DATA ANALYST TECHNICAL TEST

BANKING USER BEHAVIOR ANALYSIS

Prepared By: Nabila Yasmine Az Zahra

SQL I README I DASHBOARD I PRESENTATION

Task Requirement

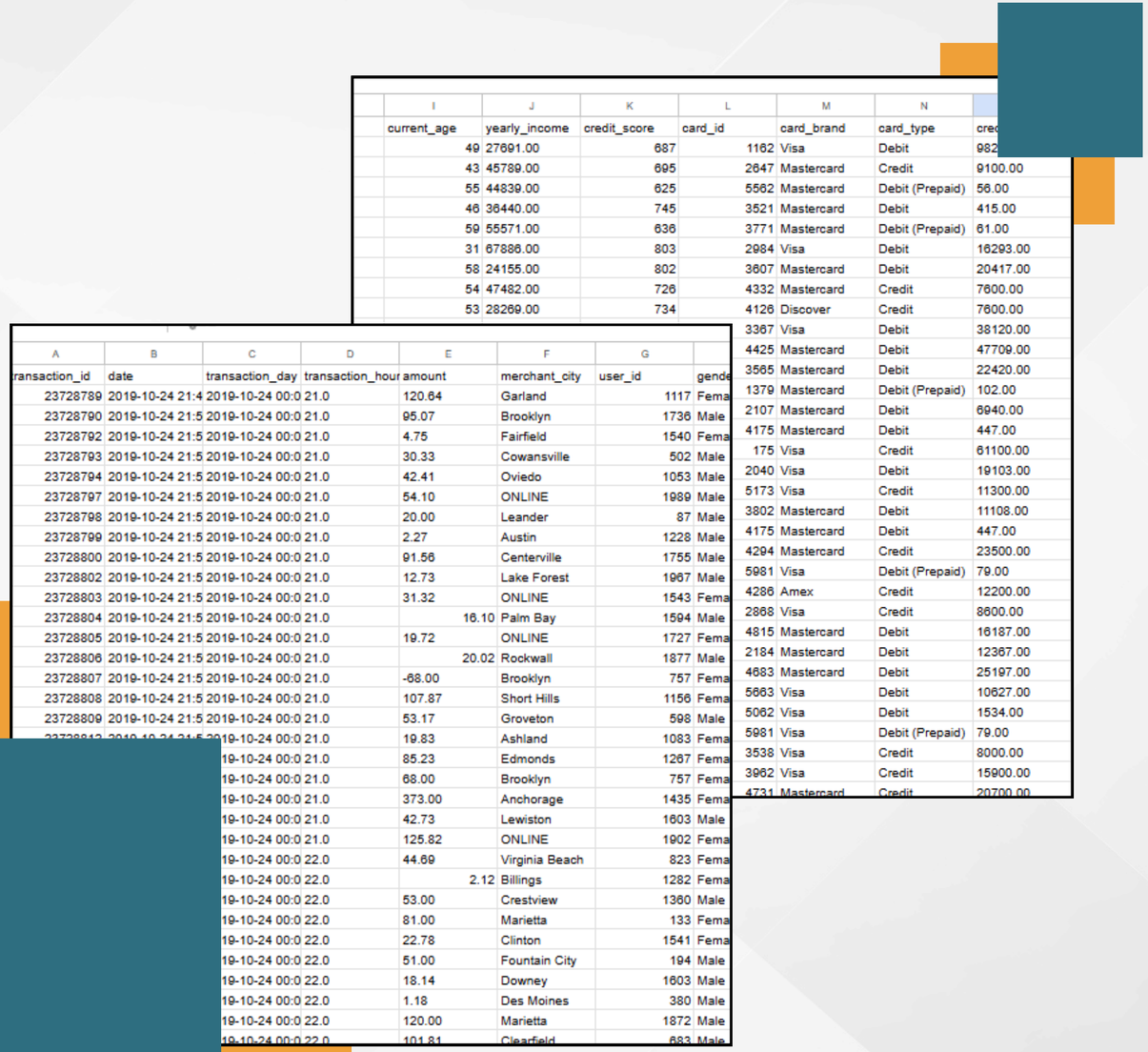
Objective: **Analyze user behavior from public banking datasets.**

01 **SQL Queries**
Exploration & Analysis

02 **README**
How to run code

03 **DASHBOARD**
Looker Studio

04 **Presentation**
Insights and Findings



	I	J	K	L	M	N	
	current_age	yearly_income	credit_score	card_id	card_brand	card_type	cred
	49	27691.00	687	1162	Visa	Debit	982
	43	45789.00	695	2647	Mastercard	Credit	9100.00
	55	44839.00	625	5562	Mastercard	Debit (Prepaid)	56.00
	46	36440.00	745	3521	Mastercard	Debit	415.00
	59	55571.00	636	3771	Mastercard	Debit (Prepaid)	61.00
	31	67886.00	803	2984	Visa	Debit	16293.00
	58	24155.00	802	3607	Mastercard	Debit	20417.00
	54	47482.00	726	4332	Mastercard	Credit	7600.00
	53	28269.00	734	4126	Discover	Credit	7600.00
				3367	Visa	Debit	38120.00
				4425	Mastercard	Debit	47709.00
				3565	Mastercard	Debit	22420.00
				1379	Mastercard	Debit (Prepaid)	102.00
				2107	Mastercard	Debit	6940.00
				4175	Mastercard	Debit	447.00
				175	Visa	Credit	61100.00
				2040	Visa	Debit	19103.00
				5173	Visa	Credit	11300.00
				3802	Mastercard	Debit	11108.00
				4175	Mastercard	Debit	447.00
				4294	Mastercard	Credit	23500.00
				5981	Visa	Debit (Prepaid)	79.00
				4286	Amex	Credit	12200.00
				2868	Visa	Credit	8600.00
				4815	Mastercard	Debit	16187.00
				2184	Mastercard	Debit	12367.00
				4683	Mastercard	Debit	25197.00
				5663	Visa	Debit	10627.00
				5062	Visa	Debit	1534.00
				5981	Visa	Debit (Prepaid)	79.00
				3538	Visa	Credit	8000.00
				3962	Visa	Credit	15900.00
				4731	Mastercard	Credit	20700.00

	A	B	C	D	E	F	G	
	ransaction_id	date	transaction_day	transaction_hour	amount	merchant_city	user_id	gende
	23728789	2019-10-24 21:4	2019-10-24 00:0	21.0	120.64	Garland	1117	Fema
	23728790	2019-10-24 21:5	2019-10-24 00:0	21.0	95.07	Brooklyn	1736	Male
	23728792	2019-10-24 21:5	2019-10-24 00:0	21.0	4.75	Fairfield	1540	Fema
	23728793	2019-10-24 21:5	2019-10-24 00:0	21.0	30.33	Cowansville	502	Male
	23728794	2019-10-24 21:5	2019-10-24 00:0	21.0	42.41	Oviedo	1053	Male
	23728797	2019-10-24 21:5	2019-10-24 00:0	21.0	54.10	ONLINE	1989	Male
	23728798	2019-10-24 21:5	2019-10-24 00:0	21.0	20.00	Leander	87	Male
	23728799	2019-10-24 21:5	2019-10-24 00:0	21.0	2.27	Austin	1228	Male
	23728800	2019-10-24 21:5	2019-10-24 00:0	21.0	91.56	Centerville	1755	Male
	23728802	2019-10-24 21:5	2019-10-24 00:0	21.0	12.73	Lake Forest	1967	Male
	23728803	2019-10-24 21:5	2019-10-24 00:0	21.0	31.32	ONLINE	1543	Fema
	23728804	2019-10-24 21:5	2019-10-24 00:0	21.0	16.10	Palm Bay	1594	Male
	23728805	2019-10-24 21:5	2019-10-24 00:0	21.0	19.72	ONLINE	1727	Fema
	23728806	2019-10-24 21:5	2019-10-24 00:0	21.0	20.02	Rockwall	1877	Male
	23728807	2019-10-24 21:5	2019-10-24 00:0	21.0	-68.00	Brooklyn	757	Fema
	23728808	2019-10-24 21:5	2019-10-24 00:0	21.0	107.87	Short Hills	1156	Fema
	23728809	2019-10-24 21:5	2019-10-24 00:0	21.0	53.17	Groveton	598	Male
			19-10-24 00:0	21.0	19.83	Ashland	1083	Fema
			19-10-24 00:0	21.0	85.23	Edmonds	1267	Fema
			19-10-24 00:0	21.0	68.00	Brooklyn	757	Fema
			19-10-24 00:0	21.0	373.00	Anchorage	1435	Fema
			19-10-24 00:0	21.0	42.73	Lewiston	1603	Male
			19-10-24 00:0	21.0	125.82	ONLINE	1902	Fema
			19-10-24 00:0	22.0	44.69	Virginia Beach	823	Fema
			19-10-24 00:0	22.0	2.12	Billings	1282	Fema
			19-10-24 00:0	22.0	53.00	Crestview	1360	Male
			19-10-24 00:0	22.0	81.00	Marietta	133	Fema
			19-10-24 00:0	22.0	22.78	Clinton	1541	Fema
			19-10-24 00:0	22.0	51.00	Fountain City	194	Male
			19-10-24 00:0	22.0	18.14	Downey	1603	Male
			19-10-24 00:0	22.0	1.18	Des Moines	380	Male
			19-10-24 00:0	22.0	120.00	Marietta	1872	Male
			19-10-24 00:0	22.0	101.81	Clearfield	683	Male

Dataset Overview

This dataset contains transactional records at the user level. Each row represents one activity, identified uniquely by a transaction ID. The most important fields we'll focus on are the date of the transaction, its type, and the value because these are the key drivers for understanding user behavior trends.

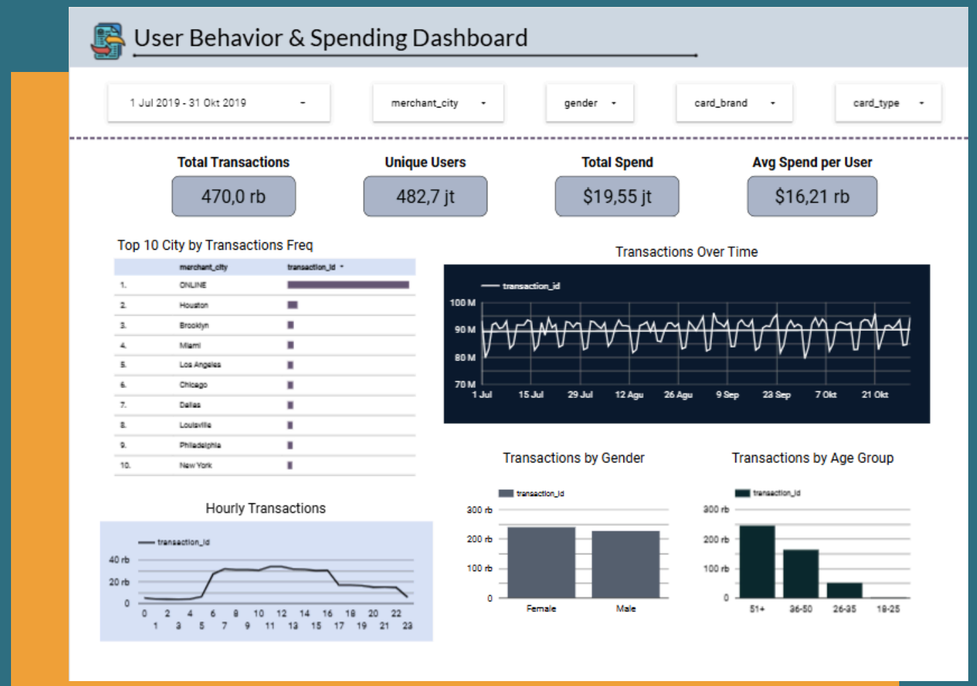
SQL Queries & Highlights

In this step, I applied SQL queries to prepare the dataset. The process included cleaning and transforming raw data, making sure date formats are consistent, and then aggregating transactions both at the user level and monthly. I also created new metrics, such as how often users transact and their average transaction value. Finally, I exported a clean fact table that serves as the basis for the Looker Studio dashboard.

```sql
-- ================
--  All Table EDA
-- ================

-- Transaction volume and amount by gender
SELECT
    u.gender,
    COUNT(t.id) AS total_transactions,
    SUM(t.amount) AS total_amount
FROM mtest.transactions t
JOIN mtest.users u ON t.client_id = u.id
GROUP BY u.gender;

-- Transaction volume and amount by age group
SELECT
    CASE
        WHEN u.current_age < 20 THEN 'Teen'
        WHEN u.current_age BETWEEN 20 AND 30 THEN 'Young Adult'
        WHEN u.current_age BETWEEN 31 AND 45 THEN 'Adult'
        ELSE 'Senior'
    END AS age_group,
    COUNT(t.id) AS total_tx,
    SUM(t.amount) AS total_amount
FROM mtest.transactions t
JOIN mtest.users u ON t.client_id = u.id
GROUP BY age_group
ORDER BY age_group;

-- Transaction analysis by card type
SELECT
    c.card_type,
    COUNT(t.id) AS total_transactions,
    SUM(t.amount) AS total_amount
FROM mtest.transactions t
.cards c ON t.card_id = c.id
.card_type
otal_amount DESC;
```

Dashboard Summary

The dashboard provides an overview of user activity, showing trends across demographics, time periods, and locations. It highlights transaction volumes, user distribution by age and gender, as well as performance across cities and sales channels. This summary enables a quick understanding of overall behavioral patterns.



Key Insights

Based on transaction data from July 1, 2019 to October 31, 2019, several key patterns in user behavior and spending can be observed from the dashboard.

CITY DISTRIBUTION

Transactions are highly concentrated in major cities, with the top contributors including Online channels, Houston, Brooklyn, and Miami. This suggests that both digital and urban markets play a central role in driving overall transaction volume.

USER DEMOGRAPHICS

Older users (aged 51 and above) dominate transaction volumes compared to younger age groups, highlighting their stronger purchasing activity. Meanwhile, transactions between male and female users are fairly balanced, showing no significant gender gap in engagement.

HOURLY TRANSACTIONS

Transaction activity follows a clear daily rhythm, peaking during the daytime hours and dropping significantly overnight. This reflects typical consumer spending habits tied to daily routines.

Recommendations

01

Focus on top cities & online channels

High transaction volume in certain cities and online indicates where digital banking services and merchant partnerships are most active. The bank can strengthen digital payment solutions and merchant network coverage in these key areas to capture more activity.

02

User Demographics (51+ Age Group)

Since older customers are the most active, the bank should design senior-friendly digital services (easy UI, strong customer support) and loyalty or rewards programs tailored for this age segment. At the same time, banks shouldn't neglect younger users, but instead focus on onboarding campaigns to grow their usage.

03

Hourly Transactions (10 AM–4 PM Peak)

Transaction peaks during business hours highlight when users rely most on banking services. Banks can optimize system capacity and service availability during these hours and consider timed promotions or alerts aligned with peak activity.

THANK YOU