The contribution of Data Analytics in predicting the future purchase intentions of consumers

Pinakshi Kalita

MSc in Management
National College of Ireland

Submitted to the National College of Ireland, August 2019

Abstract

Title of thesis: The contribution of Data Analytics in predicting the future purchase intentions of consumers.

Author: Pinakshi Kalita

In today's global market, the impact of technology can be felt in every sector and marketing is no exception to it. This study seeks to explore the role of Data Analytics in forecasting the consumers' future buying intentions.

An extensive literature review reveals that in the last couple of decades, marketing processes have undergone remarkable transformations in the way consumer data is treated and in achieving this, emerging information technologies like Data Analytics have played a significant role. The study demonstrates that meeting customers' needs is critical for companies to acquire and retain customers and to accomplish this, anticipating consumer needs ought to be a priority for companies. While the study highlights the reasons that make data-driven techniques more reliable than the traditional theories and frameworks for the study of consumer behaviour, certain gaps have been identified in the literature including lack of relevant resources pertaining to the research topic.

For the completion of this study, the qualitative research method was chosen. 15 participants from 7 different countries participated in this study through semi-structured interviews. The interviewees are Marketing and Data professionals working in the IT industry. The final outcome of this study indicates that Data Analytics is being implemented in most of the organizations to forecast consumers' buying behaviour. Specifically, the Clustering technique and Bayesian network algorithm are considered to be the most effective models for making predictions about consumer data.

The empirical findings collected through interviewing participants coupled with the extant literature contributes to the flourishing literature pertaining to the contribution of Data Analytics in forecasting consumer behaviour.

Declaration

Submission of Thesis and Dissertation

National College of Ireland
Research Students Declaration Form

Name: **Pinakshi Kalita**

Student Number: **18115179**

Degree for which thesis is submitted: **MSc in Management**

Material submitted for award:

(a) I declare that the work has been composed by myself.

(b) I declare that all verbatim extracts contained in the thesis have been distinguished by quotation marks and the sources of information specifically acknowledged.

(c) My thesis will be included in electronic format in the College Institutional Repository TRAP (thesis reports and projects)

(d) I declare that no material contained in the thesis has been used in any other submission for an academic award.

Signature of Research Student:

Date:

Acknowledgements

I would like to take this opportunity to thank all the people who have contributed to the completion of this dissertation.

At the very outset, I would like to express my appreciation and gratitude to my academic Supervisor, Fabián Armendáriz for his continuous support and motivation allowing the successful completion of this dissertation. Besides my Supervisor, I would like to acknowledge the kind support that I have received from all the participants without whom this research would not have been complete. I highly appreciate their time and engagement in this research.

I am indebted to Dr. Colette Darcy and all the professors of the National College of Ireland for their encouragement and guidance during my time at NCI.

Last but not the least, I would like to express my sincere thanks to my family and friends for their moral support throughout the Masters programme at the National College of Ireland.

Thank you all,
Pinakshi Kalita
Date:

Table of Contents

List of Figures

CHAPTER I: Introduction

1.1 Background: In today's global market, customers have sky-high expectations from any kind of business that they engage with- be it fashion brands, retail stores or online shopping. They expect to be treated as if they are 'special' (Zakaria *et. al,* 2014). While this is next to impossible and surreal, companies make every possible attempt to achieve this to stand out among its competitors. Consequently, it has become vital for companies to be able to read customers' minds and predict their expectations. It is needless to mention that customers have a diverse range of options to choose from which is why acquiring or retaining loyal customers is no longer an easy task for companies (Narayandas, 2005). On the other hand, providing exceptional customer experience is an absolute necessity for companies. Nevertheless, it is equally important for organizations to offer their products in a way that it does not diminish their profit margins. To take care of this, organizations must also target the right audience. According to the 80/20 principle, nearly 80% of a company's revenue is raised from just 20% of its customers (Dubinsky and Hansen, 1982; Banduka *et al.*, 2017). So, why not target those 20% prospective customers and try to anticipate their exact needs? However, this requires careful analysis of relevant consumer data as the existing data can act as a channel to unleash greater insights.

Studies reveal that comprehending customers' expectations is a critical problem that organizations most commonly encounter (Christensen and Overdorf, 2000). Especially, at the pace at which changes are occurring in the present times, disruption can happen any time. Moreover, businesses fail to acquire customers when they fail to meet consumer demands. This arises the need to introduce some solid techniques that can help organizations predict the expectations of their customers with precision. While several traditional marketing techniques have been in practice since ages to anticipate customers' needs, information technologies are considered to be most effective in catering to and reinforcing customers' needs (Schneider and Bowen, 1999). Several previously conducted studies have reported how different technologies have facilitated prediction of customers' purchase behaviour (Zhang *et al.*, 2019 ; Barbera, Amato and Sannino, 2016 ; Blázquez, 2014).

This study intends to explore the contribution of Data Analytics, a booming information technology, in predicting customers' future purchases. Studies carried out by Erevelles,S.,

Fukawa, N. and Swayne, L. (2015) noted that in the last couple of decades, the magnitude of data produced has been higher than what actually can be utilized by organizations. To extract only the relevant data, hidden patterns are required to be identified and to achieve this, data analytics is recognized as a very powerful technology, particularly because it helps to capture massive quantity of consumer data in real time. These findings exhibit that organizations can greatly benefit from this technology if it is correctly implemented. However, relevant research related to the current research area is broad as well as limited. Furthermore, data analytics involves a set of different data mining techniques that can be used for making predictions. Although relevant literature presenting the most effective data mining techniques can be observed, it is inadequate to make the appropriate choice.

In a nutshell, this study seeks to explore the most reliable data mining techniques for making predictions about consumer behaviour.

1.2 Structure of the study: The dissertation is split into seven chapters. This chapter provides a brief background to the study.

- The second chapter presents the existing literature on the research topic by reviewing various journal articles and relevant frameworks.
- The third chapter outlines the aim and objectives of the research and introduces the research questions of the study.
- The fourth chapter provides the detailed methodology that will be adopted to answer the research questions. This chapter justifies the use of all the procedures and methods chosen for this study.
- The fifth chapter portrays the findings received during primary data collection.
- The sixth chapter analyzes the findings in accordance with the reviewed literature and provides an in-depth discussion with a view to answering the research questions as well as outlines the research limitations and implications of the study.
- The final chapter is the conclusion of the dissertation. It also presents the recommendations for future research.

This chapter was an introductory section of the study. The following chapter will review the extant literature in accordance with the research topic.

CHAPTER II: Literature Review

2.1 Introduction

This literature review intends to highlight the key role of data analytics in predicting consumer behaviour in marketing. This chapter will provide an overview of all the theoretical concepts, frameworks and models involved in this research. With the help of reviewed journal articles and previous research, this study mainly aims at investigating the current research topic under the following three critical dimensions:

- Marketing
- The Essence of Consumer Behaviour in Marketing
- Prediction of Consumer Behaviour through Data Analytics

The undermentioned headings will explore the relevant theories and frameworks as well as the past researches on the current research topic through empirical study.

2.2 Rationale

2.2.1 Marketing: Marketing is considered as one of the most vital elements of any business and is often considered as a firm's backbone which plays a key role in integrating all the other activities of the firm (Liu, 2017). Viewing through the prism of marketing activities in an organization helps to predetermine the connection between the organization and its markets, which in turn can substantially influence the overall success of the organization. However, it is critical to note that the scope of the marketing functions is very broad and hence extremely tough to be examined as a whole. Therefore, segregating the marketing functions into narrowly defined units and then analyzing each individual unit seems to be a more feasible method (Urbanovicius, Dikcius and Kasnauskiene, 2007). This can be achieved by implementing appropriate marketing models for specific marketing function.

2.2.1.1 Need for Marketing Models: The competitive advantage attained by a firm relies on the successful analysis and visualization of relevant data pertaining to marketing knowledge. Specifically, comprehending consumer data, which is one of the key areas of discussion in this study, is one of the critical aspects in marketing. This requires a proper formulated marketing plan and formulating efficient marketing decisions requires the implementation of appropriate models.

Consequently, a number of marketing models have been listed down in the literature into different categories for different purposes.

2.2.1.2 Categories of Models: Typically, marketing models can be classified into descriptive, predictive and normative models-

i) *Descriptive models* usually tend to describe the decision-making process (Eryigit, 2017).

ii) Predictive models help in determining or predicting future events. These models are particularly used in forecasting future events.

iii) *Normative models* are mainly known to focus on end results since ages (Massy and Savvas, 1964) (Johar and Sirgy, 1989). These models are being utilized to analyse and access various substitute strategies to devise proper techniques in order to accomplish the essential objectives.

iv) Very similar to predictive models are *measurement models* which are used in gauging the demand for a certain product or service.

Furthermore, with respect to the degree of uncertainty in the output, models can be categorized into *deterministic* and *stochastic* models (Eryigit, 2017). While deterministic models consider known parameters to determine the output of a certain decision, stochastic models help in evaluating the output by using probability metrics for unknown variables.

2.2.1.3 Need for redefined models: Kumar (2015) stated:

"Over the decades, the marketing discipline has experienced changes in terms of its dominant focus, thought and practice".

Kumar (2015) points out the below mentioned factors as contributors to the evolving changes in marketing:

- Modifications in the way data is stored and processed in the current times.
- Increasing access to and easy compilation of individual consumer data.

But an important question that arises here is that with such evolution in marketing, are the same old marketing practices still feasible in the current time? Especially, considering the incredible amount of data created by businesses and consumers today.

With the advent of emerging information technologies and high performance computing, marketers can actually integrate such technologies with the relevant models to devise new models or use entirely new approaches based on these technologies. In the recent times, analytics are

widely becoming popular in executing and monitoring marketing activities. According to Tarka and Lobinski (2013), analytics has the potential to tailor marketing and bring down the overall operational costs of a firm in the long run. It is worth noting that although the initial investment in data-driven or machine-driven tools are potentially high, it's typically a one time investment. On the contrary, the labour costs are drastically minimized with time as the costs required to train the manual workforce reduces (Acemoglu and Restrepo, 2018). Furthermore, implementation of best marketing practices involves thorough analysis of consumer data. As such, the study of consumer behaviour in marketing is necessary.

2.2.2 Consumer Behaviour in Marketing: One critical aspect of the marketing discipline is consumer behaviour. It has been noted in the literature that marketing is substantially accountable for manipulating the consumption patterns of the customers (Pantelic, Sakal and Zehetner, 2016). While customers, in a survey have acknowledged that marketing is the prime reason that has influenced them to make excessive consumption, it is crucial to examine whether conventional marketing concepts are adequate enough to anticipate consumer behaviour. Potential concerns have been advocated by consumers which indicate that continuing the same traditional marketing activities could create challenges for marketers in driving consumerism for a long-term as beliefs and behaviours of consumers vary from individual to situational levels. Moreover, with respect to different tastes of consumers, the need for differentiation has become a prerequisite in marketing to build strong rapport with the customers according to their individual preferences (Shao, 2015). Hence, identifying the preferences of the consumers under diverse circumstances ought to be a priority for marketers. To gain a lucid insight, it is vital to have an in-depth knowledge about the different factors that influence customers' purchases in different phases.

.

2.2.2.1 Key phases of Consumer Behaviour: The phase of consumer behaviour extends from the point when customers become aware of the 'want' or 'desire' of a certain product/service to the transaction of the purchase (Nguyen, Leeuw and Dullaert, 2018). It includes three important phases:

- Purchasing
- Consuming
- Discarding the product or service

It is the customer who contributes as the user, payer and buyer. Understanding consumer behaviour is complicated because the attitude of each consumer towards each of the aforementioned phases is different. According to Solomon (2009), in addition to the three phases, the study of the post purchase phase of consumer behaviour is equally vital and requires intense notice of the researchers as it signifies the level of contentment of the consumers. Whilst a consumer's repeat purchase behaviour typically indicates a higher level of satisfaction, termination of usage of a product requires critical attention of marketers as it can significantly bring down sales of a company, eventually curtailing profits.

Furthermore, Myers (2001) stated that post- purchase behaviour of consumers can facilitate appropriate feedback collection on products and eventually help in retaining the standards of the products. However, fulfilling this is a complicated process since it demands a granular analysis of consumers and their preferences, and also takes into consideration the level of engagement of the consumers in the co-creation process (Tynan and McKechnie, 2009).

Nevertheless, it is unjustified to evaluate consumer behaviour solely based on preferences. It is equally important to examine other dynamics. For instance, if a customer decides to purchase a product, how much is he willing to pay for the product? Is it within his budget? The study of consumer behaviour ought to emphasize on how customers invest their resources (time, money, effort) to make their purchases. (Prasad and Jha, 2014) states that the most common decisions include the kinds of products and services to be purchased, the required quantity, the place and time of purchase and also the mode of payment for the purchases . Since a firm's primary goal of introducing a product is to cater to the audience's needs and demands, the product or service should me strategically marketed. A strategy is all about leveraging the profitability of the firm and the frequency of purchase trends (Pahwa, Taruna and Kasliwal, 2017).

2.2.2.2 Dimensions influencing Consumer Decisions: While studying consumer behaviour models, it is important to consider the various dimensions that influences consumer decisions. As noted in the literature (T.K, 2014), the below mentioned factors highly influence consumer purchases:

- **Cultural**: Cultural factors include religious norms and beliefs and thus varies from region to region. The needs and demands of the consumers are profoundly related to their culture.

This might require the marketers to segment the targeted population according to their religions, racial tribes or geographical zones.

- **Social**: Social factors mainly include family, friends, referrals and even social status. Close friends, neighbors, colleagues and relatives are considered to be the most relevant reference groups (Brown and Oplatka, 2016). Marketers strategically target families by weighing their needs against the marketed products. For instance, for selling a household item, marketers usually target the women population during advertising. Word-of-mouth is one of the strongest factors influencing purchase decisions. According to Lau and Ng (2001), consumers tend to look for product information in the initial purchase stage by indulging in intense Word-of-mouth process. Because of this, negative word-of-mouth can directly jeopardize an organization as it is directly correlated with a customer's dissatisfaction level.

- **Psychological**: According to (T.K, 2014), the psychological dimensions that drive consumer behaviour are perception, motivation, learning, belief and attitude. However, it is critical to note that the difference in perception leads the customers with similar needs to buy different products. Thus, it is essential that marketers advertise their product in a way that can make their customers believe that the product will satisfy their exact need.

- **Personal**: Personal factors include customers' lifestyle, financial condition, employment, gender, age demographics and personality. At different stages of life, from childhood to adulthood to parenthood, consumer preferences varies. Marketers tend to have very minimal or negligible control over customers' individual factors. Nevertheless, it is crucial to take note of customers' personal factors too. While making purchase decisions, individuals mainly take into consideration their disposable income. The higher is the disposable income, the higher is the capability of the individuals to buy expensive products. However, in business, such responses from consumers lead to the production of cheaper products with inferior quality to accommodate well with customers' financial needs (Brown and Oplatka, 2016).

- **Situational:** At times, situational factors influence the buying decision too. Situational factors might include weather conditions, health conditions or even a store's environment, for instance, crowding. Furthermore, at times, the unconscious state of mind urges the customers to make purchases as stated by Lau and Ng (2005).

T.K (2014) represented the nexus between the various aforesaid dimensions and the consumer's purchase decision process by the following diagram. The figure shows that pertaining to the five different dimensions, different needs of the consumers arise, for instance, social and psychological needs. While making purchases, consumers also take into consideration certain other aspects such as price, packaging, brand, advertisement, etc. which have been depicted in the figure. All of these taken together enables a customer to make purchases.
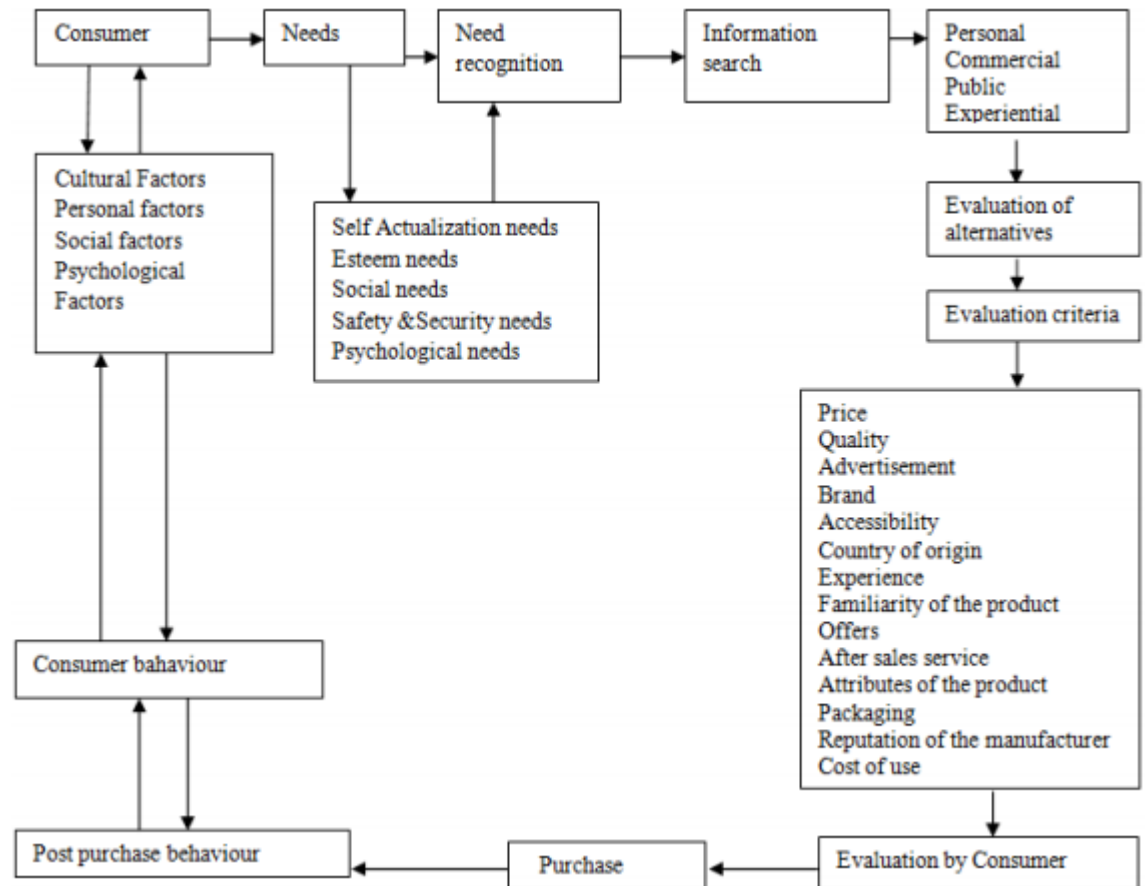
Figure 1: Consumer purchase decision framework   (T.K , 2014)

Similar to this framework, most of the traditional models have been developed by taking into account the five dimenions influencing consumer decisions which will be elaborated in the next section.

2.2.2.3 Traditional models for Consumer Behaviour Prediction: While marketing models are descriptive, predictive or normative in nature, consumer behaviour models are mainly either

descriptive or *predictive*. Moreover, consumer behaviour models are typically **stochastic** models because forecasting consumer behaviour requires working with unknown parameters in order to measure frequency of purchase. Various traditional as well as contemporary approaches have been noted in the literature for studying consumer behaviour in the past. The well-known traditional models include the economic model, learning model, psychoanalytic model and sociological model. Some of the contemporary models include the Howard-Sheth model, Engel-Kollat-Blackwell model, Solomon Model of Comparison Process, Nicosia model and Stimulus-Response model among which Engel-Kollat-Blackwell and Nicosia models seem to be the most detailed yet controversial models.

A brief overview on these two models, which are typically consumer decision-making frameworks, are provided below:

- The **Nicosia model** comprises of four primary fields (construction of Customer Attitude, Search and Purchase, Act of Purchase, Feedback). Below is a graphical illustration:
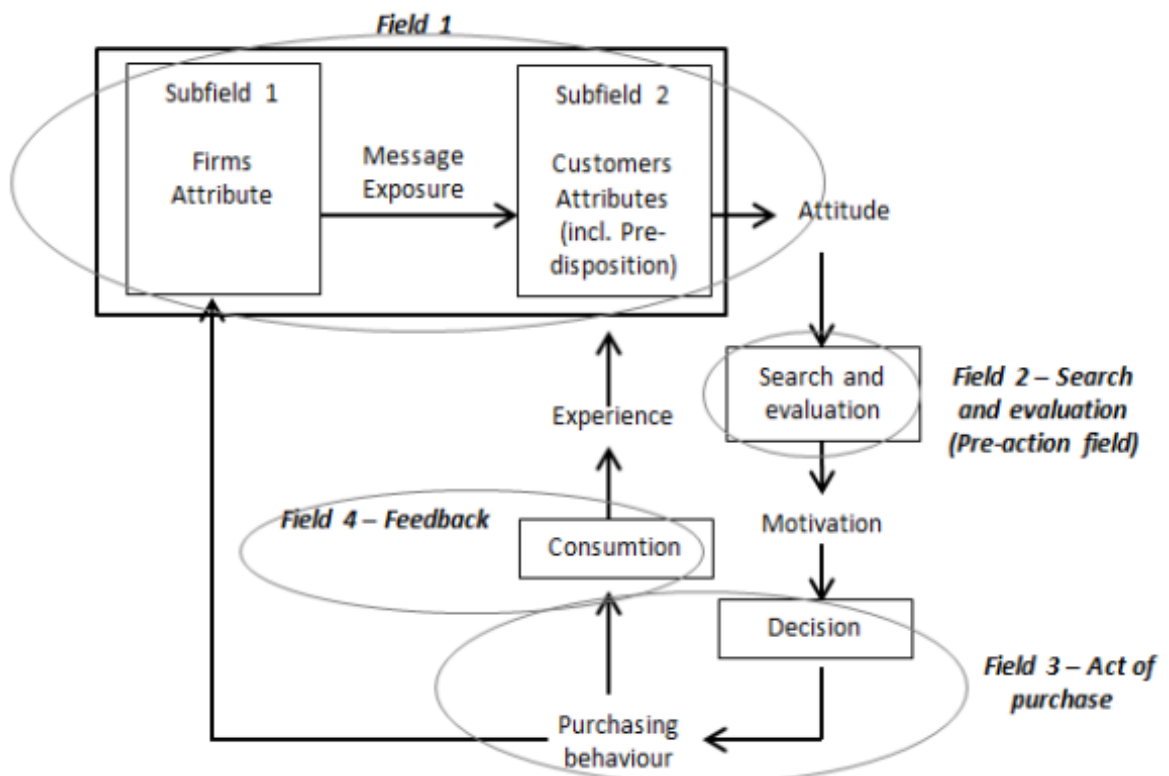
Figure 2: Nicosia Model (Viksne *et al.,* 2016)

This framework works on the principle that the values offered by an organization form the attitude of the customers (Field 1) and motivates them to thoroughly explore and evaluate (Field 2) before

making purchase decisions (Field 3). The final stage is Feedback (Field 4) and is equally important as it reflects the customers' experiences on the product's consumption.

- The ***Engel-Kollat-Blackwell*** model, developed first in 1968, is relatively complicated and considers that consumer decision process is affected by individual values, environmental impacts and situational effects (Viksne *et al.,* 2016). Below is an illustration of the model:
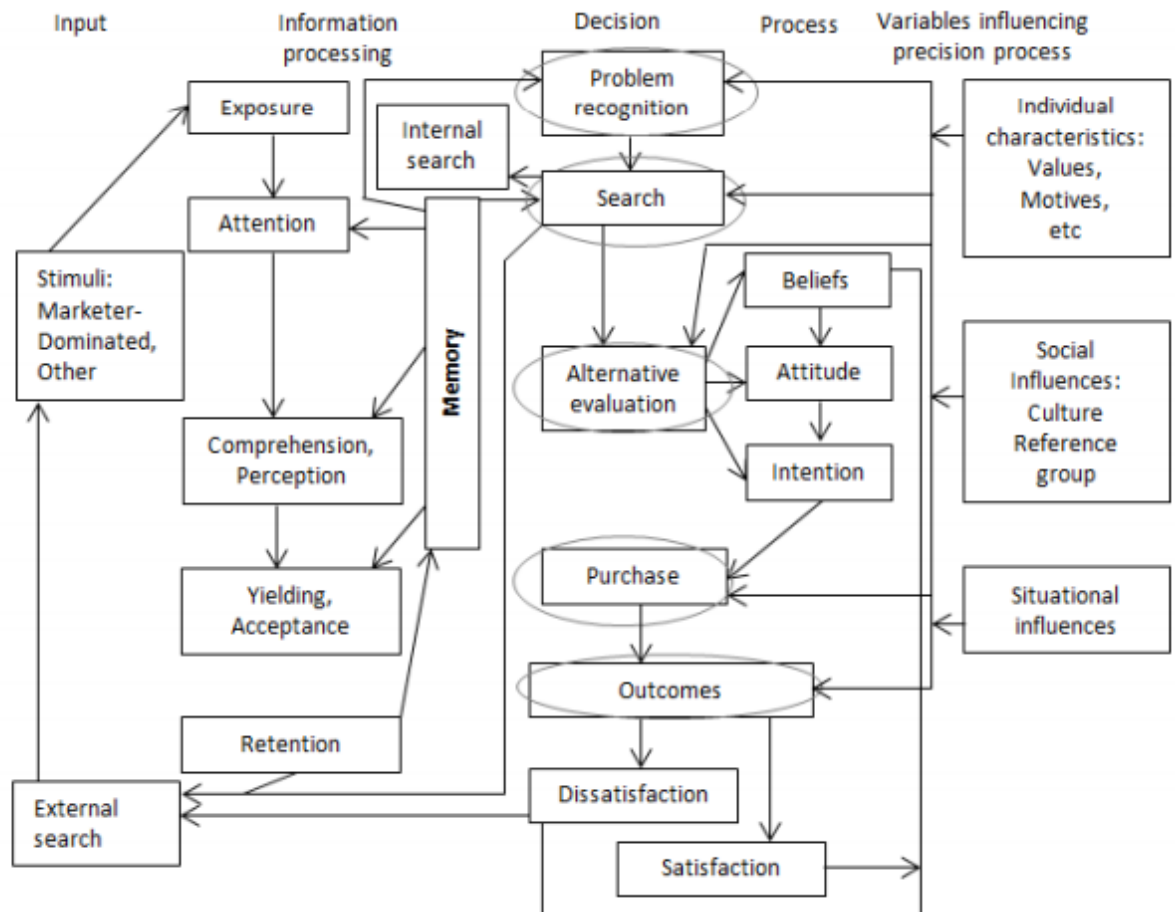
Figure 3: Engel-Kollat-Blackwell model (Viksne *et al.,* 2016)

The complexity of this model arises mainly because of the multiple fields involved in the framework. This model considers all the five dimensions influencing consumer behaviour as integral elements. While the social, cultural and situational factors are perceived to encourage customers make purchases, the psychological factors are perceived to be highly influenced by the satisfaction level of the customers as shown in the figure. In addition to this, this model anticipates that dissatisfaction enables the customers to switch to other products through external search.

Although the two discussed models are rarely used by organizations in the current time, such traditional frameworks cannot be considered as an absolute isolation from the techniques of data analytics which will be justified later in the study. Nevertheless, it is important to highlight the reasons that has minimized the use of the traditional models.

2.2.2.4 Limitations of Traditional frameworks

Foxal (1980) put forward certain limitations of these models. Firstly, the two models do not seem to be very suitable for practical applications as they tend to offer just a mechanical glimpse of individual traits of consumers. Secondly, the functional complexity of the frameworks makes it challenging to interpret the frameworks. Thirdly, the variables used for these models are not precisely defined. Such drawbacks make it tough to verify and test the models iteratively.

Furthermore, a major limitation of the Nicosia model is that it does not identify the type of consumer that this model is fit for. For instance, whether the consumer purchases the product out of habit or out of necessity? Is the customer buying the product for the first time or have there been multiple encounters between the product and the customer? According to Tordera (2013), this model is mainly suitable for studying the behaviour of new customers and does not gives a clear overview of the repurchase behaviour of the consumers which is very critical in marketing. Additionally, critics argue that the sequence of the fields in the Nicosia model is somewhat flawed. For instance, the framework assumes that customer attitude fosters motivation to make purchases. But if a customer is oblivious to marketing ads or campaigns but still makes transactions, then would the aforesaid sequence of events still hold true? Hence, the Nicosia model can be suitable as an introductory model but is not fully flexible to be used for practical purposes. Compared to this model, the Howard-Sheth model is somewhat flexible as it considers the repurchase behaviour of consumers too. Nevertheless, the integration of multiple learning concepts, stimuli-to-response processes and use of a large number of variables make the framework very complicated and time-consuming to be used in the current market context (Foxal, 1980). Although the Engel-Kollat-Blackwell model is relatively more flexible than the Nicosia and Howard-Sheth models, the sequential flow between the information processing and external search shown in the figure is very complicated and seems to be redundant and confusing. Hence, more clarity is required on the movement between the two stages. Additionally, it does not explain how the environmental factors impact the consumer behaviour.

The above analysis ultimately sheds light on the shortcomings of the traditional models and thus urges the need for the development of refined versions of the models with the help of advanced technology. Although the abundance of user information allowed marketers to benefit from individual-level transaction frameworks for marketing and making focussed decisions, the massive amount of available data can result in tremendous mess which can eventually make it challenging to derive relevant inferences from the data (Mazumdar, 2010) using traditional models. The ultimate goal of understanding consumer behaviour is to devise relevant marketing strategies in order to fulfill long term consumer satisfaction , thereby reinforcing customer retention.

Such situation instigates the usability of predicting consumer behaviour using techniques involving data analytics.

2.2.3 Data Analytics: According to France and Ghose (2018), the result of the IDG data analytics respondents indicated that the prime aim (approximately 55%) of applying analytics is to enhance 'customer relationships' which is undoubtedly the most important dimension of marketing. Predicting consumer behaviour, which is typically a direct marketing goal, is considered to be one of the key challenges for analytics. Hand (1998) acknowledged the data mining technique to be very similar to the traditional statistical approach. Moreover, increasing use of expert, intelligent systems in the field of marketing is noteworthy which has eased the decision-making process for the marketers. In the past, marketers relied on theoretical concepts of consumer behaviour to identify their customers' needs to a certain extent.

But on comparing the recent technologies of data analytics with the traditional marketing models, it can be commented that data collected through the traditional approach is bulky and difficult to maintain. In today's digital world, the amount of data that is being created yearly is rising by tenfold. Hence, implementing analytics in marketing seem to be a more relevant and convenient method.

2.2.3.1 Predictive data analytics: According to Boonsiritomachai *et al.* (2016), predictive data analytics has immense potential to optimize traditional work processes in a company, to comprehend consumer behaviour and predict unanticipated events including problems before they actually occur. In-depth evaluation of consumer behaviour leverages customer satisfaction and

enhances customer loyalty and retention as stated by (Devi and Rajagopalan, 2012). However, several controversial statements have been put forward while comparing the computing technologies to the traditional techniques. In the traditional approach, researchers constructed hypothesis based on theoretical concepts before actually performing tests and validations which indicates that traditional data analysis approaches were driven by assumptions. On the contrary, the models and frameworks used in the current times are data driven and do not rely on mere theories. Below is a simple illustration of the comparison:

Figure 1.
Traditional research approach

Theory ➡ Develop Hypotheses ➡ Test Hypotheses
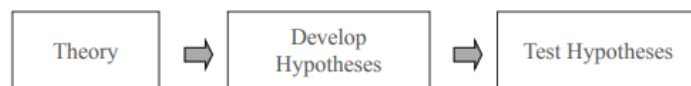
Figure 2.
Data mining and predictive analytics

Data ➡ Relationships ➡ Develop Hypotheses ➡ Model Building Test Hypotheses ➡ Model Validation
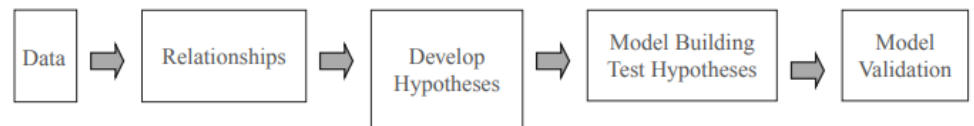
Figure 4: Traditional approach vs Analytics   (France and Ghose, 2018)

One critical question that originates here is what makes the data-driven technique comparatively more effective? Does it potentially eliminate manual work?

Figure 4 shows a comparison between the base concepts on which traditional models work and that on which data-driven models function. As depicted in the figure, the traditional approach requires the researcher to be involved in all the stages of the research which is very time-consuming when analyzing large scale of data. However, data-driven techniques require the involvement of the researcher only until a certain phase. The initial data collection phase including identifying relationships among the collected data demands active inputs from the researcher. However, once the data is fed into a model which is the fourth stage in the model, the human intervention tremendously minimizes which reduces human errors in the process, making the approach more effective and error-free.

Data analytics is complex and demands intense exploration of data. Studies conclude that the amount of data in the data warehouses are so massive that it is next to impossible to fully analyze the data and their interrelationships through manual processes (Hair Jr, 2007). This is the primary reason that has stimulated the application of predictive analytics for study of large scale of data.

According to (Hair Jr, 2007), the advancement in information technology has minimized the cost of data collection processes in the current era. On a side note, the improvement in the user interfaces have enhanced the accessibility to software or applications that performs predictive data analytics.

To obtain a detailed understanding of how predictive data analytics works, its different stages as well as the working principle of the techniques will be studied.

2.2.3.2 Phases in Data Analytics: The first crucial phase of data analytics is preparing the data which usually involves cleaning and transforming the data. Data needs to be sorted and transformed into certain format before they can be used. This ideally requires the use of specific algorithms to optimize the data. Transforming the data involves fine tuning the results by precisely and efficiently performing data mining technique (Maingi, 2015). The data preparation stage is followed by building of predictive models. Developing such models is quite labor intensive.. Moreover, to create a successfully working model, it is required to continuously train, test and validate the model. This process requires a number of iterations before finally deploying the model. The final stage in data analytics is maintaining the deployed models. Managing these models leverages performance, enabling the reuse of models and thereby helps in reducing the costs of the organization.

However, the first method that is required to be performed before conducting predictive data analytics is the data mining technique.

2.2.3.3 Data Mining: Most of the times, the terms predictive analytics and data mining are interchangeably used. However, the two approaches are independent, yet both the techniques are closely related and jointly help in determining consumer behaviour (Hair Jr, 2007). The four essential aspects that an end-to-end data mining technique requires are a standardized method, data warehouse, information technology and expertise (Pahwa, Taruna and Kasliwal, 2017). Data mining techniques help to recognize patterns in the data sets accumulated from consumers and detect potential connections among the collected data after which predictive analytics analyzes the derived relationships to anticipate future phenomenon and trends. However, depending on how much relevant data is available, there are 2 major approaches to data mining-

- **Supervised Learning:** In this type of technique, a range of historical data is utilized and used for developing predictive models based on the previous data.
- **Unsupervised Learning:** This type of technique is used when results from the past are not used, instead predictive models are created on the basis of unknown parameters. It ideally makes use of descriptive statistics to identify trends and patterns.

Comparing the two kinds of learning, it can be understood that the process of predicting consumer behaviour can be achieved through *unsupervised learning*. However, irrespective of the kind of learning used, the final outcome is an analytic framework. According to a survey that considered the feedback from 166 participants, nearly 66% of the participants confessed that predictive analytics delivers high or very high business value and 27% of them reported that it delivered moderate business value. Merely 4% of the participants acknowledged that predictive analytics delivered low or very low value (Boonsiritomachai *et al.*, 2016). Below is the detailed statistics of the survey:
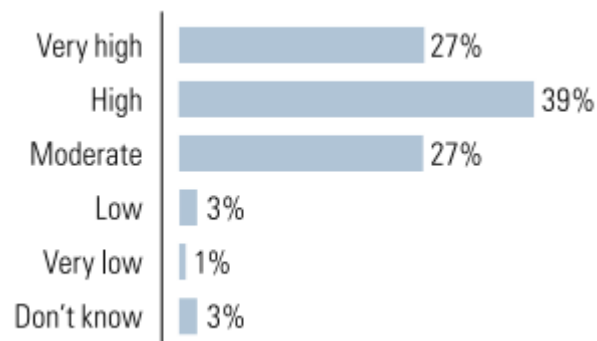
Figure 5: Contribution of analytics in enhancing business value (Boonsiritomachai *et al*., 2016)

While the figure shows the significance of predictive data analytics for businesses, a critical focus is required to explore the techniques which are the most effective in making predictions. Therefore, below is a detailed description of the different kinds of techniques.

2.2.3.4 Types of Data Mining Techniques: The most frequently used data mining techniques are clustering, classification and association as stated by (Patil et al., 2014). Although there is a huge distinction between the working principles of data mining techniques and that of the traditional models, the ground concept of data mining techniques seems to have stemmed from the traditional

frameworks. Especially, the Engel-Kollat-Blackwell model mentioned in section **2.2.2.3** has plausibly influenced the working of the clustering technique. The patterns detected by the data mining techniques are utilized to anticipate individual behaviour of customers with sharp precision. These techniques work by using statistical algorithms to detect and correlate patterns in data (Patel, Karvekar and Mehta, 2014). In technical terms, data mining techniques find correlations among all the fields contained in the vast relational databases of the company (Raorane and Kulkarni, 2011). All the information such as the time, day and place of purchase of the product, the purchased quantity and even the frequency of purchase are stored in the organization's database and each record is considered as a byte of data. Data mining techniques help to derive relationships among the various dimensions in the database and categorize it accordingly. According to Hsieh and Chu (2009), data mining finds extensive uses in market segmentation, customer categorization and risk evaluation.

A list of commonly used data mining techniques will be discussed in the following section:

2.2.3.4.1 Clustering Approach: The clustering data mining technique is used to categorize a set of entities with similar characteristics into one cluster. This technique basically follows the concept of groupism. Once similar activities are put into the same group, the reoccurrence rate of each activity is extracted with the help of complex programming algorithms.

"*Clustering is the task of segmenting a heterogeneous population into a number of more homogenous clusters* " (Devi and Rajagopalan, 2012). Categorizing customers who have purchased a specific product or service can help marketers target similar customers who have not purchased the product. This concept, in marketing, is termed as cross-selling. Typically, this technique can be utilized by marketers to perform customer segmentation. As stated by Maingi (2015), clustering technique applied on sales data can ease the task of predicting consumers' buying habits by collating related data points into specific clusters. The clustering process comprises of three distinct phases:

- Forming the clustering algorithm by selecting specific characteristics, i.e, by grouping similar trends or patterns.
- Validating the clusters to confirm their standards and precision.
- Integration of the cluster sets to interpret the final outcomes.

The ground concept or algorithm on which the clustering technique based is called the *K-Nearest Neighbors* algorithm. K-Nearest Neighbors (KNN) is a distance-based algorithm. It is a non-

parametric method which implies that it does not rely on assumptions and predominantly used to solve pattern recognition problems. It is worth noting that the working principle of K-NN algorithm has a strong resemblance to that of the traditional Engel-Kollat-Blackwell model. While comparing the two frameworks, it can be observed that each of the rectangular fields in the flowchart in Figure 5 can be defined as an independent record in the columns of a database, each assigned to a separate variable. To get a detailed understanding of how this algorithm works, let's consider a set of database records. Let's assume that there are 3 rows, each row containing details about 1 customer which implies that in total, there are records of 3 customers. Suppose there are 2 columns- one recording the frequency of purchase of each customer and the other saving the transaction details (amount of money spent). So, how does K-NN work with these records?

K-NN considers each of these data fields as a data point and assigns a particular value to the first record, say k=1. When it seams through the successive records, it picks up fields with similar features. For instance, if the frequency of purchase of Customers 1 and 3 are thrice a month while that of Customer 2 is just once in a month, it considers Customers 1 and 3 as the nearest neighbors and clubs them into the same clusters. This algorithm is the ground rule on which the clustering technique works and the concept can be represented by a simple flowchart:
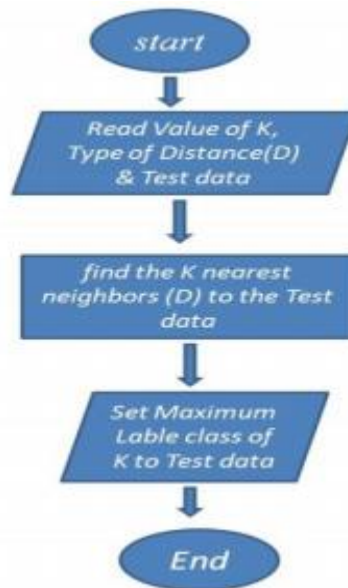
Figure 6: K-NN Algorithm Flowchart (Bazmara and Movahed, 2013)

2.2.3.4.2 Classification Approach: A close resemblance of the aforementioned clustering method is the classification technique. Like the concept of clusters in the clustering method, the

classification method classifies the database records into specific classes based on certain criteria. ***Bayesian classifiers*** are considered to be the most reliable method for data prediction (Papić-Blagojević, 2011). Bayesian networks, which are also termed as 'belief networks' are an excellent source to conduct unsupervised learning. Bayesian networks consist of graphical structures which are very suitable for comparing prior knowledge to observed data. It basically works on the concept of probability which is why it can be used for reasoning under uncertainty. This method has been noted as the most widely used technique since 2010 for predicting consumer behaviour (Ibukun.T *et al.,* 2016). As stated by Hsieh and Chu (2009), Bayesian rules were first implemented to test repeat buying behaviour in direct marketing. The positive point about this technique is that Bayesian rules rely on a belief that the probability of each individual trait stands independent from the probability value of all the other traits. Moreover, Bayesian networks are very useful for huge databases, especially because of its speed in training models.

Among all the methods, clustering and classification have been identified as the most frequently and jointly used data mining techniques. However, the ground logic behind the two techniques somewhat vary. Contrary to the clustering technique, Bayesian network algorithm is based on the concept of mathematical probability.

According to (Keller, Gray and Givens, 1985), "*Bayes decision theory gives optimal error rates*". However, in certain situations, the collected data is limited or information is incomplete. Such conditions instigate the use of a distance-based algorithm known as K-Nearest Neighbors (KNN) algorithm.

On the other hand, the probability distribution theorem on which ***Bayesian*** rule works is defined as:

$P(A/B) = [P(B/A) P(A)] / P(B)$, for two events A and B.

P(A/B) indicates the occurrence of event A provided event B occurred and P(B/A) indicates the occurrence of event B provided event A occurred. From this formula, the probability ratio can be calculated which would give the probability of occurrence of the events.

The K-NN and Bayesian rules can be used together to derive the final prediction outcome. Irrespective of the increasing use of Bayes theorem, the K-NN algorithm seems to be a better predictive algorithm. This is mainly because Bayesian are a set of 'belief' networks which implies that it is somewhat based on assumptions. In fact, Gelman (2008) argued that non-Bayesian rules give more rational outputs as they are not based on assumptions. Moreover, rather than

concentrating on how to retrieve relevant information from data, Bayesian method directly moves to probabilistic computing. Nevertheless, to gain a better insight into this, further studies are necessary.

2.2.3.4.3 Association Approach: Often, the association technique is clubbed with the clustering technique to help in prediction of consumer data. It aims at establishing connections between entities which already exist together in a certain database (Raorane and Kulkarni, 2011). For instance, this technique can foretell a future trend by determining a link between two entities within the same cluster. Association technique is progressively being a part of the current marketing areas, especially in performing market basket or affinity analysis. The market basket analysis analyzes the purchasing behaviour of customers by identifying links between items that customers place in their shopping basket and usually tend to buy together. For instance, if a customer buys bread and cheese together, the frequency of purchase of the two products together can be traced to predict the probability of future purchases. Accordingly, it can be decided which items in the store should be placed next to each other to leverage sales and gain better profits. Hence, the simplest way by which the association technique can be represented is:

<p align="center">"If the antecedent, then the consequent"</p>

According to Maingi (2015), association mining technique combined with clustering has the potential to identify the commonalities between the consumer groups and their choices for products or services. Suchacka and Chodak (2017) carried out a research on an online bookstore to make some predictions on the activities of the customers. They demonstrated how association data mining techniques were used to predict customers' future buying probability on the basis of the bookstore's data. This technique is usually considered to be very convenient as it is comparatively simple to use, especially because it avoids the use of complicated models and instead relies on continuous data collection and statistical analysis which makes it less time-consuming and very convenient to be used for real-time scenarios.

Once relevant data is accumulated with the help of data mining techniques, it is beneficial to portray the data in a presentable form and hence data visualization comes into picture.

2.2.3.5 Data Visualization: As the saying goes, '***A picture is worth a thousand words***'. Hence, the final stage involves visualizing the mined data. This phase is equally important as data

presented in a haphazard manner is difficult to be interpreted. It is crucial to portray data in a graphical or legible format to make it visually presentable. The notion of using pictorial formats to represent data has been in use since ages. However, with the advancement of computing technologies, several tools have been developed to perform data visualization in the current times. Microsoft Power BI and Tableau are two most widely used tools in the current time. It is considered that data visualization has an extensive range of uses among which predicting consumer behaviour is one of the most vital ones. Gavett in (2014) stated in his paper how data visualization played a major role in extracting the information of sales in retail stores including both offline as well as online stores. Retailers, through this technique, were able to identify which divisions of the stores were the busiest. In an interview conducted by Gavett during his research, a marketing professional acknowledged that comprehending consumer behaviour and meeting consumer needs is still a challenge. Nevertheless, he accepted that the advent of data analytics and visualization tools have eased the task of marketers in predicting consumer behaviour. According to Borkin *et al.* (2013), it is critical to ascertain the elements that make visualizations intrinsically more memorable. The most notable strength about data visualization is that in addition to representing data graphically, this technique can assist in screening out irrelevant information, digging into minute details and even in altering the display of data. As commented by Few (2007), no traditional methods can match the data visualization approach in terms of the accuracy of the final outcome.

2.2.3.6 Relevance to marketing and data professionals: From the above discussion, it can be observed that data analytics can be beneficial for organizations if implemented in the appropriate way. Previous studies have also reported its benefits for marketing and data professionals. .Jain and Yadav (2017) have highlighted how information technologies are aiding marketers to communicate with their prospective customers, facilitating strong customer relationships. Dupre (2013) made a remarkable comment stating "*...if marketers let data and analytics take the lead, they can use customer insights to attract the right prospects, boost overall engagement, and increase marketing ROI*". Moreover, it has been cited that marketers are using these techniques to track their customers' behaviour and stay in sync with them. On the other hand, Zhong *et al.* (2004) has reported that clustering techniques are commonly used by the data engineers and analysts because of the high accuracy rate that it yields. This view is supported by Chengyong (2019) as

well who further adds that Bayesian network is widely being used by data engineers for making predictions from incomplete or unavailable information.

However, the relevant literature supporting these viewpoints seem to be limited. Hence, further analysis is necessary.

2.3 Literature Conclusion: The presented literature reveals how significant is the study of consumer behaviour for an organization. Although quite a few traditional theoretical paradigms for anticipating consumer behaviour have been discussed in the literature, it appears that working with such theoretical models requires a lot of human involvement or manual efforts. Especially, considering the large scale of data that is getting multiplied at every instance, manual analysis of data seems to be tedious and time-consuming. Additionally, referring to the literature, there is a clear indication that a lot has changed in the way consumer data is being collected today as well as the ways by which marketing practices are implemented today seem to differ from those applied in the past, viz. through TV ads, newspapers, word-of-mouth and pamphlets.

Therefore, the researcher perceives a critical need of identifying a more refined, specifically some technology-induced technique that can help predict consumers' purchase behaviour with a higher accuracy within a limited timeframe. To gain a lucid insight, the literature highlights some of the data mining techniques that can aid in making predictions about consumer data. Consequently, after reviewing the relevant papers and journal articles, the researcher believes that incorporating such techniques pertaining to data analytics can at least minimize, if not eradicate, the constraints (time, accuracy) encountered while predicting consumer behaviour using theoretical frameworks. Nevertheless, it is equally important to consider that dealing with such computing technologies demands a strong level of expertise as well as potent analytical skills. This view has been supported by Chung and Gray (1999) too who stated that accurate implementation of data mining techniques demand subject matter proficiency and adept analysis.

Furthermore, after examining all the data mining techniques discussed in the literature, the researcher anticipates the clustering data mining method and the Bayesian network algorithm to be the most prominent for making predictions about data. This is predominantly because of the reliability and accuracy of these techniques that have already been reported in the literature. The researcher intends to carry out further studies taking these two techniques as a baseline. However, a review of the literature also depicts that the clustering technique, which is based on the concepts

of K-NN algorithm, might be challenging to implement if the logic is unclear to the user and this might be very risky as improper handling of data can cause severe disruption. Besides, it's important to investigate the consistency of the results derived from these techniques to verify their overall viability in the field of predicting consumer behaviour. This requires an iterative process of training and validating these techniques in diverse situations which requires relevant expertise in the field.

Consumer data is sensitive as it involves personal information and handling raw inputs or data is not an easy task. It is not surprising that data-related issues have always been a concern, especially while dealing with real-time data. To get a clearer picture, the researcher feels that this study require further insights from individuals who are actually experienced at working with data-driven techniques or/and theoretical frameworks so that a clear derivation can be drawn by comparing the effectiveness of both the approaches. Overall, the researcher anticipates from the literature that information or computing technologies have been a key driver in bringing out these transformations. Nevertheless, this assumption seems to be inept without further research into how or in what ways such technologies are helping. Consequently, the researcher understands the need of examining whether working with such technologies have a challenging side too.

From the further analysis of this study, the researcher intends to mitigate the gaps in the literature by attaining in-depth insights on the effectiveness of Data Analytics involving clustering and Bayesian network techniques for predicting consumers' future purchase behaviour.

This chapter has presented an overview of the main findings from the previous literature pertaining to the research topic. The following chapter will outline this study's aim and objectives and eventually define the research questions.

CHAPTER III: Research Question

According to Parasuraman et al. (1991), understanding and fulfilling the expectations of the customers are prerequisites for delivering superior service to the customers. One of the key challenges that marketers come cross is anticipating exactly what customers expect. A majority of the firms fail to retain customers as they lack expertise in implementing a consistent approach for satisfying customer needs. Hence, understanding customers' expectations is a critical need.

3.1 Research Aim: Considering the above problem scenario, this study attempts to put forward an information technology-based idea that can help marketers follow a relatively reliable approach to mitigate the research problem. The research will aim to extract more precise and up-to-date findings from marketing professionals and data professionals from diverse parts of the world to investigate the research topic.

3.2 Research Questions: The research topic will be explored by investigating the answers to the following research questions:

i) How is data analytics contributing in anticipating consumer behaviour?

ii) What makes the data-driven or machine-driven tools comparatively more effective to study consumer behaviour as compared to the traditional theories and frameworks?

iii) Which data mining techniques are most extensively being used to interpret and recognize patterns among data? And why?

3.3 Research Objectives: To fulfill the aim of the research, the following objectives will be addressed:

- ❏ To examine whether marketers acknowledge data analytics as a strong tool for predicting consumer behaviour.
- ❏ To investigate whether similar data mining approaches are used across different parts of the world to study consumer behaviour.
- ❏ To determine if data-driven techniques are considered to be more feasible than traditional frameworks and models.

This chapter presented the research questions and clearly defined the aim and objectives of the study. The following chapter explains the overall methods adopted to answer the research questions.

4.1 Introduction: The prime focus of this chapter is to describe the procedures of accomplishing the aim and objectives of the research defined in the previous chapter.

Saunders, Lewis and Thornhill (2015) put forward "The Research Onion" to demonstrate the various stages of Research Methodology. The significance of this section has been marked in several literatures. Blaikie (2000) commented that the two most critical aspects of a research are research design and strategy. These two components are believed to help in successful completion of the research through logical planning and structure. This chapter will justify why a certain design, philosophy or approach has been selected. Considering the Research Onion as the primary point of reference, this chapter would further describe the research philosophy, research approach and eventually delineate the sampling techniques and methods of data collection and data analysis. Ethical considerations for this research will also be outlined from the viewpoint of ethical standpoint.
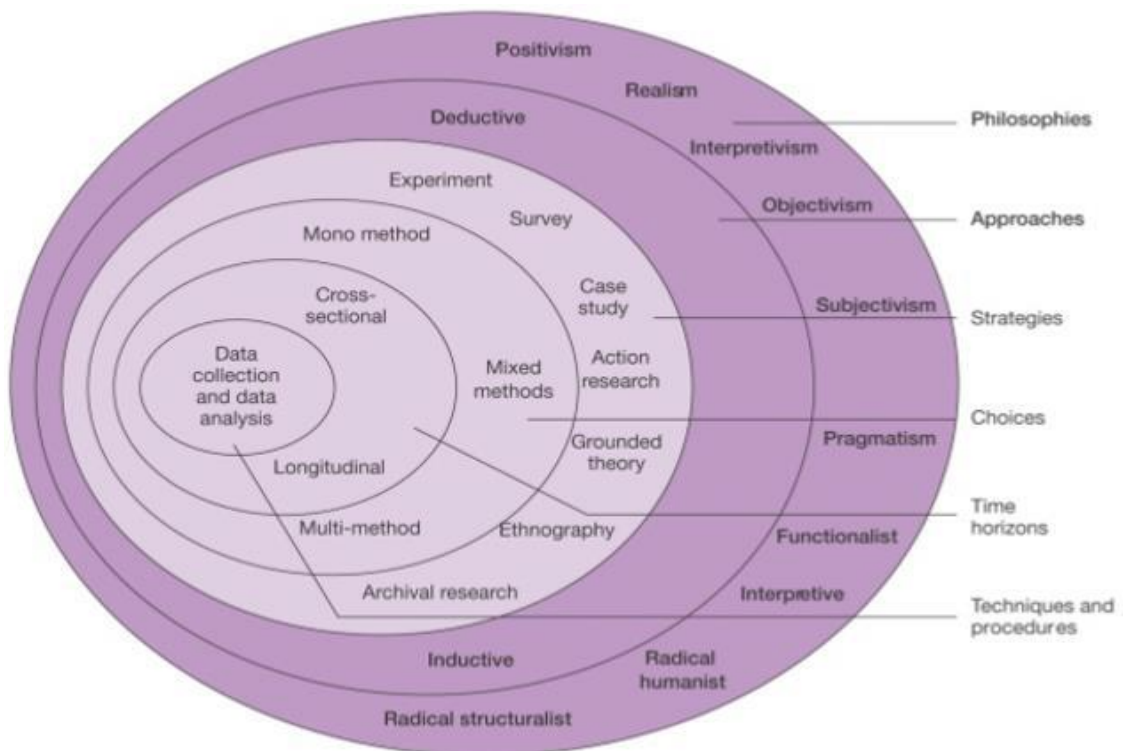
Figure 7: Research Onion (Saunders *et al.*, 2015)

4.2 Research Philosophy: As stated by Saunders et al. (2015) in the 'Research Onion', the notion of research philosophy, which is the external most layer of the Research Onion, typically refers to the way the world is viewed in order to interpret knowledge on a particular matter. It is actually a systematic way of examining diverse trends and patterns to derive a conclusion by maximizing the knowledge on a certain topic. The researcher mainly has four diverse types of research philosophies to choose from, namely- Positivism, Realism, Pragmatism and Interpretivism which are explained below:

- **Positivism**: The philosophy of 'positivism' can be associated with attaining knowledge through the implementation of scientific procedures to study real phenomenon. As stated by (Mackenzie and Knipe, 2006), this philosophical approach is most suitable for quantitative-based studies, specifically because it stresses more on the measurement of the attained knowledge through experiments, observations and logical proofs.

- **Realism:** This philosophical approach is based on the perception of individuals and makes use of the senses to interpret knowledge. Robson (2011) commented that the realism philosophy is mostly suitable for evaluating social phenomena involving a mix of qualitative and quantitative methods.

- **Pragmatism:** As stated by Lund (2012), this philosophical approach is somewhat similar to the 'realism' philosophy in a way that both are considered to be useful for studying mixed methods. It is, in fact, based on more practical experiences as compared to realism.

- **Interpretivism:** In contrast to the three philosophies, the interpretivism philosophy is quite subjective in nature as stated by Bryman and Bell (2015). In this approach, the participants of the study comprehend the set of questions put forward by the researcher and interpret in their own, unique way to derive the answers. Furthermore, Lin (1998) stated Interpretivism as a subjective approach which can ease the data collection process in unstructured or semi-structured form.

Associating with the research philosophy, Saunders et al. (2015) put forward two types of reasoning - *inductive* vs *deductive* which can be explained by the below figure:
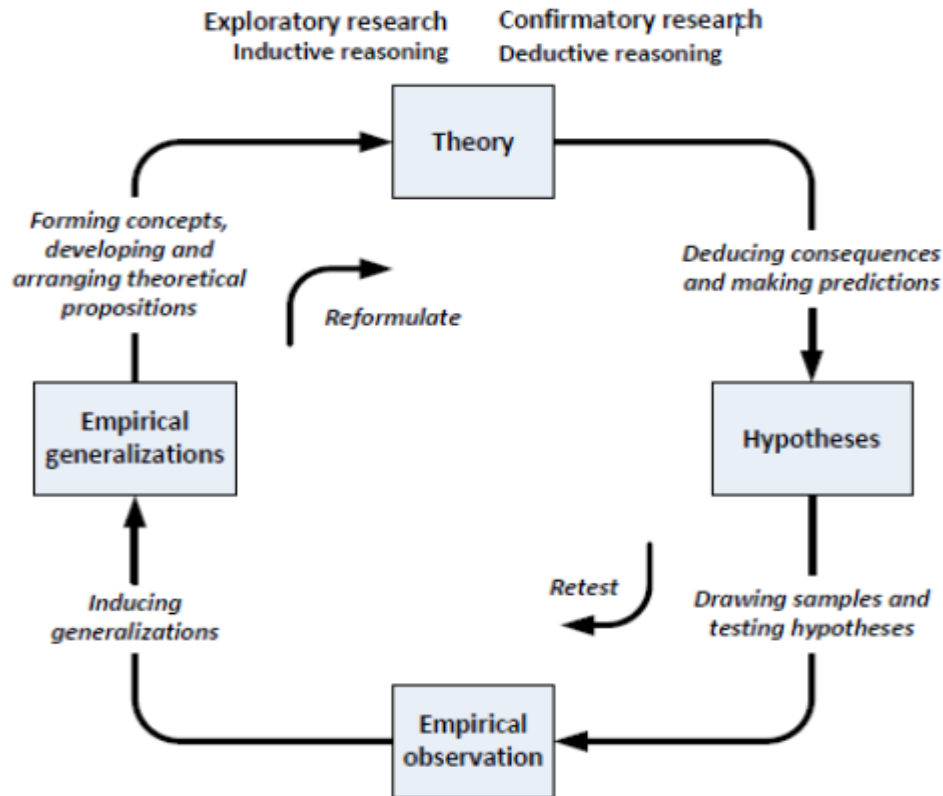
Figure 8: Wallace's Wheel of Science (Eikebrokk and Busch, 2016)

This figure, also termed as '**Wallace's 'Wheel of Science"** shows that the cycle from theory to empirical observation by testing hypotheses is achieved through deductive reasoning, whereas the inductive reasoning can be depicted by the other half cycle which takes the opposite approach of drawing generalizations from observation and then forming a notion based on that (Eikebrokk and Busch, 2016).

To link it with the philosophy, Saunders et al. (2015) stated that while the deductive reasoning can be linked with 'positivism' , inductive reasoning can be associated with interpretivism. According to Lin (1998), quantitative methods use deductive reasoning as it involves evaluating data using statistical tools. On the contrary, the qualitative method which usually takes into account the inferences from personal experiences and situations uses inductive reasoning.

4.2.1 Justification: The current research topic is subjective in nature and intends to further investigate into it by taking into considerations the beliefs and understanding of the participants from their personal experiences. By comparing all the four types of philosophies, the researcher

believes that the first three philosophies are either based on scientific or on practical methods which make it infeasible for the current research topic. On the contrary, the features of the interpretivism philosophy seems much relevant and hence is chosen for the current study.

Additionally, since this study plans to answer the research questions by drawing inferences from the participants' personal experiences and situations rather than testing hypotheses, clearly the inductive reasoning is chosen. Furthermore, since this study has chosen the interpretivism philosophy, the researcher believes that the inductive reasoning will be the most suitable fit with the research philosophy. Nevertheless, it is worth noting that proper care has to be taken during the data collection phase as few studies have stated that inductive approach at times tend to formulate biased opinions or generalizations (Bryman and Bell, 2015)

4.3 Research Method: There are 3 types of methods which the researcher can choose from- *Qualitative*, *Quantitative* and *Mixed.* This choice should ideally be in line with the research philosophy and reasoning. Alshurideh (2014) performed a similar study on customer repurchase behaviour using qualitative analysis to study the UK Mobile Phone Market. Nevertheless, it is crucial to make a comparison between the two methods to choose the most appropriate one. According to Quinlan (2013), qualitative method mainly follows subjective experiences. Unlike the quantitative method, this method involves the researcher as a key player who becomes subjectively immersed in the data collection procedure. Additionally, Bryman (2008) stated that qualitative method stresses predominantly on words rather than quantification.

Wahyuni (2012) stated that qualitative method is suitable when a study involves interpretations from participants' thorough understanding of the research problem. Contrary to this, quantitative method often relies on limited information from the participants and makes use of numerical measures and statistical tools to derive a conclusion (Zyphur and Pierides, 2017). Additionally, the qualitative method offers a flexible approach of data collection as well as to evaluate an extensive set of real-life scenarios (Miles and Huberman, 1994).

Considering the aforementioned arguments and the subjective nature of this study, the researcher opts for the qualitative method as this research intends to collect and interpret the viewpoints of the participants to understand the research problem. In addition to this, the researcher is able to link the qualitative method with the philosophy and reasoning selected for this study.

4.4 Research Strategy: Typically, there are 5 types of research strategies - case study, ethnography, narrative research, action research and grounded theory as stated by Creswell (2007). However, since this study is based on qualitative methodology, this segment will predominantly emphasize on those strategies that are usually used in qualitative studies (Saunders *et al*, 2015). Among the 5 strategies, case study, ethnography and narrative research are the most frequently used strategies for qualitative studies. However, ethnography is used when the study involves a particular cultural or ethnic group which is not feasible for this particular study. While the case study strategy is somewhat relevant to this study, it requires thorough, pragmatic analysis of the findings (single or multiple cases) which might be challenging to accomplish within the limited time period of the study that is outlined in the next section.

Therefore, the strategy of narrative research seems to be most suitable for this study. According to Creswell (2007), a narrative research is considered as a more suitable technique as it concentrates on a single individual which enables the researcher to interpret information from the individual's personal experiences and present the collected data in the form of stories. The researcher intends to portray the data received from respondents' in the form of meaningful stories so that it is consistent with the literature review as well as contributes to answering the research questions.

4.5 Research Design: Designs are an integral part of the research methodology since it helps in setting a proper plan as to how the research will be completed. This layer of the Research Onion is considered to be the most vital among all the layers as it specifically lays out a systematic blueprint for accomplishing the aim and objectives of the study. According to Yin (2014), this layer plays a major role in retaining the overall standards and consistency of the end-to-end process of accomplishing the study .

The main components of a Research Design are explained below:

4.5.1 Sample Population: This study targets to collect data from Marketing professionals who have experience in using Analytics-based applications and from Data Analysts/Engineers who have experience in developing Analytics-based applications for use in the Marketing divisions of organizations. All the selected participants will have a minimum of 2 years of work experience in the IT industry working in atleast one of the two mentioned roles.

A majority of the sample population will be from India. The rest of the participants will be from Japan, Italy, Ireland, USA, Canada and Germany. Due to the limited timeframe, the researcher was able to contact participants from only these 7 countries.

4.5.2 Sampling Techniques: Sampling techniques help in analyzing samples or populations from which data is to be collected. Probability sampling and non-probability sampling are the two different kinds of sampling techniques (Bryman and Bell, 2015). Etikan, Musa and Alkassim (2016) described probability sampling technique as where - "*every participant has an equal probability of being selected*". On the other hand, non-probability sampling techniques are considered to be suitable for subjective studies as they can be implemented quickly. The type of non-probability sampling technique that will be used for this study is ***purposive sampling***. The use of this technique has been quite frequently observed in qualitative studies, especially when it requires the researcher to get in touch with people who can provide meaningful insights from their expertise and knowledge on a specific area to answer the questions (Etikan, Musa and Alkassim, 2016).

Since participants for this study will be chosen specifically based on the two criteria mentioned in section 4.5.1 considering their proficiency and knowledge on the research topic, the purposive non-probability sampling technique is chosen for this study.

4.5.3 Sample Size: Although it is quite challenging to specify the exact number of sample size in qualitative studies, this study will attempt to gather data from 15 participants from different parts of the world. The prime aim of targeting sample population from different countries is to investigate whether the same types of data analytics-based techniques are being used in consumer behaviour prediction worldwide.

4.5.4 Data Collection Procedures: For the completion of this subjective study, the data will be compiled from multiple resources. This procedure of collecting data from more than one resources, usually termed as 'Data Triangulation' method, is considered to be most effective by Yin (2014). Patton (2001) too supported this view stating that it aids in checking the reliability of the findings. Furthermore, as stated by Mathison (1988), triangulation procedure helps in establishing logical propositions by controlling 'bias' in qualitative studies.

4.5.4.1 Pilot study: Initially, a pilot study was conducted with 3 participants with a set of interview questions to check if any adjustments will be required in the actual interviews. According to

Saunders *et al*. (2009), the significance of pilot studies is that it gives a rough overview of how the actual interview will appear to be and the kind of issues that might be encountered during the actual interview.

The researcher conducted the pilot interview to test whether the interview questions were comprehensible to the participants and to estimate how much time is required by the participants to answer all the interview questions. One of the participants of the pilot study is a Marketing Consultant while the other two are Data Engineers. The pilot interview was conducted over the telephone. Although the interviews were not recorded, the responses obtained from the 3 participants were in line with the current research topic which encouraged the researcher proceed with the actual interview with the same set of questions.

4.5.4.2 Interviews: This study will adopt semi-structured interviews to collect data from the participants. Parker (2003) stated that data collected through interviews enables the researcher to gather the knowledge transferred through conversation with industry professionals. Rather than following a rigid interview structure, semi-structured interviews help in comprehending the real-life situations faced by the participants without deviating away from the main research topic (Boeiji, 2010).

Before conducting the actual interview, participants will be informed about the research topic and mainly obtain an initial consent. The communication will be either through telephonic conversation, email or LinkedIn rather than face-to-face meetings. This is primarily due to *location constraints*. As this study will attempt to collect data from participants from different countries, greater flexibility will be provided to the participants to make them feel comfortable while answering the interview questions. The set of interview questions will be sent out to most of the participants in the form of open-ended questionnaires. However, the responses of the participants and their identity will remain anonymous. The positive side of open-ended questions is that it gives ample flexibility to the respondents to answer in their own way without having to choose from any particular set of responses (Bryman and Bell, 2015). However, a drawback of this approach is that since the selected sample size is relatively small, sometimes contradictory responses might be obtained which can make it challenging for the researcher to derive a specific conclusion.

In addition to this, 4 participants agreed to share their responses through audio recordings. The researcher recorded the interview questions and forwarded the audio clips via email to the participants. The participants responded back via email, sharing their responses through audio

recordings. However, before sending out the audio recordings of the interview questions, an informal call was made to each of the 4 participants to introduce them to the research topic as well to discuss the consent form. The participants were also given a chance to ask any doubts/concerns regarding the interview process.

4.5.5 Data Analysis: Data analysis provides meaningful insights to the assembled raw data through an iterative process of observation and drawing illations. Wahyuni (2012) stated that data can be analyzed by following a systematic procedure comprising of - Data Storage, transcription of interviews and filtering out irrelevant data. Furthermore, Saunders *et al*. (2016) put forward that in order to derive comprehensible insights from the collected data, a meaningful connection between the data collection and data analysis methods should be drawn in a way that appropriate themes or patterns can be formulated categorically. According to Kvale and Brinkmann (2009), the very first step while transcribing the interviews will be to filter out the irrelevant information. This will be followed by identification of themes such that specific patterns can be drawn out of it. Shukla, Wilson and Boddy (2014) stated that thematic analysis can be very closely associated with the narrative research as both the approaches are diverse and flexible. Riessman (2003) stated:

"*A key feature of narration is its performative and communicative nature: stories are told in interaction*"

The fact that narrative approach is typically analytically interpretive and enables the participants provide their insights through story-telling makes the researcher believe that the narrative approach in association with thematic analysis will be the best fit for this particular study. From the raw data collected, the researcher will analyze the commonalities among the participants' viewpoints and then apply thematic analysis to extract relevant themes pertaining to the current research topic.

The below described systematic guideline of thematic analysis, originally put forward by Braun and Clarke (2006) , will be adhered to as the main point of reference:
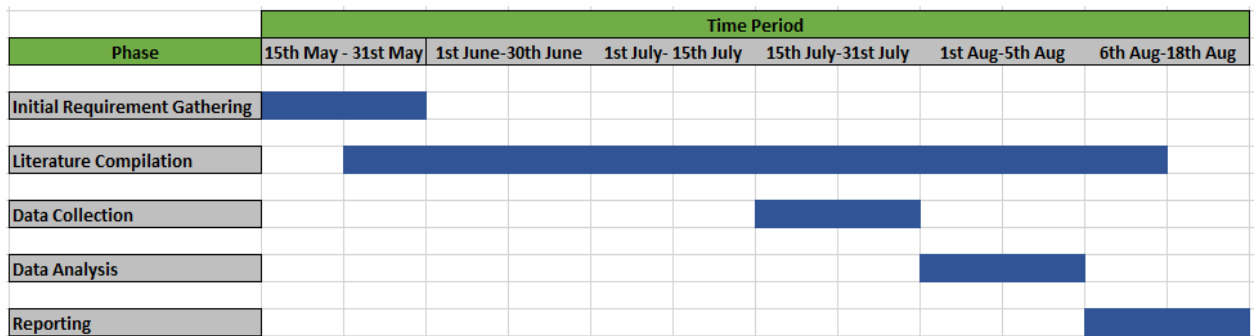
- The researcher is required to acquaint with the data by reading the collected information actively, analytically and critically.
- The next step involves formulating codes within the transcripts by identifying meaningful ideas within the collected data.
- It is followed by developing themes by identifying areas of similarity and clustering them.

- The next step is an iterative process of reviewing the created themes to check if it's coherent and distinct.
- This step involves naming the themes by identifying the essence of each of the themes.
- The final step is producing the report from the themes.

4.5.6 Time Horizon: The Research Onion put forward by Saunders *et al*. (2015) also focuses on incorporating a suitable time horizon for the study. Time horizons can either be cross-sectional or longitudinal in nature. While studies conducted under cross-sectional time horizon can be completed within a short span of time, longitudinal time horizon is particularly used when the study requires going through developments and modifications over a certain timeframe. It is usually concerned that a majority of the academic studies adopt the cross-sectional time horizon because of the time constraint.

The timeline set for completing the current study is: ***15th May-18th August***. This period includes literature compilation, methodology selection, data collection, data analysis and the overall reporting. Considering the limited timeframe allotted for the completion of the entire study , a cross-sectional time horizon is chosen. The below chart depicts the timeline for each phase.

Phase	Time Period					
	15th May - 31st May	1st June-30th June	1st July- 15th July	15th July-31st July	1st Aug-5th Aug	6th Aug-18th Aug
Initial Requirement Gathering	▰					
Literature Compilation		▰▰▰▰▰				
Data Collection				▰		
Data Analysis					▰	
Reporting						▰

Figure 9: Timeline for the study

4.6 Ethical Considerations: According to Connelly (2014), "informed consent is an essential part of the research process". It is extremely important to protect the data of the respondents to ensure that privacy is maintained and there are no harmful effects. As stated by Robson (2011), ethical considerations intensifies the engagement of the participants with the study by assuring them that the research will not pose any kind of threat. Saunders *et al.* (2009) highlighted the importance of

the researcher's behaviour towards the participants stating that the researcher should respect the rights of the participants.

A prime consideration that needs special attention is that the data collected from the participants comprises of information that are sensitive to their respective organizations as the information reveals about the kind of technical practices currently being carried out. Hence, proper care will be taken to keep the participants' personal as well as business data anonymous and confidential. Prior to collecting data, consent for voluntary participation will be noted. The consent will explicate all the rights and necessary information about the research. Finally, it will be the prime aim of the researcher to make the participants feel comfortable during their participation.

4.7 Limitations: A key limitation of this study is that the participants could not be interviewed through face-to-face interviews because of the location constraints. The participants of this study are residents of 7 different countries - India, Ireland, USA, Italy, Japan, Germany and Canada. Another limitation is the small sample size that was interviewed. Although it is regarded that inputs from fifteen participants were adequate for this study, a greater sample size would have provided stronger insights into the research topic.

Overall, this chapter described how the various procedures and methods will be conducted in accordance with the research objectives and questions. The following chapter will present the findings of the study that will help in answering the research questions.

CHAPTER V: Findings and Analysis

5.1 Introduction: The prime aim of this chapter is to portray the primary collected data through semi-structured interviews, investigate and discuss the findings within the context of this study. This section mainly attempts to make sense of all the findings by combining the observations from pre-existing literature and methodology. While transcribing the interviews, it has been verified twice so that the input data is accurate.

5.2 Participants' Data Presentation: Since this study has attempted to collect information from participants belonging to different parts across the globe, the primary data collection was accomplished through two ways:

- One-to-one interviews through voice recordings of the participants. 4 participants: one from Germany, one from Japan and two from India recorded their responses in their own voices and sent it through email.
- By floating out open-ended questionnaire to the participants. Around 11 participants preferred this medium as they felt it comparatively more convenient. The 11 participants belonged to India, Ireland, USA, Italy and Canada.

However, all the participants were selected based on *two specific criteria*:

- Participants working in the marketing division who have a minimum of 2 years experience working with analytics-based applications or tools.
- Participants working in Data Analytics department who have a minimum of 2 years experience developing analytics-based applications or tools for marketing division.
-

Below is an illustration of the personal and professional information of the participants:

Serial No.	Medium	Username of Participant	Country of Residence	Professional Role	Years of work experience
1	Voice Recordings	ani@jap	Japan	Principal Engineer and Co-founder	3
2		pta@ger	Germany	Marketing Analyst	3
3		eka@ind	India	Data Engineer	3
4		dhi@ind	India	Data Engineer	4
5	Open-ended questionnaire	kum@jap	Japan	IT Analyst	4
6		EOG@IND	India	Senior Data Analyst	4
7		kar@ind	India	Analyst	2
8		hah@usa	USA	Soon to join a startup	25
9		ARI@IND	India	Marketing Consultant	2+
10		too@can	Canada	Marketing Consultant	2.5
11		tiwind	India	Marketing Consultant	3
12		gam@ind	India	Analyst	2+
13		eka@IRE	Ireland	Data Analyst	3
14		ani@rom	Rome, Italy	Data Scientist	3
15		don@ita	Italy	Junior Data Scientist	2

Figure 10: Participants' Personal details

Figure 11: Countries of residence

As represented in the above figure , a majority of the participants are working professionals from India. Also, nearly an equal number of participants working in each of the division (marketing and data analytics) were interviewed for this study.

5.3 Thematic Analysis of Findings: The transcribed data was categorized into 4 distinct themes with a view to answering the research questions.

5.3.1 Meeting consumer needs, a crucial factor: Customers are the most precious asset of an organization. The key to the success of any organization relies on its capability of acquiring and retaining loyal customers. In today's global market, more and more businesses are emerging and competing with one another. To be successful, an organization must stand out from its competitors. This can be achieved by targeting more customers to use their products. But to acquire new customers is not an easy task and requires the organization to meet the expectations of the customers. Especially, in this era where digitalization has dramatically reduced the barriers and made communications easier, organizations should be able to read customers' minds to interpret what exactly it is that can encourage customers to purchase their products. This is essential to consider because customers make purchases only when their demands can be fulfilled with the products offered by the organizations.

Findings from this study indicate that '*meeting consumer needs*' was perceived as a critical factor by the respondents. As a prerequisite, '*understanding their exact needs*' was considered vital, yet complex. Respondents of this study from their own experiences emphasized on the importance of these two factors.

For instance, participant **gam@ind** made a remarkable statement stating:

> "*....If the company is building what it thinks it should and not what the customer wants, why would the customer even use it? It should always be customer centric....*".

In line with this view, participant **ani@jap** put forward his opinion citing an example from the tech giant Walmart. When asked about his opinion on the significance of meeting customers' demands, he remarked:

> "*....I do believe that this particular thing is an absolute necessity. There is a great deal of advantage and a great deal of success which a company can obtain if they understand what are the particular, let's just call them objects, which are related to each other that customers are most likely to buy.....*".

Participant eka@IRE made a well-versed statement arguing that acquiring customers is not easy for an organization:

"From acquiring a customer to engaging and retaining or to make an individual a loyal customer is not an easy task in today's time. Gone are the days when a customer hardly changes the brand of choice for life. There is competition out there everywhere. Understanding customer's purchase behavior is the key point for an organization's product success. "

Notably, participant eka@ind commented that meeting consumers needs can eventually help an organization grow as a brand. According to him :

"If the organization wants to grow itself as a brand in the market, then it is really important for the organization to understand the consumer behaviour so that it will help the organization to broaden its customer base"

Overall, the participants felt that for a business to be successful, the products offered should be customer-centric, especially this is a necessity for the long-term functioning of the company. Therefore, a common view was reflected from the responses of the participants, highlighting the essence of identifying and satisfying consumer demands for an organization's success.

5.3.2 Data Analytics in enhancing customer relationship: The impact of information technologies can be observed in almost every sphere in the recent times. Data Analytics is one of the rapidly emerging information technologies that is seen as a powerful tool in handling huge scale of data. Consequently, its uses can be found extensively in the field of Marketing, specifically, to deal with the massive amount of consumer data generated by businesses and consumer transactions. The pace at which data creation is accelerating makes it impossible to handle data manually by humans which has urged the organizations to adopt data analytics for accomplishing their work. While the importance of understanding consumers' needs has already been highlighted by the participants of this study, it is essential to note that to attain this requires organizations to be able to anticipate customers' expectations. Companies, now-a-days, are implementing data analytics to predict customers' needs and expectations which has eventually helped companies nurture their relationships with the customers by offering them exactly what they need.

During the interviews, each of the respondents highlighted that the inclusion of data analytics by their companies have helped them leverage their relationships with their customers. Majority of the respondents emphasized on three aspects - '*Continual increase in the amount of data generated*', '*heightened customer satisfaction*' and '*forecasting customer needs*'. For instance, Participant pta@ger, who is a Marketing Consultant from Germany said that

> "*Data Analytics has made it possible to correctly define consumer demands and behaviour for a long, foreseeable future in a way that businesses can plan their long-term strategies around such predictions. With the enormous amount of data being collected every minute and consumer behaviour very tangible and endurant conclusions can be drawn and how it might change over the course of the next decades and even more*".

A similar view was expressed by participant eka@IRE who has been working as a Data Scientist since the last 3 years:

> "*Data is the new oil……. Technologies like Data Analytics and related tools help an organization to understand their customers like never before, which helps in enhancing customer satisfaction.*"

In line with this, participants gam@ind and ARIIND respectively supported stating-

> "*...It's just that now the data we have is no longer limited. We have enormous amounts of data….*" and

> "*With data being collected at every instance from the consumers the role of data analytics powered with Artificial intelligence and machine learning will help companies forecast consumers behaviour….*"

Highlighting the essence of such technologies in building rapport with customers, participants too@can and tiwind have respectively commented that

> "*...Compared to earlier times, data analytics have made it easier for marketing analyst to help improve costumer relation.*" and "*These technologies help to forecast the needs of the customer and also identify prospective customers. Data analytics also help businesses to understand where they are strong and weak, and modify their operation similarly to achieve success.* ".

It is worth noting that Data Engineer, eka@ind stressed on the extensive uses of Data Analytics stating:

"Well, it's quite tough to mention one specific role of data analytics while studying consumer behaviour because the applications of data analytics are quite extensive. However, I would definitely like to mention that the advent of data analytics has significantly narrowed down the gap between the needs and expectations of the consumers and what is actually being offered to them."

Speaking about the role of Data analytics in making predictions, Respondent ani@jap said:

"...I have seen companies predict a lot of things with the data that they have collected, just like the example of Walmart, where predicting what a consumer might need beforehand is a process of data collection activity and utilizing that data to answer specific questions….just because of the scale of data that is being collected that manual processes tend to fail….."

In a nutshell, almost all the respondents felt that as the amount of data being yielded each day is rising dramatically, computing technologies like data analytics is contributing in predicting the needs of the customers by identifying patterns from the accumulated data which has consequently helped the companies in developing products and services that suit consumer demands, thereby leveraging customer satisfaction.

5.3.3 Accuracy and Effectiveness of data-driven techniques: The prime reason that has enabled the organizations to adopt technology-based techniques like Data Analytics is because they are considered to yield accurate results when compared to traditional techniques which are purely manual processes. Typically, clustering and Bayesian network algorithm are two frameworks that are considered to be most reliable.

The responses from the interviewees indicated that nearly all the respondents considered data-driven techniques to yield relatively accurate results. Participants mainly quoted the terms *'error free'*, *'high accuracy rate'* and *'reliable approach'* while answering the questions.

In support of why data-driven techniques are mostly preferred, participant EOG@IND responded-

"....I believe data-driven models help us make market predictions much faster, and more accurately."

Participant Tiwind emphasized on the benefits of data-driven techniques stating-

> *"Definitely data-driven. With the pace at which changes take place, disruption can happen any time and i, believe that no theoretical model can give as effective solution as a data-driven model, due to its factual and realistic feature. "*

Participant ani@rom put forward a solid reason as to why he prefers working with data-driven techniques:

> *"I prefer to work on data driven models because it requires less manual work and work can be completed error free using technologies like data analytics. Whereas, theories require more time to evaluate the information using traditional frameworks and result might not be accurate."*

Although the respondents primarily preferred data-driven methods because of its aforementioned features, few participants highlighted that such methods when used together with theoretical frameworks produce more precise results. For instance, don@ita commented:

> *"I think the best setting would be a mix of both. There has been many successful cases where theories , concepts and intuitions have inspired successful data driven algorithms."*

gam@ind too expressed a similar opinion stating -

> *".....The theories and concepts have or had been formulated on the basis of the data they have. So, indirectly they are also data driven. It's just that it's not live data and not customized as per your needs. I would like models to be first prepared based on theories/concepts and then improve them by feeding the data we have making it data-driven......"*

Another participant, too@can said-

> *"I would say both are equally important, Data Driven models can sometimes prove to be inaccurate and not reliable. hence its advisable to use both data driven models and theories and concepts. Although technology has advanced, there is always a scope of error. "*

Participants also mentioned that in certain cases, they have observed contradictions in the results derived through data-driven techniques and that through manual processes. For example, participant kum@jap highlighted a dark side of machine-driven techniques by stating-

> *"Yes there are so many cases where sudden changes in any analysis can be predicted by human brain but machine will only give you output whatever is fed to it. Although some AI tools have come up but still there is a long way to go."*

On the contrary, participant ani@rom commented:

> *"Yes, I have observed difference between results from data driven models and human predictions several times. Results from data-driven is fast, more accurate and error free. Sometimes it feels impossible to achieve accurate results with human predictions when looking on a big scale."*

Having experienced such differences between the results derived from human predictions and data-driven tools, few respondents considered a mix of both as the most reliable approach.

For instance, participant too@can stated:

> "*.....data driven models have helped understand customer behaviour more easily. But then again unless there is 100% accuracy rate with data driven model, human predictions should also be considered when predicting results*". Tiwind added further: "*A human prediction adds a depth to the results but the exact, quantifiable result gives better clarity"*

Participants who are working as Data Analysts or Data Engineers mentioned the different data mining techniques that they mostly work with. Although diverse responses were received from the participants, the most common answers included clustering and classification (Bayes network algorithm) techniques.

Participant eka@IRE briefly explained about the implementation of the process:

> "*As a Data Analyst we apply multiple set of techniques. The first starts with exploratory data analysis, understanding the data and decrease the anomalies. Many other statistical techniques are then used for understanding the relations. Algorithms like Bayesian network, clustering etc are good to go for any data analyst."*

Similarly, dhi@ind also stated that he works with clustering and Bayesian network algorithm:

> "*....I have worked upon various kinds of data mining processes but I have worked more upon the techniques which is a combination of both clustering and Bayes network algorithm."*

Some other responses from participants included:

ARIIND: "*Have used Regression analysis, to understand the relationship between two variables. For eg. Impact of product pricing on future purchase intentions.*"

kar@ind: "*Quality scrubbing and top down mining techniques are being used.*"

EOG@IND: "*Since I deal mainly with images, my data mining involves a lot of scraping from websites, and internal file storage systems.*"

Overall, it can be derived that the accuracy of data-driven techniques has proved to be useful in studying consumer behaviour and among all the techniques, clustering and Bayesian network algorithm are the most popular ones.

5.3.4 Evolution of marketing techniques in the era of technology: A number of participants expressed their opinion on how they felt about the contribution of advanced technology in the evolution of the marketing techniques over the years. Firstly, the researcher identified some of the top-most marketing processes that have gained popularity over the years with the help of the reviewed literature. Considering the expertise that the participants have in their respective job roles, all of them were asked to express their thoughts on this area.

The marketing processes that the researcher highlighted were 'Customer profiling', 'predicting future purchases' and 'Social Media Marketing'. It is interesting to note that a majority of the participants believed that 'predicting future purchases' is the most popular one among all.

In support of this ,participant pta@ger stated:

> "*Traditionally I believe a more personal but time-consuming approach where analysts actively interacted with the target customers was productive and fruitful but ever since the advent of data-driven predictive analysis of future purchasing trends, more and more marketing executives are using it in their daily work.*"

She further highlighted the impact of technology stating:

> "*I have always used data analytical tools in my work, be it for analyzing the customer once or the shift in target customers for different industries. Before data-driven technologies using data mining and machine learning became the most common and effective aids to marketing, it was the traditional feedback and word-of-mouth techniques that most marketing executives used….*"

A statement made by participant dhi@ind said:

> "*Although all the three options that are provided are increasingly becoming popular, but what i feel that is predicting the future with the help of machine learning and artificial*

intelligence holds a good stand in the market and it's the next big thing with the help of big data."

Moreover, a majority of the respondents have highlighted the strong impact of technology over these processes in their responses.

According to participant ARIIND,

"Marketing techniques have evolved radically over. Technology has been a crucial driver to influence the marketing practices worldwide."

Comparing the traditional and modern marketing processes, participant gam@ind stated:

"From my experience, I can clearly say that marketing processes are no longer just about manual feedbacks or physical hoardings, pamphlet etc. It's more about being present everywhere where the end user spends his/her time in i.e in the digital world. Wherever you go, they follow…..All this has been made possible via integrating modern tech….."

In support of this view, respondent eka@IRE explained:

"…..Marketing is the key thing that an organisation must focus on while there are competitors surround. A decade back or two, customers were mostly targeted mostly using TV ads, posters etc….But if I say about now, everything has changed. With the rise of data analytics, the marketing scenario has changed"

Highlighting one major benefit of information technologies, participant eka@ind stated:

"With advancement in digitalization and information technology, I have seen tremendous changes in the way marketing techniques are applied today, mainly the manual work has been reduced dramatically"

Going through all the responses, it can be anticipated that the respondents considered information technologies as a powerful tool that has brought significant transformations in the marketing processes over the period. A key aspect that was highlighted in the responses is '***reduced manual work***' According to most of the participants, using computing technologies have automated the tasks, significantly reducing manual work. Moreover, the participants believed that the old practices of marketing are rare to be seen now-a-days. On the contrary, technology-based practices have gained a wider attention and contributed to these transformations.

The responses from the participants indicate that predicting future purchases is opted more than customer profiling and social media marketing by the participants as illustrated in the figure.
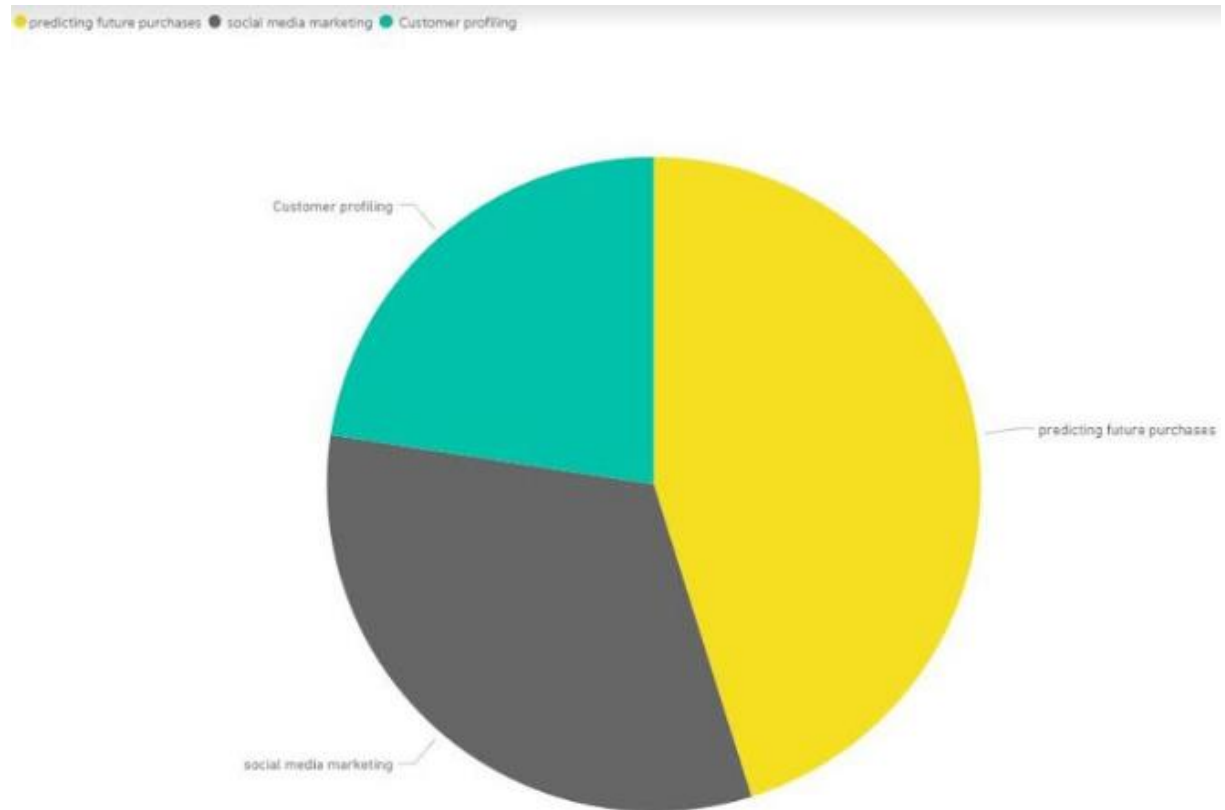
Figure 12: Participants' choice of top-most marketing process

This chapter presented all the findings assembled through primary data collection. The next chapter will discuss all the findings in accordance to the extant literature.

CHAPTER VI: Discussion

6.1 Introduction

To derive a conclusion and answer the research questions, further discussion is necessary. This chapter intends to bring together the findings and analysis of the data from the interviews and the open-ended questionnaires, and link it up to the relevant literature presented in CHAPTER II. The insights obtained from the respondents' personal experiences indicated that their views on the contribution of data analytics in studying consumer behaviour were very much in line with the existing literature.

Although responses have been received from a good number of participants through semi-structured interviews, certain aspects within the study require further research which will be discussed later in the study.

6.2 Analysis of themes

6.2.1 Meeting consumer needs, a crucial factor: As discussed in the literature review, customers make purchases only when their needs and demands tend to be fulfilled. The researcher's findings indicate that the participants articulated a similar view. It is very evident from the findings that participants highlighted two critical aspects in their responses: '***understanding customers' needs***' and '***meeting customers' needs***'.

While some participants, for instance, gam@ind and ani@jap considered meeting consumer demands as an absolute necessity, participants like tiwind felt that satisfying consumer needs is essential for the long-term functioning of an organization.

Participant pta@ger acknowledges customers as so powerful that she comments that a customer '*..can make or break a venture overnight..*'. This view can be supported by (Prasad and Jha, 2014) who briefly explained why it is essential to pay heed to the interests of the consumers including how they prefer to spend the available resources (time, money, effort) to make their purchases.

Notably, participants also explained that the whole process from acquiring customers to retaining them as loyal customers is a complicated task, which is why it should be given utmost attention by the organizations. (Tynan and McKechnie, 2009) expressed a similar opinion arguing that fulfilling consumer demands is not easy as it requires granular analysis of the preferences of the customers.

In a nutshell, the findings reflect that according to the participants' viewpoints, a customer is satisfied when the expectations of the customers exactly matches the products offered by the organization which is why meeting consumer needs is the prime target of their respective organizations. This is clearly supported in the literature by (Shao, 2015) who opined that identifying the preferences of the consumers ought to be a priority.

6.2.2 Data Analytics in enhancing customer relationship: It is very apparent from the findings of this study that data analytics is perceived to be one of the most powerful computing technologies that has helped in nurturing customer relationships. A majority of the respondents, from their experiences, conveyed that implementing data analytics has made it easier to correctly define consumers' needs and focus on long-term strategies for their organizations. Participants have acknowledged the fact that the amount of data created each day is rising incredibly. To handle this massive level of data, computing technologies like data analytics is being utilized by their companies. Such experiences shared by the participants can strongly be supported by the findings of (France and Ghose, 2018) already cited in the literature which reported that the key intent of applying analytics is to leverage customer relationships.

Furthermore, (Boonsiritomachai *et al.*, 2016) has reported that predictive analytics of data is deemed to have been delivering higher business value. Similar views were expressed by participants pta@ger, ARIIND, too@can and ani@jap which have been illustrated in the findings. They explained from their own experiences how data analytics is quickening the process of making predictions by drawing correlation among huge datasets. In support of this, one of the participants, ani@jap highlighted that manual processes are failing in handling the large scale of data.present today. While responding to the interview questions, participants mainly remarked that as data analytics is making it easier to make future predictions about customers' expectations, companies are able to increase customer satisfaction and retain high standards by developing products and services that customers demand. This view is supported by Myers (2001) too which have been cited in the literature review. Additionally, Solomon (2009) quoted that greater level of satisfaction encourages the customers to repurchase the product.

Citing real-life examples, participant ani@rom explained how the two tech giants Google and Amazon are using data analytics and machine learning to identify the needs of the customers from the collected information.
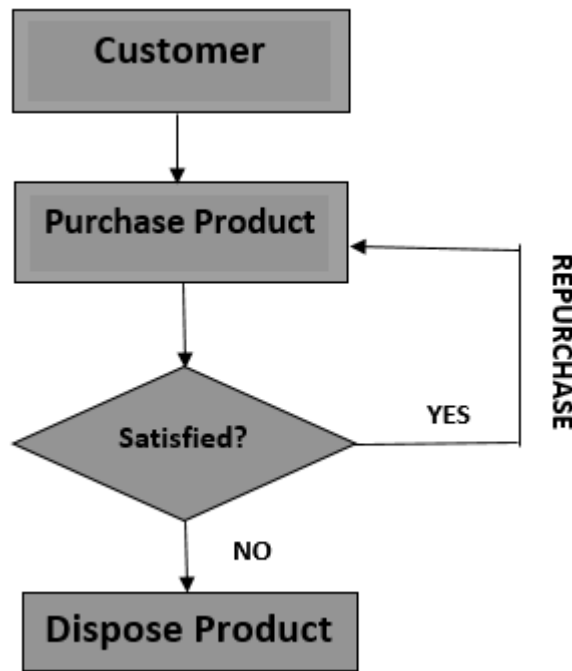
Figure 13: Cycle of customers' purchase behaviour

This figure is a simple illustration of the above discussed phenomenon. It clearly depicts that a satisfied customer tends to repurchase the product. On the contrary, if a customer is dissatisfied, it leads to termination of the consumption of the product.

6.2.3 Accuracy and Effectiveness of data-driven techniques: During the interviewees, a majority of the participants reasonably explained why their preferences were towards the data-driven techniques. All the interviewees (Marketing as well as Data professionals) had some level of experience in using data-driven tools or techniques in their professional career. From their experiences, they felt that such methods are constantly being adopted by their companies as they produce real-time based results. As stated by the respondents, data-driven or machine-driven methods yield comparatively accurate results and substantially reduces the error rate that arises during the manual work by humans. Furthermore, 5-6 respondents stressed on the point that working with theoretical frameworks is quite time-consuming, whereas data-driven methods have potential to complete massive amount of tasks in a short timeframe. Similar observations were derived by Foxal (1980) who criticized the traditional theoretical frameworks stating that they

require a huge amount of time to accomplish a specific activity. However, it is interesting to note that although most of the participants opined that theoretical models are not much effective to use, few respondents suggested that a mix of the two approaches can produce finer outcomes. For instance, participants too@can, don@ita and gam@ind explained that data-driven models can initially be built upon the concepts of theoretical frameworks, and should later be improvised to feed in more data. This approach can supposedly help in building more reliable techniques.

In addition to this, participants put forward their viewpoints explaining how they have come across contradictions between results derived from human predictions and that from machine-driven models. While a majority of them focused on the use of only data-driven models, few respondents highlighted the need to use human predictions along with such models. They felt that unless a machine-driven model can yield 100% accurate data, human intervention should always be considered. Moreover, it is crucial to note that such models predict or showcase data based on what is actually fed into it, whereas there are no boundaries to human thoughts and processes.

Finally, participants mentioned some of the data mining techniques that they frequently use in their respective organizations. While a diverse range of techniques were noted such as regression analysis, quality scrubbing, pattern tracking, top down mining and graph databases, the most popular responses were clustering, classification and Bayesian network methods . Participants have addressed that the use of clustering and classification procedures help in identifying relations among data as well as contributes in devising effective marketing strategies. Such views from the participants can be directly related to the statements from (Raorane and Kulkarni, 2011) in the literature who explained that these data mining techniques help in drawing correlations among the accumulated data. Furthermore, the comments made by participants eka@IRE and ani@rom can be well associated to that made by Maingi (2015) in the literature on how the clustering technique can help in predicting sales or consumer's purchasing habits.

In addition to this, the interviewees' recommendations on the use of Bayesian network algorithm seems to be justified. Several researchers have previously emphasized on the significance of this algorithm stating that it is a very reliable method for data prediction (Papić-Blagojević, 2011; Hsieh and Chu, 2009 ; Ibukun.T *et al.,* 2016 ; Keller, Gray and Givens, 1985).

6.2.4 Evolution of marketing techniques in the era of technology: Findings from this study indicate that the respondents felt that technology played an integral role in bringing out noteworthy

transformations in the marketing processes over the years. Participants mentioned about high computing technologies like Data Analytics, Machine Learning and Artificial Intelligence during the interviews that have brought into such changes. While responding to the interview questions, a majority of the participants made a comparison between the traditional ways of marketing and the recent marketing techniques in order to support their viewpoints. Some respondents, for instance, eka@IRE, pta@ger and gam@ind mentioned traditional marketing techniques such as advertisements through physical hoardings, TV ads, distributing pamphlets , ads in newspapers and even word-of-mouth as being widely used in the earlier times. This view has been supported by (Lau and Ng, 2001) who stated that consumers tend to get immersed in the initial product purchase stage through word-of-mouth process. Additionally, Brown and Oplatka (2016) stated that individuals typically get involved in word-of-mouth process with neighbors, colleagues, close friends or family. However, according to the participants, most of these techniques do not seem relevant or applicable in the current times. On the contrary, the marketing processes functioning in the current times seem to have some level of technological influence within it.

Furthermore, few participants, for instance, ani@jap have highlighted that the use of computing technologies has minimized the manual work or processes which are usually quite time-consuming. He made a remarkable comment stating that the biggest changes were influenced by the psychological and behavioural attributes that consumers tend to perceive. A similar viewpoint was shared by (T.K, 2014) who highlighted few psychological attributes that immensely impacted consumer behaviour.

The views of the respondents on the strong impact of information technology can be supported by the statement made by Jain (2017) :

"There's no denying that technology had changed the way that global citizens receive, interpret and react to information….the rapid evolution of devices that allow quick and easy access to its millions of portals, consumers are finding new ways to interact with companies and with products."

As described by the participants, Jain too highlights that traditional practices of advertisements through newspapers, magazines and TV ads compete with the recent technology-induced practices.

6.3 Limitations of the study: The current study is subject to certain restrictions considering its subjective and interpretive nature. The researcher considers the following as some of the limitations:

- One of the main constraints is that although the researcher puts efforts to conduct one-to-one interviews over the telephone, the geographical location constraint and differences in timezone (for 7 different countries) encouraged the participants to share their responses either by filling out the open-ended questionnaire or by sending out their voice clips over email.

- Despite the fact that the participants provided descriptive answers to the interview questions and shared their personal experiences, it is possible that the participants might have misinterpreted the interview questions or perceived the questions from altogether a different viewpoint.

- This study has also some time constraints. On the researcher's part, the data collection and data analysis had to be completed within a very limited timeframe. On the other hand, since all the participants are working professionals having certain kind of burden, it prevented the respondents to deliver developmental explanations.

- A major drawback of this kind of study is that participants sometimes tend to form biased opinions or judgement. It might be possible that participants have shared something inaccurate based on mere assumptions or due to limited knowledge on the interview questions.

6.4 Implications of the study: The empirical findings of this study contributes to the theories of consumer behaviour involving the practice of data analytics. Although the results of this research affirm that data analytics is increasingly being adopted by the companies to make predictions on consumers' buying trends, the relevant literature associated with the current research topic is quite limited. Additionally, while the findings seem consistent with the literature review, representing clustering and Bayesian network algorithm as the most popular and effective data mining techniques for making future predictions, participants have also brought into light some alternative data mining techniques which are currently being implemented in their organizations to make future predictions with consumer data.

Therefore, the results of the study and the outlined limitations imply that further research into the current research subject will be fruitful to examine the effectiveness of data-driven techniques in predicting consumers' purchase patterns. Additionally, while it can be inferred from the findings that technology is acting as a key driver in bringing more refined and reliable transformations into the ways of implementation of the marketing processes, including prediction of future purchases, it seems equally important to do some research on the darker side of technology and investigate further into what makes such technology-induced techniques better than the traditional ones for the prediction of consumer behaviour.

CHAPTER VII: Conclusion and Recommendations

7.1 Conclusion: This chapter is the overall conclusion of the current research study. The key intent of this chapter is to revisit the study's research objectives and examine if the findings of the study relate to or answer the research questions. This qualitative study was conducted within interpretivism and the participants' viewpoints assembled through semi-structured interviews were interpreted to answer the research questions. The findings from the interviews along with journal articles and extant literature have contributed to the conclusions delineated in this section. It is evident from the findings that it contributed to fulfilling the research objectives of this study mentioned in section **3.3**.

One distinct finding that emerged from the collected data is that all the participants (marketing as well as data professionals) considered data analytics as a powerful tool in anticipating consumer's purchase behaviour. This finding is also well supported in the literature explaining how data analytics is increasingly being used by these professionals in the recent times. Additionally, the findings indicated a similarity between the kinds of data mining techniques that are generally being used in different countries. After analyzing the data from respondents working in 7 different countries, it can be derived that the clustering and Bayesian network algorithm are regarded as the most efficacious methods for making future predictions. These two methods have been reported as the most efficient ones in the literature review too. Finally, the findings also bring into light the reasonings put forward by the participants as to why a majority of them preferred to work with data-driven techniques rather than traditional theories and frameworks which is in line with the reviewed literature. Nevertheless, it is worth mentioning that few participants recommend a combination of data-driven techniques and traditional theories in order to derive more precise results.

Taken together, the researcher concludes that the empirical findings are consistent with the reviewed literature and contributed to accomplishing the research objectives of the study. The findings also indicate that most of the organizations are using Data Analytics to accelerate their work as well as to deliver superior results to customers. Despite the fact that the findings of the study was sufficient to answer the research questions, it is important to note that the sample size interviewed was quite small. Hence, there is still scope for further research.

7.2 Recommendations for Future Research: Based on the overall study conducted and the outlined limitations, the following are the primary recommendations for future research:

- Since this study considered only 7 countries while collecting information, it will be beneficial and more insights can be obtained if similar researches are undertaken with a greater sample size , targeting participants belonging to other countries.

- One critical point to consider while conducting such studies is related to the privacy of consumers' data. While the researcher has not attempted to examine whether implementing data-driven or machine-driven techniques will breach customers' personal data, this aspect should be given prior attention in further research. It is important to perform an in-depth research on whether implementation of such techniques can lead to some constraints on the ethical side.

- Additionally, while this study was conducted within the qualitative context, further studies could attempt to adopt the quantitative method too or even a mix of the two methods to ensure that the findings have greater accuracy and no biased opinions influence the final outcome of the study.

- Finally, the researcher believes that further research can also include customers as the participants to enquire about their thoughts on how they think advanced information technology has made their personal data vulnerable.

References

Acemoglu, D. and Restrepo, P. (2018) 'The Race between Machine and Man: Implications of Technology for Growth, Factor Shares and Employment', *American Economic Review,* 108(6) pp. 1488-1542. doi:10.1257/aer.20160696.

Alshurideh, M.T. (2014) 'A Qualitative Analysis of Customer Repeat Purchase Behaviour in the UK Mobile Phone Market', *Journal of Management Research*, 6(1), doi:10.5296/jmr.v6i1.4659.

Banduka, N. *et al.* (2017) 'Using 80/20 principle to improve decision making at PFMEA', *Annals of DAAAM & Proceedings*, pp. 487. doi: 10.2507/27th.daaam.proceedings.073.

Barbera, F.L., Amato, M. and Sannino, G. (2016) Understanding consumers' intention and behaviour towards functionalised food : The role of knowledge and food technology neophobia' (2016) *British Food Journal*, (4), p. 885. doi: 10.1108/BFJ-10-2015-0354.

Bazmara, M. and Movahed, S.V. (2013) 'KNN Algorithm for Consulting Behavioral Disorders in Children', *Journal of Basic and Applied Scientific Research*, 3(12), pp.352-358. Available at: https://www.researchgate.net/publication/259468508_KNN_Algorithm_for_Consulting_Behavioral_Disorders_in_Children [Accessed 25th July].]

Blaikie, N.W.H. (2000), Designing social research: The logic of anticipation. UK: Blackwell Publishing.

Blázquez, M. (2014) 'Fashion Shopping in Multichannel Retail: The Role of Technology in Enhancing the Customer Experience', *International Journal of Electronic Commerce*, 18(4), pp. 97–116. doi: 10.2753/JEC1086-4415180404.

Boeije, H. (2010), Analysis in Qualitative Research, Sage Publications, London.

Boonsiritomachai, W., McGrath, G.M and Burgess, S. (2016) 'Exploring Business intelligence and its depth of maturity in Thai SMEs', Cogent Business & Management, 3(1). doi: 10.1080/23311975.2016.1220663.

Borkin, M.A., Vo, A.A., Bylinskii, Z., Isola,P., Sunkavalli, S., Olivia, A. and Pfister, H. (2013) 'What Makes a Visualization Memorable?' (2013) *IEEE Transactions on Visualization and Computer Graphics, Visualization and Computer Graphics, IEEE Transactions on, IEEE Trans. Visual. Comput. Graphics*, (12), pp. 2306. doi: 10.1109/TVCG.2013.234.

Bryman, A. and Bell, E. (2015) Business Research Methods, 4th ed. Oxford: Oxford University Press.

Christensen, C. M. and Overdorf, M. (2000) 'Meeting the Challenge of Disruptive Change. (cover story)', *Harvard Business Review*, 78(2), pp. 66–76. Available at: http://search.ebscohost.com/login.aspx?direct =true&AuthType=ip,cookie,shib&db=bth&AN=2839608&site=eds-live&scope=site (Accessed: 17 August 2019).

Chung, H. M. and Gray, P. (1999) 'Special Section: Data Mining', *Journal of Management Information Systems*, 16(1), pp. 11–16. doi: 10.1080/07421222.1999.11518231.

Connelly, L. M. (2014) 'Ethical considerations in research studies', *MedSurg Nursing*, 23 (1), pp. 54. Available at: http://search.ebscohost.com/login.aspx?direct=true&Auth Type=ip ,cookie,shib&db=edsgao&AN=edsgcl.360608990&site=eds-live&scope=site&custid=ncirlib (Accessed: 25 June 2019).

Creswell, J.W (2007) Qualitative Inquiry and Research Design, 2nd edn. US: Sage Publications

Dawson, J. A. and Sparks, L. (1986) 'New Technology in UK Retailing: Issues and Responses', *Journal of Marketing Management*, 2(1), pp. 7–29. doi: 10.1080/0267257X.1986.9963997.

Devi, P. and Rajagopalan, S.P. (2012) 'Analysis of Customer Behavior using Clustering and Association Rules', *International Journal of Computer Applications,* 43 (23). Available at: https://www.semanticscholar.org/paper/Analysis-of-Customer-Behavior-using-Clustering-and-Devi-Rajagopalan/31062736f81c9f5911fdd8effae1b460bff5c08b

Dubinsky, A. J. and Hansen, R. W. (1982) 'Improving Marketing Productivity: The 80/20 Principle Revisited', *California Management Review*, 25(1), pp. 96–105. doi: 10.2307/41164996.

Dupre, E. (2013) 'Let data take the lead: data and analytics help marketers find their lead generation rhythm', *DM News*, p. 32. Available at: http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip, cookie,shib&db=edsggo&AN=edsgcl.332655941&site=eds-live&scope=site (Accessed: 2 August 2019).

Eikebrokk, T.R. and Busch, P.A (2016) 'Progress and Stewardship in Information Systems Research: Addressing Barriers to Cumulation through Active Process Ownership', pp.1-10. Available at: https://www.researchgate.net/publication/301870765_Progress_and_Stewardship

_in_Information_Systems_Research_Addressing_Barriers_to_Cumulation_through_Active_Process_Ownership

Erevelles,S., Fukawa, N. and Swayne, L. (2015) 'Big Data consumer analytics and the transformation of marketing', *Journal of Business Research,* 69 (2016), pp. 897-904, doi:10.1016/j.jbusres.2015.07.001

Eryigit, C. (2017) 'Marketing models: A review of the literature', *International Journal of Market Research*, 59(3), pp. 355–381. doi: 10.2501/IJMR-2017-028.

Etikan, I., Musa, S.A. and Alkassim, R. S (2015) 'Comparison of Convenience Sampling and Purposive Sampling', *American Journal of Theoretical and Applied Statistics,* 5(1), pp.1-5. doi: 10.11648/j.ajtas.20160501.1

Few, S. (2007) 'Data Visualization Past, Present and Future', *Cognos Innovation Center*, pp. 2-10. Available at: http://mail.perceptualedge.com/articles/Whitepapers/Data_Visualization.pdf

France, S. L. and Ghose, S. (2019) 'Marketing analytics: Methods, practice, implementation, and links to other fields', *Expert Systems With Applications*, 119(1), pp. 456-475. doi: 10.1016/j.eswa.2018.11.002.

Gavett, G. (2014) 'How Data Visualization Answered One of Retail's Most Vexing Questions', *Harvard Business Review Digital Articles*, pp.2–6. Available at: http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,cookie,shib&db=bth&AN=118647267&site=eds-live&scope=site (Accessed: 29 July 2019).

Gelman, A. (2008) 'Objections to Bayesian statistics', *International Society for Bayesian Analysis,* 3(3), pp. 445-450. doi: 10.1214/08-BA318

Hair Joe F. (2007) 'Knowledge creation in marketing: the role of predictive analytics', *European Business Review*, (4), p. 303. doi: 10.1108/09555340710760134.

Hemsley-Brown, J. and Oplatka, I. (2016) 'Personal Influences on Consumer Behaviour'. doi: 10.1007/978-1-137-49720-8_3.

Hsieh, N. and Chu, K. (2009) 'Enhancing Consumer Behavior Analysis by Data Mining Techniques', *International Journal of Information and Management Sciences,* 20(1), pp. 39-53. Doi: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.579.9503&rep=rep1&type=pdf

Ibukun.T, A., Oladipupo, O., Worlu, R.E. and Akinyemi, I.O. (2016) 'A Systematic Review of Consumer Behaviour Prediction Studies', *Covenant Journal of Business & Social Sciences,* 7 (1). Available at: http://eprints.covenantuniversity.edu.ng/7543/#.XUCYFfkzbDd

Johar, J. S. and Sirgy, M. J. (1989) 'Positioning Models in Marketing: Toward a Normative-Integrated Model', *Journal of Business & Psychology*, 3(4), pp. 475–485. doi: 10.1007/BF01020715.

Johnson, R. B., & Stefurak, T. (2013). Considering the evidence-and-credibility discussion in evaluation through the lens of dialectical pluralism, *New Directions for Evaluation*, 2013(138), pp.37-48. doi:10.1002/ev.20056

Keller, J.M., Gray, M.R. and Givens, J.A. (1985) 'A Fuzzy K-Nearest Neighbor Algorithm', *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS,* 15(4). Available at: https://ieeexplore.ieee.org/document/6313426

Kumar, V. (2015) 'Evolution of Marketing as a Discipline: What Has Happened and What to Look Out For', *Journal of Marketing*, 79(1), pp. 1–9. doi: 10.1509/jm.79.1.1.

Kvale, S. and Brinkmann, S. (2009), Interviews: Learning the Craft of Qualitative Research Interviewing. 2nd ed. Beverly Hills, California: Sage Publications.

Lau, G.T. and Ng, S. (2001) 'Individual and Situational Factors Influencing Negative Word-of-Mouth Behaviour', *Canadian Journal of Administrative Sciences (Canadian Journal of Administrative Sciences)*, 18(3), p. 163. Available at: http://search.ebscohost.com/login.aspx ?direct=true&AuthType=ip,cookie,shib&db=bth&AN=5710136&site=eds-live&scope=site (Accessed: 20 July 2019).

Lin, A.C. (1998), 'Bridging positivist and interpretivist approaches to qualitative methods', *Policy Studies Journal*, 26(1) pp. 162-180. doi: 10.1111/j.1541-0072.1998.tb01931.x

Liu, R. (2017) 'A Reappraisal on Marketing Definition and Marketing Theory', *Journal of Eastern European and Central Asian Research*, (2). doi: 10.15549/jeecar.v4i2.170.

Lund, T. (2012) 'Combining Qualitative and Quantitative Approaches: Some Arguments for Mixed Methods Research', *Scandinavian Journal of Educational Research*, 56(2), pp. 155–165. doi: 10.1080/00313831.2011.568674.

Mackenzie, N. and Knipe, S., (2006). 'Research dilemmas: Paradigms, methods and methodology', *Issues in Educational Research*, 16(2), pp.193-205. Available at:

https://www.researchgate.net/publication/273135528_Research_dilemmas_Paradigms_methods_and_methodology

Maingi, M.N. (2015) 'A Survey on the Clustering Algorithms in Sales Data Mining', *International Journal of Computer Applications Technology and Research,* 4(2), pp.126-128. Available at: https://pdfs.semanticscholar.org/d895/930b3fb948d9009ef809e2d18925dc3001eb.pdf

Massy, W. F. and Savvas, J. D. (1964) 'Logical Flow Models for Marketing Analysis', *Journal of Marketing*, 28(1), pp. 30–37. doi: 10.2307/1249223.

Mathison, S. (1988) 'Why triangulate?', *Educational Researcher*, 17(2), doi: 10.3102/0013189X017002013.

Mazumdar, A. (2010) 'Predicting customer purchase in an online retail business, a Data Mining approach', pp.1-41. Available at: https://www.researchgate.net/publication/43806993 _Predicting_ customer_purchase_in_an_onlineretail_business_a_Data_Mining_approach.

Miles, M. B. and Huberman, M. (1994) Qualitative data analysis: an expanded sourcebook. 2nd ed. London: Sage publication.

Morse, J.M., Barrett, M., Mayan, M., Olson, K. and Spiers, J. (2002) 'Verification strategies for establishing reliability and validity in qualitative research', *International Journal of Qualitative Methods*, 1(2): pp. 1-19. doi:10.1177/160940690200100202.

Narayandas, D. (2005) 'Building Loyalty in Business Markets', *Harvard Business Review*, 83(9), pp. 131–139. Available at: http://search.ebscohost.com/login.aspx?direct =true&AuthType=ip,cookie ,shib&db=bth&AN=18301648&site=eds-live&scope=site (Accessed: 17 August 2019).

Newman, I. and Houchins, D. E. (2018) 'Conceptualizing Mixed Methods Questions in Special Education Research', *Research in the Schools*, 25(2), pp. 23–33. Available at: http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,cookie,shib&db=ehh&AN=13 5047243&site=eds-live&scope=site (Accessed: 29 July 2019).

Nguyen, Dung H., Sander de Leeuw, and Wout E. H. Dullaert. (2018) "Consumer Behaviour and Order Fulfilment in Online Retailing: A Systematic Review." *International Journal of Management Reviews,* 20(2), pp. 255–76. doi:10.1111/ijmr.12129.

Pahwa, B., Taruna, S. and Kasliwal, N. (2017) 'Role of Data mining in analyzing consumer's online buying behavior', *International Journal of Business and Management Invention,* 6(11), pp. 45-51. Available at: https://www.researchgate.net/publication/324595081_Role_of_Data_mining _in_analyzing_consumer's_online_buying_behavior

Pantelic, D., Sakal, M. and Zehetner, A. (2016) 'Marketing and sustainability from the perspective of future decision makers', *South African Journal of Business Management*, 47(1), pp. 37–47. doi: 10.4102/sajbm.v47i1.51.

Papić-Blagojević, N. (2011) 'ANALYSIS OF THE INFLUENCE ON CONSUMER BEHAVIOR BY
APPLYING BAYESIAN NETWORK', *International Symposium Engineering Management And Competitiveness',* pp.347-352 Available at: https://pdfs.semanticscholar.org/ae79/955eace 73313eecd7ad86 1bc81a658b2dd48.pdf.

Parasuraman, A., Berry, L.L and Zeithaml, V.A (1991) 'Understanding Customer Expectations of Service', *Sloan Management Review*, 32(3), pp.39-46.

Park, J, and Park, M (2016) 'Qualitative versus Quantitative Research Methods: Discovery or Justification?' *Journal of Marketing Thought,* 3(1) pp. 1-7. doi:10.15577/jmt.2016.03.01.1.

Parker, L.D. (2003), 'Qualitative research in accounting and management: The emerging Agenda', *Journal of Accounting and Finance*, 2: pp. 15-30. Available at: http://eprints.gla.ac.uk/132785/

Patel, M., Karvekar, S. and Mehta, Z (2014) 'Customer Behaviour Model Using Data Mining', *International Journal of Advanced Technology in Engineering and Science,* 2(1), pp.709-716. Available at: https://pdfs.semanticscholar.org/3c42/43c73176235dd46ebeca1185f90d7cee0c1 f.pdf.

Patil, P.H., Thube, S., Ratnaparkhi, B. and Rajeswari, K. (2014) 'Analysis of Different Data Mining Tools using Classification, Clustering and Association Rule Mining', *International Journal of Computer Applications,* vol 93(8), pp. 35-39. Available at: https://pdfs.semanticscholar.org/7d74/1e72e44a98 9e55432c359e3c4ec84592ded3.pdf.

Patton, M.Q. (2002), Qualitative research and evaluation methods, 3rd ed., Thousand Oaks, CA: Sage Publications

Piotr Tarka and Maciej Łobiński (2014) 'Decision Making in Reference to Model of Marketing Predictive Analytics – Theory and Practice', *Management and Business Administration. Central Europe*, (872348), pp. 60. Available at:

http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,cookie,shib&db
=edsdoj&AN=edsdoj.883a91d1646c40dc8387eaae33a67a85&site=eds-live&scope=site
(Accessed: 31 July 2019).

Prasad, R.K. and Jha, M.K. (2014) 'Consumer buying decisions models: A descriptive study', *International Journal of Innovation and Applied Studies,* 6(3), pp.335-351. Available at: https://www.researchgate.net/publication/305727378_Consumer_buying_decisions_models_A_d escriptive_study

Quinlan, C. (2011) Business Research Methods. UK: Hampshire.

Raorane, A. and Kulkarni, R.R. (2011) 'Data Mining Techniques: A Source for Consumer Behavior Analysis', pp.1-15. doi: 10.5121/ijdms.2011.3304.

Robson, C. (2011) Real World Research, 3rd ed. West Sussex: John Wiley & Sons Limited

Robson, C. (2011), Real world research, 3rd ed., Wiley-Blackwell: Oxford publisher.

Saunders, M. Lewis, P. and Thornhill, A. (2015), Research Methods for Business Students, New York: Pearson Education

Schneider, B. and Bowen, D. E. (1999) 'Understanding Customer Delight and Outrage', *Sloan Management Review*, 41(1), pp. 35–45. Available at: http://search.ebscohost.com/login.aspx?direct =true&AuthType=ip,cookie,shib&db=bth&AN=2421815&site=eds-live&scope=site (Accessed: 12 August 2019).

Shao, X.-F. (2015) 'Product differentiation design under sequential consumer choice process', *International Journal of Production Research*, 53(8), pp. 2342–2364. doi: 10.1080/00207543.2014.951091.

Solomon, M., Bamossy, G., Askegaard, S. and Hogg, M.K. (2009) Consumer Behaviour, 8th ed. New Jersey: Pearson Education.

Suchacka, G. and Chodak, G. (2017) 'Using association rules to assess purchase probability in online stores', *Information Systems and e-Business Management*, (3), pp. 751. doi: 10.1007/s10257-016-0329-4.

T.K., Jisana (2014) 'CONSUMER BEHAVIOUR MODELS: AN OVERVIEW', *Sai Om Journal of Commerce & Management,* 1(5), pp.35-43. doi: 10.1.1.916.4415.

Tordera, D.A. (2013) 'On Contemporary Models of Consumer Behaviour', *MBA Marketing and Financial Management.* Available at: https://danieltordera.wordpress.com/2013/11/21/on - contemporary-models-of-consumer-behavior-2012/.

Tynan, C. and McKechnie, S. (2009) 'Experience marketing: a review and reassessment', *Journal of Marketing Management*, 25(5–6), pp. 501–517. doi: 10.1362/026725709X461821.

Urbonavičius, S., Dikčius, V. and Kasnauskienė, G. (2007) 'Importance of Marketing Functions in a Company', *Engineering Economics*, 53(3), pp. 71–77. Available at: http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,cookie,shib&db=bth&AN=25 805024&site=eds-live&scope=site (Accessed: 30 July 2019).

Viksne, K., Salkovska, J., Gaitniece, E. and Puke, I. (2016) 'Comparative Analysis of Customer Behaviour Models', *International Conference "ECONOMIC SCIENCE FOR RURAL DEVELOPMENT"*, (43), pp.231-238. Available at: http://llufb.llu.lv/conference/economic_ science_rural/2016/Latvia_ESRD_43_2016-231-238.pdf.

Wahyuni, D. (2012), 'The Research Design Maze: Understanding Paradigms, Cases, Methods and Methodologies', *Journal of Applied Management Accounting Research*, 10(1), pp. 69. Available at: https://www.researchgate.net/publication/256024036_The_Research_Design _Maze_Understanding_Paradigms_Cases_Methods_and_Methodologies.

Wan, Y., Liu, C. and Qiao, W. (2019) 'An Safety Assessment Model of Ship Collision Based on Bayesian Network' *2019 European Navigation Conference (ENC), Navigation Conference (ENC), 2019 European*, p. 1. doi: 10.1109/EURONAV.2019.8714177.

Yin, R. K. (2014) Case Study Research: Design and methods. 5th ed. London: Sage publication

Zakaria, I. *et al.* (2014) 'The Relationship between Loyalty Program, Customer Satisfaction and Customer Loyalty in Retail Industry: A Case Study', *Procedia - Social and Behavioral Sciences*, 129, pp. 23–30. doi: 10.1016/j.sbspro.2014.03.643.

Zhang, T. *et al.* (2019) 'The role of virtual try-on technology in online purchase decision from consumers' aspect', *Internet Research*, (3), p. 529. doi: 10.1108/IntR-12-2017-0540.

Zhong, S., Khoshgoftaar, T.M. and Seliya, N. (2004) 'Analyzing software measurement data with clustering techniques', *IEEE Intelligent Systems, Intelligent Systems, IEEE, IEEE Intell. Syst*, (2), pp. 20. doi: 10.1109/MIS.2004.1274907.

Zyphur, M, and Pierides, D. (2017), 'Is Quantitative Research Ethical? Tools for Ethically Practicing, Evaluating, and Using Quantitative Research', *Journal of Business Ethics*, 143(1) pp.1-16. doi:10.1007/s10551-017-3549-8.

Appendices

Appendix I: Consent Form

Title of the Study: The contribution of Data Analytics in predicting the future purchase intentions of consumers.

Invitation to Participate: I would like to invite you to take part in a research study. Before taking part, please take your time to read the following information.

Introduction and Purpose: My name is Pinakshi Kalita and I am a final semester MSc Management student at the National College of Ireland (NCI). Currently, I am researching on the role of Data Analytics in predicting the future purchase intentions of consumers for the purpose of my dissertation, to be completed as a part of my postgraduate degree of Master of Science in Management.

Confidentiality: The data will be collected through audio recording interview or through an open-ended questionnaire. All the data will remain confidential and anonymous, will only be accessible to me and my Supervisor and will be deleted following the completion of my thesis as per the guidelines of National College of Ireland.

Consent for Participation: I voluntarily agree to participate in this research study. I understand that even if I agree to participate now, I can withdraw at any time or refuse to answer any question without any consequences of any kind. I have had the purpose and nature of the study explained to me and I have had the opportunity to ask questions about the study for any clarification. I understand that all the information that I share will be treated confidentially. I understand that in any form of report on the results of this study, my identity will remain anonymous. I understand that a transcript of my interview in which all identifying information has been removed will be retained for two years

By signing this, I agree to voluntarily take part in this study.

Signed by Participant: Date:

Contact Details: Pinakshi Kalita

Email: x18115179@student.ncirl.ie

Appendix II: Interview Questions

Background Questions:

1. Please create a username for yourself.

2. What is your current role in your organization? Specify the number of years of work experience you have in your current role/industry.

Essence of Consumer Behaviour:

1. What, according to you, have been the top most important marketing processes over the years?

2. Tell me briefly about your experience with regards to changes in the marketing processes or techniques over the years?

3. For a company's success, should more interest be drawn towards identifying and understanding relationships among customers' purchases? Why?

4. How important is the study of consumer behaviour for an organization?

Contribution of Data Analytics:

1. What is the role of recent information technologies like data analytics in anticipating consumer behaviour?

2. Would you prefer to work on data-driven models or on theories and concepts to derive results? Why?

3. Do you think data-driven or machine-driven techniques have made it easier in enriching the consumer experience today? Or is it overhyped?

4. Do you find any challenges working with new techniques or tools? Can you describe a situation from your own experience.

5. Which data mining techniques do you frequently work with and why?

6. Have you ever come across a situation where you observed a contradiction between results from data-driven models and human predictions? Please explain.

7. Would you like to share anything from your end?

Appendix III: Sample of Participants' Responses

Participant's username	Background Question
	What is your current role in your organization? Specify the number of years of work experience you have in your
kum@jap	IT analyst. 4 years
EOG@IND	I am a senior data scientist at HDFC Life. I have 4 years of overall work experience. I work mostly on building computer vision based application for the company.
kar@ind	Analyst (2years)
ARIIND	Marketing Consultant (2 + years of Experience)
too@can	marketing consultant. I have been working in this field for 2.5 years
Tiwind	Marketing Consultant
gam@ind	Analyst. 2+ years
eka@IRE	Data Analyst
ani@rom	Data Scientist with 3 years experience

Essence of Consumer Behaviour		
Questions	Participants	
	pta@ger	eka@IND
What, according to you, have been the top most important marketing processes over the years?	Traditionally I believe, a more personal but time-consuming approach where analysts actively interacted with the target customers was productive and fruitful. But ever since the advent of data-driven predictive analysis of future purchasing trends, more and more marketing executives are using it in their daily work.	Although all the three options are really popular nowadays but among all, I think predicting future purchases holds a strong stand in the market.
Tell me briefly about your experience with regards to changes in the marketing processes or techniques over the years?	I have always used data analytical tools in my work, be it for analyzing the trends of the customers once or the shift in target customers for different industries. Before data-driven technologies using data mining and machine learning became the most common and effective aids to marketing, it was the traditional feedback and word-of-mouth techniques that most marketing executives used. The most striking difference or rather the advantage that predictive analysis has brought upon marketing research is the huge amount of data it encapsulates and analyzes in a very very small timeframe which is humanely just not possible. With such a sure amount of data backing up your analysis, a much more wider and varied analysis is possible. And data today is the most important and valuable source in the entrepreneurial world. Being able to fully unleash its potential into predicting the next generation of marketing trends is something that every marketing executive needs to consider and apply. This has been my experience and observation regarding the astronomical jump made from the traditional marketing processes to the latest ones.	With advancement in digitalization and information technology, I have seen tremendous changes in the way marketing techniques are applied today, mainly the manual work has been reduced dramatically.

For a company's success, should more interest be drawn towards identifying and understanding relationships among customers' purchases? Why?	Yes, it is the customers to whom every business caters. Customers can make or break a venture overnight, so the ever varying changes to customer purchases and what influences a customer to make a particular purchase needs to be scrutinized and analyzed with great detail, only then can a tangible correlation be made between the external factors that affect a customer's purchase and the products.	Definitely. Understanding customers tastes and preferences is really crucial to a business. A business will be successful in reaching out to their customers only if it is customer-centric.
How important is the study of consumer behaviour for an organization?	Customer behaviour holds great sway for an organization. It is one of the key influences upon any entrepreneurial venture. With great social media awareness, customers are now active participants in promoting products. So, marketing executives need to carefully examine, analyze and introspect what it is that a customer wants and what it is that he can relate to. His behaviour and responses with respect to parameters such as the current political, sociological or environmental situations need to be explored deeply.	If the organization wants to grow itself as a brand in the market, then it's really important for that organization to understand consumer behaviour so that it will help the organization to broaden its customer base

Questions	Contribution of Data Analytics	
	Participants	
	eka@IRE	pta@ger
What is the role of recent information technologies like data analytics in anticipating consumer behaviour?	Data is the new oil - A very recently famous technical proverb. With the rapid span of digitization, everybody is leaving digital footprints every single day. Technologies like Data Analytics and related tools help an organization to understand their customers like never before, which helps is enhancing customer satisfaction.	Data analytics has made it possible to correctly define consumer demands and behaviour for a long, foreseeable future in a way that businesses can now plan their long-term strategies around such predictions. With the enormous amount of data being collected every minute, consumer behaviour being very tangible and endurant conclusions can be drawn on how it might change over the course of the next decade or even more.
Would you prefer to work on data-driven models or on theories and concepts to derive results? Why?	Data driven models I would vote for. They are real-time based and result oriented. It has the ability to generalize itself with the changing patterns of the data and tell the story whenever required.	I have been working in predicting emerging global markets for past 3 years, totally based on data-driven models. Multi-dimensional data-driven models results have helped me to grasp and forecast what is the next big thing of if you look at a miniscule level, how influential are the social media influencers. From my experience, I would continue to work on data-driven approach in the future too.
Do you think data-driven or machine-driven techniques have made it easier in enriching the consumer experience today? Or is it overhyped?	Yes. We must agree that machine driven techniques have made it easier in enriching consumer behavior. Without data analytics big shopping events like Amazon Prime Day would not have been successful. I don't think they are over-hyped.	Data driven techniques are definitely not overhyped. If anything, it is being underutilized. Data is now a very precious commodity and it is the analysis of this billions of bytes of data that has enabled corporate giants such as Google, Amazon or Facebook to profit millions of dollars, strategize themselves for the next couple of decades and as such they have been so successful in reaching out to every nook and corner of a consumer's life. They have made very distinct predictions of customer preferences successfully and I think every venture should exploit data in a similar manner.
Do you find any challenges working with new techniques or tools? Can you describe a situation from your own experience.	Learning a new technology is exciting as well as challenging. There are very experienced professionals in the market which makes integration of a new tool into an organisation ecosystem not that smooth. Periodic training from the software vendors have mitigated this challenge till a certain extent though.	I think it's human nature to fear the unknown but as far as challenges to working with new technologies or tools go , I personally did not find anything so challenging such as the ones offered by Tableau which I use quite extensively,are very user-friendly and very easy to learn and to use.

Which data mining techniques do you frequently work with and why?	As a Data Analyst we apply multiple set of techniques. The first starts with exploratory data analysis, understanding the data and decrease the anomalies. Many other statistical techniques are then used for understanding the relations. Algorithms like Bayesian network, clustering etc are good to go for any data analyst.	I believe it mostly depends on the data but I have found that the traditional clustering method and its newer variations to give a very good analysis of how customer behaviour can be mapped to the present market segments and as such how such behaviour and future markets might change. Clustering helps me to distinctly define more or less clear boundaries to different customers , their inherent preferences and their hold upon a particular market.
Have you ever come across a situation where you observed a contradiction between results from data-driven models and human predictions? Please explain.	Don't remember such.	Till date, I have not come across any such situation where the predicted results were diametrically opposite to that of a human analysis. Instead, the data-driven models have aided in pushing the boundaries, opening up newer vistas in addition to the human predictions.
12. Would you like to share anything from your end? (optional)		No and thank you very much.

Submission of Thesis to Norma Smurfit Library, National College of Ireland

Student name: **Pinakshi Kalita**   Student number: **18115179**

School: **Business**   Course: **Management**

Degree to be awarded: **MSc in Management**

Title of Thesis: **The contribution of Data Analytics in predicting the future purchase intentions of consumers**.

One hard bound copy of your thesis will be lodged in the Norma Smurfit Library and will be available for consultation. The electronic copy will be accessible in TRAP (http://trap.ncirl.ie/), the National College of Ireland's Institutional Repository. In accordance with normal academic library practice all theses lodged in the National College of Ireland Institutional Repository (TRAP) are made available on open access.

I agree to a hard bound copy of my thesis being available for consultation in the library. I also agree to an electronic copy of my thesis being made publicly available on the National College of Ireland's Institutional Repository TRAP.

Signature of Candidate:

For completion by the School: The aforementioned thesis was received by_____

Date:_____

This signed form must be appended to all hard bound and electronic copies of your thesis submitted to your school.