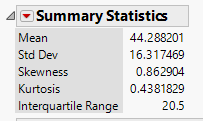
**NAME:KUSHAL SHAH**

**STUDENT ID:A20207420**

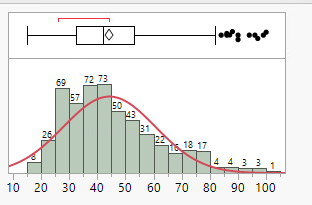
**EMAIL ID:kusshah@okstate.edu**

**Q1)B)**

****

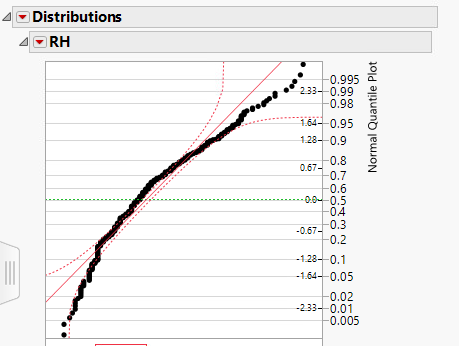
Above figure shows Mean, Std deviation, Skewness, kurtosis, and interquartile range for RH variable.

**Q1C)**

****

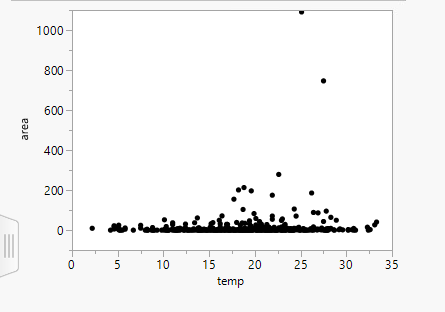
Above figure shows histogram for RH with normal curve overlaid on histogram. From the histogram, we can conclude that distribution of RH deviate from standard normal curve. Shape of the curve suggest that distribution is skewed on right hand side with median value less than mean value.

**Q1D)**

****

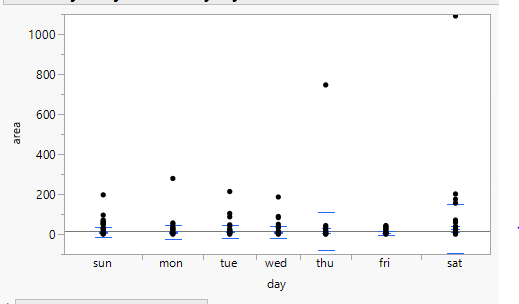
Here, normal prabability plot for RH distribution has been shown in the figure. Streight red line suggest normal distribution and curve made by joining black descrete points is actual distribution of RH variable with percentile scale from 0 to 1. Green line suggest median value. Here,curve made by joining black dotes deviates from red streight line.Distribution above and below green line in normal probability plot suggest distribution to be right skewed due to inverted c shape of distribution and deviation from red streight line

**Q2)A)**

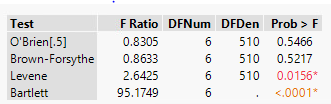
****

Above shown figure is graph between two quantitative variables Area(Y) and temperature(X). Here, Area is Response variable and Temperature is Explanatory variable.It is hard to say that there is any relationship between two variables(the area of forest burnt with temperature) by means of scatter plot. Observation of data indicates that most of the records is concentrated within specific range of Area variable(0ha-75ha) where temperature varies without following any specific trend. There are less points with Area>100 ha where we can identify no such trend between Area and Temperature.

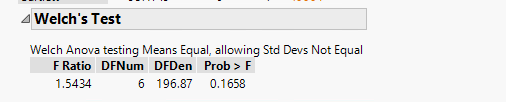
**Q2)B**

****

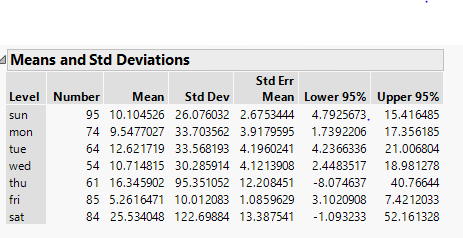
Here, Graph between Area(Quantitative Response variable) and Day(Categorical Explanatory variable) has been shown above by grouped means.(grouped by days.)

****

Above table shows test for equal variance considered. Two test shows f Ratio near 1 and high prob suggesting less variance between groups. 3rd and 4th test shows opposite results. Therefore it could be more controvercial to comment on variance between groups.

****

If we assume that variance between groups is not equal,(which implies that standard deviation is also not equal), then according to welch test, it suggest that means of each group are equal(Assuming significance level of 0.05). which implies that the mean value of burned area of forest on each day(category wise) can be considered as equal. Which further suggest relationship between Area and Days assuming variance is not equal for each group.

****

Observation of data from above table shows more deviation in the observations of area of forest burnt on Thursday and Saturday with respect to mean value of total forest area burnt on respective day. On Friday,there is least deviation with repsect to mean value and On Sunday,Monday,Tuesday,Wednesday there is somewhat same amount of deviation from respective mean value regarding area of forest burnt. Saturday has highest mean value of area of forest burnt wherein Friday has the least.