mandiri sekuritas

User Behavior Analysis

Technical Test

Analyzing the past, understanding the present, and forecasting a stronger financial future.

2025

Presented by: **Rizal Rahman R.**

Table of Contents

www.reallygreatsite.com

Dataset Overview

mandiri sekuritas

users_data

2000 Rows

Column	Data Type	Description
id	INT	Unique identifier for each user.
current_age	INT	Current age of the user.
retirement_age	INT	Expected retirement age of the user.
birth_year	INT	Year of birth of the user.
birth_month	INT	Month of birth of the user.
gender	STRING	Gender of the user (Male/Female).
address	STRING	Residential address of the user.
latitude	FLOAT	Latitude coordinate of the user's location.
longitude	FLOAT	Longitude coordinate of the user's location.
per_capita_income	STRING	Per capita income of the household (string with currency symbol).
yearly_income	STRING	Annual income of the user (string with currency symbol).
total_debt	STRING	Total outstanding debt of the user (string with currency symbol).
credit_score	INT	Credit score of the user.
num_credit_cards	INT	Number of credit cards owned by the user.

cards_data

6146 Rows

Column	Data Type	Description
id	INT	Unique identifier for each card.
client_id	INT	User ID (foreign key linking to users_data.id).
card_brand	STRING	Card brand (e.g., Visa, Mastercard, etc.).
card_type	STRING	Type of card (Debit, Credit, Prepaid).
card_number	INT	Unique card number.
expires	STRING	Card expiration date (MM/YYYY).
cvv	INT	Card CVV security code.
has_chip	BOOLEAN	Indicates if the card has a chip (TRUE/FALSE).
num_cards_issued	INT	Number of cards issued for the user.
credit_limit	STRING	Credit limit of the card (string with currency symbol).
acct_open_date	STRING	Date when the card account was opened.
year_pin_last_changed	INT	Year the card PIN was last changed.
card_on_dark_web	BOOLEAN	Whether the card was found on the dark web (True/False).

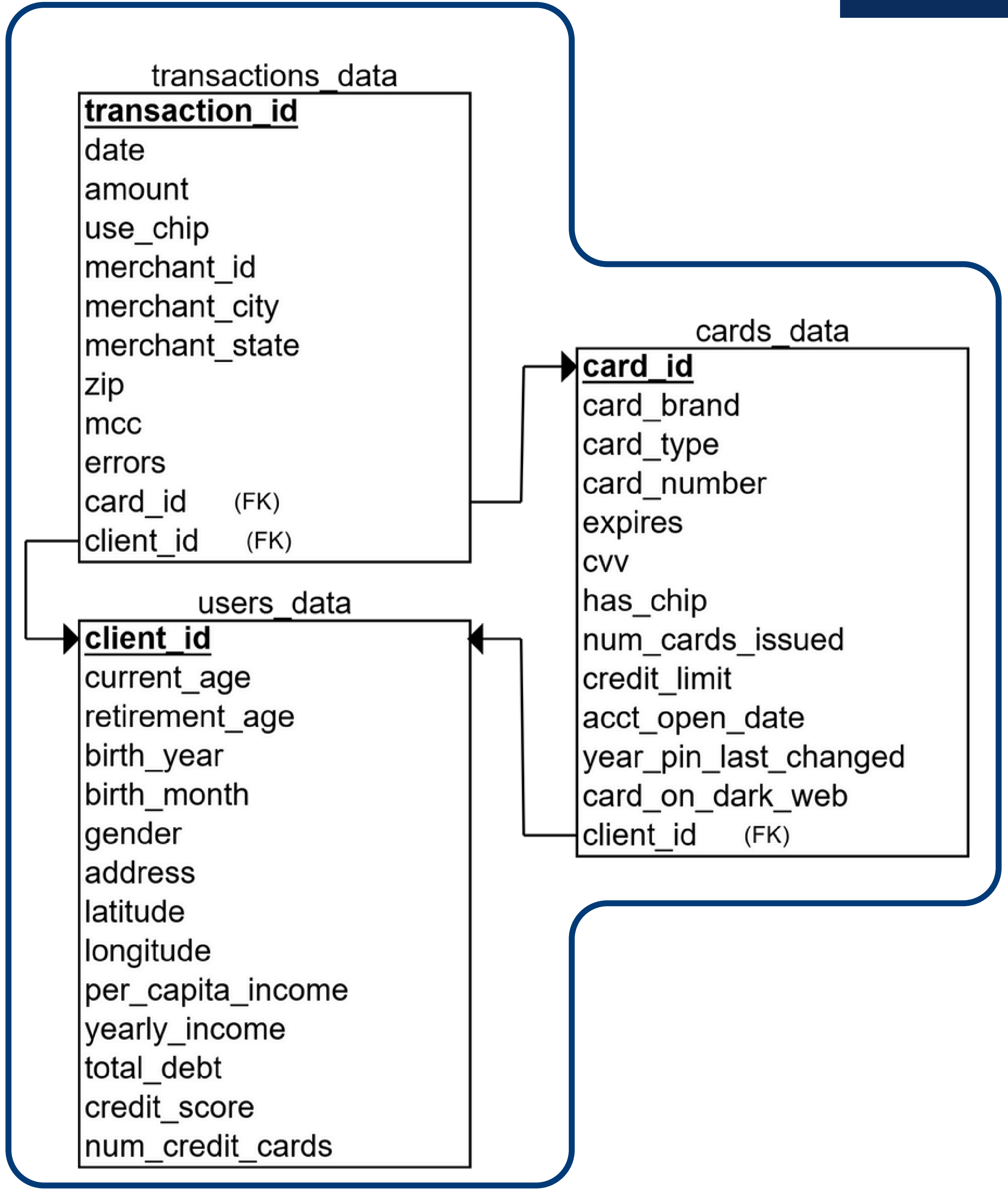
transactions_data

13,305,915 Rows

Column	Data Type	Description
id	INT	Unique identifier for each transaction.
date	STRING	Date and time of the transaction.
client_id	INT	User ID (foreign key linking to users_data.id).
card_id	INT	Card ID (foreign key linking to cards_data.id).
amount	STRING	Transaction amount (positive = purchase, negative = refund).
use_chip	STRING	Transaction method (Swipe Transaction, Chip Transaction, Online Transaction).
merchant_id	INT	Identifier of the merchant.
merchant_city	STRING	City where the merchant is located.
merchant_state	STRING	State where the merchant is located.
zip	FLOAT	Merchant ZIP code.
mcc	INT	Merchant Category Code (business classification).
errors	STRING	Error details if a transaction failed (mostly NULL).

Dataset Overview

Database Schema

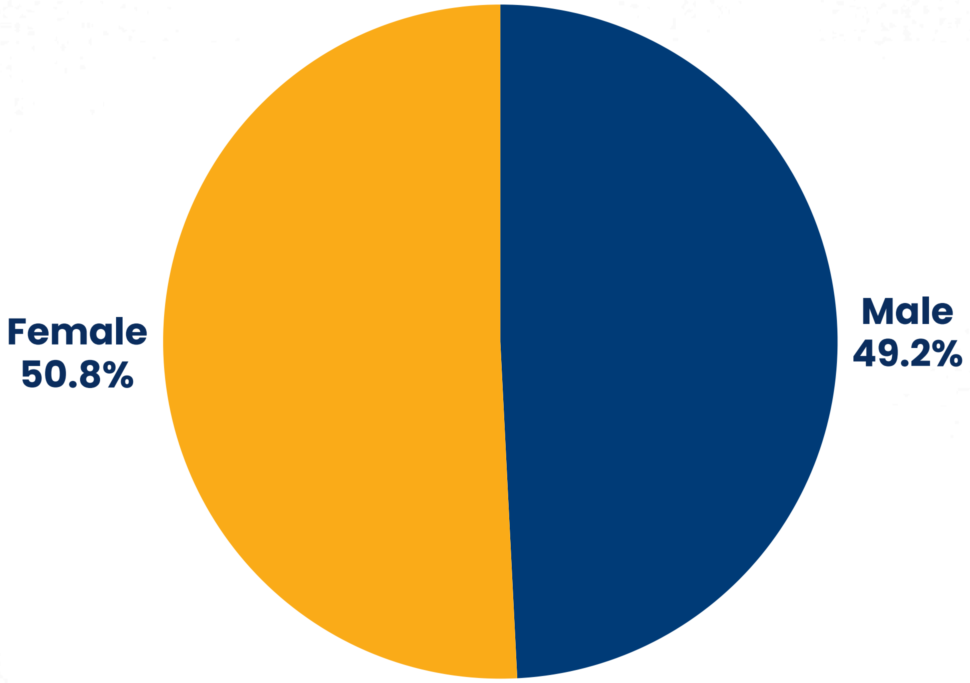


The schema shows the relationships between the tables:

- Each user can have multiple cards.
- Each card can be linked to multiple transactions.

User Demographic

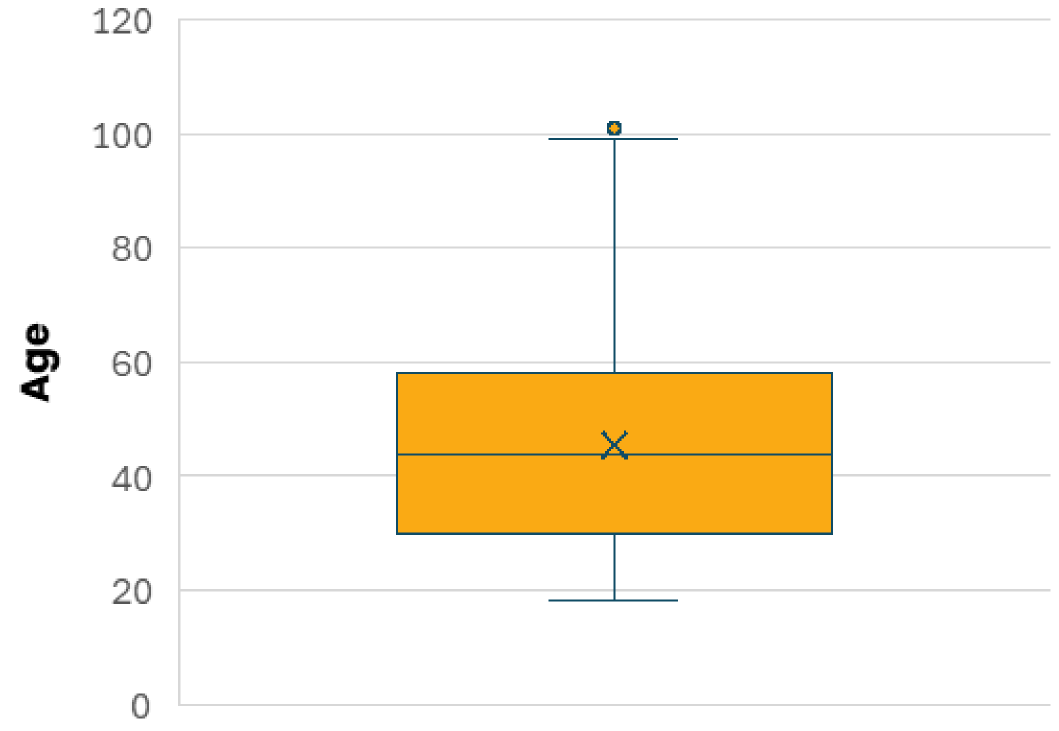
Gender ratio of each user



User Age Statistics

min_age	max_age	avg_age
18	101	45,39

User age distribution

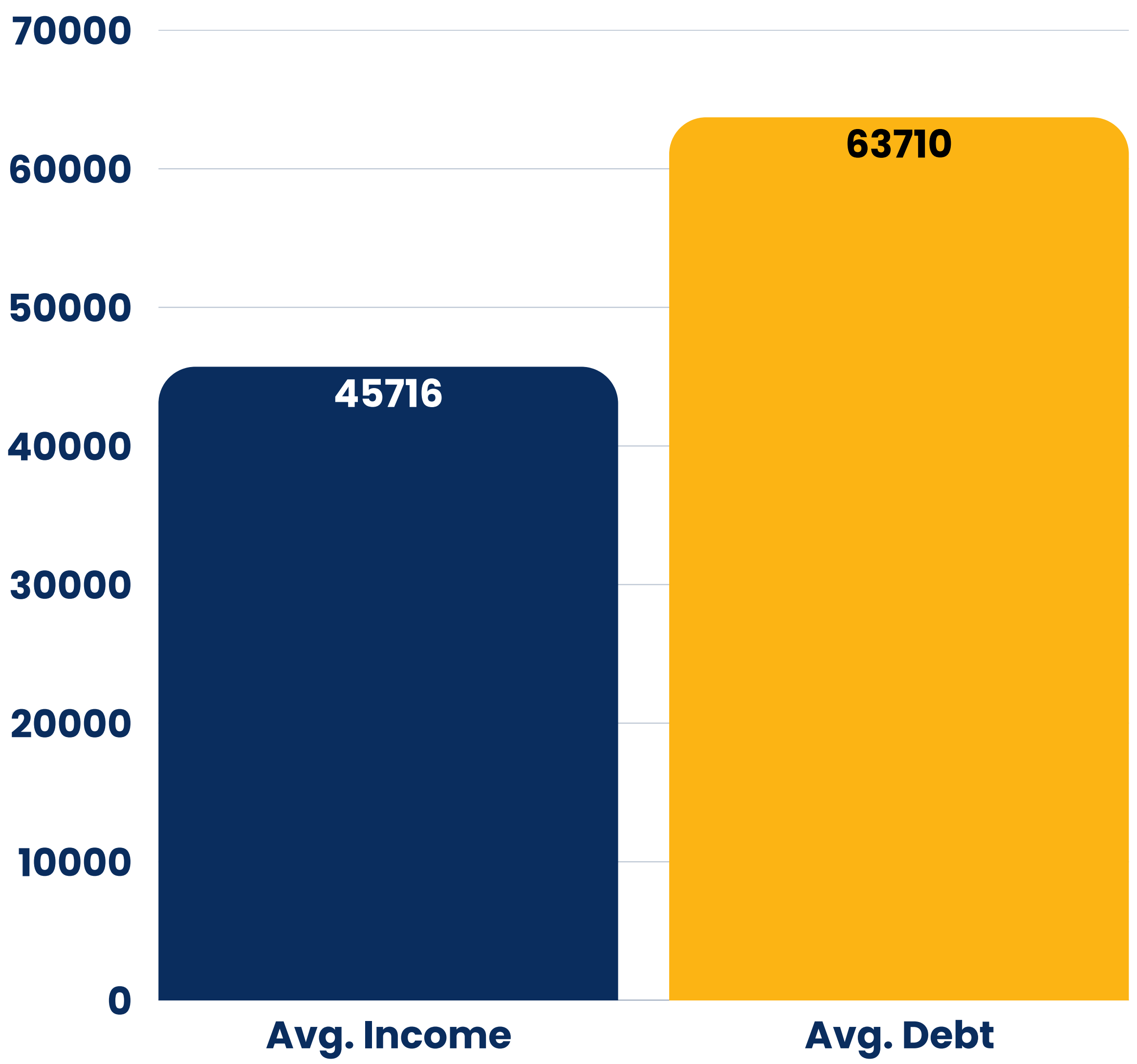


The majority of users are in the productive age group of 30–50, a potential segment for credit products. With this analysis, we can move on to lifestyle segmentation, allowing for more personalized product offerings.

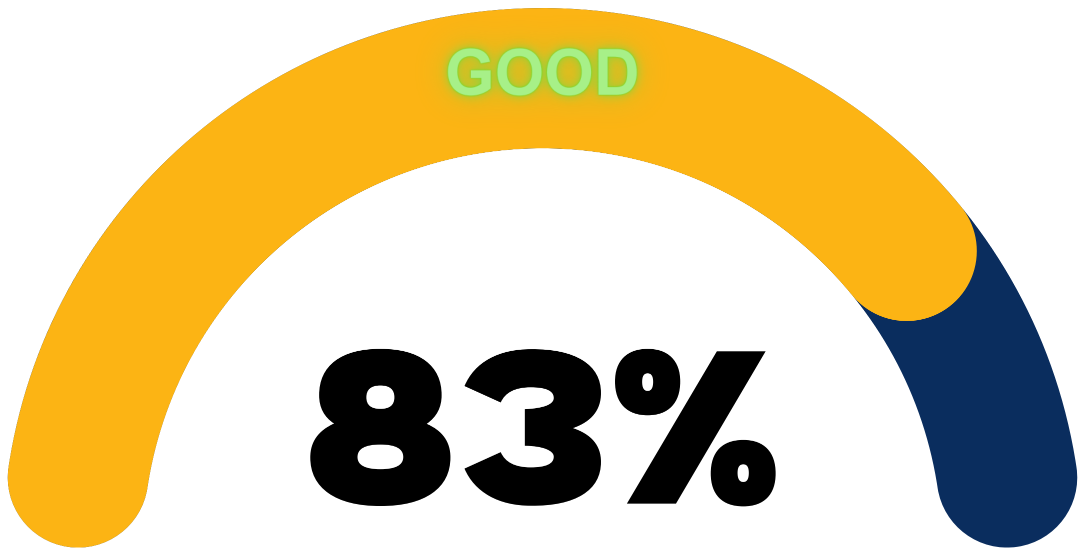
Financial *Profile*

mandiri sekuritas

Average income vs debt



Average credit score



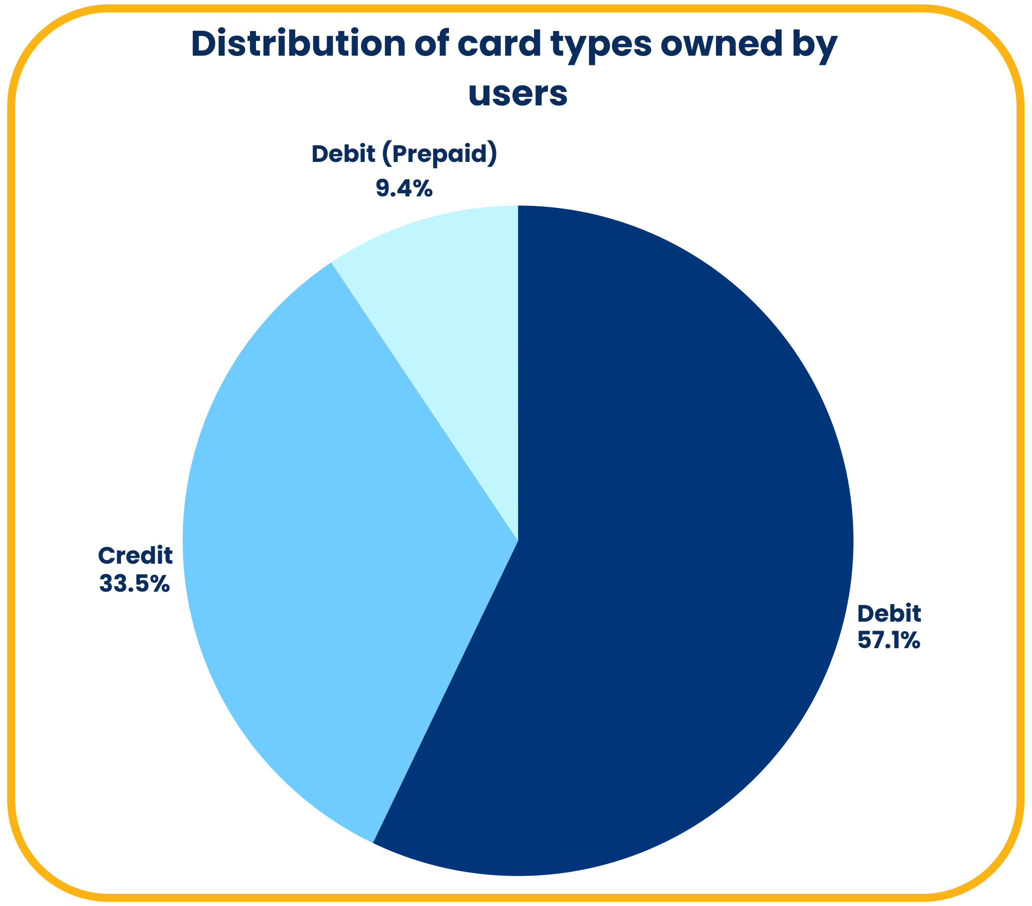
Avg. Credit Score	Max. Credit Score (by. Fair Isaac Corporation)
709.73	850

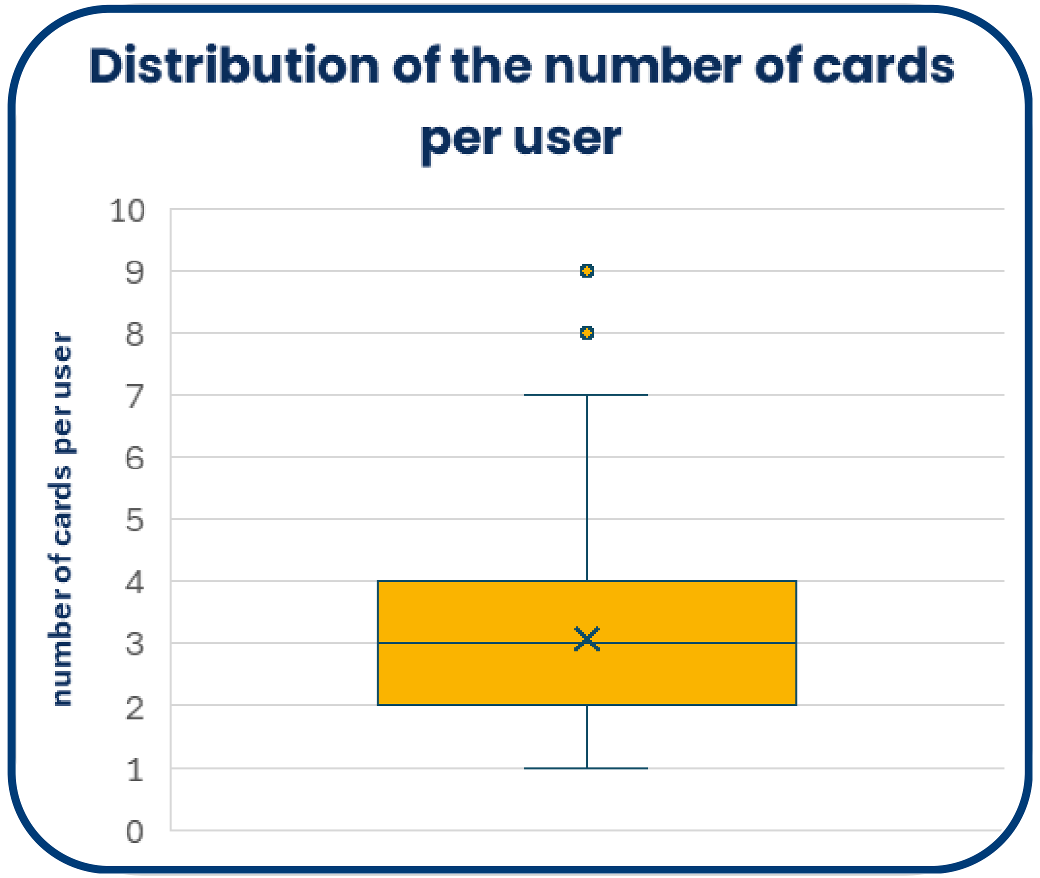
Top 10 users with the highest Debt-to-Income Ratio

No.	user_id	yearly_income	total_debt	debt_to_income_ratio
1	926	41403	206131	4,98
2	1091	53890	252106	4,68
3	1751	61581	241312	3,92
4	1319	48582	189348	3,90
5	1684	27861	108313	3,89
6	1088	34982	134631	3,85
7	1655	34856	127313	3,65
8	1005	39485	144101	3,65
9	920	28704	103891	3,62
10	74	41957	151027	3,60

The average user's income is lower than their total debt, but their credit score is still relatively good at around 710. From here, the analysis can be continued to measure the risk of default and develop a risk management strategy.

Cards *Usage*





CLV (Top 10 users with the highest value)

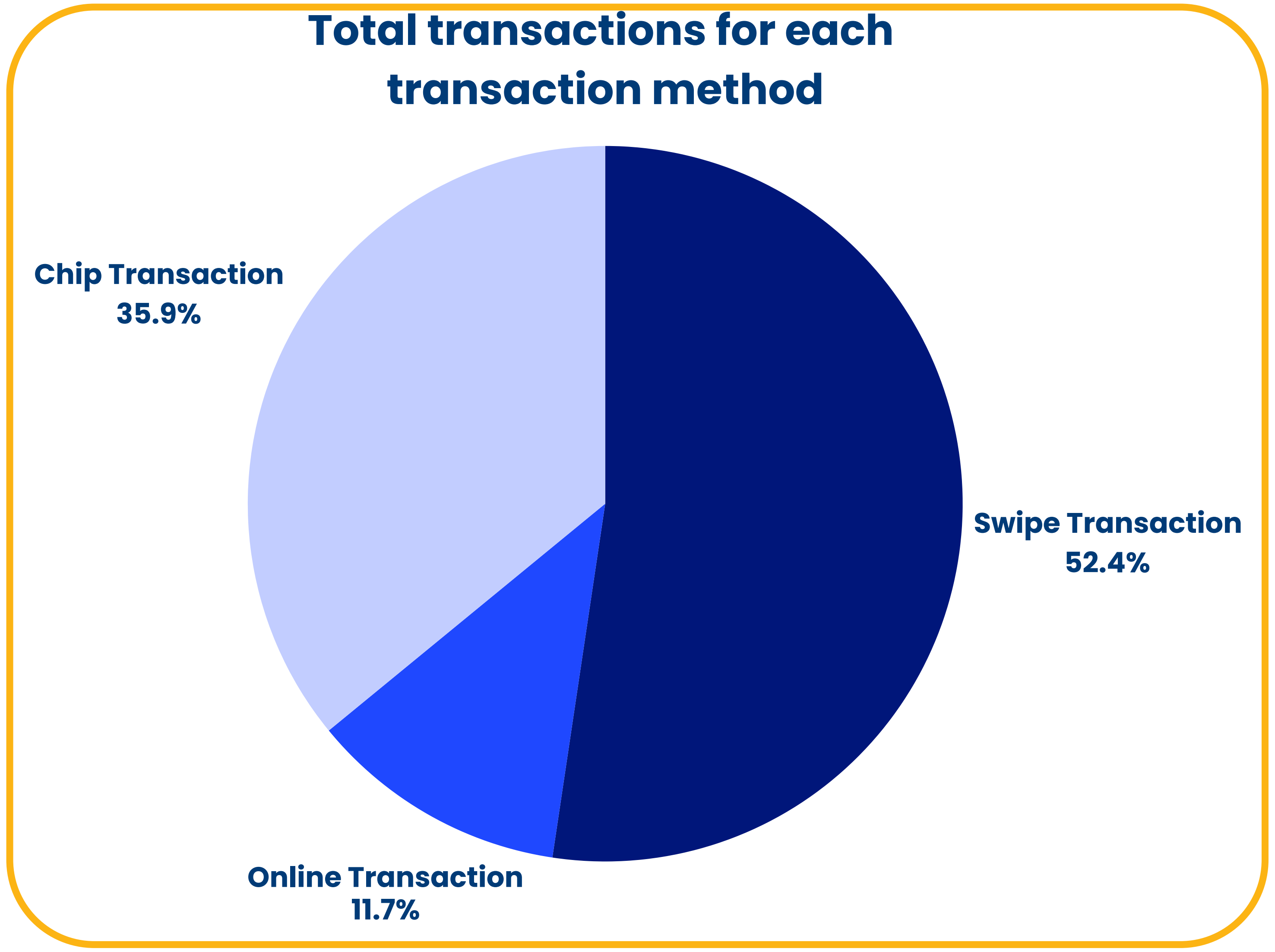
CLV = take average spending per transaction
×
number of cards owned.

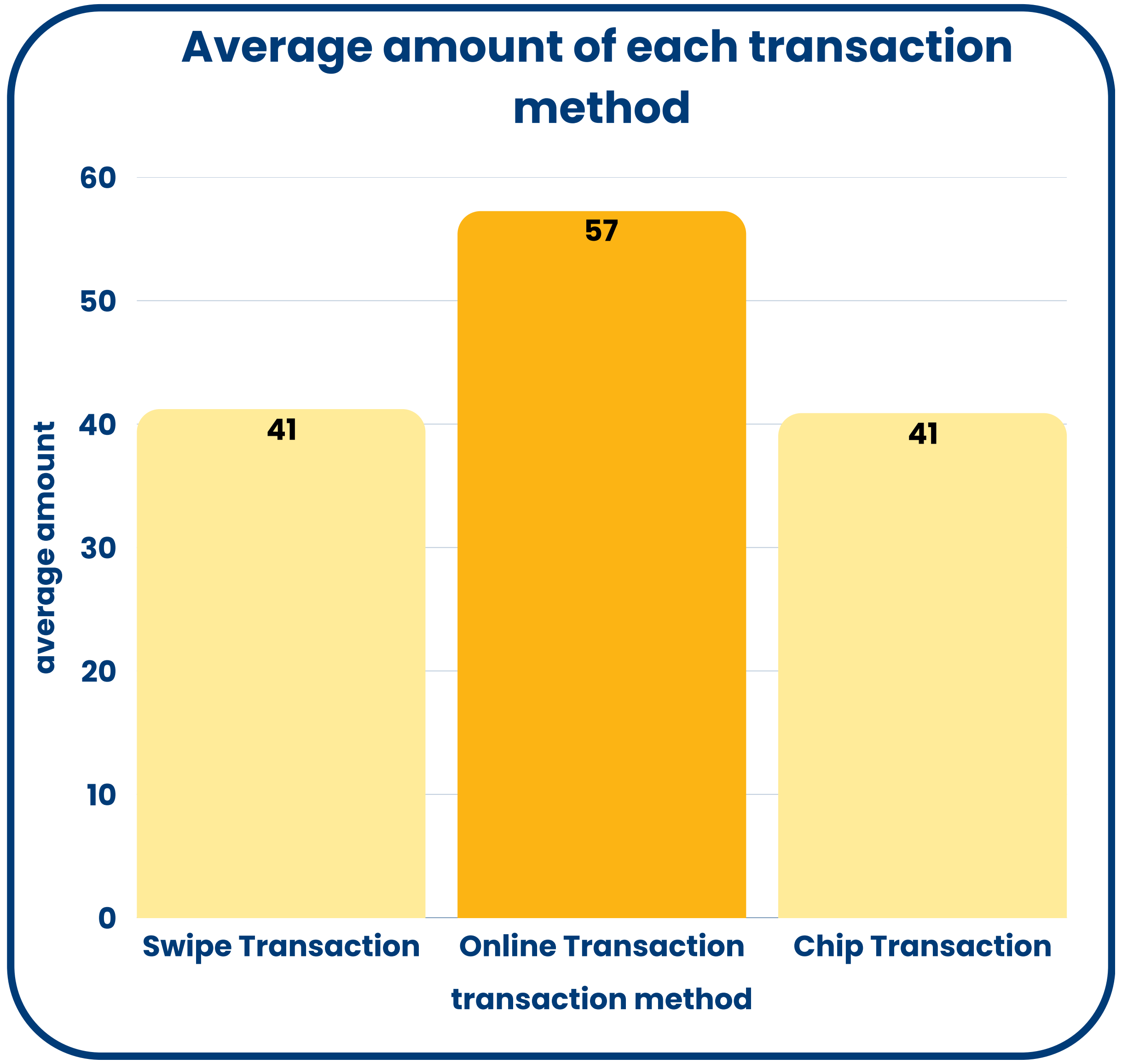
No	user_id	yearly_income	total_cards	estimated_clv
1	989	113514	8	1.020,12
2	840	88432	6	777,61
3	944	119308	6	621,34
4	77	51110	8	588,45
5	696	110570	8	572,79
6	1303	61937	7	553,93
7	1741	24960	9	551,42
8	1444	66885	7	535,56
9	1259	192269	6	508,97
10	708	249925	4	504,25

Most users use debit cards, but users with multiple cards tend to have a higher CLV. With this insight, we can further analyze the situation to identify cross-selling opportunities in the multi-card segment.

Transaction *Behavior*

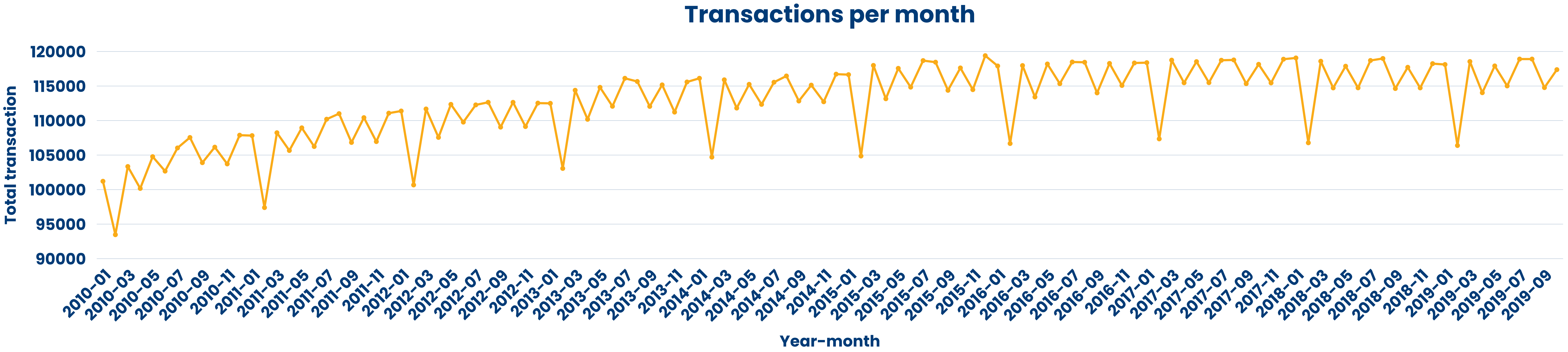
mandiri sekuritas





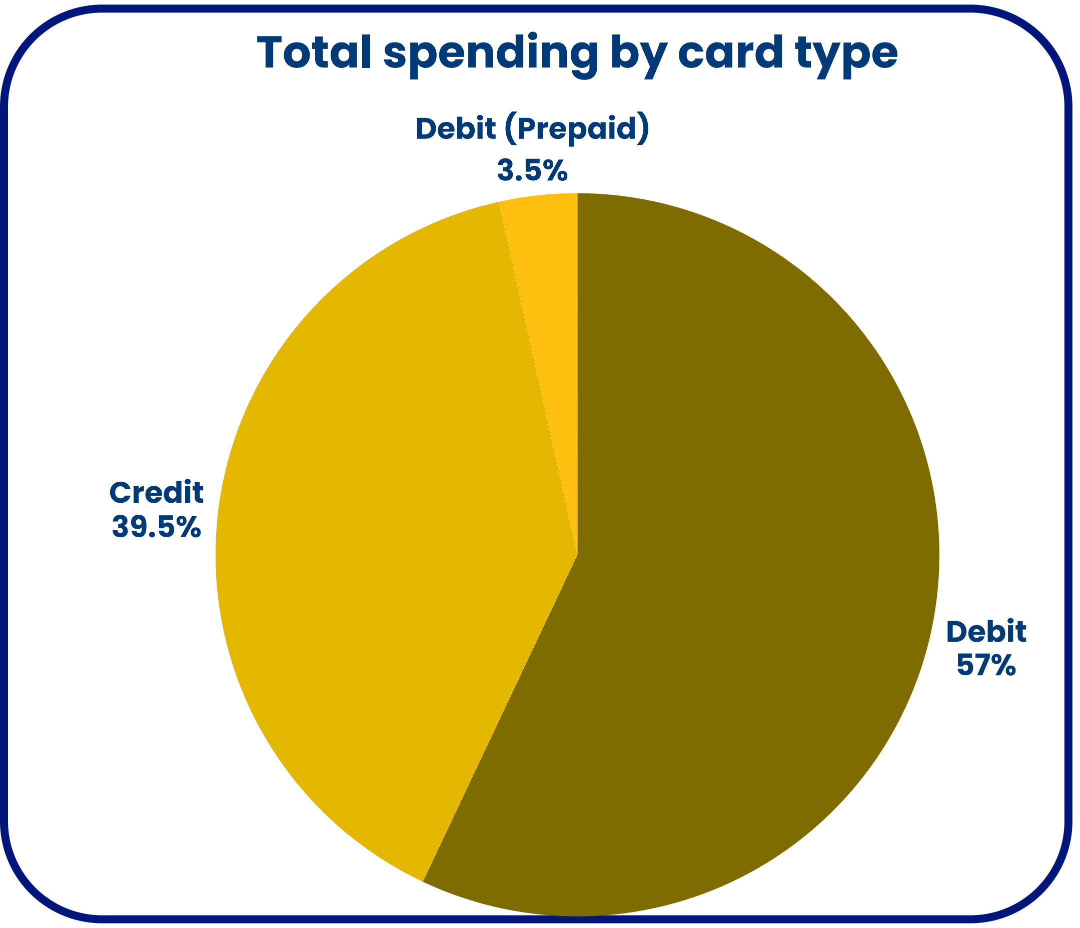
Online transactions are fewer in number, but the average amount is higher than other methods. From this pattern, we can further analyze which online shopping categories contribute the most.

Transaction *Behavior*

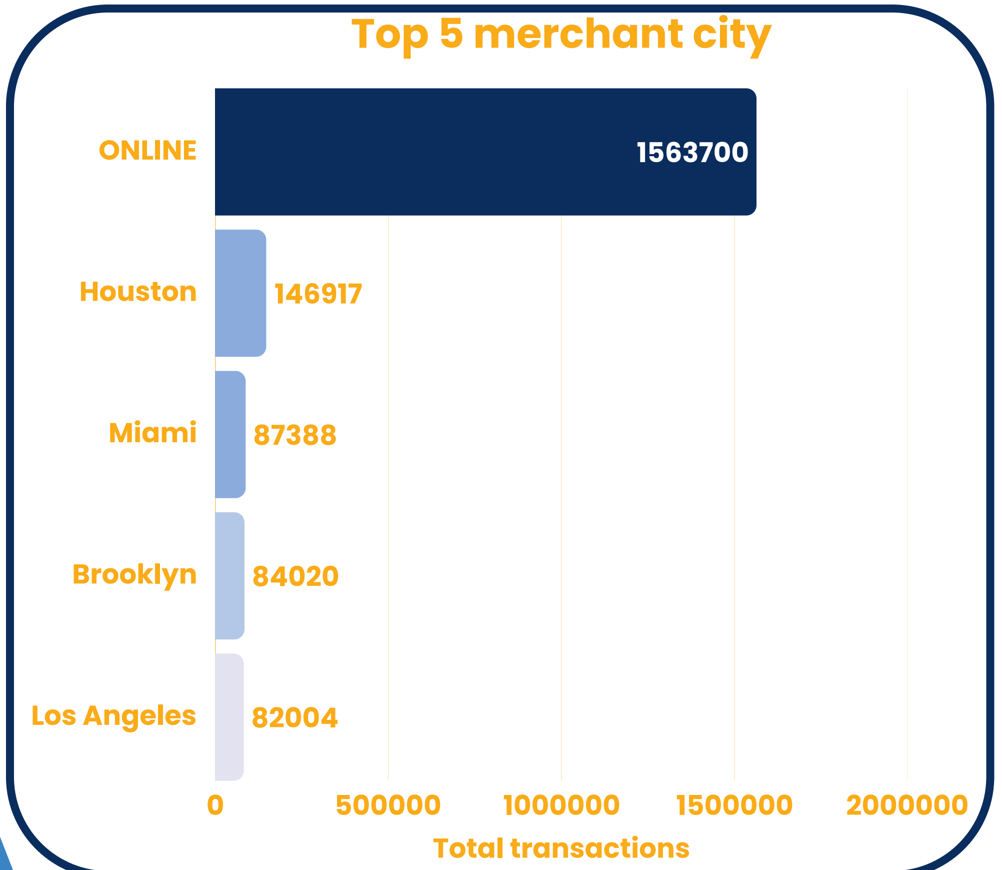


The number of transactions appears consistent and increases year over year. However, **there appears to be a seasonal pattern, particularly a decline in the second month of each year**. With this trend, we can proceed to forecasting analysis to predict spikes and anticipate declines in transactions during certain periods.

Transaction *Behavior*



The majority of users still rely on debit cards, accounting for 57%, while credit cards account for nearly 40% and offer significant potential for growth through promotions or rewards programs.



Location-wise, online transactions dominate significantly over brick-and-mortar cities like Houston, Miami, Brooklyn, and Los Angeles, demonstrating a shift in shopping behavior toward digital channels. This analysis allows us to develop strategies to optimize online services while simultaneously strengthening credit card penetration.

Top Clients

mandiri sekuritas

Top 10 high-value clients (highest total spending)

id	gender	current_age	total_spent
96	Female	69	2.445.773,2
1686	Female	54	2.167.880,9
1340	Male	53	2.039.921,2
840	Female	68	1.956.340,8
464	Male	36	1.882.901,4
490	Male	86	1.711.482,7
704	Female	51	1.635.022,0
285	Male	62	1.615.459,0
488	Male	34	1.611.114,4
1168	Male	51	1.590.822,8

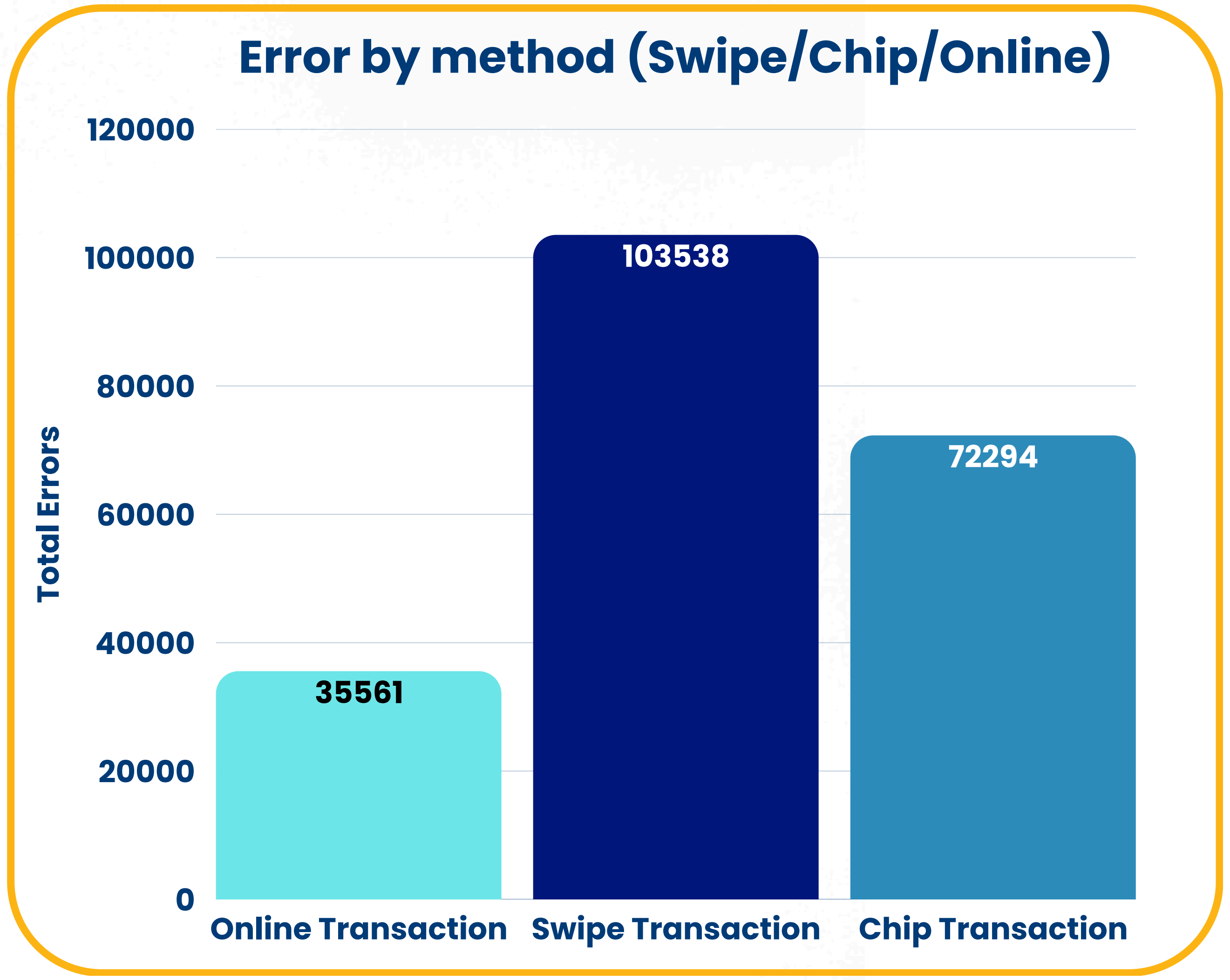
Ada sejumlah pengguna dengan pengeluaran sangat tinggi, bahkan ada yang lebih dari 2 juta. Dengan temuan ini, kita bisa melanjutkan analisis untuk memprediksi risiko churn serta merancang program loyalitas khusus bagi mereka.

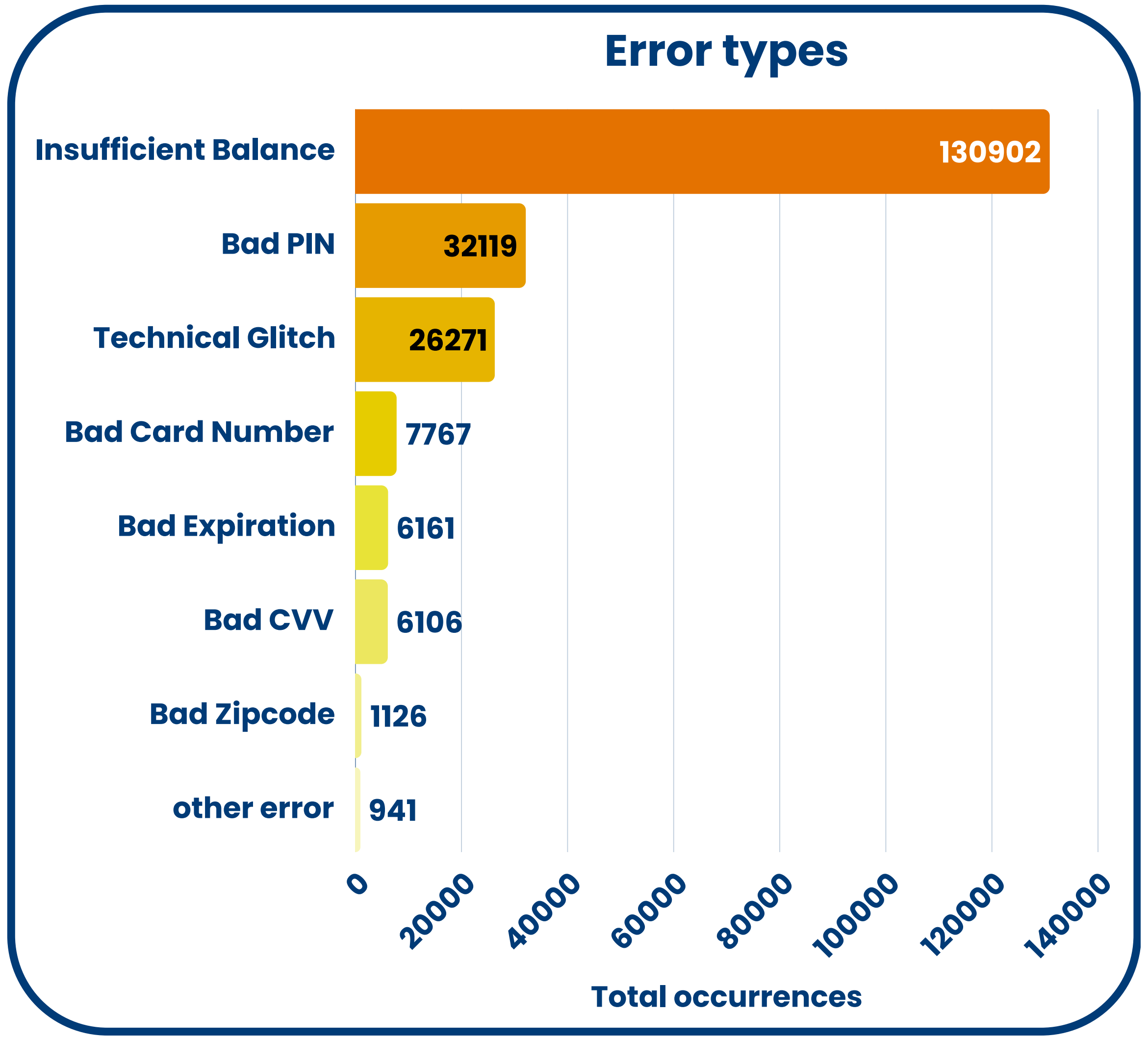
Fraud *Monitoring*

There are no transactions from dark web cards.

the dataset is relatively safe, **but still requires regular monitoring.**

Error & Reliability



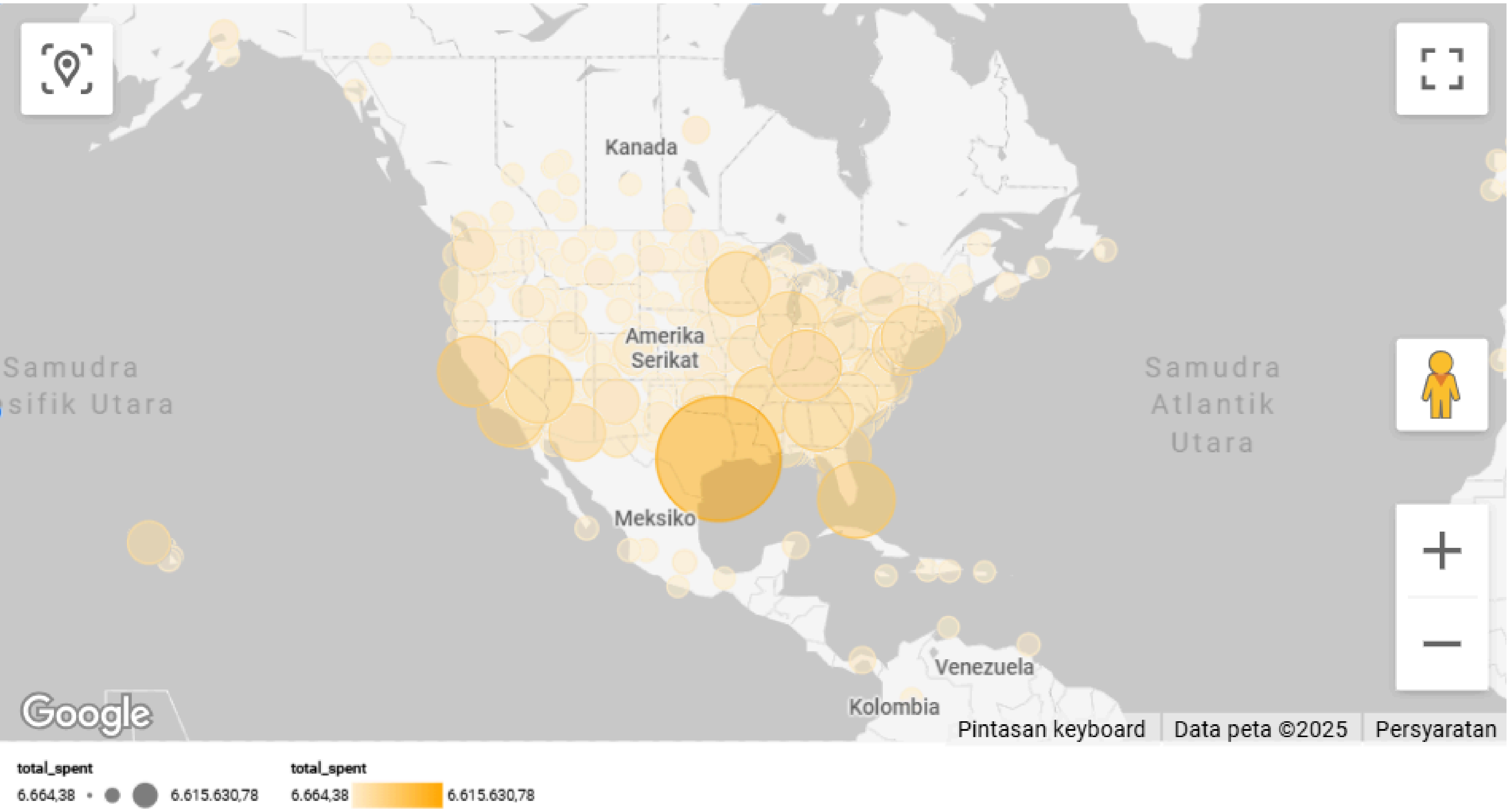


Error analysis shows that swipe transactions generate the majority of technical issues, including glitches and incorrect data entries.

This insight supports a business recommendation to encourage chip or contactless transactions, improving reliability and customer trust

Geographic *Spending*

mandiri sekuritas



merchant_city	merchant_state	total_spent
ONLINE	-	88.896.665,1
Houston	TX	6.608.055,3
Miami	FL	3.480.604,4
San Francisco	CA	3.184.666,9
Memphis	TN	3.078.223,2
Las Vegas	NV	2.973.865,8
Los Angeles	CA	2.965.266,4
Atlanta	GA	2.950.725,5
Minneapolis	MN	2.815.878,9
Orlando	FL	2.722.901,5

Online transactions top the list, while major cities like Houston and Miami also contribute significantly. With this analysis, we can move on to mapping regional trends to determine where to base expansion strategies.

Business Recommendations

Recommendation	Reason (Based on Data)
Focus on productive age users (30–50 years)	Majority of users are in this age group → strong potential for credit products.
Boost credit card usage with promos & rewards	Debit dominates 57%, but credit already holds 39.5% → clear growth opportunity.
Strengthen online transaction channels	Online spending is the highest (≈ $89M) compared to physical cities.
Prioritize high-value clients	Some users spend > $2.4M → require loyalty programs & personalized services.
Enhance error & fraud monitoring	Errors (insufficient balance, wrong PIN) remain high, though fraud risk is low.

Conclusion

mandiri sekuritas

The analysis shows that most users are in the productive age group, with relatively healthy financial profiles despite a high debt-to-income ratio. Debit cards remain dominant, but credit cards and online transactions have strong growth potential. High-value clients are a crucial segment that requires special attention, while system errors still need to be addressed to ensure user satisfaction and trust.

To achieve more strategic business goals, further analysis is needed, such as:

- User risk assessment to predict the probability of default for each customer.
- User segmentation for more personalized product offerings.
- Analysis of partnerships with top-performing merchants to expand collaboration.
- Churn risk prediction to retain high-value clients.
- Transaction forecasting to anticipate growth and seasonal fluctuations.

By combining current insights with advanced analytics, companies can strengthen decision-making, drive revenue growth, and enhance long-term customer loyalty.

mandiri sekuritas

Thank You

For Your Attention!

https://www.linkedin.com/in/rizalrahmanrizkika/

rizalrahmanr@gmail.com

+62 895-8042-20277