Effectiveness of Market Segmentation techniques using Data Sharing in the Telecom industry

Master Thesis submitted to

Delft University of Technology, The Netherlands

in partial fulfillment of the requirements for the degree of

Master of Science

in Management of Technology

Faculty of Technology, Policy & Management

by

Maruthi Kashyap Gudipati

(Student Number: 4740246)

To be defended in public on Monday, September 30th, 2019 at 13:30 hrs.

Graduation Committee

Chairperson: Dr.ir. G.A. (Mark) de Reuver,
Section of Information & Communication Technology

First Supervisor: Dr. -Ing. T. (Tobias) Fiebig,
Section of Information & Communication Technology

Second Supervisor: Dr.ir. H. (Hadi) Asghari,
Section of Organisation & Governance

This page is intentionally left blank.

Acknowledgement

First and foremost, I wish to thank my Chairperson Dr.ir. G.A. (Mark) de Reuver and my supervisors Dr. -Ing. T. (Tobias) Fiebig and Dr.ir. H. (Hadi) Asghari for giving me an opportunity to work on my Master thesis under their supervision and guidance. Without their tremendous support and constant guidance, I would not have been able to complete this thesis.

Next, I would like to thank Mr.Wirawan Agahari for his constant support in helping me find respondents for this thesis study. In addition, I would like to thank all respondents who took time out of their busy schedules to share their views and experiences. Without their inputs, this thesis could not have been completed.

Last but not least, I wish to thank my parents and God for giving me an opportunity to study at this fantastic university that has taught me a lot in the last 2 years.

Maruthi Kashyap Gudipati

Delft, September 2019

Executive Summary

Problem Statement

Since the start of the 21st century, the amount of captured data has been continuously increasing in this digital age. With almost 2.5 quintillion bytes of data being generated and captured every day (Liang et al., 2018), researchers and companies have a strong interest in exploring the value that can be created with this data, called big data analysis. Also, since companies strive to be market leaders, they constantly evaluate methods/approaches to discover hidden trends/ potential opportunities. One method of finding hidden trends is through market segmentation, a process which can be defined as a division of a heterogeneous market into several smaller homogeneous markets to precisely understand the desires of consumers. Identifying and targeting the right consumers through market segmentation is highly dependent on the collected data. Due to the usage of obsolete data collection methods and privacy regulations, most often, companies only possess siloed data. If siloed data is used, then companies might not be effective with their segmentation strategies.

One way to ensure that data is complete and consistent might be through data sharing in a 'data market' between players to holistically understand the consumers. With this thought in mind, this thesis considers the telecom industry as an example and explores the effectiveness of market segmentation using shared data. The main research question of this thesis study is Before going deeply into the aim of the thesis, let's first understand the current problems of the telecom industry. Traditionally, telecom firms have generated revenue via three streams i.e. voice, messaging and data. However, over the past decade, the market has witnessed an emergence of Over the top content players such as Netflix, YouTube, and Amazon Prime. These players do not need any association with the telecom firms to provide their services and thereby have impacted traditional telecom companies' voice and messaging revenue streams. In addition to this with new content frequently being updated in these OTT services, customer preferences are constantly changing, and telecom firms are finding it hard to predict these varying needs with the siloed data present in their databases. This has therefore resulted in low average revenue per user (ARPU) levels for these telecom firms. As the first step in the thesis, we performed a literature review and identified four common segmentation techniques used by the telecom industry. The four techniques are customer value segmentation, customer behavior segmentation, customer lifecycle segmentation, and customer migration segmentation. These techniques are customer-centric and are heavily reliant on data for their effectiveness. To observe if these identified techniques are employed at the industry level and to contemplate the viewpoints of experts on data sharing in market segmentation, we interviewed market segmentation experts from the industry. The following subsections provide an overview of the interviews, questions asked in the interview and findings from the interviews.

Data Collection & Analysis

In order to understand state-of-the-art market segmentation techniques in the industry and to hear the opinion of the experts, we conducted 10 semi-structured interviews with market segmentation experts located across both developing and developed countries in the world. As part of these interviews, we asked

the experts' 8 questions out of which 6 questions were related to market segmentation and 2 of them were related to data sharing. All the interviews were recorded, transcribed, coded and analysed manually. To perform effective coding and find patterns in the interview responses, we first reduced the collected data and then visualized it to draw conclusions. The next subsection explains the findings obtained from the interviews.

Findings

Some of the salient findings by analysing the interview response are:

1. All the four market segmentation techniques reviewed from literature are also widely used in real-time in the telecom industry. Behavioural Segmentation is predominantly used by the majority of the telecom firms. The fundamental reason for employing this technique is because this method analysis customer behaviour thereby providing detailed insights about customer's wants, needs, and desires.

2. Due to the fierce competition present in the marketplace, it is imperative that firms build accurate predictive models. This has led to telecom firms employing a combination of market segmentation techniques to define precise segments. The choice of the combination of segmentation is often dependent on four factors. Size of customer base, the company's aim, and vision, the position of the company and current market trends are the main factors that determine the combination of market segmentation techniques. For e.g. a leading telecom operator in The Netherlands adopted needs-based and lifecycle-based segmentation to perform market segmentation. Such a combination helped them gain sustainable competitive advantage in the market as the combination provided a holistic understanding of customer needs.

3. All the respondents stated that internal data acquired via customer contracts, traffic data and billing size was predominantly used to perform market segmentation. Few companies used external data in a limited manner to validate their results of market segmentation and understand the current market trends. The reasons for not using external data is twofold. After conducting pilot projects using external data, few respondents felt that the value added by external data was very less in comparison to the high investment costs that were required to acquire the external data. The other reason for the limited usage of external data is related to trust and veracity. As the quality of data acquired plays a pivotal role in implementing an effective segmentation campaign, firms lacked resources and technology that could aid them in validating the veracity of the acquired data from 3rd parties.

4. Most of the interview respondents confirmed that challenges related to data cause roadblocks to the implementation of market segmentation techniques. In data-related challenges, companies witness incomplete data sets, lack of sufficient real-time data and scattered data. On the other hand, few respondents also stated that in addition to the data-related challenges, firms faced technical difficulties such as lack of state-of-the-art technological tools and lack of data analysis during the implementation of market segmentation techniques. Both these types i.e. data related and

technological challenges hamper the quality of the end result which is identifying precise customer segments.

5. Most of the respondents believed that data sharing has the potential to improve the effectiveness of market segmentation. This is because by pooling customer data from different sources can provide information from a 360-degree perspective and help telecom firms identify critical market trends. Few respondents disagreed with the concept of data sharing due to failure pilot project conducted in this regard. The main reason for the failure of pilot projects on data sharing was that the external data acquired did not provide significant value when combined with the internal data. Therefore, it is important that data shared should be of utmost quality and there should be an overlap between the data.

After analysing the interview responses, the next subsection explains our reflections from these findings along with ethical considerations and an application for a use case scenario.

Discussions

To provide a comprehensive understanding of data sharing, this thesis study illustrated a use case scenario between a bank and a telecom firm. We assume that both the actors are in collaboration and have provided access to each of their CRM databases to identify areas of cross-selling. Apart from the demographic and geographic data points, we identified 5 data points that a bank needs to possess for an effective data sharing with the telecom firm. These 5 data points are data based on customer spending/saving information, data regarding automatic transfers, data from ATM usage, data based on transaction records and data explaining the customer's credit scores. As this thesis primarily focussed on technical standpoints, to complement this view we analysed the ethical considerations of data sharing. This is because many companies try to use a "catch-all-you-can" approach to collect as much data as possible from the individuals with an intention to quantify aspects of everyday life in an attempt to benefit companies doing the analytics. Secondly, this data that is captured from individuals moves from one company to another until there is no further value left in the captured data. Thirdly, companies use algorithms on captured data to profile individuals on their race, gender, social status, date of birth, information in bank account which might conflict with the ethical aspects of right to freedom of individuals. Considering these three points we identified crucial ethical considerations that need to be contemplated from three perspectives organization, society, and individual.

Recommendations

By using the findings and implications that were obtained from this thesis, we recommend the following points:

1. In chapter 2 we identified 5 relevant market segmentation techniques for modern age telecom firms. As these techniques are customer-centric and allow customers to be a part of multiple segments, we advise telecom firms to employ any 2 techniques out of the 5 techniques that were explained. By employing them in combination we strongly feel that a firm will understand the value and retention drivers in a comprehensive manner and will be able to perform effective market segmentation.

2. Given the changing business environment in the telecom sector caused due to globalization and the emergence of OTT players, we recommend that modern age telecom firms should adopt the concept of data sharing in market segmentation. This is because by acquire large amounts of customer data, from various organizations, telecom firms will be able to:
    a. Identify critical market trends
    b. Reduce the customer segment size
    c. Develop novel products and services
    d. Improve productivity
    e. Strengthen business ecosystems
    f. Create novel business models
    g. Improve customer relationship management

3. With strict legislation like GDPR being in practice and as customer data involves sensitive information, if telecom firms fear trading of such data might pose a great risk of a privacy breach, we advise the firms to adopt Privacy enhancing technologies (PET's). PET's are a coherent system of ICT measures that protect privacy by eliminating or reducing personal data or by preventing unnecessary and/or undesired processing of personal data, all without losing the functionality of the information system. Examples of these technologies are Multi Party Computation (MPC) – which helps the data owners to compute a function over their inputs while keeping their inputs private.

Table of Contents

List of Figures

List of Tables

List of Equations

1. Introduction

One of the most common analogies that is very often used today is – "big data is the new oil" (Brownlow, Zaki, Andy, & Urmetzer, 2015). It is estimated that the total volume of data generated and captured every day is 2.5 quintillion bytes (Liang et al., 2018). In order to investigate the main reason behind firms collecting huge volumes of data, IBM conducted a survey across 26 industries over 95 countries. The results from the survey stated that integrating data into decision making can increase productivity, develop novel products and services, and aid in cost reduction (Schroeck et al., 2012). This would in turn help firms identify new areas of opportunities, monetize higher profits and target the right customers. For an efficient data-driven decision making it is imperative that firms possess the right data with utmost quality at the right time.

The potential benefits of applying data-driven decision making can be illustrated for various industries, however, the primary focus of the thesis is on the telecom industry. Over the past few decades, telecom firms across the globe have been constantly updating their business strategies and technologies. Until the past decade, telecom firms acquired revenue from two streams i.e. voice and messaging (Sujata et al., 2015). However, the rapid advancements of technology and the explosion of internet have provided customers easy access to high-speed internet. This has resulted in a shift of revenue streams from voice and messaging to data. The main reason for this shift can be attributed to the emergence of over the top players such as Netflix, YouTube, and Amazon Prime. These players do not require any partnerships with telecom operators in order to utilize their services as the only way to use OTT services is via a data pack. Recent studies conducted by Deloitte state that the global monthly data consumption is expected to reach 3.2 exabytes (EB) per month by the end of 2019 (Joshi, 2018). With this rapid growth of OTT players such as Netflix and YouTube, customer preferences are constantly changing and telecom companies are finding it hard to predict the appropriate customer trends (Syed, Halaoui, Hicham, Mahmoud, & Mansour, 2014). This results in telecom companies experiencing low average revenue per user and an elevated customer churn rate. This can be explained by taking the case of China Mobile, the world's largest telecom operator. Researchers state that the profits of China mobile have reduced by 9.4% (Sujata et al., 2015). The main reason for a decline in profits and an increase in customer churn is because telecom companies possess data in silos which do not provide a holistic view about customer behavior. Since these discrete data sets don't explain the complete picture, incomplete customer segments are identified and as a result, telecom firms fail to create a compelling customer experience. It is therefore imperative for telecom companies to embrace new methods of deriving information from data that will help them overcome these problems by identifying micro, precise and distinct segments. By doing this, telecom companies will be in a position to offer personalized and relevant services to the identified customers (Syed et al., 2014). Having understood the value of data in the decision-making process, let's get an understanding of the context of this research study.

1.1 Context of the Research

This thesis research is conducted as part of the Safe Data-Enabled Economic Development (herein: Safe-DEED) project. In this project, several companies from telecom, cryptography, data science, business model innovation and legal domains located across various countries in Europe collaborate together to foster a

data-driven economy in Europe (Lupu, 2018). This project is funded by the European Union's Horizon 2020 Research & Innovation Program. The main goal of this project is to develop a competitive environment in Europe where every individual or a firm is well informed about the value of data they possess and feels secure to use, share and operate this data (Lupu, 2018). Data markets play a crucial role in providing the right information but fail to get adopted due to inherent privacy and trust issues (Lupu, 2018). Safe-DEED believes that privacy-enhancing technologies will play a dominant role by the year 2025 and aims to develop these technologies based on a mathematical concept of multi-party computation (MPC). With the help of MPC, encrypted data sets can be computed. This cutting-edge technology can potentially overcome all the barriers that impede large scale adoption of data marketplaces and enable a decentralized data marketplace.

1.2 Scientific Problem & Knowledge Gap

In the earlier sections, we illustrated the benefits of data-driven decision making and stated that data marketplaces play a pivotal role in helping acquire the right information at an appropriate time as the main mechanism involved in these marketplaces is data sharing. We also discussed that the concept of market segmentation has been a widely used concept/tool among organizations in segregating target market and satisfying customer needs. Currently, there is a great amount of academic literature that explains the concept of market segmentation and how these segmentation techniques can be implemented using numerous data mining algorithms such as cluster analysis, factor analysis et cetera. For e.g. (Ye, Qiuru, Haixu, Yijun, & Guangping, 2013) explain how segmentation techniques can be improved in the telecom industry using cluster analysis. On the contrary, there is no research that has been conducted that explain how segmentation can be improved from the perspective of shared data. Typically, the telecom industry is considered as a data-intensive sector and market segmentation is widely used to try and identify new customer segments. Although there are a lot of research papers and articles on segmentation techniques in the telecom industry, there is no academic literature that explains the effectiveness of pooled customer data between organizations, on market segmentation for the telecom industry. Therefore, this thesis aims to fill this knowledge gap by researching the benefits of data sharing for market segmentation in the telecom industry. As part of this research we aim to contribute to the following subject areas

Data Marketplaces: Relating to the explanation we provided in Section 1.1, we understood that the main mechanism that occurs in data market places is data sharing or data trading. Data market places are a recent phenomenon and are slowly gathering interest among scholars to conduct research. Consequently, there exist numerous gaps in the literature that are associated with data marketplaces. For instance, there is no literature that explains the benefits of data sharing between organizations. To increase the adoption rate of data market places, it is imperative that there exists such literature that explains the benefits associated with data sharing as we realize that such would be extremely beneficial for companies inclining to adopt data marketplaces. For this very reason, we are conducting this research with an aim to fill this gap to the best of our ability.

Market Segmentation: The concept of market segmentation was introduced by Smit (1956). Since then there has been a great amount of literature being published that explain the fundamentals such as various types of segmentation, the process of segmentation and criteria to evaluate the segmentation solution. Over

the past decade with an increased presence of big data in our daily lives, researchers have published numerous kinds of literature explaining the implementation of segmentation techniques using different data mining concepts. However, there is no literature that elucidates the impact of data sharing on market segmentation. As the effectiveness of segmentation techniques are heavily reliant on data, we feel that it is important to understand and analyze the relationship between data sharing and market segmentation.

1.3 Research Objective

As argued in earlier sections, the fundamental reason by adopting obsolete methods of data collection is forcing the telecom firms to employ an ineffective market segmentation. In order to overcome this challenge, in this section, we explain our main goal that will be achieved in this thesis study. The main objective of this thesis is to evaluate the current market segmentation techniques employed by telecom firms and understand whether the concept of data sharing can help improve the effectiveness of market segmentation.

1.4 Research Questions

To meet the research objective as stated in Section 1.3, it is important to understand the concept of market segmentation and thereby identify relevant market segmentation technique in the telecom industry. Subsequently, it is necessary to explain the fundamentals of data sharing and the benefits associated with it. Finally, analyzing the impact of data sharing on the various aspects of market segmentation such as effectiveness, granularity et cetera. In this section, we list the main research question for this thesis and break down the main research question into smaller questions, known as sub research questions. Breaking it down in such a manner would help us focus on different aspects of the main research question, finally aiding us to answer the main research question comprehensively.

1.4.1 Main Research Question

By converting the above-mentioned research objective (Section 1.3) into a research question, the main research question for this thesis study is: **Can telecommunication companies improve the effectiveness of their segmentation techniques by sharing customer data with other organizations?**

1.4.2 Sub Research Questions

In order to comprehensively answer the main research question, we break it down into the following sub research questions and provide motivation for each of them.

The concept of segmentation was first introduced in 1956 and from then on there has been growing interest in adopting this concept in their daily operations by the telecom firms. Until the recent past, segmentation techniques such as demographic segmentation, geographic segmentation, psychographic segmentation have been traditional segmentation techniques that have been used. However, in the current era of big data and its analytics, there is an increased need for telecom firms to adopt customer-centric segmentation techniques and shift away from the traditional approaches. Therefore, it is important to relevant segmentation techniques and discusses aspects such as benefits, procedure, and challenges that are associated with these techniques. This has given rise to the first sub research question for this thesis study:

Sub Research Question (SQ) 1: What are the relevant market segmentation techniques for telecom firms?
As stated earlier, this thesis study focuses on two central topics data sharing and market segmentation. The aim of this study is analyzing the impact of data sharing on market segmentation. As SQ 1 dealt with market segmentation, it is now important to discuss the concept of data sharing and the benefits associated with pooling customer data from different organizations on market segmentation. Therefore, the second sub research question of this thesis study is:

SQ2: What are the benefits of sharing customer data between organizations for market segmentation?
The findings of SQ1 and SQ2 are derived from literature. As the next step in this research, it is important to investigate to what extent these theoretical findings are employed at the industry level. This has given rise to the third sub research question of this thesis study:

SQ3: Which segmentation techniques are used in the telecom industry and what kind of challenges were observed during their implementation?
To provide a holistic understanding to the readers on data sharing for market segmentation it is important to apply the findings obtained from SQ3 and illustrate a use case scenario between a telecom firm and an organization. As stated in Section 1.2, this thesis study is conducted as a part of Safe-DEED. This research project (Safe-DEED) partners with several organizations specialized in telecom, legal, data science across various countries in Europe. In order to help one of the partners of Safe DEED and solve their business problem, we chose a bank as one of the actors for data sharing with the telecom firm. Therefore, the fourth research question of this thesis study is:

SQ4: Which data structures should a bank possess for effective data sharing with a telecom firm?
The answers to the first four sub research questions are from a technical perspective. As sharing of customer data involves trading of personally identifiable information (PII) that has the potential to identify an individual. With strict regulations such as GDPR in practice now, it important to discuss the ethical considerations that need to be contemplated when organizations trade customer data for business purposes. This has given rise to the last sub research question of this thesis study:

SQ5: What are the ethical considerations for data sharing and market segmentation?
Therefore, the findings obtained from the technical perspective in the first four sub research questions will help us satisfy the formulated research objective and answer our main research question in a comprehensive manner. With so many ethical issues currently happening across the globe, the answers to the first four sub research questions along with SQ 5 will provide a complete understanding from a technical perspective and the ethical perspective.

1.5 Research Methodology

This section explains the highlights regarding the methodology that was adopted to answer the research questions and to achieve the objective of this research. An extensive explanation regarding the methodology employed in this thesis is in chapter 3. In this section we explain the methodology used to answer each of our sub research question in a brief manner.

SQ1: What are the relevant market segmentation techniques for telecom firms?

While conducting preliminary research on the current techniques employed by telecom firms, we found the paper "Customer Segmentation in the Telecommunications Industry" written by (Bayer, 2010). As this is an experienced-based journal, the information obtained from this literature is reliable, fair and useful. Therefore, this paper was used as a reference paper to find more literature that could help answer this question. The author of this paper identifies four commonly used techniques employed by telecom firms i.e. Customer Value Segmentation, Customer Behavior Segmentation, Customer Life Cycle Segmentation & Customer Migration Segmentation. Every dimension of these four techniques is analyzed and explained extensively in Chapter 2 of this thesis document. Relevant literature explaining the four techniques was found using keywords such as "life cycle segmentation", "behavior segmentation", "value segmentation", "migration segmentation", "data warehouse", "customer management", "customer segmentation" and "data-driven segmentation". The method of selection of relevant literature involved three rounds. The first round involved reading and analyzing the title, abstract and citation counts if they could be relevant and if found relevant, relevant contents of the selected literature were skimmed through, in the second round. In the last round, relevant skimmed literature that could aid literature review was read completely and selected. For this literature review we used three databases i.e. Scopus, Science Direct & Google Scholar to find relevant literature.

SQ2: What are the benefits of sharing customer data between organizations for market segmentation?

To answer this question, we analyzed the published literature on data sharing. The keywords that were used to identify relevant literature were "data sharing", "benefits", "data trading", "cross-selling" and "data merging" and "B2B data sharing". While selecting relevant literature, the same procedure explained in sub research question 1 was also employed for this question.

SQ3: Which segmentation techniques are used in the telecom industry and what kind of challenges were observed during their implementation?

This question was answered by interviewing segmentation experts across the globe. In Chapter 3 we explain in a detailed manner on why we chose interviews as the research methodology.

The theoretical findings derived from SQ 1 and SQ 2 were used as inputs for the interviews. The interviews were conducted in a semi-structured manner i.e. common questions were asked in all the interviews and specific questions were asked depending on the interest and expertise of the interviewee. The fundamental reason for choosing a semi-structured manner was to keep the conversation with the experts open-ended as this helps generate new ideas. All the interviews were recorded and manually transcribed. Analyses and interpretation of the patterns from the transcribed qualitative data is crucial to answering this sub research question.

Grounded theory and Qualitative content analysis are the two most common techniques used to interpret qualitative data (Cho, Lee, Ji Young, & Eun-Hee, 2014). Grounded theory is generally employed when the goal of the research is to develop a theory (Cho et al., 2014). Qualitative data analysis involves three steps

i.e. data reduction, data display and drawing conclusions (Sekaran & Bougie, 2016). The first step involves the reduction of data by assigning labels. This process is known as coding. There are three approaches namely grounded theory, tight approach and middle ground to perform coding. We chose middle ground approach as we intend to first develop a set of code from literature and validate them empirically. We analyzed the interview transcripts of each respondent and reduced the data by assigning keywords/labels to sentences and paragraphs. In this step we were able to validate our initial code list and update them with the responses provided. After completing the first step i.e. data reduction, we visualized the data in tabular format, thereby performing the second step i.e. data visualization. Finally, with the help of the reduced and visualized data we drew conclusions and reflected our findings to our research objective. Microsoft Excel was predominantly used perform the process of coding.

SQ 4: Which data structures should a bank possess for effective data sharing with a telecom firm?

This is a reflective question that applies the findings of the interview responses into a use case scenario. This is the use case scenario - with an aim to identify cross-selling opportunities, a telecom operator and a bank are in a collaboration to share each other's CRM data. As the telecom firms want to continually increase the market share, they want to identify such customers to understand their behavioral patterns and thereby target those customers with potentially interesting plans and offers. With the help of this use case scenario and SQ3 results we aim to identify data points that a bank needs to possess for a data trade with a telecom firm.

SQ 5: What are the ethical considerations for data sharing and market segmentation?

To analyze this question, we reviewed literature using the keywords "ethics", "golden rule of ethics" and "ethics in big data analytics". Like SQ1, the same procedure was used to find and read the most relevant literature. As data originates from individuals and companies use this data to segment markets, we identified and explained ethical aspects that are important for individuals, companies, and organizations. From the eyes of individuals, the ethical considerations are privacy, trust, awareness, and choice. For companies, the most important ethical aspects are data trading, ethical governance, reputation, and data quality. As there might be differences in the ethical interpretation between companies, it is important to have a regulatory body/ society to ensure that all stakeholders are aware and follow ethical considerations.

1.6 Research Framework

Figure 1 explains the research framework of the thesis study.

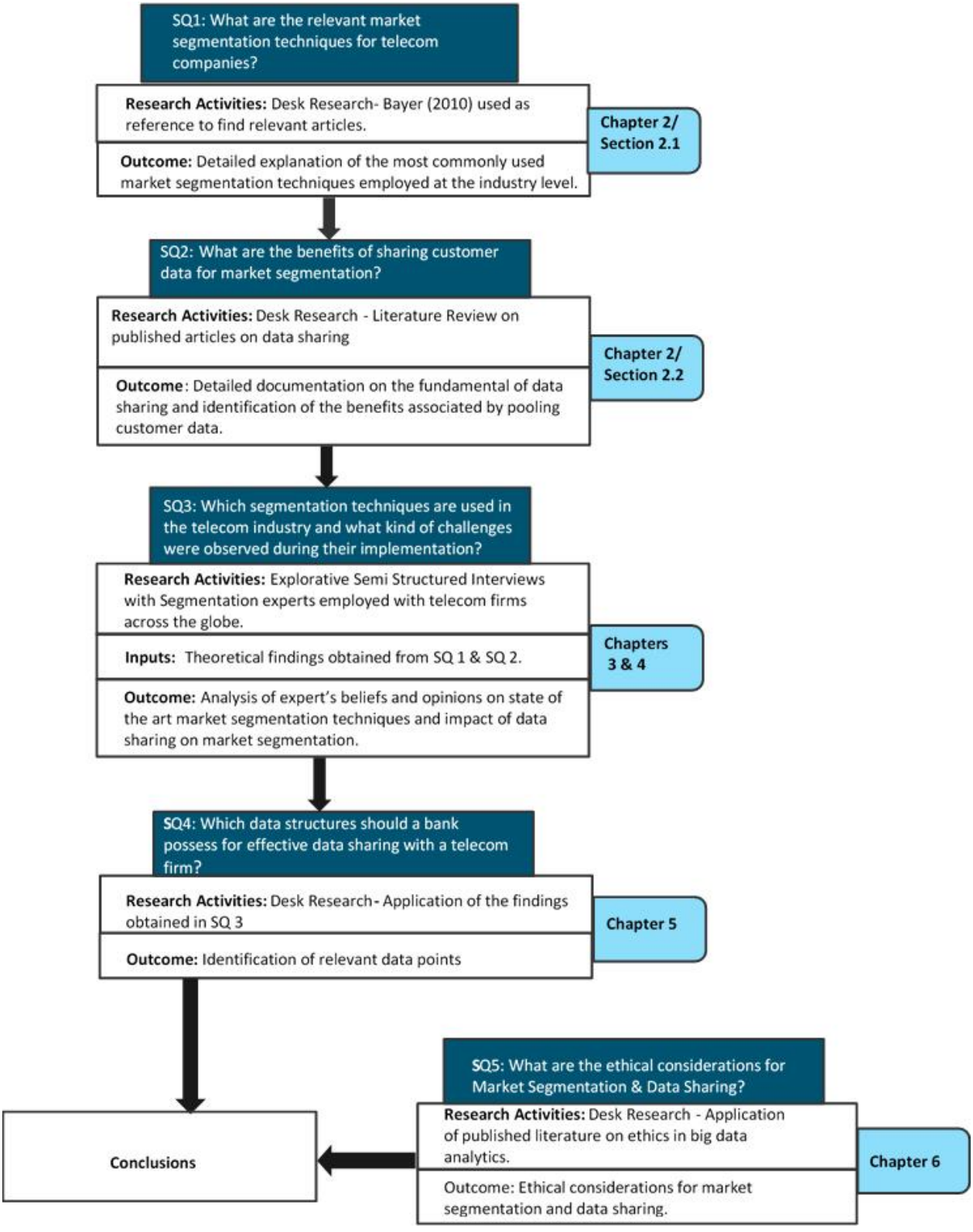

Figure 1: Research Framework Diagram

1.7 Structure of the document

This section provides an overview of the structure of this thesis document.

In **Chapter 2**, Theoretical background, we perform a literature review on the fundamentals of market segmentation, benefits of data sharing and relevant market segmentation techniques i.e. for telecom firms. **Chapter 3**, Research Methodology, explains the research methods used in this thesis. It also talks about the need for using interviews as a data collection method from industry experts and the types of questions that were asked in these interviews. These interview responses are analyzed and resulting inferences are concluded from these responses in **Chapter 4**, Findings. We apply the findings obtained in Chapter 4 and discuss an illustrative data sharing between a telecom firm and a bank, in **Chapter 5**. As data sharing involves sensitive, we describe the ethical considerations that need to be contemplated in **Chapter 6**. Finally, in **Chapters 7 & 8,** we conclude and reflect on our thesis.

2. Theoretical Background of Market Segmentation & Data sharing

Customers play the most important role in every business. In order to succeed and acquire high monetary gains, it is imperative for firms to understand the customers in every possible detail (Hamka, Bouwman, De Reuver, & Kroesen, 2014). This has led firms to embrace the concept of market segmentation in order to manage customers. However, globalization and increased competition in the market have forced firms to ideate new techniques to manage customers (Roland Ortt & Smits, 2006). In the era of big data, modern age firms find it extremely hard to survive in the market for a prolonged period of time using traditional market segmentation techniques (Goyat, 2011). Due to this, there has been a growing interest among researchers to analyze the concept of data sharing and its impact on market segmentation. In this chapter we explore the fundamentals of market segmentation, market segmentation approaches in the telecom industry and benefits and challenges of data sharing. The aim of this section is to answer the first sub research question: ***What are the relevant market segmentation techniques for telecom firms****?*

2.1 Concept of Market Segmentation

Marketing is ubiquitous and every firm invests a significant amount of resources to maintain and improve their marketing plans (Goyat, 2011). Companies perform marketing with a sole intention to identify customer's needs and desires. Such an identification helps firms manufacture products and services that aid in fulfilling customer needs. Employing a feasible, efficient & a practical marketing plan is imperative for a firm to generate higher returns (Dolnicar, Grün, & Leisch, 2018). Organizations have been employing traditional marketing approaches for a prolonged period. Email and television are some of the examples of traditional mass marketing techniques that have been used to reach out to a large target audience. Markets consist of a large number of customers with varying needs and desires and by employing mass marketing techniques, firms cannot satisfy the customer's needs comprehensively. Moreover, due to cut-throat competition in today's market, identifying a mass audience with the above-mentioned techniques does not provide significant returns to firms (Goyat, 2011). To sustain the competition, stay relevant and enhance their market share, modern age firms incline themselves to divide the market by identifying customers' with similar needs and focus on customers within their capacity (Syed et al., 2014). W. R. Smith, (1956) was the first researcher to introduce the concept of market segmentation. His definition is widely used and is also adopted in this study. Smith, (1956) defines market segmentation as "dividing a heterogeneous market into a number of smaller homogeneous markets in response to differing product preferences among important market segments. It is attributable to the desires of consumers or users for more precise satisfaction of their varying wants". Recent literature (Dolnicar et al., 2018) define market segmentation as "the process of grouping consumers into naturally existing or artificially created segments of consumers who share similar product preferences or characteristics".

Since its inception, there has been growing research and literature published on the benefits of market segmentation. With the help of market segmentation, companies can completely understand and evaluate

their customer satisfaction in the market, estimate and design their future strategic plans with ease and efficiency. This can be termed as the most important benefit of market segmentation (Dolnicar et al., 2018). As stated earlier, every customer's needs, wants and demands are distinctively different from other customers (Kotler & Zeithaml, 2000) explain that by embracing market segmentation, firms will be in a better position to identify customer's needs, wants and demands precisely and develop a value proposition that can satisfy all the varied customer preferences (Neeley, Hair, & Mcdaniel, 2011). Apart from identifying the needs, (Goyat, 2011) explains that by segmenting markets, companies can enhance their brand value, have a better customer reach and reduce sales resistance. (Evgeniou & Niessing, 2015) add on to explain that market segmentation is an iterative process that helps in identifying hidden trends and missed business opportunities for new products and services.

We also agree with the above-mentioned benefits and believe that segmentation can help firms unlock hidden market potential, thereby creating probable profits. It is therefore not an understatement to state that market segmentation is a basic building block for efficient marketing (Dolnicar et al., 2018). To holistically learn about market segmentation, it is important to understand the steps involved in segmenting a market.

2.1.1 Process of Market Segmentation

The process of identifying potential market segments is explorative in nature and is performed using statistical tests (Dolnicar et al., 2018). As every firm has different goals and objectives to achieve, methods and variables employed to segregate the market and identify potential market segments are subjective (Goyat, 2011). However, the process of market segmentation is the same irrespective of the methodology employed and the variables used. In Table 1, we describe the popular 10 step market segmentation process proposed by (Dolnicar et al., 2018).

Table 1: 10 step Market Segmentation Process

Step Number	Name of the step	Definition of the processes in the step
Step 1	Deciding to segment	Current market conditions need to be researched and companies need to evaluate if long term positive results can be generated.
Step 2	Specifying the ideal target segment	Companies need to describe their ideal market segment. This step also involves choosing segmentation bases.
Step 3	Data collection	Various characteristics (e.g. demographics, behavioral, geographical, psychographic, etc.) of customer data are collected using interviews, surveys, focus groups, etc. Collected data forms the foundation to identify market segments and hence the quality of customer data is very important.
Step 4	Exploration of data	Collected data needs to be preliminarily explored to gain insights into market segmentation approaches. This step helps in identifying the most suitable market segmentation algorithm.
Step 5	Extraction of segments	Similar data sets are clustered together, and common patterns need to be observed. This would be beneficial in extracting customer segments.

Step 6	Profiling segments	Key characteristics with respect to segmentation bases of the extracted market segments need to be identified. By identifying key characteristics, differences in the segmentation variables identified in Step 2 can be determined. Profiling can be done individually for a segment or by comparing it.
Step 7	Describing segments	Tasks in Step 6 and Step 7 are similar. However, with the key characteristics identified in Step 6, a more detailed description of the segments is done using additional information.
Step 8	Selecting the target segments	This step involves strategic decision making. Companies chose their target segments from the list of extracted segments.
Step 9	Customizing the market matrix	Companies use the inputs of Step 6 and Step 7 to understand the needs of customers and prospective customers. With this data, they develop a relevant and attractive marketing mix for the target segment.
Step 10	Evaluation and monitoring	As market conditions are volatile, companies evaluate the effectiveness of the market segment solution by validating the solution with the help of the market segmentation criteria.

From the above-mentioned points, we can infer that each of the 10 steps involves numerous tasks. The result of each of these tasks determines the quality of the identified market segment. Therefore, it is important that firms perform each of the steps diligently.

Having understood the need for market segmentation and 10 fundamental steps of any market segmentation, as part of this thesis, let's focus the commonly cited market segmentation techniques relevant to the telecom industry. The telecommunications market is often considered as the most volatile and dynamic and highly competitive market (Namvar, Ghazanfari, & Naderpour, 2018). The next section aims to identify relevant market segmentation techniques for the telecom industry.

2.2 Market Segmentation Techniques in Telecom Industry

Competition is relevant in all industries, and it can be attributed to factors such as advancement in technologies, increased awareness among customers and economic globalization (Xevelonakis, 2005). With a severe competition and dynamic nature of the market, the concept of market segmentation has been practiced in the telecommunication market for several years (Bayer, 2010). Over the last decade, with an intention to increase their customer base, telecom firms have estimated the aggregate demand, a term defined as the total number of services and products demanded by the customers in a given time period, of the market and segmented their customers using traditional approaches such as demographics, geographical, psychographic segmentation (Syed et al., 2014). However, recent studies show that by understanding the aggregate demand and employing traditional approaches, telecom companies are unable to understand distinct customer needs in a comprehensive manner. Therefore, telecom firms are inept to sustain the growth levels and are facing low average revenue per customer (ARPU) levels (Namvar, Ghazanfari, Naderpour, 2018).

In order to overcome the above-mentioned challenges, there has been a growing interest among modern age telecom firms to employ Customer Relationship Management (CRM) as a principal approach when

performing market segmentation (Kim, Jung, Suh, & Hwang, 2006). CRM "is the strategic application of people, processes, and technology in an organization-wide focus on improving the profitability of customer relationships"(Fadhil Mohammed, 2013). However, literature published on the current market segmentation trends in the telecommunications industry is sparse. Bayer, (2010) is the only researcher to publish an experience-based article on the current segmentation trends in the telecommunications sector. and she identifies four customer segmentation techniques in the telecom industry. The four techniques are:

1. Customer Value Segmentation
2. Customer Behaviour Segmentation
3. Customer Lifecycle/ Life stage Segmentation
4. Customer Migration Segmentation

2.2.1 Customer Value Segmentation

Over recent times, "customer loyalty" has been termed as the most crucial aspect of marketing. Customer loyalty can be defined as "a deeply held desire/ commitment by a customer to re-buy or re-patronize a preferred product consistently in the future" (Lin & Wang, 2006). The main reason that can be attributed to the significant growth of this term among modern age firms is the 80/20 marketing costs rule. The 80/20 rule explains that "80% of a firm's profits are contributed by top 20% profitable customers and 80% of a firm's costs are contributed by top 20% of unprofitable customers" (Kim et al., 2006). Hence, firms invest significant time and effort to retain customers (Namvar et al., 2018). Recent articles by Harvard Business Review support this claim and explicate that if a firm increases their retention rate by 5%, the profits generated have the potential to increase from 25% to 95% (Fadhil Mohammed, 2013). For these reasons, telecom firms incline themselves to segment profitable customers by calculating a value measure for every customer which is termed as customer value (ShuiXiu & Shuihua, 2012), (Lin & Wang, 2006) and thereby adopt customer value segmentation (Kim et al., 2006). Different researchers use different terminology to explain the term customer value. Some of the most commonly used terms are lifetime value, customer equity and customer profitability (Kim et al., 2006). The crucial steps involved in customer value segmentation are explained below:

Step 1: The initial step involves calculating the customer value measure for every customer. Before calculating this value measure, firms evaluate the past value and predict the future value of every customer (Han, Lu, & Leung, 2012). The past value is calculated by adding all the monetary contributions made by the customer to the overall profits of the firm until a specified time frame. The formula employed to calculate the past is shown below:

$$Past\ Value = \sum_{t=1}^{t=n} overall\ contribution$$

Equation 1: Past Value Formula

In computing the future value of customer 'i', the firm evaluates the current relationship of the customer with the firm and predicts the future monetary contributions a customer can make using below formula:

$$Future\ Value_i = \sum_{j=1}^{j=n} prob_{ij} \times profit_{ij}$$

Equation 2: Future Value Formula

The first term in the above-mentioned equation i.e. 'prob$_{ij}$ 'calculates the probability a customer 'i' will use 'j' number product or services from the 'n' offered products or services by the company. The second term,' profit$_{ij}$ ' calculates the profits the customer *i* will contribute to the company when using j number of products and services (Kim et al., 2006).

Summing up both the values i.e. past value and future value, firms determine the customer value associated with every customer (Bayer, 2010). Equation 3 is the formula employed to compute customer value.

$$Total\ Value = Past\ Value + Future\ Value$$

Equation 3: Customer Value Formula

Step 2: Decile analysis is a statistical tool that divides the entire customer base into 10 equal-sized groups, also known as deciles (Barry Harmsen, 2010). After conducting the first step, using the computed customer value for each customer, firms adopt decile analysis to split the entire customer base (Bayer, 2010). The topmost decile consists of the top 10% of profitable customers. The last decile consists of 10% of unprofitable customers (Bayer, 2010). Figure 2 illustrates the segmentation of the customer base using customer value and decile analysis.

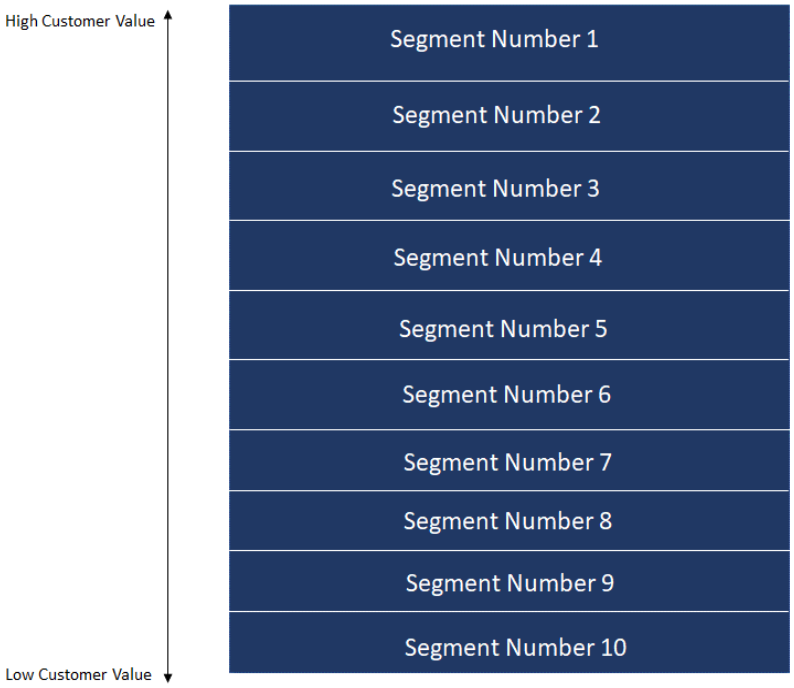

Figure 2: Division of Customer Base using Customer Value

Step 3: Firms use the identified deciles from step 2 to describe customer profiles in detail. Such a description helps a firm to carefully monitor the needs of every decile and develop a relevant marketing mix that has the potential to satisfy the needs of various decile segments. (Bayer, 2010).

Frameworks and ontologies have been used for many years to help firms develop comprehensive strategies based on customer value. (Kim et al., 2006)'s framework is widely accepted across the telecom industry and is segregated into three phases. The first phase involves data collection and illustrates the steps that are required to be performed by firms in acquiring good quality data. The second phase deals with mathematical operations that are required to calculate the customer value measure for every customer. The final phase is a strategic phase and explains the customer analysis and marketing strategies that are required to help the firm develop a compelling marketing mix. Figure 3 illustrates the framework proposed by (Kim et al., 2006).

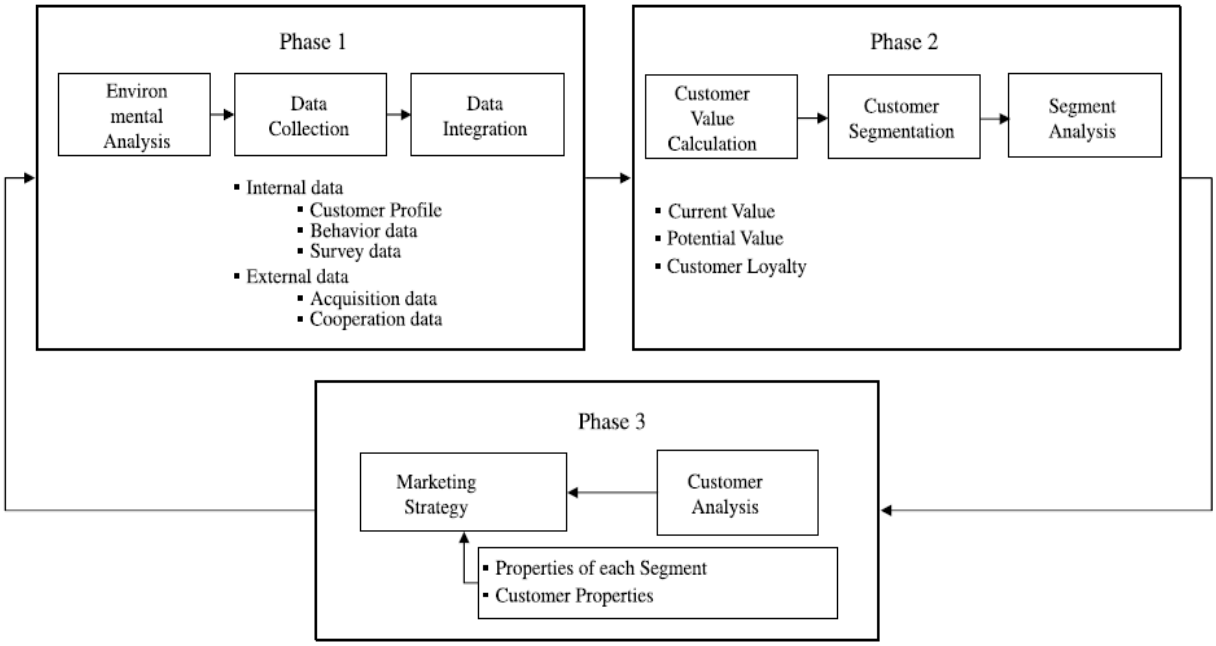

Figure 3: Lifetime Value Framework

While analysing the huge benefits of customer value segmentation, we believe that there are three main challenges that can impact the result of customer value segmentation. First, as the collected customer data is principally used to conduct this segmentation, the success of customer value segmentation is dependent on the availability and quality of data possessed by the firm. Good quality customer data can yield precise calculations of customer value and bad quality data can hamper the end solution. Second, for effective implementation of customer value segmentation, it is imperative for firms to clearly define the meaning of their profits, marketing costs, revenue et cetera. The third and last challenge is associated with the usage of future value when computing the customer value measure (Bayer, 2010). Recent studies conducted by PwC revealed that the telecommunications sector is reaching its saturation point rapidly. Studies also explain that telecom firms are finding it hard to compete in the dynamic volatile market conditions(Syed et al., 2014), therefore it is imprudent to calculate future value for every customer and think too far ahead (Bayer, 2010). Section 2.2.2 describes Customer Behaviour Segmentation, a segmentation technique based on the behaviour of the customers.

2.2.2 Customer Behavior Segmentation

Over the past 20 years, the number of people using telecom services users such as the internet, cable TV, the mobile phone has drastically increased. Recent studies prove this statement by revealing that there are more than 4 billion telecom users around the globe (Verbeke, Dejaeger, Martens, Hur, & Baesens, 2012). To manage such massive amounts of customers and understand them in a comprehensive manner, telecommunication firms (hereby: telecom firms) are capturing every action performed by customers over the internet, SMS, call et cetera and are building huge repositories containing behavioural data of customers. As telecom services are used extensively by people in their daily lives, telecom firms are tapping on the ability to determine customer needs and develop relevant strategies from the collected customer behavioural data (Guest, Struthers, Jones, & Schachtel, 2015). Usage of customer behavioural data in market segmentation is termed as customer behavioural segmentation. The key aspects of Customer Behaviour Segmentation are explained below:

- As stated above, firms adopting customer behavioural segmentation use historical behavioural data stored in the database as a primary resource. For telecom firms, the sources for behavioural data are event detail records and billing data (Maji & Sen, 2016). Event detail records or EDR is a piece of detailed documentation evidence that is generated in real-time when a customer makes calls, sends SMS and uses the internet from their devices over the network. An EDR contains various fields such as call duration, no of comments made on social media, location of user et cetera. In other words, an EDR provides complete information regarding occurred event on the network. Depending on the action performed by the customer over the network, EDRs can be divided into three subgroups such as call detail records, SMS details records, and internet usage records (Maji & Sen, 2016). Unlike EDR, Billing data is generated on a monthly basis (Maji & Sen, 2016) and contains invoices regarding service usage by customers.

- Upon collecting behavioural data of a customer, the next step employs data mining techniques on the behavioural data. Data mining can be defined “as a process that uses mathematical, statistical, artificial intelligence and machine learning techniques to extract and identify useful information and subsequently gain knowledge from databases” (Femina & Sudheep, 2015). To comprehensively understand the customer, these data mining techniques play a pivotal role (Ziafat & Shakeri, 2014). The most commonly used data mining techniques by telecom firms are clustering and classification (Ziafat & Shakeri, 2014). These techniques provide in-depth information about customer behaviour, customer satisfaction, customer usage (Ziafat & Shakeri, 2014). By employing data mining techniques, firms obtain crucial differences in customer behaviour helping them to identify novel patterns and relationships in the collected data (Sammour, Schreurs, & Vanhoof, 2009). Data mining also helps in identifying novel customer segments (Bayer, 2010). A typical data mining solution is shown in Table 2.

Table 2: An Example of a Data Mining Solution

Segment	Characteristics	Unique Segment Name
Segment 1	High voice usage, Low message, and data usage	Talker
Segment 2	Low voice and message usage, High data usage	Generation Z
Segment 3	High usage of the message, Low voice, and data usage	Message Maniacs
Segment 4	High voice, message, and data usage	All Rounders

- By segregating customers as shown in Table 2, each customer can be identified based on their needs and desires. This will help telecom firms to develop and employ numerous varied strategies based on the characteristics of the segments. Secondly, due to the comprehensive understanding regarding customer needs, firms will be in a better position to offer personalized services, use appropriate distribution channels and optimally allocate resources (Ziafat & Shakeri, 2014).

2.2.3 Customer Life Cycle/Stage Segmentation

As stated earlier, to compete in the market, modern age telecom firms have to employ customer-centric market segmentation techniques rather than rely on traditional methods of segmentation. (Ansell, Harrison, & Archibald, 2007). To adopt a viable technique, understanding customer behaviour, improving customer loyalty with the firm plays a fundamental role in determining the success of the chosen market segmentation technique (Shanks & Tay, 2001). As time progresses, people undergo changes in their social roles. The transitions between the social roles have an impact on the behaviour and loyalty exhibited by the person. Therefore, it is important to identify and understand the current social roles of customers. The third technique of market segmentation i.e. customer life stage segmentation considers a snapshot of the current life stage of the customer and performs market segmentation by analysing the needs and interests in the chosen life stage (Bayer, 2010). As the life stage of a customer can take various forms, researchers have formulated two different types of customer life stages in this type of segmentation. First type analyses the different stages a customer undergoes during their association with a firm. This approach is known as Customer life cycle segmentation (Bayer, 2010). The second type ponders over the phase/ stage a customer is currently undergoing in their individual lives (Guest et al., 2015). It is known as Customer life stage segmentation (Bayer, 2010). Regardless of the two different types, the premise of this segmentation technique lies in segmenting the varying customer behaviour exhibited at different life stages. The two types are explained in detail in the following subsections.

- #### Customer Life Cycle Segmentation

  The concept of 'life cycle' is not new in marketing and has been introduced in the 1930s (Wells & Gubar, 2006). Firms across various industries have embraced this concept to analyse sales growth when a new product is introduced in a marketplace (Levitt, 1965). The concept of customer lifecycle

and product life cycle are similar, however the former analyses the behaviour exhibited by the customer during the association with a firm (Wells & Gubar, 2006). To analyse customer behaviour at different stages, telecom firms split their customer base. Most of the telecom firms employ a customer life cycle model that divides the entire customer base into four segments i.e. New, Growth, Maturity, Decline. For e.g. when a firm acquires a customer, he/she is termed as new customer and their lifecycle with the firm is initiated. As time progress the customer progresses to be a growth customer to a matured customer and then finally a declined customer. At different stages in the customer lifecycle, telecom firms understand and analyse the needs, drivers, service usage levels and relationships of the customer. With the help of this analysis, firms design and develop appropriate services, products and offer. Generally during the decline stage, a firm develops and provides personalised services to the customer to ensure less customer attrition. By dividing the customer base into 4 segments based on their relationship, helps firms to identify needs accurately and satisfy the customer demand's comprehensively (Bayer, 2010). Figure 4 illustrates the customer life cycle model used by telecom firms.

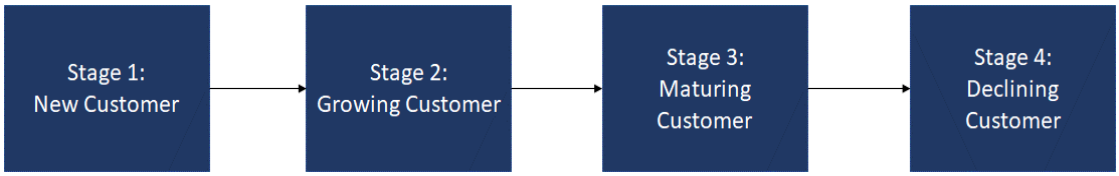

Figure 4: Four Stages of Customer Lifecycle

- ## Customer Life Stage Segmentation

The word 'change' is inevitable in one's life. As time progresses, every person experiences phases in their lives such as becoming a student, acquiring a new job, getting married, turning parents, etc. (Guest et al., 2015). Telecom firms define these phases of life as life stages. Prominent life stage events trigger a change in various aspects such as surroundings, priorities, buying patterns, needs, desires, and interests et cetera for a human being (Bandyopadhyay, 2016). These varying requirements at different life stages impact the usage of services and products provided by the firms and thereby create new marketing opportunities for firms (Mindi, 2016). To summarise, the primary focus of this segmentation technique is on the current life stage of an individual person. This is because life stages of humans predict and explain the requirements, needs and purchasing behaviour comprehensively than simple demographic data such as age, gender, family status (Brennan, 2017). Researchers have conducted empirical research to analyse the changing behaviour at different life stage. Recently, Deloitte conducted a survey to analyse the telecom service usage levels by women. It was found that their telecom services usage levels such as watching TV, surfing Internet reduced after turning into a parent (Guest et al., 2015). By obtaining such vital trends helps firms develop relevant services and products.

2.2.4 Customer Migration Segmentation

Firms across various industries suffer from high voluntary customer churn or customer attrition rates (Kuusik & Varblane, 2009). Customer churn occurs when a customer decides to stop using products or services offered from a company and switches his/her loyalties to other competing companies available in the market place (Molly Galetto, 2016). Due to the increased competition in the telecommunications industry, telecom companies are experiencing a high customer churn rate (Ahn, Han, & Lee, 2006). Studies show that, the annual churn rate in this sector range between 10% to 67% (Kuusik & Varblane, 2009). Such high levels of customer attrition have negative impacts on various aspects of a telecom company. Some of the adverse impacts include loss of average revenue per user, decreasing sales and profit and difficulty in acquiring new customers (Ahn et al., 2006). Owing to these reasons, modern age telecom firms incline themselves to understand and analyse the plausible factors that cause customer defection. This has led to the growing importance of the concept – customer migration behaviour (Nimako, 2016). The key aspects of this type of segmentation are explained in points below:

- There have been variations in the literature regarding the explanation of the concept -customer loyalty. Researchers have focussed on behaviour aspects such as customer purchasing patterns, frequency of purchase, etc. and emotional aspects like attitude towards brands in order to explain the concept of customer loyalty (Kuusik & Varblane, 2009). This segmentation technique uses the concepts of customer value segmentation, described in 2.2.1 as a preliminary point to identify loyal customers (Bayer, 2010). As a short summary, customer value segmentation focuses on calculating a value measure (customer value) for every customer and employs decile analysis to segment customers based on the calculated customer value. In customer migration segmentation, the customer value measure of each customer is observed at various time frames.

- As time progresses, it is very common for customers to increase or decrease their loyalty. By analysing at different time frames, customer value measure can increase or decrease i.e. customers can migrate between different segments, customer satisfaction, loyalty patterns can be identified (Bayer, 2010). Such identification of satisfaction and loyalty patterns can help firms to predict churn before it occurs. Therefore, telecom firms can design and develop compelling activities for customer segments who have a high probability to switch loyalties to other competing companies (Ascarza, 2018).

The most important challenge associated with this type of segmentation is the choice of time frames. This is because in choosing short time intervals, telecom companies would not be able to identify comprehensive migration patterns of customer between segments. Choosing long time intervals can deprive telecom companies to strategize and develop an attractive action to reduce customer churn. Therefore it is important for telecom companies to select time intervals with utmost care (Bayer, 2010).

Bayer, (2010) goes on to further add that with volatile challenges in today's marketplace, it is imperative for telecom firms to adopt multiple types of segmentation techniques. Practising such a vast number of segmentation techniques will result in two advantages for the telecom firms. First, the firms will be able to

identify numerous microsegments of customers. Such identification will help telecom firms target the right customer at the right time. Second, customers and prospects can be a part of various microsegments. By being a part of various groups, there is a greater probability that distinct value creation and customer retention drivers are identified. This identification would help firms develop customer-centric products and services. Arguing on similar lines, recent research conducted by Deloitte explains the need for telecom firms to investigate deeper to understand the factors that affect a customer's purchase. These insights will help the telecom operators identify microsegments of potential customers, personalise services, develop products that cater to the customer needs and thereby increase revenues (Guest et al., 2015) (Bayer, 2010). In this section, four market segmentation techniques for telecom firms were introduced and explained in a detailed manner, thereby providing a holistic theoretical knowledge. Later, in Chapter 4, findings of this section will be used as inputs to analyze state-of-the-art techniques employed by companies in the real world.

2.2.5 Challenges in Market segmentation

As described in section 2.1, the fundamental purpose of market segmentation is to identify people who have similar needs and desires and develop an appropriate marketing mix that can satisfy these needs. Product, price, place, and promotion are referred to as marketing mix (Kotler & Zeithaml, 2000). There is a lot of interest among all industries to perform market segmentation as a means of realizing the potential. This has caused the main premise among managers to shift from pondering whether segmentation should be adopted to identifying the appropriate technique to segment customers (Bayer, 2010). However, managers and firms face numerous challenges when performing market segmentation (Syed et al., 2014).

Identifying and targeting the right customer with an attractive and relevant marketing mix is highly dependent on the empirical data that is collected by firms (Harting, Sommer, & Schlosser, 2017). Empirical data often consists of customer behavior, purchasing history, interests, demographic variables, etc. Presently, the widely used technique by companies to gather data is through surveys (Dolnicar et al., 2018). Although it is the easiest method to collect data, there is a high possibility that such data can be unreliable in explaining customer behavior and can also suffer from biased opinions. In other words, most of the companies possess siloed information regarding their customer in their databases (Syed et al., 2014). Legacy systems (outdated computer systems, application programs, technologies that are being used are referred as legacy systems) (Brown, Kanagasabai, Pant, & Serpa Pinto, 2017) and privacy regulations such as General Data Privacy Regulation (Arnaut, Pont, Scaria, Berghmans, & Lecont, 2018) can be attributed to the problem of siloed database. These issues can hamper the quality of solutions generated (Dolnicar et al., 2018). As predictive models and strategies rely heavily on data, due to these siloed databases, firms fail to develop comprehensive strategies. Due to the lack of attractive and efficient strategies, firms experience high churn rate (Brown et al., 2017) and firms often end up identifying large and indistinct customer segments (Syed et al., 2014). Since every customer has varied needs and desires, a weak marketing mix fails to satisfy all the needs of customers comprehensively.

Digital platforms such as data marketplaces are one viable option that can help companies acquire relevant information about customers. As data marketplaces deal with data sharing, let's understand the concept of

data sharing and benefits of data sharing for market segmentation in the following sections. The next section aims to answer the sub research question 2: ***What are the benefits of sharing customer data between organizations for market segmentation?***

2.3 Data sharing and Benefits of Data Sharing

We live in a customer-centric world and it is highly imperative for firms to leverage customer data in order to understand the customer completely and develop a compelling marketing mix (Desai, 2019). Customer data collection and sharing between firms has been practiced for many years (Norman, Pepall, Richards, & Tan, 2016). Behavioral data, market research data is often shared between firms. Recent surveys conducted by the European Commission show that 32% of the firms in Europe have been sharing data for over 10 years (Arnaut et al., 2018). The major difference between the past and current trend in data sharing is the amount of data that is captured and shared. With an increasing dependence on smartphones, household devices, smart homes, and autonomous cars large amounts of data are being captured in the modern age (Brownlow, Zaki, Andy, & Urmetzer, 2015). The European Commission defines data sharing as "the process by which a company makes data available to another company that is neither a direct market competitor nor a sub-contractor and is interested in these data for its own business purposes. The company that shares data may either do so willingly or as a result of a legal obligation, and the transaction can either be made for free or entail some remuneration or compensation, including the provision of a service" (Arnaut et al., 2018). Actors who collect, aggregate and maintain millions of information regarding every minute aspect of an individual customer's behavior and transactions are known as data brokers (Norman et al., 2016). Google search engine and Facebook are the most common examples of data brokers (Clavorà Braulin & Valletti, 2016). With huge volumes of data being shared, this has resulted in an increasing demand for data brokers (Norman et al., 2016) and in emergence of business models that involve data-driven decision making (Clavorà Braulin & Valletti, 2016).

Recent studies show that "30 billion pieces of information are shared on Facebook every month" (Manyika et al., 2011). The main reason for abundant sharing can be attributed to two reasons. First, availability of modern technology allows firms to manage and act on the data captured and shared. Second, with the need to sustain the competition among rivals and stay relevant in the marketplace, firms incline themselves to capture maximum information about customers (Smit, 2018). As stated earlier customer data must be employed to develop efficient strategies, innovative solutions and increase productivity growth. In a recent survey conducted by McKinsey, it was found that leveraging customer data to extract insights of customers can help firms to improve their sales growth by 85% (Brown et al., 2017). Manyika et al, (2011) state that big data creates value for companies in many ways such as enhances transparency which results in cost reduction and quality improvement of the goods and service offered, segments customer preferences which help companies develop customizable actions, enables companies to experiment in new areas enabling them to create differentiation and helps innovate business models to capture value more efficiently. (Tan, 2017) agrees with the advantages that were mentioned above but also adds on to explain that analyzing customer data revolutionizes the way companies employ marketing strategies. This will pave way to techniques such as smart marketing and precision advertising thereby enabling companies target their potential customers

in a more efficient manner. Numerous benefits are associated when data is pooled together from different sources. These benefits are explained in points below:

- **Identification of Critical Patterns-** Behavioural data of customers can be captured from various sources such as banks, cafés, travel agencies, telecom operators, supermarkets et cetera. By pooling customer data from varied sources, firms develop a complete narrative about the behaviour of an individual customer. These narratives can be used to develop novel innovative solutions, formulate attractive pricing mechanism and reallocate budgets efficiently among various departments across a firm. These enriched customer profiles can also be used for data monetization (Brown et al., 2017).

- **Recognizing Missed Business Opportunities-** The term 'missed business opportunities' can be described as the potential business areas that were unexploited by a firm due to lack of relevant data. However, studies show that when data is shared among firms and third parties it is more likely for a firm to develop and deliver novel products and services using innovative methods to the target market (Arnaut et al., 2018). This is because companies can analyse gathered data and identify customer needs and wants with precision. Development of such innovative solutions can have a positive impact on the business model of a firm.

- **Reduction of Customer Segment Size –** As stated in earlier that due to the lack of access to relevant data, organisations often find random solutions to market segmentation and identify large, incoherent segments of customers. However, as data sharing helps in pooling relevant customer data, firms can develop a detailed profile of their customers and prospects. Such detailed profiling will help in identifying small and coherent customer segments (Neumann, 2017). This allows firms to develop an attractive marketing mix which can satisfy the needs and wants of customers. The helps to build stronger relationships with customers for a prolonged time period (Ernst & Dolnicar, 2018).

- **Eliminating Road Blocks for Better Sales & Services-** By gathering information regarding customers, a firm can overcome the obstacle of siloed databases. Comprehensive databases will be developed which will allow firms to develop efficient predictive models to group customers (Neumann, 2017). For example, if a bank provides access to their customer relationship management database with a telecom firm, the telecom operator can be in a better position to understand and identify the places where rich people reside. Such an identification would help the telecom operator to put relevant flyers in that locality and target prospective customers with premium plans. With the help of a comprehensive database, firms can reduce wastage of resources and enrich their sales and services (Neumann, 2017).

- **Improved Productivity -** Data integration helps in resource optimization. As stated earlier, leveraging customer data can improve productivity by 85%, this is the most crucial benefit of data sharing. By merging data from varied sources helps firms to devote less time to tracing of information and focus more on analysing the captured information.

From the above-mentioned points, we can summarize that data sharing helps firms identify microsegments of existing and potential future customers. Such identification can help a firm to improve its customer retention ratio & customer satisfaction, decrease churn rate and cost of service and increase the lifetime value of customers.

2.4 Conclusion

As mentioned in Section 2.1, in this subsection we understood the basics of market segmentation, data sharing and the benefits of data sharing for market segmentation. After reviewing the literature published, we learnt that market segmentation is a technique that divides the entire target market into small clusters containing customers with identical customer needs. We also understood the fundamentals of data sharing and found 5 important benefits (1) identification of missed business opportunities (2) eliminating roadblocks for sales and services (3) micro-targeting (4) identification of critical patterns (5) improved productivity. From these findings we develop a conceptual model as shown in Figure 5.

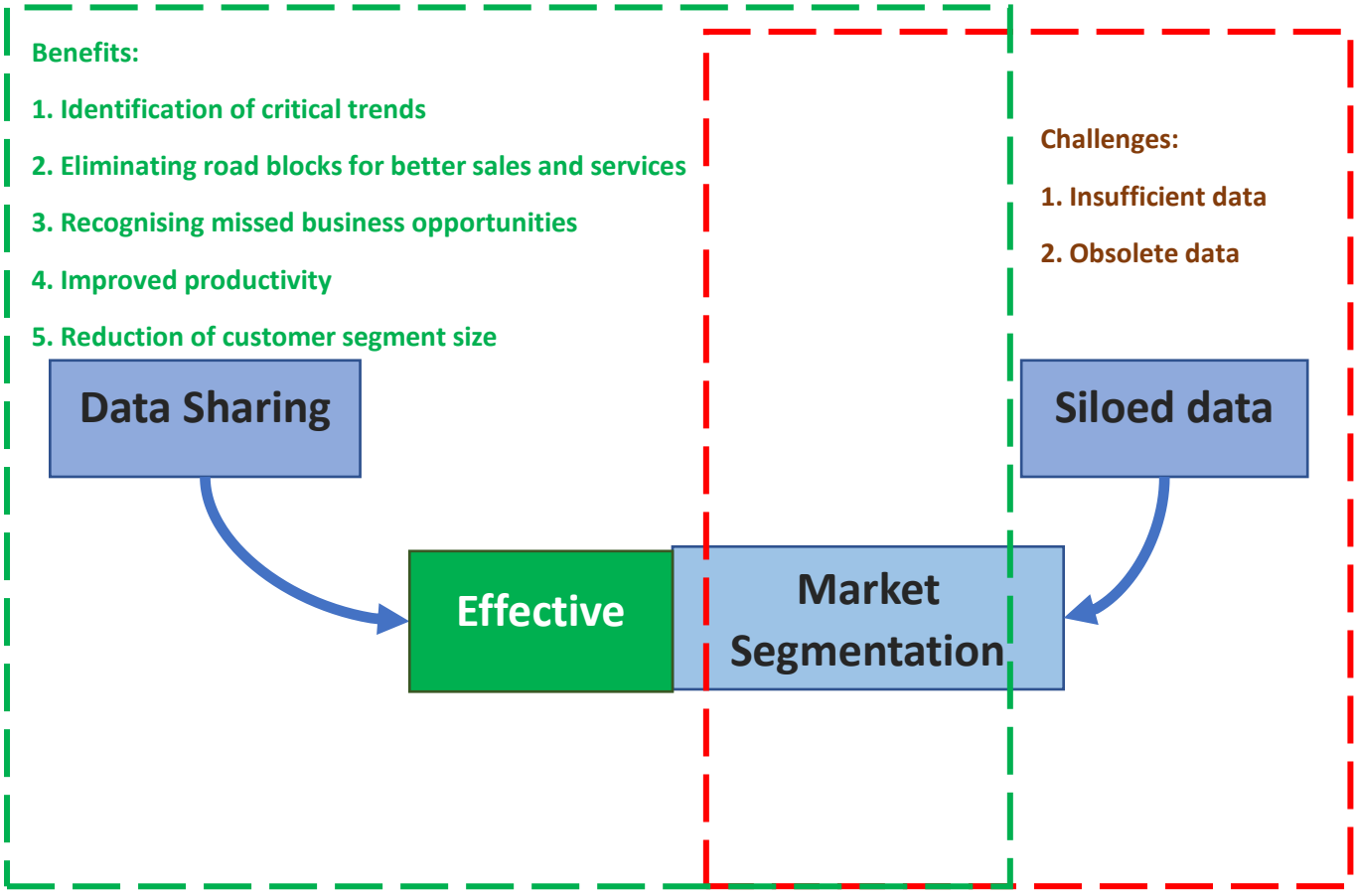

Figure 5: Conceptual Model

3. Research Methodology

3.1 Introduction

Every research topic contains an objective, which explains the purpose of the study. Based on the purpose of research, studies can be grouped into three types i.e. explorative study, descriptive study and explanatory study (Bhattacherjee, 2012). An explorative study aims to investigate a novel concept/phenomenon/problem and understand underlying reasons and factors for the occurrence of this novel concept/phenomenon/problem. Typically, explorative studies answer 'why and how' type of research questions. Descriptive research studies aim to identify crucial observations in the chosen area of research and create exhaustive documentation for the same. Research studies whose fundamental premise lies in explaining the main reasons for the observed phenomena are termed as explanatory studies (Bhattacherjee, 2012). As mentioned in the introduction chapter (Chapter 1), goal of this study is to better understand if market segmentation techniques can be used using the concept of data sharing. As the notion of sharing customer data in market segmentation is a new phenomenon, this study can be termed as an ***"explorative study".*** As a next step, we outline the methodology employed to research our goal of understanding effectiveness of data sharing in market segmentation.

Research methodology can be defined as "strategy or architectural design by which the researcher maps out an approach to problem-finding or problem-solving" (Jamshed, 2014). Research methodologies are divided into two categories i.e. quantitative and qualitative methodologies. The choice of methodology is dependent on the objective of the study and the research questions (Sekaran & Bougie, 2016). The researcher aims at exploring this concept in-depth in order to achieve the goal and answer the research questions in a comprehensive manner. Therefore, this research study employed a ***"qualitative methodology".***

3.2 Interviews as a Research Method

The most common data collection techniques in a qualitative methodology are interviews, focus groups, and workshops (Alshenqeeti, 2014). Out of these methods, data collection via interviews is often termed as "the most powerful method" (Kvale, 2006). In this thesis study we employed. **Expert Interviews** as a qualitative data collection method. Interviews are defined as "a conversation, whose purpose is to gather descriptions of the interviewee with respect to interpretation of the meanings of the described phenomena" (Alshenqeeti, 2014). There are two reasons that explain why interviews are "the most powerful" data collection method:

- In an explorative research study, interviews allow a researcher to examine the views of the respondent in a detailed manner (Kvale, 2006). Such an investigation helps the researcher to develop novel hunches/ideas in the area of study (Alshenqeeti, 2014).

- Interviews enable a medium of conversation in a natural setting. Such a medium of interaction aid the actors in the interview i.e. interviewee and the respondent to express their questions, thoughts, opinions, and beliefs in a cohesive manner thereby permitting the actors to have a detailed

conversation regarding the topic of discussion. Conducting a detailed conversation helps the researcher to avoid bias and acquire accurate data for analysis. (Alshenqeeti, 2014).

As part of this thesis, we interviewed many experts in the telecom industry and collected data and opinions from them with an aim to understand most commonly used market segmentation techniques. In addition, we also discussed the practical challenges faced by these market segmentation techniques. As part of the interviews, our aim was to cover the following three major sections:

- Due to a lack of a lot of academic literature explaining the current market segmentation techniques used by companies, the first aim for conducting interviews was to understand the state-of-the-art techniques employed by telecom firms across the globe and analyse our theoretical findings obtained in Chapter 2 with practical information.
- The second aim was to recognize the challenges faced by telecom firms during the implementation of their chosen market segmentation techniques.
- The last aim was to comprehend the expert's subjective opinion and beliefs on how sharing customer data can help overcome the encountered challenges and improve the effectiveness of the market technique.

Having understood the reasons for selecting interviews as a method of data collection and our aim for these interviews, let's now learn more about interview strategy which also explains the type of the interview, interview structure, and questions that were asked to the respondents.

3.3 Interview Strategy

3.3.1 Type of Interviews

This section explains the type of interview employed in this research study. In order to understand the type of interviews, as a first step, let's understand the underlying reasons that can affect the quality of interviews. There are two main aspects that determine the quality of the interviews. They are (1) quality of the questions (2) flow of the interview (Alshenqeeti, 2014). Questions that are discussed in an interview should be designed in a such a manner that maximum information regarding the topic of discussion is acquired thereby enabling the researcher to address the aim of the interview in a comprehensive manner (Gill, Stewart, Treasure, & Chadwick, 2008). The purpose of the interview and the number of questions to be discussed determine the type of interview chosen. Interviews in a qualitative study consist of three types i.e. unstructured, semi-structured and structured interviews (Sekaran & Bougie, 2016). Structured type of interviews contains a predetermined set of questions and provides a minimal degree of freedom for the researcher to put forward follow up questions. Structured interviews do not help the researcher to investigate the area of research in a detailed manner(Gill et al., 2008), (Alshenqeeti, 2014). On the other hand, unstructured interviews provide a greater degree of freedom as they do not contain predetermined questions. As this type of interview lack pre-decided questions, conducting interviews in a cohesive manner is a difficult task for the researcher (Gill et al., 2008). Due to the above-mentioned disadvantages associated with the two type of interviews, this thesis study used ***semi-structured interviews***. The main reasons for employing this type of interviews is that semi-structured interviews consist of a set of predetermined

questions that are generated after reviewing published literature in addition to providing freedom for the researcher to enquire follow up questions to the respondents (Alshenqeeti, 2014). As mentioned earlier, the purpose of the interviews was to understand the state-of-the-art market segmentation techniques employed and the challenges encountered during the implementation of segmentation technique in the telecom industry. By employing a semi-structured type, we were in a position to examine market segmentation techniques in telecom, in a detailed manner and ask follow-up questions to investigate an idea, thought, belief or an opinion that was generated during the course of the conversation, thus helping us build a compelling storyline. The next section describes the structure of the interview and the questions that were put forward to the market segmentation experts.

3.3.2 Interview Structure

This research study deals with two central topics in the telecommunications industry i.e. Market Segmentation and Data Sharing. In order to satisfy all the three aims of interviews that were mentioned in Section 3.2, the interview structure and the questions were developed in such a manner that both topics could be analyzed in depth.

In order to build the structure of the interview, we developed an interview protocol. Interview protocol explains the methodology used in every interview. For our interviews, we used this protocol:

- 4 days prior to the interview, an overview document containing information such as the aim of the interview, problem statement, and questions to be discussed were sent to every respondent.
- Every interview started with an introduction session where we discussed aspects such as academic background, research background and aim the interview and the interview structure.
- After the introduction session, we put forward the questions to the respondents and sought clarification during cases of misinterpretation.
- In the later stages of the interview, we explained the concept of data sharing in market segmentation using a use-case and sought the respondent's views on the same.
- At the end of the interview, we provided an opportunity to ask questions to the researcher and share feedback regarding the interview session.
- During the last stages of this thesis study we interview we sought consent from each of the 10 respondents by sending the consent form that is attached in Appendix A. All the 10 respondents granted consent for the following aspects:
    a. To record and transcribe the interview conversation.
    b. To compile and publish anonymously the results obtained from the responses.

In the following section, we elucidate the questions that were discussed during the interview session.

3.3.3 Interview Questions

Table 3: Interview Questions put forward to Respondents

Topic	Question	Reasoning for the question
Market Segmentation	Which segmentation technique is currently employed at your company?	To understand the state of the techniques and validate the theoretical findings obtained in Chapter 2.
	Which are the customer data structures used in this technique?	The answers generated from this question will help us in identifying the missing data points in a firm's database and analyze if these data points can be acquired via data sharing.
	How granular are your target market segments?	As one of the benefits of data sharing is identifying small microsegments of customers, it is important to understand the current size of market segments. This will help us in analyzing if the size could be reduced by pooling data from various sources.
	Are you satisfied with the results generated when this technique is used?	We asked this question to understand the inclination of respondents to improvements and innovation in the current techniques.
	What are the challenges encountered when this technique is used?	The resulting information from this question will help us in analyzing if the benefits of data sharing (obtained in Chapter 2) could overcome the challenges faced in implementing the segmentation technique
	What are various methods that are used by your company now in acquiring customer data?	There are two main reasons for asking this question. First, to understand the if challenges faced are due to the data collection methods. Second to check if data sharing has already been adopted.
Data Sharing	What do you feel on this concept of pooling customer data from various sources (data sharing)? Do you believe by adopting this concept, the effectiveness of segmentation techniques can be improved?	Data sharing for market segmentation is a novel concept. It is important to comprehend the expert's belief and opinions on whether data sharing could improve the effectiveness of market segmentation.
	Do you think that privacy preserving technologies such as Multi Party Computation will solve the segmentation challenges that were mentioned earlier?	As customer data involves personally identifiable information, it is important to comprehend the expert's view on the role of privacy preserving technologies in data sharing so that there is no violation of the legislation.

Following a semi-structured interview pattern, questions were developed in such a manner that both the research topics i.e. Market Segmentation and Data Sharing were investigated in a detailed manner and all

the three purposes of interviews were satisfied. A total of 9 questions were discussed during the interview sessions. Out of the 9 questions, seven questions investigated the market segmentation topic and two questions explored the data sharing concept. Table 3 describes the questions that were asked per topic.

3.4 Recruitment & Interviewer Selection

From the earlier sections in this chapter, we learnt that interview structure i.e. interview protocol, type of interview, questions discussed determine the quality of the data collected. However, there is another important aspect i.e. interviewee selection which also plays an important role in determining the quality of collected data. Owing to this, we enlisted a set of requirements for our interviews and we diligently followed these imperatives while selecting respondents for the interviews. The requirements are listed below:

1. Respondents should be associated with the telecommunication industry or must have conducted research in the telecom sector.
2. Respondents should have expertise in either market segmentation, business development, customer value management, market research analysis or product development and data analysis.
3. The researcher must ensure there is a diversity in respondents i.e. the researcher must aim to find, and interview experts associated with different telecom companies across the globe.
4. The researcher should aim to interview not more than one expert from a single company.

Following the above-mentioned requirements, the author used the contacts of the chairperson (Prof. Mark de Reuver) in this graduation committee and LinkedIn, to find and connect with experts from the above-mentioned areas of expertise. As part of this study, we conducted a total of ten semi-structured interviews. The average duration of each interview was 60 minutes. Due to the time and budget limitations the researcher conducted two interviews (Interview code: B and H) in person at the respondent's workplace. Remaining eight interviews were conducted on Skype. Table 4 explains the profile of all the ten experts who participated in this research study.

Table 4: Profile of the participated Telecom Experts

Code	Location	Role	Size of the Company
A	Belgium	Product Manager	Large Company
B	The Netherlands	Commercial Director	Small & Medium Enterprise
C	Greece	Data Analyst	Large Company
D	Indonesia	Segmentation Expert	Small & Medium Enterprise
E	Indonesia	Customer Service Manager	Large Company
F	India	Marketing Manager	Large Company
G	Finland	Researcher	University
H	The Netherlands	Consultant	Large Company
I	Indonesia	Sales Head	Large Company
J	USA	Retired. Managing Director	Large Company

3.5 Coding Technique & Analysis

Typically, data collected in a qualitative study is in the form of words. Such data can be found in interview/focus group transcripts, journals, and magazines (Sekaran & Bougie, 2016). 'Coding' is a technique that is used to analyze the collected data and identify patterns. In this process of coding, short phrases/ words are tagged along with large units of texts. These short phrases are known as codes (Saldana, 2009). Traditionally, there are three approaches to perform coding in a qualitative study. They are called as grounded theory, middle ground approach and tight approach (de Reuver, 2019). In grounded theory, codes are generated from the collected data. However, while analyzing, the collected risk factor is high in this approach as it does not provide a clear understanding of the researcher on what to look for. In the middle ground approach the initial code list is generated theory and subsequently the code list develops gradually while analyzing the data (de Reuver, 2019). It is often stated that "the excellence of the research rests in large part on the excellence of coding" (Saldana, 2009). The third approach i.e. tight approach contains a predefined code list and does not allow the researcher to improve the code list after analyzing the data. During the analysis stage, there is a great amount of risk associated as the data can be out of context. As this thesis is an exploratory study, we feel that middle ground approach is our best option and therefore, we employed it to perform coding. The steps involved in coding and analyzing is explained in points below:

1. As stated earlier, this thesis study deals with two central topics i.e. Market Segmentation and Data Sharing. The expert interviews involved questions pertaining to both these topics. Therefore these two topics were developed as themes for this thesis study. For sake of simplicity, Market Segmentation was allocated Theme 1 and Data Sharing was given Theme 2.

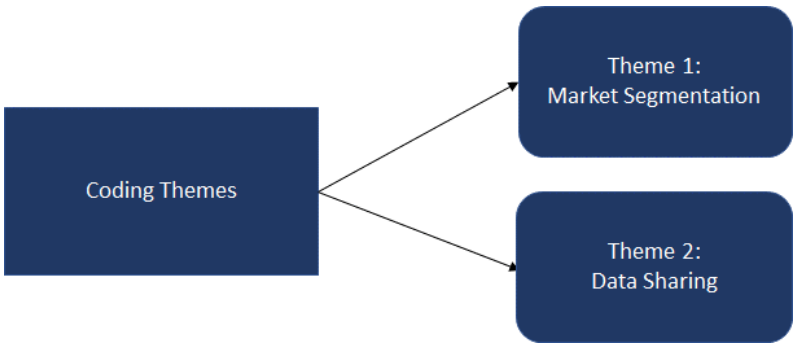

Figure 6: Themes for Coding

2. Establishing questions for both the themes helped us in generating categories. For every question that was discussed in the interviews, a category was developed. As an example, let's look at question numbers 1 and 5. In question.1, ***Which segmentation technique is employed at your firm?*** As this question relates to the different segmentation techniques employed by firms, we developed the category **'Segmentation technique'** and for question 5 ***What are the challenges encountered when this technique is used****?* the category **'Challenges encountered'** was developed. The same procedure was followed for all the remaining questions. Totally, 8 categories were developed for both themes. Out of the 8 categories, 6 categories were related to the theme Market Segmentation and 2 categories

were associated with the theme Data Sharing. Table 5 enlists the categories developed for both the themes i.e. market segmentation and data sharing.

Table 5: Categories for themes Market Segmentation & Data Sharing

Market Segmentation	Data Sharing
Segmentation Technique	Perception on Data Sharing
Data Structures Used	Possible Barriers
Type of Data/ Current Method of Acquiring	
Segment Size	
Result Satisfaction on the current segmentation technique	
Challenges Encountered	

3. As mentioned earlier, this thesis study employed a middle ground approach. By reviewing the literature published for the theme market segmentation, an initial set of code list was generated for the six categories associated with this theme. The initial set of codes for all the categories are listed and explained in Chapter 4.We stated earlier that a total of 10 interviews were conducted in this thesis study. Generating initial set of codes before the interviews helped us gain sufficient background knowledge and our direction for the interviews and post interview, transcribing the interview responses into small phrases of text helped us extract key points to analyze the acquired content comprehensively and effectively. Table 6 provides a snapshot of how coding was performed.

Table 6: Coding Process

Q.No	Question	Theme	Category	Coding Example
1.	Which segmentation technique is currently employed at your company?	Market Segmentation	Segmentation Technique	**Interview Text:** **"**We have multiple segmentation techniques. All of them are based on models/ statistical models and they are user-oriented. They are named as *behavioural segmentation, value segmentation and churn prediction technique "*(Interview C). **Assigned Labels:** Behavioural Segmentation Value Segmentation Churn Prediction
2.	What are the challenges encountered when this technique is used?	Market Segmentation	Challenges Encountered	**Interview Text:** "From my experience as a marketing analyst, It's never a lack of data or the unavailability of data that was always a hindrance for the success of the segmentation technique we used. The only challenge I faced with data is the

				unavailability of resources related to data analysis" (Interview B). **Assigned Labels:** Technical Challenges- Lack of data analysis resources **Interview Text:** "The main challenge that I think many companies in India face is the actual data itself. The acquired data is never good enough to provide us and help us understand critical patterns in the customer needs and due to this our segmentation has become flat and ineffective" (Interview F)**.** **Assigned Labels:** Data Related Challenges- Incomplete data sets
3.	What do you feel on this concept of pooling customer data from various sources (data sharing)? Do you believe by adopting this concept, the effectiveness of segmentation techniques can be improved?	Data Sharing	Perception of Data Sharing	**Interview Text:** "I really feel this concept will give firms 360 degrees regarding customer needs because the information is acquired from different perspectives" (Interview A). **Assigned Labels:** Benefits- 360-degree perspective generated

3.6 Summary

This section provides a summary of all the points discussed in this chapter. From the earlier sections in this chapter, we understood that this is an explorative study that employs a qualitative methodology. We used semi- structured explorative interviews as a data collection method, and middle ground approach to analyse the collected data.

4. Findings

The previous chapters discussed the fundamentals of market segmentation, data sharing and the research methodology employed in this thesis study. In this chapter, we interpret the opinions provided by the experts on each of the two themes i.e. market segmentation and data sharing and analyze if markets can be segmented more effectively with data sharing.

Chapter 3 (Research Methodology) explained the methodology used to conduct interviews and questions asked to the industry experts on the topics of market segmentation (six questions) and of data sharing (two questions). In this chapter, the focus will be on the answers of the experts for the two themes. We divide this chapter into two sections, the first section focuses on the topic Market Segmentation and the second half focuses on the topic Data Sharing. The sections are divided into subsections according to the categories developed in Chapter 3. Every subsection is split into three parts i.e. first part explains the initial list of codes developed for the category and the second part analysis the expert responses. Finally, we draw conclusions for each of the subsection.

4.1 Theme 1- Market Segmentation

In section 4.1 and its sub-sections, we discuss the answers of the respondents and our views on them, for the six questions listed in Table 3 related to market segmentation.

4.1.1 Category 1 - Segmentation Techniques used in Telecom industry

In Chapter 2, we defined market segmentation as "dividing a heterogeneous market into a number of smaller homogeneous markets in response to differing product preferences among important market segments. It is attributable to the desires of consumers or users for more precise satisfaction of their varying wants" (Smith, 1956). In the context of telecom industry and with excessive use of big data, we understood that segmentation techniques employed by telecom firms are data-driven. By reviewing the published literature, we found five relevant market segmentation techniques that are commonly used by telecom firms. The initial list of codes developed in this category was:

- Value Segmentation
- Behavioural Segmentation
- Lifecycle Segmentation
- Life stage Segmentation
- Migration Segmentation

In the next subsection 4.1.2, we analyse the responses provided by the experts for this category and try to understand if our theoretical findings are adopted in practice.

4.1.1.1 Results & Analysis

Armed with a theoretical understanding of the segmentation techniques and the initial list of codes, we conducted interviews with segmentation experts in the telecom industry and the main points inferred from the discussions are summarized below:

- All the respondents provided detailed answers regarding segmentation techniques employed in their companies ranging from *life cycle segmentation, customer behaviour segmentation, value segmentation, churn prediction technique, tower-based segmentation, needs-based segmentation, psychographic segmentation, and geographic segmentation*. These techniques are pictographically represented in figure 8 and explained in a brief manner in Table 7. From the theoretical background of market segmentation (Chapter 2), we identified four market segmentation techniques commonly used by telecom firms and we termed them as initial codes for this category in chapter 3. The four techniques were *value segmentation, behavioural segmentation, lifecycle segmentation, and migration segmentation*. From the respondents' answers and our theoretical background, we can deduce that all four segmentation techniques that we reviewed from literature are also actively used in the industry. However, one point to note is that in Chapter 2 we reviewed customer migration technique to predict the customer churn rate. In Chapter 2 we stated out of all the segmentation techniques reviewed, migration segmentation was the least preferred technique among firms. Telecom companies in the industry also predict customer churn rate, but instead of using the segmentation name as customer migration, they call it a churn prediction rate. At the industry level, this segmentation technique is equally preferred in comparison with the other techniques.

- During the analysis of the interview responses, an interesting trend was observed – most of the telecom companies employed two segmentation techniques in parallel (Interviews A, B, C, E, G, and I). For instance, a leading telecom operator in The Netherlands, employed needs-based segmentation and lifecycle-based segmentation to perform market segmentation. Such a combination was chosen because the company realized that by analysing the customer needs and the current status of the subscription contract of the customer in parallel would help them build accurate predictive models. The respondent stated that usage of a single technique did not provide wholistic results (Interview B). Similarly, a market leader in Indonesia used bandwidth and behavioural segmentation in combination. In bandwidth segmentation the firm segmented their customers based on their bandwidth usage and in behavioural segmentation they analysed the customer behaviour exhibited over the network (Interview E). The reasons attributed to such a usage are twofold: (1) due to fierce competition in Indonesian telecom market, the firm inclines to be customer-centric as it imperative to combat volatile market trends (2) with a customer base of 6 million users, both the techniques provided wholistic understanding to the firm regarding customer needs (Interview E). Furthermore, a Finnish academic researcher who participated in this thesis study stated that telecom firms employ multiple segmentation techniques

parallelly as it helps them to analyse the current market trends comprehensively and develop new offerings (Interview G).

Table 7: Market Segmentation Techniques employed at Industry level

Name of segmentation technique used in the industry	Description	Employed by Respondent(s)
Behavioural Segmentation	It is a technique with which firms analyse the customer behaviour patterns exhibited over the networks and segments customers with similar behavioural trends.	A, C, E, F, G, H, I & J
Needs based Segmentation	In this technique, firms analyse the needs of a customer and segment customers with similar needs.	B
Value Segmentation	'Customer Value'- a value measure is calculated for each of the customers. Depending on the estimated customer value, customer is segmented.	A and C
Churn Prediction Technique	Is a technique that analyses the churn risk factor of customers and performs market segmentation.	C
Bandwidth Segmentation	In this technique, the firm analyses the bandwidth usage of customers and divides them into two groups i.e. Low usage and High Usage.	E
Geographic Segmentation	This type of segmentation divides a target market using geographical variables such as country, states, cities, regions, areas, streets, postal code et cetera.	A & F
Psychographic Segmentation	In this type, customers and prospects are segregated using their personality profile, interests, values and lifestyle profiles.	G
Life-cycle segmentation	The current status of a customer contract is analyzed, and customers are segmented into three groups: new customer, matured customer and decline.	B
Tower Based Segmentation	In this segmentation the firm identifies the local tower providers to acquire information regarding the customers and performs segmentation.	D

- In chapter 1 and 2, we stated that market segmentation is based on data-driven decision making. However, during the interviews, we learnt that there are exceptions to this statement in a few cases. Few firms employ segmentation techniques that don't rely a lot of data. For e.g. a small medium telecom enterprise with a base of 10,000 customers in Indonesia aims to increase its market share and compete with the market leader. The firm employs tower-based segmentation which is a segmentation technique that doesn't rely on data. Instead, the company identifies tower providers and generates details about kind of services provided in a building and the infrastructure/operator needed to provide that service. By identifying the tower providers, this telecom firm tries to stitch a partnership and perform market segmentation (Interview D).

> ***"For segmentation we don't use a lot of data, but we partner with local tower providers. For this we need to identify the tower provider and generate details for example which tower provider provides service in a building and then which infrastructure that they use like do they use fibre optics or wireless systems and which operator do they employ. Once we get all these details, we can stitch a comfortable partnership with tower providers".***
>
> ***-Interview D respondent***

4.1.1.2 Drawing Conclusions

In this section, we combine our theoretical understanding and the responses of the experts to draw conclusions and update our code list for this category. The conclusions are listed in points below:

- From analysing the responses, we can infer that segmentation techniques play a pivotal role in helping a firm acquire profits. Our initial code list consisted of 5 segmentation techniques and we observed that all these techniques were employed by firms. We found that firms used different terminologies to name these segmentation techniques without changing the fundamental premise. For instance, a respondent used named 'migration segmentation' as 'churn prediction technique' and another respondent named 'behavioural segmentation' as 'needs-based segmentation' to perform market segmentation. Apart from these theoretically identified segmentation techniques, we also found that telecom firms employ different types of segmentation techniques such as geographic segmentation, psychographic segmentation, tower-based segmentation, needs-based segmentation, and bandwidth segmentation. Our final code list consists 10 segmentation techniques.

- After critically analysing the interview responses, an interesting inference that can be made is with respect to behavioural segmentation. Despite the usage of different terminologies used to name the segmentation techniques, most of the respondents employed segmentation that analysed the customer behaviour exhibited over the network. 7 out of the 10 respondents employed this segmentation technique. The reasons for the dominant usage of behavioural segmentation is two-fold. Firstly, this segmentation is implemented using internal data and does not require external data. By using the internal data kind of data, firms have more confidence in the result generated and do not fear about the veracity factor of data that is used. Secondly, in this technique every action exhibited by a customer is recorded in a detailed manner. With the help of such detailed data, firms can generate sharp insights regarding the needs of the customer. Therefore, these are main reasons why firms feel this segmentation technique is the most reliable method to segregate the customers.

- Having understood that different types of segmentation techniques are employed by the telecom firms, we observe the choice of the segmentation technique is dependent on four factors. They are"

    1. **Company's aim/vision** – Every firm has different goals and targets to achieve. Depending on the aims to achieve, firms choose segmentation techniques that are relevant to achieve these goals.
    2. **Products/Plans**- Telecom firms develop different types of products and plans. To help these plans reach the customers in a comprehensive manner, firms employ different types of segmentation techniques.
    3. **Current market trends**- It is a known fact that market trends are never constant and keep varying frequently. To combat these market trends in an effective manner, firms employ different types of segmentation technique.
    4. **Size of the company-** Firms can be of different sizes such as a large company, medium company, and small & medium company. These sizes determine the choice of segmentation technique.
    5. **Position of the company in the market**- Telecom firms always aim to gain sustainable market share. Position of the firm such as market leader, follower, new entrant plays a pivotal role in determining the segmentation technique.

The above-mentioned factors are the main factors that help the company choose a segmentation technique that is relevant to the organization. Figure 7 summarizes these factors that influence the selection of the technique.

- An interesting aspect that was not identified in the literature is regarding the employment of a combination of market segmentation techniques. With the amount of competition present in the telecom sector, we agree with the deployment of two segmentation techniques simultaneously. As the target market for majority of companies that participated in this thesis study is big, such a deployment will the telecom firms will ensure revenue streams to be maintained in a constant manner.

- We recommend that a combination of segmentation techniques should be done in such a manner that there exists an overlap between the techniques. For e.g. by combining the geographical segmentation and behavioural segmentation will provide insights regarding the behavioural trends exhibited by the customers located at different geographical locations. In this combination, the main drivers that i.e. customer behaviour in combination with location details overlap to a great extent. On contrary a combination consisting of tower-based segmentation and geographical segmentation we not reap significant benefits as there is no overlap present.

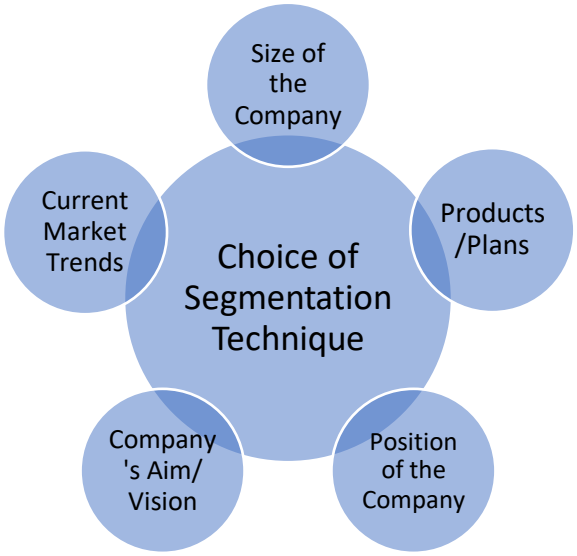

Figure 7: Factors Affecting the Selection of Segmentation Technique

4.1.2 Category 2 - Data Structures and Their Uses

For this thesis study, we consider data as a set of information or as evidence that explains the attributes of a customer, thereby enabling firms to make accurate analytical and data-driven decisions (Arnaut et al., 2018). The headings present in the database that helps telecom firms to gain insights such as demographics, behavioral information of a customer are known as data structures. From our theoretical understanding and from the database shared by one of the partners in Safe-DEED, our initial code list consisted these codes:

- **Demographic Data Points:** Age, Gender, Income & Nationality
- **Behavioural Data Points**: Data points in CDR's i.e. source number, destination number, duration of the call, type of call name of the website visited, duration of data usage, amount of data consumed & comments made.
- **Geographical Data Points:** Postal code & address

In the following subsection, we analyze the interview responses for this category.

4.1.2.1 Results & Analysis

Goals and targets established by a telecom firm are subjective and are dependent on the market the firm operates and the availability of resources. Despite these differences, this thesis study found that majority of the data structures employed by the telecom firms are similar. The data structures that are commonly used by telecom firms to perform market segmentation are *age, gender, postal code, profession, name of the city, billing size, household information, average revenue per user (ARPU), behavioral data headings such as no of calls made, destination calling number, type of call, duration of a call, number of SMS sent, name of the website visited, amount of internet data used, migration behavior* etc. Table 8 lists the data structures commonly used by telecom firms. For instance, a respondent working for a leading telecom operator in India stated that age, gender, city, postal code, profession, amount of data consumed, number of incoming and outgoing calls were

the common data structures used by them to perform behavioral segmentation (Interview F). Furthermore, the respondent from a Belgium based telecom operator also employed similar data structures to perform a combination of behavioral and life-cycle segmentation (Interview A).

The usage of demographic and behavioral data structures among many telecom firms can be attributed to two reasons. The first reason relates to the effectiveness of behavioral segmentation. As behavioral segmentation is one of the most effective segmentation techniques that helps managers understand the customer needs accurately, many firms have been employing data structures that explain the demographic and behavioral aspects of a customer (Interview A, F, B). The second reason relates to the veracity aspect. Respondents stated that as behavioral data structures are acquired internally, the degree of veracity of these data structures is high and this helps firms build reliable predictive models. For instance, the respondent from a leading telecom firm stated that -

> ***" The main reason for employing the behavioral and demographic data structures is that since they are acquired by our own systems, this gives us more confidence in building our predictive models"***
>
> ***-Interview C respondent***

Table 8: List of Data Structures Employed by Telecom Firms

Type	Data Structures
Demographic	Age,Gender, Profession,Household Information, Income, Billing size, Average revenue per user
Geographical	Postal Code, Area of residence, City of residence, Type of environment (such as town, village, city)
Behavioural	Number of calls made, Duration of a call, Type of call, Amount of data consumed, Name of the website visited, Number of SMS's sent, Migration Behaviour

4.1.2.2 Drawing Conclusions

By analysing the interview responses in this category, we can infer the following:

- Comparing our initial code list stated in Section 4.1.2 and final code list in Table 12, we observe all the data structures that were identified in our initial code list were employed at the industry level. Additionally, firm use numerous other data structures that explain the demographic, geographic and behavioural aspects of the customers. For instance, in our initial code list we identified four data points

i.e. age, gender, nationality, and income in the demographic data type. However, we observed that firms also used information that explained a customer's billing size and average revenue per user in the demographic data type. Similar inference can also be made with respect to the geographical and behavioural data type. The main reason attributed to the vast usage of data structures is the need to gain sustainable competitive advantage in today's market place. Due to the increased competition present in the market place and the availability of numerous products for customers, it is imperative that telecom firms gain sustainable advantage to stay relevant. Therefore, firms are inclining themselves to capture maximum data regarding customers in order to develop a robust market segmentation technique.

- From the previous sub section 4.1.2.1**,** we identified the data structures employed by the telecom companies and understood the benefits of using some data structures like demographic, geographical and behavioral ones. After performing content analysis in this category, we can conclude that found that none of the respondents used data structures such as wallet size, credit rating, personnel preferences such as favorite sport, football club, television channel, movie genre, inclination to travel et cetera. Collecting data that explain personal preferences can enable telecom firms to understand a customer's needs and preferences in a holistic manner and develop customized services and products. By targeting the right customer at an appropriate time will help telecom firm to increase their average revenue per user and decrease the rate of churn. Moreover, this observation was agreed upon by one of the respondents (Interview C).

- In the above mentioned we identified and explained the missing data structures in the database of a telecom firm. Reflecting on our research objective, this thesis study states that these missing data structures can be potentially acquired by sharing customer data. For e.g. a telecom operator who is interested to sell sports subscription packages can partner with local sports organizations and acquire information on questions such as Which football club does the customer support? Which football league is the customer interested. With the help of such data strcutures firms will be in a better position to understand the needs of the customers and develop a compelling marketing mix.

4.1.3 Category 3- Types of Data/ Data Collection Methods

In chapter 2, we discussed the importance of collecting good quality customer data. It is imperative for telecom firms to collect good quality data that contain demographical, behavioral, geographical and other aspects of a customer, as it forms the foundation in helping firms identify accurate market segments and perform effective segmentation. By reviewing the published literature, our code list was:

- **Type of Data:** Internal Data
- **Data Collection Methods:** Interviews, surveys & web research

In this next section, we will analyze the interview responses for this category and refine our initial code list.

4.1.3.1 Results & Analysis

In order to comprehend the above-mentioned findings, content analysis was performed to understand the views of experts for this category and these observations made:

- Traditionally there are two kinds of data i.e. internal and external data available to telecom firms to perform market segmentation. Telecom firms acquire internal data from provisioning systems, set up boxes and customer contracts. External data is often acquired by collaborations with accredited third-party agencies who specialize in market research analysis. Majority of the respondents stated that internal data was predominantly used to perform market segmentation. For instance, a respondent stated that 90% of the data used for market segmentation was internal data. This data was acquired from customer contracts, provisioning systems and setup boxes (Interview C). Similar views were expressed by the other respondents. The respondents stated that the most common methods to acquire internal data are: *customer contracts, call detail records, usage data generated when a customer initiates an action over the network, conducting surveys (*Interview A, B, F*).* Apart from these traditional methods of data acquisition, a leading telecom operator in The Netherlands employed a unique method termed 'lunch and learn'sessions with customers and prospects to acquire data in an internal manner. This firm organized 'lunch and learn' sessions by inviting customers for a free lunch and interacted with them to understand their need's and wants from telecom products and services (Interview A).

> ***"90% of our data is internal data. We fundamentally use data such as Call Detail Record's generated by our customers, the viewership data generated by the setup boxes and the demographic data that comes from the other provisioning systems and our CRM systems"***
>
> ***-Interview C respondent***

> ***"An interesting thing that we adopted was 'lunch and learn lectures' for customers. In events like these we invited customers for a free lunch and spoke to them about their needs and desires. This was an effective initiative and we acquire so much relevant information"***
>
> ***-Interview A respondent***

- From an earlier point, we understood that internal data is the most preferred option in the telecom industry. The reasons for the non-usage of external data are twofold. After conducting pilot projects by using external data, a respondent felt that the value added by external data was very less in comparison

to the high investment costs that were required to acquire the external data (Interview B). The other reason for the limited usage of external data is related to trust and veracity. As the quality of data acquired plays a pivotal role in implementing an effective segmentation campaign, an expert from Indonesia stated that there was a lack of resources and technology to validate the veracity of the acquired data from 3rd parties (Interview E).

> ***"We only use internal data and not external data. This because we have trust issues and always suspect the veracity of the external data. In addition to this we lack technological tools that can verify the veracity of such huge volumes of data".***
>
> ***-Interview E respondent***

- Although companies preferred usage of internal data, some companies didn't completely do away with external data, but they still employed external data on a limited scale. For instance, a telecom firm partnered with an organization named 'Blayedos' to acquire information about newborn infants in Belgium. This information helped the firm to enrich its database and perform effective lifestyle segmentation (Interview A). Furthermore, the respondent from enunciated that they acquired external data to validate their understanding of the market trends and did not combine their internal data with the external data (Interview C). In the 10 step market segmentation process steps that we explained in Chapter 2 we identified that evaluation and monitoring as the final step in market segmentation. Given the volatility in market and the swift change in the needs of the customers, we reason that companies should use data from market research specializing companies as this will help them evaluate the effectiveness of their current offerings. Such an evaluation will provide vital leads in developing future strategies.

4.1.3.2 Drawing Conclusions

In this section, we aim to draw conclusions by combining our insights from theory and the viewpoints expressed by the respondents.

- From our theoretical understanding, we stated interviews, surveys, and web research was the main methods employed by the telecom firms to collect data. In practice, all these methods that were identified from literature are being used by the firms. Apart from these methods' interactive sessions such as lunch and learn are also being conducted. We also agree with arguments stated by this respondent and believe that all the firms need to embrace and organize such techniques. This is because organizing such interactive techniques do not require huge monetary costs and also provide a valuable opportunity for telecom firms to directly interact with customers on a one to one manner and understand their requirements accurately.

- In the argumentation we provided in Chapter 2 we stated that firms used internal data predominantly to perform market segmentation due to the veracity aspect. However, from the interview findings, we observe that this was only the reason for the employment of internal data. We found three other factors that force firms to employ external data in a limited manner. The overlap factor between external and internal data was one factor that forced firms to employ internal data. Apart from these factors, huge investments costs involved in acquiring the external data and longer durations of time were required to clean the data the external were the other reasons why firms employed internal data.

- In Chapter 2, we opined based on theoretical literature that the quality of the acquired data determines the effectiveness of the end results in market segmentation. These statements were also supported by respondents who participated in this thesis study. This can be highlighted by comparing the arguments in Chapter 2 and the opinions expressed by a respondent (Interview A). For instance, in chapter 2 we argued that if telecom firms' possess siloed information about customer behaviour, it would hinder companies to develop a compelling marketing mix and target the right customers at the appropriate time. Supporting this claim, the Interview A respondent also realised that obtaining bad quality data or data only from a single perspective will have significant impact in identifying high value customers and thereby affect the results of segmentation (Interview A).

> ***"From my expertise after working for several firms, I realise that bad quality data and data acquired from a single perspective can have detrimental effects on segmentation. For e.g. such data hinders us from identifying high valued customer and generating poor segmentation outcomes"***
>
> ***-Interview A respondent***

4.1.4 Category 4- Segment Size

This section analyses the responses provided by the experts regarding the size of each market segment. In order to analyze this category comprehensively, it is important to understand two terms i.e. market size and size of market segment. Market size is defined "as the number of individuals in a certain market who are potential buyers/current users of the products and services offered by a firm" (Will Kenton, 2019). From the perspective of a telecom firm, there are three main factors that determine the market size of each firm. The three factors are local government regulations, degree of competition and current market trends. In order to combat these three factors, telecom firms have been dividing the market in order to define and identify relevant market segments. We define market segments as a group of customers who share common needs, wishes that exhibit similar behavior (Dolnicar et al., 2018). Apart from three factors that were mentioned above, size of a market segment for a telecom firm is dependent on the goals/vision the company wishes to achieve.

After reviewing the published literature in Chapter 2 our initial understanding of the segment size was not quantified, but we understood that modern age telecom firms identify large indistinct customer segments and perform market segmentation. Having described the two terms market size, market segment and our initial code list we shall now understand the answers provided by respondents for the question: **How granular are your market segments?**

4.1.4.1 Results & Analysis

- For large telecom companies, their customer base consists of multitude of customers. We shall elaborate in detail the above-mentioned statement by taking the examples of large companies that participated in this thesis study. India is one of the largest countries in the world with a population of ~1.3 billion. The Telecom Regulatory Authority of India (TRAI), an apex telecom regulatory body in India has divided the telecom market in India into 22 circles. The main reasons for such a division can be attributed to geographical factors and a need to make communication a basic necessity for every citizen. The Indian telecom operator addresses all the 22 circles and possesses a market share of 13%. Each circle consists of 100 segments and each segment has approximately 50-60 thousand customers (Interview F). Similarly, a market leader in Greece address 16% of the Greek telecom market and has 600,000 customers. To address their needs efficiently, the firm divides its customer base into 8-9 segments. By performing basic multiplication, we can infer each segment consists of 75 thousand customers (Interview C). One of the top players in the Netherlands also divides its entire customer base into 9 segments. Each segment consists of 100, 000 customers (Interview A).

> ***"Well as I said the Indian market is massive, I can only provide approximate numbers. Each of the 22 circles has about 100 segments of customer. The segment size would depend on the circle and geographical factors. In order to accommodate the entire population of India which is. On an average each segment size would contain around 50k-60k customers. So about 5 million customers are present in every circle".***
>
> ***-Interview F respondent***

- In the earlier point, we understood the statistics of each segment for top-performing firms. In order to gain complete knowledge on granularity aspect of segments, it is also important to understand the segment size for small telecom firms. Small and medium enterprises have a significantly smaller customer base when compared to large companies. This can be explained by analyzing the small-medium enterprise in Indonesia. This SME solely focuses on offering services for landline phones. With such a limited focus, the firm owns a customer base of 10,000 customers and the size of each segment is 10 to 20 customers (Interview D).

> ***"We are a small and medium enterprise that operate only with landline telephones and not mobile phones. We have a customer base of 10,000. Approximately the size of every segment could be 10 or 20 customers and one customer can contain a lot of telephone numbers she/he are using simultaneously".***
>
> ***-Interview D respondent***

4.1.4.2 Drawing Conclusions

- In chapter 2 we did not quantify any specific number regarding customer segment size. However, we stated that modern age firms identify large customer segments due to presence of siloed database. From the interactions with experts we observe that this was not the only reason for identification of large customer segments. Lack of management resources and lack of enough monetary budgets were also the factors that forced companies to identify large customer segments.

  - From the above-mentioned points and by analysing the responses from experts we can deduce three main factors that determine the granularity of a market segment for a telecom firm: The main factors are:

  - **Size of the company-** The size of a telecom company has a direct impact on determining the size of a market segment. Large companies and top telecom players have a larger customer base and therefore to accommodate all the customers these firms divide their market segments in large numbers.
  - **Human Resources**- This factor plays a vital role in determining the size of a segment. Examples of human resources are managers, marketing personnel, database systems operators et cetera. The more human resources a telecom firm employs, management of segments becomes an easier process. Due to the lack of human resources, we observed that the Greek telecom operator was not inclined to reduce the current market segment size as they feel management of all the segments would be difficult (Interview C).

> ***"However, for instance, when we're talking about behavioural segmentation, we don't want to have a very niche segment or segments with less no of customers. This is because we don't have that the capacity in terms of Human Resources another and commission cost resources to target a very nice segment. So, we try to maintain the number of eight to nine segments, from the gathered several customers so that we can target them and share the same characteristics".***
>
> ***-Interview C respondent***

- **Vision/ Goals-** The same reasons provided in Section 4.1.1 can be attributed here as well. The vision, goals that a firm aims to achieve in a given period of time determine the size of a market segment.

- The granularity of customer segments plays an important in helping firms identify customer needs. After analyzing the viewpoints of experts, we observe that all the telecom firms have employed traditional methods of segmentation and allocated a customer to single customer segment. However, we recommend that telecom firms have to go beyond the traditional methods and allow customer to be a part of multiple customer segments. Such multiple allocations will ensure that the needs of customers are not neglected and will also help the firm understand the retention and value drivers for each customer in a comprehensive manner.

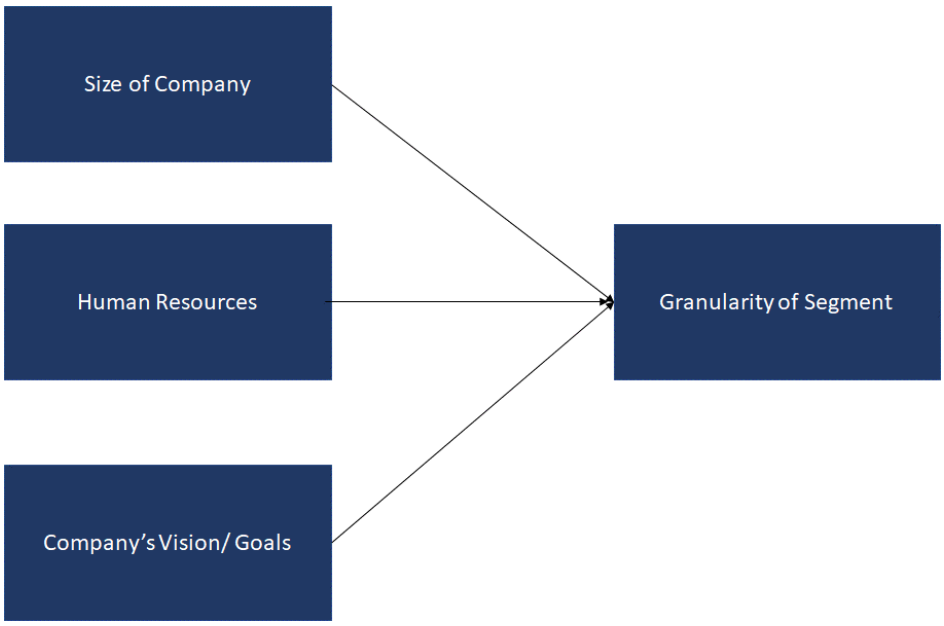

Figure 8: Factors Affecting Segment Size

4.1.5 Category 5 - Challenges of Market Segmentation

One of the most common analogies that are often used is 'competition' and 'telecom industry'. The main reasons for association of the two phrases are globalization and increased modes of communication for customers. As telecom firms operate in a highly complicated environment with volatile market trends it would not be incorrect to associate the phrase 'challenges' with the telecom industry. In theory we identified the following challenges faced by telecom firms when performing market segmentation:

- Incomplete Data Sets
- Obsolete Data

In the next section, we validate if these challenges are faced by telecom firms and try to understand the other types of challenges faced by telecom firms.

4.1.5.1 Results & Analysis

This section identifies the common challenges encountered by modern age telecom firms. Analyzing the responses provided by respondents, we identified two types of challenges that are often faced by modern age telecom firms. The two types are data-related challenges and technical challenges. They are explained in detail as points below:

1. **Data Related Challenges:** In the earlier chapters, we discussed that market segmentation is data-driven. For an efficient data-driven decision making it is imperative that firms possess the right data with utmost quality at the right time. However, many modern age telecom firms encounter data-related challenges that hamper their market segmentation. The common challenges are:
    a. **Incomplete data sets:** Recent studies show that as the even-though huge volumes of data are being collected, many organizations are lacking state of art technological tools, expertise and methods to strategically exploit the big data potential (Deichmann, Heineke, Reinbacher, & Wee, 2016). By talking to experts, we found that these research findings are common in the telecom sector. Unavailability of important drivers in customer data that help companies in understanding the needs and desires of customer segments is one of the most common challenges faced by telecom firms across the globe. Telecom firms irrespective of their size i.e. large enterprises, small and medium enterprises are facing this problem (Interview J, F, H). For instance, the Indian respondent stated that the acquired data does not provide the complete details regarding the needs of customers. Lack of sufficient data has a great potential to reduce the effectiveness of the segmentation technique. The Indian respondent adds on to state that due to lack of important drivers, the firm predicts the needs of the customers from their behavioral patterns exhibited during the usage of products and services offered by the firm. Majority of these predictions fail as customer needs keep changing and therefore the marketing plans are ineffective (Interview F). *'5 Why's'* is one of the most famous interrogative techniques done in an iterative manner. This technique is used by one of the leading companies in USA to understand the customer needs. However, due to the lack of sufficient customer data forces the firm to stop in the second iteration, thereby failing to understand the needs comprehensively (Interview J). The academic respondent who participated in our thesis study also states that lack of connectors is the main problem faced by telecom across the globe.

> ***"The main challenge that I think many companies in India face is the actual data itself. The acquired data is never good enough to provide us and help us understand critical patterns in the customer needs and due to this our segmentation has become flat and ineffective. I can explain this with an example of telecom data we use. In the current practices we acquire behavioral data of the customer, but this data does not provide details regarding the user's interests. All we do is try and predict the user's interests by the behavior they exhibit in the network. Many a times I have seen our predictive models going completely wrong. These incomplete data sets are a big challenge".***
>
> ***-Interview F respondent***

b. **Lack of sufficient real-time data**: In Chapter 1 we stated that with the advancements in technologies and emergence of over the top services, customer needs and wishes constantly would change. This has increased the importance of telecom firms to collect and analyze real-time data. However, many telecom firms lack real-time data in their database. For instance, a respondent stated that the only real-time data they possessed was regarding the behavioral trends exhibited during the usage of telecom products (Interview C).
c. **Obsolete data:** This is also one of the common challenges faced by telecom firms. During the process of acquiring customers, documents such as customer acquisition forms are used to collect data that explain the various aspects of a customer. The common data points that are collected from these documents are age, gender, postal code, name of the area household information, preferences et cetera. However, with a given period, this data can get outdated. Due to lack of methods that update the data, firms still use the old data to perform segmentation. Such lack of updated data can be attributed as the main reason for failure of predictive models.
d. **Scattered Data:** Firms such as Telekom Indonesia employ an old customer relationship management system. These old CRM systems have low storage capacity and therefore firms store different aspects of customer data in varied sources. With the usage of old CRM, the firm fails to aggregate the collected data in a comprehensive manner (Interview E).
e. **The veracity of the collected data:** In section 4.1.3 we stated that due to this challenge many firms are inclined to use internal data rather than external data. As firms employ various methods to acquire customer data, there are often unable to verify the veracity aspect and therefore fail to develop accurate predictive models. For instance, a respondent stated data acquired from survey and loyalty programmes did not give the firm enough confidence to develop offers from the responses provided. This is because the firm felt that such data possessed a great probability for biased opinions. This thesis study finds that such challenges can be succeeded if firms participate in data sharing platforms. As the data acquired will be from trusted partners the data would have a high probability of being correct and relevant.

Due to these above-mentioned data-related challenges, telecom companies possess data in silos which do not provide a holistic view regarding customer behavior. These discrete data sets are being used and large, indistinct customer segments are being identified. Hence, the telecom firms employ an ineffective market segmentation technique and fail to create a compelling customer experience.

2. Technical Challenges

The technical challenges faced by modern age telecom firms are:

a. **Lack of state-of-the-art technological tools:** As mentioned in Chapter 1, modern age telecom firms are capturing massive volumes of customers. On an average 2.5 quintillion, bytes are being captured daily. Even though huge volumes of data are being collected few top-performing telecom firms such as lack state of the art technological tools that can store these huge amounts of data. A respondent stated that:

> ***"The main challenge that we face is with respect to CDR (Call Detail Records). We have the CDR's coming from the internet, telephony, communications and TV. They generally occupy huge data spaces. We don't have that much of data space. So, we are forced most of the times to aggregate data. I feel using aggregated data for a period is okay, especially for the telecommunication data. But that's not pretty much okay for our TV viewers. It doesn't provide all the aspects that we want to investigate".***
>
> ***-Interview C Respondent***

b. **Lack of data analysis resources:** A respondent stated that despite possessing state of the art technological tools, they lacked in analyzing these massive volumes data. Inadequate monetary funds and a rigid organizational structure are the main reasons for lack of data analysis resources (Interview B).

4.1.5.2 Drawing Conclusions

From analyzing the viewpoints of experts for this category we draw the following conclusions:

- In Chapter 2, by reviewing the literature published on challenges in market segmentation we only identified that incomplete data sets and obsolete data are the main challenges faced by telecom firms. However, after analyzing the above-mentioned viewpoints our initial code list developed and we understood that firms face numerous types of data related challenges such as scattered data, verifying the veracity of the data and lack of real-time data. Apart from challenges that were associated with data, we also understood that few firms faced technical challenges that hampered the effectiveness of market segmentation techniques. The common technical challenges faced were lack of data analysis resources and lack of state-of-the-art technological tool that can help firm store large volumes of customer data.

- In Chapter 2, we explained the 10 steps involved in Market Segmentation process. Step 5 of the 10-step process relates to 'exploration of the collected data'. Exploration of data to discover critical trends is highly reliant of collected data and efficient exploration is only possible if the acquired data is complete. In order to stay competitive and relevant in the market place, telecom firms are in the need to constantly roll out attractive and efficient offers plans frequently (weekly or 10 day period). With such a constant pressure to update their offerings, acquiring complete customer data that provide detailed information regarding the factors that determine the customer's intention to buy telecom products. Therefore, this increases the need to acquire data that can overcome siloed databases. We agree with the experts about data-related challenges. We add to their arguments that in order to combat the volatile market trends it is important that telecom firms constantly evaluate their key performance indicators. Evaluating vital parameters constantly will help telecom firms to develop efficient strategies and therefore aid in creating a clear road map for future business development.

- Most of the respondents stated that data related challenges were the main hurdles that were encountered during the implementation of market segmentation. Apart from these challenges However, few respondents also stated that technology-related challenges such as lack data analysis resources, lack of storage systems as the main barriers. As our primary focus in this thesis to analyze how data sharing can overcome the current challenges faced by telecom firms, we feel that technology-related challenges cannot be nullified by data sharing, therefore, treat these challenges as out from our scope of study.

4.1.6 Category 6- Result satisfaction on chosen market segmentation technique

Pareto's principle of factor sparsity or the 80/20 rule has been a widely used thumb rule in explaining the distribution of profits for a firm. The rule describes that 20% of customers generate 80% of a firm's profit (Michael Schrage, 2017). However, as explained in the first chapter, the telecom industry has been experiencing tumultuous times in the recent past. An increased level of competition and decreased levels of average revenue per user are a few examples that explain the current state of the telecom industry. In such a scenario with volatile market conditions, the telecom industry has supercharged the 80/20 rule (Michael Schrage, 2017). Therefore, analyzing the satisfaction levels from the perspective of a telecom manager in terms of the profits generated from the chosen segmentation technique becomes imperative. From our theoretical review, we understood that due increased presence of competition in the market place, telecom operators are facing lower average revenue per user (ARPU) levels and are therefore unsatisfied with the result generated. In the next sub section, we shall analyze the viewpoints of respondents stated for this category and validate if our initial theoretical understanding is correct.

4.1.6.1 Results & Analysis

This section aims to describes the important aspect 'Satisfaction of Results' by analyzing the responses provided by experts for the question: **Are you satisfied with the results generated with the chosen market segmentation technique?**

- With the current state of the market, most of the respondents expressed dissatisfaction over the profits generated and their inability to predict customer needs accurately. For instance, the Indonesian respondent stated that by employing tower-based segmentation, finding prospective customers has been a major concern. He stated that the emergence of smartphone and voice over internet protocol applications such as Skype and WhatsApp as fundamental reason for low ARPU levels (Interview D). Another respondent also cited the emergence of Over the top players as the main reason for dissatisfaction (Interview K). Arguing on the similar lines, the academic researcher stated that current state of the art market segmentation technique fails to identify critical behavioral patterns. This has led to many telecom firms failing to create a compelling marketing mix (Interview G).

- Few respondents who participated in the discussion expressed that they were satisfied with the current results but due to presence increased competition present in the market place, they aim to embrace innovation and seek improvement. This can be explained by analysing the viewpoints of the respondents associated with market leaders in Greece and Indonesia. The respondent from Greece stated that their current results were on par with their competitors but aimed to embrace the concept of data sharing to gain sustainable market share (Interview C). The respondent from Indonesia employed bandwidth and behavioral segmentation. The expert stated that the firm was satisfied with the results generated from the bandwidth segmentation but aimed to improve its performance on behavioral segmentation(Interview E).

4.1.6.2 Drawing Conclusions

The main conclusion that we can infer from the above-mentioned viewpoints are the following:

- While reviewing the published literature in this category, we understood the telecom firms across the globe have been developing ineffective predictive models due to the lack of sufficient data. Additionally we also stated that due to the emergence of OTT players, telecom operators are failing to understand the needs of the customers in a comprehensive manner. During the empirical analysis phase, we observed that these are the issues are creating unsatisfaction among telecom operators across the globe.

- From the above-mentioned viewpoints, we can infer the following points. Traditionally, telecom firms have generated profits via three streams i.e. voice, data and messaging. Due to the growth of Over the top content services (OTT) such as Netflix, WhatsApp, Skype and YouTube have disrupted revenues from two of the above mentioned three streams i.e. voice and messaging. Recent studies show that more 45%

of the smartphone users across the globe are communicating by the means of an OTT messaging application such as WhatsApp, Viber et cetera (Sujata et al., 2015). Even though OTT service has led to massive data consumption levels, the profits generated from the data stream fail to surpass the loss the generated from the voice and messaging streams. Additionally, due to a constant upgrade in the content provided in the OTT services, telecom firms do have the data explain the value and retention drivers of customer are, therefore, failing to create a compelling marketing mix.

4.1.7 Conclusions for Theme 1- Market Segmentation

As stated earlier, this thesis study deals with two central topics. In this section analyzed the findings for the theme Market Segmentation. This theme contained 6 categories and we explained our initial code list developed for each category, analyzed the viewpoints stated by the experts and drew conclusions. The main aim of this chapter was to answer our third sub research question **Which segmentation techniques are used in the telecom industry and what kind of challenges were observed during their implementation?** From the findings of this section we observe that all the five segmentation techniques i.e. value segmentation, behavioral segmentation, lifecycle segmentation and life-stage segmentation that were identified in Chapter 2 were employed by telecom firms. Apart from these segmentation techniques, telecom firms employed needs-based segmentation, tower-based segmentation, geographic segmentation, psychographic segmentation, and bandwidth segmentation as well. Due to the volatile market trends, majority of the firms employed a combination of segmentation techniques to perform effective market segmentation. An important finding from this section is related to behavioral segmentation. Most of the respondents stated that behavioral segmentation is the most efficient segmentation technique as it provides a detailed understanding of the customer trends and needs. In Table 9 we summarize the final code list and important findings obtained in each category and in Table 10 we visualize the codes generated.

Table 9: Final Code List & Summary of Findings for Market Segmentation

Category	Final Code List	Summary of Important Findings
Segmentation Technique	Tower based Segmentation, Needs-based Segmentation, Bandwidth Segmentation, Geographic Segmentation, Psychographic Segmentation, Value Segmentation, Behavioral Segmentation	1. Majority of the respondents employ Behavioural Segmentation. 2. Due to volatility in the market, segmentation techniques are employed in combination. 3. Not all segmentation techniques are heavily reliant on data for e.g. tower-based segmentation.
Data Structures	Demographic, Geographical and Behavioral Data	1. Three types of data structures are used. 2. With the need to stay competitive firms are embracing the attitude 'catch-all you can' data regarding customers.
Type of Data	Internal Data, External Data	1. All of the respondents used internal data to perform market segmentation. 2. Reason for the limited usage of external data is due to lack of overlap data.
Current Method of Acquiring	Contract data, Interactive sessions with customers, Usage data, Billing Size & Loyalty Programmes	1. Firms collect data by themselves. They are not inclined to collaborate with market research companies. 2. Firms are more interested to conduct interactive sessions with customers to understand their needs.
Segment Size	Less than 100 customers per segment, Less than 100k customer per segments, 5 & less than 5 segments, 10 & less than 10 segments, 100 &Less than 100 segments	1. Size of segments for each company is dependent on these factors a) size of the company b) human resources c) vision/goals. 2. Most of the respondents employed large segment size due to incomplete customer data sets.
Result Satisfaction of currently used segmentation techniques	Satisfied, Unsatisfied, Satisfied but seek Improvements	1. Majority of the respondents were not satisfied, few of the respondents were satisfied but sought improvements to improve their segmentation techniques. 2. The main reason for unsatisfaction is emergence of OTT players.
Challenges Encountered	Data related challenges: Incomplete data sets Lack of sufficient real-time data, Obsolete data. Scattered Data, Lack of knowledge on the veracity of collected data Technical related challenges: Lack of state-of-the-art technological tools & Lack of data analysis resources	1. Most of the respondents faced data-related challenges. Incomplete data sets were the most frequently faced challenge. 2. Due to incomplete data sets, firms do not possess important value and retention drivers regarding customers, thereby develop ineffective segmentation techniques. 3. Few respondents also faced technology related challenges, caused due to allotment of insufficient budget for market segmentation.

Table 10: Data Visualization for the theme Market Segmentation

Codes /Interview Code	A	B	C	D	E	F	G	H	I	J
Category - Segmentation Technique										
Needs based Segmentation		✓								
Lifestyle Segmentation	✓									
Bandwidth Segmentation					✓					
Tower based Segmentation				✓						
Behavioural Segmentation	✓		✓		✓	✓		✓	✓	✓
Lifecycle Segmentation		✓								
Geographical Segmentation										
Psychographic Segmentation							✓		✓	
Churn Prediction Technique			✓				✓			
Value Segmentation			✓							
Category- Type of Data										
Internal Data	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
External Data	✓		✓							
Category- Current Method of Acquiring										
Surveys	✓	✓				✓	✓	✓		✓
Contract Data	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Interviews/Focus groups/Interactive Sessions with Customers	✓	✓					✓	✓		✓
Usage Data- Call Detail Records	✓	✓	✓		✓	✓			✓	✓
Category- Result Satisfaction										
Satisfied		✓			✓				✓	
Satisfied but seek Improvement			✓					✓		
Not satisfied	✓			✓	✓	✓	✓			✓
Category- Challenges Encountered										
Data related Challenges	✓		✓		✓	✓	✓	✓		✓
Technology related Challenges		✓	✓	✓					✓	
Category- Data Structures										
Demographic Data	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Behavioural Data	✓	✓	✓	✓	✓	✓		✓	✓	✓
Geographical Data	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Category-Segment Size										
Less than 100 customer per segment				✓	✓				✓	
Less than 100k customer per segments	✓		✓			✓				
5 & less than 5 segments		✓					✓			✓
10 & less than 10 segments								✓		
100 &Less than 100 segments										

4.2 Theme 2- Data Sharing

In section 4.2 and its sub-sections, we discuss the answers of the respondents and our views on them, for the two questions (Table 3.3.3) related to data sharing. As similar to the earlier section 4.1, for every category developed in this theme, we first explain our initial code list, analyze the interview responses and finally draw conclusions.

4.2.1 Category 1 - Perception of Data Sharing

In chapter 2 we stated that data sharing is a process in which data is traded between the data buyer and data seller for business purposes. We also understood that data buyer and data seller taking part in this process of trade are not direct opponents. By reviewing the literature published on data sharing, we identified 5 benefits associated with this concept and these were the initial codes generated for this category. The initial codes were:

- Identification of Critical Market Trends
- Reduction of Customer Segment Size (Micro Targeting)
- Elimination of Road Blocks for Better Sales and Services
- Improved Productivity
- Recognition of Missed Business Opportunities

4.2.1.1 Results & Analysis

To comprehend the above-mentioned findings empirically and understand the perception of data sharing among experts, the researcher put forward the question to the respondents: **What do you feel about the concept of pooling customer data from various sources (data sharing)? Do you believe by adopting this concept, the effectiveness of segmentation techniques can be improved?** The findings obtained and inferences identified in this category are explained below:

- In chapter 2 we discussed that telecom firms have reached their saturation point and are facing deteriorating ARPU levels. The main reasons associated with these volatile market trends are advancements in technology, increased competition, and emergence of OTT players such as YouTube, Instagram, Netflix et cetera. 8 out of the 10 respondents who participated in this thesis study agreed that with the adoption of data sharing, telecom firms can reap positive benefits and tackle the volatile market conditions effectively. For example, a respondent stated that as data buyer would acquire customer data from different organizations, this would enable the buyer to analyze the data from different perspectives. Such a detailed analysis would help the buyer to understand the needs of the customers and prospects in a wholistic manner thereby create a positive impact on the market segmentation process (Interview A). Arguing on similar lines, another expert felt that such detailed understanding of customer needs would also help in improving customer relationship management (Interview C). The academic researcher who participated in the discussions felt that by acquiring data from different sources, telecom firms will be in a better position to evaluate their current market position by comparing their offers with the needs of the customer. The respondent felt that such an evaluation will provide clear roadmaps to telecom manager for developing future strategic plans (Interview G).

> ***"As most of our data is behavioral data, we would like to adopt data sharing as it would help us answer these questions such as What is your favourite team? Which is your spending habit? Are you going to go this weekend to see this specific match? and questions like are there any frequent visitors over the stadium with having come along long-range tickets. Answers to these questions would certainly provide useful information to have to target better our clients. So far, we don't have them but, in the future, if we have access to such information, we could really utilize such answers and augment our views we have with respect to the behavior of subscribers. Therefore, by definition, I feel it will be useful"***
>
> ***-Interview C Respondent***

Similarly, the expert from India who participated in this research agreed on the above-mentioned viewpoints and explained the benefits for the Indian telecom market. He stated that the market leader in the Indian telecom market provides unlimited voice and data subscription plans at a monthly price of €2, and in order to stay competitive, rival telecom firms are forced to comply and offer services at the same price. He argued that the only way to combat the market leader and gain sustainable market share is by creating a compelling marketing mix along with the unlimited offerings. For example, he stated that with the help of data sharing if the firm acquires data regarding a user's interests in OTT services such as Amazon Prime, Netflix, Hulu or health insurance plans, the firm might be in a better position to combine these add-ons with the unlimited plans and thereby create an attractive marketing mix (Interview F).

- An interesting benefit regarding how data sharing can compete with a modern technology called 'deep packet inspection (DPI)' was provided by one of the respondents. There has been a growing interest among telecom firms to adopt this technology. This is because DPI performs advanced analytics on customer behavioral patterns generated. The working of DPI is illustrated below: In the current scenario, telecom firms are only able to understand the behavioral trends exhibited by a user in a limited manner i.e. if a user is watching a movie on Netflix, details such as name of the website, duration of internet usage are generated. However, with the help of DPI, a telecom firm can understand novel insights regarding a user's interest and desires and provide answers to questions such as *Which is your preferred movie genre? Who is your favourite superhero? et cetera*. The respondent stated that as DPI is in its early stages of research, high investment costs are associated with it which can only be afforded by big telecom firms. Additionally many deep packet inspection methods that are in practice perform the analysis is slow manner. The respondent felt that a similar value proposition that is offered by DPI and can also be generated from the concept of data sharing with lower investment costs. In other words, data sharing is a like to like replacement for deep packet inspection that can afforded by all the telecom firms across the globe (Interview F).

- Partnerships with non-rival firms and complementors are necessary for a firm to succeed in the market place. Exchange of information, resources and knowledge are often the catalysts for building and improving partnerships. The American respondent who participated in this study stated that as the concept of data sharing involves exchange of information, the concept can compound the understanding between actors. This increased in understanding between actors will develop new business ecosystems. The same respondent stated that as e-commerce firms such as Facebook and Amazon are tracking user interests in a detailed manner, telecom firms partnering with them would yield maximum benefits (Interview K).

- From the earlier points, we understood the perceived benefits of data sharing stated by the experts. However, to acquire complete knowledge on what expert observe on this concept it is also important to understand the viewpoints of experts who were not in favor of this concept. We earlier stated that 2 out of the 10 respondents rejected this idea of adopting data sharing in market segmentation. For instance, a respondent stated that they had adopted this concept of data sharing on a pilot basis, by partnering with a trusted 3rd party market research firms to understand the needs of customers. However the respondent faced low success rate as the data acquired did not overlap with the internal data possessed by the telecom firm. The respondent stated that for efficient data sharing it is imperative the actor's (data buyer and data seller) core business operations should be similar. Arguing on similar lines, the other respondent felt that telecom companies in Europe perceive this concept as intrusive to a user/customers private and hence are reluctant to adopt the concept (Interview A).

4.2.1.2 Drawing Conclusions

In this section, we shall analyze the viewpoints of experts from the previous subsection and update our theoretical findings. In chapter 2 we identified five benefits i.e. identification of critical market trends, recognizing missed business opportunities, reduction of customer segment size, improved productivity and eliminating roadblocks for better sales and services. Most of the respondents agreed to these benefits and additionally stated that by pooling customer data from various sources the following benefits are possible business ecosystem development, novel business models, creative marketing mix and improved customer relationships. Therefore, the total benefits identified are 9 which is the final code list for this category. 2 of the 10 respondents were not in favour of the concept and stated that the above-mentioned benefits are only possible when there is exists an overlap between the data seller's data and the data buyer's data. This is an important observation that we can infer from this finding. Telecom companies inclining to adopt the concept of data sharing should make sure that there is an overlap of data.

4.2.2 Category 2- Possible Barriers

From Chapter 2 we understood that data sharing has many benefits associated but every new method or a technology faces threats from potential barriers that can impede the adoption rate. Customer data involves

personally identifiable information, with stricter legislation such as GDPR in practice many firms have privacy and trust issue when sharing sensitive data. These were the main barriers that we found in Chapter 2. In the next subsection, we verify these findings empirically.

4.2.2.1 Results & Analysis

This section aims to analyze the viewpoints of experts on the possible barriers for data sharing.

- As the concept of data sharing involves the transfer of personally identifiable information of customers such as an address, gender, postal code et cetera most of the respondents expressed that with strict legislations such as General Data Privacy Regulation (GDPR) as main barrier that could hinder the adoption of this concept. For instance, a respondent stated that with hefty fines could be incurred in case of data breach, many telecom managers might not prefer to adopt this concept (Interview H). The same views were also expressed by respondents in Interview A and C. However, when we explained the role of privacy-enhancing technologies as coherent system that control protect privacy by eliminating or reducing personal data or by preventing unnecessary and/or undesired processing of personal data, all without losing the functionality of the information system all the respondents agreed to the and expressed their interest in this concept.

- Apart from GDPR, the Greek respondent who stated that lack of maturity in the market place to share data as another barrier affect the adoption rate. He stated the following:

> ***"The Greek market is not mature, exchanging data, especially between different market segments, such as financial data, operators, data and so forth. I don't know what's happening to more mature countries like France or the United Kingdom, but this type of data exchange is not happening in Greece".***
>
> ***-Interview C Respondent***

- The viewpoints expressed by the European respondents regarding GDPR were different from the views expressed by Asian respondents on the same issue. This is because privacy legislation in Asia are lenient and do not pose a big threat. For instance, the respondent India disagreed that privacy was the main barrier that could impede the adoption of this concept (Interview F).

4.2.2.2 Drawing Conclusions

Combining the insights obtained from theory and analyzing the viewpoints of experts we draw the following conclusions for this category.

- In chapter 2 we stated that with strict legislations such as General Data Privacy Regulation being enforced, companies are fearing hefty fines in case of privacy breach. As customer data involves sensitive

information, many firms are not inclined to perform market segmentation by pooling customer data from different sources. This argument was also supported by many of the respondents during our empirical analysis. It is therefore important to look for viable alternatives like enhancing the privacy of data before they are available in the data market, algorithms need to be used to judge the quality of data to provide a level playing field of all players in the market. We recommend the adoption of privacy-enhancing technologies (PET's) in sharing customer data. PET's are a coherent system of ICT measures that protect privacy by eliminating or reducing personal data or by preventing unnecessary and/or undesired processing of personal data, all without losing the functionality of the information system. examples of these technologies are Multi-Party Computation (MPC) – which helps the data owners to compute a function over their inputs while keeping their inputs private.

- Few respondents stated that in developing countries such as Greece, India and Indonesia lack of maturity among firms to share data is an important barrier. The main reason for the barrier due to the lack of research that analyse the impact of data sharing on market segmentation. Given the numerous benefits associated with data sharing, we feel the only way to help increase the maturity is conduction of pilot projects that share data in a small scale manner. By adopting the concept in a small scale manner will help the firms to understand and evaluate the overlap factor between the internal data and data acquired from sharing and also confirm the benefits associated with trading customer data.

4.2.3 Conclusion for Theme 2- Data Sharing

In the first theme, we aimed to understand the state of the market segmentation techniques and the challenges faced by firm during the implementation of the segmentation techniques. In this theme, we aimed to analyze if the challenges faced by the firm in implementing market segmentation could be nullified by pooling customer data from various sources. Majority of the respondents agreed that data sharing has the potential to improve the effectiveness of market segmentation techniques. A total of 9 benefits have been identified in this category. We also observed that all the benefits of data sharing are possible only if there exists an overlap of data between the actors participating in data sharing. Moving to the possible barriers, as data sharing involves sharing of customer data, privacy and trust were considered as the main barriers that could impede the adoption of data sharing. These factors can be controlled by Privacy Enhancing Technologies.

Table 11: Final Code List & Summary of Findings in Data Sharing

Category	Final Code list	Summary of Important Findings
Perception of Data Sharing	Identification of critical trends, Missed business opportunities, Reduction of segment size, Improved productivity, Business ecosystem development, Novel business model, Eliminating roadblocks for better sales and services, Improved customer relationship management & Creative marketing mix	1. Most of the respondents accepted that data sharing would help improve the effectiveness of segmentation techniques. 2. Few respondents expressed disagreement with data sharing due to failure of pilot projects that were done few years ago.
Possible Barriers	Strict legislation such as GDPR & Lack of maturity in the market	1. As customer data contains private information, most of the respondents expressed concerns regarding privacy legislations. 2. When we proposed the idea of Privacy Enhancing Technologies, all the respondent expressed desire to adopt them for sharing customer data in a secure manner. 3. As privacy is not a serious issue outside Europe, few Asian respondents stated that lack of maturity among companies could be a possible barrier.

Table 12: Visualization of Expert Responses in Data Sharing

Code/Interview Code	A	B	C	D	E	F	G	H	I	J
Category - Perception on Data Sharing										
Beneficial	✓		✓	✓	✓	✓	✓		✓	✓
Identification of Critical Trends	✓		✓		✓	✓	✓		✓	✓
Missed Business Opportunities			✓	✓		✓	✓		✓	✓
Reduction of Segment Size	✓		✓	✓		✓	✓		✓	✓
Improved Productivity			✓	✓	✓	✓	✓		✓	✓
Business Ecosystem Development			✓			✓	✓		✓	✓
Novel Business Models			✓			✓	✓		✓	✓
Eliminating Roadblocks for better Sales & Services	✓		✓		✓	✓	✓		✓	✓
Improved Customer Relationship Management	✓		✓	✓	✓	✓	✓		✓	✓
Creative Marketing Mix	✓				✓	✓				
Not beneficial		✓						✓		
Category - Possible Barriers										
Privacy issues due to General Data Privacy Regulation	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lack of maturity in the market			✓							

5. Illustration- Data sharing between Telco and a bank

Similar to the telecom industry, rapid advancements in technology and the explosion of the internet have increased the customer expectations on how banks engage with customers. This has forced banks to employ big data analytics in their daily operations (Ben Morgan, Geert Defreyn, 2017). Big data analytics refers to the examination of large data sets to discover hidden patterns, find correlations and gain valuable insights. Recent studies conducted by McKinsey across retail banks in United States of America, show that the banks' revenue could rise up to 8% if big data analytics is used (Amit Garg, Davide Grande, Gloria Macías-Lizaso Miranda, Christoph Sporleder, 2017). This is because big data and its analytics have the potential to increase productivity, develop novel services and enable faster and efficient decision making and aid in cost reduction. Findings like these have transformed the thinking process among bank managers and this has resulted in modern age banks across the globe investing a total of USD 20.8 billion on big data analytics (Vladimir Fedak, 2018).

To make banking hassle-free, banks have provided new ways of banking to customers. Digital banking, payments via plastic card such as debit and credit card and access to ATM vending machines are some of the common examples of modern banking methods (Alex Lane, 2019) As many customers have switched from traditional banking methods to modern methods, wealth of customer information is being captured by banks. A report from McKinsey reveals that more than 12 banks in Europe are employing machine learning to analyst customer behavioral patterns. Such an analysis is helping banks to profile customers in detailed manner and build accurate prediction models (James Ovenden, 2019).

From the above mentioned we can infer that banks have tools at their disposal to possess wealth of information regarding the customer. Having understood the tools and methods through which banks can possess detailed information of the customer, let's explore if the collected information can help other companies, e.g. a telecom company that intends to increase its customer base.

5.1 Relevant Data Points for Effective Data Sharing

By analyzing the interview responses in chapter 4 we understood that data sharing has numerous benefits such as micro-targeting, identifying critical patterns, eliminating roadblocks for better sales, improved productivity et cetera. However, to reap these benefits, it is imperative that there is an overlap of the data shared between the actors. To identify the relevant data points, that a bank needs to possess to help a telecom firm acquire new customer, we assume that the bank possesses two types of data i.e. demographic data and location details and are in a position to share these data in an anonymized manner.

Demographic Data - Demographics of customers residing in a particular geographical location has been widely used by firms across various industries to acquire a basic understanding of customer needs. The most common examples of demographic data that are used by telecom firms are age, gender, profession, level of education and marital status, household information and income.

Location Data – The fundamental purpose of market segmentation is to identify the right customer at the moment of time and target them with relevant offers. However, in order to reach the right customer at the right time, location data plays a pivotal role. Data that explain the geographical aspects of a customer such as name of the area, postal code, city of residence are known are location data.

With these above-mentioned assumptions and the use case, let's now look at solutions using the data from banks that telecom companies can use, to increase their revenue.

- **Data based on customer spending/saving information:** Banks store details regarding every transaction a customer initiates from his/her account. In other words, bank store information when money is debited or credited from an account. An individual receives a certain fixed income on a monthly basis. From this fixed income, the customer debit's money from the bank to spend on their daily living and other activities. Calculating the average amount of money spent in a given fixed time generates the average spending rate. If banks posses such data points, the financial inclination of a customer can be known i.e. such data can help us understand if a customer has a greater spending rate or greater saving rate. For e.g., let's assume that a bank is located near a university. In this bank, it is highly probable that most of the account holders are students and therefore, they don't have a huge spending rate. Using this information, telecom companies can come up with plans targetted towards these customers –e.g. unlimited high-speed data options/ free calls between friends, etc. So, if they put out flyers in the bank with such options, they might be able to attract many students. Thereby, data shared by banks can help the telecom firm to target the right customer segment.

- **Data from automatic transfers:** The automatic transfer service employed by a bank is an offering that transfers funds from one account to another account automatically when the customer grants permission once. Such kinds of service are used by the customers on their frequently used services or products. For e.g. people who often travel to various cities in The Netherlands, use the automated transfer service from their bank to credit money to Nederlandse Spoorwegen (NS) using their OV chipkaart card. If banks posses data points that list all the automatic debits transfer permitted by the customer, telecom firms can predict the interests of these frequent travelers and develop marketing mixes in line with the interests. For instance, let us assume that a customer has NS and Netflix in their list of automatic debit transfer. Telecom firms can make use of this information to collaborate with NS to provide free one-day travel offer across The Netherlands if the customer purchases a certain plan from the telecom operator or can provide more data usage for allowing the customers to access Netflix.

- **Data from ATM usage:** In order to make banking accessible to everyone at all possible times, banks have automated teller machines(ATMs) across various places. People make use of these ATMs to get their hard cash from their accounts. Typically, customers use an ATM's that are close to their houses or

workplaces due to ease of convenience. If bank posses data points that provide the location of the ATM most used by a segment of customers e.g. students, telecom firms can place flyers that contain offers for students in those ATMs.

- **Data based on transaction records:** Since all banks store the transaction details of their customers, telecom companies might be able to use this information to develop dedicated plans that might interest a market segment. For e.g., in today's global world, all of us travel internationally at least a few times of the year and we most often pay for those travels through digital payment methods – e.g. debit/card, bank transfer, etc. Telecom companies can use these transaction records to understand the % of frequent travelers and come up with interesting data/ minutes of calling plans abroad. To explain, usually when we go out of a country, we either need to pay exorbitant rates for global roaming or look for a local SIMcard of the country of choice which is both not our preferred options. To counter this, telecom companies can probably devise a plan that targets frequent travelers using bank information and entice them with free world roaming/ free data internationally if they use their services. Through this, telecom companies can benefit by increasing their market base and it also reduces the hassles of international travel for individuals.

- **Data from credit scores**: "A credit rating is a quantified assessment of the creditworthiness of a borrower in general terms or with respect to a particular debt or financial obligation" (Julia Kagan, 2019). This kind of data can be beneficial to telecom firms offering post-paid contracts. With the help of such information, telecom firms can first estimate the ability of a customer to pay the bills properly during the post-paid contract. Such an estimate will help telecom firms avoid customers with poor credit scores, thereby resulting in a reduction of losses for the telecom company.

5.2 Conclusions

The aim of this chapter was to help one of the partners of Safe-DEED and also answer our fourth research question. By reflecting on our main findings obtained in Chapter 4, we identified five relevant data points i.e. data based on customer spending/saving information, data from automatic transfers, data from ATM usages, data based on transactional records and finally data from credit scores. Therefore, this thesis study finds a few examples of the relevant data points that a bank should possess for an efficient data transfer with a telecom firm.

6. Ethical Aspects of Data Sharing

So far, in this thesis study, we have primarily focussed on the hidden potential of data sharing from a technical standpoint. Customer data contain personally identifiable information. By trading such data firms have the potential to trace and distinguish the identity of an individual customer and employ this data in an immoral way. Therefore, to complement this technical standpoint, we also need to look at the ethical aspects of data sharing. This chapter aims to answer our final research question **What are the ethical considerations for data sharing and market segmentation?** To answer this question in a comprehensive manner, it is important to first understand the fundamentals of ethics. The first section of this chapter we explain these fundamentals.

6.1 Basics of Ethics in a Nut-Shell

"When all of us think of ethics/morals, we think of rules to differentiate between right and wrong, such as the Golden Rule which says "Do unto others as you would have them do unto you", or as a code of professional conduct like the Hippocratic Oath that says "First of all, do no harm". The most common way of defining ethics is that there are norms for conduct that distinguish between acceptable and unacceptable behavior." (Resnik & Ph, 2011).

(Resnik & Ph, 2011) explain that most of us are taught the difference between right and wrong from childhood, moral development occurs through all stages of human life. "Although ethical norms might be considered as common sense, in reality, it is not, as each individual might interpret and apply these norms based on their individual interests, type of upbringing and personal experiences in life. "

Although most societies use laws to enforce widely accepted moral rules and ethical and legal rules use similar concepts, they are not completely the same. "In other words, an action might be legal but unethical or vice-versa, an action might be illegal but unethical". As an extension of this thought, we can use ethical concepts to criticize, evaluate, propose, or interpret laws. (Resnik & Ph, 2011). To summarize, it is important to understand the fundamental meaning of ethics and use it to look at concepts that might affect us as individuals or as a society as a whole.

6.2 Ethics in Big Data and Data Sharing

As we learned in earlier chapters, with the help of big data analytics, we can analyze large and complex data to discover hidden patterns using technological attributes such as velocity, volume, and variety. In their quest to tap into tremendous potential for big data, all companies continuously strive to extract as many benefits as possible. However, as companies mostly focus on the technological attributes, they sometimes forget/ignore social aspects of big data analytics. Based on (Someh, Davern, Breidbach, & Shanks, 2019), we understand that there are three major social processes that target and influence individuals: 1) data sourcing 2) data sharing and 3) algorithmic decision making. Firstly, many companies try to use a "catch-all-you-can" approach to collect as much data as possible from the individuals with an intention to quantify aspects of everyday life in an attempt to benefit companies doing the analytics. Secondly, this data that is captured from individuals moves from one

company to another until there is no further value left in the captured data. Thirdly, companies use algorithms on captured data to profile individuals on their race, gender, social status, date of birth, information in bank account which might conflict with the ethical aspects of right to freedom of individuals. This section is inspired by the work from (Someh et al., 2019).

As big data analytics broadly affects interactions among three major stakeholders – individuals, companies and society, let's understand their role in a detailed manner.

As all of us, as individuals, interact with devices on a day to day basis thereby generating this data. Companies use this data to identify hidden patterns/ trends and make decisions that impact individuals and on a broader sense, society. In addition, since companies have ownership of this data, they can also sell/share this data to other companies. Thirdly, societal agencies or government regulations need to control and regulate the sharing of the captured data to prevent loss of privacy/ sharing of information from the individuals. Hence, all the stakeholders need to be ethically aware of these aspects while generating, consuming and sharing data and understand the social side of big data analytics too.

(Someh et al., 2019)summarize the concepts and their relative importance to each of these stakeholders.

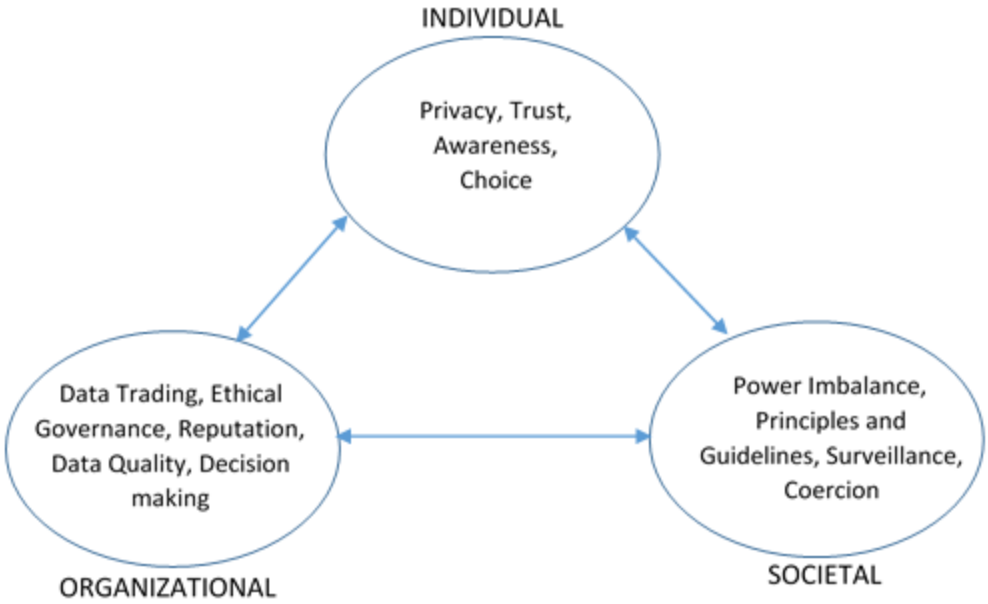

Figure 9: Ethical Concepts for Three Stakeholders

As the data is generated by individuals, let's understand the major concepts that affect the individual. They are **privacy, trust, awareness, and choice.** In the later paragraphs, we will follow this up with the concepts that affect companies (organizations) and society.

- **Privacy:** Every individual's ability to restrict and control how organizations use and disclose their personal information is known as privacy.

- **Trust:** Individuals need to understood that companies that capture the data use the data only for intended purposes and that they don't misuse this data or share it to others without the consent of the individuals.
- **Awareness:** Each individual needs to be aware of the fundamentals of big data analytics and at least get a broad overview of how companies can use the captured data to offer products and services.
- **Choice:** Since companies use the captured data to provide products or services based on their personal details, companies should not restrict the choice of options for the individuals. Individuals shouldn't be judged based on the data shared and should always have the choice to make decisions without any restrictions.

Having looked at the ethical considerations from an individual's perspective, let's now look the considerations from the eyes of an organization which is Data trading, Ethical governance, Reputation, Data quality and Algorithmic decision making.

- **Data trading:** As data is shared between shareholders, it can be treated as a commodity. Organizations get the data from individuals either when individuals use their services/products directly or by buying the data from data aggregators. The methods through which organizations source, share, and trade data can raise ethical issues like not asking for explicit consent from individuals to share data, not being transparent about companies that will use the shared data/the reason for sharing and not maintaining the anonymity of the data when shared between organizations thereby causing data misuse.
- **Ethical governance:** In order to ensure that all organizations have the same understanding of the rules and regulations of data sharing, it is important to properly define all the procedures and best-known practices. In addition, there should be regulatory bodies within the government/ society that takes care that these ethical procedures are aligned to, by all companies. Also since different countries can have different local laws concerning data sharing, companies need to be aware of them and consider them while sharing data between different entities.
- **Reputation:** This refers to the belief/feeling that individuals have about companies that acquire their data to manage and use it ethically. Usually, it is enough for one ethical violation to destroy the company's reputation and therefore, it is important for companies to underline and adhere to the ethical practices at all time and treat customer data with the utmost integrity.
- **Data quality:** Quality of captured data plays a big role in companies identifying hidden opportunities to increase their market share/revenue/ other parameters. Three main problems that can occur with quality of data are a) Since big data is captured from multiple sources, it could be that there is no alignment between the different sources thereby degrading the quality of data when aggregated. b) Even though data that is captured from the individual doesn't have any personally identifiable information, combined data obtained from various sources might still be able to provide information about individuals which were originally not intended to be captured. This is also called as "mosaic effect".

c) When organizations source data, they may not establish complete data traceability resulting in a loss of metadata.

- **Algorithmic decision making:** Companies use algorithms on captured data to determine future business decisions based on historical and subjective data. As data continuously increases in size, speed and complexity, algorithms need to be robust enough to evolve and react on growing data. However, as algorithms might have underlying assumptions, limitations, biases, and data quality issues, it is important to not completely rely on the algorithmic output. Instead, manual intervention or at least a plausibility check of the results must be done to avoid unethical/ discriminatory classification.

Finally, to understand the ethical aspects as a whole, let's discuss necessary ethical considerations at the societal level, in the subsequent paragraphs. Major concepts that affect ethical issues are Power imbalance, Principles, and guidelines, Surveillance and Coercion.

- **Power imbalance:** It concerns the power, control or influence relationships that can arise from big data analytics in society. This imbalance arises because some entities might dominate access to big data. From an ethical perspective, this restricts the freedom of choice and creates inequality among individuals. Groups, organizations or other entities might have access to a huge database of individual information and as a result, they might use this power to influence or modify individuals' behavior to generate economic or political value.
- **Principles and guidelines:** As each organization might have different principles and varying levels of guidelines on maintaining the integrity of captured data, regulatory bodies need to define a common set of guidelines. In addition, government or societal agencies must ensure that all organizations implement the guidelines to keep data safe.
- **Surveillance:** It refers to the extent to which organizations observe and monitor individual's lives. As many aspects of individuals' everyday life by the captured data, it is important for government/regulatory bodies to keep a check on the organizations and ensure that the captured data does not invade the privacy of the individuals'.
- **Coercion:** It refers to the organizations trying to convince/ force individuals to share their data. Forcing or convincing individuals to share their data without their acceptance is unethical and should not be followed. For e.g. each individual must know that his data will be captured for an explicitly defined reason and he should have the flexibility to decide if he wants his data to be captured or not captured. If individuals decide not to let their data be captured, organizations should not restrict their services to these individuals.

6.3 Implications to Chapter 4 Findings

In the previous section by taking the perspectives of three stakeholders i.e. individual, organization and society we identified a total of 12 ethical considerations that need to be contemplated when customer data is traded for market segmentation. In this section, we aim to relate these ethical considerations on our Chapter 4 findings.

1. We stated earlier that this thesis study considers data structures as a set of information to enable firms to make accurate analytical and data-driven decisions. In Chapter 4 we found that telco's use demographic, behavioral and geographical data to perform market segmentation. During the empirical analysis we observed that firms are inclining themselves to 'catch-all you can' attitude with respect to customer data. As customer data is a pivotal input for companies to perform market segmentation, telecom companies might look for options to misuse or exchange customer data between organizations without taking the customers' consent. In order to ensure that customers feel valued and understand the value of their data, customers must be informed about which types of data are being collected, for which purposes and if the data is going to be shared, with which of these partners. In addition, customers should have the flexibility to accept/reject the collection of their data and companies should not try to force customers to share their data.

2. In chapter 2 and 4 we understood the fundamentals of market segmentation. We stated that firms employ market segmentation with an aim to understand the needs of the customer segments. By understanding this needs firms try and develop relevant marketing mixes that satisfy the needs of customer segments. Similarly we found that by pooling customer data, telecom firms can create relevant and attractive marketing mixes that satisfy the needs of the customers. Taking the perspective of an individual we argue telecom firms despite understanding the needs of the customers should not offer restricted choice of options to customers that satisfy their needs. Firms should aim to provide all options to customers at all times and should allow the customer to take their own decisions.

3. We also found that behavioral segmentation was the most preferred segmentation technique by telecom firms. In this segmentation technique, every action i.e. via a call, message or data initiated by the customer is recorded and analyzed in a detailed manner with an aim to understand the behavioural trends of the customers. In the end customers with similar behavioral trends are segregated. In relating this finding to the ethical consideration we argue from two perspectives organization and society. Firstly from an organizational perspective we explicate that telecom firms should not only rely solutions provided by data mining algorithms and segregate customers. Manual intervention on the results generated by these algorithms should be doe. Secondly, By taking the perspective of society, with a constant urge and need to employ an effective behavioural segmentation, organizations should not constantly observe actions of customer over the network and collect data.

4. One of the benefits associated with pooling customer data from different sources for market segmentation is the reduction of customer segment size. When organizations perform market segmentation by pooling customer data care should be taken regarding the minimum size of each customer segment. Typically, the size of customer segment of an organization is dependent on three factors i.e. size of the company, availability of management resources and size of the customer base. For e.g. a small-medium enterprise might have a lower segment size when compared to a large organization. As customer data involves personally identifiable information, in order to avoid privacy breach and identify customer's on an individual level, we propose that a minimum size should be between 100 - 1000 customers. This is because the probability of identifying customer on an individual level is minimal for this range when compared range lying between 10-100.

5. Another important benefit associated with the sharing of customer data is identification of critical customer trends. Considering a hypothetical situation if a telecom firm pools customer data and identifies critical behavioral trends, the firm should not share such kind of information with political organizations and create an imbalance in society.

6.4 Summary

Big data analytics is a complex phenomenon that creates opportunities for further development of human lives but also has ethical implications that need to be considered. As the interactions of big data analytics affect many stakeholders, in this section we identified three main stakeholders i.e. individual, organization, and society. Privacy, trust, awareness, and choice are the main ethical consideration from an individual viewpoint whereas, from an organizational perspective data trading, reputation, ethical governance, data quality, and algorithmic decision making are the main ethical considerations. For the organizations we also explained the minimum segment size should be between 100-1000 so that the probability and risk of identifying customers on an individual level are negligible Finally for the society we identified power imbalance, principles, and guidelines, surveillance and coercion as the main ethical aspects. Therefore, these are the ethical considerations that need to be contemplated when customer data is shared.

7. Conclusions

Ever since the digital age began in the early 21st century, there has been a stupendous increase in the amount of data that is generated and captured by electronic devices around us. With huge volumes of data being easily available, researchers have coined the buzzword – big data, to extract useful information from these huge volumes of data. Companies too have realized that big data can help them comprehensively understand the needs of the consumer and as a result, have embraced the concept of market segmentation. Market segmentation is defined as a process of dividing a large heterogeneous market into smaller homogenous markets consisting of customer segments with similar wants, needs, and desires. Although companies have been regularly practicing market segmentation, some of the major problems that they have is that their segmentation techniques are not up to date or ineffective since the data that they use to segment the market is old or the data is incomplete (siloed). One way to overcome these problems is using the concept of data sharing, where datasets are sold/ shared between partners as a commodity, in a data market. However, in reality, the adoption of data markets has been slow presumably because of lack of value for companies participating in the market or because of privacy or legal concerns.

In order to increase the adoption rate of data sharing in a data market, the research project- Safe-DEED explores the idea of a decentralized data market, which is enabled by privacy enabling technologies. The adoption rate of these decentralized data markets depends on building viable business models for the marketplaces as well as the benefits associated with it. In order to identify one such benefit is analysing the impact of data sharing on market segmentation With this thought in mind, as part of this research project, we set our goal of this thesis to evaluate the effectiveness of market segmentation techniques using data sharing and for this thesis, we decided to focus on telecom industry as an example since many companies in this industry are currently facing challenges such as inability to increase revenue, decrease of customer base, etc..

As a first step in this thesis, we understood the current problems of telecom industry and learnt that the telecom industry has been witnessing a continuously changing business and technology environment due to the increased competition and emergence of Over the top service players such as Netflix and YouTube, who don't need to partner with telecom firms to provide their content. Usually, telecom firms possess two types of data i.e. internal data (e.g. customer contracts, traffic data, billing data, etc.) and external data (data mergers with business partners) regarding customers. Majority of the telecom companies employ internal data to perform market segmentation. By collecting data from a limited perspective, companies possessed incomplete data sets which in turn lead telecom firms to create large customer segments and develop ineffective plans and services, thereby reducing the average revenue per user. In addition to understanding these current-day problems of telecom industry, we also researched the current market segmentation techniques in the industry and found out the most common segmentation techniques from literature. In the third step of this thesis, we spoke to market segmentation experts in the telecom industry and understood their opinions on current market segmentation techniques and problems, and potential benefits of data sharing using semi-structured interviews.

Through the summary of these interviews with segmentation experts, we observed that by pooling customer data in an anonymized form different sources (data sharing), telecom firms can identify critical customer trends, target micro-segment of the customer, eliminate roadblocks for better sales and services and profile the customer needs in a detailed manner. As a result, these benefits can help telecom firms better understand changing customer needs and perform effective market segmentation. Therefore, through this thesis, we summarize that data sharing can help telecom firms to segment the markets more effectively.

As a means to our conclusion of this thesis, we broke down our main research question into five sub research questions that are listed below. In the following paragraphs, let's understand these questions in greater detail along with a short summary of the context, approach, and results of each of these questions.

SQ1: What are the relevant market segmentation techniques for telecommunication firms?

We answered this question by performing a literature review of the published articles on market segmentation for telecom firms. To answer this sub research question in a comprehensive manner, we first explained the fundamentals of market segmentation, listed the 10 steps involved in implementing market segmentation and then moved on to market segmentation in the telecom industry. As (Bayer, 2010) was the only paper that provided insights on the relevant market segmentation techniques in telecom industry, we took inspiration from her work and identified four market segmentation techniques i.e. customer value segmentation, customer value segmentation, customer lifecycle segmentation and finally customer migration segmentation. In the first technique -customer value segmentation, telecom firms calculate 'past value' by analyzing the history of transactions the customer had with the firm and estimate the probability of future revenue from the customer by calculating the future value. Adding both these values i.e. the past value and future value, telecom firm obtain 'customer value' for every customer. Based on the calculated customer value, decile analysis i.e. splitting into 10 equal segments is used to perform customer value segmentation. In customer behavior segmentation, every time a customer initiates an action via voice, message, data over the telecom network, characteristics such as duration of the call, type of the call, the amount of data consumed, number of SMS's are the few aspects that are analyzed. The resulting information is used to segment customers who show similar behavioral trends. In the third segmentation technique- customer lifecycle segmentation, customers are segmented into four life stages i.e. new customer, growing customer, maturing customer, and declining customer. These life stages are identified from the duration the customer was associated with the telecom firm. The premise of this segmentation technique lies in segmenting the varying customer behaviour exhibited at different life stages. Finally, in customer migration segmentation- customer loyalty is analyzed. As time progresses, it is very common for customers to increase or decrease their loyalty. By analysing at different time frames, customer value measure can increase or decrease i.e. customers can migrate between different segments and customer satisfaction and loyalty patterns can be identified (Bayer, 2010). Such identification of satisfaction and loyalty patterns can help firms predict customer churn before it occurs. Therefore, telecom firms can design and develop compelling activities for customer segments who have a high probability to switch loyalties to other competing companies (Ascarza, 2018). The common aspect of all these four segmentation techniques is that

their effectiveness and success rate is heavily reliant on the collected data. Therefore, it is imperative that firms employ data with utmost quality during the implementation.

Having understood that the data is playing a pivotal role in these segmentation techniques, it now important to understand the benefits of data sharing for market segmentation.

SQ2: What are the benefits of data sharing for market segmentation?

This question was also answered by performing a literature review on published academic journals and white paper articles on data sharing. We define data sharing "as the process by which a company makes data available to another company that is neither a direct market competitor nor a sub-contractor and is interested in these data for its own business purposes" (Arnaut et al., 2018). As market segmentation techniques are heavily reliant on the data for their effectiveness, it is important for firms to collect customer data from different sources. This helps the firm build complete profiles of customers and understand their needs and wants in a comprehensive manner. Such an understanding will help telecom firms develop compelling marketing strategies and thereby increase the average revenue per user (ARPU) levels. By pooling customer data from different sources firms will be in a position to identify critical customer patterns, eliminate roadblocks for better sales and services, recognize missed business opportunities, improve productivity and finally target micro customer segments of customers. All these benefits have the potential to improve the effectiveness of market segmentation techniques.

After identifying the challenges in market segmentation and benefits of data sharing from theory, we wanted to understand more about these findings from an industry perspective to gain a comprehensive understanding of the benefits of data sharing for market segmentation. This led to us to frame our third research question:

SQ3: Which segmentation techniques are used in the telecom industry and what kind of challenges were observed during their implementation?

We used interviews as a research method to answer this question. Judgment sampling, a type of purposive sampling was used to identify respondents (associated with telecom firms across the globe) with expertise in market segmentation. Semi-structured interviews were adopted to discuss a set of 9 questions in every interview. A total of 10 interviews were conducted in this thesis study. From the interview responses, we found that telecom firms employ various segmentation techniques ranging from behavioral segmentation, bandwidth segmentation, tower-based segmentation, geographical segmentation, psychographic segmentation, value segmentation, life cycle segmentation, and customer churn prediction technique. During our discussion with the experts, we learned that the main challenge commonly faced during the implementation of the market segmentation technique was incomplete customer data sets and obsolete data present in the database. Due to these challenges', companies were unable to employ effective segmentation techniques and thereby witness low average revenue per user (ARPU) levels. These respondents stated that by pooling customer data from various different sources, the data related challenges could be superseded. However, few respondents stated that technical challenges such as lack of state-of-the-art technological tools to store large amounts of data and

inadequate data analytics resources were also faced. In this regard, the respondents stated that as the challenges were technical, data sharing was not a useful mechanism to improve the effectiveness of segmentation techniques.

As the notion of using data sharing to segment markets in the telecom industry is new, we wanted to use the results obtained by analysing the interview responses to illustrate a use case scenario between two organizations. This led to our fourth research question:

SQ4: Which data structures should a bank possess for an effective data sharing with a telecom firm?

We reflected the knowledge gained from interview responses and literature review to answer this question. With an aim to identify cross-selling opportunities, we used a use case scenario of a telecom operator and a bank that are in a collaboration to share each other's CRM data. To explain the use case - As the telecom firms want to continually increase the market share, they want to identify bank's customers to understand their behavioral patterns and target those customers with potentially interesting plans and offers. From this use case we identified that data about customer's spending/saving pattern, data from automatic transfers, data from ATM usage, data based on transaction records, data from credit scores along with the demographic and geographical data as some of the necessary data points that a bank needs to possess for an effective data transfer with the telecom firm.

As sharing of customer data involves personally identifiable information, we also discussed the ethical aspects that need to be considered by organizations when embracing the concept of data sharing.

SQ5: What are the ethical considerations for data sharing and market segmentation?

With this rise of big data as a socio-technical phenomenon, researchers have only explained the concept from a technological perspective. However, as there was not a lot of focus on the ethical aspects of big data, we looked at it through the eyes of three major stakeholders namely individuals, organizations and society. From the perspective of an individual, we identified privacy, trust, awareness, and choice as the main ethical considerations. Moving to the perspective of an organization we identified data trading, ethical governance, reputation, and data quality as the ethical considerations. Finally, from the ethical considerations at the societal level we identified power imbalance, principles, and guidelines, surveillance and coercion. By considering each of these ethical considerations for all these stakeholders, we would also be able to appreciate the potential of data sharing in a data market.

Therefore, from the main research question and from these five research questions, we strongly feel that using data sharing to effectively segment a market can become a reality soon.

8. Reflections

A vital part of every thesis is a self reflection of the thesis to understand and present the contributions and limitations and future work of the work to the scientific community. Therefore, in this chapter, we will self introspect them. We divide this chapter into three sections. The first section explains of theoretical and practical contributions of this thesis study. In the second section we describe the limitation that was faced during the thesis study. Finally in the last section we provide future research directions for researchers.

8.1 Contributions

In the earlier chapter, we stated that that telecom industry is encountering volatile market conditions and facing low ARPU levels from customers. This is because the telecom operators are unable to perform effective market segmentation. Additionally, we found that there was not a lot of literature that explained the appropriate market segmentation techniques that could be employed by telecom firms. Combining our theoretical findings and empirical findings we identified 10 relevant market segmentation techniques for telecom firms.

Secondly, we identified and explained the benefits of data sharing for market segmentation. In this research, we introduced the fundamentals of market segmentation and data sharing. Subsequently, during our interviews, we pitched our ideas of benefits of data sharing and the interviewers also shared our opinions and felt that sharing data can help companies increase their market position/revenue. This goes on to prove that there is a strong interest in the real world about need for market segmentation techniques for our research problem.

In the process of discussing our findings, we evaluated a use-case scenario where a telecom company was looking for ways to improve its customer base using the shared data between it and a bank. As part of this use-case scenario, we provided 5 viable solutions that the telecom company could use from the data points in a bank, to target potential customer segments and thereby, improve this base. Although the evaluation in the use-case scenario was only between two entities, it can also be extrapolated to multiple entities which results in a lot more data points finally leading into more precise customer segments.

As customer data contains sensitive information, in addition to the technological perspective we also identified ethical aspects that need to contemplated before trading customer data. In the next subsection, we will address the limitations that were encountered as part of the thesis study.

8.2 Limitations

As we used interviews as a data collection technique, there is an inherent bias involved since each of the respondents might have their opinion to some of the questions that we asked. Hypothetically, if we were to interview two market segmentation experts from the same company, they might have different opinions on some aspects which would result in a difference in the collected data.

Out of the 10 respondents who participated in this research, 8 interviews were conducted using Skype video conference as the respondents lived in different geographical locations such as Indonesia, India, Belgium, and Singapore. 2 of the 10 interviews were conducted in a face to face manner as the respondents lived close to Delft. As there are a lot more advantages of having a face to face conversation in a natural setting, having to use Skype for the interviews turned out to be a limitation especially in cases where people were not comfortable with the usage of Skype.

We used judgemental sampling to find experts. Due to busy work commitments, many experts who possessed great expertise in market segmentation refused to participate. Had this thesis study interviewed these experts, we could have probably received more valuable insights which in turn might have led to more effective results.

Duration of each interview was 60 minutes. However, for a few interviews, the time was not sufficient, as the respondents provided answers that were out of the scope of the study. For instance, many respondents gave insights regarding segmentation in the business market along with the consumer market.

8.3 Future Work

For future work, we think that these could be potential starting points for future researchers interested in conducting research in this field.

- We interviewed 10 respondents and sought their opinion on the uses of data sharing for market segmentation. Future researchers can conduct the same study by increasing the sample size to validate the results of this study.
- As we used a qualitative method through interviews to validate our theory, the next topic is to also validate this theory using quantitative approaches. This would help in achieving more confidence about the benefits of data sharing
- There is enough literature published on market segmentation in tourism and pharmaceuticals industry. Future research can be done to verify the effectiveness of data sharing in those sectors.

References

Ahn, J. H., Han, S. P., & Lee, Y. S. (2006). Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry. *Telecommunications Policy*, *30*(10–11), 552–568. https://doi.org/10.1016/j.telpol.2006.09.006

Alex Lane. (2019). How Banks Are Using Customer Data To Drive Success | Articles | Analytics | Innovation Enterprise. Retrieved August 28, 2019, from https://channels.theinnovationenterprise.com/articles/how-banks-are-using-customer-data-to-drive-success

Alshenqeeti, H. (2014). Interviewing as a Data Collection Method: A Critical Review. *English Linguistics Research*, *3*(1), 39–45. https://doi.org/10.5430/elr.v3n1p39

Amit Garg, Davide Grande, Gloria Macías-Lizaso Miranda, Christoph Sporleder, and E. W. (2017). Analytics in banking: Time to realize the value | McKinsey. Retrieved August 28, 2019, from https://www.mckinsey.com/industries/financial-services/our-insights/analytics-in-banking-time-to-realize-the-value

Ansell, J., Harrison, T., & Archibald, T. (2007). Identifying cross-selling opportunities, using lifestyle segmentation and survival analysis. *Marketing Intelligence and Planning*, *25*(4), 394–410. https://doi.org/10.1108/02634500710754619

Arnaut, C., Pont, M., Scaria, E., Berghmans, A., & Lecont, S. (2018). *Digital Single Market Study on data sharing between companies in Europe*. https://doi.org/10.2759/354943

Ascarza, E. V. A. (2018). Retention Futility : Targeting High-Risk Customers Might Be Ineffective, *LV*(February), 80–98. https://doi.org/10.1509/jmr.16.0163

Bandyopadhyay, P. (2016). Best Practices of Campaign Management for a Telcom Operator. *Journal of Telecommunications System & Management*, *5*(2). https://doi.org/10.4172/2167-0919.1000135

Barry Harmsen. (2010). Decile analysis - The Qlik Fix! The Qlik Fix! Retrieved May 22, 2019, from http://www.qlikfix.com/2010/10/08/decile-analysis/

Bayer, J. (2010). Customer segmentation in the telecommunications industry. *Journal of Database Marketing and Customer Strategy Management*, *17*(3–4), 247–256. https://doi.org/10.1057/dbm.2010.21

Ben Morgan, Geert Defreyn, J. E. (2017). *Re-envisioning the customer experience Fundamentally changing customer relationships with new ways to create value*. Retrieved from www.deloitte.com/digitalbank

Bhattacherjee, A. (2012). *Introduction to Research, Social Science Research: Principles, Methods, and Practices*. *Pure and Applied Chemistry*. https://doi.org/10.1351/pac198961091657

Brennan, B. (2017). A Millennial Life Stage "Reality Check" For Marketers. Retrieved June 1, 2019, from https://www.forbes.com/sites/bridgetbrennan/2017/06/30/a-millennial-life-stage-reality-check-for-marketers/#433f6db46c72

Brown, B., Kanagasabai, K., Pant, P., & Serpa Pinto, G. (2017). Capturing value from your customer data |

McKinsey. Retrieved May 7, 2019, from https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/capturing-value-from-your-customer-data

Cho, J. Y., Lee, E.-H., Ji Young, C., & Eun-Hee, L. (2014). The qualitative report reducing confusion about grounded theory and qualitative content analysis: Similarities and differences. *Qualitative Report*, *19*(32), 1–20. https://doi.org/http://www.nova.edu/ssss/QR/QR19/cho64.pdf

Clavorà Braulin, F., & Valletti, T. (2016). Selling customer information to competing firms. *Economics Letters*, *149*, 10–14. https://doi.org/10.1016/j.econlet.2016.10.005

de Reuver, M. (2019). MOT2312 Qualitative Data Analysis Lecture Slides.

Deichmann, B. J., Heineke, K., Reinbacher, T., & Wee, D. (2016). Creating a successful Internet of Things data marketplace. *Accenture*, (October), 1–8.

Desai, G. (2019). How to Use Customer Behavior Data to Drive Revenue (Like Amazon, Netflix & Google). Retrieved May 7, 2019, from https://www.pointillist.com/blog/customer-behavior-data/

Dolnicar, S., Grün, B., & Leisch, F. (2018). *Market Segmentation BT - Market Segmentation Analysis: Understanding It, Doing It, and Making It Useful*. https://doi.org/10.1007/978-981-10-8818-6_1

Ernst, D., & Dolnicar, S. (2018). How to Avoid Random Market Segmentation Solutions. *Journal of Travel Research*, *57*(1), 69–82. https://doi.org/10.1177/0047287516684978

Fadhil Mohammed, M. (2013). Customer Relationship Management (Telecommunication Industry) Comparison Between (Airtel) And (Zain). *International Journal of Business and Management Invention ISSN*, *2*(11), 52–58. Retrieved from www.ijbmi.org

Femina, B. T., & Sudheep, E. M. (2015). An efficient CRM-data mining framework for the prediction of customer behaviour. *Procedia Computer Science*, *46*(Icict 2014), 725–731. https://doi.org/10.1016/j.procs.2015.02.136

Gill, P., Stewart, K., Treasure, E., & Chadwick, B. (2008). Methods of data collection in qualitative research: Interviews and focus groups. *British Dental Journal*, *204*(6), 291–295. https://doi.org/10.1038/bdj.2008.192

Goyat, S. (2011). The basis of market segmentation: a critical review of literature. *European Journal of Business and Management*, *3*(9), 45–54. Retrieved from http://www.iiste.org/Journals/index.php/EJBM/article/view/647

Guest, M., Struthers, T., Jones, D., & Schachtel, J. (2015). A people-based telecom business: Breathing new life into segmentation strategies, (March). Retrieved from https://www2.deloitte.com/uk/en/pages/technology-media-and-telecommunications/articles/a-people-based-telecom-business.html

Hamka, F., Bouwman, H., De Reuver, M., & Kroesen, M. (2014). Mobile customer segmentation based on smartphone measurement. *Telematics and Informatics*, *31*(2), 220–227. https://doi.org/10.1016/j.tele.2013.08.006

Han, S. H., Lu, S. X., & Leung, S. C. H. (2012). Segmentation of telecom customers based on customer value by

decision tree model. *Expert Systems with Applications*, *39*(4), 3964–3973. https://doi.org/10.1016/j.eswa.2011.09.034

Harting, A., Sommer, S., & Schlosser, D. J. (2017). Data Driven Marketing How efficient and personalized customer dialog will work in future ? *Deloitte White Paper*.

James Ovenden. (2019). The Future Of Data In Banking | Articles | Big Data | Innovation Enterprise. Retrieved August 28, 2019, from https://channels.theinnovationenterprise.com/articles/the-future-of-data-in-banking

Jamshed, S. (2014). Qualitative research method - interviewing and observation, *5*(4), 4–5. https://doi.org/10.4103/0976-0105.141942

Joshi, H. (2018). Digital Media: Rise of On-demand Content, 5–7. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/in/Documents/technology-media-telecommunications/in-tmt-rise-of-on-demand-content.pdf

Julia Kagan. (2019). Credit Rating. Retrieved August 28, 2019, from https://www.investopedia.com/terms/c/creditrating.asp

Kim, S. Y., Jung, T. S., Suh, E. H., & Hwang, H. S. (2006). Customer segmentation and strategy development based on customer lifetime value: A case study. *Expert Systems with Applications*, *31*(1), 101–107. https://doi.org/10.1016/j.eswa.2005.09.004

Kotler, P., & Zeithaml, V. (2000). Marketing Management , Millenium Edition. *Marketing Management*, *23*(6), 188–193. https://doi.org/10.1016/0024-6301(90)90145-T

Kuusik, A., & Varblane, U. (2009). How to avoid customers leaving: The case of the Estonian telecommunication industry. *Baltic Journal of Management*, *4*(1), 66–79. https://doi.org/10.1108/17465260910930458

Kvale, S. (2006). Dominance through interviews and dialogues. *Qualitative Inquiry*, *12*(3), 480–500. https://doi.org/10.1177/1077800406286235

Levitt, T. (1965). Exploit the Product Life Cycle. Retrieved June 1, 2019, from https://hbr.org/1965/11/exploit-the-product-life-cycle

Lin, H. H., & Wang, Y. S. (2006). An examination of the determinants of customer loyalty in mobile commerce contexts. *Information and Management*, *43*(3), 271–282. https://doi.org/10.1016/j.im.2005.08.001

Lupu, M. (2018). Safe-DEED :Safe Data Enabled Economic Development. *SafeDEED Project Proposal*, *1*.

Maji, G., & Sen, S. (2016). Data warehouse based analysis on CDR to retain and acquire customers by targeted marketing. *2016 5th International Conference on Reliability, Infocom Technologies and Optimization, ICRITO 2016: Trends and Future Directions*, 221–227. https://doi.org/10.1109/ICRITO.2016.7784955

Manyika, J., Chui, M., B., B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). Big data: The next frontier for innovation, competition and productivity. *McKinsey Global Institute*, (May).

Michael Schrage. (2017). AI Is Going to Change the 80/20 Rule. Retrieved August 28, 2019, from https://hbr.org/2017/02/ai-is-going-to-change-the-8020-rule

Mindi, C. (2016). Why marketers are failing to target consumers at key life events – Marketing Week. Retrieved June 1, 2019, from https://www.marketingweek.com/2016/03/04/why-marketers-are-failing-to-target-consumers-at-key-life-events/

Molly Galetto. (2016). What is Customer Churn? Definition and How to Reduce It – NGDATA. Retrieved June 4, 2019, from https://www.ngdata.com/what-is-customer-churn/

Namvar, A., Ghazanfari, M., & Naderpour, M. (2018). A customer segmentation framework for targeted marketing in telecommunication. *Proceedings of the 2017 12th International Conference on Intelligent Systems and Knowledge Engineering, ISKE 2017*, *2018-Janua*, 1–6. https://doi.org/10.1109/ISKE.2017.8258803

Neumann, N. (2017). The power of big data and algorithms for advertising and customer communication. *2016 International Workshop on Big Data and Information Security, IWBIS 2016*, 13–14. https://doi.org/10.1109/IWBIS.2016.7872882

Nimako, S. G. (2016). Consumer Switching Behaviour: A Theoretical Review and Research agenda. *Research Journal of Social Science & Management*, *02*(3), 74–82.

Norman, G., Pepall, L., Richards, D., & Tan, L. (2016). Competition and consumer data: The good, the bad, and the ugly. *Research in Economics*, *70*(4), 752–765. https://doi.org/10.1016/j.rie.2016.09.002

Resnik, D. B., & Ph, D. (2011). What is Ethics in Research & Why is it Important ?, 1–10.

Roland Ortt, J., & Smits, R. (2006). Innovation management: Different approaches to cope with the same trends. *International Journal of Technology Management*, *34*(3–4), 296–318. https://doi.org/10.1504/IJTM.2006.009461

Saldana, J. (2009). *The Coding Manual for Qualitative Researchers - Johnny Saldana - Google Books*. *SAGE Publications Ltd*. https://doi.org/10.1017/CBO9781107415324.004

Sammour, G., Schreurs, J., & Vanhoof, K. (2009). Data Mining and Marketing Approach To Track Customer Movements. *Euromedia'2009*, (February), 5–9.

Schroeck, M., Shockley, R., Dr. Smart, J., Prof. Dolores, R.-M., Prof. Tufano, P., Balboni, F., … White, K. (2012). IBM Global Business - The real-world use of big data. Retrieved from https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?subtype=XB&infotype=PM&appname=GBSE_GB_BT_USEN&htmlfid=GBE03519USEN&attachment=GBE03519USEN.PDF

Sekaran, U., & Bougie, R. (2016). Research Methods For Business- A Skill Building Approach.

Shanks, G., & Tay, E. (2001). Global Co-Operation in the New Millennium THE ROLE OF KNOWLEDGE MANAGEMENT IN MOVING TO A CUSTOMER-FOCUSED ORGANISATION. *The 9th European Conference on Information Systems*, 721–729. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.107.9795&rep=rep1&type=pdf

ShuiXiu, L., & Shuihua, H. (2012). Classifying the segmentation of Telecom Customer Value Using Decision Tree Model. *Expert Systems with Applications*, *39*(4), 3964–3973. https://doi.org/10.1016/j.eswa.2011.09.034

Smit, A. (2018). Making The Best Use Of Customer Data. Retrieved May 6, 2019, from https://www.forbes.com/sites/anthonysmith/2018/12/20/making-the-best-use-of-customer-data/#9c79f943eea1

Smith, W. R. (1956). Product Differentiation and Market Segmentation as Alternative Marketing Strategies Author (s): Wendell R . Smith Published by : American Marketing Association Stable URL : http://www.jstor.org/stable/1247695 PRODUCT DIFFERENTIATION AND MARKET SEGMENTAT. *American Marketing Association*, *21*(1), 3–8. https://doi.org/10.2307/1247695

Someh, I., Davern, M., Breidbach, C. F., & Shanks, G. (2019). Ethical Issues in Big Data Analytics: A Stakeholder Perspective. *Communications of the Association for Information Systems*, *44*, 718–747. https://doi.org/10.17705/1cais.04434

Sujata, J., Sohag, S., Tanu, D., Chintan, D., Shubham, P., & Sumit, G. (2015). Impact of Over the Top (OTT) Services on Telecom Service Providers. *Indian Journal of Science and Technology*, *8*(S4), 145. https://doi.org/10.17485/ijst/2015/v8is4/62238

Syed, S., Halaoui, H., Hicham, F., Mahmoud, M., & Mansour, M. (2014). Analytical imperatives for telecom marketers in emerging markets, *11*(2), 19–36.

Tan, J. (2017). Industrial IoT Technologies and Applications, *202*, 3–11. https://doi.org/10.1007/978-3-319-60753-5

Verbeke, W., Dejaeger, K., Martens, D., Hur, J., & Baesens, B. (2012). New insights into churn prediction in the telecommunication sector: A profit driven data mining approach. *European Journal of Operational Research*, *218*(1), 211–229. https://doi.org/10.1016/j.ejor.2011.09.031

Vladimir Fedak. (2018). Big Data analytics in the banking sector - Data Driven Investor - Medium. Retrieved August 28, 2019, from https://medium.com/datadriveninvestor/big-data-analytics-in-the-banking-sector-b7cb98d27ed2

Wells, W. D., & Gubar, G. (2006). Life Cycle Concept in Marketing Research. *Journal of Marketing Research*, *3*(4), 355. https://doi.org/10.2307/3149851

Will Kenton. (2019). Target Market Definition. Retrieved August 28, 2019, from https://www.investopedia.com/terms/t/target-market.asp

Xevelonakis, E. (2005). Developing retention strategies based on customer profitability in telecommunications: An empirical study. *Journal of Database Marketing & Customer Strategy Management*, *12*(3), 226–242. https://doi.org/10.1057/palgrave.dbm.3240259

Ye, L., Qiuru, C., Haixu, X., Yijun, L., & Guangping, Z. (2013). Customer Segmentation for Telecom with the k-means Clustering Method. *Information Technology Journal*, *12*(3), 409–413. https://doi.org/10.3923/itj.2013.409.413

Ziafat, H., & Shakeri, M. (2014). Using Data Mining Techniques in Customer Segmentation Hasan. *Journal of Engineering Research and Applications*, *4*(9), 1–195. https://doi.org/10.1057/9781137016386

Appendix A: Informed Consent Form

- I…… 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 understand that I can withdraw permission to use data from my interview within three months after the interview, in which case the material will be deleted.
- I have had the purpose and nature of the study explained to me in writing and I have had the opportunity to ask questions about the study.
- I understand that I will not benefit directly from participating in this research.
- I agree to my interview being audio-recorded.
- I understand that all information I provide for this study will be treated confidentially.
- I understand that in any report on the results of this research my identity will remain anonymous. This will be done by changing my name and disguising any details of my interview which may reveal my identity or the identity of people I speak about.
- I understand that disguised extracts from my interview may be quoted in the dissertation, conference presentation, published papers, etc.
- I understand that if I inform the researcher that myself or someone else is at risk of harm they may have to report this to the relevant authorities - they will discuss this with me first but may be required to report with or without my permission.
- I understand that under freedom of information legalization I am entitled to access the information I have provided at any time while it is in storage as specified above.
- I understand that I am free to contact any of the people involved in the research to seek further clarification and information. Names, degrees, affiliations and contact details of researchers (and academic supervisors when relevant).

Signature of researcher

Signature of participant

Appendix B: Interview Transcript (Code A)

Interview Code: A

Medium: Skype Video Conference

Q. Which segmentation technique is employed at your company?

I will answer this question from the perspective of two markets:

Business Market

Market segmentation techniques are targeted for two markets i.e. business market and consumer market. Currently, with my experience at Proximus (Belgian based firm) the segmentation technique is mainly focussed on the business market. In order to target the business market efficiently and diligently two segmentation techniques are being applied now. The first one is based on *company size.* In this technique we divide the companies into SME (small-medium enterprise), medium corporations and large corporations. The fundamental factor is identifying the company size is no of subscriptions, no of employees (the focus is more on the endpoint). For example, let's take the case of IKEA in Belgium. Inter IKEA which is the holding of the IKEA is considered a small, medium enterprise because of the number of employees in Belgium is relatively low. Overall, if you look at IKEA internationally, it's a large corporation. Segmentation in the business market is easier when compared to the consumer market. But segmentation in the business market does have issues as well for example in order to find the apt company size is a big issue in the above-mentioned technique. For instance, if you look at, we have both telecom business and IT business within boxers, a small telecom company can be a large IT company. For example, look at a payment company. A payment company can have, for example, one 200 employees, which is considered to be medium enterprise. But their systems are very important, and they need to be secured. So probably from an IT perspective, they're a large company. These are some of the issues that we face from this segmentation technique.

The second technique used in the business market is called *vertical segmentation.* In this technique, we divide the company on the type of business they perform. For e.g. healthcare, government, banking, insurance et cetera. This technique is easy to implement and does not require a lot of data. The fundamental premise of this technique is to evaluate whether identified segments have similar needs or requirements or are the needs differ.

Consumer Market

For the consumer market, my experience is more with KPN Mobile and Orange. Segmentation in this market was predominantly done with respect to the profiles i.*e. on lifestyle profiles*. So, for example, we differentiated people on age and profession. For instance, students, the young couples and singles, and couples without kids, families (with kids), empty nesters (parents whose children are old and have left home), elderly people without children. In short in this technique we followed the lifecycle of the customer and performed segmentation.

The next segmentation technique was based on the nature of customer i.e. introvert vs extrovert these kinds of things. This technique was always done at an experimental level because we always faced lack of relevant data that explained the nature of customers. For this technique the data was majorly obtained from contract data, loyalty programmes are another great option to acquire data. Apart from these two methods of acquiring data we also used behavioural data. We had an issue with the contract data that was collected. For example, the contracts for the subscription had to be signed by people older than 18 years of age. While our target market was youth i.e. age group between 12 to 25. Parents of kids aged between 45 and 55 signed the contract therefore with these demographics we could not perform effective segmentation.

Generally, telecom firms combine two or more types of segmentation techniques. The selection of combining the techniques depends on the goal the company wishes to reach. For example, my firm currently, we are rolling out fibre. In order to reap significant gains from this technique we are really looking at a combination of geographical segmentation along with our traditional segmentation technique (life cycle and type of customer). The reason for choosing geographical segmentation for example, because for local implementation this type of segmentation provides us fundamental knowledge on how to perform local marketing, identify types of household such as full of students, full of family et cetera.

At my firm we combined life cycle and customer value segmentation for our marketing purposes. We did the combination to identify certain target customer for our propose offering and we wanted to enhance our ability to pinpoint these customers in our database. At the same at my firm we parallelly adopted the value segmentation technique. In this technique we predicted to future revenue from a customer based on their past transactions. For example, we calculated the subsidy factor per customer on two factors i.e. past revenue and churn estimation. We calculated both these factors and evaluated the subsidy factor for each customer. Generally, a loyal customer received a low subsidy factor.

Q. Which are the customer data structures used in this technique?

For consumer markets we used age, ARPU, location, gender. For lifecycle segmentation we had a tuff time in acquiring pivotal data regarding the customer. At my firm, we always tried to satisfy a set of criteria by obtaining a customer label. However, we failed to reach that point several points. Thanks to advanced technologies such as machine learning we can reach those criteria.

Q. How granular are your target market segments?

I think at both the firms I worked it's always a multitude of millions of customers. So, a segment is 100,000 customers. This is just an approximate value. But with the rise of social media I think with social media, is it much easier to perform targeted advertising. For example, we utilised this option on Facebook to target people between the age group of 25 and 35 with an interest in Game of Thrones. A few weeks back Proximus just announced that they are going to replace advertisements on TV with segmented advertisements. We are looking to make our television platform interactive in order to acquire relevant information

Q. Are you satisfied with the results generated when this technique is used?

From my experience as being at senior positions in firms so far, the results were never as expected as we hoped it would be. But sometimes we generated important results which helped us in our target campaigns in a more efficient manner. We never had a set benchmark in terms of satisfaction and our target results always varied with the various types of campaigns we initiated. For instance, with one campaign at KPN where we were targeting customers who had a high value in terms of usage outside of their subscription (basically customers who had high usage levels outside their plan). To target these customers, we used a combination of statistics-based segmentation based on these factors i.e. usage, subscription plan, amount of money paid on top of their subscription et cetera. This was a very successful campaign using segmentation as we could easily and accurately pinpoint customers. We conducted a qualitative research methodology in this segmentation and invited these customers as peer groups to know and understand their behaviour (aspects such as their issue, why they paid more outside their subscription rather than buy a suitable subscription. Based on these results we generated we developed and implemented an outbound telemarketing campaign, and this helped us achieve high adoption and success rates. These examples are few and I feel there could be much more improvement. If I look, for example, I'm currently involved in security business where segmentation is a bit different from the standard market. The technology that is commonly used in the security market is machine learning. However, we have models in this regard to identify whether a file or whether a threat is malicious or benign. So, I would expect that machine learning and related technologies are becoming more and more handy and useful in the segmentation as well.

Q. What are the challenges encountered when this technique is used?

As stated earlier the biggest challenge in the business market was identifying the company size. For the consumer market (when we tried to segment customers based on their nature) Unavailability of sufficient data was one of the most common challenges I faced. For instance, any customer acquiring services or products needs to sign a contract. The contract can be only signed by people above the age of 18. For many of our plans that were targeted for young people, the contract was signed by their parents. Therefore, we acquired demographic data of parents, but usage was done by their children. Therefore, due to these reasons we could not employ an effective segmentation technique.

Q. What are various methods that are used by your company now in acquiring customer data?

There are various ways to acquire information regarding the customer. Majority of the data used by telecoms is internal data i.e. data that is acquired when a customer's signs a contract with the firm. We also undertook surveys to know more about customer need, wishes and desires. An interesting thing that KPN adopted was "lunch and learn lectures" for customers. In events like these KPN invited customers for a free lunch and spoke to them about their needs and desires. This was an effective initiative. External data was also used. For example, we used to buy information from an accredited company named "*Blayedos"* to identify new parents. This accredited company provided information *such as number of children in a family, the number people living in a*

house, birthdate of children, gender of the child et cetera. These types of information were not available with my firm and they partnered with such organizations to enrich their database.

Q. What do you feel on this concept of pooling customer data from various sources (data sharing)? Do you believe by adopting this concept, the effectiveness of segmentation techniques can be improved?

From your explanations and use cases presented I believe it can indeed help in great regard for market segmentation. From my past experiences, telecom companies in Europe are generally reluctant to adopt this concept like how the Chinese government is prohibiting the usage of Facebook and WhatsApp in China. Companies believe and perceive this concept is intrusive to a user/customer's private life. In the Netherlands I have always experienced my manager and my fellow colleagues raising questions on privacy and thereby affecting the adoption of this concept With GDPR being so strict after May 25, 2018, companies are fearing more and not inclining themselves to research more on this topic.

I really feel this concept will give firms 360 degrees regarding customer needs because the information is acquired from different perspectives. In the current methods, we are generally looking from 1 perspective and this has a great probability for a bias. Obtaining such bias will affect the results of segmentation and will hinder the firm in identifying high-value customers. From my expertise after working for several firms, I realise that bad quality data and data acquired from a single perspective can have detrimental effects on segmentation. For e.g. such data hinders us from identifying high valued customer and generating poor segmentation outcomes. For example, we had some high valued customers in our needs-based segmentation. We had a great probability to bias them in thinking to think that they are really looking for a subscription basically, or whether we should provide them subscription they were most of times their motives were different

Q. Do you think that privacy-preserving technologies such as Multi-Party Computation will solve the segmentation challenges that were mentioned earlier?

Now such technology can really help overcome the above-mentioned challenges with ease. As market, where telecom companies are operating, is competitive, it imperative that a good segmentation is required. I think the development of a privacy-preserving technology will help firms to develop and implement will help telecoms to acquire 360-degree profile about customers and perform effective segmentation. I am stressing on the need for an effective segmentation because a cold campaign where you don't know whether there is interest from your customer is generally, we high. So, in these lines to target specific customers will make life easy and interesting for telecoms. To give an example. We had a discussion at my firm on whether to use URL's used by the customer in our segmentation technique. We felt that this data structure will help us enrich our database with valuable information. However, we decided to not use this data structure because we felt it was too much of privacy breach.

Appendix C: Interview Transcript (Code B)

Interview Code: B

Medium: Face to Face conversation

Q. Which segmentation technique is employed at your company?
Generally, telecom companies operate in two types of market i.e. Business Market and Consumer Market. Well, segmentation in the business markets is rather straightforward. Generally, the business market is segmented on the company size. Of course, there's a difference in demand per industry or demand per company size, but relatively segmentation in the business market is easy to perform.

For consumer market there are lifecycles for segmentation. In other words, if I can explain it was about the lifecycle of the handset sold. The moment I started as a segmentation expert, it was more based on characteristics in the database of our customers So, it was like customer characteristics in terms of data usage or the behaviour patterns (behavioural data), market trends (like price levels, sim only information handset subscriptions, history of transactions). These market trends were explained by the premium brands where the latest handsets are sold and discount brands where the medium or lower level handsets are sold with subscriptions. Some of examples of premium brands are T mobile, KPN, Vodafone and discounted brands are Ben, Tele2 et cetera. This is only tariff information. So, it's all about what type of handsets and what kind of tariffs and information regarding the latest flagship handsets, like when there's a new iPhone, it's introduced by the premium brands. But if you have the iPhone SE got the older type (discounted handsets/old handsets), then it's sold with the discount brands.

During my time as a segmentation expert, every parent company had subsidiaries because as the market was saturated, it was imperative for firms to start milking and reducing fixed costs and ensure revenues are constantly maintained. The other reason related to having subsidiaries is that the market was big for instance the game was only on voice and then shifted focus to SMS and then it shifted to Data. For instance, KPN had a second brand called *Telforth*. However, half a year ago announced that they will dish or second brands and only do KPN brands. This was basically done to make sure that telecom does not have too much fixed costs, because when you have different brands in your markets, you have to do all the media spending for two brands

As I mentioned earlier, lifecycle segmentation had focus only on the handset and not on the personal lives. We used market data and behavioural data to perform this segmentation.

During my term with T mobile I did a lot of work on *Needs based segmentation*. We did a large market research along with an accredited 3rd party company. This was basically done to understand the kind of behavioural data present or the predictors that are available to apply this segmentation on the database. As you cannot do to market research for all your customers so there are a lot of predictive modelling that is done in this technique. This basic essence of doing predictive modelling was to create a link between the market research and the

behavioural data for your customer. So that becomes your bridge for you to develop a new marketing mix or create new offers. When we have the right content with the right links, we could segment on innovation drive in terms of innovators versus laggards or people, who don't care about a new handset, or don't care about a really high data bundle. We were also able to segment people based on their communication demands such as people who interact a lot vs people who don't interact a lot. We had some, like five to six or segments and we linked it to the database (the behavioural data) and then we started modelling as well because. In Netherlands many contracts are done on a yearly basis soon as you a customer reaches the end of his contract it becomes interesting, what kind of next subscription you will buy. For instance, the innovators are mainly driven by data bundles, etc. So, they will go to the premium brands of segments and the laggards or the low communicator will go to the discount brands as they have a low usage of data bundle.

As there we adopted two types of segmentations these techniques were used like a pendulum. At one point in time when I started, it was more focused on behavioural data that was already in the database and we used generic market research reports or generic data describing the trends in the market. At some point in time, we only used like really market specific segments like sim only users or headset or regular subscription users or age So in other words when I started it was just demographic segmentation then it evolved to needs based segmentation and then to premium vs discount. So, it's a mix of everything, that it's not just one and depending upon the market trends, we kept shifting. It also depended on the budget allocated to do segmentation. One important point I wish to explain is that there was a lot of overlap between the approaches in all the three types of segmentation i.e. demographic segmentation, needs based segmentation and behavioural data. This is because the drivers (influencing factors) behind this thinking are the same.

There are two perspective among segmentation experts at various telecom companies i.e. marketing perspective and data-base perspective. Customer Value segmentation is one technique that is used from a data base perspective. However, I take the view of the marketing perspective as I feel in the end, you need to communicate with customers. and you don't say, hey, you're a high value customer these days. It is imperative that you connect to the needs of the customer and then target them with the right message. For instance, most of the time, the high value customer wants the newest headset and they're willing to pay. So, their willingness to pay is high and then you have the right message and it is easy to win over the customer.

Q. Which are the customer data structures used in this technique?

Gender, Age, postal area, bill size (amount of billing amount), Average Revenue Per User (ARPU), type of handset used, usage data like no of SMS'S, amount of internet consumed, number of minutes utilised in voice. These were the basic data structures that where are commonly used in this technique. We only used internal data rather than external data as we felt that the value added by the external data was very less in compared to the investments that were needed to be done to acquire the external data and aggregate the collected internal and external data. This is the very reason why we used external data only to a limited extent

Q. How granular are your target market segments?

It is generally 15% to 20 % of our customer base. I understand that the size of these segments is high. Due to these reasons we did micro segmentation but in the latter parts of our analysis we understood that micro segmentation did not add much value to our previous message. One major problem with micro segmentation

is management. There are benefits of micro segmentation. For instance, is very important to have the right message at the right time for the customer. For example taking data usage as the dividing factor among customers, micro segmentation helps telecoms address customer with the right information on how to use it or on how can you make sure that you don't get a bill shock and help customer's identify specific instances in their life such as when you use too much data and What do you do with data in foreign countries, etc. There is another benefit when we adopt micro segmentation. Suppose the telco intends to want to align with the customer when they are in their last stages of their contract, the telco needs to target them with attractive offerings. Micro segmentation helps in the alignment with the company's portfolio. We used micro segmentation for to address and market our next best offering for e.g. we adopted personalised campaigning, where our agents marketed our offers with specific customer needs and convince them on why it is interesting to buy elaborate on their specific likes and desires etc. In order to test the effectiveness of micro segmentation we ran pilot projects and analysed how failure might be. We understood the micro segmentation requires more effort in terms of data cleansing, data analysis and data management and on later stages we found different or new angles that differed from our vision. Therefore, then we started aggregating them to manageable campaigns, etc. Taking my perspective as business management expert (perspective vary with the profession done), I feel micro segmentation can be used to learn more about the customer data base and find novel trends and patterns.

Q. Are you satisfied with the results generated when this technique is used?

In terms of the effort and time we invested, we were really satisfied with our results. We put efforts for example we trained our customer service in such a manner to align our services and offers that satisfy customer needs. We also invested on training our marketing professionals, for communication purposes et cetera we use this kind of segmentation. I feel that had we dug deeper with a lot of customer data it would have been too much to handle and manage and might messed up our implementation part. So, I think this high-level segmentation was really helpful to get a better picture of the customer, because from my experience and the trends in consumer markets mass marketing technique is highly preferred and embraced. I understand that as a segmentation expert one would always want to have a more targeted approach, but it is impossible to adopt an individual marketing technique due to the management issues.

Q. What are the challenges encountered when this technique is used?

From a personal point of view, I really liked the needs-based segmentation. However, the success of this segmentation is dependent on these factors: *time of targeting, what specific niche can be segmented on et cetera. These* factors are never static and are always dynamic. Factors influencing these factors are uncontrollable. Additionally, this technique is purely a data driven and requires a lot of effort such as conducting the market research, data cleansing, analysing the data and then identifying patterns in our database and finally implementation. All these steps involve a lot of time and effort to make sure this technique is profitable. Taking into consideration the effort one teams needs to put the biggest challenge was budget and research. In summary I feel it is more of operational challenges rather than data availability challenges.

Q. Does your company have enough customer data that can be used to perform this type of segmentation?

From my experience as a marketing analyst It's never a lack of data or the unavailability of data that was always a hindrance for the success of the segmentation technique we used. The only challenge I faced with data is the unavailability of resources related to data analysis. We always had more data than we wished for.

Q. What are various methods that are used by your company now in acquiring customer data?

We basically acquired data from customer contracts, conducted surveys and organised some interactive discussion with our customers. We never relied on third party agencies as we felt that the value added is less. Plus with the amount of data is being generated we tracked the usage data of our customers.

Q. What do you feel on this concept of pooling customer data from various sources (data sharing)? Do you believe by adopting this concept, the effectiveness of segmentation techniques can be improved?

I think when you start partnering, it's, of course, it's interesting. But when you look at like, there are three telecom operators in their lives, but there are there are a lot of different brands and overlap between a customer base of a telecom operator and a partner might be very small. There is not too much relation between or correlation between being a T Mobile customer and being a Barcelona fan. However, this concept would be interesting if relevant add on's are sold with relevant offers. We undertook many pilot projects in this area, but we had a very small success rate. Me and my time investigated the reason for this low success rate and found out that the overlap of this data with the core business bundles was very small. The effort we put in to combine the data sets was high. Therefore, these are two main reasons why we felt this concept was not much of use to us. I know and understand that there are many use cases to find. This concept can be of great value when the add ons or the additional information is close to the core business operations. For example, taking the case of KPN. Hypothetically thinking suppose KPN acquires information regarding handset insurances. Then such kind of information becomes relevant and the entire success rate will be much better than what we experienced earlier. In short, the more the additional services are away from core business operations then the success rate tends to decrease. I would like to add one more point here. Generally, many telecom firms are not clear for what they are segmentation is being done. If the aims and goals are clear, then relevant add on information will be acquired thereby making this concept successful

Q. Do you think that privacy preserving technologies such as MPC will solve the segmentation challenges that were mentioned earlier?

Yeah. I think it's looking at it from the point of privacy. Yeah, I think you want to give full control to the, to the customer in the end about her personal data. Yeah. So you need to know what kind of data is stored where etc Yes, and also can use for library it's used for for specific kind of purposes, and you have to make sure that it's organized. YES I do feel privacy preserving technologies are a valuable asset.

Appendix D: Interview Transcript (Code C)

Interview Code: C

Medium: Skype Video Conference

Q. Which segmentation technique is employed at your company?

As we said in the written documents (that was sent a few weeks ago), We have multiple segmentation techniques. All of them are based on models/ statistical models and they are user-oriented. They are named as *behavioural segmentation, value segmentation and churn prediction technique*. Churn prediction is also considered as segmentation because it comes into two or three different groups as an output but begins as a propensity model. So, me and my team use algorithms to train models and then to score our customer base. We apply those models to retail customers, not to corporate customers. The reason behind this is because the corporate customers have other complexities and don't operate as individuals, another large group of contracts, so there's no need for them to have for instance user segmentation. Also, any offers we do are directly negotiated with the person in charge of the company in with the retail customers, the consumer customers have we have B to C relationship. We keep updating our customers and the models once a month, the beginning of its month where the new we update our databases. Yes, we use all three models together. We generally have three different outputs one for value segmentation, one for churn and one for behavioural segmentation.

Q. Which are the customer data structures used in this technique?

Generally, we use data related to use of advertisement's and like average minutes of outgoing calls per month, average minutes of incoming phones times a percentage of usage during weekdays or weekends, calls to mobile to mobile destinations, or for instance, months that the commitment of the customer ends, and this is very important when we talk about churn prediction or there. We use specific variables for a specific technique. For instance, the average revenue per month, which is used for the value segmentation. And we have many other variables that we use to train the model, and then we also use it to score the customer through the algorithm. There are many more headings, but we consider it least important to discuss here. In summary, we basically understand how we try to analyse the behaviour and we score its birth our base to. We don't do something different than other companies. As a telecom company in Greece and a TV provider company, we have access to specific data. This is what we used to try and predict or analyse our customers behaviour. And really, when we're talking about the model, the variables us there might be over 50. The main reason for employing the behavioural and demographic data structures is that since they are acquired by our own systems, this gives us more confidence in building our predictive models"

Q. How granular are your target market segments?

In our past we had a really tuff time targeting customers, however, nowadays we are okay with our segment size. As I told you before we operate in a very dynamic and challenging environment and therefore, we wish to reduce our segment size.

However, for instance, when we're talking about behavioural segmentation using segmentation, we don't want to have a very niche segment or segments with less no of customers. This is because we don't have that the capacity in terms of Human Resources another and commission cost resources to target a very nice segment. So, we try to maintain the number of eight to nine segments, from the gathered several customers so that we can target them and share the same characteristics.

Regarding churn prediction which is one of the most effective ways to segment also which does not require a high-end statistics analysis, rather than imposing specific commercial rules and that can help you narrow down the number of the customers you want to target. And this perhaps helps me to give an example is, for instance, you can target to high churners with high revenue, this is what we want or and right now we have added some other data that allows us to be more effective on the propensity. So, we are in a better state right now definitely, but the thing is that started the company started quite a few years ago with of the self-models. What about that, and then early use the body Intelligence and of course, corporate intelligence with some assumptions and heuristics. The size of our customer base is 600,000 users and splitting that 8/9 segments would provide you the size of each segment. Just to have an order of magnitude of the market, the market in Greece is like this, we have for 4.5 million households and we have addressed 16% of the market. So, if you could do some maths, you can see the size of the clients we offer our services. And this is what he has is not totally accurate, but you can see the impact of this. So, this is what that segmentation team at Forth-Net uses to do this segmentation.

Q. Are you satisfied with the results generated when this technique is used?

Yes, I'm quite satisfied with the results after considering all those constraints I mentioned I think we are close to what is happening. But yes, wat I feel is there is a space for improvement. And I'm also happy because I know that we are the same level of our competitors are so when we're not falling behind our competitors because the market we operate is a constantly changing market. And there are new offers coming out, not on a daily level So we must cope with them.

Q. What are the challenges encountered when this technique is used?

Well, the main challenge that we face is with respect to CDR (Call Detail Records). We have the CDR's coming from the internet, telephony, communications and TV. They generally occupy huge data spaces. We don't have that much of data space. So, we are forced most of the times to aggregate data. I feel using aggregated data for a period is okay, especially for the telecommunication data. But that's not pretty much okay for our TV viewers. It doesn't provide all the aspects that we want to investigate. So first, this is one of the main challenges we use face and we are forced to compromise i.e. *with a certain level of aggregation.* So, in brief, we can summarize it, is the challenges are the lack of high-performance computing and high memory. So, access to such computing resources, especially high performers, probably the next era like Artificial Intelligence resources. would be very,

very useful for to advance our statistics analysis, and of course, access to external data of behavioural data. No matter if they are anonymous, but not this could give better insight and much, much more ideas to do our job.

The second challenge we face is that these data are not real-time data. For instance, today is 5th July, and I would have to wait until the 10th of July to get the CDRs, the users and the revenue data from last month's activity. This does not allow me to have for instance to employ a more effective churn prediction model or a model with a smaller window of prediction. I have faced this challenge when I was employed with other companies as well.

The third challenge relates to the lack of predictors. These predictors are really important to understand the needs of the users. This also another reason why we do aggregation of data.

Q. What are various methods that are used by your company now in acquiring customer data?

90% of our data is internal data. We fundamentally use data such as CDR's generated by our customers, the viewership data generated by the setup boxes and the demographic data that comes from the other provisioning systems and our CRM systems. We use ZBR as our CRM system. So, we employ mostly internal data but, in some instance, when we want to verify our acquired data with the market situation we use external data. Generally, the external data is acquitted from trusted, govt verified 3rd party organizations. However, we don't combine the results i.e. we don't combine internal and external data. I would like to add one more important point in this regard. The Greek market is not mature, exchanging data, especially between different market segments, such as financial data, operators, data and so forth. I don't know what's happening to more mature countries like France or the United Kingdom, but this type of data exchange is not happening in Greece. Moreover, we do not have systems that would assist in this process. So, it's time we may want to exchange data, we must think all these questions: *How much we want to exchange? What is competition? Which are the GDPR restrictions?* As these questions are hard and along with situational factors, we use internal data rather than external data.

Q. What do you feel on this concept of pooling customer data from various sources (data sharing)? Do you believe by adopting this concept, the effectiveness of segmentation techniques can be improved?

In principle, YES. As most of our data is behavioural data, we would like to adopt data sharing as it would help us answer these questions such as *What is your favourite team? Which is your spending habit? Are you going to go this weekend to see this specific match*? and questions like *are there any frequent visitors over the stadium with having come along long-range tickets.* Answers to these questions would certainly provide useful information to have to target better our clients. So far, we don't have them but, in the future, if we have access to such information, we could really utilize such answers and augment our views we have with respect to the behaviour of subscribers. Therefore, by definition, I feel it will be useful.We would also like to have some more data regarding the household. This is because we are a company where our services are not to individuals but to households. We would like to know the members of the households or their ages or if they have children or what age are, they, children? As we are focussed on providing TV packages and sports-related content and we wish to know their favourite team and some other lifestyle a lifestyle attribute. Another aspect we would like to know what type of mobile device they use. As of now, we don't have that knowledge. But if we know that they are using a high-end smartphone, we would have a better position to evaluate the initiative we are doing.

I would like to explain one use case about data sharing that me and my team could possible apply. Hypothetically we could have access to even anonymized Excel data sets, e.g. coming from a bank, with territorial knowledge and postcode rollers, then my team could somehow correlate using territorial keys to understand better the behaviour of which building block or its or its territory. So, such type of external databases, set of data could be useful, actually, to brief you both early doesn't know perhaps he knows. This is the type of data that my team could seek and could use in order to better understand the behaviour of our clients. Imagine that we would like we want to build the profile a general profile of our customers in order to be able to, to communicate with them and sell them more effectively. But and our main source of information on you know, market research, but it's difficult to attribute the outcome of the market research and link it to our database, because it's, it's generic, it can be linked to database, especially specific contract or client As I explained earlier the lack of maturity between firms from various sectors to share customer data. I feel in Greece we're not pioneers. We do have advanced with analytics, for instance, France is a lot more advanced in using data mining techniques, so I believe it's a it's a matter of time for us to find the right the bridges and construct the bridge is between us in order to advance and make use of data sharing So far, we lack some global rules that can be known to all market segments. For instance, a third-party party that could overlook such data exchange process and generate trust that one might want to change data should establish again and again, the processes to do show. So, it's not really very useful to do that. It's time. And you face quite a lot of barriers to that. So, the lack of this maturity, and the trusted third party will really increase this propensity or willingness to exchange data. The lack of maturity can also be related from the government point of view or the governmental services that they open the data. There is a movement towards this trend, we believe that we could have access to public data, public open data in a couple of years. And this could give us some additional insight on some behaviours we will see but still we are working as a country towards this data exchange in data openness, direction, we are not there yet. So, we have to cope with what have inside the company for the time being and try to search for surveys to augment our data. So, this is more or less the practice we do currently. And we hope to establish such a trust environment to change data in the near future. It needs a major step that the government also needs to take. And this could be a good paradigm for the rest of the segments and market segments to follow and to exchange that data as well, because they would find value in doing that.

Q. Do you think that privacy-preserving technologies such as MPC will solve the segmentation challenges that were mentioned earlier?

According to me, privacy is an element that everyone should consider. If you take it into account, it is not a prohibition factor. Generally, marketing experts or segmentation would never say privacy as a prohibition factor perhaps the IT guys might say no, for instance in the United States of America it’s easy to identify a person as they have a certain social security number, which is given to every service either private or public. In Greece, we have such ideas that can help us link persons between link the same person from a company to another. But it’s not as is not as unique as the USA’s social security number. We have made a lot of progress in that field. And we have I can think of right now two identifiers, I can hand pass link and track down a certain into an individual, for example, one can combine it the data from the bank, data from the telecommunications company and the data from the supermarket that has a loyalty card.

Appendix E: Interview Transcript (Code D)

Interview Code: D

Medium: Skype Video Conference Call

Q. Which segmentation technique is employed at your company?

Our segmentation techniques are very different from the advanced companies in the market. We provide the OTT solutions and a local PSTN number and don't involve in providing infrastructure and related services. For segmentation we don't use a lot of data, but we partner with local tower providers. For this we need to identify the tower provider and generate details for example which tower provider provides service in a building and then which infrastructure that they use like do they use fibre optics or wireless systems and which operator do they employ. Once we get all these details we can stitch a comfortable partnership with tower providers. In summary our segmentation technique involves in identifying the number of tower providers in Indonesia. It is a simple process we use which does not involve a lot of data. I can explain this technique of segmentation. Every city in Indonesia has a central tower where all sub tower in a city are connected to the central tower. Every tower has a base transfer station that helps with the telecommunication services.

Q. Which are the customer data structures used in this technique?

I am involved in the voice calls only and we employ a private branch exchange system in our organization. For voice calls we employ these data structures: *source number, destination number, starting time, stopping time, duration of the call, rating group for example call a person from Indonesia to India we rate the call as international call or if a person is calling locally we rate the call as local call, billing size and the last one is postal code.* For SMS, we use data structures such as source number, destination number, number of SMS'S, time of SMS's sent et cetera. We fundamentally use all the field of a typical CDR (Call Detail Record).

Q. How granular are your target market segments?

We are a small and medium enterprise that operate only with landline telephones and not mobile phones. We have a customer base of 10,000. Approximately the size of every segment could be 10 or 20 customers and one customer can contain a lot of telephone numbers she/he are using simultaneously.

Q. Are you satisfied with the results generated when is technique is used?

Not at all. With the current techniques it is really tuff to find a new customer. Approaching new customers who might be profitable has always been an issue. Many people now are using smartphones and WhatsApp for calling purposes rather than traditional telephone services.

Q. What are the challenges encountered when this technique is used?

The biggest challenge that I have faced is identification of fraudulent cases when performing data analysis. Corruption is rampant in Indonesia and identification of fraud is always a troublesome process. With regards to market segmentation as we don't rely on a lot of data I have not faced any major challenges as such.

Q. Does your company have enough customer data that can be used to perform this type of segmentation?

As I earlier said, we have a different market segmentation technique where we don't rely of data for execution of our process. To perform our technique, we have enough data that provides me and my team details regarding the customer and their needs, desires and wants.

Q. What are the various methods used by your company now in acquiring customer data?

As I told you before we employ PBX systems in our company. PBX stands for Private Branch Exchange, which is a private telephone network used within a company or organization. The users of the PBX phone system can communicate internally (within the company) and externally (with the outside world), using different communication channels like Voice over IP, ISDN or analog. A PBX also allows you to have more phones than physical phone lines (PTSN) and allows free calls between users. Additionally, it provides features like transfer calls, voicemail, call recording, interactive voice menus (IVRs) and call queues. Acquiring data is done from by generating a query request. It is an easy process. We don't use external data.

Q. What do you feel on this concept of pooling customer data from various sources (data sharing)? Do you believe by adopting this concept the effectiveness of segmentation can be improved?

Well I have never heard about this concept until today. From my expertise I feel it is a great concept. I think it will be helpful in marketing analytics i.e. we can understand the behaviour of a customer not only from internal data but also from external data. I also feel that with the availability of more data we can reduce the segment size even further and this would help us personalise our offers. From the use cases provided by you I feel such a concept is important. I am not sure if such a platform is already put in to use. However, I feel the biggest barrier to data sharing in Indonesia might be hampered by corruption. Mafia groups are so common here and even the government cannot do much about it. So, I feel implementing this concept here might take a long time than you think. Budget allocation is generally ruined, and this might affect the implementation.

Q. Do you think privacy preserving technologies such as Multi party computation will solve the segmentation challenges that were mentioned earlier?

Well I don't think privacy is an issue here at least for Indonesian market that my company operates in. From my understanding I believe that Indonesian people are more public with their activities and would not mind sharing details about their lifestyles are various other aspects. As I earlier said tackling corruption is the only barrier I feel that could hamper market segmentation.

Appendix F: Interview Transcript (Code E)

Interview Code: E

Medium: Skype Video Conference Call

Q. Which segmentation technique is employed at your company?

For the segmentation my company has employed several techniques and, but we employ 2 techniques that are presently used are: bandwidth segmentation and behavioural segmentation. In the first technique i.e., bandwidth segmentation we segment our customer by their bandwidth. So, we have bandwidths of measures 10 Mbps, 20 Mbps, 30 Mbps et cetera. We fundamentally divide our customer base into two groups, the higher segment is with the higher bandwidth of 50 Mbps and 100 Mbps and the second group is below 40 Mbps. The main difference lies with our services that we provide. For instance, for the 50 and 100 Mbps customers (group 1) we give them such special treatments like we have agents who will call these customers regularly to check if they have issues regarding their usage and so on.

We realise that the segmentation as mentioned above we employ now is very simple, so we also adopt a technique that is more customer centric (fit for the customer) not just using bandwidth. In the second technique we use their bandwidth as well as we use a customer's length of stay, behaviour via telephone/TV/internet, billing behaviour. We employ K means clustering data mining technique to get vital leads. By employing these we predict their future using these vital leads. We are still working on our second technique. We use both these techniques in combination.

Q. Which are the customer data structures used in this technique?

For the first approach, we only use the bandwidth. So, the 50 Mbps and 100 Mbps become higher segments and the rest for below than 40. We also use the ARPU column in this technique as it provides information regarding the bandwidth usage. For the second approach, we use a lot of parameters such as a few of them is telephone usage like destination of calling, duration of calling, TV usage like the name of channel viewed, duration of view, type of channels mostly viewed and Internet usage like website visited, amount of bandwidth consumed. We also use their billing data such as billing cycle, payment date.

Q. How granular are your target market segments?

For the first approach which involves segmentation using bandwidth we have two groups i.e. 50 Mbps or 100 Mbps and below 40 Mbps. With a customer base of 6 million users, the first group contains around 500 to 1000 people and the rest 5.5 million are in the second group.

With the second approach, which is using a lot of data, we plan to map the entire customer base into 6 segments based on their behaviour for instance, maybe as data spender as one segment. The size of each segment may be varied depending on their behaviour like one segment can have 1 million users and the other can have 10,000 et cetera.

I know that the size of each segment is large, and I feel if we have more data then we can have more segments with lesser size. This might lead to management issues but it's a trade-off.

Q. Are you satisfied with the results generated when this technique is used?

For the first technique, using bandwidth which is direct we failed to analyse the customer behaviour and failed to create opportunities to create compelling offers for instance we offered a mini pack movie to customer who doesn't have interest in movies. But in the second segmentation technique is executed well we could afford to overcome this issue because we segment the customer based on their needs. In this technique we predict accurately the interests of customers such as do they love internet or TV or phone. The conversion rate to upsell is high in the second segmentation technique.

Q. What are the challenges encountered when this technique is used?

The biggest challenge as I mentioned earlier is with respect to scattering of data. Due to the scatter of data we are not able to unify or aggregate data in a comprehensive manner. In order to overcome this issue, we would need to replace our current CRM system that is deployed. Our current CRM stores only limited information like billing data, customer contact information but does not store information regarding customer behaviour patterns. Customer behaviour patterns are basically stored in network division, TV division. In order to aggregate the data, we would need to replace our existing CRM system which is really next to impossible. Updating a CRM would be to involve large amounts of time and therefore, we would need to build a whole new CRM system. Building a new CRM involves great amount of time and investment. With these constraints and with the competition in the market we cannot afford to keep customer value at stake. We have been using our present CRM from the past 15 years.

I think the data is enough and we have all the data needed to predict the behaviour of the customer. Data aggregation is one of the main problems that I have faced. But the problem is that the information is scattered in various systems. For instance, telephonic data is in one system and internet system is in another system. Therefore, data aggregation is a problem.

Q. What are various methods that are used by your company now in acquiring customer data?

We acquire data from billing data and we also acquire using behaviour of customers. We only use internal data and not external data. This because we have trust issues and always suspect the veracity of the external data. In addition to this we lack technological tools that can verify the veracity of such huge volumes of data.

Q. What do you feel on this concept of data sharing? Do you believe by adopting this technique the effectiveness of the segmentation techniques can be improved?

Yeah, I think a concept is good. We earlier try to do adopt to this concept. But this pilot failed because a telecom firm is not just one company but has many subsidiaries. For instance, we have a lot a lot of subsidiaries like telecom cell, telecom sigma com sigma, info media et cetera. Specifically, if we consider telecom cell, our subsidiary is the largest cellular provider in Indonesia and a market leader with a customer base of more than

100 million. With such strong hold our telecom group planned to adopt and implement the concept of data sharing between our parent telecom company and telecom cell. The main reason was that telecom cell has the best data best in Indonesia. It said that it is even better than Google or Facebook, but in scope of Indonesia. Telecom cell has data that provides complete information to build comprehensive profiles of customers. For instance, they know where our customer connects to which BTS we know, you know, now people are tending to use smartphone, because all that and smartphone telecom cell could track because he is the provider. The main reason for this pilot to fail is that they were reluctant to share customer data due to privacy and trust issues. Telecom cell did not want to be involved even though in hindsight the benefits were really very high. Some of the examples of possible benefits are: we could know where the customer is, is there is in which city in which district because no, because the cell phone is always connected to weather is right to the tower, so we could predict where they are, we could predict their movement, which app knowing where they are at night, where they are in morning, where they are also their browsing history. In the smartphone, we could predict also what application maybe Facebook, Instagram, and how many hours they spend in those application that they have. As these types of information are private information, we did not have the right approach to do this. We feared a possible data breach or data leak. So, if the concept of data sharing, implemented, I think it will really enrich our future and help target our customer.

Q. Do you think that privacy preserving technologies such as MPC will solve the segmentation challenges that were mentioned earlier?

From my experience, I feel if there are tools or technology that overcome trust issues and ensures that data breach is next to impossible or can control data breach then I think it is a very good concept and can act as a catalyst to enhance segmentation techniques concept for segmentation.

Appendix G: Interview Transcript (Code F)

Interview Code: F

Medium: Telephonic Conversation

Q. Which segmentation technique is employed at your company?

The Indian Telecom Market is massive. Telecom regulatory authority of India (TRAI) has divided the entire market into 22 circles. These telecom circles are divided through geographical boundaries. Competition is high in the telecom market and 50% of the market is captured by the company *Reliance Jio.* With a such large number of circles, market segmentation is imperative. Demographic and Behavioural segmentation are combined and predominantly used to understand the customer behaviour in a comprehensive manner. This is because Indian consumes as much as data that China, Korea and Singapore consume together. Every circle performs its own segmentation and the customer data is updated by the last working day of the week i.e. Saturday. Behavioural segmentation I mean here is telecom behaviour and has nothing to do with the private behaviour of every customer.

Q. Which are the customer data structures used in this technique?

The common data structures that are used are: Age, Gender, City, Postal code, Profession (Student, working professional, housewife), Usage levels such as amount of data consumed, number of incoming calls, number of outgoing calls, number of SMS's sent et cetera.

Q. How granular are your target market segments?

Well as I said the Indian market is massive, I can only provide approximate numbers. Each of the 22 circles has about 100 segments of customer. The segment size would depend on the circle and geographical factors. In order to accommodate the entire population of India which is. On an average each segment size would contain around 50k-60k customers. So about 5 million customers are present in every circle.

Q. Are you satisfied with the results generated when this technique is used?

With the current market trends, I would say a No. In the past we were the market leader in India but in recent times with the influx of Reliance Jio in the market we are facing losses. Well if you compare the current market dynamics and rate at which new technologies are being built. I would say we are doing pretty fine but as a telecom operator we are always looking to improve. In short, I would say we need to improve and develop novel offers in order to compete with Reliance Jio. The other aspect of which I am not unsatisfied in the amount of time, budget and market research we spend on developing new offers. Approximately behind the scene we develop around 10,000 different plans but after many rounds of brain storming and analysing the current market trends 50-60 of these plans come into the market. This ratio must improve, and this is where I am not satisfied.

Q. What are the challenges encountered when this technique is used?

The main challenge that I think many companies in India face is the actual data itself. The acquired data is never good enough to provide us and help us understand critical patterns in the customer needs and due to this our segmentation has become flat and ineffective. I can explain this with an example of telecom data we use. In the current practises we acquire behavioural data of the customer, but this data does not provide details regarding the user's interests. All we do is try and predict the user's interests by the behaviour they exhibit in the network. Many a times I have seen our predictive models going completely wrong. These incomplete data sets are a big challenge.

Sometimes I have faced challenges such as lack of analytics resources as well, but this challenge is very small when you compare it with the first challenge. Therefore, according to me I would say unviability of high-quality data that provides us vital information regarding the customer as the biggest challenge.

The other challenge that I have faced is about obsolete data. The demographic data we have are not updated, therefore we are always out of place when there is a need to understand the demographical aspects of a customer

Q. What are various methods that are used by your company now in acquiring customer data?

I have been associated with many telecom firms in India and I have never seen any of the telecom's acquiring data from trusted 3rd parties. They only rely on Internal data. One major method is Customer acquisition forms or the CAF's. These are the forms that customers need to fill when get associated with a telecom for the first time. The form contains basic demographic details of the customer. We also undertake surveys to get a basic understanding of the customer needs. Apart from these two methods we also use usage data of the customer to understand their behaviour.

Q. What do you feel on this concept of pooling customer data from various sources (data sharing)? Do you believe by adopting this concept, the effectiveness of segmentation techniques can be improved?

The are 4 main players in the Indian telecom market. As I earlier said, Reliance Jio has eaten around 50% market share. Every reason why many customers are inclining themselves to Jio is because the operator provides free unlimited calls with 56GB of data per month for a price of 149 Rupees (2 Euros). Due to this reason the other operators have been forced to shell out unlimited plans to every customer and has this led to market segmentation losing relevance and importance. However, as an expert I feel data sharing will surely play a vital role in helping firms gain significant market share. For instance, with this concept as an expert with Airtel if I acquire information regarding a customer's life insurance plan, OTT services interests, hobbies et cetera we will be in a better position combine relevant adds on offers with our unlimited plan and therefore create a compelling marketing mix. I have another offer that my company Airtel has created in the recent past: Home, Work, Play. We plan to predict to a user's home location and his work location and his usual hangouts places. By understanding these three aspects of a customer, we plan to collaborate with Insurance agencies to provide relevant offers to the customers. For instance, let's suppose a person's work place and his house is about 1km. We can then provide a travel insurance to him at a discount rate of 50% because this person's chances of buying

an insurance would be very less due to minimal distance between his work place and house. Therefore, I feel that this concept of data sharing will help us acquire relevant information and develop an effective Home, Work, Play plan for our customers.

There is another benefit associated with this concept. I have attended many conferences regarding an interesting technology called Deep package inspection. This technology would give complete information regarding user behaviours. I can explain this by comparing the present situation and our visualised situation. Let's suppose that a person X is watching Netflix today. With the current technology all we can understand is the amount of usage this person has done but don't have insights regarding the movie or the web series he has watched. DPI plays an important in acquiring this information. However, from my understanding is that this technology will be hard to sell due to GDPR restrictions and cannot be afforded by small and medium telecom firms due to high costs Therefore, this concept of data sharing can be a like to like replacement for this technology.

Q. Do you think that privacy preserving technologies such as Multi Party Computation will solve the segmentation challenges that were mentioned earlier?

I completely understand this question. Since you guys are in Europe, with the GDPR being stricter now, the legislation is playing a huge role and affecting many operations for business for several industries. However, coming to the Indian telecom market we do not have any stringent laws in practise. The Indian constitution does not state privacy as a fundamental right. There are several rounds of discussion going among policymakers for the development of a law similar to GDPR. Therefore, I would say that privacy is not a barrier for the adoption of data sharing in India. My knowledge on technologies that preserve privacy is limited.

Appendix H: Interview Transcript (Code G)

Interview Code: G

Medium: Skype Video Conference Call

Q. Which segmentation technique is employed at your company?

As you know that I am an academic researcher and from this perspective I understand that they do use a number of segmentation techniques and adopt intensive efforts to segment their customers. Typically, I feel psychographic segmentation, geographic segmentation is predominantly used. From my level of experience, I think these two techniques employed are marginally useful as both these techniques do not provide wholistic picture of customer needs. From my academic research experience, I feel that behavioural data structures along with demographical I realise what they would need would be this usage data and from my research I learnt that they have not been doing much of this. I think usage data would enable them to make segmentation based on real user behaviour, like behavioural segmentation, and that I think that would be more valuable. I think geographical segmentation, for example, is very limited in terms of utility. This technique does provide insights for marketing purposes mainly but not something that will enable them to better direct their offerings and develop new services. The selection of segmentation technique solely lies in the vision and goals the management of a telecom firms wishes to achieve.

Q. Which are the customer data structures used in this technique?

For both these techniques data plays an important role in deciding the success rate of the technique. In geographical segmentation the typical data headings used are postal codes, type of environment such as is it a rural area, or urban area, bigger city, smaller city et cetera.

The data headings for psychographic segmentation are a bit tricky as most of the telecom companies across the globe have geographical information but lack on psychographic data. So probably, they would then collect a subset, for example, 1000 of their customers. From what I have witnessed, an external consulting firm would do a customer survey by calling these 1000 customers and would either ask questions based on some predefined psychographic segmentation tool or they might have some of their own structures. I am afraid of the data quality aspect in this technique. So I think they are really, I think there are many question marks built in this method of data collection. I have been doing surveys in my research, but I have come to the thinking that surveys, for example, are not very good data sources and does not provide you real time data. Data acquired from these surveys are unreliable. For instance, and you ask people about their behaviour, they will not remember. I mean, it's very difficult for anyone to report with accuracy om what they do, and how much of it. So, I strongly feel in order to compete with the market this is greater need for companies to acquire real time data.

Q. What are the challenges encountered when this technique is used?

As I have mentioned earlier the challenges related to quality of data acquired. The other main problem that I have seen is with respect to the lack of connectors to the customer data. Yes, I think the other main problem is that you cannot connect that data to customer data. For instance, let's say that a telecom company collects psychographic data to perform this segmentation. Due to unreliable methods of data collection there is no way that a company can connect it to the customer data. So due to this lack of connectors this does not help them to enhance the psychographic data. However, I feel for geographic segmentation, there is a lesser risk in acquiring connectors. Despite acquiring connectors, I think this would only help them to answer simple questions. So, in short, I can say that lack of connectors or lack of data along with data quality is a main issue faced by telecom firms.

Q. Are you satisfied with the results generated when this technique is used?

Given the present circumstances in the telecom industry, my answer would be a big NO. My understanding is that they would want to use segmentation tools and techniques that are more useful. This is because they are experiencing lower ARPU levels with the current techniques and they also fail to get answers to many questions, and this also does not enable them to think differently. With the high level of competition present in the market for these telecom firms, I feel that they should constantly on the search for something better. So, in short, I don't think telecom managers are satisfied with the results that generated presently.

Q. How granular are your market segments?

The granularity size is again dependent on the vision of the company. With my 10 years of academic experience I have seen firms dividing their entire customer base into 5 segments and 20 segments as well. But if u ask me an optimum number I would 7-8 segments would be an ideal number. This is because having more segments will involve greater efforts to manage and this may not be feasible for all the companies. Size of each segment is again dependent on the market share of the company. Generally top performing companies who have a greater market share have considerable number of customers belonging to a single segment. I would firms now have two big segments of 25 % t o 30 % of customer and remaining segments have a smaller number of approximating it to 10- 15% of customers. I understand that having more granular segments is always an attractive option as the firm can completely understand the needs of the customer accurately. However, I feel it becomes too complicated and I have seen many telecom operators failing to work with these huge number of segments. Its much easier when you have like, five, six segments, and then to maybe have them will be these very sizable ones, where most of your customers will be.

Q. What are the various methods that are used by your company now in acquiring customer data

As we discussed earlier, data collection is an important aspect for the success of a segmentation. I think telecom firms apply variety of methods to gain data from customers, apart from what they have in their own databases. The most important methods I know are focus groups, surveys and interviews. These methods help them achieve both qualitative and quantitative types of data collection.

Q. What do you feel on this concept of pooling customer data from various sources (data sharing)? Do you believe by adopting this concept, the effectiveness of segmentation techniques can be improved?

Well from the use case you presented I feel that yes it plays a huge role in improving the effectiveness of market segmentation. However, I do have a few concerns in this regard I'll just start by saying that I when there is data sharing between companies, and whether they have the sort of proper permission and privacy regulations, how that plays in into the figure. Because if they are going to give data on specified customers then of course it that becomes a major concern. For instance, whether the bank for example, what kind of data can they share to another to another company, and the football club? What can they share? But if the data is anonymised it's not possible to make it to link it to specific person, then is concept of privacy-controlled data sharing will be useful.

The other aspects where I feel data sharing is beneficial is that, firms can be better evaluate their current position in the market by comparing their present offers with the needs of the customers. Such an evaluation will provide clear for developing future strategies.

Q. Do you think that Multi-Party Computation will solve the segmentation challenges that were mentioned earlier?

I would like to answer this question with a cautious Yes. I think, because what we see with customers these days is that they have a common perception. The common perception is that customers these days are not privacy conscious. Typically customer feel that if they are getting value for their personal data then they are more interested in sharing their data .So for example, people might be quite happy to share their food shopping data, if they then get good, sort of targeted advertisements in return where they get advertisements for the products that they are actually interested in. So, I think that's sort of so the technology is one thing, but it is possible needs to be sort of works, it needs to be very reliable. But then the other part is that getting this customer sort of opt in, so that they have because I think they more and more. When we talk about data sharing, we should sort of bringing the customer into this equation, take into account that they are not only sort of merchandise here, but they are actors, and that we need to be mindful of telling them what kind of why is it beneficial for them to participate in this and also, then to make sure that they are getting something that that might be valuable for them. And it doesn't have to be much more expensive, but still something that they need to feel like they are gaining from this as well. So yes, I think my answer is that technology would help and then also to take into account these aspects.

Appendix I: Interview Transcript (Code H)

Interview Code: H

Medium: Face to Face Conversation

Q. Which segmentation technique is employed at your company?

Well as you know we are one of the leading companies in the Netherlands and we focus both on the business market and the consumer market as we develop numerous products for both these markets. Segmentation in the business market is really simple whereas segmentation in the consumer market is difficult. As I said earlier, we offer numerous products for our customers, we adopt a behavioural segmentation. In this segmentation we analyse the behaviour of the customers, from their usage pattern when they use our products. We do not do a detailed segmentation, but we perform a segmentation technique by aggregating the data. We also employ churn segmentation where we analyse the data we collect and estimate the churn the customer. The choice of selection is done by the management and it is completely dependent on the mission or the goals we wish to achieve.

Q. Which are the customer data structures used in this technique?

Well as I said, we only look at the aggregated level due to the GDPR restrictions. The basic data we look to perform segmentation is by analyzing the usage trends. Now to analyse these usage we employ data structures such as age, gender, type of environment (city, town, village) profession, postal code, no of people for a particular subscription, which subscription plane the customer uses (we have internal codes for various plans) name of website visited, duration of the call , voice and internet , destination and source numbers during the call. Using this data, we try to identify the churn patterns as well. So, in short, I would like to say we don't use different data structures for different segmentation technique.

Q. What are the challenges encountered when this technique is used?

I have faced two types of challenges in segmentation. Both are related to the data. The first challenge is about data aggregation. As we have in different system, we often face challenges in regard to the getting all this data into level. The other I would be lack of important identifiers for instance the success to identifying of the customers is to always ask ourselves 'why' is he doing this, why is he doing that, why not this. For example, I have customer use our data plans irregular usage patterns such one month the usage of data would be high and the other month the usage of data would be low. We lack information as to why there was this irregularity in usage. Therefore, due these lacks identifiers we try and predict their needs and many time we go wrong.

Q. Are you satisfied with the results generated when this technique is used?

This is a difficult question to answer. Looking at the quarterly figure we have different bases to estimate our success. In general, for the past 3 -4 years we are happy with our performances. But since we have a great

reputation on an international level, I would say we would love to improve and show results. This want can also be related to the industry i.e. telecom markets. It's a volatile industry so we always aim to improve and not lag behind our rivals.

Q. How granular are your target market segments?

Well I would not be in a position to tell you the exact number of each of our segment because there so many customers in our data base. I would like to add that the no of people in a segment keeps changing because of the churn technique we use. In this technique when we see that a person might churn then we put him into another segment. Overall to keep it simple for you, 8-10 segments is what we divide our entire customer base.

Q. What are various methods that are used by your company now in acquiring customer data?

We don't use external data i.e. we don't buying data, or we are selling data to anyone of the firms or collaborators. We use only internal data. This is basically use survey's, focus groups and invite customers to our offices. These are the basic methods.

Q. What do you feel on this concept of pooling customer data from various sources (data sharing)? Do you believe by adopting this concept, the effectiveness of segmentation techniques can be improved?

I have read a few articles regarding this, but we don't do it now because data we have or the data, we need to acquire has personable identifiable information. I think as telecom firm we roll out so many new offers and services on a weekly basis more often than no we always wish to update our customer data base. So, in that regard with the needs changing from time to time. I feel this concept will be of utmost help to us provided there is not trade of personal data. But as a manager I will not encourage this concept because even though you anonymize the data there is a great probability that the data sets can be used to identify people. And if this happens will be sued by the law. Therefore, in spite of the benefits I don't think this concept will be useful.

Q. Do you think that Multi-Party Computation will solve the segmentation challenges that were mentioned earlier?

Well I am a business manager and don't have expertise in these fields. But from the information given by you if two aspects that if privacy is controlled and if the trust is increased that I think this is a good technology but again the fact is I don't think it can be useful in segmentation because of the mosaic effect that I told you earlier. I think the benefits would be for different purposes rather than market segmentation.

Appendix J: Interview Transcript (Code I)

Interview Code: I

Medium: Skype Video Conference

Q. Which segmentation technique is employed at your company?

Telekom Indonesia International focuses both on B2B and B2C markets. For B2B markets our market is predominantly with the wholesale business. We perform segmentation in this market by dividing companies on the type of business they operate on. For instance, we divide companies such as consumer supplies, agriculture, healthcare et cetera.

With advancements in the usage of data, For B2C markets we predominantly employ a combination of two methodologies i.e. Behavioural Segmentation and Geographical Segmentation. We employ this combination as it aims to acquire a complete picture on the trends exhibited at different parts of Indonesia. Cultural differences among people who live across different cities and towns in Indonesia are different and this is the very reason why we employ geographical segmentation. As the utility is less, we combine it with behavioural segmentation.

Q. Which are the customer data structures used in this technique?

I would like to answer from the perspective of both the markets. This is because different types of data are used in different markets. For B2B market the common data structures that we use are type of company, location of business and size of business. For B2C market we have three kinds of data: demographical, geographical and behavioural type of data. In the first type we use age, gender and the size of household. For geographical data we use postal code, and city. In the last category we use billing data, usage data such as no of SMS's sent, no of calls made, amount of internet data consumed, duration of the call, duration of internet usage frequency in calling or using data.

Q. What are the challenges encountered when this technique is used?

The challenges are quite different for both the markets. In the B2B market as I told you earlier, we operate in different regions across the globe. The main challenge we face is with respect to time. For instance, we are basically based in our headquarters which is located in Jakarta, Indonesia and most of our customers are in different regions. I would term this challenge as a technical challenge technical for the challenge, I mean, we are selling our product for the premium quality. The other problem is also related to price vs quality aspect. This problem is faced by all the companies irrespective of which industry they are a part off. It is always difficult to achieve an optimum price with a significant quality product.

In the B2C market we again face the challenge with respect to price vs quality. However, in addition to this we face another challenge with respect to data. I earlier stated that in Indonesia due to the cultural differences present, people exhibit different kinds of behaviours at different intervals of time. We are often faced with situations where we do not have enough data to understand this varying behavioural trend. We also do not

have systems that can store huge data sets. Therefore, in simple terms I would term this as unavailability of data and out dated storage systems as the biggest challenges in this market.

Q. Are you satisfied with the results generated when this technique is used?

I have been employed by this firm from 2016. I have made progress regarding the roles I have worked as. I started as technical expert and later switched to the sales department. Until today, luckily, we always achieved our target, even per team per company or per personal target per individual target, we always exceed our targets. In other words that I'm quite satisfied with the performance of our company so far, even though there are many articles or people out there say the telco industry is happy, like decrease of the revenue, but luckily our company we can still run the revenue. However, coming to the B2C focus it is more challenging due to the rise of OTT services. Nowadays this market is inclining towards being a digitally oriented one. Due to these reasons say we are satisfied as of now but due to the volatile market conditions we seek to improve.

Q. How granular are your target market segments?

Again, the answer to this question will be like previous one. I will give perspective for both the markets.

In B2B markets we have 3 regions i.e. America and Europe, South East Asia and Asia and Middle East. We divide our customer based on these three regions. 60% of our customers belong to the America and Europe and the rest are equally divided. The remaining customers are divided equally for other two regions. Size of segment for each region depends on the number of customers present in each region. I can only provide you an approximate number and size of each segment is around 30 customers.

In the B2C market we are improving our services and our market share. Currently we have around 10 segments and the size of segments been the approximately the same unlike the B2B market. We have around 10,000 customers for each segment.

Q. What are various methods that are used by your company now in acquiring customer data?

The methods employed to collect data regarding the customers is totally different in both the markets.

In B2C markets many of colleagues call the customers and seek their responses. We also conduct surveys organise focus group workshops to understand the customer Given the rise in the usage levels of data, we also collect usage data as this provides us a solid platform to understand the customers. There are the main methods of acquiring customer data in this market.

In B2B markets as I have mentioned earlier that we focus only on whole sale business. We undertake surveys in this market. The major positive aspect in this market is that there are international conferences conducted frequently. Our representatives attend these conferences and try connecting with all the companies who participate in these conferences. By following this method, we have acquired more than 60% of our present customers.

Q. What do you feel on this concept of pooling customer data from various sources (data sharing)? Do you believe by adopting this concept, the effectiveness of segmentation techniques can be improved?

This is an interesting use case example you provided. Of course, I do believe that is concept will work in the Indonesian Market. Again, as I told you that people exhibit different behavioural trends in this country, adopting this concept will be of utmost help as we can predict the character of customers in an accurate manner. By doing this I do agree that the effectiveness in segmentation will improve by a significant amount. As you said this concept will allow us to use our resources efficiently and target the correct customer at the correct time.

Q. Do you think that privacy preserving technologies such as Multi-Party Computation will solve the segmentation challenges that were mentioned earlier?

My knowledge on privacy is quite low and I don't deal with this. Privacy in general is a big thing with this concept. However, with the regulations by government I feel anonymising the data will help in adoption of this concept.

Appendix K: Interview Transcript (Code J)

Interview Code: J

Medium: Skype Video Conference

Q. Which segmentation technique is employed at your company?

The telecom industry is now at saturation. So many players are present in the market. This has pushed the price points to come down on a drastic manner. 10 years ago, voice and data plans were around USD 30, but now you can get packs for USD 13-15. The main reason for this is that companies are trying their best in to increase the ARPU levels. We use behavioural segmentation. The main reason for employing such a segmentation is that in USA people from all nationalities live. These different nationalities exhibit different behavioural patterns. To capture these behavioural patterns, we employ behavioural segmentation. For instance, the Indians living in USA have generated the most revenue for us because of their inclination to talk. Therefore, we use this segmentation technique predominantly.

Q. Which are the data structures that are used to employ this segmentation?

Just like any other company we use basic demographic headings such as age, gender, profession and geographic headings like postal code, name of the area, city of residence. These basic headings are needed in whichever segmentation technique are these are basic identifiers where customers and who they are. Apart from this we use behavioural data points such as amount of data used, duration of the call, type of call, frequency of call and no of SMS's sent. We also check where the calling number is located to trace the type of call.

Q. How granular are your target market segments?

As a general rule of thumb, the more granular a segment can get, the better it is. But given the legislation laws on privacy we only get granular to particular tipping point. Even today I feel most of the of the telecom firms have not used analytics as it is meant to be used this can be attributed as the main reason for large customer segments. Generally, we have 5 segments we divide our entire customer base. Each of these 5 segments have around 2 micro segments. So, in total we have around 10 segments. I am not supposed to tell you the size of the segment due to internal privacy matters, but I can tell you that these are large segments and they can be granulated if we complete data.

Q. Are you satisfied with the results generated when this technique is used?

The success of a firm or the monetary gains are a result of all the actions taken by everyone in the company. Given that one team doesn't perform well our results get affected. Market segmentation is completely data driven, that is the success is dependent on the data we acquire and how the quality is. We were able to complete our data sets and due to this reason, we always faced lower levels of output. We had to do so much backroom activity to develop predictive models. But we were never accurate in it and our results were always low. The world might think we are a leading company, but the internal satisfaction of the results was never there.

Q. What are the challenges encountered when this technique is used?

The competition is really high these days and we ought to be on our toes to tackle in any situation. The biggest challenge according to me incomplete data. The data we have is incomplete and does not provide us deep insights the customer needs. When I was heading one division at my firmT one of my colleague and I decided to incorporate 5 why interrogation technique. It works like this; the segmentation expert analyses the data by asking 5 why. We thought this technique would be really useful in understanding the customer needs of the customer. But due to the incomplete data sets, we were forced to stop at the second iteration. This is a simple example to show you how much we lacked complete data.

The second challenge we faced was obsolete data. Typically, we have demographic data that comes from our customer acquisition forms. But in general a customer's demographics change from time to time . As I earlier said demographic data provides an important link to identify customers. Lack of updated data hampered our solutions and led to cold campaigns.

Q. What are various methods that are used by your company now in acquiring customer data?

Even though there are so many companies available with the help of which you can buy data, we always used internal data. We felt that this data was more reliable for our segmentation technique. Survey's, customer contracts, focus groups are the most common methods we used.

Q. What do you feel on this concept of pooling customer data from various sources (data sharing)? Do you believe by adopting this concept, the effectiveness of segmentation techniques can be improved?

Given the use case I feel the concept is good. The obsolescence rate is really high these days. For instance, a video on YouTube will be no longer valid after a few days. So, with this rate obsolescence it is important we know our customers in and out. If we acquire data from different perspectives it would be really useful. I understand that the use was more an identifying area for cross selling but to think of data sharing along from a single perspective I feel telecom companies should share data with the E commerce firms such as Facebook and Amazon. These firms are growing rapidly these days and they acquire so much data. For e.g. Facebook's business model only work when you provide consent to collect data. Therefore, I feel this concept is really good, but I think it is effective more if these firms tie up.

Q. Do you think that Multi-Party Computation will solve the segmentation challenges that were mentioned earlier?

Yes, I do understand that personal information is shared with data sharing. Europeans as such are more sensitive to private data being shared when compared to Americans and Indians. My expertise on privacy is really low but I feel this concept will be really helpful and increasing the adoption rate.