New Product Development Based on Customer Knowledge Management

Zeinab Rezvani

Luleå University of Technology
Master Thesis, Continuation Courses
Marketing and e-commerce
Department of Business Administration and Social Sciences
Division of Industrial marketing and e-commerce
Abstract

Human beings have always been tried to make new, innovative things to benefit from. The new things have been improved from a wheel to a smart car and the benefits have been improved from survival to survival plus fitness, happiness, sense of belonging, security and etc. Companies which have been following the human nature, and thinking of new, innovative products and consumers’ benefits cohesively, could gain competitive advantages. But the process of new product development has always been costly and due to the high failure rate it is risky. Academics and practitioners alike agree that one of the new product development failure reasons is consumers’ nonacceptance. So they try to enhance the new product development (NPD) process to reduce the consumers’ nonacceptance. One of the most famous tools used is Knowledge Management (KM). KM focuses on employees’ knowledge but has little systematic attention to the customers. In order to overcome the KM limitation, the concept of Customer Knowledge Management (CKM) introduced. CKM by combining KM and customer relationship management (CRM), more clearly extract knowledge ‘for’ customers, knowledge ‘about’ customers, and knowledge ‘from’ customers, so that a more beneficial product can be delivered to the right group of customers, to prevent product failure and to ensure commercial success.

In this research, review on NPD, KM and CKM are provided and new product development based on customer knowledge management has been implemented. To develop new products based on customer knowledge management framework, the customers’ knowledge should be elicited and converted to a pattern of consumers’ need towards the products attributes. The need patterns and consumers’ characteristics have to be segmented to be more meaningful for marketers to target and communicate. This reduces the customer fuzziness dimension in the NPD process and elevates the likelihood of success.

For eliciting customer knowledge, Means-end chain theory has been adapted, and with data mining clustering algorithms, the consumers’ segments has been derived based on their means-end chains toward the product (Malt Beverage or None Alcoholic Beer - N.A.B).

Data collection is done with a survey of 660 university people. The questionnaire consists of two sections: 1-measuring means-end chains with Association Pattern Technique (APT), 2-questions about buying behavior and demographic characteristics to better profile the segments. Data analysis of means-end chains with clustering algorithms, results in three segments of consumers, three groups of preferred product attributes and
the proper communication message for each segment. Data analysis of second part of the questionnaire, help to profile and describe the segments found.

The main objectives of this research include finding and implementing systematic approaches to extract customer knowledge, and finding frameworks for more involvement of customer and using customer knowledge in the idea generation phase of NPD process.

The outcome of this research is a practical framework for “idea generation phase of new product development process based on customer knowledge”. In continue, the mentioned framework implemented in a part of Iran N.A.B market and result in segmenting and profiling this market. Also, the critical new product attributes and bases of communication message and promotion campaigns extracted. We have also contribution, in implementing text mining techniques for clustering algorithms.

The framework proposed in this dissertation, can be used in any consumer industry in Iran. Having a framework for precisely delineating the process of customer knowledge management for new product development can lead Iranian companies and consumers to more satisfaction.

**Keywords:** Customer Knowledge Management, New product development, Means-end Chain Theory, Data Mining, Clustering, Segmentation, Latent Semantic Indexing
Table of Content

Abstract A
Table of Content B
List of Figures C
List of Tables C

Chapter 1: Introduction ................................................................. 1
  1-1. Background ................................................................. 2
  1-2. Problem Description ...................................................... 3
  1-3. Importance of Topic ....................................................... 5
  1-4. Research purpose and question identification ....................... 5
  1-5. Delimitation ..................................................................... 5
  1-6. Terms & Phrases ............................................................ 6
  1-7. Abbreviations .................................................................. 6
  1-8. Overview of the entire thesis ............................................. 6

Chapter 2: Literature Review ............................................................ 7
  2-1. Introduction to New Product Development ......................... 8
  2-2. Concurrent New Product Development (CNPD) .................. 10
  2-3. Integrated PVQ and NPD Model ...................................... 11
  2-4. New Product Development Dynamic Model (NPDDM) ...... 12
  2-5. Generic Model of New Process Design and Development ...... 14
  2-6. Introduction to Knowledge Management ........................ 16
  2-7. What Is Knowledge? ...................................................... 17
  2-8. What is Knowledge Management? ................................... 19
  2-9. KM in NPD .................................................................... 20
  2-10. Introduction to CKM .................................................... 21
  2-11. A model for CKM .......................................................... 22
  2-12. E-CKM a model for NPD based on CKM ....................... 24
  2-13. Market Segmentation, Main Component of E-CKM ......... 27
  2-15. Major Segmentation Variables for Consumer Markets ....... 30
  2-16. Segmentation by Data Mining ....................................... 30

Chapter 3: Research Methodology .................................................... 35
  3-1. Introduction ................................................................. 36
  3-2. Research design ............................................................ 36
  3-3. Designing the Qualitative Research Process ..................... 38
    3-3-1. Means-end Theory .................................................. 39
    3-3-2. Revealing the Means-end chains .............................. 41
  3-4. Designing the Quantitative Research Process ................. 43
    3-4-1. Designing the Questionnaire ................................... 44
    3-4-2. Data Collection ...................................................... 45
    3-4-3. Data Analysis .......................................................... 46
## List of Tables

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1-1: NPD Problems</td>
<td>3</td>
</tr>
<tr>
<td>Table 2-1: Main characteristics of the managed versus the entrepreneurial economy</td>
<td>16</td>
</tr>
<tr>
<td>Table 2-2: Four kinds of interaction between tacit &amp; explicit knowledge</td>
<td>18</td>
</tr>
<tr>
<td>Table 2-3: Different kinds of knowledge being created</td>
<td>19</td>
</tr>
<tr>
<td>Table 2-4: CKM versus Knowledge Management &amp; Customer Relationship Management</td>
<td>22</td>
</tr>
<tr>
<td>Table 5-1. Segments’ characteristics</td>
<td>85</td>
</tr>
<tr>
<td>Table 5-2. Product and promotion for each segment</td>
<td>86</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1-1. The fuzzy front end and success of new product development: a causal model</td>
<td>4</td>
</tr>
<tr>
<td>Figure 2-1. New Product characterizations</td>
<td>8</td>
</tr>
<tr>
<td>Figure 2-2. Traditional NPD Process</td>
<td>9</td>
</tr>
<tr>
<td>Figure 2-3. Use of tools and techniques for CNPD applications</td>
<td>10</td>
</tr>
<tr>
<td>Figure 2-4. Integrated NPD and PVQ process</td>
<td>11</td>
</tr>
<tr>
<td>Figure 2-5. Innovation-Driven Learning in NPD</td>
<td>12</td>
</tr>
<tr>
<td>Figure 2-6: Causal loop diagram of the dynamics of NPD</td>
<td>13</td>
</tr>
<tr>
<td>Figure 2-7 Generic model of NPD</td>
<td>15</td>
</tr>
<tr>
<td>Figure 2-8. The recursive relations between data, information and knowledge</td>
<td>17</td>
</tr>
<tr>
<td>Figure 2-9. Explicit knowledge is just the tip of the iceberg.</td>
<td>18</td>
</tr>
<tr>
<td>Figure 2-10. The stages of knowledge management</td>
<td>20</td>
</tr>
<tr>
<td>Figure 2-11. The knowledge value chain</td>
<td>23</td>
</tr>
<tr>
<td>Figure 2-12. CKM Process Model</td>
<td>24</td>
</tr>
<tr>
<td>Figure 2-13. The E-CKM model as applied in customer knowledge management for innovative product development.</td>
<td>25</td>
</tr>
<tr>
<td>Figure 2-14. Data mining scheme</td>
<td>30</td>
</tr>
<tr>
<td>Figure 3-1. The research design</td>
<td>37</td>
</tr>
<tr>
<td>Figure 3-2. General research process</td>
<td>38</td>
</tr>
<tr>
<td>Figure 3-3. The qualitative research process</td>
<td>38</td>
</tr>
<tr>
<td>Figure 3-4. Means-end chain examples</td>
<td>39</td>
</tr>
<tr>
<td>Figure 3-5. Ideal Means-end Chain</td>
<td>40</td>
</tr>
<tr>
<td>Figure 3-6. The quantitative research process</td>
<td>43</td>
</tr>
<tr>
<td>Figure 3-7. General process of data mining</td>
<td>46</td>
</tr>
<tr>
<td>Figure 4-1. AC (Attributes of Product-Consequences of consumption) Matrix</td>
<td>55</td>
</tr>
</tbody>
</table>
Figure 5-22. Regular Newspaper Reading?  81
Figure 5-23. Reading Family Magazines?  81
Figure 5-24. Education  82
Figure 5-25. Job  82
Figure 5-26. Monthly Cost of living of the family  83
Figure 5-27. Number of family members  83
Figure 5-28. Marriage Status  84
Figure 5-29. Number of Kids  84
Figure 6-1. Our proposed CKM framework for NPD  92
Chapter 1: Introduction

1-1. Background
1-2. Problem Description
1-3. Importance of Topic
1-4. Research purpose and question identification
1-5. Delimitation
1-6. Terms & Phrases
1-7. Abbreviations
1-8. Overview of the entire thesis
1-1. Background
New product development process allows firms to deal with progressively intensive competition when facing challenges from a rapidly changing market condition. Companies try to gain sustainable competitive advantages with continues innovation. These efforts are critical for the companies because introducing a new product to the market is expensive and challenging due to acquiring new knowledge and technologies and implementing new processes. At the other side, many firms have tried hard to enhance their customer-relating capability with knowledge management tools. But successful implementations of such initiatives are still exceptional.
Knowledge Management has long been known to managers and researchers and its importance in the innovation has been properly recognized (Grover and Davenport, 2001). Many works shows the effectiveness of KM concepts and result of its application in firm functional areas including NPD. According to the new insights of KM, a creative knowledge worker can contribute to face the problems that need new kinds of decision, the situations that require innovative approaches, and the relationships that can be discovered in the more complex markets where firms are operating (Carneiro, 2000).
Cooper (2003) declared flourishing new product development requires effective strategies for reducing risk. Knowledge management could be beneficial in risk reduction, e.g. by gathering and processing significant information and summarize knowledge from a variety of internal and external sources. Knowledge management works as a coordinating mechanism. Empirical evidence supports that a firm with a knowledge management capability will use resources more efficiently and so will be more innovative and efficient. Knowledge management components including -knowledge acquisition, knowledge dissemination and responsiveness to knowledge- were found to have a straight effect on innovation (Darroch, 2005).
The value of knowledge management in the innovation of product technology has been identified; yet, the potential for customer knowledge management has not been studied deeply (Soo et al., 2002).
Meanwhile, Customer relationship management (CRM) is a modern management tool. CRM manages the relationship with customers by employing up-to-date information technology (IT) such as on-line data analysis, data-mining and database management in order to know, communicate with, and to catch the attention of customers. Its goal is to please and retain customers (Dyche, 2002). Growing the productivity of knowledge work and managing customers’ knowledge in order to understand their needs, enables a firm to gain a competitive advantage in the market.
Lately, the ‘customer knowledge management (CKM)’ model has attracted much attention by the combining of both the technology-driven and data-oriented approaches in CRM and the people-oriented approach in KM, with a view to utilize their synergy potential (Garcia-Murillo and Annabi, 2002). The prospect result from this effort is to more eloquently describe knowledge ‘for’ customers, knowledge ‘about’ customers, and knowledge ‘from’ customers, so that a more beneficial product can be delivered to the right group of customers, therefore new product failure risk would be reduced.
1-2. Problem Description

Even though vast investments of firms on new product development, many of them face with weak market acceptance and fail (Su et al, 2006).

New product development is a costly and risky process. And in today's competitive business environment, firms have to introduce new products which could survive and make profit.

Research on NPD has recognized a number of factors that influence NPD process: technology, product characteristics, project structure, team member characteristics and patterns, team processes, organizational context, and external environment, and internal resources needed like budget, knowledge and skills (Cooper, 2003). We can conclude that the more difficult the project in terms of scale, new technology, and complexity, the more risky it is mismatches in the team, organization, and external environment, and the greater the need for knowledge acquisition and development for the project to thrive.

A significant challenge in new product development projects is how to acquire knowledge and manage sources of ambiguity in order to reduce the risk of failure of either the process or the resulting product. To clarify it see table 1-1. As shown in the table, the new product can fail due to process fail or products fail and also due to its intrinsic variable or extrinsic variables

<table>
<thead>
<tr>
<th>Problems</th>
<th>Product failure</th>
<th>Process failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrinsic variables</td>
<td>A: Does not meet performance, reliability, safety, or other requirements for the proposed environment</td>
<td>B: Violating resource constraints (e.g. cost, schedule)</td>
</tr>
<tr>
<td>Extrinsic variables</td>
<td>C: Unfavorable reception in market; regulatory change</td>
<td>D: Competition develops product first</td>
</tr>
</tbody>
</table>

Adopted from: (Cooper, 2003)

Each of these problems needs considerable efforts in terms of time and profession. According to the time allocated to master thesis and the major in marketing, the problem of NPD projects in the case of market/customer acceptance that locates in cell C, conjunction of product and extrinsic variables, have been chosen. Customers are outside the firm and they should accept the product features but how could we know their acceptance before introducing a new product to them? How can we know what they want from our products? What Is the systematic solution?

Consider this problem from another point of view: Most projects do not fail at the end; they fail at the beginning (Zhang & Doll, 2001). The product development success depends on the performance of the front-end activities. See figure 1-1. Because these front-end activities are the final gate before the team decides to invest in designing and producing the products, they need to be well managed. Otherwise, both time and money may be wasted in producing the wrong Products. The front-end activities include idea generation, assessment of market, technology and competition and product definition, project justification and action plan (Zhang & Doll, 2001). At these stages, project team may not be sure what
customers they really want, what competitors are producing, and what product and process technologies should be utilize. Environmental ambiguity related to market changes, emerging technological developments, and the evolving competitive situation is defined as “front-end fuzziness”.

Product development managers perceive at least three sources of uncertainties in front end activities:

1. The customer’s requirements;
2. The nature of competition; and
3. The changing technology.

In this research, the focus is on one specific source of uncertainty: customers. The debate is that modern tools, CRM and KM techniques and processes, have some constraints. In brief, CRM acquire customer knowledge but it doesn’t propose how to apply it and KM is about sharing knowledge but its focus is on intrinsic variables. CKM with combining these two, demonstrates how and what elicit from outside the firm, specifically customers, and how to apply it as an input to the new product development process.
1-3. Importance of Topic
By developing a framework to involve customers in the new product development process to assure that customers will accept the new product, the following advantages will be gained:
- The risk that inherently exists in NPD will be reduced.
- Introducing new successful products to market will improve firms ROI, one of the most important performance indicators.
- The process of NPD will be more structured.
- The consumers will be more satisfied with the new introduced product.
From the manager’s viewpoint each of these advantages are valuable.

1-4. Research purpose and question identification
The main purposes of this research include:
- Introducing new models of NPD process
- Finding and implementing systematic approaches to extract customer knowledge
- Finding a framework for more involvement of customer and using customer knowledge in NPD process.
- Implementing the proposed framework to verify its applicability
- Promoting usage of mixed quantitative and qualitative marketing methods

To reach these purposes, a framework should be developed to provide an answer to our main question:

*RQ- How firms can understand preferences and requirements of customers in order to have more successful new product development?*

In detail, what are the motives behind consumer’s preferences? Why they accept some products and why they do not accept some other products? How can we identify customers’ requirements to be considered in NPD process? How customers can participate in NPD Process?

Here we try to introduce and improvise a recently developed concept: Customer Knowledge Management (CKM), and to provide a framework based on CKM concepts showing the method for acquiring customer knowledge, so that the product will meet the customers’ needs, in order to secure market acceptance.

1-5. Delimitation
New product development process has many dimensions and involves almost all parts of the firm. In this study, the focus is on:
- Customer fuzziness of NPD fuzzy front end, not technology fuzziness or competition fuzziness (based on fuzzy front end model of Zhang and Doll, 2001)
- Opportunity identification stage of NPD process not development or optimization (product testing) stages (based on stages of new product development of Van Kleef, 2006)
- NPD failures originated from extrinsic and product related variable in NPD not Intrinsic or process related variable (based on NPD problems model of Cooper, 2003)

The conceptual framework can be conducted in many different industries. Due to intensive competition, low variety of products, access to resource for data collection from consumers, food industry was selected which is naturally a B2C industry. **The product in this research is Non Alcoholic Beer, N.A.B or Malt Beverage.**

### 1-6. Terms & Phrases

**New Product Development:** process of transforming customer needs to product characteristics and production of the product based on those characteristics.

**Customer Knowledge Management:** a new developed concept combining KM and CRM, specially developed to elicit and manage knowledge for customers, about customers and from customers.

**Data Mining:** The sum of methods for extraction of hidden predictive information from large databases.

**Means-end Chain Theory:** The sequence of attributes–consequences–values is called a *means–end* chain. Means–end chains are a way to describe how consumers perceive products. Attributes are the tangible aspect of products. Benefits are the positive functional consequences and values are the ends that a person wants to achieve in his/her life.

### 1-7. Abbreviations

NPD: New Product Development  
KM: Knowledge Management  
CKM: Customer Knowledge Management  
DM: Data Mining  
N.A.B: None Alcoholic Beverage/Beer  
APT: Association Pattern Technique  
LSI: Latent Semantic Indexing

### 1-8. Overview of the entire thesis

CKM is a combining concept that relates concepts including Segmentation, KM, CRM and NPD. Regarding our pre-mentioned problem, at first we review the literature of NPD, KM, CKM, and segmentation with clustering in chapter two. In the third chapter, the methodology of research is provided which include stages of data collection and data analysis and description of selected quantitative and qualitative methods. In chapter four, empirical data which used in analysis and the questionnaire used to collect data is presented. In chapter five, results are shown in graphical diagrams and three identified clusters will be described in detail. The last chapter, chapter six, is the conclusion of research including our proposed framework, limitations of the research, implications and future research areas.
Chapter 2:

Literature Review

2-1. Introduction to New Product Development
2-2. Concurrent New Product Development (CNPD)
2-3. Integrated PVQ and NPD Model
2-4. Model of Innovation-Driven Learning in New Product Development
2-5. New Product Development Dynamic Model (NPDDM)
2-6. Generic model of NPD
2-7. Introduction to Knowledge Management
2-8. What Is Knowledge?
2-9. What is Knowledge Management?
2-10. KM in NPD
2-11. Introduction to Customer knowledge Management
2-12. A Model for Customer Knowledge Management
2-13. E-CKM Model, a model for NPD based on CKM
2-14. Market Segmentation, A Core Component of E-CKM
2-15. Process of Market Segmentation
2-16. Major Segmentation Variables for Consumer Markets
2-17. Segmentation by Data Mining
2-1. Introduction to New Product Development

New product development is an ordered and determined set of tasks and steps that describe the method by which a company repeatedly converts undeveloped ideas into commercial products or services. There are distinct categories of new products (PDMA Handbook). Some products are new to the market, some are new to the company, and some are totally new and make totally new markets. From another point of view, some new product concepts are minor modifications of existing products while some are completely innovative to the company. These differences are displayed in the following figure.

Source: (PDMA Handbook)

Figure 2-1: New Product characterizations

New product development is one of the most influential but complex activities in business (Clark and Wheelwright, 1995). Business managers and marketing academics alike agree that a fundamental factor of an organization’s long-term survival is success in new product development. The developments of excellent products not only opens new markets and attracts new customers, but also influence existing assets and expand an organization’s capabilities (Zhan, 1998). The vast literature on new product success and failure reveal some factors that are critical to be a winner at this game (Snelson and Hart, 1991):

(1) The need for interdisciplinary input: usually the tasks involved fall within the field of different functions, based on professional education, training and certification. To be successful these various professions should actively and effectively collaborate through the NPD process.

(2) The need for quality inputs to the process: it means both technical and marketing information, two building blocks of NPD, have to be accurate and timely, and must be constantly revised during the development.
(3) The need for speed in the process: The NPD process has to be managed in such a way as to be quick enough to capitalize on the new product opportunity before competitors do so.

A formal, but conventional new product screening model was created, with its ratings, checkpoints, etc. The following seven points summarize the model:

(1) Strategy;
(2) Idea generation;
(3) Assessment;
(4) Business plans;
(5) Development;
(6) Testing;
(7) Commercialization.

Figure 2-2 shows a traditional NPD process. In this model new product development is disconnect from product validation and qualification.

These traditional models are valid where uses are known, and there are standard homogenous market segments. Traditional models are unable to provide
multidisciplinary, high quality, and quick inputs for NPD process (Coates and Harry Robinson, 1995).
For today situation with modern products and for diverse dynamic markets, traditional NPD process is not enough, it is impossible to reach the development target without using new framework and method. Next sections propose some of new frameworks and methods reviewed in the literature.

2-2. Concurrent New Product Development (CNPD)
One way to change an organization is to control the scheduling of activities in the NPD process so that:

(1) activities that were once carried out in series are now carried out in parallel – e.g. product engineering and production engineering phases;
(2) People are involved throughout the life of the development project working in cross-functional teams.

The above way of working is “concurrent new product development” (CNPD) (Maylor and Gosling, 1998).

Some tools have been developed to apply the concept of CNPD.
Figure 2-3 shows the top eight techniques and their level of usage in firms. Responds are base on a scale of 0 (no usage) to 10 (extensive usage) for each technique.

![Bar chart showing the level of usage for each technique.](chart.png)

Source: (Maylor and Gosling, 1998)

As can be seen, project management is the most used technique. Involvement of key suppliers ranks second. Use of multi-functional teams was highly rated – as was employing design for manufacture and assembly. Customer involvement (the active participation of customers in the design process) was still in the top group of tools and techniques, though not as highly rated as one might think from the literature on quality management. Similarly, regardless of much exposure, the use of quality function deployment (QFD) scores relatively low.

Since people are taken out of their functional duties to work in multi-disciplinary teams, the most difficulty in CNPD implementation is resistance to functional change (Maylor and Gosling, 1998).
2-3. Integrated PVQ and NPD Model

PVQ stands for product validation and qualification which means making sure that product has a particular level of quality and reliability.

In this integrated NPD model shown in Figure 2-4, there are “Gates” between stages 1, 2, 3 and 4. Each “Gate” allows the NPD team to meet each other and review each stage. A checklist will be used to review if all tasks or activities were done before taking the next stage actions. (Aw, 2005)

Stage 1: Opportunity assessment stage is where new product or concept will be initiated which is named as project initialization. At this stage a draft business case will be written and then technical assessment will be carried out. This stage usually involves the marketing and business development people.

Stage 2: Upon completion of stage 1, a product development team from various departments will be arranged and stage 2 begins. The product development team usually comprises of members from design, quality and reliability, product engineering, test engineering, manufacturing, planning, marketing, and after sale service departments. A project plan and budget will then be drafted after the completion of the product’s specification. Finally, a project budget will be presented to the management team to get their approval.

Stage 3: At this stage, the actual product design activities will begin to produce a few working prototypes by the pilot production line. At this stage the various activities are as mechanical design, electronic design, software design, manufacturing review, and material review.

Stage 4: At this stage, samples and also as PVQ’s test samples will be produced. At this stage, several evaluations and reviews will be carried out.

Stage 5: This stage will permit the product to be mass-produced. Therefore, it is essential that the product has all the necessary customers’ requirements.
Gates: These gates are in place to ensure that the product development team will meet and review the checklist to ensure all requirements or condition for each stage were completed and approved before allowed to continue to the next stage. There are a total of four gates; namely Gate A, B, C and D which are used to review stage 1, 2, 3 and 4, correspondingly. These are essential to ensure that all identified problems are fixed before proceeding to the next stage. This integrated new product development process provides greater quality and consistency sensitivity to product development team and also among the employees. Moreover, it makes the development cycle shorter (Aw, 2005).

2-4. **New Product Development Dynamic Model (NPDDM)**

In NPDD Model the principles of system dynamics (SD) methodology have been used in developing causal loop diagrams, flow diagrams, and the governing equations. The SD model has been used to analyze the factors such as change in scope since the development of new technology would require such change. See Figure 2-6.

Under the ideal situation, the project begins with initial scope of work. The hiring of potential project staffs (PSs) starts according to the divergence of workforce. The PSs are trained to communicate know-how. In the course of the project, rework occurs because of the quality standards, human errors, and work obsolescence. This rework is recognized and a part of capable PSs are assigned to complete the rework. Customer satisfaction increases when the project is completed on time, on budget, and with the expected quality. In reality, change in scope occurs due to changes in business such as change in technology or in market situation. Customers often demand those changes to be incorporated in the project when they occur (Lewlyn et. al., 2006).
The results indicate that effective knowledge management (KM) comprises the main controlling factor that has a significant control over project dynamics.
2-5. **Generic model of NPD**

New product development (NPD) can initiate from new technology or new market opportunities (Eliashberg, Lilien and Rao, 1997 cited on van Kleef, 2006). But independent of where opportunities start off, when it comes to successful new products, the consumer has the critical vote (Cooper and Kleinschmidt, 1987; Brown and Eisenhardt, 1995 cited on van Kleef, 2006). Thus, to develop successful new products, companies should gain a deep understanding of ‘the voice of the consumer’ (van Kleef, 2006). Consumer research can be fulfilled during each of the basic stages of the NPD process:

1. Opportunity identification,
2. Development,
3. Testing, and

Despite the importance of the customer research in later stages, it is increasingly recognized that successful NPD strongly depends on the quality of the opportunity identification stage and consumer research in this stage (Cooper, 1988; 1998; McGuinness and Conway, 1989 cited on van Kleef, 2006). The aim of this stage is to search on new areas of opportunities, which usually involve the unmet needs and wants of consumers (van Kleef, 2006). Consumer research is difficult at this stage because there are uncertainty about what to ask consumers. Most practitioners argue that asking consumers what they want is worthless, because they do not know what they want (Ulwick, 2002 cited on van Kleef, 2006). Even though consumers may not always be able to express their wants, it is important to understand how they perceive products, how their needs are formed and how they make product choices Based on them. In this way, it helps to avoid working on a new product that has a low chance of success (Rochford, 1991 cited on van Kleef, 2006). Additionally, it guards against potential winning product concepts being overlooked. As a result, performing consumer research in this stage is inexpensive compared to the risk of product failure (van Kleef, 2006).

Figure 2-7 shows the four typical major stages in NPD and appropriate representative consumer research methods for each stage (van Kleef, 2006).
Despite the significant attention paid to methods like Quality Function Deployment (QFD) and product testing methods, analysis of strengths and weaknesses of consumer research methods for opportunity identification stage has gotten only little attention (van Kleef, 2006).

In this research, we focus in the phase idea generation or opportunity identification with the specific attention to understanding consumers’ needs.
2-6. **Introduction to Knowledge Management**

As our economy has gradually changed from a managed economy into an entrepreneurial economy (Audretsch and Thurik, 1997 cited on Beijerse, 1999), see table 2-1, it is more commonly referred to a knowledge-based economy. Thus, one of the main assets of companies has become their supply of knowledge. Where more traditional economies focused on land, labor and capital as their main production factors, in this knowledge-based economy knowledge is becoming the main production part on which competitive advantage gains (Beijerse, 1999). Both business and academic communities believe that by supplying knowledge, an organization can keep up long-term competitive advantages (Bhatt, 2001).

| Table 2-.1: Main characteristics of the managed versus the entrepreneurial economy |
|-------------------------------------------------|---------------------------------|
| Managed economy          | Entrepreneurial economy         |
| Input                    | Land, labour and capital        | Knowledge                     |
| Output                   | Manufactured products           | Knowledge                     |
| Features                 | Certainty                       | Uncertainty                   |
|                        | Easy to shift around the world   | Costly to transact            |
|                        | Symmetric across people          | Asymmetric across people      |
| Government               | Control                         | Enabling                      |
| Economy                  | Economies of scale              | Economies of diversity        |
| Industrial structure     | Large corporations              | Small enterprises             |

Source: (Beijerse, 1999)

In the knowledge-based economy it is clear that the challenges are making information productive (Drucker, nd cited by Beijerse, 1999), handling the ambiguity of knowledge in a globalized world (Audretsch and Thurik, nd cited by Beijerse, 1999) and coming to terms with the growing importance of consumers and their individual needs and wants (Jacobs, nd cited by Beijerse, 1999). In knowledge-based economy, organizations have to deal with issues like:

- A rising complexity of products and processes;
- A larger pool of related knowledge, both technical and non-technical;
- Increasing competition because of shorter product life cycles,
- An increased focus on the core competencies of the enterprise (Kalff et al., 1996 cited by Beijerse, 1999).

Knowledge management can enable companies to deal with the complexities of the knowledge-based economy. (Beijerse, 1999) Moreover, organizations can improve efficiency; improve the market position by operating more intelligently on
the market; **Optimize the interaction between products development and marketing**; and more and more advantages.

## 2-7. What Is Knowledge?
Defining data, information, and knowledge is not easy. Only through external means or from a user's point of view, can one differentiate between data, information, and knowledge. In general, data are considered as raw facts, information is regarded as an organized set of data, and knowledge is perceived as meaningful information. Data and information are distinguished based on their "organization", and information and knowledge are differentiated based on the "interpretation" (Bhatt, 2001).

Therefore, we argue that knowledge is an organized combination of data, incorporated with a set of rules, procedures, and operations learnt through experience. From a point of view, knowledge is a "meaning" made by the mind (Marakas, 1999 cited by Bhatt, 2001). Without meaning, knowledge is information or data. It is only throughout meaning that information becomes knowledge. Thus, the difference between information and knowledge depends on users' perspectives (Marakas, 1999 cited by Bhatt, 2001).

*Tacit and explicit knowledge*

Another way of defining knowledge is to make a difference between "tacit" and "explicit" knowledge (Polyani, 1966 cited by Martensson, 2000):

- Explicit knowledge is documented and public; structured, externalized, and mindful. Explicit knowledge is what can be captured and shared through information technology.
Tacit knowledge lives in the human mind, behavior, and perception. Tacit knowledge develops from people's interactions and requires skill and practice. Tacit is personal, undocumented, dynamically created and derived, internalized and experience-based. Knowledge is the product of the communication of explicit and tacit knowledge (Nonaka and Takeuchi, nd cited by Martensson, 2000). The process of creating knowledge starts with people sharing their internal tacit knowledge by socializing with others or by capturing it in digital or analogue form. Other people then internalize the shared knowledge, and this process produces new knowledge. These people have the newly created knowledge, then share this knowledge with other people and the process begins again. This process could be expressed as innovation (Hibbard, 1997 cited by Martensson, 2000)

![Image of iceberg with explicit and tacit knowledge]

Figure 2-9: Explicit knowledge is just the tip of the iceberg.

The interaction between tacit and explicit knowledge can run in four different directions (Beijerse, 1999) see Table 2-2.

1. Socialization: The exchange of experiences though personal knowledge is being created in the form of mental models. Examples of situations are master-fellow-relationships, on-the-job-training, trial-and-error-policy, constructive brainstorm sessions, etc.

<table>
<thead>
<tr>
<th>Table 2-2: Four kinds of interaction between tacit &amp; explicit knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Into tacit knowledge</strong></td>
</tr>
<tr>
<td>From tacit knowledge</td>
</tr>
<tr>
<td>From explicit knowledge</td>
</tr>
</tbody>
</table>

Source: (Nonaka & Takeuchi, 1995 cited on Beijerse, 1999)
(2) Externalization: Personal or tacit knowledge is made explicit in the form of descriptions, analogies, hypotheses and models. The externalization has been found the key process in knowledge renovation because at this stage from tacit knowledge new and explicit designs are created (Nonaka & Takeuchi, 1995 cited by Beijerse, 1999).

(3) Combination: New knowledge can also be created through the restructuring of on hand information by categorizing, adding, combining and categorizing explicit knowledge. Combination is the kind of knowledge creation which we typically encounter in education and training. Examples of combination are knowledge and information systems.

(4) Internalization: it is process in which explicit knowledge becomes part of tacit knowledge. This happens through learning by- doing, and documented knowledge can play a supportive role in this process. Internalization happens when experienced managers or technicians give lectures or when authors decide to write the biography of an entrepreneur or enterprise. Four different kinds of knowledge being created in the four kinds of interaction between tacit and explicit knowledge, as shown in the table below:

<table>
<thead>
<tr>
<th>From tacit knowledge</th>
<th>Into tacit knowledge</th>
<th>From explicit knowledge</th>
<th>Into explicit knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-Socialize sympathy</td>
<td>2-Externalize conceptual knowledge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-Combine system knowledge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-Internalize operational knowledge</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2-3: Different kinds of knowledge being created

Source: (Nonaka and Takeuchi, 1995 cited by Beijerse, 1999)

2-8. What is Knowledge Management?
KM is usually described as a management tool. More precisely, it is described either as an operational tool or as a strategically focused management tool (Martensson, 2000).

Considering KM as an information handling tool, knowledge is often regarded as an information handling problem. It deals with the creation, management and exploitation of knowledge (Martensson, 2000). In the literature some scholars define KM as separate but related stages. The first step in the process is the acquisition of information. In the second stage, the information is the input of a storage system and organized logically. Almost every definition of knowledge management includes the storage of knowledge. KM is about acquisition and storage of employees’ knowledge and making information available to other employees within the organization. This is often achieved by using various technologies such as Internet and databases, and is a translation of tacit knowledge to explicit knowledge (Martensson, 2000). Once the information is stored in the various databases, the third stage is initiated. In the third stage, the stored information is made accessible to as many employees as possible within the organization. It is about distributing it into the hands of the right end users at the right time and where it can be of best use. The final stage which is use of the information, is about utilization of information. This process begins with people
sharing and distributing knowledge by talking and socializing with one another or by exchanging information in digital or analogue form (Martensson, 2000).

<table>
<thead>
<tr>
<th>Collecting information</th>
<th>Storing information</th>
<th>Making the information available</th>
<th>Use of the information</th>
</tr>
</thead>
</table>

Source: (Martensson, 2000)

Figure 2-10: The stages of knowledge management

2-9. KM in NPD

Traditionally, Innovation processes were united with technical aspects of new product development and were considered a linear series of related activities in the domains of science, technical research and commercialization. Now, they are currently being considered processes that integrate information flows and bases of knowledge created within and externally to the boundaries of the firm. Since knowledge is the foundation of innovation, NPD and knowledge management are strongly linked (Pérez-Bustamante, 1999).

The value of KM for reducing the risk in NPD, confirms by collecting data from internal and external sources and extracting relevant information in order to prevent product failure. The internal problems affecting product failure include the company’s ability to meet product performance, reliability, or cost requirements, while the external problems include: unsuccessful product acceptance in the market, changing regulations, and so on (Cooper, 2003).

KM accelerate dissemination of knowledge in the NPD process, thus can prevent rework by documenting and disseminate lessons learned and forming organizational memory, depict map of tacit knowledge of the organization and shorten the product lead time. But from another pint of view, KM mostly focuses on employees’ knowledge and has little systematic attention to the knowledge of customers, which is company’s most important partner in the value creation process (Su et al, 2006). Today the role that KM plays in the NPD process is better understood. yet, the role has resulted only in making a contribution to in-company NPD outcomes such as product/service quality, reduced cost, and deliverables to market, and not in terms of market outcomes such as sales, customer satisfaction, and return on investment. It is valuable that both in-company and market outcomes linked together in order to achieve real success in NPD (Su et al, 2006).

The concept, Customer Knowledge Management will be introduced in the next section.
2-10. Introduction to Customer Knowledge Management

Customer knowledge is a main asset for all businesses. It is the basis of most improvements in customer value. (Rowly, 2002) With stress on knowledge as a major competitive factor in the global economy, corporations may be overlooking a key element: customer knowledge.

Customer relationship management (CRM) and knowledge management (KM) initiatives are heading for the same target: the delivery of continuous enhancement towards customers. Initiatives resulting from this effort have been labeled 'customer knowledge management' (CKM) or 'knowledge-enabled CRM'. In the literature, it is conceptualized CKM as the make use of of knowledge for (e.g. product information) from (e.g. their ideas about product improvements) and about customers (e.g. their requirements and expectations) to enhance the customer-relating capabilities of firms. (Saloman, et. al., 2005)

At first glance, CKM may look like another name for Customer Relationship Management, or Knowledge Management. But customer knowledge managers need a different approach (Gibbert et. al., 2002).

Here are two basic conceptual differences:

- Customer knowledge managers, first and foremost focus on knowledge from the customer (i.e. knowledge residing in customers), rather than focusing on knowledge about the customer, as characteristic of customer relationship management.

- Traditional KM satisfies a critically important role so that knowledge becomes the key value-added resource in companies. Yet, this knowledge is usually shared, expanded and leveraged among employees, or between companies, with little systematic attention accorded to what could be the company’s most important partner in the value creation process, customer. In contrast to KM’s proper focus on ‘if only we knew what we know,’ CKM suggests an additional aspect: ‘if only we also knew what our customers know.’

For more precise compare of CRM, KM and CKM see table 2-4.
### Table 2-4: CKM versus Knowledge Management & Customer Relationship Management

<table>
<thead>
<tr>
<th>Knowledge sought in</th>
<th>KM</th>
<th>CRM</th>
<th>CKM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Axioms</td>
<td>'If only we knew what we know.'</td>
<td>'Retention is cheaper than acquisition.'</td>
<td>'If only we knew what our customers know.'</td>
</tr>
<tr>
<td>Rationale</td>
<td>Unlock and integrate employees’ knowledge about customers, sales processes, and R&amp;D.</td>
<td>Mining knowledge about the customer in company’s databases.</td>
<td>Gaining knowledge directly from the customer, as well as sharing and expanding this knowledge.</td>
</tr>
<tr>
<td>Objectives</td>
<td>Efficiency gains, cost saving, and avoidance of re-inventing the wheel.</td>
<td>Customer base nurturing, maintaining company’s customer base.</td>
<td>Collaboration with customers for joint value creation.</td>
</tr>
<tr>
<td>Metrics</td>
<td>Performance against budget.</td>
<td>Performance in terms of customer satisfaction and loyalty.</td>
<td>Performance against competitors in innovation and growth, contribution to customer success.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Role of customer</th>
<th>Corporate role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passive, recipient of product.</td>
<td>Encourage employees to share their knowledge with their colleagues.</td>
</tr>
<tr>
<td>Captive, tied to product/service by loyalty schemes.</td>
<td>Build lasting relationships with customers.</td>
</tr>
<tr>
<td>Active, partner in value-creation process.</td>
<td>Emancipate customers from passive recipients of products to active co-creators of value.</td>
</tr>
</tbody>
</table>

Source: (Gibbert et al., 2002)

### 2-11. A Model for Customer Knowledge Management

The customer knowledge management process model as introduced by Gebert et al. (2003), aims at integrating the two concepts of CRM and KM. See figure 2-12. 

**CRM process model:** Marketing, sales, and service are major functions with a high degree of direct customer interaction and knowledge intensity. We derive process model by detailing these functions into relevant business processes, which may be cross-functional. Such processes are either initiate by the customer for receiving information or services or by the enterprise with the aim of delivering information or services to customers. Each process handles a particular business object which distinguishes it from other processes. Based on the three major processes and the definition of sub-processes and ordinary tasks in CRM literature, six processes identified: Campaign management, lead management, offer management, contract management, complaint management, and service management. These are the six relevant CRM business processes (Rowly, 2002).
Campaign management is a core marketing process which accomplishes the idea of interactive, individualized contacts in contrast to traditional transaction marketing. It deals with the planning, realization, control and monitoring of marketing activities towards known recipients. Marketing campaigns are individualized (one-to-one marketing) (Peppers and Rogers, 1993) or segment specific.

Lead management is the strengthening, and prioritization of contacts with prospective customers. The aim is to provide sales staff with a qualified and prioritized list of valuable prospects.

Offer management is the core sales process. Its purpose is to deliver individualized, binding offers.

Contract management is the design and continuance of contracts for the supply of products and services.

Service management is the planning, realization and control of measures for the services. A service is an intangible output of a firm that deals directly with customers.

Within the scope of complaint management, expressed dissatisfaction of customers is gathered, processed, and communicated into the firm. The aim is to improve customer satisfaction.

Integrating CRM and KM: To provide a solution for the process of the customers, firms need to focus on three sorts of knowledge in CRM processes. They make up what is considering being customer knowledge.

(1) They should understand the needs of customers in order to address them. This is “knowledge about customers”.

(2) The information needs of the customers in their interaction with the firm are “knowledge for customers”.

(3) At last, customers have knowledge about the products and services they use. This “knowledge from customers” is valuable as it brings into measures to improve products and services.
2-12. E-CKM Model, a model for NPD based on CKM  
As mentioned previously, customer knowledge management is very important for NPD project. In order to address this issue, a conceptual framework entitled the E-CKM model by incorporating IT into the CKM model, as shown in Figure 2-13, has been proposed (Su et al, 2006). This model connects the above-mentioned fragmented concepts into a framework. In the E-CKM model, the CKM process includes four stages which are supported by the applications of information technology, by using e-mail and web survey.
The implementation plan is as described below:

1. Product features/benefit identification
   ‘Benefit’ is defined as the needs a customer wants satisfied from purchasing and using a product. After a new product idea is conceived, the firm identifies potential product benefits in terms of a customer’s perceived value, in the form of features, functions, and other attributes which can be communicated to the customers. This means, at this stage, the firm delivers product knowledge ‘for’ the customers. Each customer may make response based on her/his own attitudes toward these features or benefits, via a bi-directional communication channel like personal interviews, phone interviews, self administered questionnaires, etc. Among them, a mail survey by sending a mail with self-administered questionnaires has been assessed as an acceptable method (Cooper and Schneider, 2001 cited by Su et al, 2006)
   Information technology can improve communication efficiency and effectiveness by including product knowledge in the self-administered questionnaires and e-mailed it to the customers, or by sending an e-mail to customers and persuading them to answer the survey instrument posted on a predetermined website. All the customers' responses should be collected in a database. This leads to identify what product knowledge the customers already have about the presented product features (Su et al, 2006).
2. Customers’ needs categorization
A questionnaire provides responses from customers about their attitude, preference, needs, and perceived value for product features offered. If a company wants to suggest products to target customers, a method should be developed for segmenting customers in the study, based on their pattern of needs towards a set of potential benefits existing in the product. In the E-CKM model, with the aid of IT and using an interval rating scale instrument, an electronic survey is accessible on a website. By the means of the questionnaire the primary data of personal background of a customer, as well as the level of which he/she desires each product attribute, in order to elicit customers’ needs will be collected. These needs are then categorized into homogeneous groups which mean segmenting customers based on their need pattern (Su et al, 2006). Among the several response methods in the fields of market research, the ‘multiple rating list scale’ is the most suitable one for a web-based questionnaire, since it provides a layout of easy visualization (Cooper and Schneider, 2001 cited by Su et al, 2006). The total of the responses of the questionnaires collected from all customers, first undergoes a data cleaning process, and then becomes the database ready for data mining. The next step would be categorization of the customers’ needs. The firm obtains knowledge ‘about’ the customers by understanding the customers’ background, needs, and preference pattern toward product attributes (Su et al, 2006).

3. Market segmenting for converting tacit customer knowledge into codified knowledge
Recent studies shows benefit-based or needs-based segmentation methodology is the most appropriate if taking into account the overall performance (Kamakura and Wedel, 2000 cited by Su et al, 2006). The point is that demographic segmentation simply describes customers’ behavior, but fails to enlighten the reason why it is; while the benefits a product delivers can change the customers’ attitude towards the product (Haley, 1985 cited by Su et al, 2006).
Through communication by a web-based survey, a company is able to make use of knowledge ‘for’ customers and knowledge ‘about’ customers, and perform the appropriate market segmentation mission. After the segments are formed through data mining techniques, each segment’s pattern of needs toward product attributes is defined. Now the different characteristics of each segment can be recognized and analyzed. At this stage the tacit customer knowledge, discrete among the individual customers is elicited and gathered, and it can be converted into explicit customer knowledge desired by the company (Su et al, 2006).

4. Customers’ needs pattern extraction
After finishing segmentation, the characteristics of the customers’ needs in each segment should be studied in order to extract the needs patterns in each segment. Therefore, the knowledge ‘from’ customers enables the firm to target the right market segments. It also allows them to make the right strategic business decisions. Knowledge from customers helps the company to fine-tune the original definition of the product, put priorities for product attributes to be
developed, advance the functionality of the desired product features, and rule out product features in which the customers show no interest (Su et al, 2006).

5. Data-mining technique to implement market segmentation
In the E-CKM model presented by Su et al, 2006, Three data-mining methods are applied: K-means; the unsupervised Self Organizing Feature Map (SOM) neural network by Kohonen, 1990; and a network based on the Fuzzy Adaptive Resonance (FuzzyART) theory by Carpenter et al., 1991.

The E-CKM framework by Su et al, (2006), will be the conceptual model in this research. But the limitation of this research is to collect data electronically via e-mail or web.

A market segment consists of customers who share a similar set of wants (Kotler, 2003). Market segmentation defines as the process of forming groups of consumers, in a way that groups are homogeneous in terms of demand elasticity and are available via marketing strategies (Kotler, 1980, cited by Kiang, 2006). A marketing manager can select and successfully perform segment-specific marketing mixes. The value of performing market segmentation analysis includes positioning the product in the marketplace correctly, identifying the appropriate segments for target marketing, finding opportunities in existing markets, and gaining competitive advantage through product differentiation. The result of all mentioned values could be increased profitability (Kiang et al, 2006).

Businesses from all industry sectors use market segmentation in their marketing and strategic planning. The underlying reason and logic is in our changing environment, Customers needs are becoming more and more varied. Thus, these needs cannot be satisfied by a mass marketing approach anymore. Businesses can face with this diversity by grouping customers with similar requirements and buying behavior into segments which are homogeneous in needs and buying behavior and treat each segment specially. Choices about which segments are the most appropriate to serve are crucial to best utilizing the company resources (Dibb, 1998). The marketer identifies the segments and decides which one(s) to target so the company can create a more fine-tuned product or service offering and prices it appropriately for the target segment, also easily distribute and communicate it (Kotler, 2003).

Segmentation builds on an admirable perceptive of customers and competitors which can result in fewer direct conflict with competitors and the design of more appropriate marketing program. Segmentation also helps firms to allocate financial and other resources more effectively and play to their strength (Dibb et al, 2002)

So many companies are turning to micromarketing instead of mass marketing, at four levels: segments, niches, local areas, and individuals (Kotler, 2003).

A niche is a closely defined group seeking a distinctive mix of benefits. Niches are identified by dividing a segment into sub-segments (Kotler, 2003). In niche marketing, all customers are not equal and similar (they have different lifetime...
value or purchase behaviors), even if they buy identical products or services. Firms are more and more understanding the importance of the customer lifetime value or CLV (Liua and Shih, 2004). Generally thinking, recency, frequency, and monetary (RFM) methods are useful to measure CLV (Miglautsch, 2000 cited by Liua and Shih, 2004). CLV is normally used to recognize profitable customers and target them. Measuring RFM is a recognized method for CLV. The RFM are: (1) R (Recency): period since the last purchase; lower recency value indicates more profitable customer (2) F (Frequency): number of purchases made within a certain period; higher frequency imply greater loyalty; (3) M (Monetary): the money spent during a certain period; a higher value reveal that the firm should focus more on that customer (Liua and Shih, 2004).

Local marketing defined as tailoring marketing programs to the needs and wants of local customer groups which have been targeted (Kotler, 2003) Individual customer marketing is equivalent to “segments of one”, “customized marketing” or “one-to-one marketing” (Kotler, 2003).

2-13. Process of Market Segmentation
Market segments can be built up in 3 ways (Kotler, 2003):
1. Homogeneous preferences: all customers have approximately the same preferences.
2. Diffused preferences: customers vary significantly in their preferences. The first brand that enters the market is likely to attract the most people. A second competitor must position next to first to battle for the market share or far from to attract the unhappy group with first brand.
3. Clustered preferences: the market distinct preference clusters, called natural market segments. Companies should try to develop different brands to cover all clusters and make entry difficult for another companies or just position in the main cluster.

Steps in market segmentation are (Kotler, 2003):
1. Needs-based segmentation: grouping customers into segments based on similar needs and benefits sought by customer in using a product or service.
2. Segment identification: for each needs-based segments, determining which demographic, lifestyle and buying behaviors make the segment different and particular (actionable).
3. Segment attractiveness: using predetermined segment attractiveness criteria (such as market growth, competitive intensity, and market access) determining the overall attractiveness of each segment.
5. Segment positioning: for each segment, creating a value proposition and product price positioning strategy based on that segment’s unique needs and characteristics.
6. Segment “acid test”: creating "segment storyboard" to test the attractiveness of each segment’s positioning strategy.
7. Marketing-mix strategy: develop segment positioning strategy to include all aspects of the marketing mix: product, price, promotion and place.

2-14. Major Segmentation Variables for Consumer Markets
Two wide-ranging groups of variables are used to segment consumer markets (Kotler, 2003). First, segmentation by looking at consumer characteristics: geographic, demographic and psychographics. And after using these features in segmentation, examine the existence of different needs or product responses in the segments. Second, segmentation by looking at consumer responses to benefits, use occasions, or brands. Once the segments are created, the researcher sees whether special characteristics are associated with each consumer-response segment or not.

As mentioned, forceful competition is forcing companies to develop innovative marketing activities to capture customer needs and improve customer satisfaction and retention and segmentation has been used as a fundamental activity to achieve these matters. Moreover, the use of the internet and the explosive growth of e-commerce provide the opportunity to gather large volumes of customer data available for analysis. Businesses can advantage extensively from analyzing customer data to find out their customers’ preferences and consequently improve marketing decision support (Liua and Shih, 2004). Lately, IT has been utilized to help firms retain competitive advantage (Stone and Good, 2001 cited by Liua and Shih, 2004). In addition, Data mining techniques are widely used information technology for extracting marketing knowledge and promote marketing decisions support (Shaw et al, 2001 cited by Liua and Shih, 2004). Such data mining applications include market basket analysis, retail sales analysis, and market segmentation analysis.

2-15. Segmentation by Data Mining
Data mining is ‘the extraction of hidden predictive information from large databases’, a cutting-edge technology with great potential to help companies dig out the most important trends in their giant database (Thearling, 1999 cited by Hung et al, 2006). Rising data mining tools can answer business questions that have been conventionally too time-consuming to solve. Data mining techniques make the transformation of raw data into business knowledge actionable (Lejeune, 2001 cited by Hung et al, 2006). Data mining is applying data analysis and discovery algorithms to identify patterns in data for prediction and description.

Having adequate database size and quality, data mining technology can provide business intelligence to generate knowledge and detect opportunities (Lau, Wong, Hui, & Pun, 2003; Su, Hsu, & Tsai, 2002 cited by Hung et al, 2006). Data mining processes can be divided to six chronological, repetitive steps:
1) Problem definition;
2) Data acquisition;
3) Data preprocessing and survey;
4) Data modeling;
5) Evaluation;
6) Knowledge deployment.
Each step is essential: The problem defines what data are used and what a good explanation is. Modeling makes it feasible to apply the results to new data. On the other hand, data modeling without good understanding and careful preparation of the data leads to troubles. Finally, the whole mining process is to make use of new knowledge (Pyle, 1999 cited by Vesanto and Alhoniemi, 2000).

**Hypothesis testing Vs. Data mining**

Hypothesis testing is positively useful, but in some cases it is not sufficient (Berry and Linoff, 2003). Data miners often bounce back and forth between two approaches: Data mining and hypothesis testing. What they do is thinking up possible explanations for observed behavior (often with the help of business experts) and letting those hypotheses order how to analyze the data. Then, letting the data recommend new hypotheses to test (Berry and Linoff, 2003). The data mining techniques create new models based on data. Generally, a model is an explanation or description of how something works. Models take a set of inputs and produce an output (Berry and Linoff, 2003).

![Figure 2-14 Data mining scheme](image)

Source: Berry and Linoff, 2003

Data mining techniques can make three kinds of models for three kinds of questions: descriptive profiling, directed profiling, and prediction. The distinctions are not always clear. Descriptive models describe what is in the data. The output is one or more charts or numbers or graphics that clarify what is happening in the data. On the other hand, Hypothesis testing often produces descriptive models. Both directed profiling and prediction have a goal in mind when the model is being built. The difference between these two is in time frames. In profiling models, the target is from the same time frame as the input. In predictive models, the target is from a later time frame. Prediction means finding patterns in data from one period and being able to explain outcomes in a later period.

Profiling is a familiar approach in solving many questions. It necessarily would not involve any sophisticated data analysis. Surveys, for example, are one common method of developing customer profiles. Surveys show what customers and prospects look like. Profiles are built often based on demographic variables, such as geographic location, gender, and age (Berry and Linoff, 2003).

Profiling has serious limitations. The inability to distinguish cause and effect is one limitation. As long as the profiling is based on familiar demographic variables, it will not be so useful. If men buy more malt beverage than women, we do not have to wonder whether malt beverage drinking might be the cause of maleness. It seems acceptable that the link is from men to malt beverage and not vice versa. With behavioral data only, the direction of causality is not always obvious (Berry and Linoff, 2003).
Clustering, tools for segmentation in data mining
Cluster analysis is a means for exploring the data structure. The core of cluster analysis is the process of grouping objects into clusters such that the items from the same cluster are similar and items from different clusters are dissimilar (Soman et al, 2006).
Cluster analysis can be implemented on customer databases in order to identify homogeneous subpopulations of customers. These clusters or segments could represent individual target groups for the purpose of marketing. (Han and Kamber, 2006)

A cluster is a set of data items that are similar to one another within the same cluster and are dissimilar to the items in other clusters. (Han and Kamber, 2006)

Not like classification, clustering does not work based on assumptions about class labels that mark items with past identifiers. Therefore, clustering is an unsupervised learning technique while classification is a supervised technique.
The fact is, classification is a valuable means for making groups or classes of items distinctive, but it requires costly and time-consuming collection and labeling of a large set of training patterns, which the classifier uses to model each group. For segmentation, it is regularly more desirable to proceed in the reverse direction: first partition the set of data into groups based on data similarity by clustering and then appoints labels to the quite small number of groups (Han and Kamber, 2006).
Clustering is called data segmentation in some applications since it partitions large data sets into groups based on their similarity.
The need to structure and learn from the dynamically growing amount of data has been a force for utilizing clustering in many research areas. Human beings cannot simply learn and create knowledge from the excess of information in databases without the help of summarization techniques. Basic statistics or histograms only provide an initial understanding of the data. However, cluster analysis can reveal more intricate relationships among the items, among the features and between both. Clustering is important in the mining process because it summarize data to a manageable level by forming, for example, groups of customers with similar profiles (Soman et al, 2006).
In clustering analysis inputs required are similarity measures or data from which similarities can be computed. Webster’s dictionary defines similarity as the quality or level of being similar; likeness; resemblance; as, a similarity of features. The real meaning of similarity is a little bit philosophical, but in data mining we have to do a practical approach so we can measure the similarity. We define similarity based on features. Most of the times, in order to measure similarity we have to do preprocessing as:
- Generate features: generate new features based on existing ones
- Clean features: cleaning noises or outliers
- Normalize features
- Transforming features
- Reduce features: in case of too many features to avoid the curse of dimension:

There are two major ways of measuring similarity: feature projection and edit distance.
In feature projection, data is projected into feature space; the distance in feature space is appropriately measured and becomes similarity. So the similarity between items depends on the features we measure and also the distance measure.

In edit distance, one item is transformed into another item and the effort to do this is measured. The measure of the effort or rather the cost of this transformation becomes the similarity. To measure similarity, we measure relevant features (which is specific for each problem) of the object and then arrive at a numerical measure of similarity or distance.

There are several formulas for measuring distance: Euclidean distance, city-block (Manhattan) distance, chebychev distance, power distance, percent disagreement (Soman et al, 2006).

When features are binary, items cannot be represented by meaningful p-dimensional measurements. Thus pairs of items are often compared on the basis of the presence and absence of definite characteristics. Similar items have more characteristics in common than dissimilar items. The presence or absence of a certain characteristics can be described mathematically by introducing a binary variable, which assume value 1 if the characteristic is present and value 0 if the characteristic is absent.

The distance measure in this situation is mostly Euclidean distance, the same as the distance in this research.

\[
d_{Euclidean}^{(x,y)} = \sqrt{\sum_i (x_i - y_i)^2}
\]

**Basic types of clustering**

Broadly thinking, there are two major types of clustering (Soman et al, 2006):

1. Partitional
2. Hierarchical

1- Hierarchical clustering (which has not been used in this research): in Hierarchical algorithm, we create a hierarchical decomposition of the set of objects using some criterion.

Hierarchical trees can be built in two ways:

- Bottom-up (agglomerative): In bottom-up hierarchical clustering, we assume that all items belong to a separate cluster, then, we find the best pair to merge into a new cluster. It is the most commonly used hierarchical method.
- Top-down (divisive): in top-down hierarchical clustering, we start with all the data in a single cluster; consider all the possible choices to divide the cluster into two. Choose the best division and recursively operate on both sides.

In both methods, there is no need to specify the number of clusters at the beginning.
2- Partitional clustering (which has been used in this research): in particular, we construct various partitions and then evaluate them by some standard.

Generally, Partitional clustering can be divided into two methods (Berry and Linoff, 2003):

- **K-Means**: is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem. The procedure follows an easy way to classify a given data set through a certain number of clusters (k clusters). The main idea is to define k centroids, one for each cluster. These centroids should be placed in a cunning way because of different location causes different results. So the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given dataset and associate it to the nearest centroids. When no point is pending, the first step is completed and an early clustering is done. Then, we re-calculate k new centroids and a new binding has to be done between the same dataset points and the nearest new centroids. A loop must be generated. As a result of this loop, k centroids change their location step by step until no more changes are done and they don’t move.

The algorithm is made of following steps:

1. Determine a value for k (number of clusters)
2. Initialize the k clusters centers, randomly if necessary
3. decide the class memberships of the N objects by assigning them to the nearest clusters centers
4. re-estimate the k clusters centers, by assuming the membership found above are correct
5. If none of the objects changed membership in the last try, exit. Otherwise go to 3

Although it can be proved that the procedure will always terminate, the k-means algorithm does not necessarily find the most optimal configuration. It is also sensitive to the initial randomly selected clusters centers. The k-means algorithm can be run multiple times to reduce this effect.

This algorithm has some weaknesses like:

- The way to initialize the means was not specified
- The results depend on the value of k. One popular way to start is to randomly choose k of the samples

Using k-means method for clustering has advantages and disadvantages.

Advantages are:

- With a large number of variables, k-means may be computationally faster than hierarchical clustering, if k is small
• k-means may produce tighter clusters than hierarchical clustering
Disadvantages are:
• Difficulty in comparing quality of clusters produced for different k values or for different initial partitions
• Fixed number of clusters can make it difficult to predict what k should be
Chapter 3:

Research Methodology

3-1. Introduction
3-2. Research Design
3-3. Designing qualitative research process: Exploration
   3-3-1. Means-end Theory
   3-3-2 Revealing Means-end chains
3-4. Designing the quantitative research process
   3-4-1. Designing the questionnaire
   3-4-2. Data collection
   3-4-3. Data Analysis
3-1. Introduction
In order to gain customer knowledge according to our conceptual framework, the CKM framework, we should capture three kind of knowledge:
- Knowledge for customer
- Knowledge about customer
- Knowledge from customer
To gain all these three aspects of customer knowledge, a mix of quantitative and qualitative methods implemented.
We identified Means-end Chain theory a remarkable base for eliciting both “knowledge for” and “knowledge about” customer. Techniques developed based on Means-end which used in this study are Association Pattern Technique (APT), Laddering and Triadic Sorting. APT is quantitative method to measure means-end chain and the other two are qualitative methods to reveal each element of means-end chain.
To gain knowledge from customer, we segmented customers based on their means-end chains by clustering with data mining techniques. This stage is totally qualitative.

3-2. Research Design
The research design from marketing viewpoint is a framework or blueprint for conducting the research project that specifies the procedures necessary to obtain the information needed to structure and/or solve the research problem (Malhotra, 1996).
There are two major types of research designs: exploratory and conclusive. Conclusive research design can be further classified as: descriptive and causal. Exploratory research is to explore the problem situation by gaining ideas and insight into the problem. Conducting in-depth interviews or focus groups could provide valuable insights but generally followed by further exploratory research. Conclusive research assists the decision maker in determining, evaluating and selecting the best course of action for a given situation (Malhotra, 1996).
Descriptive research is to describe something usually market characteristics like target market profiles, the relationship between product use and perception of product characteristics and etc. Descriptive research assumes that the researcher has the prior knowledge about the situation that is the major difference between descriptive and exploratory research (Malhotra, 1996).
Surveys, panels, secondary data analyzed quantitatively are the major data collection techniques in this type of research.
Descriptive research could be in direct connection to exploratory research, since researchers might have started off by wanting gain insight to a problem, and after having stated it; their research becomes descriptive (Saunders et al., 2000).
There are two main research approaches to choose from when conducting research: qualitative and quantitative method (Yin, 1994).

In this study, an exploratory qualitative research has been conducted to design the questionnaire and run a descriptive quantitative research. See figure 3-1 on the next page.
In stage one of research, elicitation of attributes and benefits of the product and values of consumers, which are elements of means-end chain, with in-depth interviews based on repertory grid technique and laddering technique has been done. This stage is a qualitative exploratory research.

In stage two of research, designing the questionnaire, data collection and data analysis for market segmentation has been done. Data gathered by survey of 660 persons, the questionnaire designed based on Association Pattern Technique (APT) and then analyzed by data mining clustering algorithm to obtain market segments. This stage has been a quantitative descriptive research.

Figure 3-1. The research design
The general research process is figured below:

![Diagram of research process]

In the next sections, we will go through all parts of research process in details.

**3-3. Designing qualitative research process: Exploration**

In the figure 3-3, the process of qualitative research is presented:

- **Interviews with experts to elicit product attribute**
  - Conducting in-depth interviews with 3 experts of NAB or malt beverage production and sales team to understand product attributes (an open questionnaire guideline was used)

- **Interviews with consumers to elicit**
  - Conducting in-depth interviews with 10 consumers (sample representatives) with the help of Kelly’s triadic sorting (repertory grid) technique

- **Interviews with consumers to elicit means-end chains**
  - Conducting laddering technique, with the same group of previous step, to reveal the consequences (benefits) of consuming malt beverage sought by consumers related to the product attributes and their life values which could be satisfied with the product benefits.
In fact, all the steps of figure 3-3 are designed to reveal means-end chains of ten consumers to use in design of the questionnaire in the qualitative research.

3-3-1. Means-end Theory
The basic assumption of means–end theory is that consumers do not demand products for their own sake, but because of the consequences utilization of these products will have for them (Gutman, 1982). Products can be described in terms of their attributes, but it is not these attributes, but their consequences that are of interest. Consequences of the consumption of food products may be having a good taste experience, eating in a healthier way, refreshment, etc. These consequences in turn are regarded as desirable because they are related to the attainment of basic life values like fun & enjoyment, protection & security for oneself and the family, sharing pleasure with loved ones, etc. The sequence of attributes–consequences–values is called a means–end chain, because it shows, how — in the consumer mind — a product is regarded as a means to reach an end (Grunert and Valli, 2001).

Figure 3-4 demonstrates two hypothetical examples of means–end chains in the food area, meat product adopted from Grunert and Valli (2001).

<table>
<thead>
<tr>
<th>value</th>
<th>security</th>
<th>fun and enjoyment</th>
</tr>
</thead>
<tbody>
<tr>
<td>consequence</td>
<td>good for health</td>
<td>good taste</td>
</tr>
<tr>
<td>attribute</td>
<td>low fat</td>
<td>red colour</td>
</tr>
</tbody>
</table>

Source: (Grunert and Valli, 2001)

Let’s get deeper in this content. There are three types of product knowledge that consumers could have (Peter et al, 1999):

- Knowledge about the attributes or characteristics of product
- Knowledge about the positive consequences or benefits of product consumption
Knowledge about the values that product helps to achieve

Consumers can see bundles of attribute in products. Do they memorize and use all of the attributes in making purchase decisions? No for sure. But we have to know the most important attributes for them and also the meaning of those attributes. Meanings derive from the consequences or benefits of using the product. Consequences are the results or outcomes that happen when the product is purchased and used or consumed (Peter et al, 1999). Consequences of product consumption could be divided into Functional and Psychological (Emotional). Functional consequences are tangible outcomes such as good taste or refreshment in drinking a beverage. Psychological or emotional consequences are internal, personal outcomes such as how the product makes you feel. For example drinking a refreshing beverage make you feel happy and pleased. Consequences could be positive or negative in consumers mind. Benefits are the positive and desirable consequences that consumers seek when buying and using products (Peter et al, 1999).

Values are people’s broad life goals, like being successful, secured, and happy. Values are the ends that people seek to achieve in their lives. Values have been shown to be a powerful force in governing the behavior of individuals in all aspects of their lives (Rokeach, 1968, Yankelovich, 1981 cited on Gutman, 1982).

Consumers can combine the three types of product knowledge to form a simple associative network which named: Means-end chain. A knowledge structure that links consumers’ knowledge about product attributes with their knowledge about consequences and values (Peter et al, 1999). It suggests that consumers think: “what is this attribute good for? What does this attribute do for me?” Consumers are seeking to reach the end that could be a benefit or an abstract value.

![Means-end chain diagram](image-url)

Source: (Peter et al, 1999)

Figure 3-5. Ideal Means-end Chain

In fact, the actual product knowledge in consumers’ means-end chains does not necessarily contain each of the four levels of product meanings shown in the figure 3-5. And also some product attributes may have multiple means-end chains. An important fact here is that because means-end chains contain consumers’ personally relevant meanings for products, they are unique to each consumer’s background and personal interests. Thus different consumers are likely to have different means-end chains for the same product, although there usually are some similarities (Peter et al, 1999). This fact leads us to find the segments of consumers based on their means-end chains similarities. Consumers could be segmented in
clusters that have the most similarities within a cluster and the most dissimilarities between the clusters. This will be discussed completely later, in the clustering section.

In order to develop new, differentiated products in a consumer-led way, we need not only what product features consumers want the most but also an understanding of what motivates some consumers to demand other products than others. Means–end theory is a useful instrument in this context (Grunert and Valli, 2001).

3-3-2 Revealing Means-end chains

Various interview methods have been developed for measuring means-end chains. They all involve two basic steps (Peter et al, 1999):

1. Identifying or eliciting the product attributes that are most important to consumers
2. An interview process called laddering, to understand how consumers link product attributes to more abstract consequences and values, to reveal their means-end chains.

Techniques that apply to the first step are: Direct elicitation, Free-sort task, Triadic sorting task, and Ranking task (Peter et al, 1999, Van Kleef, 2006).

In this research, before conducting one of these techniques, an interview with three experts in non-alcoholic beverage (N.A.B.) and specially Malt beverage industry has been conducted to elicit the Malt beverage features from the experts’ point of view. After this, the elicited attributes checked by interviews with a small sample of consumers, ten persons: 5 men and 5 women, aged 20 to 35, have university degrees, which are representatives of our sample. Interviews were based on Triadic Sorting Technique which is the very first step of Kelly Repertory Grid Technique.

- **Triadic Sorting or Kelly’s Repertory Grid** is a technique developed by Kelly (1955) with the purpose of mapping cognitive structures. Kelly’s repertory grid is a personal interviewing technique used to elicit the constructs by which consumers interpret a product category (van Kleef, 2006). With the use of this method, Constructs for example attributes of products, can be elicited by repetitively presenting the respondent with triads of products drawn from a large set and asking which two products are alike and different from a third. For each triple group, the respondent is continually asked for “an important attribute on which two products are alike and at the same time different from the third” until he/she has no idea of comparing group items (Bech-Larsen, 1999 cited on van Kleef, 2006).

In this research, respondents were confronting with triads of malt beverages drawn from a set of pervading Iranian or foreign brands of malt beverages in Tehran market. Both the packages and transparent glass were on the table, in a quiet room provided by Kalleh Company. Respondents were asked to open the package and pour the malt beverage to the glass. Then, they were asked: ‘in what important way are two of these similar and different from the third? ’ this question was repeated with different
statements and also with different set of triads, to elicit the attributes of malt beverage from the consumers point of view.

It has been tried to choose the brands that, according to the last market research conducted in Kalleh Company, have the higher market share in Iran, specifically Tehran. Ones that have the 1st to 4th rank in the market share.

For the second step, which is elicitation of product benefits and consumer values, in-depth interview with the same sample of consumers in the first step conducted. The technique used in this step is laddering technique.

- **Laddering techniques** have become admired as a means of understanding consumers’ motivations for product choice especially for food products.

Laddering is a one-to-one in-depth interviewing technique applying a series of directed probes to reveal how subjects’ link product attributes to their own underlying life values (Russell et al, 2004).

The basis of this method is the assertion that lower levels involve the presence of higher levels, so that product attributes can cause to consequences that lead to life value satisfaction. The aim is to determine the chain, or “ladder”, of linkages between the attributes, consequences and values in relation to product choices made (Reynolds and Gutman, 1988 cited on Baker, 2002).

The consumer is continuously probed with some form of the questions all begin with why, like: ‘why is that important to you?’ This way of questioning forces the respondent up the ‘ladder’ of abstraction, until s/he cannot go further. The end will not always be at the value level. The result is a chain of concepts, which are called ladders.

In fact, sometimes in some articles laddering is equated with means-end chain theory; although the theory should be considered distinct from the methodology. In the classical laddering process described above, the natural speech of the consumer is not restricted as much as possible. This kind of laddering is referred to as ‘soft laddering’. On the other hand, some researchers use ‘hard laddering' approach, allowing less freedom in the answers of consumers and forcing consumers to pursue one ladder at a Time in which each subsequent answer is on a higher level of abstraction (Grunert and Grunert, 1995 cited on Hofstede et al, 1998).

In this research, for each important attribute, the researcher asks the respondents:' why is that important to you?' This question must be repeated until the ends are expressed by consumers. The ends could be emotional benefits or life values. An important notice here is that these “why” questions are hard to answer and also time consuming. So the interview environment and researcher behavior play an important role to make the consumers relax and motivated. Researcher must make a warm and convenient atmosphere so the respondent could express him/her self.

Outcomes of benefits and values were refined by literature review for example Grunert and also an expert opinion.

Although laddering is the most famous and widely applied technique to measure means-end chains and it has performed as a very useful qualitative technique to
reveal means-end chains, but it has some limitations that affect the objective of this research. Laddering interviews are time-consuming, and necessitate well-trained interviewers. As a result, it is hard, if not impossible, to apply laddering technique to a large-scale representative sample that is normally required for market segmentation (Hofstede et al, 1998). Moreover, laddering is an expensive data collection technique. It could make respondents boring and thus an acceptable result cannot be obtained (Steenkamp and Van Trijp, 1996 cited on Hofstede et al, 1998).

In sum, laddering is not appropriate as a method to be used in large representative samples, nor was it projected to be used in this situation. In reaction to laddering limitations there have been several attempts to quantify means-end chains in large-scale studies using other methods (Hofstede et al, 1998) which will be discussed in the next section.

In this research, based on literature, laddering technique has been used just for revealing the means-end chains of 10 people of sample, therefore a list of product attributes, benefits and consumer values could be ready to design the questionnaire and run a survey to measure consumers’ means-end chain in a quantitative way as will be discussed in detail later.

3-4. **Designing the quantitative research process**

In figure 3-6, the process of qualitative research is presented:

![Diagram showing the process of qualitative research](image)

- **Designing the Questionnaire**
  - Based on literature and exploration stage, a questionnaire including 2 sections designed:
    - Section 1- Related to AC and CV matrixes for segmentation
    - Section 2-Some questions on buying behavior for profiling the segments

- **Data Collection**
  - A Survey of 660 respondents from university people
  - The sample population was based on APT literature, 5 times of ACV matrix cells.

- **Data Analysis**
  - Data mining techniques used including:
    - Preprocessing data based on LSI and TF-IDF
    - Clustering and validation based on Silhouette and K-means
    - Feature selection for profiling clusters based on Random Forest

Figure 3-6. The quantitative research process
As mentioned earlier, the aim of quantitative research in here is to have segments of malt beverage consumers based on their means-end chains towards this product and therefore determine the product attributes they want and their motives for wanting the product. So consumers’ cognitive structures are needed to segment consumers, therefore, it is a part of data collected. And also some behavioral data is needed to better describe and profile the segments, thus the second part of data is on consumers’ buying behavior. It is obvious that all these data cannot be provided from secondary sources and should be collected. So this research data are primary.

As shown in the qualitative research process, a method should be applied to measure means-end chains quantitatively. For this, Association Pattern Technique (APT) method selected and used for designing the questionnaire and making data base of consumers means-end chains.

The analysis method that results in segmentation is data mining /clustering.

In the following sections, each step of quantitative research will be described in details.

3-4-1. Designing the questionnaire

As described before, we need a method to measure means-end chains quantitatively. Association Pattern Technique (APT) is the selected method. The idea of using this method was from the literature, especially Grunert and Valli (2001) and Hofstede et al (1998) which strongly recommend APT to measure means-end chains and reveal the segments of consumers based on it. We found APT easy to conduct and the data generated by this technique, easy to cluster in comparison to other methods (in fact, APT was not conducted easily for us! because the nature of APT and Means-end, is a little bit complex and unfamiliar for consumers and for us. but in comparison to other methods in literature, e.g. Hofstede et al (1998), it is easy!).

- **Association Pattern Technique (APT)** is a quantitatively oriented technique to assess and measure means-end chains (Hofstede et al, 1998). The association pattern technique is inspired by Gutman (1982). Gutman proposed that, for measurement purposes, the means-end chain can be seen as a series of connected matrices. In APT an AC-matrix (attribute-consequence matrix) and a CV-matrix (consequence-value matrix) are distinguished. In the AC-matrix, a priori defined attributes and consequences are listed in the columns and rows, correspondingly, resulting in a table of all combinations of attributes and consequences. In the same way, the CV-matrix includes all possible combinations of the consequences and values. For each column in the AC-matrix (CV-matrix), respondents indicate that which consequence (values) comes from that attribute (consequence). This results in a dataset of binary observations. Note that, unlike laddering technique that tries to reveal the chain without any given attribute, consequence and value, APT works based on the attributes, consequences and values that has to be provided by the
researcher prior to the first meeting with the respondent. Therefore, to acquire the relevant chain, pre-testing is inevitable when secondary sources are lacking. Considering the limitations of laddering, the association pattern technique could perform as a practical supplement to laddering for measuring means-end chains. APT successfully overcomes the above mentioned limitations of laddering, specifically, the data-collection process is structured and it can be used in large-scale studies (Hofstede et al, 1998).

The outcome of APT is a binary ACV matrix which is used to segment consumers. Now we have the tool to collect Part 1 of data. In addition, for part 2 of data, a number of background variables regarding purchase behavior, personality traits and demographics must be collected to be used for profiling the segments found (Grunert and Valli, 2001). In this research, most background and behavioral variables were selected from literature e.g. Grunert and Valli (2001), Peter et al (1999) and also confirmed with expert opinion.

A two sectioned questionnaire designed. For the first section, we put the 17 elicited product attributes in the columns and the 14 benefits (consequences) in the rows. Therefore the AC matrix was built. And for CV matrix, we put the 14 benefits in the columns and the 9 values in the rows.

For section 2, we designed 26 questions. A pilot test with 30 people, sample representatives, conducted to see how the questionnaire performs. We described to them how to complete the questionnaire and we were there to answer their questions while completing the questionnaire. After completing the questionnaire, we spoke to the respondents and asked them which parts were complex or meaningless for them. Generally, the ACV matrixes were interesting but hard to answer for them. Consequently, we decided to reduce the ACV matrixes dimensions to reduce the complexity. Some attributes were removed by expert opinion and only very important ones remained. We tried to not to express benefits and values in terms but make sentences with them to make them as clear as possible. This result in more meaningful ACV matrixes. The questions on behavioral variables were designed simple and they were meaningful and easy to answer for respondents.

The questionnaire has illustrated in chapter 4.

3-4-2. Data collection
Data collection is done with a survey of 660 respondents in Tarbiat Modares University. The respondents were university people: students, professors and employees. The data collection took 5 days. A table and chairs were positioned in the self service restaurant of university so when people came for lunch, which is their free time, we asked them if they drink N.A.B, have a sit and give some minutes to complete the questionnaire. Since the questionnaire takes minutes to think and complete, we could not be anywhere for survey. Sits was necessary. This may put some limitation in our results. The pictures of data collection are in appendix 1.

After coding and data entry, two data is ready to analysis:
1- ACV matrix data: a binary dataset of consumers' means-end chains towards N.A.B. if a link exist, the value is one, if not the value is zero. This dataset has 126 features and 660 records.

2- Profiling data: the coded second part of questionnaire, including buying behavioral questions that result in 57 features.

3-4-3. Data Analysis
As mentioned earlier, data mining is the tool for data analysis in this research. The software which has been used is R (version 2.6.2). R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. R has several packages for data mining purpose. R and its extra packages are available for download at www.R-project.org.

The general process of data mining in this research is figured below:

Figure 3-7 General process of data mining in this study
3-4-3-1 Exploratory Data Analysis
At the beginning of data analysis, we should know our Data. An easy way to getting to know the data is different kind of diagrams like histogram, box plot and MDS plot and also statistical summaries of data. This step especially helps in understanding what kind of preprocessing is needed. In this research, MDS plot of ACV matrix data and statistical summaries of both ACV matrix and profiling data is used to understand data and think of the ways of dealing with noises, outliers and overall analysis.

- Multi-Dimensional Scaling (MDS) is a method of inserting the
distance information of a multi-variant data-set, in a reduced
dimension L, by looking for a set of vectors \( \{\mathbf{x}\}_{i=1}^{n} \in \mathbb{R}^L \) that
procreate these distances (Larose, 2005).

3-4-3-2 Preprocessing the Data
Much of the raw data enclosed in databases is unpreprocessed, incomplete, and noisy. For example, the databases may contain (Larose, 2005):
- Fields those are obsolete or redundant
- Missing values
- Outliers or noise
- Data in a form not suitable for data mining models
- Values not consistent with policy or common sense

For reaching data mining purposes, the databases need to undergo preprocessing, in the form of data cleaning and data transformation. Although in this research, the data is not from the secondary sources and it is primary, we had missing values, outliers therefore we did data preprocessing as follow:

- Data Cleaning
  Data cleaning algorithms try to fill in missing values, smooth out noise as identifying outliers, and correct inconsistencies in the data (Han and Kamber, 2006).

  - Replacing Missing Values
    A general method of managing missing values is simply to delete them from the records or fields. However, this may be dangerous, since the pattern of missing values may in fact be systematic, and simply deleting records with missing values would cause to a biased subset of the data. Therefore, data analysts have created methods that would substitute the missing value by a value according to various criteria. They may:
      1. Replace the missing value with some constant, specified by the analyst (Larose, 2005).
      2. Replace the missing value with the field mean (for numerical variables) or the mode (for categorical variables) (Larose, 2005).
3. Replace the missing values with a value generated at random from the variable distribution observed (Larose, 2005).
4. Use the attribute mean for all samples belonging to the same class (Han and Kamber, 2006).
5. Use the most probable value to fill in the missing value; this may be determined with regression, inference-based tools using a Bayesian formalism, or decision tree induction (Han and Kamber, 2006).

In this research, we have no missing values in the data used for clustering, since the data for clustering is binary and it means if a consumer perceive a link between an attribute and a benefit or a benefit and a value or not. So missing values does not make sense. But in the profiling data, the missing values exist. In this research method No.5 is used for filling missing values. Methods No.2 to 5 make the data biased. The filled-in value may not be correct. Method No.5, however, is a popular strategy. In comparison to the other methods, it uses the most information from the present data to predict missing values (Han and Kamber, 2006). It find the ten most similar records to the missed values in the corresponding cluster then the missing value is filled with mean of this ten records. The similarity is calculated based on Euclidean distance.

- **Removing Noises and Outliers**

  **Noise** is a random error or variance in a measured variable (Han and Kamber, 2006). Our noises were rooted mostly from mistakes in data entry by human. Which by a statistical summary of ACV matrix data and profiling data, they showed. In ACV matrix data, we assume any value more than one, is one. Because in the binary dataset, if a link exist, the value is one so any other value more than one can be assume as the existence of linkage that has been mistaken in the data entry by the operator. So we replaced all the values more than one with one. In profiling data, we assume noises as missing values, and then fill them as described in previous section.

  **Outliers** are extreme values that lie near the limits of the data range or act against the trend of the remaining data (Larose, 2005). Identifying outliers is important because they may represent errors in data entry. Also, even if an outlier is a valid data point and not an error, certain statistical methods are sensitive to the presence of outliers and may deliver unstable results. Two-dimensional scatter plots or histograms are the general ways off revealing the outliers graphically (Larose, 2005). There are also some numerically methods to remove outliers. A numerical way to remove outliers is clustering. Outliers may be detected by clustering, where similar
values are organized into groups, or clusters. Naturally, values that drop out of the set of clusters could be considered outliers (Han and Kamber, 2006).

In this research, since the data is primary and has been collected by a survey, which was not conducted electronically, some outliers were data entry mistakes (which were removed as noises) and some not.

In this research, outliers were removed by clustering. In our clustering, similarity is measured based on Euclidean distance so very small clusters or individual cluster (cluster with one member) were removed as outliers. So, outliers are computed numerically which could reduce the errors of removing outliers graphically.

- **Data Transformation and Reduction**

After the data has been assembled and major data problems fixed, the data might be prepared for the analysis. This includes adding derived fields to bring information to the shell. This stage also involves grouping classes for categorical variables, applying transformations such as logarithms, turning counts into proportions, and etc (Berry and Linoff, 2003).

In this research, the need to transform data appeared after initial clustering. We found that clusters are not distinguished very well. But the question was how to transform data? With past experiences in text mining, we found data in here very similar to text mining data. ACV matrix data seems as spars as text data. Thus we decided to implement two text mining techniques for ACV matrix data transformation and reduction: TF-IDF and LSI

- **TF weighting**

This method differentiates each term in a document by weighting them based on their frequency (TF is the abbreviation of Term Frequency). It creates a measure for relative importance of words in the term frequency; each word weight represents the frequency of that word in the document and is calculated as below (Janssens, 2007):

\[ w_{i,j} = \frac{f_{i,j}}{\max_{i} f_{i,j}} \]

Where \( w_{i,j} \) represents the weight of index term \( t_i \) in document \( d_j \). \( f_{i,j} \) is the Number of occurrences of \( t_i \) in \( d_j \). The rationale behind this approach is that Words with high frequency are important and define the content of a document accurately.

- **TF-IDF weighting**

The TF-IDF (Term Frequency - Inverse Document Frequency) weighting method has shown to be very effective in information recovery for determining the most relevant documents to a user’s query and the most important terms in documents. It represents the relevance or importance of terms in a document by counting the frequency of every word (similar to the TF model), but by also taking into account the occurrence of a particular word in the entire document collection. TF-IDF values are computed as follows:
\[ w_{ij} = f_{ij} \cdot \log \left( \frac{N}{n_i} \right) \]

Where \( f_{ij} \) is the term frequency, i.e., the number of occurrences of term \( t_i \) in document \( d_j \), \( N \) represents the total number of documents, and \( n_i \) is the number of documents containing term \( t_i \). The TF-IDF weight of a term in a document is high if the term frequently occurs in that document, but only if it occurs in just a few other documents as well, i.e., having a low document frequency and consequently a high IDF. As a result, terms that occur in a lot of documents are considered common terms and are down-weighted.

In our ACV matrix data, ‘term’ corresponds to feature and ‘document’ corresponds to record or consumer. So by TF-IDF each ACV matrix cell value is weighted therefore data will be transformed to a better format for clustering purposes.

- **Latent Semantic Indexing (LSI)**

  The curse of dimensionality is defined as the exponential growth of the ‘hyper volume’ with increasing dimensionality (Janssens, 2007). This results in a decline in the performance of algorithms to discriminate between records (customer) that are close to a given record and the majority of other records in a collection. When the dimensionality increases, distance measures become more and more meaningless as all objects seem to be almost equal in distance from each other. The search for nearest neighbors becomes very unbalanced (Janssens, 2007).

  Latent Semantic Indexing (LSI) is a mathematical technique that assumes there is some underlying or latent structure in word usage that is partially hidden by variability in word choice (Deerwester, 1990 cited on Janssens, 2007). LSI makes it possible to compose a matrix \( A_k \) that is an optimal approximation of \( A \) (in least squares sense), but with rank \( k \) much lower than the term or document dimension of \( A \). LSI is a ‘feature transformation technique’ that creates factors \( (k) \) from linear combinations of the original term dimensions. Instead of the huge number of rows in the matrix \( A \) (equal to the total number of terms), only \( k \) statistically derived orthogonal indexing factors remain in \( A_k \) (Janssens, 2007).

  However, according to Deerwester et al., the number of retained factors \( (k) \) might be critical to the success of LSI. Selecting too few factors might result in loss of important information, whereas too many factors might cause to over fitting of the model. We believe that selecting an appropriate value for the number of latent semantic factors might seriously improve the performance of subsequent clustering tasks (Janssens, 2007). We investigate clustering performance for various numbers of clusters and LSI factors and provide some insight into the relation between number of LSI factors, number of clusters, and clustering performance.
3.4.3.3 Clustering and Validation
As mentioned earlier in the literature, clustering analysis is the most popular way for segmentation. But there are important issues in clustering. Choosing the best method of clustering and the best number of clusters are high importance challenges.

A variety of measures aimed at validating the results of a clustering analysis and determining which clustering algorithm performs the best for a particular experiment have been proposed (Kerr and Churchill, 2001; Yeung et al., 2001; Datta and Datta, 2003 cited on Brock et al, 2008). This validation can be based solely on the internal properties of the data or on some external reference, and on the expression data alone or in conjunction with relevant biological information (Gibbons and Roth, 2002; Gat-Viks et al., 2003; Bolshakova et al., 2005; Datta and Datta, 2006 cited on Brock et al, 2008).

For internal validation, we select measures that reflect the compactness, connectedness, and separation of the cluster partitions. Connectedness relates to what extent observations are placed in the same cluster as their nearest neighbors in the data space, and is here measured by the connectivity (Handl et al., 2005, cited on Brock et al, 2008). Compactness assesses cluster homogeneity, usually by looking at the intra-cluster variance, while separation quantifies the degree of separation between clusters (usually by measuring the distance between cluster centroids). Since compactness and separation demonstrate opposing trends (compactness increases with the number of clusters but separation decreases), popular methods combine the two measures into a single score. The Dunn Index (Dunn, 1974) and Silhouette Width (Rousseeuw, 1987) are both examples of non-linear combinations of the compactness and separation.

The Silhouette value $S(i)$ for a record (consumer) $i$ ranges from -1 to +1 and measures how similar the record is to record in its own cluster vs. record in other clusters. $S(i)$ is defined as follows (Janssens, 2007):

$$s(i) = \min B(i, C_j) - W(i)/\max \left(\min \left(B(i, C_j)\right), W(i)\right)$$

Where $W(i)$ is the average distance from record i to all other records within its cluster, and $B(i, C_j)$ is the average distance from record i to all records in another cluster $C_j$. The mean Silhouette value over all records in a cluster is an indication of cluster quality. The average for the complete data set gives an intrinsic measurement of the overall quality of the clustering result. As Silhouette values are based on distances, depending on the applied distance measure different Silhouettes can be calculated.

With text mining experiences and literature, we found silhouette more convenient, new and also applicable to finding number of clusters, number of LSI factors and method of clustering in one try (Janssens, 2007). For feeling the effect of TF-IDF and LSI, we run silhouette validation before TF-IDF, after TF-IDF and then after LSI to see how and in what extend these two methods affect the results. And with silhouette method, the number of clusters, method of clustering and number of LSI factors revealed. These procedures executed on ACV matrix data.
3.4.3.4 Feature Selection
After clustering ACV matrix data, we should begin to profile them by profiling data, but there are 57 features in profiling data which is too many features. In fact, since we want to derive profiles from the statistical summaries of clusters, it is not impossible to profile with these too many but it is confusing. The clusters are not differentiating in all of these features so we need to find the most differentiated features. Since we have found clusters from clustering ACV matrix data, we can assume them as classes for profiling data and try variable selection methods by adding the number of clusters for each record in profiling data.
Random forest (RF) introduced in 2005 by Uriarte and Andres for variable selection in gene expression data. Random forest is an algorithm for classification, also provides feature importance (Breiman, 2001). Random forest returns a measure of error rate based on the out-of-bag cases for each fitted tree, the OOB error. Using the OOB error as minimization criterion, carry out variable elimination from random forest by successively eliminating the least important variables (with importance as returned from random forest). Random forests for Variable selection use backwards variable elimination. If, a starting point with all features is chosen, the algorithm successively removes what are considered irrelevant features; this is known as backward elimination (Al-Shahib et al., 2005).

3.4.3.5 Profiling
As mentioned, in this research, the task of profiling and describing the segments is done with the profiling data, which has different features from ACV matrix data. After selecting the proper features of profiling data, an easy way to use them in describing the segments, is to have their statistical summaries and according to these statistical summaries, describe the similarities and differences of segments relative to each other.
The other method of profiling which is frequent item set examined but the results were not good and useful. We find this method inappropriate for this matter.
Chapter 4: Empirical Data

4-1. Introduction
4-2. Attributes-consequence-value (ACV) matrix
4-3. Profiling data, for profiling the identified segments
4-4. Sample
4-1. Introduction
In this chapter, we provide an outline of the collected data. Two kinds of data collected:
1- ACV matrix data for doing the segmentation based on it
2- Profiling data for profiling the identified segments
Details of each data type, the questionnaire designed to collect data, data entry and the final dataset used to analysis are provided in this chapter.

4-2. Attributes-consequence-value (ACV) matrix
The ACV matrix data is a binary dataset. It is a matrix which rows indicate the records/consumers (660 records) and columns indicate the values of AC and CV matrix (126 features). It combines AC and CV matrixes in one. These matrixes, including list of product attributes, consequences (benefits) and values, have been shown in figures 4-1 and 4-2.
Both AC- CV matrixes are part one of the questionnaire. The respondents were asked to checkmark the conjunction cell of any attribute and benefit or benefit and value that the linkage between them has meaning for them. For data entry, we made the ACV matrix, which rows are number of questionnaire that means each row is a consumer and columns are all the cells of AC- CV matrixes. We name all the cells (126 cells totally) and if a cell check marked, the value of that cell in ACV matrix is one otherwise it is zero. For our modeling, if the linkage between product attributes and benefits in the AC matrix exist that means the vector connecting the attribute and the benefit has the value equal to one, otherwise it will remain zero. This is the same for CV matrix. We name all the cells of AC and CV matrixes with two indexes: the code of its related attributes and consequence or its related consequence and value. The attributes coded as a1 to a11, the benefits coded as b1 to b7 and the values coded as v1 to v7. Thus for example the conjunction of a1 with b1 named a1b1. And if there was a checkmark in a1b1 cell, the value of it in ACV matrix is one, if not the value is zero. The final ACV matrix data is on the CD attached to this document.
<table>
<thead>
<tr>
<th>(N.A.B.) Malt Beverage Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benefits of consumption</td>
</tr>
<tr>
<td>Sugar Free or light</td>
</tr>
<tr>
<td>Hygienic Consumption</td>
</tr>
<tr>
<td>Good for Health</td>
</tr>
<tr>
<td>Drinking a Noble &amp; High Quality N.A.B.</td>
</tr>
<tr>
<td>Replenishment</td>
</tr>
<tr>
<td>Refreshment</td>
</tr>
<tr>
<td>Good Taste</td>
</tr>
<tr>
<td>Plastic or P.E.T Package</td>
</tr>
<tr>
<td>Bottled Package</td>
</tr>
<tr>
<td>Canned Package</td>
</tr>
<tr>
<td>Sweet Taste</td>
</tr>
<tr>
<td>Bitter Taste</td>
</tr>
<tr>
<td>Full of Foam,</td>
</tr>
<tr>
<td>Dark Golden Color</td>
</tr>
<tr>
<td>Light Golden Color</td>
</tr>
<tr>
<td>Fruit Flavor (including lemon)</td>
</tr>
</tbody>
</table>

Figure 4-1. AC (Attributes of Product-Consequences of consumption) Matrix
<table>
<thead>
<tr>
<th>Values</th>
<th>Good Taste</th>
<th>Refreshment</th>
<th>Replenishment (Taking water and energy)</th>
<th>Drinking a Noble and High Quality N.A.B.</th>
<th>Good for Health</th>
<th>Hygienic Consumption</th>
<th>Convenient Consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy Living</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fun &amp; Enjoyment in life</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comfortable life</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Being Well Respected in society</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharing pleasure with loved ones</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self Confidence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of life</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4-2. CV (Consequences of consumption - Values of Consumers) Matrix
4-3. Profiling data, for profiling the identified segments
The second part of the questionnaire is responsible for collecting the profiling data. The second part consists of questions on buying behavior and demographic variables. We coded the answers and make the profiling data. The questions are as above:

<table>
<thead>
<tr>
<th>Question</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How much you consume Malt Beverage every week?</td>
<td>cc</td>
</tr>
<tr>
<td>(A normal glass or bottle is equal to 300 cc)</td>
<td></td>
</tr>
<tr>
<td>2. What type of Malt Beverage do you drink?</td>
<td>Flavored Classic</td>
</tr>
<tr>
<td>3. What flavor do you prefer most? (flavors including: pineapple, apple,</td>
<td>1- …………</td>
</tr>
<tr>
<td>strawberry, peach, pomegranate, mango, pear, banana)</td>
<td>2- …………</td>
</tr>
<tr>
<td></td>
<td>3- …………</td>
</tr>
<tr>
<td>4. When you drink Malt Beverage mostly? at night (after meal □</td>
<td>even □ before lunch □ with main meals</td>
</tr>
<tr>
<td>5. Where do you purchase Malt Beverage?</td>
<td>Grocery Lots □ Chain stores □</td>
</tr>
<tr>
<td>6. Where do you Drink Malt Beverage?</td>
<td>Restaurant/coffee □ Sport Club □</td>
</tr>
<tr>
<td></td>
<td>Home □ Restaurant/coffee □ Sport Club □</td>
</tr>
<tr>
<td></td>
<td>Home □ Restaurant/coffee □ Sport Club □</td>
</tr>
<tr>
<td></td>
<td>Restaurant/coffee □ Sports Club □</td>
</tr>
<tr>
<td>7. Which packaging do you prefer for Malt Beverage?</td>
<td>bottle □ Can □ P.E.T □</td>
</tr>
<tr>
<td></td>
<td>bottle □ Can □ P.E.T □</td>
</tr>
<tr>
<td></td>
<td>bottle □ Can □ P.E.T □</td>
</tr>
<tr>
<td>8. Do you consider the price for purchase?</td>
<td>Yes □ No □</td>
</tr>
<tr>
<td>9. Do you purchase one brand regularly?</td>
<td>Yes □ No □</td>
</tr>
</tbody>
</table>
10. If your answer to Q9 is yes, do you consider yourself loyal to that brand?
Yes □        No □

11. Do you consume Iranian brand of malt beverage more or foreign brands?
Iranian □    Foreign □

12. How many times you go for grocery purchase?
    Everyday □    every two days □    Two times a week □
    Every two weeks □    rarely □

13. Do you like to experience new places for shopping?
Yes □        No □

14. Do you like testing new food products?
Yes □        No □

15. How do you persuade to test a new food product?
   My curiosity □    others’ proposing □
   because of Ads □

16. Do you see TV ads?
   Yes □        No □

17. Do you read newspaper regularly?
   Yes □        No □

18. What kind of magazine do you read regularly?
   Family □    Social □    Health □    □
   Coding □    Athletic □    Art □
   Politic □    economic □    Women □
   Professional □
   Youth □

19. Do you use internet daily?
   Yes □        No □

20. Your education
Figure 4-3 The buying behavior and demographic questions

The final profiling data is on the CD attached to this document.

4-4. Sample
As mentioned, the sample is university people, which include students, professors and employees. Since completing part one of the questionnaires, the AC-CV matrixes is not a common questionnaire and needed to be explained even during the completing, and it takes time, the place of questioning could not be anywhere. We had to provide respondents seats, table, pencil and also some thing to drink during completing the questionnaire. Grunert and Valli (2001) mailed the questionnaire to the sample but in this research due to time limitation, we collected the data in the field. According to Hofstede et al (1998), the sample size for conducting a survey with APT must be at least five times of total cells in AC-CV matrixes. Therefore our sample size must be at least five times of 126, which is equal to 630. We gathered 660 data.
Chapter 5:

Data Analysis and Results

5-1. Quantitative research results
  5-1-1 Exploratory Data Analysis Results
  5-1-2 Data Preprocessing Results
  5-1-3 Clustering Validation and Data Transformation- Reduction Results
  5-1-4 Clusters and their Means-end chains
  5-1-5 Feature Selection Results (on profiling data)
  5-1-6 Profiling the clusters
  5-1-7 Segments
5-1. Quantitative research results

Results have been presented according to the methodology sections.

5-1-1 Exploratory Data Analysis Results
As mentioned in the methodology chapter, the MDS plot with raw ACV matrix data (before preprocessing) helped to understand the data and showed the outliers and noises. The Figure 5-1 is the primary MDS plot of ACV matrix data, which has not been processed. The MDS plot created in R software with cmdscale. The scripts are:

```r
curmds=cmdscale(daisy(data),k=2)
plot(curmds,main="MDS (Primary)",xlab="",ylab="")
data is the ACV matrix data which has not been preprocessed.
```

![MDS (Primary)](image)

Figure 5-1. Primary MDS plot, before preprocessing the ACV matrix data. There are outliers in data.

Since the MDS plot is based on the distances between records (consumers) in ACV matrix data, and shows the records relative to each other, therefore the outliers are small or individual clusters (clusters with one member).
5-1-2 Data Preprocessing Results

- Removing Noises and Outliers

The outliers removed by clustering. The small or individual clusters were removed. The MDSD plot after removing the outliers by clustering is Figure 5.2.

![MDS (Outliers removed: % 8)](image)

**Figure 5-2.** MDS plot after removing outliers of ACV matrix data with clustering. 8 percent of data removed as outliers.

The scripts in R software:

```r
# ----- delete outliers: small clusters ----- #
{{ curdata=as.data.frame(data)
num=15
for (cnt in 1:2){
    clar=clara(curdata,num)
    clust=clar$clustering
    clus_size=clar$clusinfo[,1]
    clus_size=clus_size/sum(clus_size)
    my_index=1:num
```
curdata=cbind(clust, curdata)
curdata=curdata[curdata$clust %in% my_index[clus_size>0.02],]
curdata=curdata[,1]
}
data2=curdata
percentremoved=round((nrow(data)-nrow(data2))/nrow(data)*100)
curmds=cmdscale(daisy(data2),k=2)
plot(curmds,main=paste("MDS (Outliers removed: 
","percentremoved,"%),xlab="",ylab="")
savePlot(filename = paste(mypath,"analysis_plots\\"","MDS (Outliers removed)",".emf",sep=""),
  type = "emf",device = dev.cur(),restoreConsole = TRUE )}

data2 or curdata is data (ACV matrix data) after removing outliers which MDS plot of Figure 5-2 is based on it.

- Filling Missing values in profiling data
The scripts in R software used for filling missing values in profiling data are:

```r
#----- filling missing value for profiling data -----#
{{ dataset=profiledata
dist.mtx <- as.matrix(daisy(dataset,stand=T))
central.value <- function(x) {
  if (is.numeric(x)) median(x,na.rm=T)
  else if (is.factor(x)) levels(x)[which.max(table(x))]
  else {
    f <- as.factor(x)
    levels(f)[which.max(table(f))]
  }
}
for(r in which(!complete.cases(dataset))){
  dataset[r,which(is.na(dataset[r,]))] <-
  apply(data.frame(dataset[!is.na(names(sort(dist.mtx[r,]))[2:11])]),
        which(is.na(dataset[r,]))), 2,central.value)
} profiledata=dataset }}
```

Therefore, as described in the methodology chapter, the missing values in profiling data filled with the same value of the nearest and similar record.

**5-1-3 Clustering Validation and Data Transformation- Reduction Results**
As described in methodology chapter, two methods, TF-IDF and LSI, used to transform and reduce the dimension in ACV matrix data to have more differentiated clusters. In addition, Silhouette value introduced and selected to validate the number of clusters and method of clustering.
For understanding how effective the TF-IDF and LSI methods are, the silhouette validation was calculated before TF-IDF, after TF-IDF before LSI and after LSI with different numbers of LSI factors. The plots are figured below:

**Figure 5-3. Silhouette plot before any data transformation and reduction**

The scripts in R software:
```
# ---------- Cluster validation (Primary) ----------#
{val=clValid(data2, 2:maxclustnumb, clMethods = c("hierarchical","clara","kmeans"),
  method=clusmethod,validation="internal",maxitems = nrow(data))
plot(val, legend = FALSE,measures=c("Silhouette"),main="Cluster validation (Primary)"
legend("topright", clusterMethods(val), col = 1:9, lty = 1:9, pch = paste(1:9))
savePlot(filename = paste(mypath,"analysis_plots\","Cluster validation (Primary)\",".emf",sep=""),
  type = "emf",device = dev.cur(),restoreConsole = TRUE) }
```
Hierarchical, Clara and k-means are three methods of clustering. Clara is similar to k-means.

The primary silhouette plot is with ACV matrix data without any transformation and reduction. The maximum silhouette value in this primary plot is about .06. Comparing to 1, the ideal silhouette value, this is low. And shows that clusters with any number and any method of clustering (K-means, Clara, Hierarchical) will not be very well similar within and dissimilar between.

Now, see the figure below which is silhouette plot after data transformation with TF-IDF method.

![Cluster validation (TFIDF)](image)

It is obvious from comparing the two silhouette plots, that with TF-IDF weighting, the max silhouette value, .5, has been improved about 44%. So we can conclude that the TF-IDF weighting has been effective in transforming data in a way that better and differentiated clusters achieved. To our knowledge, this is the contribution of this research in clustering means-end chains. Since the ACV matrix data is spars like text data, we examine text mining methods for transforming our data to achieve better differentiated clusters, and by comparing these two last figures, we can claim that this TF-IDF method works for our data.
The scripts in R software:
# ---------- TFIDF (Vector space model) ----------#
{{ data3=as.matrix(data2)
dimnames(data3)[[1]]=rownames(data2) # for using row names in clustering result: record number
knormal=rowSums(data3) # because all are binary rowsum equals number of nonempty elements
a=as.matrix(data3+0.001)
b=as.matrix(colSums(a))
for (cnt in 1:nrow(a)){a[cnt,]=a[cnt,]/b}
idf=as.matrix(1+1/log(nrow(a))*colSums(a*log(a)))
for (cnt in 1:nrow(data3)){data3[cnt,]=data3[cnt,]*idf}
for (cnt in 1:ncol(data3)){data3[,cnt]=data3[,cnt]/knormal}
val=clValid(data3, 2:maxclustnumb, clMethods =
c("hierarchical","clara","kmeans"),
method=clusmethod,validation = "internal",maxitems = nrow(data))
plot(val, legend = FALSE,measures=c("Silhouette"),main="Cluster validation (TFIDF)")
legend("topright", clusterMethods(val), col = 1:9, lty = 1:9, pch = paste(1:9))
savePlot(filename = paste(mypath,"analysis_plots\","Cluster validation (TFIDF)".emf",sep=""),
type = "emf",device = dev.cur(),restoreConsole = TRUE) }}

In the next figures you can see silhouette plots after running LSI, with two to seven factor numbers, which according to literature and experiences is enough number of factors, has affected silhouette value.
The scripts in R software:
# ---------- SVD ----------#
{{ (s <- svd(data3))
for (xfactor in 2:7){
  D <- diag(s$d[1:xfactor])
data4=s$u[,1:xfactor] %*% D %*% t(s$v[,1:xfactor]) # X = U D V'
  val=clValid(data4, 2:maxclustnumb, clMethods =
c("hierarchical","clara","kmeans"),
  method=clusmethod,validation = "internal",maxitems = nrow(data))
  plot(val, legend = FALSE,measures=c("Silhouette"),main=paste("Cluster validation (SVD with ",xfactor," factors)"))
  legend("topright", clusterMethods(val), col = 1:9, lty = 1:9, pch = paste(1:9))
  savePlot(filename = paste(mypath,"analysis_plots\",paste("Cluster validation (SVD with ",xfactor," factors)",".emf",sep=""),
type = "emf",device = dev.cur(),restoreConsole = TRUE) }}
}}
Figure 5-5. The silhouette value after dimension reduction with LSI, factors no.2

Figure 5-6. The silhouette value after dimension reduction with LSI, factors no.3
Figure 5-7. The silhouette value after dimension reduction with LSI, factors no. 4

Figure 5-8. The silhouette value after dimension reduction with LSI, factors no. 5
Figure 5-9. The silhouette value after dimension reduction with LSI, factors no.6

Figure 5-10. The silhouette value after dimension reduction with LSI, factors no.7
From figures 5-5. To 5-10 we can infer that the silhouette value does not increase when the numbers of factors increase. In factor no.2 and 4 we have the maximum silhouette value which is .65 and .6 but from factor no.5 the silhouette value starts to decrease.

So the optimum factors number is 2 or 4. Also, for factor no.2, the maximum silhouette is resulted from k-means clustering method, with k=3 clusters. For factor no.4 the maximum silhouette value is resulted from k-means clustering method but with k=2 clusters. Since in clustering methods, 2 clusters are the minimum number of clusters and mostly two clusters does not reveal good information, the k-means clustering method with k=3 clusters selected.

We can conclude that the dimension reduction with LSI, improve the silhouette value even better that TF-IDF data transformation. To our knowledge, this is another contribution of this research. Since the application of text mining methods of data transformation-reduction in data mining has not been implemented before. In this research we found applying TF-IDF and LSI methods effective for clustering means-end chains resulting from APT method. These methods examines because the ACV matrix data is as spars as text data. With these two methods of data transformation-reduction, clusters will be more similar inside and dissimilar from each others.

5-1-4 Clusters and their Means-end chains

The clustering algorithms work based on similarity and the similarity measure in this research is Euclidean distance. The MDS plot of three clusters is demonstrated in Figure 5-11 (below).
In the Figure 5-11 the MDS plot of final three clusters, represents the data of each cluster with the same color. Thus we can see how data of each cluster are stick together, also their borders are near and may some similarities exist in clusters.

The scripts in R software for clustering with k-means algorithm, k (number of clusters which is bestnum in script) =3 and drawing the MDS plot of final clustering:

```r
# ----------- best parameters for cluster numbers and factors -----------#
{{ xfactor=2
  D <- diag(s$d[1:xfactor])
data4=s$u[,1:xfactor] %*% D %*% t(s$v[,1:xfactor]) # X = U D V'
data4=as.data.frame(data4)
dimnames(data4)[[1]]=rownames(data2)

bestnum=3
algorithm='K-means' #K-means','Clara','Hierarchical (ward)'
if (algorithm=='K-means') {km=kmeans(data4,bestnum);clust=km$cluster}
if (algorithm=='Clara') {clar=clara(data4,bestnum);clust=clar$clustering}
if (algorithm=='Hierarchical (ward)') {hc <- hclust(dist(data4),
  method=clusmethod);clust <- cutree(hc, k=bestnum)}

curmds=cmdscale(daisy(data4))
plot(curmds,main="MDS (Final clustering)",
  sub=paste("Clusters=",bestnum,"  -  Factors=",xfactor,"  -  Algorithm="algorithm),
  xlab="",ylab="",col=clust,pch=clust+3)
#legend("topright", 1:bestnum, col = 1:bestnum, lty = 1:9, pch = paste(1:9))
savePlot(filename = paste(mypath,"analysis_plots\","MDS (Final clustering)",".emf",sep=""),
type = "emf",device = dev.cur(),restoreConsole = TRUE)}
```

Now with three clusters; we should determine their mean-end chains. We calculate the mean amount of each feature for each cluster. Then we determine a threshold to accept or refuse a feature, which is a linkage between attribute and benefit or benefit and value, for a cluster. For determining the threshold, a histogram which has means values on x axis and the frequency of them on y axis, figured. See Figure 5-12
We did not find any suggested threshold in literature. We consider threshold as .25. We take the features of ACV matrix data that their mean is equal/above .25. This result in a binary matrix of all 126 features of ACV matrix data that for the means more that the threshold, the value of feature is one and for means less than threshold, the value of feature is zero. You can find this matrix is on the CD attached to document. The R scripts:

```r
#-------- hist_for_treshold -------------#
{{ allcolmeans=matrix(nrow=bestnum,ncol=ncol(data))
for (cnt in 1:bestnum) {allcolmeans[cnt,]=colMeans(data[clust %in% cnt,])} 
hist(as.vector(allcolmeans))
savePlot(filename =
paste(mypath,"analysis_plots\\","hist_for_treshold",".emf",sep=""),
    type = "emf",device = dev.cur(),restoreConsole = TRUE)
fivenum(allcolmeans)
boxplot(allcolmeans)
treshmeans=allcolmeans
treshmeans[treshmeans<.25]=0
treshmeans[treshmeans>0]=1
write.csv(treshmeans,paste(mypath,"analysis_plots\\","treshmeans.csv",sep=""))
allcolmeans }}
```

Therefore, the clusters’ means-end chains towards malt beverage are as follow:
Figure 5-13 Cluster number one means-end chain
Cluster 2

Values

1: Healthy Life  
2: Fun & Enjoyment in life

Benefits

3: Good for health  
4: Drinking a noble and high quality Drink  
5: Refreshment  
6: Replenishment  
7: Drinking a tasty beverage

Attributes

1: Sugar-free  
2: Including health and nutrition facts  
3: P.E.T  
4: Drinking a noble and high quality Drink  
5: Bottled  
6: Replenishment  
7: Drinking a tasty beverage  
3: Bitter  
11: Fruit Flavored

Figure 5-14 Cluster number two means-end chain
Cluster 3

Values

1: Healthy Life  5: Shared Pleasure  2: Fun & Enjoyment in life

Benefits

3: Good for health  2: Hygienic consumption  1: Convenient consumption  7: Drinking a tasty beverage

Attributes

1: Sugar-free  5: Bottled  3: P.E.T  6: Sweet  11: Fruit Flavored

Note: The blue arrows indicate the common linkages in all 3 clusters. The green arrows indicate the unique linkages in each cluster. The attribute and value which colored in red is the unique one for that cluster.

The three clusters are not fundamentally different especially clusters one and two. There are many similarities among them which are shown in the previous figures with thicker lines. “Healthy Life” and “Fun & Enjoyment in life” are two common values which are very important for members of all three clusters. In continue we describe the resemblance and difference of clusters.

At first let’s describe the value of living healthy. We can say that the healthy life is identical among three clusters. The interpretation of healthy life is as follows: Hygienic consumption (benefit 2), having a drink which is good for health (benefit 3), having refreshment by consumption (benefit 5), and drinking a noble and high
quality drink (benefit 4). The only difference among three clusters is that members of cluster 2 perceive the printing of nutrition facts and health information on the label as a sign of healthy beverage. And about the value of fun and enjoyment in life, it is also identical in three clusters from viewpoint of functional benefits of beverage. Drinking a tasty beverage and replenishment (providing water and energy for body by drinking a non-alcoholic beer) point out to having fun and joy in life. Going through tasty beer, fruity flavor in the level of functional attributes is a preference for all 3 clusters. They interpret a fruit flavored beverage as a delectable one. But a difference is between perceiving bitterness and sweetness of the beer. Members of two clusters 1 and 3 like sweet taste but members of cluster 2 like bitter taste. Variation between three clusters is as follows: clusters 1 and 3 have one extra value.

Members of Cluster 1 also have the value of comfortable life. Satisfying this value in our domain, beverage attributes, means convenient consumption in the level of functional attributes or benefits. NAB should be easy to open and easy to carry and preservation.

And in the level of product attributes cluster one members consider bottle and PET identical for convenient consumption. Members of cluster 3 have extra value of shared pleasure which means they favor drinking NAB with friend or family members. Tendency to drink with others rather than drinking alone make cluster 3 very different with the other two clusters. This value makes promotion of NAB for this cluster different. Here we should have a more group orientation rather than individualistic approach.

5-1-5 Feature Selection Results (on profiling data)
According to the random forest method described in the methodology chapter, the scripts in R software which are demonstrated below cause in selecting 14 features of profiling data out of 57 features, which has been used in profiling. Our three clusters should be differentiated in all of these 14 features but since no algorithm is free of errors, thus not all of these 14 selected features are discriminating clusters.

The scripts in R software to select features of profiling data:

```r
# ------ variable selection for profiling data -----#
library(randomForest)
library(varSelRF)
maxtimes=10
features<- matrix(nrow=maxtimes,ncol=NROW(colnames(profiledata)))
data.frame()
for(i in 1:maxtimes){
  featureselect <- varSelRF(profiledata,factor(clust),mtryFactor = 1 , ntree =200,ntreeIterat = 100,vars.drop.frac = 0.2,whole.range =FALSE, verbose = TRUE)
  for(k in 1:NROW(colnames(profiledata))){
    for(j in 1:featureselect[[5]]){
      if(colnames(profiledata)[k]==featureselect[[3]][j])
        features[i,k] = 1
      else
        features[i,k] = 0
    }
  }
}
```
features[i,k]<-"yes"
}
}
}

# ---------- counting and selecting features ----------#
{{ features2=as.data.frame(features)

for (cnt in 1:ncol(features2)){
    levels(features2[,cnt])<-c(1,0)
    features2[is.na(features2[,cnt]),cnt] <-
    vector(mode='numeric',length=NROW(which(is.na(features2[,cnt]))))
}
features3=2-data.matrix(features2)
dimnames(features3)[[2]]=colnames(profiledata)
write.csv(features3,paste(mypath,",""features.csv",sep=""))
mean(rowSums(features3))
#sort(colSums(features3),decreasing=TRUE)
profiledata=profiledata[,colSums(features3)>6] }}

5-1-6 Profiling the clusters
In continue the summary of selected features (14 features) form profiling data has been plotted for each cluster. In all plots, X axis is clusters no. and Y axis is the percentage of each feature. Figures 5-16 to 5-29 shows the 14 selected features for each three clusters.
Figure 5-16. Consumption. This figure shows the level of consumption for each cluster.

Figure 5-17. Preference for Beverage Type (Fruity or Bitter or both)
Figure 5-18. The Most Preferred Fruit Flavor, including lemon (flavor 1)

Figure 5-19. The second most preferred flavor (flavor 2)
Figure 5-20. Drinking with or without food (When 1)

Figure 5-21. Tendency to test new food products?
Figure 5-22. Regular Newspaper Reading?

Figure 5-23. Reading Family Magazines?
Figure 5-24. Education

Figure 5-25. Job
Figure 5-26. Monthly Cost of living of the family

Figure 5-27. Number of family members
Figure 5-28. Marriage Status

Figure 5-29. Number of Kids
The scripts in R software:

#--------output for each feature of profile ------#
{{ dev.off()
for (cnt1 in 1:dim(allprofile)[2]){
    thisattribute=unlist(allprofile)[,cnt1]
    thisattribute=thisattribute[!is.na(thisattribute)]

    discreetnames=vector(length=length(thisattribute))
    for (cnt2 in 1:length(thisattribute)){
        discreetnames[cnt2]=strsplit(thisattribute,"\:" )[[cnt2]][1]
        discreetnames[cnt2]=sub("[:space:]+","",discretnames[cnt2])
    }

    thisatt=matrix(data=0,nrow=bestnum,ncol=length(thisattribute))
    dimnames(thisatt)[[2]]=discretnames
    for (cnt2 in 1:bestnum){
        tempprofile=clusprofiles[[cnt2]][,cnt1]
        tempprofile=tempprofile[!is.na(tempprofile)]
        for (cnt3 in 1:length(tempprofile)){
            temp=tempprofile[cnt3]
            temp=unlist(strsplit(temp,"\:"))
            tempname=sub("[:space:]+","",temp[1])
            tempvalue=as.numeric(sub("[:space:]+","",temp[2]))
            if (tempname %in% dimnames(thisatt)[[2]])
                {thisatt[cnt2,tempname]=tempvalue}
            else
                {thisatt[cnt2,“(Other)”]=thisatt[cnt2,“(Other)”]+tempvalue}
        }
    }
    dimnames(thisatt)[[1]]=1:bestnum
    sumofrows=1/rowSums(thisatt)
    for (cnt2 in 1:bestnum){thisatt[cnt2,]=thisatt[cnt2,]*sumofrows[cnt2]*100}
    win.metafile(filename =
paste(mypath,"analysis_profile\\",names(profiledata)[cnt1],".emf",sep=""),
    width = 7, height = 7, pointsize = 12,restoreConsole = TRUE)
    barplot(t(thisatt),col = c(“lightblue”, “mistyrose”, “lightcyan”,
                        “lavender”, “cornsilk”, “red”, “black”),legend = colnames(thisatt),
    main=names(profiledata)[cnt1], font.main = 4, xlim =
c(0,NROW(thisatt)+3))
    dev.off() }}
}}
5-1-7 Segments
Based on the elicited means-end chains of clusters and profile data we can define products for each segment. Attributes and benefits are very illustrative for this. Consequently we can also identify the promotion needed for each product. For this, values are more useful. The results based on our empirical data are shown in tables 5-1. And 5-2.
## Table 5-1. Segments’ characteristics

<table>
<thead>
<tr>
<th>Perception of malt beverage based on mean-end analysis</th>
<th>Segment profile</th>
</tr>
</thead>
</table>
| **Segment one: comfort seekers**  
They find bottled or P.E.T package convenient for consuming malt beverage in any where, and this convenient consumption mean comfortable life for them.  
They also perceive bottled package as hygienic (due to glass transparency the inside can be seen, due to lid, point of contact with mouth is covered and clean) and therefore leads to a healthy life.  
They like sweet taste, But also want it sugar free because of healthy living. | They are normal and much drinkers of malt beverage prefer mostly flavored ones especially lemon and after that, pineapple, peach and apple. Some like strawberry despite of other segments. They sometimes enjoy classic malt beverage too. They are mostly graduate students and single.  
They pay attention to label information.  
They mostly drink malt beverage with meals. And enjoy testing new foods. |
| **Segment two: health conscious**  
From having the health and nutrition facts on the package, they think this product is good for health and leads them to a healthy life.  
They like bitter (classic) taste and they will like it with fruit (including lemon) flavor. Both bitter and fruit flavored, mean tasty and therefore having fun and enjoyment in life.  
They like sweet taste, But also want it sugar free because of healthy living. | They drink a lot of malt beverage, at least 3-5 bottles per week, they like bitter taste and also enjoy fruit (including lemon) flavor. Their first flavor is lemon and after that, they enjoy apple, pineapple. They enjoy pomegranate flavor more than other segments. They are mostly single. |
| **Segment three: social drinkers**  
They like sweet malt beverage because they think it is tasty and with a tasty malt beverage, they can share pleasure with their loved ones, like family and friends while drinking a tasty malt beverage with each other.¹ | They are mostly married and have at least one kid.  
They are mostly employees and have less education and less monthly cost of living than other segments.  
They read family magazines more than other segments. They enjoy flavored malt beverage. Lemon and pineapple are the most pleasant flavors for them.  
They are more likely sensitive to price. |

¹ In Iranian culture, in most sharing pleasure cases, a sweet, not bitter, food product like sweet drinks, confectionaries and even sweet meals are served.
### Table 5-2. Product and promotion for each segment

<table>
<thead>
<tr>
<th>Segment</th>
<th>Product 2</th>
<th>Promotion 3</th>
</tr>
</thead>
</table>
| **Comfort seekers** | Product A: Packaging type: Bottled, PET  
Volume: packed in individual portions (250 or 300 cc)  
MB taste: Sweet, Fruit flavored (all especially lemon, Pineapple, apple, peach, strawberry)  
Others: Sugar-free, label should include health & nutrition free, Extra cover for the lid for insuring of hygienic condition, easy open, easy to carry | Communication Message: mostly emotional, consumers should infer freedom, joy & convenience in life from drinking product A. they should also conclude that consumption is good for health. The ads itself should be entertaining & fun atmosphere.  
Price: Market price |
| **Health conscious** | Product B: Packaging type: Bottled, PET  
Volume: packed in individual portions (300 cc and 500 cc)  
MB taste: Bitter, Fruit flavored (all especially lemon, Pineapple, apple, peach minus strawberry)  
Others: Sugar-free, label should include health & nutrition free | Communication Message: emotional & rational, consumers should infer joy, excitement and refresh from drinking product B. they should also conclude that consumption is good for health. Healthy life with product B, should be emphasized by providing some information.  
Price: Market price |
| **Social drinkers**  | Product C: Packaging type: Bottled, PET  
Volume: packed in family portions (1000 cc and 1500 cc)  
MB taste: Sweet, Fruit flavored (lemon, Pineapple)  
Others: Sugar-free, easy to carry | Communication Message: mostly emotional, consumers should infer group (family) happiness & safety from health view.  
Price: Lower price comparing A and B. |

---

2 Based on attributes and benefits (consequences)

3 Based on values
Chapter 6:

Conclusion

6-1. Synopsis
6-2. Findings
6-3. Managerial Implications
6-4. Implications for Theory
6-5. Limitations of research and Implications for Future research
6-1. Synopsis

New product development is risky and costly but could provide great competitive advantage for a company. If NPD success, then the revenue comes but if it does not success, the loss could be disruptive. The common failure rate in new food product development is between 60 to 80 percent (Grunert and Valli, 2001) which is noticeable. But how to reduce the risk of failure in NPD? What factors could cause in success?

According to literature, one of the check boxes of NPD success is the consumers’ check box. If they are involved in NPD process in a systematic way, at least one source of risk has been faded.

The first phase in NPD process, idea generation and developing the product concept is important to the success of NPD and acceptance of new product from consumer. Since this is the first step, it is the most proper point to focus. We here also target this phase. And try to provide a framework to systematic involvement of consumers in the first step of NPD.

6-2. Findings

The research question was: How firms can understand preferences and requirements of customers in order to have more successful new product development?

With summing up related literature, we found the e-CKM model (Su et al, 2006), is state of the art framework for this research. The e-CKM model has wisely divided the three types of customer knowledge and identified the processes for acquiring them, and exploiting the acquired knowledge for customer oriented new product development. But for implementation this framework in real business environment, the operational procedure and methods are not clearly defined. We covered this gap by providing an operational “CKM framework for NPD”, which is presented in figure 5-1.

6-3. Managerial Implications

CKM is a combination of KM and CRM concepts. Therefore implementing a CKM framework in an organization is largely related to their processes. For successful implementation of CKM some conditions should be provided, may be reengineering of some processes is necessary:

- The organization should understand the value of customer knowledge in organizational knowledge generation. Customer should be considered as a partner and a main source of knowledge. For this, change in processes and norm of the organization is inevitable. (concept of customer-orientation in BPR).
- The customer should participate actively in knowledge creation process. Main motivation can be considering his/her needs in developing of new products (concept of personalization of product in CRM).
- The knowledge generated from CKM should be used in next stages of NPD. Diffusion of this knowledge and applying it is dependent to KM culture and infrastructure of the organization (knowledge sharing process of KM).
6-4. Implications for Theory
Since the ACV matrix data is sparse (scattered and not dense with many dimensions), using k-means clustering algorithm without any transformation, may not result in appropriate clusters. In this research, we experienced this problem. The solution is transforming data. Due to similarity of ACV matrix data with text data, we used text mining transformation technique, TF-IDF (Term Frequency – Inverse Document Frequency). To avoid curse of dimensionality, also another method of text mining, LSI (Latent Semantic Indexing) is used to reduce ACV matrix data dimension.
According to our knowledge, data transformation method, TF-IDF and dimension reduction method, LSI, which we used in this research for better clustering, is the contribution of this research to the domain of means-end chain data analysis. The mentioned methods resulted in better silhouette value.

6-5. Limitations and Implications for Future research
Our proposed framework has proved its usefulness for B2C organizations, for extension of the findings, examining framework of CKM for NPD in B2B settings is a valuable research issue.
The final stage of the framework can be checking the outcomes (product for each cluster) with participants. This stage was not addressed in this research. But it seems critical since all research findings can be validated.
For eliciting new product attributes we asked respondents to compare existing products. But for New products that have not yet prototyped or presented to market, this elicitation method is not applicable. There are some elicitation methods which are not based on product but customers requirements. Categorization and characterization of these methods could be a new research.
### Categories of Customer Knowledge

<table>
<thead>
<tr>
<th>Knowledge for Customer</th>
<th>Knowledge about customers</th>
<th>Knowledge from customers</th>
</tr>
</thead>
</table>

### Customer Knowledge Management Process

- **Eliciting "knowledge for customers"** including: product attributes, benefits and values, from themselves. Based on Means-end Chain Theory
- **Communicating the elicited "knowledge for customers"**
- **Collecting knowledge about customers' buying behavior and demographics info.**
- **Collecting customers’ knowledge about product attributes, benefits and values (means-end chains)**
- Tacit knowledge conversion

### Methods and Tools

- Any attribute elicitation method like Kelly’s repertory grid technique. Laddering to elicit benefits and values.
- Association Pattern Technique (APT) to build ACV matrixes
- By a survey of considerable and representative sample of customers
- By the same survey of ACV matrixes

- Transforming data by text mining methods: TF-IDF and LSI.
- Segmentation by clustering algorithm of data mining.

**Figure 6-1.** Our proposed CKM framework for Idea Generation Phase of NPD
References


Han and Kamber (2006) Data mining concepts and techniques, chapter 1,2,7
Janssens F. (2007) CLUSTERING OF SCIENTIFIC FIELDS BY INTEGRATING TEXT MINING AND BIBLIOMETRICS, PhD dissertation
Sally Dibb, Philip Stern, Robin Wensley, (2002), Marketing Knowledge and Value of Segmentation, Marketing Intelligence & Planning, 20:2, pp. 113-119
Appendix one
Pictures of Data gathering

Zeinab