Data Visualization in Research Report

Pre - Ph.D. Course Work

Paper I

Prof. (Dr.) Anju Srivastava
Department of Home Science
Patna University
srivastava.anju@yahoo.com
9334121405

Data Visualization in Research Report

Effective data visualization is the crucial final step of data analysis because:

1. **Raw data does not always tell the most compelling story**
2. **Words do not always paint the clearest picture**
3. **Human mind is very receptive to visual information**
4. **A visual summary of information makes it easier to identify patterns and trends instead of looking through thousands of rows on a spreadsheet**
5. **A picture will tell thousand words so a good chart or graph can show as much as several paragraphs of words**
6. **Humans are biologically hardwired to engage and pay attention to great visuals therefore, images speak louder than words**
7. **Human mind processes visuals more effectively and efficiently than words**

Data visualization

- Is the visual presentation of data
- What cannot be touched, smelled or tasted can be represented visually
- Involves presenting data in graphical or pictorial form which makes the information easy to understand
- A tool that provides an accessible way to see and understand trends, outliers and patterns in data
- Refers to transforming figures and raw data into visual objects
- Is the practice of converting data from raw figures into a graphical representation
- Is a key part of data analysis

Significance of Data Visualization

1. They give a birds eye view of the entire data, therefore, the information presented can be clearly understood
2. They have a great memorizing effect. The impressions created by diagrams last much longer than those created by figures presented in a tabular form
3. Can be creative and pleasing to look at
4. Makes data is engaging and easily digestible
5. Highlights the important points/parts in a set of data
6. Easy to understand – diagrams are usually attractive and impressive which makes data easier to understand

Cont…

7. **Simplified presentation – large volumes of complex data can be represented in a simplified and intelligible form**
8. **They facilitate comparison of data. They help in making quick and accurate comparison of data**
9. **Well designed data graphics are usually the simplest and at the same time most powerful**
10. **Universally accepted – almost all fields of study like business, economics, social institutions, administration etc. use diagrams. Therefore, they have universal acceptability**
11. **As we have more information at our fingertips than ever before, the importance of data visualization has never been greater than it is right now**

General Principles of Diagrammatic Presentation of Data

A diagrammatic presentation is a simple and effective method of presenting the information that any statistical data contains. Following are some principles which can make them more effective tool of understanding that data:

1. **Write a suitable title on top which conveys the subject matter in brief and unambiguous manner. More details about the title can be mentioned in the footnote.**
2. **You must construct the diagram in a manner that immediately attracts the viewer. Ensure that you draw it neatly with an appropriate balance between its length and breadth.**
3. **Further, make sure that the diagram is neither too large nor too small**
4. **You can also use different colors or shapes to emphasize different aspects of the problem**

Contd…

5. **Select the design of a diagram carefully keeping in view the nature of the data and also the objective of investigation.**
6. **If you use different shades of color to depict different characteristics in the diagram, then ensure that you provide an index explaining them**
7. **If you are using a secondary source, then ensure that you specify the source of data**
8. **Try to keep your diagram as simple as possible**

Techniques of Data Visualization

- **To create a story with your data there are many types of visualizations that you can choose. Your choice of visualization should depend on what you are trying to convey.**
- **Choosing the right type of visualization depends on what you need to show (comparison, distribution, composition, or relationship), how much details the audience needs and what information the audience needs in order to understand the report**
- **Data can be represented in many ways. The five most common are: Line graph, Bar graph, Histogram, Pie chart, Pictograph**

Bar Graph

- **A bar chart or bar graph presents categorized data with rectangular bars with height or length proportional to the value that they represent**
- **The bars can be plotted vertically or horizontally**

Uses:

1. **Simple and most straight forward way to compare various categories**
2. **When bars are stacked next to each other, viewers can compare the different bars or values at a glance**
3. **Works great for visually presenting nearly any type of data**
4. **Useful in showing relationship between independent and dependent variables**
5. **Ideal for comparing any sort of numerical value, including group sizes, inventories, ratings and survey responses**

Bar Graph

Figure 5: Example of a Bar Graph

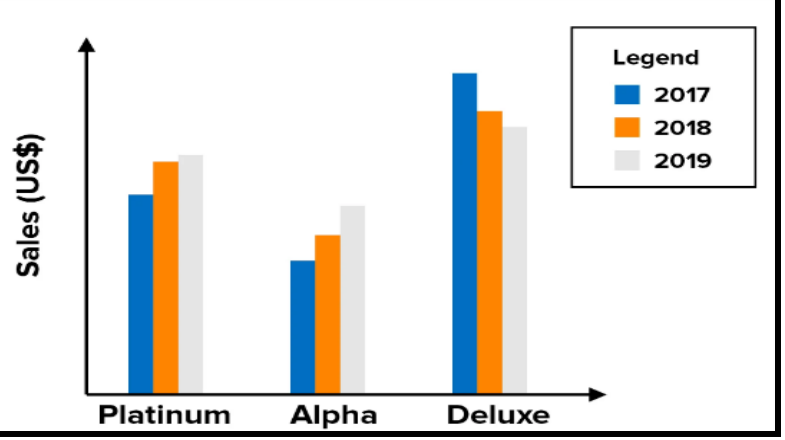
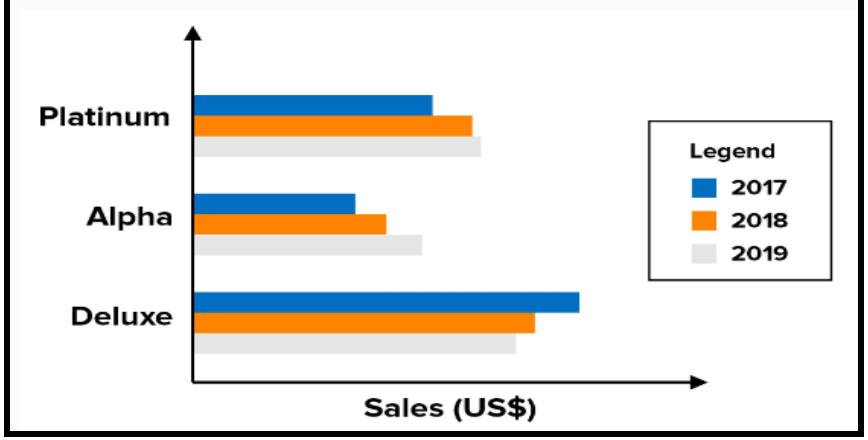
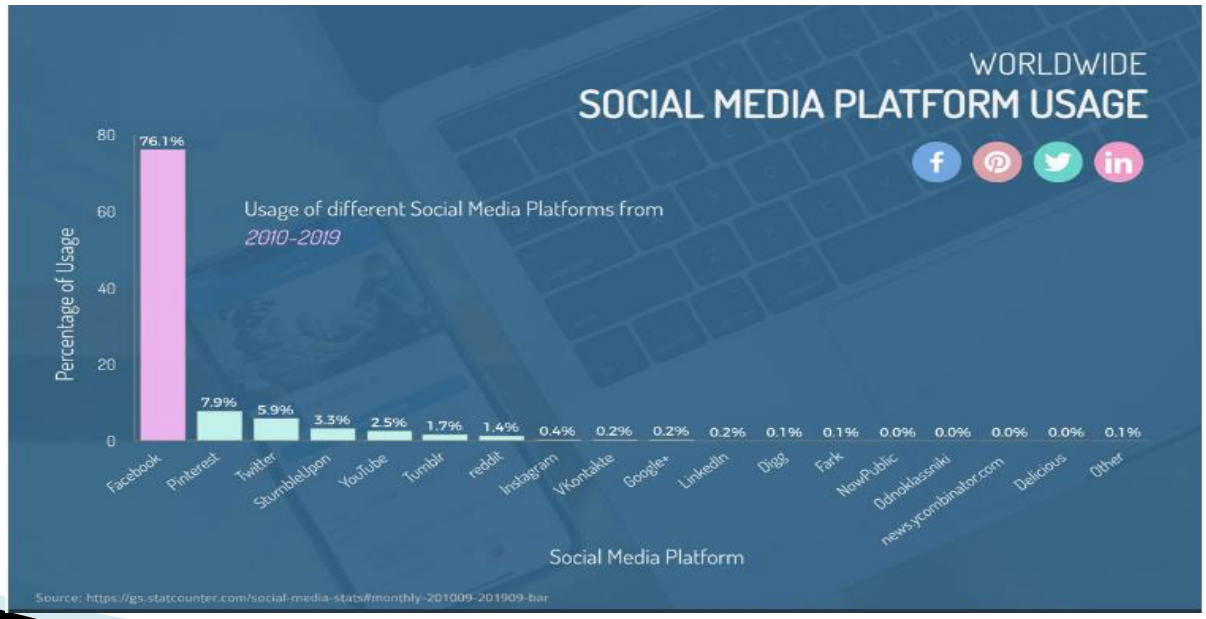
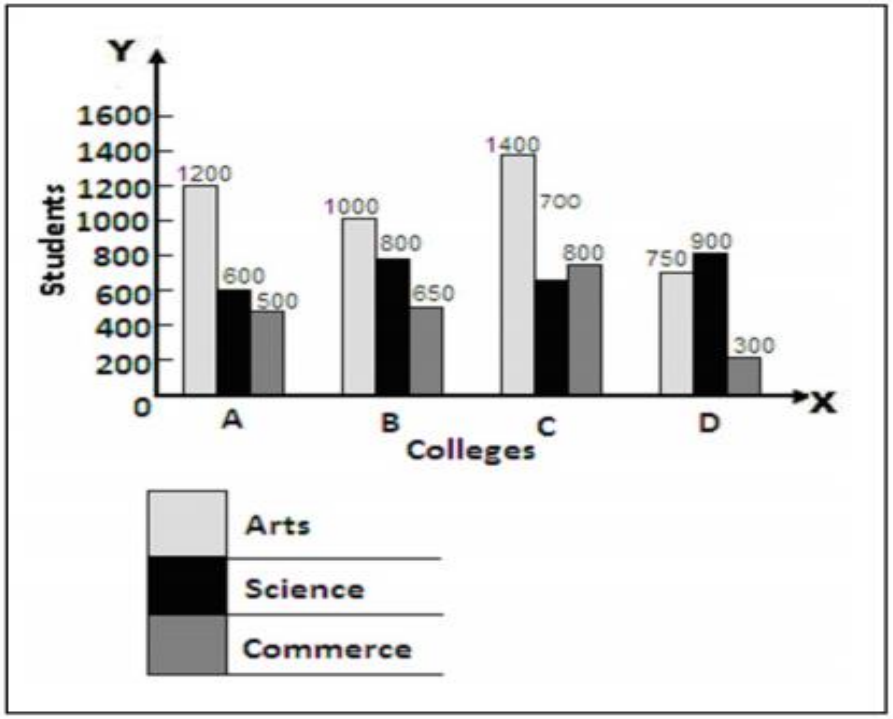


Figure 7: Example of a Horizontal Bar Graph

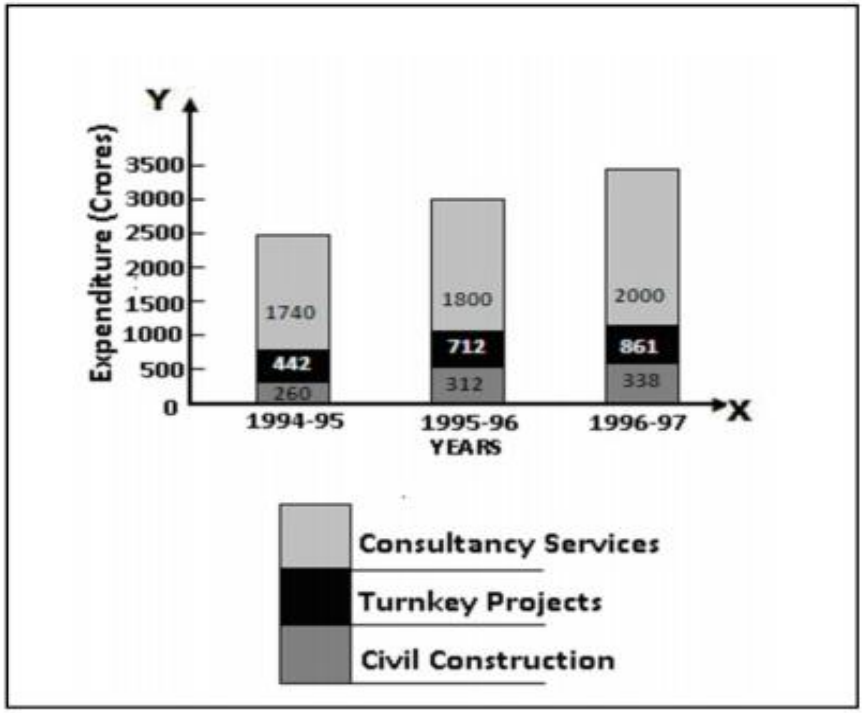




Bar Graph Contd...



Multiple Bar Diagram



Component Bar Diagram

Line Graph

- **A graph with points connected by a line to show how something changes in value**
- **Displays information as a series of data points called ‘markers’ connected by straight line segments**

Uses:

1. **Mostly used when you want to show trends. Multiple trends can be compared by plotting lines in different colors**
2. **When you want to make predictions**
3. **When comparing two or more different variables, situations and information over a given period of time**

Line Graph

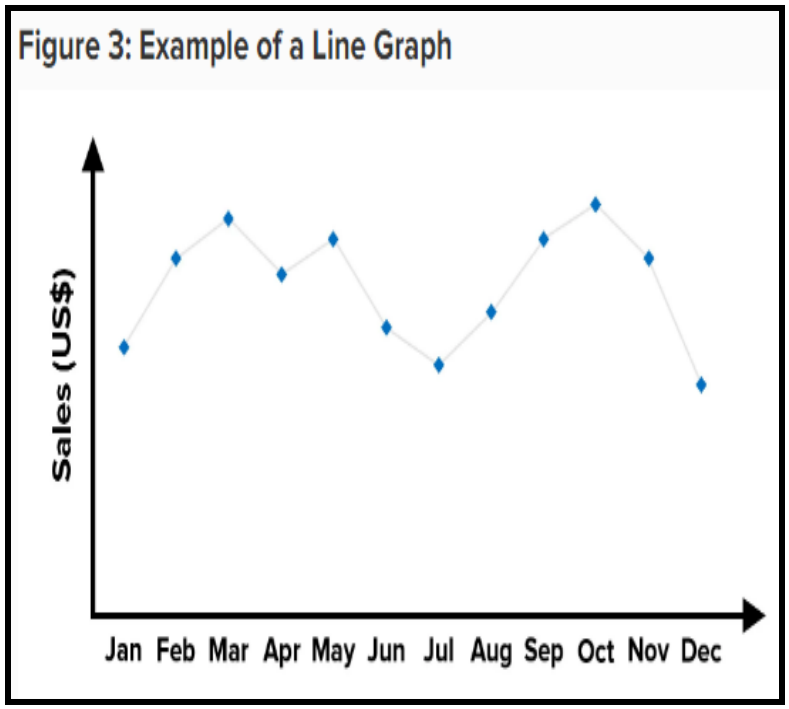


Figure 3: Example of a Line Graph
Sales (US$)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

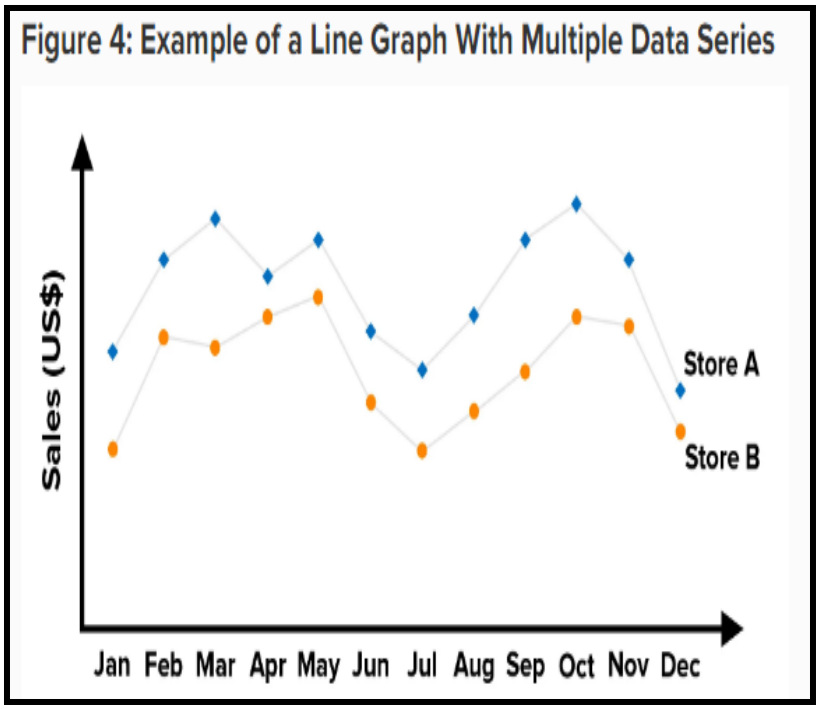


Figure 4: Example of a Line Graph With Multiple Data Series
Sales (US$)
Store A
Store B
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

Pie Chart

- A pie chart is a circular graph that is broken into segments (i.e. slices or pie)
- The chart as a whole represents the sum of all its data; individual slices show each piece of data as a percentage of the whole

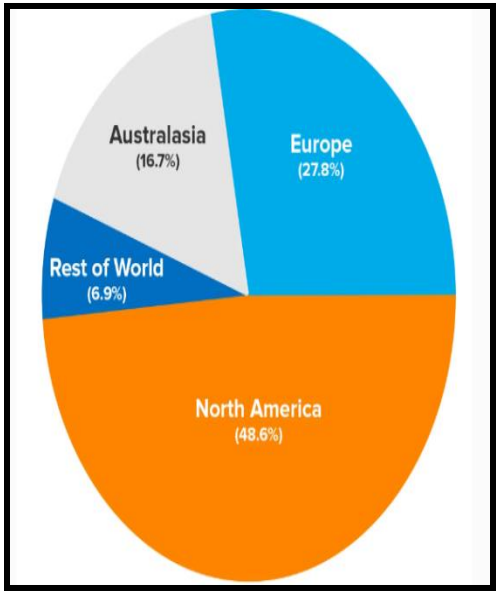
Uses:

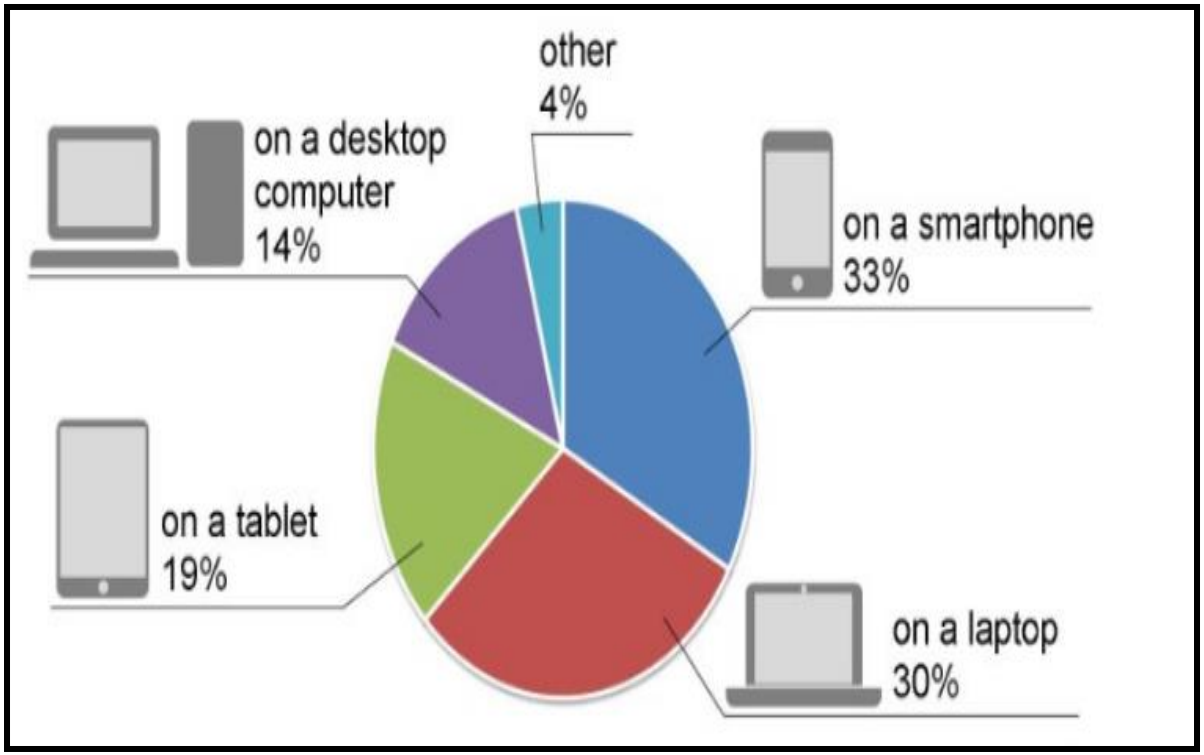
1. Pie chart is the simplest and most effective visual tool for comparing parts of a whole
2. Suitable when trying to represent composition of something
3. To show percentage or proportional data
4. Very useful for displaying nominal or ordinal categories of data

Tips for using Pie Chart:

1. Be careful not to use too many segments in your pie chart. More than six and it gets far too crowded
2. Works well when you have one set of data
3. Zero values should also be avoided
4. If you want to emphasize one of the segments, you can detach it a little from the main pie

Pie Chart


Australasia
(16.7%)
Europe
(27.8%)
Rest of World
(6.9%)
North America
(48.6%)


other
4%
on a desktop
computer
14%
on a smartphone
33%
on a tablet
19%
on a laptop
30%

Histogram

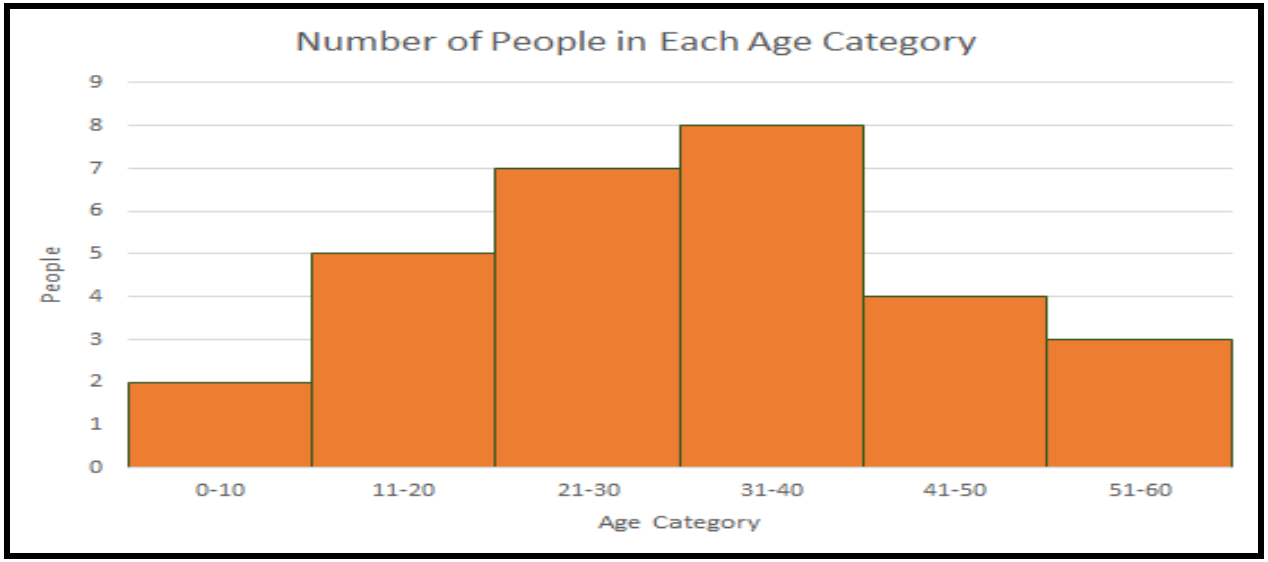
- **A histogram is an approximate representation of distribution of numeral data**
- **Shows the relationship between independent and dependent variables where independent variable is continuous rather than discreet. This means each bar represents a range of values rather than a single observation.**

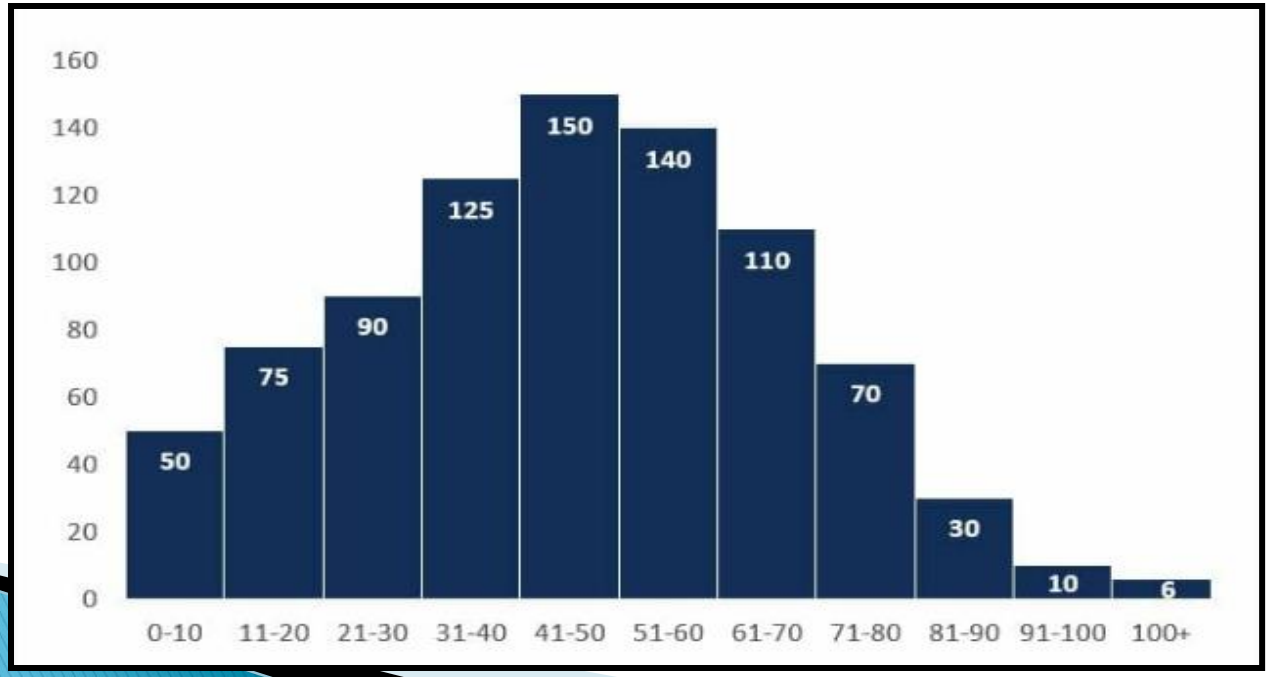
  e.g., distribution of exam scores for students in a class
  e.g., age distribution of the people living in city

Uses:

1. **When data is continuous**
2. **When you want to represent the shape of data distribution**
3. **To summarize large datasets**
4. **To communicate data distribution quickly**

Histogram


Number of People in Each Age Category
People
9
8
7
6
5
4
3
2
1
0
0-10
11-20
21-30
31-40
41-50
51-60
Age Category


160
140
120
100
80
60
40
20
0
50
75
90
125
150
140
110
70
30
10
6
0-10
11-20
21-30
31-40
41-50
51-60
61-70
71-80
81-90
91-100
100+

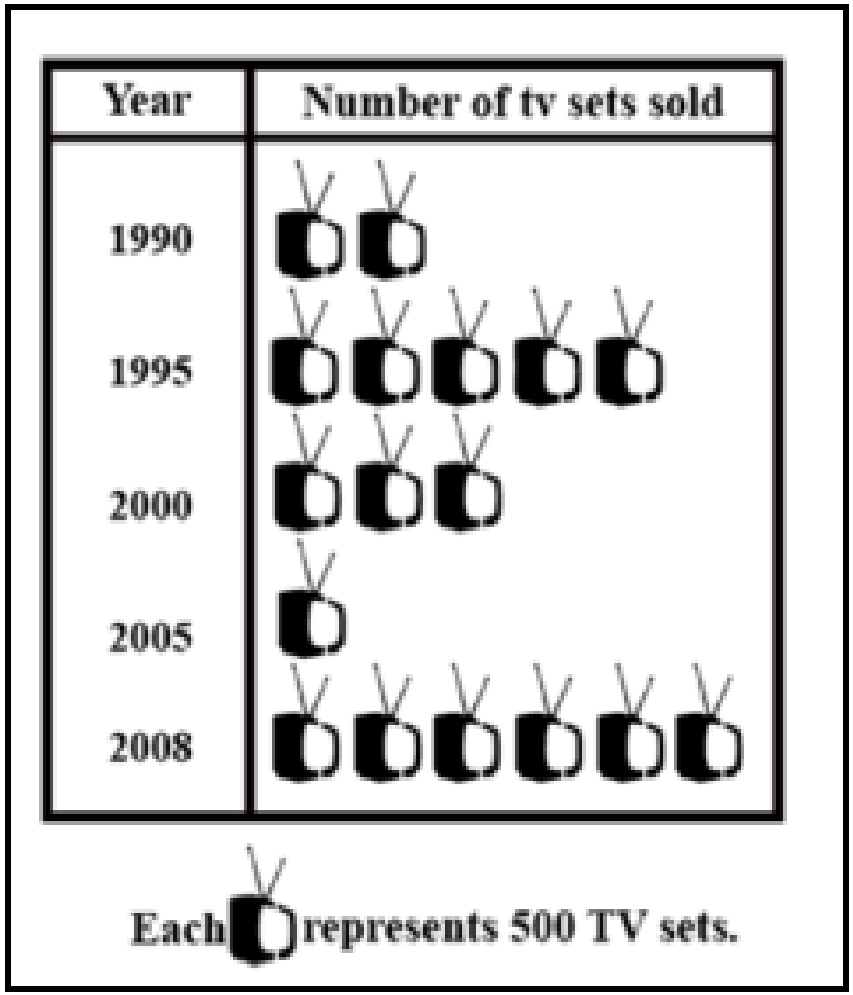
Pictograph

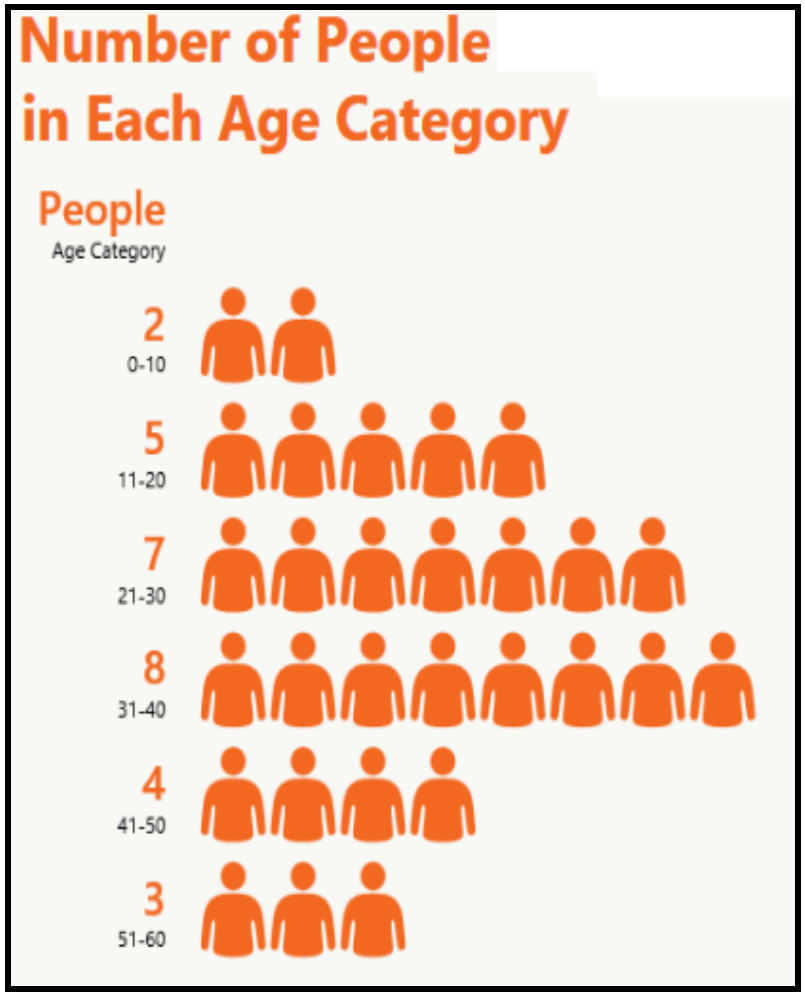
- **A pictograph is a presentation of data using images**
- **It is one of the simplest (and most popular) forms of data visualization**
- **It uses pictures or symbols to display data instead of bars. The picture represents a certain number of items**

Uses:

1. **Express large amount of information or data in a simple form**
2. **Since they make use of symbols, pictographs attract attention**
3. **Easy to read since all the information is available in one glance**
4. **Since pictograph are universally used they do not require a lot of explanation**

Pictograph


Year
Number of tv sets sold
1990
1995
2000
2005
2008
Each represents 500 TV sets.


Number of People
in Each Age Category
People
Age Category
2
0-10
5
11-20
7
21-30
8
31-40
4
41-50
3
51-60

References

- https://corporatefinanceinstitute.com/resources/excel/study/types-of-graphs/
- https://www.toppr.com/guides/business-economics-cs/descriptive-statistics/diagrammatic-presentation-of-data
- http://ecoursesonline.iasri.res.in/mod/page/view.php?id=34200
- https://venngage.com/blog/data-visualization/
- https://visme.co/blog/types-of-graphs/
- https://www.mindtools.com/pages/article/Charts_and_Diagrams.htm
- https://preply.com/en/blog/2018/08/17/charts-graphs-and-diagrams-in-the-presentation/#scroll-to-heading-7
- https://www.skillsyouneed.com/num/graphs-charts.html
- https://www.toppr.com/guides/maths/data-handling/pictographs/
- https://www.mathsisfun.com/data/pie-charts.html