DESCRIPTIVE ANALYTICS, INFERENTIAL ANALYTICS AND DATA VISUALIZATION USING PYTHON

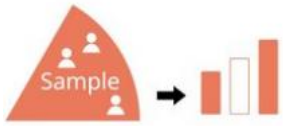
PRESENTED BY – REMYA R S

Analyzing Data for a Business

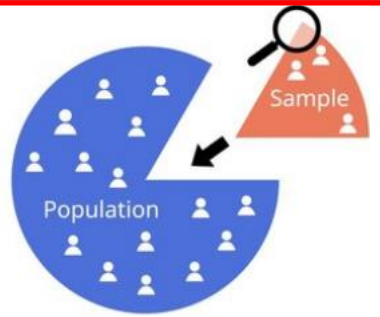
Understanding Problem Statement

Data Collection

Data Cleaning



Descriptive Analysis



Inferential Analysis

Data Visualization

Predictive Analysis

Descriptive Analysis

10,20,20,30,40,50,50,50,60

Count =9

Mean (Average):

Mean= (10+20+20+30+40+50+50+50+60) / 9= 36.67

Median, middle value of the sorted list = 40

Mode, value that appears most frequently =50

Maximum=60

Minimum=10

Descriptive Analysis

10,20,20,30,40,50,50,50,60

$$\text{Standard Deviation} = \sqrt{\frac{\sum(x_i - \mu)^2}{N}}$$

$$\mu = \frac{10 + 20 + 20 + 30 + 40 + 50 + 50 + 50 + 60}{9} = \frac{330}{9} \approx 36.67$$

N = 9

$$\sum(x_i - \mu)^2 = (10 - 36.67)^2 + (20 - 36.67)^2 + (20 - 36.67)^2 + (30 - 36.67)^2 + (4$$

$$= (-26.67)^2 + (-16.67)^2 + (-16.67)^2 + (-6.67)^2 + (3.33)^2 + (13.33)^2 + (13.33)^2 \cdot$$

$$= 711.11 + 277.78 + 277.78 + 44.45 + 11.09 + 177.69 + 177.69 + 177.69 + 544.45$$

$$= 2399.73$$

Standard deviation = 16.10

Descriptive Analysis

10,20,20,30,40,50,50,50,60

A box plot is a standardized way of displaying the distribution of data

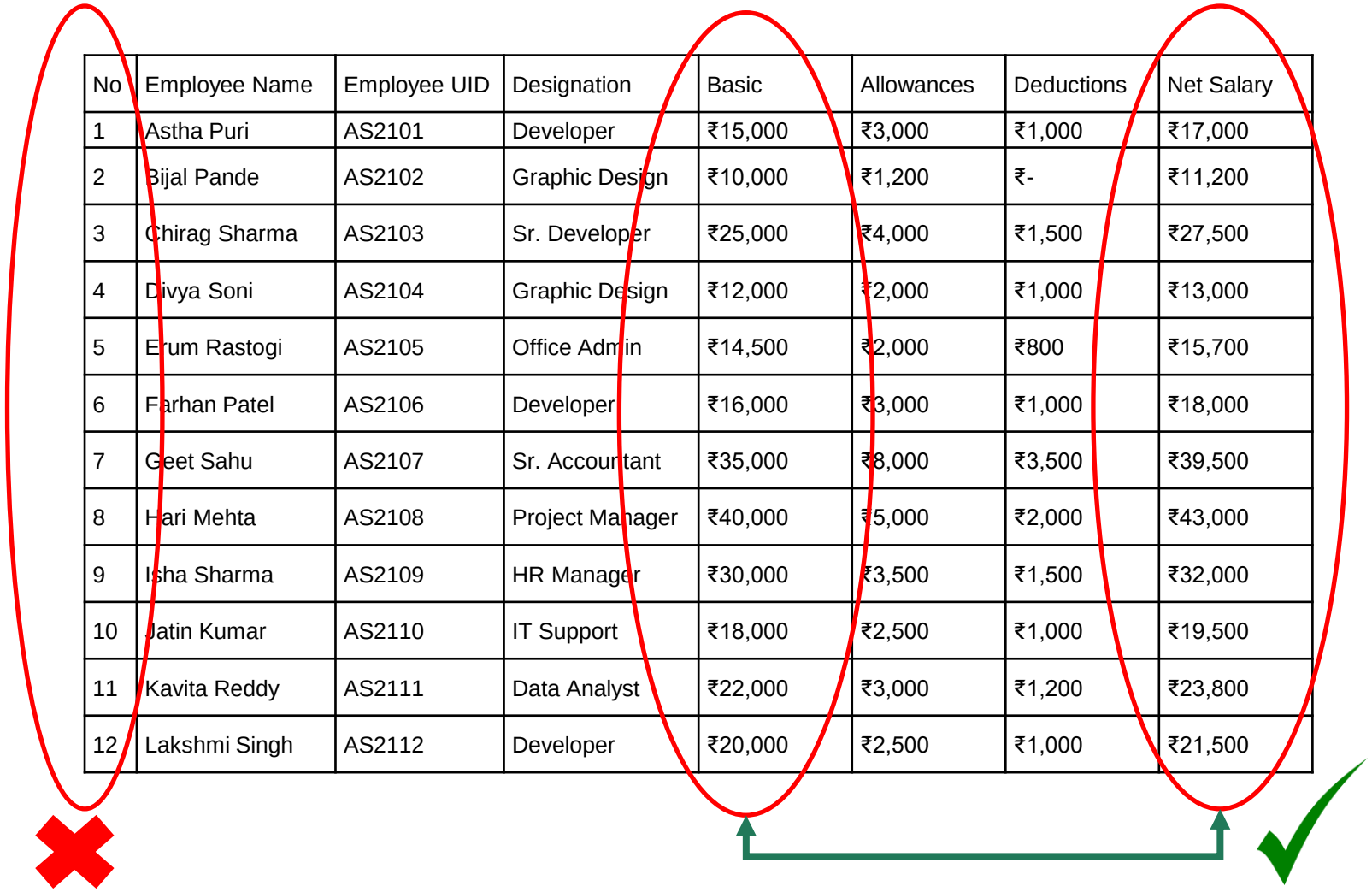
Q2 → rep. middle value and 50 % of data lies below it

Q1 → 25 % of data lies below it

Q3 → 75 % of data lies below it

Importance of Inferential Analysis

No	Employee Name	Employee UID	Designation	Basic	Allowances	Deductions	Net Salary
1	Astha Puri	AS2101	Developer	₹15,000	₹3,000	₹1,000	₹17,000
2	Bijal Pande	AS2102	Graphic Design	₹10,000	₹1,200	₹-	₹11,200
3	Chirag Sharma	AS2103	Sr. Developer	₹25,000	₹4,000	₹1,500	₹27,500
4	Divya Soni	AS2104	Graphic Design	₹12,000	₹2,000	₹1,000	₹13,000
5	Erum Rastogi	AS2105	Office Admin	₹14,500	₹2,000	₹800	₹15,700
6	Farhan Patel	AS2106	Developer	₹16,000	₹3,000	₹1,000	₹18,000
7	Geet Sahu	AS2107	Sr. Accountant	₹35,000	₹8,000	₹3,500	₹39,500
8	Hari Mehta	AS2108	Project Manager	₹40,000	₹5,000	₹2,000	₹43,000
9	Isha Sharma	AS2109	HR Manager	₹30,000	₹3,500	₹1,500	₹32,000
10	Jatin Kumar	AS2110	IT Support	₹18,000	₹2,500	₹1,000	₹19,500
11	Kavita Reddy	AS2111	Data Analyst	₹22,000	₹3,000	₹1,200	₹23,800
12	Lakshmi Singh	AS2112	Developer	₹20,000	₹2,500	₹1,000	₹21,500



Some Tests Inferential Analysis

Test	To check relation between
Pearson Correlation	Two numerical columns
Chi Square Test	Two categorical columns
T Test	A numerical column and a categorical column with two categories.
ANOVA	A numerical column and a categorical column multiple categories.

When p<0.05, there is a statistically significant association between features / columns

Visualization

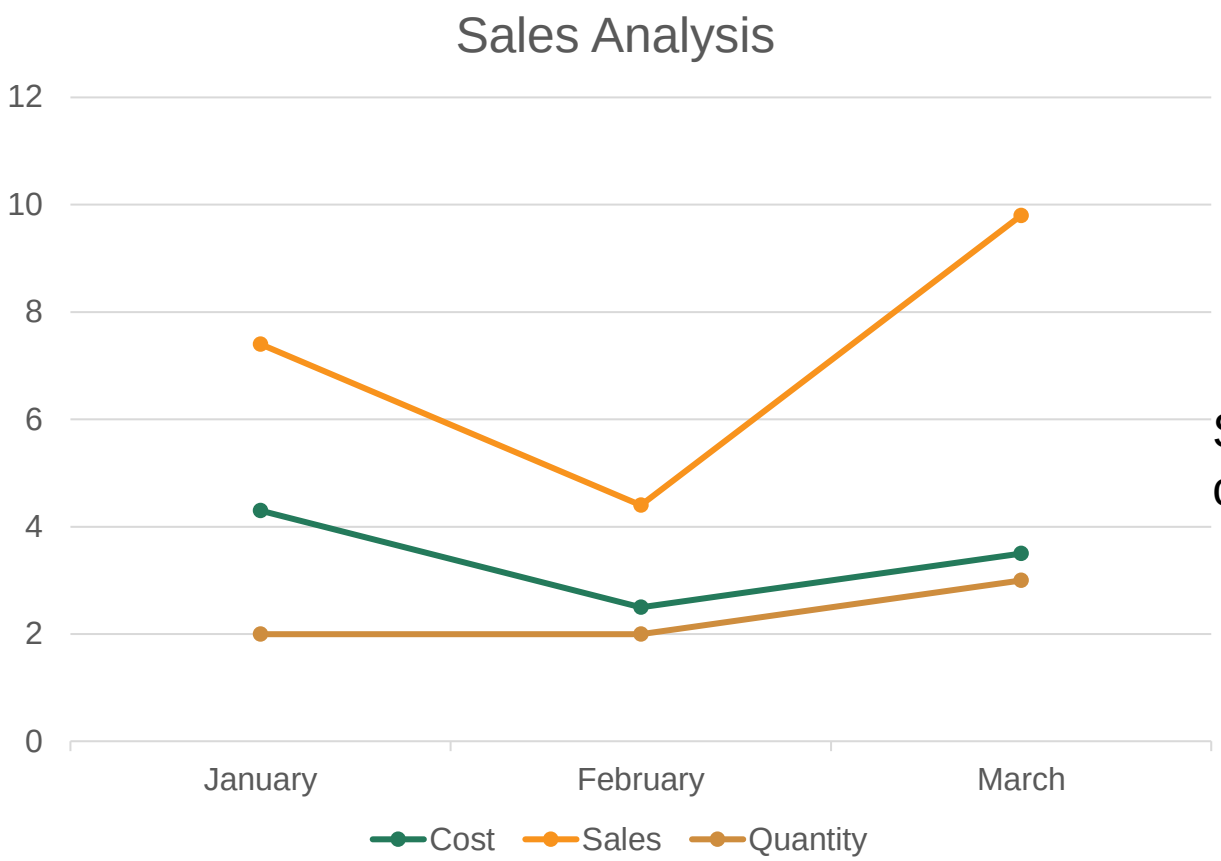
Chart : graphical representation of data or information

Its designed to visually communicate patterns, trends, and relationships that might be less clear from raw data alone.

Charts provide a visual summary of data, making it easier to grasp patterns and trends at a glance.

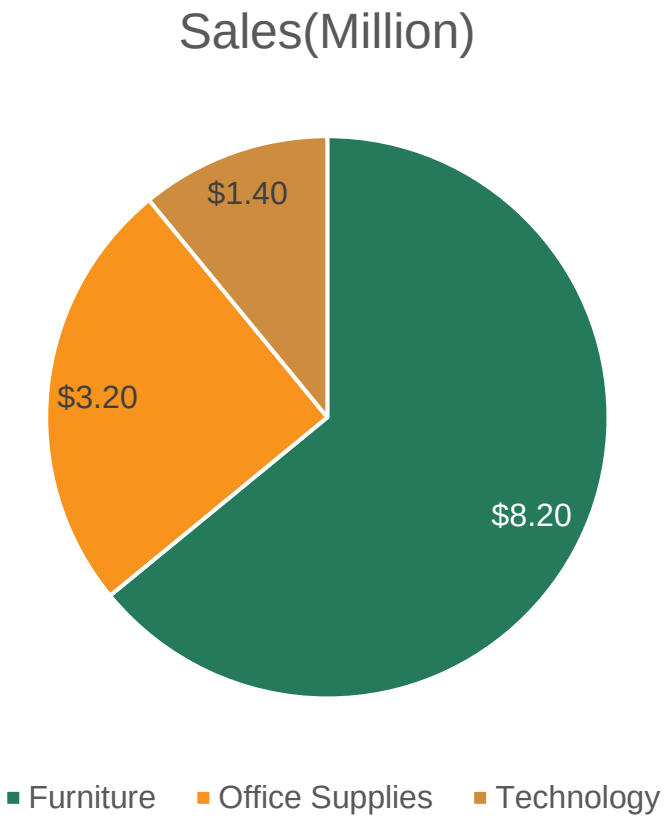
Different types of charts (such as bar charts, line graphs, pie charts, etc.) allow for different types of analysis.

Line Chart



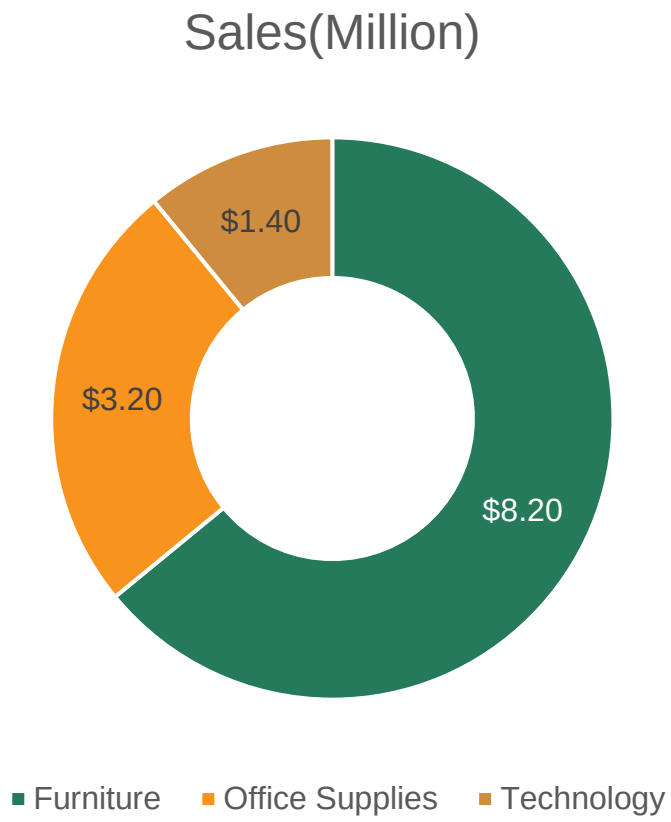
Shows trends over time or progression of data points in chronological order.

Pie Chart



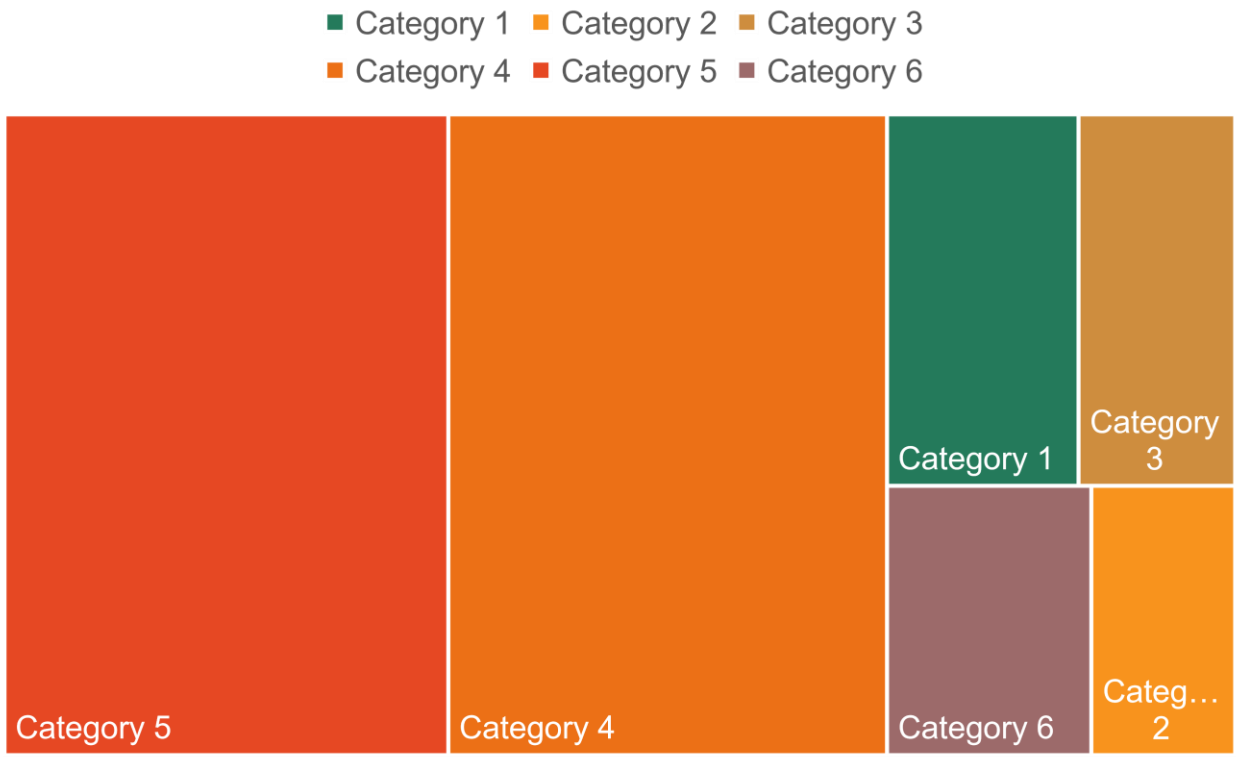
Displays proportions or percentages of a whole.

Donut Chart



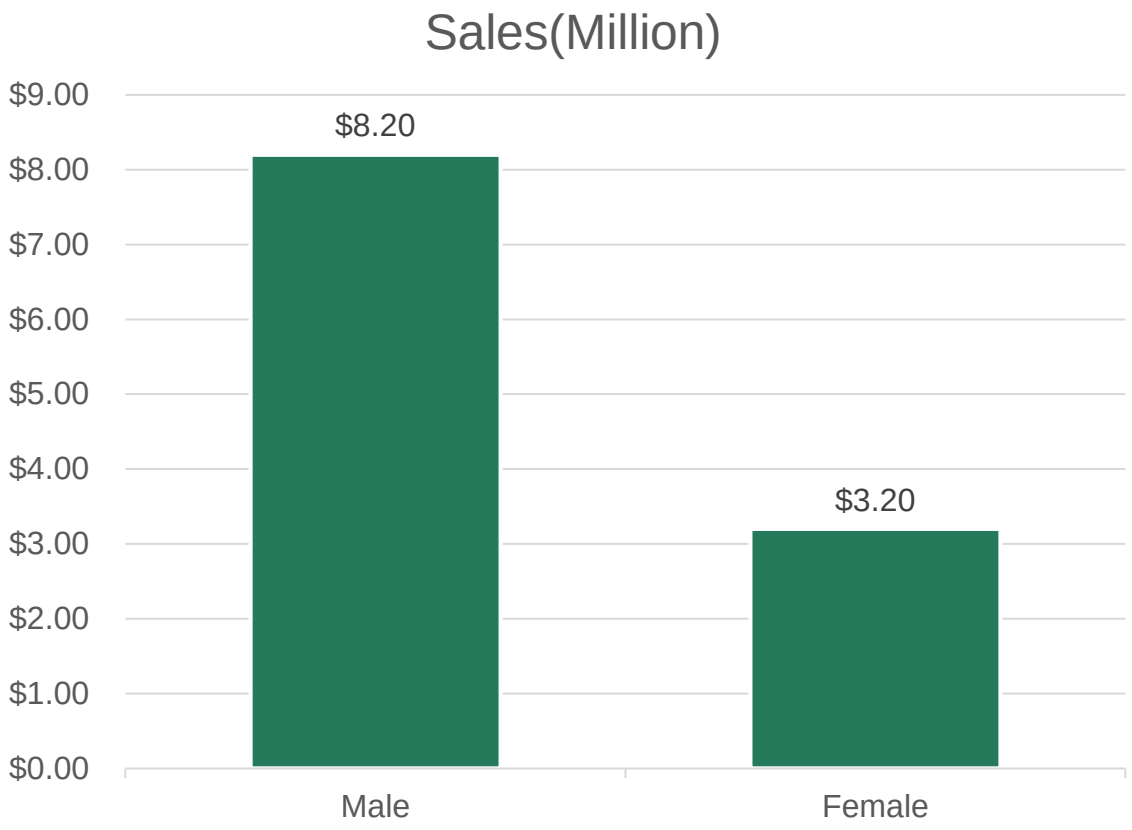
Displays proportions or percentages of a whole.

Tree map



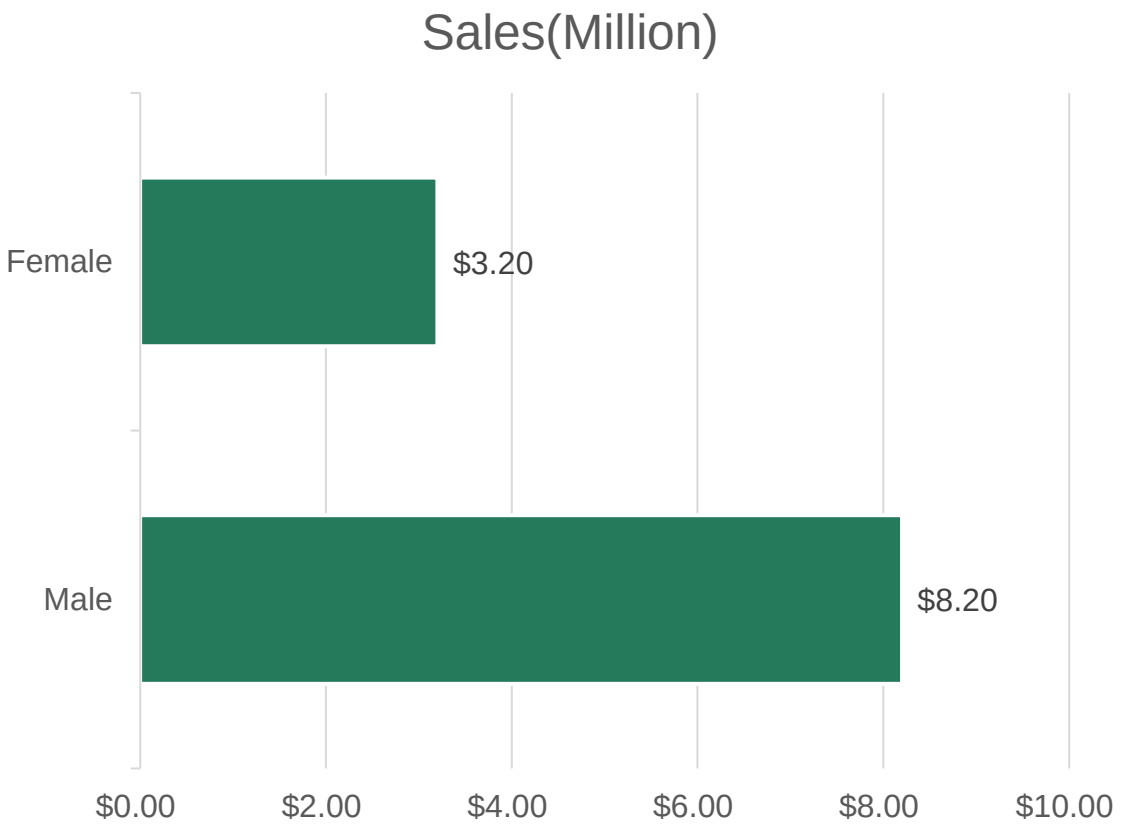
Used to visualize data with multiple levels of categories,

Column Chart



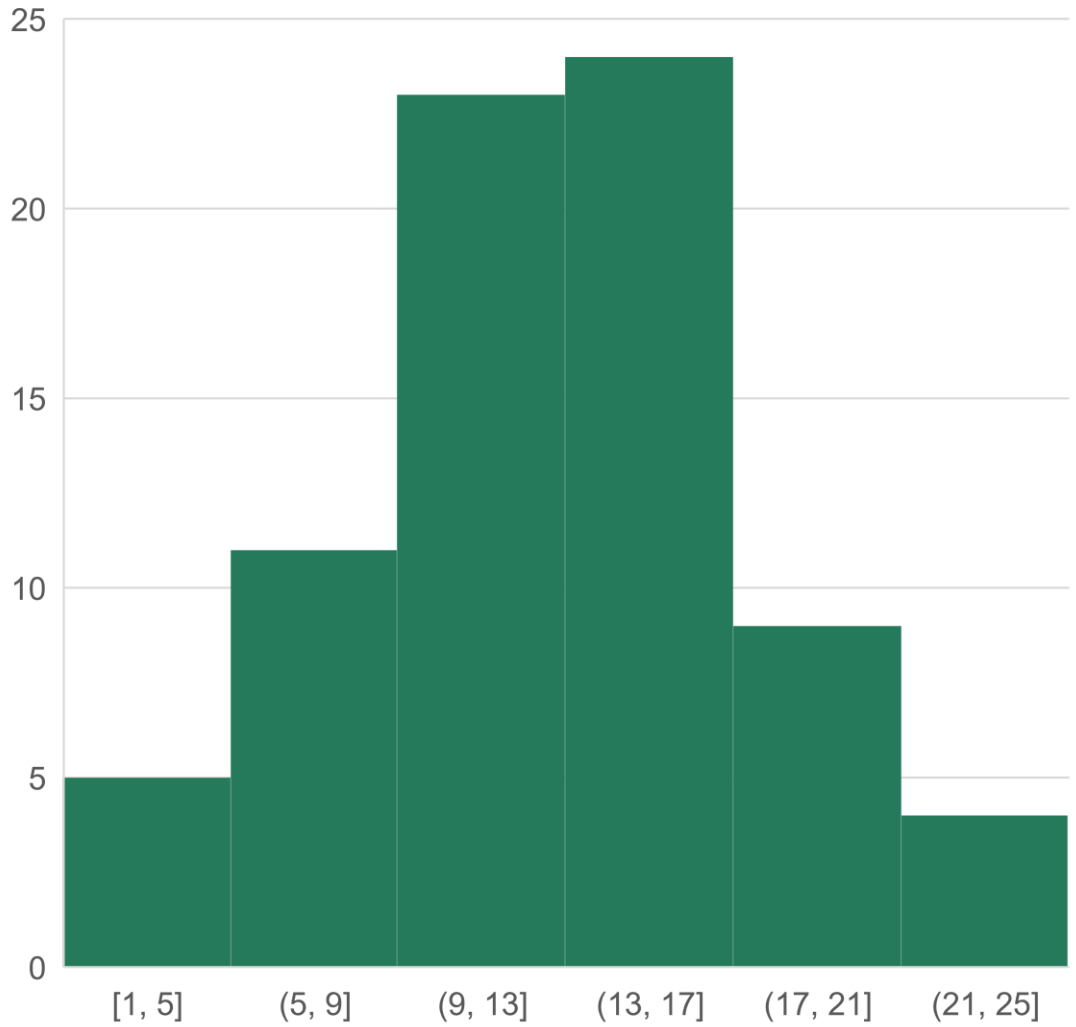
Used to compare quantities of different categories or groups.

Bar Chart



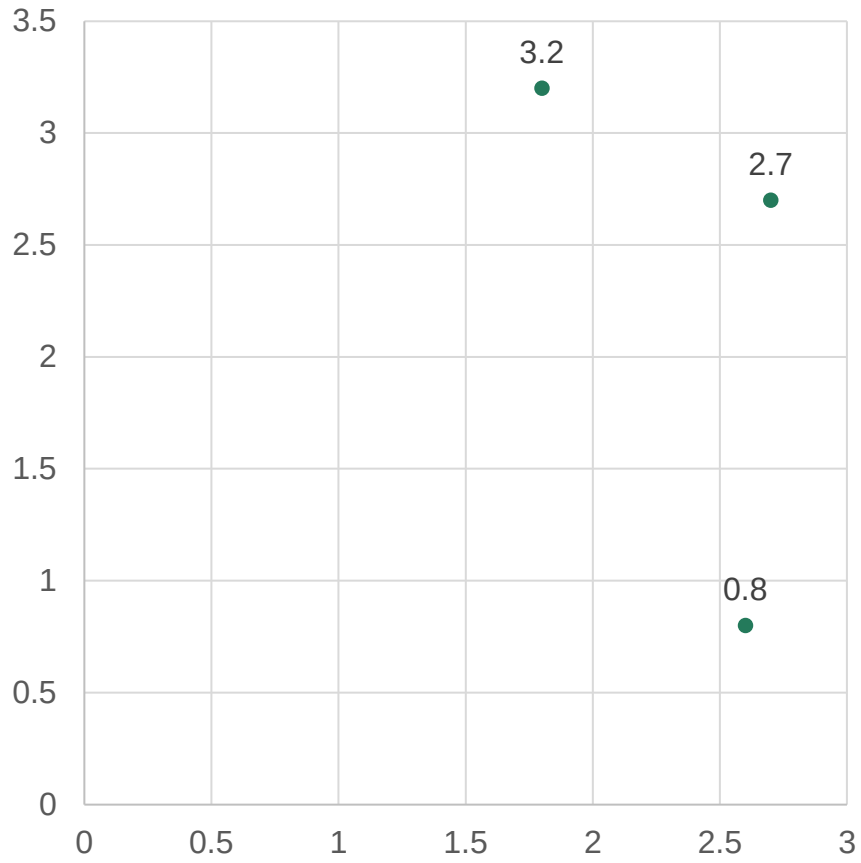
Used to compare quantities of different categories or groups.

Histogram



Illustrates the distribution of data points across a continuous interval or range.

Line Chart



Shows relationships or correlations between two variables.

Thank
You