Stat 587, section 2 – Lab 4 self assessment, part 2

The data in tomato.txt are the data from my small backyard experiment on the effect of a new fertilizer on tomato yield. The two variables are treatment (a = new fertilizer, b = control) and yield (lbs of tomatoes).

Tomato.txt is a text file with spaces between values.

R users: You will need to use read.table() with header=T to read the data correctly.

JMP users: you will need to use ‘best guess’ to read the file. The default setting does not include spaces.

1) Draw a side-by-side box plot of the observations in the two treatment groups. Your answer is the plot.

2) What feature(s) of the box plot suggest that a non-parametric test may be more appropriate than a T-test?

Remember that most software shows unusual values individually in the box plot.

3) Use a Wilcoxon rank-sum test to test the null hypothesis that the two groups have the same median yield. Use an exact p-value calculation, not the normal approximation. Report your p-value.

My answers:

1)



2) Both treatments include an unusually large value.

Also, it really looks like the two groups have different variances

3) p = 0.053