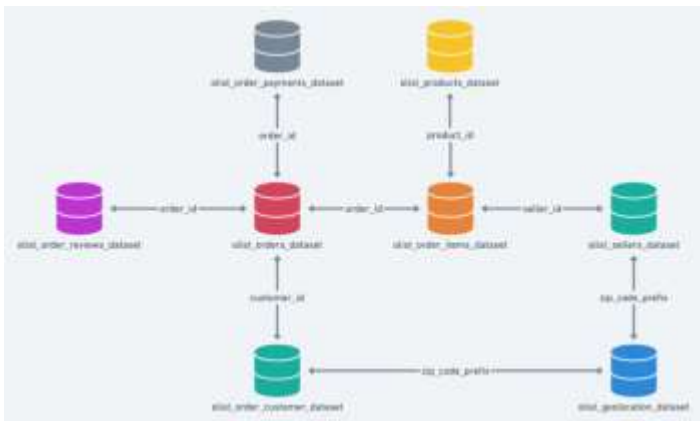


Semester: August 2021 – December 2021		
Examination: ESE Examination		
Programme code: 54 Programme: B.TECH	Class: SY	Semester: III (SVU 2020)
Name of the Constituent College: K. J. Somaiya College of Engineering	Name of the Department: COMP	
Course Code: 116h54C301	Name of the Course: Data Visualization	
Duration : 1 Hour 45 Minutes (15 minutes extra for uploading)	Maximum Marks : 50	
Instructions: 1)Draw neat diagrams 2) Assume suitable data if necessary 3) Apply All Visualization principle. 4) Apply visualization technique where ever required and give interpretation for each visualization along with screenshots of visualization.		

Question No.		Max Marks
Q1	<p>This dataset has information about the customer and its location. This is a Brazilian ecommerce public dataset of orders made at Olist Store. The dataset has information of 100k orders from 2016 to 2018 made at multiple marketplaces in Brazil. Its features allow viewing an order from multiple dimensions: from order status, price, payment and freight performance to customer location, product attributes and finally reviews written by customers. This is real commercial data, it has been anonymised, and references to the companies and partners in the review text have been replaced with the names of Game of Thrones great houses.</p> <p>Each order is assigned to a unique customerid. This means that the same customer will get different ids for different orders. The purpose of having a customer unique_id on the dataset is to allow you to identify customers that made repurchases at the store. Otherwise you would find that each order had a different customer associated with.</p> <p>Please refer to the data schema:</p>  <p>Link of dataset: https://drive.google.com/file/d/1abYUkHSvYyT5IK-P1PB-Vx_AujDrigBr/view </p>	

	Please answer the following on given dataset	
	A. Connect Brazilian ecommerce dataset and pre-process the dataset	07
	B. Verify the quality of given dataset using various visualization technique also provide its interpretation in words.	08
	C. Apply Visualization techniques and perform the Sales Prediction OR C. Apply Visualization techniques and perform the Delivery Performance	10
	D. Find out the Product Quality which is more prone to the customer's dissatisfaction.	05
Q2	The data.csv includes latest edition FIFA 2019 players attributes. Every player featuring in FIFA 18 70+ attributes Player and Flag Images Playing Position Data Attributes based on actual data of the latest EA's FIFA 18 game Attributes include on all player style statistics like Dribbling, Aggression, GK Skills etc. Player personal data like Nationality, Photo, Club, Age, Wage, Salary etc. Link of dataset: https://drive.google.com/file/d/1UbDyTRZRZmU7EAoev5SINApF9TOes7CM/view	
	A. Connect Above dataset and pre-process the dataset	05
	B. Verify the quality of given dataset using various visualization technique also provide its interpretation in words.	05
	C. Analyse which Club or National Team has the best-rated players OR C. Assess the strength of a team at a particular position.	04 04

	<p>D. Attempt Any TWO of the following</p> <ul style="list-style-type: none"> i. Co-relate between Age and Overall rating ii. Co-relate between Age and Nationality iii. Co-relate between Age and Potential 	06
--	--	----