NAME: _____ PERIOD: _____ DATE: _____

Homework Problem Set

An advertising agency researched the ages of viewers most interested in various types of television ads. Consider the following summaries:

Ages	Target Products or Services
30–45	Electronics, home goods, cars
46–55	Financial services, appliances, furniture
56–72	Retirement planning, cruises, health-care services

1. The mean age of the people surveyed is approximately 50 years old. As a result, the producers of the show decided to obtain advertisers for a typical viewer of 50 years old.

   A. According to the table, what products or services do you think the producers will target? * 50 is in age range 46-55

   Financial Services, Appliances, Furniture

   B. Based on the sample, what percent of the people surveyed about the *Fact or Fiction* show would have been interested in these commercials if the advertising table is accurate?

   11 out of 80 or about 14%

   * Use dot plot from page 65.

2. The show failed to generate the interest the advertisers hoped. As a result, they stopped advertising on the show, and the show was cancelled. Kristin made the argument that a better age to describe the typical viewer is the median age.

   A. What is the median age of the sample?

   Median: 60 years old

   B. What products or services does the advertising table suggest for viewers if the median age is considered as a description of the typical viewer? * in 56-72 age range

   Retirement planning, cruises, healthcare services

   C. What percent of the people surveyed would be interested in the products or services suggested by the advertising table if the median age were used to describe a typical viewer?

   31 out of 80 or about 39%

3. A. What percent of the viewers have ages between Q1 and Q3?

50%.

B. The difference between Q3 and Q1, or Q3 − Q1, is called the **interquartile range**, or **IQR**. What is the IQR for this data distribution?

IQR = 70-40 = 30 years

4. Do you think producers of the show would prefer a show that has a small or large interquartile range? Explain your answer.

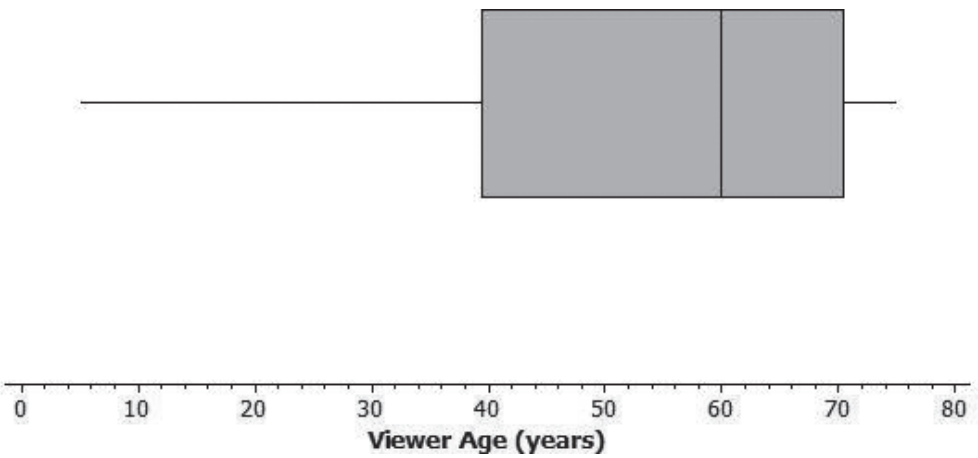
Smaller IQR is less variability, so it may be easier to target speaific audiences for advertisements.

Larger IQR is more variability. Even though it hits a wider age range, it may be harder to target advertising.

5. Do you agree with Kristin's argument that the median age provides a better description of a typical viewer? Explain your answer.

The median is a better description since this data is skewed.

6. Which ages, if any, do you think are outliers for the viewer ages in the box plot below?



OUTLIERS

IQR: 30 years

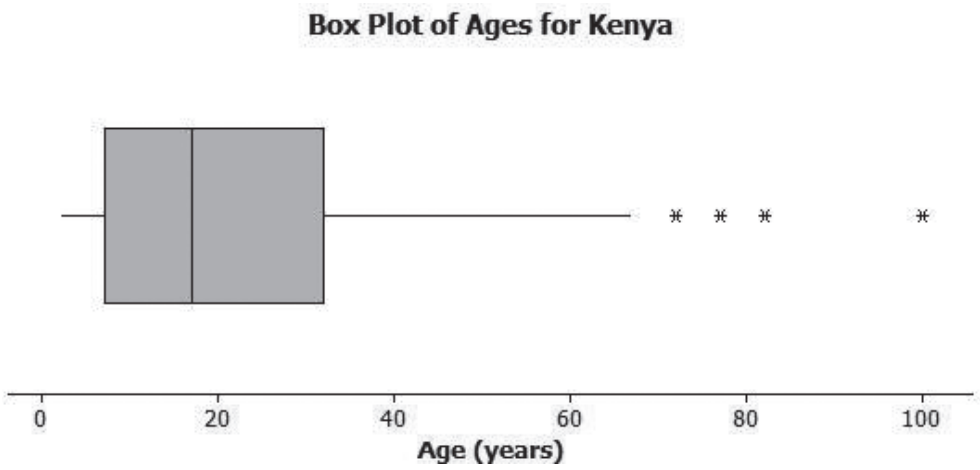
Upper boundary: 70 + (1.5 × 30) = 115

Lower boundary: 40 - (1.5 × 30) = -5

Possible Answer

I don't think there are any outliers. The IQR is almost the same length as the lower whisker. The lower whisker would need to be 1.5 times longer than IQR.

Students at Waldo High School are involved in a special project that involves communicating with people in Kenya. Consider a box plot of the ages of 200 randomly selected people from Kenya.



The four "*"s in the box plot represents the ages of four people from this sample. Based on the sample, these four ages were considered outliers.

7. Estimate the values of the four ages represented by an *.

possible answers: 70, 78, 82, 100 years old.

8. A. What is the median age of the sample of ages from Kenya?

Approximately 18 years old.

B. What are the approximate values of Q1 and Q3?

Q1: about 7 years old.
Q3: about 32 years old.

C. What is the approximate IQR of this sample? IQR: Q3-Q1

32-7=25 IQR: 25

D. Multiply the IQR by 1.5. What value do you get?

1.5 x 25 = 37.5 years

E. Add 1.5 × (*IQR*) to the third quartile age (Q3). What do you notice about the four ages identified by an *? upper boundary: Q3 + (1.5 x IQR)

32+37.5=69.5 years

* The 4 ages identified as outliers are more than 69.5 yrs.

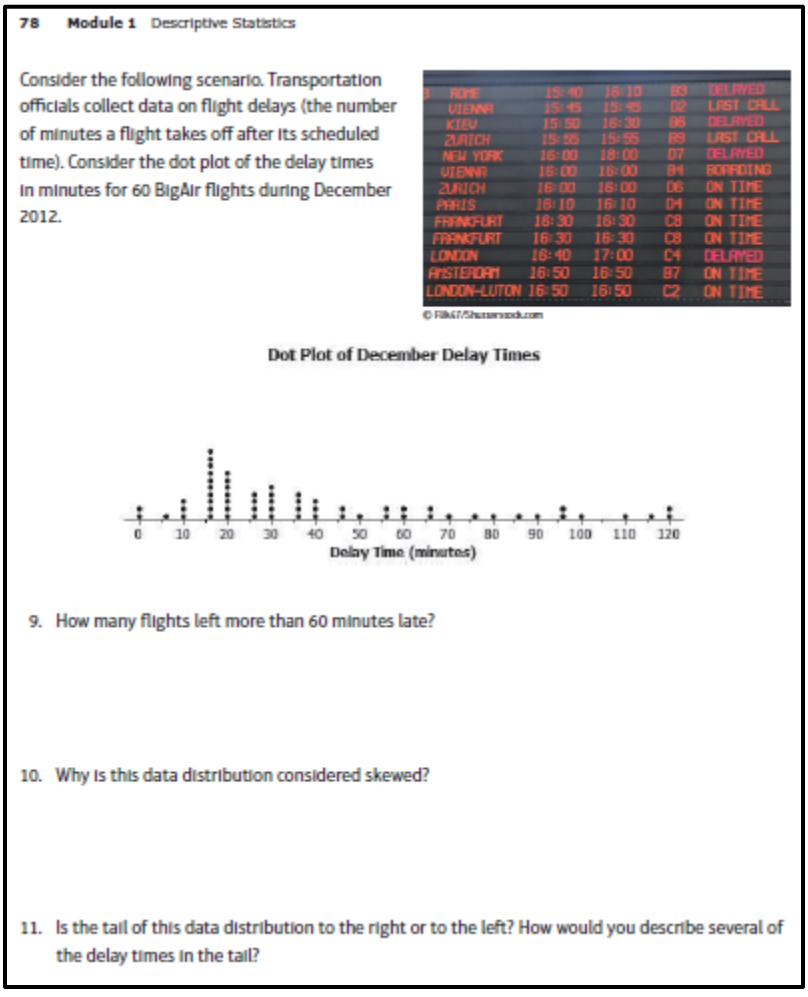
F. Are there any age values that are less than $Q1 - 1.5 \times (IQR)$? If so, these ages would also be considered outliers. lower boundary: Q1 - (1.5 x IQR)

7-37.5=-30.5 years

There are no ages less than this value.

G. Explain why there is no * on the low side of the box plot for ages of the people in the sample from Kenya.

An outlier on the lower end would be a negative age which is not possible.



78 **Module 1** Descriptive Statistics

Consider the following scenario. Transportation officials collect data on flight delays (the number of minutes a flight takes off after its scheduled time). Consider the dot plot of the delay times in minutes for 60 BigAir flights during December 2012.

Dot Plot of December Delay Times

9. How many flights left more than 60 minutes late?

10. Why is this data distribution considered skewed?

11. Is the tail of this data distribution to the right or to the left? How would you describe several of the delay times in the tail?

9. 14 flights left more than 60 minutes late.

10. This is a skewed distribution because there is a stretch of flights located to the right.

11. The tail is to the right. The delay times in the tail represent flights with the longest delays.

12. See below
13. The IQR is approximately $60 - 15$, or 45 minutes.
14. The mean value of 42 minutes is not a good description of a typical flight delay. It is pulled upward to a larger value because of flights with the very long delays.
15. Students should include a summary of the data in their reports. Included should be the median delay time of 30 minutes and that 50% of the flights are delayed between 15 minutes to 60 minutes, with a typical delay of approximately 30 minutes.
16. 14 flights were delayed more than 60 minutes, or 1 hour. These 14 flights represent approximately 23% of the flights. This is not a large number, although the decision of whether or not 23% is large is subjective.

12. Draw a box plot over the dot plot of the flights for December.

13. What is the interquartile range, or IQR, of this data set?

14. The mean of the 60 flight delays is approximately 42 minutes. Do you think that 42 minutes is typical of the number of minutes a BigAir flight was delayed? Why or why not?

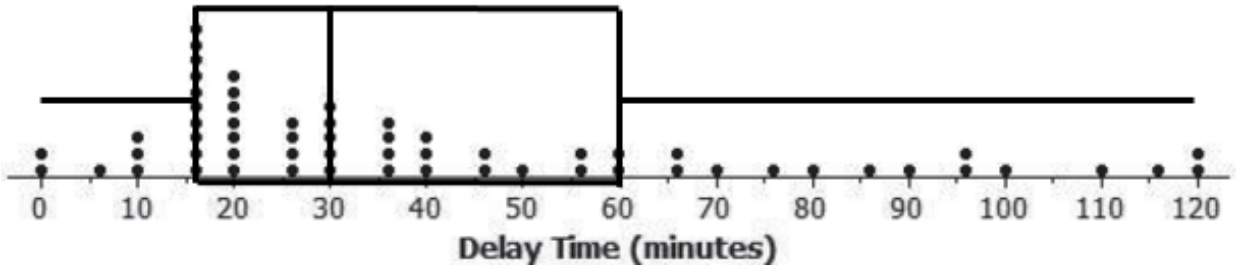
15. Based on the December data, write a brief description of the BigAir flight distribution for December.

16. Calculate the percentage of flights with delays of more than 1 hour. Were there many flight delays of more than 1 hour?

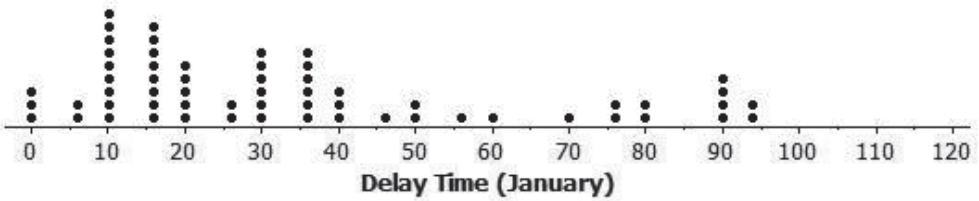
17. BigAir later indicated that there was a flight delay that was not included in the data. The flight not reported was delayed for 48 hours. If you had included that flight delay in the box plot, how would you have represented it? Explain your answer.

17. A flight delay of 48 hours would be much larger than any delay in this data set and would be considered an extreme value, or outlier. To include this flight would require an extension of the scale to $2{,}880$ minutes. This flight might have been delayed due to an extreme mechanical problem with the plane or an extended problem with weather.

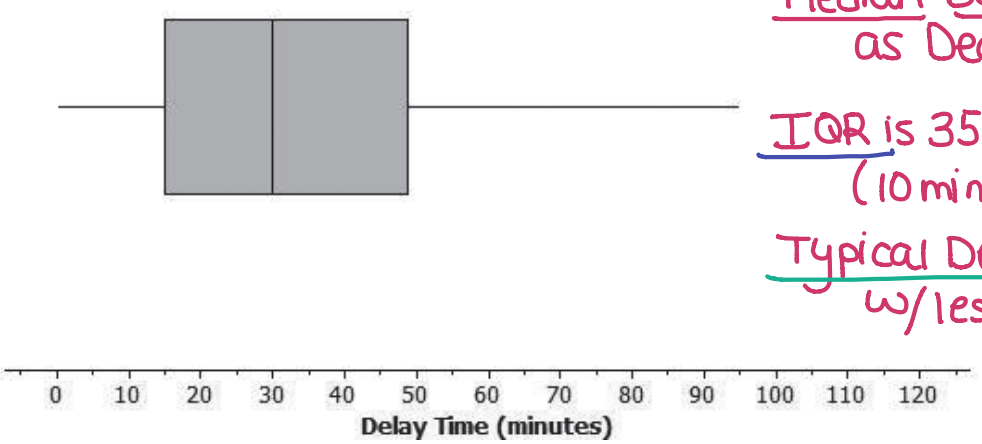
12.



18. A. Consider a dot plot and the box plot of the delay times in minutes for 60 BigAir flights during January 2013. How is the January flight delay distribution different from the one summarizing the December flight delays? In terms of flight delays in January, did BigAir improve, stay the same, or do worse compared to December? Explain your answer.



Box Plot of January Delay Times



Median Delay is same as December (30 min)

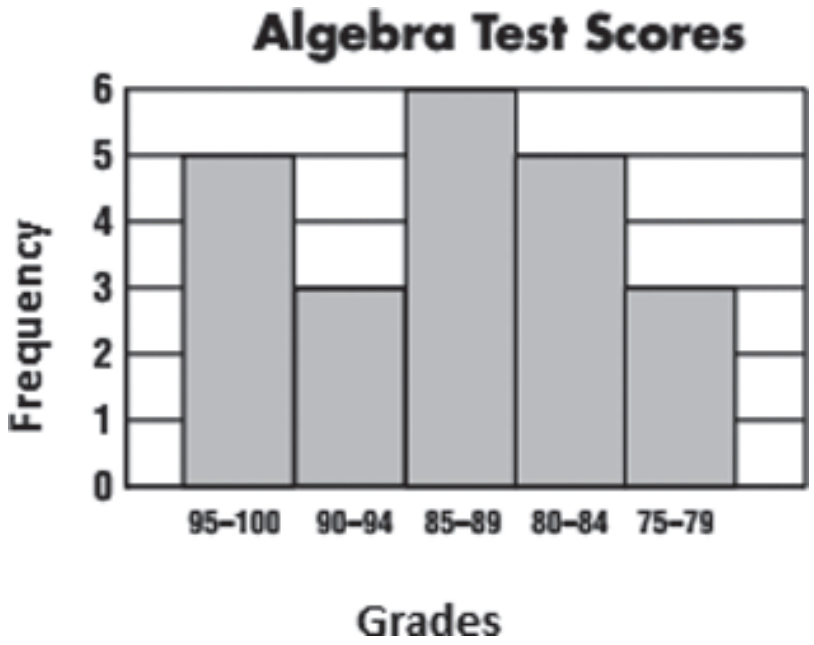
IQR is 35 minutes (10 min less than Dec.)

Typical Delay is 30 min w/ less variability.

B. Do you think this data set contains any outliers? Explain your thinking.

The IQR doesn't seem to be short enough to have an outlier in the upper whisker. The lower whisker is too short for outlier

Spiral REVIEW—Histograms



19. How many students took the algebra test?

22 students

20. Which grade has the most test scores?

85-89

21. Which grades have the same number of test scores?

95-100 & 80-84 → 5 each

90-94 & 75-79 → 3 each

22. How many more students earned 85–89 than earned 80–84?

One more student

23. How is this histogram different from the ones you studied in Lessons 2 and 3?

This histogram shows the lower and upper units for each bar. The other histograms only showed the upper limit for each bar. This histogram is also showing larger numbers on the left with smaller numbers on the right.