**Statistician: Ryan Welch Version: B**

**Directions**

The final exam will consist of several application-type questions related to the following topics we’ve covered this semester – univariate EDA (quantitative & categorical), bivariate EDA (quantitative & categorical), linear regression, one-sample t-test, two-sample t-test, and chi-square. On the final exam, you will be asked to answer each question from results that you have prepared prior to the exam using R.

The data that you will examine is introduced below, including actual questions that will be on the exam. You should load these data into R and create output that can be used to answer each question. Your R input and output should be printed and brought to the exam to be used to answer the exam questions.

The R output document that you bring to the exam must meet the following strict criteria:

* The document must be produced by you and you alone!! You may not ask anyone but me (including the tutors) for specific help on using R. If you have questions, I will be available in my office or via e-mail.
* The document can only contain R commands related to expressions, assignments, functions, or objects; R output; or R created graphics. You may not type or write ***any other*** material on the document (including labeling figures, tables, output, or sections). You may not type any “notes” (i.e., non-R-related expressions, assignments, functions or objects) as “R code.” You cannot use R comments. The document should contain no code that results in errors.

Failure to follow all of these criteria will result in a 0 for the final exam!!

**Data Set -- BatMorph**

The Hawaiian hoary bat, *Lasiurus cinereus semotus*, is an endemic subspecies of the North American hoary bat; the mainland subspecies is *L. c. cinereus*. Researchers in Hawaii and on the mainland recorded a variety of morphological (body) characteristics on bats collected from three different types of habitat. The data are in [**Batmorph.csv**](https://raw.githubusercontent.com/droglenc/NCData/master/Batmorph.csv) and information about the data are in the [**Batmorph\_meta.txt**](https://github.com/droglenc/NCData/blob/master/Batmorph_meta.txt) file. You should prepare results for each of these items …

1. Univariate EDA for bodymass, skull length, wingspan, and habitat.

2. Bivariate EDA for **all pairs** of quantitative variables. [*This may be done with one graph and one table.]*

3. Bivariate EDA for **all pairs** of categorical variables.

4. Linear regression results (equation results and r2) for predicting the height of the coronoid process from the height of the canine tooth.

5. Results for testing the following research hypotheses (use 5% level for each)

a. The mean height of the canine tooth is less than 0.33 cm.

b. The mean height of the canine tooth is different between the two subspecies of bats.

c. The distribution of individuals into the three habitats differs between the two subspecies.

d. The mean wingspan differs between bats captured in habitat A and those captured in habitat B.

e. The mean body mass of the *L. c. cinereus* subspecies is greater than 35 g.

1. Describe the importance of statistics (as a field of study or a collection of methods). Among other things make sure you describe the two major goals of statistics, identify at least three major concepts or ideas of statistics, and identify how some of the “tools” you have learned this semester illustrate or are related to why you think statistics is important. [*This question is general and is not specific to these data.*]