2Market's Customer Purchase Behaviour

Hamza Zidan
19/12/2022

Table of Contents

Background/context of 2Market ***.....3***

Analytical Approach ***.....4***

Dashboard design and development ***.....7***

Patterns, trends, and insights ***.....9***

Background/context of 2Market

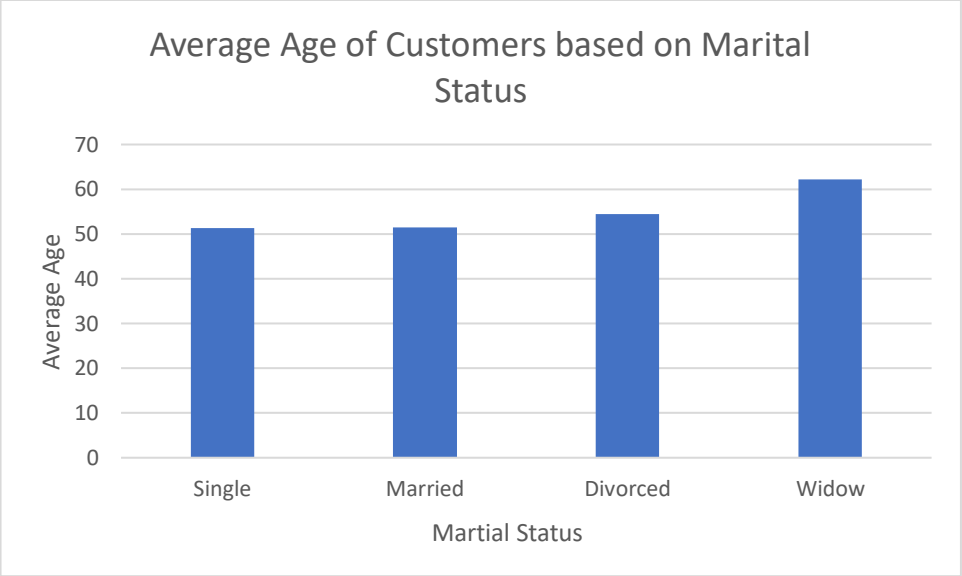
2Market is a global supermarket which sells products online and in-store in different parts of the world. Currently, they are on a mission to learn about the purchase behaviour of their customers. Why? The business would like to learn more about the demographics of its customers. Why? To understand if products and advertising channels vary based on demographic. Why? So that the correct products are advertised through the appropriate channels to each customer demographic group. Therefore, the business problem is that the marketing team would like to improve their approaches to attract more customers for a stronger return on investment.

Analytical Approach

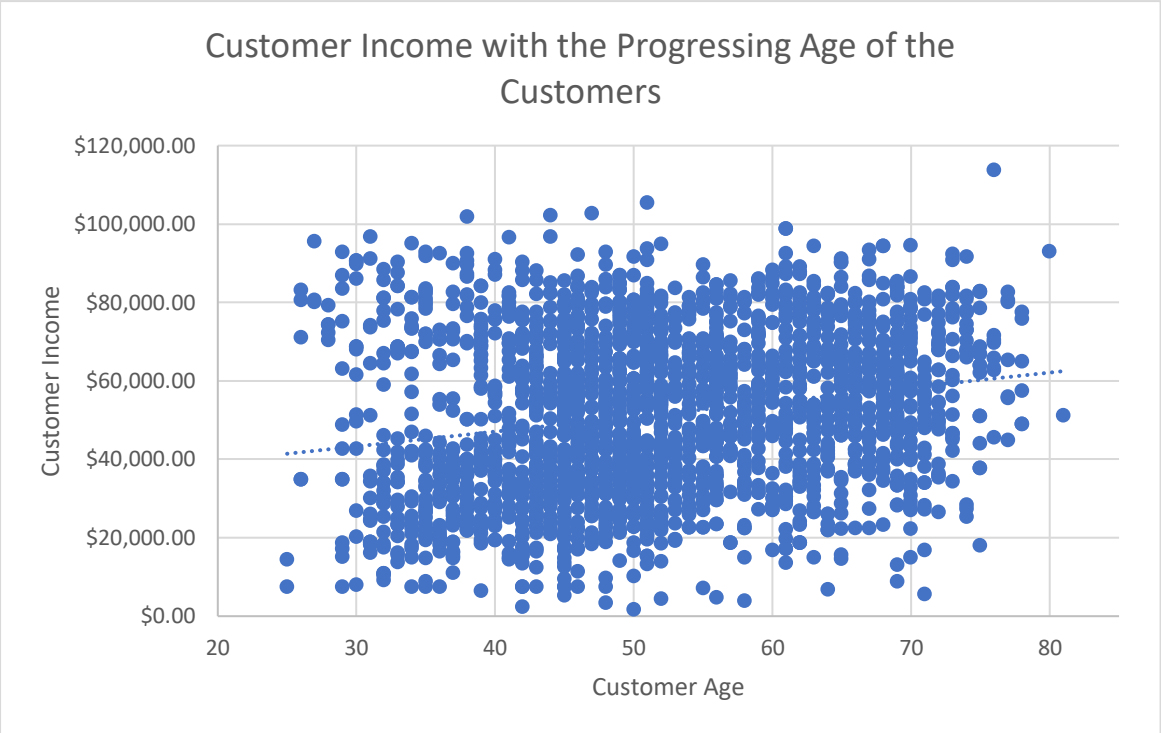
After receiving the raw marketing data, cleaning needed to be in process. This process was conducted through Microsoft Excel. Before cleaning, I ensured that there were no missing data.

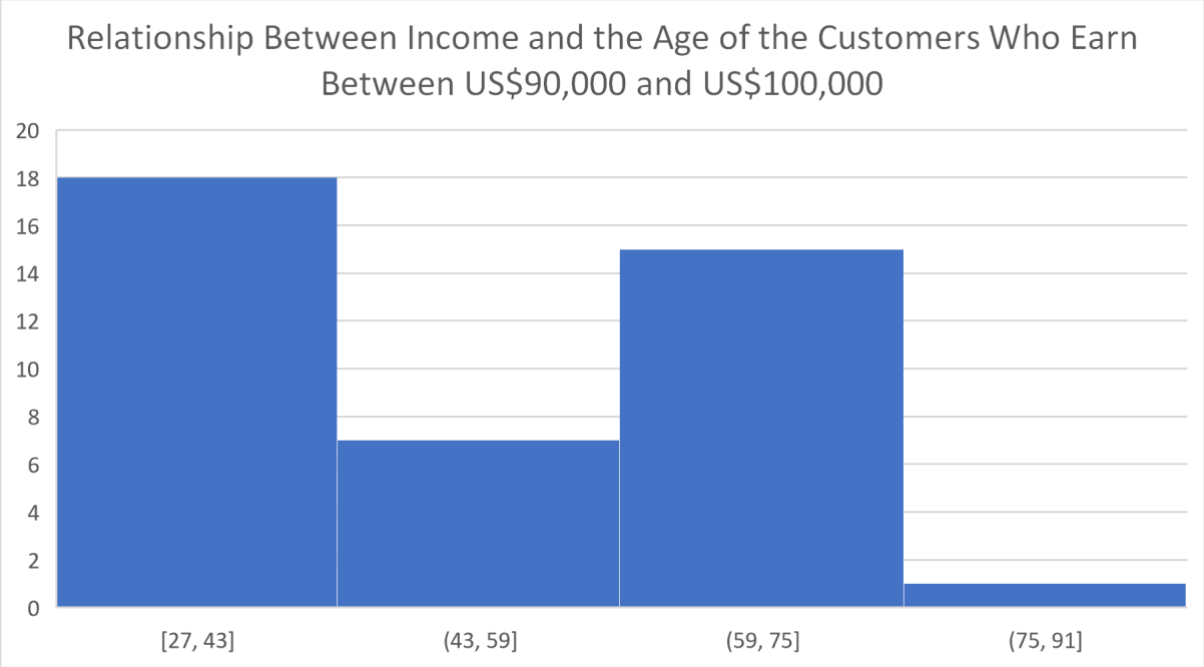
However, many of the data such as the date were not inputted uniformly. My first task was to make all dates in DD/MM/YYYY format. Also, customers have inputted random categories for marital status such as “Absurd”, “Alone”, “Together”, and “YOLO”, rather than just “Single (Unmarried)”, “Married”, “Divorced”, and. “Widow”.

The next step was to find the age of each customer, this was done by taking the current year and subtracting it from their birth year. Finding out the ages made it simpler to identify any outliers that may negatively affect the analysis by finding out the lower limit (17 years old) and upper limit (89 years old). As shown below, this made it possible to understand the average of each of the customers in each marital status category.



In addition, the figures in the income column were seen as strings rather than currency figures. After converting them to currency figures, I used the lower limit (-$14,431.50) and upper limit ($117,908.50) to remove outliers. The data was then put in a scatter display comparison among categories of income and the progressing age of the customers, and the trend was shown that there is a positive correlation between the increase in age and income with some exceptions to other customers as shown below.





After cleaning the data, I transferred the data to PostgreSQL with the use of pgAdmin to conduct further analysis. I have created a database where I was able to add both market and advertisement data and created tables ensuring each field is correctly categorised for the values to be added.

My first query was to find out how much each country spent in total, which resulted in Spain having the most products sold to customers. My next query is to find out the total spend per product per country.

Finally, I have done a left join with the files marketing data and ad data respectively, to understand which are the most effective advertising channels for each country and marital status.

Dashboard design and development

The thought process of constructing my dashboard was that it could allow anyone from the marketing team, or a senior stakeholder could look at it with the business problem in mind and be able to formulate an answer after using the filters. Having this in mind, I needed to make sure that the most relevant information needed to be added with filters that will make the accessibility as smooth as possible.

After querying the data with PostgreSQL, I have decided that the first part of the dashboard will be a map with the countries that 2Market has customers in to be highlighted, which reveals the total quantity of products sold when hovering a country on the map. The rationale behind this segment of the dashboard is to give a stronger idea of how many customers are in each country, and understand which country holds the highest amount to project more focus on there.

The next visual added to the dashboard shows the sum of the number of products purchased by customers in each age range, with the range being 10 years. The graph also has a filter so that a stakeholder can easily look through the performance of each product in each age range. This graph was created by pivoting the columns of products. Hence, making two new columns which one shows the range of products and the other shows the quantity sold. The rationale is to understand if age plays a role in the decision-making of different purchasing products.

The third visual added to the dashboard shows the average income customers in different age ranges make. This was added as income links to purchasing power and consumer confidence. Thus, it would help give an idea to the marketing team on what type of product they could advertise whether it is on the higher or lower end, depending on what most of their customers could afford. In addition, this could also guide them to know if older

customers make more money than younger ones, hence giving them a clearer idea of how to target different audiences.

Finally, the last visual added are the different social media platforms and how often customers in different categories of marital statuses use them. The rationale is to understand which platforms are widely used and which method of advertising channels should be dropped.

To add, the dashboard has four filters which control all the visuals, they are Country, Marital Status, Income Range, and Age Range. This allows users to go more in-depth to access more knowledge about their customers. For instance, the dashboard can be filtered to find out more about Spanish customers who are married and buy liquor. As a result, the user will see the number of customers, the amount of liquor and income within each age range, and the most used social media platforms.

Patterns, trends, and insights

I have concluded Spain has the highest amount of total spend per product for each product and Montenegro has the lowest total spend per product for each product. In addition, the data showed that liquor is the most popular product for each country and based on each marital status, to be more specific Spanish singles and married couples have the highest total purchase for liquor which is 155,906 and 127,625 respectively.

These results have made me question why Spain has the most products bought. After querying the number of customers each country has, I have noticed a significant difference in customer numbers with Spain having 1,089, South Africa with the second highest number of 334, and Montenegro with the lowest number of 3. This is a limitation in data and making marketing assumptions, as it may not seem that 2Market only has 3 customers in 1 country.