**ICA**

**BIG DATA AND BUSINESS INTELLIGENCE**

**NAME: NWODO UCHENNA DONALD**

**STUDENT ID: B1163786**

**TABLE OF CONTENTS**

**SECTION 1: BUSINESS INTELLIGENCE DESIGN**

DATASET

The dataset used in this work is a fictional dataset created by Microsoft for the sole purpose of learning. The dataset is called Adventure works and is provided on Kaggle for public use. The link to the dataset is shown below:

<https://www.kaggle.com/ukveteran/adventure-works?select=AdventureWorks_Customers.csv>

This dataset is from Adventure Works company, a large multinational manufacturing company involved in the manufacturing and sales of bicycles and its components, as well as sports wears and other accessories to North American, European and the Pacific commercial markets. The company has a customer base of about 18148 customers comprising of both males and females, with each customer having a unique customer ID number for identification across the different regions.

Once a customer places an order from any of the regions, an order number is generated and order date is noted which will be used to calculate the estimated time of delivery to the customer. The order number is used to generate an invoice and prepare the order for delivery to the customer as required. The invoice and the ETA is further sent to the customer through email. Also, on a quarterly basis, a satisfaction survey is sent to all customers through email to get feedback from the customers about their purchase experience within the quarter. The dataset consists of customers’ purchase order information for Adventure Works Company for a period of three years, 01/01/2015 to 30/06/2017.

OBJECTIVES AND APPROACH

This Business Intelligent project employed Power BI application and its report to provide a means for Adventure Works management team to analyse the company's yearly performance over a three financial year period. This analysis is primarily concerned with the issues below:

* To provide Sales trend visibility on monthly, Quarterly and yearly basis.
* To check the company’s performance and how it can be improved based on predefined Key Performance Indicators
* To know and understand the customers behaviours. i.e., Who are our customers?
* To Understand product performance i.e., What are the customers buying?
* To help the Marketing Department to know the customers to be targeted for promotion amongst others to maximize profit and minimize cost.
* To help the company in inventory management

**Scope**

This project scope comprises the above approach, but the focus is on data pre-processing, data cleaning, and data modelling in section 1. In section 2, the priority is to deliver an analytical report and interactive dashboard to support performance evaluation based on the above questions

**Rationale**

This dataset is chosen for this project because of two reasons. Firstly, Power BI appears as the most appropriate Business Intelligence (BI) tool option for this dataset as it is a friendly-user and cost-effective app to create intelligent visualisation for the company performance analysis. The reports and dashboards can be customised to suit the company's needs and can be easily shared within the organisation. Secondly, the dataset contains real commercial data with a large volume (nearly 100,000 transactions) and complex data modelling, like implementing big data and analytics applications in the real-world scenario. Therefore, it is an excellent opportunity to enhance the skill learnt from this module to apply in future jobs.

**Outcomes**

At the end of this BI project, the following skills can be demonstrated:

• Data cleaning and data pre-processing skills through Power BI, such as importing data, adding/removing a column, removing NAs or errors, changing data types.

• Data modelling by creating and managing relationships.

• Using M language in Power Query

• Performing a calculation using DAX

• Critically analysing data using different visualisations and Artificial Intelligence tools

• Designing and publishing report

• Creating an interactive dashboard from the report

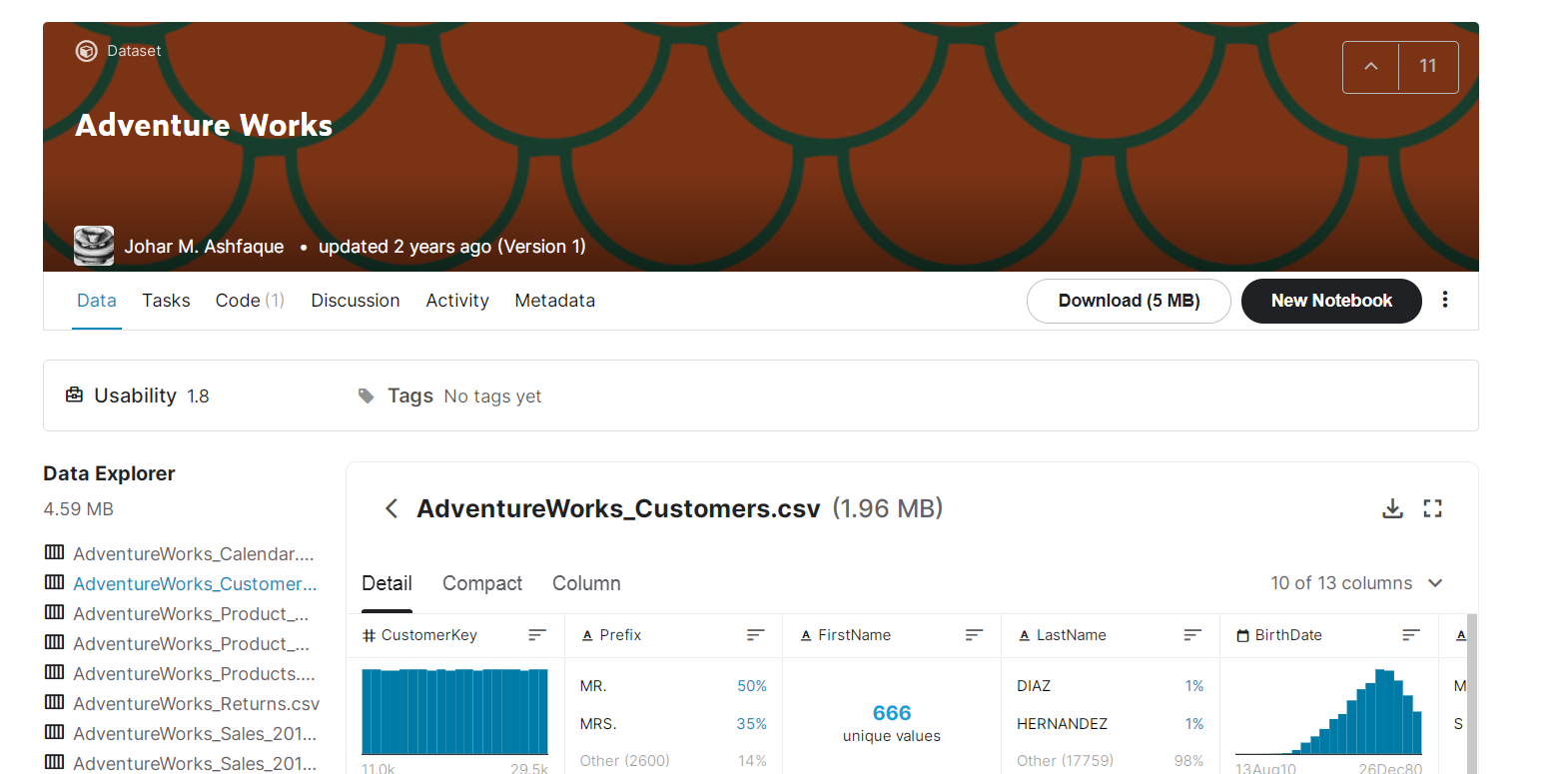


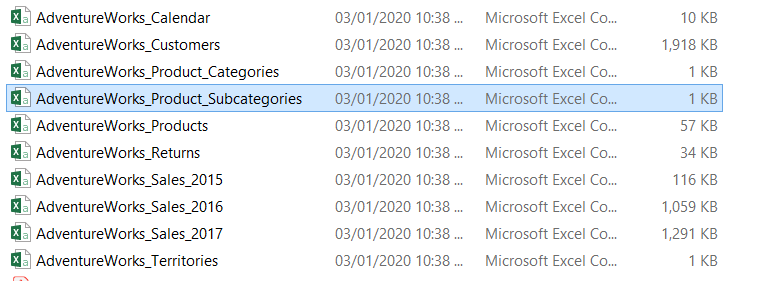
Fig. 1: Screenshot of data source

Data Description

The dataset has 10 tables out of which nine of them were used. The calendar table which contains the same information as the order dates in the sales table was deleted because it was irrelevant to the dataset. The nine tables were summarised in a tabular form as below:

|  |  |  |  |
| --- | --- | --- | --- |
| **No** | **Table** | **Column** | **Definition** |
| **1** | AdventureWorks\_Customers\_Dataset 13 Columns 18148 Rows | Customer Key | Customer ID Number |
| Prefix |  |
| FirstName | Firstname of Customers |
| LastName | Lastname of Customers |
| BirthDate | Customers date of Birth |
| Marital Status | Customers marital status |
| Gender | Customers Gender |
| Email Address | Customers Email address |
| AnnualIncome | Yearly income of customers |
| TotalChildren | Number of children each customer has |
| EducationLevel | Customers level of education |
| Occupation | Customers Occupation |
| Homeowner | Customers that own houses |
| **2** | AdventureWorkd\_Product\_  Categories\_Dataset 2 Columns 4 Rows | ProductCategoryKey | Product category ID |
| CategoryName | Name of each product category |
| **3** | AdventureWorks\_Product\_SubCategories\_Dataset 3 Columns 37 Rows | ProductSubCategoryKey |  |
| SubCategoryName |  |
| ProductCategoryKey |  |
| **4** | AdventureWorks\_Products\_  Dataset 11 Columns 293 Rows | ProductKey |  |
| ProductSubCategoryKey |  |
| ProductSKU |  |
| ProductName | Name of each product |
| ModelName | Name of the model for each product |
| ProductDescription | A full description of each product |
| ProductColor | Color of each of the products |
| ProductSize | Size of the products |
| ProductStyle | Style of the products |
| ProductCost | Production cost of each product |
| ProductPrice | Selling price of each product |
| **5** | AdventureWorks\_Returns\_  Dataset 4 Columns 1,809 Rows | ReturnDate | Date Product was returned |
| TerritoryKey | Territory Key Number |
| ProductKey | Product Key Number |
| ReturnQuantity | Quantity of Product returned |
| **6** | AdventureWorks\_Sales\_2015\_Dataset 8 Columns 2,630 Rows | OrderDate |  |
| StockDate |  |
| OrderNumber |  |
| ProductKey |  |
| CustomerKey |  |
| TerritoryKey |  |
| OrderLineItem |  |
| OrderQuantity |  |
| **7** | AdventureWorks\_Sales\_2016\_Dataset 8 Columns 23,935 Rows | OrderDate |  |
| StockDate |  |
| OrderNumber |  |
| ProductKey |  |
| CustomerKey |  |
| TerritoryKey |  |
| OrderLineItem |  |
| OrderQuantity |  |
| **8** | AdventureWorks\_Sales\_2017\_Dataset 8 Columns 181 Rows | OrderDate |  |
| StockDate |  |
| OrderNumber |  |
| ProductKey |  |
| CustomerKey |  |
| TerritoryKey |  |
| OrderLineItem |  |
| OrderQuantity |  |
| **9** | AdventureWorks\_Territories\_Dataset 4 Columns 10 Rows | SalesTerritoryKey | ID numbers for locations where the products have been sold. |
| Region | Regions where the products have been sold |
| Country | Countries where the products have been sold |
| Continent | Continents where the products have been sold |

The screenshot of the raw data files in a CSV file format can be shown in Fig. 2 below



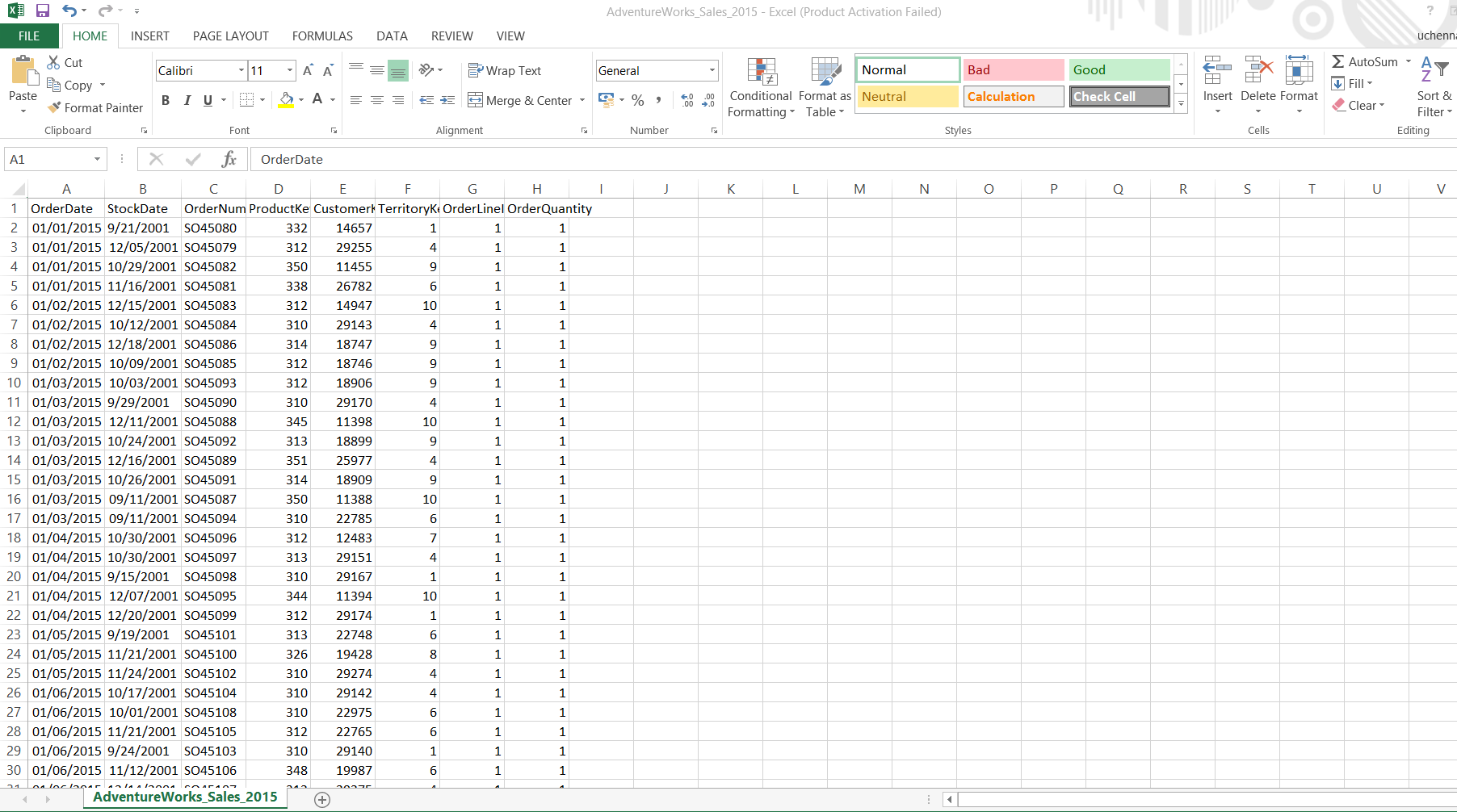


Fig.2 Screenshot of the Raw data files in CSV

BI DATA PREPROCESSING AND DATA CLEANSING

Data Loading

There are two options for loading or connecting data into the power BI, either by Importing (Getting Data) or by Direct Query method. However, in this project I will consider importing the data and this method involves loading the data or data sets from the source file, folder or database to the power bi.

By clicking on the Get **Data** button on the **home ribbon** as shown in the figure below, a drop-down window showing **common data source** options for loading the data is displayed. As highlighted above, for this project the data for this analysis is saved in csv file extension, therefore we will be using the **Text/CSV option** from thedropdown window. After clicking the Text/CSV option a new window open and the data file was selected.

Graphical user interface, application, Word, website

Description automatically generated

Figure 3: Screenshot of loading the dataset into the power BI

After selecting the file, in this case the *Adventureworks\_Calender* file, a new window opened as shown below, **Transform Data** option was selected, and the data was loaded into the **power query**.

Graphical user interface, application, Word

Description automatically generated

Figure 4: A screenshot showing how the file to be loaded was selected

Graphical user interface, text, application

Description automatically generated

Figure 5: A screenshot showing the window page to load or transform the dataset.

The power query window opened showing the loaded data table as shown in figure 6 below.

Graphical user interface, application, Word

Description automatically generated

Figure 6: Opened power query window with the loaded file.

Then the table name was changed from *Adventureworks\_Calendar* to *Calendar* on then Query settings pane by the right-hand side on the power query window just as shown in figure 7 below.

Graphical user interface, application, Word

Description automatically generated

Figure 7: A screenshot of changing the table name

The other CSV files were loaded into the power query by selecting **New Source** on the **Home tab** ribbon and the same process was repeated to load the remaining CSV files into the power query as shown in the figure 8.

Graphical user interface, table

Description automatically generated

Figure 8: A screenshot of how the other data tables were loaded into the power query

Data Pre-processing and Data Cleaning:

In order to put the queries in proper shape, there is a need to pre-process (Data preprocessing is a technique which is used to transform the raw data in a useful and efficient format) and clean the data. To do this I checked each column of queries to check for errors, missing values and for headers that are not properly promoted.

Dealing With Missing Data

While checking through all the tables for missing values, the **Prefix** column on the *Customers* table has missing value as shown in the figure below.

Graphical user interface, table

Description automatically generated

Figure 9: A screenshot showing blank or null value in the Prefix column

On the **Customers** table the **Prefix** columnhas missing values and is not needed in the analysis, therefore it was removed by selecting the **Prefix** column and right clicking and then choose **Remove** as shown below. This process was also repeated for other queries where a column is needed to be removed.

Graphical user interface, application, table, Excel

Description automatically generated

Figure 10: A screenshot of how the prefix column was removed

On the Gender column still on the **Customers** table it was also discovered that the column contains an unknown value **NA,** which was then filtered off by unchecking the **NA** as can be seen on the Figure below.

Graphical user interface, application, table

Description automatically generated

Figure 11: A screenshot showing how to filter the Gender column.

**Merging Columns**

The *FirstName* and the *LastName* column on the *Customers* table can be merged to become *FullName* column. This was done by holding the Ctrl key and selecting the *FirstName* and *LastName* columns and then right click on any of the selected columns and then I selected Merge Columns as shown below, then a new window popped up where I selected space as the separator and then changed the new column name to FullName.

Graphical user interface, table

Description automatically generated

Figure 12: A screenshot of how two columns were merged

**Replacing Values in the Columns**

The values on the *MaritalStatus* column were replaced with their full text, for example **Married** for **M** and **Single** for **S**. To do this I selected and right clicked on the *MaritalStatus* column to select **Replace values** amidstthe options, the Values to find and to be replaced were then inputtedas shown in the figure below.

Graphical user interface, application, table, Excel

Description automatically generated

Figure 13: A screenshot showing how to replace values in a column.

This process is repeated for other columns like *Gender* and *HomeOwner*.

**Appending Tables**

In power query there is the possibility of appending tables to another to either form a new one. The sales table for 2015, 2016 and 2017 were appended. **Append Queries** was selected from the **Home tab** ribbon, the 3 queries were appended as a new query as show below. The new query name was changed from **Append1** to **Sales.**

Graphical user interface, application, table, Excel

Description automatically generated

Figure 14: A screenshot showing how Sales\_2015, Sales\_2017 and Sales\_2016 were appended.

Table

Description automatically generated

Figure 15: A screenshot showing the appended table that has been renamed.

**Merging Queries**

To check the product information and reduce the data model's complexity, I merged the Product\_Categories and the Product\_Subcategories table into the Products table using the **merge queries** button on the **Home tab**. The Product\_subcategories table was first merged with the Products table, as described and shown below.

* I first selected the **products** table and then selecting **Merge Queries** on the **Home tab** ribbon and a Merge window will pop up,
* On the Merge window the table to be merged and their matching columns were selected and then expanded, this step was also repeated for the merging of Products table and the Product\_categories table.

Graphical user interface, application, table, Excel

Description automatically generated

Figure 16: A screenshot showing how to merge tables

Table

Description automatically generated

Figure 17: A screenshot showing how to expand the column of the merged tables.

**Creating Additional Columns**

In order to create additional columns on the calendar table, I selected the calendar table and then the **Date** column, On the **Add Column** tab ribbon under **From Date & Time** group I selected **Date** and then Year to get the Year column, Month🡪 Month Name etc.

Table

Description automatically generated

Figure 18: A screenshot showing how to create new date columns

**Grouping the Queries in Folders**

Since some of the tables have either been merged or appended to form a new one, I decided to group the tables. This was done by right clicking on any of the tables selecting **Move To Group,** **New Group** and then named the groupsas **Other queries** (Sales 2015, Sales 2016, Product\_Categories, Product\_subcategories), **Facts** (Sales, Returns) and **Dimensions** (calendar, customers, products etc.). The related tables were dragged into their respective groups. This is shown in the figure below.

Table

Description automatically generated

Figure 19: A screenshot on how to group the tables/ queries

Table

Description automatically generated

Figure 20: A screenshot showing the grouped tables

After grouping the queries, the applied steps on the power query are applied by selecting close **& apply** on the **home tab** ribbon as shown below.

Graphical user interface, application, table, Excel

Description automatically generated

Figure 21: A screenshot showing how to Close and apply the queries

**BI Data Modelling**

After loading and pre-processing data, Power BI will automatically generate the proper relationship between tables by matching columns with the same name on each table. The picture below shows the entity-relationship diagram of raw unsupervised data and its interconnection.

Graphical user interface, application

Description automatically generated

Figure 22: A screenshot showing the automatically generated relationship between the tables

Next, we want to modify the above data schema by deleting the automatic relationship created by Power BI by manually connecting the dimensions tables to the Facts tables. The Sales and Return tables are the fact table in the above schema. The other tables are called the dimension table, which provides descriptive information.

**Adding and managing relationships**

After deleting the automatic relationship created by the Power BI, we then started creating the relationships from the dimension tables to the facts tables, this was done by dragging the primary key columns from the dimension table to the related foreign key column in the facts tables. For Example, The **ProductKey** from the Products table was dragged to the top of the **ProductKey** on the **Sales** and **Returns** tables to form **one-to-many** relationship with each of the tables. The **Date** column in the **Calendar** table is dragged on the **OrderDate** on the **Sales** table and on the **ReturnDate** on the Return **table.**

Graphical user interface, application

Description automatically generated

Figure 23: A screenshot showing the unmanaged and unconnected tables

Graphical user interface

Description automatically generated

Figure 24: A screenshot showing one-to-many relationship between the **Products** tableand the **Sales** table.

The Following Process is repeated with the other dimension tables to form a **Star Schema** as shown below

Graphical user interface, application, Word

Description automatically generated

Figure 24: A screenshot showing a Star Schema of the one-to-many relationships between the dimension tables and the Facts tables

**New measures and Calculated Columns**

**Age Column**

To calculate the Age of each customer at the date of their last transaction and then group them, we computed a calculated column on the customers table on the **Data** view by selecting **New column** on the **Table tools** tab, the column was created using the DAX function.

Table

Description automatically generated

Figure 25: A screenshot showing how to create a calculated column

Table

Description automatically generated with medium confidence

Figure 26: A screenshot showing the Age column created using DAX function

Age =

VAR vLastD=YEAR(

MAXX(Sales,

Sales[OrderDate]

)

) //This variable returns the year in the OrderDate

RETURN

vLastD-YEAR(Customers[BirthDate])

Where:

* Age is the name of the Calculated Column
* MAXX() returns the maximum date in the orderDate column in the Sales table for each customer
* YEAR() function returns the year of the date

The same process but with different DAX functions were used in creating other Calculated columns such as Age\_group, Income\_group.

**Creating a New Measure**

In order to calculate the Total revenue, the **Total Sales** measure was created as shown in the figure below

Graphical user interface, application, email

Description automatically generated

Figure 28: A screenshot on how to create a measure

Graphical user interface, application, Word

Description automatically generated

Figure 29: A screenshot showing the Total Sales measure created

Total Sales =

SUMX(

Sales,

Sales[OrderQuantity]\*Sales[ProductPrice]

)

Where:

* Total Sales is the total Revenue generated over time.
* SUMX() is an iterative DAX function that returns the sum of an expression evaluated for each row in a table

SECTION 2:

**ADVENTURE WORKS CYCLE PERFORMANCE ANALYSIS**

STUDENT ID:

STUDENT NAME:

**EXECUTIVE SUMMARY**

Adventure Works Cycles is a large and rapidly growing multinational manufacturer and seller of bicycles and accessories to commercial markets. This report Analyses the company’s performance for the Financial Year 2017 (FY17) started from 01/01/2015 to 31/06/2017, to evaluate the current marketing strategy’s effectiveness and identify the for-marketing process improvements.

The report addresses the following questions:

* To understand how the business revenue is generated and the evaluate performance and trend
* To understand the product performance and hence inventory management
* To provide sales team information about customers' profiles and preferences.
* To dynamically segment and understand the customers behaviours
* To understand the return rate of the products based on geographical location and product category

**Findings and Conclusions –**

* The Total Revenue generated in the first half of FY17 was $9,122,607, with a year-over-year growth of 211**%. 92% of the Total Revenue generated is from Bike category.** Monthly revenue increased by 44.11% and trended up between January and June 2017 and is forecasted to have grown between **$2.5M** and **$2.6M** by December of FY17 from the current month.
* The **Accessories** category with a **gross margin percentage** of **63%** is more profitable than the **Bikes** category with **gross margin %** of **41%**
* Based on the product ranking by their profitability (gross margin %), **Accessories** and **Clothing** products are at the Top 10 products that are more profitable. The top 10 least profitable products were also from the **Clothing** productcategory.
* There are **293** products in the company’s inventory, **191** of these products have not been sold before and of which **132** of these products are from the **Components** category, which have never been sold before.
* Products in the **Accessories category** with gross margin percentage of **63%** are more profitable than products in the **Bikes category** contributed to **93%** of the Total revenue in FY17
* Customers within the age of 40 to 80 years with **bachelor’s degree** or **Partial College level of education**, who are **Professionals** or **skilled Manually** earning at most **$100,000** with 2 or a smaller number of children either **male** or **female** that owns a house are more likely to buy the company’s products.
* Customers who bought Bikes are more likely to return between 10 to 13 months after their first purchase
* There are a total of **17,293** customers, **36.5%** (6,312) are **new customers, 29.8%** (5,160) are **Promising Customers** and **24.6%** (4,262) are **Champions.**
* **Shorts and Vests** subcategories had a high return rate in **Europe** and **Pacific.** Customers are more likely to return products in the Q1 compared to other quarters of the year. The United States and Australia had the highest return based on the quantity returned. France although with the least sales revenue has the highest return rate of **2.26%. Water Bottle – 30 Oz.** being the highest returned quantity followed by **Patch kits/8 patches** and **Mountain Tire Tubes.**

**Recommendations for improving performance**

* Instead of spending resources trying to gain new customers, the company should focus on upselling or cross-selling current customers. This is significantly more effective and cost-efficient as the current customers are already acquainted with the products and are therefore more likely to purchase. A customer appreciation gesture such as special discounts will lead past customers and to their next purchase.
* Customer engagement and loyalty to the company should be prioritized.
* A market basket analysis should be conducted to know complementary products and which product to discount to drive the sales of the other
* Focus should be on United States, Australia and United Kingdom
* Focus should be on Bikes as larger percent of the revenue comes from this category.
* Effective marketing Strategy should be used to drive the sales of most of the products without sales. Based on a strategic plan develop targeted promotions to hit specific customers with ad messages and promotional offers.
* Leverage on varieties of marketing strategies and channels to get your product noticed.
* Regularly review the company inventory to identify redundant products and products with high return rate.

1. **INTRODUCTION**

Adventure Works Cycles is a large and rapidly growing multinational manufacturer and seller of bicycles and accessories to commercial markets in North America, Europe, and Asia. The Headquarter of Adventure Works Cycles is Bothell, Washington, where the company employs 500 workers. Following a successful fiscal year, Adventure Works Cycles now wants to broaden its market share by targeting advertising to its best customers, extending product availability through an external Web site, and reducing the cost of sales by reducing production costs.

This report Analyses the company’s performance for the Financial Year 2017 (FY17) started from 01/01/2015 to 31/06/2017, to evaluate the current marketing strategy’s effectiveness and identify the for-marketing process improvements. The report was based on the dataset described in section 1.

Question 1: To understand how the business revenue is generated by Product, geographical location, customer demography and evaluate company’s performance and Sales Trend. How did the company perform in Sales in Previous years compare to this year and what is the likely sales trend in the next two quarters of the current year (Forecast the sales for the next six months?).

Question 2: To understand the product performance. Which of the products are top sales and which of the products are not selling at all or bottom sales? This is to help the management identify product performance by product category, product subcategory and geographical location.

Question 3: Who are our customers, what do they buy and how often do they patronize us? This is to provide sales manager more information about customers' profiles based on the customer’s demographic attributes.

Question 4: To dynamically segment and understand the customer segments behaviors based on their Recency, Frequency and Monetary value

Question 5: To understand the return rate of the products based on geographical location and product category. To have in-dept understanding of what contributes to the high returning rate. This will reveal the frequently returned products based on Continent, Product category.

**2. KEY INSIGHTS**

A dashboard was set up to address the company’s performance and to help make possible business decisions as demonstrated in the figure, as demonstrated in the figure 1 below.

**2.1** **Revenue Analysis:** **Understanding how the business revenue is generated, compare company’s performance and sales trend**

Companies or businesses that aims to achieve its set goals and objectives, must carefully analyze its strengths and weaknesses and there are many parts of the business that requires attentions. Among such parts is the revenue stream or profits of the company. Revenue analysis identifies the factors such as the costs of products and which areas of the company needs an increase in revenue. This analysis helps the business owners to ensure that their plans and strategies does not deviate from the set goal(s). An analysis on the revenue was performed to provide an overview landscape of the company’s effectiveness and to understand product, category, subcategory and geographical contributions to the total annual revenue generation. Figure 1 illustrates the report page, in which order-related data was visualized to address the revenue trend.

Graphical user interface

Description automatically generated

Figure 1: The Dashboard for the Revenue Analysis of the FY17 Performance

The page also consists of some metrics such as the total sales, Gross Margin, Gross Margin percentage, Total customers, Total order and the total number of return orders. These metrics allows us to monitor the performance of the business briefly to determine the healthiness of the business. These metrics are shown on all the pages.

Some buttons were added to the report page to support easy navigation of the dashboard by the user.

|  |  |
| --- | --- |
| Button | Function |
|  | This button opens the pane where all the filters or slicers are |
|  | This navigates to the home page |
|  | This navigates to the Revenue analysis page, the green colour signifies that this is the current page on the screen |
|  | This navigates to product analysis page |
|  | This navigates to the return analysis page |
|  | This navigates to the customer analysis page |
|  | Ask a question about the data |
|  | Switch between total sales trend and YOY % |

As data visualization helps to make correlations, it also helps to understand the trends and patterns in a dataset in order to make sense of the data. A properly designed dashboard as shown above coupled with data storytelling uncovers valuable insights to the business decision-makers. The detail visuals in the revenue analysis page are described below.

A screenshot of a computer

Description automatically generated with low confidence

Figure 2: Total Sales by Category

A screenshot of a computer

Description automatically generated with medium confidence

Figure 3: Total Sales by Subcategory

**By Product Category and Subcategory**

Referring to the figure 2 and 3 above shows the FY17 revenue by the product category and subcategories. The Total Sales Category and Top 10 Subcategory by Total Sales are donut chart and clustered bar chart respectively created using the **CategoryName** and the **SubcategoryName** column in the product table as the axis and the **Total Sales** measure created as the value. On the donut chart reveals that the Bike category is the highest performing where more profitable sales were recorded from the Mountain Bikes, Road Bikes and Touring Bike subcategory. The least profitable sales were recorded in Clothing category. With a year-over-year growth of over 186% Bike category has improved in sales compared to the previous year.

The Total Sales Measure and the YOY % (Year-over-year growth %) were created with the DAX function Illustrated below.

**Total Sales** =

CALCULATE (

SUMX (Sales,

Sales[OrderQuantity]\*Sales[ProductPrice]

),

Customers

)

**YOY Growth %** =

VAR vDiff=[Total Sales]-[Total Sales LY]

RETURN

DIVIDE (

vDiff,

[Total Sales LY]

)

**Total Sales LY** =

CALCULATE (

[Total Sales],

DATEADD (

'Calendar'[Date],

-1,

YEAR

)

)

Where:

* SUMX () returns the sum of an expressions evaluated in each row in a table
* DIVIDE () is a function that divides with the ability to handle divide by zero error
* DATEADD () moves the given set of dates by a specific interval

Map

Description automatically generated

Figure 4: Sales by Country

**By Continent and Country**

Referring to Fig. 4 above, a total of 10 thousand customers across North America, European and Pacific were recorded as the customer base where the more sales were obtained from the United States and Australia. The USA is the highest performing country with more profitable sales recorded from the northeast and southern part of the country. The least profitable sales were recorded in Canada, UK, Germany, and France. Even though USA contributes the highest sales, the gross margin is 1.43% lower than Australia, thus proposes to do priority sales based on the regions.

The sales by country were visualized on a map visual created using the **Continent** and the **Country** column in the **territory** table as the axis and the **Total Sales** measure created as the value (The DAX is as explained above).

Chart, line chart

Description automatically generated

Figure 5: Sales Trend in the Current year

A screenshot of a computer screen

Description automatically generated with medium confidence

Figure 6: A comparison of the Total Sales Last year and the Current year

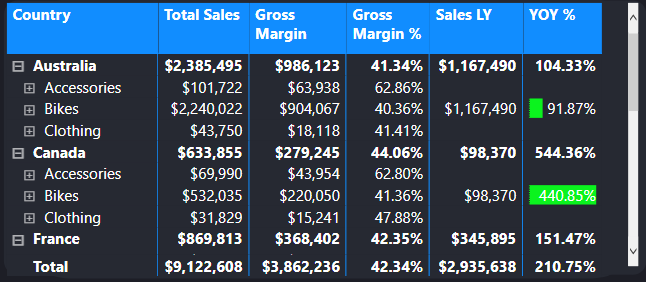


Figure 7: A matrix visual comparing sales growth based on country and product category

**Sales trend and year-over-year growth comparison**

Looking at Figure 5 above, ﻿Monthly revenue increased by 44.11% and trended up between January and June 2017 and is forecasted to have grown between $2.5M and $2.6M by December of FY17 from the current month. The forecast line and its confidence interval in this chart were created in the forecast section in the analytics tab with the following parameters: the forecast length of 6 months and 95% confidence. The forecast shows that there will be a lot of sales in the next 6 months, this will help the Production manager and Sales department in inventory management.

Also figure 6 is a line and clustered column chart, with the lines showing the monthly growth rate compared to the previous year and the clustered column showing the current year monthly sales. The growth rate over the FY16 is at 210.75% compared to that of the previous year which was about 45% over FY 15, when this increase was analyzed or investigated, it was discovered that this is due to the accessories and clothing categories introduced starting from July 2016.

Figure 7 shows the sales performance of each product categories in the different geographical region comparing sales in FY16 to FY17 by country and product category. It was interesting to discovered that though most of the revenue were from the Bikes category however the **Accessories** category with a **gross margin %** of **63%** is more profitable than the **Bikes** category with **gross margin %** of **41%** as shown in the figure 8 below.

A screenshot of a computer

Description automatically generated with medium confidence

Figure 8: A matrix visual comparing sales growth based on product category

This matrix visual were created using the **Country** column on the territory column and the **Category** column on the product table on the visual axis and then the **Total Sales, Gross Margin, Gross Margin %, Sales LY and YOY growth %.** The Gross Margin, Gross Margin % and Sales LY are measures created with the DAX functions shown below.

**Cost of Goods** =

CALCULATE(

Sumx(Sales,Sales[OrderQuantity]\*Sales[ProductCost]),

customers

)

**Gross Margin** =

[Total Sales]- [Cost of Goods]

**Gross Margin %** =   
Divide ([Gross Margin],[Total Sales])

Where:

* CALCULATE () evaluates an expression in a context modified by filters
* SUMX () returns the sum of an expressions evaluated in each row in a table
* DIVIDE () is a function that divides with the ability to handle divide by zero error.

By right clicking on any of the fields such as **Month name**, **Continent**, **country**, **Product category**, **product subcategory** in the dashboard, you could drill through to the product analysis, return analysis and the customer analysis page. For example, as seen in the Figure 9 below. By right clicking on the Bikes category on the donut chart and selecting **Drill through** and then select **Product Analysis.** This takes us to the drilled through page (Product analysis) as shown in figure 10 gives further information about the Bikes Category that was clicked on. We could see on the page that though Bikes category had the major contribution to the revenue, 37 products out of 97 unique products from this category have never been purchase by any customers before.

The matrix on the left shows the products without sales and the table on the right shows the products ranking based on the revenue generated. You can then return to the main page by clicking on the arrow at the top right corner of the page.

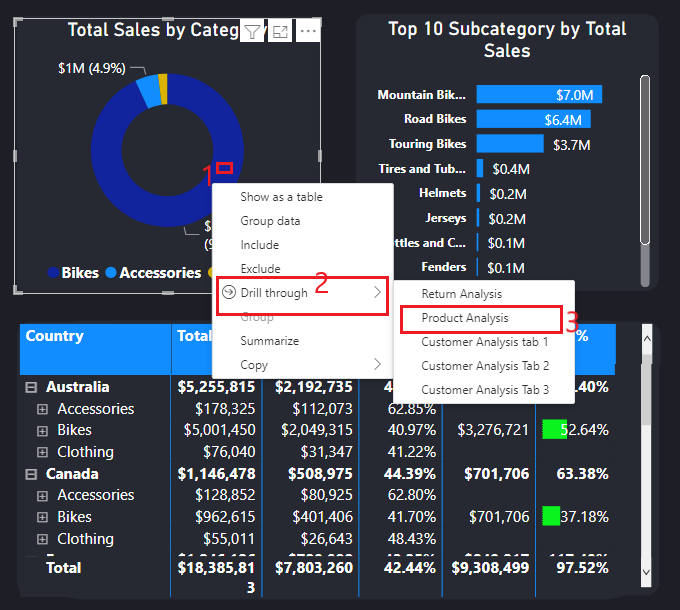


Figure 9: A screenshot shot of how to drill through another page

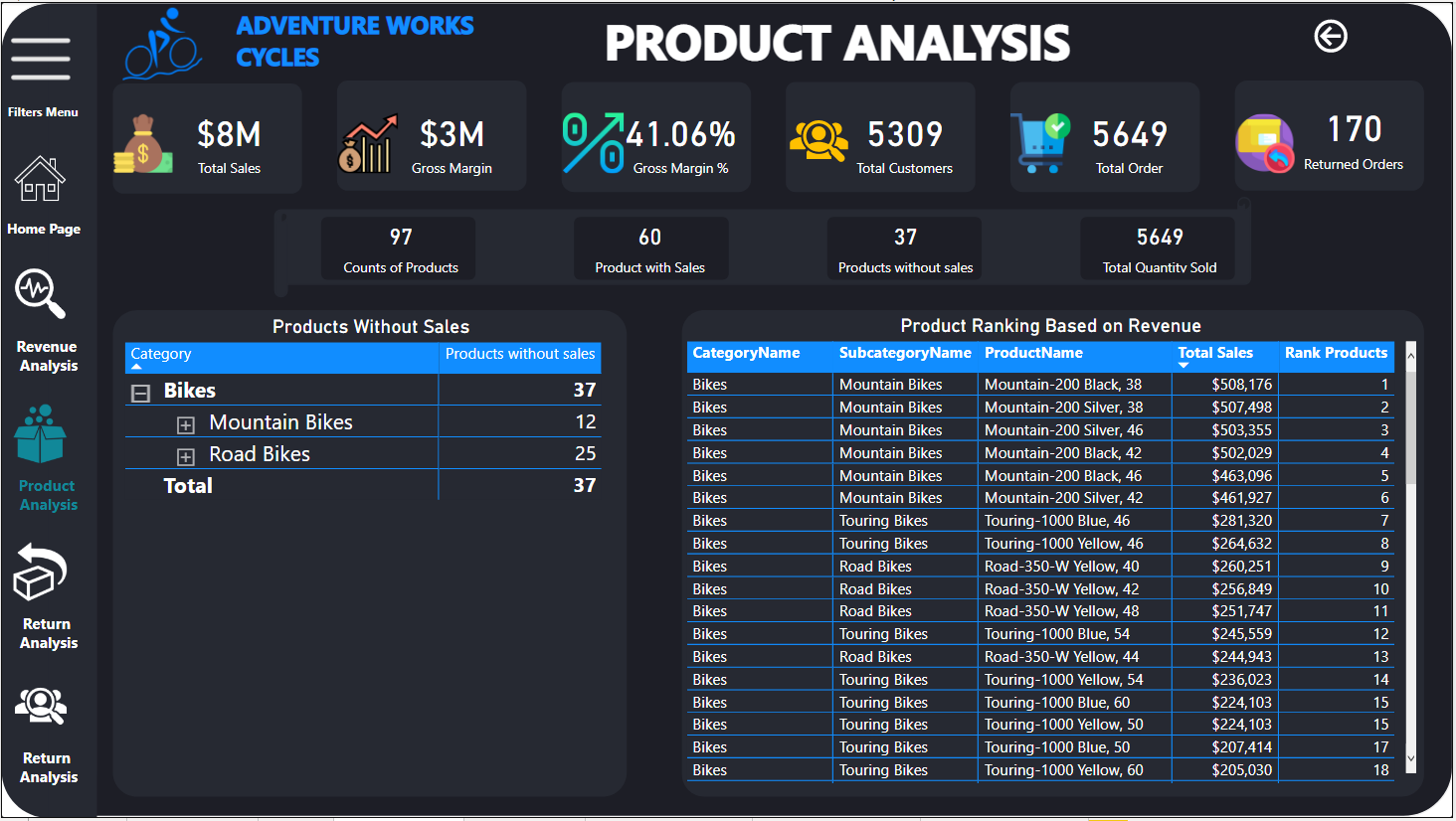


Figure 10: A screen shot of the Drilled through Product Analysis

**2.2 Product Sales Analysis: To understand the Product performance**

Product sales analysis is a judgment on the market performance of a product. For each product sold by the business, it is recommended that a product sales analysis is performed to compare the profit contribution of different products and if certain products do not perform well in generating profits, the organization can decide whether to stop or reduce the investment in that product. On the product sales analysis page, product related data was visualized to determine top performing products and products that are not performing well.

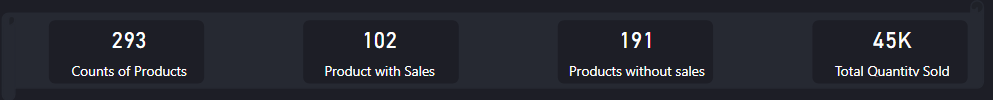


Figure 11: Product Sales Analysis metrics or scorecards

The scorecards illustrated above displayed the **Counts of Products, Products with sales, product without sales** and **Total Quantity of product sold.**

Table

Description automatically generated with medium confidence

Figure 12: Products category and subcategory without sales

A screenshot of a computer

Description automatically generated with medium confidence

Figure 13: Product with sales vs product without sales

Figure 12 shows the Product category, product subcategory and product name of the products that have not been sold, it is insightful to note that there are 191 unique products without sales of which 132 of these products are from the **Components** category. This poses the question of why they are not sold? Is there a need for promotion for these products etc. These questions are beyond the scope of this analysis due to the data available to us. Figure 13 further simplify the comparison between the products with sales and those without sales, from the visual we can deduce that about 76% of the products in the **Accessories** category had sales compare to 62% in **Bikes** category. It was also discovered that no product has been sold from the **Components** category and therefore the marketing department and production section should take necessary actions to review why this category has not been sold.

A picture containing text, scoreboard, black, cabinet

Description automatically generated

Figure 14: Product Ranking Based on the Gross Margin %

Figure 14 shows the Products ranking of all the 102 products with sales based on their gross margin %, it was observed that the **Accessories** and **Clothing** category are at the Top 10 products that are more profitable. The top 10 least profitable products were from the product **Clothing** category.

In summary, although products in the **Bikes category** contributed to **93%** of the Total revenue in FY17, products in the **Accessories category** with gross margin percentage of **63%** are more profitable in the same year. The measures created with DAX functions are illustrated below.

**Rank Products** =

VAR vRank=

RANKX(ALLSELECTED(

Products[ProductName],

Products[CategoryName],

Products[SubcategoryName]

),

[Gross Margin %]

)

RETURN

IF(

ISINSCOPE(Products[ProductName]),

vRank,BLANK()

)

**2.3 Customer Analysis: To understand the Customer profile based on their demographic attributes**

Customer analysis helps businesses to identifies target customers, ascertain the needs of these customers and then specifies how the product satisfies these needs. Customer analysis can be divided into a behavioral profile and a demographic profile. The behavioral profile tells why your product matches a customer’s lifestyle and the demographic describes a customer’s demographic attributes. The customer analysis page on the report, in which customer related data was visualized to determine top performing customer demography contain the following visuals.

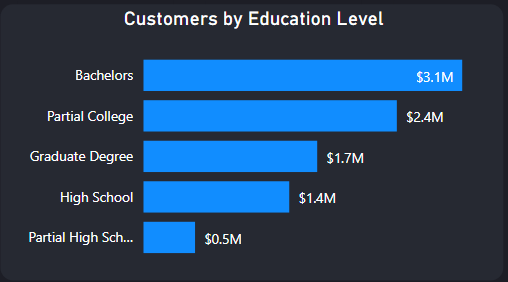


Figure 15: Sales by Customer Education level

Figure 15 above is a clustered bar chart were created using the **Education Level** columnin the customer table as the axis and **Total Sales** as values. Bachelor’s degree holders have the highest sales amount, followed by Partial College, while Partial High School has the lowest sales amount. These findings suggest that customers with bachelor’sdegree and Partial College are more likely to buy our products and this should be taken into consideration by the marketing department during future promotion campaign.

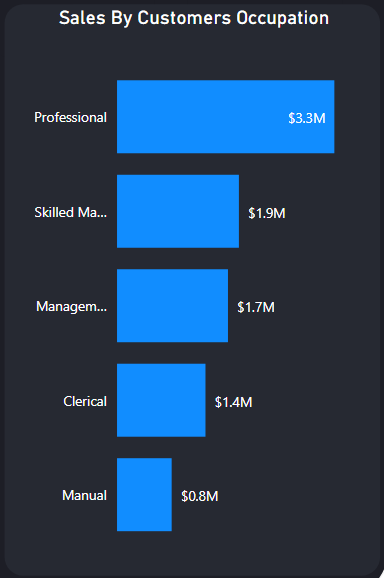


Figure 16: Sales by Customers Occupation

Figure 16 is a clustered bar chart created using the **Occupation** columnin the customer table as the axis and **Total Sales** as values. Professional Customers had the highest sales amount ($3.3M), followed by Skilled Manual customers ($1.9M) and the customers with the lowest sales amount are the Manual customers ($0.8M). These findings suggest that professional and skilled manual customers are more likely to buy our products, and this should be taken into consideration by the marketing department during future promotion campaign.

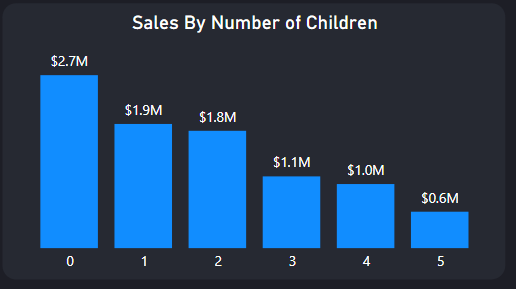


Figure 17: Sales by Number of children

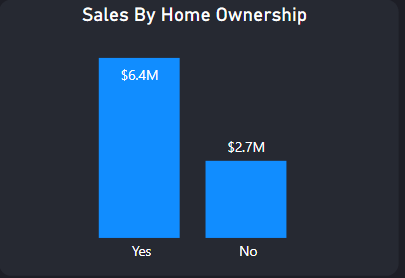


Figure 18: Sales by Home Ownership

Figure 17 revealed that the customers without children or with least number of children are more likely to purchase our product. The figure shows that customers without children have the highest sales amount, followed by those with 1 child as the number of children increases the sales amount decreases. Figure 18 also illustrate sales by home ownership and clearly revealed that Customers that owns their home have the highest sales amount. The customers in this category should be targeted. Figures 17 and 18 are both clustered column chart and were created by using **Number of children** and **Home Ownership** columns respectively on the axis of each and the **Total Sales** measure explained above as the values.

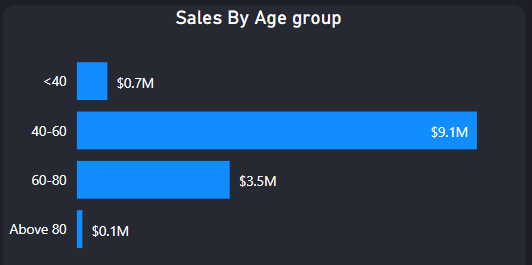


Figure 19: Sales by Age group

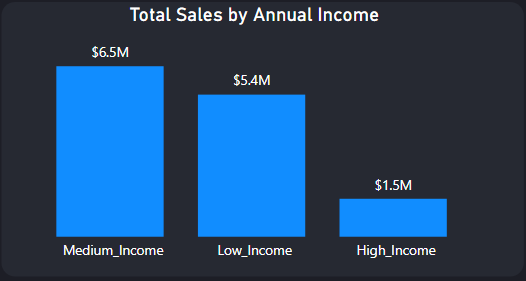


Figure 20: Sales by Annual Income

Figure 19 is a clustered bar chart and was created by using customer **Age\_group** column on the axis (this column is a calculated column grouping the customer age into four categories as explain in section 1) and the **Total Sales** on the values. The visual clearly showed that customers in the age group 40-60 had the highest sales amount followed by customers in the age group 60-80. Figure 20 is a clustered column chart and was created by using customer **Annual Income Group** column on the axis (this column is a calculated column grouping the customer annual income into three categories as explain in section 1) and the **Total Sales** on the values. The figure also clearly shows that customers with medium income had the highest sales amount, followed by low-income earners and then high-income earners with the lowest sales amount. The low-income earners are customers earning less than or equal to $50,000, medium income customers earn between $50,000 to $100,000 and the High-income customers earn above $100,000.

In summary, the figures above have clearly shown the customer demographic profile with the attribute group to focus on or target for sales and the ones that needs promotion campaign. This analysis helps the management to see through who their customers are but however does not give clarity about the customers purchasing behaviors like how frequent does the customer shops? Is the customer a big spender? When was the last time this customer purchased any of our products? This is addressed in the figure 21.

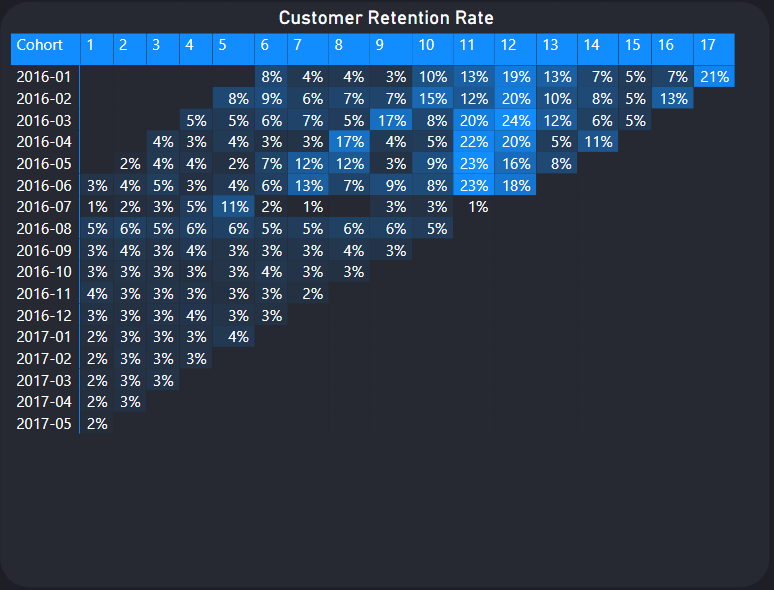


Figure 21: Customer Retention Rate

The true success of marketing is not just to enable a single transactional sale, but in building a customer relationship that spans for as long as possible and that makes the customer retention a high priority for the marketing department in the organization. One of the reasons why this analysis is very important is because new customer acquisition is five times more costly when compared to the cost of retaining existing customers. Figure 21 was created using a matrix visual with the customer’s **First Order Date** column on the Sales tableas the customers **cohort** and is used on the axis, **Months** is a column on the **Month After Acquisition** table (a calculated table created for the purpose of this visual) was used on the columns, and the **Customer retention rate** measure was used as the values. The visual was color coded to emphasize higher rate compared to low rate, this visual displayed customer acquired in FY16 and FY17 and it is seen that none of the customers acquired January 2016 were retained until 6 months. You will also observe that customer cohort between January- June 2016 came back at a higher rate between 10 to 13 months later, as the only product sold for these customers are **Bikes**. These insights will enable the marketing department to use cohort analysis to identify or spot when the rate drop has been significant. The drop can then be traced back to specific activities carried out during the month. Below is the DAX function for calculation the customer retention rate.

**Customer Retention %** =

VAR CurrentMonthAfter=SELECTEDVALUE('Months After Customer Acquisition'[Months])

VAR CurrentFirstOrderMonth=SELECTEDVALUE(Sales[First Order Date(EOM)])

RETURN

DIVIDE (

CALCULATE (

DISTINCTCOUNT(Sales[CustomerKey]),

Customers,

FILTER (

Sales,

EOMONTH(Sales[OrderDate],0)=EOMONTH(CurrentFirstOrderMonth,CurrentMonthAfter)

)

),

CALCULATE(

DISTINCTCOUNT(Sales[CustomerKey]),

Customers

)

)

**Months After Customer Acquisition** =

GENERATESERIES (1,

DISTINCTCOUNT (Sales[First Order Date(EOM)])-1,

1

)

**2.4 Customer analysis: To dynamically segment and understand the customer segments based on their Recency, Frequency and Monetary value**

A screenshot of a computer

Description automatically generated with medium confidence

Figure 22: Customer Segmentation

Customer segmentation is an effective tool for businesses to closely align their strategy and tactics with, and better target, their customers. Every customer is different, and every customer journey is different, so a single approach often isn’t going to work for all. This is where customer segmentation becomes a valuable process. These customer segmentation groups can also be used to begin discussions of building a marketing persona. This is because customer segmentation is typically used to inform a brand’s messaging, positioning and to improve how a business sells – so marketing personas need to be closely aligned to those customer segments in order to be effective. In this customer segmentation, RFM model was adopted to segment the customers. The RFM Model of customer value uses proven marketing principles to help businesses differentiate between marketing to existing and new customers and helps them create relevant and personalized messaging by understanding user behavior. The model allows the business to segment its customers based on three criteria based on an existing customer’s transaction history, namely:

* Recency (When was the last time a customer purchased a product?)
* Frequency (How often did the customer purchase in a fixed time period?)
* Monetary Value (How much money has the customer spent on your brand so far?)

Figure 22 is a table showing the customer segments with other demographic attributes, the customer score and segment are measures created using DAX functions as shown below.

**Recency\_value** =

CALCULATE (

DATEDIFF (

[Last\_transaction date],

max(Sales[OrderDate]),

Day

),

Customers

)

**Frequency\_Value** = DISTINCTCOUNT (Sales[OrderNumber])

**Monetary\_value** = [Total Sales]

**Customer Segmentation Table** =

summarize(

Sales,

Sales[CustomerKey],

"Recency Value",

[Recency\_value],

"Frequency Value",

[Frequency\_Value],

"Monetary Value",

[Monetary\_value]

)

Graphical user interface, application

Description automatically generated

Figure 23: Customers by segments

Graphical user interface, text, application

Description automatically generated with medium confidence

Figure 24: A tooltip for customer segment

Figure 23 gives a clearer view of the segments existing amidst the **17,293** customers, it revealed that **36.5%** (6,312) of our customers are **new customers,** followed by **Promising Customers 29.8%** (5,160) and **24.6%** (4,262) are our **Champions**. The Figure 24 above is a tooltip that shows the activity and actionable tips for each of the customer segments. It also displays information on the number of customers on this segment and their gross margin %.

**2. 5 Return Analysis: What quantity of the products sold are returned?**

For some retail businesses, high rate of returns is a daily struggle as these returns can severely impact their profits or gross margin. Not only are returns reducing your sales, but they also cost money in form of expenditures. Because even with a powerful returns management software, shipping and handling returns is never cheap. Therefore, reducing return rate can lead to a significantly healthier bottom line. Therefore, in-depth return analysis is so important to know where the returns come from and why they happen, it’s impossible to tackle them successfully.

A screenshot of a computer

Description automatically generated with medium confidence

Figure 25: Continent, Product Category by Returned Orders

Figure 25 is a matrix visual illustrating the return rate at different continent expanded to the product subcategory returned within the continent**. Shorts** subcategorywith a return rate of **9.09%** and **8.82%** in **Europe and Pacific** continent respectively have the highest and alarming return ratefollowed by **Vest** subcategory with a return rate of **6.15%** and **6.67%** also **Europe and Pacific** continent respectively. This insight informs the business management on the need to check what could be the probable reason for the high rate of product return in this region. Figure 25 display the product return rate trend, a higher return rate was observed in the Q1 of the FY17, while trying to compare with Q1 in FY16, I discovered that there was also a high rate of product return as well, this suggest that the high rate could be seasonal and there is therefore a need for the sales department to deliberate action in order to reduce this next Q1.

A screenshot of a computer screen

Description automatically generated with medium confidence

Figure 26: Return Rate Trend

Figure 27 showed that United State had the highest return rate in the FY17 based on the total quantity returned, followed by Australia and Canada at the bottom. However, **France** although with the least sales revenue has the highest return rate of **2.26%.** Figure 28 illustrated the top 10 products returned based on quantity. **Water Bottle – 30 Oz.** being the highest returned quantity followed by **Patch kits/8 patches** and **Mountain Tire Tubes.**

A screenshot of a computer

Description automatically generated with medium confidence

Figure 27: Country by the Total amount of the returned orders

A screenshot of a computer

Description automatically generated with medium confidence

Figure 28: Top 10 Product returned based on Quantity returned

The measures used in this subsection and their DAX functions are below.

* **Total Amount Returned** =

SUMX (

Returns,

Returns[ReturnQuantity]\*Returns[ProductPrice]

)

* **Quantity Returned** =

SUMX (

Returns,

Returns[ReturnQuantity]

)

* **Total Quantity Sold =**

**CALCULATE (**

**sum(Sales[OrderQuantity]),**

**customers**

**)**

* **Return Rate** =

**DIVIDE** (

[Quantity Returned],

[Total Quantity Sold]

)

**2.6 Report support information**

The analytical insight reported is delivered with a Power BI report (\*.pbix). The report includes 8 pages, where each page addresses an aspect in performance evaluation of Adventure Works Cycle Company. This report was published online in Power BI Service platform for every stakeholder in the organization that would require this report. By clicking on the **Publish** button in the **Home** ribbon and selecting a destination (see Figure 26). Figure 27 shows the windows that publish processing and successful publish.

**Sir, YOU HAVE TO PROVIDE THIS FROM YOUR END**

Figure 29:

**3 CONCLUSIONS**

The key Findings are summarized below:

* The Total Revenue generated in the first half of FY17 was $9,122,607, with a year-over-year growth of 211%
* 92% of the Total Revenue generated is from Bike category. With a year-over-year growth of over 186% Bike category has improved in sales compared to the previous year.
* The USA is the highest performing country with more profitable sales recorded from the northeast and southern part of the country. Though USA contributes the highest sales, the gross margin is **1.43%** lower than Australia.
* ﻿Monthly revenue increased by 44.11% and trended up between January and June 2017 and is forecasted to have grown between **$2.5M** and **$2.6M** by December of FY17 from the current month.
* Though most of the revenue were from the Bikes category however the **Accessories** category with a **gross margin percentage** of **63%** is more profitable than the **Bikes** category with **gross margin %** of **41%**
* Based on the product ranking by their profitability (gross margin %), **Accessories** and **Clothing** products are at the Top 10 products that are more profitable. The top 10 least profitable products were also from the **Clothing** productcategory.
* There are **293** products in the company’s inventory, **191** of these products have not been sold before and of which **132** of these products are from the **Components** category, which have never been sold before.
* Products in the **Accessories category** with gross margin percentage of **63%** are more profitable than products in the **Bikes category** contributed to **93%** of the Total revenue in FY17
* Customers within the age of 40 to 80 years with **bachelor’s degree** or **Partial College level of education**, who are **Professionals** or **skilled Manually** earning at most **$100,000** with 2 or a smaller number of children either **male** or **female** that owns a house are more likely to buy the company’s products.
* Customers who bought Bikes are more likely to return between 10 to 13 months after their first purchase
* There are a total of **17,293** customers, **36.5%** (6,312) are **new customers, 29.8%** (5,160) are **Promising Customers** and **24.6%** (4,262) are **Champions.**
* **Shorts and Vests** subcategories had a high return rate in **Europe** and **Pacific.** Customers are more likely to return products in the Q1 compared to other quarters of the year.
* The United States and Australia had the highest return based on the quantity returned. France although with the least sales revenue has the highest return rate of **2.26%. Water Bottle – 30 Oz.** being the highest returned quantity followed by **Patch kits/8 patches** and **Mountain Tire Tubes.**

Based on the above findings of this research, the following conclusions were drawn:

* There is a growing trend in the revenue generated over time with a tremendous year over year growth. Majority of the company revenue is generated from the sales of Bikes with US as the country with the highest revenue generated followed by Australia. Accessories are more profitable than other product categories.
* Although Bikes category contributes to the larger shares of the revenue generated, products in accessories and clothing categories are more profitable and hence the top performing categories based on profit margin. Products in components category have never been sold.
* Majority of the customers that purchased our products bachelor’s degree holder within the age of 40 to 80 years who are without children and are professionals with there own house. They are low or middle annual income earners. Larger percentage of our customers are new customers and Customers that purchased bikes are more likely to return a year after their first purchase.
* Return rate is higher in January to March (Q1) compared to other months in both FY116 and FY17, Shorts and Vests subcategories are returned the most in Europe and Pacific continent. There are more returns in USA and Australia and with France with the highest return rate. Water Bottle – 30 Oz., Patch kits/8 patches and Mountain Tire Tubes are the top 3 products returned

Personal Concision and Challenges:

**Sir, YOU HAVE TO PROVIDE THIS FROM YOUR END**

**4 RECOMMENDATIONS**

The following are recommended for adventure works company,

* Instead of spending resources trying to gain new customers, the company should focus on upselling or cross-selling current customers. This is significantly more effective and cost-efficient as the current customers are already acquainted with the products and are therefore more likely to purchase. A customer appreciation gesture such as special discounts will lead past customers and to their next purchase.
* Customer engagement and loyalty to the company should be prioritized.
* A market basket analysis should be conducted to know complementary products and which product to discount to drive the sales of the other
* Focus should be on United States, Australia and United Kingdom
* Focus should be on Bikes as larger percent of the revenue comes from this category.
* Effective marketing Strategy should be used to drive the sales of most of the products without sales. Based on a strategic plan develop targeted promotions to hit specific customers with ad messages and promotional offers.
* Leverage on varieties of marketing strategies and channels to get your product noticed.
* Regularly review the company inventory to identify redundant products and products with high return rate.