Module 1 Assessment: 2Market Business Case Study
05/09/2022

2Market is a global supermarket chain seeking to better understand their consumers. The business commissioned a report on their customer demographics, the effectiveness of their advertisement campaigns, and the popularity of their products across consumers. 2Market provided historical customer and advertisement data, which contained two datasets corresponding to all customers globally. This data was analysed to capture insights and extrapolate trends, which were used to generate strategic marketing recommendations. These insights will improve 2Market's understanding of their customers, including those responsible for the largest revenue contribution, and customer engagement with advertising campaigns.

To ascertain what insights would be relevant to 2Market, the '5W' problem-solving framework was applied, which converted 2Market's mandate into a series of analytical questions. As the business did not posit a specific issue, it was assumed that the research would support marketing activities, based on the combination of questions about customer attributes and advertisement campaigns. Each of the three questions were subdivided into broader, actionable queries. 2Market first sought insights pertaining to 'customer demographics', which was expanded to encompass comparative revenue generation. Next, as the business requested that their advert campaigns be assessed, the effectiveness of campaigns in relation to revenue was also queried. Lastly, to determine what products were the most popular, both individual and regional demographics were scrutinised. This report was generated with a Chief Marketing Officer in mind, with all results pertaining to customers and revenue generation.

After the analytical framework was developed, the data was cleaned and wrangled in Excel. There were several outlying values across the data in income, age, and marital status. In customer income data, a box-and-whisker plot in Tableau was used to determine the distribution of values. The chart determined that salary values above $113,734 were considered outliers and were omitted. Additionally, there were several marital status datapoints that contained illogical values, such as 'YOLO'. These values were incorporated into a miscellaneous column so that corresponding values in other columns could still be factored into analysis. An additional column for the age of each customer was generated based on their year of birth, which also produced some outliers. Customers above 110 years old were removed from the data set on the basis that they would skew analysis.

Once cleaned, descriptive statistics of the dataset were calculated to determine the average consumer profile. The average 2Market customer is 51 - 52 years old. They have an average income of $51,623 and median income of $51,371. 2Market's customers are most likely to be married, representing 39% of the dataset, and customers are most likely to have an Undergraduate degree (50%).

With the benchmark statistics established, consumer behaviour was investigated based on demographic attributes, such as marital status, age, or education. Surprisingly, there was no significant relationship between customer attributes and items purchased. There was no pattern between purchases and age, marital status. Furthermore, there was only a slight difference in purchases between different education levels, with customers that possessed a PhD purchasing slightly more alcohol than their counterparts. To establish whether there were any patterns in products purchased, regional activity was next analysed. This determined that Spain is the largest market, with 1,092 customers and a revenue of $657,704. In Spain, the most popular products are fish, accounting for $40,049 revenue for the region. In Germany, Canada, India, Australia, the USA, and Montenegro, the most popular items are alcoholic beverages. In South Africa, the most popular items are meat products.

To determine which advert campaigns are the most effective the table of advertisement data and customer data was joined and queried using SQL (Appendix A). The popularity of campaigns differ across regions, with Instagram being the most popular in Spain and Australia. In Canada and Germany, Twitter is the most popular platform. In India and the USA, email campaigns are the most

popular, and in South Africa email campaigns are as popular as Twitter. In Montenegro there is no engagement across any platforms except 1 email conversion. The data implies a relationship between revenue per region and the amount of lead conversions, with Spain having both the largest revenue and the most clicks. Further exploration is required to establish whether lead conversions and revenue are connected, but the relationship between the two is strongly correlated in this data.

After exploring the data patterns became apparent, which were visualised using Tableau. With the framework of 2Market's core questions in mind, the data was investigated thoroughly, and a dashboard was built to support this. The dashboard is divided into three segments corresponding to the three primary questions 2Market requested insights on.

As the first line of enquiry regarding customer demographics was open-ended, the question was expanded to determine which of 2Market's customers were the most valuable in revenue contribution. In the upper-left corner of the dashboard is a scatterplot showing the interaction between income and revenue contribution, which shows a strong positive correlation between the two. Intriguingly, there is a negative correlation between the number of children a customer has and the amount they spend, which the user can observe by toggling the children filter on this chart. This chart utilises a sequential colour scheme which corresponds to an increase in revenue contribution. In the middle of the upper row is a box-and-whisker plot showing the relationship between educational attainment and revenue contribution. This chart follows the theme present in the first chart, that higher education leads to a higher income, which correlates to an increased spend in 2Market. The rationale for including this plot was to showcase an alternative customer grouping that 2Market can target with marketing activities. In the upper-right side of the dashboard is a histogram quantifying the number of times consumers in different income brackets visited a shop, purchased from the shop, or purchased on the internet. This showcases an interesting shift in behaviour across consumers, with those on lower incomes visiting the supermarket more often but purchasing less than those with higher income. These three charts collectively provide insight on 2Market's customer demographics and their behaviour.

2Market additionally sought information regarding the effectiveness of their ad campaigns. To demonstrate how well these campaigns perform in each region bar chart was incorporated in the lower-left side of the dashboard, which is intended to be analysed alongside the charts directly below. Holistically, these charts show the popularity of different campaigns in each region, but also imply a cause-effect relationship between successful lead conversion in each region and revenue generation. For example, Spain has the largest amount of lead conversions across all platforms and also contributes the most revenue.

Lastly, 2Market also wanted to learn about which products are the most popular and whether this differs by customer demographics. While testing the data there was little evidence to suggest customers with different attributes purchased different types of products, but it was discovered that there were preferences the regional level. For this reason, in the lower-right corner of the dashboard there is a map of revenue across each region, which can be filtered by different products. This map enables the user to isolate each product line and assess its performance by region, expanding upon the insights shared in this report.

From the information collated in the analysis, recommendations were also generated to apply these findings. Firstly, as 2Market's largest revenue-contributing customers are high-earners or customers with fewer children, campaigns can be tailored to their preferences. A potential area for 2Market to explore further would be to survey these customers in order to establish what items they prefer as this study did not find any product preferences by income. Additionally, as alcoholic products are the most popular among customers and across 6 out of 8 regions, promotional activities on alcoholic products are likely to be a popular choice. Lastly, marketing activity and consumer preferences differ by region. Based upon the insights presented in this report, advert budgets should be adjusted according to the popularity of different platforms in each region, and a secondary study should be conducted to determine the return on investment for each of these platforms. This study also implies

a positive relationship between lead conversions and revenue, with regions with the most clicks per ad showing the strongest revenue. Additional data is needed to robustly quantify the effectiveness of each campaign beyond lead conversion, and to determine whether the relationship between lead conversion and revenue is causal, but the findings of this report strongly suggest a relationship.

With an ambition to support 2Market's marketing activities, a rigorous and multi-platform analysis was conducted. This report captures consumer information and identify patterns that give a clear depiction of 2Market's customers, their preferences, and how to tailor advertising campaigns towards them both at an individual and regional level. To ensure the findings were relevant and actionable, strategic recommendations were implemented and advice on suggested additional investigation that could bolster the findings of this analysis was included.

Appendix A. SQL Syntax using Inner Join and Tabulated Findings:

Syntax:

```
SELECT "Country",
SUM("Twitter_ad") AS Twitter_sum,
SUM("Instagram_ad") AS Instagram_sum,
SUM("Facebook_ad") AS Facebook_sum,
SUM("Brochure_ad") AS Brochure_sum,
SUM("Bulkmail_ad") AS Bulkmail_sum,
SUM ("AmtLiq" + "AmtVege" + "AmtNonVege" + "AmtPes" + "AmtChocolate" + "AmtComm") AS
Totals
FROM public.marketing_data4
INNER JOIN public.ad_data USING ("ID")
GROUP BY "Country"
```

Output:

"Country"	“twitter_ sum”	"instagram _sum"	"facebook_sum "	"brochure_ sum"	"bulkmail_ sum"	“Totals”
"SP "	87	**88**	76	16	83	657704
"CA "	**24**	21	18	6	18	167403
"IND "	10	6	7	2	**13**	77741
"AUS "	6	**12**	7	0	9	85576
"US "	6	5	7	0	**8**	67546
"ME "	0	0	0	0	**1**	3122
"SA "	20	**21**	20	4	**21**	210987
"GER "	**11**	8	7	2	10	73198

This chart was used to determine the revenue for each country in addition to how many successful lead conversions there are. The chart shows both the most popular advertisement channel in bold, and shows a general trend between the amount of lead conversions and the revenue for the region.