Programme	Computer Science & Engineering	Semester	V
Course	Artificial Intelligence & Machine Learning	Max Marks	30
Course Code	20CS51I	Duration	4 hours
Name of the course coordinator	Mrs. Nagaveni Kadakol	CIE	1

Note: Answer one full question from each section.

<table>
<tr><th>Q.No</th><th>Questions</th><th>CL L3/L4</th><th>CO</th><th>PO</th><th>Marks</th></tr>
<tr><td colspan="6">Section 1(Theory) 10 Marks</td></tr>
<tr><td>1.a)</td><td>What is the difference between Supervised and unsupervised machine learning? Explain with examples</td><td>L4</td><td>1</td><td>1</td><td>5</td></tr>
<tr><td>b)</td><td>Create two series as shown using pd.series() function.
Series_A = [1,2,3,4,5,6] Series_B = [4,5,6,7,8,9] Get the items not common to both.</td><td>L3</td><td>1</td><td>4</td><td>5</td></tr>
<tr><td>2.a)</td><td>How AI Software Development life cycle differs from traditional software development</td><td>L4</td><td>1</td><td>1,3,4</td><td>5</td></tr>
<tr><td>b)</td><td>Create a data frame with following data
<table>
<tr><th>Ename</th><th>Type</th><th>Department</th><th>Experience</th><th>Salary</th></tr>
<tr><td>Roshan</td><td>Regular</td><td>CS</td><td>10</td><td>50000</td></tr>
<tr><td>Amar</td><td>Adhoc</td><td>CS</td><td>20</td><td>15000</td></tr>
<tr><td>Ashwini</td><td>Regular</td><td>EC</td><td>5</td><td>30000</td></tr>
<tr><td>Lohith</td><td>Adhoc</td><td>EC</td><td>14</td><td>15000</td></tr>
<tr><td>Mohan</td><td>Contract</td><td>CS</td><td>9</td><td>10000</td></tr>
<tr><td>Pramod</td><td>Regular</td><td>EC</td><td>8</td><td>40000</td></tr>
</table>
1. Make a pivot table that shows the average salary of each employee for each department.
2. Make a pivot table that shows the sum and mean of the salaries of each type of employee and the number of employees of each type.</td><td>L3</td><td>1,5</td><td>2,3,4</td><td>5</td></tr>
<tr><td colspan="6">Section 2 (Practical) 20 Marks</td></tr>
<tr><td>3.a)</td><td>Consider the credit card dataset which contains the following columns:
• Create a bivariate plot to find if there is a correlation between credit card limit and average purchase made on the card.
• Visualise the distribution of values for credit card limit and average purchase made on the card. Also, identify the outliers in the data, if any.
• Provide a visual representation of the number of customers in each income group using a bar chart.
• Plot the frequency distribution of the total transaction amount.
• Graphically represent the percentage of customers retained and those attrited. Highlight the latter by slicing it apart from the main pie.</td><td>L3</td><td>2,5</td><td>2,3,4</td><td>15</td></tr>
<tr><td>b</td><td>1. Find a list of squares of all the numbers in a given list using lambda and map function.
2. Find the odd numbers from a given list using a filter
3. Compute a sum of the first five integers using reduce function.</td><td>L3</td><td>1,5</td><td>2,3,4</td><td>5</td></tr>
<tr><td>4.a)</td><td>Use the 'mtcars.csv' dataset to answer the above questions.
Create the following plots to visualize/summarize the data and customize it appropriately.
• Histogram to check the frequency distribution of the variable ‘mpg’ (Miles per gallon) and note down the interval having the highest frequency.</td><td>L3</td><td>2,5</td><td>2,3,4</td><td>15</td></tr>
</table>

	• scatter plot to determine the relationship between the weight of the car and the mpg • bar plot to check the frequency distribution of transmission type of cars. • Box plot of mpg and interpret the five-number summary. Create a git repository and push source code to the repo.				
b)	Consider the rainfall dataset. This data contains region(district) wise rainfall across India. Perform the following operations for the dataset 1. Find the district that gets the highest annual rainfall. 2. Drop the columns 'Jan-Feb', 'Mar-May', 'Jun-Sep', 'Oct-Dec'. 3. Display the state-wise mean rainfall for all the months using a pivot table.	L3	1,5	2,3,4	5

Course Coordinator **HOD** **IQAC Coordinator**