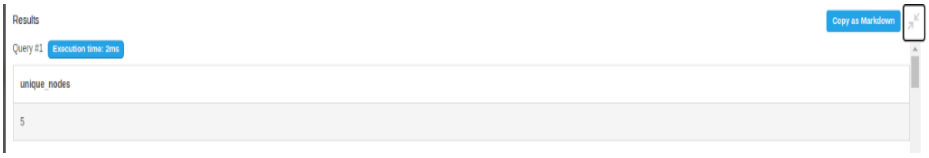
# Case Study-4

Thursday, February 3, 2022 4:59 PM

## Section A

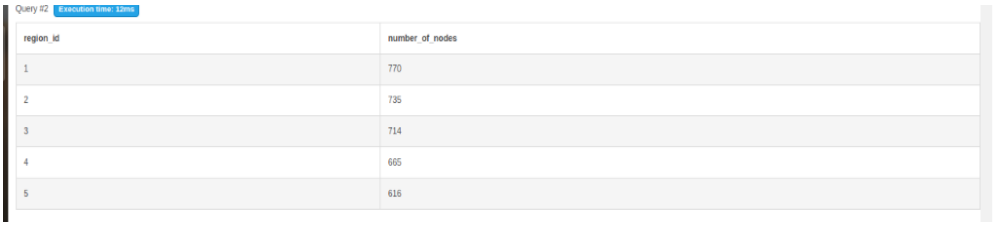
1. How many unique nodes are there on the Data Bank system?

```
select count(Distinct node_id) as unique_Nodes
    from data_bank.customer_nodes;
```




2. What is the number of nodes per region?

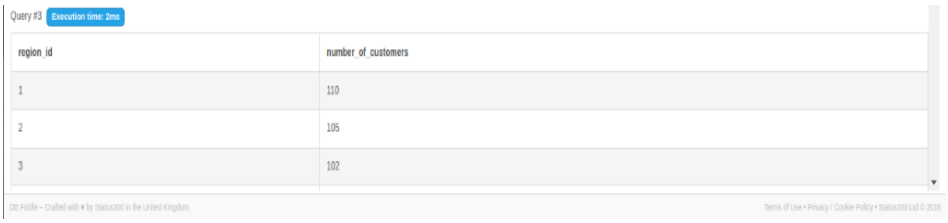
```
select Distinct(region_id),
    count(node_id) As Number_of_Nodes
    from data_bank.customer_nodes
    group by region_id
    order by region_id;
```




| region_id | number_of_nodes |
|---|---|
| 1 | 770 |
| 2 | 735 |
| 3 | 714 |
| 4 | 665 |
| 5 | 616 |

3. How many customers are allocated to each region?

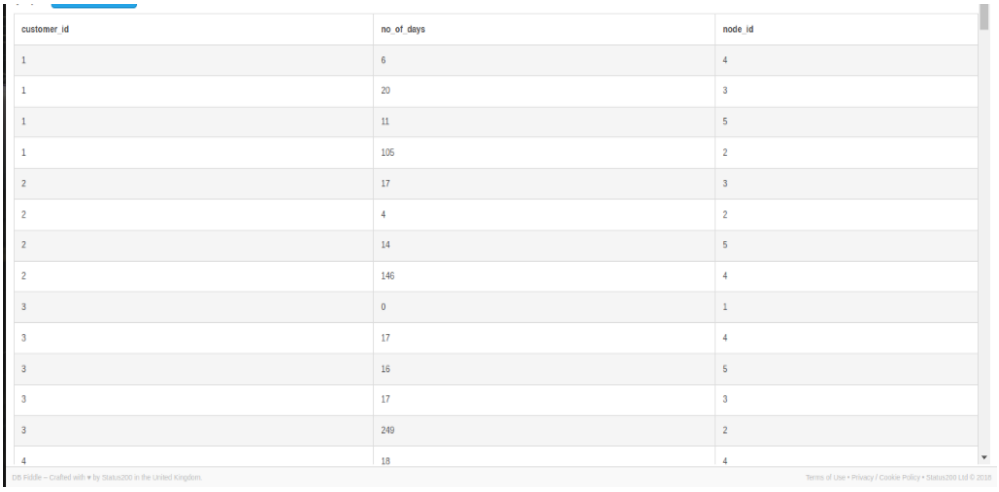
```
select Distinct(region_id),
    count(distinct customer_id) As Number_of_Customers
    from data_bank.customer_nodes
    group by region_id
    order by region_id;
```




| region_id | number_of_customers |
|---|---|
| 1 | 110 |
| 2 | 105 |
| 3 | 102 |

4. How many days on average are customers reallocated to a different node?

```
select customer_id,
    round(avg(((extract(doy from end_date)::int)-(extract(doy from start_date)::int))),0) as No_Of_Days,
    node_id from data_bank.customer_nodes
    group by customer_id,node_id
    order by customer_id;
```



| customer_id | no_of_days | node_id |
|---|---|---|
| 1 | 6 | 4 |
| 1 | 20 | 3 |
| 1 | 11 | 5 |
| 1 | 105 | 2 |
| 2 | 17 | 3 |
| 2 | 4 | 2 |
| 2 | 14 | 5 |
| 2 | 146 | 4 |
| 3 | 0 | 1 |
| 3 | 17 | 4 |
| 3 | 16 | 5 |
| 3 | 17 | 3 |
| 3 | 249 | 2 |
| 4 | 18 | 4 |

5. What is the median, 80th and 95th percentile for this same reallocation days metric for each region?

```
select
        region_id,
        PERCENTILE_CONT(0.5) WITHIN GROUP(ORDER BY days) as MEDIAN,
        PERCENTILE_CONT(0.80) WITHIN GROUP(ORDER BY days) as Eighty_Percentile,
        PERCENTILE_CONT(0.95) WITHIN GROUP(ORDER BY days) as NightyFive_Percentile
from
        (select customer_id,
                region_id,(extract(doy from end_date)::int-extract(doy from start_date)::int) As days from
                data_bank.customer_nodes
                where end_date != '9999-12-31'
                group by customer_id,region_id,days order by region_id) as day group by region_id;
```

| region_id | median | eighty_percentile | nightyfive_percentile |
|---|---|---|---|
| 1 | 15 | 23 | 28 |
| 2 | 15 | 23 | 28 |
| 3 | 15 | 24 | 28 |
| 4 | 15 | 23 | 28 |
| 5 | 15 | 24 | 28 |

## Section B

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

```
select count(distinct customer_id),
        txn_type,sum(txn_amount) As total_amount
        from data_bank.customer_transactions
        group by txn_type;
```

| count | txn_type | total_amount |
|---|---|---|
| 500 | deposit | 1359168 |
| 448 | purchase | 806537 |
| 439 | withdrawal | 793003 |

2. What is the average total historical deposit counts and amounts for all customers?

```
with CTE as
        (select distinct(customer_id) as customers,
                count(txn_type) as counts,
                txn_type,
                sum(txn_amount) as amount
                from data_bank.customer_transactions
                where txn_type='deposit'
                group by customer_id,txn_type)

    -- select * from CTE;
select customers,
        round(avg(counts),0) as avg_deposit_counts,
        amount from CTE
        group by   customers,amount order by customers;
```

| customers | avg_deposit_counts | amount |
|---|---|---|
| 1 | 2 | 636 |

| | | |
|---|---|---|
| 2 | 2 | 610 |
| 3 | 2 | 637 |
| 4 | 2 | 848 |
| 5 | 4 | 2910 |
| 6 | 9 | 4722 |
| 7 | 7 | 4588 |



3. For each month - how many Data Bank customers make more than 1 deposit and either 1 purchase or 1 withdrawal in a single month?

```
with Month_CTE As(
        select extract(month from txn_date) As month,
            to_char(txn_date,'month') as month_name,
            txn_type,count(txn_type) as counts,customer_id
            from data_bank.customer_transactions
            group by customer_id,txn_type,month,month_name
            order by customer_id)

select month,month_name,sum
    (case when (txn_type='deposit' and counts > 1) and (txn_type='purchase' and counts=1) or (txn_type='withdrawal'
    and counts=1) then 1 else 0 end) as Total_Customers
    from Month_CTE group by month,month_name
    order by month;
```

| month | month_name | total_customers |
|---|---|---|
| 1 | january | 123 |
| 2 | february | 163 |
| 3 | march | 161 |
| 4 | april | 105 |

4. What is the closing balance for each customer at the end of the month?

```
select Distinct(customer_id),
    date_trunc('month',txn_date) + interval '1 month' - interval '1 day' as end_month,
    to_char(txn_date,'month') as month_name,
    sum(txn_amount) As amount
    from data_bank.customer_transactions
    group by customer_id,end_month,month_name order by customer_id;
```

| customer_id | end_month | month_name | amount |
|---|---|---|---|
| 1 | 2020-01-31T00:00:00.000Z | january | 312 |
| 1 | 2020-03-31T00:00:00.000Z | march | 1600 |
| 2 | 2020-01-31T00:00:00.000Z | january | 549 |
| 2 | 2020-03-31T00:00:00.000Z | march | 61 |
| 3 | 2020-01-31T00:00:00.000Z | january | 144 |
| 3 | 2020-02-29T00:00:00.000Z | february | 965 |
| 3 | 2020-03-31T00:00:00.000Z | march | 401 |
| 3 | 2020-04-30T00:00:00.000Z | april | 493 |
| 4 | 2020-01-31T00:00:00.000Z | january | 848 |
| 4 | 2020-03-31T00:00:00.000Z | march | 193 |
| 5 | 2020-01-31T00:00:00.000Z | january | 2606 |
| 5 | 2020-03-31T00:00:00.000Z | march | 5137 |

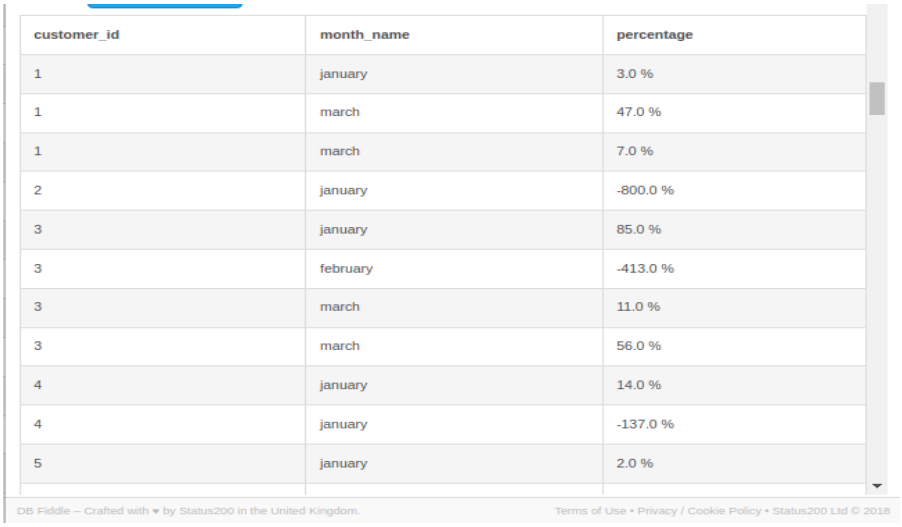


5. What is the percentage of customers who increase their closing balance by more than 5%?

```
With CTE as(
    select Distinct(customer_id),
        date_trunc('month',txn_date) end_month,
        to_char(txn_date,'month') as month_name,
```

```
        txn_amount As amount from
        data_bank.customer_transactions
        group by customer_id,end_month,month_name,amount
        order by customer_id)

select
    customer_id,
    month_name,
    ROUND((Next_row-amount)*100 /Next_row,1) || ' %' as  Percentage from (
        select *,
            LEAD(amount,1) OVER (
            PARTITION BY customer_id ORDER BY end_month) Next_row
            from CTE order by customer_id) as t where Next_row is not null;
```



| customer_id | month_name | percentage |
|---|---|---|
| 1 | january | 3.0 % |
| 1 | march | 47.0 % |
| 1 | march | 7.0 % |
| 2 | january | -800.0 % |
| 3 | january | 85.0 % |
| 3 | february | -413.0 % |
| 3 | march | 11.0 % |
| 3 | march | 56.0 % |
| 4 | january | 14.0 % |
| 4 | january | -137.0 % |
| 5 | january | 2.0 % |

DB Fiddle – Crafted with ♥ by Status200 in the United Kingdom. Terms of Use • Privacy / Cookie Policy • Status200 Ltd © 2018

```
with CTE AS(
 select date_trunc('month',txn_date) + interval '1 month' - interval '1 day' as end_month,sum(txn_amount) as amount from
data_bank.customer_transactions group by end_month),
 CTE2 AS(
   select date_trunc('month',txn_date) + interval '1 month' - interval '1 day' as end_month,
   txn_amount,LAG(txn_amount,1) OVER (order by end_month) previous_amount from CTE)
select end_month,
txn_amount,
      previous_amount
        (previous_amount-txn_amount) as variance
          from CTE2;


select Distinct(customer_id),date_trunc('month',txn_date) + interval '1 month' - interval '1 day' as
end_,txn_amount,LAG(txn_amount,2) OVER (
-- ORDER BY extract(month from txn_date)
-- ) previous_amount
-- -- sum(
-- --   CASE when txn_type='deposit' then txn_amount
-- --       ELSE -txn_amount
-- --   END)as txn
--    from data_bank.customer_transactions group by customer_id,end_,txn_amount order by customer_id;
```