

PUSASQF405
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 calculator is allowed.

Q1. Attempt All questions.

A. Write a python program to find the average from given list of numbers.
using a function **5 Marks**
numbers = [23,25,28,29,30,46,55]

B. Write a program in python to calculate the interest earned by a customer.
from the bank after depositing a principal amount for 5 years. Take input.
from the customer for principal amount and the interest provided by the
the bank was at 6.75%. **5 Marks**
(Formula: principal_amount*intrest_rate*years)

C. Create a dataframe 2 dataframes, one for students and another for
academic details. Store them in df1 and df2. Split the data based on condition.
Result and merge the data based on name and on students whose result was P in data
post-split. **5 Marks**

student_details:

```
data = {  
    "Name": ["Alice", "Bob", "Charlie", "David"],  
    "Age": [30, 40, 25, 50],  
    "Gender": ["F", "M", "M", "M"],  
    "City": ["New York", "Los Angeles", "Chicago", "Houston"]  
}
```

academic_details:

```
data = {  
    "Name": ["Alice", "Charlie", "David", "Bob", "Alice", "David"],  
    "Marks": [75, 80, 55, 78, 40, 67],  
    "TransactionType": ["P", "P", "P", "P", "F", "P"]  
}
```

}

Q2. Attempt All questions.

- A.** Generate a python list of all palindromes between 10 to 1000

5

Marks

- B.** Write a program to print list containing the running multiplication of the below.
list

input_list = [22,22,33,45,90,100,101]

Hint: Running multiplication = current_item*previous_multiplication

5 Marks

- C.** Create three functions calculate_mean, calculate_variance and
calculate_standard_deviation

5 Marks

Formulas:

Average:

$\text{sum}(\text{numbers}) / \text{len}(\text{numbers})$

Variance:

$\text{mean} = \text{calculate_mean}(\text{numbers})$

$\text{squared_differences} = [(x - \text{mean})^2 \text{ for } x \text{ in numbers}]$

$\text{variance} = \text{sum}(\text{squared_differences}) / \text{len}(\text{numbers})$

Calculate Mean, Variance and Standard Deviation for the below list of numbers.

Numbers = [4,7,10,13,16]

Q3. Attempt question 3A or question 3B.

A. Load the built in diabetes dataset in python. Predict the progression of the disease using linear regression. 30 Marks

- a. To Read the built-in dataset use the below code:
from sklearn.datasets import load_diabetes

diabetes = load_diabetes() (7)

- b. Split diabetes data in X and Y such that, X consists of the data and Y consists of the target. (7)

- c. Perform Linear Regression (8)

- d. Provide the values for coefficient and intercept. (8)

OR

B. Load the train.csv dataset. The dataset contains a list of survivors in the Titanic, predict whether there are any chances of survival. Predict using Logistic Regression. 30 Marks

- a. Read the dataset. (2)

- b. Drop unnecessary columns Passenger Id, Name, Ticket, Cabin (4)

- c. Fill the missing values for age and embarked. Fill missing values of Age column with it's median value and for embarked fill it with its mode value (4)

- d. Convert the categorical columns Sex and embarked using one hot encoding. (5)

- e. Split the data into train and test with test size as 20% (5)

- f. Run the Logistic Regression on training model. (3)

- g. Predict the values for test dataset. (4)

- h. Find the accuracy of the model. (3)