

DATA BANK

BY ANUSH BHARATHWAJ L

customer_nodes					
customer_nodes.csv					
customer_id	region_id	node_id	start_date	end_date	
1	3	4	2020-01-02	2020-01-03	
2	3	5	2020-01-03	2020-01-17	
3	5	4	2020-01-27	2020-02-18	
4	5	4	2020-01-07	2020-01-19	
5	3	3	2020-01-15	2020-01-23	
6	1	1	2020-01-11	2020-02-06	
7	2	5	2020-01-20	2020-02-04	
8	1	2	2020-01-15	2020-01-28	
9	4	5	2020-01-21	2020-01-25	
10	3	4	2020-01-13	2020-01-14	
11	2	5	2020-01-19	2020-01-25	
DuckDB 3 ms (1 hour ago) 5 columns · 3,500 rows					

customer_transactions				
customer_transactions.csv				
customer_id	txn_date	txn_type	txn_amount	
429	2020-01-21	deposit	82	
155	2020-01-10	deposit	712	
398	2020-01-01	deposit	196	
255	2020-01-14	deposit	563	
185	2020-01-29	deposit	626	
309	2020-01-13	deposit	995	
312	2020-01-20	deposit	485	
376	2020-01-03	deposit	706	
188	2020-01-13	deposit	601	
138	2020-01-11	deposit	520	
373	2020-01-18	deposit	596	
DuckDB 1 ms (1 hour ago) 4 columns · 5,868 rows				

regions		
regions.csv		
region_id	region_name	
1	Australia	
2	America	
3	Africa	
4	Asia	
5	Europe	
DuckDB 1 ms (1 hour ago)		

A) How many different nodes make up the Data Bank network?

Unique Nodes Network

```
select
  count(distinct (node_id)) as Unique_Nodes
from
  customer_nodes;
```

Unique_Nodes
5

DuckDB

5 ms (29 minutes ago)

1 column · 1 row

B) How many nodes are there in each region?

Nodes By Region

```
select
  r.region_name,
  count(cn.node_id) as count_no
from
  customer_nodes cn
  inner join regions r on cn.region_id = r.region_id
group by
  r.region_name
order by
  r.region_name;
```

region_name	count_no
Africa	714
America	735
Asia	665
Australia	770
Europe	616

DuckDB

4 ms (28 minutes ago)

2 columns · 5 rows

C) How many customers are divided among the regions?

Customers By Region's

Select
 r.region_name,
 count(cn.customer_id) as count_no
from
 customer_nodes cn
 inner join regions r on cn.region_id = r.region_id
group by
 r.region_name
order by
 count_no desc;

region_name	count_no	
Australia	770	
America	735	
Africa	714	
Asia	665	
Europe	616	

DuckDB 2 ms (26 minutes ago) 2 columns · 5 rows

D) Determine the total amount of transactions for each region name?

Amount of Transactions by Each Region's

select
 r.region_name,
 sum(txn_amount) as amount
from
 regions r
 inner join customer_nodes cn on r.region_id = cn.region_id
 inner join customer_transactions ct on ct.customer_id = cn.customer_id
group by
 r.region_name
order by
 r.region_name;

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

DuckDB 11 ms (25 minutes ago) 2 columns · 5 rows

E) How long does it take on an average to move clients to a new node?

Average Client Migration Time to New Node

```
select
  round(avg(datediff('DAY',start_date, end_date)), 2) as Avg_Client
from
  customer_nodes
where
  end_date != '9999-12-31';
```

Avg_Client
14.63

DuckDB 8 ms (21 minutes ago) 1 column · 1 row

F) What is the unique count and total amount for each transaction type?

Transaction Summary by Type

```
select
  txn_type,
  count(txn_type) as No,
  sum(txn_amount) as Total_Transaction
from
  customer_transactions
group by
  txn_type;
```

txn_type	No	Total_Transaction
deposit	2671	1359168
withdrawal	1580	793003
purchase	1617	806537

DuckDB 2 ms (19 minutes ago) 3 columns · 3 rows

G) What is the average number and size of past deposits across all customers?

Deposit Statistics

```
select
  round(
    count(customer_id) / count(distinct (customer_id)),
    2
  ) as AVG_COUNT,
  concat('$', round(avg(txn_amount), 2)) as AVG_DEPOSIT
from
  customer_transactions
where
  txn_type = 'deposit';
```

AVG_COUNT	AVG_DEPOSIT	
5.34	\$508.86	

DuckDB 2 ms (17 minutes ago) 2 columns · 1 row

H) 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?

Monthly Active Customers with Multiple Deposits and Transactions

```
with
  trans_cte as (
    select
      monthname (txn_date) as Months,
      customer_id,
      sum(if (txn_type = 'deposit', 1, 0)) as Deposit,
      sum(if (txn_type = 'withdrawal', 1, 0)) as Withdrawal,
      sum(if (txn_type = 'purchase', 1, 0)) as Purchase
    from
      customer_transactions
    group by
      Months,
      customer_id
  )
select
  Months,
  count(customer_id) as NO
from
  trans_cte
where
  Deposit > 1
  and (
    Withdrawal = 1
    or Purchase = 1
  )
group by
  Months
order by
  Months;
```

Months	NO	
April	50	
February	108	
January	115	
March	113	

DuckDB 6 ms (13 minutes ago) 2 columns · 4 rows

Sources And Refrences

Query Dataset: <https://drive.google.com/drive/folders/10ELDsgZbwZQxtPdrgieqvC5fKApwWg2C>

CSV Dataset: <https://drive.google.com/drive/folders/1zcWPOK4OgsePIHpK95r023t1vQzaflgd?usp=sharing>

My Portfolio: <https://anushbharathwaj.w3spaces.com/>

Report: <https://drive.google.com/file/d/1RLhiBUw70xdJiWkBuEdOvqAhcqNrYotd/view?usp=sharing>