



Exploratory Analysis and Presenting Insights



Antonio Arantes

April 2024

Index

1. Background of the business.....	3
2. Analytical approach	3
3. Dashboard design and development.....	7
4. Patterns, trends, and insights	10
Appendices	12
Appendix A – The Five Whys framework	12
Appendix B – Data cleaning	13
Appendix C – Excel exploration.....	14
Appendix D - SQL data validation.....	15
Appendix E - SQL queries and further analysis	18
Appendix F – Tableau calculated fields	21

1. Background of the business

2Market, an international supermarket operating through direct, in-store, and online channels, is experiencing a decline in profits. The company aims to increase sales by comprehensively understanding customer demographics, top-selling products, and advertisement channels. However, lacking insight into those areas hampers the company's ability to optimize its marketing efforts and allocate resources effectively. Improved data analytics support is needed to address these challenges and enhance the company's advertising and promotion strategies.

This problem statement was drawn with the assumption that the profits of 2Market are declining, and it was developed using the Five Whys Framework (Appendix A).

2. Analytical approach

2Market team provided two datasets: *marketing_data* and *ad_data*. The first step was to clean some data with *marketing_data* (Appendix B). The dataset has no missing values. The next step was to investigate duplicates in the data. All IDs are unique, so I did not consider that for the duplicates. In total, 47 duplicates were removed using the *Remove Duplicates* tool in Excel.

The next stage was to check outliers in *Age* and *Income* data. The quartiles 1 and 3, the Interquartile range, and the Lower and Upper Limit were calculated for that. Three records were removed due to *Age*, and eight were removed due to *Income*.

	Age	Income
Quartile 1	46	35388
Quartile 3	64	68487
IQR	18	33099
Lower Limit	19	-14260.5
Upper Limit	91	118135.5

Figure 1 - Results of Quartile 1, Quartile 3, IQR, Lower Limit and Upper Limit for Age and Income.

It was noticed that the dataset has six records with 0 purchases made online and 0 purchases made in-store. Since all customers need at least one purchase made, which

seemed inconsistent, those records were deleted from the dataset *marketing_data*. I would have to go back to the 2Market team and ask for clarification on those records.

I finished this process with the *marketing_data* dataset with 2152 records by 24 variables and *ad_data* with 2216 records by 6 variables. Since I will work with *ad_data* combined only with *marketing_data*, there was no need to remove the records that were removed in the second dataset for this case.

Excel was used to explore attributes such as Age, Marital Status, and Income. Different ranges were created for the third variable (more in Appendix C) to compare age by income.

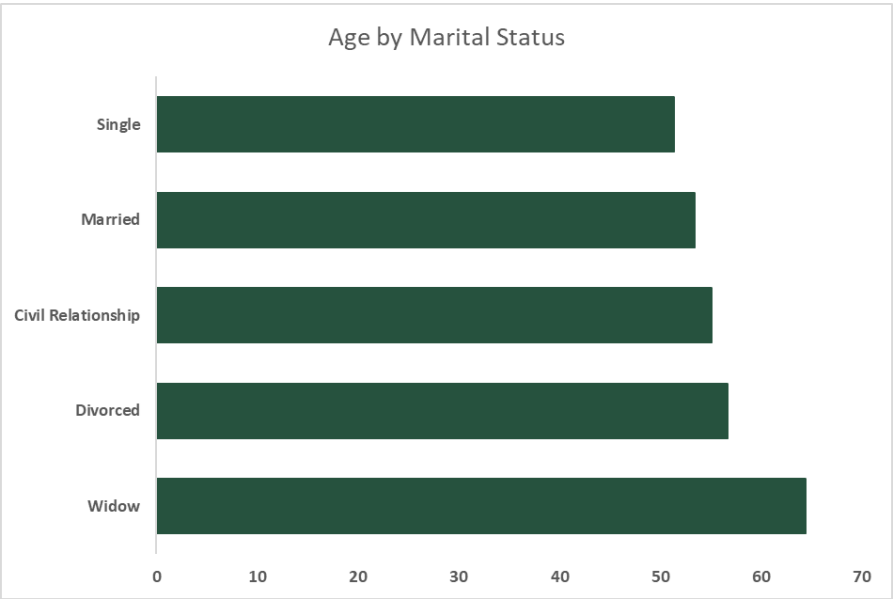


Figure 2 - Average Age by Marital Status

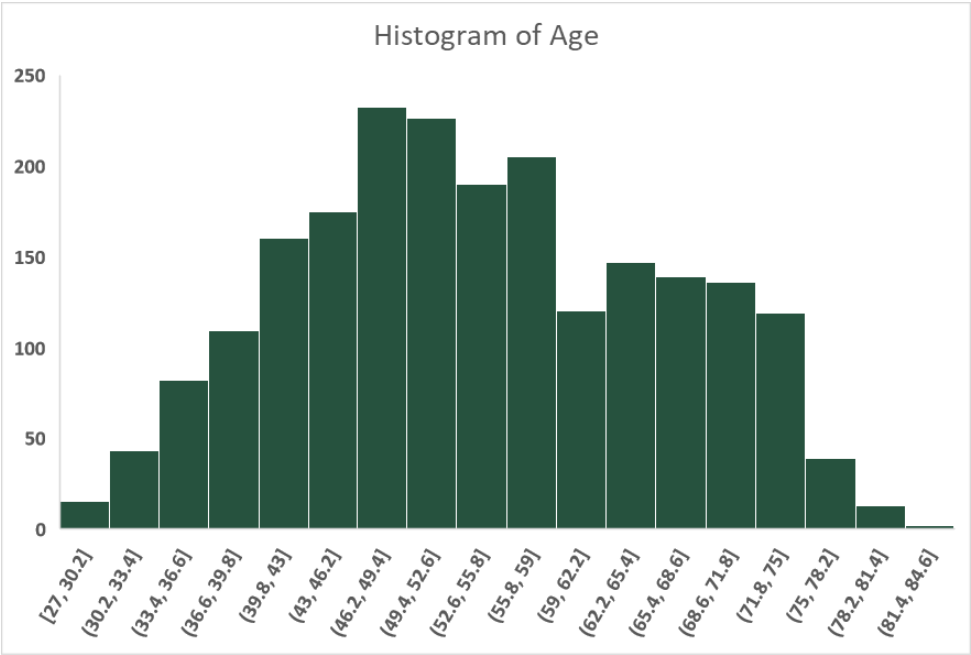


Figure 3 - Histogram of Age

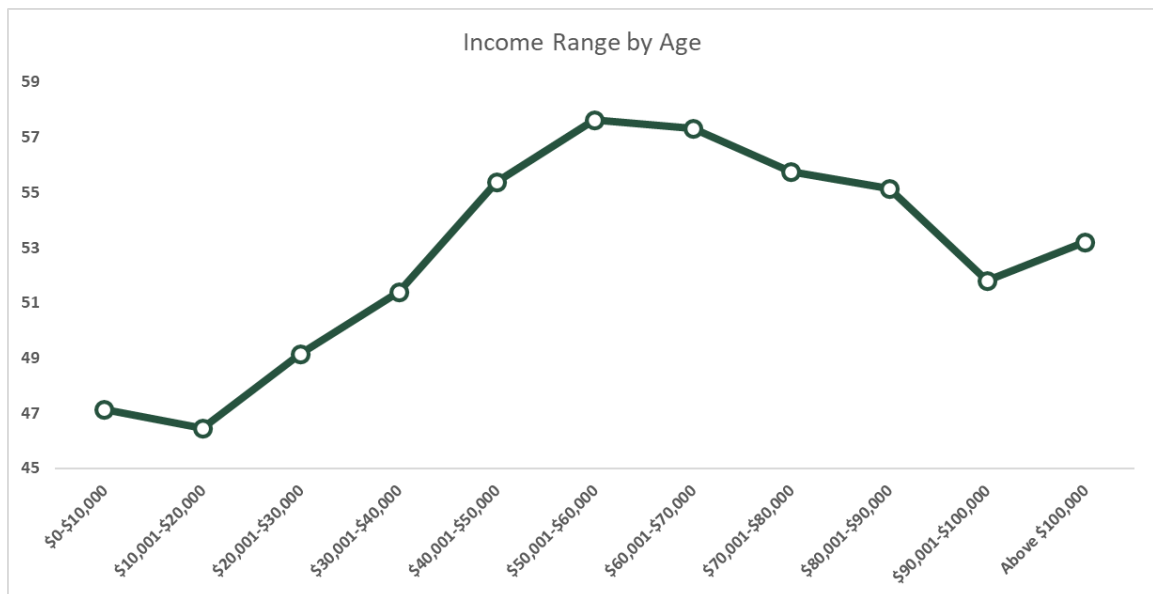


Figure 4 - Income Range by Average Age

After Excel, the initial exploratory analysis continues in SQL. Exploring data in SQL allows for more detailed analysis compared to Excel. It helps uncover more profound insights and makes manipulating and summarizing large datasets easier.

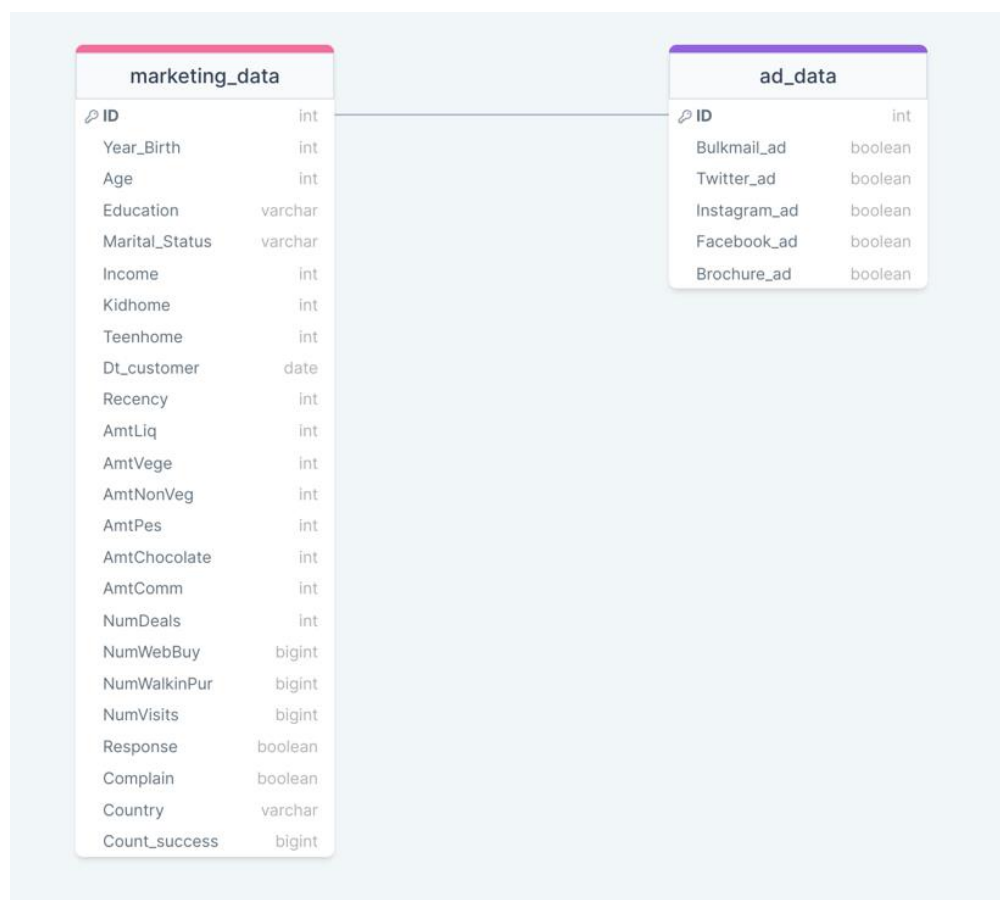


Figure 5 - Entity Relationship Diagram (ERD)

	Country character varying (30)	totalspend bigint
1	Spain	633931
2	South Africa	207525
3	Canada	167069
4	Australia	84970
5	India	74894
6	Germany	73198
7	United States of America	64099
8	Montenegro	3122

Figure 6 - Total Sales by Country

	Country character varying (30)	totalliquid bigint	totalvege bigint	totalnonveg bigint	totalpes bigint	totalchocolates bigint	totalcomm bigint	totalspend bigint
1	Spain	323340	26942	171690	38298	28989	44672	633931
2	South Africa	105014	8872	56233	13610	8919	14877	207525
3	Canada	83891	7677	45853	9962	7591	12095	167069
4	Australia	42424	3680	22204	5534	4105	7023	84970
5	India	35867	3585	21761	4661	3090	5930	74894
6	Germany	36776	2980	20272	4601	2801	5768	73198
7	United States of America	32158	3017	16838	4393	2859	4834	64099
8	Montenegro	1729	8	817	226	122	220	3122

Figure 7 - Total sales by category per Country

	bulkmail_leads numeric	twitter_leads numeric	instagram_leads numeric	facebook_leads numeric	brochure_leads numeric
1	7.39	7.53	7.39	6.51	1.35

Figure 8 - Effectiveness (%) of each advertising method

	Country character varying (30)	total_leads_percentage numeric	customers bigint
1	Montenegro	33.33	3
2	Canada	32.82	262
3	Germany	32.76	116
4	Spain	32.50	1046
5	India	26.57	143
6	South Africa	25.98	331
7	United States of America	24.76	105
8	Australia	23.29	146

Figure 9 - Effectiveness (%) of all types of leads by Country

	Country character varying (30)	customers bigint	bulkmail_leads numeric	twitter_leads numeric	instagram_leads numeric	facebook_leads numeric	brochure_leads numeric
1	Spain	1046	7.65	8.13	8.22	7.07	1.43
2	South Africa	331	6.34	6.04	6.34	6.04	1.21
3	Canada	262	6.49	9.16	8.02	6.87	2.29
4	Australia	146	6.16	4.11	8.22	4.79	0.00
5	India	143	9.09	6.99	4.20	4.90	1.40
6	Germany	116	8.62	9.48	6.90	6.03	1.72
7	United States of America	105	7.62	5.71	4.76	6.67	0.00
8	Montenegro	3	33.33	0.00	0.00	0.00	0.00

Figure 10 - Effectiveness (%) of each advertising type by Country

The SQL queries for the previous figures and more SQL analysis can be consulted in Appendix D.

The Excel and SQL analysis highlighted the importance of investigating product and advertisement shares across client categories to address the problem effectively.

3. Dashboard design and development

Tableau visually depicted correlations as the final step in our data exploration, simplifying pattern recognition and trend understanding and offering valuable insights crucial for decision-making.

The four dashboards were designed with the same approach to maintain consistency and make it easier for the stakeholders to use them. The colour scheme was built around the same colour as the fictional logo created for 2Market and shown on the cover of this report.

2Market consumers were split into three types of clients to do a customer segmentation based on the last time that a purchase was made. That information is present in the column *Recency*. If a purchase was made less or equal to 30 days, the client is considered *Current Customer*, 31 to 60 days *Irregular Customer*, and 61 days or more *Lost Customer*.

The first dashboard is called *Demographics*. Bar charts make it easy to identify the typical consumer of 2Market, including age, income, country, marital status, and education. The combination of filters allows the selection of specific variables to investigate the customer database.

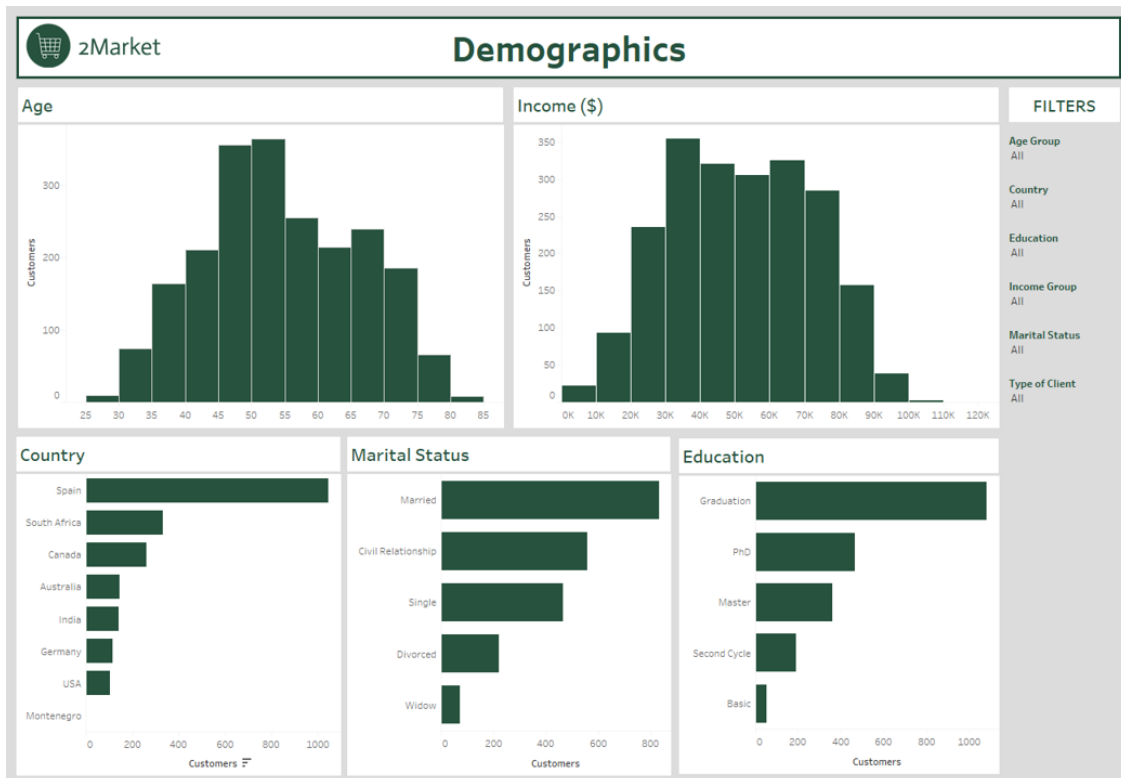


Figure 11 - First dashboard

The second dashboard is also about the customer and is called *Customer Profile*. The purpose of this dashboard is to analyse the average client of 2Market across several categories related to sales and purchases. It also includes a map with the number that displays the number of customers per Country, and the flags are an easy way to filter each territory.

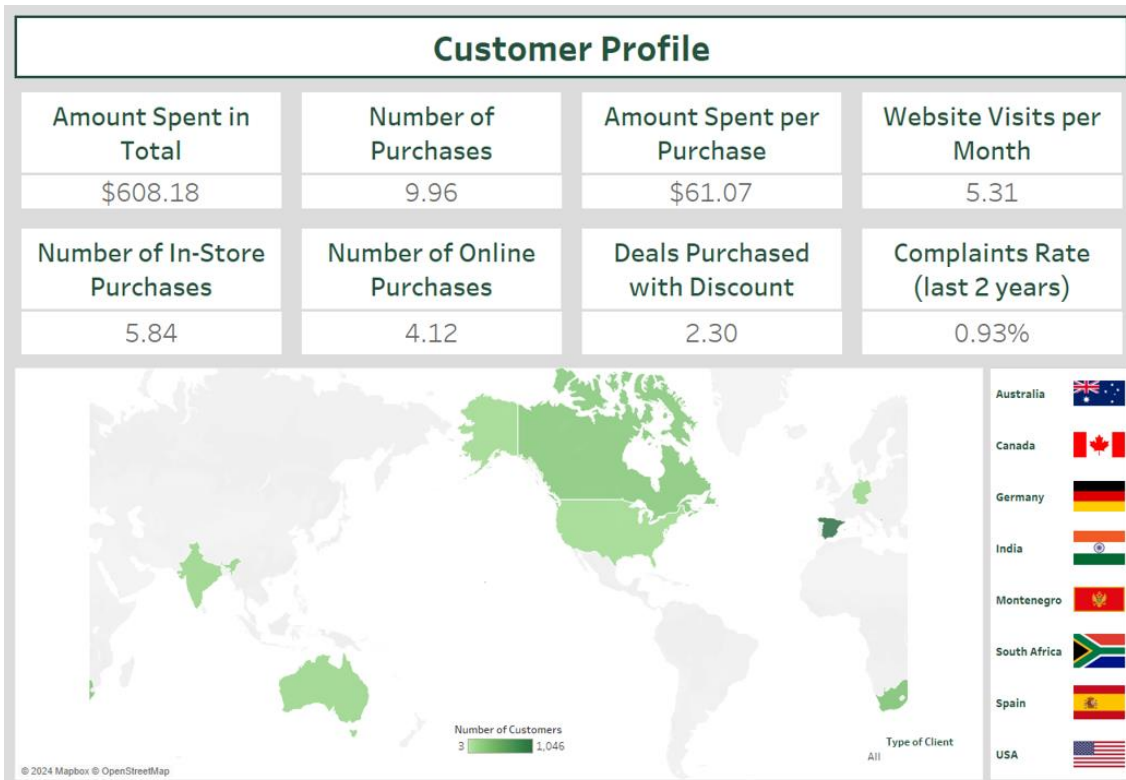


Figure 12 - Second dashboard

The third dashboard is labelled *Products*. The purpose is to see the share of each product type in general and across various demographic categories. It is easily identifiable which categories impact the product share and which ones don't have any impact. All the graphics work as a filter in case they need to be filtered by any specific country, age, marital status, or income.

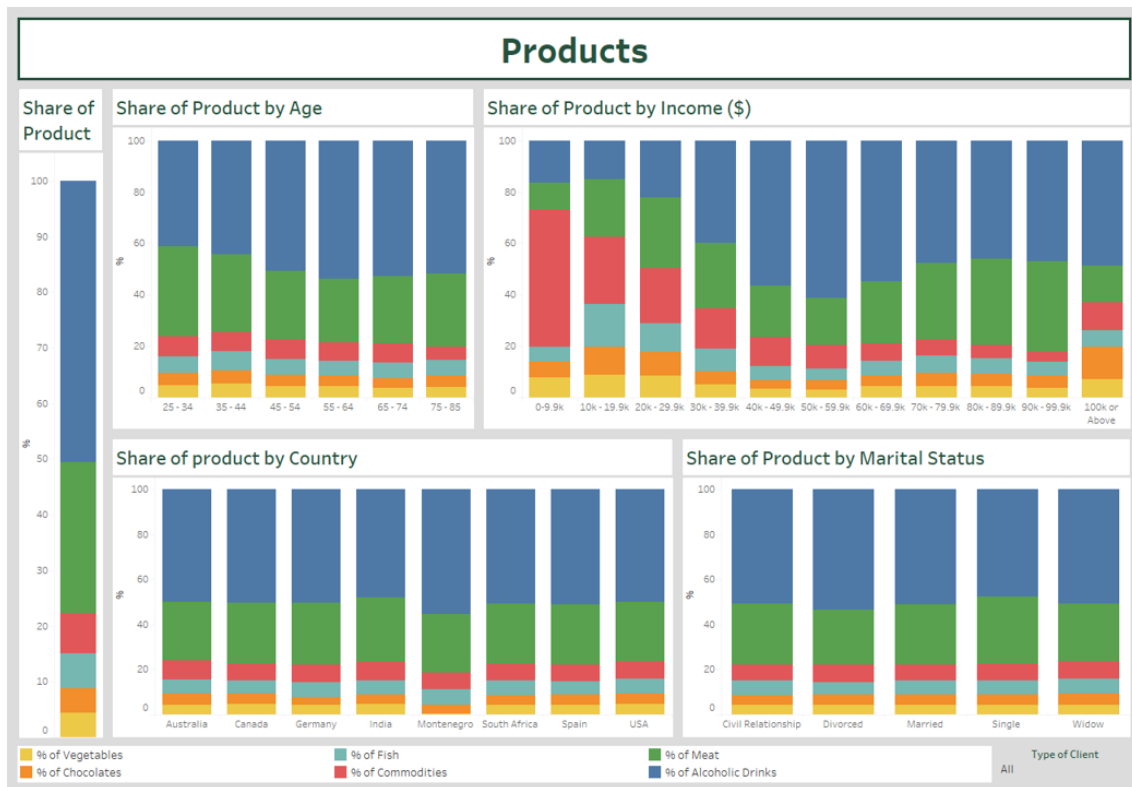


Figure 13 - Third dashboard

The fourth dashboard is called *Advertisement*. It displays the average number of successful leads per age group and income range and the effectiveness of each type of advertisement per country. Above 7% is considered the success factor of each type, and a reference line effortlessly portrays it.

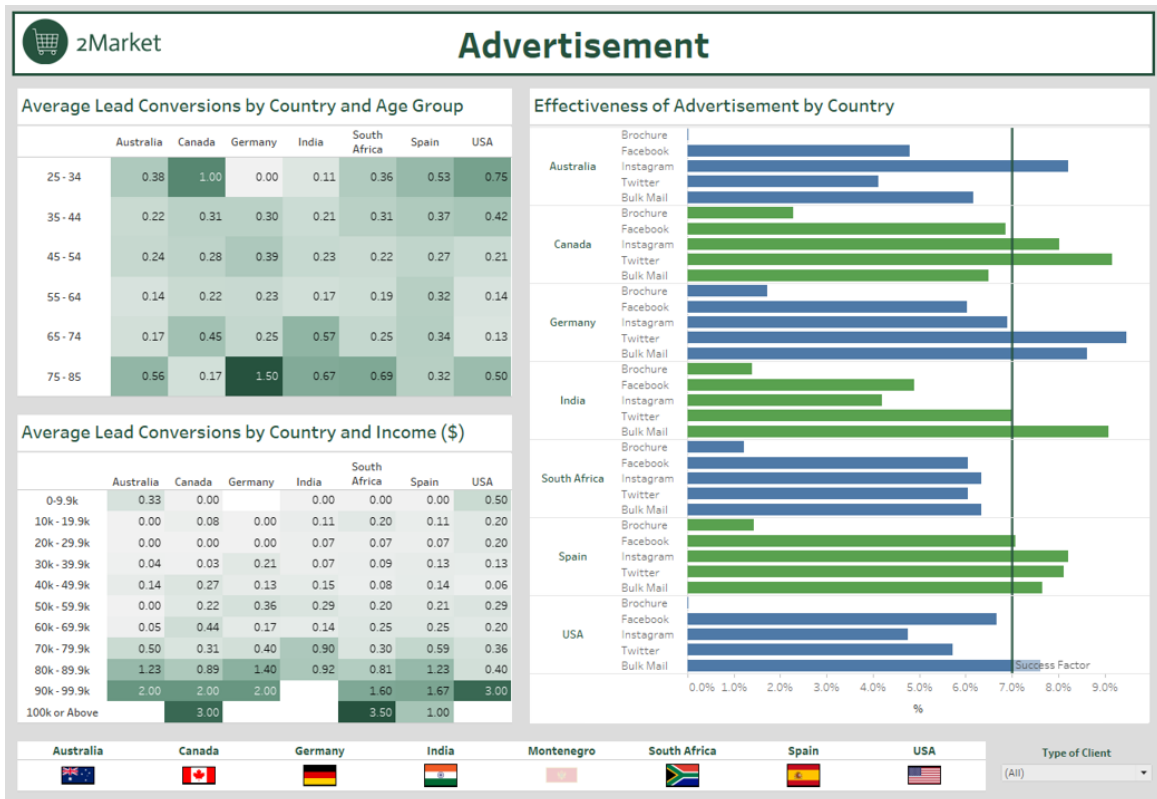


Figure 14 - Fourth dashboard

4. Patterns, trends, and insights

Our analysis identified several key patterns, trends, and insights that shed light on the decline in profits at 2Market and provided actionable recommendations for improvement.

Firstly, we discovered that the typical consumer at 2Market is from Spain, aged between 45 and 55 years, married, and holds a graduate degree. This demographic profile provides valuable insight into the target audience, enabling more targeted marketing efforts.

Furthermore, our analysis revealed that the average customer makes approximately 9.96 purchases per year, with 5.84 occurring in-store and 4.12 online, spending an average of \$61.07 per purchase. Understanding these purchasing habits allows more effective resource allocation and tailored marketing strategies.

Interestingly, the top-selling product categories are alcoholic drinks and meat. Additionally, product share varies significantly according to age and income demographics, with alcohol consumption increasing with age but decreasing with lower income levels.

Moreover, our analysis indicates that older age groups and higher income brackets are more responsive to advertising leads. Also, prioritizing advertising channels

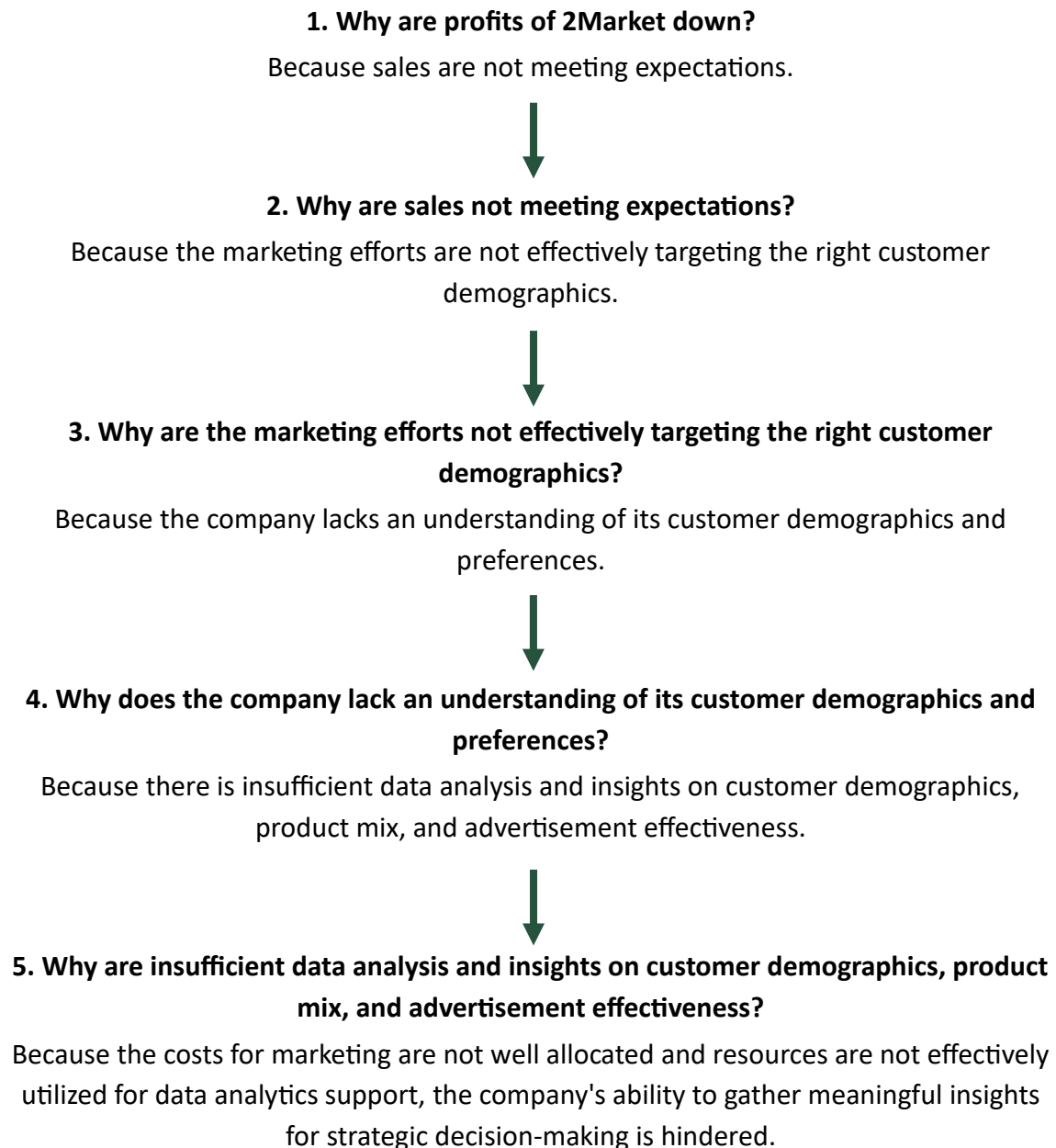
with success rates above 7% can optimize marketing efforts and reduce costs while increasing sales.

For further exploration, it would be beneficial to analyse customer order data over time to identify trends and seasonal variations in purchasing behaviour. This additional dataset would provide valuable insights into the effectiveness of marketing campaigns and promotional strategies throughout the year, enabling more informed decision-making and strategic planning.

Appendices

Appendix A – The Five Whys framework

Assumption – 2Market profits are declining.



Appendix B – Data cleaning

- The first step was correctly setting up each column with the correct data type.
- I replaced the dollar sign from the column "Income" and changed the data type to *Currency*.
- Through *Text to Columns*, *Dt_customer* was standardized with the DD/MM/YYYY format.
- Each *Country's abbreviation* was changed to the full name (for example, AU was changed to Australia).
- *Alone*, *Absurd*, and *Yolo* were considered *Single* in the *Marital Status* column. Also, *Together* was considered *Civil Relationship*.
- *ID 0* was converted to 2, so all data in that column equals or exceeds 1. This process was done in *marketing_data* and *ad_data*.

Appendix C – Excel exploration

This section will show some other pivot tables and graphics created in Excel.

Row Labels	Average of Age
Civil Relationship	55.15
Divorced	56.66
Married	53.43
Single	51.36
Widow	64.40
Grand Total	54.13

Figure 15 - Pivot used to create the graph of Age by Marital Status

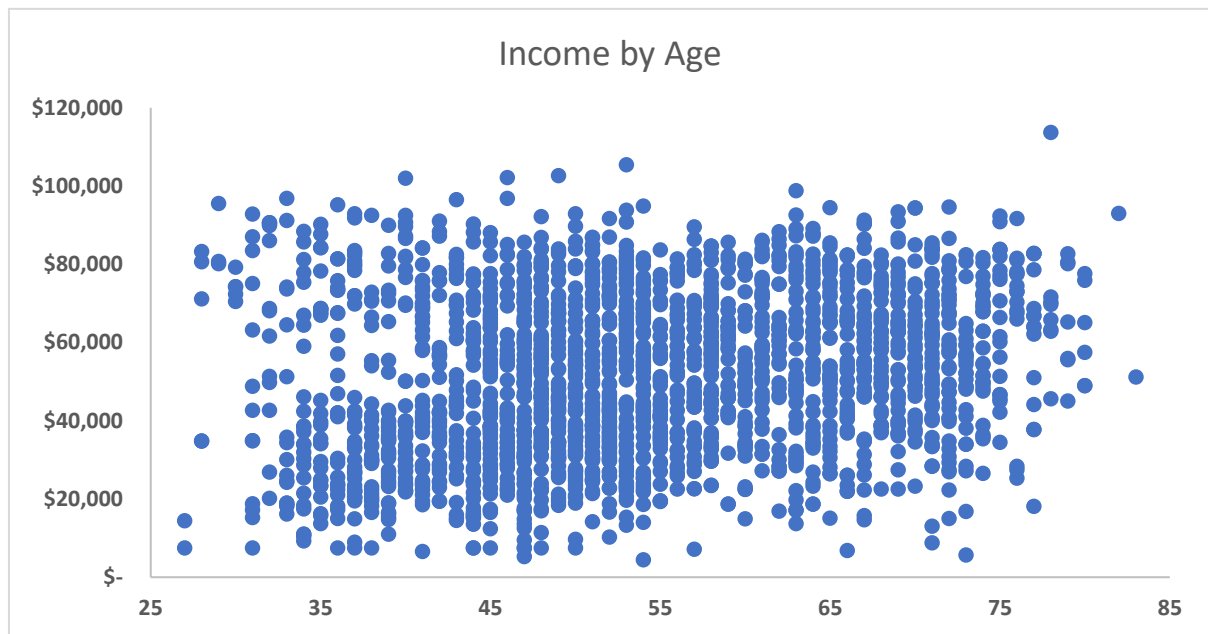


Figure 16 - Scatterplot of Income by Age before creating the Income Bins

Row Labels	Average of Age
\$10,001-\$20,000	46.45
\$20,001-\$30,000	49.14
\$30,001-\$40,000	51.41
\$40,001-\$50,000	55.40
\$50,001-\$60,000	57.64
\$60,001-\$70,000	57.31
\$70,001-\$80,000	55.75
\$80,001-\$90,000	55.14
\$90,001-\$100,000	51.79
Above \$100,000	53.20
\$0-\$10,000	46.57
Grand Total	54.13

Figure 17 - Pivot used to create the graph Income Range by Age

Appendix D - SQL data validation

```
1 SELECT *
2 FROM public.marketing_data
3 WHERE "ID" IS NULL OR
4 "Year_Birth" IS NULL OR
5 "Age" IS NULL OR
6 "Education" IS NULL OR
7 "Marital_Status" IS NULL OR
8 "Income " IS NULL OR
9 "Kidhome" IS NULL OR
10 "Teenhome" IS NULL OR
11 "Dt_customer" IS NULL OR
12 "Recency" IS NULL OR
13 "AmtLiq" IS NULL OR
14 "AmtVege" IS NULL OR
15 "AmtNonVeg" IS NULL OR
16 "AmtPes" IS NULL OR
17 "AmtChocolates" IS NULL OR
18 "AmtComm" IS NULL OR
19 "NumDeals" IS NULL OR
20 "NumWebBuy" IS NULL OR
21 "NumWalkinPur" IS NULL OR
22 "NumVisits" IS NULL OR
23 "Response" IS NULL OR
24 "Complain" IS NULL OR
25 "Country" IS NULL OR
26 "Count_success" IS NULL
```

Figure 18 - Null check in marketing_data

```
1 SELECT
2     "ID",
3     "Bulkmail_ad",
4     "Twitter_ad",
5     "Instagram_ad",
6     "Facebook_ad",
7     "Brochure_ad"
8 FROM public.ad_data
9 WHERE "ID" IS NULL OR
10     "Bulkmail_ad" IS NULL OR
11     "Twitter_ad" IS NULL OR
12     "Instagram_ad" IS NULL OR
13     "Facebook_ad" IS NULL OR
14     "Brochure_ad" IS NULL
```

Figure 19 - Null check in ad_data

```

1 SELECT "ID",
2        COUNT(*) as ordercount
3 FROM public.marketing_data
4 GROUP BY "ID"
5 HAVING COUNT(*) > 1
6

```

Figure 20 - Primary Key check in marketing_data

```

1 SELECT
2     "ID",
3     COUNT (*) AS ordercount
4 FROM public.ad_data
5 GROUP BY "ID"
6 HAVING COUNT(*) > 1;

```

Figure 21 - Primary Key check in ad_data

```

1 SELECT "ID", "Year_Birth", "Age", "Education", "Marital_Status", "Income ", "Kidhome",
2        "Teenhome", "Dt_customer", "Recency", "AmtLiq", "AmtVege", "AmtNonVeg", "AmtPes",
3        "AmtChocolates", "AmtComm", "NumDeals", "NumWebBuy", "NumWalkinPur", "NumVisits",
4        "Response", "Complain", "Country", "Count_success",
5        COUNT(*) as duplicate_count
6 FROM public.marketing_data
7 GROUP BY "ID", "Year_Birth", "Age", "Education", "Marital_Status", "Income ", "Kidhome",
8        "Teenhome", "Dt_customer", "Recency", "AmtLiq", "AmtVege", "AmtNonVeg", "AmtPes",
9        "AmtChocolates", "AmtComm", "NumDeals", "NumWebBuy", "NumWalkinPur", "NumVisits",
10       "Response", "Complain", "Country", "Count_success",
11 HAVING COUNT(*) > 1

```

Figure 22 - Duplicates check in marketing_data

```

1 SELECT
2     "ID",
3     "Bulkmail_ad",
4     "Twitter_ad",
5     "Instagram_ad",
6     "Facebook_ad",
7     "Brochure_ad",
8     COUNT (*) AS duplicate_count
9 FROM public.ad_data
10 GROUP BY "ID",
11         "Bulkmail_ad",
12         "Twitter_ad",
13         "Instagram_ad",
14         "Facebook_ad",
15         "Brochure_ad"
16 HAVING COUNT(*) > 1;

```

Figure 23 - Duplicates check in ad_data


```
1 SELECT "ID",  
2     COUNT(*) as ordercount  
3 FROM public.marketing_data  
4 GROUP BY "ID"  
5 HAVING COUNT(*) > 1
```

Figure 24 - Primary Key Duplicates check in marketing_data

```
1 SELECT  
2     "ID",  
3     COUNT (*) AS duplicate_count  
4 FROM public.ad_data  
5 GROUP BY "ID"  
6 HAVING COUNT(*) > 1;
```

Figure 25 - Primary Key Duplicates check in ad_data

Appendix E - SQL queries and further analysis

```

1 SELECT "Country",
2     SUM("AmtLiq") + SUM("AmtVege") + SUM("AmtNonVeg") + SUM("AmtPes") + SUM("AmtChocolates") + SUM("AmtComm") AS TotalSpend
3 FROM public.marketing_data
4 GROUP BY "Country"
5 ORDER BY TotalSpend DESC;

```

Figure 26 - SQL query related to Figure 6

```

1 SELECT "Country",
2     SUM("AmtLiq") AS TotalLiquid,
3     SUM("AmtVege") AS TotalVege,
4     SUM("AmtNonVeg") AS TotalNonVeg,
5     SUM("AmtPes") AS TotalPes,
6     SUM("AmtChocolates") AS TotalChocolates,
7     SUM("AmtComm") AS TotalComm,
8     SUM("AmtLiq") + SUM("AmtVege") + SUM("AmtNonVeg") + SUM("AmtPes") + SUM("AmtChocolates") + SUM("AmtComm") AS TotalSpend
9 FROM public.marketing_data
10 GROUP BY "Country"
11 ORDER BY TotalSpend DESC;

```

Figure 27 - SQL query related to Figure 7

```

1 SELECT
2     ROUND(SUM(CAST(a."Bulkmail_ad" AS INT)) * 100.0 / COUNT(*),2) AS bulkmail_leads,
3     ROUND(SUM(CAST(a."Twitter_ad" AS INT)) * 100.0 / COUNT(*),2) AS twitter_leads,
4     ROUND(SUM(CAST(a."Instagram_ad" AS INT)) * 100.0 / COUNT(*),2) AS instagram_leads,
5     ROUND(SUM(CAST(a."Facebook_ad" AS INT)) * 100.0 / COUNT(*),2) AS facebook_leads,
6     ROUND(SUM(CAST(a."Brochure_ad" AS INT)) * 100.0 / COUNT(*),2) AS brochure_leads
7 FROM ad_data a
8 INNER JOIN marketing_data m USING("ID")

```

Figure 28 - SQL query related to figure 8

```

1 SELECT
2     m."Country",
3     ROUND(SUM(CAST(a."Bulkmail_ad" AS INT)
4         + CAST(a."Twitter_ad" AS INT)
5         + CAST(a."Instagram_ad" AS INT)
6         + CAST(a."Facebook_ad" AS INT)
7         + CAST(a."Brochure_ad" AS INT)) * 100.0 / COUNT(*), 2) AS total_leads_percentage,
8     COUNT(m."ID") as customers
9 FROM ad_data a
10 INNER JOIN marketing_data m USING("ID")
11 GROUP BY m."Country"
12 ORDER by total_leads_percentage DESC;

```

Figure 29 - SQL query related to figure 9

```

1 SELECT
2     m."Country",
3     COUNT(m."ID") as customers,
4     ROUND(SUM(CAST(a."Bulkmail_ad" AS INT)) * 100.0 / COUNT(*),2) AS bulkmail_leads,
5     ROUND(SUM(CAST(a."Twitter_ad" AS INT)) * 100.0 / COUNT(*),2) AS twitter_leads,
6     ROUND(SUM(CAST(a."Instagram_ad" AS INT)) * 100.0 / COUNT(*),2) AS instagram_leads,
7     ROUND(SUM(CAST(a."Facebook_ad" AS INT)) * 100.0 / COUNT(*),2) AS facebook_leads,
8     ROUND(SUM(CAST(a."Brochure_ad" AS INT)) * 100.0 / COUNT(*),2) AS brochure_leads
9 FROM ad_data a
10 INNER JOIN marketing_data m USING("ID")
11 GROUP BY m."Country"
12 ORDER BY customers DESC;

```

Figure 30 - SQL query related to table 10

	Marital_Status character varying (30)	totalliquid bigint	totalvege bigint	totalnonveg bigint	totalpes bigint	totalchocolates bigint	totalcomm bigint	totalspend bigint
1	Married	251634	21471	131189	29662	22445	35990	492391
2	Civil Relationship	172198	14215	91154	21913	14719	23947	338146
3	Single	136891	12776	86058	18324	12617	20753	287419
4	Divorced	73075	6071	33621	7755	5956	10585	137063
5	Widow	27401	2228	13646	3631	2739	4144	53789

Figure 31 - Total sales by category per Marital Status

```

1 SELECT "Marital_Status",
2       SUM("AmtLiq") AS TotalLiquid,
3       SUM("AmtVege") AS TotalVege,
4       SUM("AmtNonVeg") AS TotalNonVeg,
5       SUM("AmtPes") AS TotalPes,
6       SUM("AmtChocolates") AS TotalChocolates,
7       SUM("AmtComm") AS TotalComm,
8       SUM("AmtLiq") + SUM("AmtVege") + SUM("AmtNonVeg") + SUM("AmtPes") + SUM("AmtChocolates") + SUM("AmtComm") AS TotalSpend
9 FROM public.marketing_data
10 GROUP BY "Marital_Status"
11 ORDER BY TotalSpend DESC;

```

Figure 32 - SQL query related to figure 31

	household_type text	totalliquid bigint	totalvege bigint	totalnonveg bigint	totalpes bigint	totalchocolates bigint	totalcomm bigint	totalspend bigint
1	No Kids or Teens	301583	32015	226003	47024	32753	39739	679117
2	With Kids or Teens	359616	24746	129665	34261	25723	55680	629691

Figure 33 - Total sales by having kids or teens or not

```

1 SELECT
2     CASE
3         WHEN "Kidhome" > 0 OR "Teenhome" > 0 THEN 'With Kids or Teens'
4         ELSE 'No Kids or Teens'
5     END AS Household_Type,
6     SUM("AmtLiq") AS TotalLiquid,
7     SUM("AmtVege") AS TotalVege,
8     SUM("AmtNonVeg") AS TotalNonVeg,
9     SUM("AmtPes") AS TotalPes,
10    SUM("AmtChocolates") AS TotalChocolates,
11    SUM("AmtComm") AS TotalComm,
12    SUM("AmtLiq") + SUM("AmtVege") + SUM("AmtNonVeg") + SUM("AmtPes") + SUM("AmtChocolates") + SUM("AmtComm") AS TotalSpend
13 FROM public.marketing_data
14 GROUP BY Household_Type
15 ORDER BY TotalSpend DESC;

```

Figure 34 - SQL query related to table 33

	bulkmail_leads bigint	twitter_leads bigint	instagram_leads bigint	facebook_leads bigint	brochure_leads bigint
1	159	162	159	140	29

Figure 35 - Effectiveness (number of leads) of each advertising method

```

1  SELECT
2      SUM(CAST(a."Bulkmail_ad" AS INT)) AS bulkmail_leads,
3      SUM(CAST(a."Twitter_ad" AS INT)) AS twitter_leads,
4      SUM(CAST(a."Instagram_ad" AS INT)) AS instagram_leads,
5      SUM(CAST(a."Facebook_ad" AS INT)) AS facebook_leads,
6      SUM(CAST(a."Brochure_ad" AS INT)) AS brochure_leads
7  FROM ad_data a
8  INNER JOIN marketing_data m USING("ID")

```

Figure 36 - SQL query related to table 35

	Marital_Status character varying (30) 🔒	bulkmail_leads bigint 🔒	twitter_leads bigint 🔒	instagram_leads bigint 🔒	facebook_leads bigint 🔒	brochure_leads bigint 🔒
1	Civil Relationship	37	42	43	32	12
2	Married	62	61	65	61	6
3	Widow	3	10	7	5	1
4	Single	39	31	32	31	5
5	Divorced	18	18	12	11	5

Figure 37 - Effectiveness (number of leads) of all types of leads by Marital Status

```

1  SELECT
2      m."Marital_Status",
3      SUM(CAST(a."Bulkmail_ad" AS INT)) AS bulkmail_leads,
4      SUM(CAST(a."Twitter_ad" AS INT)) AS twitter_leads,
5      SUM(CAST(a."Instagram_ad" AS INT)) AS instagram_leads,
6      SUM(CAST(a."Facebook_ad" AS INT)) AS facebook_leads,
7      SUM(CAST(a."Brochure_ad" AS INT)) AS brochure_leads
8  FROM ad_data a
9  INNER JOIN marketing_data m USING("ID")
10 GROUP BY m."Marital_Status";

```

Figure 38 - SQL query related to table 37

Appendix F – Tableau calculated fields

The following calculated fields were made in Tableau to help build the dashboards presented.

Age Bin Range - $\text{STR}(\text{INT}([\text{Age}]/5)*5) + "-" + \text{STR}(\text{INT}([\text{Age}]/5)*5+5-1)$

Age Group - IF [Age] >= 25 AND [Age] <= 34 THEN "25 - 34"
ELSEIF [Age] >= 35 AND [Age] <= 44 THEN "35 - 44"
ELSEIF [Age] >= 45 AND [Age] <= 54 THEN "45 - 54"
ELSEIF [Age] >= 55 AND [Age] <= 64 THEN "55 - 64"
ELSEIF [Age] >= 65 AND [Age] <= 74 THEN "65 - 74"
ELSEIF [Age] >= 75 AND [Age] <= 85 THEN "75 - 85"
END

Income Bin Range - "\$" + $\text{STR}(\text{INT}([\text{Income}]/10000)*10000) + "-" + \text{STR}(\text{INT}([\text{Income}]/10000)*10000+10000-1)$

Income Group - IF [Income] >= 0 AND [Income] <= 9999.99 THEN "0-9.9k"
ELSEIF [Income] >= 10000 AND [Income] <= 19999.99 THEN "10k - 19.9k"
ELSEIF [Income] >= 20000 AND [Income] <= 29999.99 THEN "20k - 29.9k"
ELSEIF [Income] >= 30000 AND [Income] <= 39999.99 THEN "30k - 39.9k"
ELSEIF [Income] >= 40000 AND [Income] <= 49999.99 THEN "40k - 49.9k"
ELSEIF [Income] >= 50000 AND [Income] <= 59999.99 THEN "50k - 59.9k"
ELSEIF [Income] >= 60000 AND [Income] <= 69999.99 THEN "60k - 69.9k"
ELSEIF [Income] >= 70000 AND [Income] <= 79999.99 THEN "70k - 79.9k"
ELSEIF [Income] >= 80000 AND [Income] <= 89999.99 THEN "80k - 89.9k"
ELSEIF [Income] >= 90000 AND [Income] <= 99999.99 THEN "90k - 99.9k"
ELSEIF [Income] >= 100000 THEN "100k or Above"
END

Type of Client - IF [Recency] <= 30 THEN "Current Customer"
ELSEIF [Recency] > 30 AND [Recency] <= 60 THEN "Irregular Customer"
ELSEIF [Recency] > 60 THEN "Lost Customer"
END

% of Alcoholic Drinks - $(\text{SUM}([\text{Amt Liq}]) / \text{SUM}([\text{Total Sales}])) * 100$

% of Chocolates - (SUM([Amt Chocolates]) / SUM([Total Sales])) * 100

% of Commodities - (SUM([Amt Comm]) / SUM([Total Sales])) * 100

% of Fish - (SUM([Amt Pes]) / SUM([Total Sales])) * 100

% of Meat - (SUM([Amt Non Veg]) / SUM([Total Sales])) * 100

% of Vegetables - (SUM([Amt Vege]) / SUM([Total Sales])) * 100

% Brochure - SUM([Brochure ad]) / COUNT([Brochure ad])

% Bulkmail - SUM([Bulkmail ad]) / COUNT([Bulkmail ad])

% Facebook - SUM([Facebook ad]) / COUNT([Facebook ad])

% Instagram - SUM([Instagram ad]) / COUNT([Instagram ad])

% Twitter - SUM([Twitter ad]) / COUNT([Twitter ad])

% Leads - SUM([Count success]) / COUNT([Count success])

Number of Purchases - ROUND([Num Walkin Pur] + [Num Web Buy], 2)

Sales per Purchase - SUM([Total Sales]) / SUM([Number of Purchases])

Total Leads - [Brochure ad] + [Bulkmail ad] + [Facebook ad] + [Instagram ad] + [Twitter ad]

Total Sales - [Amt Chocolates] + [Amt Comm] + [Amt Liq] + [Amt Non Veg] + [Amt Pes] + [Amt Vege]