STAT1011 Assignment 1 Solution

Section 3-1

4.

The numbers of observers for the top 10 states:

Mean $= \bar{X} = \frac{\sum_{i=1}^{10} X_i}{10} = 380.4$

Median $= \frac{352+378}{2} = 365$

Mode: no mode.

Midrange $= \frac{302+484}{2} = 393$

The numbers of visits for the top 10 states:

Mean $= \bar{Y} = \frac{\sum_{i=1}^{10} Y_i}{10} = 276.9$

Median $= \frac{219+194}{2} = 206.5$

Mode: no mode.

Midrange $= \frac{634+114}{2} = 374$

Mean, median and midrange of the first group of data are larger than those of the second group of data, respectively. Both groups of data do not have mode. (More reasonable answers are acceptable.)

Section 3-2

16.

Range $= 2786 - 65 = 2721$

Variance $= \frac{\sum (X_i - \bar{X})^2}{n-1} = 355427.6$

Standard deviation $= \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}} = 596.1775$

38.

Let X be the number of trials took by a sample of mice to learn.

$\mu = 12, \sigma = 3.$

From 4 to 20 $\Rightarrow \frac{8}{3}$ standard deviations from the mean $\Rightarrow \mathrm{k} = \frac{8}{3}$

$$\mathrm{P}(4 \le \mathrm{X} \le 20) = \mathrm{P}\left(12 - \frac{8}{3} \times 3 \le X \le 12 + \frac{8}{3} \times 3\right) \ge 1 - \frac{1}{k^2} = 1 - \frac{1}{\left(\frac{8}{3}\right)^2} = \frac{55}{64}$$

∴ The minimum percentage of data values that will fall in the rage of 4-20 trials is 85.94%.

40.

Let X be the fruit amount consumed by an American per year.

$\mu = 26.8, \sigma = 4.2.$

Since the distribution is bell-shaped and hence symmetric,

$$\mathrm{P}(\mathrm{X} > 31) = \mathrm{P}(\mathrm{X} > \mu + \sigma) = \frac{1}{2}(1 - P(\mu - \sigma \le X \le \mu + \sigma)) = \frac{1}{2} \times (1 - 0.68) = 0.16$$

∴ 16% Americans consume more than 31 pounds of citrus fruit per year.

Section 3-3

16.

a) $\mathrm{z} = \frac{3.2-4.6}{1.5} = -0.93$

b) $\mathrm{z} = \frac{630-800}{200} = -0.85$

c) $\mathrm{z} = \frac{43-50}{5} = -1.4$

∴ A score of 630 on a test with $\bar{X} = 800$ and $\mathrm{s} = 200$ indicates the highest relative position.

22.

percentile rank of 12 $= \frac{0+0.5}{7} \times 100\% = 7.14\%$

percentile rank of 28 $= \frac{1+0.5}{7} \times 100\% = 21.43\%$

percentile rank of 35 $= \frac{2+0.5}{7} \times 100\% = 35.71\%$

percentile rank of 42 $= \frac{3+0.5}{7} \times 100\% = 50\%$

percentile rank of 47 $= \frac{4+0.5}{7} \times 100\% = 64.29\%$

percentile rank of 49 $= \frac{5+0.5}{7} \times 100\% = 78.57\%$

percentile rank of 50 $= \frac{6+0.5}{7} \times 100\% = 92.86\%$

$\mathrm{c} = \frac{7\times 60}{100} = 4.2 \Rightarrow$ round up $\Rightarrow$ the 5$^{\text{th}}$ number corresponds to the 60$^{\text{th}}$ percentile.

$\therefore$ 47 corresponds to the 60$^{\text{th}}$ percentile.

28.

Arrange the data in ascending order:

0, 3, 6, 7, 8, 10, 11, 37, 48

Median is 8.

$Q_1 = \frac{3+6}{2} = 4.5,\ Q_3 = \frac{11+37}{2} = 24$

30.

a) $Q_1 = 84,\ Q_3 = 97,\ IQR = \ Q_3 - Q_1 = 97 - 84 = 13$

$1.5IQR = 1.5 \times 13 = 19.5$

$Q_1 - 1.5IQR = 84 - 19.5 = 64.5,\ Q_3 + 1.5IQR = 97 + 19.5 = 116.5$

Since all values fall within this range (64.5 to 116.5) $\Rightarrow$ no outliers.

b) $Q_1 = 118,\ Q_3 = 125,\ IQR = \ Q_3 - Q_1 = 125 - 118 = 7$

$1.5IQR = 1.5 \times 7 = 10.5$

$Q_1 - 1.5IQR = 118 - 10.5 = 107.5,\ Q_3 + 1.5IQR = 125 + 10.5 = 135.5$

The only value which falls outside the range of 107.5 to 135.5 is 145, which is identified as the outlier.

c) $Q_1 = 14.5,\ Q_3 = 23.5,\ IQR =\ Q_3 - Q_1 = 23.5 - 14.5 = 9$

$1.5IQR = 1.5 \times 9 = 13.5$

$Q_1 - 1.5IQR = 14.5 - 13.5 = 1,\ Q_3 + 1.5IQR = 23.5 + 13.5 = 37$

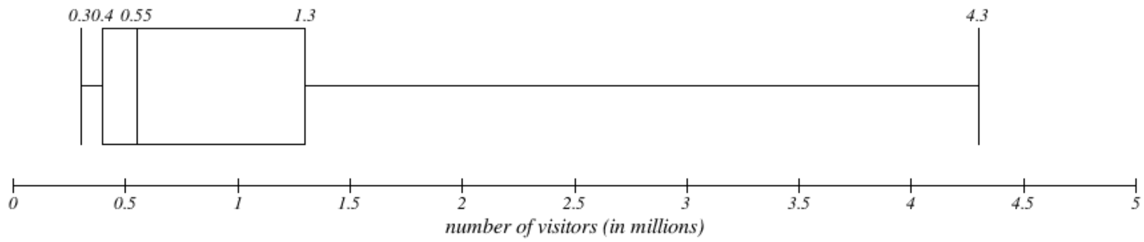
All values fall inside (1, 37) ⇒ no outliers.

Section 3-4

14.

Min = 0.3, $Q_1 = 0.4$, median = 0.55, $Q_3 = 1.3$, Max = 4.3

Boxplot:



The boxplot indicates that the data are positively skewed.