

Business Analysis Report

Xue Ming Wang

Introduction

The data set is detailed order information from X company. By analyzing this data set, the business insights will be made in the direction of the sales and profit of the company and the customer behavior.

Data

The data set contains five excel tables, where 9994 orders have been recorded from different customers, locations in the USA from the beginning of 2014 till the end of 2017. It provided the orders information, customers information, products information, and returned order information. The order information includes row IDs, order IDs, order dates, ship dates, and four ship modes (same day, first-class, second class, standard class) locations, sales, quantities, discounts, and profits. The locations are detailed to 531 different cities, 49 different states, four different regions (south, central, east, west), and postal codes. The customer's information contains customer ID, 3 segments (home office, corporate, consumer), and customer name. The products consist of three main categories, and 17 sub-categories, and all of the product names. Meanwhile, the data set also contains a returned order table, where 297 orders return products returned to the company.

Overall view of sales and profit of the company

Data Analysis:

The annual sales and profit: The sales (Fig.1) and profit are constantly increasing year by year, except for the sales between 2014 to 2015, but there is a massive increase in sales between 2015 to 2016.

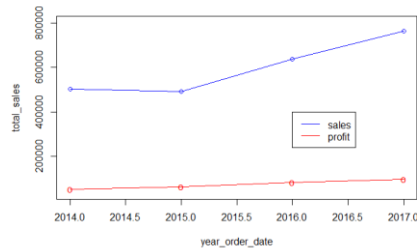


Fig.1 Annual Sales and Profit

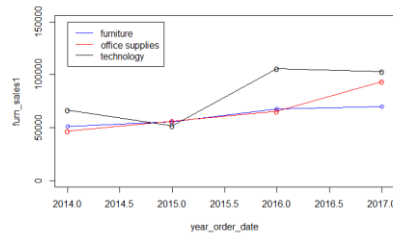


Fig.2 First half year sales

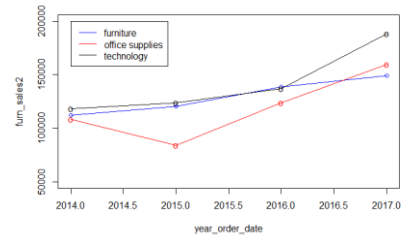


Fig.3 Second half year sales

Half year comparison: When the time from January to June compared to the time from July to August, the first half of every year had a much lesser sales and profit when compared to the second half of all years. The number of customers is similar as the first half-year has 704 and the second half-year has 783, which only increased by a little. However, the number of orders increased from 1770 to 3239. When looking separately at each category with a half-year range, the first and second half-year sales are similar in categories, and the second half-year (Fig.3) is generally higher than the first half-year (Fig.2). The technology has the highest sales, office supplies is the second-highest sales, and furniture has the lowest sales most of the time, not too many changes. For the profit, the furniture is the lowest for both half-year. In the first half of the year of 2014 and 2015, office supplies have a higher profit than technology, but the situation changed back to technology as the highest profit in 2016 and 2017. In the second half-year, technology is still the highest except for 2016. Categories are not the reason for high sales and profit in the second half-year. However, I noticed that the customers (total of 2969 orders) that returned for more shopping happened mainly in the second half-year rather than the first half-year (total of 1293 orders). That is the reason for higher sales and profit in the second half-year.

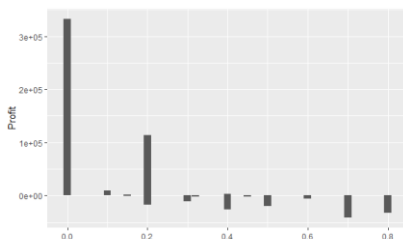


Fig.3 Profit in different amount of discounts

The reason for low profit: Discounts is the most serious factor for low profit. The discount would bring up the sales but lower the profit. Out of all the orders, 4961 orders do not have discounts, and 5370 orders have discounts. With the orders that have a discount,

the most frequent discount is 0.2, which is 20% (Fig.4). For the discounts greater than 20%, the discounts all had a negative profit. For the sales and profit trends every year in discounts, the trend showed that when discount applied, the sales were increasing every year, whereas profit was going the other direction, which produced negative profit as more discounts were given out to the customers.

Business Insights:

The annual sales and profit: From the annual sales and profit analysis, the company's business is gradually becoming stable in an increasing trend.

Half-year comparison: For the first half of every year, the sales and profits are lower than the second half of every year. The high sales and profit in the second half of the year are because the customers usually returned to shop again in the second half-year. Therefore, the customer behavior is meaningful and could be considered to do a deeper analysis.

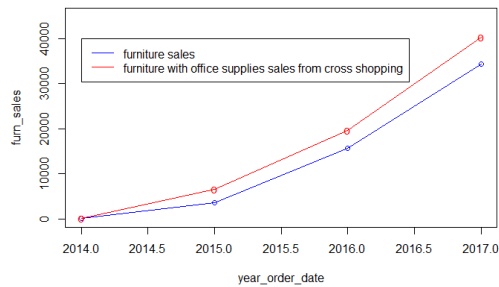
The reason for low profit: The largest problem for low profit is the distribution of discounts. The best discount to give out that could increase the sales and profit at the same time is 0.2, 20%. It does not worth giving out a discount greater than 30%, which would decrease the profit significantly. In order to increase the profit, fewer discounts should be offered next year. As there are more discounts provided to the customers, the profit could be lesser, resulting in a huge amount of loss and hardly any profit. Therefore, the amount of discount has to be handled carefully to catch more customers and to increase the profit at the same time.

Analysis of customer behavior

Data Analysis:

The product in three categories: The X company sells the products in three categories: furniture, technology, and office supplies. There are 17 sub-categories, where 4 sub-categories in furniture, 4 sub-categories in technology, and 9 sub-categories in office supplies.

The annual sales and profit in categories: Over the years, all three categories increased by a certain amount in sales, profit, and the number of orders. The furniture category always has low sales and profit overall, and technology always has the least orders over the four years out of all categories.



←Fig.5 Line graph of the sales of furniture itself and the total of furniture and office supplies sales from furniture cross-shopping.

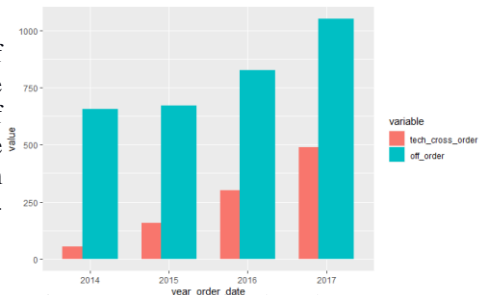


Fig 6 Bar graph of orders that cross to office supplies from technology.

Cross category shopping: Looking from another perspective, many customers started shopping in one category, but they crossed over and bought products from another category. With the segment of costumers in consumers as it is the largest segment, there are 409 customers, and 400 had the experience of cross-shopping. When crossing from furniture to office supplies, more profits are achieved (Fig.5). The technology category has the least orders, but 138 customers started with technology and then went over to office supplies. Also, the technology category guided consumers to office supplies to order more (Fig.6). The customers generally returned sooner from technology to office supplies rather than from furniture to office supplies. Meanwhile, most customers would come back within one year to buy in the office supplies category for both furniture and technology categories.

The overview of the office supplies category: The office supplies category deserves an in-depth decomposition as it is the best out of all categories. The total sales, total profit, total customers, total orders, average expending per customer, and average orders ordered by customers are all increasing, where 2017 has the highest out of other years.

Group by the number of orders: When grouping by orders, the customers could divide by the number of orders they have brought. Based on the number of orders, the customers could divide

into four groups; group 1 is between 1 to 5 orders, group 2 is between 6 to 10 orders, group 3 is between 11 to 15 orders, group 4 is greater than 15 orders.

Group by the quality of the customers with the RFM method: Additionally, the RFM method could conduct a deeper analysis of the customers. R stands for recency, which means the most recent of the last consumption time, F stands for frequency, which means the frequency of customer purchases, and M stands for monetary, which means the total amount over a period of time, over four years here. By building the RFM model, the value of recency, frequency, and monetary could be rated for each customer. When adding each individual value of recency, frequency, and monetary all up, the total value is the raw score; the higher the score is, the better the customers.

Detailed analysis with RFM values: A more detailed analysis could be done with the customers when examining RFM values with an aggregate score of $\text{recency} * 100 + \text{frequency} * 10 + \text{monetary}$. Then, the customers are divided into eight groups.

Business Insights:

Cross shopping: Out of all categories, the office supplies category is the best out of all three. Even though the furniture category had low sales and profit, and technology does not have as much order as office supplies, they are the categories that are essential to keep in order to expand the office supplies. Thus, both categories could not be eliminated.

The overview of the office supplies category: The sales, profit, and order numbers of the office supplies category are increasing every year. When grouping by the number of orders that each customer ordered, group 2 (6-10 orders) which is the customers who ordered 6-10 orders, is generally better than other groups in most of the aspects, such as sales, profit, number of orders, etc. These customers are good in shape, and the company does not need to pay any more attention to them. On the other hand, in group 4 (greater than 15 orders), the customers who ordered greater

than 15 orders have the worst performance out of all groups since there are only 44 customers there. However, each customer's average sales and profit are higher than other groups, which means each customer in this group ordered more and spent more than other groups. It might be beneficial if the company could find more customers like that, so this group's performance could be improved.

Group by the quality of the customers with RFM method: With RFM raw score, all customers are put into four groups: best customers, important customers, general customers, and at-risk customers. The customers from different groups should be treated differently as the best customers are the top priority, which needed special attention; next is important customers, and then the general customers and at-risk customers.

Detailed analysis with RFM values: More profoundly, the customers are divided into eight groups with the RFM aggregate score to analyze their qualities. All types of customers and the action that the company needs to take to these customers are all in the table 1.

Table 1. Types of customers and the actions for company.

Type of customers	Needed actions for the company
Low spending inactive old customers	Shopped more than 5 orders or last order is during 2016 and 2017: keep in touch, others: ignore.
High spending inactive old customers	Re-engage based on their preferences from previous and give discounts.
Low spending active old customers	Incentivize to increase spending, such as giving out discounts.
High spending active old customers	Re-engage based on their preferences from previous and give discounts.
Low spending inactive new customers	Encourage them to spend more and shop more by introduce other products in another category.
High spending inactive new customers	Entertain with products or introduce them to another categories.
Low spending active new customers	Incentivize to increase spending, such as giving out discounts.
High spending active new customers	Maintain and make them feel they are the top priority.