

SCHOOL OF COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE		DEPARTMENT OF COMPUTER SCIENCE ENGINEERING	
Program Name: B. Tech		Assignment Type: Lab	Academic Year: 2025-26
Course Coordinator Name		Dr.K.Deepthi	
Instructor(s) Name		Dr.J.Bhavana, Dr.M.Ranjeeth Kumar, Dr.N.Venkatesh, Dr.Sudersan Beheran, Dr.Hitesh Vijay Kumar P, Dr.B.Girirajan, Mr.D.Sravan Kumar, Mr.A.Vijay Kumar, Mr.K.Arunima, Ms.P.Nagalaxmi, Chandra Prakash, Mounika	
Course Code	24CS201PC210	Course Title	Data Analytics and Data Visualization
Year/Sem	II/I	Regulation	R24
Date and Day of Assignment	4-08-2025	Time(s)	09:00AM -05:00PM
Duration	2 Hours	Applicable to Batches	24CSBTB03, 24CSBTB22 24CSBTB42, 24CSBTB31 24CSBTB04, 24CSBTB37 24CSBTB27
Assignment Number: 02/12(Week2)-Monday			
Q. No.	Question		Expected Time to complete
1	<p>You are working as a data analyst for a large retail company that wants to analyse its sales and customer behaviour to optimize its marketing strategies. You are given three datasets:</p> <p>customers.csv: customer_id, name, gender, city 101, Alice, F, New York 102, Bob, M, Chicago 103, Clara, F, Boston 104, Dan, M, New York</p> <p>orders.csv: order_id, customer_id, order_date, amount 1001, 101, 2023-11-01, 250 1002, 102, 2023-11-01, 180 1003, 101, 2023-11-03, 75 1004, 104, 2023-11-05, 400</p> <p>products.csv: product_id, order_id, product_name, category P001, 1001, Laptop, Electronics P002, 1002, Mouse, Electronics</p>		

	<p>P003, 1003, Book, Stationery P004, 1004, Headphones, Electronics Perform the following data manipulations and answer the questions below:</p> <p>Merging & Joining: Merge the orders and customers datasets to get a complete view of customer orders. Join the resulting dataset with products using appropriate keys to get full order details including product name and category.</p> <p>Grouping: Group the merged dataset to find total sales (amount) by city. Group by customer name and product category to find how much each customer spent in each category.</p> <p>Reshaping: Pivot the grouped data to create a summary table with customer name as rows, product category as columns, and amount as values (fill missing with 0).</p> <p>Questions: Which city had the highest total sales? Which customer spent the most in the "Electronics" category? Provide a reshaped table that compares spending across product categories for each customer. How would the result change if one order contained multiple products? How would you handle such a scenario in your join?</p>	
--	--	--

SCHOOL OF COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE		DEPARTMENT OF COMPUTER SCIENCE ENGINEERING	
Program Name: B. Tech		Assignment Type: Lab	Academic Year: 2025-26
Course Coordinator Name		Dr. K. Deepthi	
Instructor(s) Name		Dr. J. Bhavana Dr. M. Ranjeeth Kumar Dr. N. Venkatesh Dr. Sudersan Beheran Dr.Hitesh Vijay Kumar P Dr.B.Girirajan Mr.D.Sravan Kumar Mr.A.Vijay Kumar Mr.K.Arunima, Ms.P.Nagalaxmi, Chandra Prakash, Mounika	
Course Code	24CS201PC210	Course Title	Data Analytics and Data Visualization
Year/Sem	II/I	Regulation	R24
Date and Day of Assignment	05-08-2025	Time(s)	09:00AM -05:00PM