AUTOMOBILE PARTS MANUFACTURER

MRA PROJECT 1

BY GIRIDHARAN VELMURUGAN

GREAT LEARNING

PROBLEM STATEMENT

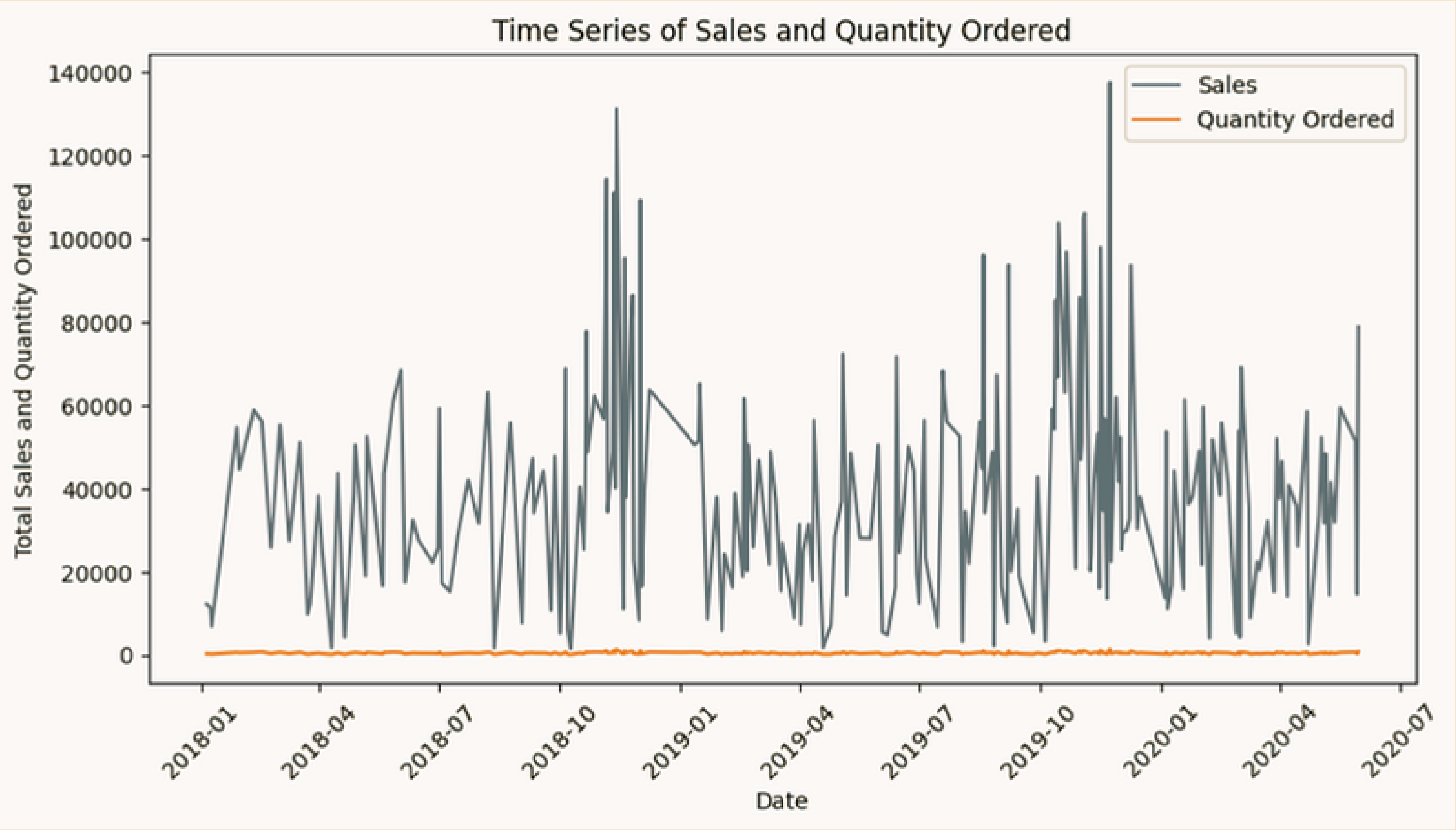
An automobile parts manufacturing company has collected data on transactions for 3 years. They do not have any in-house data science team, thus they have hired you as their consultant. Your job is to use your data science skills to find the underlying buying patterns of the customers, provide the company with suitable insights about their customers, and recommend customized marketing strategies for different segments of customers.

DATA OVERVIEW

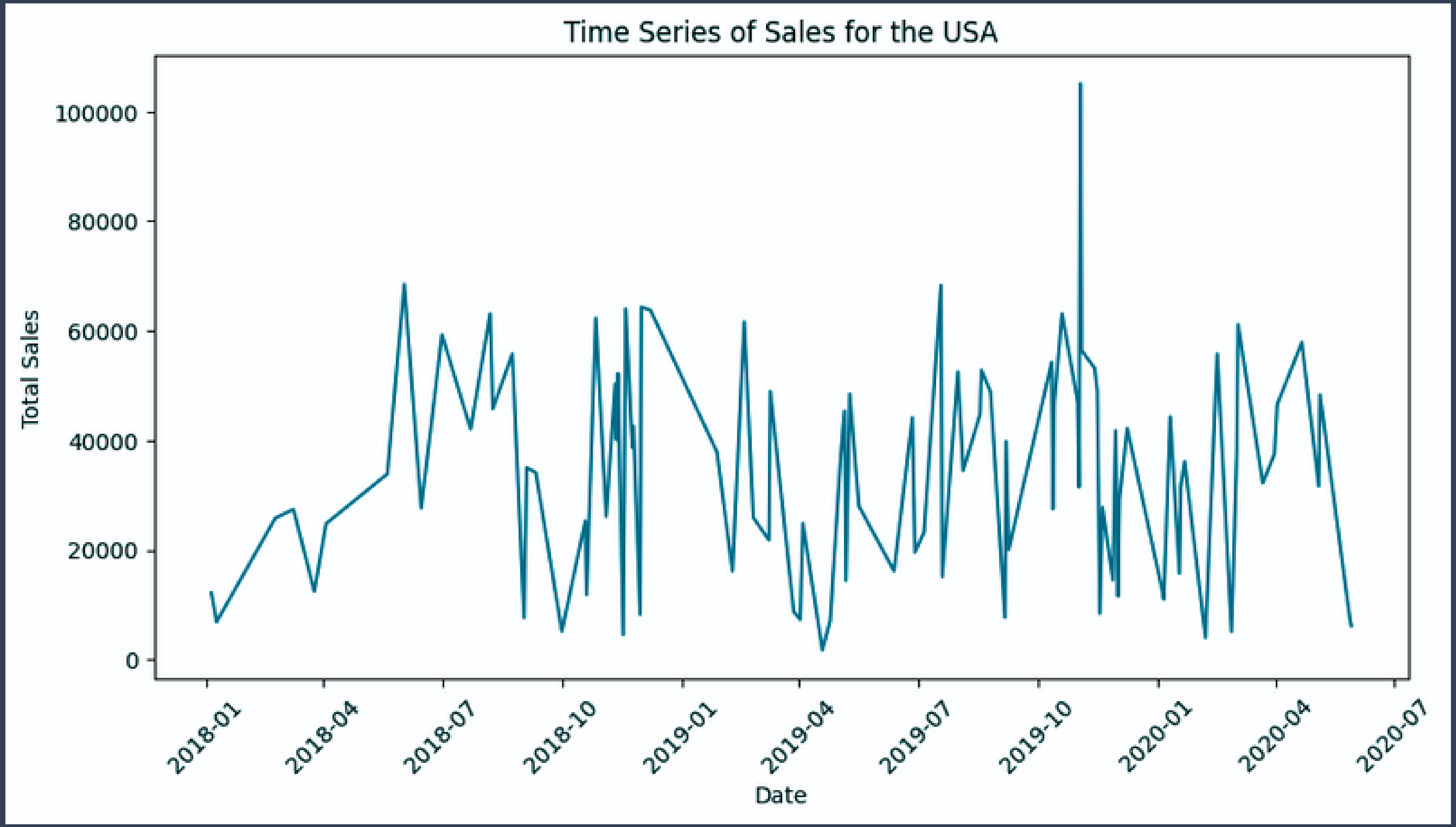
- Dataset has only int(64), float(64) and Object datatypes.
- There were no duplicates in the dataset, which is a boon.
- There are no null values within the features, all the columns have data and there are a total of 2747 instances per feature
- There are a total of 20 features(columns)
- Our data spans the whole globe, covering important cities and countries
- There are multiple products from covering land, air and water with corresponding product codes
- There are products in all types of shipping status.
- We have customers who have placed multiple orders.

EXPLORATORY DATA ANALYSIS

- The image on the right is a time-series showing sales that has spanned over the years.
- We can see variations throughout, what is very apparent is that as the holiday season approaches, the sales increase in folds.
- Post new year, we see a dip in sales.
- The line in orange is quantity ordered over the same timespan, added for reference.
- The peak sales is clearly at December at close to 140000, second place is also very close to that.



SALES ACROSS USA(our Largest market)

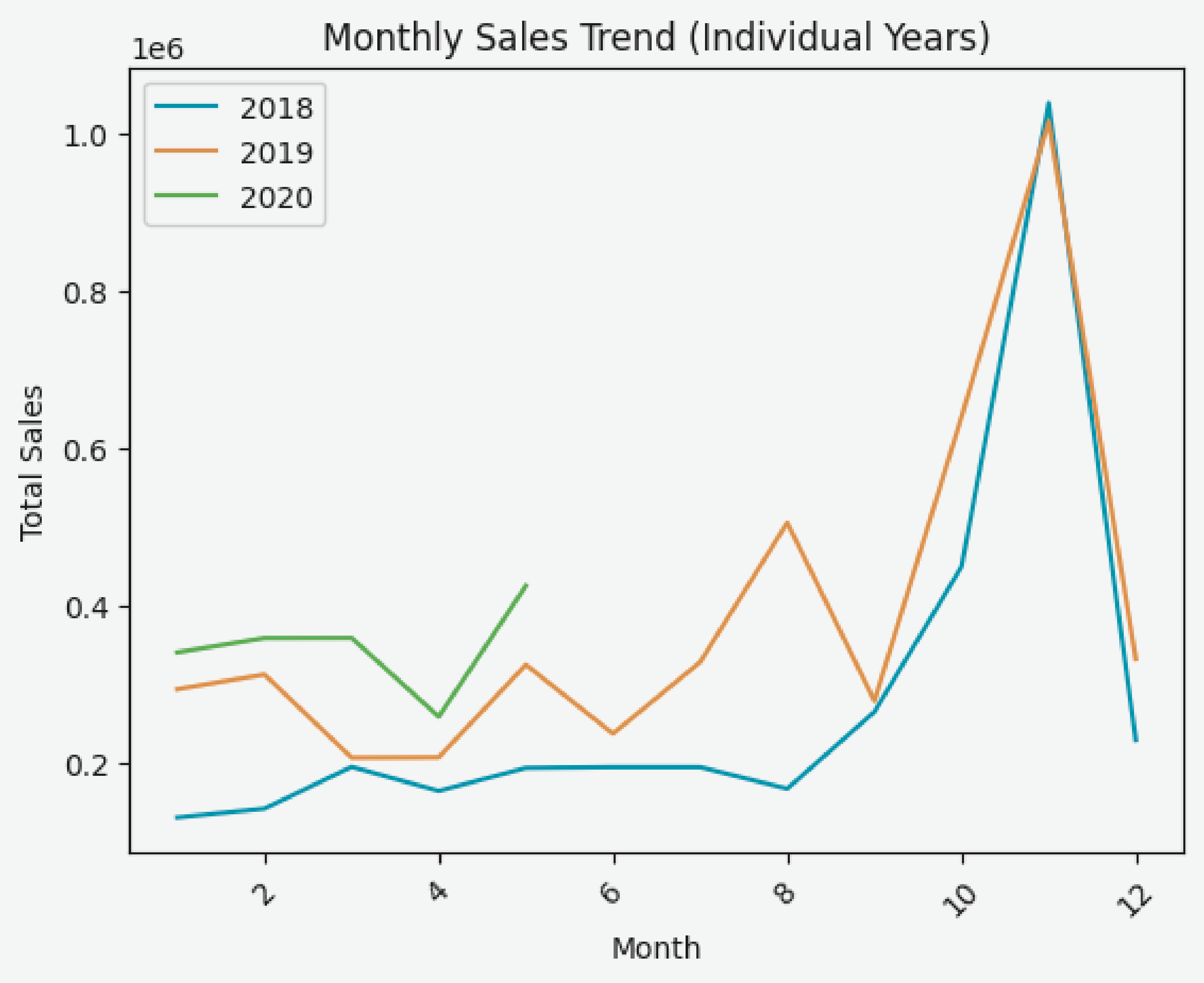


- The country USA takes the top spot close to 930 orders and sales in the region of 3.5M.
- The trend is very close to the yearly curve.
- There seems to be a downfall post May(end of summer), but it again picks up around October.

- USA, being our largest market share we should prioritize them and make sure to bring in new products that specifically caters to their needs.

MONTHLY SALES

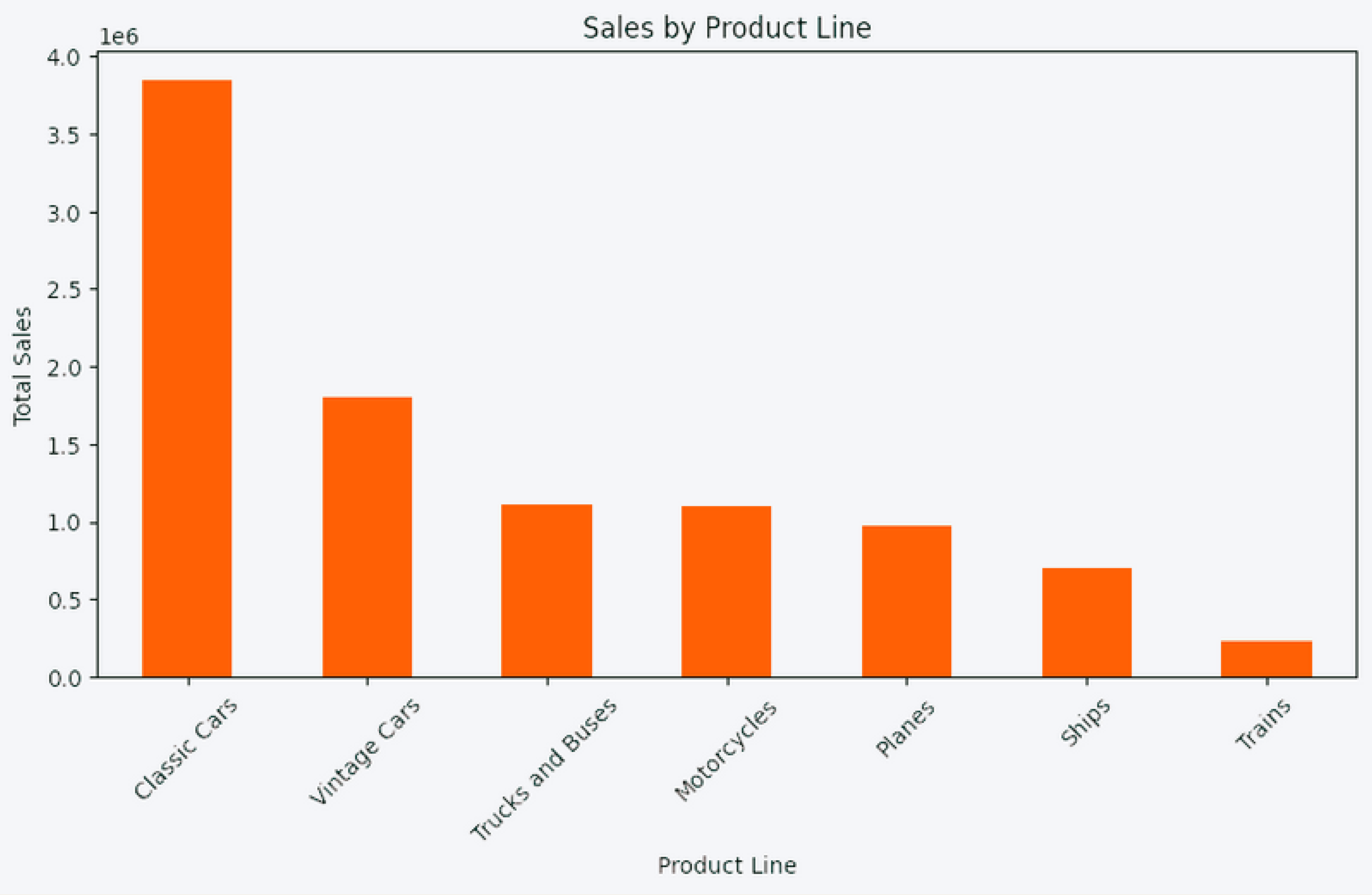
- The monthly trend further cements our findings that as the year approaches the holiday season and its end we can see a massive spike in sales of parts.
- There seems to be a dip here and there throughout and the overall terrain is never smooth.
- The above situation has to be understood even deeper to find the ‘why’ so that we can improve



- There seems to be a growth irrespective of the month, as if we compare January of 2018 to 2019 and through 2020, we see upward trend.
- No drop in sales is less than the same month in the previous year.
- 2020 has data only for a few months hence we see a cut-off in the trend line.

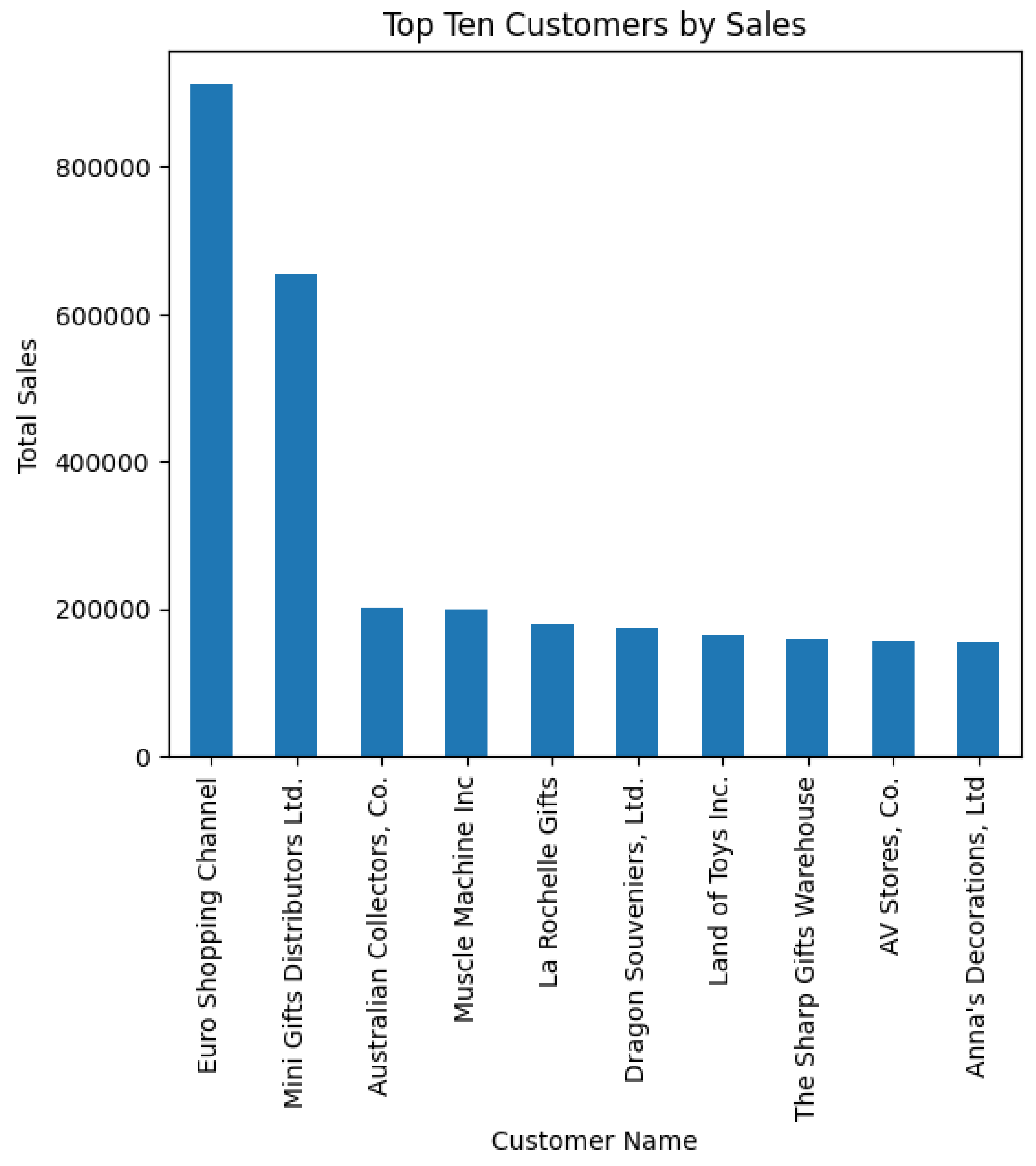
PRODUCT LINE - SALES

- Classis cars is almost double of vintage cars which is the second one in line.
- Our least product category is trains which is okay.
- Surprisingly, trucks and busses are neck and neck with motorcycles. This has to be studied further.
- The category of cars alone(vintage plus classic) make a majority of our sales and this must be prioritized.



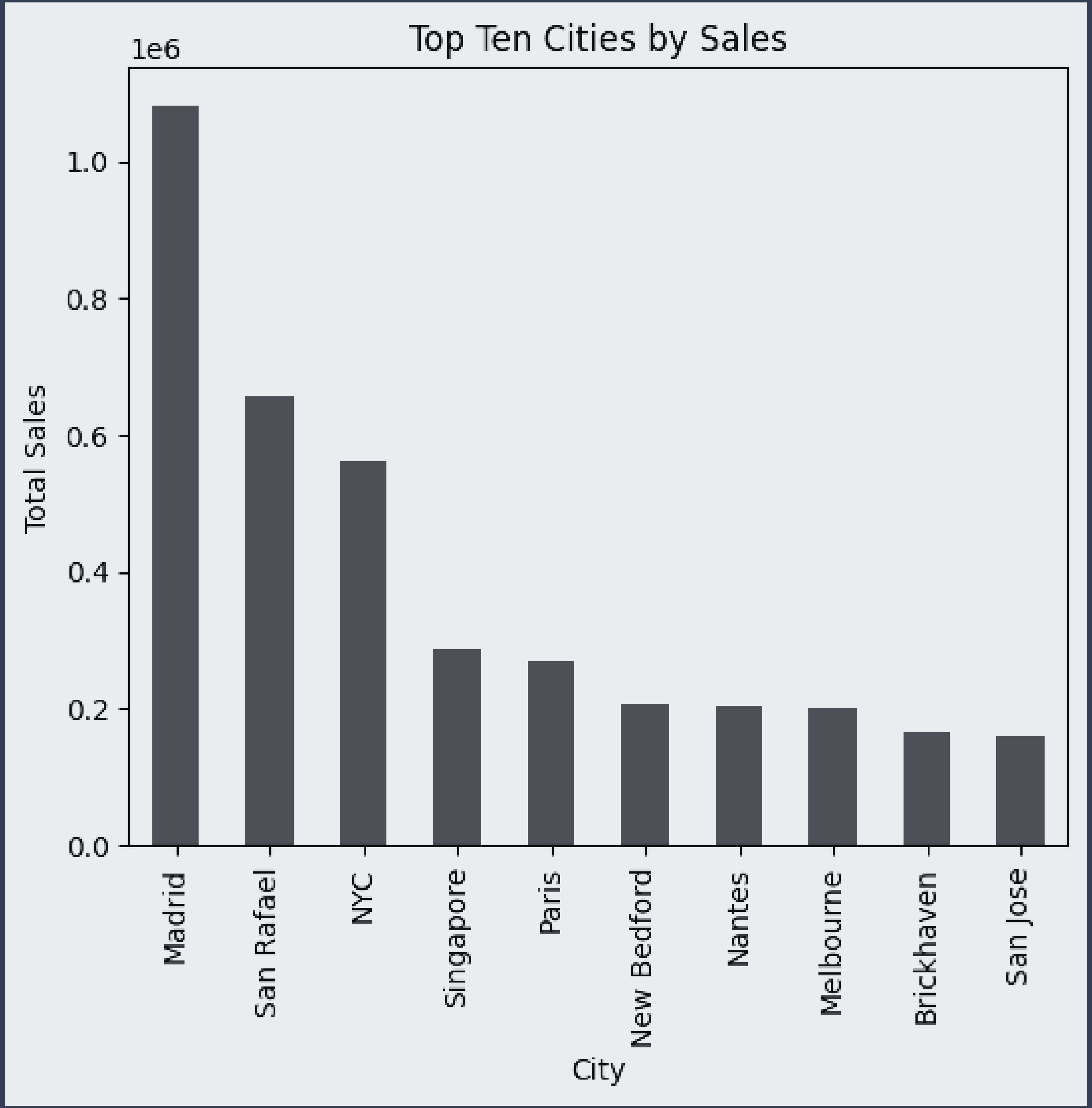
TOP TEN CUSTOMERS

- The first two in the given list alone account for more than 1.4 million plus.
- Rest of our customers do not cross more than 200k.
- We must analyze why this is the case and see what can be done about this.
- Strategies to bring improve our stake to other clients is mandatory.



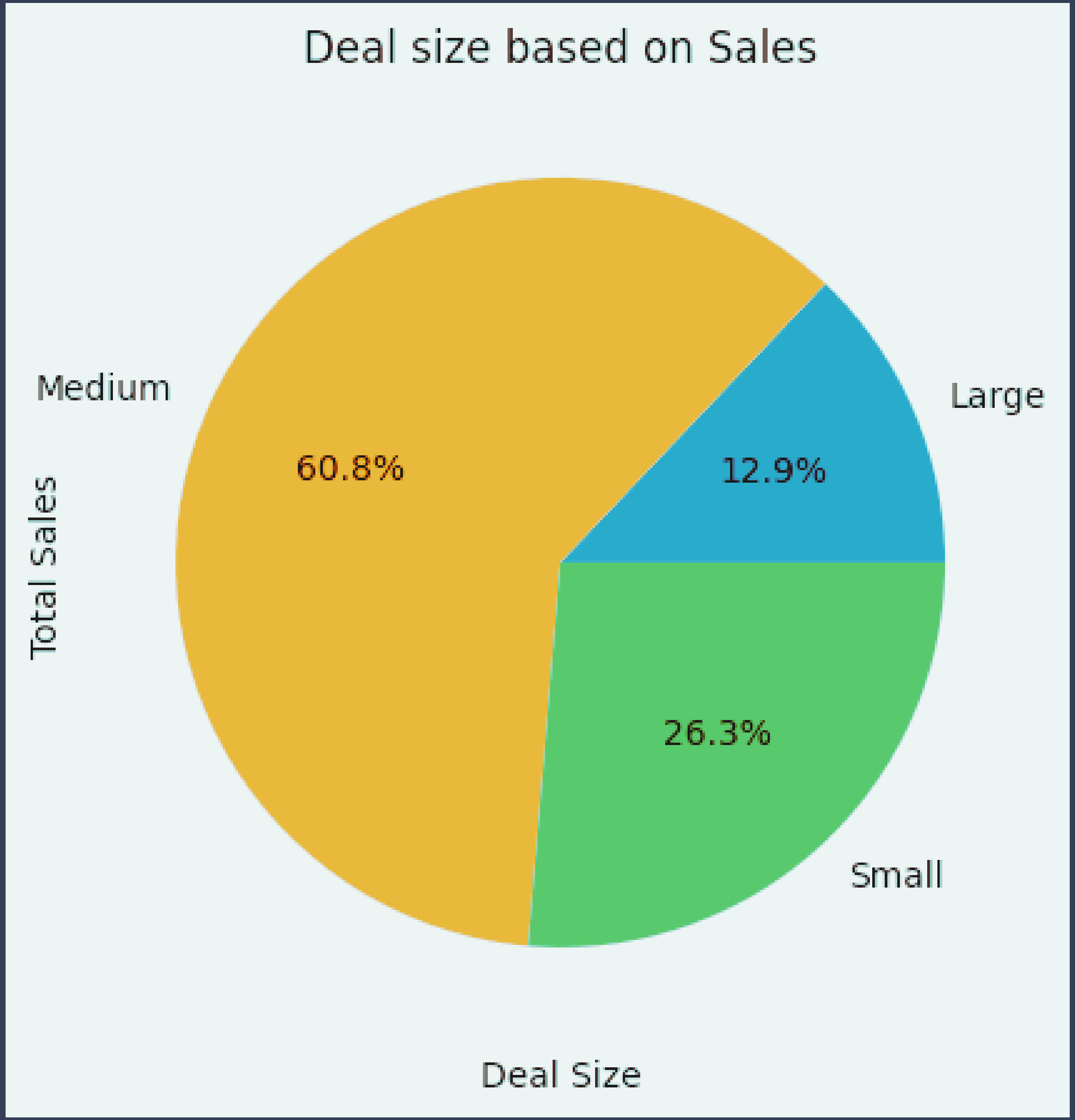
TOP TEN CITIES

- Madrid is our top city followed by San Rafael(California).
- 5 cities from USA have are in the top 10 globally, this is in line with USA being our top destination.
- The surprising thing to see is that Singapore although being small in population has mustered 4th place.
- There is a difference of around 300k between first and second places.

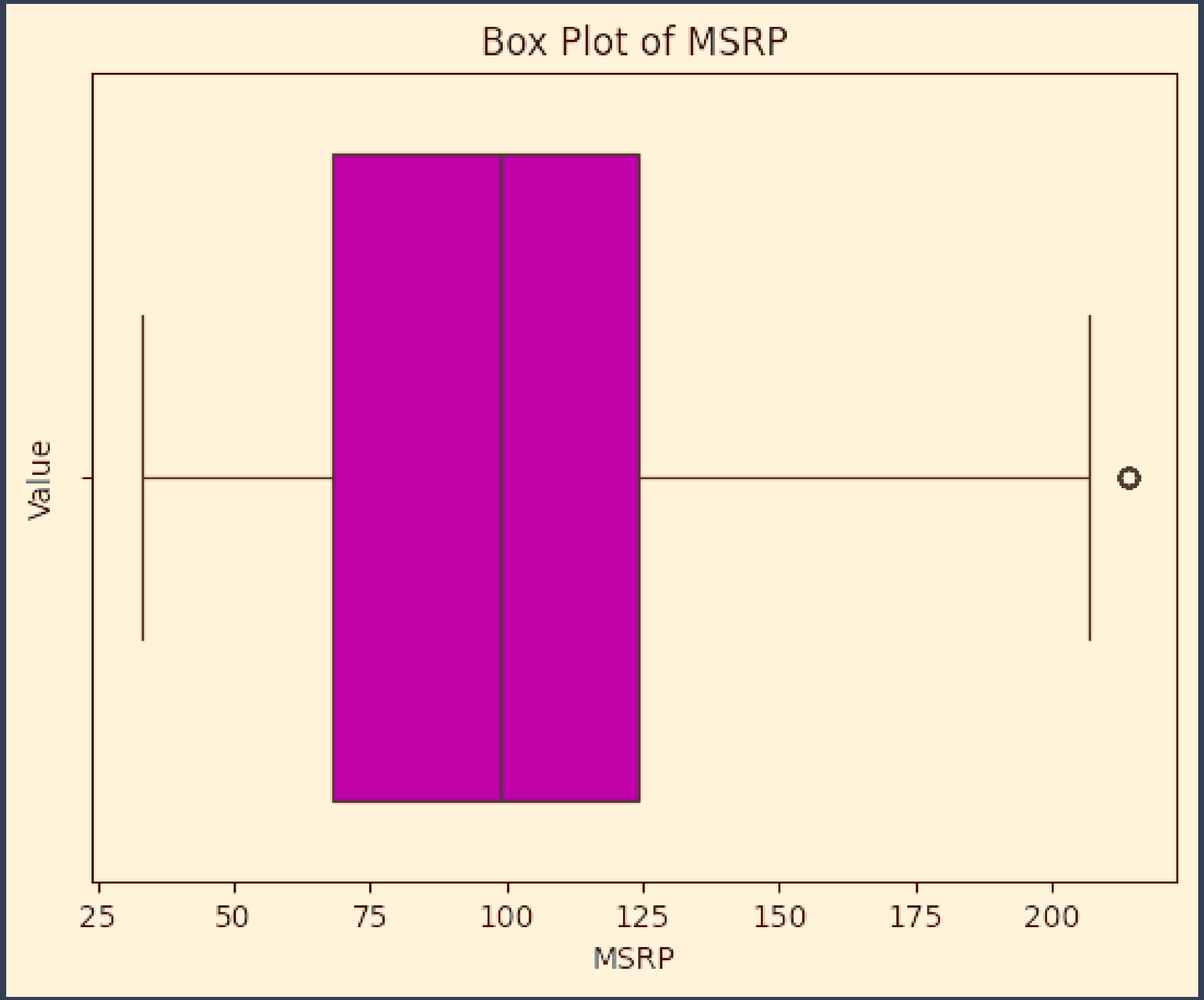


DEAL SIZE IN TERMS OF SALES

- Medium has the lion share at just over 60%.
- Second is Small at 26%.
- The above two alone rakes in more than 87% in market share.
- Our large orders are very less and we need to dig in deeper to see what is lacking in large orders over medium.



MSRP - Its edges and whiskers



- MSRP is surprisingly on the low end with the mean at only 100.
- With the max limit at just above 200, we have only one outlier.
- Our lower whisker stands at above 25(or close 27).
- The majority of our products are binned around the 60 to 125 range.

WHAT IS "RFM" ANALYSIS

RECENCY

FREQUENCY

MONETARY

RFM stands for Recency, Frequency, and Monetary value, It's a customer segmentation technique used in marketing to group customers based on their purchasing behavior across these three dimensions

THE THREE ELEMENT'S

RECENCY

Time elapsed since the customer's last purchase.
It's a measure of time since the last purchase and implies the freshness of the customer's engagement.

FREQUENCY

Number of purchases made by the customer within a defined period.

It's a count of transactions within a certain period.

MONETARY

Total amount spent by the customer across all purchases.

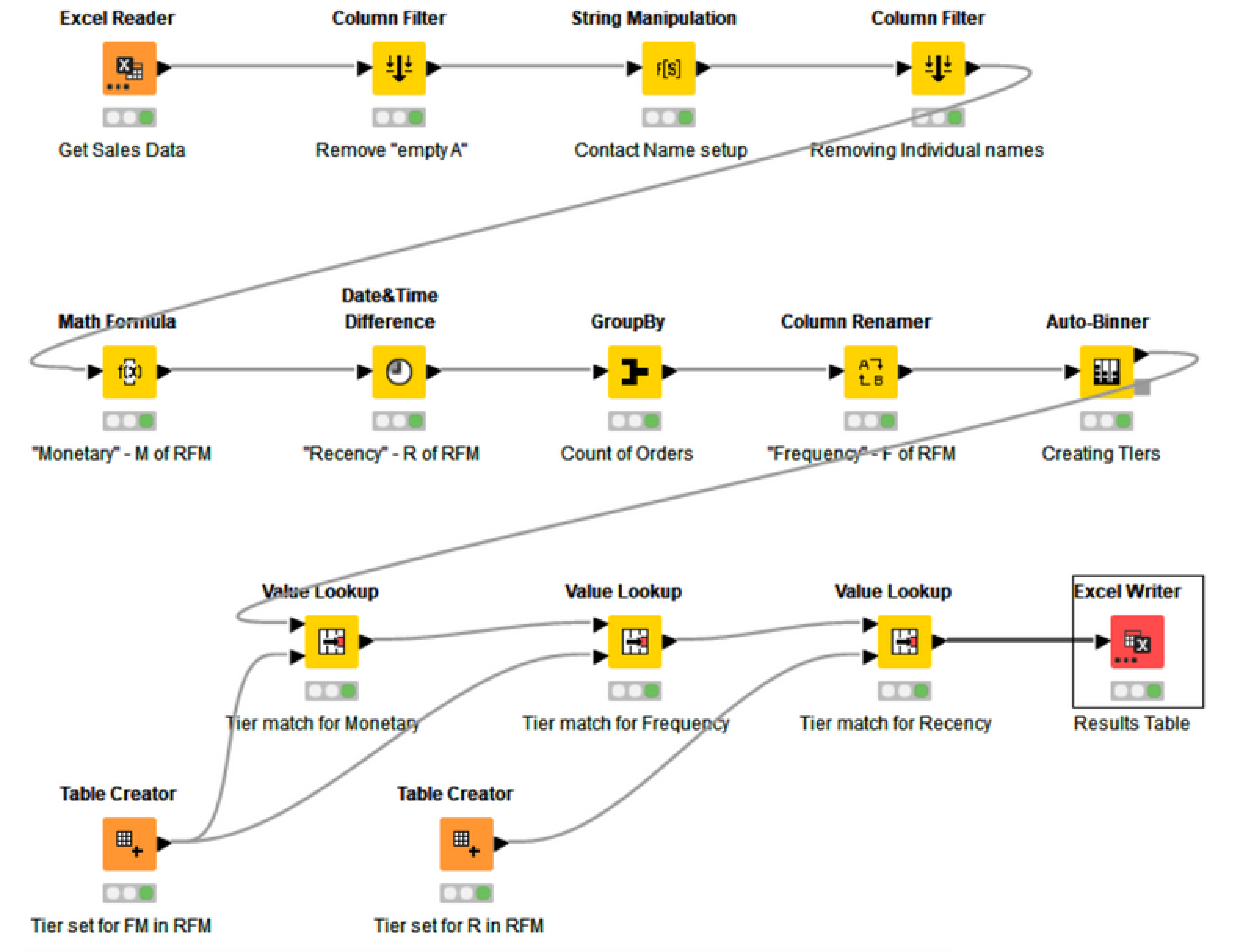
It's a sum of all transaction values from a customer.

ASSUMPTIONS IN “ RFM ”

1. ***Customers who recently made a purchase are more likely to buy again*** compared to those who haven't bought in a sometime. This targets marketing efforts on users who were relatively recent.
2. ***Customers who buy frequently are more engaged and potentially more loyal*** than those who purchase less often. They're seen as having a stronger relationship with the brand.
3. ***Customers who spend more are more valuable*** to the business from a revenue perspective. This can help in identifying high-value customers for specific offers.

“RFM” USING KNIME

- KNIME is a user friendly, powerful, flexible Data science tool, which is used here to perform RFM.
- With KNIME and its Nodes, we have brought in our dataset which holds the required values. With some preprocessing we have(like “contact Name setup) primed the data.
- Next up we have started to create the Three Dimensions that are CORE to RFM.
- The final result is the output table from the node “Results Table”.
- The following slide shows the output.



RESULT'S TABLE - SNAP SHOT

CONTACT NAME	FREQUENCY	MONETARY	RECENCY	FREQUENCY [Binned]	MONETARY [Binned]	RECENCY [Binned]
Yu, Kwai	4	164069.44	1545	Tier 4	Tier 4	Tier 2
Henriot, Paul	5	135042.94	1409	Tier 4	Tier 4	Tier 4
Da Cunha, Daniel	3	78570.34	1422	Tier 2	Tier 2	Tier 4
Young, Julie	3	104561.96	1487	Tier 2	Tier 3	Tier 3
Hirano, Juri	4	120783.07	1494	Tier 4	Tier 4	Tier 3
Rance, Martine	2	69052.41	1812	Tier 1	Tier 1	Tier 1
Oeztan, Veysel	3	111640.28	1618	Tier 2	Tier 3	Tier 1
Perrier, Dominique	3	93170.66	1401	Tier 2	Tier 3	Tier 4
Ferguson, Peter	5	200995.41	1531	Tier 4	Tier 4	Tier 3
Frick, Michael	3	88041.26	1555	Tier 2	Tier 3	Tier 2
Brown, William	3	83228.19	1405	Tier 2	Tier 2	Tier 4
King, Julie	3	101894.79	1373	Tier 2	Tier 3	Tier 4
Labrune, Janine	4	180124.9	1347	Tier 4	Tier 4	Tier 4
Hernandez, Marta	2	103080.38	1578	Tier 1	Tier 3	Tier 1
Karttunen, Matti	3	111250.38	1459	Tier 2	Tier 3	Tier 3
Bergulfsen, Jonas	4	116599.19	1555	Tier 4	Tier 3	Tier 2

NOTE - The above results table only includes columns that are absolutely necessary for RFM. Contact name has been prioritized instead of Company name(aka Customer name). Contact name has also been restructured to suit requirements.

INFERENCES FROM RFM

TOP 5 BEST CUSTOMERS

The table on the right shows us the top 5 customers.

The sorting method was “Higher Frequency”, “Higher Monetary value” and “Most Recent”. These customers are not only spending a lot but they are also engaged recently. Keeping them in with us is top priority always.

Rank	Contact Name	Frequency	Monetary Value	Recency (Days)
1	Diego Freyre	26	$912,294.11	1347
2	Valarie Nelson	17	$654,858.06	1349
3	Labrune, Janine	4	$180,124.90	1347
4	Frick, Sue	4	$160,010.27	1386
5	Huxley, Adrian	4	$151,570.98	1349

INFERENCES FROM RFM

TOP 5 CUSTOMERS ON THE VERGE OF CHURN

The table on the right shows us the top 5 customers who are about to fall of from our boat.

The sorting method was “Higher Frequency”, “Higher Monetary value” and “Oldest in terms of engagement”. These customers “were” spending a lot but they have for have become quite in recent times. If we focus and nudge them in, there is significant probability that they can grow to be more valuable.

Rank	Contact Name	Frequency	Monetary Value	Recency (Days)
1	Ferguson, Peter	5	$200,995.41	1531
2	Young, Jeff	4	$197,736.94	1529
3	Natividad, Eric	5	$172,989.68	1437
4	Yu, Kwai	4	$164,069.44	1545
5	O'Hara, Anna	4	$153,996.13	1430

INFERENCES FROM RFM

CUSTOMERS WHOM WE HAVE LOST

The table on the right shows us the top 5 customers who have left us long ago.

Their spending and other metrics are opposite of what your would generally want to see.

These customers are not significant enough to be spent more money to keep them in our ring of customers.

Rank	Contact Name	Frequency	Monetary Value	Recency (Days)
1	Donnermeyer, Michael	1	$34,993.92	1606
2	Hardy, Thomas	2	$36,019.04	1842
3	Tseng, Kyung	2	$36,163.62	1736
4	Thompson, Steve	2	$46,084.64	1835
5	Fernandez, Jesus	2	$49,642.05	1786

INFERENCES FROM RFM

OUR MOST LOYAL CUSTOMERS

The table on the right shows us our most loyal customers. Interesting thing to see is our best customers happen to be our most loyal customers too.

We must make sure they never leave us and have contingency retention policies in place.

Rank	Contact Name	Frequency	Monetary Value	Recency (Days)
1	Freyre, Diego	26	$912,294.11	1347
2	Nelson, Valarie	17	$654,858.06	1349
3	Labrune, Janine	4	$180,124.90	1347
4	Frick, Sue	4	$160,010.27	1386
5	Huxley, Adrian	4	$151,570.98	1349
6	Pipps, Georg	4	$149,798.63	1361
7	Petersen, Jytte	5	$145,041.60	1393
8	Henriot, Paul	5	$135,042.94	1409
9	Yu, Kyung	4	$122,138.14	1348
10	Shimamura, Akiko	4	$120,562.74	1386

INFERENCES FROM RFM

With the help of RFM we have found our best, loyal, the churners and also people who were once our customers.

Now we must develop targeted marketing campaigns that resonate with the distinct needs and behaviors of each RFM segment, like for our most loyal customers, consider loyalty programs or exclusive offers to reward their patronage.

Consider Re-Engagement Initiatives for those identified as at risk of churn, it can be a bit more personal. We must get surveys, to understand their needs and preferences better.

Optimize Product and Service Offerings, this will ensure that we have covered the breath and depth needed to navigate out customer base.

Invest in customer experience to strengthen our commitment to providing an exceptional customer experience at every interaction level, leveraging data to anticipate and meet customer needs proactively.

RFM analysis is not a one goal victory scenario, the goal posts keep on changing and from time to time we have to change our strategies and methods to find the best, loyal, churner and also people who have left us.

INFERENCES FROM RFM

On average our top customers have garnered around $36,500.00 per order for 4 to 5 times. This would automatically promote them to the top tier(provided they are recent). If we do find customer who have say ordered for more than said amount, we can edge them with more offerings and personalized experiences from the get go to grab their attention and loyalty.

The above per order size is also more or less consistent through out our customer base, determining their needs, analyzing frequency by how quickly they are coming back can help us understand how we can help them.

We will have some one time customers who make a ad-hock purchase, they can be weeded out in due course of time, but while that happens we can attempt to get their experience with us and what we have missed. This would be vital in determining if they can become loyalist’s in the future.

Customers whose frequency is at 3 are well placed to become our top customers, their future disposition will depend on how we guide them, hence real-time promotions and offerings must be put in place.

To maintain top customers and loyalists, we must partner up with companies that are our direct extensions in all directions, this not only increases traffic but also provides them the confidence of familiarity and comfort.

While the above methods can help us to retain existing customers and also present us with new ones, we must think of the bigger picture always and drive decisions and innovations in the direction.

GREAT LEARNING

THANK YOU

BY GIRIDHARAN VELMURUGAN