

Birla Institute of Technology & Science, Pilani
Work Integrated Learning Programmes Division
Second Semester 2022-2023

Mid-Semester Test
(EC-2 Regular)

Course No. : DSECLZG555
Course Title : Data Visualization and Interpretation
Nature of Exam : Open Book
Weightage : 30%
Duration : 2 Hours
Date of Exam : 15 July 2023 (FN)

No. of Pages	= 03
No. of Questions	= 05

Note to Students:

1. Please follow all the *Instructions to Candidates* given on the cover page of the answer book.
2. All parts of a question should be answered consecutively. Each answer should start from a fresh page.
3. Assumptions made if any, should be stated clearly at the beginning of your answer.

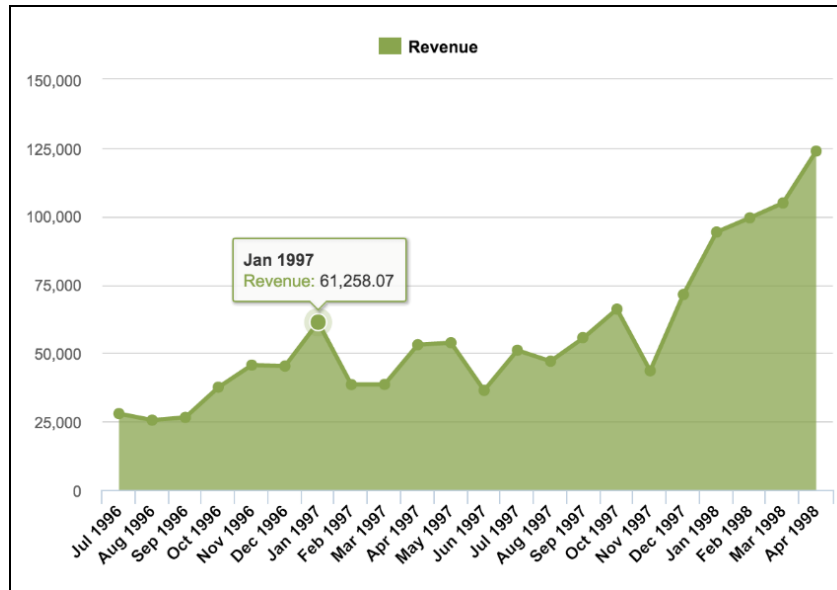
Please Note: Choice of chart type, Decluttering, Effective use of Gestalt principles, Strategic use of pre-attentive attributes, Logical ordering, and effective use of text -contribute towards the credit. Most effective chart fetches maximum credit. [Pen and paper drawing would suffice. No additional credit for usage of any Viz Tools for creating visuals]

- Q.1.** You work as a data analyst for a marketing agency, and you have been tasked with analyzing the effectiveness of a recent social media marketing campaign. The campaign aimed to promote a new product launch and increase brand awareness among the target audience. After conducting a comprehensive analysis of the campaign data, you have discovered some interesting insights and want to present your findings to the marketing team.
- Using the 3 minutes story telling approach, craft a compelling narrative (2 marks) and identify the big idea (2 marks) that encapsulates the key findings of the social media marketing campaign analysis. **[4 Marks]**

Answer Key:

- **3 Minutes story telling:** For a social media marketing campaign, the 3 minutes story telling narrative ideally includes components such as, campaign goals, data sources and metrics, key findings, audience characteristic analysis, impact of storytelling narrative, recommendations followed by a conclusion.
(The most effective narrative fetches full marks – 2M)
- **Big idea:** The big idea statement typically conveys the “So what” boiled down further to a single sentence statement. It should articulate your point of view, convey what is at stake and be written in one complete sentence.
Example: *“The stake in analyzing the effectiveness of the recent social media marketing campaign lies in uncovering how leveraging authentic engagement and personalization not only boosts brand awareness but also holds the potential to maximize the impact of the new product launch among our target audience”*
(The most effective statement fetches full marks – 2M)

- Q.2.** When analyzing the provided area chart visualization, how effectively can the gestalt principles and pre-attentive attributes of visualization be applied and improve the visualization context? Outline the steps and explain how each step aligns with the gestalt principles (e.g., proximity, similarity, closure, enclosure, and focal point) (2 marks) and pre-attentive attributes (2 marks) (e.g., color, size, and position). **[4 Marks]**



Answer Key:

- At least 2 gestalt principles (1M for each) and 2 pre-attentive attributes (1M for each) should be suggested.

For Example,

Gestalt principles:

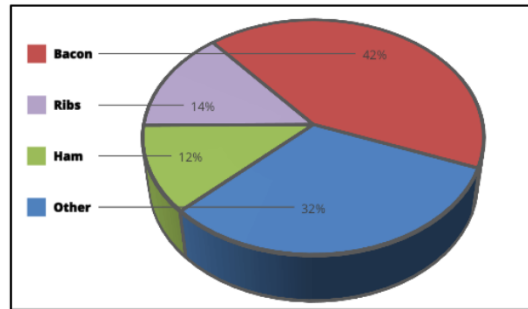
- **Proximity** - Data points representing revenue for each year are placed close to each other along the x-axis, emphasizing their relation in time.
- **Focal Point** - The highest or lowest point in the area chart (peak or trough) acts as a focal point, highlighting significant revenue fluctuations.

Pre-attentive attributes:

- **Color** - Introducing color gradients or shading can visually convey revenue patterns and trends over the years.
- **Position** - Placement of text labels or annotations directly on or near the data points of interest

(The most appropriate principles and attributes fetches full marks – 4M)

- Q.3.** Based on the given pie chart, which illustrates the distribution of meat preferences among different individuals, each slice represents a specific meat category. The size of each slice indicates the proportion of preference allocated to that meat category. Larger slices indicate a higher level of interest in that preference. [6 Marks]



- A.** State the potential limitations or challenges associated with using a pie chart for this type of data visualization [3 Mark]
- B.** Propose and draw an alternative effective chart that would be more effective in presenting the meat preferences with a justification [3 Mark]

Answer Key:

- A. At least 3 potential limitations or challenges associated with pie chart (1M for each)**

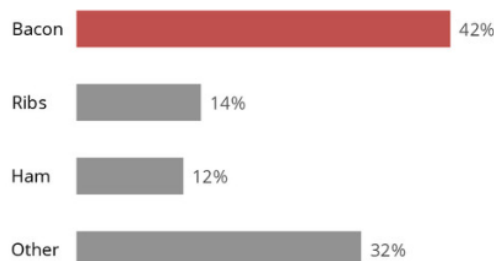
For example,

- **Difficulty comparing small segments:** Viewer struggles to differentiate and compare segments when they are numerous or small in size.
- **Limited representation of data:** Pie charts work best with a few distinct categories; they become cluttered with too many categories or similar proportions.
- **Inefficient for trends over time or multiple groups:** Pie charts are static and don't effectively display data changes across time or different groups; other chart types are more suitable.

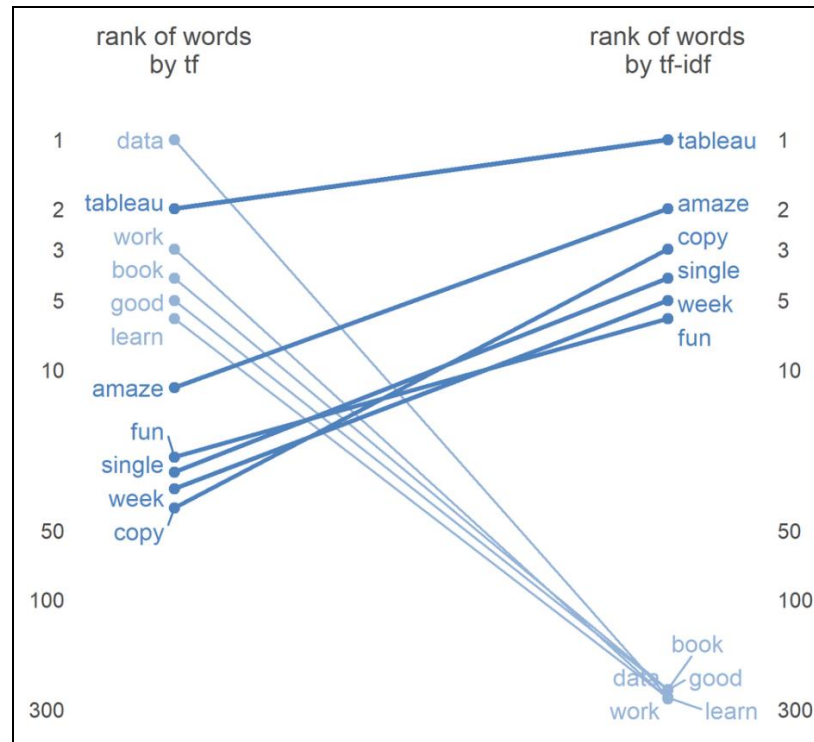
- B. A most effective alternative for a 3D pie chart (chart selection – 1M), reason/justification for chart selection (1M), and the alternative visual – 1M.**

For example,

- **Bar Chart** - A bar chart can be used to compare the meat preferences side by side. Each meat category is represented as a separate bar, and the length of each bar indicates the proportion or count of individuals preferring that meat type. This chart allows for easy comparison and precise evaluation of preferences.



- Q.4.** Based on the concept of term frequency (tf) and term frequency–inverse document frequency (tf-idf), which measure the importance of words in a document collection or corpus, the provided visual displays the ranking of the most common words associated with data visualization. To gain deeper insights and enhance data analysis, **propose and draw four alternative visuals** (2 Mark for each visual) that effectively represent and analyze the data. [8 Marks]



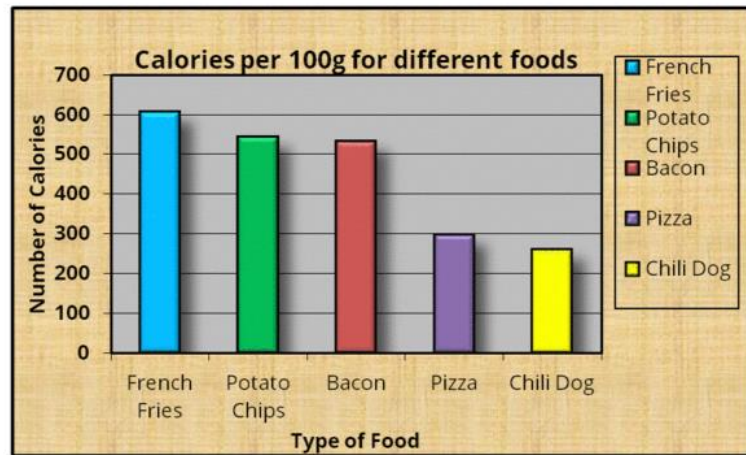
Answer key:

1. The slope graph is a useful data visualization tool for comparing two time periods or points of comparison, but it also has some limitations. However to avoid some limitations such as comparing diverging trends, data precision, subjective interpretation, alternative visuals shall be used.
2. At least 4 alternative visuals (bar, stacked bar, area, line, etc.,) should be suggested and justified effectively.

(2 marks for each visual – (1M for visual + 1M for Justification))

(The most effective alternative visuals and justification fetches full marks – 8M)

Q.5. Based on the data related to calories per 100g for different food of people interest, the provided visual is cluttered and requires improvements. Considering the concepts learned in class, propose the necessary changes (at least 5 steps – 5 marks) to **declutter the graph** and present a refined version of the visual (3 marks). **[8 Marks]**



Answer key:

By removing these superfluous elements, the given chart can become much cleaner and more focused on its primary goal to compare the calorie values of bacon with other foods. A minimalist design would enhance data readability and ensure that viewers can quickly grasp the essential information without being distracted by unnecessary visual elements.

- **At least 5 changes – (5 marks each)**
- **Step-by-step visuals (2M) to final visual (1M)**

For Example,

Changes for decluttering:

1. **Tan Background Border:** The tan background border around the chart is unnecessary as it adds visual clutter and does not provide any relevant information. A simple white background or no border would be more appropriate to keep the focus on the data.
2. **Grey Background Color:** The grey background color behind the bars is unnecessary and may distract from the main purpose of the chart. A plain white background would be sufficient, ensuring the bars stand out clearly.
3. **3-D Effect on the Bars:** The 3D effect applied to the bars is unnecessary and can distort the visual perception of the data. It adds unnecessary depth and can make it challenging to accurately compare the bar lengths.
4. **Legend:** Since the bars are already labeled with the food items they represent, the legend becomes redundant and adds unnecessary visual elements. Removing the legend would streamline the chart and make it more concise.
5. **Colors of Bars:** If the colors of the bars do not signify any specific meaning or grouping, they can be seen as superfluous. Using a simple, consistent color scheme or even a single color for all bars would suffice for the purpose of comparing calories.

Refined Visual:

