2MARKET PROJECT
DATA ANALYSIS REPORT

Executive Summary:

2Market is a consumer-focused global supermarket selling products online and in-store seeking to enhance its product and marketing strategies by gaining insight into its customer's purchasing behaviors. With its global reach, 2Market aims to investigate their customer demographics, the advertising channels that seem to be the most effective, and which products resonate most with specific segments. Through an analysis of its customer's demographic composition, advertising exposure, and product purchasing behavior, the business aims to find applicable knowledge that can inform future business campaigns, product recommendations, and regional strategies.

Analytical Approach:

In order to address the business questions, a structured analytical workflow was employed using Excel and SQL before finalizing data to be visualized via Tableau.

First, the customer demographic data was cleaned to handle missing values, fix formatting issues, and standardize entries as needed. Initial data analysis was then performed to add an *Age* column and aggregate the average age of customers belonging to each type of marital status and as well as their income brackets. This was particularly helpful in identifying and visualizing the relationships in the customer demographic to better understand the target audience.

SQL queries were then used to understand the structure of the data and interpret results to answer which products seem to sell best and whether or not that varies by customer demographics. The *total spending per country, total spending per product per country*, the *most popular product in each country*, and *the most popular product based on marital status* were all queried in SQL to generate results that helped gain further insight on the current business scheme of 2Market. Consequently, SQL was also utilized to aid in understanding which is the most effective method of advertising in each country. Appropriate SQL queries to LEFT join the customer demographic data and the advertising data were executed to aggregate functions to perform the necessary calculations which were in turn filtered and ranked by *country*, *marital status*, and *platform* to derive insights. All queries used aliases for clarity and the COALESCE function ensured that NULL *"ad"*

conversion values were counted as zero. Thus, this SQL-driven analysis allowed for a detailed, structured foundation to inform effective dashboard creation.

Dashboard Design

The Tableau dashboard was designed with clarity, interactivity, and stakeholder usability in consideration. It includes three main sections, with each aiming to provide analytical overview and answers to the business questions.

The first section, *Customer Demographics* comprises of a stacked and grouped bar chart to show distribution by country, marital status, and income ranges. The color-coded bars allow users to immediately spot which segments dominate the customer base. Filters for each distribution was also available to slice data and explore demographic and region-specific patterns.

The second section, Advertising Channel Effectiveness, summarizes the total reach of each advertising channel including Facebook, Twitter, Instagram, Brochure and Bulkmail using a simple yet comprehensive bar chart. It highlights which platforms reached the most customers in each income range to assess whether high-value customers are being effectively targeted. This color-coded advertising channel chart can also be segmented per country to allow the stakeholders to see their client's current demographic locations, and which advertising platform is the best one to use to target each client segment.

The final section of the dashboard comprises of a series of bar charts that visualizes spending on three major categories – *Customer's Product Spending per Income Group, Customer's Product Spending per Country,* and *Customer's Product Spending per Marital Status*. These charts show a breakdown of the data that allows the stakeholders to see which products are more profitable, which income groups have the maximum buying trends, which countries are more commercial, and which buyers have the highest purchasing power. Filters for custom analysis were also included to highlight tables or stacked bars and show comparative values for further analysis. Throughout the dashboard, calculated fields such as *Income Range* were used to support grouping and filtering and mark labels were added to improve readability and easier comprehension.

The design choices for the dashboards were intentional to show optical coherence. Different colors were used to highlight and segregate information clearly and interactivity was introduced via filter to empower the user to explore insights without needing much technical proficiency.

Trends and Insights

The analysis revealed a number of meaningful trends across customer demographics, advertising effectiveness, and product purchasing behavior, each of which offers important direction for 2Market's future marketing and sales strategies.

Demographic analysis showed that a large proportion of the customer base is *Married* with the majority residing in *Spain* and *Saudi Arabia*. Most of these customers fall between the "Mid" to "High" income brackets, thus representing a particularly valuable customer segment due to their higher levels of disposable income and willingness to spend. Thus, this profile – **Married, high earning individuals from Spain and Saudi Arabia** - represents a major pool of the customer base 2Market needs to continuously target in their sales and promotion schemes.

Advertising data indicated that **Instagram** and **Facebook** are the most effective channels in targeting these **high-earning** individuals across all target countries. It is also worth noting that on **lower-income** and **mid-income** targets, **Bulkmail** seems to be the most effective advertisement tool, which should be a viable consideration when deciding on which advertising channel to employ in each segmented income bracket of 2Market's clientele.

On the product side, **Liquor** and **Non-Vegetarian** items had the highest customer spending values, especially among **High-income** and **Married** customers. Consequently, the countries where these products are also most profitable are in **Spain** and **Saudi Arabia**, showing a clear view of the primary demographics of 2Market's products. **Commodities** and **Chocolates** also performed well but showed more fluctuation depending on income levels.

These insights point towards a strong business case for personalized marketing campaigns that target **high-income married** customers in **Spain** and **Saudi Arabia**, emphasizing products like **Liquor** and **Non-Vegetarian** food. Through an alignment of advertising efforts with these key demographics and preferred platforms, 2Market can improve its customer engagement and overall marketing return on investment.

APPENDICES

(SQL QUERIES)

```sql
CREATE TABLE marketing_data (
    id INT PRIMARY KEY,
    year_birth INT,
    education VARCHAR(50),
    marital_status VARCHAR(50),
    income NUMERIC,
    kidhome INT,
    teenhome INT,
    dt_customer DATE,
    recency INT,
    amtliq NUMERIC,
    amtvege NUMERIC,
    amtnonveg NUMERIC,
    amtpes NUMERIC,
    amtchocolates NUMERIC,
    amtcomm NUMERIC,
    numdeals INT,
    numwebbuy INT,
    numwalkinpur INT,
    numvisits INT,
    response INT,
    complain INT,
    country VARCHAR(50),
    count_success INT,
    age INT
);
```

Query for creating the marketing_data Table in SQL

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

Query for creating the ad_data Table in SQL

```sql
SELECT
    country,
    SUM(amtliq + amtvege + amtnonveg + amtpes + amtchocolates + amtcomm) AS total_spend
FROM
    marketing_data
GROUP BY
    country
ORDER BY
    total_spend DESC;
```

Query for Total Spend per Country in SQL

```sql
SELECT country, 'Liquor' AS product, SUM(amtliq) AS total_spend FROM marketing_data GROUP BY country
UNION ALL
SELECT country, 'Vegetables', SUM(amtvege) FROM marketing_data GROUP BY country
UNION ALL
SELECT country, 'Meat', SUM(amtnonveg) FROM marketing_data GROUP BY country
UNION ALL
SELECT country, 'Fish', SUM(amtpes) FROM marketing_data GROUP BY country
UNION ALL
SELECT country, 'Chocolates', SUM(amtchocolates) FROM marketing_data GROUP BY country
UNION ALL
SELECT country, 'Communication', SUM(amtcomm) FROM marketing_data GROUP BY country
ORDER BY country, total_spend DESC;
```

Query for Total Spend per Product per Country in SQL

```sql
WITH product_spend AS (
    SELECT country, 'Liquor' AS product, SUM(amtliq) AS total_spend FROM marketing_data GROUP BY country
    UNION ALL
    SELECT country, 'Vegetables', SUM(amtvege) FROM marketing_data GROUP BY country
    UNION ALL
    SELECT country, 'Meat', SUM(amtnonveg) FROM marketing_data GROUP BY country
    UNION ALL
    SELECT country, 'Fish', SUM(amtpes) FROM marketing_data GROUP BY country
    UNION ALL
    SELECT country, 'Chocolates', SUM(amtchocolates) FROM marketing_data GROUP BY country
    UNION ALL
    SELECT country, 'Communication', SUM(amtcomm) FROM marketing_data GROUP BY country
),
ranked_products AS (
    SELECT *,
            RANK() OVER (PARTITION BY country ORDER BY total_spend DESC) AS rank
    FROM product_spend
)
SELECT * FROM ranked_products WHERE rank = 1;
```

Query for Most Popular Products in Each Country in SQL

```sql
WITH product_spend AS (
    SELECT marital_status, 'Liquor' AS product, SUM(amtliq) AS total_spend FROM marketing_data GROUP BY marital_status
    UNION ALL
    SELECT marital_status, 'Vegetables', SUM(amtvege) FROM marketing_data GROUP BY marital_status
    UNION ALL
    SELECT marital_status, 'Meat', SUM(amtnonveg) FROM marketing_data GROUP BY marital_status
    UNION ALL
    SELECT marital_status, 'Fish', SUM(amtpes) FROM marketing_data GROUP BY marital_status
    UNION ALL
    SELECT marital_status, 'Chocolates', SUM(amtchocolates) FROM marketing_data GROUP BY marital_status
    UNION ALL
    SELECT marital_status, 'Communication', SUM(amtcomm) FROM marketing_data GROUP BY marital_status
),
ranked_products AS (
    SELECT *,
            RANK() OVER (PARTITION BY marital_status ORDER BY total_spend DESC) AS rank
    FROM product_spend
)
SELECT * FROM ranked_products WHERE rank = 1;
```

Query for Most Popular Products by Marital Status in SQL

```
CREATE OR REPLACE VIEW joined_data AS
SELECT
    m.id,
    m.country,
    m.marital_status,
    m.kidhome,
    m.teenhome,
    m.amtliq,
    m.amtvege,
    m.amtnonveg,
    m.amtpes,
    m.amtchocolates,
    m.amtcomm,
    COALESCE(a.twitter_ad, 0) AS twitter_ad,
    COALESCE(a.instagram_ad, 0) AS instagram_ad,
    COALESCE(a.facebook_ad, 0) AS facebook_ad
FROM
    marketing_data m
LEFT JOIN
    ad_data a ON m.id = a.id;

SELECT * FROM joined_data LIMIT 10;
```

*Query for LEFT JOIN marketing_data
and ad_data in SQL*

```
SELECT
    country,
    SUM(twitter_ad) AS twitter_conversions,
    SUM(instagram_ad) AS instagram_conversions,
    SUM(facebook_ad) AS facebook_conversions
FROM
    joined_data
GROUP BY
    country
ORDER BY
    (SUM(twitter_ad) + SUM(instagram_ad) + SUM(facebook_ad)) DESC;
```

Query for Country with the most conversions per platform in SQL

```
SELECT
    marital_status,
    SUM(twitter_ad) AS twitter_conversions,
    SUM(instagram_ad) AS instagram_conversions,
    SUM(facebook_ad) AS facebook_conversions
FROM
    joined_data
GROUP BY
    marital_status
ORDER BY
    (SUM(twitter_ad) + SUM(instagram_ad) + SUM(facebook_ad)) DESC;
```

Query for which platform is most effective by Marital Status in SQL

```sql
SELECT
    country,
    'Twitter' AS platform,
    SUM(amtliq + amtvege + amtnonveg + amtpes + amtchocolates + amtcomm) AS total_spend
FROM
    joined_data
WHERE
    twitter_ad = 1
GROUP BY
    country

UNION ALL

SELECT
    country,
    'Instagram',
    SUM(amtliq + amtvege + amtnonveg + amtpes + amtchocolates + amtcomm)
FROM
    joined_data
WHERE
    instagram_ad = 1
GROUP BY
    country

UNION ALL


SELECT
    country,
    'Facebook',
    SUM(amtliq + amtvege + amtnonveg + amtpes + amtchocolates + amtcomm)
FROM
    joined_data
WHERE
    facebook_ad = 1
GROUP BY
    country

ORDER BY
    total_spend DESC;


SELECT
    country,
    SUM(twitter_ad) AS twitter_conversions,
    SUM(instagram_ad) AS instagram_conversions,
    SUM(facebook_ad) AS facebook_conversions
FROM
    joined_data
GROUP BY
    country
ORDER BY
    country;
```

Query for which platform(s) are per country in terms of total spending in SQL