2Market Technical Report by Will Burton

Contents

Background/context of the business

2Market is an international global supermarket chain with a dual presence, both online and in physical stores.

2Market hopes to understand consumer habits, demographics, and the factors driving sales across its international stores and boost sales through informed decision-making.

The business problem I have identified by using the Five Whys (*Appendix 1a*) which centres on suboptimal lead conversion due to a lack of understanding of customer demographics. Alternatively consider the Ishikawa (*Appendix 1b*) diagram, showing a hypothetical business problem which considers a 10% drop in 2Market's global customers, due to a lack of people, technology and processes to analyse their own data.

This report will consider the following questions:

- *"Who are the customers?"*
- *"How effective have current ad strategies been?"*
- *"How do customers interact with products?"*

Analytical approach

The data cleaning process for the given dataset involved: addressing issues related to accuracy, completeness, consistency, uniqueness, and timeliness.

- In terms of accuracy, outliers in income and age categories were identified and flagged for attention, acknowledging the possibility of legitimate but extreme values. (Appendix 2a)
- Incorrect or invalid data types were rectified, and columns were appropriately formatted. (Appendix 2b)
- The completeness check ensured no blanks or partial entries, and some columns were created. (Appendix 2c)

- Data consistency was achieved through precision, standardised structure, and case sensitivity adjustments.
- Uniqueness considerations involved merging certain marital status categories for more meaningful analyses. (Appendix 2d)
- No duplicates were found.
- Overall, the dataset is in good condition, but awareness of outliers and potentially out of date data. (Appendix 2e)

The process of SQL for the given dataset involved: examining metadata, establishing "2Market" database, formatting and verifying data for optimal data quality, and executing a comprehensive backup

- Initial Data Review of the metadata and data dictionary provided me with excellent data structure insights.
- Established the 2Market database and created “marketing_data” table and “ad_data” table. I designated “id” as the primary key.
- Renamed and adjusted data types of columns for user-friendly queries. (Appendix 3a)
- Opted for the "money" data type for currency, adjusting parameters for the dollar sign and aligned DateStyle with the American convention (M-D-Y).
- (Appendix 3b)
- Set up the "ad_data" database in a similar manner. (Appendix 3c)
- Successfully imported and meticulously verified cleaned data for accuracy against the original file. Executed a comprehensive backup, naming it "2MarketBackup".

Below are the outputs of queries asked by 2Market. Please find the SQL query in the respective appendices.

The total spend per country (Appendix 4a)

	country character varying (25)	total_spend money
1	SP	$659,557.00
2	SA	$211,071.00
3	CA	$167,403.00
4	AUS	$85,576.00
5	IND	$77,806.00
6	GER	$73,198.00
7	US	$67,546.00
8	ME	$3,122.00

The total spend per product per country (Appendix 4b)

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money	totalspend money
1	SP	$336,392.00	$28,288.00	$178,409.00	$40,153.00	$30,134.00	$46,181.00	$659,557.00
2	SA	$105,918.00	$8,937.00	$58,398.00	$13,670.00	$9,019.00	$15,129.00	$211,071.00
3	CA	$84,066.00	$7,681.00	$45,925.00	$9,980.00	$7,607.00	$12,144.00	$167,403.00
4	AUS	$42,752.00	$3,689.00	$22,328.00	$5,546.00	$4,129.00	$7,132.00	$85,576.00
5	IND	$36,236.00	$3,788.00	$23,729.00	$4,818.00	$3,221.00	$6,014.00	$77,806.00
6	GER	$36,776.00	$2,980.00	$20,272.00	$4,601.00	$2,801.00	$5,768.00	$73,198.00
7	US	$32,214.00	$3,034.00	$20,185.00	$4,411.00	$2,863.00	$4,839.00	$67,546.00
8	ME	$1,729.00	$8.00	$817.00	$226.00	$122.00	$220.00	$3,122.00

Which products are the most popular in each country? (Appendix 4c)

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money
1	SP	$336,392.00	$28,288.00	$178,409.00	$40,153.00	$30,134.00	$46,181.00

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money
1	SA	$105,918.00	$8,937.00	$58,398.00	$13,670.00	$9,019.00	$15,129.00

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money
1	CA	$84,066.00	$7,681.00	$45,925.00	$9,980.00	$7,607.00	$12,144.00

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money
1	AUS	$42,752.00	$3,689.00	$22,328.00	$5,546.00	$4,129.00	$7,132.00

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money
1	IND	$36,236.00	$3,788.00	$23,729.00	$4,818.00	$3,221.00	$6,014.00

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money
1	GER	$36,776.00	$2,980.00	$20,272.00	$4,601.00	$2,801.00	$5,768.00

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money
1	US	$32,214.00	$3,034.00	$20,185.00	$4,411.00	$2,863.00	$4,839.00

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money
1	ME	$1,729.00	$8.00	$817.00	$226.00	$122.00	$220.00

Which products are the most popular based on marital status? (Appendix 4d)

	marital_status character varying (25)	total_liquor money	total_non_vegetables money	total_commodities money	total_fish money	total_chocolates money	total_vegetables money
1	Together	$176,715.00	$95,374.00	$24,754.00	$22,383.00	$15,031.00	$14,612.00
2	Married	$256,976.00	$137,888.00	$36,719.00	$30,395.00	$22,926.00	$21,981.00
3	Widow	$27,902.00	$14,085.00	$4,245.00	$3,793.00	$2,878.00	$2,422.00
4	Divorced	$75,364.00	$34,848.00	$10,739.00	$8,130.00	$6,222.00	$6,363.00
5	Single	$137,217.00	$87,064.00	$20,397.00	$18,262.00	$12,751.00	$12,840.00

Which products are the most popular based on whether or not there are children or teens in the home? (Appendix 4e)

Kids at home only

	kid_home smallint	total_liquor money	total_non_vegetables money	total_commodities money	total_fish money	total_chocolates money	total_vegetables money
1	2	$3,312.00	$1,384.00	$776.00	$316.00	$179.00	$278.00
2	1	$92,550.00	$43,560.00	$19,613.00	$10,894.00	$7,639.00	$7,576.00
3	0	$580,221.00	$325,119.00	$77,038.00	$72,195.00	$52,078.00	$50,551.00

Teens at home only

	teen_home smallint	total_liquor money	total_non_vegetables money	total_commodities money	total_fish money	total_chocolates money	total_vegetables money
1	2	$18,083.00	$6,316.00	$2,317.00	$1,191.00	$786.00	$891.00
2	1	$307,060.00	$103,811.00	$43,365.00	$26,227.00	$20,628.00	$19,268.00
3	0	$350,940.00	$259,936.00	$51,745.00	$55,987.00	$38,482.00	$38,246.00

Which social media platform (Twitter, Instagram, or Facebook) is the most effective method of advertising in each country? (Appendix 4f)

	country character varying (25)	most_effective_ad_platform text
1	SP	instagram
2	CA	twitter
3	IND	twitter
4	AUS	instagram
5	US	facebook
6	ME	twitter
7	SA	instagram
8	GER	twitter

Which social media platform is the most effective method of advertising based on marital status? (Appendix 4g)

	marital_status character varying (25)	most_effective_ad_platform text
1	Together	instagram
2	Married	instagram
3	Widow	twitter
4	Single	twitter
5	Divorced	twitter

But if you include other ad platforms then we see new insights: (Appendix 4h)

	marital_status character varying (25)	most_effective_ad_platform text
1	Together	instagram
2	Married	instagram
3	Widow	twitter
4	Single	bulkmail
5	Divorced	bulkmail

Which social media platform(s) seem(s) to be the most effective per country? (In this case, assume that purchases were in some way influenced by lead conversions from a campaign). (Appendix 4I)

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money	totalSpend money	twitter_total bigint	instagram_total bigint	facebook_total bigint	total_lead_conversions bigint
1	SP	$336,392.00	$28,288.00	$178,409.00	$40,153.00	$30,134.00	$46,181.00	$659,557.00	87	89	76	252

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money	totalSpend money	twitter_total bigint	instagram_total bigint	facebook_total bigint	total_lead_conversions bigint
1	CA	$84,066.00	$7,681.00	$45,925.00	$9,980.00	$7,607.00	$12,144.00	$167,403.00	24	21	18	63

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money	totalSpend money	twitter_total bigint	instagram_total bigint	facebook_total bigint	total_lead_conversions bigint
1	SA	$105,918.00	$8,937.00	$58,398.00	$13,670.00	$9,019.00	$15,129.00	$211,071.00	20	21	20	61

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money	totalSpend money	twitter_total bigint	instagram_total bigint	facebook_total bigint	total_lead_conversions bigint
1	GER	$36,776.00	$2,980.00	$20,272.00	$4,601.00	$2,801.00	$5,768.00	$73,198.00	11	8	7	26

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money	totalSpend money	twitter_total bigint	instagram_total bigint	facebook_total bigint	total_lead_conversions bigint
1	IND	$36,236.00	$3,788.00	$23,729.00	$4,818.00	$3,221.00	$6,014.00	$77,806.00	10	6	7	23

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money	totalSpend money	twitter_total bigint	instagram_total bigint	facebook_total bigint	total_lead_conversions bigint
1	AUS	$42,752.00	$3,689.00	$22,328.00	$5,546.00	$4,129.00	$7,132.00	$85,576.00	6	12	7	25

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money	totalSpend money	twitter_total bigint	instagram_total bigint	facebook_total bigint	total_lead_conversions bigint
1	US	$32,214.00	$3,034.00	$20,185.00	$4,411.00	$2,863.00	$4,839.00	$67,546.00	6	5	7	18

	country character varying (25)	total_liquor money	total_vegetables money	total_non_vegtables money	total_fish money	total_chocolates money	total_commodities money	totalSpend money	twitter_total bigint	instagram_total bigint	facebook_total bigint	total_lead_conversions bigint
1	ME	$1,729.00	$8.00	$817.00	$226.00	$122.00	$220.00	$3,122.00	0	0	0	0

However, as with marital status, it is more insightful to open that query up to all ad platforms: (Appendix 4J for all the metrics in one place)

	country character varying (25)	most_effective_ad_platform text
1	SP	instagram
2	CA	twitter
3	IND	bulkmail
4	AUS	instagram
5	US	bulkmail
6	ME	bulkmail
7	SA	instagram
8	GER	twitter

(Appendix 4K for this table)

Dashboard design and development

This project is primarily designed towards 2Market's marketing team to provide comprehensive analysis featuring key demographic, ad and product metrics with an aim to provide data driven recommendations.

With the dashboard design I wanted to focus on the three core data segments:

"Who are the customers?" (Demographics dashboard)

"How do they interact with ads?" (ad data dashboard)

"How do they interact with products?" (Product dashboard)

This natural linear narrative allows for a seamless progression into several key data driven actionable insights that interconnect. This is highlighted by the demographic total chart that remains in each dashboard throughout the story to ground the user.

I wanted this chart to be entirely customisable with relevant metrics, so that at each stage of the journey, the user would have a top down, editable, view of the customer demographic relationship to both ad data and product.

Finally there is a real emphasis that each sheet on the respective dashboard allows for easy granular transitions from macro to micro by highlighting areas that you want to explore further. While intuitive, I am sure to demonstrate it in the presentation so that stakeholders with less technical experience can have full access to the data behind the insights.

This dashboard also aims to be accessible as possible following the Web Accessibility Initiative (WAI) guidelines:

- a reversed Orange-Blue Diverging colour palette
- Shapes instead of colour to aid colour deficient users
- Annotations on graphs to highlight key points
- Captions on all sheets for screen readers to aid users with vision impairment
- Clear and intuitive focus order

Patterns, trends, and insights

I also wanted to ensure that all insights would be useful to marketing and sales, by showing impact on the bottom line.

Demographic highlights :

- Focus on the cluster of 58-68 year olds (middle-income group), as there is a demonstrable correlation between higher median income and higher product spend. (Appendix 5A)

Ad Data Highlights:

- Twitter and Instagram perform well globally but recognise regional dynamics like Facebook opportunity in the US. (Appendix 5B)
- Reevaluate the effectiveness of brochures. (Appendix 5C)
- Don't rely solely on social media. Focus on bulk emails, particularly in South Africa, India, and the U.S. (Appendix 5D)
- Look into diversifying ad platforms (Appendix 5E)

Product Analysis highlights:

- Liquor and meat product categories lead the way globally by a significant margin. (Appendix 5F)
- Target customers without children segment, given their significantly higher spending patterns. (Appendix 5G)

Appendix 1 **Background Context**

1A . Let's apply the Five Whys structured approach to problem solving.

Why are lead conversions not optimal for the supermarket?

- Because our marketing efforts are not effectively targeting potential customers.

Why are our marketing efforts not effectively targeting potential customers?

- Because we are using a one-size-fits-all approach in our marketing strategy.

Why are we using a one-size-fits-all approach in our marketing strategy?

- Because we lack the data and insights to personalise our marketing efforts for different customer segments.

Why do we lack the data and insights to personalise our marketing efforts?

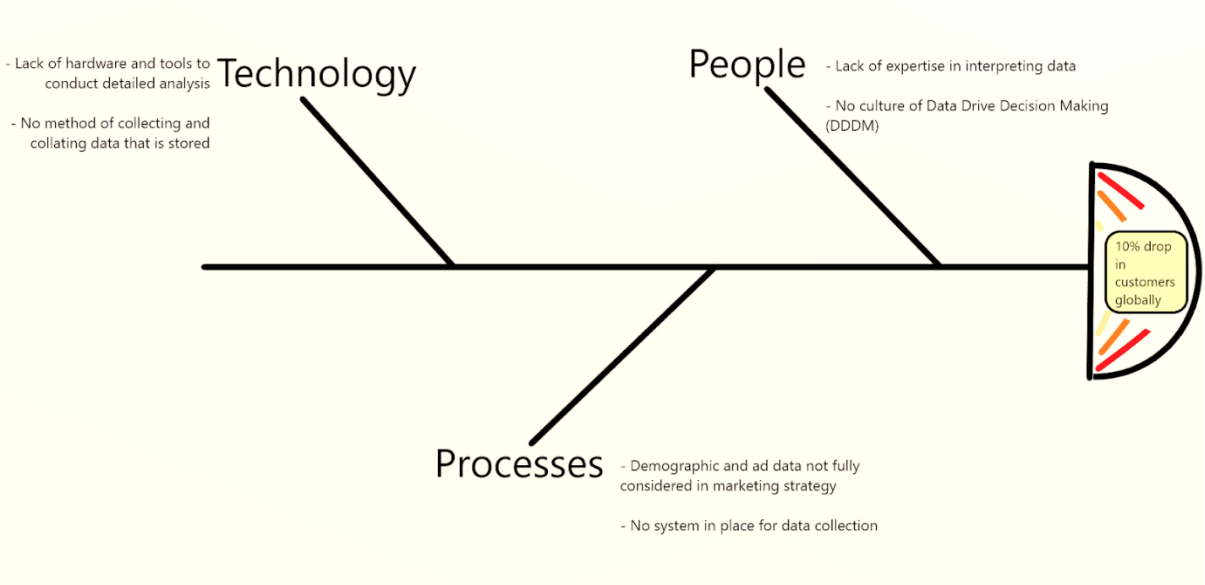
- Because our data collection methods are outdated, and we have not invested in customer segmentation analysis.

Why have we not invested in customer segmentation analysis and updated data collection methods?

- Because there is a lack of awareness within the organisation about the potential benefits of personalization and the available tools and techniques for data-driven marketing.

—

1B . An alternative problem-solving structure is the Ishikawa Fishbone Diagram.



This basic diagram hypothesises that 2Market's problem is a 10% drop in global customers. Then it seeks to find systematic cause and effect, demystifying the business problem and making it more accessible.

Beyond internal solutions such as systems upgrades, clear processes and staff training; one solution is to use external help to solve your issue. This is where our role as external analysts can come in.

Appendix 2 **Analytical Approach** – *Data Cleaning*

2A. *Values out of range:*

There are two outliers in the income category:

- **ID 6862** earns $1,730 annually. This is a low income, but it is plausible given that the individual may not work full-time or have a high-paying job.
- **ID 9432** earns $666,666 annually. This is a very high income and may be an outlier. However, it is also possible that the individual is a high-earner, such as a CEO or doctor.

There are three outliers in the Age category:

- **ID 11004**, born in 1894, with an age of 120 at the time of registration.
- **ID 1150**, born in 1900, with an age of 113 at the time of registration.
- **ID 7829**, born in 1901, with an age of 112 at the time of registration.
    - There is a 40 year gap before the next ID's continue in a expected, linear fashion. Ages at time of registration begin then at 72. It is possible that these are errors in the data, or that these individuals are centenarians, which is relatively rare.
- As such, all these outliers have been left in the data set, but they have been highlighted in red so that they can be easily identified.

2B. *Incorrect or invalid data types*:

- The data types of all columns have been and formatted using "Format Cells". The following have been corrected:
    - The "Dt_-customer" column was reformatted to date type M-D-Y format.
    - The "Income" column was reformatted to currency type to include the dollar symbol.

2C. *Missing values:* I create two columns I feel will be useful in the future:

- Age: I subtract "Year_Birth" from the "Dt_customer" column.
    The earliest date we have available to us is the Dt_Customer (Date of customer's registration with the company). As such this will be the data used for finding out the customer's youngest possible age.
- I include a "AmtTotal" column to detail the total amount spent by each customer.

2D. *Different words but with the same meaning:*

- The marital status categories “Absurd”, “Yolo”, and “Alone” have been amalgamated into a new category named "Inapplicable". This is because they will skew certain analyses.
- They are highlighted red with the other outliers but left in the data set as they still have other relevant data.

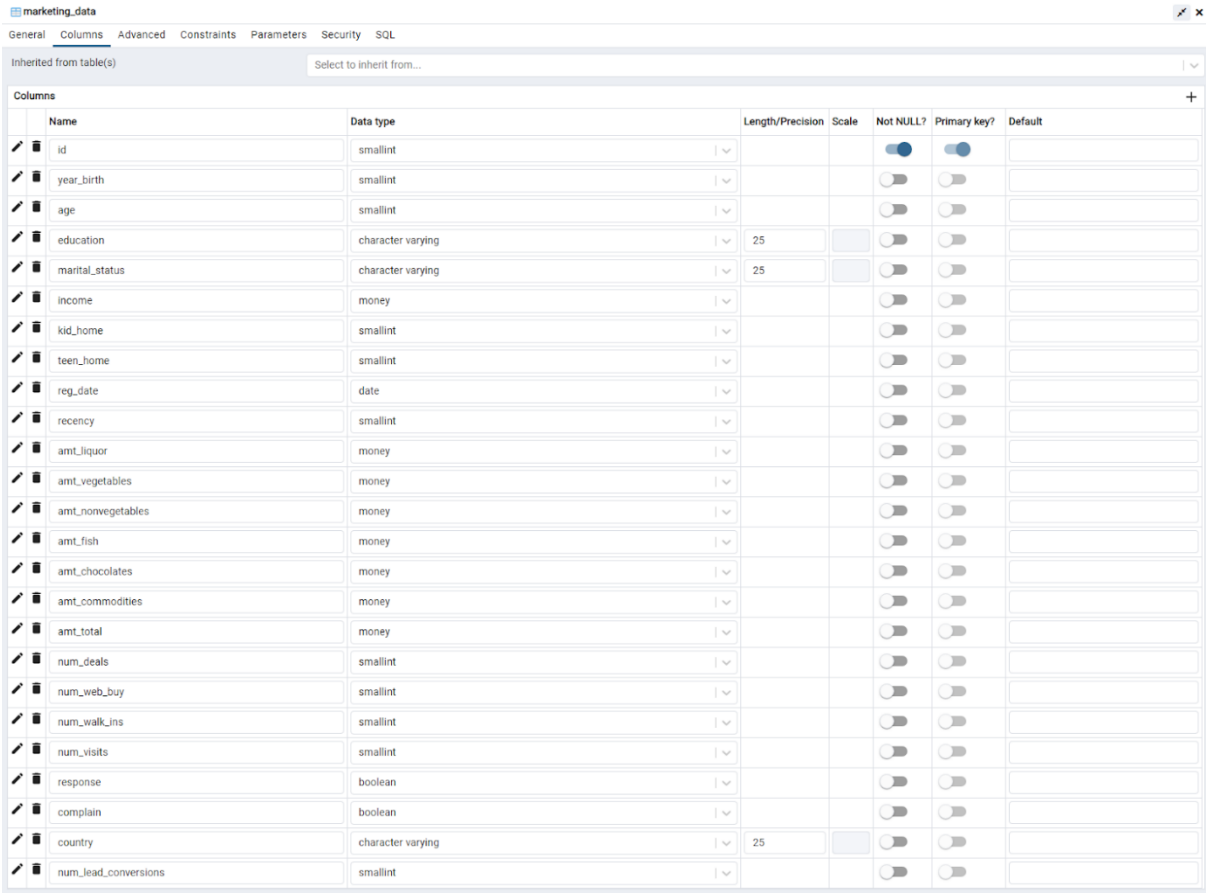
2E. *Dates being correct:* The dates in the data set are correct.

- Columns that are related to a point in time: It is difficult to judge the relevance of the dates with the information available. For example, the Dt_Customer column shows the date when the customer registered with the company, but it is not clear how recent this date is.
- This could prove problematic for creating current strategies.

Appendix 3 **Analytical Approach** – *SQL*

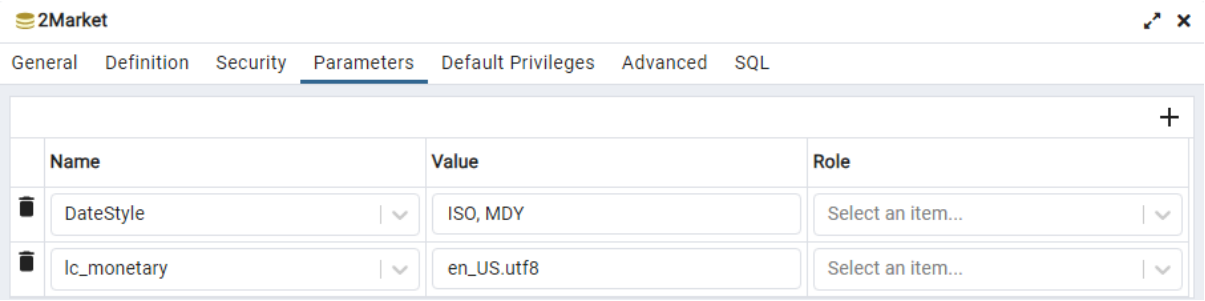
3A. Focus on keeping restraints as succinct and tidy as possible.

- Added columns as lower case in order to streamline future queries.
- Renamed columns for ease of use, see below:



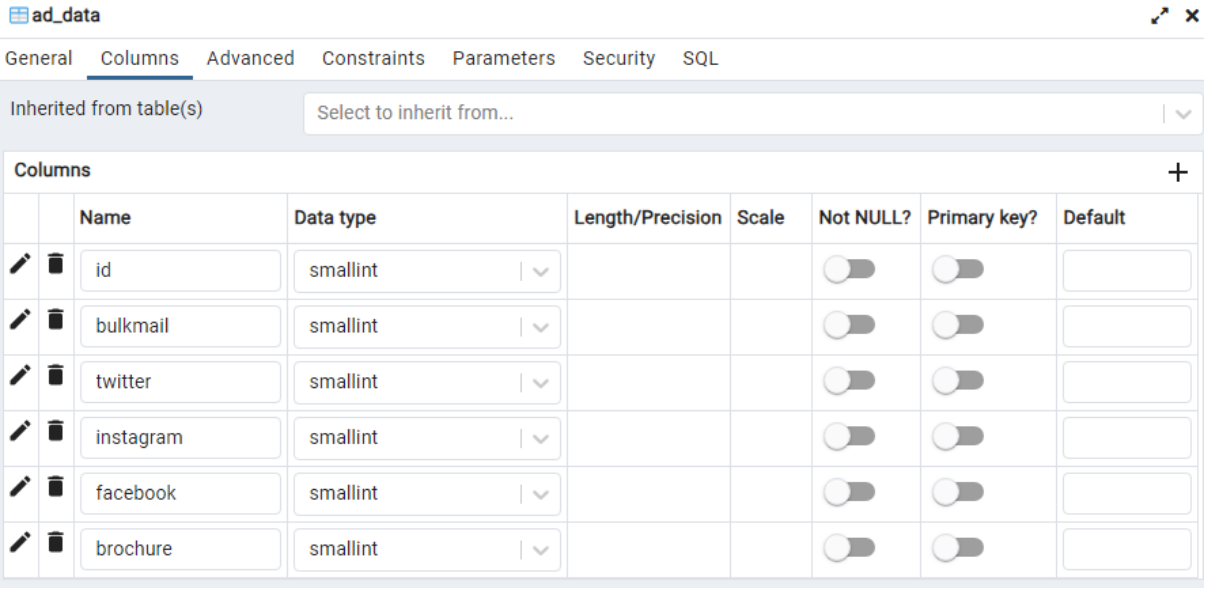
3B. As the currency in the data has no decimal places rounding accuracy is not an issue, so I decide to use the "money" data type.

- This requires a change in the database's parameters to accommodate the dollar sign.



*DateStyle is also changed to reflect the American convention of M-D-Y

3C. ad_type data types.



Appendix 4 **Analytical Approach** – *SQL Queries*

4A. **The total spend per country?**

```
SELECT "country",
SUM("amt_liquor" + "amt_vegetables" + "amt_nonvegetables" + "amt_fish" +
"amt_chocolates" + "amt_commodities")  AS Total_Spend
```

```
FROM public.marketing_data

GROUP BY "country"

ORDER BY "total_spend" DESC;
```

4B. **The total spend per product per country?**

```
SELECT "country",
SUM("amt_liquor") AS total_liquor,
SUM("amt_vegetables") AS total_vegetables,
SUM("amt_nonvegetables") AS total_non_vegtables,
SUM("amt_fish") AS total_fish,
SUM("amt_chocolates") AS total_chocolates,
SUM("amt_commodities") AS total_commodities,
SUM("amt_liquor" + "amt_vegetables" + "amt_nonvegetables" + "amt_fish" +
"amt_chocolates" + "amt_commodities") AS "totalspend"

FROM marketing_data

GROUP BY "country"

ORDER BY "totalspend" DESC;
```

4C. **Which products are the most popular in each country?**

```
SELECT "country",
SUM("amt_liquor") AS total_liquor,
SUM("amt_vegetables") AS total_vegetables,
SUM("amt_nonvegetables") AS total_non_vegtables,
SUM("amt_fish") AS total_fish,
SUM("amt_chocolates") AS total_chocolates,
SUM("amt_commodities") AS total_commodities

FROM  marketing_data

WHERE "country" = 'INPUT COUNTRY CODE'

GROUP BY  "country";
```

4D. **Which products are the most popular based on marital status?**

```
SELECT "marital_status",
SUM("amt_liquor") AS total_liquor,
SUM("amt_nonvegetables") AS total_non_vegetables,
SUM("amt_commodities") AS total_commodities,
SUM("amt_fish") AS total_fish,
SUM("amt_chocolates") AS total_chocolates,
SUM("amt_vegetables") AS total_vegetables

FROM marketing_data

WHERE "marital_status" IN ('Widow', 'Together', 'Married', 'Divorced', 'Single')

GROUP BY "marital_status";
```

* I choose to omit the outliers of “Absurd”, “Yolo” and “Alone” as the sample size is so small that it is distracting.

4E. **Which products are the most popular based on whether or not there are children or teens in the home?**

Kids at home only

```
SELECT "kid_home",
SUM("amt_liquor") AS total_liquor,
SUM("amt_nonvegetables") AS total_non_vegetables,
SUM("amt_commodities") AS total_commodities,
SUM("amt_fish") AS total_fish,
SUM("amt_chocolates") AS total_chocolates,
SUM("amt_vegetables") AS total_vegetables

FROM marketing_data

GROUP BY "kid_home"

ORDER BY "kid_home" DESC:
```

Teens at home only

```
SELECT "teen_home",
SUM("amt_liquor") AS total_liquor,
```

```
SUM("amt_nonvegetables") AS total_non_vegetables,
SUM("amt_commodities") AS total_commodities,
SUM("amt_fish") AS total_fish,
SUM("amt_chocolates") AS total_chocolates,
SUM("amt_vegetables") AS total_vegetables

FROM marketing_data

GROUP BY "teen_home"

ORDER BY "teen_home" DESC;
```

4F. **Which social media platform (Twitter, Instagram, or Facebook) is the most effective method of advertising in each country?**

*I decide that an inner join would be the most appropriate actions here using the the "id" column:

```
SELECT m."country",
CASE GREATEST(

        SUM(a."twitter"),

        SUM(a."instagram"),

        SUM(a."facebook")
   )
WHEN SUM(a."twitter") THEN 'twitter'
WHEN SUM(a."instagram") THEN 'instagram'
ELSE 'facebook'
END AS most_effective_ad_platform

FROM public."marketing_data"

INNER JOIN public."ad_data" a ON m."id" = a."id"

GROUP BY m."country"
```

4G. **Which social media platform is the most effective method of advertising based on marital status?**

*I am still choosing to exclude the outliers of “Absurd”, “Yolo” and “Alone”

```
SELECT m."marital_status",
CASE GREATEST(

        SUM(a."twitter"),

        SUM(a."instagram"),

        SUM(a."facebook")
         )

WHEN SUM(a."twitter") THEN 'twitter'
WHEN SUM(a."instagram") THEN 'instagram'
ELSE 'facebook'
END AS most_effective_ad_platform

FROM public."marketing_data" m

INNER JOIN public."ad_data" a ON m."id" = a."id"

WHERE m."marital_status" IN ('Widow', 'Together', 'Married', 'Divorced', 'Single')

GROUP BY m."marital_status"
```

4H. **Which *ad platform* is the most effective method of advertising based on marital status?**

```
SELECT m."marital_status",
CASE GREATEST(
   SUM(a."twitter"),

   SUM(a."instagram"),

   SUM(a."facebook"),

   SUM(a."bulkmail"),
```

```
    SUM(a."brochure")
    )

WHEN SUM(a."twitter") THEN 'twitter'
WHEN SUM(a."instagram") THEN 'instagram'
WHEN SUM(a."bulkmail") THEN 'bulkmail'
WHEN SUM(a."brochure") THEN 'brochure'
ELSE 'facebook'
END AS most_effective_ad_platform

FROM public."marketing_data" m

INNER JOIN public."ad_data" a ON m."id" = a."id"

WHERE m."marital_status" IN ('Widow', 'Together', 'Married', 'Divorced', 'Single')

GROUP BY m."marital_status";
```

4I. **Which social media platform(s) seem(s) to be the most effective per country?**

```
SELECT m."country",
SUM("amt_liquor") AS total_liquor,
SUM("amt_vegetables") AS total_vegetables,
SUM("amt_nonvegetables") AS total_non_vegtables,
SUM("amt_fish") AS total_fish,
SUM("amt_chocolates") AS total_chocolates,
SUM("amt_commodities") AS total_commodities,
SUM("amt_liquor" + "amt_vegetables" + "amt_nonvegetables" + "amt_fish" +
"amt_chocolates" + "amt_commodities") AS "totalSpend",

SUM("twitter") AS "twitter_total",
SUM("instagram") AS "instagram_total",
SUM("facebook") AS "facebook_total",
SUM("twitter" + "instagram" + "facebook") AS "total_lead_conversions"

FROM public."marketing_data" m

LEFT JOIN public."ad_data" a ON m."id" = a."id"

WHERE m."country" = ---'COUNTRY CODE---
```

```
GROUP BY m."country"

ORDER BY "total_lead_conversions" DESC
```

4J. **Which "*ad platform(s)*" seem(s) to be the most effective per country?**

```
SELECT m."country",
SUM("amt_liquor") AS total_liquor,
SUM("amt_vegetables") AS total_vegetables,
SUM("amt_nonvegetables") AS total_non_vegtables,
SUM("amt_fish") AS total_fish,
SUM("amt_chocolates") AS total_chocolates,
SUM("amt_commodities") AS total_commodities,
SUM("amt_liquor" + "amt_vegetables" + "amt_nonvegetables" + "amt_fish" +
"amt_chocolates" + "amt_commodities") AS "totalSpend",

SUM("bulkmail") AS "bulkmail_total",
SUM("twitter") AS "twitter_total",
SUM("instagram") AS "instagram_total",
SUM("facebook") AS "facebook_total",
SUM("brochure") AS "brochure_total",
SUM("bulkmail" + "twitter" + "instagram" + "facebook" + "brochure") AS
"total_lead_conversions"

FROM public."marketing_data" m

LEFT JOIN public."ad_data" a ON m."id" = a."id"

GROUP BY m."country"

ORDER BY "total_lead_conversions" DESC
```

4K. **Which "*ad platform(s)*" seem(s) to be the most effective per country? (table)** SELECT m."country",

```
CASE GREATEST(
SUM(a."twitter"),
SUM(a."instagram"),
SUM(a."facebook"),
SUM(a."brochure"),
SUM(a."bulkmail")
    )
WHEN SUM(a."twitter") THEN 'twitter'
WHEN SUM(a."instagram") THEN 'instagram'
WHEN SUM(a."facebook") THEN 'facebook'
WHEN SUM(a."bulkmail") THEN 'bulkmail'
ELSE 'brochure'
END AS most_effective_ad_platform
FROM public."marketing_data" m
INNER JOIN public."ad_data" a ON m."id" = a."id"
GROUP BY m."country"
```

1

Appendix 5 **Patterns, trends, and insights**

5A . "Average Income by Customer Age" - sheet in Excel

Age Bands	Average of Income
26-30	$56,266
31-35	$46,880
36-40	$44,920
41-45	$49,871
46-50	$49,372
51-55	$50,931
56-60	$54,602
61-65	$56,234
66-70	$57,487
71-75	$58,615
76-80	$63,822
81-85	$72,084

5B . "Ad Conversion_Rates" - sheet in Excel

Conversion Rates	Country	Bulkmail_total	Twitter_total	Instagram_total	Facebook_total	Brochure_total	Total_lead_conversions
	SP	24%	25%	25%	22%	5%	100%
	CA	21%	28%	24%	21%	7%	100%
	SA	24%	23%	24%	23%	5%	100%
	GER	26%	29%	21%	18%	5%	100%
	IND	34%	26%	16%	18%	5%	100%
	AUS	26%	18%	35%	21%	0%	100%
	US	31%	23%	19%	27%	0%	100%
	ME	100%	0%	0%	0%	0%	100%

5C . "Ad Conversion_Rates" - sheet in Excel

Conversion Rates	Country	Bulkmail_total	Twitter_total	Instagram_total	Facebook_total	Brochure_total	Total_lead_conversions
	SP	24%	25%	25%	22%	5%	100%
	CA	21%	28%	24%	21%	7%	100%
	SA	24%	23%	24%	23%	5%	100%
	GER	26%	29%	21%	18%	5%	100%
	IND	34%	26%	16%	18%	5%	100%
	AUS	26%	18%	35%	21%	0%	100%
	US	31%	23%	19%	27%	0%	100%
	ME	100%	0%	0%	0%	0%	100%

5D . "Ad Conversion_Rates" - sheet in Excel

Conversion Rates	Country	Bulkmail_total	Twitter_total	Instagram_total	Facebook_total	Brochure_total	Total_lead_conversions
	SP	24%	25%	25%	22%	5%	100%
	CA	21%	28%	24%	21%	7%	100%
	SA	24%	23%	24%	23%	5%	100%
	GER	26%	29%	21%	18%	5%	100%
	IND	34%	26%	16%	18%	5%	100%
	AUS	26%	18%	35%	21%	0%	100%
	US	31%	23%	19%	27%	0%	100%
	ME	100%	0%	0%	0%	0%	100%

- Email marketing remains effective, with 51% of marketers considering it the most effective channel (HubSpot, 2023[1]).

5E . **Diversifying Platforms:**

- TikTok generated substantial revenue in Q4 2022, outperforming Facebook, Instagram, Twitter, and Snapchat combined. (FORBES, 2023[2])

Recommendation:

- Continuous exploration of emerging platforms is essential, considering the dynamic landscape of digital advertising.
- Explore potential ad platforms beyond current ones, with particular emphasis on emerging platforms like TikTok.

5F . "Product spend by Country" – sheet in Excel

Row Labels	Sum of AmtVege	Sum of AmtChocolates	Sum of AmtPes	Sum of AmtComm	Sum of AmtNonVeg	Sum of AmtLiq
AUS	3,689	4,129	5,546	7,132	22,328	42,752
CA	7,681	7,607	9,980	12,144	45,925	84,066
GER	2,980	2,801	4,601	5,768	20,272	36,776
IND	3,788	3,221	4,818	6,014	23,729	36,236
ME	8	122	226	220	817	1,729
SA	8,937	9,019	13,670	15,129	58,398	105,918
SP	28,288	30,134	40,153	46,181	178,409	336,392
US	3,034	2,863	4,411	4,839	20,185	32,214
Grand Total	**58,405**	**59,896**	**83,405**	**97,427**	**370,063**	**676,083**

5G. "Kids at Home vs Total Spend" - sheet in Excel

Kidhome	0

Row Labels	Count of Kidhome	Sum of AmtTotal
Divorced	140	124732
Married	488	423835
Single	266	253054
Together	327	300049
Widow	58	52299
Grand Total	**1279**	**£ 1,153,969.00**

Kidhome	1

Row La	Count of Kidhome	Sum of AmtTotal
Divorced	88	16772
Married	349	79686
Single	195	34380
Together	234	47198
Widow	18	3026
Grand Tot	**884**	**£ 181,062.00**

Kidhome	2

Row Labels	Count of Kidhome	Sum of AmtTotal
Divorced	4	162
Married	20	3364
Single	10	1097
Together	12	1622
Grand Total	**46**	**£ 6,245.00**

KIDS

Row Labels	Sum of AmtLiq	Sum of AmtVege	Sum of AmtNonVeg	Sum of AmtPes	Sum of AmtChocolates	Sum of AmtComm	Count of ID	Total Spend	Total Spend per head
0	$ 580,221	$ 50,551	$ 325,119	$ 72,195	$ 52,078	$ 77,038	1,283	$ 1,157,202	$ 901.95
1	$ 92,550	$ 7,576	$ 43,560	$ 10,894	$ 7,639	$ 19,613	887	$ 181,832	$ 141.72
2	$ 3,312	$ 278	$ 1,384	$ 316	$ 179	$ 776	46	$ 6,245	$ 4.87
(blank)									
Grand Total	**$ 676,083**	**$ 58,405**	**$ 370,063**	**$ 83,405**	**$ 59,896**	**$ 97,427**	**2,216**		

[1] Perricone, C. (2023) *Email marketing: The ultimate guide (+ expert tips)*, *HubSpot Blog*. Available at: https://blog.hubspot.com/marketing/email-marketing-guide (Accessed: 16 October 2023).

[2] Belle Wong, J.D. (2023) *Top social media statistics and trends of 2023*, *Forbes*. Available at: https://www.forbes.com/advisor/business/social-media-statistics/#source (Accessed: 16 October 2023).