BIKES SALES ANALYSIS USING SQL AND POWER BI

Data Analyst Personal Project

This is a personal project aimed at demonstrating proficiency in SQL and Power BI through a comprehensive analysis of bicycle sales data. The project involves developing a detailed SQL database to manage and query the sales data, and creating insightful visualizations and dashboards in Power BI. The goal is to showcase the ability to handle data, derive meaningful insights, and present findings effectively using these tools.

Jefferson Alexander Romero Márquez
romerojefferson1999@gmail.com

Table of Contents

Introduction 1

1. Project Overview 1

2. Objectives 1

3. Scope 1

Methodology 2

1. Data Collection from Kaggle 2

2. Importing Data into SQL 4

3. Dataset Preparation in SQL 5

4. Exploratory Data Analysis in SQL 6

5. Exporting Cleaned Data from SQL 8

6. Importing Data into Power BI 9

7. Creating Interactive Visualizations in Power BI 11

Analysis 13

1. Age Distribution 13

2. Yearly Financial Trends 14

3. Product Profitability 16

4. Monthly Sales Trends 18

5. Financial Performance by Country 19

Conclusions 22

1. Key Insights from the Bicycle Sales Data Analysis 22

2. Reflections on the Project 23

Appendices 24

Introduction

1. Project Overview

This project is a personal initiative designed to showcase my skills in data analysis using SQL and Power BI. The primary focus is on analyzing bicycle sales data to uncover trends, patterns, and insights that can inform business decisions. By developing a SQL database and leveraging Power BI for visualization, this project aims to demonstrate the ability to manage data, perform detailed analysis, and present findings effectively.

2. Objectives

The main objectives of this project are:

- To design and implement a SQL database that effectively stores and manages bicycle sales data.
- To perform detailed data queries and manipulations using SQL to prepare the data for analysis.
- To create informative and interactive visualizations in Power BI that highlight key trends and insights from the sales data.
- To compile a comprehensive report that summarizes the analysis and presents actionable findings.

3. Scope

The scope of this project includes:

- Data collection and preparation from available bicycle sales datasets.
- Database design and development using SQL, including schema creation and query development.
- Data analysis and visualization using Power BI, focusing on trends, performance metrics, and geographic insights.
- Documentation of the entire process, including SQL code, Power BI dashboards, and key findings.

Methodology

1. Data Collection from Kaggle

The first step in this project was to obtain the dataset from Kaggle, a popular platform for data science competitions and datasets. Kaggle provides a vast repository of datasets across various domains, allowing users to download data for analysis and model building.

For this project, we downloaded a CSV (Comma-Separated Values) file from Kaggle. A CSV file is a simple text file format used to store tabular data, where each line represents a record and each field within a record is separated by a comma. This format is widely used due to its ease of use and compatibility with many data processing tools.

The specific dataset used is available at [this Kaggle link](). The CSV file contains bicycle sales data that will be utilized for further analysis and processing in SQL. This file will serve as the primary data source for building and querying the SQL database.

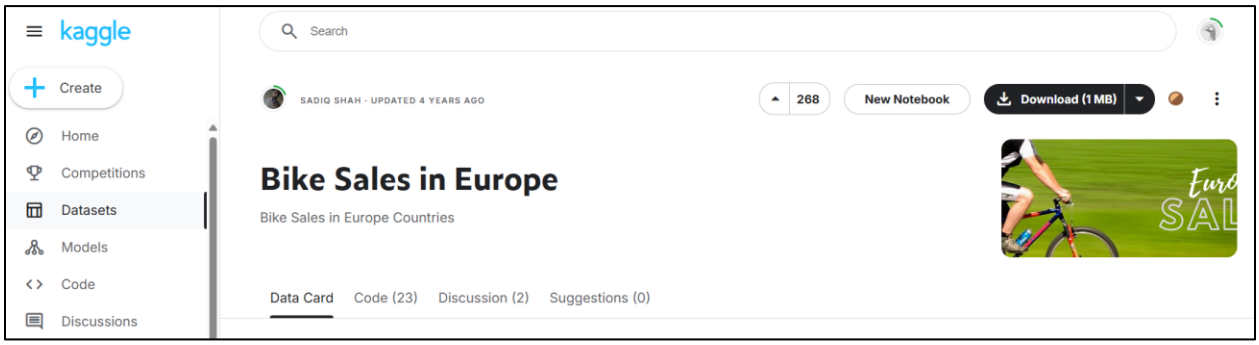


Figure 1: Kaggle Dataset Download Page

To download the dataset, click on the "Download" button located in the dataset interface. This will initiate the download of the CSV file to your local machine, which can then be used for further analysis and processing.

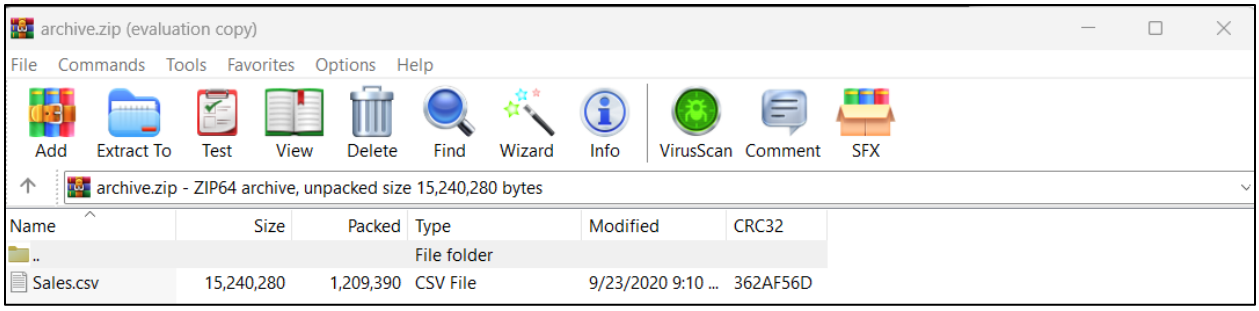


Figure 2: File Downloaded and csv Extraction

After downloading the dataset from Kaggle, the file is saved as a ZIP archive. To access the CSV file, locate the ZIP file on your machine, right-click it, and select "Extract All" to unzip the contents. This will create a folder containing the extracted files, including the CSV file named "Sales.csv," which will be used for further analysis in SQL.

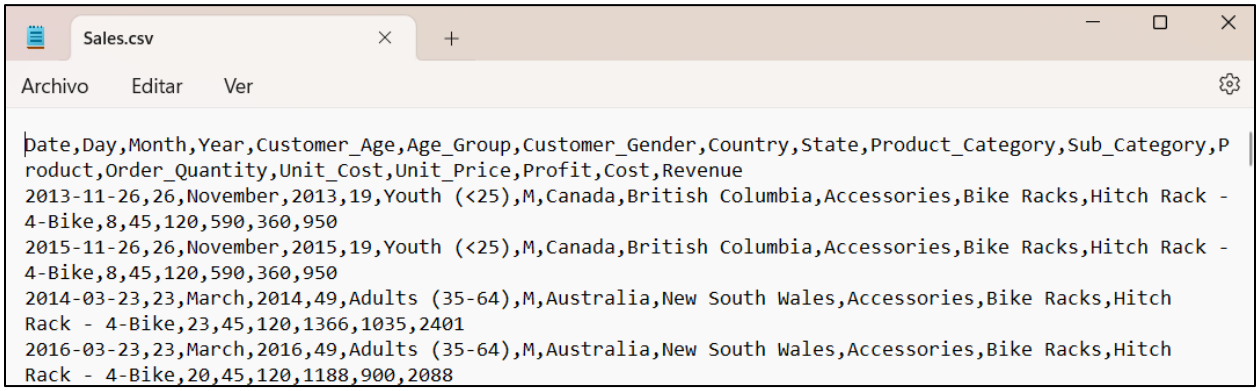


Figure 3: csv File "Sales"

The "Sales.csv" file is a Comma-Separated Values (CSV) file that contains structured tabular data with records and fields separated by commas. This file will be imported into the SQL database for further processing and analysis, enabling insights to be derived from the bicycle sales data.

2. Importing Data into SQL

This is a summarized explanation of how to import data into SQL:

- **Create a New Database:** Start by creating a new database in SQL Server where the data will be stored. You can do this using the SQL Server Management Studio (SSMS) by right-clicking on the "Databases" folder and selecting "New Database." Name the database appropriately (e.g., BikesSales).
- **Import the CSV File:** Use the "Import Data" wizard in SSMS to import the CSV file into a new table. Right-click on the database you just created, go to "Tasks," and select "Import Data." Follow the wizard's steps to choose the CSV file (Sales.csv) and specify the table name where the data will be stored.
- **Map the Columns:** During the import process, ensure that the columns in the CSV file are correctly mapped to the columns in the SQL table. The wizard will show a preview of the data and allow you to make any necessary adjustments.
- **Execute the Import:** Once everything is set up, complete the wizard to import the data. After the process finishes, the data from the CSV file will be available in the specified table within your SQL database.

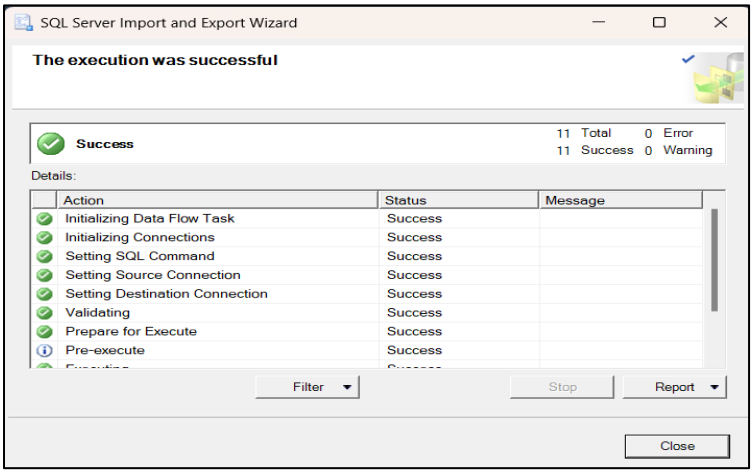


Figure 4: Successful Data Import into SQL

The "Sales.csv" file has been successfully imported into the SQL database without any errors, confirming that the data is now ready for further processing and analysis.

3. Dataset Preparation in SQL

The first step involves visualizing the entire dataset to identify any potential errors. This process is carried out using an SQL query that generates comprehensive tables, allowing us to examine the data closely. By doing so, we can detect and address any inconsistencies or formatting issues that may affect further analysis. The SQL query used for this step is named "Bike_Sales_SQL" and the complete query is available in the project's repository for reference.

Results Messages

	Day	Month	Year	Customer_Age	Age_Group	Customer_Gender	Country	State	Product_Category	Sub_Category	Product	Order_Quantity	Unit_Cost	Unit_Price	Profit	Cost	Revenue
28044	6	May	2016	34	Young Adult...	F	Canada	British Columbia	Clothing	Gloves	"Half-Finger Gloves	M"	29	9	24	428	261,689
28045	15	May	2014	34	Young Adult...	F	Canada	British Columbia	Clothing	Gloves	"Half-Finger Gloves	M"	26	9	24	384	234,618
28046	15	May	2016	34	Young Adult...	F	Canada	British Columbia	Clothing	Gloves	"Half-Finger Gloves	M"	25	9	24	369	225,594
28047	16	March	2014	33	Young Adult...	M	Australia	Tasmania	Clothing	Gloves	"Half-Finger Gloves	S"	7	9	24	73	63,136
28048	16	March	2016	33	Young Adult...	M	Australia	Tasmania	Clothing	Gloves	"Half-Finger Gloves	S"	9	9	24	94	81,175
28049	24	October	2013	33	Young Adult...	M	Australia	Queensland	Clothing	Gloves	"Half-Finger Gloves	M"	12	9	24	134	108,242
28050	24	October	2015	33	Young Adult...	M	Australia	Queensland	Clothing	Gloves	"Half-Finger Gloves	M"	12	9	24	134	108,242
28051	11	July	2014	33	Young Adult...	M	Australia	Queensland	Clothing	Gloves	"Half-Finger Gloves	M"	1	9	24	11	9,20
28052	11	July	2016	33	Young Adult...	M	Australia	Queensland	Clothing	Gloves	"Half-Finger Gloves	M"	1	9	24	11	9,20
28053	1	October	2013	25	Young Adult...	M	Germa...	Nordrhein-Westf...	Clothing	Gloves	"Half-Finger Gloves	S"	6	9	24	71	54,125
28054	1	October	2015	25	Young Adult...	M	Germa...	Nordrhein-Westf...	Clothing	Gloves	"Half-Finger Gloves	S"	5	9	24	59	45,104
28055	24	October	2013	19	Youth (<25)	M	Australia	Victoria	Clothing	Gloves	"Half-Finger Gloves	L"	13	9	24	129	117,246
28056	24	October	2015	19	Youth (<25)	M	Australia	Victoria	Clothing	Gloves	"Half-Finger Gloves	L"	15	9	24	149	135,284
28057	13	December	2013	19	Youth (<25)	M	Australia	Victoria	Clothing	Gloves	"Half-Finger Gloves	L"	10	9	24	100	90,190
28058	13	December	2015	19	Youth (<25)	M	Australia	Victoria	Clothing	Gloves	"Half-Finger Gloves	L"	10	9	24	100	90,190

Figure 5: Dataset Overview Showing Initial Data

The dataset shows an error that occurred during data import, affecting the last six columns: ***Order_Quantity***, ***Unit_Cost***, ***Unit_Price***, ***Profit***, ***Cost***, and ***Revenue***. The problem originated from the ***Product*** column, which contained commas. As a result, the system mistakenly assigned the text after each comma to the next column, leading to a misalignment. This caused ***Order_Quantity*** to display sizes, and the data intended for each subsequent column was shifted incorrectly. By the time it reached the ***Revenue*** column, it contains two values separated by a comma. This misalignment is evident in the dataset.

To resolve this issue using SQL, the following steps were taken:

- ***Creating New Columns:*** New columns (`Product_Fixed`, `Order_Quantity_Fixed`, etc.) were added to hold corrected values without altering the original data.
- ***Removing Quotation Marks:*** Quotation marks were removed from the `Product` and Order_Quantity` columns to clean up the data.
- ***Shifting and Concatenating Data:*** Data that had shifted into incorrect columns due to commas in the `Product` column was realigned, and relevant data was concatenated to restore proper order.

- ***Filling Missing Values:*** Missing values were filled by shifting adjacent data to ensure that all columns contained accurate information.
- ***Dropping Unnecessary Columns:*** Once the data was corrected, the original misaligned columns were removed (`Product`, `Order_Quantity`, `Unit_Cost`, `Unit_Price`, `Profit`, `Cost`, `Revenue`). Additionally, other irrelevant columns (`Date`, `Day`, `Unit_Cost_Fixed`, and `Unit_Price_Fixed`) were dropped to simplify the dataset for exploratory data analysis.

For a more detailed exploration, the complete SQL code is available in the repository under the file named "Bike_Sales_SQL".

Results Messages

	Month	Year	Customer_Age	Age_Group	Customer_Gender	Country	State	Product_Category	Sub_Category	Product_Fixed	Order_Quantity_Fixed	Profit_Fixed	Cost_Fixed	Revenue_Fixed
28044	May	2016	34	Young Adult...	F	Canada	British Columbia	Clothing	Gloves	Half-Finger Gloves- M	29	428	261	689
28045	May	2014	34	Young Adult...	F	Canada	British Columbia	Clothing	Gloves	Half-Finger Gloves- M	26	384	234	618
28046	May	2016	34	Young Adult...	F	Canada	British Columbia	Clothing	Gloves	Half-Finger Gloves- M	25	369	225	594
28047	March	2014	33	Young Adult...	M	Australia	Tasmania	Clothing	Gloves	Half-Finger Gloves- S	7	73	63	136
28048	March	2016	33	Young Adult...	M	Australia	Tasmania	Clothing	Gloves	Half-Finger Gloves- S	9	94	81	175
28049	October	2013	33	Young Adult...	M	Australia	Queensland	Clothing	Gloves	Half-Finger Gloves- M	12	134	108	242
28050	October	2015	33	Young Adult...	M	Australia	Queensland	Clothing	Gloves	Half-Finger Gloves- M	12	134	108	242
28051	July	2014	33	Young Adult...	M	Australia	Queensland	Clothing	Gloves	Half-Finger Gloves- M	1	11	9	20
28052	July	2016	33	Young Adult...	M	Australia	Queensland	Clothing	Gloves	Half-Finger Gloves- M	1	11	9	20
28053	October	2013	25	Young Adult...	M	Germa...	Nordrhein-Westf...	Clothing	Gloves	Half-Finger Gloves- S	6	71	54	125
28054	October	2015	25	Young Adult...	M	Germa...	Nordrhein-Westf...	Clothing	Gloves	Half-Finger Gloves- S	5	59	45	104
28055	October	2013	19	Youth (<25)	M	Australia	Victoria	Clothing	Gloves	Half-Finger Gloves- L	13	129	117	246
28056	October	2015	19	Youth (<25)	M	Australia	Victoria	Clothing	Gloves	Half-Finger Gloves- L	15	149	135	284
28057	December	2013	19	Youth (<25)	M	Australia	Victoria	Clothing	Gloves	Half-Finger Gloves- L	10	100	90	190
28058	December	2015	19	Youth (<25)	M	Australia	Victoria	Clothing	Gloves	Half-Finger Gloves- L	10	100	90	190

Figure 6: Dataset Overview After Data Cleaning

Following the data cleaning process, the dataset now displays accurately aligned columns, with all quotation marks and misalignments removed. Each value is correctly placed in its respective field, ensuring that the dataset is prepared for further analysis and exploration.

4. Exploratory Data Analysis in SQL

With the dataset cleaned, we proceeded to perform exploratory data analysis to gain insights from the data. Five key analyses were conducted to better understand various aspects of the bike shop's operations. The complete SQL code for these analyses can be reviewed in the "Bike_Sales_SQL" file in the repository.

1. Age Distribution

This analysis examines the total number of customers based on age groups and gender. The `Age_Group` column provides a breakdown of the number of female and male customers in

each age group, along with the total for each group. This insight helps in understanding customer demographics and tailoring marketing strategies.

2. Yearly Financial Trends

We analyzed financial performance by year to observe trends in total cost, profit, and revenue. This analysis highlights how these financial metrics have evolved over time, revealing patterns such as peaks and troughs in performance. Understanding these trends is crucial for identifying periods of growth or decline. The SQL code for this analysis provides a comprehensive view of annual financial performance.

3. Product Profitability

This analysis focuses on the profitability of products, grouped by product category and sub-category. It includes:

- **Total Profit by Product Category:** Summarizes profit by each product category.
- **Total Profit by Sub-Category:** Details profit for each sub-category within product categories.
- **Profit by Individual Products:** Breaks down profit by each product within its category and sub-category.

These insights help identify which product categories and specific products are most profitable. The SQL code allows for a detailed view of profitability at various levels of granularity.

4. Monthly Sales Trends

We examined total sales per month to understand sales performance throughout the year. This analysis provides insights into seasonal trends and helps in forecasting future sales. The SQL code includes a breakdown of sales by month and a total for the year, which will be visualized more effectively in Power BI.

5. Financial Performance by Country and State

This analysis assesses financial performance based on geographical location, with a focus on revenue. It includes:

- **Total Revenue by Country:** Summarizes revenue by each country.
- **Revenue by State within Each Country:** Provides revenue details for each state within the countries.

Understanding geographic performance is essential for regional market analysis and strategic planning. The SQL code offers a clear view of revenue distribution across different locations. The results of these analyses, along with their visualizations, are presented in the tables and charts provided later in this report.

5. Exporting Cleaned Data from SQL

Once the exploratory analysis is complete, it's time to export the cleaned file, which we will name "Sales_Fixed." Here is a summarized explanation of how to export data from SQL:

- **Prepare the Data for Export:** Ensure that the data in your SQL table is clean and correctly formatted. Perform any necessary transformations or fixes before exporting.
- **Export the Data:** Use SQL Server Management Studio (SSMS) to export the data. Right-click on the database containing the cleaned table, go to "Tasks," and select "Export Data." The Export Wizard will guide you through the process.
- **Select the Export Destination:** Choose the destination format for the exported file. You can select options like CSV, Excel, or others. Specify the file name (e.g., Sales_Fixed.csv) and the location where you want to save the file.
- **Map Columns and Complete Export:** Ensure that the columns in the SQL table are correctly mapped to the columns in the export file. Complete the Export Wizard to finalize the export process. After finishing, the cleaned data will be saved in the chosen file format and location.

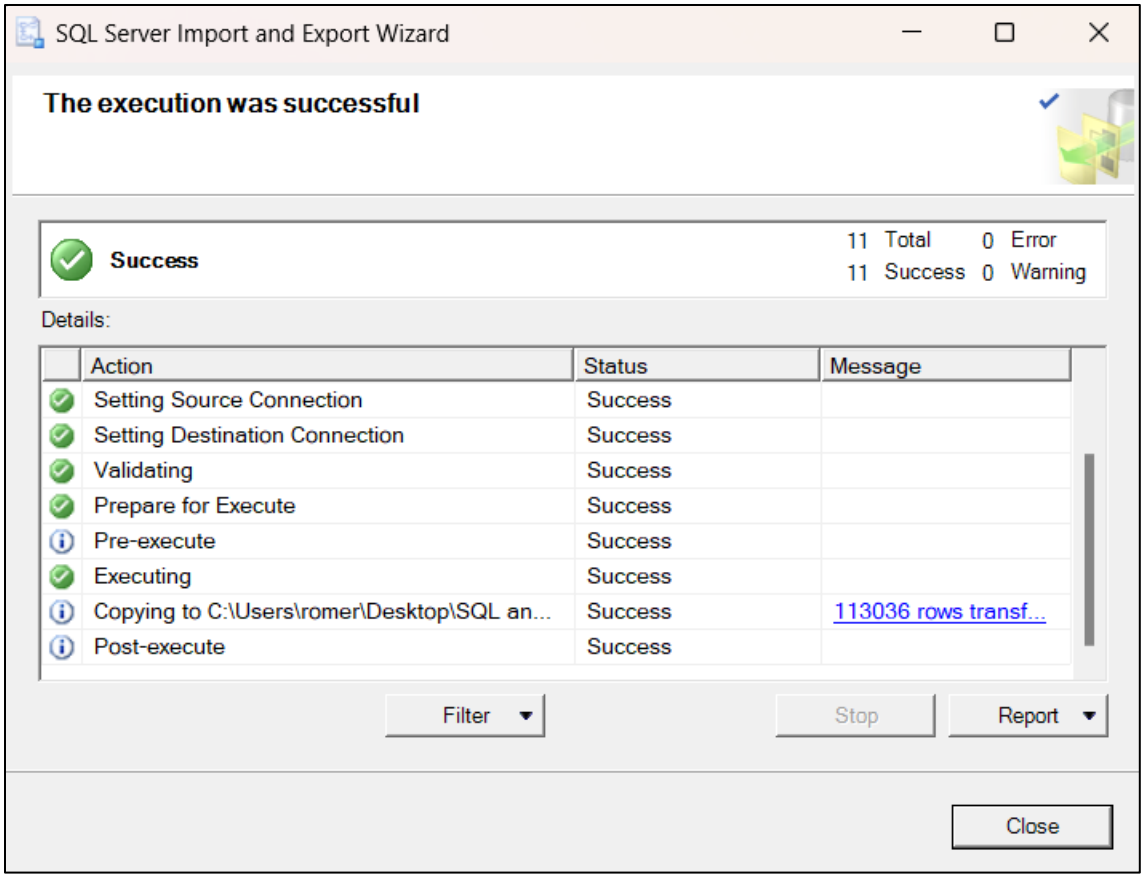


Figure 7: Successful Data Export from SQL

The "Sales_Fixed.csv" file has been successfully exported from the SQL database without any errors, confirming that the data is now clean and ready for further processing. It is now prepared to be imported into Power BI for advanced analysis and visualization.

6. Importing Data into Power BI

In this section, we will import the cleaned data file, "Sales_Fixed.csv," into Power BI to prepare for creating visualizations. Power BI is a powerful analytics tool that allows us to transform raw data into interactive and insightful reports.

To import the cleaned data into Power BI, follow these steps:

1. Open Power BI Desktop: Start by launching the Power BI Desktop application. If it's not already installed, you can download it from the Power BI website.

2. Start a New Project: Click on "File" in the top-left corner and select "New" to create a new Power BI project.

3. Import Data:

- Go to the "Home" tab on the ribbon and click on "Get Data."
- In the "Get Data" window, choose "Text/CSV" and click "Connect."
- Navigate to the location where the "Sales_Fixed.csv" file is saved, select it, and click "Open."

4. Transform Data:

- In the "Navigator" window, click on "Transform Data" instead of "Load." This will open the Power Query Editor.

- In the Power Query Editor, ensure that numeric columns are correctly identified as numbers and that columns intended to be strings are formatted as text. You can adjust column data types by selecting the column header, clicking on the data type icon, and choosing the appropriate type from the dropdown menu.

- After confirming the data types, click "Close & Load" to import the data into your Power BI project.

By following these steps, you will successfully import and prepare the cleaned dataset in Power BI, ensuring that the data types are accurate for creating meaningful visualizations and analyses in the next stages.

Month	Year	Customer_Age	Age_Group	Customer_Gender	Country	State	Product_Category	Sub_Category	Product_Fixed	Order_Quantity_Fixed	Profit_Fixed	Cost_Fixed	Revenue_Fixed
February	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	5	5	5	10
February	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	7	7	7	14
March	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	5	5	5	10
March	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	3	3	3	6
March	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	27	26	27	53
April	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	10	10	10	20
April	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	10	10	10	20
April	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	11	11	11	22
April	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	22	21	22	43
April	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	4	4	4	8
April	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	9	9	9	18
April	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	15	14	15	29
June	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	7	7	7	14
June	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	7	7	7	14
June	2014	45	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	12	12	12	24
January	2014	64	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	9	9	9	18
June	2014	64	Adults (35-64)	M	United States	California	Accessories	Tires and Tubes	Patch Kit/8 Patches	20	19	20	39

Figure 8: Successful Data Import into Power BI

The dataset shown in the figure illustrates the successful import of the "Sales_Fixed.csv" file into Power BI. It highlights the data preview in the Navigator window, which confirms that the file has been imported correctly. The Power Query Editor is visible, demonstrating that the data has been accurately loaded and is ready for further preparation and analysis. This step is essential for ensuring that the data is in the right format for creating effective visualizations and reports in Power BI.

7. Creating Interactive Visualizations in Power BI

In this section, we focus on creating interactive dashboards in Power BI to effectively visualize and analyze the insights derived from the SQL exploratory analysis. The dashboard integrates the following visualizations, each corresponding to the SQL queries performed:

1. Age Distribution

This visualization presents the total number of customers segmented by age groups and gender, as analyzed in SQL. By showcasing the Age_Group distribution, it helps understand customer demographics and tailor marketing strategies. The SQL analysis provided a detailed breakdown of customers by age and gender, which is now visually represented.

2. Yearly Financial Trends

The dashboard visualizes the financial performance over different years, including total cost, profit, and revenue. This visualization reflects the yearly trends identified in the SQL analysis, which highlighted fluctuations and patterns in financial metrics over time. This helps in recognizing periods of growth or decline based on historical data.

3. Product Profitability

This visualization focuses on the profitability of products, grouped by product category and sub-category. It includes:

- Total Profit by Product Category: Summarizes profit across each product category.
- Total Profit by Sub-Category: Breaks down profit within product categories.

- Profit by Product: For simplicity and clarity, the dashboard aggregates data by category and sub-category, omitting individual product details to avoid clutter and ensure effectiveness. The SQL analysis originally included detailed product-level data, but the dashboard prioritizes aggregated insights for clearer visualization.

5.Financial Performance by Country

The dashboard assesses financial performance by geographical regions, focusing on total revenue by country. It reflects the SQL analysis which summarized revenue at the country level. For clarity, state-level details are excluded to avoid overwhelming the dashboard, ensuring a more focused view on regional performance.

The dashboard features interactive slicers that allow users to filter and explore the data dynamically. These visualizations complement the exploratory analysis performed in SQL, providing a comprehensive view of the data. For a detailed exploration of the visualizations and interactive features, refer to the Power BI file available in the repository. This section outlines the creation process, with in-depth analysis and visualizations discussed in the subsequent sections.

Analysis

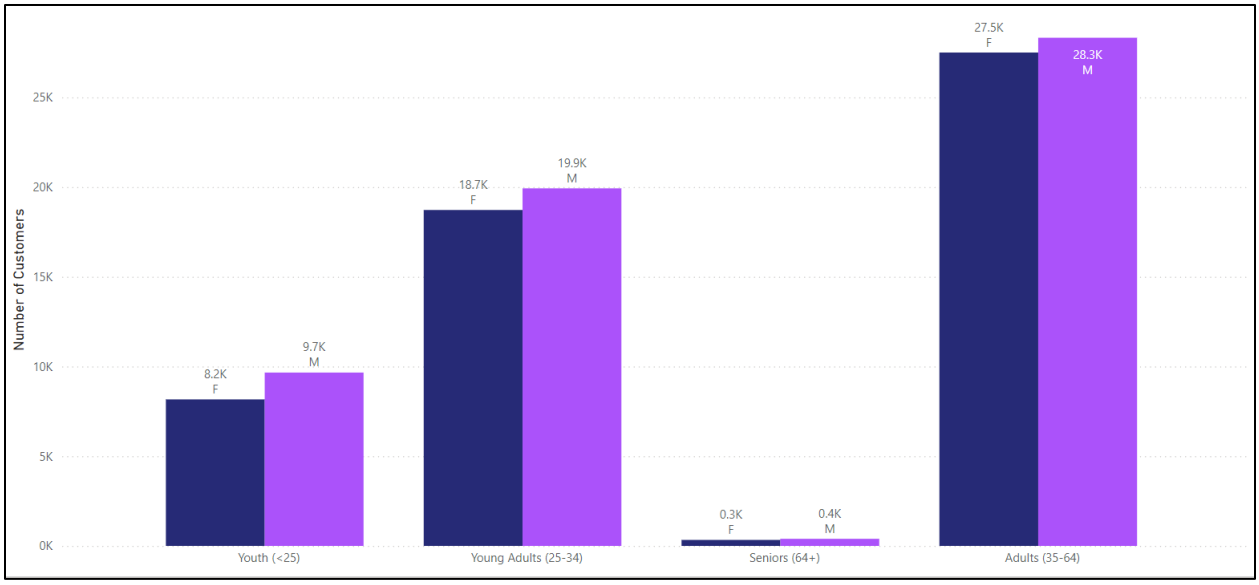
1. Age Distribution

	Age_Group	Female	Male	Total
1	Youth (<25)	8166	9662	17828
2	Young Adults (25-34)	18724	19930	38654
3	Seniors (64+)	334	396	730
4	Adults (35-64)	27500	28324	55824
5	Total	54724	58312	113036

The age distribution of customers provides the following insights:

- **Youth (<25)** make up approximately **15.8%** of the total customer base. This group is relatively small compared to others, which might suggest that younger customers are less prevalent or less engaged with the products offered.
- **Young Adults (25-34)** represent about **34.2%** of the total. This segment is significant, indicating that young adults form a large part of the customer base.
- **Seniors (64+)** account for around **0.6%** of the total customer base. This small percentage shows that older adults are a minor segment, possibly due to different purchasing habits or interests.
- **Adults (35-64)** constitute approximately **49.4%** of the total. This is the largest group, indicating that middle-aged individuals are the predominant consumers of the products.

In summary, the majority of customers fall within the Adults and Young Adults age groups, with a notable decline in the Youth and Seniors segments. This distribution suggests that the marketing efforts should be focused on the larger Adult and Young Adult groups while considering ways to engage the smaller Youth and Senior segments. Gender distribution is relatively balanced, with **females** representing **48.4%** and **males 51.6%** of the total customer base.



2. Yearly Financial Trends

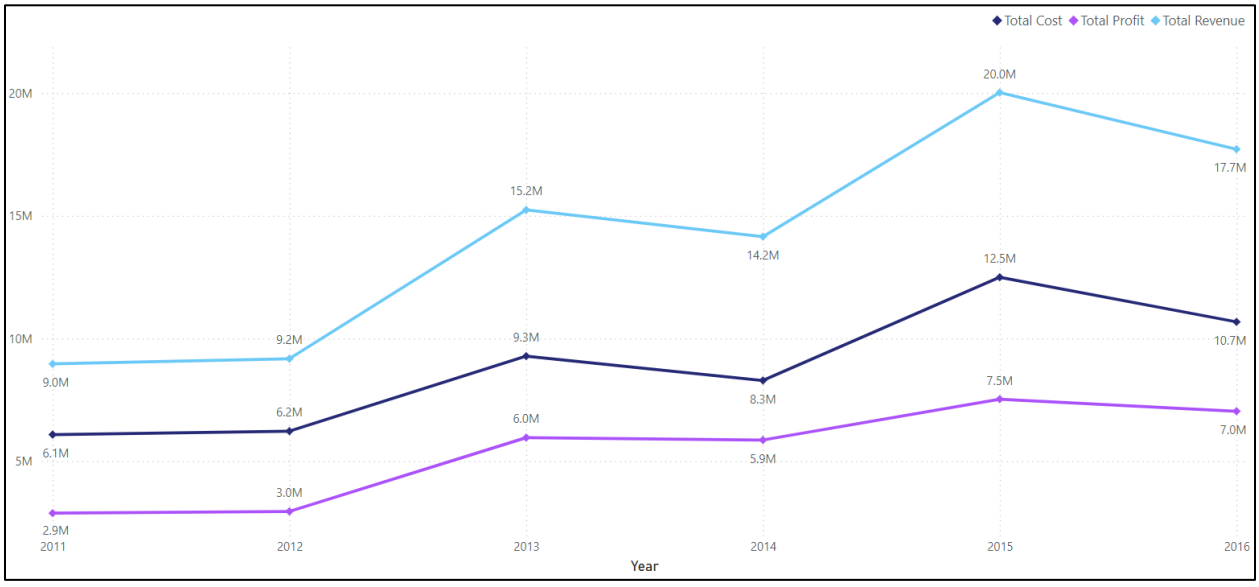
	Year	Total_Cost	Total_Profit	Total_Revenue
1	2011	6083587	2881301	8964888
2	2012	6223990	2951993	9175983
3	2013	9280829	5959208	15240037
4	2014	8288637	5864087	14152724
5	2015	12495428	7528563	20023991
6	2016	10677437	7035948	17713385
7	Total	53049908	32221100	85271008

The following percentages represent each year's contribution to the total financial metrics across the years 2011 to 2016.

Year	Cost Percentage	Profit Percentage	Revenue Percentage
2011	11.5%	8.9%	10.5%
2012	11.7%	9.2%	10.8%
2013	17.5%	18.5%	17.9%
2014	15.6%	18.2%	16.6%
2015	23.6%	23.4%	23.5%
2016	20.2%	21.8%	20.8%

- **Significant Increase in 2013:** The year 2013 stands out with the highest values in cost, profit, and revenue, representing 17.5%, 18.5%, and 17.9% of the total respectively. This year likely saw exceptional performance or expansion.
- **2015 Highs:** The year 2015 recorded the highest total revenue and total profit, contributing 23.5% and 23.4% to the overall totals. This indicates strong performance or a major growth period.
- **Consistency in Costs:** Costs remained relatively stable from 2011 to 2014 but saw a significant increase in 2015, reflecting possibly higher investments or operational costs.
- **Profit Margins:** Despite fluctuations in total revenue, profit margins were highest in 2013 and 2015, showing that these years were not only high-grossing but also highly profitable.
- **Revenue Trends:** The revenue trend shows a peak in **2015,** aligning with the peak in profits. This suggests a strong correlation between revenue and profitability in that year.

These insights provide a comprehensive view of financial trends and performance, highlighting key years of growth and areas of interest for further analysis.



3. Product Profitability

	Product_Category	Total_Profit
1	Accessories	8862377
2	Bikes	20519276
3	Clothing	2839447
4	Total	32221100

The total profit is $32,221,100. Of this, the Bikes category contributes the largest share at 63.7%, indicating its primary role in overall profitability. Accessories follow with a significant contribution of 27.5%, highlighting its importance as a profitable category. Clothing, on the other hand, contributes a smaller portion of 8.8%. These percentages reveal that the majority of the profit is concentrated in the Bikes category, suggesting a strong performance in this area and potential opportunities for strategic focus or investment.

- **Dominance of Bikes:** The Bikes category is the most profitable, contributing 63.7% to the total profit. This indicates that Bikes are the key driver of profitability in the dataset.
- **Accessories' Contribution:** Accessories follow with a substantial contribution of 27.5% to the total profit, suggesting that this category also plays a significant role in overall profitability.
- **Clothing's Share:** The Clothing category contributes 8.8% to the total profit, which is relatively small compared to Bikes and Accessories. This might indicate lower margins or lesser sales volume in this category.

	Product_Category	Sub_Category	Total_Profit
1	Accessories	Bike Racks	304455
2	Accessories	Bike Stands	201935
3	Accessories	Bottles and Cages	810598
4	Accessories	Cleaners	115891
5	Accessories	Fenders	748914
6	Accessories	Helmets	3380203
7	Accessories	Hydration Packs	572668
8	Accessories	Tires and Tubes	2727713
9	Bikes	Mountain Bikes	8160463
10	Bikes	Road Bikes	10078875
11	Bikes	Touring Bikes	2279938
12	Clothing	Caps	77921
13	Clothing	Gloves	511654
14	Clothing	Jerseys	603568
15	Clothing	Shorts	1034342
16	Clothing	Socks	93171
17	Clothing	Vests	518791

In the **Accessories category**, the total profit distribution is as follows:

- Helmets lead with $3,380,203.
- Tires and Tubes follow with $2,727,713.
- Bottles and Cages contribute $810,598.
- Other sub-categories like Bike Racks, Bike Stands, Cleaners, Fenders, and Hydration Packs have smaller contributions.

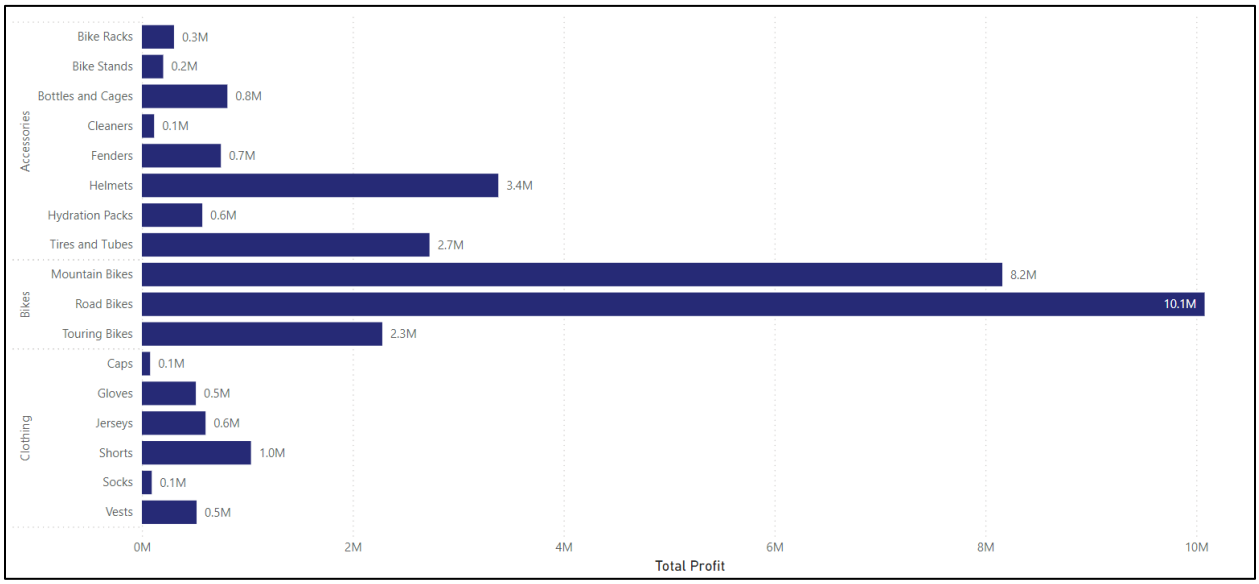
In the **Bikes category**:

- Road Bikes are the most profitable with $10,078,875.
- Mountain Bikes follow with $8,160,463.
- Touring Bikes contribute $2,279,938.

For the **Clothing category:**

- Shorts generate the highest profit of $1,034,342.
- Jerseys follow with $603,568.
- Gloves contribute $511,654.
- Caps, Socks, and Vests have smaller contributions.

This analysis highlights that Helmets and Road Bikes are the top profit generators, indicating their significant impact on the overall profitability.



4. Monthly Sales Trends

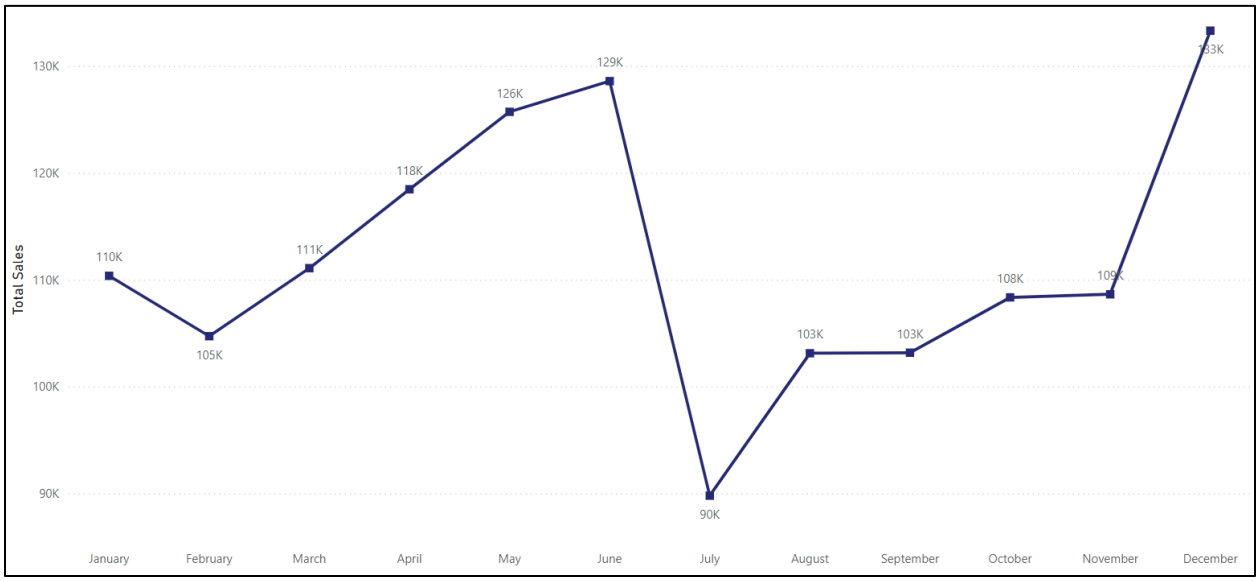
	Month	Total_Sales
1	January	110367
2	February	104717
3	March	111085
4	April	118467
5	May	125715
6	June	128591
7	July	89787
8	August	103119
9	September	103171
10	October	108348
11	November	108637
12	December	133312
13	Total	1345316

In the monthly sales data, the total sales distribution is as follows:

- **December**, with $133,312, accounts for 9.9% of the annual sales.
- **June** follows with $128,591, representing 9.6%.
- **May** contributes $125,715, which is 9.3%.
- **April** and March account for $118,467 (8.8%) and $111,085 (8.2%), respectively.

- **January and October** have sales of $110,367 (8.2%) and $108,348 (8.1%).
- **November's** sales total $108,637, which is 8.1%.
- **September and August** show sales of $103,171 (7.7%) and $103,119 (7.7%).
- **February** has $104,717, representing 7.8%.
- **July**, with $89,787, contributes 6.7%.

The total sales for the year amount to $1,345,316. This data reveals a notable peak in sales during December, with a steady performance throughout the middle and end of the year, while sales in July are comparatively lower.



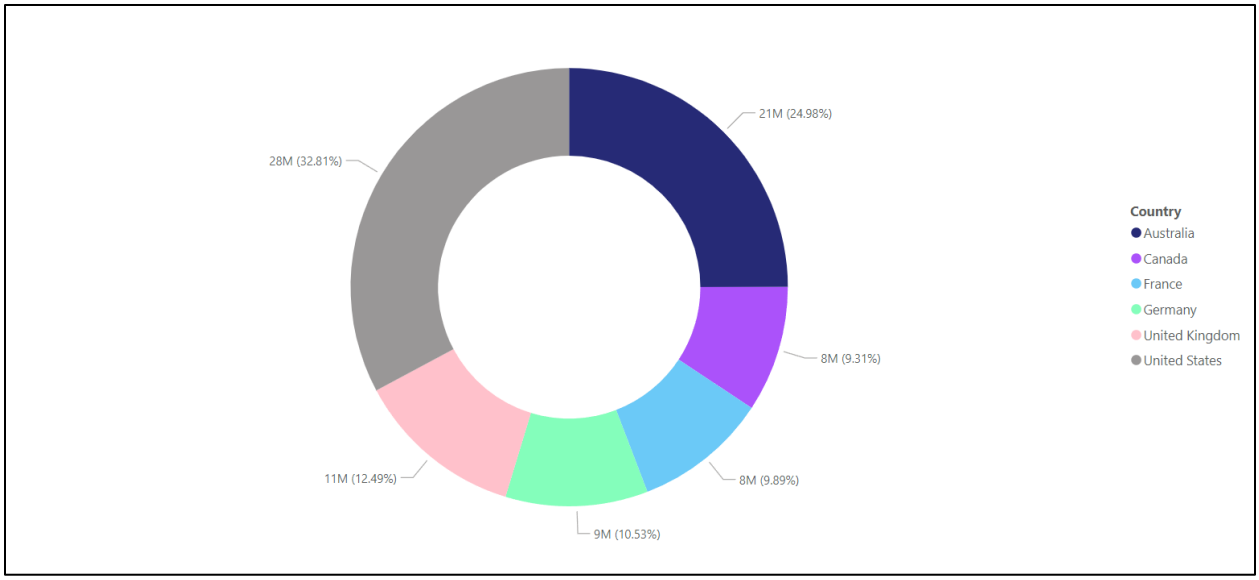
5. Financial Performance by Country

	Country	Total_Revenue
1	Australia	21302059
2	Canada	7935738
3	France	8432872
4	Germany	8978596
5	United Kingdom	10646196
6	United States	27975547
7	Total	85271008

The revenue distribution by country reveals several key insights:

- **United States:** With $27,975,547 in total revenue, the United States is the largest revenue contributor, accounting for 32.8% of the overall revenue. This substantial figure highlights the importance of the U.S. market in the business's revenue generation.
- **Australia:** Australia follows with $21,302,059, representing 25.0% of the total revenue. The strong performance in Australia indicates a significant market presence and potentially robust sales in the region.
- **United Kingdom:** The United Kingdom contributes $10,646,196, which is 12.5% of the total revenue. This reflects a considerable market share and suggests a strong foothold in the U.K. market.
- **Germany:** Germany's total revenue stands at $8,978,596, making up 10.5% of the total. This indicates a notable presence in the German market, contributing a significant portion to the overall revenue.
- **France:** France generates $8,432,872, or 9.9% of the total revenue. This shows a solid revenue stream from the French market, although it is slightly less compared to Germany.
- **Canada:** Canada contributes $7,935,738, representing 9.3% of the total revenue. While it is the smallest among the listed countries, it still plays an important role in the overall revenue picture.

In summary, the United States stands out as the dominant revenue source, with Australia and the United Kingdom also making substantial contributions. Other countries like Germany, France, and Canada contribute smaller, yet significant, portions of the total revenue, emphasizing the diverse geographic spread of the business's revenue sources.


21M (24.98%)
8M (9.31%)
8M (9.89%)
9M (10.53%)
11M (12.49%)
28M (32.81%)
Country
Australia
Canada
France
Germany
United Kingdom
United States

Conclusions

1. Key Insights from the Bicycle Sales Data Analysis

- **Age Distribution:** The customer base is predominantly composed of adults aged 35-64 and young adults aged 25-34, who together account for the majority of sales. This suggests that marketing efforts should primarily target these groups, while considering strategies to engage the smaller youth and senior segments.
- **Yearly Financial Trends:** The financial data from 2011 to 2016 shows that 2015 was a peak year for both revenue and profit, reflecting a period of significant growth. The year 2013 also stood out for its strong financial performance, indicating a possible expansion or successful campaign during that time.
- **Product Profitability:** The Bikes category is the most profitable, contributing nearly two-thirds of the total profit, with Road Bikes leading this segment. Accessories also play a crucial role in profitability, particularly helmets and tires. Clothing, while contributing less overall, still represents a notable portion of profits, with shorts being the most profitable item.
- **Monthly Sales Trends:** December emerged as the top month for sales, likely due to holiday demand, while the summer months, particularly July, saw a decline in sales. This suggests a seasonal pattern in consumer purchasing behavior that could inform inventory and marketing strategies.
- **Financial Performance by Country:** The United States is the largest revenue generator, followed by Australia and the United Kingdom. The diverse geographic distribution of revenue highlights the importance of these key markets, with substantial contributions also coming from Germany, France, and Canada.

2. Reflections on the Project

- This project has effectively demonstrated the integration of SQL and Power BI to analyze and visualize bicycle sales data. The development of a well-structured SQL database allowed for efficient data storage, querying, and manipulation, ensuring that the data was ready for detailed analysis.
- Through the use of Power BI, interactive visualizations were created, highlighting key trends and insights within the sales data. These visualizations not only made the data more accessible but also provided valuable information that could inform future business decisions.
- The comprehensive report compiled during this project showcases the ability to manage data, conduct in-depth analysis, and present findings in a clear and impactful way. This work emphasizes the value of combining technical expertise with data-driven insights to achieve meaningful outcomes.

Appendices

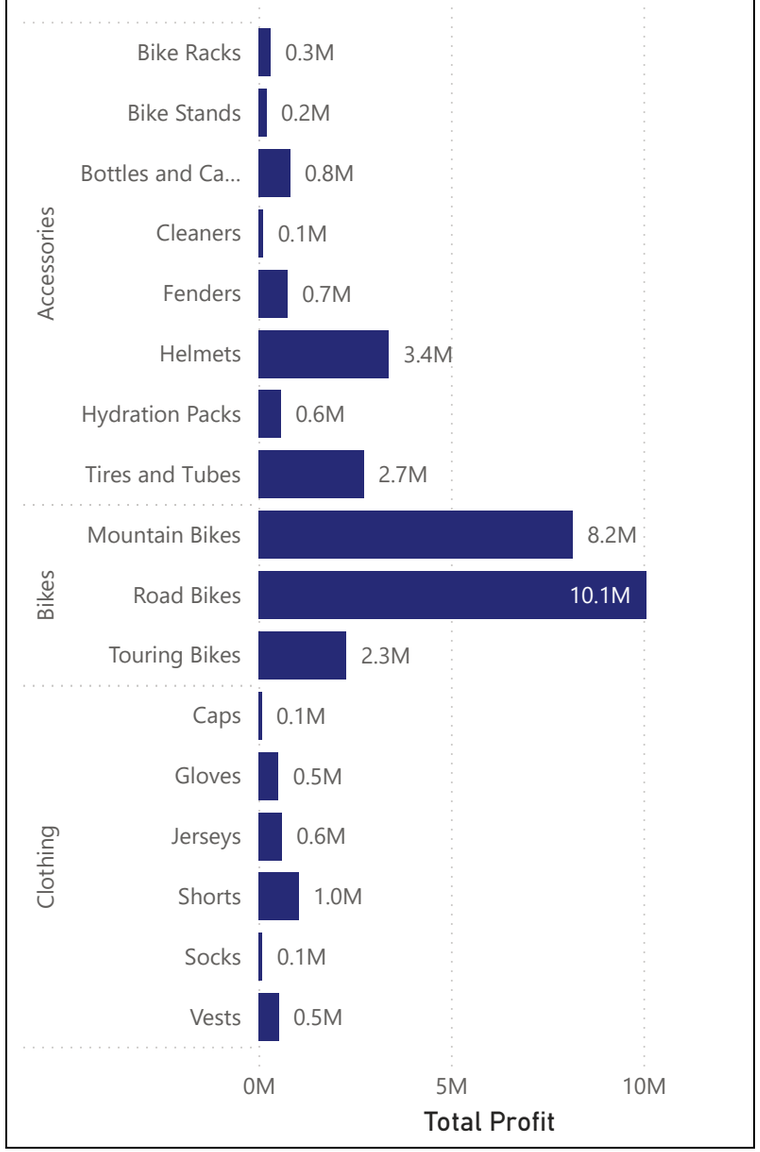
- **Product Category:** Evaluating the performance of the best-performing product category, Bikes, against the lowest-performing category, Clothing, provides the following insights: Bikes, with total sales of 36K, incurred a total cost of $41,263K and generated $61,782K in total revenue, resulting in a substantial profit of $20,519K. This highlights the significant profitability and strong market demand for Bikes. In contrast, Clothing, with total sales of 255K, had a total cost of $5,531K and generated $8,371K in total revenue, leading to a much lower profit of $2,839K. This comparison underscores the difference in performance and profitability between these two categories.
- **Month:** Comparing the performance of December, the highest-performing month, with July, the lowest-performing month, reveals the following insights: December achieved total sales of $133K, with total costs amounting to $5,677K and total revenue of $9,087K, resulting in a notable profit of $3,410K. This indicates December's strong performance and profitability. Conversely, July, despite having total sales of $90K, incurred total costs of $3,582K and generated $5,721K in total revenue, leading to a significantly lower profit of $2,140K. This comparison highlights the considerable disparity in performance and profitability between these two months.
- **Year:** Evaluating the performance between the year 2011, the initial year, and 2016, the most recent year with data up to July, provides the following insights: In 2011, total sales amounted to $5K, with total costs of $6,084K and total revenue of $8,965K, resulting in a profit of $2,881K. In contrast, by 2016, total sales had risen to $371K, with total costs increasing to $10,677K and total revenue reaching $17,713K, yielding a significantly higher profit of $7,036K. This comparison highlights a substantial growth in sales, costs, revenue, and profit over the years, demonstrating improved performance and profitability by 2016.

- **Age Group:** Comparing the performance between the age group Adults (35-64), which represents the highest performance, and Seniors (64+), which is the lowest, yields the following insights: For the Adults group, total sales amount to $672K, with total costs of $26,263K and total revenue of $42,584K, resulting in a substantial profit of $16,322K. In contrast, the Seniors group shows significantly lower figures with total sales of $10K, total costs of $170K, total revenue of $308K, and a profit of just $138K. This comparison underscores the dominant performance of the Adults group in generating higher sales and profit, while the Seniors group contributes minimally in terms of sales, revenue, and profit.
- **Country**: When comparing the performance between Canada, which has the lowest performance, and the United States, which has the highest performance, the following insights emerge: For Canada, total sales amount to $192K, with total costs of $4,218K and total revenue of $7,936K, resulting in a profit of $3,717K. On the other hand, the United States exhibits significantly higher figures, with total sales of $478K, total costs of $16,902K, and total revenue of $27,976K, leading to a substantial profit of $11,074K. This comparison highlights the substantial disparity in performance, with the United States generating significantly higher sales, revenue, and profit compared to Canada.

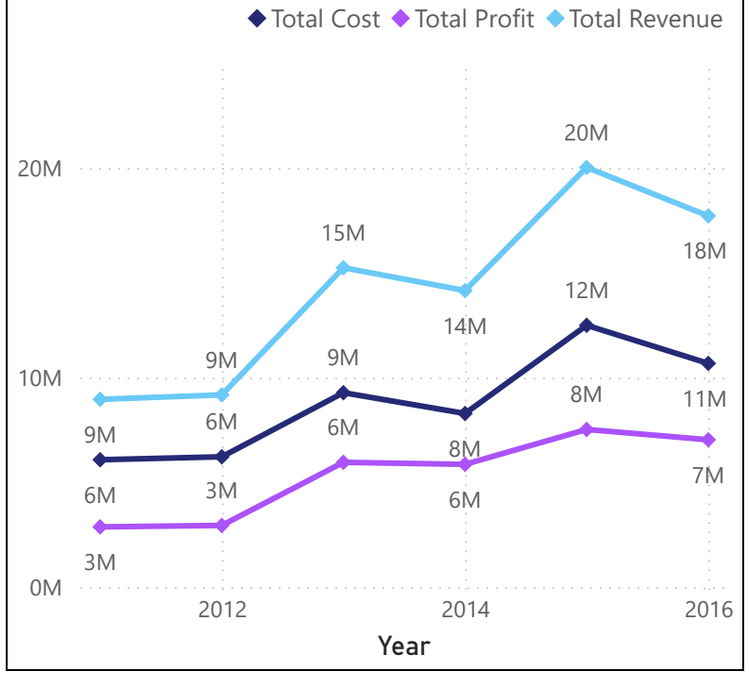
Bike Sales Dashboard

Product: All | Category: All | Month: All | Year: All | Age Group: All | Gender: All | Country: All

Product Profitability

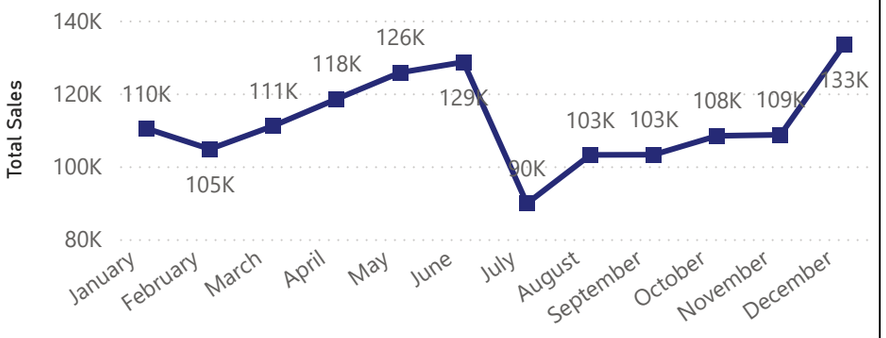


Yearly Financial Trends

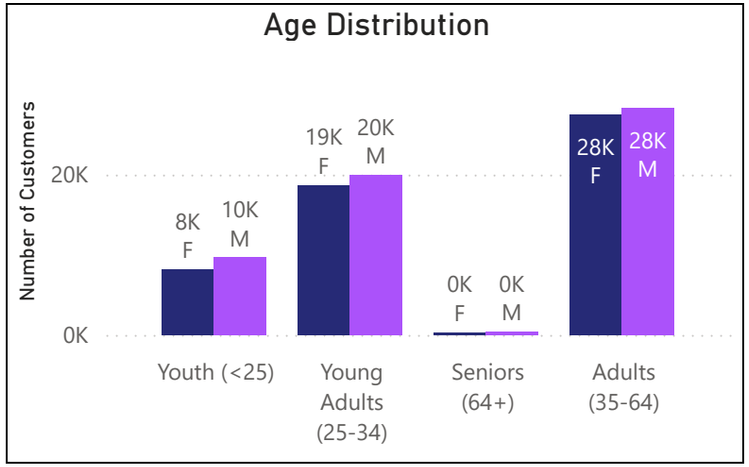


1345K	53050K
Total Sales	Total Cost
85271K	**32221K**
Total Revenue	Total Profit

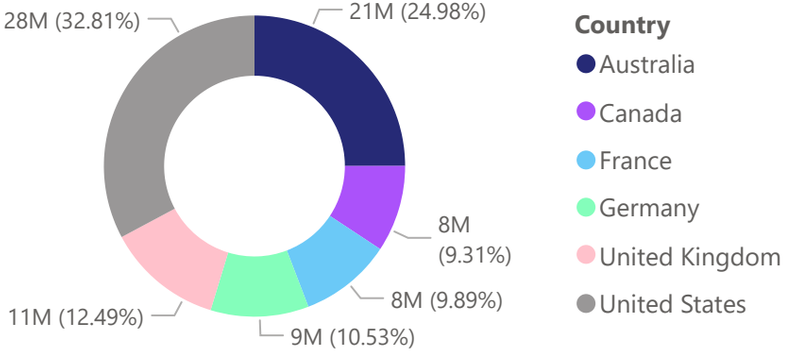
Monthly Sales Trends



Age Distribution



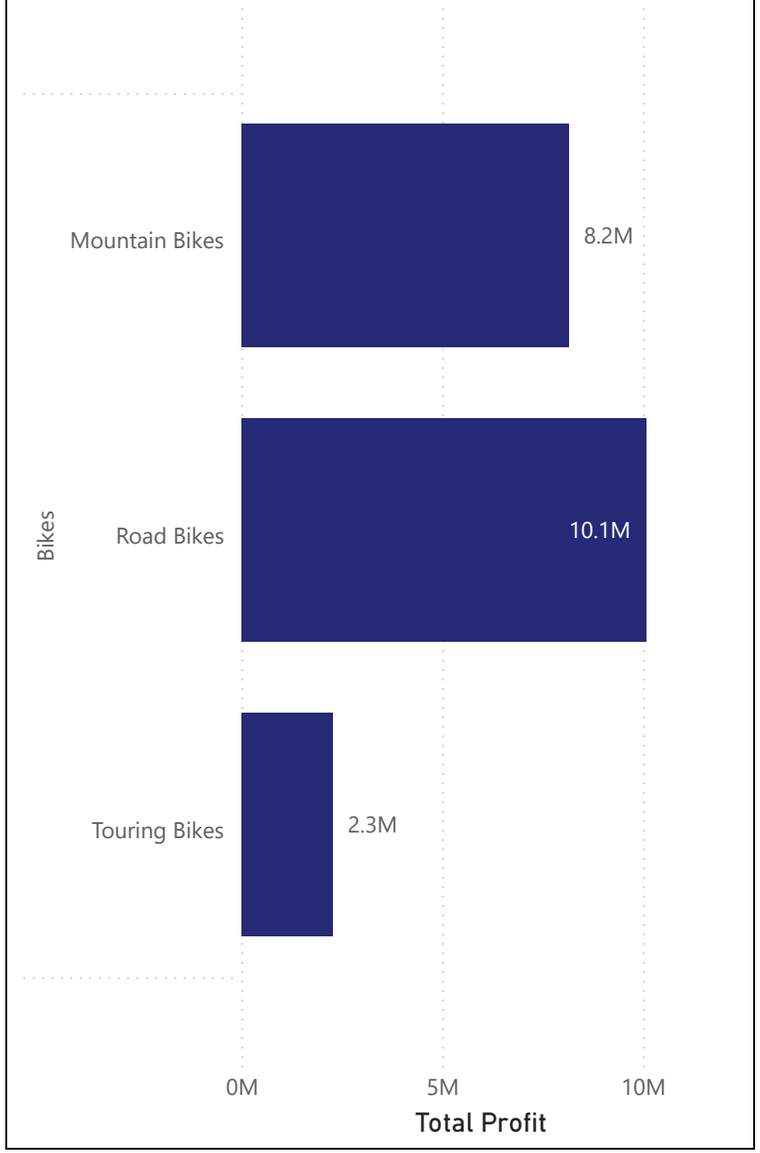
Financial Performance by Country



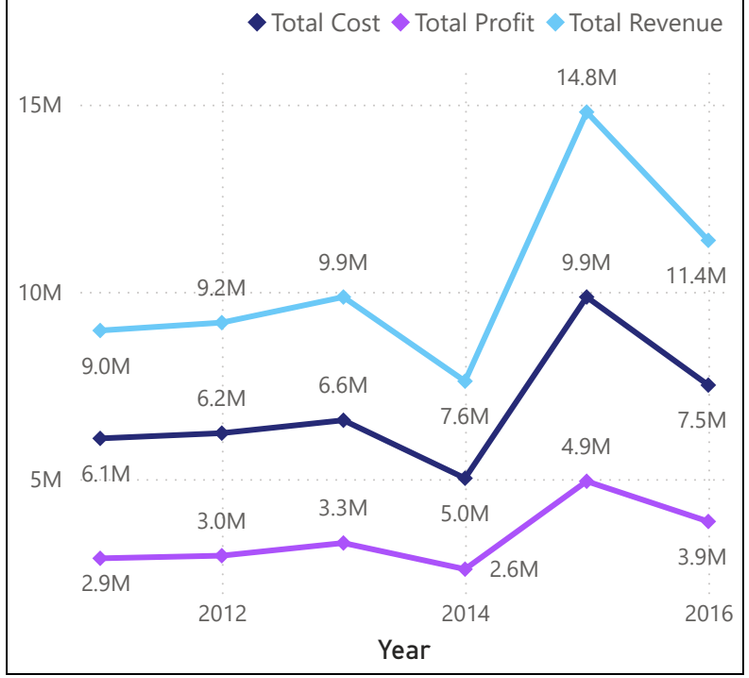
Bike Sales Dashboard

Product	Category	Month	Year	Age Group	Gender	Country
All	Bikes	All	All	All	All	All

Product Profitability

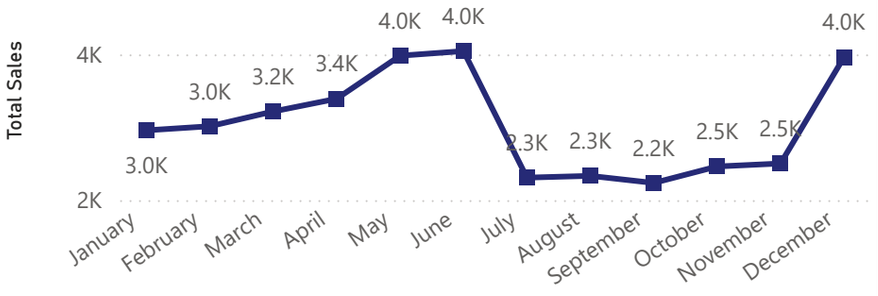


Yearly Financial Trends

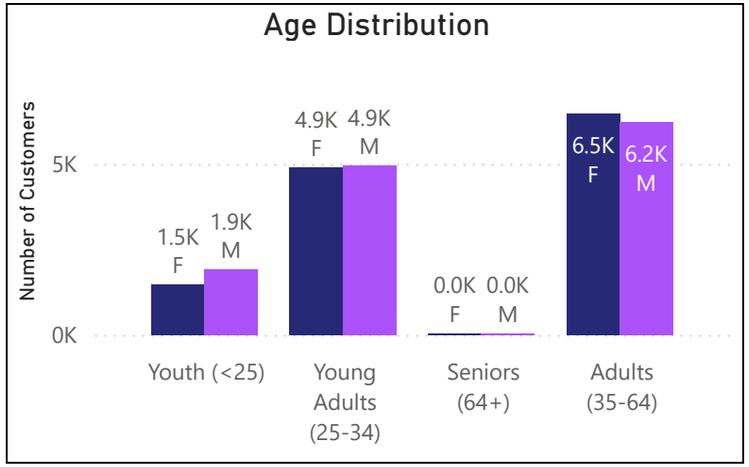


36K	41263K
Total Sales	Total Cost
61782K	**20519K**
Total Revenue	Total Profit

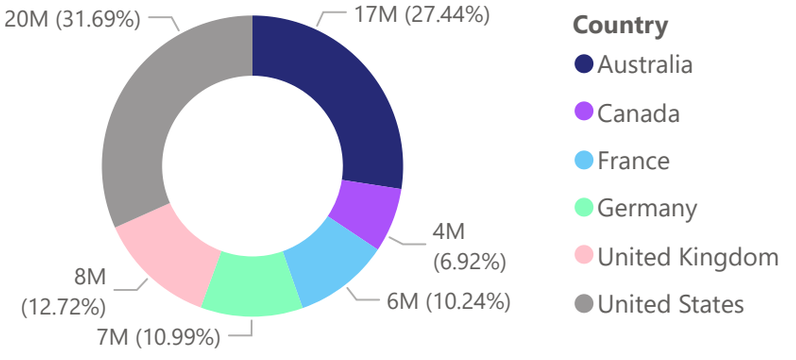
Monthly Sales Trends



Age Distribution



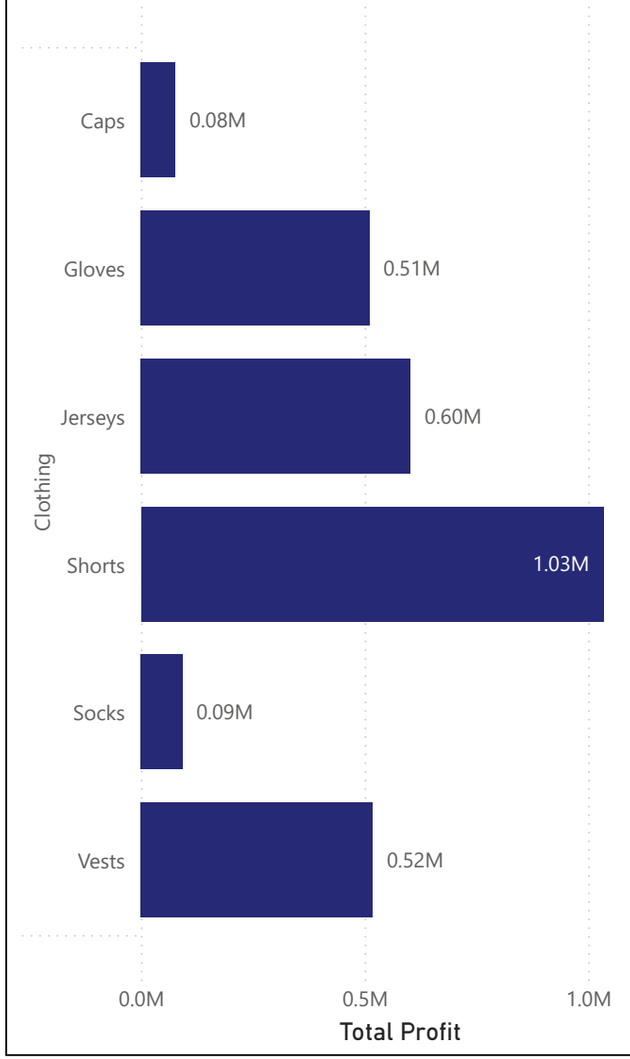
Financial Performance by Country



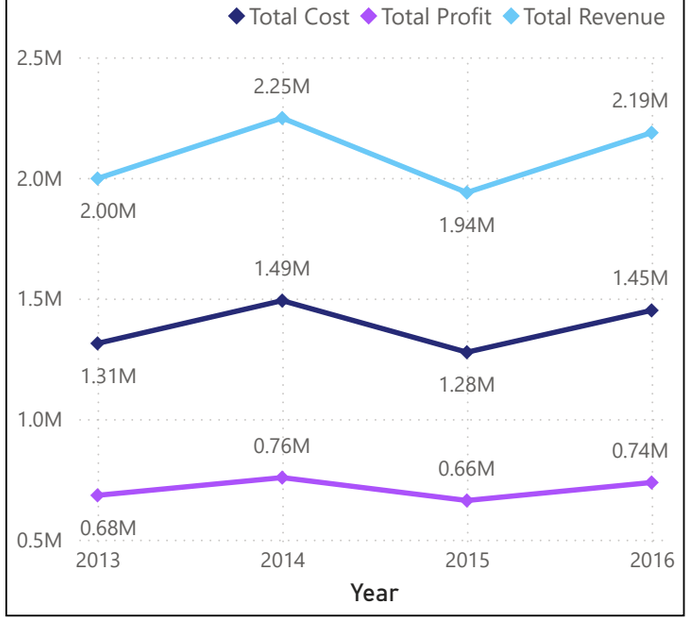
Bike Sales Dashboard

Product	Category	Month	Year	Age Group	Gender	Country
All	Clothing	All	All	All	All	All

Product Profitability

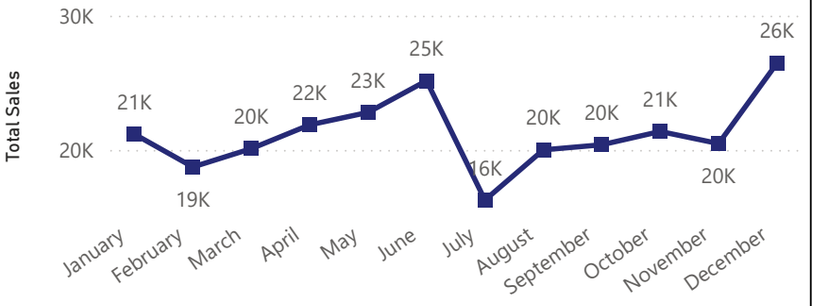


Yearly Financial Trends

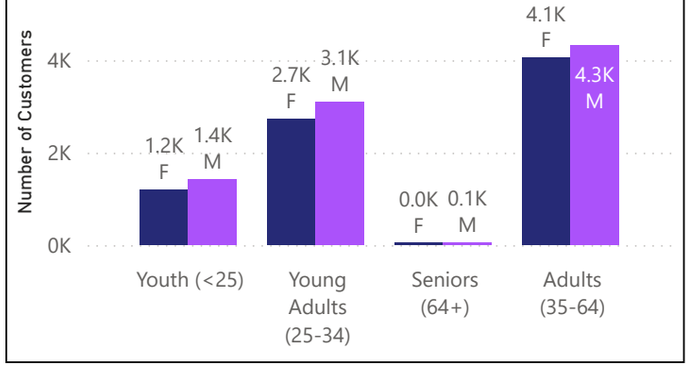


255K	5531K
Total Sales	Total Cost
8371K	**2839K**
Total Revenue	Total Profit

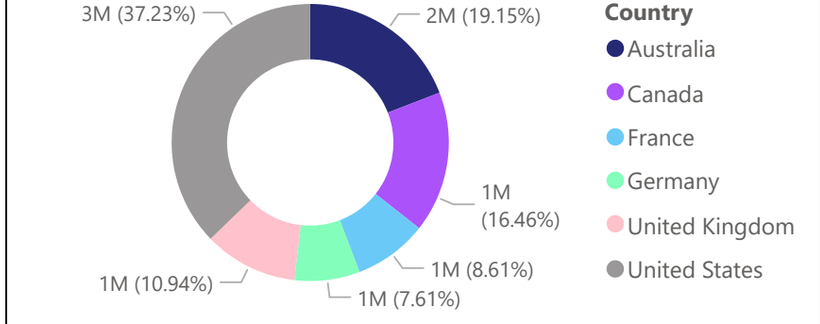
Monthly Sales Trends



Age Distribution



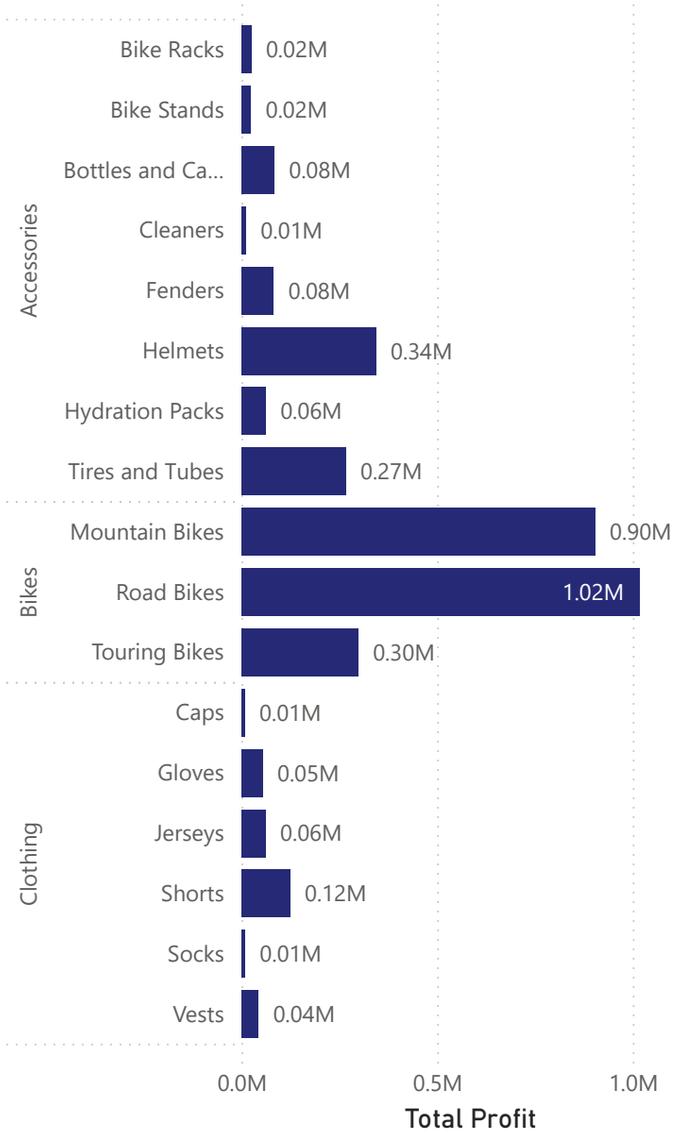
Financial Performance by Country



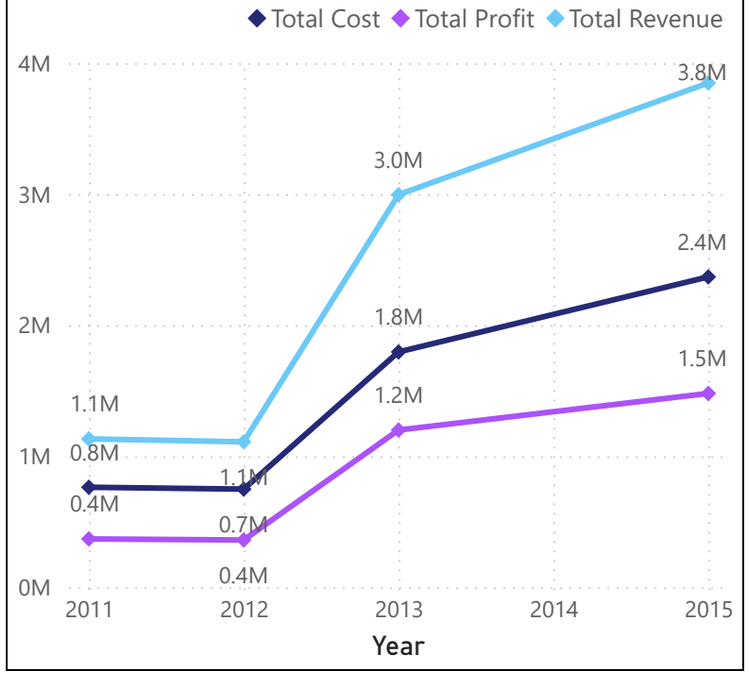
Bike Sales Dashboard

Product	Category	Month	Year	Age Group	Gender	Country
All	All	December	All	All	All	All

Product Profitability

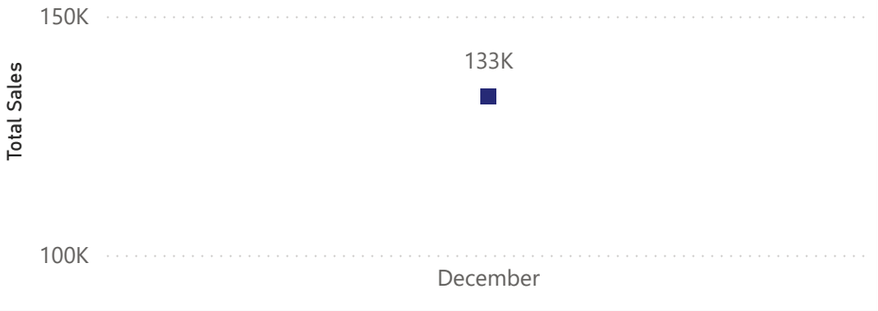


Yearly Financial Trends

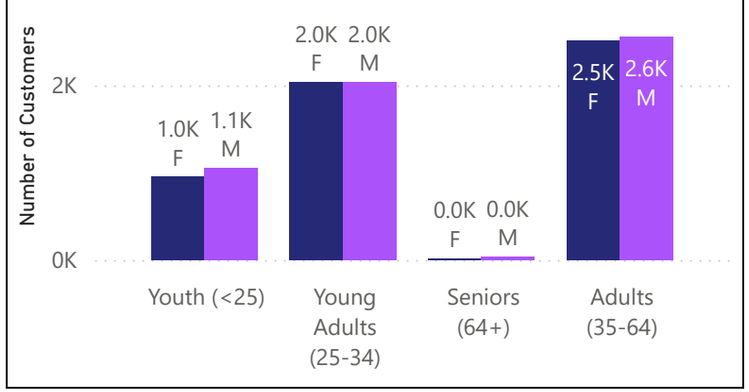


133K	5677K
Total Sales	Total Cost
9087K	**3410K**
Total Revenue	Total Profit

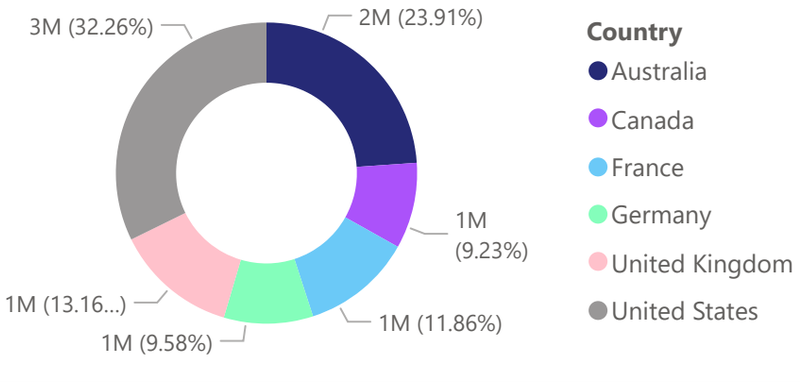
Monthly Sales Trends



Age Distribution



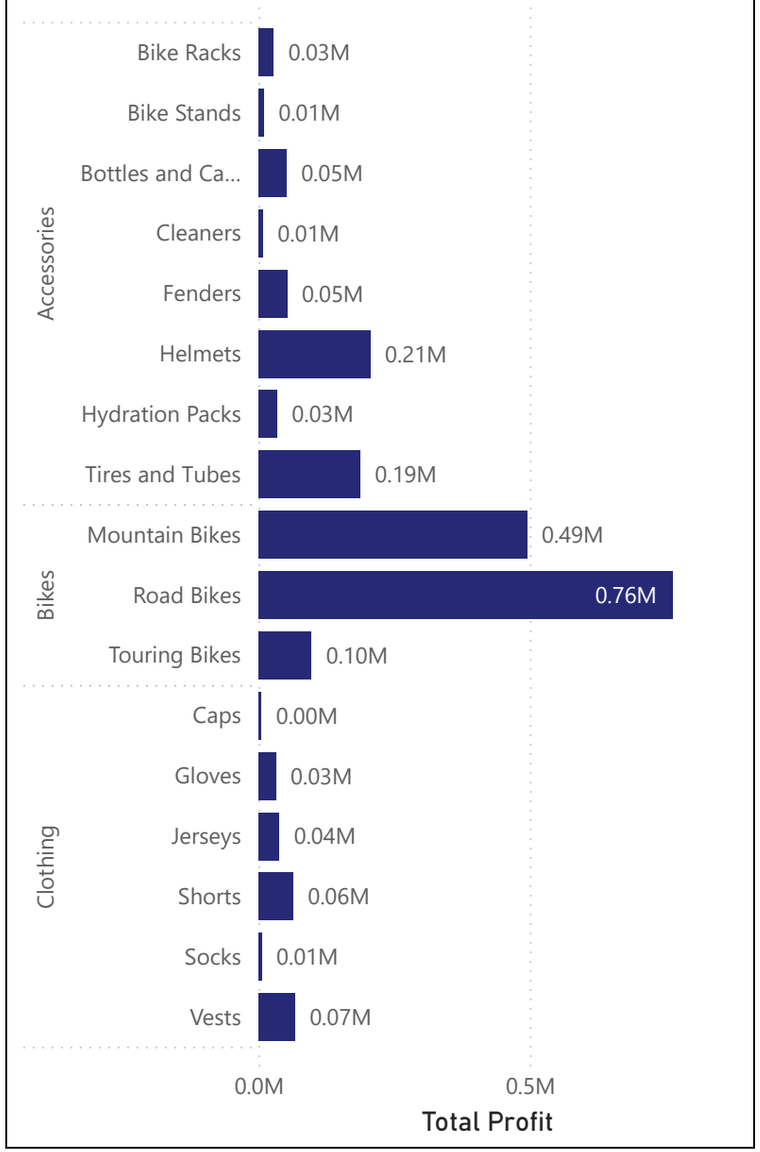
Financial Performance by Country



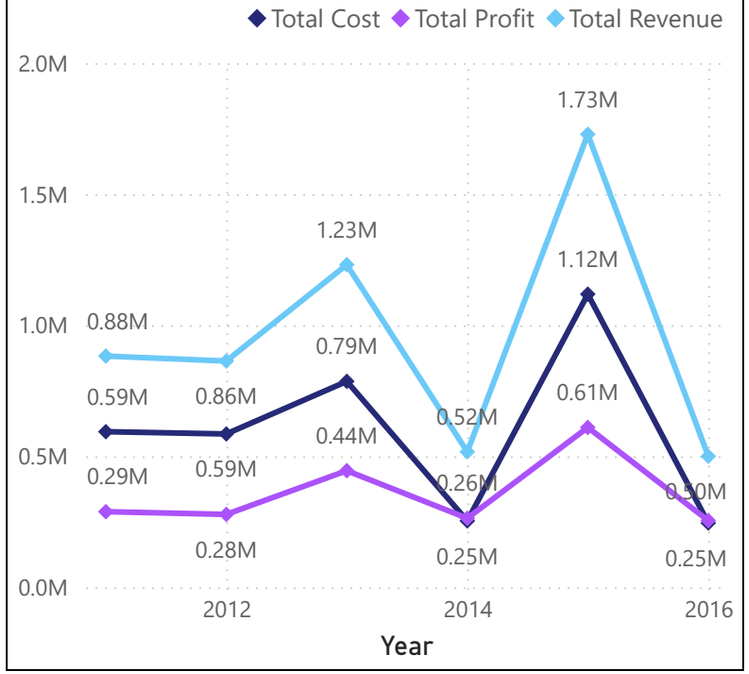
Bike Sales Dashboard

Product	Category	Month	Year	Age Group	Gender	Country
All	All	July	All	All	All	All

Product Profitability

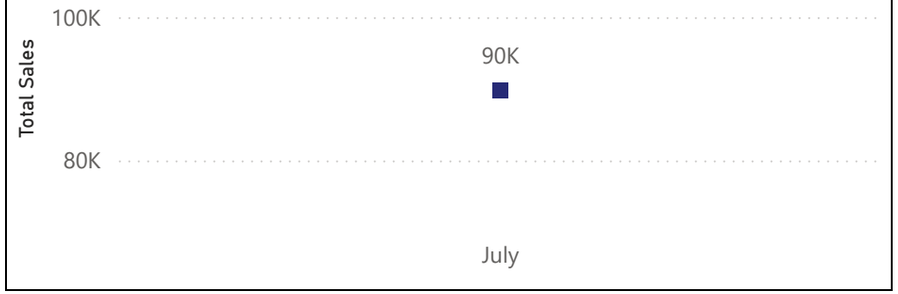


Yearly Financial Trends

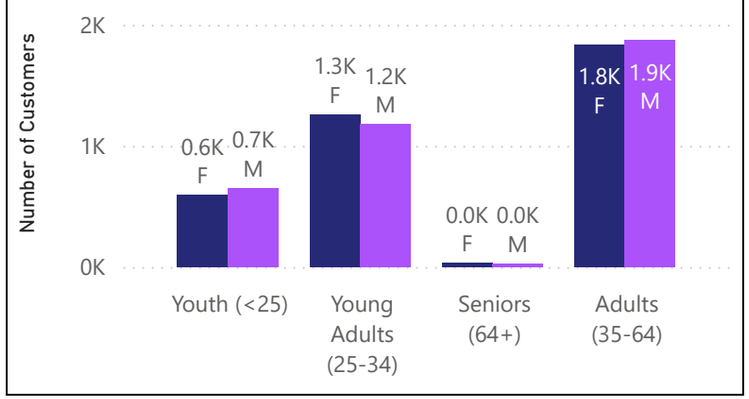


90K	3582K
Total Sales	Total Cost
5721K	**2140K**
Total Revenue	Total Profit

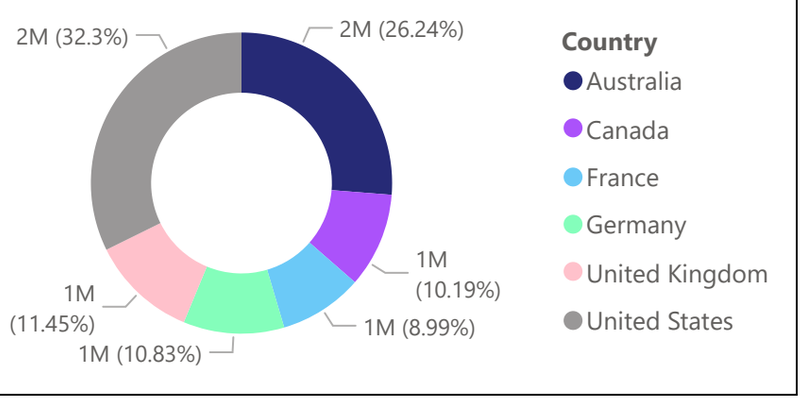
Monthly Sales Trends



Age Distribution



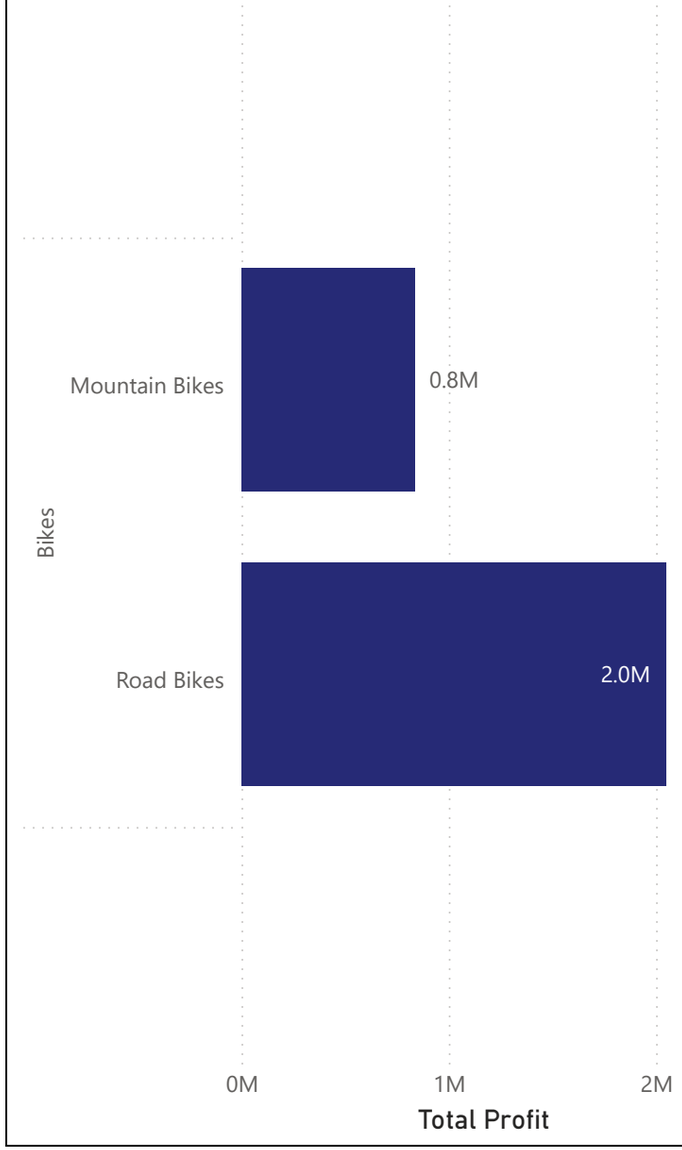
Financial Performance by Country



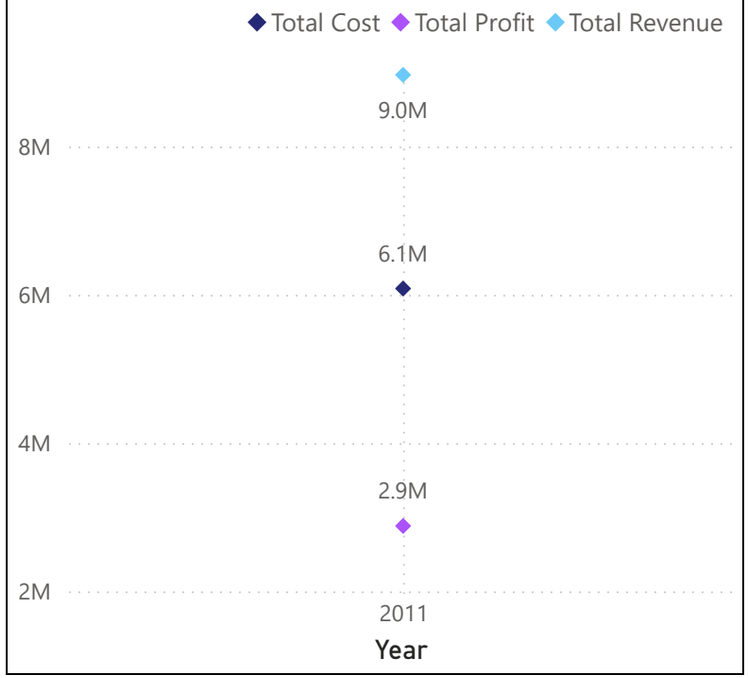
Bike Sales Dashboard

Product: All | Category: All | Month: All | Year: 2011 | Age Group: All | Gender: All | Country: All

Product Profitability

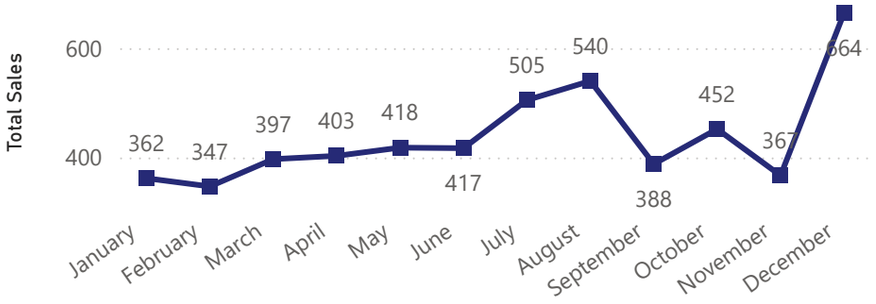


Yearly Financial Trends

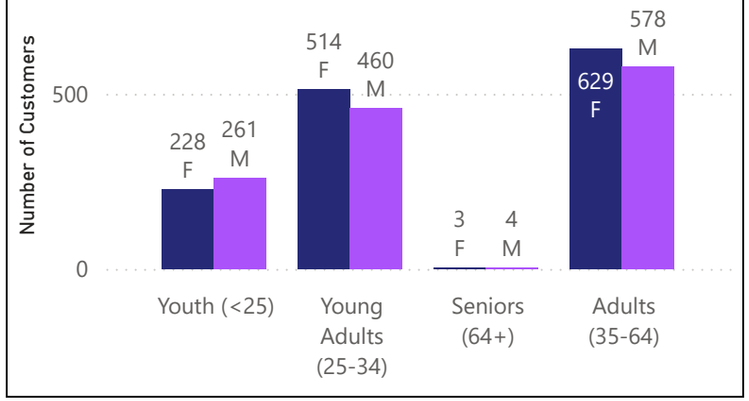


5K	6084K
Total Sales	Total Cost
8965K	**2881K**
Total Revenue	Total Profit

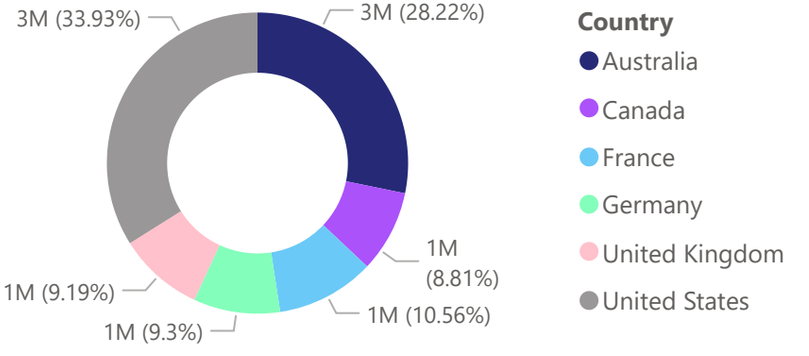
Monthly Sales Trends



Age Distribution



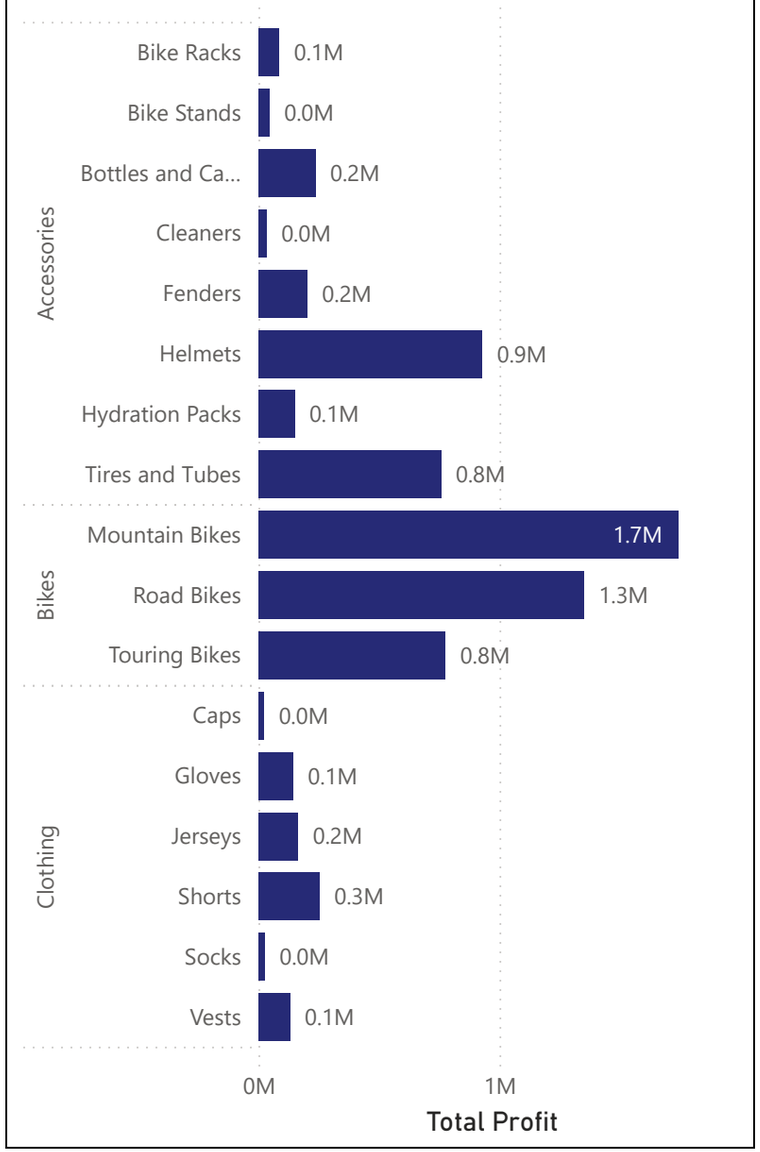
Financial Performance by Country



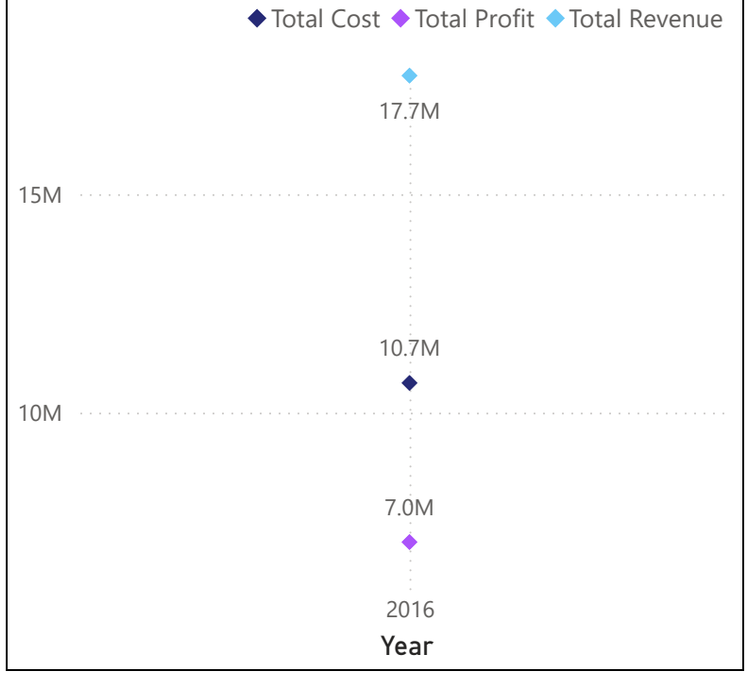
Bike Sales Dashboard

Product	Category	Month	Year	Age Group	Gender	Country
All	All	All	2016	All	All	All

Product Profitability

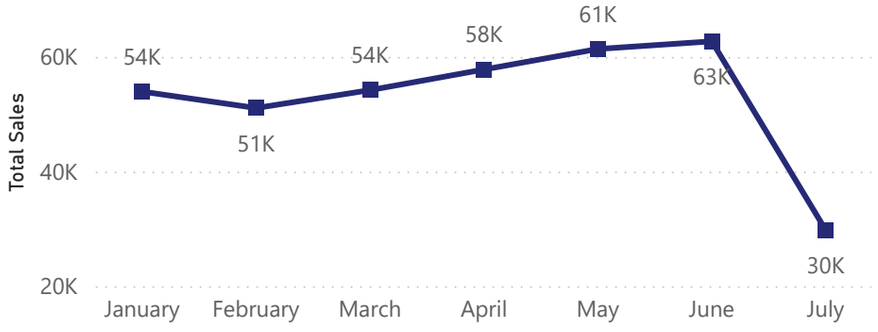


Yearly Financial Trends

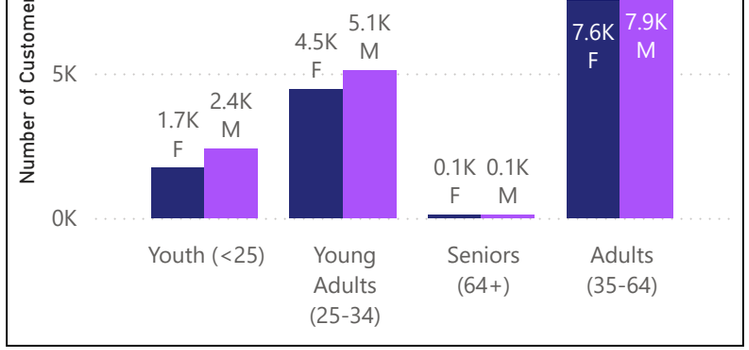


371K	10677K
Total Sales	Total Cost
17713K	**7036K**
Total Revenue	Total Profit

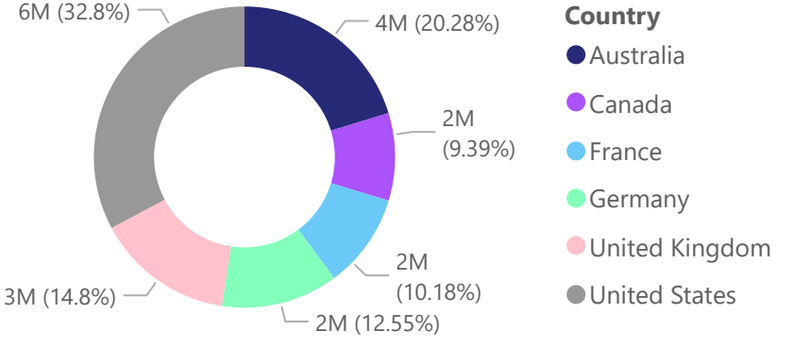
Monthly Sales Trends



Age Distribution



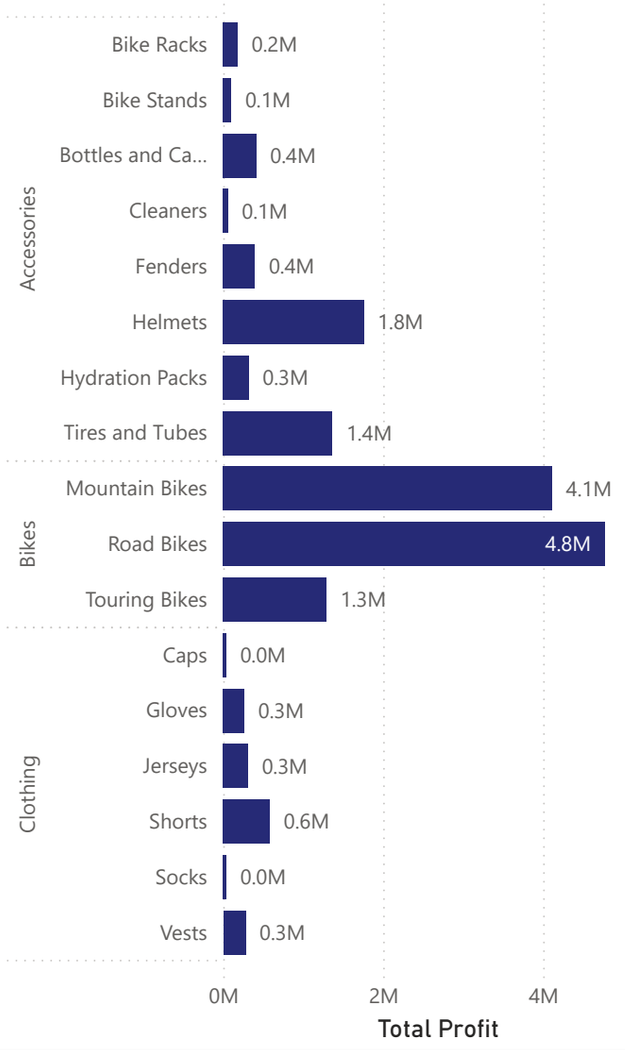
Financial Performance by Country



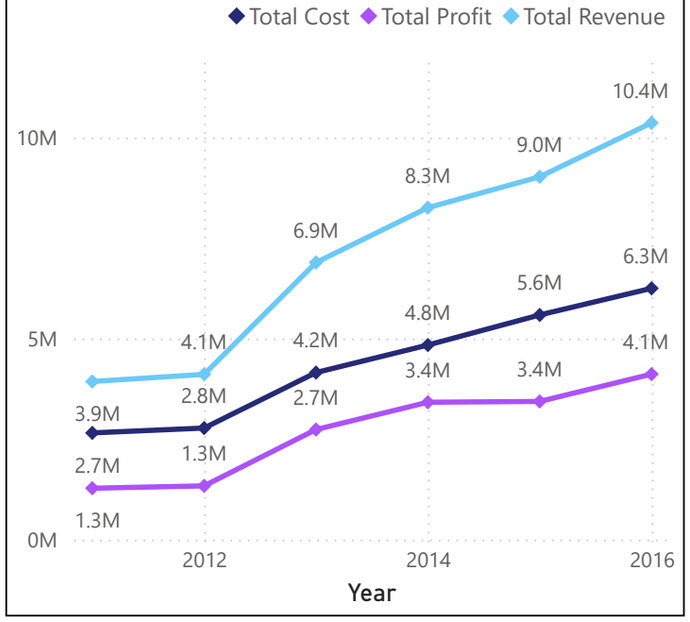
Bike Sales Dashboard

Product: All | Category: All | Month: All | Year: All | Age Group: Adults (35-6... | Gender: All | Country: All

Product Profitability

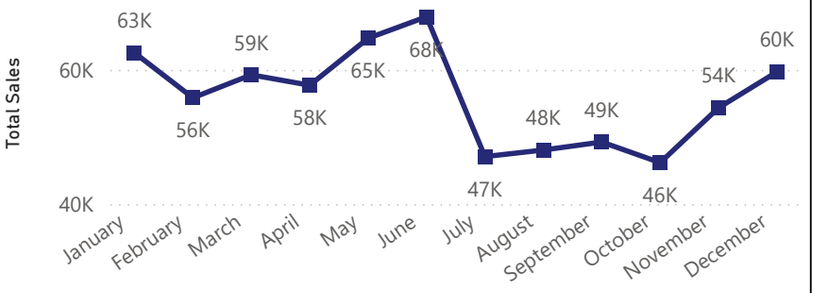


Yearly Financial Trends

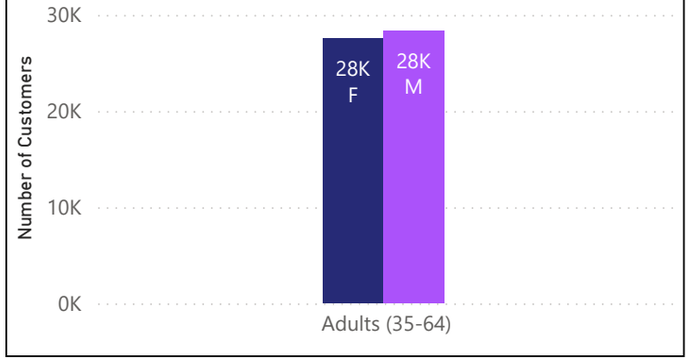


672K	26263K
Total Sales	Total Cost
42584K	**16322K**
Total Revenue	Total Profit

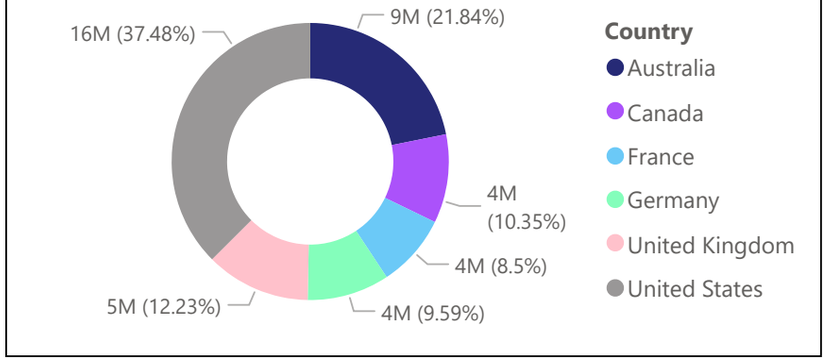
Monthly Sales Trends



Age Distribution



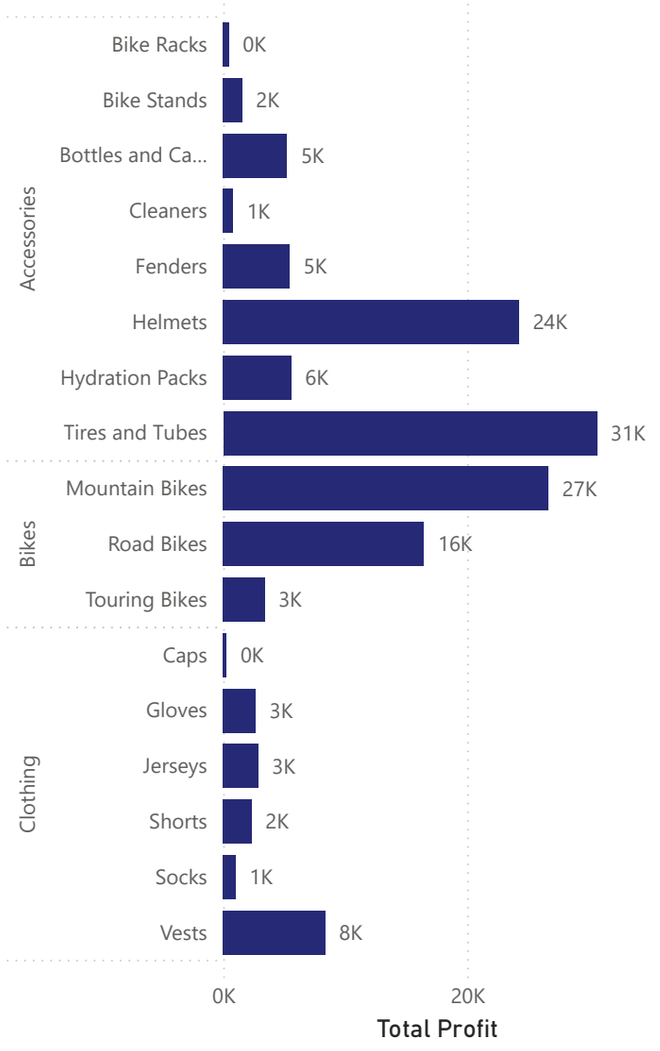
Financial Performance by Country



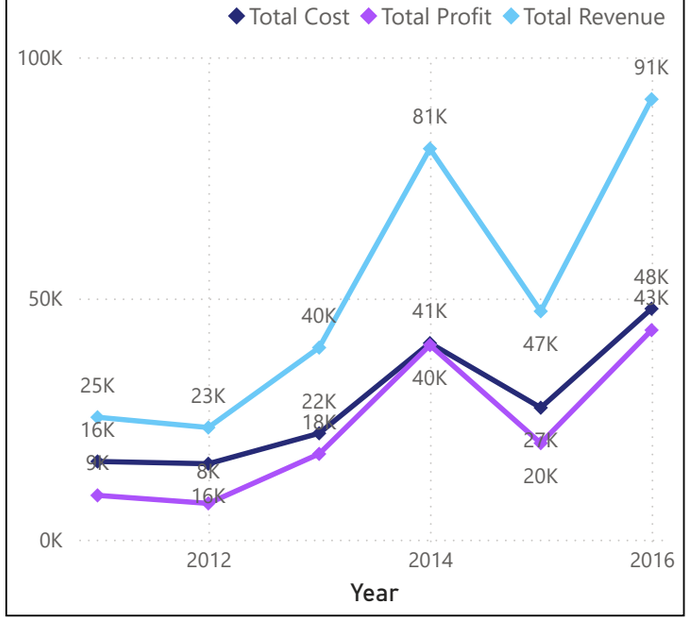
Bike Sales Dashboard

Product: All | Category: All | Month: All | Year: All | Age Group: Seniors (64+) | Gender: All | Country: All

Product Profitability

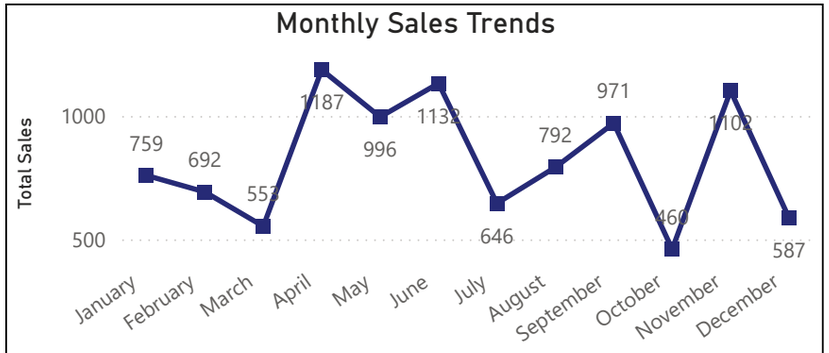


Yearly Financial Trends

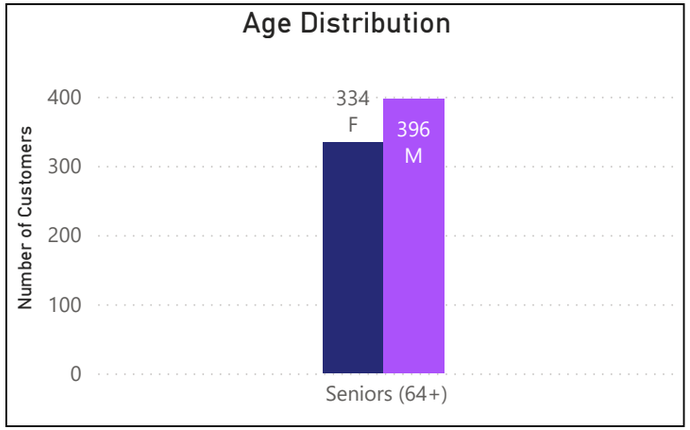


10K	170K
Total Sales	Total Cost
308K	**138K**
Total Revenue	Total Profit

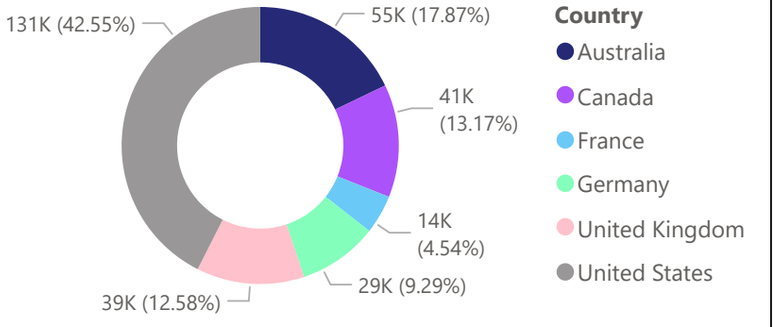
Monthly Sales Trends



Age Distribution



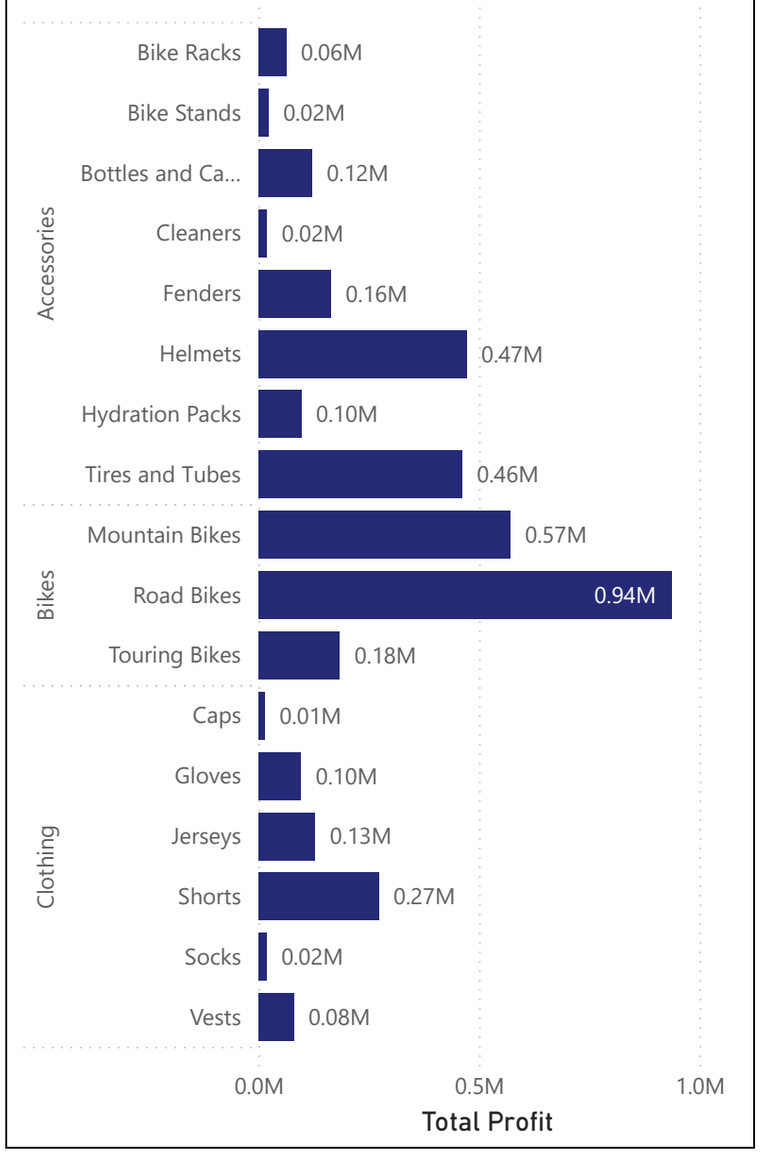
Financial Performance by Country



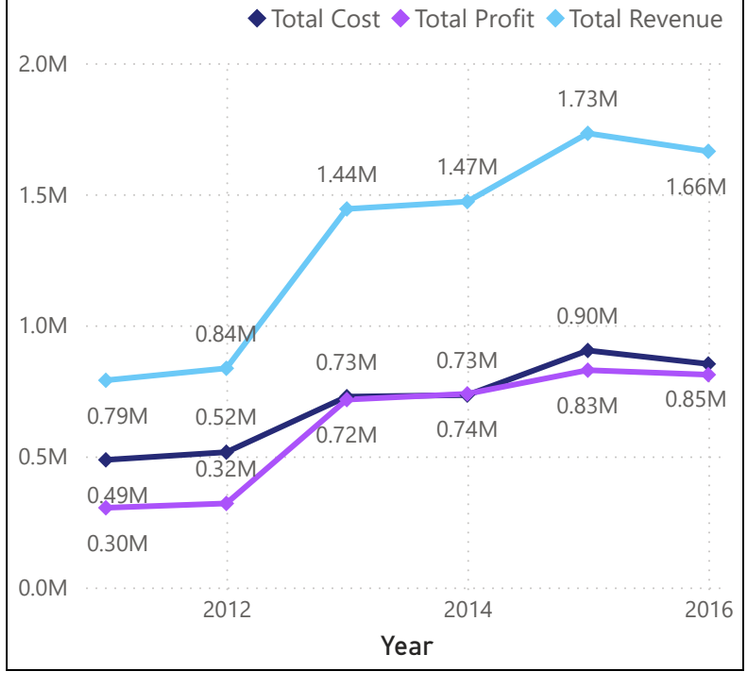
Bike Sales Dashboard

Product	Category	Month	Year	Age Group	Gender	Country
All	All	All	All	All	All	Canada

Product Profitability

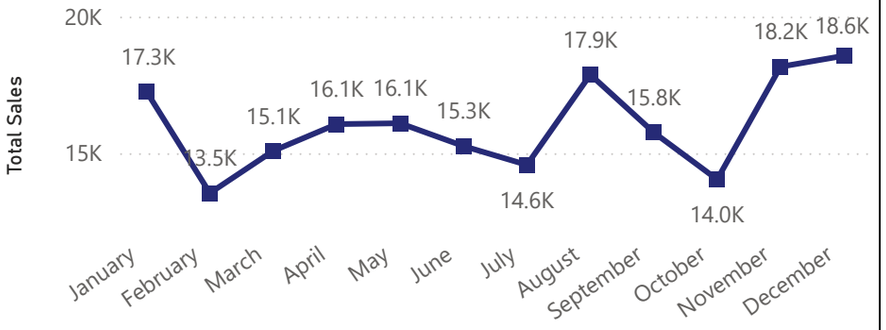


Yearly Financial Trends

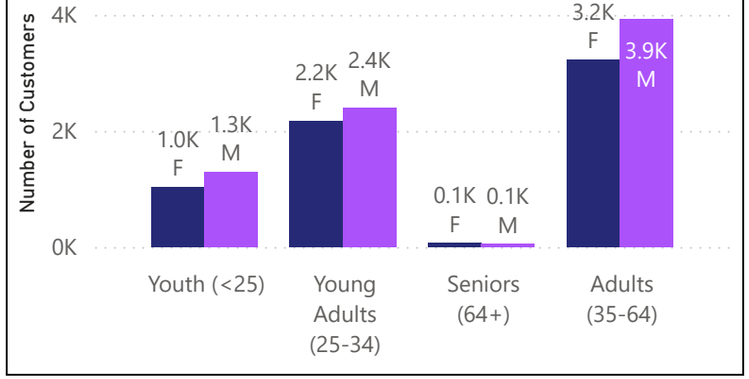


192K	4218K
Total Sales	Total Cost
7936K	**3717K**
Total Revenue	Total Profit

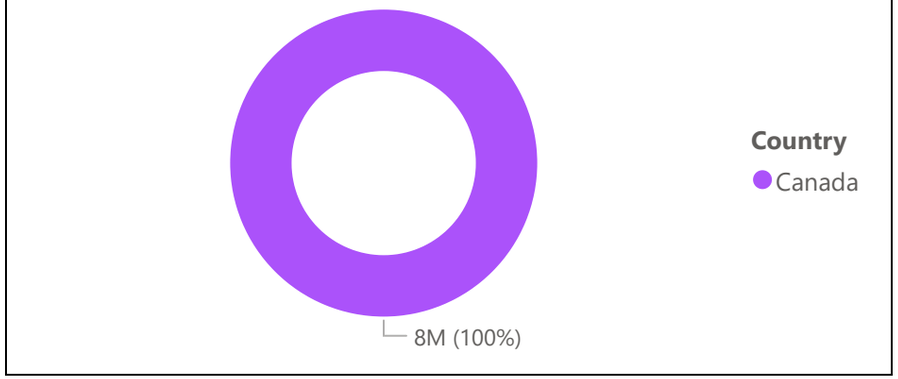
Monthly Sales Trends



Age Distribution



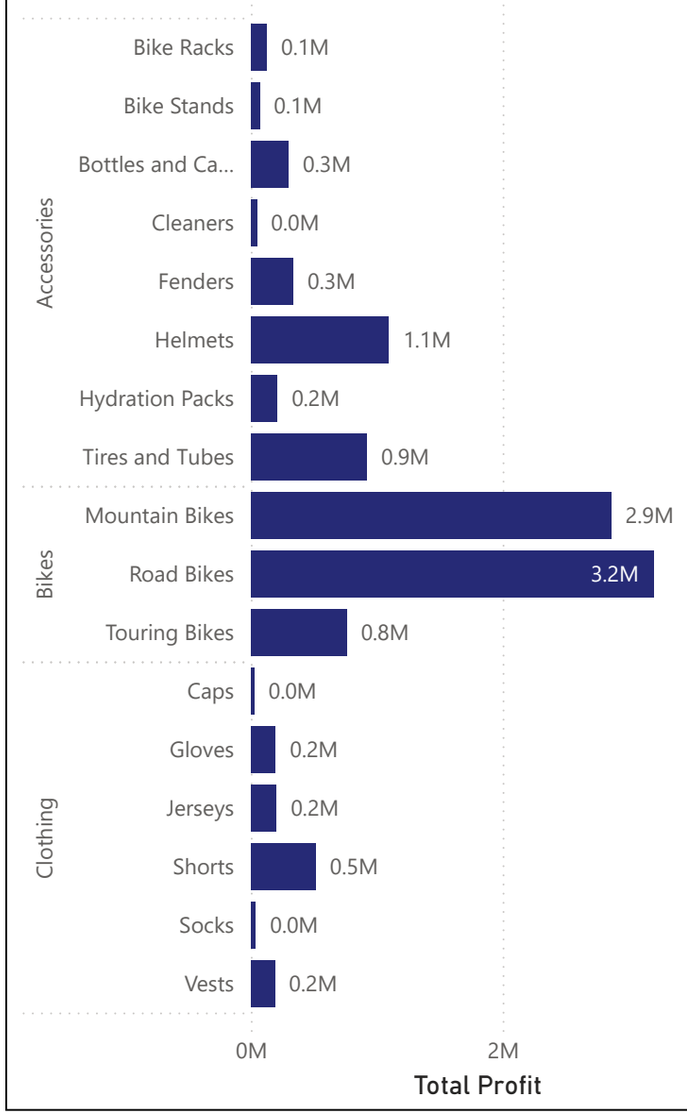
Financial Performance by Country



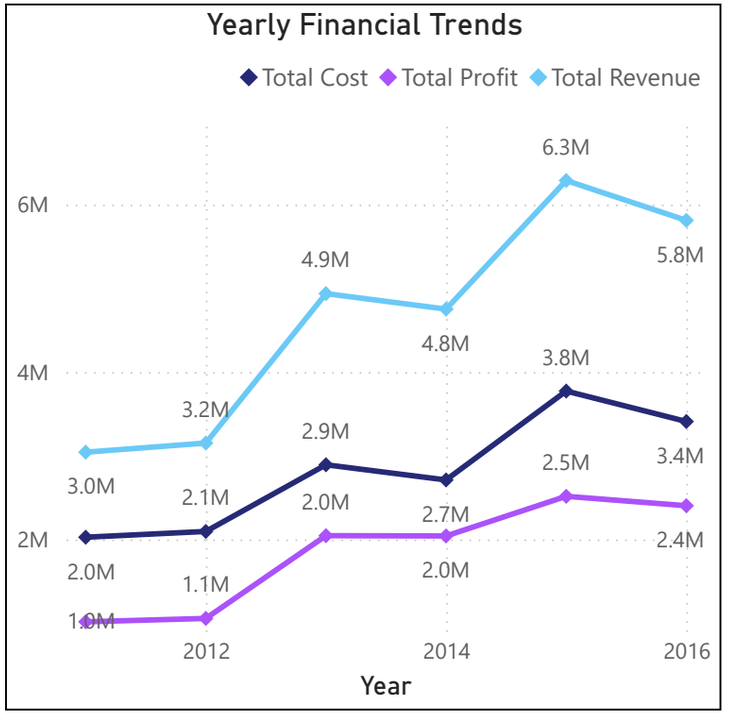
Bike Sales Dashboard

Product: All | Category: All | Month: All | Year: All | Age Group: All | Gender: All | Country: United Stat...

Product Profitability

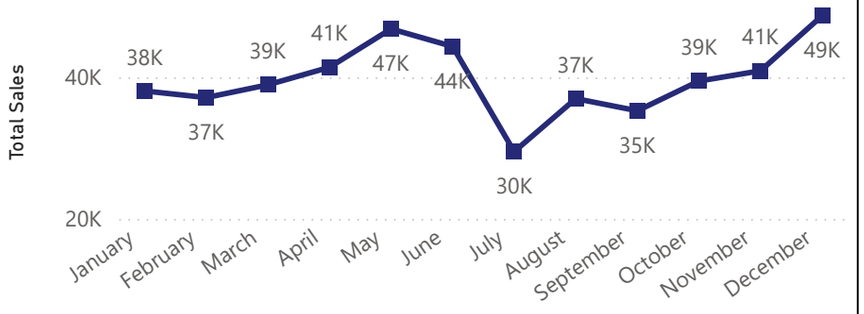


Yearly Financial Trends

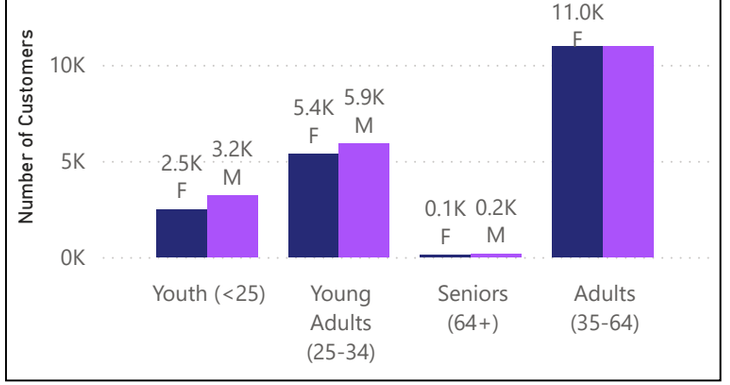
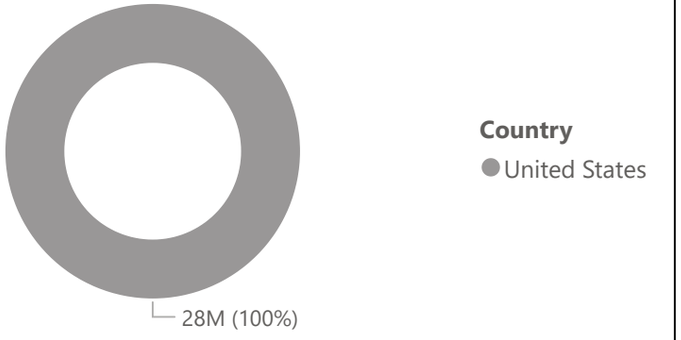


478K	16902K
Total Sales	Total Cost
27976K	**11074K**
Total Revenue	Total Profit

Monthly Sales Trends



Age Distribution



Financial Performance by Country