REPORT

2MARKET

DATA ANALYSIS.

Background/context of the business

The business that the report talks about is a global supermarket chain and the analysis done in the following sections using EXCEL,SQL and TABLEAU goes on to highlight certain key points that will prove helpful for the business to understand their customer base. The specific analysis that is done in this assignment concentrates on some major business problems which are as follows:

1. The demographic of their major customer group.
2. The advertising channels which are the most effective for their business
3. The best selling products and the demographics of customers buying them.
4. The buying behaviour of customer that varies according to each country
5. What sort of marketing efforts can be taken to improve their business.

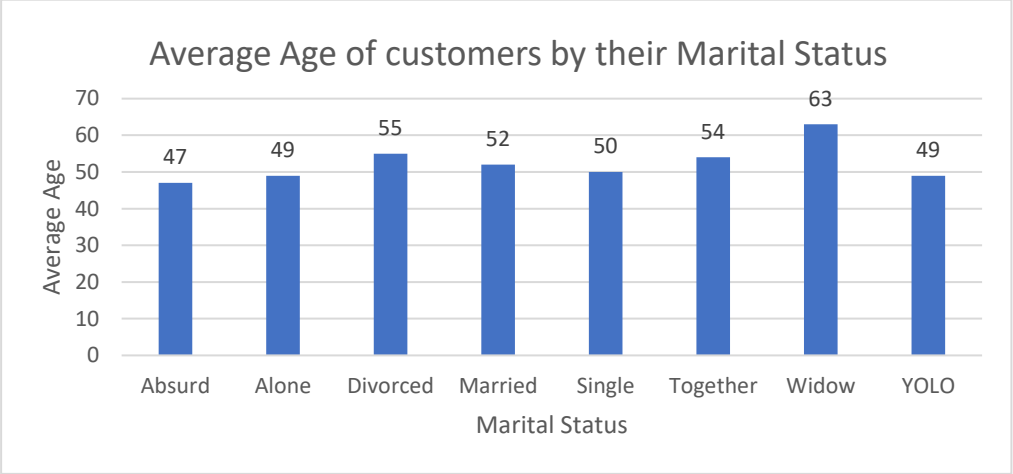
For the purpose of understanding the above dimensions of the business, analysis was done and the process of each of them have been given below along with the trends and patterns identified.

Analytical approach

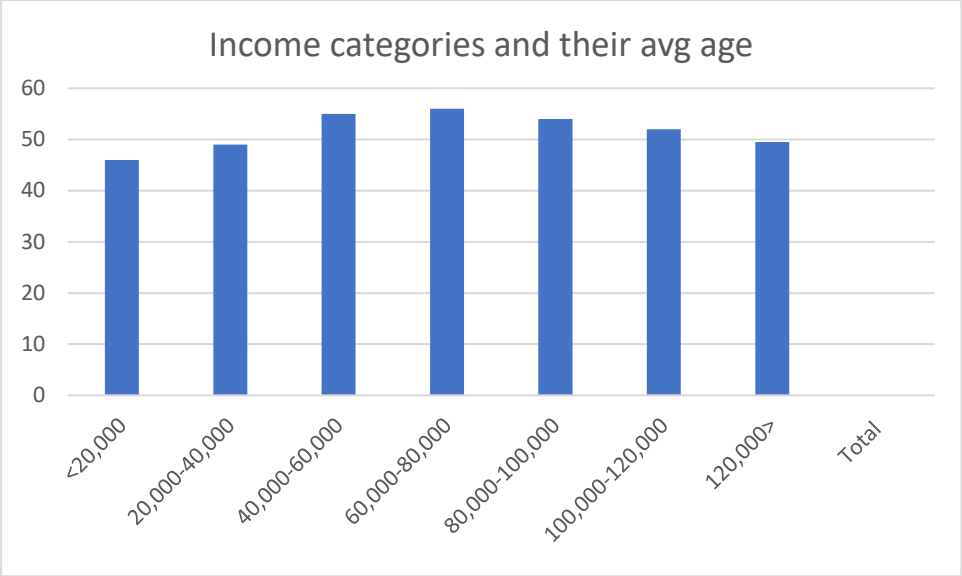
Excel:

The process I conducted involved the analysis of provided data using both Excel and SQL. In the Excel analysis: I commenced by downloading and extracting the given data onto my personal computer. The data files were initially in .csv format, and I transformed them into Excel files for easier analysis. The initial data cleaning process included tasks such as spell checking, removing error values, and eliminating blank spaces. I also applied conditional formatting to remove duplicate values throughout the dataset, which resulted in the removal of approximately 2000 or more redundant entries.
Subsequently, I conducted statistical analysis on the data, which began with calculating the average age of the customers. This involved determining the difference between the current year (assumed to be 2023) and their year of birth. Then, I used the AVG function on the age column to find that the average customer age was 54 years. The analysis also involved creating data visualizations to better understand the provided data. This preliminary analysis included.

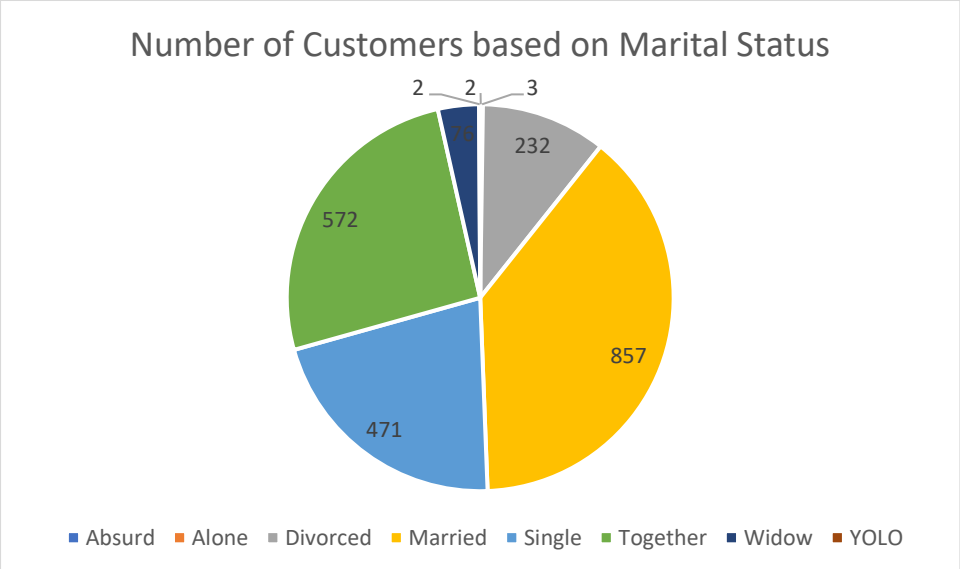
Graph 1: Examining the correlation between Marital Status and the average age of the customer group.



Graph 2: Investigating the relationship between income and different age groups.



Graph 3: Proportion of the total customers belonging to different Marital Status.



From these visualizations, I derived the following trends:

Trend 1: Married individuals appeared to be a significant portion of the customer base.

Trend 2: Married people had the highest average income, and their average age was approximately 53 years.

SQL SYNTAX

Step 1: Created 2 Tables marketing data and ad data

Step 2: Built Columns based on datatype and queries were made to find the following information.

- Total Spend by each country,
- Spend per each product per each country.
- Best selling product in each country.

Syntax is as follows.

```
SELECT
    Country,
    SUM(AmtLiq + AmtVeg + AmtNonveg + AmtPes + AmtChocolates + AmtComm)
FROM public.marketing_data
GROUP BY Country;

SELECT
     Country,
     SUM(AmtLiq) AS TotalSpendLiq,
     SUM(AmtVeg) AS TotalSpendVeg,
     SUM(AmtNonveg) AS TotalSpendNonveg,
     SUM(AmtPes) AS TotalSpendPes,
     SUM(AmtChocolates) AS TotalSpendChocolates,
     SUM(AmtComm) AS TotalSpendComm
FROM Spending
GROUP BY Country;
WITH TotalSpendPerProduct AS (
```

```sql
SELECT
    Country,
    SUM(AmtLiq) AS TotalSpendLiq,
    SUM(AmtVeg) AS TotalSpendVeg,
    SUM(AmtNonveg) AS TotalSpendNonveg,
    SUM(AmtPes) AS TotalSpendPes,
    SUM(AmtChocolates) AS TotalSpendChocolates,
    SUM(AmtComm) AS TotalSpendComm
FROM Spending
GROUP BY Country
```

```sql
Country,
CASE
    WHEN TotalSpendLiq >= GREATEST(TotalSpendVeg, TotalSpendNonveg, TotalSpendPes, TotalSpendChocolates, TotalSpendComm) THEN 'Lic
    WHEN TotalSpendVeg >= GREATEST(TotalSpendLiq, TotalSpendNonveg, TotalSpendPes, TotalSpendChocolates, TotalSpendComm) THEN 'Veg
    WHEN TotalSpendNonveg >= GREATEST(TotalSpendLiq, TotalSpendVeg, TotalSpendPes, TotalSpendChocolates, TotalSpendComm) THEN 'Nor
    WHEN TotalSpendPes >= GREATEST(TotalSpendLiq, TotalSpendVeg, TotalSpendNonveg, TotalSpendChocolates, TotalSpendComm) THEN 'Pes
    WHEN TotalSpendChocolates >= GREATEST(TotalSpendLiq, TotalSpendVeg, TotalSpendNonveg, TotalSpendPes, TotalSpendComm) THEN 'Cho
    ELSE 'Comm'
END AS MostPopularProduct
 TotalSpendPerProduct;
```

TABBLEAU DASHBOARD DESIGN:

The dashboard design and development process for this assignment was interesting and required a step by step breakdown of the objectives that the analysis wanted to achieve. The steps that were used for the process are given as follows and the rationale behind each element included in the dashboard design have also been included. In the data analysis process, several key steps were taken. Initially, an Excel file was imported to serve as the primary data source. Subsequently, the Data Interpreter tool was utilized to clean and preprocess the data, ensuring its accuracy and reliability for further analysis. In the realm of data visualization, several notable trends emerged.

The dashboard has 5 graphs and each are different from each other in terms of design, however it is important to note the rationale behind it is interesting. The supermarket that is being analysed here is a global chain and because of that the dashboard recurring theme is the country category. In order to understand the logistics of the customer base it is important to understand how it varies across these countries which are given here. The types of charts that were chosen were done so to make visualization more interactive and at the same time simple to understand. Some of the trends that were noticed from the Tableau dashboards were as follows.

Firstly, Spain was identified as the company's largest market. Secondly, it was observed that married individuals constituted the highest-earning group within the dataset. Additionally, the analysis revealed that married individuals tended to have more children on average. Notably, Spain, Canada, and Saudi Arabia emerged as the countries with the highest average income. Furthermore, it was discovered that Spain had the highest number of walk-in customers. Lastly, an intriguing trend was identified – a correlation between income and deals that was anticipated to be negative, yet India was the exception, displaying a positive relationship. This insight into consumer behaviour suggests that India's consumers may be price-sensitive, making this trend a noteworthy point of interest in the analysis.

Patterns and Trends Identified from the Analysis:

Spain as the Largest Market: The data indicates that Spain is the company's largest market, suggesting a strong customer presence in the Spanish market. This could be due to various factors, such as population size, economic activity, or the company's marketing efforts in Spain.

Marital Status and Income: Married individuals stand out as a significant customer segment. They not only have the highest average income but also tend to be in their early 50s on average. This implies that the company's products or services may be particularly appealing to married individuals in their middle years, potentially due to factors like stability, disposable income, or specific product preferences.

Family Size and Income: The analysis also reveals that married individuals tend to have more children on average. This suggests that the company's offerings might cater to families or individuals with dependents, influencing both their income and purchasing behaviour.

Country-Specific Insights: Spain, Canada, and Saudi Arabia are the countries with the highest average income. This could be due to economic factors, but it also suggests that the company's products or services resonate well with the consumer base in these countries.

Walk-In Customers in Spain: Spain's high number of walk-in customers might indicate a preference for physical retail experiences in this market. Understanding this trend could help the company optimize its sales and marketing strategies in Spain.

Income-Deals Correlation, with an Indian Exception: The intriguing finding of a positive income-deals correlation in India, while a negative correlation was anticipated, suggests that Indian consumers may be price-sensitive. This trend is a unique point of interest, as it deviates from the expected pattern and could have implications for pricing and promotions in the Indian market.

In summary, the data reveals that the company's largest market is Spain, with a strong presence of married individuals who have the highest average income. Understanding country-specific insights, family dynamics, and the unexpected income-deals correlation in India can help the company tailor its strategies to better serve its diverse customer base.

.