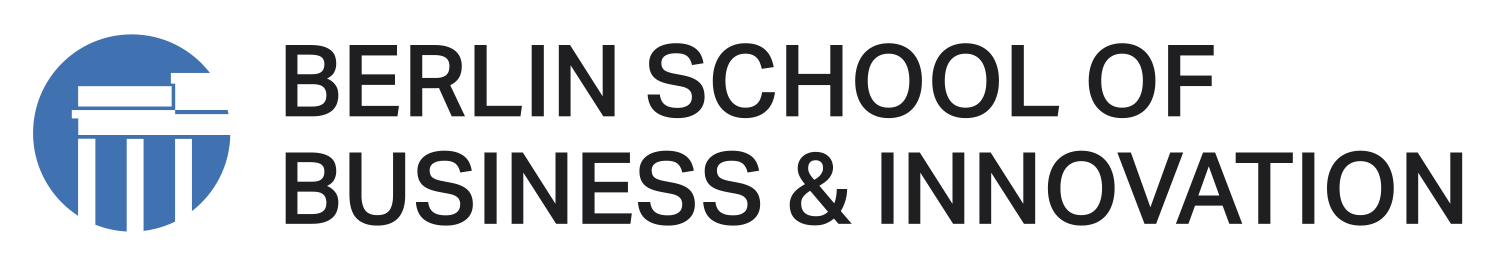
Shape

Description automatically generated with medium confidence



**Essay / Assignment Title:** Designing a sales Tableau Dashboard for a chosen company

**Programme title:MSc of Data Analytics**

**Name:Turan Can Gün (Q1055616)**

**Year:2024 (2023 October Intake)**

CONTENTS

[INTRODUCTION 3](#_Toc464502685)

[CHAPTER ONE EXPLAINING THE PROBLEM and THE DATASET 4](#_Toc109639966)

[1.1 Specify of The Problem and Explanation of The Dataset 5](#_Toc1975699163)

[1.2 Strenghts and Constraints 6](#_Toc1214526315)

[CHAPTER TWO DEFINING ATTRIBUTES of Tableau 7](#_Toc1088578452)

[CHAPTER THREE IMPLEMENTATION 9](#_Toc734176251)

[CHAPTER FOUR DASHBOARD DESIGN and ANALYSIS 15](#_Toc1974543276)

[CONCLUDING REMARKS and DISCUSSIONS 19](#_Toc1801133964)

[BIBLIOGRAPHY 20](#_Toc1369196149)

[APPENDIX (if necessary) 21](#_Toc1715646118)

**Statement of compliance with academic ethics and the avoidance of plagiarism**

I honestly declare that this dissertation is entirely my own work and none of its part has been copied from printed or electronic sources, translated from foreign sources and reproduced from essays of other researchers or students. Wherever I have been based on ideas or other people texts I clearly declare it through the good use of references following academic ethics.

(In the case that is proved that part of the essay does not constitute an original work, but a copy of an already published essay or from another source, the student will be expelled permanently from the postgraduate program).

Name and Surname (Capital letters): Turan Can Gün

.......................................................................................................................................

Date: 21/10/2024

# INTRODUCTION

Problems with statistical analyses were well documented in the scientific literature ( Weissgerber , T. , 2021).Then when topic comes into trading at international level according to the International Monetary Fund (IMF) is expected to global trade in 2024-2025 annually also global online traidng platform market size will be reach to 10.15 billion USD end of year 2024.

About technological advancements in online trading in terms of financial situation this can be example over %85 of global forex market transactions happen on automated traidng system.This system which was built conceives all these currency pairs and analyzes them best. During the development and testing work vast majority of time the system was run on these pairs and showed positive results ( Roskladka , A. ; Baiev R. , 2021 ).

Data analysis tools such as PowerBI and Tableau have critical roles in that scenario to handle tihs vast amount pure data to extract meaningful insights for business as well as customers.Moreover in this study a sample scenario is handled to demonstrate implementation of Tableau.

# CHAPTER ONE EXPLAINING THE PROBLEM and THE DATASET

## 1.1 Specify of The Problem and Explanation of The Dataset

Tthe scenario of this study there is an online commercial company and they have been got a request from an online shopping platform to create special advertisements based on some product categories at the same time this platform is specailzed based on bicycles and byproducts.After the request the platform has been provided required dataset regarding to this aim then about the dataset that can be described in the dataset products are able to categorized across 3 different parts as accessories , bikes and clothing.Moreover these3 categories are divided into itself sub products.For instance under accessories category there are helmets , tires ,bike racks so on.Then under bikes category there are mountain – touring – road types.About clothes there are gloves shorts etc.

The dataset for this scenario is obtianed from Kaggle datasource and which is created by Jehanzaib Bhatti.There are key features about the dataset like temporal (details of date) , product level information and customer demographics and further they will be explained with more details by visualizaitons.

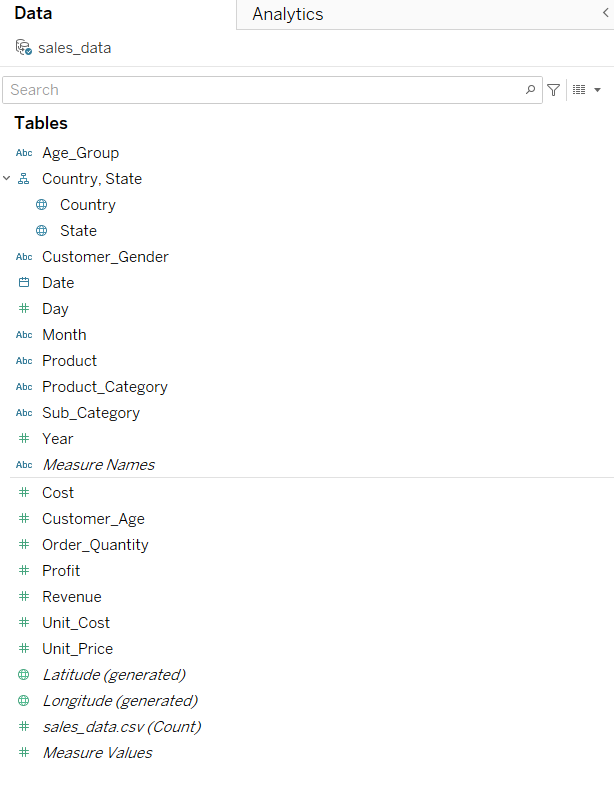


Figure 1.1 Features of The Dataset

Figure 1.1 indicates all categorical and numerical values which belongs the dataset also we can make categorize as table ;

|  |  |  |
| --- | --- | --- |
| **Categorical** | **Numerical** | **Visually** |
| Age Group | Cost | Latitude |
| Country-State | Customer Age | Longitude |
| Customer Gender | Order Quantity |  |
| Date | Profit |  |
| Month | Revenue |  |
| Product | Unit Cost |  |
| Product Category | Unit Price |  |
| Sub Category | Sales \_data.csv(count) |  |
| Measure Names | Measure Values |  |
|  | Year |  |
|  | Day |  |

Table 1.1 Types of Feature of the Dataset

To explain the features might be benefitical for further steps.Under categorical variables , age group claasifies customers based on young – youth – adult – senior parts.Country - state feature consists UK-Canada-USA-Germany-France-Australia and their states.Product segment (product category – sub category) that includes accessories , bikes and clothing with their subproducts.

Under numerical variables unit cost and cost feaatures are related production expanditure of goods.Order quantity indicates number of ordered products,revenue is about income also profit is able to demonstrate net gain from selling.

## 1.2 Strenghts and Constraints

When working with any dataset understanding limitation and advantages is very crucial to obtain meaningful insights from the data.The dataset which is used in this study has some various advantages.For instance cıstomer demography can be invested from many aspects such as age category , customer gender as well their nationality as a result of this advantage customer behavior or purchase behavior can be extracted from there.

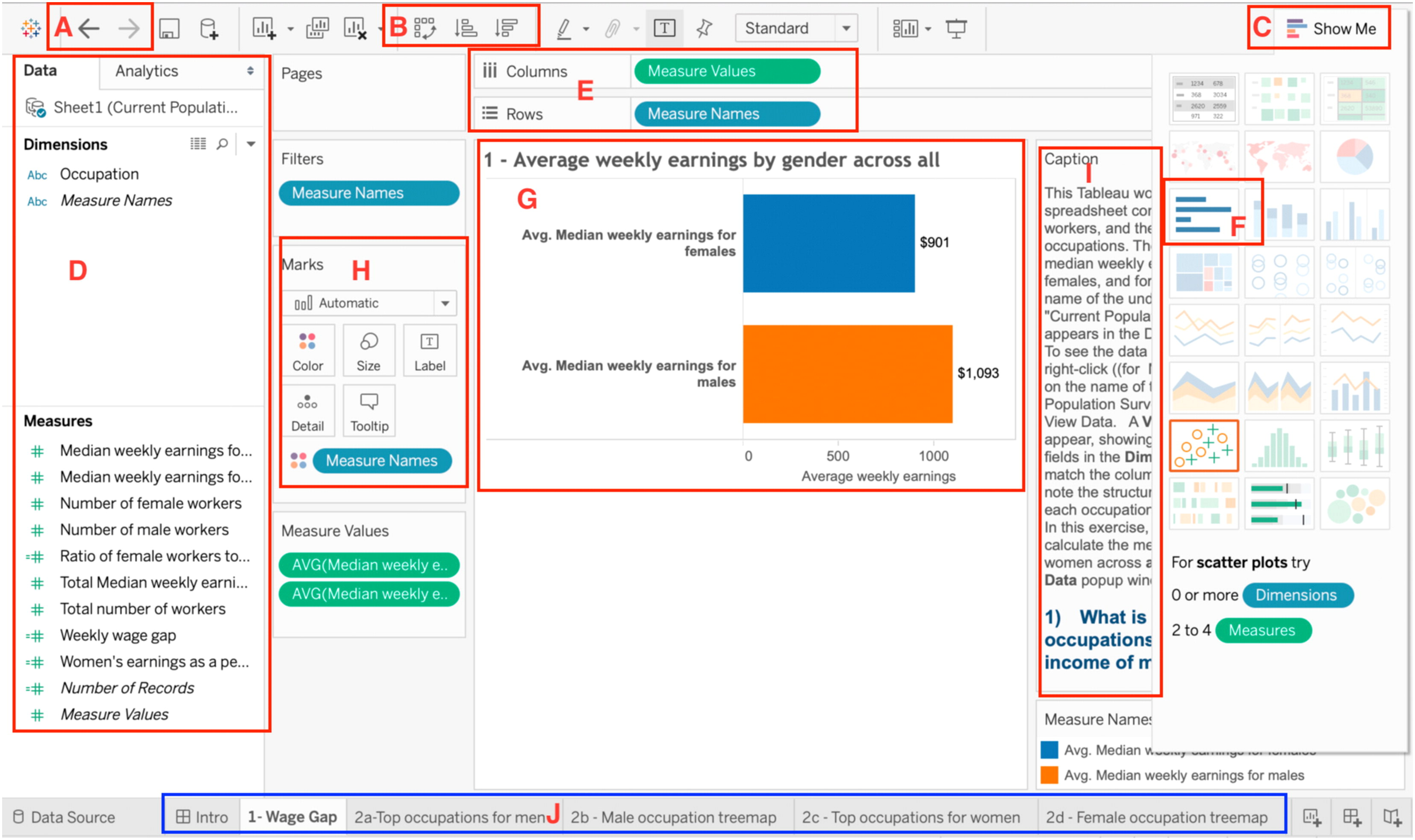
Another advantage is of the dataset related to profit as an example from features which are cost,revenue,unit cost and so on the dataset is able to provide meaningful inference to maximize customer satisfaction as well as annual profit of the online trading platform in our scenario.

Potential impact of missing or inconnsistent points in a dataset especially if this situation occurs in customer details such as demographics or purchase details which leads to prevent to get brightful insgihts.However in this study this affect is unlikley to emerge as the data has enough knowledge regarding to customers.Moreover about the general constraints there is no any impactful either incosistent or missing points.

# CHAPTER TWO DEFINING ATTRIBUTES of Tableau

Tableau is free visualizaiton tool which is used in business sectors especially by data analysts ( Steven B. , 2020).Also it is more popular then PowerBI since Tableau is able handle more large scale data relatively to PowerBI.

In Tableau there are mainly 2 sections these are called as “ Sheets “ and “ Dashboards “. In sheaats you can create any visualizaitons like bar chart , graphs etc. based on dataset and preferences and dashboard can be described as main display of unify of sheets.

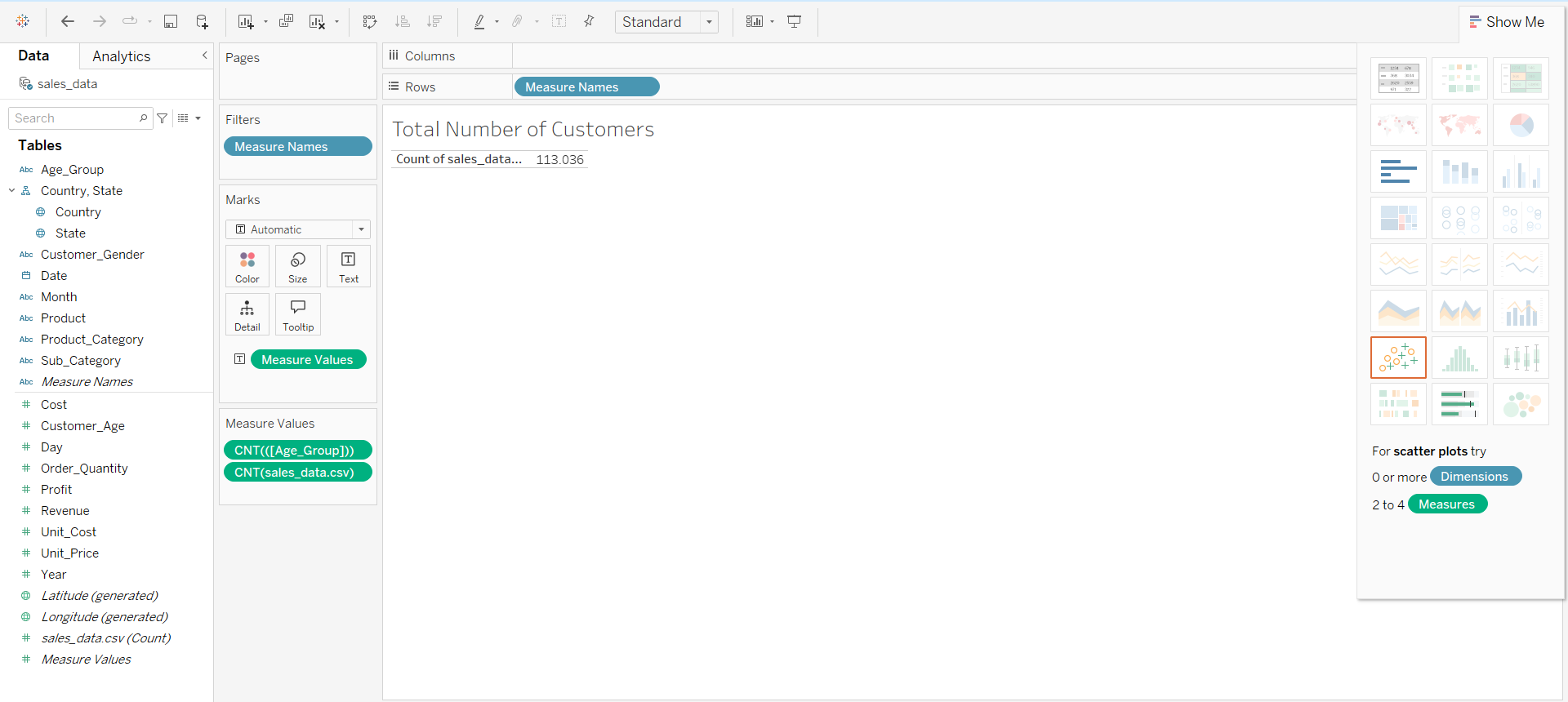
Table 1.2 Demonstration of Tableau Interface ( Bat , S. , 2020)

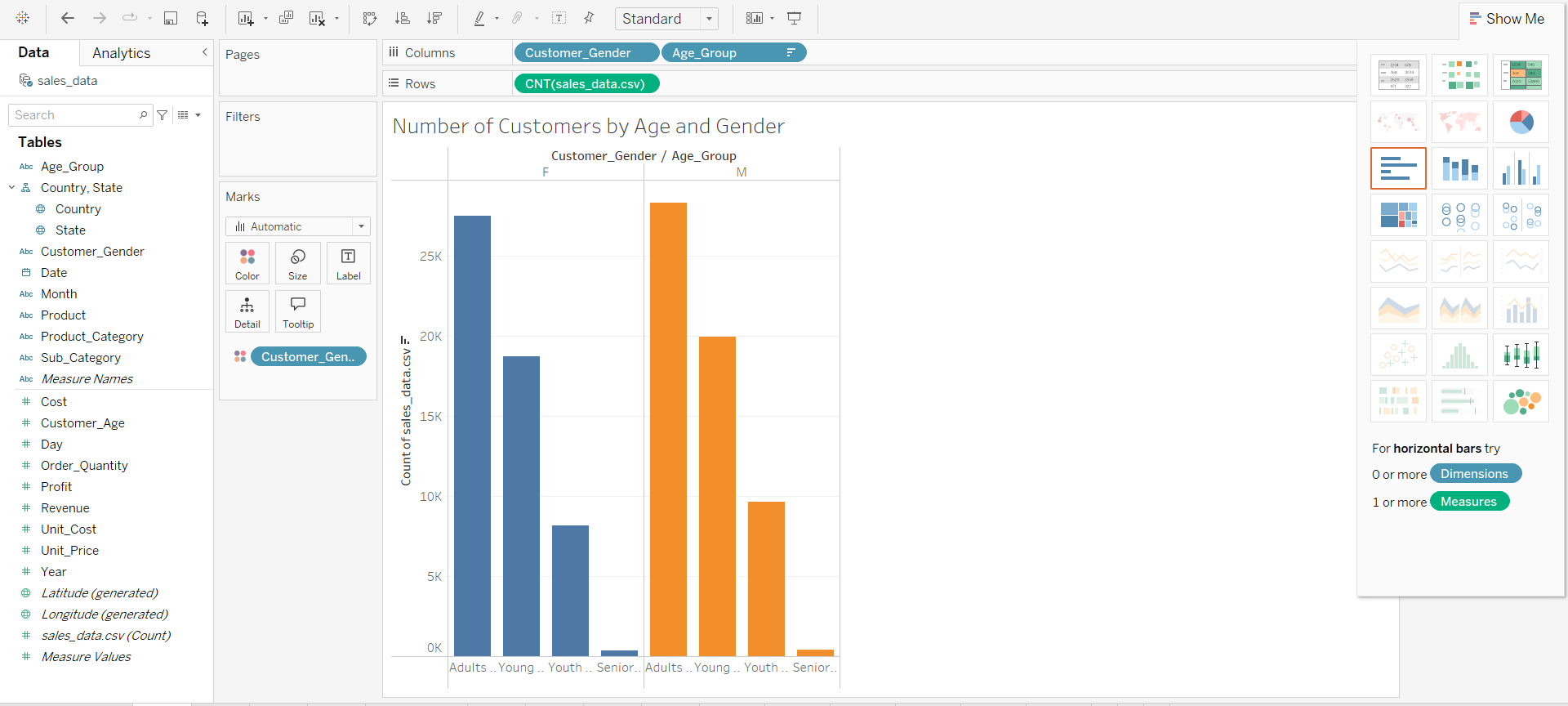
* In section A either undo or rendo can be realized and seciton B is about sorted of the dataset variables based on ascending or descending order.
* Section C is important related to select appropraite visualizaiton types according to used attributes in rows and columns.
* Section D might be useful to analyzing of attributes of the dataset also which indicates categorcial or numerical attributes.Section J is able switch betweeen sheets as weel as dashboards.
* Seciton E is used across substutition of attributes into rows and columns on the other side section H has properties to mark variables based on their attributes.For instance countries attribute can be color like USA red , UK orange etc. Another example when size mark used this make different radius circles of each variable under a certain attribute according to attribute size.
* Seciton G is display of sheets towards to selected attributes for row and column.
* Also calculated field is another crucial property in Tableau which is useful when calculate specific variables in dataset.
* In this scneario bar chart , line chart , map , bubble chart and text tables are used.

|  |  |
| --- | --- |
| Visual Element | Usage Area |
| Bar Chart | Comparision of cateogrical variables. |
| Line Chart | Variation of variables in a time interval |
| Bubble Chart | Comprasion of categories by different bubble size |
| Map | Illustration of geographical data |
| Text Tables | Standard table visualizaiton |

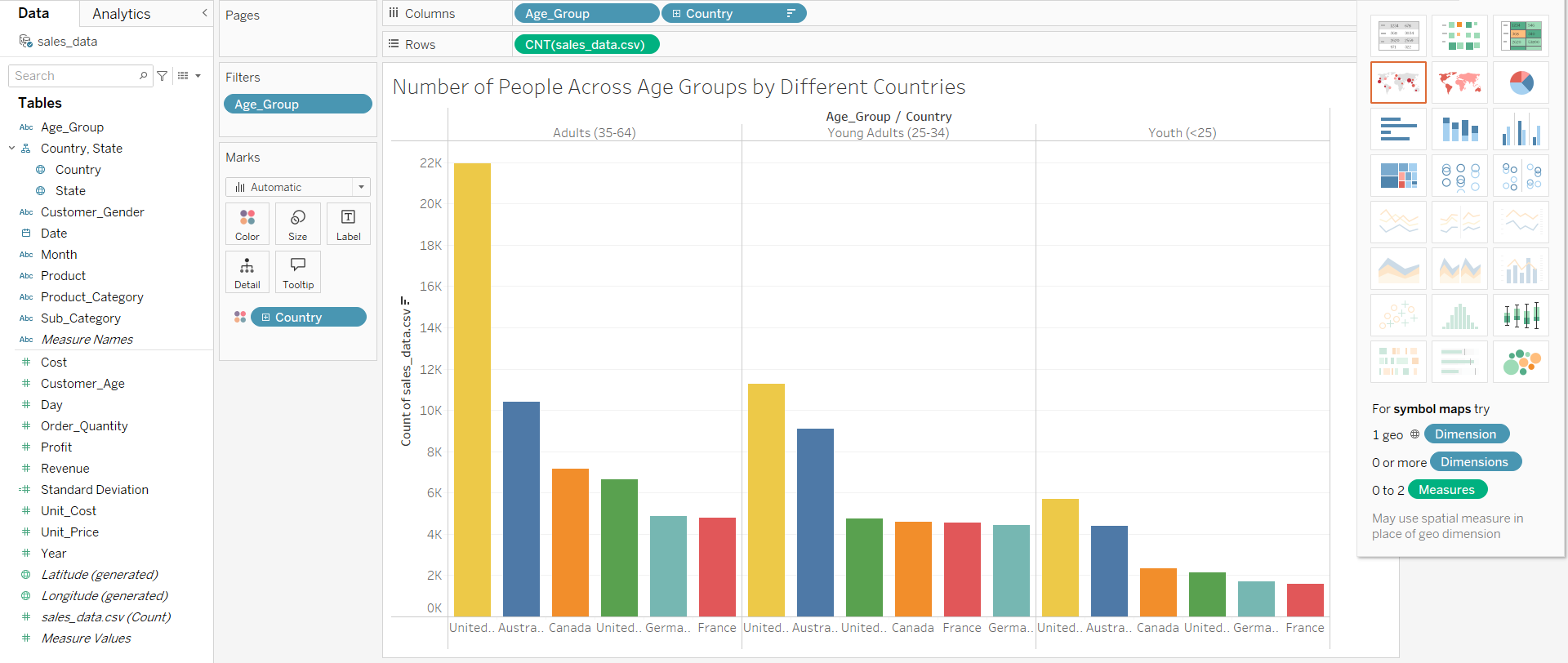
Table 1.3 Comparision of Visual Elements

# CHAPTER THREE IMPLEMENTATION



Figure 3.1 Visualazitaion of Number of Customers by Age and Gender

For implementation step analysing of customer demographics is first stage as Figure 3.1 indicates that total people in terms of customer for this company is 113.036 and the bar chart demosntarates number of customers based on their age as well as gender.To create bar chart in columns Customer\_Gender and Age\_Group attributes are used then in rows CNT attributes is used.Also mark feature is utilized to color the graph according to the gender.

Figure 3.2 Number of People Across Age Groups by Different Countries

Second in Fig 3.2 aims to exploring of number of people across age group then categorized by countries.To visualize that appropraite method is use to bar chart but there is an important point senior age group is filtered due to low number of people.Also the mark is enabled to differs countries from each other by colours.

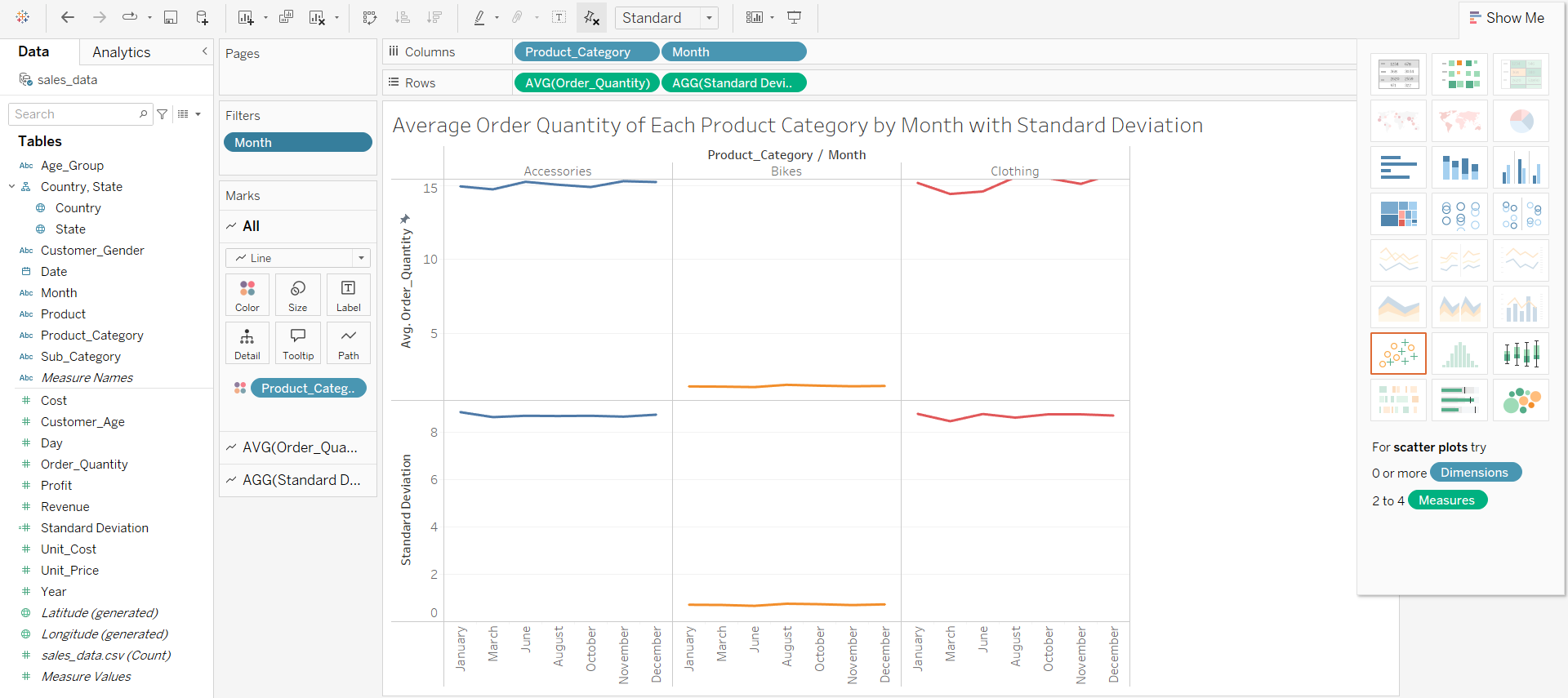
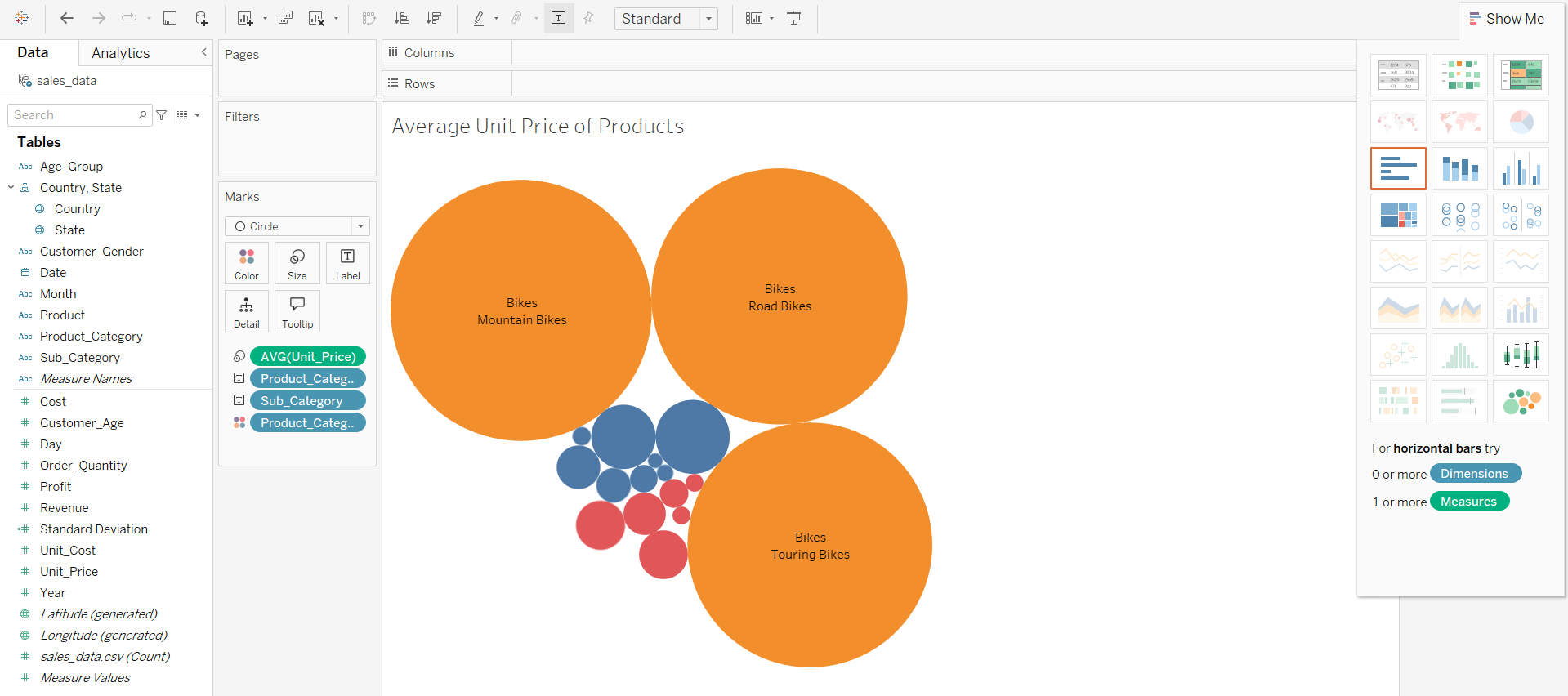


Figure 3.3 Average Order Quantity of Each Product Category by Month with Standard Deviation

Third step in Fig 3.3 computes that average order quantity by month with standard deviation.To implement this analysis in columns Product\_Category – Month attributes are used on the other side in rows there is Order\_Quantity with average function and standard deviation feature.Afterwards to utilize from standard deviation STDEV([Order\_Quantity]) calculation field is applied.Furthermore there is a filter to exclude some months (April – May – July \_ September) and still as previous implementation mark (colour) is enable to color for differs product category.

Fig 3.4 Share of Sum of Ordered Products

Fourth step in Fig 3.4 represents number of ordered products by their size and to visualize this statistic treemap is used.Moreover darker colour is an indciatior towards to higher number of order or bigger share within in all prodcuts.There are Product\_Category and Sub\_Category attributes to make this visualization.

Figure 3.5 Average Unit Price of Product Categories and Their Subcategory

Fifth step in Fig 3.5 implies unit price of product categories and their sub categories by usinn packed buubles.Actually this analysis was visualized by bar chart but trasnform into packed bubble makes it more effecitve to understand unit price difference especially between product category as well as sub categories.To realize that there are Product\_Category and Sub \_Category then there is AVG (Unit-Price).As everytime mark is enable to categorize product categories by colours.

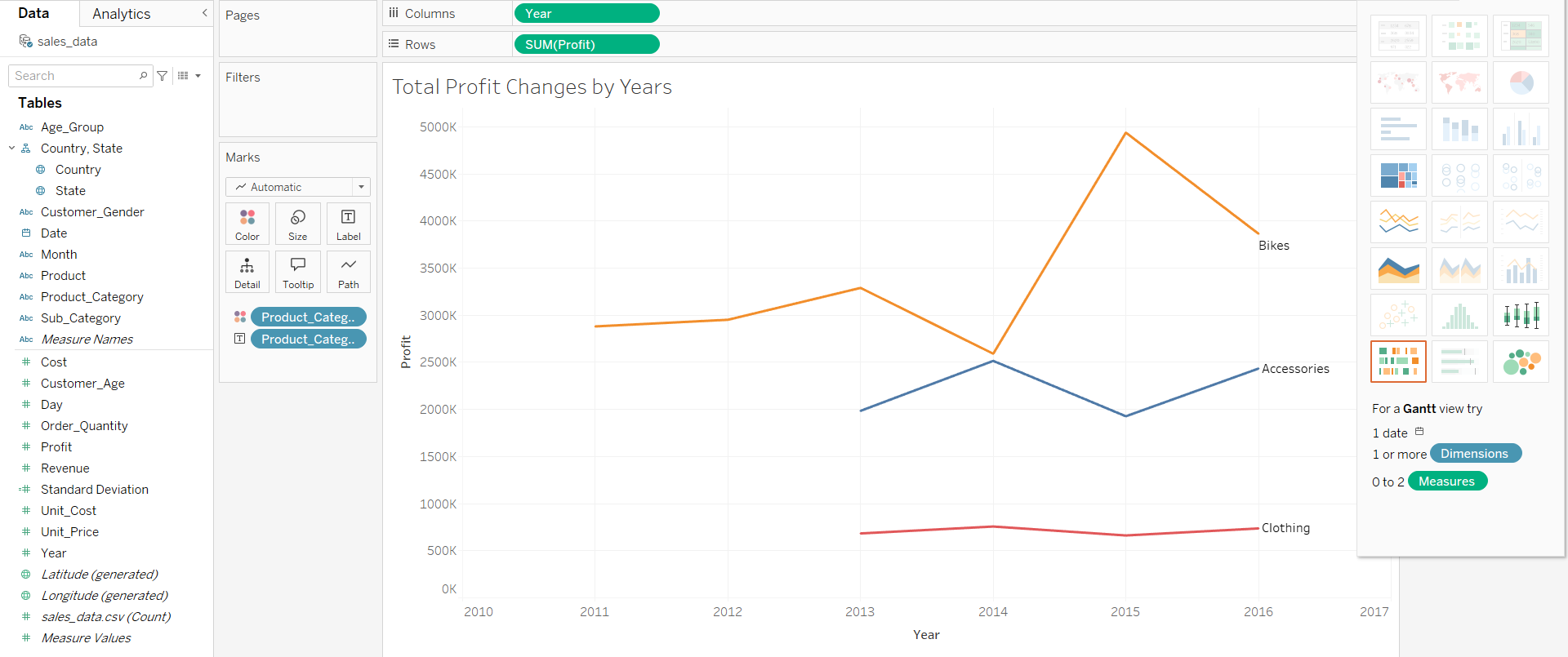


Figure 3.6 Total Profit Changes by Years

After step 5 , analysing profit can be meaningful step to get insights.Fig 3.6 demonstrates changing of total profits according to years between 2011 and 2016.To visualize it line graph can be used and marks are avaliable to differs product category by colour and labels.Columns contains Product\_Category feature then rows includes Sum (Profit) attribute.

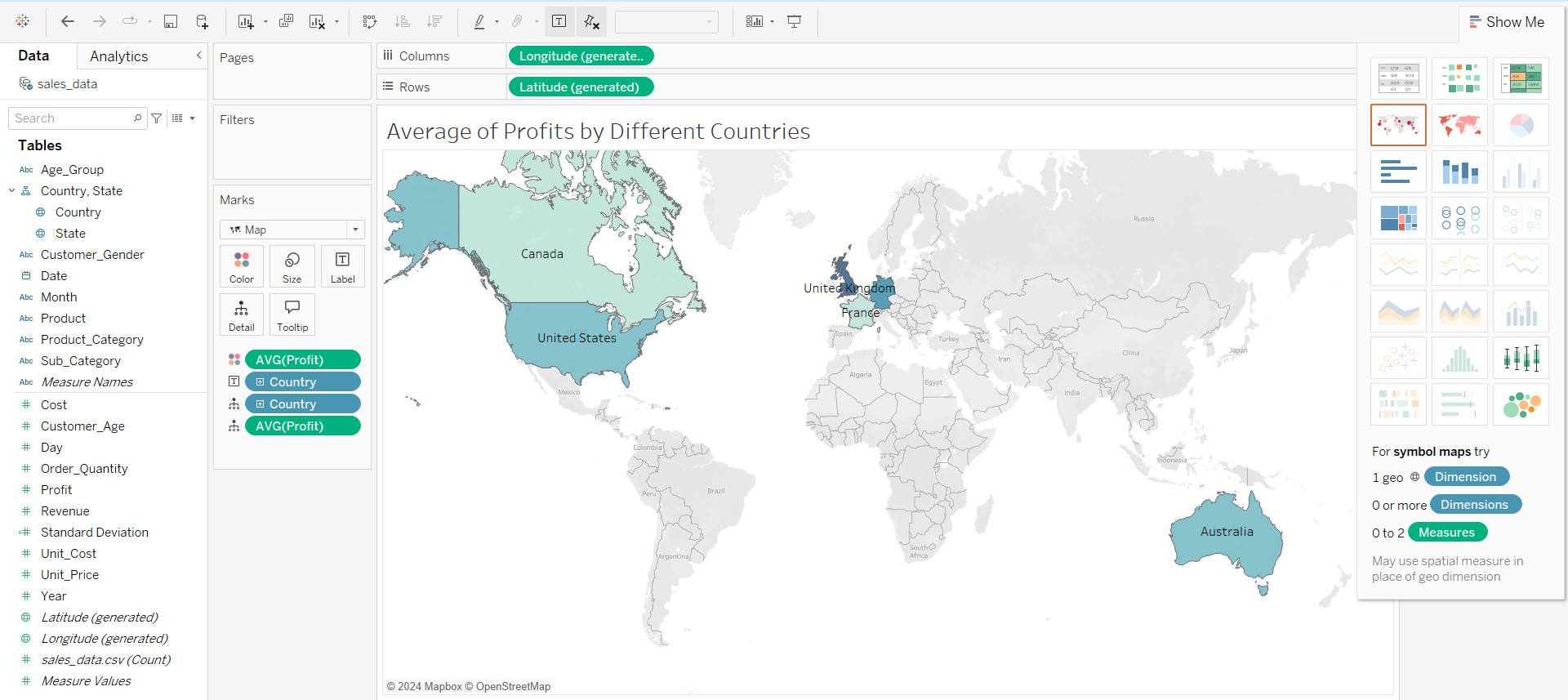


Figure 3.7 Average of Profits by Different Countries

Next step in Fig 3.7 is related to obtained profit from products by various countries.To analysis this part more sophisticated visualizaiton technique is used that is called “Maps Graph”.Then in as Figure 3.4 shows darker blue is indicator regarding to higher profit from products by different countries.To applied this visualizaiton there are Country and AVG (Profit) and countries have their name labels which is able to enable in marks section.

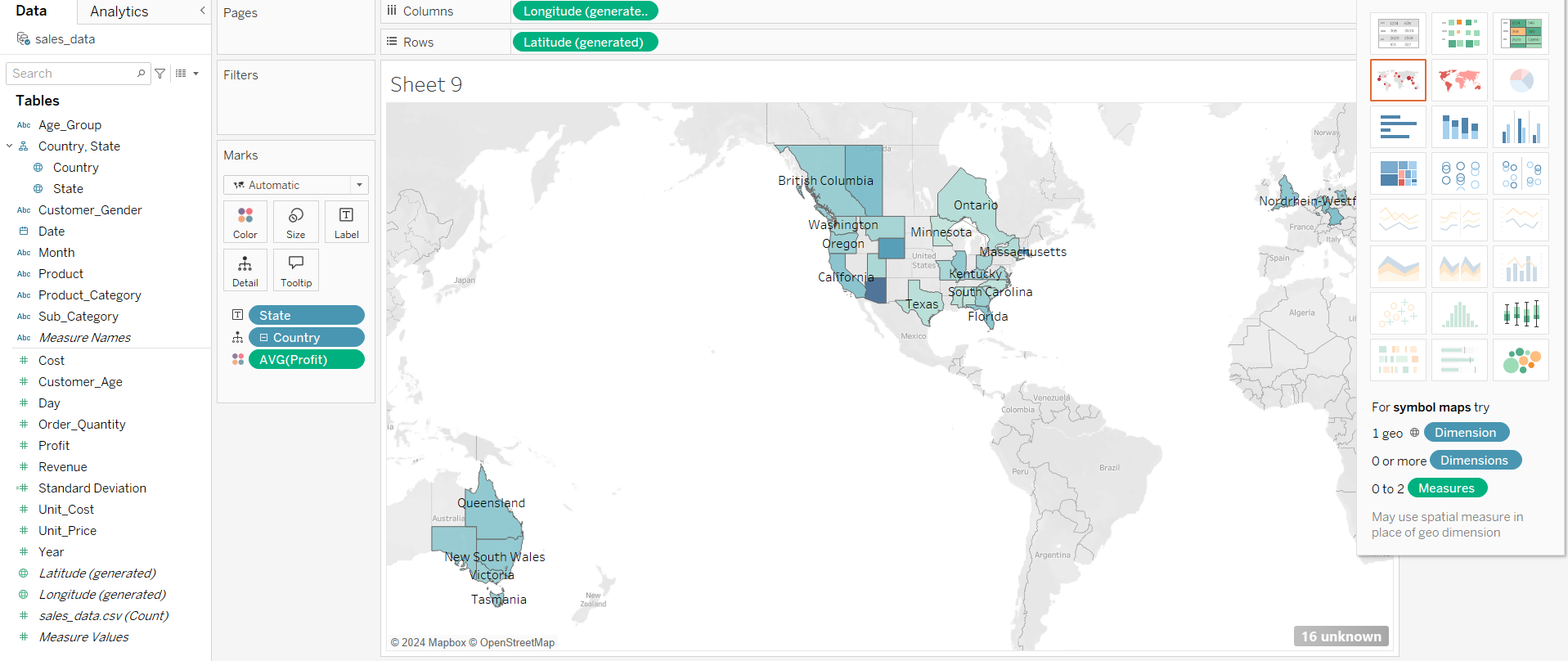
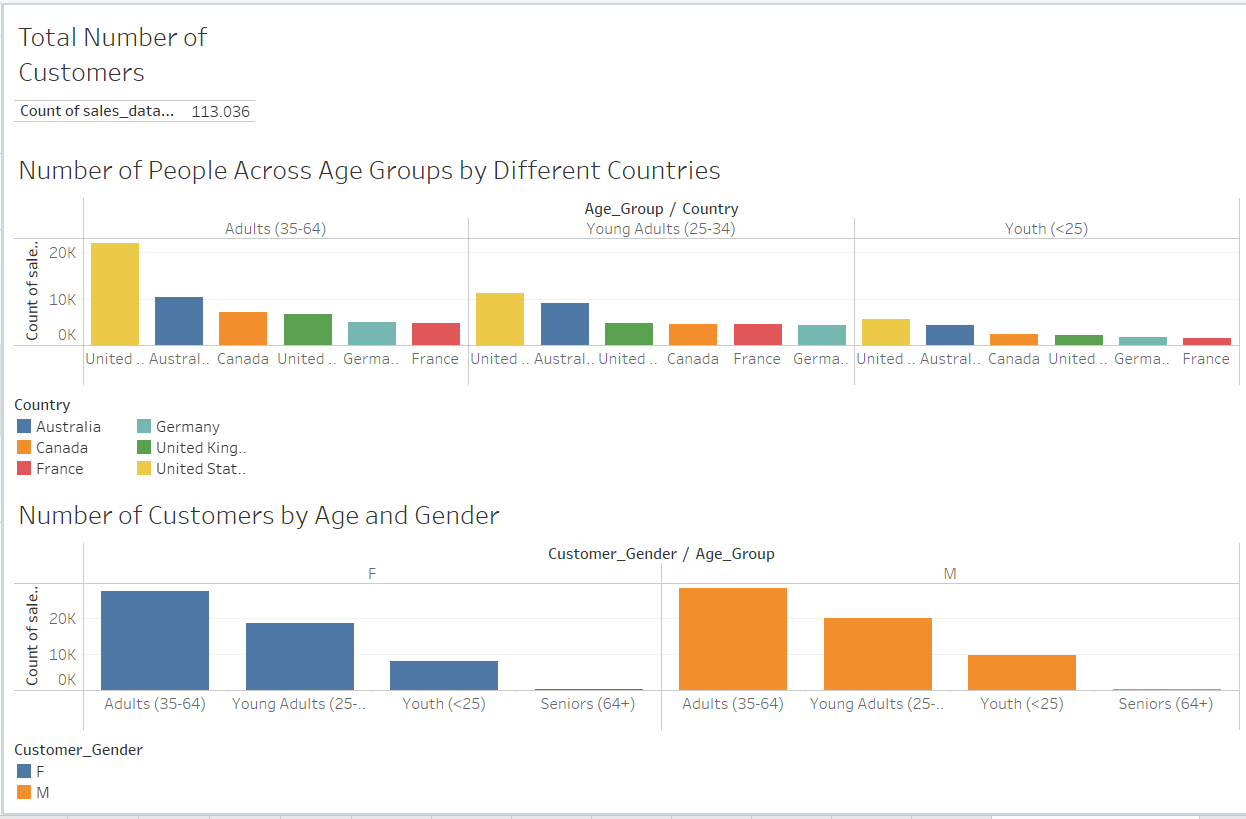


Fİgure 3.8 Average of Profits by Countries States

Apart from implementation of Fig 3.8 at the same time average can be analyzed based on countries states when State attribute is used instead of Country attribute in columns.

# CHAPTER FOUR DASHBOARD DESIGN and ANALYSIS

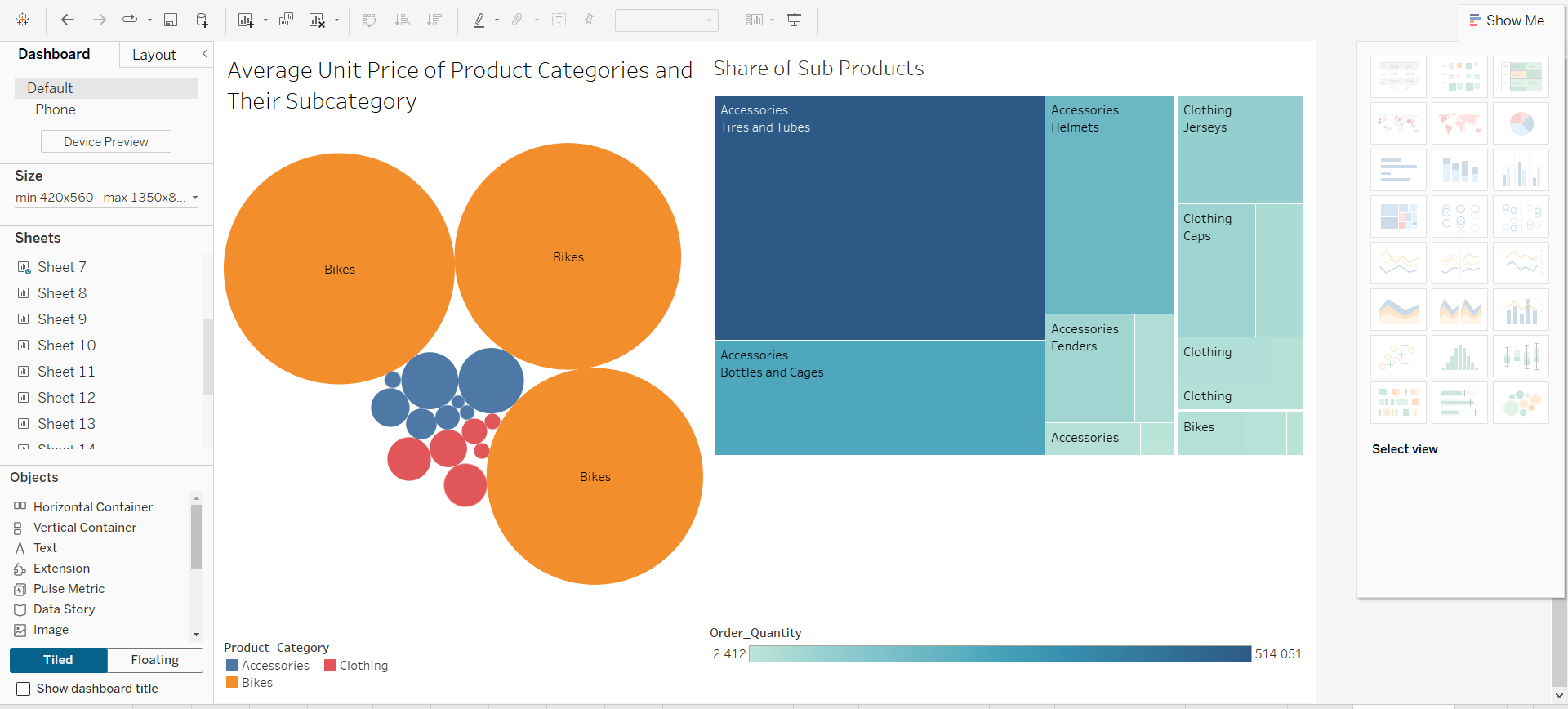
Dashboard in Tableau can be described as a general review unify of created sheets.In this study there are 3 different dashboards and which are categorized based on 3 different categories like customer demographics – order details – financial situation then all of these three category is analyzed to get benefitical insights from the dataset.

Figure 4.1 Dashboard 1 for Customer Demography

In Fig 4.1 there is the first dashboard which is related to customer demography which means analysing specific population based on size , age , countries or origin and so on.

In this dashboard there are 3 segments after these segments are able analyse as ;

* There are total over 113.000 customers in this specific order details.
* Across all age groups highest number of people belongs to United States of America and this isn’t surprisingly becuase in among all other spercific countries USA has highest population which is around 336 million people in 2023 ( The Demographic Outlook 2023 to 2053 , 2023).Moreover statitistically this situation has to be realized.On the other side about age groups which can be extracted highes number of customers are coming from adults (35-64) then it is followed by young adults (25-34) and youth people( < 25 ).
* At the same time previous analysis is still valid when explored according to both genders.

Fig 4.2 Dashboard 2 for Average Unit Price and Share of Ordered Products

In Fig 4.2 there is the second dashboard which is related to average unit price of goods and share of ordered products within the dataset.

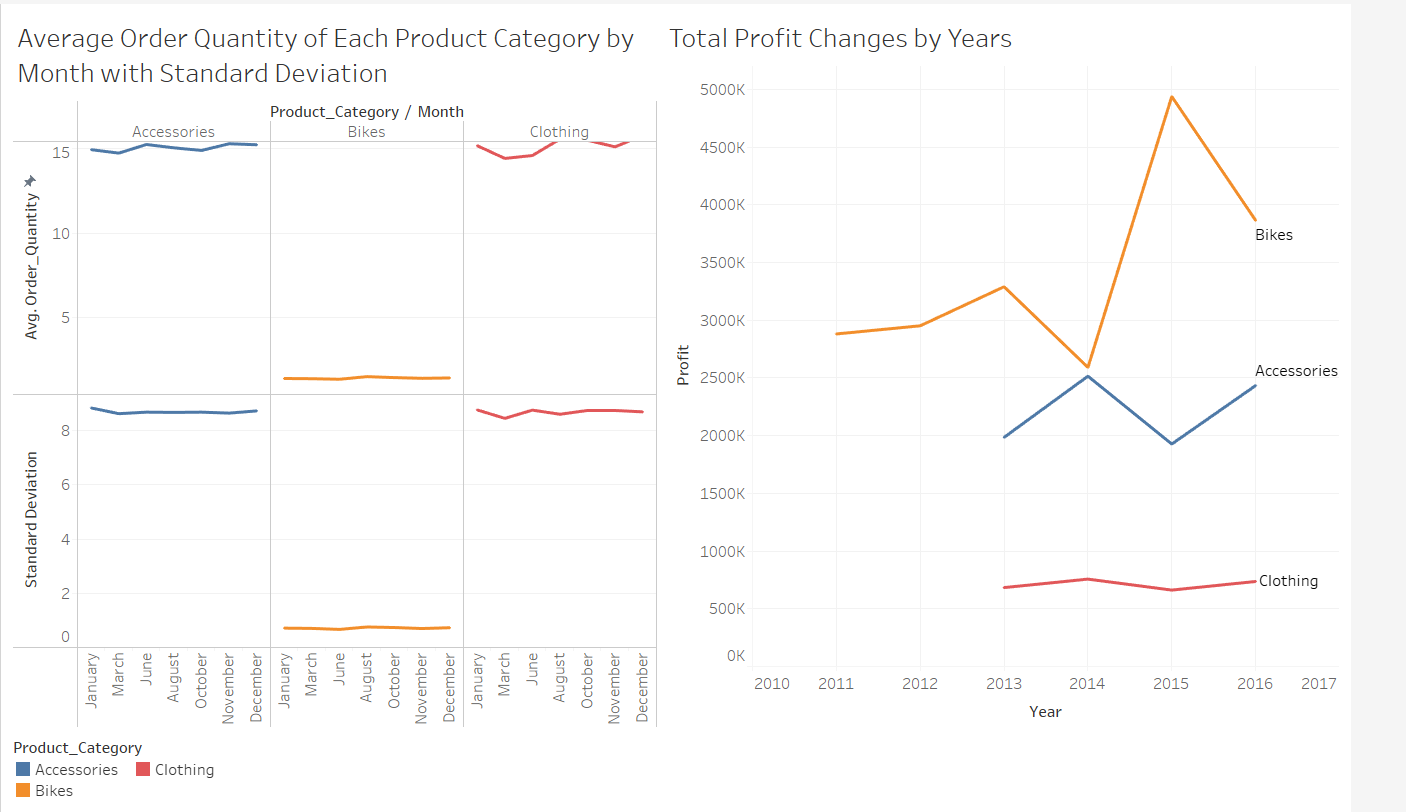
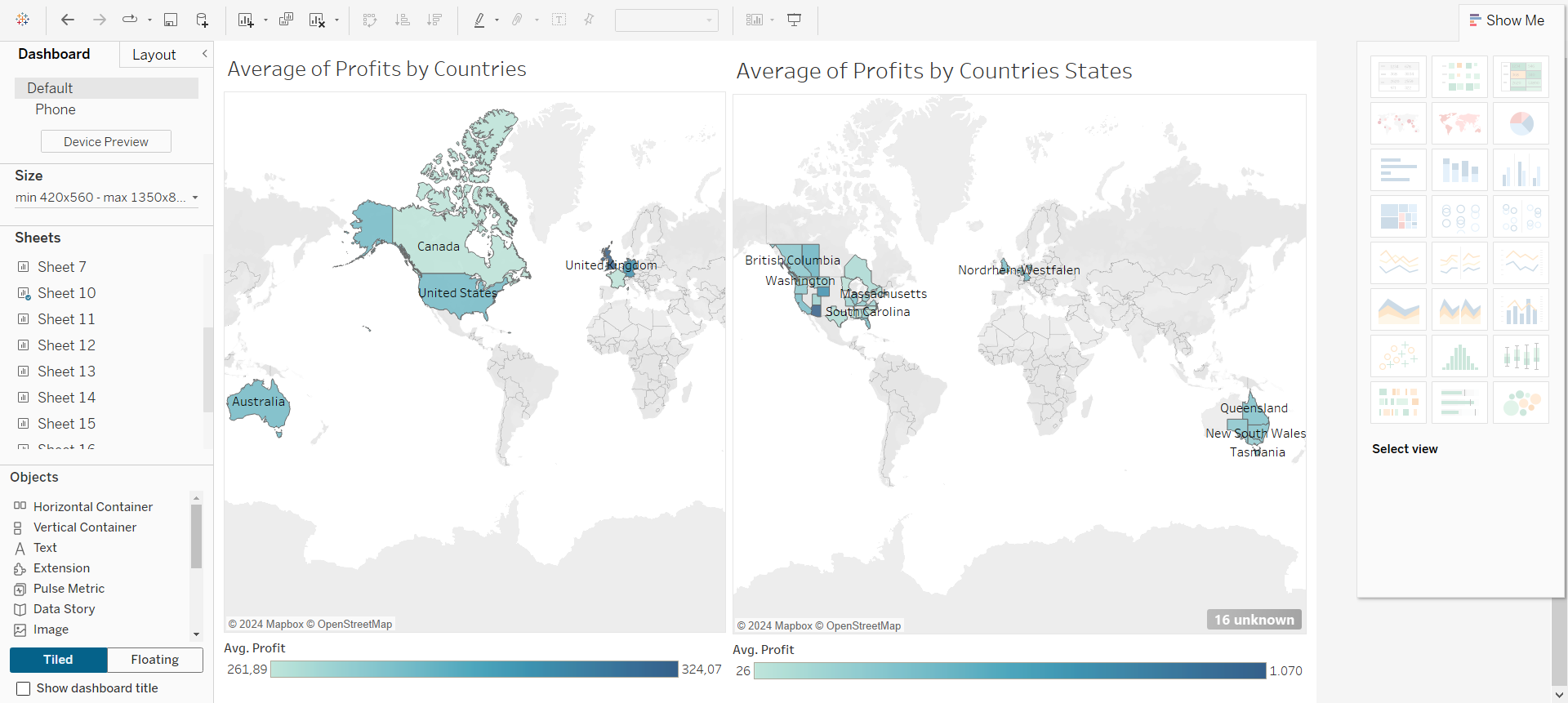
* About unit<price of prodcut highest price matches across bikes with its subcategories then its followed by accessories and clothing respectively.
* In contrast , share of ordered products demonstrate another picture because in that case accessories with its sub products are ordered at highest level rather than others.

Fig 4.3 Dashboard 3 for Order Quantity and Profit

In Fig 4.3 there is the third dashboard which is related to average order quantity by months and total profit changes throughout years between 2011 and 2016.

In this dashboard there are 2 segments then these segments are able analyse as ;

* About average ordered quantity of products that can be said peak months for accessories – bicycle – clothing are June – August –August respectively.This situation might be explained people are more tend to use bicycle during the summer days rather other seasons.
* Regarding to total profit changes throughout betwwen 2011 and 2016 among three product categories highest profit belongs to bicycle then it is followed by accessories and clothing.Moreover peak years of profits for bicycle – accessories – clothing are 2015 , 2014 , 2014 respectively.

Fig 4.4 Dashboard 4 for Average Profits by Countries and States

In Fig 4.4 there is the last dashboard which is related to average profits by countries and states.

In this dashboard there are 2 segments then these segments are able analyse as ;

* The highest profit is coming from United Kingdom it is about 324.000 then it is followed by Germany , Australia , USA , Canada and France respectivley.
* In terms of states of countries where has highest profit are England (United Kingdom ) Bayern (Germany) Queensland (Australia) , Arizona (USA) , Alberta (Canada) and Essonne (France) based on descending order.Moreover reason of this order might be these mentioned states can be marked as either riched society or high population.

# CONCLUDING REMARKS and DISCUSSIONS

In conclusion , this study handles a case scenario is handled based on specific needs which is a request from advertisement corporation for online trading company that purchases bikes and other stuff regarding to it.For this reason the company provided the dataset which includes benefitical databae regarding to demography information customers , details of ordered products as well as financial situation connected to these goods.

Extracted insights from the dataset can be very benefitical for advertisement company to intend potential customers correctly.However regarding to the dataset there is a concern which is if the dataset has more compherensive information across geographical location such as more countries in that case outcome would be more accurately.Moreover in 2032 online trading platform size will be increased to 16.5 billion USD ( Dhapte , A. , 2024) this is a sign of data analysis tools like Tableau will be more critical to handle this vast amount of information from these platforms.

# BIBLI**OGRAPHY**

1. Batt, S. *et al.* (2020) *‘Learning Tableau: A data visualization tool’*, *The Journal of Economic Education*, 51(3–4), pp. 317–328. doi: 10.1080/00220485.2020.1804503.
2. Biratti , J. (2022) , *Sales Data* available at : [Sales Data (kaggle.com)](https://www.kaggle.com/datasets/jehanzaibbhatti/sales-data) [ Accessed on 1 October ,2024]
3. Congressional Budget Office (2023) *The Demographic Outlook:2023 to 2053*
4. Dhapte , A. (2024) *Online Trading Platform Market Research Report Information By Interface Type*
5. International Monetary Fund (2024). *World Economic Outlook Update.* Database.
6. Roskladka, A., & Baiev, R. (2021). *Digitalization of data analysis tools as the key for success in the online trading markets.* Access to science, business, innovation in digital economy, 222-233.
7. Weissgerber, T. L. (2021). *Learning from the past to develop data analysis curricula for the future.* PLoS Biology, 19(7), e3001343.

# APPENDIX (if necessary)

IMF : INTERNATIONAL MONETARY FUND

UK : UNITED KINGDOM

USA : UNITED STSTAES of AMERICA

USD : UNITED STATES DOLLARS