Answering Questions in Data with SQL

Question 1

### Questions

[Deposits.csv](https://drive.google.com/file/d/1ml-1Iqae3ZTyWTKjUjV1M-uCVu8KqKL6/view?usp=sharing)

The table has the first 1000 rows of approx. 100k records of deposits made on the Bamboo platform by its users since inception. Given the following documentation, answer the questions below in the form of a query

* id: Transaction id
* amount: Money paid by the user in $
* status: Status of deposit (Note: 'Settlemented' refers to completed/accepted)
* user\_id: Id of user
* payment\_method\_id: Id of payment type (up to 7)
* inserted\_at: DateTime of the transaction first captured by the database
* updated\_at: DateTime of transaction last updated by database
* amount\_paid: If in Naira, the amount is converted in local currency
* fee: Transaction fee
* exchange\_rate: The exchange rate benchmarked to the dollar
* deposit\_type: Type of deposit (Standard or Instant)
* dollar\_instant\_deposit\_fee: Fee if deposit type in dollar
* dollar\_processing\_fee: Processing fee in dollar
* currency: USD or Naira

**Queries**

1. What is the median/average deposit value of those using payment method 3 in 2020?
2. When, in terms of which month of which year, did we observe the peak and valley (max and min) of deposit volume for Bamboo since inception?
3. Who are our power (most frequent) depositors by volume in 2021? Retrieve the user\_id and the total volume of those users
4. We define High Net Worth individuals as people depositing $10k or above. Retrieve the user\_id and the total deposits for users in that segment
5. What's the most used payment method among all users by volume?
6. What's the average transaction fee?

### Answers

1. Median deposit value of those using payment method 3 in 2020:

SET @row\_index := -1;

SELECT

    AVG(sub.amount\_paid) AS Median

FROM

    (SELECT

        @row\_index:=@row\_index + 1 AS row\_index, amount\_paid

    FROM

        deposits

    WHERE

        YEAR(inserted\_at) = 2020

            AND payment\_method\_id = 3

    ORDER BY amount\_paid) AS sub

WHERE

    sub.row\_index IN (FLOOR(@row\_index / 2) , CEIL(@row\_index / 2));

b) Average deposit value of those using payment method 3 in 2020:

SELECT

    AVG(amount\_paid) AS Average

FROM

    deposits

WHERE

    YEAR(inserted\_at) = 2020

        AND payment\_method\_id = 3;

1. Peak and Valley of deposit volume since inception:

SELECT

    MAX(amount\_paid) AS Peak\_Valley,

    MONTH(inserted\_at),

    YEAR(inserted\_at)

FROM

    deposits

UNION ALL SELECT

    MIN(amount\_paid) AS min,

    MONTH(inserted\_at),

    YEAR(inserted\_at)

FROM

    deposits;

1. Most Frequent depositors by volume in 2021:

select user\_id, Total

from(

SELECT DISTINCT

    (user\_id), sum(amount\_paid) AS Total, count(amount\_paid) as freq

FROM

    deposits

WHERE

    YEAR(inserted\_at) = 2021

GROUP BY user\_id

ORDER BY freq DESC) as subq;

1. High Net worth individuals:

SELECT

    user\_id, SUM(amount) AS Total

FROM

    deposits

GROUP BY user\_id

HAVING Total >= 10000

ORDER BY Total DESC;

1. Most used Payment method by volume:

SELECT

    payment\_method\_id

FROM

    (SELECT

        payment\_method\_id, COUNT(payment\_method\_id) AS frequency

    FROM

        deposits

    GROUP BY payment\_method\_id

    ORDER BY frequency DESC) AS sub

LIMIT 1;

1. Average Transaction Fee:

SELECT

    AVG(fee)

FROM

    deposits;

Question 2

### Questions

[Trades.csv](https://drive.google.com/file/d/12OYqGyN63OePRFE8P3ULyh-SGc51oJk3/view?usp=sharing)

The following table has the first 1000 rows of approx. 900k records of trades made on the Bamboo platform by its users since inception. Given the following documentation, answer the questions below in the form of a query

* id: Transaction Id
* price\_per\_share: Share price
* side: Trade side (BUY/SELL)
* stock\_symbol: Tick Symbol of Stock
* transaction\_value: Dollar amount for transaction
* type: Transaction type
* user\_id: id of user
* quantity: Number of shares
* inserted\_at: DateTime first captured by database
* updated\_at: DateTime last updated by database
* naira\_fee: 1.5% commission on the transaction value (in $) then converted in Naira
* dollar\_fee: 1.5% commission on the transaction value (in $)
* status: Trade's stage of approval (Note: Filled refers to accepted/completed)

**Queries**

1. Obtain the month-over-month change in absolute value and percentage for all completed (where status=Filled)Trades in 2020?
2. Retrieve the list of user\_id for those traded above 300x in 2020?
3. Find the best performing week, by total transaction value, from 2019 until today?
4. What are the top 5 bought stocks by volume?
5. What are the trade commissions for each side?
6. Which stock generated the most commissions from the BUY side?

### Answers

1. Month-over-month change for completed trades in 2020:

select abs((trades-lag(trades,1) over(order by month))/

lag(trades,1) over(order by month)) as value,

(100\*abs((trades-lag(trades,1) over(order by month))/

lag(trades,1) over(order by month))) as percent\_change

from(

select month(inserted\_at) as month,

sum(transaction\_value) as trades

from trade

where status='Filled' and month(inserted\_at)=2020

group by month) as sub;

1. Users who traded 300x in 2020:

SELECT

    user\_id

FROM

    (SELECT

        user\_id,

            100 \* MAX(transaction\_value) / MIN(transaction\_value) AS percent\_ratio,

            MAX(transaction\_value),

            MIN(transaction\_value)

    FROM

        trade

    WHERE

        MONTH(inserted\_at) = 2020

    GROUP BY user\_id

    HAVING percent\_ratio >= 300) AS subq;

1. Best Performing week from 2019 till date:

SELECT

    YEAR(inserted\_at) AS year,

    WEEK(inserted\_at) AS week,

    SUM(transaction\_value) AS transaction\_value

FROM

    trade

WHERE

    YEAR(inserted\_at) > 2018

GROUP BY year , week

ORDER BY transaction\_value DESC

LIMIT 1;

1. Top 5 bought stocks by volume

SELECT

    stock\_symbol

FROM

    (SELECT

        stock\_symbol, SUM(quantity) AS quant

    FROM

        trade

    GROUP BY stock\_symbol

    ORDER BY quant DESC) subq

LIMIT 5;

1. Trade commissions for each side:

SELECT

    side,

    SUM(naira\_fee) AS Com\_Naira,

    SUM(dollar\_fee) AS Com\_Dollar

FROM

    trade

GROUP BY side;

1. Stock with most commissions in BUY:

SELECT

    stock\_symbol

FROM

    trade

WHERE

    naira\_fee = (SELECT

            MAX(naira\_fee)

        FROM

            trade

        WHERE

            side = 'BUY');

Question 3

### Questions

* Count all users who deposited since bamboo began but haven't traded
* Count all users with deposits under $1k
* Count all users who had 4 or more trades before 2021 but have not traded in 2021
* Count all users who had 2 or more deposits before 2021 but have not deposited in 2021

### Answers

1. Deposited but never traded:

SELECT

    COUNT(DISTINCT (d.user\_id))

FROM

    deposits d

        INNER JOIN

    trade t ON d.user\_id = t.user\_id

WHERE

    t.status != 'FILLED';

1. Count users with deposits under $1k

SELECT

    COUNT(DISTINCT (user\_id)) AS Num\_of\_users

FROM

    deposits

WHERE

    amount < 1000;

1. Count all users who had 4 or more trades before 2021 but have not traded in 2021

SELECT

    COUNT(sub.user\_id)

FROM

    (SELECT DISTINCT

        (user\_id), status

    FROM

        trade

    WHERE

        status = 'FILLED'

            AND YEAR(inserted\_at) < 2021

    GROUP BY user\_id

    HAVING COUNT(status) >= 4) AS sub

        INNER JOIN

    (SELECT DISTINCT

        (user\_id), status

    FROM

        trade

    WHERE

        status != 'FILLED'

            AND YEAR(inserted\_at) = 2021) AS sub1 ON sub.user\_id = sub1.user\_id;

1. Count all users who had 2 or more deposits before 2021 but have not deposited in 2021:

SELECT

    COUNT(sub.user\_id)

FROM

    (SELECT DISTINCT

        (user\_id)

    FROM

        deposits

    WHERE

        amount\_paid IS NOT NULL

            AND YEAR(inserted\_at) < 2021

    GROUP BY user\_id

    HAVING COUNT(amount\_paid) >= 2) AS sub

        INNER JOIN

    (SELECT DISTINCT

        (user\_id)

    FROM

        deposits

    WHERE

        amount\_paid IS NOT NULL

            AND YEAR(inserted\_at) = 2021) AS sub1 ON sub.user\_id = sub1.user\_id;