Assignment 1 Solutions

1-6. The order of the solutions of Questions 1-6 may be different than the order in which which the questions were presented to you.

The variable **swimmer** is categorical and nominal.

The variable **medal** is categorical and ordinal.

The variable **time** is quantitative.

The variable **race** is categorical and nominal.

The variable **lane #** is categorical and ordinal.

The variable **placing** is categorical and ordinal.

7. We first find the position of the median. The median is in position $(n+1)/2 = 103/2 = 51.5$ of the ordered data values (i.e., the average of the 51^{st} and 52^{nd} ordered values).

   There are 51 data values before the median and 51 data values after the median. The third quartile is the median of the last 51 data values, so $Q3$ is in position $(51+1)/2 = 26$ among the last 51 data values. In other words, $Q3$ is in position $51 + 26 = 77$ since there are 51 data values before the median (note that the median is not one of the data values).

   There are 6 data values ≤ 25. There are $6 + 14 = 20$ data values ≤ 50. There are $6 + 14 + 17 = 37$ data values ≤ 75. There are $6 + 14 + 17 + 25 = 62$ data values ≤ 100. There are $6 + 14 + 17 + 25 + 15 = 77$ data values ≤ 125.

   The 77^{th} ordered data value must therefore be contained between 100 and 125. Even though we don't know the exact value of the third quartile, we know it must lie in the $100 - 125$ interval.

   Note that, because $Q3$ is in position 26 among the last 51 data values, we can also find the position of $Q3$ as the 26^{th} largest data value. There is 1 data value ≥ 250, $1 + 1 = 2$ data values ≥ 225, $1 + 1 + 3 = 5$ data values ≥ 200, $1 + 1 + 3 + 5 = 10$ data values ≥ 175, $1 + 1 + 3 + 5 + 6 = 16$ data values ≥ 150, $1 + 1 + 3 + 5 + 6 + 9 = 25$ data values ≥ 125, and $1 + 1 + 3 + 5 + 6 + 9 + 15 = 40$ data values ≥ 100. Therefore, the 26^{th} largest data value must lie in the $100 - 125$ interval.

8. We first need to find the median of this data set to help us find $Q1$ and $Q3$. The median is in position $(n + 1)/2 = 55/2 = 27.5$, i.e., the average of the 27^{th} and 28^{th} ordered values. Therefore, the median is $(162 + 164)/2 = 163$.

   The first quartile is the median of the first 27 ordered data values, so $Q1$ is in position $(27 + 1)/2 = 14$. Therefore, $Q1 = 157$.

The third quartile is the median of the last 27 ordered data values, so $Q3$ is in position $(27+1)/2 = 14$ above the median. Equivalently, we could count 14 positions down from the maximum. Therefore, $Q3 = 168$.

Next, we calculate the lower and upper fences:

$LF = Q1 - 1.5IQR = Q1 - 1.5(Q3 - Q1) = 157 - 1.5(168 - 157) = 157 - 16.5 = 140.5$

$UF = Q3 + 1.5IQR = Q3 + 1.5(Q3 - Q1) = 168 + 1.5(168 - 157) = 168 + 16.5 = 184.5$

Any data value less than 140.5 or greater than 184.5 will be labeled as an outlier. We see that there are 2 outliers on the left (137 and 140) and one outlier on the right (193).

9. The left whisker extends to the lowest data value which is not an outlier (i.e., 142). The right whisker extends to the highest data value which is not an outlier (i.e., 184).

10. (a) Standard deviation is a measure of spread around the mean. Data set C has just one observation right at the mean (10), two observations far below the mean, and two observations far above the mean. So most of the values in data set C are spread very far away from the mean, more so than in data sets A and B.

    (b) Since the mean is 10 and all observations are equal to 10, all of the observations are concentrated right at the mean, i.e., there is *no* spread around the mean and the standard deviation is zero.

11. John's final percentage grade is calculated as a weighted average, using the weights given to different assessment items. We are given the desired weighted average (70%), we have all of the weights needed in the calculation, and we need to solve for the score on the missing assessment item (the final exam):

$$\overline{x}_W = \frac{x_1 w_1 + x_2 w_2 + x_3 w_3}{w_1 + w_2 + w_3}$$

$$70 = \frac{(63)(35) + (82)(15) + (x_3)(50)}{35 + 15 + 50}$$

$$7000 = 2205 + 1230 + 50x_3$$

$$x_3 = 71.3$$