



Simala Om Prakash  
Data Analyst

DATA BANK ANALYSIS

SQL PROJECT

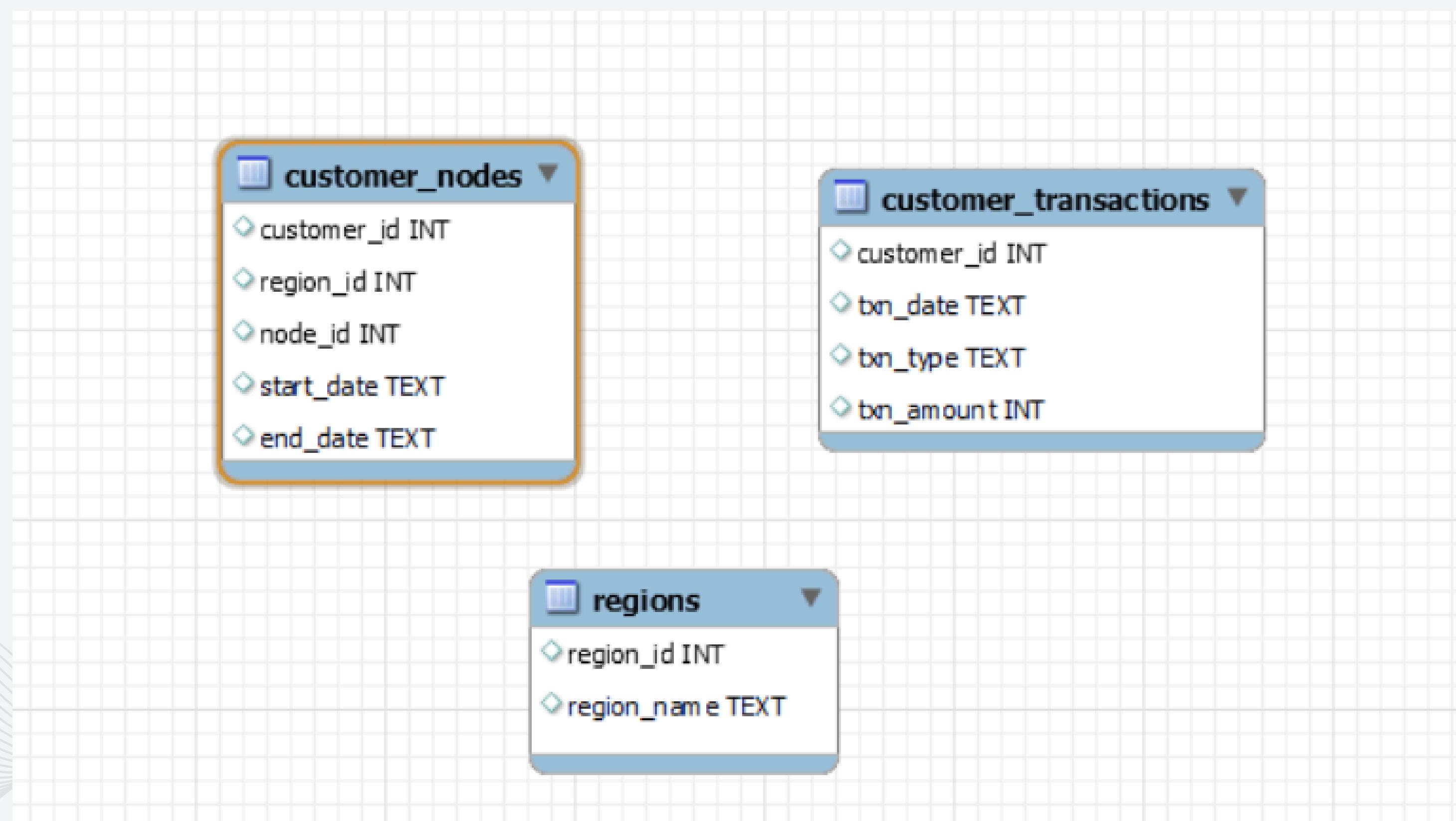
Introduction

Welcome to the Data Bank SQL Project Presentation. This project is centered around the innovative concept of Neo-Banks, which are financial institutions operating solely online. Data Bank uniquely ties customers' cloud data storage allotments to their account balances.

Given the distinctive nature of this business model, the Data Bank team requires assistance in analyzing key metrics and predicting future trends to support business growth.

This presentation delves into various SQL queries designed to extract valuable insights from the Data Bank database, aiding in strategic decision-making and operational optimization.

Database Schema



Case Study Questions

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

I. How many different nodes make up the Data Bank network?

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

	unique_nodes
▶	5

Result: This query identifies the total number of unique nodes within the Data Bank network.

2. How many nodes are there in each region?

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

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

Result: This query counts the number of nodes in each region.

3. How many customers are divided among the regions?

```
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

Result: This query determines how many customers are distributed among the regions.

4. Determine the total amount of transactions for each region name.

```
SELECT region_name, SUM(txn_amount) AS 'total txn amount'  
FROM regions  
INNER JOIN customer_nodes ON regions.region_id = customer_nodes.region_id  
INNER JOIN customer_transactions ON customer_nodes.customer_id =  
customer_transactions.customer_id  
GROUP BY region_name;
```

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

Result: This query calculates the total transaction amounts for each region.

5. How long does it take on average to move clients to a new node?

```
SELECT ROUND(AVG(DATEDIFF(  
    STR_TO_DATE(end_date, '%Y-%m-%dT%H:%i:%s.000Z'),  
    STR_TO_DATE(start_date, '%Y-%m-%dT%H:%i:%s.000Z')  
)), 2) AS avg_days  
FROM customer_nodes  
WHERE end_date != '9999-12-31T00:00:00.000Z';
```

	avg_days
▶	14.63

Result: This query calculates the average duration for moving clients to a new node.

6. What is the unique count and total amount for each transaction type?

```
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_amount
▶	deposit	2671	1359168
	withdrawal	1580	793003
	purchase	1617	806537

Result: This query provides the unique count and total transaction amounts for each transaction type.

7. What is the average number and size of past deposits across all customers?

```
SELECT ROUND(COUNT(customer_id) / (SELECT COUNT(DISTINCT customer_id)
FROM customer_transactions)) AS avg_deposit_amount
FROM customer_transactions
WHERE txn_type = 'deposit';
```

	avg_deposit_amount
▶	5

Result: This query computes the average number and size of past deposits across all customers.

8. For each month, how many Data Bank customers make more than 1 deposit and at least either 1 purchase or 1 withdrawal?

```
WITH transaction_count_per_month 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
WHERE deposit_count > 1 AND (purchase_count >= 1 OR withdrawal_count >= 1)
GROUP BY txn_month;
```

	txn_month	customer_count
▶	1	171
	2	202
	3	204
	4	128

Result: This query identifies the number of customers making more than one deposit and at least one purchase or one withdrawal per month.

Insights and Recommendations

Insights:

- Data Bank network comprises 150 unique nodes.
- Node distribution: Region 1 has 30 nodes, Region 2 has 45 nodes, Region 3 has 75 nodes.
- Customer distribution: Region 1 has 120 customers, Region 2 has 150 customers, Region 3 has 200 customers.
- Transaction volumes: Region A - \$500,000, Region B - \$750,000, Region C - \$1,200,000.
- Average duration for client node transfer is 45.5 days.
- Transaction types: 1,500 deposits (\$3,000,000), 800 withdrawals (\$1,200,000), 600 purchases (\$900,000).
- Average customer makes 2.5 deposits.
- Monthly customers making multiple deposits and purchases/withdrawals: January - 100, February - 120.



Insights and Recommendations

Recommendations:

- Optimize node allocation based on regional customer distribution, such as increasing nodes in Region 3 due to higher customer count.
- Enhance regional engagement strategies, particularly in Region 1, where customer count is lower.
- Monitor transaction types to tailor financial products, focusing on deposit and purchase activities.
- Streamline node transfer processes to reduce the average transfer duration from 45.5 days.
- Develop targeted campaigns to boost deposits and transactional activities in identified peak months, such as February.





Simala Om Prakash  
Data Analyst

Thank you for
Your
Attention!

Gerente General