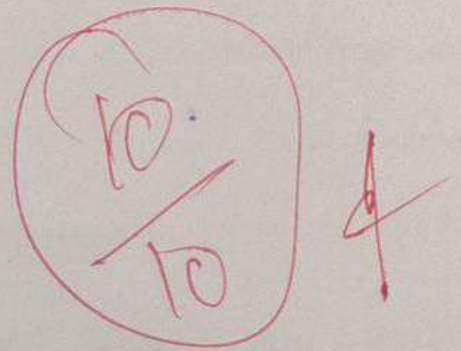


DATA VISUALIZATION ASSIGNMENT

NAME : J. Akshaya

VTU : 21751

SLOT NO : S4L6



- 1) Explain how human perceptual processing models and Gestalt principles influence the effectiveness of data visualization. Discuss with suitable examples how visualization designers can minimize information overload and maximize information clarity using concepts such as Gibson's Affordance theory, data abstraction, and appropriate dataset representation.

Human perceptual processing Models in Data Visualization:-

Human perception plays a critical role in how we interpret visual information. Data visualizations are effective only if they align with how our brain naturally processes patterns, contrasts and structures.

* Pre-attentive processing:- The human brain can instantly recognize visual features like color, size, and shape without conscious effort.

Example:- A red dot in a sea of blue dots immediately stands out.

* Gestalt principles:- These describe how humans group and interpret visual elements.

→ Proximity:- Elements close to each other are perceived as related.

Example:- Grouping bar chart categories together helps users see relationships.

→ Similarity:- Similar shapes/colors are seen as part of a group.

Example:- Using consistent color for one category across charts improves clarity.

→ Continuity:- The eye prefers continuous patterns.

Example:- Line charts are effective in showing trends because the eye follows the path smoothly.

→ Closure:- Incomplete shapes are seen as complete.

Example:- Circular progress charts are effective because the brain "closes the gap".

Minimizing Information overload & Maximizing clarity

To avoid overwhelming the user, designers apply several cognitive and visualization principles:-

a) Gibson's Affordance Theory:-

* Affordances are cues in design that suggest how something can be used.

* In visualization:-

→ A scrollable timeline affords exploring sequential data.

→ A slider bar affords filtering a range of values.

→ Interactive buttons afford drilling down into details.

3

Example:- In a COVID-19 dashboard, a slider to select the data range affords temporal exploration, reducing clutter by not showing the entire dataset at once.

b) Data Abstraction:-

* The transforming raw data into a simpler form for easier understanding

* Levels of abstraction:-

→ Data → Information → Knowledge → Insight.

* Instead of raw numbers, abstract into visual summaries

→ Example:- Instead of displaying millions of sales records, show a heatmap or boxplot to abstract patterns.

c) Appropriate Dataset Representation.

* choosing the right visualization for the data type is critical:-

* Categorical data → Bar charts, Treemaps.

* Continuous trends → Line graphs.

* Hierarchical data → Sunburst or Tree diagrams

* Geospatial data → Maps.

Example:- Using a pie chart to show time series is ineffective. A line chart communicates it better by aligning with perceptual continuity

4

3) With the help of suitable datasets, compare and contrast different visualization techniques used in Univariate, Bivariate, and Multivariate analysis. Explain how the choice of visualization (e.g., bar chart, scatterplot fit line, violin plot, heatmap etc) depends on the type of data and the number of variables being analyzed. Provide at least one practical example for each analysis type.

D Univariate Analysis:-

Understand the distribution, central tendency and spread of a single variable.

suitable visualizations:-

categorical data \rightarrow Bar chart, Pie chart

continuous data \rightarrow Histogram, Boxplot, violin plot, Density plot.

* Dataset Example:- Student Exam scores

* variable: Math score (0-100)

* Visualization:-

* Histogram \rightarrow shows frequency distribution

* Boxplot \rightarrow Highlights median, quantiles, and outliers

* Violin plot \rightarrow shows both distribution shape & density, more informative than a boxplot.

Example:- In an e-commerce dataset, plotting a histogram of "order delivery Time" helps find whether delays are common.

2) Bivariate Analysis:-

Explore relationships, correlation, or comparisons between two variables.

* Suitable visualizations:-

* Categorical vs Categorical \rightarrow Grouped Bar chart, Mosaic plot.

* Categorical vs Continuous \rightarrow Boxplot, Violin plot, Bar chart.

* Continuous vs Continuous \rightarrow Scatter plot, Scatterplot with fit line, Hexbin plot.

* Dataset Example:- sales Dataset

* Variables: Advertising spend vs Sales Revenue

* Visualization:-

* Scatterplot \rightarrow shows linear positive correlation

* Scatterplot with regression line \rightarrow Adds clarity to the trend.

* Variables:- product category vs sales Revenue

* Boxplot \rightarrow shows how sales vary by category.

Example:- In healthcare, comparing BMI vs Blood pressure using a scatterplot helps see correlation; while Gender vs BMI can be explored using violin plots.

Multivariate Analysis:-

Examine interactions and combined effects among multiple variables.

* Suitable Visualizations:-

- * Categorical + continuous \rightarrow Grouped / stacked Bar charts

- * Continuous \rightarrow Scatterplot Matrix, Heatmap, Parallel coordinates plot.

- * Mixed data \rightarrow Bubble charts, Treemaps.

* Dataset Example:- Smart Farming IoT dataset

- * Variables: Soil Moisture, Temperature, Crop type.

* Visualization:-

- * Heatmap \rightarrow correlation matrix between soil Moisture, temperature, rainfall.

- * Scatterplot Matrix \rightarrow Explore relationships across multiple sensor readings.

- * Bubble chart \rightarrow Soil moisture (x), Crop yield (y), Bubble size = Temperature.

Example:- In marketing analyzing customer Age spending score and gender can be visualized using a scatterplot with color for gender and bubble size for spending score.