Graphical user interface

Description automatically generated

Master of Science (MS) in Data Science

**ITC6004A1 Data Visualization**

**07/04/2023**

FINAL PROJECT

**Mediterra Market: Marketing Department**

Professor: Dr. Bloutsos Konstantinos

**Students**

Lykos Stamatis - 273820

Miloš Lazarevic - 272036

Karatsiolis Christos - 270254

The American College of Greece

**WINTER SEMESTER 2023**

Table of Contents

[1. Introduction 3](#_Toc131625264)

[2. Overview 4](#_Toc131625265)

[2.1 Our dataset 4](#_Toc131625266)

[2.2 PowerBI Context 5](#_Toc131625267)

[3. PowerBI Analysis 6](#_Toc131625268)

[3.1 Pre-processing 6](#_Toc131625269)

[3.2 Customer Demographics 7](#_Toc131625270)

[3.3 Customer Purchasing Behavior 10](#_Toc131625271)

[3.4 Sales Channel Analysis 12](#_Toc131625272)

[3.5 Campaign Response Analysis 14](#_Toc131625273)

[3.6 Correlation Heatmap 15](#_Toc131625274)

[3.7 Meat – Wine Correlation (Linear Regression) 16](#_Toc131625275)

[4. Executive summary 18](#_Toc131625276)

[5. Extra conclusions 19](#_Toc131625277)

[6. Appendix 20](#_Toc131625278)

Table of figures

[Figure 1: Dataset overview 4](#_Toc131624771)

[Figure 2: Gender distribution 7](#_Toc131624772)

[Figure 3: MS distribution and Gender/MS relationship 8](#_Toc131624773)

[Figure 4: Education distribution 8](#_Toc131624774)

[Figure 5: Age distribution 9](#_Toc131624775)

[Figure 6: Kids and teen summary 9](#_Toc131624776)

[Figure 7: Stores locations map 9](#_Toc131624777)

[Figure 8: Product categories summary 10](#_Toc131624778)

[Figure 9: Average of products sold per store location 10](#_Toc131624779)

[Figure 10: Last month transactions pie chart 11](#_Toc131624780)

[Figure 11: Scatter plot between age and products sold 11](#_Toc131624781)

[Figure 12: Clustered column chart between number of products sold and store location 12](#_Toc131624782)

[Figure 13: Purchases made through channels 12](#_Toc131624783)

[Figure 14: Number of complains (2 years) 13](#_Toc131624784)

[Figure 15: Scatter plot between total number of complains and customers 13](#_Toc131624785)

[Figure 16: Overview of customer participation in campaigns 14](#_Toc131624786)

[Figure 17: Response rate for each campaign per store location 14](#_Toc131624787)

[Figure 18: Customers participation per campaign 15](#_Toc131624788)

[Figure 19: Heatmap representing correlations between various features 16](#_Toc131624789)

[Figure 20: Meat and wine correlation scatter plot 17](#_Toc131624790)

[Figure 21: Alpha and Beta calculations 17](#_Toc131624791)

[Figure 22: Executive summary 18](#_Toc131624792)

**Mediterra Market**

Icon

Description automatically generated

# Introduction

Market analysis is a crucial component of any business report as it helps in understanding the current trends and dynamics of the market. It involves collecting and analyzing data related to the market size, growth rate, competition, customer behavior, and other factors that affect the industry. The insights derived from market analysis can be used to develop effective marketing strategies, identify potential opportunities and threats, and make informed business decisions.

In this report, we will conduct a comprehensive market analysis, aiming to gain a deeper understanding of the market landscape by using data-driven insights to improve marketing efforts, boost profitability of customers, increase sales revenue, and provide valuable insights for our stakeholders.

# Overview

The following two subsections provide a brief overview of the dataset[[1]](#footnote-1), and the visualization of its contents for the analysis.

Graphical user interface, application, website

Description automatically generated

Figure 1: Dataset overview

Our retail company operates stores in two locations in Greece, Athens and Thessaloniki, with a customer base of 2,236. Our supermarket provides premium quality products across six distinct categories: Fish, Meat, Wine, Fruits, Sweet, and Household. Over the course of the past two years, we have conducted four marketing campaigns - two per year, one in winter and one in summer. Our overall objective through analysis is to identify the critical factors that influence customers to accept our campaign offers and make purchases.

## Our dataset

The dataset used for the purposes of our analysis, contained both numerical (continuous) and categorical (discrete) features that can be depicted below as follows:

|  |  |
| --- | --- |
| Column name | Explanation |
| Customer ID | Customer identification number |
| Gender | Gender identity |
| Education | Customer’s level of education |
| Marital Status (MS) | Customer’s marital status |
| Store Location (SL) | Athens or Thessaloniki |
| Age | Customer’s Age |
| Income | Customers yearly income |
| Kidhome | Number of small children in customer’s household |
| Teenhome | Number of teenagers in customer’s household |
| Recency | Number of days since the last purchase |
| Last month transaction | If each customer purchased something the last month |
| MntFishProducts | Amount spent on fish products in the last 2 years |
| MntMeatProducts | Amount spent on meat products in the last 2 years |
| MntFruits | Amount spent on fruits products in the last 2 years |
| MntSweetProducts | Amount spent on sweet products in the last 2 years |
| MntWines | Amount spent on wine products in the last 2 years |
| MntGoldProds | Amount spent on gold products in the last 2 years |
| MntTotal | Total amount in spent in all the categories the last 2 years |
| NumDealsPurchases | Number of purchases made with discount |
| NumCatalogPurchases | Number of purchases made using catalogue |
| NumStorePurchases | Number of purchases made directly in stores |
| NumWebPurchases | Number of purchases made through company’s web site |
| Total Purchases | Total number of purchases |
| AvgAmountPurchase | Average amount per purchase |
| NumWebVisitsMonth | Number of visits to company’s web site in the last month |
| Complain | If customer complained in the last 2 years (1=yes) |
| AcceptedCmp1 | If customer accepted the offer in the 1st campaign |
| AcceptedCmp2 | If customer accepted the offer in the 2nd campaign |
| AcceptedCmp3 | If customer accepted the offer in the 3rd campaign |
| AcceptedCmp4 | If customer accepted the offer in the 4th campaign |
| Date | Date purchased for 2 years |

Table 1: Dataset’s columns reference

## PowerBI Context

For the purposes of facilitating our work and to present the results of our analysis, a PowerBi presentation has been created as described below:

1. **Cover page**: Includes the title of the dashboard, the name of the company, and any relevant images or logos.
2. **Agenda**: Provides an overview of the dashboard's purpose, the parts of the analysis and the key findings that will be presented.
3. **Customer Demographics:** Provides a visualization of the customer demographics.
4. **Customer Purchasing Behavior:** Shows statistical data on customer purchasing behavior; allows filtering by product.
5. **Sales Channel Analysis & Complains:** Conducts an analysis on purchasing channels/mediums and provides analytical data on purchase types by year and visualizes frequency of customer complaints.
6. **Campaign Response Analysis:** Measures the customer response to different marketing campaigns by type, year, and customer demographics.
7. **Correlations:** Measures the correlations between our column’s characteristics.
8. **Meat – Wine Correlation:** Measures the relationship between wines and meat categories
9. **Summary/Conclusion page**: Summarizes the key findings of the analysis and provides recommendations for improving marketing efforts and increasing sales revenue.

# PowerBI Analysis

The Power BI analysis for the marketing campaign involves using the data visualization and analysis tool **(goal)** to gain insights into customer behaviour and preferences to inform marketing strategies and improve customer satisfaction. Based on the above, our task was to identify the key factors that influence customers to accept the company's campaign offers and make purchases.

## Pre-processing

To process our data, we used Exploratory Data Analysis (EDA)[[2]](#footnote-2) which is the critical process of performing initial investigations on data sets to discover patterns, to spot anomalies, to test hypothesis and to check assumptions with the help of summary statistics and graphical representations. In other words, EDA is an important step in the data analysis process, as it helps to identify any issues with the data and provides a solid foundation for further analysis.

More specifically, to construct a dataset suitable for further analysis, we performed the following steps:

* Detected and dropped any duplicates
* Detected outliers/imbalances and missing values (if any)
* Summarized descriptive characteristics
* Created 3 new columns:
* MntTotal, which represents total amount of products each customer bought
* Total Purchases, which illustrates the total number of purchases made by customers
* AvgAmountPurchase, which outlines the average number of products purchased
* Deleted 4 rows which included 2 marital statuses (YOLO and Absurd)
* Created data visualizations to better comprehend the data such as scatter plots, histograms, box plots, crosstabs and more.
* Uncovered patterns by calculating the Pearson correlations and plotted the correlation coefficient heatmap
* Performed a linear regression technique by calculating alpha and beta, which are two important parameters in linear regression that help to estimate the linear relationship between two variables and define the equation of the regression line.

## Customer Demographics

This section is dedicated to identifying customer demographics which are a key component in our analysis. A picture containing toy, doll

Description automatically generated

We created a pie chart to show the distribution of customers by gender. Our customer base is predominantly female, with a total of 1342 female customers, accounting for 60.02% of the total. Conversely, the remaining 39.98% comprises 894 male customers.

Chart, pie chart

Description automatically generated

Figure 2: Gender distribution

We created a pie chart to show the distribution of customers by marital status. Our consumer base can be classified into several categories based on their marital status. However, the two most significant groups are married or in a relationship, which together account for approximately 64% of our customer base. Specifically, customers who are married represent the largest group, comprising 38.64%, followed by those in a relationship at 25.94%. Singles, although a smaller group, still account for a noteworthy 21.47% of our customer base, which may warrant consideration in our upcoming campaigns. In addition, we generated a crosstab (contingency table), to summarize and compare the distribution between gender and marital status variables, where the aforementioned information can be depicted.

Chart, pie chart

Description automatically generated Graphical user interface, table

Description automatically generated

Figure 3: MS distribution and Gender/MS relationship

Furthermore, we produced a bar chart to showcase the distribution of customers by education level. Our customers can be classified into five distinct educational categories, with the most prevalent group being those who have completed graduation, representing roughly 50% of our customer base. A significant proportion of customers hold a Ph.D. or a master’s degree. It is noteworthy that only approximately 10% of our customers possess a basic education level, which is a positive indicator of the high quality of our products and their appeal to a more educated customer base.

Chart, bar chart, funnel chart

Description automatically generated

Figure 4: Education distribution

Additionally, we created a histogram to show the distribution of customer age and identify our target audience. At the beginning, the age range lies between 27 and 80. The largest customer segment falls within the age range of mid-40s to late-50s, indicating that most of our customers are middle-aged. However, there is also a noteworthy number of customers in their 40s and 60s. These age distribution patterns align with our unique selling point of providing high-quality and relatively expensive products, which may appeal more to middle-aged and older customers who prioritize quality over price. Therefore, these data are not unexpected.

Chart, bar chart, histogram

Description automatically generated

Figure 5: Age distribution

Moreover, we have generated two tables displaying the distribution of the number of kids and teens at home among our customer base. The analysis revealed that approximately 40% of our customers have one child under the age of 12, while roughly 45% have one teenager at home. This indicates that a significant portion of our customer base consists of families. As sweets are one of our main product categories and are frequently consumed by underage individuals, this is an essential factor to consider. It is worth noting that only a small percentage, around 2-3% of total population, have two kids or teens at home though.

Table

Description automatically generated Table

Description automatically generated

Figure 6: Kids and teen summary

We created a map to visualize the distribution of customers by store location. The company's customers are primarily located in two cities, Athens and Thessaloniki where our stores are located. Most of our customers are based in Athens, accounting for approximately two-thirds of our total customer base. The remaining third of our customers are based in Thessaloniki.

Map

Description automatically generated

Figure 7: Stores locations map

## Customer Purchasing Behavior

This section provides crucial statistical information used to further describe the consumer base and identify the main source of revenue.

Graphical user interface, text, application

Description automatically generated

Figure 8: Product categories summary

The generated table that can be represented below is useful for analyzing sales trends across different product categories and identifying areas for potential growth or investment. It's important to note, however, that sales per product alone does not necessarily reflect profitability or overall success of a product category, and other factors such as production costs and market demand should also be taken into consideration when making business decisions.

In our case, each row represents a product category per store location, while the first column illustrates the average amount of products bought for each category. Upon further analysis of the table, it is evident that the categories of Household, Fruit, Sweet, and Fish products have lower sales in comparison to the Wine and Meat categories, with only around 10% or less of our customers making purchases from these four categories. Further examination of the table reveals that the average sales of Household and Fish products are higher than that of Fruits and Sweets. This indicates that the majority of our customers prioritize purchasing products for their everyday needs rather than adhering to a strictly healthy or unhealthy lifestyle.

Table

Description automatically generated

Figure 9: Average of products sold per store location

It is also noteworthy that 100% of our customers have purchased at least one wine product within the past two years, with a total of 680,816 wine products sold. This category stands out from the other product categories. Moreover, although only one out of four (25%) customers have purchased meat products, 373,968 meat products have been sold, which is more than half the quantity of wine products sold. Considering the above table, it can be inferred that our customers exhibit a proclivity towards purchasing substantial amounts of meat products, as evidenced by the information displayed in the table depicting the average expenditure per product below.

Furthermore, a pie chart can be illustrated below, showcasing the distribution of customers that have purchased products within the last month. Based on the given data, it seems like customer retention should be a key focus of the new marketing campaign, especially considering that only 30,89% of the customer base has made a purchase within the last month. Checking for any complaints could be useful in identifying potential areas for improvement, but it's important to note that low purchase frequency could also be influenced by factors such as competition or changing customer preferences. Therefore, it may be useful to conduct additional research to gain a better understanding of why customers are not making purchases more frequently, and to tailor the marketing campaign accordingly.

Chart, pie chart

Description automatically generated

Figure 10: Last month transactions pie chart

Additionally, we produced a scatter plot outlining the relationship between customer age and the amount spent (how many products bought) on different product categories. Based on the data, it can be observed that customers in their late 20s to early 40s and those above 70 years of age tend to spend more on average. This trend may be attributed to the smaller number of customers in these age groups compared to the 40 to 60 age group, which comprises the largest portion of the customer base. However, it would be advantageous to focus on the younger audience, as they tend to have a greater influence on others and consume more even when not necessarily required. Moreover, considering that the majority of customers are either married or in a relationship, younger customers may eventually have children, which would lead to an increase in their overall consumption as a family.

Chart, scatter chart

Description automatically generated

Figure 11: Scatter plot between age and products sold

Apart from the scatterplot, a clustered column chart to show the products bought on different product categories by store location. Our sales data indicates that the majority of purchases are made in our Athens store, with approximately two-thirds of sales coming from Athens and the remaining from Thessaloniki. Nevertheless, the average amount spent on all categories is nearly identical, with customers in Thessaloniki buying more Meat and Fruit products on average. Consequently, we conclude that increasing our customer base in Thessaloniki could result in higher sales for the company.

Chart, bar chart

Description automatically generated

Figure 12: Clustered column chart between number of products sold and store location

## Sales Channel Analysis

Diving further into our analysis, it was identified that the purchase channels of the company’s products are useful when determining the right marketing strategy and analysis. For this reason, we have generated a donut chart (represented below) showing the percentage of purchases made by customers through different channels. Approximately 65% of our company's sales have been generated through store and web purchases, making these two channels our primary focus for marketing campaigns. Of these, 38.97% of sales were from in-store purchases and 27.49% from web purchases. These figures suggest that opening a new physical store and partnering with online platforms such as Efood or Wolt to sell our products may be worth considering. On the other hand, sales through catalogues accounted for 17.90% of our total sales, indicating that we should not prioritize this channel. Despite being the least significant source of sales at 15.64%, it is still a respectable figure given that we are a new supermarket with high-quality products and limited resources to provide many special deals or offers during the first two years. Hence, incorporating new deals, preferably for new products, into our upcoming campaigns could be a viable approach.

**Chart

Description automatically generated**

Figure 13: Purchases made through channels

Moreover, a pie chart has been generated, depicting the distribution of customers who complained in the last 2 years.

The fact that only 2.5% of our customers have filed complaints within the preceding 24 months serves as a compelling indicator of the exceptional quality of our products and the competence of our store personnel.

Diagram

Description automatically generated

Figure 14: Number of complains (2 years)

We have generated a scatter plot showing the relationship between complaints and the amount spent. A considerable proportion of customers who registered complaints have demonstrated a tendency to purchase a relatively low total value of products. It is plausible to infer that their grievances may have impacted their inclination towards our supermarket. However, it is noteworthy that a small subset of customers who had reported issues did not exhibit any corresponding shift in their purchase patterns, despite having bought a substantial quantity of products.

Chart, scatter chart

Description automatically generated

Figure 15: Scatter plot between total number of complains and customers

## Campaign Response Analysis

The main propagation of our analysis is campaign response section, where the data shown here is a determining factor in our marketing strategy improvement. The bar chart as presented below, provides an overview of customer participation in each of the four marketing campaigns conducted by the supermarket in Athens and Thessaloniki. The data indicates that Campaign 3 (Winter) garnered the highest participation rate in Athens, whereas Campaign 4 (Summer) had the highest participation rate in Thessaloniki. These findings shed light on the campaigns' effectiveness at a geographical level, with Athens exhibiting greater campaign effectiveness overall than Thessaloniki.

Chart, bar chart

Description automatically generated

Figure 16: Overview of customer participation in campaigns

We have also generated a stacked bar chart: to show the response rate for each campaign per store location. The graphs show that Campaign 2 during the summer of 2021 was the least successful while Campaign 3 during the winter of 2022 was the most successful. By examining the characteristics of these campaigns, we can identify factors that contributed to their success or failure and use them to inform our future marketing strategies. It is important to learn from the shortcomings of Campaign 2 and replicate the strengths of Campaign 3 to increase the effectiveness of our future campaigns.

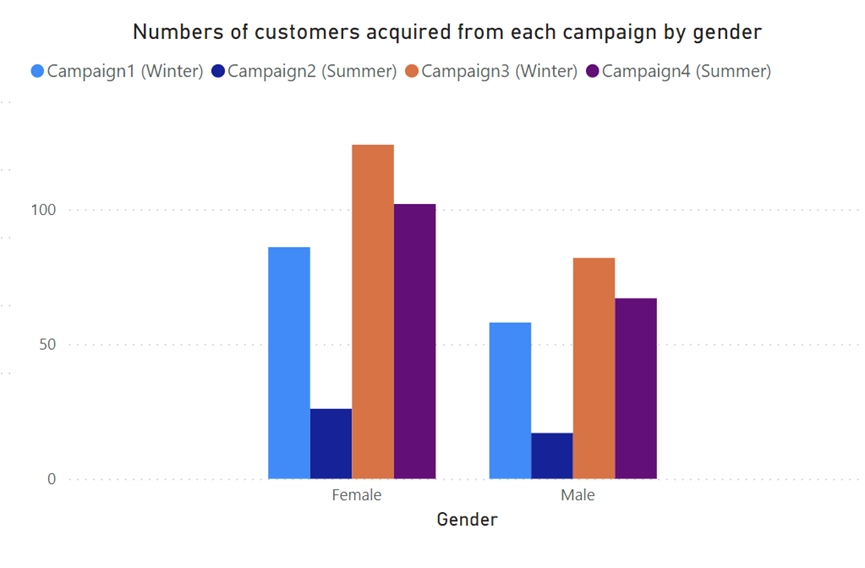


Figure 17: Response rate for each campaign per store location

In the table below, it can be observed that none of the campaigns conducted by our company within the last two years were deemed highly successful, as none of them managed to attract more than 10% of our customer base. It is recommended to not base future campaigns solely on the outcomes of previous ones, but instead analyze the differences between the third campaign and the rest to determine potential areas of improvement. As our sales have shown an increase in the past year, it is imperative to allocate additional funds towards our upcoming campaign to gain a better understanding of our customer demographics and attract new patrons to our store locations.

Graphical user interface

Description automatically generated with medium confidence

Figure 18: Customers participation per campaign

Furthermore, we plotted a line chart to depict the participation of all customers for each campaign based on month and year. The chart has two axes - the horizontal axis, which represents time, and the vertical axis represents the number of customers who participated in the campaigns. There are two lines on the chart, each representing a different year and as the campaigns progress, it is observed that both years, the participation in January is almost identical with December even though the participation in 2022 was slightly increased. At some points (ex. mid-January, mid-June, mid-September etc) , the lines cross one another, and this is an indication that the campaigns of 2021 are more popular than the campaigns of 2022 at that time and vice versa.

Chart, line chart

Description automatically generated

## Correlation Heatmap

In this section, a Correlation Heatmap [[3]](#footnote-3)(type of plot that visualize the strength of relationships between numerical variables) has been created to observe the correlation in between all our factors. This heatmap was produced by plotting it in Python and we can observe a strong positive correlation between wine and meat products, which are the highest purchasing categories. Positive correlations are also observed between the amount of meat and wine purchases and all other product categories, indicating that increasing sales in these two categories may have a cascading effect on other categories.

Chart, treemap chart

Description automatically generated

Figure 19: Heatmap representing correlations between various features

Additionally, there is a strong correlation between the amount of wine products purchased and income, as well as web and in-store purchases. Conversely, most customers prefer to purchase meat products in-store rather than online. Another strong positive correlation that can be depicted is between the number of purchases (meat and wine products) made through catalogs (type of marketing collateral that groups important product details, which can be helpful for buyers making a purchase decision.

Further analysis should be conducted to take into consideration more features and leverage it in our upcoming marketing campaign. Notably, we observe weak correlations between age and all other factors, including both product categories and campaigns.

## Meat – Wine Correlation (Linear Regression)

In linear regression[[4]](#footnote-4), alpha and beta are the parameters used to estimate the linear relationship between two variables.

* Beta (b) is the slope of the regression line and represents the change in the dependent variable for every one-unit change in the independent variable. In other words, it shows how much the dependent variable changes for every unit increase in the independent variable. A positive b value indicates a positive relationship between the two variables, while a negative b value indicates a negative relationship.
* Alpha (a) is the y-intercept of the regression line and represents the value of the dependent variable when the independent variable is zero. Alpha is also known as the constant term or the intercept. Alpha is used to adjust the position of the regression line along the y-axis.

Together, alpha and beta help to define the equation of the regression line, which is used to predict the values of the dependent variable based on the values of the independent variable.

The equation of the regression line is:

**y = α + βx**

where y is the predicted value of the dependent variable, x is the value of the independent variable, (β) is the slope of the line, and (α) is the y-intercept.

Chart, scatter chart

Description automatically generated

Figure 20: Meat and wine correlation scatter plot

Based on the aforementioned information, a linear regression was performed to model the relationship between wine products and meat products, which tend in general to be supplementary products. This linear regression analysis suggests that there is a positive relationship between wine sales and meat sales in the supermarket, with wine sales increasing as meat sales increase. This information could be useful for our marketing analysis in determining how to optimize sales of both wine and meat in the supermarket.

The y-intercept (α) term in the equation, 163,61 represents the predicted value of wine sales when meat sales are equal to zero. In other words, if the supermarket did not sell any meat, the model predicts that the supermarket would still sell 163 units of wine.

The slope term (β) in the equation, 0.84, represents the change in wine sales for every one-unit increase in meat sales. In other words, for every additional unit of meat sold, the model predicts that wine sales will increase by 0.84 units.

A picture containing background pattern

Description automatically generated

Figure 21: Alpha and Beta calculations

# C-Level Executive Summary

Based on the analysis, here are some factors and details to consider for the upcoming winter campaign in 2023:

1. Target the highly educated, married or in a relationship individual, with a majority of female customers aged between 40-60. These customers are the largest proportion of our customer base.
2. It is recommended to consider the location of the customer base in developing a marketing strategy. As the majority of sales are currently generated from Athens, there may be an opportunity to expand the customer base in Thessaloniki, given that customers there tend to spend more on average. This expansion could lead to increased sales and may offer potential for further business expansion.
3. It is advisable to prioritize online sales in the upcoming marketing campaign, given that they currently constitute a significant proportion of our sales and have a high potential for growth. This is attributed to the increasing trend of people opting to purchase goods through online channels, due to the convenience it offers, especially considering their busy schedules.[[5]](#footnote-5)
4. Learn from the previous campaigns, where Campaign 3 in winter 2022 was the most successful. Analyze the factors that contributed to this success and try to replicate them in our upcoming campaign.
5. Pay special attention to wine and meat products as they have a strong positive correlation with higher sales.
6. Develop a strategy to encourage customers who have not made purchases in the last month to visit the store and increase our customer retention.
7. Create an engaging campaign that will excite and motivate customers to make purchases.

By focusing on these factors and details, a targeted and effective marketing campaign for Mediterra supermarket's winter 2023 season can be created.

Chart, bar chart

Description automatically generated

Figure 22: Executive summary

# Extra conclusions

The findings below allow us to draw an analytical conclusion on the best practices that could be used to improve the marketing efforts of the company, forecasting an increase in sales and revenue supported by the data we have gathered.

1. Since the sales are higher in web-based mediums, it can be assumed that customers are tech-savvy and thus the marketing strategy can focus on catering to modern clients.
2. The number of web visits greatly exceeds the number of purchases, which could indicate that the web(page) marketing is succeeding, but not in the product marketing per se.
3. As the amount spent increases, complaints decrease. The cause may be that customers that buy in bulk are less attentive to individual issues. On the other hand, it could be brand loyalty.
4. To increase engagement, the campaign may include a call-to-action that encourages potential customers to increase participation in the campaigns. The campaign may also leverage customer testimonials and reviews to build trust and credibility.

# Appendix

This form has been filled out by all team members after the completion of the group assignment. The team leader has been chosen upon agreement and is responsible to upload the group assignment after its completion and deal with any technical and other issues that might arise during the submission process.

Team Leader name: Lykos Stamatios

Team Leader ID: 273820

Team member name: Lykos Stamatios

Team member ID: 273820

***I herewith express my agreement with the submission of this final version of the group project by the team leader.***

Date: 06/04/2023

Team member Signature: \_\_Lykos Stamatios

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Team member name: Lazarević Miloš

Team member ID: 272036

***I herewith express my agreement with the submission of this final version of the group project by the team leader.***

Date: 06/04/2023

Team member Signature: \_\_\_Lazarević Miloš\_

Team member name: Karatsiolis Christos

Team member ID: 270254

***I herewith express my agreement with the submission of this final version of the group project by the team leader.***

Date: 06/04/2023

Team member Signature:



1. https://www.kaggle.com/datasets/rodsaldanha/arketing-campaign [↑](#footnote-ref-1)
2. https://en.wikipedia.org/wiki/Exploratory\_data\_analysis [↑](#footnote-ref-2)
3. https://vitalflux.com/correlation-heatmap-with-seaborn-pandas/ [↑](#footnote-ref-3)
4. https://en.wikipedia.org/wiki/Linear\_regression [↑](#footnote-ref-4)
5. https://www.weforum.org/agenda/2021/07/global-consumer-behaviour-trends-online-shopping/ [↑](#footnote-ref-5)