## STAT 205 Homework 1 Solutions

The plots below are a histogram and a boxplot of the counts of the number of dendritic segments coming out of 36 nerve cells.



We can see from both the histogram and the boxplot that the data are asymmetric, and they are skewed to the right, i.e. there is a long tail of larger values in the distribution. From the histogram the distribution has a single large mode (or unimodal) at just under 30 segments. The boxplot also indicates right skewness (or positive skewness) because the median is closer to Q1 than Q3, and the upper whisker is longer.

From the boxplot there is a large outlier at 57 segments which lies outside the upper fence (Q3+1.5xIQR) and is indicated by a small circle in the boxplot.

Using R the mean is 33.78 segments and the median is 30.5 segments, an so we see that the mean is larger than the median and has been pulled in the direction of the right skewness. The five number summary is (min,Q1,Median,Q3,max) and is given by (20,27,30.5,37.75,57) and so the inter-quartile range (IQR) is 37.75-27 = 10.75, and the range is 57-20 = 37 segments.

NOTE: there are actually many different ways of calculating the quartiles. The function in R has a possible 9 different methods that could be used! The method given in the book for Q1 as the median of the ordered data up to and including the median gives 27, and Q3 is calculated by hand as 38.5. So, although Q3 is different it is quite similar to the default method in R.