04/10/2024

CUSTOMER360 ANALYSIS

REPORT

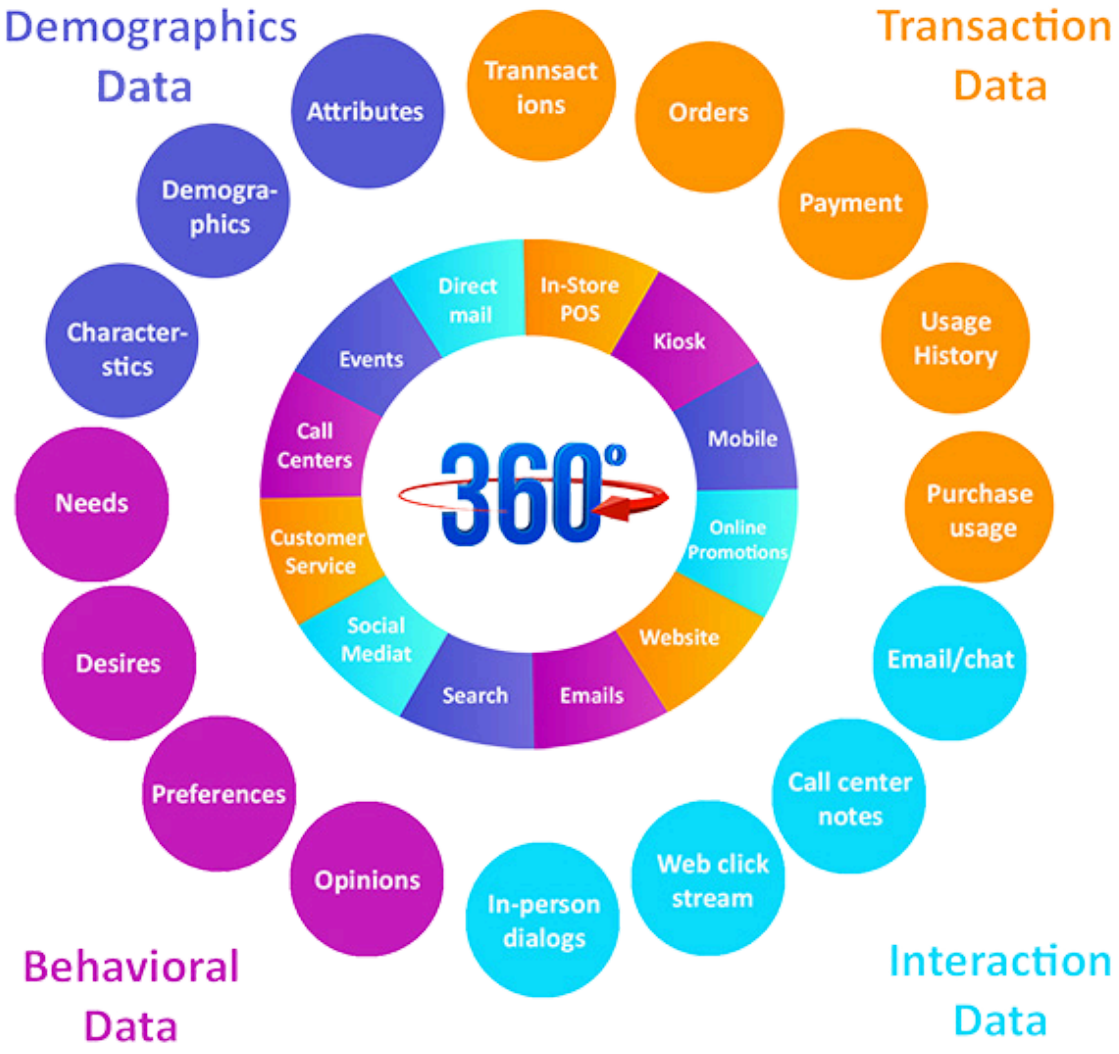
Prepared by:
NGUYEN MINH KHOI

INTRODUCTION

The Customer360 Analysis Report is a comprehensive report that analyzes customer data based on theories and models such as the Customer360 framework, RFM model, IQR method, and BCG matrix. It provides insights into grouping and analyzing customer groups.

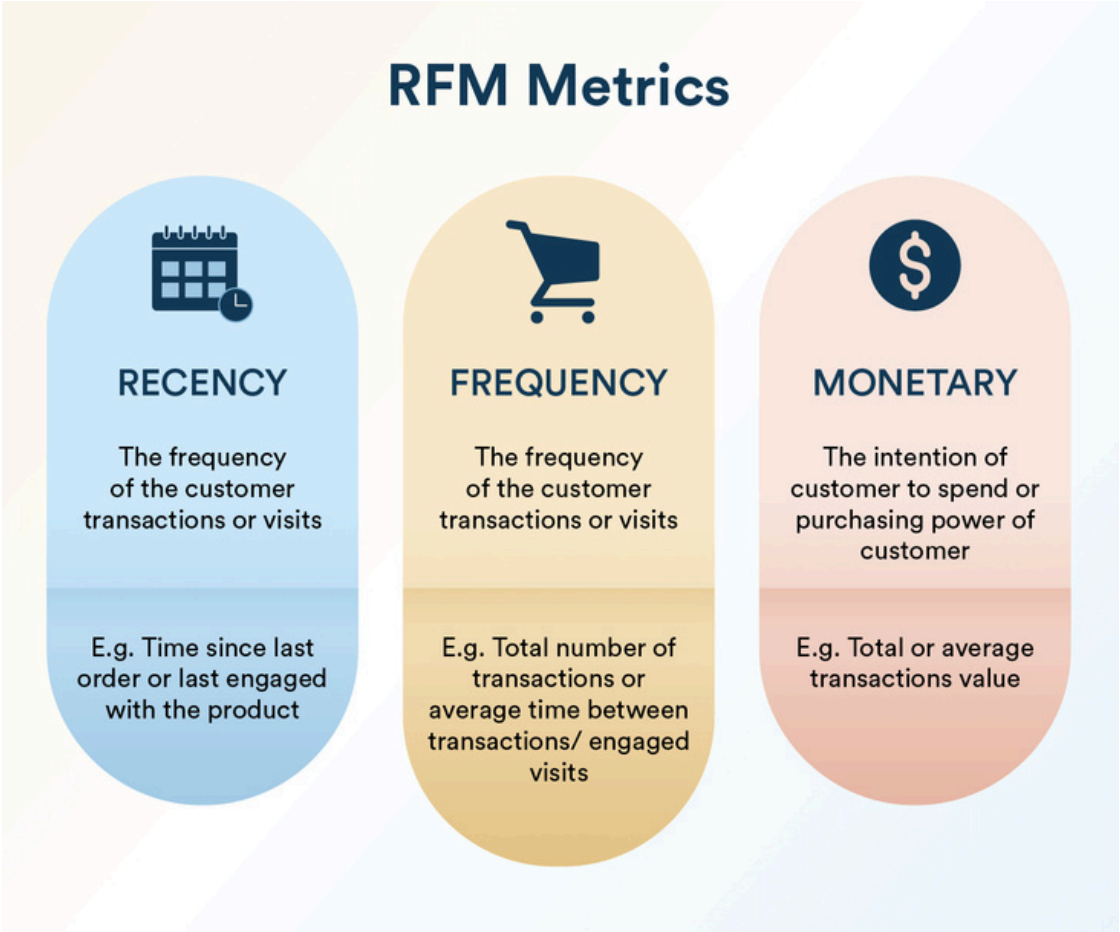
OVERVIEW

WHAT IS CUSTOMER360?



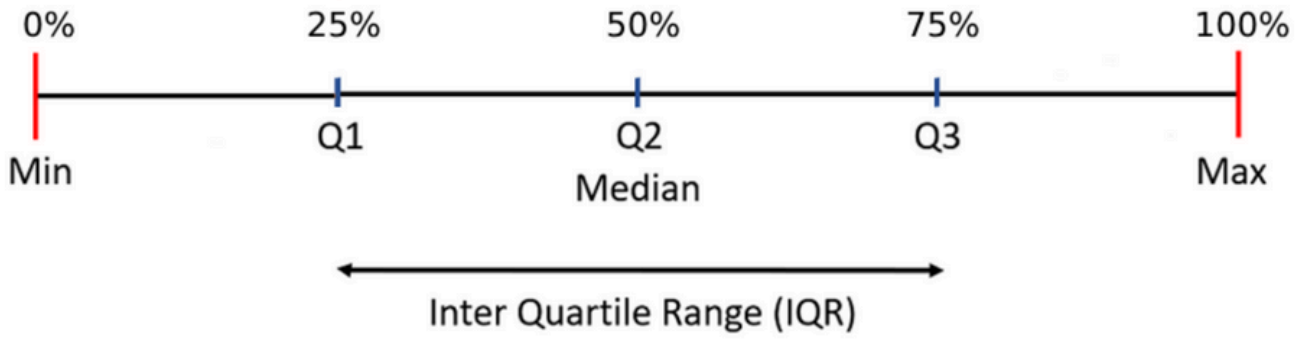
- Customer360 is a framework designed to give businesses a holistic, 360-degree view of customers across all touchpoints and interactions.
- It involves collecting, integrating, and analyzing diverse customer data from various sources and consolidating it into a cohesive profile of each customer

WHAT IS RFM METRICS?



- RFM (Recency, Frequency, Monetary) is a marketing analysis tool used to identify a company's best customers based on their purchasing behavior. It helps businesses segment their customer base and tailor marketing strategies to different customer groups.
- RFM stands for:
  - Recency: How recently a customer made a purchase.
  - Frequency: How often a customer makes purchases.
  - Monetary Value: How much money a customer spends on average.

WHAT IS IQR METHOD?

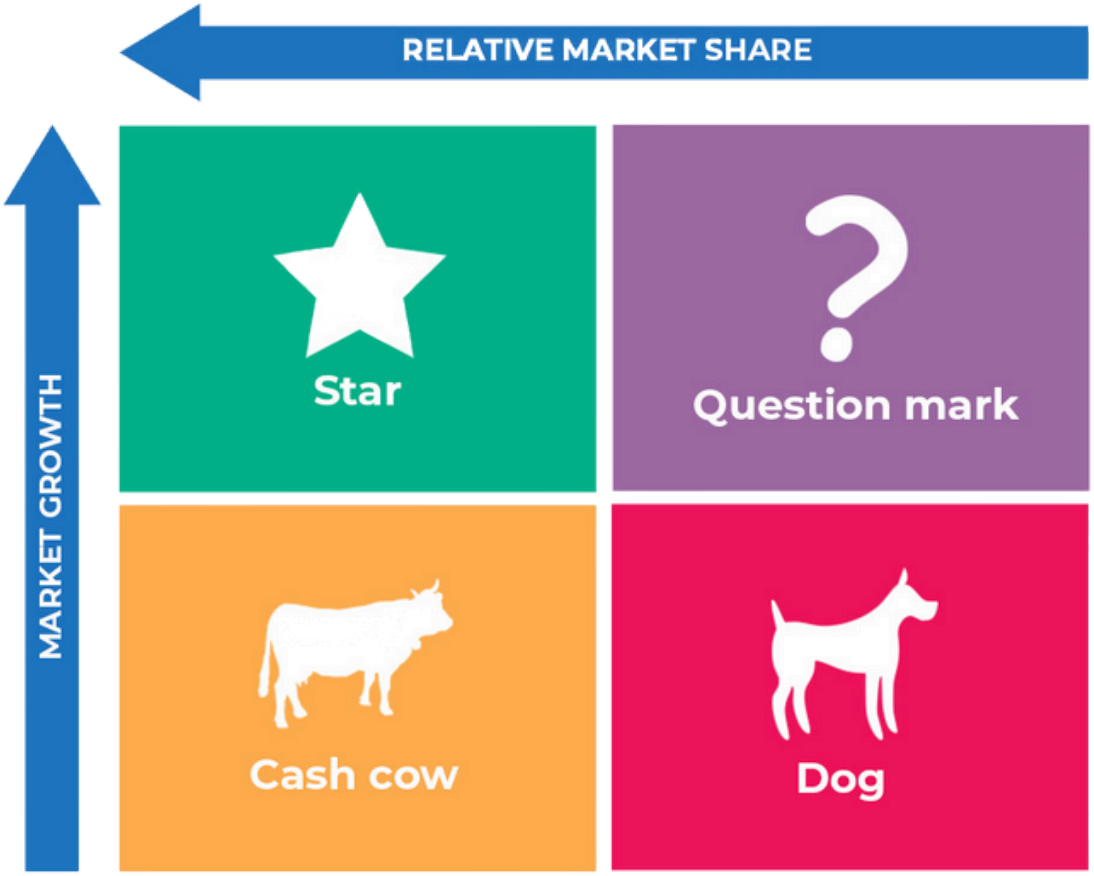


- The IQR is a statistical measure that characterizes the spread and variability of a dataset. It is calculated as the difference between the upper quartile (Q3) and the lower quartile (Q1) values. In essence, is creating a box-plot with the data, where the box composed of data that falls between quartiles 1 and 3, the wires are the min and max.

- Interestingly, despite it gives you information on the distribution of the data, is one of the methods that is not too affected by the non-normality or non-parametricity of the data.

- It is calculated as the difference between the upper quartile (Q3) and the lower quartile (Q1) values. In essence, is creating a box-plot with the data, where the box composed of data that falls between quartiles 1 and 3, the wires are the min and max.

WHAT IS BCG MATRIX?



- The BCG Matrix (Boston Consulting Group Matrix) is a tool used to analyze the prospects of a company's existing products or services based on their growth rate and market share. Accordingly, they are categorized into four corresponding quadrants: Stars, Cash Cows, Dogs, and Question Marks.

ANALYSIS

DESCRIPTIVE DATA

- We have customer data stored in two tables:
  - Table 1: Customer_Registered: Stores customers' personal information.
  - Table 2: Customer_Transaction: Stores customers' transaction information.

Customer_Registered

Column	Data Type	Describe
ID	int	Customer code
Contract	varchar(50)	Contract code
LocationID	int	Location code
BranchCode	int	Branch code
Status	int	Status
Created_date	datetime	Registration date
Stopdate	datetime	Cancellation date

Customer_Registered data sample

	ID	Contract	LocationID	BranchCode	Status	created_date	stopdate
0	0	SGDN00215	8	1	0	2011-11-25 00:00:00.000	2012-01-05 00:00:00.000
1	1	SGDN00214	8	1	0	2012-06-14 00:00:00.000	
2	2	SGD374348	8	1	0	2012-11-01 00:00:00.000	
3	3	SGD022064	8	1	2	2011-06-22 00:00:00.000	2013-05-29 00:00:00.000
4	4	SGD041015	8	5	2	2011-12-17 00:00:00.000	2014-11-11 00:00:00.000
5	5	SGDN00211	[NULL]	[NULL]	2	2015-06-09 00:00:00.000	2015-09-09 00:00:00.000

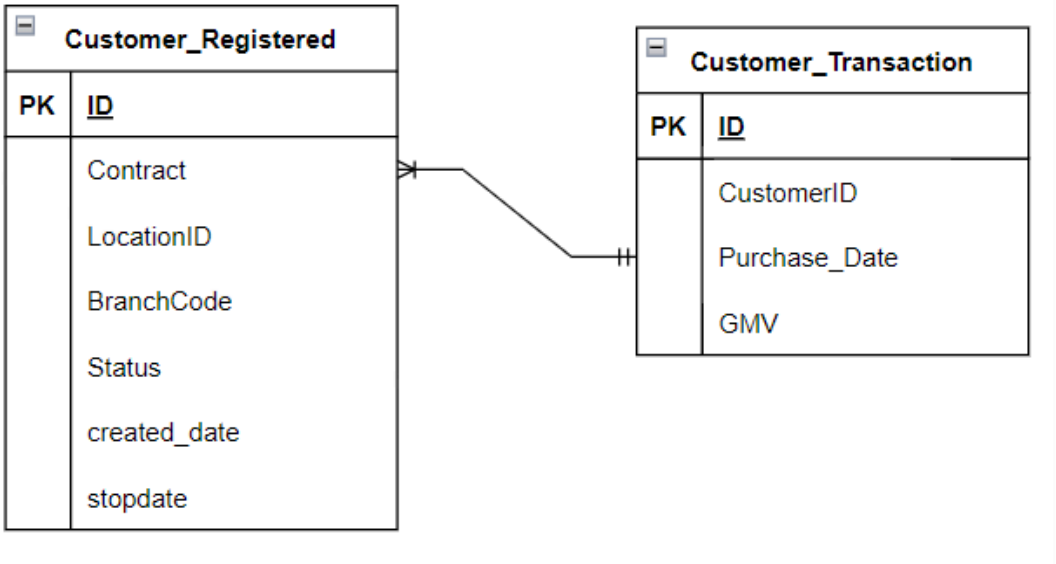
Customer_Transaction

Column	Data Type	Describe
ID	int	Transaction code
CustomerID	int	Customer code
Purchase_Date	datetime	Transaction time
GMV	int	Transaction value

Customer_Transaction data sample

	ID	CustomerID	Purchase_Date	GMV
0	0	1,327,813	2022-06-01 00:00:00.000	95,000
1	1	1,157,830	2022-06-01 00:00:00.000	75,000
2	2	873,915	2022-07-01 00:00:00.000	95,000
3	3	3,505,071	2022-07-01 00:00:00.000	90,000
4	4	2,930,918	2022-07-01 00:00:00.000	109,091
5	5	899,882	2022-06-01 00:00:00.000	105,000

RELATIONAL DATABASE MODEL



DATA PROCESSING

- Based on the RFM model, each customer is scored on three factors: R, F, and M, specifically:
  - Recency = Report date – Most recent purchase date
  - Frequency = Number of purchases / Contract age (to avoid bias toward customers with frequent but low-value purchases compared to newer customers with high-value purchases)
  - Monetary = Total purchase value / Contract age

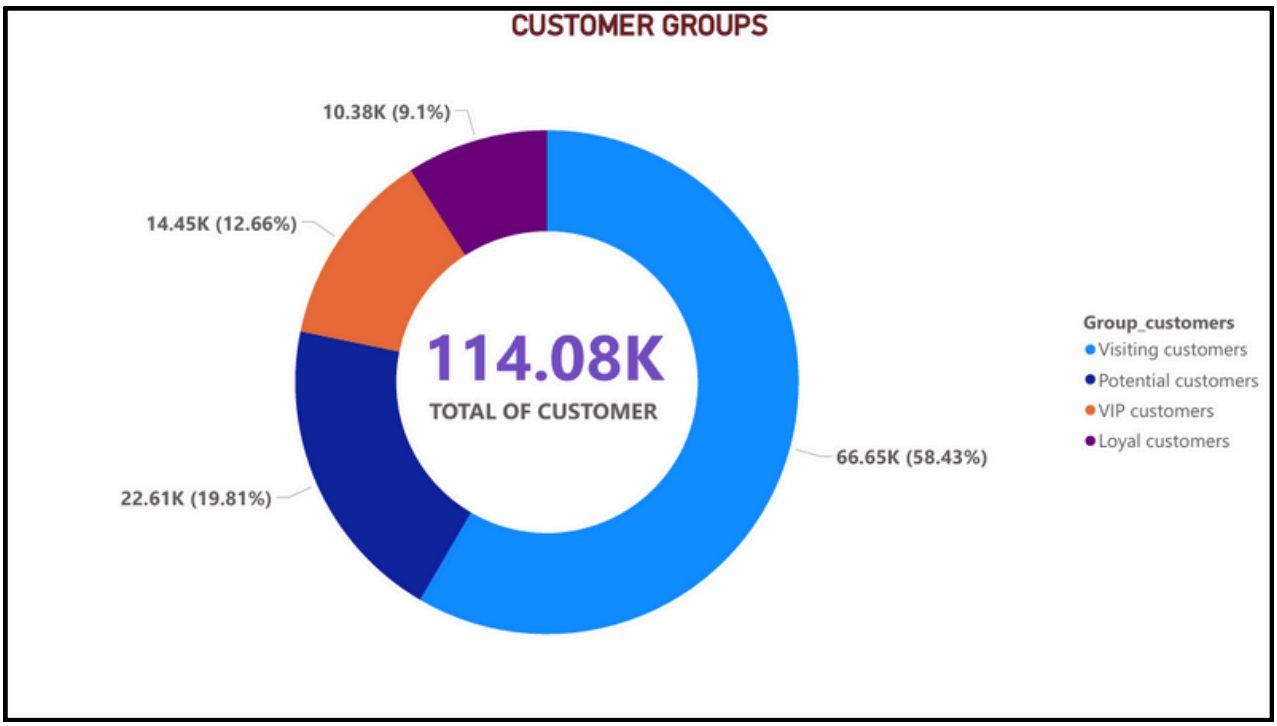
- After using the IQR method, the R, F, M scores are determined as shown in the table below:

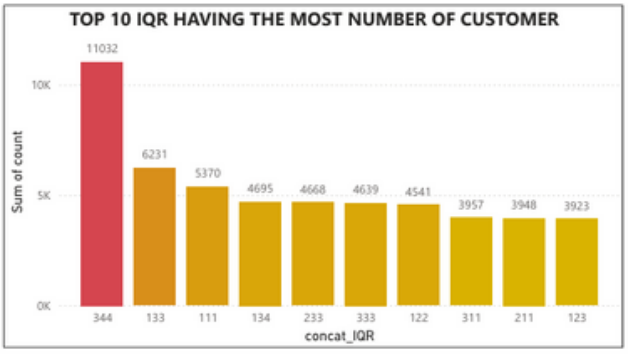
IQR	1	2	3	4
R	92	91-62	61-31	30-01
F	0.0005	0.0006	0.0007	0.0008-0.0033
M	0-42.8082	42.8081-52.534	52.533-65.2921	65.2920-419.0133

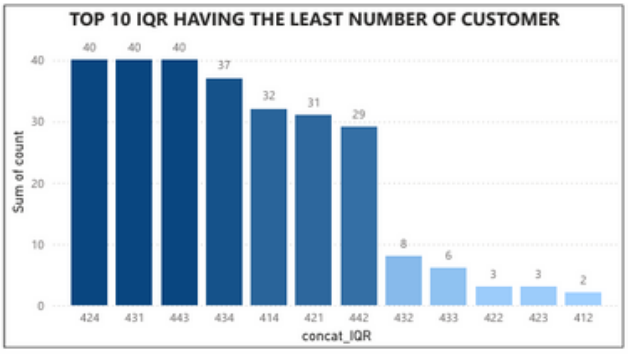
- Combine the three scores into a single RFM index and classify customers into groups based on the BCG Matrix:
  - Stars – VIP Customers: High market share, high growth rate.
  - Question Marks – Loyal Customers: Low market share, high growth rate.
  - Cash Cows – Potential Customers: High market share, low growth rate.
  - Dogs – Occasional Customers: Low market share, low growth rate.

Customer group		Description	RFM index
Stars	VIP customers	Are the customers buy regularly and large order value	444, 443, 434, 344, 334
Question Marks	Loyal customers	Are customers who buy frequently and have a fairly high order value	433, 424, 423, 414, 413, 343, 333, 324, 323,314, 313
Cash Cows	Potential customers	Are the customers Buy less and value normal order	442, 432, 422, 412, 342, 332, 322, 312, 242, 233, 244, 243, 234, 224, 223, 214, 213
Dogs	Visiting customers	Are customers already long time purchase or customer New item, single value low row	441, 431, 421, 411, 144, 143, 142, 141, 134, 133, 132, 131, 124, 123, 122, 121, 114, 113, 112, 111, 341, 331, 321, 311, 241, 232, 231, 222, 221, 212, 211

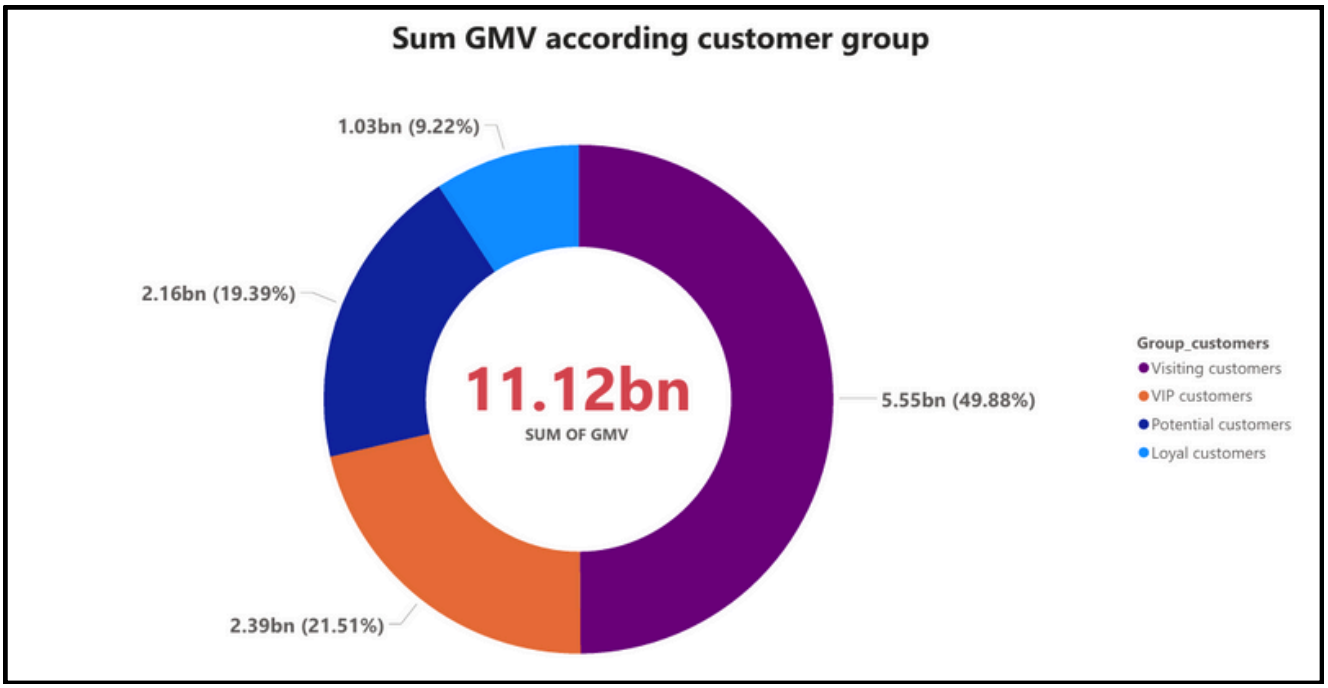
CONCLUSION



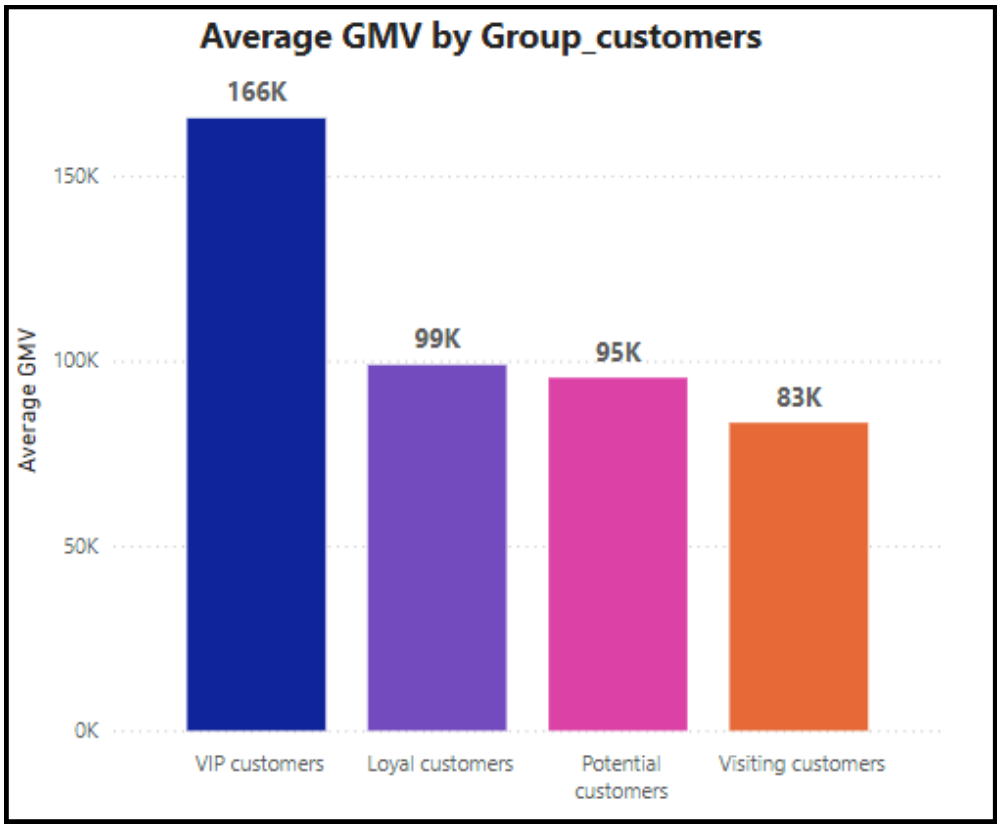




- Overview:
  - The total number of customers is **114.08 thousand**, with more than half (**66.65 thousand customers**) classified as **Visiting Customers**. The second largest group is **Potential Customers** (**22.61 thousand customers**), followed by **VIP Customers** (**14.45 thousand customers**), and finally **Loyal Customers** (**10.38 thousand customers**).



- The company's **total revenue** is approximately **11.12 billion**. With the majority of customers falling into the **Visiting Customers group**, nearly **half of the revenue (5.55 billion)** comes from this segment. Both **VIP Customers** and **Potential Customers** contribute **over 2 billion in revenue** each, while the **Loyal Customers** group generates the lowest revenue, **around 1 billion.**



- **VIP customers:**
  - Although ranking third in the number of customers, the VIP customer segment generates the second-highest revenue. On average, each customer in this segment contributes 166,000, significantly outperforming other customer groups.
  - This customer group exhibits high spending and purchase frequency, making it essential to focus on nurturing and retaining them through dedicated care and attention.

- **Loyal customers:**
  - This is the customer group with the lowest number of customers and the least revenue contribution to the store. However, on average, customers in the Loyal segment spend significantly at the store. This segment represents a high-potential customer base with strong spending capacity, but additional measures and campaigns are needed to drive growth in customer numbers within this group.

- **Potential customers:**
  - This customer group has a proportional balance between its size and revenue contribution. On average, customers in the Potential segment spend nearly as much as Loyal customers. With effective strategies to encourage this group, it is highly likely to drive increased revenue for the store in the future.

- **Visiting customers:**
  - This is a unique customer group, as they typically make purchases with low order frequency and value. However, they account for the largest proportion of the customer and revenue. This indicates that while the store has a significant influx of new customers, it has yet to effectively retain them. Implementing policies, programs, and incentives to convert these new customers into loyal ones is both necessary and urgent.

SQL CODE

```sql
-- Choosing the report date
SET @report_date = '2022-09-01'
-- rfm calculation
CREATE TABLE rfm_project.rfm_statistics (WITH rfm_table AS (
SELECT ct.customerid
-- "recency" calculation
,DATEDIFF(@report_date, MAX(ct.purchase_date)) AS 'recency'
-- "frequency" calculation
, round(1.0 *(count(DISTINCT
ct.purchase_date))/datediff(@report_date,cr.created_date),4) AS 'frequency'
-- "monetary" calculation
, round(1.0* sum(gmv)/DATEDIFF(@report_date,cr.created_date),4) AS 'monetary'
       FROM customer_transaction ct
    JOIN customer_registered cr
    ON ct.customerid = cr.id
-- Filter out customers whose stopdate is different "NULL"
    WHERE cr.stopdate = ""
    GROUP BY ct.customerid),
-- rfm_statistics
rfm_statistics AS (
SELECT rfm.customerid, rfm.recency, rfm.frequency, rfm.monetary
-- arrange the order of R
,ROW_NUMBER() OVER (ORDER BY rfm.recency DESC) AS 'rank_recency'
-- arrange the order of F
,ROW_NUMBER() OVER (ORDER BY rfm.frequency ASC) AS 'rank_frequency'
-- arrange the order of M
,ROW_NUMBER() OVER (ORDER BY rfm.monetary ASC) AS 'rank_monetary'
FROM rfm_table rfm
GROUP BY rfm.customerid)
SELECT * FROM rfm_statistics);
-- Divided by IQR(interquartile range)
WITH RFM_IQR AS (
SELECT r.CustomerID, r.recency, r.frequency, r.monetary,
-- Recency_IQR
CASE
       WHEN r.recency <= (SELECT max(rs.recency) FROM rfm_statistics rs)
              AND r.recency >= (SELECT rs.recency FROM rfm_statistics rs WHERE
rs.rank_recency = (SELECT round(count(rs.customerid)*0.25,0) FROM rfm_statistics rs))
              THEN 1
       WHEN r.recency < (SELECT rs.recency FROM rfm_statistics rs WHERE rs.rank_recency
= (SELECT round(count(rs.customerid)*0.25,0) FROM rfm_statistics rs))
              AND r.recency >= (SELECT rs.recency FROM rfm_statistics rs WHERE
rs.rank_recency = (SELECT round(count(rs.customerid)*0.5,0) FROM rfm_statistics rs))
              THEN 2
       WHEN r.recency < (SELECT rs.recency FROM rfm_statistics rs WHERE rs.rank_recency
= (SELECT round(count(rs.customerid)*0.5,0) FROM rfm_statistics rs))
              AND r.recency >= (SELECT rs.recency FROM rfm_statistics rs WHERE
rs.rank_recency = (SELECT round(count(rs.customerid)*0.75,0) FROM rfm_statistics rs))
              THEN 3
       ELSE 4 END AS "R",
-- Frequency_IQR
CASE
       WHEN r.frequency >= (SELECT min(rs.frequency) FROM rfm_statistics rs)
              AND r.frequency <= (SELECT rs.frequency FROM rfm_statistics rs WHERE
rs.rank_frequency = (SELECT round(count(rs.customerid)*0.25,0) FROM rfm_statistics rs))
              THEN 1
       WHEN r.frequency > (SELECT rs.frequency FROM rfm_statistics rs WHERE
rs.rank_frequency = (SELECT round(count(rs.customerid)*0.25,0) FROM rfm_statistics rs))
```

```
            AND r.frequency <= (SELECT rs.frequency FROM rfm_statistics rs WHERE
rs.rank_frequency = (SELECT round(count(rs.customerid)*0.5,0) FROM rfm_statistics rs))
            THEN 2
       WHEN r.frequency > (SELECT rs.frequency FROM rfm_statistics rs WHERE
rs.rank_frequency = (SELECT round(count(rs.customerid)*0.5,0) FROM rfm_statistics rs))
            AND r.frequency <= (SELECT rs.frequency FROM rfm_statistics rs WHERE
rs.rank_frequency = (SELECT round(count(rs.customerid)*0.75,0) FROM rfm_statistics rs))
            THEN 3
       ELSE 4 END AS "F",
-- Monetary_IQR
CASE
       WHEN r.monetary >= (SELECT min(rs.monetary) FROM rfm_statistics rs)
            AND r.monetary <= (SELECT rs.monetary FROM rfm_statistics rs WHERE
rs.rank_monetary = (SELECT round(count(rs.customerid)*0.25,0) FROM rfm_statistics rs))
            THEN 1
       WHEN r.monetary > (SELECT rs.monetary FROM rfm_statistics rs WHERE
rs.rank_monetary = (SELECT round(count(rs.customerid)*0.25,0) FROM rfm_statistics rs))
            AND r.monetary <= (SELECT rs.monetary FROM rfm_statistics rs WHERE
rs.rank_monetary = (SELECT round(count(rs.customerid)*0.5,0) FROM rfm_statistics rs))
            THEN 2
       WHEN r.monetary > (SELECT rs.monetary FROM rfm_statistics rs WHERE
rs.rank_monetary = (SELECT round(count(rs.customerid)*0.5,0) FROM rfm_statistics rs))
            AND r.monetary <= (SELECT rs.monetary FROM rfm_statistics rs WHERE
rs.rank_monetary = (SELECT round(count(rs.customerid)*0.75,0) FROM rfm_statistics rs))
            THEN 3
            ELSE 4 END AS "M"
       FROM rfm_statistics r),
-- CONCAT RFM IQR
CONCAT_RFM_IQR AS (SELECT concat(r.R,r.F,r.M) AS 'concat_IQR', count(r.CustomerID) AS
'count'
            FROM RFM_IQR r
            GROUP BY r.R,r.F,r.M)
-- Group customers
SELECT CASE
       WHEN c.concat_IQR IN ('444', '443', '434', '344', '334')
            THEN 'VIP customers'
       WHEN c.concat_IQR IN ('433', '424', '423', '414', '413', '343', '333', '324',
'323','314', '313')
            THEN 'Loyal customers'
       WHEN c.concat_IQR IN ('442', '432', '422', '412', '342', '332', '322', '312',
'242', '233', '244', '243', '234', '224', '223', '214', '213')
            THEN 'Potential customers'
       ELSE 'Visiting customers'
       END AS 'Group_customers', sum(c.count) AS 'count'
       FROM CONCAT_RFM_IQR c
       GROUP BY Group_customers;
```

THE END