**Exercise – Sales Data**

**The attached dataset shows the bank’s customers in different segments, together with the latest account balance, how long they have been a customer, and which sales representative is servicing the customer. Using the sample data, please answer questions below:**

1. Possible Data anomalies?
2. Which Sales representative has better performance? R1/R2/ Data insufficient? Please explain
3. Write a program using SQL/Python/Pandas (any one) to identify (based on Balances):

* The most profitable customer for each representative
* The least profitable customer for each representative
* The second most profitable customer for each representative
* Repeat the above 3 for each representative in each Customer Segment

1. Write a program using SQL/Python/Pandas (any one) to identify the most profitable sector between Hospitals, Banks, Mining? What are some different ways of assessing this performance metric (KPI)?
2. If a new Banking customer is interested in opening an account with us, what is a reasonable account balance we can expect that customer to maintain? Please explain.
3. Write a python/SQL/pandas program to sort customers from highest balance to lowest :

* For each Sales representative
* For each Sales representative in the below CUSTOMER\_SINCE buckets:

1 – 5yrs; 6 – 10 yrs; 11- 20 yrs; Greater than 20 yrs