

Python for Data Science
Question Paper for Hybrid Participants

Duration: 3 hours
Maximum Marks: 60

Instructions:

- Attempt all questions
- Read each question carefully before answering.
- All answers should be written in a Python script.

Question 1 (20 Marks)

You are a botanist studying a dataset of flowers (iris.csv) containing information about their species, petal length, petal width, sepal length, and sepal width.

- **a) (10 marks)** Write a Python script that:
 1. Imports the library.
 2. Reads the dataset.
 3. Prints the first 5 rows.
 4. Calculates the average petal length for each flower species.
- **b) (10 marks)** Extend your script to:
 1. Create a new Data containing only rows where the petal length is greater than 5 cm.
 2. Display the descriptive statistics (mean, standard deviation, etc.) for the "sepal length" and "sepal width" columns in this new Data.

Question 2 (20 Marks)

An e-commerce company wants to analyse its customer purchase data (data.csv). Perform the following tasks:

- Import the necessary libraries and load the dataset.
- Find the total number of orders placed by each customer.
- Write a function that calculates the average order value for each customer.
- Create a new Data showing the top 5 customers with the highest average order value.

Question 3 (20 Marks)

You are analysing the World Happiness Report (whr.csv) to understand factors influencing happiness. Perform the following tasks:

- Import the necessary libraries and load the dataset.
- Create a new column named "Happiness Score Category" based on the following criteria:
 - a) Low (Score < 5)
 - b) Medium (5 <= Score < 7)
 - c) High (Score >= 7)

- Write a function that takes a country name as input and returns its happiness score category.
- Find the top 5 countries with the highest happiness score and print their country names and categories.