National College of Ireland

Project Submission Sheet – 2021/2022

Student Name:	DEEP MAYUR THAKAR, JUAN PABLO ANDRES PALMA BUSTOS, PRASHANT DIGAMBAR WAGHELA, TANMAY SHRIKANT PARANJAPE, VISHAL GAJANAN PATWARDHAN
Student ID:	x20208111, x20225776, x20207786, x20101015, x18190839
Programme:	Msc. Data Analytics **Year:** 2021-22
Module:	Business Intelligence and Business Analytics
Lecturer:	Prof. Vikas Sahni
Submission Due Date:	14th December 2021
Project Title:	BIBA Project Specification Report
Word Count:	2583

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the references section. Students are encouraged to use the Harvard Referencing Standard supplied by the Library. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action. Students may be required to undergo a viva (oral examination) if there is suspicion about the validity of their submitted work.

Signature:	DEEP MAYUR THAKAR
Date:	14th December 2021

PLEASE READ THE FOLLOWING INSTRUCTIONS:

1. Please attach a completed copy of this sheet to each project (including multiple copies).
2. Projects should be submitted to your Programme Coordinator.
3. **You must ensure that you retain a HARD COPY of ALL projects**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. Please do not bind projects or place in covers unless specifically requested.

4. You must ensure that all projects are submitted to your Programme Coordinator on or before the required submission date. **Late submissions will incur penalties.**
5. All projects must be submitted and passed in order to successfully complete the year. **Any project/assignment not submitted will be marked as a fail.**

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Business Intelligence and Business Analytics Project Specification Report

Deep Mayur Thakar
MSc. Data Analytics
National College of Ireland
Dublin, Ireland
x20208111@student.ncirl.ie

Juan Pablo Andres Palma Bustos
MSc. Data Analytics
National College of Ireland
Dublin, Ireland
x20225776@student.ncirl.ie

Prashant Digambar Waghela
MSc. Data Analytics
National College of Ireland
Dublin, Ireland
x20207786@student.ncirl.ie

Tanmay Shrikant Paranjape
MSc. Data Analytics
National College of Ireland
Dublin, Ireland
x20101015@student.ncirl.ie

Vishal Gajanan Patwardhan
MSc. Data Analytics
National College of Ireland
Dublin, Ireland
x18190839@student.ncirl.ie

***Abstract*—The paper provides a detailed study about the business analysis of Staples. The company has corporate clients across the globe and is best in class in generating sales in furnishing products, home office accessories, electronics, and hard goods. The business model is analysed using the sales data of 2 years based on the unit orders generated by the employees, its gross profit margins, and country-wise distribution of earnings. To understand the business structure, an entity-relationship diagram is created for connecting matching data fields across various tables. The E-R diagram plays a crucial role in learning the data structure precisely. All the sections of the diagram are explained in a tabular format, which provides detailed information about the major data fields. The above-mentioned factors help implement the sales analysis of the organization.**

I. History and Marketplace of Staples Inc.

Staples Inc. is a retail company based in America which primarily involved in office supply products. Staples Inc. started its first retail store in 1986 with the aim of doing retail business which later expanded to online and direct contract sales. The company presently works as a b2b (business to business) model, providing a supply of office accessories to its corporate clients across the globe via a web-based portal, retail stores, and direct shopping. The company's primary focus after the establishment was supply-chain management to fulfill the product demand of its corporate clients hence, making the retail and e-commerce companies its target audience. To provide better service and market expansion Staples Inc. extended its business across 6 different countries. The company has various warehouses and distribution centers and its headquarters is in the United States.

II. Understanding the Business of Staples Inc.

Staples Inc. was established in 1986 grew rapidly and showed substantial growth in its revenue. By 1996, they got listed in the Fortune 500 companies. The business model of Staples Inc. was quite innovative and it looked upon fulfilling corporate demands related to office accessories via its retail stores across the United States. However, post-2011, Staples Inc. experienced a gradual decline in its sales numbers which resulted in major losses. The rise of other companies like Amazon, Walmart who provided solutions in a more innovative way made them the market leaders and eventually resulted in the loss of clients for Staples Inc. After 2013, the trend to go online was realized and Staples Inc. started providing e-commerce services for order placement. From Fig. 01., it is understood that after 2011, the company experienced major losses. The paper will provide a detailed study to find the reason for these losses by studying its sales data for the years 2020 and 2021. Solutions will be given using various Business intelligence methods that need to be implemented to retain its customer base and become profitable to some extent.

Fig. 01. Staples Inc. Revenue Trends 2009-16 [2]

III. Significant Insights of the Organisation

Staples Inc. provides solutions in wide areas like furniture, major appliances, home electronics, bedding, hard goods, home office accessories to their corporate clients. They generate and maintain business contracts with corporate organizations to fulfill their day-to-day office needs. The primary order methods through which Staples Inc. accepts purchase orders from their clients include online, direct, and retail orders. Exploring the order data from January 2020 to June 2021 concludes that, the retail order method plays an important role in generating higher revenue than direct and online order methods. Between Jan 2020 and June 2021, the company generated a revenue of $1,456M with a gross profit of $275.43 M. Staples Inc. has been more successful to provide better supply to the US-based clients which contributed to a profit of 41.91%. According to the company's selling trends, Office Furniture has generated the highest revenue with a profit of $66.62 M.

IV. BI System Design

Business Intelligence systems are about taking raw data from various sources, cleaning that data, and converting it into valuable information, and then using that information to make valuable decisions in the organization. All the processes involving data cleaning, data transformation to output data occur in BI Systems.

How the organization's BI system will work depends upon the efficacious architecture used.

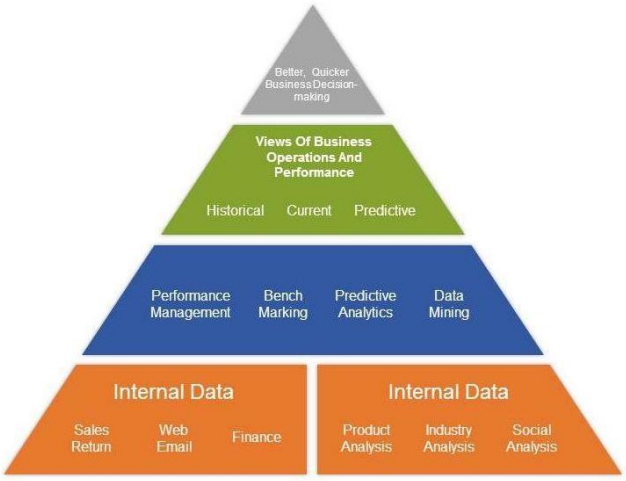

Fig. 02. Generalized Business Intelligence System Diagram [3]

The wide and in-depth knowledge and industry experience are very important in the construction of a full-fledged BI system. The cost of BI system architecture may vary with the size of the business and the data that needs to be handled. Inside the business intelligence architecture, the first component is Data Capture and Integration. Business data needs to be captured from various branches, warehouses through catalogs and reports. The data is then loaded into various tables in databases where a relationship can be built between multiple entities. The size of data depends upon the size of the business. To maintain data and transactions it is loaded into data warehouses to which all the databases are getting connected. The data gathered from various sources need to undergo cleaning and transformation in order to get meaningful information for the business. The clients can also be connected to data warehouses via web applications and hosted websites. The data cleaning and transformation is a continuous work, to fit into changing market requirements so that the business strategies can be adopted. Once data is organized in an appropriate manner it is very easy to do forecasting because it gives the ability to make business decisions and develop data-oriented strategies. Past data is stored because it helps to generate periodic patterns and trends. Forecasting allows organizations to take proactive decisions.

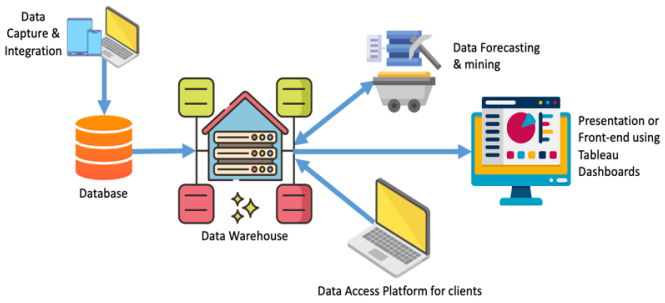

Fig. 03. BI System Design of Staples Inc.

The business intelligent system of Staples Inc. work as follows:

1. Business data is getting captured from various countries with respect to various products and associated clients.
2. Data is stored in MySQL relational databases.
3. The databases are connected to huge data warehouses, which are up and running 24*7.
4. Data is then cleaned in order to remove inappropriate entries to perform financial adjustments.
5. To find periodic trends and pattern forecasting is done on the data by fetching it using various BI tools.
6. Appropriate reports are generated and sent to management who provide insights and feedback. The updated data is again stored back into data warehouses.
7. Clients and End-users also have access to a few tables in order to find out about various products of the company.
8. From data warehouses bulk data gets fetched using various tools such as SQL developer which is then stored into Excel sheets.
9. Since it's quite tedious to work on an excel sheet, Staples Inc. is using the Tableau tool for visualization purposes.
10. Various reports and dashboards are generated using Tableau so that, it can be easier for Managers to take appropriate decisions. From various visualizations, various flows in business can be identified. The performance of each employee can also be tracked here. The business in various countries can also be evaluated.

V. DATABASE DESIGN

A. *Entity Relationship Diagram*

An Entity-Relationship (ER) Diagram is a kind of flowchart which illustrates that how entities like people, concepts, and things inside the database or any system interact with each other. In all the sectors like education, research, business information systems, and software engineering, ER diagrams are mostly used to troubleshoot and build relational databases. There are set of symbols used like ovals, diamonds, rectangles, diamonds, and lines to show the relationship of the entities, properties, and the relationships which is also known as ER models or ERDs.

The uses of the E-R diagram are as explained below.

- Database Design

An ER Diagram helps to build the relational models with keeping the business understanding and rules in place. In software engineering, it is mostly used to implement the structure of the databases having a relationship through the various entities which is mainly done with the concept of primary and foreign keys.

- Database Troubleshooting

Database troubleshooting often becomes the best way to analyze the problems in the entire database system by drawing the ER diagrams.

- Business Information Systems

Diagrams are used to analyze and initiate the business infrastructure which can help to understand the relationship between any modules or department. It also helps to streamline the process and disclose the information much more easily and quickly.

- Business Process Re-engineering

ER diagrams help to model new database setups and also analyses the database system used in re-engineering business process.

- Education

ER Diagrams prove to be more valuable in the field of education which helps to create a method to store the relational data for later use.

- Research

Research mostly focuses on the structured data-set in which ER diagram plays an important role in setting up useful databases to get insights into the data.

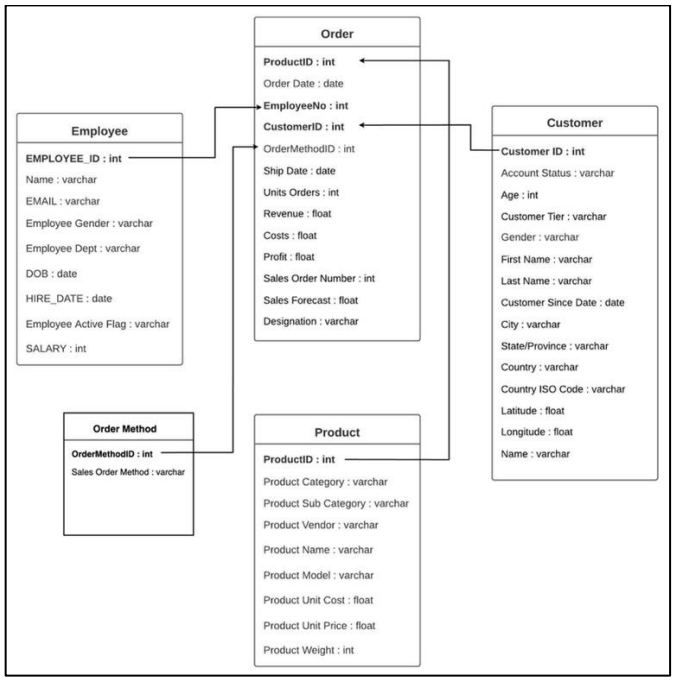

Fig. 04. E-R Diagram of Staples Dataset

Fig. 04. shows the database structure of the sales process of Staples Inc. Rectangular blocks with the heading Order, Employee, Customer, Order Method, and Product are the entities present in the database. Each entity has its separate set of attributes and the line connecting different entities shows the relation between them and the way they interact with each other. 'Order' table is a master table which comprises of attributes present in relation with other entities like 'ProductID', 'EmployeeNo', 'CustomerID', and 'OrderMethodID'. They are known as foreign keys. These relations are achieved by drawing a relational line between the attributes of different tables.

VI. DATASET DESCRIPTION

The Staples dataset consists of Customer data, Employee data, Product data, Order Methods, and Order data from January 2020 to June 2021. A detailed explanation of each of these datasets is as mentioned below.

A. Customer Data Table

The Customer table contains data of each corporate client. This table also includes the region in which the client operates. Additionally, the table classifies each client into tiers based upon the business they generate along with their activity status. There are 15 data columns in the table which can be referred to from Fig. 05.

Field Name	Data Type	Description	key	Mesure or Dimension
Customer id	int	Unique Customer id	Primary key	Dimension
Account Status	varchar	customer's account status	NA	Dimension
Age	int	Age of customer	NA	Dimension
Customer Tier	varchar	Catagoery of customer based on business done.	NA	Dimension
Gender	varchar	Gender of customer	NA	Dimension
First Name	varchar	Intial name of customer	NA	Dimension
Last Name	varchar	Last name of customer	NA	Dimension
Customer Since Date	date	Customer enrollment date	NA	Dimension
City	varchar	City of customer	NA	Dimension
State/Province	varchar	State of customer	NA	Dimension
Country	varchar	Country of customer	NA	Dimension
Country ISO Code	varchar	International recognized country code	NA	Dimension
Latitude	float	Geographic coordinate of country	NA	Measure
Longitude	float	Geographic coordinate of country	NA	Measure
Name	varchar	Name of Company	NA	Dimension

Fig. 05. Customer Table Data Field Information

B. Employee Data Table

The employee table has comprehensive data about the employee of Staples Inc. based on employee hierarchy, their business units, and the department in which they work. The table also represents whether an employee is a manager or not. Employee classification is based upon the business sector they work such as finance, customer service, corporate finance, retail, sales, and other service lines.

Field Name	Data Type	Description	Key	Measure or Dimension
Employee ID	int	ID of employee	Primary Key	Dimension
Name	varchar	Name of employee	N/A	Dimension
EMAIL	varchar	EMAIL of employee	N/A	Dimension
Employee Gender	varchar	Gender of employee	N/A	Dimension
Employee Dept	varchar	Department of employee	N/A	Dimension
DOB	date	DOB of employee	N/A	Dimension
HIRE DATE	date	Joining date of employee	N/A	Dimension
Employee Active Flag	varchar	Activity status of employee	N/A	Dimension
Salary	int	Salary of employee	N/A	Dimension

Fig. 06. Employee Table Data Field Information

C. Product Data Table

The Product table has all the information about the sectors in which Staples Inc. has its business. The table has information about various Product categories which include Office accessories, Furniture, Hard Goods, Home Electronics, Home office accessories, and Major appliances. Each product category has numerous subcategories which include a large number of products. Moreover, the table contains information about the model of the product, its unit cost to the company, and the final unit sale price. This table includes crucial information that drives the business model.

Data Field	DataType	Description	Primary/Foreign	Type
ProductID	integer	Unique ID for each product.	Primary Key	Dimension
Product Category	varchar	Primary product categories.	N/A	Dimension
Product Sub Category	varchar	Sub-categories of each product	N/A	Dimension
Product Vendor	varchar	Unique product vendor key.	N/A	Dimension
Product Name	varchar	Name of the product.	N/A	Dimension
Product Model	varchar	Model name of the product.	N/A	Dimension
Product Unit Cost	float	Cost of each unit product.	N/A	Dimension
Product Unit Price	float	Sale value of each product.	N/A	Dimension
Product Weight	integer	Weight of the product.	N/A	Dimension

Fig. 07. Product Table Data Field Information

D. Order Method Table

The Order Method table has categorical information about how a company generates its orders. It includes Retail, Direct, and Online Sales orders.

Data Field	DataType	Description	Primary/Foreign	Type
OrderMethod_ID	integer	Unique ID for each order method	Primary Key	Dimension
Order Method	varchar	Order methods classified as retail, direct and online.	N/A	Dimension

Fig. 08. Order Method Table Data Field Information

E. *Order Data Table*

The Order data table is the master table of Staples Inc. The table includes references from Customer, Employee, Product, and Order Method table. After studying the Order table, the company's revenue, costs, and profit can be computed. This table has important data points about the number of units that have been ordered and how much gross profit is generated from it. Sales forecast based upon previous financial years is given. The employee designation data for which the company generated its highest profit and unit orders can be computed from specific data fields. The detailed insights about companies' business strategies, their strengths, and weaknesses can be evaluated from this dataset. It represents a business catalog of Staples Inc.

Data Field	DataType	Description	Primary/Foreign	Type
ProductID	integer	Unique ID for each product.	Foreign Key	Dimension
Order Date	date	Date when order is placed.	N/A	Dimension
EmployeeNo	integer	Unique Employee ID.	Foreign Key	Dimension
CustomerID	integer	Unique Customer ID.	Foreign Key	Dimension
OrderMethodID	integer	Types of order methods.	Foreign Key	Dimension
Ship Date	date	Date when order is shipped.	N/A	Dimension
Unit Orders	integer	Number of unit order sales.	N/A	Dimension
Revenue	float	Total revenue for unit orders.	N/A	Dimension
Costs	float	Total cost to company for unit orders.	N/A	Dimension
Profit	float	Gross profit from unit orders.	N/A	Dimension
Sales Order Number	integer	Unique order number for each sale.	N/A	Dimension
Sales Forecast	float	Forcasted sales value of unit orders.	N/A	Dimension
Designation	varchar	Designation of the employee who generated unit sale orders.	N/A	Dimension

Fig. 09. Order Table Data Field Information

VII. REQUIREMENT ANALYSIS

The requirement analytics is a crucial process in a company's business model. It is a detailed study of what changes are required in an existing business, how customers' needs are fulfilled, and what major changes need to be implemented to improve the business performance. The requirement analytics should necessarily consider the conflicts in requirements of various stakeholders to manage and streamline the process accordingly. Various events need to be accounted for a detailed study, documentation, validation, and later management. One of the most important aspects of requirement analytics is information management. It is a way of making an inventory so that it can be used and referred to at any stage of the business. Information management needs to have well-standardized policies and protocols.

A. *Business Requirements:*

Business requirements are a critical process in all organizations. These are the activities that are necessarily required to perform and meet the goals and objectives of a business. To understand the business requirements, customers' expectations and needs must be analyzed. This facilitates communication medium between business and customer. It also gives useful insights into the subsequent phase of business. Customers' expectations need to be fulfilled to run a business efficiently. For example, gathering information from various loyalty programs and populating the inputs from it, and using those to improve business.

B. *Requirements for architecture and design:*

In order to document the detailed study of business requirements and their associated factors, it is important to put this in a proper framework and structure. The design requirement in the Staples Inc. model is an E-R diagram. In order to build architecture, connections can be established between tables by use of Primary and Foreign keys. It provides a meaningful relationship between unique keys to create database understanding.

VIII. OBJECTIVE OF THE PROJECT

The objective of the project is to implement an end-to-end business solution for Staples Inc. which is scalable, feasible, and cost-efficient. The obtained sales data will be studied and all the processes that need to be implemented on it are listed below.

- Creating a star schematic in Tableau to generate relationships between different matching entities among separate data tables to join them all in a logical manner.
- Implement data cleaning and filtering in Tableau on necessary columns to eliminate junk values.
- Identifying a set of most profitable products.
- To identify country-wise profit and unit sales trends.
- Understanding sales trends and comparing it among different financial years.
- Evaluating aggregate percentage profit on all the products and comparing these results in a month-wise manner.
- Identifying top clients that bring maximum business to the organization.
- Understanding the sales numbers based on various order methods to identify where the company is falling short.
- To analyze all the above points and provide business optimization solutions to improve market-performance of Staples Inc.

All these points in a combined way will play a crucial role in evaluating the strengths and weaknesses of Staples Inc. The listed points need to be visualized in a meaningful way to generate business insights.

REFERENCES

[1] https://www.kaggle.com/worldopendata/sales-data-2021

[2] https://www.statista.com/statistics/316986/staples-inc-gross-profit/

[3] https://www.slideteam.net/example-of-business-intelligence-system-diagram.html

[4] https://www.liveplan.com/blog/what-is-a-swot-analysis-and-how-to-do-it-right-with-examples/

[5] https://www.isixsigma.com/implementation/project-selection-tracking/business-requirements-document-high-level-review/#:~:text=Business%20requirements%20are%20the%20critical,of%20customer%20needs%20and%20expectations.

[6] https://disruptiveadvertising.com/business/b2b-companies/