**Applied Data Analysis Techniques for Businesses**

**Assignment Report**

Data Analytics for Consumer Product Development for Dyson Company Limited

**Student Name**: Jones Kwame Osei

**Student Code**: s4112106

**Module Title**: CT7203 Applied Data Analysis Techniques for Businesses

**Module Tutor**: Professor Shujun Zhang

School of Computing and Technology

University of Gloucestershire

10/06/2022

Contents

[Introduction 3](#_Toc108774616)

[Part One 4](#_Toc108774617)

[1.1 Business Case – Dyson Company Limited 4](#_Toc108774618)

[1.2 Business Profile 4](#_Toc108774619)

[1.3 Business Strategy 4](#_Toc108774620)

[1.4 Critical Analysis of Business Strategy and Recommendation 6](#_Toc108774621)

[1.5 Data Requirements 6](#_Toc108774622)

[1.6 Challenges of Data Analysis 7](#_Toc108774623)

[Part Two 8](#_Toc108774624)

[2.1 Discussion of the Data Analysis Techniques 8](#_Toc108774625)

[2.2 Data Collection 8](#_Toc108774626)

[2.3 Data Cleaning 9](#_Toc108774627)

[2.4 Data Exploration Analysis 10](#_Toc108774628)

[2.5 Communication and Visualisation 11](#_Toc108774629)

[Part Three 11](#_Toc108774630)

[3.1 Critical Analysis and Evaluation of the Applications of the Techniques 11](#_Toc108774631)

[3.2 Application of Data Collection 11](#_Toc108774632)

[3.3 Application of Data Cleaning 12](#_Toc108774633)

[3.4 Application of Data Exploration Analysis. 13](#_Toc108774634)

[3.5 Application of Communication and Visualisation 13](#_Toc108774635)

[Part Four 14](#_Toc108774636)

[4.1 Critical Review of the Commercial and Open-source Software 14](#_Toc108774637)

[4.2 Software for Data Collection 14](#_Toc108774638)

[4.3 Software for Data Cleaning 15](#_Toc108774639)

[4.4 Software for Data Exploration Analysis 16](#_Toc108774640)

[4.4 Software for Communication and Visualisation 16](#_Toc108774641)

[Conclusion 17](#_Toc108774642)

[References 18](#_Toc108774643)

# Introduction

This assignment critically examines how data analysis techniques and their underlying features are applied in the business world. Specifically, four data analytic techniques will be explored, crucially analysed and evaluated to understand and maximise their applications to enhance production for customers by Dyson Company Limited.

* Part 1 of the report covers what Dyson Company is about, its current value and position in the market, its competitors in the industry, business strategy, data requirements, and then further addresses the data analytic related problems of the business case.
* Part 2 entails the four relevant data analysis techniques that Dyson uses to achieve its goals.
* Part 3 of the report looks at the critical analysis and evaluation of the four data analysis techniques being used by Dyson.
* In the Part 4, the report throws more light on the commercial and open-source software for each of the four data analysis techniques and critically review them as well as the conclusion.

# Part One

## Business Case – Dyson Company Limited

### 

### 1.2 Business Profile

Dyson Limited, commonly known as Dyson is a United Kingdom multinational technology company specialising in the sales and services of domestic appliances and has its headquarters currently moved to St. James Power Station across the Island of Sentosa, Singapore(Min 2022, 2022 ). About thirty years ago, on the 8 July 1991, Dyson was established and founded by Sir James Dyson the inventor of the dual cyclonic bagless vacuum cleaner at Malmsbury in England(Gov.uk). This company designs and manufactures products including its famous vacuum cleaners, hair and hand dryers, bladeless fan, heaters and lights.

Presently, Dyson operates in an extremely competitive business environment(Prasad 2009). Competitiveness in the industry has helped the company under study to implement modern business strategies to have a competitive advantage over its competitors and by far, Dyson is he market leader in upright, cylinder and robotic vacuum categories whiles Vax is the main competitor in the market. The race for market supremacy and leadership in the industry has tremendously increased over the years for Dyson from other industry contenders such Hoover, MediaMarkt, Lotus Electronic, Electrolux(Motohashi 2015). If competition is high, the only possible way is to strategically adopt, embrace and initiate new technological tools as contention from other companies remains a reality for Dyson.

### 1.3 Business Strategy

Dyson as compared to today’s business entity, depends hugely on the collection of data for better consumer experience as well as the growth of the company. The collection of data is generated from both external and internal sources. External data collection encompasses the experiences and feedback of customers whilst internally, thus within Dyson, information on employee’s productivity and target achievements as well as yearly appraisals are recorded. These data from consumers and staff help the firm to basically understand and use them (data) to enhance products for the satisfaction of consumers and leverage to motivate and increase staff productivity respectively. Notwithstanding the relevance of both sources, the information from outside, thus, from customers should always be prioritised. This is because the experiences of consumers comprises all facets of the company’s contribution including the quality and durability of products, packaging and even customer service(Gentile, Spiller et al. 2007). Having seen a paradigm shift in modern businesses where consumers now interact with firms through myriad touch points in several channels and avenues as customer experiences are exceedingly more social in nature., these changes require a company like Dyson to integrate and activate multiple business functions for its sustainability(Lemon and Verhoef 2016). To achieve this goal, our case study will examine existing business strategies of Dyson Company Limited, critique it and suggest why it is appropriate to use new technologies to modify them.

The company formulates and implements a wide range of business strategies to focus on user experiences to remain competitive. Dyson uses a product differentiation strategy(Porter 2008) which they seek to produce home appliance that are different from the similar products from its competitors. Obviously, the goal of the company is to achieve a competitive advantage by manufacturing a completely unique products to consumers. This uniqueness of the product in the market due to differentiation makes it more aesthetically attractive to customers. The company’s robotic vacuum cleaner one of most popular product in the market because of its distinctiveness. Even though this production strategy can potentially maximise Dyson’s profit, the differences made in the products must always reflect the values of customers.

The company to cut down production waste also incorporates the lean production model(Holweg 2007). Dyson uses this strategy to eliminate waste, where waste can be anything that does not add value for the customers. This justifies that in as much as the company seeks to increase profit, customer values and experiences should be the centre of focus in production. The implementation of this model has helped this firm to save production cost(Alves, Dinis‐Carvalho et al. 2012) and not result to low cost strategy, which is not good strategy for their image and market position. Low cost strategy denotes businesses attempting to reduce its production costs be it raw material, research, or promotion. Dyson does not give room to this policy because they lose their market supremacy and incur a lot of losses.(Murni 2017).

### 1.4 Critical Analysis of Business Strategy and Recommendation

It is tangible Dyson focuses on products refinements and differentiation. However, doing this without adequate deployment policies can lead to huge decrease in profit. Dyson has not clearly exhibited the involvement of technology in their attempt to modify and improve their products for customers. Data analytics toolkits can be useful in their effort to achieve that(Sedkaoui 2018). Using data analysis techniques to know customer behaviour can help support decisions of enhancing products because the values of the consumers will be captured in the data analysis. These insights can further enable the company to reduce drastically, production waste.

There are several strategies that be recommended to Dyson. The company should continue with the differentiation strategy; however, it will be more beneficial to strongly adapt and embrace various data analytic technique to and predict when a product should be developed or modified.

### 1.5 Data Requirements

Today, in over sixty-five countries, Dyson machines and home appliances are being sold around the world. This means the company has many customers across the globe and requires huge amount of data. These are mostly customers experiences and personal data. It is necessary for the company to collect personal details of its customers because these data enable Dyson to provide and improve services, product and experiences that are expected from a top tier firm. The company collects its data from their website or app, by email or phone, or in person at their shops. Information from customers connected to Dyson products – products that are intelligent or smart devices is stored as data. Also, information from their various shops and events across the world such as Close Circuit Television (CCTV), film footage and photographs are stored as data. These data can be seen as big data and Dyson needs to manage, protect and handle it appropriately without breaking data protection laws. Hence, Dyson use appropriate technical and institutional measures, including encryption, to protect personal data and privacy. Big data including personal data from customers are very sensitive and must be carefully handled as any breach of it will affect the trust and confidence customers have in the company. Consequentially, it can impact on sales and profits.

### 1.6 Challenges of Data Analysis

It can be agreed that Dyson is a successful competitor in the electric and home appliance industry. This firm’s powerful strategies have over the years attracted new customers, produce innovative products, and address some changing demands of its consumers. However, the company has continuously been affected by several data analytic problems and these have had a negative effect on its performance in the industry.

Notwithstanding the company’s efforts of data collection, there has been concerns of acquiring useful or real time data and even the available data collected have not been used to its maximum. Again, the manual collection of big data for an internationally recognised company like Dyson has made it difficult for the firm to gather and acquire up-to-date data on the contemporary needs of their customers. Another problem of the firm is its long data response. There are delays in getting real-time insights because their system is designed for batch processing. This is the method which computer systems complete programmes known as jobs in batches(Carbone, Katsifodimos et al. 2015). Batch processing inhibits availability of data since it is still being gathered or pre-processed.

Furthermore, Dyson’s worrying challenge is poor data quality and messy data visualisation. Poor quality of source data stems from reliance of data that has defects, errors, and incomplete. Data quality determines the quality of outcomes and obviously, dirty data produces poor and inaccurate results. This problem may bring all efforts of supporting decision making through data visualisation to naught. Uncleaned data produces messy visualisation which gives incorrect insights about the business. This is a case of “garbage in, garbage out”. Data quality management and responsible data validation process, automating data collection and the appropriate use of data collected can help solve the issues associated with data analytics in this firm. Hence, we will address them through technology by exploring four data analysis techniques: Data collection, Data cleaning, Data exploration and Communication and Visualisation.

# Part Two

## 2.1 Discussion of the Data Analysis Techniques

### 2.2 Data Collection

Data Scientist, Clive Humby in 2006 declared that “Data is the New Oil” , an assertion that means the world’s most valuable resource is no longer oil but data(Hirsch 2013). Although Humby said this about sixteen years, it is the reality today. This is because there is huge money to be made if data and for that matter, big data, is handled well. On the other hand, a lot of money can be lost if data is recklessly handled. Data are proofs or statistics that are generated by technically observing underlining factors(Sapsford and Jupp 1996).

Companies today including Dyson are conscious of the relevance of big data and hence, the need to gather and collect useful data that support their decision making. This is not about just sourcing for any data but real time data. Real Time Data (RDT) is the smart way of distributing information as soon as it is collected and it is mostly used for tracking(Croushore 2011). They are statistical facts derived directly from the primary source as swiftly as possible. These kinds of data are always up-to-date and any changes in the original source will be seen immediately. A firm like Dyson needs to collect real time data about the experiences of its clients or customer activity status to know their ever changing demands. Knowing these about their consumers will be beneficial to production improvement, saving time and cost of production. The acquisition of these kinds of data can be derived from data mining. Data mining, which is an embodiment of data collection will be the focal point when we are critically analysing the application of this technique by our case study.

### 2.3 Data Cleaning

Michael Palmer, the American author in expanding Clive Humpy’s “Data is the New Oil” assertion, he explained that “Data is just like crude. It must be refined into gas, petrol, diesel for it to be valuable”(Palmer 2006). It is no doubt that cleansing of data is an important technique that the firm under study should pay much attention to. In 1.6, we identified that a serious concern of the company is poor data quality. Obviously, to get an unblemished data, it must be cleaned and pre-processed well to remove any noises, biases and outliers. Data cleaning, also called data cleansing or scrubbing, deals with detecting and removing errors and inconsistencies from data to improve the quality of data(Chai 2020). This should be the first task of the company’s data analysts should begin with. Without acknowledging errors in data makes the analysis vulnerable and susceptible to failure.

Detecting and repairing dirty data is one of the perennial challenges in data analytics, and failure to do so can result in inaccurate analytics and unreliable decisions(Natarajan, Li et al. 2010). Dyson’s business, services and decision-making hugely depends on storing and obtaining big amounts of data. These datasets can support and improve management decisions as well as exponentially optimise services and production.

Notwithstanding this importance, the quality of data continues to be a greater concern for the firm and can make results unrealistic subsequently affecting decision support. To avoid this, the company has a duty to effectively detect the very errors in their data and subsequently repair or correct them. Some of these blunders are as a result of data entry (if manually collected) especially if those making the entry are untrained to carry out that task. Some common mistakes like typo errors, adding commas to figures, incorrect date format, duplicate entries and many more can render a data uncleaned.

Due to the forgone, there is the need to data to have something more meaningful to make tangible inferences from it. The conversion of raw data into a form that will make it less difficult to understand and interpret solely depends on a clean data. This can be achieved by rearranging, ordering, and manipulating data to provide more insightful information for the firm.

### 2.4 Data Exploration Analysis

After the data has been collected and processed to a clean one, it must be explored to give a more intuitive information about the data. Exploratory Data Analysis (EDA) is usually intended to generate hypotheses and not to lead to conclusions based on the results of the study (Shinde, Majumder et al. 2022). This is done to know and understand the data that will be used to make more in-depth analysis. It is a good step even before data cleaning and after data cleaning before any data analysis is done.

Data analysis is the process of extracting insights from data. Data is heterogeneous in all ways, and processing such data is a challenge. Before performing an analysis to any data to obtain business insight, it is necessary to understand the data and its associated problems if any, deal with the missing values and noise, visualize the dataset, and select the machine learning model to analyse the data(Taylor and Cihon 2004, Shinde, Majumder et al. 2022). Data exploration in this report will be critically analysed in the subsequent parts how Dyson will utilise this technique. The technique leverages on both basic and advanced data exploration as well statistical evaluations and inferences.

### 2.5 Communication and Visualisation

Finally, communication and visualisation are the last techniques to be introduced. Dyson endeavours to promote and enhance customer’s experiences in real time. To do this, they need to make crucial business decisions on the bases of analysis made. Without visualising and communicating results to decision makers, there can not be tangible business decision made. Visualising data is serious phase in the business intelligence process since it factors raw data, transform or models it and delivers insightful information support production decisions.

Data Visualisation simply means visually demonstrating information in the form of pictures, images, diagrams and graphics(Chen, Härdle et al. 2007). Visualising data is an important tool, however, it only becomes relevant to the company if results are appropriately communicated to the stakeholders. The means of sending meaningful information and insightful results about the data visualisation to support decision is seen as an art of communication. Our study looks at how effectively Dyson make good use of these tools.

# Part Three

## 3.1 Critical Analysis and Evaluation of the Applications of the Techniques

### 3.2 Application of Data Collection

## 

Understanding the company’s culture, dynamism and conducting an inhouse and external evaluations about the happenings within and outside can be more beneficial before collecting the data. This means, Dyson needs to identify related customer problems which needs to be resolved. In other words, they need to set goals for the reason of collecting the data. Just like any industrial company, setting the right goals is key to success. This goal should be achievable and realistic for the business. Once this has been clearly identified, data collection can proceed. Dyson needs to collect real time data to know consumer needs to improve its production, hence, it can collect its needed data through data mining. Dyson can use data mining to extract more valuable data for the company. Data mining can be defined as a collection of technological skill sets, processes and analytical approaches combined to uncover insightful knowledge and information in data(Roiger 2017). The company through automation can collect data in the form of quantitative and qualitative data. Quantitative data is in the form of numbers like the number of customers who bought cordless hoover whilst qualitative data is descriptive in nature and non-numeric such as the review of the cordless hoover by customers. Collecting these kinds of data can put the company in the right place to follow, know and understand customers behaviour in real time.

### 3.3 Application of Data Cleaning

Dyson’s insights and analysis will only be useful if their data is also good. As explained above, garbage data produce garbage analysis and will lead to garbage results. While the techniques used for data cleaning may be case by case specific, there should always be justification for doing this.

The data must be cleaned to have a good data to derive a realistic analysis to obtain a result that can effectively support decisions. Generally, data cleaning reduces errors and improves data quality(Natarajan, Li et al. 2010). Some of the data cleaning techniques that can be adopted may include removing an irrelevant entry, handling missing values, converting data type, standardising capitalisation and wording, renaming, clear formatting, and fixing errors and duplication. Entries that are not important to the analysis will only make it slow and might even confuse the analysis to be done. Ensuring only what is relevant and necessary before the analysis is key to success. Also, duplicates and unwanted observations can be removed from the dataset. Duplicates most at times are as a result of manual data entry. Again, client’s data from multiple sources can create duplication. Outliers should be filtered and missing values can be replaced or removed when it is not significant in the analysis. The justification for the integration and data cleaning is that sometimes, the data is provided in parts, and it is desirable to implement the analysis wholistically.

### 3.4 Application of Data Exploration Analysis.

Dyson realising that data can play a significant role in improving its productivity as well as efficiency and support decision that can lead to achieving higher profits, sales and revenue at this point will want to explore its data after being collected and cleansed. It has now access to huge amount of data but can only be valuable to the business if factual information has been analysed through data exploration as a way to grow the business.

Luckily for Dyson, the era of technological dispensation has brought computer systems and innovative ways of business making good use of data. The company can use these tools to control data sources efficiently and effectively, assisting its data scientists to look for answers the company seeks by discovering the patterns in the data, detecting anomalies, confirming and testing hypothesis.

The exploratory data analysis can be further broken into two, basic data and advance exploratory analysis. The basic exploratory data analysis is done to know the data set to be used. It encompasses exploring the shape of the data, info of the data, its structure, dimension, type, as well as generating its statistical summary. Here, further investigation about missing values can be done. The advance data exploratory analysis is when data scientists of the firm dig deeper into the variables in the data set(Tukey 1977). Dyson, that operates online sales for example, can perform an exploratory data analysis to ascertain the different sales trends graphically and visualise data related to some of the best selling product’s classifications, the demographics and consumer preferences, patterns of customer spendings and units of products sold within a particular period. Dyson can only come into terms with this through exploratory data analysis.

### 3.5 Application of Communication and Visualisation

Dyson has a strong team of data scientists to visualise various components such as univariate variables either presenting a complete or non-graphical picture of the data or multivariate variables also either showing a full image or non-image of the data. Dyson could maximise the power of visualisation by setting a workflow design to gain a better understanding of the data by appropriately using visuals to present the various data categories. Continue variables can be measured or presented with a histogram whilst a categorical variable could be displayed with a bar chart. Covariation can also be measured with symbol size or boxplots. To determine the distribution of the data, the boxplot can again be used and this helps to known how skewed the data distribution is. The skewness will further inform how the missing values should be replaced. Either by using the mean, mode or median.

All the results of the exploratory data analysis and visualisation only support Dyson’s business decisions if they are clearly and properly communicated to the decision makers. The company’s data scientists must use piece of software and platform to facilitate internal and external communication(McNamara 2014). For stakeholders, the appropriate channel is to present the findings to them by face-to-face presentation like a board or stakeholders meeting. An importance issue here is to consider the technical background of the stakeholders. If they are non-information technology persons, what is shared with them should be comprehensible and brief. This means a summary of the findings in a layman’s terms should be presented to them.

# Part Four

## 4.1 Critical Review of the Commercial and Open-source Software

### 4.2 Software for Data Collection

Data collection software are tools that assist businesses collect and save data. The means of collecting data by Dyson depends on what type of data they need and what they want to use the data for. There are several software for data collection including Hadoop, Apache Storm, MongoDB and Konstanz Information Miner.

The Apache Hadoop software library is a framework allowing large data to be distributed across clusters of computers(O’Driscoll, Daugelaite et al. 2013). It is efficiently designed to scale up from single servers to thousands of computer systems and it gives users the opportunity to write and test distributed systems competently and at the same time automates work across the systems. Apache Storm is a free distributed real-time computation system that enables real tie processing of huge streams of data easy for real time dispensing and integrates very well with any programming language.

Another open source software is MongoDB which is a NoSQL, document biased database written C, C++, Python and JavaScript. Konstanz Information Miner is a tool used for Enterprise report reporting, customer relational management, data mining, data integration, text mining, and data analytics. It supports all operating systems, integrates smartly with other technologies and programming languages. This open source tool allows users to automate a lot of manual work, in an excellent and organised workflow environment.

All the above software are efficient and Dyson can use any of them. However, Dyson has a company that has huge data requirement and real time data, the Apache Storm is the most appropriate and highly recommend tool for data collection and processing.

### 4.3 Software for Data Cleaning

### 

Data mining, exploration and visualisation is all about identifying patterns in big data. However, to get the most out of it, the data must be clean without any dirt, noise and outliers. Tools like Microsoft Excel, Python and R can be used to scrub data but there are also other data cleaning tools available. These include OpenRefine, Trifacta Wrangler and Data Ladder (Datamatch Enterprise).

While OpenRefine allows it users to transform data between different format and ensures that data is structurally cleaned, Trifacta Wrangler also allows its users to transform data carryout analyses and produce visualisations.

Data ladder is a visually-driven cleansing software which focuses on customer data. It goes further to resolve data quality issues within data that are already in a poor state. Users can create anything from database tables to Excel spreadsheets or simple report after transforming their data. It is no doubt that Dyson’s data cleaning should adopt the Data Ladder software which has a wide range of functionalities and also focuses on customers data. Python and R can be used as well.

### 4.4 Software for Data Exploration Analysis

### 

The software for exploratory data analysis are the two main open source high level programming languages, R for statistics and Python. Both can serve same purpose spanning from data cleaning and pre-processing to machine learning. The Python pandas’ package can be used clean data. The matplotlib and seaborn packages is sued for plotting graphical images and its Scikit-learn package is used for both supervised and unsupervised machine learning which can predict the factors affecting products sales, the product to be sold most and optimisation of customer products deliveries.

R is specifically for statistics and has all the packages as the python. It dplyr package is used to clean pre-process data as well as generating a summary statistic for the data. The GGplot2 package is used to visualise and plot data and the rpart, mlr3 and kernLab are used for machione learning.

Dyson can adopt both software in all areas of working with its data because they care both big data friendly.

### 4.4 Software for Communication and Visualisation

Again, Python and R programming languages as mentioned earlier can be used for both visualisation and communication. Dyson currently uses Microsoft Power BI which a low level programming language. Although is it good with handling simple relationships between tables in a data model, it short falls is seen when the relationships between tables come with complexities. Also, it does not give many opportunities to configure their visuals as per their requirements. The company has limited options for what they can add or change in the visuals. Its number one short fall is it inability to handle the company’s voluminous data. Notwithstanding, it has interactive visuals and interface but sometimes becomes crowded.

Due to the foregone, Python or R best suits the requirements of Dyson. Both can handle big data, good with handling the relationships between dataframes or tables with complexities and the company can have the options of adding or changes visuals due to their flexible formulae or coding.

Both have an inbuilt markdown software that can be used to present reports and findings to decision makers. This inbuilt markdown can be knit in R or transform in Python into either HTML, Word or PDF to communicate results to support decision. R goes beyond of using its Shinny package to be build interactive dashboards to communicate results. Also, Python through the streamlit packages can create dashboard for presentations to communicate findings to Dyson decision makers.

# Conclusion

Throughout this case study, it has been identified that Dyson has state-of-the-art facilities and skills for its operations. However, it inadequately uses or explores data analysis techniques. The company will be able to get useful and real time data having executed the data analytic tools and skills discussed in this report. Dyson will now make use of the data collected as modern technologies have been explored to refine data, visualise and communicate results to support decision.

The company will continue to lead the electrical and home appliances industry as they have come to terms and abreast with technology to satisfy their customers needs because through the adaptation of the analytical tools and skills identified in the case study, consumer products will be improved.

# References

(2022 March 25). "Dyson opens new global headquarters at the historic St James Power Station." Retrieved June 15, 2022, from <https://www.dyson.com.sg/newsroom/st-james-power-station-opening>.

Alves, A. C., et al. (2012). "Lean production as promoter of thinkers to achieve companies' agility." The Learning Organization **19**(3): 219-237.

Carbone, P., et al. (2015). "Apache flink: Stream and batch processing in a single engine." Bulletin of the IEEE Computer Society Technical Committee on Data Engineering **36**(4).

Chai, C. P. (2020). "The importance of data cleaning: Three visualization examples." Chance **33**(1): 4-9.

Chen, C.-h., et al. (2007). Handbook of data visualization, Springer Science & Business Media.

Croushore, D. (2011). "Frontiers of real-time data analysis." Journal of economic literature **49**(1): 72-100.

Gentile, C., et al. (2007). "How to sustain the customer experience:: An overview of experience components that co-create value with the customer." European management journal **25**(5): 1.

Gov.uk. "DYSON LIMITED

Company number 02627406." Retrieved June 15, 2022, from <https://find-and-update.company-information.service.gov.uk/company/02627406>.

Hirsch, D. D. (2013). "The glass house effect: Big Data, the new oil, and the power of analogy." Me. L. Rev. **66**: 373.

Holweg, M. (2007). "The genealogy of lean production." Journal of Operations Management **25**: 420-437.

Lemon, K. N. and P. C. Verhoef (2016). "Understanding customer experience throughout the customer journey." Journal of marketing **80**(6): 69-96.

McNamara, J. E. (2014). Technical aspects of data communication, Digital Press.

Min, A. H. (2022, May 25, 2022). "Dyson opens new global HQ at St James Power Station, to invest S$1.5 billion in Singapore operations over next 4 years." Retrieved June 15, 2022, from <https://www.channelnewsasia.com/singapore/dyson-global-headquarters-hq-st-james-power-station-2585726>.

Motohashi, K. (2015). Global Business Strategy: Multinational Corporations Venturing into Emerging Markets.

Murni, T. (2017). "The Effect of Entrepreneurial Orientation to Low Cost Strategy, Differentiation Strategy, Sustainable Innovation and Performance of Small and Medium Enterprises (Studies at Batik Small and Medium Enterprises in East Java Province, Indonesia)." European Journal of Business and Management **9**(20): 137-153.

Natarajan, K., et al. (2010). Data mining techniques for data cleaning. Engineering Asset Lifecycle Management, Springer**:** 796-804.

O’Driscoll, A., et al. (2013). "‘Big data’, Hadoop and cloud computing in genomics." Journal of biomedical informatics **46**(5): 774-781.

Palmer, M. (2006, Nov 3, 2006). "Data is the New Oil." Retrieved June 20, 2022, from <https://ana.blogs.com/maestros/2006/11/data_is_the_new.html>.

Porter, M. E. (2008). "The five competitive forces that shape strategy." Harv Bus Rev **86**(1): 78-93, 137.

Prasad, K. S. (2009). Strategic Management: Text and Cases.

Roiger, R. J. (2017). Data mining: a tutorial-based primer, Chapman and Hall/CRC.

Sapsford, R. and V. Jupp (1996). Data collection and analysis, Sage.

Sedkaoui, S. (2018). "How data analytics is changing entrepreneurial opportunities." International Journal of Innovation Science **10**: 274-294.

Shinde, G. R., et al. (2022). Exploratory Data Analysis. Quality of Work-Life During Pandemic, Springer**:** 97-105.

Taylor, J. K. and C. Cihon (2004). Statistical techniques for data analysis, Chapman and Hall/CRC.

Tukey, J. W. (1977). Exploratory data analysis, Reading, MA.

Chai, C. P. (2020). "The importance of data cleaning: Three visualization examples." Chance **33**(1): 4-9.

Natarajan, K., et al. (2010). Data mining techniques for data cleaning. Engineering Asset Lifecycle Management, Springer**:** 796-804.

Shinde, G. R., et al. (2022). Exploratory Data Analysis. Quality of Work-Life During Pandemic, Springer**:** 97-105.