BIG DATA AND BUSINESS INTELLIGENCE

OLADEINBO OLUTAYO TOSIN

B1232321

Executive Summary

The report provides a detailed analysis of Dataco's revenue, profit, customer base, distribution of orders across regions and product categories, major factors influencing product orders and profit, forecast for revenue, profit and orders for the next 3 months, and other information such as the company's most loyal customers and top-selling products. From the data presented in the report, it is evident that Europe generated the highest percentage of revenue, that January had the highest revenue and December had the lowest, the top-selling product category is Fishing, the company's profit margin is 12.14%, and the Consumer segment of customers generates the highest profit. Additionally, Estados Unidos is the country with the highest number of orders, Field & Stream Sportsman 16 Gun Fire Safe is the top-selling product, and Mary Smith is the company's most loyal customer. From the report, it is clear that Dataco can optimize its supply chain operations and improve financial performance through effective management of its supply chain.

INTRODUCTION

Supply chain management is a critical aspect of any business, as it involves the procurement, production, and distribution of goods and services. An efficient supply chain can lead to reduced costs, improved customer satisfaction, and increased profitability. However, managing a supply chain can be a complex task, as it involves coordinating the activities of multiple stakeholders, including suppliers, manufacturers, logistics providers, and customers.

In this report, we will explore the various components of a supply chain, including sourcing, production, logistics, and distribution. We will also discuss the role of technology in modern supply chain management and the challenges and opportunities that businesses face in optimizing their supply chains. By understanding these key concepts, we can gain insight into how businesses can effectively manage their supply chains and improve their overall performance.

DATASET DESCRIPTION

The analysis for this report was conducted using a dataset of supply chains from Dataco Global. The dataset includes detailed information on the major areas of the company, including Provisioning, Production, Sales, and Commercial Distribution. It consists of 180,520 rows and 53 columns and covers daily transactions over a 4-year period. By using this dataset, we were able to track Total Orders, Total Revenue, and Total Sales across various Market Regions. This data allows us to gain a thorough understanding of the company's supply chain operations.

The data was gotten from Kaggle:

https://www.kaggle.com/datasets/shashwatwork/dataco-smart-supply-chain-for-big-data-analysis

The table below provides a description of the columns in the "DataCo Supply Chain Dataset.

S/N	FIELDS	DESCRIPTION
1	Type	Type of transaction made
2	Days for shipping (real)	Actual shipping days of the purchased product
3	Days for shipment (scheduled)	Days of scheduled delivery of the purchased product
4	Benefit per order	Earnings per order placed
5	Sales per customer	Total sales per customer made per customer
6	Delivery Status	Delivery status of orders
7	Late_delivery_risk	Categorical variable that indicates if sending is late (1), it is not late (0).
8	Category Id	Product category code
9	Category Name	Description of the product category
10	Customer City	City where the customer made the purchase
11	Customer Country	Country where the customer made the purchase
12	Customer Email	Customer's email
13	Customer Fname	Customer name
14	Customer Id	Customer ID
15	Customer Lname	Customer lastname
16	Customer Password	Masked customer key
17	Customer Segment	Types of Customers: Consumer , Corporate , Home Office
18	Customer State	State to which the store where the purchase is registered belongs
19	Customer Street	Street to which the store where the purchase is registered belongs
20	Customer Zipcode	Customer Zipcode
21	Department Id	Department code of store
22	Department Name	Department name of store
23	Latitude	Latitude corresponding to location of store
24	Longitude	Longitude corresponding to location of store
25	Market	Market to where the order is delivered : Africa , Europe , LATAM , Pacific Asia , USCA
26	Order City	Destination city of the order
27	Order Country	Destination country of the order
28	Order Customer Id	Customer order code
29	order date (DateOrders)	Date on which the order is made
30	Order Id	Order code
31	Order Item Cardprod Id	Product code generated through the RFID reader
32	Order Item Discount	Order item discount value

33	Order Item Discount Rate	Order item discount percentage
34	Order Item Id	Order item code
35	Order Item Product Price	Price of products without discount
36	Order Item Profit Ratio	Order Item Profit Ratio
37	Order Item Quantity	Number of products per order
38	Sales	Value in sales
39	Order Item Total	Total amount per order
40	Order Profit Per Order	Order Profit Per Order
41	Order Region	Region of the world where the order is delivered : Southeast Asia ,South Asia ,Oceania ,Eastern Asia, West Asia , West of USA , US Center , West Africa, Central Africa ,North Africa ,Western Europe ,Northern , Caribbean , South America ,East Africa ,Southern Europe , East of USA ,Canada ,Southern Africa , Central Asia , Europe , Central America, Eastern Europe , South of USA
42	Order State	State of the region where the order is delivered
43	Order Status	Order Status : COMPLETE , PENDING , CLOSED , PENDING_PAYMENT ,CANCELED , PROCESSING ,SUSPECTED_FRAUD ,ON_HOLD ,PAYMENT_REVIEW
44	Product Card Id	Product code
45	Product Category Id	Product category code
46	Product Description	Product Description
47	Product Image	Link of visit and purchase of the product
48	Product Name	Product Name
49	Product Price	Product Price
50	Product Status	Status of the product stock :If it is 1 not available , 0 the product is available
51	Shipping date (DateOrders)	Exact date and time of shipment
52	Shipping Mode	The following shipping modes are presented : Standard Class , First Class , Second Class , Same Day

|	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Type	Days for sl	Days for sl	Benefit pe	Sales per c	Delivery St	Late_deliv	Category I	Category N	Customer	Customer	Customer	Customer	Customer	Customer	Customer	Customer	Customer	Customer
2	DEBIT	3	4	91.25	314.64	Advance s	0	73	Sporting G	Caguas	Puerto Ric	XXXXXXXX	Cally	20755	Holloway	XXXXXXXX	Consumer	PR	5365 Nobl
3	TRANSFER	5	4	-249.09	311.36	Late deliv	1	73	Sporting G	Caguas	Puerto Ric	XXXXXXXX	Irene	19492	Luna	XXXXXXXX	Consumer	PR	2679 Rusti
4	CASH	4	4	-247.78	309.72	Shipping o	0	73	Sporting G	San Jose	EE. UU.	XXXXXXXX	Gillian	19491	Maldonad	XXXXXXXX	Consumer	CA	8510 Rour
5	DEBIT	3	4	22.86	304.81	Advance s	0	73	Sporting G	Los Angele	EE. UU.	XXXXXXXX	Tana	19490	Tate	XXXXXXXX	Home Offi	CA	3200 Amb
6	PAYMENT	2	4	134.21	298.25	Advance s	0	73	Sporting G	Caguas	Puerto Ric	XXXXXXXX	Orli	19489	Hendricks	XXXXXXXX	Corporate	PR	8671 Iron
7	TRANSFER	6	4	18.58	294.98	Shipping c	0	73	Sporting G	Tonawand	EE. UU.	XXXXXXXX	Kimberly	19488	Flowers	XXXXXXXX	Consumer	NY	2122 Hazy
8	DEBIT	2	1	95.18	288.42	Late deliv	1	73	Sporting G	Caguas	Puerto Ric	XXXXXXXX	Constance	19487	Terrell	XXXXXXXX	Home Offi	PR	1879 Gree
9	TRANSFER	2	1	68.43	285.14	Late deliv	1	73	Sporting G	Miami	EE. UU.	XXXXXXXX	Erica	19486	Stevens	XXXXXXXX	Corporate	FL	7595 Cott
10	CASH	3	2	133.72	278.59	Late deliv	1	73	Sporting G	Caguas	Puerto Ric	XXXXXXXX	Nichole	19485	Olsen	XXXXXXXX	Corporate	PR	2051 Dust
11	CASH	2	1	132.15	275.31	Late deliv	1	73	Sporting G	San Ramo	EE. UU.	XXXXXXXX	Oprah	19484	Delacruz	XXXXXXXX	Corporate	CA	9139 Blue
12	TRANSFER	6	2	130.58	272.03	Shipping c	0	73	Sporting G	Caguas	Puerto Ric	XXXXXXXX	Germane	19483	Short	XXXXXXXX	Corporate	PR	4058 Quie
13	TRANSFER	5	2	45.69	268.76	Late deliv	1	73	Sporting G	Freeport	EE. UU.	XXXXXXXX	Freya	19482	Robbins	XXXXXXXX	Consumer	NY	3243 Shad
14	TRANSFER	4	2	21.76	262.2	Late deliv	1	73	Sporting G	Salinas	EE. UU.	XXXXXXXX	Cassandra	19481	Jensen	XXXXXXXX	Corporate	CA	131 Sunny
15	DEBIT	2	1	24.58	245.81	Late deliv	1	73	Sporting G	Caguas	Puerto Ric	XXXXXXXX	Natalie	19480	Mcfadden	XXXXXXXX	Corporate	PR	2531 Wish
16	TRANSFER	2	1	16.39	327.75	Late deliv	1	73	Sporting G	Peabody	EE. UU.	XXXXXXXX	Kimberley	19479	Sharpe	XXXXXXXX	Corporate	MA	6417 Silve
17	DEBIT	2	1	-259.58	324.47	Late deliv	1	73	Sporting G	Canovana:	Puerto Ric	XXXXXXXX	Sade	19478	Lancaster	XXXXXXXX	Corporate	PR	257 Harve
18	PAYMENT	5	2	-246.36	321.2	Late deliv	1	73	Sporting G	Canovana:	Puerto Ric	XXXXXXXX	Brynne	19477	Giles	XXXXXXXX	Corporate	CA	7342 Hazy
19	CASH	2	1	23.84	317.92	Late deliv	1	73	Sporting G	Paramoun	EE. UU.	XXXXXXXX	Ciara	19476	Bird	XXXXXXXX	Corporate	CA	7787 Lazy
20	DEBIT	2	1	102.26	314.64	Late deliv	1	73	Sporting G	Caguas	Puerto Ric	XXXXXXXX	Bo	19475	Griffin	XXXXXXXX	Consumer	PR	5136 Rusti
21	PAYMENT	0	0	87.18	311.36	Shipping o	0	73	Sporting G	Mount Pro	EE. UU.	XXXXXXXX	Kim	19474	Simon	XXXXXXXX	Consumer	WS	1723 Tawr

OVERVIEW OF BUSINESS QUESTIONS

The main aim of this business intelligence project is to provide Dataco's managers, CEO, and founders with a thorough analysis of their supply chain system and recommend areas for improvement. The focus of the analysis includes:
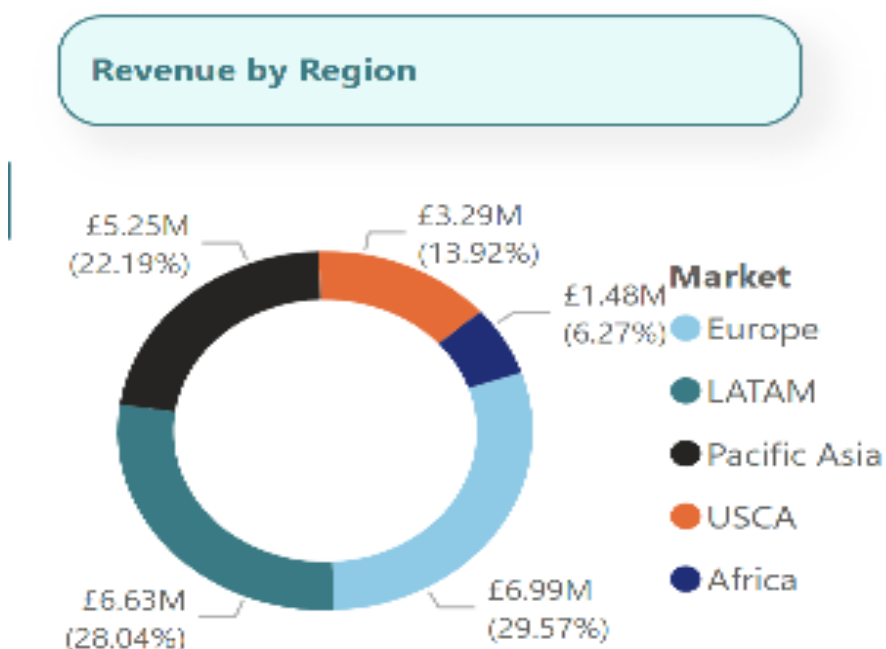
1. What has been the company's revenue, profit, and customer base over the years?
2. Who are the company's most loyal customers?
3. How are orders distributed across regions?
4. How is revenue distributed across regions?
5. What are the top-selling products and product categories?
6. What are the major factors influencing product orders?
7. What are the major factors influencing profit?
8. What is the forecast for revenue, profit, and orders for the next 3 months?"

Overall, the goal of this report is to help Dataco optimize its supply chain operations and improve its financial performance through effective management of its supply chain.
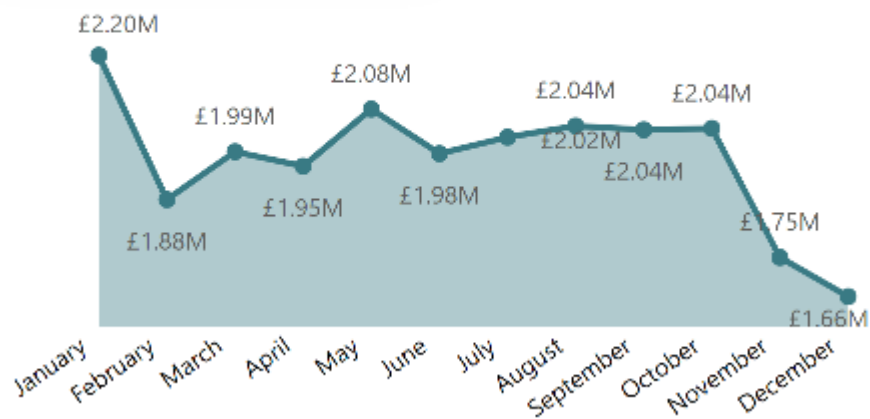
FINDING BASED ON ANALYSIS AND EVALUATION

Revenue	**Net Profit**
£23.64M	£2.87M
Total Order	**Total Customer**
65,752	20,652
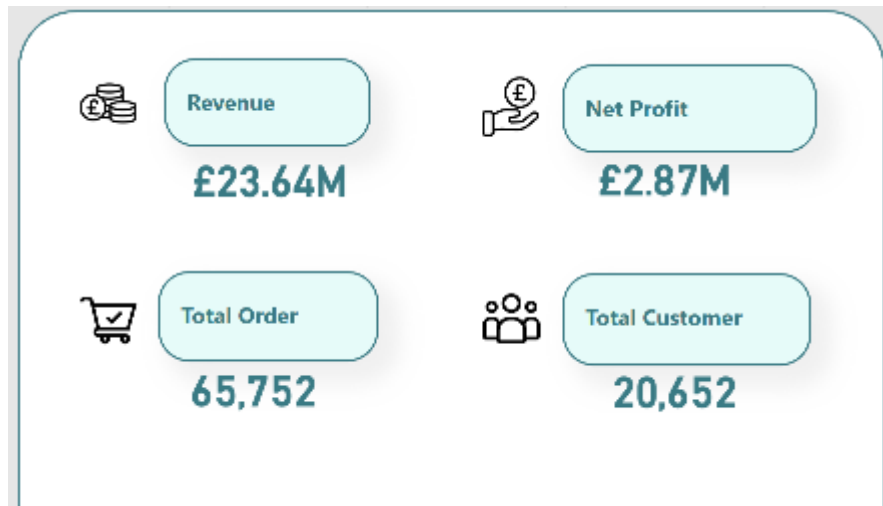
Dataco has achieved a total revenue of £23.64m and a profit of £2.87m, as well as 65,752 orders from its 20,652 customers.

Revenue by Region

£5.25M (22.19%)

£3.29M (13.92%)

£1.48M (6.27%)

£6.63M (28.04%)

£6.99M (29.57%)

Market
- Europe
- LATAM
- Pacific Asia
- USCA
- Africa

The donut chart shows the split of Dataco's revenue across several market regions, with Europe making up the largest portion at 29.57% and Latin America being next with 28.04%.

Monthly Revenue Performance

£2.20M

£2.08M

£1.99M

£2.04M £2.04M

£2.02M

£2.04M

£1.88M

£1.95M £1.98M

£1.75M

£1.66M

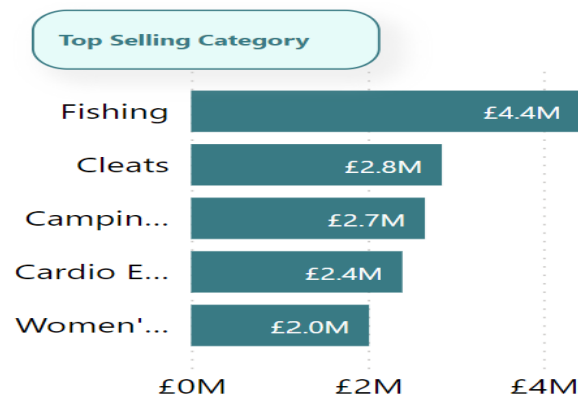January February March April May June July August September October November December

The area line chart details the overall revenue of Dataco each month, with January being the highest and December the lowest.
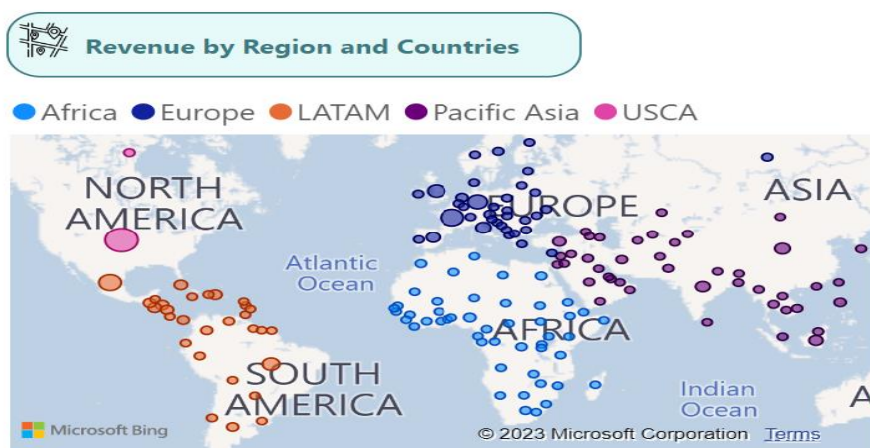
Total Orders per Month

| 0 | 100 | 200 | 300 | 400 | 500 |

Australia 543

India 375

China 361

Indonesia 286

Filipinas 123
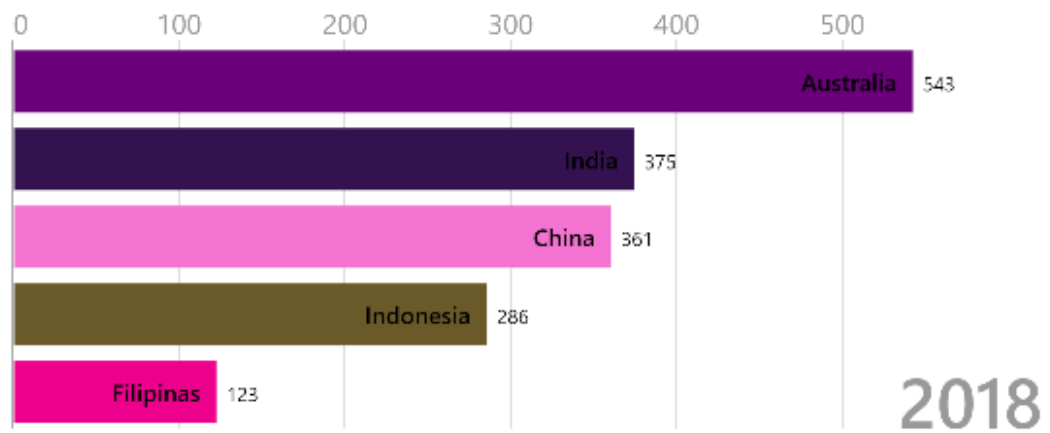
2018

This animated bar chart illustrates the top 5 countries with the highest number of orders from 2014 to 2018. In 2018, Australia had the most orders with a total of 543.

Top Selling Category

Fishing	£4.4M
Cleats	£2.8M
Campin...	£2.7M
Cardio E...	£2.4M
Women'...	£2.0M

£0M   £2M   £4M

This horizontal bar chart reveals the top 5 product categories, with Fishing being the highest-selling one.

Total Cost of Sales

£21M

Profit Margin

0%   12.14%   100%

- Dataco's profit margin stands at 12.14%, while the cost of sales is £21m..

Revenue by Region and Countries

● Africa ● Europe ● LATAM ● Pacific Asia ● USCA

The map visualization reveals that Dataco has a smaller presence in North America compared to other regions, however North America generated the most orders. This data can be used to identify potential growth and expansion opportunities in various regions and nations.

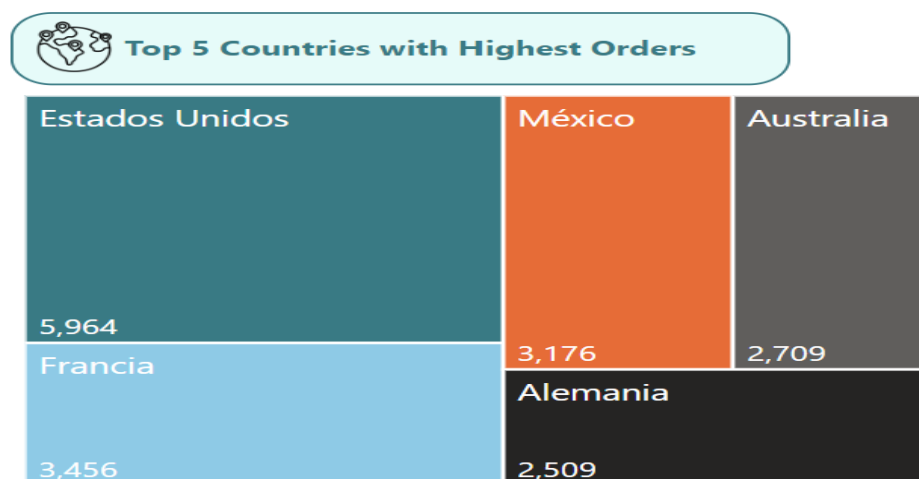| Department Name	Revenue		Cost of Sales	Net Profit		Profit Margin
⊞ Apparel	⚠	£5,134,238.88	£4,492,804.01	↘	£641,434.87	12.49%
⊞ Book Shop	◆	£7,785.25	£7,079.27	↓	£705.98	9.07%
⊞ Discs Shop	◆	£148,071.92	£130,041.88	↓	£18,030.04	12.18%
⊞ Fan Shop	●	£10,952,636.09	£9,639,330.04	↑	£1,313,306.05	11.99%
⊞ Fitness	◆	£256,978.36	£223,812.51	↓	£33,165.85	12.91%
⊞ Footwear	◆	£2,594,912.45	£2,288,112.47	↓	£306,799.98	11.82%
⊞ Golf	◆	£2,947,998.92	£2,592,040.15	↘	£355,958.77	12.07%
⊞ Health and Beauty	◆	£73,120.82	£65,445.05	↓	£7,675.77	10.50%
⊞ Outdoors	▲	£814,887.32	£714,117.49	↓	£100,769.83	12.37%
Total		**£23,642,312.64**	**£20,771,561.74**		**£2,870,750.90**	**12.14%**

Matrix tables (or cross-tabulation tables or pivot tables) are a type of visualization tool used in PowerBI that allow you to summarize, compare and analyse data across multiple dimensions. They enable you to see how metrics like revenue, cost of sales and profit margin differ between product categories and departments, as well as quickly group and aggregate data. Slicing and filtering your data allows you to focus on specific subsets and trends for further analysis. This is demonstrated in the table above, which shows the departments along with their overall revenue, cost of sales, net profit, and profit margin, plus the ability to view the list of product categories within each department and their financial breakdowns. Using matrix tables can help identify areas for improvement and make informed decisions about your business.

Product Category Name	Total Order(s)	Revenue		Profit Margin
⊞ Accessories	1,753	◆	£87,040.44	14.38%
⊞ As Seen on TV!	68	◆	£12,282.23	-2.21%
⊞ Baby	207	◆	£7,408.42	13.25%
⊞ Baseball &	628	◆	£61,948.03	13.67%
⊞ Basketball	64	◆	£19,328.51	6.04%
⊞ Books	405	◆	£7,785.25	9.07%
Total	**65,752**		**£23,642,312.64**	**12.14%**

The matrix table above displays the different Product Categories along with their Orders, Revenue, and Profit Margin. Using the dropdown menu, you can view the list of each product within the category and their individual financial breakdowns.

Top Customers with Highest Orders

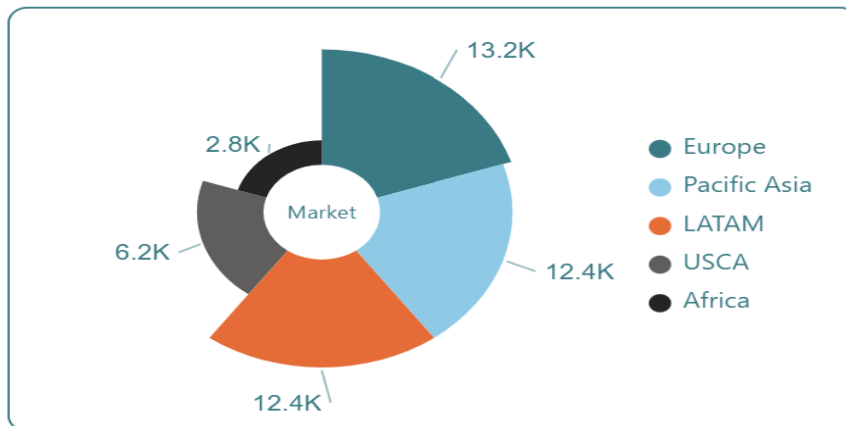| Customer	Orders
Mary Smith	40
Jerry Smith	15
Ashley Smith	14
Deborah Hu...	14
Jacob Guerr...	14
Mary Rios	14
Amanda Ma...	13
Debra Smith	13
Denise Cohen	13
Dorothy Frie...	13
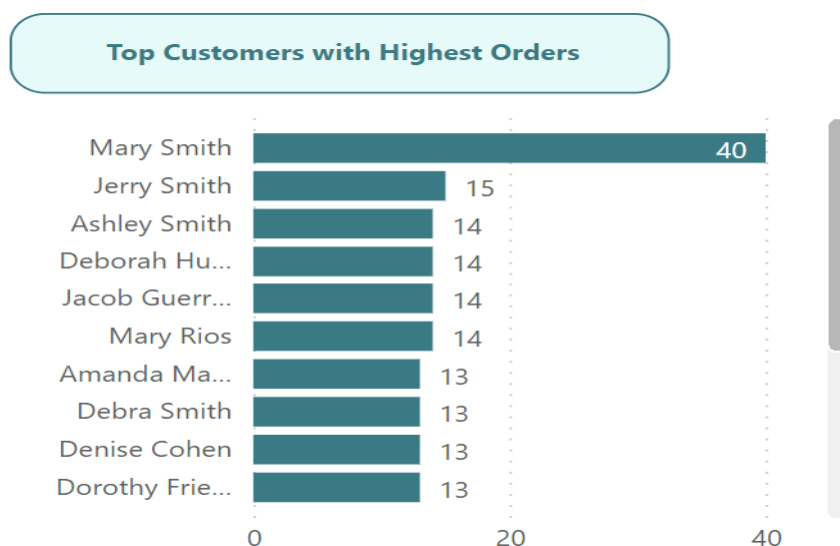
This list of customers shows those with the highest number of orders. Mary Smith is the company's most loyal customer, having placed the highest number of orders.

Top 5 Countries with Highest Orders

Country	Orders
Estados Unidos	5,964
México	3,176
Australia	2,709
Francia	3,456
Alemania	2,509

The treemap below displays the Top 5 countries in terms of orders, with Estados Unidos leading the pack with over 5,964 completed orders.

13.2K

2.8K

6.2K

Market

12.4K

12.4K

- Europe
- Pacific Asia
- LATAM
- USCA
- Africa

The aster plot illustrates the distribution of orders across market regions, with Europe having the highest number of orders at 13.2k, while Pacific Asia and LATAM are tied for second place at 12.4k.

Customer Segments with highest Profit

11.78K
(17.91%)

19.86K
(30.2%)

34.12K
(51.89%)

- Consumer
- Corporate
- Home Office
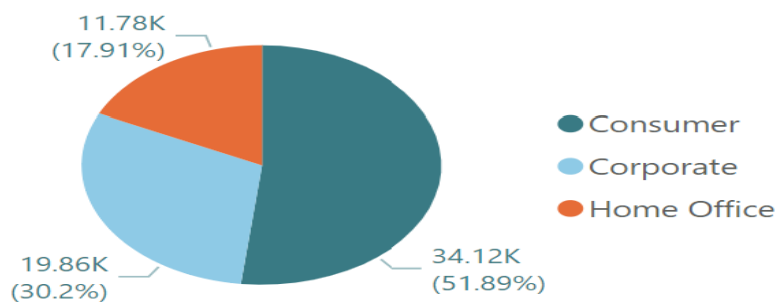
The pie chart showcases the different customer segments of the company and reveals that 51.89% of the company's profit was contributed by the Consumer segment.

Top 5 Selling Products by Profit

Field & Stream Sportsman 16 Gun ...	Diamondback ...	Nike Men's Fr...
£536,959.20	£309,113.92	£284,429.78

Perfect Fitness Perfect Rip Deck — £362,636.43

Nike Men's Dri-FIT Victory Golf P... — £250,497.68

The tree map displays the top 5 company products in terms of orders and reveals that Field & Stream Sportsman 16 Gun Fire Safe has the highest net profit of £536,959.20 with a total of 15,164 orders.
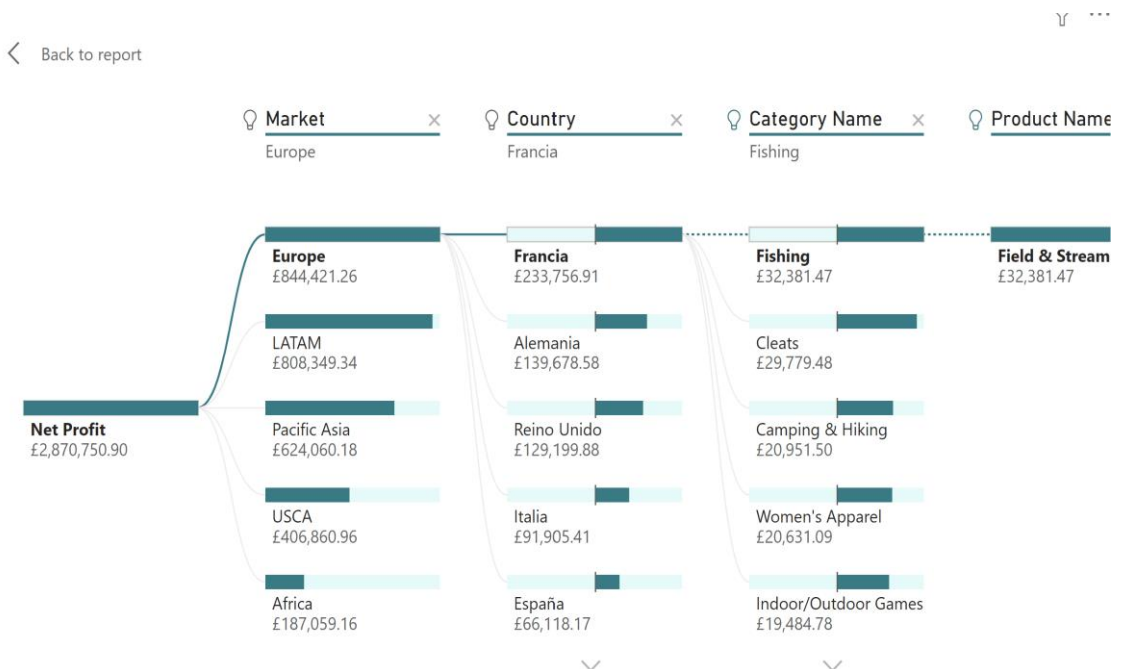
AI Absolute Analysis Decomposition Tree is a type of data analysis tool that uses AI to decompose a metric or value into its constituent parts. It helps to identify which factors have the greatest influence on a given metric or value, by breaking down the metric into its most important components. This can then be used to inform decisions, such as how to allocate resources or optimize strategies.

AI Insight Decomposition Tree Absolute Analysis is a type of data analysis tool that uses AI to break down a metric or value into its most influential components. It helps to identify which factors have the greatest impact on the result, allowing businesses to allocate resources more effectively and optimize strategies for maximum effectiveness. Additionally, it can be used to find hidden intractable relationships between different variables in order to uncover new insights.
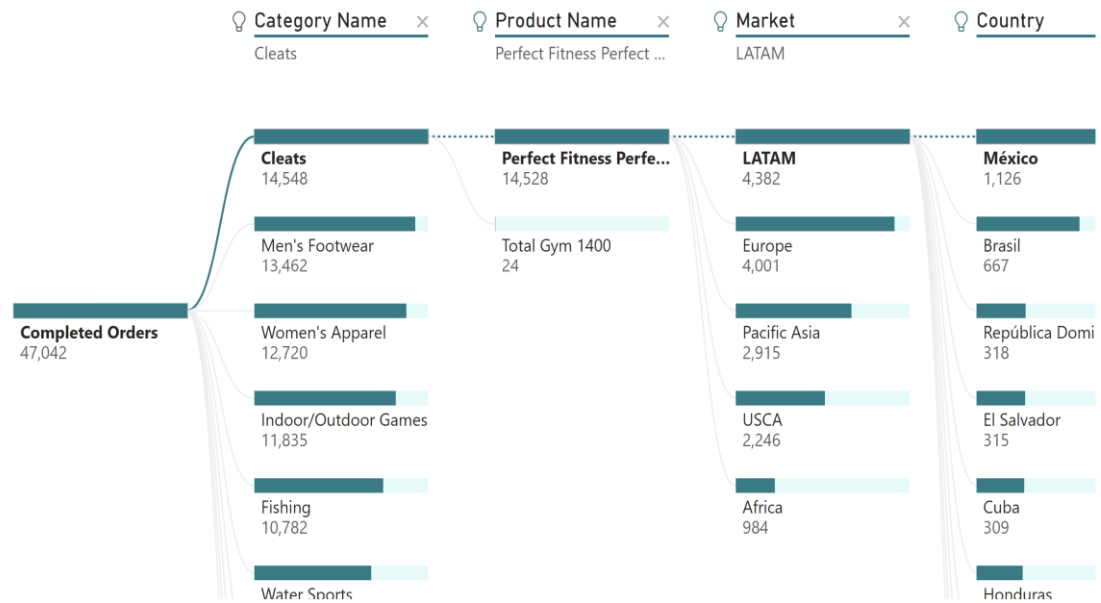
Using AI Insight Absolute Analysis, it can be seen that market has been the most influential factor on profit. This is further broken down to show that country, followed by the category name, and then the product name are the major contributors to this effect.
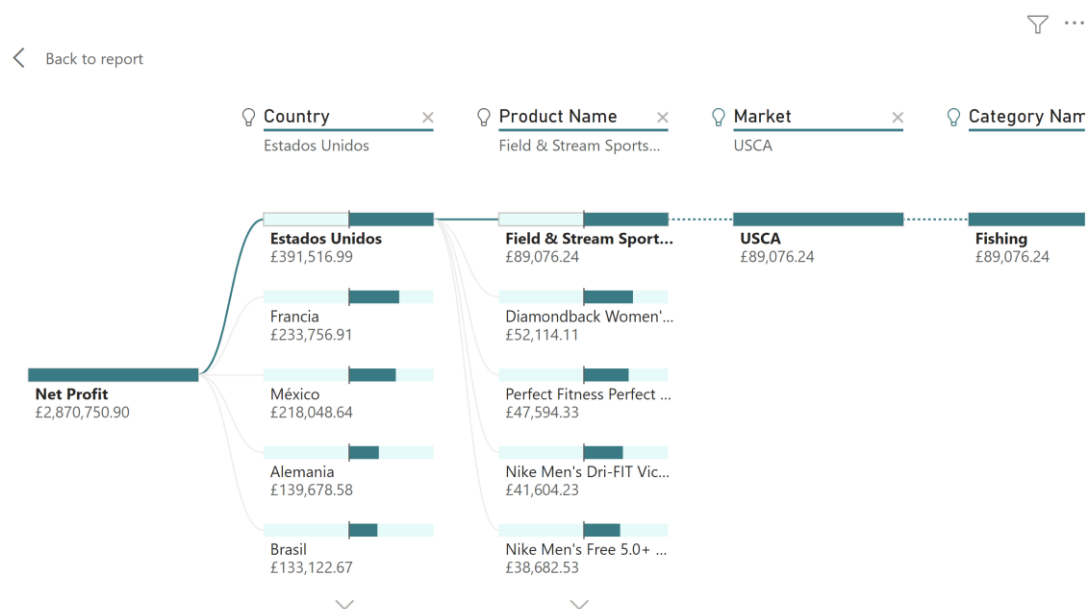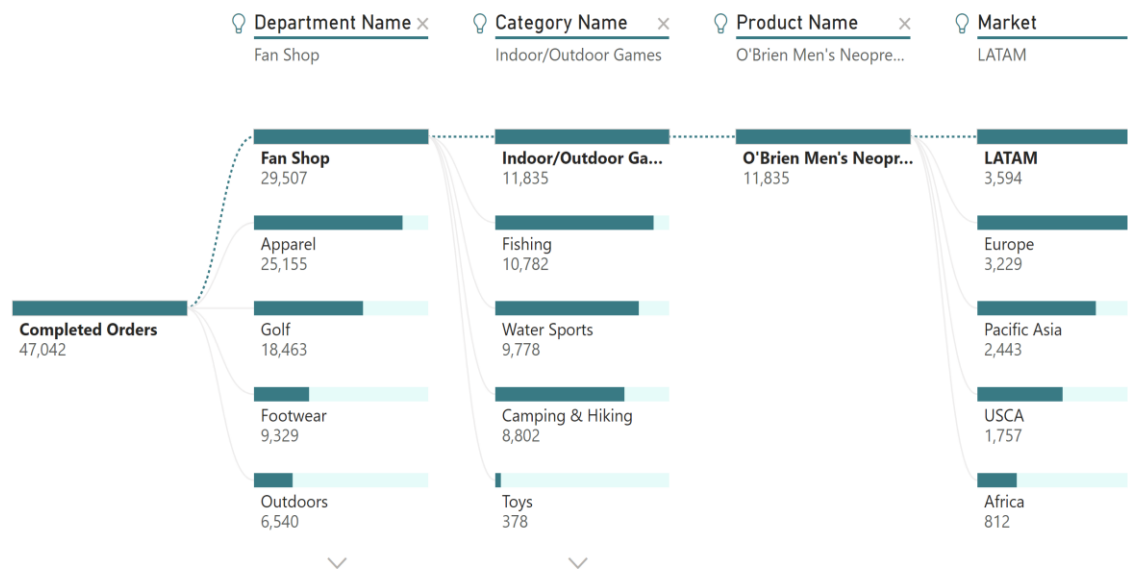

AI Insight Decomposition Tree Relative Analysis is a type of data analysis tool that uses AI to break down a metric or value into its components and measure the relative impact of each component. It helps identify which factors have the most influence on the result and allows businesses to adjust their strategies accordingly. Additionally, it can be used to uncover hidden relationships between different variables, so that new insights can be discovered.

AI Insight Relative Analysis reveals that the factor with the most influence on Profit is the country, followed by the product, then the market and category Name.

Category Name ×	Product Name ×	Market ×	Country
Cleats	Perfect Fitness Perfect ...	LATAM	

Cleats 14,548	**Perfect Fitness Perfe...** 14,528	**LATAM** 4,382	**México** 1,126
Men's Footwear 13,462	Total Gym 1400 24	Europe 4,001	Brasil 667
Women's Apparel 12,720		Pacific Asia 2,915	República Domi 318
Indoor/Outdoor Games 11,835		USCA 2,246	El Salvador 315
Fishing 10,782		Africa 984	Cuba 309
Water Sports			Honduras

Completed Orders 47,042

AI Insight Absolute Analysis indicates that the Product Category has the greatest influence on Completed Orders, with the product, market, and country following.

Department Name ×	Category Name ×	Product Name ×	Market
Fan Shop	Indoor/Outdoor Games	O'Brien Men's Neopre...	LATAM

Fan Shop 29,507	**Indoor/Outdoor Ga...** 11,835	**O'Brien Men's Neopr...** 11,835	**LATAM** 3,594
Apparel 25,155	Fishing 10,782		Europe 3,229
Golf 18,463	Water Sports 9,778		Pacific Asia 2,443
Footwear 9,329	Camping & Hiking 8,802		USCA 1,757
Outdoors 6,540	Toys 378		Africa 812

Completed Orders 47,042

AI Insight Relative Analysis shows that the factor with the most impact on Completed Orders is the Product Department, followed by the Product Category, then Product Name, and finally the market.
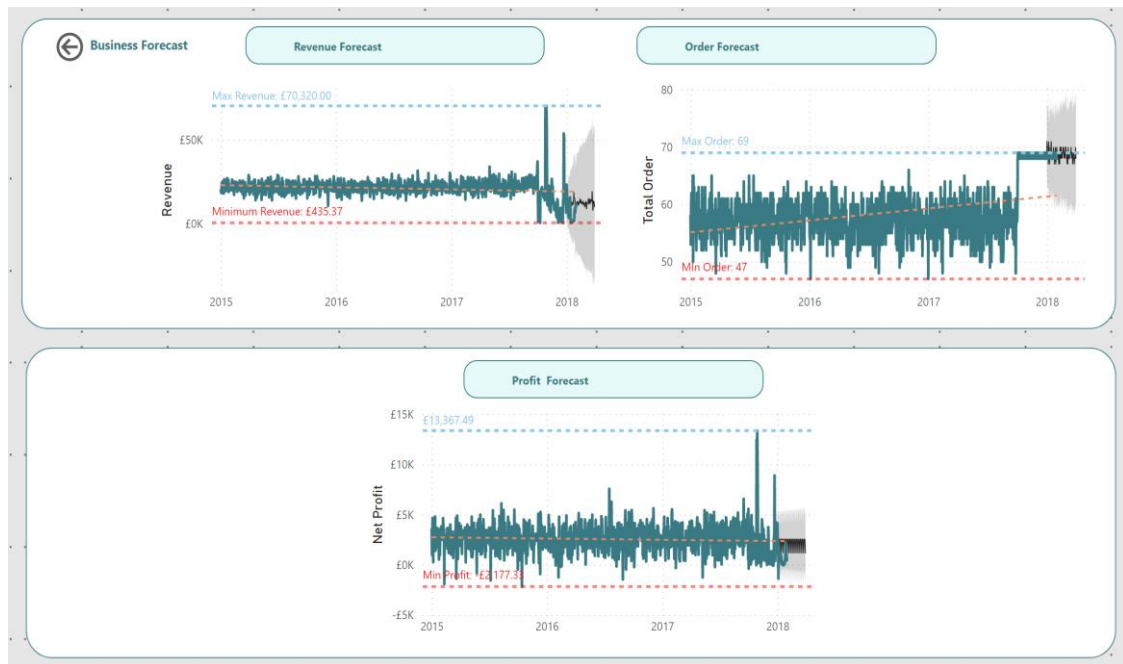
The Business forecast utilizes three line charts to forecast the company's revenue, order count, and profit for the next three months. Given this short timespan, three months was chosen as it enables greater accuracy compared to longer durations.

CONCLUSION AND RECOMMENDATIONS

RECOMMENDATION

Based on the analysis of Dataco's supply chain management, the following recommendations can be made to improve the company's operations and performance:

1.   Develop a clear plan for distribution of orders among different market regions and countries.
2.   Invest in customer loyalty initiatives and promotions to benefit from repeat customers and boost sales.
3.   Identify the best-selling product categories and prioritize inventory and resources accordingly.
4.   Utilize analytics and data-driven insights to optimize product pricing, inventory levels, and distribution strategies.
5.   Leverage AI technology to gain greater visibility into the supply chain process and identify areas of improvement.
6.   Monitor customer feedback and implement strategies to improve customer service and satisfaction.

CONCLUSION

In conclusion, it is clear that Dataco has great potential to optimize its supply chain operations and improve its financial performance. With the right strategies in place, its overall revenue, profit, customer base, distribution of orders across regions and product categories, and other metrics can be improved. Utilizing data-driven insights and AI technology can assist in gaining a deeper understanding of the supply chain system and identifying areas of improvement. Implementing the recommendations outlined above can lead to better management of the supply chain process and greater success for the company.

APPENDIX - BI DESIGN

Data Pre-processing or Data Cleansing

1. To begin pre-processing the DataCo supply chain dataset, it is first imported into Microsoft Power BI.
2. The data was loaded into Microsoft Power BI successfully without any errors.
3. The first step in this process was to identify and review each column in the dataset, which contained a large number of transactions with 52 columns per transaction. We checked for any similar columns with the same values that were represented separately. Following this thinking process, we continued with the cleaning procedure. Thanks to the column description provided by the company, it was easy for us to identify the columns and remove them as part of the cleaning process.
4. We removed any columns that we deemed unnecessary for our analysis. This reduced the dimensionality of the table, making it easier to work with. By simplifying the data in this way, we can more easily focus on the relevant information for our analysis.
5. Upon review, we noticed that there were missing values in the customer's Lastname column. Since the percentage of missing values was less than 1%, we resolved this issue by combining the FirstName and LastName columns into a new column called Fullname.
6. To create the customers table, we duplicated the *fact_transactions* table and selected the relevant columns from the list. The next step is to remove the duplicates from the *CustomerID* column. After the duplicates are removed, we now have unique Customers.
7. The next stage of modelling is to create the Products table, we do this by duplicating the *fact_transactions* table and selected the relevant columns from the list. The next step is to remove the duplicates from the *ProductID* column. After the duplicates are removed, we now have unique Products
8. The next step in the modelling process is to create the *Product_Categories* table. We do this by duplicating the *fact_transactions* table and selected the relevant columns from the list. The next step is to remove the duplicates from the *CategoryID* column. After the duplicates are removed, we now have unique Product Categories
9. The next step in the modelling process is to create the *Product_Department* table. We do this by creating a duplicate of the *fact_transactions* table and selected the relevant columns from the list. The next step is to remove the duplicates from the *ProductDepartmentID* column. After the duplicates are removed, we now have unique Product Department.
10. The next step in the modelling process is to create the Market table. We do this by duplicating the *fact_transactions* table and repeating the steps we have taken

previously. However, we notice that the Market column in the ***fact_transactions*** table does not have an ID. To address this, we remove all duplicates from the Market column in the Market table, then add an index column and rename it Market_Id. Next, we join the Market table with the ***fact_transactions*** table on the Market_Column. This allows us to introduce a new column to the ***fact_transactions*** table called Market_ID, which we will use to merge the ***fact_transactions*** table with the Market_ID. It is generally preferred to use a numerical ID column to join tables because numerical values are smaller in size and faster to compare than text values. Numerical ID columns are also less likely to contain spelling mistakes or formatting errors, which can cause problems when joining tables based on a text column. Using a numerical ID column can improve the performance and reliability of our data analysis.

11. The same process in Step 10 was carried out to create the Region table.
12. The same process in Step 10 was carried out to create the Countries table.
13. The same process in Step 10 was carried out to create the State table.
14. The same process in Step 10 was carried out to create the City table.

DATA MODELLING
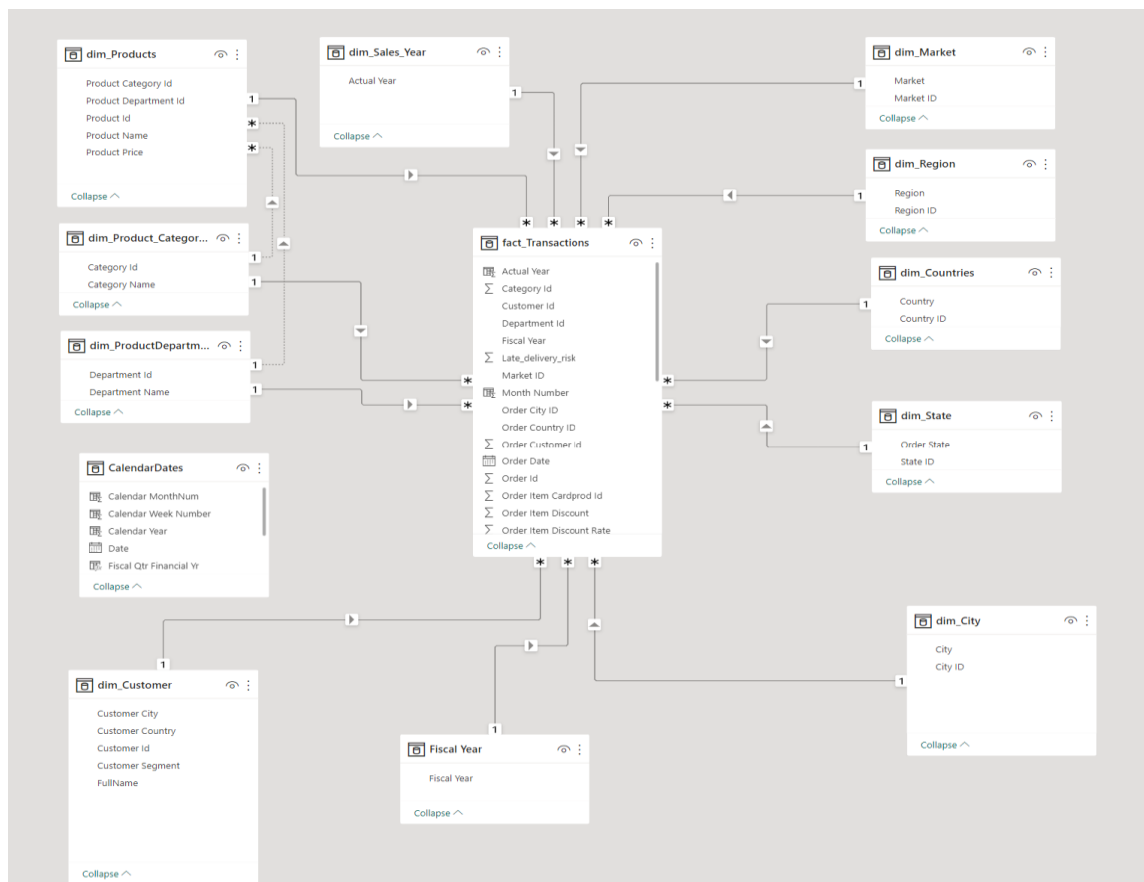
The data modeling process involves normalizing the dataset into tables and organizing them according to a snowflakes schema. This type of database design is optimized for fast query performance and efficient data storage and retrieval, making it ideal for frequently updated and accessed databases such as data warehouses and large-scale online databases. The model contains 13 tables including fact_transactions, dim_Region, dim_Countries, dim_Market, dim_product_categories, dim_productDepartments, dim_products, dim_state, Fiscal Year, dim_sales_year, CalendarDates, dim_customers, and dim_city. These tables provide structured formats for storing and accessing data quickly and easily, allowing for fast insights into the data that can be used to make informed decisions.

Utilizing a calculated table like CalenderDate, we derived the following formula in order to generate specified data points. This allowed us to gain further insights into the data in an efficient manner. By creating such calculated tables, we gained an even more detailed view of the dataset, enabling us to make informed decisions backed by accurate information.

```
CalendarDates = CALENDARAUTO()
```

The formula "CalenderDates = CalendarAuto()" is a function that automatically generates the calender dates based on the given parameters. This means that the user does not have to manually input the data points, saving time and energy by streamlining the process. The generated calender is then used to display the data in an organized structure, enabling easier understanding and analysis of the information.

We went ahead to create some calculated columns

```
Calendar Week Number = WEEKNUM(CalendarDates[Date])
```

The formula is a function that can be used to determine which week a specified date falls into. This formula helps to make data analysis much more organized and efficient, allowing users to quickly obtain valuable insights into the data. With this formula, it's now easier than ever to make informed decisions based on reliable information.

```
Calendar MonthNum = MONTH(CalendarDates[Date])
```

The formula enables users to quickly determine the numerical month corresponding to a specified year. This is especially useful for organizing data in order to gain deeper insights into the information, providing users with all the necessary information needed

to make informed decisions. This formula helps make data analysis much more efficient, enabling users to save time and energy when dealing with large datasets.

```
Calendar Year = YEAR(CalendarDates[Date])
```

The formula enables users to extract the year from the date column and store it into a separate calendar year column. This helps to save time and energy when dealing with large datasets and enables users to gain even deeper insights into the information. The resulting data is more organized, allowing users to make decisions quickly with accurate and reliable information.

```
FISCAL YEAR =
var _FiscalMonthStart = 9
return
IF(
    CalendarDates[Calendar MonthNum]>= _FiscalMonthStart,
    CalendarDates[Calendar Year] + 1,
    CalendarDates[Calendar Year]
    )
```

The formula enables users to determine the fiscal year corresponding to a specified date. For this project, the chosen fiscal year starts from September and ends in August. This formula checks the month a date falls under and if it is greater than or equals to nine (the beginning of a new fiscal year), then 1 is added to the year, resulting in the output being displayed in a new column. If the month is less than nine, then the original year value remains.

```
Fiscal WeekNum =
var _FiscalWeekStart = 36
return
IF(
    CalendarDates[Calendar Week Number]>=_FiscalWeekStart,
    (CalendarDates[Calendar Week Number]) - (_FiscalWeekStart - 1),
    52 + (CalendarDates[Calendar Week Number]) - (_FiscalWeekStart -
1)
    )
```

The formula is used to calculate the week in which a specified date falls under the fiscal year. By first specifying the week that the fiscal year should begin (the 36th week, which is the week starting the 9th month), this formula checks if the date's week number is greater than or equal to the specified week. If so, then the week is restart at one, indicating that the new fiscal week cycle has begun. This formula helps to streamline the process of sorting and organizing data according to the fiscal year, making data analysis much more efficient and accurate.

```
Fiscal Quarter =
SWITCH( TRUE(),
 CalendarDates[Calendar MonthNum] = 1, "Q2",
```

```
    CalendarDates[Calendar MonthNum] = 2, "Q2",
    CalendarDates[Calendar MonthNum] = 3, "Q3",
    CalendarDates[Calendar MonthNum] = 4, "Q3",
    CalendarDates[Calendar MonthNum] = 5, "Q3",
    CalendarDates[Calendar MonthNum] = 6, "Q4",
    CalendarDates[Calendar MonthNum] = 7, "Q4",
    CalendarDates[Calendar MonthNum] = 8, "Q4",
    CalendarDates[Calendar MonthNum] = 9, "Q1",
    CalendarDates[Calendar MonthNum] = 10, "Q1",
    CalendarDates[Calendar MonthNum] = 11, "Q1",
    CalendarDates[Calendar MonthNum] = 12, "Q2",
    BLANK())
```

This formula enables users to check the month number column in the calendar table and output the corresponding quarter value. This formula helps to make data analysis easier and more efficient by organizing the data into quarters, leading to a better understanding and improved decision-making based on reliable information. Additionally, this also helps to provide insights into when certain activities or events are likely to occur and how the data may be affected by seasonal changes. By creating such calculated tables, users can gain deeper insights into the data and make informed decisions with greater accuracy.

```
Fiscal Qtr Financial Yr = CalendarDates[Fiscal Quarter] & "-" & CalendarDates[Fiscal
                                          Year]
```

The formula combines the fiscal quarter and fiscal year values together. This is useful for quickly understanding which year a particular quarter falls into, making data analysis more efficient and accurate. By creating such calculated tables, users are able to gain a deeper understanding of the data, helping them make informed decisions with greater confidence.

```
Fiscal Year = ALLNOBLANKROW(fact_Transactions[FISCAL YEAR])
```

This formula creates a new table that contains the distinct values from the fact_transactions table. This helps to organize and structure the data more efficiently, making it easier to gain valuable insights into the information. Additionally, this also allows users to quickly identify which years are represented in the data, providing greater accuracy when making decisions.

```
The following are the list of measures we created

    Net Profit =
    CALCULATE(SUM(fact_Transactions[Profit]),FILTER(fact_Transactions,
    AND(fact_Transactions[Order Status]<>"ON_HOLD",
        AND(fact_Transactions[Order Status]<>"CANCELLED",
        AND(fact_Transactions[Order Status]<>"PAYMENT_REVIEW",
```

```
        AND(fact_Transactions[Order
    Status]<>"PENDING_PAYMENT",fact_Transactions[Order
    Status]<>"SUSPECTED FRAUD"))))))
```

This is a measure that calculates the net profit of a company. It does this by ignoring orders with status "ON-HOLD", "CANCELLED", "PAYMENT REVIEW", "PENDING PAYMENT", or "SUSPECTED FRAUD" since these are not considered to be successful transactions. Through this calculation, users are able to gain a more accurate view of the company's overall profits, helping them make better decisions with reliable information.

```
Expected Profit = SUM(fact_Transactions[Profit])
```

This formular calculates the expected profit by simply summing the total profit without taking into account the order status. This calculation provides users with more accurate information, enabling them to make more informed decisions backed by reliable data. By creating such calculated tables, users can gain deeper insights into the data and can quickly obtain valuable insights into the company's overall performance.

```
Revenue = CALCULATE(SUM(fact_Transactions[Sales after
Discount]),FILTER(fact_Transactions,
AND(fact_Transactions[Order Status]<>"ON_HOLD",
    AND(fact_Transactions[Order Status]<>"CANCELLED",
    AND(fact_Transactions[Order Status]<>"PAYMENT_REVIEW",
    AND(fact_Transactions[Order
Status]<>"PENDING_PAYMENT",fact_Transactions[Order
Status]<>"SUSPECTED FRAUD"))))))
```

The formula is a measure used to calculate the revenue of a company. It does this by ignoring orders with status "ON-HOLD", "CANCELLED", "PAYMENT REVIEW", "PENDING PAYMENT", or "SUSPECTED FRAUD" since these are not considered to be successful transactions. This calculation enables users to accurately assess the company's financial performance by providing reliable information about their revenues. By creating such calculated tables, users can gain deeper insights into the data and make more informed decisions that are based on accurate information.

DASHBOARD DESIGN

TITLE PAGE

To kickstart the project, I have included a title page featuring the logo of the company and my information along with the school logo. This serves as an introduction to the project and provides a visual representation of the scope of the project. It also serves as a reminder of my commitment to excellence in all my work.

OVERVIEW DASHBOARD

The overview page contains a combination of charts, cards, and page navigations, providing an organized view of the data. Additionally, a slicer is included to enable users

to easily set the year they want to view information. This allows for quick and efficient data analysis, giving users the ability to gain valuable insights into the data. By including such features, users are able to make more informed decisions about their data.

FINANCIAL DASHBOARD
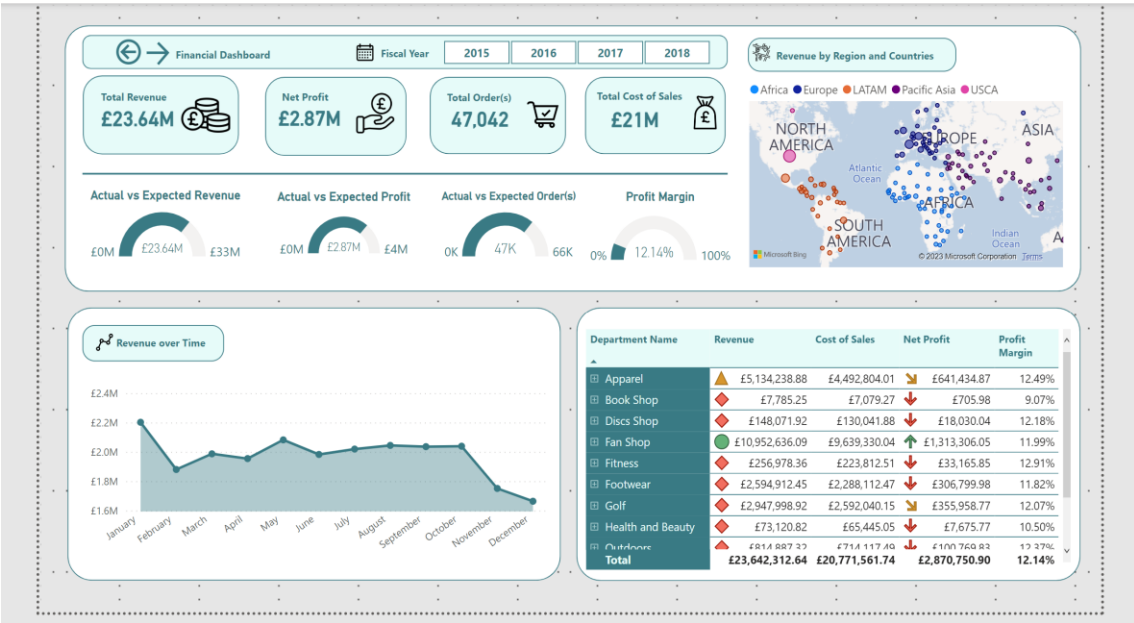
The financial dashboard offers an organized view of the data in the form of charts, cards, and page navigations. To enable users to quickly and efficiently gain insights into the company's finances for a given year, a slicer has been included. The matrix allows the user to evaluate financial details for each product category and drill down further using the drop-down menu.

TRANSACTION DASHBOARD

The transaction dashboard consists of a range of charts that illustrate the company's order count and net profit. It also contains a matrix that enables users to drill down on individual products by category to view their financial statements, along with other charts. To enable the user to easily narrow down the data to the desired year, a slicer tool has been included. Furthermore, the user is able to select any region on the chart and obtain an analysis of that region.

PRODUCT DASHBOARD

The product dashboard offers a combination of different charts, cards, and navigations. It also presents an overview of the top-selling products and order statuses. Additionally, a matrix with a drop-down menu has been included to enable users to view the financial statement for each product within its respective category.

AI INSIGHTS DASHBOARD
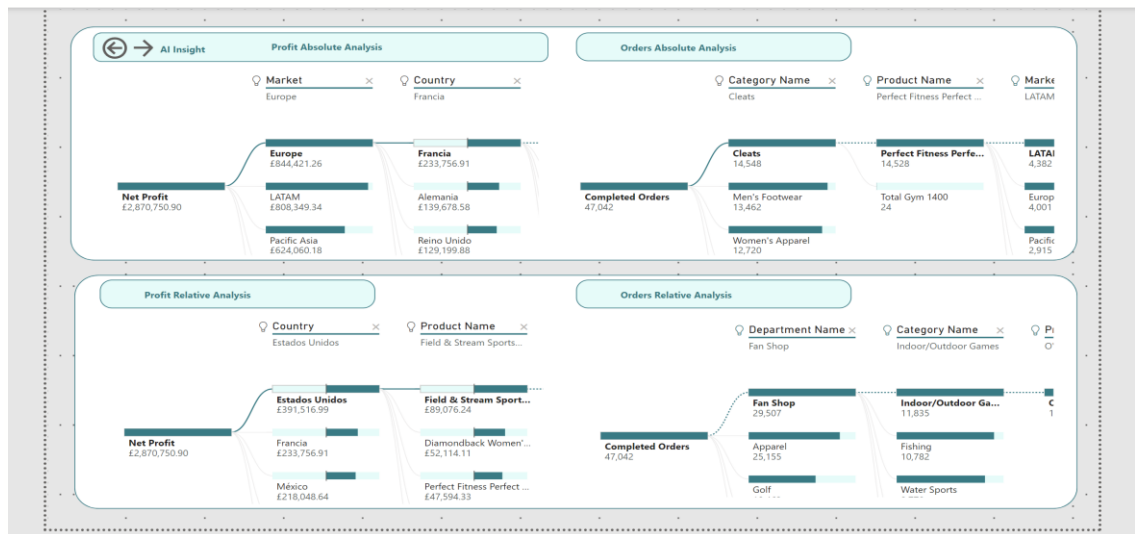
The AI insight dashboard comprises four decomposition trees that offer insightful analyses on various topics.

BUSINESS FORECAST DASHBOARD
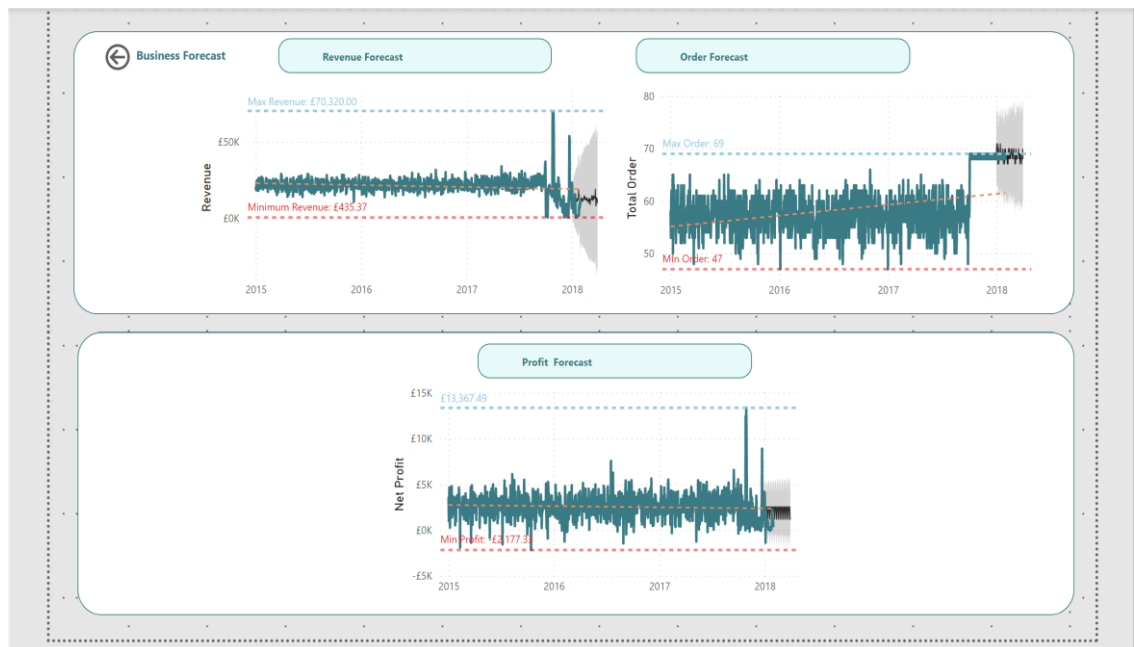
The Business forecast dashboard utilizes three line charts to forecast the company's revenue, order count, and profit for the next three months. Given this short timespan, three months was chosen as it enables greater accuracy compared to longer durations.

Report Section	Description	
Report Structure	The report is well-written, and it contains all the relevant sections	95
Data Pre-processing and Data Modelling	Many pre-processing steps have been applied. The data model is well-structured	93
Dax and M language	Both DAX and M Language have been **extensively** used in the report	95
Dashboard Design	The dashboard contains a variety of charts, including advanced ones not covered in the module.	97
Average		**Add below the average of the four cells above: 95**