

CHAPTER 4: DATA MANAGEMENT



Objectives:

- a. Classify organizing data and presentation of data
- b. Present data collected using textual method, tabular method and graphical method.
- c. Apply the basic ideas and concerns about the processes used to obtain sample data.

Lesson 1: Data Gathering and Organizing Data

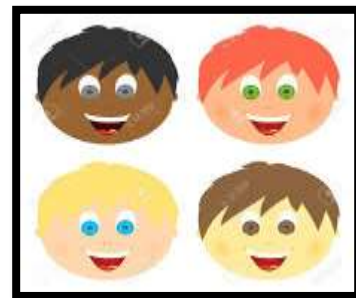
Data

- is a collection of facts, such as numbers, words, measurements, observations or just descriptions of thing.

Two Categories of Data

- Qualitative Data
 - Also called attribute data are typically associated with inherent characteristics of the items under consideration. However, attribute data are not easy to analyze using numerical techniques.

Example: If we are studying human beings, the following characteristics may be considered attributes: color of the hair, race, color of the eyes, types of car driven, etc.



- Quantitative Data
 - Also called as numerical data result from a process that quantifies which includes counts of (how many) or measurements (length, weight, and so on).

Classification of Quantitative data

- **Discrete numerical data**
 - data that represents counts

Example: The number of people living in Brgy. San Isidro
 The score in a basketball game
 The number of books in ICCT Cainta Library

- **Continuous numerical data**
 - data that represents measurement

Example: IQ scores of the students in Ms. Berces' Math class
The height of the basketball player
The volume of weekly garbage produced by Brgy. 123

Classification of Data using the four levels of measurement

- **Nominal Level**

- Is characterized by data that consist of names, labels, or categories only. The data cannot be arranged in an ordering scheme.

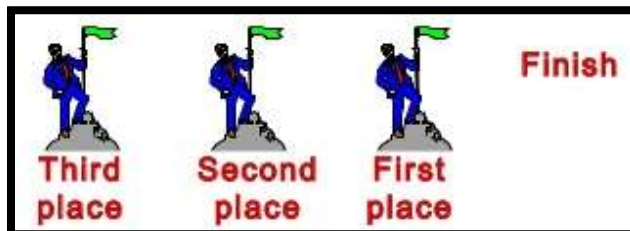
Example: Category of shirt in terms of color



- **Ordinal Level**

- Involves data that may be arranged in some order but differences data values either cannot be determined or are meaningless.

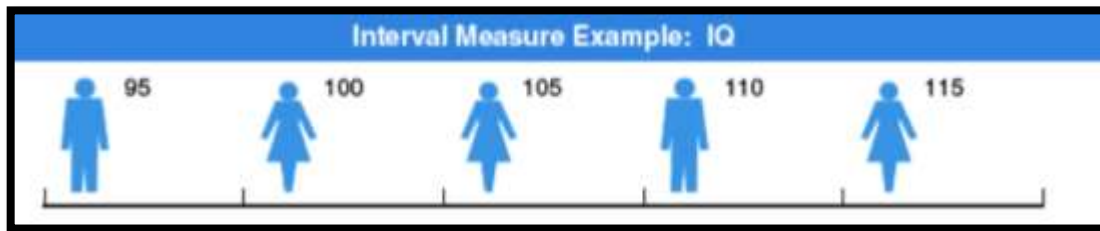
Example: Rank order of winners



- **Interval Level**

- Is like the ordinal level but meaningful amounts of differences between data can be determined. It has no inherent (natural) zero starting point where the none of the quality is present.

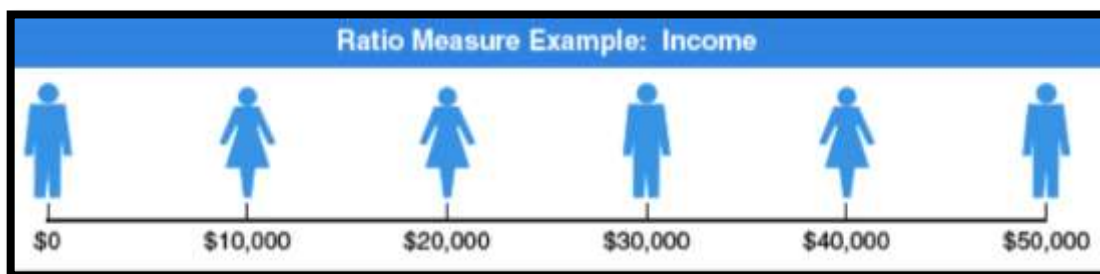
Example: Interval of Student's IQ



- **Ratio Level**

- Is the interval level modified to include the inherent zero starting point where zero indicates that none of the quantity is present.

Example: Income of employee



Sources of Data

The analysis of data begins with the acquisition of data. For the analysis of data to be meaningful, the data must be a representative of the population under consideration. The art of gathering data is called **sampling**.

Most Common Sampling from a Population:

1. Random Sampling

- In this sampling technique, each member of the population has the same chance of being chosen, with no one member being favored over any other.

Example:



In a card game, random samples of cards are distributed to each of player by means of shuffling deck.

2. Systematic Sampling

- This is typically used when a list of the population is readily available. The samples are determined by choosing every n^{th} item on the list until the desired number of samples are obtained.

In determining the n^{th} item we have to:

- Number the item from 1 to N ;
- Decide on the sample size you want;
- Get the interval, that is, the n^{th} item. Divide the population by the

$$n^{\text{th}} = \frac{N}{m},$$

Where N is the population and m is the desired number of sample

- Randomly select the sample taken every n^{th} item.

Example:

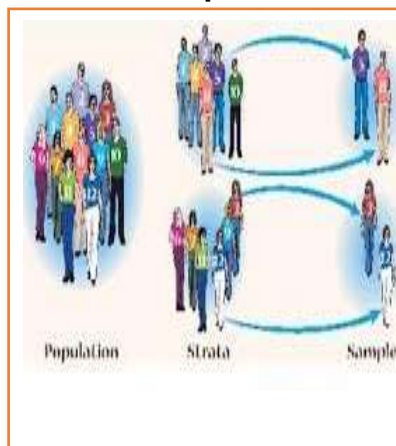


If ABC Corp. wanted to conduct survey of its 15000 employees, it could begin with a complete roster, then select every 100th employee to obtain a sample size of 150.

3. Stratified Random Sampling

- In this technique, the population is divided into sub-groups or strata and then a small sample is taken from each stratum. Each subsample is then drawn in equal numbers from each stratum. An extension of this method is the *proportional* method where the number of observations chosen from each stratum is proportional to the size of the stratum.


Example:



Suppose you want to draw a sample size of 36 from a college course which has an enrolment of 900 students. Assume that this population of 900 students comprises 200 taking marketing majors, 300 economics major and 400 accounting majors. Then using the proportion method, you would choose 8 marketing majors, 12 economic majors, and 16 accounting majors from the respective groups to form a sample.

4. Cluster Method

- The population is divided into strata, except that now certain strata are selected randomly, and the sample is chosen from only those strata.

Example:

In a pre-election poll, we randomly select 50 election precincts and then survey all the people from each of those chosen precincts.

Collection of Data

The validity of the conclusion drawn from a sample is determined based on how the data are collected or gathered. These data are taken in several ways, some are listed below:

1. Interviews and Focus Groups

Interviews and focus groups are often used to gather detailed and qualitative descriptions. An interview is conducted on one-on-one basis while focus groups are conducted in small groups. In these methods, questions are generally open-ended and responses are documented in thorough and detailed notes or transcriptions.

**Advantages:**

- a. It provides a rich data that can picture a broader aspect.
- b. It may highlight issues, which are not previously considered, or information that are useful to further interpretation.
- c. It may increase the comfort level of the respondents.

Disadvantages:

- a. It can be difficult to elicit participation from an individual who has time constraints.
- b. Interviews and large groups may intimidate some participants.
- c. Documentation and analysis can be time consuming and may even require experts for a certain qualitative analysis.

2. Surveys and questionnaires

These are indirect and commonly used methods of gathering data. Surveys and questionnaires are collected by the evaluator to gather specific information from respondents. These are usually administered through papers in a structured or semi-structured format which can be done through mail, over the phone or via internet.

Respondents often choose from the set of forced choices or provided responses. These includes yes/no or scaled responses.



Advantages:

- a. It is less time consuming and less expensive to administer.
- b. It can be administered to a larger group of individuals.
- c. It is effective for assessing program satisfaction.

Disadvantages:

- a. Data entry and analysis can be time consuming.
- b. It is quite difficult to receive complete survey.
- c. Answers of the respondents may not be reliable.

3. Observations

Observation are generally unobtrusive method for gathering information about a certain concern or interest. They are usually conducted by external evaluators or researchers and are often used to verify and supplement information gathered through other methods.

Observation can be highly structured with protocols for recording specific behavior at specific times. They are most reliable when they are conducted over a period of time to minimize the chances of observation days being unusual.



Advantages:

- a. Trained evaluators and researchers may provide less biased descriptions.
- b. Provide a highly detailed information from an external perspective on what actually occurs.

Disadvantages:

- a. It can be time consuming, labor intensive, and expensive.
- b. Observers must be trained and be consistent with one another.

4. Registration Method

This method of gathering data or information is enforced by certain laws. In this method, the information are kept, systematized, and made available to all for future references. Examples are the registration of births, marriage, death, vehicles, licenses, and many others.



5. Experiment Method

This method is used when the objective is to determine the cause-and-effect relationship of certain phenomena under controlled conditions. An equal number of respondents with relative similarities are examined to distinguish the effects of whatever are done to the experimental group. This is usually conducted on scientific researches.

Experimental method is sometimes difficult to conduct because it is not that easy to consider respondents with the same characteristics.





For more knowledge about Data Gathering and Organizing data, please check the link provided;

<https://www.showme.com/sh/?h=W6CI5Uu>

<https://www.siyavula.com/read/maths/grade-10-mathematical-literacy/data-handling/12-data-handling-05>

REMEMBER



- Data is categorized into two; qualitative and quantitative.
- Data is classified into four level of measurement: nominal, ordinal, interval and ratio.
- Gathering data is done through sampling: Random sampling, Systematic sampling, Stratified random sampling, and cluster method.
- Collecting data is done through: Interviews, Surveys and questionnaires, Observations, Registration method and Experiment method.



APPLICATION

ACTIVITY: Getting to know me

Choose your partner and ask your partner about her personal information; including family background, likes and dislikes.

Lesson 2: Presentation of Data

Collected data must be organized for effective presentation to add analysis and interpretation. There are basic rules which should be followed once a set of data has been collected and is ready for presentation. The person who is reading the figures should be convinced that the data presented are both factual and relevant.

These collected data that can be presented in several forms.

1. Textual or Narrative Presentation

The textual presentation combines text and figures in a statistical report. It aims to direct the reader's attention to some data which need particular emphasis and probably comparison.

As a method of presenting data, the textual presentation is not that effective since it may not even give the reader a good grasp of the meaning of the quantitative relationship in the presentation.

Example:

Data Collected:

Consider an investigation by a large pharmaceutical company which needs information regarding the smoking habits of its male and female employees according to their ages—under or over 25. Information is to be collected from 8 690 employees based in the Visayas and from the 4270 employees based in Mindanao.



Narrative or textual form:

“Of the 2570 male employees in the Visayas who smoked, 1200 were under 25 and 1370 were 25 or over. The figures for the 2000 female Visayas employees were 850 under 25 and 1150 aged 25 or over. In the Mindanao laboratory, of the 1470 male smokers, 850 were under 25 and for the 90 female smokers 720 were under 25. In the Visayas there were 2300 nonsmoking males, of whom 1500 were aged under 25, and 1820 nonsmoking females, of whom 970 were aged under 25. In Mindanao there were 1200 nonsmoking males, of whom 480 were under 25 and 630 nonsmoking females, of whom 350 were under 25.”

2. Tabular Presentation

This method is better than the textual method because it provides numerical facts in a more concise and systematic manner.

Statistical tables are used to facilitate comparison and analysis of relationships. There is an orderly arrangement of numerical facts in columns and rows. Although most tables are self-explanatory, a clear and brief statement is necessary to call attention and convey understanding.

Some rules to be followed in constructing a table.

1. Every table should have a clear and unambiguous number. It is a common practice to use a decimal numbering system such as 9.1, 9.2, 9.3, ..., 9.9 for tables which are either in the ninth chapter of your work or if there are several tables which relate to the same topic.
2. A title must be given that includes what the table is all about.
3. Row and column labels should be concise and unambiguous.
4. Categories should be exhaustive and mutually exclusive.
5. The units of measurement must be clearly stated.
6. Label any subheadings clearly.
7. Combine any unimportant figures or omit them.
8. Show any totals, subtotals, percentages, etc.
9. If relevant, check that the column and row grand totals are equal.

Example:

The figures for the data in Example 1 when presented in narrative form are clearly incomprehensible. The patterns among them are almost impossible to distinguish. The same data, if presented in tabular form would resemble in the table below.

Table: Smoking Habits of the Employees in the Visayas and Mindanao Laboratories

AGE	SMOKERS		NONSMOKERS	
VISAYAS LAB	MALE	FEMALE	MALE	FEMALE
Under 25	1200	850	1500	970
25 and over	1370	1150	800	850

TOTAL	2570	2000	2300	1820
MINDANAO LAB	MALE	FEMALE	MALE	FEMALE
Under 25	850	720	480	350
25 and over	620	250	720	280
TOTAL	1470	970	1200	630

3. Graphical Presentation

Graphs, in general, do not present as much data in detail as do tables. When data are shown in terms of visually interpreted illustrations, the reader easily sees essential facts and relationships. The significant proportions, differences or terms can easily be grasped.

Advantages:

- a. It attracts the attention of readers more effectively than tables.
- b. The use of colors and pictorial diagram make the list of figures more meaningful.
- c. It gives a comprehensive view of quantitative data.
- d. Graphs can be grasped and understood easily. In fact, other relations not readily seen are easily discovered.

Limitations:

- a. It does not show as much information that maybe needed.
- b. Graph does not show accurate data as tables do.
- c. Charts require more skill, more time and are expensive to prepare.
- d. Graphs cannot be quoted in the same way as tabulated data.
- e. Graphs can be made only when data have been tabulated.

Pertinent information can be obtained from a graph if only we know how to interpret it. The following are some basic rules to follow when illustrating data by means of a graph.

1. The graph should have a clear title.
2. Axe should be clearly labeled.

3. Do not put too many curves on the same graph.
4. Quote all sources of data.
5. If possible, accompany the graph with the table of data it represents.
6. A zero line for the vertical axis should be used if possible.

A graph is visual representation of a table. These presentations are designed to meet particular needs for which they are constructed. Graph also has type which can be used in presentation of data.

A. Line Graph

A **line graph** is used to represent changes in data over a period of time. Data like changes in temperature, income, population and the like can be represented by a line graph. In line graph, data are represented by points and are joined by a line segment. A line graph may be curved, broken or straight.

Generally, the horizontal axis is used as the time axis and the vertical is used to show the changes in the other quantity.

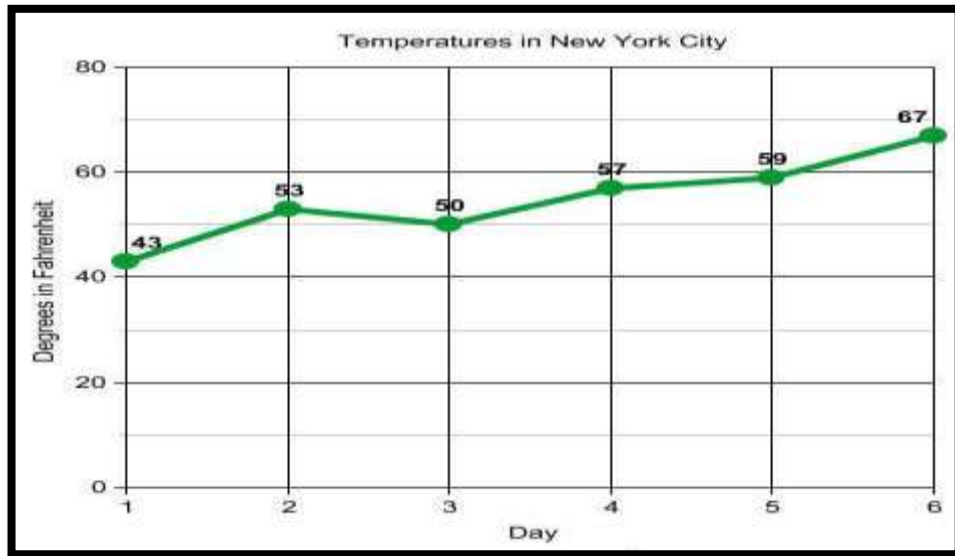
Here are some guidelines to follow in constructing a line graph.

1. The title of the graph must be descriptive and clearly stated.
2. Label the x- and y-axes with the unit these variables and scale represent.
3. The zero point should be an indicator.
4. For a multiple line graph, one line should be distinguished from the other. Using different colors, line forms, or making a legend would facilitate the understanding of the graph.
5. The graph should not be loaded by too much curves.

Example: The table below shows daily temperatures for New York City, recorded for 6 days, in degrees Fahrenheit.

Temperatures In NY City	
Day	Temperature
1	43° F
2	53° F
3	50° F
4	57° F
5	59° F
6	67° F

The data from the table above has been summarized in the line graph below.



B. Bar Graph

Bar graphs consist of series of rectangular bars where the length of the bar represents the magnitude to be demonstrated. A *simple* bar graph is one in which each bar represents one unit only. A *multiple* bar graph is one where different units are used on the same diagram in order to emphasize comparisons. Bar graphs may be vertical or horizontal.

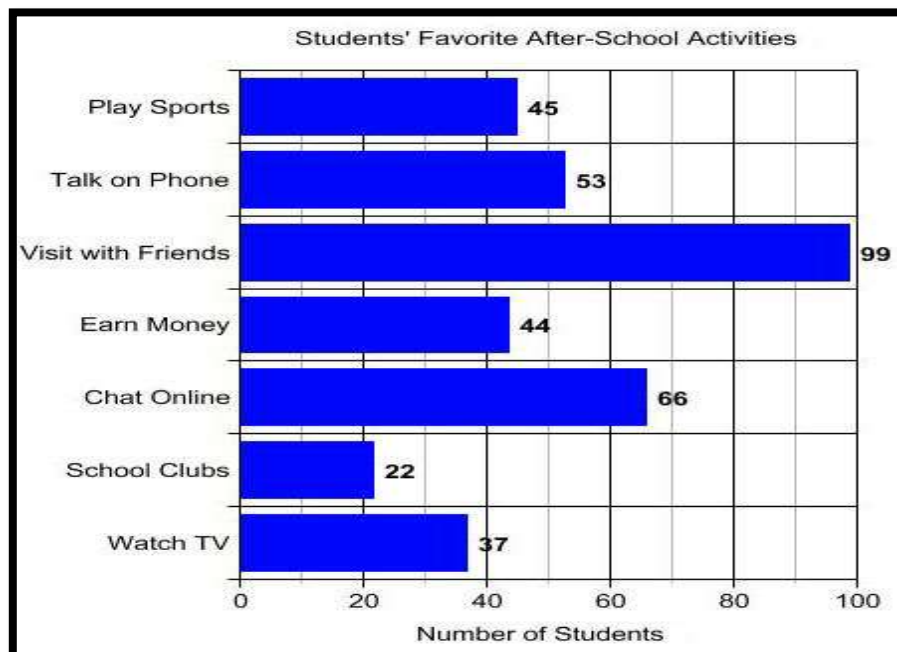
Some Guidelines in Constructing a Bar Graph

1. The title must be clearly stated. It is usually written on top of the graph.
2. Determine the longest and smallest number in the given data and choose a unit of length such that the largest bar is within the space of the paper.
3. The space between any two consecutive bars must be the same but not less than half of the width of the bar.
4. The bar must be of equal width and the lengths must be proportional to the numbers they represent.
5. All designations or notations on the vertical and horizontal axes must be indicated.
6. Show the zero point. If in case a zero point and a point on the scales are irrelevant, you may omit those numbers by breaking the scale. The symbol $_$ is used instead.

Example: A survey of students' favorite after-school activities was conducted at a school. The table below shows the results of this survey.

Students' Favorite After-School Activities	
Activity	Number of Students
Play Sports	45
Talk on Phone	53
Visit With Friends	99
Earn Money	44
Chat Online	66
School Clubs	22
Watch TV	37

Note that since the data in this table is not changing over time, a line graph would not be a good way to visually display this data. Each quantity listed in the table corresponds to a particular category. Accordingly, the data from the table above has been displayed in the bar graph below.



C. Circle Graph/ Pie Chart

A circle graph or pie chart consists of a circle divided into sections whose sizes are proportional to the magnitudes or percentages they represent. The circle or pie chart aims to show percent distribution of a whole into its component part. It is very effective in presenting financial data. The effect of pie chart is particularly strong if the relation between the various components is needed for comparison. The slices or sectors in a pie chart allow easy visual impression about the relative sizes of the component parts.

The following should be observed in constructing a pie chart or circle graph.

1. Express each component part as a certain percent of the whole and multiply the result by 360 degrees.
2. Mark off the desired number of degrees on the circumference of a circle using a protractor. Connect this point to the center in order to produce pie-shaped areas which make up the whole circle.
3. Label the sectors horizontally with the percentages indicated.
4. Arrange the sectors clockwise with the largest sector first. This sector should begin by drawing a vertical line from the center of the circle to the twelve o'clock position on the circumference.
5. Avoid overloading the chart by showing too many categories.

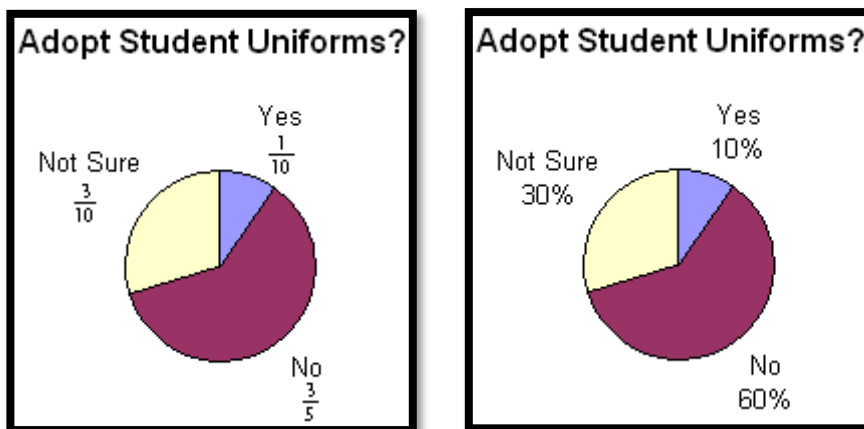
Example: At a private school, 300 students and faculty voted on adopting uniforms for students. The results are shown in the table below. Display the results of this vote in a circle graph.

Adopt Student Uniforms?	
Response	Number
Yes	30
No	180
Not Sure	90

Analysis: In order to draw a circle graph, we need to represent the number for each response as a fraction or as a percent.

Adopt Student Uniforms?			
Response	Number	Fraction	Percent
Yes	30	$\frac{1}{10}$	10%
No	180	$\frac{3}{5}$	60%
Not Sure	90	$\frac{3}{10}$	30%

Solution: The results of this vote have been displayed in the two circle graphs below. In the graph on the left, fractions are used to label the data. In the graph on the right, percents are used to label the data.



As you can see, a circle graph is easier to read when a percent is used to label the data.

D. The Picture Graph or Pictograph







The pictograph sometimes called the pictogram is another device in presenting data by means of pictures or symbols. It is an adaptation of the bar graph. Its symbol called **isotype** represents specific number of people or objects. Its purpose is to catch the attention of readers and to convey a clear analysis of the numerical facts.

How to construct a pictograph:

1. The symbols used must be self explanatory. If the data pertain to people, then the pictures of people may be used.
2. Use a legend to represent the number of units in a picture.

3. Round off numbers to eliminate fractions. If the fractional parts are to be presented they must be done proportionally.
4. Use symbols that are of the same size and space them equally.

Example: Number of red colored boxes sold by William, a shopkeeper, in six days of a week. See the picture graph or pictograph to answer the questions.

Days	Number of Red-Boxes Sold
Monday	
Tuesday	
Wednesday	
Thursday	
Friday	
Saturday	

Information gathered from the above table:

(i) Number of red boxes sold:

Monday – 4, Tuesday – 2, Wednesday – 3, Thursday – 5, Friday – 8,
Saturday – 1

Therefore, sale during the week = 23

REMEMBER



- Presenting of data is done through; Textual presentation, tabular presentation and graphical presentation.
- Graphical presentation includes bar graph, line graph, and circle or pie graph.



APPLICATION

ACTIVITY:

In a 1 whole paper, create a presentation of the following data:

- a. Pie chart of how many percent you spent in:
 - Transportation
 - Food
 - Clothes
 - Electric bills
 - Internet/load

Use different color to represent the data.



REFERENCES

Math Power textbook, p.365

<https://www.khanacademy.org/math/pre-algebra/pre-algebra-math-reasoning/pre-algebra-representing-data/v/ways-to-represent-data>