Visualization of Data using Tableau

1. Introduction

Big data analysis is achieving increasing attention in many industries, including Marketing. Visualization plays a significant role not only in showing the results of data analysis but also in whole process of data collecting, data reducing, selecting the most important factors, analyzing and sharing data.

In this note, we present a visualization of a dataset using Tableau as a business intelligence tool. The example data is not real because of data privacy protection and shows sales, profit, discount in different countries in three continents Europe, Asia and Africa in years 2010 to 2022. We can see the first 5 records of dataset as follows:

```
Data = pd.read_csv('Book1.csv')
print(Data.head())
```

```
  Continent  Country  Discount  2010sale  2011sale  2012sale  2013sale  \
0     Europ  Germany      0.02      2100      2190      3266      3500
1     Europ    Italy      0.03      4000      3500      3700      4100
2     Europ   France      0.06      1200      2100      2150      2700
3     Europ    Spain      0.01      1100      1700      2500      2400
4     Europ  Denmark      0.01      2000      4000      5000      5500

   2014sale  2015sale  2016sale  2017sale  2018sale  2019sale  2020sale  \
0      3100      5900      6000      6250      7100      7600      8000
1      4600      4000      3900      3100      4900      5000      5100
2      3600      4900      5700      6400      7100      9500     12000
3      2900      3000      3400      3500      3900      4200      4800
4      5600      5900      6300      6600      6700      7150      7340

   2021sale  2022sale  Profit
0      8300     10000  148000
1      6000      6400   98000
2     14500     15800  165000
3      5200      6000  105000
4      7590      8000  131000
```

2. Connecting to sample data

The sample data in our work includes 19 records during 2010-2022. This sample is Excel formatted data and can be connected via the connections section to the left of the Tableau window.

3. Results

Firstly, we are going to show a view of the position of the countries in our dataset (Figure1). By labeling the name of countries we can see the name of countries in the map.

Figure1. Position of the countries

Next, we want to pivot data from columns (2010sale-2022sale) to rows (pivot field sales and pivot field name). The new column replace the original columns that we selected to create the pivot variable. Sometimes, data is stored in a cross tab format and we need to pivot. Here, we see that the amount of sales during the years 2010-2022 has an increasingly behavior generally (Figure2).

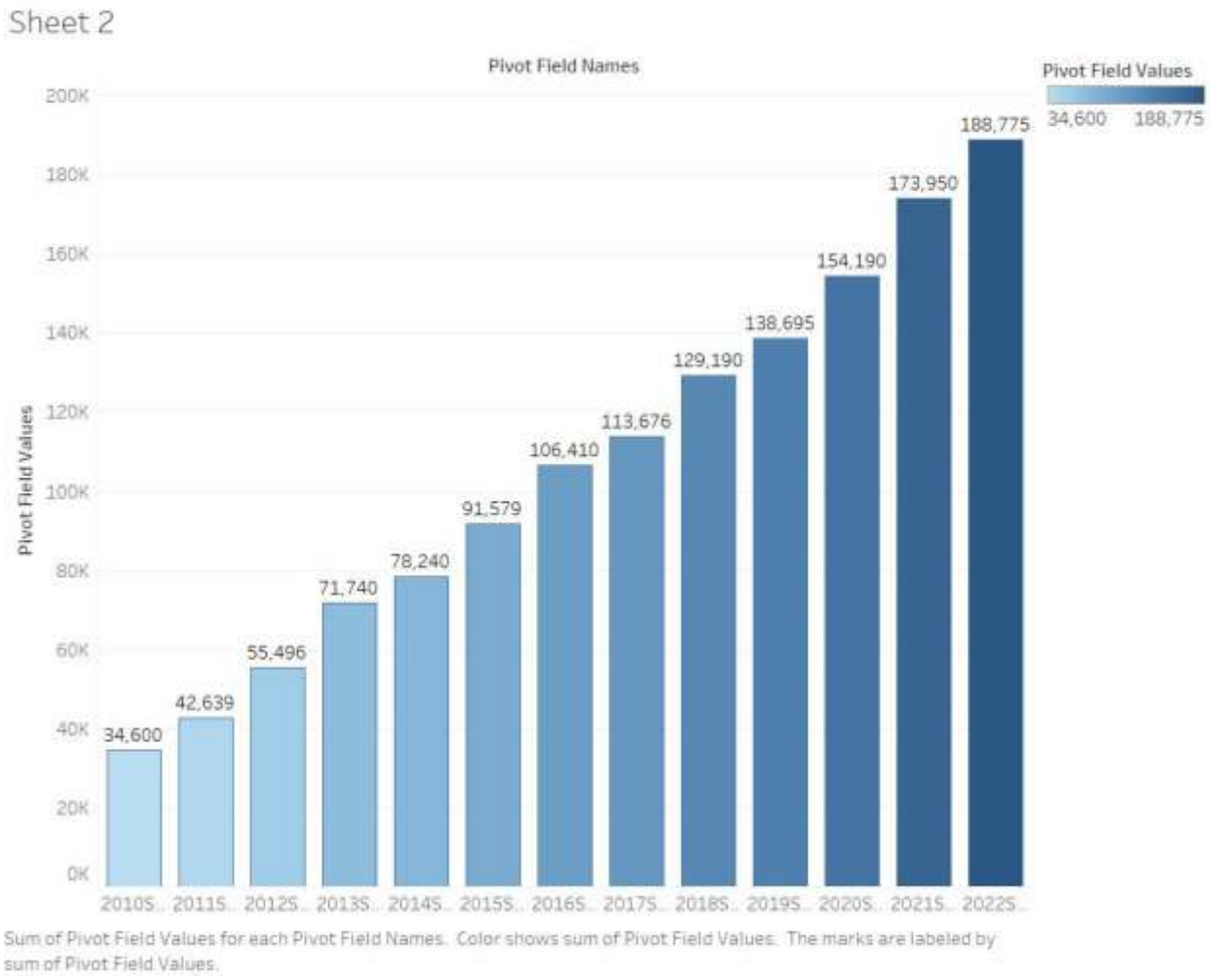

Figure2. Sales during 2010-2022

Consequently, we have provided a Bar chart in order to show that amount of sales in all countries in three different continents in the year 2010. As a result of labeling the variable 2010sale we get China and Italy have the maximum sales between countries. Whereas, Kenya has the least amount of sales equal to 270 (Figure3).

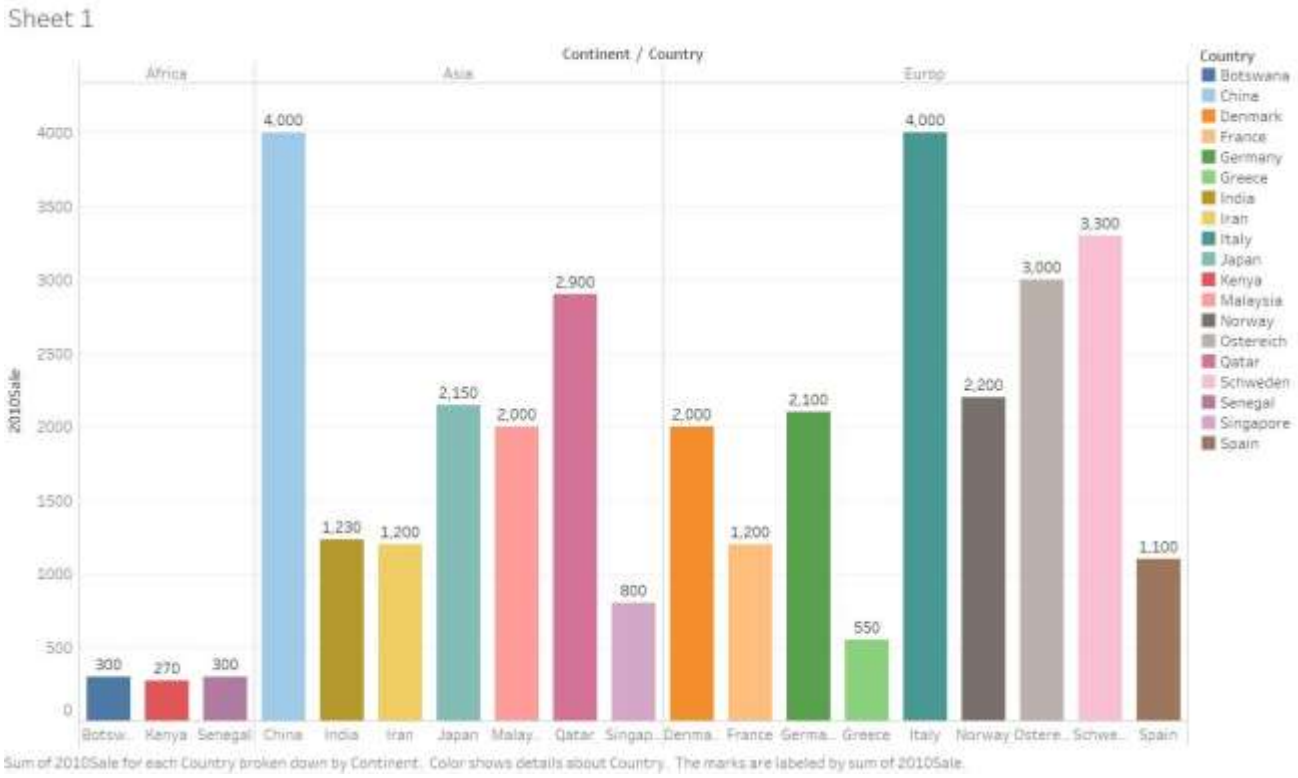

Figure3. The amount of sale in all countries in the year 2010

In Addition, Figure4 displays three Bar chart about the amount of sales, profit, discount in all countries together in the year 2010. It is interesting that Norway and Schweden had the largest profit, while they had not the largest discount. In other word, Norway and Schweden have 0.11 and 0.10 respectively. Also, it is remarkable that Kenya has not only the lowest amount of sale but also profit. Meanwhile, India gives the maximum discount. But it has not a good sale and profit (Figure4).

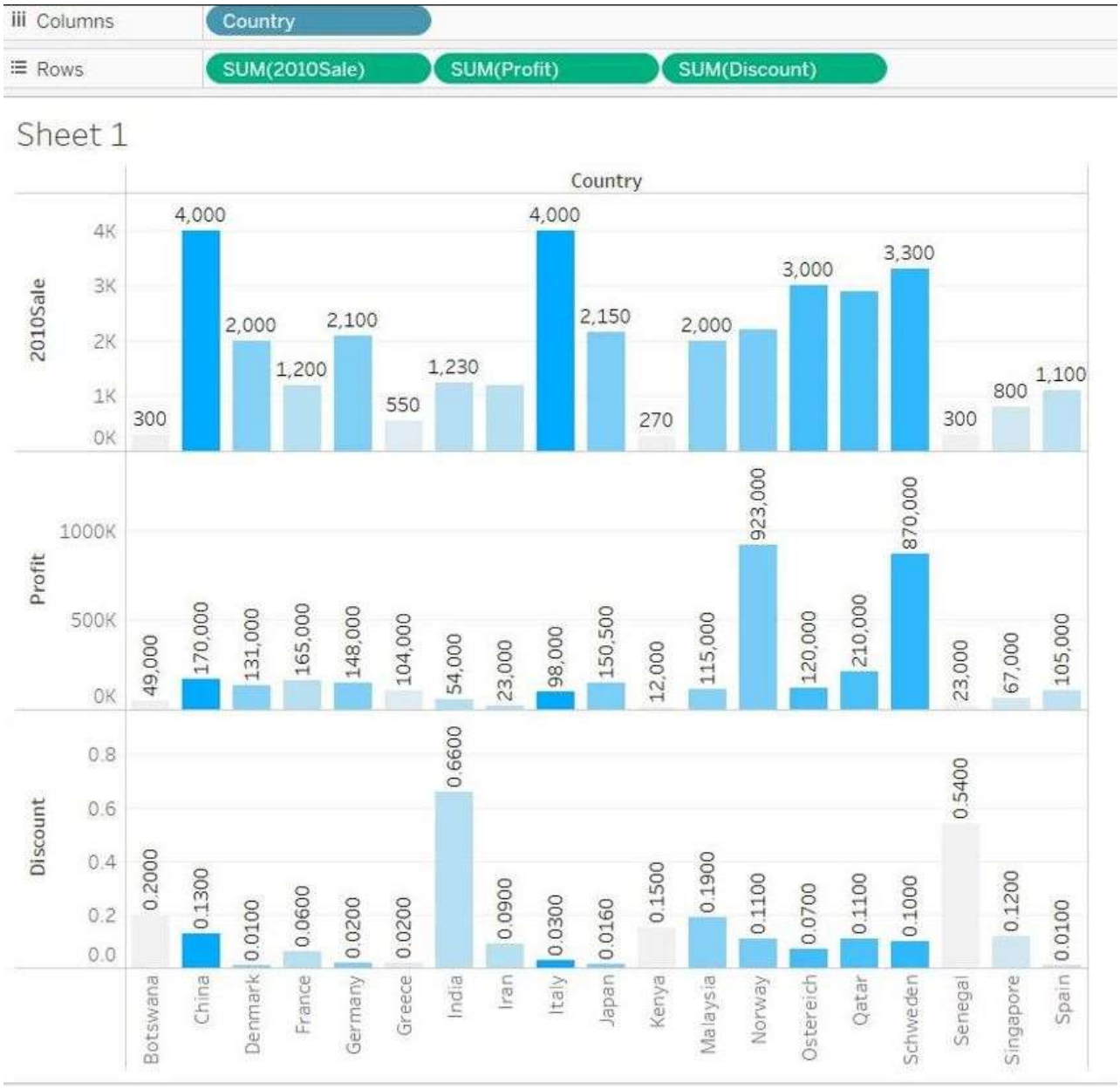

Figure4. Comparing sale, profit and discount in all countries in the year 2010

Last but not least, Figure5 shows an other Presentation of Figure4 in form of line chart.

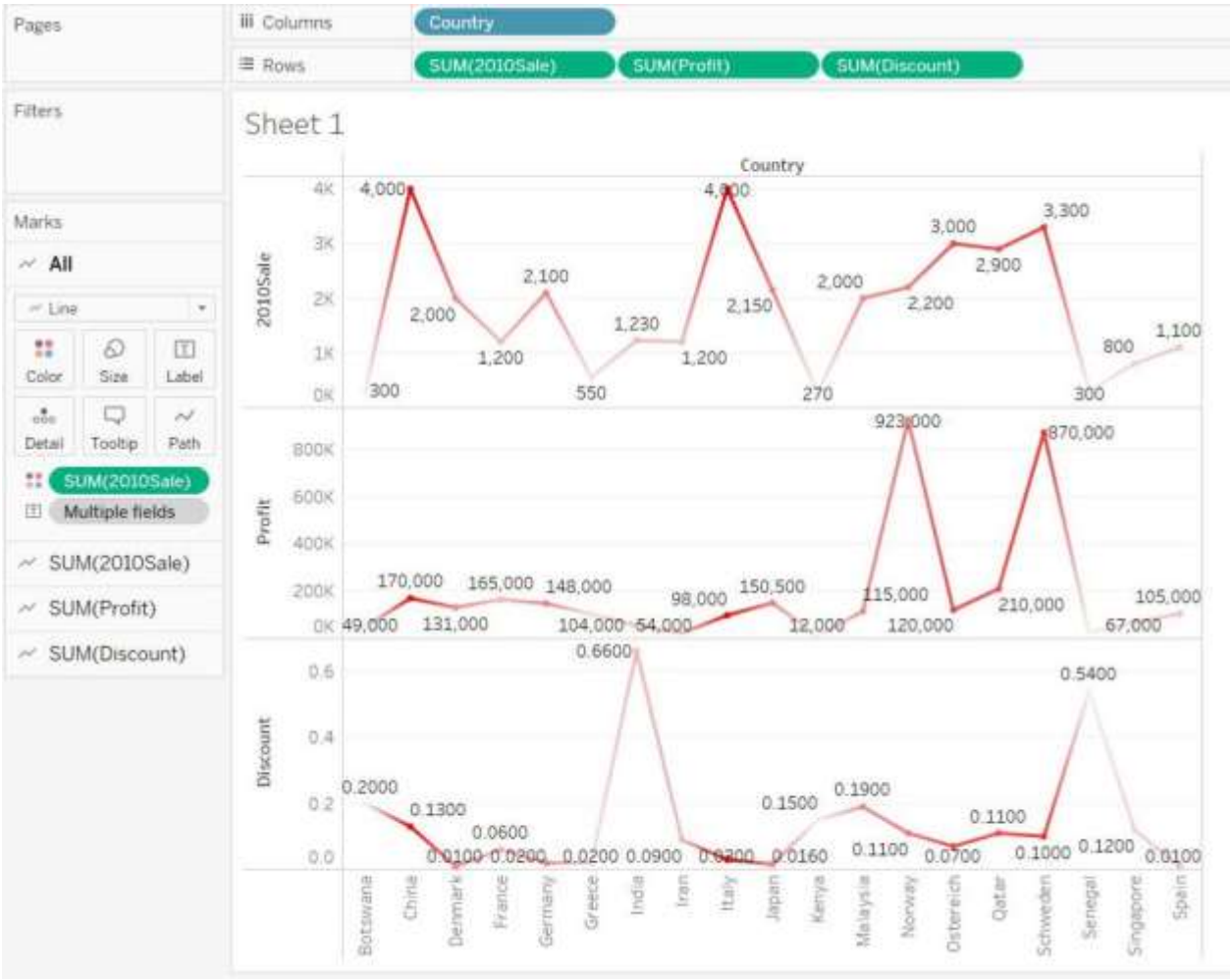

Figure5. Line chart of sale, profit and discount in the year 2010