



BOOMBIKES BIKE SHARING

Data Visualization using Tableau
Internal Assessment

Sample data: 2011-2016

Avadhut Dilip Varvatkar
Roll no: 36
Set no: 3836384

B.K. Birla College of Arts, Science & Commerce, Kalyan



BoomBikes Bike Sharing

Introduction

The dataset consists sales data of the Boom-Bikes bike sharing of each year from 2011 to 2016. The dataset was downloaded from kaggle. It consists of information like product category, Sub-Category, Product, Order Quantity, unit price, unit cost, Revenue, Profit, Country, State, Age Group, Customer age, Order Date, Customer Gender. Date Cleaning done Using Python Pandas Library And the Visualizations were created on Tableau Desktop.

Reading Data in Pandas

Jupyter Untitled Last Checkpoint: Yesterday at 10:16 PM (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3 (ipykernel) O

In [1]: `import pandas as pd`
`import numpy as np`

In [2]: `df = pd.read_csv("Sales.csv")`

In [4]: `df.head()`

Out[4]:

	Date	Day	Month	Year	Customer_Age	Age_Group	Customer_Gender	Country	State	Product_Category	Sub_Category	Product	Order_Quantity
0	2013-11-26	26	November	2013	19	Youth (<25)	M	Canada	British Columbia	Accessories	Bike Racks	Hitch Rack-4-Bike	8
1	2015-11-26	26	November	2015	19	Youth (<25)	M	Canada	British Columbia	Accessories	Bike Racks	Hitch Rack-4-Bike	8
2	2014-03-23	23	March	2014	49	Adults (35-64)	M	Australia	New South Wales	Accessories	Bike Racks	Hitch Rack-4-Bike	23
3	2016-03-23	23	March	2016	49	Adults (35-64)	M	Australia	New South Wales	Accessories	Bike Racks	Hitch Rack-4-Bike	20
4	2014-05-15	15	May	2014	47	Adults (35-64)	F	Australia	New South Wales	Accessories	Bike Racks	Hitch Rack-4-Bike	4

In [5]: `df.columns`

Checking For NULL Values

Jupyter Untitled Last Checkpoint: Yesterday at 10:16 PM (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3 (ipykernel) C

In [5]: `df.columns`

Out[5]:

```
Index(['Date', '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'],
      dtype='object')
```

In [9]: `df.isna().sum()`

Out[9]:

Date	0
Day	0
Month	0
Year	0
Customer_Age	0
Age_Group	0
Customer_Gender	0
Country	0
State	0
Product_Category	0
Sub_Category	0
Product	0
Order_Quantity	0
Unit_Cost	0
Unit_Price	0
Profit	0
Cost	0
Revenue	0

dtype: int64

Metadata When Connected to Tableau

Fields

Type	Field Name	Physical Table	Remote Field Name
Date	Date	Sales.csv	Date
#	Year	Sales.csv	Year
#	Customer Age	Sales.csv	Customer_Age
Abc	Age Group	Sales.csv	Age_Group
Abc	Customer Gender	Sales.csv	Customer_Gender
⊕	Country	Sales.csv	Country
⊕	State	Sales.csv	State
Abc	Product Category	Sales.csv	Product_Category
Abc	Sub Category	Sales.csv	Sub_Category
Abc	Product	Sales.csv	Product

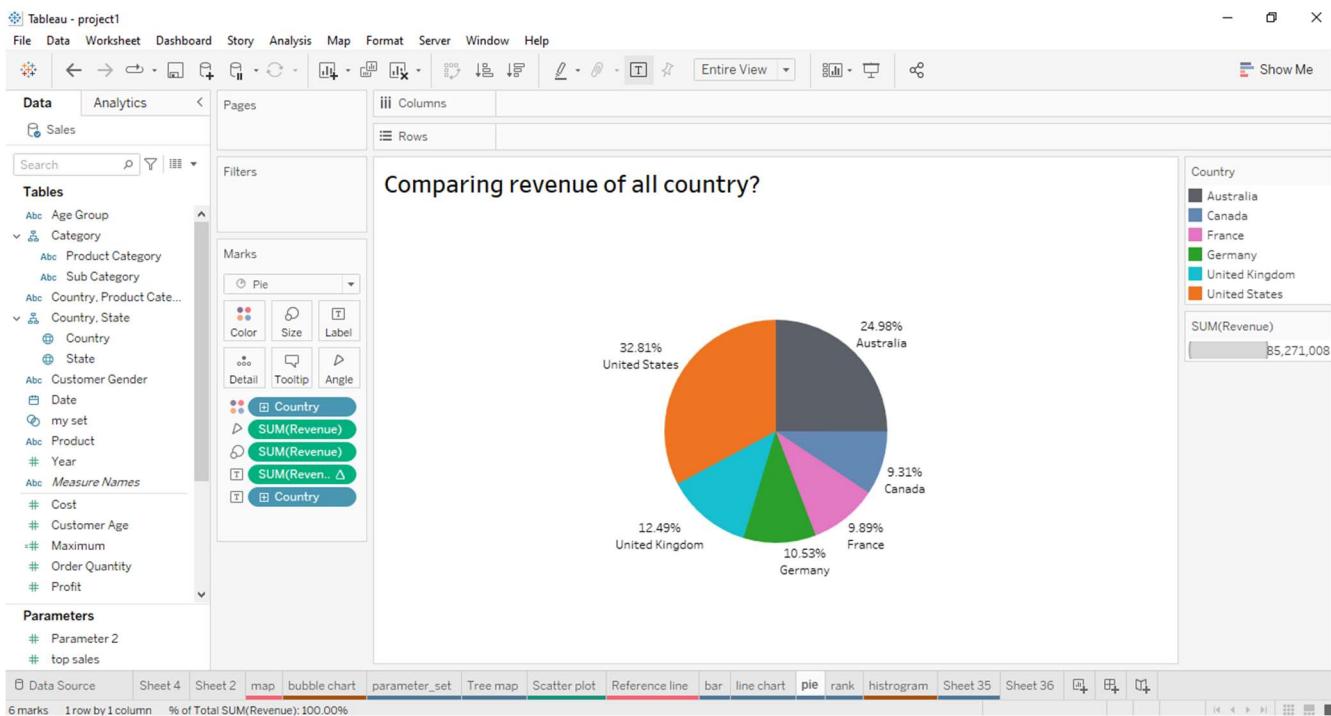
When the dataset is connected to Tableau, Tableau automatically assigns certain role to the data. **Product Name, Category and Sub-Category, Customer Gender, Age Group, Product, Country, State** are **Dimensions** whereas **Sales, Revenue, Price, cost, Unit Price, Unit Cost** are **Measures**. Year was considered as a Number but we changed its data type to Date.



1. Comparing revenue of all country?

- In mark card **Drag** the **SUM(Revenue)** in label and Drag Country in color shelf
- In **mark card click** menu and select pie chart.
- Drag **SUM(Revenue)** in label and select percent of total in **Quick table calculation**
- The pie chart is created.

Revenue distribution across countries



Revenue distribution across countries

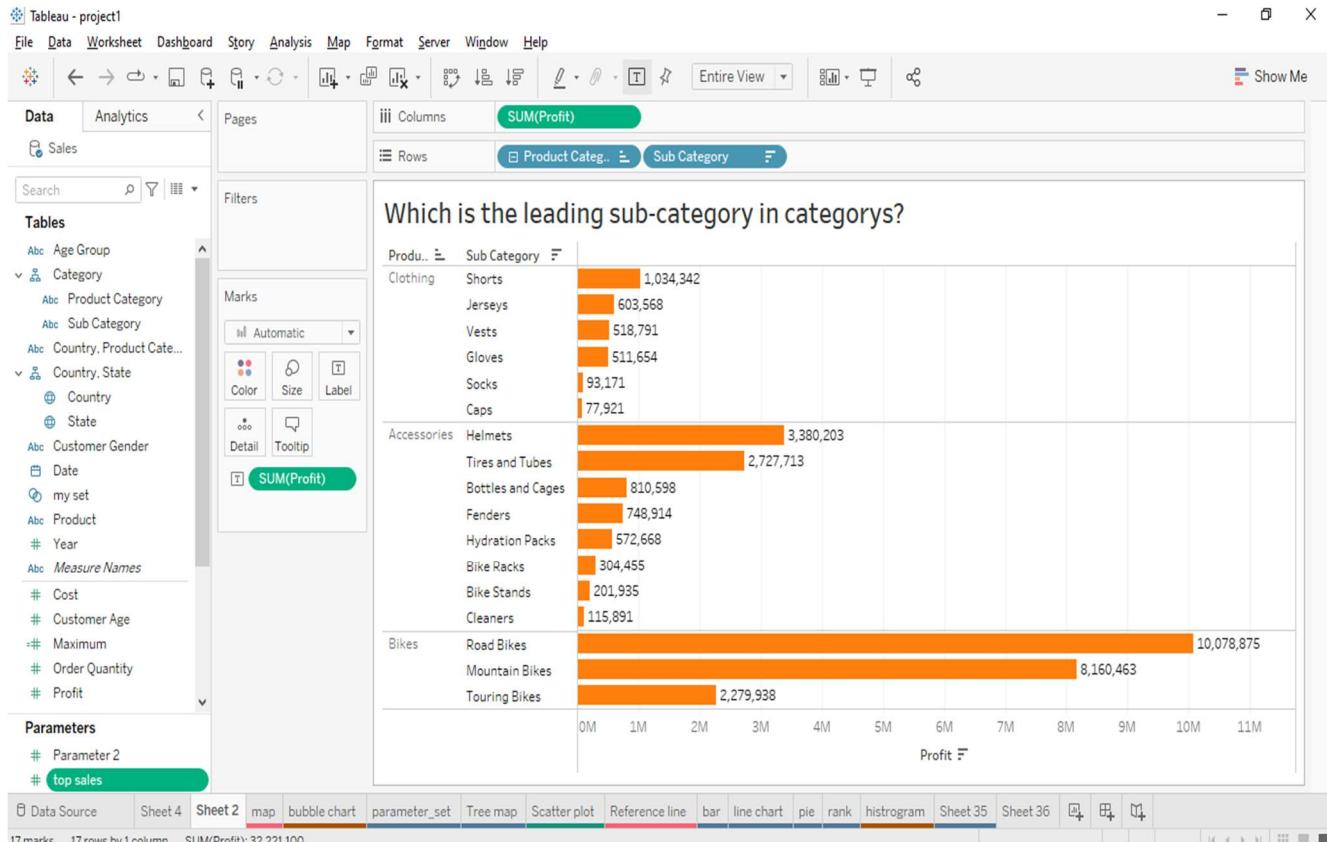
From the above visualization it is clearly visible that USA has the highest revenue collection.

More than 50% of the revenue comes from USA and Australia

Canada has the lowest share in revenue.

2. Which is the leading sub-category in category?

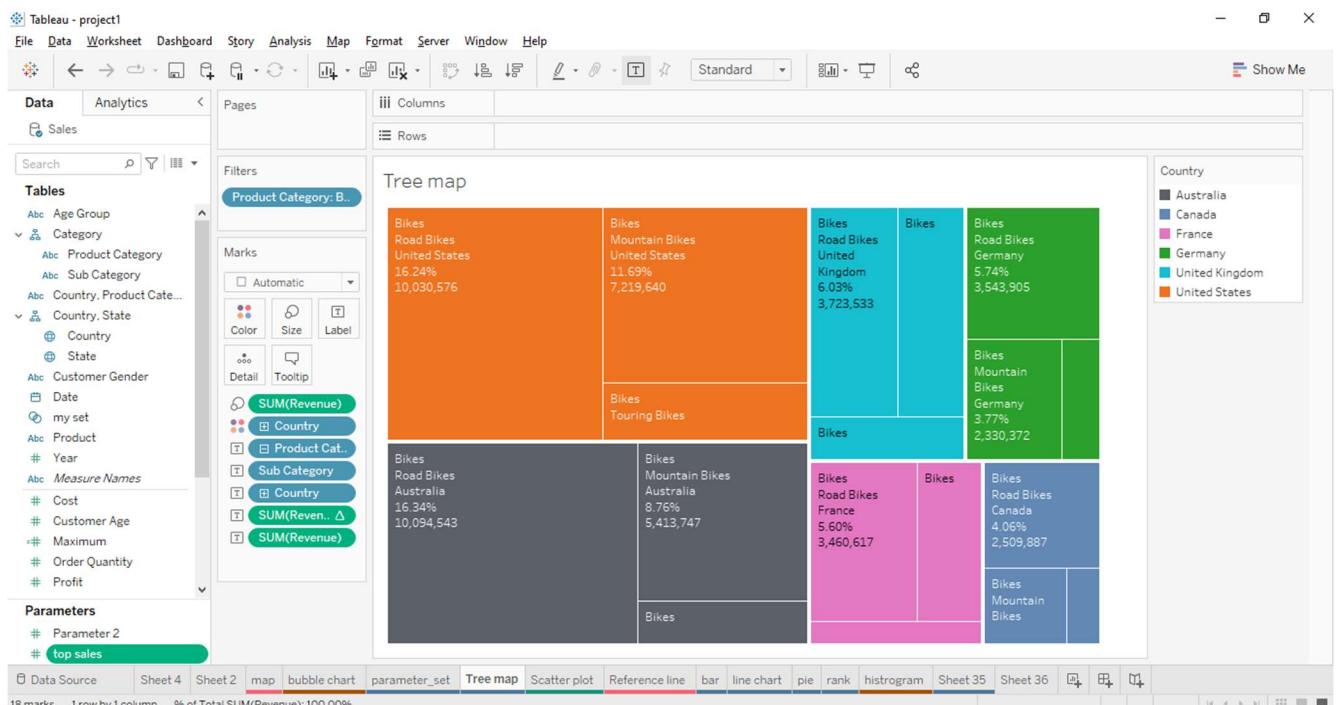
- Drag the **product category** and **Sub-Category** in row shelves and Drag the **SUM(Profit)** in Row shelf.
- In the Marks Card dropdown menu, select **Bar char**
- In Menu Bar click **ascending** button the date arrange in ascending order.
- In mark card drag the **SUM(Profit)** in the **label** and **color**



3. Which bike model is more revenueable in all country?

- Drag **sub-category** in row shelf and in **mark card** drag **SUM(Revenue)** in Shapes if in case **Bubble chart** is not plot then **click show me** and select the **bubble chart**
- In **mark card** Drag **Revenue** in label And Select present of Total in **quick table calculation**,
- Drag **sub-category** in color And **SUM(Revenue)** in label.

Bike sub-category wise revenue distribution across countries

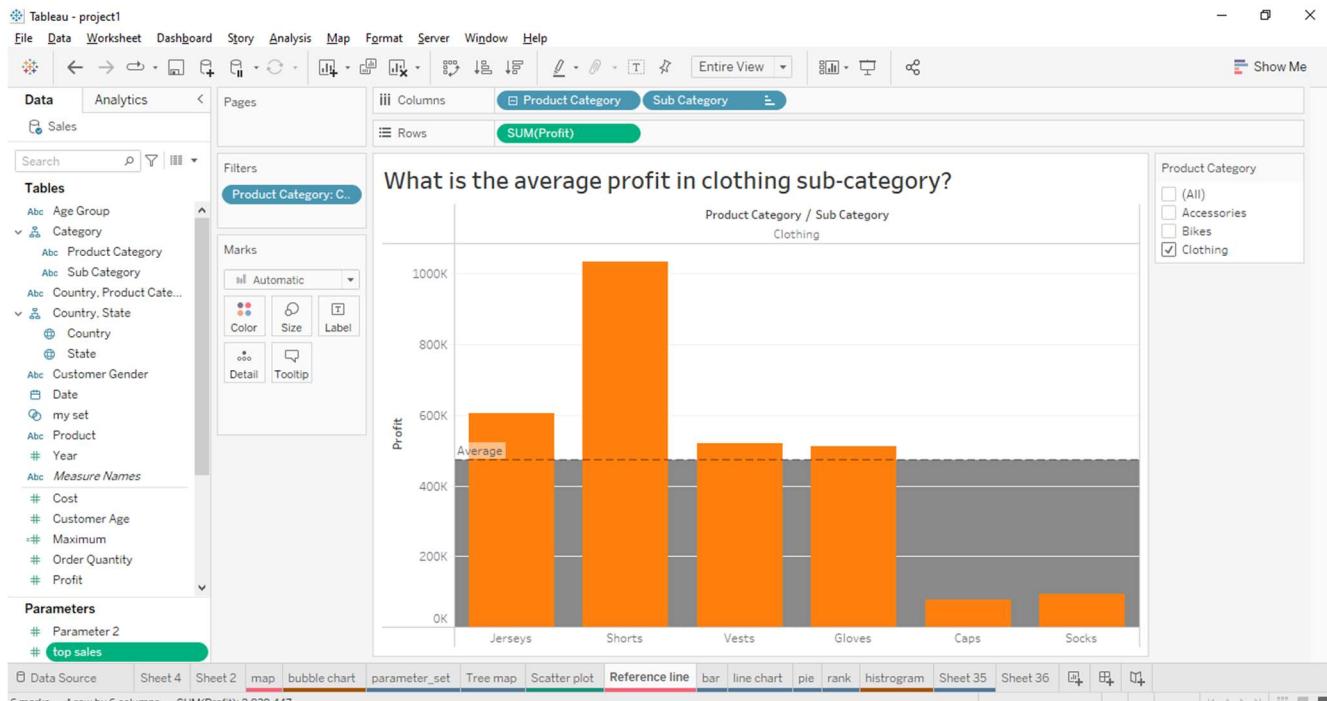


Bike sub-category wise revenue distribution across countries As understood from the previous visualization bike has the highest contribution in the revenue so it is essential to know sub category wise bike revenue distribution for a better understanding to take better business driven decision As we can see Road bike contribute more then 50% revenue share in bike category across countries whereas mountain bike contribute the least in revenue share.

4. What is the average profit in clothing sub-category?

- Drag the **(Product category)** and **(Sub-Category)** in Column Shelves And **SUM(Profit)** in Row Shelf.
- Drag **(Product Category)** in **Filter** and In general Section Select only clothing **(sub-Category)** click apply and **(OK)** the filer is created.
- Right Click on X-axis and **click Create Reference line**,
In Scope select per-pane, In Line Select Value **Sum(profit)** Average and Format the line as you want then Click apply **(OK)**, The Reference line is

Taking average as a reference line



Taking average as a reference line

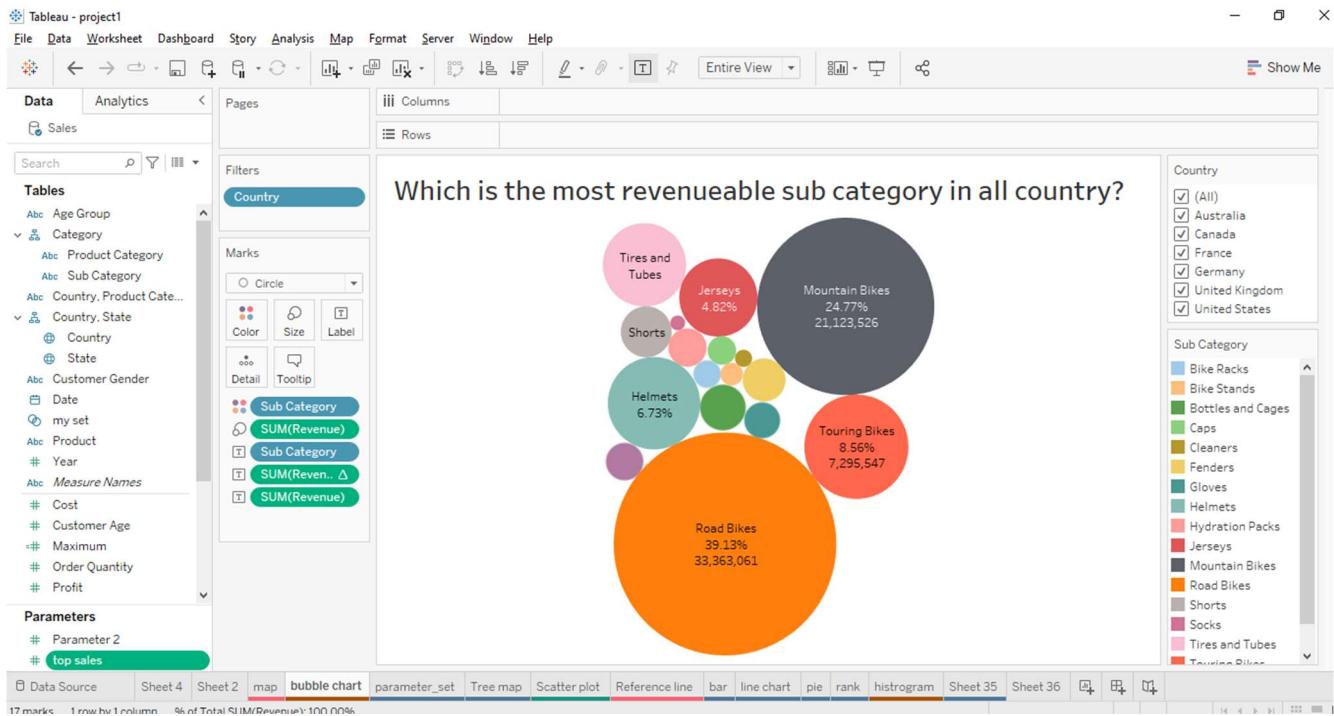
The above visualization depicts the profit share from clothing sub categories across countries.

The average of profit from clothing sub categories is taken as reference line. From the above visualization caps and socks lie below the reference line which indicates the share of profit from socks and caps is less as compared to the other sub categories.

5. Which is the most revenue sub category in all country?

- Drag sub-category in row shelf and in mark card drag **SUM(Revenue)** in Shapes if in case Bubble chart is not plot then **click show me** and select **the bubble chart**
- In **mark card** Drag **Revenue** in **label** And Select present of **Total in quick table calculation,**
- Drag sub-category in color And **SUM(Revenue)** in label.

Sub-category wise revenue distribution across countries



Sub-category wise revenue distribution across countries

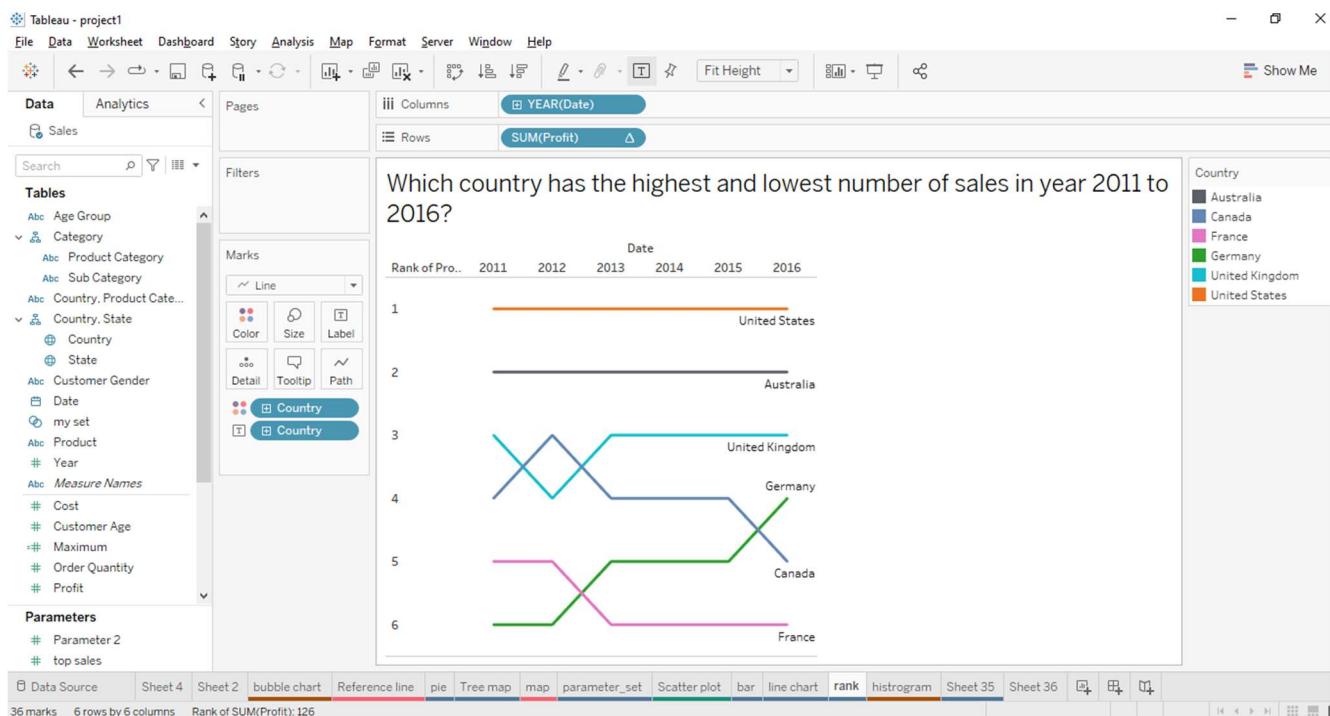
As visible in the bubble chart road bike contributes the highest revenue share whereas socks contribute the least in the revenue share.

Bigger the size of the bubble higher the revenue share of that particular sub-category

6. Which country has the highest and lowest number of sales in year 2011 to 2016?

- Drag the **YEAR(DATE)** in Column Shelf and Drag the (**CNT(Sales)**) in Row shelf.
- Click **Sales** and select **rank in Quick** table calculation, and convert **continuous series** to **discrete series**.
- Click one more time in **sales** and click (**Computing Using**) and select the Count the chart is created.

Year wise profit distribution across countries in terms of ranking



Year wise profit distribution across countries in terms of ranking

From the above visualization it is clear that in terms of profit between 2011 to 2016 United States and Australia is at 1st and 2nd position respectively.

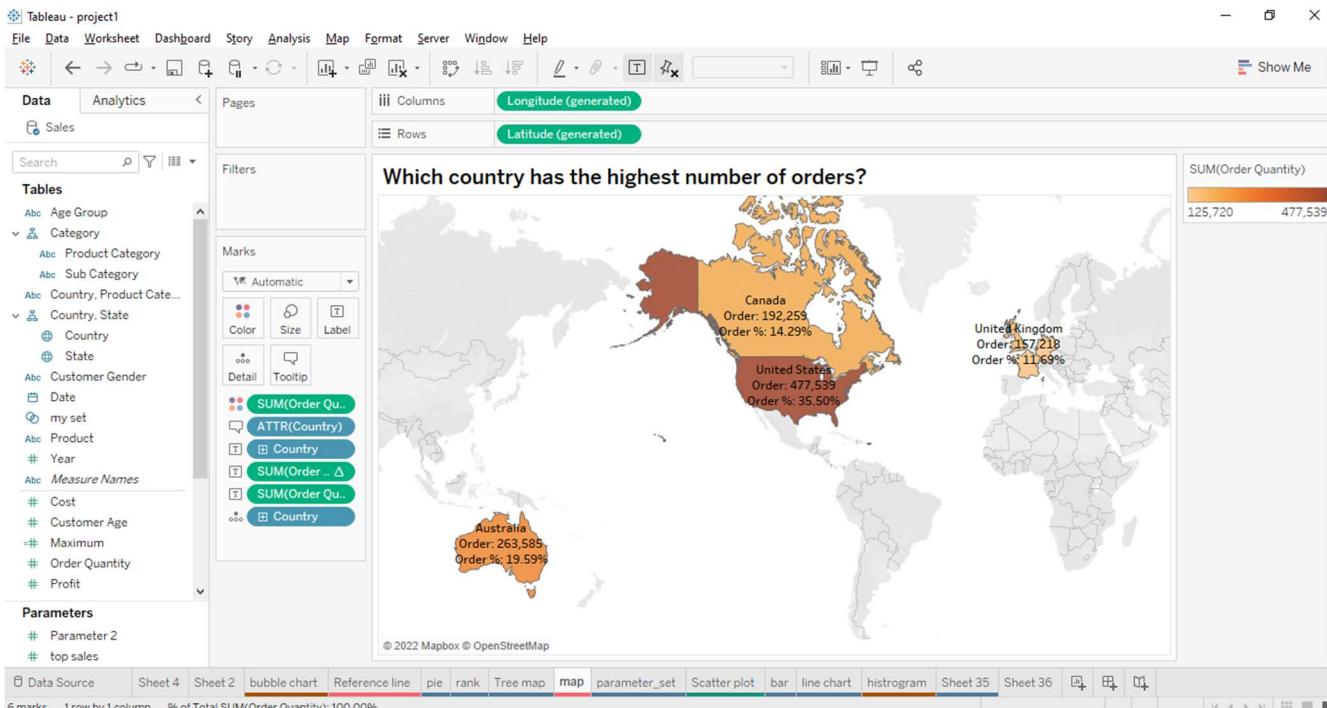
Whereas the profit in Canada is declineing from 2012 till 2016

There is a continuous increase in profit of France from 2012 till 2016

7. Which country has the highest number of orders?

- Steps: double click on **country**, then drag **SUM(Order quantity)** in labels
And select percent of total in Quick Table Calculation
- In **Mark card** Drag order In **Color**
- Click on **labels** and **edit Text**. Give name **order Percentage**

Which country has the highest orders?



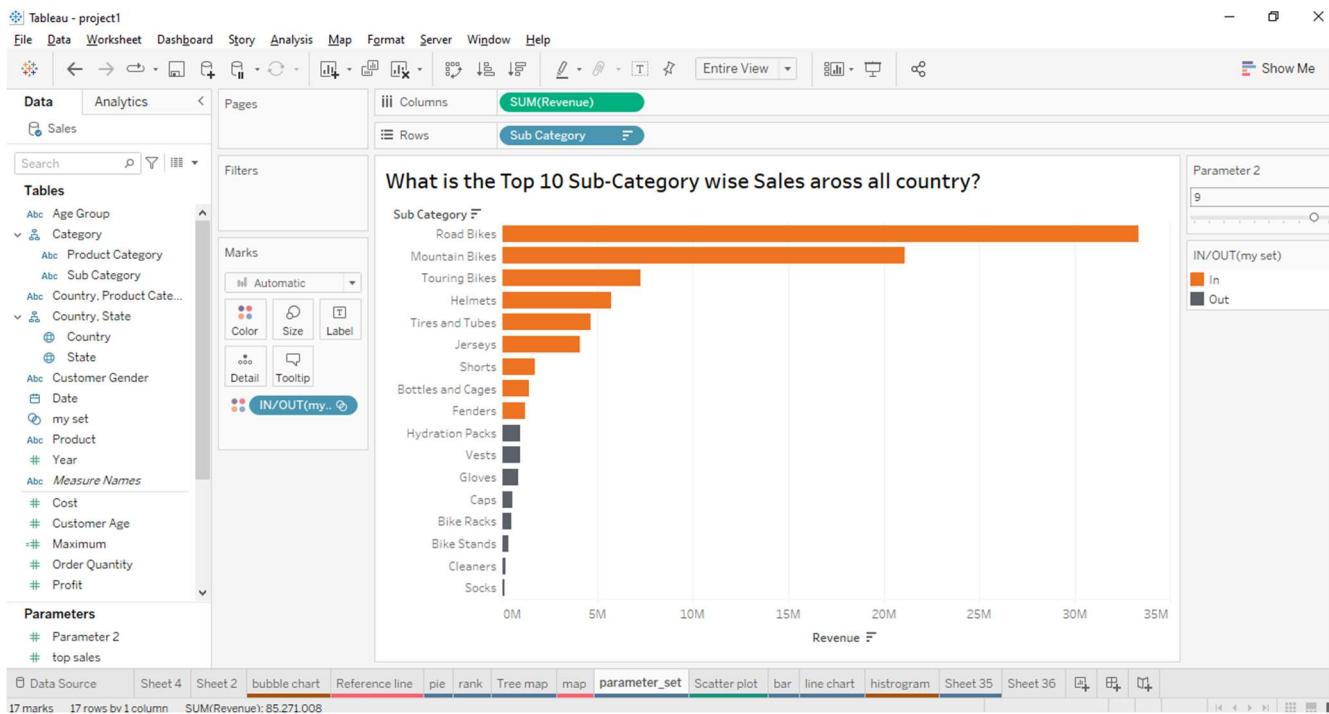
Which country has the highest orders?

United States gets the highest order compared to other countries

8. What is the Top 10 Sub-Category wise Sales across all country?

- Drag **SUM(Revenue)** in color shelf and **sub- category** in Row shelf. In menu bar **click short button** and short in **ascending order**.
- Click on **sub – category** create set and go to **Top Section** And select By Field and select Top 10 by **SUM(revenue)** Give One Better Name As **(Top 10 Sub-Category)** apply **And ok**, The new **parameter fields** is created.
- Click on **(Top 10 Sub-Category)** and Select Show parameter, (**IN and OUT**) Created

Sub category wise sales distribution across countries.



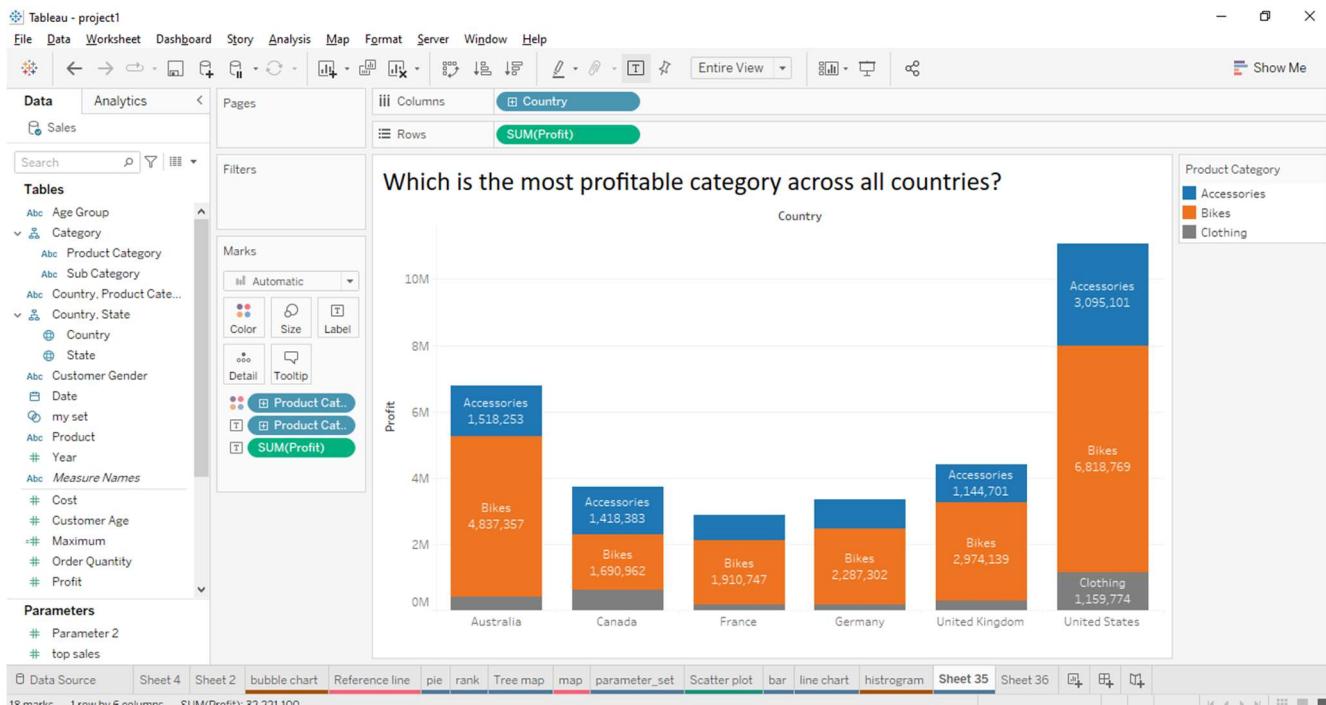
By using parameter set we can sort all the sub category into ascending and descending order as per the sales and highlight the top sub category by assigning parameter

From the visualization we can conclude that road bikes and mountain bikes have the highest sale whereas cleaners and socks has the least sale.

9. Which is the most profitable category across all countries?

- Drag the (Country) in Column Shelf and Drag the SUM(Profit) in Row shelf.
- In mark card Drag (Sub-Category) in label and press control and Drag (Sub-Category) in color.
- Chick show me button and select Stacked bar chart, the bar is created.
- In mark card drag SUM(Profit) in Label section.

Category wise profit distribution across countries



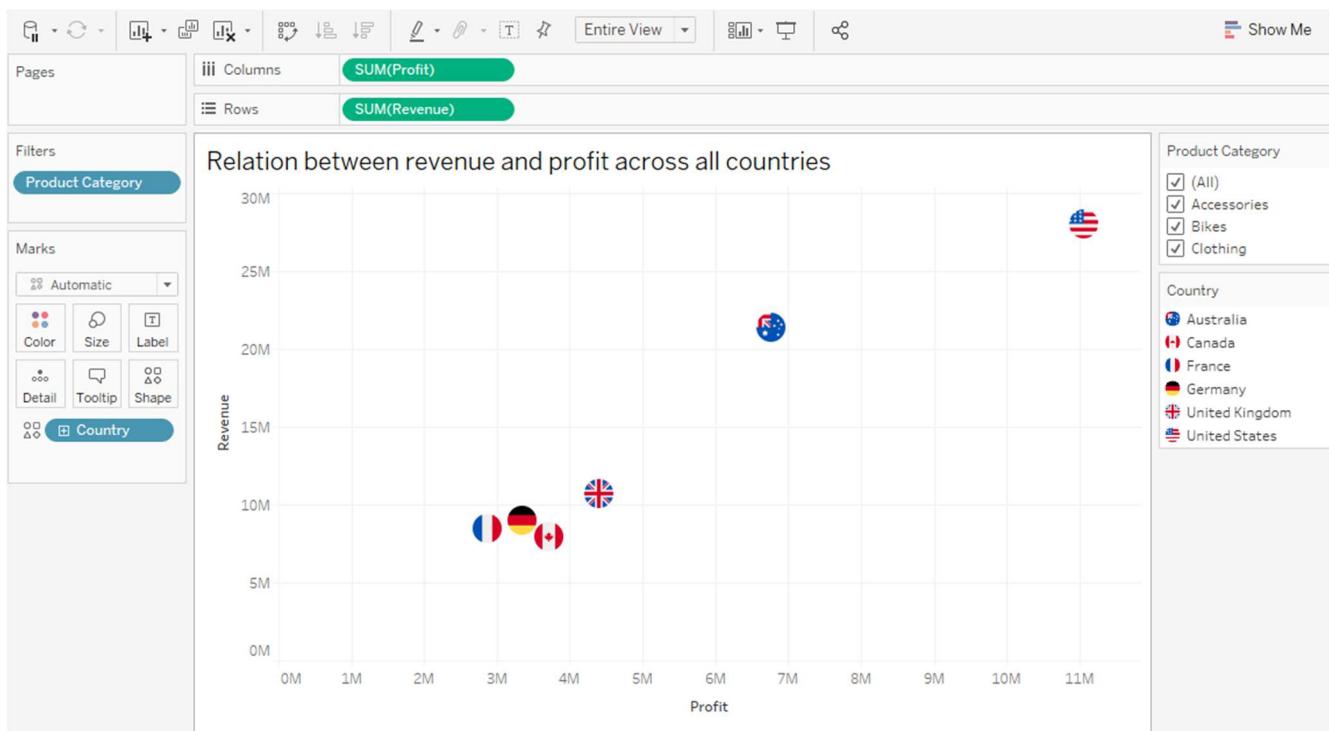
Category wise profit distribution across countries

From the above visualization we can see for all the countries the major profit comes from bike and the least profit comes from clothing.

10. Relation between revenue and profit across all countries

- Drag **SUM(profit)** in Row Shelf and **SUM(Revenue)** in Column shelf, In Mark card Drag **country** in **Shape** The scatter plot is created for better **visualization**
- Click on **Shape** and **Select Shape palette** as **Flags**.

Revenue and profit difference between countries



Revenue and profit difference between countries

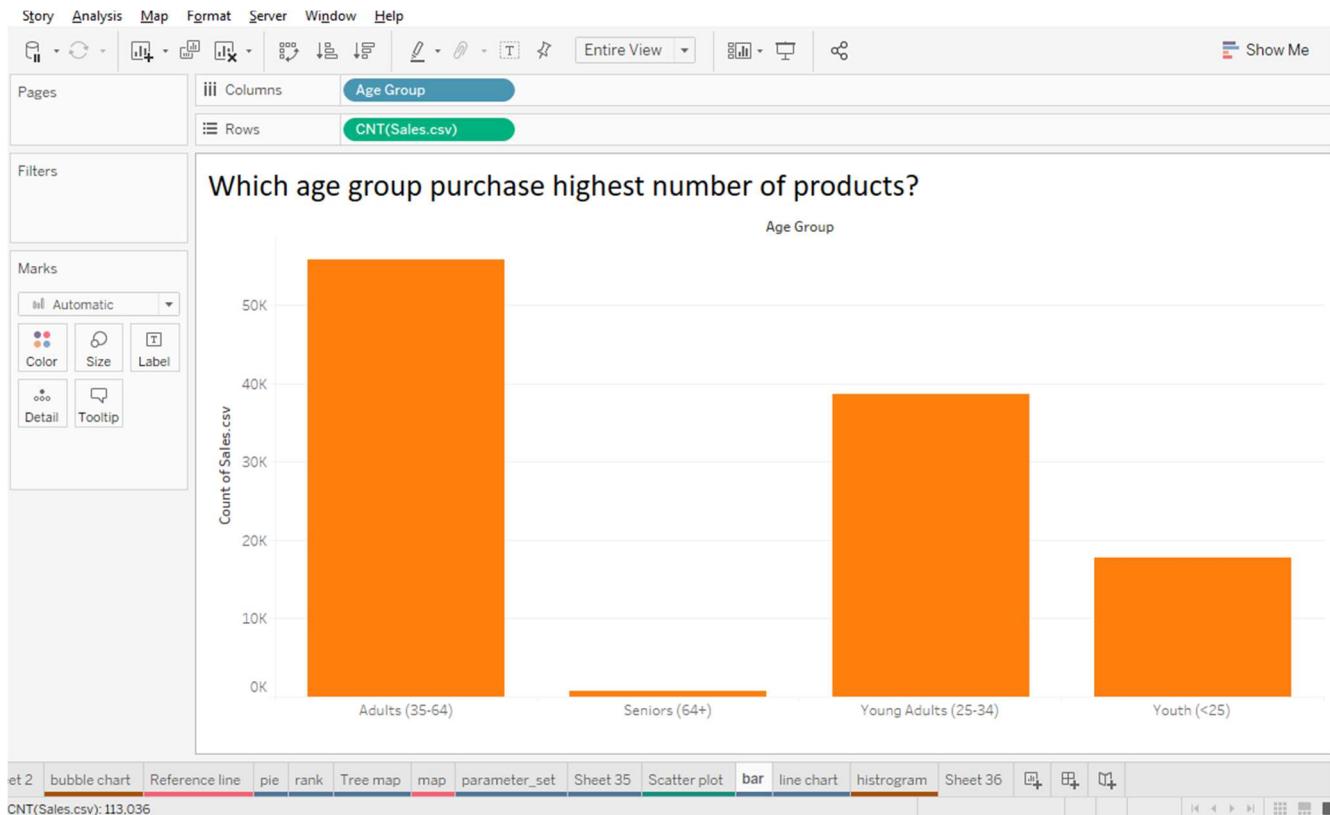
From the above visualization we can see USA and Australia has more revenue and profit compared to other Countries.

We can see Germany gets more revenue than Canada still Canada has more profit then Germany though it has less revenue than Germany which means Canada has more profit margins as compared to Germany.

11. Which age group purchase highest number of products?

- Drag the **Age Group** in **Color** shelf and Drag the **CNT(Sales)** in Row Shelf
- Click on **color** in mark card and select a suitable **color**, the Bar chart is created.

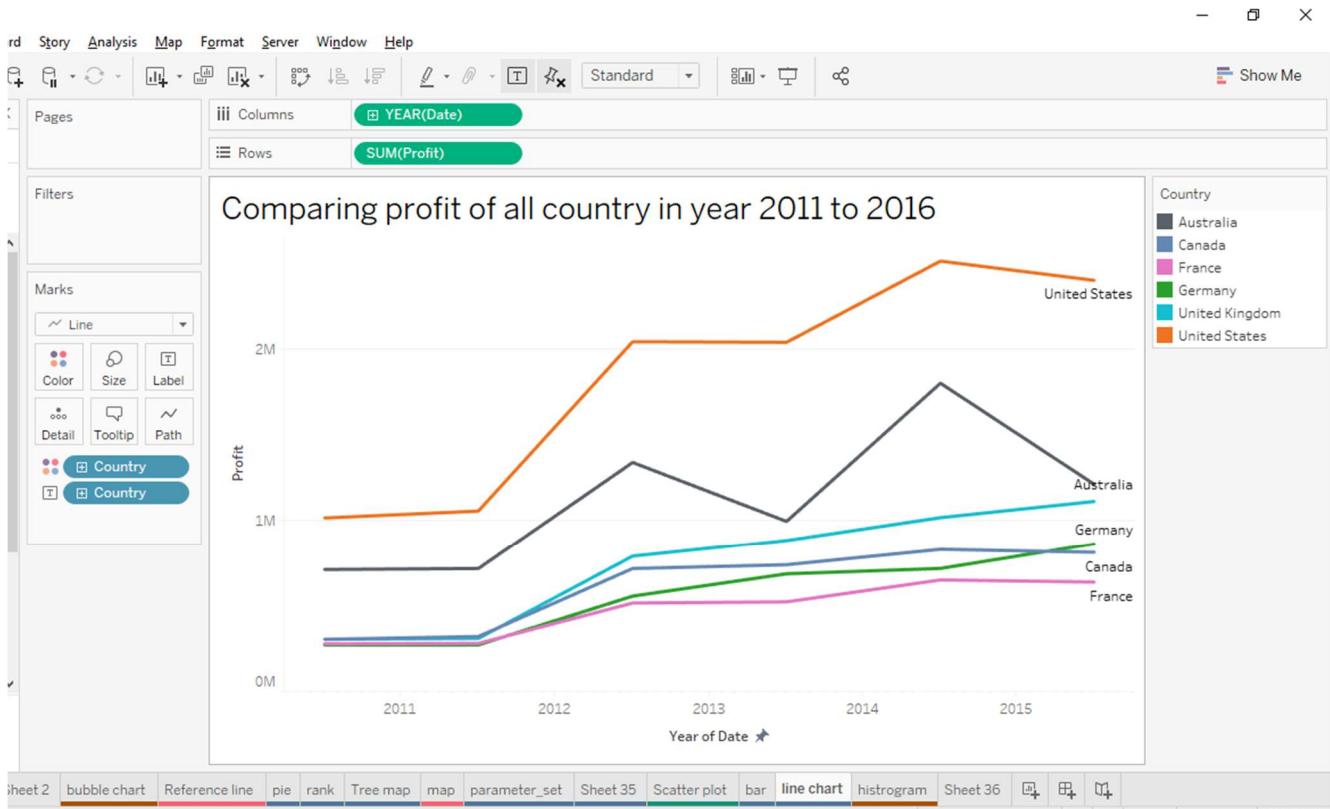
Which age group purchase the highest number of products?



Age group between 35-64 purchase the highest number of product

12. Comparing profit of all country in year 2011 to 2016

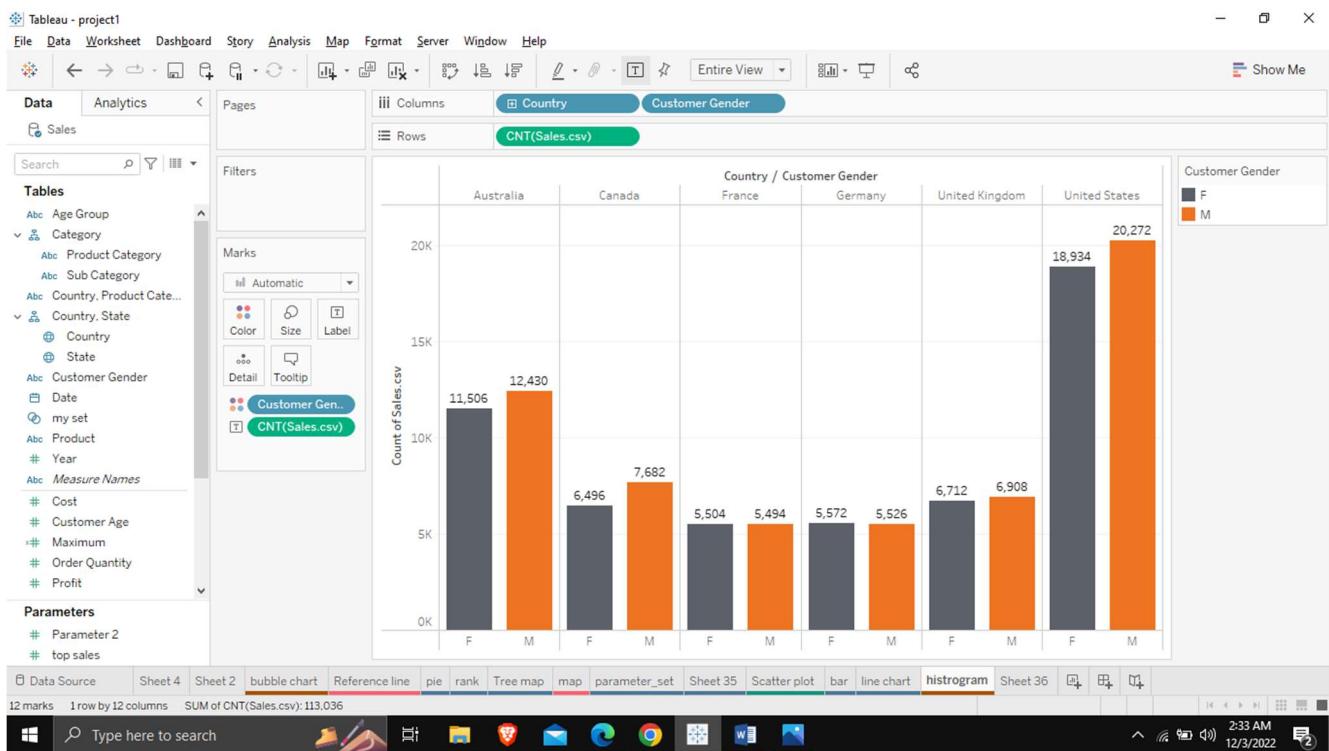
- Drag the **YEAR(DATE)** in column Shelf and **SUM(Profit)** in the Row Shelf.
- In Mark Card Drop the **Menu** and select the **line chart**.
- In Mark Card Drag **(Country)** in **Color** and Press Control and Drag **Country** in **labels**.



13. Compare male and female sales in all country's

- Drag the (Country) and (Customer Gender) in Column shelves and Drag the Sales CNT(Sales) in row shelf.
- Click show me button and select the Histogram.
- In mark card drag the customer Gender in Color and CNT(Sales) in labels

Gender wise sale distribution across countries

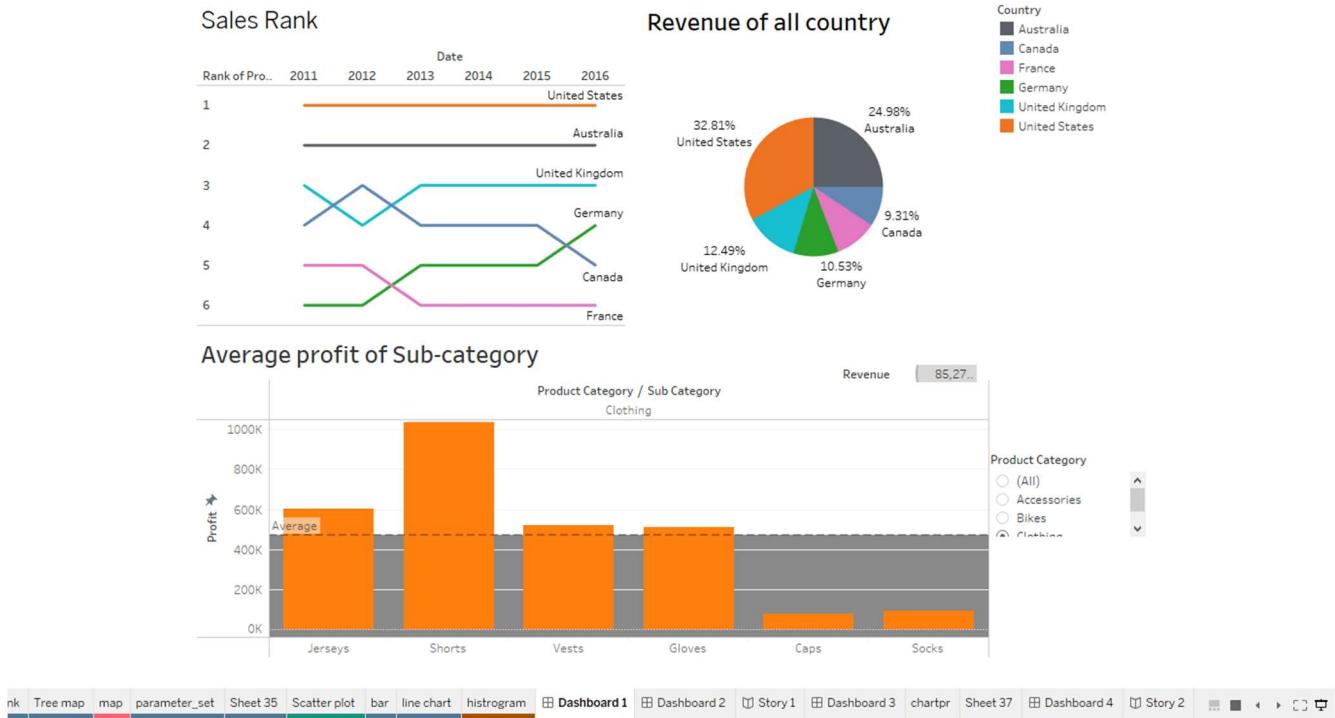


Gender wise sale distribution across countries

From the above visualization we can see a very slight difference of sales between male and female across all the countries.

France and Germany get most of the sales from female population whereas Australia, Canada, UK and USA get most of the sales from male population.

14. Dashboard



15. Story

