Eka Pramudianzah

Data Analyst Candidate

PT Mandiri Sekuritas -
PT Indocyber Global Teknologi

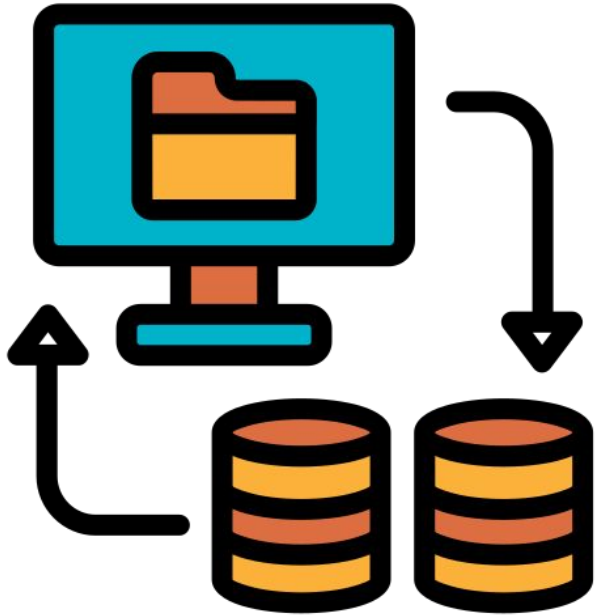
PROFESSIONAL EXPERIENCE

- PT Sinar Roda Utama, Contract
  (November 2024 - Present)
- Data Scientist, Self-Employed Freelance
  (June 2021 - July 2024)
- Danone Indonesia, Contract
  (April – July 2024)
- PT Multi Bintang Indonesia Tbk, Contract
  (March - December 2023)

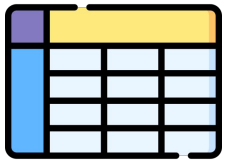
Table of Contents

01

Dataset Background

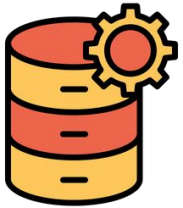


Dataset Introduction



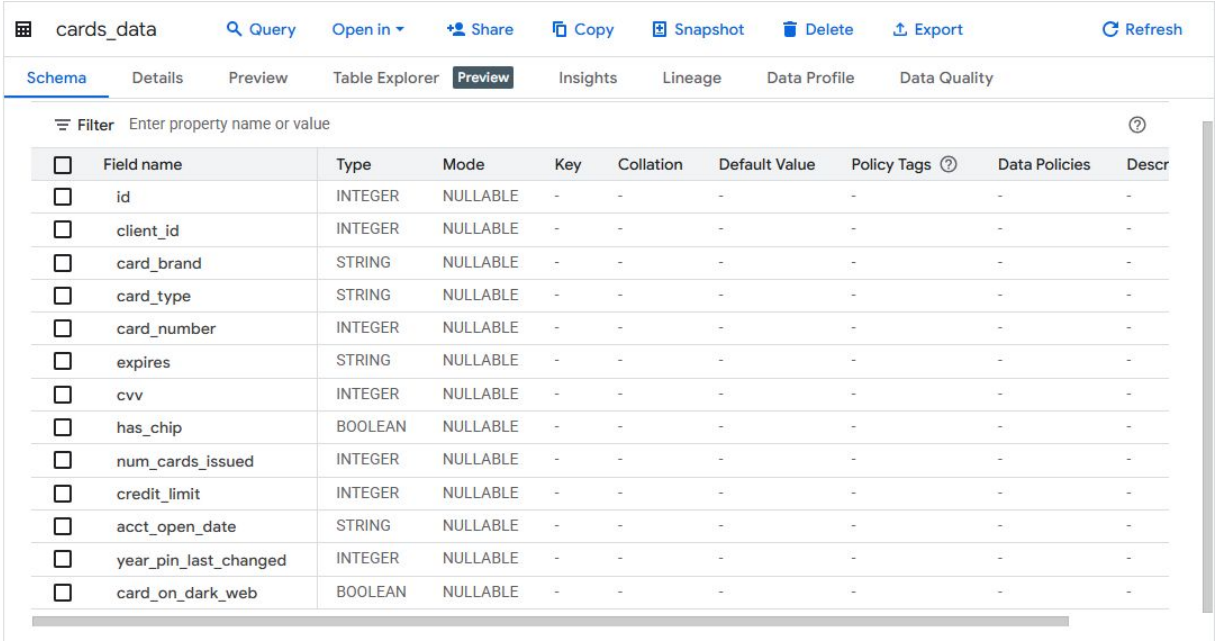
Three Datasets File that will be Used and Analyzed

The Format of Datasets File is CSV (Comma Separated Value)



The Name of 3 Datasets File is “cards_data.csv”, “transactions_data.csv”, and “users_data.csv”

First Dataset Structure (cards_data.csv)



Field name	Type	Mode	Key	Collation	Default Value	Policy Tags	Data Policies	Descr
id	INTEGER	NULLABLE	-	-	-	-	-	-
client_id	INTEGER	NULLABLE	-	-	-	-	-	-
card_brand	STRING	NULLABLE	-	-	-	-	-	-
card_type	STRING	NULLABLE	-	-	-	-	-	-
card_number	INTEGER	NULLABLE	-	-	-	-	-	-
expires	STRING	NULLABLE	-	-	-	-	-	-
cvv	INTEGER	NULLABLE	-	-	-	-	-	-
has_chip	BOOLEAN	NULLABLE	-	-	-	-	-	-
num_cards_issued	INTEGER	NULLABLE	-	-	-	-	-	-
credit_limit	INTEGER	NULLABLE	-	-	-	-	-	-
acct_open_date	STRING	NULLABLE	-	-	-	-	-	-
year_pin_last_changed	INTEGER	NULLABLE	-	-	-	-	-	-
card_on_dark_web	BOOLEAN	NULLABLE	-	-	-	-	-	-

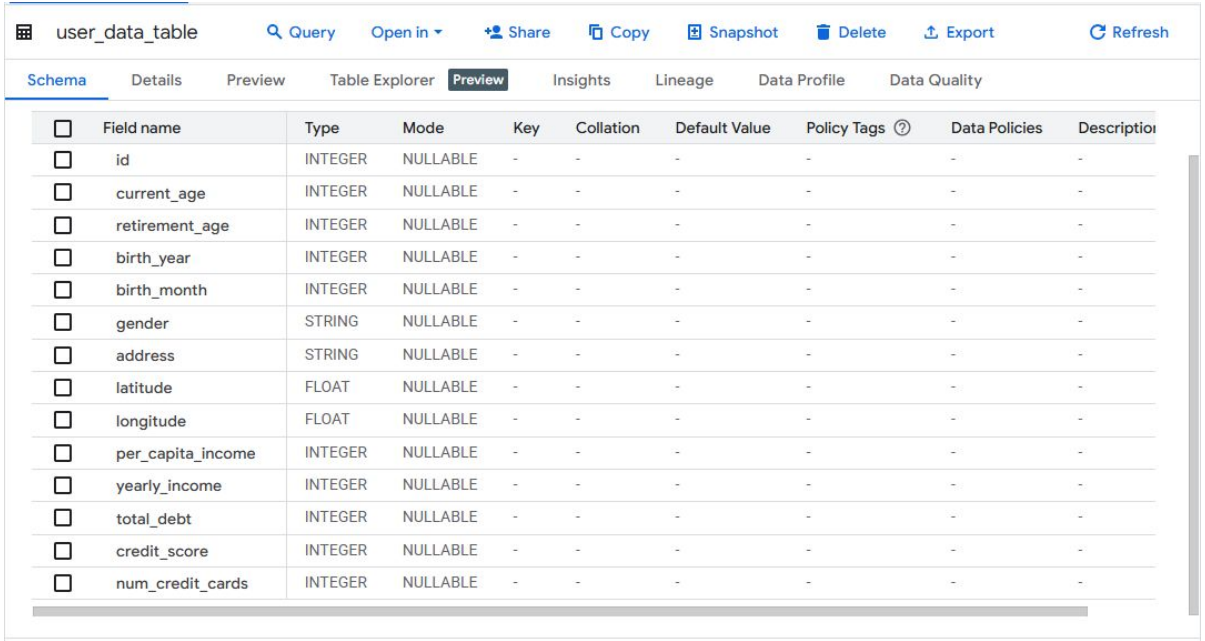
Total Columns “cards_data.csv” is 13 data columns

Second Dataset Structure (transactions_data.csv)

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 13305915 entries, 0 to 13305914
Data columns (total 12 columns):
 #   Column          Dtype
---  ------          -----
 0   id              int64
 1   date            object
 2   client_id       int64
 3   card_id         int64
 4   amount          object
 5   use_chip        object
 6   merchant_id     int64
 7   merchant_city   object
 8   merchant_state  object
 9   zip             float64
 10  mcc             int64
 11  errors          object
dtypes: float64(1), int64(5), object(6)
memory usage: 1.2+ GB
```

Total Columns
“transactions_data.csv” is
14 data columns

Third Dataset Structure (users_data.csv)



Field name	Type	Mode	Key	Collation	Default Value	Policy Tags	Data Policies	Descriptio
id	INTEGER	NULLABLE	-	-	-	-	-	-
current_age	INTEGER	NULLABLE	-	-	-	-	-	-
retirement_age	INTEGER	NULLABLE	-	-	-	-	-	-
birth_year	INTEGER	NULLABLE	-	-	-	-	-	-
birth_month	INTEGER	NULLABLE	-	-	-	-	-	-
gender	STRING	NULLABLE	-	-	-	-	-	-
address	STRING	NULLABLE	-	-	-	-	-	-
latitude	FLOAT	NULLABLE	-	-	-	-	-	-
longitude	FLOAT	NULLABLE	-	-	-	-	-	-
per_capita_income	INTEGER	NULLABLE	-	-	-	-	-	-
yearly_income	INTEGER	NULLABLE	-	-	-	-	-	-
total_debt	INTEGER	NULLABLE	-	-	-	-	-	-
credit_score	INTEGER	NULLABLE	-	-	-	-	-	-
num_credit_cards	INTEGER	NULLABLE	-	-	-	-	-	-

Total Columns “users_data.csv” is 14 data columns

02

Dataset Preparation

Preparation before Analysis in Python Notebook

01

Check for Missing and Duplicate Values from the 3 Dataset Files

02

Change the Data Type of the "date" column from String to Datetime, then Extract the Year Data and Create a New Column

03

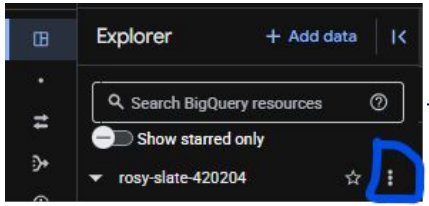
Delete the "errors" and "zip" columns, then Remove Missing Values in the "merchant_state" Column from the "transactions_data.csv" Dataset File

04

Select only the 2019 Transaction Data, then Extract it to a New CSV Dataset File and Upload it to the Google BigQuery Console

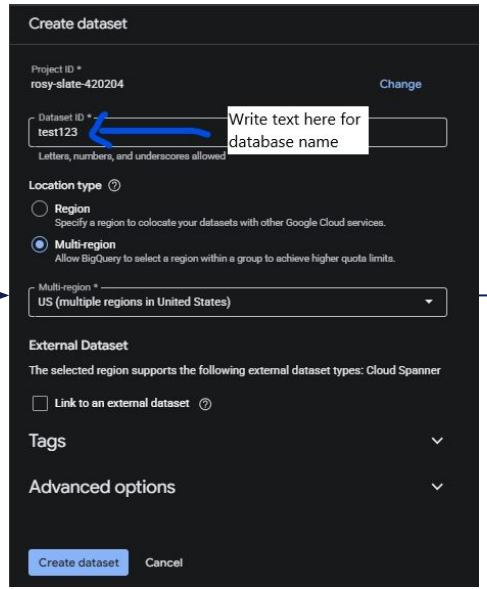
Preparation before Analysis in Google BigQuery (1)

01



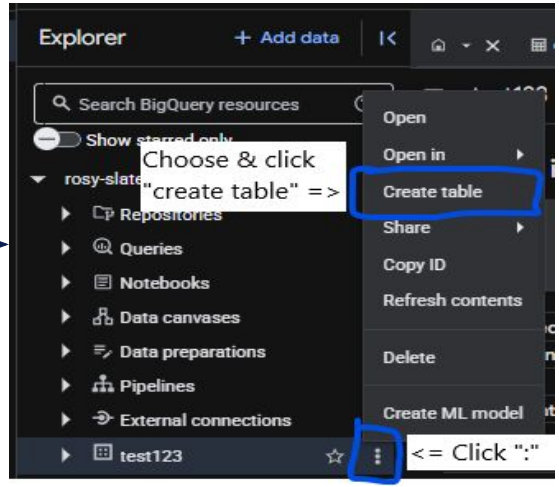
Click “⋮” on the right of project name

02



Write Database name and Click **“Create dataset”** button

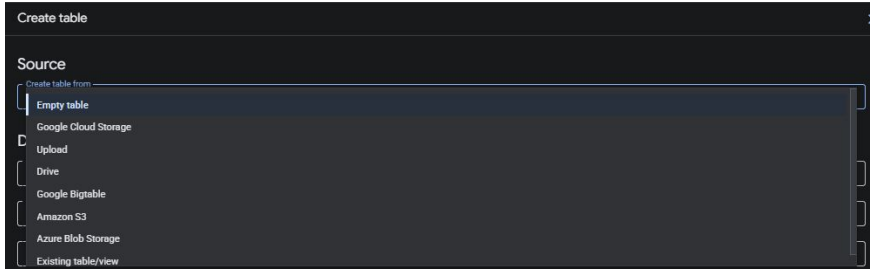
03



After the database is successfully created, click “⋮” then choose **"Create table"** option

Preparation before Analysis in Google BigQuery (2)

04

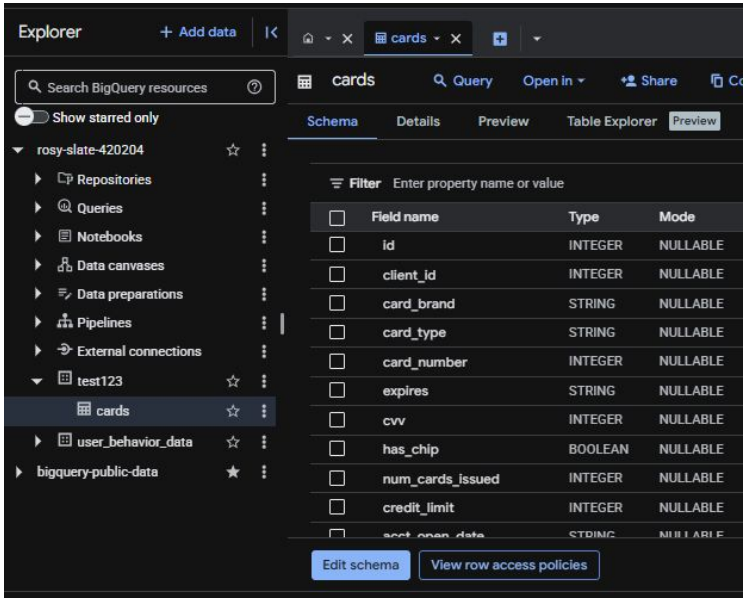


After you choose "Create table" option, you can choose the data source that you want to use, either Google Cloud Storage, a dataset by link stored in Google Drive, or you can upload your dataset from internal storage, etc

After that, you can write table name in “Table” column then click checkbox “Auto detect” for Schema

Preparation before Analysis in Google BigQuery (3)

05



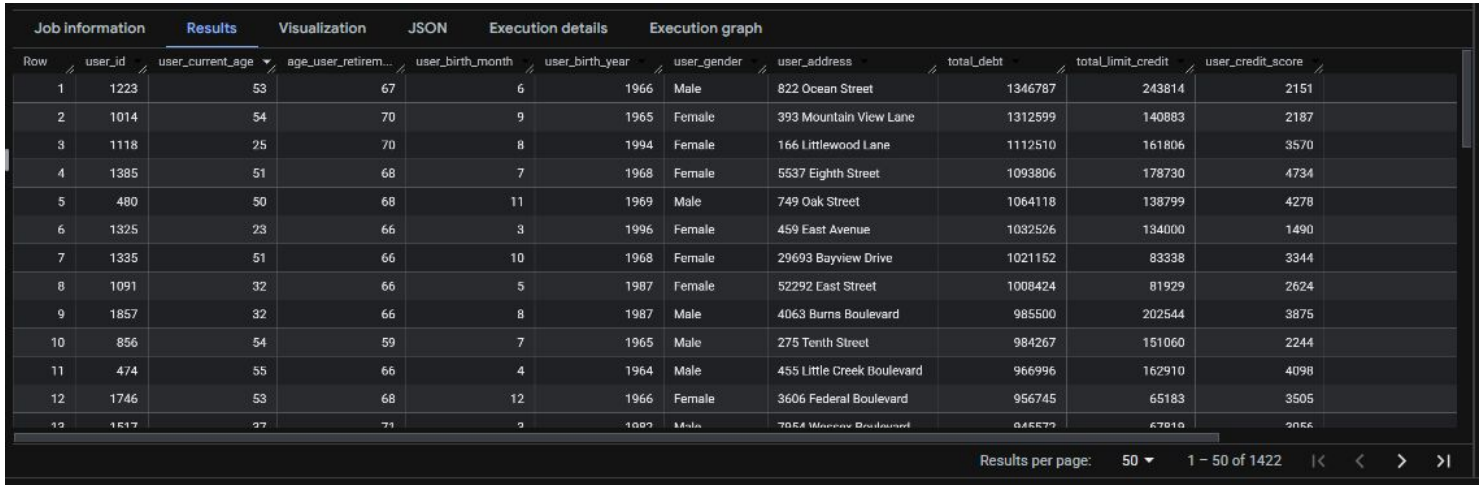
The first table of the "cards_data.csv" dataset has been successfully created in the sample database named "test123"

Next, repeat the steps to upload "transactions_data.csv" and "users_data.csv"

03

Data Analysis with SQL Queries

Displays User Profiles with Total Debt > Total Credit Limit, Age ≤ 60 Years, and not exceeding their respective Retirement Age

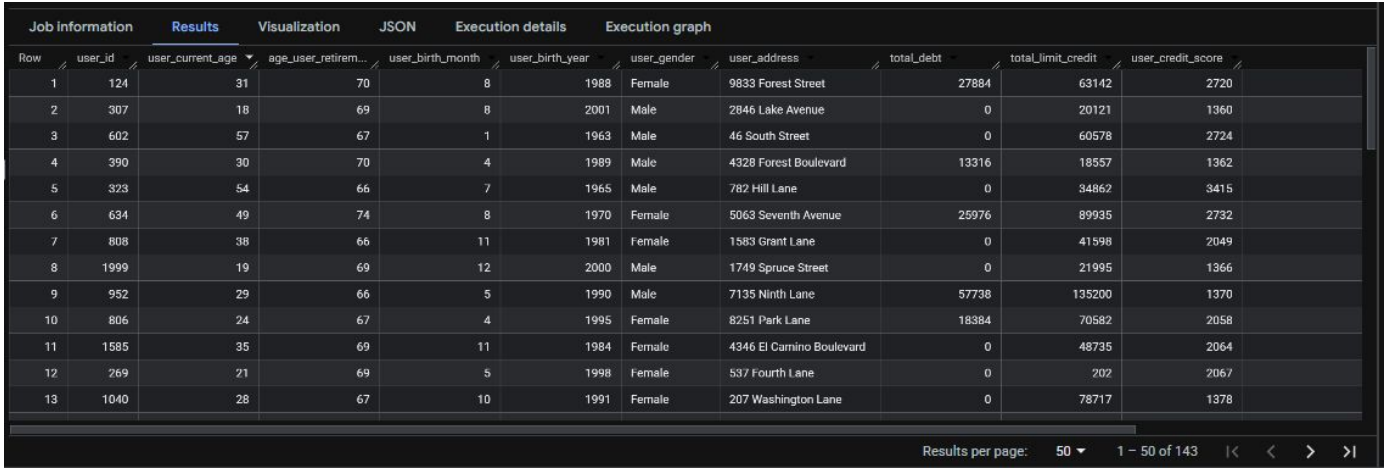


Job information | Results | Visualization | JSON | Execution details | Execution graph

Row	user_id	user_current_age	age_user_retirem...	user_birth_month	user_birth_year	user_gender	user_address	total_debt	total_limit_credit	user_credit_score
1	1223	53	67	6	1966	Male	822 Ocean Street	1346787	243814	2151
2	1014	54	70	9	1965	Female	393 Mountain View Lane	1312599	140883	2187
3	1118	25	70	8	1994	Female	166 Littlewood Lane	1112510	161806	3570
4	1385	51	68	7	1968	Female	5537 Eighth Street	1093806	178730	4734
5	480	50	68	11	1969	Male	749 Oak Street	1064118	138799	4278
6	1325	23	66	3	1996	Female	459 East Avenue	1032526	134000	1490
7	1335	51	66	10	1968	Female	29693 Bayview Drive	1021152	83338	3344
8	1091	32	66	5	1987	Female	52292 East Street	1008424	81929	2624
9	1857	32	66	8	1987	Male	4063 Burns Boulevard	985500	202544	3875
10	856	54	59	7	1965	Male	275 Tenth Street	984267	151060	2244
11	474	55	66	4	1964	Male	455 Little Creek Boulevard	966996	162910	4098
12	1746	53	68	12	1966	Female	3606 Federal Boulevard	956745	65183	3505
13	1517	37	71	3	1982	Male	7954 Waverly Boulevard	945572	67819	3056

Results per page: 50 1 – 50 of 1422

Displays User Profiles with Total Debt < Total Credit Limit, Age ≤ 60 Years, and not exceeding their respective Retirement Age



Job information | Results | Visualization | JSON | Execution details | Execution graph

Row	user_id	user_current_age	age_user_retirem...	user_birth_month	user_birth_year	user_gender	user_address	total_debt	total_limit_credit	user_credit_score
1	124	31	70	8	1988	Female	9833 Forest Street	27884	63142	2720
2	307	18	69	8	2001	Male	2846 Lake Avenue	0	20121	1360
3	602	57	67	1	1963	Male	46 South Street	0	60578	2724
4	390	30	70	4	1989	Male	4328 Forest Boulevard	13316	18557	1362
5	323	54	66	7	1965	Male	782 Hill Lane	0	34862	3415
6	634	49	74	8	1970	Female	5063 Seventh Avenue	25976	89935	2732
7	808	38	66	11	1981	Female	1583 Grant Lane	0	41598	2049
8	1999	19	69	12	2000	Male	1749 Spruce Street	0	21995	1366
9	952	29	66	5	1990	Male	7135 Ninth Lane	57738	135200	1370
10	806	24	67	4	1995	Female	8251 Park Lane	18384	70582	2058
11	1585	35	69	11	1984	Female	4346 El Camino Boulevard	0	48735	2064
12	269	21	69	5	1998	Female	537 Fourth Lane	0	202	2067
13	1040	28	67	10	1991	Female	207 Washington Lane	0	78717	1378

Results per page: 50 | 1 – 50 of 143

Most Frequently Used Debit & Credit Card Brand

Row	card_brand	card_type	count_of_card_us...	total_credit_limit
1	Mastercard	Debit	2191	40050847
2	Visa	Debit	1320	25105900
3	Visa	Credit	811	9160700
4	Mastercard	Credit	635	6967000
5	Amex	Credit	402	4597400
6	Mastercard	Debit (Prepaid)	383	24810
7	Discover	Credit	209	2260600
8	Visa	Debit (Prepaid)	195	12441

If the purpose of using the card is Debit, then Mastercard brand is most used rather than Visa because Mastercard debit provide the highest total credit limit compared to Visa

But if the purpose of using is Credit, then Visa card brand is most used rather than Mastercard because Visa in Credit provide the highest total credit limit

Comparison of Total Owned Credit Card

Row	total_user_with_have_more_than_1_credit_cards	total_user_with_have_only_1_credit_card
1	1584	416

The results above show that the most contrasting difference in user behavior is the total number of credit cards owned

More than 1.500 users (1.584 users) have more than 1 credit cards. Meanwhile, only 416 users have 1 credit card

List of Users Detail who have more than 1 Credit Cards

Row	user_id	card_brand	user_card_number	card_type	total_debt	start_date_of_account_user	end_of_expired_date_account_user	credit_limit
1	0	Mastercard	5802759460691737	Credit	36199	09/2007	12/2019	1760
2	0	Mastercard	5691786126216800	Debit	36199	02/2020	06/2021	3100
3	0	Mastercard	5472791324252431	Debit	36199	02/2020	04/2020	2555
4	0	Mastercard	5050211780967429	Debit	36199	02/2011	04/2021	3149
5	1	Visa	4419924074647230	Credit	14587	09/2007	12/2021	1280
6	1	Visa	4843491272960882	Debit	14587	08/2012	01/2017	1810
7	1	Visa	4417513283605637	Credit	14587	07/2002	04/2014	1090
8	2	Mastercard	5617443939610345	Debit	80850	12/2012	08/2024	1649
9	2	Mastercard	5458954548704072	Debit	80850	06/2015	07/2024	900
10	2	Mastercard	5191615123045322	Credit	80850	06/2005	11/2020	740
11	2	Mastercard	5077942665919872	Debit	80850	04/2008	08/2023	966
12	2	Visa	4809764397537147	Credit	80850	05/2006	12/2013	780
13	3	Mastercard	5175412661393752	Debit (Prepaid)	18693	09/2012	08/2023	6
14	3	Mastercard	5568042569458348	Debit (Prepaid)	18693	08/2010	04/2023	8

Results per page: 50 1 – 50 of 5730

A total of 5,730 rows of data with a distribution of 1,584 unique users have more than 1 credit card with different total credit limit

List of Users with Productive Age (23 - 60 Years) who Start Date Accounts after 2015 with Expired after 2020

Row	user_id	user_current_age	start_date_of_account_user	end_of_expired_date_account_u...
1	628	57	02/2020	02/2024
2	315	54	02/2020	03/2023
3	421	53	02/2020	06/2023
4	107	50	01/2020	09/2021
5	1376	49	02/2020	06/2021
6	1981	48	06/2016	04/2021
7	1954	47	03/2019	05/2024
8	185	47	07/2016	07/2024
9	839	45	05/2016	04/2024
10	308	44	06/2018	08/2023
11	1586	43	02/2016	02/2021
12	455	41	07/2017	10/2022
13	675	40	08/2016	06/2021
14	1589	39	09/2017	12/2024

Results per page: 50 1 – 50 of 573

Row	total_unique_user
1	436

A total of 436 unique users with a productive age range of 23 - 60 years produced 573 rows of data spread with the month + year of credit account creation above 2015 and the month + year of account expiration above 2020

Top 25 Users with the Most Transactions in 2019

Row	user_id	count_of_user_transaction
1	1098	4104
2	1888	3303
3	909	3266
4	96	3242
5	114	2995
6	1776	2707
7	285	2620
8	208	2393
9	1559	2306
10	1654	2271
11	1963	2253
12	1727	2240
13	1575	2130
14	1168	2095
15	1452	2080
16	379	2026
17	1604	2020
18	464	1990
19	1966	1981
20	112	1970
21	997	1958
22	488	1932
23	502	1931
24	598	1877
25	1382	1864

User with ID 1098 has the highest total number of transactions among the 24 other users who had the highest total transaction in 2019, 4,104 transactions

This indicates that user ID 1098 has spending behaviors, such as purchasing products/goods

Top 21 Merchant Cities with Total Transaction > 1,500 during 2019

Row	merchant_city	address_of_transaction	unique_count_of_transaction	sum_of_total_transaction_amount
1	Olympia	215 Ocean View Drive	3094	52517.31
2	Yorba Linda	20 Oak Street	2834	145794.5
3	North Hollywood	112 Elm Avenue	2238	46220.99
4	Farmington	7854 Plum Avenue	2211	65963.64
5	Columbia	31 Oak Avenue	2129	42608.5
6	Vacaville	468 Spruce Street	2129	64549.73
7	Acworth	607 George Boulevard	2060	53983.52
8	Nashport	130 Lincoln Drive	2004	34086.03
9	Houston	524 Ocean Drive	1921	61381.81
10	Orlando	2352 Bayview Boulevard	1849	132650.22
11	Brooklyn	2479 Valley Avenue	1747	36432.96
12	Dallas	16 Eighth Lane	1730	41509.29
13	Mobile	837 Lincoln Avenue	1594	32941.49
14	Albuquerque	9385 Birch Street	1579	21526.26
15	Berkeley Heights	207 Ocean View Street	1577	50550.23
16	Columbus	878 Forest Street	1552	46483.22
17	Houston	545 Eighth Avenue	1521	44427.88
18	Newport	385 Pine Drive	1520	58050.58
19	Tucson	153 Tenth Lane	1519	65731.72
20	Riverside	691 Ocean View Street	1513	61848.1
21	Paragould	3835 Norfolk Street	1503	23619.28

Olympia as one of the merchant cities had the highest number of transaction values among other cities. A total of 3.094 number transactions were conducted in Olympia first top merchant cities if we looking up from number of transaction

Furthermore, Olympia wasn't among the merchant cities with the highest sum of total transaction amount, only 52.517 transactions

Instead, Yorba Linda, despite only having 2,834 transactions, managed to achieve a total transaction value of 145.794

List of User Profiles with Age ≤ Average Overall Age and Total Transactions ≥ Average Total Transactions in 2019

Job information | Results | Visualization | JSON | Execution details | Execution graph

Row	user_id	user_gender	user_current_age	total_amount_of_user_transaction_2019
1	0	Male	33	44.43
2	1	Female	43	45.66
3	11	Female	39	50.87
4	28	Female	39	56.9
5	34	Male	41	44.95
6	37	Female	41	26.43
7	38	Female	45	39.59
8	51	Female	29	47.2
9	57	Female	28	38.33
10	58	Female	38	67.2
11	61	Female	44	57.58
12	63	Male	40	31.29
13	64	Female	39	51.56
14	68	Male	42	56.84

Results per page: 50 1 – 50 of 415

A total of 415 users meet these two conditions, where they (the users who transact) have an age below or equal to the average age of all users and these 415 users have a total transaction value greater than or equal to the total average transaction value as a whole.

Credit Card Brand with Highest Total of Credit Limit and the Number of Uses

Row	card_brand	total_credit_limit	count_of_card_uses
1	Discover	2260600.0	209
2	Amex	4597400.0	402
3	Visa	34279041.0	2326
4	Mastercard	47042657.0	3209

Mastercard credit card brand has the highest number of users with a total of 3.209

The main reason why Mastercard credit card has highest number of users because Mastercard offers and provide the highest credit limit among the other three credit card brands, Visa, Amex, and Discover, 47.042.657 total credit limit from Mastercard

Total Debt, Total Amount of Transaction Value, and Count Distinct of Transaction per Each Month in 2019

Row	month_name	total_debt	total_amount_of_transaction	number_of_transaction
1	January	6023195792.0	4211984.83	103639
2	February	5435896291.0	3815276.02	93455
3	March	6114192926.0	4246403.65	104300
4	April	5839672236.0	4146296.78	100243
5	May	6061940672.0	4244264.63	103398
6	June	5888455727.0	4177797.3	100922
7	July	6082033711.0	4230436.17	104381
8	August	6107036491.0	4219926.11	104495
9	September	5876245810.0	4096403.89	100997
10	October	6032904273.0	4197883.96	103065

1. March 2019 has the highest total debt among other months with total number of debt is 6.114.192.926.0
2. Still on March 2019 has the highest total amount of transaction value with number of 4.246.403.65
3. August 2019 has the highest of number transaction with total number is 104.495

Total Number of Transactions from Each Credit Card Brand by Months

Row	month_name	card_brand	number_of_transaction
1	January	Mastercard	55649
2	January	Visa	38685
3	January	Amex	6528
4	January	Discover	2777
5	February	Mastercard	50050
6	February	Visa	35036
7	February	Amex	5930
8	February	Discover	2439
9	March	Mastercard	55936
10	March	Visa	39263
11	March	Amex	6464
12	March	Discover	2637
13	April	Mastercard	53918
14	April	Visa	37323
15	April	Amex	6327
16	April	Discover	2675
17	May	Mastercard	55419
18	May	Visa	38733
19	May	Amex	6578
20	May	Discover	2668
21	June	Mastercard	54448
22	June	Visa	37474
23	June	Amex	6353
24	June	Discover	2647
25	July	Mastercard	56444
26	July	Visa	38807
27	July	Amex	6397
28	July	Discover	2733
29	August	Mastercard	56068
30	August	Visa	39197
31	August	Amex	6423
32	August	Discover	2807
33	September	Mastercard	54084
34	September	Visa	38001
35	September	Amex	6235
36	September	Discover	2677
37	October	Mastercard	55488
38	October	Visa	38452
39	October	Amex	6638
40	October	Discover	2487

In 10 months of transactions in 2019 (January - October), Mastercard credit card brand always dominated the market because users had used almost more than 50.000 different transactions in each month

Total Number of Debit/Credit Account PIN Code Changes for Users with Ownership of > 1 Credit Card

Row	user_id	user_card_brand_owned	card_type	count_of_total_year_pin_last_changed
1	303	Amex	Credit	2
2	303	Mastercard	Debit	2
3	303	Visa	Credit	2
4	1659	Mastercard	Debit	3
5	1659	Visa	Credit	2
6	1135	Amex	Credit	2
7	1135	Visa	Debit	2
8	1576	Mastercard	Debit	3
9	1576	Mastercard	Debit (Prepaid)	2
10	1576	Visa	Debit	2

This analysis is particularly useful for observing the behavior of users with more than one credit card, including how many times they have changed their PIN codes. For example, user ID 1659, who owns 2 different card brands (Mastercard and Visa), has changed PIN code five times, 3 changes for Mastercard and 2 changes for Visa

10 User Profiles with Low Total Credit Scores

Row	user_id	user_gender	user_current_age	user_address	user_credit_score	total_debt	total_credit_limit	difference_between_credit_limit_and_total_debt
1	1801	Female	18	3371 Madison Boulevard	480	70064	7284.0	-62780.0
2	1987	Male	63	786 12th Drive	488	24997	32200.0	-17794.0
3	1163	Male	45	2270 Sixth Lane	489	114251	49146.0	-293607.0
4	559	Male	56	593 Valley Stream Drive	490	53185	26982.0	-79388.0
5	630	Male	60	1583 Grant Avenue	491	61861	34668.0	-150915.0
6	1733	Female	47	6613 Hill Drive	498	54290	898.0	-53392.0
7	80	Female	82	362 Martin Luther King Street	498	1597	43272.0	35287.0
8	1818	Female	85	23 11th Avenue	500	1741	76420.0	71197.0
9	1436	Female	21	729 Wessex Avenue	500	54189	48789.0	-167967.0
10	897	Female	30	847 Martin Luther King Lane	501	33332	23660.0	-43004.0

By knowing the profile of the 10 users who have the lowest credit score, we can anticipate by reject all forms of credit if 10 users want to apply for additional credit when their annual income doesn't increase/decrease. This explains that the user does not have good payment capabilities

10 User Profiles with Good Total Credit Scores

Job information | Results | Visualization | JSON | Execution details | Execution graph

Row	user_id	user_gender	user_current_age	user_address	user_credit_score	total_debt	total_credit_limit	difference_betwe...	yearly_income	difference_between_yearly_income_and_total_debt
1	984	Female	63	642 Ocean View Lane	850	80794	44943.0	-359027.0	32928	-239330.0
2	678	Female	63	861 Martin Luther King Boulevard	850	56355	43857.0	-181563.0	29477	-107512.0
3	929	Female	63	899 Grant Lane	850	59198	36200.0	-22998.0	56102	-3096.0
4	729	Female	80	499 Lafayette Lane	850	1828	84584.0	73616.0	24884	138336.0
5	103	Female	41	8 Tenth Drive	850	78060	63477.0	-326823.0	34774	-216430.0
6	21	Male	48	819 Park Boulevard	850	0	34572.0	34572.0	43084	129252.0
7	1705	Male	50	952 Valley Avenue	850	2070	17274.0	11064.0	27952	77646.0
8	1104	Male	91	120 Lafayette Boulevard	850	805	91322.0	86492.0	40141	236016.0
9	1369	Female	59	535 Grant Boulevard	850	49101	23716.0	-74486.0	35854	-26494.0
10	893	Male	18	4485 Plum Lane	850	18962	28129.0	-28757.0	34716	47262.0

By identifying the profiles of 10 users with good credit score, we can consider applying for a loan. Although, based on the difference between annual income and total debt, some users have a negative value, if they have a good credit score, we can consider them because they have a good repayment history

04

Conclusion

Final Conclusion based on Analysis using SQL Queries (1)

01

The majority of users have more than 1 credit cards with a ratio of 416 users only have 1 credit card and 1.584 users have more than 1 credit cards

02

Olympia as a merchant city has the highest number of transactions (3.094 number of transactions) but in terms of the sum of the total amount of transactions, Yorba Linda is the merchant city with transaction of 145.794

03

Mastercard is the most widely used credit card (3.209 users) because it provides the largest credit limits among the other 3 credit cards (47,042,657 total credit limit)

04

March 2019 was the month that produced the highest total debt amount and had the highest sum of total transaction amount among other months

Final Conclusion based on Analysis using SQL Queries (2)

05

The users credit score is very important and must be taken into account by the financial & banking industry in providing loans to users because it is the main factor in whether the user is worthy and able to repay the loan or not

05

Data Report in Looker Data Studio

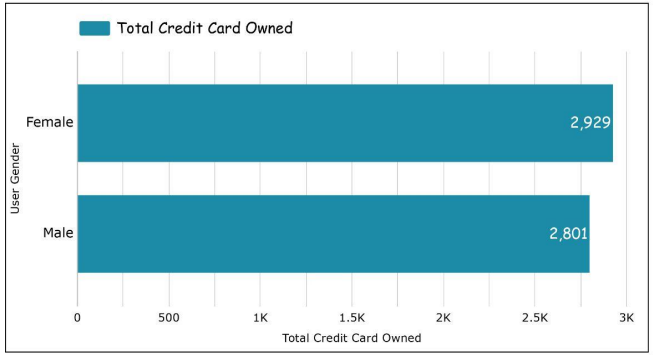
First Page in Dashboard Design

Summary of User Profile Overview

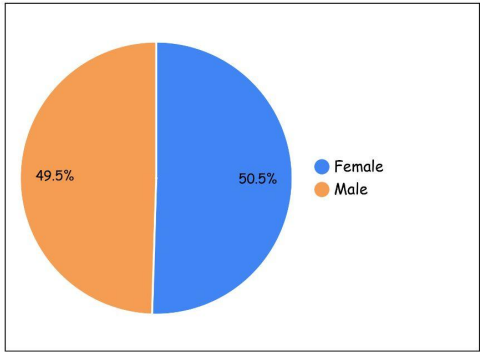
User Gender

User ID

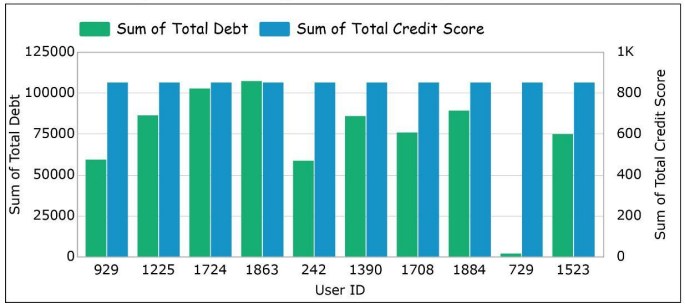
Comparison of User Gender based on Total Credit Card Owned > 1



% of Comparison Total Debt from Gender



Top 10 User with Highest Credit Score plus Total Debt



10 Users with Income per Capita and Yearly Income

	User ID	Income per Capita	Yearly Income
1.	929	44,106	56,102
2.	1225	30,819	62,837
3.	1724	18,036	36,775
4.	1863	21,639	44,124
5.	242	21,365	43,560
6.	1390	29,868	60,897
7.	1708	22,195	45,254
8.	1884	28,092	57,281
9.	729	16,421	24,884
10.	1523	19,453	39,661

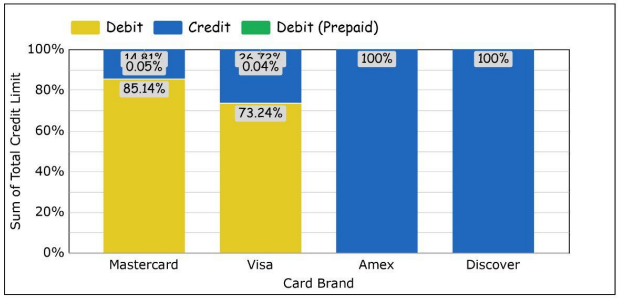
Second Page in Dashboard Design

Summary of Card User Used Overview

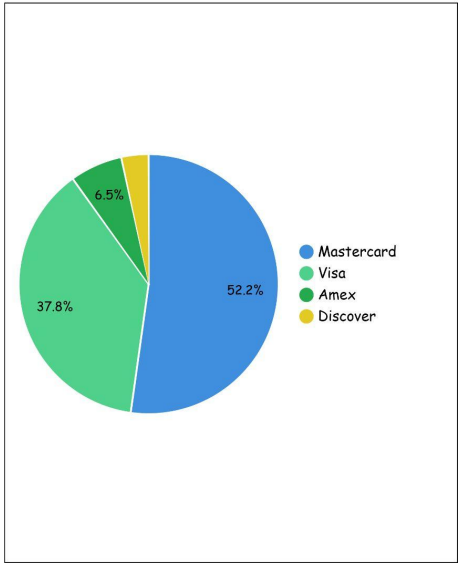
Card Type

Unique of Card Brand	Unique of Card Number	Total Credit Limit	Total Pin Last Changed
4	6,146	88179698	1240

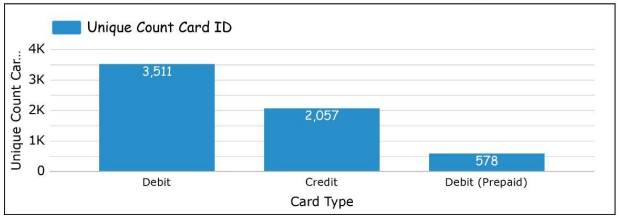
% of Total Credit Limit based on Card Brand and Card Type



% of Count Card Brand Uses



Unique of Total Count Number based on Card Type



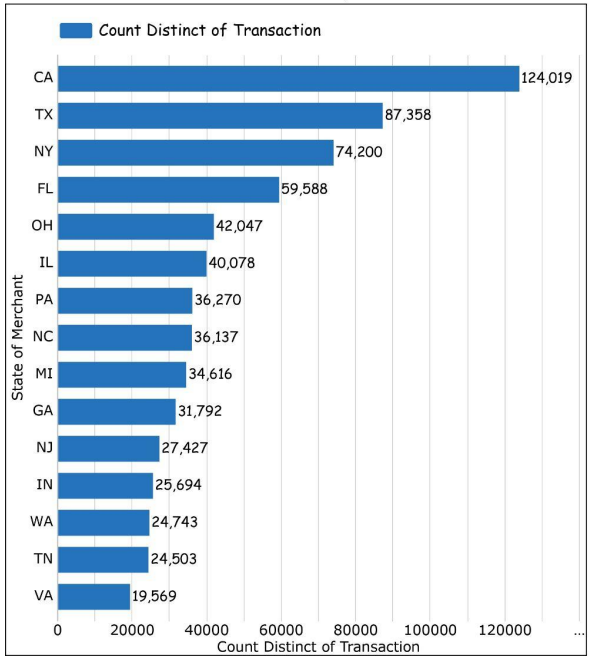
Third Page in Dashboard Design

Summary of History Transaction User Overview

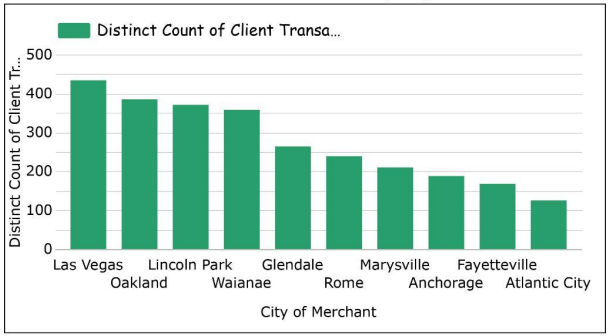
Month of Transaction

Total Rows	Total Card No for Transaction	Distinct Count of Date Transaction	Year Date
1,018,895	3,436	339,439	2019

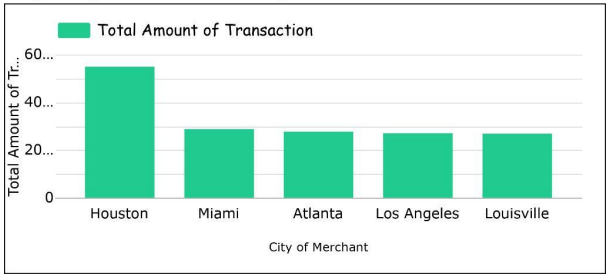
Count Distinct of Transaction by State of Merchant



Distinct Count of Client Transaction by City of Merchant



Top 5 City of Merchant with Highest of Total Amount of Transaction



THANK YOU!

Any Question?
Let's discuss