User Behaviour Analytics

—

Technical Test 1 - Data Analysis

Firyal Anis Safiyah

Objective

The company aims to better understand **user behavior and transaction patterns**.

Objectives:

- Profile users (gender, age groups, credit score)
- Analyze transaction trends (monthly, hourly, city, chip usage, card brand)
- Identify top spenders and potential churn users
- Detect suspicious transactions

Data Overview

Datasets used:

1. `users_data` – demographics, income, credit score, number of cards
2. `cards_data` – card brand and details
3. `transactions_data` – date, amount, city, chip usage, etc.

Total Users: 2,000

Total Transactions: $43M

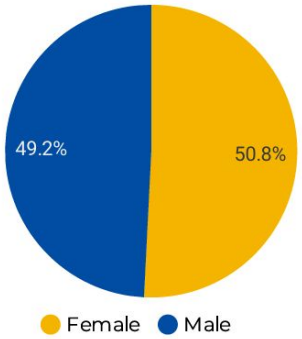
Total Cards : 6146

Data Period: Jan–Oct 2010

User Profile Analysis

Gender Distribution

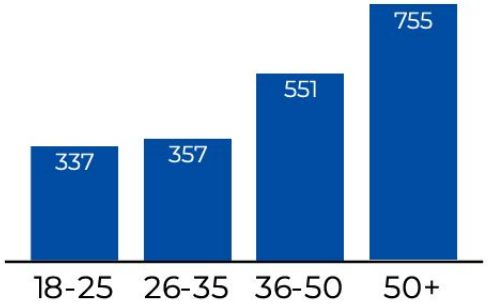
```
SELECT gender, COUNT(*) AS
total_users
FROM users_data
GROUP BY gender;
```



The distribution of users by gender is fairly balanced, with only a 0.8% difference.

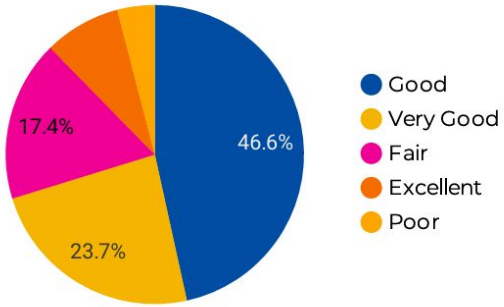
Age Groups

```
SELECT FORMAT_DATE('%Y-%m', DATE(date))
AS month,
       COUNT(*) AS total_transactions,
       SUM(amount) AS total_value
FROM transactions_data
GROUP BY month
```



The 50+ age group had the highest number of transactions, with 755 transactions, indicating that older users are more active in conducting transactions.

User Profile Analysis



Most users fall under the "Good" credit score category (46.6%), followed by "Very Good" (23.7%) and "Fair" (17.4%), indicating a generally healthy credit profile among users.

Credit Score

```
SELECT CASE
         WHEN credit_score < 580 THEN 'Poor'
         WHEN credit_score BETWEEN 580 AND
669 THEN 'Fair'
         WHEN credit_score BETWEEN 670 AND
739 THEN 'Good'
         WHEN credit_score BETWEEN 740 AND
799 THEN 'Very Good'
         ELSE 'Excellent'
       END AS credit_category,
       COUNT(*) AS user_count
FROM users_data
GROUP BY credit_category
ORDER BY user_count DESC;
```

Transaction Trend Analysis (Time)

hour	id
00	10,310
01	8,693
02	8,883
03	7,954
04	8,655
05	13,837
06	56,435
07	67,598
08	65,436
09	64,812
10	65,349
11	70,926
12	72,048
13	67,022
14	66,846
15	64,972
16	65,046
17	36,651

Hourly Transactions

```
SELECT EXTRACT(HOUR FROM
date) AS hour_of_day,
       COUNT(*) AS
total_transactions
FROM transactions_data
GROUP BY hour_of_day
ORDER BY hour_of_day;
```

Transaction volume peaks between 11 AM and 1 PM, indicating that midday is the busiest period for user activity.

Monthly Transactions

```
SELECT FORMAT_DATE('%Y-%m', DATE(date)) AS month,
       COUNT(*) AS total_transactions,
       SUM(amount) AS total_value
FROM transactions_data
GROUP BY month
ORDER BY month;
```

Transaction amounts remain relatively stable from March to September, with a sharp decline observed in October.

Transaction Analysis (Details)

Top 5 Users

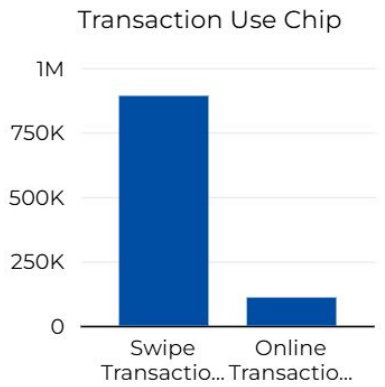
```
SELECT u.id AS user_id,
       u.current_age,
       SUM(t.amount) AS
total_spent
FROM transactions_data t
JOIN users_data u ON
t.client_id = u.id
GROUP BY u.id, u.current_age
ORDER BY total_spent DESC
LIMIT 5;
```

	client_id	amount
1.	96	206,985.98
2...	1686	177,742.05
3...	1340	176,558.1
4...	840	164,120.59
5...	464	155,547.7

User with ID 96 is the highest spender with a total transaction amount of 206,985.98, followed closely by users 1686 and 1340.

Chip Usage

```
SELECT use_chip,
       COUNT(id)
AS
transaction_total
FROM
transactions_data
GROUP BY use_chip;
```

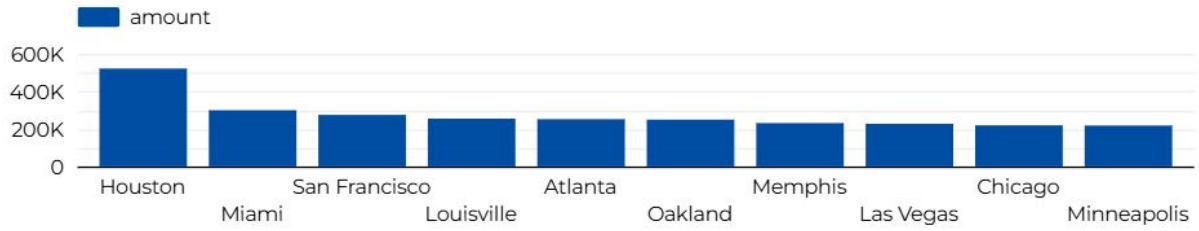


Swipe transactions are significantly more common than online transactions, indicating user preference for physical card usage.

Merchant & Card Analysis

Top Cities

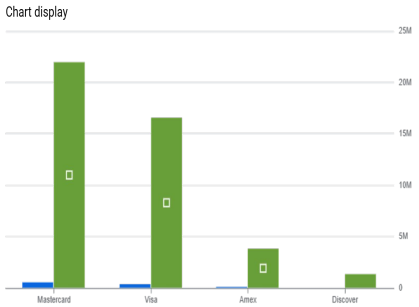
```
SELECT merchant_city,
       COUNT(id) AS transaction_total
FROM transactions_data
WHERE merchant_city != 'ONLINE'
GROUP BY merchant_city
ORDER BY transaction_total DESC;
```



Houston recorded the highest number of face-to-face transactions, far surpassing other cities, amounting to $522k.

Card Brand

```
SELECT c.card_brand,
       COUNT(t.id) AS
total_transactions,
       SUM(t.amount) AS total_value
FROM transactions_data t
JOIN cards_data c ON t.card_id = c.id
GROUP BY c.card_brand
ORDER BY total_value DESC;
```



Mastercard is the most used card both in terms of total transactions and transaction value, followed by Visa.

Customer Status & Fraud Check

Customer Status

```
SELECT customer_status, COUNT(*) AS
total_customers
FROM (
  SELECT client_id,
         CASE
           WHEN DATE_DIFF(DATE
'2010-10-23', DATE(MAX(date)), DAY) > 90
             THEN 'Potential Churn'
           ELSE 'Active'
         END AS customer_status
  FROM transactions_data
  GROUP BY client_id)
GROUP BY customer_status;
```

client_id	user_status
1223	Potential Churn
534	Potential Churn
1753	Potential Churn
1403	Active

1 - 10 / 1126

To detect inactive customers, marked as "Potential Churn," which indicates that they have not made a transaction in more than 90 days, there are 3 people.

Suspicious Transactions

```
SELECT client_id,
       DATE(date) AS transaction_date,
       COUNT(DISTINCT merchant_city) AS
distinct_cities
FROM transactions_data
WHERE merchant_city != 'ONLINE'
GROUP BY client_id, transaction_date
HAVING COUNT(DISTINCT merchant_city) > 1
ORDER BY transaction_date DESC;
```

Row	client_id	transaction_date	distinct_cities
1	1519	2010-10-23	3
2	1344	2010-10-23	2
3	1815	2010-10-23	2
4	1966	2010-10-23	3
5	213	2010-10-23	2

There are users conducting transactions in three different cities on the same day. Although this does not necessarily indicate fraud, it could be an early warning sign that needs to be monitored depending on the user profile and context.

Key Findings

- **User Profile**: Gender is balanced; most active users are **aged 50+**.
- **Credit Score**: Majority of users have a "**Good**" score, indicating healthy credit behavior.
- **Card Usage**: **Mastercard** dominates both in usage and transaction value.
- **Top Cities**: **Houston** leads offline transaction volumes.
- **Transaction Time**: Peak activity occurs at **11 AM–1 PM;** transactions dip sharply in October.
- **User Activity**: Out of 2,000 users, only 1,126 have made transactions, leaving **874 users inactive** and **3 at risk of churn**.
- **Fraud Signal**: There are users transacted in **3 cities** on the same day, requires monitoring.

Recommendations

- **Re-engage inactive users (874)** with targeted campaigns.
- **Follow up churn-risk users (3)** to prevent full disengagement.
- **Leverage peak hours (11 AM–1 PM)** for promos and notifications.
- **Plan campaigns around August**, the month with highest transaction volume.
- **Monitor multi-city same-day transactions** for potential fraud signals.
- **Strengthen Mastercard partnerships**, as it's the top used card.

🎯 Focus on user retention, fraud prevention, and optimizing peak activity windows.

Dashboard


USER BEHAVIOUR ANALITYCS DASHBOARD
Overview of user demographics, activity patterns, and transaction behaviour
Total Users
2,000
Total Transactions
$43,797,460.97
Average Transactions
$43.80
Total Active Cards
6,146
Data Period : January - October 2010
card_type
transactions_date
Users by Gender
49.2%
50.8%
Female
Male
Users by Age
337
357
551
755
18-25
26-35
36-50
50+
Top 5 Users by Transaction Value
client_id
amount
1. 96 206,985.98
2... 1686 177,742.05
3... 1340 176,558.1
4... 840 164,120.59
5... 464 155,547.7
Credit
17.4%
46.6%
23.7%
Good
Very Good
Fair
Excellent
Poor
Transactions by Hour
ho...
id
00 10,310
01 8,693
02 8,883
03 7,954
04 8,655
05 13,837
06 56,435
07 67,598
08 65,436
09 64,812
10 65,349
11 70,926
12 72,048
13 67,022
14 66,846
15 64,972
16 65,046
17 36,651
18 35,621
19 34,453
20 31,993
Total Transaction by Month
amount
5M
4.5M
4M
3.5M
3M
January
February
March
April
May
June
July
August
September
October
November
Dece...
Transaction Use Chip
1M
750K
500K
250K
0
Swipe Transacti...
Online Transacti...
User Status
client_id
user_status
1223 Potential Churn
534 Potential Churn
1753 Potential Churn
1403 Active
1 - 10 / 1126
Transaction by City
amount
600K
400K
200K
0
Houston
Miami
San Francisco
Louisville
Atlanta
Oakland
Memphis
Las Vegas
Chicago
Minneapolis