Data Visualization

Introduction

The Adventure Works database supports standard online transaction processing scenarios for a fictitious bicycle manufacturer (Adventure Works Cycles). Scenarios include Manufacturing, Sales, Purchasing, Product Management, Contact Management, and Human Resources.

Industry: Manufacturing
Company Name: Adventure Works Cycles
Function area of Interest:

1. Sales Analysis
2. Cost Analysis
3. Customer Analysis
4. Promotion Analysis
5. Product Analysis

Objectives and scope: Evaluate and understand various aspects of Adventure Works Cycles to make informed decisions. This analysis provides insights into our company's financial health, understanding of the strengths and weakness of products, effectiveness of promotion, operational efficiency, and overalls performance.
Business Processes to improve:

1. Revenue growth
2. Product Development
3. Resources Utilization
4. Marketing Strategies

Note: All the name of the authors of the visualization is in the sheets name or dashboard names

Business Assumptions

- Our company located in North America
- Our company purchase part from vendors, does manufacturing and then sells them to stores (B2B) & customers (B2C).
- Our current datetime will be the last quarter of the data.
- All Employees work 192 hours per month.
- There are 4 levels of management in the company.
- Cost data I have: BOM (Bill of Material), Manufacturing Cost, Salary & Shipping.

ETL

https://github.com/Microsoft/sql-server-samples/releases/download/adventureworks/AdventureWorks2022.bak

First, I download the .bak database file from the above link address, then I use the Microsoft SQL Management Service to extract the required data source files from the database, and the following would be the list of data source file that I extracted:

Address.csv	ProdcutReview.csv
AdressType.csv	ProductSubCategory.csv
BillOfMaterial.csv	PurchaseOrderDetail.csv
Customer.csv	PurchaseOrderHeader.csv
EmployeeDeparment.csv	SalesOrderDetails.csv
EmployeeDeparmentHistory.csv	SalesOrderHeader.csv
EmployeePayHistory.csv	SalePerson.csv
Employee.csv	SalesTerritory.csv

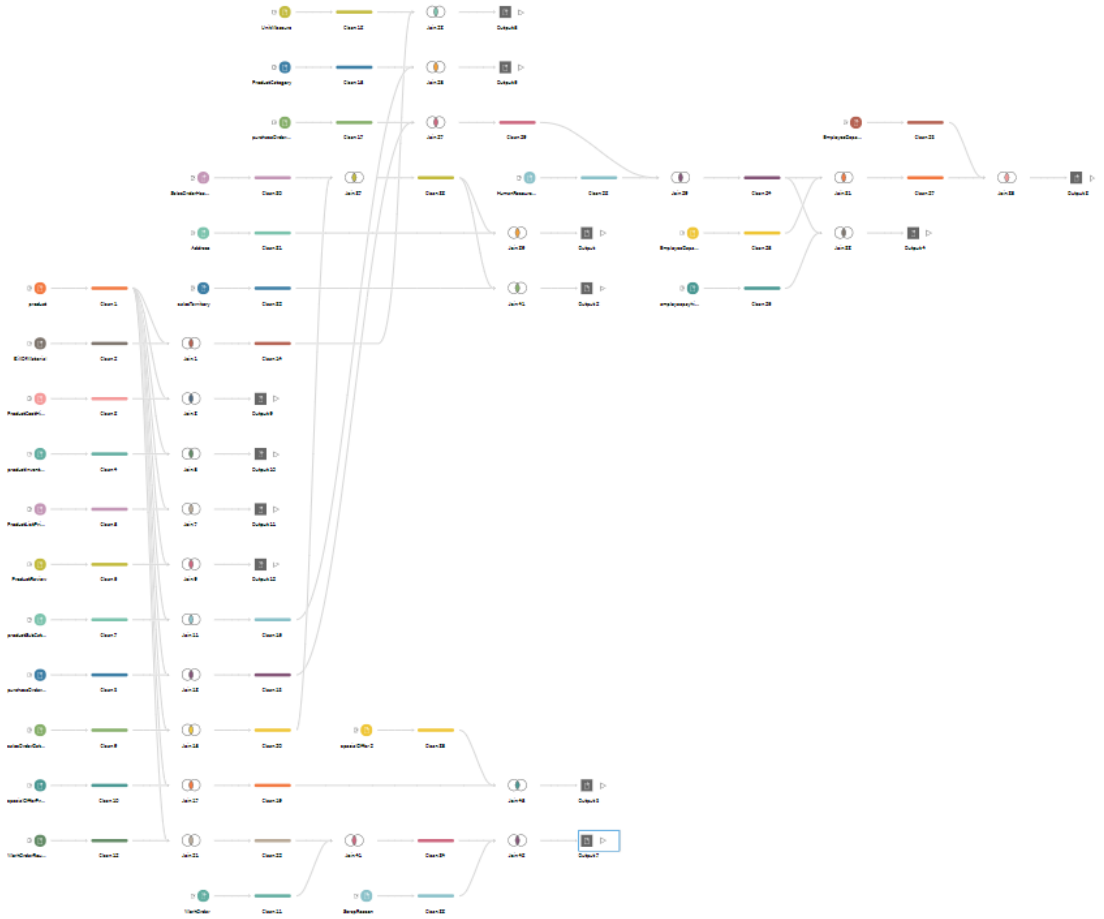
Product.csv	ScrapReason.csv
ProductCategory.csv	SpecialOffer.csv
ProductInventory.csv	SpecialOfferProdcut.csv
ProductListPriceHistory.csv	UnitMeasure.csv
	WorkOrder.csv
	WorkOrderRouting.csv

Then during the transforming I clean the data with removing null values and reformatting the date time to match the Tableau format.

2 Set of Union Table was created:

- WordOrderRouting.csv UNION purchaseOrderHeader.csv for cost distribution calculations
- WordOrderRouting.csv UNION purchaseOrderHeader.csv UNION SalesOrderHeader.csv for cost & sales calculations

After extracting the required data and transforming, I load the data into Tableau Prep Builder to generate all the hyper files, and the data flow of our transformed database is shown as the following. (The flow file is also attached in the assignment submission.) In total I have generated 12 tableau hyper files.



Graph 1. Structure of the Data Flow in Tableau Prep

Note: If there is data source error when you open the dashboard worksheets, please kindly select the correct data source.

The above parts are the same as the Lab 2 reports.

Dashboard

Name	Definition
Linear Regression Forecast on Sales (Filter: Region)	Regional Sales are forecasted by Linear Regression
Pareto Analysis on Product (Filter: Region)	Perform Pareto Analysis on how different products contribute to regional sales
Normal Distribution on Building of Materials	Perform Normal Distribution on the costs of building materials, to focus on the distribution.
Sales vs Quarter	Basis Statistics of sales among each quarter within the selected year

Sales Performance scorecard – KPIs, Comparison & Trends

KPI	Description
1. Sales Quota KPI	The overall total sales quota, comparing with the previous quarter for the growth / decay
2. B2B Sales Growth KPI	Measuring the B2B sales growth / decay when comparing to the previous year
3. B2C Sales Growth KPI	Measuring the B2C sales growth / decay when comparing to the previous year
4. Best Salesperson	Showing the top three salesperson with the greatest sales (year to date)
5. Region Sales Trend (Quarter Comparison)	The total sales in different region,

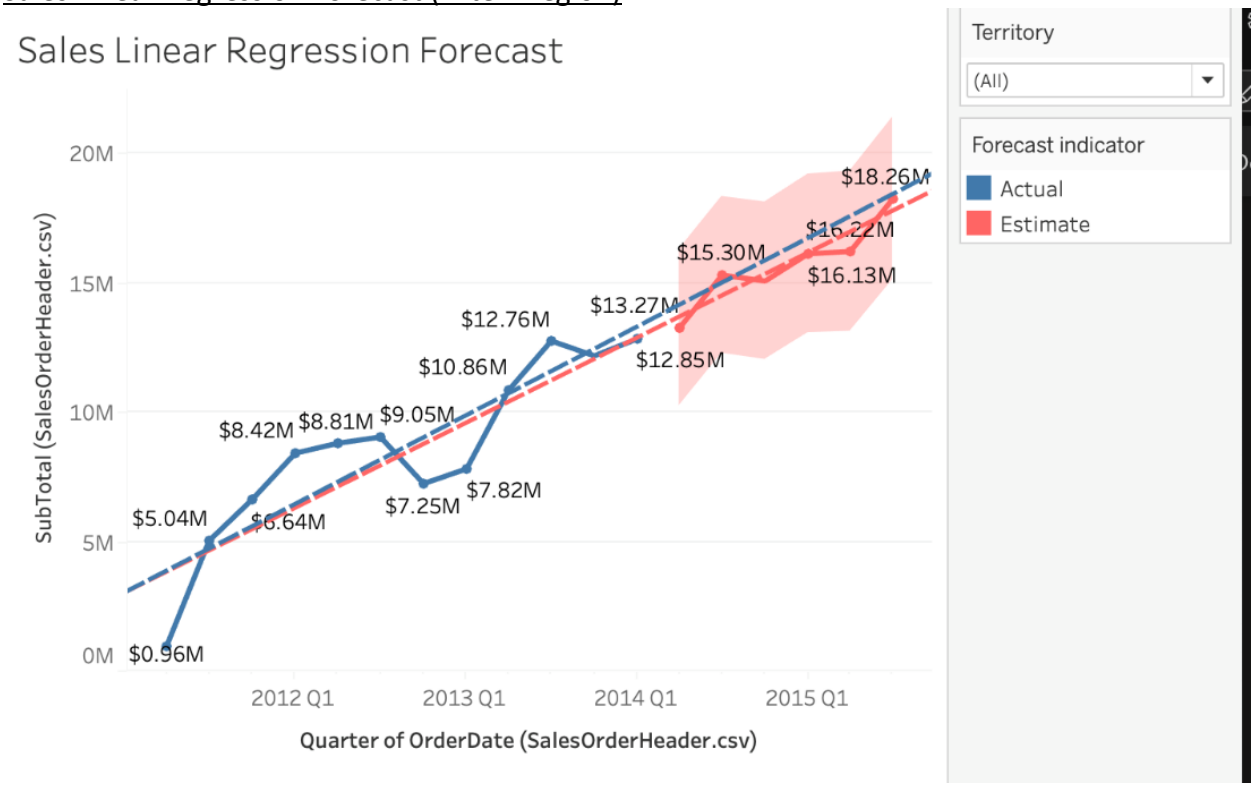
	comparing with the previous quarter for the growth / decay
6. Pareto Analysis on differences in revenue by region	Applying the 80/20 Pareto principle to analyze regional revenue changes between 2012-2013 provides visibility into product sales trends projected for Q3 and Q4 in 2014.

KPI - Visualization	Description
7. Conversion Correlation on Product Level	The correlation of purchase cost and sales amount on product level
8. Top 5 Selling Products	Top five products contributing to the total sales with their share percentage
9. Total Sales by Quarter **-> What–If Analysis, Forecast Sales in Region**	The quarterly total sales in different

	region, projecting forecast sales with what-if analysis
10. Pareto Analysis on Repeating Customer Contribution	Analyze the business model of the manufacturer based on sales revenue from different repeat purchase frequencies.

Dashboard Visualization

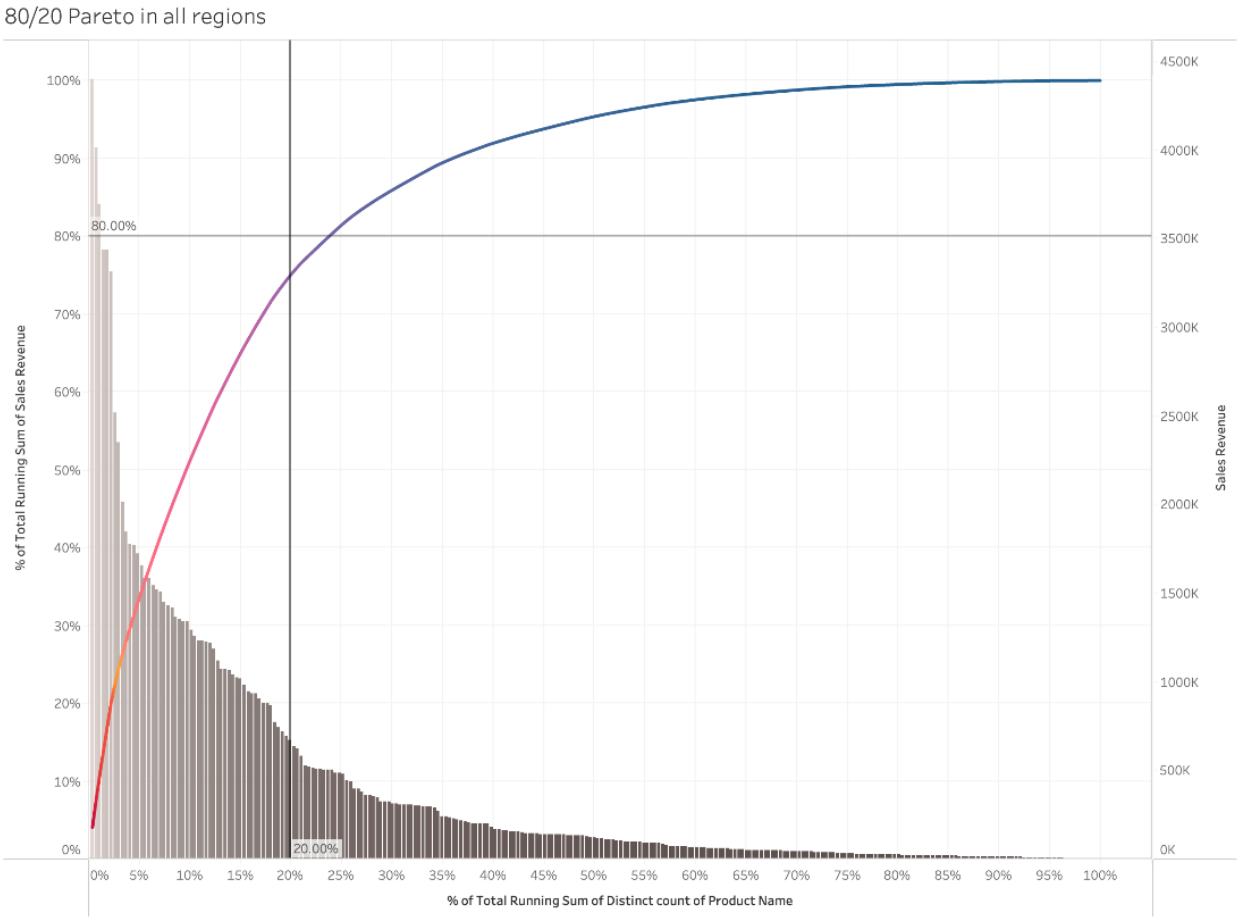
Sales Linear Regression Forecast (Filter: Region)



By applying linear regression on the sales quarter plot, I get the increasing trend line with R-Squared=0.8 and p-value<0.0001 which means the linear regression is accurate. After that, I applied a forecasting model and it returns a different predicted linear regression for 2015 Q1.

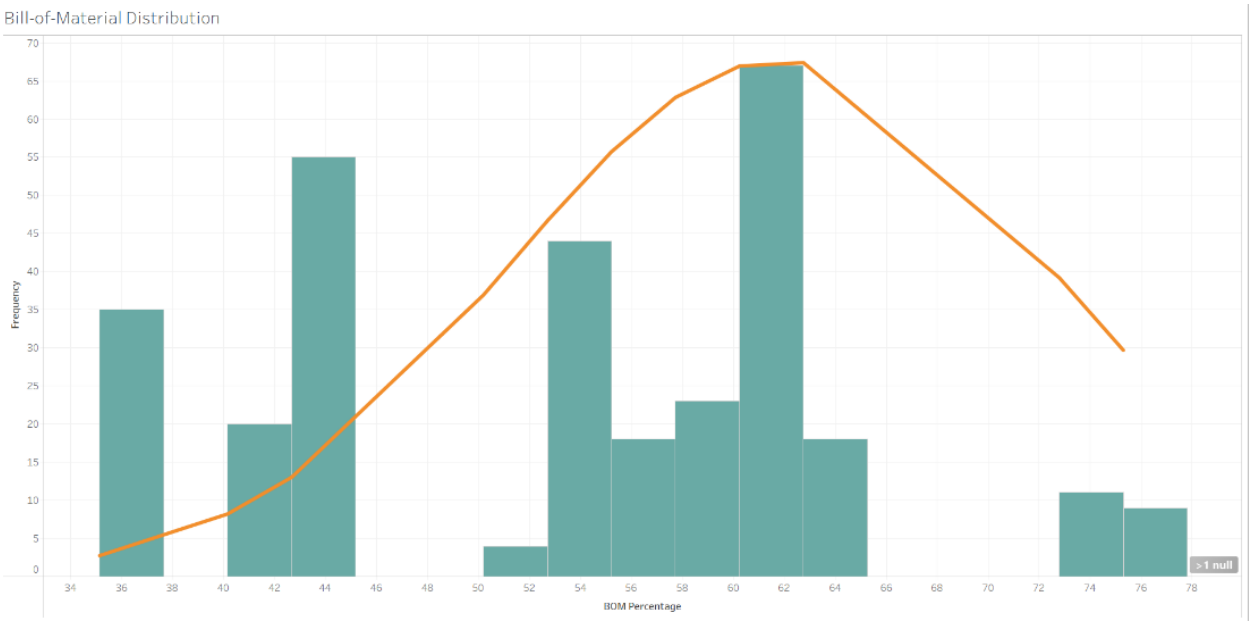
The estimated line states that the sales amount has less increase than that of the actual one. However, both trend lines indicate the sales amount will increase in 2015 Q1. To conduct a deeper analysis, I picked territory as the filter. And I found some of the territories are showing having less sales amount in the forecast. Applying the root cause analysis, economic factors and inflation could cause the sales to slow down in the upcoming quarter.

Pareto Analysis on Product (Filter: Region)



Using the Pareto technique, I divided product revenue according to the 80/20 principle. The results show that 20% of products generate most of the profit for our company, while the remaining 80% are less profitable.
I applied this analysis to each region to identify the most trending and profitable 20% of products by area. Targeting these regional top sellers allows us to efficiently focus our efforts on the products poised to drive the most revenue growth for the overall business.

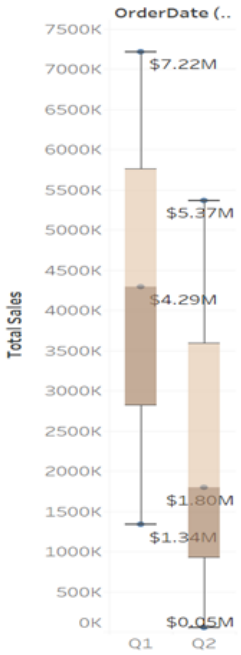
Normal Distribution on Building of Materials



The Bill-of-material distribution is left-skewed. It usually means that the mean < median < mode. There are 35 products for which BOM is between 35%-37.5%, which is much more than I expected compared to the normal graph. This causes the distribution to become left-skewed.

Here I would like to perform the root cause analysis of this finding. As the company tries to replace some of the materials to lower costs, the BOM of some of the products has been reduced to a lower level.

Sales vs Quarter



Sales by quarter (2013 & 2014)

Generally, the company performance varies directly to the seasonality. In 2012 and 2013, the company performed the best in q3 while they perform the best in q4 in 2011. However, in 2014, the performance of q2 is better than q1. Overall, they performed better in winter.
Here I would like to perform the root cause analysis to this founding in our sales,
As our company is the manufacturer who sales and manufactures bikes, I always use black Friday and Boxing Day discounts as the marketing strategies to enhance the sales, which usually happens to be the biggest discount throughout the year, which will be the best timing to purchase from us with a customer perspective. Moreover, consider our company locate in North US, most of the winter time will have the snow, and bike could not be ride on slippery road with snows, therefore I will have the clearance sale to attract customers which can help us to spare some of the storage spaces, on the other hand, customer will most likely attracted by the price, as they know there is no expiration date for the bike, and you can always save the bike for later, meanwhile saving a great amount of money on the same bike. Therefore, the total sales of bikes are usually higher in winter.

Sales Performance Scorecard KPI

Sales Quota KPI

Quarter..	
2011 Q2	3,886,000
2011 Q3	5,627,000 1,741,000▲ 44.8%▲
2011 Q4	4,750,000 -877,000▼ -15.6%▼
2012 Q2	10,537,000 5,469,000▲ 107.9%▲
2012 Q4	5,913,000 -2,741,000▼ -31.7%▼
2013 Q1	8,039,000 2,126,000▲ 36.0%▲
2013 Q2	13,733,000 5,694,000▲ 70.8%▲
2013 Q3	11,097,000 -2,636,000▼ -19.2%▼
2013 Q4	8,051,000 -3,046,000▼ -27.4%▼
2014 Q1	10,359,000 2,308,000▲ 28.7%▲

Generally, the sales quota decreased in winter. There is a drop on sales quota in Q4 of all years and 2013 q3, especially in 2012 Q4, which is –31.7%. In contrast, the most significant increase of sales quota is in 2012 Q2, which is +107.9%.

Here I would like to perform the root cause analysis of this finding, considering our company is located in the North US, and most of the wintertime will have snow and under zero degrees which is very cold. Customers would prefer not to choose bikes as their transportation instead of buses or the subway. Therefore, the total sales quota of bikes is usually lower in winter.

B2B Sales Growth KPI

The B2B sales are increasing in 2012 and 2013 while they are decreasing in 2014. The B2C sales of 2012 increased 281.49% compared to 2011. The B2C sales of 2013 increased 110.10% compared to 2012. The B2C sales of 2014 decreased 32.20% compared to 2013.

B2C Sales Growth KPI

The B2C sales are increasing in 2012 and 2013 while they are decreasing in 2014. The B2C sales of 2012 increased 150.39% compared to 2011. The B2C sales of 2013 increased 152.67% compared to 2012. The B2C sales of 2014 decreased 70.92% compared to 2013.

Best Salesperson

Busines..	Ranking_sale..	Sales YTD
276	1	4,251,369
289	2	4,116,871
275	3	3,763,178

From the above visualization, I have the top three salespersons which have 4.2, 4.1 and 3.7 million sales as of today. The third top salesperson tremendously outstands the fourth salesperson by around 600k.

Differences in revenue between 2012-2013 in Canada by applying the 80/20 Pareto principle:

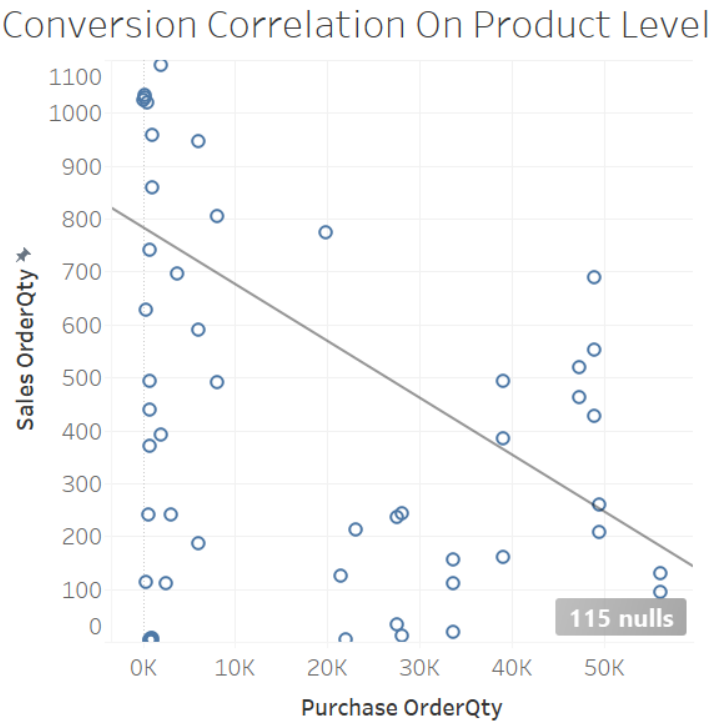
Differences in revenue between 2012-2013 in Canada by applying the 80/20 Pareto principle

	Canada	
80/20 Revenue	2012	2013
80% revenue	193.2% ▲ $3,377,109.5	-10.6% ▼ -$545,087.9
20% revenue	481.1% ▲ $609,565.5	124.1% ▲ $913,548.0

Drilling down by using clustering techniques, I created a dataset consisting of the top 20% selling products from each region. I then visualized the sales trends for these 'clustered' products over time in a KPI table. Given limitations in the completeness of the 2014 data at the time of analysis, I focused the profitability analysis on 2012 and 2013. These two full years provide clearer visibility into emerging trends versus the partial 2014 data.
Focusing on the specific Canada regions, there was a 193.2% upward trend for the top 20% of products from 2011 to 2012. Sales then dipped 10.6% from 2012 to 2013 but remained strong overall. This minor decrease does not significantly impact on the growth potential for these clearly popular products. I can continue prioritizing and promoting these momentum-driving products in the Canada region into 2014 and beyond.

Sales Performance Scorecard Visualization

Conversion Correlation on Product Level



Conversion Correlation on Product Level (2013 & 2014)

The scatter plot shows a negative relationship between purchase cost and sales amount in 2013 and 2014. It shows that the higher the cost that the company spent, the lower the sales amount achieved, which is one of the main reasons for having a negative profit margin. For 2011 and 2012, the relationship between purchase cost and sales amount is positive.

Here I would like to perform the root cause analysis. From the above analysis, I know that there are some new products that have been added since 2013. Starting from 2013, the profit started to have a trend of decrease, which means there is a loss on the profit. The reason could be the marketing team make decision without historical sales data, as the missing of sales data, I can only predict the sales of the new product by the other similar product sales trend, however the market reaction are not good as expected, and the result is our company overstock these products, and gigantic increase of the purchasing cost and lead to loss of profit.

Top 5 best-selling products

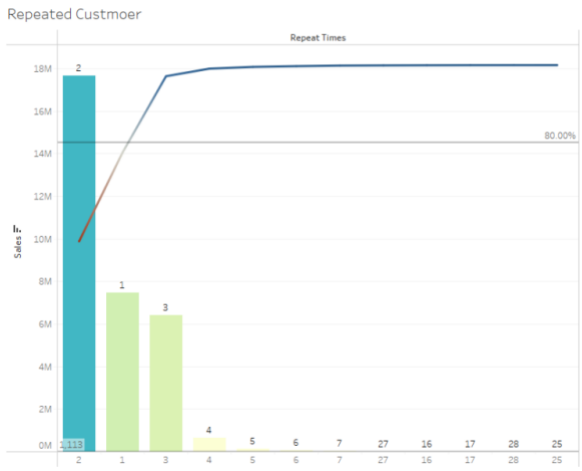
From this visualization, I can see that the top 5 best-sellers of our company are all the mountain-200 but in different colors and sizes. Among different sizes and colors, Black with 38 inches tires is the all-time best-seller no.1 with about 20% of our sales of all time, and then the color black with 42 inches tires is the 2nd best-seller of all time with about 18.5% of total sales. Furthermore, the color silver with 3 different sizes of tires, the 38, 46 and 42 occupy the 3rd to 5th places with 17.3%, 16.5% and 16.3%. So that our main sales rely on our main product, the bikes.

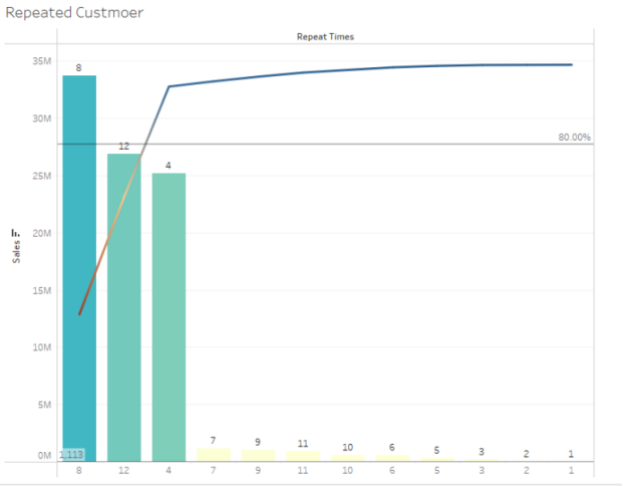
Applying the root cause analysis to this founding, that all of our best 5 top sellers are all the bikes Mountain-200 in different colors and with different tires, there are various reasons to explain about, first our product quality is good, that I have a good reputation among the markets and customers, so that they would buy our products, on the other hand, the another explanation would be that our marketing strategies are mainly focusing on the bikes, so that whenever I have the promotions and special offers, bikes will be a part of the event, so that the sales of the bikes could be boosted.

Total Sales by Quarter – What-If Analysis

Applying what-if analysis to the total sales by quarter plot, I can see how the forecast lines react to the percentage growth. To have a better analysis, I set the percentage growth filter to be between –100% and 100%. If I apply 100% to the filter, the forecast sales will be 100% higher than the actual sales and vice versa. In this case, it helps us to make decisions in different situations.

Pareto Analysis on Repeating Customers





Repeating customers are divided into B2B and B2C. In the B2B charts, customers who made repeat purchases 4 times, 8 times, and 12 times accounted for the majority, totaling 95% of sales performance. It is noteworthy that the sales performance of customers who made repeat purchases between 4 to 8 times was significantly better than those who purchased 8 to 12 times. Customers who purchased the product 12 times indicate that there is no major issue with the product itself; the problem might be due to other companies developing more competitive or lower-priced products. Therefore, a feasible strategy could be to develop entry-level, advanced, and professional products based on the original product, offering more diverse choices.

In the B2C charts, purchases made from one to three times ranked in the top three, with two-time purchases accounting for the majority. This suggests that the public really likes the product after purchasing and using it. Hence, the strategy could focus on marketing the brand, attracting more unknown users to the product, and thereby increasing market share.