QUESTION BANK FOR INTERNAL-2

Q1. What do you understand by Data Visualisation?

Q2. Explain some of the Tools and Techniques involved in Data Visualisation.

Q3. State some of the important applications of Data Science.

Q4. What do you understand by Mapping Variables into encodings. Explain with examples.

Q5. What do you understand by Retinal Variables. List some of them along with their applications.

Q6. Explain the recent trends and technologies in Data Collection and Analysis.

Q7. State some recent developments in the Application Development methods used in Data Science.

Q8. What is Bokeh in Python. Explain with its applications.

Q9. What do you understand by Visual Encodings. Explain its importance.

Q10. What are some common visual encodings and their appropriate uses?

Q11. How retinal variables are used in different types of charts?

Q12. What are the key principles of effective data visualization?

Q13. Explain the role of storytelling in effective data visualization.

Q14. Discuss how advancements in machine learning have impacted data analysis.

Q15. What are some recent developments in data visualization techniques?

Q16.Explain the role of artificial intelligence in modern data science applications.

Q17.Describe the current trends in application development methods for data science.

Q18.How has big data analytics changed the landscape of data science?

Q19.How is data science applied to solve business problems using data visualization?

Q20.Explain the key technologies used for creating interactive data visualizations.

Q21.How has big data analytics changed the landscape of data science?

CASE STUDY BASED QUESTIONS:

Case Study 1:

You have been given a dataset that contains sales data for a retail company. The dataset includes the following columns:

Date: The date of the sale

Product Category: The category of the product sold (e.g., Electronics, Clothing, Home Goods)

Sales Amount: The total sales amount for the transaction

Region: The region where the sale was made (e.g., North, South, East, West)

Units Sold: The number of units sold

Based on the provided dataset, you need to create a set of visualizations to provide insights into the company's sales performance. Answer the following:

1. Describe the types of visualizations you would use to answer these questions and explain why you chose each type.

2. Discuss how you would handle any potential challenges or issues with the data in your visualizations.

Case Study 2: Data Collection and Preprocessing

Scenario: A multinational retail company wants to launch a loyalty program based on customer data from multiple regions, with data coming from online and physical stores. However, the data from the online stores is well-structured while the data from the physical stores is semi-structured and has missing entries.

Question:

Explain the steps you would take to collect, clean, and manage this data from different sources. How would you handle missing data and ensure data integrity before analysis? Discuss the potential challenges of integrating multiple data sources and how to overcome them.

Case Study 3: Data Analysis for Business Decision-Making

Scenario: A food delivery startup is trying to analyze customer data to optimize delivery routes. The company has historical data on delivery times, customer feedback, and traffic conditions. They want to understand which factors most affect delivery times and improve customer satisfaction.

Question:

2. Describe how you would use data analysis techniques, including statistical methods and machine learning algorithms, to find the factors that affect delivery times. Propose a model that can predict delivery time based on historical data, and explain how you would validate the accuracy of the model.

Case Study 4: Data Visualization for Insights

Scenario: A healthcare organization wants to visualize patient data to identify patterns in the spread of seasonal diseases. The data includes patient age, geographic location, symptoms, and treatment effectiveness. They need a dashboard to help doctors quickly assess the severity and spread of illnesses.

Question:

3. Design a data visualization strategy for this healthcare organization. Discuss which types of visualizations would be most effective for showing trends in disease spread and treatment success. Explain how you would map different data types to visual encodings to create an actionable dashboard for doctors.

Case Study 5: Machine Learning Model Development

Scenario: A bank wants to implement a predictive model to identify potential loan defaulters based on features such as income, credit history, and spending behavior. The goal is to minimize financial risk by identifying risky loan applicants early.

Question:

4. Discuss how you would apply machine learning algorithms like Logistic Regression and Decision Trees to predict loan default risk. How would you choose between these algorithms, and what factors would you consider when evaluating the performance of the

models? Provide insights on feature selection and model validation.

Case Study 6: Recent Trends in Data Science

Scenario: A company wants to stay ahead of the competition by adopting the latest trends in data collection and visualization. They are particularly interested in using IoT (Internet of Things) data from smart devices to monitor machine performance in their manufacturing plants.

Question:

5. Explain how recent trends in data collection techniques, such as IoT data, can improve the company's manufacturing efficiency. Discuss how you would collect, process, and visualize this data to provide real-time insights into machine performance. What technologies would you recommend for managing and visualizing large volumes of IoT data?