Key for Short Problem Set 1

Due at the beginning of our next class. All calculations to be completed and all figures drawn by hand. Be prepared to share your work in class.

Problem 1. Listed below are the number of yellow M&Ms found in seven samples of 1.69-oz bags of plain M&Ms:

$$23\ 15\ 16\ 16\ 18\ 7\ 17$$

Report the mean, the median, the variance, the standard deviation, the interquartile range, and the median absolute deviation for these seven results. You may check your answers using a calculator, but must show all of your work.

To find the mean we sum up the results for the individual samples and divide by the number of samples; thus

$$\text{mean} = \frac{23 + 15 + 16 + 16 + 18 + 7 + 17}{7} = \frac{112}{7} = 16$$

To find the mean we order the samples from smallest-to-largest

$$7\ 15\ 16\ 16\ 17\ 18\ 23$$

and, as the number of samples is odd, report the middle value, or 16.

To find the variance, we sum the squares of the differences between each result and the mean and divide by one less than the number of samples; thus

$$\text{variance} = \frac{(7-16)^2 + (15-16)^2 + (16-16)^2 + (16-16)^2 + (17-17)^2 + (18-16)^2 + (23-16)^2 +}{7-1} = \frac{136}{6} = 22.7$$

To find the standard deviation, we just take the square root of the variance, or 4.76.

To find the interquartile range, we divide the samples into a lower half

$$7\ 15\ 16$$

and an upper half

$$17\ 18\ 23$$

discarding the middle value as the number of samples is odd. The median of the lower half is 15 and the median of the upper half is 18, so the IQR is 3.

To find the median absolute difference we first subtract the median of 16 from each result and take the absolute value of each

$$9\ 1\ 0\ 0\ 1\ 2\ 7$$

order the resulting differences from smallest-to-largest

$$0\ 0\ 1\ 1\ 2\ 7\ 9$$

and report the median, which is 1.

Problem 2. A larger set of results for 30 1.69-oz bags of M&Ms gives the following number of yellow M&Ms:

$$23\ 15\ 16\ 16\ 18\ 7\ 17\ 8\ 16\ 17\ 13\ 23\ 13\ 10\ 8$$

$$15\ 13\ 14\ 18\ 22\ 15\ 5\ 12\ 16\ 16\ 14\ 19\ 14\ 14\ 8$$

for which the summary statistics are as follows:

1

mean: 14.5
median: 15
variance: 19.9
standard deviation: 4.46
F_U: 17
F_L: 13
IQR: 4
MAD: 2

Construct a histogram and a boxplot for the yellow M&Ms in these 30 samples. For the histogram use the following bins with right endpoints: 5–10, 10–15, 15–20, and 20–25. In two to three sentences, what do your two plots suggest to you about the distribution of the data?
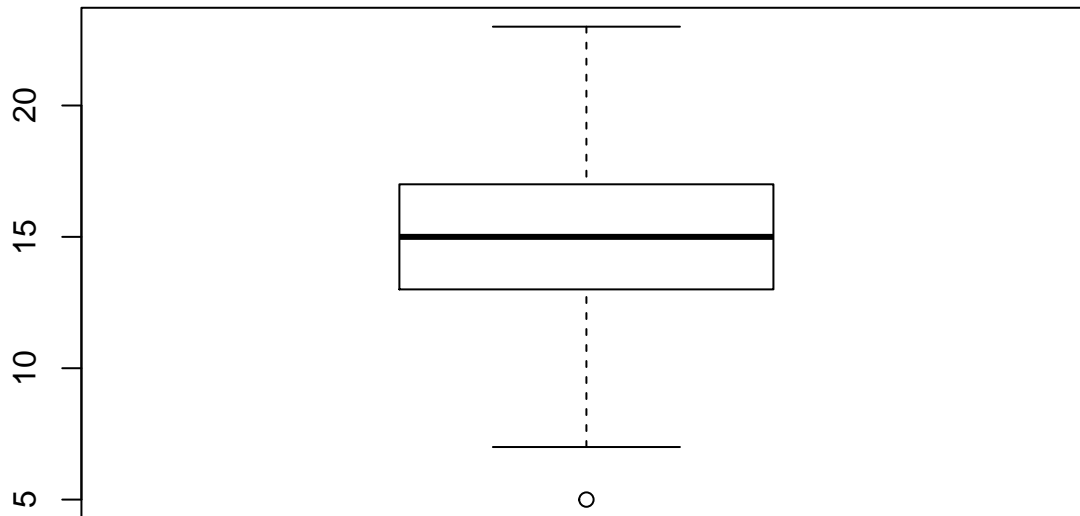
To construct the histogram we first need to bin the samples, which gives the following table of results:

bin	number
5–10	6
10–15	11
15–20	10
20–25	3

The resulting histogram will consist of four bars with heights of 6, 11, 10, and 3.

Histogram of yellow
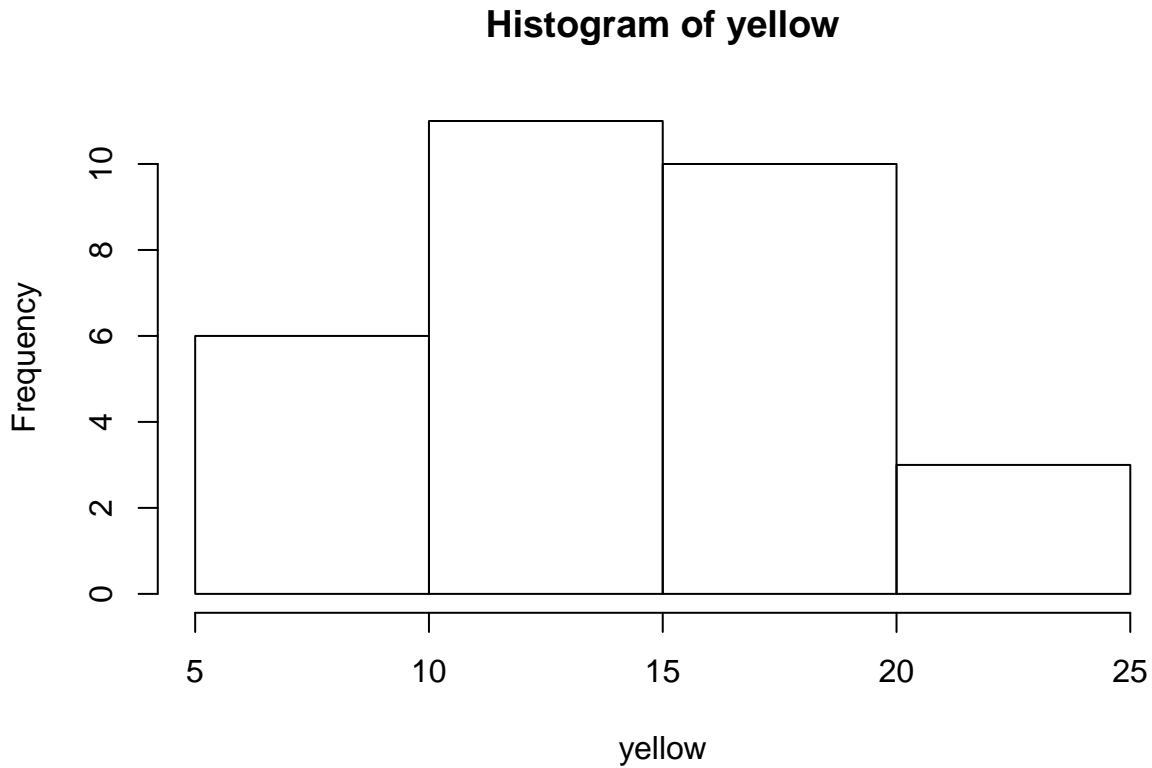
The boxplot is shown here

Both the histogram and the boxplot suggest a symmetrical distribution of the data. Both also suggest that results have a small skew toward lower values (more samples in the 5–10 bin than in the 20–25 bin for the histogram and a result that falls below the lower whisker in the box plot); however, with just 30 samples we should not be surprised to find such deviations.