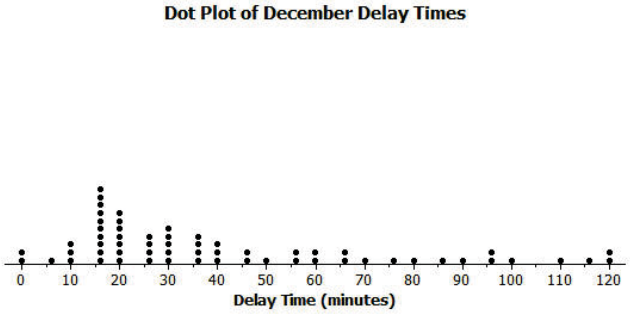
S.43

Homework Problem Set Sample Solutions

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:



1. **How many flights left more than 60 minutes late?**

   14 flights left more than 60 minutes late.

2. **Why is this data distribution considered skewed?**

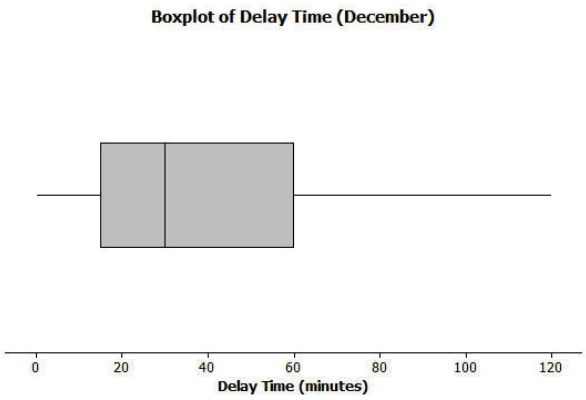
   This is a skewed distribution because there is a stretch of flights located to the right.

3. **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?**

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

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

   A box plot of the December delay times is as follows:



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

   The IQR is approximately $60 - 15$, or 45 minutes.

6. **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?**

   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.

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

   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.

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

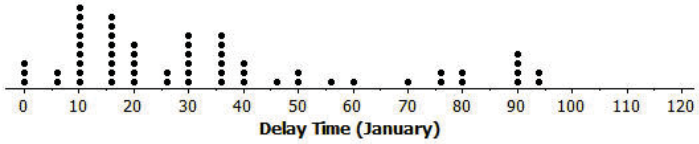
   14 flights were delayed more than 60 minutes, or 1 hour. These 14 flights represent approximately 23% of the flights. This is not a large number, although the decision of whether or not 23% is large is subjective.

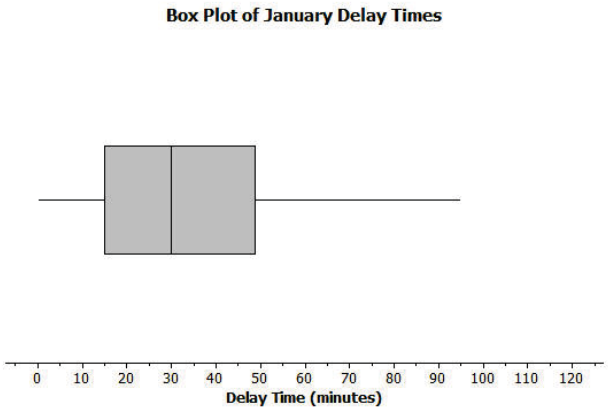
9. **BigAir later indicated that there was a flight delay that was not included in the data. The flight not reported was delayed for 48 hours. If you had included that flight delay in the box plot, how would you have represented it? Explain your answer.**

   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.

10. **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.





The median flight delay is the same as in December, which is 30 minutes. The IQR is less, or approximately 35 minutes. The maximum is also less. In general, this indicates a typical delay of 30 minutes with less variability.