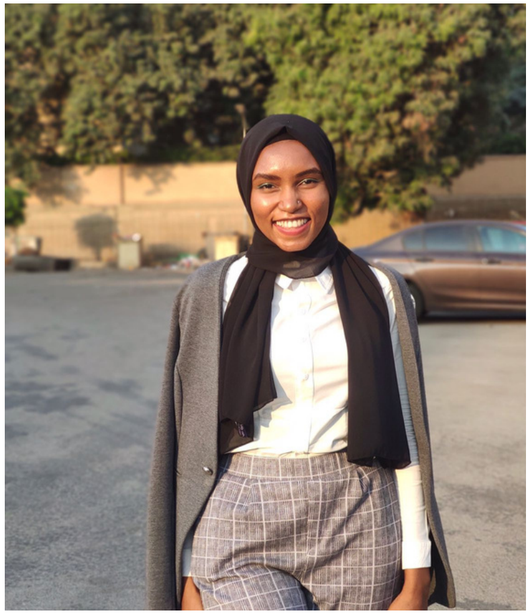
KYC

KNOW YOUR CUSTOMER

Taking your company to the next level

Team Members



Asmaa Adel
(Data Scientist)

Mohamed Essam
(Data Analyst)

Rech Raymondo
(Data Analyst)

Mohamed Hany
(Data Analyst)

Agenda

- Problem Definition
- Work Flow of the project
- Recap on phase 1
- RFM Analysis
- Machine Learning
- Dashboard
- STP Model
- Future Work
- Conclusion
- Time Frame

The Problem

Customer Segmentation is the process of dividing customers into distinct group according to their similar characteristics.

This process makes it easier to target specific groups of customers with tailored products, services, and marketing strategies.

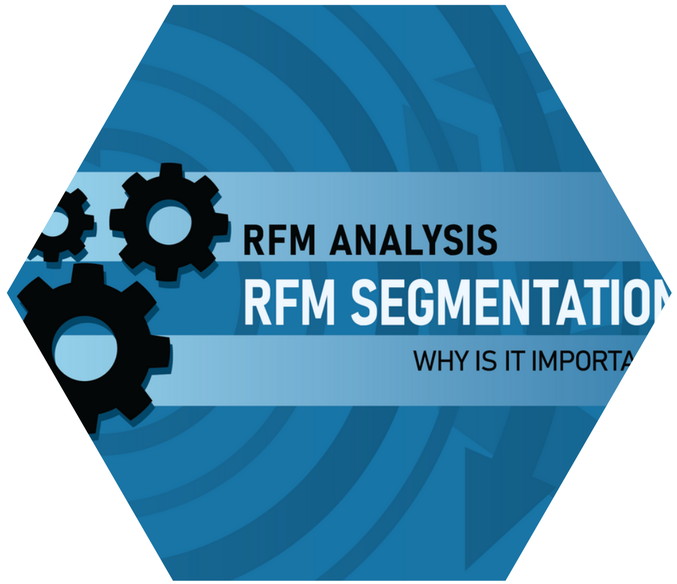
31%

Business fail due to ignoring their customers.

By using Data Science and Data Analysis Techniques we solved this problem.

Exploratory Data Analysis

Collect and Analyze data



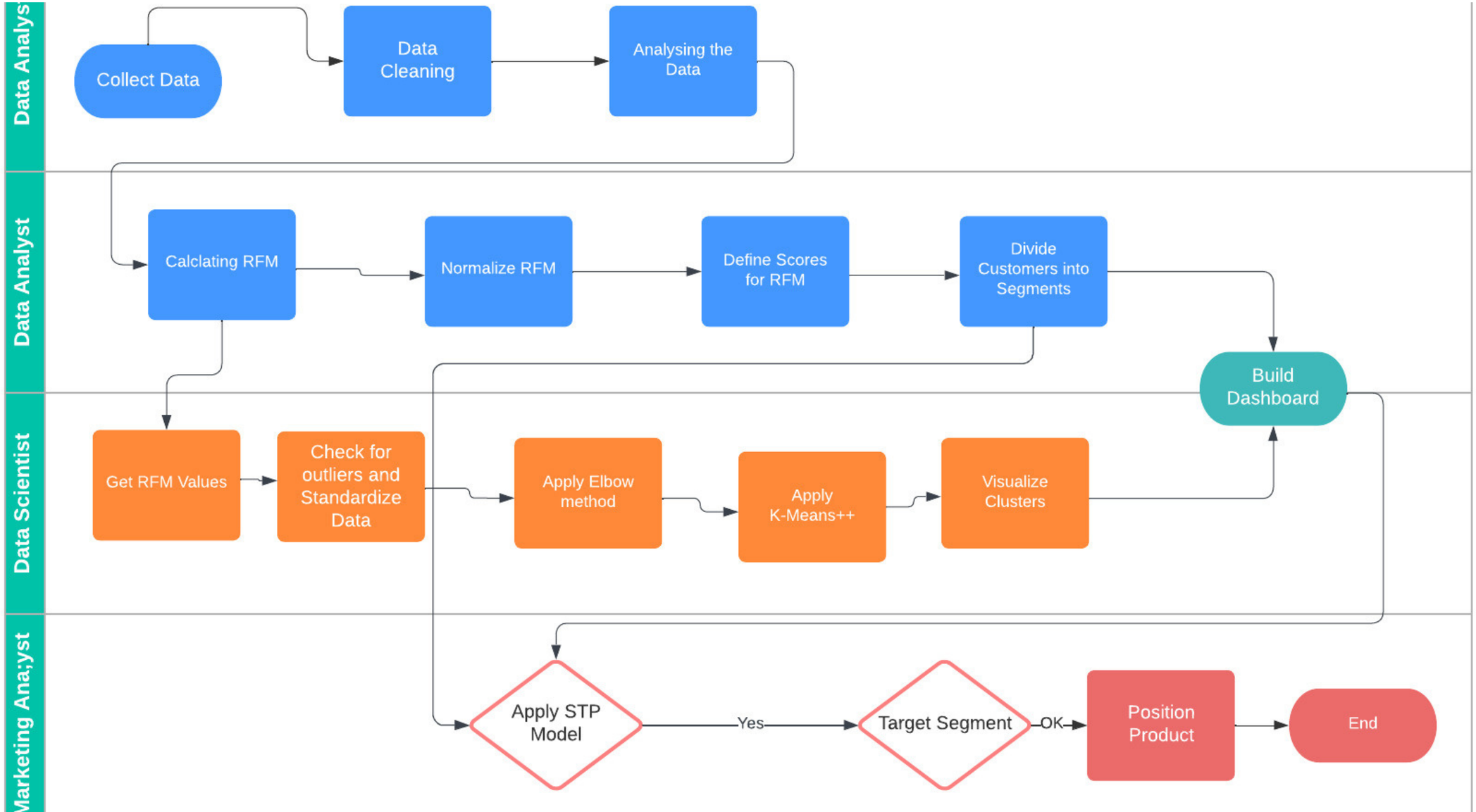
RFM Analysis

Perprocessing on data

Data Science

Apply Unsupervised Machine Learning

Work Flow of the project



Phase 1

- We applied Data Analysis and part of RFM Analysis.

- Data Collection
- Data Cleaning
- Analysing the data
- Calculate Recency, Frequency, Monetary
- Normalizing RFM values

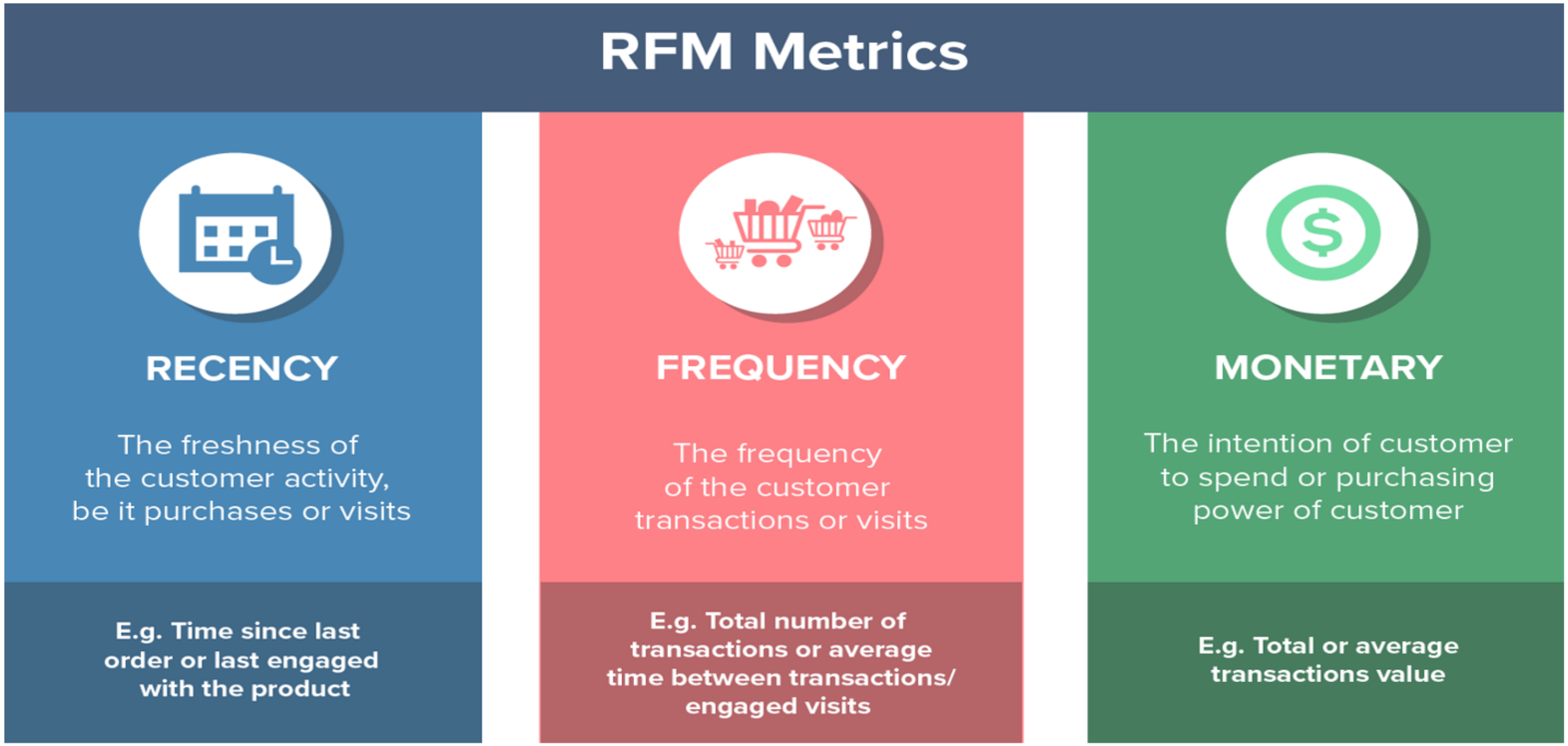
E-comerce Dataset

```
E_data.head()
```

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
0	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6.0	12/1/2010 8:26	2.55	17850.0	United Kingdom
1	536365	71053	WHITE METAL LANTERN	6.0	12/1/2010 8:26	3.39	17850.0	United Kingdom
2	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8.0	12/1/2010 8:26	2.75	17850.0	United Kingdom
3	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6.0	12/1/2010 8:26	3.39	17850.0	United Kingdom
4	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6.0	12/1/2010 8:26	3.39	17850.0	United Kingdom

Data Shape (541910,8),

RFM Metrics


RFM Metrics
RECENCY
The freshness of the customer activity, be it purchases or visits
E.g. Time since last order or last engaged with the product
FREQUENCY
The frequency of the customer transactions or visits
E.g. Total number of transactions or average time between transactions/ engaged visits
MONETARY
The intention of customer to spend or purchasing power of customer
E.g. Total or average transactions value

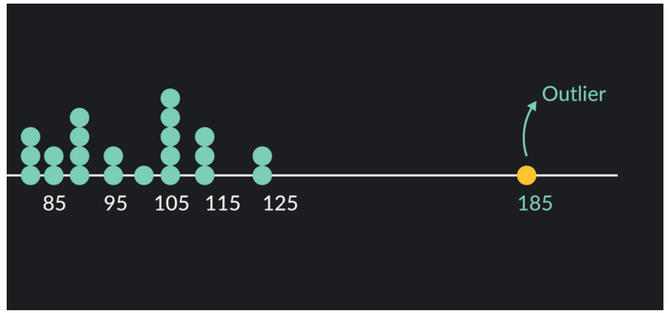
Create Segments according to scores

'...	index	CustomerID	Recency	Frequency	Monetary	R	F	M	RFMGroup	RFMScore	RFM_Loyalty_Level
0	0	12346.0	325	1	77183.60	4	4	1	441	9	Silver
1	1	12347.0	1	182	4310.00	1	1	1	111	3	Platinum
2	2	12348.0	74	31	1797.24	3	3	1	331	7	Gold
3	3	12349.0	18	73	1757.55	2	2	1	221	5	Platinum
4	4	12350.0	309	17	334.40	4	4	3	443	11	Bronze
5	5	12352.0	35	85	2506.04	2	2	1	221	5	Platinum

Machine Learning

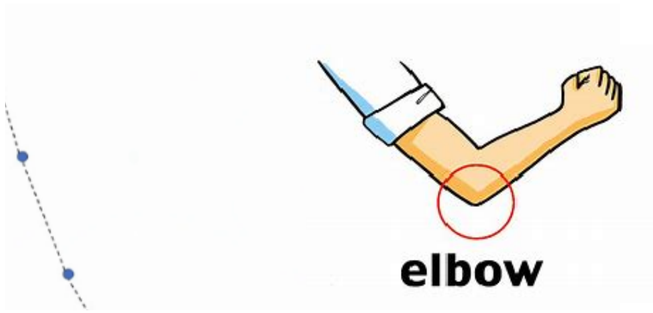
We used K-Means algorithm to cluster customers based on their purchasing behaviour, Clusters are based on 3 factors: Recency, Frequency, Monetary.

Machine Learning Model Steps



Check for outlier at the RFM Values

Use StandardScaler



Use Elbow method to detect number of K.

We got 3 clusters

Cluster 0

It contains the platinum who got 3,4 at score and gold that got 6 in score.

Cluster 1

It contains platinum with 5 at score and gold with 7 at score and silver with 9 at score and bronze with 11 at score

Cluster 2

It contains gold with 8 at score and silver with 10 at score and bronze with 12 at score.

Validation part

- We are focusing on internal cluster validation index
- We used Calinski Harabasz Index also known as Variance ratio criterion
- We got 3299.77 score

Dashboard

STP Model

Take your company to a higher level.

Segmenting

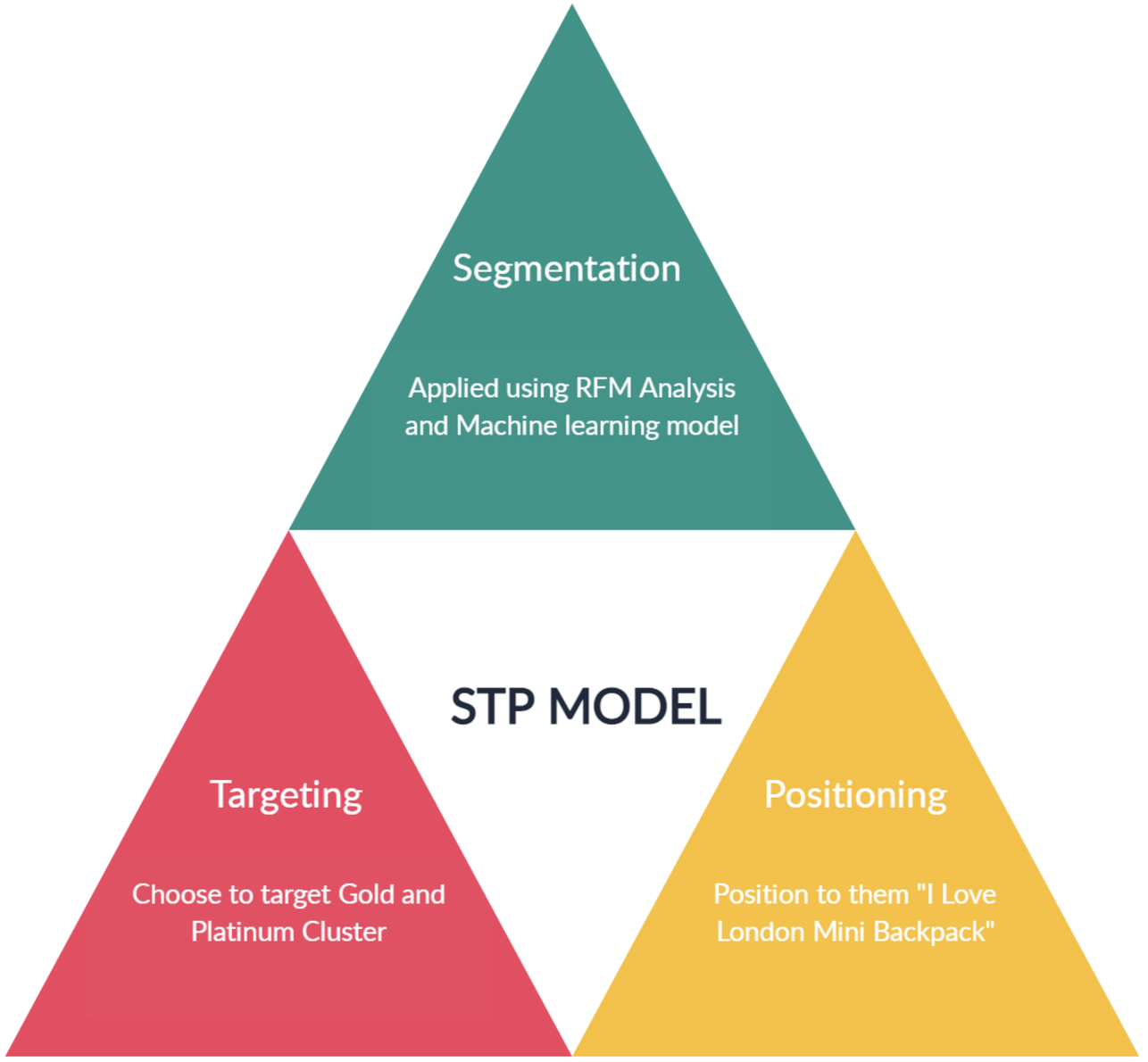
Targeting

Positioning

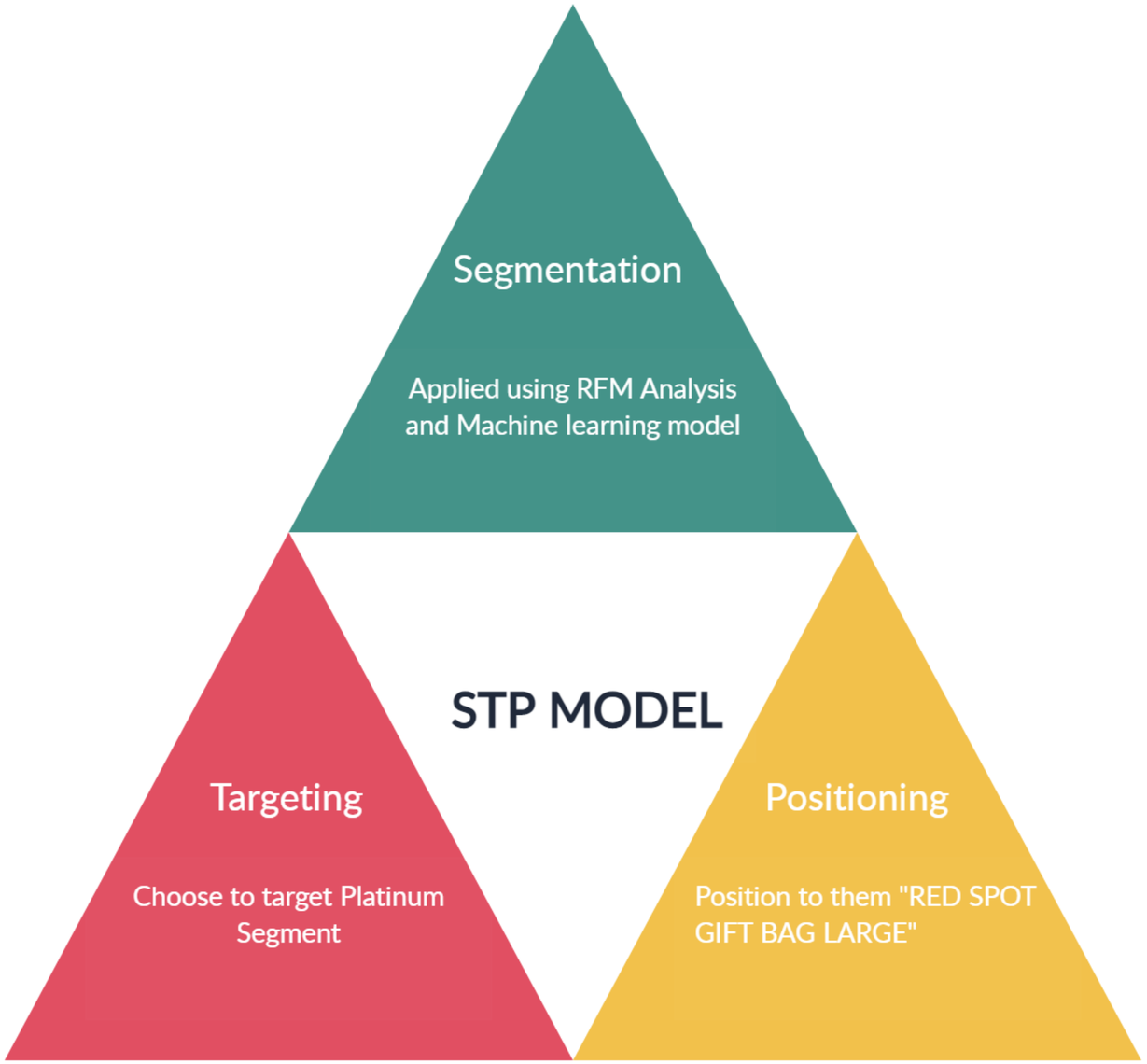
Sender

We have created 2 STP Models examples

First Model for gold and platinum cluster



Second model is for Platinum segment



Future work

Deploy our model

To be used for its intended purpose

Convert it into Web Application

To be submitted in Google developer solution Challenge as it serve two of the 17 UN sustainable development goals.

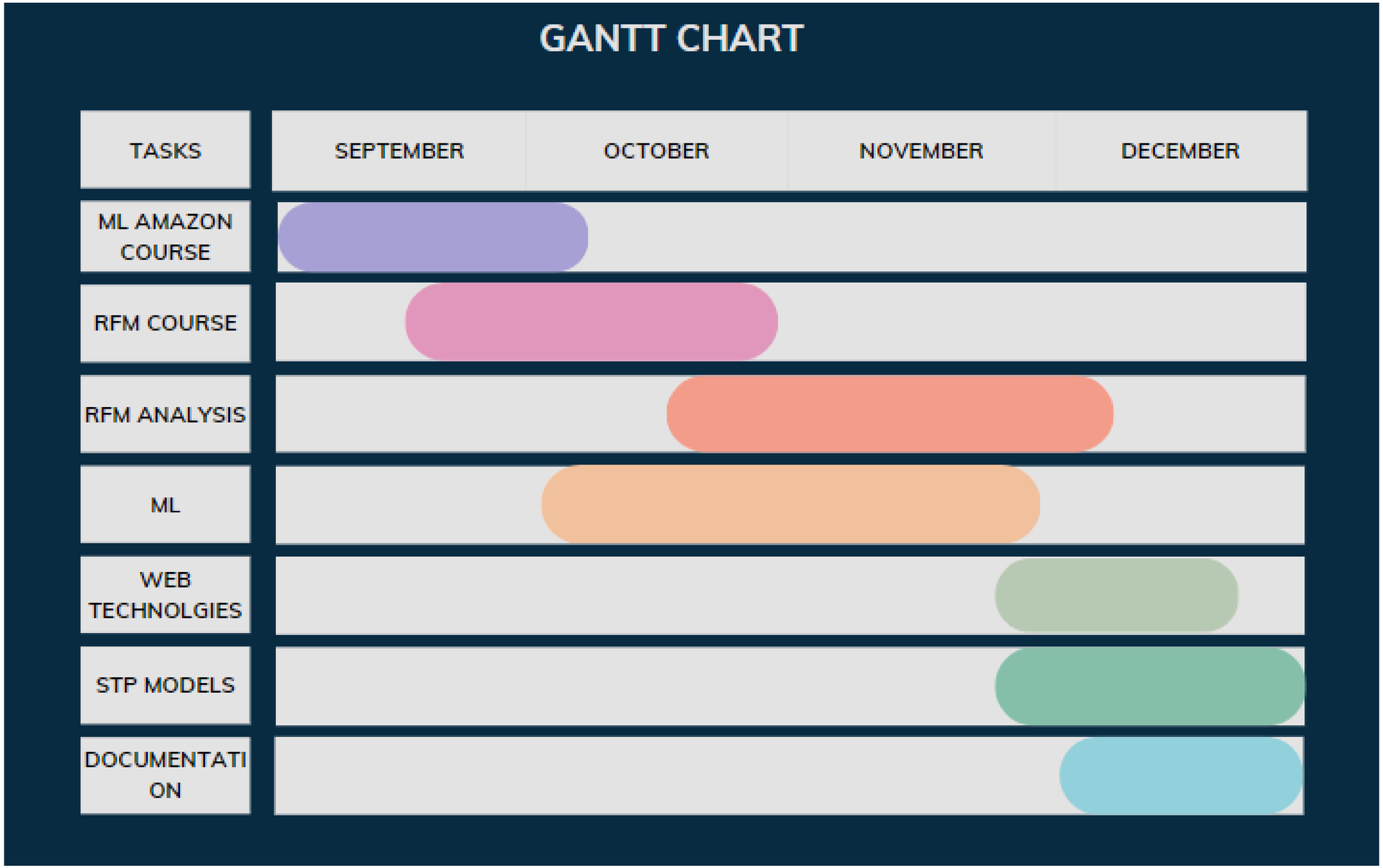
Create a business model

To be submitted in Google developer solution Challenge as it serves two of the 17 UN sustainable development goals.

Conclusion

Overall, the analysis shows that the company has a solid customer base, but there are areas for improvement in customer retention and satisfaction. By utilizing the insights gained from the analysis, the company can make data-driven decisions to improve its products, customer service, and marketing strategies to increase customer engagement and ultimately drive revenue growth.

Project Time Frame



Thank You

For you time.