

Programming with Python

Module-1 Graded Project-1

Learning Objective:

Part - I (10 Points)

- Variables, Data Structures
- Loops, Conditional statements
- Functions

Part - II (10 points)

- Numpy
- Pandas

Part - I (10 Points)

1. Answer the following questions. (2 points)

- a. Given the string 'great', use an appropriate index command that returns 'a' from the string. Given s = 'great'. (0.5 point)

Expected output: 'a'

- b. Write a program to find out the second smallest number in the list given below. (0.5 point)

li=[32, 3, 5, 8, 33, 7] & Expected output: 5

- c. Find out the unique values of the list given below. (0.5 point)

mylist=[1,1,5,3,3,6,11,11] & Expected output: [1,5,3,6,11]

- d. Write a program to print cube root of 216 using appropriate syntax. (0.5 point)

Expected output: 6

2. Replace the value 'hello' in the below nested list with 'bye'. (1 point)

a = [1, 2, [3, 4, 'hello']] & Expected output: [1, 2, [3, 4, 'bye']]

3. Given the below dictionary, write a program that returns ['object0']. (1 point)
dict = {'a':9, 'b':(2,3,1), 'c':['object0','object1','object2']}
Expected output: ['object0']
4. Write a program to calculate the area of a circle. The program should ask for an appropriate input and print the computed area. (2 points)
- a. Input text: Please enter the radius of the circle.
 - b. If the user enters a value of 5. The output should be: The area of the circle is 78.54
 - c. This program should indicate that the input is invalid if the user enters a character instead of a number as input. For instance if the user enters a value of 'a' instead of 5 in the previous example the program should prompt: Invalid input, please enter a number: _
5. Write a program which will take two lists X & Y as input and it will return a list having all the odd numbers present in the two lists. (Note:- common elements can be repeated) (2 points)
Input list: X=[1,2,3,4,5] & Y=[5,6,7,8,9]
Expected Output: [1,3,5,5,7,9] or [1,3,5,7,9] both are accepted.
6. Write a function that prints the integers from 1 to 15. But for multiples of three print "GREAT" instead of the number, and for the multiples of five print "LEARNING". For numbers which are multiples of both three and five print "GREATLEARNING". (2 points)

Part-II (10 Points)

Dataset: Online_sales.csv

Domain: Retail

Objective: Analyzing sales data to understand sales trend

1. Import necessary libraries and read the provided dataset (online_sales.csv) and check the top 5 and random 5 samples of the dataframe. (0.5 point)
2. Check info of the dataframe and write your observations. Comment on data types and shape of the dataset. (0.5 point)
3. Check for null values and report the percentage of null values of each column. And drop the rows having null values in it. (1 point)
4. Check the statistical summary of the dataset. (Use inbuilt functions) (0.5 point)
5. Drop the instances having quantity less than zero. (1 point) (Note:- refer to the 'Quantity' feature)
6. Check unique values of the country and report the name of the country that has the highest number of instances/rows. (0.5 point)
7. Create a new column with the name as 'sales' having total sales. The total sales is defined as Quantity*UnitPrice. (1 point)
8. Report the top 5 countries in terms of sales. (2 points)
 - a. Consider the size of sales.
 - b. Consider the mean value of sales.
9. Report the top 5 products which bring the highest sales. Use StockCode for product information. (1 point)
10. Convert the 'InvoiceDate' into a date format and report the month on which the maximum sales occur? (2 points)