

LB/DCN/18/09

AWARENESS OF DATA WAREHOUSING AND DATA MINING IN THE COMMERCIAL BANKING SECTOR OF SRI LANKA

LIBRARY
UNIVERSITY OF MORATUWA, SRI LANKA
MORATUWA

By

S.L Kulatilake



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Supervised By

Dr. A.S. Perera

University of Moratuwa



93340

The Dissertation was submitted, to the Department of Computer Science & Engineering of the University of Moratuwa, in partial fulfilment of the requirement, for the Degree of Master of Business Administration.

Department of Computer Science & Engineering

University of Moratuwa

December 2008

93340


004 "08"
004(043)

93340

TH

DECLARATION

"I certify that this thesis does not incorporate, without acknowledgement, any material previously submitted for a degree or diploma in any university. To the best of my knowledge and belief, it does not contain any material previously published, written or orally communicated by any other person or me, except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, to be made available for photocopying and for interlibrary loans, and for the title and summary to be made available to outside organizations"


.....

Signature of the Candidate

22nd Dec, 2008
.....

Date



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

To the best of my knowledge, the above particulars are correct.

UOM Verified Signature


Supervisor

22/12/08



ABSTRACT

Being a survivor in today's competitive business world, where information has become the most important asset, is indeed a colossal task, in itself. Many industries worldwide have taken forward strides by investing heavily in data technology. The banking industry in Sri Lanka is one industry which is highly information centric. This study is focused on measuring the awareness of Data Warehousing and Mining, within the local banking sector and finding out the factors that affect awareness. Data was initially gathered through interviews. A questionnaire was also given out to a sample of commercial bank employees from both the public and the private sectors. The questionnaire measures awareness levels in Data Warehousing and Mining as well as the factors that affect the awareness, which are Knowledge levels of employees, Business Requirements, Influence by higher management which affects the socio technical aspects and cross industry awareness, of Data Warehousing and Mining.

 The results of this study indicate, that, in terms of business requirements, basic knowledge and industry knowledge, awareness in the local commercial banking sector is high. It was also found that business practices and management initiatives have a significant impact on the awareness. Socio-technical aspects affected awareness, only where there was a data warehouse already implemented. The study concludes with a list of recommendations that the researcher sees, as suitable for the commercial banking sector in Sri Lanka to adopt, in order to improve awareness and improve usage of Data Warehousing and Mining and thereby improve on the service provided.



ACKNOWLEDGEMENT

First I would like to thank my supervisor, Dr. Shehan Perera for accepting my proposal and always being receptive, supportive and flexible. His constant confidence in me was definitely a driving factor to make this endeavour a success.

I would also like to acknowledge the support and guidance provided by the staff of the Department of Computer Science and Engineering, our head of department, Mrs. Vishaka Nanayakara and Dr. Chandana Gamage, our co-ordinator for the programme. Your patience was invaluable when time was running out.

If not for the cooperation from the banking industry, this research would never have been in existence. I wish to thank all the senior officials who readily granted interviews and to all the employees who took time to fill in the questionnaire.

Special thanks go to my fellow batchmates, Sunari, Nuwan, Malanie and Jagath who helped me in obtaining interviews and in the distribution and collection of data while completing their own research.

There were times when the path was uphill. Thank you Duminda for unconditionally taking time to give advice and show that analysis is in itself an art.

Thank you Wickum, for all the valuable guidance and encouragement, as I travelled along the path of this research study.

A big thank you goes to my family and in-laws who were always receptive and tolerant of the sacrifices that had to be made, in order to make this, a reality.

And last but not least, I wish to thank Milindu, for had it not been for the constant support and encouragement from him, whilst doing his own research, this would not have materialized

TABLE OF CONTENTS

| | |
|--|-------------|
| DECLARATION..... | I |
| ABSTRACT..... | II |
| ACKNOWLEDGEMENT | III |
| TABLE OF CONTENTS | IV |
| LIST OF FIGURES | VIII |
| LIST OF TABLES | IX |
| CHAPTER 1: INTRODUCTION..... | 1 |
| 1.1 Background..... | 1 |
| 1.1.1 Introduction | 1 |
| 1.1.2 Motivation | 1 |
| 1.2 Research Topic | 3 |
| 1.2.1 Problem Statement..... | 3 |
| 1.2.2 Research Objectives..... | 3 |
| 1.2.3 Significance of the Study..... | 3 |
| 1.3 Research Methodology..... | 4 |
| 1.4 Nature and Form of Results..... | 5 |
| CHAPTER 2: LITERATURE SURVEY | 6 |
| 2.1 The Case for Data Warehousing and Mining | 6 |
| 2.1.1 Improved decision making | 6 |
| 2.1.2 Better quality of information | 7 |
| 2.1.3 Ease of operations..... | 8 |
| 2.1.4 Performance of systems | 8 |
| 2.1.5 Better practicality..... | 9 |
| 2.1.5.1 Improving Response Time..... | 9 |
| 2.1.5.2 Improving data quality..... | 9 |
| 2.1.5.3 Archival reasons..... | 9 |
| 2.1.5.4 Access Control purposes..... | 10 |
| 2.2 A brief Introduction to data warehousing, Data Marts and Data Mining | 10 |
| 2.2.1 The Data Warehouse..... | 11 |

| | | |
|-------------------------------------|---|-----------|
| 2.2.2 | Building the Data Warehouse..... | 12 |
| 2.2.2.1 | Extracting Data..... | 12 |
| 2.2.2.2 | Cleaning and Transforming the data | 13 |
| 2.2.2.3 | Aggregation and Analysis..... | 13 |
| 2.2.3 | The Data Mart..... | 13 |
| 2.2.4 | The Difference between Data Warehouses and Data Marts..... | 15 |
| 2.2.5 | Data Mining..... | 17 |
| 2.3 | Data Warehousing and Mining Applied to Business Domains | 18 |
| 2.4 | Data Warehousing and Mining in Commercial Banking..... | 21 |
| 2.4.1 | Achieving Operational performance..... | 21 |
| 2.4.2 | Being customer centric..... | 23 |
| 2.4.3 | Moving towards multi-channel systems..... | 23 |
| 2.4.4 | Regulatory requirements | 24 |
| 2.4.5 | Other operational areas | 25 |
| 2.5 | Critical factors in Data Warehousing and Mining | 25 |
| CHAPTER 3: METHODOLOGY | | 28 |
| 3.1 | Theoretical Background | 28 |
| 3.1.1 | Awareness studies in General..... | 28 |
| 3.1.2 | Awareness studies in ICT usage in an industrial setting..... | 29 |
| 3.2 | Research Approach | 31 |
| 3.2.1 | Conceptual framework | 31 |
| 3.2.2 | Research Hypotheses | 33 |
| 3.3 | Research Design | 36 |
| 3.3.1 | Research Instruments | 36 |
| 3.3.2 | Operationalising the variables | 36 |
| 3.4 | Sampling Design..... | 37 |
| 3.4.1 | The Population..... | 37 |
| 3.4.2 | Sampling procedure of banks | 38 |
| 3.4.3 | Sampling procedure of employees..... | 38 |
| 3.5 | Data Collection | 39 |
| 3.5.1 | Interviews | 39 |
| 3.5.2 | Questionnaires | 39 |
| 3.6 | Data Analysis..... | 39 |

| | | |
|---|---|-----------|
| 3.6.1 | Checking Reliability and Validity | 40 |
| 3.6.2 | Descriptive Analysis | 40 |
| 3.6.3 | Statistical Analysis..... | 40 |
| CHAPTER 4: DATA ANALYSIS & RESULTS | | 41 |
| 4.1 | Reliability and Validity | 41 |
| 4.1.1 | Reliability | 41 |
| 4.1.2 | Validity | 42 |
| 4.2 | Descriptive Analysis | 42 |
| 4.2.1 | Composition of the Sample | 42 |
| 4.2.1.1 | Composition by Sector of Bank | 42 |
| 4.2.1.2 | Composition by Bank | 43 |
| 4.2.1.3 | Composition by whether Data Warehousing and Mining is implemented or not | 43 |
| 4.2.2 | Descriptive Analysis of Concept and Construct Variables | 43 |
| 4.2.2.1 | Business Practices | 43 |
| 4.2.2.2 | Management Initiatives | 45 |
| 4.2.2.3 | Socio-technical Aspects | 47 |
| 4.2.2.4 | Awareness of Business Practices for overall awareness of Data Warehousing and Mining | 49 |
| 4.2.2.5 | Awareness of Industry practices for overall awareness on Data Warehousing and Mining | 50 |
| 4.2.2.6 | Awareness of basic knowledge for overall awareness of Data Warehousing and Mining | 51 |
| 4.2.2.7 | Overall Awareness of Data Warehousing and Mining..... | 52 |
| 4.2.3 | Overall Summary of Research Constructs | 53 |
| 4.3 | Statistical Analysis..... | 53 |
| 4.3.1 | Awareness of Data Warehousing and Mining..... | 53 |
| 4.3.1.1 | Awareness of Business Practices | 54 |
| 4.3.1.2 | Awareness of Basic Knowledge on Data Warehousing and Mining | 55 |
| 4.3.1.3 | Awareness of Industrial Practices | 56 |
| 4.3.2 | Relationship between Business Requirements and Awareness..... | 57 |
| 4.3.3 | Relationship between Management Initiatives and Awareness | 58 |
| 4.3.4 | Relationship between Socio-technical Aspects and Awareness..... | 59 |

| | | |
|---|--|-----------|
| 4.3.4.1 | Relationship when data warehousing is implemented..... | 59 |
| 4.3.4.2 | Relationship when data warehousing is not implemented..... | 60 |
| 4.3.5 | Relationship between Skill levels and Awareness..... | 60 |
| 4.3.6 | Relationship between the fact that there is a Data Warehouse already implemented and Awareness of Data Warehousing and Mining | 61 |
| 4.4 | Overall Results of the Study | 62 |
| CHAPTER 5: DISCUSSION & RECOMMENDATIONS..... | | 63 |
| 5.1 | Discussion on Awareness of Data Warehousing and Mining | 63 |
| 5.2 | Discussion on Management Initiatives for Data Warehousing and Mining Awareness | 64 |
| 5.3 | Recommendations on Improving Management Initiatives | 65 |
| 5.4 | Discussion on Socio-Technical Aspects for Data Warehousing and Mining Awareness | 67 |
| 5.5 | Recommendations on Improving Socio-Technical aspects..... | 69 |
| 5.6 | Discussion on Data Warehousing and Mining Awareness with respect to Business Practices..... | 70 |
| 5.7 | Recommendations on Improving Business Practices | 71 |
| 5.8 | Recommendations on Technical perspectives for Data Warehousing and Mining..... | 73 |
| 5.9 | Recommendations on having Good Awareness Programs | 74 |
| CHAPTER 6: CONCLUSION | | 76 |
| 6.1 | Significance of the Study | 76 |
| 6.2 | Research Findings | 77 |
| 6.3 | Limitations of this Study..... | 78 |
| 6.4 | Further Research..... | 78 |
| REFERENCES..... | | 80 |
| APPENDIX..... | | A |

LIST OF FIGURES

| | |
|---|----|
| Figure 2.1: Compliance and Regulatory issues that drove IT expenditure in 2007..... | 24 |
| Figure 3.1: Framework adapted by the International benchmarking study..... | 30 |
| Figure 3.2: Conceptual framework between Awareness on Data Warehousing & Mining and factors affecting awareness..... | 32 |



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

LIST OF TABLES

| | |
|---|----|
| Table 1.1: Important statistics of the banking industry of Sri Lanka as at end 2007..... | 2 |
| Table 3.1: Break down of types of banks in Sri Lanka | 37 |
| Table 3.2: List of local commercial banks | 38 |
| Table 4.1: Reliability Analysis of constructs..... | 41 |
| Table 4.2: Validity Analysis of awareness | 42 |
| Table 4.3: Summary table for sector wise frequency of respondents | 42 |
| Table 4.4: Summary table for bank wise frequency of respondents..... | 43 |
| Table 4.5: Summary table for bank wise frequency of respondents..... | 43 |
| Table 4.6: Descriptive statistics of Business Practices | 44 |
| Table 4.7: Descriptive statistics of quality of information | 44 |
| Table 4.8: Descriptive statistics of churn analysis using historical data | 44 |
| Table 4.9: Descriptive statistics of profitability analysis using historical data | 45 |
| Table 4.10: Descriptive statistics of Management Initiatives..... | 45 |
| Table 4.11: Descriptive statistics of importance of a Data Warehouse being stressed | 46 |
| Table 4.12: Descriptive statistics of stressing the use of a warehouse against whether there is a DWH implemented | 46 |
| Table 4.13: Descriptive statistics of initiatives on providing training | 46 |
| Table 4.14: Descriptive statistics of importance of recognition | 47 |
| Table 4.15: Frequencies on providing training on DWH&M | 47 |
| Table 4.16: Frequencies on whether the provided training was related to day-to-day operations | 48 |
| Table 4.17: Frequencies on whether motivated to use the DWH solution..... | 48 |
| Table 4.18: Frequencies on whether comfortable and assured with the information from DWH..... | 48 |
| Table 4.19: Frequencies on whether awareness training was provided | 48 |
| Table 4.20: Frequencies on anticipation of a data warehousing solution | 48 |
| Table 4.21: Descriptive statistics on Awareness of Business Practices..... | 49 |
| Table 4.22: Descriptive statistics on overall Awareness of Business Practices | 49 |
| Table 4.23: Frequency on whether DWH is used in the banking industry in general . | 50 |
| Table 4.24: Frequency on awareness of other industries using DWH&M | 50 |
| Table 4.25: Frequency on whether a DWH will help achieve BASEL II compliance | 50 |
| Table 4.26: Frequency on knowledge of common DWH&M packages used in the | |

| | |
|--|----|
| banking industry | 50 |
| Table 4.27: Descriptive statistics on overall awareness on industrial knowledge | 51 |
| Table 4.28: Frequency on overall awareness on industrial knowledge | 51 |
| Table 4.29: Frequency on basic knowledge on data warehousing and mining | 52 |
| Table 4.30: Descriptive statistics on overall basic knowledge on data warehousing and mining | 52 |
| Table 4.31 Descriptive statistics of research constructs..... | 53 |
| Table 4.32 One sample t-test results on awareness of business practices mean..... | 54 |
| Table 4.33 One sample t-test results on awareness of basic knowledge mean. | 55 |
| Table 4.34 One sample t-test results on awareness of industrial knowledge mean. | 56 |
| Table 4.35 Correlation results between Awareness and Business Requirements | 57 |
| Table 4.36 Correlation results between Awareness and Management Initiatives | 58 |
| Table 4.37 Correlation results between Awareness and Socio-technical Aspects in the presence of Data Warehouse | 59 |
| Table 4.38 Correlation results between Awareness and Socio-technical Aspects in the absence of Data Warehouse | 60 |
| Table 4.39 Factors for measuring employee skill levels..... | 60 |
| Table 4.40 ANOVA results on awareness between skill definition groups | 61 |
| Table 4.41 ANOVA results on awareness between groups where there is a DWH and there is no data warehouse. | 62 |