PBR VISVODAYA INSTITUTE OF TECHNOLOGY & SCIENCE

SUBJECT: DATA VISUALIZATION

QUESTION BANK & BIT BANK

UNIT – I (10 Marks Questions)

1. What is Data visualization? How it is useful to present the data? Give an example.
2. Explain the difference between bad graph & Good Graph Give any three examples.
3. What is the meaning of storytelling in Data visualization? Briefly explain the steps involved in storytelling.
4. Which one is better among exploratory, explanatory analysis? Justify
5. What is the importance of context in story selling? Explain the words Who, What and How in context.
6. What are the different visual types in Data Visualization? Explain each type with example.
7. What is Bar graph which one is better among non-zero baseline, zero base line graphs? Justify.

UNIT – II (10 Marks Questions)

1. What is clutter in data visualization? Give any 2 examples for clutter.
2. Why do we need to eliminate clutter? With an example explain the steps to remove the clutter from a presentation.
3. What are the Gestalt principles of visual perception? Explain each principle with an example.
4. Are you seeing with your eyes or brain? Write a brief story on memory?
5. What are the different preattentive attributes in text? Explain with an example.
6. What are the different preattentive attributes in the graph? Explain with an example.
7. How to construct a story in story telling? Explain.

UNIT – III

1. Describe steps in the data communication process and explain its importance in transmitting information effectively.
2. Identify and discuss three common types of communication problems that organizations may face. Provide examples for each type and suggest strategies for overcoming them.
3. Explain six principles that are essential for effective data communication. Discuss how each principle contributes to better data communication practices.
4. Explain the process of connecting to data sources in Tableau. Provide step-by-step instructions for connecting to a CSV file and a SQL database.
5. Differentiate between communicating "how much" and "how many" in a business context. Give examples of situations where each type of communication is more appropriate and explain why.
6. Define what a ratio is and how it is commonly used in business and finance. Provide an example of a financial ratio and explain its significance in decision-making.

PBR VISVODAYA INSTITUTE OF TECHNOLOGY & SCIENCE

SUBJECT: DATA VISUALIZATION

QUESTION BANK & BIT BANK

UNIT – IV

1. Explain the concept of proportions and provide an example of how they are used to compare a part to a whole. Describe the significance of this in various fields, such as finance or business analytics.
2. Compare and contrast the use of percentages to represent current-to-historical data and actual-to-target data. Give examples of situations where these percentage comparisons are valuable in decision-making.
3. Define the mean and median in statistics. Discuss their respective advantages and limitations as measures of central tendency. Provide an example where the mean and median may lead to different insights.
4. Explain the concept of the "outlier" in data analysis and how it can affect the mean and median differently. Discuss strategies for handling outliers when calculating and interpreting these measures.
5. Describe the importance of respecting variation in quality control and process improvement. Discuss how control charts are used to monitor variation over time and maintain consistency in a manufacturing or service process.
6. Explain the concept of uncertainty in decision-making and data analysis. Discuss the difference between aleatory and epistemic uncertainty and provide examples of how each type of uncertainty can impact outcomes.

UNIT – V

1. Describe the purpose and utility of scatterplots in data visualization. Provide an example of a situation where a scatterplot is an effective way to represent multiple quantities and relationships.
2. Discuss the concept of regression and trend lines in data visualization. How can regression analysis help in understanding relationships between multiple quantities? Provide an example of a scenario where a trend line is valuable.
3. Describe what a quadrant chart is and how it is used to visualize multiple quantities. Provide a real-world example where a quadrant chart would be appropriate and explain why.
4. Describe the connected scatterplot and its applications in representing changes over time. Provide a clear example of a situation where a connected scatterplot is a better choice than other time-based charts.
5. Discuss the concept of a slope graph and how it can be used to visualize changes or trends over time. Provide an example of a scenario where a slope graph is a useful tool in data communication.
6. Describe one type of special map used in data visualization and explain its unique characteristics and applications. Provide an example where this special map type is more effective than traditional maps.

PBR VISVODAYA INSTITUTE OF TECHNOLOGY & SCIENCE

SUBJECT: DATA VISUALIZATION

QUESTION BANK & BIT BANK

UNIT – I (2 Marks Questions)

1. What is the role of an introduction in data visualization, and how can it engage the audience?
2. Explain the significance of providing context in data visualization. How does it help the audience understand the data better?
3. Why is it important to carefully select the appropriate visual representation for your data?
4. In data visualization, what are the primary considerations when choosing a visual format?
5. How does the nature of the data, such as categorical or numerical, influence the choice of a visual representation?
6. Describe a scenario where a bar chart is a more effective choice than a pie chart for visualizing data.
7. What are some key factors to consider when deciding whether to use a line chart or a scatterplot to present data?
8. How does the choice of color and style in a visual affect its effectiveness in conveying information?
9. Explain why it's important to match the visual complexity to the complexity of the data when choosing a visualization type.
10. How can the audience's familiarity with specific types of visuals influence the choice of representation in data visualization?

PBR VISVODAYA INSTITUTE OF TECHNOLOGY & SCIENCE

SUBJECT: DATA VISUALIZATION

QUESTION BANK & BIT BANK

UNIT – II (2 Marks Questions)

1. Can you explain why clutter is often considered a significant challenge in data visualization, and how it can impact the effectiveness of data communication?
2. Share an example of a data visualization that you believe suffers from clutter. Describe the elements causing clutter and suggest how it could be improved.
3. What techniques or strategies can you employ to effectively draw your audience's attention to specific data points or areas of interest in a visualization?
4. In your opinion, what role do visual cues play in guiding the audience's attention, and can you provide examples of effective use of visual cues in data visualization?
5. How does incorporating storytelling enhance the effectiveness of data communication, and can you provide an example of a data presentation where storytelling was used effectively?
6. Explain what you understand by the term "data narrative." Why is it important in data visualization, and what are the key components of a compelling data narrative?
7. Share a situation in which data storytelling could be particularly valuable. Describe the type of data, the target audience, and the key elements of the narrative.
8. Discuss the potential challenges of incorporating storytelling in data presentations. How can these challenges be overcome for more effective communication?
9. Clutter is often seen as a barrier to effective data communication. Can you share specific guidelines or best practices to minimize clutter in data visualizations, and how do these practices contribute to clarity?
10. When presenting data to an audience, what factors should you consider to determine which aspects of your data should be emphasized, and how can you ensure that your audience's attention is directed to the most relevant information?

PBR VISVODAYA INSTITUTE OF TECHNOLOGY & SCIENCE

SUBJECT: DATA VISUALIZATION

QUESTION BANK & BIT BANK

UNIT – III (2 Marks Questions)

1. Can you describe one critical step in the process of communicating data and explain how it impacts the overall effectiveness of data communication?
2. What are the three types of common communication problems that can arise in data communication, and can you provide examples of each?
3. Explain one of the six principles of communicating data and how it can be applied in real-world data communication scenarios.
4. Briefly explain the main Tableau products and their specific use cases. How do these products contribute to more effective data visualization?
5. What are the key steps and considerations when connecting to external data sources in Tableau for data analysis and visualization?
6. Share your thoughts on the future of data visualization tools like Tableau and their role in the evolving data-driven landscape.
7. When it comes to communicating "how much" and "how many," can you provide an example of when each type of communication is more appropriate and explain your reasoning?
8. How do the methods of communicating "how much" and "how many" differ when dealing with numerical data versus categorical data?
9. Define the concept of a ratio in data visualization and provide an example of how ratios are used to convey meaningful information.
10. Explain the significance of using rates in data communication, particularly in scenarios involving comparisons or benchmarks.

PBR VISVODAYA INSTITUTE OF TECHNOLOGY & SCIENCE

SUBJECT: DATA VISUALIZATION

QUESTION BANK & BIT BANK

UNIT – IV (2 Marks Questions)

1. Explain the concept of representing "part to whole" using proportions and percentages in data visualization. Provide an example of a real-world application where this is particularly valuable.
2. Describe the significance of comparing "current to historical" data in data visualization, and how does it help in understanding trends or changes over time?
3. How do you effectively communicate "actual to target" percentages to stakeholders or decision-makers in a way that highlights both achievements and areas for improvement?
4. Share an example of a data visualization where proportions and percentages were used to provide insights that might not have been apparent from absolute values alone.
5. Define the mean and median as measures of central tendency in statistics. Explain how they differ and provide scenarios where each is more suitable.
6. Discuss how the choice of mean or median can be influenced by the distribution of data. When might you prefer one measure over the other based on the data distribution?
7. Explain why it is important to present both the mean and median in certain data visualizations and provide an example of when this dual representation is beneficial.
8. Describe the concept of "respecting variation" in quality control and process improvement. How does this concept impact the decision-making process?
9. Explain the difference between aleatory and epistemic uncertainty. Can you share an example of each type of uncertainty and its implications in decision-making?
10. What are some common strategies to address and minimize uncertainty in data analysis and visualization, and how can transparency in communicating uncertainty be improved?

PBR VISVODAYA INSTITUTE OF TECHNOLOGY & SCIENCE

SUBJECT: DATA VISUALIZATION

QUESTION BANK & BIT BANK

UNIT – V (2 Marks Questions)

1. Describe the primary purpose of using scatterplots in data visualization and provide an example of when scatterplots are the most effective choice for representing multiple quantities.
2. Explain how stacked bar charts are used to convey information about multiple quantities. Can you provide an example where a stacked bar chart is more suitable than other chart types?
3. How do regression and trend lines help in understanding the relationship between multiple quantities? Share a real-world application where these lines played a crucial role.
4. Describe the significance of time-based charts and their role in understanding changes over time. What are the key differences between time-based charts and other chart types?
5. How does the choice between a line chart and a connected scatterplot depend on the type of data and the story you want to convey? Can you provide an example that highlights this?
6. In data visualization, what is the importance of considering date fields and seasonality in time-based data? Share a practical case where seasonality played a critical role in data analysis.
7. Can you provide an example of a situation where a slope graph was used to effectively visualize changes over time, and what made it particularly useful in that context?
8. Describe the purpose of using a circle map in data visualization. How does it enhance the representation of location-based data?
9. Explain the key characteristics of filled maps and how they differ from other map types in data visualization. Share an example where a filled map effectively conveyed information.
10. What are the primary advantages of using dual-encoded maps in data visualization? Can you provide a specific use case where dual-encoded maps were employed successfully?

PBR VISVODAYA INSTITUTE OF TECHNOLOGY & SCIENCE

SUBJECT: DATA VISUALIZATION

QUESTION BANK & BIT BANK

UNIT – I (BIT BANK)

1. What is the primary purpose of data visualization? [ ]
   A. To hide data B. To make data more difficult to understand
   C. To present data in a visual format that is easy to understand.
   D. To replace data analysis.
2. Which of the following is NOT a common type of data visualization? [ ]
   A. Bar chart B. Scatter plot C. Spreadsheet D. Line chart
3. What is the context of data visualization referring to conveying a narrative or message that engages and informs the audience. [ ]
   A. Exploration and Discovery B. Communication
   C. Decision-Making D. Storytelling
4. What is the primary purpose of considering context when creating data visualizations?
   A. To make the visualization more colorful.
   B. To ensure that the data is presented accurately and meaningfully.
   C. To hide data that is not relevant.
   D. To add decorative elements to the visualization.
5. In data visualization, what does the term "contextual information" refer to? [ ]
   A. Information that is presented in isolation from the data.
   B. Additional information that helps viewers understand the data and its significance.
   C. Data that is irrelevant to the visualization
   D. Information that is intentionally kept hidden from the audience.
6. What is the primary goal of exploratory data analysis (EDA)? [ ]
   A. To summarize and describe the main features of a dataset.
   B. To communicate insights and findings to a non-technical audience.
   C. To test hypotheses and make predictions.
   D. To prepare data for machine learning algorithms.

7. Which of the following techniques is commonly used in exploratory data analysis to identify patterns and relationships in data? [ ]

   A. Hypothesis testing. B. Data visualization. C. Linear regression. D. Machine learning.

8. In explanatory data analysis, what is the emphasis on? [ ]

   A. Discovering hidden patterns and relationships.
   B. Communicating insights and findings clearly to others.
   C. Preparing data for statistical modeling.
   D. Collecting raw data from various sources

9. How can you ensure effective communication in data visualization for a non-technical audience? [ ]

   A. By using complex jargon and technical terms.
   B. By adding as much data as possible to make it comprehensive.
   C. By providing clear labels, legends, and explanations.
   D. By minimizing the use of colors and shapes.

10. In data visualization “Who” Indicates [ ]

    A. Presenter B. Tone C. Action D. Mechanism

11. In data visualization “What” Indicates [ ]

    A. Presenter B. Audience C. Action D. Tone

12. Which refers to a document that combines elements of both a traditional written document and a presentation slide deck. [ ]

    A. Slideument B. Word Document C. Presentation D.E-mail

13. Which of the following represents a hierarchy? [ ]

    A.  B.  C.  D. 

14. Which of the following represents a Relationship? [ ]

A.  B.  C.  D. 

15. How many visual types are available in data visualization? [ ]

1. 1 2. 2 3. 3 4. 4

16. What is the primary purpose of incorporating simple text in data visualizations?

A. To make the visualization more colorful. B. To add decorative elements.

C. To provide context and explanations. D. To replace visual elements entirely.

17. In data visualization, which role does simple text often play? [ ]

A. To confuse the audience. B. To add complexity to the visualization.

C. To annotate, label, and clarify visual elements. D. To replace numerical data.

18. Why is it important to use concise and clear text in data visualizations? [ ]

A. To make the visualization more visually appealing.

B. To confuse the audience and challenge their understanding.

C. To ensure that viewers understand the data and its significance.

D. To replace the need for visual elements.

19. In a table, what do the rows typically represent? [ ]

A. Data categories or variables B. Individual data points or observations

C. Numeric values only D. Color-coded legends

20. How are columns usually used in a data visualization table? [ ]

A. To separate rows from each other

B. To display data labels and titles

C. To provide additional information or attributes for each data point

D. To represent data in the form of bars and lines

21. When should you consider using a table as a data visualization tool? [ ]
    A. When you want to simplify data presentation for a non-technical audience
    B. When you need to convey trends and patterns visually
    C. When you have a large dataset with multiple variables and precise values
    D. When you want to create an artistic and decorative visualization
22. What is the primary purpose of a heat map table in data visualization? [ ]
    A. To display numerical data in alphabetical order
    B. To represent data using colors to highlight patterns and variations.
    C. To replace traditional bar charts and line graphs
    D. To add decorative elements to a visualization
23. How are data values typically represented in a heat map table? [ ]
    A. Using 3D effects and shadows B. Using varying shades of color
    C. With animated transitions D. With different font styles and sizes
24. What do the colors in a heat map table indicate? [ ]
    A. The temperature in the room where the data was collected.
    B. The significance of each data point
    C. The value or intensity of the data at each cell
    D. The type of data being displayed.
25. What is the primary purpose of a scatter plot in data visualization? [ ]
    A. To display data in a series of vertical bars.
    B. To represent hierarchical data relationships.
    C. To show the distribution and relationship between two numerical variables.
    D. To replace textual data with visual elements.
26. What is the primary visual element used in a scatter plot to represent data points?
    A. Bars B. Dots or points C. Lines D. Pie slices
27. What does the position of a data point on a scatter plot indicate? [ ]
    A. The color of the data point B. The size of the data point
    C. The values of the two variables are being compared. D. The category of the data point.

28. What is the primary purpose of a line graph in data visualization? [ ]

A. To display data using rectangular bars.

B. To represent data in a scatterplot format.

C. To show the trend or change in data over a continuous interval or time.

D. To replace textual data with visual elements.

29. In a line graph, what type of data is typically plotted on the x-axis (horizontal axis)?

A. Categorical data B. Discrete data

C. Time or a continuous variable D. Textual data

30. What does the height or length of a bar in a bar graph typically represent? [ ]

A. The color of the bar B. The size of the bar

C. The values of the data being compared. D. The category of the bar

UNIT – II (BIT BANK)

31. What is data clutter in the context of data visualization? [ ]
    A. An intentional design choice to make visualizations more engaging.
    B. The use of vibrant colors and decorative elements in a chart
    C. Excessive or irrelevant visual elements hinder data comprehension.
    D. The inclusion of detailed labels for every data point
32. What do the Gestalt principles of visual perception describe? [ ]
    A. Mathematical calculations are involved in visual perception.
    B. The innate human ability to see in color.
    C. The principles that govern how we perceive and interpret visual elements.
    D. The history of art and its impact on perception.
33. When multiple objects are aligned in a straight line, which Gestalt principle is at play?
    A. Closure B. Continuity C. Similarity D. Proximity
34. Which Gestalt principle suggests that people tend to group elements that are close to each other in space? [ ]
    A. Proximity B. Similarity C. Closure D. Continuity
35. Which Gestalt principle involves perceiving objects as whole figures, even when they are composed of smaller parts? [ ]
    A. Similarity B. Proximity C. Continuity D. Closure
36. What is the first step in decluttering a data visualization? [ ]
    A. Adding more data points and labels. B. Removing all visual elements except for the data.
    C. Choosing a color palette for the visualization. D. Creating a cluttered and complex design.
37. What is iconic memory primarily associated with? [ ]
    A. Short-term memory. B. Visual sensory information
    C. Long-term memory consolidation. D. Auditory sensory information
38. How long does iconic memory typically last? [ ]

A. A few seconds B. Several minutes C. Up to 30 minutes D. Less than a second

39. What is the primary function of iconic memory in the information processing system? [ ]
    A. To store information for long-term use B. To hold and process information for immediate use.
    C. To consolidate memories during sleep. D. To store emotional experiences.
40. How long does short-term memory typically last without rehearsal? [ ]
    A. A lifetime. B. About 24 hours C.30 seconds to a minute D. Several years
41. What cognitive process is often used to maintain information in short-term memory for longer periods? [ ]
    A. Encoding B. Retrieval C. Rehearsal D. Consolidation
42. Which type of memory is responsible for the storage of information over an extended period, potentially a lifetime? [ ]
    A. Iconic memory. B. Short-term memory. C. Long-term memory. D. Sensory memory
43. What are preattentive attributes in data visualization? [ ]
    A. Visual cues that capture viewers' attention quickly and automatically
    B. Elements that require in-depth analysis and focus to understand.
    C. Data points that are irrelevant for the visualization
    D. The visual representation of categorical data
44. What does the preattentive attribute "enclosure" refer to in data visualization? [ ]
    A. The use of bold colors to draw attention.
    B. The practice of placing data points far apart.
    C. Grouping related data points with a boundary or container.
    D. The use of distinct fonts for text labels
45. Which of the following is an example of a preattentive attribute that can be used to encode ordinal data, such as low, medium, and high values? [ ]
    A. Color B. Shape C. Proximity D. Texture

46. What is the term for the preattentive attribute that involves the alignment of text to the left, right, or center of a visual element? [ ]

    A. Orientation B. Proximity C. Position D. Color

47. Which preattentive attribute of text is often used to convey a sense of hierarchy or importance in headings and subheadings? [ ]

    A. Color B. Texture C. Size D. Position

48. What is the term for the preattentive attribute that involves the arrangement of data points in a specific order or sequence within a graph? [ ]

    A. Texture B. Proximity C. Position D. Orientation

49. In data visualization, what is the primary advantage of utilizing preattentive attributes of graphs? [ ]

    A. They require significant time and effort to process.
    B. They can help viewers quickly identify patterns and trends in the data.
    C. They are primarily used for decorative purposes.
    D. They are only suitable for representing text labels.

50. In data visualization, what does the preattentive attribute "length" refer to when applied to graphs? [ ]

    A. The distance between data points
    B. The horizontal arrangement of bars in a bar chart
    C. The use of bold lines in line graphs
    D. The vertical size of data markers

UNIT – III (BIT BANK)

1. What is the first step in the process of communicating data?
    a) Data collection
    b) Data analysis
    c) Data visualization
    d) Data interpretation
2. Which model of communication represents the sender, message, channel, and receiver?
    a) Linear model
    b) Circular model
    c) Interactive model
    d) Transactional model
3. What are the three common types of communication problems in data communication?
    a) Noise, feedback, redundancy
    b) Interference, context, tone
    c) Clarity, speed, accuracy
    d) Encoding, decoding, understanding
4. Which of the following is NOT one of the six principles of communicating data?
    a) Relevance
    b) Simplicity
    c) Consistency
    d) Complexity
5. What is Tableau primarily used for?
    a) Video editing
    b) Data visualization and analytics
    c) Web development
    d) Social media management

6. Which of the following is NOT a Tableau product?
    a) Tableau Desktop
    b) Tableau Server
    c) Tableau Mobile
    d) Tableau Calculator

7. What is the purpose of "Connecting to data" in Tableau?
    a) Social networking
    b) Data analysis and visualization
    c) Online gaming
    d) Graphic design

8. What does "communicating how much" typically refer to?
    e) Measuring quantities
    f) Describing color
    g) Explaining processes
    h) Comparing sizes

9. What is a ratio in mathematics?
    a. A fraction that represents the relationship between two quantities
    b. A whole number
    c. A percentage
    d. A decimal number

10. Which of the following represents a rate?
    a. 3:1
    b. 5%
    c. 0.25
    d. 2/3

QUESTION BANK & BIT BANK

11. If a car travels 300 miles in 5 hours, what is its rate in miles per hour?
    a. 60 mph
    b. 150 mph
    c. 45 mph
    d. 35 mph
12. In the linear model of communication, communication flows in which direction?
    a) Unidirectional
    b) Bidirectional
    c) Multidirectional
    d) Random
13. What is the key advantage of Tableau Server?
    a) Offline data storage
    b) Cloud-based data analysis
    c) Collaboration and sharing of Tableau workbooks
    d) Real-time data collection
14. Which of the following is an example of "communicating how many"?
    a) The temperature is 75 degrees Fahrenheit.
    b) There are 7 red apples in the basket.
    c) The stock price increased by 20%.
    d) The car traveled 300 miles.
15. True or False: The transactional model of communication only involves one-way communication.
16. True or False: Redundancy is a desirable quality in data communication.
17. True or False: Tableau is primarily used for data analysis and visualization.
18. True or False: Tableau Desktop is a Tableau product.
19. True or False: "Connecting to data" in Tableau is unrelated to data analysis.
20. True or False: "Communicating how much" is about describing qualities.

UNIT – IV (BIT BANK)

1. What type of relationship does "part to whole" typically represent in data visualization?
    a) Ratio
    b) Percentage
    c) Proportion
    d) Mean

2. When comparing data from the present to historical data, what kind of analysis are you performing?
    a) Comparative analysis
    b) Trend analysis
    c) Descriptive analysis
    d) Predictive analysis

3. When you evaluate how close actual data is to a predefined target, you are examining:
    a) Part-to-whole relationships
    b) Current-to-historical data
    c) Actual-to-target comparisons
    d) Mean and median values
4. What does the mean represent in a set of data?
    a) The middle value
    b) The most frequently occurring value
    c) The average value
    d) The highest value
5. When respecting variation in data visualization, what principle are you following?
    a) Reducing all variation
    b) Ignoring all variation
    c) Respecting and understanding variation

d) Maximizing variation

6. Control charts are commonly used to monitor:
    a) Historical data
    b) Uncertainty in data
    c) Variation over time
    d) Mean values

7. What is "uncertainty" in the context of data visualization?
    a) The absence of data
    b) The range of possible values
    c) A constant value
    d) The mean value

8. True or False: Proportions represent the percentage of the whole that a part comprises.

9. True or False: When comparing current data to historical data, you are performing predictive analysis.

10. True or False: When actual data matches a predefined target, no further analysis is required.

11. True or False: The mean is the same as the median in any dataset.

12. True or False: Control charts help in monitoring variation over time.

13. True or False: Respecting variation means eliminating all variation in the data.

14. True or False: Uncertainty in data refers to the range of possible values.

15. True or False: The median is the value that occurs most frequently in a dataset.

16. Section 5: Multiple Choice
17. In a part-to-whole relationship, if a part is 25% of the whole, what is the proportion represented as?
    a) 0.25
    b) 25
    c) 1/4
    d) 4

18. Which of the following is a measure of central tendency in a dataset?
    a) Range
    b) Variance
    c) Mean
    d) Standard deviation

19. Control charts are primarily used for:
    a) Comparing data to historical data
    b) Analyzing proportions
    c) Monitoring variation over time
    d) Calculating mean values

20. When examining variation over time, what type of chart is typically used?
    a) Line chart
    b) Pie chart
    c) Scatter plot
    d) Bar chart

UNIT – V (BIT BANK)

1. What type of chart is commonly used to visualize the relationships between two numerical variables?
    a. Line chart
    b. Scatterplot
    c. Stacked bar chart
    d. Radar chart

2. In a stacked bar chart, what does each bar represent?
    a. Individual data points
    b. The total sum of all data points
    c. Data points over time
    d. Relationships between data points

3. What does regression analysis help you determine when used with scatterplots?
    a. Patterns and trends in data
    b. The number of data points
    c. The proportion of data
    d. The width of data points

4. The quadrant chart divides data into how many quadrants?
    a. Two
    b. Three
    c. Four
    d. Five

QUESTION BANK & BIT BANK

5. What is the origin of time charts primarily used for?
    a. Visualizing relationships
    b. Displaying data across a timeline
    c. Showing proportions
    d. Analyzing scatterplots
6. In a line chart, what is typically displayed on the x-axis?
    a. Time
    b. Categories
    c. Numeric values
    d. Geographical locations
7. What does a dual-axis line chart allow you to do?
    a. Display two different datasets with the same scale
    b. Create a 3D line chart
    c. Compare two unrelated variables
    d. Show multiple line charts on one axis
8. In a connected scatterplot, what connects the data points?
    a. Lines
    b. Curves
    c. Arrows
    d. Dots
9. What is the purpose of the "date field type" in data visualization?
    a. Categorize data points by dates
    b. Calculate regression lines
    c. Display proportions
    d. Group data points by colors
10. What does the term "seasonality" refer to in data visualization?
    a. The concept of time intervals in data
    b. The use of multiple seasons in visualizations
    c. The relationship between data and the weather
    d. The division of data by quarters

11. What type of chart is used to show the change in values between two points in time?
    a. Slopegraph
    b. Radar chart
    c. Scatterplot
    d. Pie chart
12. What is the primary purpose of a circle map in data visualization?
    a. Display data points with circles
    b. Show geographical locations
    c. Visualize proportions
    d. Connect data points with lines
13. What does a filled map represent in data visualization?
    a. Geographic regions filled with colors based on data values
    b. Points on a map connected by lines
    c. A 3D representation of geographical data
    d. Maps with no data displayed
14. In dual-encoded maps, how are two types of data represented on the same map?
    a. Through different colors and shapes
    b. By using two separate maps
    c. With a single color scheme
    d. By encoding them in the legend
15. What type of map can represent data using various shades or patterns to convey information?
    a. Choropleth map
    b. Circle map
    c. Filled map
    d. Radar map

16. In a scatterplot, what type of relationship can be observed when data points cluster closely together?
    a. A positive correlation
    b. A negative correlation
    c. No correlation
    d. A perfect correlation
17. What is the primary purpose of a slopegraph in data visualization?
    a. Display geographical data
    b. Show the change in values between two time points
    c. Visualize data distribution
    d. Present proportions within a timeline
18. What is the key feature of a choropleth map?
    a. Use of circles to represent data points
    b. Display of data points as stacked bars
    c. Representation of data with colors based on geographical regions
    d. Connection of data points with lines on a map
19. In a dual-encoded map, what is one common method of encoding data values?
    a. Different line styles
    b. Different marker sizes
    c. Different map projections
    d. Different color schemes

20. When using a radar chart in data visualization, what does each axis typically represent?
    a. Time intervals
    b. Geographical locations
    c. Different data categories or variables
    d. Proportions

QUESTION BANK & BIT BANK
ASSIGNMENT – I & II

CASE STUDY 1

- Imagine you are working as a data analyst for an e-commerce company. You have been provided with a dataset containing information about customer demographics, product details, and sales transactions over the past year. Your goal is to create a meaningful data visualization that can help the company's marketing team understand customer purchasing behavior.

  Task:
- Using the provided dataset, create a data visualization that answers the following question:
- **"What are the top three product categories that generate the highest revenue, and how does their sales performance vary over the course of the year?"**
- In your visualization, consider factors such as:

1. Product category sales trends over the past year.
2. Monthly variations in sales within each of the top three product categories.
3. Any seasonality or patterns in customer purchasing behavior related to these categories.

- You may choose any type of data visualization (e.g., line chart, bar chart, heatmap, etc.) and ensure that it is visually appealing, informative, and effectively communicates the insights related to the question. Include appropriate labels, legends, and a title to make your visualization easy to understand.
- Additionally, provide a brief written interpretation of your visualization, highlighting the key insights and any recommendations that can be derived from the data.

- Dataset: (Provide a sample dataset with relevant columns such as product category, sales amount, date of purchase, etc., or specify where students can access such data for analysis.)

CASE STUDY 2

- You are tasked with analyzing a dataset containing information about patient outcomes for different medical procedures and treatments at a hospital. The dataset includes patient characteristics (e.g., age, gender), the type of procedure or treatment performed, and the associated outcomes (e.g., recovery time, complication rates). Your goal is to create a data visualization that helps the hospital make informed decisions about treatment effectiveness.

Task:

- Using the provided dataset, create a data visualization that addresses the following question:
- **"How does the age of patients impact the recovery time for two specific medical procedures, Procedure A and Procedure B?"**
- In your visualization, consider the following:

1. Create separate scatter plots for Procedure A and Procedure B.
2. Plot patient age on the x-axis and recovery time on the y-axis.
3. Use different colors or markers to distinguish between the two procedures.
4. Add appropriate labels, titles, and axes to make the visualization clear and informative.

- Additionally, calculate and display any relevant statistical measures (e.g., mean recovery time) for each procedure and age group if necessary.

- After creating the visualization, provide a brief written interpretation, including any insights about the relationship between patient age and recovery time for Procedures A and B. Include recommendations for the hospital based on your findings.

- Dataset: (Provide a sample dataset with columns like patient age, procedure type, recovery time, and any other relevant variables, or specify where students can access such data for analysis.

CASE STUDY 3

- You are working with a dataset that contains various environmental indicators related to climate change, including temperature records, CO2 levels, and deforestation rates, spanning the past few decades. Your goal is to create a data visualization that helps communicate the relationship between temperature changes and CO2 levels.

  Task:
- Using the provided dataset, create a data visualization that addresses the following question:
- **"How have global average temperatures and CO2 concentrations changed over time, and is there a discernible relationship between these two variables?"**

- In your visualization, consider the following:
  1. Plot global average temperatures on one axis (e.g., y-axis) and CO2 concentrations on the other axis (e.g., x-axis).
  2. Use different colors or markers to distinguish between temperature and CO2 data points.
  3. Time should be on the x-axis (e.g., years or decades).
  4. Ensure that the visualization is clear, labeled, and includes a title.
- Additionally, you may want to include annotations or trendlines to help interpret any patterns or relationships in the data.
- After creating the visualization, provide a brief written interpretation, including any insights about the relationship between global temperatures and CO2 concentrations over time. Discuss any potential implications for climate change policy or action based on your findings.
- Dataset: (Provide a sample dataset with columns like year, global temperature, CO2 concentration, or specify where students can access such data for analysis.)

CASE STUDY 4

- Imagine you are working as a financial analyst, and you have access to historical stock price data for a selected set of companies over the past five years. Your goal is to create a data visualization that helps investors assess the risk and return associated with different stocks in the portfolio.

  Task:

- Using the provided dataset, create a data visualization that addresses the following question:
- **"How does the volatility (standard deviation of daily returns) of three different stocks compare over the past five years, and how does this relate to their average annual returns?"**
- In your visualization, consider the following:

1. Select three different stocks from the dataset for comparison.
2. Calculate the standard deviation of daily returns for each selected stock over the five-year period.
3. Calculate the average annual return for each stock.
4. Use a scatter plot or bubble chart to visualize the relationship between volatility (standard deviation) and average annual return. Place volatility on one axis and average annual return on the other.
5. Ensure that the visualization includes clear labels, a title, and any necessary legends.

- Additionally, you may want to include trendlines or annotations to highlight any notable findings or trends.
- After creating the visualization, provide a brief written interpretation, including any insights about the risk-return trade-off among the selected stocks. Offer recommendations for investors based on your findings.
- Dataset: (Provide a sample dataset with columns like stock symbol, date, daily returns, or specify where students can access such data for analysis.)

CASE STUDY 5

- You are working as a data analyst for a school district, and you have access to a comprehensive dataset containing information about student demographics, standardized test scores, teacher qualifications, and classroom resources across multiple schools. Your task is to create a data visualization that helps school administrators identify factors influencing student performance.

  Task:

- Using the provided dataset, create a data visualization that addresses the following question:
- "How does the student-to-teacher ratio in different schools within the district correlate with the average standardized test scores of students in those schools?"
- In your visualization, consider the following:

1. Plot the student-to-teacher ratio on one axis (e.g., x-axis) and the average standardized test scores on the other axis (e.g., y-axis).
2. Use a scatter plot or a similar visualization type to display the relationship between these two variables.
3. Differentiate schools within the district using colors or markers.
4. Ensure that the visualization includes appropriate labels, a title, and any necessary legends.

- Additionally, you may want to calculate and display correlation coefficients to quantify the strength and direction of the relationship.
- After creating the visualization, provide a brief written interpretation, including any insights regarding the impact of the student-to-teacher ratio on student performance. Offer recommendations for school district administrators based on your findings.
- Dataset: (Provide a sample dataset with columns like school name, student-to-teacher ratio, average standardized test scores, or specify where students can access such data for analysis.)

PBR VISVODAYA INSTITUTE OF TECHNOLOGY & SCIENCE

SUBJECT: DATA VISUALIZATION

QUESTION BANK & BIT BANK