Framework for the calculation of Customer Lifetime Value in the Hotel industry

Group 1

Customer Lifetime Value (CLV)

Total amount of revenue a business can earn from a customer, throughout the entire business relationship between the parties.

Revenue x Relationship Duration (i.e. timeframe of days, months, years)

Basis in making impactful business decisions: overall corporate strategies, cost-cutting measures, new product/business, pricing, marketing and advertising budgets, customer segmentation.

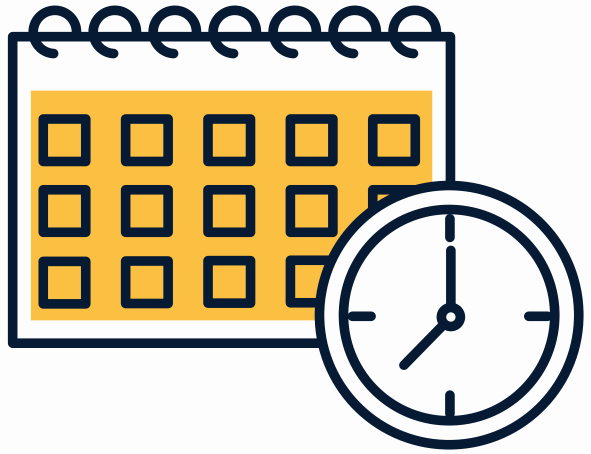
CLV in the Hotel Industry

Focus: Customer Segmentation

Identify the most valuable guests to understand overall profitability and sustainability of the business

Considerations:

Transaction Timing



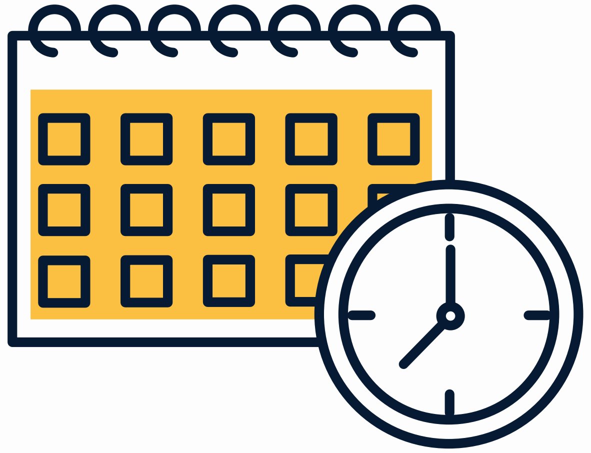
Repeat Business

Guest Spend

Proposed Two-Step Approach

Recency, Frequency, Monetary (RFM)

Transaction Timing (R)



Repeat Business (F)

Guest Spend (M)

A metric score from 1 to 5 is assigned per category, and a total RFM Score is computed

Proposed Two-Step Approach

Customer Lifetime Value (CLV)

Customers are classified based on RFM Score:

Low - Value

Mid - Value

High - Value

Clusters are useful for:

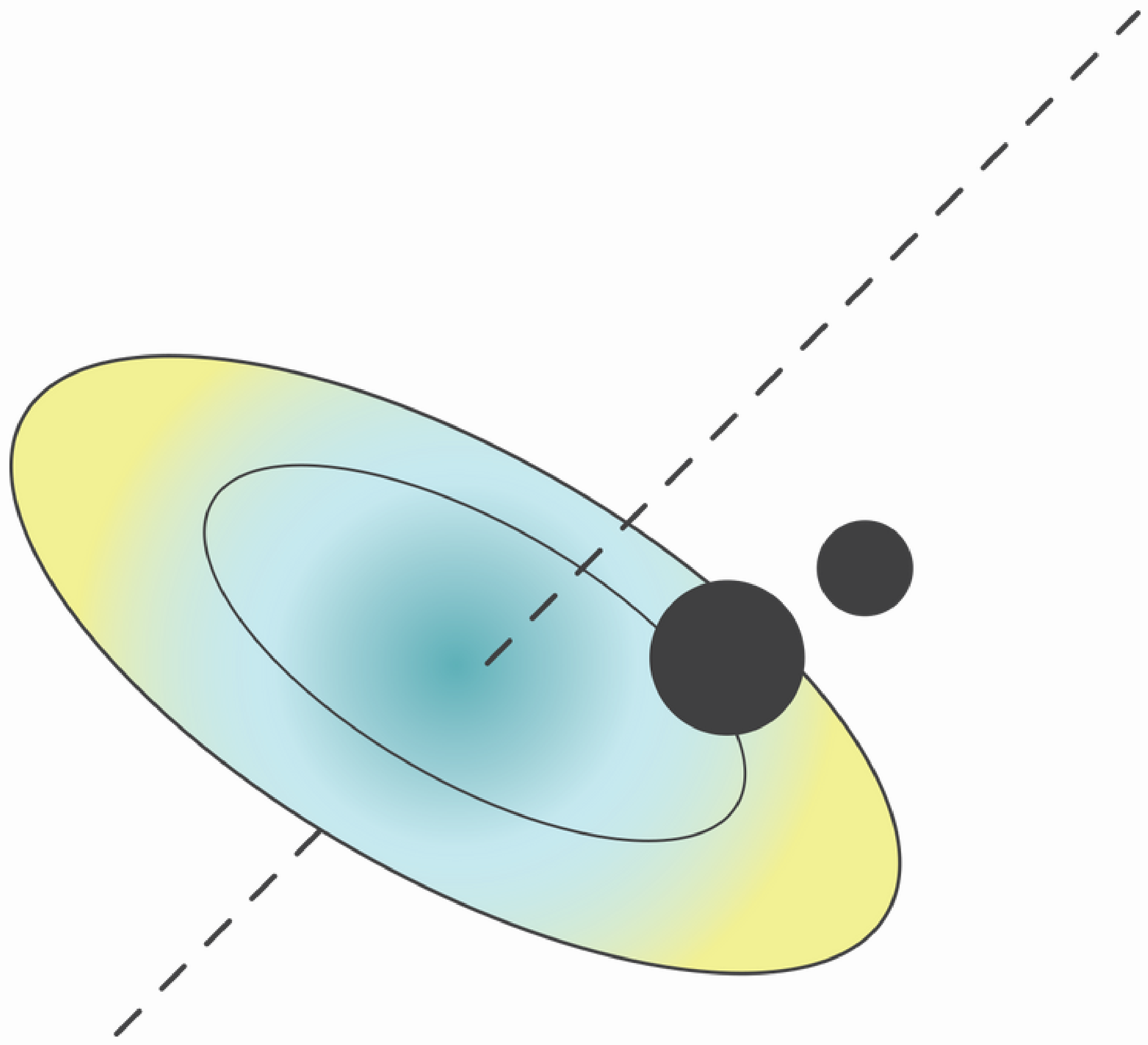
Demographic Segmentation

Behavioral Segmentation

CLV Assignment

Dataset

And Data Cleansing



Our dataset comprises of ~83,000 unique customers, with 28 features defined for each of them.

However, some features could result redundant and potentially increase recurring computational costs.

For instance, of the 28 initial features only 16 have been kept in the DataFrame exploited for the definiton of our framework.

Feature engineering

Computing Recency, Frequency and Monetary scores and RFM

```
data['RFM'] = data['Recency'] + data['Frequency'] + data['Monetary']
```

Clustering Customers on RFM

Exploiting K-Means (unsupervised) algorithm

	RFM	Cluster
25722	6	1
33557	7	2
75859	11	2
3297	6	1
2329	0	0

Model Output

Based on the computed RFM score, each customer is assigned to a cluster

Average RFM values by Cluster

Cluster	Recency	Frequency	Monetary	RFM
0	0.013959	0.013959	0.000000	0.027918
1	1.191377	1.005311	2.909623	5.106311
2	1.864928	1.096542	4.612058	7.573528

Minimum and Maximum RFM values by Cluster

Minimum

Cluster	Recency	Frequency	Monetary	RFM
0	0	0	0	0
1	0	1	0	3
2	1	1	0	7

Maximum

Cluster	Recency	Frequency	Monetary	RFM
0	1	1	0	2
1	5	4	4	6
2	5	5	5	15

Deriving CLV values

For each customer:

(Monetary * Frequency) * RecencyPenalty

Assigned Penalties by Recency

Recency	5	4	3	2	1	0
Penalty	1.0	0.9	0.8	0.7	0.6	1.0

Average CLV by Cluster

Cluster	CLV
0	1.147241
1	159.703562
2	607.328488

Automation

01

Ingestion

Data on new customers is ingested in batch at day-end in Apache NiFi

02

Storage

Data is stored in a SQL Relational Database, to be customized ad-hoc for each client

03

Data processing

The stored data is analyzed and specific RFM, CLV and AC values are computed and a Cluster is assigned

04

Targeted marketing

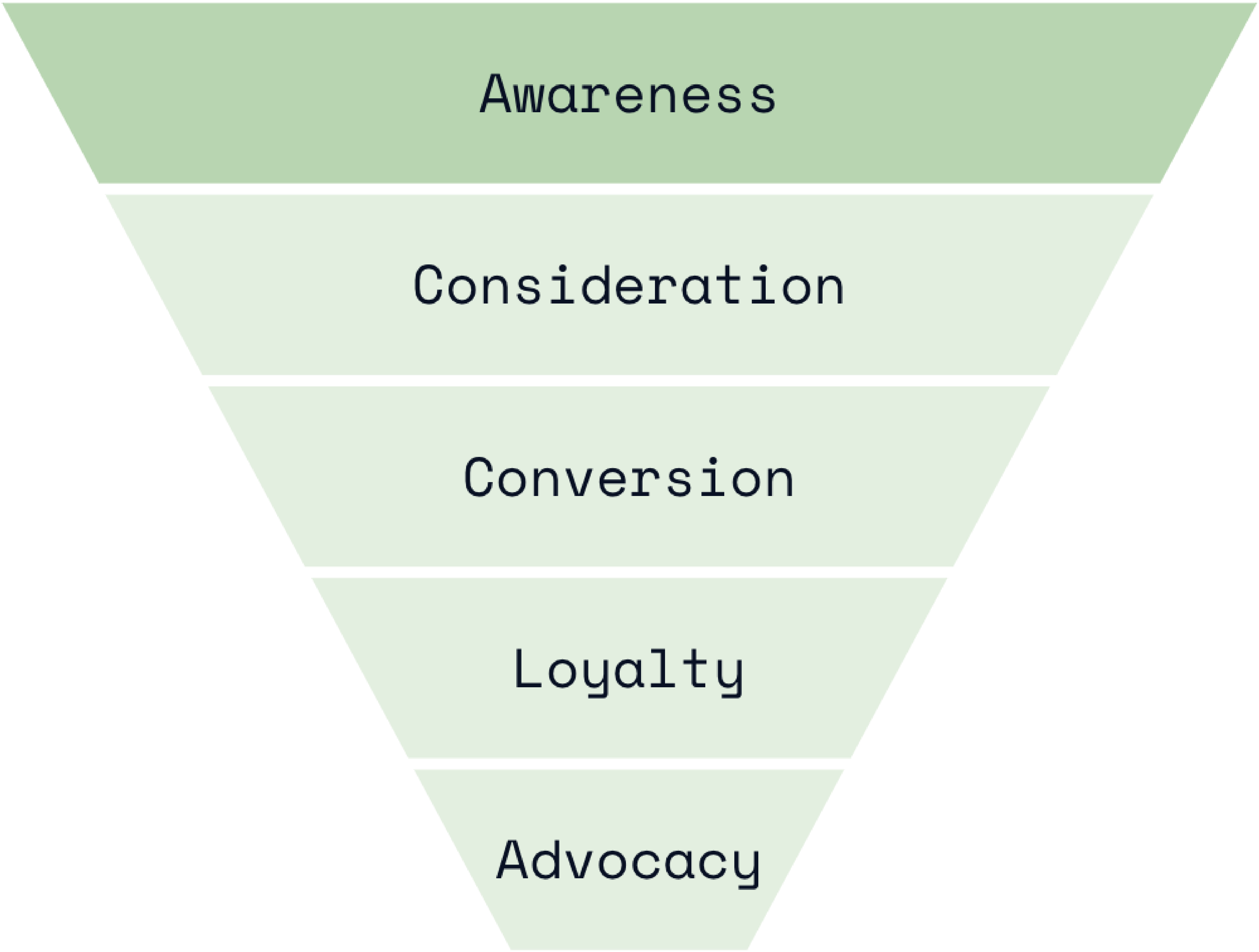
Based on the computed values and Clusters, specific marketing campaigns can be automatically applied

Value for Hotels and Recommendations

With proper segmentation and CLV per segment, hotels can:

- Understand the behaviors distinct to each and create a customer persona
- See which segment offers a higher CLV on average
- Focus on a particular segment as the main target for the hotel and strategize
    - Business hotel
    - Family hotel
    - Luxury Hotel
- Adjust the Customer Acquisition Cost accordingly (maximum 3 times the CLV)

CO: Low-value customers



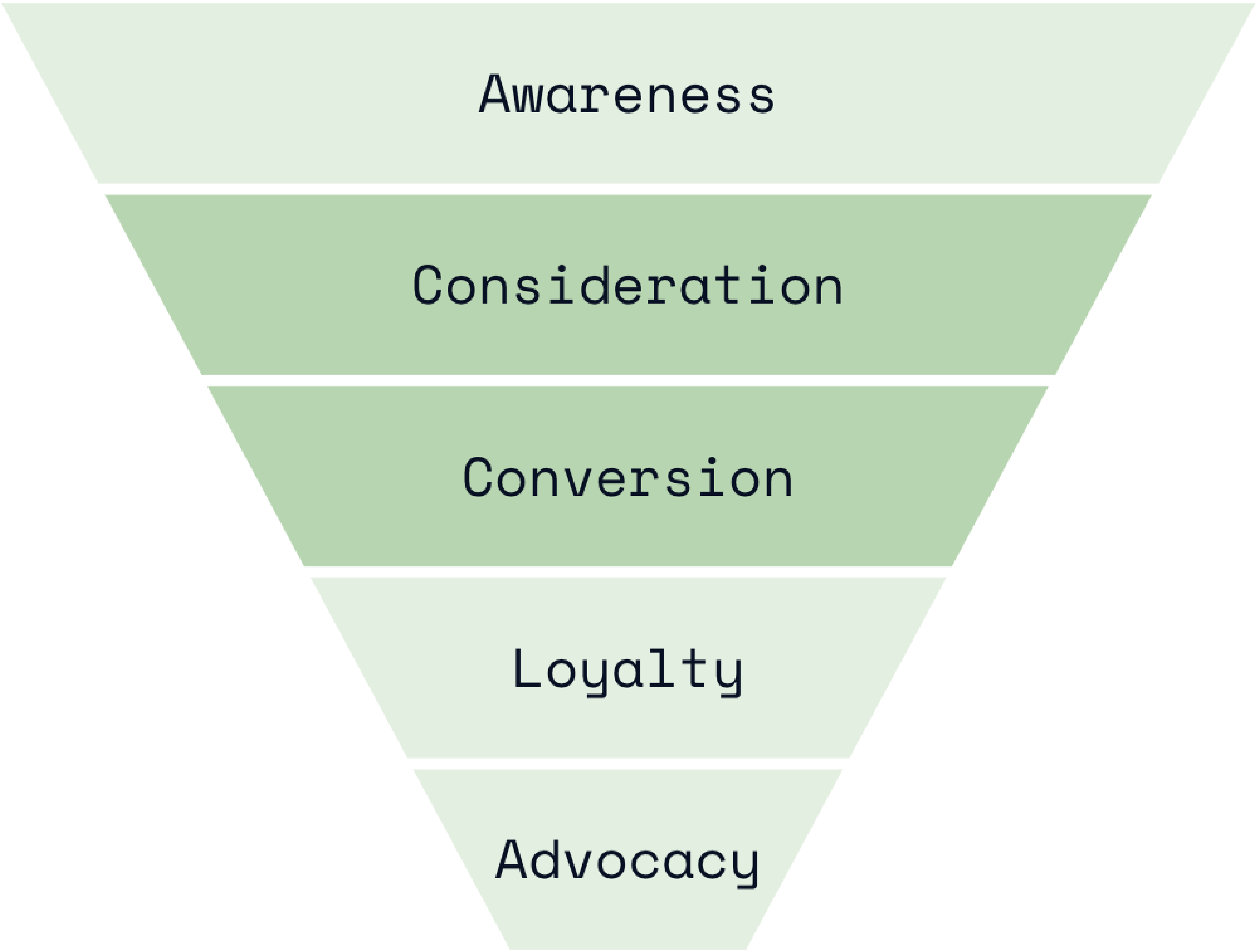
C0: Low-value customers

- **Offering discounts or packages** can be an effective way to incentivize low-value and incoming customers to book with the hotel*

- **Providing excellent customer service** is crucial for retaining low-value customers and encouraging incoming customers to become repeat customers*

*("Pricing Strategies for Hoteliers", 2016)

C1: Mid-value customers

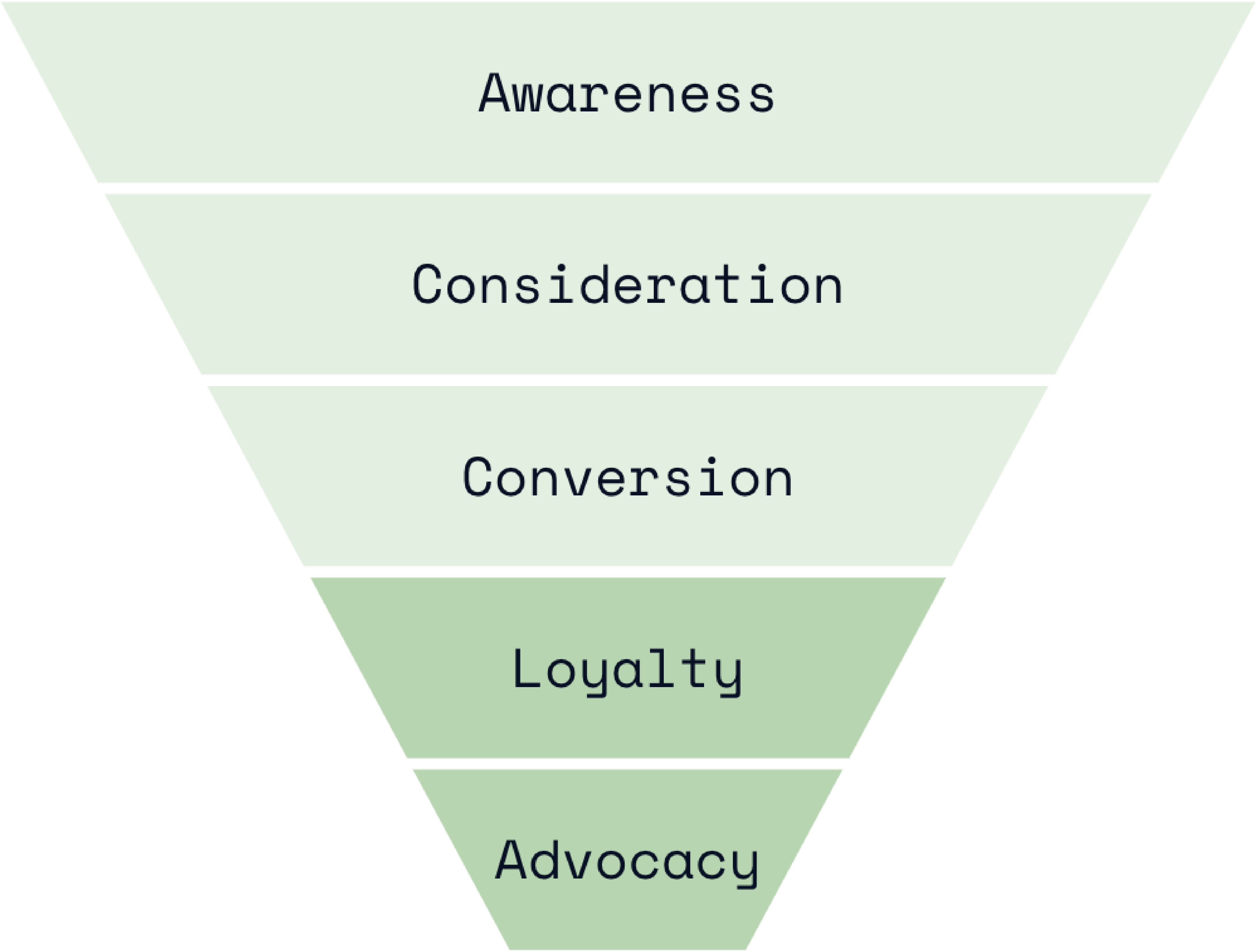


C1: Mid-value customers

- **Personalized recommendations** and experiences can be a key driver of customer satisfaction among mid-value customers** since their RFM scores can result from a variety of different factors
    - If they score high in Monetary and Recency, we should target Frequency by offering loyalty programs
    - If the score is high on Monetary and Frequency, we should target Recency by **upselling additional services or amenities***

*("Upselling Practices in the Hospitality Industry: An Exploratory Study", 2016)
** ("Oracle Hospitality Consumer Research 2019: The Generational Divide", 2019)

C2: High-value customers



C2: High-value customers

This segment must be rewarded and targeted on the basis of their loyalty, signaled by their high RFM scores.

- **Offering exclusive perks or amenities** is an effective way to increase customer loyalty among high-value customers*
- **Personalized service** and attention to customers' needs and preferences is a key factor in driving customer satisfaction among high-value customers**

*("Hotel Loyalty Programs: Customers' Perspectives and Implications for Hoteliers", 2016)
** ("North America Hotel Guest Satisfaction Index Study", 2019)

Next Steps

Applying additional layers of analysis

Adjusting thresholds

Creating Cluster-based Customer Personas

Thank you.

References

- Cornell University. (2016). Pricing Strategies for Hoteliers. Retrieved from https://scholarship.sha.cornell.edu/cgi/viewcontent.cgi?article=1179&context=chrreports
- Journal of Hospitality Marketing & Management. (2016). Upselling Practices in the Hospitality Industry: An Exploratory Study. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/19368623.2015.1065804
- Oracle Hospitality. (2019). Oracle Hospitality Consumer Research 2019: The Generational Divide. Retrieved from https://www.oracle.com/a/ocom/docs/hospitality-2019-consumer-research-study-en.pdf
- International Journal of Hospitality Management. (2016). Hotel Loyalty Programs: Customers' Perspectives and Implications for Hoteliers. Retrieved from https://www.sciencedirect.com/science/article/pii/S0278431916300182
- J.D. Power. (2019). North America Hotel Guest Satisfaction Index Study. Retrieved from https://www.jdpower.com/sites/default/files/file/2019-07/2019083 North America Hotel Guest Satisfaction Index Study SM.pdf

Appendix

Sample of Monetary column computation

```
data["Total_revenues"] = data['LodgingRevenue'] + data['OtherRevenue']
data['Monetary'] = np.select(
    [
        data['Total_revenues'] >= 500,
        data['Total_revenues'].between(301, 499),
        data['Total_revenues'].between(201, 300),
        data['Total_revenues'].between(101, 200),
        data['Total_revenues'].between(1, 100),
        data['Total_revenues'] < 1
    ],
    [
        5,
        4,
        3,
        2,
        1,
        0
    ]
)
data = data.drop("Total_revenues", axis=1)
```

Instantiating and assigning clusters

```
from sklearn.cluster import KMeans

# Set the number of clusters
num_clusters = 3

# Create a new dataframe with only the "RFM" column
rfm_data = data[['RFM']]

# Create a KMeans model and fit it to the data
kmeans = KMeans(n_clusters=num_clusters, random_state=42)
kmeans.fit(rfm_data)

# Add the cluster labels to the original dataframe
data['Cluster'] = kmeans.labels_
```