STAT 217 Project 1: Due Monday, September 22, 2014

Write up the results using a word processor and include appropriate plots and output in your report. Part of the grade is based on the write-up, so you need to explain your results as if you are talking to a classmate who knows some statistics but does not know anything about computer output. Your final write up should **READ LIKE A REPORT! Refer to the guidelines posted on D2L to carry out a full analysis and *communicate your results clearly!***

Dataset: Dr. Spock Trial

(E.8 (pg 57) Practicing Statistics, Kupier) Dr. Benjamin Spock was a well-known pediatrician who faced trial in 1968. He was charged with conspiring to violate the Selective Service Act by encouraging young men to resist the draft. As part of his defense, his counsel claimed that women were underrepresented on the jury. Women tend to be more sympathetic toward war protesters than men do. The defense counsel claimed that the judge had a history of choosing juries on which women were systematically underrepresented. At that time, jury members in Boston were chosen from a venire (a group of 30 to 200 individuals preselected from the population by the judges clerk). By law, people were supposed to be selected for a venire at random. For Dr. Spock's trial, the judges venire had 100 people and only 9 women, none of whom were selected to be on the actual jury. Dr. Spock's defense counsel collected data on the percentages of women in venires from this judges recent trials (Judge 1) together with those of 6 other judges in the Boston area (Judges 2-7).

Getting the data ready for analysis:
The dataset is available on D2L content under Project 1. Read the data into RStudio using the data import wizard. Notice that judge is numeric right now, but it is the grouping variable so it needs to be categorical. Run the following code to make `judge` a categorical variable.

```
spock$judge <- as.factor(spock$judge)
```

R code for the analysis is provided below. ***It is your responsibility to put the relevant output from these lines of code together into a report that tells a full story.***

```
require(mosaic)
favstats(percent~judge, data = spock)
boxplot(percent~judge, data = spock, ylab = "% Women in Venire")
points(percent~judge, data = spock, col = as.numeric(judge), pch = 20)
fit.spock<-lm(percent~judge, data = spock)
anova(fit.spock)
par(mfrow = c(2,2))
plot(fit.spock)
require(multcomp)
pairs <- glht(fit.spock, linfct=mcp(judge = "Tukey"))
confint(pairs)
plot(pairs)
plotmeans(percent~judge, data=spock, mean.labels = T, digits = 2)
text(c(1:7),c(10,37,37,32,30,30,30),c("a", "b", "b","b", "ab", "b", "b")
 ,col=c("blue",rep("green", 3), "red", "green","green"),cex=2)
```