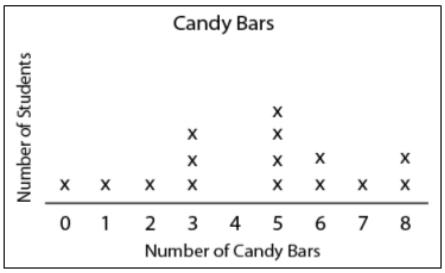
NAME: _____ DATE:_____ SECTION:5___

STUDY GUIDE: DATA

Data are facts and information collected something. Just about data can be anything, but the *way* we choose to represent data is important so that we may read it accurately and easily. We've explored a few ways to collect and record data, so let's look at some examples.

Colour	Tally	Frequency
Red	𝍸 𝍸 III	13
Blue	𝍸 IIII	9
White	𝍸 𝍸 𝍸 𝍸 IIII	24
Black	𝍸 𝍸 II	12
Other	𝍸 IIII	9

A. *Frequency Table*

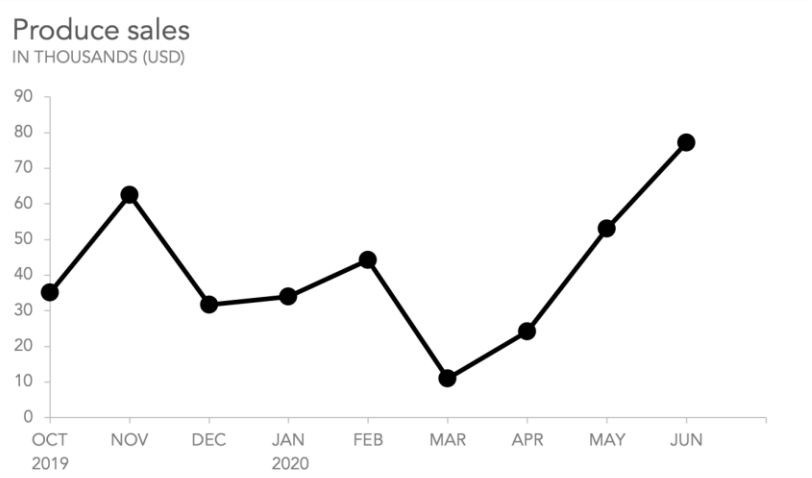


B. *Line Plot*

Above are two examples we have explored which are composed of individual pieces of data or recordings we call ***data points***. On the left is a **frequency table**, where each tally mark represents one data point, and on the right a line plots where a cross represents a data point. In the left table, the column the tally is it represents the category or *bucket* the data falls into, and on the right, numbers do. It's important to **very** carefully read every piece of table or plot – in the right graph, the numbers don't represent people, but according to the horizontal **scale**, represents the number of candy bars sold by how many students, the vertical scale. The number of times each type of data appears is known as the **frequency**.

We also looked at **line graphs,** not to be confused with line plots, which are put onto coordinate planes, which best represent continuous data that changes either over time or the relationship between some input and output. The example on the next page relates each month to the produce sales in thousands of USD (American Dollars).

NAME: _____ DATE:_____ SECTION:5___



This line graph is made of a few parts – what we have learned importantly is that it has two **scales,** or the type and unit of number representing the axes. The horizontal scale here is in months, and its **interval,** or difference between each tick is **one month.** Similarly for the vertical, its scale is in thousands of USD, which an interval of 10. We connect data points, which in this case are coordinates, with a line. For example, our coordinates are (October, 35), (November, 60), (Dec, 30), etc. etc.

So how can we use and interpret this data? We have three **measures of center** we can use to look at the approximate middle of these data. Let's turn figure B from the first page into a list of data from least to greatest, called a **data set.**

$$0, 1, 2, 3, 3, 3, 3, 5, 5, 5, 5, 6, 6, 7, 8, 8$$

We can calculate the **mean** (**μ**) by summing all these data points' values, then dividing by how many we have, in this case 15, and round to the nearest tenth.

$$\mu = \frac{0,1,2,3,3,3,3,5,5,5,5,6,6,7,8,8}{15} = \frac{67}{15} \approx 4.5$$

Next is the **median**, or the center value in a data set. We can find this by eliminating one piece of data from each end of the **ordered** set until we have one or two data points left. If only one data point is left, that data point is the mean. However, if there are **two** data points left in the center, then we must add those two together and divide by 2. Let's try it on this same data set, step by step.

NAME: _____ DATE:_____ SECTION:5___

0,1,2,3,3,3,5,5,5,5,6,6,7,8,8
0,1,2,3,3,3,5,5,5,5,6,6,7,8,8
0,1,2,3,3,3,5,5,5,5,6,6,7,8,8
...
0,1,2,3,3,3,5,5,5,5,6,6,7,8,8

As we can see in this example, our median is 5.

Finally, we have the **mode**. The **mode** is simply the data point that occurs **the most frequently** in a data set. Data can have one mode, multiple if more than one number occurs equally as much, or even no mode if none of the numbers repeat. In our data set, the number that occurs the most is 5, therefore it is our mode as well.

Another piece of information that's helpful to know is our data's **range.** This is **NOT** a measure of center, but is useful so that we can understand how large or small our data sets are. The range is a simple calculation: it is the **difference** between the **largest** and **smallest** data points in our sets. In our data set, the smallest value is 0, and the largest is 8. Therefore, our range is 8…

$$Range = 8 - 0$$
$$Range = 8$$

NAME: _____ DATE:_____ SECTION:5___

Your Turn

Use the data below to answer questions 1 through 3.

The list of data below shows the ages of a small 5th grade classroom's students.

9, 9, 9, 9, 9, 9, 10, 10, 10, 10, 10, 10, 10, 10, 10, 11, 11, 11, 11, 12

1. Use the space below to create a **line plot** of the ages.

2. Calculate the mean (round to the nearest tenth if necessary), median, mode, and range of the data.

3. Is "what is your age" a survey question or observation? Do you believe this to be a good / reasonable question, or a bad question? Why?

   _____

   _____

   _____

   _____

   _____

NAME: _____ DATE:_____ SECTION:5___

The following table shows how the highest yearly water level in a lake over the course of a few years. Use the data below to answer questions 4 through 6.

Year	2015	2016	2017	2018	2019
Depth (ft.)	$10\frac{1}{2}$	$7\frac{1}{3}$	9	$9\frac{2}{3}$	11

4. Between which two years did the water level change the most?

5. Calculate the mean and range of the water level in the lake over the years listed.

6. Use the blank coordinate grid on the next page to create a line plot of the data listed above. Label your axes with the appropriate scales and choose reasonable intervals for each axis. **Before plotting your points,** list the coordinates on the lines below.

_____

_____

_____

_____

_____

_____

_____

_____

NAME: _____ DATE:_____ SECTION:5___