SNS COLLEGE OF TECHNOLOGY

Coimbatore-37.
An Autonomous Institution

COURSE NAME : 19CSE301 INTRODUCTION TO DATA SCIENCE

III YEAR/ V SEMESTER

UNIT – II
Topic: Types of Data Representation

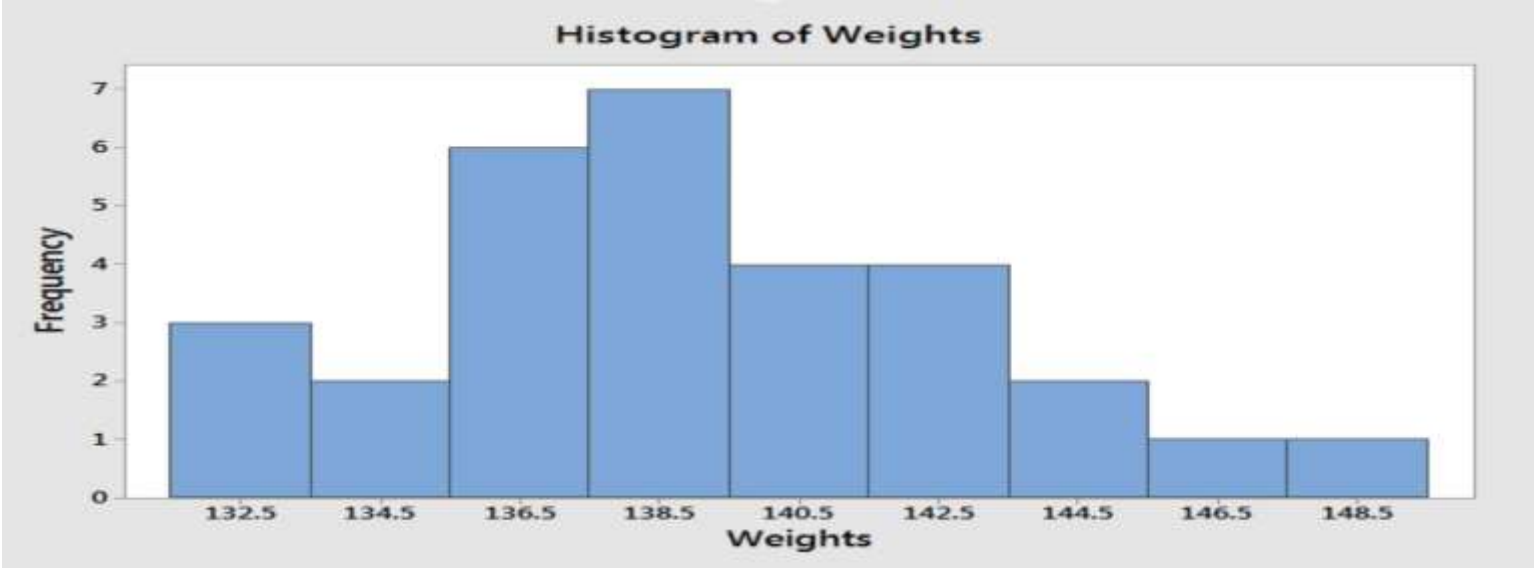
Ms.B.Sumathi
Assistant Professor
Department of Computer Science and Engineering

Variety of Data Representation

- Table
- Vertical bar chart
- Pictograph
- Circle
- Dot plots
- Histograms
- Line/stick Graph
- Frequency Polygens
- Scatter Diagrams

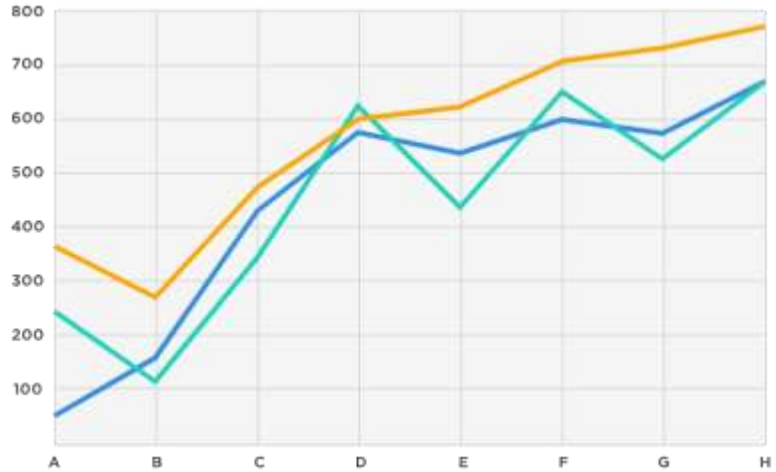
Histogram

- **Histograms** divide the horizontal axis into equal-sized intervals and use the heights of the bars to show the count or percent of data within each interval.
- By convention, each interval includes the lower boundary but not the upper one.
- Histograms show only totals for the intervals, not specific data points.



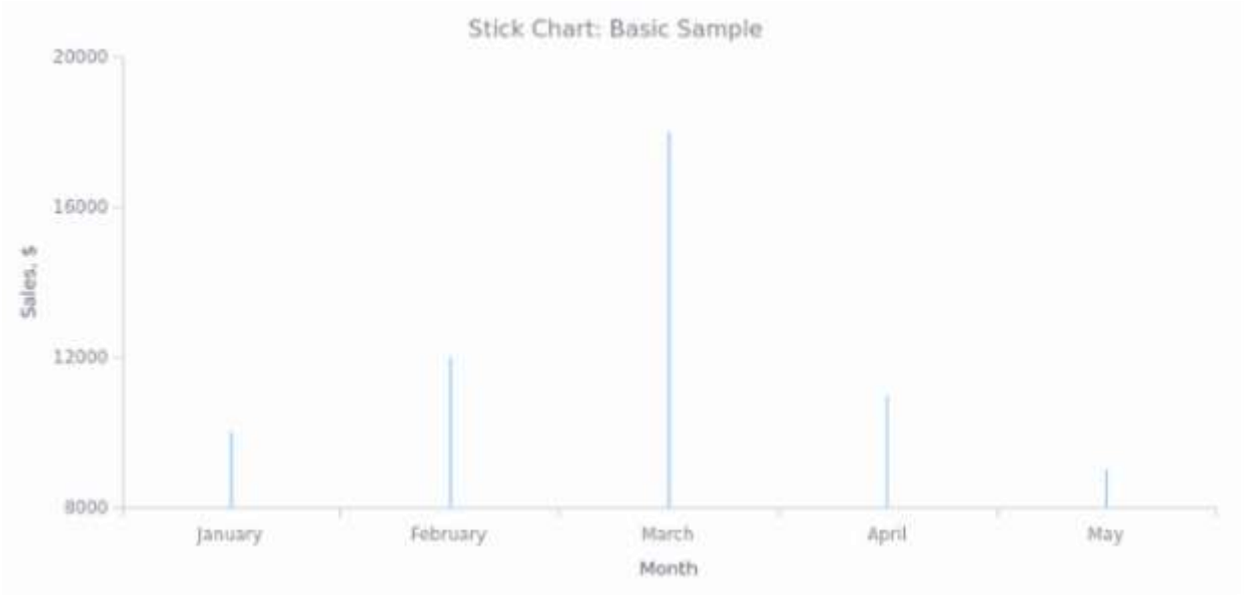
Line Graph

- A **line chart** provides the clearest graphical representation of time-dependent variables.
- It is also the preferred mode of representing trends or variables over a period of time.
- People are familiar with this simple chart, which is made up of data values plotted as points along the X and Y axes and are connected using line segments.
- Usually, time is plotted along the X-axis, and the Y-axis represents some metric of interest in the context of the period being tracked.



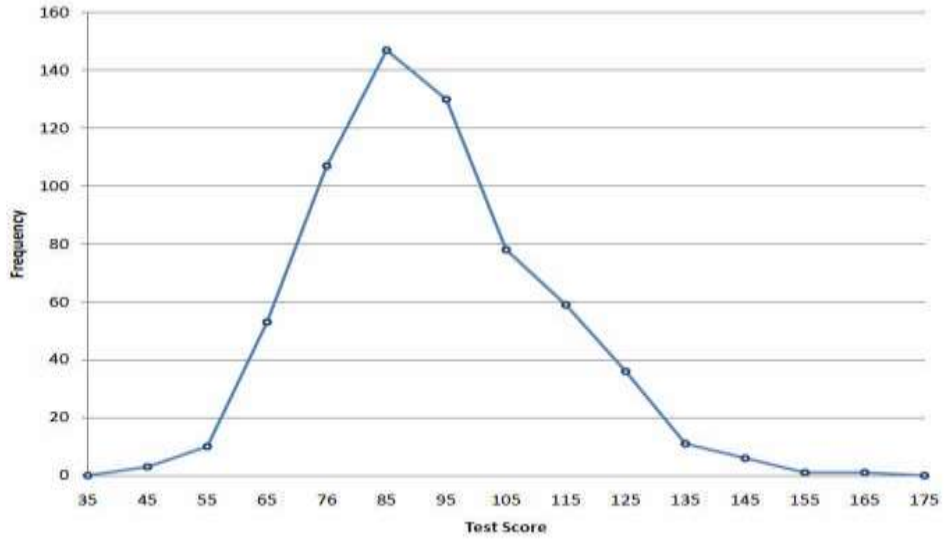
Stick Graph

- Stick Charts look like Column Charts with no width. Sticks are good at demonstrating some discrete data.
- Stick Chart is a variation of a Column, so when a Stick Chart is created it is necessary to create a Column Chart first and set the series of a Stick type.



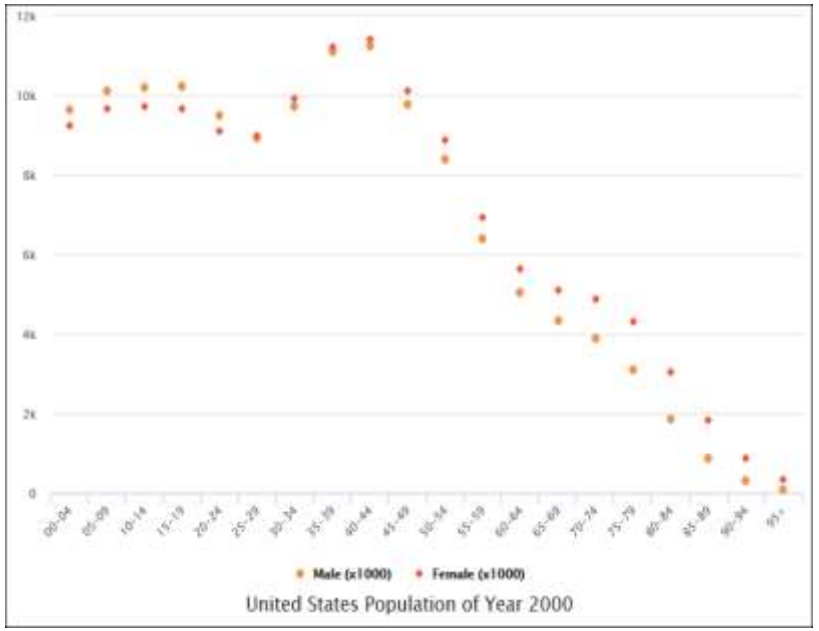
Frequency Polygens

- Frequency polygons are a graphical representation of data distribution that helps in understanding the data through a specific shape.
- Frequency polygons are very similar to histograms but are helpful and useful while comparing two or more data.
- The graph mainly showcases cumulative frequency distribution data in the form of a line graph.



Scatter Diagram

- A scatter diagram is a tool for analyzing relationships between two variables for determining how closely the two variables are related.
- One variable is plotted on the horizontal axis and the other is plotted on the vertical axis. The pattern of their intersecting points can graphically show relationship patterns.



References

- Tom M. Mitchell, “Machine Learning”, McGraw-Hill Education (India) Private Limited, 2013.
- 2Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, “An Introduction to Statistical Learning: with Applications in R”, Springer; First Edition 2013.


Thank
You