CODE BASICS

/codebasics

KEY TAKEAWAYS

CHAPTER TITLE | **Data & Visualization Basics**

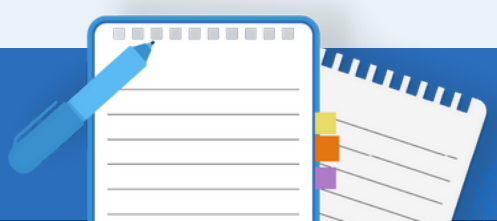
LECTURE TITLE | **Types of Data**

1. Structured data is organized, follows a clear format, and is easy to work with. (Ex: Table data, JSON, XML)

2. **Unstructured data** lacks a specific structure, is more diverse in content, and requires specialized tools and techniques for analysis. (Ex: Audio, Video, Free Form Text)

3. **Categorical data** represents categories, labels, or distinct groups. It is used to classify and categorize items or observations based on certain characteristics or attributes.

4. **Numeric data** consists of numerical values that represent measurable quantities or continuous variables.



KEY TAKEAWAYS

CHAPTER TITLE | **Data & Visualization Basics**

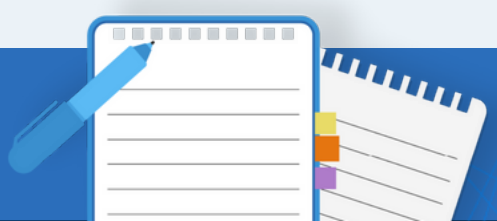
LECTURE TITLE | **Types of Data**

5 **Nominal data** consists of categories or labels with no inherent order or ranking.

6 **Ordinal data**, unlike nominal data, has a specific order or ranking among categories.

7 **Continuous data** encompasses an infinite range of precise values, often with decimals.

8 **Discrete data**, on the other hand, comprises distinct, countable numerical values.



CODE BASICS

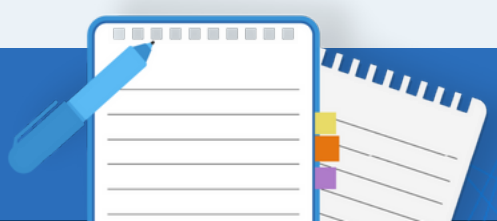
/codebasics

KEY TAKEAWAYS

CHAPTER TITLE | **Data & Visualization Basics**

LECTURE TITLE | **Pie and Bar Chart**

1. Use a **bar chart** when you have benchmark values to compare with.
2. Use a Horizontal bar chart when **category labels** are long.
3. Use a Horizontal bar chart instead of a pie chart when the number of categories is **more than 5.**
4. Use a vertical bar chart for **time series data.**
5. Consider a **Clustered Column** chart for comparing multiple data series within distinct categories.



KEY TAKEAWAYS

CHAPTER TITLE | **Data & Visualization Basics**

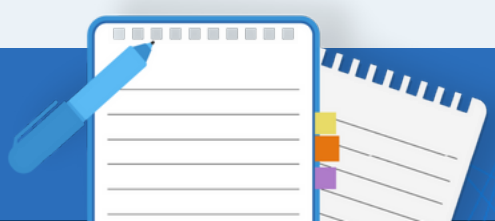
LECTURE TITLE | **Histogram and Line Chart**

1. Histograms are primarily used to show the **frequency distribution** of a continuous or discrete dataset.

2. In a histogram, all bins are of **equal size.**

3. A line chart is useful in presenting the **trend or change** in data over a period.

4. A stacked column chart used to represent and compare multiple categories in a single bar, while also displaying the overall total.



CODE BASICS

/codebasics

KEY TAKEAWAYS

CHAPTER TITLE	Data & Visualization Basics

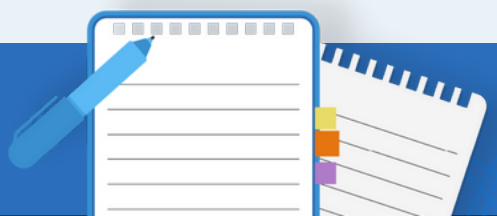
LECTURE TITLE	Scatter and Bubble Plot

1. **A scatter plot** is a graph that uses a grid to present data values, typically involving two variables within a dataset.

2. Scatter plots help to visualize the **relationship between two variables.**

3. **A Bubble chart** is a variation of a Scatter Plot that allows you to represent a third variable through the size of the bubbles.

4. Scatter plots and Bubble Charts help to detect **outliers,** visualize the **relationship** between two variables, and also Identify **trends.**



info@codebasics.io

codebasics.io

KEY TAKEAWAYS

CHAPTER TITLE | **Data & Visualization Basics**

LECTURE TITLE | **Univariate vs Bivariate vs Multivariate analysis**

1. **Univariate analysis** = Analysis of a **single** variable
2. **Bivariate analysis** = Analysis of **two** variables
3. **Multivariate analysis** = Analysis of > **two** variables