One-Way ANOVA test assumptions

- Random samples of three or more groups, each measuring the same quantitative variable.
- Groups are independent of each other.
- Check if all the samples are at least 30 or nearly normal.
- Populations have same variance. Check if the largest sample standard deviation is less than or equal to twice the smallest standard deviation.

ANOVA Test Statistic \( F = \frac{\text{Variance Between the groups}}{\text{Variance Within the groups}} \)

Directions: Copy and paste the ANOVA test data sets into Statcrunch and answer the following questions.

1. Random samples of black bears were weighed at various times of the year. Some of the bears were weighed in April through July. Others were weighed in August and September or October and November.
   a) Create a side by side box plot for the three data sets. Draw the boxplot below and describe the graph. Also find the mean average, standard deviation and variance for each of the three data sets. How do they compare.

   b) Does the data meet the assumptions necessary to do an ANOVA test? (If the data set is small, be sure to check the nearly normal assumption by making a histogram.)

   c) Use an ANOVA test to test the claim that the average weight of black bears is different depending on what time of year they are measured. (Can you think of a reason why this might be true?) Give the null and alternative hypothesis, the F-test statistic and the P-value. Write a sentence describing the meaning of the F-test statistic and another sentence describing the meaning of the P-value. Give the degrees of freedom between the groups and the degrees of freedom within the groups. Did you reject the null hypothesis or fail to reject? Write a conclusion for your test.

   d) Follow up questions:

      i.) Was the variance between the groups significantly higher than the variance within the groups? (Explain how you know.)

      ii.) Is it likely or unlikely that the sample data occured by random chance from groups that have the same mean average? (Explain how you know.)
2. Now we are going to look at the relationships between how much sleep Math 075 students get and how many units they have completed at COC. Since the Math 075 data is census data we can assume it is representative of the population of all 075 students. Julie thinks that the average number of units will be the same no matter how much sleep the person gets. Analyze the data for Julie. The number of units have been broken up into three data sets (less than 6 hours, 6-8 hours, more than 8 hours).

a) Create a side by side box plot for the three data sets. Draw the boxplot below and describe the graph. Also find the mean average and variance for each of the three data sets. How do they compare.

b) Does the data meet the assumptions necessary to do an ANOVA test? (If the data set is small, be sure to check the nearly normal assumption by making a histogram.)

c) Use an ANOVA test to test the claim that the average number of units completed at COC is the same regardless of how much sleep a person gets. Give the null and alternative hypothesis, the F-test statistic and the P-value. Write a sentence describing the meaning of the F-test statistic and another sentence describing the meaning of the P-value. Give the degrees of freedom between the groups and the degrees of freedom within the groups. Did you reject the null hypothesis or fail to reject? Write a conclusion for your test.

d) Follow up questions:
   i.) Was the variance between the groups significantly higher than the variance within the groups? (Explain how you know.)

   ii.) Is it likely or unlikely that the sample data occurred by random chance from groups that have the same mean average? (Explain how you know.)
3. Let’s look again at the Math 075 data and explore the relationship between political party and how much alcohol someone drinks. Since the Math 075 data is census data we can assume it is representative of the population of all 075 students. The amount of alcohol drunk has been separated into four data sets corresponding to four political affiliations (democrat, republican, independent, other). Is the average amount of alcohol drunk different depending on political affiliation?

a) Create a side by side box plot for the three data sets. Draw the boxplot below and describe the graph. Also find the mean average and variance for each of the three data sets. How do they compare.

b) Does the data meet the assumptions necessary to do an ANOVA test? (If the data set is small, be sure to check the nearly normal assumption by making a histogram.)

c) Use an ANOVA test to test the claim that the average amount of alcohol a person drinks differs depending on political affiliation. Give the null and alternative hypothesis, the F-test statistic and the P-value. Write a sentence describing the meaning of the F-test statistic and another sentence describing the meaning of the P-value. Give the degrees of freedom between the groups and the degrees of freedom within the groups. Did you reject the null hypothesis or fail to reject? Write a conclusion for your test.

d) Follow up questions:

i.) Was the variance between the groups significantly higher than the variance within the groups? (Explain how you know.)

ii.) Is it likely or unlikely that the sample data occurred by random chance from groups that have the same mean average? (Explain how you know.)
4. Let’s look at the math 140 survey from fall 2015. We are analyzing the amount of minutes per week spent on social media. We separated the data into the type of social media being used. Test the claim that the amount of minutes spent is the same no matter what social media was being used.
   a. Create a side by side box plot for the three data sets. Draw the boxplot below and describe the graph. Also find the mean average and variance for each of the three data sets. How do they compare.
   b. Does the data meet the assumptions necessary to do an ANOVA test? (If the data set is small, be sure to check the nearly normal assumption by making a histogram.)
   c. Use an ANOVA test to test the claim that the amount of minutes spent is the same no matter what social media was being used. Give the null and alternative hypothesis, the F-test statistic and the P-value. Write a sentence describing the meaning of the F-test statistic and another sentence describing the meaning of the P-value. Give the degrees of freedom between the groups and the degrees of freedom within the groups. Did you reject the null hypothesis or fail to reject? Write a conclusion for your test.
   d. Follow up questions:
      i.) Was the variance between the groups significantly higher than the variance within the groups? (Explain how you know.)
      ii.) Is it likely or unlikely that the sample data occurred by random chance from groups that have the same mean average? (Explain how you know.)