

Lesson 3: Distributions and the Stories They Tell



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.

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

Statistics Vocabulary	Definitions
1. Max	A. The greatest value in the data set
2. Min	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. St. Dev.	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:

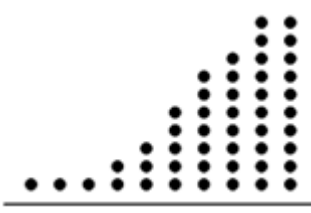


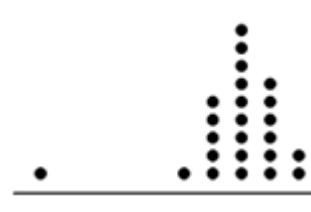
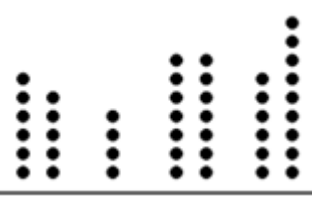

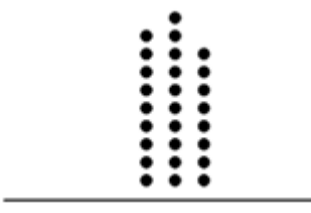
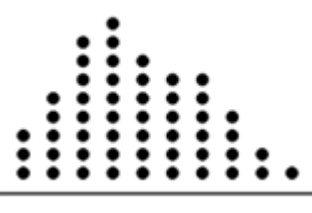

Mode	Mean	Standard Deviation	Outliers
Minimum	Range	Maximum	Median

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.

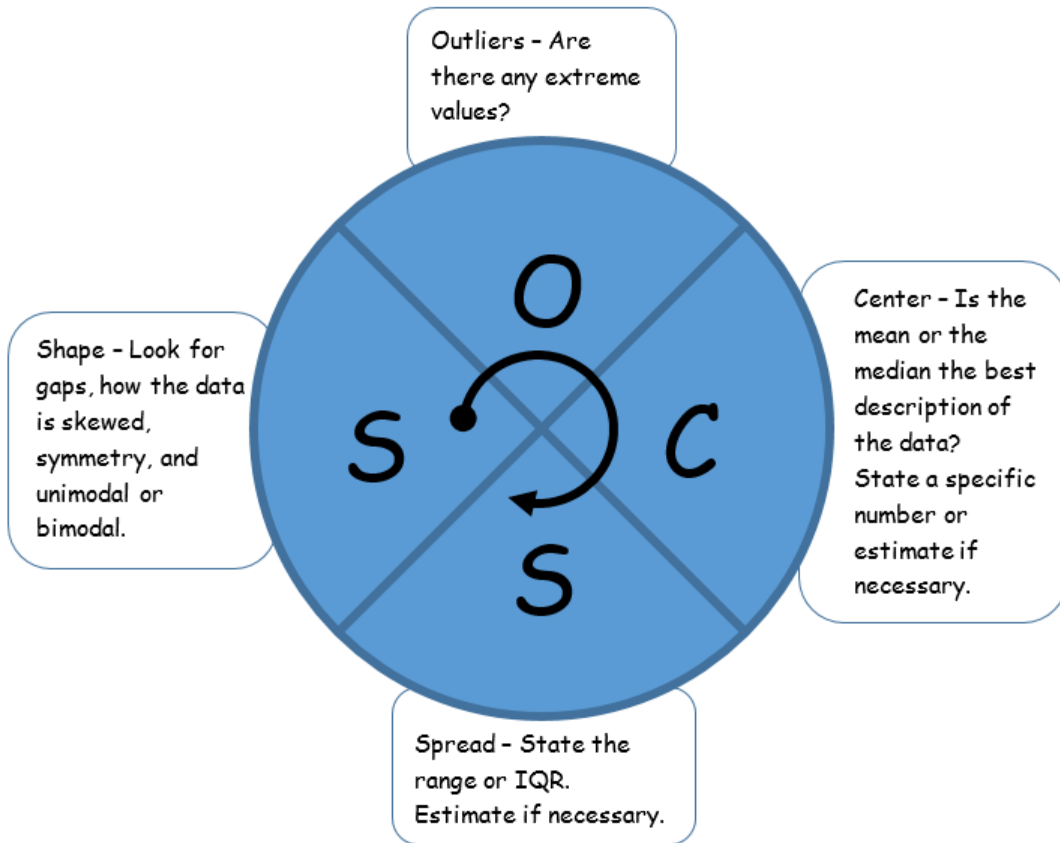
With your group use the word bank to match the statistics vocabulary to the correct picture.

Word Bank:

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

9. <u>Left Skewed</u> 	10. <u>Right Skewed</u> 	11. <u>Symmetric</u> 
12. <u>Outlier</u> 	13. <u>Gaps</u> 	14. <u>Wide IQR</u> 
15. <u>Narrow IQR</u> 	16. <u>Unimodal</u> 	17. <u>Bimodal</u> 

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

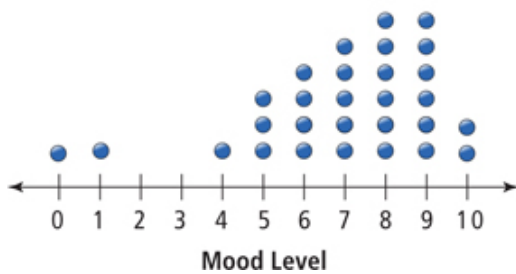


18. Use the words from the exploration to describe each dot plot below. Be sure to address all the parts of SOCS.

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.

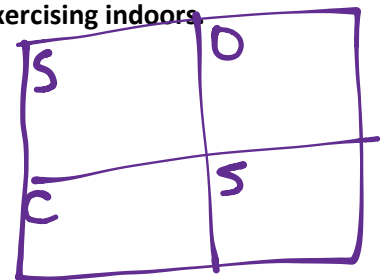
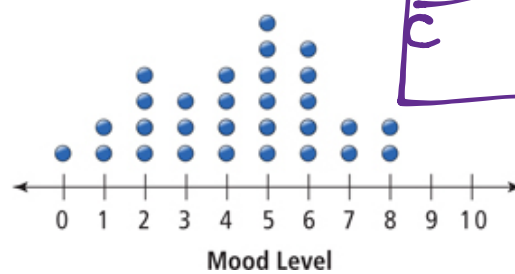
A.

Outdoor Group

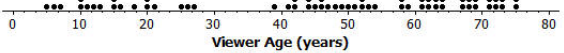
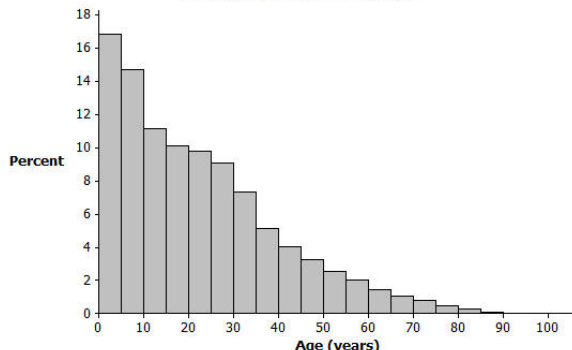


B.

Indoor Group



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>19. DOT PLOTS</p> <p>Advantages: <i>All data present. Easy to see clusters and gaps.</i></p> <p>Challenges: <i>Tedious</i></p>	<p>20. HISTOGRAMS</p> <p>Advantages: <i>Data is in an easy to read format.</i></p> <p>Challenges: <i>Can't see specific data. Gaps and clusters hidden in intervals.</i></p>
<p>DOT PLOTS: A plot of each data value on a scale or number line. These plots are used with quantitative (numerical) data.</p> <p style="text-align: center;">Dot Plot of Viewer Age</p> 	<p>HISTOGRAMS: 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 style="text-align: center;">Histogram of Ages for Kenya</p> 

21. BOX PLOTS

Advantages:

Good for large sets of data. Easy to see 5 number summary.

Challenges:

Not all data points shown so you can't see gaps and clusters.

22. BAR GRAPHS

Advantages:

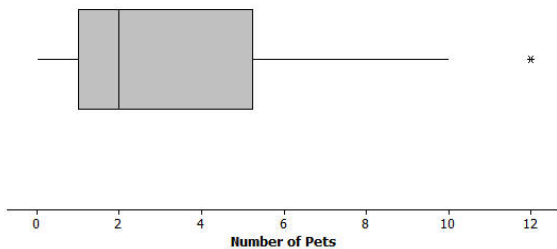
Can use categorical data. Good for comparing categories.

Challenges:

Can't see individual data points.

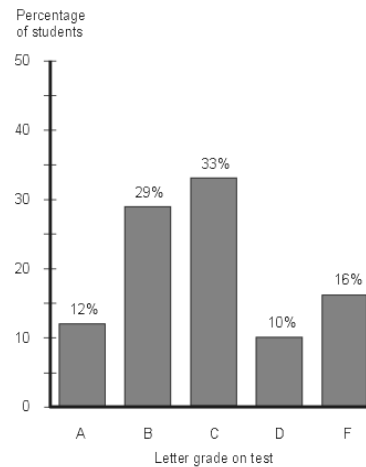
BOX PLOTS: A graph that provides a picture of the data ordered and divided into four intervals that each contains approximately 25% of the data. These plots are used with quantitative (numerical) data.

Boxplot of Number of Pets



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.

Bar Graph of Percentage of Students' Letter Grade



23. Let's look at the data from the first lesson when you estimated the length of a minute.
- A. With your group determine which display would be best for this type of data. Explain your groups' reasoning.
- B. Create your display in the space below. Be sure to clearly label the axes.

Station 6 Data: How Long is a Minute?

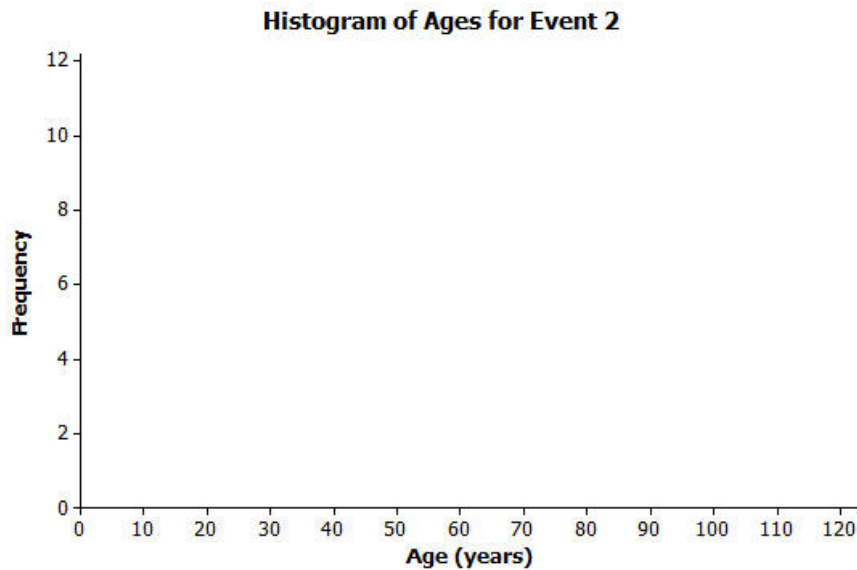
- C. Use the vocabulary in this lesson to describe the display of data.

Homework Problem Set

1. A different forty people were also 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, 49, 51, 52, 52, 53,
53, 53, 54, 55, 56, 57, 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.

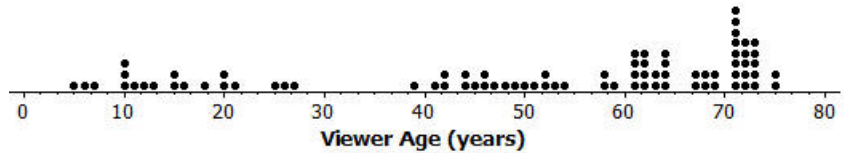
c. Identify a typical age of the forty people.

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.

Dot Plot of Viewer Age



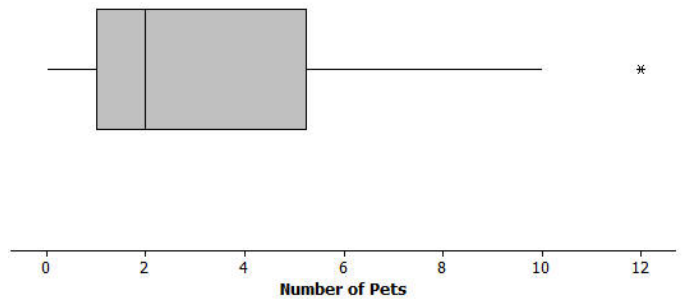
A. What do you think this graph is telling us about the ages of the eighty viewers in this sample?

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.

Boxplot of Number of Pets



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

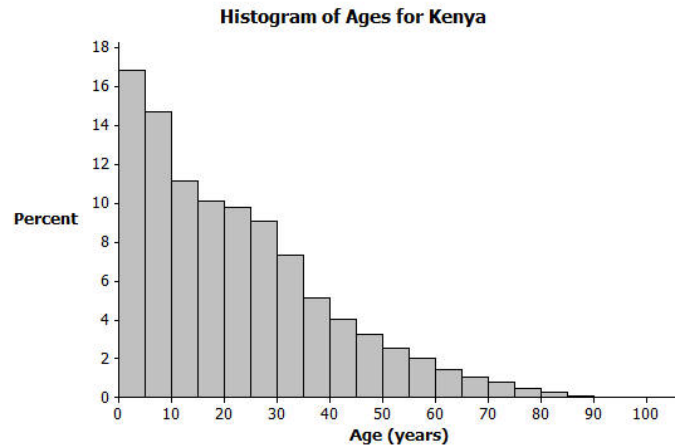
B. Why might understanding the data behind this graph be important?

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?

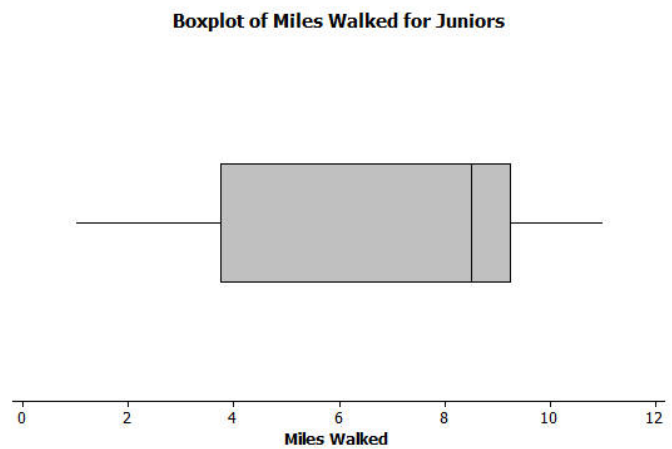
B. Why might we want to study the data represented by this graph?

C. Based on your previous work with histograms, would you describe this histogram as representing a symmetrical or a skewed distribution? Explain your answer.



5. Twenty-two juniors from River City High School participated in a walkathon to raise money for the school band. The following box plot was constructed using the number of miles walked by each of the twenty-two juniors.

A. What do you think the box plot tells us about the number of miles walked by the twenty-two juniors?



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

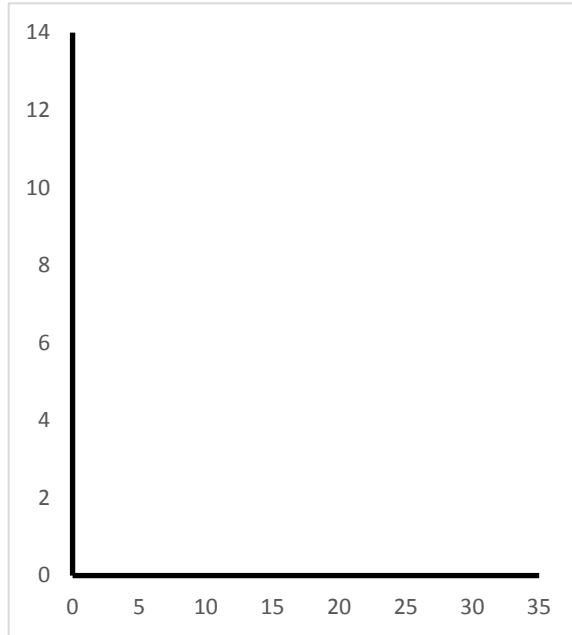
6. Use the data below to create a histogram and a dot plot.

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

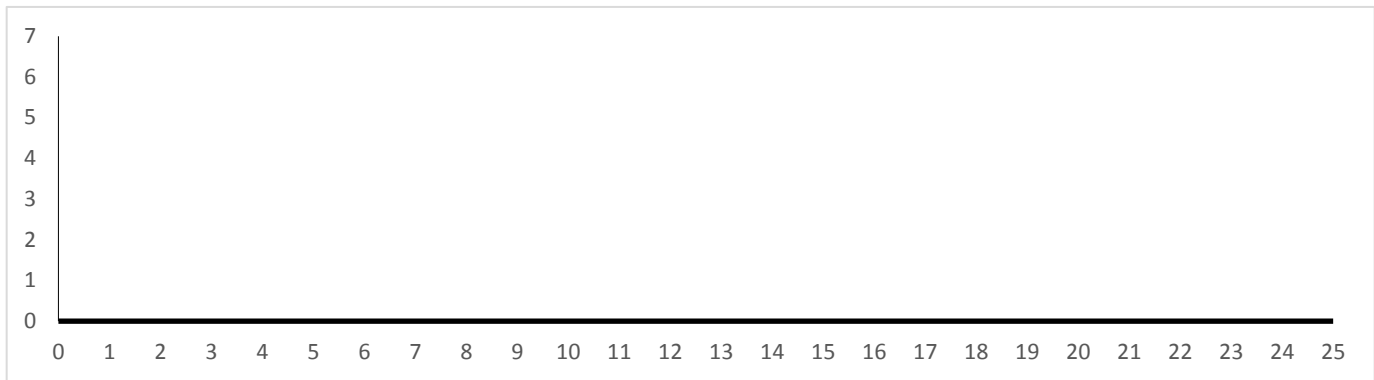
3, 3, 4, 4, 4, 4, 5, 6, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7, 7, 16, 17, 22, 22, 25.

A.

HISTOGRAM



DOT PLOT

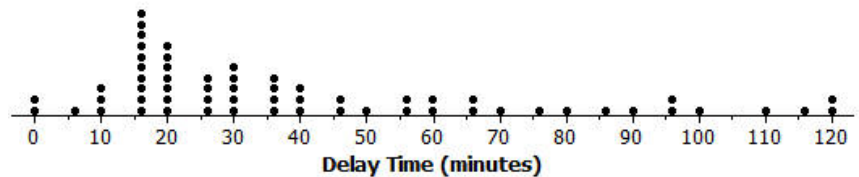


B. Which graph do you think is the most representative of the data? Why?

- C. Why weren't you asked to create a bar graph?
- D. Would you describe the graphs as symmetrical or skewed? Explain your choice.
- E. Identify a typical age of the twenty-five people.
- F. What event do you think the twenty-five people were attending? Use a graph to justify your conjecture.
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.

Dot Plot of December Delay Times



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