Part-II Pandas Questions

Assignment: Introduction to Pandas in Python
Objective:
To understand the basics of Pandas for data manipulation, analysis, and visualization, focusing on Series, DataFrames, and essential operations.

Assignment Tasks
Task 1: Working with Series

1. Create a Pandas Series from the following list:
   data = [25, 30, 35, 40, 45]
2. Write a Python program to:
   - Assign custom indices ['A', 'B', 'C', 'D', 'E'] to the Series.
   - Display the first three elements of the Series.
   - Calculate the mean, median, and standard deviation of the Series.

Task 2: Creating and Inspecting DataFrames

1. Create a Pandas DataFrame with the following data:

Name	Age	Gender	Marks
Alice	20	Female	85
Bob	22	Male	78
Carol	19	Female	92
David	21	Male	74
Eve	20	Female	88

2. Write a Python program to:
   - Display the first two rows of the DataFrame.
   - Display the column names, data types, and summary statistics.
   - Add a new column Passed that contains True if Marks >= 80 and False otherwise.

Task 3: Data Selection and Filtering

1. Using the DataFrame from Task 2, write a Python program to:
   - Select and display the Name and Marks columns.
   - Filter and display records where Marks > 80.
   - Display the record of the student with the highest marks.

Task 4: Handling Missing Data

1. Modify the DataFrame from Task 2 by introducing missing values:
2. df.loc[1, 'Marks'] = None
3. df.loc[4, 'Age'] = None
4. Write a Python program to:
    - Identify missing values in the DataFrame.
    - Fill missing values in the Marks column with the column's mean.
    - Drop rows where the Age column has missing values.

Task 5: Grouping and Aggregation

1. Using the DataFrame from Task 2, write a Python program to:
    - Group the data by Gender and calculate the mean age and marks for each gender.
    - Count the number of students in each gender group.

Task 6: Reading and Writing Data

1. Write a Python program to:
    - Save the modified DataFrame from Task 4 to a CSV file named students_data.csv.
    - Read the CSV file into a new DataFrame.
    - Display the first five rows of the newly loaded DataFrame.

Task 7: General

1. Download a sample dataset from Kaggle or use any public dataset.
2. Write a Python program to:
    - Load the dataset using Pandas.
    - Perform exploratory data analysis (EDA) by summarizing key statistics, checking for missing values, and visualizing data trends using Matplotlib or Seaborn.
    - Document your findings in Markdown or comments.

Coding Blocks | Online