Marketing Campaign Case Study *-Walkthrough by Eslam Morsy*

# Target Audience

This dashboard is intended to be viewed/used by marketing personnel, data analysts and data scientists. It could also be viewed by Marketing leads, however, it needs some sharpening in order to be more concise to be able to present it to C level managers.

The dashboard contains filters that are synced across all charts. The viewer can mix and match and answer questions that will give a clearer understanding of the impact of the different marketing campaigns, what good did we do that we could repeat, and how can we customize our next marketing campaigns for better results.

# Chart Documentation

### Assumptions:

1. I assumed that the current time in the case study is Jan 1st, 2015. I found that if I took today’s date, some customers’ age would have exceeded 120 years.
2. I assumed that column “**Response**” referred to a sixth marketing campaign other than the five campaigns mentioned.
3. I assumed that if a customer made a purchase in the last 30 days, then he/she’s considered an *Active* customer. If a customer made a purchase between 30 and 60 days ago, then he/she’s considered a *Semi-active* customer. If a customer made a purchase more than 60 days ago, then he/she’s considered a *Dormant* customer.

### Technical Details and Features

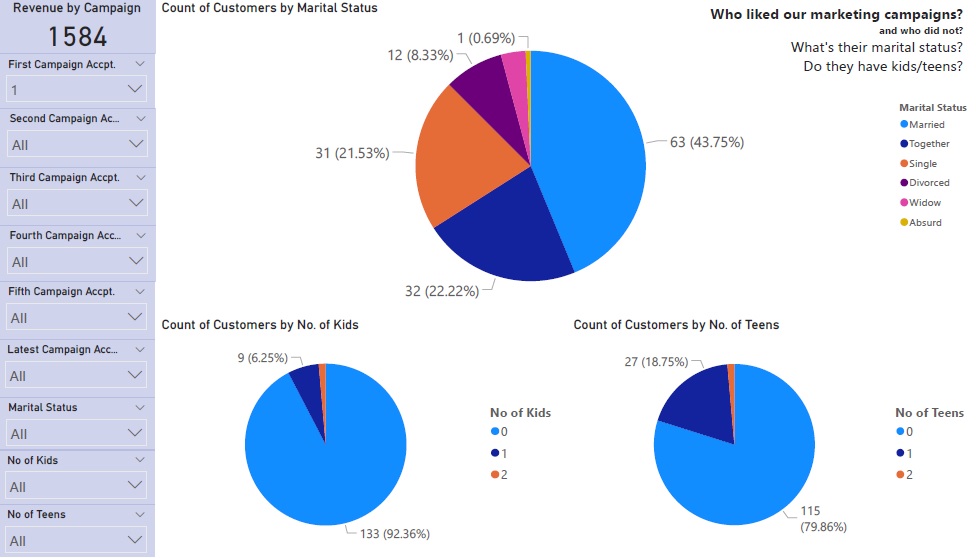
1. I’m working on Power Bi version: 2.92.943.0 64-bit (April 2021).
2. The first five pages in the attached PDF represent the first dashboard. The sixth and last page represent the second dashboard.
3. The filters on the left of each page of the first five pages are synced through the first dashboard. Meaning, if you change a filter on the first page, it will automatically reflect on the rest of the pages in the dashboard.
4. The last page *(the second dashboard)* has its own set of filters.
5. You can also monitor the “**Revenue by Campaign**” on the top left corner in the first dashboard.

### Walkthrough

1. The First Dashboard *(the first five pages in the PDF)*

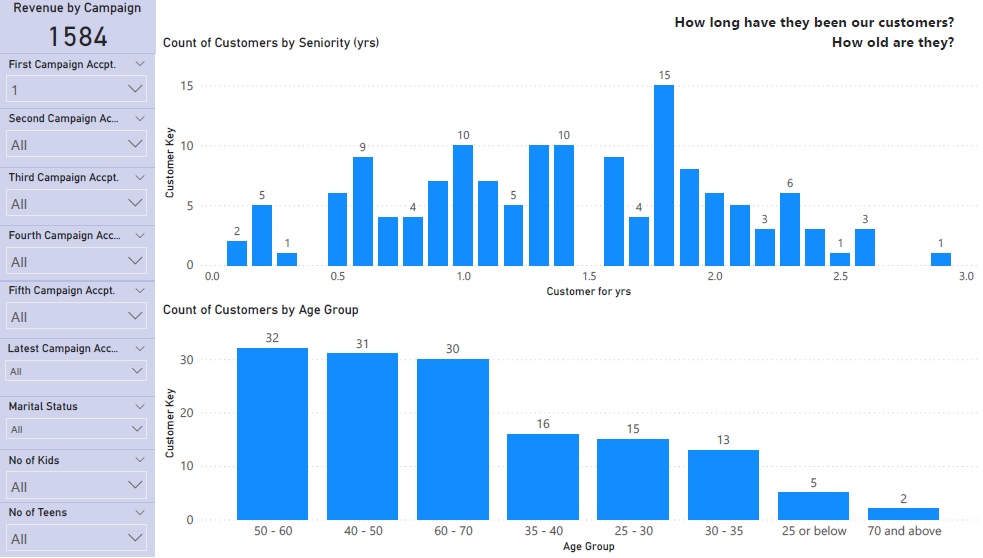
“… *different marketing campaigns work better for specific clusters of customers*.” I took that sentence and built the entire first dashboard on it. I asked myself a couple of questions that could tell me more about each group of customers:

1. Who liked our marketing campaign/s? Are they married or what? Do they have kids? Do they have teens?



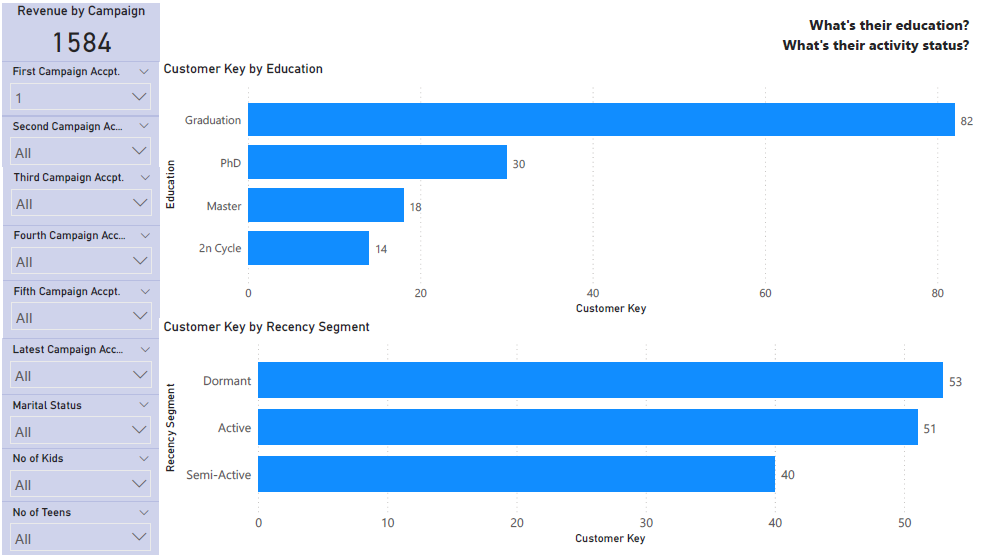
Here for example, I wanted to know who liked our first campaign, I found out that the majority of them are couples *(Married or Together).* I also found out that the majority doesn’t have kids or teens at home.

1. How long have they been our customers? And how old are they?



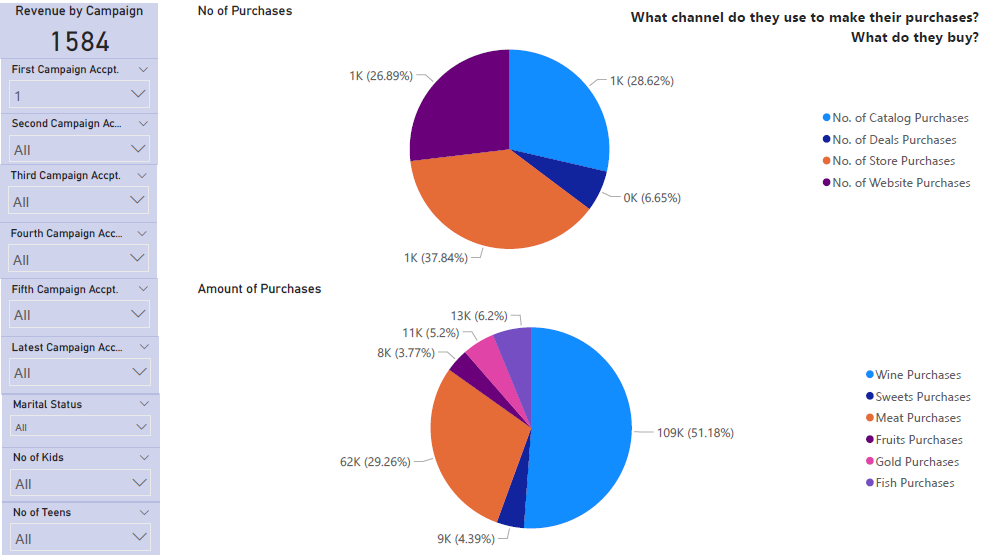
I found out that most of the customers who liked our first campaign were our customers for at least six months. The majority aged between 40 and 70 years old.

1. What’s their education? And what’s their activity status?



I found out that the majority of customers who liked our first campaign were graduates then PhD holders. And to my surprise, the majority were *Dormant* customers which is supercool, because these customers were on the verge of churn. Chapeau to the campaign for this point.

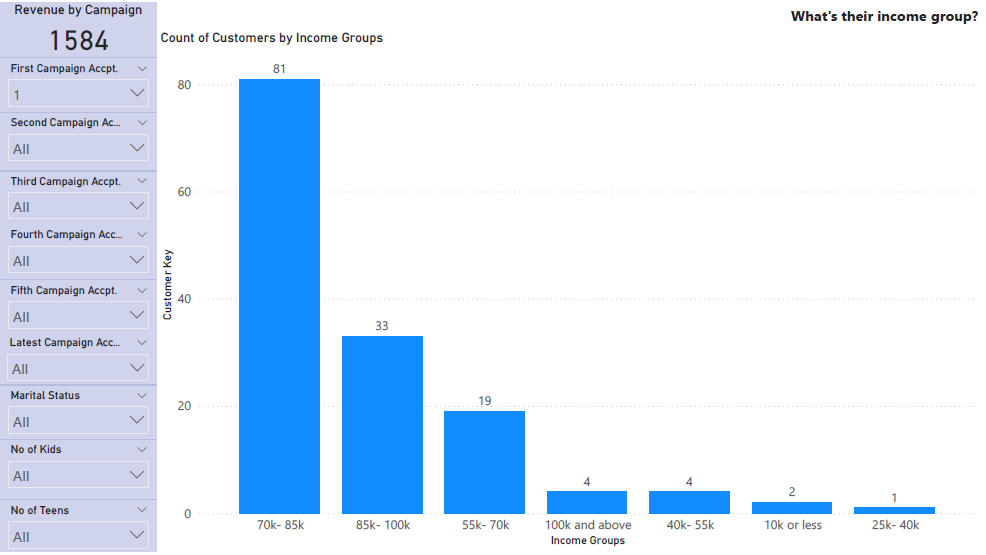
1. What channel do they use to make their purchases? What do they buy?



I found out that the majority of customers who liked our first campaign preferred in-store purchases then catalog purchases and website purchases respectively. This majority also buys a lot of wine and meat.

This could help the marketing team focus on products that drive sales as well as potential products. This could also help in customizing the offers per selling channels and products.

1. What’s their income group?



Here you can notice that the majority of customers who liked our first marketing campaign are in the income group 70k - 85k. This could give us a hint about the social status of those customers *(A class, Middle class, etc…)*

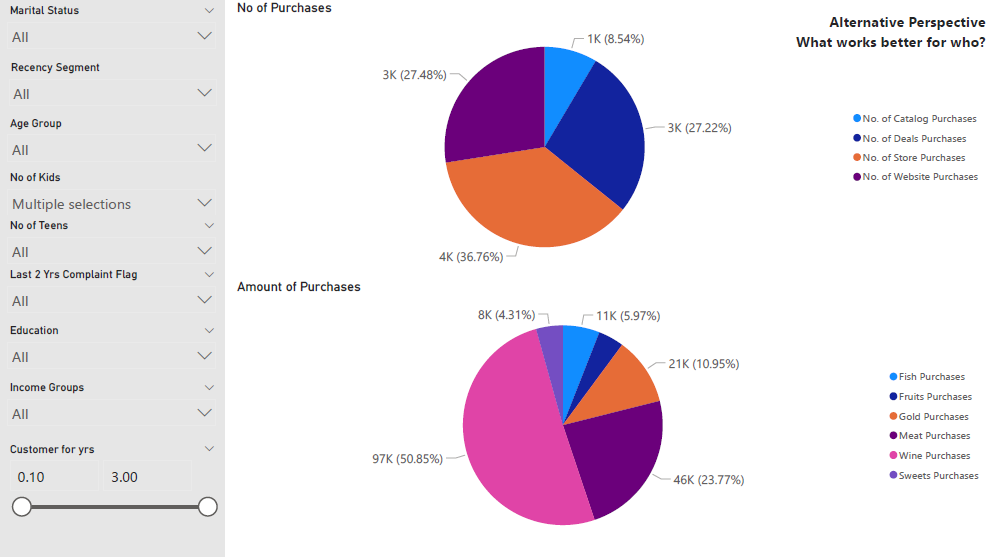
1. The Second Dashboard *(the sixth page in the PDF)*

In this dashboard, I wanted the marketing personnel to customize the clusters they want to monitor *(based on all/some of the dimensions i.e. marital status, age group, etc…)*

Also, I’m focusing on sales channels and how they can act as marketing campaigns too *(like discounts, coupons, website offers and in-store offers).*

I can’t use the “**Revenue by Campaign**” here *(since I’m disregarding campaign acceptance as they are separate campaign),* so I decided to use the amount of purchases as an indicator for revenue.

The main question I’m asking here is “**What works better for who?**”



In the above two charts, I wanted to see the number and amount of purchases made by customers who have kids in the household. Kids like to go the store, so I expected the in-store purchases to have the biggest portion, and it did. People who have kids also look for discounts and doesn’t always have the time to go to the store, that’s why he website and deal purchases came in the second place.

### Challenges and Limitations

1. I think we could have measured the responsiveness of customers to the marketing campaigns better if there were measures before, during and after the campaign *(data across time).*
2. I know that this is dummy data, but some customers didn’t respond to any marketing campaign and still there was value in the “**Revenue**” column. I should have handled this during data preparation.
3. I couldn’t share the live dashboard because I don’t have a subscription to a premium Power BI account, but at least Power BI lets you use the desktop version for free.
4. I couldn’t use Tableau because, unlike Power BI, Tableau gives you a 10-14 days free trial.

# Data Preparation Documentation

I used pandas dataframes on Jupyter notebook for data cleaning and transformation.

By taking a quick look at the data, I needed to do the following:

* Check the datatypes, especially the date column/s.
* Trim customer IDs and Response columns to numbers only.
* Replace\remove NULL values with zeroes.
* Renaming column names for better readability *(I did that using power BI).*
* Adding a derived column that calculates the age of each customer *(assuming we’re in Jan 1st, 2015).*
* Adding a derived column that calculates how long each customer has been our customer in years.
* Adding a derived column based on the “**Recency**” column to calculate whether the customer status is *Active, Semi-active or Dormant*based on this rule:
  + If the customer made a purchase in the last 30 days, then the customer is considered *Active.*
  + If the customer made a purchase in the last 30 to 60 days, then he\she is considered *Semi-active.*
  + If the customer made a purchase more than 60 days ago, then he\she is considered *Dormant.*

Below you would find code snippets for each step and a brief explanation of what happens in this step.

*# first, started with importing pandas and numpy*

import pandas as pd

import numpy as np

*# reading csv as a dataframe*

df\_mrktng = pd.read\_csv('E:\\After ITI\\vois BI Engineer\\marketing\_campaign.csv')

*# checking data types and null values*

df\_mrktng.info()

*# creating another instance of the dataframe*

df\_mrktng1 = df\_mrktng

*# converting the data type of Dt\_Customer column from object to datetime*

df\_mrktng1['Dt\_Customer'] = pd.to\_datetime(df\_mrktng1['Dt\_Customer'])

*# creating another instance of the dataframe as a checkpoint*

df\_mrktng2 = df\_mrktng1

*# creating a function that iterates over every row in a column and returns numbers only*

def keep\_number(text):

clean\_text = ''.join([item for item in text if item.isdigit()])

return clean\_text

*# applying the keep\_number function on the X.ID and Response columns respectively*

df\_mrktng2['X.ID'] = df\_mrktng2['X.ID'].apply(keep\_number)

df\_mrktng2['Response'] = df\_mrktng2['Response'].apply(keep\_number)

*# changing datatypes of Response and X.ID columns to int*

df\_mrktng2['Response'] = df\_mrktng2['Response'].astype('int64')

df\_mrktng2['X.ID'] = df\_mrktng2['X.ID'].astype('int64')

*# creating another instance of the dataframe as a checkpoint*

df\_mrktng3 = df\_mrktng2

*# Replacing NULL values in Income column with zeroes.*

df\_mrktng3['Income'].fillna(0, inplace = True)

*# adding derived age column*

df\_mrktng3['Age'] = 2015 - df\_mrktng['Year\_Birth']

*# creating a function that calculates status of a customer based on Recency*

def recency\_seg (number):

if number <= 30:

seg = 'Active'

elif number > 30 and number <= 60:

seg = 'Semi-Active'

else:

seg = 'Dormant'

return seg

*# adding derived column “Recency Segment” based on “Recency” column.*

df\_mrktng3['Recency Segment'] = df\_mrktng['Recency'].apply(recency\_seg)

*# calculating how long a customer has been our customer in days and years*

df\_mrktng3['Tenurity'] = '2015-01-01'

df\_mrktng3['Tenurity'] = pd.to\_datetime(df\_mrktng3['Tenurity'])

df\_mrktng3['Customer for'] = df\_mrktng3['Tenurity'] - df\_mrktng3['Dt\_Customer']

df\_mrktng3['Customer for yrs'] = (df\_mrktng3['Customer for'] / np.timedelta64(1,'Y')).round(1)

df\_mrktng3.drop('Tenurity', axis = 1, inplace = True)

*# saving dataframe as a csv*

df\_mrktng3.to\_csv(r'E:\\After ITI\\vois BI Engineer\\marketing\_campaign\_cleansed\_II.csv')