Credit Card Risk Analysis

A Project By
Prasenjit

Introduction

Credit Card Risk Assessment: Insights into Spending, Risk, and Fraud Detection

Overview:

• Credit card transactions generate vast amounts of **financial data** that can be analyzed to understand **spending behavior, credit risk, and fraud patterns**.

• This project leverages **SQL queries** to extract meaningful insights from transaction data, helping financial institutions with **risk management, customer segmentation, and fraud prevention**.

• The goal is to **identify high-risk customers, detect fraud patterns, and optimize credit strategies** for better financial decision-making.

Dataset Description

Tables Used (Provide a brief description of the datasets):

- **Users Data** (Client info: age, income, credit imit, etc.)

- **Transactions Data** (Purchase details: amount, merchant, date, etc.)

- **Cards Data** (Credit card details: brand, limit, number of cards)

Key Data Fields

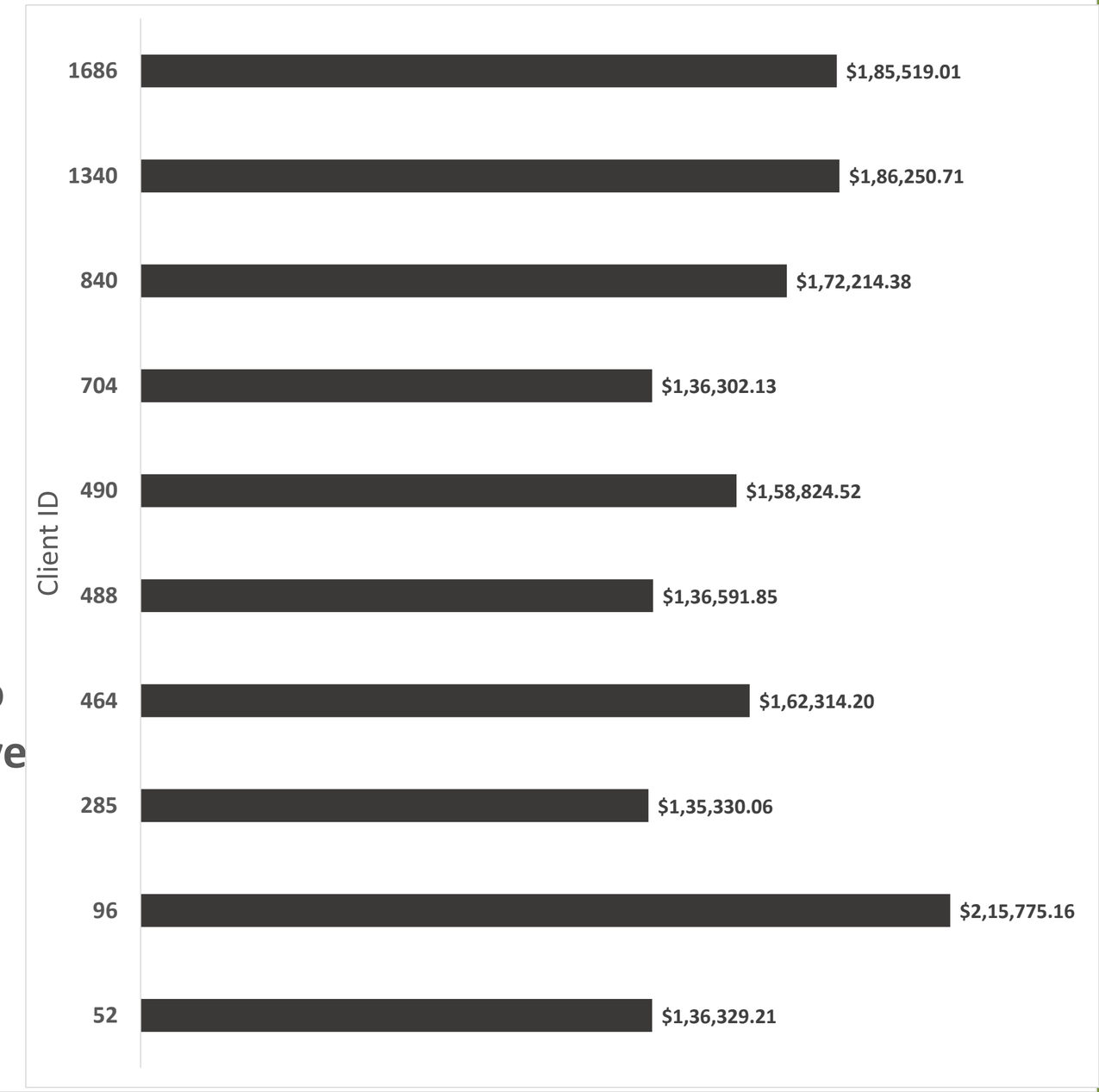
- client_id, yearly_income, credit_limit, amount, merchant_id, date, errors, mcc, etc.

Any Data Cleaning Steps Taken?

- Removed duplicates and null values
- Standardized date formats
- Merged relevant tables

Top Revenue Generating Customers

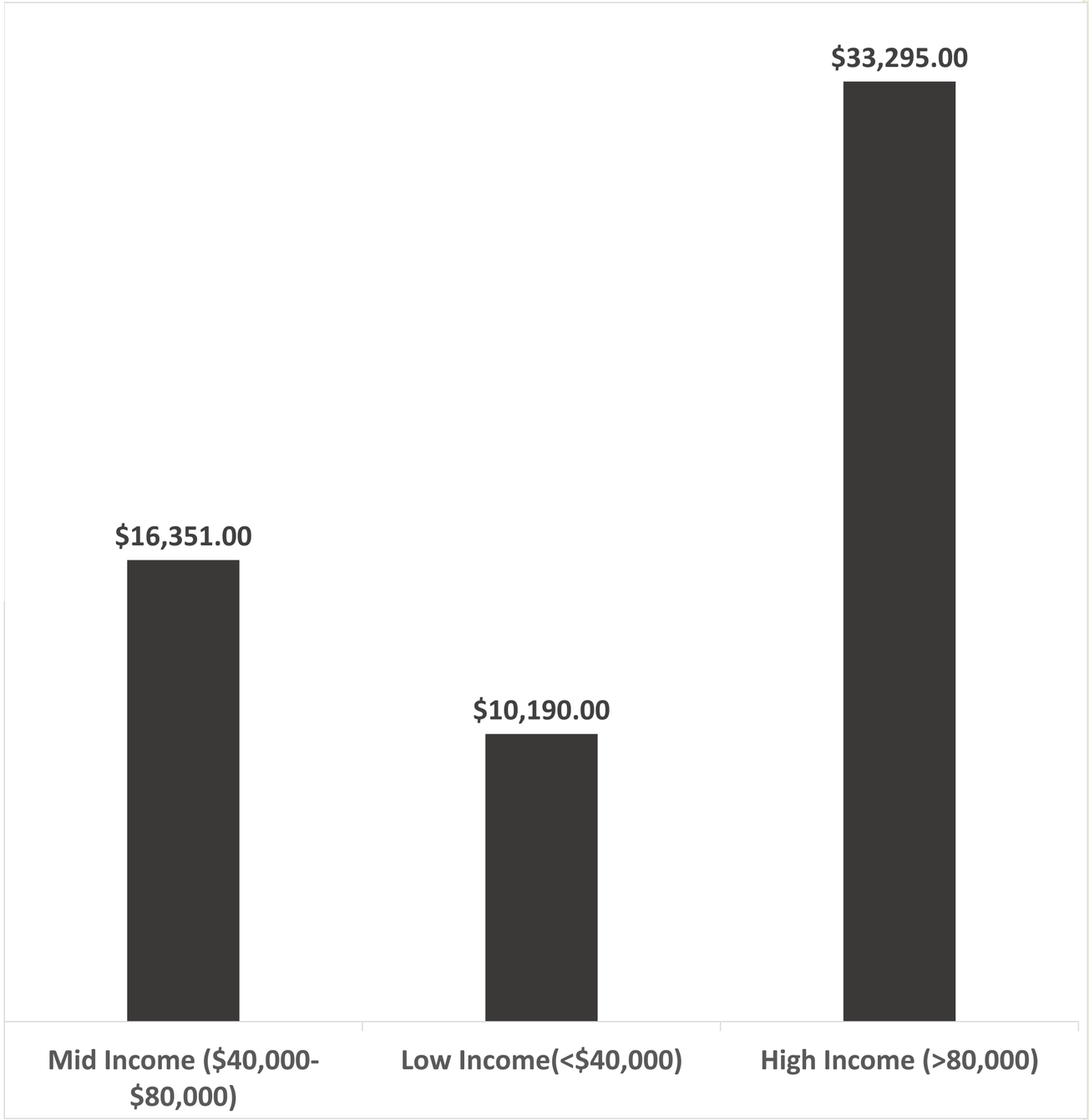
```sql
select td.client_id, round(sum(td.amount),2) as Total_spending
from transactions_data td inner join users_data ud using(client_id)
group by td.client_id
order by Total_spending desc
limit 10;
```



Insight –
High Value Customers who contribute significantly to revenue and transaction volume are ideal for exclusive loyalty programs, personalized offers, and premium service. By Offering personalized rewards may encourage further spending.

Average Credit Limit by Income Bracket

```sql
select
    case
        when yearly_income < 40000 Then "Low Income(<$40,000)"
        when yearly_income between  40000 and 80000
        Then "Mid Income ($40,000-$80,000)"
        when yearly_income > 80000 Then "High Income (>80,000)"
        end AS Income_Bracket,
        round(avg(c.credit_limit),2) as Avg_Credit_Limit
from users_data u join cards_data c using(client_id)
group by Income_Bracket;
```

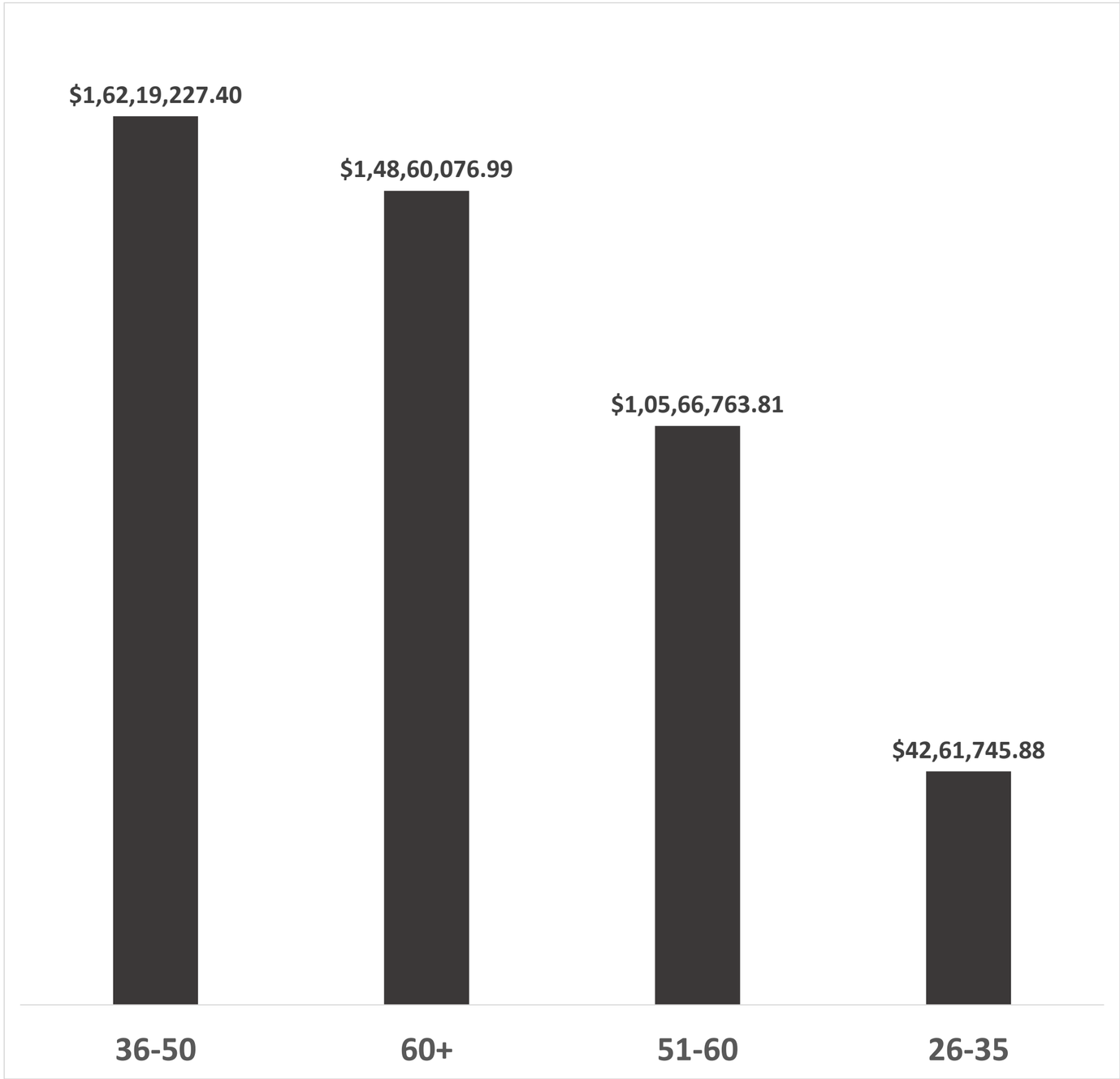


Insight –

High-income customers have high credit limits, it indicates several business opportunities and risks. These customers can be targeted with exclusive benefits and personalized financial products to boost engagement and spending. Encouraging high-limit customers to utilize more credit can maximize profitability. Offering perks like cashback, rewards, and low-interest EMI options can further incentivize spending.

Total Credit Card Spending by Age Group

```sql
select
      case
          when current_age between 18 and 25 then "18-25"
          when current_age between 26 and 35 then "26-35"
          when current_age between 36 and 50 then "36-50"
          when current_age between 51 and 60 then "51-60"
          else "60+"
          end as Age_Group,
          round(sum(t.amount),2) as Total_Spent
from users_data u join transactions_data t using(client_id)
group by Age_Group
order by Total_Spent desc;
```



Insights –

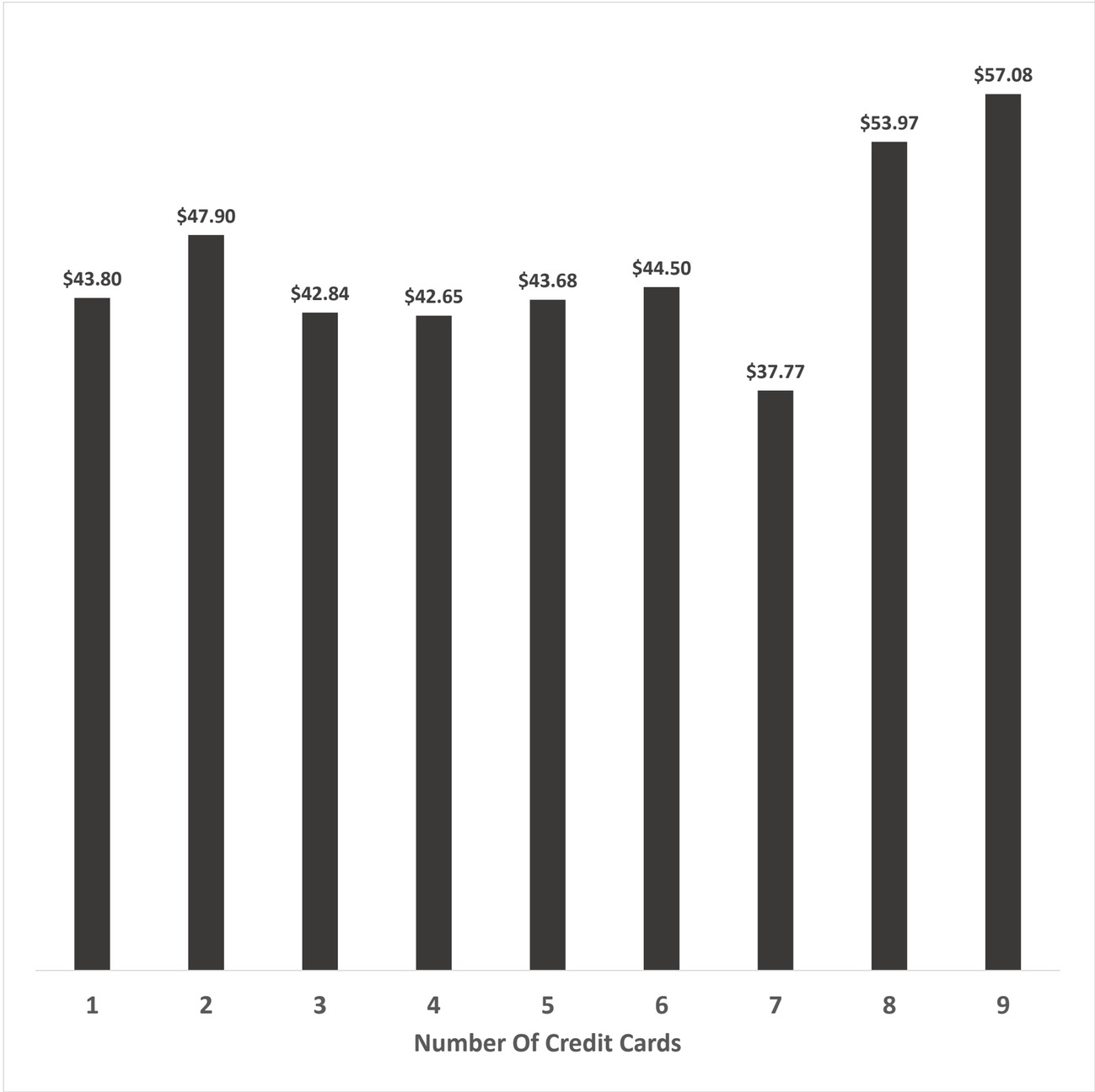
Since the Age group 36 to 50 and 60+ people are spending more. Premium credit cards with benefits tailored to their needs can be offered :

✅ **For 36-50: Higher credit limits, family travel perks, business rewards.**

✅ **For 60+: Medical expense benefits, cashback on essentials, fraud protection.**

Credit Card Ownership vs. Total Spending

```sql
select u.num_credit_cards, round(avg(t.amount),2) as Average_Transaction
from users_data u left join transactions_data t using(client_id)
group by u.num_credit_cards
order by Average_Transaction desc;
```



Insight –

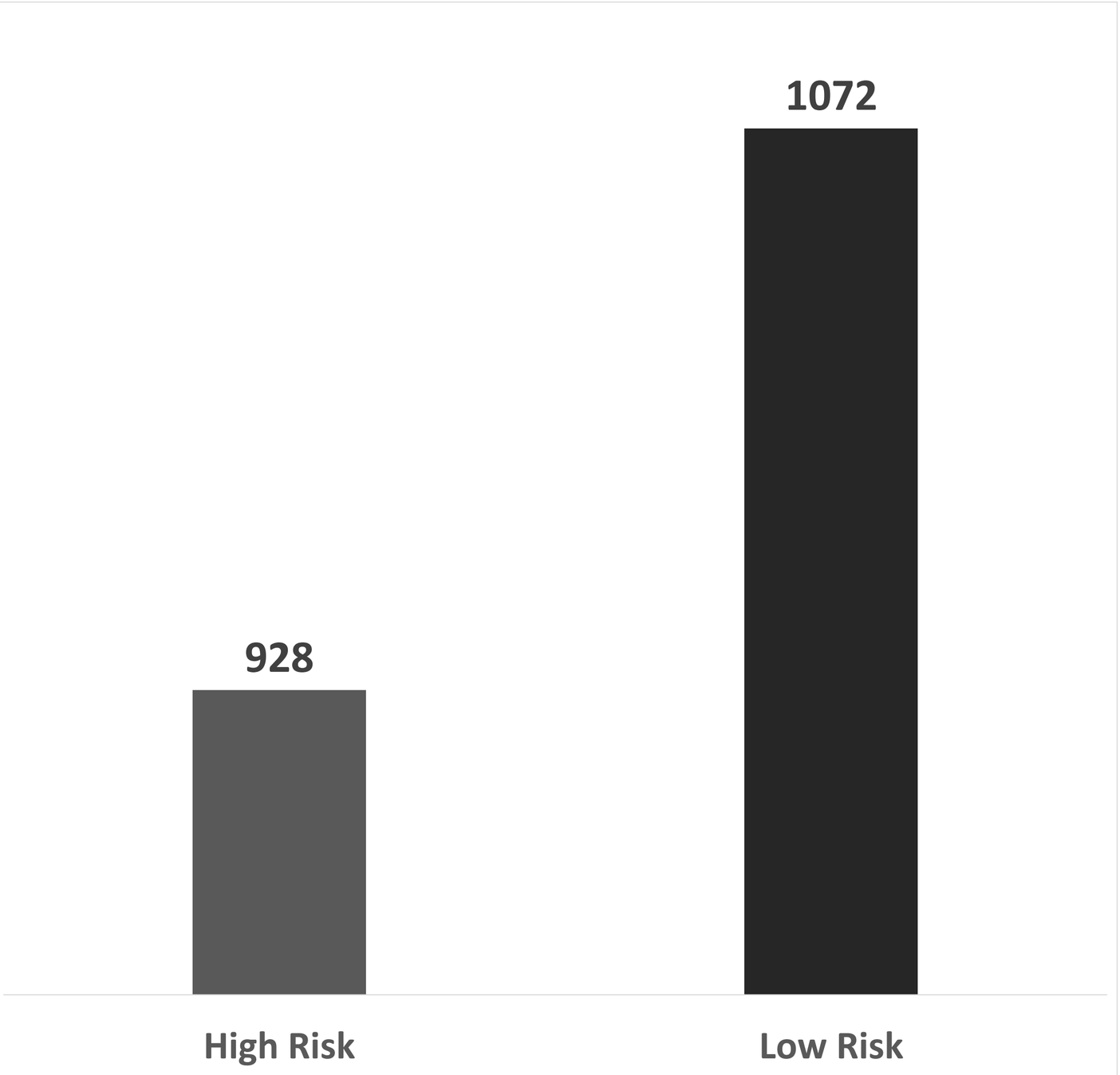
Customers with multiple credit cards contribute significantly to transaction fees, interest income, annual fees are likely financially active & high spenders. These customers can be targeted with exclusive premium credit cards, higher limits, and VIP benefits ,Encouraging higher spending by offering tier-based rewards (e.g., platinum, elite, black cards).

High Risk Clients

```sql
with Risk_Analysis as(
select client_id,yearly_income,total_debt,
case
    when total_debt > yearly_income * 1.5 Then "High Risk"
    else "Low Risk"
end as Risk_level-- debt is 150% or more of income
from users_data)

select Risk_level,count(*) as Risk_Count
from Risk_Analysis
Group by Risk_Level;
```

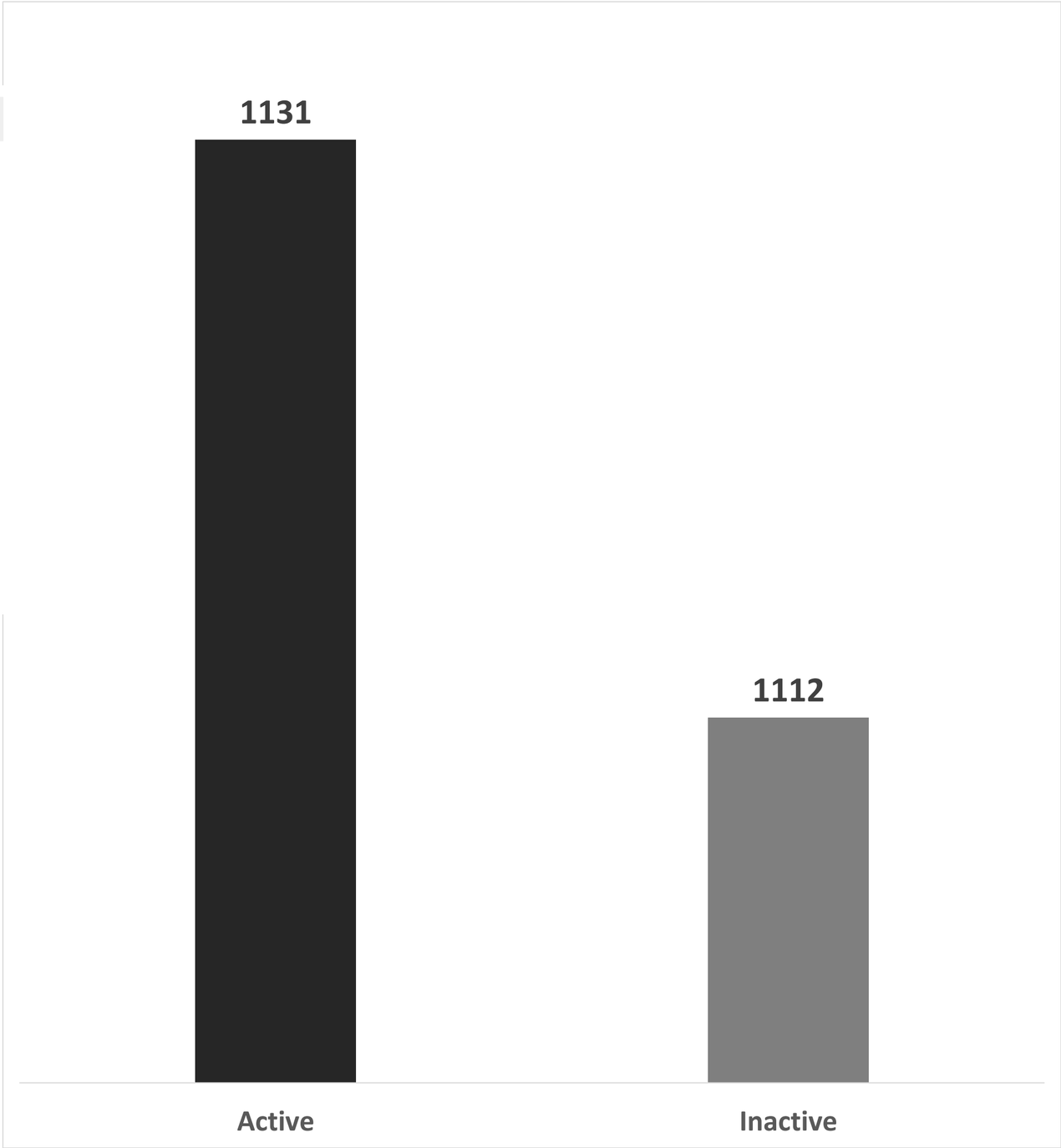
Insight –
Customers with low income but high debt may struggle to repay their loans or credit card balances. So they are Flagged as High Risk Customers.
They can be Offered low-interest debt consolidation loans to help them manage payments. "Income-Based Repayment Plans" can be given to ease their financial burden.
Credit policies should be adjusted to prevent over-lending to high-risk customers and Set up automated reminders & payment flexibility to reduce missed payments.



Inactive Users

```sql
with Inactive_state as(select distinct client_id,
    case
        when str_to_date(`date`,"%d-%m-%Y %H:%i") <
        (SELECT DATE_SUB(MAX(STR_TO_DATE(date, '%d-%m-%Y %H:%i')), INTERVAL 6 MONTH)
        FROM transactions_data)
        then "Inactive" else "Active" end as Status
from transactions_data)

select status,count(*) as Number_Of_Customers
from Inactive_state
group by Status;
```



Insight –
Identifying the inactive users can provide valuable insights into customer retention, engagement, and potential revenue loss. Users who haven't transacted in 6+ months may be at risk of churning (leaving the platform or switching to competitors).

Customers at Risk of Credit Exhaustion

```sql
select c.client_id,c.card_id,credit_limit,round(sum(t.amount),2) as Total_Spent,
       round(credit_limit - sum(t.amount),2) as Remaining_limit
from cards_data c join transactions_data t using(client_id)
group by c.client_id,c.card_id,credit_limit
having Remaining_limit < (credit_limit * 0.1) and Remaining_limit > 0
-- less than 10% limit left
order by Remaining_limit;
```

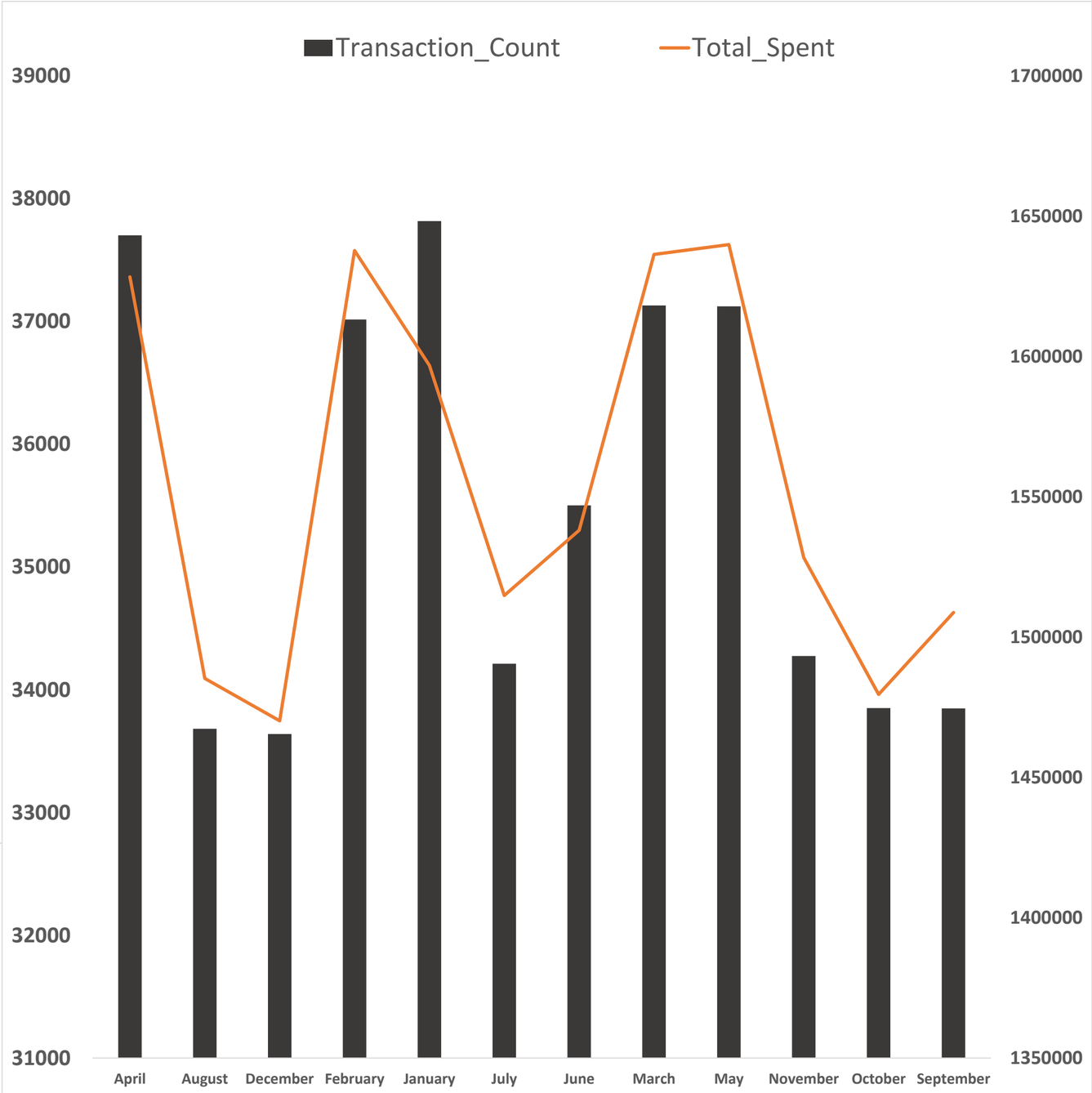
client_id	card_id	credit_limit	Total_Spent	Remaining_limit	Credit Utilisation %
875	2380	13788	13784.01	3.99	**100%**
634	3541	37620	37578.45	41.55	**100%**
1876	3315	21552	21498.98	53.02	**100%**
1219	3466	32903	32848.27	54.73	**100%**
855	3670	15975	15865.72	109.28	**99%**
618	2282	11591	11479.45	111.55	**99%**
678	1300	8600	8460.49	139.51	**98%**
1394	3468	19597	19445	152	**99%**
1360	2184	12367	12123.69	243.31	**98%**
1794	5205	4392	4121.01	270.99	**94%**
988	4384	17100	16828.46	271.54	**98%**
609	2290	19436	19109.04	326.96	**98%**
1600	4251	48754	48414.86	339.14	**99%**
494	1847	29335	28979.37	355.63	**99%**
410	2801	15000	14569.11	430.89	**97%**
875	193	14300	13784.01	515.99	**96%**
1024	3564	33103	32578.56	524.44	**98%**
520	2020	34278	33748.2	529.8	**98%**
697	1298	20195	19649.18	545.82	**97%**
1193	4023	38290	37740.57	549.43	**99%**
357	3706	41221	40667.77	553.23	**99%**
1437	1942	30404	29832.99	571.01	**98%**
1556	412	11200	10624.34	575.66	**95%**
849	5121	16640	16020.47	619.53	**96%**
480	2929	51051	50360.21	690.79	**99%**
1461	4535	23754	23022.35	731.65	**97%**
11	1336	51012	50150.61	861.39	**98%**
1303	6085	25292	24368.73	923.27	**96%**

Insight –

- Customers with low remaining credit limits are more likely to miss payments or default.
- Implementing early alerts for at-risk customers to encourage timely payments.
- Offer structured repayment plans or debt consolidation options.

Seasonal Trend shows Increase in Credit Card Transactions and $ spent during Festivals

```sql
select monthname(str_to_date(date,"%m-%d-%Y %H:%i")) as Month,
    count(*) as Transaction_Count,
    round(sum(amount),2) as Total_Spent
    FROM transactions_data
where monthname(str_to_date(date,"%m-%d-%Y %H:%i")) <> ""
group by Month
order by Month;
```



Insight –

Increase spending during Festivals means customers are more receptive to promotions .Consumers may exceed their credit limits or take personal loans to fund purchases.

BNPL (Buy Now, Pay Later) options gain popularity during festivals. Banks & lenders can offer special festival loan rates or pre-approved credit limit increases. Promote EMI conversion options to encourage big-ticket spending.

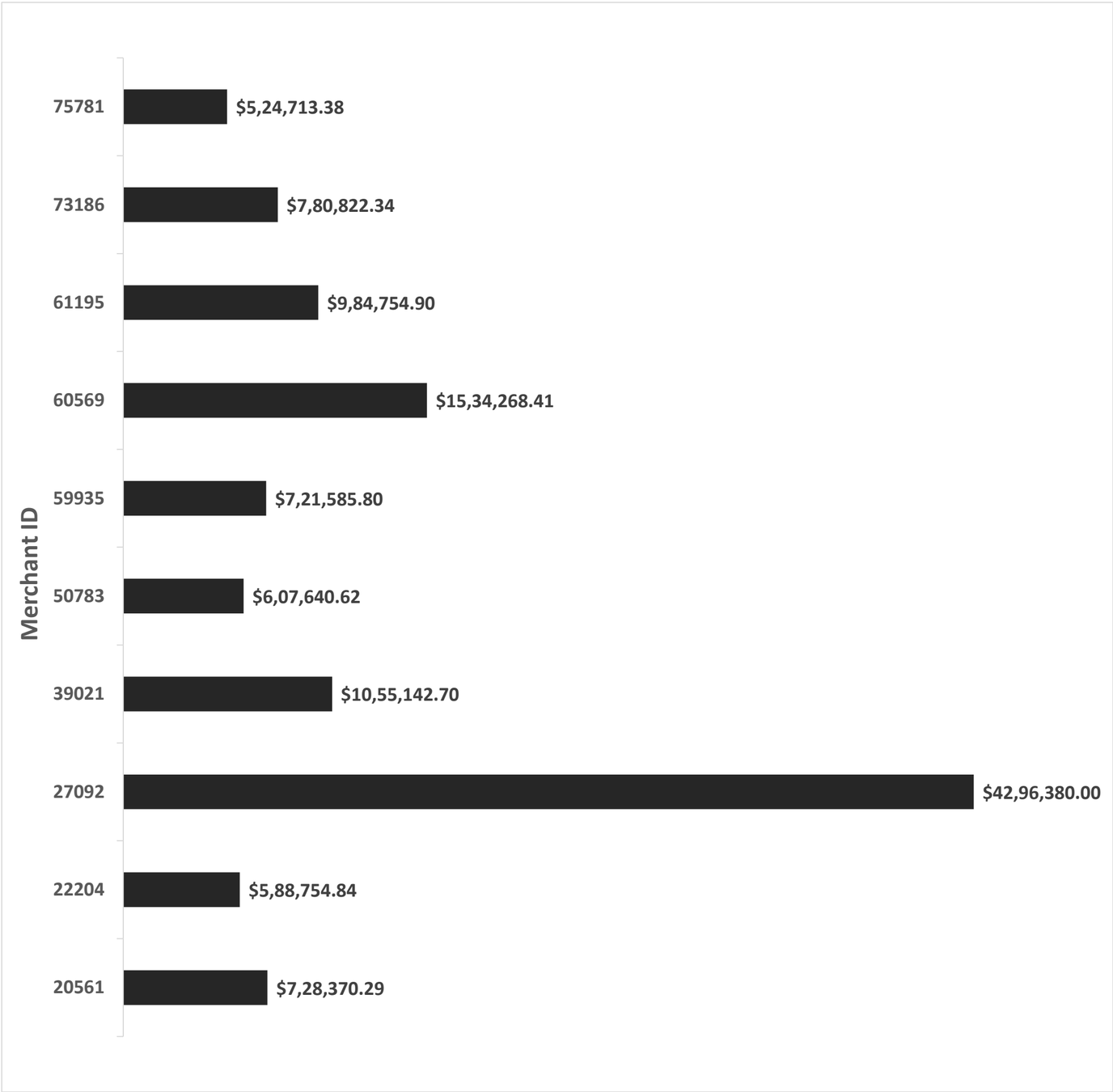
Top Merchants by Transaction Volume

```sql
select merchant_id,round(sum(amount),2) as Transaction_Volume
from transactions_data
group by merchant_id
order by Transaction_Volume desc
limit 10;
```

Insight –
Identifying top merchants helps analyze which industries dominate spending. Retail, e-commerce, travel, fuel, groceries, and entertainment may have different transaction trends.

✅ **Tailor marketing efforts based on popular merchant categories (e.g., more credit card rewards for travel during holidays).**

✅ **Encourage new businesses to enter high-performing sectors with insights on consumer demand.**



ONLINE VS OFFLINE TRANSACTIONS

```sql
select
    count(transaction_id) as Total_transaction,
    sum(case when use_chip = "Online Transaction" then 1 else 0 end)
    /count(transaction_id)  as Percent_Online_Transactions,
    sum(case when use_chip = "Swipe Transaction" then 1 else 0 end)
    /count(transaction_id)  as Percent_Offline_Transactions
from transactions_data;
```

Total_ transaction	Percent_ Online_ Transactions	Percent_ Offline_ Transactions
1048575	11%	89%

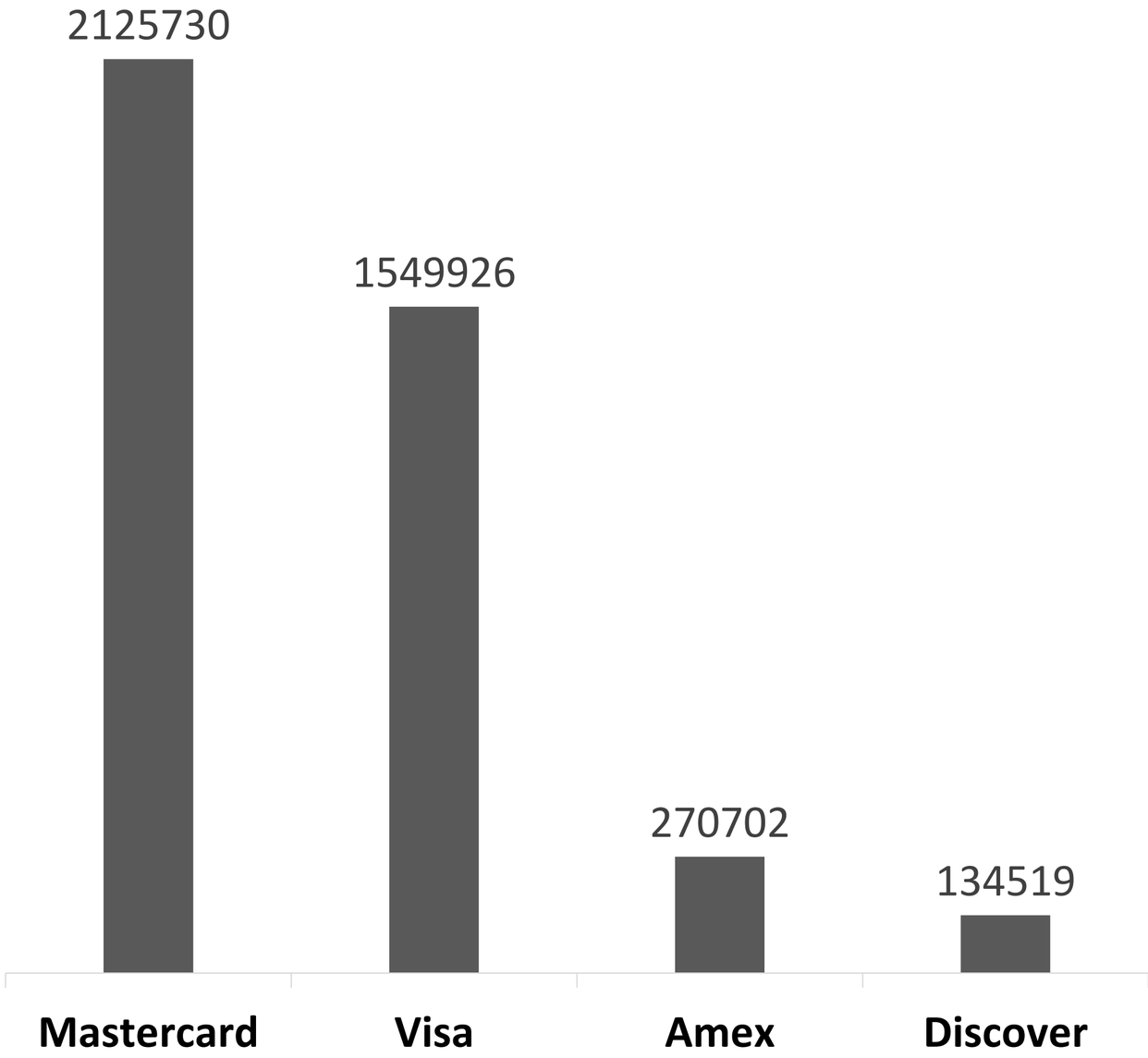
Insight –

Offline transactions dominating means it's an opportunity to boost online payments via:

- ✅ **Digital incentives** 📲
- ✅ **Secure payment methods** 🔐
- ✅ **Merchant partnerships** 🤝

Transaction Numbers Shows Customers Prefers Master Card over other Credit Card Brands

```sql
select c.card_brand,count(t.transaction_id) as Total_Transaction
from cards_data c join transactions_data t using(client_id)
group by c.card_brand
order by Total_Transaction desc;
```



Insight –
More transactions on a specific card brand indicate **stronger customer preference and trust** in that network.It may be due to better acceptance, customer benefits, or established reputation.

✅ **Banks can strengthen partnerships** with the leading card brand.

✅ **Merchants can prioritize promotions for the most-used card** (e.g., “10% off on all Visa transactions”).

Merchants with High Transaction Error Rates

```sql
select merchant_id,count(*) as Total_Transactions,
       sum(case when `errors` <> "" then 1 else 0 end) as Failed_Transactions,
       (sum(case when `errors` <> "" then 1 else 0 end)/count(*)) * 100 as Failure_rate
       from transactions_data
group by merchant_id
having Failure_rate > 60 -- filtering merchants with more than 5% failure rate
order by Failed_Transactions desc;
```

merchant_id	Total_Transactions	Failed_Transactions	Failure_rate
55284	3	2	67
25514	3	2	67
39886	2	2	100
33189	3	2	67
25947	2	2	100
26531	2	2	100
7682	3	2	67
821	2	2	100
29188	1	1	100
68167	1	1	100
27521	1	1	100
82546	1	1	100
60472	1	1	100
35003	1	1	100
39475	1	1	100
43232	1	1	100
87719	1	1	100
88410	1	1	100
79061	1	1	100
45647	1	1	100
4869	1	1	100
85595	1	1	100

Insight –

Frequent transaction failures may indicate **technical glitches in payment systems** at specific merchants.

Possible reasons:

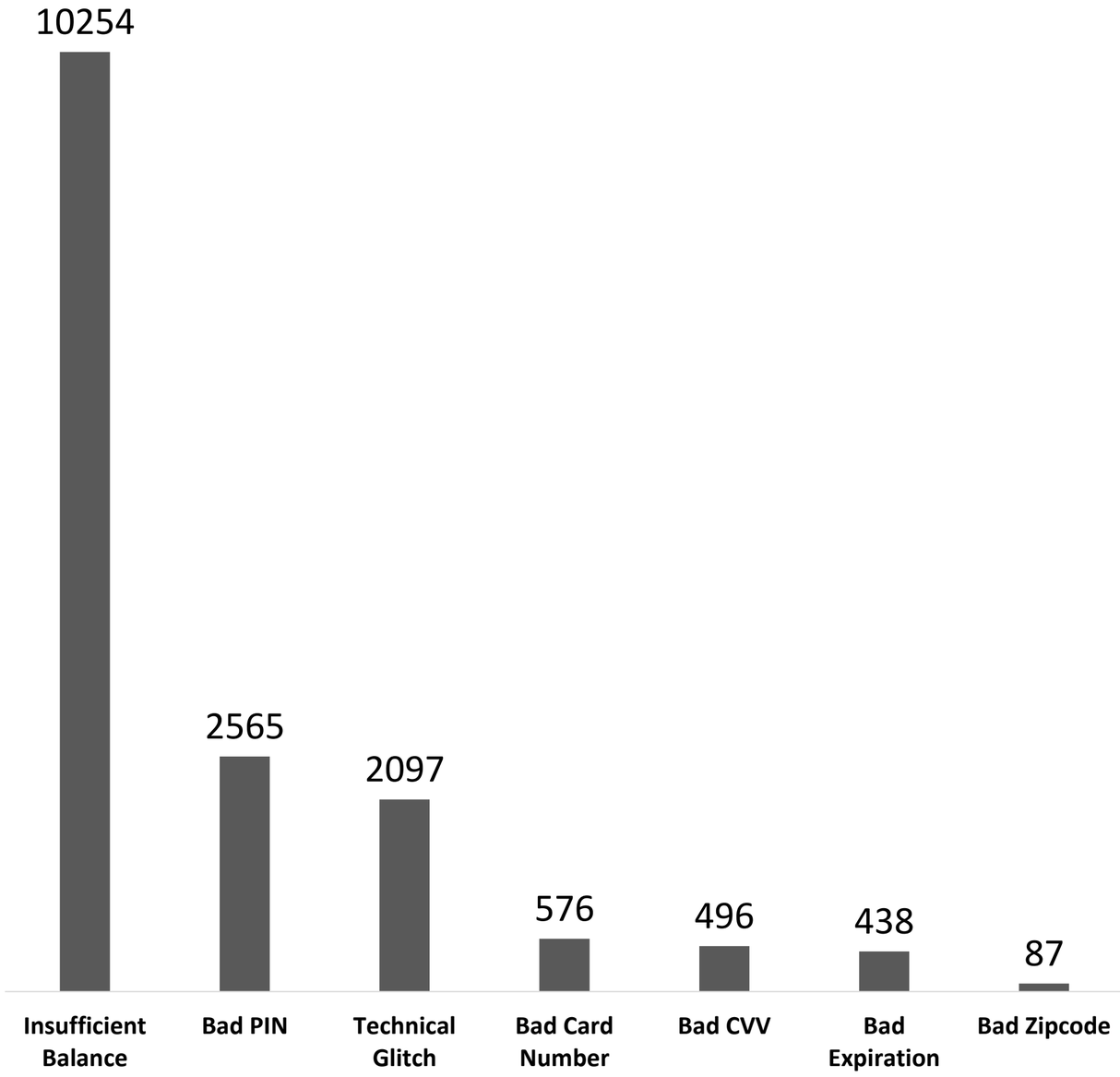
Outdated POS terminals 🏧 , Network connectivity issues 🌐 , Bank/payment gateway failures 🚫

✅ **Merchants need to upgrade their payment systems** to prevent revenue loss.

✅ **Banks & payment processors should investigate recurring technical issues** with these merchants.

Top Error Types In Transactions

```sql
select `errors`,count(*) as Error_count
from transactions_data
where `errors` not like ""
group by `errors`
order by Error_Count desc;
```



Insight –
Customers frequently facing insufficient balance error may benefit from short-term loans or Offering installment-based payments or Buy Now, Pay Later options.

Users making transactions in multiple cities within the same day.

```
233 •  select client_id,date,count(distinct(merchant_city)) as Unique_Cities
234    from transactions_data
235    group by client_id,date
236    having Unique_Cities > 1;
237
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

client_id	date	Unique_Cities
0	28-02-2010 19:17	2
1	13-10-2010 12:14	2
1	20-01-2010 11:47	2
4	09-10-2010 08:35	2
11	08-04-2010 17:18	2
16	21-03-2010 11:43	2
21	07-03-2010 13:17	2
34	03-08-2010 02:12	2
38	01-02-2010 06:42	2
38	17-10-2010 06:40	2
38	29-09-2010 06:54	2
39	17-04-2010 10:38	2
39	25-02-2010 11:25	2
39	29-10-2010 11:45	2
44	08-05-2010 08:25	2
46	01-07-2010 06:32	2

Insight –

As there are many cards showing transactions made in a single day in different cities it may indicate –

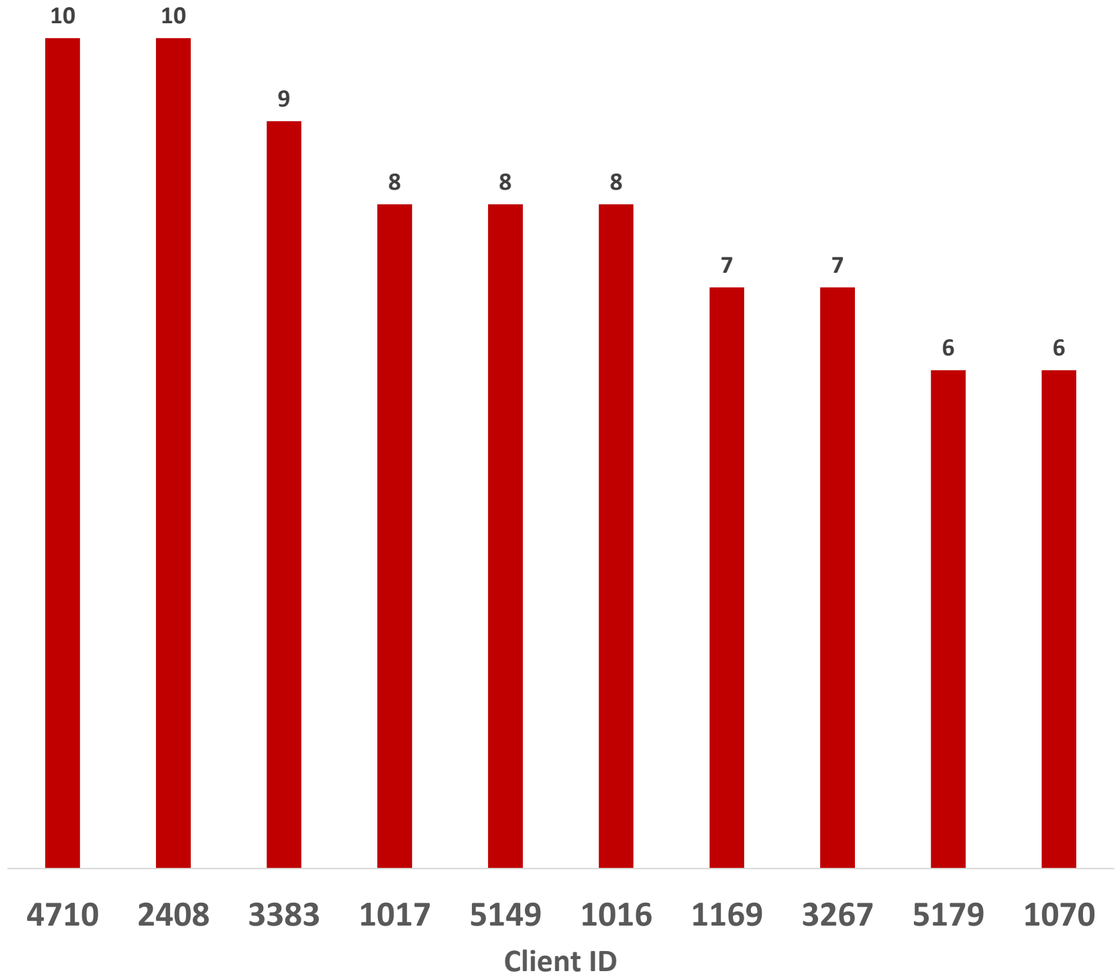
Card Testing Fraud: Fraudsters often test stolen or fake cards by making multiple small transactions in a short time to check if they are valid before making larger purchases.

Bot or Automated Transactions:
If a single card is making high-frequency transactions at the same merchant, it could indicate automated transactions or a compromised account.
Money Laundering Risks: Some fraudsters move money through quick transactions between merchants to avoid detection.

Clients Making Transactions With High Risk Merchants Frequently

```sql
WITH transactions_time AS (
    SELECT card_id ,
        STR_TO_DATE(date, '%d-%m-%Y %H:%i') AS transaction_time
    FROM transactions_data
),
transaction_counts AS (
    SELECT t1.card_id, t1.transaction_time,
        COUNT(*) OVER (PARTITION BY t1.card_id
        ORDER BY t1.transaction_time RANGE INTERVAL 20 MINUTE PRECEDING) AS transaction_count
    FROM transactions_time t1
)
SELECT DISTINCT card_id, COUNT(*) as total_transactions
FROM transaction_counts
WHERE transaction_count > 5  -- Threshold for 'numerous transactions'
GROUP BY card_id
having total_transactions > 5
ORDER BY total_transactions DESC;
```

The query helps detect potential fraud by identifying users making unusually high transactions within a short time, enabling better security measures and fraud prevention



Key Insights -

•**Potential Fraud Detection** – Identifies users making unusually high transactions in a short time span (more than 5 transactions within 20 minutes), which may indicate fraudulent activity.

•**Suspicious Spending Patterns** – Highlights card IDs with excessive transaction volumes, signaling potential card misuse or unauthorized access.

•**Operational Insights for Banks & Merchants** – Assists in refining fraud detection algorithms and improving customer authentication processes.

•**Policy Optimization** – Can be used to set automated alerts or implement transaction limits per time frame to reduce fraud risks.

Problem Statement & Importance

What are the key challenges in credit card risk assessment?

- **High-risk customers may default on** payments
- Fraudulent transactions lead to financial losses
- Understanding customer behavior can improve financial services

Why does this need data-driven insights?

- Manual risk assessments are inefficient
- Banks need automated systems to detect risks & fraud
- Data helps in **better decision-making** for credit allocation

Key Findings

Customer Spending & Segmentation

- High-income users have higher credit limits and spending.
- Multiple cardholders show higher average transactions.
- Customers near credit limits can be offered increases or financial advice.

Merchant & Transaction Analytics

- Few merchants account for most transactions—useful for partnerships.
- Online vs. offline transaction split reveals digital adoption trends.
- High failure rates at some merchants indicate potential payment issues.

Fraud Detection & Risk Assessment

- Transactions in multiple cities on the same day may indicate fraud.
- Many small transactions at the same merchant/IP suggest card testing.
- High-risk MCCs (gambling, crypto) require strict monitoring.
- Users with multiple declined transactions may need fraud review.

Conclusion

- The credit card risk assessment project provided **valuable insights into customer spending behavior, credit risk, and fraud detection**.

- By analyzing **spending patterns, credit usage, and transaction errors**, we can improve **credit allocation strategies** and **customer retention efforts**.

- **Fraud detection queries** helped uncover **suspicious transactions**, including high-risk MCC transactions and users making purchases in multiple cities within a short timeframe.

- The insights can aid financial institutions in making **data-driven decisions** to enhance **credit risk management and fraud prevention**.

Thank you for your valuable Time.