

UFS  
UV



BUSINESS SCHOOL  
SAKESKOOL  
BE WORTH MORE  
WEES MEER WERD

**THE IMPACT OF POOR DATA QUALITY ON THE SUCCESSFUL  
IMPLEMENTATION OF A CUSTOMER-RELATIONSHIP MANAGEMENT SYSTEM  
IN THE BANKING ENVIRONMENT –**

**AN APPLICATION TO A MAJOR SOUTH-AFRICAN BANK**

by

**Zanele Jafta  
(B.Com. Econ)  
Field Study**

Submitted in partial fulfilment of the degree  
**MASTER'S IN BUSINESS ADMINISTRATION**  
in the  
**BUSINESS SCHOOL**  
**FACULTY OF ECONOMIC AND MANAGEMENT SCIENCES**  
at the  
**UNIVERSITY OF THE FREE STATE**  
**BLOEMFONTEIN**

**Supervisor: Dr Johan Coetzee**

**20 November 2014**

UNIVERSITY OF THE  
FREE STATE  
UNIVERSITEIT VAN DIE  
VRYSTAAT  
YUNIVESITHI YA  
FREISTATA



## **ACKNOWLEDGEMENTS**

I would like to thank the following people for their contributions:

- My supervisor, Dr Johan Coetzee for his guidance, support and patience.
- Duduzile Ndlovu, for her assistance with data processing and analysis.
- My family and friends for their support and encouragement.
- To all the staff at the University of Free State Business school who were always available and ready to assist.

Above all, I wish to thank God, who gave me strength and perseverance to complete.

## **DECLARATION**

I declare that the field study hereby handed in for the qualification Master's in Business Administration at the UFS Business Scholl at the University of the Free State is my own independent work and that I have not previously submitted the same work, either as a whole or in part, for a qualification at/in another university/ faculty. I also hereby cede copyright of this work to the University of the Free State

Name: Zanele Jafta

Date: 20 November 2014

**Abstract**

*Data quality is critical in business decisions, especially in formulating marketing, customer-relationship and sales strategies. Consistency and the quality of customer data are critical pillars for any business that is looking at introducing a client relationship management (CRM) system. For any successful CRM implementation, it is very important to have a single, integrated view of the customer even though the source data might come from different databases. This study explores the impact of poor data quality on the successful implementation of the customer-relationship management (CRM) system in one of the major banks in South Africa. The study will also determine data improvement and data-management frameworks in CRM systems.*

## TABLE OF CONTENTS

Title	
Declaration	
Abstract	
Table of contents	
List of tables	
List of figures	
LIST OF TABLES .....	8
LIST OF FIGURES .....	8
CHAPTER 1 .....	9
<b>INTRODUCTION AND SCOPE OF STUDY .....</b>	<b>9</b>
1.1. Introduction .....	9
1.2. Background to the banking sector in South Africa .....	9
1.3. Background of Bank X .....	10
1.4. Background to the research problem .....	11
1.5. Research problem.....	12
1.6. The research questions .....	13
1.7. Research objectives.....	13
1.7.1. Primary objective .....	13
1.7.2. Secondary objectives .....	13
1.8. Conclusion .....	14
<b>CHAPTER 2 .....</b>	<b>15</b>
<b>LITERATURE REVIEW .....</b>	<b>15</b>
2.1. Introduction .....	15
2.2. CRM investments and data-quality overview .....	15
2.3. Data quality .....	16
2.3.1. Definition of data .....	16
2.3.2. Sources of data.....	17
2.3.3. Quality of data .....	18
2.3.4. Dimensions of quality data .....	18
2.3.5. The importance of improved data quality .....	21
2.3.6. Poor-quality data.....	22
2.4. Data-quality framework .....	23
2.5. Data-quality management .....	26

2.6.	A framework for data-quality management .....	27
2.6.1.	Data consistency.....	28
2.6.2.	Data validation.....	28
2.6.3.	Data-error correction.....	28
2.6.4.	Data monitoring .....	28
2.7.	Data governance.....	29
2.7.1.	Data-security management.....	30
2.7.2.	System protection.....	30
2.7.3.	Data protection.....	31
2.7.4.	Data-security accessibility.....	31
2.7.5.	Data-forensics readiness.....	31
2.7.6.	Data-compliance management .....	31
2.8.	Customer-relationship management .....	32
2.8.1.	Definition of customer-relationship management .....	32
2.8.2.	Benefits of CRM .....	33
2.9.	Conclusion .....	33
<b>CHAPTER 3 .....</b>		<b>34</b>
<b>RESEARCH DESIGN AND METHODOLOGY .....</b>		<b>34</b>
3.1	Introduction .....	34
3.2.	Research-design strategy .....	34
3.3.	Data-collection strategy .....	35
3.4.	Questionnaire design .....	37
3.6.	Sampling.....	39
3.7.	Ethical consideration.....	40
3.8.	Data management and analysis .....	42
	Reliability of the measuring instrument.....	42
<b>CHAPTER 4 .....</b>		<b>44</b>
<b>PRESENTATION OF RESULTS .....</b>		<b>44</b>
4.1.	Introduction .....	44
4.2.1.	Reliability of the results .....	45
4.2.2.	T-test and post-hoc analysis.....	45
4.2.3.	Analysis of variance (ANOVA) .....	47
4.3.	Interpretation of results .....	48
4.3.1.	Reliability of the scale used .....	48
4.3.2	Biographical information .....	49
4.3.2.1.	Gender .....	50
4.3.2.2.	Age .....	50
4.3.2.3.	Tenure of employment.....	52
4.3.2.4.	Ethnicity.....	53
4.3.2.5.	Primary job roles.....	54
4.3.2.6.	Primary type of data used.....	56
4.3.2.7.	Average scores.....	57

4.4. Perceptions on data quality .....	58
4.5. Perception on data quality management .....	62
4.6. Perceptions on Client Relationship Management .....	64
4.7. Average scores.....	67
4.8. Correlations .....	67
4.9. Conclusion.....	69
<b>CHAPTER 5 .....</b>	<b>70</b>
<b>FINDINGS AND RECOMMENDATIONS.....</b>	<b>70</b>
5.2. Summary .....	70
5.3.1. Findings and recommendations from biographical information.....	70
5.3.2. Findings and recommendations on perceptions of data quality .....	71
5.3.3. Findings on perceptions of DQM.....	72
5.3.4. Findings on perceptions of the value of CRM .....	74
5.3.5. Findings and recommendation on the correlation between all these variables.....	74
5.4. Limitations of the study .....	75
5.5. Future research .....	75
5.6. Conclusion.....	76
<b>ANNEXURES .....</b>	<b>85</b>
Annexure 1 .....	85

## LIST OF TABLES

Table 1: Dimensions of data quality (source: adapted from lehohla 2010:9).....	20
Table 2: Cronbach's alpha coefficient scores.....	47
Table 3: T-test results for different age groups.....	49
Table 4: Post-hoc analysis amongst different age groups.....	49
Table 5: T-test results concerning different tenure of employment.....	50
Table 6: T-test results for different ethnic groups.....	51
Table 7: Post-hoc analysis results for different ethnic groups.....	52
Table 8: T-test of results in different primary job roles.....	52
Table 9: Post-hoc analysis of results amongst different ethnic groups.....	53
Table 10: T-test results for different types of data used.....	54
Table 11: Average scores per demographic group.....	56
Table 12: Average scores.....	65
Table 13: Correlations.....	66

## LIST OF FIGURES

Figure 1: Main problems affecting data quality.....	12
Figure 2: Data lifecycle.....	17
Figure 3: Data-quality framework.....	24
Figure 4: Interaction matrix.....	25
Figure 5: Value, cost and quality of information cycle.....	26
Figure 6: Deming cycle.....	26
Figure 7: Data-quality test scenarios.....	29
Figure 8: Data-governance framework.....	30
Figure 9: Customer-relationship management systems.....	32
Figure 10: Percentage of respondents in different gender groups.....	50
Figure 11: Percentage of respondents in different age groups.....	50
Figure 12: Percentage of respondents in tenure of employment.....	52
Figure 13: Percentage of respondents in different ethnic groups.....	53
Figure 14: Percentage of respondents in different primary job roles.....	54
Figure 15: Percentage of respondents in different types of data used.....	56
Figure 16: Perceptions of data quality – 1.....	59
Figure 17: Perceptions of data quality – 2.....	60
Figure 18: Perceptions of data quality – 3.....	61
Figure 19: Perceptions of data-quality management – 1.....	62
Figure 20: Perceptions of data-quality management – 2.....	63
Figure 21: Perceptions of data-quality management – 3.....	64
Figure 22: Perceptions of client-relationship management – 1.....	65
Figure 23: Perceptions of client-relationship management – 2.....	66

# CHAPTER 1

## INTRODUCTION AND SCOPE OF STUDY

---

### 1.1. Introduction

This research investigates the impact of poor data quality for the successful implementation of Client Relationship Management in one of the major banks in South Africa (SA). Due to ethical considerations, I refer to this bank as Bank X for the purposes of this study so as to maintain the integrity and brand image of the bank.

This chapter looks at the reasons for conducting the research and the framework that will be followed. It will give an overview of the banking sector in SA as well as a brief background of Bank X. The research problem with objectives will also be stated, and this will be followed by literature review. This chapter will briefly explain the research design, ethical considerations and the demarcation of the study. It will then conclude with the complete layout of the research.

### 1.2. Background to the banking sector in South Africa

South Africa has developed an efficient and regulated banking sector which compares favourably with other first-world countries. The South-African banking sector has four major banks in terms of assets, namely Absa, First National Bank, Nedbank and Standard Bank. However, Capitec bank, which dominates in the retail market, is also contesting to be one of the major banks. (Country profile 2007:59)

The banking sector is regulated by the Banks Act, 94 of 1990, and the Mutual Banks Act, 124 of 1993, which looks at an efficient and sound banking system in South Africa. The South-African banking environment is seen to be very stable, and this is evidenced by the number of foreign banks establishing branches in the country and others acquiring stakes in deals with major banks such as the Barclays and ABSA. According to the latest World Economic Forum Competitive Survey 2012/13, South African banks are rated second out of 144 countries for soundness whilst the country

was rated third for financial-sector development (Banking Association of South Africa 2012:5).

### **1.3. Background of Bank X**

Bank X is one of the largest and oldest banks in South Africa, and it has its headquarters in Johannesburg. It also has large operational centres in Durban and Cape Town with over 750 staffed outlets. It has 28 000 employees with over 6 million customers. The bank has five main divisions namely: Capital, Business Banking, Corporate, Retail and Wealth, which are similar to the structures of the other big banks.

A tough competitive environment and technological advancement has seen an increase in companies adopting Client Relationship Management systems (CRM), a system that provides business with information and processes necessary to understand and track customers' behaviours. CRM system is an analysis technique and sophisticated software tool, developed to exploit the potential information contained in databases of customer details and activity (Curtis & Cobham 2002:166). It is often refined into customer profiling and categories which assist in predicting customer behaviour. CRM is also used to maintain relationships in order to increase customer loyalty and profitability.

Many companies think that data quality is an information technology (IT) problem. However, Bitterer (2007:7) argues: "Data quality is not an IT problem. IT can help fix it, but the business must own the problem." Data quality is about the availability, integrity, validity and completeness of a business' sets of data. The consistency and quality of customer data are critical pillars for the successful implementation of a CRM system that enables effective client management and marketing strategies. In many organisations, customer information is stored in multiple touch points. These need to be integrated into one, accurate, current record so that organisations can segment their customers consistently and reliably in order to reveal true purchasing behaviour and patterns to forecast future sales accurately.

Companies invest large amounts of money in CRM, and if the project does not succeed due to data quality, the business stands to lose much money. According to the Ovum report, (Sheina 2010:2), an independent analyst suggests that poor-quality data is costing United States (US) businesses around \$700 billion a year or 30% of the average company's revenue.

#### **1.4. Background to the research problem**

With the introduction of legislative requirements like Treating Customers Fairly (TCF), there is pressure on banks to ensure that they have good-quality customer data that are secured safely. The main purpose of TCF is a regulatory and supervisory approach, designed to ensure that specific, clearly articulated outcomes concerning fairness for financial-services consumers are delivered by regulated financial firms. This has now put pressure on banks to invest in good CRM systems. It means that the quality of data that is migrated into the system must be high. The banks cannot afford to have data of a poor quality as they will not be able to gain a better understanding of their customers. The major challenge of poor data quality is that the bank will not be able to build and maintain customer relationships, and this will have a negative effect on increasing customer loyalty.

Even though the Treating Customers Fairly Bill has not yet been promulgated, banks have now started investing in CRM systems to increase customer loyalty and profitability. However, they are also faced with the problems of data quality. Statistics South Africa (Stats SA), which is a custodian of data in South Africa, is also faced with the same challenge of ensuring that the data quality is good. According to the framework guidelines for South African Statistical Quality Assessment (Lehohla 2010:1), Stats SA is responsible for the collection and dissemination of data concerning official statistics, and it plays a central role in the evaluation and improvement of data quality. It is important for agencies like Stats SA to ensure quality as the government relies on the information to make decisions and allocate budgets effectively.

Data quality is a huge problem in many organisations across all industries. Poor quality data cost money and reduce productivity. A global study on master-data

management was conducted by PricewaterHouseCoopers in 2011. In the study, 49 companies from 12 countries participated in eight different fields, and five of these were financial services. Some of the findings were documented as follows:

- 75% of the companies admitted that defective data had a negative financial impact on their business.
- One survey respondent reported that data problems caused an \$8 million loss in a single fiscal year. (PricewaterHouseCoopers 2011:23)

The companies that were surveyed were also asked to describe the main challenges in maintaining data quality. Companies indicated that their data quality do not meet the data-quality standards as reflected in Figure 1 below. The biggest challenge was data currency.

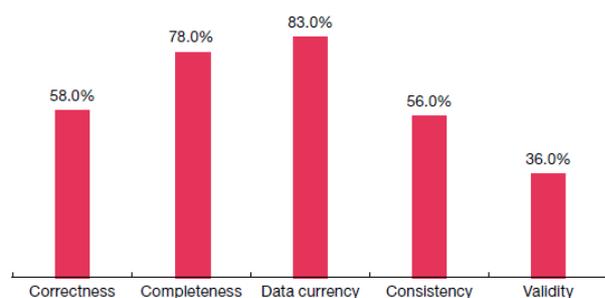


Figure 1: Main problems affecting data quality (Source: PricewaterHouseCoopers 2011:23)

As part of the Client Management Strategy, Bank X has invested over R3 million in a CRM system as they would like to increase and maintain client relationships. This will give them a competitive advantage which will have a positive effect on customer loyalty. It is therefore critical that data that is migrated and used in the CRM system is of good quality so that all the business objectives are realised. As is common in most businesses, banks look at this project from an IT perspective only. It is very important that they consider the business impact if the data quality is not of a high standard.

### 1.5. Research problem

The research problem for this bank is as follows:

Bank X has been using different systems to create and maintain data. It neither specified the key data that should be captured nor ensured that the data is validated. Data creators would put anything into a form to ensure that there are no empty or null fields, for example, email addresses such as noemail@email.co.za. The omission of identity-document (ID) numbers is also a problem. It is therefore important to have proper guidelines on how to create and validate data. There are no strategies to maintain client data and to keep it current. This makes it difficult for the business to implement effective marketing, sales and customer-relationship strategies to increase profitability and loyalty due to unreliable data. Due to the fact that data are kept in different systems with different formats, it becomes a challenge for management to segment customers for differentiated value propositions due to inaccurate data.

## **1.6. The research questions**

The research problem needs to be addressed by looking at the following questions:

- What are the guidelines or framework for creating good-quality customer data in the bank?
- What is the perception concerning the quality of existing data and its degree of reliability and consistency?
- What is the perception concerning maintaining good-quality customer data?
- What is the perception on the value of client-relationship management?

## **1.7. Research objectives**

### *1.7.1. Primary objective*

The primary objective of this study is to investigate the impact of poor data quality on the successful implementation of CRM at one of the major South African banks.

### *1.7.2. Secondary objectives*

The primary objective will be supported by the following secondary objectives:

- determine guidelines for creating good quality data
- identify a data-improvement strategy using a data-management framework

- determine the relationship between data quality, data-quality management and client-relationship management.

## **1.8. Conclusion**

Chapter one introduced the research topic by outlining the background to the banking sector in South Africa and also focused on the background to Bank X. Research problem, research questions and research objectives were also discussed in this chapter.

## **CHAPTER 2**

### **LITERATURE REVIEW**

---

#### **2.1. Introduction**

This chapter presents all the variables that are involved in this study. It will explain in depth the concepts of data quality (DQ), data-quality management (DQM) and client-relationship management (CRM). This chapter analyses the dimensions of data quality and the requirements for good-quality data. The chapter also explains the data-management framework in detail as well as client-relationship management and its benefits to the bank.

#### **2.2. CRM investments and data-quality overview**

Companies have massive amounts of information as a result of digital technology. This has led to a major challenge to store, collect and ensure the appropriate use of this data. Companies use customer-relationship management (CRM) systems to implement integrated and profitable marketing-communication strategies. It is therefore critical for companies to be specific and clear about the types of data that they need to collect in order for them to have a better understanding of their customers.

There has been an increase in the amount of data that is collected by companies from different sources, and this requires the development and maintenance of this large quantity of data to ensure that the data is of a high quality. Companies have to look at the inconsistencies and disparity between these databases as the marketers will find it difficult to trust the quality of the underlying data. The Gartner Group indicates that, despite substantial CRM investments, less than 10% of enterprises have a single view of their customers, and this is a critical stepping stone towards customer loyalty. Customer relationships are dependent on obtaining accurate, reliable and consistent customer data. Bad data includes incorrect and outdated values and missing data like contact details and inconsistent formats (Livesey 2010:6).

Data quality has always been perceived to be an information technology (IT) problem. However, there has been a shift towards businesses taking ownership of data quality. IT is merely an enabler and support in supplying the systems to manage and maintain the data. The business has the responsibility of identifying processes and systems that will assist in capturing customer data and also increase knowledge about the customers. These processes and systems should also maintain the consistency and the quality of data.

The ongoing maintenance of accurate and good-quality customer information is critical to ensure successful CRM systems. This will assist in managing campaigns and also determine customer value. In any company, the analysis and organisation of every customer touch point are critical as these will allow companies to prevent unnecessary expenditure on low-value customers whilst increasing the focus on promotions and spending on lifetime customers of a high value.

## **2.3. Data quality**

### *2.3.1. Definition of data*

Different authors have different way of defining data. Data are raw, specific, undigested and therefore largely meaningless. (Jankowicz 2005:171). Information, in contrast, is what you get when data has been arranged in such a way that data uncertainty is lessened, queries resolved and questions answered.

Mosley and Brackett (2010:2) define data as the representation of facts as text, numbers, figures, images, sound or video. Facts are captured, stored and expressed as data. We use data as raw material to interpret and create information. Data is the foundation of knowledge and information. Information is the intelligent manipulation of data in a context that contributes to knowledge (Mosley & Brackett 2010:2). Data is a valuable asset, and it is important to understand that data is created, stored and maintained for use by the data consumer for business strategies like sales and marketing. When that data is used, it is translated into information which can be analysed and utilised by the data consumer. Like any other asset, data has its own lifecycle which describes the processes performed to manage data assets.

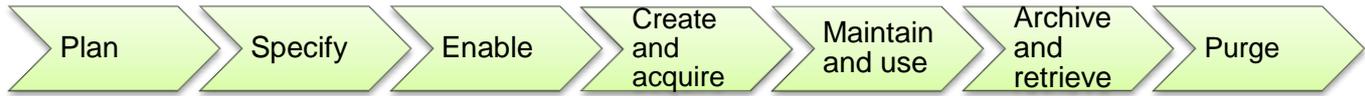


Figure 2: Data lifecycle (Source: Adapted from Mosley & Brackett 2010:4)

Data is collected, stored and exchanged in information systems that are used by companies to provide services to customers and to enhance business processes.

There are many terms that are used in relation to data as explained below:

- data quality: the accuracy, reliability and usefulness of data
- data strategy: internal policies and processes to maintain data quality over time
- big data: unstructured data that is not easily organised or interpreted by traditional databases
- single customer view (SCV): an organisation's single, coherent view of all of a customer's touch-points across all channels and business areas (Global Research report 2014:4).

### 2.3.2. Sources of data

Data can be stored in different formats, depending on the type of data. There are four main types of data:

- electronic: refers to any original or any non-identical copies of digital or other programs and any other data that can be accessed on the computer
- non-electronic: this includes any paper-based data
- structured: represents data that is used by formal applications or machines
- unstructured: can best be described as data domains (types) which cannot be normalised and stored in a typical relational, network or hierarchical database.

The classification of data into different types makes it easy to define quality rules of the data of businesses. There are three main categories of data:

- Stable data: This is data that is unlikely to change. The data will remain stable even though there are additions. The older data remains unchanged.

- Long-term changing data: This is data that does not change frequently, in other words, it has a low change frequency. For example, hotel prices do not change frequently. They would change only during certain periods in a year
- Frequently changing data: This is data that changes frequently. These changes can be random or within a defined frequency, for example, real-time traffic information.

### *2.3.3. Quality of data*

According to Juran (2008:6), data is regarded as high-quality data "... if they are fit for the intended uses in operations, decision making and planning. Poor data quality can cost companies 15% to 25% of their operating profit (Olson 2003:9).

MacGilvray (2008:56) defines information quality as the degree to which information and data can be a trusted source for any or all required uses. It is about having the right set of correct information, at the right time, in the right place, for the right people to use to make decisions, to run the business, to serve customers and to achieve company goals. For companies to determine data quality, they need to look at the following questions:

- How can they create and maintain a database of consistent, high-quality customer information?
- How can consistent, integrated customer interactions be enabled across multiple channels?
- How will they evolve their analytical systems to reveal key customer insights?
- How can customer insights be profitably applied to customer interactions?
- How do they ensure privacy for customers' information?

### *2.3.4. Dimensions of quality data*

Data quality can be defined in different ways as everyone has a different view on this topic. It is mostly dependent on the user of the data. –It is important to understand that data quality is associated with the processes involved in managing data assets. Data quality is defined in terms of the prerequisites of quality as well as the seven dimensions, namely relevance, accuracy, timeliness, accessibility, interpretability, coherence and integrity (Lehohla 2010:7).

There are specific characteristics or dimensions to which data have to adhere so that it can be classified as good-quality data. Some authors look at only six dimensions whilst others refer to eight dimensions (Lehohla 2010:7) and sometimes 12 dimensions (Mosley & Brakett 2010:296-297). For this research, I am looking at the seven dimensions which prescribe the following standards.

According to Peppers and Rogers Group (2003:1), there are a number of challenges that need to be addressed first in order to unlock the value of customer data:

- How do I deal with multiple and disparate data formats?
- How do I consolidate duplicate information?
- How do I ensure accuracy?
- How do I mine and protect the data?

According to a Friedman 2002: 2, a Gartner report titled, “A strategic approach to improving data quality”, 50 percent of enterprises undertaking a CRM strategy are unaware of significant data-quality problems in their environment.

Collecting customer data is not enough to achieve effective CRM. Cleaning and organising data can allow a company to enrich customer profiles with additional information, enabling the organisation to develop unique views on each customer. The Peppers and Rogers Group (2003:3) identify six obstacles to data quality:

- **Interpretability:** In most instances, stored data exists in a wide variety of formats. Before any data-quality processes can be implemented, the source data must be parsed into the appropriate components so that it can be understood by data-quality systems.
- **Consistency:** The standardisation of all data elements is necessary to achieve the best possible results for matching and consolidation.
- **Accuracy:** Incorrect information about customers’ names and addresses is a common and costly mistake that companies make. This has a negative impact of marketing campaigns as the mail will be undeliverable and the printing and postage costs will increase.

- **Completeness:** Incomplete data is often an organisation's biggest challenge in data quality. Information gaps within one database or scattered amongst several customer databases contribute to rising costs.
- **De-duplication:** The unnecessary duplication of data can also become a tremendous resource burden since it requires stakeholders to sift through vast quantities of redundant data for information that should be readily available.

For this research there seven dimensions that will be explored and they prescribe the following standards:

1. **completeness:** a measure by which all required data are included so that there are no missing records
2. **consistency:** a measure by which data adhere to a common definition for its meaning and use
3. **validity:** a measure by which data adhere to defined business rules, accepted values and accepted formats
4. **accuracy:** a measure by which data contain correct values
5. **timeliness:** a measure by which data are current and up-to-date at the time of release or use
6. **accessibility:** a measure by which data can be accessed when required and by the appropriate people
7. **relevance:** a measure by which to provide a clear representation of the activity that the data were meant to measure.

As already indicated, there are other researchers who look at eight dimensions as shown in Table 1 below (Lehohla 2010:9).

Table 1: Dimensions of data quality (Source: Adapted from Lehohla 2010:9)

<b>DATA DIMENSIONS</b>	<b>DESCRIPTION</b>
<b>Relevance</b>	➤ Data should be relevant for the purposes for which it is used and should meet the real needs of clients or the users.
<b>Accuracy</b>	➤ Data should provide a clear representation of the activity it was meant to measure.

<b>Timeliness</b>	➤ Data should be collected and recorded as quickly as possible after the event or activity. It must be released frequently and timeously.
<b>Accessibility</b>	➤ It refers to the ease with which data can be obtained.
<b>Interpretability</b>	➤ It refers to the ease with which data use can understand statistical information through the provision of metadata.
<b>Comparable and coherence</b>	➤ It reflects the degree to which, for analysis, the data could be brought together with other statistical information.
<b>Methodological soundness</b>	➤ It refers to the application of all standards, guidelines and practices to produce statistical outputs.
<b>Integrity</b>	➤ It refers to values and related practices that maintain the user's confidence in the statistical product.

### 2.3.5. *The importance of improved data quality*

Information and data are strategic resources and should form part of business management. Good-quality information gives companies a competitive advantage and enables businesses to be effective. High-quality, well-integrated customer data are the cornerstone of a successful CRM project.

Companies invest large amounts of money in buying the best CRM systems but put very little financial resources into improving the quality of their data. Good-quality data enable CRM strategies to be more effective and efficient. In many cases, CRM's potential to improve customer service and increase revenue has been prevented by inconsistent data sets that prevent the organisation from having a clear, unified profile of each customer. Data quality gives an indication on how well the company matches up with the real world at any given time as they need to have accurate customer data. There are many sources of 'dirty data'. According to Petersen (2003:1), Cutter Consortium identified the following sources:

- poor data entry, which includes misspellings, typographical errors and transpositions and variations in spelling or naming
- data missing from database fields
- lack of company-wide or industry-wide data-coding standards

- multiple databases scattered throughout different departments or organisations with the data in each of these structured according to the rules of that particular database
- older systems that contain poorly documented or obsolete data

There are many benefits for the businesses in keeping clean data:

- Correcting inaccuracies in customer data will lead to a reduction on high printing, mailing and marketing production costs. The elimination of duplicate records can prevent that customers become frustrated from receiving multiple mailings of the same marketing piece.
- Quality data contributes to operational savings.
- The deletion of redundant data eliminates the cost of maintaining duplicate records in databases.
- Clean and integrated customer records are essential to a CRM initiative. An effective database is vital to CRM success. A single and accurate view of the customer will spur all retention and acquisition efforts.

Loshin (2001:10-15) also explains the reasons why businesses should take care of data quality by looking at the following benefits of good-quality data:

- improves the effectiveness of marketing, sales and customer service
- enhances new-customer acquisition
- protects customer relationships from the brand damage – marketing fatigue
- increases return on investment in information technology.

#### *2.3.6. Poor-quality data*

There are different causes of poor data quality. These may be caused by process, human and systems issues. There are also issues of using different systems to capture the data, which will result in the data not being in a consistent format.

According to MacGilvray (2008:5), there are a few indicators of poor data quality:

- A high return rate in posted mail as a result of incorrect or incomplete addresses
- multiple delivery of the same documents or advertising material which can be tracked through customer complaints

- incorrect forms of address and address lines, for example, addressing a female client as Mr Carla and vice versa
- reprocessing of orders
- dealing with customer complaints.

The effects of poor data quality are financial and also lead to business inefficiencies. The financial cost is always underestimated by companies, and they tend to blame information systems. Poor data:

- leads to operational inefficiency in the business
- constrains decision making
- leads to customer attrition
- leads to a breakdown in organisational confidence
- restricts system- and data-migration projects.

#### **2.4. Data-quality framework**

Every organisation should have a strategy for data quality relating to collection, storage, security and maintenance. The strategic objectives for data quality should be linked to the success of the company and its organisational goals (MacGilvray 2008:4). Even though the goals and strategies might be company specific, there are common goals for CRM systems, namely:

- improving communication
- creating multiple touch points to serve the customer
- a strong relationship with enterprise resource planning (ERP)
- reducing waste in the mailing budget
- complying with data regulations.

It is not easy to achieve the level of integrated, consistent customer data necessary to support CRM. In most cases, CRM implementations are forced to use customer information sourced from multiple data stores. In order to ensure the correct use of the data from these different sources, the following four basic questions must be asked:

- Where do data reside?

- What format does it have?
- Where are duplicated data?
- Do the overlapping data have incremental value?
- Which data sources are the most reliable?

To ensure that there is good-quality data, a data-quality framework is necessary. It has seven main sections as shown in the diagram below.

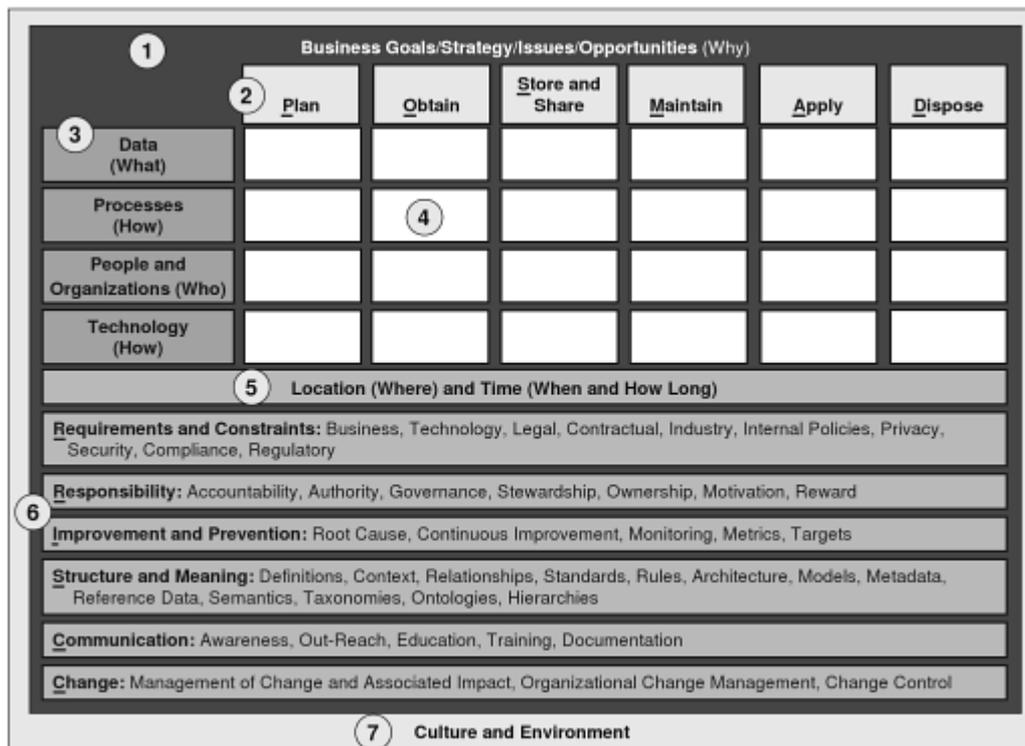


Figure 3: Data-quality framework (Source: MacGilvray 2008:84)

Figure 3 looks at the data-quality framework that must be followed to ensure that data quality is maintained. The numbered paragraphs below refer to the numbers one to five in the figure.

- 1) The main section refers to the business goals and strategy which drives all decisions in any organisation.
- 2) This section deals with the information life cycle, which entails planning, obtaining, storing and sharing, maintaining, applying and disposing of the information once it has served its purpose (POSMAD life cycle).

3) There are four components that affect the data throughout its life cycle, namely:

- data (what)
- processes(how)
- people and organisation (who)
- technology (how).

Each of these components has to be accounted for in each of the phases of the life cycle discussed above.

4) The matrix shows the phases of the information life cycle and the key components of data, processes, people and technology as shown in the diagram below. This is known as the interaction matrix. All the phases of the lifecycle (POSMAD) have a cost attached to them. This could be financial or business cost such as time wasted on poor data.

	PLAN	OBTAIN	STORE AND SHARE	MAINTAIN	APPLY	DISPOSE
DATA						
PROCESS						
PEOPLE AND ORGANISATION						
TECHNOLOGY						

Figure 4: Interaction matrix (Source: Adapted from MacGilvray 2008:84)

5) Location and time: It is important to know where the information is coming from, when will it be available and for how long it will be available.

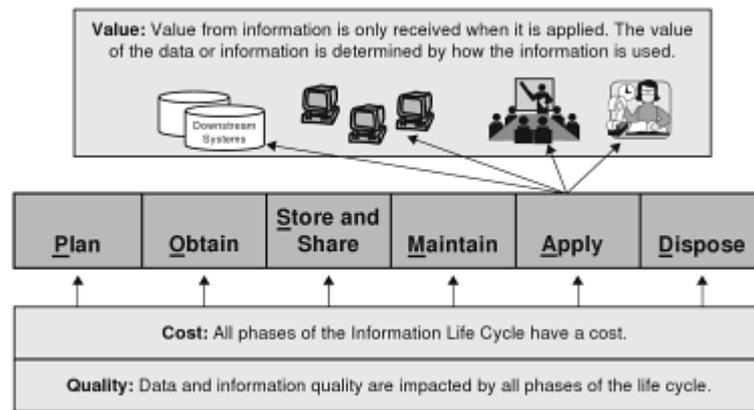


Figure 5: Value, cost and quality of information cycle (Source: MacGilvray 2008:92)

## 2.5. Data-quality management

Data-quality management (DQM) involves analysing the quality of data, identifying data anomalies and defining business requirements and corresponding business rules for asserting the required data quality (Mosley & Brackett 2010:291).

Data-quality management goes through a cycle. It follows the Deming cycle of “plan-do-check-act”.

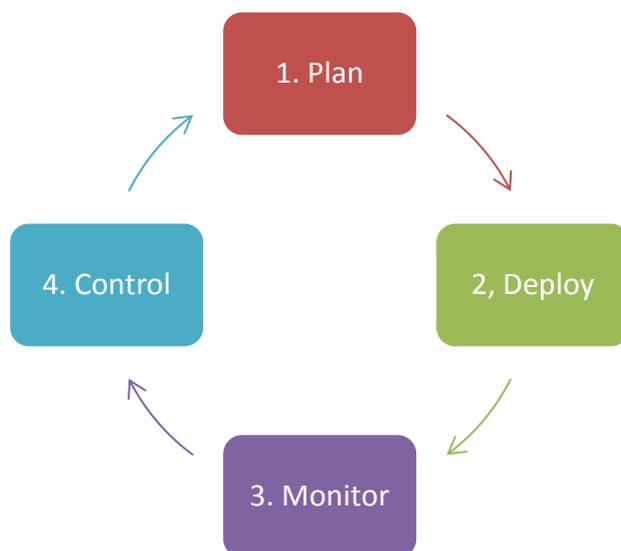


Figure 6: Deming cycle (Source: Adapted from Mosley & Brackett 2010: 292)

Mosley and Brackett (2010:292) explain the Deming cycle which they have adapted to “plan-deploy-monitor-act” as follows:

- Plan: It refers to all activities related to the definition of a strategy and approach used to develop the opportunities that an organisation wishes to pursue.
- Document (deploy): These activities deal with the formal documentation of the processes and procedures executed as part of the daily operations of the organisation. These capabilities help to build a clear context around the way in which the activities of an organisation are defined and maintained for operational excellence.
- Execute (monitor): These activities look at the practical execution of processes within the organisation and the areas of attention that should be considered to support uniformity in the fulfilment of tasks by different areas.
- Control (act): It relates to the continuous control and measuring of results and the impact of the actions taken by the organisation. It also supports the assessment of possibilities for continuous improvement (GS 1 2010:7).

This adaptation makes sense for application in the DQM framework because it is important to start by identifying the issues that relate to the current state and measurement of the current data quality. Improvement principles can only be applied once the integrity of the data is known. Monitoring the maintenance of data and resolving any issues will be the last part of the cycle.

## **2.6. A framework for data-quality management**

The data-quality framework is based on a protocol for data quality in every industry that consists of:

- a system for data-quality management to validate the existence and effectiveness of key data-management business processes
- an inspection procedure to physically validate product attributes (GS 1 2010:7).

The main component of the data-quality management framework (DQMF) is the data-quality management system (DQMS), which is defined as a series of documented, periodically reviewed procedures that are implemented within an

organisation to maintain and support the production of good quality data (GS 1 2010:9).

In order for the DQMF to be implemented successfully, the following points must be taken into consideration. :

#### *2.6.1. Data consistency*

Data profiling within source systems allows inaccuracy and inconsistency issues concerning the data to be exposed and resolved, resulting in trustworthy, high-quality data. Data profiling is used to identify incorrectly uploaded data or false information originating from user-input errors or errors within acquired data.

#### *2.6.2. Data validation*

Routines, often called "validation rules" or "check routines", are used to verify the correctness, meaningfulness and security of the data that originates from input throughout the entire data path. The rules should be implemented through automated facilities, making use of data dictionaries or the inclusion of explicit application program-validation logic. Incorrect data-validation processes can lead to data corruption or security vulnerability.

#### *2.6.3. Data-error correction*

Data-error correction is the process which corrects errors in data quality that were identified during the data-profiling and data-validation processes. This process does not only correct false information but also alerts the organisation when information is manipulated. Ultimately the process to correct errors in data ensures that future data errors are prevented.

#### *2.6.4. Data monitoring*

Data monitoring is the process of continuously monitoring data quality and categorising data domains (types) into categories such as critical, important and insignificant to business users (Nedbank 2011:14). Tests are put in place to continuously check the quality of data. Through data monitoring, the organisation will be able to rapidly respond to data-quality violations and vulnerabilities. It will furthermore allow the organisation to identify structural changes in the data.

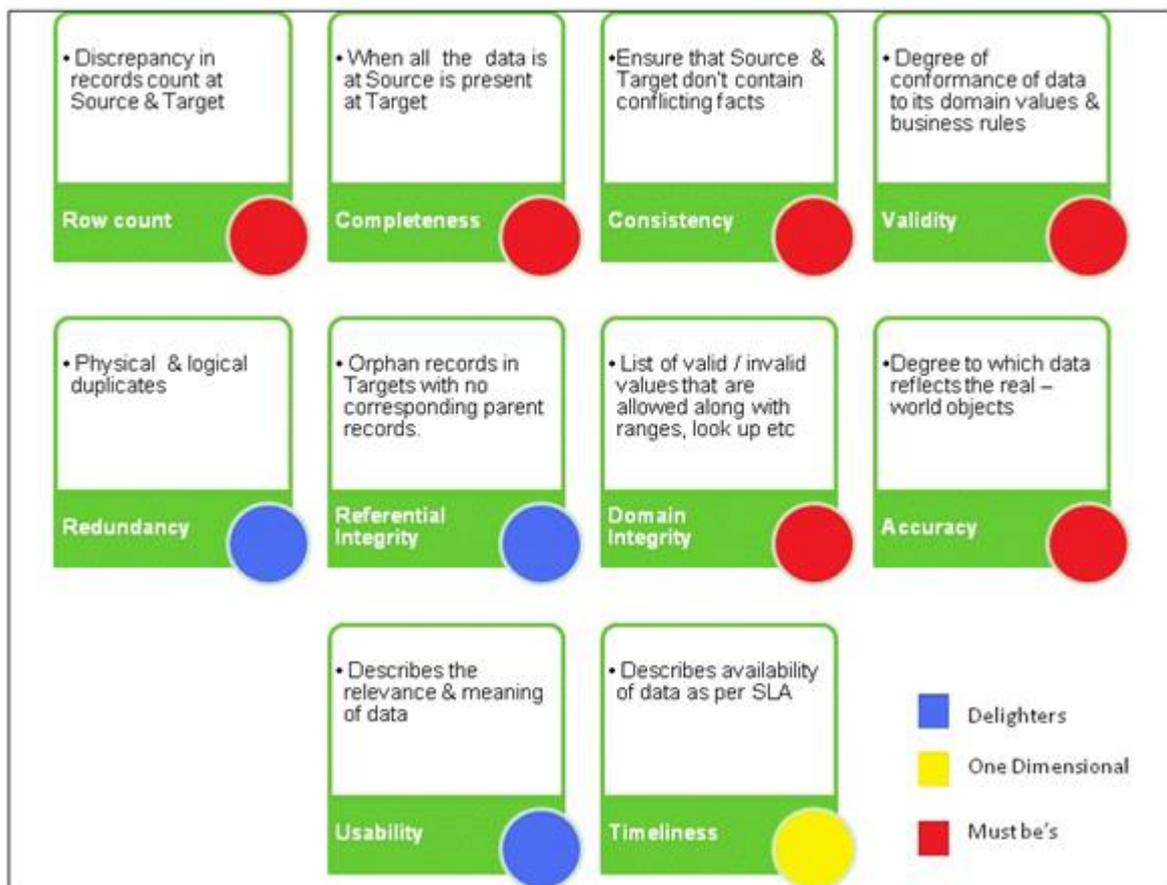


Figure 7: Data-quality test scenarios (Source: Microsoft Developer Network n.d.)

As part of monitoring, data-quality testing can be done at certain periods to ensure that the data are at the required level as determined by the company.

## 2.7. Data governance

Nedbank's Data Management Master Policy (Nedbank 2011) states:

Data Governance can be defined as the people (organisation), processes, procedures, policies, standards and technologies that ensure data value (alignment), quality improvement (information), single shared (accepted) definitions, consistency across application systems and business processes, audit ability, security, accessibility and availability at the right time to the right people. (p. 14)

The aim of data governance is to develop and embed a culture of accountability and responsibility through clear policies and procedures.

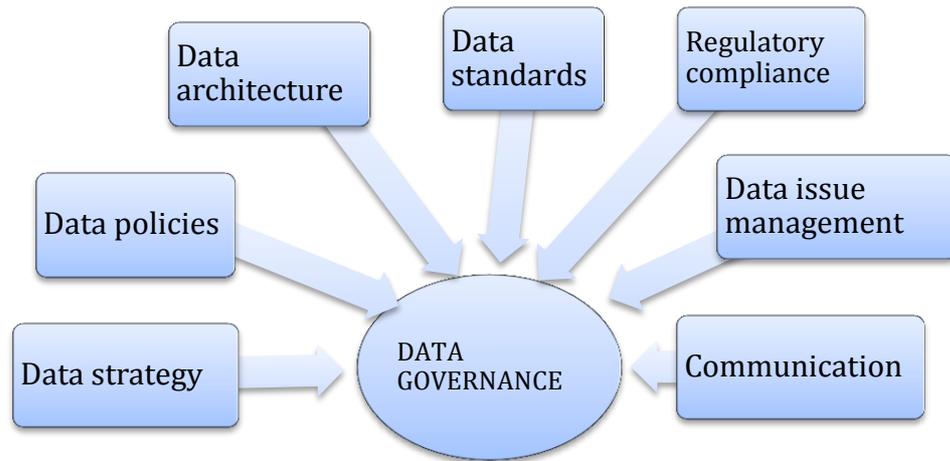


Figure 8: Data-governance framework (Source: Nedbank 2011:14-15)

In this study, the researcher only looks at the regulatory compliance as it is a big component of the bank. As part of data governance, it is important to know what kind of data is available to whom and for what purpose. Therefore data-security management is a very important component of data governance. This is to ensure that companies comply with the legislative requirement like BASEL III.

### *2.7.1. Data-security management*

Data security represents the protection of the organisation's data. There are many ways of ensuring that the data is secure.

### *2.7.2. System protection*

System protection is put in place to ensure that the data which is contained within the organisation's systems are protected from various security risks. These protection mechanisms typically include firewalls, anti-spyware, antivirus software and data encryption.

System protection should not be regarded as a once-off process but rather as a continuous function with includes regular updates of all software programs (this includes all software programs which are used to ensure the security of the data residing in various systems). Systems protection should make use of the correct wireless protocols and keep up-to-date with the latest security risks. Included in system protection are the processes which dictate the disaster protection measures that should be in place. The processes should ensure that disaster recovery

functions optimally in a secure environment. A further consideration concerning system protection is the physical protection measures that should be in place to protect the systems where data resides. These measures include access control to buildings (server room), alarm systems, close-circuit television, cooling systems, operating-system updates and similar measures. Typically these measures would dictate that data servers are housed in physically secure environments which comply with the vendor's requirements for environmental conditions such as the temperature, humidity and power sources. Power sources need to be secure in that back-up power is available and sufficient to ensure that servers are not damaged during power outages.

### *2.7.3. Data protection*

Data protection represents the protection of data against unauthorised disclosure, alteration, restriction and destruction. This implies the protection of clients' rights and ensuring that information remains confidential, thus ensures the protection of sensitive data. Furthermore, unauthorised access, modification and denial of access are covered under data protection. Data-protection measures include access rights according to relevant levels of authorisation, password policies, rules and regulations which dictate the process of the deletion and alteration of data items, disclosure requirements, non-disclosure agreements and the establishment of disciplinary processes to discourage the fraudulent disclosure of information.

### *2.7.4. Data-security accessibility*

Data must be readily available to authorised personnel. It should also include passwords so as to identify the fraudulent misuse of authority.

### *2.7.5. Data-forensics readiness*

When a serious incident concerning data security occurs, forensic investigations of digital evidence should be employed as a response to such an incident.

### *2.7.6. Data-compliance management*

Data-compliance management relates to the compliance of data as specified by regulations which impact any organisation. Audit trails are the main drivers for data-compliance management.

## 2.8. Customer-relationship management

### 2.8.1. Definition of customer-relationship management.

Customer-relationship management (CRM) refers to the tactical management of the customer's interaction with the company, the moments of truth, by proactively using tools such as campaign-management software (Batini & Scannapieco 2006:49). Customer-relationship management is explained as a core business strategy, implemented by firms to maintain valuable relationships with their clients (Rootman 2006:87). CRM deals with the segmentation of customers, the analysis of the data to ensure targeted marketing and effective communication with the target customers. It uses information technology, including data warehousing to deliver and extract good customer value.



Figure 9: Customer-relationship management systems (Source: Benefits of CRM n.d.)

CRM deals with aspects of the work with clients that is concerned with service-related issues or marketing procedures.

CRM is increasingly associated with cost-saving and streamline processes. According to a survey by Gartner Inc., the world's largest technology research and advisory firm, more than 50% of CRM implementations are considered failures from the company's point of view. One of the most important contributors is poor data quality. This only becomes apparent when the CRM project is already underway.

The collection of customer data is an important step in understanding and learning about customers, segmentations and the analysis of customer value. When developing the CRM system, one has to ensure that the data are of good quality as this will provide the company with competitive advantage in establishing and managing customer relationships. According to organisational learning theory, the customer-information process involves, firstly, acquiring the data and, then, transforming the data into customer information.

### *2.8.2. Benefits of CRM*

When companies look at implementing CRM systems, they consider the benefits of these systems as these are very costly systems to acquire and maintain. Here are some benefits of using CRM systems:

- Decrease the cost of customer acquisition: Data enable companies to correctly identify their target audience and to focus all of the marketing efforts on that particular group of people.
- Increase sales: It analyses current customer-service practices and can be used to pinpoint shortcomings and areas that need improvement. Better customer service will ultimately result in higher volumes of sales.
- Increase efficiencies: Employees can access important information quickly, and the process can be automated.
- Better and more accurate data: Analytics and reporting allow the sales, marketing and customer-service teams to work together and introduce improvements.
- Data security: Not only does a CRM manage your data, but security allows you to control who has access to certain data and features (Benefits of CRM n.d.:6).

## **2.9. Conclusion**

This chapter looked at data quality and its dimensions. It is also important to know how to maintain data quality and the framework used for execution. The chapter was closed with explaining customer-relationship management.

# CHAPTER 3

## RESEARCH DESIGN AND METHODOLOGY

---

### 3.1 Introduction

Chapter three focuses on research design, methodology, including data collection, questionnaire design, sampling, ethical considerations and a summary of data management and analysis. Cooper and Schindler (2011:139) argue that there is no single definition of research design. On their part, they define it as the plan and structure of investigation so conceived as to obtain answers to research questions. It also constitutes the blueprint for the collection, measurement and analysis of data.

Research design is the plan according to which the research will be conducted and needs to be done so that the evidence obtained will enable the researcher to answer the initial question as unambiguously as possible. This research will be done to understand the impact of quality data on the successful implementation of the client-relationship management tool.

### 3.2. Research-design strategy

A research design provides a framework for the collection and analysis of data. A choice of research design reflects decisions about the priority given to a range of dimensions of the research process. These include the importance attached to:

- expressing causal connections between variables
- generalising to larger groups of individuals than those actually forming part of the investigation
- understanding the behaviour and the meaning of that behaviour in its specific social context
- having a temporal (i.e. over time) appreciation of social phenomena and their interconnections (Bryman & Bell 2007).

This study will be a descriptive study, following a quantitative methodology. A descriptive study is concerned with finding answers to the questions of who, what,

where, when or how much (Cooper & Schindler 2011:141). This means that the researcher will look at the current environment as it is and not make any changes to the environment. The variables will not be manipulated but are measured as they occur. A descriptive study does not identify the cause of differences or relationships but measures whether there are such difference or relationships. The goal is to determine the relationship or association between an independent variable and a dependent variable. The descriptive study will be conducted to demonstrate the impact of quality data on the implementation of the CRM system.

The researcher will follow a quantitative methodology. This methodology explains phenomena by collecting numerical data that are analysed, using mathematically based methods. Quantitative research deals in numbers, logic and the objective, focusing on logic, numbers and unchanging static data and detailed, convergent reasoning rather than divergent reasoning (Cooper & Schindler 2011:144). The data are usually gathered by using more structured research instruments like questionnaires. This method uses a larger sample size, which is representative of the population.

The main advantage of quantitative research is that it allows for greater objectivity and accuracy of results. Generally, quantitative methods are designed to provide summaries of data that support generalisations about the phenomenon under study. Quantitative research also has its disadvantage in that, even though it is more efficient and able to test hypotheses, it may miss contextual detail. The results provide less detail on behaviour, attitudes and motivation. Results are limited as they provide numerical descriptions rather than detailed narrative, and they generally provide less elaborate accounts of human perception (ACAPS n.d.:5).

### **3.3. Data-collection strategy**

The quantitative data were collected in the form of a structured questionnaire, which I developed myself. The questionnaire is most frequently a very concise, preplanned set of questions designed to yield specific information to meet a particular need for research information about a pertinent topic. (Key.1997: 1)

Questions need to be simple and easy to understand and the questionnaire has to be clear and easy to complete because no interviewer is available to assist the respondent. (Constantinos 284: 4)

Surveys using self-completion questionnaires have some distinct advantages over face-to-face interviews:

- They are cheap to administer. The only costs are those associated with printing or designing the questionnaires, their postage or electronic distribution.
- They allow for a greater geographical coverage than face-to-face interviews without incurring the additional costs of time and travel. Thus they are particularly useful when carrying out research with geographically dispersed populations.
- Using self-completion questionnaires reduces biasing error caused by the characteristics of the interviewer and the variability in interviewers' skills.
- The absence of an interviewer provides greater anonymity for the respondent. When the topic of the research is sensitive or personal it can increase the reliability of responses.

This questionnaire approach has been chosen because of simplicity and the ease of data collection. The questionnaire was a self-administered questionnaire due to the flexibility of use and follow-up. This approach also had lower costs than other methods, and it was also possible to reach a wider geographical area. The self-administered questionnaires were sent *via* email as the researcher had access to all the emails of the sample, and it would have been easy to follow up on responses. The self-administered questionnaire offered the participant time to think about questions in the questionnaire. The main disadvantage of this method is that it could have a low response rate due the fact that the respondents can respond in their own time.

The self-administered questionnaires were collected from employees of the bank's Wealth Division who are client-facing as well as from business-support services. The

seven-point Likert scale was used to measure the impact of poor data quality. The biggest advantage of the Likert scale is that it is easy and quick to construct. According to Cooper and Schindler (2006:339), using a Likert scale will probably provide more reliable data than many other scales. The Likert scale also assists in the data analysis process because the researcher will use the frequency-distribution model for data analysis.

### **3.4. Questionnaire design**

The questionnaire was designed following the basic principles of simplicity, conciseness and ease of navigation (Gendall 2011:7). The questions were kept short, simple and concrete. Closed questions were asked throughout the questionnaire. The layout and presentation of a questionnaire should help the respondent to “navigate” through it and encourage the respondent to respond (Gendall 2011:3). The questionnaire was formulated and divided into four sections to ensure that it is easier for the respondents to navigate and understand.

Section one of the questions focused on biographical information whilst the other three sections focused on the perceptions of the respondents concerning data quality, data-quality management and the value of CRM.

Section two focused on the perceptions of data quality: A total of 11 questions were formulated to test the perception of the respondents on data quality in the bank. The questions were based on the data-quality dimensions, namely, accuracy, reliability, timeliness, accessibility and relevance. These principles have been described in the literature review in Chapter Two of this research. Six of the eleven questions asked tested whether or not the existing data adheres to the principles of data quality and were phrased as follows:

- The information is consistently presented in the same format.
- The information is easily retrievable.
- The information is accurate.
- The information is easily accessible.
- The information is reliable.

- The data is relevant to my work.

Section three focused on the perceptions concerning data-quality management (DQM). A total of ten questions have been formulated to test the perceptions of DQM. The questions concentrated on determining whether the respondents are aware of the existence of a DQM strategy and whether the bank has the necessary tools to manage data quality now and in the future. All the questions were formulated based on the DQM framework which has been discussed in the literature review in Chapter Two of this research. Eight of the ten questions were asked to determine whether the tools for DQM exist in the bank. The two self-developed questions determine the respondents' perceptions of the special focus on DQM and were formulated as follows:

- There are people whose primary job it is to ensure the quality of information.
- The bank has a specific business unit responsible for information quality.

Section four focused on client relationship management (CRM). A total of ten questions were formulated as a basis of measuring the perception concerning the value of CRM with the focus on the following questions:

- The value of the CRM tool is well understood by me.
- The CRM tool will assist the bank with cross-selling opportunities.
- The value of the CRM tool has been well communicated.
- CRM will assist the bank to broaden its marketing initiatives, capabilities and efficiencies.

A Likert scale was used in the questionnaire. This is a very useful type of questionnaire when you want to obtain an overall measurement of a particular topic, opinion or experience, and it also assist in collecting specific data on contributing factors (Bertram 2011:8). The Likert scale has been chosen as it is easy to construct and also produces highly reliable data. A Likert scale assumes that the strength or intensity of experience is linear, i.e. on a continuum from strongly agree to strongly disagree, and it makes the assumption that attitudes or opinions can be measured (McLeod 2008:2).

A seven-point Likert scale was used as follows: (1) Very strongly disagree, (2) Strongly disagree, (3) Disagree, (4) Neutral, (5) Agree, (6) Strongly agree, (7) Very strongly agree. Six questions gathered the biographical data of the respondents in the first section of the questionnaire, namely the gender, age, years of service at the bank, ethnicity, and primary job role, primary type of data used. The questionnaire is included as Annexure 1.

### **3.5. Testing and evaluation**

On quantitative surveys, a small number of pilot (test) interviews should ideally be conducted to test the feasibility and appropriateness of the questionnaire (MRS 2011:21). For the purposes of pilot interviews, the questionnaire was given to five colleagues who were not involved in the research at the bank and do not work directly with the bank's data as they are involved in strategy formulation and project management. This pilot group was involved to test the following:

- Navigability: Was it easy to navigate and use?
- Ambiguity: Were there any questions that could be interpreted differently from the researchers' intentions?
- Duration: Is the questionnaire too long, and how long does it take to complete the questionnaire.

Feedback received from this pilot group indicated that the questionnaire was easy to navigate, the questions were clear, and it would take a maximum of fifteen minutes to complete the questionnaire.

### **3.6. Sampling**

The researcher used the entire population of 120 staff members for sampling purposes, which were comprised of members from client-facing and business-support services. A response rate of 56% was achieved as 67 questionnaires were useable for analysis. This sample is based at the bank's head office in

Johannesburg. A sample of the entire population can be chosen for any of the following reasons (Chaturvedi 2007:6):

- if your population is small
- when you have extensive resources
- when you do not expect a very high response.

The whole population was used due to the size of the population as it was small, and it also increased the chances of a higher response rate.

The sample frame was a list of email addresses for all the participants that would have been participants in the survey. The researcher ensured that the sample frame was correct as these were all email addresses of participants that are still in the employment of the bank. A sampling-frame error was eliminated as the researcher had all the business email addresses of the participants. Since the whole population was selected, all the staff members had an opportunity of being selected in the sample, and therefore, probability sampling had been used. Probability sampling is a controlled procedure that ensures that each population element is given a known non-zero chance of selection (Cooper & Schindler 387:6). The self-completion questionnaires were distributed to the entire population. As indicated above, a response rate of 56% was achieved, as 67 questionnaires were useable for analysis.

### **3.7. Ethical consideration**

Ethics is defined as what is deemed to be acceptable or unacceptable in human conduct (Tustin & Martins 2005:43). It is about doing what is right and good beyond law and regulation. Ethics cannot be regarded as optional extras. According to Jankowicz (2005:326), there is a general ethics rule that applies to research: The researcher must treat the responses from the respondents anonymously unless they gave permission to be mentioned as providers of that information.

It is crucial to obtain informed consent from all of the participants who are willing to complete the questionnaire. The participants and the management team were made

aware of the type of information required and what it was needed for. They were also informed of the purpose of the study

Family, educational and religious institutions play a big role in determining the subset of values held by society. Businesses have now entrenched ethics as a normal business practice, and it is an important aspect of research. Good ethics build a good foundation for professionalism which is also a requirement for a researcher. Two key ethical issues that should be considered in any project are consent and confidentiality:

### *3.7.1 Consent*

All participants consented voluntarily to participate without being coerced or pressurised unfairly. They were well informed about the purpose of the study and what their participation entails.

### *3.7.2 Confidentiality*

It is essential to protect the identity of the participants, and this was communicated to all the participants.

For this research, the following ethical standards were followed:

- Permission was sought from the company, and the research topic was approved by the General Manager of the business unit and the Chief Operations Officer of Wealth Management.
- Permission was also sought from the company's risk and compliance division and signed confidentiality documents as the bank's data and information would have been utilised.
- Transparency and openness was displayed as the necessary introductions were made to the participants.
- The purpose of the questionnaire and the benefits of the study was thoroughly explained to the participants. This was done when the email with the questionnaire was sent out to the participants.
- It was also explained to the participants that completion of the questionnaire was not compulsory but voluntary.

All responses from the participants were treated with confidence, and the participants were informed as such.

### **3.8. Data management and analysis**

All quantitative data from the respondents was collated and electronically captured on Microsoft Excel. Exploratory data analysis was used as the researcher had the flexibility to respond to the patterns revealed in the data. The data was analysed using the Statistical Package for Social Science (IBM, SPSS statistics 22). The results were summarised by using frequency distribution figures to identify the percentage of respondents for each question and biographical information. Other statistical tests, namely Analysis of Variance (ANOVA) and T-tests, were performed in order to establish perceptions and the relationships the between variables.

#### **Reliability of the measuring instrument.**

Validity and reliability are two important elements in the evaluation of a measurement instrument and they are closely associated. Validity is concerned with the extent to which an instrument measures what it is intended to measure and reliability is concerned with the ability of an instrument to measure consistently (Tavakol and Dennick 2011: 1). Internal reliability of the instrument used to test the variables was assessed using Cronbach Alpha. Cronbach Alpha provides a measure of the internal consistency of a test or scale and it is expressed as a number between 0 and 1 (Tavakol and Dennick 2011: 1). It describes the extent to which all the items in a test measure the same concept or construct. Cronbach Alpha coefficient scores were calculated using the SPSS computer program. The Cronbach Alpha coefficient scores and a detailed discussion on the reliability of the measuring instrument will be discussed in Chapter four.

### **3.9. Conclusion**

In this chapter, the research design and methodology, including data collection, questionnaire design, sampling, ethical considerations and a summary of data management and analysis were discussed. The assessment questionnaire was

circulated via email. A cover letter was attached. It contained a brief summary of the basic scope of the study involved and permission to conduct the study. Testing and evaluation of the questionnaire and the reliability of the instruments used were also discussed.

## **CHAPTER 4**

### **PRESENTATION OF RESULTS**

---

#### **4.1. Introduction**

Chapter four focuses on the analysis of the results. The results are provided in five different categories, namely, the internal reliability of the measuring instrument, biographical data of the respondents, perceptions on data quality, perceptions on data quality management, perceptions on the value of client relationship management and the correlations between the three variables. Average scores on the biographical data and on all the respondents will also be discussed. As already explained in Chapter three, frequency distribution graphs were used to depict the results and other statistical tests like the T-test and ANOVA results are depicted in tables. The types of statistical tests used and the analysis of the data will also be explained in this chapter.

#### **4.2. Analysis of results and test instruments used**

The results from biographical information are presented in frequency-distribution figures. Frequency distribution is a numerical display, showing the number and percentage (the relative frequencies) of cases that correspond to each value or group of values of a variable (Chambliss & Schutt 2012:175). This has been used to show the variations amongst the different groups of respondents.

The findings from each of the variables, namely, perceptions of DQ, perceptions of DQM and the perceptions of the value of CRM have been presented separately where T-tests and ANOVA were used to describe the statistical relationship between them. The results were presented in tabular format and also depicted in figures. Finally, the correlation between all these variables is presented. Descriptive statistics were used to interpret the biographical data as it simply describes what is going on in the data. Inferential statistics were used to interpret the rest of the sample data as the whole population was used as a sample. It will make it easier to draw inferences

from our data to more general conditions. Inferential tests will provide you with a 'p-value'.

The response rate was 56%, and the results indicate a good representation of the targeted audience in terms of gender, age, time period in the position (experience), ethnicity, primary job roles and primary type of data used.

#### *4.2.1. Reliability of the results*

As already mentioned in Chapter three, Cronbach Alpha coefficient scores were calculated in order to assess the internal reliability of the measuring instrument. Cronbach alpha was calculated so as to determine the reliability of the variables constructed. *The Cronbach alpha determines the reliability of the scale used as well as to test whether or not the items used for a construct or variable actually measure that particular construct.* Cronbach Alpha coefficient scores with a value of more than 0.70, as recommended by Nunnally and Bernstein (1994:264-265), were recorded for each of the six factors. The measuring instrument can therefore be considered as reliable (see Table 6.1).

#### *4.2.2. T-test and post-hoc analysis*

A T-test is used when you want to see if the group means of interval variables differ from one another, and it is only used when two groups are to be compared (Institute for Digital Research and Education n.d.:3). For conclusions on whether the perceptions are different or not, we looked at the p-value and compared it to a value of 0.1 (10%).

### **Tests of Normality**

In order to do a t-test and analysis of variance, a Kolmogory-Smirnov test was done to test whether the three variables, data quality, data quality management and client relationship management are normally distributed. This test has to be done because the variables have to be normally distributed for one to use the t-test and ANOVA. The test used is the Kolmogorv-Smirnov test. We look at the p-value and compare it

with a significance level of 0.1. If the p-value is less than 0.1 we conclude that the variable is not normally distributed.

Table 2: Tests of Normality

Variables	Kolmogorov-Smirnov		
	Statistic	Degree of freedom	P-value
Data quality	0.062	64.000	0.200
Data quality management	0.100	64.000	0.141
Client relationship management	0.093	64.000	0.200

All p-values are greater than 0.1 which means that all 3 variables are normally distributed. This means that we can use the t-test and ANOVA to test for significant differences among groups.

A subjective judgement regarding the degree of risk that you are willing to accept for being wrong needs to be made by selecting a confidence level. Typical confidence levels are 90%, 95% or 99% (ArcGis Resource Centre n.d.:4). The p-value associated with a 90% confidence level is 0.1, which indicates the risk the researcher is willing to accept for being wrong.

If the p-value is less than 0.1, there are different perceptions of data quality, data-quality management and client-relationship management. If the p-value is greater than 0.1, there are no different perceptions of data quality, data-quality management and client-relationship management.

If there are differences after the T-test is done, post-hoc analysis is done to determine which variables displayed the most differences. Where the researcher has found significant differences, a Tukey test needs to be done to determine the factors that showed the most differences (DeCoster 2006:38).

## Steps

1. Calculate an analysis of variance (e.g. one-way between subjects ANOVA).
2. Select two means and note the relevant variables (means, mean square within and number per condition or group).
3. Calculate Tukey's test for each mean comparison.
4. Check to see if Tukey's score is statistically significant with Tukey's probability or critical-value table, taking into account the appropriate  $df_{\text{within}}$  and number of treatments (Hall 1998a:4).

### 4.2.3. Analysis of variance (ANOVA)

Analysis of variance (ANOVA) is also used when you want to see if the group means of interval variables differ from one another, but it is used when you have three groups or more to compare (DeCoster 2006:47). An ANOVA test was done to determine whether or not there are significant differences in perceptions of DQ, DQM as well as value of CRM. ANOVA was used as it assists with identifying relationships between three variables. In this case, it was used to establish relationship between DQ, DQM and the value of CRM. The results are all depicted in tables and figures.

Analysis of variance (ANOVA) tests the hypothesis that the means of two or more populations are equal. ANOVA assesses the importance of one or more factors by comparing the response variable means at the different factor levels (DeCoster 2006:38). The null hypothesis states that all population means (factor-level means) are equal whilst the alternative hypothesis states that at least one is different. Once ANOVA indicates that the groups do not all have the same means, a T-test is used to compare them two by two.

To perform an ANOVA, one must have a continuous response variable and at least one categorical factor with two or more levels. ANOVAs require data from approximately normally distributed populations with equal variances between factor levels. However, ANOVA procedures work quite well even if the normality assumption has been violated unless one or more of the distributions are highly skewed or if the variances are quite different. Transformations of the original dataset may correct these violations.

Each F-statistic is a ratio of mean squares. The numerator is the mean square for the term. The denominator is chosen such that the expected value of the numerator mean square differs from the expected value of the denominator mean square only by the effect of interest (Quinn & Keogh 2001:32).

The effect for a random term is represented by the variance component of the term. The effect for a fixed term is represented by the sum of squares of the model components associated with that term divided by its degrees of freedom. Therefore, a high F-statistic indicates a significant effect (Quinn & Keogh 2001:32).

#### *4.2.4. Correlation*

Correlation is a statistical calculation which is used to describe the relationship between two variables consisting of interval or ratio data (Cohen & Brooke 2004:221). A correlation can indicate whether there is a relationship between two variables and whether the direction of the relationship is positive or negative. It also determines the strength of the relationship. It does not mean that one variable is the cause of the other, in other words, it does not explain causation (Jaffe 2010:3). It is important to establish the relationships between the variables used in this research.

### **4.3. Interpretation of results**

#### **4.3.1. Reliability of the scale used**

Alpha coefficient scores were calculated in order to assess the internal reliability of the measuring instrument. Cronbach Alpha coefficient scores with a value of more than 0.70, as recommended by Nunnally and Bernstein (1994:264-265), were recorded for each of the six factors. The measuring instrument can therefore be considered as reliable (see Table 4.1).

Three variables were constructed namely, data quality, data quality management and client relationship management. They were constructed by summing up all the responses of the relevant questions associated with each of the three variables. For

data quality, DQ1 - DQ11 questions were used. For data quality management, DQM1 - DQM8 questions were used. For client relationship management, CRM1 - CRM10 questions were used. Cronbach alpha was then calculated so as to determine the reliability of the variables constructed. The Cronbach alpha determines the reliability of the scale used as well as to test whether or not the items used for a construct or variable actually measure that particular construct.(DeCoster 2000:13) A value that is above 0.7 is acceptable as already indicated in chapter three in order for us to conclude that we have a reliable scale.

Table 3. Cronbach's alpha

<b>Variables</b>	<b>Cronbach's Alpha</b>	<b>Number of questions</b>
Data quality	0.834	11
Data quality management	0.891	8
Client relationship management	0.936	10

### Conclusion

From the constructed variables, the Cronbach's alpha are all greater than 0.7. Therefore, this indicates that the scales are reliable.

### 4.3.2 Biographical information

Frequency distribution graphs were used to interpret the data. Average scores for each demographic groups (gender; ethnicity; primary function; primary type of information; age and tenure were also constructed to determine their perceptions on each of the variables, namely, data quality, data quality management and client relationship management.

Out of the responses obtained, there was an equal spread of gender at 50% males and 50% females. The majority of respondents were from the sales staff at 34% and customer services at 28%. Management was represented by 17% of the

respondents whilst Finance was at 13% and Information Technology (IT) was at 8% representation.

#### 4.3.2.1. Gender

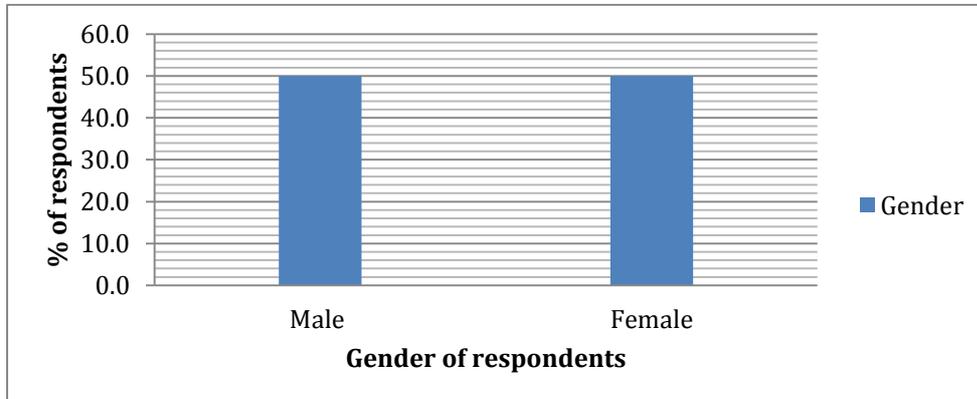


Figure 10: Percentage of respondents in different gender groups

The respondents were split equally between the different gender groups, with 50% male and 50% female respondents.

#### 4.3.2.2. Age

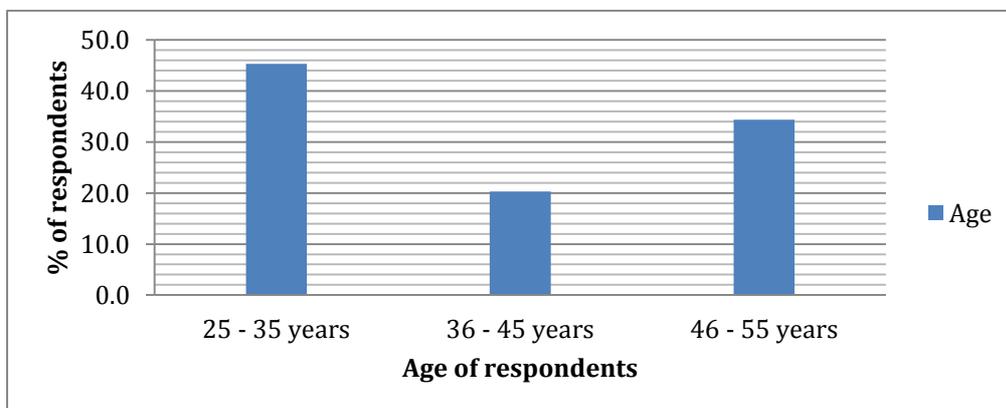


Figure 11: Percentage of respondents in different age groups

The largest group of the respondents, 45%, was between the ages of 25 and 35 years, followed by the respondents between the ages of 46 and 55 years (35%). The age group of 36 to 45 years constituted only 20% of the respondents. A T-test was also done to determine whether or not there were significant differences in

perceptions of data quality, data-quality management and client-relationship management amongst the different age groups.

Table 4: T-test results for different age groups

Variables		F statistic	P-value
Age	Data quality	0.711	0.495
	Data-quality management	3.239	0.046
	Client-relationship management	0.005	0.995

The p-value for data-quality management was 0.046, which is less than 0.1. This indicates that there were significant differences amongst the age groups. It means that the perceptions of data-quality management differed amongst the age groups. The p-values for data quality and client-relationship management were 0.495 and 0.995, which are greater than 0.1. This indicates that there were no significant differences amongst the age groups on these issues. It means that the perceptions of data quality and client-relationship management did not differ amongst the age groups.

Because the researcher found significant differences for data-quality management, a post-hoc analysis needed to be done to determine which of the age groups have significant differences. The test used was the Tukey test.

Table 5: Post-hoc analysis amongst different age groups

Data-quality management	Age		Mean Difference (I-J)	P-value
	25-35 years	36-45 years		
			6.355	0.043

The difference in the perceptions of data-quality management was between respondents aged 25 to 35 years and those aged 36 to 45 years. This means that these age groups perceived data-quality management differently.

### 4.3.2.3. Tenure of employment

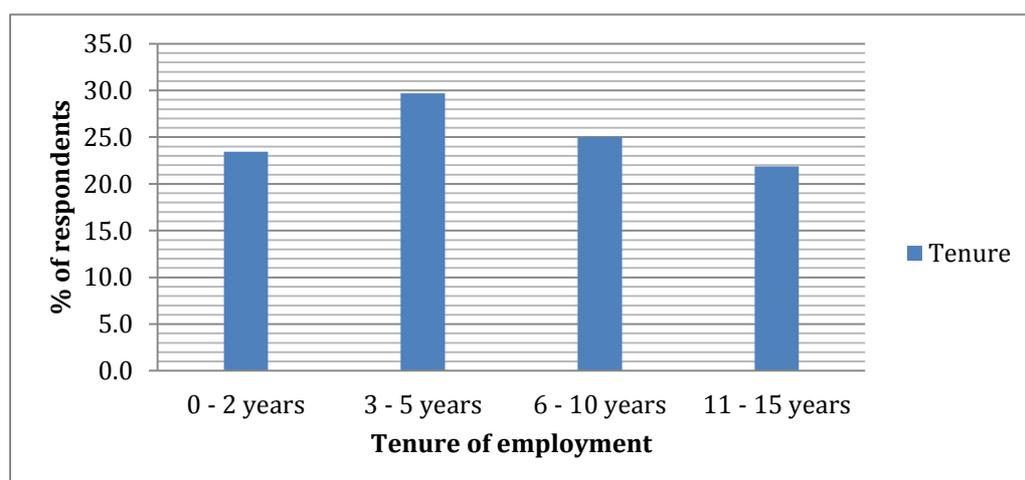


Figure 12: Percentage of respondents in tenure of employment

The largest number of respondents (30%) has been working for between 3 and 5 years, followed by the group with 6 to 10 years of tenure (25%). Furthermore, 22% of the respondents had tenure of 11 to 15 years at the bank, and 23% of respondents had tenure of less than 2 years.

T-test was also done to determine whether or not there were significant differences in perceptions of data quality, data-quality management and client-relationship management in terms of tenure of employment.

Table 6: T-test results concerning different tenure of employment

Variables		F statistic	P-value
Tenure of employment	Data quality	1.017	0.391
	Data-quality management	1.052	0.376
	Client-relationship management	0.885	0.454

The p-values for data quality, data-quality management and client-relationship management were 0.391, 0.376 and 0.454, which are greater than 0.1. This indicates that there were no significant differences in terms of the tenure of employment. Therefore, the perceptions of data quality, data-quality management and client-relationship management did not differ amongst the tenure groups.

#### 4.3.2.4. Ethnicity

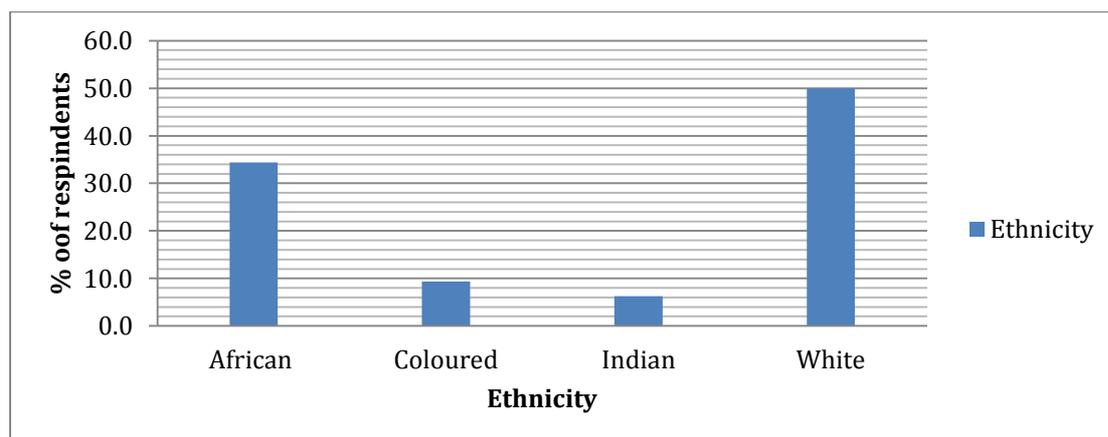


Figure 13: Percentage of respondents in different ethnic groups

The majority of the respondents were White people (50%), and the minority were Indian people (6%). African people comprised 35% and Coloured people 9% of the respondents. A T-test was also done to determine whether there were differences concerning the perceptions of data quality, data-quality management and client-relationship management amongst the ethnic groups.

Table 7: T-test results for different ethnic groups

Variables		F statistic	P-value
Ethnicity	Data quality	1.399	0.252
	Data-quality management	2.352	0.081
	Client-relationship management	0.448	0.720

The p-value for data-quality management was 0.081, which is less than 0.1. This indicates that there were significant differences amongst the ethnic groups. It means that the perceptions of data-quality management differed amongst the ethnic groups. The p-values for data quality and client-relationship management were 0.252 and 0.720, which are greater than 0.1. This indicates that there were no significant differences amongst the ethnic groups. It means that the perceptions of data quality and client-relationship management did not differ amongst the ethnic groups.

Because significant differences were found for data-quality management, a post-hoc analysis was done to determine which of the ethnic groups have significant differences. The Tukey test was used.

Table 8: Post-hoc analysis results for different ethnic groups

Data-quality management	Ethnicity		Mean Difference (I-J)	P-value
	White	Indian		
			10.250	0.082

The difference in perceptions of data-quality management occurred between the White and Indian ethnic groups. This indicates that the White and Indian ethnic groups had different perceptions of data-quality management.

#### 4.3.2.5. Primary job roles

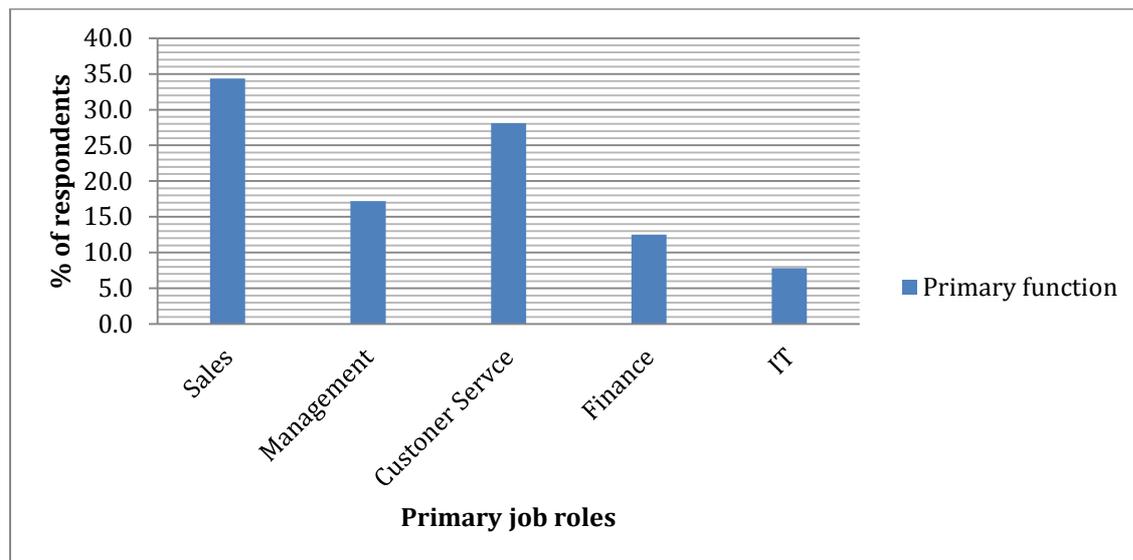


Figure 14: Percentage of respondents in different primary job roles

Most of the respondents were in the sales environment (34%), followed by staff in the customer service environment (28%). The remainder of the respondents came from management (17%), finance (13%) and IT (8%). A T-test was also done to determine whether or not there were significant differences in perceptions of data quality, data-quality management and client-relationship management amongst the respondents from different primary job roles.

Table 9: T-test of results in different primary job roles

Variables		F statistic	P-value
Primary job roles	Data quality	2.219	0.078
	Data-quality management	0.743	0.567
	Client-relationship management	0.302	0.875

The p-value for data quality was 0.078, which is less than 0.1. This indicates that there were significant differences amongst the primary function groups. It means that the perceptions of data quality differed amongst the respondents in different primary job roles. The p-values for data-quality management and client-relationship management were 0.567 and 0.875, which are greater than 0.1. This indicates that there were no significant differences amongst the primary function groups. It means that the perceptions of data-quality management and client-relationship management did not differ amongst the respondents from the primary job roles.

Because significant differences for data quality were found, a post-hoc analysis was done to determine which of the primary function groups had significant differences. The test used was the Tukey test.

Table 10: Post-hoc analysis of results amongst different ethnic groups

Data quality	Primary function		Mean Difference (I-J)	P-value
	Sales	Management	8.000	0.071

The difference in perceptions of data quality was between the primary job roles of sales and management. This gives an indication that people in sales and in management viewed data quality differently.

#### 4.3.2.6. Primary type of data used

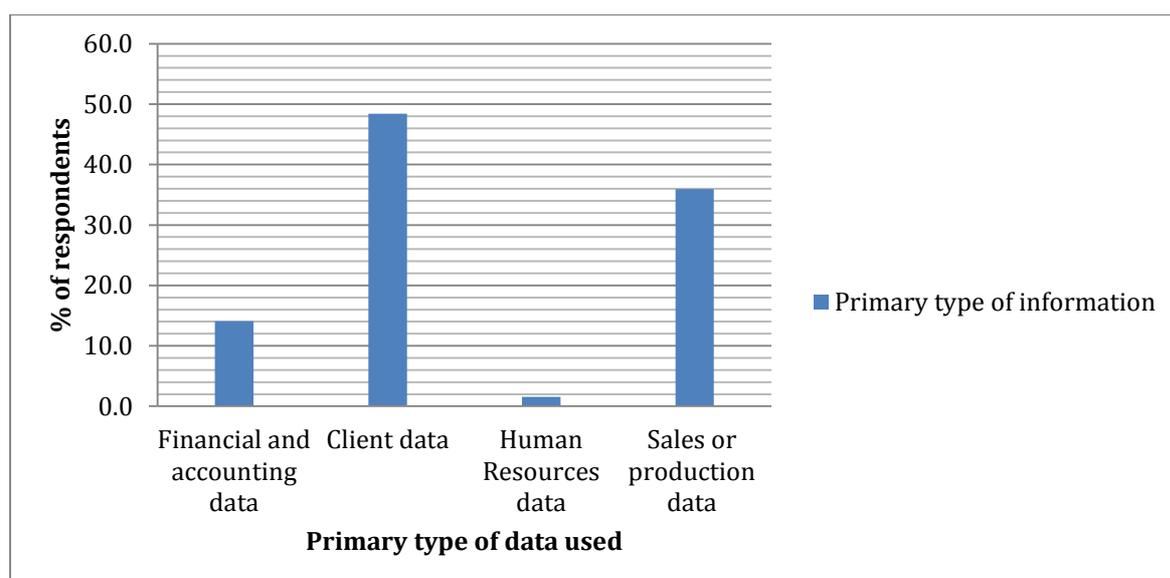


Figure 15: Percentage of respondents in different types of data used

The data type most used was client data (48%), followed by sales or production data (36%). The use of financial or accounting data constitutes only 14% and human-resources data only 2% of responses. A T-test was also done to determine whether or not there were significant differences in perceptions of data quality, data-quality management and client-relationship management, depending on the type of primary data used.

Table 11: T-test results for different types of data used

Variables		F statistic	P-value
Primary type of data used	Data quality	1.217	0.303
	Data-quality management	0.466	0.630
	Client-relationship management	0.133	0.876

The p-values for data quality, data-quality management and client-relationship management were 0.303, 0.630 and 0.876, which are all greater than 0.1.

This indicates that there were no significant differences amongst the groups of primary type of information used. This means that the perceptions of data quality,

data-quality management and client-relationship management did not differ amongst the primary type of information groups.

#### 4.3.2.7. Average scores

Average scores for each demographic group were also constructed to determine the perceptions on each of the three variables. These are calculated as a percentage, the closer the value is to 100 the higher the perception on each of the variables (data quality, data quality management and client relationship management).

Table 2: Average scores per demographic group

Variables	Groups	Data quality	Data quality management	Client relationship management	N
Gender	Male	62.500	66.406	75.536	32
	Female	60.268	62.779	75.893	32
Ethnicity	African	63.932	69.643	74.610	22
	Coloured	56.710	63.095	76.905	6
	Indian	51.299	51.339	68.214	4
	White	61.769	63.058	77.188	32
Primary function	Sales	67.532	68.425	75.649	22
	Management	57.143	60.065	76.364	11
	Customer service	59.452	64.087	74.048	18
	Finance	55.682	61.830	74.464	8
Primary type of information	IT	59.740	63.929	82.571	5
	Financial	56.277	62.698	77.143	9
	Client data	60.871	63.422	76.267	31
	HR & sales	63.961	66.815	74.464	24
Age	25 - 35 years	63.502	68.904	75.517	29
	36 - 45 years	59.441	57.555	76.044	13
	46 - 55 years	59.740	63.068	75.779	22
Tenure	0 - 2 years	60.779	60.595	70.381	15
	3 - 5 years	59.193	67.199	78.872	19
	6 - 10 years	66.153	67.746	75.893	16
	11 - 15 years	59.555	61.735	76.939	14

Male respondents have the highest score on DQ and DQM whereas the female respondents have the highest score on CRM. African respondents have the highest

scores on DQ and DQM compared to the Whites respondents who have the highest score on CRM. Respondents in the sales function and those who use the sales and human-resources data have the highest score on DQ and DQM whereas respondents in IT and those who use the financial data have the highest score in CRM. The respondents aged between 25 and 35 years have the highest score on data quality as well as data-quality management whereas the respondents aged between 36 and 45 years have the highest score on client-relationship management. The respondents who have six to ten years tenure also showed the highest score on DQ and DQM, and those with three to five years tenure show the highest score on CRM

#### **4.4. Perceptions on data quality**

Eleven questions were asked to the participants to determine their perception on data quality, and a seven point Likert scale was used as discussed in Chapter Three.

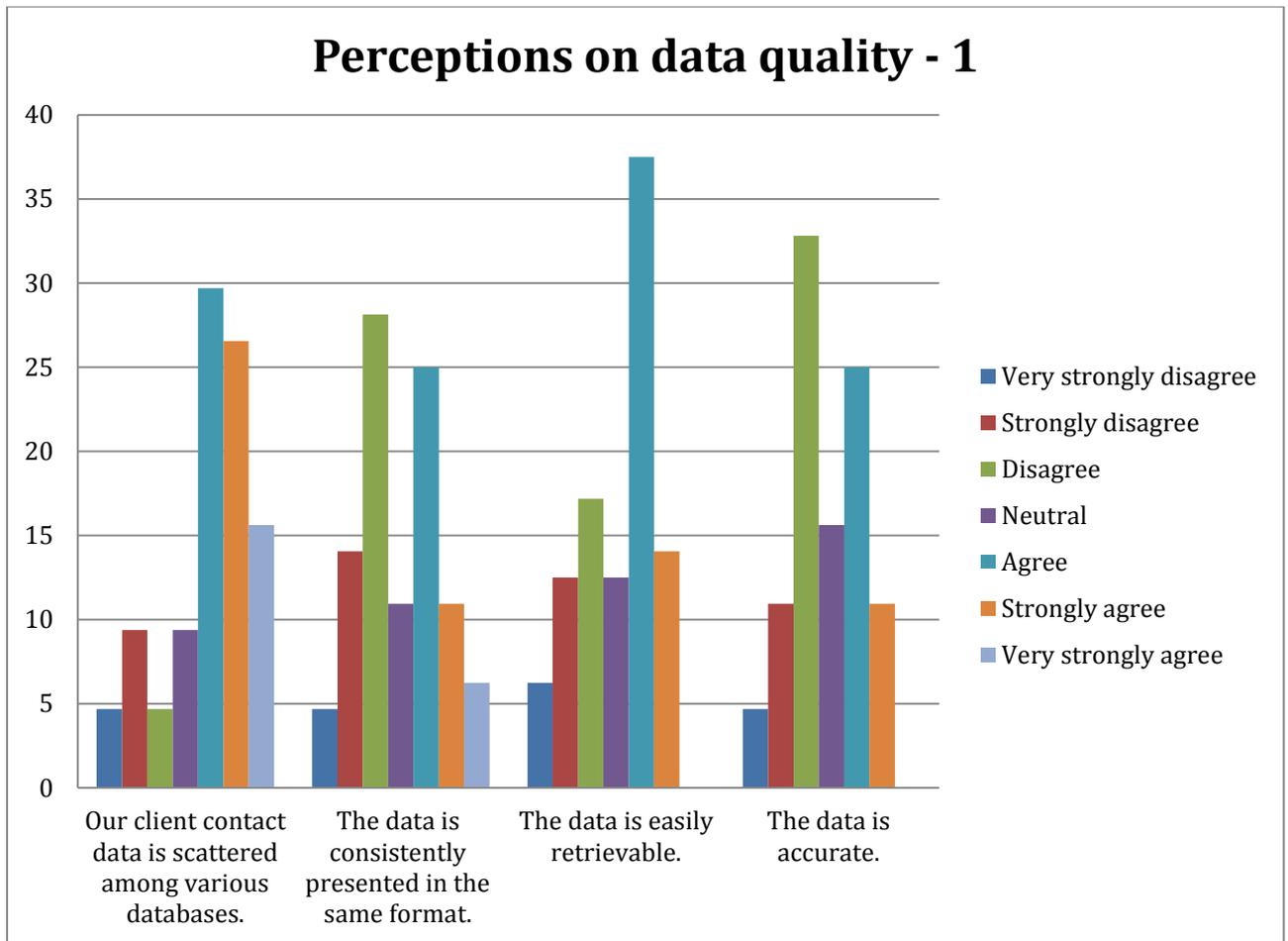


Figure 16: Perceptions of data quality – 1

In the research, 30% of the respondents agreed that the client data is scattered amongst various databases, and 28% of respondents disagreed with the statement that the data is consistently presented in the same format. Furthermore, 38% of respondents agreed that the data is easily retrievable. However, 33% of respondents disagreed with the statement that the data is accurate.

A total of 73% of the respondents agreed that the client data is scattered amongst various databases, compared to 18% who disagreed. The balance of 9% was neutral.

A higher percentage of the respondents (47%) disagreed with the statement that the data is consistently presented in the same format whilst 42% agreed, and the balance of 10% was neutral.

In total, 52% of respondents agreed that the data is easily retrievable, and 12% were neutral whilst 36% disagreed.

A higher percentage of 49% disagreed that the data is accurate, and 15% were neutral compared to 36% who agreed.

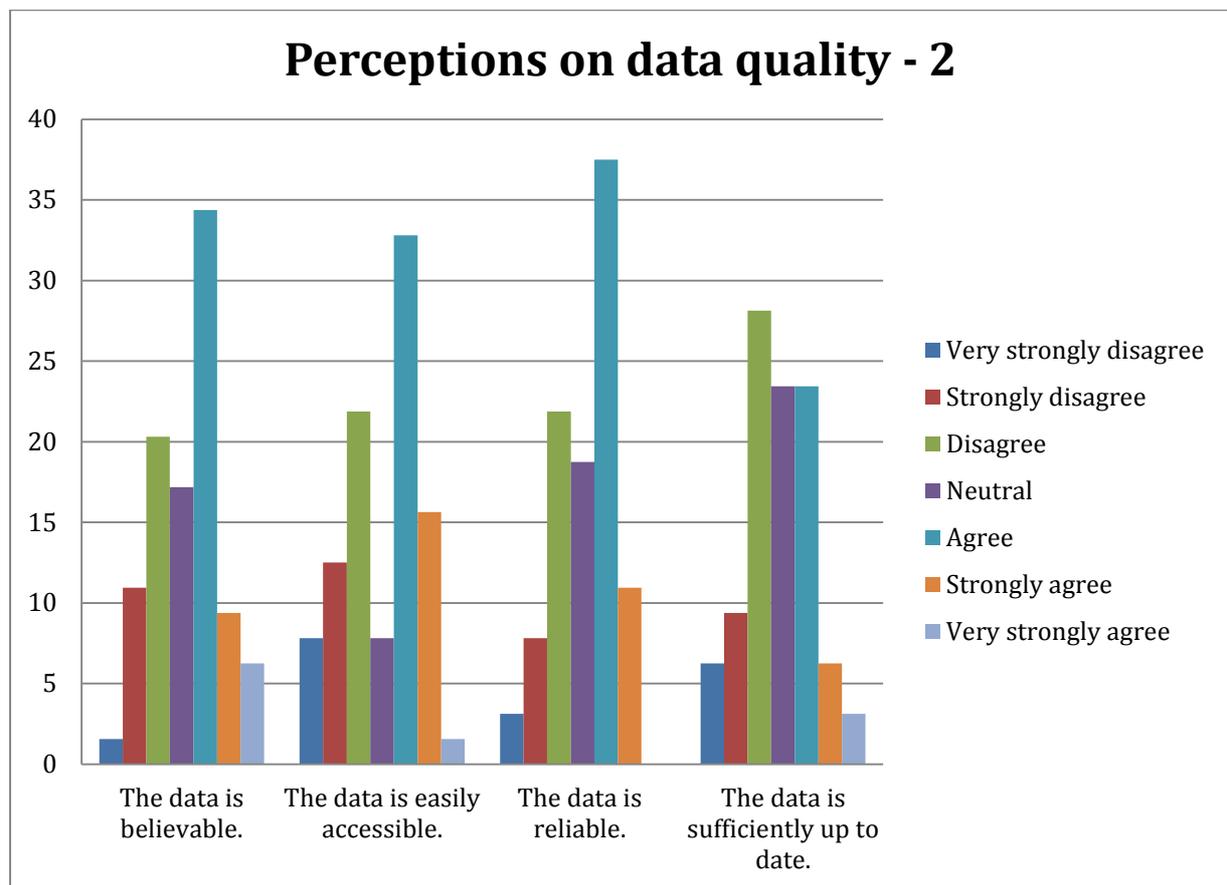


Figure 17: Perceptions of data quality – 2

Figure 17 shows that 34% of respondents agreed that the data is believable, and 33% also agreed that the data is easily accessible. Furthermore, 38% agreed that the data is reliable whilst 28% disagreed with the statement that the data is sufficiently up to date.

In their responses, 49% of respondents overall agreed that the data is believable whilst 33% disagreed and 16% were neutral. A higher total percentage of 51% agreed that the data is easily accessible, and 43% disagreed. The balance of 7% was neutral. A total 29% of all respondents indicated that they agreed that the data is

reliable with 18% indicating that they were neutral. Overall, 43% of respondents disagreed with the statement that the data is sufficiently up to date.



Figure 18: Perceptions of data quality – 3

In the research, 34% of the respondents strongly agreed that the data is protected with adequate security. A high percentage of 35% also agreed that the data is relevant to their work. Furthermore, 30% indicated that the data is sufficiently current.

A very high percentage of 70% of respondents indicated that they agreed that the data is protected with adequate security whilst 16% were neutral. In addition, 73% of respondents also indicated that the data is relevant to their work, and about 56% agreed that the data is sufficiently current.

## 4.5. Perception on data quality management

A total of 10 questions were included in the questionnaire to determine the perceptions of data-quality management in the bank

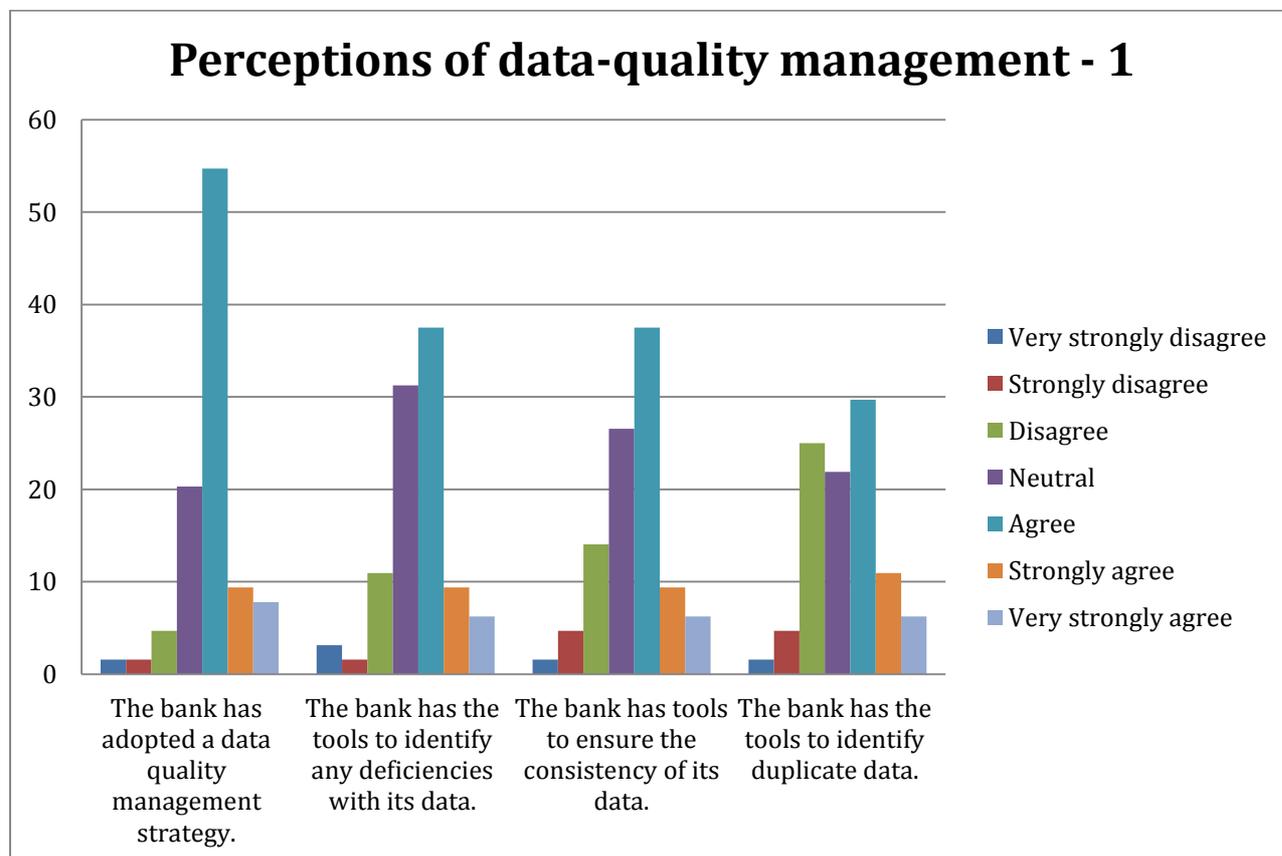


Figure 19: Perceptions of data-quality management – 1

A very high percentage of the respondents (55%) agreed that the bank has adopted a data-quality management strategy. In addition, 38% of respondents indicated that the bank has the tools to identify any deficiencies with its data, and 31% were neutral on this matter, which indicates that they were not sure whether the bank has the tools to identify any deficiency in the data. Furthermore, 38% of respondents also agreed that the bank has the tools to ensure the consistency of its data whilst 27% were neutral on this matter. Figure 19 above shows that 30% of respondents agreed that the bank has the tools to identify duplicate data whilst 25% disagreed, and about 22% were neutral.

The number of respondents who were neutral is high in these four questions compared to all the other questions asked to determine perceptions of data-quality management.

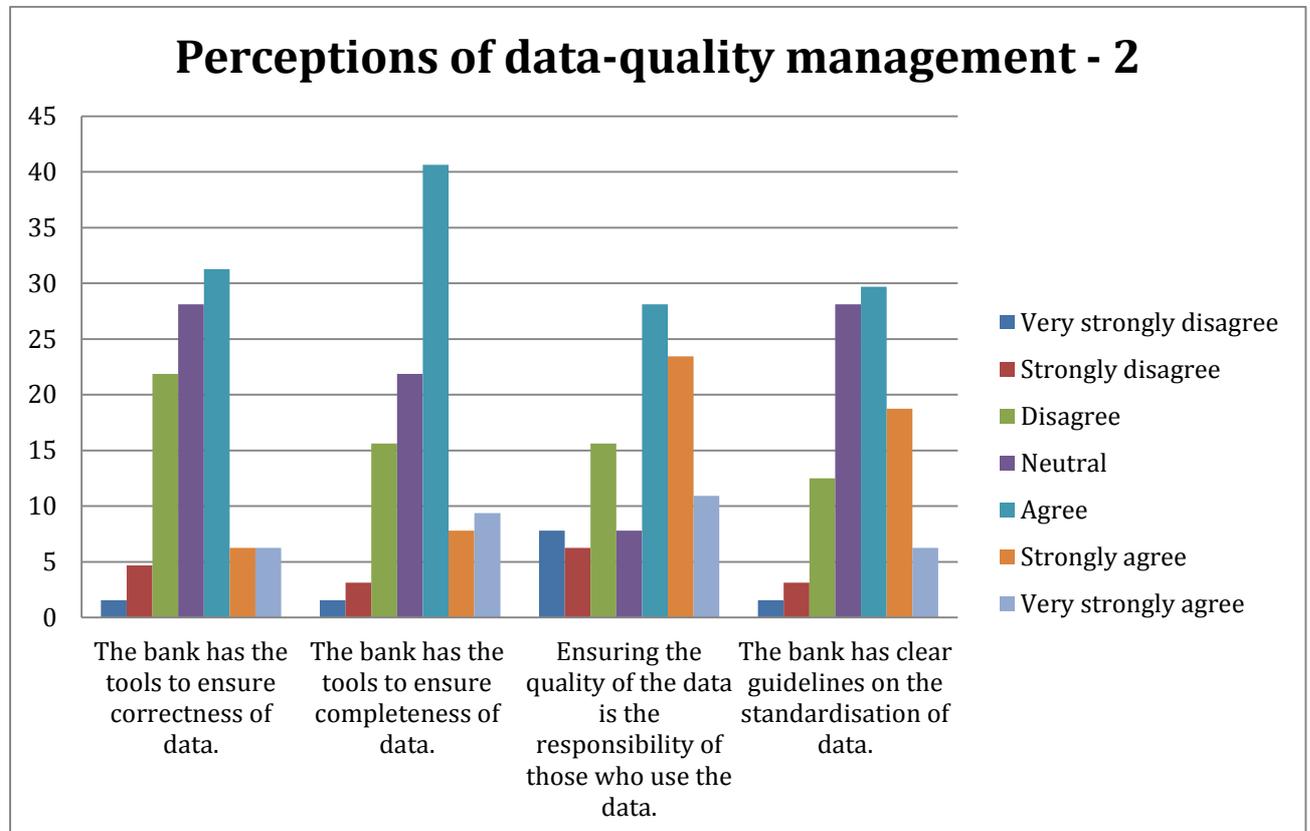


Figure 20: Perceptions of data-quality management – 2

The figure (20) shows that 31% of the respondents agreed that the bank has the tools to ensure correctness of the data, but 28% were neutral whilst 22% disagreed. A higher percentage of 41% of respondents agreed that the bank has the tools to ensure completeness of data whilst only 16% disagreed. In addition, 28% of respondents indicated that ensuring data quality is the responsibility of those who use the data, and in this category, 23% of respondents strongly agreed compared to 16% who disagreed. Concerning the next question 30% of respondents indicated that the bank has clear guidelines on the standardisation of data whilst 28% were neutral, and 19% strongly agreed.

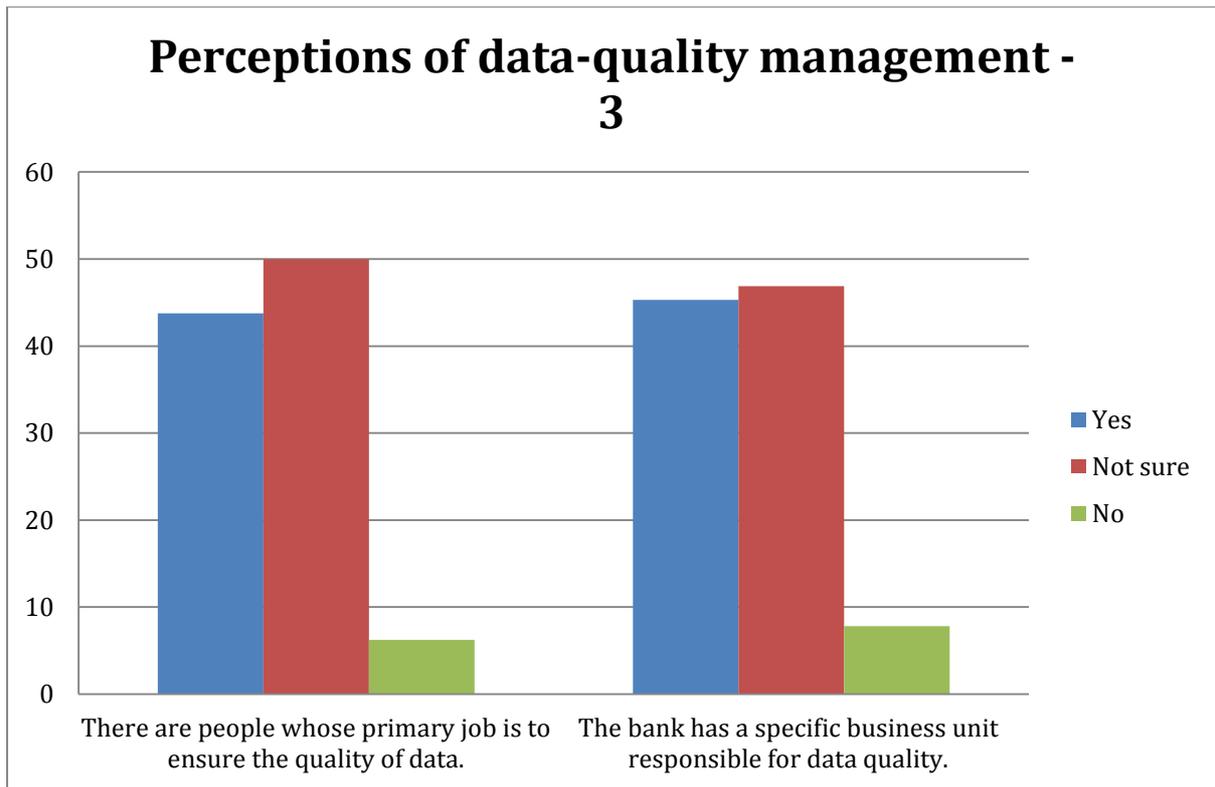


Figure 21: Perceptions of data-quality management – 3

Quite a high percentage of respondents (50%) were not sure whether there are people whose primary job it is to ensure the quality of data, and 44% indicated that there are people who ensure the quality of data. Concerning the questions whether the bank has a specific business unit that is responsible for data quality, 47% of respondents were not sure, and 45% indicated that the bank has a specific business unit that is responsible for data quality. A high number of respondents, 50% and 47% respectively, were neutral on these questions, which indicate that they were not sure whether there are people whose primary job it is to ensure the quality of data and whether the bank has a specific business unit that is responsible for data quality.

#### 4.6. Perceptions on Client Relationship Management

A total of 10 questions were included in the questionnaire to determine the perceptions of the respondents on client-relationship management (CRM)

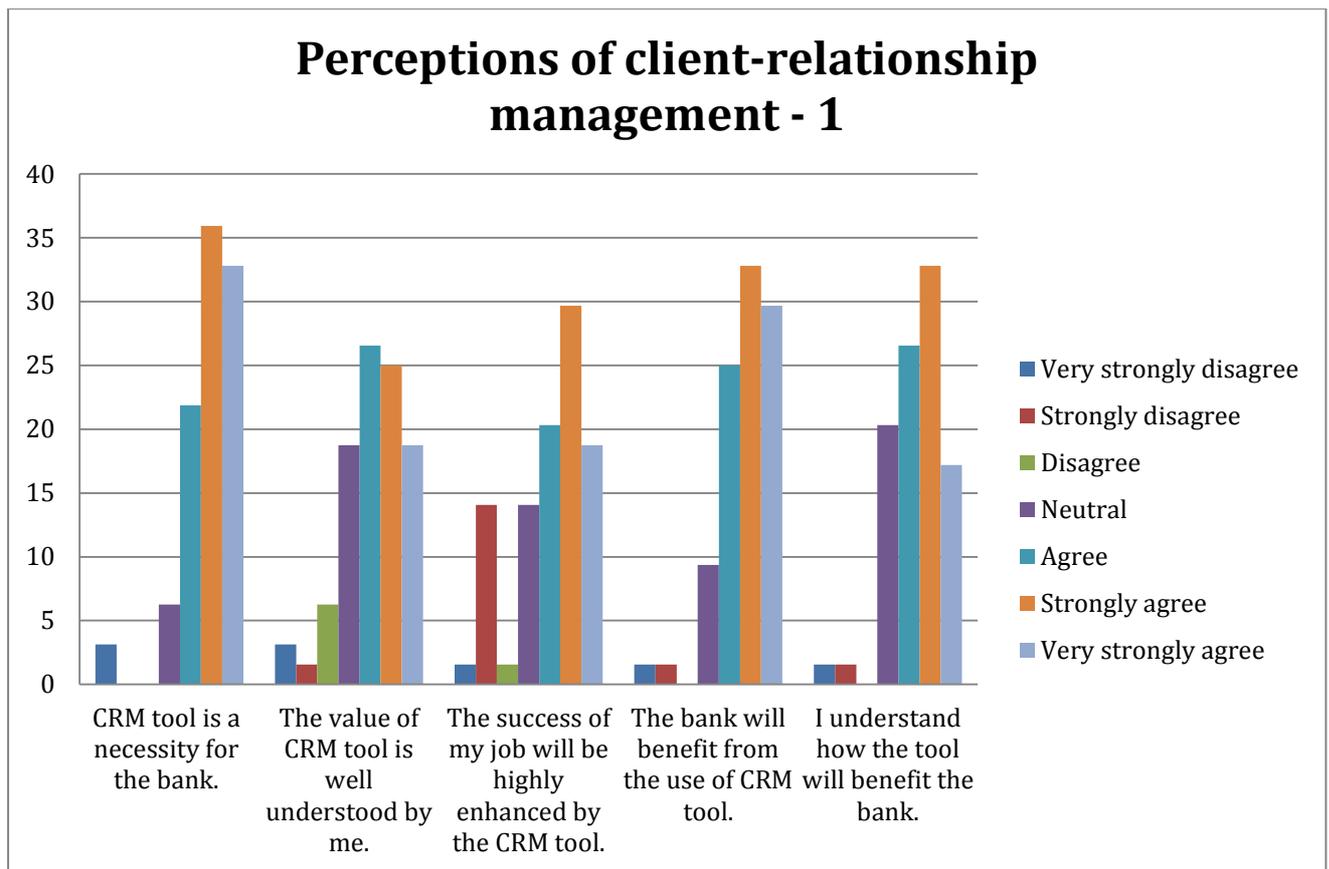


Figure 22: Perceptions of client-relationship management – 1

In the study, 36% of the respondents strongly agreed that the CRM tool is a necessity for the bank with 33% of the respondents who very strongly agreed. Furthermore, 27% of respondents indicated that they understood the value of the CRM tool followed by 25% who strongly agreed and 19% who very strongly agreed. On the next question, 33% of respondents strongly agreed that the bank will benefit from the use of the CRM tool and that they understand how the tool will benefit the bank. There is a trend of high percentages overall from respondents who agreed with the given statements:

- 69% agreed that the CRM tool is a necessity for the bank.
- 71% agreed that they understand the value of the CRM tool very well.
- 69% indicated that the success of their job will be highly enhanced by the CRM tool.
- 88% agreed that the bank will benefit from the use of CRM tool.
- 77% agreed that they understand how the tool will benefit the bank.

## Perceptions of client-relationship management - 2

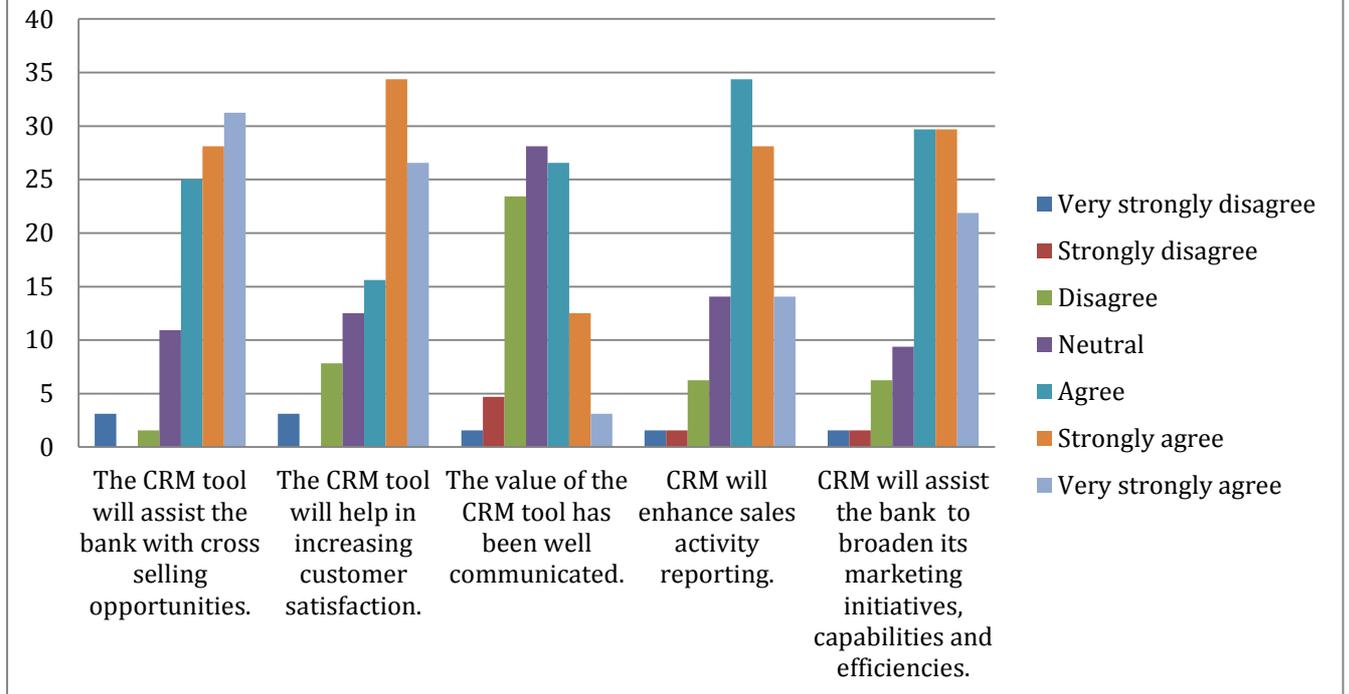


Figure 23: Perceptions of client-relationship management – 2

A high percentage of 31% of respondents very strongly agreed that the CRM tool will assist the bank with cross-selling opportunities whilst 28% strongly agreed. Also, 34% of respondents strongly agreed that the CRM tool will help in increasing customer satisfaction followed by 27% who responded with 'very strongly agree'. The concern is the 27% of respondents who were neutral as they were not sure whether the value of the CRM tool has been well communicated and the 23% who disagreed with the statement. Only 27% agreed that the value of the CRM tool has been well communicated. A higher percentage (34%) agreed that the CRM will enhance sales-activity reporting whilst 28% strongly agreed. In addition, 30% of the respondents indicated that they agree and strongly agree that the CRM will assist the bank to broaden its marketing initiatives, capabilities and efficiencies.

Overall, higher percentages of above 75% of respondents indicated that they agree with most of the statements. Only 43% of respondents indicated that the value of the CRM has been well communicated whilst 27% were neutral and the balance of 30% disagreed.

#### 4.7. Average scores

Average scores for the overall sample were also constructed to determine the perception of the respondents on each of the variables, namely, data quality, data quality management and client relationship management. These are calculated as a percentage, the closer the value to 100% the higher the perception on each of the variables.

Table 3: Average scores for the overall sample

Variables	Groups	Data quality	Data quality management	Client relationship management	N
Overall		61.384	64.593	75.714	64

From Table 12, the following can be deduced:

- Of all respondents, 61% expressed their satisfaction regarding the data that they were collecting and consuming or had in their custody.
- An even higher percentage of 65% of respondents were satisfied that the bank had adopted some form of data-quality management strategy.
- A total of 75% or respondents indicated that they understand the value of CRM in the bank.

These results suggest that data quality has not received enough attention by the bank, and this is supported by the very low scores of 28% to 38% of respondents who believed that the data was believable, accessible and reliable.

#### 4.8. Correlations

Testing was also done to determine whether or not there is a correlation between data quality and data-quality management between data quality and client-relationship management as well as between data-quality management and client-relationship management.

The researcher looked at the p-value and compared it with a significance level of 0.1. If the p-value was less than 0.1, I concluded that a relationship exists between the variables. If a positive relationship was found between the two variables, the correlation coefficient would be positive, and it would range between 0 and 1. The closer the value is to 1, the stronger the relationship between the 2 values. In contrast, if a negative relationship was found, the correlation coefficient would be negative, and it would range between -1 and 0. The closer the value was to -1 the stronger the relationship between the 2 values.

Table 4: Correlations

<b>Variables</b>		<b>Coefficient</b>	<b>P-value</b>
<b>Data quality</b>	Data-quality management	0.682	0.000
	Client-relationship management	0.162	0.200
<b>Data-quality management</b>	Client-relationship management	0.259	0.039

The correlation coefficient between data quality and data-quality management was 0.682. The p-value was 0.000, which is less than 0.1. This means that there was a significant positive relationship between data quality and data-quality management. This means that, as data quality improves, data-quality management also improves.

The correlation coefficient between data quality and client relationship management was 0.162. The p-value was 0.200, which is greater than 0.1. This means that there was an insignificant positive relationship between data quality and client-relationship management. The correlation coefficient between data-quality management and client-relationship management was 0.259. The p-value was 0.039, which is less than 0.1. This means that there was a significant positive relationship between data-quality management and client-relationship management. This means that, as data-quality management improves, client-relationship management also improves.

## **4.9. Conclusion**

The analysis and interpretation of the results were discussed in this chapter, using statistical tests like ANOVA and T-tests to test the perceptions of the respondents concerning data quality, data-quality management and the value of client-relationship management. The results from correlations tests between the three variables were also discussed. The results were provided in six different categories, namely the internal reliability of the measuring instrument, the biographical data of the respondents, perceptions of data quality, perceptions of data-quality management, perceptions of the value of client-relationship management and the correlations between the three variables. Average scores based on the demographical information and overall responses concerning the perceptions of these variables were also discussed in this chapter

## **CHAPTER 5**

### **FINDINGS AND RECOMMENDATIONS**

---

#### **5.1 Introduction**

This chapter will first present the findings of this research in relationship to the research objectives. Recommendations will then be made, based on the findings of the research. The recommendations will assist the bank to make decisions regarding data quality and its impact on the successful implementation of CRM in the bank. Limitations of the study will also be included in this chapter.

#### **5.2. Summary**

The primary objective of this research was to determine and investigate the impact of data quality on the successful implementation of the customer-relationship management at a major South African bank.

The primary objective was supported by the following secondary objectives:

- determine guidelines for creating good-quality data
- identify a data-improvement strategy, using a data-management framework.
- determine the relationship between data quality, data-quality management and client-relationship management.

Statistical tests like ANOVA and t-tests were conducted on all the variables, namely DQ, DQM and CRM. The relationships between DQ, DQM and CRM as they relate to the specific biographical information, namely gender, tenure of employment, job profiles and type of data used, were also evaluated.

#### **5.3 Findings and recommendations**

##### *5.3.1. Findings and recommendations from biographical information*

- a) Different age groups: The difference in perceptions concerning data-quality management was between respondents aged 25 to 35 years and those aged

36 to 45 years. This means that these age groups perceive data-quality management differently. There were no differences in those aged 46 to 55 years.

- b) Ethnicity: The difference in perceptions of data-quality management was between White people and Indian people. This indicates that White and Indian people have different perceptions concerning data-quality management
- c) Primary job role: The difference in perceptions of data quality was between the primary job roles of sales and management. This gives an indication that people in sales and in management viewed the quality of data differently.

### **Recommendations**

- a) The findings indicate that the differences are mainly in DQM with respect to the different age groups and ethnicity. Focused training on DQM should be targeted at the identified age and ethnic groups to ensure that there is consistency and a common understanding of DQM.
- b) Respondents from sales and management also showed different perceptions regarding DQ. It is important that has the same understanding of DQ as they have to work towards a common goal. Training is also recommended for these two groups as it important that the bank should adopt the dimensions of data quality, namely completeness, consistency, validity, accuracy, timeliness, accessibility and relevance. This will also make it easier for employees to know how to manage this data going forward.
- c) DQ and DQM must be added on the sales and management balance scorecard so as to ensure that it is their responsibility to ensure that the dimensions of data quality and the DQM framework are adhered to.
- d) Management should also include DQ in their quarterly risk report as this will encourage them to take DQ seriously if they see that senior executives also monitor DQ at business-unit level.

#### *5.3.2. Findings and recommendations on perceptions of data quality*

In the study, 47% of the respondents disagreed with the statement that the data is consistently presented in the same format, and 49% also disagreed with the statement that the data is accurate. However, 49% of respondents overall agreed that the data is believable, and a higher total percentage of 51% agreed that the data

is easily accessible. In addition, 43% of respondents overall disagreed that the data is sufficiently up to date.

A very high percentage of 70% responded that they agreed that the data is protected with adequate security, and a total of 73% overall indicated that the data is relevant to their work.

The findings indicate that the major data-quality issues entail the fact that data is not consistently presented in the same form, and it is not accurate. However, the respondents indicated that the data is easily accessible even though it is not sufficiently up to date. The respondents also indicated that there is adequate security and that the data is protected. An average of about 36% of the respondents expressed satisfaction with the data they were collecting, consuming or had in their custody. The low percentages suggest that the issue of data quality has not been given enough attention by the bank

### **Recommendations**

Data-quality initiatives involve activities that aim at improving the quality of the bank's data by setting guidelines and criteria for data input.

- a) The bank needs to use the data-quality dimensions as guidelines and criteria for data input. This will ensure that data quality is controlled at the input stage and that the bank does not wait for data users to report inaccuracy.
- b) The bank should implement a mechanism where employees would be able to give feedback on any data-quality deficiencies.
- c) The bank has to train all the employees on the principles of data quality and give guidelines on the criteria for data quality.

#### *5.3.3. Findings on perceptions of DQM*

A relatively low percentage of 33% of respondents said that their organisation had adopted some form of data-quality management, which indicates that the respondents do not think that the bank has a data-quality management strategy or framework. The number of respondents who were neutral was the highest at 28% in the three questions listed below compared to all the other questions asked to determine their perceptions of data-quality management.

- The bank has tools to identify deficiency in the data.
- The bank has the tools to ensure consistency in the data.
- The bank has tools to identify duplicate data.

These responses indicate that the employees were not sure if these tools exist or not. A high percentage, 50%, were not sure whether there are people whose primary job it is to ensure the quality of data, and 47% of respondents were not sure whether the bank has a specific business unit that is responsible for data quality.

### **Recommendations**

- a) The bank has to adopt a DQM strategy and ensure that it is properly communicated to all its employees. The bank needs to develop a data-quality management framework that must be used in the bank. The foundation of this framework could be a checklist that ensures that the data quality is maintained and checked at all times. The bank needs to adopt a standard framework to be used across the bank to ensure that consistency in data-quality management. Handling data-quality management from a strategy perspective will ensure that it becomes one of the primary focus areas of the bank.
- b) The bank needs to ensure that there is a business unit that is primarily responsible for data quality. This group of people should develop policies about acceptable data-quality standards and data-quality management. It is important that there are pre-defined data criteria and that these are well communicated to the business. This business unit must be fully responsible for collecting all the feedback and updating or acting upon low-quality data. Sufficient resources must be provided to ensure that this unit functions effectively.
- c) Training needs to be provided to all employees of the bank to encourage reporting of any unsatisfactory quality of data.
- d) Operational and management report within various divisions should include a data-quality score-card and improvement reports on key data elements that are strategic to operational results.

- e) The bank must introduce information systems that are implemented in the bank to support the data-quality management framework. Systems and processes will need to be reviewed to drive and ensure that the quality of data is enforced from the input stage and maintained within the entire bank. The information systems should include compulsory fields that are required for primary client information where the user cannot continue if the information is missing. There must also be an automatic-feedback functionality that will immediately alert the user of any deficiency in any predefined quality criteria.

#### *5.3.4. Findings on perceptions of the value of CRM*

In the study, 69% of respondents agreed that the CRM tool is a necessity for the bank, and 71% agreed that they understand the value of the CRM tool very well. Furthermore, 69% of respondents indicated that the success of their job will be highly enhanced by the CRM tool, and 88% agreed that the bank will benefit from the use of CRM tool. A total of 77% of respondents agreed that they understand how the tool will benefit the bank. These findings indicate that the respondents understand the value of the CRM in the bank.

### **Recommendations**

The respondents seem to have a good understanding on the value of CRM, and therefore, the bank has to ensure that the employees utilise the CRM to its full potential.

#### *5.3.5. Findings and recommendation on the correlation between all these variables*

Testing was also done to determine whether or not there is a relationship between data quality and data-quality management, between data quality and client-relationship management as well as between data-quality management and client-relationship management. The findings indicate the following:

- A significant positive relationship was found between data quality and data-quality management. This means that, as data quality improves, data-quality management also improves.

- An insignificant positive relationship was found between data quality and client-relationship management. This means that as data quality improves, it will have an insignificant effect on CRM.
- A significant positive relationship was found between data-quality management and client-relationship management. This means that the existence of DQM will have a positive effect on CRM.

### **Recommendations**

The findings indicate that both DQ and CRM have a significant relationship with DQM. DQM is therefore a critical element in ensuring that the DQ is maintained and that CRM implementation is successful. The bank has to put much focus on determining and implementing a DQM strategy.

The findings indicate that data is the most significant element that will have an impact on the successful implementation of DQM. However, the bank cannot maintain data that is of poor quality. The bank needs to put much focus and resources into ensuring that their data is of good quality and then formulate proper data-quality management strategies so that the CRM system is implemented successfully in the bank.

### **5.4. Limitations of the study**

The limitation of the study is that it was conducted in one major South African bank with the focus on the wealth management division in the Gauteng province. A follow-up study could be done to include all the major South African banks.

### **5.5. Future research**

A follow-up study on this topic could be extended to include the other major South African banks. It can also be extended to other financial services like insurance services, and the findings can then be compared to those of this study.

## **5.6. Conclusion**

Data quality should be everybody's problem, from the data users to the people that input the data. Data quality and its consistency remain crucial to the operational effectiveness of a bank, and any bank needs to include quality issues in their risk reports if these are found to be deficient. Addressing the data-quality issue then becomes a strategic input in an organisation-information management and reporting strategy.

For this to be successful, it will need the support of all executive members, from the chief executive officer right down to the data consumer. Everybody in the bank must be made aware of the importance of data quality and the value that CRM will bring to the bank. The effectiveness and successful implementation of the data are dependent on data quality and data-quality management strategies.

## **BIBLIOGRAPHY**

ACAPS, n.d., *Qualitative and quantitative research techniques for humanitarian needs assessment*, May 2012, viewed 18 October 2014, from <http://www.acaps.org/img/documents/q-qualitative-and-quantitative-research.pdf>

ArcGIS Resource Centre, n.d., *What is a z-score? What is a p-value?*, viewed 6 November 2014, from <http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//005p00000006000000>

Bobbie, E.R., 2010, *The practice of social research*, 12th ed., Wadsworth Cengage, Belmont, CA.

Banking Association South Africa, 2012, *South African banking sector overview 2012*, viewed February 2014, from <http://www.banking.org.za/index.php/our-industry/2012-south-african-banking-sector-overview/>

Basic Research Concepts, n.d., *Research designs*, viewed 22 July 2014, from [http://ori.hhs.gov/education/products/sdsu/res\\_des1.htm](http://ori.hhs.gov/education/products/sdsu/res_des1.htm)

Batini, C. & Scannapieco, M., 2006, *Data quality: Concepts, methodology and techniques*, Springer, New York.

BCPS, n.d., *Planning the methodology: The quantitative pathway*, viewed 15 September 2014, from [http://www.bcps.org/offices/lis/researchcourse/develop\\_writing\\_method\\_quantitative.html](http://www.bcps.org/offices/lis/researchcourse/develop_writing_method_quantitative.html)

Benefits of CRM, n.d., *Benefits of CRM systems*, viewed 10 April 2014, from <http://www.benefitsofcrm.com/>

Bertram, D., 2011, 'Likert scales are the meaning of life', viewed 5 November 2014, from <http://poincare.matf.bg.ac.rs/~kristina/topic-dane-likert.pdf>

Bitterer, A., 2007, “*Dirty data*” is a business problem, not an IT problem’, viewed 28 April 2013, from <http://www.gartner.com/newsroom/id/501733>

Bryman, A. Bell, E. *Business Research Methods*. 2007. New York. Oxford University Press

CCS, n.d., *Qualitative research methods: A data collector’s field guide: Family health International*, viewed 5 May 2014, from <http://www.ccs.neu.edu/course/is4800sp12/resources/qualmethods.pdf>

Chambliss, D.F. & Schutt, R.K., 2012, *Making sense of the social world: Methods of investigation*, Sage, New York.

Chaturvedi, K., 2007, ‘Sampling methods’, viewed 3 October 2014, from [http://www.google.co.za/url?url=http://www.pitt.edu/~super7/43011-44001/43911.ppt&rct=j&frm=1&q=&esrc=s&sa=U&ei=g2FHVP7VPM7T7Abx8YHIBg&ved=0CBMQFjAA&usg=AFQjCNEIAYK-6bWg\\_Q4h3UQBWE5MLqmsEw](http://www.google.co.za/url?url=http://www.pitt.edu/~super7/43011-44001/43911.ppt&rct=j&frm=1&q=&esrc=s&sa=U&ei=g2FHVP7VPM7T7Abx8YHIBg&ved=0CBMQFjAA&usg=AFQjCNEIAYK-6bWg_Q4h3UQBWE5MLqmsEw)

Cohen, B.H. & Brooke, L.R., 2004, *Essentials of statistics for the social and behavioural sciences*, John Wiley & Sons.

Constantinos, N.P. Alice, B. & Seal, C. *Structured methods: Interviews, questionnaires and observation*, viewed 01 February 2015 from [http://www.sagepub.com/upm-data/47370\\_Seale\\_Chapter\\_11.pdf](http://www.sagepub.com/upm-data/47370_Seale_Chapter_11.pdf)

Cooper, D. Schindler, P. *Business Research Methods*. 2006. McGraw-Hill Irwin.

Cooper, D. & Schindler, P. *Business research methods*, 2011. McGraw-Hill International, New York.

Country profile. 2007, *South Africa*. viewed 4 February 2015 from [http://edu.care.org/Documents/Country%20Profile\\_South%20Africa-ECR.pdf](http://edu.care.org/Documents/Country%20Profile_South%20Africa-ECR.pdf)

Creswell, J., 2012, *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*, 4th end., Pearson Education, Upper Saddle River, NJ.

Curtis, G. & Cobham, D. (2002) *Business information systems: analysis, design and practice*. 4th ed. Harlow: Prentice Hall

Daniels, R., 2012, 'A framework for investigating micro data quality, with application to South African labour market household surveys', viewed 8 May 2013, from [http://www.opensaldru.uct.ac.za/bitstream/handle/11090/181/2012\\_90.pdf?sequence=1](http://www.opensaldru.uct.ac.za/bitstream/handle/11090/181/2012_90.pdf?sequence=1)

Database trends and applications (DBTA), n.d., *Data quality*, viewed 13 April 2014, from <http://www.dbta.com/Categories/Data-Quality-331.aspx>

DeCoster, J. *Scale construction notes*. 2000. Viewed 6 November 2014 from <http://www.stat-help.com/scale.pdf>

DeCoster, J. 'Testing group differences using T-tests, ANOVA, and nonparametric measures'. 2006. viewed 6 November 2014, from <http://www.stat-help.com/ANOVA%202006-01-11.pdf>

Financial Services Board, n.d., *Treating customers fairly*, viewed 9 February 2014, from <https://www.fsb.co.za/feedback/Pages/tcfhome.aspx>

Fischer, A.S., n.d., 'Mixed methods', viewed 22 July 2014, from [http://www.fischlerschool.nova.edu/Resources/uploads/app/35/files/ARC\\_Doc/mixed\\_methods.pdf](http://www.fischlerschool.nova.edu/Resources/uploads/app/35/files/ARC_Doc/mixed_methods.pdf)

Friedman, T. *A Strategic Approach to Improving Data Quality*. 2002. Viewed 13 March 2013 from <https://www.gartner.com/doc/361077/strategic-approach-improving-data-quality>

Gendall, P., 2011, 'Questionnaire design', viewed 5 November 2014, from <http://owll.massey.ac.nz/pdf/questionnaire-design.pdf>

Global Research Report. *Making your data work for you*.2014. Experian data quality. Viewed 18 March 2013 from [http://cdn.qas.com/media/marketing/downloads/pdf/whitepapers/uk/v6\\_WP\\_Landscape-12pp-Ext-Jan13.pdf](http://cdn.qas.com/media/marketing/downloads/pdf/whitepapers/uk/v6_WP_Landscape-12pp-Ext-Jan13.pdf)

GS1 data quality framework: Version 3.0, 2010.

Hall, R., 1998a, 'Psychology world: One-way ANOVA', viewed 3 October 2014, from <http://web.mst.edu/~psyworld/anovadescribe.htm>

Hall, R., 1998b, 'Psychology world: Tukey's HSD, post hoc tests', viewed 3 October 2014, from <http://web.mst.edu/~psyworld/anovadescribe.htm>

Hofmann, U., 2005, 'Quantitative methods in education research', viewed 15 September 2014, from <http://www.edu.plymouth.ac.uk/resined/quantitative/quanthme.htm>

Institute for Digital Research and Education, n.d., *Statistical analyses using Stata*, viewed 6 November 2014, from <http://www.ats.ucla.edu/stat/stata/whatstat/whatstat.htm>

I school, 1997, *The case study as a research method*, viewed 11 May 2014, from <https://www.ischool.utexas.edu/~ssoy/usesusers/l391d1b.htm>

Jaffe, A., 2010, 'Correlation, causation, and association: What does it all mean?', viewed 28 August 2014, from <http://www.psychologytoday.com/blog/all-about-addiction/201003/correlation-causation-and-association-what-does-it-all-mean>

Jankowicz, A.D., 2005, *Business research projects*, 4th ed., Thomson Learning, London.

Key, J.P. 1997, *Research design in Occupational Education*, Oklahoma State University. Viewed 4 February 2014, from <http://www.okstate.edu/ag/agedcm4h/academic/aged5980a/5980/newpage16.htm>

Lehohla, P., 2010, 'South African statistical quality assessment framework', viewed 2 May 2013, from [http://www.statssa.gov.za/inside\\_statssa/standardisation/SASQAF\\_OpsGuidelines\\_Edition\\_1.pdf](http://www.statssa.gov.za/inside_statssa/standardisation/SASQAF_OpsGuidelines_Edition_1.pdf)

Levesey, K. *Bad data costing US businesses \$700 billion a year*. 2011. Viewed 13 March 2013 from <http://about.datamonitor.com/media/archives/4871>

Loshin, D., 2011, *Enterprise knowledge management: The data quality approach*, Academic Press, London.

MacGilvray, D., 2008, *Executing data quality projects*, Morgan Kaufmann Publishers, Burlington.

McLeod, S. *Qualitative, Quantitative*. 2008. Viewed 11 May 2013 from <http://www.simplypsychology.org/qualitative-quantitative.html>.

Microsoft Developer Network, n.d., *Resurrecting the prodigal son: Data quality*, viewed 27 April 2013, from <http://msdn.microsoft.com/en-us/library/gg261774.aspx>

Minitab, n.d., *What is a general linear model?*, viewed 3 October 2014, from <http://support.minitab.com/en-us/minitab/17/topic-library/modeling-statistics/anova/basics/what-is-a-general-linear-model/>

Mosley, M. & Brackett, M., 2010, *DAMA: Guide to the data management body of knowledge*, Technics Publications, Bradley Beach.

Nedbank, 2011, *Nedbank enterprise data management master policy*.

Nunnally JC & Bernstein IH. 1994. *Psychometric Theory*. 3rd ed. New York: McGraw-Hill

Olson, J.E., 2003, *Data quality: The accuracy dimension*, Morgan Kaufmann publishers, Burlington.

Patton, M.Q. & Cochran M., n.d., 'A guide to using qualitative research methodology', viewed 5 May 2014, from <http://fieldresearch.msf.org/msf/bitstream/10144/84230/1/Qualitative%20research%20methodology.pdf>

Peppers and Rogers Group. *Maximizing Data quality to build profitable customer relationships*.2003. Viewed 18 March 2013 from <http://www.pbinsight.com/files/resource-library/resource-files/maximizing-data-quality.pdf>

Peterson,T. *Data Scrubbing*. 2003. Viewed 16 October 2014 from <http://www.computerworld.com/article/2580604/business-intelligence/data-scrubbing.html>

Pettier, J. & Sashay, D., 2011, 'Building the foundation for customer data quality in CRM systems for financial services firms', *viewed 27 April 2013*, from [www.palgrave-journals.com/dbm/](http://www.palgrave-journals.com/dbm/)

PriceWaterHouseCoopers. *PwC's global study on master data management*. 2011. viewed 11 March 2013 from [http://www.pwc.com/en\\_US/us/increasing-it-effectiveness/assets/global-study-master.pdf](http://www.pwc.com/en_US/us/increasing-it-effectiveness/assets/global-study-master.pdf)

Qualitative methods, n.d., *Research designs: Qualitative methods*, viewed 28 April 2013, from [http://www.charlesdennishale.com/books/eets\\_ap/9\\_Qualitative\\_Research\\_Designs.pdf](http://www.charlesdennishale.com/books/eets_ap/9_Qualitative_Research_Designs.pdf)

Quinn, G. & Keogh, M., 2001, '*Comparing groups or treatments: Analysis of variance*', viewed 6 November 2014, from <http://bio.classes.ucsc.edu/bio286/MlcksBookPDFs/QK08.PDF>

Reid, A. & Cattail, M., 2005, '*Invisible data quality issues in a CRM implementation*', viewed 27 April 2013, from [www.palgrave-journals.com/dbm/](http://www.palgrave-journals.com/dbm/)  
Sage Publications, 2010, 'Introduction to quantitative research', viewed 22 July 2014, from [http://www.sagepub.com/upm-data/36869\\_muijs.pdf](http://www.sagepub.com/upm-data/36869_muijs.pdf)

Rootman C. 2006. *The influence of customer relationship management on the service quality of banks*. MCom dissertation. Port Elizabeth: Nelson Mandela Metropolitan University

Sheina, M., 2010, 'Bad data costing US businesses \$700 billion a year', viewed 9 May 2013, from [http://ovum.com/press\\_releases/bad-data-costing-us-businesses-700-billion-a-year/](http://ovum.com/press_releases/bad-data-costing-us-businesses-700-billion-a-year/)

Shuttle worth, M., 2008, *Quantitative research design*, viewed 22 July 2014, from <https://explorable.com/quantitative-research-design>

Syed, N. & Liaoyang, L., 2012, 'The linkage between knowledge management practices and company performance: Empirical evidence', viewed 9 May 2013, from [www.jiem.org](http://www.jiem.org)

Tulip, F., 2010, 'The basis of distinction between qualitative and quantitative research in social science: Reflection on ontological, epistemological and methodological perspectives', viewed 11 May 2013, from <http://www.ajol.info>

Tustin, L. & Martins, V-W., 2005, *Marketing research in practice*, Unisia Press, Pretoria.

USC Libraries, 2014, *Organizing your social sciences research paper, quantitative methods*, viewed 15 September 2014, from <http://libguides.usc.edu/content.php?pid=83009&sid=615867>

Weber, n.d., *Types of research: Quantitative research*, viewed 15 September 2014,  
from  
<http://www.weber.edu/wsuiimages/jmitchell/MEDUC%206355/TYPES%20OF%20RESEARCH.pdf>

## ANNEXURES

---

### Annexure 1

#### **THE IMPACT OF DATA QUALITY ON THE SUCCESSFUL IMPLEMENTATION OF A CLIENT-RELATIONSHIP MANAGEMENT (CRM) SYSTEM**

##### **Instructions:**

Please read the statements below and show the extent to which you agree or disagree with each one. The statements relate to data quality, data-quality management and the value of CRM in the bank. There are no right or wrong answers – please merely tick on the column that best reflects your views. Thank you for your participation.

##### **GENERAL AND DEMOGRAPHIC INFORMATION**

Please select your option from the dropdown list by clicking on the word “**select option**”.

1. Gender

**Select Option**

2. Age

**Select Option**

3. How long have you worked for the bank

**Select Option**

4. Ethnicity

**Select Option**

5. My primary function is in

**Select Option**

6. The primary type of information that I use is mainly

**Select Option**

## PERCEPTIONS OF DATA QUALITY

	Statements	Very strongly disagree	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Very strongly agree
		1	2	3	4	5	6	7
1.	Our client contact information is scattered amongst various databases.	<input type="checkbox"/>						
2.	The information is consistently presented in the same format.	<input type="checkbox"/>						
3.	The information is easily retrievable.	<input type="checkbox"/>						
4.	The information is accurate.	<input type="checkbox"/>						
5.	The information is believable.	<input type="checkbox"/>						
6.	The information is easily accessible.	<input type="checkbox"/>						
7.	The information is reliable.	<input type="checkbox"/>						
8.	The information is sufficiently up to date.	<input type="checkbox"/>						
9.	The information is protected with adequate security.	<input type="checkbox"/>						
10.	The data is relevant to my work.	<input type="checkbox"/>						
11.	The information is sufficiently current.	<input type="checkbox"/>						

## PERCEPTIONS OF DATA-QUALITY MANAGEMENT

	Statement	Very strongly disagree	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Very strongly agree
		1	2	3	4	5	6	7
1.	The bank has adopted a data-quality management strategy.	<input type="checkbox"/>						
2.	The bank has the tools to identify any deficiencies with its data.	<input type="checkbox"/>						
3.	The bank has tools to ensure the consistency of its data.	<input type="checkbox"/>						
4.	The bank has the tools to identify duplicate data.	<input type="checkbox"/>						
5.	The bank has the tools to ensure correctness of data.	<input type="checkbox"/>						
6.	The bank has the tools to ensure completeness of data.	<input type="checkbox"/>						
7.	Ensuring the quality of the data is the responsibility of those who use the data.	<input type="checkbox"/>						
8.	The bank has clear guidelines on the standardisation of data.	<input type="checkbox"/>						

		Yes	Not sure	No
9.	There are people whose primary job it is to ensure the quality of information.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.	The bank has a specific business unit responsible for information quality.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**PERCEPTIONS OF THE VALUE OF CLIENT-RELATIONSHIP MANAGEMENT  
(CRM)**

	Statement	Very strongly disagree	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Very strongly agree
		1	2	3	4	5	6	7
1	The CRM tool is a necessity for the bank.	<input type="checkbox"/>						
2	The value of the CRM tool is well understood by me.	<input type="checkbox"/>						
3	The success of my job will be highly enhanced by the CRM tool.	<input type="checkbox"/>						
4	The bank will benefit from the use of the CRM tool.	<input type="checkbox"/>						
5	I understand how the tool will benefit the bank.	<input type="checkbox"/>						
6	The CRM tool will assist the bank with cross-selling opportunities.	<input type="checkbox"/>						
7	The CRM tool will help in increasing customer satisfaction.	<input type="checkbox"/>						
8	The value of the CRM tool has been well communicated.	<input type="checkbox"/>						
9	CRM will enhance sales-activity reporting.	<input type="checkbox"/>						
10	CRM will assist the bank to broaden its marketing initiatives, capabilities and efficiencies.	<input type="checkbox"/>						