

Assignment 1: 2Market Performance Report

14 March 2022
Hazel Chan

LSE Career Accelerator Data Analytics

Table of Contents

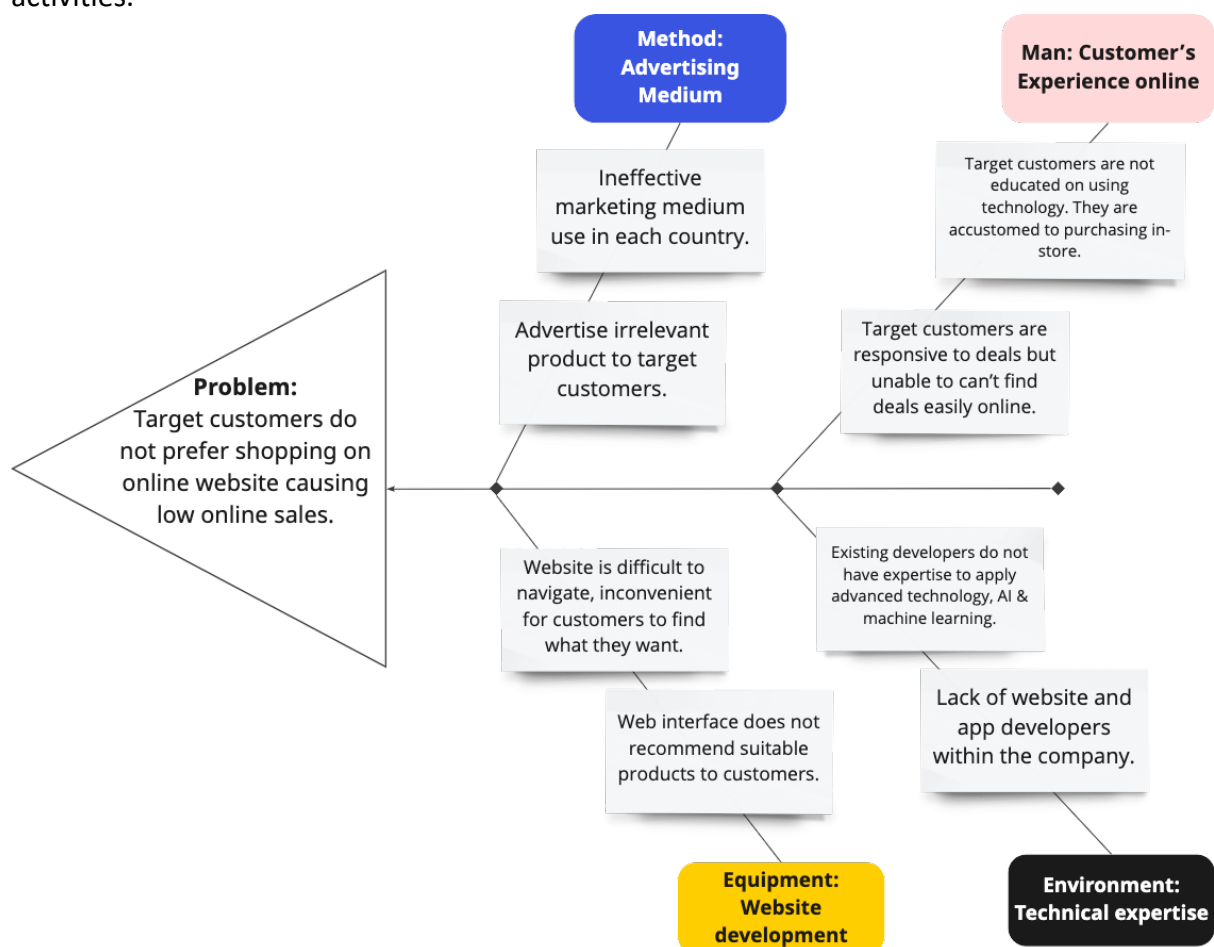
<i>Background</i>	<i>3</i>
<i>Analytical Approach</i>	<i>4</i>
<i>Dashboard Design and Development</i>	<i>8</i>
<i>Patterns and Insights</i>	<i>9</i>

Background

With the emerging trend of online grocery shopping, global supermarket, 2Market, is struggling with low online sales. Based on historic data, 95% of conversions was from online advertisements, however, only 41% of customers purchased online.

Fishbone diagram is used to identify three underlying causes: irrelevant products are advertised to customers, ineffective advertising medium used in each country, lack of understanding customer demographics.

To adapt to the digital market, the business is devising a marketing campaign with a data driven approach. Data analysis on the sales by product, average revenue of advertising medium and demographics of top spending customers will be conducted to inform marketing activities.



Analytical Approach

Excel and SQL are used in the pre-analysis and analysis stage to clean, transform, explore, and analyse market and adtype datasets.

Firstly, data formatting is cleaned in Excel to validate the integrity of data. Conditional formatting is used to remove duplicates in customer ID column, special search function is used to check errors and empty cells, and text to columns is used to remove symbols in numerical column.

Secondly, descriptive analysis was conducted using box and whisker to identify outlier of customer ID 1501 who spent \$1717 and only purchase once with 2Market.



Thirdly, SQL is used to import and create marketdata and adtype tables. Then, SQL statement is written to effectively perform multiple actions including joining tables, aggregating total spend, sum of online_ad and offline_ad conversions, grouping by non-aggregated variables and order by id. Statement is written with lowercase for names, uppercase for commands to minimise syntax errors and comment is used to document objective results for future reference.

Query Editor

Query History

```

103 --joined marketdata and adtype table to analyse revenue generated by adtype
104 SELECT
105     m.id, country, marital_status, amtliq, amtvege, amtnonveg, amtpes, amtchocolates, amtcomm,
106     numdeals, numwebbuy, numwalkinpur, SUM(amtliq+amtvege+amtnonveg+amtpes+amtchocolates+amtcomm) AS totalspend, numvisits,
107     a.bulkmail_ad, twitter_ad, instagram_ad, facebook_ad, brochure_ad,
108     SUM(bulkmail_ad + twitter_ad + instagram_ad + facebook_ad) AS online_ad, SUM(brochure_ad) AS offline_ad
109 FROM public.marketdata m
110 LEFT JOIN public.adtype a
111 USING(id)
112 GROUP BY
113     m.id, country, marital_status, amtliq, amtvege, amtnonveg, amtpes, amtchocolates, amtcomm,
114     numdeals, numwebbuy, numwalkinpur, numvisits,
115     a.bulkmail_ad, twitter_ad, instagram_ad, facebook_ad, brochure_ad
116 ORDER BY id;
117

```

Data Output

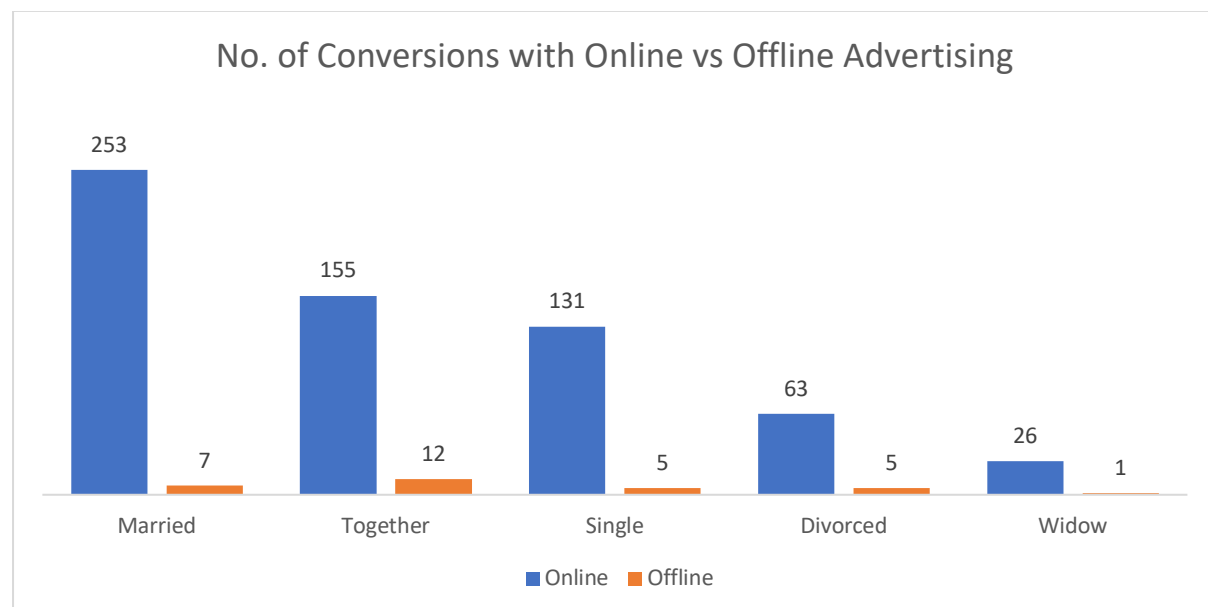
Explain

Messages

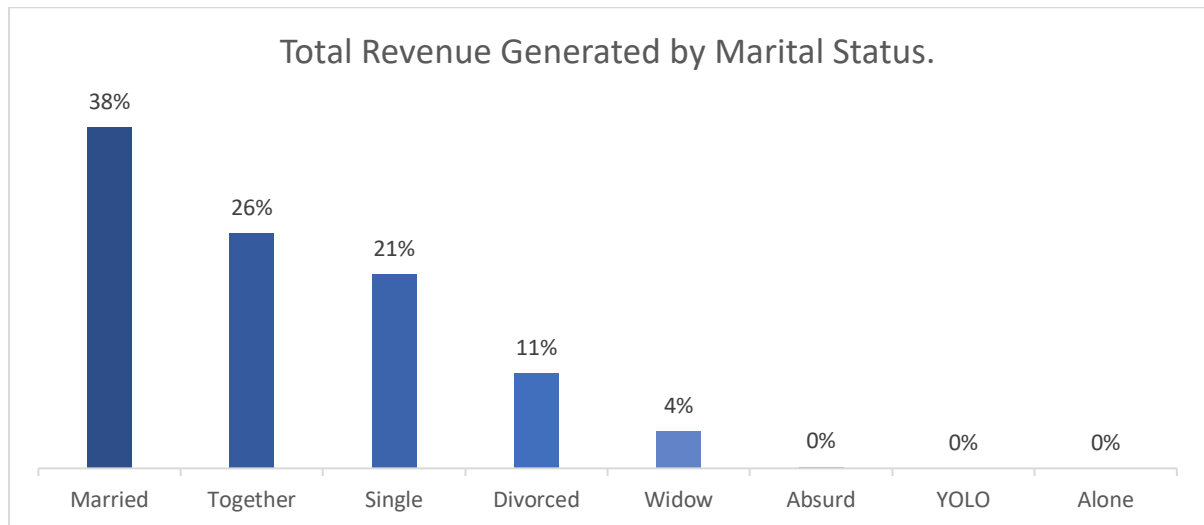
Notifications

	id bigint	country character (50)	marital_status character (50)	amtliq bigint	amtvege bigint	amtnonveg bigint	amtpes bigint	amtchocolates bigint	amtcomm bigint	numdeals bigint	numwebbuy bigint	numwalkin bigint
1	0	SA	Married	239	10	554	254	87	54	1	3	
2	1	CA	Single	464	5	64	7	0	37	1	7	

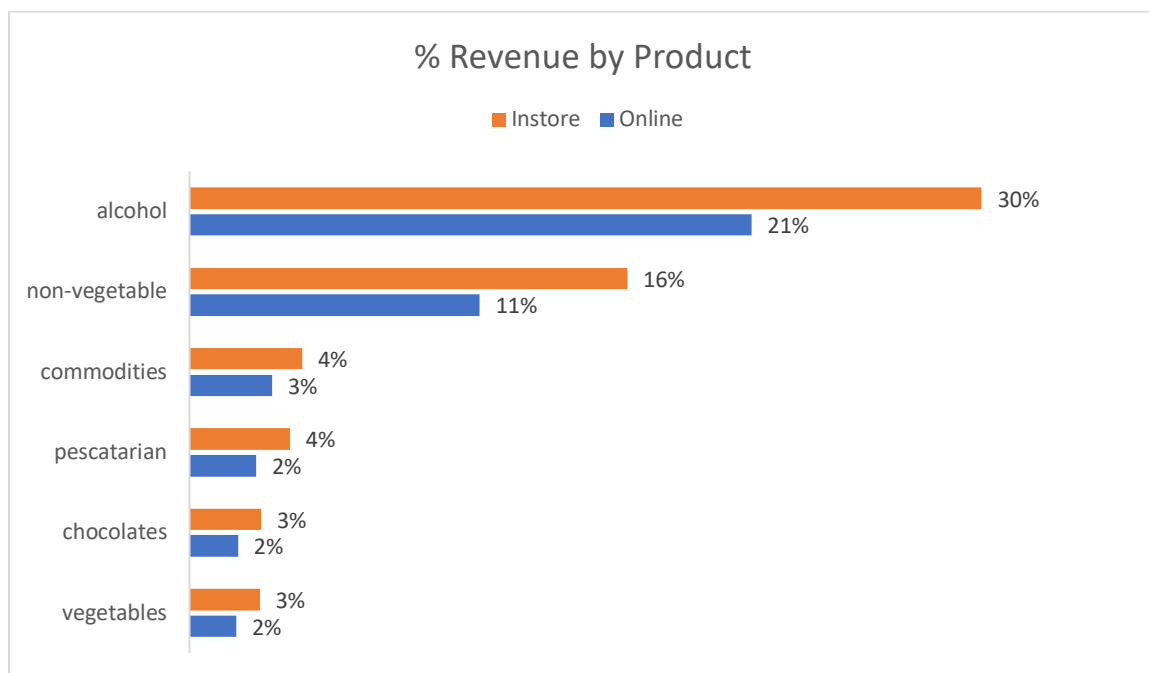
Fourthly, exploratory analysis is conducted using Excel for quick visualisation to discover insights. Pivot table and double-bar graph are used to analyse most effective advertising medium for each marital status. Below shows that majority is responsive to online advertisements.



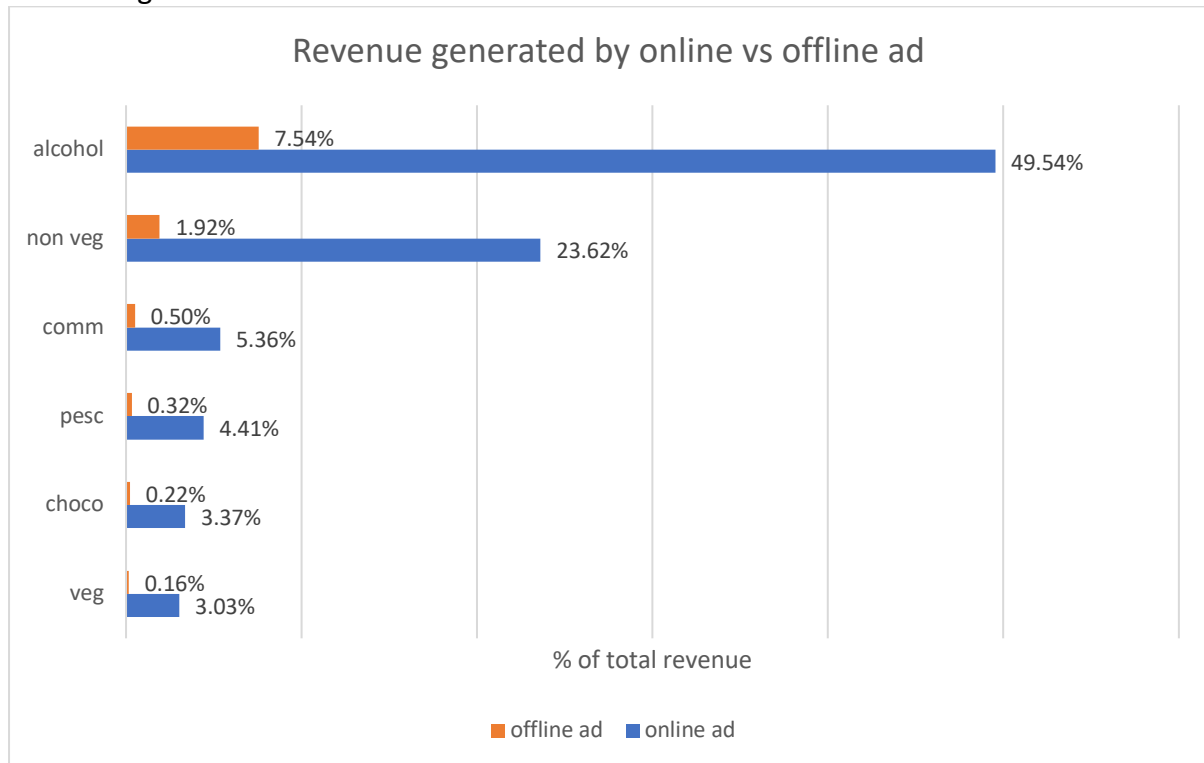
Pivot table and bar graph are used to aggregate revenue and analyse top spending customers. In the graph below, married and together customers account for 64% of total sales revenue.



Instore and online revenue for each product is calculated based on the ratio $\frac{[\text{NumWebBuy or NumWalkin}]}{[\text{TotalNumPurchase}]} \times [\text{amt spent on each product}]$ in the marketdata table. Then, double-bar graph was used to compare instore and online revenue by product. This analysis shows that alcohol generates 51% of total sales revenue and 30% is generated instore. As a supermarket, its food categories are not performing well which may signify lack of brand awareness. However, results may be due to higher price point of alcohol than other food categories resulting in higher sales performance. Calculating with sales quantity data may lead to more accurate comparison.



Then, assuming online purchase is influenced by online advertisements, online and offline ad revenue for each product are calculated in the same table based on the ratio: $\left[\frac{\text{number of online_ad conversions}}{\text{numwebbuy}} \right] * [\text{online revenue of each product}]$. 89% of total advertising revenue is from online.



Lastly, this excel worksheet is exported as `adtype_rev_cleandata.csv` including *instore vs online sales revenue* and *online_ad* and *offline_ad* revenue by each product.

Dashboard Design and Development

Tableau is used in the post-analysis phase to communicate insights to internal and external stakeholders. Two files, `clean_marketdata_SQL.csv` and `adtype_rev.csv`, are connected as data source and joined together using ID column. Each variable is grouped by folder under data pane, then graphs are created in independent worksheets and added to the dashboard.

The dashboard layout is divided into three sections using colours allowing stakeholders to perceive relevant and important insights from left to right: green represents revenue generated by product and advertising medium, blue represents customers shopping behaviours and orange represents effectiveness of online advertising medium. Offline advertising is omitted as the business problem is focused on improving online sales.

As this dashboard is intended for board of directors and investors, insights are kept at high level and details are displayed using tooltips when hovering over a mark. Interactive filters are added on the right for country managers to view by country and marital status. Instructions and objective of the dashboard are displayed by hovering over information icon next to dashboard title.

In the green section, total revenue is displayed in large numbers with a breakdown of the revenue by product, country, advertising medium displayed in charts or tables underneath. Values are sorted in descending order to easily identify highest revenue generating variable. Online and offline advertising revenue values were aggregated as average per advertisement for accurate comparison as there is substantially more online ads conversions than that of offline ads.

In the blue section, customer demographics and behaviour are displayed for stakeholders to develop campaign messages targeted to top spending customers. Total sales revenue by customer marital status is formatted as percentage with sequential highlight for ease of reading and tooltip includes number of deals purchased as element to attract customers. Percentage of instore versus online purchases variable is created using calculated field and is displayed using pie chart with labels. Average spend on website is displayed using a line chart to identify bestselling product online.

In the orange section, the number of conversions for each online advertising is presented with a horizontal bar graph for comparison. Since 95% of conversions is from online advertisements, brochure ad is omitted from this graph to focus on increasing online sales.

Patterns and Insights

The business is generating 50% of total revenue from selling alcohol with Spain as its largest market. Target customers of the campaign are married and together customers who accounts for 64% of total revenue. Although alcohol is performing well online and instore globally, it signifies customers perceive the brand as a liquor store more than a supermarket for daily necessities as its food categories perform significantly lower.

Moreover, online ad average revenue is \$24 higher than that of offline advert at \$3 but only 41% of customers prefer to shop online. This implies that customers are fluent in technology as majority of advertising revenue is from social media platforms but are unattracted to the online website.

In conclusion, it is recommended for the business to develop a marketing campaign to drive food product sales by targeting married customers through appropriate online ad medium to build multiple revenue streams and increase brand awareness. To increase online sales, it is suggested to invest in AI and machine learning technologies to improve online shopping experience with personalised product recommendations. Though further data is required to analyse website engagement to make decisions on improving website design.