

Quantitative Modelling of Business Problems: Analysing Sales and Marketing Data

I was assigned to work on 'get to know your products' questions 3, 4, and 5, as well as 'Understand your marketing efforts' question 1. The questions can be found in the file 'Mini Project 1: Quant. Modelling Of Business Problems [PDF]'.

To solve the questions regarding products, I had to refer to the data in the provided Excel file 'sales', while for the 'marketing' questions, I had to refer to the Excel file 'marketing'. The information about each column header can be found in 'Mini Project 1: Quant. Modelling Of Business Problems [PDF]'.

Get to know your products:

3. What is the distribution of sales by distribution channel?

For this question, I needed to determine the purchasing method for each product. According to the provided data, there are five products (wines, fruits, meats, seafood, and sweets) and three ways of purchase: website, store, and catalogs.

I started a new sheet called 'Answers' to answer the remaining questions. Then, I created a pivot table of 'sales' on the 'Answers' sheet to find out which channel each product was sold through.

I began with Wines. From the PivotTable fields, I selected Wines, Website, Catalog, and Store, resulting in a table with four columns: Sum of Wines, Sum of Website, Sum of Catalog, and Sum of Store. However, the results were not accurate because the sum of purchase channels showed the total of all products, even when Wines were not included. So, I went back to the PivotTable fields and added wines to the filters. From the filter drop-down menu, I unchecked the values -621 and 0 to exclude them from the sum when wines were not being purchased.

I repeated the same process for the remaining products: Fruits, Meat, Seafood, and Sweets.

From the tables, I concluded that the highest number of purchases was for Wines through in-store purchases, and the least purchases were for wines through catalogs. However, the product with the highest sales across all channels was Wines.

4. What's the average time since the last purchase?

To calculate the average, I used the AVERAGE function. In the dataset, there is a column called 'Recency,' which represents the number of days since the last purchase.

On the 'Answers' sheet, I typed the function =AVERAGE(sales!B:B), which calculates the average from column B, representing 'Recency' in the 'sales' sheet. The calculation resulted in an average of 49.025 days.

5. Based on your understanding of the sales data, do you have an idea of which products the company may want to focus on?

The expected answer would be to focus on the products with lower sales. However, as I do not have experience in marketing or running a business, I researched on the internet and found slightly different ideas. It is suggested that a business should focus on the products that generate the most sales to ensure substantial revenue. However, they should also address the products with lower sales to identify the reasons behind their poor performance. For example, it could be due to high prices compared to product quality, a lack of proper marketing plans, or a lack of variety.

In our case, Wines were the highest generating product, while Sweets had the lowest sales. I would suggest that the company take steps to increase sales of Sweets.

Understand your marketing efforts:

Before answering this question, I combined all the sheets (sales, marketing, and customer) into one Excel sheet. To do this, I copied the 'sales' sheet into a new sheet and added the headers. Then, I used the VLOOKUP function to fill each column with the appropriate data based on the ID column. For example, to copy the 'Recency' column, I entered the function =VLOOKUP(A2, sales!\$1:\$1048576, 13, FALSE) in the header cell.

1. Which campaigns generated the most and least interest?

For this question, I did not need to use the 'Combined data' sheet. Instead, I calculated the number of purchases for each campaign to determine which one generated the most and least interest. I used the COUNTIF function, which counts cells based on a condition. In this case, I counted how many cells had the value 1 (indicating a customer made a purchase based on the campaign) or 0 (indicating no purchase based on the campaign).

To count purchases based on Campaign 1, I used the function =COUNTIF(Marketing!E:E, 1), which resulted in 143. I performed the same calculation for Campaigns 2, 3, 4, and 5.

The calculations showed that Campaign 3 and Campaign 4 generated the most interest equally, while Campaign 2 generated the least interest.

Summary:

This project demonstrated the importance of understanding data and how it can provide valuable insights for a business's progress and problem avoidance. For example, in question 3, 'What is the distribution of sales by distribution channel?', we were able to determine the preferred purchase channels. In fact, by analysing the 'Combined data,' we could identify the preferred purchase channels based on age or education level.

This project highlights the power of data. It is not surprising that some companies spend a fortune to acquire consumer data from third parties.