

1. XYZ Bank has multiple customers. Each customer may hold multiple accounts. Each account may have multiple account holders (customers). How would you model this?

Customer_Account_Master_Table					
CustomerID	AccountNo	AccountType	AccountBAL	Currency	AccountOpenDate
1	100789001	SAVINGS	780.00	Pounds	1/8/2020
1	100789002	CURRENT	500.00	Pounds	1/8/2020
2	100789003	SAVINGS	123.00	Pounds	6/8/2020
2	100789004	CURRENT	54.00	Pounds	6/8/2020
2	100789005	SAVINGS_JOINT	99.00	Pounds	6/8/2020
7	100789008	CURRENT	2500.00	Pounds	6/9/2020
7	100789009	CURRENT_COLLECTION	890000.00	Pounds	6/10/2020
7	100789010	SAVINGS	210.00	Pounds	6/11/2020
8	100789011	SAVINGS	234.56	Pounds	6/12/2020
3	100789012	SAVINGS	259.12	Pounds	6/13/2020
4	100789013	SAVINGS	283.68	Pounds	6/14/2020
5	100789014	SAVINGS	308.24	Pounds	6/15/2020
6	100789015	SAVINGS	332.80	Pounds	6/16/2020

Account_Link_Sig_TABLE			
AccountNo	CustomerID	Designation	Signature
100789005	2	PARTNER	
100789005	6	PARTNER	
100789008	7	DIRECTOR	
100789008	1	DIRECTOR	
100789009	7	PARTNER	
100789010	7	PARTNER	
100789011	8	ONLY_SIGNATORY	
100789012	3	ONLY_SIGNATORY	
100789013	4	ONLY_SIGNATORY	
100789014	5	ONLY_SIGNATORY	
100789015	6	ONLY_SIGNATORY	

Customer_Table						
CustimerID	Custome Na	Address	PostCode	Province	Country	Status
1	John Smith	22 XYZ road	xx1 yyh	london	ENGLAND	Active
2	Victoria Ade	33 abc road	dg2 kkk	manchester	ENGLAND	Active
3	Jones Leich	34 abc road	gg4 1gh	birmigham	ENGLAND	Active
4	Willy Clinton	35 abc road	jjk 34g	new castle	ENGLAND	Active
5	Craig Doyle	36 abc road	888 vfg	southampton	ENGLAND	Active
6	Johnson Ade	33 abc road	dg2 kkk	manchester	ENGLAND	Active
7	Sally Superr	22 keneth ro	hh4 7aj	london	ENGLAND	Active
8	Peter James	37 abc road	hh4 7aj	london	ENGLAND	Active



## Data Model for the Above Conceptual schema

**Customer\_Table**

ColumnName	DataType
CustomerID	Integer
Customer Name	VARCHAR(30)
Address	VARCHAR(50)
PostCode	VARCHAR(10)
Province	VARCHAR(20)
Country	VARCHAR(20)
Status	VARCHAR(10)

Primary\_Key

**Customer\_Account\_Master\_Table**

ColumnName	DataType
CustomerID	Integer
AccountNo	varchar(10)
AccountType	varchar(20)
AccountBAL	NUMERIC(16,2)
Currency	varchar(10)
AccountOpenDate	DATE

Foreign\_Key

Primary\_Key

**Account\_Link\_Sig\_TABLE**

ColumnName	DataType
AccountNo	varchar(10)
CustomerID	Integer
Designation	varchar(20)
Memo_Text	Varchar(50)

Foreign\_Key

Foreign\_Key



2. Using your model, write a SQL query to return the total account balance for each customer where number of accounts is > 1

There are two scenario to this problem:

- a. Considering the retail account without the interlinked customer
- b. Considering the corporate account with more than one signatories: this implies partner/couples account, corporate account

***select***

***a.customerID***

***,a.name***

***,Sum(b.accounbalance)***

***,Count(b.AccountNo)***

***from Customer\_Table a,***

***Customer\_Account\_Master\_Table b,***

***Account\_Link\_Sig\_TABLE c***

***where a.customerID = b.customerID***

***and b.customerID = c.customerID***

***and b.AccountNo = c.AccountNo***

***group by a.customerId, a.name***

***having Count(b.AccountNo) >1***

3. Using your model, write a SQL query to return all the customers with no accounts

**SELECT Q.customerID , Q. CustomerName**

**FROM (select**

**a.customerID as customerID**

**,a.name as CustomerName**

**,Sum(b.accounbalance)**

**,Count(b.AccountNo)**

**from Customer\_Table a,**

**Customer\_Account\_Master\_Table b,**

**Account\_Link\_Sig\_TABLE c**

**where a.customerID = b.customerID**

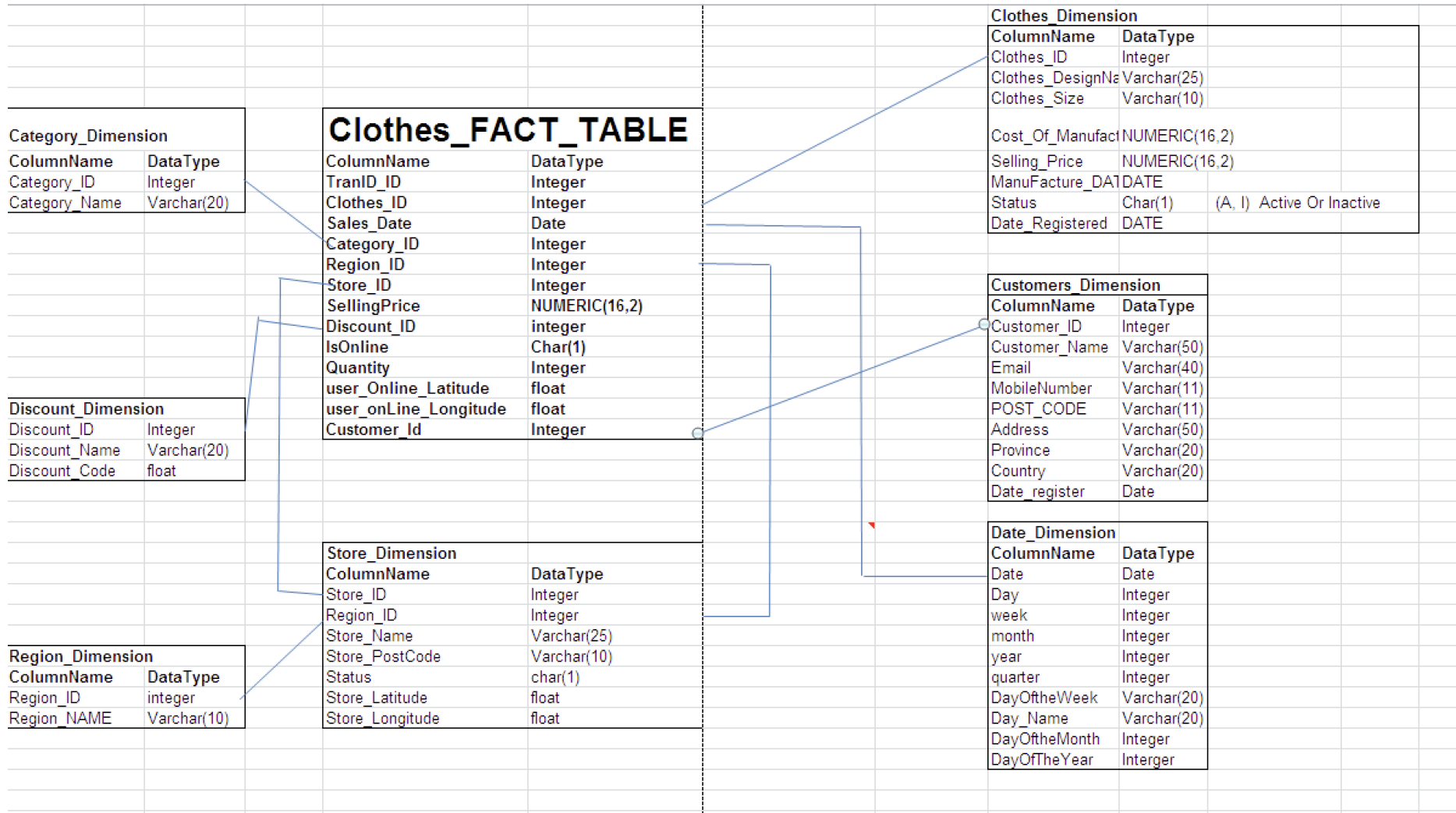
**and b.customerID = b.customerID**

**and b.AccountNo = c.AccountNo**

**group by a.customerId, a.name**

**having Count(b.AccountNo) < 1) Q**

4. ABC Ltd sells clothing via internet and in store. They want to determine the most profitable product categories across various regions. You need to take into account cost of product manufacture and any discounts applied to items sold. How would you model this (in a data warehouse)?



5. A data stage process was previously taking 10 minutes to run. Now it's taking an hour to complete. Detail the possible causes and how you would determine which was responsible
1. ***DataStage Jobs hang or have slow performance when temp directories have many files. The Temporary folder needs to be purged after backing it up.***
  2. ***The Infosphere Information Server log could be be set is not set to create another log after it reaches the optimal threshold say 1024 kilobytes. The antidote to the above to is to log onto the IIS Administrator and change the configuration to create new system.out log when it reaches 1024 KB.***
  3. ***The storage my might be full depending in the server infrastructure whether it is Converge Infrastructure or baremetal server. The VBlock needs to be check by the System Administrator and correct the server replication pattern/configuration***
  4. ***May be the OLTP system of records database log might be near above 80% capacity nearing 100%, this could make the underline SQL in the datastage job at specific stage to be slow. The Developer can peek into the job to determine the stage where the issues lies. The developer can implement parallel processing pattern after the Database administrator pruning the database log of the OLTP system of records.***
  5. ***Also if the IIS server is running on Websphere Application Server Network Deployment webserver. We would check the all the jvm of the WAS instance whether there are WAS heap dump issue. The log is analyzed using IBM HeapAnalyzer , Then we apply the recommendation of the heap analyser.***

