In [1]:
```
greetings = "Assalam-o-Alaikum!"
print(greetings)
```

```
Assalam-o-Alaikum!
```

Import Libraries

In [2]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

Import Dataset

In [5]:
```
df = pd.read_csv("marketing_campaign.csv", delimiter = "\t")
df
```

Out[5]:

	ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recency	MntWines	...	NumWebVisitsMonth
0	5524	1957	Graduation	Single	58138.0	0	0	04-09-2012	58	635	...	7
1	2174	1954	Graduation	Single	46344.0	1	1	08-03-2014	38	11	...	5
2	4141	1965	Graduation	Together	71613.0	0	0	21-08-2013	26	426	...	4
3	6182	1984	Graduation	Together	26646.0	1	0	10-02-2014	26	11	...	6
4	5324	1981	PhD	Married	58293.0	1	0	19-01-2014	94	173	...	5
...	...	...	...	...	...	...	...	...	...	...	...	...
2235	10870	1967	Graduation	Married	61223.0	0	1	13-06-2013	46	709	...	5
2236	4001	1946	PhD	Together	64014.0	2	1	10-06-2014	56	406	...	7
2237	7270	1981	Graduation	Divorced	56981.0	0	0	25-01-2014	91	908	...	6
2238	8235	1956	Master	Together	69245.0	0	1	24-01-2014	8	428	...	3
2239	9405	1954	PhD	Married	52869.0	1	1	15-10-2012	40	84	...	7

2240 rows × 29 columns

In [6]:
```
df.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
```

In [7]:
```
df.describe()
```

Out[7]:

	ID	Year_Birth	Income	Kidhome	Teenhome	Recency	MntWines	MntFruits	MntMeatProducts	MntFis
count	2240.000000	2240.000000	2216.000000	2240.000000	2240.000000	2240.000000	2240.000000	2240.000000	2240.000000	2
mean	5592.159821	1968.805804	52247.251354	0.444196	0.506250	49.109375	303.935714	26.302232	166.950000	
std	3246.662198	11.984069	25173.076661	0.538398	0.544538	28.962453	336.597393	39.773434	225.715373	
min	0.000000	1893.000000	1730.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	
25%	2828.250000	1959.000000	35303.000000	0.000000	0.000000	24.000000	23.750000	1.000000	16.000000	
50%	5458.500000	1970.000000	51381.500000	0.000000	0.000000	49.000000	173.500000	8.000000	67.000000	
75%	8427.750000	1977.000000	68522.000000	1.000000	1.000000	74.000000	504.250000	33.000000	232.000000	
max	11191.000000	1996.000000	666666.000000	2.000000	2.000000	99.000000	1493.000000	199.000000	1725.000000	

8 rows × 26 columns

Data Analysis

Average Income Of Each Employee

In [10]:
```
income = df["Income"].agg("mean").round(2)
print("The Average Income of Employee = $" + str(income))
```

```
The Average Income of Employee = $52247.25
```

Average Income across Employee`s Education

In [14]:
```
Education = df.groupby("Education")["Income"].agg("mean").round(2).to_frame().reset_index()
Education
```

Out[14]:

	Education	Income
0	2n Cycle	47633.19
1	Basic	20306.26
2	Graduation	52720.37
3	Master	52917.53
4	PhD	56145.31

Who spent more amount on wine product also find ID's Marital status.

In [41]:
```
amount = df.groupby(["ID", "Marital_Status"])["MntWines"].agg("max").to_frame().reset_index().sort_values("MntW
amount = amount.iloc[0]
print("ID", "'"+str(amount["ID"])+"'", "has spent more Amount on Wine Product = $" + str(amount["MntWines"]))
print("ID", "'"+str(amount["ID"])+"'", "has Martial Status is " + str(amount["Marital_Status"])+".")
```

```
ID '737' has spent more Amount on Wine Product = $1493
ID '737' has Martial Status is Married.
```

What is the minimum and maximum Number of Recency?

In [45]:
```
Min = df["Recency"].min()
Min
```

Out[45]:
```
0
```

In [46]:
```
Max = df["Recency"].max()
Max
```

Out[46]:
```
99
```

In [48]:
```
print("Min Recency = " + str(Min))
print("Max Recency = " + str(Max))
```

```
Min Recency = 0
Max Recency = 99
```

How Many Costumers have Same Year of Birth?

In [57]:
```
Costumers = df.groupby("Year_Birth")["ID"].agg("count").to_frame().reset_index()
Costumers.columns = ["Year Birth", "Costumers"]
Costumers
```

Out[57]:

	Year Birth	Costumers
0	1893	1
1	1899	1
2	1900	1

3	1940	1
4	1941	1
5	1943	7
6	1944	7
7	1945	8
8	1946	16
9	1947	16
10	1948	21
11	1949	30
12	1950	29
13	1951	43
14	1952	52
15	1953	35
16	1954	50
17	1955	49
18	1956	55
19	1957	43
20	1958	53
21	1959	51
22	1960	49
23	1961	36
24	1962	44
25	1963	45
26	1964	42
27	1965	74
28	1966	50
29	1967	44
30	1968	51
31	1969	71
32	1970	77
33	1971	87
34	1972	79
35	1973	74
36	1974	69
37	1975	83
38	1976	89
39	1977	52
40	1978	77
41	1979	53
42	1980	39
43	1981	39
44	1982	45
45	1983	42
46	1984	38
47	1985	32
48	1986	42
49	1987	27
50	1988	29
51	1989	30
52	1990	18
53	1991	15
54	1992	13
55	1993	5
56	1994	3
57	1995	5

58	1996	2

In [ ]: