

DATABASES AND DATA VISUALIZATION FINAL GROUP PROJECT
GROUP 16

Marharyta Dzhafarova (13443925)

Miras Kanatzhanov (13075012)

Shantanu Motiani (13757830)

Felicia Otto (13053949)

Megu Tsujibayashi (13503499)

Case Description

This project is aimed at improving the overall performance of the company specializing in selling bikes, biking accessories, clothing, and gear. At the beginning of the project, we were provided with the company's dataset "salesdeveloperdb" (see *Figure 1*). Initially, the dataset contained 15 tables with data, providing perspective on the company's employees, suppliers, products, and various details about the orders.

The "product" table specifies the physical characteristics of each product as well as details about the target stock levels of each product and the sales periods of each product. The table also contained the column StandardCost, which was useful in determining the Total Cost of Goods Sold and consequently Gross Profit. The "vendor" table specifies the details of the suppliers. The CreditRating column was used to assess the vendors' performance as part of one of the KPIs. The "productvendor" table served as a junction table between "product" and "vendor", but it wasn't used in further analysis. The "employee" table contained information about the background as well as data about the base rate salary and the number of vacation hours and sick leave hours that were needed for our further analysis. However, no foreign key connection of this table was recognized to any other table. The "salesorderheader" specified the details of each order, including the customer, for whom the order was shipped and the territory, where the shipping occurred, connecting it to the "salesteritory" and "customer" table. The "salesorderdetail" table was very important to our analysis as it contained the columns UnitPrice and OrderQty, which helped derive the Total Sales and consequently Gross Profit of the company. The table "address" contained the details of the territories, where the business operates. The column CountryRegionCode was especially useful for our analysis when displaying the performance indicators per region.

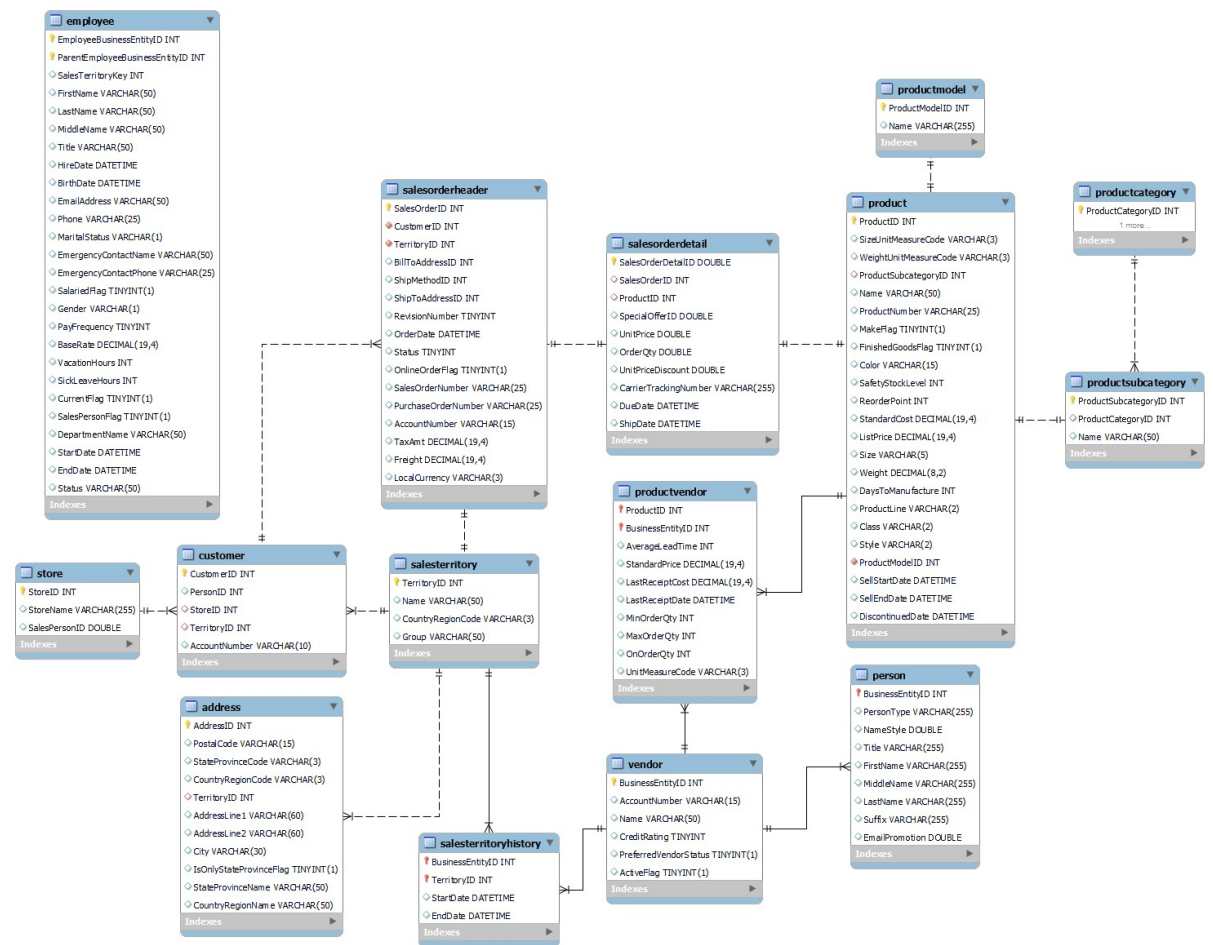


Figure 1. "salesdeveloperdb" database

Key Performance Indicators

We created two KPIs for each perspective of the balanced scorecard, totalling 8 goals. *Figure 2* shows the full overview of our Balanced Scorecard.

Financial Perspective	Customer Perspective	Learning and Growth Perspective	Internal Business Processes Perspective
<p>1- Gross Profit Margin by Product Category and by Region</p> <ul style="list-style-type: none"> Assess the profitability of different product categories, providing insights into which categories contribute the most to the gross profit margin. Measure the profitability of each region by evaluating the percentage of sales retained as gross profit. Formula: $(\text{Total Sales} - \text{Total Cost of Goods Sold}) / \text{Total Sales}$ <p>2- Average Transaction Value</p> <ul style="list-style-type: none"> Measure the average value of a transaction. Provide insights into the purchasing behavior of customers and can help identify opportunities to increase the average transaction value. Formula: $(\text{Total Sales} / \text{Number of Transactions})$ 	<p>1- Customer Retention Rate</p> <ul style="list-style-type: none"> Measure the percentage of customers who continue to make purchases, indicating the effectiveness of customer retention efforts. Formula: $(\text{Number of Repeat Customers} / \text{Total Number of Customers}) * 100$ <p>2- Region Order Frequency</p> <ul style="list-style-type: none"> Calculate how frequently, on average, an order is placed in each region. This KPI helps in understanding region engagement and loyalty. Formula: Average number of orders per day per country 	<p>1- Gender Pay Equity Index by Department</p> <ul style="list-style-type: none"> Assess gender pay equity by comparing the average pay for female employees to the average pay for male employees by department. A value close to 100 indicates greater pay equity, while deviations may indicate potential salary gaps. Formula: $(\text{Average Pay for Females} / \text{Average Pay for Males}) * 100$ <p>2- Employee Well-being Index by Department</p> <ul style="list-style-type: none"> Measure the percentage of total work hours that employees spend on sick leave or vacation. A lower index indicates a healthier work-life balance and potentially higher employee satisfaction. Formula: $((\text{Total Sick Leave Hours} + \text{Total Vacation Hours}) / \text{Total Work Hours}) * 100$ 	<p>1- Order Fulfillment Cycle Time</p> <ul style="list-style-type: none"> Measure the average time it takes for the vendor to fulfill orders, helping assess the efficiency of internal order processing. Formula: $(\text{Average Ship Date} - \text{Order Date})$ <p>2- Vendor Credit Rating Index</p> <ul style="list-style-type: none"> Measure the average credit rating across all vendors. A higher index indicates, on average, better credit ratings among vendors. Formula: (Average Credit Rating)

Figure 2. Balanced Scorecard

Financial Perspective

Gross Profit Margin by Product Category and by Region

Gross Profit Margin is one of the most important measures of financial performance for a retail company. It is calculated by dividing the difference between Total Sales and Total Cost of Goods Sold by Total Sales. This gives us a perception of the average profitability of each sale, as this is the main part of each retail business process. The Gross Profit Margin per product category helps the business compare different items they are considering stocking based on their profitability and set a price so that it generates a profit, attracts clients, and offsets the expenditures incurred by the company (“What Is”, 2022).

In the case of the researched company, bike accessories turned out to be the most profitable product category with a gross profit margin of about 50%, the least profitable product category being bikes and bike components with gross profit margins of 8.86% and 4.20% respectively (see *Figure 6*). Although the high profitability of accessories might be related to low manufacturing costs and therefore low and therefore low wholesale prices, we would still advise the management to leverage the success of accessories to inform marketing strategies, such as including a bigger share of accessories in the product mix. As for the low

profit margins for the bikes and bike components, we recommend evaluating the supply channels of the bikes and the efficiency of the current cost structures.

The research company is operating in 6 countries on four different continents, therefore we found it essential to investigate the profitability of operations in every country. The highest gross profit margin was observed in Australia (32.39%), while the lowest gross profit margins are seen in the biggest regions - the United States and Canada (see *Figure 7*). Such a big difference is a prompt for the company to consider pricing strategies, targeted to each region, based on the level of competitiveness and the number of target customers. We would also advise to look at the efficiency of having operations in this region. It might be more beneficial to concentrate the operations in fewer regions, which might increase operational efficiency and reduce expenses.

Average Transaction Value

The average transaction value indicates how much customers spend on each purchase. Because it helps to determine which price ranges are generating the most volume at checkout, the average transaction value is a very useful measure for retailers. This information is useful in making decisions about pricing and product strategy (“What Is”, 2023).

The results of this calculation were surprising as they were not in line with the gross profit margin. The regions with the highest transaction value, also have the lowest gross profit margin (US and Canada) (see *Figure 7*). This might be related to the fact that more expensive products, such as bikes and gear are bought in those regions (average transaction value \$5,421 and \$4,254), while in Australia (average transaction value \$1,601) cheaper products, such as Accessories are more popular, they, in turn, contribute to the biggest profit margin. We recommend the company take this data into account when making decisions about product distribution between different regions of their operations.

Customer Perspective

Customer Retention Rate

Customer Retention Rate has gained recognition due to its ability to produce superior relationship economics—that is, the cost of retaining customers is lower than that of acquiring new ones. Moreover, this is one of the best ways to build loyalty, which is essential for a competitive retail industry. Therefore, it became one of the primary goals of relationship marketing (Ahmad & Buttle, 2002).

We have calculated the customer retention rate per region, by dividing the number of customers, who have placed the order more than 2 times, by the total number of customers, who have placed an order. Following the gross profit margin trend, Australia shows the biggest customer retention rate of 37.84%, while the United States portrays the lowest rate of 17.43% (see *Figure 8*). The rates for the European regions are fluctuating near 20%. The reason for the low retention rate in the US and Canada might be also a highly competitive market. Those branches should consider conducting customer surveys and investigating competitors' strategies regarding customer retention. For European regions, it is advised to consider the strategies that the Australian branch is using, this might include initiatives and loyalty programs for existing customers.

Region Order Frequency

The order frequency KPI is an important metric for a retail business as it helps to control the number of personnel and the efficiency of front office operations in general. This measure was calculated by taking the average amount of orders each region has per day.

This metric is mostly dependent on the density of the population in a certain area. Therefore, the highest region order frequency is observed in the US (334 orders per day). The second place is taken by Australia (191 orders per day), this is logical as Australia is a leading branch regarding gross profit margin and customer retention rate (see *Figure 8*). France and Germany have significantly lower metrics (74 and 73 orders per day respectively). Although this might be related to the fact that both of those markets are relatively small in size, the management should still consider improving targeted marketing strategies to improve customer engagement.

Learning and Growth Perspective

Gender Pay Equity Index by Department

The gender pay gap has proven to be an enduring issue in the contemporary work environment. Gender inequality is a form of discrimination and can hinder businesses' decision-making processes. *Figure 9* shows the proportions of wages between female and male employees by department. The table shows clear income differentiation for positions at an executive level where male compensation is higher than female compensation. Furthermore, the Product Control and Quality Assurance departments do not consist of

female employees. This implies that in general, this company consists of fewer female workers than male workers.

Research by Hultin and Szulkin (2003) suggests that structural discrimination is more common in companies like this. There may be higher intrinsic motivation by male executive members to keep gender bias when distributing power and responsibility to other employees. However, the same gender bias also applies inversely, where female power distribution is greater for organizations with higher female representation. While it is difficult to measure distinct discrimination or equality, recruiters and employees in the HR department should be aware and take into consideration that the company has a higher risk of gender bias from empirical research. This will prevent future problems in which high-quality workers are overlooked because of biased decisions.

Employee Well-being Index by Department

The employee well-being KPI was calculated based on employees' sick leave hours to the total working hours (160 hours per month). This KPI is especially important because it can give an insight into employees' well-being as well as workforce productivity and its effect on business operations. Literature research reports an increase in illness-related absenteeism throughout the decade and the major costs associated with it to employers (Wolvetang et al., 2022). The study presents a direct influence of sick leaves on cost management in businesses and indicates the average costs for stress-related absenteeism to build up to 799 million Euros annually.

The top right table in *Figure 9* shows the results from the employee well-being index by department. It can be seen that the "shipping and receiving" and "facilities and maintenance" departments scored the worst. This score can also be indicative of potential issues with the work environment contributing to higher employee dissatisfaction and should be investigated for further improvements in the business environment. As another option, the company can decrease extrinsic motivation for employees to take more sick hour leaves than necessary by reducing sick hour compensation. Although the company should be careful with ethical limitations to this, studies do show a decrease in the number of sick leaves when compensation for it is reduced (Wolvetang et al., 2022).

Internal Business Processes Perspective

Order Fulfillment Cycle Time

To analyze the efficiency of the organization's internal business processes, the average amount of time it takes for the vendor to fulfil its orders before the due date was investigated (Order Fulfillment Cycle Time). On average, it was calculated that orders from suppliers typically take around 2.4 days to ship their products. Furthermore, the average shipping date per product was computed in *Figure 3*. It can be seen that product number 744 was shipped significantly earlier compared to other products. Further investigation shows that the products that were shipped comparatively higher than the average (product numbers 744, 710, 811, 943, 911, 942, 927, 946) were all products associated with mountain bike parts. It can be assumed that the reason for faster shipping can be based on its popularity and its perceived value from consumers. However, further investigation is required for specified implications that the company can take due to the possibility that the shipping process can depend highly on the employees' skills, which is beyond the scope of the currently given dataset. *Figure 10* illustrates the average shipping time before the due date per region. It can be inferred that the average time does not differ significantly depending on the region a store is located. European regions take slightly longer to ship their products, which can be explained by cultural differences.

Data shows that the company uses a make-to-stock technique where a business has product inventory readily made available. This assumption is made because there are columns associated with stock levels and reorder points in the dataset. This is a good inception for the company, as research suggests that customer satisfaction and reliability can be improved with advanced order fulfilment techniques implemented in which products are constantly available through inventory (Okongwu, 2011). To improve the operation process, stock reorder points should be revised according to consumer consumption rate. Alternatively, a system can be placed in which the order is shipped immediately after it is placed to maximize efficiency if it depends on employee response time.

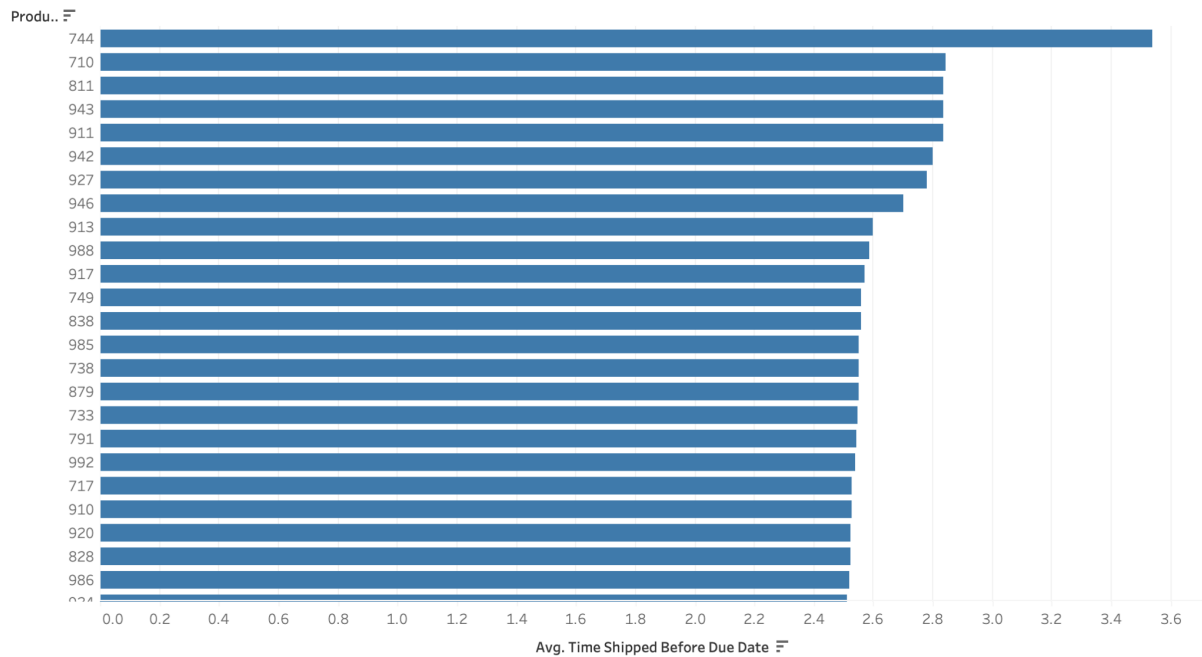


Figure 3. Products and Shipping Time

Vendor Credit Rating Index

Finally, the internal business operations inspected the difference between the expected and actual vendor credit ratings (vendor credit rating index). The credit score ranged from one to five. In this situation, the team assessed that a higher credit score equates to better credit. This assessment was established based on the Fair Isaac Corporation (FICO) score where a higher score also indicated a better financial position (Hayes, 2023). Generally, studies show that companies face lower returns when interacting with other organizations with a lower credit rating (Bissoondoyal-Bheenick & Brooks, 2015). Additionally, organizations that experience a downgrade in credit rating also induce greater future risk in returns.

Figure 4 shows the vendors' credit rating while the black vertical line indicates the company's expected credit rating. The company's standard expected credit rating is one, which is the lowest rating there is. It is recommended that the company increase its vendor credit criteria to reduce future financial risks. It can also be seen that Victory Bikes and Merit Bikes exceed the company's expectations significantly in *Figure 4*. Thus, the company should continue or expand its business operations with these vendors. For future improvements, the company can also record any downgrades of credit rating in vendors to prevent unnecessary risk-taking activities. Moreover, we recognize that the expected credit rating may be considered as a redundant variable if the overall standard was 1, thus visualization was

improved by removing the lines in *Figure 4* and only displaying the actual credit rating scores of vendors (see *Figure 10*).

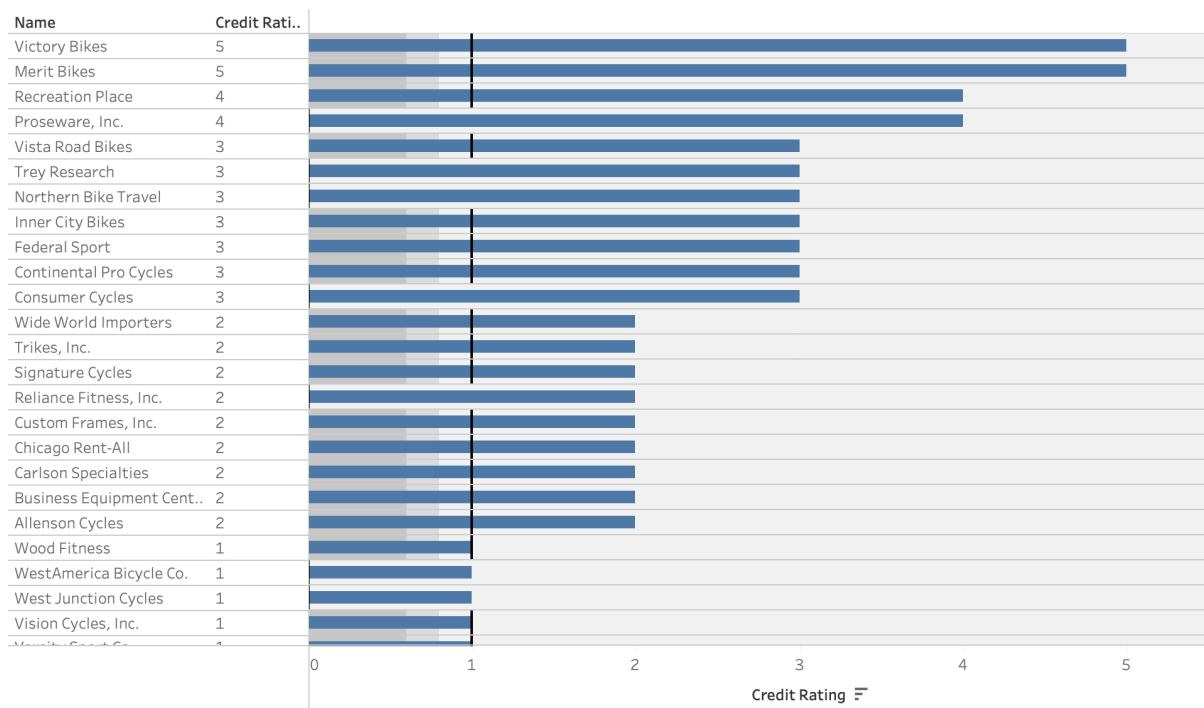


Figure 4. Vendors' Credit Rating

The Data Warehouse Design

A data warehouse is a large, centralised data storage used to support business intelligence and reporting. It is designed to store and manage huge volumes of structured and sometimes unstructured data from various sources within an organization. The primary purpose of a data warehouse is to provide a unified and consistent representation of an organization's data, allowing decision-makers to analyze and derive information to make more informed business decisions.

The data warehouse consists of a fact table surrounded by dimension tables. The fact table includes quantified data about the measurements (facts) and a set of all foreign keys from dimension tables. Dimension table has a primary key and a set of descriptive attributes about the measurements.

Future data warehouse design should have reflected all the KPIs that we constructed earlier. For that reason, we excluded data from the “salesdeveloperdb” database that was useless or did not represent metrics needed for the KPIs. The final version of the data

warehouse, shown in *Figure 5*, consisted of 10 tables in total. 1 fact table connected with 8 dimension tables and 1 separate table for measurements about the employees. Our fact table includes all foreign keys and UnitPrice, OrderQty, UnitPriceDiscount, StandardCost, and a Credit Rating as fact metrics. Dimension tables consist of primary keys and descriptive attributes about, for instance, order location, product details, shipment dates, vendor info etc. We created a separate table called “employee_table” which is disjointed from the rest schema because it does not contain any key connections with other tables. However, measurements in that table are crucial for our KPI - “Learning and Growth Perspective”.

Our data warehouse resembles a star schema with a “dimension_product” table having another connection with “dimension_productsubcategory” and “dimension_productcategory”. According to the requirements, All dimension tables have a “one-to-many” relationship with the fact table.

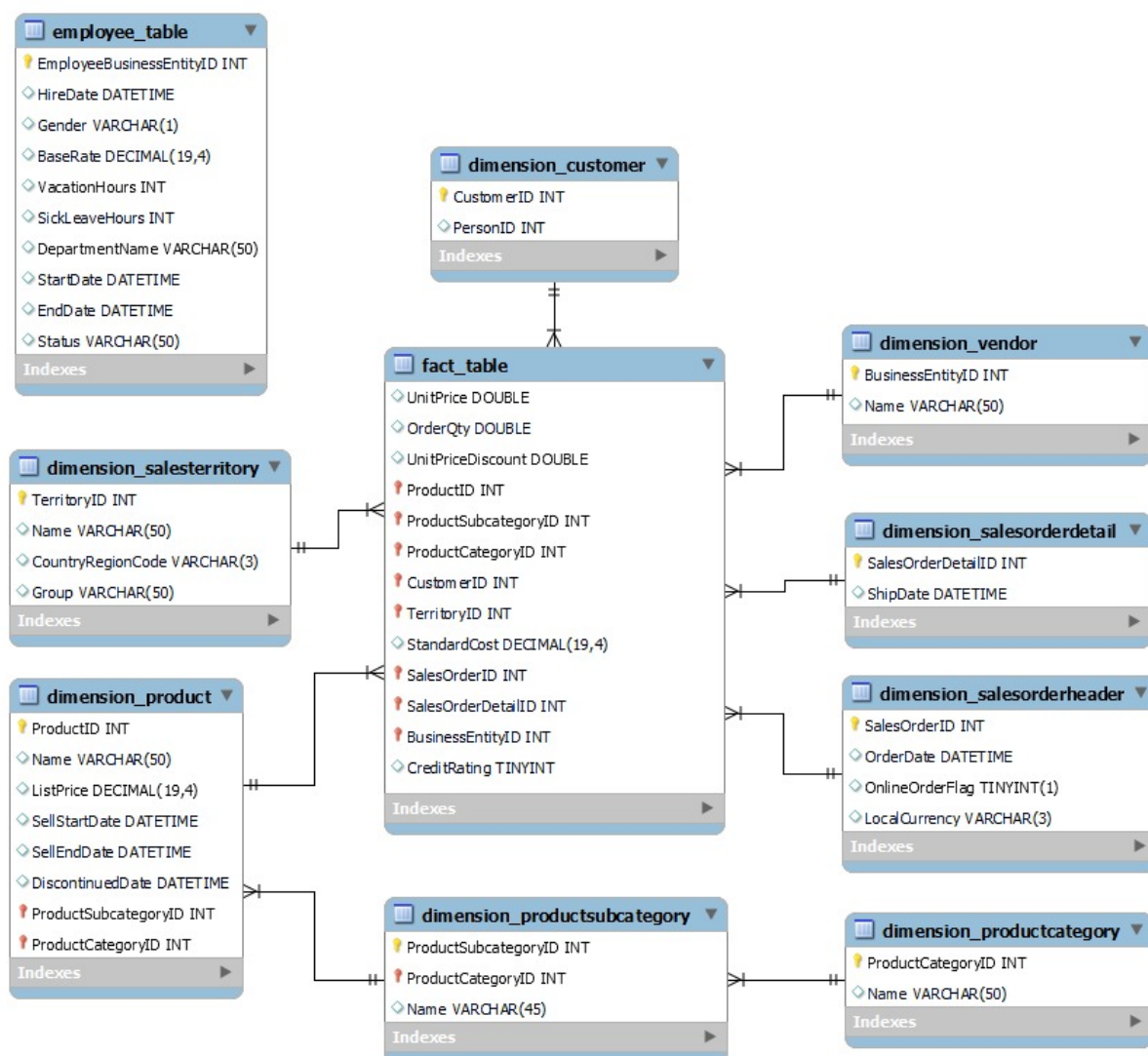


Figure 5. Data Warehouse

Tableau Dashboard

In total, we compiled 4 dashboards, one for each perspective.

Financial Perspective

Financial perspective

Gross Profit Margin and
Average Transaction Value per Region

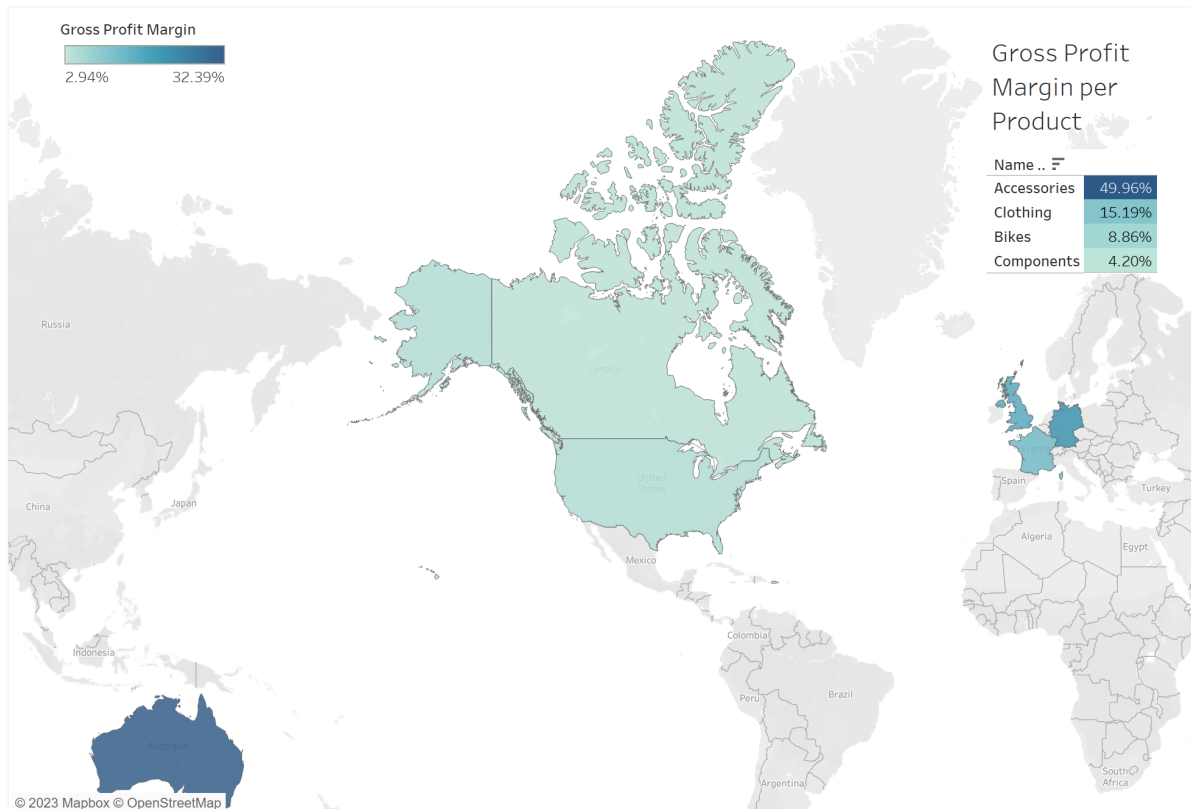


Figure 6. Financial Perspective

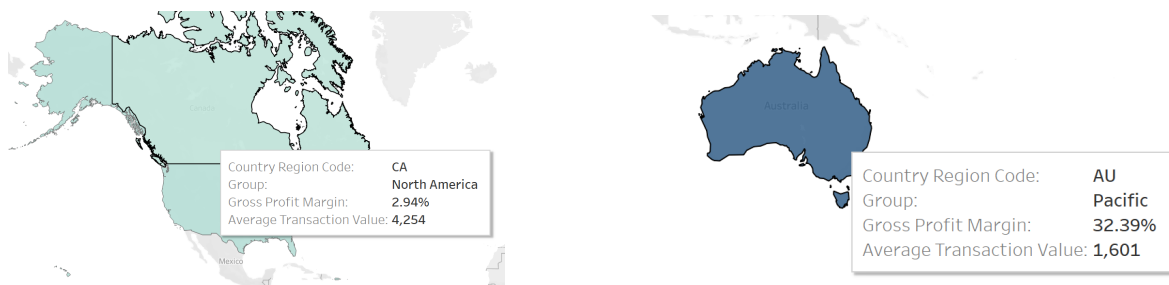


Figure 7. Financial Perspective (labels)

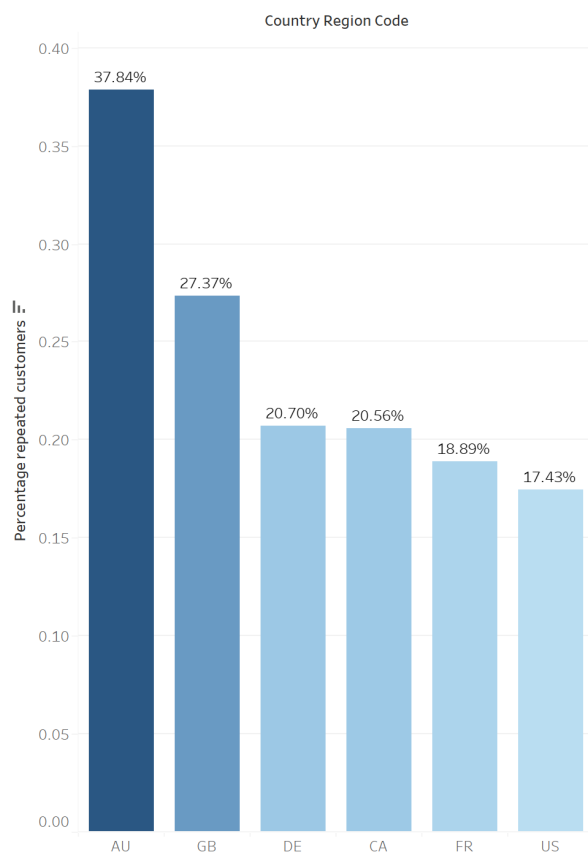
This dashboard summarizes the gross profit margin per region and per product and average transaction value per region. The colour on the map shows the difference in the gross

profit margin, therefore the lighter the colour of the region, the smaller the gross profit margin. Consequently, the North American region has the smallest gross profit margin, while Australia has the biggest gross profit margin. The gross profit margin per product is summarized in a table in the right top corner.

Customer Perspective

Customer Perspective

Customer Retention Rate



Orders Per day

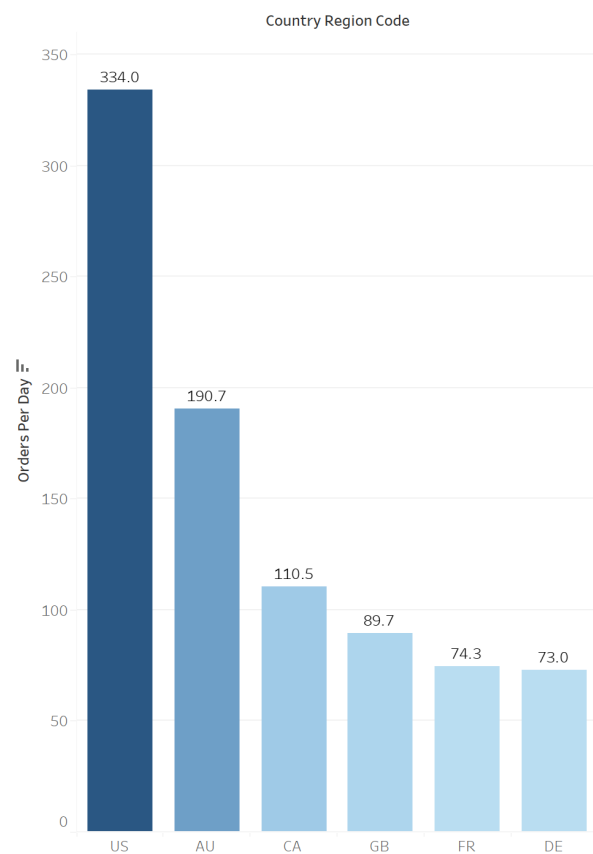


Figure 8. Financial Perspective

The dashboard summarizes customer retention rate and orders per day. Both measures are pictured using bar charts, where the x-axis is the measure and the y-axis is the code for the region.

Learning and Growth

Learning and Growth Perspective

Gender Pay Equity

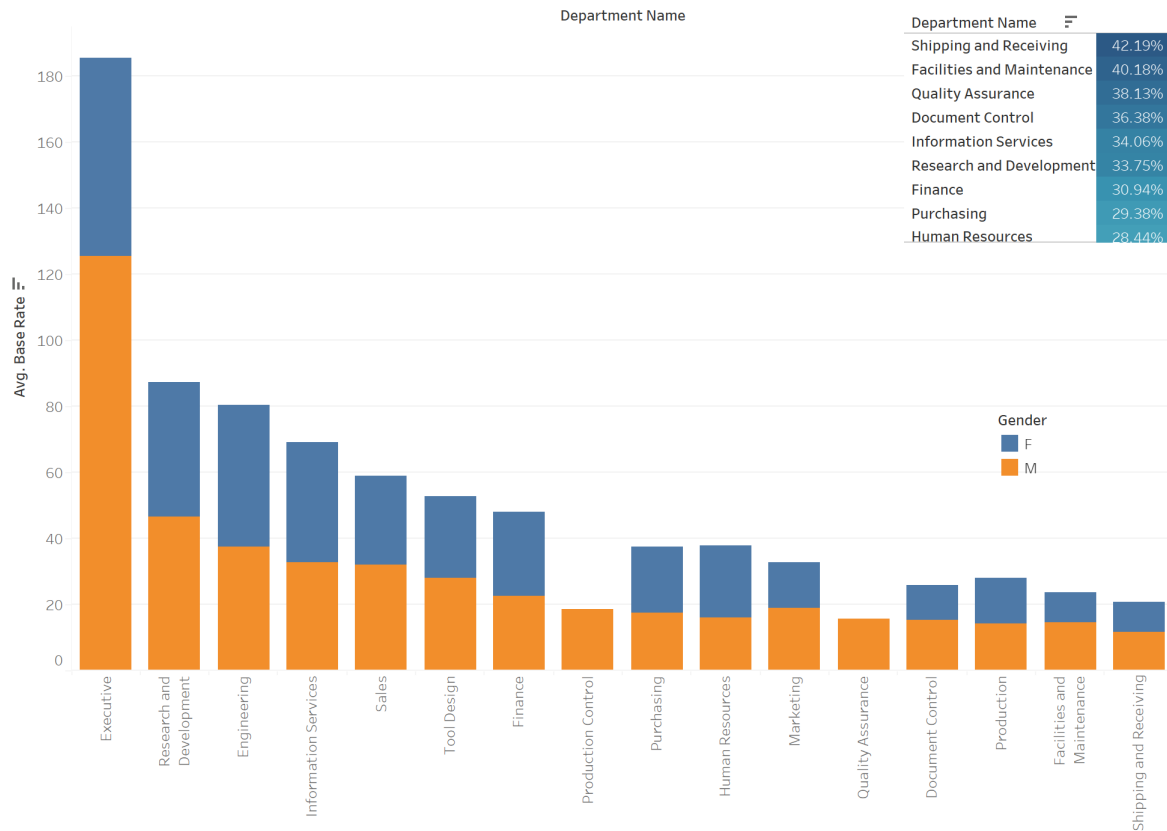


Figure 9. Learning and Growth Perspective

The dashboard summarizes Gender Pay Equity and Employee Well-Being. Gender Pay equity is summarized with a stacked bar chart, where the blue part shows the average base rate for females and the blue part for males. The y-axis corresponds to the base rate, while the y-axis shows the departments. Employee Well-Being is summarized in a table in the top right corner.

Internal Business Process Perspective

Internal Business Process Perspective
Order Fulfillment Cycle Time

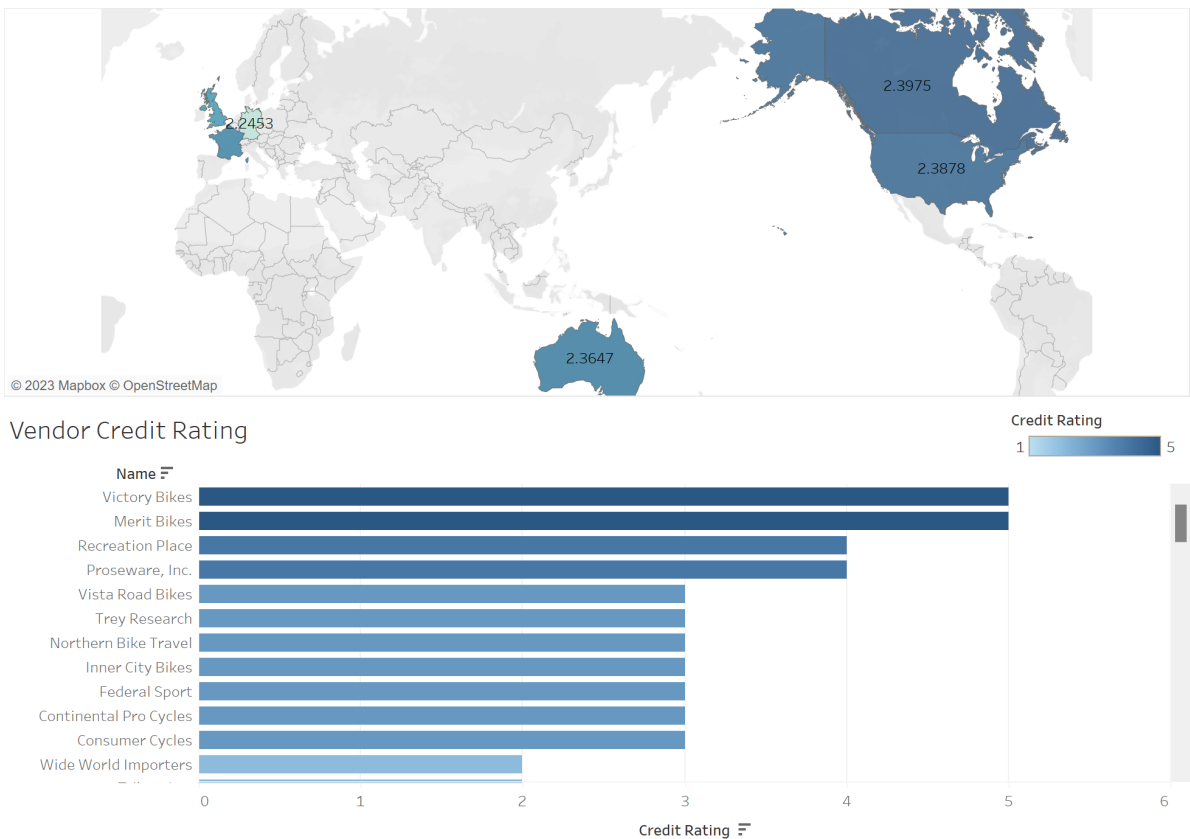


Figure 10. Internal Business Process Perspective

The dashboard summarizes Order Fulfillment Cycle Time and Vendor Credit Rating. The order fulfilment cycle is summarized with a map, where the regions are labelled with the average amount of days it takes for the order to be fully processed and delivered before the due date. A larger score indicates better efficiency of the order fulfilment cycle. This is because a larger score means the order is fulfilled earlier than the due date. The vendor credit rating is summarized with a reversed bar chart, where the vendors are listed according to their credit rating in descending order.

Implications

One of the most significant fixes that our dataset can undertake is the improvement of the 'employee' table. Currently, the 'employee' table is isolated without any connections to other tables via a foreign key. This is a problem with our database integrity. One way to improve it is to connect 'employee' to other tables, preferably to 'orderdetail' table. That way we can trace which orders were processed by which employees. Another problem with the dataset is the multivalued 'StartDate' attribute. For the sake of data integrity and consistency, it should be normalized in the future.

In conclusion, we undertook an analysis in order to discover hidden economic patterns and facts throughout our dataset. Financially speaking, bike accessories are the most profitable products the company sells. However, our report shows the importance of specialized treatment in each region where the company operates. One-for-all strategy is too robust and impractical to keep up with the sales worldwide.

The financial KPIs give valuable insight into customer values for the company's products as well as each regional branch's performance based on competition. When assessing the KPIs from the customer perspective, it was made clear that the regions in the U.S. and Australia did significantly better than European stores. It is recommended that European regional branches take into consideration Australia's loyalty and customer-retaining programs and focus on creating better marketing strategies tailored to the market culture.

The realization of learning and growth KPI showed us the differences in gender pay. The company should focus its attention on the fairness of pay between its workers and retaining quality human capital. We identified that the dataset company has a higher ratio of male to female workers. The possibility of gender bias should be taken into consideration in future recruitment processes. In addition, compensation for sick leaves should also be reviewed to decrease extrinsic motivation for unnecessary absenteeism.

After accessing the results of the implementation of internal business operations KPI, we got an insight into order fulfilment cycles and vendor credit ratings. In order to increase customer satisfaction the company should focus on keeping a larger inventory of popular products like mountain bike parts. The company should also prioritize business with organizations with a better credit rating and track any downgrades of credit ratings in the future.

All in all, this report's goal was to dive into an informational database to get deeper insights. To help with the goal we came up with a broad spectrum of KPIs. The

implementation of the recommendations that we have written above would boost the company's overall performance, sustainable growth, customer satisfaction, and comfortable workplace for its employees.

Reference list

- Ahmad, R. & Buttle, F. (2002). Customer retention management: a reflection of theory and practice. *Marketing Intelligence & Planning* (20)3, 49-161.
<https://doi.org/10.1108/02634500210428003>
- Bissoondoyal-Bheenick, E., & Brooks, R. (2015, November 30). *The credit risk–return puzzle: Impact of credit rating announcements in Australia and Japan*. ScienceDirect. <https://doi.org/10.1016/j.pacfin.2014.09.001>
- Hayes, A. (n.d.). *FICO Score*. investopedia.
<https://www.investopedia.com/terms/f/ficoscore.asp#toc-how-to-improve-your-fico-score>
- Hultin, M. (2003). *Mechanisms of Inequality. Unequal Access to Organizational Power and the Gender Wage Gap*. academic.oup.com.
<https://academic.oup.com/esr/article-abstract/19/2/143/537107?redirectedFrom=fulltext>
- Uche Okongwu , Matthieu Luras matthieu.luras@mines-albi.fr , Lionel Dupont & V  rane Humez (2012) A decision support system for optimizing the order fulfillment process, *Production Planning & Control*, 23:8, 581-598, DOI: [10.1080/09537287.2011.566230](https://doi.org/10.1080/09537287.2011.566230)
- What is Average Transaction Value (ATV) and why is it important? (2023, Apr 16). Checkout. Retrieved from
<https://www.checkout.com/blog/what-is-average-transaction-value>
- What Is Retail Margin? (With Steps and Tips to Calculate It). (2022, June 25). *Indeed Editorial Team*. Retrieved from
<https://www.indeed.com/career-advice/career-development/what-is-retail-margin#:~:text=One%20of%20these%20essential%20metrics,ensure%20your%20business%20operates%20efficiently.>
- Wolvetang, S., van Dongen, J.M., Spekl  , E. *et al.* Sick Leave Due to Stress, What are the Costs for Dutch Employers?. *J Occup Rehabil* **32**, 764–772 (2022).
<https://doi.org/10.1007/s10926-022-10042-x>