TABLE OF CONTENTS

1.0 Introduction

1.1 Project Introduction

The internet, as it exists now, is a vast space to which enormous quantities of new information are always being uploaded. According to an infographic that was created by IBM, there are currently 2.7 zettabytes worth of data in existence in the digital realm. As a result of the high level of activity that takes place on social networks and the findings of this study, it is estimated that 35 zettabytes of data will be generated yearly by the year 2020. There is a correlation between the society in which we live and the significance of large data sets and analyses of large data sets. Today, we live in what is known as an Informational Society, and we are steadily progressing toward a Knowledge-Based Society. We require a greater quantity of data in order to extract superior information from our research. In a civilization known as the Society of Information, information plays a significant part in many aspects of societal life, including the political, cultural, and economic spheres (Ularu et al., 2017).

Business intelligence is a broader term that encompasses several techniques for gathering, storing, and analysing data from business processes. Most of this data is idle and essentially underused. With the aid of business intelligence, businesses can extract actionable data-driven insights from a huge amount of data. With that said, the ability to gain customer insights, improved visibility of business operations, improved efficiency across the organisation, improved marketing efforts, real-time data availability, and a competitive advantage for the business are some of the benefits that can be gained from using business intelligence in businesses. Therefore, more and more companies are hiring data analysts and data scientists to extract data-driven insights from the huge amount of data generated by the company everyday.

There are four general categories of business intelligence that are distinguished by the results they produce, namely descriptive analysis, diagnostic analysis, predictive analysis, and prescriptive analysis (Erl et al., 2016). These four types of analyses will be employed in this assignment in order to provide effective business solutions to the company targeting their business problems. The dataset that will be used in this assignment is the sales dataset of GBI, a company that produces high-end professional and deluxe touring bikes, as well as men's and women's off-road bikes. This assignment aims to identify the business problems that exist in

GBI and propose effective business solutions to solve the problems. SAP Lumira Discovery, a data visualisation tool will be employed in this assignment to facilitate better business analysis for this project.

1.2 Organisation Introduction

Global Bike Incorporation, also known as GBI, is a company that was established with the mission of providing innovative, high-performance bicycles to riders who place the highest demand on their equipment. GBI was founded by two co-CEOs, John Davis from America and Peter Schwarz from Germany. John Davis was a world-renowned cyclist and champion in mountain bike racing, and he was victorious in a great number of downhill and cross-country championships. John discovered early on that the available mass-produced bicycles were inadequate for the type of racing he was partaking in. John disassembled four of his old bicycles and reassembled them into a single "Frankenstein" bicycle, which he rode to victory in the national championship. In 1990, while John was recovering from an injury, he began creating the first batch of Frankenstein bicycles in his garage; each one was custom-built using parts from other bicycles. John successfully transitioned Frankenstein Bikes from his garage into a full-fledged production plant in Dallas and began building custom trail bikes, which he sold through a network of specialist bike dealers around the nation, as the number of orders increased.

On the other hand, Peter Schwarz was an engineering student in Heidelberg, Germany and he enjoyed participating in regional touring races during the weekends. Peter began designing and constructing road bikes based on an ultra-light composite frame he had developed for his engineering classes. Peter's creative use of carbon composite materials enabled him to construct a frame that was much stronger and ten times lighter than rival frames. As a student with limited financial resources, Peter formed a partnership with a local company that produced his frame designs as a contract manufacturer. Shortly, Peter founded Heidelberg Composites in order to market and design frames that would be produced by a contract manufacturer and utilised by racers around Europe.

Peter and John met in 2000 by chance and immediately identified their shared passion for performance and complementary company models. Each was seeking for a partner in a

different racing discipline and in a different market. They rapidly understood that merging their companies would be highly synergistic, and that merging their product lines and regional distribution routes would yield substantial efficiencies. Therefore, Heidelberg Composites and Frankenstein Bikes combined in 2001 to establish GBI. As co-CEOs, John and Peter are responsible for overseeing GBI's expanding organisation. From an organisational reporting standpoint, John is in charge of sales, marketing, service and support, information technology, finance, and human resources, while Peter is in charge of research, design, procurement, and manufacturing. The organisation structure of GBI is shown in Figure 1.

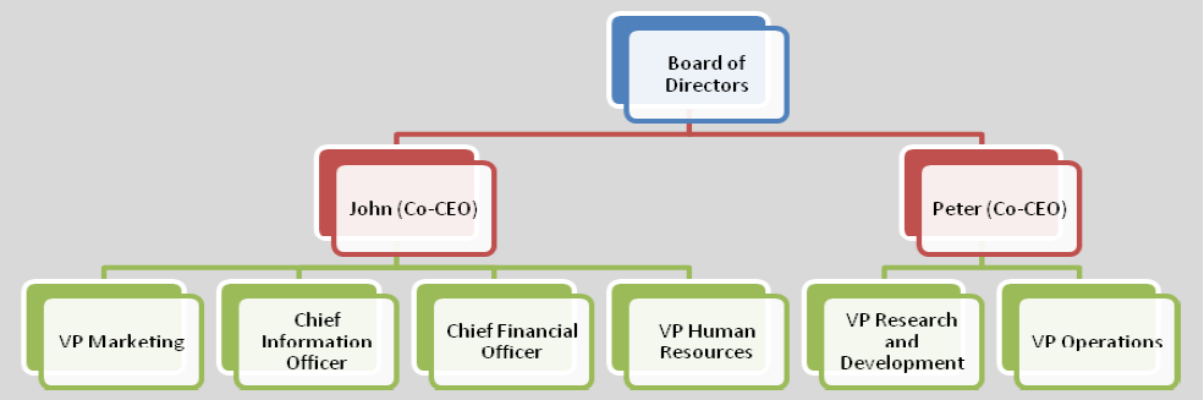


Figure 1. The organisation structure of GBI.

Due to a variety of tax and export concerns, GBI's headquarters are located in Dallas, and the corporation is registered as a US entity adhering to US GAAP accounting rules. In addition, GBI operates GBI Europe, a subsidiary headquartered in Heidelberg and conforms to IFRS accounting standards and German tax legislation.

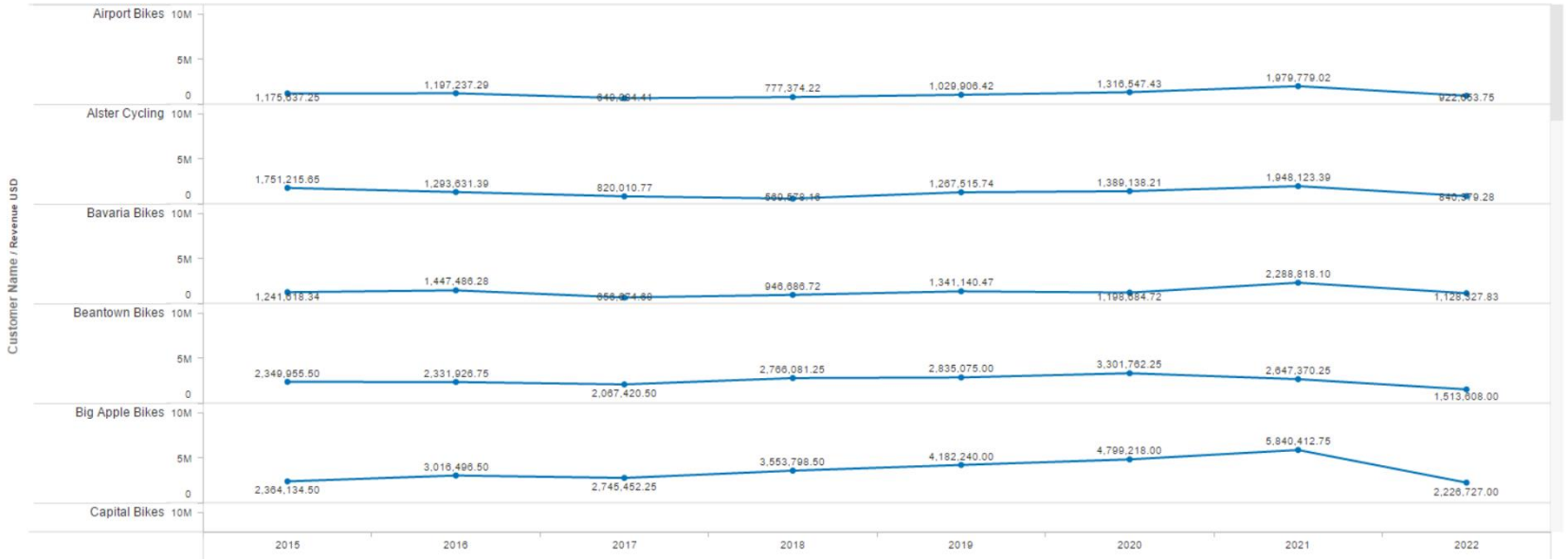
GBI mainly produces touring bikes and off-road bikes. GBI's distinctive road bikes are available in deluxe and professional editions. The primary distinction between the two models is the material used for the wheels. The wheels on the entry-level model are made of aluminium, while the wheels on the professional model are made of a carbon composite material. GBI also manufactures two models of off-road bicycles, one for men and one for women. The primary distinction between the two versions is the diminutive size and ergonomic design of the women's frame. Additionally, GBI also offers a line of accessories that includes helmets, T-shirts, and other riding equipment. GBI only partners with suppliers of accessories of the highest quality to increase the performance and comfort of GBI bike users.

GBI utilises two different locations for its manufacturing operations: Dallas and Heidelberg. Each factory has three assembly lines and has the capacity to produce around one thousand bicycles annually. The annual capacity of production is approximately 6000 bicycles, although it can be boosted by 15 to 20 percent through the use of overtime and part-time labour. Due to the extremely specialised nature of GBI's bicycles and the individualised needs of riders, GBI distributes its bicycles exclusively through reputable Independent Bicycle Dealers (IBDs). These dealers employ tour racing and off-road professionals to assist customers in selecting the GBI bike and accessories most suited to meet their specific requirements.

In 2009, GBI moved all of its information technology operations into the Dallas office and implemented a shared services model for those operations. In conjunction with this transition to centralised information technology, GBI also deployed SAP ERP. Before this, each division was responsible for maintaining its own unique set of application environments. The centralisation of all enterprise resource planning (ERP) functions was done with the primary objectives of lowering costs and providing the most cutting-edge technology to all worldwide divisions as possible. This centralised approach provides GBI with a sophisticated business platform that operates within a highly controlled environment. This helps GBI to maintain operational consistency and the integrity of its processes in all regions of the world (Magal et al., 2014).

1.3 Descriptive Data Analysis

Descriptive analysis is a sort of data analysis that assists in describing, displaying, or summarising data points in a constructive manner so that patterns may develop that satisfy all of the data's conditions. It is a crucial stage in undertaking statistical data analysis. It provides a conclusion about the distribution of your data, allows you to find typos and outliers, and identifies commonalities between variables, preparing you for future statistical analysis. The goal of conducting descriptive analytics is to provide answers to questions regarding events that have already taken place. This method of analytics places data inside its proper context in order to produce information (Erl et al., 2016). Data visualisation of the dataset of GBI from the sales department is carried out and investigated. After data visualisation, descriptive data analysis is carried out in this section. GBI is a bicycle company which produces touring bikes and off-road bikes, and its mission is to provide customers with the highest quality bikes with high performance, and the bikes that are most suitable for their needs.




Customer Name / Revenue USD
Capital Bikes
Cruiser Bikes
DC Bikes
Drahtesel
Fahrpott
Furniture City Bikes
Motown Bikes
Neckarad
Northwest Bikes
Ostseerad
Peach Tree Bikes
10M
5M
0
771,605.23
782,149.35
450,499.14
1,113,454.52
1,281,730.84
1,174,579.62
946,277.98
807,738.16
985,309.54
1,163,518.00
2,017,595.80
1,885,237.25
1,328,232.50
1,033,043.78
687,565.11
317,872.15
739,401.36
916,399.82
1,554,041.06
765,219.71
341,903.86
399,749.23
485,274.48
957,911.44
926,403.49
2,866,188.75
2,427,646.75
1,833,026.25
2,773,044.75
3,143,160.75
3,539,226.50
4,219,961.00
2,087,169.00
1,475,572.25
1,140,477.00
1,714,726.49
1,826,273.14
1,964,455.37
2,016,279.44
1,702,907.58
1,477,114.76
1,629,993.94
3,759,370.00
3,446,022.75
3,924,169.50
2,397,724.25
2,896,392.25
5,422,186.00
7,688,899.50
4,220,482.50
797,519.72
379,393.71
483,234.85
523,989.47
250,347.77
241,414.94
665,037.81
2015
2016
2017
2018
2019
2020
2021
2022

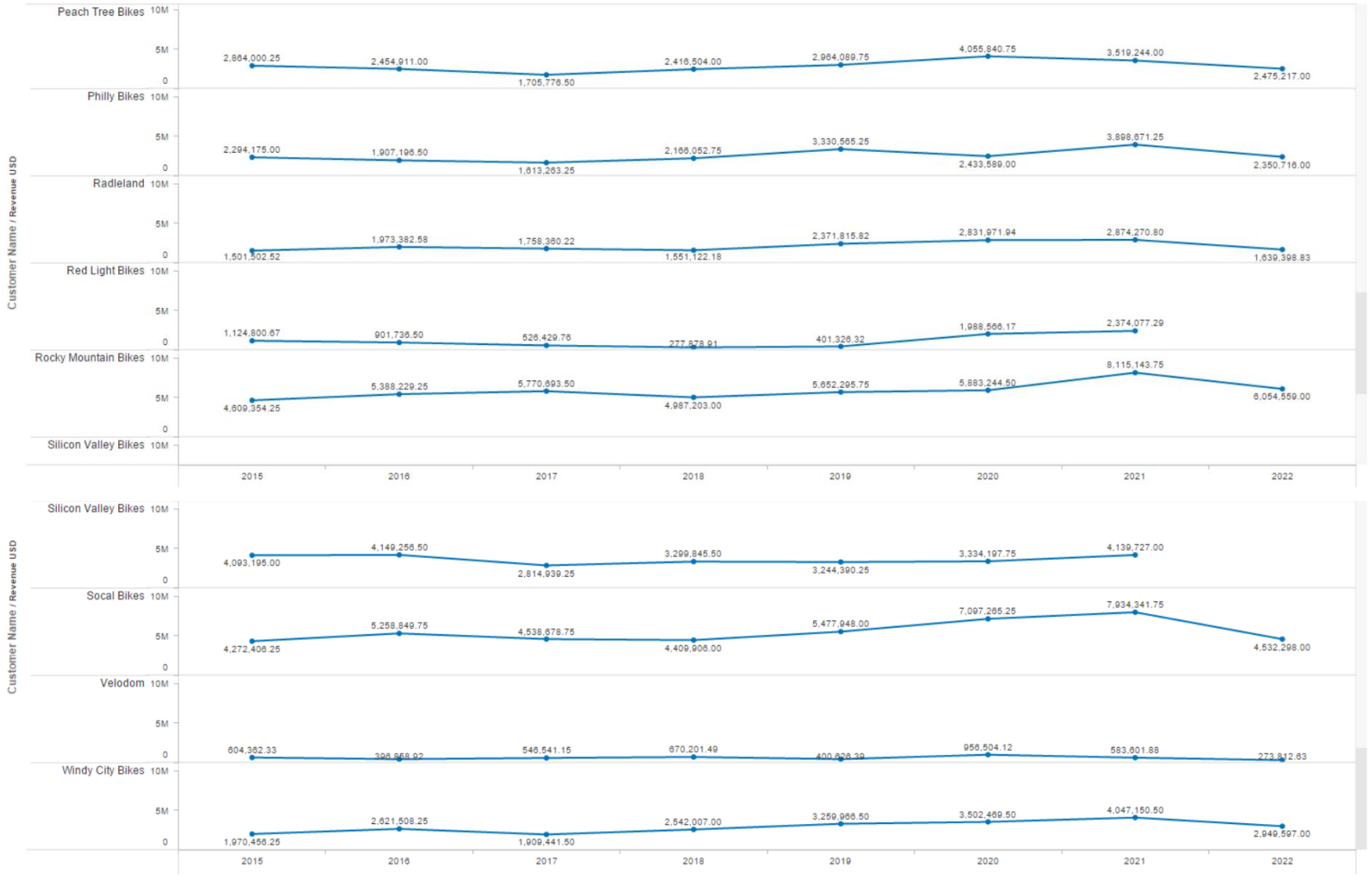


Figure 2 – 6. The line charts of Revenue vs Year by Customer Name.

Figures 2 to 6 show the line charts of revenue generated by GBI by selling its bicycles and accessories to the customers. Based on the figures above, it can be seen that almost all of the customers have a stable trend line in the revenue generated. However, it can also be seen that DC Bikes and Motown Bikes did not purchase any bicycles or accessories since 2017. On the other hand, Red Light Bikes and Silicon Valley Bikes also did not make any purchases from GBI in 2022. This phenomenon indicates that the customers might have left GBI and purchased from another competitor's company, or the customers have stopped their businesses. This will be further analysed in the data analysis section.

1.4 Problem Statement / Question

1. DC Bikes, Motown Bikes, Red Light Bikes and Silicon Valley Bikes left GBI.
2. Why did the customers leave GBI?
3. How to prevent the customers from leaving GBI?

1.5 Objectives

1. To investigate the reasons for the customers leaving GBI.
2. To propose a suitable solution to prevent more customers from leaving GBI in the future.

2.0 Organisational Memory and Integration

2.1 Organisational Memory

Organisational memory is the initial skill that builds the groundwork for the BI solution. It is the organisation's historical collection of information and knowledge. A data warehouse is one of the most common instruments for this capability. Before this repository's data can be used, it must be taken from its source, converted, and placed in the warehouse (Kuilboer et al., 2010). Online analytical processing (OLAP) and online transactional processing (OLTP) are both involved in the business of GBI as part of the ERP. OLAP is a system for analysing massive volumes of data in several dimensions at fast rates. These data typically originate from a data mart, data warehouse, or another centralised data repository. OLAP is great for data mining, business intelligence, complicated analytical calculations, and business reporting applications such as financial analysis, budgeting, and sales forecasting (IBM, 2021). On the other hand, OLTP permits the real-time execution of a large number of database transactions, generally via the Internet, by a large number of individuals. Many of our daily transactions, such as ATM withdrawals, in-store sales, and hotel reservations, rely on OLTP systems. OLTP can also be used to facilitate non-financial transactions, such as password changes and SMS messaging. A relational database is utilised by OLTP systems (IBM, 2021). Before 2009, GBI divisions operated various, separate application environments, hence increasing the organisation's expenses. In 2009, GBI implemented a shared services model for all IT operations as a result of the company's high expenses. The primary objective of implementing a centralised system is to reduce expenses and provide the most up-to-date technology to all divisions worldwide; concurrently, the data may be stored safely because the business platform operates in a tightly regulated environment (Magal et al., 2014). The data created by GBI are stored on SAP ERP server situated in Queensland, Australia, with SAP HANA serving as its big data platform. SAP HANA is a multi-model database that keeps data in memory as opposed to on a disc. This makes data processing orders of magnitude faster than disk-based data systems, enabling advanced, real-time analytics. The architecture of SAP HANA is shown in Figure 7 (SAP, 2021).

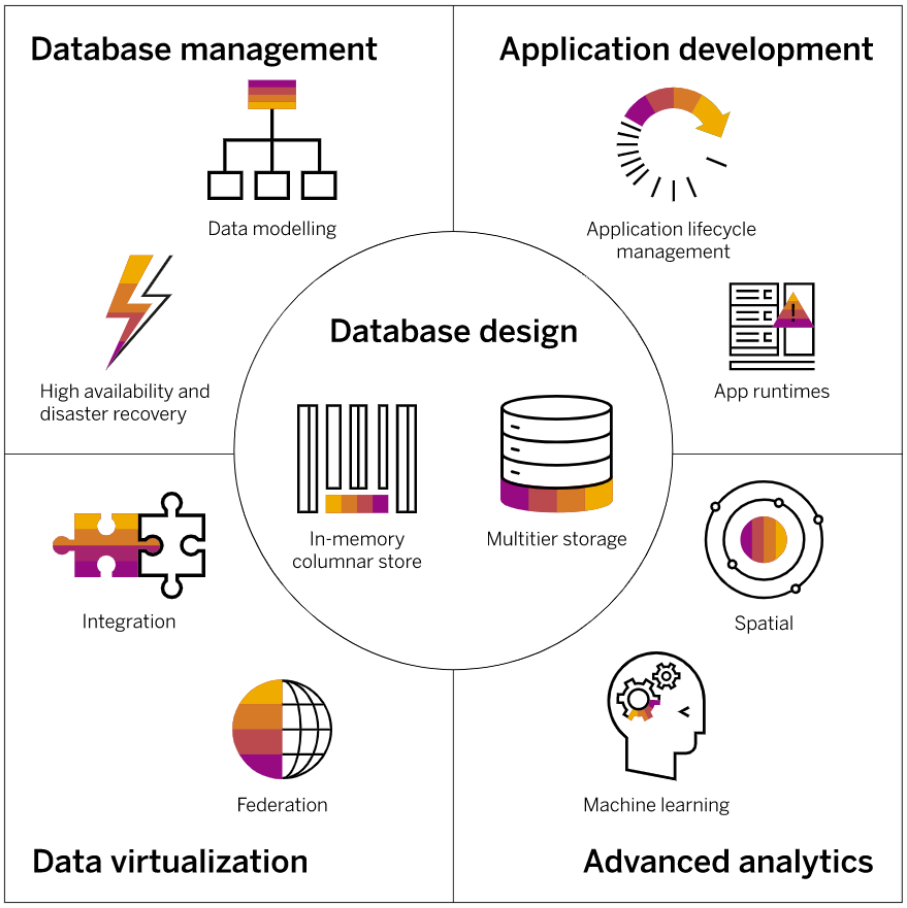


Figure 7. The architecture of SAP HANA.

2.2 Data Integration

Data integration is the process of merging data obtained from a variety of sources into a single, coherent whole. The ingestion process is the first step of the integration process, which also includes other steps like cleansing, ETL (Extract, Transform, Load) mapping, and transformation. Integration of data is ultimately what makes it possible for analytical tools to generate useful and relevant business knowledge (Lenzerini, 2002). SAP Business Warehouse, which is also utilised by GBI, is one of the solutions able to acquire data from the SAP ERP system. SAP Business Warehouse allows a robust analysis approach, involving the collection and analysis of an organisation's operational data from different sources or departments in order to enhance decision-making. SAP Business Warehouse is widely regarded as one of the best data warehouses available, and it is interoperable with SAP ERP systems and other SAP systems. In addition, SAP Business Warehouse is able to expedite the data management process by utilising the most recent in-memory and computational technologies. Using SAP Business Warehouse, anyone can quickly get historical data and analyse trends and patterns within the data, as well as extract data from other databases. SAP BusinessObjects Business Intelligence allows access to numerous SAP Business Warehouse sources in order to enhance analysis, reporting, and process planning. The retrieval of GBI's historical sales data from the enterprise resource planning (ERP) system is a vital step for enhancing GBI's decision-making since GBI can examine the data for corporate operations and goals to detect trends or patterns. This process is costly and superfluous to the GBI's operation. Certain businesses, however, may choose real-time observations of sales data because it may increase performance. GBI does not need a real-time data tracking system, but it must ensure that the data may be transferred without interruption for those who need it for analysis and the discovery of new insights or data. In addition, analytical methodologies can affect the conclusions made by decision-makers; consequently, the GBI must apply the right analysis approaches. Figure 8 displays the many data integration pathways that can be utilised to gather GBI's data.

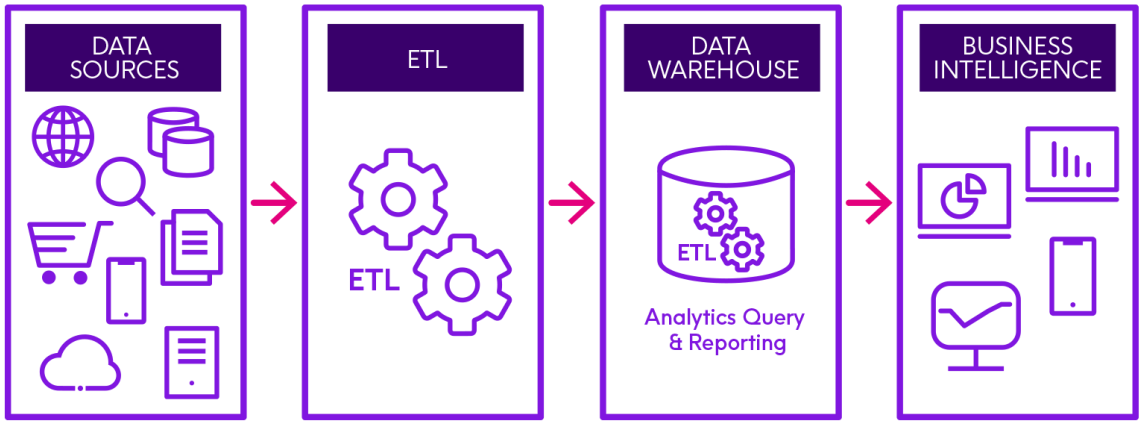


Figure 9. The process of data integration.

3.0 Analysis / Data Mining & Presentation / Visualisation

3.1 Data Mining Methodology

Data mining is the process of extracting insights and predictions from massive data collections. Using this method, data experts can perform an analysis of the data and then transform the results into relevant information. This allows businesses to make better and more accurate decisions. Data mining also helps to produce smart market decisions, conduct accurate campaigns, and make forecasts, among other things. Using data mining, businesses can evaluate customer behaviours and their perspectives that may contributes to tremendous success and data-driven business (Niakšu, 2015). CRISP-DM is one of the most powerful data mining methodologies and it will be utilised in this assignment to solve GBI's business issues. CRISP-DM stands for cross-industry standard process for data mining and it is practically very powerful, flexible and it is highly customisable when it is employed for analytics. The life cycle model is comprised of six phases, including business understanding, data understanding, data preparation, modelling, evaluation, and deployment, with arrows representing the most significant and common dependencies between phases. There is no rigid order between the phases (IBM, 2021). The life cycle of CRISP-DM is illustrated in Figure 9.

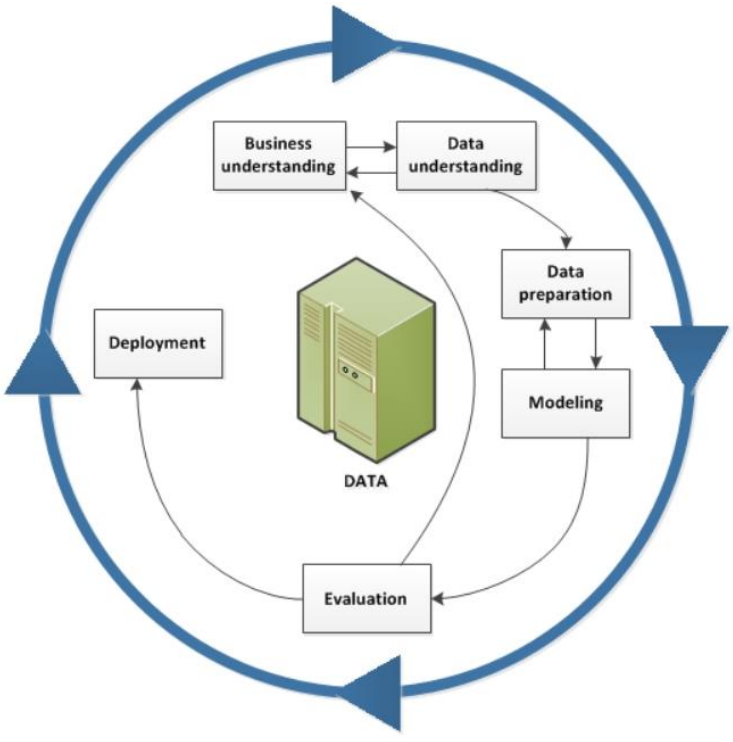


Figure 9. The life cycle of CRISP-DM.

3.1.1 Business Understanding

Before finding the solution for a business problem, business understanding is the utmost important step in a data mining methodology cycle. Understanding the business's vision, mission, and goals is essential to developing a solution that is effective for the business problem. Besides, it is also crucial to assess the situation and dive into the details of the issues associated with the business objectives. This includes finding the facts and developing a more comprehensive explanation of the matters described in the business objectives assignment. For instance, for the problem that we will be addressing in this assignment, which is four customers had left GBI, we must first find out the reasons that caused the customers to leave GBI and understand the purchasing behaviour of the customers before GBI.

3.1.2 Data Understanding

The data understanding phase complements the business understanding phase by encouraging the focus on identifying, assembling, and analysing the datasets that can assist data experts in meeting project objectives. This phase includes collecting the data, describing the data, exploring the data and verifying the quality of the data. For this assignment, firstly, the sales data from GBI is needed to be collected and description of data is needed to be done by identifying the dimension of the dataset. Next, basic statistics will be computed using the dataset and any missing data or abnormalities in the dataset will also be identified.

3.1.3 Data Preparation

The process that is tightly related to data understanding is data preparation. This stage, which is commonly referred to as "data wrangling" or "data munging", aims to provide the final dataset for modelling. It includes all processes required to generate the final dataset from initial raw data. There is a high probability that data preparation tasks will be performed many times and out of order. This includes the selection of tables, records, and attributes, as well as the transformation and cleansing of data for modelling tools.

3.1.4 Modelling

Modelling is a step where the desired model is selected and applied to the dataset, and the performance of the model is assessed and evaluated. In most cases, many models will compete against one another, and it will be necessary for the data scientist to interpret the outcomes of the models based on their knowledge of the domain, the pre-defined performance objectives, and the design of the test (Hotz, 2022). There are several modelling techniques that can be used such as regression models, classification models, and clustering models. Each models serves different purposes, hence the evaluation of the models in the next step of CRISP-DM is carried out.

3.1.5 Evaluation

Earlier evaluation steps focused on aspects such as the model's precision and applicability. During this step, the degree to which the model fits the business goals will be evaluated and the attempt to identify any business reasons why this model is inadequate will be carried out. Another option is to evaluate the trained model against real datasets inside a production environment, and the data mining outcomes are evaluated based on the underlying business goals. For this reason, test data sets are constructed using the methods described in the “Data Preparation” and “Modelling” phases apart from the labelling step (Huber et al., 2019).

3.1.6 Deployment

The final step of the CRISP-DM methodology is deployment. A customer’s inability to view the output of the model reduces the use of the model significantly. There is a wide range of possible degrees of difficulty for this phase. In order for the information obtained to be useful to the client, it ought to be organized and presented in a certain way. The deployment step, on the other hand, can be as straightforward as producing a report or as complicated as putting in place a data mining process that can be repeated across the company. This, of course, will depend on the requirements.

3.2 Data Preparation and Data Understanding

ETL, which is an acronym for extract, transform, and load, is a data integration process that integrates data from numerous data sources into a single, consistent data store that is then stored into a data warehouse or other desired system (IBM, 2020). The information that was gathered from GBI was entered into SAP ERP, which is an OLTP system, and it would be kept in SAP Business Warehouse (Wang et al., 2011). The GBI dataset that will be used for this assignment was acquired from SAP's centralised database system in the form of an excel spreadsheet, and OLAP analysis will be performed using the dataset. The architecture of a data warehouse is shown in Figure 10.

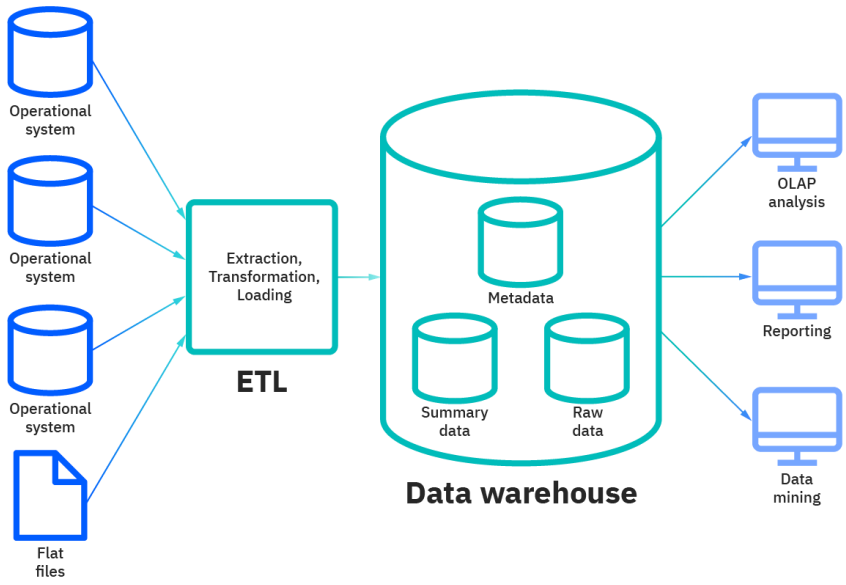
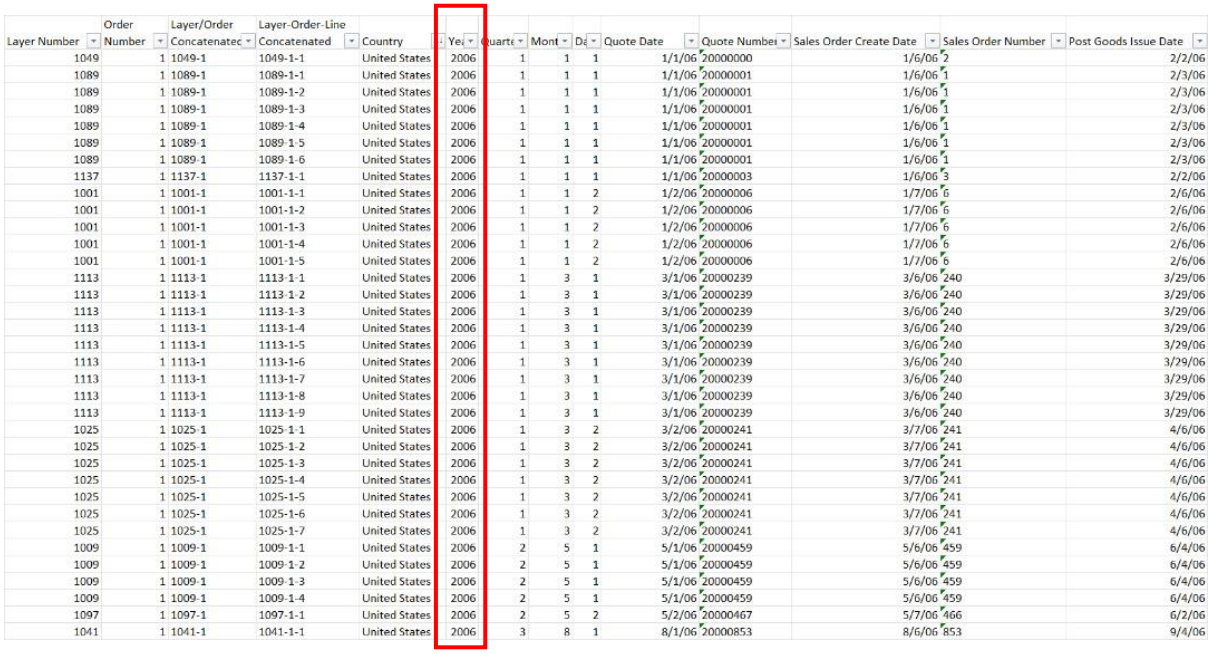
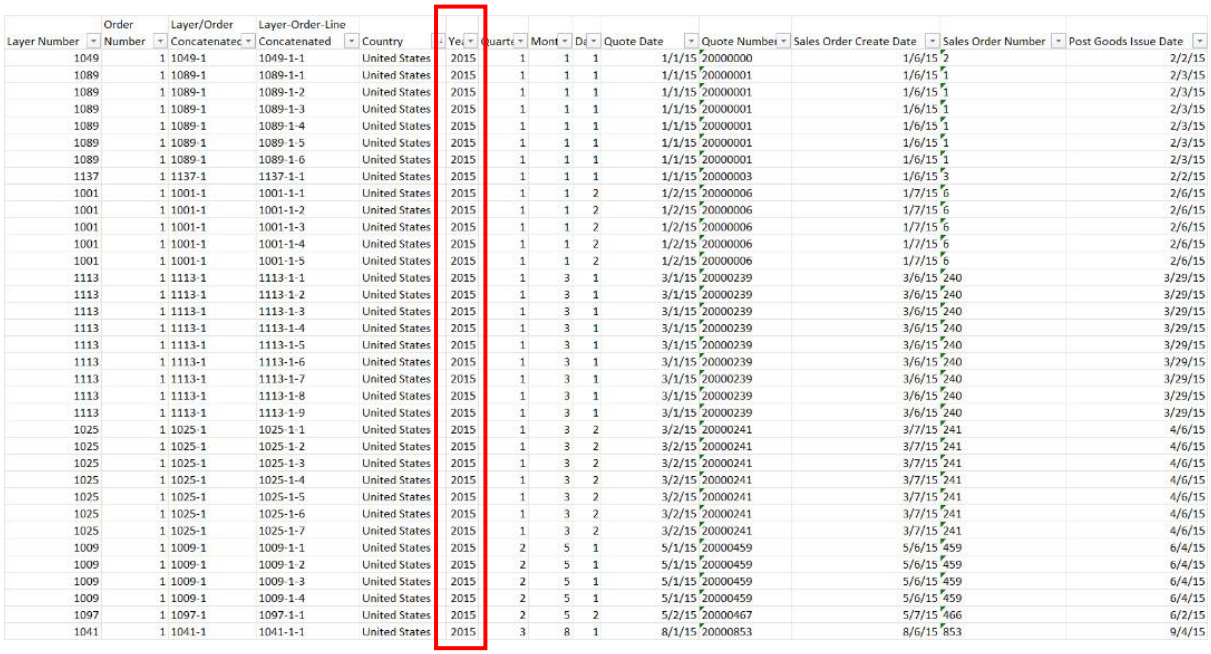


Figure 10. The architecture of a data warehouse.



Layer Number	Order Number	Layer/Order Concatenated	Layer-Order-Line Concatenated	Country	Ye	Quarte	Mont	Da	Quote Date	Quote Numbe	Sales Order Create Date	Sales Order Number	Post Goods Issue Date
1049	1	1049-1	1049-1-1	United States	2006	1	1	1	1/1/06	20000000	1/6/06	2	2/2/06
1089	1	1089-1	1089-1-1	United States	2006	1	1	1	1/1/06	20000001	1/6/06	1	2/3/06
1089	1	1089-1	1089-1-2	United States	2006	1	1	1	1/1/06	20000001	1/6/06	1	2/3/06
1089	1	1089-1	1089-1-3	United States	2006	1	1	1	1/1/06	20000001	1/6/06	1	2/3/06
1089	1	1089-1	1089-1-4	United States	2006	1	1	1	1/1/06	20000001	1/6/06	1	2/3/06
1089	1	1089-1	1089-1-5	United States	2006	1	1	1	1/1/06	20000001	1/6/06	1	2/3/06
1089	1	1089-1	1089-1-6	United States	2006	1	1	1	1/1/06	20000001	1/6/06	1	2/3/06
1137	1	1137-1	1137-1-1	United States	2006	1	1	1	1/1/06	20000003	1/6/06	3	2/2/06
1001	1	1001-1	1001-1-1	United States	2006	1	1	2	1/2/06	20000006	1/7/06	6	2/6/06
1001	1	1001-1	1001-1-2	United States	2006	1	1	2	1/2/06	20000006	1/7/06	6	2/6/06
1001	1	1001-1	1001-1-3	United States	2006	1	1	2	1/2/06	20000006	1/7/06	6	2/6/06
1001	1	1001-1	1001-1-4	United States	2006	1	1	2	1/2/06	20000006	1/7/06	6	2/6/06
1001	1	1001-1	1001-1-5	United States	2006	1	1	2	1/2/06	20000006	1/7/06	6	2/6/06
1113	1	1113-1	1113-1-1	United States	2006	1	3	1	3/1/06	20000239	3/6/06	240	3/29/06
1113	1	1113-1	1113-1-2	United States	2006	1	3	1	3/1/06	20000239	3/6/06	240	3/29/06
1113	1	1113-1	1113-1-3	United States	2006	1	3	1	3/1/06	20000239	3/6/06	240	3/29/06
1113	1	1113-1	1113-1-4	United States	2006	1	3	1	3/1/06	20000239	3/6/06	240	3/29/06
1113	1	1113-1	1113-1-5	United States	2006	1	3	1	3/1/06	20000239	3/6/06	240	3/29/06
1113	1	1113-1	1113-1-6	United States	2006	1	3	1	3/1/06	20000239	3/6/06	240	3/29/06
1113	1	1113-1	1113-1-7	United States	2006	1	3	1	3/1/06	20000239	3/6/06	240	3/29/06
1113	1	1113-1	1113-1-8	United States	2006	1	3	1	3/1/06	20000239	3/6/06	240	3/29/06
1113	1	1113-1	1113-1-9	United States	2006	1	3	1	3/1/06	20000239	3/6/06	240	3/29/06
1025	1	1025-1	1025-1-1	United States	2006	1	3	2	3/2/06	20000241	3/7/06	241	4/6/06
1025	1	1025-1	1025-1-2	United States	2006	1	3	2	3/2/06	20000241	3/7/06	241	4/6/06
1025	1	1025-1	1025-1-3	United States	2006	1	3	2	3/2/06	20000241	3/7/06	241	4/6/06
1025	1	1025-1	1025-1-4	United States	2006	1	3	2	3/2/06	20000241	3/7/06	241	4/6/06
1025	1	1025-1	1025-1-5	United States	2006	1	3	2	3/2/06	20000241	3/7/06	241	4/6/06
1025	1	1025-1	1025-1-6	United States	2006	1	3	2	3/2/06	20000241	3/7/06	241	4/6/06
1025	1	1025-1	1025-1-7	United States	2006	1	3	2	3/2/06	20000241	3/7/06	241	4/6/06
1009	1	1009-1	1009-1-1	United States	2006	2	5	1	5/1/06	20000459	5/6/06	459	6/4/06
1009	1	1009-1	1009-1-2	United States	2006	2	5	1	5/1/06	20000459	5/6/06	459	6/4/06
1009	1	1009-1	1009-1-3	United States	2006	2	5	1	5/1/06	20000459	5/6/06	459	6/4/06
1009	1	1009-1	1009-1-4	United States	2006	2	5	1	5/1/06	20000459	5/6/06	459	6/4/06
1097	1	1097-1	1097-1-1	United States	2006	2	5	2	5/2/06	20000467	5/7/06	466	6/2/06
1041	1	1041-1	1041-1-1	United States	2006	3	8	1	8/1/06	20000853	8/6/06	853	9/4/06

Figure 11. The "Year" variable in GBI dataset before transformation.



Layer Number	Order Number	Layer/Order Concatenated	Layer-Order-Line Concatenated	Country	Ye	Quarte	Mont	Da	Quote Date	Quote Numbe	Sales Order Create Date	Sales Order Number	Post Goods Issue Date
1049	1	1049-1	1049-1-1	United States	2015	1	1	1	1/1/15	20000000	1/6/15	2	2/2/15
1089	1	1089-1	1089-1-1	United States	2015	1	1	1	1/1/15	20000001	1/6/15	1	2/3/15
1089	1	1089-1	1089-1-2	United States	2015	1	1	1	1/1/15	20000001	1/6/15	1	2/3/15
1089	1	1089-1	1089-1-3	United States	2015	1	1	1	1/1/15	20000001	1/6/15	1	2/3/15
1089	1	1089-1	1089-1-4	United States	2015	1	1	1	1/1/15	20000001	1/6/15	1	2/3/15
1089	1	1089-1	1089-1-5	United States	2015	1	1	1	1/1/15	20000001	1/6/15	1	2/3/15
1089	1	1089-1	1089-1-6	United States	2015	1	1	1	1/1/15	20000001	1/6/15	1	2/3/15
1137	1	1137-1	1137-1-1	United States	2015	1	1	1	1/1/15	20000003	1/6/15	3	2/2/15
1001	1	1001-1	1001-1-1	United States	2015	1	1	2	1/2/15	20000006	1/7/15	6	2/6/15
1001	1	1001-1	1001-1-2	United States	2015	1	1	2	1/2/15	20000006	1/7/15	6	2/6/15
1001	1	1001-1	1001-1-3	United States	2015	1	1	2	1/2/15	20000006	1/7/15	6	2/6/15
1001	1	1001-1	1001-1-4	United States	2015	1	1	2	1/2/15	20000006	1/7/15	6	2/6/15
1001	1	1001-1	1001-1-5	United States	2015	1	1	2	1/2/15	20000006	1/7/15	6	2/6/15
1113	1	1113-1	1113-1-1	United States	2015	1	3	1	3/1/15	20000239	3/6/15	240	3/29/15
1113	1	1113-1	1113-1-2	United States	2015	1	3	1	3/1/15	20000239	3/6/15	240	3/29/15
1113	1	1113-1	1113-1-3	United States	2015	1	3	1	3/1/15	20000239	3/6/15	240	3/29/15
1113	1	1113-1	1113-1-4	United States	2015	1	3	1	3/1/15	20000239	3/6/15	240	3/29/15
1113	1	1113-1	1113-1-5	United States	2015	1	3	1	3/1/15	20000239	3/6/15	240	3/29/15
1113	1	1113-1	1113-1-6	United States	2015	1	3	1	3/1/15	20000239	3/6/15	240	3/29/15
1113	1	1113-1	1113-1-7	United States	2015	1	3	1	3/1/15	20000239	3/6/15	240	3/29/15
1113	1	1113-1	1113-1-8	United States	2015	1	3	1	3/1/15	20000239	3/6/15	240	3/29/15
1113	1	1113-1	1113-1-9	United States	2015	1	3	1	3/1/15	20000239	3/6/15	240	3/29/15
1025	1	1025-1	1025-1-1	United States	2015	1	3	2	3/2/15	20000241	3/7/15	241	4/6/15
1025	1	1025-1	1025-1-2	United States	2015	1	3	2	3/2/15	20000241	3/7/15	241	4/6/15
1025	1	1025-1	1025-1-3	United States	2015	1	3	2	3/2/15	20000241	3/7/15	241	4/6/15
1025	1	1025-1	1025-1-4	United States	2015	1	3	2	3/2/15	20000241	3/7/15	241	4/6/15
1025	1	1025-1	1025-1-5	United States	2015	1	3	2	3/2/15	20000241	3/7/15	241	4/6/15
1025	1	1025-1	1025-1-6	United States	2015	1	3	2	3/2/15	20000241	3/7/15	241	4/6/15
1025	1	1025-1	1025-1-7	United States	2015	1	3	2	3/2/15	20000241	3/7/15	241	4/6/15
1009	1	1009-1	1009-1-1	United States	2015	2	5	1	5/1/15	20000459	5/6/15	459	6/4/15
1009	1	1009-1	1009-1-2	United States	2015	2	5	1	5/1/15	20000459	5/6/15	459	6/4/15
1009	1	1009-1	1009-1-3	United States	2015	2	5	1	5/1/15	20000459	5/6/15	459	6/4/15
1009	1	1009-1	1009-1-4	United States	2015	2	5	1	5/1/15	20000459	5/6/15	459	6/4/15
1097	1	1097-1	1097-1-1	United States	2015	2	5	2	5/2/15	20000467	5/7/15	466	6/2/15
1041	1	1041-1	1041-1-1	United States	2015	3	8	1	8/1/15	20000853	8/6/15	853	9/4/15

Figure 12. The "Year" variable in GBI dataset after transformation.

Figure 11 and Figure 12 shows the "Year" variable in the dataset before and after transformation. After the data is extracted from SAP Business Warehouse, data cleansing, data preparation and data understanding steps have to be carried out. There is no missing value in the GBI sales dataset obtained, hence the data cleansing step is not necessary. Next is the data preparation step. Originally, the "Year" variable in the dataset ranges from 2006 – 2013. In the data preparation step, the "Year" variable is transformed to the range of 2015 – 2022 to ease

the data analysis phase. After data preparation, a brief data understanding is carried out. The dataset consists of 51 attributes and 47,940 tuples. The dataset also consists of the profit, profit margin, and revenue of the company in both USD and Euro currency, as well as the customers' name, customers' city, sales organisation and so on. It has already been mentioned that the dataset that will be used in this assignment is GBI's sales dataset, hence, most details and records that are related to the customers and sales transactions were recorded in this dataset. After the data cleansing, data preparation and data understanding phase, the dataset is loaded into SAP Lumira Discovery for data visualisation and data analysis.

3.2.1 ETL Process

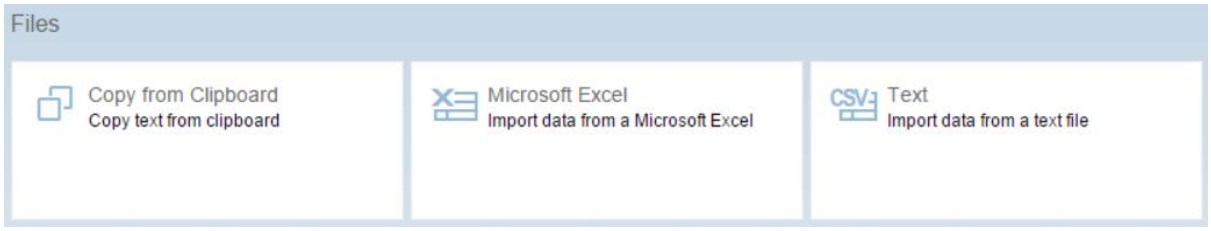
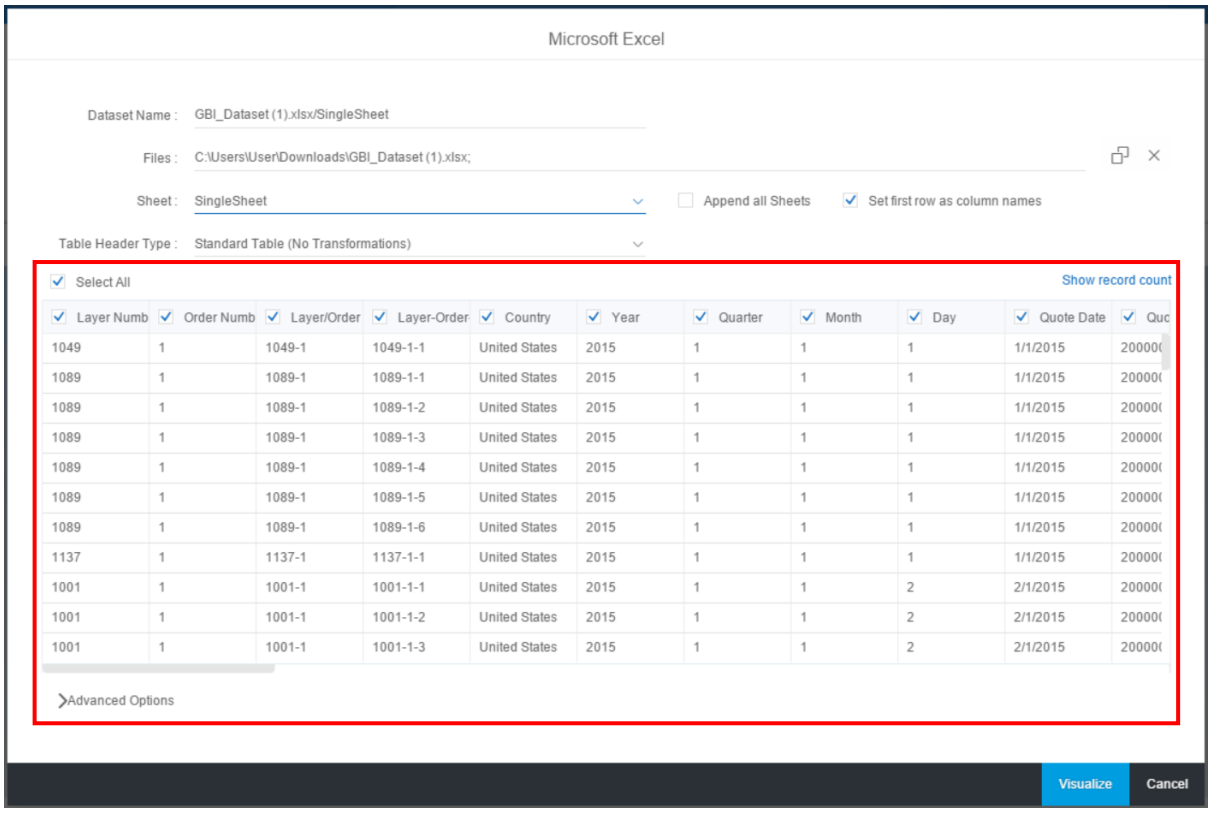


Figure 13. The "Microsoft Excel" in the Files option in SAP Lumira Discovery.



Layer Numb	Order Numb	Layer/Order	Layer-Order	Country	Year	Quarter	Month	Day	Quote Date	Quc
1049	1	1049-1	1049-1-1	United States	2015	1	1	1	1/1/2015	20000(
1089	1	1089-1	1089-1-1	United States	2015	1	1	1	1/1/2015	20000(
1089	1	1089-1	1089-1-2	United States	2015	1	1	1	1/1/2015	20000(
1089	1	1089-1	1089-1-3	United States	2015	1	1	1	1/1/2015	20000(
1089	1	1089-1	1089-1-4	United States	2015	1	1	1	1/1/2015	20000(
1089	1	1089-1	1089-1-5	United States	2015	1	1	1	1/1/2015	20000(
1089	1	1089-1	1089-1-6	United States	2015	1	1	1	1/1/2015	20000(
1137	1	1137-1	1137-1-1	United States	2015	1	1	1	1/1/2015	20000(
1001	1	1001-1	1001-1-1	United States	2015	1	1	2	2/1/2015	20000(
1001	1	1001-1	1001-1-2	United States	2015	1	1	2	2/1/2015	20000(
1001	1	1001-1	1001-1-3	United States	2015	1	1	2	2/1/2015	20000(

Figure 14. The process of loading the GBI dataset into SAP Lumira Discovery.

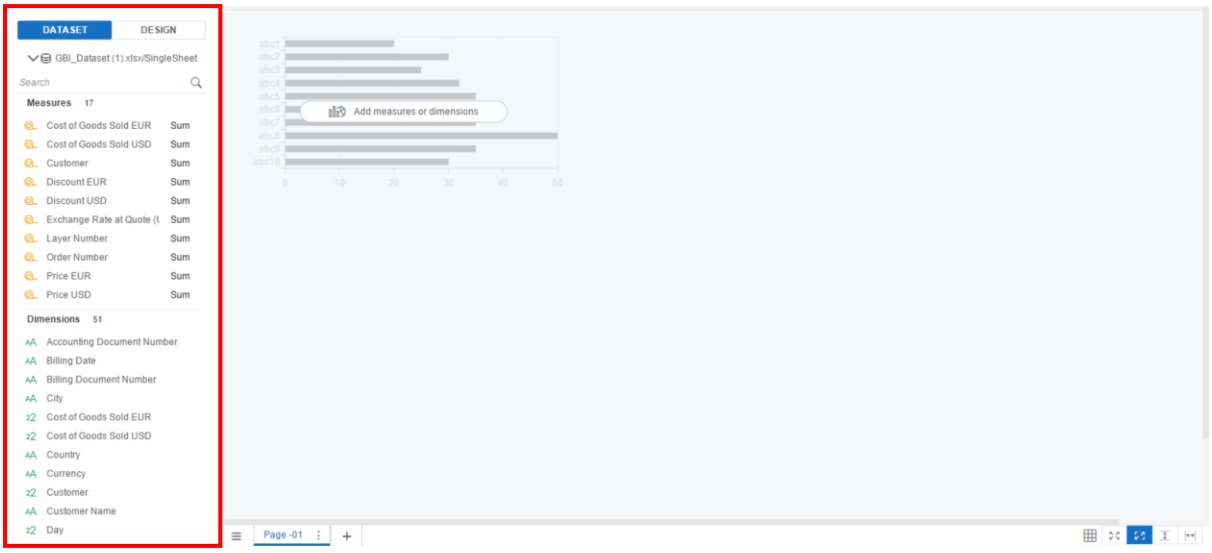


Figure 15. The attribute of the dataset is shown on the left panel of the software.

After launching SAP Lumina Discovery, the first step is to select “Microsoft Excel” on the homepage as shown in Figure 13 and choose the dataset from the correct destination. This is due to the GBI sales dataset being stored in an Excel spreadsheet. After the dataset is selected, there will be a panel showed by SAP Lumira Discovery and this panel shows the metadata of the dataset and the first 11 rows of the dataset (as shown in Figure 14). When this panel pops up, the “Visualize” button is clicked and the dataset is successfully loaded into the software. After that, the data visualisation dashboard is ready to be used. The measures and dimensions of the dataset is also shown on the left panel of the dashboard so that the users can drag the desired attribute into the charts to start visualising and exploring the dataset.

3.3 Data Analysis

3.3.1 Online Analytical Processing (OLAP) Cubes

OLAP is used to perform multidimensional analysis on business data. It also gives users the capacity to execute complex computations, analyse trends, and create sophisticated data models. It serves as the basis for a wide variety of applications for use in business (OLAP, 2022). The OLAP cube is the core of the majority of OLAP databases, allowing you to efficiently query, report, and analyse multidimensional data. The OLAP cube expands the row-by-column schema of a conventional relational database and adds levels for various data dimensions (IBM, 2021).

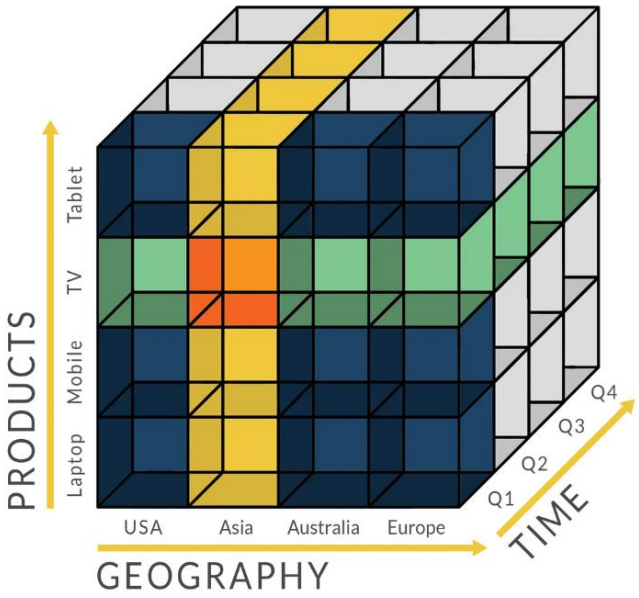


Figure 16. An example of an OLAP cube (OLAP, 2022).

Due to the OLAP cube being constructed using a dimensional model, it has the advantage of allowing analysts to obtain the data-related information they need in a shorter amount of time. Additionally, the dimensional structure of the cube enables analysts to intelligently access a variety of possible data perspectives (S. Jensen & Snodgrass, 2017). Now that we have understood what an OLAP cube is, as mentioned previously, Figure 15 above illustrates the different measures and dimensions in the GBI sales dataset was loaded automatically into SAP Lumira Discovery. Therefore, it is understood that the combination of the measures and

dimensions makes up an OLAP cube in which, due to its properties, it can considerably enhance the analysis process in this assignment.

3.3.2 Presentations / Visualisations

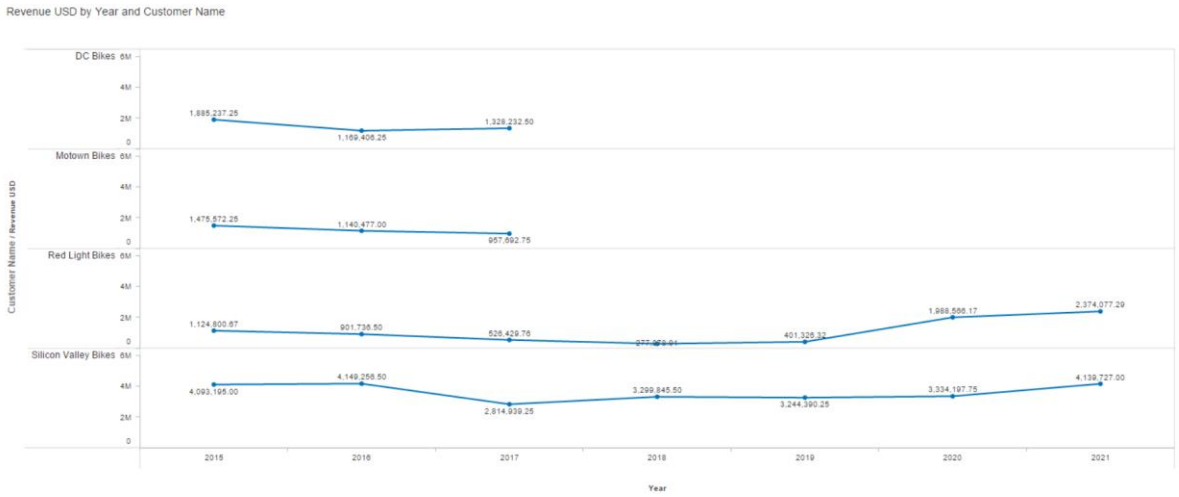


Figure 17. Revenue USD by Year and Customer Name (DC Bikes, Motown Bikes, Red Light Bikes and Silicon Valley Bikes.

As mentioned in Section 1.3, Descriptive Data Analysis, it was seen through the visualisation that DC Bikes, Motown Bikes, Red Light Bikes and Silicon Valley Bikes have stopped doing business with GBI. As shown in Figure 17, DC Bikes and Motown Bikes left in 2018 and Red Light Bikes and Silicon Valley Bikes left at 2022. This will be our main business problem to address and tackle in this assignment. Thus, in order to provide efficient business solution to GBI, further deeper analysis is needed to be carried out using the dataset to gain more historical business insights of GBI.

However, from the figure above, an obvious increase in revenue in the year 2021 is seen in Red Light Bikes and Silicon Valley Bikes. The revenue of Red Light Bikes increased from 1,988,566.17 in 2020 to 2,374,077.29 in 2021; whereas for Silicon Valley Bikes, the revenue increased from 3,334,197.75 to 4,139,727.00. Hence, it is important to find out the reasons why did these customers leave GBI.

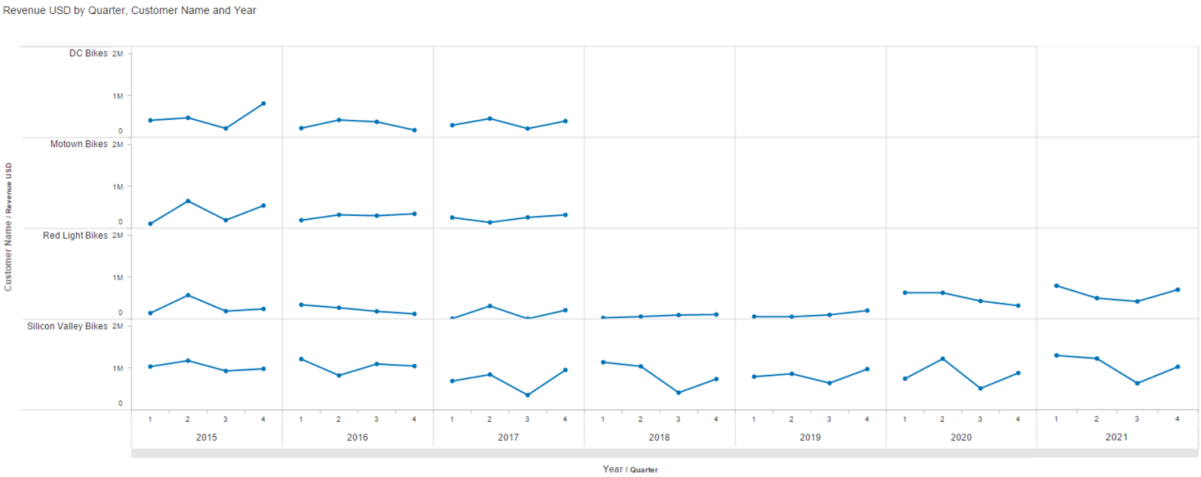


Figure 18. Revenue USD by Quarter, Customer Name and Year.

Since only the sales records of the first half of the year are included in 2022, it is important to identify the purchase behaviour of the customers throughout the previous years. Based on Figure 18, the trend of purchasing of Motown Bikes is decreasing in 2017 compared to 2015 and 2016. However, for Red Light Bikes and Silicon Valley Bikes, the customers purchased the most bikes and accessories from GBI in the first quarter of 2021. Besides, the trend of purchase of these two customers also increased in the fourth quarter of 2021. Hence, it remains a mystery that Red Light Bikes and Silicon Valley Bikes did not purchase any bikes or accessories from GBI is the first half of 2022.

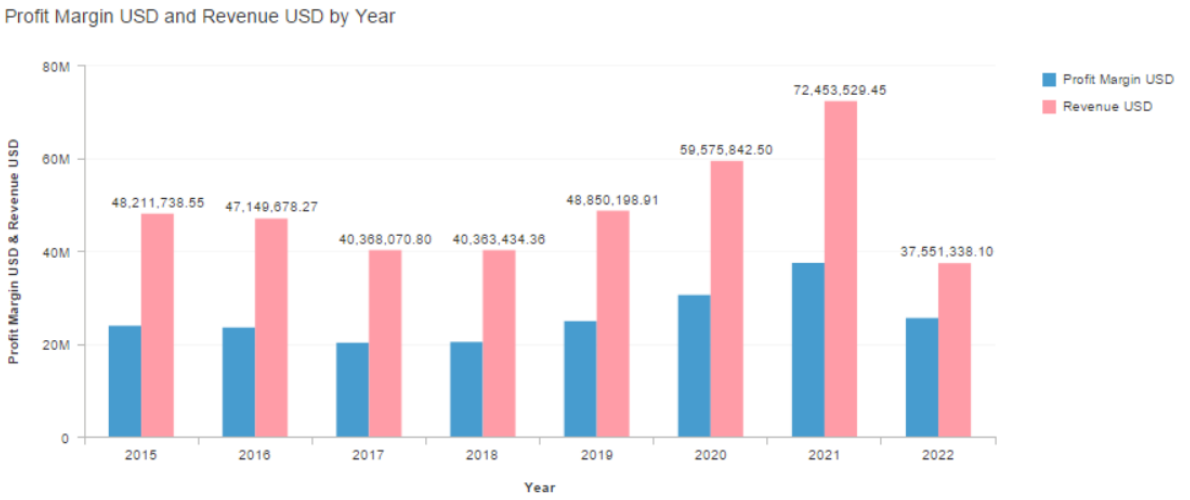


Figure 20. Profit Margin and Revenue by Year.

According to Figure 20, the profit margin and revenue of GBI is lowest in 2017 and 2018. The profit margin and revenue in 2017 and 2018 is approximately 20 million and 40 million, respectively. The profit margin dropped around 13.85% and the revenue dropped around 14.38% in 2017 compared to 2016. After the year 2018, the profit margin and revenue gradually increase and peak at 38 million and 72 million in 2021, respectively.

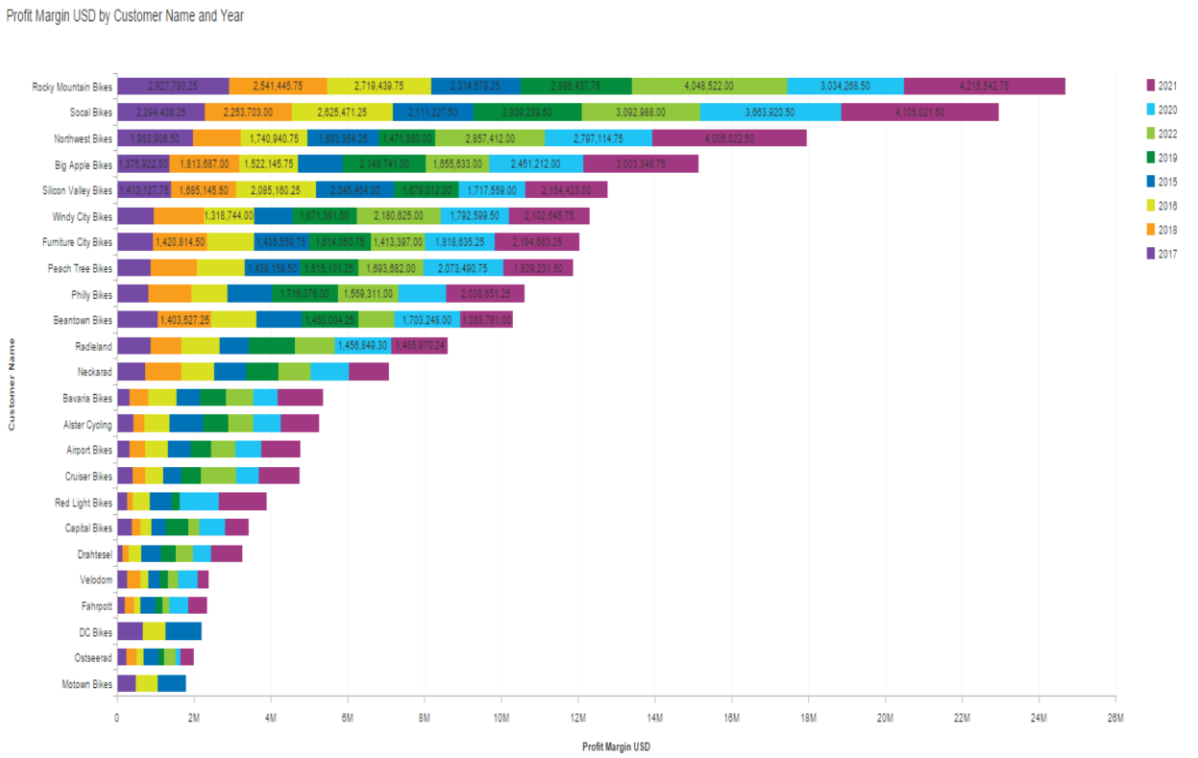


Figure 19. Stacked bar chart of Profit Margin by Customer Name and Year in descending order.

Figure 19 shows the stacked bar chart of Revenue by Customer Name and Year. Based on Figure 19, Rocky Mountain Bikes had purchased the most goods from GBI for the past 7.5 years, followed by Social Bikes and Northwest Bikes. Besides, it can also be seen that Silicon Valley Bikes ranks number 5 in purchasing goods from GBI, which is a fairly high ranking as well; whereas for Red Light Bikes, the company did not purchase many goods from GBI since it started doing business with GBI. Therefore, it can be assumed that Red Light Bikes is a small to medium size company, whereby it does not need to keep many goods to be sold to direct users. Motown Bikes and DC Bikes are at the bottom due to them stopped doing business with GBI after 3 years, since 2015.

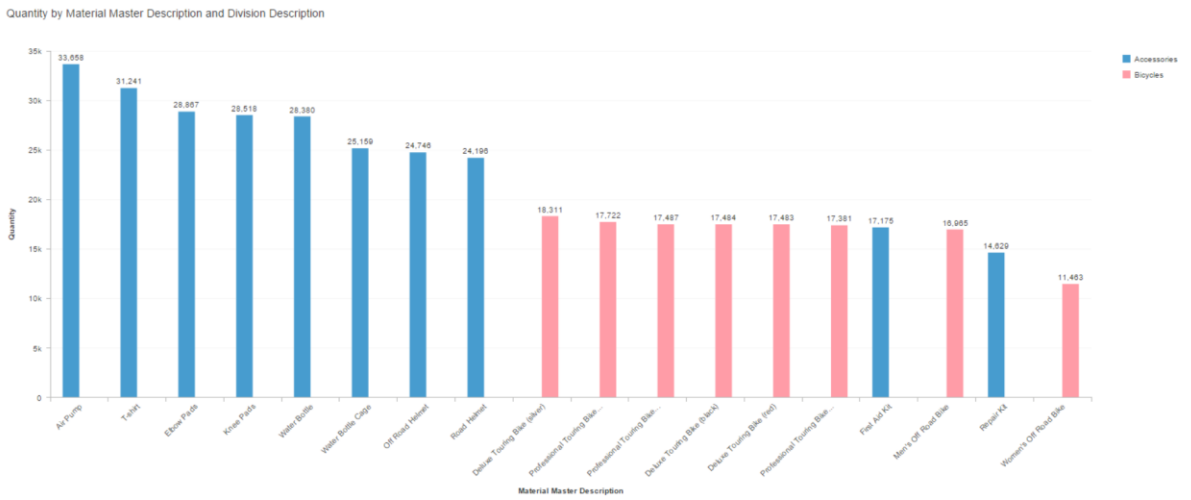


Figure 20. Quantity by Material Master Description in descending order.

Figure 20 shows the quantity of bikes and accessories purchased by the customers in descending order. Based on the figure, it can be seen that air pump has the highest number of sales, followed by T-shirts and elbow pads. As for the bikes category (coloured in pink), the bike with the highest sales is silver deluxe touring bikes, followed by red professional touring bikes and black professional touring bikes. The number of air pumps, T-shirts and elbow pads sold is 33,658, 31,241 and 28,867, respectively; whereas the number of silver deluxe touring bikes, red professional touring bikes, and black professional touring bikes sold is 18,311, 17,722 and 17,487, respectively.

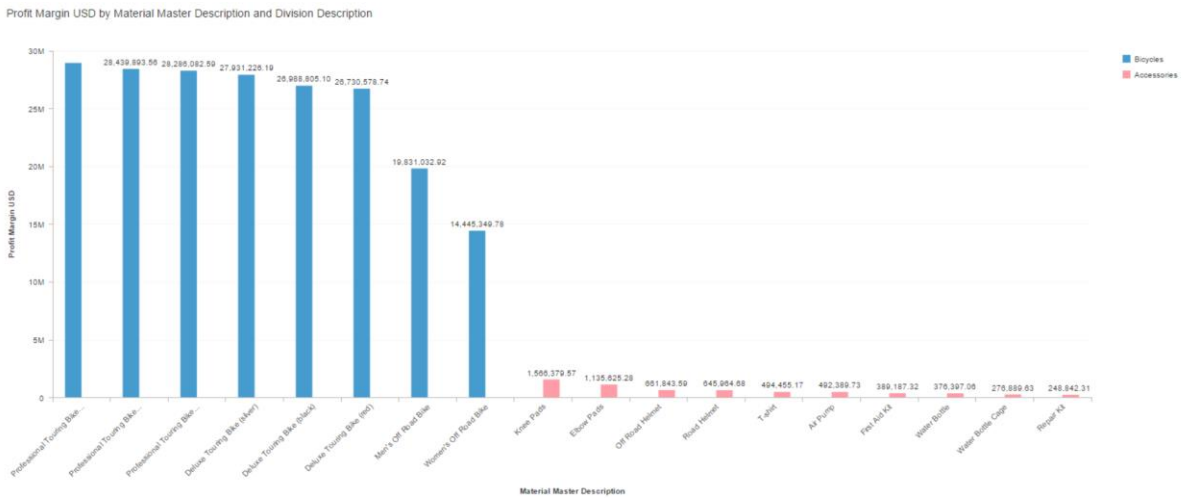


Figure 21. Profit Margin USD by Material Master Description and Division Description.

Based on Figure 21, it can be seen that the total profit made by professional touring bikes is the highest, followed by deluxe touring bikes, men's off-road bikes, and lastly women's off-road bikes. The total profit made by professional touring bikes is 85,681,642 USD. On the other hand, it is seen that the profits made by accessories is far not as much as the profits made by the bikes. The total profit made by the accessories is 6,257,970 USD. However, from the previous figure illustrated, it is seen that air pumps, T-shirts and elbow pads are the goods with the highest number of sales.

Based on Figure 20 and Figure 21, it can also be concluded that most customers prefer to buy touring bikes than off-road bikes as there were only 16,965 units of men's off-road bikes sold and 11,463 units of women's off-road bikes sold. Moreover, these two types of bike are one of the lowest saleable goods of GBI.

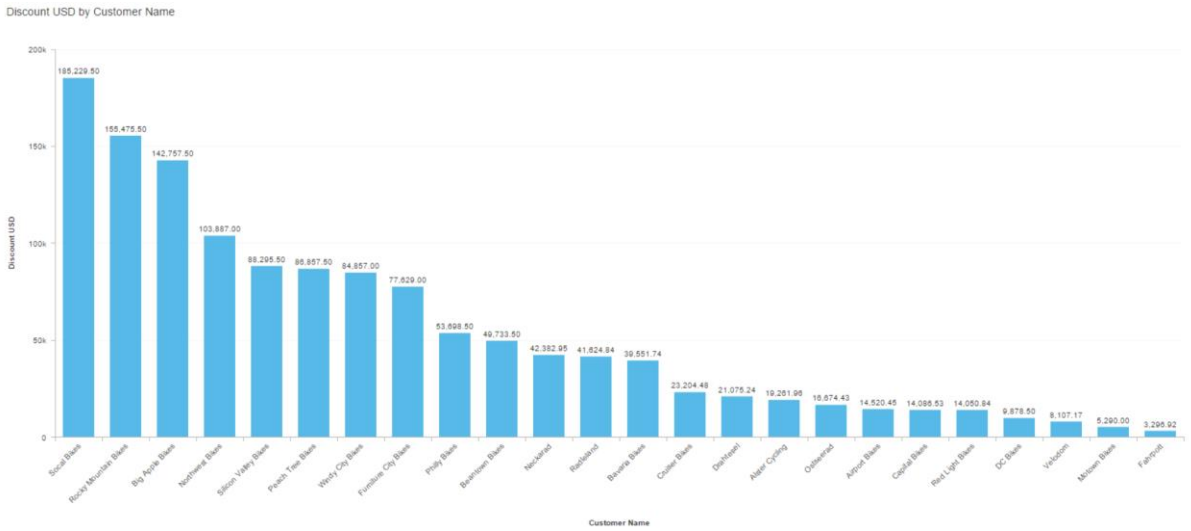


Figure 22. Discount USD by Customer Name.

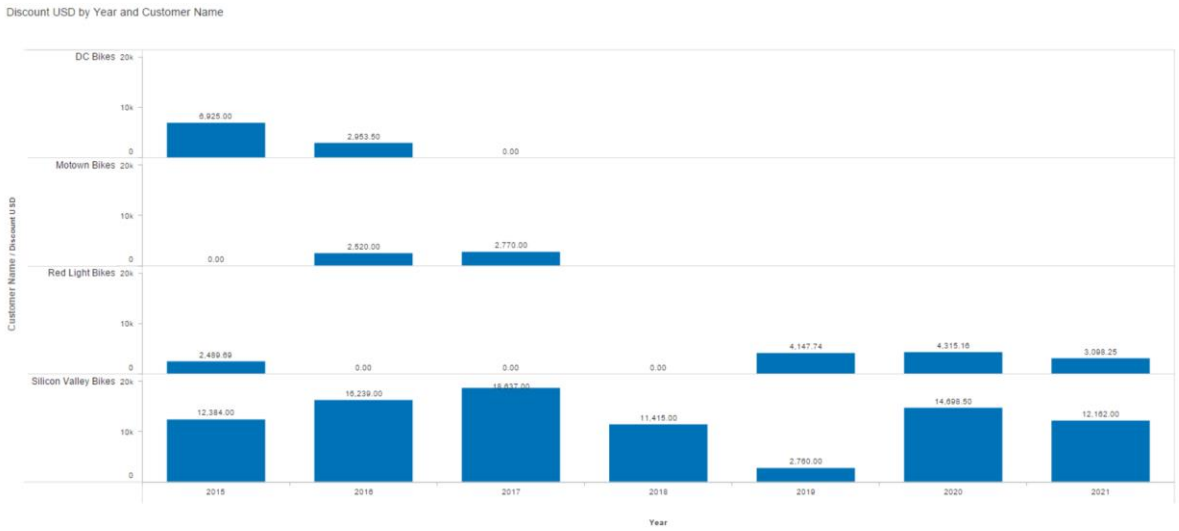


Figure 23. Discount USD by Customer Name and Year.

Based on Figure 22, it can be seen that Socal Bikes received the most discounts from GBI, followed by Rocky Mountain Bikes and Big Apple Bikes. It is also noticed that Silicon Valley Bikes ranks the fifth place in receiving discounts from GBI, with a total of 88,295 USD. On the other hand, Red Light Bikes also often got discount from GBI in 2019, 2020 and 2021. However, based on Figure 23, the discount that Silicon Valley Bikes and Red Light Bikes received in 2021 is lower than the discount that they got in 2020. As for DC Bikes, it did not get any discount in 2017, the year before it left GBI. And for Motown Bikes, it did not receive any discount in 2015, but got discount if both 2016 and 2017.

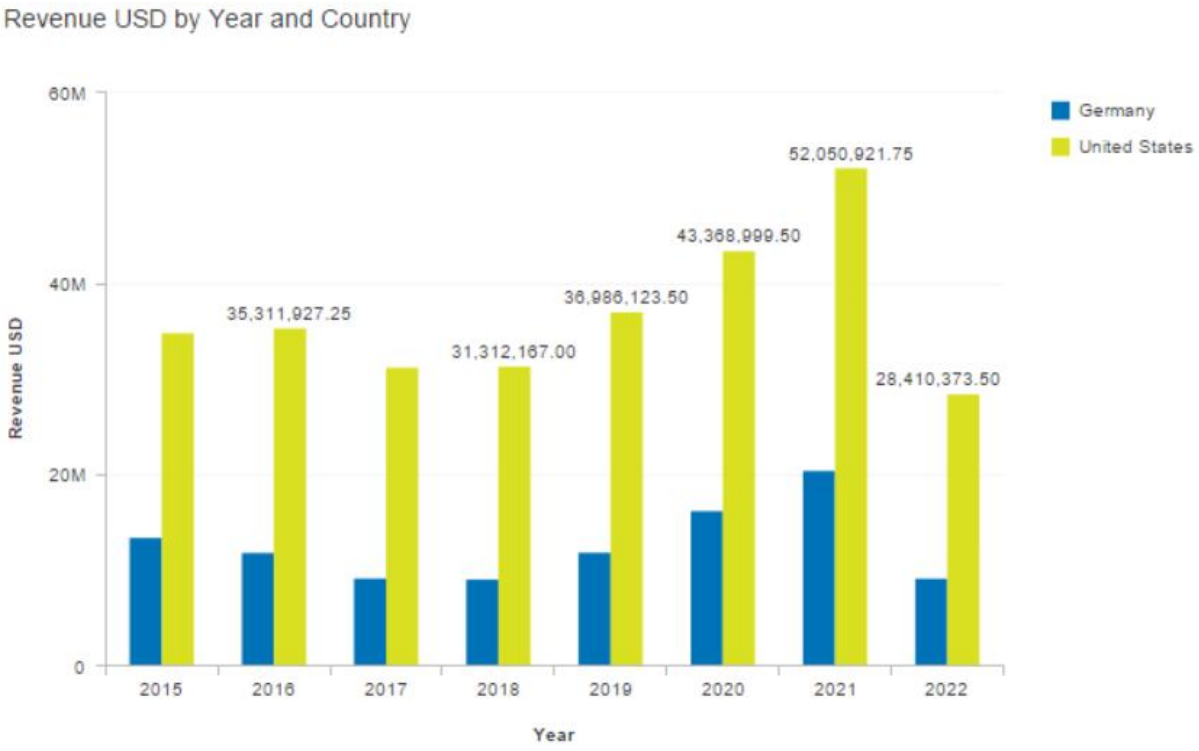


Figure 24. Revenue by Year and Country.

Based on Figure 24, the revenue generated by the German GBI is significantly lesser as compared to the American GBI. It can be seen that the revenue generated by the German GBI is more than 50% lower than the American GBI. This is also an important problem to be issued to avoid the company from losing money in the future.

3.3.3 Information Analysis

After a deeper understanding of the dataset through data analysis, information analysis can now be carried out. The aim of information analysis is to assess and evaluate what are the factors causing a certain business problem. In our case, we would like to evaluate and tackle the factors that are causing customers to leave GBI.

First and foremost, through data analysis, it can be seen that DC Bikes, Motown Bikes, Red Light Bikes and Silicon Valley Bikes had left GBI. DC Bikes and Motown Bikes left GBI in 2018, whereas Red Light Bikes and Silicon Valley Bikes left GBI in 2022. The purchasing behaviour of the customers was visualised, and it is seen that all of the customers purchased bikes and accessories throughout the year. This step is necessary as only the sales records of the first half of 2022 is recorded, this is to ensure that those customers do not only purchase on the second half of the year, but always purchase on the first half of the year as well. It was also noticed that Red Light Bikes and Silicon Valley Bikes purchased the greatest number of goods in the first quarter of the year, hence it can be confirmed that these two customers have already left GBI in 2022.

Other analyses were also carried out and it was also noticed that Silicon Valley Bikes was actually one of the top customers who had been purchasing a lot of goods from GBI, and GBI had earned a significant amount of profits by doing business with Silicon Valley Bikes. Therefore, it is kind of odd that Silicon Valley Bikes had left GBI. Besides, it was also noticed that Silicon Valley Bikes had consistently been receiving discounts from GBI, and thus proving that Silicon Valley Bikes has a good business relationship with GBI.

With that being said, we will be addressing about Silicon Valley Bikes first. According to the dataset, Silicon Valley Bikes is located in Palo Alto, California, USA. The biggest and most famous cycling events that is held in USA is Tour Divide, which is one of the most challenging mountain bike trial, starting from Canada to the border of Mexico. However, due to COVID-19, the Tour Divide was cancelled and resumed in 2021 (New York Times, 2021). Although the race immediately resumed in 2021, the business of Silicon Valley Bikes must have been impacted.

Moreover, although the race has been resumed and everything seemed to be normal, there were a lot of new variants for COVID-19 in 2020, such as the Delta and Omnicron variants. These variants had caused a surge in the COVID-19 cases in USA and most of the residents chose to do self-quarantine and did not want to get out from the house. Therefore, there might not be a lot of users purchasing bikes from Silicon Valley Bikes, causing Silicon Valley Bikes having a lot of goods left in their storage. Hence, they stop purchasing from GBI in 2022 and may continue purchasing from GBI when they have cleared their stock left in their storage.

Furthermore, it was reported that the Tour Divide will be resume on the 10th of June, 2021, but the course of the race will be shortened due to the borders of Canada and Mexico were not fully opened yet (New York Times, 2021). Hence, this might have caused the users and cyclists to lose interest in purchasing the bikes as most of the cyclists would prefer a more challenging race. Cyclists might also not purchase new bikes due to the course being less challenging, the cyclists do not need bikes that are as high end.

One of the most important multi-stage road bicycle race in Germany is the Deutschland Tour. In 2008, it was announced that the Deutschland Tour will be halted without any further notice on the date that the race will be resumed (Cyclingnews, 2008). In 2018, the Deutschland Tour resumed but the quantity of goods sold by GBI, Germany remains low. Also, the tour was cancelled again in 2020 due to the German government's ban on mass gatherings through August 31 in an effort to manage the COVID-19 pandemic (Cyclingnews, 2020). Based on the dataset, Red Light Bikes was located in Hamburg, Germany. Hence, the frequent cancellation of the Deutschland Tour might have a big impact on the bicycle business for Red Light Bikes.

According to World Bank (2022), compounding the effects of the COVID-19 epidemic, Russia's invasion of Ukraine has exacerbated the global economic slowdown, which may lead to a prolonged period of weak growth and high inflation. This increases the possibility of stagflation, which might be detrimental to economies of low to medium income levels. Global growth is projected to decline from 5.7% in 2021 to 2.9% in 2022. Therefore, almost all businesses are affected by the economic crisis in 2022. This might cause the shutdown of a few businesses. As mentioned above, an inference was made that Red Light Bikes might be a small to medium size business as it did not purchase many goods from GBI since it has started doing business with GBI. Hence, there might be a high chance that Red Light Bikes was greatly affected by the economic crisis, and their business was forced to be shutdown.

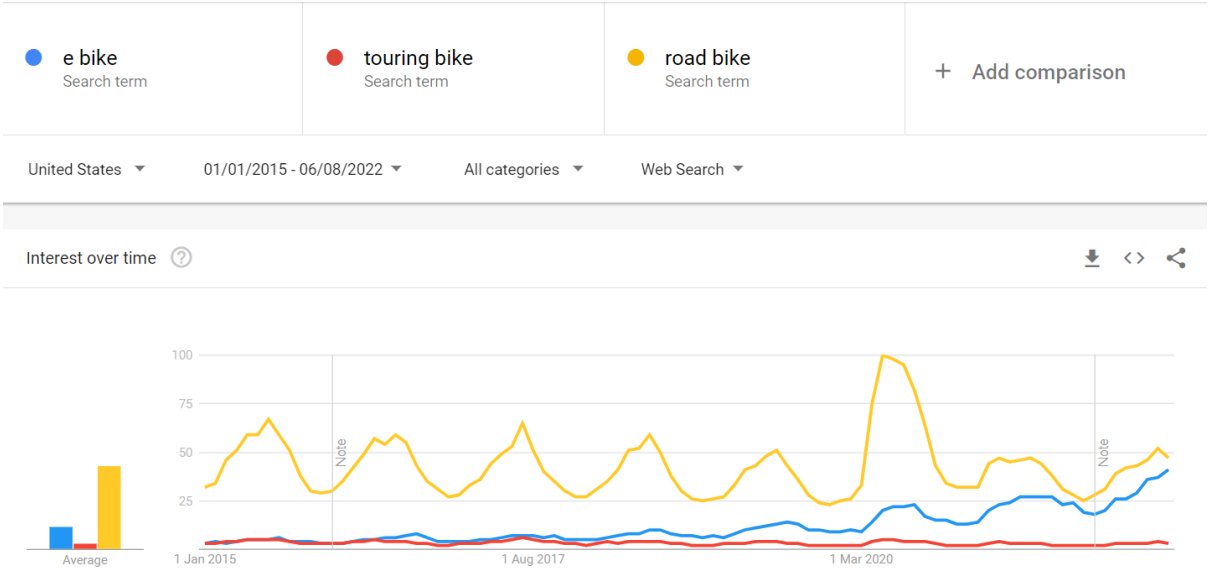


Figure 25. The search trends of e-bikes, touring bikes, and road bikes in the United States.

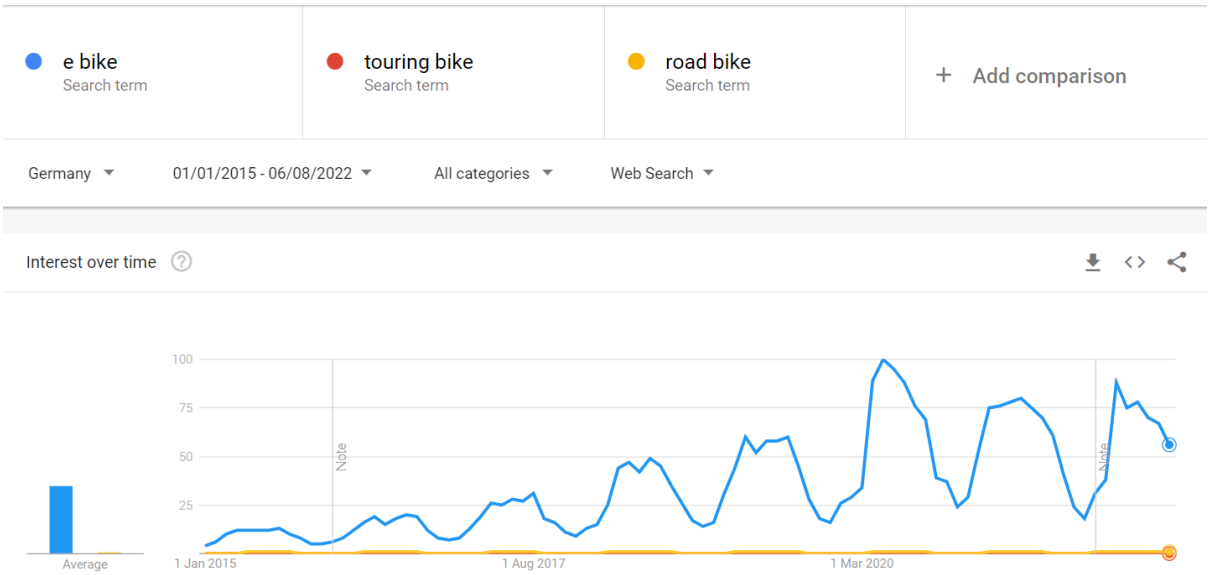


Figure 26. The search trends of e-bikes, touring bikes, and road bikes in Germany.

Figure 25 and Figure 26 show the search trends of e-bikes, touring bikes, and road bikes in the United States and Germany since 2015. Based on Figure 25, it can be seen that road bikes had gained a lot of interests since March 2020. As we know, in the beginning of 2020 is when COVID-19 was the most serious and the surge of COVID-19 cases happened in the first half of 2020. Hence, due to movement control order (MCO), residents in the United States were not allowed to have mass gatherings or dine-in in restaurants, as well as go to the gym. Thus, the alternative exercise that the cyclists can do is to cycle on the road, leading to a high number of

searches for road bikes. This can also be seen in Figure 24, where the revenue of GBI increased in 2020.

Based on Figure 26, it can be seen that the search trend of e-bikes gradually increases in Germany and also peaked in March 2020. This might be due to the same reason as mentioned above in the scenario of United States, except for the Germans preferred e-bikes over road bikes. E-bikes have been said to be able to improve your fitness, help you keep up with your meters, and it is easier to get up hill (Bike Radar, 2021). Besides, several types of e-bikes are now available for rental all over Europe, especially Germany. Hence, this might be the reason that the revenue of the German GBI is much lower than the revenue of the American GBI. This might also be one of the reasons that Red Light Bikes had left GBI. There might also be a chance where Red Light Bikes stopped selling touring bikes and off-road bikes, and started doing e-bikes businesses.

3.3.4 Knowledge Analysis

The two pillars of any business are its primary objectives and the method for achieving them. The third component is the actualization of the plan. Without a strategy, the path to accomplishing goals is not well-defined, and the business will encounter hurdles for which there are no immediate solutions to move forward (Lazzari, 2019). The research of Kassim et al. (2009) has proven that a good customer relationship marketing (CRM) strategy is effective in retaining customers. Businesses should focus and prioritise on components that greatly boost their business performance, particularly by retaining loyal customers and acquiring new ones in order to increase their relationship marketing strategies. In CRM, the satisfaction of customers is the most important variable. As a result, GBI's CRM strategy should place greater emphasis on customer satisfaction. The second most important factor is employee dedication, followed by customer confidence and customer loyalty. It was shown that loyalty is not a significant relationship marketing strategy aspect. This does not imply, however, that GBI should disregard this factor. GBI must still prioritise acquiring long-standing, devoted customers.

Kormin and Baharun (2016) suggested that businesses can use social media platforms and relationship marketing strategies to retain customers. This research demonstrates that organisations utilise numerous social media channels to engage with their customers. Additionally, they routinely update their Facebook with an average of one post each day, and they share posts that are aligned with customers' goals. In addition, this survey demonstrates that brands utilise images more often than video and text-only content. Moreover, in terms of marketing performance, brands acquire a good rate of engagement on average; hence, social media has been demonstrated to be a viable alternative marketing technique that must be utilised by all brands and businesses in order to obtain free publicity with low expense and effort.

The research of McCoy et al. (2019) has demonstrated that there is a significant relationship between bicycles and cycling infrastructure, and that the development of cycling infrastructures will have a direct impact on the sales of related businesses if they are not developed well over time. This is due to the fact that when cycling infrastructures are not developed well, the risk of injury also rises, which decreases the demand for cycling and the likelihood of citizens purchasing bicycles. Thus, the government plays a crucial part in the development of a

country's plan, and many economic, political, and social variables must be taken into account so that the government can effectively and efficiently address diverse situations.

From the organisation understanding section, we have understood that GBI is running a business-to-business (B2B) style, whereby GBI distributes its bicycles exclusively through reputable Independent Bicycle Dealers (IBDs). This may be a good strategy as then GBI does not have to store a huge number of bikes in their storage but can directly distribute to IBD according to orders made. However, it is also good that if GBI can do business-to-customer (B2C) business as this can improve customer's experience. For instance, GBI can provide a test ride to direct users when they wish to purchase a bicycle. This can allow the direct users to choose and purchase the bikes that are most suitable for them and tailored to their needs. Besides, when there is any technical problem associated with the bikes, the users can also send the bikes directly to GBI workshop for repair, instead of having to wait for at least a week to get their bikes back by sending to IBDs, as IBDs need to send back the bikes to GBI's headquarter for repair.

The desire for environmentally friendly transportation solutions can be connected to the increased awareness of environmental challenges. Studies on consumer behaviour have shown that consumer innovativeness and green perceptions are important factors in the process of implementing environmentally friendly developments (Flores & Jansson, 2021). Today, e-bikes are seen as a more environmentally friendly substitute for traditional automobiles. The e-bike offers all of the advantages that are available from traditional bicycles, in addition to being able to travel further, more quickly, and in more comfort, all while requiring less effort from the rider (Apostolou et al., 2018). Based on Figure 26, it is seen that the search trend of e-bikes increased tremendously in Germany, indicating that most of the residents in Germany are now using e-bike. Besides, e-bikes are on high demand all over Europe. Therefore, it is suggested for GBI to start inventing e-bikes in order to attract more new customers, as well as increasing the chance to retain existing customers. This is because existing customers can also consider selling e-bikes to direct users to fulfil the interest and market's demand.

4.0 Insights Extraction

After performing information analysis and knowledge analysis against the dataset, a lot of data-driven insights can be extracted from the analyses. First and foremost, it is noticed that there were a few customers who left GBI, namely DC Bikes and Motown Bikes left in 2018; whereas Red Light Bikes and Silicon Valley Bikes left in 2022. It is important to tackle this phenomenon by performing analysis and proposing suitable business solutions in order to avoid losing more customers in the future. Not only retaining customers, but it is also important to attract more customers to partner with GBI to ensure that the business can grow stronger.

From the data analysis, it is shown that Silicon Valley Bikes was one of the top customers of GBI, according to Figure 19. Based on Figure 22, it is also shown that Silicon Valley Bikes also received quite a lot of discounts from GBI throughout the years. These indicate that Silicon Valley Bikes has a good business relationship with GBI. Hence, it is important to find out why did Silicon Valley Bikes stop purchasing bikes and accessories from GBI. It is suggested for the sales manager of GBI to pay Silicon Valley Bikes a visit to understand what is the reason that caused Silicon Valley Bikes to leave GBI and try to persuade or offer more discounts to Silicon Valley Bikes to sustain the relationship between the two companies. Besides, if Silicon Valley Bikes is not satisfied with anything related to GBI, i.e., product quality, product price, and employees' attitude, GBI should immediately address the problem and solve the problem. Of course, not only for Silicon Valley Bikes, but GBI should also pay a visit to Red Light Bikes as these two companies have been doing business with GBI for a very long time and they only left GBI recently.

Besides, the outbreak of COVID-19 at the end of 2019 has left a major impact on all businesses. Hence, as suggested in the knowledge analysis section, GBI can utilise several social media platforms to engage with their customers. They can routinely update their Facebook and Instagram pages with attractive images and video content. This can greatly boost the popularity of their brand and knowability to potential customers and business partners. Moreover, in terms of marketing effectiveness, brands gain a strong rate of engagement on average; therefore, social media has been proved to be a viable alternative marketing approach that all brands and enterprises must employ in order to obtain free publicity with less money and effort.

Additionally, a lot of countries had experienced lockdown and movement control order in order to manage the COVID-19 situation, as the number of cases is going up every day. Thus, during this period of time, mass gatherings and sports events were not allowed to be carried out. The Divide Tour in US and the Deutschland Tour in Germany were affected and was forced to be cancelled. Hence, during this time, the customers of GBI might be greatly affected as there will not be many cyclists who will purchase new bicycles. It is suggested that GBI can come out with several marketing campaigns to boost the number of sales. Besides, GBI can also grab this chance to promote cycling activity outdoors since residents are also unable to go to the gym during the lockdown. Hence, GBI can encourage the residents to do cycling as an alternative exercise.

Based on Figure 22 in the data analysis section, it can be seen that some customers get a lot of discounts from GBI while some only got very less discounts. Hence, this is a very important issue to be addressed. It is seen that Socal Bikes, Rocky Mountain Bikes and Big Apple Bikes received the most discounts, and according to the profit gained from the customers, these three customers rank in the top 5 list. This indicates that the revenue generated by the customers is directly proportional to the amount of discounts given. Thus, GBI needs to replan their discount scheme in order to give a fair amount of discounts to every customer. This will significantly increase the satisfaction of the customers, and at the same time, it will boost the relationship between GBI and the customer.

As mentioned above, it is crucial to appreciate the loyalty of the customers that have been doing business with GBI for many years so that the foundation of the business relationship between the companies can be more rigid. GBI can organise company events occasionally and invite the customers to express GBI's appreciation towards the customers. Besides, they can also offer the customers bundle price or discount coupons to encourage the customers to continue purchasing from GBI. This can significantly improve a good business relationship with the customers.

GBI can also consider including the business-to-customers (B2C) business strategy in order to improve users' experience. If GBI includes the B2C strategy, they can provide a test ride on the bikes for users so that the users can choose the bikes that are most suited to their needs. Besides, if there is something wrong with their bikes, they can immediately send the bikes to

GBI workshops instead of having to wait for one to two weeks to get their bikes back. Other than that, the B2C business strategy can sustain more customers and increase customer loyalty.

In this era, as the interest and demand of users for e-bikes is increasing, GBI should grab the chance to develop and produce e-bikes. As shown in Figure 26, the search trend for e-bikes in Germany increased and peaked in March 2020. Hence, GBI may benefit from producing and selling e-bikes to customers. This can also attract more potential customers and business partners in purchasing e-bikes from GBI. Not only this can increase the revenue of GBI, but it is also in line with GBI's mission, which is to provide innovative, high-performance bicycles to riders who place the highest demand on their equipment.

With that being said, GBI must also always ensure that the bikes that they produce have the highest quality and highest performance. GBI should not compromise on the quality of their bikes and implement strict quality control (QC) and quality assurance (QA) on the bikes that they produced. This is because when the quality of the bikes decreases, it will leave a big impact on the customers, and the customers will lose confidence in the company. GBI must always remember the vision and mission of the company in order to gain customer confidence, customer loyalty and customer satisfaction.

5.0 Final Deliverables

After performing data analysis, information analysis and knowledge analysis to tackle the problem of customers leaving GBI, several effective business solutions can be proposed to GBI. The business solutions are summarised and listed in point forms below:

- To pay a visit to Red Light Bikes and Silicon Valley Bikes to understand the reasons causing them to leave GBI.
- To utilise social media platforms to engage with their customers in order to boost the popularity of their brand.
- To promote the advantages of cycling and organise several cycling events to spark interest in the users in cycling through marketing campaigns.
- To rework their discount scheme and give a fair discount rate to every customer to increase the satisfaction of the customers and improve customer loyalty.
- To maintain good relationships with the customers that have been doing business with GBI for a long time by organising company events occasionally and inviting loyal customers.
- To include a business-to-customers (B2C) business strategy to gain more customers.
- To provide repair services to users when the bikes have technical errors or when the bikes are broken. The users should be able to send their bikes directly to GBI's workshop for repair.
- To adapt to the changes in the market post-COVID by carrying out R&D and producing e-bikes to fulfil the interest and demands of the market and attract new potential customers.
- To always adhere to the company's vision and mission and to always produce bikes with the highest quality and highest performance.

6.0 Conclusion

In this assignment, an in-depth data analysis had been carried out against GBI sales dataset to identify potential business problems that should immediately be addressed. Through the descriptive data analysis, one of the business problems that we have identified is there were a few customers, namely DC Bikes, Motown Bikes, Red Light Bikes, and Silicon Valley Bikes stopped doing business with GBI. Hence, data analysis and information analysis were carried out using a visualisation tool, SAP Lumira Discovery, to identify and understand the reasons that caused the customers to leave GBI. After that, knowledge analysis and insights extraction were carried out to tackle the business problem and effective business solutions were proposed. Through the visualisations and analyses, the aims and objectives of this assignment were achieved. It is important for businesses to actively identify business problems and immediately look for effective business solutions in order to provide better quality service to customers, and also maintain a good relationship with customers, subsequently leading to the success of the business. It is dangerous to be indifferent to even a small business problem and this would eventually cause the business to gradually go downhill. Hence, data analysts and data scientists play an important role in a business and should actively be identifying problems from every department's point of view.

REFERENCES

1. *14 benefits of riding an electric bike*. BikeRadar. (2022). Retrieved 6 August 2022, from https://www.bikeradar.com/features/ebike-benefits/.
2. *2009 TOUR OF GERMANY CANCELLED AFTER DOPING SCANDALS*. cyclingweekly.com. (2022). Retrieved 6 August 2022, from https://www.cyclingweekly.com/news/latest-news/2009-tour-of-germany-cancelled-after-doping-scandals-92244.
3. *A 2,700-Mile Cycling Race Is Now Even More Extreme*. Nytimes.com. (2022). Retrieved 5 August 2022, from https://www.nytimes.com/2022/07/04/sports/tour-divide-race-bikepackers.html#:~:text=This%20year's%20Tour%20Divide%20was,of%20pandemic%2Drelated%20border%20closures.
4. *August's Deutschland Tour cancelled*. cyclingnews.com. (2022). Retrieved 6 August 2022, from https://www.cyclingnews.com/news/augusts-deutschland-tour-cancelled/.
5. Apostolou, G., Reinders, A., & Geurs, K. (2018). An overview of existing experiences with solar-powered e-bikes. *Energies*, *11*(8). https://doi.org/10.3390/en11082129
6. Education, I. (2022). *What is ETL (Extract, Transform, Load)?*. Ibm.com. Retrieved 4 August 2022, from https://www.ibm.com/cloud/learn/etl.
7. Erl, T., Buhler, P., & Khattak, W. (2016). *Big Data Fundamentals Concepts, Drivers & Techniques* (pp. 30 - 35). ServiceTech Press.
8. Flores, P. J., & Jansson, J. (2021). The role of consumer innovativeness and green perceptions on green innovation use: The case of shared e-bikes and e-scooters. *Journal of Consumer Behaviour*, *20*(6), 1466–1479. https://doi.org/10.1002/cb.1957
9. Huber, S., Wiemer, H., Schneider, D., & Ihlenfeldt, S. (2019). DMME: Data mining methodology for engineering applications - A holistic extension to the CRISP-DM model. *Procedia CIRP*, *79*, 403–408. https://doi.org/10.1016/j.procir.2019.02.106
10. Hotz, N. (2022). *What is CRISP DM? - Data Science Process Alliance*. Data Science Process Alliance. Retrieved 4 August 2022, from https://www.datascience-pm.com/crisp-dm-2/.
11. *IBM Docs*. Ibm.com. (2022). Retrieved 3 August 2022, from https://www.ibm.com/docs/en/spss-modeler/saas?topic=dm-crisp-help-overview.
12. Kassim, K. M., Bahari, A., Kassim, N., Rashid, N. R. N. A., & Jusoff, K. (2009). Retaining Customers through Relationship Marketing in an Islamic Financial Institution in Malaysia.

International Journal of Marketing Studies, *1*(1), 66–72. https://doi.org/10.5539/ijms.v1n1p66

13. Kormin, K., & Baharun, R. (2016). Social media and relationship marketing strategies of malaysian car brands. *Indian Journal of Science and Technology*, *9*(46). https://doi.org/10.17485/ijst/2016/v9i46/107138
14. Kuilboer, J., Russ, H., & Ashrafi, N. (2010). Business intelligence as an enabler of organizational agility. *Beachvolleyball.Guidedbees.Com*, 1–10. http://beachvolleyball.guidedbees.com/~ruben/EDSI/papers/kuilboer_business.pdf
15. Lazzari, Z. (2019). *What Is Strategic Perspective?*. Small Business - Chron.com. Retrieved 6 August 2022, from https://smallbusiness.chron.com/strategic-perspective-14365.html.
16. Lenzerini, M. (2002). Data integration. *Proceedings of the Twenty-First ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems - PODS '02*, 233. https://doi.org/10.1145/543613.543644
17. Magal, S., Weidner, S., & Word, J. (2014). Global Bike Inc. Retrieved July 25, 2021, from http://www3.cis.gsu.edu/dtruex/courses/CIS8670/SAP%20homework%20and%20exercises/Exercise%201-%20GBI%20Intro/Intro_ERP_Using_GBI_Story[Letter]_en_v2.30.pdf
18. McCoy, R., Poirier, J. A., & Chapple, K. (2019). Bikes or Bust? Analyzing the Impact of Bicycle Infrastructure on Business Performance in San Francisoco. Journal of the Transportation Research Board, 2673(12), 277-289.
19. Niakšu, O. (2015). CRISP Data Mining Methodology Extension for Medical Domain. *Baltic J. Modern Computing*, *3*(2), 92–109.
20. *OLAP vs. OLTP: What's the Difference?*. Ibm.com. (2022). Retrieved 3 August 2022, from https://www.ibm.com/cloud/blog/olap-vs-oltp.
21. S.Jensen, C., & Snodgrass, R. T. (2017). Valid Time. In *Encyclopedia of Database Systems*. https://doi.org/10.1007/978-1-4899-7993-3_1066-2
22. *Stagflation Risk Rises Amid Sharp Slowdown in Growth*. The World Bank. (2022). Retrieved 5 August 2022, from https://www.worldbank.org/en/news/press-release/2022/06/07/stagflation-risk-rises-amid-sharp-slowdown-in-growth-energy-markets.
23. ULARU, E. G., PUICAN, F. C., APOSTU, A., & VELICANU, M. (2017). Contents :Perspectives on Big Data and Big Data Analytics Elena. *Database Systems Journal*, *III*(4), 1–79.

24. Wang, M., Zhang, X., & Hwang, D. (2011). OBJECT DATA REUSE AND INTEGRATION IN ERP. *Issues In Information Systems*, *12*(1), 392–400. https://doi.org/10.48009/1_iis_2011_392-400
25. *What is SAP HANA? | IBM*. Ibm.com. (2022). Retrieved 3 August 2022, from https://www.ibm.com/my-en/topics/sap-hana.
26. *What is SAP HANA?*. SAP. (2022). Retrieved 3 August 2022, from https://www.sap.com/sea/products/technology-platform/hana/what-is-sap-hana.html.
27. *What is the Definition of OLAP? OLAP Definition*. OLAP.com. (2022). Retrieved 5 August 2022, from https://olap.com/olap-definition/.