Business Intelligence and Analytics Mid Term Exam.

Set A

Maximum Marks 25 Time: 60 mins

Instructions:

1. All questions are compulsory.
2. Please be very specific with your answers. Marks will be proportional to sense the answers make and not the length of the answers.

Question 1: Explain with an example any practical case where you have seen analytics being used for getting better outcomes. (3 Marks)

Question 2 : Take a look at the raw data captured in the table below. Create ETL rules so that all the data is cleansed and consistent. Explain why the rule was necessary. (6 Marks)

Roll No.	Name of the Student	Date of Birth	Address	Placement Status	Company Placed in	Salary Package
87	Akash Kore	2002/1/13		Placed	BCCI	12 Lakh
BT18CSE109	FADIA AAYUSH AMEET	7th February		Not Placed	Google	75,000 per month
CSE52	Diplesh	07/12/2002		No	JPMC	
BT18CSE128	Sagar M	02/10/2001		Yes		$100,000 p.a.
	MantriYash	30/02/2002		PPO		
003	Harshal F	9 Sep 02		Opted out		N.A.

Question 3: Emma owns a franchisee of Baskin Robins. She captures all the billing data such as Billing Date, Billing Time, Total Bill Amount, Loyalty Card Number, Item Code, Price per Item, Number of Items purchased, and Discount per Item. Using this data what kind of analyses can she perform? Specify 4 analyses and 3 **actionable business insights** she can derive from this data. (6 Marks)

Question 4: Take a look at the table and answer the questions:

Transaction ID	Purchased Items
001	{1,2,5,6}
002	{1,8,9}
003	{1,2,3,5}
004	{2,9}
005	{1,7,9}
006	{2,6,8}

007	{3,5,7}
008	{1,2,5,7,8)

a. How many items have support more than 50%
b. What is the confidence of the rule 1 -> 5 (if Item 1 purchased, then item 5 is also purchased)
c. Which association rule can you create based on this dataset? (6 Marks)

Question 5: State True or False with reasons/examples (4 Marks)

A. Data in a data warehouse is in most normalized form.
B. OLTP systems are optimized for storage whereas OLAP systems are optimized for querying.
C. In context of Market Basket Analysis, confidence and support can be more than 1.
D. Apriorialogorithmassumes the anti-monotonicity of frequency i.e. if an itemset is frequent all its subsets will not be frequent.