**19AD605 - DATA VISUALIZATION & NLP LABORATORY**

**Lab - Viva questions**

**DATA VISUALIZATION**

**Experiments 1: Acquiring and plotting data**

**General Data Acquisition**

1. What are the common methods to acquire data for visualization purposes?
2. How do you handle missing or incomplete data during acquisition?
3. What are the key steps involved in cleaning data before visualization?
4. What is the difference between structured and unstructured data? Provide examples of each.
5. Explain the role of APIs in data acquisition. Can you name some commonly used APIs for data extraction?

**Data Plotting and Visualization**

1. What are the different types of plots available for data visualization, and when would you use each?
2. What is the significance of labeling axes and adding titles to your plots?
3. How can you enhance the readability of a plot for audiences unfamiliar with the dataset?
4. What are the advantages of interactive visualizations compared to static ones? Can you name a tool or library used for interactive plotting?
5. What is the importance of scaling and normalization in data visualization? How does it affect the output plot?

**Experiments 2: Time-series analysis – stock market**

**General Concepts**

1. What is time-series data, and how is it different from other types of data?
2. Why is time-series analysis important for stock market data?
3. What are the key components of a time-series dataset (e.g., stock prices)?  
   *(Hint: Trend, seasonality, and noise)*

**Stock Market Analysis**

1. What are OHLC data points, and how are they used in stock market analysis?  
   *(Open, High, Low, Close prices)*
2. What are candlestick charts, and why are they commonly used in stock market visualization?
3. What is a moving average, and how does it help in analyzing stock trends?
4. What are some common indicators used in stock market time-series analysis?  
   *(e.g., RSI, Bollinger Bands, MACD)*

**Tools and Techniques**

1. Which libraries or tools would you use to perform time-series analysis and visualization in Python?  
   *(Hint: Matplotlib, Pandas, Plotly, Seaborn)*
2. What is the role of data resampling in time-series analysis? Can you provide an example?  
   *(e.g., converting daily data to weekly data)*
3. How do you detect and handle anomalies or outliers in time-series data?

**Experiments 3: Visualization of various massive dataset - Finance - Healthcare - Census - Geospatial**

**General Questions**

1. What is data visualization, and why is it important in handling massive datasets?
2. How does interactivity in visualizations enhance data analysis? Provide examples.

**Finance**

1. Which visualization technique would you use to analyze stock market trends, and why?
2. How can you use heatmaps for financial correlation analysis? Explain with an example.

**Healthcare**

1. What visualization techniques can help monitor the spread of diseases in different regions?
2. How would you visualize patient recovery rates and detect anomalies in the data?

**Census**

1. Explain how population density can be represented effectively. Which visualization technique would you use and why?
2. How would you compare literacy rates across multiple regions using data visualization?

**Geospatial**

1. What are geospatial maps, and how are they used in visualizing environmental or urban data?
2. Describe how to integrate satellite data with visualizations for disaster management.

**Experiments 4:**

**Visualization on Streaming dataset (Stock market dataset, weather forecasting)**

**General Questions**

1. What are streaming datasets, and how do they differ from static datasets in visualization?
2. What challenges arise when visualizing streaming data, and how can these be addressed?

**Stock Market Dataset**

1. Which visualization techniques are best suited for real-time stock market analysis? Provide examples.
2. How would you use candlestick charts to represent stock price movements? Explain their components.
3. What role do dashboards play in monitoring stock market trends in real time?

**Weather Forecasting**

1. How can heatmaps be used for visualizing temperature variations in real-time weather forecasting?
2. Explain how line charts or area charts can help display hourly or daily weather trends.

**Streaming Data Processing and Visualization**

1. How would you handle data latency or missing values in streaming datasets for visualization purposes?
2. What tools and frameworks are commonly used for visualizing streaming data, such as stock prices or weather metrics?

**Advanced and Domain-Specific Questions**

1. How can geospatial maps be integrated with weather forecasting data to visualize storms or rainfall intensity in real time?

**Experiments 5:**

**Text visualization using web analytics**

**General Questions**

1. What is text visualization, and how does it differ from traditional data visualization?
2. What types of textual data are typically analyzed in web analytics? Give examples.

**Tools and Techniques**

1. What is a word cloud, and in what scenarios is it most useful?
2. Explain how sentiment analysis can be visualized using bar charts or pie charts.
3. What is keyword co-occurrence, and how can it be represented visually?

**Applications in Web Analytics**

1. How can text visualization help in understanding user behavior on a website?
2. Describe how topic modeling techniques (e.g., LDA) are used to visualize themes in textual data.

**Advanced Questions**

1. What are the advantages of using hierarchical tree maps for analyzing website content categories?
2. How would you visualize the evolution of trending topics over time using web analytics data?

**Ethical and Practical Considerations**

1. What challenges might arise when visualizing text data from web analytics, and how can they be mitigated?