Business Intelligence System for Banking Credit Finance

Prepared By:

- Yousif Ahmed Alamin Yousif Yassin 09-612
- Osman Adil Osman Mohamed 09-607
- Abdulrahman Jaffer Bashir 08-607

Supervisor name:

- Mohammed Omer Ahmed Aburaid
ABSTRACT

Banking sectors and central banks are trying to find new solutions and technologies to face the credit finance risks. Therefore, banks want solutions that give them the ease of making decisions. However, data information systems, Data warehousing and business intelligence solutions become very important. They help the banks to make better decisions based on accurate data collected from many different sources.

A leading industry like Banking needs BI credit finance because of the importance of its data that grows up rapidly. Another reason is variety of banking business environment that frequently updated and restructured. This project aimed at developing a suitable BI credit finance that could fulfill the needs of the banking sector in Sudan and face the problems that prevent using BI credit finance in this sector.

The project has concentrated on management credit finance highlighted report and dashboard in the bank where the case study were made, and focused on some new intelligent method for building such banking data warehouse.

At this project, there is a discussion about the basics of the Business intelligence solution, the management of the data warehouse environment and the ETL processes.

Finally, a subset of the BI solution were delivered using Oracle Hyperion reporting tools and a developed system for the purpose of configuration of the fact and dim table (Push data) configuration of the data source, configuration of the destination database (Data warehouse)
**Table of Contents**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Abbreviations</td>
<td>5</td>
</tr>
<tr>
<td>List of figures</td>
<td>6</td>
</tr>
<tr>
<td>Acknowledgment</td>
<td>7</td>
</tr>
<tr>
<td>Declaration</td>
<td>8</td>
</tr>
<tr>
<td><strong>Chapter one</strong></td>
<td></td>
</tr>
<tr>
<td>1.1 Motivation</td>
<td>10</td>
</tr>
<tr>
<td>1.2 Problem</td>
<td>10</td>
</tr>
<tr>
<td>1.3 Project Aim &amp; Objectives</td>
<td>11</td>
</tr>
<tr>
<td>1.3.1 Aim</td>
<td>11</td>
</tr>
<tr>
<td>1.3.2 Objectives</td>
<td>11</td>
</tr>
<tr>
<td><strong>Chapter two</strong></td>
<td></td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>12</td>
</tr>
<tr>
<td>2.2 Business Intelligence</td>
<td>13</td>
</tr>
<tr>
<td>2.3 Business Intelligence History</td>
<td>14</td>
</tr>
<tr>
<td>2.4 Benefits of BI</td>
<td>15</td>
</tr>
<tr>
<td>2.5 Future of BI</td>
<td>17</td>
</tr>
<tr>
<td>2.6 BI Components</td>
<td>18</td>
</tr>
<tr>
<td>2.6.1 What is a Data Warehouse?</td>
<td>18</td>
</tr>
<tr>
<td>2.6.2 Data Warehouse Architectures</td>
<td>19</td>
</tr>
<tr>
<td>2.6.2.1 Data Warehouse Architecture (Basic)</td>
<td>19</td>
</tr>
<tr>
<td>2.6.2.2 Data Warehouse Architecture (with a Staging Area)</td>
<td>20</td>
</tr>
<tr>
<td>2.7 Reasons for Staging Area</td>
<td>20</td>
</tr>
<tr>
<td>2.8 Data Warehousing Schemas</td>
<td>21</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2.8.1 Star Schemas</td>
<td>22</td>
</tr>
<tr>
<td>2.9 Data Warehousing Objects</td>
<td>23</td>
</tr>
<tr>
<td>2.9.1 Fact Table</td>
<td>24</td>
</tr>
<tr>
<td>2.9.2 Dimension Tables</td>
<td>24</td>
</tr>
<tr>
<td>2.10 The Extract, Transform and Load</td>
<td>25</td>
</tr>
<tr>
<td>2.11 ETL Processes</td>
<td>26</td>
</tr>
<tr>
<td>2.11.1 Extraction in Data Warehouses</td>
<td>27</td>
</tr>
<tr>
<td>2.11.3 Data load</td>
<td>30</td>
</tr>
<tr>
<td>2.12 Chapter Summary</td>
<td>33</td>
</tr>
<tr>
<td><strong>Chapter three</strong></td>
<td>34</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>35</td>
</tr>
<tr>
<td>3.2 What is Microsoft SQL Server?</td>
<td>35</td>
</tr>
<tr>
<td>3.2.1 SQL Server License</td>
<td>35</td>
</tr>
<tr>
<td>3.3 SQL Server Database Engine (Server)</td>
<td>35</td>
</tr>
<tr>
<td>3.4 Oracle database</td>
<td>36</td>
</tr>
<tr>
<td>3.4.2 What does Oracle Database (Oracle DB) mean?</td>
<td>36</td>
</tr>
<tr>
<td>3.4.3 Oracle Database Architecture</td>
<td>36</td>
</tr>
<tr>
<td>3.5 Visual studio definition</td>
<td>36</td>
</tr>
<tr>
<td>3.6 C# Programming Language</td>
<td>38</td>
</tr>
<tr>
<td>3.7 Hyperion Reporting Tools</td>
<td>38</td>
</tr>
<tr>
<td>3.7.1 Hyperion Definitions</td>
<td>39</td>
</tr>
<tr>
<td>References</td>
<td>41</td>
</tr>
</tbody>
</table>
List of Abbreviations

<table>
<thead>
<tr>
<th>BA</th>
<th>Business Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI</td>
<td>Business Intelligence</td>
</tr>
<tr>
<td>DW</td>
<td>Data Warehouse</td>
</tr>
<tr>
<td>DSS</td>
<td>Decision support system</td>
</tr>
<tr>
<td>ETL</td>
<td>Extract Transform and load</td>
</tr>
<tr>
<td>IT</td>
<td>Information Technology</td>
</tr>
<tr>
<td>MIS</td>
<td>Management Information System</td>
</tr>
<tr>
<td>OLAP</td>
<td>On-Line Analytical Processing</td>
</tr>
</tbody>
</table>
List of figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2.1:</td>
<td>data, information, knowledge</td>
<td>14</td>
</tr>
<tr>
<td>Figure 2.2:</td>
<td>PRESENT AN UNDERSTANDING OF BI</td>
<td>17</td>
</tr>
<tr>
<td>Figure 2.3:</td>
<td>DW ARCHITECTURE</td>
<td>20</td>
</tr>
<tr>
<td>Figure 2.4:</td>
<td>DW ARCHITECTURAL WITH STAGING AREA</td>
<td>21</td>
</tr>
<tr>
<td>Figure 2.5:</td>
<td>Star schema</td>
<td>22</td>
</tr>
<tr>
<td>Figure 2.7:</td>
<td>Dim And Facts</td>
<td>23</td>
</tr>
<tr>
<td>Figure 2.8:</td>
<td>Overview of ETL</td>
<td>25</td>
</tr>
<tr>
<td>Figure 2.9:</td>
<td>ETL Process</td>
<td>26</td>
</tr>
</tbody>
</table>
Acknowledgment

At the beginning, we would love to thank the Allah the Almighty for giving us the strength, power, inspiration, courage, health, determination and patience to succeed in the partial of this project.
We would like to thank our supervisor T. Mohammed Omer Ahmed Aburaid for his advice, assistance and comments, which we always used to receive when we were just in need for them.
Declaration

"We understand the definition of plagiarism. If it is shown the material has been plagiarized, or we have attempted to obtain an unfair advantage for others, or ourselves we understand that we may face couselled in accordance with the policies of the university. A mark of zero may be awarded and the reason for that mark will be recorded and big mistake in my future come"
*Chapter One*

*Introduction*
Introduction:
Nowadays, with the competitive business environment, a better, fast and fact-based business decision-making methodology had been really needed. To achieve this aim, decision makers need more than an opinions and self-made experience.

Since technology evolution and increasing, specialists in all different fields and sources started to develop computer-based technologies using their own knowledge and experiences. One example is the business field where many new technologies faced and some of other technologies adapted to use such computer and control systems. This is more like driving the technology wave toward business and seeing the business through the technology future appearance.

Banks nowadays became one of the largest users of information systems and a lot of information-based technologies like Business Intelligence (BI), Data warehousing (DW) and Decision Support Systems (DSS). Using analytical data to from normal databases that designed basically to serve operational systems is the hard part but maintaining a clean and ready to use analytical data by find new fact-based information.

BI provides a single consistent source of management information for reporting and analysis across the organization.

1.1 Motivation:-
The Sudanese-banking sector has already automated most of its business processes through acquiring core-banking systems such as IMAL.

The static highlighted reports produced by such systems do not effectively extract the most useful information out the massive data, which will prop up the decision-making.

1.2 Problem:-
Need to have a better methodology for the analysis and the reporting processes to support better decision-making for credit finance.
1.3 Aims & Objectives:-

1.3.1 Aim:-

- Provide an optimal methodology for tracking credit finance processes by using common functions of BI technologies for analytics and highlighting reports that can serve the banking sector analytical needs based on own made facts and dimensions tables by the manager or the user of the system (BI Configurator).

- To fully integrate all the bank information.

Objectives:-

- Interviews with bankers.

- Gather the required information to start building the Bi configurator (FACT, DIM, CONNECT)

- After pushing the data the Report Sector can work on the table created by the system to build their reports.

- Testing the efficiency of the analytics and the highlighted reports and how it improved the current business flow.
*Chapter Two*

*Literature Review*
2.1 Introduction:

"Business Intelligence is the process of transforming data into information and through discovery transforming that information into knowledge." – Gartner Group [1]

Before we talk about Business intelligence to begin with, it is better to define the data, information and knowledge.

Data:
As datum (Federal Enterprise Architecture), is "A value or set of values representing a specific concept or concepts". In computer field, alphanumeric, image, audio, and video are types of data. [2]

Information:
Conceptually, information has many meanings depend on the field that you are investigating on. Ralph Stair and George Reynolds define information at their Fundamentals of Information Systems as "Information: collection of facts organized in such a way that they have additional value beyond the value of the facts themselves" "Data become "information" when analyzed and possibly combined with other data in order to extract meaning, and to provide context. The meaning of data can vary depending on its context". [2]

Knowledge:
Knowledge refers to meaning and understanding that results from processing information by users. That mean knowledge is derived from information in same way information is derived from data; it is a person’s range of information.
2.2 BUSINESS INTELLIGENCE:-

Nowadays, the business world competitive modified an important methodology about business decision making more than their self-experience and opinions to achieve this goal.

Since its beginnings, information technology (IT) has been transforming the nature of products, processes, companies, Industries, and even competition itself (Porter & Millar, 1985). For today's organizations, in order to succeed, it is important to understand how this technology can create substantial and sustainable competitive advantages. It not only affects how individual process activities are performed, but, Through new information flows, it is also greatly enhancing a company's ability to exploit linkages between activities, both inside and outside the company (Porter & Millar, 1985, p. 152).[3]

Since the information age get started till the days we live now the information getting used, so that general business had to gather their data from not automated sources and storages, companies even used this information and made their decisions primarily on the intuition and then analyze the data after business lacked their resources.
More and more information and data became available after systems started to automating their systems. However, collection made a challenge to a lack of initiation and infrastructure for exchanging data between systems, analyze of these data get months to generate which gathered, collected and reports. Such reports allowed informed long time strategic of decision-making.

In general BI is:

Is a set of theories, methodologies, architectures, and technologies that transform raw data into meaningful and useful information for business purposes. BI can handle large amounts of information to help identify and develop new opportunities. Making use of new opportunities and implementing an effective strategy can provide a competitive market advantage and long-term stability.

2.3 Business Intelligence History:

Hence, the need for business intelligence, a term and a definition which date back to an October 1958 IBM Journal article by Hans Peter Luhn entitled A Business Intelligence System. [3]

Luhn wrote:

“Business is a collection of activities carried on for whatever purpose, be it science, technology, commerce, industry, law, government, defense, et cetera. The communication facility serving the conduct of a business in the broad sense may be referred to as an Intelligence system. The notion of intelligence is also defined here, in a more general sense, as “the ability to apprehend the interrelationships of presented facts in such a way as to guide action towards a desired goal.”
As general, BI understanding have combination of data warehouse (DW) and decision support system (DSS).

This above picture presents an understanding of BI. A BI system in other words is a combination of data warehousing and decision support systems. The figure also reveals how data from disparate sources can be extracted and stored to be retrieved for analysis. The basic BI functions and reports are shown in the above picture. The primary activities include gathering, preparing and analyzing data. The data itself must be of high quality. The various sources of data is collected, transformed, cleansed, loaded and stored in a warehouse [4].

2.4 Benefits of BI:-

BI have many benefits to business environment world, it can increase doing their activities and work, decrease the guesswork and enable these companies within the business environment to respond faster and quickly to changes such as financial conditions, supply chain operations and customer development and preferences.
BI improves the performance of the companies using it. When one of the company can make decisions based on accurate and time of information, this company’s performance can improve. In addition, it can also improve the customer experience, by allowing the time and appropriate respond for these customers and their problems.

Employees and workers can easily now convert their knowledge by using the BI tools to solve their many issues, like internet marketing campaigns and increasing the e-mail and telephone response rate.

2.5 Future of BI:-

The Gartner Group makes a couple predictions about the future of BI. First, between 2009 and 2012, more than 35 per cent of the top 5,000 global companies will regularly fail to make insightful decisions about significant changes in their business and markets (Gartner Group). Faced with the need to make better decisions companies will seek to increase their investments in BI systems. [1]

Another important prediction is that business units will control at least 40 percent of the total budget for BI despite the fact that IT departments are responsible for building and maintaining the BI infrastructure (Gartner Group). The reason for this situation is that business users have lost confidence in the ability of IT departments to deliver the information they need to make decisions (Gartner Group). Faced with some problems it is no wonder that the business unit will take it into their own hands to create the BI infrastructure [1].

2.6 BI Components:-

2.6.1 Data Warehouse:-

"An enterprise structured repository of subject-oriented, time-variant, historical data used for information retrieval and decision support. The data warehouse stores atomic and summary data." [2]
A data warehouse is a database that is designed for analysis and query rather than using it in transaction processing. Usually, it contains historical data, which get from transaction data and it can include data from other sources. In the data warehouse, data is stored based on common subject areas (for example, customer, product, and so on) for ease of access.

Data warehouse (DW) separates analysis workload from transaction workload and enable organizations to make data by unit it from several sources.

"Data warehouse environment includes an extraction, transformation, and loading (ETL) solution, client analysis tools, online analytical processing (OLAP), and other applications which use for managing the process of gathering and delivering data." [5]

Subject Oriented and Integrated is ways of introducing the data warehouse as:

Subject Oriented:

Database is organized in which that the data elements which relating to the same object are linked together. That means the data in the DW designed to help employees analyze data.

Integrated:

Integration is almost related to subject orientation. Data warehouses put data from different sources into a consistent one. "They must resolve such problems as naming and Inconsistencies in units of measure" [5]. After we finish this, we can say the system is integrated.

2.6.2 Data Warehouse Architectures:-

Data warehouses have Tow common architectures:

2.6.2.1 Data Warehouse Architecture (Basic):-

A simple one for a data warehouse in which end users access directly data derived from the source systems through DW.
2.6.2.2 Data Warehouse Architecture (with a Staging Area):-

DW architecture in which transformation happening or takes place.

2.7 Reason for Staging Area:

- Recoverability: Stage immediately after major processing and the data when have been extracted from the source.
- Auditing: Linking between the indicating transformations source data before loading to the DW.
- Backup: use to reload the DW from the staging tables without return to the sources.
- Example for recoverability (cleaning, transformation, etc.).
2.8 Data Warehousing Schemas:-

A schema is a collection of database objects, including tables, views, indexes, and synonyms.

There is a variety of ways of arranging schema objects in the schema models designed for data warehousing. One data warehouse schema model is a star scheme.
2.8.1 Star Schemas:-

A star schema model can be depicted as a simple star: a central table contains fact data, and multiple tables radiate out from it, connected by database primary and foreign keys. Unlike other database structures, a star schema has denormalized dimensions.

Star model is easy to understand by the users because the structure is so simple and straightforward which provides fast response to queries with optimization and reductions in the physical number of joins required between fact and dimension tables. Contains simple metadata, supported by many front-end tools.

"The star schema is emerging as the predominant model for data warehouses or data marts and a slow to build because of the level of demoralization." [2]

![Star Schema Diagram]

**Figure 2.5:** Star schema
2.9 Data Warehousing Objects:

Fact tables and dimension tables are the two types of objects commonly used in dimensional data warehouse schemas. Fact tables are the large tables in DW, it contains facts and keys (foreign) to dimension tables, fact tables represent numeric and additive that can be tested. Examples: profit, cost, and sales.

Dimension tables (reference tables), contain static data in DW, it stores the information which use represents textual and describing data. Example: products or customers.

![Diagram of Facts and Dims]

**Figure**: Facts and Dims
2.9.1 Fact Table:-

The fact table is sometimes defined as the central table of the star schema of the data warehouse. It stores quantitative information for the purpose of analysis and it is denormalized.

A fact table is connected to the dimension table and they work together, the fact table as we said holds the data to be analyzed, and the dimension table stores data about the ways in which the data in the fact table can be analyzed. According to this the fact table consists of two types of columns, those that contain the stored data to be analyzed and the foreign key column that allows joins to dimension tables. A fact table may contain one of the following facts: detail-level facts or aggregated facts sometimes called (Summary tables). Usually a fact table contains facts that are in the same level of aggregation. Facts can be classified into: additive facts that can be aggregated by using simple arithmetical addition operation. Example are (sales). The second is non-additive facts that cannot be added at all. Example is (average). The last is semi-additive facts that are aggregated with some dimensions but not all of the dimensions. Example is (inventory levels) where you cannot know what dose a level mean by just looking at it.

2.9.2 Dimension Tables:-

A dimension is a structure that is used to categorize data in a way that helps and enable users to analyze and answer business questions. Dimension information is stored in a table that is called a dimension table. Dimensions have attributes that is used to describe the value of the dimension. Dimension data are always descriptive, and in the form of text. Examples of commonly used dimensions are (Customers, time and product). When start to collect dimension data its usually started at the lowest level of detail, then we reach the higher level by aggregating the data. The generation of the high level is useful in analysis. The aggregations of data are called Hierarchies within the dimension table.
2.10 The Extract, Transform and Load:-

Extraction, Transformation, loading (ETL): Is the process of extraction of data from several sources, cleansing and transforming the data to fit the business needs, finally the process of loading of this data into a data warehouse.

Overview of ETL

Data warehouses need to be loaded regularly to make business analysis easy and good. So to achieve this data from multiple sources and different operational systems should be extracted and copied into the warehouse. This process is commonly called ETL which stands for: Extraction, transformation and Loading. Different systems and applications need to share data and trying to establish integration between them, by giving them similar picture of the world. The data sharing and integration processes were mostly done by mechanisms similar to what we call now ETL.

![Figure 2.9: Overview of ETL](image-url)
2.11 ETL Processes:

Extract, transform and load (ETL) is the main process of data integration and it is associated with data warehousing. In the extraction phase, data is located from a chosen source, then it is transformed into a new form based on business rules, and then in the last phase it is loaded into the target data structure. After the loading process now, users can use the system and generate highlighted reports and dashboards that help them in their decision-making. Even though ETL process consumes roughly 70% of computing resources, they are hardly visible to the end users [6].

![ETL Process Diagram]

*Figure 2.10: ETL Process*
2.11.1 Extraction in Data Warehouses:

Overview of Extraction in Data Warehouses

Extraction is the process of extracting data from different data sources to be used in the data warehouse. Data extraction is the first function in the ETL process, after the extraction the data is then transformed and loaded into the data warehouse. However, there can be more than data source that are different from each other and they are located in different locations. Therefore the extraction process is not easy because the source systems could be very complex and not fully documented, for the reason it will be difficult to choose which data need to be extracted.

The extraction process should be done several times not only once, so that to supply all the changes to the data warehouse and keep it up to date. In the next section, we introduce you to some of the extraction methods in a data warehouse.

Extraction Methods in Data Warehouses

Selecting the type of the extraction method is not an easy decision to make, you should analysis and understand the source system structure and the business needs of the target data warehouse, on order to choose the suitable extraction method. First you should extract the data logically using (logical methods), then the physical extraction using (physical methods).
Logical Extraction Method

The extraction process needs to be efficient so that your extracted data are clean and from the right sources so the decisions will be made base on right data. The data from the sources may change every period or maybe updated, for this reasons data should be planned logically before the actual extraction. The logical extraction has two types:

- Full extraction
- Incremental extraction

- Full Extraction

In the full extraction, the data is completely extracted from the source systems.

- Incremental Extraction

In the incremental extraction only the data that has been changed since the last extraction or since a specific event in the history. However, some source systems have change-capture mechanism, that they have the ability to track the changed data. However, other data warehouses do not use the change-capture mechanism as part of their extraction process. In this case, the extraction is done by comparing new data extracted from the source with old ones, and then identifying the changed data.

Physical Extraction Methods

The physical extraction has two mechanisms to use depending on the choose of the logical extraction method and the source system constrains. The first mechanism is to extract the data online from the source system (online extraction); the second is to extract the data from an offline structure (offline extraction).
- **Online Extraction**

In the online extraction, the data is extracted from the source system itself directly or intermediate system where data is stored. The extraction is done by connecting directly to the source system data tables.

- **Offline Extraction**

On the offline extraction the data is extracted from existing data, structures for example (redo logs, archive log or table spaces). There is no direct connection with the source system. There are many data structures where you can use the offline extraction on, here are some of the common structures to consider:

**Flat File**

Is an operating system file that contains large amount of data (binary or text).

**Dump files**

Oracle-specific format. Information about the containing objects is included. [7]

**Redo and archive logs**

Information is in a special, additional dump file. [7]

**Web log sources**

Web log is a log used by internet companies to store visitor's information.

**Change Data Capture**

It is a mechanism for keeping in track new data and the data that has been changed since the last extraction process in the source system. When a data warehouse uses the incremental extraction or change data capture it only extracts the changed data, thus the extraction process and the other processes in the ETL process will be better in less time consuming because only small amount of volume of data will be extracted and dealt with.
2.11.3 Data load:

The loading process is the process where loading the data that have been worked on the previous phases, which is validated and integrated and transformed through the transform functions to the DW (Data warehouse). This data must be staged first before the loading process in temporary tables in the database (SA).

When loading data from staging area in to data warehouse if the load process fails due to only the load process can be restarted avoiding going through integrating and transformation processes all over again.

There are two types of data load

1- Historic load
2- Incremental load.

Historic load:

Aka full load and we mean by full load a fully refreshment of the data that is based on the historical load which is stays in the data warehouse, and this loading process have to be to controlled by a specific duration of load based on user requirements data in the data warehouse, this duration can take various times (years-decades), when data is loaded regularly in to the warehouse ETL process might break the data and fixing it would take several hours to several days.
Incremental load:

It generally means only loading into the warehouse the records that have changed since the last load, the main advantage is it reduce the data transfer to the data warehouse, incremental load is a regular interval load of the data warehouse and it load the most recent data from the OLTP.

This incremental process runs from time to time (daily, weekly, monthly, yearly, scheduled time) until the end of the warehouse’s life.

Every incremental load have a loading window indicates to the beginning and the ending the loading process, after that the users can work on the data after the ending of the load process.

Either loading a fact tables or a dimension tables, the type of the table determine the appropriate operation for the loading, as describes below the operations for the load processing.

**Insert operation:** inserts data into the warehouse, if data exist in table the insert operation fails.

**Update operation:** updates the existing records in the DW, if it didn’t find records to update it automatically become inserting operational.

**Upset operation:** This operation first executes update operation and if that fails as mentioned above then it inserts records into warehouse.

**Delete insert operation:**

This operation first executes delete operation and then inserts the new source records into data warehouse.

**Insert update operation:** This operation first executes insert operation and if that fails then it updates existing records in warehouse. Insert update operation is preferred to upset operation since it is more efficient [8].
**Bulk load operation:** Bulk load is a utility provided by major ETL vendors these days which are faster and efficient in loading huge amounts (hundred millions) of data into warehouse [8][9].
2.12 Chapter Summary:

This chapter provides an introduction about BI concept, future and benefits, BI components that provide information about data warehouse, DW architectures, DW schemes and DW objects. The chapter also provides definitions about ETL processes: Extract, Transform and Load.
*Chapter three*

*modeling & tools*
3.1 Introduction:

In this chapter discussing the different types of databases that is used in the system, which is the oracle database, and the sql server database and its main features and the advantages of each and overview information about the two different database and comparison between them.

Also discussing the visual studio 2010 that been used for building the interfaces for the desired system and its capabilities and unique features that been used in the building process.

3.2 What is Microsoft SQL Server?

Microsoft SQL Server is a relational database management system (RDBMS). Its primary query language is Transact-SQL, an implementation of the ANSI/ISO standard Structured Query Language (SQL) used by Microsoft and Sybase.

The architecture of Microsoft SQL Server is broadly divided into three components: SQL Server that implements the basic services required by SQL Server, including thread scheduling, memory management and I/O management; the Relational Engine, which implements the relational Database components including support for databases, tables, queries and stored procedures as well as implementing.

3.2.1 SQL Server license:

SQL Server is a relational database server that runs on all versions of Microsoft Windows. The server software is licensed from Sybase and there is a high degree of compatibility with large-scale Sybase servers.
3.3 SQL Server Database Engine (Server):

SQL Server's database engine is the primary server application of the SQL server package. Its main functions are:

- Provide reliable storage
- Rows of data are stored as pages, each 8kb in size.
- Provide a means to rapidly access the data, this is done through utilizing indexes both clustered and non-clustered to search for data, which removes the need for all data to be scanned from the database tables.
- Ensure consistent access to the data, Consistent access basically means only allowing one client to modify/ changing the data at any time
- Implement security, Microsoft SQL Server has multiple security levels, which is Server Level, database level and database object level. Access to the server can be controlled by a Username or password or through Windows security in LAN/Networked environment.
- Enforce data integrity, ensure the data stays consistent.\[10\]

3.4 Oracle database:

3.4.1 Introduction:

3.4.2 What does Oracle Database (Oracle DB) mean?
Oracle database (Oracle DB) is a relational database management system (RDBMS) from the Oracle Corporation. Originally developed in 1977 by Lawrence Ellison and other developers, Oracle DB is one of the most trusted and widely used relational database engines. The system is built around a relational database framework in which data objects may be directly accessed by users (or an application front end) through structured query language (SQL). Oracle is fully scalable relational database architecture and is often used by global enterprises, which manage and process data across wide and local area networks. The Oracle database has its own network component to allow communications across networks. Oracle DB is also known as Oracle RDBMS and, sometimes, just Oracle.

3.4.3 Oracle Database Architecture:

An Oracle database is a collection of data treated as a unit. The purpose of a database is to store and retrieve related information. A database server is the key to solving the problems of information management. In general, a server reliably manages a large amount of data in a multi-user environment so that many users can concurrently access the same data. All this is accomplished while delivering high performance. A database server also prevents unauthorized access and provides efficient solutions for failure recovery.

Oracle Database is the first database designed for enterprise grid computing, the most flexible and cost-effective way to manage information and applications. Enterprise grid computing creates large pools of industry-standard, modular storage and servers. There is no need for peak workloads, because capacity can be easily added or reallocated from the resource pools as needed.
The database has **logical structures** and **physical structures**. Because the physical and logical structures are separate, the physical storage of data can be managed without affecting the access to logical storage structures.

### 3.5 Visual studio definition:

Microsoft Visual Studio is an (IDE) from Microsoft. It is used to develop console and graphical user interface applications along with Windows Forms or WPF applications, web sites, web applications, and web services in both native code together with managed code for all platforms supported by Microsoft Windows, Windows Mobile, Windows CE, .NET Framework, .NET Compact Framework and Microsoft Silverlight.

### 3.6 C# Programming language:

C# is an elegant and type-safe object-oriented language that enables developers to build a variety of secure and robust applications that run on the .NET Framework. The user can use C# to create traditional Windows client applications, XML Web services, distributed components, client-server applications, database applications, and much, much more. Visual C# provides an advanced code editor, convenient user interface designers, integrated debugger, and many other tools to make it easier to develop applications in visual studio. C# syntax is highly expressive, yet it is also simple and easy to learn. The curly-brace syntax of C# will be instantly recognizable to anyone familiar with C, C++ or Java. Developers who know any of these languages are typically able to begin to work productively in C# within a very short time. The C# build process is simple compared to C and C++ and more flexible than in Java. There are no
separate header files, and no requirement that methods and types be declared in a particular order. A C# source file may define any number of classes, structs, interfaces, and events[11]

In this project, the system is built using the c# language and visual studio c# windows form to code the program that runs the system.

3.7 Hyperion reporting tools:

3.7.1 Hyperion definition:

Is a database that allows you to access the data very quickly? The company was bought by Oracle in 2007 and Oracle has continued to improve the product.

The name Hyperion is Business intelligence and targets to support better business decision-making. In addition, BI system can be called a decision support system (DSS). In addition, the term business intelligence is sometimes used as a synonym for competitive intelligence, because they both support decision-making. Hyperion has two main products, although there are many alternative software products:

- Hyperion Essbase

Is a multidimensional database management system (MDBMS) that provides a multidimensional database platform upon which to build analytic applications?

Many users work with Essbase data using as their interface an add-in for Microsoft Excel.

In 2005, Hyperion began to offer a visualization tool called Tableau under the name "Hyperion Visual Explorer". Tableau originated at Stanford University as a
government-sponsored research project to investigate new ways for users to interact with relational and OLAP database

**Essbase products:**

- Oracle Essbase Studio
- Oracle Essbase Administration Services
- Oracle Essbase Integration Services
- Oracle Essbase Provider Services
- Oracle Hyperion Smart View for Office
References:-

- [8] Ralph Kimball, Margy Ross, "The Kimball Group Reader: Relentlessly Practical Tools for Data Warehousing and Business Intelligence"

- [10] A COURSEWARE ON ETL PROCESS