

Statistics 12, Homework 2

Michael Wu
UID: 404751542

January 27th, 2019

Chapter 3, Problem 20

a) I would expect the mean to be higher since this is a right skewed distribution. There are a few students who sent a high amount of emails that would increase the mean but not affect the median.

b) This distribution is heavily right skewed. The median is 1 email. The range of this distribution is 20, with a minimum of 1 and a maximum of 21. There is one outlier who sent 21 emails, and without this outlier the minimum and maximum of this distribution would be 1 and 8 emails, respectively.

c) I would use the median and interquartile range in order to summarize shape and spread for this distribution, since it captures the skew of this data better than mean and standard deviation.

Chapter 3, Problem 24

a)

$$\bar{x} = \frac{130 + 94 + 40 + 40 + 555 + 54 + 35 + 38 + 67 + 81 + 125 + 21 + 45 + 553 + 70 + 54}{16} = 125.125$$

b) Arranging in ascending order yields the following.

21, 35, 38, 40, 40, 45, 54, 54, 67, 70, 81, 94, 125, 130, 553, 555

The two innermost numbers are 54 and 67, so their average which is 60.5 gives us the median. The first quartile is 40, since those are the values halfway from the start to the median. The third quartile is the average of 94 and 125, which is 109.5 since those are the values halfway from the median to the end.

c) The range of this distribution is $555 - 21 = 534$. The interquartile range of this distribution is $109.5 - 60.5 = 49$.

Chapter 3, Problem 30

I would choose mean and standard deviation to summarize the center and spread of this data. This is because the data has a unimodal distribution that is fairly symmetric, so mean and standard deviation would describe it very well.

Chapter 4, Problem 16

a) On average, coffee is the most expensive commodity.

b) This data does not show that the newspaper is always more expensive than a ride on public transportation. There is one city where a ride of public transportation is \$3, while the minimum price of a newspaper is around \$1.75. So there is at least one newspaper that costs less than one ride on public transportation.

c) No, the presence of outliers does not affect my conclusions for the earlier parts. Even with the outliers removed, the distribution for coffee tends to have the highest price and there is at least one newspaper that costs less than one ride on public transportation.

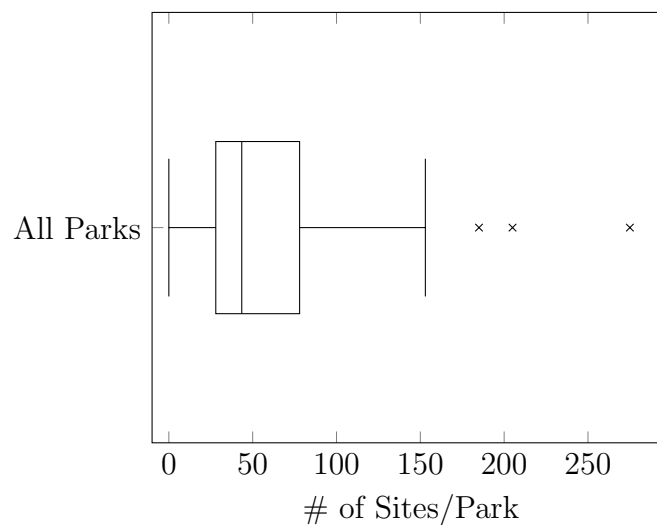
Chapter 4, Problem 22

a) I would use the median and interquartile range to describe the center and spread of this distribution. This is because the data is skewed right,

so the mean and standard deviation cannot capture the properties of the distribution as well.

b) The interquartile range is 50, so the outlier formula would put an upper bound at $78 + 1.5 \times 50 = 153$. Since the two parks in the range from 150 to 160 might be within this range, I would not say they are outliers. Then there are three parks outside of this range that must be outliers for sure.

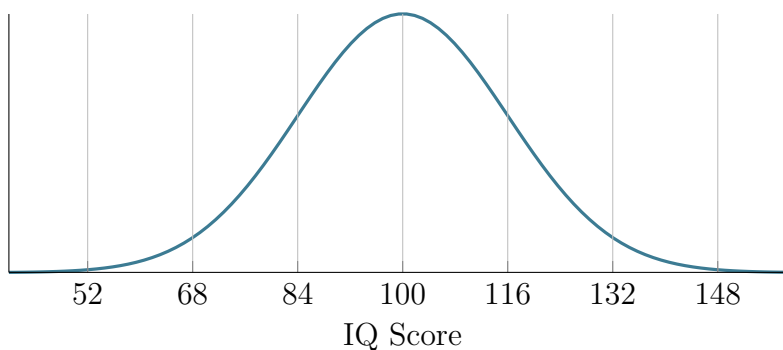
c)



d) This data is skewed right with a median of 43.5 and an interquartile range of 50. It has a minimum of 0, a first quartile at 28, third quartile at 78, and a maximum at 150. There are three outliers, one between 180 and 190, one between 200 and 210, and one at 275.

Chapter 5, Problem 8

a) The distribution of IQ scores is shown below. The ticks are spaced one standard deviation apart, so the 68-95-99.7 rule indicates that 68% of scores will be within one standard deviation, 95% will be within two standard deviations, and 99.7% will be within three standard deviations.



b) I would expect the central 95% of scores to be found between 68 and 132.

c) About 16% of people should have an IQ score above 116.

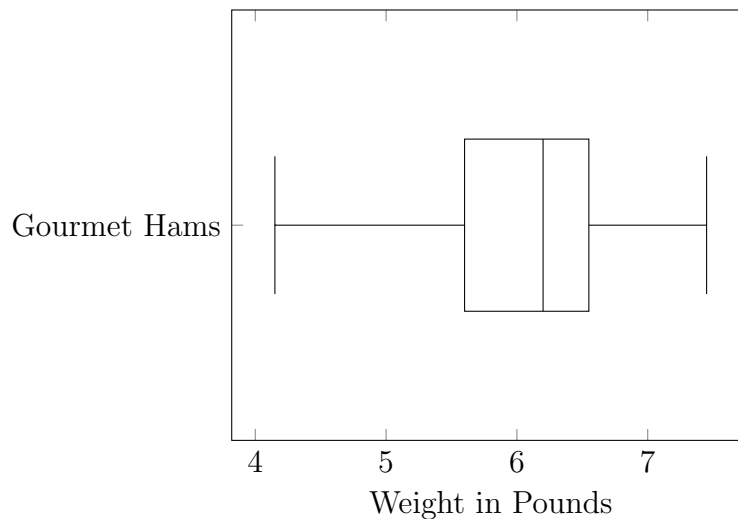
d) About 13.5% of people should have an IQ score between 68 and 84.

e) About 2.5% of people should have an IQ score above 132.

Chapter 5, Problem 14

a) The range is $7.45 - 4.15 = 3.3$ pounds. The IQR is $6.55 - 5.6 = 0.95$ pounds.

b) The distribution of the weights is slightly skewed left. I came to this conclusion by looking at the following plot.



c) The mean would be 96 ounces, the standard deviation would be 10.4 ounces, the first quartile is 89.6 ounces, the median is 99.2 ounces, the third quartile is 104.8 ounces, the IQR is 15.2 ounces, and the range is 52.8 ounces.

d) The mean would be 126 ounces, the standard deviation would be 10.4 ounces, the first quartile is 119.6 ounces, the median is 129.2 ounces, the third quartile is 134.8 ounces, the IQR is 15.2 ounces, and the range is 52.8 ounces.

e) The first quartile, median, third quartile, and IQR might not change.

Chapter 5, Problem 20

a) They have z-scores of 2.794, 3.865, and 3.183, respectively.

b) The most extraordinary wind event is the 3.93 mph wind in June.