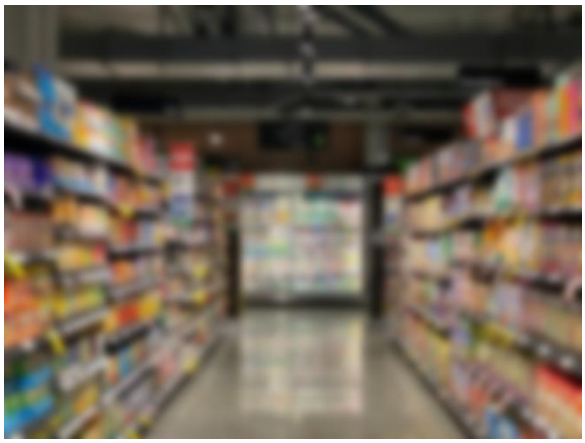

UNDERSTANDING YOUR CUSTOMERS

2Market Customer Demographic & Marketing Analysis

AUTHOR: DHRUVAL PATEL

DATE: 01 APR 2025

Table of Contents

Background/context of the business

2Market is a global supermarket operating both online and in-store. Their goal is to understand customer purchase patterns and behaviour. This includes analysis and presentation of their customer demographic range, the effectiveness of different advertising channels (online and print), and bestselling products across different customer demographics. This analysis would help 2Market target high-value customers with personalised offers, improve advertising effectiveness, adjust product offerings to match purchase trends, and re-engage inactive customers.

Analytical Approach

Accurate and meaningful analysis requires a structured approach to examining, cleaning and processing data sets in Excel before conducting deeper analysis using SQL. In Excel, I focused on identifying and correcting errors, standardizing formats and ensuring data consistency across columns.

2.1 Wrangling & Formatting of the Dataset

After importing the marketing_data.csv file into Excel, I applied the Text to Columns feature to correctly format integer columns and remove the currency symbol ($) from income, converting it to numeric data. Additionally, the registration date column contained mixed formats (DD/MM/YYYY and M/D/YY). Sorting the column alphabetically isolated the two formats, allowing me to restructure M/D/YY values into DD/MM/YYYY format by using Text to Columns. This standardised all registration dates to one format.

I used interquartile ranges (Figure 1, Appendix) to identify extreme outliers. By sorting the salary column, I found no negative values but identified an unrealistic salary entry of $666,666 which I removed. Salaries reaching $162,397 were retained as they were plausible. This approach was applied for all numerical columns and only erroneous entries such as ages over 120 were removed. Removing no more than 10 entries in total, ensuring that 2Market are provided analysis which is accurate and relevant to their business problems.

2.2 Initial Insights Developed from Pivot Tables

Once cleaned, Pivot Tables (Figure 2, Appendix) revealed key insights, such as liquor and non-veg products were the top-selling categories, and customer who responded to ads, spent more suggesting campaign effectiveness. I then imported the cleaned data into PostgreSQL, creating two tables: marketing_data and ad_data (Figures 3 & 4)

2.3 PostgreSQL Integration

To further analyse the data set, I utilised PostgreSQL by creating a database with two tables, marketing_data and ad_data (Figure 3 and 4, Appendix). I imported the CSV data into these tables and utilised Tableau's custom SQL feature for analysis. This includes filtering out rows with NULL values and pivoting data. In addition, PostgreSQL helped in pivoting purchase categories, streamlining visualisation for each customer ID.

Using custom SQL in Tableau, I created multiple tables which combined ad data with success counts for each customer ID, and pivoted purchase methods for each customer. The decision to pivot this data creates the possibility to display all information in a single graph, rather than separate graphs for each category enabling accurate filtering and readability. Figures 5, 6 and 7 in the appendix show the creation of these tables.

Dashboard Design and Development

The primary goal for my Tableau dashboard was to create a clear and accessible interface for 2Market, focussing on easily interpretable results and tweak-ability of demographics. The use of two dashboards, one for customer demographic and one for advertising effectiveness allows 2Market to focus on each area separately. KPI tiles displayed at the top immediately highlight the most relevant metrics of interest, such as the most and least popular product, along with most common age. As the business is interested in their customer demographics and which products sell best, KPIs work effectively in displaying this.

3.1 Visualisation Types

These were followed by visualisations which explore purchasing behaviour, demographic patterns and advertising effectiveness. Bar charts were used to compare categorical data such as purchasing methods and advertising channels. The use of heatmaps to indicate which products sell best according to age groups, allow 2Market to identify which demographic are most interested in their products. For my marketing dashboard, the use of a pie-chart indicating whether or not customers responded to their previous campaigns, show immediately whether or not campaigning has been successful to a particular demographic. The filters present on the dashboard apply to all visualisations shown which displays information relevant to the user's request. This involves adjusting many variables which serve information about particular customer groups.

3.2 Dynamic Filtering

The addition of an ‘apply filters’ button allows the user to adjust multiple filters before applying them, which saves time as filters usually dynamically update as each one is changed. This means desired criteria can be applied without being interrupted after changing each variable, creating a more streamlined experience.

The dashboard adopts a colourblind-friendly palette to enhance accessibility while using distinct colours to differentiate key metrics. This applies to the heatmap, where areas of higher spending can be identified instantly. Consistent font sizes, and alignment help avoid clutter by limiting each area to one clear insight, allowing users to focus on elements they wish.

3.3 PostgreSQL usage for Visualisations

By using Custom SQL in Tableau, I adjusted the marketing dashboard by replacing a static chart showing the different purchase types into a pivoted version that better highlighted each route of purchase (Figure 6, Appendix)

Overall, each design choice from layout, to colour, to interactivity was motivated by the aim of delivering accurate, streamlined, and business-relevant analysis.

Patterns, trends, and insights

4.1 Demographic of Customers

Analysis of 2Market’s customer data shows the most active customer group falls between the ages of 50-54, with the majority holding a graduate level education or higher. This suggests the business’ key customer base are educated and middle-aged. Spain accounted for the highest overall spending making it a region for future campaigning and promotion.

4.2 Product Engagement & Methods of Purchase

Liquor and non-veg items were amongst the top-selling products, whereas vegetables were the least popular. This indicates customers are more drawn to lifestyle-based purchases over essentials. Purchase channel analysis shows that in-store and online purchases are more successful than deal-based transactions, indicating a limited response to discount-led marketing. This trend was apparent over many different countries. 2Market can utilise this by increasing promotions and product range for these two categories.

4.2 Advertising Effectiveness

Advertising performance shows Twitter achieved the highest overall success rate, sparking most engagement from customers aged 40-44 and 65-69, indicating a strong responsiveness to marketing efforts within these age groups. Social media platforms show success in prompting customers to make subsequent purchases, whereas Bulkmail had performed well but only prompted customers once. As a result, efforts into social media marketing may return higher engagement.

Future exploration to understand which specific ad channel resulted in the highest spending, and product preferences by region would provide finer analysis into customer engagement. Additionally, knowing exactly which campaign were customers offered last, could identify if advertising channels are successful. Additionally, analysing why deal-based purchases underperform may improve future campaign design.

Appendix

Salary	
Q1	35303
Q3	68522
IQR	33219
Lower Limit	-14526
Upper Limit	118351

Figure 1 – Upper and lower limits for salary

Category	
Vegetables	$58,220
Chocolates	$59,558
Fish Products	$83,257
Commodities	$97,154
Non Veg	$362,813
Liquor	$675,213
Sum of Total Spent	$1,336,215

Response to Advertising	Average of Total Spent
0	$538
1	$986

Row Labels	Count of Count_success	Average of Total Spent
0	1749	$478
1	322	$914
2	81	$1,410
3	44	$1,706
4	11	$1,502
Grand Total	**2207**	**$605**

Age Group	Number of Individuals
18 to 34	61
35 to 55	1156
Over 55	990
Grand Total	**2207**

Age Group	Sum of NumDeals	Sum of NumWebBuy	Sum of NumWalkinPur
18 to 34	92	219	351
35 to 55	2596	4342	6254
Over 55	2409	4455	6238
Grand Total	**5097**	**9016**	**12843**

Figure 2 – Excel pivot tables showing initial analysis of 2Market's customers.

```sql
CREATE TABLE marketing_data (
    id BIGINT PRIMARY KEY,
    year_birth INTEGER,
    age INTEGER,
    Age_group VARCHAR(15),
    education VARCHAR(20),
    marital_status VARCHAR(10),
    income_$ INTEGER,
    kidhome INTEGER,
    Teenhome INTEGER,
    dt_customer_month CHAR(3),
    dt_customer_year INTEGER,
    recency INTEGER,
    amtliq INTEGER,
    amtvege INTEGER,
    amtnonveg INTEGER,
    amtpes INTEGER,
    amtchocolates INTEGER,
    amtcomm INTEGER,
    total_spent INTEGER,
    numdeals INTEGER,
    numwebbuy INTEGER,
    numwalkinpur INTEGER,
    numvisits INTEGER,
    response INTEGER,
    complain INTEGER,
    country CHAR(3),
    count_success INTEGER);
```

	id [PK] bigint	year_birth integer	age integer
1	2968	1944	81
2	5956	1949	76
3	5329	1950	75

```sql
CREATE TABLE ad_data (
    id BIGINT PRIMARY KEY,
    bulkmail_ad INTEGER,
    twitter_ad INTEGER,
    instagram_ad INTEGER,
    facebook_ad INTEGER,
    brochure_ad INTEGER);
```

Figure 3 – PostgreSQL query input to create marketing_data table

Edit Custom SQL

```sql
SELECT
    ad_data.id,
    Bulkmail_ad AS Bulkmail,
    Brochure_ad AS Brochure,
    Instagram_ad AS Instagram,
    Twitter_ad AS Twitter,
    Facebook_ad AS Facebook,
    count_success
FROM marketing_data
INNER JOIN ad_data ON marketing_data.id = ad_data.id
ORDER BY ad_data.id ASC
```

Figure 4 – PostgreSQL query producing ad_data table

# Custom S id	# Custom SQL Query3 bulkmail	# Custom SQL Query3 brochure	# Custom SQL Query3 instagram	# Custom SQL Query3 twitter	# Custom SQL Query3 facebook	# Custom SQL Query3 count_success
0	0	0	0	0	0	0
1	0	1	0	0	0	1
9	0	0	0	0	0	0
13	0	0	0	0	0	0
17	0	0	0	0	0	0

Figure 5 – Ad data table queried with count success for each customer

```
SELECT
        id,
        'Bulkmail' AS ad_type,
        bulkmail_ad AS amount
FROM ad_data
UNION ALL
SELECT
        id,
        'Twitter' AS ad_type,
        twitter_ad AS amount
FROM ad_data
UNION ALL
SELECT
        id,
        'Instagram' AS ad_type,
        instagram_ad AS amount
FROM ad_data
UNION ALL
SELECT
        id,
        'Facebook' AS ad_type,
        facebook_ad AS amount
FROM ad_data
UNION ALL
SELECT
        id,
        'Brochure' AS ad_type,
        brochure_ad AS amount
FROM ad_data
ORDER BY id ASC
```

# Custom SQL Query1 id (ad pivoted)	Abc Custom SQL Query1 ad_type	# Custom SQL Query1 amount
0	Bulkmail	0
0	Instagram	0
0	Twitter	0
0	Brochure	0
0	Facebook	0
1	Brochure	1
1	Bulkmail	0
1	Facebook	0
1	Twitter	0
1	Instagram	0

Figure 6 – Ad data table pivoted for each customer

```
SELECT
        id,
        'Deals' AS purchase_type,
        COALESCE(NumDeals, 0) AS amount
FROM marketing_data
UNION ALL
SELECT
        id,
        'Online' AS purchase_type,
        COALESCE(NumWebBuy, 0) AS amount
FROM marketing_data
UNION ALL
SELECT
        id,
        'In-store' AS purchase_type,
        COALESCE(NumWalkinPur, 0) AS amount
FROM marketing_data
UNION ALL
SELECT
        id,
        'Web Visits Per Month' AS purchase_type,
        COALESCE(NumVisits, 0) AS amount
FROM marketing_data
ORDER BY id ASC
```

# Custom S id (	Abc Custom SQL Query2 purchase_type	# Custom SQL Query2 amount (Custom SQL Que
0	Online	3
0	Deals	1
0	In-store	9
0	Web Visits Per Month	1
1	Web Visits Per Month	5
1	Deals	1
1	In-store	7
1	Online	7
9	Online	3
9	Web Visits Per Month	8

Figure 7 – Purchasing method pivoted by customer ID