Data Visualization

chart   diagram   tables   map   graph

→ drawn on graph paper

→ proper scale is used

→ drawn without using graph paper
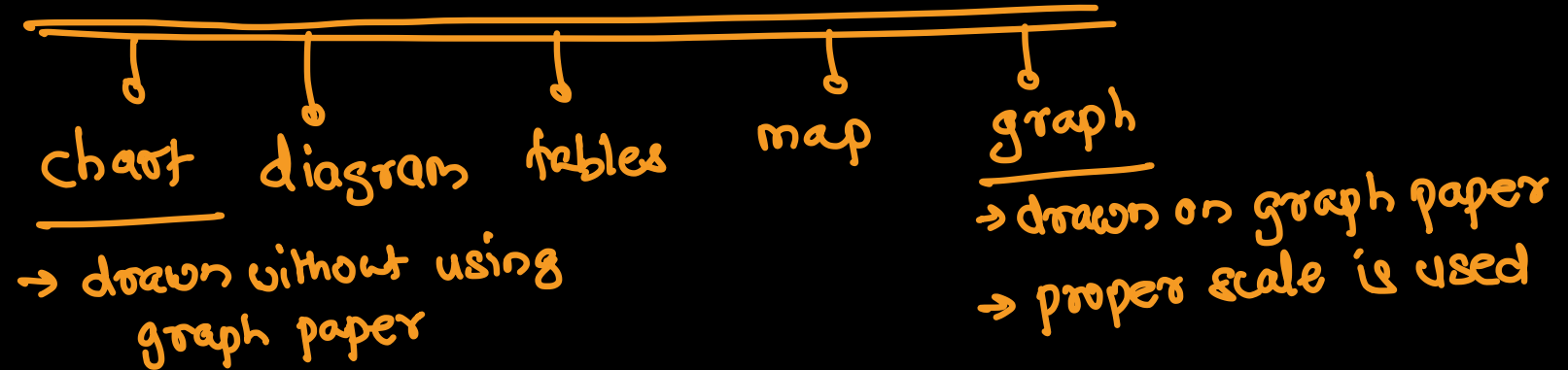
What is Data Visualization?

- Data visualization is the graphical representation of information and data using visual elements like charts, graphs, and maps.

- **Key Purpose**
  - Make complex data accessible and understandable
  - Reveal patterns, trends, and insights
  - Enable faster decision-making
  - Communicate findings effectively

Why Data Visualization Matters

- ## The Human Brain and Visuals
  - Processes visual information 60,000x faster than text
  - 90% of information transmitted to brain is visual
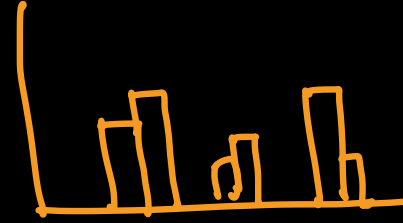 - Visual memory lasts longer than textual memory

- ## Business Impact
  - Reduces time to insight by 67%
  - Increases comprehension by 400%
  - Improves decision-making speed by 5x
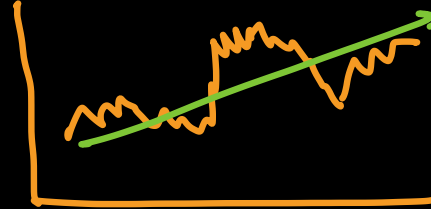
Types of Data Visualization

- **Comparison Chart**
  - Bar charts, column charts
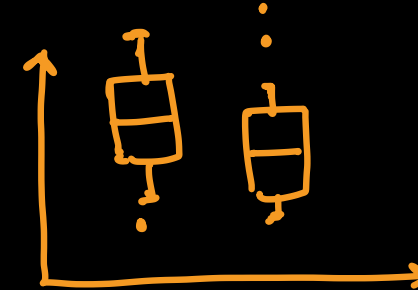 - Best for: Comparing quantities across categories

- **Trend Analysis**
  - Line charts, area charts
  - Best for: Showing changes over time

- **Distribution**
  - Histograms, box plots, scatter plots
  - Best for: Understanding data spread and relationships

- **Composition**
  - Pie charts, stacked bars, treemaps
  - Best for: Showing parts of a whole

Choosing the Right Chart Type

Data Type	Best Visualization	Use Case
Categorical	Bar Chart	Compare Categories
Time Series	Line Chart	Show trends over time
Correlation	Scatter Plot	Relationship between variables
Distribution	Histogram	Frequency of Values
Hierarchical	Treemap	Nested Categories
Geographic	Map	Location based data

Design Principles

- **Clarity First**
  - Remove unnecessary elements (chartjunk)
  - Use clear, descriptive titles
  - Label axes and data points

- **Color with Purpose**
  - Use color to highlight key insights
  - Maintain accessibility (colorblind-friendly)
  - Limit color palette (3-5 colors maximum)

- **Consistency**
  - Standardize fonts, colors, and styles
  - Maintain consistent scales across similar charts

Color Psychology in Data Visualization

- **Color Meanings**
  - **Red**: Urgency, danger, negative values
  - **Green**: Growth, positive values, go-ahead
  - **Blue**: Trust, stability, neutral information
  - **Orange**: Warning, attention, moderate priority
  - **Gray**: Neutral, secondary information
- **Best Practices**
  - Use diverging colors for positive/negative data
  - Apply sequential colors for ordered data
  - Employ categorical colors for distinct groups

Common Visualization Mistakes

- **Wrong Chart Type**
  - Pie charts for too many categories (>5)
  - 3D effects that distort data
  - Dual y-axes without clear purpose

- **Misleading Scales**
  - Truncated y-axis without indication
  - Inconsistent intervals
  - Logarithmic scales without labeling

- **Information Overload**
  - Too many data series
  - Cluttered legends
  - Excessive annotations

Tools and Technologies

- **Enterprise Solutions**
    - Tableau - Advanced analytics and visualization
    - Power BI - Microsoft ecosystem integration
    - Qlik Sense - Associative data modeling
- **Programming Libraries**
    - Python: Matplotlib, Seaborn, Plotly
    - R: ggplot2, Shiny, Leaflet
    - JavaScript: D3.js, Chart.js, Highcharts
- **Simple Tools**
    - Excel/Google Sheets - Basic charts
    - Canva - Design-focused visuals