Exercise 2.2: Graph Analysis with Matplotlib

#1: Introduction of the Dataset

The Dataset selected for this exercise is "Marketing Campaign" dataset from Kaggle. A response model can provide a significant boost to the efficiency of a marketing campaign by increasing responses or reducing expenses. The objective is to predict who will respond to an offer for a product or service via a detailed analysis of this dataset which might provide good insights for the business to make an informed decision. The dataset has a total of 29 attributes covering customers info, products, promotions and sales. The data set can be found at : https://www.kaggle.com/datasets/rodsaldanha/arketing-campaign

Description of the data : AcceptedCmp1 - 1 if customer accepted the offer in the 1st campaign, 0 otherwise AcceptedCmp2 - 1 if customer accepted the offer in the 2nd campaign, 0 otherwise AcceptedCmp3 - 1 if customer accepted the offer in the 3rd campaign, 0 otherwise AcceptedCmp4 - 1 if customer accepted the offer in the 4th campaign, 0 otherwise AcceptedCmp5 - 1 if customer accepted the offer in the 5th campaign, 0 otherwise Response (target) - 1 if customer accepted the offer in the last campaign, 0 otherwise Complain - 1 if customer complained in the last 2 years DtCustomer - date of customer's enrolment with the company Education - customer's level of education Marital - customer's marital status Kidhome - number of small children in customer's household Teenhome - number of teenagers in customer's household Income - customer's yearly household income MntFishProducts - amount spent on fish products in the last 2 years MntMeatProducts - amount spent on meat products in the last 2 years MntFruits - amount spent on fruits products in the last 2 years MntSweetProducts - amount spent on sweet products in the last 2 years MntWines - amount spent on wine products in the last 2 years MntGoldProds - amount spent on gold products in the last 2 years NumDealsPurchases - number of purchases made with discount NumCatalogPurchases - number of purchases made using catalogue NumStorePurchases - number of purchases made directly in stores NumWebPurchases - number of purchases made through company's web site NumWebVisitsMonth - number of visits to company's web site in the last month Recency - number of days since the last purchase.

#2: Question - Who are the potential customers to respond to an offer ?

Examine and understand the Data

In [1]:
```
# Import required libraries

import warnings
import numpy as np
import pandas as pd
from datetime import date
import plotly as py
import seaborn as sns
import plotly.express as px
```

```
import plotly.graph_objs as go
warnings.filterwarnings("ignore")
pd.set_option('display.max_columns', None)
import matplotlib.lines as lines
import matplotlib.pyplot as plt

#Load the "Marketing Campaign" dataset from the downloaded file (Source: Kaggle)
Capmaign_Data = pd.read_csv("/Users/Jagadeesh/Desktop/SAI_MS/SAI_550/week-2/marketing_c
```

In [2]:

```
pip install plotly==5.10.0
```

```
Requirement already satisfied: plotly==5.10.0 in c:\programdata\anaconda3\lib\site-packa
ges (5.10.0)
Requirement already satisfied: tenacity>=6.2.0 in c:\programdata\anaconda3\lib\site-pack
ages (from plotly==5.10.0) (8.0.1)
Note: you may need to restart the kernel to use updated packages.
WARNING: Ignoring invalid distribution - (c:\programdata\anaconda3\lib\site-packages)
WARNING: Ignoring invalid distribution -equests (c:\programdata\anaconda3\lib\site-packa
ges)
WARNING: Ignoring invalid distribution -arkupsafe (c:\programdata\anaconda3\lib\site-pac
kages)
WARNING: Ignoring invalid distribution - (c:\programdata\anaconda3\lib\site-packages)
WARNING: Ignoring invalid distribution -equests (c:\programdata\anaconda3\lib\site-packa
ges)
WARNING: Ignoring invalid distribution -arkupsafe (c:\programdata\anaconda3\lib\site-pac
kages)
WARNING: Ignoring invalid distribution - (c:\programdata\anaconda3\lib\site-packages)
WARNING: Ignoring invalid distribution -equests (c:\programdata\anaconda3\lib\site-packa
ges)
WARNING: Ignoring invalid distribution -arkupsafe (c:\programdata\anaconda3\lib\site-pac
kages)
WARNING: Ignoring invalid distribution - (c:\programdata\anaconda3\lib\site-packages)
WARNING: Ignoring invalid distribution -equests (c:\programdata\anaconda3\lib\site-packa
ges)
WARNING: Ignoring invalid distribution -arkupsafe (c:\programdata\anaconda3\lib\site-pac
kages)
WARNING: Ignoring invalid distribution - (c:\programdata\anaconda3\lib\site-packages)
WARNING: Ignoring invalid distribution -equests (c:\programdata\anaconda3\lib\site-packa
ges)
WARNING: Ignoring invalid distribution -arkupsafe (c:\programdata\anaconda3\lib\site-pac
kages)
WARNING: Ignoring invalid distribution - (c:\programdata\anaconda3\lib\site-packages)
WARNING: Ignoring invalid distribution -equests (c:\programdata\anaconda3\lib\site-packa
ges)
WARNING: Ignoring invalid distribution -arkupsafe (c:\programdata\anaconda3\lib\site-pac
kages)
```

In [3]:

```
Capmaign_Data.head()
```

Out[3]:

	ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recency	M
0	5524	1957	Graduation	Single	58138.0	0	0	2012-09-04	58	
1	2174	1954	Graduation	Single	46344.0	1	1	2014-03-08	38	
2	4141	1965	Graduation	Together	71613.0	0	0	2013-08-21	26	
3	6182	1984	Graduation	Together	26646.0	1	0	2014-02-10	26	

	ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recency	M
4	5324	1981	PhD	Married	58293.0	1	0	2014-01-19	94	

In [4]:
```
Capmaign_Data.tail()
```

Out[4]:

	ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recen
2235	10870	1967	Graduation	Married	61223.0	0	1	2013-06-13	
2236	4001	1946	PhD	Together	64014.0	2	1	2014-06-10	
2237	7270	1981	Graduation	Divorced	56981.0	0	0	2014-01-25	
2238	8235	1956	Master	Together	69245.0	0	1	2014-01-24	
2239	9405	1954	PhD	Married	52869.0	1	1	2012-10-15	

In [5]:
```
Capmaign_Data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2240 entries, 0 to 2239
Data columns (total 29 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 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
```

```
dtypes: float64(1), int64(25), object(3)
memory usage: 507.6+ KB
```

Find the different Martial status from the data

In [6]:
```
Capmaign_Data.Marital_Status.unique()
```

Out[6]:
```
array(['Single', 'Together', 'Married', 'Divorced', 'Widow', 'Alone',
       'Absurd', 'YOLO'], dtype=object)
```

Simplify further and categorize the data into 2 categories - "In Relation" and "No Relation"

In [7]:
```
mar_status = {
    'Single' : 'No Relation',
    'Together': 'In Relation',
    'Married' : 'In Relation',
    'Divorced' : 'No Relation',
    'Widow' : 'No Relation',
    'Alone' : 'No Relation',
    'Absurd' : 'No Relation',
    'YOLO' : 'No Relation'
}

Capmaign_Data.Marital_Status = Capmaign_Data.Marital_Status.map(mar_status)
```

In [8]:
```
Capmaign_Data.tail()
```

Out[8]:

	ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recen
2235	10870	1967	Graduation	In Relation	61223.0	0	1	2013-06-13	
2236	4001	1946	PhD	In Relation	64014.0	2	1	2014-06-10	
2237	7270	1981	Graduation	No Relation	56981.0	0	0	2014-01-25	
2238	8235	1956	Master	In Relation	69245.0	0	1	2014-01-24	
2239	9405	1954	PhD	In Relation	52869.0	1	1	2012-10-15	

Merge 'Kidhome' and 'Teenhome' into one column 'Children' for simplicity

In [9]:
```
Capmaign_Data['Children'] = Capmaign_Data['Kidhome'] + Capmaign_Data['Teenhome']
```

In [10]:
```
Capmaign_Data.tail()
```

Out[10]:

	ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recen
2235	10870	1967	Graduation	In Relation	61223.0	0	1	2013-06-13	
2236	4001	1946	PhD	In Relation	64014.0	2	1	2014-06-10	
2237	7270	1981	Graduation	No Relation	56981.0	0	0	2014-01-25	
2238	8235	1956	Master	In Relation	69245.0	0	1	2014-01-24	

	ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recen
2239	9405	1954	PhD	In Relation	52869.0	1	1	2012-10-15	

Based on Marital Status, calculate the Family_Size for every row in the dataset.

In [11]:
```
Capmaign_Data['Family_Size'] = 0
for i in range(len(Capmaign_Data)):
    if Capmaign_Data['Marital_Status'][i] == 'In Relation':
        Capmaign_Data['Family_Size'][i] = Capmaign_Data['Children'][i]  + 2 # Add 2 par
    else:
        Capmaign_Data['Family_Size'][i] = Capmaign_Data['Children'][i] + 1 # Add single
```

In [12]:
```
Capmaign_Data.tail()
```

Out[12]:

	ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recen
2235	10870	1967	Graduation	In Relation	61223.0	0	1	2013-06-13	
2236	4001	1946	PhD	In Relation	64014.0	2	1	2014-06-10	
2237	7270	1981	Graduation	No Relation	56981.0	0	0	2014-01-25	
2238	8235	1956	Master	In Relation	69245.0	0	1	2014-01-24	
2239	9405	1954	PhD	In Relation	52869.0	1	1	2012-10-15	

In [13]:
```
# Examine the Data availability (FINAL Date)

pd.to_datetime(Capmaign_Data['Dt_Customer']).unique().max()
```

Out[13]:
```
numpy.datetime64('2014-06-29T00:00:00.000000000')
```

In [14]:
```
pd.to_datetime(Capmaign_Data['Dt_Customer']).unique()
```

Out[14]:
```
array(['2012-09-04T00:00:00.000000000', '2014-03-08T00:00:00.000000000',
       '2013-08-21T00:00:00.000000000', '2014-02-10T00:00:00.000000000',
       '2014-01-19T00:00:00.000000000', '2013-09-09T00:00:00.000000000',
       '2012-11-13T00:00:00.000000000', '2013-05-08T00:00:00.000000000',
       '2013-06-06T00:00:00.000000000', '2014-03-13T00:00:00.000000000',
       '2013-11-15T00:00:00.000000000', '2012-10-10T00:00:00.000000000',
       '2012-11-24T00:00:00.000000000', '2012-12-24T00:00:00.000000000',
       '2012-08-31T00:00:00.000000000', '2013-03-28T00:00:00.000000000',
       '2012-11-03T00:00:00.000000000', '2012-08-08T00:00:00.000000000',
       '2013-01-06T00:00:00.000000000', '2012-12-23T00:00:00.000000000',
       '2014-01-11T00:00:00.000000000', '2013-03-18T00:00:00.000000000',
       '2013-01-02T00:00:00.000000000', '2013-05-27T00:00:00.000000000',
       '2013-02-20T00:00:00.000000000', '2013-05-31T00:00:00.000000000',
       '2013-11-22T00:00:00.000000000', '2014-05-22T00:00:00.000000000',
       '2013-05-11T00:00:00.000000000', '2012-10-29T00:00:00.000000000',
       '2013-08-29T00:00:00.000000000', '2013-12-31T00:00:00.000000000',
       '2013-09-02T00:00:00.000000000', '2014-02-11T00:00:00.000000000',
```

```
'2013-02-01T00:00:00.000000000', '2013-04-29T00:00:00.000000000',
'2013-03-12T00:00:00.000000000', '2013-11-05T00:00:00.000000000',
'2013-10-02T00:00:00.000000000', '2014-06-28T00:00:00.000000000',
'2012-11-09T00:00:00.000000000', '2013-05-24T00:00:00.000000000',
'2014-01-01T00:00:00.000000000', '2012-11-08T00:00:00.000000000',
'2014-05-12T00:00:00.000000000', '2012-08-11T00:00:00.000000000',
'2014-06-07T00:00:00.000000000', '2013-06-12T00:00:00.000000000',
'2012-11-19T00:00:00.000000000', '2013-04-02T00:00:00.000000000',
'2014-04-28T00:00:00.000000000', '2013-06-17T00:00:00.000000000',
'2014-03-03T00:00:00.000000000', '2013-07-04T00:00:00.000000000',
'2012-09-07T00:00:00.000000000', '2013-02-18T00:00:00.000000000',
'2013-06-11T00:00:00.000000000', '2013-12-06T00:00:00.000000000',
'2013-05-21T00:00:00.000000000', '2014-05-11T00:00:00.000000000',
'2014-03-19T00:00:00.000000000', '2013-09-27T00:00:00.000000000',
'2013-04-08T00:00:00.000000000', '2012-09-11T00:00:00.000000000',
'2012-09-14T00:00:00.000000000', '2013-03-17T00:00:00.000000000',
'2013-04-05T00:00:00.000000000', '2014-04-30T00:00:00.000000000',
'2012-12-19T00:00:00.000000000', '2012-08-27T00:00:00.000000000',
'2012-10-12T00:00:00.000000000', '2013-09-04T00:00:00.000000000',
'2012-08-29T00:00:00.000000000', '2013-06-23T00:00:00.000000000',
'2013-07-03T00:00:00.000000000', '2014-02-25T00:00:00.000000000',
'2013-08-11T00:00:00.000000000', '2013-07-16T00:00:00.000000000',
'2014-05-28T00:00:00.000000000', '2014-01-21T00:00:00.000000000',
'2014-05-27T00:00:00.000000000', '2013-11-23T00:00:00.000000000',
'2014-03-23T00:00:00.000000000', '2014-05-24T00:00:00.000000000',
'2012-11-22T00:00:00.000000000', '2013-04-11T00:00:00.000000000',
'2013-12-01T00:00:00.000000000', '2013-06-20T00:00:00.000000000',
'2013-07-23T00:00:00.000000000', '2014-03-30T00:00:00.000000000',
'2013-04-20T00:00:00.000000000', '2013-05-17T00:00:00.000000000',
'2014-05-08T00:00:00.000000000', '2013-12-10T00:00:00.000000000',
'2013-09-24T00:00:00.000000000', '2013-02-02T00:00:00.000000000',
'2012-12-07T00:00:00.000000000', '2013-11-02T00:00:00.000000000',
'2012-11-10T00:00:00.000000000', '2014-06-25T00:00:00.000000000',
'2012-12-12T00:00:00.000000000', '2013-09-10T00:00:00.000000000',
'2014-01-24T00:00:00.000000000', '2013-02-19T00:00:00.000000000',
'2013-07-11T00:00:00.000000000', '2013-11-14T00:00:00.000000000',
'2014-02-24T00:00:00.000000000', '2013-04-22T00:00:00.000000000',
'2013-04-21T00:00:00.000000000', '2014-06-08T00:00:00.000000000',
'2014-05-10T00:00:00.000000000', '2013-09-30T00:00:00.000000000',
'2013-03-10T00:00:00.000000000', '2013-08-14T00:00:00.000000000',
'2013-07-07T00:00:00.000000000', '2014-05-19T00:00:00.000000000',
'2013-07-28T00:00:00.000000000', '2012-10-19T00:00:00.000000000',
'2013-04-19T00:00:00.000000000', '2013-08-28T00:00:00.000000000',
'2013-03-19T00:00:00.000000000', '2012-10-18T00:00:00.000000000',
'2012-10-28T00:00:00.000000000', '2012-08-22T00:00:00.000000000',
'2014-06-21T00:00:00.000000000', '2014-04-24T00:00:00.000000000',
'2014-03-07T00:00:00.000000000', '2012-12-14T00:00:00.000000000',
'2012-10-09T00:00:00.000000000', '2013-07-08T00:00:00.000000000',
'2013-09-12T00:00:00.000000000', '2013-10-17T00:00:00.000000000',
'2013-08-20T00:00:00.000000000', '2013-07-01T00:00:00.000000000',
'2012-11-05T00:00:00.000000000', '2014-01-05T00:00:00.000000000',
'2013-05-01T00:00:00.000000000', '2014-03-01T00:00:00.000000000',
'2013-11-13T00:00:00.000000000', '2013-11-18T00:00:00.000000000',
'2014-04-21T00:00:00.000000000', '2013-07-13T00:00:00.000000000',
'2014-01-30T00:00:00.000000000', '2014-04-04T00:00:00.000000000',
'2012-09-12T00:00:00.000000000', '2012-12-16T00:00:00.000000000',
'2014-05-23T00:00:00.000000000', '2014-06-24T00:00:00.000000000',
'2013-09-28T00:00:00.000000000', '2014-01-22T00:00:00.000000000',
'2014-06-15T00:00:00.000000000', '2012-12-05T00:00:00.000000000',
'2013-08-02T00:00:00.000000000', '2013-02-23T00:00:00.000000000',
```

```
'2012-09-09T00:00:00.000000000', '2013-09-26T00:00:00.000000000',
'2013-05-30T00:00:00.000000000', '2013-12-29T00:00:00.000000000',
'2014-05-13T00:00:00.000000000', '2013-09-19T00:00:00.000000000',
'2013-01-17T00:00:00.000000000', '2014-03-31T00:00:00.000000000',
'2014-06-29T00:00:00.000000000', '2013-12-09T00:00:00.000000000',
'2013-10-14T00:00:00.000000000', '2012-11-02T00:00:00.000000000',
'2013-07-17T00:00:00.000000000', '2013-12-08T00:00:00.000000000',
'2013-05-13T00:00:00.000000000', '2012-09-10T00:00:00.000000000',
'2013-03-31T00:00:00.000000000', '2014-03-18T00:00:00.000000000',
'2012-10-05T00:00:00.000000000', '2013-01-21T00:00:00.000000000',
'2013-05-04T00:00:00.000000000', '2014-04-01T00:00:00.000000000',
'2013-08-31T00:00:00.000000000', '2012-11-14T00:00:00.000000000',
'2014-04-11T00:00:00.000000000', '2014-04-14T00:00:00.000000000',
'2013-01-05T00:00:00.000000000', '2012-09-08T00:00:00.000000000',
'2013-03-26T00:00:00.000000000', '2012-10-25T00:00:00.000000000',
'2012-12-09T00:00:00.000000000', '2014-02-16T00:00:00.000000000',
'2013-03-20T00:00:00.000000000', '2013-05-15T00:00:00.000000000',
'2014-04-16T00:00:00.000000000', '2013-03-23T00:00:00.000000000',
'2013-12-04T00:00:00.000000000', '2013-10-04T00:00:00.000000000',
'2013-12-26T00:00:00.000000000', '2014-05-17T00:00:00.000000000',
'2013-10-28T00:00:00.000000000', '2013-07-31T00:00:00.000000000',
'2013-11-28T00:00:00.000000000', '2012-08-17T00:00:00.000000000',
'2014-06-05T00:00:00.000000000', '2013-05-20T00:00:00.000000000',
'2013-01-09T00:00:00.000000000', '2013-09-16T00:00:00.000000000',
'2013-10-27T00:00:00.000000000', '2012-09-22T00:00:00.000000000',
'2012-10-13T00:00:00.000000000', '2012-10-16T00:00:00.000000000',
'2012-10-22T00:00:00.000000000', '2013-06-04T00:00:00.000000000',
'2013-05-22T00:00:00.000000000', '2014-06-17T00:00:00.000000000',
'2012-11-23T00:00:00.000000000', '2013-02-03T00:00:00.000000000',
'2013-03-14T00:00:00.000000000', '2014-06-26T00:00:00.000000000',
'2012-10-15T00:00:00.000000000', '2013-06-19T00:00:00.000000000',
'2014-03-20T00:00:00.000000000', '2014-02-04T00:00:00.000000000',
'2014-04-06T00:00:00.000000000', '2013-02-06T00:00:00.000000000',
'2013-02-11T00:00:00.000000000', '2014-04-17T00:00:00.000000000',
'2013-07-12T00:00:00.000000000', '2013-01-29T00:00:00.000000000',
'2013-01-08T00:00:00.000000000', '2013-06-13T00:00:00.000000000',
'2013-12-27T00:00:00.000000000', '2012-10-31T00:00:00.000000000',
'2014-01-15T00:00:00.000000000', '2012-08-23T00:00:00.000000000',
'2013-03-07T00:00:00.000000000', '2013-01-13T00:00:00.000000000',
'2013-01-12T00:00:00.000000000', '2014-03-17T00:00:00.000000000',
'2013-10-12T00:00:00.000000000', '2014-04-13T00:00:00.000000000',
'2012-09-18T00:00:00.000000000', '2014-03-05T00:00:00.000000000',
'2013-04-27T00:00:00.000000000', '2014-01-18T00:00:00.000000000',
'2013-06-03T00:00:00.000000000', '2013-12-17T00:00:00.000000000',
'2014-03-11T00:00:00.000000000', '2013-07-29T00:00:00.000000000',
'2012-08-14T00:00:00.000000000', '2013-08-23T00:00:00.000000000',
'2014-02-09T00:00:00.000000000', '2013-02-07T00:00:00.000000000',
'2013-01-11T00:00:00.000000000', '2013-07-05T00:00:00.000000000',
'2013-07-02T00:00:00.000000000', '2013-11-07T00:00:00.000000000',
'2013-05-09T00:00:00.000000000', '2013-02-13T00:00:00.000000000',
'2013-04-16T00:00:00.000000000', '2013-09-11T00:00:00.000000000',
'2013-04-03T00:00:00.000000000', '2013-01-10T00:00:00.000000000',
'2013-06-30T00:00:00.000000000', '2012-12-06T00:00:00.000000000',
'2012-11-12T00:00:00.000000000', '2013-03-03T00:00:00.000000000',
'2012-08-10T00:00:00.000000000', '2013-12-07T00:00:00.000000000',
'2013-08-15T00:00:00.000000000', '2013-11-10T00:00:00.000000000',
'2014-06-16T00:00:00.000000000', '2012-12-25T00:00:00.000000000',
'2014-01-03T00:00:00.000000000', '2012-10-27T00:00:00.000000000',
'2012-12-22T00:00:00.000000000', '2013-11-29T00:00:00.000000000',
'2013-10-08T00:00:00.000000000', '2012-09-28T00:00:00.000000000',
```

```
'2014-03-22T00:00:00.000000000', '2012-12-28T00:00:00.000000000',
'2012-08-21T00:00:00.000000000', '2013-03-16T00:00:00.000000000',
'2012-11-17T00:00:00.000000000', '2012-12-01T00:00:00.000000000',
'2014-04-22T00:00:00.000000000', '2012-11-11T00:00:00.000000000',
'2013-06-22T00:00:00.000000000', '2012-08-18T00:00:00.000000000',
'2012-12-30T00:00:00.000000000', '2013-06-14T00:00:00.000000000',
'2013-10-16T00:00:00.000000000', '2012-08-30T00:00:00.000000000',
'2014-05-04T00:00:00.000000000', '2013-04-18T00:00:00.000000000',
'2013-10-06T00:00:00.000000000', '2012-09-15T00:00:00.000000000',
'2012-09-27T00:00:00.000000000', '2013-03-11T00:00:00.000000000',
'2013-10-22T00:00:00.000000000', '2014-06-09T00:00:00.000000000',
'2014-05-30T00:00:00.000000000', '2012-10-17T00:00:00.000000000',
'2013-03-30T00:00:00.000000000', '2013-01-23T00:00:00.000000000',
'2013-11-20T00:00:00.000000000', '2014-02-14T00:00:00.000000000',
'2013-02-22T00:00:00.000000000', '2013-03-05T00:00:00.000000000',
'2014-05-06T00:00:00.000000000', '2013-04-13T00:00:00.000000000',
'2014-04-05T00:00:00.000000000', '2013-04-25T00:00:00.000000000',
'2013-11-25T00:00:00.000000000', '2014-02-02T00:00:00.000000000',
'2013-06-21T00:00:00.000000000', '2013-12-21T00:00:00.000000000',
'2013-09-07T00:00:00.000000000', '2013-08-22T00:00:00.000000000',
'2013-12-20T00:00:00.000000000', '2013-08-06T00:00:00.000000000',
'2013-10-09T00:00:00.000000000', '2013-06-09T00:00:00.000000000',
'2013-12-22T00:00:00.000000000', '2013-05-02T00:00:00.000000000',
'2013-02-16T00:00:00.000000000', '2012-08-20T00:00:00.000000000',
'2013-04-01T00:00:00.000000000', '2014-05-25T00:00:00.000000000',
'2012-09-25T00:00:00.000000000', '2013-09-22T00:00:00.000000000',
'2014-01-28T00:00:00.000000000', '2013-04-23T00:00:00.000000000',
'2014-03-06T00:00:00.000000000', '2012-08-15T00:00:00.000000000',
'2013-03-27T00:00:00.000000000', '2013-10-10T00:00:00.000000000',
'2012-08-04T00:00:00.000000000', '2014-05-18T00:00:00.000000000',
'2014-05-09T00:00:00.000000000', '2012-10-23T00:00:00.000000000',
'2013-12-30T00:00:00.000000000', '2012-08-19T00:00:00.000000000',
'2013-10-30T00:00:00.000000000', '2013-06-26T00:00:00.000000000',
'2014-01-26T00:00:00.000000000', '2012-09-24T00:00:00.000000000',
'2012-12-03T00:00:00.000000000', '2013-11-27T00:00:00.000000000',
'2012-08-06T00:00:00.000000000', '2012-12-29T00:00:00.000000000',
'2012-08-24T00:00:00.000000000', '2012-11-07T00:00:00.000000000',
'2013-12-02T00:00:00.000000000', '2013-07-19T00:00:00.000000000',
'2013-05-25T00:00:00.000000000', '2013-10-20T00:00:00.000000000',
'2013-04-14T00:00:00.000000000', '2014-01-27T00:00:00.000000000',
'2013-08-16T00:00:00.000000000', '2012-09-30T00:00:00.000000000',
'2012-12-04T00:00:00.000000000', '2012-08-12T00:00:00.000000000',
'2013-11-03T00:00:00.000000000', '2013-02-15T00:00:00.000000000',
'2013-09-18T00:00:00.000000000', '2014-06-18T00:00:00.000000000',
'2013-04-24T00:00:00.000000000', '2014-02-15T00:00:00.000000000',
'2013-08-08T00:00:00.000000000', '2013-12-23T00:00:00.000000000',
'2012-11-28T00:00:00.000000000', '2013-08-07T00:00:00.000000000',
'2014-01-07T00:00:00.000000000', '2013-12-03T00:00:00.000000000',
'2014-03-29T00:00:00.000000000', '2014-04-08T00:00:00.000000000',
'2014-02-07T00:00:00.000000000', '2013-11-06T00:00:00.000000000',
'2013-03-01T00:00:00.000000000', '2013-08-03T00:00:00.000000000',
'2013-06-16T00:00:00.000000000', '2013-06-25T00:00:00.000000000',
'2013-09-13T00:00:00.000000000', '2013-08-01T00:00:00.000000000',
'2013-01-31T00:00:00.000000000', '2013-07-09T00:00:00.000000000',
'2013-04-15T00:00:00.000000000', '2014-04-03T00:00:00.000000000',
'2014-04-18T00:00:00.000000000', '2012-09-23T00:00:00.000000000',
'2013-09-14T00:00:00.000000000', '2012-10-26T00:00:00.000000000',
'2013-10-19T00:00:00.000000000', '2013-02-21T00:00:00.000000000',
'2013-09-17T00:00:00.000000000', '2013-01-26T00:00:00.000000000',
'2014-05-07T00:00:00.000000000', '2013-08-19T00:00:00.000000000',
```

```
'2013-07-18T00:00:00.000000000', '2012-12-11T00:00:00.000000000',
'2013-07-26T00:00:00.000000000', '2013-05-06T00:00:00.000000000',
'2013-07-14T00:00:00.000000000', '2014-03-10T00:00:00.000000000',
'2013-12-28T00:00:00.000000000', '2013-01-20T00:00:00.000000000',
'2012-09-21T00:00:00.000000000', '2012-09-06T00:00:00.000000000',
'2012-12-18T00:00:00.000000000', '2013-11-19T00:00:00.000000000',
'2014-06-13T00:00:00.000000000', '2013-05-07T00:00:00.000000000',
'2013-08-18T00:00:00.000000000', '2013-12-13T00:00:00.000000000',
'2014-02-17T00:00:00.000000000', '2012-08-03T00:00:00.000000000',
'2014-01-04T00:00:00.000000000', '2013-10-21T00:00:00.000000000',
'2013-12-25T00:00:00.000000000', '2013-01-01T00:00:00.000000000',
'2013-05-16T00:00:00.000000000', '2013-09-25T00:00:00.000000000',
'2013-11-08T00:00:00.000000000', '2013-12-16T00:00:00.000000000',
'2012-10-30T00:00:00.000000000', '2014-01-02T00:00:00.000000000',
'2013-10-18T00:00:00.000000000', '2013-07-25T00:00:00.000000000',
'2013-10-24T00:00:00.000000000', '2013-12-19T00:00:00.000000000',
'2014-05-16T00:00:00.000000000', '2012-10-04T00:00:00.000000000',
'2013-08-25T00:00:00.000000000', '2013-05-29T00:00:00.000000000',
'2014-03-26T00:00:00.000000000', '2014-04-09T00:00:00.000000000',
'2014-03-12T00:00:00.000000000', '2014-04-29T00:00:00.000000000',
'2014-06-27T00:00:00.000000000', '2013-04-10T00:00:00.000000000',
'2013-11-21T00:00:00.000000000', '2012-10-21T00:00:00.000000000',
'2013-05-14T00:00:00.000000000', '2012-08-01T00:00:00.000000000',
'2014-01-25T00:00:00.000000000', '2013-10-11T00:00:00.000000000',
'2012-08-07T00:00:00.000000000', '2013-03-29T00:00:00.000000000',
'2014-06-03T00:00:00.000000000', '2012-08-09T00:00:00.000000000',
'2013-03-06T00:00:00.000000000', '2013-05-12T00:00:00.000000000',
'2014-03-09T00:00:00.000000000', '2013-09-15T00:00:00.000000000',
'2014-01-29T00:00:00.000000000', '2013-10-03T00:00:00.000000000',
'2013-02-28T00:00:00.000000000', '2014-05-02T00:00:00.000000000',
'2013-02-09T00:00:00.000000000', '2013-03-15T00:00:00.000000000',
'2013-05-10T00:00:00.000000000', '2012-09-05T00:00:00.000000000',
'2013-09-01T00:00:00.000000000', '2013-05-26T00:00:00.000000000',
'2013-12-12T00:00:00.000000000', '2012-12-08T00:00:00.000000000',
'2013-08-30T00:00:00.000000000', '2013-12-18T00:00:00.000000000',
'2013-06-18T00:00:00.000000000', '2014-05-29T00:00:00.000000000',
'2013-01-15T00:00:00.000000000', '2013-10-29T00:00:00.000000000',
'2013-03-02T00:00:00.000000000', '2013-12-14T00:00:00.000000000',
'2013-01-19T00:00:00.000000000', '2013-04-06T00:00:00.000000000',
'2013-11-04T00:00:00.000000000', '2013-03-22T00:00:00.000000000',
'2014-05-31T00:00:00.000000000', '2013-02-04T00:00:00.000000000',
'2013-05-05T00:00:00.000000000', '2013-02-14T00:00:00.000000000',
'2014-03-21T00:00:00.000000000', '2013-08-05T00:00:00.000000000',
'2012-11-29T00:00:00.000000000', '2013-07-27T00:00:00.000000000',
'2014-01-13T00:00:00.000000000', '2012-09-29T00:00:00.000000000',
'2013-10-23T00:00:00.000000000', '2012-11-06T00:00:00.000000000',
'2014-05-26T00:00:00.000000000', '2014-03-04T00:00:00.000000000',
'2014-01-06T00:00:00.000000000', '2014-06-11T00:00:00.000000000',
'2014-06-19T00:00:00.000000000', '2012-11-26T00:00:00.000000000',
'2013-12-11T00:00:00.000000000', '2014-01-14T00:00:00.000000000',
'2014-01-23T00:00:00.000000000', '2013-03-24T00:00:00.000000000',
'2013-06-29T00:00:00.000000000', '2013-01-25T00:00:00.000000000',
'2013-11-17T00:00:00.000000000', '2013-05-18T00:00:00.000000000',
'2013-04-17T00:00:00.000000000', '2013-01-24T00:00:00.000000000',
'2013-08-09T00:00:00.000000000', '2013-06-10T00:00:00.000000000',
'2013-07-20T00:00:00.000000000', '2013-05-28T00:00:00.000000000',
'2013-03-04T00:00:00.000000000', '2012-10-01T00:00:00.000000000',
'2013-02-10T00:00:00.000000000', '2012-10-14T00:00:00.000000000',
'2014-02-03T00:00:00.000000000', '2014-06-22T00:00:00.000000000',
'2013-10-25T00:00:00.000000000', '2014-02-18T00:00:00.000000000',
```

```
'2013-11-09T00:00:00.000000000', '2013-08-13T00:00:00.000000000',
'2012-08-25T00:00:00.000000000', '2014-04-19T00:00:00.000000000',
'2014-04-07T00:00:00.000000000', '2013-01-03T00:00:00.000000000',
'2012-11-04T00:00:00.000000000', '2013-01-22T00:00:00.000000000',
'2014-02-26T00:00:00.000000000', '2014-04-23T00:00:00.000000000',
'2013-07-30T00:00:00.000000000', '2013-10-13T00:00:00.000000000',
'2012-11-27T00:00:00.000000000', '2014-05-01T00:00:00.000000000',
'2012-11-30T00:00:00.000000000', '2013-08-17T00:00:00.000000000',
'2013-01-04T00:00:00.000000000', '2014-05-03T00:00:00.000000000',
'2014-04-26T00:00:00.000000000', '2014-06-02T00:00:00.000000000',
'2013-10-26T00:00:00.000000000', '2014-03-24T00:00:00.000000000',
'2013-02-25T00:00:00.000000000', '2013-09-20T00:00:00.000000000',
'2013-01-16T00:00:00.000000000', '2013-12-24T00:00:00.000000000',
'2012-11-18T00:00:00.000000000', '2012-08-05T00:00:00.000000000',
'2013-05-03T00:00:00.000000000', '2013-11-12T00:00:00.000000000',
'2013-07-06T00:00:00.000000000', '2013-07-21T00:00:00.000000000',
'2012-10-20T00:00:00.000000000', '2013-05-23T00:00:00.000000000',
'2012-08-26T00:00:00.000000000', '2014-01-12T00:00:00.000000000',
'2014-02-13T00:00:00.000000000', '2014-02-27T00:00:00.000000000',
'2014-04-20T00:00:00.000000000', '2013-06-28T00:00:00.000000000',
'2013-04-28T00:00:00.000000000', '2013-03-21T00:00:00.000000000',
'2013-08-27T00:00:00.000000000', '2012-07-30T00:00:00.000000000',
'2013-12-05T00:00:00.000000000', '2013-02-08T00:00:00.000000000',
'2012-11-21T00:00:00.000000000', '2013-02-26T00:00:00.000000000',
'2012-09-20T00:00:00.000000000', '2013-06-08T00:00:00.000000000',
'2013-09-05T00:00:00.000000000', '2014-04-12T00:00:00.000000000',
'2013-06-24T00:00:00.000000000', '2014-02-19T00:00:00.000000000',
'2012-12-27T00:00:00.000000000', '2012-11-01T00:00:00.000000000',
'2013-07-22T00:00:00.000000000', '2014-01-31T00:00:00.000000000',
'2013-08-26T00:00:00.000000000', '2013-10-01T00:00:00.000000000',
'2014-06-10T00:00:00.000000000', '2014-02-01T00:00:00.000000000',
'2012-09-03T00:00:00.000000000', '2012-09-26T00:00:00.000000000',
'2012-07-31T00:00:00.000000000', '2012-12-02T00:00:00.000000000',
'2014-05-05T00:00:00.000000000', '2012-10-06T00:00:00.000000000',
'2014-03-02T00:00:00.000000000', '2014-05-15T00:00:00.000000000',
'2012-12-10T00:00:00.000000000', '2012-08-13T00:00:00.000000000',
'2014-02-23T00:00:00.000000000', '2014-03-16T00:00:00.000000000',
'2013-09-21T00:00:00.000000000', '2012-10-24T00:00:00.000000000',
'2013-10-15T00:00:00.000000000', '2013-02-17T00:00:00.000000000',
'2012-08-28T00:00:00.000000000', '2014-02-12T00:00:00.000000000',
'2014-02-08T00:00:00.000000000', '2012-09-01T00:00:00.000000000',
'2012-11-16T00:00:00.000000000', '2013-02-12T00:00:00.000000000',
'2012-08-02T00:00:00.000000000', '2012-08-16T00:00:00.000000000',
'2012-11-25T00:00:00.000000000', '2014-04-27T00:00:00.000000000',
'2013-04-30T00:00:00.000000000', '2012-11-20T00:00:00.000000000',
'2014-06-14T00:00:00.000000000', '2014-03-28T00:00:00.000000000',
'2013-03-08T00:00:00.000000000', '2012-09-02T00:00:00.000000000',
'2014-04-15T00:00:00.000000000', '2014-04-25T00:00:00.000000000',
'2013-06-27T00:00:00.000000000', '2013-06-01T00:00:00.000000000',
'2013-10-05T00:00:00.000000000', '2014-02-22T00:00:00.000000000',
'2013-04-12T00:00:00.000000000', '2013-09-23T00:00:00.000000000',
'2013-07-24T00:00:00.000000000', '2014-06-23T00:00:00.000000000',
'2014-01-10T00:00:00.000000000', '2013-03-25T00:00:00.000000000',
'2014-01-17T00:00:00.000000000', '2014-06-04T00:00:00.000000000',
'2014-03-25T00:00:00.000000000', '2012-12-31T00:00:00.000000000',
'2012-09-19T00:00:00.000000000', '2013-02-27T00:00:00.000000000',
'2012-10-02T00:00:00.000000000', '2013-04-09T00:00:00.000000000',
'2014-04-02T00:00:00.000000000', '2013-01-07T00:00:00.000000000',
'2013-11-26T00:00:00.000000000', '2013-08-04T00:00:00.000000000',
'2012-12-17T00:00:00.000000000', '2014-06-06T00:00:00.000000000',
```

```
       '2013-05-19T00:00:00.000000000', '2013-08-12T00:00:00.000000000',
       '2013-12-15T00:00:00.000000000', '2012-10-07T00:00:00.000000000',
       '2013-09-06T00:00:00.000000000', '2013-07-15T00:00:00.000000000',
       '2013-01-27T00:00:00.000000000', '2014-02-05T00:00:00.000000000',
       '2012-12-26T00:00:00.000000000', '2014-02-06T00:00:00.000000000',
       '2013-01-14T00:00:00.000000000', '2014-06-20T00:00:00.000000000',
       '2013-08-24T00:00:00.000000000', '2014-02-28T00:00:00.000000000',
       '2013-04-07T00:00:00.000000000', '2014-04-10T00:00:00.000000000',
       '2014-06-12T00:00:00.000000000', '2013-11-30T00:00:00.000000000',
       '2013-03-09T00:00:00.000000000', '2014-03-27T00:00:00.000000000',
       '2012-12-15T00:00:00.000000000', '2012-09-17T00:00:00.000000000',
       '2013-06-02T00:00:00.000000000', '2012-12-21T00:00:00.000000000',
       '2013-11-01T00:00:00.000000000', '2013-08-10T00:00:00.000000000',
       '2012-10-11T00:00:00.000000000', '2012-12-20T00:00:00.000000000',
       '2014-01-09T00:00:00.000000000'], dtype='datetime64[ns]')
```

In [15]:

```
# Calculate Customer's Age based on end date of dataset from above and Calculate Total

Capmaign_Data['Age'] = 2014 - Capmaign_Data['Year_Birth']

Capmaign_Data['Total_purchases_sum'] = Capmaign_Data['MntWines'] + Capmaign_Data['MntFr
+ Capmaign_Data['MntMeatProducts'] + Capmaign_Data['MntFishProducts']
+ Capmaign_Data['MntSweetProducts']+ Capmaign_Data['MntGoldProds']

Capmaign_Data['Total_purchases_amount'] = Capmaign_Data['NumCatalogPurchases'] + Capmai

Capmaign_Data['TotalCmp'] = Capmaign_Data['AcceptedCmp1'] + Capmaign_Data['AcceptedCmp2
Capmaign_Data['AcceptedCmp3'] + Capmaign_Data['AcceptedCmp4'] +Capmaign_Data['AcceptedC
```

Out[15]:

```
0       1
1       0
2       0
3       0
4       0
       ..
2235    0
2236    0
2237    1
2238    0
2239    1
Length: 2240, dtype: int64
```

In [16]:

```
Capmaign_Data.tail()
```

Out[16]:

	ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recen
2235	10870	1967	Graduation	In Relation	61223.0	0	1	2013-06-13	
2236	4001	1946	PhD	In Relation	64014.0	2	1	2014-06-10	
2237	7270	1981	Graduation	No Relation	56981.0	0	0	2014-01-25	
2238	8235	1956	Master	In Relation	69245.0	0	1	2014-01-24	
2239	9405	1954	PhD	In Relation	52869.0	1	1	2012-10-15	

In [17]:

```
# Drop redundant columns or the columns that are not required like ID, MntWines, Kidhom
```

```
Capmaign_Data.drop(columns=['ID', 'MntWines', 'MntFruits',
                            'MntMeatProducts', 'MntFishProducts',
                            'MntSweetProducts', 'MntGoldProds',
                            'NumCatalogPurchases', 'NumStorePurchases', 'NumWebPurchases',
                            'Kidhome', 'Teenhome', 'Marital_Status',
                            'AcceptedCmp3', 'AcceptedCmp4', 'AcceptedCmp5',
                            'AcceptedCmp1','AcceptedCmp2'], inplace=True)
```

In [18]:

```
# see the unnecessary columns are removed from the Data Frame
Capmaign_Data.tail()
```

Out[18]:

	Year_Birth	Education	Income	Dt_Customer	Recency	NumDealsPurchases	NumWebVisitsMonth
2235	1967	Graduation	61223.0	2013-06-13	46	2	5
2236	1946	PhD	64014.0	2014-06-10	56	7	7
2237	1981	Graduation	56981.0	2014-01-25	91	1	6
2238	1956	Master	69245.0	2014-01-24	8	2	3
2239	1954	PhD	52869.0	2012-10-15	40	3	7

In [19]:

```
# Examine missing/null values or in other words, or find outliers to clean the data

Capmaign_Data.isnull().sum()
```

Out[19]:

```
Year_Birth              0
Education               0
Income                 24
Dt_Customer             0
Recency                 0
NumDealsPurchases       0
NumWebVisitsMonth       0
Complain                0
Z_CostContact           0
Z_Revenue               0
Response                0
Children                0
Family_Size             0
Age                     0
Total_purchases_sum     0
Total_purchases_amount  0
TotalCmp                0
dtype: int64
```

In [20]:

```
# From the above result that Income has 24 missing values
# Replace missing values for income with mean

Capmaign_Data['Income'].fillna(Capmaign_Data['Income'].mean(), inplace=True)
```

In [21]:

```
# for the Income missing values are replaced with Mean Income from the campaign data.
Capmaign_Data.isnull().sum()
```

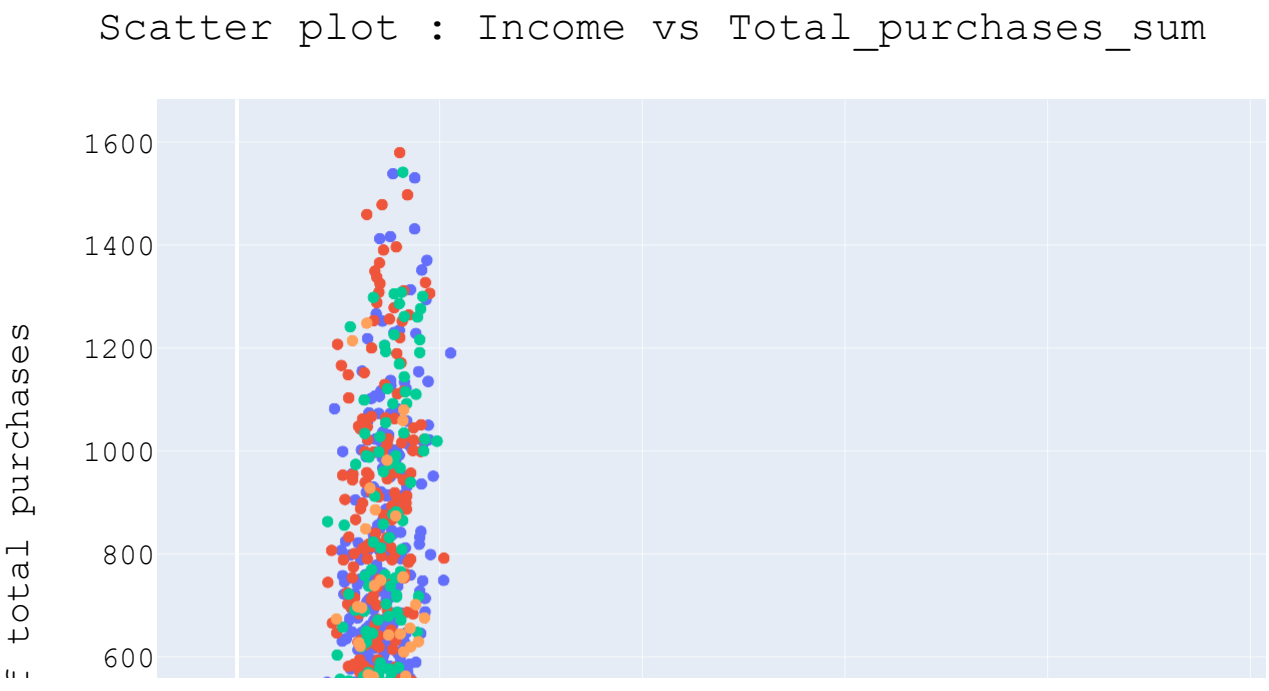
Out[21]:
```
Year_Birth              0
Education               0
Income                  0
Dt_Customer             0
Recency                 0
NumDealsPurchases       0
NumWebVisitsMonth       0
Complain                0
Z_CostContact           0
Z_Revenue               0
Response                0
Children                0
Family_Size             0
Age                     0
Total_purchases_sum     0
Total_purchases_amount  0
TotalCmp                0
dtype: int64
```

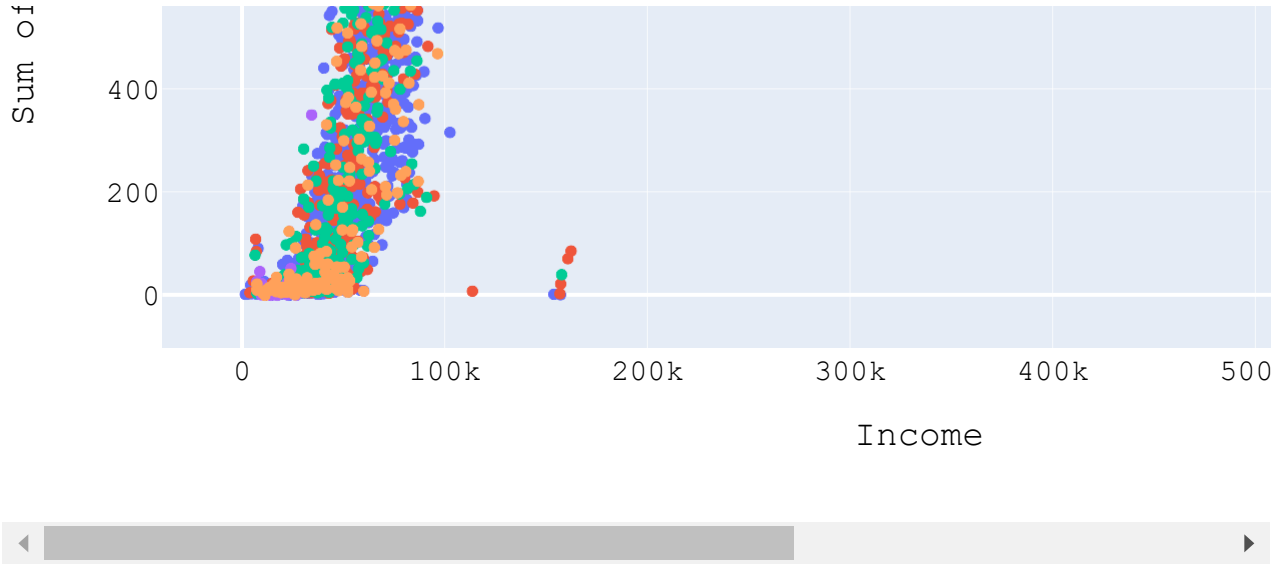
Question#3: Graph1 - Plot a Scatter plot between 'Income' and 'Total_purchases_sum' (Amount spent).

In [22]:
```
#scattered plot
#X axis : Income, Y axis : Total_purchases_sum
fig = px.scatter(Capmaign_Data, x='Income', y='Total_purchases_sum',
                 color = 'Education')

fig.update_layout(height=600, width=1000, title_text='Scatter plot : Income vs Total_pu
                  yaxis_title="Sum of total purchases",font=dict(
        family="Courier New, monospace",
        size=15,
        color="black"
    ))

fig.show()
```





#: Graph1 Observations:

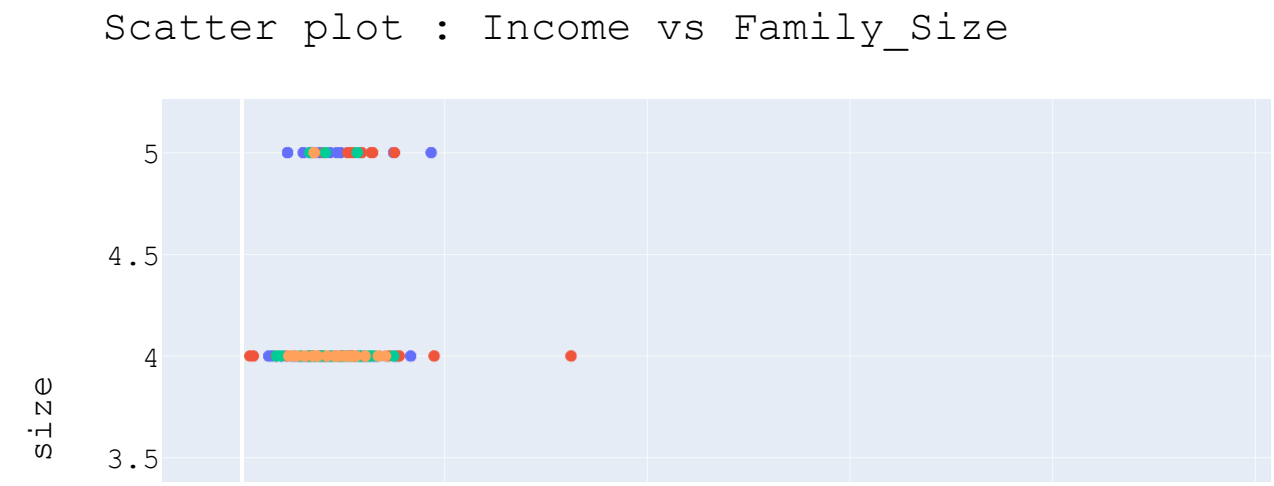
Above graph suggests that there's a clear pattern between income and total purchases which are directly propotional (i.e., As income increases, purchases increase and vice-versa). However, same is not evident with respect to Eduation and Income.

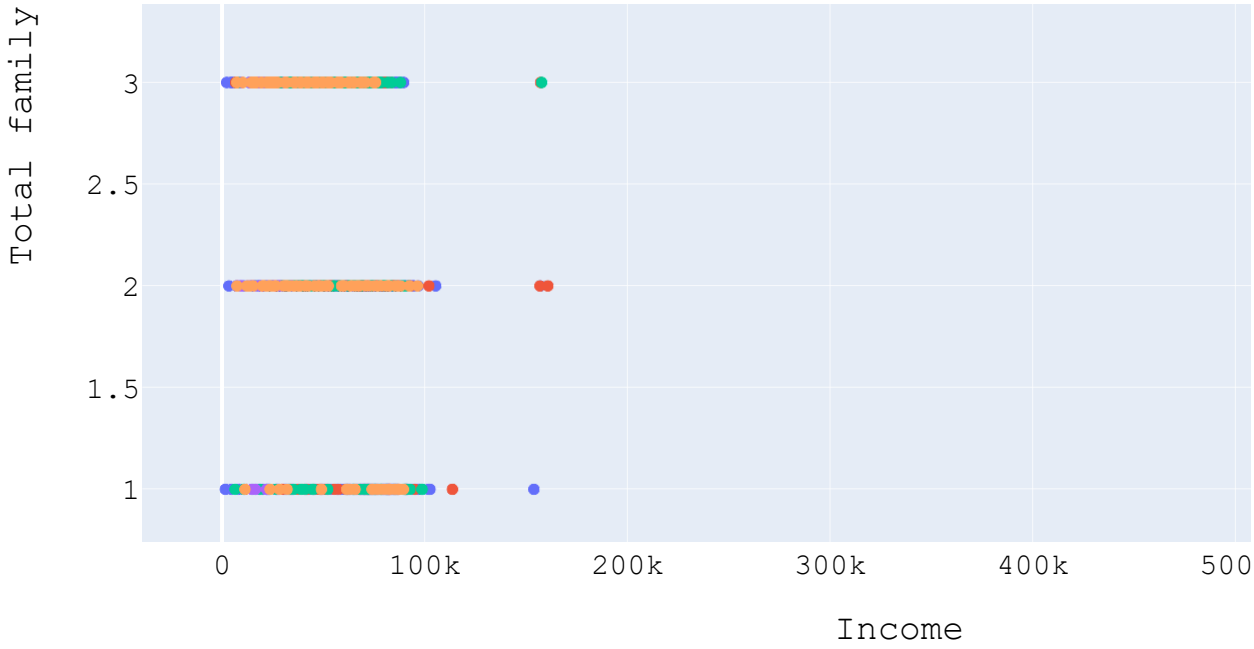
Question#3: Graph2 - Plot a Scatter plot between 'Income' and 'Family_Size' (Amount spent).

In [23]:
```
#scattered plot
#X axis : Income, Y axis : Total_purchases_sum
fig = px.scatter(Capmaign_Data, x='Income', y='Family_Size',
                 color = 'Education')

fig.update_layout(height=600, width=1000, title_text='Scatter plot : Income vs Family_S
                  yaxis_title="Total family size",font=dict(
        family="Courier New, monospace",
        size=15,
        color="black"
    ))

fig.show()
```





#: Graph2 Observations:

Above graph suggests that there's a clear pattern between income and Family size which are directly propotional (i.e., As family size increases, income increases and vice-versa). However, same is not evident with respect to Eduation.

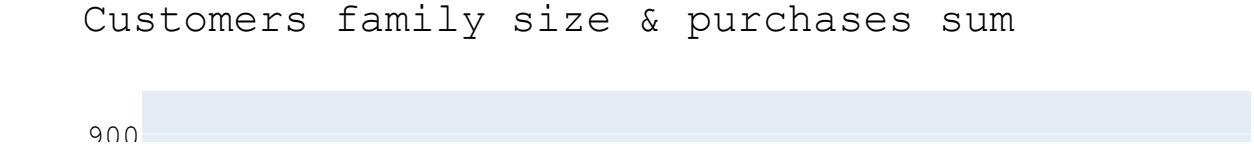
#3: Graph3 - Plot a Bar Graph using Family_Size, Education and Total Purchases

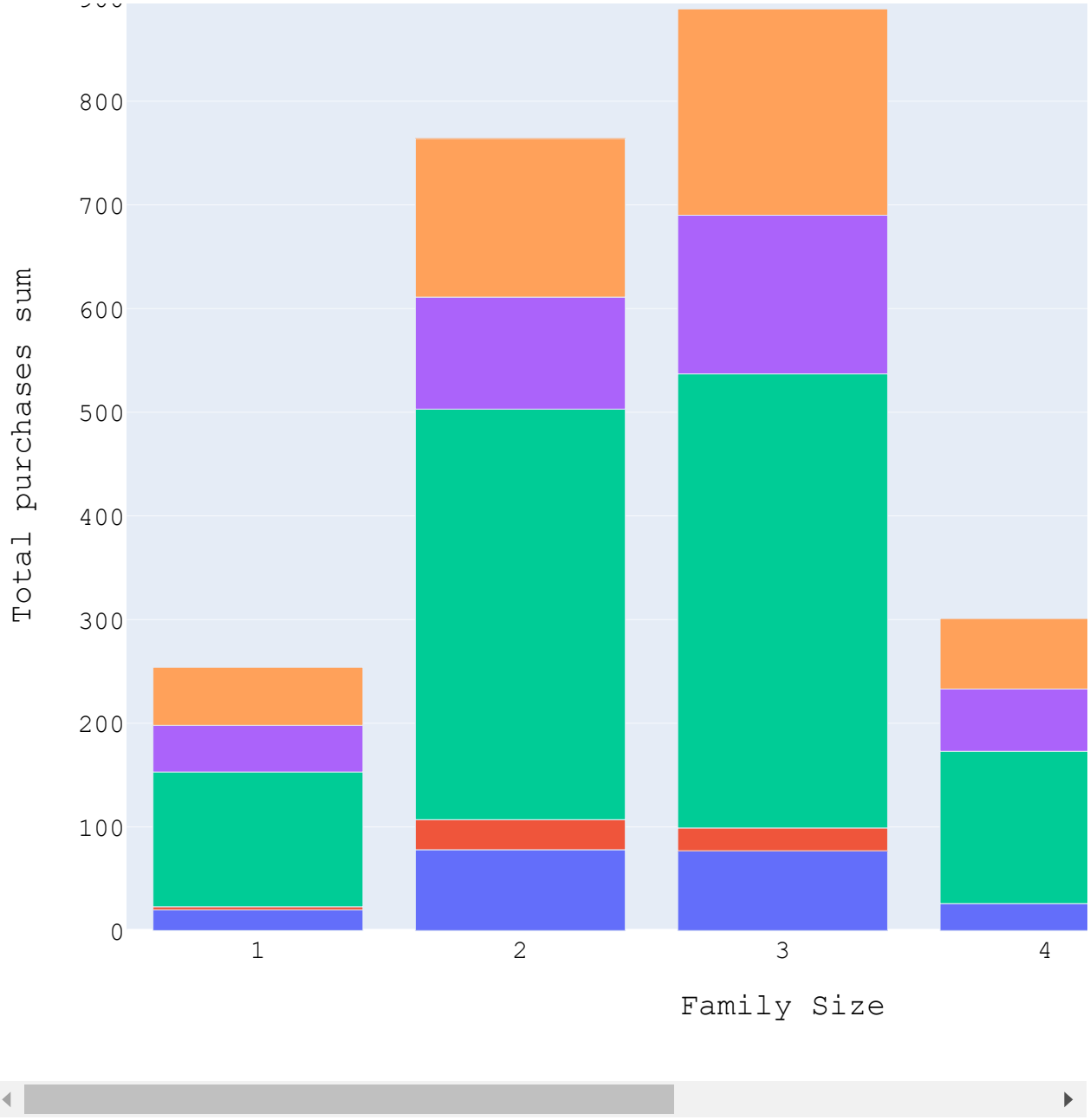
In [24]:
```
total_family_size = Capmaign_Data.groupby(['Family_Size', 'Education']).count()['Total_

fig = px.bar(x = total_family_size.index.get_level_values(0) ,y = total_family_size.val
             color = total_family_size.index.get_level_values(1),width = 1000, height =
             labels=dict(x="Family Size", y="Sum of total purchases", color="Education"
)

fig.update_layout(title_text='Customers family size & purchases sum',
                  xaxis_title='Family Size',
                  yaxis_title='Total purchases sum',
                  legend_title = 'Education',
                 font=dict(
        family="Courier New, monospace",
        size=15,
        color="black")
)

fig.show()
```





Graph3 Observations:

Above graph suggests that the potential customers are most likely the couples with either no kids(count 2) or a single kid(count 3) at max and most of them usually have a graduation degree.

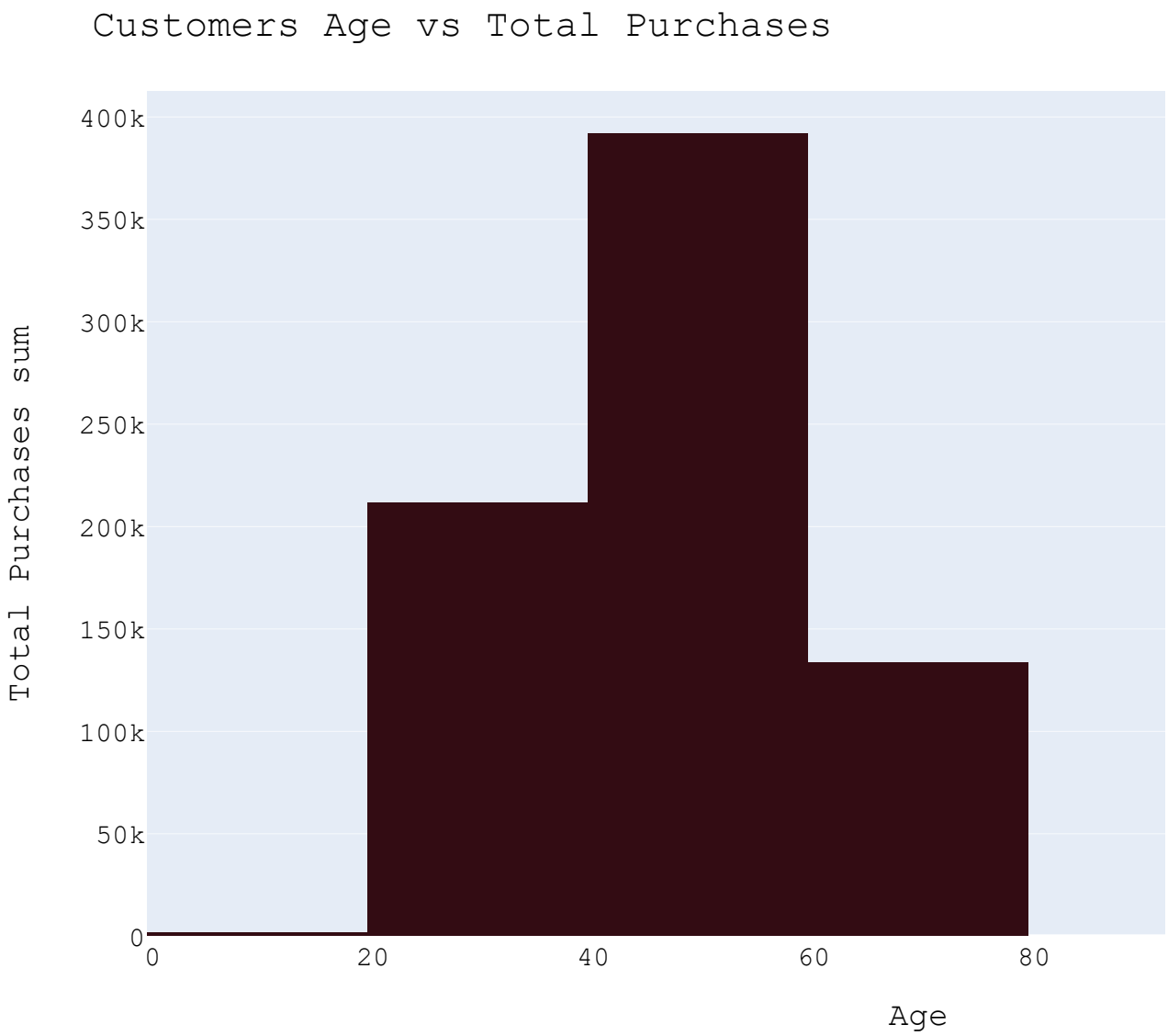
#4: Graph4 - Plot a Bar Graph using Age and Total Purchases

In [25]:
```
fig = px.histogram(data_frame=Capmaign_Data, x='Age', y='Total_purchases_sum',height=60
                   color_discrete_sequence=['#330C13'])

fig.update_layout(title_text='Customers Age vs Total Purchases',
                    xaxis_title='Age',
                    yaxis_title='Total Purchases sum',
                  font=dict(
        family="Courier New, monospace",
        size=15,
        color="black")
```

```
)

fig.show()
```



Graph4 Observations:

Above graph suggests that the customes of Age 20-80 are the actual people who purchases the most.

Conclusion:

Based on above 4 graphs, it is observed to conclude that the potential customers who are more likely to respond to the campaigns/offers are 1. The couple 2. Between Age 20-80 3. with no kids or atmost 1 kid 4. have a graduation degree.

In [ ]: