PES University, Bengaluru (Established under Karnataka Act No. 16 of 2013)	**UE20CS901**

Aug 2023: END SEMESTER ASSESSMENT (ESA)
M TECH DATA SCIENCE AND MACHINE LEARNING_ SEMESTER I

UE20CS901 - Python for Data Science

Time: 3 Hrs	Answer All Questions	Max Marks: 100

INSTRUCTIONS

- All questions are compulsory.
- Section A should be handwritten in the answer script provided.
- Periodically save Jupyter notebook, after successful run of each cell.
- Section B and C are coding questions which have to be answered in the system.

SECTION A – 20 MARKS

1	a)	list down any 2 differences between list and tuple.	2
	b)	What is Seaborn's main function in data visualization, and how does it enhance the appearance of statistical plots?	2
	c)	What is the difference between append() and extend() methods?	2
	d)	What are *args and **kwargs in Python functions?	2
	e)	What are different approaches for typecasting in Python?	2

2	a)	How to merge dataframes in pandas?	2
	b)	Define significant features of Numpy library.	2
	c)	How to rank the rows of Dataframe?	2
	d)	What are map and reduce functions in Python?	2
	e)	Describe the characteristics and purpose of dictionaries in Python.	2

SECTION B – 40 MARKS

3	a)	i) You are working on a project to develop a simple online shopping application. You need to implement various data structures in Python to manage the shopping cart, product inventory, and user information efficiently. You have been given a list of products with their names and prices. Write a Python function named calculate_total_cost that takes a list of products and their corresponding quantities in the shopping cart and returns the total cost of the items.	5
		ii) You need to maintain an inventory of products in the online store. Create a Python dictionary named product_inventory to store the product names as keys and their corresponding quantities as values.	2
		iii) You have received a new shipment of products. Write a Python function named update_inventory that takes the current product_inventory and a dictionary representing the new shipment with product names and quantities. The function should update the inventory with the new shipment quantities.	6

		iv) You want to offer discounts on certain products. Write a Python function named apply_discount that takes the current product_inventory and a dictionary representing the discounted products with their new prices. The function should update the inventory with the discounted prices.	7
	b)	Create a scientific calculator using python. The calculator should give a menu to the user asking the type of mathematical operation to be performed (1. Arithmetic 2. Trigonometric 3. Logarithmic). According to the choice of the user, operations are performed on the numbers entered. Any innovation in this basic calculator model will be given more marks. You can add multiple operators/features. Refer jupyter notebook for sample input/output.	12
	c)	You have been given a string called sentence. Write a Python function named reverse_words that takes this string as input and returns a new string with the order of the words reversed.	3
	d)	You are a data analyst working on a top-secret project for a government agency. Your task is to analyze a series of encrypted messages. Each encrypted message is a combination of letters and digits, but you suspect that it might contain hidden information in the form of special alphanumeric codes. Write a Python function named find_special_codes that takes an input string containing the encrypted messages and returns a list of special codes. Special codes are defined as words or combinations of letters and digits that do not contain any special characters or symbols.	5

SECTION C – 40 MARKS

4	a)	Read the dataset(Ecommerce Purchases1.csv) and answer the below questions
		  I. What are the 5 most common Job Titles? (2 MARKS)
		  II. How many people have American Express as their Credit Card Provider and made a purchase above $95 ?(3 MARKS)
		  III. How many people have a credit card that expires in 2025? (3 MARKS)
		  IV. What are the top 5 most popular email providers/hosts (e.g. gmail.com, yahoo.com, etc...) . Support your answer with suitable graph. (7 MARKS)
		  V. Someone made a purchase that came from Lot: "90 WT" , what was the Purchase Price for this transaction?(3 MARKS)
		  VI. What is the email of the person with the following Credit Card Number: 4926535242672853(2 MARKS)
			20

4	b)	Read the dataset athelete_events.csv and regions.csv and answer the following questions. 1. Merge the dataset athletes and regions. (2 Marks) 2. Show the distribution of the age of gold medalists (6 Marks) 3. Display the number of athletes who are gold medalists and whose age is greater than 50 with their info. (6 Marks) 4. Women who can play in summer (2 Marks) 5. Display the top 5 countries who won the most medals .Support your answer with suitable plot.(4 Marks)	20