Assignment - Product Analyst | Sinjini Roy

PART 1:

Link:
https://docs.google.com/spreadsheets/d/1bo-LCDx1cUHh2CGelWiPY_WdivI7USy9dlysuoq50OA/edit?usp=sharing

The exploratory data analysis on the Consumer personal data is mainly focussed on the features:

- Education
- Marital Status
- Income: The income feature has been distributed into 3 buckets.

Income Bucket (Rs)
<=10000
>10000 to <=50000
>50000

KEY INSIGHTS:

- 47% of the entire population who purchase **Wine** are in the Education Category (Graduation), 64% are either married or together and 89% of people are from the income bracket more than Rs 50000.

- 59% of the entire population who purchase **Fruits** are in the Education Category (Graduation), 84% are either single, married or together and 86% of people are from the income bracket more than Rs 50000.

- 76% of the entire population who purchase **Dairy Products** are in the Education Category (Graduation or PhD), 87% are either single, married or together and 89% of people are from the income bracket more than Rs 50000.

- 57% of the entire population who purchase **Fish Products** are in the Education Category (Graduation), people other than Alone, Absurd and YOLO category evenly consume Fish Products,85 %of people are from the income bracket more than Rs 50000 and 13% are in the bracket of Rs 10000-Rs 50000.

- Channel Wise Visit distribution are as follows.

1.

Marital Status	Channel Wise Visits (%)			
	WebPurchases	CatalogPurchases	StorePurchases	WebVisitsMonth
Single	20.32%	20.93%	20.87%	21.29%
Together	25.87%	26.03%	25.65%	25.62%
Married	38.60%	38.03%	38.97%	38.92%
Divorced	10.93%	10.40%	10.41%	10.69%
Widow	3.89%	4.29%	3.81%	3.16%
Alone	0.16%	0.03%	0.09%	0.16%
Absurd	0.08%	0.25%	0.10%	0.03%
YOLO	0.15%	0.03%	0.09%	0.13%

2.

Education	Channel Wise Visits (%)			
	WebPurchases	CatalogPurchases	StorePurchases	WebVisitsMonth
Graduation	50.81%	51.52%	50.66%	50.04%
PhD	23.50%	24.20%	22.71%	21.46%
Master	16.31%	15.95%	16.82%	16.09%
Basic	1.11%	0.44%	1.19%	3.12%
2n Cycle	8.27%	7.90%	8.62%	9.30%

3.

Income Bucket	Channel Wise Visits (%)			
	WebPurchases	CatalogPurchases	StorePurchases	WebVisitsMonth
<=10000	0.97%	0.70%	0.38%	2.43%
>10000 to <=50000	29.40%	13.23%	28.33%	57.27%
>50000	68.57%	85.33%	70.40%	39.28%

- 96% of the Customer Segment who purchased in the first campaign are with Income greater than Rs 50000.
- The % conversion for each promotional event is as follows.

1.

Income Bucket	Promotion Details (%)						
	DealsPurchases	Cmp1Status	Cmp2Status	Cmp3Status	Cmp4Status	Cmp5Status	Response
<=10000	2.13%	0.00%	0.00%	1.23%	0.00%	0.00%	0.90%
0000 to <=500	46.06%	2.08%	13.33%	53.37%	10.18%	0.00%	36.23%
>50000	50.67%	96.53%	86.67%	45.40%	88.02%	99.39%	62.57%

2.

Marital Status	Promotion Details (%)						
	DealsPurchases	Cmp1Status	Cmp2Status	Cmp3Status	Cmp4Status	Cmp5Status	Response
Single	19.64%	21.53%	16.67%	23.31%	19.76%	19.02%	31.74%
Together	25.88%	22.22%	40.00%	22.70%	25.75%	26.99%	17.96%
Married	39.69%	43.75%	23.33%	38.65%	37.72%	40.49%	29.34%
Divorced	10.85%	8.33%	16.67%	12.27%	10.78%	7.98%	14.37%
Widow	3.46%	3.47%	3.33%	2.45%	5.99%	4.91%	5.69%

3.

Education	Promotion Details (%)						
	DealsPurchases	Cmp1Status	Cmp2Status	Cmp3Status	Cmp4Status	Cmp5Status	Response
Graduation	49.98%	56.94%	53.33%	47.85%	48.50%	52.76%	45.51%
PhD	22.16%	20.83%	33.33%	24.54%	26.95%	23.93%	30.24%
Master	17.24%	12.50%	6.67%	14.72%	18.56%	17.18%	17.07%
Basic	1.86%	0.00%	0.00%	3.68%	0.00%	0.00%	0.60%
2n Cycle	8.76%	9.72%	6.67%	9.20%	5.99%	6.13%	6.59%

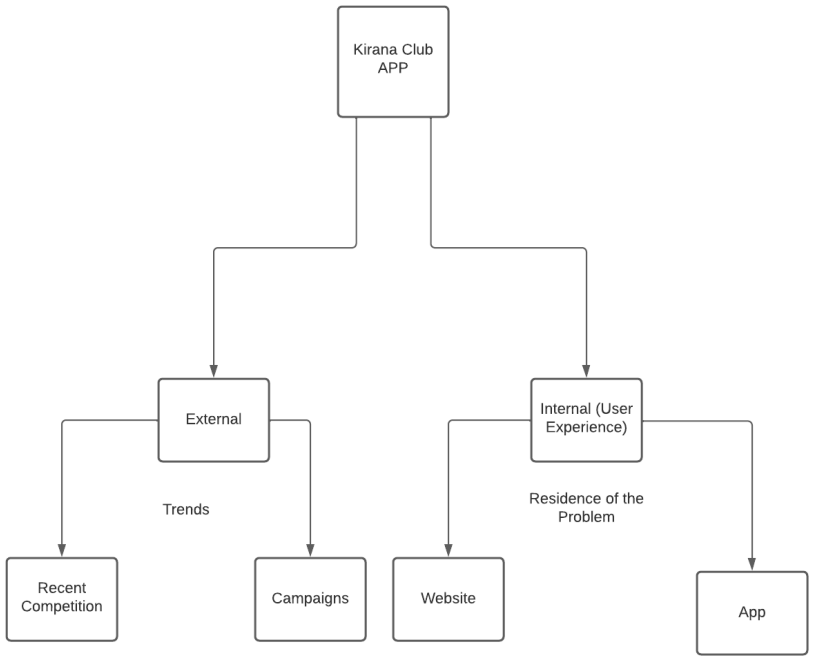
- The Customers can be tagged on the basis of the Recency. The Recency has been divided into 3 buckets.
    - If the Number of Days since customers last purchase is less than 20 days, they can be tagged as **“ACTIVE”** users.
    - If the Number of Days since customers last purchase is between 20 to 30 days, they can be tagged as **“IN ACTIVE”** users.
    - If the Number of Days since customers last purchase is greater than 50 days, they can be tagged as **“CHURNED”** users.

Recency Bucket	Consumer Count

0-20	477
20-50	667
>50	1096

- The Consumers can be clustered based on Income Bracket and Recency Bucket or can be Categorized into Age Groups/User Persona. Apart from this, I have tried to perform a ClusteringModel on the given Data set to categorize the users on various parameters, the Algorithms used are KMeans and KPrototype, since there are a lot of Categorical Variables present in the given data set.
  Link :
  https://colab.research.google.com/drive/1Dk34ku9t-NV34thT2jyBN9-m4cHFG8J-?usp=sharing

  PART 2:



- **Kirana Club** : If the 15 day uninstall rate is spiked to 70%, which means out of all the users who installed the app in the last 15 days, 70% users have

already uninstalled, we can deep dive into the internal and the external factors and check the previous trends based on various metrics like

- Daily Active Users
- D0 Active Users
- D1 Active Users
- D15 Active Users
- Average Engagement time