Assignment SU8 – Instructions

CMPG111 – Pandas and Matplotlib

SU8_1   Filtering and Selecting

Write a Python program to filter and select data from the given CSV file using Pandas.

- Read the provided CSV file into a Pandas DataFrame.
- 8.1.1. Filter rows where the course is "Databases" and print the resulting DataFrame.
- 8.1.2. Filter rows to identify students who are at-risk of failing, meaning any one of their grades are below 40%. Print the resulting DataFrame.
- 8.1.3. Filter rows to identify students with a distinction in "Programming", meaning that they achieved 75% or greater. Print the resulting DataFrame.

Submit your Python script (*.py) named: *SU8_1.py*

SU8_2   Grouping and Aggregating

Write a Python program to group data by course and calculate aggregate statistics.

- Read the provided CSV file into a Pandas Dataframe.
- Group the data by the "Course" column and calculate the average grades for each course. Print the resulting aggregated DataFrame.

Submit your Python script (*.py) named: *SU8_2.py*

SU8_3   Plotting

Imagine you're a student managing your monthly expenses. Your budget for various categories is as follows:

- Leisure: R500
- Rent: R3500
- Internet: R550
- Food: R2000
- Petrol: R1200

Using matplotlib, visualize this budget allocation. Write a Python code snippet that utilises matplotlib to create a pie chart representing the budget distribution. Ensure that the chart has a suitable title and labels for each slice.

Submit your Python script (*.py) named: *SU8_3.py*

GENERAL REQUIREMENTS:

- Add comments to both your scripts that concisely explain what they do.
- Add your name, surname, and student number as a comment on the first line of both scripts.
- Make sure you supply appropriate headings / sub-headings / labels to ensure that your input and output make sense.
- **The use of AI generated code is strictly prohibited.**

8.1.1. Example Output

```
Course = Databases
     Student ID      Course   Grade
3       23456789   Databases      36
7       23456790   Databases      88
11      23456791   Databases      59
15      23456792   Databases      12
19      23456793   Databases      80
..           ...         ...     ...
315     23456867   Databases      98
319     23456868   Databases      62
323     23456869   Databases      12
327     23456870   Databases      55
331     23456871   Databases      33
```

8.1.2. Example Output

```
Grade < 40
     Student ID                   Course   Grade
3       23456789                Databases      36
4       23456790              Programming      39
14      23456792   User Interface Design      39
15      23456792                Databases      12
17      23456793          Systems Analysis     31
..           ...                     ...     ...
323     23456869                Databases      12
324     23456870              Programming       1
326     23456870   User Interface Design       9
330     23456871   User Interface Design      39
331     23456871                Databases      33
```

8.1.3. Example Output

```
Course > 75
     Student ID      Course    Grade
12      23456792   Programming      76
28      23456796   Programming      90
32      23456797   Programming      91
40      23456799   Programming      86
48      23456801   Programming      86
60      23456804   Programming      81
104     23456815   Programming      89
108     23456816   Programming      86
180     23456834   Programming      89
200     23456839   Programming      88
216     23456843   Programming      80
224     23456845   Programming      84
240     23456849   Programming      80
244     23456850   Programming      75
248     23456851   Programming      94
268     23456856   Programming      81
272     23456857   Programming      96
276     23456858   Programming      82
284     23456860   Programming      92
**************
```

8.2. Example Output

```
Course
Databases                48.469880
Programming              47.204819
Systems Analysis         49.783133
User Interface Design    47.337349
```

8.3. Example Output