Homework 1

Due: Tuesday February 16 at 11:59pm Champaign (US Central) time

See general homework tips and submit your files via the course website.

For all exercises, use the **skulls** data set defined in **HW1Data.sas** in the **Homework 1** folder on the course website. The data set is based on the **skulls.dat** file from the data files accompanying our textbook *A Handbook of Statistical Analyses Using SAS, Third Edition*. The original Egyptian skulls data is described on page 98 of that textbook. Measurements are assumed to be in mm. The variables in the **skulls** data set are:

- **mb**: Maximum breadth
- **bh**: Basibregmatic height
- **bl**: Basialiveolar length
- **nh**: nasal height
- **epoch**: epoch defined in the book, with 1 being earliest and 5 being latest

Exercise 1:
a) Obtain basic location, spread, and skew statistics for maximum breadth (**mb**) for all the data. Comment on typical values, skew, and spread or range for maximum breadth for Egyptian skulls based on the sample.
b) Repeat the analysis in part **a) by epoch** for epochs **1, 3, and 5**. Comment on what this tells us about similarities and differences of maximal breadths across the first, middle, and last time periods in the data.

Exercise 2:
a) For maximum breadth (**mb**), visually (with graphics) and quantitatively (with appropriate hypothesis tests) check if an assumption of normality would be reasonable and state your conclusion.
b) Repeat the analysis in part **a) by epoch** for epochs **1 and 3**. Based on the results, comment on how reasonable an assumption of normality in maximum breadth is for epoch 1 and how reasonable that assumption would be for epoch 3.

Exercise 3:
a) A claim is made that a typical (mean or median) modern day maximum breadth (**mb**) is about 136mm. Test whether the typical (mean or median) maximum breadth for the population this sample came from is significantly less than 136mm. Consider your test for normality in **Exercise 2** to choose the appropriate location test. State your conclusions.
b) Perform a hypothesis test of whether **epoch 1** tended to have significantly greater maximum breadths **epoch 3**, and state your conclusions. Again, consider the hypothesis tests in **Exercise 2** when choosing your test for difference in maximum breadth.

Exercise 4:
a) Perform a correlation analysis for all of the data. State what this tells us about relationships between the four skull measurements in the data sampled, and what we might infer about relationships in the population it was sampled from.
b) Perform the same correlation analysis by epoch for the earliest (**epoch=1**) and latest (**epoch=5**) time periods in the data. Comment on how the relationships between characteristics differ between the earliest and latest epoch, and note any differences with what you found in part **a)** for the combined data.