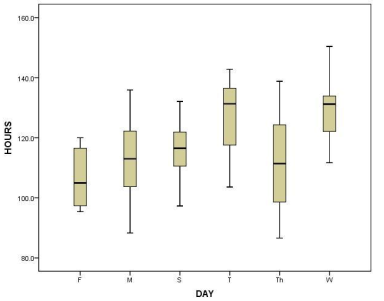
**HW5**

**Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_**

let’s use the “clerical\_Q2.txt” data.

A store manager noticed that the busiest days for clerical staff are Wednesdays and Tuesdays. See enclosed box plot. The manage tries to compare the group means in hours by different days



a). [10] Observe the box plot. Can you confirm that the hours in Tuesday is the highest? Why?

b). [20] Write down your hypothesis in the ANOVA to compare the group means in hours by different days

b).[30]Using R to build the ANOVA regression model, and help the manager to make the decision whether the group means in hours or different days are the same or not.

c). [20] Try to interpret the coefficients you got in the ANOVA regression model from part b).

d). [20] practice for data preprocessing: create N-1 dummy variables for the variable ‘DAY’. Convert the variable “mail” to nominal variable by creating 4 groups. Again, past the codes and snapshots