Percentiles, Box and Whisker Plots

Percentiles, Box and Whisker Plots

- We are in sections 2.3, 2.4 of the textbook
- The $p-th$ percentile is the data point which is greater than or equal to $p\%$ of the data.
  - Write the percentile as a decimal
  - Multiple by the size of the data set N
  - Round up. Take this data point.
- The *lower quartile*, Q1, is the 25th percentile. Compute this as the median of the data below the median (don't use the median).
- The *upper quartile*, Q3, is the 75th percentile. Compute this as the median of the data above the median (don't use the median).
- The **five number summary** consists of: Q0 = Min, Q1, Q2 = Median, Q3, Q4 = Max.
- The **interquartile range** (IQR) is Q3 − Q1. This is another measure of variation.
- *Box plots* (box and whisker plots) display quantitative data. Box plots visually display the five-summary.
  - Box plots help you compare several data sets
- *Outliers* are data points that are
  - less than 1.5 * IQR below Q1
  - more than 1.5 * IQR above Q3
- In SPSS, outliers are marked with an open circle. SPSS denotes extreme outliers, more than 3 * IQR from Q1 or Q3, with an asterisk.

1. Compute the five number summary and IQR for this data set:

   2 4 4 6 7 8 9 10 10 11

   Min _____ Q1 _____ Median: _____ Q3 _____ Max _____

   IQR _____

2. Compute the five number summary and IQR for this data set:

48 50 52 52 60 78 84 90 100

Min _____ Q1 _____ Median: _____ Q3 _____ Max _____

IQR _____

3. Consider the following data set of 10 data points:

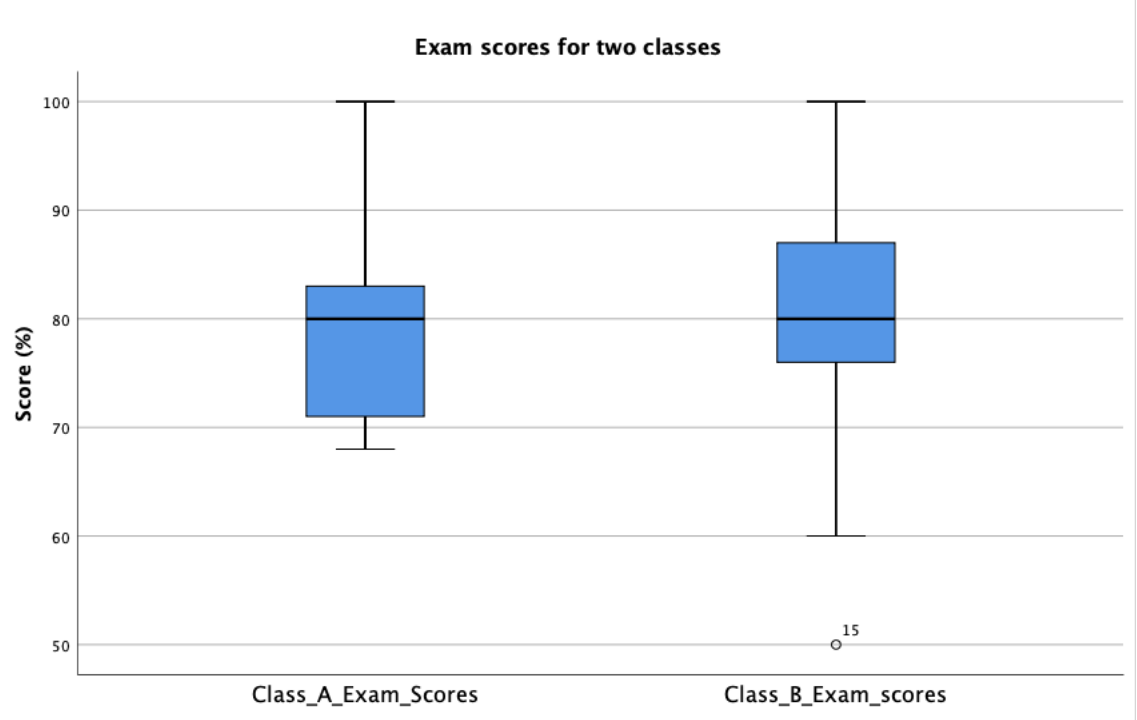
15 18 22 25 29 31 33 38 42 45

(a) What value corresponds to the 100th percentile?

(b) What value corresponds to the 40th percentile?

(c) What value corresponds to the 25th percentile?

4. The following diagram from SPSS shows box and whisker plots for the exam scores of two classes. We want to compare the classes to see which did better.



(a) Fill in the following table.

Class	Min	Q1	Median	Q3	Max
Class A					
Class B					

(b) Use the information in the box and whisker plot to decide which class you think did better overall. Write an explanation using several sentences that explains how you analyzed the data, which class you think did better, and why you think that class did better.