MATH 140 **Name** _____ **Date:**
Worksheet: ANOVA

The Scene: Suppose we are interested in comparing three or more population means. We wish to test these hypotheses:

H_o: All population means are equal

H_a: Not all population means are equal

ANOVA (which is short for analysis of variance) is a common method for doing inference on several means. The point of this activity is to get to know ANOVA by looking at a particular problem.

In this worksheet we consider the following **Research Question**

> To examine the effect of pets in stressful situations, researchers recruited 45 people who identified themselves as dog lovers. The subjects were randomly assigned to three groups to do a stressful task alone, with a good friend present, or with their dog present. Each group had 15 subjects. The subject's mean heart rate during the task is one measure of the effect of stress.

---

1. Import the data into an RStudio session and answer these questions:
   (a) How many observations are there? Call this value N. $N =$ _____
   (b) How many treatments are there? Call this value k. $k =$ _____
   (c) Fill out the table below, giving the name of each of the treatments, the sample size within each group, the sample mean heart rate in each group, and the sample standard deviation of the heart rates in each group.

treatment	sample size	sample mean	sample st. dev.

2. Construct side-by-side boxplots for heart rate during the task. Considering this plot and the summary statistics for each group above, are the ANOVA assumptions met here? Explain.

3. Regardless of your answer to the previous question, run the ANOVA test in RStudio and answer these questions:

   (a) What are the numerator degrees of freedom $(k-1)$?

   (b) What are the denominator degrees of freedom $(N-k)$?

   (c) What is the value of the F test statistic?

   (d) According to the anova print out, what is the p-value for the test?

   (e) run '1-pf(F,num df, denom df)' to confirm that this gives the p-value that appears in the print out.

4. State your conclusion in the context of the actual experiment. Do these data provide significant evidence that the choice of companion influences stress levels?