Visual Analytics with Tableau

Business Problem

The director of Classic Models company, Diane Murphy is very much concerned about the decline of the company sales in the recent years and she is hoping to reach a solution through data analytics. The objective of this project is to analyse the sales data by conducting in-depth visual analytics with Tableau to uncover meaningful insights that will help to improve the sales revenue of the company. To meet the above objectives, several key questions need to be answered and clarified:

1. How were the previous year sales?
2. What are the best performing products and worst performing products?
3. What are the most profitable products/product lines?
4. Which customers have not been making second purchase for a long period of time?
5. How is the sales performance in all branches?
6. Does number of sales representative affect number of order/total sales?

Data Source

The data sets used for the analysis is a small subset from the database of Classic Models, a retailer of scale models of classic cars. The database contains sales information from the year of 2003 until 2005. It consists of the following tables:

Customers: contains customer informations (13 columns, 122 rows)

Products: contains car model informations (9 columns, 110 rows)

ProductLines: contains a list of product line categories (4 columns, 7 rows)

Orders: contains sales orders placed by customers including customer number and order date (7 columns, 326 rows)

OrderDetails: contains sales order details for each order including product code, price and quantity ordered (5 columns, 2996 rows)

Employees: contains employee information (8 columns, 23 rows)

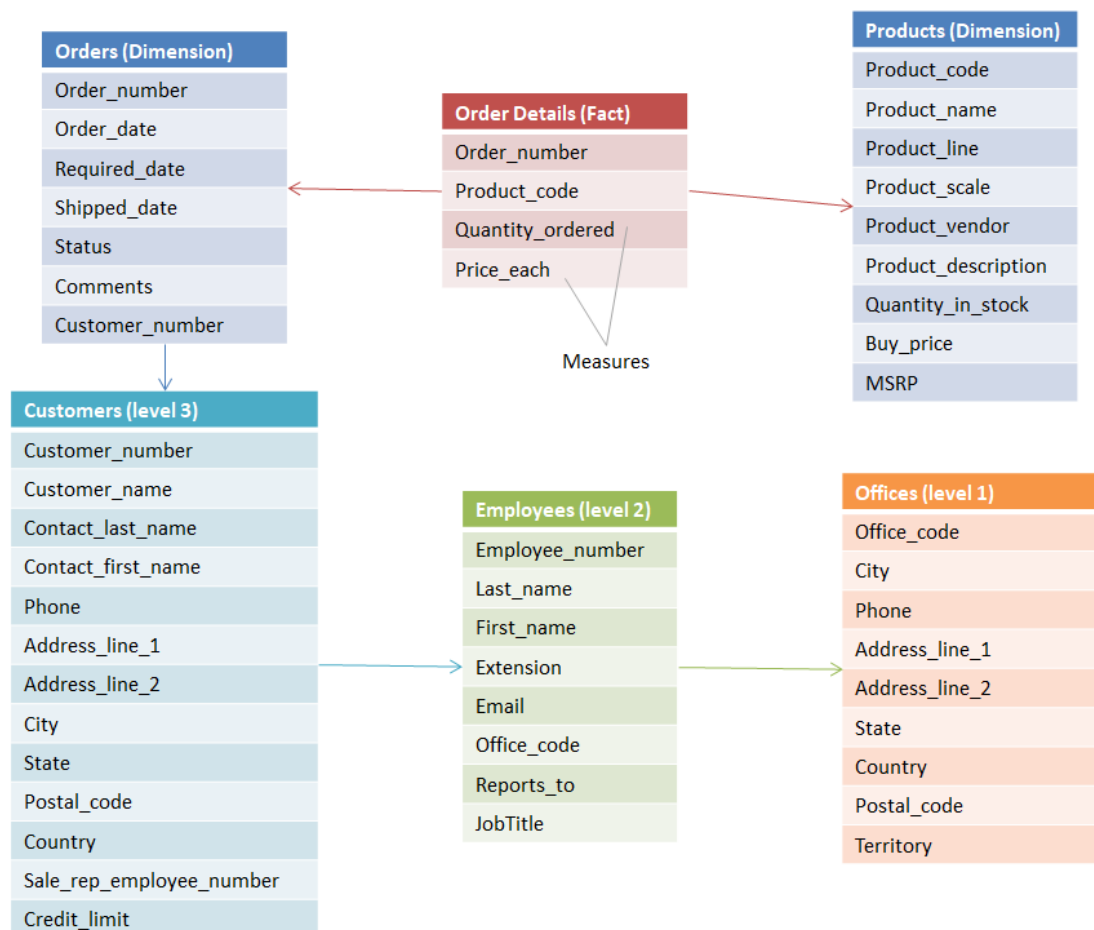*Offices*: contains sales office information (9 columns, 8 rows)

Snowflake Schema

Orders (Dimension)
Order_number
Order_date
Required_date
Shipped_date
Status
Comments
Customer_number

Order Details (Fact)
Order_number
Product_code
Quantity_ordered
Price_each

Measures

Products (Dimension)
Product_code
Product_name
Product_line
Product_scale
Product_vendor
Product_description
Quantity_in_stock
Buy_price
MSRP

Customers (level 3)
Customer_number
Customer_name
Contact_last_name
Contact_first_name
Phone
Address_line_1
Address_line_2
City
State
Postal_code
Country
Sale_rep_employee_number
Credit_limit

Employees (level 2)
Employee_number
Last_name
First_name
Extension
Email
Office_code
Reports_to
JobTitle

Offices (level 1)
Office_code
City
Phone
Address_line_1
Address_line_2
State
Country
Postal_code
Territory

Figure 1: Snowflake schema

Data Visualization and Analysis

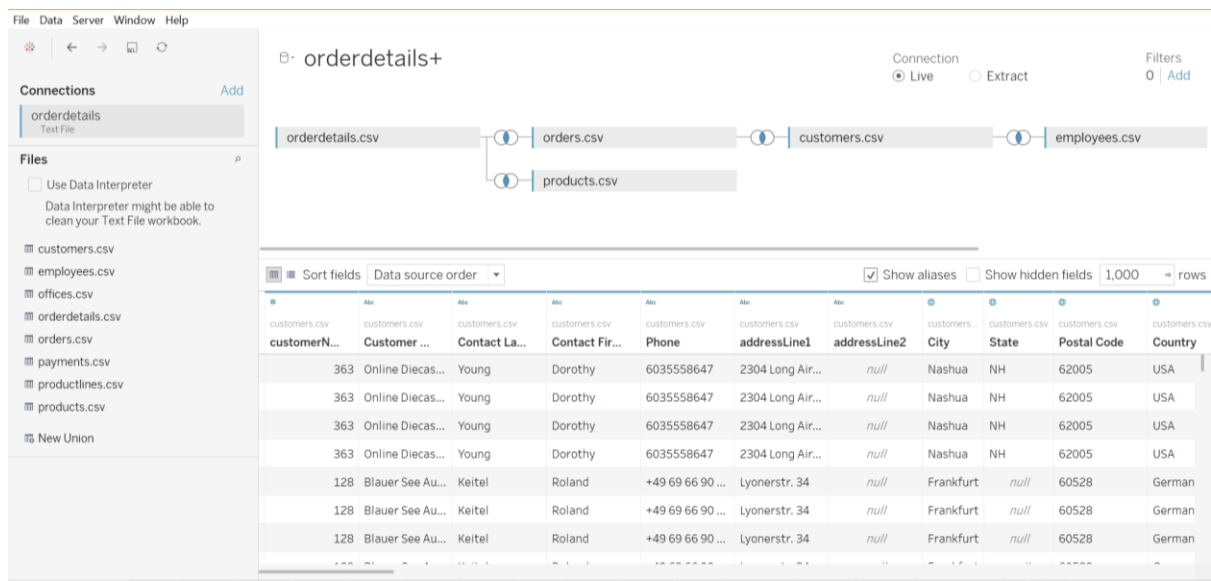A preview of the loaded data sets:

Figure 2: Data preview

How were the previous year sales?

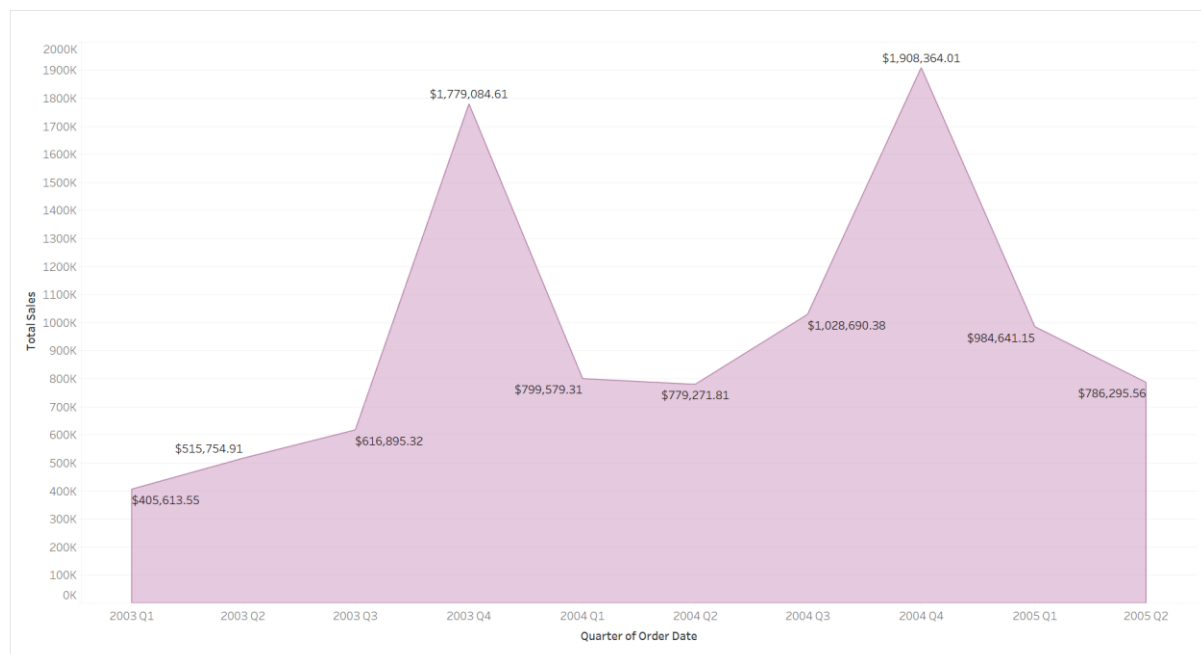An area chart of total sales by quarters of year was plotted as shown in below figure.

Figure 4: Total sales by quarters

Adding up the numbers from first quarter to fourth quarter would give total sales of $3,317,348.39 in 2003 and $4,515,905.51 in 2004. In the first half year of 2005, the company has earned a total sales value of $1,770,936.71 which indicates that the company need to earn at least 400K in the next half year to achieve the target of 30% sales growth from 2004. An interesting trend is also

discovered from the chart where the sales rose dramatically during the fourth quarter of both year 2003 and year 2004.

What are the best performing products and worst performing products?

We want to see how well or how bad the products are performing in overall for future sales strategy planning. Figure 5 is a packed bubbles chart of total quantity ordered for each product line. Every bubble and colour represents a different product line category.
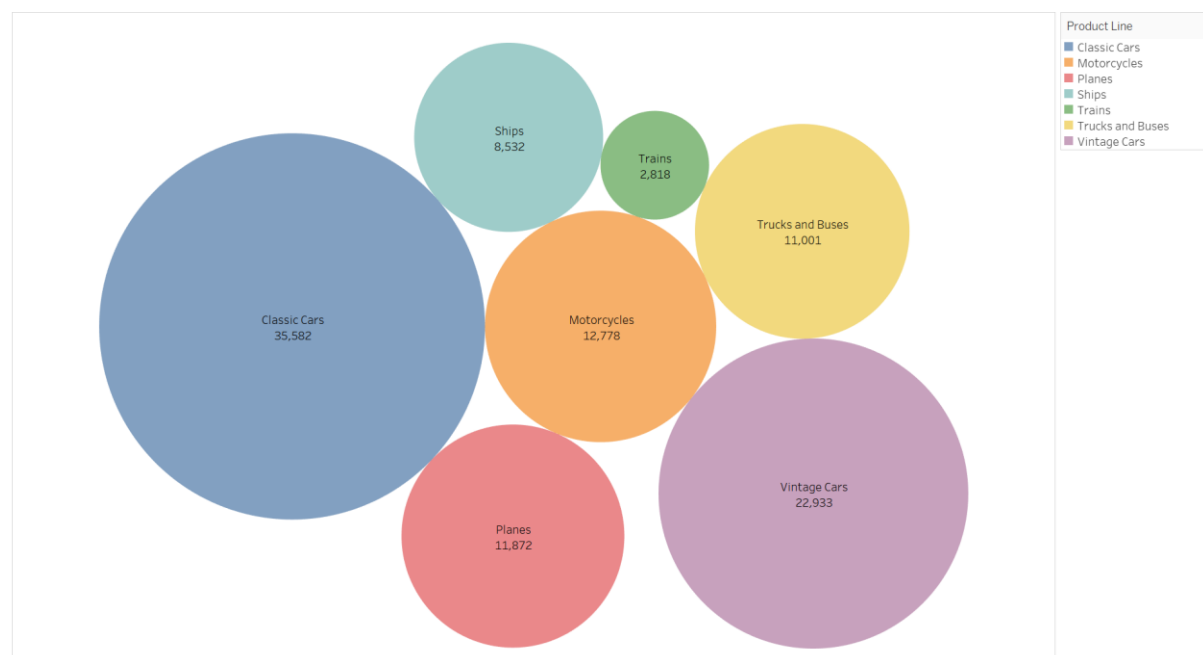

Figure 5: packed bubbles chart of product line with total quantity ordered

According to the bubble chart above, the classic car and vintage car product line performed the best with total 35,582 and 22,933 products ordered, followed by motorcycles (12,778) and planes (11,872), while the worst performing product line is trains with only 2,818 product ordered in total.

A horizontal bar graph is a great way to see which products are the most popular among the customers. A spotlight named 'sales spotlight' was created to separate data into two categories using a threshold of 950, where products above the threshold will be in 'good' category and below it will be in 'bad' category. With the use of color mark, the 'good' results are highlighted in orange indicating a higher popularity and 'bad' results are highlighted in blue indicating a lower popularity. In figure 6, the result of the graph has been sorted in descending order so that product with a higher quantity ordered will be placed on top.
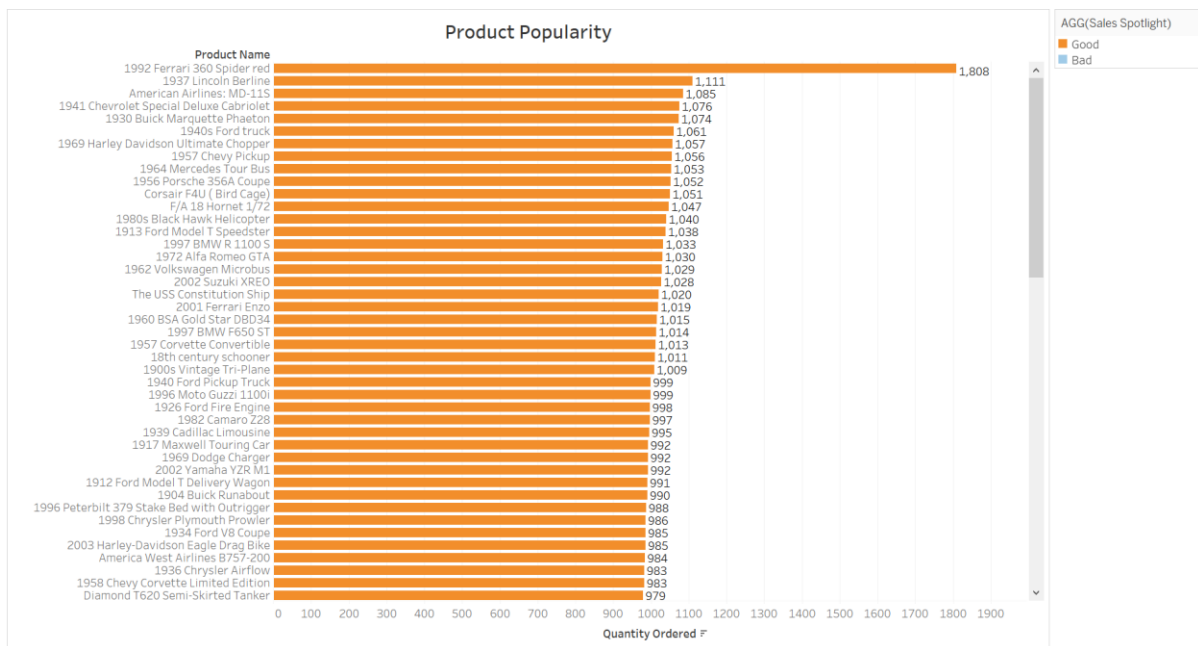
Figure 6: Product Popularity

1992 Ferrari 360 spider red is customers' favourite car model with a great amount of 1808 sold, while the second most popular car model, 1937 Lincoln Berline has 1,111 sold, which is only 63% of the first. There hasn't been much difference in the popularity of the rest of the products.

To view the products that are not performing well, the result was sorted again in ascending order to show the less popular products on top as shown in figure 7.
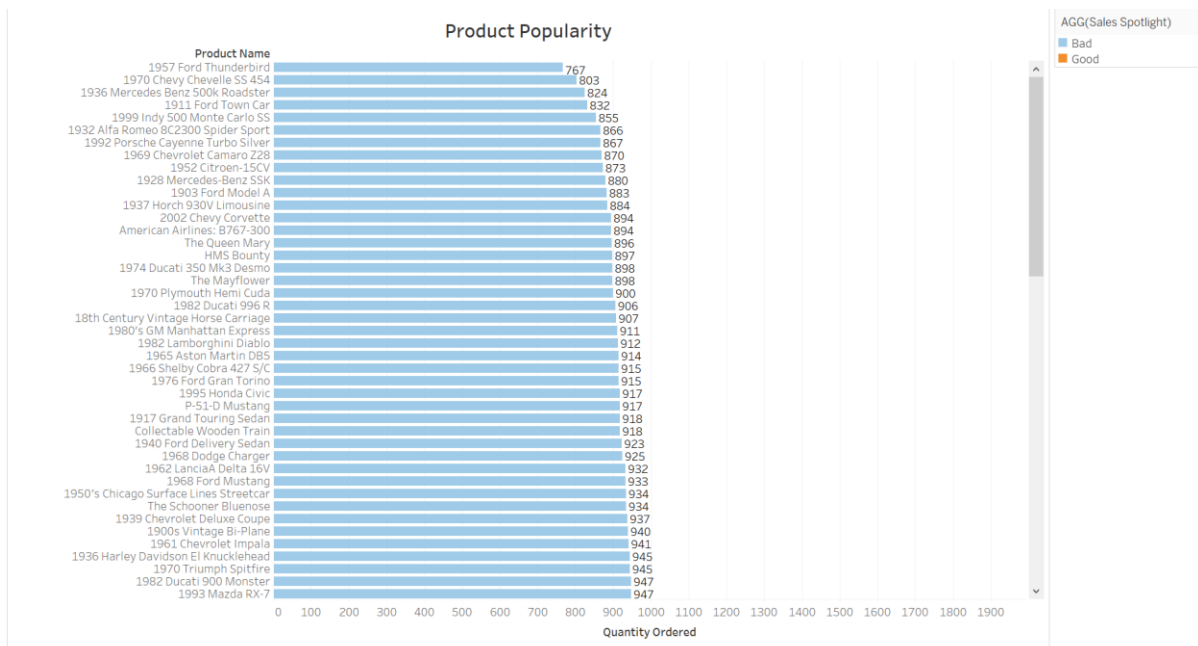

Figure 7: Product Popularity

The 1957 Ford Thunderbird, 1970 Chevy Chelivie SS 454 and 1936 Mercedes Benz 500k Roadster are the top three least popular products among the customers, with an average sale of 798 units.

In any business, it is important to know how much stock should be kept to avoid stockout or overstock. Stock to sales ratio is the ratio of quantity in stock for sale versus the quantity sold. A calculated field was created to calculate the stock to sales ratio. In figure 8, a tree map was plotted to identify products that are having high stock to sales ratio or having large quantity in excess of demand. Products are displayed in rectangles proportional to the amount of ratio they represent. A darker and bigger rectangle indicates a higher ratio, and a lighter and smaller rectangle indicates a lower ratio.
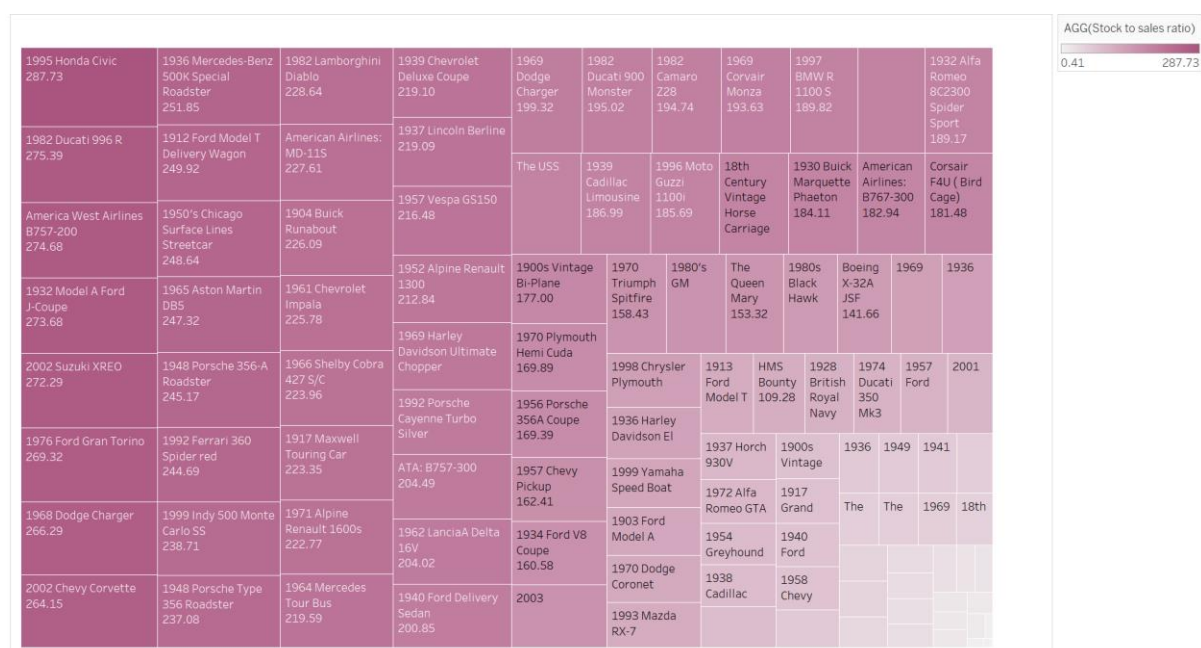
Figure 8: Stock to sales ratio tree map

There are total of 71 products with stock to sales ratio above 100 and 33 products with ratio above 200, with 1995 Honda Civic having the highest ratio which is 287.73. These products are classified as slow-moving products.

What are the most profitable products?

Both revenues and profits are significant measures of a business's success. To maximize both, the company should track not only profit, but also profit margin and focus on selling high profitable products.

A bullet graph was plotted to explore average cost and average selling price for each product, where the grey area representing average selling price and the blue area representing average cost of the product. Labels were added to show the profit margin.
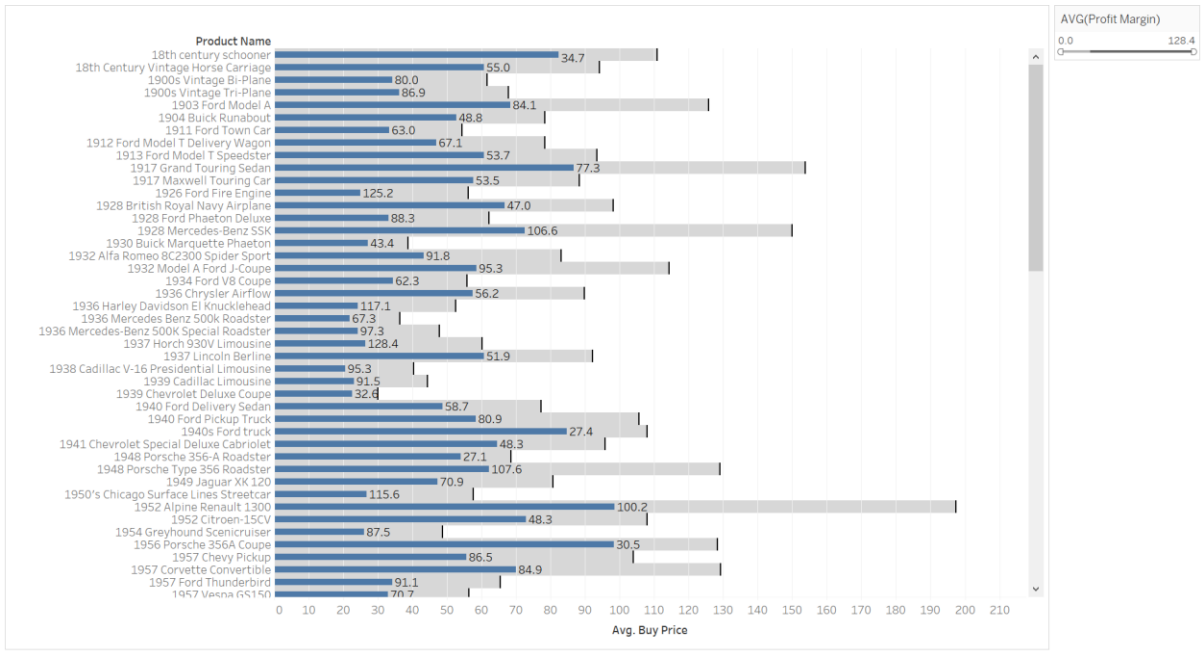

Figure 9: Bullet graph of average cost and average selling price

In figure 10, profit margins are visualized in a highlight table and the result is sorted in descending order and filtered to display only products with profit margin of at least 90%. Rows in darker highlight represent higher profit margins.
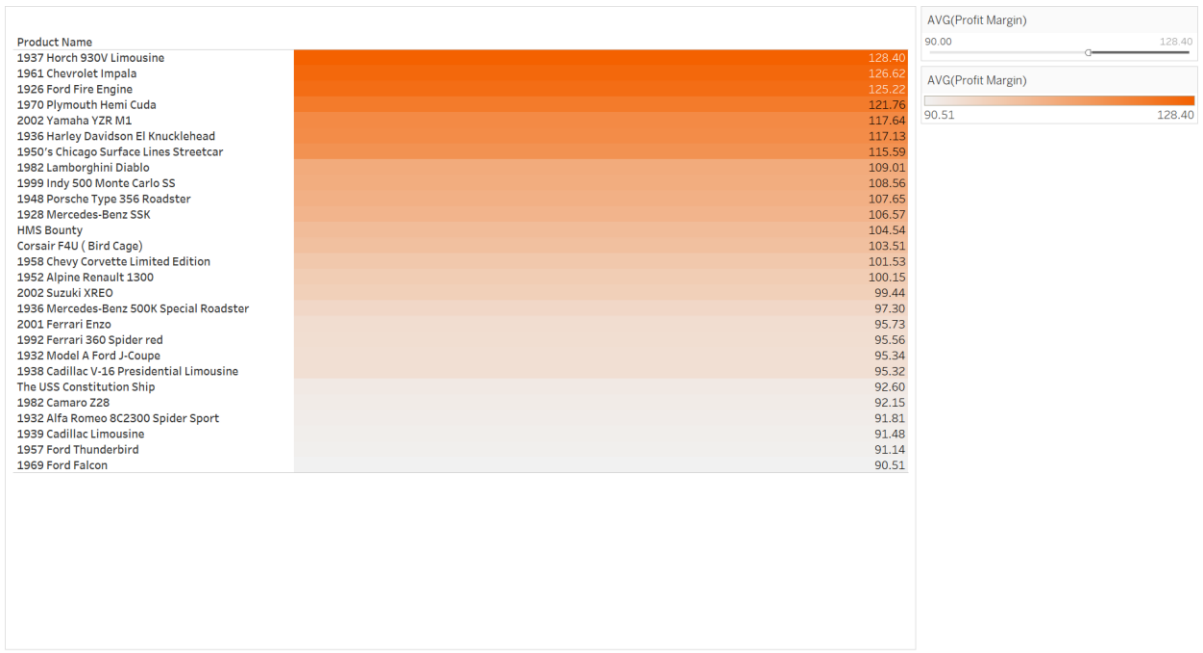
Figure 10: profit margin highlight table

Three products with the highest profit margin are 1937 Harch 930V Limousine (128.40%), 1961 Chavrolet Impala (126.62%) and 1961 Ford Fire Engine (125.22%).

Apart from high profitable products, we want to filter out products that are both slow-moving and low profitable. To achieve this, slow-moving products was grouped into a set using create set function by highlighting the products with stock to sales ratio above 200 from the tree map as shown in figure 8.
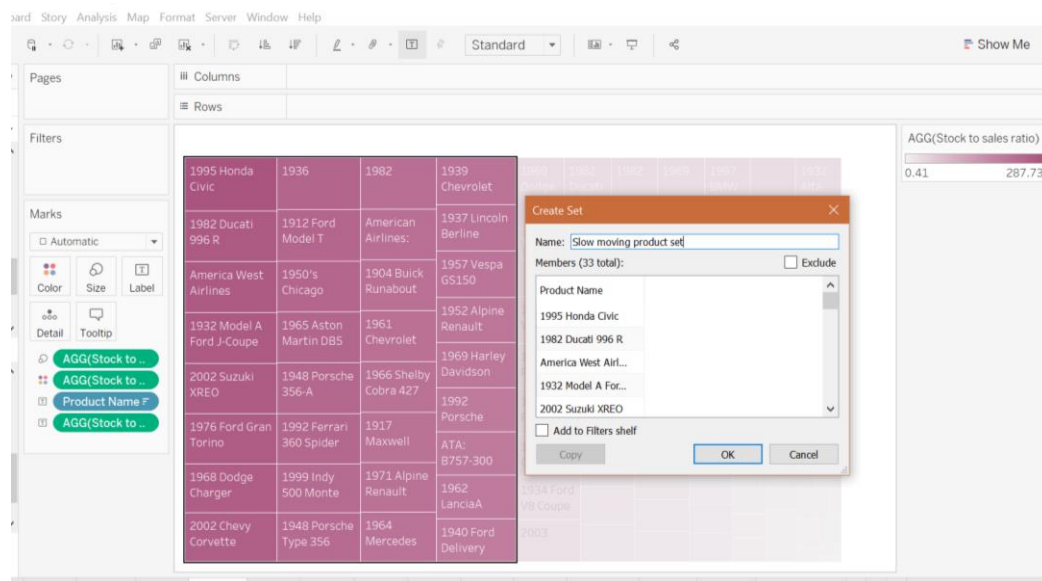

Figure 11: create slow-moving product set

Using the same highlight table in figure 10, the product name row was replaced by slow-moving product set. The result was sorted and filtered to display products only with profit margin below 50%.
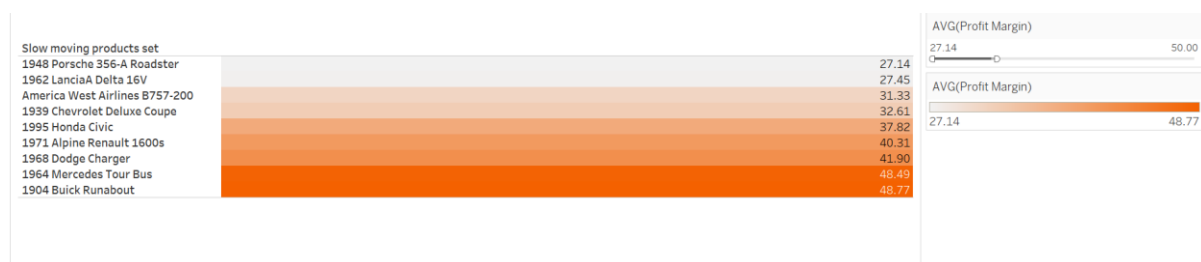

Figure 12: profit margin highlight table of slow-moving products

There are a total of 9 slow-moving and low profit products.

Which customers have not been making second purchase for a long period of time?

Since acquiring new customers can be expensive, we want to ensure that existing customers are making repeat purchases. The highlight table in figure 13 was plotted to show how many quarters

elapsed before the customers make another purchase. Each colored rectangle represents the number of distinct customer and a darker color indicates larger number of customers. The value lapsed indicates the customer did not make a second purchase.

FIXED expression was used to find the first and second purchase dates for each customer and derive the number of quarters to make a repeat purchase from them.
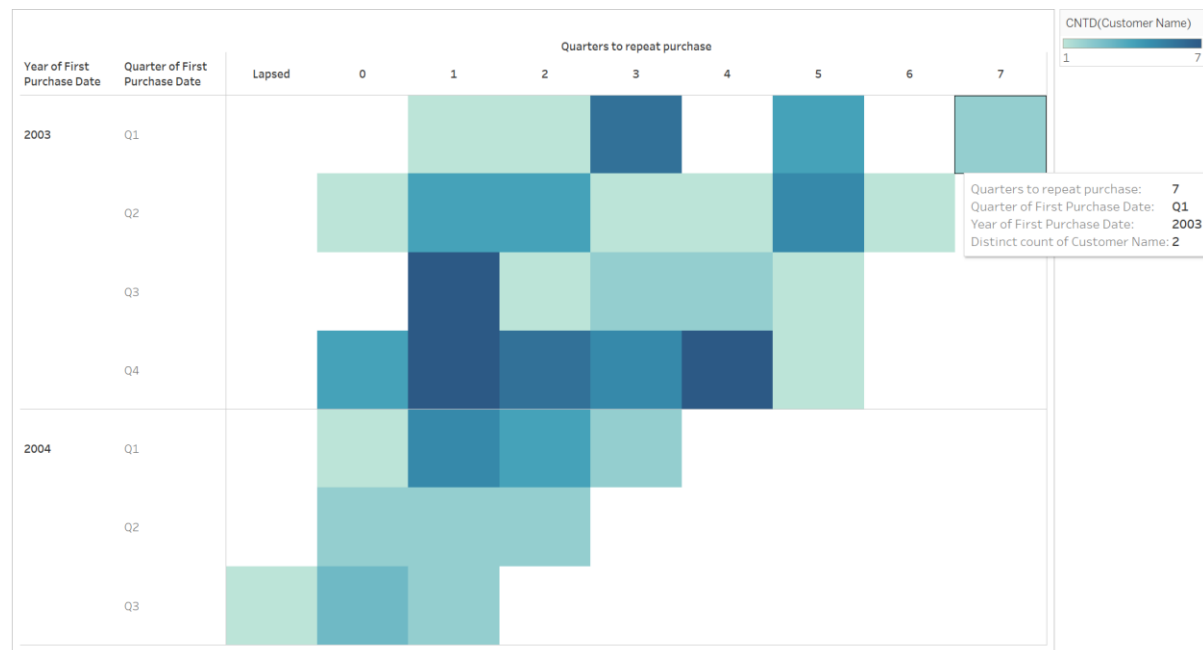

Figure 13: highlight table of customer repeat purchase

2 customers who have made their first purchase in first quarter of 2003 have waited 7 quarters to make their second purchase. 1 customer who made his first purchase in the third quarter of 2004 did not make any repeat purchase since then. There is no new customer acquired since the third quarter of 2004. It can be concluded that most customers make their second purchase after 3 months to one year.

How is the sales performance in all branches?

A map view in figure 14 shows each country's revenue contribution to the overall sales of the company. The countries are colored by the total sales value.
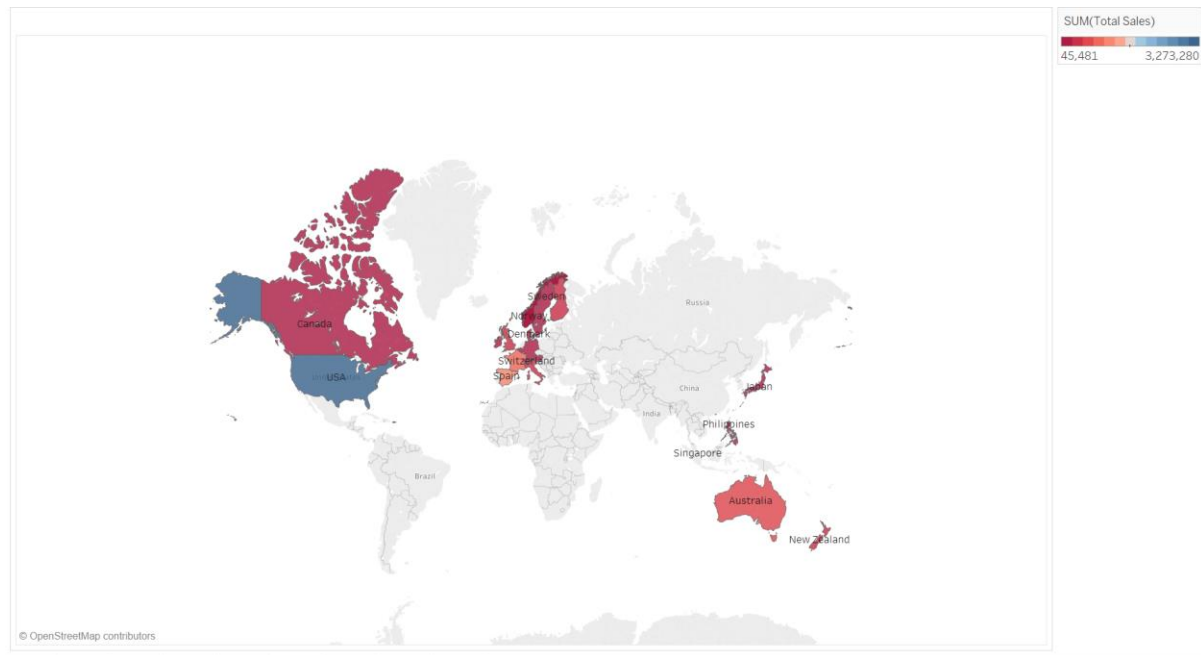
Figure 14: map view of sales contribution by country

It is noticeable that USA has the highest sales contribution and Asia has lower sales contribution in overall compared to other continents.
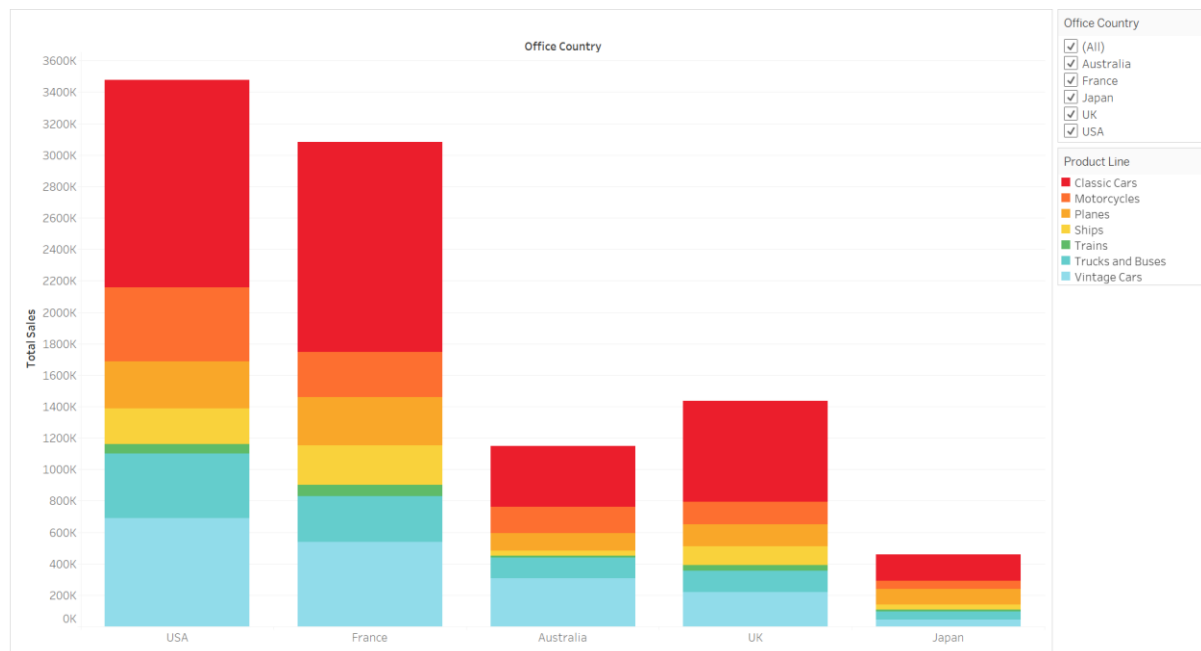

Figure 15: Histogram of sales performance by office branch

The sales representatives in USA branch have generated the highest sales, followed by France, while sales representatives in Japan performed the least satisfactory with only about 500K of sales.

Does number of sales representative affect sales?

The text table in figure 16 shows the number of sales representatives in each branch and their sales performance in 2003 and 2004. Unsurprisingly, countries with higher number of sales representatives generates more sales.

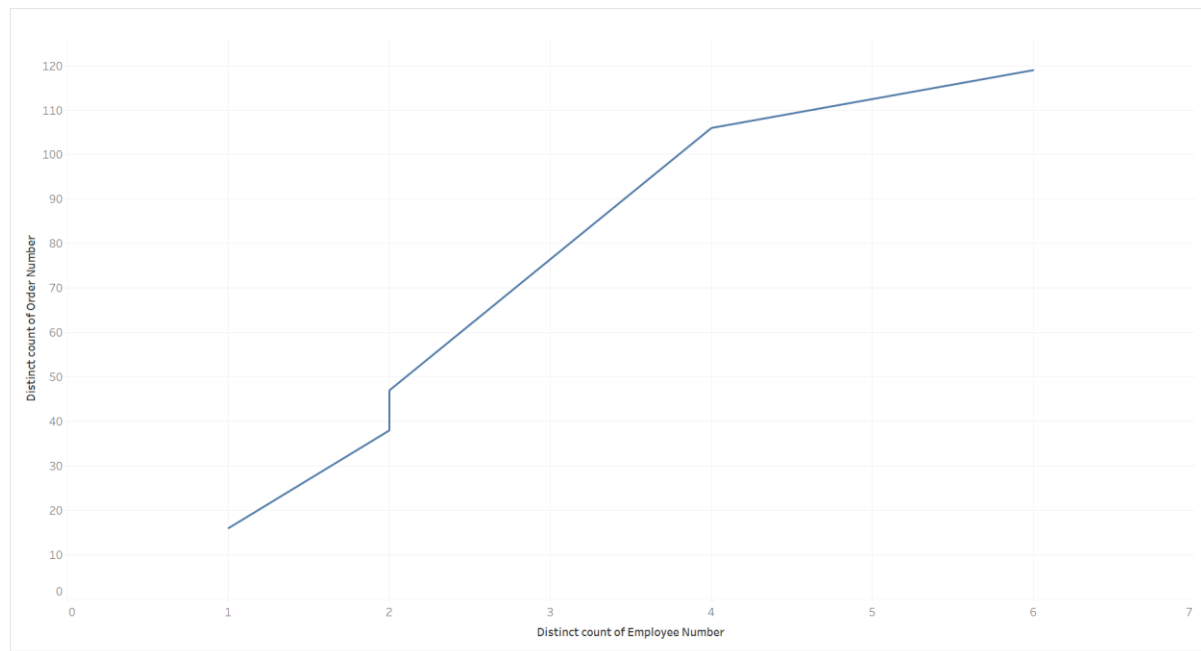|	Order Date			
	2003		2004	
Office Co..	Distinct count of Employee Number	Total Sales	Distinct count of Employee Number	Total Sales
Australia	2	304,949	2	542,996
France	4	969,960	4	1,465,230
Japan	1	267,249	1	151,761
UK	2	549,552	2	706,015
USA	6	1,225,638	6	1,649,904

Figure 16: text table

Figure 17: line graph of number of orders versus number of sales representative

Seem likes there is a positive linear relationship between the number of sales representative and the amount of deals closed.
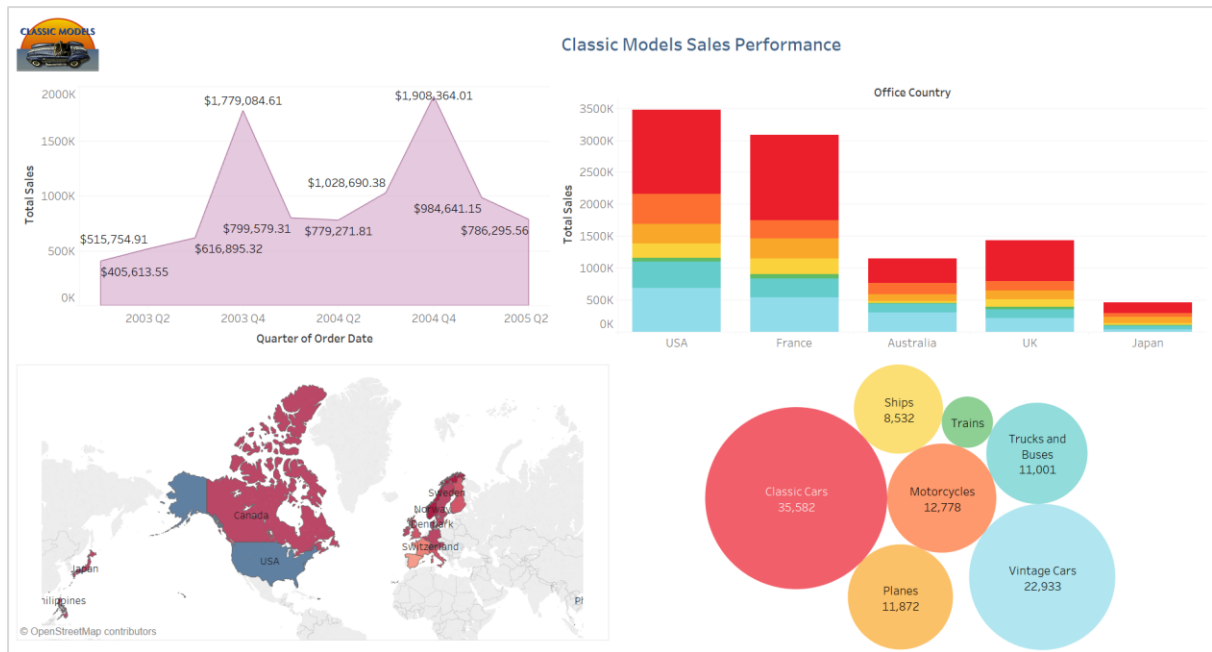
Tableau Dashboard

Figure 18: Tableau Dashboard

The dashboard presents a glance view of Classic Models company's sales performance.

Conclusion

The findings from the above analysis may be helpful in cultivating better marketing strategies if they are interpreted and used wisely. For instance, by looking at the stock to sales ratio tree map, the company could take note of the slow-moving products and keep them at the minimum to maintain proper balance of inventory. The company could also consider getting rid of products that is slow-moving and low profit and focus on products with high profit, such as 1937 Harch 930V Limousine and 1961 Chavrolet Impala. Sales representatives are responsible in increasing customer retention by following up on customers who have not been making repeat purchase for a long period of time. Allocating more sales representatives in sales offices with fewer employees especially Japan branch will help in reaching out to more potential customers, thus driving more revenues to the company.