

BIKE SALES IN EUROPE

ANALYSIS OF BIKE SALES IN EUROPE COUNTRIES FROM 2011 TO 2015

TABLE OF CONTENT

5 TEAM MEMBERS AND ROLES

6 INTRODUCTION

7 PROJECT AIMS

8 PROJECT SCOPE

12 OBJECTIVES

13 METHODOLOGY

22 DASHBOARD AND PROTOTYPE

29 DATA INSIGHT AND DISCUSSION

43 CONCLUSION

LIST OF FIGURES

Figure 1: Milestones of the Project.....10
Figure 2: Data Science Life Cycle.....13
Figure 3: A Boxplot of Order Quantity.....17
Figure 4: A Clustered Column of Profit.....17
Figure 5: A Boxplot of Customer Age.....17
Figure 6: Derivative of New Profit Attribute.....19
Figure 7: Derivative of New Revenue Attribute.....20
Figure 8: Elements of Dashboard Page 1.....23
Figure 9: Elements of Dashboard Page 2.....24
Figure 10: Reading Flow and Objectives of Dashboard.....25
Figure 11: Dashboard's Filter.....27
Figure 12: A Clustered Column of Customer by Year.....32
Figure 13: A Pyramid Chart of Customer by Country.....33
Figure 14: A Pyramid Chart of Customer by Gender and Age Group.....34
Figure 15: A Map of Customer by State.....36
Figure 16: A Bubble Chart of Bikes Item.....37
Figure 17: A Bar Chart of Bike Sales by Sub Category.....39
Figure 18: A Stacked Bar Chart of Customer Transaction by Year.....41

LIST OF TABLES

Table 1: Gantt Chart of Project Statement of Work.....8
Table 2: Description of Attributes.....14
Table 3: Data Dictionary.....15
Table 4: Description of Elements of Dashboard Page 1.....23
Table 5: Description of Elements of Dashboard Page 2.....24
Table 6: Title of Chart and Description.....30
Table 7: Title of Chart and Purpose.....31

TEAM MEMBERS AND ROLES



NURUL AMIRA BINTI ROHI'SHAM
Head of Data Scientist

SITI HAJAR BINTI MUHAMMAD JOHAN
Data Scientist

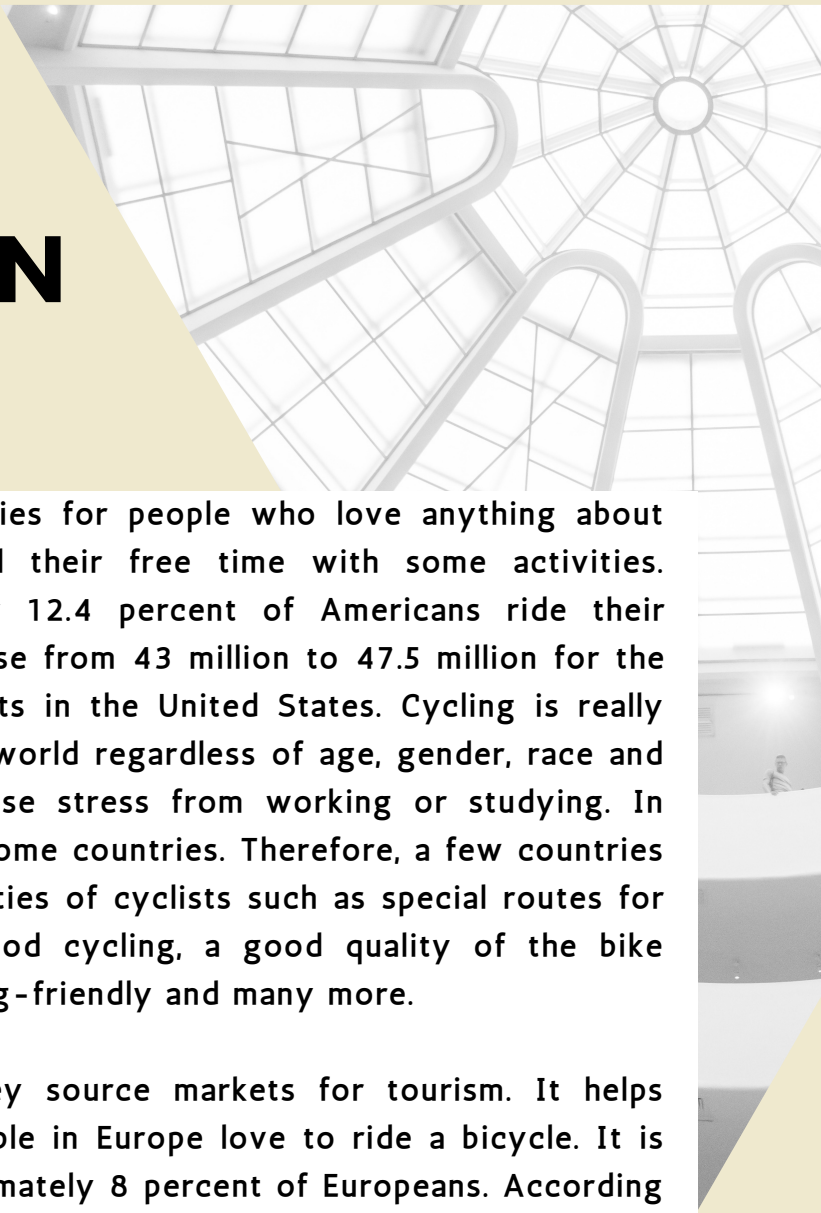




NURUL HAZLIN BINTI ZULKIFLI
Data Analyst

Youtube Link:
https://www.youtube.com/watch?v=e0dpPG2foCY

1.0 INTRODUCTION



Cycling is one of the popular activities for people who love anything about cycling or people who want to fill their free time with some activities. According to Statista, approximately 12.4 percent of Americans ride their bicycle on a daily basis. It has increase from 43 million to 47.5 million for the past three years of number of cyclists in the United States. Cycling is really popular among people all around the world regardless of age, gender, race and many more. It helps people to release stress from working or studying. In addition, cycling is an attraction for some countries. Therefore, a few countries make a decision to increase the facilities of cyclists such as special routes for safe cycling, an infrastructure of good cycling, a good quality of the bike components, accommodation of cycling-friendly and many more.

In Europe, cycling is one of the key source markets for tourism. It helps increase the markets since many people in Europe love to ride a bicycle. It is also one of the transports for approximately 8 percent of Europeans. According to the Ministry of Foreign Affairs, in 2012, the rate of average annual growth is 21 percent and by 2017, an interest in bikes has increased to 142 percent and approximately 20 million bicycles were sold in Europe for every year. One example of a popular bike company is Cycleurope. Cycleurope has evolved to be a global attraction with Europe as their main market since many people in Europe like to ride a bicycle. Hence, most companies in Europe sell a different type of component of bikes to make sure people have a good material for their bikes, especially for people who love cycling as their hobby. Not just that, it will also help people to ensure there will be no accidents that involve bikes due to bikes being too old and components already rusty.

Therefore, the data about bike sales in Europe is very valuable for the sales manager of a bike company in Europe to analyze which items are the most popular among customers. The dataset is a secondary dataset obtained from Kaggle website in format of CSV and the data have been recorded from 2011 to 2016 in Europe.

2.0 PROJECT AIMS

The aim of this project is to create a visualization of the data set for the users to see the trend of sales performance of bike sales in Europe from 2011 to 2015. This project also aims to analyze the pattern of each attribute to see the relation of each other and to give ideas to the users to help them reach the sales goals based on the cost, profit, and revenue of bike sales.



3.0 PROJECT SCOPE

A PROJECT STATEMENT OF WORK (SOW)

The Gantt chart below represents a total of 9 weeks involving four stages which are introductory, design, development, and finishing in the project statement of work (SOW).

Task name	Start Date	Progress	WEEK 7	WEEK 8	WEEK 9	WEEK 10	WEEK 11	WEEK 12	WEEK 13	WEEK 14	WEEK 15
Phase 1: Introductory											
1. Identify Project Title	14/12/2021	5%									
2. Data Preparation	14/12/2021	8%									
Identify User Requirement		12%									
Project Objective		21%									
Design Requirement	15/12/2021	23%									
Mini Project Proposal Submission	23/12/2021	28%									
Phase 2: Design											
1. Design Data Presentation	25/12/2021	34%									
Mock Design		39%									
2. Develop Data Presentation	27/12/2021	41%									
3. Project Objective Review	30/12/2021	44%									
4. Simulation and Analysis	02/01/2022	48%									
Phase 3: Development											
Chit-Chat Session	05/01/2022	55%									
2. Dashboard Critic	10/01/2022	58%									
3. Mini Project Discussion	21/01/2022	65%									
Phase 4: Finishing											
1. Mini Project Report Discussion	28/01/2022	72%									
2. Mini Project Report Writing	04/02/2022	91%									
3. Final Project Submission	12/02/2022	95%									
Video Project Submission	12/02/2022	100%									

Table 1: Gantt Chart of Project Statement of Work

At the beginning of the project planning, the researcher has spent 6 weeks in planning and designing the dashboard by using the first data that was obtained from the Kaggle website. Due to some issues, the researcher has to change to another data which is a bike sales dataset. The data used in this project is different from what was planned, and the researcher started the progression for this project at week 7. However, the researcher is trying to keep on track to complete within the allocated time.

CONSTRAINTS

To do research, it is important to ensure that the data that has been used is reliable and accurate. Any constraints and limitations might negatively impact the outcome of the project. In this project, data constraints arise during data preparation. There are two attributes that consist of an error observation value due to error in calculation profit and revenue. Other than that, timing in completing this project is the main constraint for the researcher since the researcher has to exchange the data and to get the same progression with other groups.

SCOPE EXCLUSION

The bike sales dataset consists of 113,037 observations and 18 attributes that explain product details with customer profile within six countries from year 2011 until 2015. However, the researcher is unable to cover all observations for this project since the data does not provide a complete month for several years. For instance, the data only provide the information from January until July for 2014 and 2016. It is difficult for the researcher to define the sales, revenue and cost if these two years are being included during data analysis. Hence, year 2014 and 2016 are excluded from being used in this project since that does not give a complete month of the data.

MILESTONES

There are four stages in milestones for this project. First stage is data acquisition, defined as the process of getting the data. Define data is the second stage, where the researcher has to define each of the attributes in the dataset. Next stage is dashboard mock design and the last stage in milestones is complete design and generate a report. The figure 1 below shows the precise date for each stage in milestones in order for this project.

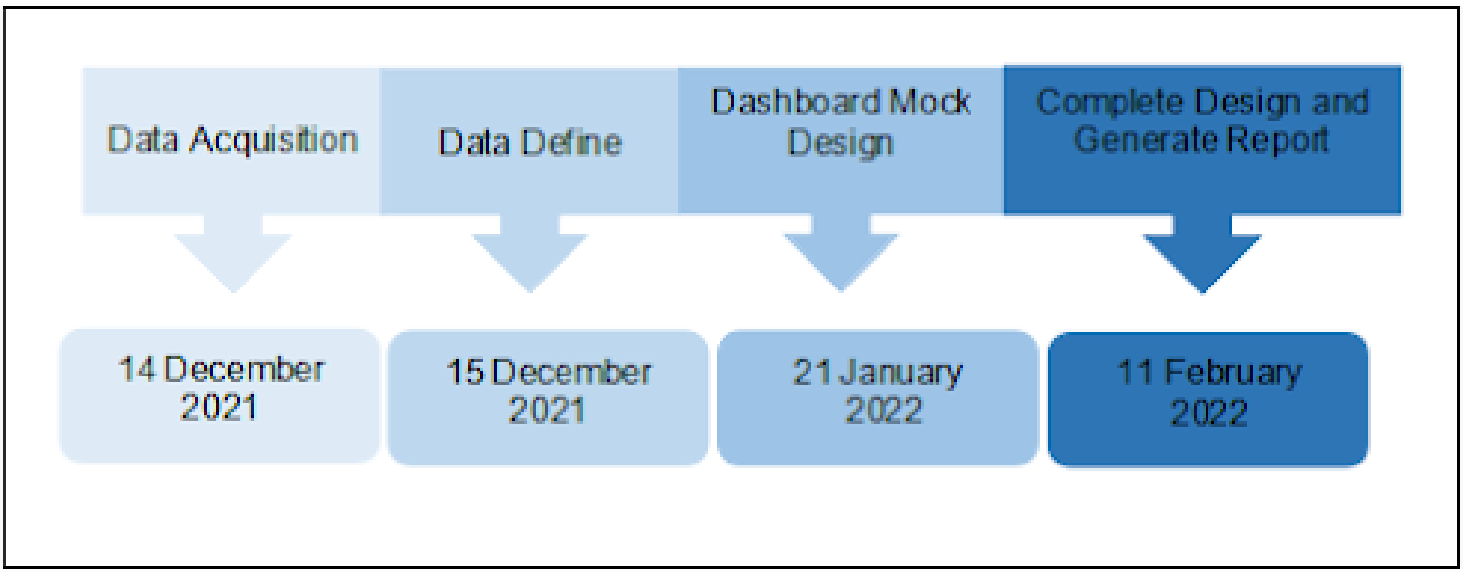


Figure 1: Milestones of the Project

FINAL DELIVERABLES

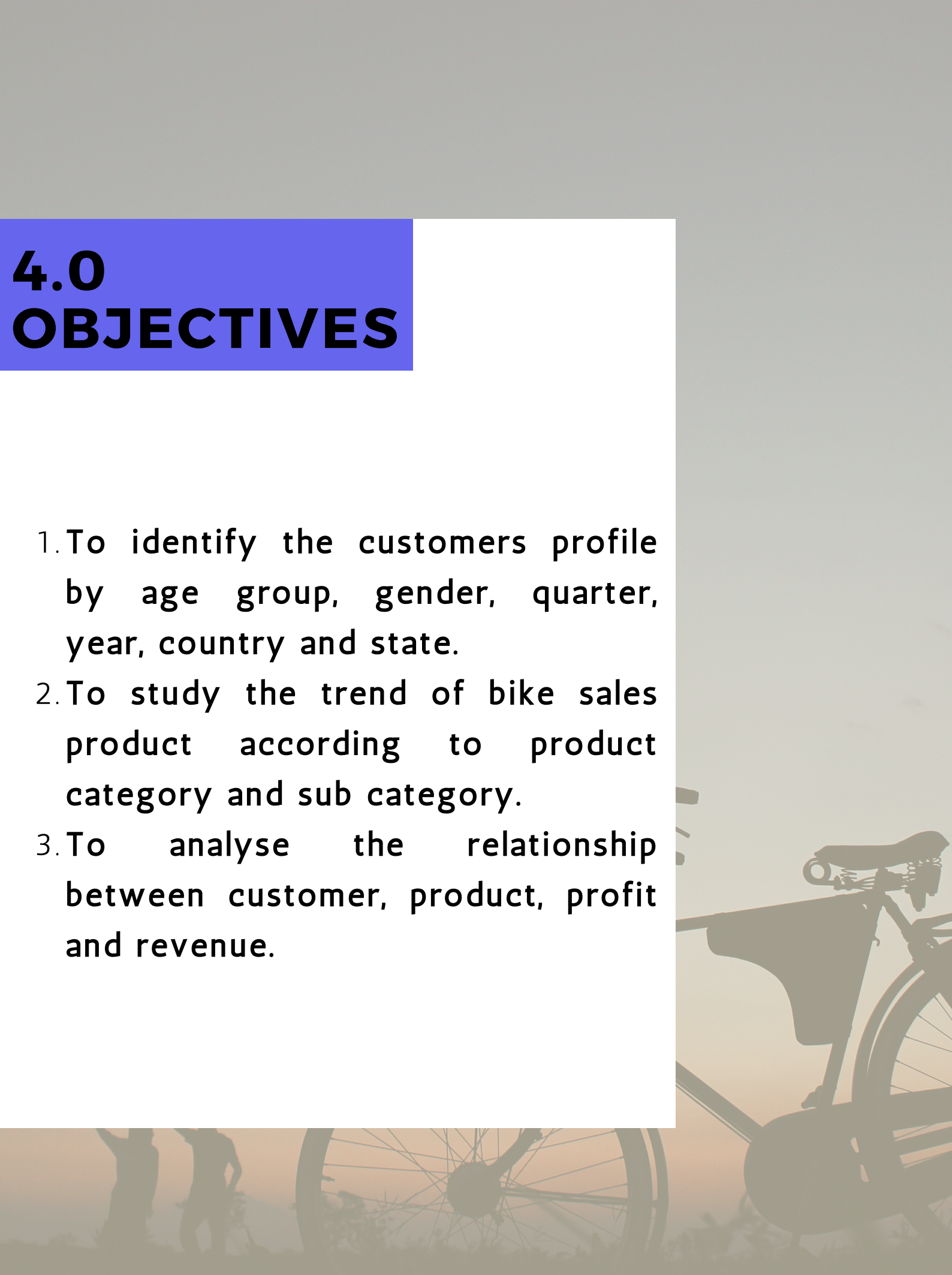
The final deliverables will be provided to the user to make the user more understandable about the data and information that the researcher wants to deliver. The outcome for this project is a clean data with derivatives attributes that has been done in Microsoft Excel. Furthermore, dashboard, report and video presentation are also being provided to the user. The dashboard that has been visualized by using Power BI comes with a few charts in order to see the pattern of the study for this project. Eventually, formal reports and video presentations are also being given to ensure the user reaches the level of understanding that the researcher wants and to avoid any misunderstanding about this project.

ACCEPTANCE CRITERIA

This project manages to be completed in 9 weeks with a good teamwork of the researchers and the help of the experts. In the meantime, this project achieved all objectives and the aim of the project had been accomplished. The conclusion for this study also can be derived based on all aspects. Therefore, the acceptance criteria are accomplished since the points and goals of this consideration are met.

4.0 OBJECTIVES

1. To identify the customers profile by age group, gender, quarter, year, country and state.
2. To study the trend of bike sales product according to product category and sub category.
3. To analyse the relationship between customer, product, profit and revenue.



5.0 METHODOLOGY

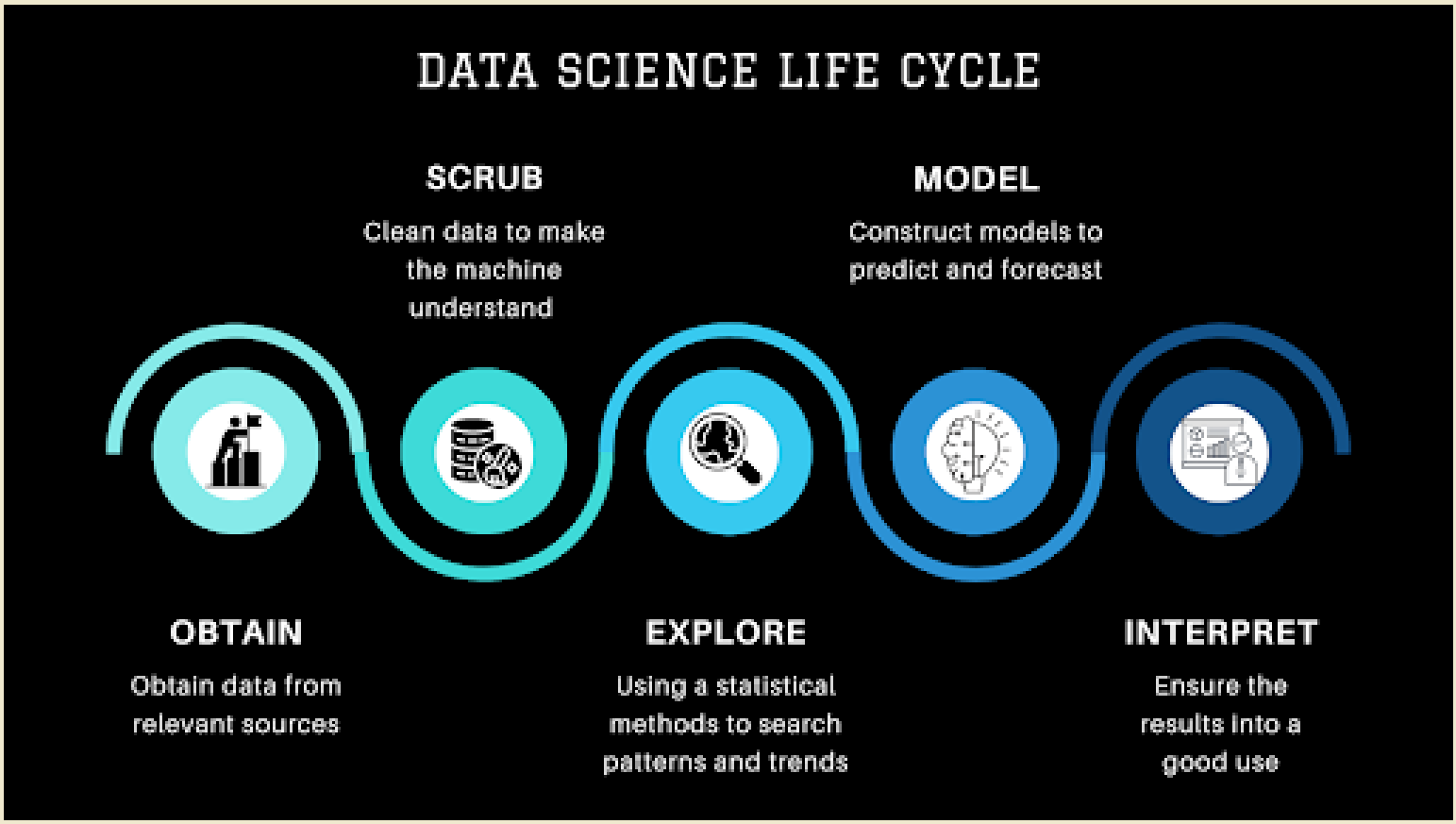


Figure 2: Data Science Life Cycle

Data science life cycle is a series of data science steps that you go through to complete a project or analysis. There are 5 steps in the data science life cycle that need to go through before successfully completing the project which are obtain, scrub, explore, model and interpret. The first step which is obtain is to obtain data from relevant sources. To generate the data, MySQL is used but if want to make it easier, Microsoft Excel is the best option for file formats. Other than that Python and R also is the simplest since it directly imports the datasets into the program. The second step is scrub which is to clean data to make the machine understand. By that, it actually cleans the raw data to make sure the data that has been processed is relevant data. Next step is explore which is using statistical methods to search patterns and trends. In these steps, from the data that has been cleaned, the data will be studied and examined since it has different types of data. To test the significant variables, descriptive statistics are computed. By exploring the data, it will create a better narration of storytelling for the data. After that, the next step is to construct models to predict and forecast. In this step, any data that is not useful will be removed so that correct data will be represented. Last step is interpret which is to ensure the results into a good use so that users can analyze the data and be able to achieve their objectives and meet the requirements needed. Hence the data will be visualized in the dashboard to make it useful for the future.

5.1 DATA DESCRIPTION AND DATA DICTIONARY

Table 2 below shows description of attributes:

Attribute	Meaning
Date	The date of transaction made that consists of day, month and year
Day	The day of transaction made
Month	The month of transaction made
Year	The year of transaction made
Customer Age	The age of the customer
Age Group	Customer's age group
Customer Gender	The gender of the customer
Country	The name of country that the customer purchased the product
State	The name of state that the customer purchased the product
Product Category	Category of product
Sub Category	Sub category of product
Product	The details of the product such as name and size of product
Order Quantity	The quantity of product per order
Unit Cost	Cost of product per unit
Unit Price	Price of product per unit
Profit	Total of profit
Cost	Total of cost
Revenue	Total of revenue

Table 2: Description of Attributes

Table 3 indicates data dictionary of bike sales in Europe dataset:

No	Variable Type	Variable Name	Data Type	Level of Measurement
1.	Independent Variable	Date	Date	Interval
2.	Independent Variable	Day	Numeric	Interval
3.	Independent Variable	Month	Alphabetic	Nominal
4.	Independent Variable	Year	Numeric	Interval
5.	Independent Variable	Customer Age	Numeric	Interval
6.	Independent Variable	Age Group	Alphabetic and Numeric	Ordinal
7.	Independent Variable	Customer Gender	Alphabetic	Nominal
8.	Independent Variable	Country	Alphabetic	Nominal
9.	Independent Variable	State	Alphabetic	Nominal
10.	Independent Variable	Product Category	Alphabetic	Nominal
11.	Independent Variable	Sub Category	Alphabetic	Nominal
12.	Independent Variable	Product	Alphabetic and Numeric	Nominal
13.	Dependent Variable	Order Quantity	Numeric	Interval
14.	Dependent Variable	Unit Cost	Numeric	Ratio
15.	Dependent Variable	Unit Price	Numeric	Ratio
16.	Dependent Variable	Profit	Numeric	Ratio
17.	Dependent Variable	Cost	Numeric	Ratio
18.	Dependent Variable	Revenue	Numeric	Ratio

Table 3: Data Dictionary

5.2 STAGES AND STEPS

Stage 1: Data and context

Data collection was involved as a first step in the data and context stage. Secondary data has been used in this project that was obtained from Kaggle website in CSV format. The title of the dataset is Bike Sales in Europe: Bike Sales in Europe Countries. The data source URL is shown as below:

https://www.kaggle.com/sadiqshah/bike-sales-in-europe

Data description, data examination and data preparation comes as a second step of this stage. Data description is the information about the dataset that describes each attribute as mentioned in 5.1. This stage is considered as a really important part to the researcher, as it is to understand the real meaning of the dataset and to make a connection among attributes. Next, data examination is performed on the data to determine the possible decisions that can be made to the data. Through the data used in this project, the researcher had identified two attributes that consist of an error observation value. The more information will be explained through data issues in 5.4. However, the other attributes do not acquire any critical decisions to make. Data preparation is a data pre-processing stage that is also part of the cleaning process. The purpose is to ensure that the data are being readied and to produce consistent results. No eliminating action has been taken on the data since the data is correctly formatted and there is no missing value.



Exploratory Data Analysis (EDA) is shown as below:



Figure 3: A Boxplot of Order Quantity

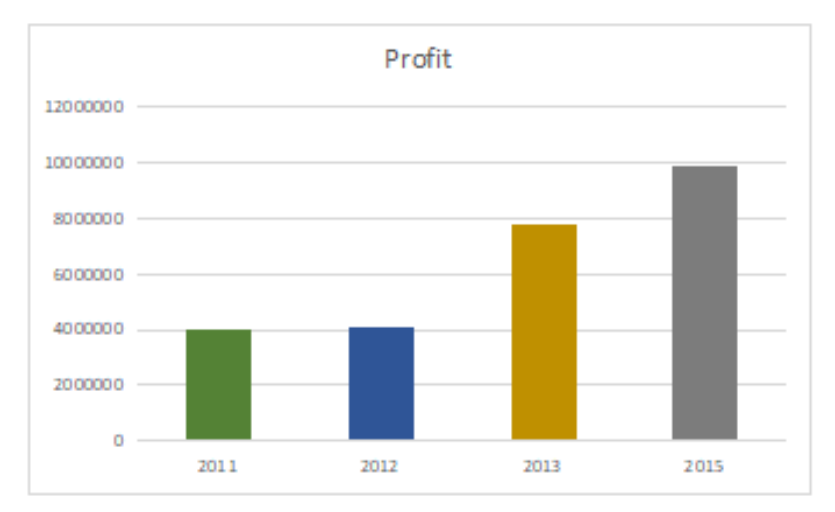


Figure 4: A Clustered Column of Profit



Figure 5: A Boxplot of Customer Age

The third step is identifying the target user. The target user for this project is the sales manager of the company. It helps in recognizing KPIs and data trends, by providing optimal data on the business's performance with real-time data. It allows the management to make strategic decisions to summarize and track actionable KPIs that are important in daily operations.

The fourth step is objectives and scope identification. The objectives are focused on customer profile, bikes sales product and financial arrangement. The scope of the project is also identified. These all the elements have been stated in 3.0 and 4.0.

Project making plans is the final step, where the undertaking plan for every week is defined to ensure that the plan is kept on track and any problem throughout this project can be resolved in the right time. The planning of this project is illustrated in the Gantt chart on project scope.

Stage 2: Visual design

The visual design or data visualization is done by using Microsoft Power BI. The objective of visual design is to extract actionable insights, identify trends and patterns. Hence, a good dashboard should be represented clearly, which has direct consequences in the use of filters in displaying the information and interactive that can guide the user to understand more into certain trends and metrics. Types of chart and choosing the colors of the dashboard are also important in data visualization. The aim is to choose the right type of chart that suits the information that the researcher wants to convey and to be consistent in using color when designing a dashboard.

Stage 3: Storytelling narration

Storytelling narration is the main thing that should be focused on during making the dashboard. This is because it is a process of presenting data to depict the whole narrative of overall key performance and approaches in creating interactive dashboards in real-time. The researchers gave insights presentations and made a storytelling from each dashboard to present the trends and patterns. In this project, the data has been sorted from the year 2011 until 2015 to define the pattern in the number of customers according to each quarter. This could assist the audience to apprehend the stories conveyed through the charts. Ultimately, the researchers derived the conclusion after summarizing based on charts in each dashboard.

5.3 DERIVATIVES ATTRIBUTES

There are 2 attributes had been derived for this project which are new profit and new revenue.

i.New Profit

Profit attribute is calculated by deducting all direct and indirect costs from total sales that are used to derive a new profit attribute. New profit attribute is derived to recalculate the profit attribute and make a correction to the data since the profit attribute is miscalculated. The function that is used to derive this attribute is:

```
=S2-Q2
```

where S2 refers to the first cell of the new revenue attribute while Q2 refers to the first cell of cost in the data. The result is shown as below:

T2 =S2-Q2

	M	N	O	P	Q	R	S	
1	Unit_P	Order_	Unit_C	Profit	Cost	Reven	New Revenu	New Prof
2	120	8	45	590	360	950	960	600
3	120	8	45	590	360	950	960	600
4	120	8	45	398	360	758	960	600
5	120	8	45	398	360	758	960	600
6	120	7	45	349	315	664	840	525
7	120	7	45	349	315	664	840	525
8	120	5	45	369	225	594	600	375
9	120	7	45	517	315	832	840	525
10	120	2	45	148	90	238	240	150
11	120	1	45	74	45	119	120	75
12	120	3	45	221	135	356	360	225
13	120	1	45	74	45	119	120	75
14	120	6	45	443	270	713	720	450
15	120	5	45	369	225	594	600	375
16	120	9	45	524	405	929	1080	675
17	120	9	45	524	405	929	1080	675
18	120	25	45	1845	1125	2970	3000	1875
19	120	23	45	1697	1035	2732	2760	1725
20	120	10	45	726	450	1176	1200	750
21	120	10	45	726	450	1176	1200	750
22	120	8	45	581	360	941	960	600
23	120	10	45	726	450	1176	1200	750
24	120	2	45	148	90	238	240	150
25	120	1	45	74	45	119	120	75
26	120	12	45	886	540	1426	1440	900
27	120	14	45	1033	630	1663	1680	1050
28	120	3	45	160	135	295	360	225
29	120	1	45	53	45	98	120	75
30	120	8	45	427	360	787	960	600
31	120	8	45	427	360	787	960	600
32	120	28	45	1697	1260	2957	3360	2100
33	120	27	45	1636	1215	2851	3240	2025

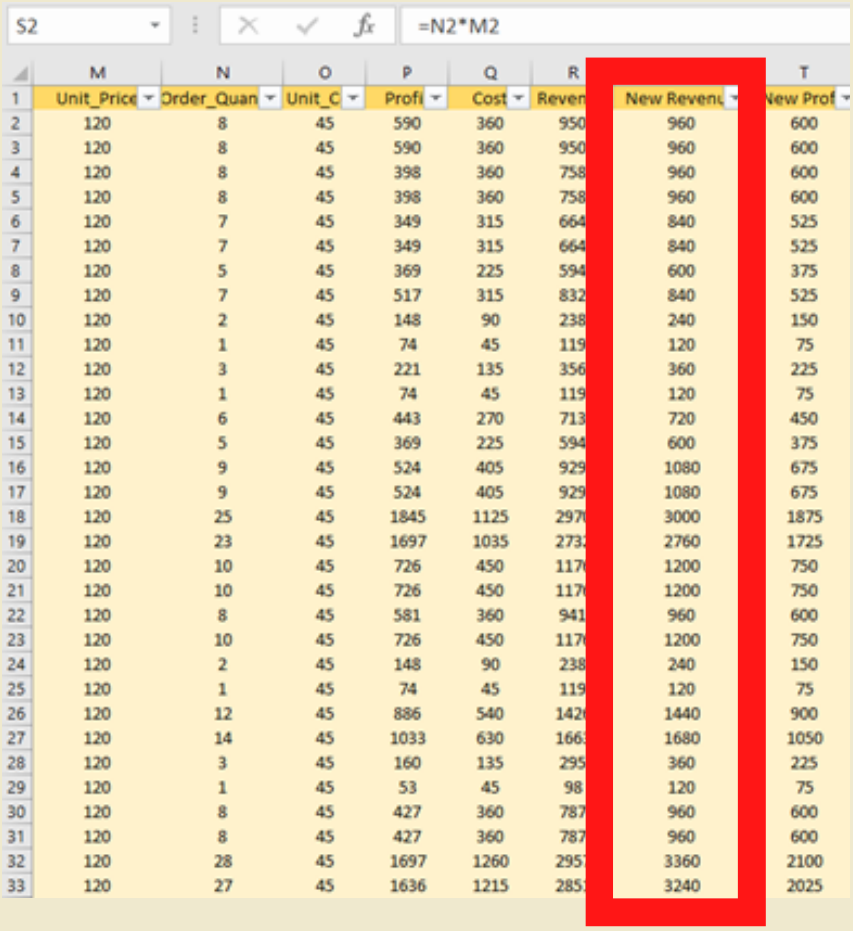
Figure 6: Derivative of New Profit Attribute

ii.New Revenue

Revenue attribute is the total amount of revenue generated through the sale of goods or services connected to the company's principal business is referred to as revenue that are used to derive a new revenue attribute. New revenue attribute is derived to recalculate the revenue attribute and make a correction to the data since the revenue attribute is miscalculated. The function that is used to derive this attribute is:

```
=N2*M2
```

where N2 refers to the first cell of the order quantity attribute while M2 refers to the first cell of unit price in the data. The result is shown as below:



S2 =N2*M2

	M	N	O	P	Q	R	New Revenu	T
1	Unit_Price	Order_Quan	Unit_C	Profi	Cost	Reven	New Revenu	New Prof
2	120	8	45	590	360	950	960	600
3	120	8	45	590	360	950	960	600
4	120	8	45	398	360	758	960	600
5	120	8	45	398	360	758	960	600
6	120	7	45	349	315	664	840	525
7	120	7	45	349	315	664	840	525
8	120	5	45	369	225	594	600	375
9	120	7	45	517	315	832	840	525
10	120	2	45	148	90	238	240	150
11	120	1	45	74	45	119	120	75
12	120	3	45	221	135	356	360	225
13	120	1	45	74	45	119	120	75
14	120	6	45	443	270	713	720	450
15	120	5	45	369	225	594	600	375
16	120	9	45	524	405	929	1080	675
17	120	9	45	524	405	929	1080	675
18	120	25	45	1845	1125	297	3000	1875
19	120	23	45	1697	1035	273	2760	1725
20	120	10	45	726	450	117	1200	750
21	120	10	45	726	450	117	1200	750
22	120	8	45	581	360	941	960	600
23	120	10	45	726	450	117	1200	750
24	120	2	45	148	90	238	240	150
25	120	1	45	74	45	119	120	75
26	120	12	45	886	540	142	1440	900
27	120	14	45	1033	630	166	1680	1050
28	120	3	45	160	135	295	360	225
29	120	1	45	53	45	98	120	75
30	120	8	45	427	360	787	960	600
31	120	8	45	427	360	787	960	600
32	120	28	45	1697	1260	295	3360	2100
33	120	27	45	1636	1215	285	3240	2025

Figure 7: Derivative of New Revenue Attribute



5.4 DATA ISSUES

Data issues is to examine the issues that can arise in the data such as duplicate data, unstructured data, missing data, incompatible data formats, or trouble accessing the data. This session is used to examine the data issues for bike sales in Europe dataset. Before beginning to visualize data, it is critical to ensure that the dataset is accurate and completeness. During the data preparation process, a few data concerns develop. The issues are as follows:

i. Error of calculation in the dataset

In this dataset, the first issue that has been identified in the dataset is there is an error of calculation in two attributes which are profit and revenue. Users may miscalculate the total bike sales in Europe as a result of the inaccuracy also can potentially lead to data misinterpretation by the user. This is due to the fact that the original data employed the formula incorrectly. Thus, to overcome this issue, we were able to add two new attributes for new profit and new revenue by recalculating the calculation with the right formula.

ii. There are insufficient months in the dataset

This data is partly because there are two years, 2014 and 2016, that do not have enough months and only provide information from January to July in both years. As a result, analyzing the data to identify real sales by year is challenging for the user due to the inability to calculate the exact reading each year. Hence, to overcome these issues, the years 2014 and 2016 are not included in this project because they do not provide a complete month of data. As a result, we will remove all data in 2014 and 2016 from the dataset. This data will also reduce from 113,037 to 54240.





6.0 DASHBOARD AND PROTOTYPE

6.1 DESIGN AND IMPLEMENTATION

The target user for this dashboard is the sales manager of a bike company in Europe. It helps in recognizing KPIs and data trends, by providing optimal data on the business's performance with real-time data. It allows the management to make strategic decisions to summarize and track actionable KPIs that are important in daily operations. Strategic dashboard is presented for this dashboard since it is updated on a daily basis. Additionally, it assists the company in tracking and managing top-line key performance indicators.



Elements of Dashboard Page 1

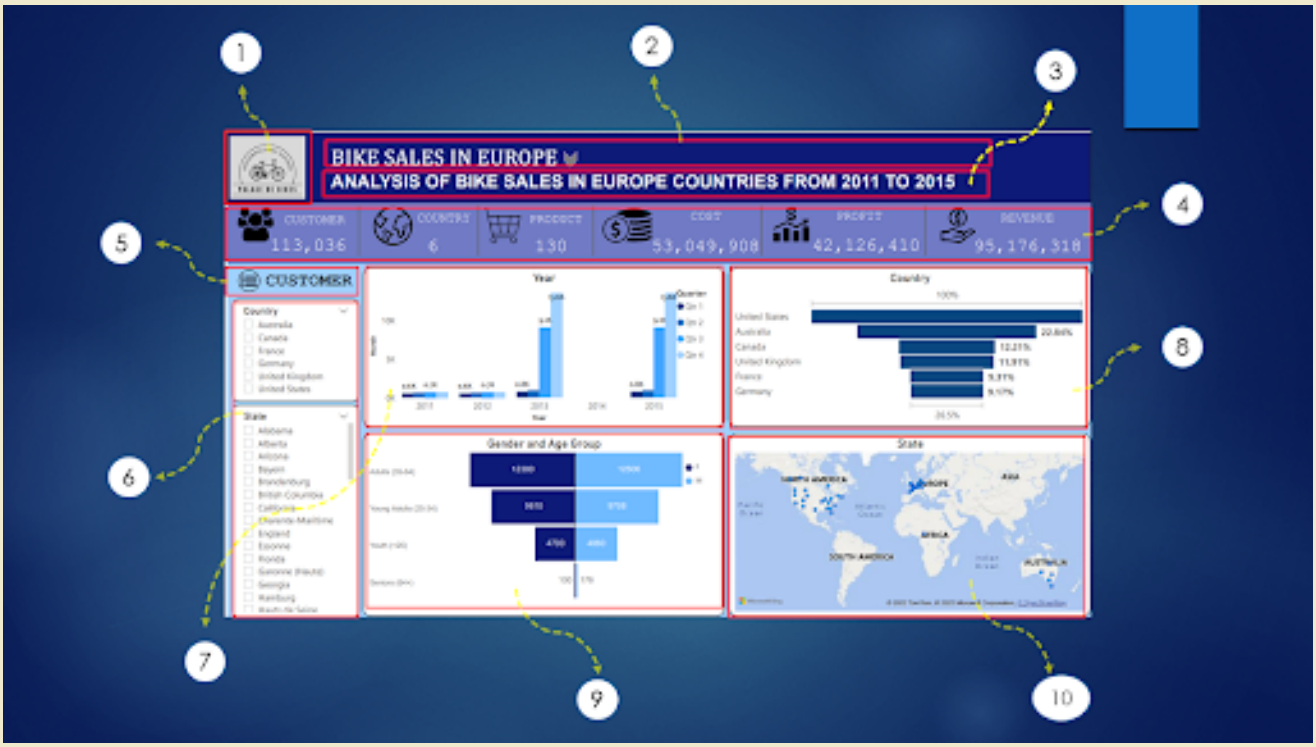


Figure 8: Elements of Dashboard Page 1

In the first page, there are 10 elements:

Label	Description
1	Logo and name of the company
2	Title of the dashboard
3	Description of the dashboard
4	Card for displaying the total of customer, country, product, unit cost, profit and revenue of bike sales in Europe
5	Card displaying customer profile
6	Filters of country and state in Europe
7	A clustered column chart of the total customer by year
8	A funnel chart of the total customer by country
9	A tornado chart of the total customer by gender and age group
10	A maps of customer by state in Europe

Table 4: Description of Elements of Dashboard Page 1

Elements of Dashboard Page 2



Figure 9: Elements of Dashboard Page 2

In the first page, there are 6 elements:

Label	Description
1	Logo and name of the company
2	Title of the dashboard
3	Description of the dashboard
4	A bubble chart of bikes sales product according to product category
5	A clustered column chart of bikes sales by sub category
6	A stacked bar chart of transaction between customer, product and sales by year

Table 5: Description of Elements of Dashboard Page 2

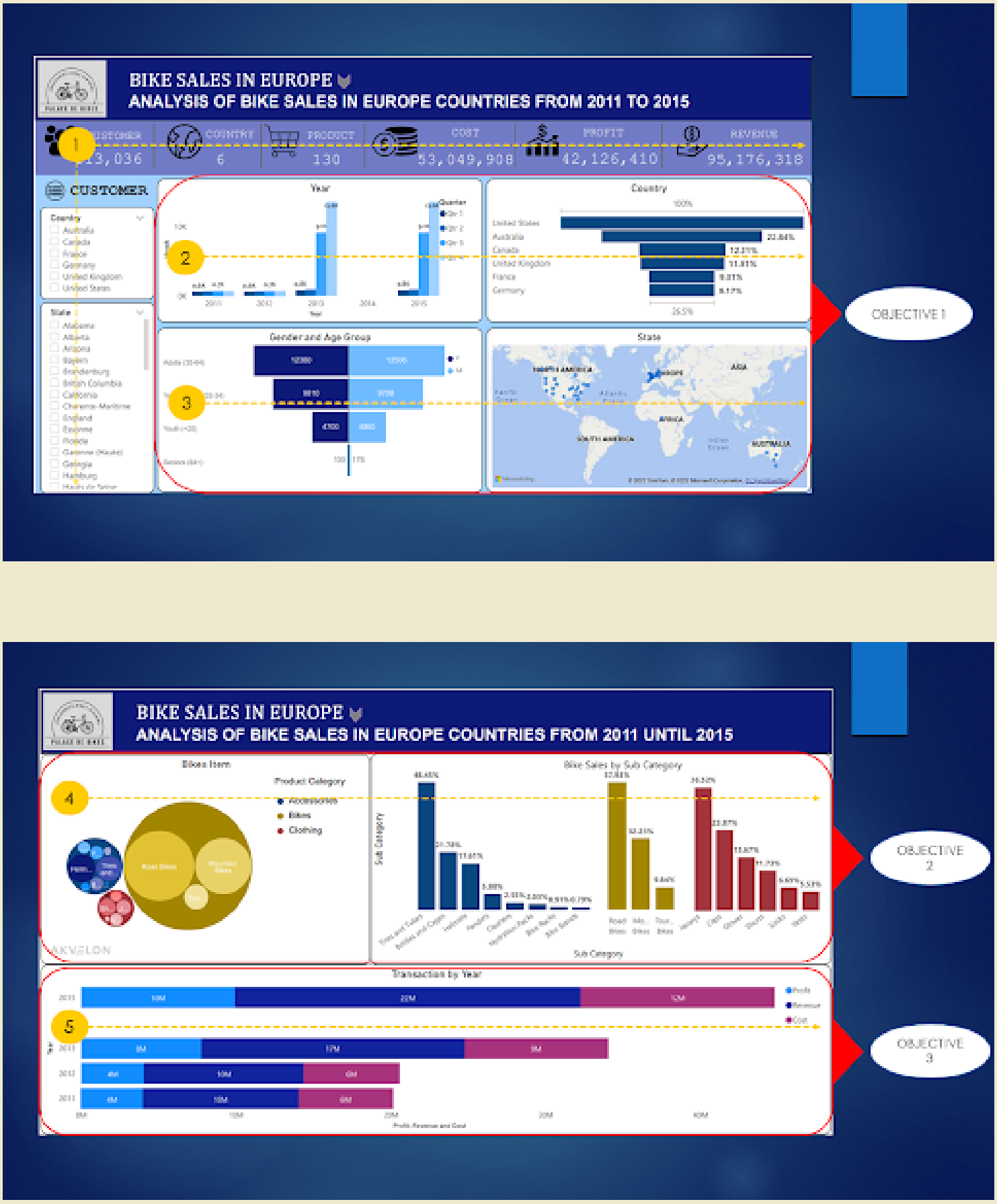


Figure 10: Reading Flow and Objectives of Dashboard



Based on the figure 5, shows the reading flow and objectives of the dashboard where there are 2 pages in this study. The reading flow of the dashboard starts with page 1. In the first row, the user will first look at the introduction of this project which is to the analysis of bike sales in Europe countries from 2011 to 2015. In the intro of this project, the user will look at the first row that displays six cards that analyze the total number of customers, country, product, cost, profit, and revenue in Europe. Also, the user will look at the down arrow for a filter that is used to navigate the data that they want. After that, the user will move to the next row which are row number 2 and number 3 representing objective 1 which is to identify the customers profile by age group, gender, quarter, year, country, and state. In row number 2, the user will look at the first chart where the clustered column chart of the total customer by year then will continue to the funnel chart of the total customer by country. Lastly on page 1, the user will look at the third row that displays the tornado chart of the total customer by gender and age group and then will look at the maps of the customer by the state in Europe.

Next in page 2, there are 2 rows which are row numbers 4 and 5 that represent objectives 2 and 3 respectively. The user will look at the fourth row on page 2 which is for objective 2 that to study the trend of bike sales product according to product category and sub category. This row consists of 2 columns which are the bubble chart of bikes sales product according to product category and the next is the clustered column chart of bikes sales by sub category. Lastly, the user will look at the last row which is row 5 that to analyze the relationship between customer, product, profit, and revenue. In this row, there will be one column only for users used which is the stacked bar chart of a transaction between a customer, product, and sales by year.

Dashboard link:
https://app.powerbi.com/links/pWZ1Czqtto?ctid=cdcbb0e2-9fea-4f54-8670-672707797ada&pbi_source=linkShare



6.2 INTERACTION FEATURES

Additional interactive tools enabled users to interact with the data in real time, such as creating and assigning tasks, or quickly generating and filtering reports, allowing them to gain a better knowledge of the dashboard. Moreover, this allows users to access data to express results quickly, save time, and increase overall productivity. Some of the interaction tools that have been implemented in the dashboard are described below.

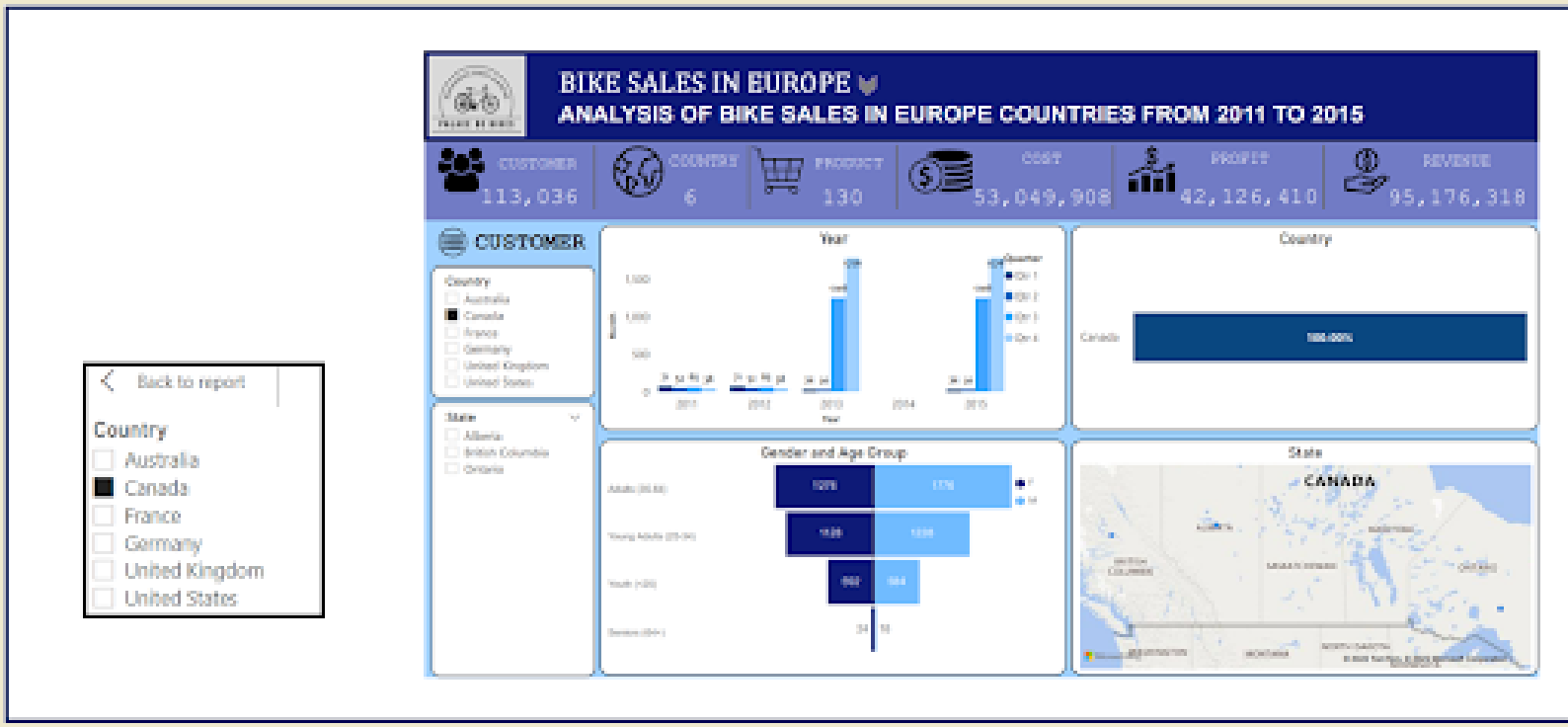


Figure 11: Dashboard's Filter

1. Filter

Filter is one the main features when using a dashboard since it will make it easier for the user to look into more details about the chart. Filter can sort the data with a specified condition that has been selected. It is easier to use by using slicer and selecting which data that need to look deep into. This will instantly provide new perspectives of the data on the dashboard. For example that shown below, the slicer shows that Canada is the country that has been selected so that the user can look deeper into. Hence, the dashboard will show the specific data of Canada for any charts that are related. Furthermore, it is also easier to go back to the unfiltered slicer so that the user can choose another piece of data that needs to look deep into with only a few fast clicks.



2. Color

The choices of color are very important since it helps the researcher to deliver and communicate the things that hold meaning to the story of the dashboard more effectively. Choosing the right tone of the color guides attention to the user besides helping in understanding the main key of the dashboard. In this project, the researcher tends to use soft prominent colors such as blue, red and purple to engage the user and capture their attention quickly. Font colors that have been applied on the dashboard are in black and white color, and mostly it consists of consistent uses of the color.





7.0 DATA INSIGHT AND DISCUSSION

7.1 VISUAL DESCRIPTION AND PURPOSES

No	Title of Chart	Description
1.	Quarter by Year	Use three attributes which are year, quarter and month to determine which year based on every quarter has the highest bike sales in 4 years from 2011 to 2015.
2.	Country	Use only one attribute which is country to identify which country is the highest percentage in purchasing the products.
3.	Gender and Age Group	Use two attributes which are gender and age group to identify which gender that group by age group has the highest number to purchase the products.
4.	State	Use only one attribute which is state to study which state has the most buyers purchasing the products.
5.	Bikes Item	Use two attributes which are product category and sub category to identify what product category and sub category are the most purchased from the buyers.
6.	Bike Sales by Sub Category	Use two attributes which are product category and sub category to determine which sub category has the highest percentage sales group by product category.
7.	Transaction by Year measured by Euro	Use four attributes which are year, revenue, cost and profit to identify the sales transaction of bike sales according to revenue, cost and profit in 4 year from 2011 to 2015.

Table 6: Title of Chart and Description

No	Title of Chart	Purpose
1.	Quarter by Year	To determine which year based on every quarter has the highest bike sales in 4 years from 2011 to 2015.
2.	Country	To identify which country is the highest percentage in purchasing the products.
3.	Gender and Age Group	To identify which gender that group by age group has the highest number to purchase the products.
4.	State	To study which state has the most buyers purchasing the products.
5.	Bikes Item	To identify what product category and sub category are the most purchased from the buyers.
6.	Bike Sales by Sub Category	To determine which sub category has the highest percentage sales group by product category.
7.	Transaction by Year measured by Euro	To identify the sales transaction of bike sales according to revenue, cost and profit in 4 year from 2011 to 2015.

Table 7: Title of Chart and Purpose

7.2 DATA STORIES

Chart 1 (a): Year

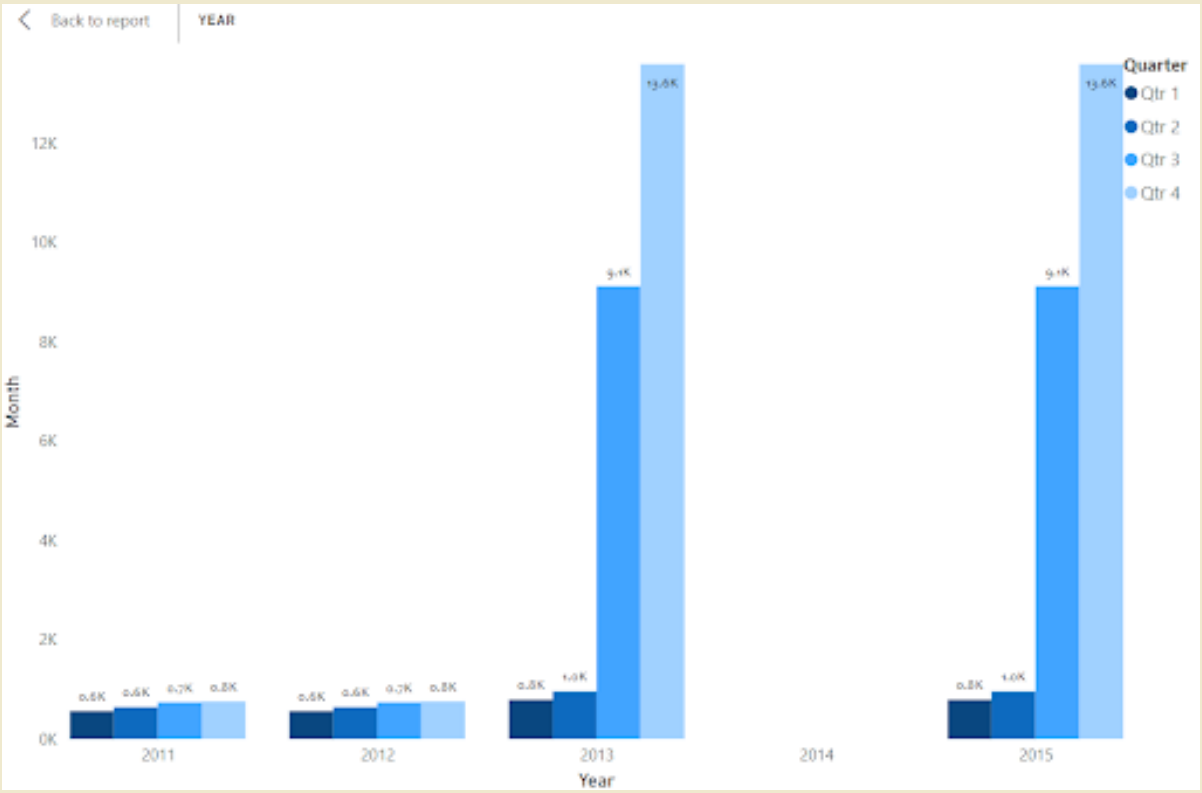


Figure 12: A Clustered Column of Customer by Year

Figure 12 represents a clustered column of customers purchasing bikes by year from 2011 until 2015. Each year can be classified into four quarters and being as a vertical axis while horizontal axis consists of the number of customers for every month.

Based on the chart, it showed an increasing pattern in Quarter 1 to Quarter 4 for the year 2011 until 2015. The company has the highest number of customers during Quarter 4 for every year, for instance there are 13590 customers in 2013 and 2015 during Quarter 4, and 752 customers during Quarter 4 in 2011 and 2012. Year 2013 and 2015 are the year where the company has the highest number of customers with 24443 customers.

As evidence, the number of customers in 2013 and 2015 are 9115, 950 and 788 customers, besides there are 732, 635 and 558 customers in 2011 and 2012 for Quarter 3 down to Quarter 1. Besides, Quarter 1 comes as the lowest number of customers for each year followed by Quarter 2 and Quarter 3. The total of customers in Quarter 3 consists of 19694 customers, where it is higher compared to 3170 and 2692 customers in Quarter 2 and Quarter 1. On the other hand, 2013 and 2015 are the year that had the highest number of customers with 24443 customers compared to 2677 customers for 2011 and 2012.

In conclusion, the company had most of the customers during Quarter 4 and the least in Quarter 1 for every year. Therefore, the company should always maintain good quality of the products due to the volatility in demand.

Chart 1 (b): Country

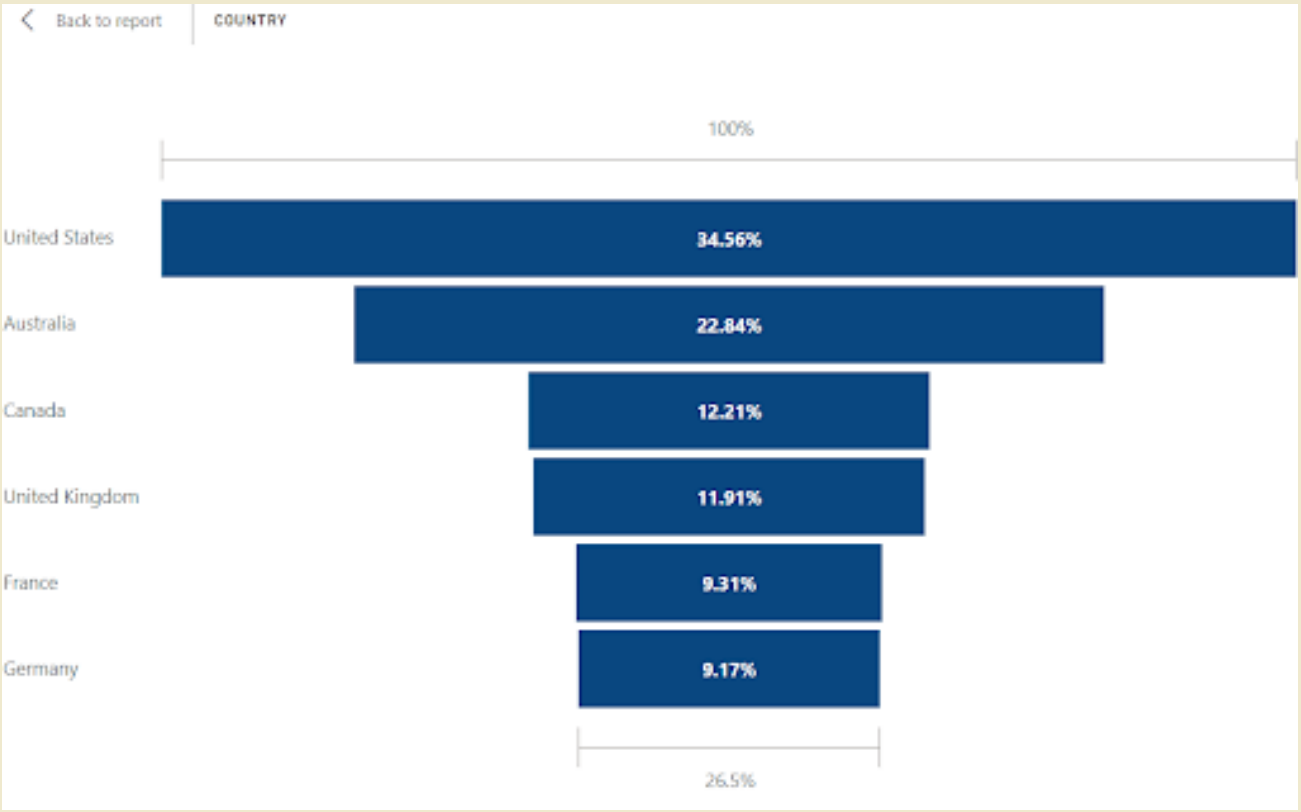


Figure 13: A Pyramid Chart of Customer by Country

The chart above illustrates the number of customers by country. The country consists of six countries in Europe which are Australia, Canada, France, Germany, United Kingdom and United States.

It shows a decreasing pattern by referring to the chart, starting from the United States and Australia, followed by Canada, United Kingdom, France and Germany. As stated, most of the customers are from the United States, hence that country takes first place in contributing in the number of customers with 34.56 percent of customers among countries. In contrast, Germany took last place with 9.17 percent of customers.

In other words, the United States consists of 18744 customers, while there are 12388 customers in Australia where it is contributing about 22.84 percent in number of customers. Next, Canada comes as the third country that consists of most of the customers about 12.21 percent with 6624 customers followed by the United Kingdom with 6462 customers about 11.91 percent in number of customers. France contributes about 9.31 percent and makes it to 5048 customers, while Germany is in last place among countries with 4974 customers and about 9.17 percent in number of customers.

In conclusion, the company must be more focused in the United States since it has the highest number of customers and hence can gain more profit from it. Other than that, the company also has to define the reason and problem that arise due to the least number of the customers in Germany.

Chart 1 (c): Gender and Age Group

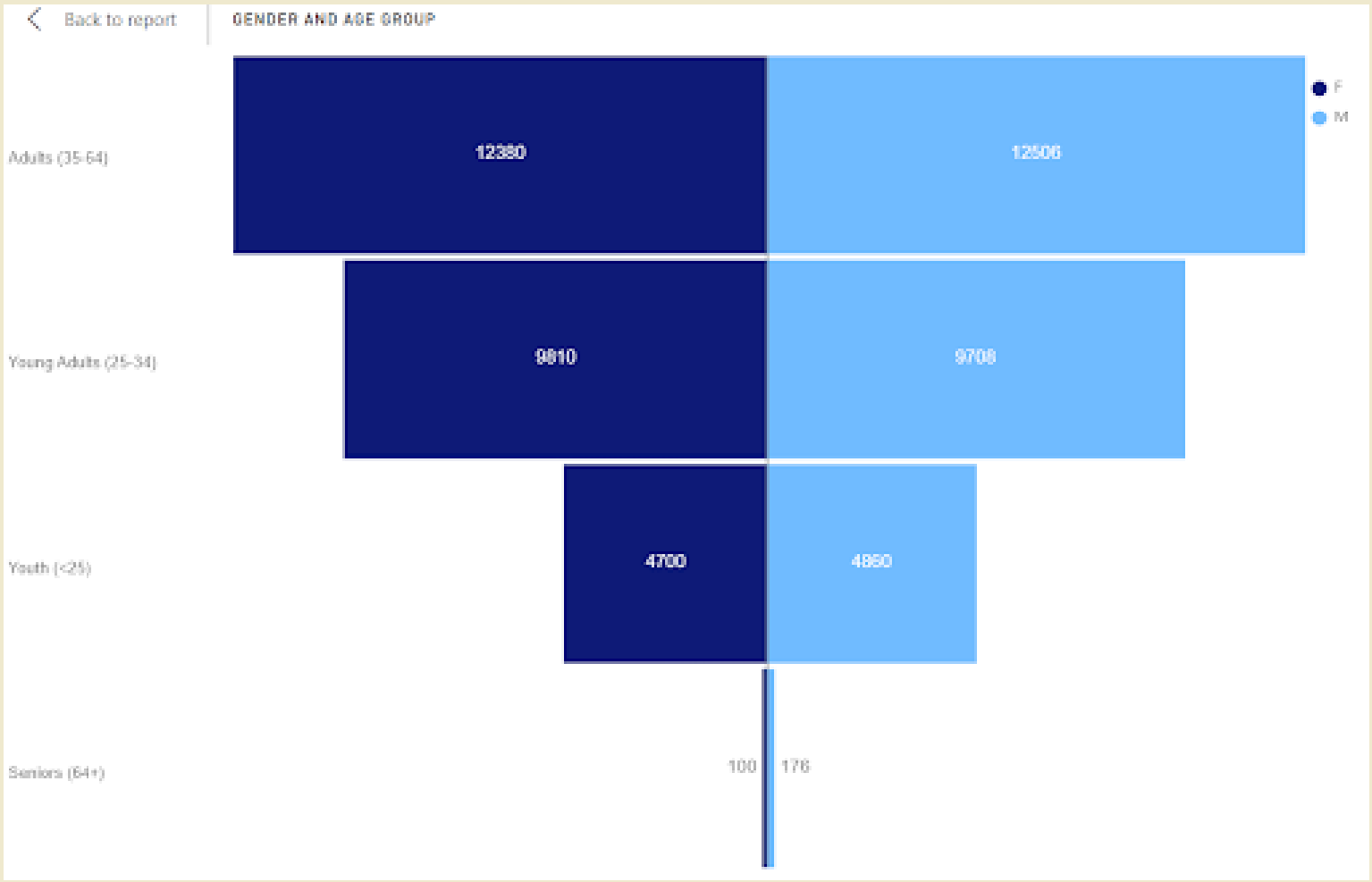


Figure 14: A Pyramid Chart of Customer by Gender and Age Group

The pyramid chart above shows the number of customer by gender and age group, where the gender can be classified into male and female, while age group can be categorized into four categories which are adults (35-64 years old), young adults (25-34 years old), youth (less than 25 years old) and seniors (above 64 years old).

Based on the chart, gender among male appears as the highest customer in purchasing bikes with 27250 customers in total compared to females. On the other hand, the age group among adults (35-64 years old) comes in first place in purchasing bikes with 24886 customers in total compared to other categories. In contrast, there are only 276 customers in total among seniors whose age group is more than seniors (64 years old) and make it appear as the lowest in purchasing bikes. It shows a declining pattern of combination gender and age group of the customer, where adults (35-64 years old) customers have the highest purchasing bikes for both genders with 12380 females and 12506 males.

Customer gender among males contributes 50.24 percent while female customers have 26990 customers and makes it 49.76 percent in the total of the customer. For customer according to age group, there are about 45.88 percent of the customer comes from among adults (35-64 years old), 35.98 percent from young adults (25-34 years old), 17.63 percent from youth (less than 25 years old) and only 0.51 percent comes from seniors (more than 64 years old). The result of a combination gender and age group, adults (35-64 years old) customers takes the first place in purchasing bikes followed by young adults (25-34 years old) with 9810 females and 9708 males, and youth (less than 25 years old) with 4700 females and 4860 males. Seniors (more than 64 years old) have the least in total number of customers where there are 100 females and 176 males.

In conclusion, most of the customers are among male customers with a 0.48 percent more compared to female customers. Other than that, customers among adults (35-64 years old) have contributed the most in bike sales with 45.37 percent more compared to seniors (more than 64 years old). Male customers also are the most influential in each age group which are among adults (35-64 years old), seniors (more than 64 years old) and youth (less than 25 years old).

Chart 1 (d): State

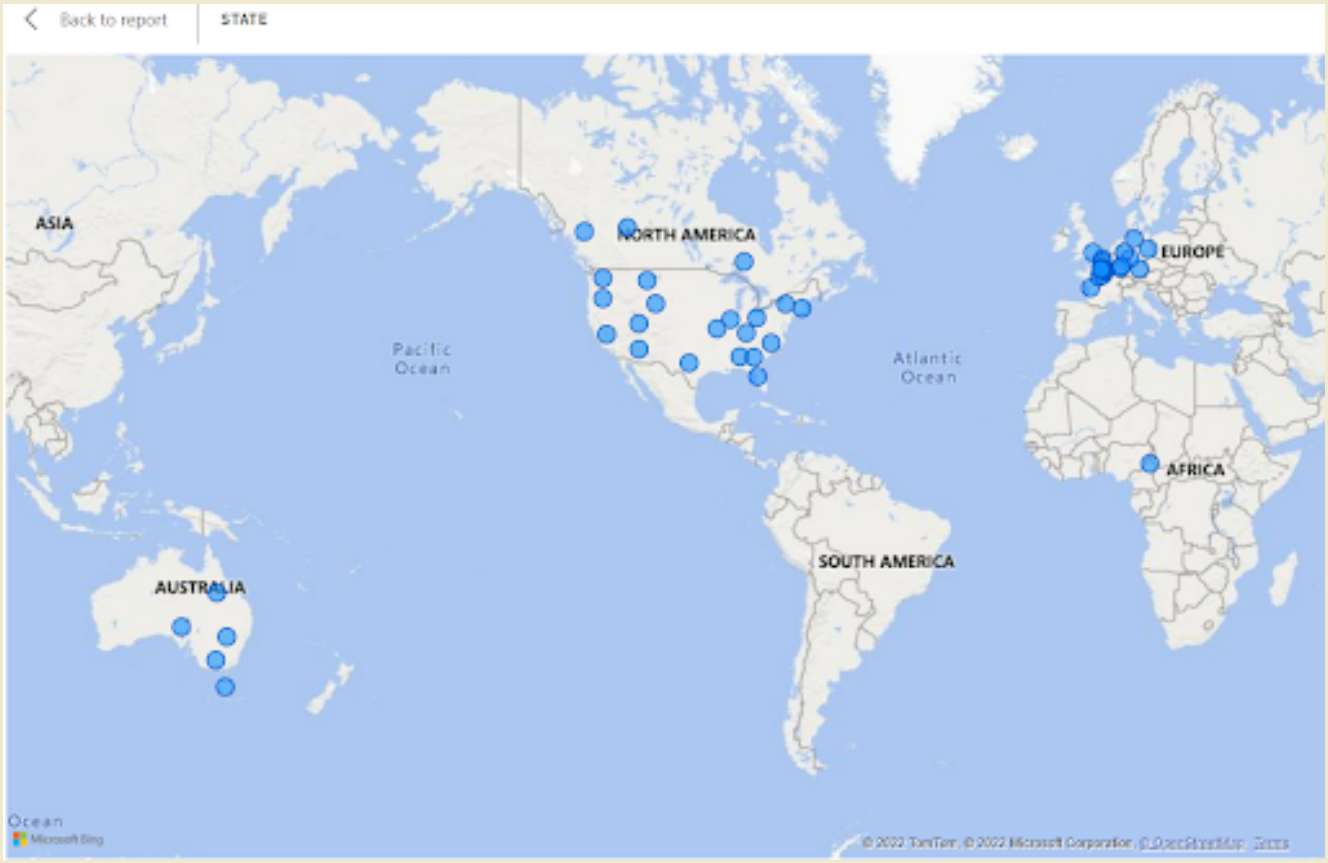


Figure 15: A Map of Customer by State

The map shows the number of customers by state, where the state has been grouped into continents. There are about 49 states involved and continents are used as the references.

Based on the map, most of the blue dots gathered on the North America continent, where blue dots represent the customer. North America, which consists of 21 states, comes in as the highest in total number of customers with 25368 customers. California state that also on the North America continent takes the first place for the highest number of customers within states with 10846 customers. The least blue dots are detected on the Australia continent with only 12388 in total.

Despite California being the highest number of customers within states, it only contributes about 0.43 percent on the North America continent and about 20 percent in the United States country itself. Besides, England has the highest number of customers on the Europe continent with 6462 customers and contributes about 11.91 percent to the United Kingdom country. To the Europe continent, it contributes about 0.39 percent, 22.84 percent is contributed from Australia as it is the only one from their own continent.

In conclusion, the least number of states in each continent absolutely can affect the number of customers other than sales. The company can focus on developing their business to other states and continents in future.

Chart 2 (a): Bikes Item

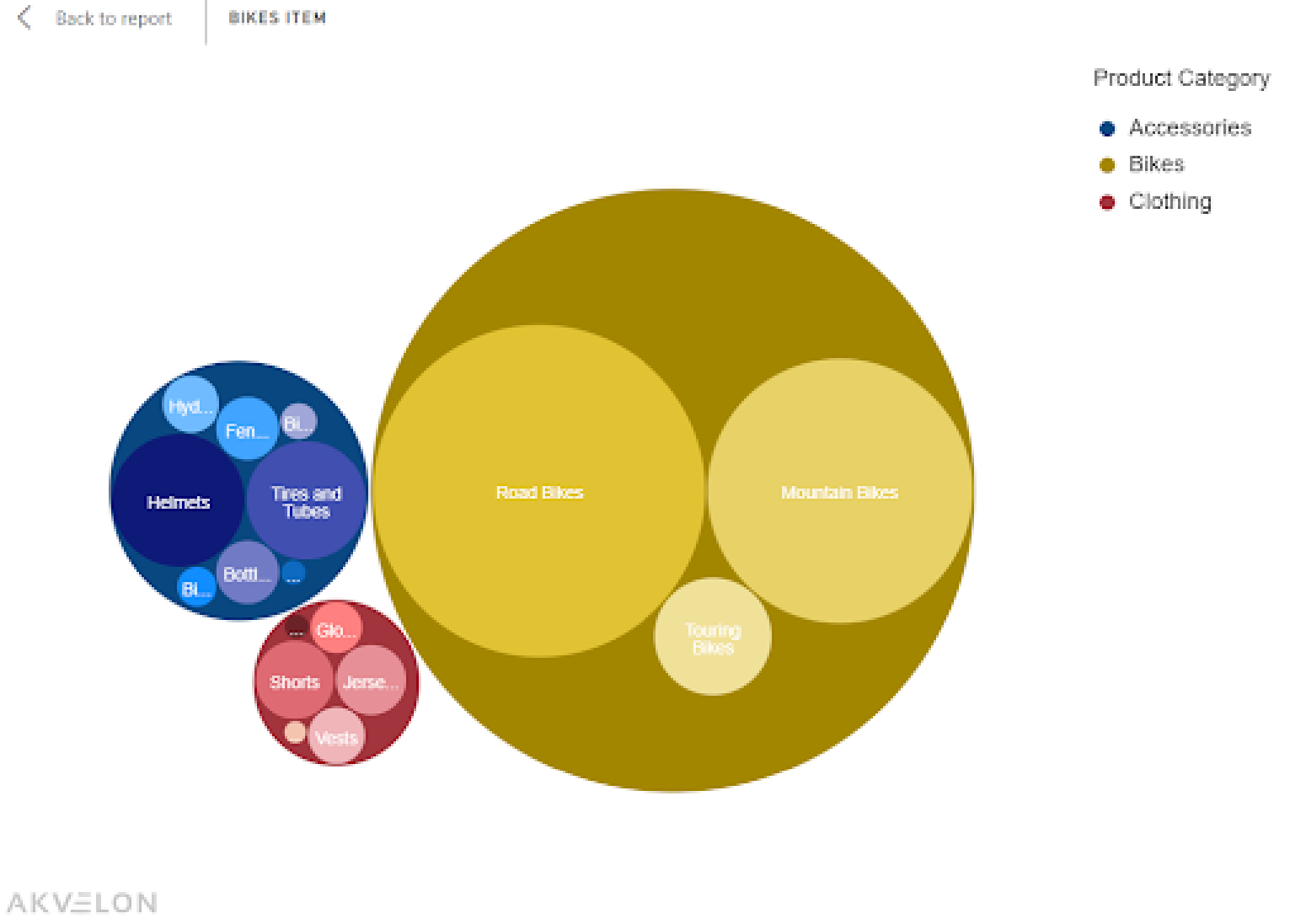


Figure 16: A Bubble Chart of Bikes Item

Bubble chart shows bikes items of bike sales in Europe from 2011 to 2015. There are three product categories that consist of accessories, bikes and clothing. For each product category, it consists of a sub category for every item. For accessories, it consists of 8 sub category which are bike racks, bike stands, bottles and cages, cleaners, fenders, helmets, hydration packs and tires and tubes. For bikes, there are only three sub category which are mountain bikes, road bikes and touring bikes. Lastly for clothing, it consists of 6 sub category which are caps, gloves, jerseys, shorts, socks and vests.

As we can see bikes are the highest product category sold since the bubble chart is biggest compared to the others which are accessories and clothing. In bikes, road bikes are the highest sub category sold with 9718 sold followed by mountain bikes with 5406 sold and the least sold in bikes is touring bikes with only 1650 sold. For the second most sold product category which is accessories, tires and tubes is the highest sub category sold with 14558 sold. Followed by bottles and cages with 6546 sold, helmets with 5292 sold, fenders with 1766 sold, cleaners with 766 sold, hydration packs with 610 sold, bike racks with 274 sold and the least is bike stands with only 238 sold. Lastly, in clothing, jerseys are the highest sub category sold with 2708 sold. Followed by caps with 1770 sold, gloves with 1162 sold, shorts with 870 sold, socks with 496 sold and the least is vests with only 410 sold.

For road bikes, it contributes 57.93 percent sales for bikes compared to mountain bikes and touring bikes which only contribute 32.23 percent and 9.84 percent respectively. Following that, 48.45 percent of tires and tubes contribute to the sales of accessories followed by bottles and cages with 21.78 percent. While for helmets and fenders, it contributes 17.61 percent and 5.88 percent respectively. For the other 4 sub category of accessories, it has the least percentage of sold with only 2.55 percent, 2.03 percent, 0.91 percent and 0.79 which are cleaners, hydration packs, bike racks and bike stands respectively. Last but not least for clothing which contributes the least sold of bike items, it contributes only 36.52 percent of jerseys followed by caps with 23.87 percent, gloves with 15.67 percent, shorts with 11.73 percent, socks with 6.69 percent and vests with5.53 percent.

Therefore, it can be concluded that bikes are the highest product category sold with road bikes the most sold while touring bikes are the least. Following that is accessories with tires and tubes is the most sold while the least is bike stands and the least product category sold which is clothing with jerseys is the most sold compared to vests which is the least sold.

Chart 2 (b): Bike Sales by Sub Category

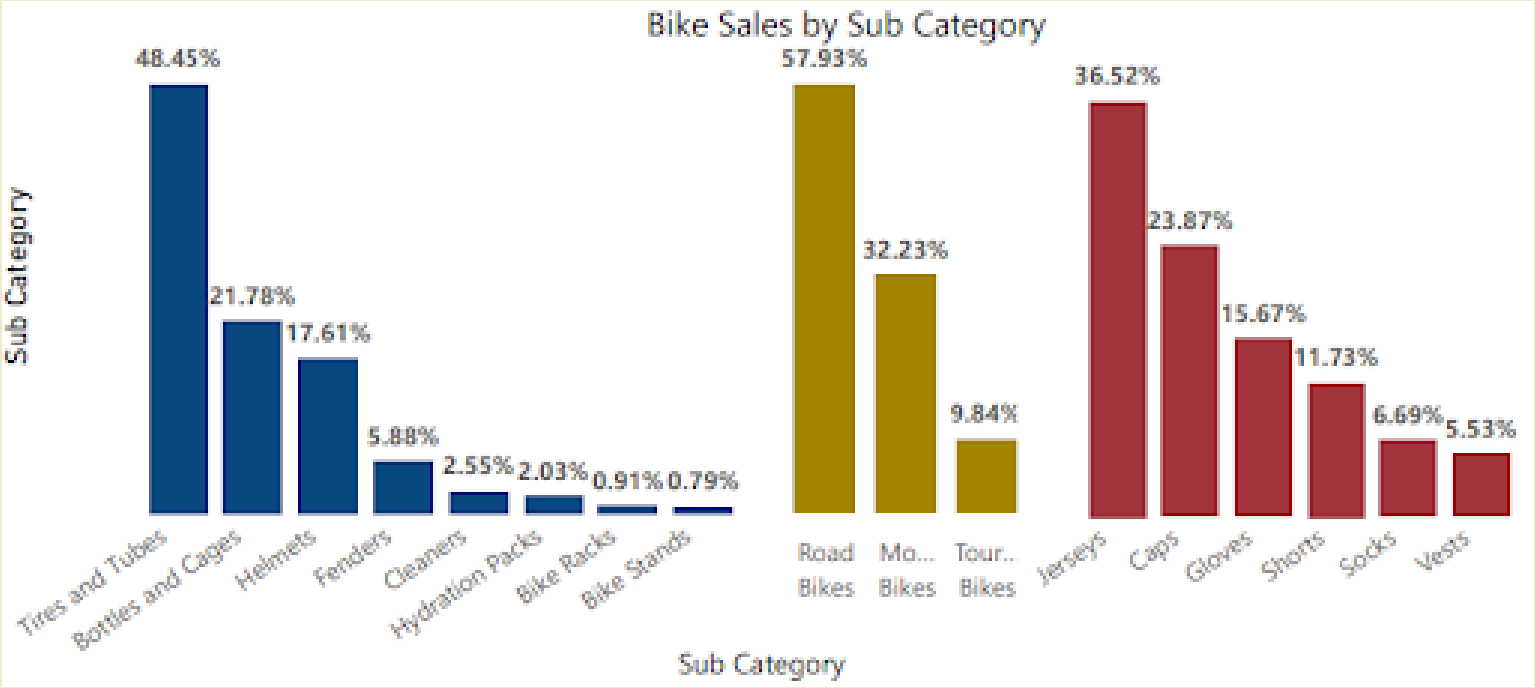


Figure 17: A Bar Chart of Bike Sales by Sub Category

Bar chart shows the percentage bike sales by sub category group within product category of bike sales in Europe from 2011 to 2015. Tires and tubes, bottles and cages, helmets, fenders, cleaners, hydration packs, bike racks and bike stands are items in the product category of accessories. While for road bikes, mountain bikes and tour bikes are items in sub category of product category in bikes and lastly, jerseys, caps, gloves, shorts, socks and vests are items in sub category for product category of clothing.

It shows tires and tubes are the highest percentage sold for accessories with almost half of that product category which is accessories followed by bottles and cages is the second, third is helmets, fenders is the fourth and followed by cleaners, hydration packs, bike racks and bike stands respectively. For bikes, it shows road bikes have the highest percentage that more than half of the sales of bikes in the product category sold followed by mountain bikes and tour bikes. Next for clothing, even though none of the sub category sold more than half, it shows that jerseys is the highest percentage sold in clothing followed by caps, gloves, shorts, socks and the least percentage sold is vests.

As we can see for accessories, 48.45 percent of tires and tubes sold followed by 21.78 percent of bottles and cages, 17.61 percent of helmets and 5.88 percent of fenders. However, there are 4 sub category of accessories that have the least percentage sold which are cleaners, hydration packs, bike racks and bike stands with only 2.55 percent, 2.03 percent, 0.91 percent and 0.79 percent respectively. For bikes, it shows that more than half of bike sales of product category bikes contributed by road bikes with 57.93 percent. Followed by mountain bikes with 32.23 percent and touring bikes with 9.84 percent. Last but not least, as mentioned that there are no sub category of clothing sold more than half of the percentage, jerseys has the highest percentage sold in clothing with 36.52 percent followed by 23.87 percent of caps, 15.67 percent of gloves, 11.73 percent of shorts, while this two are the least contribute in sales of clothing which only 6.69 percent and 5.53 percent which are socks and vests respectively.

In conclusion, tires and tubes, road bikes and jerseys are the highest percentage sold for every product category which are accessories, bikes and clothing respectively. Nevertheless, the least percentage of bike sales are bike stands, touring bikes and vests which are in the product category of accessories, bikes and clothing respectively.

Chart 3: Transaction by Year

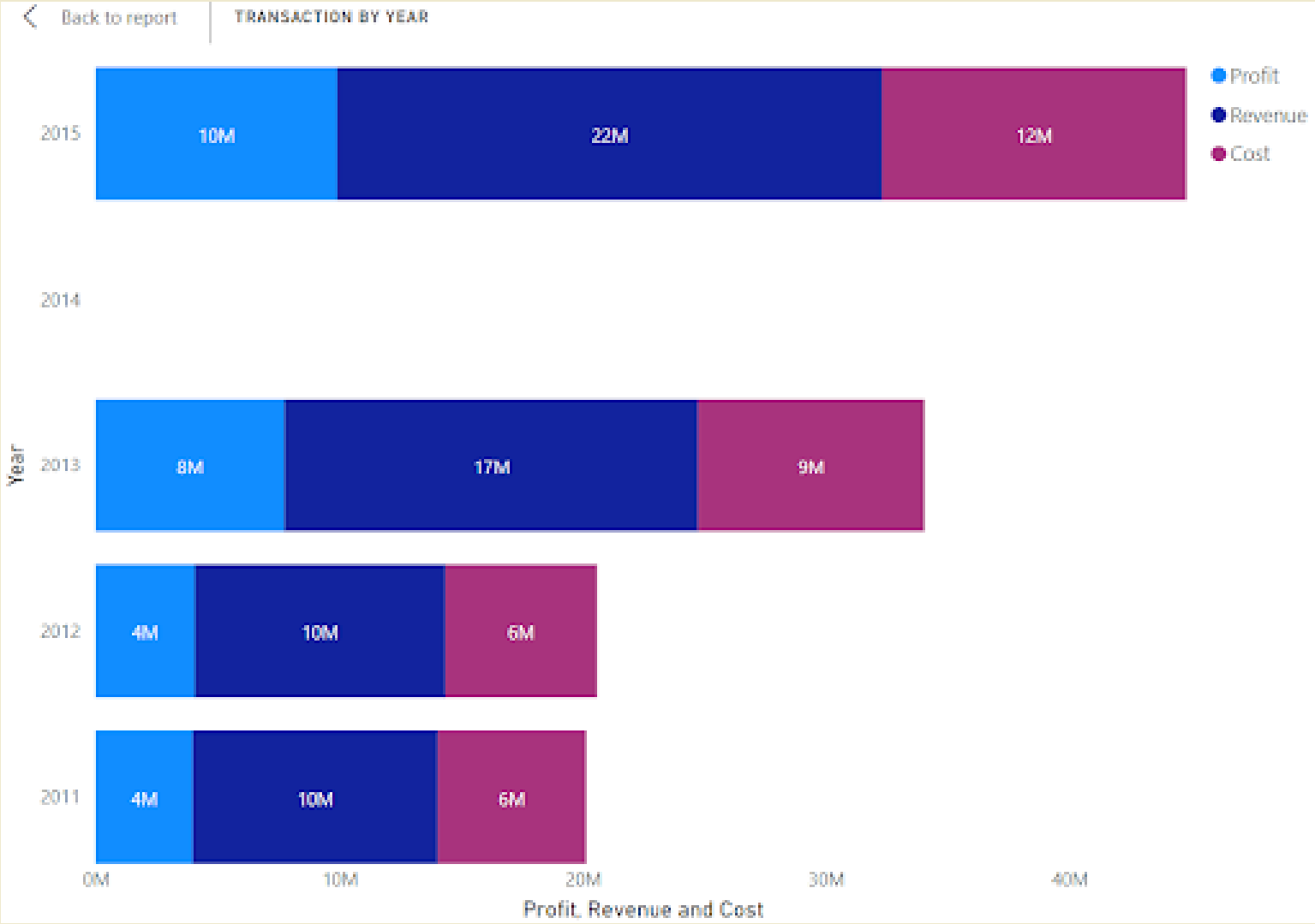


Figure 18: A Stacked Bar Chart of Customer Transaction by Year

The stacked bar chart above shows the customer transaction by year in European countries from 2011 until 2015. Based on the figure, it shows the three categories of the transaction to know the relationship between customer profile, product, and sales of bikes based on monthly and year of bike sales. The three categories are inclusive of cost, profit, and revenue.

Based on this chart, it shows the number of transactions of products sold by the customer in Europe by year. In the first and second year bike sales in Europe got 4 Million, 10 Million and 6 Million of profit, revenue, and cost respectively. There is increasing trend of the total trade of cost, profit and revenue from year 2012 to 2015 where the value increases from 4 Million to 10 Million for profit of bike sales also increase in the revenue with 10 Million to 22 Million. Same goes to the cost of bikes sales in Europe there are also shows the increasing trend with 6 Million to 12 Million.

However, European countries had the highest cost, profit, and revenue in 2015 with about 10 Million of cost, 12 Million of profit, and 22 Million of revenue. Hence, 2011 and 2012 had the lowest of total bikes sales since it was the first and second years of opening a bikes shop. From the chart, it can be seen that the total sales of bikes increased from 2012 to 2015. Europe can achieve a total sales of more than 22 Million in 2015 if companies in six countries in Europe maintain the strategies to increase their sales. Other than that, the chart indicates that the highest number of customers who purchase the bike's product can make the highest sales in companies. So, to get the highest of sales it depends on the total of the customer. As a result, shown in the chart, it can be concluded that the highest transaction of customer purchase was in 2015 since it was the highest of the total cost, profit, and revenue.

In conclusion, the stacked bar chart used to show a trend over time sales patterns in Europe in four years which are in 2011, 2012, 2013, and 2015. The trends show flat sales in 2011 and 2012 since there is the same total of cost, profit, and revenue and started to rise after 2012 to 2015. From the stacked bar chart, Europe has the highest cost, profit and revenue in 2015 with about 12 Million of cost, 10 Million of profit and 22 Million of revenue. In the chart of the transaction purchase of customers are generally increasing over time since there is no trend showing a declining pattern.



8.0 CONCLUSION

In conclusion, all the project's major objectives were met with remarkable success. This project is to achieve the three objectives which are to identify the customers profile by age group, gender, quarter, year, country and state to study the trend of bike sales product according to product category and sub category and lastly to analyse the relationship between customer, product, profit and revenue. Most of the customers of bike sales in Europe are among male where males contribute 50.24 percent while female customers have 26990 customers and make it 49.76 percent in the total of the customer. However, every year, the company has the most customers in Quarter 4 where there are 13590 customers in Quarter 4 in 2013 and 2015, and 752 customers in Quarter 4 in 2011 and 2012. The company must be more focused in the United States since it has the highest number of customers that contribute among the Europe countries with about 34.56 percent and hence can gain more profit from it. Because of the fluctuation in demand, the company must always maintain good product quality.

For the product category for the bike sales in Europe for six countries, we can see that the highest product category sold is road bikes, contributing 57.93 percent sales for bikes compared to another category. The main purpose of this project is to analyze the bike sales in Europe countries from 2011 until 2015 with the measure in Euro for the total cost, profit, and revenue. From 2011 to 2015, the trade value for cost, profit, and revenue has increased from 4 Million to 10 Million for profit, 10 Million to 22 Million for revenue, and 6 Million to 12 Million for cost. As a result, the largest cost, profit, and revenue for 4 years is in 2015 with about 12 Million of cost, 10 Million of profit, and 22 Million of revenue. Europe can achieve a total sales of more than 22 Million in 2015 if companies in six countries in Europe maintain the strategies to increase their sales. Other than that, the stacked bar chart indicates that the highest number of customers who purchase the bike's product can make the highest sales in companies. So, to get the highest of sales it depends on the total of the customer.





ACCESSORIES.BIKES.CLOTHING

PALACE DE BIKES

PALACE DE BIKES

Email: palacedebikes@gmail.com
Phone Number: +603-8423-5383

