Vellore Institute of Technology

SCOPE

FDA (CSE1006) - Slot – L22+L23

DA -IV

1. Customer transactions dataset as follows.

Customer_ID	Transaction_Type	Amount
201	Deposit	2000
202	Withdrawal	1000
204	Deposit	3000
206	Transfer	500

Customer_ID	Name	Account_Balance
201	John	5000
202	Alice	12000
203	Robert	8000
204	Emma	15000
205	David	9500

Using merge function to combine both the datasets to perform

a) Merge with Left Join to retain all accounts
b) Fill missing transaction data
c) Add Customer_Status to merged data frame which relay on Account_Balance (if Account_Balance >10000 then Mark as “VIP” customers else “Regular).
d) Add Risk_Flag based on Customer_Status as VIP customers along with Transaction_Type “withdrawals” "High Attention" and “Normal”
e) Add Account_Flag accounts with no transactions & balance < 7000 as "Dormant" or “Active”.

2. Hospital maintains the patients’ data with following.

Patient_ID	Name	Age	Gender	Disease	Bill_Amount
101	Alice	25	F	Diabetes	500
102	Bob	30	M	Hypertension	1200
103		200	M	Asthma	300
104	David	45	M		-50
105	Eve	NA	F	Heart Disease	2000
106	Frank	33	M	Cancer	NA
107	Grace	27		Flu	400
108	<NA>	29	F	<NA>	750
109	Henry	150	M	Covid-19	1000000
110	Ivy	40	F	Stroke	950
101	Alice	25	F	Diabetes	500

a) Keep only unique PatientID rows
b) Replace Empty string with “unknown” and NA values set to mean.
c) Remove the outliers with (PatientAge > 120 | PatientAge < 0) replace with Median of PatientAge and (Bill > 50000 | Bill < 0) replace with Bill with Median.
d) Create column Age_Group with PatientAge < 30, "Young", or PatientAge < 60, "Middle-Aged" Patient_Age>60, "Senior".
e) Create a column to classify patients by hospital expense level: Bill < 500 ~ "Low", Bill >= 500 & Bill < 5000 ~ "Medium", Bill >= 5000 ~ "High.
f) Find only “High” billing patients above **40 years old**
g) Find patients with Diabetes or Hypertension paying over **$1000**
h) Group and summaries the data frame based on Bill category.

Note:

- ✓ Mention your name, roll number, and assessment subject details.
- ✓ Document should be in PDF/word with your registration number.
- ✓ In program execution shot 1st display your name and reg no followed by output in each screen shot.

****