0.720 | X | X |
| Sq.ft | 0.806 | 0.842 | 0.868 | X |
| Price | 0.857 | 0.687 | 0.757 | 0.930 |

According to the data in the matrix, we definitely do have to worry about multicollinearity. The R values are most very high for example between Sq.ft and price (0.930) or Bathrooms and Sq.ft.

**Hypothesis Testing:**

1. Ho: B1 = B2 = B3 = B4 = 0

Ha: Bj ≠ 0 for at least one j.

1. F-Value = 47.93
2. P-Value ≈ 0
3. P < 0.05.

Therefore we reject Ho and conclude that at least of the coefficients is significantly different from zero therefore a regression model exists.

Slope Hypothesis testing

Acres

1. Ho: B1 = 0

~~Ha: B~~~~1~~ ~~≠ 0~~

1. P > 0.05. Therefore we fail to reject Ho and conclude that the slope of Acres is not significantly different from zero and is therefore not a good predictor.

Bedrooms

1. ~~Ho: B~~~~2~~ ~~= 0~~

Ha: B2 ≠ 0

1. P < 0.05. Thus we reject Ho and conclude that the slope of Bedrooms is significantly different from zero and is therefore a good predictor.

Bathrooms

1. Ho: B3 = 0

~~Ha: B~~~~3~~ ~~≠ 0~~

1. P > 0.05. Therefore we fail to reject Ho and conclude that the slope of Bathrooms is not significantly different from zero and is therefore not a good predictor.

Sq.ft

1. ~~Ho: B~~~~4~~ ~~= 0~~

Ha: B4 ≠ 0

1. P < 0.05. Thus we reject Ho and conclude that the slope of Sq.ft is significantly different from zero and is therefore a good predictor.

**Removing a variable**

Since Bathrooms has the highest P value, we’ll drop it and recalculate and conduct a new hypothesis test.

1. Ho: B1 = B2 = B3 = B4 = 0

Ha: Bj ≠ 0 for at least one j.

1. F-Value = 47.93
2. P-Value ≈ 0
3. P < 0.05. Therefore we reject Ho and conclude that at least of the coefficients is significantly different from zero therefore a regression model exists.

**Equation of line**

Below is a linear equation that can be used to predict prices.

Price ($1,000s) = 56.2 + 267 Acres - 77.0 Bedrooms + 0.2122 Sq. ft.

This equation has a y value of 56.2 when all variables are zero. The equations denotes that; for every 1 point increase in Acres price, increases by 267, for every 1 point increase in Bedrooms, price decreases by 77 and lastly for every 1 point increase in sq.ft, price increases by 0.2122.