Walmart sales Analysis

*Divya Velugula, 11340384*

*Department of Data Science, University of North Texas*[*DivyaSreeVelugula@my.unt.edu*](mailto:DivyaSreeVelugula@my.unt.edu)

**INTRODUCTION:**

**Problem:**

Walmart is the world’s biggest retailer with over 20,000 stores in 28 countries. Walmart is the largest retail corporation of discount department and warehouse stores in the world. Though it now operates in 27 countries, Walmart had humble beginnings in the 1960s in Bentonville, Arkansas. The retail chain offers low prices and a wide selection of products, which has given Walmart an edge over the competition. But they are still running behind few companies in sales and profits.

Data Set Used: WMartEU2.csv (Downloaded from Kaggle)

**Motivation**:

The reason I opted this problem is that I personally like Walmart shopping. Nowadays, I see Walmart is not as busy as it used to be. So, I studied about it in various websites . Then I observed that their sales & profits are not up to the mark & decided to work on this issue to find the reason behind this.

**Hypothesis**:

To analyze the problem, few hypothesis questions were built using Exploratory Data Analysis. They are listed below:

1. Report which sub-category is having the maximum quantity and in which region.
2. Report which state has the maximum and minimum profit ratio.
3. Report forecasting for the sales and profit of year 2015.

**RELATED WORK:**

A complete review of visual quality measures can be found in :

1. The work of Edward Tufte and Graphic Press – “Edward Tufte is a statistician and artist, and Professor Emeritus of Political Science, Statistics, and Computer Science at Yale University. He wrote, designed, and self-published 4 classic books on data visualization. The New York Times described ET as the "Leonardo da Vinci of data," and Bloomberg as the "Galileo of graphics." He is now completing a book Meaning, Space, Models, Data, Truth, and constructing a 234-acre tree farm and sculpture park in northwest Connecticut, which will show his artworks and remain open space in perpetuity. He founded Graphics Press, ET Modern Gallery/Studio, and Hogpen Hill Farms.” [1]
2. [Information Visualization Research Projects that Would Benefit Practitioners](http://www.perceptualedge.com/blog/?p=2258)[2]
3. Process and Pitfalls in Writing Information Visualization Research Papers – “This paper is structured around a chronological model of the information visualization research process. I argue that a project should begin with a careful consideration of the type of paper that is the desired outcome, in order to avoid the pitfalls of unconvincing validation approaches. Research projects that involve the design of a new visual encoding would beneﬁt from checking for several middle-stagepitfallsinunjustiﬁedorinappropropriateencodingchoices.Another critical checkpoint is the late stage of the project, after the bulk of the work is done, but before diving into writing up results. At this point, you should consider both strategic pitfalls about the high-level structure of the entire paper, tactical pitfalls that aﬀect one or a few sections, and possible pitfalls in the speciﬁcs of your approach to the results section. At a ﬁnal stage, when there is a complete paper draft, you can check for lower-level pitfalls of writing style and avoid submission-time pitfalls.” [3]

**METHODS & RESULTS:**

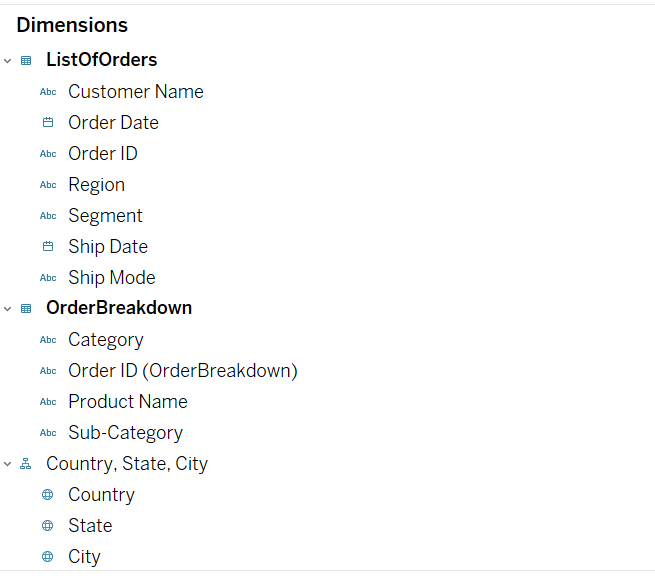
**Data Cleansing:**

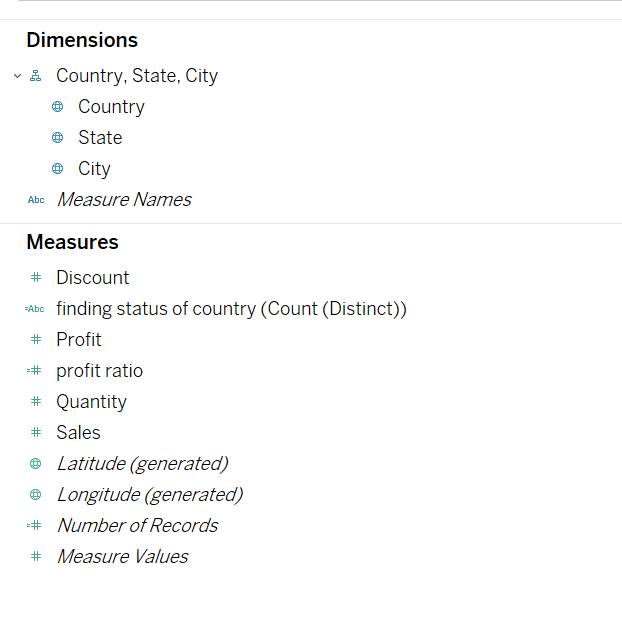
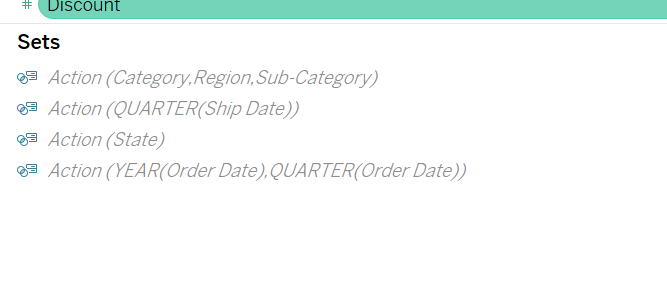
* Initial phase I used an excel sheet containing the raw data and directly uploaded the same on the tableau to build the Dash boards.
* I used data source filter to filter the raw data
* I have performed inner join to extract.
* I have created Measures, Dimensions & some Sets for easy analysis.
* I did some transformations(Aggregations, Sets & Parameters).

**Charts used:**

* Bar and Line
* Waterfall Chart
* Dual Axis Graph
* Geo Map

**Dimensions & measures:**

****

** **

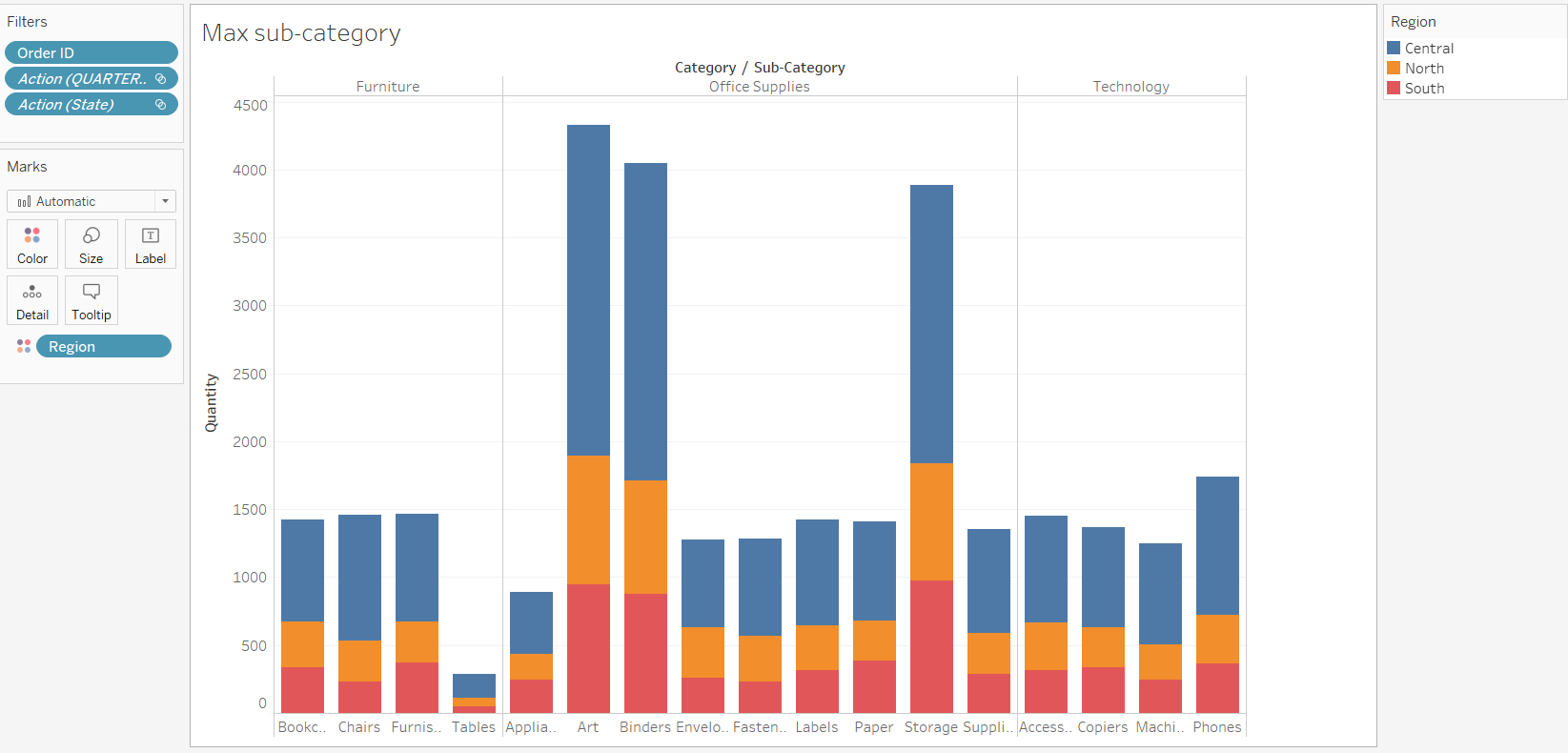
**Data analysis using Visualization:**

1. Report which sub-category is having the maximum ordered quantity and in which region.

* I developed Bar graph on a single chart.
* Bar Graph depicts region wise sub-category having maximum order quantity .
* Colors along with the color legend present on the right corner helps us to differentiate between the region.

**Legends Used:**

* Legend Name: Region
* Legend Type : Colors
* Legend UI : Box
* Legend position: Right hand top corner



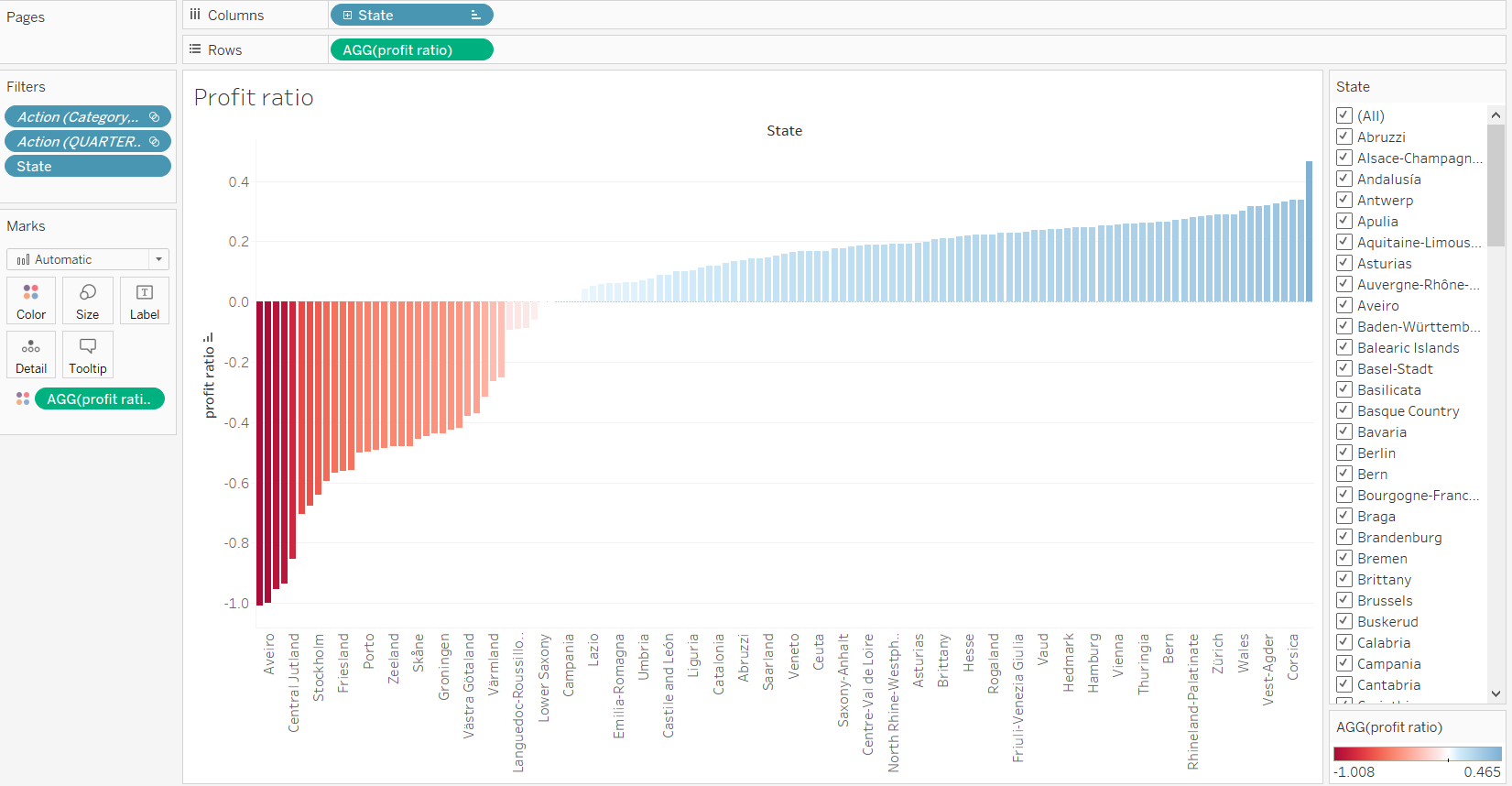
To access the above worksheet please open the .twbx file with sheet name “max sub category” on the bottom left hand corner using tableau. Here, Office supplies are the most sold items amongst all and tables in furniture category are the least sold items. In most of the cases, the Central region has higher sales.

1. Report Which state has the maximum and minimum profit ratio?

* I developed a Bar chart with both positive and negative value .
* This graph depicts profit ratio state wise from negative to positive ratio.
* I used Red-blue-white diverging palette to represent the chart.

**Legends Used:**

* Legend Name: AGG(Profit ratio)
* Legend Type : Colors
* Legend UI : Box
* Legend position: Right hand top corner



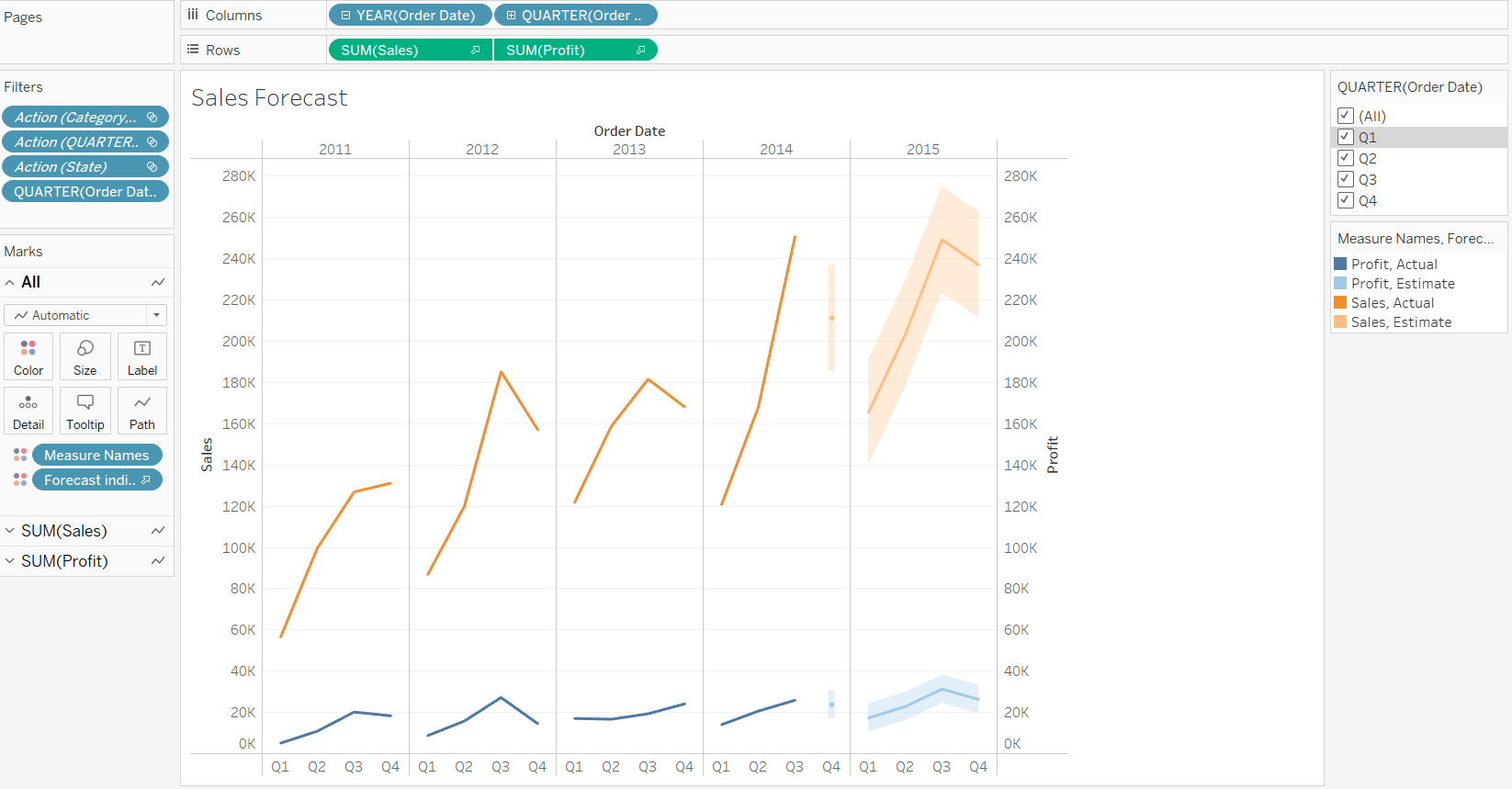
* Here, in the above figure, we can observe an interactive visualization of profit ratios with respect to different states. The states listed in the right corner can be selected as per the requirement and we can see the profit ratios of the selected states only. This graph illustrates the profit ratio for different states.
* Least profit ratio: Zealand with -1.008
* Highest profit ratio: Buskerud with 0.465
* To access the above work sheet please open the wmart\_Analysis.twbx file with sheet name “profit ratio” on the bottom left hand corner using tableau.

1. Report forecasting for the sales and profit of year 2015.

* I developed line graph on a single chart using dual axis where I used 2 measure Profits& Sales.
* 2 forecast indicator Actual & Estimate indicator I have used in this chart.
* For 2 different indicators, I have used 2 different color palettes.
* To make visualization more effective, I segregate the result into quarter wise.

**Legends Used:**

* Legend Name: Measure Names, Forecast
* Legend Type : Colors
* Legend UI : Box
* Legend position: Right hand top corner

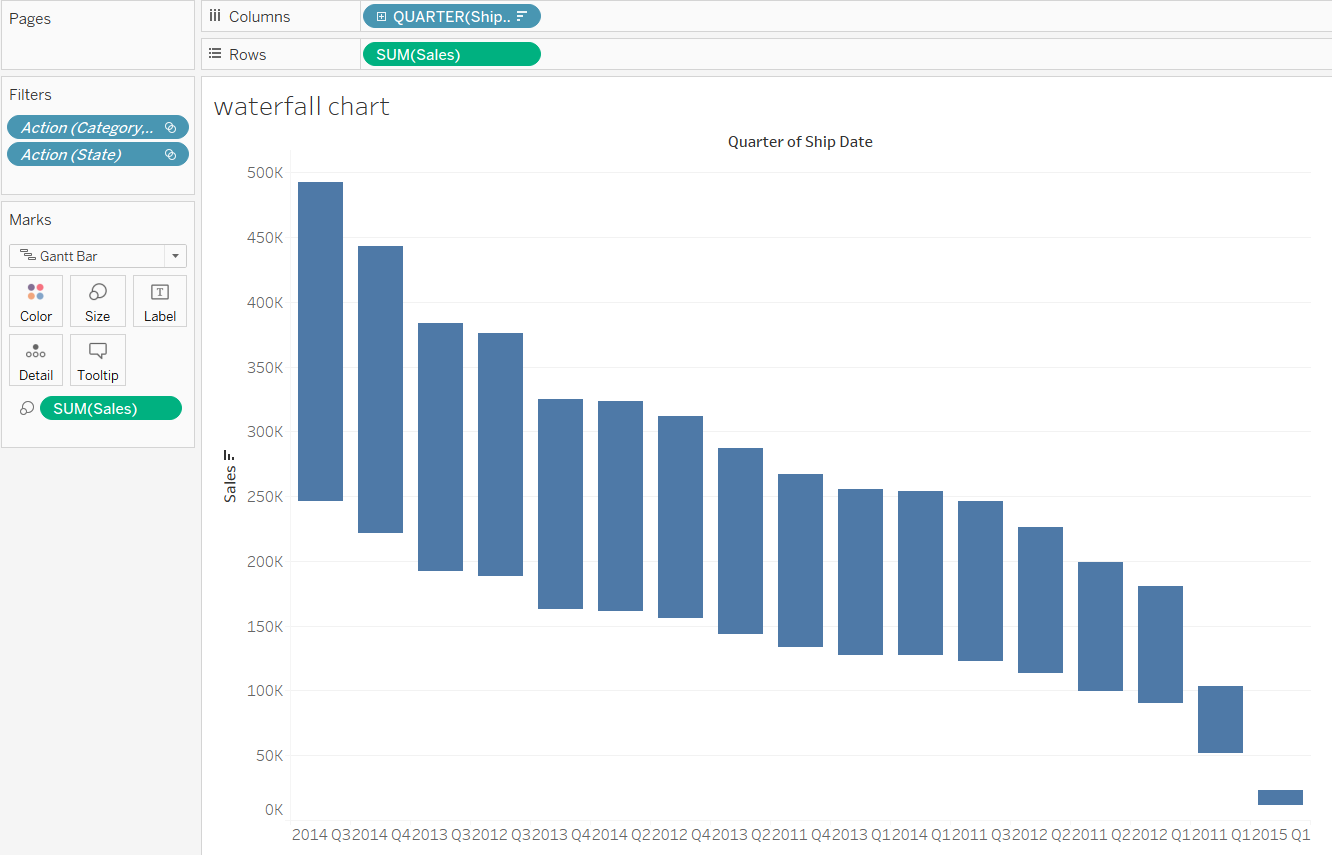


The above graph can be visualized quarter-wise interactively to observe the sales & profits every year. To access the above work sheet please open the wmart\_Analysis.twbx file with sheet name “Sales Forecast” on the bottom left hand corner using tableau. To access the above work sheet please open the wmart\_Analysis.twbx file with sheet name “Sales Forecast” on the bottom left hand corner using tableau.

* Actual Sales: High in the year 2014 for 3rd quarter
* Actual Profits: High in 2012 for 2nd quarter
* Sales are increasing every year, but profits are nearly constant.

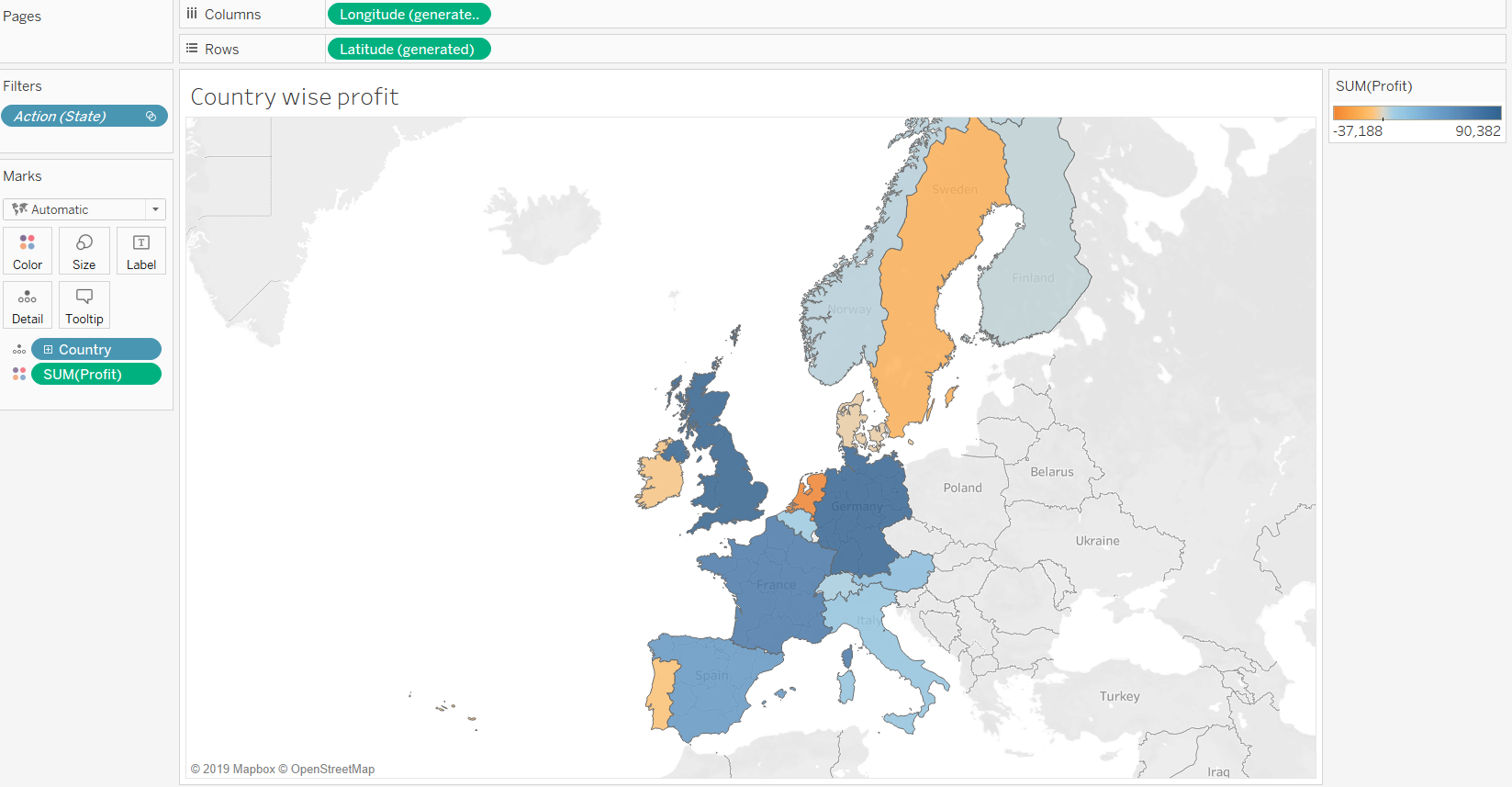
Apart from the above-mentioned hypothesis, I have plotted 2 more chart for better analysis.

1. For analysis of Quarter wise sales data, I have used waterfall chart.



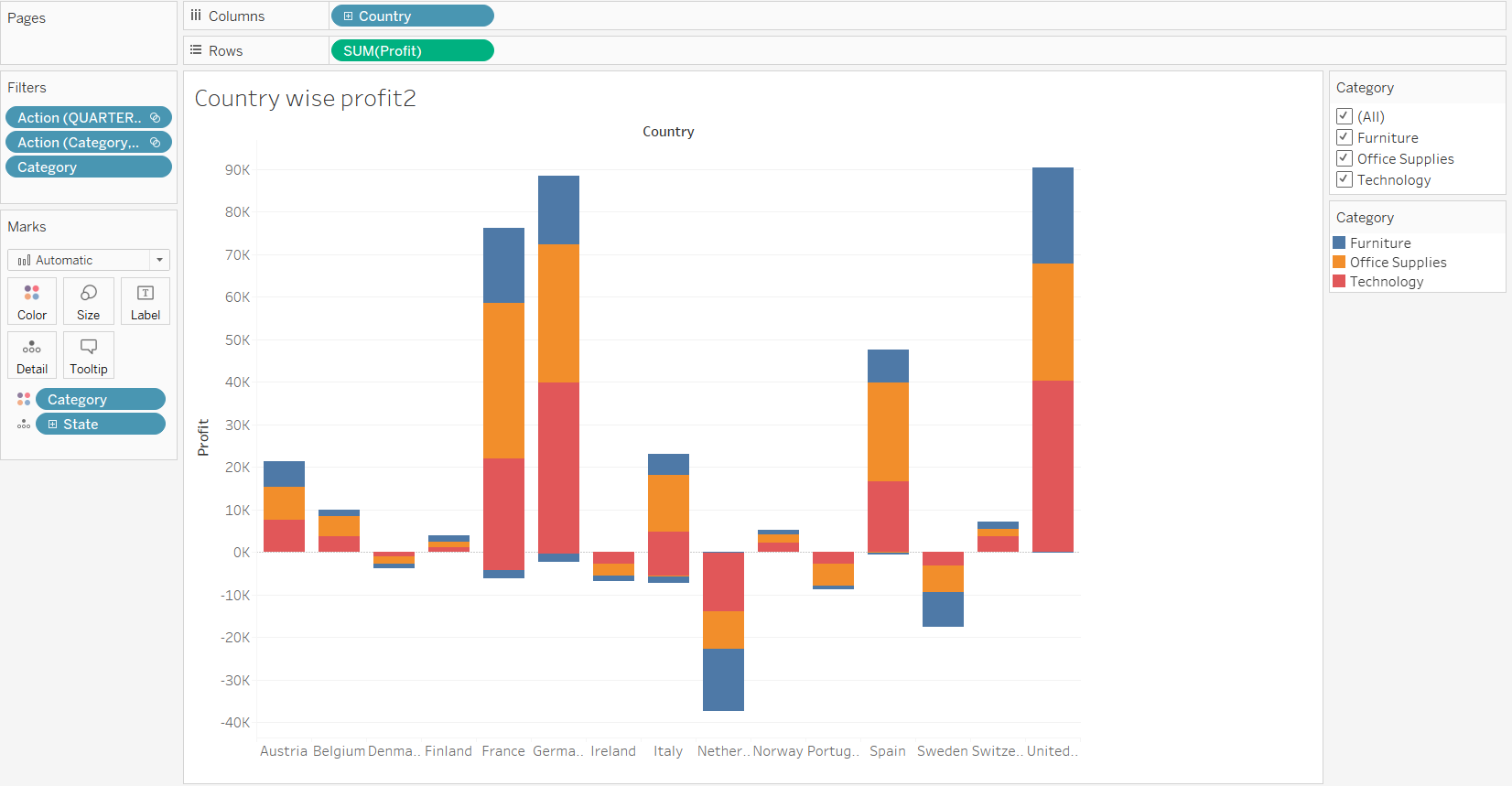
In 2014, it has the highest sales of 246,243 for third quarter and least sales of 11,420 in the year 2015 first quarter.

1. I have used Symbols map with Divergence color palette for Country wise profit calculation.

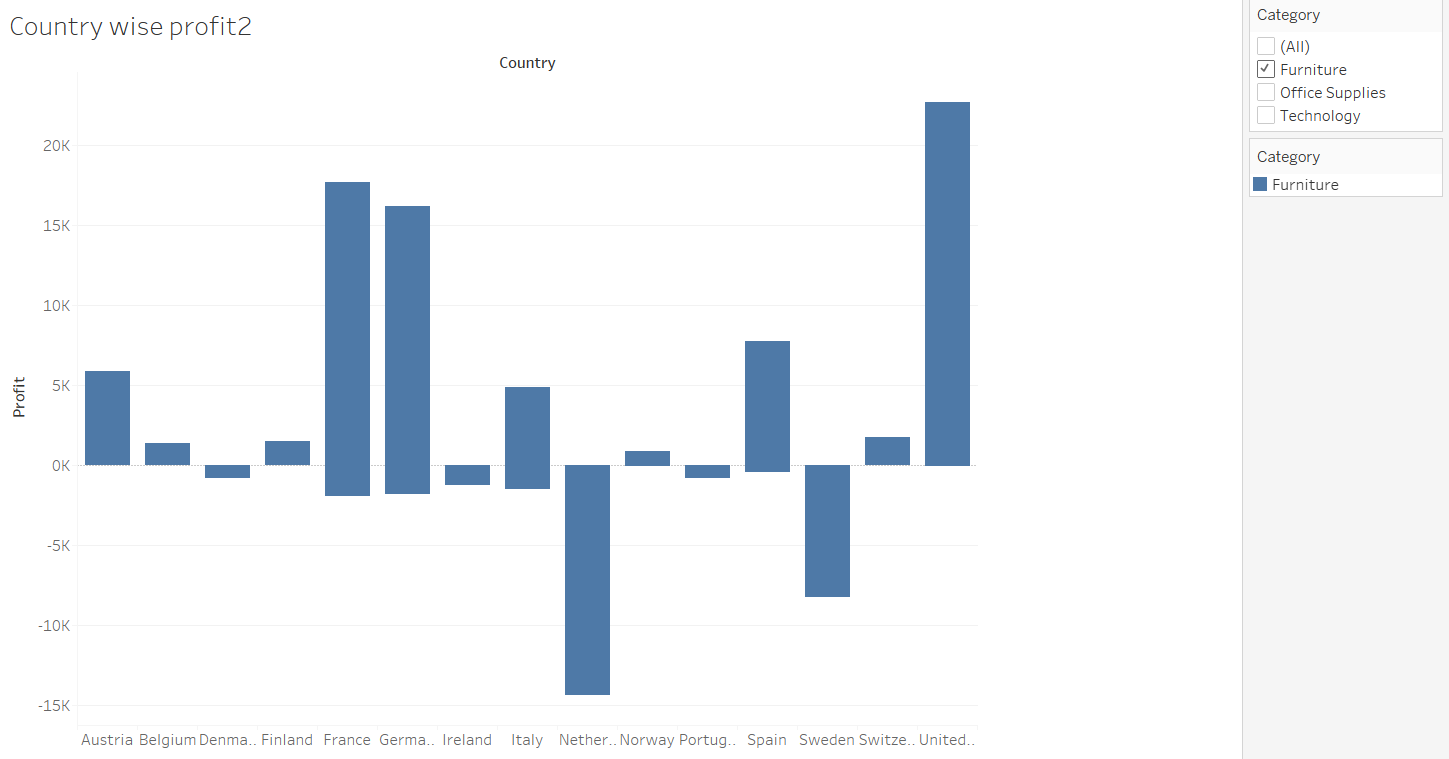


From the above graph, we can observe the results that the severity of profit is indicated from orange to blue as shown in the right most panel. Based on this, we can conclude that UK has the highest profits of around 90,000 and Netherlands has the least profits of around -37,000.

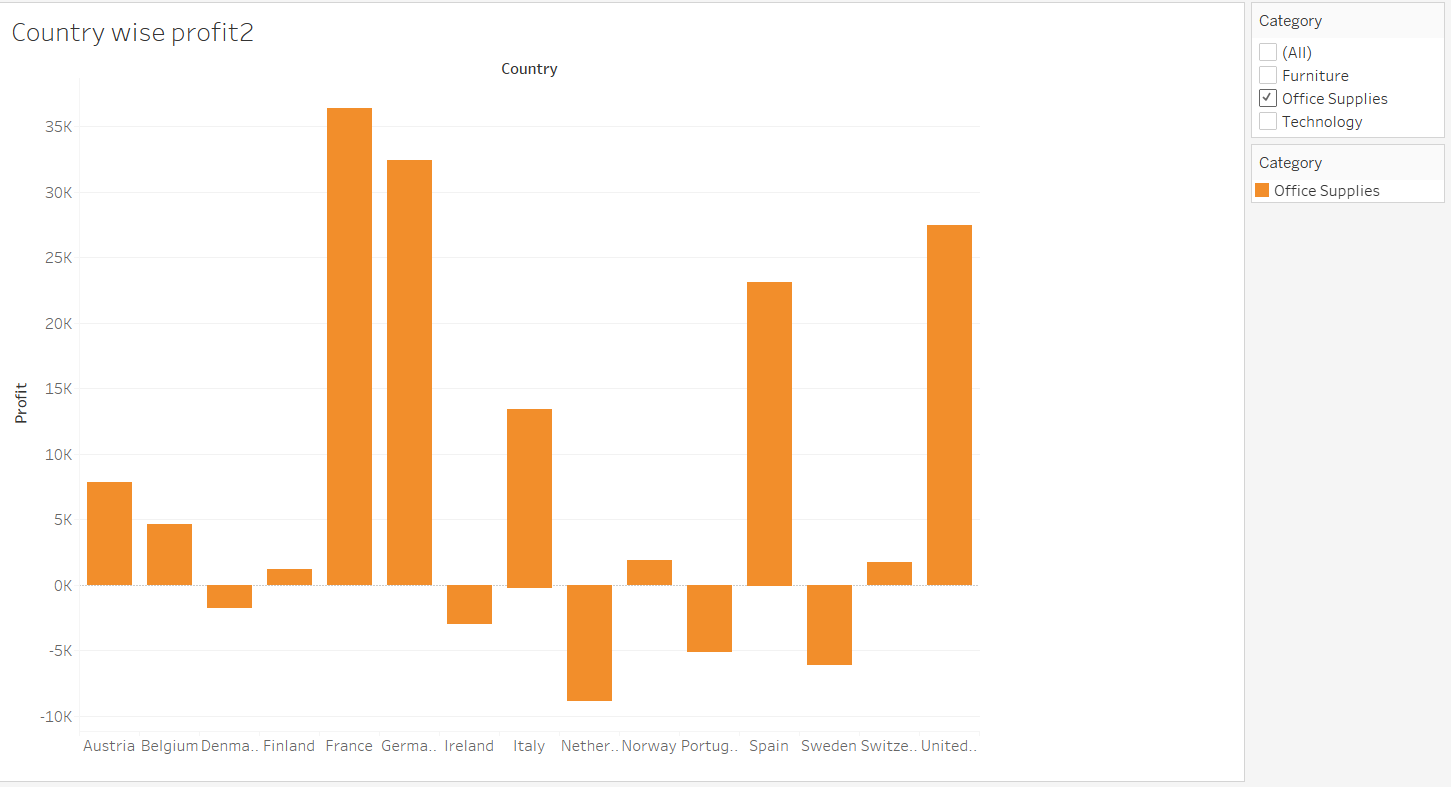
1. I have used stacked bar graph to get Country-wise profits for different categories of products.



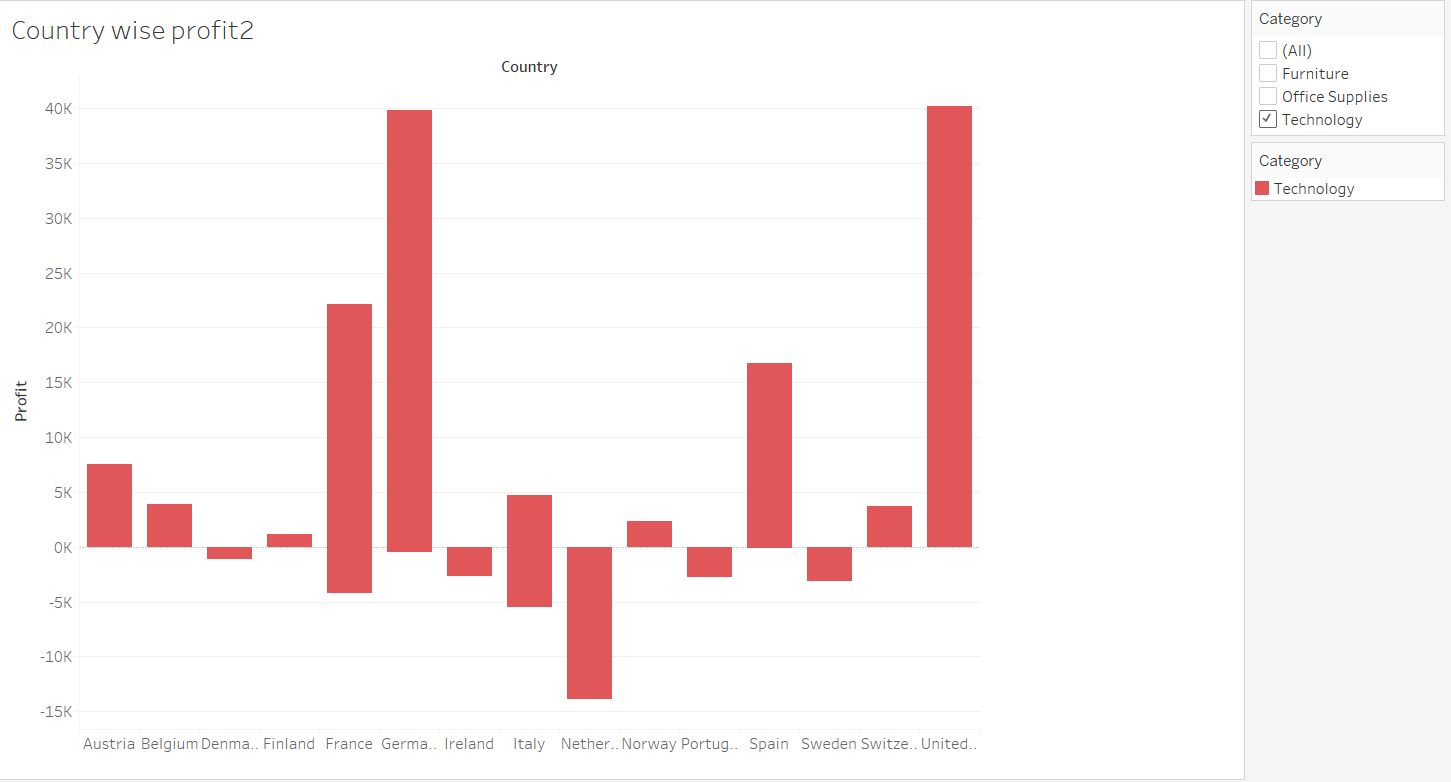
From the above graph, we can clearly say that UK has the highest profits overall and Netherlands has the least profits overall. When we observe the results category-wise, UK & Germany has highest sales for technology supplies, France has the highest sales for office supplies and UK has the highest sales for Furniture. We can interactively select the different category products and view the concerned results.



*Fig: Interactive results of Country-wise profits for Furniture*

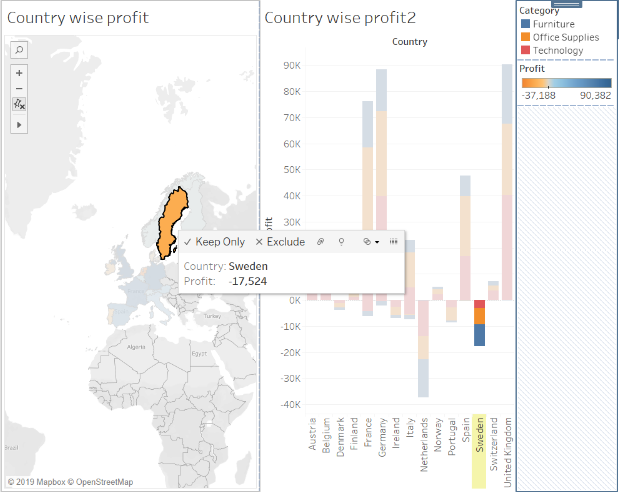


*Fig: Interactive results of Country-wise profits for Office Supplies*



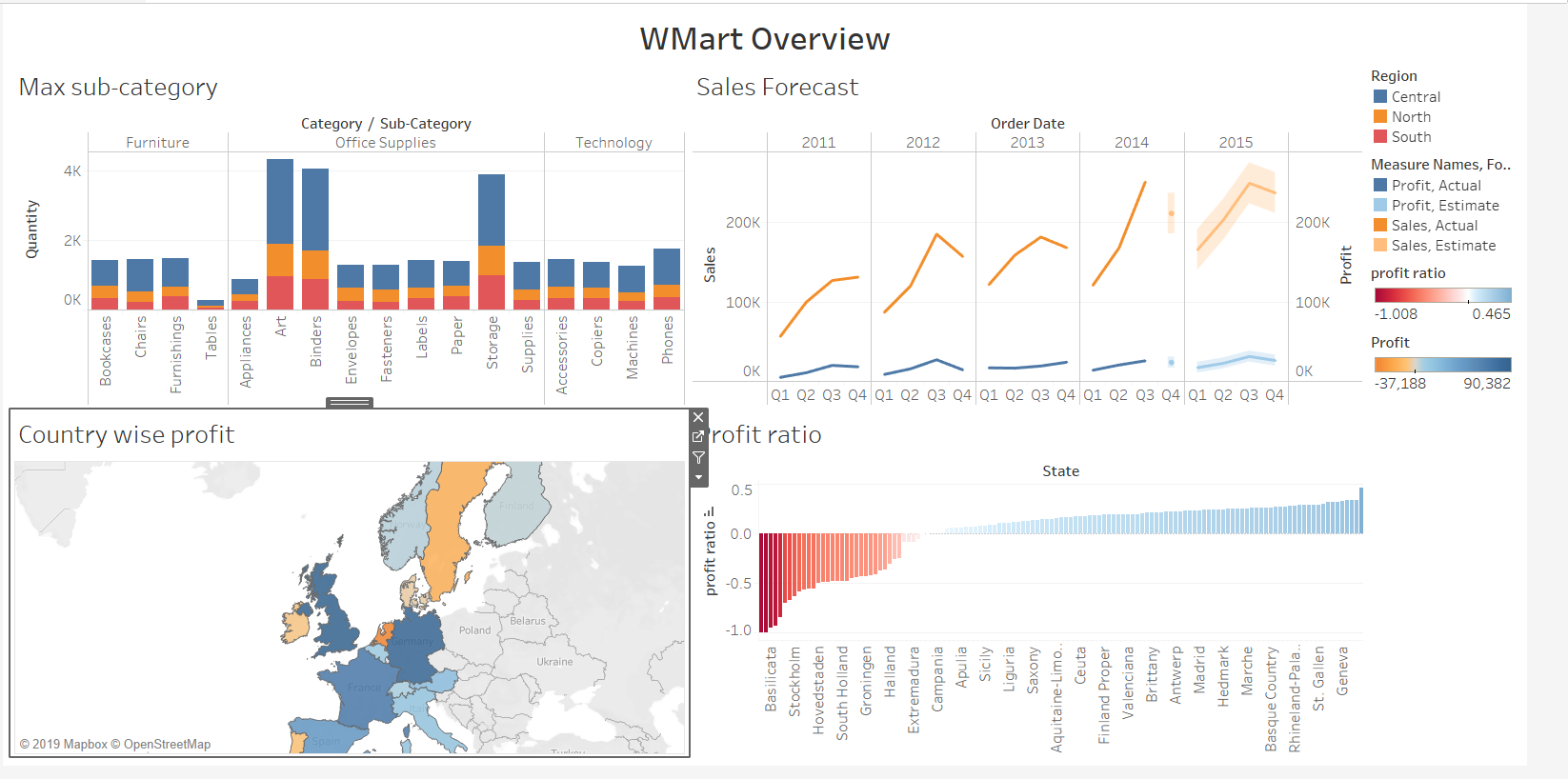
*Fig: Interactive results of Country-wise profits for Technology*

1. Finally, I made an interactive dashboard to view the results of profits of different products for a specific country.



From the above figure, the first sheet is known as source sheet and the second one is the target sheet. To get the interactive results, we can select a specific country from the source sheet and the corresponding profits will be highlighted from the target sheet.

**FINAL DASHBOARD:**



**DISCUSSION & FUTURE WORK:**

* From the visualizations, we can easily identify the problems of organizational sales based on specific product in a specific quarter & year, also based on regions.
* As Walmart has higher sales for Office supplies, it should enhance the availability for these products and provide more offers on the supplies with less sales, which improves the overall sales & profits.
* With more data, we need to analyze the requirements for the countries with less sales and should focus on increasing those particular products.

# **REFERENCES:**

|  |  |
| --- | --- |
| [1] | E. Tufte, "The work of Edward Tufte and Graphics Press," 2019. [Online]. Available: https://www.edwardtufte.com/tufte/. |
| [2] | S. few, "Visual Business Intelligence," 2019. [Online]. Available: http://www.perceptualedge.com/blog/?p=2258. |
| [3] | T. Munzner, "Process and Pitfalls in Writing Information Visualization Research Papers," p. 20, 2019. |