Remy Lagrois

Final Section 402

1. Using PROC POWER and the given parameters a sample size of 7 would be needed.
2. a)

b) The men were randomly assigned to each treatment group which means the different weights of the men should be pretty evenly distributed between the groups. One group won’t have mostly really heavy men while the other has mostly less heavy men, the distribution of weight within the groups should look very similar to each other.

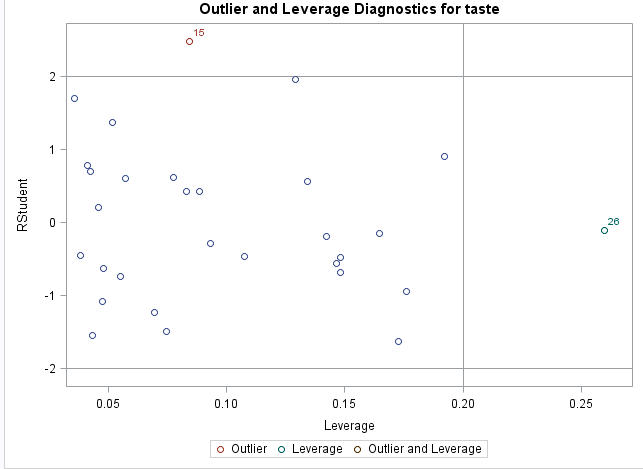
2. There are no situations where you can use your data to assess lack of model fit to the data.
3. **As long as the sample size is large enough, the F-Test for the equality of variance is robust to normality.**
4. **Adding more variables to the model will always increase the R2.**
5. An r = 0 indicates that that there is no relationship between x and y.
6. A causal relationship can be concluded if the p-value associated with the test is <0.001.
7. Type 1 error can be controlled by the researcher while type 2 error cannot.
8. **The coefficient of a categorical variable with two levels represents the difference between the group means.**
9. Paul Minton was the first chair for the Statistics Department. SMU alumni include Lamar Smith who currently serves in the House of Representatives and the award winning actress Kathy Bates.
10. a) The response variable is taste

b) taste = -26.94 + 3.80(acetic) + 5.15(H2S)

c) Yes the overall regression equation is significant, the F value of 18.81 gives a p-value of <.0001 so we reject the null hypothesis that all slopes are zero.

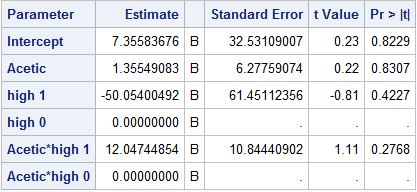
d) The slope for H2S is significant but the acetic acid slope is not. The H2S slope has a p-value of 0.0002 while the acetic slope’s is 0.4062. Therefore we can reject the null hypothesis that the slope is zero for H2S but not for acetic.

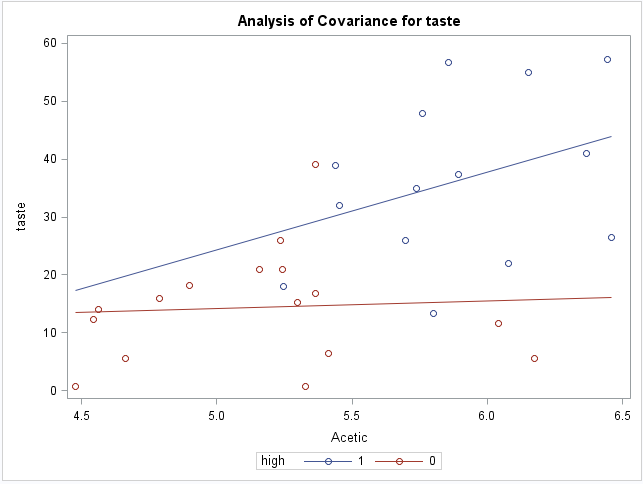
e)The slope for H2S is 5.15 with a 95% CI of 2.66-7.62. Since the slope H2S is positive the taste score will increase as the concentration of H2S increases; for each 10% increase in the concentration of H2S the taste score should increase by 0.491 points. Acetic acid’s slope of 3.80 is also positive, however it’s 95% CI ranges from -5.44 to 13.045 and the p-value indicates it isn’t significantly different from zero. However if we assume the 3.80 number is right then increasing the acetic acid concentration also increases the taste score; a 10% increase in acetic acid concentration increases the taste score by .362 points.

f)

Looking at the studentized residuals plotted against leverage we can see that case 26 has high leverage which means its predictors have a value much different than the average value for the dataset. However it also has a residual of nearly zero and a low Cook’s D which indicates it doesn’t have undue influence on the model.

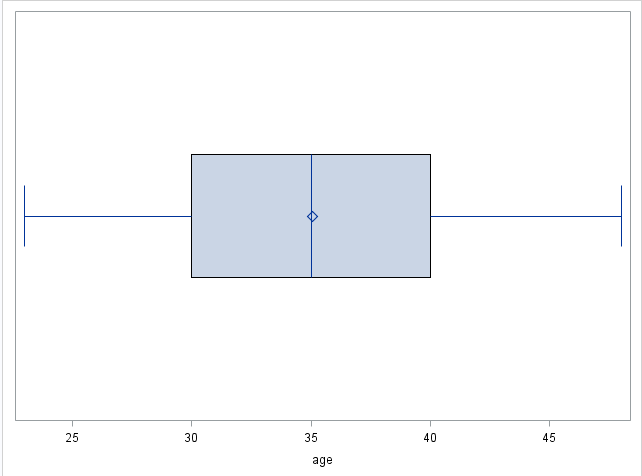
g) The predicted taste score would be 28.767 and the 95% PI is 17.885 to 39.649

h)Yes, the acetic acid slope with high H2S is much larger than without



1. a) A rank of 9.5 means that the values of the 9th and 10th places of the ordered data were the same (in this case the salary for the NYY left and center fielders which is 13). When that happens, the score assigned to the ranks with equal values is the average of those ranks. So, instead of one 13 having a score of 9 and the other 10, they both get 9.5.

b)

1.  This is a box plot of the age of the salesmen. It shows that the mean and median are both around 35, the minimum is just under 25 and the max is just under 50. Half of them (ie 50%) are between the ages of 30 and 40; it appears the remaining 50% are evenly split between younger than 30 and older than 40. The ages appear to be normally distributed.

For this data a boxplot or a histogram would make sense. I chose the boxplot because you can still get an idea of the distribution while also getting more exact information as to the mean, median, min, and max (as well as 1st and 3rd quartiles). The histogram would give a better idea of the relative frequency of each age or age group but I don’t think that’s worth sacrificing the information gained from a boxplot in this case.