

ACKNOWLEDGMENT

In the beginning, I thank Allah for all his grace also I want to thank Dr. AdilYousif for overseeing this research .I would like to thank my parents, whose pray and avoidance are with me during the study period and providing me all support I need .I wish to thank my colleagues who provide valuable assistance, which made me complete the research.

Abstract

It's difficult to make a decision, without enough information because the databases in the branches of Sales department of National Paints factory are not integrated for management of Information and decisions making. So for this reason we prepared a study on designing of data warehouse model to support forecasting and decision making of paint sales. The main objective of this research is to collect those data in one repository. Data warehouse is responsible for consistency of information. The aim of data warehousing is to organize the gathering data and store it in single repository. Data warehousing solve the problem and provide technology which enables the user or decision maker to process the data in short time. With the help of data warehousing, manager extract the Knowledge in real time and it helps the manager in the decision making process. The research used SQL Server Data Tool that contains SQL Server Integration Services (SSIS) due to consistency and integrating data. Also, it used SQL Server Analysis Services (SSAS) to analyze data and build a dimensional model. Manager can understand the data easily by used cubes of dimensional model. Also, manager can drill down or up by adding or removing attributes from their analyses with excellent performance. Moreover, the researcher explored data by using Time Series Algorithm and prediction of the quantity during the first quarter will be in January and February between 1-50 products only, whereas its increase to 101 products in begins of March. As well as, the researcher used SQL Server Reporting Services (SSRS) to review reports after analysis to answer to questions of manager and support him to make decisions.

المستخلص

يصعب اتخاذ أي قرار دون الحصول على معلومات كافية وذلك لان قواعد بيانات فروع مبيعات مصنع الوطنية للدهانات غير متكاملة لإدارة المعلومات واتخاذ القرار اتلذلك قمنا بدراسة بناء مستودع البيانات لدعم اتخاذ القرار والتنبؤ بمبيعات الدهانات وكان الهدف الرئيسي من البحث هو إنشاء مستودع بيانات لتجميع كل البيانات فيه .مستودع البيانات هو المسئول عن تناسق المعلومات . الهدف من مستودع البيانات هو تنظيم البيانات المجمعة وتخزينها في مستودع واحد.مستودع البيانات حل مشكلة البحث ووفر تقنية تساعد المستخدم او صانع القرار لمعالجة البيانات في زمن قصير .بمساعدة مستودع البيانات يستطيع المدير استخراج المعرفة في الزمن المناسب التي تساعد المدير في عملية اتخاذ القرار .تم استخدام أداة خدمات التكامل في خادم قاعدة البيانات لتكامل البيانات والتأكد من تجانسها . كما تم استخدام خدمات التحليل في خادم قاعدة البيانات لتحليل البيانات وبناء نموذج الأبعاد يمكن للمدير فهم البيانات بسهولة باستخدام مكعباتنموذج الأبعاد.كما يمكن للمدير أن يضيف أو يزيل اي سمة من التحليل معالحفاظة على الأداء الممتاز. كما تم التنقيب في البيانات باستخدام خوارمية التسلسل الزمني والتنبؤ بأن الكمية المباعة خلال الربع الأول سوف تكون في شهر يناير وفبراير تتراوح بين 1-50 منتج فقط بينما تتزايد في بداية شهر مارس الي 101 منتج. وتم استخدام خدمات التقرير في عرض تلك التقارير بعد تحليلها للإجابة على اسئلة المدير ودعمه في اتخاذ القرار.

TABLE OF CONTENTS

الآية.....	I
Acknowledgment.....	II
Abstract.....	III
المستخلص	IV
Table of Contents.....	V
List of Figures.....	VIII
List of Tables.....	IX
List of Abbreviations.....	X
CHAPTER1: Introduction	
1.1 Introduction.....	1
1.2 Research Background.....	1
1.3 Problem Statement.....	2
1.4 Research Objectives.....	3
1.5 Propose Solution	3
1.6Scope.....	4
1.7 Thesis Structure.....	4
CHAPTER 2: Literature Review	
2.1. Introduction.....	5
2.2. Background Theory.....	5
2.2.1 The data warehouse concept.....	5
2.2.2 Main components of the data warehouse.....	6
2.2.3.Difference between data warehouse and database.....	9
2.2.4 Data warehouse modelling.....	10
2.2.5 Star schema.....	10
2.2.6 Dimensional model.....	11
2.2.7 Fact tables.....	12
2.2.8 Dimension tables.....	13
2.2.9 Data Marts.....	13

2.2.10 Data warehouse VS Data mart.....	14
2.2.11 Data mining.....	14
2.2.12 Microsoft Time Series Algorithm.....	14
2.2.13 Data mining Exetesion.....	15
2.3 Related Studies.....	16
2.4 About Company and Data Set.....	18

CHAPTER3: ResearchMethodology

3.1 Introduction.....	19
3.2 Research Design.....	19
3.2.1System Analysis.....	20
3.2.2 Preparation to Design of a dimensional model.....	21
3.2.3The Data Mining Process.....	21
3.2.4 Data Required for Time Series Models.....	22
3.2.5 How the Algorithm Works.....	22
3.3 Tool Used in Methodology.....	23

CHAPTER4: The Proposed Data warehouse Implementation

4.1 Introduction.....	24
4.2 System Design.....	24
4.3 Logical Models.....	25
4.3.1 Fact and Dimension Tables.....	25
4.4 Design of the Physical Database.....	26
4.4.1.Fact and Dimension Tables.....	27
4.5 Design of the ETL Process.....	28
4.5.1.Dimension Tables Incremental Processing.....	28
4.6 SQL Server Analysis Services (SSAS).....	32
4.7 Cube.....	32
4.8SQL Server Reoprt Services (SSRS).....	38
4.8.1 Design Report.....	38

4.8.2 Preview Report.....	39
4.9 Data Mining using Microsoft Time Series Algorithm.....	43
4.9.1 Viewing Time Series Model.....	44
4.9.2 Create Time Series Predictions.....	45
CHAPTER 5: Conclusions	
5.1 Conclusion.....	47
5.2 Recommendations.....	47
References.....	48

List of Figures

Figures Number	Figures Name	Page Number
Figure(2.1)	Components of the data warehouse	7
Figure(2.2)	Star Schema	11
Figure(3.1)	Research Design	20
Figure(4.1)	High –Level Data Staging Plan Schematic	25
Figure (4.2)	Logical Model of Dimensional Model	26
Figure (4.3)	The Scripts to Build D W in SQL Server Management Studio	27
Figure (4.4)	Dimension Tables and Fact Table Diagram	28
Figure (4.5)	ETL product data	29
Figure (4.6)	ETL Date data	30
Figure (4.7)	ETL Store data	30
Figure (4.8)	load fact table	31
Figure (4.9)	Browsing Dimension of Store	32
Figure (4.10)	Browsing Dimension of Date	33
Figure (4.11)	Browsing Dimension of Product	33
Figure (4.12)	Cube Structure	34
Figure (4.13)	Calculation in a cube	35
Figure (4.14)	Cube named sets	35
Figure (4.15)	Cube Browser	36
Figure (4.16)	Ad-hoc pivot table reports	37
Figure (4.17)	Design Report with Wizard	38
Figure (4.18)	Report after Design	39
Figure (4.19)	Preview Report	40
Figure (4.20)	The Report of Sold Quantity in Year	41
Figure (4.21)	The Report of Best Selling in two Location by Quarter	42
Figure (4.22)	The Report of Lowest Selling in two Location by Quarter	43
Figure (4.23)	Chart of Quantity	45
Figure (4.24)	Extract predict value of Quantity ,SalesCost	46

List of Tables

Table Number	Table Name	Page Number
Table(2.1)	Comparison between database and data warehouse	9
Table(2.2)	Differences between Data warehouse VS Data mart.	14
Table(4.1)	sample for data in data warehouse used in data mining	43
Table(4.2)	sample of output of query	46

List of Abbreviations

Abbreviations	Explanation
DB	Database
DBMS	Database Management System
DM	Dimensional Modeling
ER	Entity Relationship Modeling
ARTXP	Autoregressive Tree Models with Cross Prediction
ARIMA	Autoregressive Integrated Moving Average
DMX	Data Mining Extensions
DDL	Data Definition Language
DML	Data Manipulation Language
SSAS	SQL Server Analysis Services
SM	Sales Management
SQL	Structured Query Language
SSIS	SQL Server Integration Services
SSRS	SQL Server Report Services
DW	Data Warehouse
SSDT	SQL Server Data Tools
ETL	Extract Transform and Load
SCD	Slowly Changing Dimension
GUI	Graphical User Interface
BI	Business Intelligence