**BISM 2202 Data Analytics and Information Management**

|  |  |
| --- | --- |
| Ashar Artani | 44279284 |
| Junting Li | 44033048 |
| Mosa Zuley del Carmen Nakada Araujo | 44432397 |
| Chenxin Yang | 44620420 |
| Yuru Zhou | 45213931 |

Table of Contents

[1. Executive Summary 3](#_Toc527297102)

[2. Business Analytics Framework 3](#_Toc527297103)

[3. Design of a Dimensional Mode 5](#_Toc527297104)

[3.1 Sales 5](#_Toc527297105)

[3.2 Purchase 6](#_Toc527297106)

[4. Implementation of the Dimensional Model 8](#_Toc527297107)

[4.1 Data integration (Pentaho) 8](#_Toc527297108)

[4.2 Data visualisation (Qlik) 9](#_Toc527297109)

[5. Generate Business Insights and Strategies 12](#_Toc527297110)

[5.1 Who are the key customers 12](#_Toc527297111)

[5.2 Which products are the most profitable 13](#_Toc527297112)

[5.3 Which sales territories are the most profitable 14](#_Toc527297113)

[5.4. Which time periods are the most profitable 15](#_Toc527297114)

[5.5 Which vendors should be considered for replacement 16](#_Toc527297115)

[6. Reference 17](#_Toc527297116)

[7. Appendix 18](#_Toc527297117)

[7.1 Data Dictionary 18](#_Toc527297118)

[7.2 Work Breakdown 19](#_Toc527297119)

# **Executive Summary**

This report provides an analysis and evaluation of the current and prospective profitability of WorldView Bike Company. With the use of ETL and dashboard, results of data analysed show that the most profitable product, territory, time period, key customers and the less important vendors. The dashboards also illustrate relationship and results between different dimensions and measures clearly.

The report finds the prospects of the company in its current position are not as profitable as the previous year as it has a decreasing revenue each year. The major areas of weakness require further investigation and remedial action by management. Recommendations discussed includes reduce unfavorable products, focus on the sales to our key customers and increase the inventory for most profitable products sold in each territory.

In conclusion, BA project is recommended as it is used by companies committed to data-driven decision-making. It helps WorldView Bike Company to gain insights that inform business decisions and can be used to automate and optimize business processes. Data-driven companies treat their data as a corporate asset and leverage it for a competitive advantage. Successful business analytics depends on [data quality](https://searchdatamanagement.techtarget.com/definition/data-quality), skilled analysts who understand the technologies and the business, and an organizational commitment to data-driven decision-making.

# **Business Analytics Framework**

According to Galetto (2016), business analytics is used to study data through the use of statistical and operational analysis which helps to predict the formation of models and the application of optimization techniques that allows organisations to communicate these results to customers and work partners. In other words, business analytics (BA) is the practice of iterative, methodical exploration of an organization's data, with an emphasis on statistical analysis.

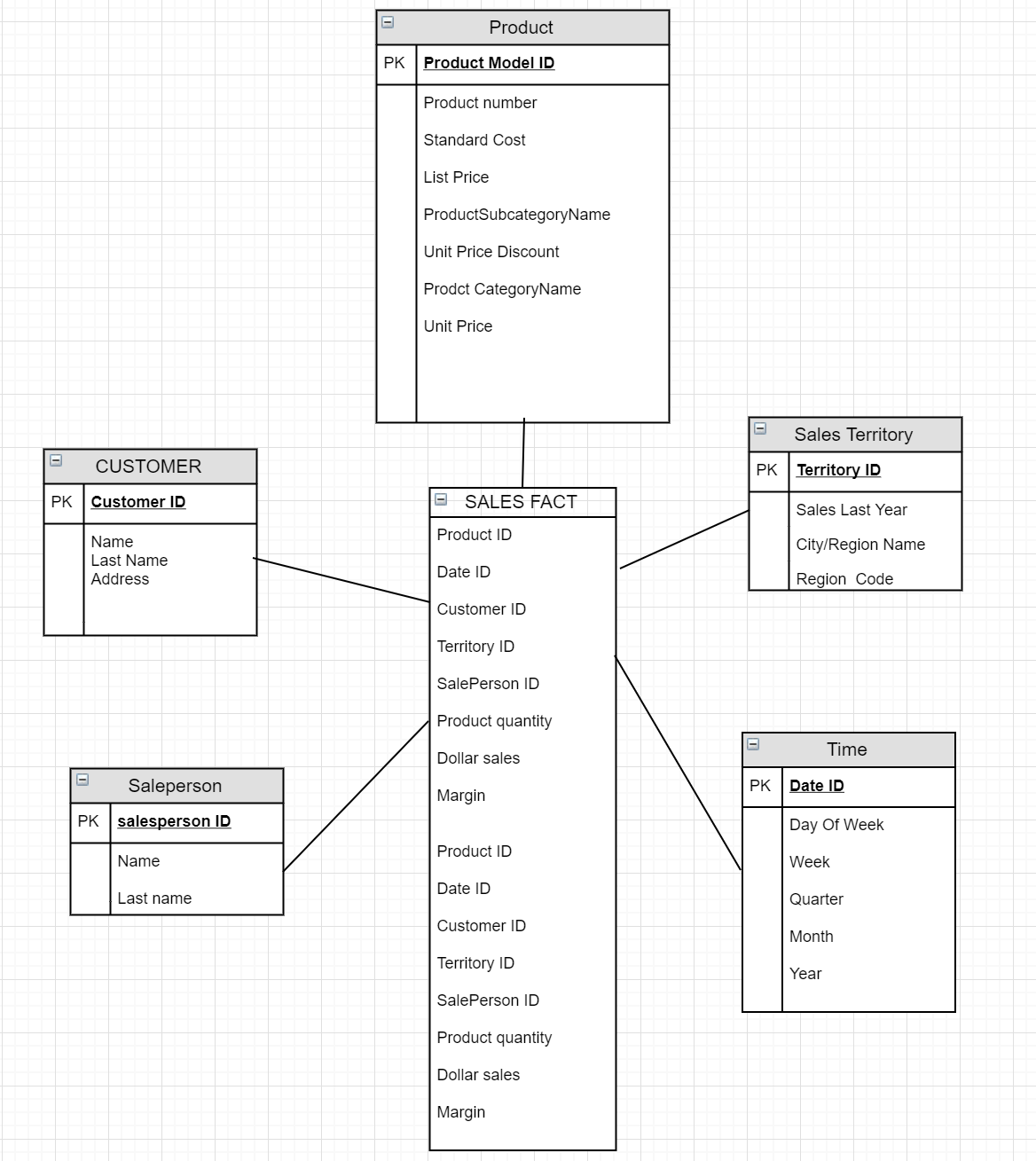
Business analytics is also used by companies committed to data-driven decision-making. Firstly, it provides clear insight data visualization that allows business data to be extracted, analyzed and organized, then visualized by a graph or chart. This makes the raw data easier to understand for the shareholders, managers or customers, and this in turn reduces the communication cost due to requiring specialization knowledge. Secondly, this methodology helps to quantify business values. Each company has its own mission statement to serve as the basis for market strategy and development, which can make it difficult to determine whether the mission has been followed during the implementation of strategy. BA can help to quantify the business values of an organisation and define their goals for employees to follow, for instance, using staff evaluations. This helps a company by improving their analytical processes as well as making better decisions. Thirdly, BA are used to ingrain smart decision-making into the standard operation of a business. When a company continually sources quality and relevant data, it definitely contributes towards making the best possible decision available at the time. If a company strives for optimized use of analytics and shares data with the team, then is data analysis can serve a large variety of business areas and unify the organisation.

Therefore, BA brings many advantages for the Worldview bike company to enhance strategy, such as clearer insight and quantifying business values. Designing a Business Analytics Framework would greatly boost the capabilities of Worldview and allow internal and external stakeholders to measure the effectiveness of strategies and decisions.

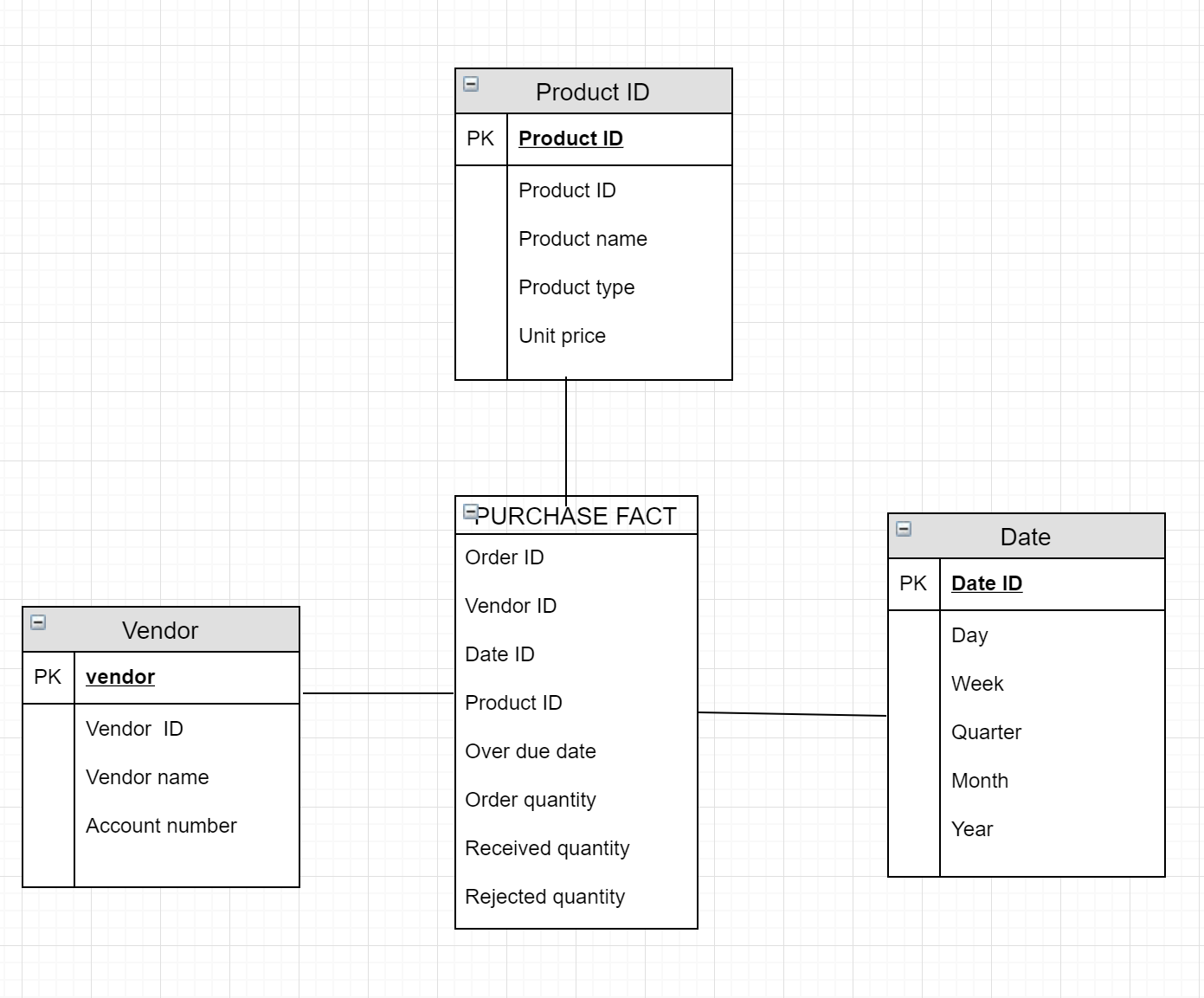
Worldview bikes has two operational systems, a purchasing system and a sales system, which means they separately have two databases for data storage. The next step is data stage, and ETL process applies in this area. Operational data from Worldview will be extracted, transformed and loaded into Data Warehouse to be integrated and cleansed. Then, Data Warehouse provides data for analytics and reporting to satisfy different analytics needs for different stakeholders, such as OLAP, data mining and query systems.

# **Design of a Dimensional Mode**

## **3.1 Sales**



## **3.2 Purchase**



As (Adamson, 2013) stated, the star schema groups the information into business categories being an effective data warehouse model, making the process of navigation and extraction simple, reducing errors during the process of analysis, and using “business intelligence,” which will better inform business decision making.

Our group chose to make two star schemas on two of the basic operational activities of the business; one of sales and the other purchases. These form “central tables” on their respective schema (Fact table), and both of these tables contain “numerical measurements” (Adamson, 2013) on the respective business activity that occur over time. As an example, in the first case the object of analysis is the sales (bicycles and accessories), our Fact table represents the profit of the company “Worldview bike” in Dollars (total sales minus total costs). On the other hand, in our second star schema “Purchases,” in the Fact table we find other numerical measures such as overdue date (due date of delivery minus the actual delivery date).

Each Fact table will also have dimensional keys, which represent each dimensional table created. As Chaudhuri & Dayal (1997) states, dimensional keys works as foreign keys in the fact table and are primary keys inside every dimensional table, called “surrogate keys”. In this sense, the first and most important step is to define the level of detail our tables require, granularity (Adamson, 2013),.

The granularity of fact table is shaped by the “list of dimensions,” in both of our star schemas the granularity is in a “transaction level” (Martyn, 2004).

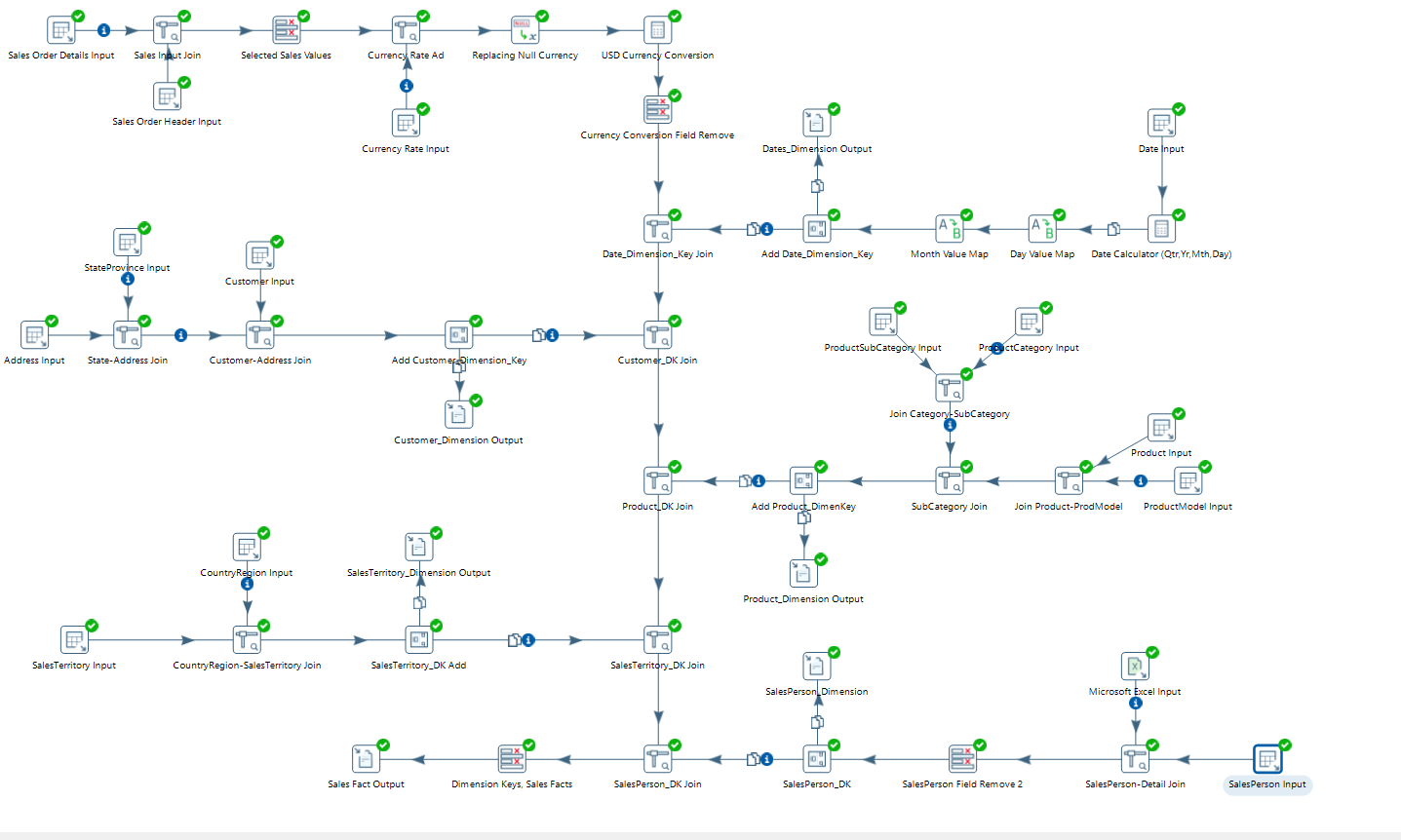
On our first star schema, labelled “Sales”, we have the dimensions; “Product”, “Sales territory”, “Date”, “Customer” and “Salesperson”, As Adamson (2013) states, Dimensions are points of view from which the activity can be analysed. These are categories that describe the context in which the measures (sales, profit) will be analysed, so we want to find sales results by which consumers they originated, and in which territory and by which seller they were made. Every single one of these dimensions have a primary key which is the dimensional key in the Fact table. In the second inside the “purchase fact” we have “purchase order ID”, “product ID”, “vendor ID” and “date ID”, these are at the same time the dimensions.

Each of these dimensions have attributes or levels of granularity which contribute more detail. For example, for a customer dimension table we have categorized data with all the information we need; name and address. In “Sales territory” the granularity is defined up to city. This reduces data redundancy, for example in the fact table we will only have Sales territory ID and not “City ID” or “country ID” Because they are included in “Sales territory ID.”

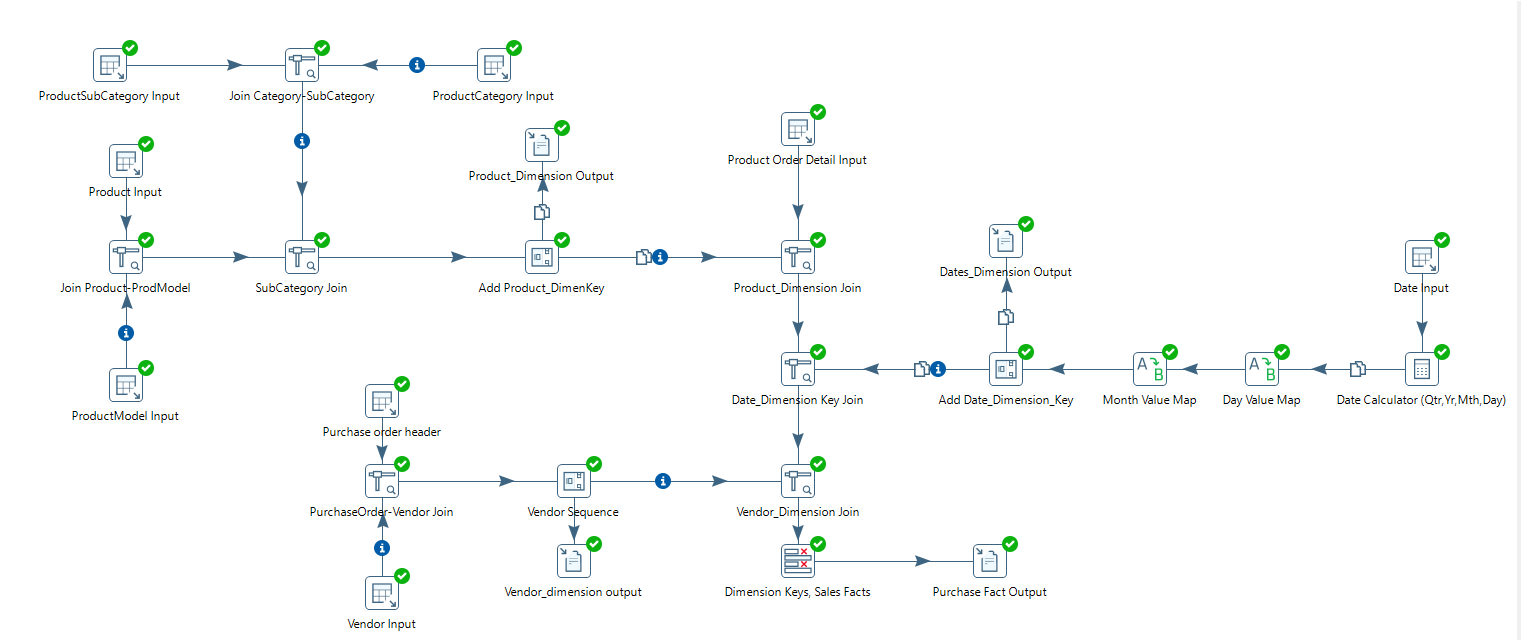
In both star schemas we need a time dimension and as Chaudhuri & Dayal (1997) stated its granularity and hierarchy depends on the dynamics of the business. For “Sales”, we consider the minimum level of granularity to be a “day quota”, since for the analysis of some data like “online sales”, we need information not about individual transactions/sales but of the daily sales. That being said, it is possible that other information on sales quotas will be needed at the level of quarterly or monthly intervals.

# **Implementation of the Dimensional Model**

## **4.1 Data integration (Pentaho)**

****

A table input was used to connect the university database to retrieve data for the project. For Sales ETL we retrieved Sales Person, Product, Customers, Sales Territory and Date information as they were a part of our star schema and required that information for getting dimension key.



For Purchases ETL the group retrieved Product, Vendors and Date information.

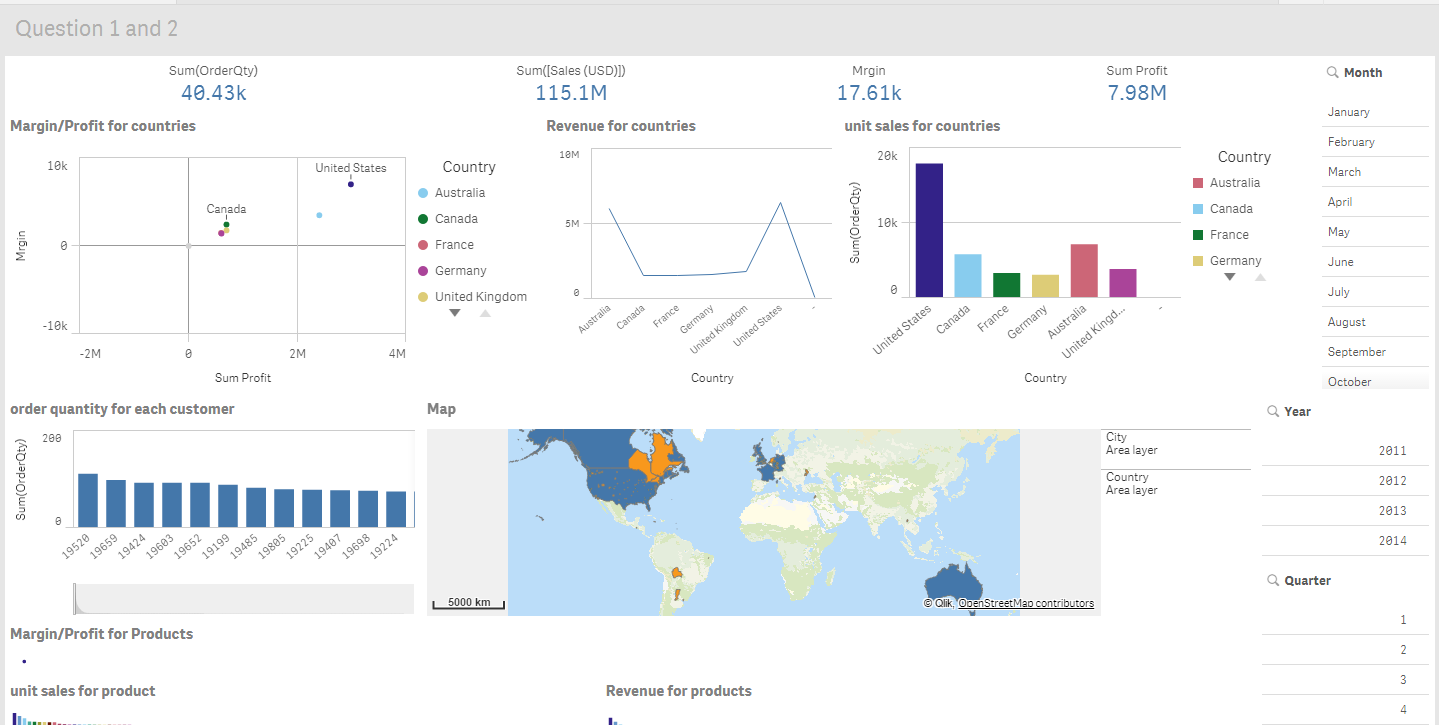
Once data was retrieved for the specific dimensions it was further joined with other input data. For instance, products input was joined with subcategory and category via add sequence tool which allowed us to get information on what products was under each category. And same procedure was applied to retrieve saleseperson name but instead of input data table we used input excel file as the sales person name was not specified in the database.

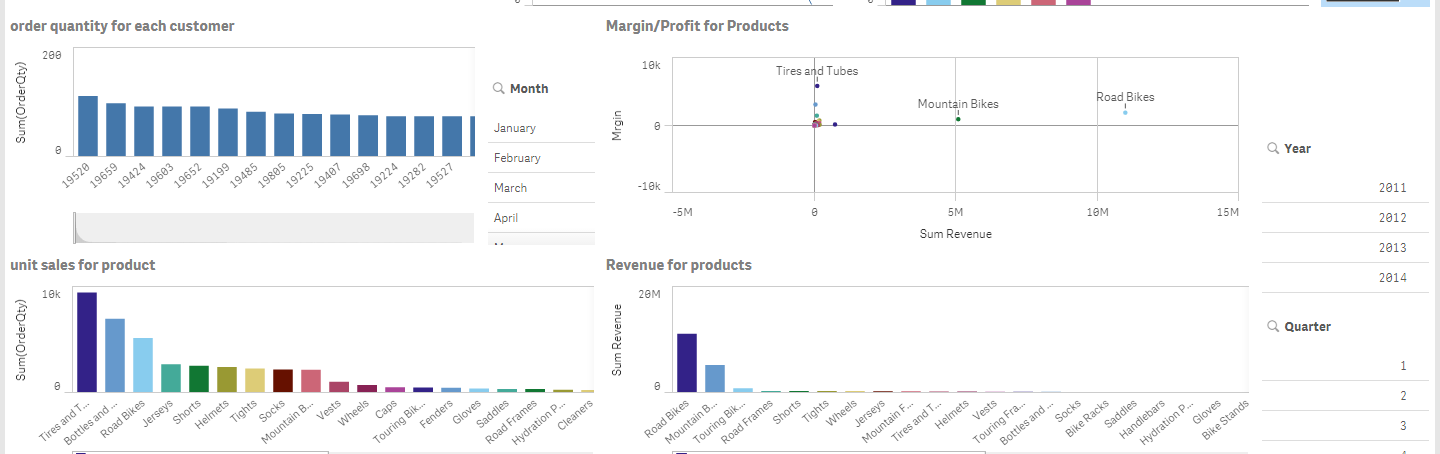
Once all the required information was joined via add sequence tool, dimension keys were added, and a link was created to export required information to a CSV file which will be used for data visualization.

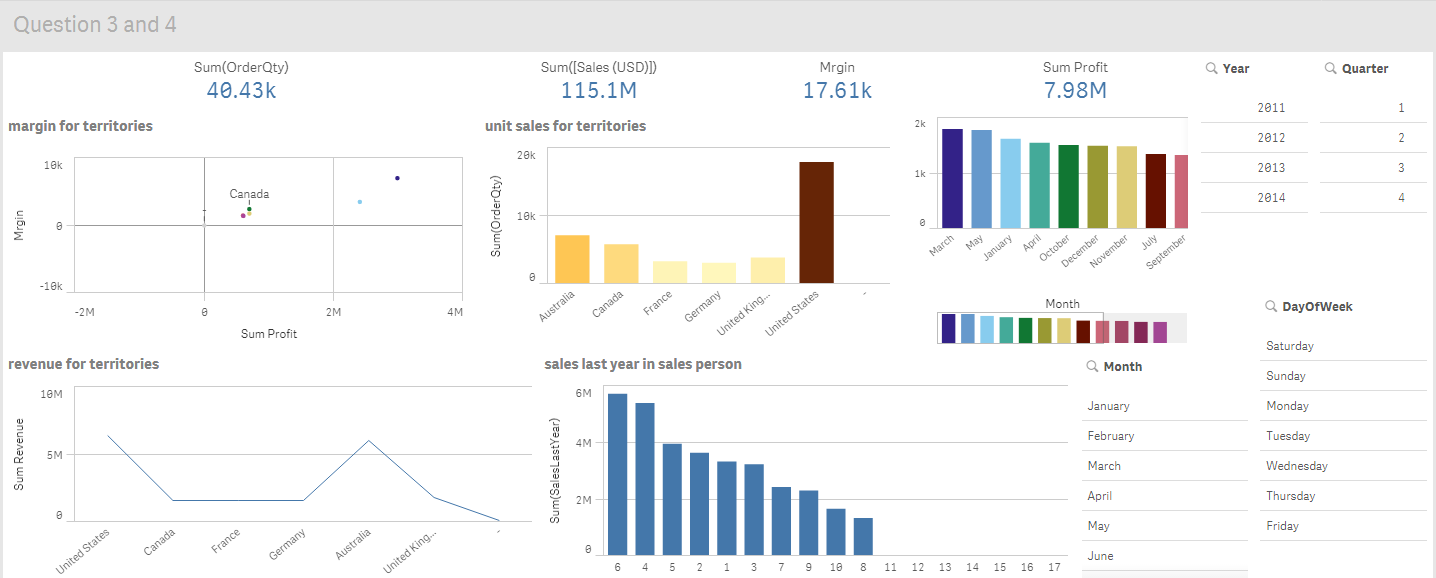
Each dimension was joined to each other to get a Sales/Purchase Fact output file but before exporting the file the information went through a filtering tool which allowed us to select the information that we required for the data visualization.

## **4.2 Data visualisation (Qlik)**

We combined question 1 and question 2 in the first dashboard, question 3 and question 4 in the second dashboard and the last question in a new app. The time filter placed on the right side of the dashboard. Overall, we used same graph layout and charts selection for every question.





For who’s the key customer question, we created three bar charts and one scatter plot to analysis and present data. Firstly, the scatter plot shows margin and profit for different locations. We can easily spot that the customer with the highest margin and profit. I calculated margin, the sum of profit and the sum of revenue by using Sum((ListPrice \* OrderQty - standardCost) - (ListPrice \* OrderQty), Sum((ListPrice \* OrderQty - standardCost) and Sum(ListPrice \* OrderQty) separately. The rest three bar chart present the unit sales, margin and the order quantity for each and customer. The customer's location showed by map which have both country and city information. We also displayed margin and the sum of profit on there to make it easier to read. Similarly, the fifth scatter plot directly shows the most profitable product. The following two tables present the unit sales and margin for product as well.   




(due to the laptop screen size, some of information can’t be displayed in one screenshot for one dashboard. So I screenshot them in question order)

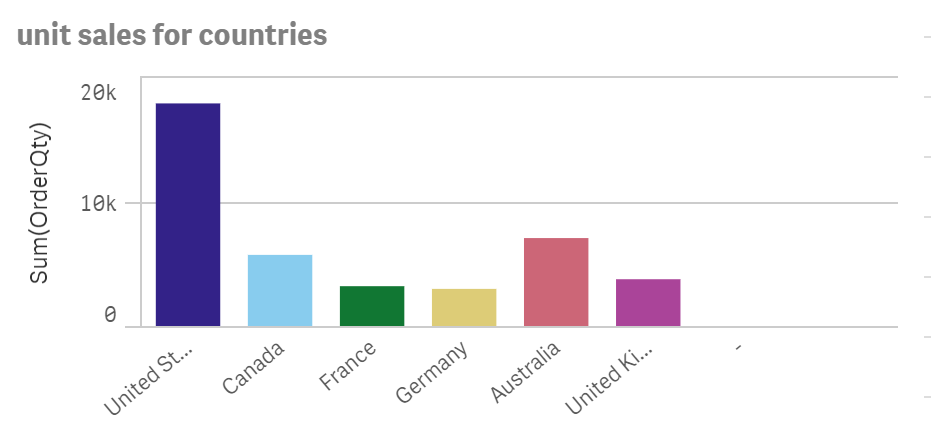
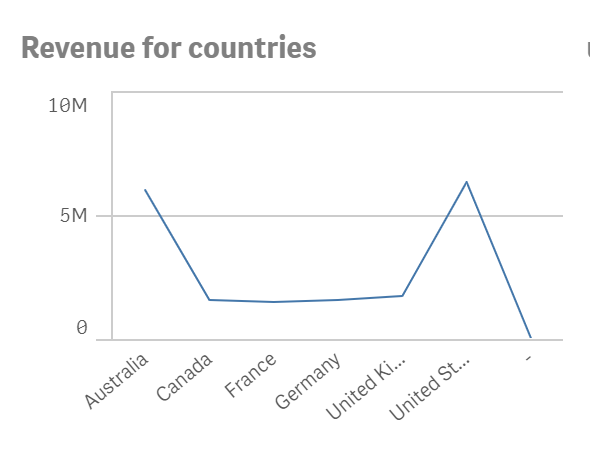
In second dashboard, we created unit sales, revenue tables for each territory and used margin and the sum of profit to show which territory is the most profitable one. We used bar chart instead of scatter plot for this, because it makes the result looks more straight forward. Besides, the salesperson’s sales in last year has been added as well. The bar chart can be present on monthly basis by click on month filter next to the bar chart. For the most profitable time period, we used unit sales as y-axis in sixth chart, revenue as y-axis in the last chart and years as x-axis in both of two charts. The same as question 4’s margin and profit charts, we used bar chart to present it. Since it only has a few variable in x-axis.



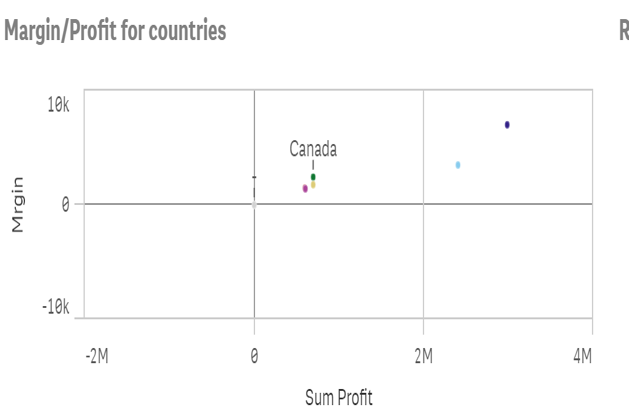
For the last question, we created a new app for purchase star scheme. We listed rejected quantity for each store and a scatter plot to show the relationship between order quantity and rejected quantity. In the end, we displayed the sum of order quantity and the sum of rejected quantity. I used Max(rejectedQty/receivedQty) to show the rejected rate.

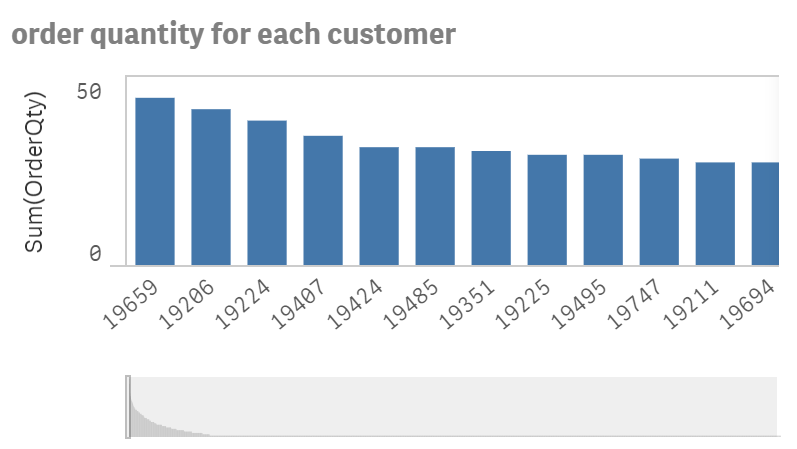
# **Generate Business Insights and Strategies**

## **5.1 Who are the key customers**

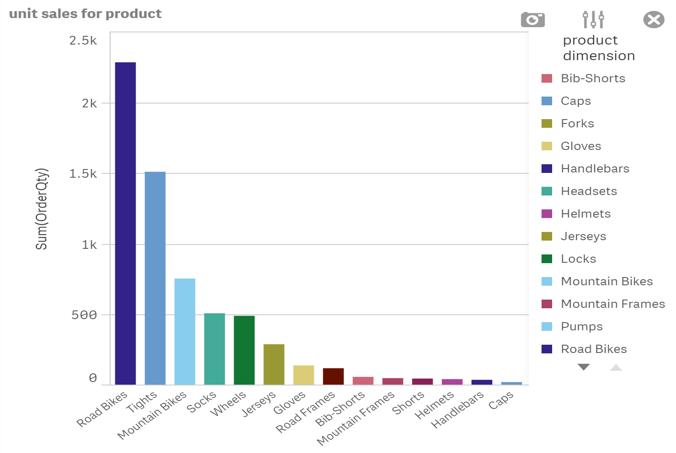


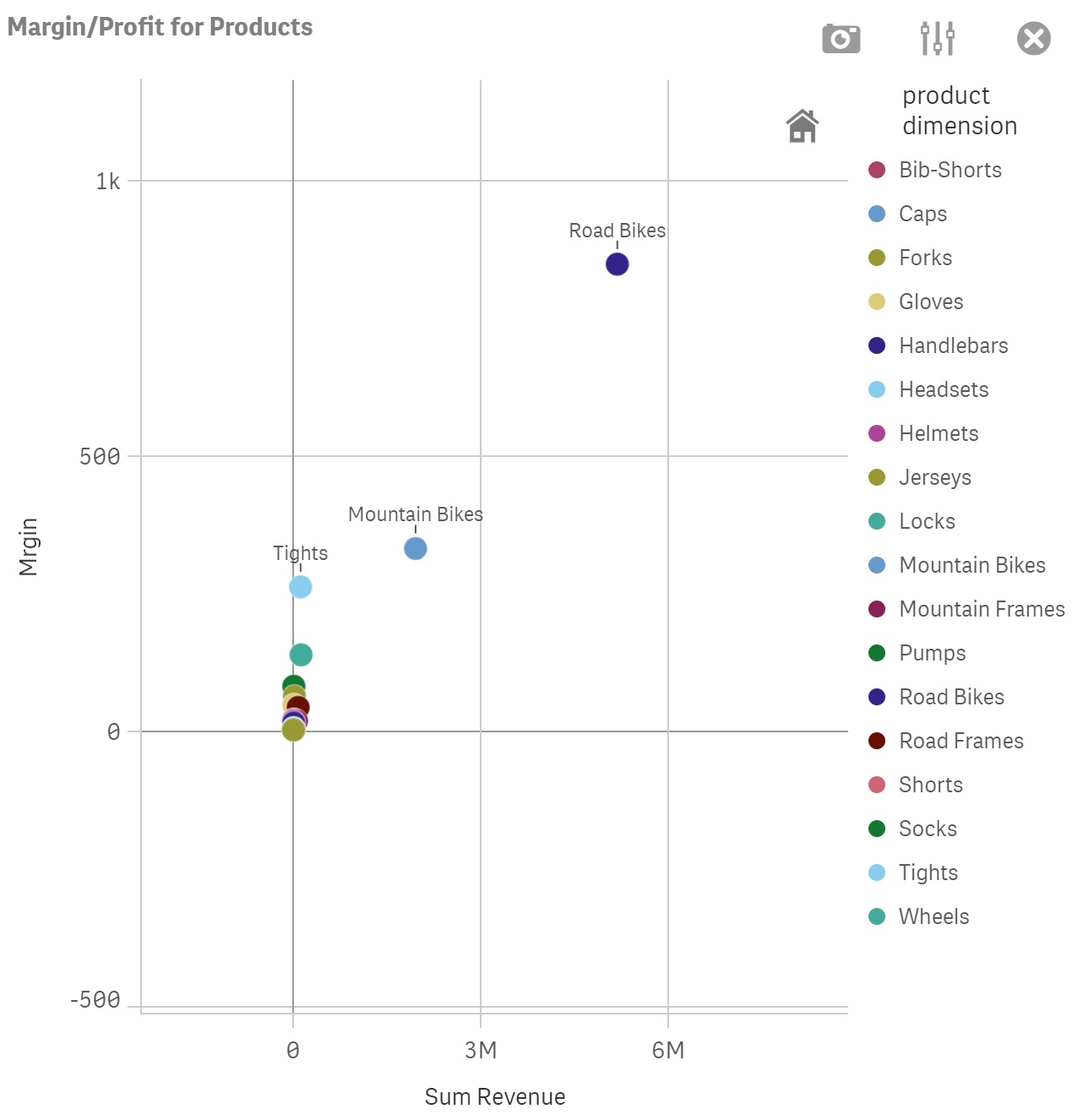
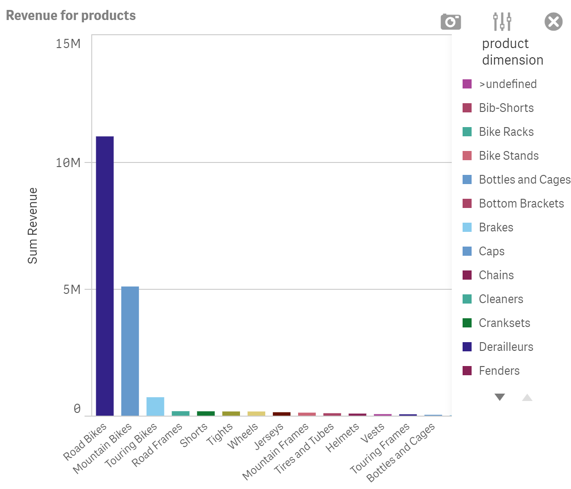
From the two tables above, Customers are mainly located in six countries, including Australia, Canada, France, Germany, the United Kingdom and the United States. Consumers from the United States contributed many sales, which is far more than the other five countries. The revenue for product table shows that from 2011-2014, there is a huge difference in unit sale, but the revenue gap between Australia and the United States is small. Even in 2011, both countries have same revenue and reached 1.54 million. However, consumers from the US are still the most profitable and ranked first in profit/margin table.

From 2011 to 2012, unit sale and revenue have shown an upward trend in 6 countries. After 2013, there has been a significant decline, especially in the United States. The reason for this situation may be that, according to the national surveys, in 2011-2012, there was a ‘cycling rejuvenation’ craze (Bollards by Reliance Foundry, 2017). At the same time, bicycles and vehicles have gradually evolved into a healthy way of exercising, as well as combined with the concept of recycle and environment. Thus, the sales has reached its peak in 2012. However, due to geographic, population distribution, and related infrastructure constraints, this boom subsided after 2012.

Through the above analysis, we can use American consumers as key customers because the high level of consumption has brought most of the profits to WV. In addition, we can divide customer group into loyal customers and ordinary consumers. Through the order quantity for each customer table, some loyal customers prefer to spend more money on Worldview bike. In conclusion, WV can focus its marketing strategy on the US market and will offer different strategies for cyclists and regular customers. For cyclists and loyal customers, WV should emphasize the quality of the products, the specialization of services and the price is not the main factor they consider. For ordinary customers, their focus the price or performance of the products. Worldview bike can offer different product line with different price, quality and service to satisfy customers various demands.

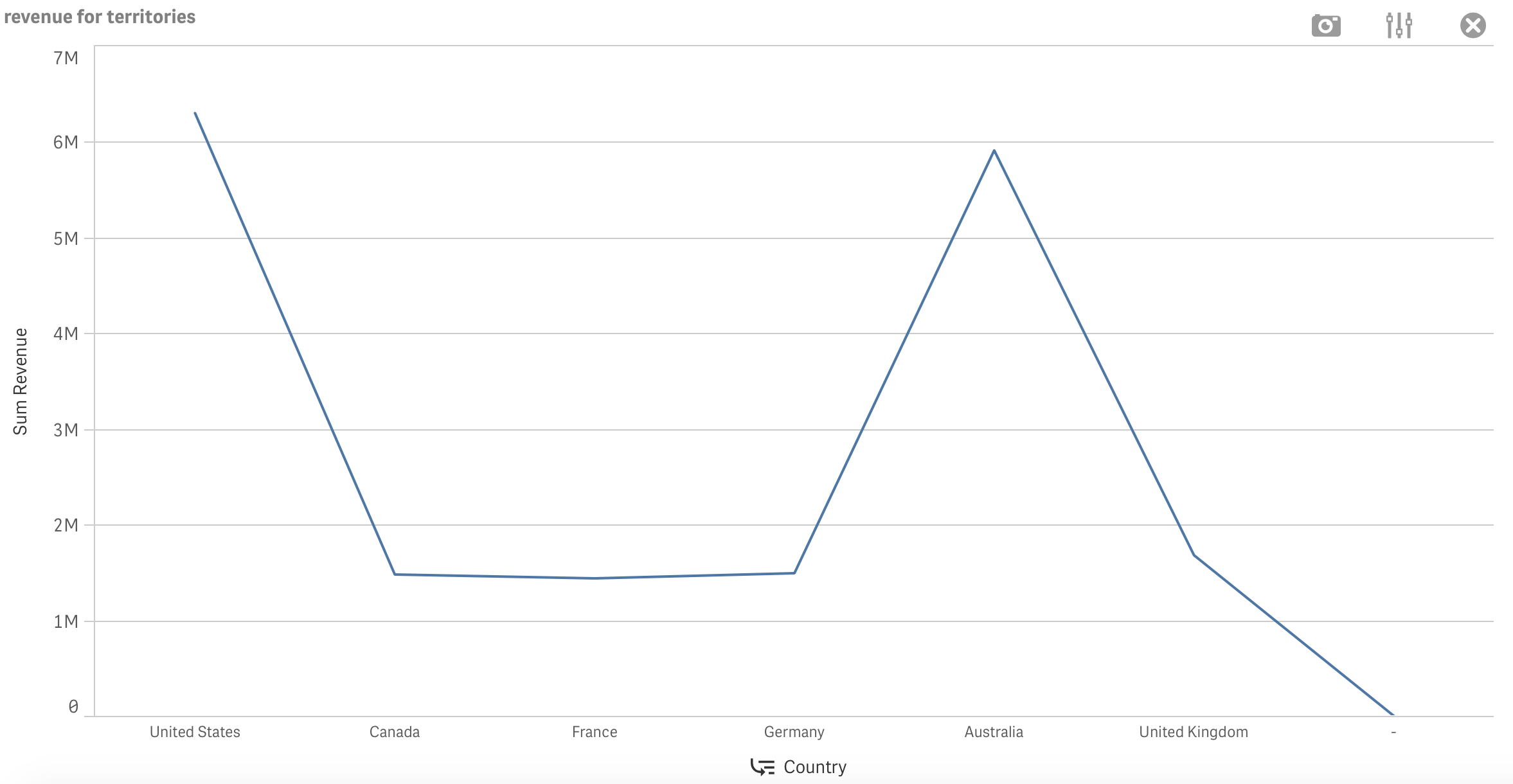
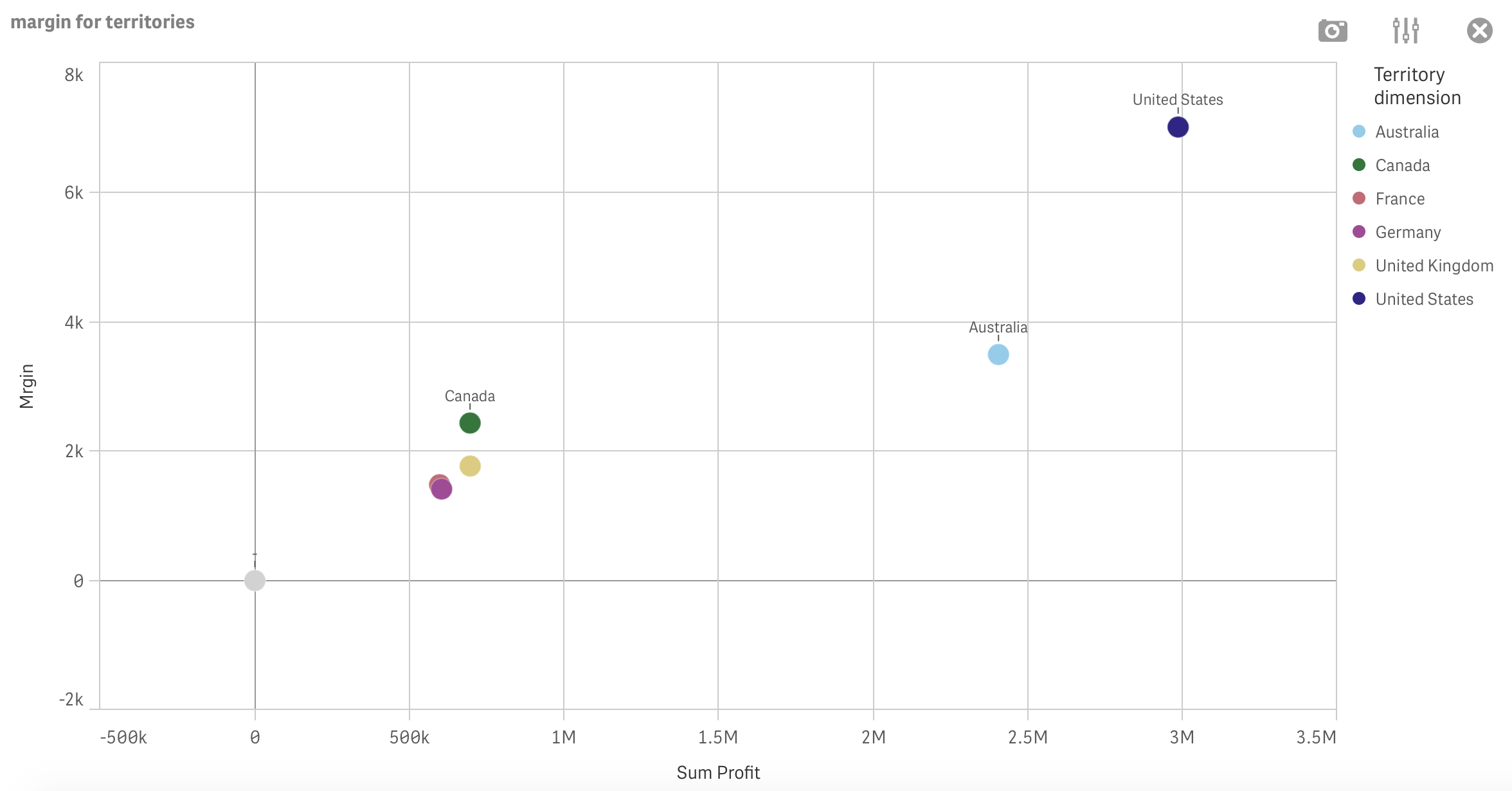
## **5.2 Which products are the most profitable**

By product unit sale table，Tires and tube become the bestselling product from 2011 to 2014 and the cumulative sales is 9.38k. Bottles and cages are behind tires and tube and Rode bike rank the third in product unit sale. In product margin table, tires and tubes have become the most profitable products in the past four years, followed by bottles and cages and then road bikes. However, Rode bike was the choice for majority customers and it also the most profitable product between 2011 and 2012. There is a down trend for rode bike from 2013 to 2104. Although Road Bike has brought a high turnover to the Worldview bike in the past four years, its profit margin is lower than that of Tires and tube and bottle cage due to the high cost of the product. In addition, because reduction of customer demand result in the unit sale decrease, the profit margin drops from 415.75k in 2011 to 133.25k in 2014.



Through the above analysis, it is found that the products with the best sales and the largest profit margins for Worldview are tires and tubes. Compared with bicycles as the main business products, some components and accessories bring actual profit for Worldveiw bike, such as tires and tubes or bottles and cages. Therefore, WV can adjust market strategy and considering consumer preferences and actual profits. The first strategy is that providing free service with fixing or installing for customer who purchase tires and tubes, which help to attract more customers. In addition, Worldview bike offer trade-in activities that allow customers to use broken tires and tubes in exchange for discounts to buy new tires and tubes. WV earns extra profits by recycling old tires and tubes for repair or reprocessing. Bottles and cages, as the second most profitable product, have brought considerable profits to WV. So managers can consider a profit upgrade for this. For example, customers can provide customized services for bottles and cages, such as customers can engrave their names on bottles or cages. Or carry out brand cooperation and launch a limited amount of special cooperation.

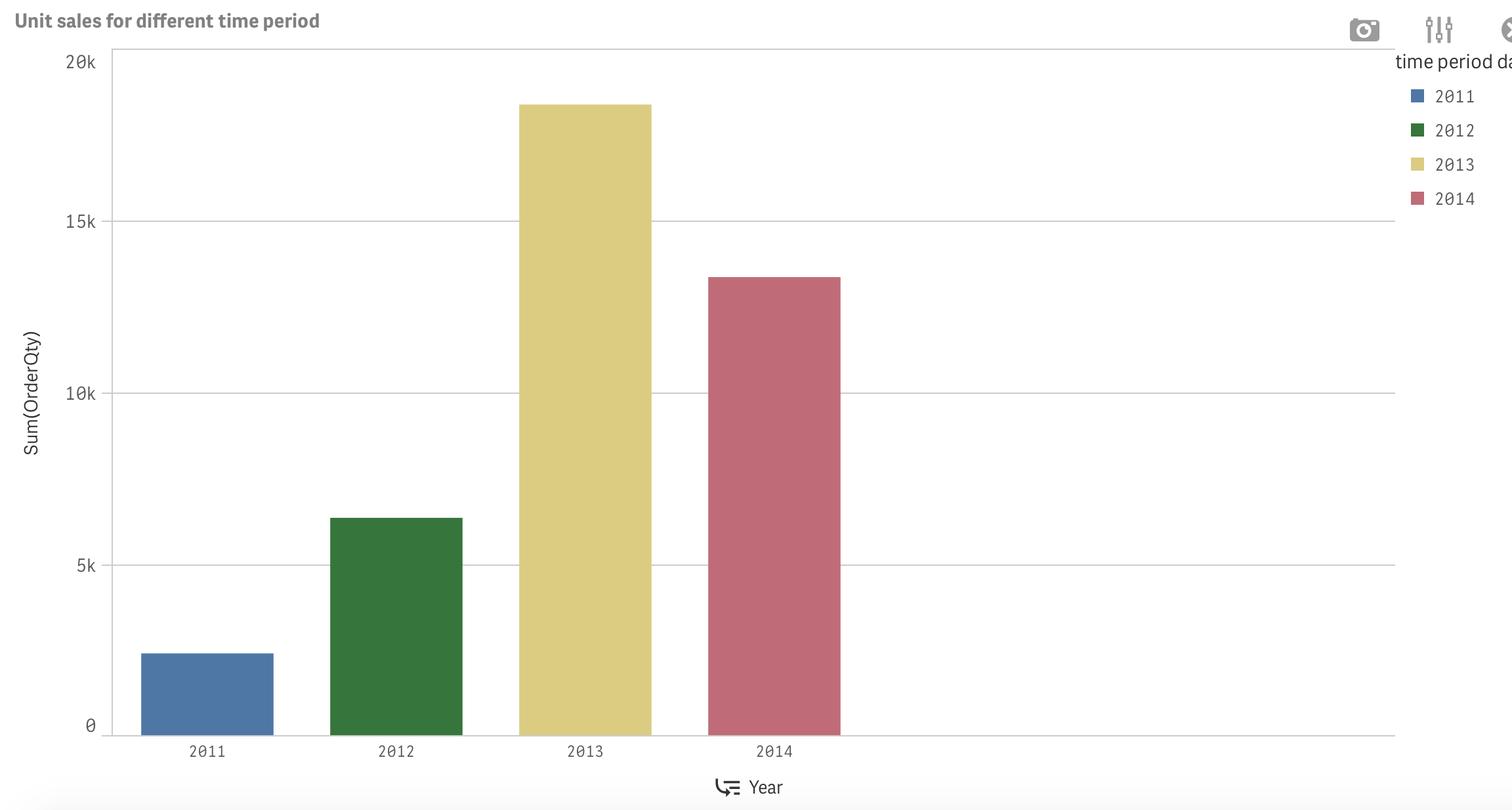
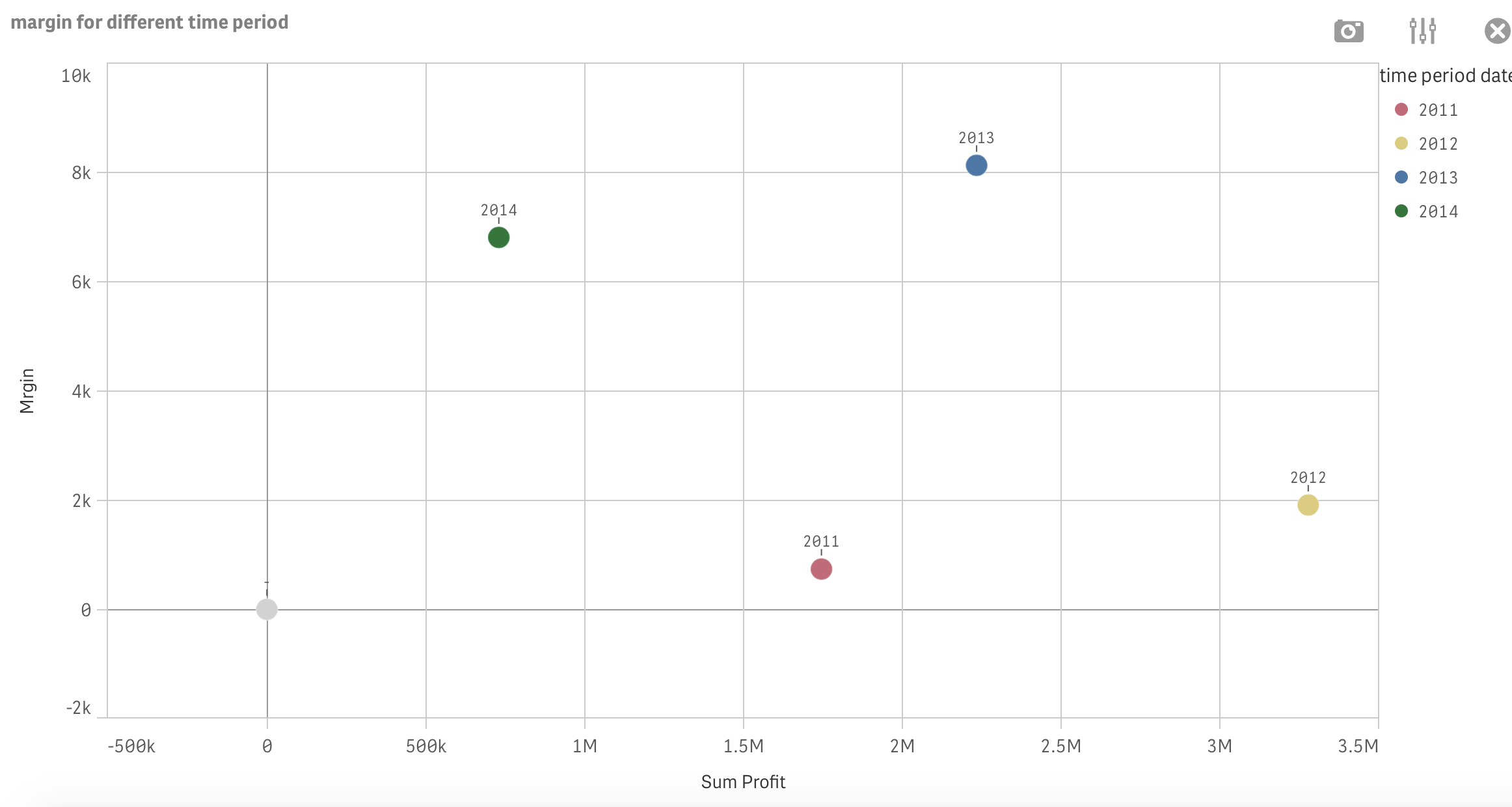
## **5.3 Which sales territories are the most profitable**

****

Base on graph above, the most profitable sales territories are belong to United State, with a total profit of 2.99M and a profitable margin of 7.01k. Moreover, United State make a total sale of 70.83M USD with only 6.28M revenue. According to the unit sales graph, we can see that United stated are way higher than the rest of the country, making a total of 17.81M order quantity and followed by Australia, a total of 7.04k. Graphs also reflected that on January 2014, United States make a margin of 472.6 which is higher than Australia’s margin of 260.6. With Sales Person 6 making a total of 5.37M and is higher than sales person 5 of 3.93M. Then a significant increase from January 2014 to March 2014 and maintained at about the same level till May. Then showed a significant decrease all the way to December 2014.

The proposed action plan is that WV should reduce their expenses in order to grow the company and for the rest of the country, maybe we can have a closer analysis on why they are not as profitable as United States. Moreover, With the analysis of which product are more profitable in each territory, we can increase the profitable product sold at each territory and reduced the amount of sales for least profitable product. In this way, company are able to make more sales base on the interest of the product.

## **5.4. Which time periods are the most profitable**



The most profitable time was on 2013 with a total profit of 2.23M and a margin of 8.14K. However, on 2012, it has a higher profit of 3.28M and a lower margin of 1.91K. While looking at the unit sales, 2013 has a higher sale of 18.37K than 2012’s sale of 6.33K and 2014’s sale of 13.34K. Moreover, 2013 makes a total of 45.79M USD sales which is also higher than 2012’s sale of 35.19M. Base on quarterly margin, we can see there is a decrease in value for first and second quarter, and there is no margin eared for the third and fourth quarter. While base on the yearly margin, it showed an increase from 2011 to 2012 then a significant decrease from 2012 to 2014. In 2013, it also makes a margin of 1.16K on all the Sunday which is higher than 2011, 2012 and 2014. Therefore, it tells us that 2013 was most profitable.

Based on the analysis above, the proposed action plan is to further analysis on 2013. For example, we can analysis to see why it is more profitable and what does it makes 2013 more profitable etc. Then we can compare both data to see what action needs to be taken and how can we improve the profitable for the next year.

## **5.5 Which vendors should be considered for replacement**

Here are a few standards when we are considering suppliers, such as quality, cost, on-time delivery, and service. Delivery on time is important because this affects the inventory turnover rate. Therefore, when selecting a supplier, WV should exclude supplier who have long delivery times, such as G&k bicycles, because the increase in time cost and transportation cost will affect the profit of Worldview bike. In addition, suppliers with high rejection rates mean problems in product quality, such as advanced bicycles. Product quality issues will improve WV and customer WV should establish long-term stable cooperative relationship with suppliers and establish an effective two-way incentive mechanism. For example, for outstanding suppliers， Worldview can offer order rewards or price incentives. (Due to data error on Pentaho, there is a lack of information for us to analysis this question)

# **Reference**

Adamson, C. (2013). Mastering data warehouse aggregates. Hoboken, N.J.: Wiley.

Bollards by Reliance Foundry. (2017). *Bike Culture: Europe vs America*. [online] Available at: https://www.reliance-foundry.com/blog/biking-usa-europe#gref [Accessed 22 Jun. 2017].

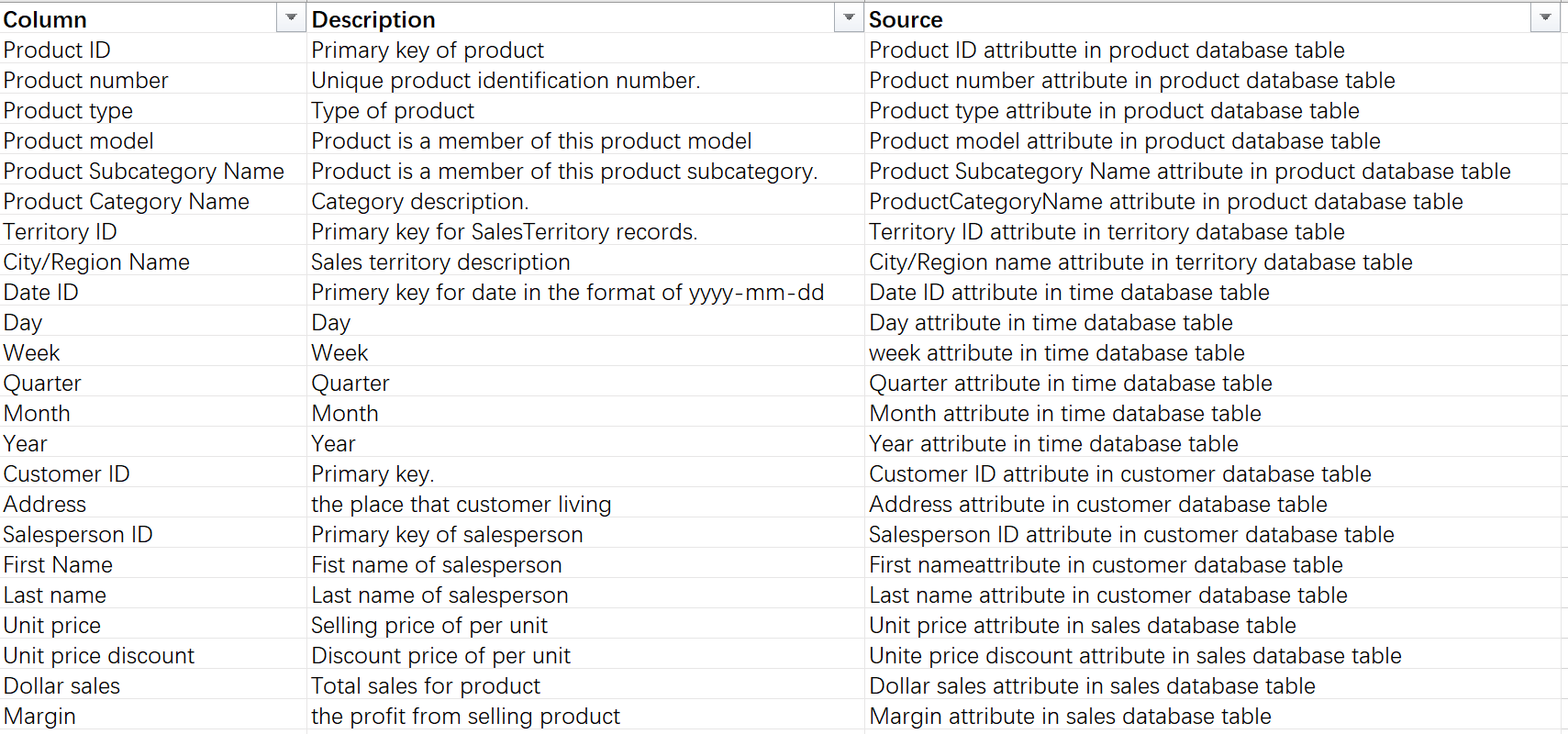
Chaudhuri, S. & Dayal, U., 1997. An overview of data warehousing and OLAP technology. *ACM SIGMOD Record*, 26(1), pp.65–74.

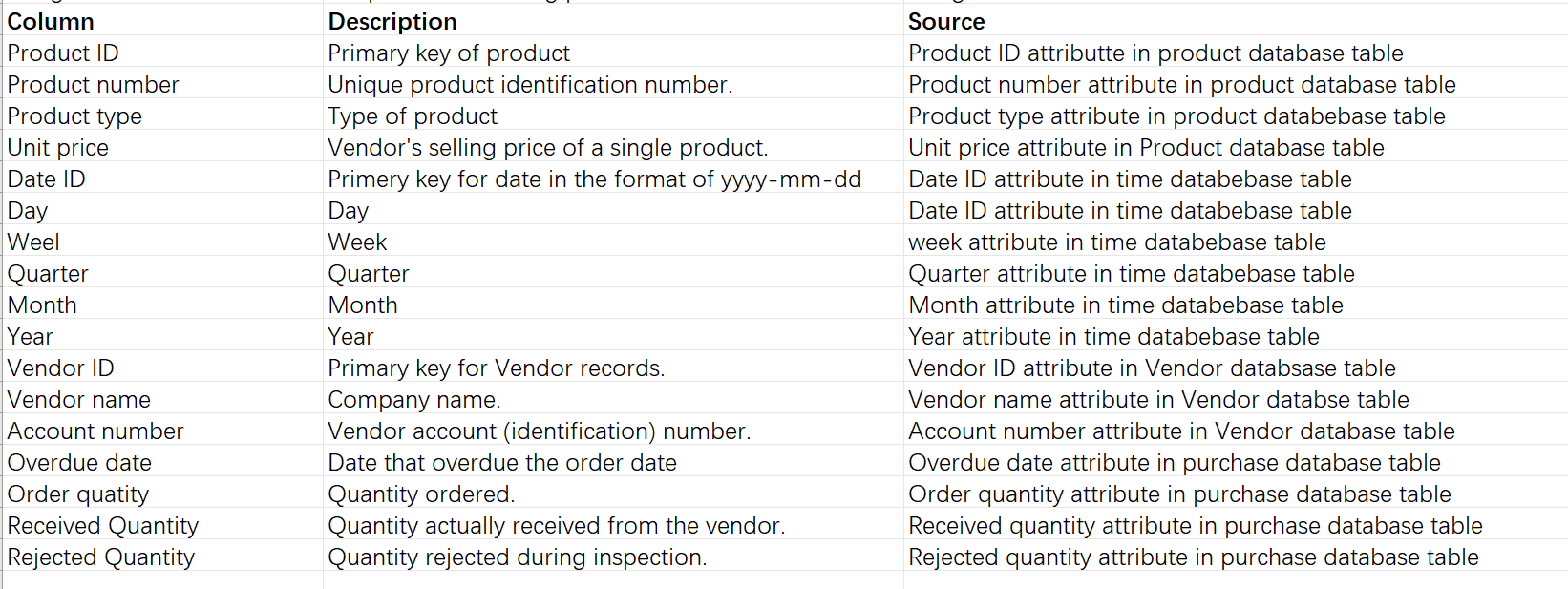
Galetto, M. (2016). *What is Business Analytics? See Benefits and Applications – NGDATA*. [online] NGDATA. Available at: https://www.ngdata.com/what-is-business-analytics/ [Accessed 16 Feb. 2016].

Martyn, T., 2004. Reconsidering Multi-Dimensional schemas. *ACM SIGMOD Record*, 33(1), pp.83–88.

# **Appendix**

## **7.1 Data Dictionary**

****

****

## **7.2 Work Breakdown**

|  |  |
| --- | --- |
| Ashar Artani | * Implementation of Pentaho |
| Junting Li | * Business Analytics Discussion * Generating business insights and strategising * Data Dictionary |
| Mosa Zuley del Carmen Nakada Araujo | * Dimensional Model Design |
| Chenxin Yang | * Implementation of Qlike |
| Yuru Zhou | * Writing of Summary * Generating business insights and strategising * Work Breakdown * Formatting |