

# **Data Analysis Task: UPSC CSE Prelims Exam Trends**

## **Objective**

The goal of this task is to evaluate your ability to analyze historical exam data, identify trends, visualize findings, and make data-driven predictions.

## **Task Description**

You have been provided with the previous year's questions of the UPSC CSE Prelims exam from 2013 to 2023. Using this dataset, you are required to perform a comprehensive analysis and present your findings. The task involves the following steps:

## **Steps and Requirements**

### **1. Data Cleaning and Preparation**

- Data Import: Import the provided dataset into a suitable data analysis tool or environment (e.g., Python, R, Excel).
- Data Cleaning: Ensure the data is clean and structured. Handle any missing or inconsistent data.
- Categorization: Categorize the questions based on topics and subjects.

### **2. Trend Analysis**

- Yearly Trends: Analyze the distribution of questions across various subjects and topics for each year from 2013 to 2023.
- Subject Trends: Identify which subjects have been consistently emphasized over the years.
- Topic Trends: Delve deeper into each subject to find out which specific topics within those subjects are frequently covered.

### **3. Visualization**

Create visualizations to illustrate the trends identified. This may include:

- Bar Charts: Showing the number of questions per subject per year.
- Line Graphs: Illustrating trends over time for major subjects and topics.
- Pie Charts: Representing the overall distribution of questions across subjects for a given year.
- Heatmaps: Highlighting the intensity of questions on specific topics over the years.

### **4. Predictive Analysis**

- Based on the historical trends, develop a predictive model to forecast the potential distribution of questions for the 2024 exam.
- Justify your predictions with logical reasoning and statistical support.

## **5. Report Writing**

- Compile a report detailing your methodology, analysis, visualizations, and predictions.
- Provide insights on how the trend might influence the preparation strategies of future candidates.
- Include any assumptions made during the analysis and potential limitations of your predictions.

## **Deliverables**

- Code/Analysis Script: Submit the code or scripts used to perform the analysis.
- Report: A detailed report in PDF format.
- Visualizations: High-quality images or interactive visualizations embedded in the report.

## **Evaluation Criteria**

- Data Handling: Ability to clean, prepare, and structure the data efficiently.
- Analytical Skills: Depth and accuracy of the trend analysis and predictions.
- Visualization: Clarity and effectiveness of the visual representations.
- Predictive Model: Soundness of the methodology and accuracy of the predictions.
- Communication: Quality of the written report and the presentation.

## **Submission Deadline**

Please submit your completed task by 3rd June 2024.