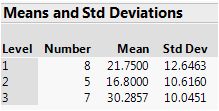
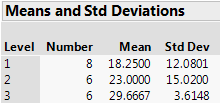
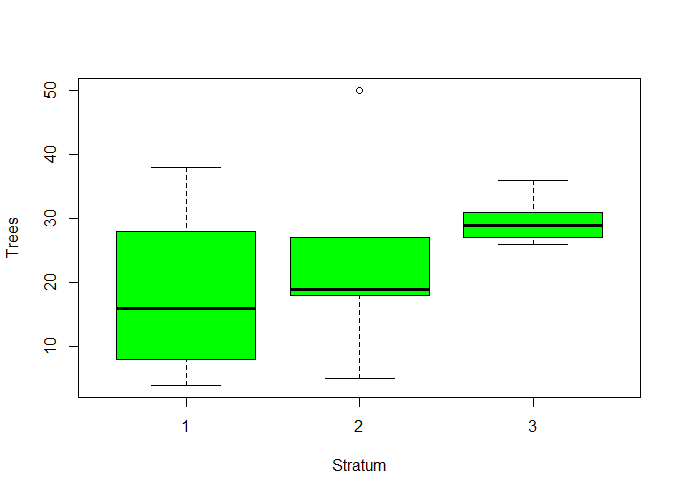
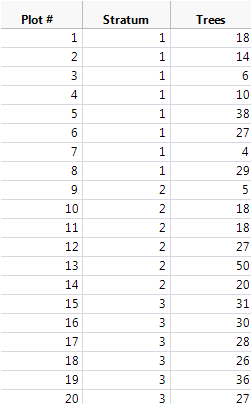
Cody Frisby

STAT 4200 Project 3

 Using the data I collected in project two and the new stratification areas, I added a strata column to my data and looked at the respective variances from each strata. As you can see on the right the data from all three strata are fairly equal but strata one does have the highest variance and thus we will most likely be taking the most samples from this strata. When reexamining my 196 pieces from project two I found a few that were mostly white space so I decided to throw them out. After doing so, N was now equal to 192. After dividing N into the Ni groups based on the provided map, N1 = 59, N2 = 66, and N3 = 67. Using Neyman allocation to determine sample size of each stratum I arrived at 8, 6, and 6 for stratum 1, 2, and 3 respectively. I then took a simple random sample of each of those numbers from their respective strata and I arrived at the samples as indicated on the map at the end of this document. Pictured left is a summary of the means and standard deviations of the trees for each stratum. As you can see, I got very similar variance for stratum 1 as before, larger variance for stratum 2, and much lower variance for stratum 3. One of the main reasons the variance for stratum 2 is so large is that one of the “forests” that are against a few of the buildings on campus was inside one of my samples (see plot 13). I didn’t really know how to treat this so I counted as many of the trees as I could. I ended up with 50 for this plot and this was most likely on the conservative side of the actual number of trees due to this “forest”. Here’s a plot of the data by stratum:

The calculation for the estimate of the total number of trees using stratified sampling is:

Nst = (59 18.25) + (66 23) + (67 29.67) 4582 trees. This is very close to my approximation from project 2, which was 4606 trees. Now, let’s see if the variance is any lower. The calculation for the estimated variance using stratified sampling is: st ) =  212,684.394. Which then gives us a bound on or error of estimation equal to 2 922.354 trees. This is a little better than the first time around where we had an approximate bound of 999 trees. One of the things that I did differently this time around was that I brought my laptop with me to campus when counting the trees. This allowed me to zoom in on the area of the map and I felt that I was better able to identify the boundaries than last time. I was surprised at how close my estimate is to my first time. I felt the first time around I may have been a little too liberal with my plot areas.

