Marketing Campaign Analysis

TianYu Gao,Shagufa Habibi, Ge Jin ,Yuetong qu, Rixin Su

Agenda

Tianyu

1.Introduction

2.Data

3.Model

4.Conclusion

Introduction

- Background
- Main Business
- Context

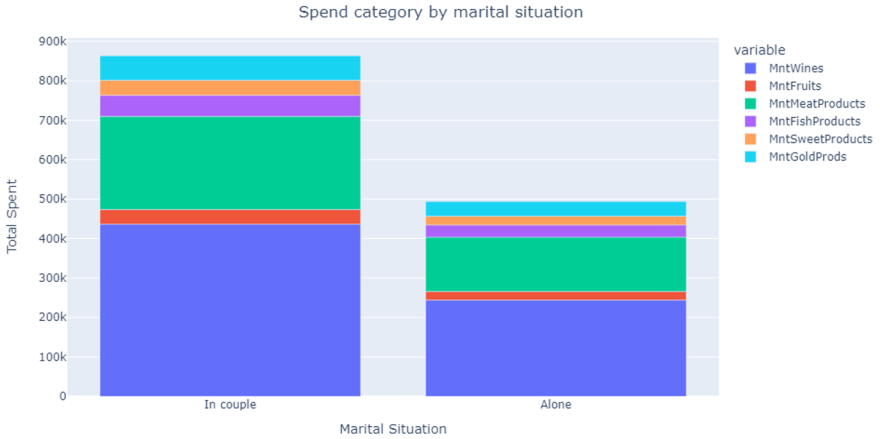
Data - Preprocessing

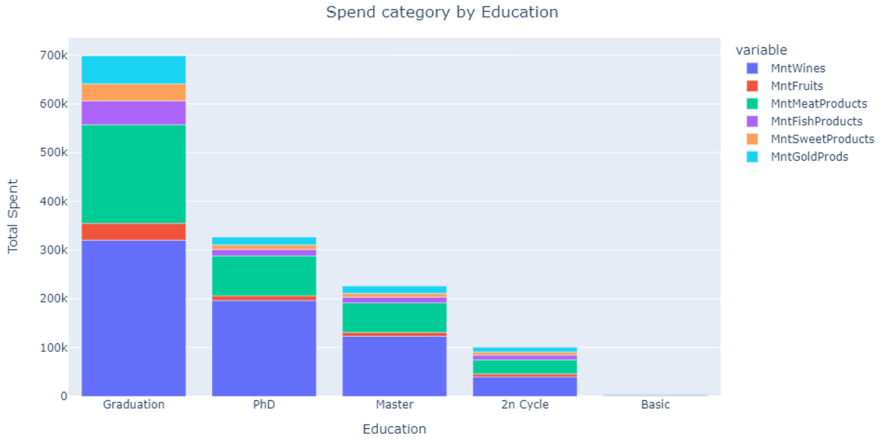
- Overview
- Outliers
- Missing values
  - Filled using regression
- Feature Engineering
  - Decimal encoding
  - One-hot encoding
  - Numeralization

Data Visualization

Yuetong

Data - Visualization

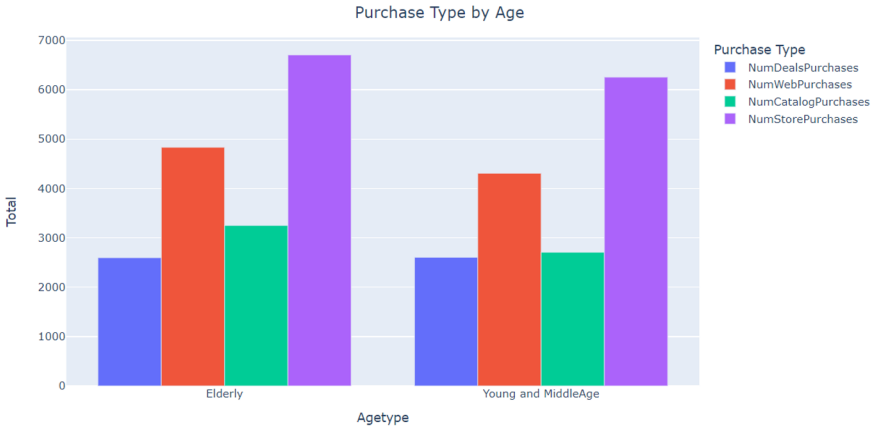


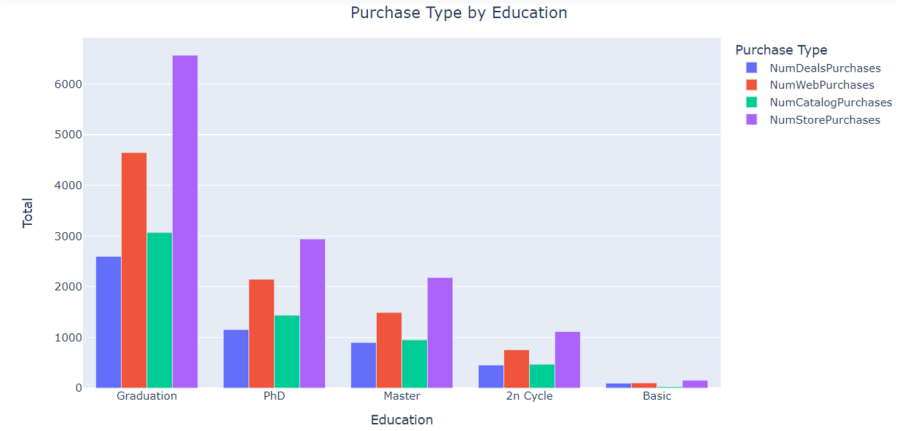


- For customers in couple and alone, the spend categories distribution are almost the same.
- Wines and meat occupy the top 2 largest proportion of purchase.

- For the first four education levels' customers, the spend categories distribution are also almost the same.
- Wines and meat are still the top 2 purchase.
- Basic consumers spend almost nothing.

Data - Visualization

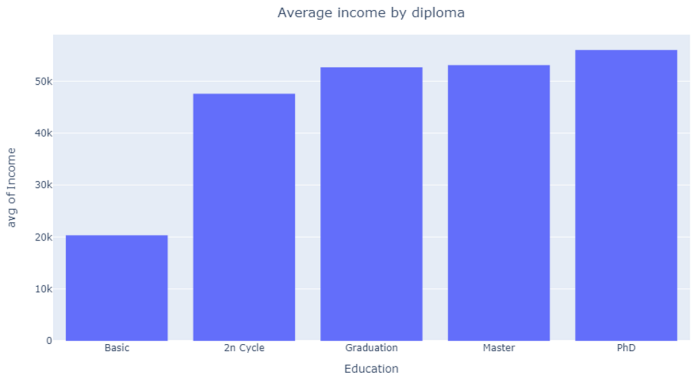


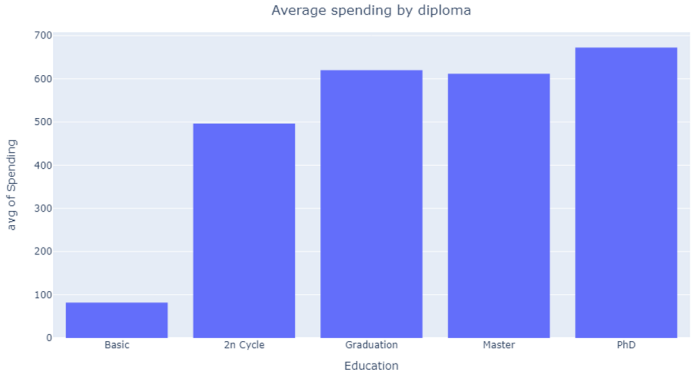


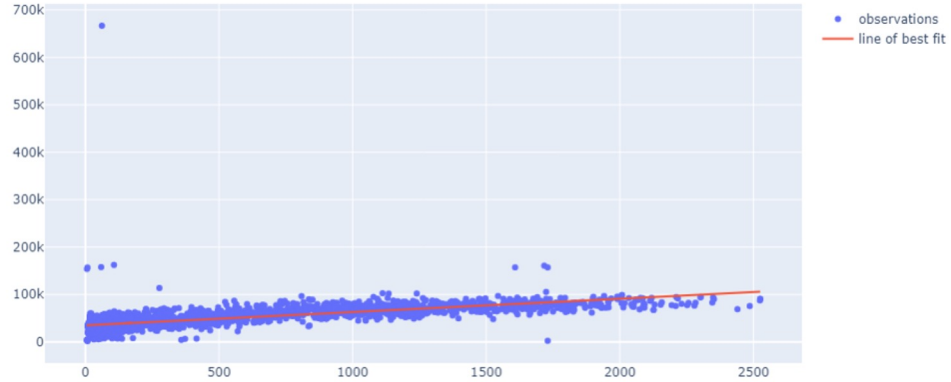
- For both elderly and young and middle aged customers, the purchase channel distribution are almost the same.
- Most customers chose purchase in store.

- The purchase channel distribution of first four education levels' customer is also almost the same. Basic customers have very little purchase. And graduation customers have the most.
- Similarly, most customers chose purchase in store.

Data - Visualization

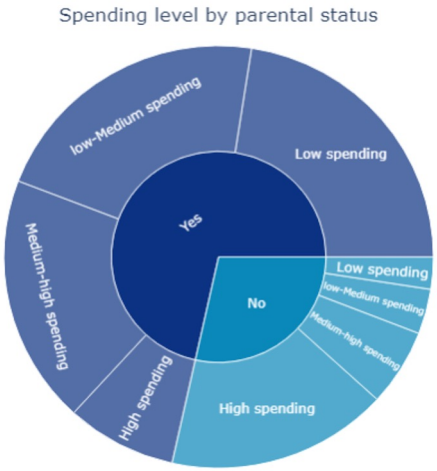


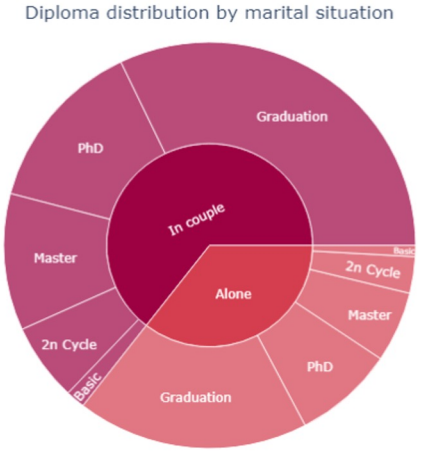


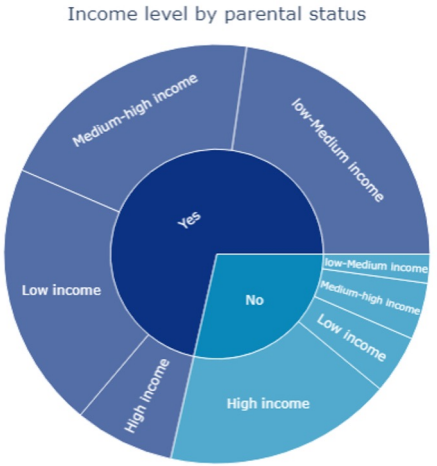


- The better the diploma is, the higher the average income. The same trend applies to average spending.
- Spending seems to be positively correlated with the income.

Data - Visualization

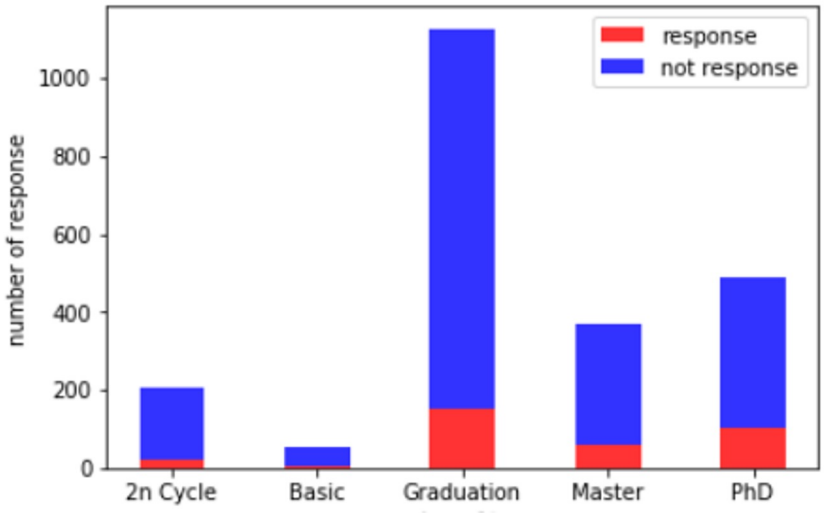


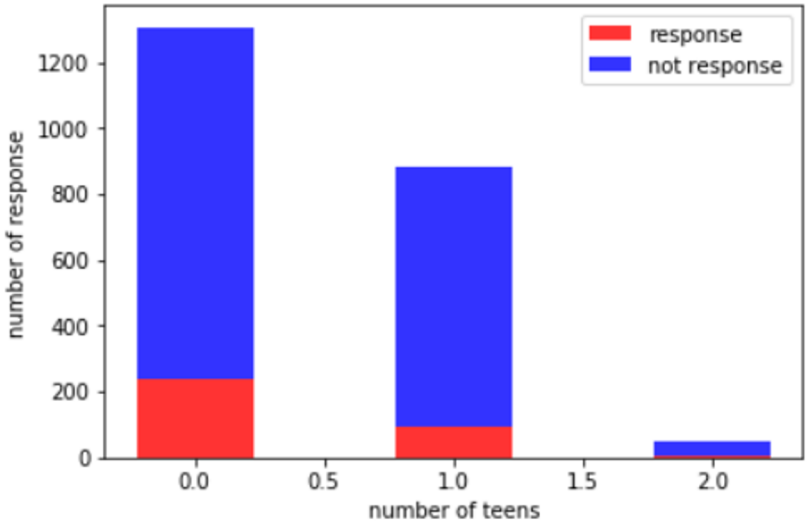




- The distribution of education seems to be identical for the two population In couple and Alone. We could be tempted to believe that there is no correlation between education and the marital status.
- People with high income and spending are largely representing the population who has no child. People having at least 1 child are mainly represented by people with low income. We could be tempted to believe that people having a high income tend to not have a child.

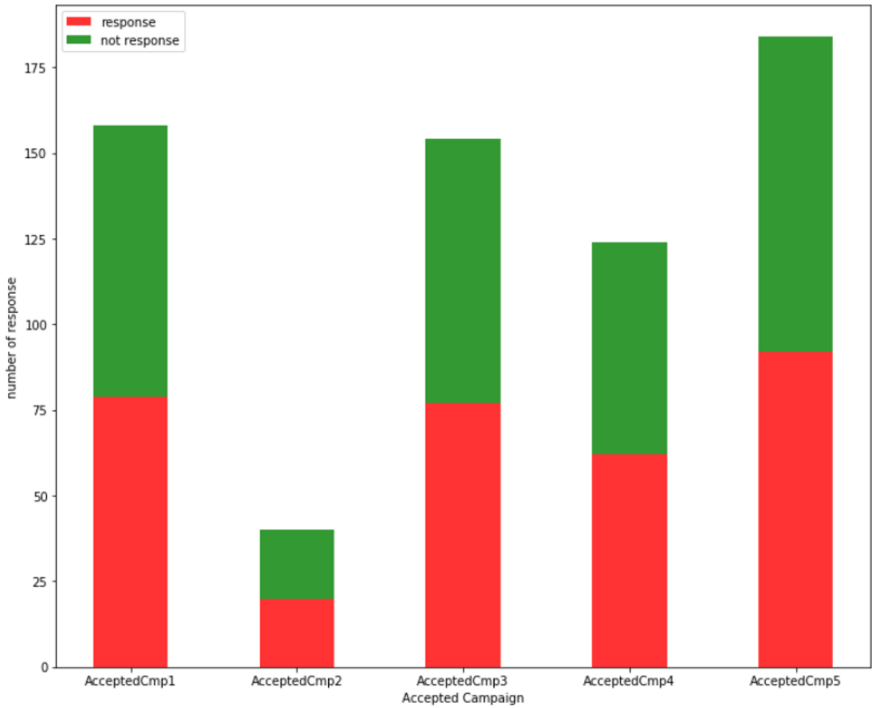
Data - Visualization





- Graduation customers have the highest response rate. And customers with basic education have the lowest.
- Customers with no teens have higher response rate. The response rate for customers with two children is almost zero.

Data - Visualization

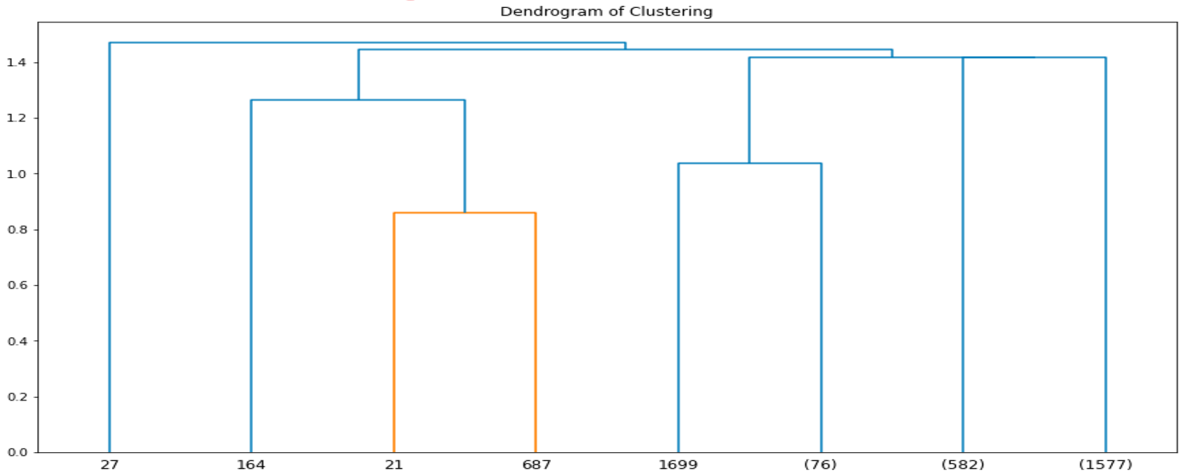


- It seems that customers accepting each campaign have nearly the same response rate.
- Customers accepting campaign 2 have the lowest number of response. Customers accepting campaign 5 have the highest number of response.
- However, this is for the whole population. Different campaigns may have different impacts on different kinds of customers. So we should first cluster our customers, and then study the impact of each campaign on customer response.

Clustering Analysis

Shagufa

Model - Clustering



Cluster	Description	Cluster	Description
1	High-income meat consumers	4	Couples with low willingness to have babies
2	Aged middle class	5	Highly interactive
3	Living together with others, not married	6	Parents of Young kids

Model - Clustering

Goal:

To find out the impact of acceptance of each campaign on the final response of the last campaign for each segmented customer cluster

Focus: **History and Campaign variables**

Process: **Two new variables lift and p_value.**

Analysis/ Interpretation:

For cluster 5, compared to observations that did not accept any campaigns, the campaign 2 leads to a lift in the response rate of the final campaign of 0.8704, and the boost is statistically significant. We suggest that stores optimize their target customers for cluster 5 with campaign 2.

For cluster 4, compared to observations that did not accept any campaigns, the campaign 3 leads to a lift in the response rate of the final campaign of 0.3615, and the boost is statistically significant. We suggest that stores optimize their target customers for cluster 4 with campaign 3.

For cluster 3, observations accepted the campaign1, campaign 5 lead to a lift in the response rate of the final campaign of 0.5, and the boost is statistically significant. We suggest that stores optimize their target customers for cluster 3 with campaign 5.

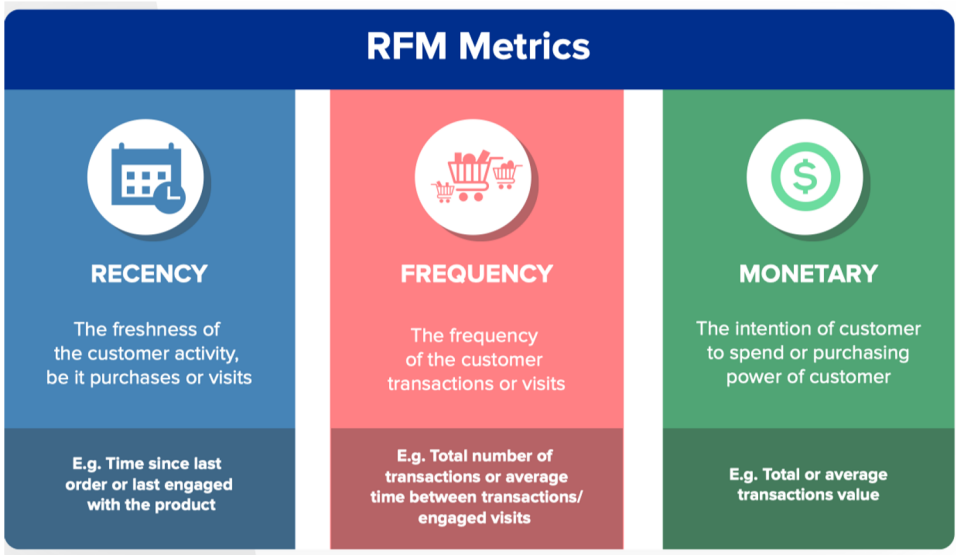
For cluster 2, compared to observations that did not accept any campaigns, campaign 1 leads to a lift in the response rate of the final campaign of 0.5287, and the boost is statistically significant. We suggest that stores optimize their target customers for cluster 2 with campaign 1.

RMF Model & Business Metrics

Ge

What is RFM Analysis?

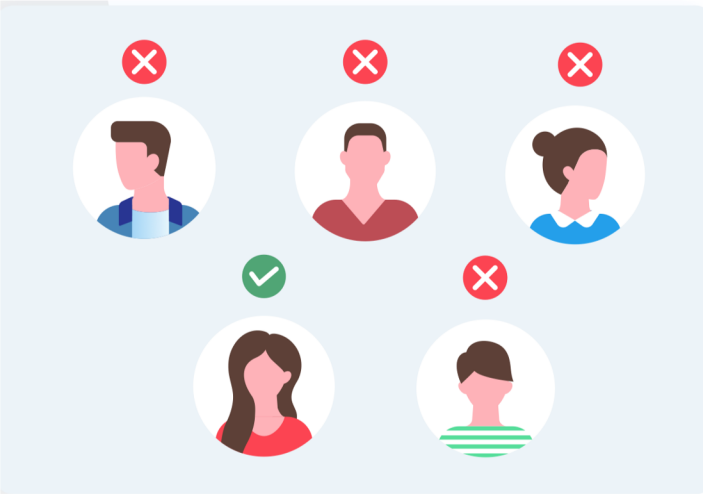
RFM (Recency, Frequency, and Monetary value) Analysis measures how recently, how often, and how much money a customer has given your brand.



How Does RFM Help?

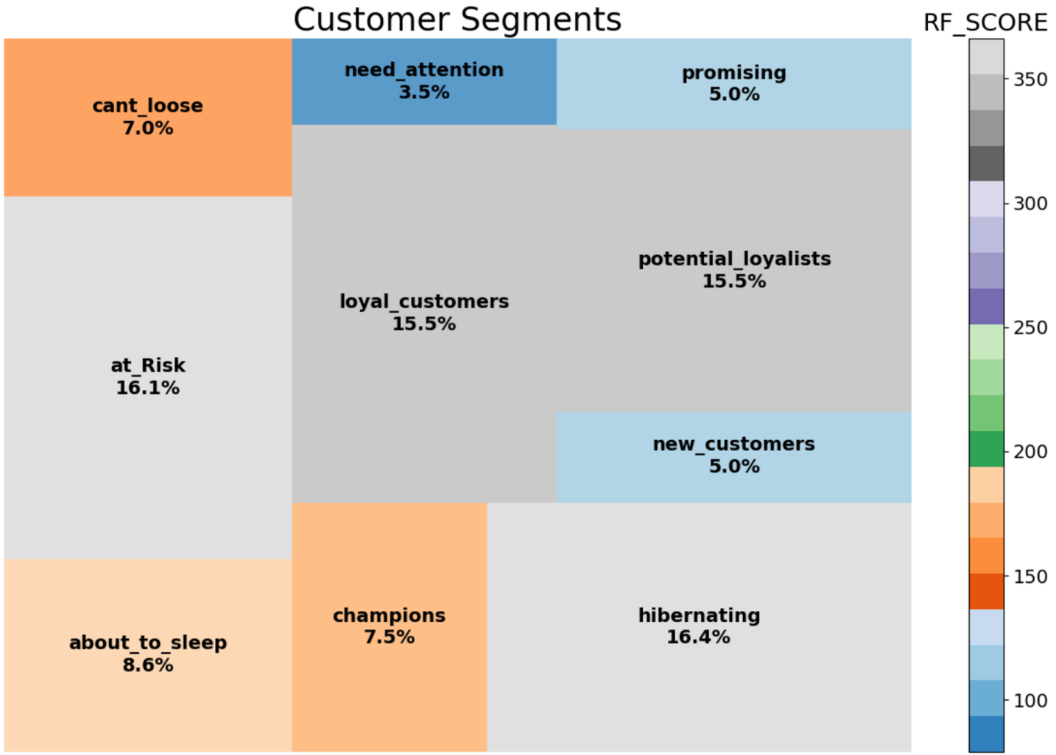
RFM analysis helps marketers answer:

- Who are their best customers?
- Which of their customers could contribute to churn rate?
- Who has the potential to become valuable customers?



RFM Analysis

Our plot uses recency and frequency scores to visualize RFM on a 2-dimensional graph. And we pared down the number of possible segments to arrive at just 10.



Analyzing RFM Segmentation

- **Hibernating (16.4%)**at the bottom end of our customer base are the Hibernating customers. Last purchase was long back, low spenders and low number of orders.Offer other relevant products and special discounts. Recreate brand value.
- **At Risk Customers (16.1%)**are our customers who purchased often and spent big amounts, but haven't purchased recently. Need to bring them back! Send personalized emails to reconnect, offer renewals, provide helpful resources.
- **Potential Loyalist (15.5%)** Recent customers, but spent a good amount and bought more than once.Offer membership / loyalty program, recommend other products.
- **Champions(7.5%)** Our Champions can help us grow. Bought recently, buy often and spend the most! Reward them. Can be early adopters for new products. Will promote your brand.

Business Metrics

1) **Average Order Value**
(average_order_value = total_price / total_transaction)

2) **Purchase Frequency**
(total_transaction / total_number_of_customers)

3) **Repeat Rate & Churn Rate**
repeat_rate = number of customers who make two or more purchases / total number of customers x 100%

churn_rate = 1-repeat_rate

	segment	average_order_value	purchase_frequency	repeat_rate	churn_rate
0	about_to_sleep	15.20	5.26	98.96	0.52
1	at_Risk	57.94	15.29	100.00	0.00
2	cant_loose	55.92	23.35	100.00	0.00
3	champions	55.00	20.44	100.00	0.00
4	hibernating	15.24	5.40	98.36	0.82
5	loyal_customers	57.23	20.45	100.00	0.00
6	need_attention	46.16	12.59	100.00	0.00
7	new_customers	11.06	4.03	99.10	0.00
8	potential_loyalists	38.98	10.12	100.00	0.00
9	promising	9.68	3.99	97.30	1.80

Conclusion

Rixin

Conclusion

- Marketing campaign content should be designed for highly educated customers and offline stores
- To improve our marketing campaign's response rate, we should focus on the customers who response campaign 2,1,3 respectively
- Loyal customers, potential loyalists and at_risk, these users need our extra attention.
- Champions and new customers together account for 12.5% of the total, which are an important source of revenue for the company.
- We need retention measures to at_risk customers.

Future Work

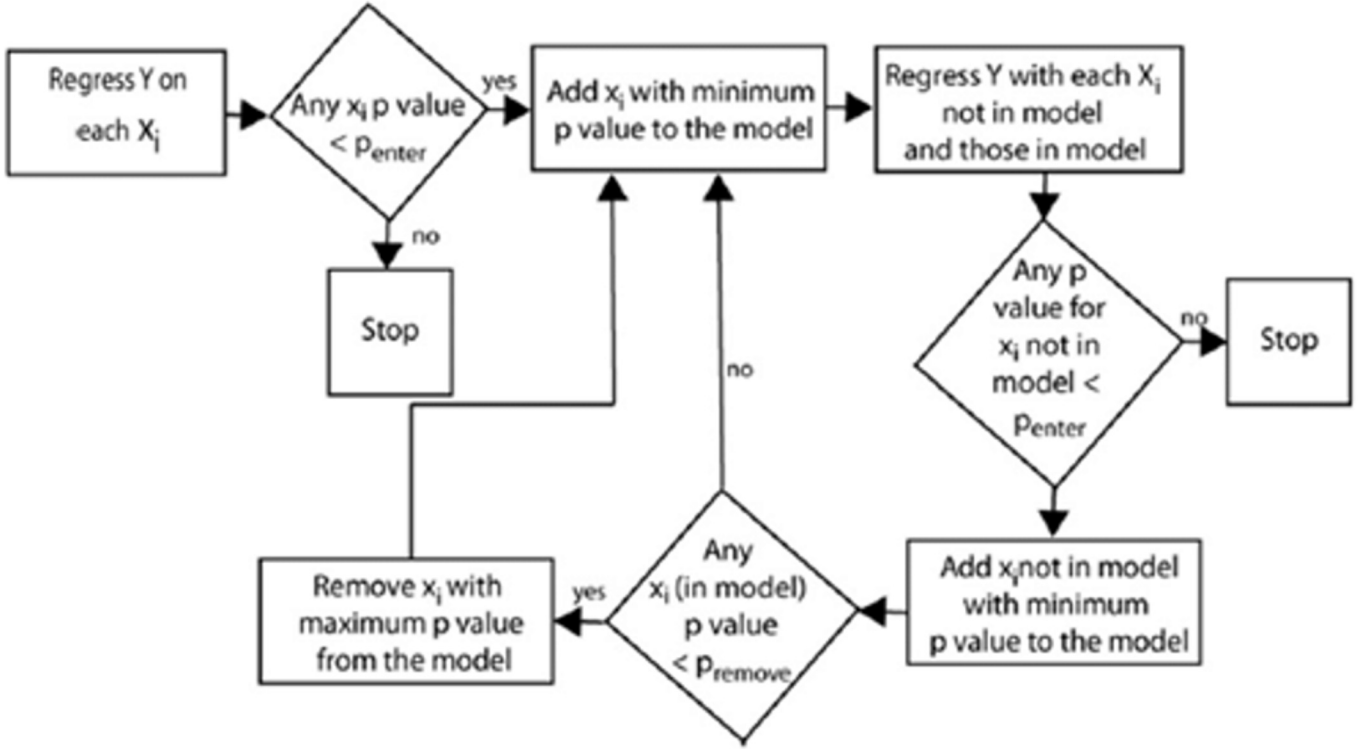
- We will do Customer Lifetime Value Analysis to estimate the lifetime value of each customer.

- By analyzing customer complaints, we can analyze the points that users most care about and take corrective actions.

- We will also evaluate the performance of campaign 1-5 by the efficiency of converting customers.

Q&A

Appendix



Graph 1. Flow chart of Stepwise Selection

Appendix

VARIABLES	Income	VARIABLES	Income
		NumWebVisitsMonth	-3,175***
MntSweetProducts	29.26***	MntWines	15.94***
Education_yrs	166.5***	Teenhome	6,392***
Recency	-13.05**	MntMeatProducts	19.56***
NumWebPurchases	1,579***	Year_Birth	-45.23***
NumCatalogPurchases	397.4***	MntFruits	22.14***
Kidhome	3,478***	NumDealsPurchases	-1,057***
NumStorePurchases	543.7***	Constant	133,418***
		R-squared	0.828

Table 1 Results of Regression

Appendix

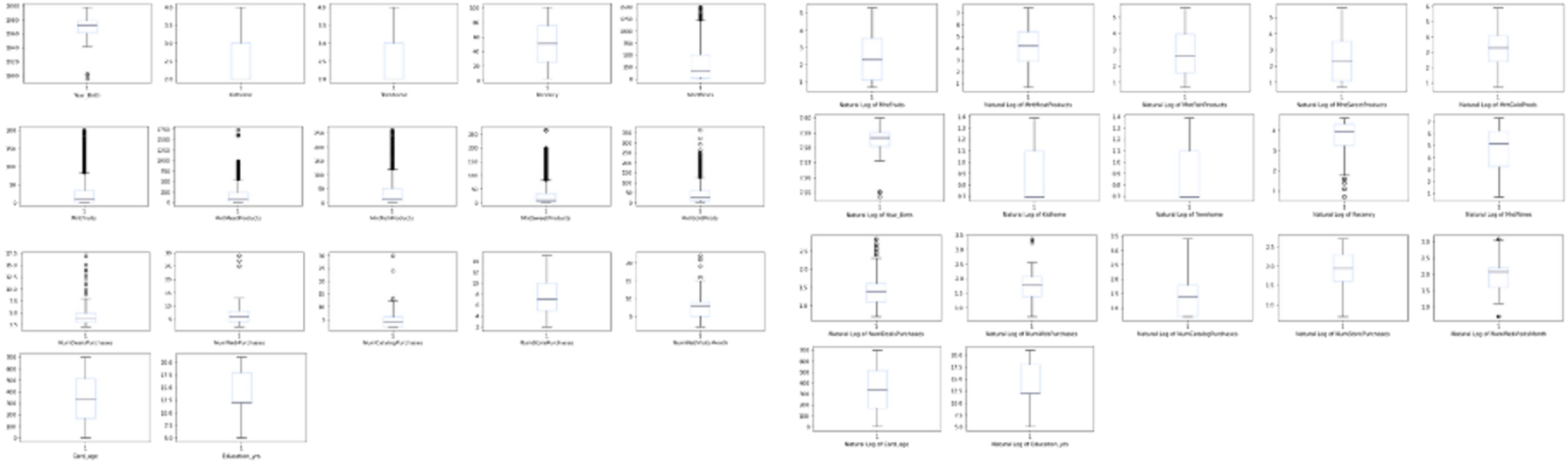


Exhibit 1 Box plots