**Assignment given to Students**

The objective of the study for which you will analyze the data was to determine if significant gender differences existed between subjects 65 years of age and older with regard to calcium, phosphorous, and alkaline phosphatase levels (Boyd et al., 1998). The researchers performed a retrospective chart review of laboratory procedures performed in 6 different physician practices. The data consisted of 178 subjects representing 92 males and 86 females age 65 or older. In the data set, there are three discrete variables, **sex**, **lab**, and **agegroup**. The coding is as follows:

|  |  |
| --- | --- |
| Sex | 1=Male; 2=Female |
| Lab | 1=Metpath; 2=Deyor; 3=St. Elizabeth's; 4=CB Rouche; 5=YOH; 6=Horizon |
| Agegroup | 1=65-69; 2=70-74; 3=75-79; 4=80-84; 5=85-89 years |

The other variables of **age**, **alkphos** - alkaline phosphatase (IU/L), **cammol** - calcium (mmol/L), and **phosmmol** – inorganic phosphorus (mmol/L), are continuous.

1. The first task of the assignment is to check the validity of the data. Determine if this is a "messy" data set with variable values that appear incorrect. Attempt to recover the correct values by looking up the true values from the actual data records. Copies of these can be found on http://academic.csuohio.edu/holcombj/clean/bigtable.htm

2. Once the data is "clean", perform a summary analysis of the three discrete variables (**sex, lab**, and **agegroup**). For the variables **alkphos, cammol** and **phosmmol**, report the mean, median, standard deviation, min and max broken down by **sex**. Also summarize the variables **alkphos, cammol** and **phosmmol** in a similar way with the factor variable as **lab**.

3. Then Construct side by side box plots of the variables **alkphos, cammol,** and **phosmmol** with the factor variable as **sex**. Then construct side by side box plots of the **alkphos, cammol,** and **phosmmol** continuous variables with the factor variable as **lab**.

4. Do you believe a significant difference exists in **alkphos, cammol**, or **phosmmol** levels with respect to **sex**? Why or why not? Do you believe a significant difference exists in **alkphos, cammol**, or **phosmmol** levels with respect to **lab**? Why or why not?

5. Suppose Mr. and Mrs. Contrarian are married and Mrs. Contrarian has lower calcium than Mr. Contrarian. She refuses to believe the results of the study that men tend to have lower calcium than women because she has lower calcium than her husband. Using your results to question #3, explain to Mrs. Contrarian the flaw in her thinking.

6. One of the objectives of this research was to propose a reference range of values that are to be considered “normal” for calcium, inorganic phosphorus, and alkaline phosphatase. Looking at the results for **cammol** alone for each of the labs, explain why a single reference range is so difficult to establish.