INTRODUCTION TO STATISTICAL REASONING

36-201 **Lab #1 -Partial Solutions**

Q.3 From the histogram for the *normal REM times* (NREML), the distribution looks unimodal and postively skewed. The mode is in the 50–75 minute interval, and we can see that the right tail is heavier (more observations) and longer than the left tail. Also, notice the gap in the right tail. This gap suggests that the two largest observations might be outliers.

Q.6–7 Notice how the scale changes the perception of the shape of the histogram. We still see one primary mode, but it is a little harder to assess shape (skewness). Still the right tail seems a bit heavier and longer than the left tail, so we might still say "skewed right". Note the gap in the left tail which isn't as large as the gap in the right tail.

Q.8 From the histograms it looks as though the median is between 60 and 75, since it looks like about half the observations are above and half the observations are below that class interval. The boxplot confirms this — it looks like the median (horizontal line through the box) is a little higher than 60 minutes.

Q.10 Recall that in the histograms there was a gap in the right tails, suggesting that the two largest observations might be outliers (unusual). The boxplot also identifies these as unusual — as being far away from the mode — by drawing two asterisks (they overlap a little) where these two data points are.

Part A From the histogram it appears that DREML (depressed patients' REM latency times) might have two or even three modes. It is not easy to talk about skewness when there is more than one mode. However, we might say, relative to the first mode this histogram is skewed right, like the ones for NREML.

Q.11 Since the median for the depressed group is about 40 minutes (using the boxplot) and for the normals it is about 60 minutes or so, it seems that the depressed patients have shorter REM latency times, at least with respect to the *median*, a measure of central tendency. Boxplots represent the whole distribution, not just the median, so a more thorough comparison can be made by comparing the boxplots. When we put the boxplots side by side it is clear that the whole distribution for REM latency times is shifted down for depressed patients, relative to the normal subjects.