**Statistician: Mason Deja Version: R**

A Master’s student collected a variety of morphological (body characteristics) data on the invasive Ruffe (*Gymnocephalus cernuus*). The data are in [**RuffeMorph.csv**](https://raw.githubusercontent.com/droglenc/NCData/master/RuffeMorph.csv) and information about the data are in the [**RuffeMorph\_meta.txt**](https://github.com/droglenc/NCData/blob/master/RuffeMorph_meta.txt) file. Code for “cleaning” the data are given in the metadata (**make sure to follow this code**!). Once the data have been “cleaned” the ***miss*** and ***out*** variables should be ignored. Use these data to complete the following tasks.

1. Perform an EDA for total length

2. Perform an EDA for sex.

3. Perform an EDA for the relationship between total length and weight of Ruffe.

4. Perform an EDA for location of collection and sex.

5. For the linear regression to predict body girth from total length …

a. Interpret the slope of the linear regression that you performed.

b. Predict the body girth of a Ruffe whose total length equals the median total length.

c. What proportion of the total variability in body girth is explained by knowing the total length?

6. Perform a complete appropriate hypothesis test for the following research hypotheses (at the 5% level)

a. The mean weight is different for Ruffe from Allouez then for Ruffe from Whaleback.

b. The distribution of Ruffe into the sexes differs among the three locations.

c. The mean upper jaw length of female Ruffe is less than 11 mm.

7. Describe the importance of statistics (as a field of study or a collection of methods). Make sure you **DESCRIBE** the two major goals of statistics, at least three major concepts or ideas of statistics, and how “tools” that you 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. An answer with <500 words is probably inadequate.*]