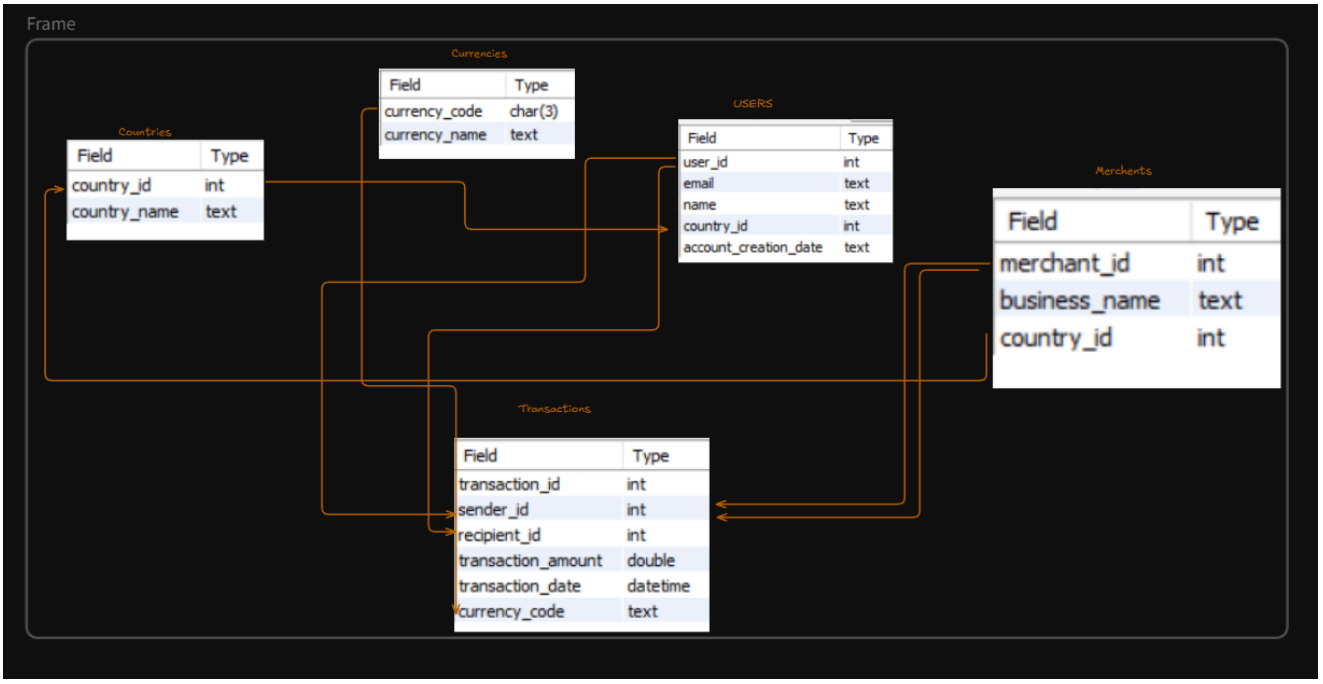
PayPal SQL Project



Problem statement

As a financial analyst at PayPal, you are tasked with analyzing transaction data to identify key markets.

```
Determine the top 5 countries by total transaction amount for both sending
and receiving funds in the last quarter of 2023 (October to December 2023).
Provide separate lists for the countries that sent the most funds and those
that received the most funds. Additionally, round the total_sent and
total_received amounts to 2 decimal places.
```

```
Select
    c.country_name,
    Round(Sum(t.transaction_amount),2) AS total_sent
from countries c
join users u
on c.country_id = u.country_id
join transactions t
on t.sender_id  = u.user_id
where t.transaction_date between '2023-10-01' AND '2024-01-01'
group by c.country_name
Order by total_sent desc
Limit 5;
```

Answer: Iceland , Zambia , Israel , Saudi Arabia , Togo has made Most of transaction in last quarter of 2023,

```
Select
    c.country_name,
    Round(Sum(t.transaction_amount),2) AS total_recevied
from countries c
join users u
on c.country_id = u.country_id
join transactions t
on t.recipient_id  = u.user_id
where t.transaction_date between '2023-10-01' AND '2024-01-01'
group by c.country_name
Order by total_recevied desc;
```

Answer: Zambia , Iceland , Saint Kitts Nevis, Timor-Leste And Guadeloupe has received most in last quarter though we can also see Israel ,Saudi and Togo who was the top sender was not in top 5 receivers.

Problem statement 2

To effectively manage risk, it's crucial to identify and monitor high-value transactions.

Find transactions exceeding $10,000 in the year 2023 and include transaction ID, sender ID, recipient ID (if available), transaction amount, and currency used.

```
Select
    transaction_id,
    Sender_id,
    recipient_id,
    transaction_amount,
    currency_code
from transactions
Where Year(transaction_date) = 2023
AND Transaction_amount > '10000';
```

Answer: We have only one transaction that is more than $10,000 and it stands around $200,000.

Problem statement 3

The sales team is interested in identifying the top-performing merchants based on the number of payments received. The analysis will help the sales team to better understand the performance of these key merchants during the specified timeframe.

Your task is to analyze the transaction data and determine the top 10 merchants, sorted by the total transaction amount they received, within the period from November 2023 to April 2024. For each of these top 10 merchants, provide the following details: merchant ID, business name, the total transaction amount received, and the average transaction amount

```
Select
    m.merchant_id,
    m.business_name,
    round(Sum(t.transaction_amount),2) AS total_recevied,
    round(avg(t.transaction_amount),2) AS avg_transaction
from transactions t
join merchants m
on t.recipient_id = m.merchant_id
where t.transaction_date between '2023-11-01' AND '2024-04-30'
group by m.merchant_id, m.business_name
Order by total_recevied desc
limit 10;
```

Answer: Most of amount received between Nov and Apr is by Gould LLC with Amount 36,380 and avg transactions stands out to be 9095.

Problem statement 3

The finance team wants to analyze the company's exposure to currency risks.

Analyze currency conversion trends from 22 May 2023 to 22 May 2024. Calculate the total amount converted from each source currency to the top 3 most popular destination currencies.

```
Select
    c.currency_code,
    Sum(t.transaction_amount) AS total_converted
from currencies c
join transactions t
on c.currency_code = t.currency_code
Where t.transaction_date between '2023-05-22' AND '2024-05-22'
Group by c.currency_code
order by total_converted desc
limit 3;
```

Answer: Euro, Japanese yuan, US Doller are top 3 currencies converted from may 2023 to may -2024

Problem statement 4

The finance team is evaluating transaction classifications.

Categorize transactions as 'High Value' (above $10,000) or 'Regular' (less than or equal to $10,000) and calculate the total amount for each category for the year 2023.

```
Select
    Case
```

```
        When transaction_amount > 10000 Then 'High Value'
        Else 'Regular' END AS Transaction_category,
    Round(Sum(transaction_amount),2) AS total_amount
from transactions
group by Case
        When transaction_amount > 10000 Then 'High Value'
        Else 'Regular' END;
```

Answer: Here Regular has most of transactions.

Problem statement 5

To meet compliance requirements, the finance team needs to identify the nature of transactions conducted by the company. Specifically, you are required to analyze transaction data for the first quarter of 2024 (January to March).

Your task is to create a new column in the dataset that indicates whether each transaction is international (where the sender and recipient are from different countries) or domestic (where the sender and recipient are from the same country). Additionally, provide a count of the number of international and domestic transactions for this period.

This classification will assist in ensuring compliance with relevant regulations and provide insights into the distribution of transaction types. Please include a detailed summary of the counts for each type of transaction.

```
Select
    Case
        When u.country_id <> u2.country_id Then 'Internataionnal'
        Else 'Domestic' END 'transaction_type',
    Count(*) as transaction_count
from users u
join transactions t
on u.user_id = t.sender_id
join users u2
on t.recipient_id = u2.user_id
Where t.transaction_date Between '2024-01-01' AND '2024-04-01'
Group by Case
        When u.country_id <> u2.country_id Then 'Internataionnal'
        Else 'Domestic' END;
```

Answer: Most of the transaction are International in 1st Q1 of 2024

Problem statement 6

To improve user segmentation, the finance team needs to analyze user transaction behavior.

Your task is to calculate the average transaction amount per user (Round up to 2 Decimal Places) for the past six months, covering the period from November 2023 to April 2024.

Once you have the average transaction amount for each user, identify and list the users whose average transaction amount exceeds $5,000.
This analysis will help the finance team to better understand high-value users and tailor strategies to meet their needs.

```
Select
    u.user_id,
    u.email,
    Round(avg(t.transaction_amount),2) AS Avg_Amount
from users u
join transactions t
on u.user_id = t.sender_id
Where t.transaction_date between '2023-11-01' AND '2024-05-01'
Group by u.user_id, u.email
having avg(t.transaction_amount) > 5000
order by  user_id asc;
```

Problem statement 7

As part of identifying top customers for a new loyalty program, the finance team needs to find the most valuable customer over the past year. Specifically, your task is to determine the user who has the highest total transaction amount from May 22, 2023, to May 22, 2024.

Please provide the details of this user, including their user ID, name, and total transaction amount. This information will help the finance team to select the most deserving customer for the loyalty program based on their transaction behavior over the specified period.

```
Select
    u.user_id,
    u.email,
    u.name,
 Round(sum(t.transaction_amount),2) AS total_amount
from users u
join transactions t
on u.user_id = t.sender_id
Where t.transaction_date between '2023-05-22' AND '2024-05-22'
Group by u.user_id , u.email, u.name
order by total_Amount desc
limit 1;
```

Problem statement 8

As part of a financial analysis, the team needs to categorize transactions based on multiple criteria. Create a report that categorizes transactions into 'High Value International', 'High Value Domestic', 'Regular International', and 'Regular Domestic' based on the following criteria:

-High Value: transaction amount > $10,000 International: sender and recipient from different countries

```
Select
    Case
        When t.transaction_amount > 10000 AND  u.country_id <> ur.country_id
then 'High Value International'
        When t.transaction_Amount < 10000 And u.country_id <> ur.country_id
then 'Regular International'
        When t.transaction_amount > 10000 AND u.country_id = ur.country_id
then 'High Value Domestic'
        When t.transaction_amount < 10000 AND u.country_id = ur.country_id
then 'Regular Domestic'
        END AS 'transaction_category',
    Count(*) as Transaction_count
from users u
join transactions t
on t.sender_id = u.user_id
join users ur
on t.recipient_id = ur.user_id
Where extract(year from t.transaction_date) = 2023
Group by
1;
```

Answer: In year 2023 Regular International has more transactions and Regular domestic and High value transaction is only 1.

Problem statement 9

The finance department requires a comprehensive monthly report for the year 2023 that segments transactions by type and nature. Specifically, the report should classify transactions into 'High Value' (above $10,000) and 'Regular' (below $10,000), and further differentiate them as either 'International' (sender and recipient from different countries) or 'Domestic' (sender and recipient from the same country).
Your task is to write a query that groups transactions by year, month, value_category, location_category, and then calculates both the total transaction amount and the average transaction amount for each group. This detailed analysis will provide valuable insights into transaction patterns and help the finance department in their review and planning processes.

```
Select
    Year(t.transaction_date) AS transaction_year,
    Month(t.transaction_date) AS transaction_month,
    Case When t.transaction_amount > 10000 Then 'High Value' ELSE 'Regular'
END AS value_category,
    Case when  u.country_id <> ur.country_id then 'International'
        When u.country_id = ur.country_id then 'Domestic' END
location_category,
```

```
    Round(Sum(t.transaction_amount),2) as total_amount,
    Round(avg(t.transaction_amount),2) as average_amount
from users u
join transactions t
on u.user_id = t.sender_id
join users ur
on t.recipient_id = ur.user_id
where extract(year from t.transaction_date) = 2023
Group by 1,2,3,4
order by 1,2,3,4;
```

Problem statement 10

The sales team wants to evaluate the performance of merchants by creating a score based on their transaction amounts. The score is calculated as follows:

- If total transactions exceed $50,000, the score is 'Excellent'
- If total transactions are greater than $20,000 and lesser than or equal to $50,000, the score is 'Good'
- If total transactions are greater than $10,000 and lesser than or equal to $20,000, the score is 'Average'
- If total transactions are lesser than or equal to $10,000, the score is 'Below Average' Write a query to assign a performance score to each merchant and calculate the average transaction amount for each performance category for the period from November 2023 to April 2024.

```
Select
    m.merchant_id,
    m.business_name,
    Round(Sum(t.transaction_amount),2) AS total_received,
    Case
        When Sum(t.transaction_amount) > 50000 Then 'Excellent'
        When Sum(t.transaction_amount) Between 20000 AND 50000 Then 'Good'
        When Sum(t.transaction_amount) Between 10000 AND 20000 Then
'Average'
        When sum(t.transaction_amount) <= 10000 Then 'Below Average'
        END performance_score,
    Round(avg(t.transaction_amount),2) AS average_transaction
from merchants m
join transactions t
on m.merchant_id = t.recipient_id
where t.transaction_date between '2023-11-01' AND '2024-05-01'
group by 1,2
Order by total_received desc;
```

Problem statement 11

The marketing team wants to identify users who have been consistently engaged over the last year (from May 2023 to April 2024). A consistently engaged user is defined as one who has made at least one transaction in at least 6 out of the 12 months during this period.

Write a query to list user IDs and their email addresses for users who have made at least one transaction in at least 6 out of 12 months from May 2023 to April 2024.

```
Select
    u.user_id,
    u.email
from users u
join (
    Select
        t.sender_id as user_id,
        count(distinct date_format(t.transaction_date ,'%Y-%m')) as
Active_month
        from transactions t
        where t.transaction_date between '2023-05-01' AND '2024-05-01'
        Group by t.sender_id) m
on u.user_id = m.user_id
Where m.active_month >= 6
order by u.user_id;
```

Problem statement 12

The marketing team wants to identify users who have been consistently engaged over the last year (from May 2023 to April 2024). A consistently engaged user is defined as one who has made at least one transaction in at least 6 out of the 12 months during this period.

Write a query to list user IDs and their email addresses for users who have made at least one transaction in at least 6 out of 12 months from May 2023 to April 2024.

```
Select
    u.user_id,
    u.email
from users u
join (
    Select
        t.sender_id as user_id,
        count(distinct date_format(t.transaction_date ,'%Y-%m')) as
Active_month
        from transactions t
        where t.transaction_date between '2023-05-01' AND '2024-05-01'
        Group by t.sender_id) m
on u.user_id = m.user_id
Where m.active_month >= 6
order by u.user_id;
```

Problem statement 13

The sales team wants to analyze the performance of each merchant by tracking their monthly total transaction amounts and identifying months where their transactions exceeded $50,000.

Write a query that calculates the total transaction amount for each merchant by month, and then create a column to indicate whether the merchant exceeded $50,000 in that month. The transaction date range should be considered from 1st Nov 2023 to 1st May 2024 (inclusive). The new column should contain the values 'Exceeded $50,000' or 'Did Not Exceed $50,000'. Display the merchant ID, business name, transaction year, transaction month, total transaction amount, and the new column indicating performance status.

```
Select
    m.merchant_id,
    m.business_name,
    Year(t.transaction_date) AS transaction_year,
    Month(t.transaction_date) AS transaction_month,
    Round(sum(t.transaction_amount),2) as Total_Amount,
Case When sum(t.transaction_amount) > 50000 Then 'Exceeded %50,000' Else
'Did Not Exceed $50,000' END AS transaction_Status
FROM transactions t
join merchants m
on t.sender_id = m.merchant_id
where t.transaction_Date between '2023-11-01' AND '2024-05-30'
Group by 1,2,3,4
order by 1,3,4;
```