SINGLE-VARIABLE DATA

BIG IDEAS:

- **Measures of central tendency** (mean, median, mode) and **measures of spread** (range, interquartile range) are used to analyze **single-variable data**
- **Box plots** provide a visual representation of a single-variable data set's central tendency and spread using its **quartiles**

LEARNING GOALS AND SKILL DEVELOPMENT:
You know you have met the goals for this lesson when you can:

EMERGING

LEARNING GOALS	ANCHOR QUESTIONS
Determine the mean, median and mode of a given data set	1
Identify a specific type of measure as a measure of central tendency or a measure of spread	3
Match the name of a type of graph with its visual representation	4
Suggest an appropriate type of graph to use for a given context	5

SKILL BUILDING QUESTIONS			
1	2	3	4
5	6		

EVOLVING

LEARNING GOALS	ANCHOR QUESTIONS
State the highest and lowest values, median, range, quartile values and interquartile range of a data set from its box plot	7
Make a box plot for a given data set	8, 9, 13
Compare data sets by analyzing their box plots	10, 15
Identify the shapes of various histograms and predict the shape of a histogram corresponding to a specific context	17, 18

SKILL BUILDING QUESTIONS			
7	8	9	10
11	12	13	14
15	16	17	18

EXTENDING

LEARNING GOALS	ANCHOR QUESTIONS
Make generalized statements about mean, median and mode for a histogram with a specified shape	19
Create a modified box plot for a given data set	20

SKILL BUILDING QUESTIONS			
19	20		

BUILD YOUR SKILLS

1. Determine the mean, median and mode for the following data set.

9 12 7 14 7 18 14 9 15 14

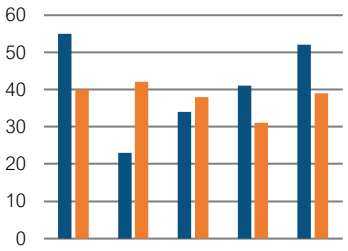
2. Explain the purpose of *measures of central tendency* and *measures of spread*.

3. Identify each of the following as either a measure of central tendency or a measure of spread.

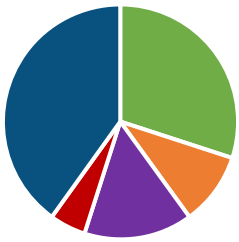
a) mean b) range c) interquartile range d) median e) mode

4. Match the name of each graph type with the corresponding diagram.

a)



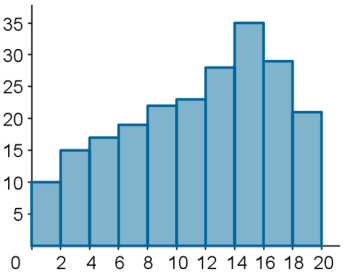
b)



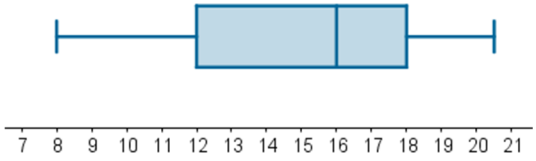
Names:

- circle graph
- double bar graph
- box plot
- histogram

c)



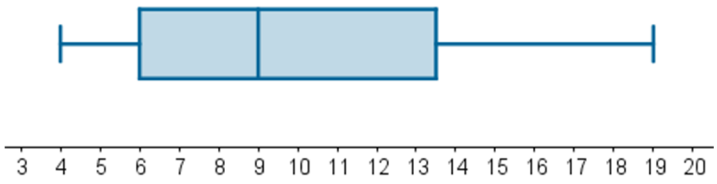
d)



5. Guneet would like to compare the amount of snowfall per month for two Canadian provinces. Which type of graph from question #4 would be most appropriate for displaying this data?

6. Explain the meaning of *single-variable data*.

7. For the data set described by the box plot below, state



a) the median.

b) the greatest value.

c) the lowest value.

d) the range.

e) the value of the first quartile.

f) the value of the third quartile.

g) the interquartile range.

8. Create a box plot for each of the following data sets.

a) 4 4 6 7 7 9 11 12

b) 12.3 8.5 9.6 7.4 10.5 5.2 9.4 8.8 11.2

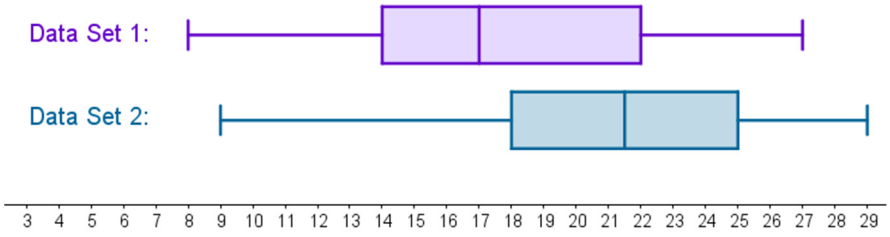
9. The final grades for a biology class are given below.

74	68	81	77	93	84	76	73	88	87	44	78
81	86	72	53	90	47	92	75	73	96	85	66

a) Create a box plot for the data set.

b) Calculate the interquartile range and explain the meaning of this value.

10. Consider the two box plots shown below.



a) Which data set contains the greatest value?

b) Which data set contains the lowest value?

c) Which data set has a greater range?

d) Which data set has a greater interquartile range?

11. Although box plots may display less information than histograms, explain why box plots might sometimes be preferred over histograms.

12. The table on the right shows the number of absences recorded for the students in a class. Create a box plot of this data.

Number of Absences	Number of Students
0	10
1	5
2	4
3	3
4	4
5	0
6	1

13. The table below shows the heights of the 30 girls who tried out for their school volleyball team. Create a box plot of this data.

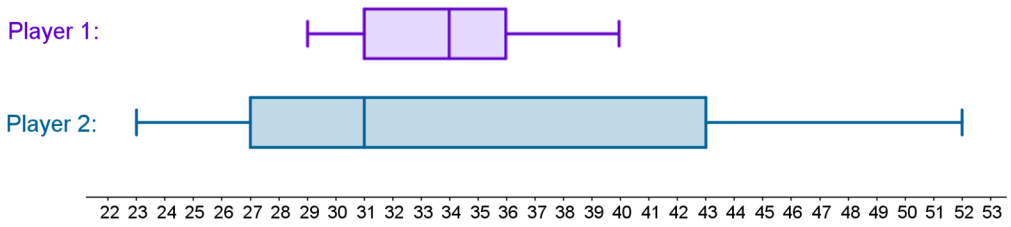
Height (cm)	Frequency
155–160	2
160–165	3
165–170	5
170–175	8
175–180	6
180–185	2
185–190	3
190–195	1

14. Max's annual net income is $75 000. He plans to use $11 250 of this income for transportation.

a) What percentage of Max's annual net income will be dedicated to transportation?

b) If a circle graph is created to display Max's budget, what central angle would be used for the transportation sector?

15. The careers of two basketball players consisted of the same number of seasons. The following box plots are based on their average points per game for each season.

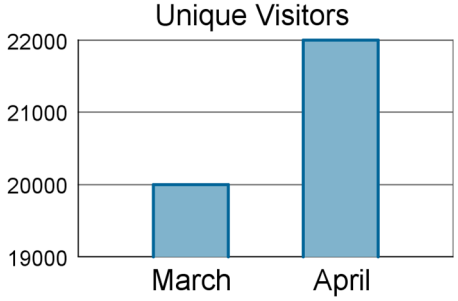


a) Which of the two players had the season with highest average points per game?

b) Which player was more consistent?

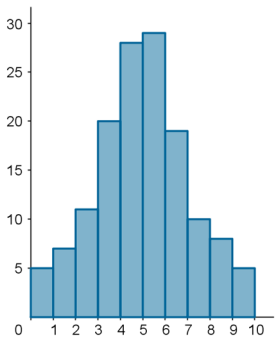
c) Which player had more seasons with an average points per game at or above 34?

16. The bar graph on the right shows the number of unique visitors to a website during the months of March and April. Explain how this graph could be misleading.

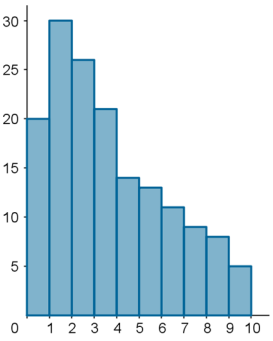


17. Match the name of each histogram shape with the corresponding diagram.

a)



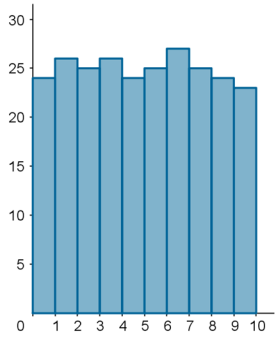
b)



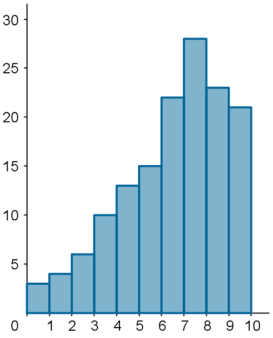
Shapes:

- uniform
- bell shaped
- left skewed
- right skewed
- bimodal

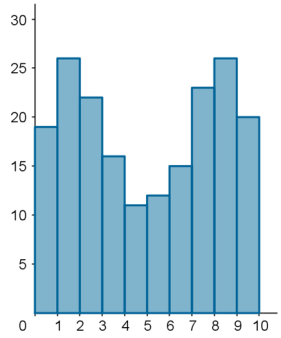
c)



d)



e)



18. Predict the shape of a histogram displaying each of the following.

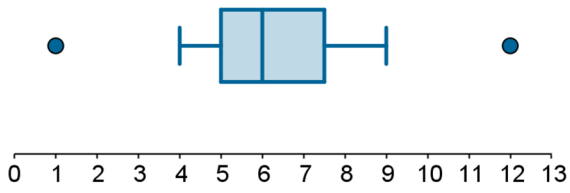
a) The weights of newborn babies.

b) The ages of death from natural causes.

c) The number of customers in a restaurant each hour from 11:00 a.m. to 9:00 p.m.

19. What is typically true about the relationship among the mean, median and mode for a data set that results in a right skewed histogram?

20. A value in a data set is considered to be an *outlier* if it is above the third quartile or below the first quartile by more than 1.5 times the interquartile range. A *modified box plot* is created the same way as a standard box plot, but outliers are displayed as points beyond the "whiskers" (see the example below). The maximum and minimum values that define the modified box plot's "whiskers" come from the data set's greatest and least values that are not outliers. Create a modified box plot for the data in question #9.



ANSWERS

CHECK YOUR UNDERSTANDING

1. mean = 11.9, median = 13, mode = 14

2. Measures of central tendency are values that describe the middle or central value of a data set, whereas measures of spread describe the dispersion or consistency of the values in a data set.

3. a) measure of central tendency b) measure of spread c) measure of spread
b) measure of central tendency e) measure of central tendency

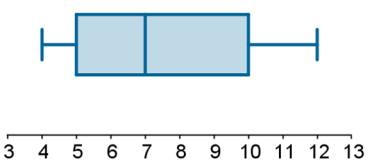
4. a) double bar graph b) circle graph c) histogram d) box plot

5. double bar graph

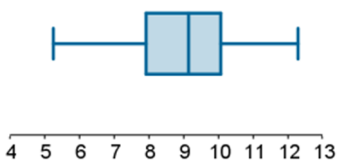
6. Single-variable data is data based on a single characteristic. It is often given as a single list of numbers. The main purpose of single-variable data is to describe an attribute, without giving consideration to the possible causes of a result.

7. a) 9 b) 19 c) 4 d) 15 e) 6 f) 13.5 g) 7.5

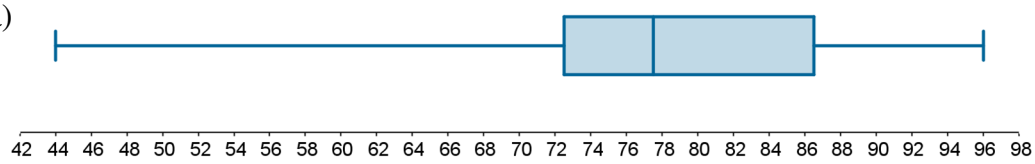
8. a)



b)



9. a)

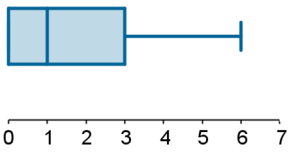


b) The interquartile range is 14, so the middle 50% of the grades have a range of 14.

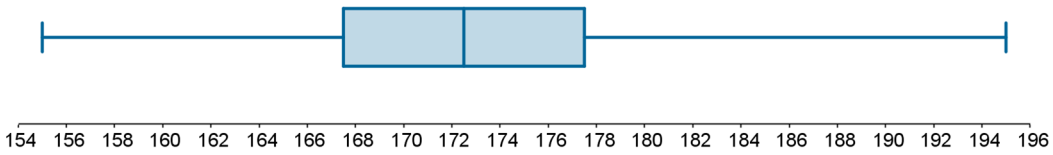
10. a) Data Set 2 b) Data Set 1 c) Data Set 2 d) Data Set 1

11. The key values shown in a box plot (median, minimum, maximum, first quartile, and third quartile) are not clearly displayed in a histogram and are usually impossible to determine exactly. Furthermore, since multiple box plots can be displayed on the same graph, it's much easier to compare several box plots than it is to compare several histograms.

12.



13.



14. a) 15% b) 54°

15. a) Player 2 b) Player 1 c) Player 1

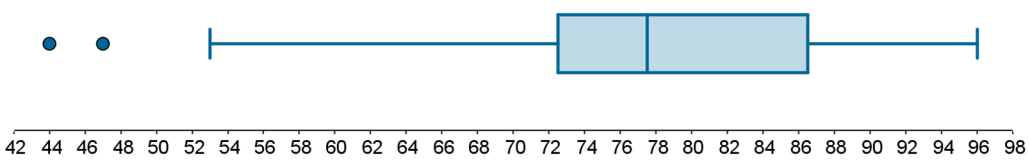
16. Starting the vertical axis at 19 000 instead of 0 makes the increase in unique visitors from March to April appear significantly greater than it actually is. The relative heights of the bars suggest that the number of unique visitors tripled from March to April, but it actually only increased by 10%.

17. a) bell shaped b) right skewed c) uniform d) left skewed e) bimodal

18. a) bell shaped b) left skewed c) bimodal

19. mode < median < mean

20.



ANSWERS