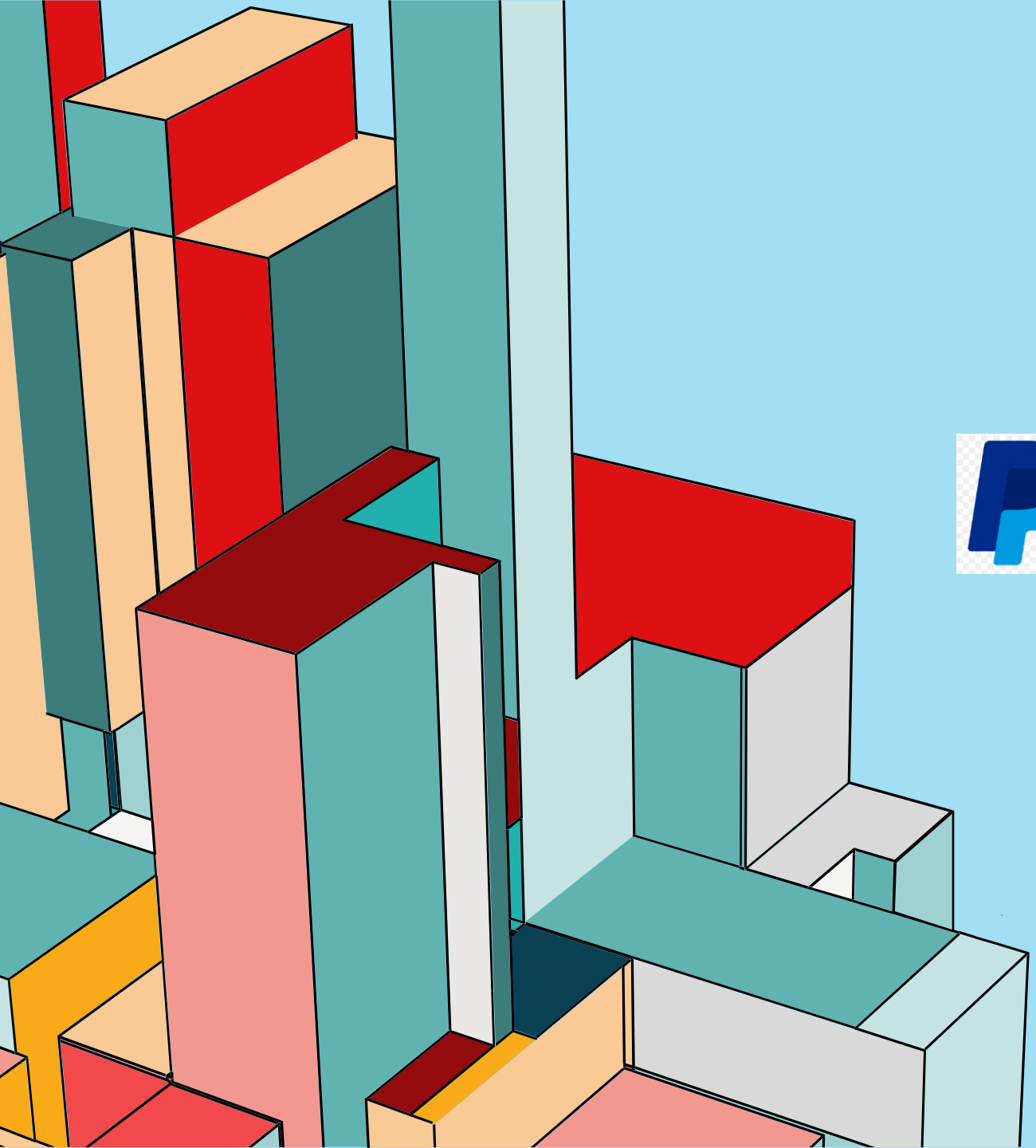
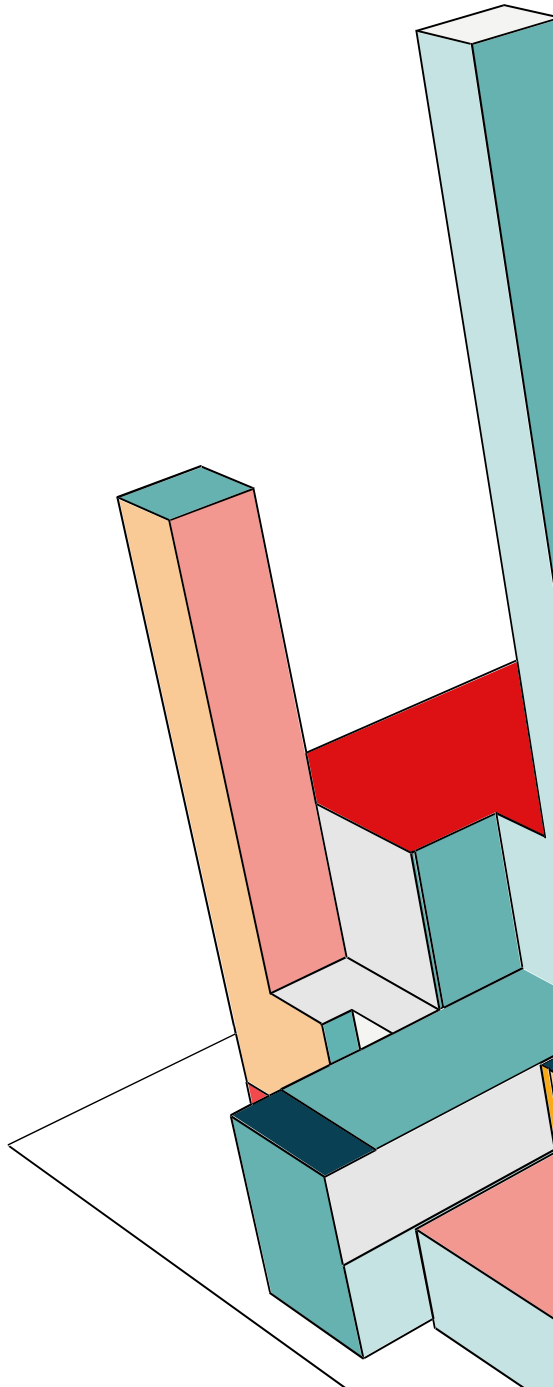


PAYPAL TRANSACTIONS ANALYSIS PROJECT

- BY SURABHI GOYAL

AGENDA

- Introduction
- Overview
- Objectives
- Dataset Details
- Data Analysis In MySQL
- Business Insights



INTRODUCING

SURABHI GOYAL

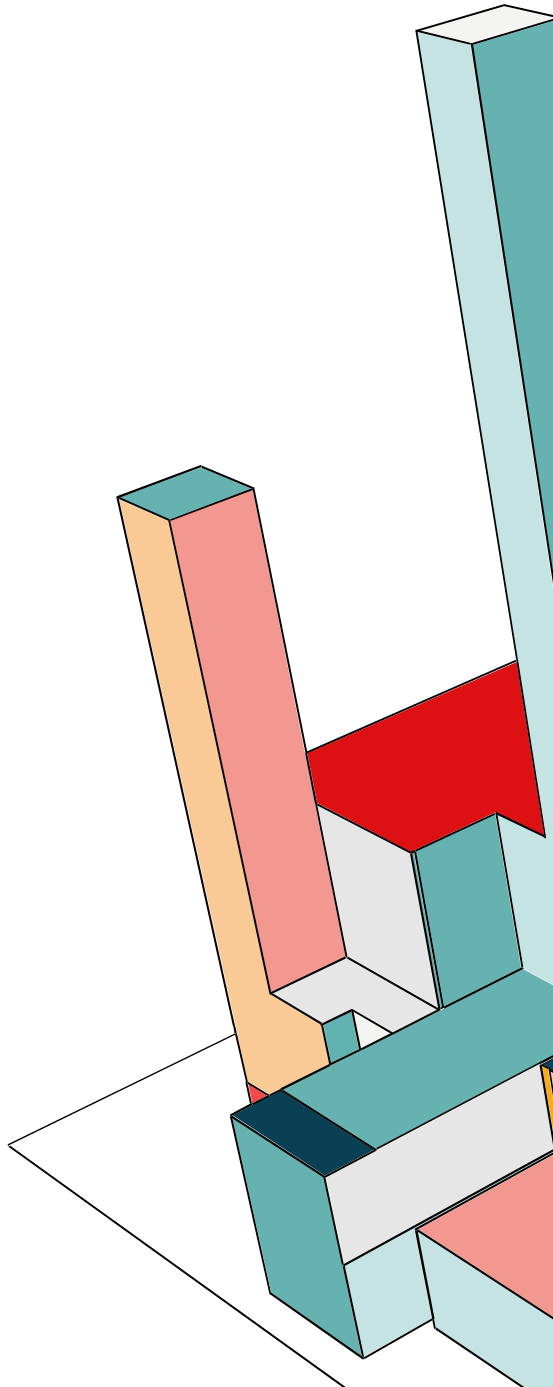
Result-oriented Data Analyst with 2+ years of experience and a proven track record of extracting valuable insights from complex datasets. Proficient in SQL, Python, and data visualization tools, I utilize statistical analysis and data mining techniques to inform and drive strategic business decisions.

OVERVIEW

The project involves leveraging extensive databases within a dynamic global digital payment platform PayPal to increase transaction volume and improve user satisfaction. The goal is to identify actionable insights from the transaction data that can drive strategic decisions, improve risk management, and ultimately increase the platform's market share and profitability in the competitive digital payment industry.

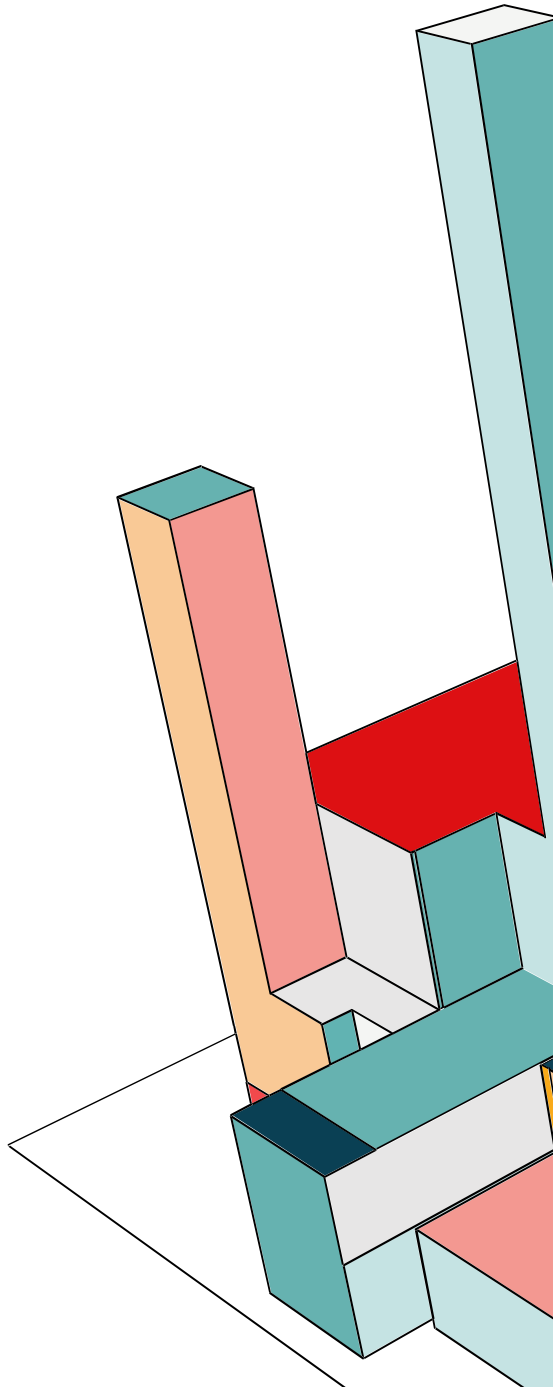
OBJECTIVES

- User Behavior and Engagement: Identify key user segments and analyze transaction patterns to enhance customer retention and value.
- Geographic Insights: Analyze transaction distribution to uncover high-performing regions and growth opportunities.
- Sales Optimization: Evaluate sales trends and customer preferences to boost platform adoption and revenue.
- Platform Performance: Monitor KPIs like transaction volume, revenue, and user growth to ensure optimal platform performance.

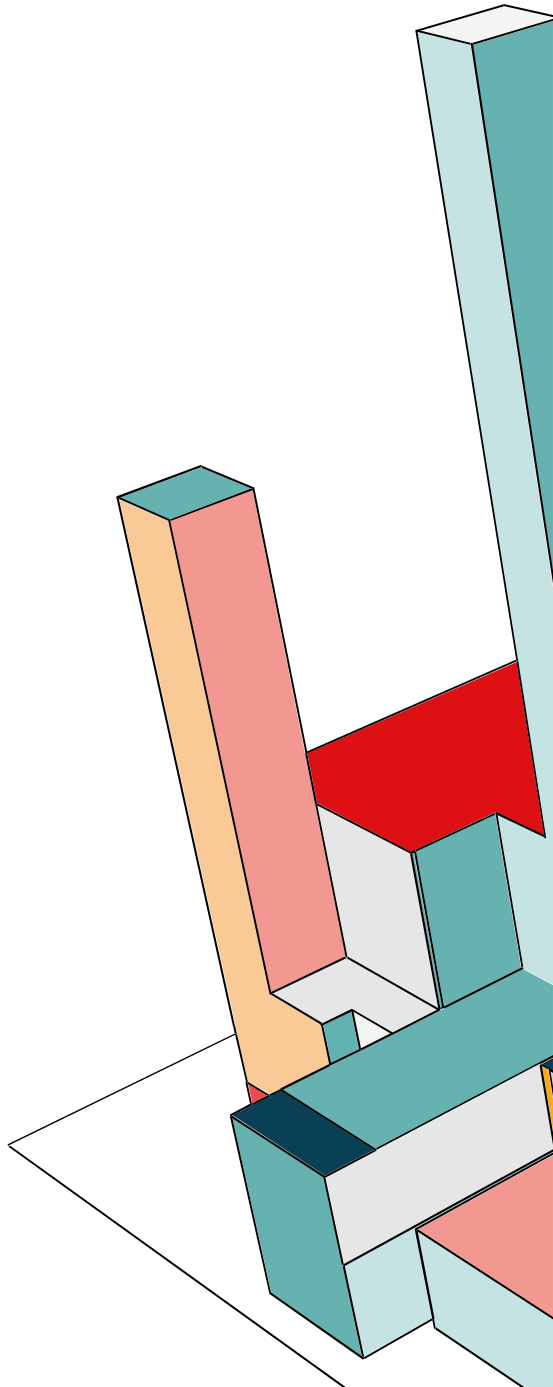


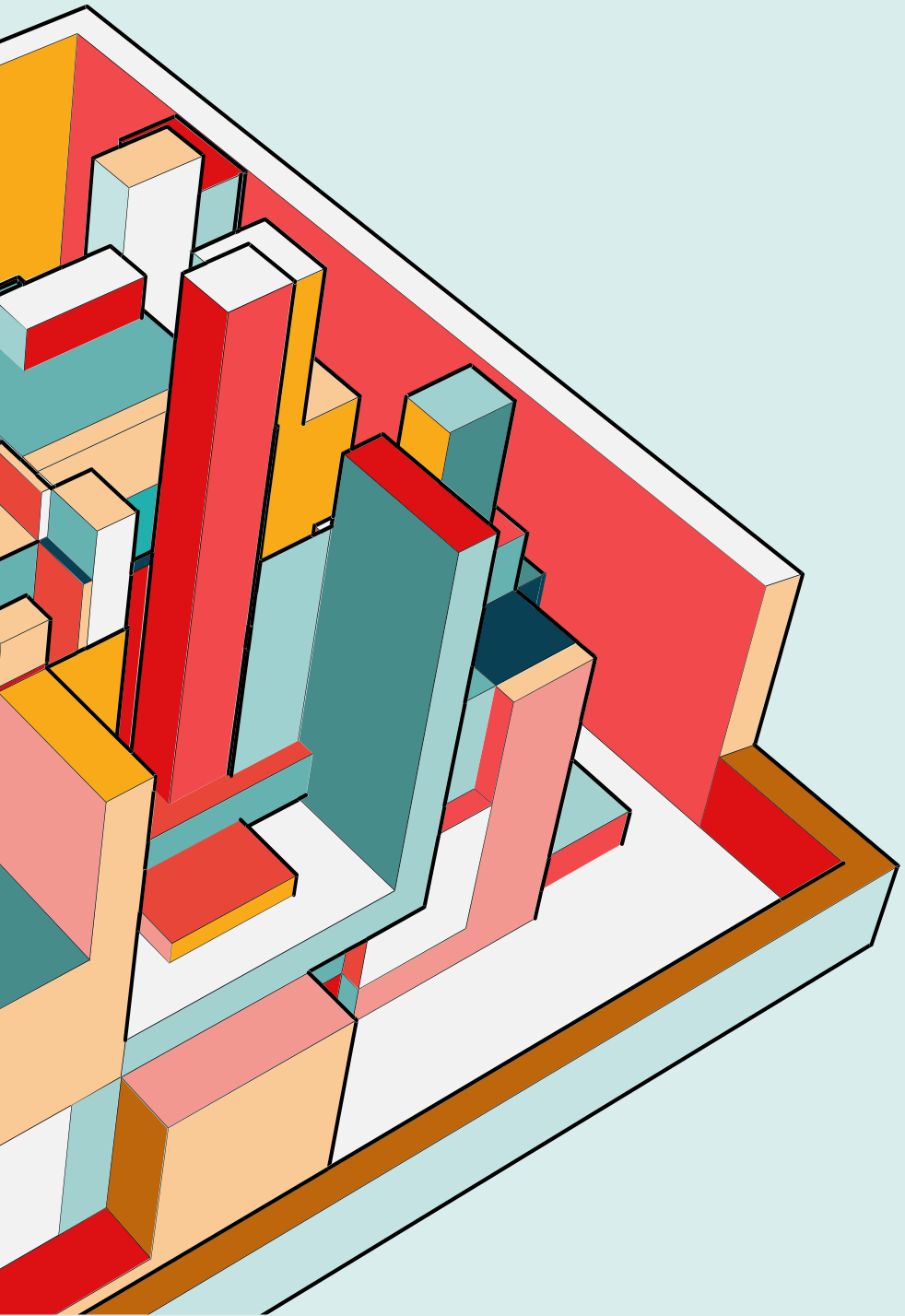
DATASET DETAILS

- Countries Table: This table contains information about different countries. It includes the name of each country to facilitate various analyses and operations.
- Currencies Table: This table provides a comprehensive list of currencies used in financial transactions. The structure allows for efficient referencing of currencies throughout the database, enabling accurate currency-related operations and reporting across various financial processes and international transactions.
- Merchants Table: This table captures essential information about businesses or individuals who accept payments through the platform. This structure enables efficient merchant management, facilitates reporting, and supports tailored services based on geographical considerations.



- Transactions Table: This data structure captures detailed information about financial exchanges within the system. The comprehensive record of transactions forms the core of the system, facilitating various financial operations, user activity tracking, and providing essential data for analytics and compliance purposes.
- Users Table: This data structure contains essential information about individuals who utilize the platform for financial transactions. This comprehensive user profile supports various functions including personalized services, security measures, and demographic analysis to inform business strategies and improve user experience.





DATA ANALYSIS IN MYSQL

Problem statement:

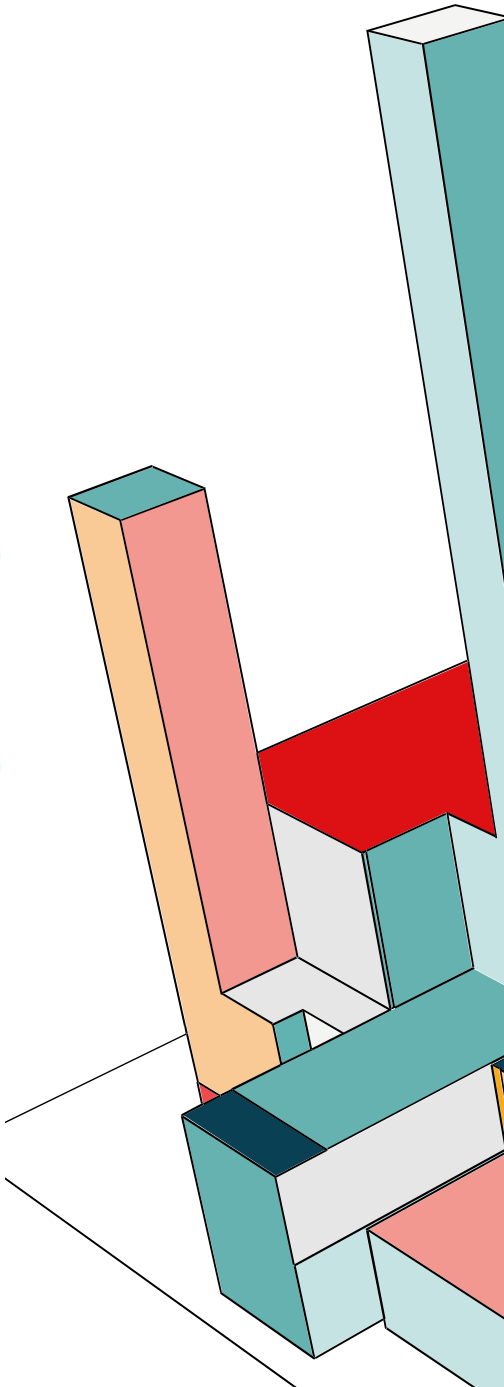
Determine the top 5 countries by total transaction amount for both sending and receiving funds in the last quarter of 2023 (October to December 2023). Provide separate lists for the countries that sent the most funds and those that received the most funds. Additionally, round the total sent and total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

```
select c.country_name, round(sum(t.transaction_amount),2) as total_sent from transactions t join users u
on t.sender_id = u.user_id join countries c on u.country_id = c.country_id
where t.transaction_date >= "2023-10-01" and t.transaction_date < "2024-01-01" group by 1 order by 2 desc limit 5;
```

```
select c.country_name, round(sum(t.transaction_amount),2) as total_received from transactions t join users u
on t.recipient_id = u.user_id join countries c on u.country_id = c.country_id
where t.transaction_date >= "2023-10-01" and t.transaction_date < "2024-01-01" group by 1 order by 2 desc limit 5;
```

country_name	total_sent
Iceland	173856.02
Zambia	162170.51
Israel	131061.51
Saudi Arabia	99142.61
Togo	94600.44

country_name	total_received
Zambia	174263.15
Iceland	140067.63
Saint Kitts and Nevis	121171.01
Timor-Leste	118107.68
Guadeloupe	98715.18



Problem statement:

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 transaction_amount > 10000 and date_format(transaction_date, "%Y") = 2023;
```

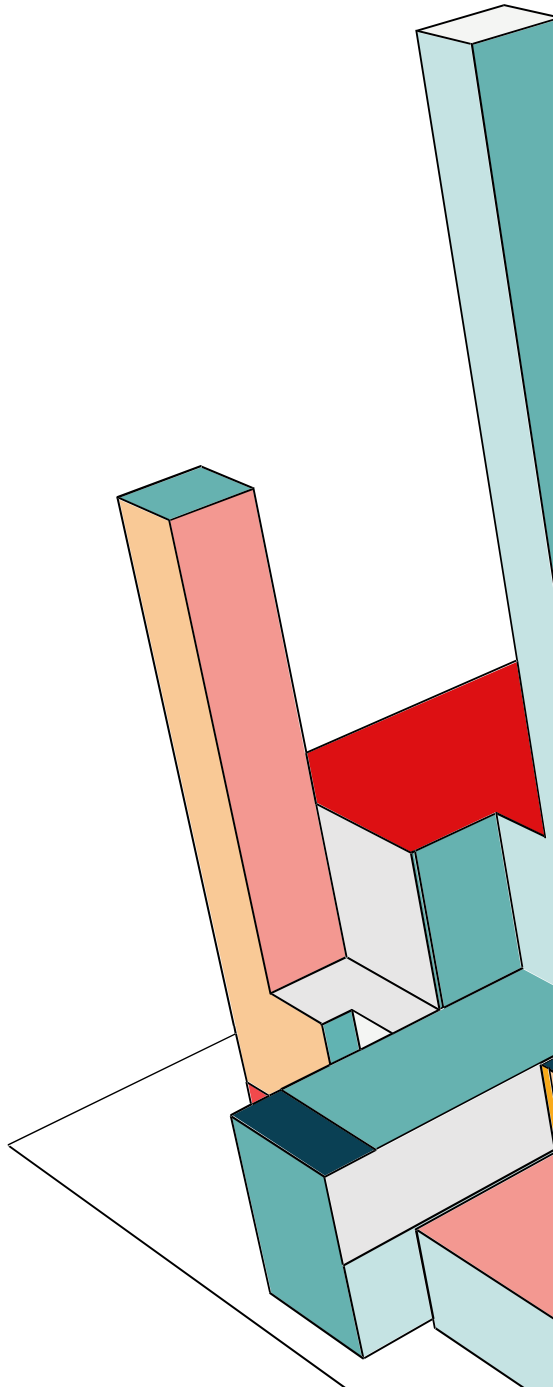
transaction_id	sender_id	recipient_id	transaction_amount	currency_code
10002	970	781	200000	EUR

Problem statement:

The finance team is analyzing currency conversion trends to manage exposure to currency risks. Which currency had the highest transaction amount from in the past one year up to today indicating the greatest exposure? (assume today is 22-05-2024)

```
select currency_code, sum(transaction_amount) as currency_sum from transactions where
transaction_date >= "2023-05-22" and transaction_date < "2024-05-23" group by 1 order by 2 desc limit 1;
```

currency_code	currency_sum
EUR	2599639.6000000015

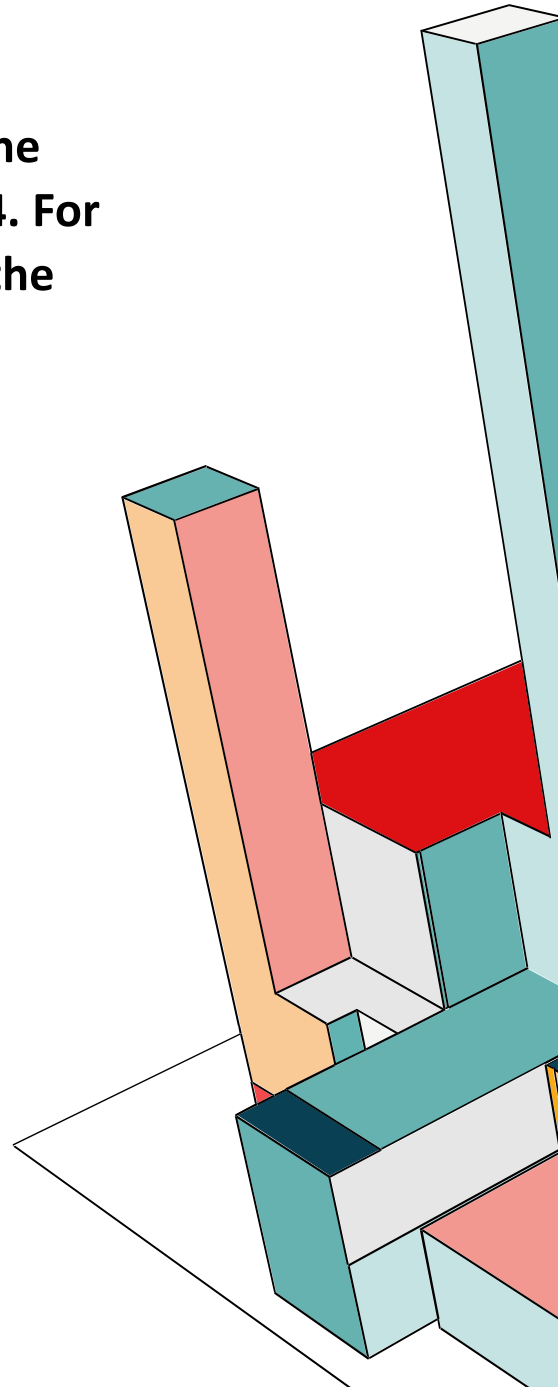


Problem statement:

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 , sum(t.transaction_amount) as total_received ,
avg(t.transaction_amount) as average_transactions from transactions t join
merchants m on m.merchant_id = t.recipient_id where t.transaction_date >= '2023-11-01'
and t.transaction_date < '2024-05-01'group by m.merchant_id , m.business_name
order by total_received desc limit 10;
```

merchant_id	business_name	total_received	average_transactions
151	Gould LLC	36380.25	9095.0625
91	Wilson-Mosley	32675.13	6535.026
129	Grant-Gallegos	29428.43	4904.738333333334
24	Simon PLC	28920.03	5784.005999999999
53	Shelton, Jones and Ferguson	28176.04	7044.01
17	Bender Ltd	26454.78	8818.26
25	Cherry and Sons	24795.120000000003	4959.024
113	Melton, Johnston and Lee	23079.199999999997	5769.799999999999
102	Knapp, Dean and Jones	21902.53	4380.505999999999
98	Lopez-Green	21896.18	5474.045



Problem statement:

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 currency_code, sum(transaction_amount) as total_converted from transactions
where transaction_date >= '2023-05-22' and transaction_date < '2024-05-22'
group by 1 order by 2 desc limit 3;
```

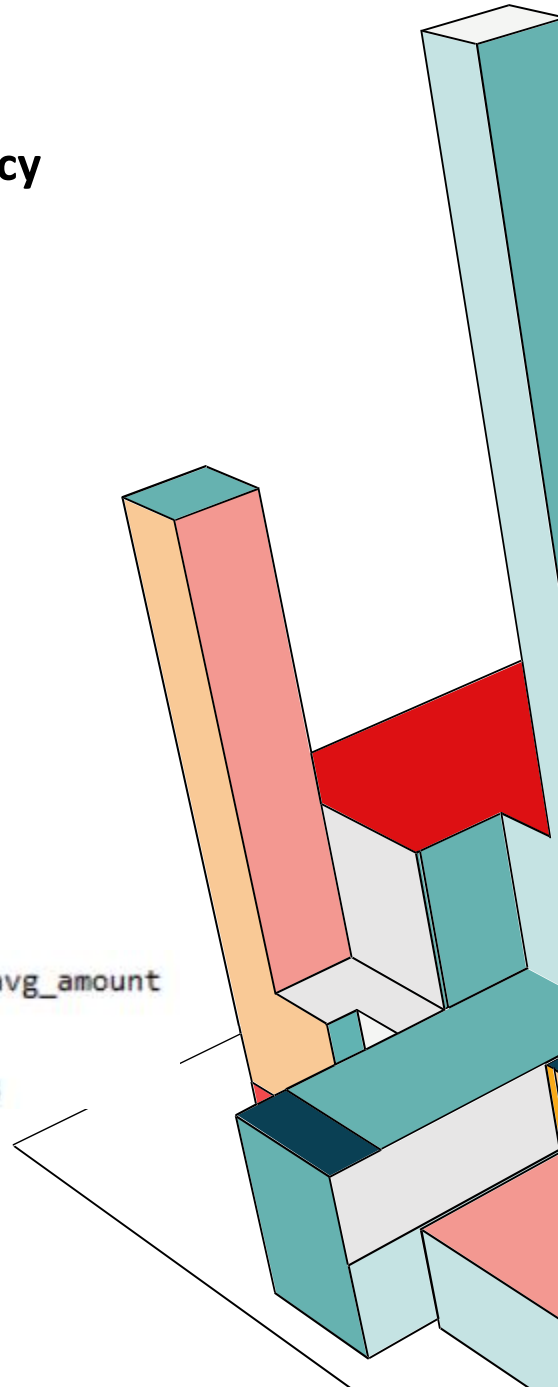
currency_code	total_converted
EUR	2583686.810000001
JPY	2326288.5899999994
USD	2251330.4000000004

Problem statement:

The sales team wants to identify top-performing merchants. Which merchant should be considered as the most successful in terms of total transaction amount received between November 2023 and April 2024?

```
select m.merchant_id, m.business_name, sum(t.transaction_amount) as total_amount, avg(t.transaction_amount) as avg_amount
 from transactions t join merchants m on m.merchant_id = t.recipient_id
where transaction_date >= "2023-11-01" and transaction_date < "2024-05-01" group by 1,2 order by 3 desc limit 1;
```

merchant_id	business_name	total_amount	avg_amount
151	Gould LLC	36380.25	9095.0625

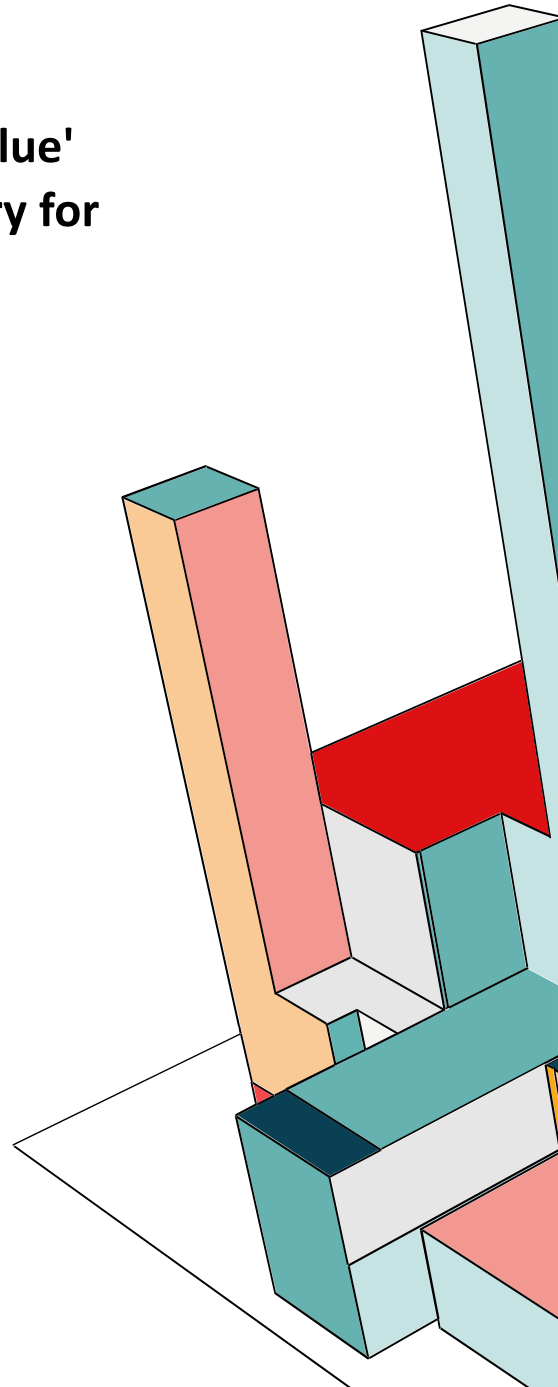


Problem statement:

The finance team is evaluating transaction classifications. Categorize transactions as 'High Value' (above $10,000) or 'Regular' (below $10,000) and calculate the total amount for each category for the year 2023.

```
select transaction_category, sum(transaction_amount) as total_amount from
(select transaction_amount, case when transaction_amount > 10000 then 'High Value'
when transaction_amount < 10000 then 'Regular' end as transaction_category from transactions
where date_format(transaction_date, '%Y') = 2023) temp
group by 1 order by 2 desc;
```

transaction_category	total_amount
Regular	11183960.760000005
High Value	200000

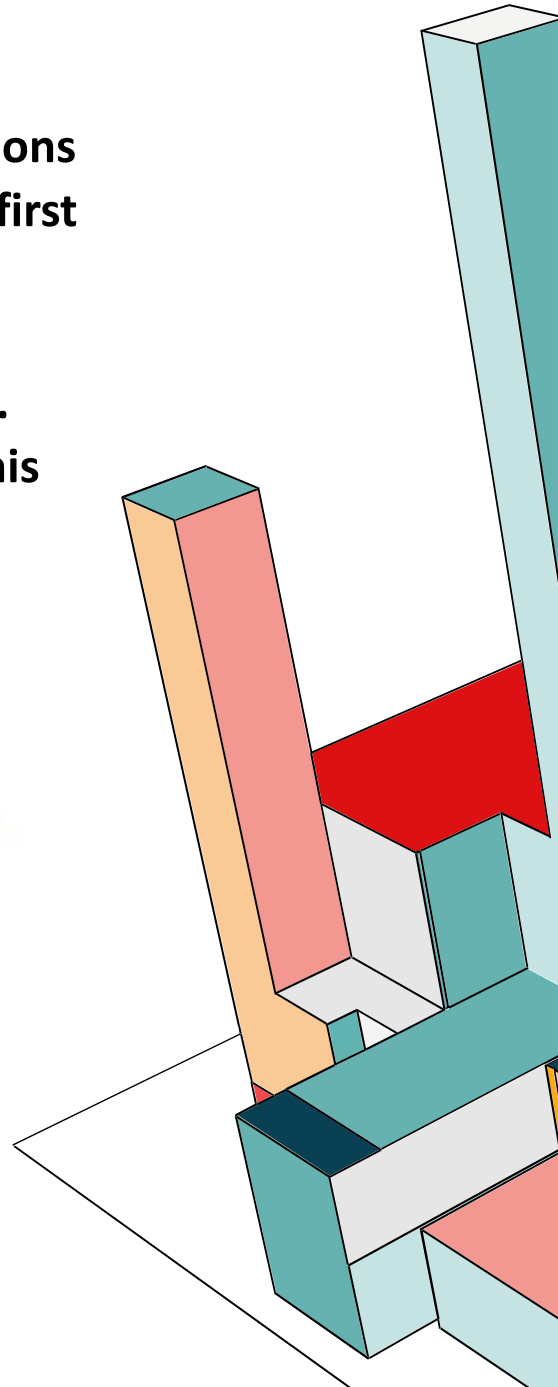


Problem statement:

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.

```
with cte as
(select t1.sender_id, u1.country_id as sender_country, t1.recipient_id, u2.country_id as receiver_country from
transactions t1 join users u1 on t1.sender_id = u1.user_id join users u2
on t1.recipient_id = u2.user_id where transaction_date >= "2024-01-01" and transaction_date < "2024-04-01")
select case when sender_country <> receiver_country then "International" when sender_country = receiver_country
then "Domestic" end as transaction_type, count(*) as transaction_count from cte group by 1 order by 2 desc;
```

transaction_type	transaction_count
International	587
Domestic	9

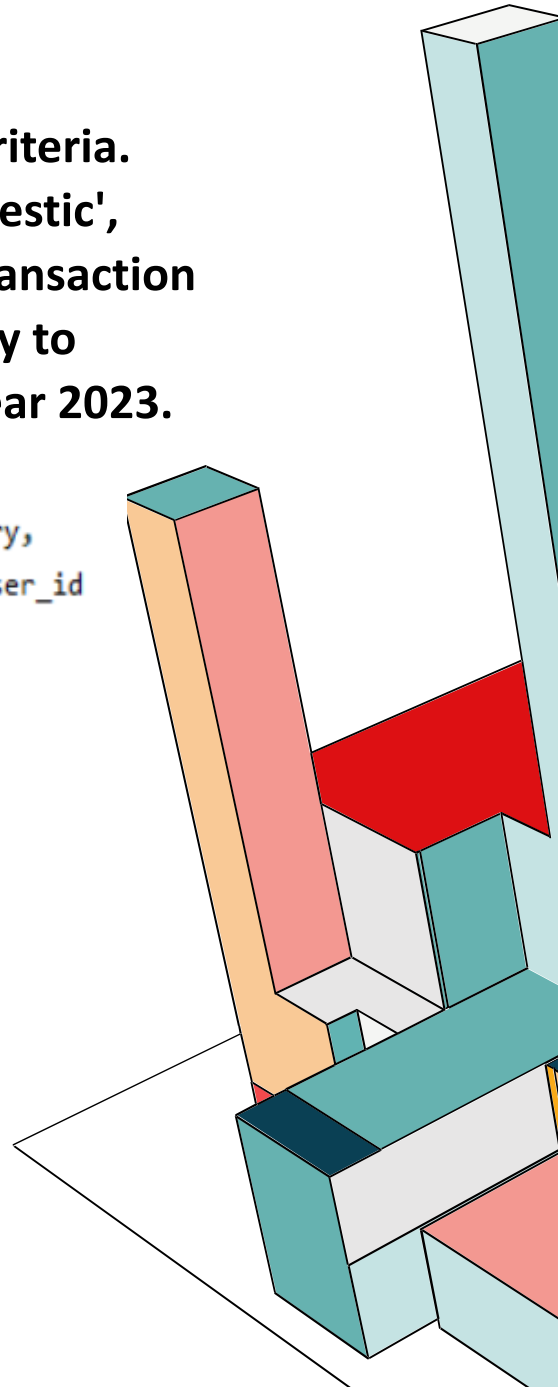


Problem statement:

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. Write a query to categorize each transaction and count the number of transactions in each category for the year 2023.

```
with cte as
(select t.transaction_date, t.sender_id, u1.country_id as sender_country, t.recipient_id, u2.country_id as receiver_country,
t.transaction_amount from transactions t join users u1 on t.sender_id = u1.user_id join users u2 on t.recipient_id = u2.user_id
where year(t.transaction_date) = 2023)
select case when sender_country <> receiver_country and transaction_amount > 10000 then "High Value International"
when sender_country = receiver_country and transaction_amount > 10000 then "High Value Domestic"
when sender_country <> receiver_country and transaction_amount < 10000 then "Regular International"
else "Regular Domestic" end as transaction_category, count(*) as transaction_count from cte group by 1;
```

transaction_category	transaction_count
Regular International	2233
Regular Domestic	45
High Value International	1



BUSINESS INSIGHTS

• 'Iceland', 'Zambia' and 'Israel' are the countries that sent the most funds and 'Zambia', 'Iceland' and 'Saint Kitts and Nevis' are those countries that received the most funds in the last quarter of 2023 so it assist in optimizing marketing strategies, tailoring financial products, and managing cross-border transaction risks.

• High-value transactions (2 lacs) were flagged, highlighting key users and regions contributing to significant monetary movement. This data is crucial for risk assessment, fraud detection, and targeted customer engagement strategies.

• 'Gould LLC' , 'Wilson-Mosley', 'Grant-Gallegos', 'Simon PLC', 'Shelton, Jones and Ferguson' are top 5 merchants with the highest transaction volumes, providing valuable information for prioritizing partnerships and customizing merchant services.

• The top 3 currencies for conversion were identified, with 'Euro' leading as the most frequently transacted currency and it helps for currency risk management and strategic planning for forex operations.

• There are '2233' Transactions which are categorized as 'regular international', while 'Regular Domestic' are '45' only and 'High Value International' with only '1' so mostly transactions are international with transaction amount < $10,000.

• In 'Oct 2023', there are highest total transactions amount compare with other months in year 2023.

• Users who transacted in at least 6 months were highlighted as consistent contributors, providing opportunities for enhanced retention strategies.

• Users with an average transaction amount exceeding $5,000 were identified, aiding in segmenting high-value customers for tailored services.

RECOMMENDATIONS BASED ON THE INSIGHTS

• **For Marketing Teams**: Focus on high-performing countries and consistently engaged users to design campaigns.

• **For Finance Teams**: Utilize high-value transaction data to manage risks and address compliance requirements.

• **For Merchant Relations**: Collaborate with top-performing merchants to sustain and grow transaction volumes.

• **For Risk Management**: Monitor currency trends and implement measures to mitigate exposure to forex volatility.

• **For Customer Retention**: Use insights from consistent and high-value user behavior to personalize offerings and improve engagement.

THANK YOU !

+91-8958936840

surabhigoyal198@gmail.com

https://github.com/SurabhiGoyal-Coder

https://www.linkedin.com/in/surabhi-goyal-36154a207