SQL CASE STUDY

DATA BANK

Neo-Banks are a recent development in the financial sector; they are new banks that solely operate online.

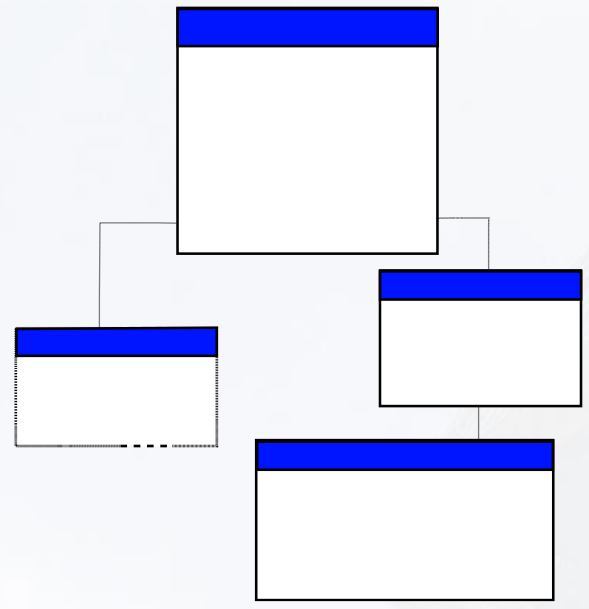
I believed that there should be some kind of connection between the digital world, these new age institutions, and cryptocurrencies.

So I made the decision to start a new project called Data Bank!

Customers of Data Bank receive cloud data storage allotments that are directly related to the balances in their accounts. The Data Bank team needs your assistance since this business model comes with some intriguing drawbacks.

This case study focuses on metrics calculations, business growth, and smart data analysis to assist the company more accurately estimate and plan for the future.

SCHEMA USED



regions	
region_id	int
region_name	varchar

customer_transactions	
customer_id	int
txn_date	date
txn_type	varchar
txn_amount	int

customer_nodes	
customer_id	int
region_id	int
node_id	int
start_date	date
end_date	date

CASE STUDY QUESTIONS

- How many different nodes make up the Data Bank network?
- How many nodes are there in each region?
- How many customers are divided among the regions?
- Determine the total amount of transactions for each region name.
- How long does it take on an average to move clients to a new node?
- What is the unique count and total amount for each transaction type?
- What is the average number and size of past deposits across all customers?
- For each month - how many Data Bank customers make more than 1 deposit and at least either 1 purchase or 1 withdrawal in a single month?

Q1. HOW MANY DIFFERENT NODES MAKE UP THE DATA BANK NETWORK?

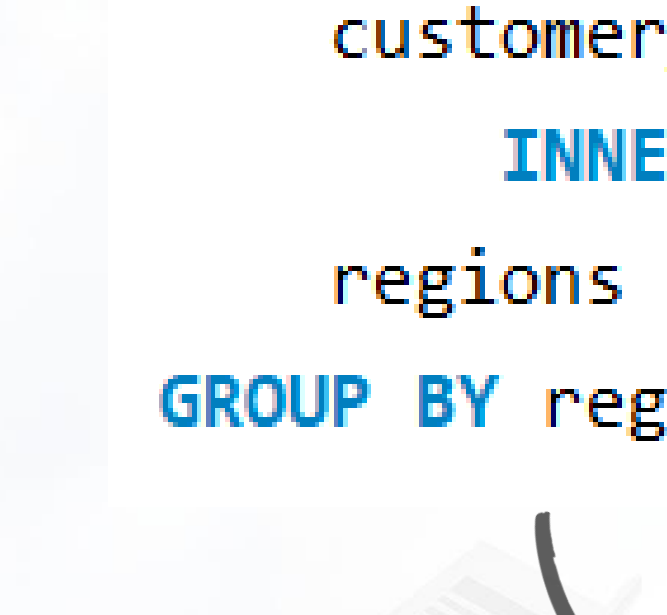
```sql
SELECT
    COUNT(DISTINCT node_id) AS unique_nodes
FROM
    customer_nodes;
```

Result Grid | Filter Rows:

	unique_nodes
▶	5

Q2. HOW MANY NODES ARE THERE IN EACH REGION?

```sql
SELECT
    region_id, COUNT(node_id) AS node_count
FROM
    customer_nodes
        INNER JOIN
    regions USING (region_id)
GROUP BY region_id;
```



	region_id	node_count
▶	1	770
	2	735
	3	714
	4	665
	5	616

Q3. HOW MANY CUSTOMERS ARE DIVIDED AMONG THE REGIONS?

```sql
SELECT
    region_id, COUNT(DISTINCT customer_id) AS customer_count
FROM
    customer_nodes
        INNER JOIN
    regions USING (region_id)
GROUP BY region_id;
```

region_id	customer_count
1	110
2	105
3	102
4	95
5	88

Q4. DETERMINE THE TOTAL AMOUNT OF TRANSACTIONS FOR EACH REGION NAME.

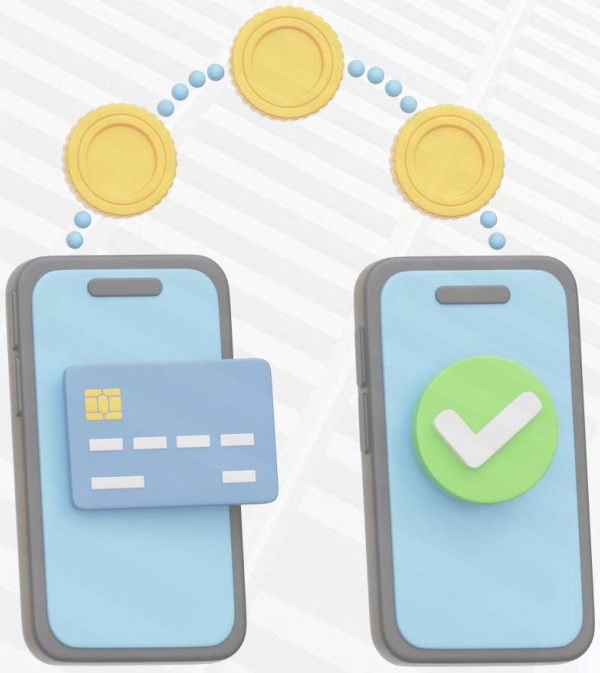
```sql
SELECT
    region_name, SUM(txn_amount) AS 'total transaction amount'
FROM
    regions,
    customer_nodes,
    customer_transactions
WHERE
    regions.region_id = customer_nodes.region_id
        AND customer_nodes.customer_id = customer_transactions.customer_id
GROUP BY region_name;
```

region_name	total transaction amount
Europe	3401552
Asia	4057879
Africa	4233481
Australia	4611768
America	4406276

Q5. HOW LONG DOES IT TAKE ON AN AVERAGE TO MOVE CLIENTS TO A NEW NODE?

```sql
SELECT
    ROUND(AVG(DATEDIFF(end_date, start_date)), 2) AS avg_days
FROM
    customer_nodes
WHERE
    end_date != '9999-12-31';
```

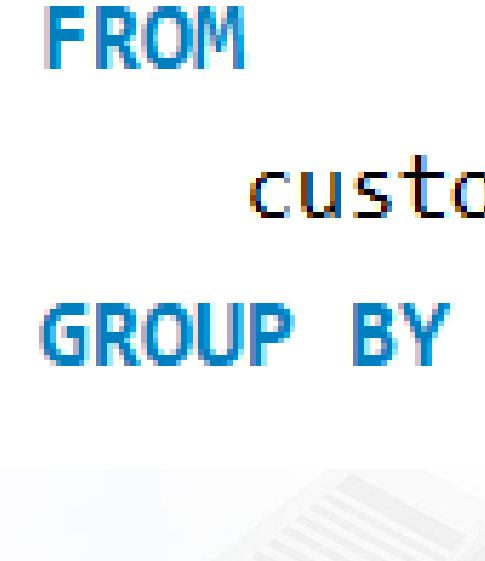
avg_days
14.63



Q6. WHAT IS THE UNIQUE COUNT AND TOTAL AMOUNT FOR EACH TRANSACTION TYPE?

```sql
SELECT
    txn_type,
    COUNT(*) AS unique_count,
    SUM(txn_amount) AS total_amount
FROM
    customer_transactions
GROUP BY txn_type;
```

	txn_type	unique_count	total_amont
▶	deposit	2671	1359168
	withdrawal	1580	793003
	purchase	1617	806537



Q7. WHAT IS THE AVERAGE NUMBER AND SIZE OF PAST DEPOSITS ACROSS ALL CUSTOMERS?

```sql
SELECT
    ROUND(COUNT(customer_id) / (SELECT
                    COUNT(DISTINCT customer_id)
                FROM
                    customer_transactions)) AS average_deposit_count,
    CONCAT('$', ROUND(AVG(txn_amount), 2)) AS average_deposit_amount
FROM
    customer_transactions
WHERE
    txn_type = 'deposit';
```

	average_deposit_count	average_deposit_amount
▶	5	$508.86

Q8. For each month - how many data bank customers make more than 1 deposit and at least either 1 purchase or 1 withdrawal in a single month?

```sql
WITH transaction_count_per_month_cte AS
  (SELECT customer_id,
          month(txn_date) AS txn_month,
          SUM(IF(txn_type="deposit", 1, 0)) AS deposit_count,
          SUM(IF(txn_type="withdrawal", 1, 0)) AS withdrawal_count,
          SUM(IF(txn_type="purchase", 1, 0)) AS purchase_count
   FROM customer_transactions
   GROUP BY customer_id,
            month(txn_date))
SELECT txn_month,
       count(DISTINCT customer_id) as customer_count
FROM transaction_count_per_month_cte
WHERE deposit_count>1
  AND (purchase_count = 1
       OR withdrawal_count = 1)
GROUP BY txn_month;
```

	txn_month	customer_count
▶	1	115
	2	108
	3	113
	4	50

Thank you!

Let's Connect

https://github.com/NikhilRoyDA

https://www.linkedin.com/in/nikhil-kumar-roy/

nikhilroy744@gmail.com