



A CONCEPTUAL FRAMEWORK OF E-COMMERCE USER'S SEGMENTATION USING BIG DATA ANALYTICS

**SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF THE
DEGREE OF MASTERS IN SOFTWARE ENGINEERING**

A MASTERS THESIS

ON

**A CONCEPTUAL FRAMEWORK OF E-COMMERCE USER'S
SEGMENTATION USING BIG DATA ANALYTICS**

**SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF THE DEGREE
OF MASTERS IN SOFTWARE ENGINEERING**

BY:
MUSA DIMA
(ID: GSR\5541\06)

Advisor:
Dr. Manoj Kumar Singh
Department of Computing, SEE&C
Adama Science & Technology University
Adama, Ethiopia.



ADAMA SCIENCE AND TECHNOLOGY UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING & COMPUTING
DEPARTMENT OF COMPUTING

June, 2015
Adama, Ethiopia

DECLARATION AND APPROVAL

1. DECLARATION

Here by I declare that this thesis is my original work and has not been submitted for award a degree in any university.

Name of the Student

Signature

Date

Musa Dima

2. Approval

A thesis presented to the school of Graduate Studies of Adama Science and Technology in partial fulfillment of the Degree of Master of Science in Software Engineering.

Approved by Advisor:

Name of the Advisor

Signature

Date

Dr. Manoj Kumar Singh

ABSTRACT

In today's competitive business environment, more and more organizations lead to move or extent their business e-commerce. Thus, there is an increasing need for organizations to build concrete e-commerce strategies in order to engage with their users. One basic step towards achieving the objectives related to e-commerce is the segmentation of e-commerce user's, based on the user's data gathered e-commerce. Since there is an onslaught of user's information collected from e-commerce sources, new techniques are required for managing and analyzing the huge amount of data; and this is where the concept of "Big Data" can play an essential role. This research sheds light on three fields: e-commerce marketing, user's segmentation, and big data analytics. The three terms are combined into one framework, which attempts to show how e-commerce marketing objectives can be supported by an effective e-commerce user's segmentation that can be implemented by techniques and tools applicable to extremely large datasets. For the creation of the framework the following steps are followed: a set of main e-commerce marketing objectives is defined; the differences among user's attributes gathered from offline and e-commerce channels are discussed and e-commerce user's segmentation categories are identified; the concept of big data is introduced and relevant techniques and tools suitable for analyzing user's segmentation categories and segmenting users effectively are described.

Keywords: - Big Data Analytics, E-commerce, User Segmentation

ACKNOWLEDGMENTS

Without the help and ingenuity of several people, this Thesis could not have been done.

Above all I would like to thank the Almighty, GOD ALLAH (SW) for giving me the strength throughout my life. Next, I am pleased to acknowledge the consistent support of my advisor **Dr. Manoj Kumar Singah**, whose valuable guidance and advice made this work to come to an end. Without his management capability and understanding of the jobs, I could not finish this project within its time limit. Thank you very much for all of your polite way of handling things. I also thank my Mother **Medina Ebu** for her love and support. My mother have always given me her encouragement and support in a non-intrusive way I appreciate deeply.

I also thank my Wife **Medina Hassen**, Lety, for her love, support and unwavering confidence in me. Her contribution in the research is greatly appreciated, but most important is her contributions in all the matters that are beyond a thesis.

Again I would like to thank ASTU ICT Software Developer Team who covered my work responsibilities in the University when I was busy with the Thesis. Again my special gratitude goes **Mr.Kassahun Gelana, Mr.Kuma Waktola** for helping me in printing my Thesis.

Finally yet importantly, I would like to express my heartfelt to **Mr.Muez Mokenin** and **Miss. Mekdes Silashi** for helping me through all those moments during my study, good and bad. I am glad to have such wonderful friends.

TABLE OF CONTENTS

DECLARATION AND APPROVAL.....	I
ABSTRACT.....	II
ACKNOWLEDGMENTS.....	III
TABLE OF CONTENTS	IV
LIST OF FIGURES.....	VII
LIST OF TABLES	VIII
ACRONYMS AND ABBREVIATIONS.....	IX
CHAPTER ONE	1
INTRODUCTION.....	1
<i>1.1 Background of the Study.....</i>	<i>1</i>
<i>1.2 Statement of the Problem.....</i>	<i>2</i>
<i>1.3 Purpose of the Study</i>	<i>3</i>
<i>1.4 The Objectives of the Study.....</i>	<i>3</i>
<i>1.5 Research Questions.....</i>	<i>4</i>
<i>1.6 Conceptual framework.....</i>	<i>4</i>
<i>1.7 Scope and Limitation of the Study</i>	<i>5</i>
<i>1.8 Significance of the Study.....</i>	<i>6</i>
<i>1.9 Organization of the paper.....</i>	<i>6</i>
<i>1.10 Problem Description.....</i>	<i>7</i>
CHAPTER TWO	8
LITERATURE REVIEW	8
<i>2. User's Segmentation</i>	<i>9</i>
2.1 E-commerce	

2.2 E-Commerce User's Segmentation	10
2.3 Big Data	16
2.4 Tools and algorithms suitable for User's Segmentation.....	19
2.4.1 Concept of Online Customer Engagement.....	19
2.5 Related work	21
CHAPTER THREE	24
METHODOLOGY	24
3.1 Initial study	24
3.2 Reviewing of previous studies.....	24
3.3 Research Design	25
3.4 Framework development.....	25
3.5 Research Strategy	26
3.6 Guidelines to Design Science Research.....	26
3.7 Design Science Research Methodology (DSRM).....	28
3.8 Research placed into DSRM	29
3.9 Steps of Methods	30
3.10 Process Deliverable Diagram (PDD).....	31
CHAPTER FOUR.....	33
FRAMEWORK OF E-COMMERCE USER'S SEGMENTATION	33
4.1 Matching E-commerce marketing goals with segmentation types.....	33
4.2 Techniques for E-commerce User's Segmentation	35
4.2.1 Increase User's Satisfaction	40
4.2.2 Increase User's Loyalty	41
4.2.3 Increase Cross-Up Sales.....	41
4.2.4 Improve Conversion Rates.....	42
CHAPTER FIVE	45
ANALYSIS & INTERPRETATION.....	45

5.1 Walk-through of the framework.....	45
5.2 Utilization of the first framework.....	56
5.3 Big Data tools	59
CHAPTER SIX	61
FINDINGS, REASERCH LIMITATION, CONCLUSIONS AND RECOMMENDATIONS	61
6.1 Findings	61
6.2 Research Limitations	64
6.3 Conclusion	65
6.4 Recommendation.....	65
6.5 Prototype GUI and Sample Code	66
6.5.1. Service Management.....	67
6.5.2 Some GUI of prototype	67
6.5.3 Sample Code	71

LIST OF FIGURES

FIGURE 1: THE V PROPERTY OF BIG DATA	17
FIGURE 2 TOOLS FOR ONLINE CUSTOMER ENGAGEMENT	20
FIGURE 3: OCEM FRAME WORK (FOTAKI ET AL.2012)	21
FIGURE 4 : FRAMEWORK OF RESEARCH APPROACH	29
FIGURE 5: PDD OF RESEARCH APPROACH (ADOPTED FROM ONLINE CUSTOMER SEGMENTATION USING BDA 2012)	32
FIGURE 6: E-COMMERCE USER’S SEGMENTATION TYPES (1ST STEP OF THE FRAME WORK)	33
FIGURE 7: TECHNIQUES FOR E-COMMERCE USER’S SEGMENTATION	35
FIGURE 8: SIZE OF E-COMMERCE USER’S SEGMENTS	52
FIGURE 9 HOME PAGE OF E-COMMERCE USER SEGMENTATION	67
FIGURE 10.USER REGISTRATION FORM	68
FIGURE 11.ONLINE AVAILABLE ITEMS	69
FIGURE 12.USER MANAGEMENT ROLE ASSIGNMENT	70
FIGURE 13.MAIN SEGMENTATION CATEGORIES	70
FIGURE 14.SEGMETATION REPORT	71

LIST OF TABLES

TABLE 1 OCEM FRAME WORK(FOTAKI ET AL.2012).....	15
TABLE 2: E-COMMERCE MARKETING OBJECTIVES WITH APPROPRIATE SEGMENTATION TYPES ...	34
TABLE 3: FRAMEWORK -TECHNIQUES FOR E-COMMERCE USER’S SEGMENTATION	40
TABLE 4: SUMMARY OF E-COMMERCE USER’S SEGMENTATION FRAMEWORK.....	44
TABLE 5:E-COMMERCE USER’S SEGMENTATION TYPES.....	47
TABLE 6:USER’S ATTRIBUTES GATHERED BY THE OCEM TOOL FOR COMPANY	48
TABLE 7:BEHAVIORAL ATTRIBUTES	49
TABLE 8: ATTITUDINAL ATTRIBUTES	50
TABLE 9: BEHAVIORAL SEGMENTATION TYPE (ADOPTED FROM E-COMMERCE USER SEGMENTATION (2012))	51

ACRONYMS AND ABBREVIATIONS

URM	User's Relation Management
KM	Knowledge Management
SRM	Supplier Relation Ship management
KPIs	Key Performance Indicators (KPIs)
OCEM	E-commerce User's Engagement Management
E-Commerce	Electronic Commerce
OCE	E-commerce User's Engagement
CLV	User's Life Value
CIV	User's Influence Value
CKV	User's Knowledge Value
CRV	User's Referral value
MLS	Multilevel Secure
HDFS	Hadoop distributed file system
DSR	Design Science Research
DSRM	Design Science Research Methodology
PDD	Process Deliverable Program
OC	E-commerce User's
BDA	Big Data Analytics
OCST	E-commerce User's Segmentation type
TOCS	Technique for E-commerce User's Segmentation
IT	Information Technology

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND OF THE STUDY

With the growing importance of the internet and other computer networks, companies have paid a great deal of attention to the User's side of their businesses. This has manifested itself in areas such as User's relationship management (URM), knowledge management (KM), and in the move from call centers to contact centers. Now many companies are applying this same emphasis on relationships to the supplier side. This has led to a new solutions area: Supplier Relationship Management (SRM). Effective governance of big data analytics facilitates quality outcomes that create economic, research and social value while still preserving the ability for individuals to define themselves, where appropriate, and avoid predestination. A common ethical framework provides the necessary foundation for organizations to build an accountable governance system for big data analytics.

A vital process for e-commerce is user's segmentation. User's segmentation constitutes the process of dividing users into distinct and homogeneous groups and is considered an effective method for managing different users with different preferences, while developing diverse marketing strategies (Chen et al. 2007, Tsitptis & Chorianopoulos, 2009). E-commerce users can be segmented according to their characteristics, which are tracked e-commerce, with the use of specific techniques and algorithms. Since the amount of user's data that are gathered e-commerce grows rapidly, the use of "Big Data" tools and techniques that are able to handle and analyses a huge amount of data on real time has become inevitable. Software tools for e-commerce and user's engagement that already exist in the market can integrate and utilize approaches in order to facilitate and enhance e-commerce user's' segmentation, as well as other

functions they might perform providing a holistic User's view. Therefore, the impact of big data techniques and technologies on e-commerce user's' segmentation needs to be explored. This research sheds light on three fields: e-commerce user's segmentation and big data analytics.

These three terms are combined into one framework that attempts to show how e-commerce marketing objectives can be supported by an effective e-commerce user's' segmentation, which can be implemented by techniques and tools applicable to large datasets. In the context of this research the terms "e-commerce user's" and "e-commerce visitor" both refer to the visitor of a website regardless of their purchasing behavior.

1.2 STATEMENT OF THE PROBLEM

Nowadays the Internet represents a big space where great amounts of information are added every day. In today's fiercely competitive business environment, organizations struggle to improve user's experience, achieve user's retention, and expand their user's database. The increasing need to keep the users satisfied and treating them as individuals by offering right products and services at the right time, has triggered interest in user's engagement. The escalation of the internet technologies as lead organizations to pay more attention on building their e-commerce strategy and focus on engaging with their user's. E-commerce engagement is defined by Molen & Wilson (2010) as "a cognitive and affective commitment to an active relationship with the brand as personified by the website or other computer-mediated entities designed to communicate the brand value".

In their effort to leapfrog the competition, many organizations have already started employing new big data analytics software tools. Such tools are capable of assisting the whole process and facilitate the improvement of e-commerce campaigns and the engagement of user's with the brand. However, objectives and Key Performance Indicators (KPIs) for e-commerce are still, for many organizations, not well defined and not yet aligned with the business strategy (Chaffey et al., 2009). Therefore, marketers still have to put much effort in defining their digital strategy, in

order to select the appropriate software tools and effectively utilize them to achieve their business goals.

1.3 PURPOSE OF THE STUDY

There were some research findings concerning the problems of e-commerce user's in different countries. That finding is various and depends on the types of business and country's economic, social and developmental situations. In their recommendation, there were some researchers who recommend big data analytics. However there is no any research findings that demonstrate the implementation of big data analytics for segmentation system for e-commerce services.

Thus, the purpose of this study is to investigate the e-commerce user's big data problem and to design user's segmentation using big data analytics. In this study attempt will be made to develop prototype for the implementation of e-commerce user's segmentation using big data analytics. This will help the organization to recommend and assist the applicability of user's segmentation using big data analytics system for the e-commerce user's services. Moreover I intended to contribute through this study findings; the implementation of user's segmentation using big data analytics.

1.4 THE OBJECTIVES OF THE STUDY

The general and specific objectives of this study are presented as follows.

1.4.1 General Objective: The major purpose of this study is to propose the framework of big data for e-commerce user's segmentation using big data analytics; in which asset utilization can be improved, reduced customer dissatisfaction and to enhance the financial management and improve collaboration.

1.4.2 Specific objective

In order to achieve the general objective, the following specific objectives are addressed.

- ❖ Review of previous researches on e-commerce user segmentation
- ❖ To assess the current status of big data for e-commerce user segmentation.
- ❖ To identify major problems and their possible causes of big data for e-commerce user's segmentation.
- ❖ To explore the survey of big data for e-commerce user's segmentation.
- ❖ To develop the prototype for big data for e-commerce user's segmentation.
- ❖ Making conclusion and indicating the way forward for future works.

1.5 RESEARCH QUESTIONS

In this study, to attain the objectives efficiently attempt will be made to answer the following basic research questions.

- ❖ What is the current status of big data analytics in e-commerce user's segmentation?
- ❖ What are the major problems and their possible causes of e-commerce user's?
- ❖ Which are the e-commerce user's segmentation types that can assist each of the business objectives regarding e-commerce marketing?
- ❖ Which big data approaches, tools & techniques can be used for each e-commerce user's segmentation?

1.6 CONCEPTUAL FRAMEWORK

The main contribution of this thesis is to give a method of answering current e-commerce problem and transferring the organization to e-commerce user segmentation. To perform this move we present a literature study regarding e-commerce, user segmentation and big data environments and services. Using literatures we gathered a lot of Information about e-

commerce user segmentation, applications, service available, and big data analytics. And for identifying organizational requirement we investigated the current e-commerce customer segmentation usage to find problems and opportunities. By mapping the requirement identified with the available e-commerce user segmentation solution we propose the framework for e-commerce user segmentation.

The framework has two phases. The first phase of the framework, which relates the e-commerce user's objectives with the appropriate segmentation types, was evaluated during the literature survey with the marketers and the users. For this purpose, the framework was presented to the; advisor and survey of related work were addressed regarding the following during literature:

- ❖ The relevance of the presented objectives to e-commerce user's segmentation.
- ❖ The relevance of the presented e-commerce user's segmentation categories.
- ❖ The usefulness of each segmentation category for each of the objectives
- ❖ The possibility of utilization of such a framework.
- ❖ The problems and the obstacles that would appear in utilizing of such a framework.

1.7 SCOPE AND LIMITATION OF THE STUDY

The study focuses on using big data analytics tools & techniques to developing a model that is capable to explore the segmentation of e-commerce user's.

Due to the time available to complete the study with regard to the limitations; particularly in the data collection process resource and related work may not exist sufficiently. This may influence the quality of the study findings.

In order to overcome such limitations, I attempted to Survey journals, reports and current issue to maximize confidentiality of findings. Regarding data required for the study and certain data that needs the observation of the researcher through related work and report minimized the possibility of limitations.

1.8 SIGNIFICANCE OF THE STUDY

The researcher assumes the findings of this study have the following significances. The result of the analysis made on current e-commerce user's segmentations created awareness to the service provider and the users. This will help them to find alternative solutions to in order to strengthen the services system of the service they provided through e-commerce.

In this study e-commerce user's segmentation using big data analytics was recommended and developed in order to be used in organization. This will help the Organization and user's to maximize security level, increase e-commerce user's satisfaction, and enhance financial management in both cases. Moreover it helps us evidence for other researcher the implementation of using big data analytics for e-commerce user's segmentation and as an impute for them in conducting farther study's in the topic.

1.9 ORGANIZATION OF THE PAPER

In this study the paper organized in six chapters.

The first chapter is the introductory part of the study; which includes back ground of the study, statement of the problem, purpose of the study, objective of the study, research specific questions, scope and limitation of the study, and significance of the study. Chapter two contains review of literature and related works. The third chapter continues with research methodology of the study including research method .The fourth chapter is framework of e-commerce user segmentation, with the literature review and adds topics related to field of study. The fifth chapter is analysis and interpretation. The sixth chapter discusses about conclusion and recommendation of the study. It also shows the implementation phase of prototype including sample codes, and user interface screen shots and close the research. At the end of the document there is a bibliography section followed by appendices.

1.10 PROBLEM DESCRIPTION

Many organizations are collecting, storing, and analyzing massive amounts of data. This data is commonly referred to as “big data” because of its volume, the velocity with which it arrives, and the variety of forms it takes. Big Data is creating a new generation of decision support data management. For the purpose of decision making in an organizations, the need of processing and analyses of large volume of data is increasing. The various operations are used for the data processing that includes the culling, tagging, highlighting, searching, indexing etc. Data is generated from the many sources in the form of structured as well as unstructured form [20]. Big data sizes vary from a few dozen terabytes to many petabytes of data. The processing and analysis of large amount of data or producing the valuable information is the challenging task. As the Big data is the latest technology that can be beneficial for the business organizations, so it is necessary that various issues and challenges associated with this technology should bring out into light. The two main problems regarding big data are the storage capacity and the processing of the data. By this research, attempt made to solve this problems and solution given in coming chapters.

CHAPTER TWO

LITERATURE REVIEW

This chapter present review of literature and related work on e-commerce user segmentation. Different points that are related to the current work are discussed. Extensive literature review is conducted on e-commerce user segmentation in order to obtain an in-depth understanding of the area and to find useful approaches for the e-commerce user segmentation.

Theories in the area of e-commerce user segmentation, and big data to understand related issues in the study area is presented. Later be used in the analysis so that legitimate conclusions can be presented. In order to answer the research questions that this research poses a literature review was conducted, regarding the following domains: e-commerce marketing, User's segmentation, and big data. The literature review was based on academic journals, books, thesis projects, analyst reports, and white papers. Sciencedirect.com, omega library and Google scholar, as well as recommendations from the supervisors of this research were the basic source of literature selection. In order to ensure that the information of this research is up-to-date, the researcher focused on selecting relatively new studies. However, valuable content from older studies was also used, when considered necessary.

Thus, theories in the area of e-commerce user's segmentation and big data to understand related issues in the study area is presented. This chapter will stand as ground for the formulation of the framework decision to choose e-commerce services.

The chapter also includes about user's segmentation, brief introduction of big data has been provided followed by the service types, tools used specially associated with e-commerce segmentation using big data analytics.

2. USER'S SEGMENTATION

The initial point of satisfactory user's segmentation is the determination of business objectives and the relationship with the Users (Tsiptsis & Chorianopoulos, 2009). Correspondingly, before proceeding to segmentation of e-commerce user's, it is essential that e-commerce marketers define a clear e-commerce strategy and all the business tactics and processes involved in it. This chapter introduces the concepts of e-commerce user's segmentation.

2.1 E-COMMERCE

The word commerce is the basic concept for electronic commerce, pertaining to buying and selling of goods while 'commercial' denotes business practice and activities intended to make profits. Electronic commerce, like any other business, deals with the exchange of money for soft or hard goods and services. Electronic Commerce (e-commerce) is electronic business. It's using the power of computers, the Internet and shared software to send and receive product specifications and drawings; bids, purchase orders and invoices; and any other type of data that needs to be communicated to Users, suppliers, employees or the public. E-commerce is the new, profitable way to conduct business which goes beyond the simple movement of information and expands electronic transactions from point-of-sale requirements, determination and production scheduling, right through to invoicing, payment and receipt.

According to Vladimir Zwass,(2011) 'Electronic commerce is sharing business information, maintaining business relationships and conducting business transactions by means of telecommunications networks'.

With the advent of the Internet, the term e-commerce began to include:

- ❖ Electronic trading of physical goods and of intangibles such as information.
- ❖ All the steps involved in trade, such as on-line marketing, ordering payment and support for delivery.

2.2 E-COMMERCE USER'S SEGMENTATION

In the following paragraphs the term “user’s segmentation” is explained and User’s segmentation types based on attributes gathered from different sources are presented and analyzed. The main segmentation types discussed are the following: Attitudinal, Behavioral, Demographics, Loyalty-based and Value based. Moreover, other overlapping segmentation types found in the literature are presented. Additionally, data found during the literature review regarding e-commerce User’s segmentation are presented.

2.2.1. What is User’s Segmentation?

User’s segmentation is the process of categorizing Users with heterogeneous characteristics into distinct, homogeneous groups based on common attributes (Hong and Kim, 2012). User’s segmentation is vital for e-commerce and User’s engagement constitutes a part of CRM. It is considered as an effective method for developing differentiated marketing strategies based on user’s characteristics. As stated by Chen et al. (2007), effective e-commerce user’s segmentation contributes in raising not only user’s satisfaction, but also the expected profits of the organization. Moreover, it assists in maximizing user’s value, promotes user’s loyalty, and facilitates the CRM processes; user’s acquisition, retention and development hence, user’s segmentation is a field of considerable interest for the researchers.

It is often the case that the definition of user’s segmentation is considered the same as that of market segmentation. However, there is a difference between the two concepts that should be underlined. As defined by Stroud (2006), market segmentation “is the process of identifying Users who comprise a homogeneous group of consumers for a specific range of goods and services”. On the other hand user’s segmentation takes place within a certain and defined market; it is the grouping of the Users within a market, which has already been defined.

E-commerce user's segmentation is performed based on data, according to which user's can be grouped. Such data can be collected from various sources, such as organizations' databases or invoices, and are most of the times unstructured and continuous. Thus, they should be analyzed and utilized to perform effective user's segmentation. There are few empirical researches available, proposing algorithms and approaches for User's so called user's segmentation. For very large databases, the application of big data tools and techniques would be advisable.

The main goal of a marketer that performs user's segmentation is to develop effective marketing strategies in order to achieve various marketing objectives. More specific, e-commerce marketers are interested in knowing their user's better and being engaged with them, by implementing suitable marketing strategies. Therefore, the wise and effective segmentation of their e-commerce user is very important for achieving their e-commerce targets and further engage with their users.

2.2.2. User's Segmentation Ttypes

There is not any one-size-fits-all segmentation scheme that is able to assist effectively all the business needs and goals. There are several segmentation approaches available, each of them suitable for supporting different business needs (Tsiptsis & Chorianopoulos. 2010). Stemming from the results of the literature review on user's segmentation, a set of segmentation types is presented and explained below. Since there is not enough literature on e-commerce user segmentation, these segmentation types are basically based on data that can be collected from different sources. Each type regards specific criteria and user's attributes according to which distinct user segments can be identifies. As it is explained below, some segmentation categories overlap each other. This means that either some of the authors refer to the same set of user's attributes- that form a certain segmentation type- with a different naming, or that they describe sub-sets of segmentation types that are included in a wider segmentation category.

Main Segmentation Categories

The most well-known and are commonly found the segmentation types described below, are:

Value-based segmentation

Value-based is a segmentation process, through which e-commerce user's are categorized according to their value (Chan 2008; Kim et al. 2006; Hosseini&Tarokh, 2011; Tsipitis & Chorianopoulos, 2009). User's value has been defined by many researchers and is also known as user's equity or user's profitability (Hosseini & Tarokh, 2011). Moreover, there are several models for estimating user's value. In order to perform value-based segmentation, a certain procedure for determining the value of the user's should be developed. The term "value" does not necessarily entails only the monetary profits derived from the User's. Kumar et al. (2010) proposed 4 value components that constitute the user's's engagement value: User's Life Time Value (CLV) determined by User's's purchase behavior; User's Referral value (CRV), that related to the acquisition of new user's; User's Influencer Value (CIV) regarding user's' behavior to influence other user's; User's Knowledge Value (CKV) regarding the value added to the firm by receiving user's' feedback.

Behavioral segmentation

Behavioral Segmentation is a process, through which Users are grouped according to usage, attitude and behavior regarding a product or promotion (Stroud, 2006). In a traditional CRM system that integrates data only from offline channels, these User's data can be retrieved from firms' data warehouse (Tsipitis & Chorianopoulos, 2009). User's attributes that can be collected for behavioral segmentation contain: product ownership, type of transactions and frequency of transactions, revenue history, payments, and product utilization.

Loyalty or Engagement Segmentation

Loyalty or Engagement Segmentation is used to determine different groupings of Users according to different degrees or loyalty to supplier or brand (Stroud, 2006). The segments for loyalty or engagement segmentation are often created by applying simple business rules and/or cluster models on survey or database information (Tsiptsis & Chorianopoulos, 2009). Such segmentation contributes into identifying the less loyal Users, so as retention actions can be implemented in order to raise Users' loyalty. Attributes related to User's loyalty that can be: Engagement/Loyalty Score, Frequency of purchases, Number of Complains, new or old User's.

Socio-demographics and Life Stage

Through Socio-demographics and Life Stage segmentation Users are grouped according to their demographic or social characteristics. This segmentation type is one of the most commonly used, as the characteristics it includes can influence the changes in User's' needs, attitudes preferences and usage behaviors. At this point, it should be underlined though, that it is quite common that Users with the similar demographic and social characteristics have different needs and preferences (Tsiptsis & Chorianopoulos, 2009). User's attributes that can be tracked and used for socio-demographic segmentation are: Age, gender, income, ethnicity, marital status, education other personal details of the consumer.

Needs or attitudinal-based

Needs or attitudinal based segmentation is performed in order to explore Users' needs that can be fulfilled by the purchase of a product or service, wants, views, attitudes, and preferences (Tsiptsis & Chorianopoulos, 2009). This segmentation type can be useful for new product development, since clients preferences that are revealed can be used as an input for the design of new communication of the brand image and the key features of products or services. In an offline

environment the data needed for this kind of segmentation can be basically gathered from external sources like market surveys, through which Users are able to express their opinions and preferences.

The results of the literature review regarding User's segmentation revealed that while there is theory on User's segmentation there is not enough literature that sheds light on the segmentation of e-commerce user's and visitors. Although, one would think that there are not any significant differences among e-commerce users, the augmenting interest in segmenting e-commerce user's indicates the opposite.

An interesting insight on how e-commerce user's segmentation is related with e-commerce marketing goals is given by Peterson, highlighting the importance of an effective segmentation of e-commerce users (2006). According to Peterson, a meaningful segmentation of the e-commerce user's, which will distill user's behaviors and create focused groups that will enhance the decision making based on the data. Such attributes are cost per conversion, cost per visit, average order value, average revenue per visitor, number of visits, new visitor, returning visitor, subscription, click depth, time spent on each visit. Some of those attributes can be straight captured and added to visitors' profiles, while others, such as average order value, require calculations.

The available literature on e-commerce user's segmentation focuses more on e-commerce buyers behavior in retailers website, and do not provide an overview of the user's data that can be tracked from website visits. Therefore, in order to further explore what characteristics can be gathered in order to create informative e-commerce user's profiles an explorative case study on the OCEM Tool was conducted and is presented in table 1

Table 1 OCEM frame work (Fotaki et al.2012)

Segmentation type	Overlapping Segmentation type	Relevant User's Attribute	Related Studies
Attitudinal	Benefit	Preferences, Interests, Usage occasion, User's needs, Motivations, Usage situation, lifestyles, personality, necessities, favorite channels, payment methods, preferable promotions, type of communication	Miguéis et al..2012; Stroud ,2006; Tsiptsis & Chorianopoulos, 2009,
	Interaction		
	Occasion		
	Psychograph		
	Usage		
Behavioral	FRM	Frequency of transactions, revenue history, payments, product ownership, product utilization, last time of Purchase, Frequency of purchase	Tsiptsis & Chorianopoulos, 2009, Stroud 2006, Leung 2009, Chou & Wu, 2010; Chan, 2005; Chan, 2008;
Demographic	Geographic Lifecycle	Age, gender, income, ethnicity, marital status, education, country region Marital status, age, employment, number of children	Leung 2009, Stroud ,2006; Tsiptsis & Chorianopoulos, 2009,
Value Based	RFM Propensity scores	User's Value, Monetary Value, Propensity scores for profitability	Chan 2008; Kim et al. 2006; Hosseini & Tarokh, 2011; Tsiptsis & Chorianopoulos, 2009, Kumar et al, 2010.

2.3 BIG DATA

Inevitably, the boom in technology and informational technology has had an impact on the business world of the data, influencing the relationship of the organizations with their User's'. As a matter of fact, organizations nowadays are provided with the capability to generate huge amounts of User's data, which reveal their characteristics, their needs, their attitudes and their preferences. Hence, there is an increasing need for using new approaches and tools, and implement new strategies in order to handle and utilize the unstructured data and transform them into useful and actionable insights.

2.3.1. Big Data Definition

While organizations struggle to attain as many information as possible, the notion of big data is currently becoming a trend and it is a topic of discussion on many scientific forums. Moreover, there are already few techniques and software tools available in the market, able to analyze such a big amount of data. Since, the term itself is quite new currently there are not many scientific researches available on the field.

A number of researchers and analysts have attempted to define "Big Data". However, there is not a standard scientific definition of the term yet.

According to Chen et al. (2012), "big data is - data sets and techniques in applications that are so large and complex that they require advanced and unique data storage, management, and analysis and visualization technologies."

Another definition provided by Forrester (2012) refers to big data as "techniques and technologies that make data handling at extreme scale affordable". However, most analysts and researchers that attempt to define big data focus on the most significant attributes or else the so-called V's of big data, which actually characterize the nature of the data (Gartner1, 2011; Russom, 2011; Forrester, 2012; O'Reilly, 2012; Sathi, 2012).

An example of such a definition is the one given by Gartner¹, which describes big data as “high volume, high velocity, and/or high variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making”. Figure 1 (a & b) illustrates the three most commonly mentioned attributes of big data; Volume, Velocity and Variety.

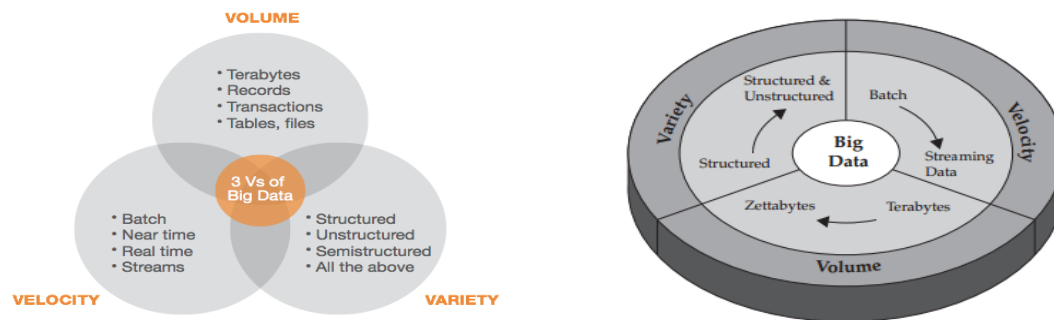


Figure 1: the V property of Big Data

Volume

Volume refers to the amount of data that are created and can be processed by big data techniques. The volume of data which are created is rapidly growing, since today petabytes of data are created in daily basis (Sathi, 2012) Handling the huge volume of data constitutes a challenge for IT since it increases the need for scalable and dynamic storage and new approaches to querying (O'Reilly, 2012)

Velocity

Velocity regards the frequency at which data are generated as well as the latency of the data (O'Reilly, 2012; Zaslavsky, Perera & Georgakopoulos, 2012; Russom, 2012). The augmenting rate of data generation requires massive parallel process and bigger space in order these data to be analyzed (Sathi, 2012).

Variety

Variety refers to the different types and sources of data. In big data systems, data are generated from various sources and can be found in diverse types. Such data might be images or text from social networks or mobile devices, web logs, streamed video or audio.

Based on the above description of Big Data, in this research we adopted the following definition of big data as:

“Data characterized by high-volume, high-velocity, and high-variety, which require advanced and exclusive tools and techniques for information management, processing and storage, analysis and visualization”.

While Big Data are becoming more and more popular, its impact on several aspects of business life and scientific fields has started being obvious. According to Chen et al. (2012) some of the main fields of applications of Big Data analytics are the following:

E-commerce and Market Intelligence: Due to the explosion of e-commerce business and the great use of social media, the amount of information that needs to be handling and analyzed has necessitated the use of big data tools and technologies. Big data can assist e-commerce marketing by applying techniques for highly targeted searches ,personalized product recommendations, text analysis etc. (Chen et al., 2012).

2.3.2 Big Data Techniques &Tools

As the volume of data increases rapidly, so does the need for using suitable algorithms, techniques or tools capable of analyzing this data and turn them into useful and usable insights for the organizations. Hence, many algorithms, techniques and tools that are able to utilize, analyze and visualize big data have been developed and already adopted by organizations. Such tools and techniques derive from several scientific fields, such as statistics, computer science, economics etc.. Those tools and techniques were afterwards adapted and further developed so as to be effectively applied in very large databases which contain wide variety of data (Manyika et al., 2011).

Although there is not enough scientific literature regarding techniques & tools that can be applied on big data there are several reports and some scientific papers that enlist and present big

data techniques and tools that are currently available.

2.4 TOOLS AND ALGORITHMS SUITABLE FOR USER'S SEGMENTATION

Since the volume of user's data is rapidly growing the issue and user's segments are at a continuous motion, the issue of how effective user's segmentation can be achieved out of the huge amount of data has been raised. However, after the literature research conducted on big data analytics field for the purpose of this research, it is apparent that there is not enough scientific. In order to understand, which techniques are most appropriate for User's segmentation a focused literature review was conducted.

The purpose was to find relevant studies that present and analyze algorithms and techniques appropriate for certain segmentation types. The literature review resulted in 12 studies on User's segmentation, as, 10 are journal articles, one study is a published master thesis research on User's segmentation, and one study is a book. All of the studies present methods for certain User's segmentation types. Moreover, all of the studies contain case studies or implementations of the proposed algorithms, in order to validate and evaluate their result. It should be underline that the book of Tsiptsis & Chorianopoulos (2009) contained several algorithms some of which were also implemented through case studies.

2.4.1 CONCEPT OF ONLINE CUSTOMER ENGAGEMENT

Customer Engagement is defined as "the behavioral manifestation from a customer towards a brand or a firm, which goes beyond purchase behavior" (Van Doorn et al., 2010). Many researchers support that online customer engagement constitutes a very important notion for Online Marketing, while others do not pay much attention on it (Moley & Wilson, 2010). Online Customer engagement according to Brodie et al. (2011) is an estimation of the degree of visitors' interaction on the site, which is measured against a clearly defined set of goals. This means that online customer engagement indicates customers' behavior towards a brand and reveals the level that the customer is engaged with a brand online. Hence, goals of online marketers are related with the achievement of customer engagement.

There are few tools in the market that are intended for OCE. These tools are different from traditional CRM tools, since they take into account customer behavior from multiple online channels creating a holistic view of the customer. Such tools are Web analytics tools, Social Media monitoring software, Web Content Management systems, Audience Targeting tools and Online Customer Engagement Management tools, such as the OCEM Tool (Fotaki et al., 2012). A summary of the characteristics of such tools is depicted in Figure2.1 below.

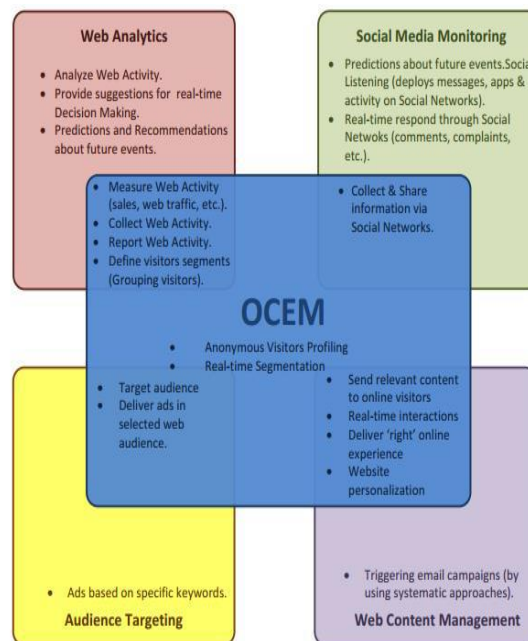


Figure 2 Tools for Online Customer Engagement

A framework for Online Customer Engagement is the OCEM (Online Customer Engagement Management) framework, which shows the steps to achieve online customer engagement and the functional requirements of a tool for OCEM (Fotaki et al., 2012). As it is illustrated in Figure 2, the OCEM framework consists of three ongoing phases, which result in Online Customer Engagement Optimization: Identification, Evaluation and Reaction. The basic tasks entail the identification of customers' behavior, the analysis of customer characteristics and the development of targeted campaigns according to visitors' online behavior. The software tools mentioned above are able to support, at least some –if not all- of the online customer engagement

phases.

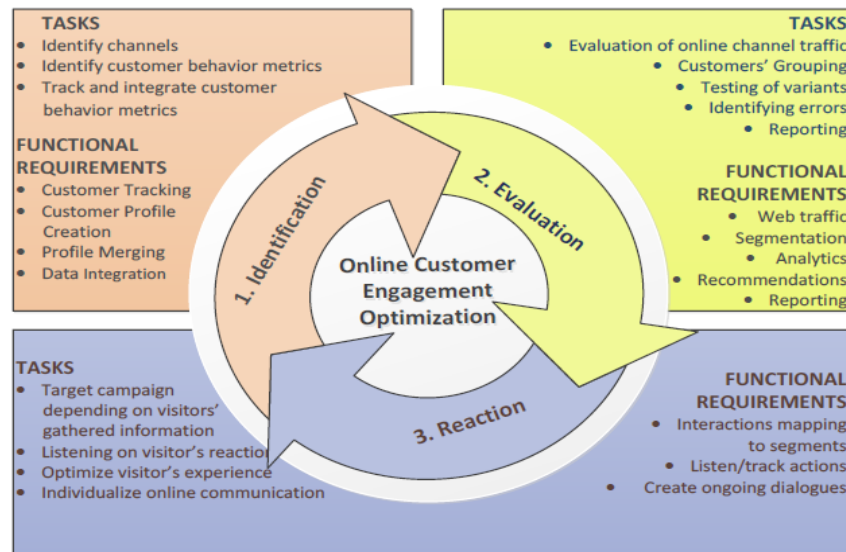


Figure 3: OCEM frame work (Fotaki et al.2012)

2.5 RELATED WORK

The researcher's review the following list of papers to best of researcher's Knowledge. In order to answer the research questions that this research poses a literature review was conducted, regarding the following domains: e-commerce user's segmentation, and Big Data.

- **John A. Keane [2] in 2013** proposed a framework in which big data applications can be developed. The framework consist of three stages (multiple data sources, data analysis and modeling, data Organization and interpretation) and seven layers (visualization/presentation layer, Service/query/access layer, modeling/ statistical layer, processing layer, system layer, data layer/multi model) to divide big data application into blocks. The main motive of this paper is to Manage and architect a massive amount of big data applications. The advantage of this paper is big data handles heterogeneous data and data sources in timely to get high performance and Framework Bridge the gap with business needs and technical realities. The disadvantage of this paper is too

difficult to integrate existing data and systems.

- **Xin Luna Dong [5] in 2013** explained challenges of big data integration (schema mapping, record linkage and data fusion). These challenges are explained by using examples and techniques for data integration in addressing the new challenges raised by big data, includes volume and number of sources, velocity, variety and veracity. The advantage of this paper is identifying the data source problems to integrate existing data and systems. The disadvantage of this paper is big data integration such as integrating data from markets, integrating crowd sourcing data, providing an exploration tool for data sources.
- **Jun Wang [17] in 2013 proposed the Data-g Grouping-Aware (DRAW) data placement** scheme to improve the problems like performance, efficiency, execution and latency. It could cluster many grouped data into a small number of nodes as compared to map reduce/hadoop. The three main Phases of DRAW defined in this paper are: cluster the data-grouping matrix, learning data grouping information from system logs and recognizing the grouping data. The advantage of the paper is improve the throughput up to 59.8%, reduce the execution time up to 41.7% and improve the overall performance by 36.4% over the Hadoop/map reduce.
- **Yaxiong Zhao [7] in 2014** proposed data aware caching framework that made minimum change to the original map reduce programming model to increment processing for big data applications using the map reduce model. It is a protocol, data aware cache description scheme and architecture. The advantage of this paper is, it improves the completion time of map reduce jobs.
- **Jian Tan [6] in 2013** author talks about the theoretical assumptions, that improves the performance of Hadoop/map reduce and purposed the optimal reduce task assignment schemes that minimize the fetching cost per job and performs the both simulation and real system deployment with experimental evolution. The advantage of this paper is improves the performance of large scale Hadoop clusters. The disadvantage of this paper is environmental factors such as network topologies

effect on a reduce task in map reduce Clusters.

- **Thuy D. Nguyen [4] (2013)** author solve the multilevel secure (MLS) environmental problems of Hadoop by using security enhanced Linux (SE Linux) protocol. In which multiple sources of Hadoop applications run at different levels. This protocol is an extension of Hadoop distributed file system (HDFS). The advantage of this paper is solving environmental problems without requiring complex Hadoop server components.

CHAPTER THREE

METHODOLOGY

This chapter elaborates on the research design choices made in order to answer the main research question and fulfill the research objective. A literature review tells you if others have used procedures and methods similar to the ones that you are proposing, which procedures and methods have worked well for them, and what problems they have faced with them. Thus select a methodology that is capable of providing valid answer to your research questions.

To answer the main research questions, following method used:-

- ❖ Initial study
- ❖ Reviewing of previous studies
- ❖ Problem understanding.
- ❖ Discussion on the problem.
- ❖ Finally e-commerce users' segmentations framework is proposed.

3.1 INITIAL STUDY

The initial study is conducted to understand the current e-commerce, customer segmentation and big data analytics utilization in e-commerce customer segmentation so as to identify the organizational need, customer need, opportunity, and problems. As part of the research conducted indirect interviews.

3.2 REVIEWING OF PREVIOUS STUDIES

Documents that presents the e-commerce user segmentation needs of organizations were reviewed. Literature review was done to find out typical functionalities of existing in e-commerce user segmentation. For the development of the system two strategies were proposed: The first strategy was to develop e-commerce user segmentation system from scratch.

The second strategy was to customize/localize existing proprietary or open source software. To choose between the two strategies the pros and cons of each strategy has been studied. Based on the study the first strategy was selected. Consequently, several different segmentation tolls and techniques have been studied and no free tools to use it with the new systems and it exceeds time frame for this study. Generally for the development of the system the following procedures have been followed:

- ❖ Reviewing available tools e-commerce user segmentation management system
- ❖ Getting experience from experts(possible advisor)
- ❖ Selecting MVC framework
- ❖ Developing concept
- ❖ Coding the prototype
- ❖ Writing a project report

3.3 RESEARCH DESIGN

Research Design constitutes a problem-solving research paradigm, which aims in providing innovative solutions in the form of artifacts to related organizational problems. This is achieved through the analysis, management, design, implementation and use of information systems (Hevner et al., 2004).

In order to answer the research questions of this research, and achieve the respective research objectives, a research method appropriate for designing frameworks should be used. In view of this I introduced Design Science Research (DSR) which appropriate when the solution provided to a specific problem is the design of an abstract artifact, such as a framework, a prototype, a models, a new method etc.

3.4 FRAMEWORK DEVELOPMENT

To come up with the framework decision the marketers and users concern used. For this study some segmentation types are identified as important by the organization during the literature review and by studying the characteristics of those segmentation types the organization

model of the framework and the available service are selected.

3.5 RESEARCH STRATEGY

The research approach used in this study is exploratory research. This type of research would be usually used in situations where the subject of the research is relatively new and insufficient knowledge on the area exists (2011). Exploratory research is useful in situations where the researcher is aiming to obtain a better understanding of the situation. Exploratory research is very flexible regarding the methods that used for collecting information and insights (2006). This research is considered as an explorative research due to its characteristics. Explorative research is aligned with this research since the very goal of the research is to investigate the employment of e-commerce users segmentation and gain enough information to propose a framework for e-commerce company.

3.6 GUIDELINES TO DESIGN SCIENCE RESEARCH

In order to conduct and evaluate design science research a certain process should be followed. Hevner et al. (2004) suggested the following set of guidelines, which constitutes a clarification of how DSR should be conducted and how is placed into this research.

❖ Design as an Artifact.

The DSR bears as a result an artifact such as a model, a method or a framework that can be applicable. This research results in suggesting a framework for e-commerce user segmentation. The framework maps e-commerce marketing objectives with relevant e-commerce User's segmentation categories, which can assist these objectives. Moreover, it entails the use of analytic techniques that can be implemented by big data tools and technologies.

❖ **Problem relevance.**

The result of the DSR under investigation should be a solution based on technology and be relevant for a specific business problem. The result of this research attempts to highlight the gaps found in the existing literature related to e-commerce user's segmentation and its effectiveness to e-commerce user. Moreover, it raises the issue of big data, which constitutes a current trend in analytics, points out differences and similarities with data mining and explains its relevance to e-commerce user's segmentation.

❖ **Design evaluation.**

The resulted artifact should be evaluated precisely in terms of its utility, quality and efficacy. The framework that this research proposes is high level, and could not be tested on real situation due to time limitation. However, its quality is ensured as it is based on a concrete theoretical background and it is evaluated by experts in terms of utility and effectiveness.

❖ **Research contributions.**

The DSR should be clear enough in the areas of design artifact, design foundations and methodologies, in order to be effective. Respectively, this research contains a detailed explanation of the scientific methods and methodologies used for the design of the proposed framework.

❖ **Research rigor.**

The construction and the evaluation of the artifact should be conducted with the application of rigorous methods. Therefore, the construction of the framework was based on a combination of literature and empirical data.

❖ **Design as a search process.**

In order to search for an effective artifact it is mandatory to utilize available means to achieve the research goals and comply with the laws of the problem environment. The design of the

framework of this research is considered effective, since it is based on existing theory and data observations which were able to build the base for the final result.

❖ **Communication of research.**

The research should be presented and explained effectively to both technology and management oriented people. The research was presented in people related to the fields of e-commerce, business, and software engineering data analytics.

3.7 DESIGN SCIENCE RESEARCH METHODOLOGY (DSRM)

The approached of DSR used for the purpose of this research is the design science research methodology, as it was proposed and defined by Peffers et al. (2008). DSRM process is considered suitable for this research, since the main objective of a DSRM is to provide a certain artifact, which contains the characteristics of the research outcomes. In this research the artifact is a framework that constitutes an indirect mapping among big data techniques for e-commerce segmentation goals.

The DSRM consists of the six activities as described below:

- 1. Problem identification and motivation:** Define the specific research problem and justify the value of a solution to this problem.
- 2. Define the objectives for a solution:** Identify the objectives for the solution to the problem.
- 3. Design and Development:** Create the artifact.
- 4. Demonstration:** Show how the artifact is capable of solving one or more instances of the problem.
- 5. Evaluation:** Measure and observe how well the artifact supports a solution to the problem.
- 6. Communication:** Communicate the problem and its importance, and indicate its effectiveness to professionals and researchers. Make suggestions for future research.

3.8 RESEARCH PLACED INTO DSRM

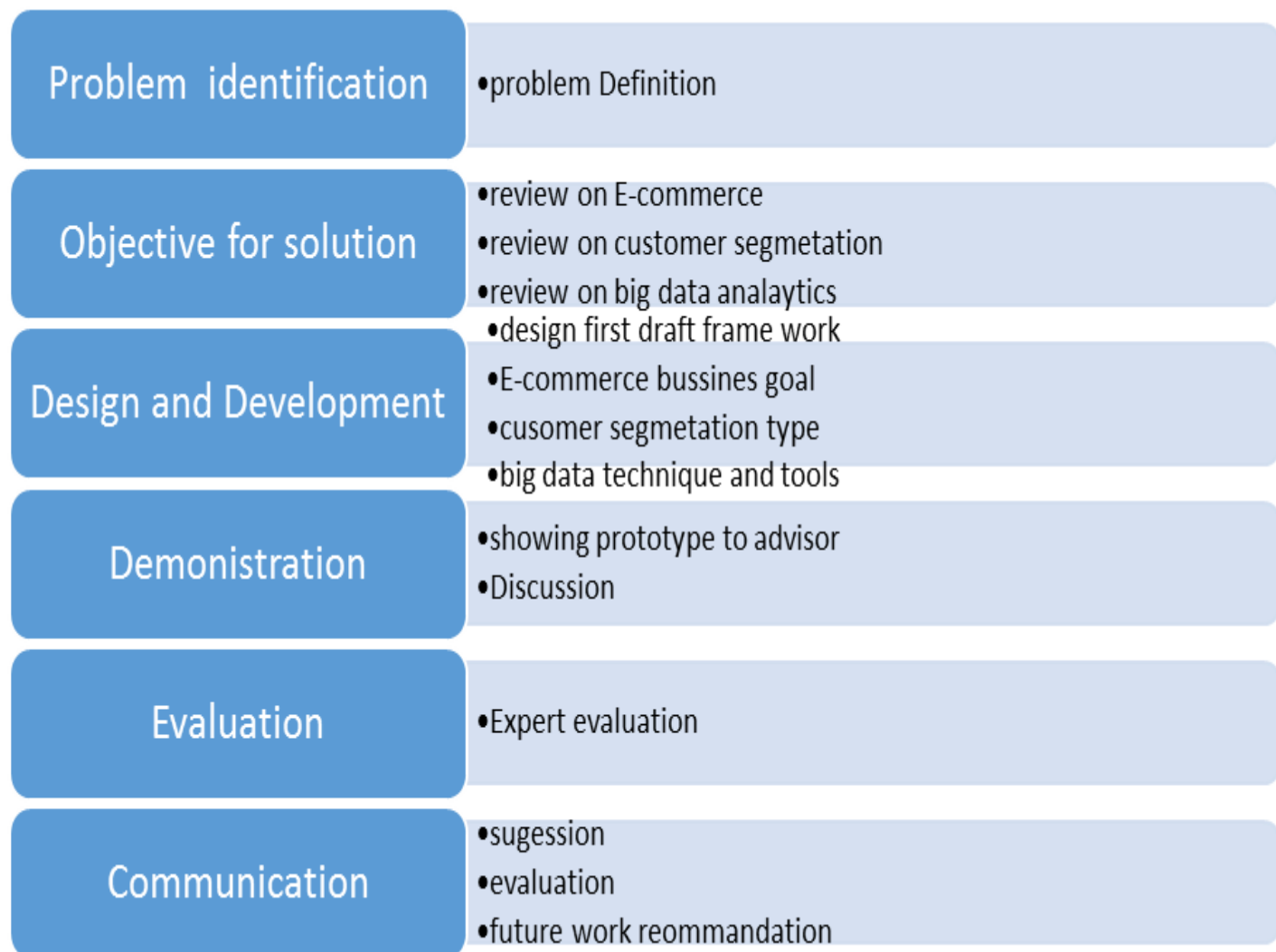


Figure 4 : Framework of Research Approach

3.9 STEPS OF METHODS

Step-1:

The method of the problem is described from both scientific and practical point of view. A literature study is performed in the domains of big data and User's segmentation, so as to discover knowledge gaps. Additionally, several discussions forum, blogs of professionals of in this field were studies prior to the start of the research in order to specify the problem.

Step-2:

The objectives for the framework are defined. At this point the literature review will focus on gathering data, the analysis of which will be able to answer each of the above-mentioned research questions.

Initially, a set of business objectives regarding e-commerce marketing determined. The set of goals stems from both the scientific theory and the data gathered. Secondly, several user's segmentation types are thoroughly described.

Step-3:

A hypothetical yet representative example of how the framework could be utilized in the OCEM tool's case is given. Since the proposed framework is high level and theoretical, an implementation on a real situation would require time that would greatly exceed the time limit of this research. Furthermore, for the evaluation, which constitutes the next step of methods the framework is validated by advisor.

Step-5:

Recommendations on how to utilize such a framework could be utilized are made. Furthermore, related gaps and limitations are discussed, while suggestions for future research are made.

3.10 PROCESS DELIVERABLE DIAGRAM (PDD)

The following Figure 5 illustrates the process described above in the form of a Process Deliverable Diagram (PDD). The processes of the DSRM approach are illustrated in the left part, while the related deliverables are illustrated in the right part. Moreover, in the right part of the diagram, the corresponding parts of the research in which the deliverables are presented are depicted.

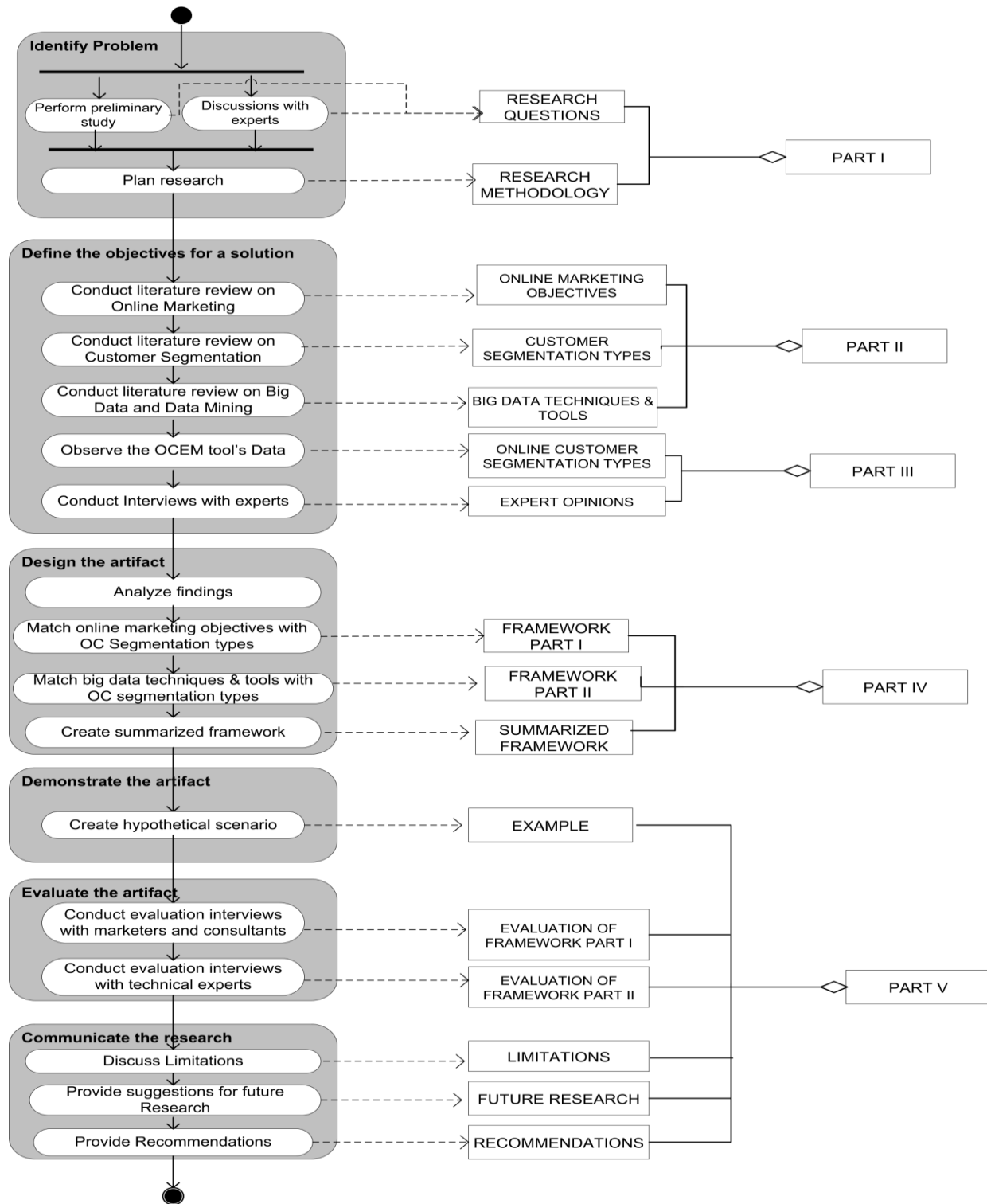


Figure 5: PDD of Research Approach

CHAPTER FOUR

FRAMEWORK OF E-COMMERCE USER'S SEGMENTATION

In this Chapter a framework for e-commerce user segmentation is proposed. As it is comprehensively described below the framework have a basic inputs the results that occurred from literature study gathering. The framework consists of two parts. In the first part the set of the objectives related to e-commerce marketing and User's engagement which are able to assist each of the business goals. In the second part, the segmentation types are matched with relevant techniques and algorithms introduced that can be applied in large data sets. The summary of the two parts is a table, which matches e-commerce marketing objectives with segmentation types and appropriate techniques and algorithms.

4.1 MATCHING E-COMMERCE MARKETING GOALS WITH SEGMENTATION TYPES

As depicted by the scheme in Figure 6. The first part of the framework for e-commerce user's segmentation that this research proposes entails the mapping of business objectives regarding e-commerce user's segmentation types, which can effectively assist each of those objectives. Regarding e-commerce user's segmentation types are defined as certain categories of User's attributes that can be gathered from e-commerce channels . Based on this segmentation types analysis can be made in order to define underlying e-commerce user's segments.

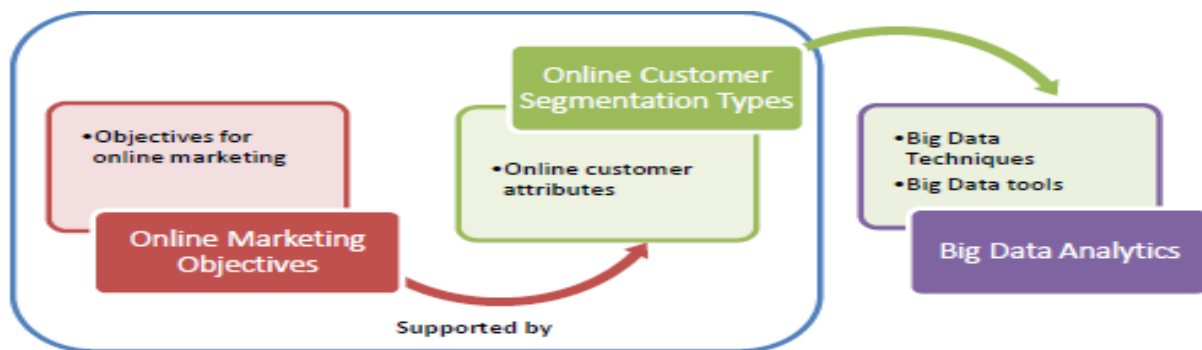


Figure 6: e-commerce User's Segmentation types (1st step of the frame work)

The first part of the framework for e-commerce user segmentation is illustrated in the form of a matrix in Table 2. As it can be seen, the matrix shows which e-commerce marketing objectives can be assisted by e-commerce User's segmentation types. Since there is not one-size-fits all segmentation for each business objective none of the User's segmentation types can be excluded during User's segmentation. However, the framework suggests, which segmentation categories are more preferable and useful for assisting each of the related e-commerce marketing objectives. In order to indicate the usefulness of each segmentation type for each of the e-commerce marketing objectives a High-Medium-Low scale is used. More specific:

- ❖ **H:** Indicates that the segmentation type is highly useful for the corresponding objective, and therefore should be preferred.
- ❖ **M:** Indicates that the segmentation type is useful in some cases and complementary to the highly useful segmentation types.
- ❖ **L:** Indicates that the segmentation type is not that useful for the corresponding business goal, especially if it's used alone.

Table 2: E-commerce Marketing Objectives with appropriate segmentation types

E-commerce marketing objective with Appropriate segmentation							
	Attitudinal	Behavioral	Demographic	Loyalty based	Referral	Technical	Value-based
Objectives							
Increase New User's Acquisition	M	M	L	L	H	H	L
Improve Churn Rates	M	H	L	L	L	L	H
Increase User's Satisfaction	H	H	L	M	L	L	L
Increase User's loyalty	L	H	L	H	L	L	L
Increase Cross – up sales	H	H	M	L	L	L	H
Improve conversion rates	H	H	L	L	M	M	M

4.2 TECHNIQUES FOR E-COMMERCE USER'S SEGMENTATION

The following figure shows the second step of the framework, which entails the techniques that are appropriate for each of the previously presented e-commerce User's segmentation types. As shown on the scheme the analysis of the e-commerce User's segmentation types requires certain techniques which can be implemented by big data tools. From the literature research on big data field, it became obvious that the most used and well known tool that can implement algorithms for data analysis in large databases is apache mahout. Apache mahout can implement various data mining techniques that can be used for segmentation in very large and complex databases.

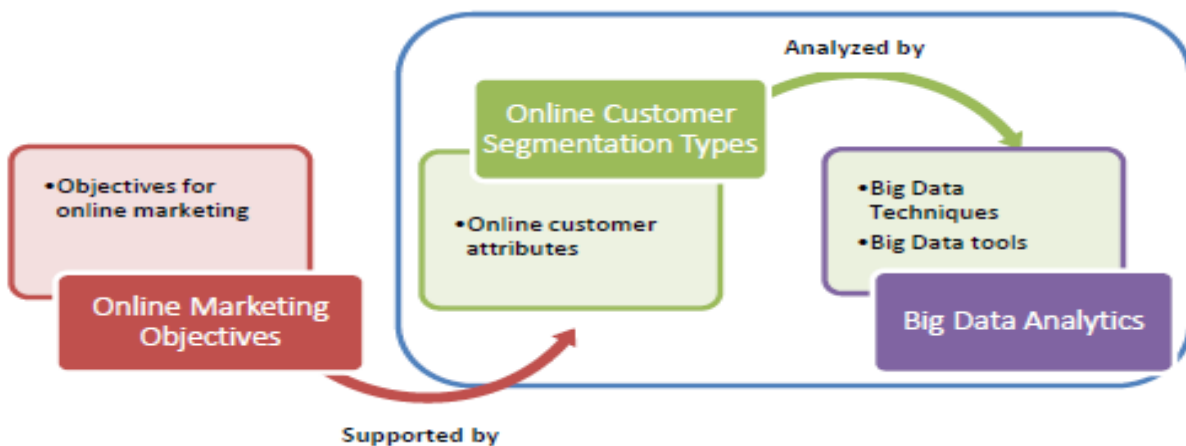


Figure 7: Techniques for E-commerce User's Segmentation

The link between e-commerce User's segmentation types and Big Data tools and techniques is mostly based on the literature review made on User's segmentation using data mining techniques. The results from the literature review are used as an input for the framework. Moreover, experts' opinions as per interviews & blogs were taken into account for the creation of the framework.

Table 3 presents the second part of the framework that this research proposes. As it can be seen the table consists of three columns, and it suggests techniques that are more suitable for analyzing each of the e-commerce user's segmentation types, as well as appropriate tools that can implement these techniques in large data bases.

Clustering

Clustering technique suits in most cases for user's segmentation is clustering. Cluster analysis belongs to the unsupervised modeling techniques. In that case all user's' attributes from a selected set can be simultaneously analyzed, in order to create User's segments with similar characteristics. A cluster model is able to manage a large number of attributes and reveal data-driven segments which are not known in advance. As it was also mentioned clustering is used when it is difficult to know the segments in advance, and thus a general algorithm is needed. Therefore, clustering is preferred for behavioral segmentation.

An example would be: If a marketer wants to create segments of user's in order to identify which of the users of the website are more active and visit the website frequently, then a cluster algorithm can be used in order to analyze behavioral data and create user's segments using behavioral data. The clustering algorithm could take as inputs a combination of behavioral attributes such as number of clicks, number of visits, time of last visit etc. The algorithm would create homogeneous groups of user's which would show which of them are more active or less active on the website. One well-known big data tool that can be used for the implementation of clustering techniques is Apache Mahout.

Classification

On the other hand, a very popular technique for user's segmentation is classification, which is used for supervised modeling. Classification is used when it is required for the segments to end in a specific result. This means that there is always a dependent variable according to which segments are created. According to the results of the literature review regarding techniques for User's segmentation, it is preferably used for loyalty-based segmentation or value-based segmentation. The amount of money spent a user's spend or the degree of loyalty could constitute the dependent variable. However, it can also be used in some cases for behavior segmentation or for attitudinal segmentation, but still is preferred for loyalty-based or value-

based segmentation. An example would be: When a marketer wants to create User's segments in order to identify the most profitable user's that are probably willing to purchase more products, based on the average amount of money they spend, then a classification technique is possible to be used. The dependent variable in this case would be "the average amount of money spent", while the independent variables would be attributes like "amount of orders", "average order value", "number of purchases" etc. A classification algorithm would reveal groups of user's that are more profitable and thus more likely to move on with more purchases. Classifications techniques can be implemented as well by the big data tool Apache Mahout.

The techniques that are described below are not straightforward segmentation approaches. However, they can assist e-commerce user's segmentation and enhance decision making, by providing holistic approach of the e-commerce user's.

Association

Association is a technique, which also belongs to unsupervised modeling and is used for recommendation engine for web page personalization .Association is not a straightforward segmentation technique, since it is basically used for uncovering relations among the data. However, some characterize it as segmentation technique since it is capable of "predicting" user's' preferences according to previous behavior so segments with recommended products for certain users can be created. Therefore, it is highly useful for segmentation based on user's' preferences, in order to find out what type of content the visitor might prefer to see or what type of product they might prefer to buy. Apache Mahout is a big data tool that provides recommendation engines able of predicting patterns of user preferences out of very large amount of data (Owen et al. 2012). Thus, it is very effective when used for products recommendations and web page personalization based on the user preferences.

Regression

Regression is a traditional statistical technique which is basically used for prediction and forecasting. Although regression analysis cannot be considered a straightforward segmentation technique, it is actually able to create User's segments by predicting the segment in which a User's belongs. It is very useful for loyalty segmentation, where the users are segmented according to their degree of loyalty (loyal or disloyal). Regression analysis is also useful for value-based segmentation where users are grouped according to their profitability, predicting if a client is profitable or not.

As mentioned before, R is an open source programming language which contains libraries for statistical analysis that can also be implemented in huge amount of data. Currently, R is a known open source tool that used for regression analysis of big data.

Visualization

Visualization is not a segmentation technique itself, but is used complementary to segmentation in order to illustrate segments and make them more understandable. Especially, in the case of a software tool that is addressed to people with limited knowledge of analytics, visualization of data becomes extremely useful for simplifying complex data patterns and thus, facilitating decision making.

For e-commerce user's segmentation, it is very important to visualize huge amounts of e-commerce User's data into segments with the use of visualization tools for big data, indicating preferences, behaviors or consumer characteristics. There are plenty of tools for visualization of Big Data such as Visual Insight & Jaspersoft.

Two new e-commerce user's segmentation types were identified based on e-commerce users attributes; 1. Technical 2. Referral.

These segmentation types are not found in the literature, and thus it is not possible to make an assumption for the big data techniques that are suitable for analyzing these data, without implementing and comparing the techniques on real data. Some examples could be: Create clusters of users according to the device they use or the keyword they searched for; Use classification in order to find which clients came from a certain affiliate website or from a certain social media channel.

In the next table 3 shows 6 high level business goals that e-commerce marketers usually have. The objectives shown stem from the literature review. Each of the goals can be assisted by certain segmentation types. Based again on the literature and on the discussions with the experts the relation between the business goals and the segmentation types are defined. More specific, in order to identify the level of usefulness of the segmentation types for each of the business goals the following factors were taken into account:

- ❖ Definition of the objective
- ❖ Measurements of the objective
- ❖ Usual actions related to the achievement of the objective
- ❖ Opinions of the interviewees regarding the objectives in relation with the users attributes, on which more attention should be paid for the achievement of the objectives.

Table 3: Framework -Techniques for e-commerce user's segmentation

Method	Segmentation types					Big Data Tools				
	Clustering	Association	Classification	Visualization	Regression	Recommender Apache Mahout	Clustering apache Mahout	Classic apache Mahout	Linear Logistic	Visual Insight & Jasper soft Attitudinal
Attitudinal	x	X		X		x	X			X
Behavioral	x			X						
Demographic	x			X						
Loyalty – based			X		X			x		
Technical	x		X	X			X	x		X
Value Based	x		X	X					x	
Referral	x		X	X			x	x		X

The Analysis of the above framework:

4.2.1 INCREASE USER'S SATISFACTION

User's satisfaction is a very important goal that e-commerce provider set. Therefore, it is crucial, that e-commerce user's or visitors get relevant information from the website and are exposed to certain interactions, such as banners, light boxes, or emails. The most effective type of segmentation in this case is attitudinal segmentation type, which includes characteristics that indicate User's preferences and interests. By using this segmentation type, e-commerce marketers can provide their e-commerce users with relevant content and relevant offers based on their needs, avoiding having them exposed to irrelevant content that in most cases becomes annoying. User's' satisfaction is difficult to be captured and measured e-commerce, since most of the times it can be obvious by a phone call or by filling in a satisfaction form. However, segmenting the users according to behavioral characteristics or their engagement/loyalty score reveals the level of satisfaction, since the frequency of visits and interactions is an indication the level of visitors' satisfaction.

4.2.2 INCREASE USER'S LOYALTY

It is essential for e-commerce marketers to gain their users trust, increasing loyalty with the brand and further engaged with their e-commerce users. According to Peterson (2004) "loyalty can be measured as the number of visits any visitor is likely to make over lifetime as a visitor", and can be typically measured by the number of visits per visitor. As explained before engagement or loyalty score can be calculated by setting certain business rules.

For example, a simple business rule could be based on the number of visits of a visitor: if a user's visits more than two times per day a website then he scores a point. The aggregated number of points results in a score that indicates user's loyalty. Such score shows how often the visitor visits the website or their level of interest for certain products or content. Loyalty-based segmentation includes attributes that indicate the loyalty and the interest of the users for the brand. Hence, it is very important to be used as a segmentation type for identifying loyal or disloyal user's and support the increase of User's loyalty. According to Peterson (2006) the frequency of visits, or the time that a visitor spends on a website also indicates user's loyalty.

Therefore, behavioral segmentation can also be used for distinguishing loyal and disloyal users, and proceed to corresponding marketing tactics to increase user's loyalty. Other User's attributes are of low importance for supporting the increase of user's loyalty.

4.2.3 INCREASE CROSS-UP SALES

The increase of cross and up sales belongs to User's development stage of user's lifecycle. It is a very important objective for e-commerce retailers and entails the selling of additional and more expensive products to users. In order to achieve the increase of cross up sales, the e-commerce marketer has to understand the needs and preferences of the Users their buying behavior. Moreover, most valuable and loyal Users should be targeted for the sale of products of higher

price. Therefore, it is suggested that e-commerce marketers should focus on the following segmentation types:

- (1) Attitudinal, Value-based and Behavioral segmentation. Attitudinal segmentation type should be used by e-commerce marketers in order to provide Users with relevant products and promotions based on their needs and lifestyles.
- (2) Behavioral segmentation is also important, since it shows the most active clients, who will be more willing to move to a purchase.
- (3) Demographics, like the age, genre or the marital status, are also valuable for delivering relevant content to the User's and increase cross-up sales. However, as it is stated by B3 "demographics category contains attributes that are not easy to be captured e-commerce, and thus are only useful in some specific business cases".

Therefore, since it can be used complementary to the attitudinal segmentation type is considered of medium usefulness. Tailored propositions can be offered to them according to their needs, so as cross-up sales can be increased.

Additionally, decisions about the design of new products or the improvement of existing ones can be made. However, it is also important to take into account how profitable the Users are and how willing they are to proceed into further purchases. Therefore, value based segmentation type should be also used for identifying and targeting the most profitable Users for the increase of up sales.

4.2.4 IMPROVE CONVERSION RATES

The improvement of conversion rates is as stated that "The performance of conversion rates is actually what e-commerce marketers are judged upon. A conversion entails the successful completion of certain activities by e-commerce user's that have a positive effect for the e-commerce organization. There are different ways to measure conversions, according to each activity. However, the high visiting frequency of the website is the basic measurement, since

conversion rates are calculated by the proportion of visitors that move on to a certain activity on a website, such as click on interactions, proceed on a purchase or on a subscription and obtain information (Winer, 2001).

As far as the meaning of conversion is concerned, “it always depends on the type of business”.

For example for an insurance company, a conversion could be an e-commerce application for insurance. For other organizations can be a purchase, a contact request, a download of a product or just a click on a banner. Therefore, for the improvement of conversion rates all segmentation types can be important, depending always on the marketers' point of view. Attitudinal and referral segmentation types can play an important role for improving conversion rates, since the first indicates preferences and the latter the incentives that lead the visitor to the website, both allowing for distributing relevant content to the visitor. Past behavior is also essential for the improvement of conversions”.

This is attributed to the fact that it shows how frequent the client visits, the number of actions they already did on the websites, indicating their willingness to move on with further actions. Therefore, behavioral segmentation is also important for e-commerce marketers, in order to identify users with lower activity level and decide upon further marketing actions to trigger those users to become more active on the website. Moreover, the adjustment of the content according to technical characteristics, can also lead the e-commerce User's to stay on the website and complete several activities, thus, technical segmentation could be also useful in some cases.

Final E-Commerce User's Segmentation framework

Table 4 constitutes an overview of the two frameworks, showing which e-commerce user's segmentation types should be preferably analyzed by certain techniques in order to assist the achievement of main e-commerce marketing objectives.

In the table shows the level of usefulness of the e-commerce user's segmentation types based on

e-commerce user's attributes gathered from e-commerce channels is indicated by different colors. In the colored cells the big data techniques suitable for analyzing each e-commerce user's segmentation type, are shown in table 4.

Table 4: Summary of E-commerce User's Segmentation framework

	E-commerce User's Segmentation Type						
	Attitudinal	Behavioral	Demographic	Loyalty-based	Referral	Technical	Value based
Objectives	Big Data Techniques (Implemented by big data tools)						
Increase New User's Acquisition	Association Clustering Visualization	Clustering Visualization	Clustering Visualization	Classification Regression	Classification Clustering Visualization	Classification Clustering Visualization	Classification Regression
Improve Churn Rates	Association Clustering Visualization	Clustering Visualization	Clustering Visualization	Classification Regression	Classification Clustering Visualization	Classification Clustering Visualization	Classification Regression
Increase User's Satisfaction	Association Clustering Visualization	Clustering Visualization	Clustering Visualization	Classification Regression	Classification Clustering Visualization	Classification Clustering Visualization	Classification Regression
Increase User's Loyalty	Association Clustering Visualization	Clustering Visualization	Clustering Visualization	Classification Regression	Classification Clustering Visualization	Classification Clustering Visualization	Classification Regression
Increase Cross –up Sales	Association Clustering Visualization	Clustering Visualization	Clustering Visualization	Classification Regression	Classification Clustering Visualization	Classification Clustering Visualization	Classification Regression

CHAPTER FIVE

ANALYSIS & INTERPRETATION

This chapter presents the results of the evaluation of the two frameworks presented in previous chapter. The first part of the framework, which matches the e-commerce user's segmentation types with e-commerce marketing goals, was evaluated by advisor or this research. The second part of the framework, which matches the e-commerce user's segmentation types with the big data techniques and tools, was evaluated. Moreover, in order to illustrate how the two parts of the framework would work in practice, a walk-through of the framework presenting an ideal scenario of how the frameworks would be used in a hypothetical situation is provided.

5.1 WALK-THROUGH OF THE FRAMEWORK

Since it was not possible to test the framework on a real situation, a hypothetical yet representative example of how the framework could work in the case of a user's that uses the tool is provided below. This example serves in better understanding how the frameworks that this research proposes could be used in practice. For the creation of this example, a sample of user's profiles created by the OCEM tool was used. The hypothetical business scenario presented below is based on a real business case of one organization-client.

Business Scenario

Company X is a TV and radio broadcaster that uses an OCEM tool provided called the OCEM tool. X owns a bunch of e-commerce channels, since apart from the main website it owns a different website for every TV or radio program, Facebook pages and twitter streams. Company X has employed the OCEM tool in order to create User's profiles by collecting e-commerce user's attributes in real time from all owned channels, understand individual users' journey,

understand users' behavior and provide more relevant content and information to the individual User's. Through the websites of company X a visitor can purchase a membership.

In general company X focuses on increasing user's satisfaction, improving churn rates and conversion rates and attracting new users. The objective that e-commerce marketers of company X should reach this year is to improve conversion rates by increasing the number of visitors that choose to move on with purchasing a membership. In order to reach the main goal, company X focuses on promoting the membership by showing to user's relevant banners according to their preferences, interests and past activities, and instead showing additional promotions, in order to avoid churn and decrease churn rates. Moreover, they want to avoid showing banners for memberships to those who are already members. Table 5 shows the e-commerce user's attributes that company X gathers by using the OCEM tool. Table 6 shows how these attributes fall under the e-commerce user's segmentation types, as they are defined in this research.

Table 5:e-commerce user's segmentation types

Categories	Online Customer Attribute gather by Company X
Person	Age, Gender, City, Job, Member, Email, Phone, Social works
Statistics	Number of Visits Interactions Clicked, interactions Viewed Average Visit Time Last Visited time Entry page Keywords Click count engagement
Other	Gift Preference Payment preference Program preference Interest Visited page Mode of payment Delivery Duration Educational status

Table 6: User's Attributes gathered by the OCEM Tool for Company

E-commerce User Segmentation	E-commerce Users Attribute Gathered by company X
Attitudinal	program preference Visitor Type Interest Entry page Visited site
Behavioral	Number of Visits Number of click per page Last visited time Click count
Demographic	Age, Gender, City, Job Email Phone number
Locality based	Engagement Score
Referral	Refer host name Keyword on website
Technical	Flash version Java version OS, browser Name
Value Based	N/A

In order to improve the conversion rates of visitors becoming subscribers and also improve the churn rates company X can do the following:

- ❖ Distinguish the non-subscribers from subscribers.
- ❖ Identify how often users visit company X websites.
- ❖ Identify e-commerce visitors' preferences
- ❖ Identify e-commerce visitors' interests.
- ❖ Promote banners based on visitor interests.
- ❖ Promote banners based on visitors' preferences.
- ❖ Show more relevant content based on visitors preferences and interests.

- ❖ Show particular banners for members.
- ❖ Do not show banners for membership to members.

An effective segmentation of the visitors of the e-commerce channels that company X owns, would help in conducting the aforementioned tasks. According to the proposed framework in order to increase conversion rates one can start segmenting their e-commerce visitors based on the behavioral and attitudinal segmentation types. Tables 7 and 8 show behavioral and attitudinal attributes with a short description of each attribute. By using a clustering technique clusters of users can be identified according to attitudinal and behavioral characteristics.

Table 7: Behavioral Attributes

Behavioral attributes	Description
Number of Visits	Total number of times that the User's has visited X's website
Clicks per visit	Number of clicks per visit.
Average visit time	Average time spent on time per visit
Click count	Total amount of clicks
Member	Visitor owns a membership or not
last visited time	The last time that an e-commerce User's visited one of X's websites.

Table 8: Attitudinal Attributes

Attitudinal attributes	Description
Interactions Clicked	Interactions on website clicked by the User's
Interactions Viewed	Interactions on website clicked by the User's
Visited page	Number of clicks per visit
Entry X page	The first X's page that the User's visited
Interests	Visitors interests in categories: music, reality, nature etc.
Program preference	Preference to a specific program (TV or radio)
Program fan	Fan of a specific program (TV or radio)

Behavioral segmentation type

For the purpose of the example a sample of 200 user's profiles were obtained, and the IBM SPSS statistic data editor was used for analyzing the data. In order to create clusters of users, the two step algorithm, which belongs to the clustering techniques, was used. The results of the segments are visualized in Table 9.

Table 9: behavioral segmentation Type (adopted from E-commerce user segmentation (2012))

Input (Predictor) Importance


Cluster	1	4	5	3	2
Label	Moderate active visitor (unknown ...)	Not a Member	Moderate active member	Highly active member	Highly active visitor (unknown ...)
Description	Unknown Membership, lower visited time and clicks	Visitor is not a member	Visitor is a member with lower visiting time and visits	Visitor is a member, with high visiting time and clicks	Unknown Membership, but highly number of clicks and visits
Size	 66,5% (799)	 16,3% (196)	 11,9% (143)	 3,6% (43)	 1,7% (20)
Inputs	ismember 0 (100,0%)	ismember 2 (100,0%)	ismember 1 (100,0%)	ismember 1 (81,4%)	ismember 0 (55,0%)
	visits 291,99	visits 249,40	visits 383,76	visits 1.414,07	visits 728,50
	averagetime 0,71	averagetime 1,03	averagetime 0,50	averagetime 1,19	averagetime 41,80
	clickcount 762,60	clickcount 773,44	clickcount 1.003,36	clickcount 4.087,53	clickcount 7.556,20
	visitclicks 2,76	visitclicks 2,96	visitclicks 2,59	visitclicks 4,33	visitclicks 26,65

As it can be seen in Table 9 revealed 5 basic e-commerce user's segments with similar behavioral characteristics. The pie chart in Figure 8 below constitutes a simple visualization of the clusters and their size. Table 9 illustrates the clusters and the average values of the specific behavioral attributes. As it can be seen the attributes are sorted and colored according to their input importance, as it was decided by the algorithm.

The largest group labeled as “moderate active visitor (unknown membership)” contains of the e-

commerce users. The e-commerce users that belong in this group for whom it is not known yet if they own a membership, and the level of activity according to their behavioral attributes (visiting frequency, number of clicks, average time spent).

The second group labeled as “Not a member” includes the e-commerce users. This group consists only of e-commerce users for who is known that they are not members.

The third group labeled as “moderate active members”, shows the members who are less active, while the fourth labeled as “highly active members”, the members who are more active during their visits on X websites., the smallest group labeled as “highly active visitor (unknown membership)” includes visitors for who it is not known if they own a membership or not, but they seem to be highly active based on their behavioral attributes.

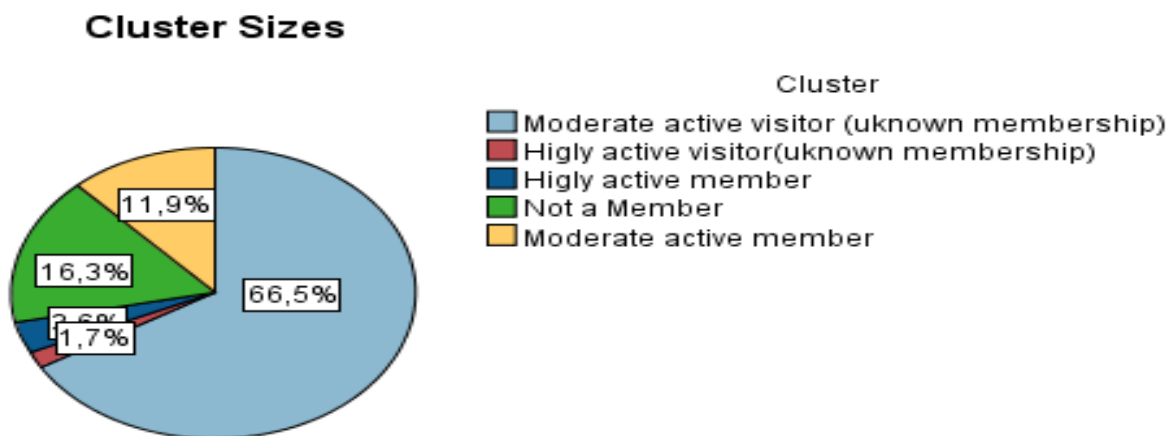


Figure 8: Size of E-commerce User's segments

For each of these groups e-commerce marketers can create different campaigns. For example, they can target e-commerce users that belong to “No Members” group, with relevant banners for membership. Moreover, they can target “highly active visitors” with relevant banners for memberships, since due to their highly visiting frequency they would be more likely to buy a membership. This strategy could help in leading more visitors to buy memberships, increasing

the related conversion rates. Moreover, e-commerce marketers could avoid showing to visitors who belong to the categories “Highly active member” and “Moderate active member” banners which promote the purchase of a membership, in order to avoid deregistration due to exposure to irrelevant content, improving churn rates.

Similarly, using a clustering technique for attitudinal attributes, segments of users with certain preferences and interests could be identified. Banners, email campaigns or other interactions could be based on the resulted clusters, in order to deliver more relevant content to the client.

The analysis was provided with a preliminary version of the framework and they were asked for their opinions on the relevance of the objectives and user's segmentation categories, and the level of usefulness of its segmentation category-

1. Relevance of e-commerce marketing objectives
2. Relevance of the e-commerce user's segmentation types
3. Usefulness of the E-commerce user's segmentation types

Description:

1. Relevance of e-commerce marketing objectives

As far as the relevance of the set of the objectives presented on the framework is concerned, I found that the objectives are all relevant to e-commerce marketing and can summarize the basic high-level objectives of an e-commerce marketer. Thus, analyze that this was relevant for the proposed framework. However, the improvement of conversion rates is at a level lower than the other objectives of the framework and the improvement of conversion rates is a very basic goal for e-commerce marketers, but it can be measured based on what a marketer considers a conversion. Although the improvement of churn rates and the improvement of loyalty are two

distinct objectives, there might be occasions that these two objectives are not easily distinguishable, and thus, they have to be well defined by marketers

Relevance of the e-commerce User's segmentation types

All segmentations types are appropriate for e-commerce User's segmentation and the e-commerce user's attributes fit perfectly to each of the categories. Categorization of attributes could also be integrated in a software product capable of capturing e-commerce user's properties and creating user's profiles, to make it easier for the marketers to define their segments. When they were asked if they would segment their clients according to referral and technical attributes all of the respondents answered positively.

2. Usefulness of the E-commerce User's segmentation types

In order to evaluate the framework in terms of the usefulness of each of the e-commerce user's segmentation types for each of the objectives, the interviewees were asked on which segmentation types they would focus more when they think of starting with a specific objective. The answers they provided confirmed the framework, while some useful remarks were made. More specific:

Users purchasing behavior, while for business sites it's easier to identify users that are more likely to churn by segmenting them according to their behavioral characteristics. After identifying user's profitability and the risk of churn according to behavioral and value-based characteristics. An important indication made that the keywords entered by the visitor indicate in a few cases the intention of the visitor to abandon the brand. "It has been observed that words "cancel" or "cancel subscription" followed by the name of the brand, is used by people who want to terminate subscription or abandon the brand. In that case segments according to the keywords users used are also important, to find out if they are about to churn or not".

When asked for the "Increase of Satisfaction" objective, all four interviewees responded that

attitudinal are the first characteristics that they have to be taken into account in order to create segments showing the preferences of the user's. As was commented by one of the interviewees "Users should be guided correctly through the channels and get what they need, in order to be satisfied. Therefore characteristics that indicate preferences are very important. However, at this point it should be underlined that satisfaction is not always easily measured only from the e-commerce behavior of the User's. "It is very important to have feedback from the clients to increase satisfaction. When a client fills in a form, or calls at the user's support center to complain or to solve a problem, then the degree of satisfaction can be determined. Therefore, it is important that information from offline channels should be integrated into the product and taken into account for user's segmentation", that it worth the case to conduct surveys for satisfaction, in order to be able categorize user's according to their degree of satisfaction, and move on with keeping the users satisfied.

When asked for the **"Increase of Loyalty"** objective, all four interviewees agreed that the loyalty-based segmentation type is very useful for the specific objective. A score that indicates the engagement, loyalty or the interest of the client in a certain brand should be always taken into account, so segments according to degree of loyalty can be created". An engagement or loyalty score can be calculated by setting specific business rules that regard behavioral characteristics. Thus, the interviewees agreed that behavioral segmentation is also important to start with, in order to enhance the increase of loyalty. The loyalty, satisfaction, and the churn rate are related, and thus, a certain segmentation type that should be taken into account for the three related objectives is the behavioral segmentation.

When asked for the **"Increase of Cross – up Sales"** objective that the first step is to see which users are most likely to buy products. Therefore, value-based segmentation and even more behavioral based segmentation is very essential in order to see whether a user's is likely to spend or whether they are willing to move on with purchases. The behavior segmentation is necessary in order to see what types of products a user's has already purchased, in order to avoid

suggesting the same products or be able to decide upon additional offers. All four interviewees mentioned that they would use attitudinal based on segmentation in order to segment users according to their preferences and suggest new products for purchase. It is worth saying that demographic segmentation important, but they believe that e-commerce user's characteristics such as age or genre are sometimes difficult or it takes time until they are captured e-commerce and added in a visitor's profile. As specifically I would use demographic segmentation when demographics are available and for specific type of products, such as insurances, for which demographic characteristics are very important.”

For the “**Increase of conversion rates**”, it depends on the e-commerce marketers' point of view and what they consider a conversion in each case. However, all agreed the attitudinal segmentation type and the behavior segmentation are important in most cases. Examples : “ Preferences and past behavioral is always the first thing you would see, when you consider as conversion visitors clicking on a specific banner, or moving on with purchasing a product or downloading content from your website. However, if somebody new enters referral or technical attributes would also be important to look into. Segmentation types like loyalty or a value-based segmentation, wouldn't add any value”. The low usefulness of value-based and loyalty-based segmentation is basically attributed to the fact that conversion rates are related with the activities that an e-commerce visitor completes while being on a website. This means that attributes that reveal visitors' preferences, intentions and behaviors are more important in order to lift conversion rates than the ones that reveal loyalty or profitability.

5.2 UTILIZATION OF THE FIRST FRAMEWORK

The indirect interviewees were also asked whether they would utilize the first part of the proposed framework for e-commerce user's segmentation and how this could be done according to their opinions.

B3 mentioned that such a framework would be useful as an overview for business consultants in order to advise their clients who make use of OCEM products upon segmentation. Moreover, he mentioned that the framework would constitute a guideline for use-cases considering campaign creations, in order to identify User's segments, to which the campaigns should target. The framework "would be useful for a quick overview when an e-commerce marketer start with specific objectives in order to know where they should focus or not, what they should do or not." He also added that "it could be given to the clients that use an OCEM software product in order to help them with segmentation. It would raise a lot of discussion and trigger them to think more on their objectives". It would be a first guideline to show how different segmentation categories or a combination of them are needed in order to help in the achievement of the objectives of e-commerce marketers. It would be also help the clients to think of their objectives based on their needs, and possibly it would fit for specific use cases including segmentation that would automatically be embedded in the product.

Moreover, all the categorization of the e-commerce attributes according to the segmentation categories would possibly fit in most cases in an OCEM tools that creates user's profiles by capturing e-commerce User's attributes. In such a case, the e-commerce user's attributes that are gathered could be also categorized in the product according to the segmentation categories.

Finally, would be the problems or the obstacles that would appear in the utilization of such a framework. Although the approach is interesting, it would require much time in order to see actual results on certain business goals. This framework would be the start on achieving a goal by effectively segmenting the users, but still there is much more that needs to be done in achieving those objectives. A standardization of segmentation types according to the proposed framework, could be integrated to a product for OCEM, but still there would be cases that would not fit exactly and, thus the product would have to be configured for the specific case.

To sum up, after analyzing, the first framework was confirmed by the experts and it was considered useful as a first step for an effective e-commerce user's segmentation that would enhance the achievement of objectives related to e-commerce marketing. Such a framework could be utilized as a guideline for effective e-commerce user's segmentation that assists business goals. However, due to the nature of e-commerce marketing it would take time to see the actual results of e-commerce user's segmentation on a specific objective.

With regards to the evaluation of the second framework, which shows which techniques -also applicable in big data- are appropriate for segmenting e-commerce users according to certain segmentation such a categorization would fit in an ocm tools, which creates e-commerce visitors profiles by collecting e-commerce attributes, in order to facilitate the segmentation process. From his point of view and his experience he mentioned that all classification, clustering, association, and regression are techniques relevant for segmenting users according to the characteristics that are gathered e-commerce.

For **clustering technique** is the prevailing and most commonly used technique for segmenting users where cluster analysis is concern. It is used when the natural groups that constitute the segments are not known in advance and a clustering algorithm is needed in order to define those groups and categorize the users .Clustering is the most commonly used for behavioral segmentation type and demographics. An example from banking sector would be: If it is required to identify the most profitable clients that are retailers, then a clustering technique should be used. In an e-commerce environment, you would always check how long the user's stayed in the website and how did he behave during his visit site." In that case clustering could be used to create user's segments that reveal their behavior in terms of visiting time, clicks, page views etc.

For Classification, "We use classification when we want the segments to end in a specific result". Classification technique is very effective for classifying users according to their loyalty score. In the case of value-based segmentation category he does not consider the classification

very effective, because normally the marketers want to identify the most profitable clients, without knowing the user's value. However, it can be used in cases for behavioral segmentation as well. An example where classification would best fit would be when through segmentation there is a need to identify clients that use more a certain product. Another example regarding the banking sector was: if it is required to identify clients that make high use of a certain banking product, then a classification tree should be used. Classification always useful when there is a dependent variable according to which segments can be created.

For **Association** is not a technique for segmentation, but there are several analysts and scientists who might treat it as such. Association is used for recommendation engines, and especially for basket analysis as it happens in the case of Amazon.com. Association is always useful to find out what visitors prefer to buy. As in an offline environment is very difficult to understand the preferences and attitudes of the clients, and therefore data from surveys or market analysis are used. In an e-commerce environment it is much easier to track some of those characteristics. In such cases association analysis is useful for characteristics that indicate user's preferences (attitudinal segmentation type). Respectively, association analysis very helpful for creating recommendation and for micro-segmentation according to the user preferences.

However, it estimates the probability that a user's belongs in a certain segment. It is often used to estimate churn probability and, thus, segment users to churners or non-churners, based on value-based characteristics.

For **Visualization** -Visualization is very important for illustrating segments providing clear view, and is similar to cluster analysis. It actually, illustrates the results of a clustering.

5.3 BIG DATA TOOLS

When an analytic tool for big data could be implemented in an OCEM software product to facilitate e-commerce user's segmentation stated: "From a theoretical point of view a big data

tool can be implemented in the product for analysis and user's segmentation. However, from a technical point of view it always takes time to integrate such a product, since technical perspectives such as performance or multi-tenancy, should be taken into account." Moreover, big data analytic tools such as Apache Mahout or language R could be used for analyzing e-commerce user's data and perform user's segmentation. These tools could also be used in an OCEM tool. However, operational problems might appear that would regard to set a tool to work in the production scale and configure it according to the performance limits and the settings they have set for the product.

Utilization of the second framework

The second framework for this research proposes, would be a starting point to look into segmentation types and implement tools , like apache mahout, that are able to implement segmentation techniques. It shows which techniques are preferred for each segmentation category. Then it could be easily check what works or not, according to each specific situation. So, the second framework that there are not one-size-fits-all segmentation techniques. However, clustering and classification are the prevailing segmentation techniques, which can be used according to the required result. The second framework was considered useful as a starting point to look into e-commerce user's segmentation types and the techniques and tools through which they can be analyzed. It constitutes an overview that one can use in order to check what can be used in each case, normally, the marketers should be able to ask for specific segmentation types, according to their needs and then data analysts are called to do the analysis and provide them with the best solution.

This opinion indicates that the second framework in combination with the first framework proposed in this research, could be used to help marketers decide on specific segmentation types which would assist them in achieving their goals, and data analysts to have an overview techniques and big data tools that would me more effective in analyzing the specific segmentation types in each case.

CHAPTER SIX

FINDINGS, REASERCH LIMITATION, CONCLUSIONS AND RECOMMENDATIONS

6.1 FINDINGS

Due to the tremendous amount of data generated daily from fields such as business that choose to e-commerce has resulted in the generation of tremendous amount of user's data. As the organizations struggle to handle and utilize effectively all the information available in order to provide better products and services and gain competitive advantage, e-commerce User's segmentation and the so called "big data" analytics are two fields that currently constitute matter of concern and discussion.

During the research a comprehensive literature review in the fields of e-commerce marketing, user's engagement, big data and data mining was conducted. The main objective of the research was to answer the main research question. The main deliverable of this research, which constitutes an answer to the main research question was a framework that shows e-commerce user's segmentation types able to assist e-commerce marketing objectives, and the techniques that can be used for big data analysis are suitable for segmenting e-commerce users according to each of, the e-commerce user's segmentation types.

The main research question is answered by three sub-research questions which are answered as follows:

RQ1: Which are the main business objectives regarding e-commerce marketing and e-commerce User's engagement?

The findings from a literature review regarding e-commerce marketing, focusing on finding main e-commerce marketing objectives. The literature review showed that organizations are currently building their e-commerce strategies and thus, their objectives are not clear and well defined in

most cases. However, they are related to the steps of user's lifecycle: acquisition, retention and development which can also be calculated by including e-commerce user's attributes that are gathered from e-commerce channels. In order to identify a set of high level business goals, and further discuss the matter a set blog and survey made on marketers. After the analysis, the following set of high-level business goals was defined:

- ❖ Increase new user's acquisition
- ❖ Increase user's satisfaction
- ❖ Increase user's loyalty
- ❖ Improve churn rates
- ❖ Increase cross-up sales
- ❖ Increase conversion rates

RQ2: Which are the user's segmentation types that can assist each of the business goals regarding e-commerce marketing?

In previous chapters the notion and the importance of an effective user's segmentation for achieving is explained, while user's segmentation types, which are based on attributes gathered from offline channels and sources are presented. The OCEM tool is taken as a case study, in order to figure out what kind of attributes can be gathered e-commerce and point out the differences from offline attributes. As it was found out the most of the attributes gathered e-commerce match with the attributes gathered offline and fall under the main user's segmentation categories, as they are defined from traditional user's segmentation theory. Two new e-commerce user's segmentation types were defined that are based on referral attributes such as keywords, or host site and on technical info such as screen resolution, operation system, browser, language. In total 7 e-commerce user's segmentation types were identified;

- ❖ Attitudinal (also in offline)
- ❖ Behavioral (also in offline)
- ❖ Demographics (also in offline)

- ❖ Loyalty based (also in offline)
- ❖ Value-based; (also in offline)
- ❖ Referral (only e-commerce)
- ❖ Technical (only e-commerce)

However, as it was concluded there is no one-size-fits-all segmentation. The main inputs of the framework were the data collected from the literature review. The framework shows the level of usefulness of each e-commerce user's segmentation type for each of the aforementioned e-commerce marketing objectives. As it was observed behavioral and attitudinal segmentation types are of high usefulness for most of the objectives.

RQ3: Which big data approaches and techniques can be used for each e-commerce user's segmentation type?

As it was found the techniques that can be applied on big data, stem from data mining and statistics. The difference is related to the characteristics of the data, while the analytic techniques that are used for big data analysis do not differ from those used for traditional data mining. The difference exists in the tools that are able to handle and analyze the large volume and the diversification of the data. As far as the tools are concerned, the most well-known open source framework for big data is Apache Hadoop, which includes tools for big data management, processing, storing and analysis. Apache Hadoop includes Apache Mahout for big data analysis using clustering, classification and recommendation engines (associations). Moreover, programming language R. Since the main techniques used for User's segmentation can also be applied in big data, a literature review on User's segmentation and data mining techniques was made. A set of 15 studies relevant to the topic was found, and trends of techniques used for certain types of user's segmentation were pointed out. The main techniques found were: clustering, classification, visualization, regression, associations. Clustering and classification are the techniques that are used more often and can be applied to the most of the user's segmentation

types, while the rest of the techniques are not straight segmentation techniques, but their results can be used for user's segmentations. All techniques can be implemented on big data by apache mahout and the programming language R.

The framework suggests which techniques can preferably be used for e-commerce user's segmentation according to each of the e-commerce user's segmentation types.

6.2 RESEARCH LIMITATIONS

The results of this research should be interpreted in regard to the following limitations:

Lack of prior research studies on the topic: Although there is prior research on the field of user's segmentation and data mining, there was no scientific research found on the opportunities that big data offer for the e-commerce user's segmentation. Furthermore, the related literature found focusing on e-commerce user's segmentation was limited. Moreover, there was no framework found that connects e-commerce user's segmentation to e-commerce marketing objectives. Therefore, the results of the research are based on a combination of literature data stemming from e-commerce marketing, traditional user's segmentation with data mining, and currently available big data theory, as well as on empirical data gathering.

Lack of available data: During the research it was not possible to interview the clients that use the ocem tool, in order to find out their objectives and needs in terms of segmentation. Since, the product is relatively new the clients still work with it under the guidance of consultants, and thus interviewing them it was not advisable. Moreover, it was not possible to observe user's data that are gathered by tools similar to the OCEM tool, in order to search for other possible e-commerce user's data that can be gathered e-commerce.

Time limitation: During the research discussions and interviews were conducted, internal documentation was reviewed, while data from the OCEM tool were observed. Due to time

limitation, it was not possible to test the proposed frameworks on a real-situation. First of all, the application of big data tool capable of implementing techniques for e-commerce user's segmentation, such as Apache Mahout on the OCEM tool, was considered time-consuming, due to the complexity of the tool and the configuration it requires. Moreover, as it was also stated by the experts testing the big data techniques in order to see the actual results in e-commerce marketing strategies , is a process that would require time that would much exceed the time limits of this research.

Moreover, at this point it should be underlined that this research focuses on techniques used for user's segmentation and not on specific algorithms that these techniques involve. This means that there are algorithms that belong to a specific technique that might perform better than the others. Moreover there are cases where algorithms that belong to a specific technique, cannot be applicable to very large datasets, basically due the fact that they are not as fast as required.

6.3 CONCLUSION

E-commerce user segmentation using big data analytics is a task of segmenting users into homogenous group based on their preference. It is a research area in the field of software engineering in relation with big data analytics. The problem to segment customers using big data analytics is e-commerce businesses are currently building their e-commerce strategies and thus they do not have well defined e-commerce marketing goals. Therefore, marketers do not provide data analysts with the appropriate information, while proceeding with valuable and effective e-commerce user's segmentation becomes difficult, therefore while segmenting we have no one-fit-size segmentation. E-commerce user's segmentation is a core process for assisting an e-commerce marketing strategy. However, there is limited scientific research related to the field. Huge amount of e-commerce user's data are continuously generated. However, there was no scientific research found for the use of big data tools in e-commerce user's segmentation. In the world of business a gap between e-commerce marketers and data analysts emerges. Normally, e-commerce marketers should be able to select combinations of e-

6.4 RECOMMENDATION

There are different research areas in e-commerce user segmentation using big data analytics user segmentations one of the issues. Customer Segmentation is an important component of many organization, and the algorithm designed in this work is expected to be a significant contribution to the field, and mainly to researchers working on various aspects of e-commerce user segmentation using big data analytics. Therefore, the researchers in the area can use the algorithm or the implemented system for processing segmentation as component in their research, mainly on big data analytics applications.

As a future work we would like suggest the following points:

- ❖ A general guideline for effective e-commerce user's segmentation is needed, while the opportunities that big data offer for e-commerce user's segmentation should be further explored.
- ❖ In feature research, the frameworks could be tested as a whole on more than one real situation. The actual usefulness of e-commerce user's segmentation types for each of the objectives, according to the first framework can be tested on a real situation.
- ❖ Starting with a certain e-commerce marketing objective certain e-commerce user's segmentation types can be selected to be analyzed for creating actionable user's segments.
- ❖ After the e-commerce user's segmentation types that should be analyzed have been chosen, the implementation of big data techniques and tools as shown in the second framework should be tested.
- ❖ More OCEM software tools, which are capable of gathering e-commerce user's attributes, should be observed, in order to explore if there are additional e-commerce user's attributes that could form new e-commerce User's.

6.5 PROTOTYPE GUI AND SAMPLE CODE

Development of the prototype system starts by identifying what is required in the intended system from the research work. In the previous chapters the problem area were

identified and understood, literature and related research works were reviewed and critiqued. Thus, some of the missing or undone parts from previous researches opened door for this research to develop a better field of study e-commerce user segmentation. Having this in mind the prototype is developed.

6.5.1. SERVICE MANAGEMENT

For managing the organization computing environment e-commerce user segmentation graphical management prototype is implemented. It provides the opportunity for all of the e-commerce products to be viewed and managed from a single support.

6.5.2 SOME GUI OF PROTOTYPE



Figure 9 home page of e-commerce user segmentation

The screenshot displays a web application interface for 'E-Commerce CUSTOMER segmentation'. The header includes a logo of a shopping basket with a red 'E' and navigation links for 'login' and 'home'. Below the header is a menu bar with 'Home', 'Registration' (the active tab), 'Items', and 'Rules And Regulations'. The 'Customer Registration' form is centered, featuring a 'preferences >' button in the top right corner. The form contains the following fields: 'Title' (a dropdown menu with '-- Select Title --'), 'First Name' (a text input field), 'Middle Name' (a text input field), 'Last Name' (a text input field), 'Gender' (a dropdown menu with '-- Select Gender --'), 'Birth Date' (a text input field), 'Email' (a text input field), 'Phone' (a text input field), and 'Country' (a dropdown menu with '-- Select Country --').

Figure 10.user Registration form

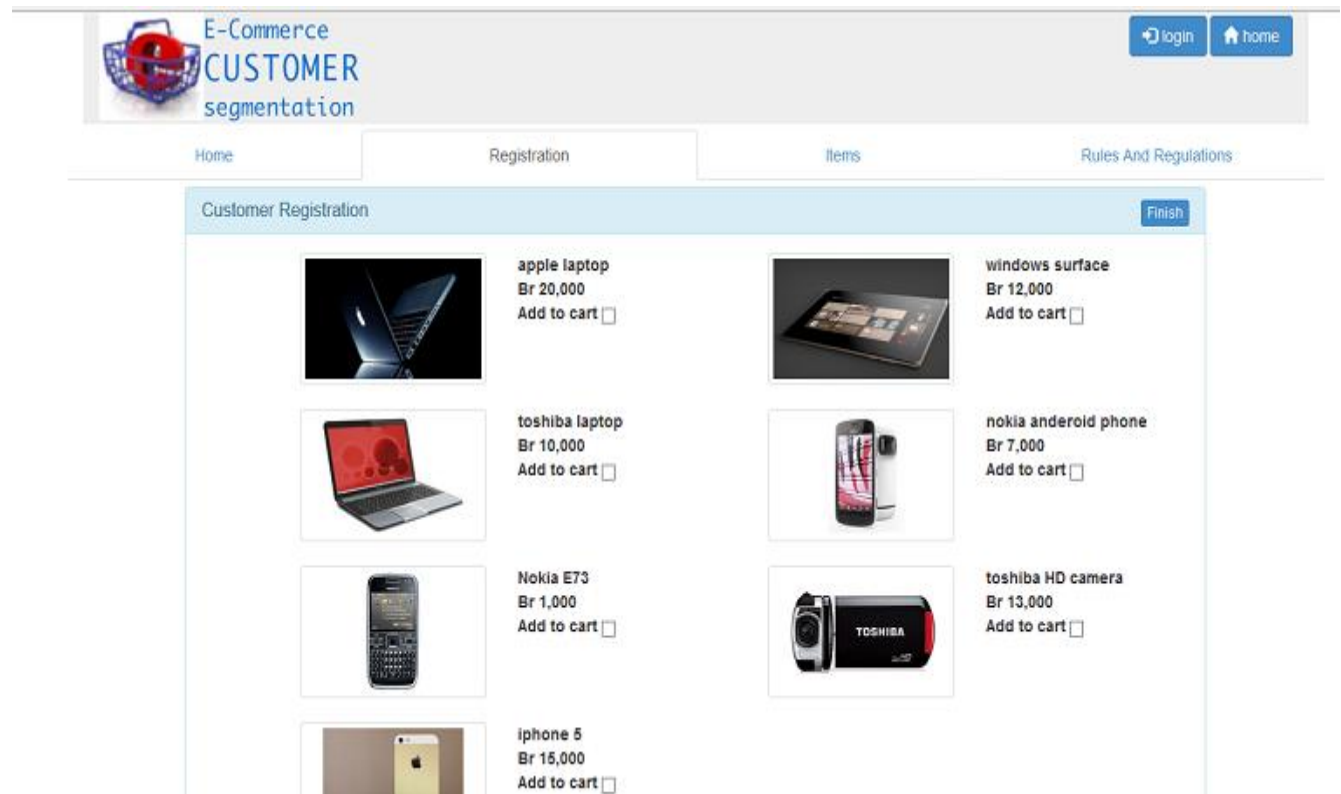


Figure 11. online available Items

A CONCEPTUAL FRAMEWORK OF E-COMMERCE USER'S SEGMENTATION USING BIG DATA ANALYTICS

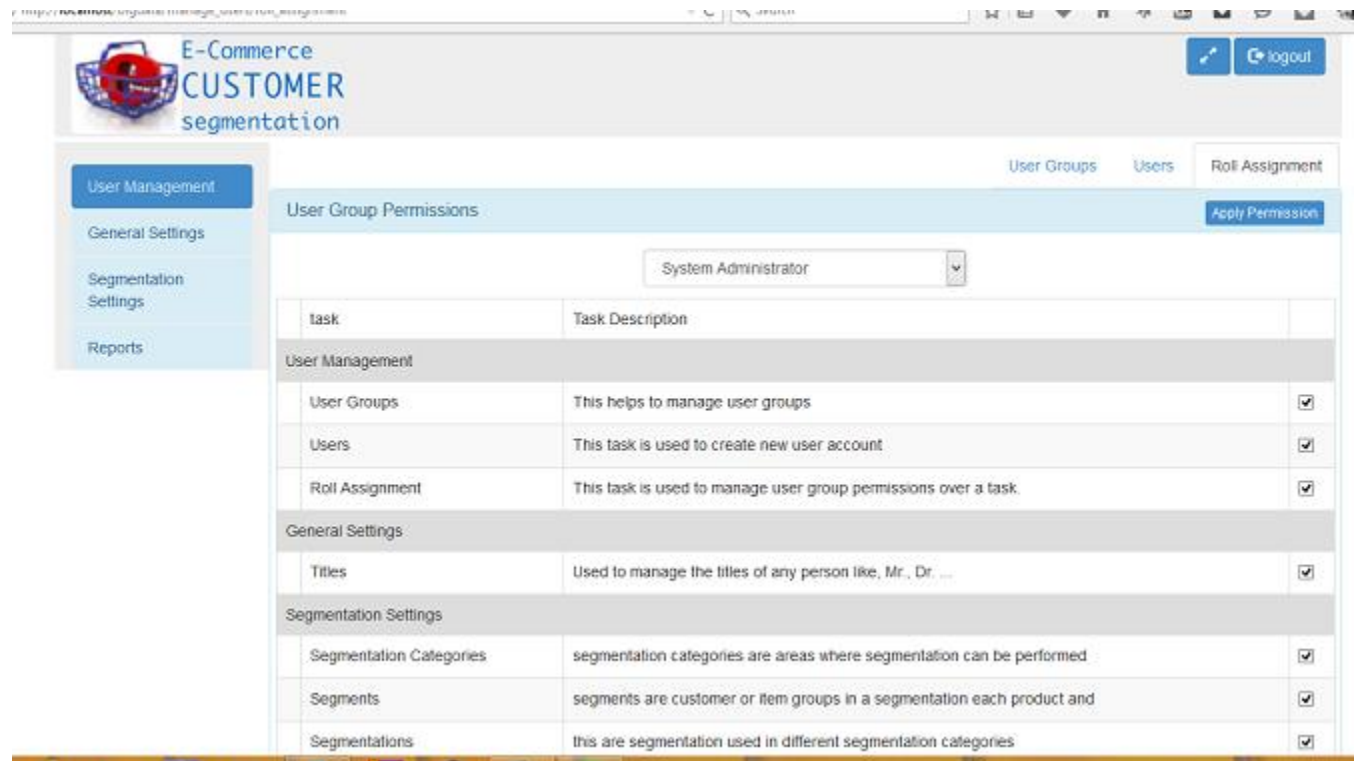


Figure 12.user management Role Assignment

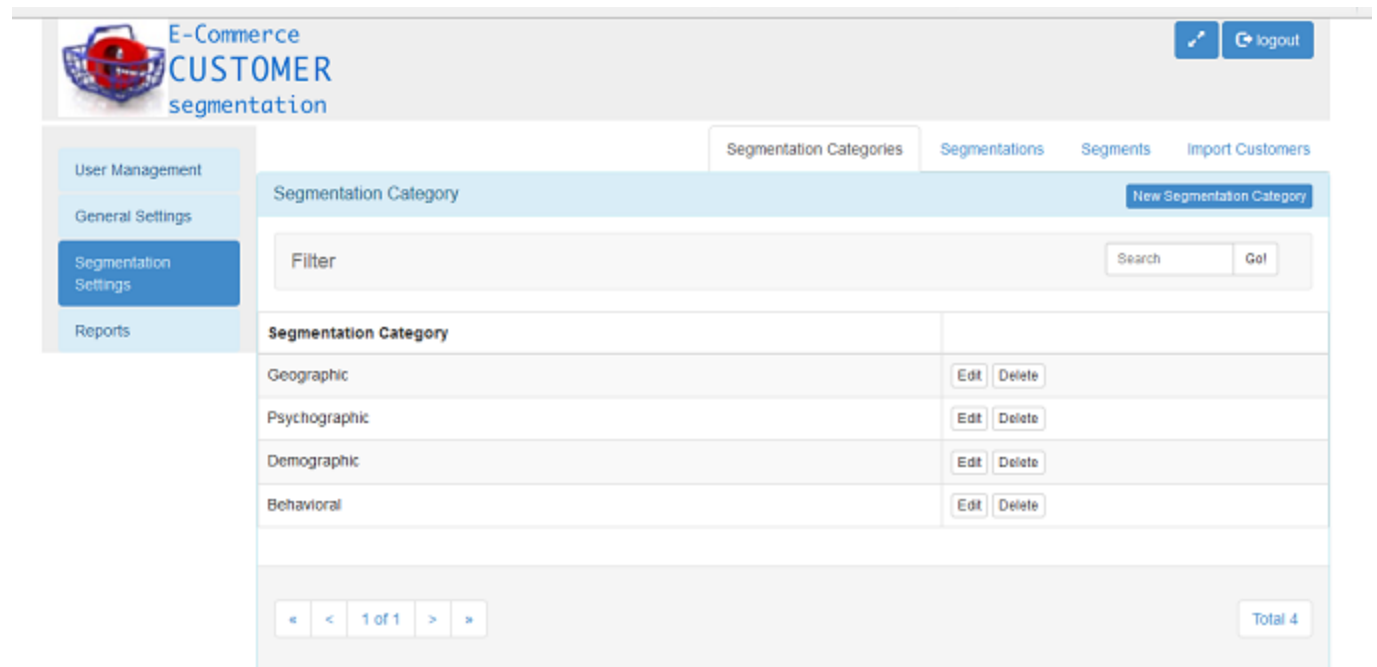


Figure 13.main segmentation categories

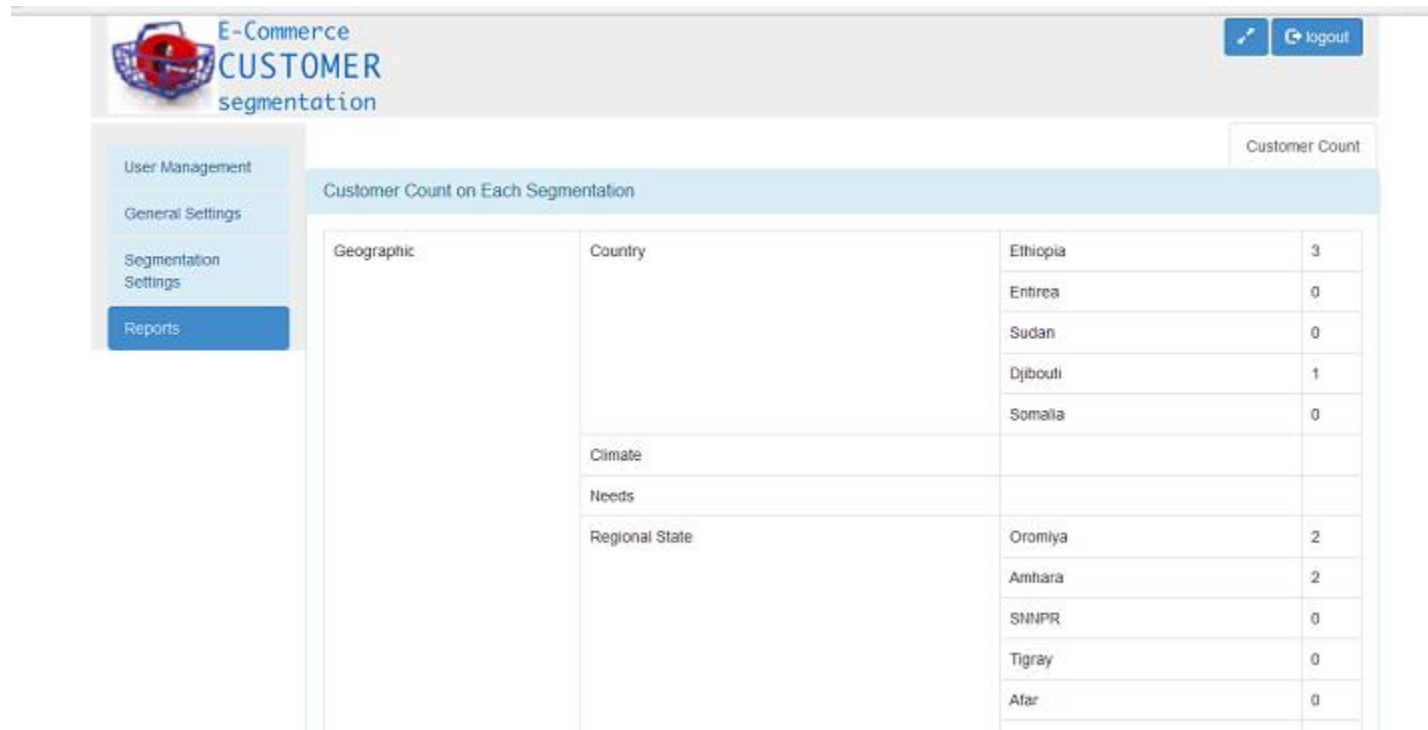


Figure 14. segmetation report

6.5.3 SAMPLE CODE

```

public function segmentation_categories() {
    $class = 'SegmentationCategory';
    if ($this->input->post('ajax')) {
        $action = $this->input->post('action');
        $this->$action($class);
    } else {
        $grid_settings = $this->$class->grid_settings;
        $data = $grid_settings;

        $data['function'] = 'segmentation_categories';

        $data['form_fields'] = $this->$class->getFormFields();
        $data['model_name'] = $this->$class->modelDisplayName();
        $data['total'] = $this->$class->countGridCollection($grid_settings);
        $students = $this->$class->gridCollection($grid_settings);
        $data['grid'] = $this->$class->toDisplayTable($students, $grid_settings['columns']);
        $this->loadView('common_views/main_panel', $data);
    }
}

```

```
}}
```

```
public function doSegmentation($f_data = null) {  
    if ($f_data) {  
        foreach ($f_data as $name => $value) {  
            $fd = explode('_', $name);  
            if ($fd[0] == 'segment') {  
                $questionnaire_segmentation = new CustomerSegmentation();  
                $questionnaire_segmentation->customer_id = $this->id;  
                $questionnaire_segmentation->segment_id = $value;  
                $questionnaire_segmentation->save();  
            }  
        }  
    }  
    foreach ($this->segmentation_fields as $field => $segmentation_name) {  
        $segment_id = $this->$field;  
        $attribute_segmentation = new CustomerSegmentation();  
        $attribute_segmentation->customer_id = $this->id;  
        $attribute_segmentation->segment_id = $segment_id;  
        $attribute_segmentation->save();  
    }  
}
```

7. Bibliography

Abraham, A., Hassanien, A. E., & Snâaésel, V. (2010). Computational social network analysis: trends, tools and research advances. Springer.

Agneeswaran, V. S. (2012). Big-Data : Theoretical, Engineering and Analytics Perspective. In Big Data Analytics (pp. 8-15). Berlin Heidelberg: Springer.

Anil, R., Dunning, T., & Friedman, E. (2012). Mahout in action. Shelter Island: Manning.

Baranov, A. 2012. Building E-commerce User's Relationship. Bulletin of the Transylvania University of Brasov, Economic Sciences. 5(54), 15-18.

Brodie, R. J., Hollebeek, L. D., Juric, B., & Ilic, A. (2011). User's Engagement: Conceptual Domain, Fundamental Propositions, and Implications for Research. Journal of Service Research, 14(3), 252–271.

Cakir, O., & Aras, M. E. (2012). A Recommendation Engine by Using Association Rules. Procedia-Social and Behavioral Sciences, 62, 452-456.

Chaffey, D. Chadwick-Ellis, F., Mayer, R., Johnston, K. (2009), Internet Marketing: Strategy, Implementation and Practice, Essex: Prentice Hall.

Chaffey, D., & Smith P.R. (2008). eMarketing eXcellence: Planning and optimizing your digital marketing. Oxford: Elsevier

Chan, C. C. H. (2005). E-commerce auction User's segmentation using a neural network model. International Journal of Applied Science and Engineering, 3(2), 101-109.

Chan, C., C., H. (2008). Intelligent value-based User's segmentation method for campaign management: A case study of automobile retailer. *Expert Systems with Applications* 34(4): 2754–2762.

Chen, H., Chiang, R., H., L., & Storey, V., C. (2012). Business Intelligence and Analytics: From Big Data to big impact, *MIS Quarterly* 36 (4), pp.1165-1188

Cheng, C.H., & Chen, Y.S. (2009). Classifying the segmentation of User's value via RFM model and RS theory. *Expert Systems with Applications*, 36(3), 4176–4184.

Chen, Y., Guozheng, Z., Dengfeng, H. & Chua, F. (2007). User's segmentation based on survival character. *Journal of Intelligent Manufacturing* 18(4): 513–517.

Chu, B., H., Tsai, M., S. & Ho, C., S. 2007. Toward a hybrid data mining model for User's retention. *Knowledge-Based Systems* 20(8): 703–718.

Dean, J., & Ghemawat, S. (2008). MapReduce: simplified data processing on large clusters. *Communications of the ACM*, 51(1), 107-113.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). "From Data Mining to Knowledge Discovery in Databases." *AI Magazine*, Vol. 17(3), pp. 37-54.

Farooqi, R., & Raza, K. (2011). A Comprehensive Study of CRM through Data Mining Techniques. *Business*, 2(i).

Fotaki, G., Gkerpini, N., Triantou, A, I., Brinkkemper, S. (2012). E-commerce User's Engagement Management. Utrecht University.

Gogia, S. (2012). The Big Deal about Big Data For User's Engagement. Forrester Research. Retrieved from: <http://www.forrester.com/The+Big+Deal+About+Big+Data+For+User's+Engagement/fulltext/-/E-RES72241>

Gupta, R., Gupta, H., & Mohania, M. (2012). Cloud Computing and Big Data Analytics: What Is New from Databases Perspective?. In Big Data Analytics(pp. 42-61). Springer Berlin Heidelberg.

Hevner, A., & Chatterjee, S. (2010). Design Research in Information Systems. 22. Springer

Hevner, A., R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information System Research. MIS Quarterly. 28(1), 75–105.

Hong, T. & Kim, E. (2010). Segmenting Users in e-commerce stores from factors that affect the User's's intention to purchase. International Conference on Information Society, 39(2), 383–388.

Hosseni, M, .B, Tarokh M. J. (2011). User's Segmentation Using CLV Elements. Journal of Service Science and Management 04(03): 284–290.

Jansen, S. M. H. (2007). User's Segmentation and User's Profiling for a Mobile Telecommunications Company Based on Usage Behavior. A Vodafone Case Study.

Kim, S.-Y., Jung, T.-S., Suh, E.-H., & Hwang, H.-S. (2006). User's segmentation and strategy development based on User's lifetime value: A case study. Expert Systems with Applications, 31(1), 101-107.

Kumar, V., Aksoy, L., Donkers, B., Venkatesan, R., Wiesel, T., & Tillmanns, S. (2010). Undervalued or Overvalued Users: Capturing Total User's Engagement Value. *Journal of Service Research*, 13(3), 297–310.

Lee, J., & Park, S. (2005). Intelligent profitable Users' segmentation system based on business intelligence tools. *Expert Systems with Applications*, 29(1), 145–152.

Leung, H., C. (2009). An Inductive Learning Approach to Market Segmentation based on User's Profile Attributes. *Asian Journal of Marketing*.

Mitra, S., Pal, S. K., & Mitra, P. (2002). Data mining in soft computing framework: a survey. *IEEE transactions on neural networks*, 13(1), 3–14.

Manyika, J., Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. 2011. "Big data : The next frontier for innovation, competition, and productivity." *McKinsey Global Institute* 364(May): 156.

Miguéis, V. L., Camanho, A. S., & Falcão e Cunha, J. (2012). User's data mining for lifestyle segmentation. *Expert Systems with Applications*, 39(10), 9359–9366.

Mollen, A., & Hugh W. (2010). Engagement, telepresence and interactivity in e-commerce consumer experience: Reconciling scholastic and managerial perspectives. *Journal of Business Research* 63(9-10): 919–925.

Ngai, E.W.T., Xiu, L., Chau, D.C.K., (2009) Application of data mining techniques in User's Relationship Management: A literature review and classification.

O'Reilly (2012). *Big Data Now*: 2012 Edition.

Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1–135.

Peterson, E. T. (2004). *Web analytics demystified: A marketer's guide to understanding how your web site affects your business*. Celilo Group Media & CafePress.

Peterson, E.T. (2006). *The Big Book of Key Performance Indicators*. Web Analytics Demystified.

Pillai, J., & Vyas, O. P. (2012). CSHURI – Modified HURI algorithm for User's Segmentation and Transaction Profitability, 2(2), 79–89.

Russom, P. (2011). *Big Data Analytics*. TDWI Best Practices Report, Fourth Quarter

Sathi, A. (2012). *Big Data Analytics: Disruptive Technologies for Changing the Game*. USA: MC Press E-commerce.

Sengamedu, S. H. (2012). Scalable Analytics–Algorithms and Systems. In *Big Data Analytics* (pp. 1-7). Springer Berlin Heidelberg.

Shukla, A., Wexler, M., Singh, V., Kan, E., Joshi, D., & Lestiyo, I. (2012). Micro-bucket testing for page optimization. U.S. Patent No. 8,126,930. Washington, DC: U.S. Patent and Trademark Office.

Srinivasa, S., Bhatnagar, V. (2012) *Big Data Analytics*. Berlin: Springer

Shaw, M. J., Subramaniam, C., Tan, G. W., & Welge, M. E. (2001). Knowledge management and data mining for marketing. *Decision Support Systems*, 31(1), 127–137.

Stroud, Dick. (2006). User's Intelligence. *Journal of Direct Data and Digital Marketing Practice* 7(3): 286–288.

Sun, H., & Heller, P.(2012). *Oracle Information Architecture: An Architect's Guide to Big Data. An Oracle White Paper in Enterprise Architecture.*

Tsiptsis, K., & Chorianopoulos, A. (2009). *Data Mining Techniques in CRM: Inside User's Segmentation.* Wiley.

Turban E., Sharda, R., Delen D., & King, D. (2010). *Business Intelligence: A Managerial Approach.*(2nd ed.) New Jersey: Prentice Hall

Van Doorn, J.,Lemon, K.,N., Mittal, V., Nass, S., Pick, D., Pirner, P., & Verhoef, P., C. (2010). User's Engagement Behavior: Theoretical Foundations and Research Directions. *Journal of Service Research* 13(3): 253–266.

Weerd, I. van de, Brinkkemper, S. (2008). Meta-modeling for situational analysis and design methods. In M.R. Syed and S.N. Syed (Eds.), *Handbook of Research on Modern Systems Analysis and Design Technologies and Applications* (pp. 38-58). Hershey: Idea Group Publishing.

Winer, R. S. (2001). A framework for User's relationship management. *California management review*, 43(4), 89-105.

Woo, J., Bae, S., & Park, S. (2005). Visualization method for User's targeting using User's map. *Expert Systems with Applications*, 28(4), 763–772.

Wu, R.S & Chou P.H. (2011). User's segmentation of multiple category data in e-commerce using a soft-clustering approach. *Electronic Commerce Research and Applications* 10(3): 331–341

Wymbs, C. (2011). Digital Marketing: The Time for a New “Academic Major” Has Arrived. *Journal of Marketing Education*, 33(1), 93–106.

Ye, L., Qiuru, C., Haixu, X., Yijun, L., & Guangping, Z. (2013). User's Segmentation for Telecom with the k-means Clustering Method. *Information Technology Journal*, 12(3), 409-413.

Zaslavsky, A., Perera, C., & Georgakopoulos, D. (2012). Sensing as a Service and Big Data. *Proceedings of the International Conference of Advances in Cloud Computing*, 8: 21-29

Zhou, Y., Wilkinson, D., Schreiber, R., & Pan, R. (2008). Large-scale parallel collaborative filtering for the 86etflix prize. In *Algorithmic Aspects in Information and Management* (pp. 337-348). Springer Berlin Heidel