2MARKET

Customer Purchase Behaviour Analysis

-Exploratory Analysis Report-

Tilani Wijamunige - Data Analyst

2MARKET - Customer Purchase Behaviours - Exploratory Analysis Report

Business Context

2MARKET is a global supermarket chain that sells products both online and in-store. They are aiming to gain better insights into customer purchase behaviours. Following questions have been asked to meet their marketing objectives.

1) How do socio-demographic factors influence customer purchase behaviour?
2) What are the best-selling products and how do they vary by region?
3) What advertising channels are the most popular?

Getting deep insights into customers' spending habits will help *2Market* to make informed decisions regarding customer experience and retention, product development, business expansion, pricing strategies, and overall business development.

Analytical Approach

Customer data and advertisement data collected from 2216 *2MARKET* customers over unspecified period were used for this analysis. These data include customer demographics data, annual income ($), spendings ($) per category for unspecified period, purchase frequency and customers response for advertisement campaigns.

During the data cleaning process, there were no duplicates found, based on the unique customer id. Few anomalies were dealt as below.

1. Marital Status**:** 4 customers listed their marital status as "Absurd" or "YOLO". These entries were reclassified as "Unknown" to maintain data integrity.
2. Age Outliers: 3 customers were reported as >100 years. These customers were included in visual analysis (in scatter plot and whisker plot) with the use of filters but excluded in analytical stages when using average age. This approach allows for the examination of these outliers while providing the option to focus on more typical data ranges when necessary.
3. Education categories: The "2n Cycle" category in the education data was retained to ensure comprehensive representation of all education levels.

Socio-Demographic Factor Analysis

Averages, percentages, and distributions of demographic variables were calculated using Microsoft Excel due to its user-friendly interface and data manipulation capabilities, and straightforward data cleaning and processing abilities.

Findings:

- The average customer age is 54 years.
- Customers those who are “married” or “together” have counted the highest marital status. This can be concluded as total of 65% of the customers are in “a partnership” in comparison to the 35% those who are not.
- Most customers have a graduate level of education (50.3%), followed by those with a PhD (21.7%) and a master's degree (16.5%). Fewer customers have 2nd cycle (9%) basic (2.4%) level education.
- The average income across all customers is $52,247 and PhD holders have the highest average income at $56,145. Master's degree holders follow closely with an average income of $52,917.53. Those with a 2nd cycle and the basic education group has the lowest average income.

Marital Satus	Count of Customers	Average Annual income($)	Average of Total Spending ($)	Average Age
2n Cycle	**200**	**47,633.19**	**494.93**	**51**
Divorced	23	49,395.13	651.17	59
Married	80	46,201.10	427.73	50
Single	36	53,673.94	553.50	49
Together	56	44,736.41	472.84	50
Widow	5	51,392.20	677.20	64
Basic	**54**	**20,306.26**	**81.80**	**46**
Divorced	1	9,548.00	29.00	47
Married	20	21,960.50	122.85	45
Single	18	18,238.67	57.72	40
Together	14	21,240.07	56.43	52
Widow	1	22,123.00	102.00	63
Graduation	**1116**	**52,720.37**	**621.69**	**53**
Alone	1	34,176.00	89.00	35
Divorced	119	54,526.04	616.41	56
Married	429	50,800.26	595.75	53
Single	246	51,322.18	615.62	51
Together	285	55,758.48	661.13	55
Unknown	1	79,244.00	1,216.00	30
Widow	35	54,976.66	677.14	61
Master	**365**	**52,917.53**	**609.77**	**56**
Alone	1	61,331.00	632.00	65
Divorced	37	50,331.95	551.30	56
Married	138	53,286.03	566.67	56
Single	75	53,530.56	746.04	52
Together	102	52,109.01	560.98	58
Unknown	1	65,487.00	1,169.00	66
Widow	11	58,401.55	817.55	69
PhD	**481**	**56,145.31**	**676.73**	**57**
Alone	1	35,860.00	49.00	50
Divorced	52	53,096.62	632.87	59
Married	190	58,138.03	718.06	56
Single	96	53,314.61	626.77	55
Together	116	56,041.42	654.80	58
Unknown	2	48,432.00	424.00	50
Widow	24	60,288.08	797.67	68
Grand Total	**2216**	**52,247.25**	**607.08**	**54**

A strong correlation between education level, marital status, and income is indicated. Higher education levels typically lead to higher incomes, with PhD holders earning the most. Being in a partnership (married or together) is associated with higher income, especially among those with higher education.

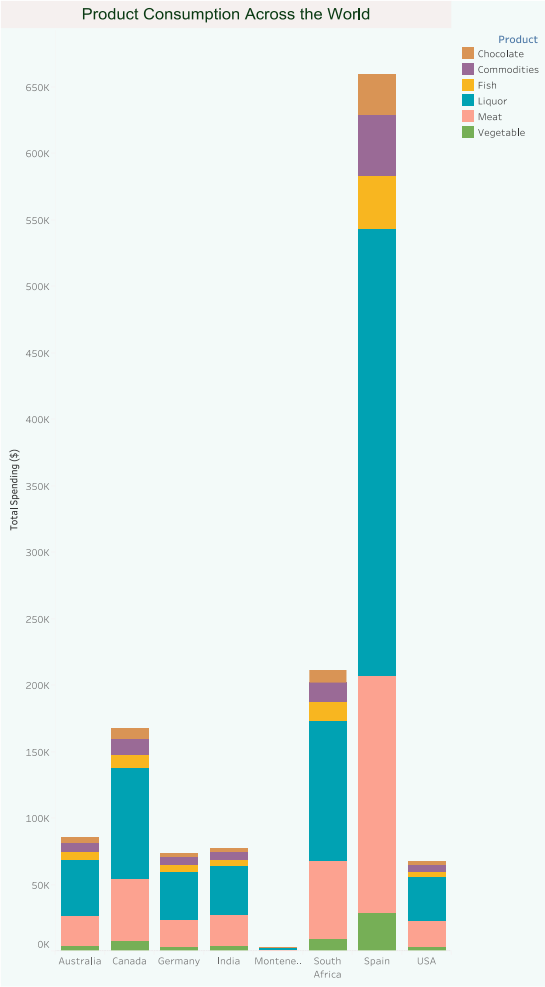
Global Product Consumption Analysis

PostgreSQL was used to create a customer database for querying and find relationship between mutable variables. Tableau was used to present and interpret the data insights visually due to its advanced analytical and visualisation features such as dashboards.

Findings:

- Spain is leading the total spending category with $659,557 spendings. South Africa has the 2nd most spendings with $211,071. Montenegro shows significantly lower spending compared to other countries, with a total spending of $3,122.
- Liquor is the category with the highest customer spending, amounting to $676,083 in sales. Meat is 2nd most popular product category with $370,063 of sales. Other significant categories include commodities ($97,427), fish ($83,405), chocolate ($59,896), and vegetables ($58,405).
- Households without children have a higher total spending ($699,622) compared to those with children ($645,657). Even though the difference ($53,965) is not substantial relative to the total spendings, Liquor is the highest spending category across both household categories.

Strategic offers and promotions should focus on liquor and meat sales in high-spending countries like Spain and South Africa, while also encouraging consumption of less popular categories like vegetables, fish, and commodities.

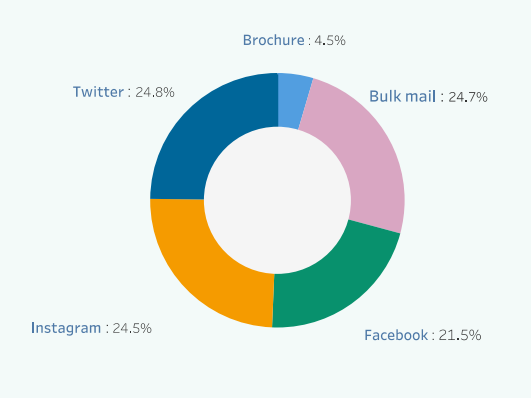


Impact of Advertising Channel Analysis

SQL was used to aggregate and analyse conversion data across advertising channels and to create RFM score and suitable segments. Tableau facilitated the clear and compelling visualisation by pivoting and joining multiple data sources.

Findings:

- Instagram is the most effective platform in Spain, South Africa, and Australia. Twitter is the most effective platform to reach customers from Canada, India, Germany, and Montenegro.
- Twitter has the highest total conversions (164), followed closely by Bulk Mail (163), Instagram (162), and Facebook (142). Brochures have the lowest total conversions (30).



Twitter should be employed for its highest number of conversions. Marketing strategies should prioritise Instagram for visual and engaging content in Spain, South Africa, and Australia, and Twitter for brief engagements in Canada, India, Germany, and Montenegro. Bulk Mail should be used for its strong performance in targeted, personalised communications. Suitability of brochures should be re-evaluated.

RFM analysis

RFM (Recency, Frequency, Monetary) score is a marketing analysis tool used to assess and predict customer purchase behaviour based on RFM values. The RFM score was calculated using SQL with the NTILE (5) function, which divides the data into five equally sized groups (1 = lowest, 5 = highest). The combined RFM score is useful for identifying important aspects of customer behaviour. This score was further used to segment customers into 10 meaningful groups and displayed on a heat map.

Findings:

- Nearly a third of the customers (29%) are “loyal”, indicating a solid foundation of satisfied customers. This group is critical for business success and positive marketing.

- A significant portion (23%) of the customer base is categorised as “needing attention”. "Almost Lost" and "Lost Customers" segments (28%) represent over a quarter of the customer base. This suggests a significant risk and highlights the importance of understanding and addressing the reasons behind customer dissatisfaction.

Dashboard design

Data sources were imported into Tableau to create a dashboard, utilising analytical insights from Excel and data structures created in PostgreSQL. Additional data pivoting and calculated fields were also used. Most relevant charts and graphs were put together to address the 3 main objectives. The dashboard features clearly labelled, colourful, and distinct charts and graphs on a neutral background, with interactive elements and filters for enhanced usability. It is optimised for an A3 Landscape (1654x1169) setting, providing the *2MARKET* marketing department a comprehensive user experience to gain valuable customer insights.

Recommendations based on the marketing objectives and analytical insights

1. Engage different age groups: The average customer age indicates a middle-aged and older demographic. Maintain engagement with them in-store discounts and dietary offers. Attract younger customers with targeted social media campaigns and influencer partnerships on Instagram and TikTok, featuring promotions like student discounts and flash sales.
2. Boost spending in Montenegro: Address low spending in Montenegro with campaigns and products tailored to local communities. Implement personalised offers, gather feedback through surveys, and utilise family and multi-buy discounts to increase popularity.
3. Appeal to couples and singles: Most customers are in partnerships. Create offers for couples, such as “meal deals for two” and “buy one-get one free” promotions. Develop strategies for single customers, like “meal for one” deals, to cater to the 35% not in partnerships.
4. Promote less popular categories: Use targeted promotions and seasonal offers to boost sales in categories like vegetables, fish, and commodities. Partner with local suppliers for region-specific items to ensure competitive pricing and freshness.
5. Collect relevant/up to date customer data: Gather more customer-focused data, including gender information, to better target products. As suggested by Bloomberg Intelligence, there is a growing market for pets. Gather more data about having pets at home to explore the market opportunities for pet food and care products.
6. Develop a mobile App: Create the *2Market* mobile app to enhance customer segmentation beyond RFM analysis. Use the app to gather more data, refine segmentation strategies, and design personalised loyalty schemes to boost customer engagement.

Reference

1. Bloomberg Intelligence 2023, 'Global Pet Industry To Grow To $500 Billion By 2030, Bloomberg Intelligence Report Finds', Bloomberg, viewed 23 June 2024,< https://www.bnnbloomberg.ca/pet-lovers-shopping-online-are-disrupting-the-industry-1.1955301.>
2. Buckland, T 2023, 'RFM Segments Based on RFM Analysis: An In-Depth Guide', MoEngage, viewed 13 June 2024, <https://www.moengage.com/blog/rfm-analysis-using-rfm-segments/>
3. Hanna, KT 2024, 'RFM analysis (recency, frequency, monetary)', TechTarget, viewed 15 June 2024, <https://www.techtarget.com/searchdatamanagement/definition/RFM-analysis>
4. Joseph, SG 2023, 'Customer Analysis using Tableau – Dashboard From Scratch', YouTube, viewed 14 June 2024, <https://www.youtube.com/watch?v=_qReGTOrKTk&t=570s>.
5. Kurniawan, A 2023, 'Superstore's RFM Analysis using PostgreSQL', Medium, viewed 15 June 2024, <https://medium.com/@arif0kurniawan/superstores-rfm-analysis-using-postgresql-part-2-1ed1b30cfd0d>.
6. Microsoft Support n.d., 'Pivot data in a PivotTable or PivotChart - Excel for Microsoft 365', Microsoft, viewed 07 June 2024, <https://support.microsoft.com/en-gb/office/pivot-data-in-a-pivottable-or-pivotchart>
7. Shenvi, T 2023, 'RFM Analysis for Customer Segmentation', CleverTap, viewed 20 June 2024, <https://clevertap.com/blog/rfm-analysis/.>
8. Tableau Help n.d., 'Build Common Chart Types in Data View', Tableau, viewed 07 June 2024, <https://help.tableau.com/current/pro/desktop/en-us/dataview_examples.htm>.

Note: Please find the attached Excel and SQL files used for the analysis in the zip file.