This assignment is intended to help you become more familiar with the various methods of assessing whether or not a given sample may be from a population that follows a particular distribution. Each group has 6 different samples of size $n = 200$ from the following distributions:

A normal distribution.

A chi-square distribution with 4 d.f.

A t distribution with 7 d.f.

A 50%-50% mixture of two normal distributions with means separated by 2 standard deviations.

You will have at least one sample from each of the above four distributions. Each of the samples has a mean of roughly 500 and a standard deviation approximately equal to 100.

Your job is to determine, as best as you can, which samples came from which distributions. Submit a VERY brief report detailing your findings and indicate what statistics, plots or histograms led you to your decision for each sample. (e.g., I believe that sample5 is from the t distribution because a boxplot of the data (figure 2) indicated — . Also, the kurtosis had a value of — indicating —. A Q-Q plot showed that ...)

Any plots that you refer to in your report should be clearly labeled (e.g., figure 1a, 1b, ... , output 2a,2b, ... where the # corresponds to the sample #) and attached at the end of the report. If you are including several plots, histograms, etc. you should place several on the same page (recall that the command *multifig(3,3)* will cause R to place the next 9 figures on the same page). Please do not print one plot per page (if you do, points will be deducted).

You can use at most one page for written description and one page for figures/plots (You should be able to fit 2 sets of 3x3 figures on one page, giving you 18 plots - which should be plenty).

Each group has its own data set named `dist#.txt`, where # is the number corresponding to your group. For example if you were in group 15 you could read the data in with the following R commands:

```
> source("http://www.uvm.edu/~rsingle/stat221/data/scripts-221.R")
> dat <- otherdata("dist15.txt")
```

Now you can access samples x1 through x6 individually using `dat$x1 ...  dat$x6`.

Indicate which distribution you think corresponds to each sample in a table like the one below at the top of your first page, under the names of the people in your group for this project. You might want to look back at "chapter 3 - part 2" of the R-code for class on the webpage.

sample 1	sample 2	sample 3	sample 4	sample 5	sample 6

Good luck!!!