23CSE101 Computational Problem Solving

Lab Evaluation – II (Set 1)

Duration: 1 hour   Maximum: 15 Marks

Instructions:

- **Programming must be done in Python IDLE (script mode). The time allotted for completing the program is 55 minutes.**
- **Save the program file using the last five digits of your roll number (e.g., 252xx) and set number as its filename (e.g., 25268_S1.py).**
- **Once the assessment is completed by the concerned teacher (ensure that your code is evaluated before your leave the lab), the file must be uploaded to the FTP server (the IP address will be shared on the spot).**

1. An online shopping platform records customer purchases inside a single long string. Each entry in this string contains a customer's name along with one product category and the corresponding amount spent. The product category may be `'electronics'`, `'fashion'`, or `'grocery'`. Since customers may buy products from multiple categories at different times, the same customer can appear many times in the string, each time with a different category or amount. Your task is **to** manually parse this string without using **eval()**, extract all purchase entries correctly, and build a purchase summary for each customer. After parsing, you must create a purchase book that groups all the spending amounts for every customer according to category, computes their **total spending**, and calculates their ==average spending per category== based on the number of categories they have purchased from. Additionally, you must compute the ==overall average spending per category== across all customers using the available entries. Based on the ==overall average spending per category==, predict which category is likely to generate the highest revenue in the next month. ==The prediction== should be written as a simple conclusion using the highest average value.

   Write a Python function `generate_purchase_report(data_string)` that takes the given string as input and, using appropriate data structures such as dictionaries and lists, processes the data to solve the problem described above.

   Sample Input:
   entries = [
   {'name': 'Arun', 'electronics': 1200},
   {'name': 'Diya', 'fashion': 750},
   {'name': 'Arun', 'grocery': 300},
   {'name': 'Rahul', 'electronics': 2200},
   {'name': 'Diya', 'grocery': 450},
   {'name': 'Arun', 'fashion': 900}
   ]

Sample Output:

'Arun': [1200, 300, 900, 2400, 800.0]
'Diya': [750, 450, 1200, 600.0]
'Rahul': [2200, 2200, 2200.0]

'Electronics': [1200, 2200, 1700.0]
'Fashion': [750, 900, 825.0]
'Grocery': [300, 450, 375.0]

Prediction: 'Electronics' category is likely to generate the highest revenue.