

## LESSON

# 3

# Distributions and the Stories They Tell

### LEARNING OBJECTIVES

- Today I am: checking to see what vocabulary I remember about statistics.
- So that I can: describe sets of data based on the shape of their graph.
- I'll know I have it when I can: analyze dot plots and histograms.

## Opening Activity

Over the last few days your class has gathered data and reviewed different ways to display data. To tell the “story” behind each display you’ll need the language of statistics.



© Jo karen/Shutterstock.com

With your group use the word bank to match the statistics vocabulary to the correct definition.

Statistics Vocabulary	Definitions
1. Maximum	A. The greatest value in the data set
2. Minimum	B. The least value in the data set
3. Median	C. The value in the middle of the data set. For a data set with an even number of numbers take the average of the two middle numbers.
4. Mean	D. The average of all data in the set. Add all the numbers and divide by the total number of numbers.
5. Mode	E. The value appearing the most in the data set.
6. Range	F. The difference between the greatest number and least number in the set.
7. Standard Deviation	G. A calculated value that represents how far apart the data is from the mean.
8. Outliers	H. Values that are far from the norm in a data set.

**Word Bank:**

Mode	Mean	Standard Deviation	Outliers
Minimum	Range	Maximum	Median

Some of the vocabulary is easier to understand using pictures. Below are 9 dot plots and 9 words that could be used to describe those graphs.

With your group use the word bank to match the statistics vocabulary to the correct picture.

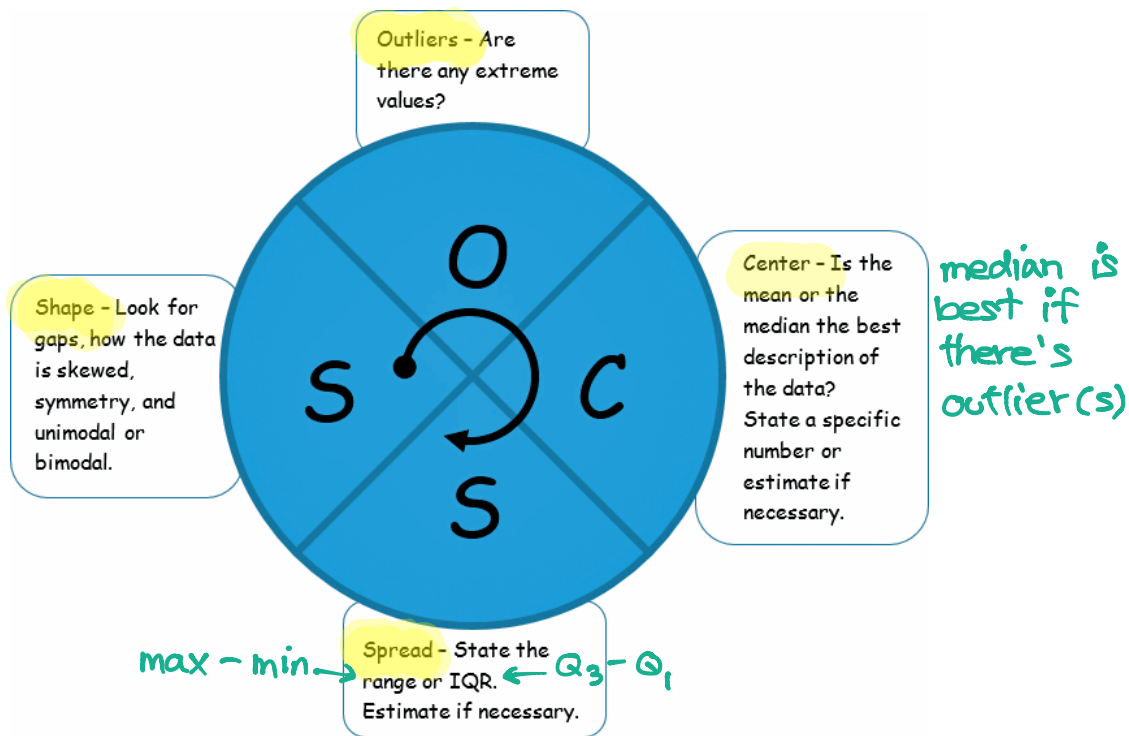
**Word Bank:**

Bimodal	Gap(s)	Left Skewed
Narrow IQR	Outlier(s)	Right Skewed
Symmetric	Wide IQR	Unimodal

<p>9. <i>left skewed</i></p>	<p>10. <i>right skewed</i></p>	<p>11. <i>Symmetric</i></p>
<p>12. <i>outlier</i></p>	<p>13. <i>Gaps</i></p>	<p>14. <i>Wide IQR</i></p>
<p>15. <i>narrow IQR</i></p>	<p>16. <i>Unimodal</i> <i>one peak</i></p>	<p>17. <i>Bimodal</i></p>

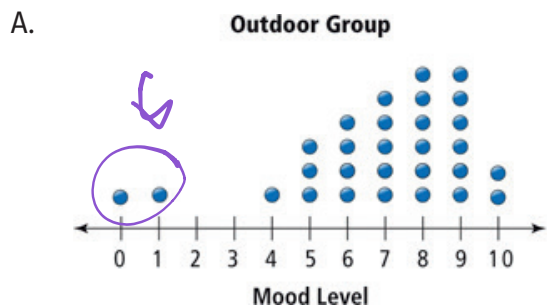
*35, 92, 94, 95*  
*95, 94, 92, 35*  
*median = 93*  
*mean = 79*

When describing graphs, look at the key features such as the **shape**, **outliers**, **center** and **spread**. These are shown in the SOCS diagram below.

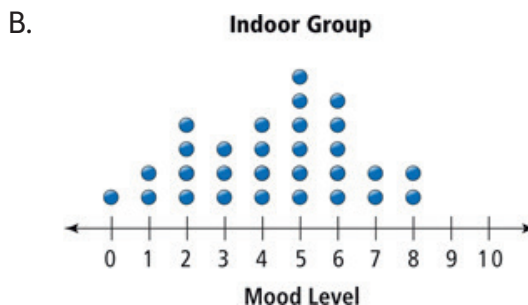


18. Use the words from the exploration to describe each dot plot below. Be sure to address all the parts of SOCS.

**There was a mood survey (0 being depressed and 10 being ecstatic) given to 2 groups of people who were exercising. One of the groups was exercising outdoors and one group was exercising indoors.**

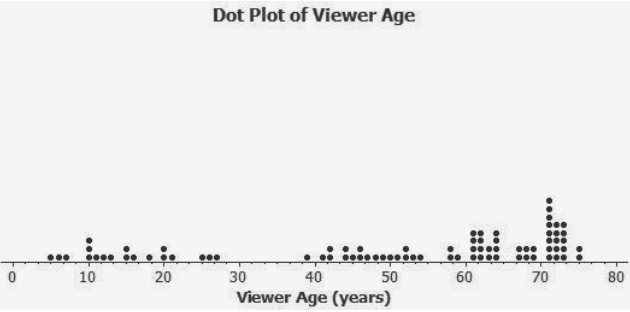
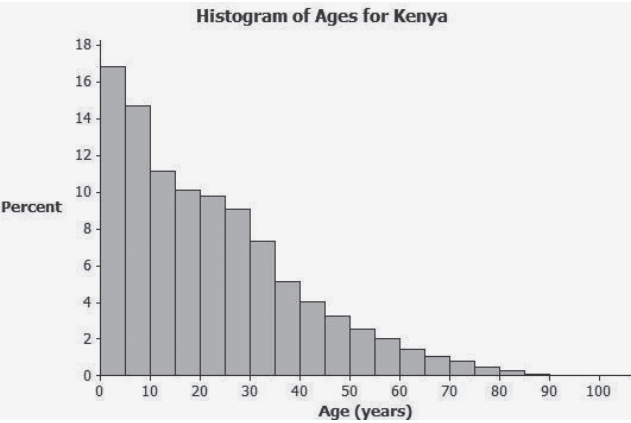


S left skewed  
 O possible  
 C median  
 S IQR



S unimodal  
 O none  
 C mean  
 S Range/IQR

**Graphs**—Data are often summarized by graphs. Graphs are the first indicator of variability in the data. Below is an example of four types of graphs you should be familiar with. For each state the advantages of challenges when using this type of graph.

<p><b>19. Dot Plots</b></p> <p>Advantages: <i>can see each data. can add easily, gaps, max &amp; min, cluster</i></p> <p>Challenges: <i>not ideal for large amount of data, can't see quartile.</i></p>	<p><b>20. Histograms</b></p> <p>Advantages: <i>can see large amount of data, patterns</i></p> <p>Challenges: <i>can not see the data (max, min, median)</i></p>
<p><b>Dot Plots:</b> A plot of each data value on a scale or number line. These plots are used with quantitative (numerical) data.</p>  <p>The dot plot shows individual data points for viewer ages. The x-axis is labeled 'Viewer Age (years)' and ranges from 0 to 80 with major ticks every 10 units. There are several small clusters of dots: a group of 3 dots at age 5, a group of 4 dots at age 10, a group of 3 dots at age 15, a group of 2 dots at age 20, a group of 3 dots at age 25, a group of 2 dots at age 30, a group of 2 dots at age 35, a group of 2 dots at age 40, a group of 2 dots at age 45, a group of 2 dots at age 50, a group of 2 dots at age 55, a group of 2 dots at age 60, a group of 3 dots at age 65, a group of 4 dots at age 70, a group of 5 dots at age 75, and a group of 2 dots at age 80.</p>	<p><b>Histograms:</b> A graph of data that groups the data based on intervals and represents the data in each interval by a bar. These plots are used with quantitative (numerical) data.</p>  <p>The histogram shows the distribution of ages for Kenya. The x-axis is labeled 'Age (years)' and ranges from 0 to 100 with major ticks every 10 units. The y-axis is labeled 'Percent' and ranges from 0 to 18 with major ticks every 2 units. The bars represent the percentage of the population in each 5-year age interval. The distribution is right-skewed, with the highest percentage (approximately 17%) in the 5-10 age group, and the percentage decreasing as age increases, reaching near zero for ages 80 and above.</p>

21. Box Plots

Advantages: *Can see range  
median, IQR*

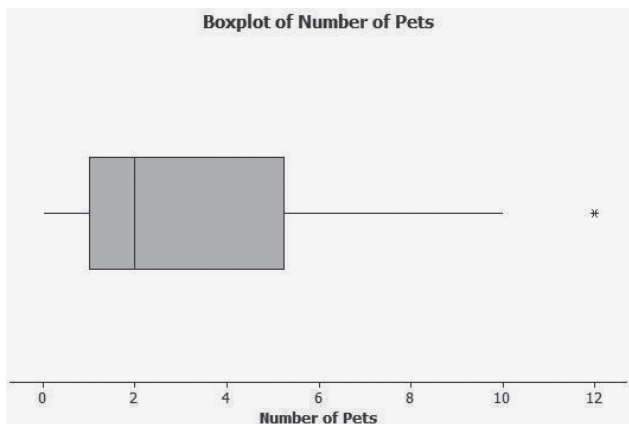
Challenges: *can not see individual  
data.*

22. Bar Graphs

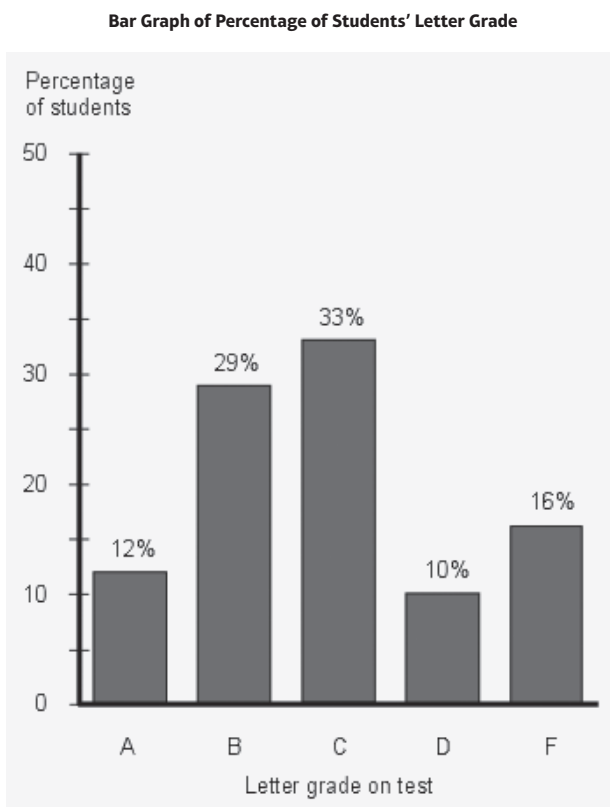
Advantages: *- good for categorical  
data, comparing*

Challenges: *- can be difficult  
to determine categories*

**Box Plots:** A graph that provides a picture of the data ordered and divided into four intervals that each contains approximately 25% of the data. These plots are used with quantitative (numerical) data.



**Bar Graphs:** A diagram in which the numerical values are represented by the length of lines of rectangles of equal width. These plots are used with categorical data.

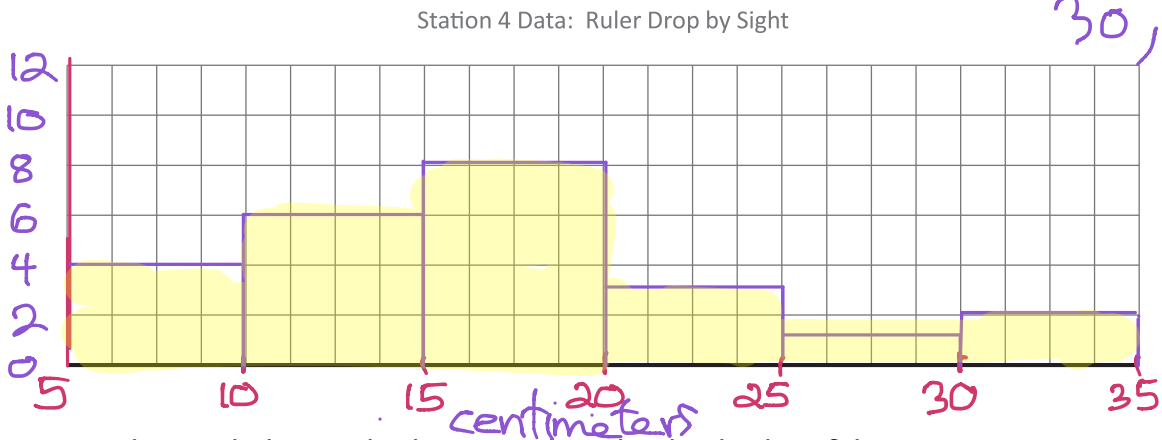


23. Let's look at the data from the first lesson when you did the Ruler Drop by Sight in Experiment 1.

A. With your group, determine which display would be best for this type of data. Explain your groups' reasoning.

5, 5.5, 6, 9, 10, 11, 13, 13, 14, 14, 15, 15, 15, 16  
 17, 18, 19, 19, 20  
 22, 24, 25  
 30, 30

B. Create your display in the space below. Be sure to clearly label the axes.



C. Use the vocabulary in this lesson to describe the display of data.

S  
 O  
 C  
 S





NAME: \_\_\_\_\_ PERIOD: \_\_\_\_\_ DATE: \_\_\_\_\_

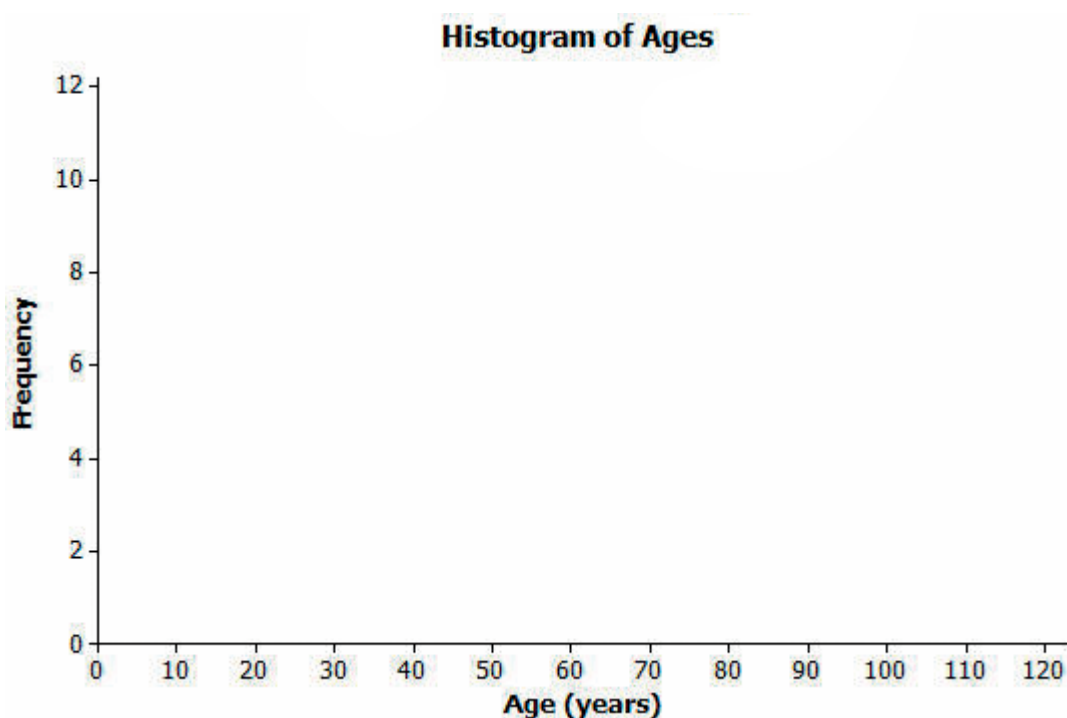
# Homework Problem Set

1. A group of forty people were attending an event. The ages of the people are as follows:

6, 13, 24, 27, 28, 32, 32, 34, 38, 42, 42, 43, 48, 49, 49, 51, 52, 52, 53,

53, 53, 54, 55, 56, 57, 60, 61, 61, 62, 66, 66, 66, 68, 70, 72, 78, 83, 97.

A. Create a histogram of the ages using the provided axes.



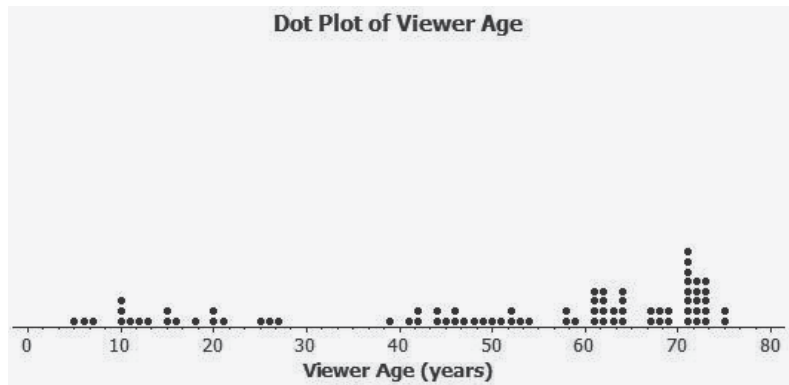
B. Would you describe your graph of ages as symmetrical or skewed? Explain your choice.

C. Identify a typical age of the forty people.

D. What event do you think the forty people were attending? Use your histogram to justify your conjecture.

E. Describe the graph using SOCS.

2. A random sample of eighty viewers of a television show was selected. The dot plot at the right shows the distribution of the ages (in years) of these eighty viewers.

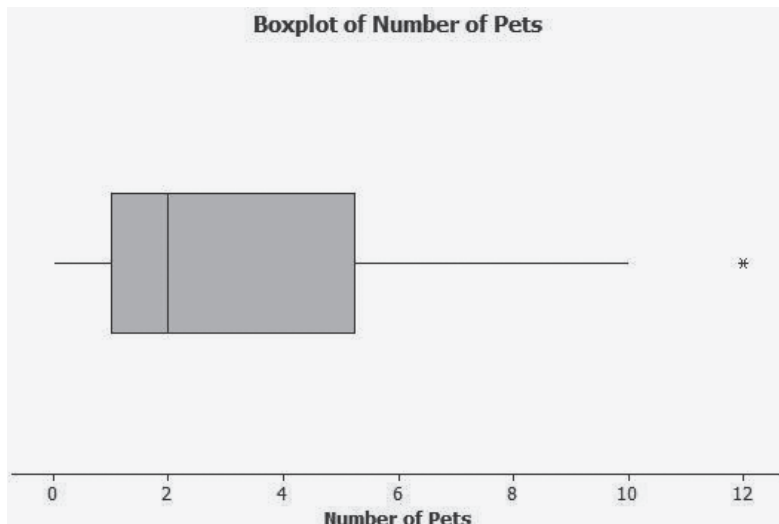


A. What do you think this graph is telling us about the ages of the eighty viewers in this sample?

B. Can you think of a reason why the data presented by this graph provide important information? Who might be interested in this data distribution?

C. Based on your previous work with dot plots, would you describe this dot plot as representing a symmetric or a skewed data distribution? Explain your answer.

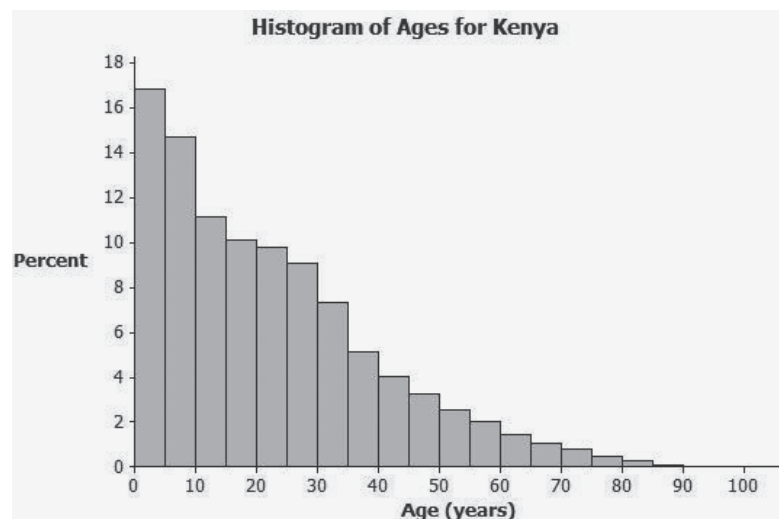
3. Thirty students from River City High School were asked how many pets they owned. The box plot was prepared from their answers.



A. What does the box plot tell us about the number of pets owned by the thirty students at River City High School?

B. Why might understanding the data behind this graph be important?

4. The histogram represents the age distribution of the population of Kenya in 2010.



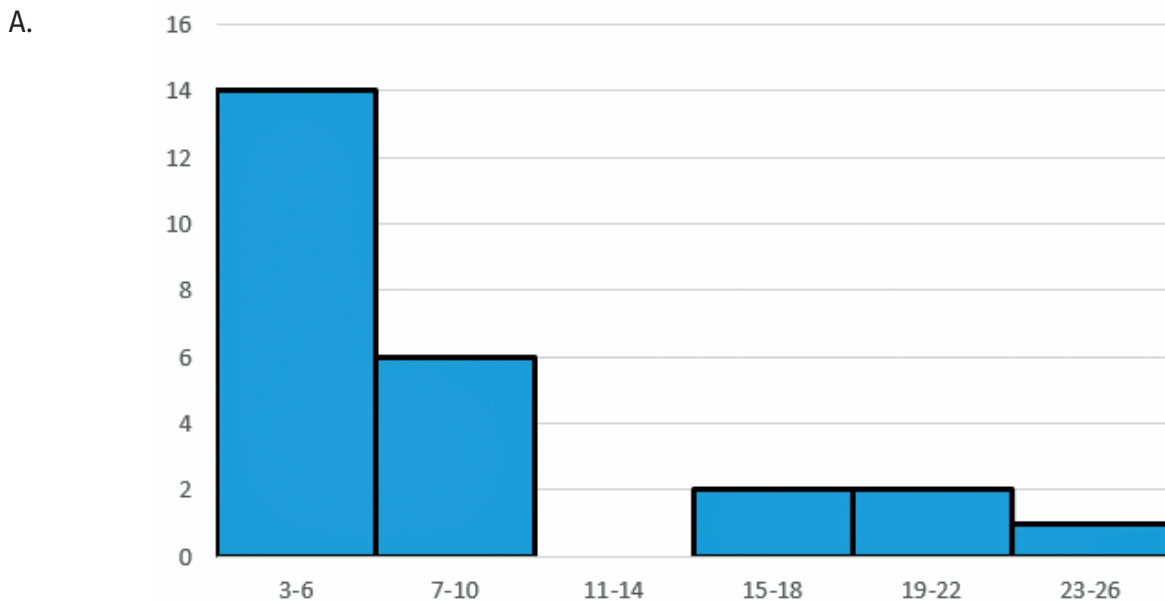
A. What do you think this graph is telling us about the population of Kenya?



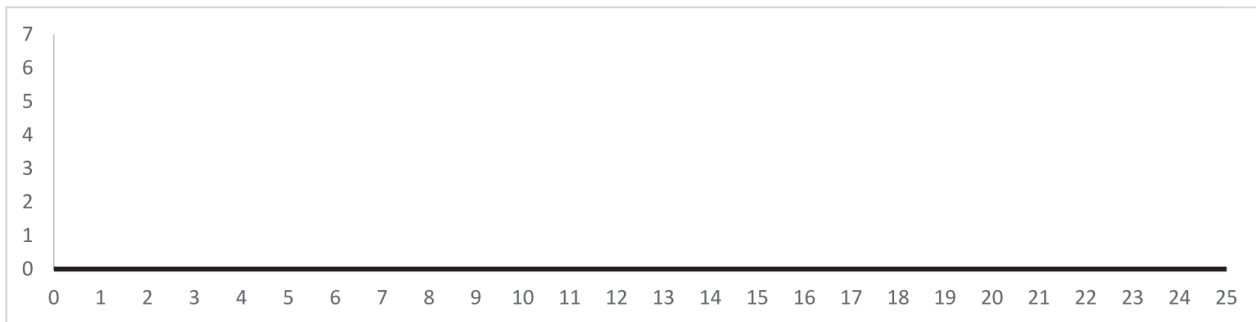
6. Use the data below to create a dot plot.

**Twenty-five people were attending an event. The ages of the people are as follows:**

**3, 3, 4, 4, 4, 4, 5, 6, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7, 7, 16, 17, 22, 22, 25.**



DOT PLOT

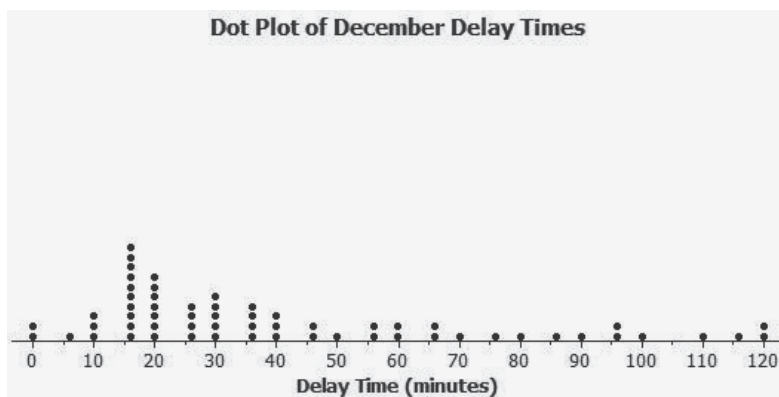


B. Which graph do you think is the most representative of the data? Why?



7. Answer the questions that accompany the graph to begin your understanding of the story behind the data.

**Transportation officials collect data on flight delays (the number of minutes past the scheduled departure time that a flight takes off). Consider the dot plot of the delay times for sixty BigAir flights during December 2012.**



- A. What do you think this graph is telling us about the flight delays for these sixty flights?
- B. Can you think of a reason why the data presented by this graph provide important information? Who might be interested in this data distribution?
- C. Based on your previous work with dot plots, would you describe this dot plot as representing a symmetric or a skewed data distribution? Use SOCS to describe this graph.

## REVIEW—Solving Equations and Simplifying Expressions with Exponents

Solve the equations below. Be sure to check your solution.

8.  $4x + 3 = 11$

9.  $2 - 2x = 12$

10.  $3x + 1 = 6x + 7$

11.  $\frac{x}{2} = \frac{5}{10}$

12.  $\frac{x}{2} + 3 = 7$

13.  $2(x + 1) = 16$

14.  $4(x - 2) = 5(x - 1)$

15.  $3(2x + 1) = 5(x - 2) + 12$

16.  $\frac{x+2}{x-1} = 2$

17.  $\frac{x}{10} - \frac{3x}{2} = 7$

18.  $7x + 4 + 2x = 2x + 3(3x - 1) + 7$

19.  $\frac{4}{x} + \frac{3}{x} = 14$

~~$\frac{7}{x} = \frac{14}{1}$~~   
 $7 = 14x$