Copenhagen Business School

MSc in Business Administration and E-Business

Business Data Processing and Business Intelligence

Lundbeck's declining Revenue and Gross Margins in Q1 for 2020

Vidyarani Achrekar: 132149

26th February 2021

Number of pages: 10 normal pages

Number of characters: 2886 words

1 Introduction

1.1 Data Mining in organization

According to Han, Kamber, and Pei (2012) Data mining is an essential process where intelligent methods are applied to extract data patterns and knowledge from large amounts of data. The data sources can include databases, data warehouses, the Web, other information repositories, or data that are streamed into the system dynamically (Nemati & Barko, 2001). As a general technology, data mining can be applied to any kind of data as long as the data are meaningful for a target application. Organizational Data Mining (ODM) is defined as leveraging data mining tools and technologies to enhance the decision-making process by transforming data into valuable and actionable knowledge to gain a competitive advantage (Nemati & Barko, 2001). ODM enables organizations to answer questions about the past (what has happened), the present (what is happening) and the future (what might happen)(Nemati & Barko, 2001). With the assistance of ODM, organizations can generate valuable knowledge from their data, which in turn enhances enterprise decisions. This decision-enhancing technology enables many advantages in various operations like faster product development, increased market share with quicker time to market, optimal supply chain management), marketing (higher profitability and increased customer loyalty through more effective marketing campaigns (Nemati & Barko, 2001). In this paper, the data of the case company (Lundbeck A/s) is utilized to identify sales and margin behavioral patterns and use those data insights to effectively visualize the data to optimize the decision of their revenue and margins. Please refer to figure 1 to see the steps involved in an organizational data mining process.

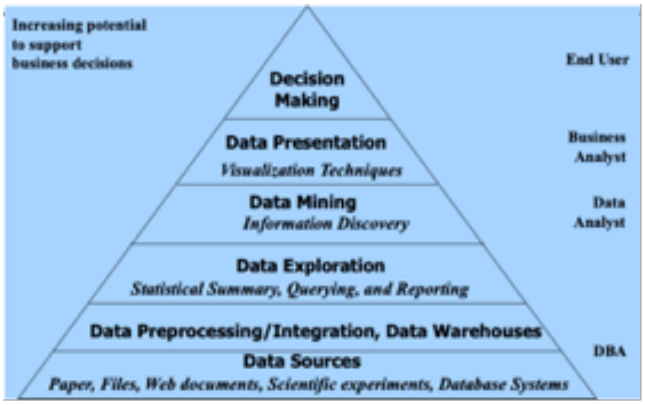


Figure 1: Hierarchical steps involved in data mining of Lundbeck's data

1.2 Case company

Lundbeck is a global pharmaceutical company highly committed to improving the quality of life of people living with brain diseases. For this purpose, Lundbeck is engaged in the research, development,

manufacturing, marketing and sale of pharmaceuticals across the world (Lundbeck.com). The company's products are targeted at the disease areas within psychiatry and neurology (Lundbeck.com).

Lundbeck has a wide range of products to treat various kinds of brain diseases and some of Lundbeck's main products are supporting to treat patients with depression, schizophrenia and Alzheimer's and Parkinson's diseases. Lundbeck markets a number of different pharmaceuticals for the treatment of brain diseases. The most recently launched products include Rexulti® (depression and schizophrenia), Brintellix® (depression), Abilify Maintena® (schizophrenia) and Vyepti® (migraine). Lundbeck has employees in more than 50 countries, and our products are registered in more than 100 countries (Lundbeck.com). Its production centres are located in Denmark, France and Italy and their research centres are located in Denmark and the US.

As the pandemic (Covid 19) hit the entire world in 2020, Lundbeck began facing challenges from the first quarter (Q1) 2020 which was evidenced by revenue and gross margins starting to trail behind budget during this period. As Lundbeck started to analyze the causes for underperformance, its top priorities during the pandemic were to preserve the health and safety of their employees and ensure business continuity in safely supplying the medicines to millions of patients worldwide.

1.3 Problem Formulation and Research Question

Lundbeck's revenue and gross margins have been behind budget during the first quarter of 2020. The Sales index was 90 vs. budget, whereas gross margins were at index 97. This was explained by an unfavorable development in the Med Aid product group.

The sales and margins developed in the wrong direction during the second quarter. Management has asked you to analyze revenue and margin performance for year-to-date May of 2020 and asked to present the findings.

Which countries/region and product line should Lundbeck's management focus on to identify the underlying causes for sale and margins under performance?

1.4 Delimitations

Due to data privacy concerns, all data provided in the dataset is fabricated for purely academic research and not representative of Lundbeck's actual performance, products or regions. Hence, the findings from this research are limited to academic purpose and have no relevance for real life business decision making.

2 Methodology

2.1 Dataset Description

The *data source* received was in the form of an excel sheet consisting of three data tables: 1. Fact table with Sales and Cost of goods sold (COGS) data (referred as Transactional table), 2. Profit Center overview (Master data for Profit center), 3. Material overview (Master data for Materials). The transactional table contains 11708 rows and 6 columns, the *Profit Center overview* table contains 71 rows and 5 columns and lastly, *the Material overview table* contains 477 rows and 3 columns. The process of data extraction, data processing, and the data visualization is depicted in the process-flowchart below.

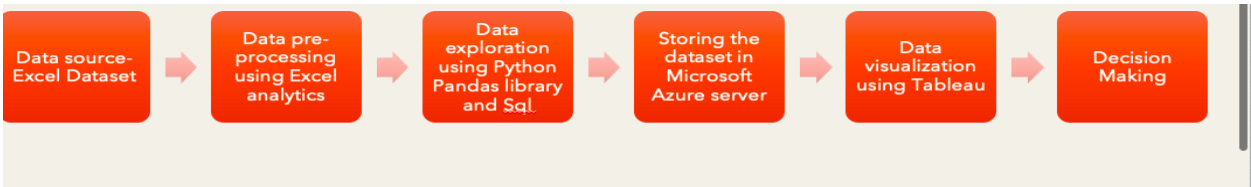


Figure 2: Data Process flow diagram

2.2 Data Collection: Methods and Tools

As mentioned above, the data was collected in the form of excel files that were used as the ***Data source*** to perform the required data analysis. Further, to assure that the Excel files can be used as the relevant data source, the data structure in the file was validated to make sure that the data was in the correct list format. This meant that the data structure consisted of one or more columns with header names. The analytical tools that were applied in this paper are ***Excel analytics*** to ***clean the data*** and ***to carry out the data normalization***, ***MS SQL server*** to store the data on the server, ***MS SQL management studio*** to perform ***data modelling*** by implementing the relational database diagram, ***Python, Jupyter Notebook*** to carry out basic data ***descriptive analysis*** and ***Tableau*** for ***data visualization.***

Further, to conduct the data analysis, a data ***process flow diagram*** (refer figure 2) was incorporated for graphical representation of data extraction, data processing, and data visualization. The ***Process Flow Diagram (PFDs)*** is a graphical way of describing a process and its constituent tasks, and the sequence presented in the flow diagram helps with brainstorming and communication of the processes under data analysis (Bijan Elahi, 2018).

2.3 Dataset Pre-processing and data analysis

The data analysis and visualization provided in this study is based on Lundbeck's revenue and cost data for the first five months of 2020 (i.e., Jan-May 2020). The initial revenue and cost data were not in a structured format to enable a meaningful data analysis. Hence, the data had to be subjected to a cleansing process,

which included the following key elements: 1)Identification and removal of duplicate data (mainly the duplications in master data). 2)Restructuring the data in a format that is conducive for exporting to a relational database system. 3)Mapping master data with transactional data for meaningful analysis.

In this paper, the steps executed for the data cleaning process with the initial three datasets: A)**Transactional dataset** (which contained revenue and cost data per month for each profit center and material). B) **Master data for Profit center.** C) **Master data for Materials**. To begin with, the Master data relating to Profit center and Material were validated for duplication and the duplicate data was removed after proper evaluation. This step was necessary to ensure that the master dataset only includes ***unique data keys*** and thus does not distort the results due to duplication. ***The Duplicate data was identified using conditional formatting functionality in MS Excel.*** The next step was to rearrange the transactional data (revenue and cost) into a meaningful structure to aid proper analysis and also enable seamless data transfer. In the initial dataset, some key data elements of the transaction data set (viz, revenue and cost data fields) were arranged as rows. This made analysis very difficult and also presented significant challenges in creating an effective data visualization. To remedy this challenge, a 'Power Pivot' function in MS Excel was put into application to separate unique transactional key data elements and rearrange them into columns. This helped in restructuring the data in an easily readable format, where the data relating to Month, Profit Centre, Actual Cost, Budgeted Cost, Actual Revenue, and Budgeted Revenue were transposed as columns. The data restructuring not only created a meaningful relationship between key elements (revenue and cost) but also provided an opportunity to simplify calculation of Actual and Budgeted Gross Profit (using a simple mathematical operation of subtracting cost from revenue). Further, the primary key column named ***DataField*** was introduced in the Transactional dataset to uniquely identify the table record. Third step involved mapping of the profit center and material keys provided in the transactional dataset with the Profit Centre and Material description in the Master data. The objective was to provide the description of each Profit Centre and Material key used in the transaction data. The initial transactional dataset contained only the unique numerical value for identifying profit center and material, without any description attached to those keys. The description and other details corresponding to each numerical value however was available in the respective master datasets. VLOOKUP functionality in MS Excel was used to populate details from master dataset into transactional dataset (using unique numerical value as the key identifier). Final outcome of data normalization process:

The normalization/cleaning process helped to clean the dataset and create a meaningful relational data structure ready for exporting to a relational database management system or to any other data analytical tool. As a double check, we imported the dataset in Python, Pandas data frame to check for any null values and remove null values using fillna method = mean for each column (refer appendix 8). And conducted descriptive analysis to validate the clean data in the dataset. This data frame was later transferred to ***MS***

SQL management studio to perform ***data modelling*** by implementing the relational database diagram. Here the various attributes were creates using "create table" function and assigning the respected data types (refer Appendix8). The set of attributes that can distinguish an entity in a table were grouped/assigned as a Primary key and two Foreign keys, the data from Excel File was imported into the existing tables in SQL Server. Later the stored data was imported in tableau for data visualization.

3 Results

After processing the data in Python and SQL, the clean dataset is ready for visualization. Tableau is an extremely powerful visualization tool that has been exercised to create dashboards to graphically represent the finding of the analysis. In this chapter, all the findings from the descriptive data analysis are outlined and examined in detail.

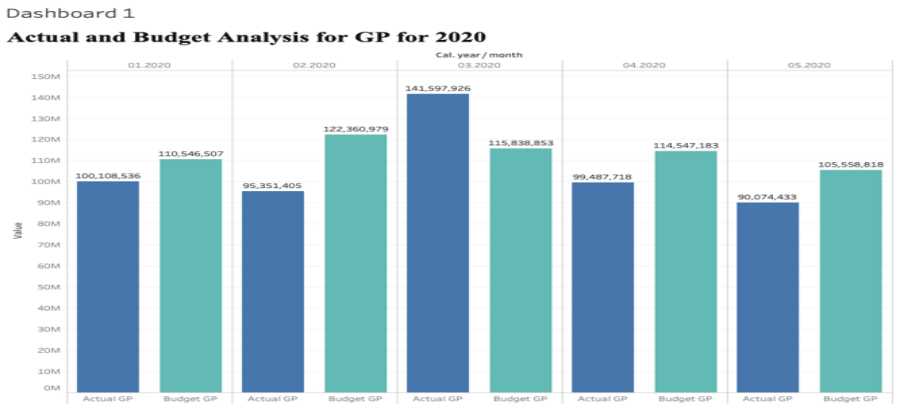


Figure 3: Actual and Budget Analysis for GP

The above bar graph (figure 3) there is a favorable variance in the Actual GP(Gross Profit) in comparison to the Budgeted GP in the month of March 2020 of the. On the other hand, the rest of the months the Actual GP is behind the budget gross margin indicating negative profit variance.

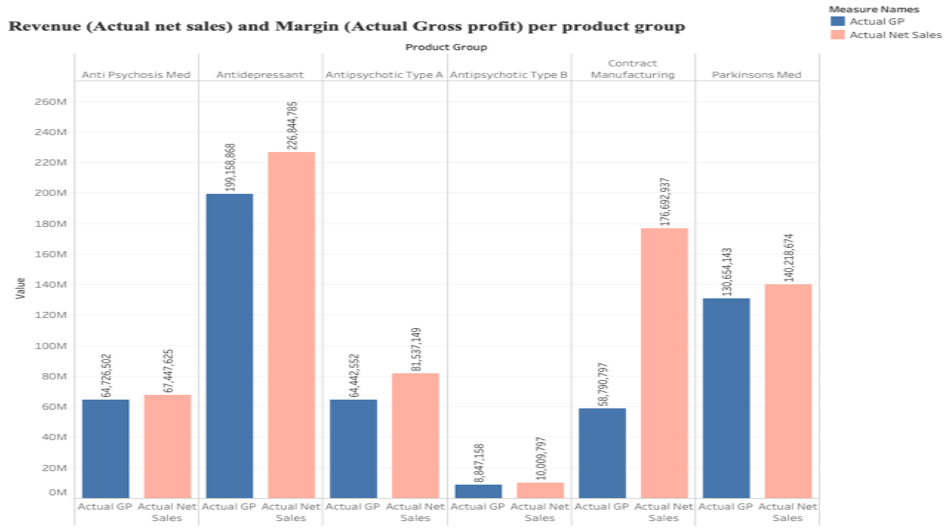


Figure 4: Revenue and Margins per product group

In figure 4, the analysis of the sales and profit margin data per product group globally reveals that the three Med aid products Anti psychosis, Antipsychotic type A and type B are underperforming as shown in the graph above. On the contrary Antidepressant is the highest revenue generating product group for Lundbeck, in my opinion this is also concerning as it reflects on the deteriorating mental health.

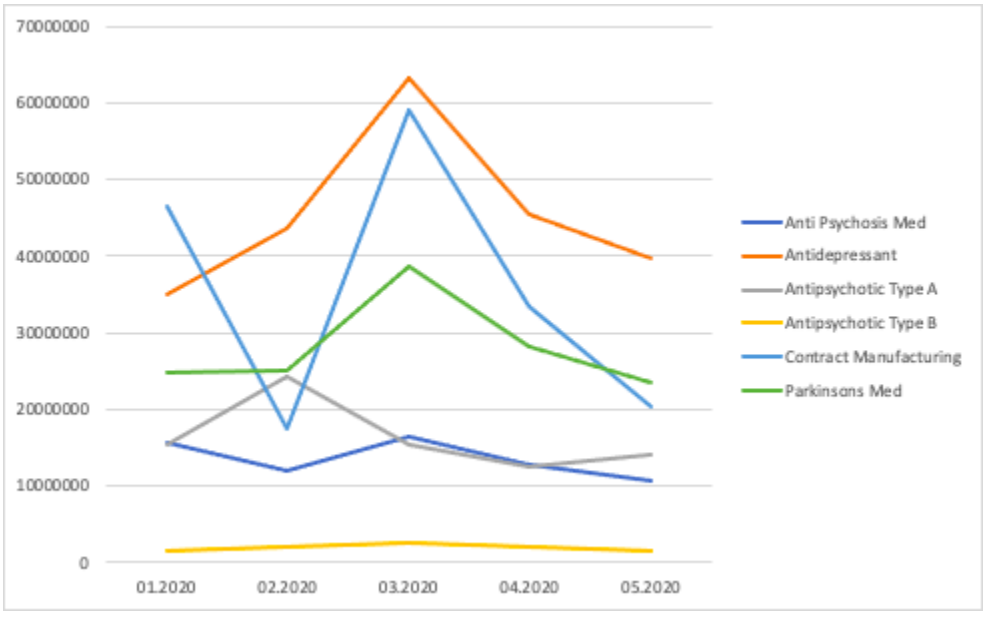


Figure 5 Sales trend for each product group per calendar month

From the above data (figure 5) it is evident that all the product groups demonstrate a decreasing trend in sales except for March 2020.

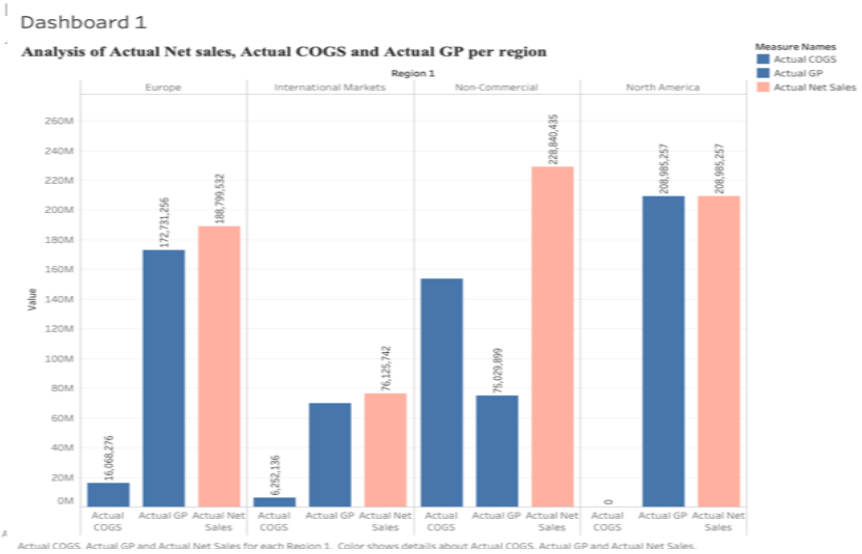


Figure 6: Analysis of Actual GP per region

In the figure 6, based on the above graph it is indicated that the International markets are making less profit and revenue and it's worth noting that North America is making higher gross margins and net sales as there is no cost attached to the goods sold. This could be because North America receives finished products and has no cost of production or the North American region is purely engaged in services. On the other hand, the Non-Commercial region stands on the top with highest revenue, this may be because some Non- commercial region is engaged in production of drugs and supply them to all other regions for small margin markup.

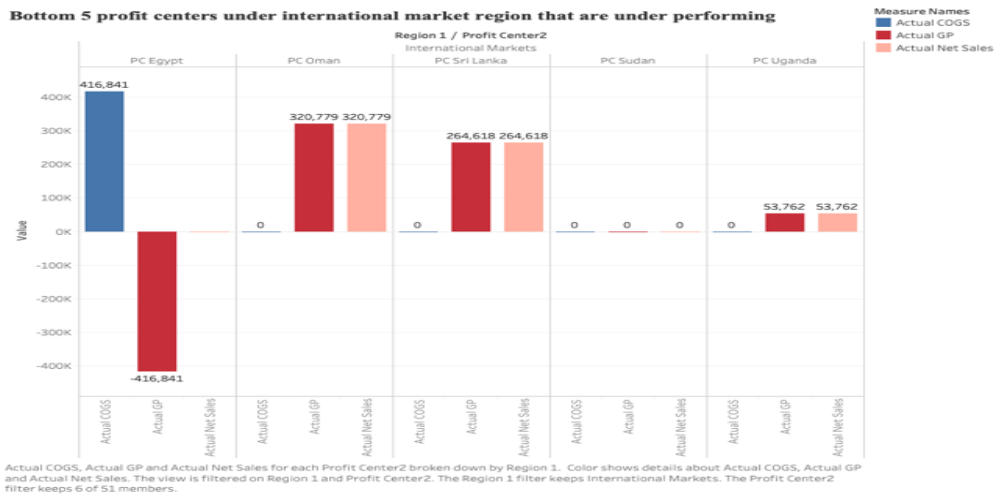


Figure 7: Bottom 5 profit centers under international market region that are under performing

There are 20 countries under the international market region, out of which the bottom 5 underperforming countries are Egypt, Sudan, Sri Lanka, Oman and Uganda as shown in the above graph (figure 7). The graph indicates that Egypt is the most underperforming country as it has a high cost of goods sold without any sales volume (resulting in negative gross profit).

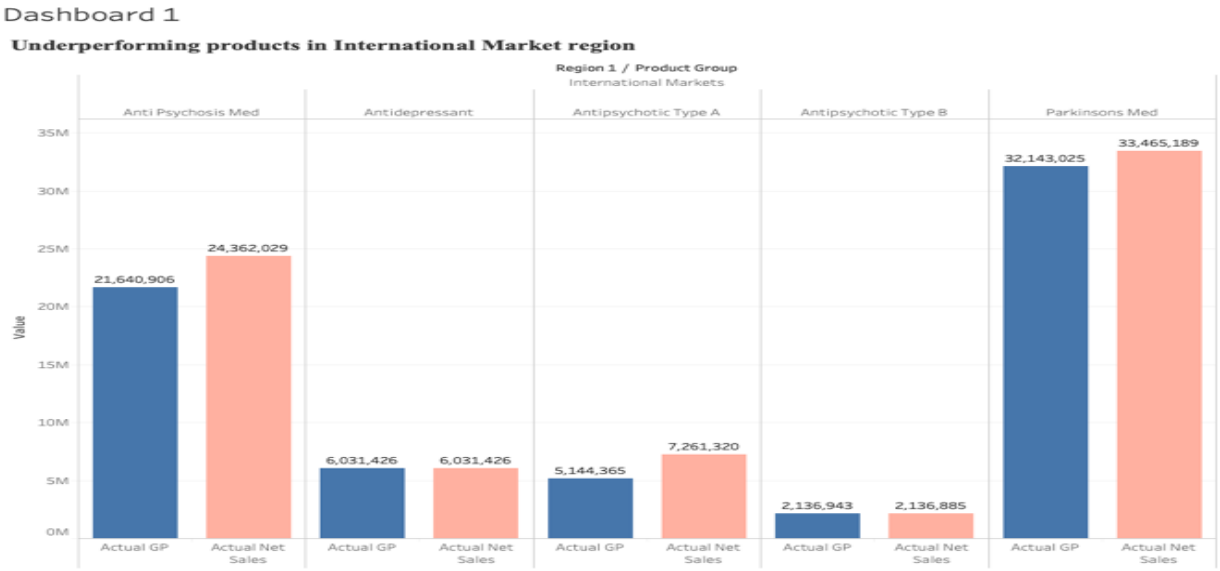


Figure 8: Products in international market region that are underperforming

In figure 8, further investigation into International Markets data revealed that the three products (Antidepressant, Antipsychotic type A and type B) were underperforming as shown in the above graph.

4 Discussion

Based on the analysis of Lundbeck's financial data from January to May 2020 confirms that the revenue and the gross margins have been behind the budget except for March 2020. Further drilling into the sales and profit margins for regions, Lundbeck should focus on the international market region as it is underperforming. In specific, Lundbeck should conduct an in-depth analysis for the countries (Egypt, Sudan, Sri Lanka, Oman and Uganda) to find the underlying reasons for their underperformance mainly relating to the three product groups with lower revenue and gross margins. Based on the results of this analysis, Lundbeck should consider developing a road map to boost their sales and increase the performance by reassessing their market strategy, sourcing strategy and customer needs and also do an in-depth competitor analysis to understand their competitors pricing strategy. Further Lundbeck should consider conducting a detailed and well-researched analysis that indicates revenue for all activities in a company. This can range from sales (products and services), costs, income, and other variables

It's normal for some products to outstrip others in sales and demand. It's not uncommon for a business to consolidate on areas that yield more revenue and profits for the business. From such analysis, Lundbeck can gain insight on how to improve the areas where products and services have failed or are deficient in. The informed decisions could include whether to increase company's revenue or increase margins with minimal expenditure. Further, it is an assumption that as the pandemic (Covid 19) hit the entire world in 2020, Lundbeck began facing challenges during the first quarter (Q1) 2020 could be one of the reasons for revenue and gross margins trailing behind budget during the same period.

5 Conclusion

In this paper the dataset provided by Lundbeck is fabricated for purely academic research and not representative of Lundbeck's actual performance. This was in adherence with the data privacy compliance. Various analytical tools were deployed to conduct the descriptive data analysis. Tools like excel analytics was used to normalize/clean and transform the data, SQL management studio for data modelling and implementing relational database and Python to replace the null values with a value of the particular column and SQL server to store the data in the server and lastly, Tableau for data visualization. The analysis of the sales and profit margin data per product group globally reveals that the three Med aid products Anti-psychosis, Antipsychotic type A and type B are underperforming. ***On the contrary, Antidepressant is the highest revenue generating product group for Lundbeck. In my opinion this is also concerning as it reflects on the deteriorating mental health.***

This study provided an in-depth insight and understanding of how various analytical tools can be used to process and transform raw data in uncovering patterns and extracting meaningful insights. Also, this provides a reflection on how organizations can harness the Data mining process to turn the raw data into useful data.

By utilizing robust data analytical tools to find patterns in large data sets, pharma companies like Lundbeck can learn more about their customers personalized needs to develop more efficient innovative products, deploy competitive business strategies, boost sales, reduce costs, identify patterns from historical patient data to improving the quality of drugs and, sourcing methods.

Further in-depth analysis and drilling of data is possible to arrive at more concrete trends for sales and margin per product, per material, per country etc. This has not been done in this study due to limitation of space to document the study result.

Link for the raw dataset

https://docs.google.com/spreadsheets/d/1tuVvewNyeCYlQC6GjdKW4Culg1bt2qVQ/edit#gid=1148538098

Link of the dataset during transformation

https://docs.google.com/spreadsheets/d/1DU2SRf_7ILDzLkH3uPoVFaXy4JHiXmN1/edit#gid=286788787

6 References

Bijan Elahi, Chapter 12 - Risk Analysis Techniques, Editor(s): Bijan Elahi, *Safety Risk Management for Medical Devices,* Academic Press, 2018, Pages 67-120, ISBN 9780128130988,

https://www.sciencedirect.com/topics/engineering/process-flow-diagram

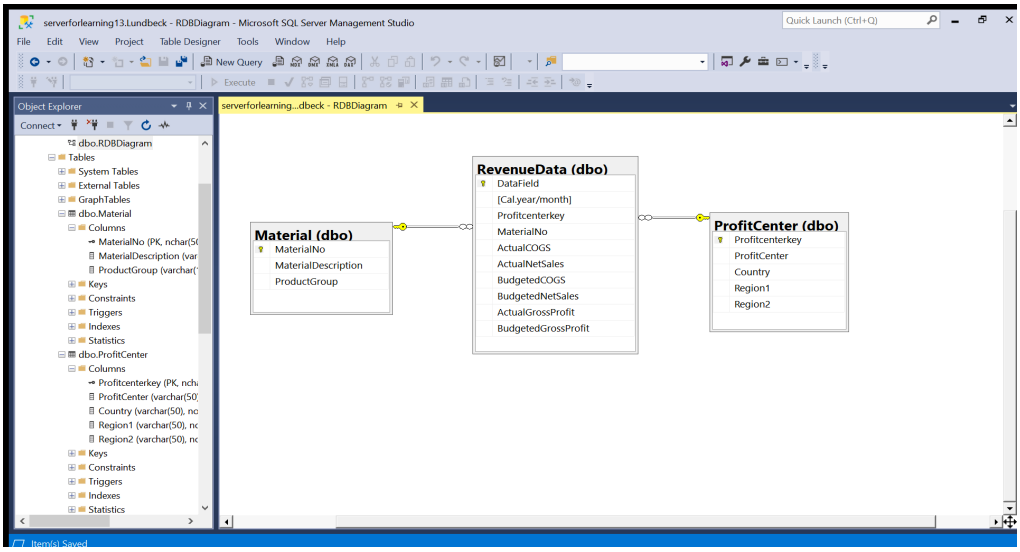
Jiawei Han, Micheline Kamber, and Jian Pei (2012). *Data-Mining Concepts and Techniques, 3rd edition*. Copyright © 2011 Elsevier Inc. All rights reserved, 2012. ISBN 978-0-12-381479-1

https://www.lundbeck.com/global

Nemati, H. R., & Barko, C. D. (2001). Issues in organizational data mining: A survey of current practices. Journal of Data Warehousing, (1), 25-36

7 Appendices

Appendix 1



Relational database diagram after normalization /restructuring the row and columns

Appendix 2

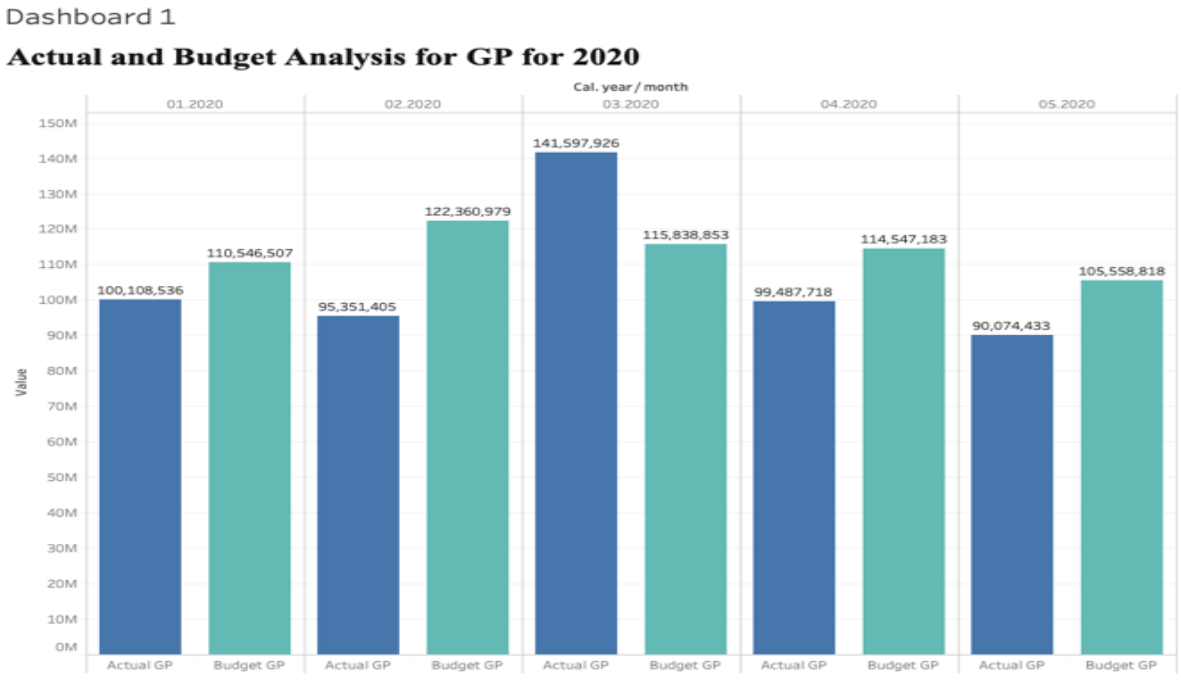


Figure 3: Actual and Budget Analysis for GP

Appendix 3

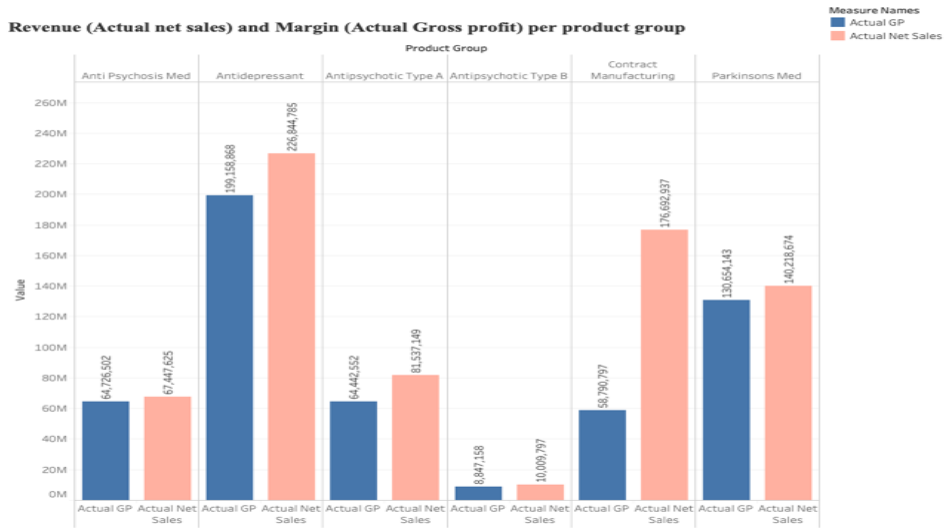


Figure 4: Revenue and Margins per product group

Appendix 4

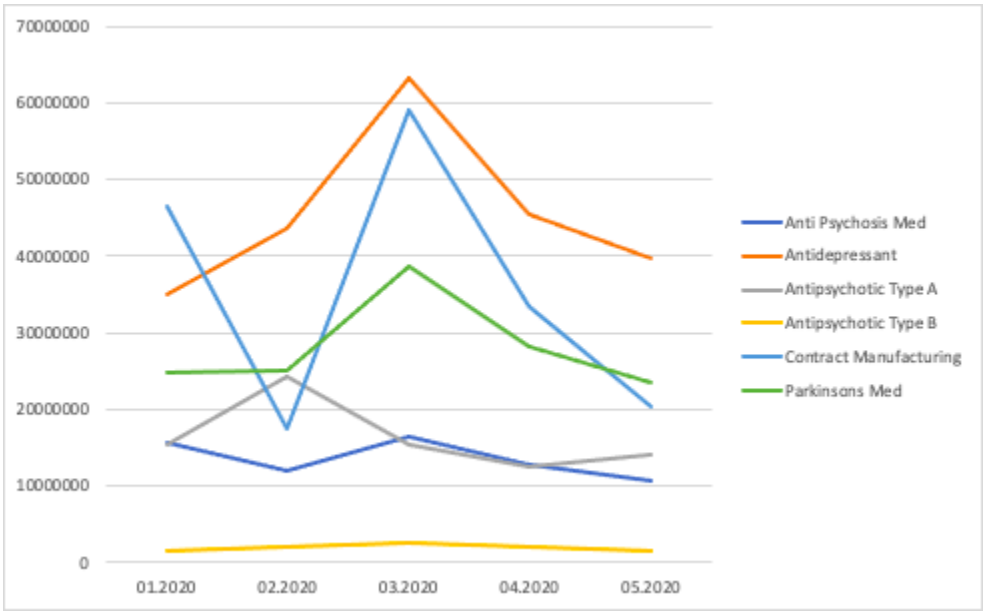


Figure 5 Sales trend for each product group per calendar month

Appendix 5

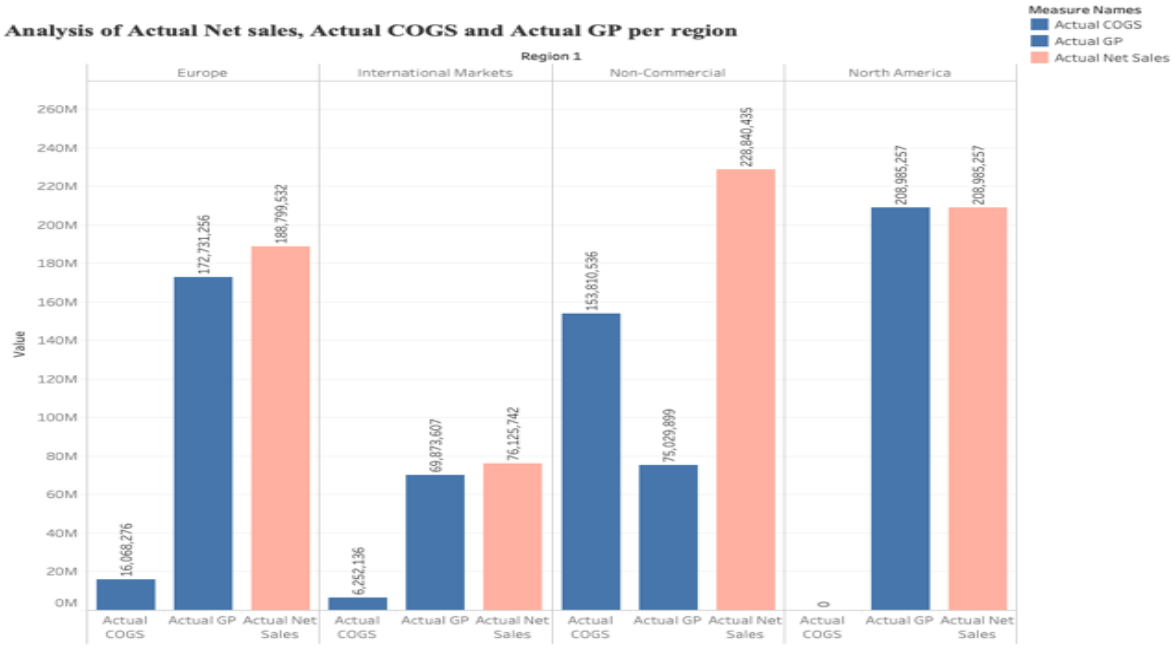


Figure 6: Analysis of Actual GP per region

Appendix 6

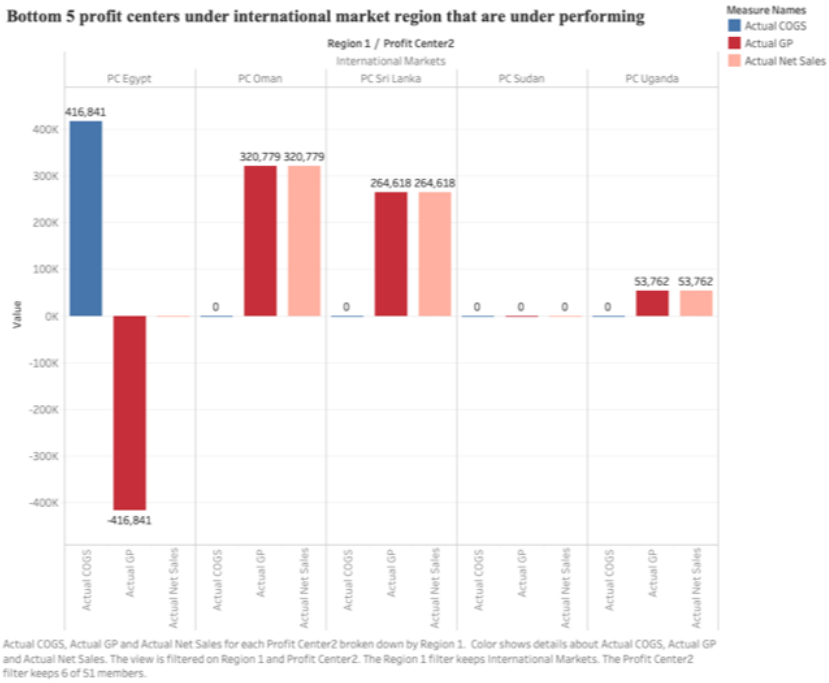


Figure 7: Bottom 5 profit centers under international market region that are under performing

Appendix 7

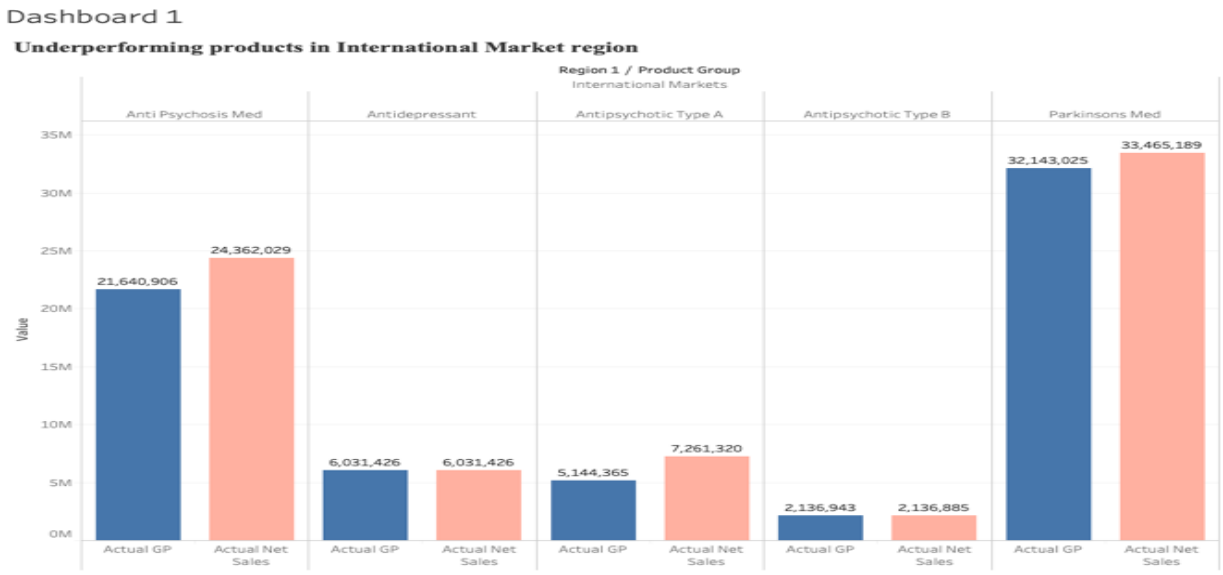


Figure 8: Products in international market region that are underperforming

Appendix 8