Lifetime Value Modelling

Frederick Jacques Van Der Westhuizen

Dissertation submitted in fulfilment of the requirements for the degree Master of Science at the Vanderbijlpark Campus of the North-West University

Supervisor: Prof, PD, Pretorius

January 2009

ACKNOWLEDGEMENTS

I wish to thank Geritha Raphela senior Statistician of First National Bank for assisting me with this project. The following people made a huge contribution towards this project and they are Willie Rudolph and Karel Oberholzer. I would also like to thank my supervisor Prof. Phillip Pretorius for all of his support on this Research Project.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS II
TABLE OF CONTENTS III
LIST OF SYMBOLS AND ABBREVIATIONS VII
LIST OF FIGURES VIII
LIST OF TABLES IX
EXECUTIVE SUMMARY X
CHAPTER 1 1
INTRODUCTION AND BACKROUND 1
- 1.1 Problem Description 2
- 1.2 Research Scope 2
- 1.3 Research Objectives 3
  - 1.3.1 Acquisition Cost 4
  - 1.3.2 Churn Rate 4
  - 1.3.3 Discount Rate 4
  - 1.3.4 Retention Cost 4
  - 1.3.5 Time Period (Tenure) 4
- 1.4. Methodology 5
  - 1.4.1 Research 5
  - 1.4.2 Explore data 5
  - 1.4.3 Calculating behaviour 6
  - 1.4.4 Data Transformation 6
  - 1.4.5 Implementation 6

CHAPTER 2 7
CUSTOMER RELATIONSHIP MANAGEMENT 7
- 2.1 Introduction 7
- 2.2 When is a customer a customer? 7
- 2.3 Consider Time 8
- 2.4 Lifetime Value (LTV) 8
- 2.5 Assumptions 9
- 2.6 Retention 10
- 2.7 Cross sell, Up-Sell 10
- 2.8 Analytic Techniques 11

CHAPTER 3 12
DATA MINING 12
3.1 What Is Data Mining? 12
3.1.1 Overview of Data Mining 12
3.1.2 Business Intelligence Using Data Mining 14
3.2 Translating The Business Question Into A Data Mining Problem ... 16
3.3 Select The Appropriate Data 17
3.4 Get To Understand The Data 18
3.5 Create A Model Set 18
3.6 Data Preparation - Fixing Problems With The Data 19
3.6.1 Univariate Normality 20
3.6.2 Multivariate Normality 20
3.6.3 Homoscedasticity 20
3.6.4 Multicolinearity 21
3.6.5 Relative Variances 21
3.6.6 Outliers 22
3.6.7 Missing data 22
3.6.8 Transformations 24
3.6.9 Binary/Dummy Coding of Categorical Variables 24
3.7 Transform Data To Bring Information To The Surface 25
3.8 Building Models 25
3.9 Assessing Models 25
3.10 Deploying Models 26
3.11 Assessing Results 26
3.12 Data mining for Customer Relationship Management 26
CHAPTER 4 28
THE NEED FOR CUSTOMER CHURN PREDICTION 28
4.1. The customer lifetime value concept 28
4.2. Customer churn 29
4.3. Why is Churn Modelling Useful? 31
4.4 Increasing customer retention 32
4.5 Increasing lifetime value 32
CHAPTER 5 34
ON-LINE ANALYTICAL PROCESSING 34
5.1 What is On-Line Analytical Processing? 34

5.2. OLAP and Data Mining 34
5.3 Strengths of OLAP 35
5.4 OLAP is a Powerful Visualisation Tool 36
5.5 OLAP Provides Fast Response 36
5.6 OLAP is good for time series 37
5.7 OLAP finds clusters and outliers 37
5.8 OLAP is supported by many vendors 37
5.9 Weaknesses of OLAP 37
5.10 Setting up a cube is difficult 38
5.11 OLAP does not handle continuous values 38
5.12 Cubes become out-of-date quickly 38
5.13 OLAP does not automatically find patterns 39
5.14 When to apply OLAP 39
CHAPTER 6 40
LITERATURE STUDY: LIFETIME VALUE 40
6.1 Introduction 40
6.2 What is Lifetime Value (LTV)? 41
6.3 Why LTV? 41
6.4 Where to start 41
6.4.1 Apply the Customer Lifetime Value Concept 41
6.4.2 Benefits from Customer Lifetime Value 41
6.4.3 Identify Categories of Customer 42
6.4.4 Calculate Lifetime Value 42
6.4.5 Refine the Calculation 43
6.4.6 Analyse the Results 43
6.4.7 Use Customer Lifetime Values to Improve Marketing Performance 43
6.4.8 Set Target Customer Acquisition Costs 43
6.4.9 Allocate Acquisition Funds 44
6.4.10 Select Acquisition Offers 44
6.4.11 Support Customer Retention Activities 44
6.4.12 Increase Value with New Offers 44
6.5. Common Mistakes 44
6.5.1 Trying to Retain the Wrong Customers 44
6.5.2 Offering Customers a Limited Range of Products 44
6.5.3 Spending Too Much on Acquiring New Customers 45

6.6 Conclusion 45
CHAPTER 7 47
CASE STUDY 47
7.1 Status of Accounts 47
7.2 Sum of Accounts 49
7.3 Profit of All Active Accounts 50
7.4 Average Profit per Accounts 51
7.5 Retention 54
7.6 Churn 56
7.7 Survival Analysis 59
7.8 Lifetime Value 61
7.9 Present Lifetime Value 64
CHAPTER 8 66
CONCLUSION 66
REFERENCES 68
APPENDIX A 71
Lifetime Value Modelling: Test 71

LIST OF SYMBOLS AND ABBREVIATIONS

BI	Business Intelligence
CRM	Customer Relationship Management
LOS	Length of Service
LTV	Lifetime Value
OECD	Organisation for Economic Cooperation and Development
OLAP	On-Line Analytical Process
SAS	Statistical Analysis Software
SOM	Self-Organising Maps
ROI	Return On Investment

LIST OF FIGURES

Figure 1: Research Objectives	4
Figure 2: Cheque Accounts Structure	6
Figure 3: The Data Pyramid	13
Figure 4: The Evolution of Business Intelligence	15
Figure 5: Step by Step Approach to Data Mining	16
Figure 6: Periods Presented in a Model Set	19
Figure 7: Average Profit per Account	52
Figure 8: Retention	55
Figure 9: Churn	57
Figure 10: Survival Analysis	60
Figure 11: Lifetime Value	62
Figure 12: Present Lifetime Value	65

LIST OF TABLES

Table 1: Examples of the churn prediction in literature	30
Table 2: Status of Accounts	47
Table 3: Sum of Accounts	49
Table 4: Profit of all active Customers	50
Table 5: Average Profit per Account	51
Table 6: Retention	54
Table 7: Churn	56
Table 8: Survival Analysis	59
Table 9: Lifetime Value	61
Table 10: Present Lifetime Value	64

EXECUTIVE SUMMARY

Given the increase in popularity of Lifetime Value (LTV), the argument is that the topic will assume an increasingly central role in research and marketing. As such, the decision to assess the state of the field in Lifetime Value Modelling, and outline challenges unique to choice researchers in customer relationship management (CRM). As the research has argued, there are an excess of issues and analytical challenges that remain unresolved. The researcher hopes that this thesis inspires new answers and new approaches to resolve LTV.

The scope of this project covers the building of a LTV model through multiple regression. This thesis is exclusively focused on modelling tenure. In this regard, there are a variety of benchmark statistical techniques arising from survival analysis, which could be applied, to tenure modelling. Tenure prediction will be looked at using survival analysis and compared with "crossbreed" data mining techniques that use multiple regression in concurrence with statistical techniques. It will be demonstrated how data mining tools complement the statistical models, and show that their mutual usage overcomes many of the shortcomings of each singular tool set, resulting in LTV models that are both accurate and comprehensible.

Bank XYZ is used as an example and is based on a real scenario of one of the Banks of South Africa.

CHAPTER 1
INTRODUCTION AND BACKROUND

A fundamental principle in customer relationship management is the potential benefit to Bank XYZ of attracting and retaining their most valuable customers. This is a simple concept if a bank knows whom their most valuable customers are. But many banks take a simplistic view of measuring customer value. To really understand what banks customers are worth, banks need to think broadly about the way in which customers add value to the Bank. Furthermore banks need to create more sophisticated approaches of quantifying the value of customer relationships. Reasons why the Bank should realise sustainable value from the Lifetime Value (LTV) analysis include:

- Recognising the values of unprofitable customers;
- The benefits of gaining intelligence of customers from the data collected from devotion programs to make intelligence planning decisions, this will lead to more successful marketing, merchandising, and operations tactic; and
- Enhance the leadership viewpoint of marketing processor software, consulting and guidance selections.

In the light of the above statements, it is understandable that one of the most effective ways to determine the value of a customer is by calculating the LTV.

LTV is defined as the net present value (NPV) of the profits from a customer's relationship with a bank. It calculates how much trade the customer is likely to do with the Bank during the lifetime of his or her relationship. But not many large banks know how much trade it does with a customer at present, not even to talk about how much they would be worth in the future, customers may for example buy several different goods from different business units.

The aim of the thesis is to observe the identification of client behavioural patterns, and how these patterns can be identified. In Chapter 2, a closer look will be taken at Customer Relationship Management (CRM) and what roles Lifetime Value (LTV) and Retention plays in the CRM environment. Business Intelligence using Data Mining will be looked at and Data Mining for CRM in the LTV arena will be discussed in Chapter 3. "What is the need for customer churn?" is the question that will get a lot of attention in Chapter 4. The LTV concept, why churn modelling is useful and the increase of customer retention and LTV is the main topics with detailed explanations are also in Chapter 4. In Chapter 5, OLAP is the main topic. An in-depth look will be taken at why OLAP was the best decision for this data mining project. The benefits and the weakness of OLAP are described in detail.

1.1 Problem Description

Currently, Bank XYZ does not have a future value for customers. South Africa Banks only know what the customer present value is and have a vague idea of LTV. The aim of this study would be to quantify this value now. A new prepared LTV model needs to be formulised in order to make more efficient direct marketing motives arrived at the "valuable customer". Properly analysing and implementing a LTV model could lead to (Bets, A., Datta, P., Drew, J., Mani, P.R., 2002):

- Effective strategising which leads to efficient marketing;
- Gaining customer intelligence; and
- Recognising unprofitable customers.

Currently no future value analysis is conducted or either on behalf of or instantly within Bank XYZ.

1.2 Research Scope

In the international market, banks regardless of size are beginning to realise that one of the fundamentals of profitable growth is ascertaining and fostering a close relationship with the customer. Businesses realise now that preserving and budding existing customers is much more lucrative than centring mainly on adding customers. To come to an understanding, techniques for Customer Relationship Management (CRM) are being developed and implemented.

These techniques should support Bank XYZ in accommodating customer requirements and expenditure patterns, and help develop promotions which are not only better customised for each customer, but are also more lucrative over a longer period.

LTV is increasingly being considered a benchmark for overseeing the CRM process in order to grant benefits to and retain prestige customers, at the same time maximising takings.

Powerful and accurate techniques for modelling LTV are crucial in order to assist CRM assignments. As the LTV assignments continue it will be noted and controlled via LTV.

A customer LTV model needs to be explained and understood to a degree before it can be adopted to facilitate CRM. LTV is composed of three self-governing mechanisms namely:

- Tenure (Product Life);
- Value (Profit); and
- Churn (Closed Status).

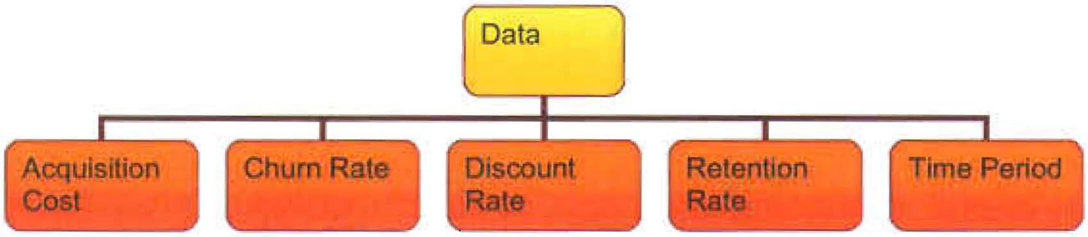
"Although modelling the profit component of LTV (which comprises of account revenue, fixed and variable costs) is a challenge in itself, experience has shown that finance departments, to a large degree, dictate this current value of a customer" (Pfeifer, P. E., Haskins, M. R., Conroy, R. M., 2005).

1.3 Research Objectives

The main objective of this research project was to determine and to develop a LTV Model through On-Line Analytical Process (OLAP).

To successfully achieve the main objective, certain sub-objectives have been identified. Refer to figure 1 which depicts those sub objectives (Japkowics N., Stephen S., 2002).

Figure 1: Research Objectives



1.3.1 Acquisition Cost

The amount of money Bank XYZ has to spend, on average, to acquire a new customer.

1.3.2 Churn Rate

The customers who end their relationship with the bank expressed as a percentage for a given time period. Churn rate typically applies to customers whose accounts have become inactive based on the Demand Deposit Account (Cheque Accounts).

1.3.3 Discount Rate

The discount rate is the cost of capital used to discount future revenue from a customer. Discounting is a sophisticated issue that is regularly ignored in customer LTV calculations. The current interest rate is sometimes used as an easy alternative for the discount rate. The discount rate won't be used in this study because our interest rate changes to often in our economy.

1.3.4 Retention Cost

The amount of money a business has to spend in a certain time frame to retain an active customer.

1.3.5 Time Period (Tenure)

Time period is the unit of time into which a "customer relationship" is calculated for examination. One year or 12 months is the most commonly used time period for all businesses who calculates the tenure. LTV is a multiple stage calculation, usually taking a minimum of 3 years and a maximum of 7 years timeframe into thought. In practice, analysis beyond this stage is seen unreliable. Time period in this thesis will be known as product life.

1.4. Methodology

In order to achieve the objectives as described above, a strategy had been identified and implemented as follows (Japkowics N., Stephen S., 2002):

1.4.1 Research

The starting block is a definition of the predicament. It is important to have a clear perceptive of what the project is about and what is expected from the outcome. Thorough research was done regarding key factors in the project. Except for better understanding of the project and problem on hand, a formal literature study was the output of this step.

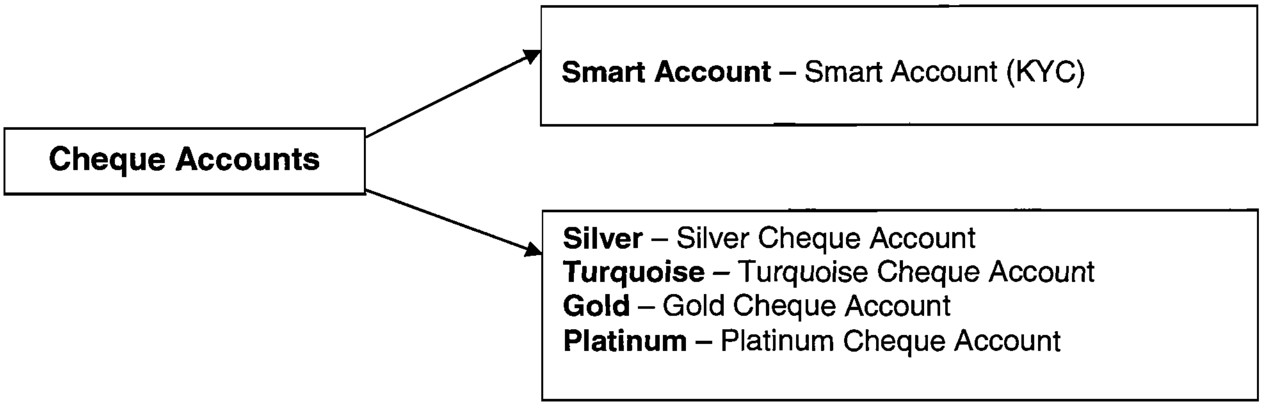
1.4.2 Explore data

This step is predominantly important since the Bank must be able to comprehend the amounts, ranges and meaning of all the quantities and codes used in the dataset.

The raw data was transferred into a format that was functional and undemanding to work with. The relevant variables were selected, cast off variables deleted and the data was summarised or rolled-up to create a summarised dataset.

To identify the problem area products needed to be assessed. Because of Cheque Accounts the product, there will now be focused on Silver (Silver Cheque Account), Turquoise (Turquoise Cheque Account), Gold (Gold Cheque Account), Platinum (Platinum Cheque Account) and Smart Account (KYC). The diagram on the next page will give the reader a better understanding.

Figure 2: Cheque Accounts Structure



1.4.3 Calculating behaviour

How is behaviour calculated? Firstly, the Bank needs to understand the data and need to know what it needs after exploring the data. To understand the behaviour of customers the Bank must look at the customers who already have churned (closed account) and only then will it see how a customer behaved. A customer churn rate was calculated and for the LTV of a customer. These calculations will be shown in the Case Study.

1.4.4 Data Transformation

Knowledge gained during data exploration was used to transform available data to fit in the theoretical foundation.

1.4.5 Implementation

Lastly, the model will be implemented by Bank XYZ.

CHAPTER 2
CUSTOMER RELATIONSHIP MANAGEMENT

2.1 Introduction

In this section the relationship marketing perspective, or marketing based on Customer Relationship Management (CRM), is discussed. The nature of relationship marketing (as compared with transaction marketing) and the strategic and tactical characteristics of a relationship and which customers are interested in relationships. The importance to service management of a relationship perspective in marketing is also described (Richards, K.A., Jones, E., 2008).

2.2 When is a customer a customer?

In a transactional approach to marketing the customer is considered a customer when he or she is the target of marketing and sales efforts. According to the relationship marketing perspective the situation is different. A relationship is an ongoing process. From time to time exchanges or transactions of goods, services, information and other utilities for money take place, but the relationship exists all the time, including in between such exchanges. Customers should continuously feel that the other party is there to help and support them, not just when they make a purchase. Therefore, once a relationship has been recognised, customers are customers on a permanent basis and they should be taken care of regardless of any miss fortune at any given time even if they make a purchase or not. Organisations which understand and truly believe in this concept and perform accordingly like this, take care of their customers as first priority (Gronroős, C., 2000).

To sum up, customers of an organisation are also customers when they do not purchase and use services, or goods, marketed by that firm. They should therefore be treated as relational customers; that is, valued customers important to the organisation. Unless customers are treated in this way the organisation does not show a genuine relational intent, even though it knows about the importance of relationship marketing and CRM in theory.

On the path to customer profitability many businesses began by attempting to accomplish product profitability. In the beginning it could be seen as an attempt at product profitability and thus better costing as an expression of an organisation's obsession on being product driven as opposed to customer focused. However, this approach was very difficult to put into operation as a product's profitability was often managed up and down, from growth throughout sales, with no measurement of costs borne latterly through a business administrative office (Richards, K.A., Jones, E., 2008).

Development had overcome this issue but the more work determined companies found it very hard because of all the products and services the business have. A personal computer may have many working pieces but it performs a special function and does not have risks attached to it. A current account service actually has a lot of tasks obtaining and bestows many of its costs day to day on the business. The considerations centre on the concepts of cost distribution and cost credit.

2.3 Consider Time

To make customer profitability energetic and to merge it into the CRM dream we need to consider the customer over a long period of time. The most important lifecycle actions that affect customer's product needs and possible profitability include the following: graduation from school, marriage, home purchase, birth of children, children leaving home to work, retirement and to finally death. Other methodologies need to be introduced to predict customer behaviours. Once the businesses receive this information, business can start integrating with predictive aptitude, this will become a valuable tools for all analytic companies. (Malthouse, E. C., Blattberg, R. C., 2005).

2.4 Lifetime Value (LTV)

The concept is derived from that a customer has a relationship of some sorts with a company over a specific period of time. There isn't any relationship with a customer if there is no relationship that is known. By taking into thought the age of a customer, the expected length of their relationship with the business, demographics and possible future products they might purchase.

To understand the cash back difficulty today, discounting techniques such as net present value (NPV) are used. For example, let's say there is a customer who produces R3000 profit every year for the following 10 years. If the present discount rate is 8%, then the customer LTV will be (Pfeifer, P. E., Haskins, M. R., Conroy, R. M., 2005):

$LTV = {}^{i=10} S_i \, 3000/1.08^{\,i} = 20{,}130.25$

$LTV = 3000/1.08 + 3000/1.08^{2} + 3000/1.08^{3} + 3000/1.08^{4} + 3000/1.08^{2} + \ldots. \; 3000/1.08^{10}$

The customer is worth R 20,130 to the company to date! (Kim, S.H., Ko, E., Kim, M., Woo, J.Y., 2007)

2.5 Assumptions

If one wants to calculate the upcoming value of a customer's tenure, assumptions need to be made. An assumption concerning the length of time, a customer is expected to linger with a company if one could include their personal information in the calculation for instance: age, their lifestyle, occupation, geographic location and income. If this information is gathered assured forecast calculations have to be made concerning the types of services the customer is likely to acquire; the profit that will be derived from those services chosen; and the fees related with marketing and providing those services. An analysis based on active customers using certain products and services, predictions are made for customers who have the same behaviour.

Current customer characteristics that will influence the analysis for future predictions are:

- Length of the association;
- Account balances;
- Default rates;
- Customer's relationship tenure;
- Product or service possessions; and
- Capability to pay.

One can start by identifying these characteristics for each customer and what actions have the prospective to change a customer's value to the business. Product development and pricing strategies are of the most important, cross-sell, up-sell and cost structures, can impact the customer's value good or bad, as can the retention treatment of a customer. Profitability methods joined with LTV calculations enable a company to improve the profile of potentially profitable customers and to identify similarities amongst its potential customers. A company can use this information they have obtained to retain its customers (Malthouse, E. C., Blattberg, R. C., 2005).

2.6 Retention

Those customers who had been identified as profitable, a company really must retain them and can commence detailed customer retention treatments because they are the blue eyed boys of the company. Why so much trouble to go trough one might ask? Because it cost five or six times as much or even more to find new customers. A company can spend a bit more on retention than on acquiring new customers.
By knowing current and profitable customers a company can amend actions to their less profitable customers (Gronroős, C., 2000).

There is a lot of competing services available currently; the focus need to be applied on customer satisfaction continuously. In the business world, a customer loyal to the company must be rewarded and customer churn needs to be minimal. The big problem with the banks is that they tend to respond after a customer decided to leave. To change the customer's mind at this time is nearly impossible. If proper data mining was done this could have been avoided. If a proper retention model was build and frequently run then one could have seen a customer's past service usage, service performance, spending and other behaviour patterns. This would have given a company a good prediction which customer was likely to churn. Data mining can predict all sorts of customer behaviour accurately (Richards, K.A., Jones, E., 2008).

2.7 Cross sell, Up-Sell

To add to their split of the customers wallet a lot of companies consider that a profundity of products taken means turnover. (Pfeifer, P. E., Haskins, M. R., Conroy, R. M., 2005) If a person spends a lot of money on unprofitable products or services that still doesn't make the business profitable. However if equipped with profitability and LTV information a business can cross-sell, up-sell and retain their customers. The possibility to move that slightly profitable customer into profit does exist but then a company needs to understand the business dynamics for this to happen.

If there is a decision or likely an action, there would follow a pattern or behaviour, out of this cash flows and risk could be predictable. As soon as a cross sell or up-sell model is build costs will start mounting up, now management need to make a decision if this is feasible enough and will the company benefit out of this. The short answer is yes (Gronroős, C., 2000).

2.8 Analytic Techniques

If all business, data and governance issues are solved then and only then may a business start to develop, organise and adjust its marketing techniques to attract more customers to the business. *Successful processes* (people and technology processes) are of the at most importance because this will use and control all the tools and techniques. This is the corner stone for the analytics before a model can be developed (Gronroős, C., 2000)

CHAPTER 3
DATA MINING

3.1 What Is Data Mining?

Data mining is the analytical discovery tool of patterns, associations, changes in data, differences, statistically significant structures and events in data. Data mining tries to gain knowledge from data.

Data mining is very different form the traditional statistics in several ways:

- Normal statistics is statement driven in the sense that a hypothesis is created and validated against the data;
- Data mining is detection driven in such a way that patterns and hypothesis are automatically extracted from data; and
- The one part of statistics that data mining resembles the most is, is exploratory data analysis.

"Data mining differs so much from traditional statistics that sometimes the goal is to extract qualitative models which can easily be translated into logical rules or visual representations; in this sense data mining is human centred and is sometimes coupled with human-computer interfaces research" (Van den Poel, D., Larivi`ere, B., 2004).

Data mining's first step is always to start with raw data that have been cleaned. Answers of the data mining process include the following:

- Insights;
- Rules; or
- Predictive models.

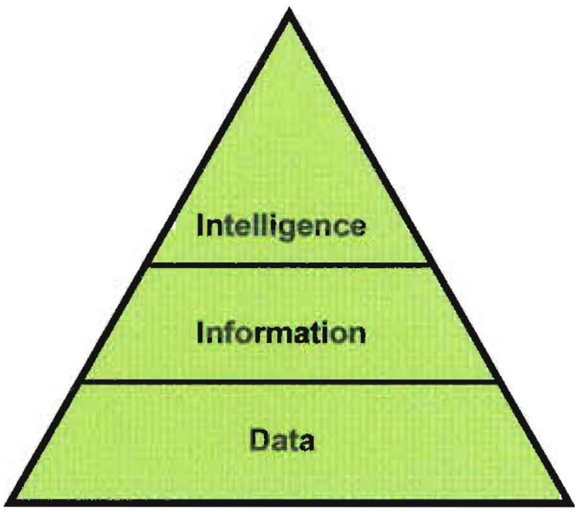
Data Mining feeds from several roots, including statistics, machine learning and databases.

3.1.1 Overview of Data Mining

To get a single answer for data from the data warehouse or data mart, many businesses are looking towards data mining, a new generation of advanced software solutions. Data mining combines several of advanced techniques to explore huge amounts of data and discover relationships and patterns that answer in-depth questions business were looking for. Companies can use these answers to do more effective marketing to increase their profit (Van den Poel, D., Larivi`ere, B., 2004).

Data mining was intended for exploiting huge amounts of data. This process would be very efficient if the Bank knows what the business problem is, and then determine the amount of data it will need to answer the questions the banks may ask. Taking this approach, data mining can solve certain business problems and the potential ROI (Return on Investment), the process will be more goal focussed.

Figure 3: The Data Pyramid



(Novo, J. 2000)

Banks employ data mining to venture and to plot relationships on a model from huge amounts of data what is supplied by the data warehouse. If there is no group of authenticated and cleansed data that a data warehouse provides, data mining would be extremely difficult and the road to finish your model in time would be long and winding. The Internet is also a very good source of data where one can found a "data warehouse". Companies might as well use the internet for findings, analyse it and distribute it via an intranet.

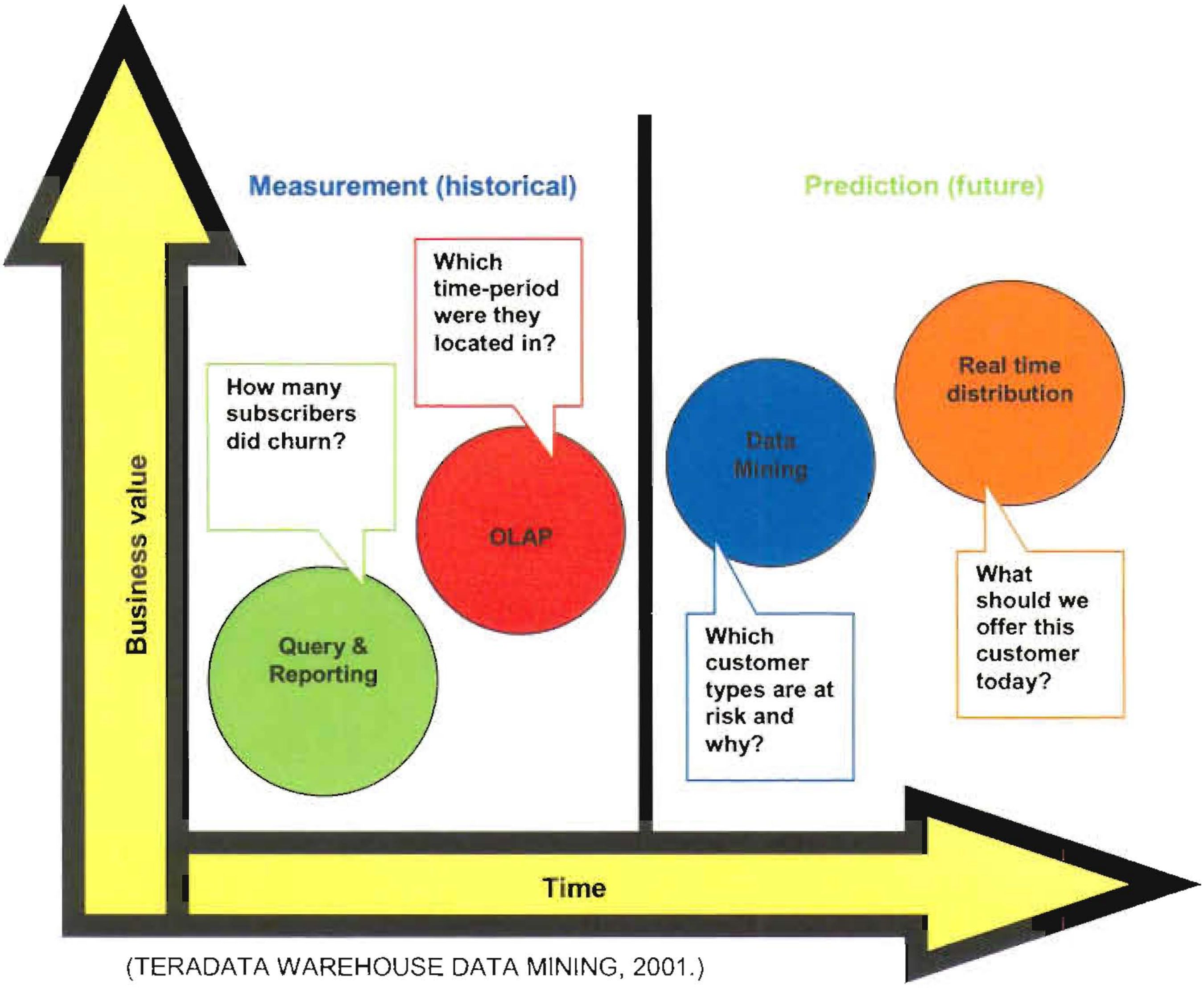
The better the computer hardware the easier the processing becomes. This would make the life of an analyst much easier and time computing will also be much faster.

Program software advances will continue the development of data mining. If all the software enhances is in order, banks can start exploiting these immense stores of data in the warehouse, new modelling tools and techniques could be developed through data visualisation, neural networks, and decision trees.

3.1.2 Business Intelligence Using Data Mining

Banks normally commence their business intelligence (BI) voyage with the spotlight on accepting and analysing the results of past decisions. But the past results can't give the bank an understanding what is going to happen in the near future bur they do give a broad view of the road behind. Banks are understanding that future predictions through business intelligence is very important to make improved decisions that unravel business problems and keep their businesses moving forward on a profitable road. Data mining can't wait to tell the Bank what is most probable going to happen giving the situation today to advance in the future.

Figure 4: The Evolution of Business Intelligence

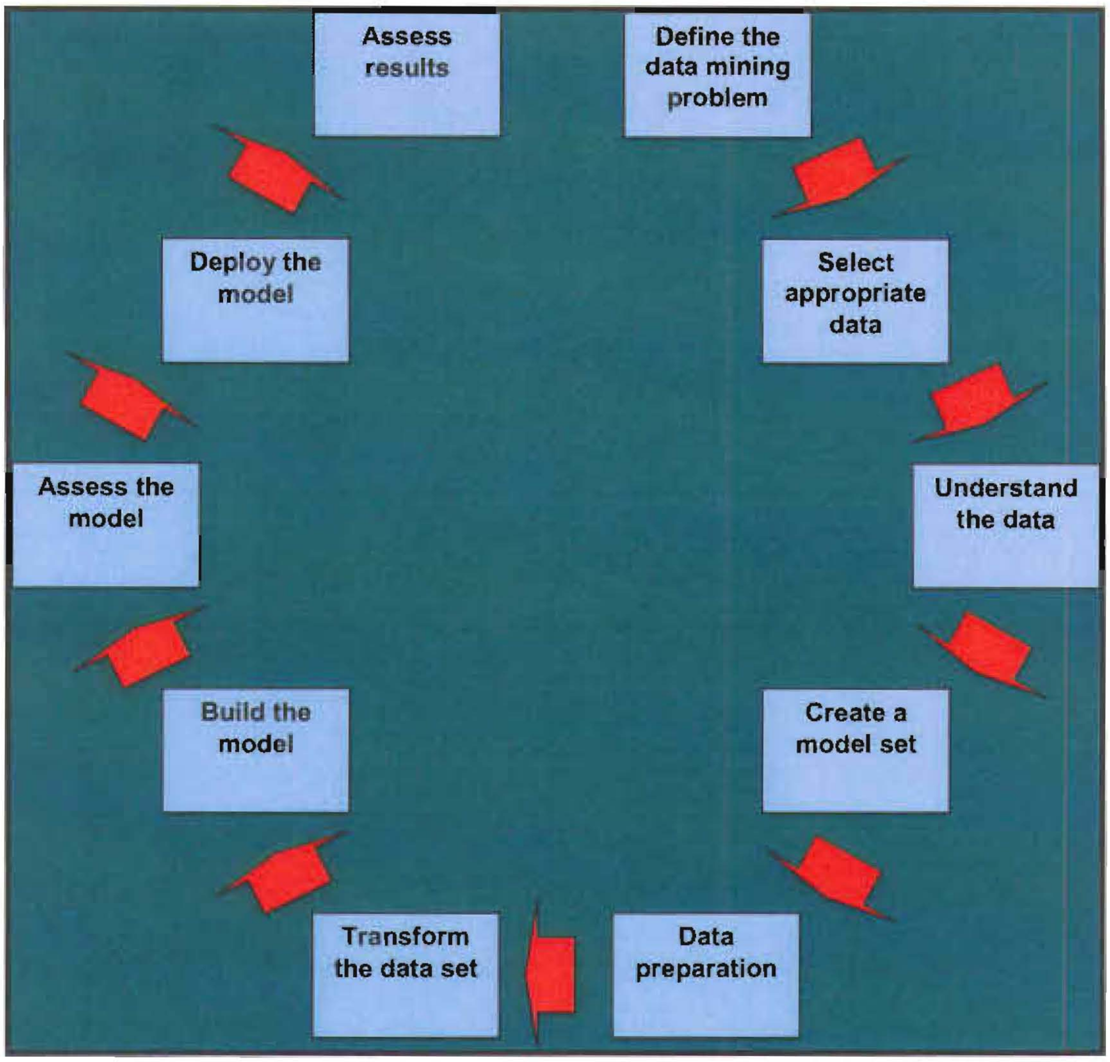


(TERADATA WAREHOUSE DATA MINING, 2001.)

Figure 4 shows how a wireless telephone company has evolved their BI, a growth to answering the questions that affect potential profits. They start with reporting that gave the company straightforward capacity. OLAP the tool is added to recover more detail in the data. Centre their BI on the vision through data mining. And finally, deploy data mining answers to their front lines to continually improve Return on Investments (ROI).

Data mining integrated all of these tools and techniques into a regular, iterative method. All of this will be shown in figure 5 below:

Figure 5: Step by Step Approach to Data Mining



<u>3.2 Translating The Business Question Into A Data Mining Problem</u>

"Without some way of recognising the destination, you can never tell whether you have walked long enough". The solution for a data mining problem is a well defined business problem. A section is included in this document detailing different business problems and how to tackle them (see section 3.3). Data mining goals for a particular project should not be stated in broad, general terms, such as "gaining insight into customer behaviour". This is an interesting task but it is hard to measure, thus general goals should be broken down to specific ones to make it easier to monitor progress in achieving them. Gaining insight into customer behaviour might turn into concrete goals like (Kim, S.H., Ko, E., Kim, M., Woo, J.Y., 2007):

- Identify customers who are likely to refinance their loans;
- Rank order all customer based on propensity to take up a loan; and
- List products whose sales are at risk if we discontinue short term loans.

To translate a business problem into a data mining problem, it should be formulated as one of the six tasks mentioned earlier. The other most important question on deciding how to best translate a business problem into a data mining problem is how the results will be used - different intended uses entail different solutions.

3.3 Select The Appropriate Data

The data sources that are useful and available vary from problem to problem and industry to industry. Often the question is regarding how much of the data is sufficient for the data mining procedure? The answer depends on the algorithm employed, the complexity of the data and the relative frequency of the target / outcome. Use as much data so that the training, validation and test sets each contain many thousands of records. Often the target variable represents something that is relatively rare.

N.B When the model set is large enough to build a good, stable model, making it larger is counter-productive - since data mining is an iterative process, the time spent waiting for results can become very long if each run of a modelling routine takes hours instead of minutes. Use a sample of the data set to build the model (Olafsson, S., Li, X., Wu, S. 2008).

The rule of thumb is that in your sampling, for every observation with a target variable in the affirmative of the experiment, select four observations for the others. An example is when one is building an attrition model and, for example, you have 400 already closed accounts, then one should randomly sample 1600 open accounts for the modelling set so as to balance the model set. Although this creates bias in your sample data set as this is not representative of the actual population, one can correct for this bias after building the model (Haenlein, M., Kaplan, A.M., Beeser, A.J., 2007).

The other question is how much history is needed for the model data set. The first question is to consider seasonality, for example distortions due to Christmas shopping. There should be enough data to capture periodic events of this sort or, otherwise, one can correct for seasonality in the data set before modelling. There is a code which corrects for seasonality that is readily available. One should be careful not to get data from the distant past as this might not be of any use due to changing market conditions. This is especially true when some external event like the N.C.A has intervened.

In terms of the number of variables to choose, data mining calls for letting the data itself reveal what is and what is not important. It takes experience to carefully choose the variables that seem unlikely or not interesting. Often, variables that have previously been ignored turn out to have predictive value when used in combination with other variables. A rule of thumb is that your data set should contain 10 times the number of observations as there are variables (Olafsson, S., Li, X., Wu, S. 2008).

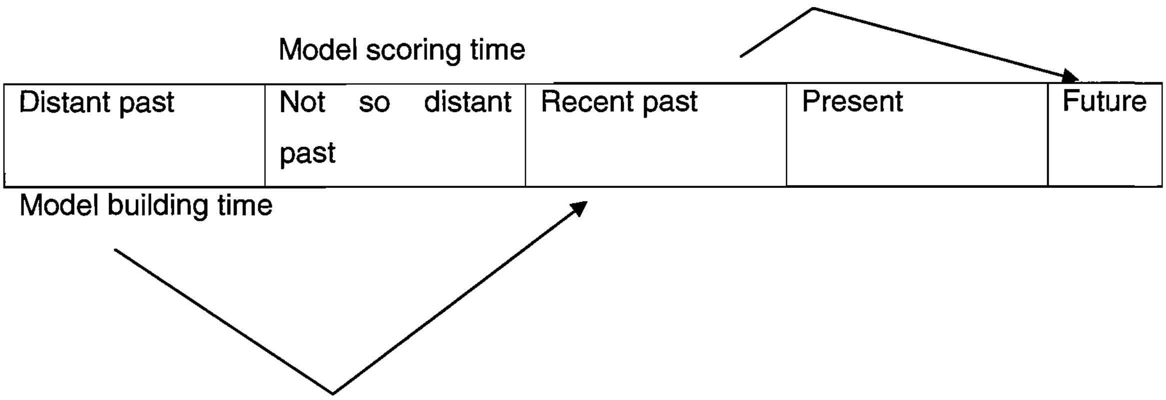
3.4 Get To Understand The Data

Spend time exploring the data - examine distributions of each variable individually and also against the target or response variable. Examine the histogram of each variable in the dataset and make notes of any outstanding features. An example is finding the gender variable with a category named African. Closely identify any unexpected patterns. Are there any missing values and how many are there? What is the likely cause of these missing values? Also pay attention to the range and the height in distribution of each variable (Kim, S.H., Ko, E., Kim, M., Woo, J.Y., 2007). Ask questions like; is the mean far from the median? Are there any negative values where there are not supposed to be? Have the variable counts been consistent over time? Look at the variables of each distribution and compare them with the description given for that variable in available documentation. Use simple cross- tabulations and visualisation tools such as scatter plots, bar graphs est., and validate assumptions about the data.

3.5 Create A Model Set

First determine the level at which the analysis will be carried out e.g. either at customer level or at account level. Also determine the type of sampling to be done to create a balanced sample. The most commonly used one is the simple random sampling. The primary goal of the methodology is to build stable models - this means that the model should work at any time of the year. Account for different time periods and seasonality so that the model generalises from the past rather than memorises.

Figure 6: Periods Presented in a Model Set



As shown above (Haenlein, M., Kaplan, A.M., Beeser, A.J., 2007), all of these three periods (past, present and future) should be represented in the model set. Predictive models are built by finding patterns in the distant past to explain the outcomes in the recent past. When the model is deployed, it is then able to use data from the recent past to make predictions about the future. Data from the immediate past is not used so as to use it for the testing of your model.

Once the pre-classified data has been obtained, the data now has to be partitioned into three sets, namely, a training data set, a validation data set and a test data set. The training set is used to build the initial model, the validation set is used to adjust the initial model to make it more general and less tied to the idiosyncrasies of the training set. The third part, the test set, is used to gauge the likely effectiveness of the model when applied to unseen data. A test set from a different time period, often called an out of time test set, is a good way to verify model stability, although such a test set is not always available.

3.6 Data Preparation - Fixing Problems With The Data

Below is a list of some data issues, common model assumptions and corrective measures (Ferreira J., Vellasco M., Pachecco M., Barbosa C., 2004). It is imperative that the assumptions associated with the statistical model intended for use are known.

3.6.1 Univariate Normality

Definition

- The data point/values of a single, continuous variable with a normal distribution.

Notes

- This is a common assumption for linear statistical models.

Diagnostics

- Kolmogorov-Smirnov test
- Informal: Distribution/Histogram plot should look like a bell.

Remedial measures

Transformations

3.6.2 Multivariate Normality

Definition

1. All univariate distributions are normal;
2. The joint distribution of any pair of variables is bivariate normal and
3. All bivariate scatter plots are linear and homoscedastic.

Notes

- It is difficult to determine.
- Fortunately, many instances of multivariate non-normality are detectable through inspection of univariate distributions.

Remedial measures

- Deletion of cases that are outliers contribute to multivariate normality.
- Transformations

3.6.3 Homoscedasticity

Definition

- Uniform variances across all levels of a variable

Notes

- It may be caused by non-no normality in variable X or Y.
- This is also a common assumption for linear statistical models.

Diagnostics

- Modified Levene test
- Breusch-Pagan test

- Informal: Evaluate through inspection of bivariate scatter plots

Remedial measures

- Transformations

3.6.4 Multicolinearity

Definition

- Happens when inter-correlations between some variables are consequently high (r_{xy}>0.85) that several mathematical operations are either impossible or unsteady since some denominators are zero or close to zero, i.e. different variable seem to measure the same thing

Notes

- This is also a common assumption for linear statistical models

Diagnostics

- Informal: Evaluated correlation matrix to find large correlation coefficients.
- Calculate squared multiple correlation between each variable and all the rest; $R_{smc}^2 > 0.9$ suggests multicollinearity
- Tolerance = 1 - R_{smc}^2 and indicates the proportion of total standardized variance that is unique (i.e., not explained by all the other variables); Tolerance < 0.1 may indicate multicollinearity
- Variance inflation factor (VIF) = 1/Tolerance = $1/(1 - R_{smc}^2)$; VIF > 10 suggest that a variable may be redundant

Remedial measures

- Eliminate redundant variables.
- Combine (i.e. sum, use principal component analysis) collinear variables into a composite variable and keep only the composite variable.

3.6.5 Relative Variances

Notes

- It may be caused by non-normality in variable X or Y.
- This is also a common assumption for linear statistical models.

Diagnostics

- Informal: Inspect covariance matrix; the largest variance of any variable should not be more that 10 times larger than the smallest variance of any other variable.

Remedial measures

- Standardise variables (i.e. (X – average)/standard deviation) or just multiplying the variable by a factor (i.e. X*15 or X*(1/15))

3.6.6 Outliers

Definition

- A dimension which lies in an extreme position compared to the other measures in the data set; two different types: univariate- and multivariate outliers.

Diagnostics

- Univariate
    - Informal: Evaluate frequency distributions and z-scores to find extreme values
    - Rule of thumb: Any value further away from the mean than 3 standard deviations is an outlier

- Multivariate
    - Mahalanobis distance (D) statistic
    - DFFITS statistic
    - Cook's Distance Statistic
    - DFBETAS Statistic

Remedial measures

- Delete record/case/observation

3.6.7 Missing data

Notes

- Attempt to understand the nature of the underlying data loss mechanism

Remedial measures

- Determine whether you want to replace the missing values or not; if not, create additional variables, one for each variable with missing value, in which you code incomplete cases with regards to that variable as value 1 and complete cases as 0; force new variable into model if associated original variable is selected, i.e. both of the variables should be selected or the new variable, not only the original variable.

- General categories of methods to impute/replace/handle missing values:
  - Available cases methods
    - In list wise deletion, cases with missing values on any variable are excluded from the analysis
    - In pair wise deletion, cases are excluded only if they have missing data on variables involved in a particular computation (this method is not recommended due to problems associated with it)

  - Single imputation
    - Mean substitution - Replaces missing data with the overall sample mean (simple, but tends to distort underlying distribution by reducing variability and making the distribution more peaked at the mean)
    - Regression-based imputation - Replace missing data with predicted value generated from multiple regression based non-missing values on other variables
    - Pattern matching - Replace missing data with a value from another case with similar profile or values across other variables
    - Random hot-deck imputation - Separate complete and incomplete cases, sort both sets of records so that cases with similar profiles on variables known to be related to the variable with the missing values are grouped together, randomly interleave the incomplete cases among the complete cases, replace the missing values with those on the same variable from the nearest complete record

  - Model-based imputation (sophisticated)
    - Replace missing values with one or more imputed (estimated) values from a predictive distribution that explicitly models the underlying data loss mechanism
    - Expectation-maximisation (EM) algorithm (related method) - In the E (estimation) step, missing values imputed by predicted values in a series

of regressions where each missing variable is regressed on the remaining variables for a particular case. In the M (maximisation) step, the whole imputed dataset is then submitted for maximum likelihood estimation. These two steps are repeated until a stable solution is reached across M steps.

3.6.8 Transformations

Definition

- Corrective measure for non-normality and other data issues; useful for dealing with outliers

Notes

- It may be necessary to try several different transformations before finding the one that works for a particular distribution

Possible transformations

- Square root ($X' = X^{1/2}$)
- Squared ($X' = X^2$)
- Cubed ($X' = X^3$)
- Logarithmic ($X' = \log X$)
- Inverse ($X' = 1/X$)
- Box-Cox – a power transformation which seeks to find the optimal power to transform a variable to be a close as possible to normality
- All transformations above can be applied on ($X - \max(X) + 1$) to remedy negative skewness.
- Odd root ($X^{1/3}$) and sine functions tend to bring outliers in from both tails of the distribution towards the mean
- Odd-power polynomial transformations (X^3) may help for negative kurtosis

3.6.9 Binary/Dummy Coding of Categorical Variables

Definition

- A method used to handle categorical variables in regression models

Notes

- If the categorical variable has N levels (possible values), only N-1 dummy variable should be included in the model, or the beta coefficients cannot be mathematically computed; the other dummy variable (any of the possible N) is used as a reference.

Method

- All values of the categorical variable is characterised in the model with an indicator variable. Every indicator or dummy variable includes only the values 1 and 0, with a value of 1 indicating that the observation related to the indicator has the given categorical value.

3.7 Transform Data To Bring Information To The Surface

Once the data has been assembled and major data problems fixed, the data must still be prepared for analysis. This may include adding derived variables to bring information to the surface. Most of the bank's data contains time series but most of the data mining algorithms do not understand time series data. Signals such as "three months of declining account balance" cannot be spotted by treating each month's observation independently. It is up to you as the data miner to bring trend information to the surface by adding derived variables e.g. ratio of balance in one month to the same month last year for a long term trend (Olafsson, S., Li, X., Wu, S. 2008).

Adding fields that represent relationships considered important by experts in the field is a way of letting the mining process benefit from that expertise.

More often, one will find data sets with counts or rand amount which might not be interesting themselves because they vary according to some underlying value e.g. comparing account balances for one person who initially deposited R10, 000 and the other who deposited R100, 000. It is more interesting to compare the proportion of account balance with respect to the deposit a person made. It is noteworthy to remember that patterns one finds determine correlation not causation.

3.8 Building Models

Please refer to sections 3.4.and 3.5 of this chapter for an overview of the alternative statistical techniques that may be utilised to build models.

3.9 Assessing Models

This is where you determine whether or not the models are working. A model assessment should answer questions like:

- How accurate is the model?
- How well does the model describe the observed data?
- How much confidence can be placed in the model's prediction?
- How comprehensible is the model?

The answers depend on the type of the model that was built. Assessment here refers to the technical merits of the model.

3.10 Deploying Models

This involves moving the model from the data mining environment into the scoring environment. To deploy the model, a programmer takes a print description of the model and recodes it in another programming language so it can be run on the scoring platform.

A common problem is that the model uses input variables that are not in the original data. This should not be a problem since the model inputs are at least derived from the fields that were originally extracted to from the model set. It is always important to keep a clean, reusable record of the transformations that were applied to the data.

3.11 Assessing Results

The measure that counts the most is the return on investment. Measuring lift on a test set chooses the right model. Profitability models based on lift will help decide how to apply the results of the model. However, it is important to measure these things as well. In a database marketing application, this requires always setting aside control groups and carefully tracking customer response according to various model scores.

3.12 Data mining for Customer Relationship Management

The expression "customer relationship management" is on the lips of every CEO and director now days. The term was solely invented for the purpose of one-on-one marketing; with this ideas were formed for high revenue sales. First class customer relationship management means a company has a bond with the customer, their needs will be put first and products and services will be supplied that will fit their pockets. CRM requires a lot of understanding, understanding the customer, understanding the customer's needs and understanding the customer's finances.

Data mining plays a very big role, if a proper model is build, this "understanding problems" can be sorted out before the customer come and see the bank manager (Kim, S.H., Ko, E., Kim, M., Woo, J.Y., 2007)

The field that has come to be called data mining has grown from several antecedents. On the academic side are machine learning and statistics. Machine learning has contributed important algorithms for recognising patterns in data. Machine learning researchers are on the bleeding edge, conjuring ideas about how to make computers think. Statistics is another important area that provides background for data mining. Statisticians offer mathematical rigor; not only do they understand the algorithms they understand the best practices in modelling and experimental design.

CHAPTER 4
THE NEED FOR CUSTOMER CHURN PREDICTION

The data in this thesis was provided by Bank XYZ. Within the personal retail banking sector a long-term customer management strategy should be adopted, as result of the fact that customers in the earlier stages of the life cycle is not reasonably profitable. In this regard, customer lifetime value analysis is created in order to understand customer behaviour and the customer sector.

4.1. The customer lifetime value concept

"The customer lifetime value is usually defined as the total net income from the customer over his lifetime." This customer analysis is done with the following in mind (Neslin, S. A., Gupta, S., Kamakura, W., Junxiang, L., Manson, C. H., 2006):

- Customer value;
- Customer lifetime value;
- Customer equity; and
- Customer profitability

The idea in LTV concept is easy and very measurable; the lifetime value calculation is straightforward after the customer relationship is over. The challenge is this: What is the concept, how is it defined and how will it be measured during the LTV, or even before, the active stage of customer relationship?

A conceptual LTV model is defined as follows:

"LTV is the total value of direct contributions and indirect contributions to overhead and profit of an individual customer during the entire customer life cycle that is from start of the relationship until its projected ending."

Most LTV models begin from a basic equation. The components of the basic LTV model are as follows (Neslin, S. A., Gupta, S., Kamakura, W., Junxiang, L., Manson, C. H., 2006):

- The customer net present value over time (revenue and cost);
- Retention rate or Length Of Service (LOS); and

- Discount rate.

Each of these components could be measured separately and afterwards this is done, it can be combined for the LTV model. The understandings of the benefits of the customer lifetime value are plentiful, as is mentioned in the following (Neslin, S. A., Gupta, S., Kamakura, W., Junxiang, L., Manson, C. H., 2006):

- Present and the future income are calculable for the customers;
- A business can also encourage customer retention and loyalty, this will mean higher profitability;
- The LTV examination will help a company on their selection of products and certain services they want to offer;
- This understanding of the customer value will help the business focus on profitable customers; and
- The LTV is not permanent; the significance can be influenced by marketing.

4.2. Customer churn

The focal point on customer churn is to establish which customers are at risk of churning and is it worth keeping these customers at the bank. The banking world and customer relationship influences the end result of how churning customers are identified. From a data analysis perspective, declining transactions is one of the main variables indicating potential attrition. On the other hand, for example, mobile industries, a customer can switch of mobile networks to another and keep their mobile number. (Au W., Chan C.C., Yao X., 2003).

The customer churn is related to the retention rate and loyalty of a customer. LTV propose that the churn rate of a customer has a brawny influence to the LTV value because this will have an influence on the length of service and the potential profit. Customer loyalty is defined plain and simple, a customer will stay with a company no matter what. Churn is described as a customer who ended his or her relationship with the company. (Van den Poel, D., Larivi`ere, B., 2004).

Modelling customer churn with not a lot of data is not suitable for LTV because the retention function tends to be "spiky" and non-smooth. Marketing is very important because this will supply the adequate information about the probability or the possibility of churn. This will enables the marketing department to contact the high probability churners first.

Table 1: Examples of the churn prediction in literature

Article	Market Sector	Case Data	Methods Used
Au et al.	Wireless telecom	100 000 subscribers	DMEL method (data mining by evolutionary learning)
Buckinx et al.	Retail business	158 884 customers	Logistic regression, ARD (automatic relevance determination), decision tree
Buckinx et al.	Daily grocery	878 usable responses	MLR (multiple linear regression), ARD, and decision tree
Ferreira et al.	Wireless telecom	100 000 subscribers	Neural network, decision tree, hierarchical systems, rule evolver
Garland	Retail banking	1 100 customers	Multiple regression
Hwang et al.	Wireless telecom	16 384 customers	Logistic regression, neural network, decision tree
Mozer et al.	Wireless telecom	46 744 subscribers	Logistic regression, neural network, decision tree

Table 1 presents examples of the churn prediction studies found in literature. The methods used for churn analysis are shown as well as the case data size and target market information (Buckinx W., Van den Poel D., 2005.) The studies above mentioned measures the loyalty and churn rate in a retail enviroment. The studies above establish which customers are likey to churn and which customers are likely to be retained. This is possible analysis because the studies focus on only on loyal customers (Buckinx W., Verstraeten G., Van den Poel D.,2005.).

The retail banking sector is the perfect market sector where the studies can conclude because the studies will show that customers is not regularly switching from one bank to another. This makes customer churn a priority for the banking sector. (Garland et al, .2005:12) has done fantastic research studies on customer profitability in retail banking.

They focus mainly on customer's value to the bank, they also focus on the length and age of customer relationship based on the profitability of the customer (Van den Poel, D., Larivi`ere, B., 2004).

4.3. Why is Churn Modelling Useful?

With a definition of churn, lots of data, and powerful data mining tool we can develop models to predict the likelihood to churn. The key to successful data mining is to incorporate the models into the business.

Because this was a real project, we can admit one of the primary business drivers was an executive who insisted on having a churn model by the end of the year. His reasoning was simply that churn is becoming a bigger and bigger problem and well-run cellular companies have churn models. He wanted his company to be the best (Kim, S.H., Ko, E., Kim, M., Woo, J.Y., 2007).

Fortunately, there are many good reasons for churn models besides satisfying the whims of executive management (even when they are right). The most obvious is to provide the lists to the marketing department for churn prevention programs. Such programs usually consist of giving customers discounts on air time, free incoming minutes, or other promotions to encourage the customers to stay with the company. For the case study, the cellular company belonged to a conglomerate, and their promotions offered products from sister companies that were not at all related to telephone usage (Hwang H., Jung T., Suh E., 2004).

Other applications of churn scores are perhaps less obvious. Churn is related to the length of time that customers are estimated to remain; that is, the customer lifetime. The idea is simple: If a group of customers have a 20 percent chance of churning this month, then we would expect them to remain customers for five months (one month divided by 20 percent). If the churn some suggested a churn rate of only 1 percent, then we would expect the customers to remain for one hundred months. The length of the customer lifetime can then be fed into models that calculate customer's lifetime revenue or profitability (also called lifetime customer value).

Churn models have an ironic relationship to customer lifetimes. If the churn model were perfect, then the scores would either be a 100 percent chance of churning in the next month, or a 0 percent chance. The customer lifetimes would then be either one month or forever (Mozer M. C., Wolniewicz R., Grimes D.B., Johnson E., Kaushansky H., 2000). However, because the churn model is not perfect, it can provide insight into the length of customer lifetimes as well.

4.4 Increasing customer retention

Traditionally, most marketing theory a practice centres on attracting new customers rather than retaining existing ones. Obtaining a new customer costs more than retaining one, the company must always remember to customer satisfaction frequently. "The golden key to customer retention is customer satisfaction. Banks must remember the following that a satisfied customer stays loyal longer, buys more products or services, pays less attention to competing products or services, is less sensitive to price, offers product or service ideas to you, and costs less to serve than new customers".

While analysing customer defection, some basic questions to ask says Kim are (Kim, S.H., Ko, E., Kim, M., Woo, J.Y., 2007):

1. What are the retention norms for our industry?
2. What is the relationship between retention rates and changes in prices?
3. Where to lost customers go for the same products or services?
4. Is there a cyclical pattern for customer defection?
5. Does defection rate vary by region or sales representative or distributor?
6. Which company in our industry has the highest retention rate?

4.5 Increasing lifetime value

By definition, customer lifetime value is the present value of the future profits. To increase customer lifetime value, one has to increase the profits generated from that customer. The most common ways to achieve that is either to up-sell or to cross-sell to the same customer, i.e., this will make your existing customers buy more products from you and buy it more often.

When a customer is satisfied, he or she will recommend he service or the product to colleagues and friends. This recommendation results will increase as well as the transfer sales. "The cost of acquiring new customers by referrals is substantially lower than traditional methods (Van den Poel, D., Larivi`ere, B., 2004)." In the long run, the company's lifetime value will increase considerably and so will the profits of a company.

CHAPTER 5
ON-LINE ANALYTICAL PROCESSING

5.1 What is On-Line Analytical Processing?

On-Line Analytical Processing (OLAP) is a technology that provides a multi dimensional function for assessing any production action, from all corners of the cube at different speeds. "Any observation, starting at the top, or anywhere within, may be drilled down to the next level of detail, and as far down as the original transactions (Michael L. Gonzales., 2005)."

OLAP is fast, supple and systematic and may obsolete the necessity for any traditional analysis programming.

5.2. OLAP and Data Mining

Data mining is about the successful exploitation of data for decision support purposes. We need to provide feedback to people and use the information from data mining to improve business processes. We need to enable people to provide input, in the form of observations, hypotheses, and hunches about what results are important and how to use those results (Agresti, A., 2002).

In the larger solution to exploiting data, OLAP clearly plays an important role as a means of broadening the audience with access to data (Michael J.A. Berry and Gordon S. Linoff, 2004). Decisions once made based on experience and educated guesses can start to be based on data and patterns in the data. Anomalies and outliers can be identified for further investigation, including applying data mining techniques. For instance, a user might discover that a particular item sells better at a particular time during the week by use of an OLAP tool. This might lead to an investigation using marker basket analysis to find other items purchased with that item. Market basket analysis might suggest an explanation for the observed behaviour – more information and more opportunities for exploiting the information.

Another problem when building cubes is determining how to make continuous dimensions discrete. We make a dimension discrete by assigning bins to ranges of values on that dimension. This begs the question of how to choose the ranges. In this chapter, we talked about using equal-sized bins, such as deciles. The information from the decision tree is useful here, too. The nodes of a decision tree determine the best breaking point for a continuous value. This information can be fed into the OLAP tool to improve the dimension (Michael L. Gonzales., 2005).

One of the problems with neural networks is the difficulty of understanding the results. This is especially true when using them or undirected data mining, such as a clustering algorithm using SOM (Self Organising Maps) networks (Michael J.A. Berry and Gordon S. Linoff, 2004). We might use SOM's to find clusters of people who are no longer using their credit cards. The inputs into the network might be account balances in the months before they left, the types of purchases made with the card, and some demographic and credit information. The SOM identifies clusters, but we do not know what they mean.

OLAP to the rescue! We have a set of data that is now enhanced with a predicted cluster and we want to visualize it better. This is a good application for a cube. By using OLAP on the same data – with information about the clusters included as a dimension – we can determine the features that distinguish clusters. The dimensions used for the OLAP tool are the inputs to the neural network along with the cluster identifier. There is a tricky data conversion problem because the neural networks require continuous values scaled between 0 and 1, and OLAP tools require discrete values. For values that were originally discrete, this is no problem. For continuous values, we can use the binning technique based on ranges.

As these examples show, OLAP and data mining complement each other. "Data mining can help build better cubes by defining appropriate dimensions, and further by determining how to break up continuous values on dimension." OLAP provides powerful visualization to better understand the results of data mining, such as clustering and neural networks. Used together, OLAP and data mining reinforce each other's strengths and provide more opportunities for exploiting data.

5.3 Strengths of OLAP

OLAP has several strengths for analysing data (Michael J.A. Berry and Gordon S. Linoff, 2004):

- It is a powerful visualisation tool.
- It provides fast, interactive response times.
- It is good for analysing time series.
- It can be useful to find some clusters and outliers.
- Many vendors offer OLAP products.

5.4 OLAP is a Powerful Visualisation Tool

OLAP best strong point is its ability to visualise data in a way what is easy to understand for all business users. OLAP can be applied across a wide variety of domains, to represent sales in retail stores, charges on a credit cart, drugs proscribed to patients, repairs to diesel engines, and so on. This format is related to the star schema mechanism popular among some database vendors for some decision-support applications.

If one makes a solid choice of scope then ones framework will be solid and flexible for presenting data. Dimensions can model hierarchies of data that are readily apparent in a particular domain. Geography is a good example, where sites, counties, states, and regions are all interesting for data analysis purposes, and OLAP tools can represent all along a single dimension. Other uses for complex dimensions are time and product codes.

5.5 OLAP Provides Fast Response

OLAP's ability to analyse data is useful because it can respond to user queries quickly, with response times on large databases measured in seconds. This is a result of pre-calculating summaries of the data that are suggested by the dimensions on the cube (Michael J.A. Berry and Gordon S. Linoff, 2004).

This response time is valuable, especially when using the OLAP tools to drill-down into the data to find patterns. Whether the cube is implemented as an MDD or in a relational database, you can usually find the specific data items used to generate a particular report. This yields much better productivity for analysis. And, because of the fast response time, gives users increased confidence in the data and in their ability to use it.

5.6 OLAP is good for time series

Time series are difficult for most tools to handle, but they fit quite well into OLAP. When using time as a dimension, support for hierarchies on the dimensions is important. Complex dimensions can support looking at dates by quarter, month, or day of the week all at the same time.

Some OLAP tools provide broader support for time series by supporting features specific to series. For instance, they can help calculate the ratio between adjacent values, to find, for instance, the rate of growth from one quarter to the next. Trying to create a similar query on a relational database is cumbersome and expensive because SQL does not readily support such inter-row calculations (Michael J.A. Berry and Gordon S. Linoff, 2004).

5.7 OLAP finds clusters and outliers

"As an acid for analysis, OLAP provides the ability to identify clusters in the data by being able to look at the data along several dimensions at the same time." For instance, al cluster of drugs prescribed by a particular physician group can be identified using the slicing and dicing features of OLAP tools. This cluster can then suggest ways to sell more drugs to this particular group. Or, using a technique like memory-based reasoning, similar physician groups might be identified for targeted sales.

Once upon a time, special reports had to be coded to identify trends and find outliers. With an OLAP tool, outliers along many dimensions can be identified quickly by business users familiar with the domain in which they are working.

5.8 OLAP is supported by many vendors

Another major strength of OLAP is that there are many vendors supporting OLAP tools – probably almost as many for this single application as for data mining in general. OLAP tools are maturing and popular features are appearing in all the best tools. In addition, training is available for end users and interfaces to data sources are getting simpler and simpler.

5.9 Weaknesses of OLAP

OLAP has some weaknesses as well:

- Setting up a cube can be difficult;
- it does not handle continuous variables well;
- cubes can quickly become out-of-date and;
- it is not data mining.

5.10 Setting up a cube is difficult

Not all domains are appropriate for setting up a cube. Sometimes the number of dimensions gets out of control. For instance, a simple retail model might have individual customer transactions as the base data. However, the market may be split between business and retail customers, have international franchises, and different approaches in upmarket and down-market areas. It only becomes more complicated if you want to include the method of payment (cash, credit card, store credit, and check) and decision trees can be helpful (Michael J.A. Berry and Gordon S. Linoff, 2004).

When setting up dimensions on a cube, the values for the different dimensions need to be comparable between records. For instance, a customer that carries an average monthly balance of R700 on a credit but has a credit limit of R5 000 is different from a customer with the same average monthly balance but with a credit limit o R750.

The cube needs to be informative to the end users. Selecting the dimensions and choosing the data for the cube are most of the work that needs to be done.

5.11 OLAP does not handle continuous values

All the proportions for OLAP are necessarily isolated, meaning that they take on a permanent list of values. This requires preconditioning continuous variables in a process called binning. So, a continuous variable like monthly balance on a credit may be split into several bins: "no monthly balance," "small monthly balance", "typical monthly balance", "high monthly balance", and "very high monthly balance".

5.12 Cubes become out-of-date quickly

The creation of a cube involves several types of information and data:

- Specific base data that goes into the cubes;
- definitions of dimensions and hierarchies on the dimensions and;
- methods to make continuous data discrete

In a rapidly changing market, all of these can change quickly. The introduction of new product lines, reorganisation geographic regions, and changes in pricing structures rapidly conspire to make current data difficult to compare to historical data. This is a common problem in data mining, but is exacerbated by all the choices that explicitly go into making a cube.

By taking advantage of relational database technology, OLAP tools make it fairly easy to incorporate new base data into a cube. In fact, updates to the cube can happen at the same time that users are analysing data. Adding a new hierarchy to a geography dimension, or a new definition of products, is simpler in a tool based on multidimensional databases.

5.13 OLAP does not automatically find patterns

As we have pointed out throughout this chapter, OLAP and data mining complement one another. It will never be substitute for data mining. It offers a much better perceptive of data, and because of the cube developed for OLAP, this will make data mining results understandable. However, the only glitch in OLAP is that it does not automatically find patterns in data.

5.14 When to apply OLAP

OLAP is an influential method to share information end users for advanced reporting needs. This will provide the capability for more users to base their decisions on data, instead of gut feelings, knowledgeable guesses, and personal experience. OLAP goes together with data mining techniques such as Self Organising Maps (SOM) networks. When using SOM networks for finding clusters, OLAP can provide the insight needed to find the business value in the identified clusters.

CHAPTER 6
LITERATURE STUDY: LIFETIME VALUE

6.1 Introduction

"Customers are not created: some are more valuable than others. Business that have the ability to determine the value of the customers have the edge in today's highly competitive business environment." (Crowder, M., Hand, D.J., Krzanowski, W 2007)

Almost any business doesn't matter how big or small, is in possession of some kind of data. What most of the businesses don't realise is that there is a wealth of knowledge hidden in this data. The immense question however is how to extract this valuable information, and what can one use it for.

Different customer groups can be identified, for example (Gupta, S., Lehmann, D. R., Stuart, J. A., 2004):

- Profitable customers;
- Non profitable customers;
- Customers who are likely to churn; and
- Loyal customers, etc.

It is the responsibility of the business to correctly identify the specific products that each of the customers acquire.

If a bank stays up to date with its custorner and customer behaviour, the bank would know when a customer behaves out of his/her norm. This is important, because it will act as an "early warning system" that the customer is not satisfied or not interested in the business anymore. In the banking environment, abnormal behaviour can even suggest something like fraud or churn (Haenlein, M., Kaplan, A.M., Beeser, A.J., 2007).

An effective method to achieve this is called LTV.

6.2 What is Lifetime Value (LTV)?

Lifetime value is a method of determining how much money the banks customers are worth over a specific time if the customer buys certain products and services. If customer retention increases then sales and profits will also increase significantly. It is important to retain customers, but not at the expense of other customers (Pfeifer, P. E., Haskins, M. R., Conroy, R. M., 2005).

6.3 Why LTV?

According to Jim Nova (2000) the complicating part of measuring LTV is deciding what a "lifetime" is. The lifetime is the total time a customer will be at a bank before leaving the bank. Details on calculating lifetime value will be discussed later on, but first, a clarification.

Jim Nova also says the following: "The lifetime value concept has been horribly abused and misunderstood from the beginning of customer relationship management. *"It is not necessary to figure out an absolute lifetime value for a customer or wait "a lifetime" to find out the value."* If a bank is new to this LTV and has not tracked the appropriate parameters, or a company is new and lacks meaningful operating history, it can look for "relative lifetime value", link it to customer behaviour, and still get tremendous leverage from using LTV in the business model." (Jim Nova, 2000).

6.4 Where to start

6.4.1 Apply the Customer Lifetime Value Concept

LTV is the technique of calculating how much the value of each customer is worth to the bank, over the length of time that they remain customers. The lifetime for customers will vary from industry to industry, and from brand to brand.

The lifetime of customers should come to an end when their contribution ceases to be profitable unless steps are taken to revitalise them.

6.4.2 Benefits from Customer Lifetime Value

Organisations experience indicates that a number of benefits apply and Pfeifer gives the following examples (Pfeifer, P. E., Haskins, M. R., Conroy, R. M., 2005):"

- A 5% increase in customer retention can create a 125% increase in profits;

- A 10% increase in retailer retention can translate to a 20% increase in sales; and
- Extending customer lifecycles by three years can triple profits per customer."

6.4.3 Identify Categories of Customer

Before calculating customer lifetime value, it is possible to analyse the Bank customers according to four key attributes. This can help to explain the analysis and can play a major part for the basis of marketing actions to advance LTV (Michael J.A. Berry and Gordon S. Linoff, 2004):

- Frequency: How often they purchase? ;
- Regency: How much time has elapsed since the last purchase? ;
- Amount: How much they spend? ; and
- Category: What sort of product they buy.

6.4.4 Calculate Lifetime Value

In a consumer business, customer lifetime value is calculated, in practice, by analysing the behaviour of a group of customers who (Gupta, S., Lehmann, D. R., Stuart, J. A., 2004):

- Have the same recruitment date;
- Recruited from the same source; and
- Bought the same types of product.

In a business-to-business environment, a similar approach can be used.

- Isolate particular customers, and examine them individually; and
- Analyse the behaviour of different groups, segmenting the customer database by factors such as industry, annual turnover, or staff numbers.

The basic calculation has three stages (Pfeifer, P. E., Haskins, M. R., Conroy, R. M., 2005):

- Identify a discrete group of customers for tracking;
- Record (or estimate) each revenue and cost for this group of customers, by campaign or season; and
- Calculate the contribution, by campaign or season.

6.4.5 Refine the Calculation

Other factors can be introduced to make the calculation more relevant. In a business-to-business environment, for example, it may be the sales representatives who generate sales.

In this case, the calculation should include the representative's "running costs" and the cost of any centrally produced sales support material.

6.4.6 Analyse the Results

A company may offer different products or brands, which are marketed under different cost centres. If a customer is a customer of more than one cost/profit centre, there is a choice of approaches (Pfeifer, P. E., Haskins, M. R., Conroy, R. M., 2005):

- Examine customers of each brand and ignore multi purchases; and
- Build a more detailed model that combines and allocates the cumulative costs as well as the cumulative profit in the appropriate proportions.

When the results are analysed, the building of financial models could commence.

6.4.7 Use Customer Lifetime Values to Improve Marketing Performance

There are four important applications (Pfeifer, P. E., Haskins, M. R., Conroy, R. M., 2005):

- Setting target customer acquisition costs;
- allocating acquisition funds;
- selecting acquisition offers; and
- supporting customer retention activities.

Equally valid may be an increase in expenditure aimed at reactivating customers, this is a classic retention activity.

6.4.8 Set Target Customer Acquisition Costs

If a customer is expected to generate more than one sale, the allowable cost can be greater than the cost allowed for the first sale, the classic loss-leader approach to customer acquisition. However, overspending on customer acquisition can also be harmful. A reasonable calculation is to recruit only from those sources that yield new customers at less than half the estimated lifetime value (Haenlein, M., Kaplan, A.M., Beeser, A.J., 2007). On that basis, the worst sources will have a cost per customer close to a LTV, while the average cost per customer should be far lower.

6.4.9 Allocate Acquisition Funds

Different recruitment sources will provide customers with different lifetime values. After identifying those values, spend more on the best sources.

6.4.10 Select Acquisition Offers

The LTV of a customer may depend on the type and value of their initial purchase. In turn, this can lead to decisions about which products and offers to use when advertising externally, or when considering how to upgrade existing customers.

6.4.11 Support Customer Retention Activities

Once the typical LTV of a group of customers is known, companies can decide how hard to work at retaining them. It is not a foregone conclusion that all customers are worth having. Activities should be tailored to the customers who are most valuable.

6.4.12 Increase Value with New Offers

A financial services company can increase customer lifetime value by cross selling a variety of products and services.

6.5. Common Mistakes

6.5.1 Trying to Retain the Wrong Customers

Customer retention costs money in terms of sales and marketing funds, so do bear in mind that not all customers are worth keeping. The Bank should carefully select the customers who are likely to yield the highest returns over a specific time and prioritise the allocation of marketing resources to these (Crowder, M., Hand, D.J., Krzanowski, W 2007).

6.5.2 Offering Customers a Limited Range of Products

When the Bank has identified the most valuable customers, it needs to have a wide variety of products or services to offer them. Cross selling and up selling are the best ways to increase customer lifetime value, but this can be difficult with a limited product range (Crowder, M., Hand, D.J., Krzanowski, W 2007).

Customers are its company's most valuable asset, think about "share of customer wallet" rather than just share of market (Michael J.A. Berry and Gordon S. Linoff, 2004).

6.5.3 Spending Too Much on Acquiring New Customers

Customer LTV analysis reinforces a traditional marketing rule of thumb, that it costs less to retain existing customers than to acquire new ones. Over emphasis on new business development could be a bad move, since existing customers are easier to sell to (Garland R., 2003).

6.6 Conclusion

To address the question "What is a consumer?" let's look at different aspects of consumer behaviour. Even within the same household, the answer is not always the same. To give a simple example, consider a young couple with two children. The choice of which movie to see or which cable channels to subscribe to might involve the whole family. On the other hand, only the parents might make the choice of which car to buy. Only one person typically makes the choice of which credit card to use for any given purchase.

Let's dive in a bit and look at these roles in some familiar industries. For example, in a bank (Pfeifer, P. E., Haskins, M. R., Conroy, R. M., 2005):

- Actions, such as depositing money or making a credit card payment, occur within product lines. So the action dimension occurs at the account level. Often, data mining is focused on the "cardholder" or on the "account". A single individual may have several relationships, but each one is treated as a different customer.
- Ownership is a legal definition for many types of accounts, and most often, corresponds to a single individual or couple. So the ownership dimension allows all the commonly owned accounts to be grouped together.
- Decision making typically occurs once within a household for any given decision. The household dimension gathers together all the accounts and transactions owned by individuals in the household.

This is a useful breakdown, even though it oversimplifies the real world. The sidebar tells the true story of a bank that rejected a credit application from the daughter of an important customer. Who is the customer here?

Understanding the dimension of customers behaviour is important because it can also influence the business problem consider churn in the wireless industry. Does a customer churn when one telephone in the household is cancelled? Does a customer churn when all telephones are cancelled? Does a customer churn when a highly profitable monthly plan is downgraded to a basic plan?

There is no right answer to these questions. The answer depends on the definition of customer. Is the customer a telephone number, a household, or the owner of a particular service plan? Understanding customers is an ongoing process. It is always worthwhile, though, remembering the full complexity of the customer relationship, even when using accounts (or some other unit) as proxies for full customer relationships

CHAPTER 7
CASE STUDY

Chapter 7, Case Study is based on six months of customer account data and profit tables. For all profit table purposes, if there is a negative before an amount, it means the institution is making a profit and if the amount is positive the institution is making a loss. See Table 11 in Appendix A for meaning of abbreviations.

7.1 Status of Accounts

Table 2 below, sample results set of an OLAP model; show the status of the accounts. OLAP clustered the for the researcher what will help for the data mining These are active accounts and/or are they closed (churned) accounts. For each sub product in a specific product life the total active accounts and total closed account are given and the total of the active and closed accounts added together are also given as total. Product life is synonym for account life.

Table 2: Status of Accounts

	Cheque Accounts					
Sum Of COUNT ACCT		**Sub product code**				
PROD LIFE in years	**STATUS**	**Silver**	**Turquoise**	**Gold**	**Platinum**	**Smart Account**
0 – 1	**ACTIVE**	17,994	27,964	8,435	3,221	359,490
	CLOSED	6,690	6,455	2,701	722	67,354
0 - 1 Total *		**24,684**	**34,419**	**11,136**	**3,943**	**426,844**
1 – 2	**ACTIVE**	23,273	29,431	9,695	3,989	327,754
	CLOSED	6,556	8,262	2,134	409	190,596
1 - 2 Total *		**29,829**	**37,693**	**11,829**	**4,398**	**518,350**
2 – 3	**ACTIVE**	16,227	9,857	6,889	2,482	186,908
	CLOSED	4,629	1,711	1,583	187	76,421
2 - 3 Total *		**20,856**	**11,568**	**8,472**	**2,669**	**263,329**

3 – 4	**ACTIVE**	14,344	5,320	3,134	2,309	150,320
	CLOSED	2,383	527	181	146	46,869
3 - 4 Total *		**16,727**	**5,847**	**3,315**	**2,455**	**197,189**
4 – 5	**ACTIVE**	14,777	4,493	2,762	2,162	131,031
	CLOSED	1,904	364	140	95	30,993
4 – 5 Total *		**16,681**	**4,857**	**2,902**	**2,257**	**162,024**
5 – 6	**ACTIVE**	12,358	3,246	2,018	1,626	93,308
	CLOSED	1,517	227	97	93	19,736
5 - 6 Total *		**13,875**	**3,473**	**2,115**	**1,719**	**113,044**
6 – 7	**ACTIVE**	11,671	2,476	2,080	1,401	67,977
	CLOSED	1,234	159	88	76	13,593
6 - 7 Total *		**12,905**	**2,635**	**2,168**	**1,477**	**81,570**
7 – 8	**ACTIVE**	10,139	1,859	2,008	1,567	49,238
	CLOSED	1,127	102	92	45	10,287
7 - 8 Total *		**11,266**	**1,961**	**2,100**	**1,612**	**59,525**
8 – 9	**ACTIVE**	12,623	1,820	2,563	1,837	49,800
	CLOSED	1,477	103	126	79	10,007
8 - 9 Total *		**14,100**	**1,923**	**2,689**	**1,916**	**59,807**
9 – 10	**ACTIVE**	18,598	2,267	4,193	3,221	56,451
	CLOSED	1,731	124	158	119	10,679
9 – 10 Total *		**20,329**	**2,391**	**4,351**	**3,340**	**67,130**
10 – 11	**ACTIVE**	22,694	2,676	4,851	3,705	66,751
	CLOSED	1,815	121	158	90	11,570
10 - 11 Total *		**24,509**	**2,797**	**5,009**	**3,795**	**78,321**

7.2 Sum of Accounts

Table 3: Sum of Accounts

	Cheque Accounts				
STATUS	ACTIVE				
Sum Of COUNT_ACCT	Sub product code				
PROD_LIFE	Silver	Turquoise	Gold	Platinum	Smart Account
0 – 1	17,994	27,964	8,435	3,221	359,490
1 – 2	23,273	29,431	9,695	3,989	327,754
2 – 3	16,227	9,857	6,889	2,482	186,908
3 – 4	14,344	5,320	3,134	2,309	150,320
4 – 5	14,777	4,493	2,762	2,162	131,031
5 – 6	12,358	3,246	2,018	1,626	93,308
6 – 7	11,671	2,476	2,080	1,401	67,977
7 – 8	10,139	1,859	2,008	1,567	49,238
8 – 9	12,623	1,820	2,563	1,837	49,800
9 – 10	18,598	2,267	4,193	3,221	56,451
10 – 11	22,694	2,676	4,851	3,705	66,751

Table 3 shows how many active accounts are there in every sub product code.

Active status will be used for all tables as a result of the fact that lifetime values for active accounts need to be calculated to see what impact current accounts have on the profit of Bank XYZ and what actions need to be implemented to retain these current accounts.

7.3 Profit of All Active Accounts

Table 4: Profit of all active Customers

	Cheque Accounts				
STATUS	**ACTIVE**				
Sum Of profit of all active accounts	**Sub product code**				
PROD_LIFE	**Silver**	**Turquoise**	**Gold**	**Platinum**	**Smart Accou**
0 – 1	-7,503	93,580	-7,844	-67,846	648,5
1 – 2	-781,673	-108,863	-678,173	-512,257	-59,5
2 – 3	-1,070,013	-246,523	-758,188	-524,449	-837,3
3 – 4	-1,130,563	-145,806	-467,950	-506,654	-667,1
4 – 5	-1,124,423	-123,407	-422,457	-517,940	-666,2
5 – 6	-889,692	-91,803	-323,773	-455,571	-461,8
6 – 7	-829,978	-69,160	-352,527	-350,420	-374,8
7 – 8	-736,008	-48,338	-308,941	-378,546	-251,5
8 – 9	-952,652	-55,232	-415,478	-467,920	-297,9
9 – 10	-1,414,398	-70,725	-653,321	-853,601	-406,2
10 – 11	-1,585,234	-67,464	-763,569	-904,650	-413,6

Table 4 is all about the sum of profit of all active accounts in each sub product. Again this is an OLAP query is implemented to calculate the profit contributed by accounts by a specific sub product in a specific product life.

As expected, all accounts in their first years are not as profitable as accounts that have been with the bank for at least a year. In addition to that the Turquoise and Smart Account run at a loss in their first year as compared to other cheque products.

7.4 Average Profit per Accounts

Table 5: Average Profit per Account

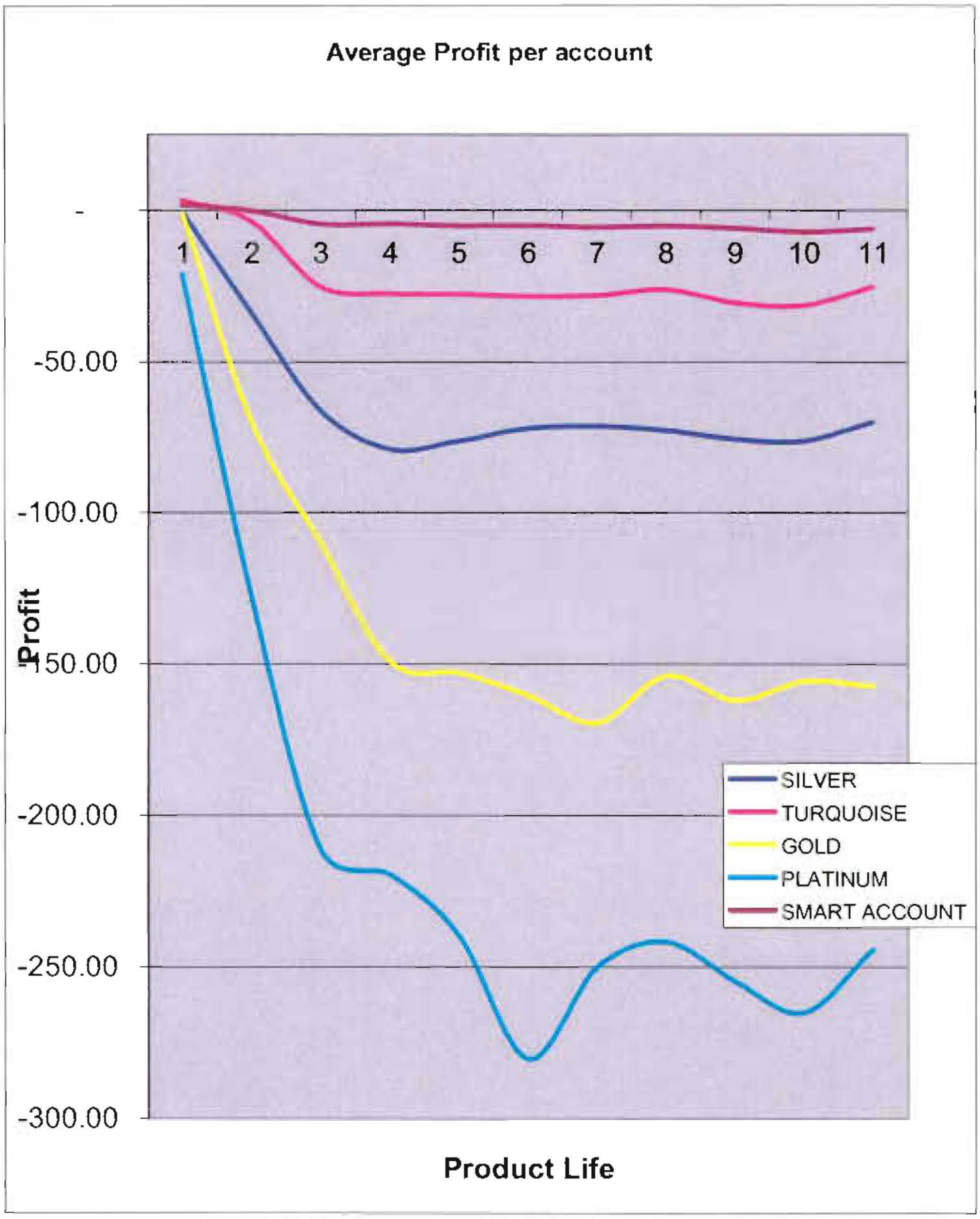
Average Profit per account					
	Sub product code				
Product life	**Silver**	**Turquoise**	**Gold**	**Platinum**	**Smart Accoun**
0 – 1	-0.42	3.35	-0.93	-21.06	1.
1 – 2	-33.59	-3.70	-69.95	-128.42	-0.
2 – 3	-65.94	-25.01	-110.06	-211.30	-4.
3 – 4	-78.82	-27.41	-149.31	-219.43	-4.
4 – 5	-76.09	-27.47	-152.95	-239.57	-5.
5 – 6	-71.99	-28.28	-160.44	-280.18	-4.
6 – 7	-71.11	-27.93	-169.48	-250.12	-5.
7 – 8	-72.59	-26.00	-153.85	-241.57	-5.
8 – 9	-75.47	-30.35	-162.11	-254.72	-5.
9 – 10	-76.05	-31.20	-155.81	-265.01	-7.
10 – 11	-69.85	-25.21	-157.40	-244.17	-6.

$$Avgerage\ \ profit\ \ per\ \ year = \frac{Sum\ \ of\ average\ \ profit\ \ per\ accounts}{Sum\ \ of\ \ accounts}$$

Let's take product life 10 - 11, in table 5 above, Silver as an example:

Accounts that have been with Bank XYZ for 10 years will contribute R69.85 of profit to Bank XYZ annually. To see the figure 7 of table 5 refer to the next page.

Figure 7: Average Profit per Account



The Bank must look with great concern at the sub products that show no increase in profit over time as in the figure 7 above. Most of the sub products accounts show an increase in profit over the first three (3) years in product life. After the three (3) years in product life the sub products show no increase but show more a tendency to stabilise over longer periods that means no increase in profit. This is a serious issue that urgently needs to be looked at.

7.5 Retention

Table 6: Retention

Retention Rate					
	Sub product code				
Product Life	Silver	Turquoise	Gold	Platinum	Smart Account
0 – 1	73%	81%	76%	82%	84%
1 – 2	78%	78%	82%	91%	63%
2 – 3	78%	85%	81%	93%	71%
3 – 4	86%	91%	95%	94%	76%
4 – 5	89%	93%	95%	96%	81%
5 – 6	89%	93%	95%	95%	83%
6 – 7	90%	94%	96%	95%	83%
7 – 8	90%	95%	96%	97%	83%
8 – 9	90%	95%	95%	96%	83%
9 – 10	91%	95%	96%	96%	84%
10 – 11	93%	96%	97%	98%	85%

$$\text{Retention} = \frac{\textit{Total of active accounts in a spesific product life}}{\textit{Total accounts of that spesific product life}}$$

Retention as one of the major elements of life time value is measured on its own to give an indication of the number of accounts that are retained in each year the product is opened. Retention in this thesis refers to how many accounts from the previous product life year are still active in the current product life year.

Of the 100% opened in the first year, only an average of 73% will be taken to the second year and of that 73% only 78% will be taken to the third year and so forth.

To see the graph of table 6 refer to figure 8 on the following page.

Figure 8: Retention

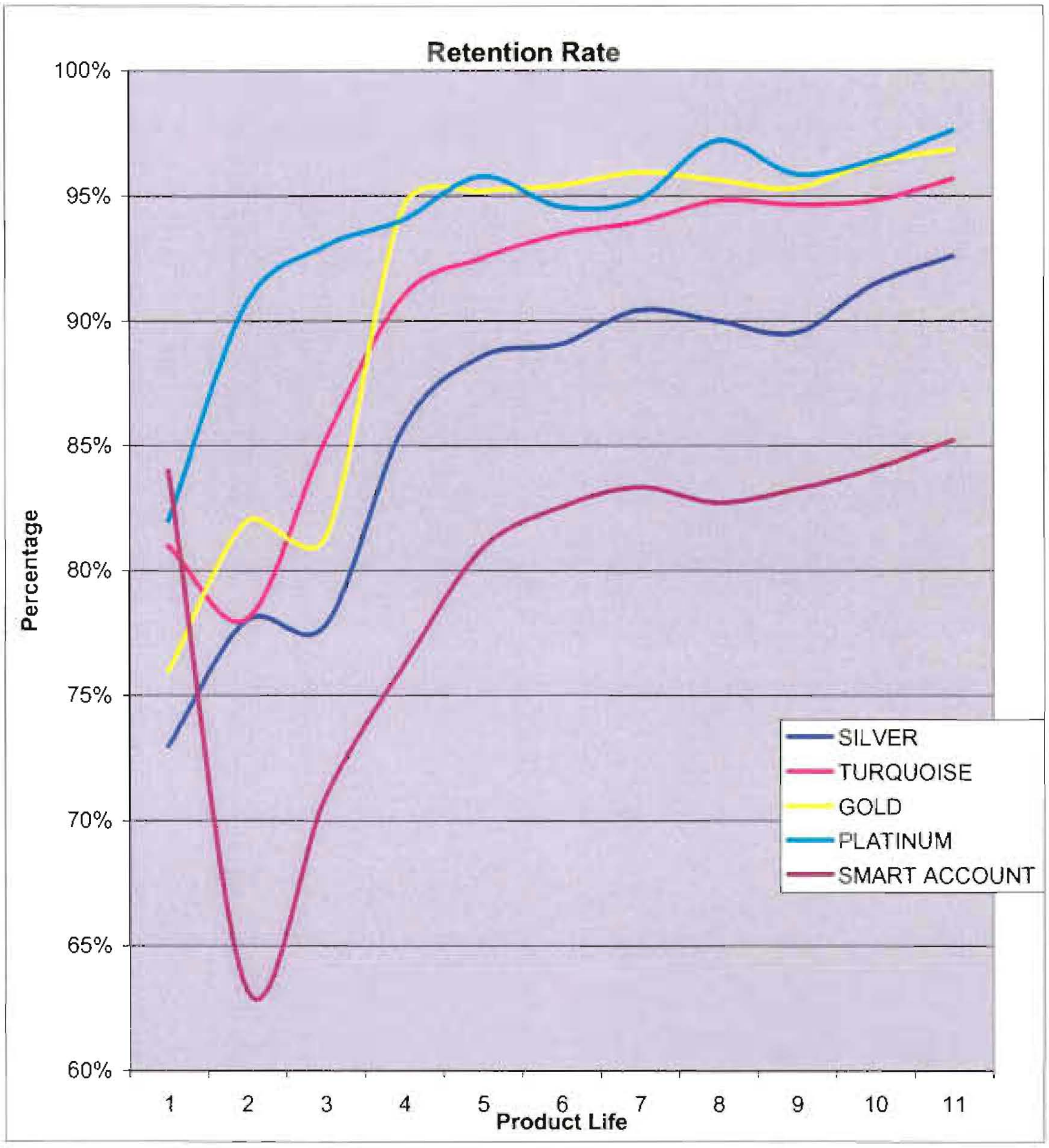


Figure 8 shows the retention rate of active accounts over time. Apart from Platinum all other monitored product indicators a declining trend in term of retaining accounts from the previous year. Smart Account shows a huge drop in retention at year two (2). As a result of the overall first three (3) years, accounts that have been with Bank XYZ for a short period of time is a concern, meaning that Bank XYZ retention of the previous product life year accounts are very low.

7.6 Churn

Table 7: Churn

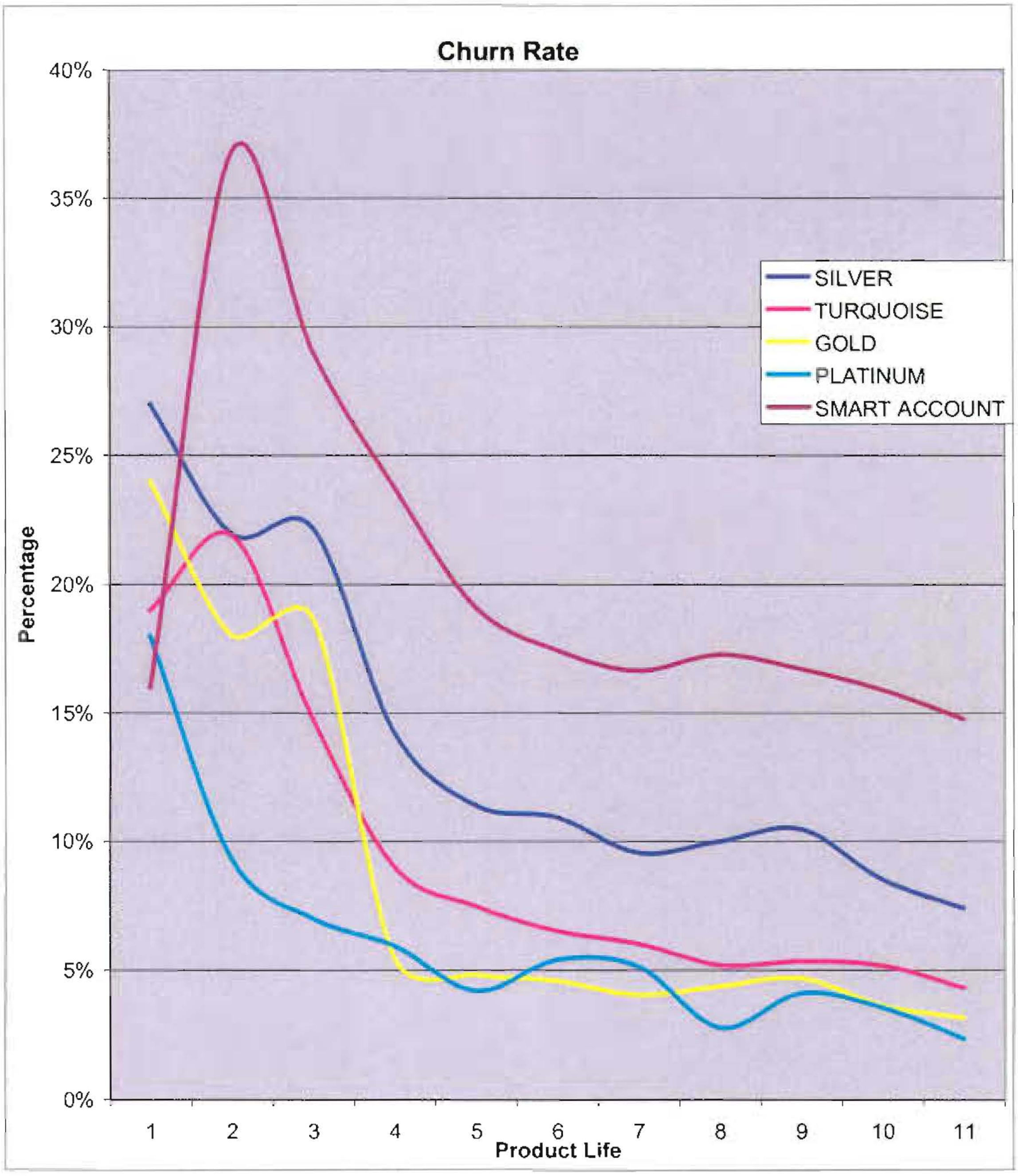
Churn Rate					
	Sub product code				
Product Life	Silver	Turquoise	Gold	Platinum	Smart Account
0 – 1	27%	19%	24%	18%	16%
1 – 2	22%	22%	18%	9%	37%
2 – 3	22%	15%	19%	7%	29%
3 – 4	14%	9%	5%	6%	24%
4 – 5	11%	7%	5%	4%	19%
5 – 6	11%	7%	5%	5%	17%
6 – 7	10%	6%	4%	5%	17%
7 – 8	10%	5%	4%	3%	17%
8 – 9	10%	5%	5%	4%	17%
9 – 10	9%	5%	4%	4%	16%
10 – 11	7%	4%	3%	2%	15%

$$Churn = 1 - \ retention \ \ rate$$

Churn can literally be defined as, how many accounts haven't survived from the previous product life year to the current product life year.

To see the graph of table 7 refer to figure 9 on the following page.

Figure 9: Churn



The churn rate depicted above in figure 9 shows a huge increase in churn for sub product Smart Account in the 2nd year of the product life as a result of this. For the other sub products there is also a huge increase in churn for the first three (3) years in product life after the third year there is a decline in churn. In summary, churn, customers who have accounts with the Bank doesn't have a lot of faith in the Bank for the first three (3) years of their product life after that the churn rate declines over time.

7.7 Survival Analysis

Table 8: Survival Analysis

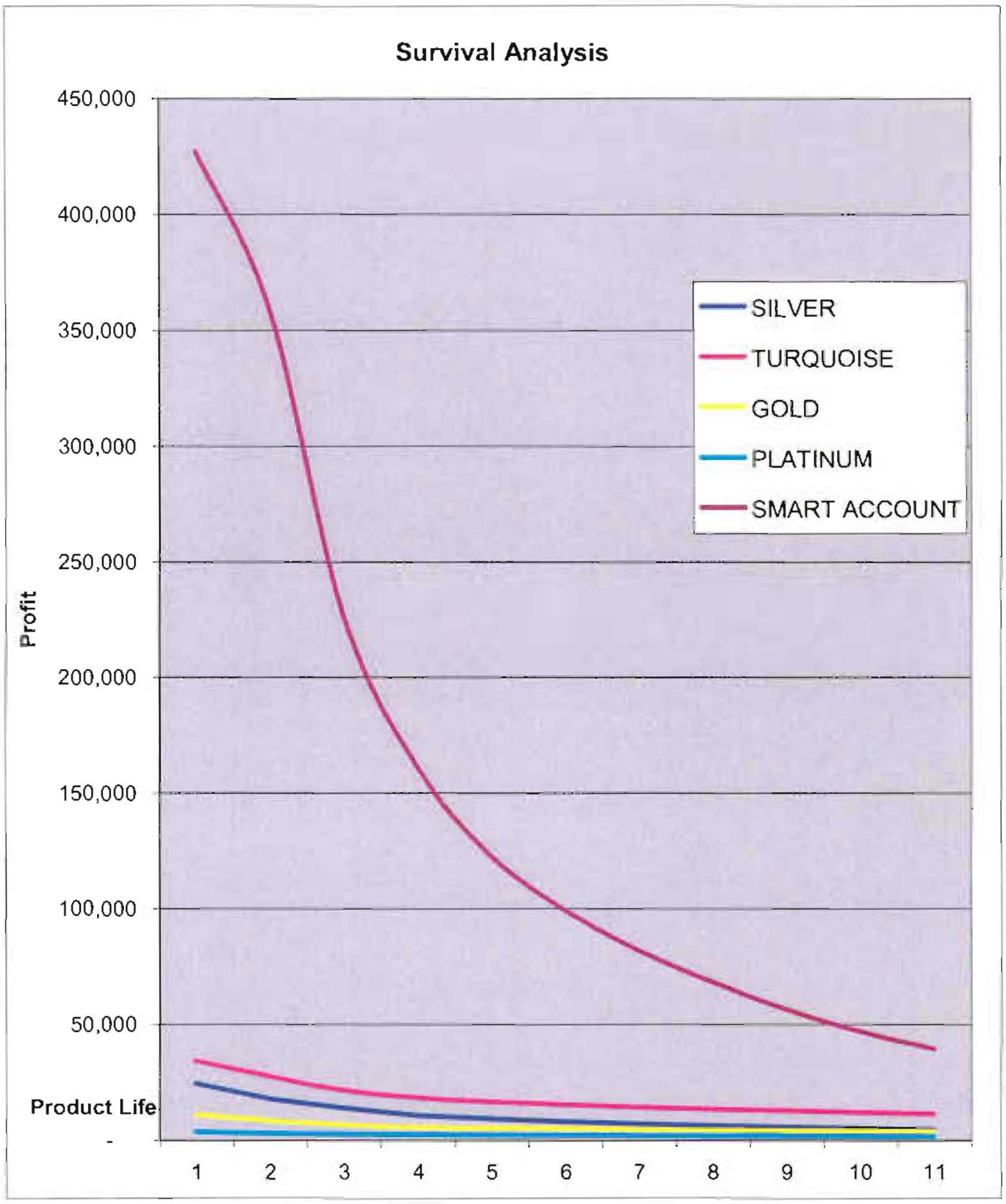
Survival Analysis of accounts that are less than a year of relationship with Bank XYZ					
	Sub product code				
Product Life	Silver	Turquoise	Gold	Platinum	Smart Account
0 – 1	24,684	34,419	11,136	3,943	426,844
1 – 2	18,019	27,879	8,463	3,233	358,549
2 – 3	14,059	21,768	6,937	2,933	226,711
3 – 4	10,939	18,549	5,640	2,727	160,917
4 – 5	9,380	16,877	5,332	2,565	122,670
5 – 6	8,310	15,612	5,075	2,457	99,204
6 – 7	7,401	14,592	4,842	2,324	81,885
7 – 8	6,693	13,711	4,646	2,204	68,239
8 – 9	6,024	12,998	4,442	2,143	56,446
9 – 10	5,393	12,302	4,234	2,055	47,002
10 – 11	4,934	11,664	4,080	1,981	39,525

$$Survival\ analysis = Total\ accounts\ in\ that\ spesific\ product\ life \times \text{Retention}\ rate\ of\ that\ specific\ product\ lif$$

Survival analysis is exactly what it tells one: How many accounts have survived from the previous product live to the current product life. From Silver, product life 0 – 1, only 4934 accounts has survived from an initial 24684 accounts.

To see the graph of table 8 refer to figure 10 on the following page.

Figure 10: Survival Analysis



Clearly Smart Account is the sub product that has the most product downward trend in figure 10 as a result of this. The focus will fall on Smart Account. Survival analysis is exactly what it says, how many accounts have survived from the previous year to the current year and that is what is plotted above in figure 10.

7.8 Lifetime Value

Table 9: Lifetime Value

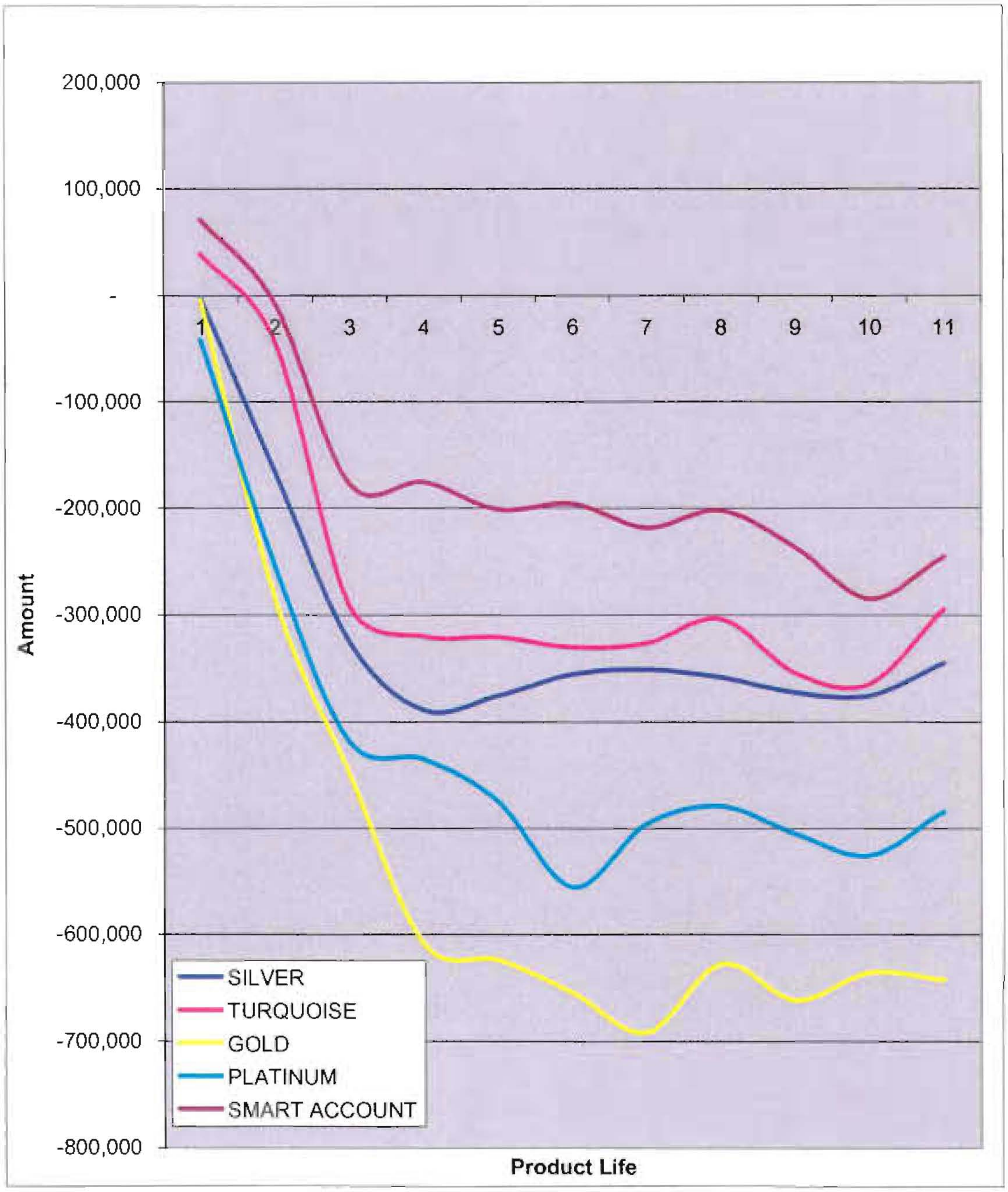
Lifetime Value for accounts					
	Sub product code				
Product Life	Silver	Turquoise	Gold	Platinum	Smart Account
0 – 1	-2,057	39,032	-3,795	-41,735	71,304
1 – 2	-165,704	-43,144	-285,430	-254,441	-7,183
2 – 3	-325,320	-291,711	-449,085	-418,664	-177,077
3 – 4	-388,852	-319,672	-609,267	-434,762	-175,423
4 – 5	-375,408	-320,365	-624,117	-474,665	-200,981
5 – 6	-355,182	-329,875	-654,678	-555,135	-195,623
6 – 7	-350,848	-325,797	-691,570	-495,580	-217,930
7 – 8	-358,135	-303,281	-627,796	-478,645	-201,938
8 – 9	-372,333	-353,967	-661,464	-504,691	-236,500
9 – 10	-375,202	-363,883	-635,784	-525,083	-284,439
10 – 11	-344,622	-294,052	-642,280	-483,789	-244,940

$$Lifetime\ \ Value = \Pr oduct\ \ life\ \ of\ \ survival\ \ analysis \times Average\ \ profit\ \ per\ \ product\ \ life$$

As one can see from the survival analysis, Silver will be used as an example again. Considering the retention rate and churn rate, for product life 0 – 1, 24684 accounts contribute to a profit of R2057. For product life 1 – 2, 18019 accounts (78% of 24684 accounts) contribute to R165 704 of profit. Turquoise and Smart Account made a loss in their first year. This needs urgent attention.

To see the graph of table 9 refer to figure 11.

Figure 11: Lifetime Value



In order to explain the trench endures from figure 11: Silver will be used as an example. Of the 24684 accounts in the first product life, only 4934 accounts will make it to the 10-product life year. Now the question is how much profit will these 4934 accounts contribute to profit over a 10-year product life cycle? R344 662 profits will these 4934 accounts generate over 10 years.

7.9 Present Lifetime Value

Table 10: Present Lifetime Value

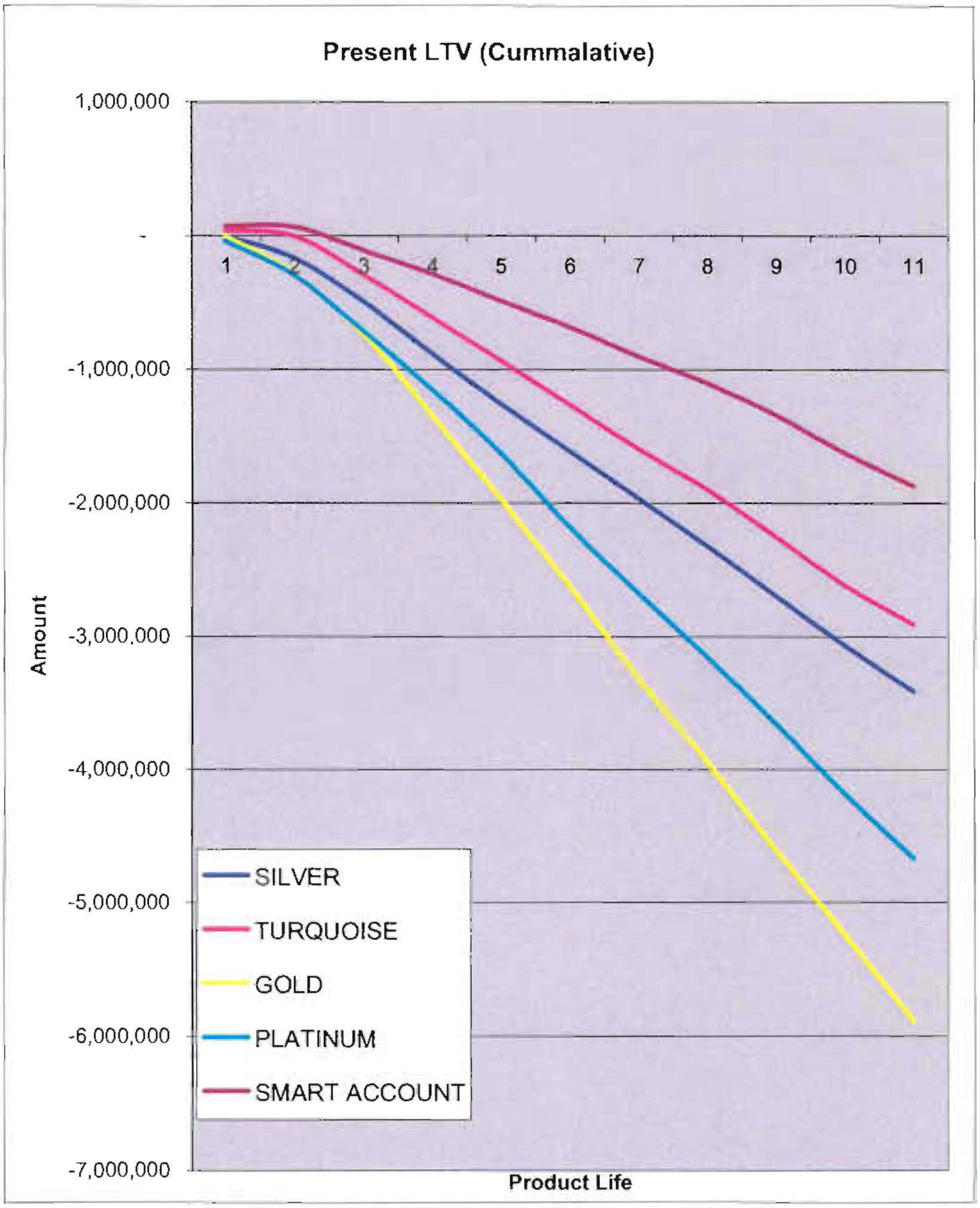
Present LTV					
	Sub product code				
Product Life	Silver	Turquoise	Gold	Platinum	Smart Account
0 – 1	-2,057	39,032	-3,795	-41,735	71,304
1 – 2	-167,761	-4,111	-289,225	-296,176	64,121
2 – 3	-493,081	-295,822	-738,310	-714,839	-112,955
3 – 4	-881,934	-615,494	-1,347,577	-1,149,601	-288,378
4 – 5	-1,257,342	-935,860	-1,971,693	-1,624,266	-489,358
5 – 6	-1,612,524	-1,265,734	-2,626,371	-2,179,402	-684,981
6 – 7	-1,963,372	-1,591,531	-3,317,941	-2,674,982	-902,911
7 – 8	-2,321,507	-1,894,812	-3,945,737	-3,153,627	-1,104,850
8 – 9	-2,693,840	-2,248,780	-4,607,201	-3,658,318	-1,341,350
9 – 10	-3,069,043	-2,612,662	-5,242,985	-4,183,401	-1,625,789
10 – 11	-3,413,664	-2,906,714	-5,885,264	-4,667,190	-1,870,728

$$\text{Present lifetime value} = \text{Cumulative lifetime value}$$

The present lifetime value is the lifetime value of each product life added together for example: Silver, product life 10 – 11 is all the amounts of each product life's lifetime value added together and total profit generated was R3 413 664.

To see the graph of table 10 refer to figure 12 on the following page.

Figure 12: Present Lifetime Value



Of the 24684 accounts in Silver only 4934 could be active accounts until the 10th year. The question the Bank needs to ask when it looks at figure 12 is: How much profit has been accumulated over a 10-year product life cycle? In this case (Silver) it is R3 413 664.

CHAPTER 8
CONCLUSION

In this thesis, a framework was provided for evaluating churner classification techniques based on a financial measure of accuracy, i.e. the profit loss incurred by a misclassification, considered from a customer lifetime value perspective.

First, using a customer-centric approach, we define a churner as someone whose LTV is decreasing in time. CRM was identified, LTV was defined and retention was given a purpose why it is so important in the CRM field. Second, we emphasize the fact that not all customers are equal, neither are all misclassifications. Data mining was looked at where it came from, where in the model it can be applied the best and how one can interpret the answers. Business Intelligence using data mining was researched and fundamental nature of the topic was laid out. Data mining for CRM was also another topic that got a lot of attention and where this fits in the model and why this fits in the model. Thirdly, the need for customer churn was identified and described how fundamentally important this was. Also, the LTV concept was derived in the customer churn process. The question that was asked was "Why churn modelling is useful?" was answered and because of this the increase of customer retention and LTV was automatically created to understand this topic much better. Lastly, in Chapter 5 it was shown how OLAP can be used to understand data mining concept and how it was used in this research project. The benefits and weaknesses of OLAP were also closely inspected to show the researcher what he should thoroughly understand upon completing a model based on this subject.

The cost-sensitive approach was showed how to achieve very good results in terms of the defined profit measure, emphasising the point that, even though it is important to achieve a good average performance, it is at least as important to correctly classify potentially profitable churners.

In an ideal world, where every management control parameter would be known, it would be possible to estimate the exact monetary value of the retention campaign: that would be the sum of changes in the customer lifetime values generated by the model implementation.

In reality, thanks to the LTV function, one could have at least an insight on what is the approximate value of a customer retention campaign.

Different topics for further research can be identified. As it has been shown, the product usage growth rate has a large impact on the LTV. In this paper, it was assumed that the growth rate to be constant. It would be interesting to allow varying growth rate and investigate the impact on our findings. Further developments could focus on a more accurate prediction of this value or a more accurate prediction of the LTV.

REFERENCES

Agresti, A., 2002. Categorical data analysis. Wiley, Hoboken, New Jersey.

Allison, P. D. 2003, Survival Analysis Using the SAS@ System, Cary, NC: SAS Institute.

Au W., Chan C.C., Yao X. 2003: A Novel evolutionary data mining algorithm with applications to churn prediction. IEEE Transactions on evolutionary computation, Vol. 7, No. 6, Dec 2003.

Bets, A., Datta, P., Drew, J., Mani, P.R., 2002. Statistics and Data Mining Techniques for Lifetime Value Modelling., Waltom.

Buckinx W., Van den Poel D., 2005. Customer base analysis: partial detection of behaviorally loyal clients in a non-contractual FMCG retail setting. European Journal of Operational Research 164

Buckinx W., Verstraeten G., Van den Poel D.,2005. Predicting customer loyalty using the internal transactional database. Expert Systems with Applications xxx

Crowder, M., Hand, D.J., Krzanowski, W 2007. On optimal intervention for customer lifetime value. European Journal of Operational Research Vol. 183 pp. 1550–1559

Gronroős, C., 2000: Service Management and Marketing: A Customer Relationship Management Approach

Ferreira J., Vellasco M., Pachecco M., Barbosa C., 2004. Data mining techniques on the evaluation of wireless churn. ESANN2004 proceedings – European Symposium on Artificial Neural Networks Bruges. ISBN 2930307048, p 483488.

Garland R., 2003. Investigating indicators of customer profitability in personal retail banking. Proceedings of the Third Annual Hawaii International Conference on Business, June.

Gupta, S., Lehmann, D. R., Stuart, J. A., 2004. Valuing customer. Journal of Marketing Research 41 (1), 7–18.

Haenlein, M., Kaplan, A.M., Beeser, A.J., 2007. A model to determine customer lifetime value in a retail banking context. European Management Journal Vol. 25, No. 3, pp. 221–234.

Hwang H., Jung T., Suh E., 2004. An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry. Expert Systems with Applications 26.

Japkowics N., Stephen S., 2002: The class imbalance problem: A systematic study. Intelligent Data Analysis Vol 6 p.429449.

Kim, S.H., Ko, E., Kim, M., Woo, J.Y., 2007: Organizational characteristics and the CRM adoption process. Journal of Business Research Vol. 61, pp. 65–74

Malthouse, E. C., Blattberg, R. C., 2005. Can we predict customer lifetime value? Journal of Interactive Marketing 19 (1), 2–16.

Michael J.A. Berry and Gordon S. Linoff, 2004. Mastering Data Mining: The art and science of CRM, 2004 p.311 – p.356, New York: John Wiley&Sons.

Michael L. Gonzales., 2005 – Getting Set for OLAP, www.Ibm.com. [Web]

Mozer M. C., Wolniewicz R., Grimes D.B., Johnson E., Kaushansky H., 2000. Predicting Subscriber Dissatisfaction and Improving Retention in the Wireless Telecommunication Industry. IEEE Transactions on Neural Networks, Special issue on Data Mining and Knowledge Representation.

Neslin, S. A., Gupta, S., Kamakura, W., Junxiang, L., Manson, C. H., 2006. Defection detection: Measuring and understanding the predictive accuracy of customer churn models. Journal of Marketing Research 43 (2), 204–211.

Novo, J. 2000. Predictive Modelling of Customer Behaviour. [Web:] http://www.jimnovo.com/graphs.htm

Olafsson, S., Li, X., Wu, S. 2008, Operations research and data mining. European Journal of Operational Research 187 (2008) 1429–1448

Pfeifer, P. E., Haskins, M. R., Conroy, R. M., 2005. Customer lifetime value, customer profitability and the treatment of acquisition spending. Journal of Managerial Issues 17 (1), 11–25.

Richards, K.A., Jones, E., 2008. Customer relationship management: Finding value drivers. Industrial Marketing Management 37 (2008) 120–130

Teradata Warehouse Data Mining, 2001. Data Mining for Business Intelligence. [Web] http://www.teradata.com

Van den Poel, D., Larivi`ere, B., 2004. Customer attrition analysis for financial services using proportional hazard models. European Journal of Operational Research 157 (1), 196–217.

APPENDIX A

Lifetime Value Modelling: Test

1. A proper LTV model leads to a couple of benefits, name them?
2. What are the three main components of a LTV model?
3. Before one can develop a LTV model, certain objectives should be considered. Name these objectives and give a short description of each?
4. How does one calculate customer behaviour?
5. Give the formula for Retention and how does this contribute to the LTV model?
6. Churning and retention contributes huge to one another. Why?
7. Give the definition for LTV?
8. Why is OLAP so crucial in LTV modelling and why is this so important for data mining?
9. What does SEMMA stand for and how does this contributes to LTV?
10. Describe how one calculates the present LTV with calculations and formulas?

Answers

1.

- Effective strategising which leads to efficient marketing;
- Gaining customer intelligence; and
- Recognising unprofitable customers.

2.

- Tenure (Product Life);
- Value (Profit); and
- Churn (Closed Status).

3.

3.1 Acquisition Costs

The amount of money Bank XYZ has to spend, on average, to acquire a new customer.

3.2 Churn Rate

The customers who end their relationship with the bank expressed as a percentage for a given time period. Churn rate typically applies to customers whose accounts have become inactive based on the Cheque Accounts

3.3 Discount Rate

The discount rate is the cost of capital used to discount future revenue from a customer. Discounting is an advanced topic that is frequently ignored in customer LTV calculations. The current interest rate is sometimes used as a simple proxy for the discount rate. The discount rate won't be used in this study because the interest rate changes to often in our economy.

3.4 Retention Cost

The amount of money a company has to spend in a given time period to retain an existing customer. Retention costs include customer support, billing, promotional incentives, etc.

3.5 Time Period (Tenure)

Time period is the unit of time into which a "customer relationship" is divided for analysis. A year is the most commonly used time period. LTV is a multi period calculation, usually taking a 3-7 year timeframe into consideration. In practice, analysis beyond this point is viewed as too speculative to be reliable. Time period in this thesis will be known as product life.

4.

Firstly, the Bank needs to understand the data and need to know what it needs after exploring the data. To understand the behaviour of customers the Bank must look at the customers who already have churned (closed account) and only then will it see how a customer behaved. A customer churn rate was calculated and for the LTV of a customer.

5.

$$\mathrm{Re}\,tention = \frac{Total\ \ of\ \ active\ \ accounts\ \ in\ \ a\ \ spesific\ \ product\ \ life}{Total\ \ accounts\ \ of\ \ that\ \ spesific\ \ product\ \ life}$$

Retention as one of the major elements of life time value is measured on its own to give an indication of the number of accounts that are retained in each year the product is opened.

6.

Churn needs retention as this is the main input function for the formula of the churn rate. The main input variable for Churn is retention.

7.

Lifetime value is a way of measuring how many the banks customers are worth over the time they buy its products and services.

8.

When one is using models with large datasets, it will be difficult or nearly impossible to work on software that can't handle huge datasets because this will make analysing and data mining very difficult. With OLAP it is much easier to data mine and to analyse. OLAP is the missing link to data mine because of the huge data sets it can handle and analyse.

9.

To clarify the data mining process, SAS Institute has mapped out an overall plan for data mining. This step-by-step process is referred to by the acronym SEMMA: sample, explore, modify, model, and assess.

LTV needs data mining to find answers for the unsolved problems that normal statistics can't do.

10.

10.1 First of all a retention rate needs to be worked out, hence the formula below.

$$\mathrm{Re}\,tention = \frac{Total\ \ of\ \ active\ \ accounts\ \ in\ \ a\ \ spesific\ \ product\ \ life}{Total\ \ accounts\ \ of\ \ that\ \ spesific\ \ product\ \ life}$$

10.2 Secondly, a churn rate needs to be derived. Churn needs a retention rate before it can be solved, therefore the formula below follows.

$$Churn = 1 - \ \ retention\ \ rate$$

10.3 A Survival analysis is the next step. One needs to see what accounts have survived from the previous year to the current year, thus the formula below:

$$Survival\ \ analysis = Total\ \ accounts\ \ in\ \ that\ \ spesific\ \ product\ \ life \times \mathrm{Re}\,tention\ \ rate\ \ of\ \ that\ \ specific\ \ product\ \ life$$

10.4 LTV, the fourth step, where the survival rate plays a major part. The survival rate multiply by the average profit per product life will equal the LTV.

$$Lifetime\ Value = \Pr oduct\ life\ of\ survival\ analysis \times Average\ profit\ per\ product\ life$$

10.5 The present lifetime value is the lifetime value of each product life added together.

$$\Pr esent\ lifetime\ value = Cumulative\ lifetime\ value$$