Introduction :

Customer Personality Analysis is a detailed analysis of a company’s ideal customers. It helps a business to better understand its customers and makes it easier for them to modify products according to the specific needs, behaviors and concerns of different types of customers.

Modeling :

I divided the dataset into three tables(customer fact, customer dim and date dim).

Customer Dim Attributes:

This section includes demographic information about the customers, such as their ID, year of birth, education level, marital status.

Date Dim Attributes:

This section includes information about the customer's enrollment date with the company(dt\_customer), I extract from that date year, quarter, month and day.

Customer Fact Attributes:

This section includes information about all measures in my dataset such as Income, Recency, Complain, Kidhome, Teenhome, MntWines, MntFruits, MntMeatProducts, MntFishProducts, MntSweetProducts, MntGoldProducts, NumDealsPurchases, NumWebPurchases, NumStorePurchases, NumCatlogPurchases, NumWebVisitsMonth, AcceptedCmp1, AcceptedCmp2, AcceptedCmp3, AcceptedCmp4, AcceptedCmp5, Response, and ID, dt\_customer as forighen keys to the dimensions tables.

Overall, this dataset can be used to analyze and understand customer behavior and preferences, as well as to identify the customer segments that are most likely to respond to specific marketing campaigns and promotions. It can also be used to identify patterns and trends in the customer data, such as the correlation between income and purchase behavior, or the effect of marketing campaigns on customer response. The insights gained from this analysis can help the company to optimize its marketing and sales strategies, as well as to improve customer satisfaction and loyalty.

Cleaning:

First, I removed the two Columns (Z\_Revenue & Z\_CostContact) Because every column of them has one value and is not useful for my data.

Then I created an age column by substructing the birth year column from 2015(I assumed that I get this data in 2015 and I am doing the analysis on the three years before(2014, 2013, 2012).

I changed the type of dt\_customer from character to date by location(United Kingdom).

Dax Calculations :

1. I made the column age label using if statement and divided the age to these labels:

* 19 – 24 young adult
* 25 – 39 adult
* 40 – 59 middle aged
* 60 – 79 senior
* 80 – 99 eldery
* 100 – 121 centenarian

Note: I didn’t remove the ages above 100 because there are people who exceed 100 and are still alive so these people may be made their purchases on the website or another person (relatives or something like that) made purchases for them from the store.

1. I made the column Income label using if statement and divided the Income to these labels:

* <= 2000 Low Income
* <= 70000 Meduim Income
* Greater than that High Income

1. I made these measures like ResponseRateCmp for all campaigns from 1 to 5 by getting the sum of AcceptedCmp and dividing by the number of customers, Total Purchases which is the sum of all purchases made by different categories (NumDealsPurchases, NumWebPurchases, NumStorePurchases, NumCatlogPurchases) and Total Products which is the sum of all amount of products.
2. I made the date hierarchy.

Story:

In this dataset, I was trying to investigate customer personalities that will help the store make the customer happier and increase its sales.

I started by finding out if there is a correlation between the income of every customer and the total purchase that he made and I found that there is a strong correlation even if the correlation not meaning causation but of course, the total purchases would affect by the income because logically the customers who earned a low income would make lower purchases (as much as they need).

Then I discovered if the level of education affects income and I found out that as the education level grows the income grows too.

After that I was curious about the total purchases that customers made through the years, I found that 2013 has the highest number of purchases between the three years with an obviously different and also in each category of the purchases, also discovered that July has the least purchases we have to investigate what is happening in that month.

The number of purchases in the store is higher than those on the website so we have to find out if there any problems with the website like a bad user interface or may the store didn’t make enough discounts on it like inside the store.

After that, I looked at the products, and the most sold product is the wine which was shocking for me in the beginning what about the underages?! Are they drinking wine too, I investigate more in this case and I found out that most of the customers who drink wine are middle-aged (between 40 and 59) and also the age of customers starts from 19 so I don’t have many under age in the data.

Also, I discovered that married people are the most customers making purchases on all categories but this is logical as most of my customers are married.

There is also something interesting that customers who had a medium-income made purchases of gold and wine more than those who had a high income.

In the final, I discovered the rate response of the five campaigns that the store made and they don’t seem like they achieve their goals, there is a very lower rate of customers that accepted the campaigns, not even reaching 4% in any campaign over the three years so the store has to work more on all campaigns with the marketing team.

On the other hand, there are a very low number of complaints from customers only 21 of the whole 2240 customers so we may say that the customer made a successful work in that and make most of the customers satisfied.