**BIOST 514/517**

**Discussion Section: Week 3**

**Introduction to R**

**R Exercises**

**Exercises:**

1. The crab.txt dataset will be used for this question. The data are available from canvas or from web: <http://faculty.washington.edu/tathornt/Biost509/DataSets/crab.txt>:  
   1. Provide a plot of the number of satellites versus width
   2. Provide a plot of the number of satellites versus weight, and color-code the points differently if their width exceeds 26cm
   3. Illustrate the distributions of number of satellites, for crabs of different colors. (There are several ways to do this)
2. The salary.txt dataset will be used for this question. The data are available from canvas or from the web: <http://faculty.washington.edu/tathornt/Biost509/DataSets/salary.txt>;
   1. Provide boxplots of salary by gender. Briefly describe any differences in the average or median distribution of salary by gender.
   2. Provide boxplots of salary by rank. Briefly describe any differences in the average or median distribution of salaries by rank.
   3. Now provide boxplots of salary for each combination of rank and gender using the *boxplot()* function. [**Hint: Use the *boxplot()* function with the interaction() function for rank and gender: boxplot(salary~interaction(gender,rank),data=salary) ].** Does it appear as if the average or median salary distribution is significantly different between males and females within each rank? Briefly explain.

* 1. Redo question, but use the *varwidth* option with the boxplot() function. [**Hint: use the help function for boxplot, i.e., ?boxplot, to learn what the varwidth option does**.]