



CUSTOMER SEGMENTATION ANALYSIS ON RETAIL INDUSTRY

Anvi Radhakrishna Naik

AGENDA

- ❖ Introduction
- ❖ Problem Statement
- ❖ Methodology
- ❖ Results and Discussion
- ❖ Conclusion and Future Work



INTRODUCTION



- ❖ Retail industry uses variety of approach to collect data; via Loyalty program, Point of Sales, Surveys, Social media, Web analytics.
- ❖ Data gathered are then analyzed and appropriate business strategies are made.
- ❖ Data consists of consumer details from demographical, transactional, consumer sentiments, geographical etc.
- ❖ Goal of consumer analysis is to understand consumer behavior and preferences in order to improve sales.





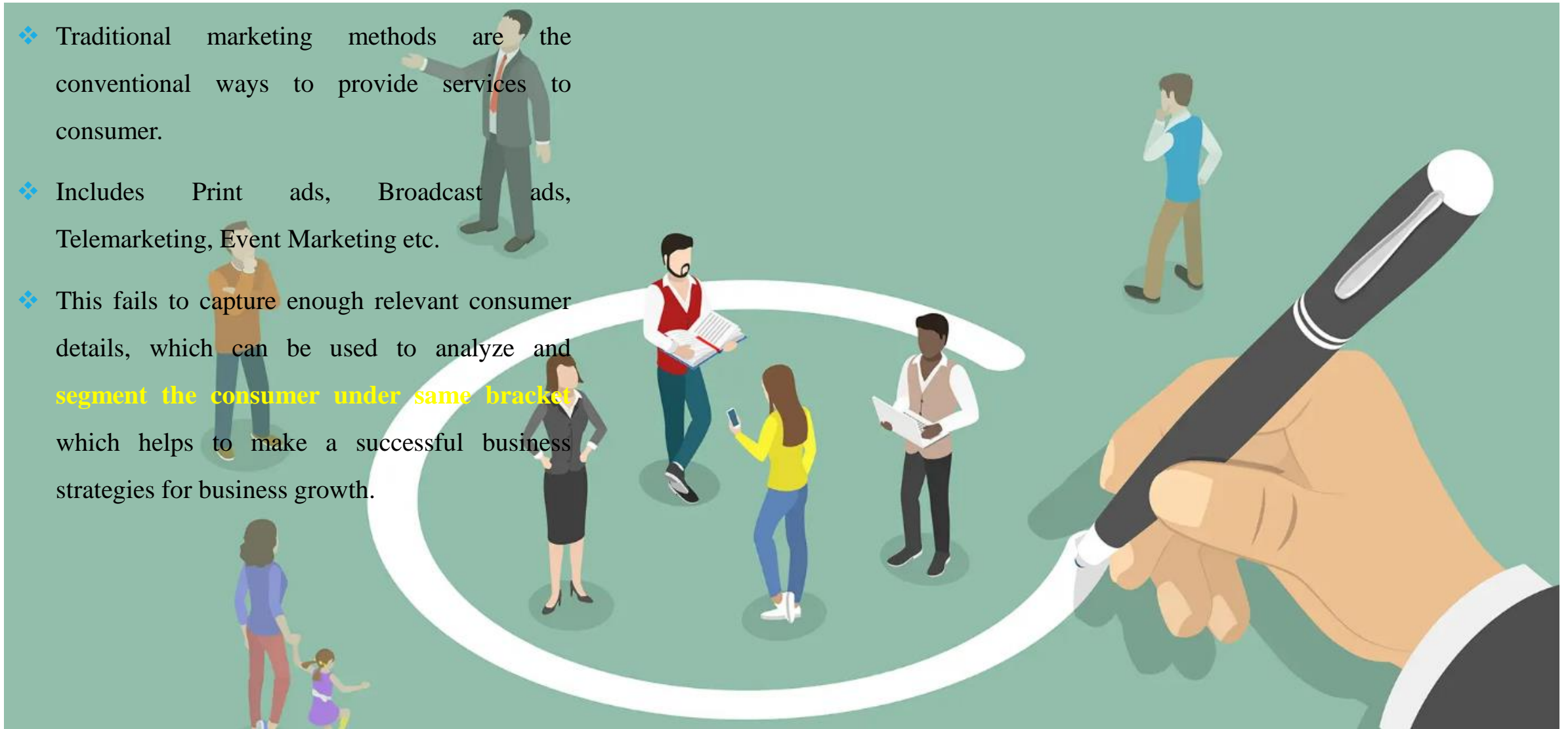
PROBLEM STATEMENT



- ❖ Traditional marketing methods are the conventional ways to provide services to consumer.

- ❖ Includes Print ads, Broadcast ads, Telemarketing, Event Marketing etc.

- ❖ This fails to capture enough relevant consumer details, which can be used to analyze and **segment the consumer under same bracket** which helps to make a successful business strategies for business growth.





METHODOLOGY



CUSTOMER SEGMENT



- ❖ A customer segment refers to a group of consumers who share similar characteristics and preferences.
- ❖ It is a process of dividing the a larger market into smaller groups of consumers, in order to create a targeted audience and make personalized marketing strategies that are more effective in engaging consumers.

4 TYPES OF Customer Segmentation



Geographic

Location-based, such as from IP address or user-provided address



Demographic

Age, marital status, sex, education level, occupation, income, religion, and nationality



Psychographic

Lifestyles, personalities, interests, beliefs, attitudes, and values



Behavioral

Online shopping habits, loyalty, usage frequency, and website actions

OBJECTIVES OF METHODOLOGY

- ❖ Identify Consumers – understand the general trends affecting
- ❖ Factors associated with consumers – RFM and K mean technique
- ❖ Stronger factor associated – propose attributes that are highly associated, which forms segments.



RFM AND K MEAN — HYBRID MODEL

RFM Analysis

- ❖ Is widely used to identify customer value and also identify valuable customer available.
- ❖ Factors that are affected are the consumers who frequently shop and also spend the most within the time frame.

K-Mean

- ❖ is used to understand their consumer segments better, by collecting similar data points together based on characteristics.
- ❖ which in turn makes data-driven decisions to improve consumer experience and drive growth.

For Hybrid model, analysis is carried out with factors on transactional data; invoice periods, total transaction and average purchases done.



RESULTS AND DISCUSSION



RESULTS

- ❖ On the overall analysis, on the factors considered, we have come with 4 clusters / segments based on the transactional data.
- ❖ We have consumers who have spent low and ordered less in the same range. Consumers who have spent high and ordered less. Consumers who are in middle range but also spend at the same level. Then we have consumers who have spent very high but they have ordered high as well.





CONCLUSION AND FUTURE WORK



CONCLUSION AND FUTURE WORK

- ❖ Based on this the business retailer can come up with the campaign strategies, which can be grouped accordingly for the profitable consumers, loyalty consumers, awareness programs can be built to acquire new consumers for conversions, etc.
- ❖ Other factors can also be explored and added to the model, which opens up a new dimensional on the clusters.





APPENDIX



VISUALIZATION OR PLOTTED GRAPHS USED IN ANALYSIS

Fig 5.2: Cluster Sum of Squared Distances VS Clusters

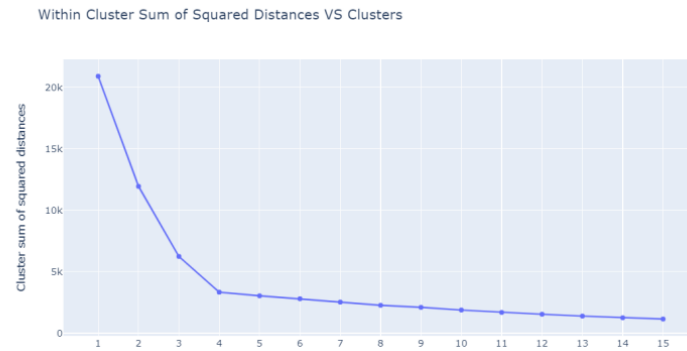


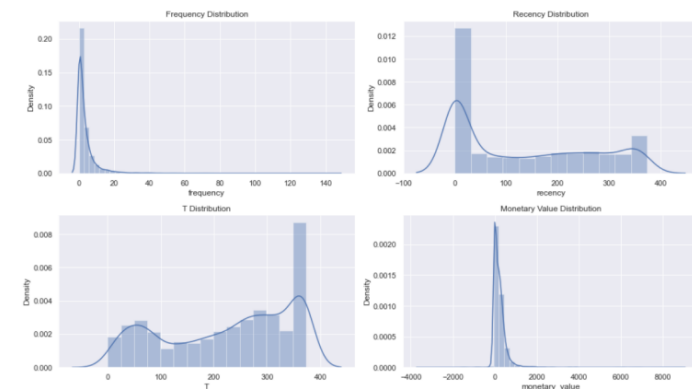
Table 5.3: RFM grouping based on the Customer purchases made

	CustomerID	frequency	recency	T	monetary_value
0	12346.0	0.0	0.0	325.0	0.000000
1	12347.0	6.0	365.0	367.0	599.701667
2	12348.0	3.0	283.0	358.0	301.480000
3	12349.0	0.0	0.0	18.0	0.000000
4	12350.0	0.0	0.0	310.0	0.000000

Table 5.4: Cluster statistics

	products_ordered	average_return_rate	total_spending	log_products_ordered	log_average_return_rate	log_total_spending	clusters
0	5.775433	158.953794	1.0	1.913303	5.074885	0.693147	0
1	61.323296	1215.651062	1.0	4.132335	7.103857	0.693147	1
2	23.772598	401.817358	1.0	3.209738	5.998483	0.693147	2
3	141.052279	4934.377218	1.0	4.956195	8.504184	0.693147	3

Figure 5.1: frequency distribution plotted in the RFM value.



VISUALIZATION OR PLOTTED GRAPHS USED IN ANALYSIS

Fig 4.12: Most of Transaction based on quantity purchased by consumers

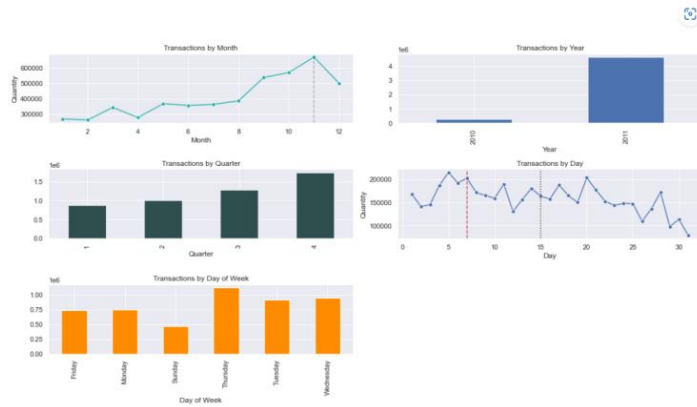


Fig 4.5: Average Price by Country

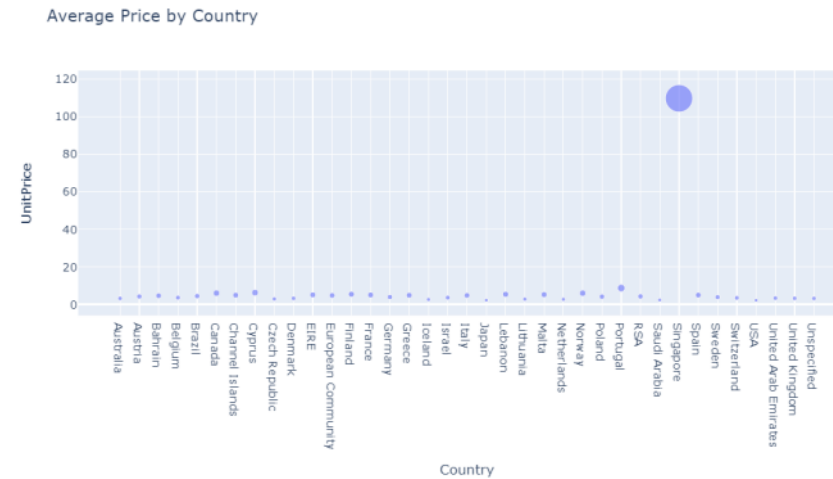


Fig 4.6: Price Distribution Country wise

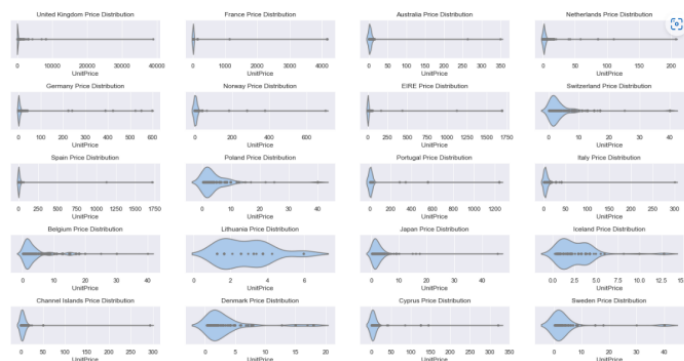


Fig 4.1: various types of department vs customer count

