

Customer Personality Analysis

1 Introduction

In this project we are going to analyze a customer personality dataset containing demographic information, spending habits, promotional responses, and purchasing behaviors. We will conduct hypothesis tests to uncover insights about customer behaviors and characteristics.

2 Data

2.1 Dataset

The data is composed of 29 columns and 2240 entries.

	Id	Year_Birth	Education	Marital_Status	...	Z_CostContact	Z_Revenue	Response
0	5524	1957	Graduation	Single	...	3	11	1
1	2174	1954	Graduation	Single	...	3	11	0
2	4141	1965	Graduation	Single	...	3	11	0

Table [1]: Train Dataset (3 rows × 29 columns) , first 3 entries and 29 features in addition of id.

2.2 Features

There are 29 feature columns. Using these features we can uncover insights about customer behaviors and characteristics.

The features are as follows:

Column	Non-Null Count	Dtype
<hr/>		
0 ID	2240 non-null	int64
1 Year_Birth	2240 non-null	int64
2 Education	2240 non-null	object
3 Marital_Status	2240 non-null	object
4 Income	2216 non-null	float64
5 Kidhome	2240 non-null	int64
6 Teenhome	2240 non-null	int64
7 Dt_Customer	2240 non-null	object
8 Recency	2240 non-null	int64
9 MntWines	2240 non-null	int64
10 MntFruits	2240 non-null	int64
11 MntMeatProducts	2240 non-null	int64
12 MntFishProducts	2240 non-null	int64
13 MntSweetProducts	2240 non-null	int64
14 MntGoldProds	2240 non-null	int64
15 NumDealsPurchases	2240 non-null	int64
16 NumWebPurchases	2240 non-null	int64
17 NumCatalogPurchases	2240 non-null	int64
18 NumStorePurchases	2240 non-null	int64
19 NumWebVisitsMonth	2240 non-null	int64
20 AcceptedCmp3	2240 non-null	int64
21 AcceptedCmp4	2240 non-null	int64
22 AcceptedCmp5	2240 non-null	int64
23 AcceptedCmp1	2240 non-null	int64

24	AcceptedCmp2	2240 non-null	int64
25	Complain	2240 non-null	int64
26	Z_CostContact	2240 non-null	int64
27	Z_Revenue	2240 non-null	int64
28	Response	2240 non-null	int64

We can also see the statistics associated with each feature as well:

Numerical features:

	ID	Year_Birth	Income	Kidhome	Teenhome \
count	2240.000000	2240.000000	2216.000000	2240.000000	2240.000000
mean	5592.159821	1968.805804	52247.251354	0.444196	0.506250
std	3246.662198	11.984069	25173.076661	0.538398	0.544538
min	0.000000	1893.000000	1730.000000	0.000000	0.000000
25%	2828.250000	1959.000000	35303.000000	0.000000	0.000000
50%	5458.500000	1970.000000	51381.500000	0.000000	0.000000
75%	8427.750000	1977.000000	68522.000000	1.000000	1.000000
max	11191.000000	1996.000000	666666.000000	2.000000	2.000000

	Recency	MntWines	MntFruits	MntMeatProducts \
count	2240.000000	2240.000000	2240.000000	2240.000000
mean	49.109375	303.935714	26.302232	166.950000
std	28.962453	336.597393	39.773434	225.715373
min	0.000000	0.000000	0.000000	0.000000
25%	24.000000	23.750000	1.000000	16.000000
50%	49.000000	173.500000	8.000000	67.000000
75%	74.000000	504.250000	33.000000	232.000000
max	99.000000	1493.000000	199.000000	1725.000000

MntFishProducts ... NumWebVisitsMonth AcceptedCmp3 AcceptedCmp4 \

count	2240.000000 ...	2240.000000	2240.000000	2240.000000
mean	37.525446 ...	5.316518	0.072768	0.074554
std	54.628979 ...	2.426645	0.259813	0.262728
min	0.000000 ...	0.000000	0.000000	0.000000
25%	3.000000 ...	3.000000	0.000000	0.000000
50%	12.000000 ...	6.000000	0.000000	0.000000
75%	50.000000 ...	7.000000	0.000000	0.000000
max	259.000000 ...	20.000000	1.000000	1.000000

AcceptedCmp5 AcceptedCmp1 AcceptedCmp2 Complain Z_CostContact \

count	2240.000000	2240.000000	2240.000000	2240.000000	2240.0
mean	0.072768	0.064286	0.013393	0.009375	3.0
std	0.259813	0.245316	0.114976	0.096391	0.0
min	0.000000	0.000000	0.000000	0.000000	3.0
25%	0.000000	0.000000	0.000000	0.000000	3.0
50%	0.000000	0.000000	0.000000	0.000000	3.0
75%	0.000000	0.000000	0.000000	0.000000	3.0
max	1.000000	1.000000	1.000000	1.000000	3.0

	Z_Revenue	Response
count	2240.0	2240.000000
mean	11.0	0.149107
std	0.0	0.356274
min	11.0	0.000000
25%	11.0	0.000000
50%	11.0	0.000000
75%	11.0	0.000000
max	11.0	1.000000

categorical features:

	Education	Marital_Status	Dt_Customer
count	2240	2240	2240
unique	5	8	663
top	Graduation	Married	31-08-2012
freq	1127	864	12

Here we see the unique values for each categorical feature:

Education

['Graduation' 'PhD' 'Master' 'Basic' '2n Cycle']

Marital_Status

['Single' 'Together' 'Married' 'Divorced' 'Widow' 'Alone' 'Absurd' 'YOLO']

Dt_Customer

['04-09-2012' '08-03-2014' '21-08-2013' '10-02-2014' '19-01-2014'
 '09-09-2013' '13-11-2012' '08-05-2013' '06-06-2013' '13-03-2014'
 '15-11-2013' '10-10-2012' '24-11-2012' '24-12-2012' '31-08-2012']

'28-03-2013' '03-11-2012' '08-08-2012' '06-01-2013' '23-12-2012'
'11-01-2014' '18-03-2013' '02-01-2013' '27-05-2013' '20-02-2013'
'31-05-2013' '22-11-2013' '22-05-2014' '11-05-2013' '29-10-2012'
'29-08-2013' '31-12-2013' '02-09-2013' '11-02-2014' '01-02-2013'
'29-04-2013' '12-03-2013' '05-11-2013' '02-10-2013' '28-06-2014'
'09-11-2012' '24-05-2013' '01-01-2014' '08-11-2012' '12-05-2014'
'11-08-2012' '07-06-2014' '12-06-2013' '19-11-2012' '02-04-2013'
'28-04-2014' '17-06-2013' '03-03-2014' '04-07-2013' '07-09-2012'
'18-02-2013' '11-06-2013' '06-12-2013' '21-05-2013' '11-05-2014'
'19-03-2014' '27-09-2013' '08-04-2013' '11-09-2012' '14-09-2012'
'17-03-2013' '05-04-2013' '30-04-2014' '19-12-2012' '27-08-2012'
'12-10-2012' '04-09-2013' '29-08-2012' '23-06-2013' '03-07-2013'
'25-02-2014' '11-08-2013' '16-07-2013' '28-05-2014' '21-01-2014'
'27-05-2014' '23-11-2013' '23-03-2014' '24-05-2014' '22-11-2012'
'11-04-2013' '01-12-2013' '20-06-2013' '23-07-2013' '30-03-2014'
'20-04-2013' '17-05-2013' '08-05-2014' '10-12-2013' '24-09-2013'
'02-02-2013' '07-12-2012' '02-11-2013' '10-11-2012' '25-06-2014'
'12-12-2012' '10-09-2013' '24-01-2014' '19-02-2013' '11-07-2013'
'14-11-2013' '24-02-2014' '22-04-2013' '21-04-2013' '08-06-2014'
'10-05-2014' '30-09-2013' '10-03-2013' '14-08-2013' '07-07-2013'
'19-05-2014' '28-07-2013' '19-10-2012' '19-04-2013' '28-08-2013'
'19-03-2013' '18-10-2012' '28-10-2012' '22-08-2012' '21-06-2014'
'24-04-2014' '07-03-2014' '14-12-2012' '09-10-2012' '08-07-2013'
'12-09-2013' '17-10-2013' '20-08-2013' '01-07-2013' '05-11-2012'
'05-01-2014' '01-05-2013' '01-03-2014' '13-11-2013' '18-11-2013'
'21-04-2014' '13-07-2013' '30-01-2014' '04-04-2014' '12-09-2012'
'16-12-2012' '23-05-2014' '24-06-2014' '28-09-2013' '22-01-2014'

'15-06-2014' '05-12-2012' '02-08-2013' '23-02-2013' '09-09-2012'
'26-09-2013' '30-05-2013' '29-12-2013' '13-05-2014' '19-09-2013'
'17-01-2013' '31-03-2014' '29-06-2014' '09-12-2013' '14-10-2013'
'02-11-2012' '17-07-2013' '08-12-2013' '13-05-2013' '10-09-2012'
'31-03-2013' '18-03-2014' '05-10-2012' '21-01-2013' '04-05-2013'
'01-04-2014' '31-08-2013' '14-11-2012' '11-04-2014' '14-04-2014'
'05-01-2013' '08-09-2012' '26-03-2013' '25-10-2012' '09-12-2012'
'16-02-2014' '20-03-2013' '15-05-2013' '16-04-2014' '23-03-2013'
'04-12-2013' '04-10-2013' '26-12-2013' '17-05-2014' '28-10-2013'
'31-07-2013' '28-11-2013' '17-08-2012' '05-06-2014' '20-05-2013'
'09-01-2013' '16-09-2013' '27-10-2013' '22-09-2012' '13-10-2012'
'16-10-2012' '22-10-2012' '04-06-2013' '22-05-2013' '17-06-2014'
'23-11-2012' '03-02-2013' '14-03-2013' '26-06-2014' '15-10-2012'
'19-06-2013' '20-03-2014' '04-02-2014' '06-04-2014' '06-02-2013'
'11-02-2013' '17-04-2014' '12-07-2013' '29-01-2013' '08-01-2013'
'13-06-2013' '27-12-2013' '31-10-2012' '15-01-2014' '23-08-2012'
'07-03-2013' '13-01-2013' '12-01-2013' '17-03-2014' '12-10-2013'
'13-04-2014' '18-09-2012' '05-03-2014' '27-04-2013' '18-01-2014'
'03-06-2013' '17-12-2013' '11-03-2014' '29-07-2013' '14-08-2012'
'23-08-2013' '09-02-2014' '07-02-2013' '11-01-2013' '05-07-2013'
'02-07-2013' '07-11-2013' '09-05-2013' '13-02-2013' '16-04-2013'
'11-09-2013' '03-04-2013' '10-01-2013' '30-06-2013' '06-12-2012'
'12-11-2012' '03-03-2013' '10-08-2012' '07-12-2013' '15-08-2013'
'10-11-2013' '16-06-2014' '25-12-2012' '03-01-2014' '27-10-2012'
'22-12-2012' '29-11-2013' '08-10-2013' '28-09-2012' '22-03-2014'
'28-12-2012' '21-08-2012' '16-03-2013' '17-11-2012' '01-12-2012'
'22-04-2014' '11-11-2012' '22-06-2013' '18-08-2012' '30-12-2012'

'14-06-2013' '16-10-2013' '30-08-2012' '04-05-2014' '18-04-2013'
'06-10-2013' '15-09-2012' '27-09-2012' '11-03-2013' '22-10-2013'
'09-06-2014' '30-05-2014' '17-10-2012' '30-03-2013' '23-01-2013'
'20-11-2013' '14-02-2014' '22-02-2013' '05-03-2013' '06-05-2014'
'13-04-2013' '05-04-2014' '25-04-2013' '25-11-2013' '02-02-2014'
'21-06-2013' '21-12-2013' '07-09-2013' '22-08-2013' '20-12-2013'
'06-08-2013' '09-10-2013' '09-06-2013' '22-12-2013' '02-05-2013'
'16-02-2013' '20-08-2012' '01-04-2013' '25-05-2014' '25-09-2012'
'22-09-2013' '28-01-2014' '23-04-2013' '06-03-2014' '15-08-2012'
'27-03-2013' '10-10-2013' '04-08-2012' '18-05-2014' '09-05-2014'
'23-10-2012' '30-12-2013' '19-08-2012' '30-10-2013' '26-06-2013'
'26-01-2014' '24-09-2012' '03-12-2012' '27-11-2013' '06-08-2012'
'29-12-2012' '24-08-2012' '07-11-2012' '02-12-2013' '19-07-2013'
'25-05-2013' '20-10-2013' '14-04-2013' '27-01-2014' '16-08-2013'
'30-09-2012' '04-12-2012' '12-08-2012' '03-11-2013' '15-02-2013'
'18-09-2013' '18-06-2014' '24-04-2013' '15-02-2014' '08-08-2013'
'23-12-2013' '28-11-2012' '07-08-2013' '07-01-2014' '03-12-2013'
'29-03-2014' '08-04-2014' '07-02-2014' '06-11-2013' '01-03-2013'
'03-08-2013' '16-06-2013' '25-06-2013' '13-09-2013' '01-08-2013'
'31-01-2013' '09-07-2013' '15-04-2013' '03-04-2014' '18-04-2014'
'23-09-2012' '14-09-2013' '26-10-2012' '19-10-2013' '21-02-2013'
'17-09-2013' '26-01-2013' '07-05-2014' '19-08-2013' '18-07-2013'
'11-12-2012' '26-07-2013' '06-05-2013' '14-07-2013' '10-03-2014'
'28-12-2013' '20-01-2013' '21-09-2012' '06-09-2012' '18-12-2012'
'19-11-2013' '13-06-2014' '07-05-2013' '18-08-2013' '13-12-2013'
'17-02-2014' '03-08-2012' '04-01-2014' '21-10-2013' '25-12-2013'
'01-01-2013' '16-05-2013' '25-09-2013' '08-11-2013' '16-12-2013'

'30-10-2012' '02-01-2014' '18-10-2013' '25-07-2013' '24-10-2013'
'19-12-2013' '16-05-2014' '04-10-2012' '25-08-2013' '29-05-2013'
'26-03-2014' '09-04-2014' '12-03-2014' '29-04-2014' '27-06-2014'
'10-04-2013' '21-11-2013' '21-10-2012' '14-05-2013' '01-08-2012'
'25-01-2014' '11-10-2013' '07-08-2012' '29-03-2013' '03-06-2014'
'09-08-2012' '06-03-2013' '12-05-2013' '09-03-2014' '15-09-2013'
'29-01-2014' '03-10-2013' '28-02-2013' '02-05-2014' '09-02-2013'
'15-03-2013' '10-05-2013' '05-09-2012' '01-09-2013' '26-05-2013'
'12-12-2013' '08-12-2012' '30-08-2013' '18-12-2013' '18-06-2013'
'29-05-2014' '15-01-2013' '29-10-2013' '02-03-2013' '14-12-2013'
'19-01-2013' '06-04-2013' '04-11-2013' '22-03-2013' '31-05-2014'
'04-02-2013' '05-05-2013' '14-02-2013' '21-03-2014' '05-08-2013'
'29-11-2012' '27-07-2013' '13-01-2014' '29-09-2012' '23-10-2013'
'06-11-2012' '26-05-2014' '04-03-2014' '06-01-2014' '11-06-2014'
'19-06-2014' '26-11-2012' '11-12-2013' '14-01-2014' '23-01-2014'
'24-03-2013' '29-06-2013' '25-01-2013' '17-11-2013' '18-05-2013'
'17-04-2013' '24-01-2013' '09-08-2013' '10-06-2013' '20-07-2013'
'28-05-2013' '04-03-2013' '01-10-2012' '10-02-2013' '14-10-2012'
'03-02-2014' '22-06-2014' '25-10-2013' '18-02-2014' '09-11-2013'
'13-08-2013' '25-08-2012' '19-04-2014' '07-04-2014' '03-01-2013'
'04-11-2012' '22-01-2013' '26-02-2014' '23-04-2014' '30-07-2013'
'13-10-2013' '27-11-2012' '01-05-2014' '30-11-2012' '17-08-2013'
'04-01-2013' '03-05-2014' '26-04-2014' '02-06-2014' '26-10-2013'
'24-03-2014' '25-02-2013' '20-09-2013' '16-01-2013' '24-12-2013'
'18-11-2012' '05-08-2012' '03-05-2013' '12-11-2013' '06-07-2013'
'21-07-2013' '20-10-2012' '23-05-2013' '26-08-2012' '12-01-2014'
'13-02-2014' '27-02-2014' '20-04-2014' '28-06-2013' '28-04-2013'

'21-03-2013' '27-08-2013' '30-07-2012' '05-12-2013' '08-02-2013'
'21-11-2012' '26-02-2013' '20-09-2012' '08-06-2013' '05-09-2013'
'12-04-2014' '24-06-2013' '19-02-2014' '27-12-2012' '01-11-2012'
'22-07-2013' '31-01-2014' '26-08-2013' '01-10-2013' '10-06-2014'
'01-02-2014' '03-09-2012' '26-09-2012' '31-07-2012' '02-12-2012'
'05-05-2014' '06-10-2012' '02-03-2014' '15-05-2014' '10-12-2012'
'13-08-2012' '23-02-2014' '16-03-2014' '21-09-2013' '24-10-2012'
'15-10-2013' '17-02-2013' '28-08-2012' '12-02-2014' '08-02-2014'
'01-09-2012' '16-11-2012' '12-02-2013' '02-08-2012' '16-08-2012'
'25-11-2012' '27-04-2014' '30-04-2013' '20-11-2012' '14-06-2014'
'28-03-2014' '08-03-2013' '02-09-2012' '15-04-2014' '25-04-2014'
'27-06-2013' '01-06-2013' '05-10-2013' '22-02-2014' '12-04-2013'
'23-09-2013' '24-07-2013' '23-06-2014' '10-01-2014' '25-03-2013'
'17-01-2014' '04-06-2014' '25-03-2014' '31-12-2012' '19-09-2012'
'27-02-2013' '02-10-2012' '09-04-2013' '02-04-2014' '07-01-2013'
'26-11-2013' '04-08-2013' '17-12-2012' '06-06-2014' '19-05-2013'
'12-08-2013' '15-12-2013' '07-10-2012' '06-09-2013' '15-07-2013'
'27-01-2013' '05-02-2014' '26-12-2012' '06-02-2014' '14-01-2013'
'20-06-2014' '24-08-2013' '28-02-2014' '07-04-2013' '10-04-2014'
'12-06-2014' '30-11-2013' '09-03-2013' '27-03-2014' '15-12-2012'
'17-09-2012' '02-06-2013' '21-12-2012' '01-11-2013' '10-08-2013'
'11-10-2012' '20-12-2012' '09-01-2014']

Here we see the number of missing values for each col. Since almost 1% of our whole data is missed, handling this by dropping the null values seems to be logical.

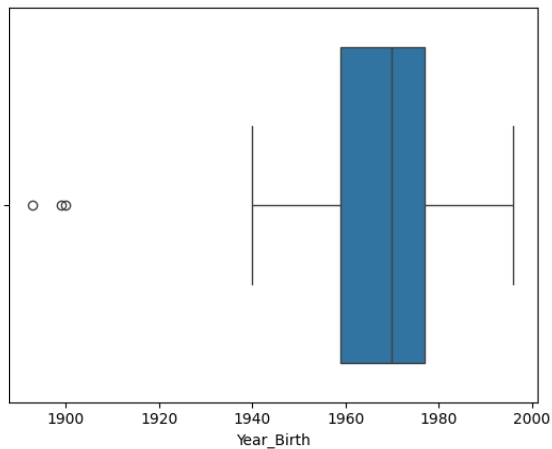
ID	0
Year_Birth	0
Education	0
Marital_Status	0
Income	24
Kidhome	0
Teenhome	0
Dt_Customer	0
Recency	0
MntWines	0
MntFruits	0
MntMeatProducts	0
MntFishProducts	0
MntSweetProducts	0
MntGoldProds	0
NumDealsPurchases	0
NumWebPurchases	0
NumCatalogPurchases	0
NumStorePurchases	0
NumWebVisitsMonth	0
AcceptedCmp3	0
AcceptedCmp4	0
AcceptedCmp5	0
AcceptedCmp1	0
AcceptedCmp2	0
Complain	0
Z_CostContact	0

Z_Revenue	0
Response	0

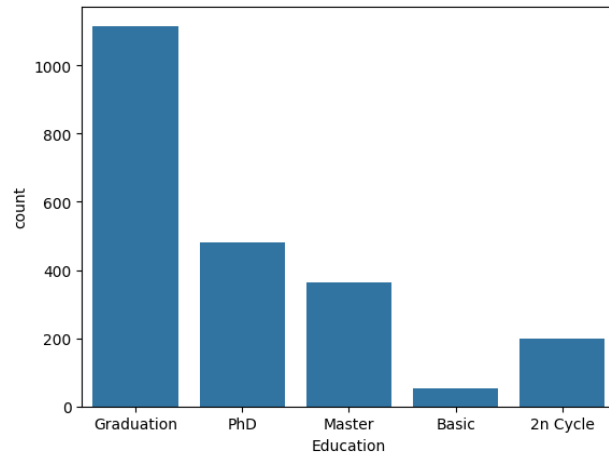
So after dropping them the number of missing values will be 0 and we will have 2216 rows and 29 cols.

2.3 visualization

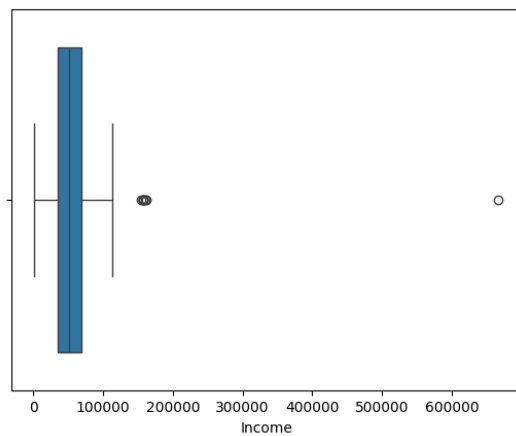
In this section we will go through some visualization related to the data that can help us to have a better sense of our data and the relationship between the features.



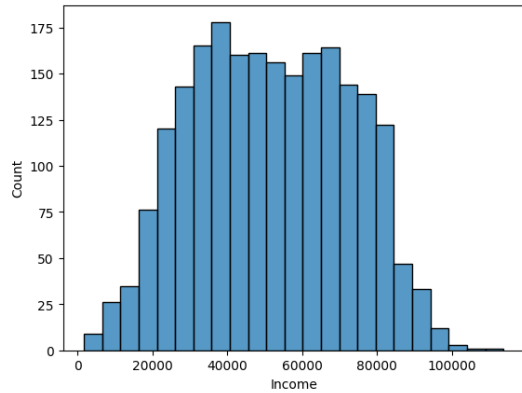
Distribution of birth year. Very few outliers exist (only 3) from the 19th century. I drop them. Most of customers were born between the 60s and the 80s



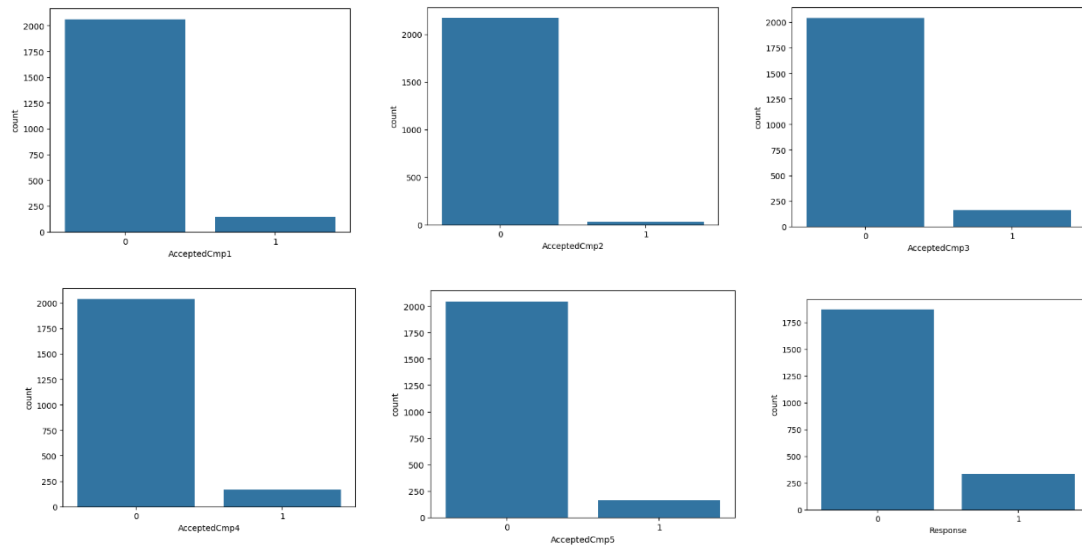
Education levels among the customers. Most customers had a graduation level education while the Least amount had Basic level.



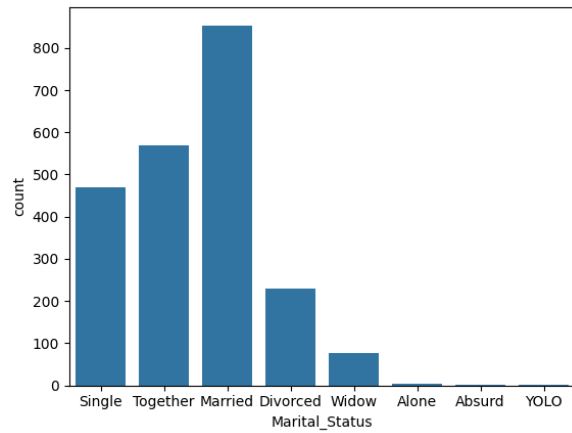
Distribution of income. Very few outliers exist. I drop them.



This plot is the distribution of income but without the outliers. We can see that Most customers have income between 40000 and 80000.



In these plots we can see the popularity of each campaign. It seems that the Last Campaign was the most successful one and the Second Campaign was the least successful one The rest were nearly equal



In this plot we can see the marital status of customers. As we can see most of the customers have partners. Absurd and YOLO seem to be very little in amount here so I'll drop them.

3 Hypothesis Tests

3.1.1 Hypotheses:

H0 (Null Hypothesis): There is no significant difference in income levels among customers with different education levels.

H1 (Alternative Hypothesis): There is a significant difference in income levels among customers with different education levels.

3.1.2 Test:

To test whether customers with different education backgrounds earn significantly different incomes, use the Kruskal-Wallis H test. Since income data may not be normally distributed, this non-parametric test is appropriate for comparing median income levels across multiple education categories.

3.1.3 Result:

Kruskal-Wallis H Statistic:

140.11991249406134

P-value:

2.6605568012607163e-29

Reject the null hypothesis: There is a significant difference in income levels among education levels.

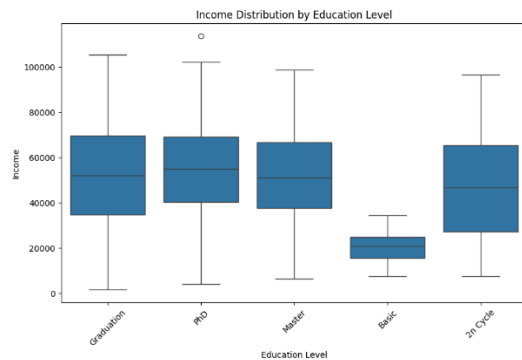
Since the p value is less than 0.05 (alpha), H_0 is rejected.

3.1.4 Visualization:

Here we can see the plot of Income Distribution by Education Level.

As we can see there is a significant relationship between income and graduation level.

The ones with higher education seem to have higher incomes.



3.2.1 Hypotheses:

H0 (Null Hypothesis): there is no significant difference in spending behavior between customers who received the campaign and those who did not.

H1 (Alternative Hypothesis): there is a significant difference in spending behavior between customers who received the campaign and those who did not

3.2.2 Test:

We Compare the spending behavior of customers who received a marketing campaign versus those who did not. First check the distribution of data and use the Mann-Whitney U Test if the data is not normally distributed, or the Independent Samples T-test if normality holds.

3.2.3 Result:

Shapiro-Wilk p-value (Campaign Group):

4.4484650500614455e-12

Shapiro-Wilk p-value (No Campaign Group):

1.7745478269729e-38

As we can see the distribution is not normal (because the p values are less than 0.05) so we use Mann-Whitney U Test.

Test Used:

Mann-Whitney U Test

Test Statistic::

428556.5

P-value:

5.552195334026083e-29

Reject the null hypothesis: The marketing campaign influences spending behavior.

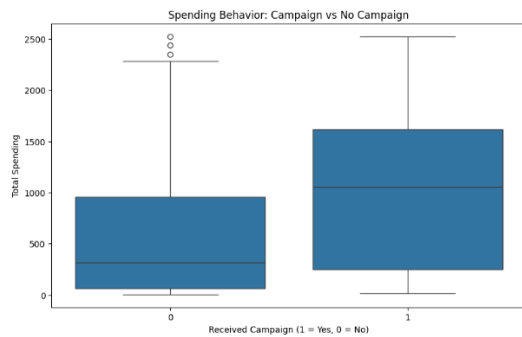
Since the p value is less than 0.05 (alpha), H_0 is rejected.

3.2.4 Visualization:

Here we can see the plot of Spending Behavior: Campaign vs No Campaign.

As we can see The marketing campaign influences spending behavior.

The ones who received the campaign have spent more than the ones who did not.



3.3.1 Hypotheses:

H0 (Null Hypothesis): there is no significant difference in spending behavior between the two groups (Customers with/without children) .

H1 (Alternative Hypothesis): there is a significant difference in spending behavior between the two groups.

3.3.2 Test:

To determine whether households with children allocate their spending differently compared to those without we use the Mann-Whitney U Test. This test compares the spending distributions of two independent groups when the assumption of normality is uncertain.

3.3.3 Result:

Mann-Whitney U Statistic:

198854.5

P-value:

5.406858767067534e-106

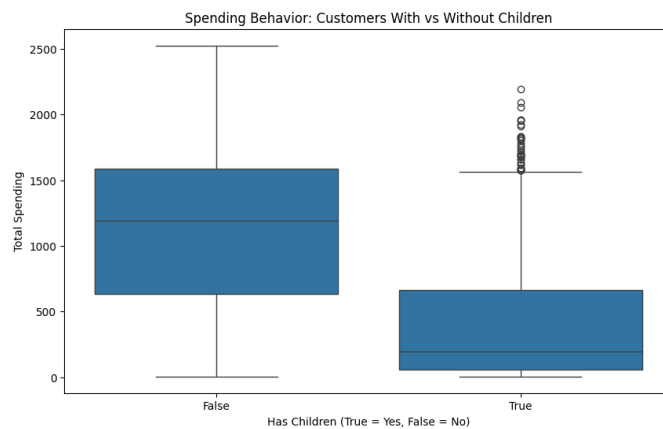
Reject the null hypothesis: Customers with children spend differently than those without children. Since the p value is less than 0.05 (alpha), H_0 is rejected.

3.3.4 Visualization:

Here we can see the plot of Spending Behavior: Customers With vs Without Children.

As we can see having or not having children influences spending behavior.

The plot likely illustrates how spending behavior differs between these two groups, And we can obviously see that the spending behavior is different between these two groups.



3.4.1 Hypotheses:

H0 (Null Hypothesis): There is no significant difference in spending across different product categories

H1 (Alternative Hypothesis): There is a significant difference in spending across different product categories

3.4.2 Test:

We Use the Friedman Test to compare customer spending across multiple product categories . This test will help determine whether customers have a preference for spending more on certain product categories over others.

3.4.3 Result:

Friedman Test Statistic:

5846.870183284754

P-value:

0.0

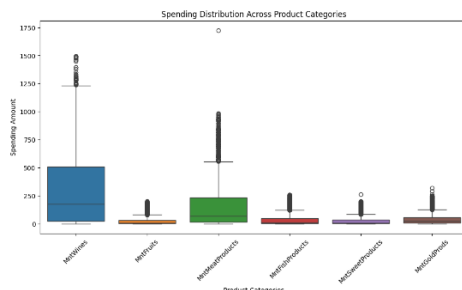
Reject the null hypothesis: There is a significant difference in spending across product categories. Since the p value is less than 0.05 (alpha), h0 is rejected.

3.4.4 Visualization:

Here we can see the plot of Spending Distribution Across Product Categories.

As we can see There is a significant difference in spending across product categories.

The plot likely shows that the spending among different products is different. Wine and meat products seem to be more popular than the others.



3.5.1 Hypotheses:

H0 (Null Hypothesis): There is no relationship between customer education level and acceptance of promotional campaigns

H1 (Alternative Hypothesis): There is a relationship between customer education level and acceptance of promotional campaigns

3.5.2 Test:

We Use the Chi-Square Test for Independence to determine whether education level is associated with customers accepting promotional campaigns. This test will analyze whether different education groups have significantly different response rates to campaign offers.

3.5.3 Result:

Contingency Table:

Response	0	1
Education		
2n Cycle	176	22
Basic	52	2
Graduation	961	151
Master	307	56
PhD	374	100

Chi-Square Statistic:

23.333703320850276

P-value:

0.00010859738991304676

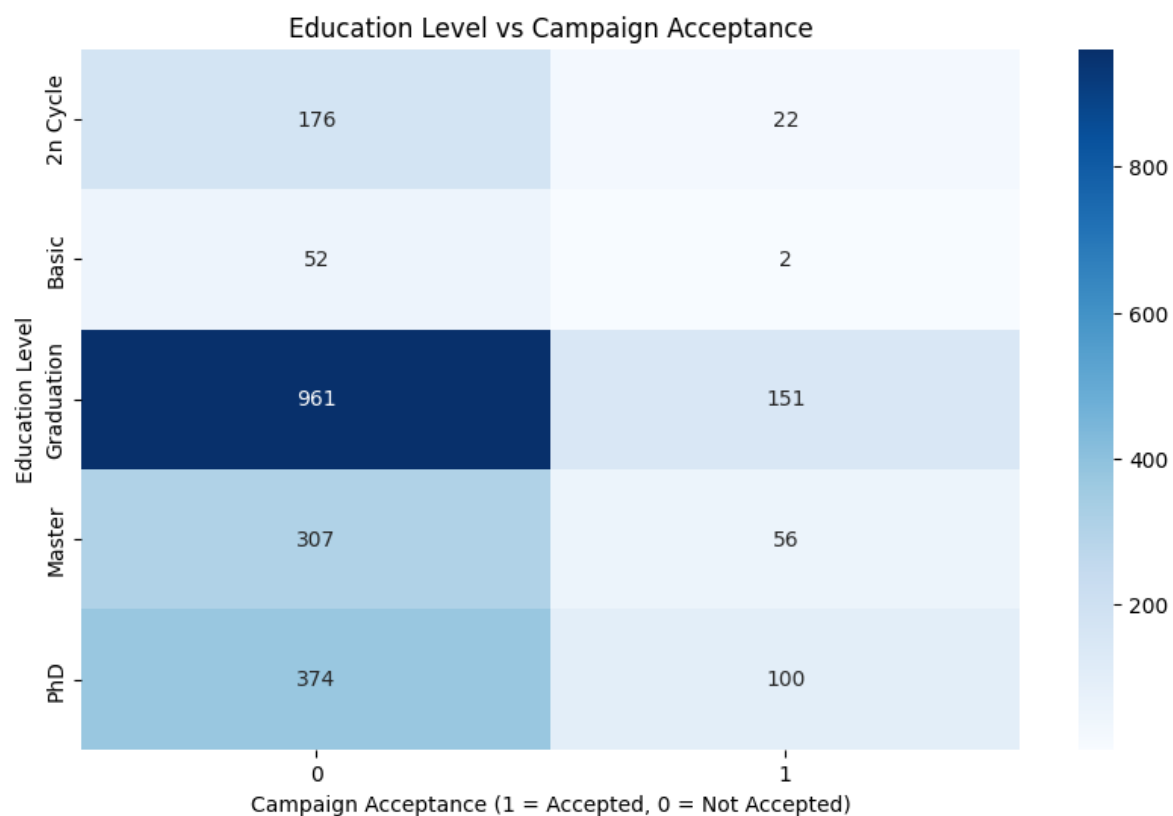
Reject the null hypothesis: There is a significant relationship between education level and campaign acceptance. Since the p value is less than 0.05 (alpha), H_0 is rejected.

3.5.4 Visualization:

Here we can see the plot of Education Level vs Campaign Acceptance

As we can see : There is a significant relationship between education level and campaign acceptance

The plot indicates that individuals with higher education levels (PhD, Master) tend to have a higher rate of campaign acceptance compared to those with a basic education level



4 Additional Tests and Visualizations

4.1.1 Hypotheses:

Relationship Between Income and Spending Behavior

Hypothesis:

H0: There is no significant correlation between customer income and total spending.

H1: There is a significant correlation between customer income and total spending.

4.1.2 Test:

Spearman's Rank Correlation Test (for monotonic relationships). because :Spearman's correlation is more robust and It provides meaningful insights even when the relationship between variables is not strictly linear.

4.1.3 Result:

Spearman Correlation:

0.8601746860697305

P-value:

0.0

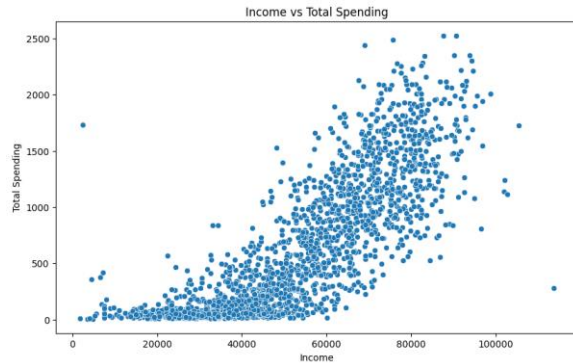
Reject the null hypothesis: There is a significant monotonic relationship between Income and Total Spending. Since the p value is less than 0.05 (alpha), h0 is rejected.

4.1.4 Visualization:

Here we can see the plot of Income vs Total Spending

As we can see : There is a significant relationship between Income level and spending

The plot indicates that individuals with higher income are likely to spend more than the ones with lower incomes.



4.2.1 Hypotheses:

Null Hypothesis (H0): There is no significant difference in the mean income between customers who accepted the marketing campaign and those who did not.

Alternative Hypothesis (H1): There is a significant difference in the mean income between customers who accepted the marketing campaign and those who did not.

4.2.2 Test:

t-test to compare the mean income of two groups, such as customers who accepted a marketing campaign versus those who did not

4.2.3 Result:

T-statistic:

-0.5144389258045258

P-value:

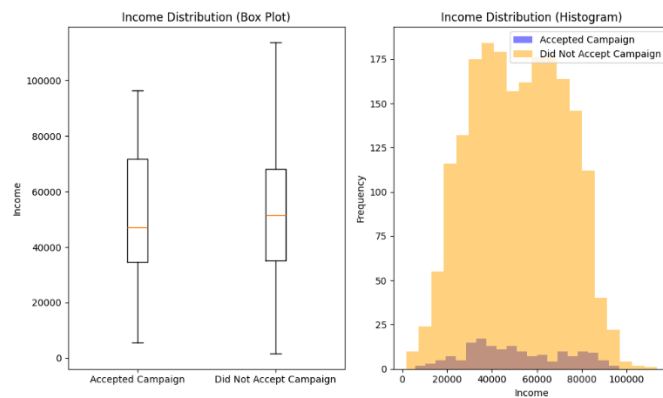
0.6069968464537736

Fail to reject the null hypothesis: There is no significant difference in mean income between the two groups.. Since the p value is more than 0.05 (alpha), we are failed to reject H_0 .

4.2.4 Visualization:

Here we can see the box and histogram plots of Income Distribution

As we can see : There is no significant difference in mean income between the two groups



5 Data Storytelling

Customer Demographics and Income Levels Key Finding:

most of customers were born between the 1960s and 1980s and most of them have a graduation level education. Incomes are mostly between 40,000 and 80,000 and people with higher education levels have higher income.

Insight: Customers with higher education levels is more likely to have higher income making them a prime target for premium products and services.

Strategy:

Marketing Strategy: Tailor marketing campaigns to target higher income and well-educated customers with premium product offerings.

Product Development: Develop luxury goods or exclusive services.

Customer Engagement: Offer loyalty programs or discounts to high income customers to increase lifetime value.

Marketing Campaign Effectiveness Key Finding:

The last marketing campaign was the most successful while the second campaign was the least effective. Customers who accepted the campaign spent significantly more than those who did not.

Insight: Marketing campaigns have a direct effect on customer spending behavior and some campaigns are more effective than others.

Actionable Strategy:

Marketing Strategy: analyzing the last campaign factors to point out why it was successful and apply the findings to new campaigns. the factors like :messaging timing or channels. And also do the same thing about the unsuccessful campaigns to discover why they were not successful and do not repeat those factors in new ones. And also those unsuccessful campaigns must be stopped and not to be continued.

Customer Engagement: Use personalised marketing strategies to target customers who are more likely to respond positively to campaigns based on their past behavior.

Product Development: continue products or services that were popular during successful campaigns to have additional sales.

Spending Behavior and Customer Segments Key Finding:

Customers with children spend differently than the ones without children and there is a significant difference in spending across product categories. Wine and meat products are the most popular on the other hand fruits and sweets are less popular.

Insight: Family status and product preferences are important in customer spending behavior.

Actionable Strategy:

Marketing Strategy: Create targeted campaigns for families, emphasizing products that are more likely to be their needs .

Product Development: Focus on promote the product lines that are more popular (wines and meat products) and reducing less popular products.

Customer Engagement: Offer family promotions or discounts to customers with children to increase their spending.

Education Level and Campaign Acceptance Key Finding:

There is a significant relationship between education and campaign acceptance. Customers with higher education levels are more likely to accept marketing campaign

Insight: Education level is a important factor to make campaign success because more educated customers are more likely accept them.

Actionable Strategy:

Marketing Strategy: Design campaigns that are appropriate to highly educated customers such as those emphasizing quality and innovation

Customer Engagement: Use educational content in marketing materials to engage this segment such as product tutorials expert reviews or industry insights.

Product Development: Develop products that are at the same level with the values of educated customers such tech products.

Conclusion

The analysis show several keys insights that can drive business growth and improve customer satisfaction. If we focus on high income well educated customers providing marketing campaigns to their preferences and optimizing product offerings based on spending behavior the business can enhance its marketing effectiveness increase customer involvement and results to revenue growth. Additionally understanding the impact of family status and product preferences allows for more targeted and personalized strategies leading to a stronger competitive advantage in the market.

References

[1] <https://www.kaggle.com/datasets/imakash3011/customer-personality-analysis/data>

Maryam Soleimani 401222075