

Excel Tasks:

1. Data Exploration:

Task 1: Create a statistical summary for numerical features

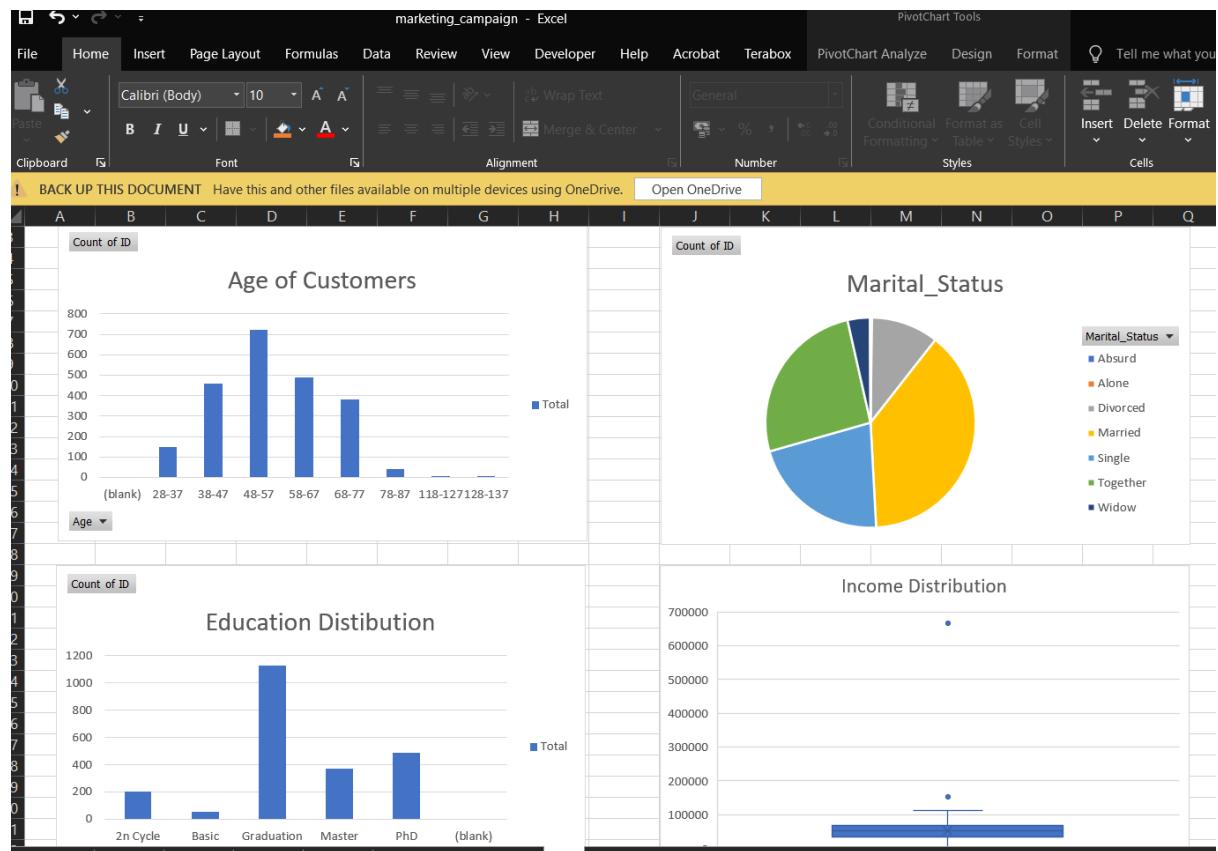
Task 2: Create a line chart for the number of enrolments by year

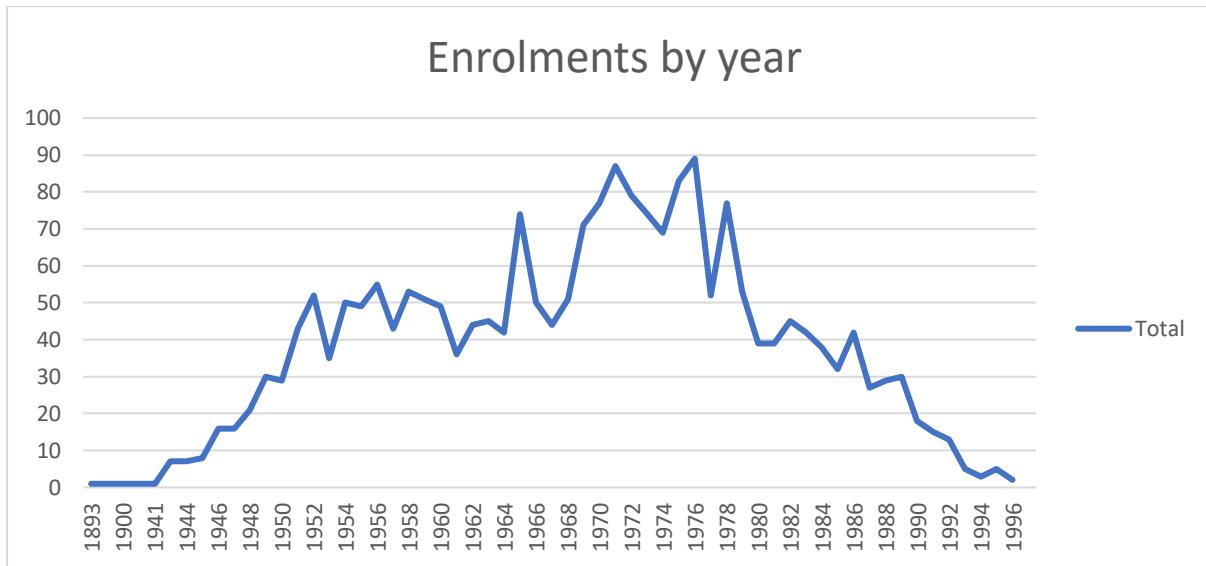
Task 3: Give a cross-tabulated count for response values against education

Task 4: Make a boxplot on income and write your observations

Task 5: Calculate the age of customers and make a histogram of that

Task 6: Visualize the response against Marital_Status





Data Loading: Create a schema named "retail_data"

Set "retail_data" as the default schema

Create tables to store the retail transaction data

Set ct_customer as the datetime field while loading the data and apply the appropriate date format

The screenshot shows the MySQL Workbench interface. In the top navigation bar, 'Local instance MySQL80' is selected. The 'Schemas' tree on the left shows several schemas like 'ceo', 'computer_manufactures', 'db_jan', 'db_movies', 'finances', and 'retail_data'. The 'marketing_campaign' table under 'retail_data' is selected. The main pane displays a SQL script with several tasks and a result grid showing data from the 'marketing_campaign' table. The result grid has columns: ID, Year_Birth, Education, Marital_Status, Income, Kidhome,Teenhome, Dt_Customer, Recency, MntWines, MntFruits, MntMeatProducts, MntFishProducts, MntSweetProducts, MntGoldProds, NumDealsPurchases, NumWebPur.

ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recency	MntWines	MntFruits	MntMeatProducts	MntFishProducts	MntSweetProducts	MntGoldProds	NumDealsPurchases	NumWebPur
5524	1957	Graduation	Single	58138	0	0	04-09-2012	58	635	88	546	172	88	88	3	8
2174	1954	Graduation	Single	46344	1	1	08-03-2014	38	11	1	6	2	1	6	2	1
4141	1965	Graduation	Together	71613	0	0	21-08-2013	26	426	49	127	111	21	42	1	8
6182	1984	Graduation	Together	26646	1	0	10-02-2014	26	11	4	20	10	3	5	2	2

1 Data Preprocessing:

Calculate the total number of customer encounters in the marketing campaign dataset

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: cape stone ff capestone RA

SCHEMAS Filter objects

- cep
- computer_manufactures
- db_jan
- db_movies
- finance
- finances
- retail_data

Tables marketing_campaign Views Stored Procedures Functions

Administration Schemas Information No object selected

```

1 • CREATE schema retail_data;
2 • USE retail_data;
3
4 #2 task 2
5 • SELECT * from retail_data.marketing_campaign;
6 #2 task 3
7
8 #3 task 1
9 • USE retail_data;
10 • SELECT count(*) AS total_transaction from marketing_campaign;
11
12 #3 task 2

```

Result Grid | Filter Rows: Export: Wrap Cell Content: Result 2 x

total_transaction
2216

Action Output

#	Time	Action	Message	Duration / Fetch
1	19:42:28	SELECT * from retail_data.marketing_campaign	2216 row(s) returned	0.016 sec / 0.015 sec
2	19:45:48	USE retail_data	0 row(s) affected	0.000 sec
3	19:45:48	SELECT count(*) AS total_transaction from marketing_campaign	1 row(s) returned	0.015 sec / 0.000 sec

Identify the top 10 most purchased products in the dataset, such as Wines, Meat Products, etc.

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: cape stone ff capestone RA

SCHEMAS Filter objects

- cep
- computer_manufactures
- db_jan
- db_movies
- finance
- finances
- retail_data

Tables marketing_campaign Views Stored Procedures Functions

Administration Schemas Information No object selected

```

26
27 -- Select from the CTE and order to find the top 10
28
29     SELECT
30         ProductCategory,
31         TotalAmount
32     FROM
33         ProductTotals
34     ORDER BY
35         TotalAmount DESC
36     LIMIT 10;
37

```

Result Grid | Filter Rows: Export: Wrap Cell Content: Result 3 x

ProductCategory	TotalAmount
Wines	676083
Meat Products	370063
Gold Products	97427
Fish Products	83405
Sweet Products	59896

Action Output

#	Time	Action	Message	Duration / Fetch
1	19:42:28	SELECT * from retail_data.marketing_campaign	2216 row(s) returned	0.016 sec / 0.015 sec
2	19:45:48	USE retail_data	0 row(s) affected	0.000 sec
3	19:45:48	SELECT count(*) AS total_transaction from marketing_campaign	1 row(s) returned	0.015 sec / 0.000 sec
4	19:47:20	CREATE schema retail_data	Error Code: 1007. Can't create database 'retail_data'...	0.031 sec
5	19:47:50	SELECT 'Meat Products', SUM(Mnt.MeatProducts) ...	Error Code: 1064. You have an error in your SQL sy...	0.000 sec
6	19:48:11	WITH ProductTotals AS (SELECT 'Wines' AS Pr...	6 row(s) returned	0.016 sec / 0.000 sec

Find the count of response values

MySQL Workbench

Local instance MySQL80 ×

File Edit View Query Database Server Tools Scripting Help

Navigator cape_stone f capestone RA ×

SCHEMAS

- cep
- computer_manufactures
- db_jan
 - Tables
 - Views
 - Stored Procedures
 - Functions
- db_movies
- finance
- finances
- retail_data
 - Tables
 - marketing_campaign
 - Views
 - Stored Procedures

Administration Schemas

Information No object selected

cape_stone f capestone RA ×

33 ORDER BY Execute the selected portion of the script or everything, if there is no selection

34 LIMIT 10;

35

36

37

38

39 #3 task 3

40 • USE retail_data;

41 • SELECT count(*) AS response from marketing_campaign;

42

43

44 #3 task 4

Result Grid | Filter Rows: Export: Wrap Cell Content: □

response
2216

Result 4 ×

Action Output

#	Time	Action	Message	Duration / Fetch
3	19:45:48	SELECT count(*) AS total_transaction from market...	1 row(s) returned	0.015 sec / 0.000 sec
4	19:47:20	CREATE schema retail_data	Error Code: 1007. Can't create database 'retail_da...	0.031 sec
5	19:47:50	SELECT 'Meat Products', SUM(MntMeatProducts)...	Error Code: 1064. You have an error in your SQL s...	0.000 sec
6	19:48:11	WITH ProductTotals AS (SELECT 'Wines' AS ...	6 row(s) returned	0.016 sec / 0.000 sec
7	19:49:41	USE retail_data	0 row(s) affected	0.000 sec
8	19:49:41	SELECT count(*) AS response from marketing_ca...	1 row(s) returned	0.015 sec / 0.000 sec

Object Info Session

Determine the distribution of customers based on their education level and marital status

The screenshot shows the MySQL Workbench interface with a query editor and a results grid.

Query Editor:

```
45 • SELECT
46     Education,
47     Marital_Status,
48     COUNT(*) AS NumberOfCustomers
49 FROM
50     marketing_campaign
51 GROUP BY
52     Education,
53     Marital_Status
54 ORDER BY
55     NumberOfCustomers DESC;
56
```

Results Grid:

Education	Marital_Status	NumberOfCustomers
Graduation	Married	429
Graduation	Together	285
Graduation	Single	246
PhD	Married	190
Master	Married	138

Identify the average income of customers who participated in the marketing campaign

MySQL Workbench

Local instance MySQL80 ×

File Edit View Query Database Server Tools Scripting Help

cape stone ff capestone RA* ×

Schemas

Filter objects

cep computer_manufactures db_jan db_movies finance finances retail_data

Tables marketing_campaign Views Stored Procedures Functions

Administration Schemas Information

Columns:

ID	int
Year_Birth	int
Education	te
Marital_Status	te
Income	int
Kidhome	int
Teenhome	int
Dt_Customer	te
Recency	int
MntWines	int
MntFruits	int
MntMeatProducts	int
MntFishProducts	int
MntSweetProducts	int
MntGoldProducts	int
NumDealsPurchases	int
NumWebPurchases	int
NumCatalogPurchases	int
NumStorePurchases	int
NumWebVisitsMonth	int

Object Info Session Duerr Completed

Query Editor

```
56
57      #3 task 5
58 •   SELECT
59          avg(Income)
60      FROM
61          marketing_campaign;
62
63
64      #3 task 7
65 •   SELECT
66          COUNT(*)
67      FROM
```

Result Grid | Filter Rows: Export: Wrap Cell Content: □

avg(Income)
52247.2514

Result 9 × Read Only

Action Output

#	Time	Action	Message	Duration / Fetch
14	20:02:42	ALTER TABLE marketing_campaign ADD COLU...	0 row(s) affected Records: 0 Duplicates: 0 Wami...	0.125 sec
15	20:02:42	UPDATE marketing_campaign SET Age1 = (EXT...	2216 row(s) affected Rows matched: 2216 Chan...	0.406 sec
16	20:03:40	ALTER TABLE marketing_campaign ADD COLU...	Error Code: 1060. Duplicate column name 'Age_gr...	0.000 sec
17	20:03:54	SELECT Age_group, AVG(NumWebVisitsMo...	4 row(s) returned	0.015 sec / 0.000 sec
18	20:08:43	SELECT COUNT(*) FROM marketing_campa...	2 row(s) returned	0.000 sec / 0.000 sec
19	20:48:46	SELECT avg(Income) FROM marketing_ca...	1 row(s) returned	0.000 sec / 0.000 sec

Identify the distribution of customers' responses to the last campaign

MySQL Workbench

Local instance MySQL80 x

File Edit View Query Database Server Tools Scripting Help

Navigator: cape stone ff capestone RA*

Schemas

#3 task 7

```
60 #3 task 7
61 • SELECT
62     COUNT(*)
63 FROM
64     marketing_campaign
65 GROUP BY
66 AcceptedCmp5;
67
68
69
70 #3 task 8
71 • SELECT
```

Result Grid | Filter Rows: Export: Wrap Cell Content:

COUNT(*)
2054
162

Result 8 x Read Only

Action Output

#	Time	Action	Message	Duration / Fetch
13	20:00:20	SELECT AVG(Kidhome) AS avg_children, ...	1 row(s) returned	0.000 sec / 0.000 sec
14	20:02:42	ALTER TABLE marketing_campaign ADD COLU...	0 row(s) affected Records: 0 Duplicates: 0 Warnings: 0	0.125 sec
15	20:02:42	UPDATE marketing_campaign SET Age1 = (EXT...	2216 row(s) affected Rows matched: 2216 Changed: 2216	0.406 sec
16	20:03:40	ALTER TABLE marketing_campaign ADD COLU...	Error Code: 1060. Duplicate column name 'Age_gr...	0.000 sec
17	20:03:54	SELECT Age_group, AVG(NumWebVisitsMo...	4 row(s) returned	0.015 sec / 0.000 sec
18	20:08:43	SELECT COUNT(*) FROM marketing_campa...	2 row(s) returned	0.000 sec / 0.000 sec

Information:

MILVINES	int
MntFruits	int
MntMeatProducts	int
MntFishProducts	int
MntSweetProducts	int
MntGoldProds	int
NumDealsPurchases	int
NumWebPurchases	int
NumCatalogPurchases	int
NumStorePurchases	int
NumWebVisitsMonth	int
AcceptedCmp3	int
AcceptedCmp4	int
AcceptedCmp5	int
AcceptedCmp1	int
AcceptedCmp2	int
Complain	int
Z_CostContact	int
Z_Revenue	int
Response	int
Age	int

Object Info Session

Calculate the average number of children and teenagers in customers' households

The screenshot shows the MySQL Workbench interface. The top bar displays "MySQL Workbench" and "Local instance MySQL80". The menu bar includes File, Edit, View, Query, Database, Server, Tools, Scripting, and Help. Below the menu is a toolbar with various icons. The main area has tabs for "Navigator" and "cape stone ff" (which is currently active). The "Schemas" pane on the left lists several databases and their objects. The central pane contains a SQL editor with the following code:

```

63     marketing_campaign
64     GROUP BY
65     Response
66     ORDER BY
67     NumberOfCustomers DESC;
68
69 #3 task 8
70 • SELECT
71     AVG(Kidhome) AS avg_children,
72     AVG(Teenhome) AS avg_teens
73 FROM marketing_campaigns;
74

```

The "Result Grid" pane below shows the output:

	avg_children	avg_teens
	0.4418	0.4418 054

The bottom pane, titled "Result 6", shows the execution history:

#	Time	Action	Message	Duration / Fetch
8	19:49:41	SELECT count(*) AS response from marketing_ca...	1 row(s) returned	0.015 sec / 0.000 sec
9	19:50:42	SELECT Education, Marital_Status, COU...	31 row(s) returned	0.031 sec / 0.000 sec
10	19:58:51	SELECT AVG(Kidhome) AS avg_children, AVG(...)	Error Code: 1064. You have an error in your SQL s...	0.000 sec
11	19:59:11	AVG(Kidhome) AS avg_children, AVG(Teenho...	Error Code: 1064. You have an error in your SQL s...	0.000 sec
12	19:59:53	SELECT AVG(Kidhome) AS avg_children, ...	Error Code: 1064. You have an error in your SQL s...	0.000 sec
13	20:00:20	SELECT AVG(Kidhome) AS avg_children, ...	1 row(s) returned	0.000 sec / 0.000 sec

Create an Age column by subtracting year_birth from the current year

- Create Age_group columns based on the below condition:

WHEN Age BETWEEN 18 AND 25 THEN '18-25'

WHEN Age BETWEEN 26 AND 35 THEN '26-35'

WHEN Age BETWEEN 36 AND 45 THEN '36-45'

WHEN Age BETWEEN 46 AND 55 THEN '46-55'

ELSE '56+'

- Determine the average number of visits per month for customers in each age group

MySQL Workbench

Local instance MySQL80 ×

File Edit View Query Database Server Tools Scripting Help

Navigator cape stone if capstone RA*

schemas

```

90     ELSE '56+'
91   END;
92 •   SELECT
93       Age_group,
94           AVG(NumWebVisitsMonth) AS AvgMonthlyVisits
95   FROM
96       marketing_campaign
97   GROUP BY
98       Age_group
99   ORDER BY
100      Age_group;
101

```

Result Grid | Filter Rows: Export: Wrap Cell Content: Result 7 × Read Only

Age_group	AvgMonthlyVisits
26-35	4.9016
36-45	5.5559
46-55	5.6603
56+	5.0282

Object Info Session

Query Completed

Task 1: General Data Overview:

Exploratory Data Analysis Check the first few rows of the dataset to understand its structure

Check the data types of each column

Check for any missing values in the dataset

localhost:8888/notebooks/Desktop/csv%20files/Untitled.ipynb? Jupyter Untitled Last Checkpoint: 19 minutes ago

Anaconda Toolbox v4.20.0

File Edit View Run Kernel Settings Help Trusted

[6]: df = pd.read_csv('marketing_campaign.csv')

[7]: print(df.head())

```
ID Year_Birth Education Marital_Status Income Kidhome Teenhome \
0 5524 1957 Graduation Single 58138.0 0 0
1 2174 1954 Graduation Single 46344.0 1 1
2 4141 1965 Graduation Together 71613.0 0 0
3 6182 1984 Graduation Together 26640.0 1 0
4 5324 1981 PhD Married 58293.0 1 0

Dt_Customer Recency MntWines ... AcceptedCmp3 AcceptedCmp4 \
0 04-09-2012 58 635 ... 0 0
1 08-03-2014 38 11 ... 0 0
2 21-08-2013 26 426 ... 0 0
3 10-02-2014 26 11 ... 0 0
4 19-01-2014 94 173 ... 0 0

AcceptedCmp5 AcceptedCmp1 AcceptedCmp2 Complain Z_CostContact \
0 0 0 0 0 3
1 0 0 0 0 3
2 0 0 0 0 3
3 0 0 0 0 3
4 0 0 0 0 3

Z_Revenue Response Age \
0 11 1 67
1 11 0 70
2 11 0 59
3 11 0 40
4 11 0 43

[5 rows x 30 columns]
```

[8]: print("b. Checking the data types of each column:")
df.info()

Program Details Salesforce Remote Online E NC Nativecamp.net Google Gemini Home Untitled

localhost:8888/notebooks/Desktop/csv%20files/Untitled.ipynb? Jupyter Untitled Last Checkpoint: 31 minutes ago

Anaconda Toolbox v4.20.0

File Edit View Run Kernel Settings Help Trusted

b. Checking the data types of each column:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2240 entries, 0 to 2239
Data columns (total 30 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  
 29  Age               2240 non-null    int64
```

```
[9]: missing_values = df.isnull().sum()
print(missing_values)
```

◆ ⊞ □ ↑ ↓ ± ⌂

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

Compute summary statistics for numerical columns (mean, median, min, max, and standard deviation)

The screenshot shows a Jupyter Notebook interface running on a local host. The title bar indicates the URL is `localhost:8888/notebooks/Desktop%2Fcsv%20files%2Fretail%20analytics.ipynb`. The notebook header includes tabs for File, Edit, View, Run, Kernel, Settings, Help, and a Trusted status indicator. Below the header is a toolbar with various icons for file operations like Open, Save, and Print.

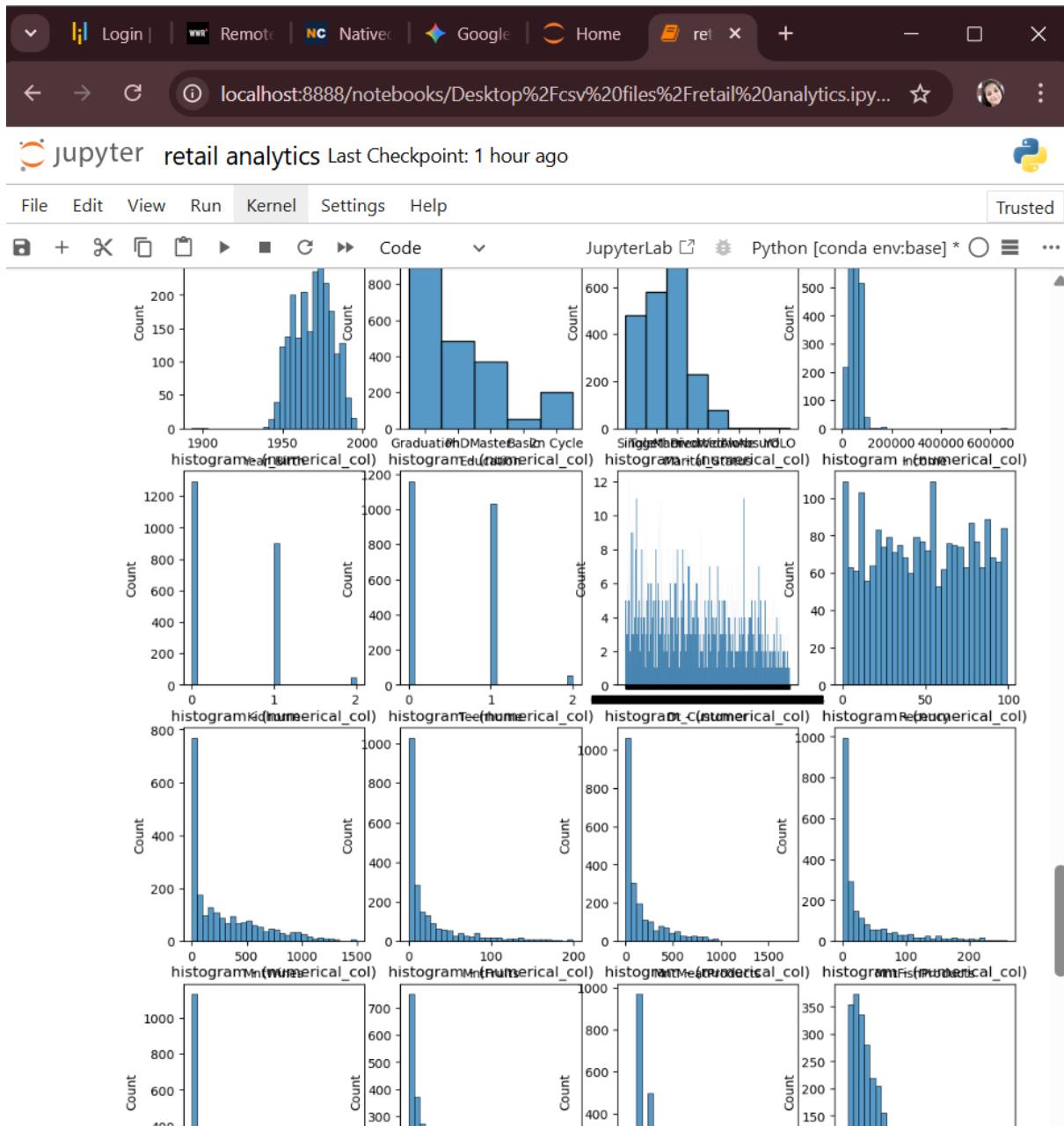
In the main workspace, cell [11] contains the following Python code:

```
[11]: summary_stats = df.describe()
print(summary_stats)
```

The output displays the descriptive statistics for the DataFrame `df`, showing the count, mean, standard deviation, minimum, 25th percentile, 50th percentile (median), 75th percentile, and maximum for each numerical column. The columns include ID, Year_Birth, Income, Kidhome, Teenhome, Recency, MntWines, MntFruits, MntMeatProducts, MntFishProducts, AcceptedCmp3, AcceptedCmp4, AcceptedCmp5, AcceptedCmp1, AcceptedCmp2, Complain, Z_CostContact, Z_Revenue, and two unnamed columns (n/a).

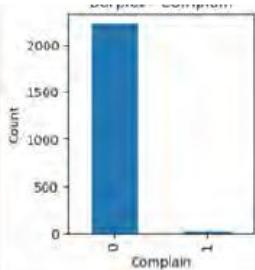
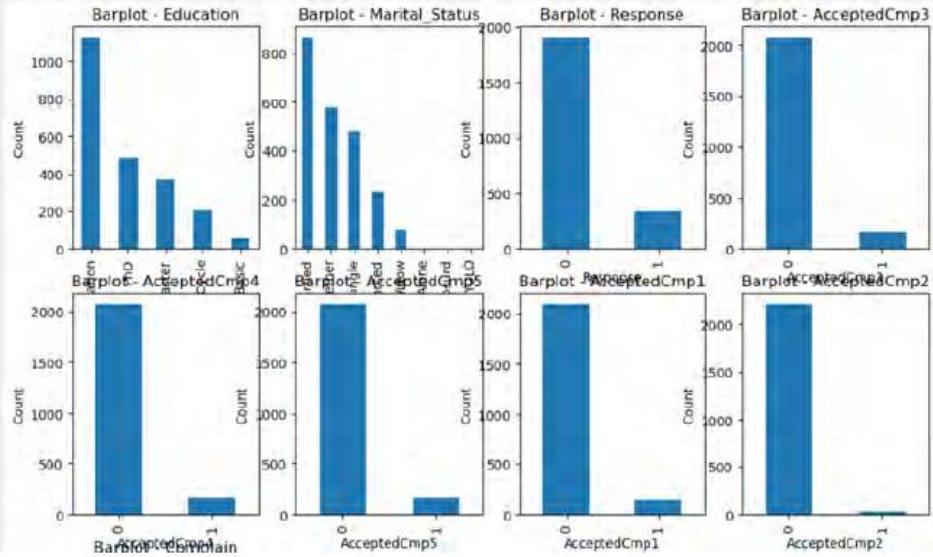
	ID	Year_Birth	Income	Kidhome	Teenhome	Recency	MntWines	MntFruits	MntMeatProducts	MntFishProducts	AcceptedCmp3	AcceptedCmp4	AcceptedCmp5	AcceptedCmp1	AcceptedCmp2	Complain	Z_CostContact	Z_Revenue	n/a	n/a
count	2240.000000	2240.000000	2216.000000	2240.000000	2240.000000	2240.000000	2240.000000	2240.000000	2240.000000	2240.000000	2240.000000	2240.000000	2240.000000	2240.000000	2240.000000	2240.000000	2240.000000	2240.000000	n/a	n/a
mean	5592.159821	1968.805804	52247.251354	0.444196	0.506250	49.109375	303.935714	26.302232	166.950000	37.525446	... 0.072768	0.074554	0.072768	37.525446	... 0.072768	0.074554	0.072768	0.074554	n/a	n/a
std	3246.662198	11.984069	25173.076661	0.538398	0.544538	28.962453	336.597393	39.773434	225.715373	54.628979	... 0.259813	0.262728	0.259813	54.628979	... 0.259813	0.262728	0.259813	0.262728	n/a	n/a
min	0.000000	1893.000000	1730.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	n/a	n/a
25%	2828.250000	1959.000000	35303.000000	0.000000	0.000000	24.000000	23.750000	1.000000	16.000000	12.000000	... 0.000000	0.000000	0.000000	24.000000	... 0.000000	0.000000	0.000000	0.000000	n/a	n/a
50%	5458.500000	1970.000000	51381.500000	0.000000	0.000000	49.000000	173.500000	8.000000	67.000000	50.000000	... 0.000000	0.000000	0.000000	49.000000	... 0.000000	0.000000	0.000000	0.000000	n/a	n/a
75%	8427.750000	1977.000000	68522.000000	1.000000	1.000000	74.000000	504.250000	33.000000	232.000000	50.000000	... 1.000000	1.000000	1.000000	74.000000	... 1.000000	1.000000	1.000000	1.000000	n/a	n/a
max	11191.000000	1996.000000	666666.000000	2.000000	2.000000	99.000000	1493.000000	199.000000	1725.000000	259.000000	... 1.000000	1.000000	1.000000	99.000000	... 1.000000	1.000000	1.000000	1.000000	n/a	n/a

Explore the distribution of numerical variables using histograms or box plots

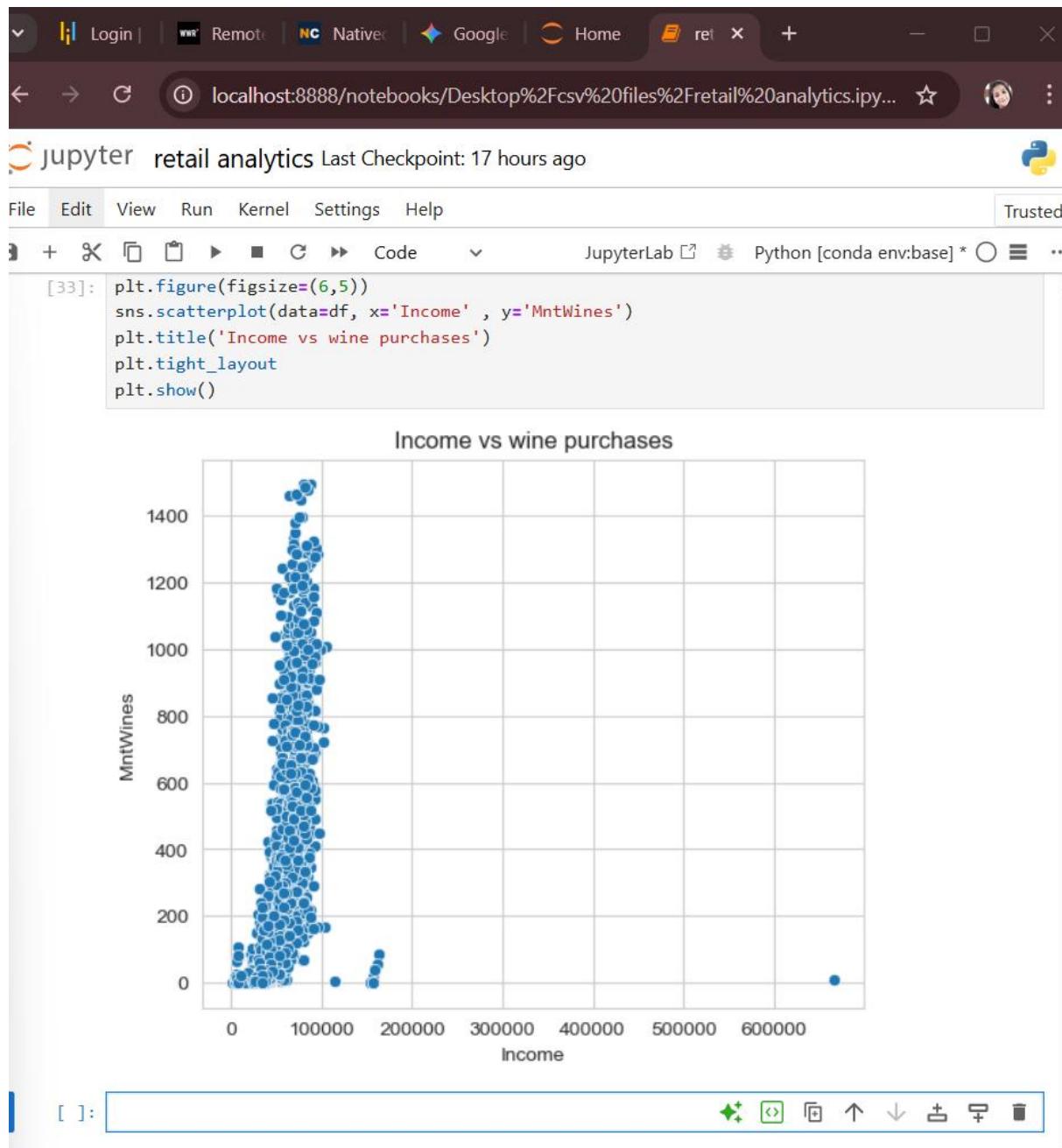


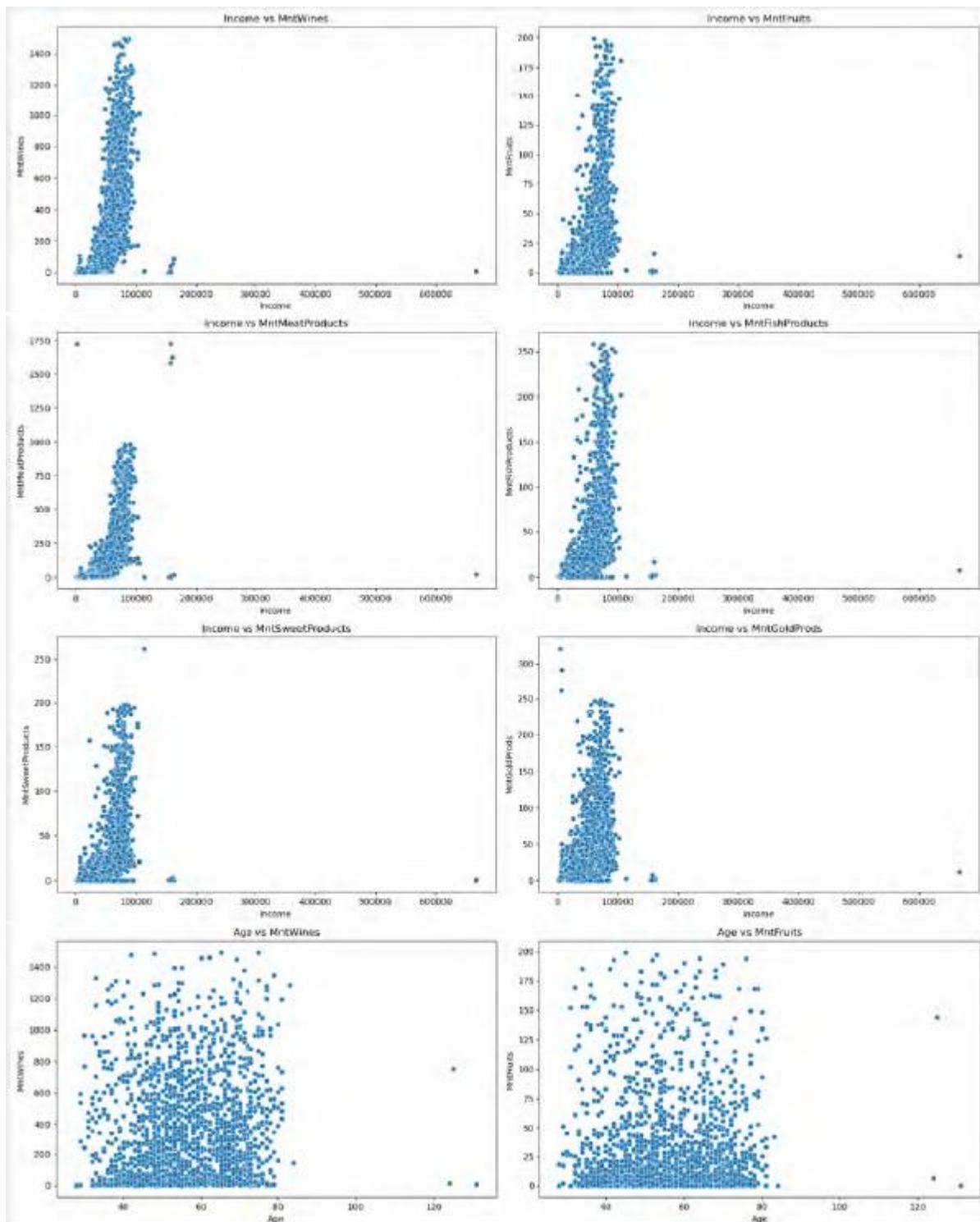
Explore the distribution of each categorical variable using bar plots or pie charts

```
# Plot a Histogram
plt.figure(figsize=(12,18))
for i , col in enumerate(categorical_col):
    plt.subplot(5,4,i+1)
    data[col].value_counts().plot(kind='bar')
    plt.title(f'Barplot - {col}')
    plt.xlabel(col)
    plt.ylabel('Count')
plt.tight_layout()
plt.show()
```



Identify outliers in numerical variables using box plots or scatter plots





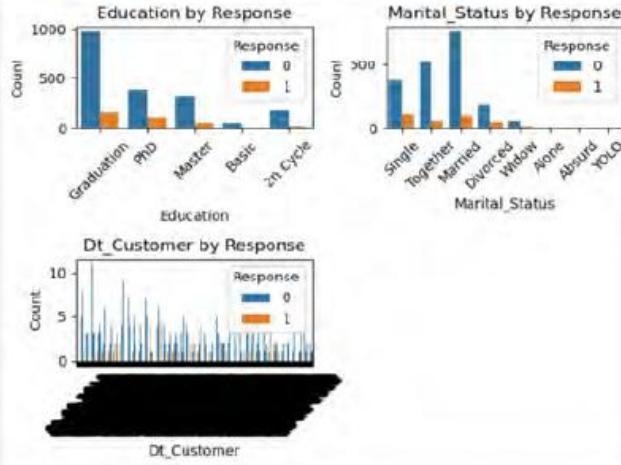
Explore the relationship between categorical variables and the target variable using bar plots or chi-square tests

```

categorical_cols = data.select_dtypes(include=['object']).columns

# Plot Response distribution by each categorical variable
for idx, col in enumerate(categorical_cols,1):
    plt.subplot(2,2,idx)
    sns.countplot(x=col, hue='Response', data=data)
    plt.title(f'{col} by Response')
    plt.xticks(rotation=45)
    plt.ylabel('Count')
    plt.tight_layout()
plt.show()

```



Explore the relationship between numerical and categorical variables using box plots or violin plots

The screenshot shows a Jupyter Notebook interface with a code cell and its corresponding output. The code cell contains:

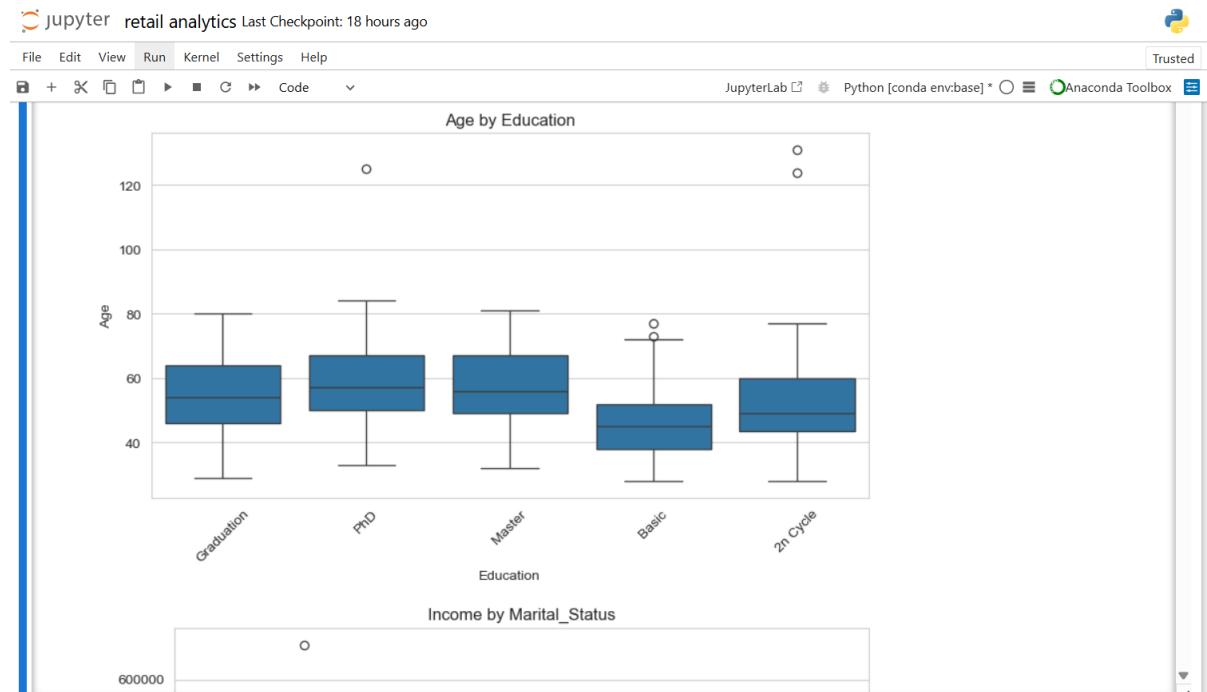
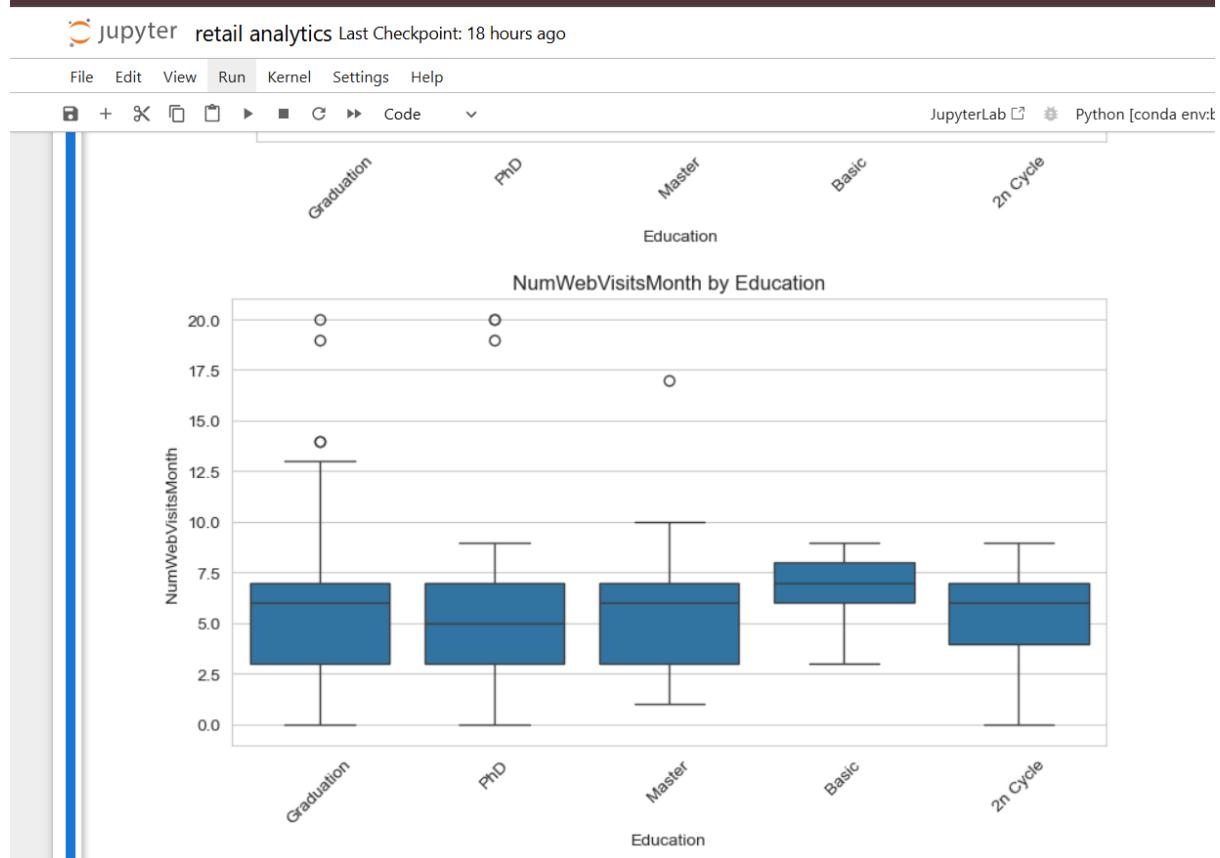
```

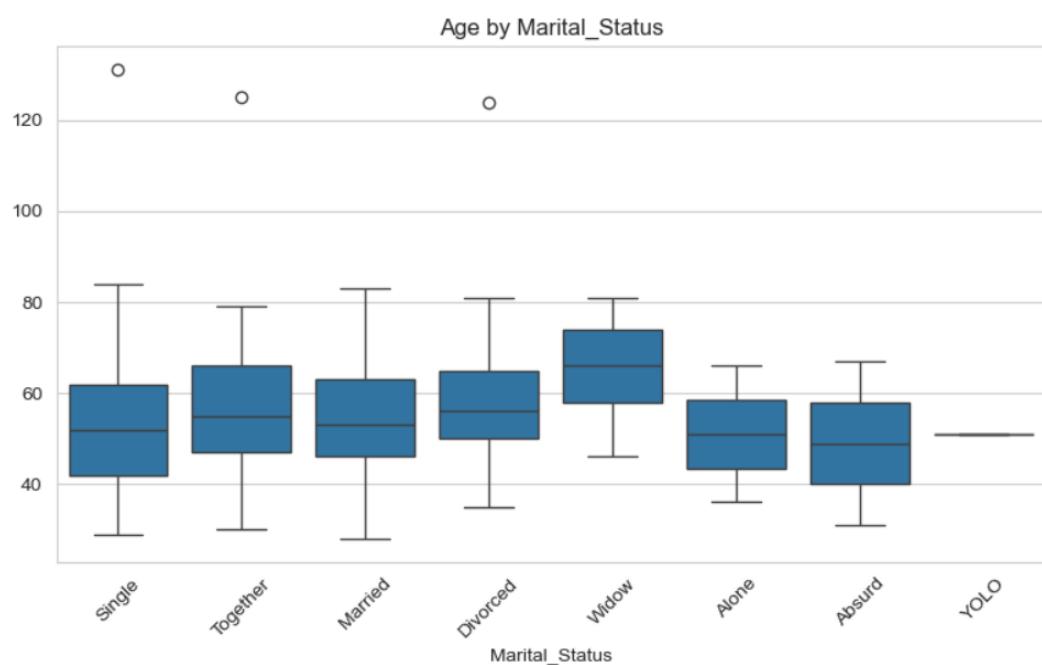
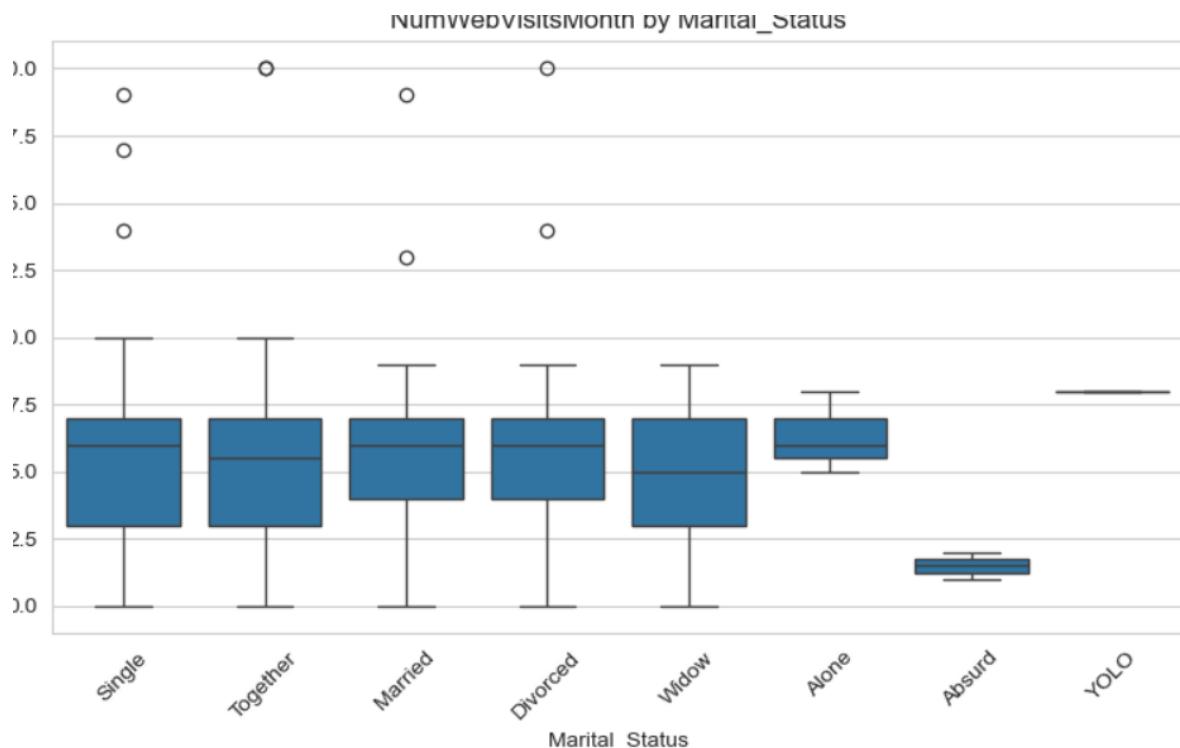
numerical_vars = ['Income', 'NumWebVisitsMonth', 'Age']

[35]: for cat in categorical_vars:
    for num in numerical_vars:
        plt.figure(figsize=(8,5))
        sns.boxplot(data=df, x=cat, y=num)
        plt.title(f'({num}) by ({cat})')
        plt.xticks(rotation=45)
        plt.tight_layout()
    plt.show()

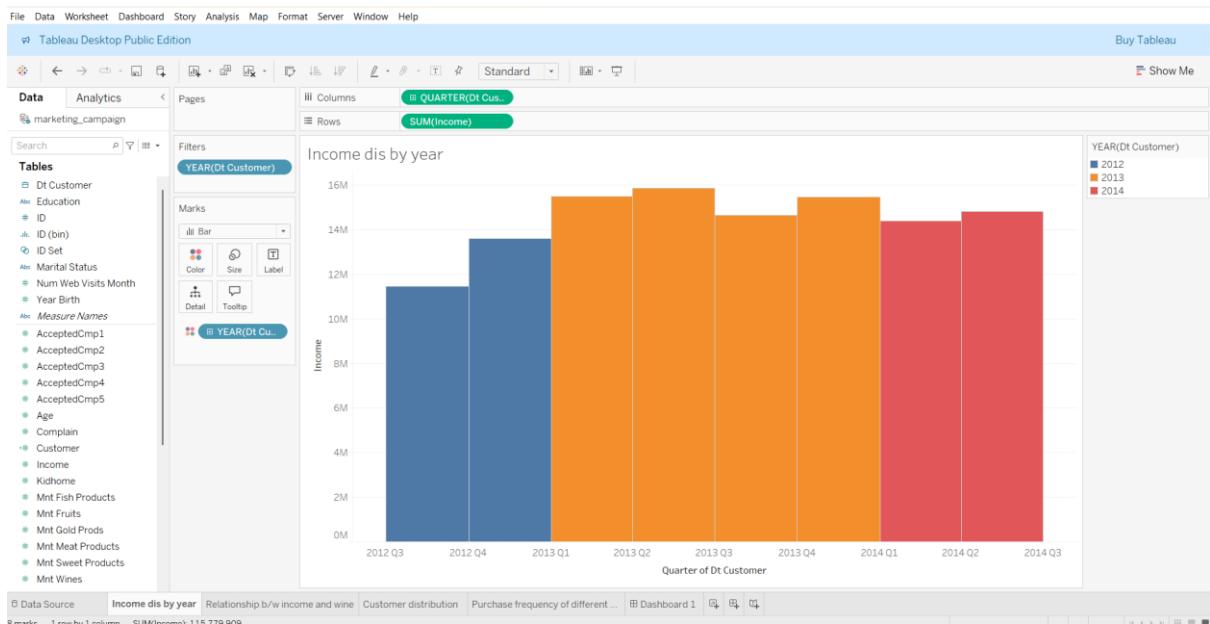
```

The output of the code is a box plot titled "Income by Education". The y-axis is labeled "Income" and ranges from 0 to 600,000. The x-axis has five categories labeled 0, 1, 2, 3, and 4. Each category has a blue box plot showing the distribution of income. The median income appears to be around 50,000-70,000 for categories 0, 1, and 2, and around 20,000 for categories 3 and 4. There are some outliers at higher income levels for categories 0, 1, and 2.

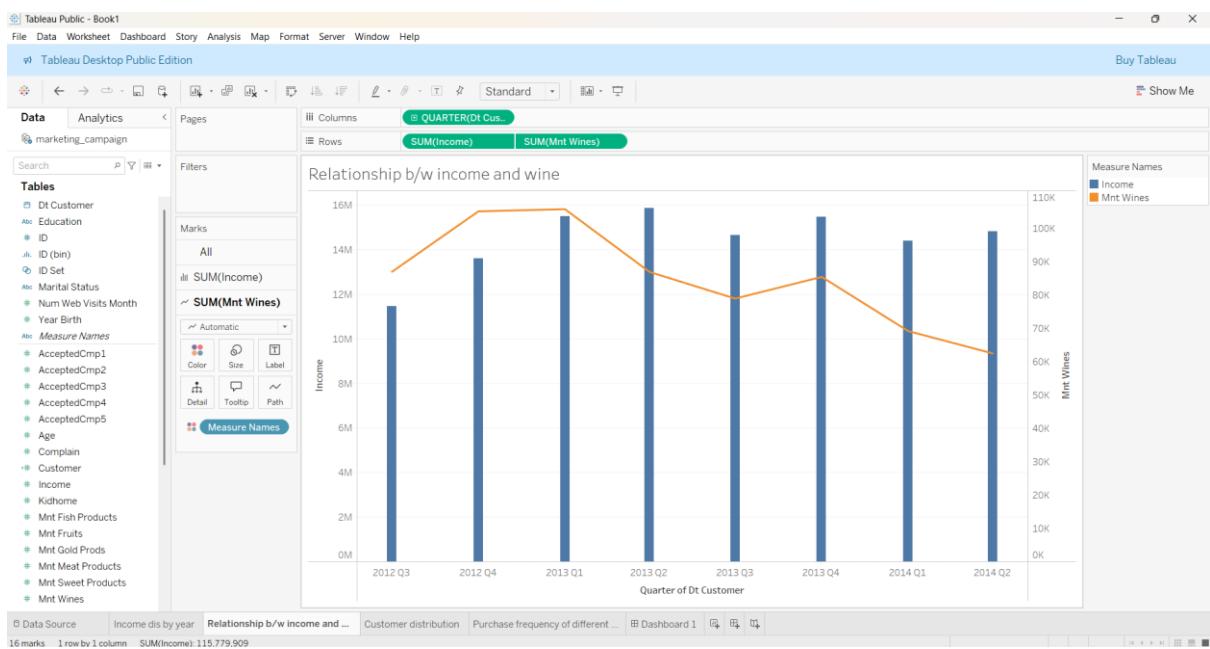




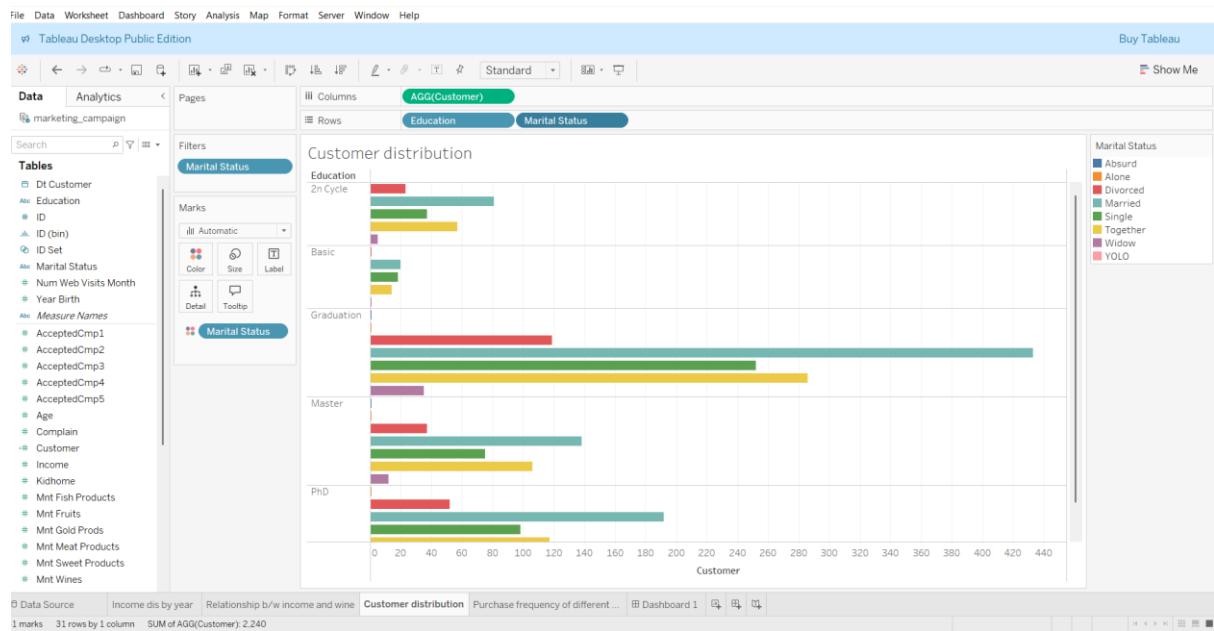
Task 1: Visualize Customer Income Distribution grouped by Year of Registration



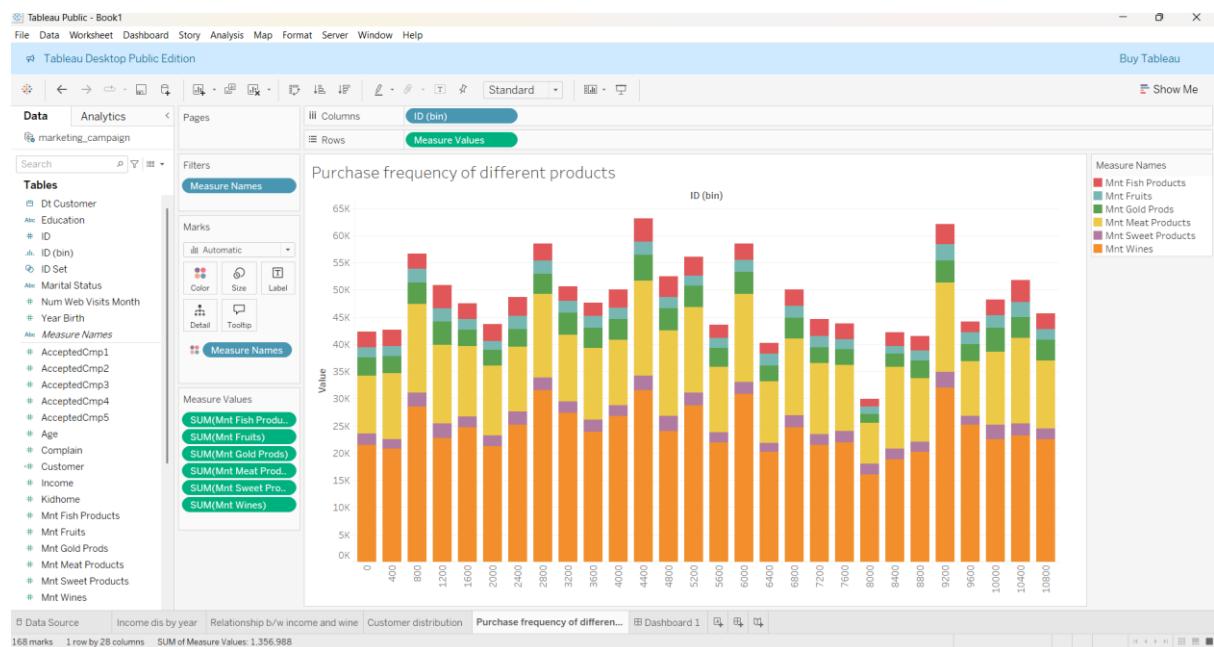
Task 2: Examine the breakdown of education levels and marital status using side-by-side circles to represent count



Task 3: Explore the relationship between Income and Wine Spending



Task 4: Analyze the frequency of purchases across various product categories



Task 5: Create a dashboard using all the visualizations

