PPSASMAJ304
APPLICATIONS OF IT - BASICS OF PYTHON

Time: 2 hours
Total Marks: 60

Note:

1. **The candidate has the option to either question 3A or question 3B. Rest all questions are mandatory.**
2. **Numbers to the right indicate full marks.**
3. **The candidates will be provided with the formula sheet and graphs (if required) for the examination.**
4. **Use of approved scientific calculators is allowed.**

Q1. Attempt All questions.

A) **5 Marks**

Given an amount of money and a list of coin denominations, what is the minimum number of coins you must select to make up that amount, assuming an unlimited number of coins from each denomination?

Input: {"coins": [1, 2, 5], "amount": 11}
Output: 3

B) **5 Marks**

Write a program which prompts the user to input principal, the rate is 5% annually,
calculate the compound interest (after 10 years)? Plot the interest amount over 9 years on a bar chart?
Formula is: CI = P(1+R/100) ^T-P

C) **5 Marks**

Given a list of integers, find out how many additions are needed for each element so that each integer appears with equal frequency.
Input: [1, 3, 1, 2, 2, 4, 5]
Output: {"3": 1, "4": 2, "5": 2}

Q2. Answer the following **15 Marks**

Answer the following questions, based on the excel file shopping.xlsx with sheet names as cust and details?

a) Read the excel file, store cust sheet in df1 and details sheet in df2. Merge the dataframe based on **cust_id.** **(5)**

b) Find the customer's name with the highest daily total order cost between 2019-02-01 to 2019-05-01. If customer had more than one order on a certain day, sum the order costs on daily basis. Output customer's first name, total cost of their items, and the date. **(5)**

c) Using pie charts analyze the order category and total order cost only for city Miami. In legends, also display the % share in total cost. **(5)**

Q3. Attempt question 3A or question 3B. **30 Marks**

A. Load the Matches.csv dataset. The dataset contains information.
about the premier league football matches. Predict the results using logistic regression.

.

a. Read the dataset and split the data in X and Y.

i. Identify independent variables, consider it X **(3)**

ii. Identify dependent variables, consider it Y **(3)**

b. Perform feature engineering, fill all missing values and Perform label encoding or One hot encoding whenever required. **(6)**

c. Perform train-test split with test ratio 15% **(3)**

d. Perform Logistic Regression on train dataset. **(6)**

e. Predict the values on the Test dataset. **(6)**

f. Provide the value for Accuracy score. **(3)**

OR

B. **30 Marks**

Load the Fuel consumption.csv dataset. The dataset contains information about Fuel consumption and factors related to it. Predict the fuel consumption with all other factors given in the dataset using Linear regression

a. Read the dataset and split the data in X and Y.

i. Identify independent variables consider it X **(3)**

ii. Identify dependent variables consider it Y **(3)**

b. Perform feature engineering, fill all missing values and Perform label encoding or One hot encoding whenever required. **(6)**

c. Create a train-test split with 20% test ratio. **(3)**

d. Run the Linear Regression model. **(6)**

e. Predict values for test dataset. **(6)**

f. Provide the RMSE & R square value **(3)**