Exercise 2

1. The following data give the odometer mileage (rounded to the nearest thousand miles) for all 20 cars that are for sale at a dealership.

27 27 28 36 38 40 40 43 43 48

50 58 62 72 77 84 86 86 90 94

   a. Calculate the mean and median. Do these data have a mode? Why or why not?

   Mean = $\sum x/N$ = 1129/20 = 56.45

   Median = (48 + 50) / 2 = 49

   Mode = 27, 40, 43, and 86 as they occur twice, whereas other values occur once.

   b. Compute the range, variance, and standard deviation for these data.

   (This is a population data)

   Largest = 94, smallest = 27, range = 94 – 27 = 67

   $\sum x = 1129, \sum x^2 = 73789, N = 20$

   $$\sigma^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N} = \frac{73789 - \frac{(1129)^2}{20}}{20} = 502.85$$

   $$\sigma = \sqrt{502.85} = 22.42$$

2. The following data give the number of driving citations received during the last three years by 11 drivers.

4 8 0 3 11 7 14 8 13 7 9

   a. Find the mean and median for these data. Do these data have a mode? Why or why not?

   Mean = $\sum x/n$ = 84/11 = 7.636

   Ranked data: 0 3 4 7 7 8 8 9 11 13 14

   Median = 8

   Mode = 7 and 8 as they occur twice whereas other values occur once.

   b. Calculate the range, variance, and standard deviation.

   (This is a sample data)

   Range = largest – smallest = 14 – 0 = 14

   $\sum x = 84, \sum x^2 = 818, n = 11$

   $$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1} = \frac{818 - \frac{(84)^2}{11}}{10} = 17.65$$

$\sigma = \sqrt{17.65} = 4.202$

3. The mean time taken to learn the basics of a software program by all students is 200 minutes with a standard deviation of 20 minutes.
   a. Using Chebyshev's theorem, find the minimum percentage of students who learn the basics of this software program in
      i. 160 to 240 minutes

      $\mu = 200, \sigma = 20$

      $\mu + k\sigma = 240$

      $200 + 20k = 240$

      $k = 2$

      $100\left(1 - \frac{1}{k^2}\right) = 75\%$

      At least 75% of students takes time between 160 to 240 minutes.

      ii. 140 to 260 minutes

      $\mu + k\sigma = 260$

      $k = 3$

      $100\left(1 - \frac{1}{k^2}\right) = 88.89\%$

      At least 88.89% of students takes time between 140 to 260 minutes.

   b. Using Chebyshev's theorem, find the interval that contains the times taken by at least 84% of all students to learn this software program.

      $100\left(1 - \frac{1}{k^2}\right) = 84$

      $k = 2.5$

      $\mu - k\sigma = 200 - 2.5(20) = 150$

      $\mu + k\sigma = 200 + 2.5(20) = 250$

      The interval is between 150 to 250 minutes.

4. The mean time taken to learn the basics of a software program by all students have a bell-shaped distribution with a mean of 200 minutes and a standard deviation of 20 minutes.
   a. Using the empirical rule, find the (approximate) percentage of students who learn the basics of this software program in

i. 180 to 220 minutes

$\mu = 200, \sigma = 20$

$\mu + k\sigma = 220$

$200 + k(20) = 220$

$k = 1$

Approximately 68% of students take time between 180 to 220 minutes.

ii. 160 to 240 minutes

$\mu + k\sigma = 240$

$200 + k(20) = 240$

$k = 2$

Approximately 95% of students take time between 160 to 240 minutes.

b. Using the empirical rule, find the interval that contains the times taken by (approximate) 99.7% of all students to learn this software program.

$k = 3$

$\mu - k\sigma = 200 - 3(20) = 140$

$\mu + k\sigma = 200 + 3(20) = 260$

Interval is 140 to 260 minutes

5. A student washes her clothes at a laundromat once a week. The data below give the time (in minutes) she spent in the laundromat for each of 15 randomly selected weeks. Here, time spent in the laundromat includes the time spent waiting for a machine to become available.

62 67 72 73 75 77 81 83

84 85 90 93 107 112 135

a. Prepare a box-and-whisker plot.

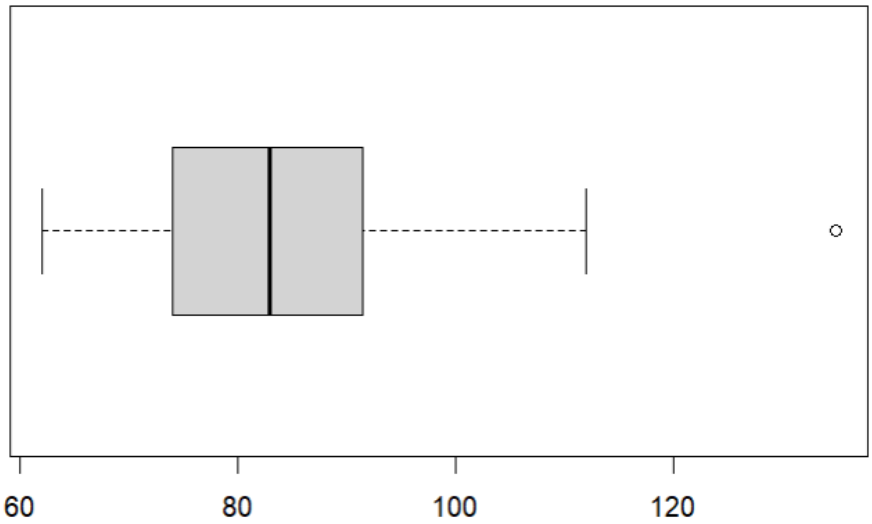
$Q_2 = 83, Q_1 = 73, Q_3 = 93, IQR = 93 - 73 = 20$

Lower inner fence = $Q_1 - 1.5 \times IQR = 73 - 1.5(20) = 43$

Upper inner fence = $Q_3 + 1.5 \times IQR = 93 + 1.5(20) = 123$

Smallest within inner fences = 62

Largest within inner fences = 112



b. Is the data set skewed in any direction? If yes, is it skewed to the right or to the left?

Skewed to the right. Longer whisker on the right.

c. Does this data set contain any outliers?

Yes, 135 is outlier.