KUMARAGURU
college of technology
character is life

M.TECH DEGREE EXAMINATIONS: NOV/ DEC 2024

(Regulation 2024)

First Semester

DATA SCIENCE

24DSI503: Data Analysis and Visualization

COURSE OUTCOMES

CO1: Understand advanced data analysis techniques and their application as tools for extracting meaningful insights from complex and large datasets.

CO2: Identify appropriate visualization methods and tools to effectively represent data, ensuring clarity, accuracy, and impact in the communication of analytical results.

CO3: Analyze the effectiveness of various visualization techniques by examining their ability to convey patterns, trends, and relationships within the data, while adhering to best practices in visual communication

CO4: Develop advanced data visualizations and interactive dashboards that integrate real-time data and custom visual components, facilitating dynamic and insightful data exploration.

CO5: Apply advanced data analysis and visualization techniques to domain-specific challenges in business, healthcare, finance, and social media, demonstrating the ability to deliver actionable insights and informed decision-making.

Time: Three Hours   **Maximum Marks: 100**

PART A (4*20 = 80 Marks)

1. Scenario: A marketing team is analyzing the network of influencers and their followers on a social media platform. The goal is to visualize how influencers are connected to their followers and how information might flow through the network.

    a) Describe how you would use NetworkX to create a network graph that illustrates the connections between influencers and their followers. What would be the key steps in constructing this network?   10   CO2   [K_3]

    b) How would you highlight important nodes and their connections in the visualization? Describe the process you would use to make the visualization informative and visually appealing.   10   CO2   [K_2]

2.	a)	Analyze the process of creating a line plot using Matplotlib to visualize and compare the sales trends of three main product categories over five years. Discuss the steps involved in preparing the data, plotting the graph, and the insights that can be drawn from the trends observed in the visualization.	8	CO3	[K4]
	b)	Analyze and compare the effectiveness of different visualization techniques in conveying patterns, trends, and relationships within a dataset.	8	CO3	[K4]
	c)	What are the different ways of plotting geographic data using Geopandas, and how do you choose between them based on the type of spatial data you have?	4	CO3	[K2]
3.	a)	When preparing a dataset for training a machine learning model, what are the key steps involved in data normalization, and how does normalization impact the model's performance?	10	CO1	[K2]
	b)	Describe the process of performing a hypothesis test and explain the role of confidence intervals in this process. How can confidence intervals provide additional insights to complement the outcomes of a hypothesis test?	10	CO1	[K2]
4.	a)	Make use your knowledge of visualization tools such as Tableau, Power BI, and D3.js to assess how each can improve data communication. Compare their key features and determine which tool is best suited for particular scenarios, such as building interactive dashboards and designing complex visualizations.	8	CO4	[K4]
	b)	Examine the ethical responsibilities of a data analyst in creating visualizations. Provide examples of practices that could mislead the audience and explain how these issues can be prevented.	8	CO4	[K3]
	c)	Explain the role of Apache Kafka in managing streaming data and how it can be integrated with visualization tools	4	CO4	[K2]

Answer any ONE Question
PART B (1*20 = 20 Marks)

5. a) How can dashboards in healthcare help track hospital performance metrics, such as patient wait times and resource utilization? Discuss the visualization techniques involved.   8   CO5   [K₃]

b) When creating a dashboard for users with hearing impairments, what accessibility features would you include to ensure all important information is conveyed effectively?   8   CO5   [K₃]

c) Articulate the challenges involved in visualizing patient data and explain the strategies you would use to overcome them.   4   CO5   [K₃]

OR

6. a) Discuss the ethical considerations when analyzing and visualizing social media data. How can data privacy and transparency be ensured?   8   CO5   [K₂]

b) A financial analyst needs to present the historical performance of a stock over the past five years to stakeholders. Identify and explain the visualization techniques would be most effective.   6   CO5   [K₃]

c) A retail company wants to launch a new marketing campaign aimed at increasing customer engagement. How can predictive modeling be applied to enhance the decision-making process for this campaign?   6   CO5   [K₃]

****