Chapter 1   Introduction

1.1   Motivation and introductory to the problem

There are many challenges that face business nowadays. These challenges grow because there is a volatility of technology and rapidly increase of changes, these changes lead to increase the quality of the products and services and help in customer demands anywhere and anytime (Shaw, Robert, 1991). Every business today is targeting a specific kind of users. Therefore, every enterprise needs to deal with markets to meet customer needs and requirements. On the other hand, technology is affecting business rapidly, so it generates high competition between competitors and their products and services. Technology is growing fast, so it plays a big role in moving business to highest degree of productivity. The first things that business process depends on is customer, and this dependability can lead to success or failure. One of the most important things that focus on customer is the Customer Relationship Management (CRM), which started in the early 1970s, when customer satisfaction was evaluated using annual surveys or by front-line asking (History of CRM Software, 2017). At that time, businesses had to rely on standalone mainframe systems to automate sales, but the extent of technology allowed them to categorize customers in spreadsheets and lists.

Recently customer relationship management systems become limited and has many problems. At the same time competition becomes very strong between competitors, so they need much more advance to take deep inside understanding of customer needs (Mithas, Sunil., 2005). Because of these changes, the world prefers to deal with Business Intelligence (BI) to meet customer needs. BI is a group of skills and technologies that help in develop and success business. One of the most important things that BI is based on is Data Ware House (DW). It is a framework that is based on Online Analytical Processing (OLAP) technology that help in providing solution in different dimensions and facilitate in empowering commercial business and make it self-service. However, data warehouse is a big environment, which has many techniques. There are small versions of DW called data mart. Data mart build to meet a special purpose. In addition, it focuses on specific target and implementing all practices of data warehouse on it.

Customer is one of the important factors in business process (Mithas, Sunil, 2005). The customer demands are the main things that affecting the market and this effect can lead to success or failure in business (Tavana, Ali Feizbakhsh, 2013), so the customer value needs

to be measured to found these successes and failures. Also customer has a big role in increasing the incomes of the enterprise and making business profitable. The focusing on customer have to be in multidimensional. The important level in customer value development (CVD) is the lifecycle of the customer in the enterprise. Therefore, the value of the customers is very crucial for business success and profit. As well, the customer can be categorized according to the length of time with the company and to the nature of the customer relationship with the enterprise. There are six cycles for customer, begin by prospective customers, new customers, active customers, repeat or loyal customers, lapsed customers and inactive or abandoned customers. Business Intelligence is depending on data warehouse in all processes. OLAP technology used to facilitate data analysis in multidimensional view. DW is a system used for reporting and data analysis, and is considered a core component of business intelligence. DWs are central repositories of integrated data from one or more disparate sources. A data mart is low level of data warehouse environment that is used to get data out to the users (Ranjan, Jayanthi, 2009). The data mart is a subset of the data warehouse and is usually oriented to a specific business and has a special purpose. On customer's side, it is important to make data mart specifically for them because of their needs and demands.

There are many limitations facing the business and these limitations focused on the data sources which have been restricted. Also limitation of the data entry and sources system capability in capturing information, also there are limitations in external data integration and in understanding to the targeted customer.

1.2 Problem Statement

Technology is developing rapidly, so its impacting all business and generate a high competition all over the world. Enterprise should measure the customer value, deal with customer lifecycle, and develop customer value to decrease the limitation of the resources. Current customer relationship management look at the customer in single dimension. There is gloomy understanding and controlling of customer's revenue in enterprise and dealing with transactional systems limit the agility in responding to quick market.BI comes to enhance the customer value development (CVD). DWH is one of the Business Intelligence technologies that based on it. Also it provides a data mart and it is a small version of data warehouse that has a special purpose and focusing in a specific area.

1.3 Importance

Measuring customer value and deal with customer in different views to serve customer better and to generate more revenue.

1.4 Research Questions

- How to measure the customer value?
- How to consider customer in multidimensional view?

1.5 Objectives

- To design or build Data mart for special purpose that allow measuring the customer value.
- To measure the customer value in multidimensional view and from different perspective.

1.6 Methodology

The research will define the problem and the literature review firstly, after that declares the method development. In the next phase, describe the data collection and data analysis. Finally, the research results and evaluations.

1.7 Research Scope

The research focus on the financial and banking services. The research work in the customer value domain and measuring these values and loyalty. The research deal with customer's data, which belong to a specific local bank. The outcomes from this research is to deal with customers in multidimensional view.

1.8 Research Structure

This research contains five chapters. Chapter one is introduction; this chapter contains motivation and introductory to the problem, problem statement, importance, related works, hypothesis, objectives, research methodology and research scope. Chapter two is research background and literature review; this chapter will contain literature review about CRM, BI and DW. Chapter three is methodology; this chapter will contain operational framework, research design, research type, case study, data collection and data analysis. Chapter four is Results; this chapter will contain proposed solution and the results. Chapter five is conclusion; this chapter will contain conclusion and future work. Finally, the references.

Chapter 2   Research Background and Literature Review

2.1 Introduction

This chapter describes the literature review of the technology development and the impact of this technology in the business process. In the next section, the chapter discusses the role of technology and its competitors in business competitions followed by the customer value development (CVD) section, the limitation of the customer relationship management (CRM) section, business intelligence techniques, data warehouses and data marts. In addition, it discusses the multidimensional systems and demonstrates multiple dimension model. Therefore, the chapter explains the literature review evolution and in the last section the chapter has been summarized.

2.2   Roles of technology and competitions in business

In these days, technology becomes the most important things all over the world. It helps people in different fields and enable them to work effectively. Also it makes tasks and activities to complete efficiently and produce high quality products. Furthermore, technology has grown fast, so it generates many challenges in business. Business becomes depending on internet based computing and network. So E-business able customers to gain global visibility and help them respond quickly to their demands and to resource shortages. Many dimensions can be examined to drive E-business to success such as information integration, synchronized planning, coordinated workflow and new business model (Lee HL, Whang S., 2004).On the other hand, technology lead to have a huge amount of knowledge, and this knowledge play a big role in business success or failure.

Nowadays there is a big competition between competitors. This competition helps customer to deliver their demands properly and in a good manner. Driven by customers that are more demanding, global competition and slow-growth economics and industries. Big organizations search for new ways to achieve and retain a competitive advantage (Woodruff RB., 1997) and to achieve greatest amount of customer demands and needs.

In addition, business depends on customer as the main factor in the business process, and the organization must focus on measuring the customer value using best solutions and services. Also all organizations must look to the customer in a multidimensional view, to

enable the organizations to deliver high quality products. In an effort to improve company operations and their results, more firms are applying the principle of "lean", not only to manufacturing but also to systems engineering process. Lean creates new problems and tensions and may not deliver expected results. Lean is not just about minimizing cost, cycle time; it is also about maximizing value (Browning TR., 2003). Also there is an investigation which has been done and it reports studies lead to the development and validation of a customer value co-creation behavior scale. The scale comprises two dimensions: customer participation behavior and customer citizenship behavior, with each dimension having four components. The elements of customer participation behavior include information seeking, information sharing, responsible behavior, and personal interaction, whereas the aspects of customer citizenship behavior are feedback, advocacy, helping, and tolerance. The scale is multidimensional and hierarchical, and it exhibits internal consistency reliability, construct validity, and nomological validity (Yi Y, Gong T., 2013) .

2.3   Customer value

Customer Value Development (CVD) is one of the important things in business. Business Intelligence plays a big role to enhance the CVD. The customer value knowledge gives targeted data for personalized markets, generating customer relationship management strategy that aims enterprises to identify and segment customers and create long-term relationship with customers and they get maximum customer lifetime value (R. qiasi,2012). Web-based customer support system provides services for customers, the quality of systems measured by the information that is provided by the systems. A lot of user's data in internet were used to test relationships between customers and enterprises. The results of study indicate that information and system quality determine effectiveness while service quality has no impact (Negash S, Ryan T, Igbaria M., 2003) .

2.4   Customer Relationship Management

Is   an   approach   to   manage   a   company's   interaction   with   current   and potential customers. It uses data analysis about customers' history with a company and to improve   business   relationships   with   customers,   specifically   focusing   on customer retention and ultimately driving sales growth (Bain, 2015).

One   important   aspect   of   the   CRM   approach   is   the   systems   of   CRM   that compile data from a range of different communication channels, including a company's

website, telephone, email, live chat, marketing materials, and more recently, social media(Shaw, Robert1991).

2.4.1 Categories of customer relationship management

There are four categories of customer relationship management and they are Strategic CRM, Operational CRM, Analytical CRM and Collaborative and the definition of every type is shown below:

- **Strategic:** Strategic CRM is focused upon the development of a customer-centric business culture (Buttle, Francis; Maklan, Stan ,2015).

- **Operational:** The primary goal of customer relationship management systems is to integrate and automate sales, marketing, and customer support. Therefore, these systems typically have a dashboard that gives an overall view of the three functions on a single customer view, a single page for each customer that a company may have. The dashboard may provide client information, past sales, previous marketing efforts, and more, summarizing all of the relationships between the customer and the firm. Operational CRM is made up of 3 main components: sales force automation, marketing automation, and service automation ("Types of CRM and Examples | CRM Software",2015).

- **Analytical:** The role of analytical CRM systems is to analyze customer data collected through multiple sources, and present it so that business managers can make more informed decisions. Analytical CRM systems use techniques such as data mining, correlation, and pattern recognition to analyze the customer data. These analytics help improve customer service by finding small problems which can be solved, perhaps, by marketing to different parts of a consumer audience differently ("Types of CRM and Examples | CRM Software", 2015).

- **Collaborative:** The third primary aim of CRM systems is to incorporate external stakeholders such as suppliers, vendors, and distributors, and share customer information across organizations. For example, feedback can be collected from technical support call, which could help provide direction for marketing products and services to that particular customer in the future (Tavana, Ali Feizbakhsh.; Fili, Saeed.; Tohidy, Alireza.; Vaghari, Reza. &Kakouie, Saed, 2013).

2.4.2 Components of Customer Relationship Management

Customer satisfaction has important implications for the economic performance of firms because it has the ability to increase customer loyalty and usage behavior and reduce customer complaints and the likelihood of customer defection (Bolton, Ruth N. ,1998) .The implementation of a CRM approach is likely to have an effect on customer satisfaction and customer knowledge for a variety of different reasons(Fornell, Claes,1992).

CRM systems include Data warehouse technology, Opportunity management, network, software as a service (SaaS), e-commerce, Customer-centric relationship management (CCRM) and Systems for non-profit and membership-based organizations.

Firstly, firms are able to customize their offerings for each customer. By accumulating information across customer interactions and processing this information to discover hidden patterns, CRM applications help firms customize their offerings to suit the individual tastes of their customers. This customization enhances the perceived quality of products and services from a customer's viewpoint, and because perceived quality is a determinant of customer satisfaction, it follows that CRM applications indirectly affect customer satisfaction. CRM applications also enable firms to provide timely, accurate processing of customer orders and requests and the ongoing management of customer accounts (Sunil.; Krishnan, M.S. &Fornell, Claes ,2005).

There are many problems that facing customer relationship management (CRM). These limitations affecting the customer value and decreasing the loyalty. These limitations found it in data quality (dirty data). The limitation has different ways such as Poor data entry, Data missing from database field, Lack of company –wide or industry-wide data coding standard, Multiple database scattered throughout different organization, Older systems that contain poorly documented or obsolete data. the solution for these problems is data cleaning or scrubbing and to bring consistency. Each example of integrating data from legacy systems show that it is not always possible to identify data quality problems. (Reid A, Catterall M., 2005). Mobile phones systems containing many different features and characteristics. some of these features are Mobility, Safety, Price, Privacy. E-CRM is considered as a technique used to serve customer. The result of E-CRM performance was positively related to customer loyalty and E-CRM performance explains a large percent of the variation in customer loyalty(Azila N, NoorNeeraj M., 2011).

2.5　Business Intelligence

Business Intelligence (BI) is a group of strategies and technologies that used by enterprises to make data analysis for business data and information (Dedić N. &StanierC. ,2016). BI technologies provide different views such as historical, current and predictive views of business operations. The common functions of business intelligence technologies that used is reporting, online analytical processing, analytics, data mining, process mining, complex event processing, business performance management, benchmarking, text mining, predictive analytics and prescriptive analytics and the enterprises use their results for decision making. BI technologies can handle large amounts of structured and sometimes unstructured data to help identify, develop and otherwise create new strategic business opportunities. They aim to allow for the easy interpretation of these big data. Identifying new opportunities and implementing an effective strategy based on insights can provide businesses with a competitive market advantage and long-term stability (Rud, Olivia ,2009).

Business Intelligence becomes one the new technologies that enhancing the customer value. BI concept has two different meaning, the first one look to intelligence as human intelligence capacity and human cognitive faculties, the second one look to intelligence as information valued, expert information, technologies and knowledge. Business intelligence strategies are collecting data, analyzing data and generating reports. Also BI has depth knowledge about some factors such as customer, competitors, business partners, economic environment and internal operation (RanjanJ., 2009).Also BI defined as a process for extracting transforming, managing and analyzing large data by mathematical model to gain knowledge to make decision. BI contains many elements such as data warehouse, data mining and decision support system (Fitriana R, Eriyatno TD, Djatna T., 2011). On the other hand, BI as implementing in enterprises is defined as an umbrella term that companies, architecture, tools, database, analytical tools, application and methodologies. CRM and ERP systems can be used in BI because of their domains and flexibility. BI solution composed from several component as resource of data, database, connectivity tools, data transformation tools, analytical tools, visualization tools, BI front end tools and measures of business impact (Horakova M, Skalska H. , 2013).

2.5.1 Technologies of Business Intelligence

There are many technologies in business intelligence which are concluded as follow (Ranjan, Jayanthi, 2009):

- **OLAP (On-line analytical processing)**

    It refers to the way in which business users can slice and dice their way through data using sophisticated tools that allow for the navigation of dimensions such as time or hierarchies. Online Analytical Processing or OLAP provides multidimensional approach and it will be discussed in section 2.6 and section 2.7, summarized views of business data and is used for reporting, analysis, modeling and planning for optimizing the business.

- **Advance Analytics**

    It is referred to as data mining, forecasting or predictive analytics, this takes advantage of statistical analysis techniques to predict or provide certainty measures on facts (Ranjan, Jayanthi, 2009).

- **Corporate Performance Management (Portals, Scorecards, Dashboards)**

    This general category usually provides a container for several pieces to plug into so that the aggregate tells a story. For example, a balanced scorecard that displays port lets for financial metrics combined with say organizational learning and growth metrics.

- **Real time BI**

    It allows for the real time distribution of metrics through email, messaging systems and/or interactive displays.

- **Data Sources**

    Data sources can be operational databases, historical data, external data for example, from market research companies or from the Internet), or information from the already existing data warehouse environment.

- **Data Warehouse (DWH)**

    The data warehouse is the significant component of business intelligence. It is subject oriented, integrated. The data warehouse supports the physical propagation of data by handling the numerous enterprise records for integration, cleansing, aggregation and query tasks (Ranjan, Jayanthi, 2009).

2.5.2 Data Warehouse Design methods

There are three types of methods for designing and they have been concluded as follow:

- **Bottom-up design**: In the bottom-up approach, data marts are first created to provide reporting and analytical capabilities for specific business processes. These data marts can then be integrated to create a comprehensive data warehouse. The data warehouse bus architecture is primarily an implementation of "the bus", a collection of conformed dimensions and conformed facts, which are dimensions that are shared (in a specific way) between facts in two or more data marts.

- **Top-down design**: The top-down approach is designed using a normalized enterprise data model. "Atomic" data, that is, data at the greatest level of detail, are stored in the data warehouse. Dimensional data marts containing data needed for specific business processes or specific departments are created from the data warehouse.

- **Hybrid design**: Data warehouses (DWH) often resemble the hub and spokes architecture. Legacy systems feeding the warehouse often include customer relationship management and enterprise resource planning, generating large amounts of data. To consolidate these various data models, and facilitate the extract transform load process, data warehouses often make use of an operational data store, the information from which is parsed into the actual DW. To reduce data redundancy, larger systems often store the data in a normalized way. Data marts for specific reports can then be built on top of the data warehouse.

2.5.3 Data Warehouse Characteristics

There are basic features that define the data in the data warehouse that include subject orientation, data integration, time-variant, nonvolatile data, and data granularity (Paulraj., Ponniah, 2010)(H., Inmon, William , 2005).

- **Subject-Oriented**: Unlike the operational systems, the data in the data warehouse revolves around subjects of the enterprise (database normalization). Subject orientation can be really useful for decision making.

- **Integrated**: The data found within the data warehouse is integrated. Since it comes from several operational systems, all inconsistencies must be removed. Consistencies include naming conventions, measurement of variables, encoding structures, physical attributes of data, and so forth.

- **Time-Variant**: While operational systems reflect current values as they support day-to-day operations, data warehouse data represents data over a long time horizon (up to 10 years) which means it stores historical data. It is mainly meant for data mining and forecasting, if a user is searching for a buying pattern of a specific customer, the user needs to look at data on the current and past purchases.

- **Nonvolatile**: The data in the data warehouse is read-only which means it cannot be updated, created, or deleted.

- **Summarized**: In the data warehouse, data is summarized at different levels.

- The user may start looking at the total sale units of a product in an entire region. Then the user looks at the states in that region. Finally, they may examine the individual stores in a certain state. Therefore, typically, the analysis starts at a higher level and moves down to lower levels of details.

2.5.4 Data Warehouse Architecture

The different methods used to construct/organize a data warehouse specified by an organization are numerous. The hardware utilized, software created and data resources specifically required for the correct functionality of a data warehouse are the main components of the data warehouse architecture. All data warehouses have multiple phases in which the requirements of the organization are modified and fine-tuned (Gupta, SatinderBal; Mittal, Aditya, 2009).

2.5.5 Evolution in organization use

These terms refer to the level of sophistication of a data warehouse:

- **Offline operational data warehouse**: Data warehouses in this stage of evolution are updated on a regular time cycle (usually daily, weekly or monthly) from the operational systems and the data is stored in an integrated reporting-oriented data.

- **Offline data warehouse**: Data warehouses at this stage are updated from data in the operational systems on a regular basis and the data warehouse data are stored in a data structure designed to facilitate reporting.

- **On time data warehouse**: Online Integrated Data Warehousing represent the real time Data warehouses stage data in the warehouse is updated for every transaction performed on the source data.

- **Integrated data warehouse**: These data warehouses assemble data from different areas of business, so users can look up the information they need across other systems.

- **Data Mart**

As described by (Inmon, 1999) data martis a collection of subject areas organized for decision support based on the needs of a given department. Finance has their data mart, marketing has theirs, and sales have theirs and so on. And the data mart for marketing only faintly resembles anyone else's data mart.

2.5.6 Benefits of Business Intelligence

BI provides many benefits to companies utilizing it. It can eliminate a lot of the guesswork within an organization, enhance communication among departments while coordinating activities, and enable companies to respond quickly to changes in financial conditions, customer preferences, and supply chain operations. BI improves the overall performance of the company using it.

2.6 Multidimensional Recommender Systems

There are many multidimensional systems in e-commerce field; these systems help companies to develop the relationship with their customers. The customers deal with these companies over different marketing channels such as e-mail, web and mobile communication. Most of old research focused on recommending items for a specific users or a group of users. The mechanisms that used in the past for recommending items are collaborative filtering, content-based filtering or a combination of the two techniques (Adomavicius G, Tuzhilin A., 2001).

In many e-commerce systems such as packages, restaurants or web contents, it may not be sufficient to recommend items to users or users to items only .So the customer can recommend a specific item in a specific time of year, and that makes the time becomes the third party of the process and this process lead to make multidimensional in the system .Also the place can be the fourth dimension in the system, so the customer can recommend the item in a particular place such as (go to Caribbean in winter, but not in summer) and the relation will be looks like (customer, item, time and place).So recommender systems supposed to support multiple dimensions .

Many applications may not recommend individual items to individual users such as watching action movies for college students in movies system, so they apply one to many

technique (one item to multiple users). In multiple dimension it proposed to support aggregation hierarchies for various dimension and provide recommendation capabilities at various level of aggregation. Some of existing recommender systems support profiles of users and items. Therefore, Multidimensional systems need extensive profiling capabilities and proposed to make these features support these systems.

Traditional two-dimensional recommender systems usually provide one or two types only and recommend N items to one user or M users to one item. So these types of multidimensional systems are "hand wired" in recommender engines by a software vendor. Therefore, Theses systems need more flexibility and extensive features in capabilities, for example preparing top three action movies are not longer than two hours, and to people who like action movies only. So to apply flexibility and extensive features a new recommendation query language (RQL) has been provided to allow users to express complex recommendations that take multiple dimensions, aggregation hierarchies and extensive profiling information.

All features that mentioned above are not stand alone but comprised of a multidimensional recommendation model and integrated into one approach. This approach based on data warehousing paradigm because it makes support for multiple dimensions and hierarchies provide recommendation capabilities at multiple aggregation levels and example for that OLAP-enabled data warehousing systems.

Multidimensional recommender systems have two components: the first component is recommender engine that estimates new recommendation ratings based on the already available set of ratings, and the second component is query engine that allows the users to express different types of recommendations "on the fly" in a flexible manner.

2.7 Multidimensional Recommendation Model

It a traditionally, collaborative, content-based, and hybrid recommender systems deal with applications that have two types of entities (users , items)(G. Adomavicius and A. Tuzhilin, 2001) .this model is used to provide recommendations, set of rating specifying how users liked items, example for that x-person like watching action movie and assign the rate of the movie 7 out of13   (set Rmovie(x-person, action movie) = 7).these ratings help recommender systems to determine the ratings of the items yet unspecified by the users by estimating the rating function R :

$$R : Users \times Items \rightarrow Ratings .$$

And it for the (user, item) pairs that have not been rated yet. So in traditional two-dimensional recommender systems, a rating function can be implemented in a matrix rating such as $R(i, j)$ of item j by user i. In some e-commerce and web content applications traditional two dimensional approach user/item do not fit well with their systems, so they need extra or multiple dimensions such as time place and other dimensions to model their inherent additional complexities.

2.7.1 Multiple Dimensions

This approach is extended form traditional-two dimensional approach. In the two-dimensional approach the function it was looks like S = Users × Items, so in multidimensional approach the functions becomes looks like S = Users × Content × Time and this function described the web content application, where a user assigns a score to a content seen at a certain time (G. Adomavicius and A. Tuzhilin, 2001). for example x-person assigned rating 8 to the stock reports on weekday evenings, Rcontent(x-person, "Stock Report", weekday evening) = 8. Other examples of multidimensional applications include recommending food to customers in restaurants, i.e., S = Users ×Restaurant ×Food, and purchasing agent applications − recommending the products to buy to customers, including when and where, i.e., S = Users × Products × Time × Place. So traditional two dimensional approach provide a recommendation only for one type per dimension such as (recommend top N items to a user) or (recommend item to M users) .but multiple dimension can offer many more possibilities , so in web content application case ,the function could ask for the "top N content items for each user/time combination" or the "top N times for each user/item combination," or the "top N user/time combinations for each item." Therefore, applying multiple dimension approach in recommender systems this allow to apply recommender technologies in a much more diverse set of applications than the traditional two-dimensional recommender systems and to obtain new types of recommendations.

R(101,7,1) = 6

6

User

Id	Name	Age
101	John	25
102	Bob	18
103	Alice	27
104	Mary	21

101
102
103
104

2   3   5   7

3
2
1

Time

Item

Id	Name	Cost
2	AB17	250.00
3	AB23	299.95
5	XY70	150.00
7	ZZ55	115.50

Id	Name
1	Weekday
2	Weekend
3	Holiday
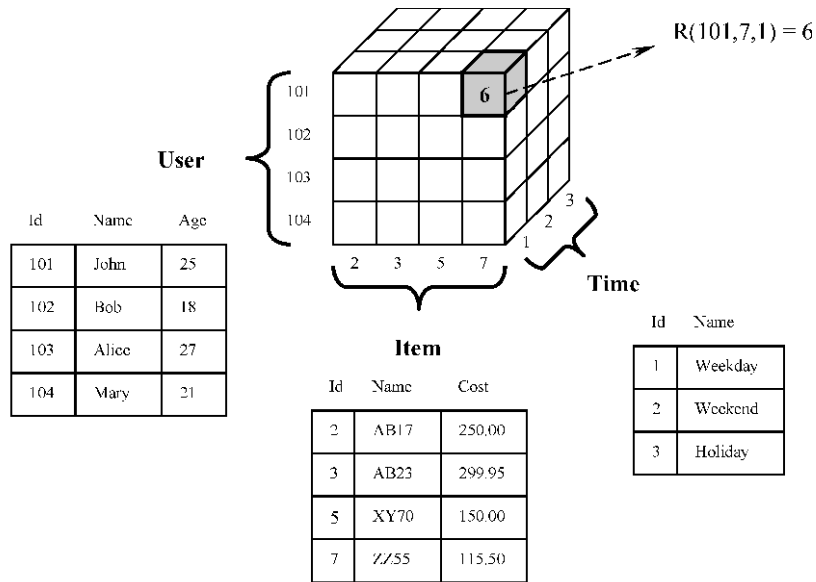
Figure 2.1 Multiple dimensions' approach

2.7.2 Profiling Capabilities

It is an approach that enables the recommender system to make profile information for every dimension in the system. In recommender system that apply a two-dimensional matrix of ratings it provides very limited profiling capabilities because unknown rating estimations were based only on the known ratings, and neither user nor item profiles have been used for this purpose (G. Adomavicius and A. Tuzhilin, 2001).

Profiling capabilities propose to be extended from recommender systems and view it as a recommendation warehouse consisting of multidimensional cubes for storing the ratings. So it provides profiles representing elements of each dimension. These profiles can contain a set of various attributes describing each dimension. For example, for the "user" dimension, a profile may include attributes such as the name, the address and the phone number of each user, as well as his/her preferences (e.g. a favorite drink), and behavioral characteristics, such as the largest purchase made at a Web site. In web content case a profile may include information about the Web content type (e.g., politics, finance, weather, sports, science), the length of the content item (e.g., how long is the news article), and the presence or absence of the important keywords in the content item. The profile can be saved as records in relational tables. Each profile stored in one record. Each dimension in a separate table. So profiling approach provides more complex recommendations.

2.7.3 Aggregation Capabilities:

These OLAP-like aggregation capabilities have been used in e-commerce and web content applications, and have not been utilized in recommender systems. So various dimensions may have hierarchies associated with these capabilities. For example, Products dimension has standard industrial product hierarchy and Time dimension has a temporal hierarchy, e.g., minutes, hours, days, months, seasons, etc (G. Adomavicius and A. Tuzhilin, 2001).

With these hierarchies, recommender systems can provide complex recommendations not only for individual items but also for group of items or other dimensions. For example, there is not only one user would like to watch one movie only Rmovie(x-person, Gladiator) = 7, but also how they may like categories of movies Rmovie(x-person, action movies) = 5. Also we can grouping the users such as Rmovie(graduate students, Gladiator) = 9.

Generally, if there is individual rating in the multidimensional cube, hierarchies can be used to compute aggregated ratings, for example if there are some movies have been grouped according to their type and x-person preferring each action movie individually, then as shown in figure (2.3) below it becomes easy to compute an overall rating of how x-person likes action movies by aggregating his individual action movie ratings.

Rmovie(x-person, action) := AGGRx.genre=actionRmovie(x-person, x) .
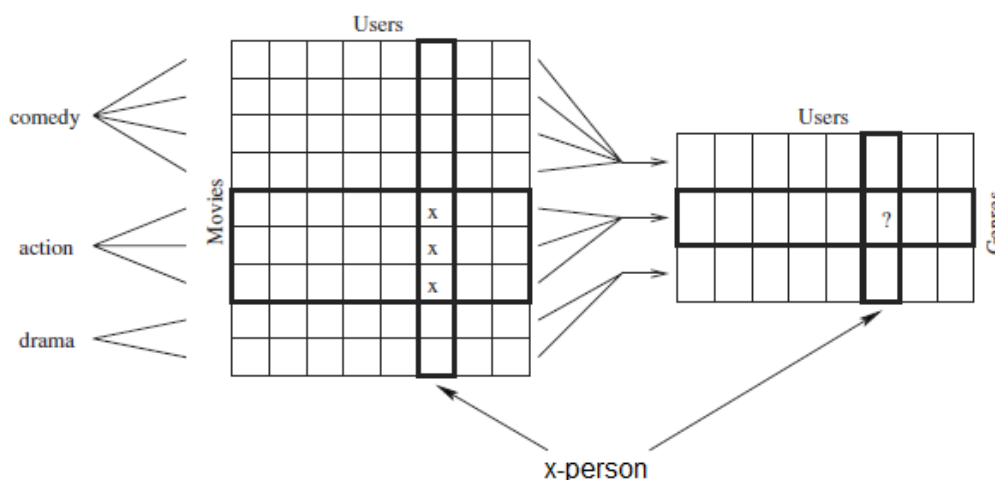
Figure 2.2Aggregation capabilities for recommender systems: aggregating the ratings

Many traditional OLAP systems use Aggregation operation and it known as roll-up and it is a simple summation of all elements in the system. But this approach is not applicable with recommender systems because ratings usually are not additive quantities. So recommender systems need more appropriate aggregation functions AGGR such as AVG, MAX, MIN, and AVG of TOP k. For example, the cumulative rating of action movies that can be computed for x-person is look like:

$$Rmovie(x\text{-}person, action) := AVGx.genre=actionRmovie(x\text{-}person, x) .$$

2.7.4 Estimating the Ratings

One of the important things in recommender system is how to estimate the rating in multidimensional systems recommendation space. There are many methods for estimating the rating in traditional two dimensional recommender systems and have been classified to into three broad categories: collaborative, content-based, and hybrid. These methods cannot be extended to multidimensional systems directly. These methods depending on a combines of information about users and items into a single hierarchical regression-based Bayesian preference model that uses Markov chain Monte Carlo techniques for exact estimation and prediction. This Bayesian preference model allows statistical integration of the following types of information that can be useful for making recommendations of items to users: a person's expressed preferences (ratings), preferences of other consumers, expert evaluations, item characteristics, and characteristics of individuals. Example for that in movie case the information can be movie ratings, gender, age of users, movie genres and movie reviews by critics. So the approach of rating that apply in traditional two dimensional recommender system can be directly extended to combine information about more than two dimensions by letting regressions and considering additional variables or attributes that describe characteristics of other dimensions, not only for users and items (G. Adomavicius and A. Tuzhilin, 2001).

Estimating the rating function depend on the multidimensional system model and their information. So different rating estimation functions may work better in different applications. And it is difficult to provide a single one-size-fits-all function for all possible applications. But domain expert can select a specific rating estimation function (2) that is most suitable for the application at hand. And that can be achieved by providing the DEFINE ESTIMATOR and the ESTIMATE commands as a part of the data definition component of

the Recommendation Query Language (RQL). These ESTIMATE commands are specified by the user, system administrator and determine a particular rating estimation function for an application at hand (DEFINE ESTIMATOR command) as well as computing the actual rating (ESTIMATE command).

2.7.5 Recommendation Query Language (RQL):

Traditional two – dimensional techniques become not suitable for recommender systems. So multiple diminutions come to provide more complex types of recommendations. This issue motivates the need for a flexible query language that allows users to express ad hoc recommendations in the same manner as SQL allows expressing ad hoc database queries.

- **Data Definition Language:**

The purpose of this language in RQL is to define the various components of a recommendation warehouse, such as dimensions, cubes, and rating estimation methods. the dimensions of recommendation such as   User, Product, and Time are defined in RQL with the DEFINE DIMENSION command.   In addition, multidimensional cubes of ratings are defined in RQL with DEFINE CUBE. Example of defining a traditional movie recommendation application, consisting of two dimensions and a matrix of ratings in RQL can be as follows (G. Adomavicius and A. Tuzhilin, 2001):

DEFINE DIMENSION User (UserId, LastName, FirstName, Gender, Age)
DEFINE DIMENSION Movie (MovieId, Title, Genre, Length, Director, Year)
DEFINE CUBE MovieRecommender(User, Movie) WITH MEASURES (Rating)

Each dimension represented by a name and is described with a list of attributes characterizing each element of the dimension. Data cube is defined using DEFINE CUBE command that utilizes previously defined Dimensions (User and Movie in the above example) and uses the Rating measure specified with the WITH MEASURES clause. Most of the traditional two-dimension recommender systems use single measures, but in multidimensional recommender systems multiple measures can be used .so in some applications like recommending a restaurant the rates can specified according to four criteria: food, decor, service, and cost. This is an example of the data warehouse for a vacation recommender system that is described with more than two dimensions and more than one measure and the definition of the dimension and the cube will be like:

DEFINE DIMENSION Customer (CustId, LastName, FirstName, Gender, Age)
DEFINE DIMENSION Vacation (VacationId, Destination, Length, Price)
DEFINE DIMENSION Time (TimeId, Month, Season)
DEFINE CUBE VacationRecommender(Customer, Vacation, Time)
WITH MEASURES (Rating, Profit)

The DEFINE DIMENSION and DEFINE CUBE commands demonstrate the structure of a recommendation warehouse. Therefore, the recommendation warehouse has to support various rating insertion and estimation methods. In addition, user ratings are inserted into the warehouse using INSERT and LOAD commands. And the function for estimating the rating is defined with DEFINE ESTIMATOR command.

2.8 Literature Review Evolution

The evolution of literature review started by the technology and the competition across business and the challenges in customer relationship management. After that describing the business intelligence techniques and the multidimensional systems. Finally, the development of data mining techniques, tasks, applications, importance and tasks.

2.9 Summary

The chapter discuss the limitation of views in enterprises for customers and describe the multidimensional systems.

In the next chapter, we will describe the Research methodology, the way of collecting the data, data analysis and implementation.

Chapter 3　　　　　Research Methodology

3.1 Introduction

This chapter describes the research methodology and structure. It describes the research framework in the next section followed by the research design section. Then it explains the interpretive research approach and the research case study. Also the chapter discusses the data sources. In the last section, we summarize the chapter.

3.2 Operational Framework

Operational Framework				
		Objectives	**Activities**	**Deliverables**
Problem Definition and literature Review	Phase 1 :	- To identify the developing of technology and competitions and their impact in business. - To identify Customer Relationship Management limitation - To identify business intelligence technologies, data warehouse methods and data marts.	- Literature review	- Analysis of systematic literature review results - Problem identification - Research questions - Research objectives - Research scope
Method Development	Phase 2 :	- **RQ1: How to measure customer value?** - **RQ2: How to consider customer in multidimensional view?**	- To propose multidimensional model - To test and validate the proposed model - To apply the proposed model - To propose the mechanism of implementing the model - To use the mechanism to generate accurate results. - To select case study - To analyze the case study to develop the model	- Declare the research type - Declare case study - Declare the data collection - Declare the data analysis - Declare the research design	- Clear vision for building and implementing multidimensional model

Phase	Objectives	Activities	Output
Phase 3: Data Collection and multidimensional model	- To collect research data - To make extract, transformation and load - To build multidimensional model	- Set the data collection criteria. - Execute case study iteration to collect the data - Build model from collected data.	- Model Results
Phase 4 : Evaluation and Validation	- To evaluate the proposed method	- Declare the way of evaluation - Evaluate the results and see impacts - Identify strength of the proposed method	- Evaluation Results
Phase 5 : Result and Discussion	- To compile research results with regard to the research questions and objectives and define future work.	- Compile the research result with regard to the research questions and objectives - Define future work	- Research results and future work

Table 3.1 operational framework

3.3 Research Design
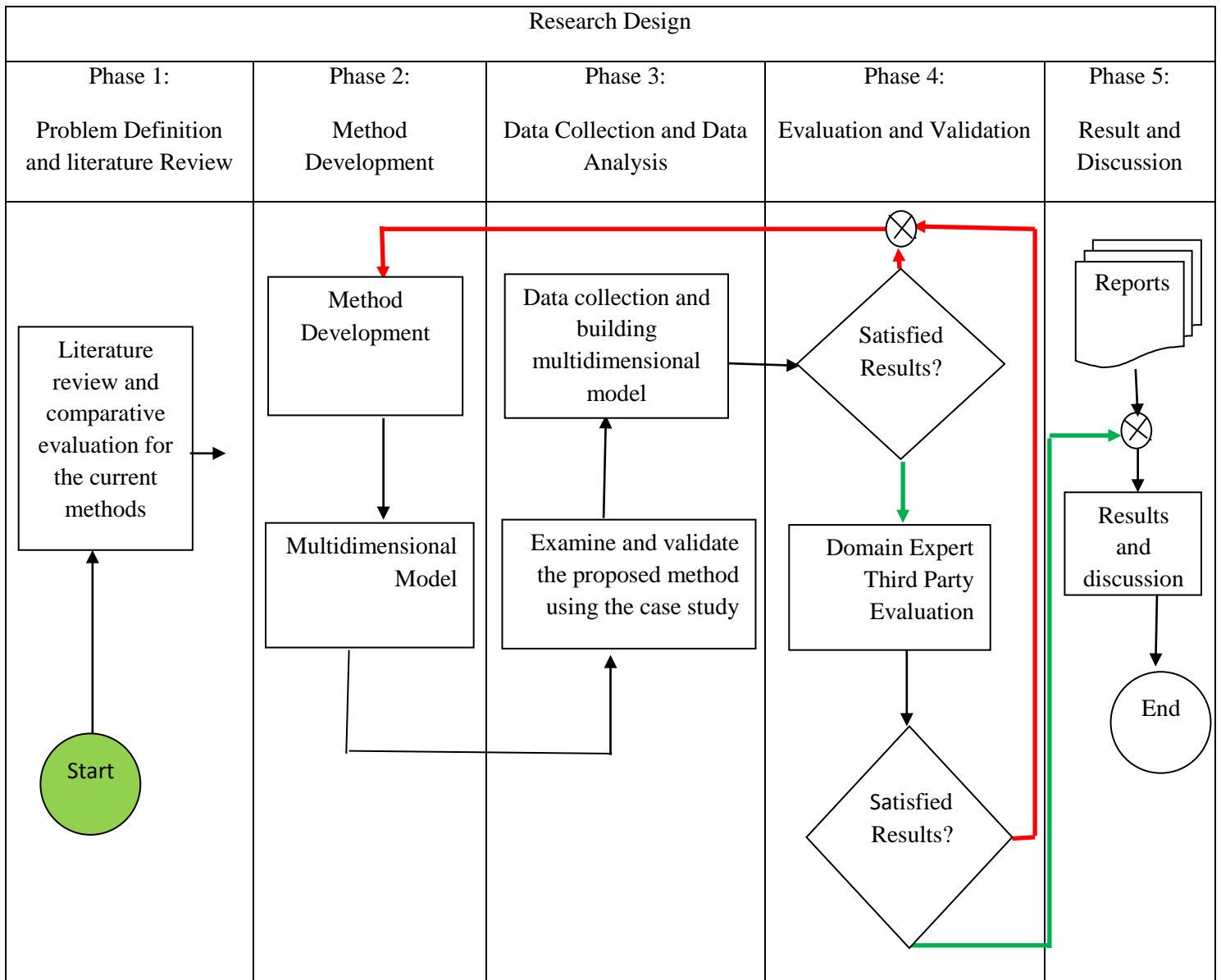
Figure 3.1 Research design

3.4 The interpretive research approach

The research classified as a quantitative research because the value of customer will increase after creating the multidimensional model. The result will reflect how is the value has increased and the parts that will serve. The research depends on dividing data warehouse to multiple data marts to generate more dimensions, which help in generating more BI queries to help in decision-making.

We use this approach to quantify data and generalize results from a sample to the population of interest and to measure the incidence of various views and opinions in a chosen sample. Usually as a sample the approach, use a large number of cases representing the population of interest. Randomly selected respondents. As a data collection the approach use a structured techniques such as online questionnaires, on-street or telephone interviews. In data analysis, Statistical data is usually used in the form of tabulations (tabs). Findings are conclusive and usually descriptive in nature. Finally, the outcome is used to recommend a final course of action.

3.5 Theoretical Multidimensional Model

We focus in our research on OLAP systems and in this case the main target is to implement a multidimensional recommendation model (shown in Figure 3.1) and it can be implemented in one of the following ways. First, as in the case of MOLAP systems, it can support proprietary data structures for cube storage and proprietary RQL query processing methods. Furthermore, MOLAP systems have many features such as (G. Adomavicius and A. Tuzhilin, 2001):

- MOLAP extends OLAP functionality to MDBMS.
- Best suited to manage, store and analyze multidimensional data.
- Proprietary techniques used in MDBMS.
- MDBMS and users visualize the stored data as a 3-Dimensional Cube i.e Data Cube.
- MOLAP Databases are known to be much faster than the ROLAP counter parts.
- Data cubes are held in memory called "Cube Cache"

Alternatively, as in the case of ROLAP systems and also they have many features and they are:

- Provides functionality by using relational databases and relational query tools to store and analyze multidimensional data.

- Build on existing relational technologies and represent extension to all those companies who already used RDBMS.
- Multidimensional data schema support within the RDBMS.
- Data access language and query performance are optimized for multidimensional data.
- Support for very large databases.

On the other hand, the multidimensional recommendation model can be implemented via two techniques the relational data model and SQL.

In this research, we will use ROLAP-based approach to test our ideas and methods because its maps the data model into the relational model, and map our multidimensional recommendation model into the relational data model and the Recommendation Query Language into SQL. It also made the implementation of our prototype multidimensional recommender system simpler than if we had to design and implement our own system for data storage and management.

To implement the system, we use two components, the first one is recommendation engine that populates the recommendation cube with ratings, and we can use various multidimensional rating estimation methods. The second component is query engine and we use this component to translate RQL queries into corresponding SQL queries and run them in the relational database.
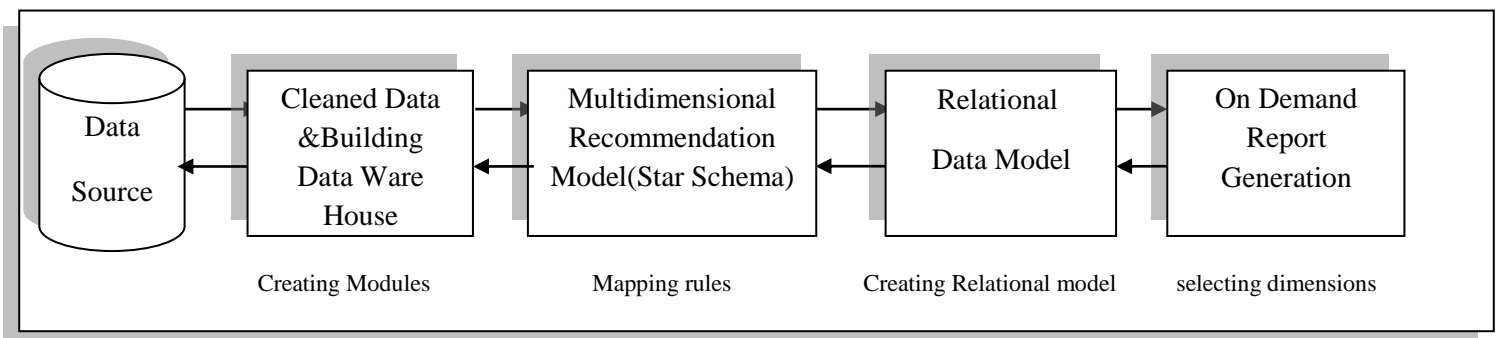
Figure 3.2 Implementing Multidimensional model

3.6 Case Study

3.6.1 Environment

The data collected from different databases and with different languages (English – Arabic). The data can be used in another environment but it needs to make some operations on it such as data integration, data cleaning and data transformation.

3.6.2   Background Information

This case study is targeting the customers of the enterprises to evaluate the customer value and loyalty. It belongs to a specific Sudanese bank and it demonstrates the financial information that belongs to this bank.

3.6.3   About the Client

The customer that has been targeted is the bank clients with their formal information, finance information, insurance information and durations information.

3.6.4   Their Challenge

The data collected randomly and after that has been processed to be clean and accessible. The bank provided transactions data with about 20,000 records for the 19 past years. The records were collected from three different databases of bank branches distributed around Sudan.

3.6.5   The Solution

As a solution, we made many operations and processes on the data that was collected and creating dimensions and generating multidimensional recommendation model (star schema).

3.6.6   Results

As a result, we solve the single dimension approach problem and change it to multidimensional approach.

3.7 Data Sources

Data collection one of the important stage in the research. According to the research, the type of the data that will be used in is a multidimensional systems data. The research is targeting the financial data. The data collected from a Sudanese bank, and we put a specific specification and rules for this data. In addition, we put on consideration different attributes that it can lead to generate dimensions.

3.7.1 Introduction about data

Firstly, the data was containing financial information only but after making some operations there are some demographic information which has been added such as (age, sex, number of children…). The bank provided all customer information accept the private information because of their confidentiality such as (name, phone and account number). Some part of the data was missed so it needs to be preprocessed, also the data contains different data types such as (numeric and characters).

3.7.2 Data specification

The attributes of data represented as:

- Customer income

   The customer salary

- Customer sex

   The gender of the customer

- Customer marital

   Status of customer

- ID type

   The document of registration

- Customer age

   Number of customer's years.

- Financial size type

   The loan size

- Financial duration in month

   Duration of paying loan

- Financial duration start

   Loan starting time

- o Financial duration end

    Loan ending time

- o Financial type

    Sector type

- o Financial mode

    Sector mode

- o Sector

    Finance sector

- o Defaulter or not

    Default sector

- o Brief name arab

    Insurance name

- o Number of children

    Numbers of customer children

- o Number of spouse

    Number of customer spouse

- o Monthly expenditure

    Expending of duration

3.8 Summary

The chapter covers the research methodology and describes the research design and the research framework that will be implemented. In addition, it describes the interpretive research approach and the type of the research which specified as a quantitative research. In the next section it explains the theoretical multidimensional model. After that it defines the case study that will be used to test the multidimensional model and data collection. In the next chapter, the multidimensional model will be executed.

Chapter 4   Results and discussions

4.1   Introduction

This chapter discusses the mechanisms that used and the implementation of the multidimensional model and customer value measurement.

4.2   Implementation

4.2.1   Microsoft sql server management studio
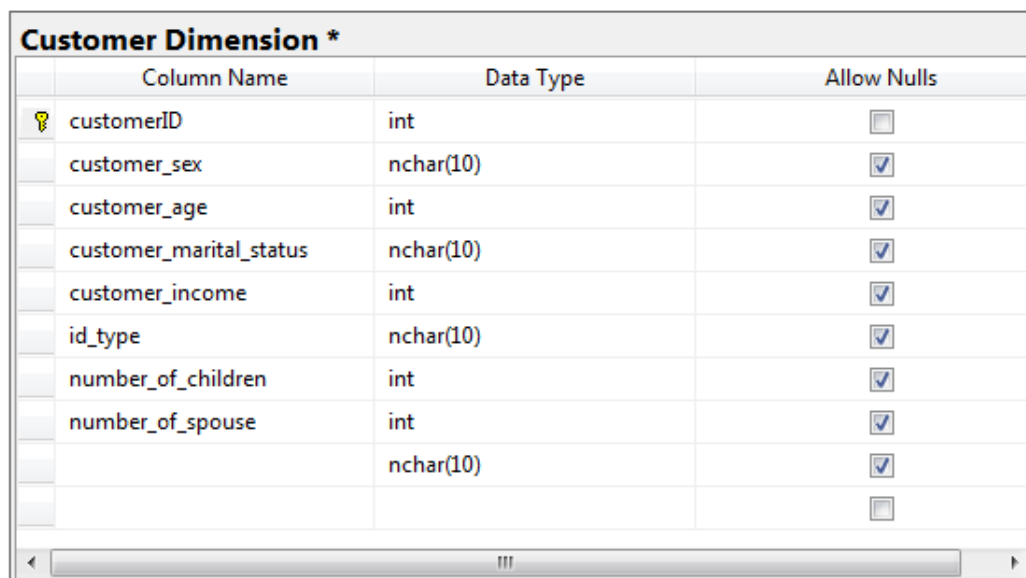
SQL Server Management Studio (SSMS) is a software application first launched with Microsoft SQL Server 2005 that used for configuring, managing, and administering all components within Microsoft SQL Server. The tool includes both script editors and graphical tools, which work with objects and features of the server.[2]

A central feature of SSMS is the Object Explorer, which allows the user to browse, select, and act upon any of the objects within the server.[3] It also shipped a separate Express edition that could be freely downloaded; however, recent versions of SSMS are fully capable of connecting to and manage any SQL Server Express instance. Microsoft also incorporated backwards compatibility for older versions of SQL Server thus allowing a newer version of SSMS to connect to older versions of SQL Server instances.

4.2.2 Multidimensional model

- **Creating the Dimensional Database and tables**

First we create the database, after that the tables and their columns.

Customer Dimension *		
Column Name	**Data Type**	**Allow Nulls**
customerID	int	☐
customer_sex	nchar(10)	☑
customer_age	int	☑
customer_marital_status	nchar(10)	☑
customer_income	int	☑
id_type	nchar(10)	☑
number_of_children	int	☑
number_of_spouse	int	☑
	nchar(10)	☑
		☐

Figure 4.1 Creating table and columns

- **Star Schema Model**

Star schema in computing is the simplest style of data mart schema and is the approach most widely used to develop data warehouses and dimensional data marts. The star schema consists of one or more fact tables referencing any number of dimension tables. The star schema is an important special case of the snowflake schema, and is more effective for handling simpler queries.

The star schema gets its name from the physical model's resemblance to a star shape with a fact table at its centre and the dimension tables surrounding it representing the star's points.
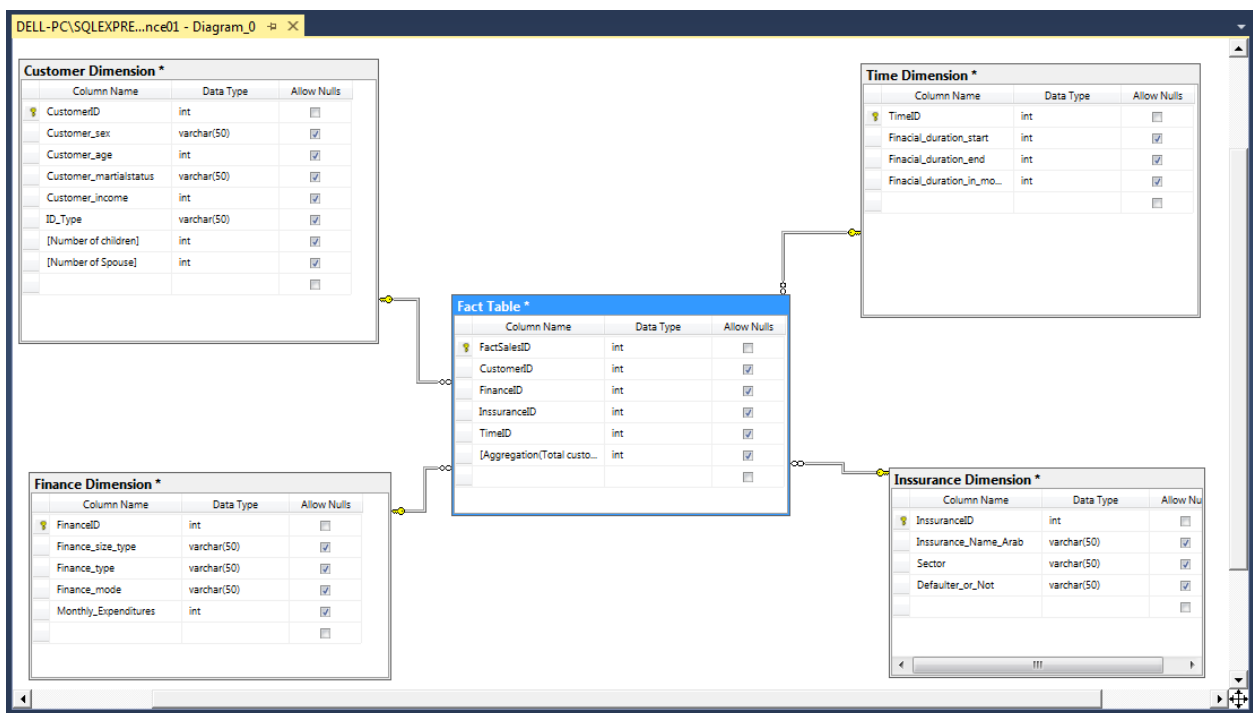
Figure 4.2 Star Schema Model

4.2.3 Data Marts Process

The star schema divided to small data marts. Each data mart has a specific task. Moreover, we have in this research four Data Marts and they are:

- Customer Data Mart
- Finance Data Mart
- Insurance Data Mart
- Time Data Mart

Consider a database of bank customer's, perhaps from a store chain, classified by Time, Finance and Insurance. Fact_Table is the fact table and there are four dimension tables Time_Dimension, Customer_Dimension,Finance_Dimension and Inssurance_Dimension.Each dimension table has a primary key on its Id column, relating to one of the of the Fact_Table table's four-column (compound) primary key (TimeID, CustomerID, FinanceID,InssuranceID). The non-primary key Aggregation(Total_customer) column of the fact table in the schema represents a measure or metric that can be used in calculations and analysis. The non-primary key columns of the dimension tables represent additional attributes of the dimensions (such as the Finance_duration_start of the Time_Dimension dimension).

4.2.4 Executing queries of creating table:

- **Finance table**
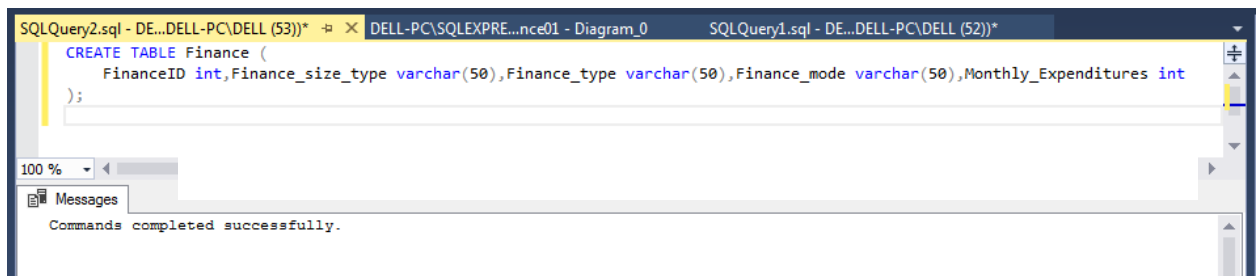
Figure 4.3 Query for creating finance table

- **Customer table**
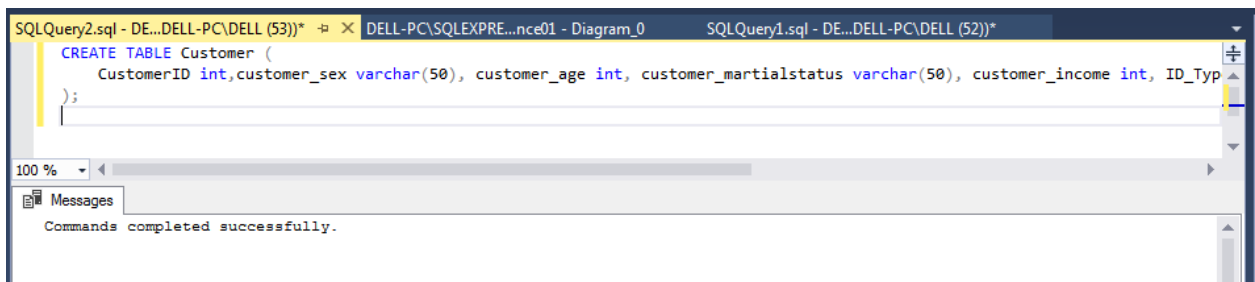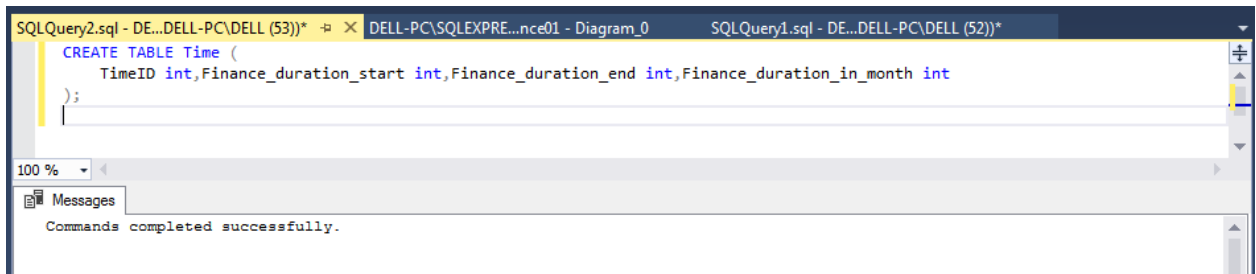
Figure 4.4Query for creating customer table

- **Time table**
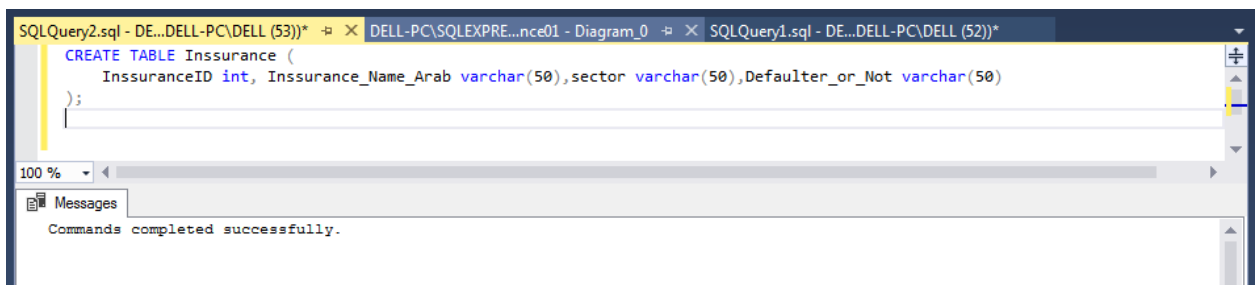
Figure 4.5 Query for creating time table

- **Insurance table**
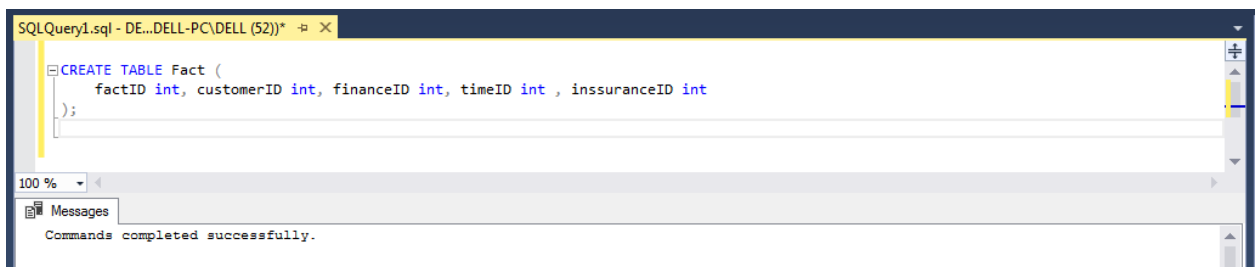
Figure 4.6 Query for creating insurance table

- **Fact table**

Figure 4.7 Query for creating Fact table

4.3 ETL Process

4.3.1 Step 1: Extract data from excel file

Four excel file has been created with a columns names. The first file for Finance and the columns are FinanceID, Finance_size, Finance_type, Finance_work monthly_expend. After that add some data as shown in below image.

C	D	E	F	G	H	I	J	K
				Finance Infromation				
		FinanceID	**Finance_size**	**Finance_type**	**Finance_work**	**monthly_expend**		
		1	SMALE	DIRECT	مرابحة زراعية	0		
		2	SMALE	DIRECT	مرابحة زراعية	0		
		3	MICRO	DIRECT	مرابحة زراعية	0		
		4	NORMAL	DIRECT	مرابحة حرفي	0		
		5	NORMAL	DIRECT	مرابحة اخرى	0		
		6	SMALE	DIRECT	مرابحة تجارة محلية	0		
		7	MICRO	DIRECT	مرابحة زراعية	0		
		8	NORMAL	DIRECT	مرابحة صناعية	0		
		9	SMALE	DIRECT	مرابحة تجارة محلية	0		
		10	NORMAL	DIRECT	مرابحة تجارة محلية	0		
		11	SMALE	DIRECT	مرابحة تجارة محلية	0		
		12	SMALE	DIRECT	سلم	0		
		13	MICRO	DIRECT	سلم	0		
		14	NORMAL	DIRECT	مرابحة اخرى	0		
		15	NORMAL	DIRECT	مرابحة صناعية	0		
		16	NORMAL	DIRECT	مرابحة صناعية	0		
		17	NORMAL	DIRECT	مرابحة صناعية	0		
		18	NORMAL	DIRECT	مرابحة صناعية	0		
		19	NORMAL	DIRECT	مرابحة صناعية	0		
		20	NORMAL	DIRECT	مرابحة صناعية	0		
		21	MICRO	DIRECT	مرابحة تجارة محلية	2000		
		22	NORMAL	DIRECT	مرابحة صناعية	0		
		23	NORMAL	DIRECT	مرابحة صناعية	0		
		24	NORMAL	DIRECT	مرابحة مواصلات	0		
		25	NORMAL	DIRECT	مرابحة مواصلات	0		
		26	NORMAL	DIRECT	مرابحة صناعية	0		
		27	NORMAL	DIRECT	مرابحة صناعية	0		
		28	NORMAL	DIRECT	استصناع انشاء	0		
		29	NORMAL	DIRECT	مرابحة تجارة محلية	0		
		30	NORMAL	DIRECT	مرابحة تجارة محلية	0		
		31	NORMAL	DIRECT	مرابحة تجارة محلية	0		
		32	NORMAL	DIRECT	مرابحة تجارة محلية	0		

Figure No. 4.8 Finance excel file

The second file for Customer and the columns are customerID, gender, age, marital_status, income, id_type, duration_in_month, no._of_children, no._of_spouse.After that add some data as shown in below image.

customerID	gender	age	marital_status	income	id_type	duration_in_months	no._of_children	no._of_spouse
					Customer Information			
1	M	37.0	N	.0	Passport	6.0	.0	.0
2	M	37.0	N	.0	PersonalCard	6.0	.0	.0
3	M	37.0	N	.0	Passport	6.0	.0	.0
4	S	18.0	N	.0	Com. Reg. No.	6.0	.0	.0
5	M	59.0	Y	.0	Passport	9.0	3.0	1.0
6	F	37.0	N	.0	PersonalCard	6.0	.0	.0
7	F	37.0	N	.0	PersonalCard	6.0	.0	.0
8	S	27.0	N	.0	Com. Reg. No.	24.0	.0	.0
9	F	37.0	N	.0	PersonalCard	6.0	.0	.0
10	M	46.0	N	.0	Passport	10.0	.0	.0
11	F	37.0	N	.0	PersonalCard	6.0	.0	.0
12	F	37.0	N	.0	PersonalCard	7.0	.0	.0
13	F	83.0	N	.0	PersonalCard	7.0	.0	.0
14	M	60.0	Y	.0	PersonalCard	6.0	.0	.0
15	S	19.0	N	.0	Com. Reg. No.	12.0	.0	.0
16	S	19.0	N	.0	Com. Reg. No.	12.0	.0	.0
17	S	19.0	N	.0	Com. Reg. No.	12.0	.0	.0
18	S	19.0	N	.0	Com. Reg. No.	12.0	.0	.0
19	S	19.0	N	.0	Com. Reg. No.	12.0	.0	.0
20	S	19.0	N	.0	Com. Reg. No.	12.0	.0	.0
21	M	63.0	Y	.0	PersonalCard	6.0	5.0	1.0
22	S	19.0	N	.0	Com. Reg. No.	12.0	.0	.0
23	S	19.0	N	.0	Com. Reg. No.	12.0	.0	.0
24	S	18.0	N	.0	Com. Reg. No.	24.0	.0	.0
25	S	18.0	N	.0	Com. Reg. No.	24.0	.0	.0
26	S	19.0	N	.0	Com. Reg. No.	12.0	.0	.0
27	S	19.0	N	.0	Com. Reg. No.	12.0	.0	.0
28	M	81.0	Y	.0	Passport	54.0	5.0	1.0
29	M	58.0	Y	.0	Passport	9.0	.0	1.0
30	M	58.0	Y	.0	Passport	9.0	.0	1.0

Figure No. 4.9 Customer excel file

The third file for Time and the columns are timeID, duration_start, duration_end, No._duration_in_month. After add some data as shown in below image.

timeID	duration_start	duration_end	No._duration_in_months
		Time Information	
timeID	duration_start	duration_end	No._duration_in_months
1	Jul 28 1998 12:00AM	Jan 28 1999 12:00AM	6.0
2	Jul 28 1998 12:00AM	Jan 26 1999 12:00AM	6.0
3	Dec 20 1998 12:00AM	Jun 20 1999 12:00AM	6.0
4	Mar 6 2002 12:00AM	Sep 6 2002 12:00AM	6.0
5	Mar 22 2002 12:00AM	Dec 9 2002 12:00AM	9.0
6	Nov 14 2002 12:00AM	May 13 2003 12:00AM	6.0
7	Jul 19 2003 12:00AM	Jan 18 2004 12:00AM	6.0
8	Nov 30 2003 12:00AM	Nov 29 2005 12:00AM	24.0
9	Jan 13 2004 12:00AM	Jul 12 2004 12:00AM	6.0
10	Feb 8 2004 12:00AM	Dec 8 2004 12:00AM	10.0
11	Feb 15 2004 12:00AM	Aug 14 2004 12:00AM	6.0
12	Aug 3 2004 12:00AM	Mar 15 2005 12:00AM	7.0
13	Aug 9 2004 12:00AM	Mar 15 2005 12:00AM	7.0
14	Jan 7 2005 12:00AM	Jul 28 2005 12:00AM	6.0
15	Apr 19 2005 12:00AM	Apr 19 2006 12:00AM	12.0
16	Apr 19 2005 12:00AM	Apr 19 2006 12:00AM	12.0
17	Jun 25 2005 12:00AM	Jun 25 2006 12:00AM	12.0
18	Jun 25 2005 12:00AM	Jun 25 2006 12:00AM	12.0
19	Jul 16 2005 12:00AM	Jul 16 2006 12:00AM	12.0
20	Jul 16 2005 12:00AM	Jul 16 2006 12:00AM	12.0
21	Sep 29 2005 12:00AM	Mar 29 2006 12:00AM	6.0
22	Oct 18 2005 12:00AM	Oct 18 2006 12:00AM	12.0
23	Oct 18 2005 12:00AM	Oct 18 2006 12:00AM	12.0
24	Oct 29 2005 12:00AM	Oct 27 2007 12:00AM	24.0
25	Oct 29 2005 12:00AM	Oct 27 2007 12:00AM	24.0
26	Nov 20 2005 12:00AM	Nov 20 2006 12:00AM	12.0
27	Nov 20 2005 12:00AM	Nov 20 2006 12:00AM	12.0
28	Nov 29 2005 12:00AM	May 29 2010 12:00AM	54.0
29	Apr 8 2006 12:00AM	Jan 8 2007 12:00AM	9.0
30	Apr 8 2006 12:00AM	Jan 8 2007 12:00AM	9.0

Figure No. 4.10 Time excel file

The fourth file for Inssurance and the columns are as inssuranceID, insurance_name, sector, defaulter_or_not. After that add some data as shown in below image.

C	D	E	F	G	H	I	J
			Inssurance Information				
		inssuranceID	inssurance_name	sector	defaulter_or_not		
		1	رهن عقاري سكني	أخرى	Not Defaulter		
		2	رهن عقاري سكني	أخرى	Not Defaulter		
		3	رهن عقاري سكني	أخرى	Not Defaulter		
		4	شهادات تخزين البضائع	أخرى	Not Defaulter		
		5	رهن عقاري سكني	أخرى	Not Defaulter		
		6	رهن عقاري تجاري	أخرى	Not Defaulter		
		7	رهن عقاري تجاري	أخرى	Not Defaulter		
		8	رهن عقاري تجاري	أخرى	Not Defaulter		
		9	رهن عقاري سكني	أخرى	Not Defaulter		
		10	رهن عائم	أخرى	Not Defaulter		
		11	رهن عقاري سكني	أخرى	Not Defaulter		
		12	رهن عقاري سكني	أخرى	Not Defaulter		
		13	ضمان شخصي	أخرى	Not Defaulter		
		14	رهن عقاري تجاري	أخرى	Not Defaulter		
		15	رهن عقاري تجاري	أخرى	Not Defaulter		
		16	رهن عقاري تجاري	أخرى	Not Defaulter		
		17	رهن عقاري تجاري	أخرى	Not Defaulter		
		18	رهن عقاري تجاري	أخرى	Not Defaulter		
		19	رهن عقاري تجاري	أخرى	Not Defaulter		
		20	رهن عقاري تجاري	أخرى	Not Defaulter		
		21	رهن عائم	أخرى	Not Defaulter		
		22	رهن عقاري تجاري	أخرى	Not Defaulter		
		23	رهن عقاري تجاري	أخرى	Not Defaulter		
		24	رهن عائم	أخرى	Not Defaulter		
		25	رهن عقاري سكني	أخرى	Not Defaulter		
		26	رهن عقاري تجاري	أخرى	Not Defaulter		
		27	رهن عقاري تجاري	أخرى	Not Defaulter		
		28	رهن عقاري سكني	أخرى	Not Defaulter		
		29	رهن عقاري سكني	أخرى	Not Defaulter		
		30	رهن عقاري تجاري	أخرى	Not Defaulter		

Figure No. 4.11 Insurance excel file

53

4.3.2 Step 2: Transform data (Convert local currency to dollar currency)

Customer excel file has column name called "Income" and that amount is in local currency. So we need to convert the local amount in that time to dollar currency amount to be able to evaluate the incomes of customer in this time because the collected data are old and cannot reflect the real value.

In transformation process, rapid miner tool has been used to perform the process and the steps of the process are:

1- Firstly, importing the data into rapid miner as shown below.

	customerID	Customer s...	Number Of ...	Number Of ...	Customer i...	Customer ...	ID_Type	Customer a...
	integer	polynominal	integer	integer	integer	polynominal	polynominal	integer
1	1	M	0	0	0	N	Passport	37
2	2	M	0	0	0	N	PersonalCard	37
3	3	M	0	0	0	N	Passport	37
4	4	S	0	0	0	N	Com. Reg. No.	18
5	5	M	3	1	0	Y	Passport	59
6	6	F	0	0	0	N	PersonalCard	37
7	7	F	0	0	0	N	PersonalCard	37
8	8	S	0	0	0	N	Com. Reg. No.	27
9	9	F	0	0	0	N	PersonalCard	37
10	10	M	0	0	0	N	Passport	46
11	11	F	0	0	0	N	PersonalCard	37
12	12	F	0	0	0	N	PersonalCard	37
13	13	F	0	0	0	N	PersonalCard	83
14	14	M	0	0	0	Y	PersonalCard	60
15	15	S	0	0	0	N	Com. Reg. No.	19
16	16	S	0	0	0	N	Com. Reg. No.	19
17	17	S	0	0	0	N	Com. Reg. No.	19
18	18	S	0	0	0	N	Com. Reg. No.	19

Figure No. 4.12 Import data source

2- After that display the data that has been imported as shown below

Figure No. 4.13 Displaying data source

3- Create the transformation expression that want to be generated

Figure No. 4.14 Create transformation expression

4- The interface shown the generated expression
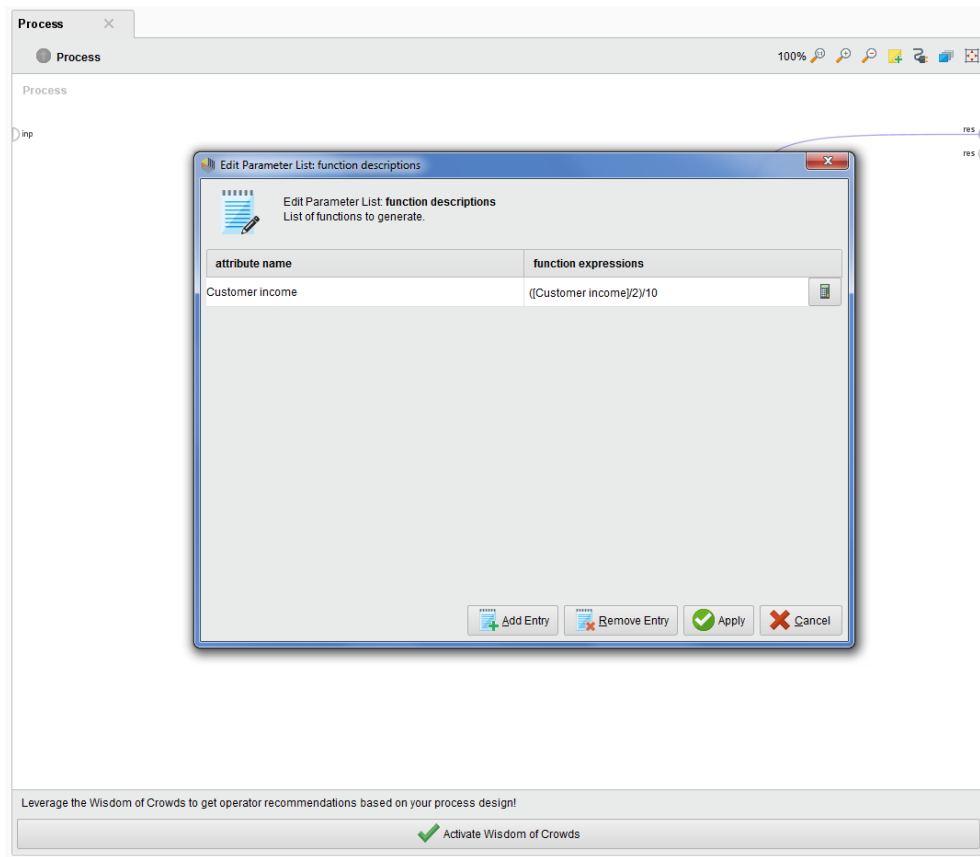
Figure No. 4.15 Generated expression

5- The full diagram of transformation process which containing (the data source, generate attributes and write excel) and complete data transformation part.
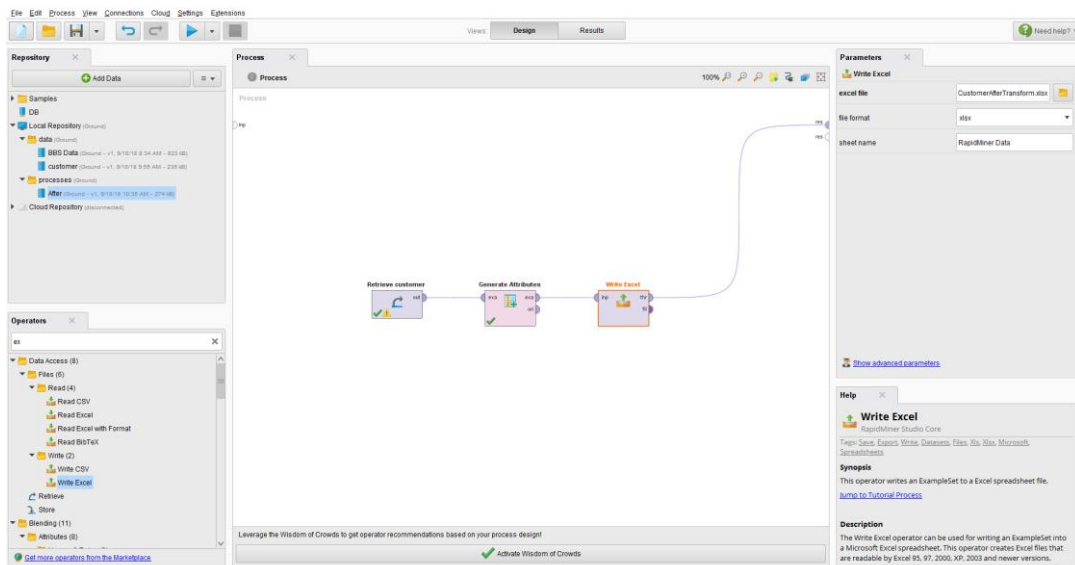
Figure No. 4.16 Full process of transformation

56
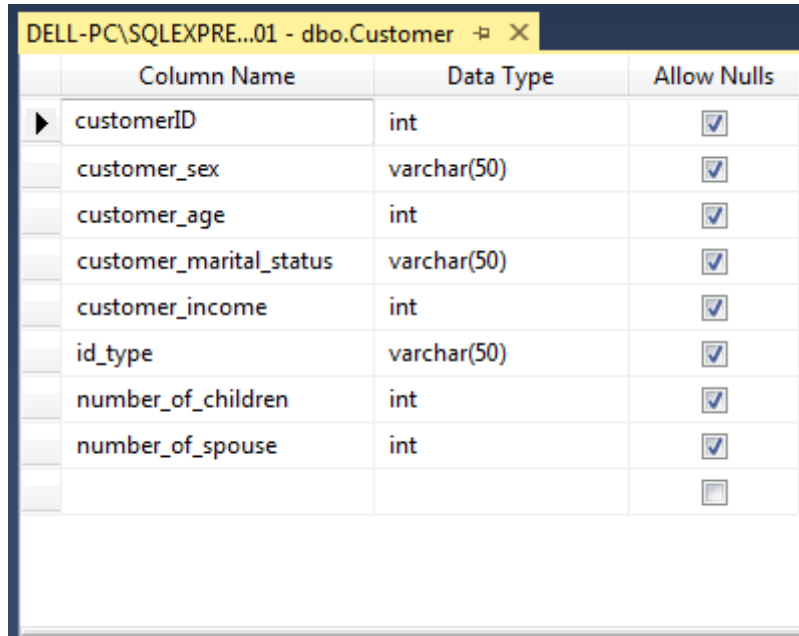
6- Data after transformation

	customerID	Customer sex	Number Of Children	Number Of Spouse	Customer marit	ID_Type	Customer age	Customer income
					Customer Information After Transformation (Convert to Dollar currency and Missing Values)			
	1.0	M	.0	.0	N	Passport	37.0	.0
	2.0	M	.0	.0	N	PersonalCard	37.0	.0
	3.0	M	.0	.0	N	Passport	37.0	.0
	4.0	S	.0	.0	N	Com. Reg. No.	18.0	.0
	5.0	M	3.0	1.0	Y	Passport	59.0	.0
	6.0	F	.0	.0	N	PersonalCard	37.0	.0
	7.0	F	.0	.0	N	PersonalCard	37.0	.0
	8.0	S	.0	.0	N	Com. Reg. No.	27.0	.0
	9.0	F	.0	.0	N	PersonalCard	37.0	.0
	10.0	M	.0	.0	N	Passport	46.0	.0
	11.0	F	.0	.0	N	PersonalCard	37.0	.0
	12.0	F	.0	.0	N	PersonalCard	37.0	.0
	13.0	F	.0	.0	N	PersonalCard	83.0	.0
	14.0	M	.0	.0	Y	PersonalCard	60.0	.0
	15.0	S	.0	.0	N	Com. Reg. No.	19.0	.0
	16.0	S	.0	.0	N	Com. Reg. No.	19.0	.0
	17.0	S	.0	.0	N	Com. Reg. No.	19.0	.0
	18.0	S	.0	.0	N	Com. Reg. No.	19.0	.0
	19.0	S	.0	.0	N	Com. Reg. No.	19.0	.0
	20.0	S	.0	.0	N	Com. Reg. No.	19.0	.0
	21.0	M	5.0	1.0	Y	PersonalCard	63.0	.0
	22.0	S	.0	.0	N	Com. Reg. No.	19.0	.0
	23.0	S	.0	.0	N	Com. Reg. No.	19.0	.0
	24.0	S	.0	.0	N	Com. Reg. No.	18.0	.0
	25.0	S	.0	.0	N	Com. Reg. No.	18.0	.0
	26.0	S	.0	.0	N	Com. Reg. No.	19.0	.0
	27.0	S	.0	.0	N	Com. Reg. No.	19.0	.0
	28.0	M	5.0	1.0	Y	Passport	81.0	.0
	29.0	M	.0	1.0	Y	Passport	58.0	.0
	30.0	M	.0	1.0	Y	Passport	58.0	.0

Figure No. 4.17 Data after transformation

57

4.3.3 Step 3: Loading data to SQL server

Before this step, SQL server management studio opened up and has been create a new database and new table with the same excel column names as shown below image.

Figure No. 4.18 Tables in sql server form excel file

Loading data into sql server has many steps and they are:

1- Choosing the data source
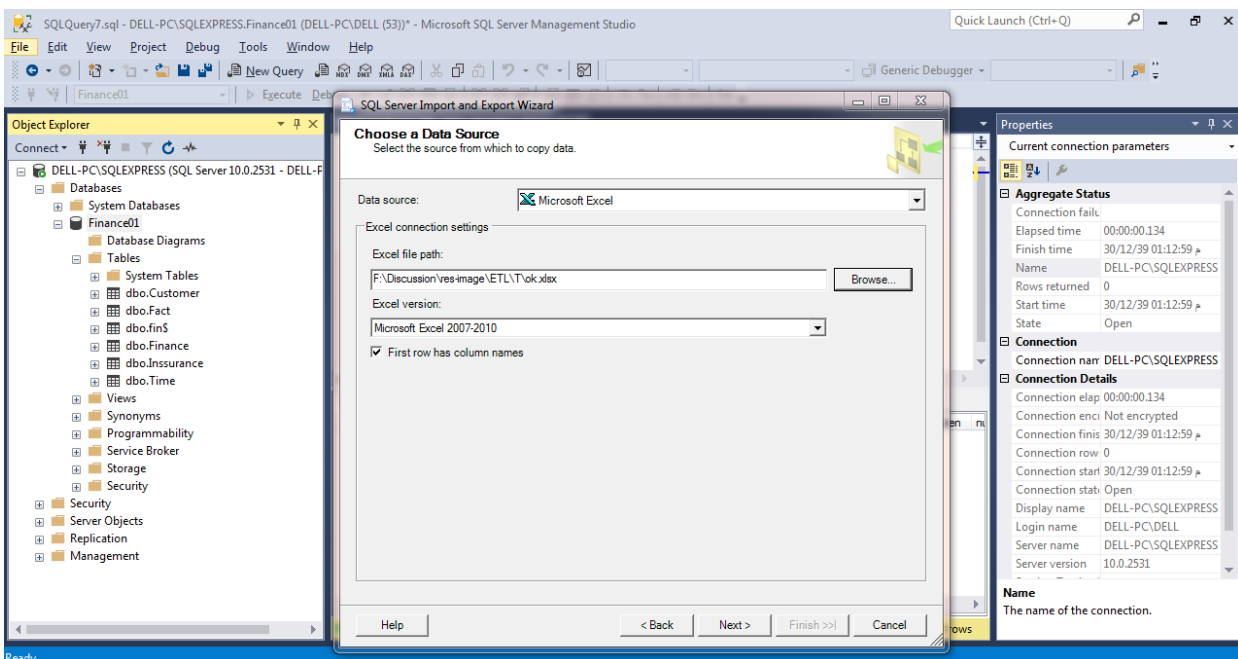
Figure No. 4.19 Choose data source

2- Choosing the destination
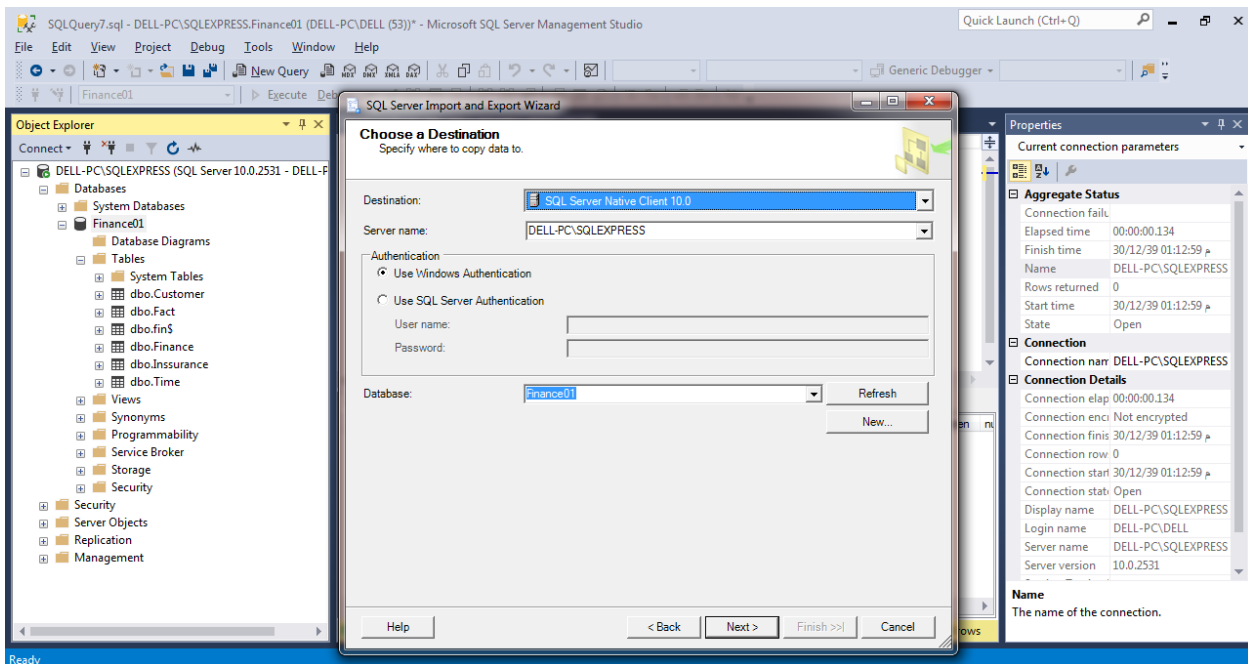
Figure No. 4.20 Choose the place

3- Selecting the source tables and views
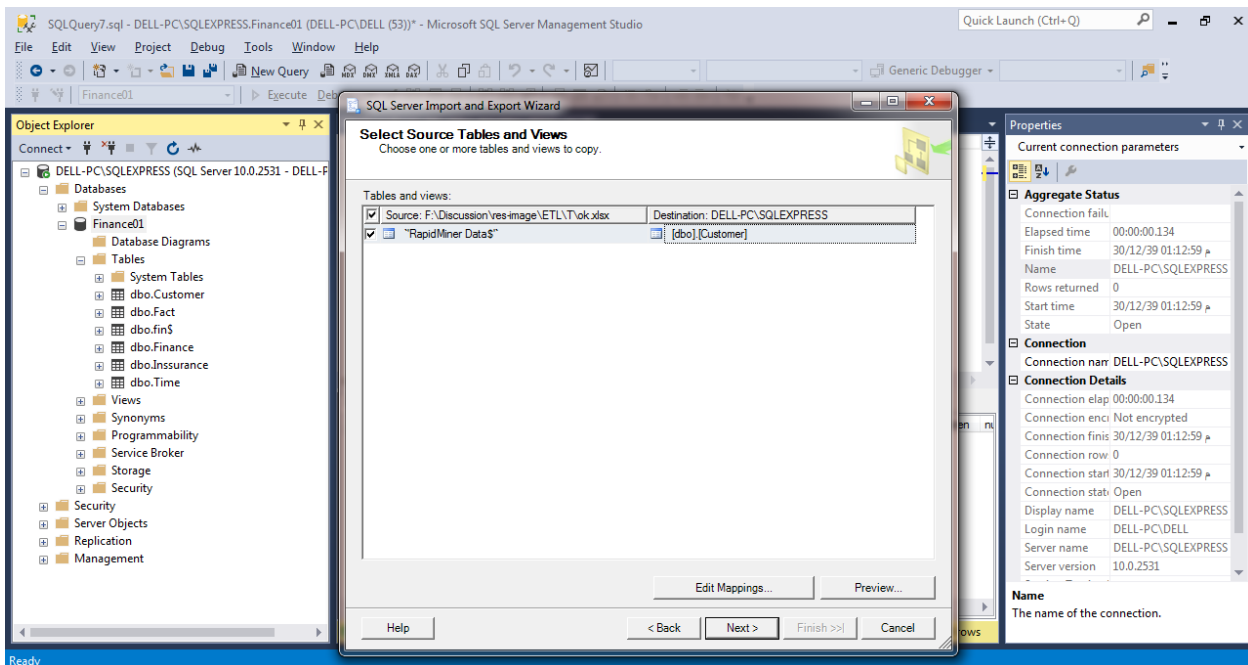
Figure No. 4.21 Choose the table in sql server

4- Reviewing data type mapping
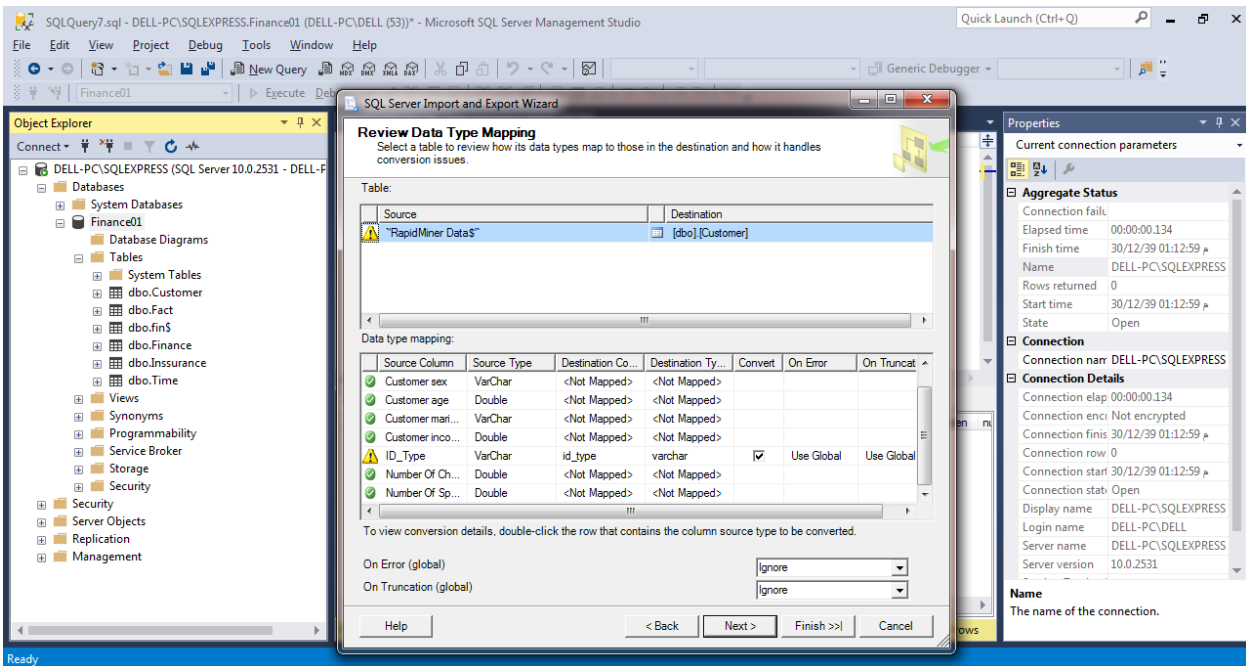
Figure No. 4.22 Review mapping

5- The execution of load process
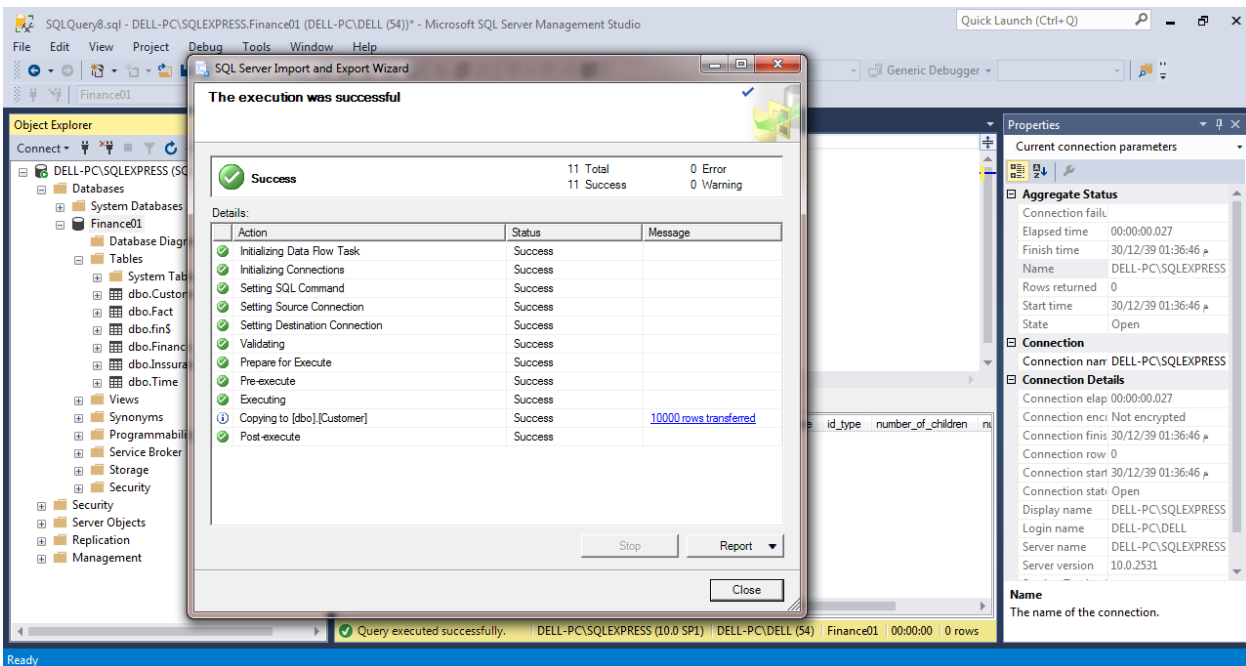
Figure No. 4.23 The execution

4.4 Join tables to apply multidimensional model

4.4.1 Join between Finance dimension and Customer dimension (Finance size and customer age:

The following query get the customer that meet the condition, there are four conditions in this query and they are:

1- Small ages and Small size of loan

Figure No. 4.24 Small ages and Small size query

2- Small ages and Big size of loan

Figure No. 4.25 Small ages and Big size query

3- Big ages and Small size of loan
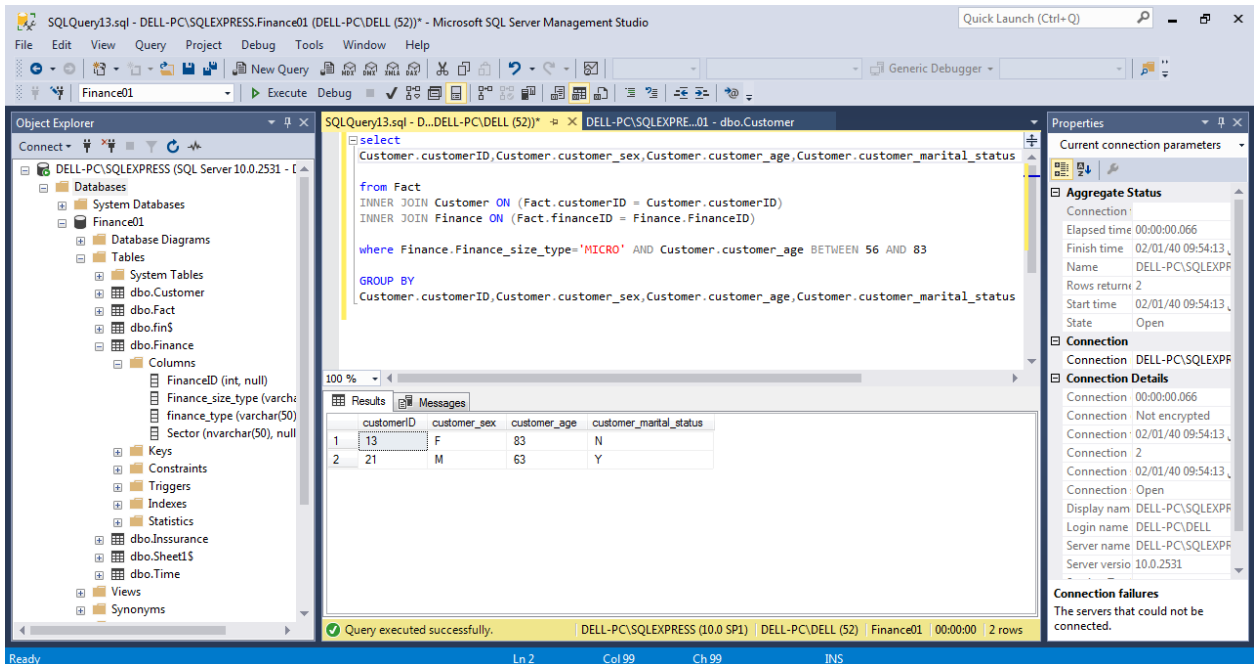
Figure No. 4.26 Big ages and Small size query

4- Big ages and Big size of loan
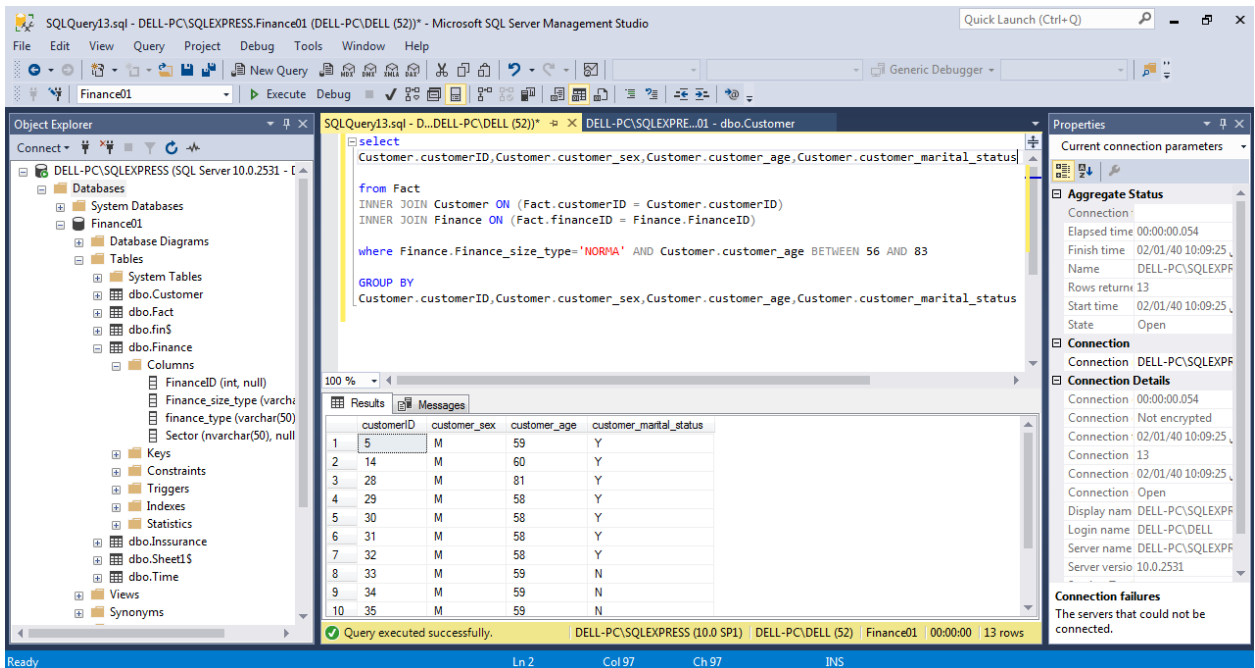
Figure No. 4.27 Big ages and Big size query

Evaluation of the first query:

Number	Age and Loan Size Dimensions Values	Number of customers
1	Small Ages and Small size of loan	2
2	Small Ages and Big size of loan	32
3	Big Ages and Small size of loan	2
4	Big Ages and Big size of loan	13

Table 4.1 Age and Loan Size evaluation
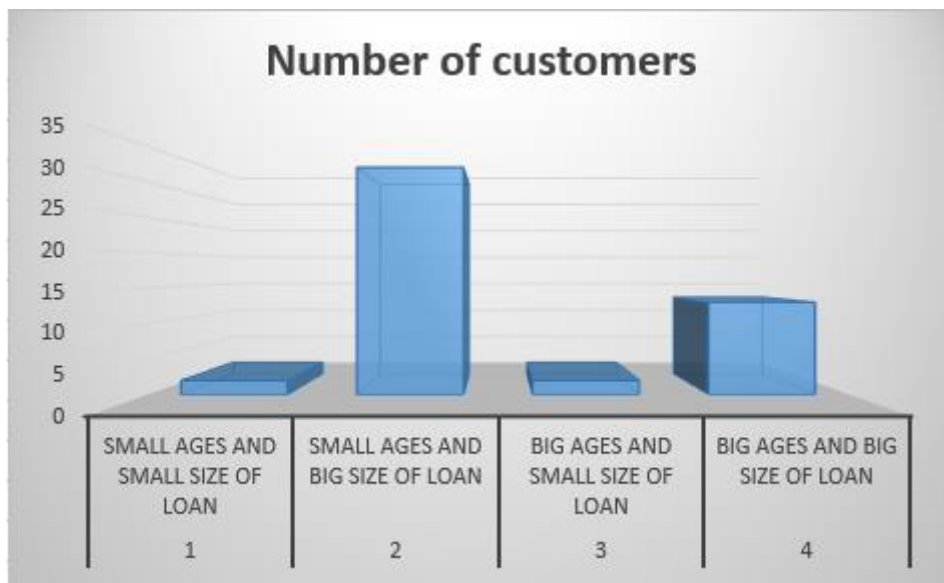
Figure No. 4.28 Evaluation of age and size queries

4.4.2 Join between Finance dimension and Time dimension (Finance sector and duration of loan in month:

The following query get the customer that meet the condition, there are two conditions in this query and they are:

1- Few Duration and Common Finance Sector

Figure No. 4.29 Few Duration and Common Sector query

2- Huge Duration and Common Finance Sector

Figure No. 4.30 Huge Duration and Common Sector query

Evaluation of the second query:

Number	Duration in month and Finance Sector Dimensions Values	Number of customers
1	Few Duration and Common Finance Sector	38
2	Huge Duration and Common Finance Sector	62

Table 4.2 Duration and Sector evaluation
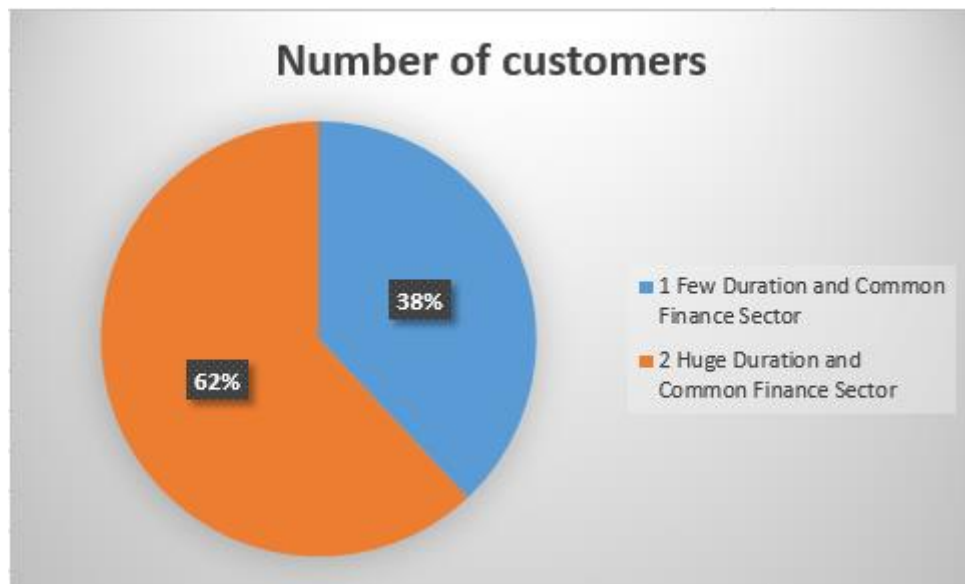
Figure No. 4.31 Evaluation of duration and sector queries

4.4.3 Join between Finance dimension and Insurance dimension (Finance sector and insurance type:

The following query get the customer that meet the condition, there are four conditions in this query and they are:

1- Common Finance Sector ، رهن عقاري سكني

Figure No. 4.32 Sector and rahn sakani query

2- Common Finance Sector ، رهن عقاري تجاري

Figure No. 4.33 Sector and rahn tgari query

3- Common Finance Sector ، رهن عائم

Figure No. 4.34 Sector and rahn aaim query

4- Common Finance Sector ، ضمان شخصي

Figure No. 4.35 Sector and personal insurance query

Evaluation of the third query:

Number	Insurance Name and Finance Mode Dimensions Values	Number of customers	Percentages
1	Common Finance Sector ، رهن عقاري سكني	14	14.43%
2	Common Finance Sector ، رهن عقاري تجاري	18	18.56%
3	Common Finance Sector ، رهن عائم	35	36.08%
4	Common Finance Sector ، ضمان شخصي	30	30.93%

Table 4.3 Insurance name and Sector evaluation
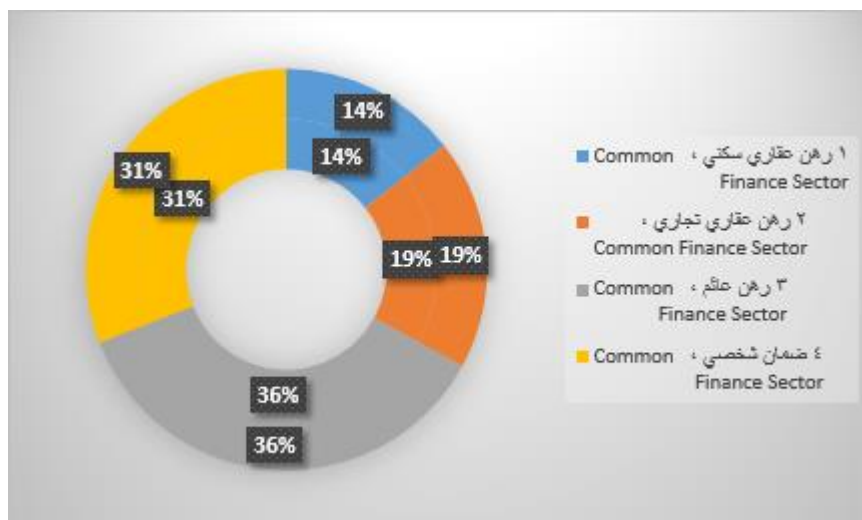
Figure No. 4.36 Evaluation of sector and insurance queries

4.4.4 Join between Customer dimension and Time dimension (customer age and duration for loan in month):

The following query get the customer that meet the condition, there are four conditions in this query and they are:

1- Small Ages and Few Duration

Figure No. 4.37 Small Ages and Few Duration query

2- Small Ages and Huge Duration

Figure No. 4.38 Small Ages and Huge Duration query

3- Big Ages and Few Duration
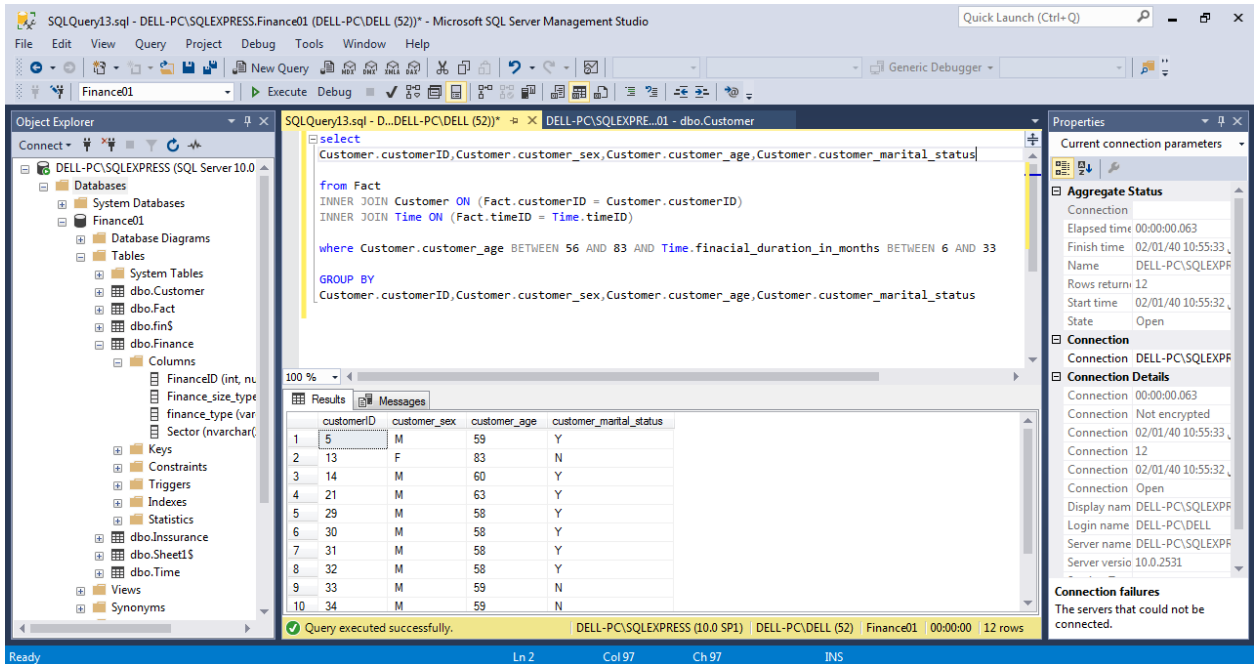
Figure No. 4.39 Big Ages and Few Duration query

4- Big Ages and Huge Duration
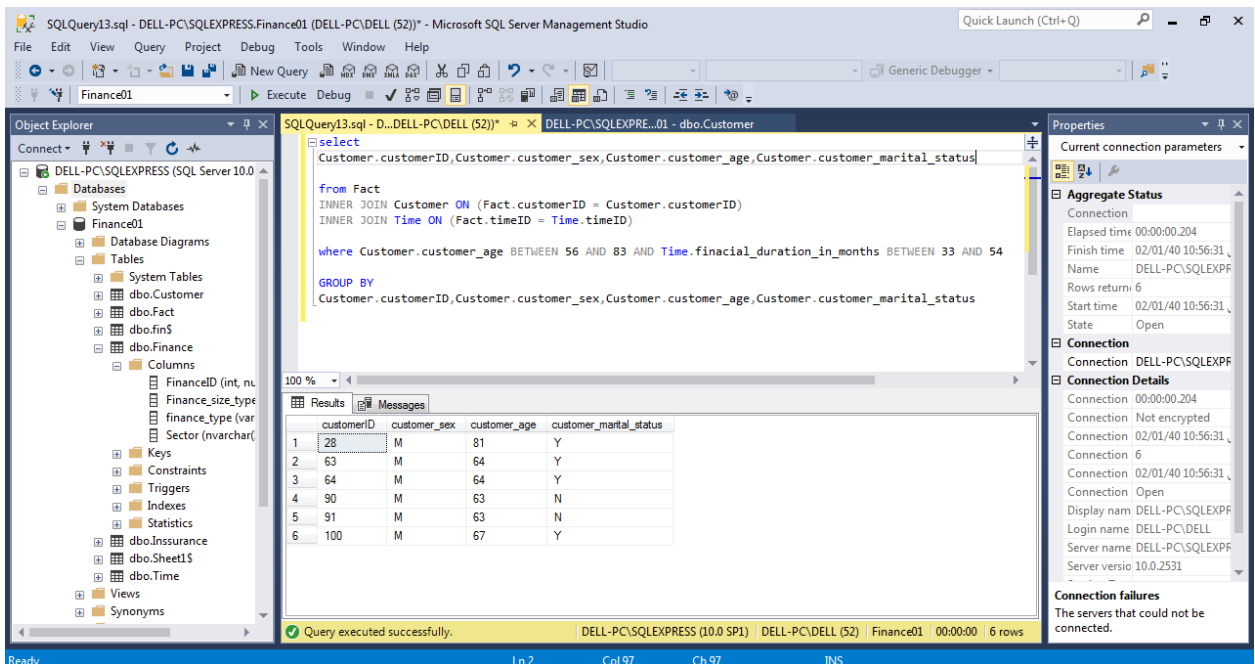
Figure No. 4.40 Big Ages and Huge Duration query

70

Evaluation of the fourth query:

Number	Age and Month Duration Dimensions Values	Number of customers	Percentages
1	Small Ages and Few Duration	26	26.00%
2	Small Ages and Huge Duration	56	56.00%
3	Big Ages and Few Duration	12	12.00%
4	Big Ages and Huge Duration	6	6.00%

Table 4.4 Age and Duration evaluation
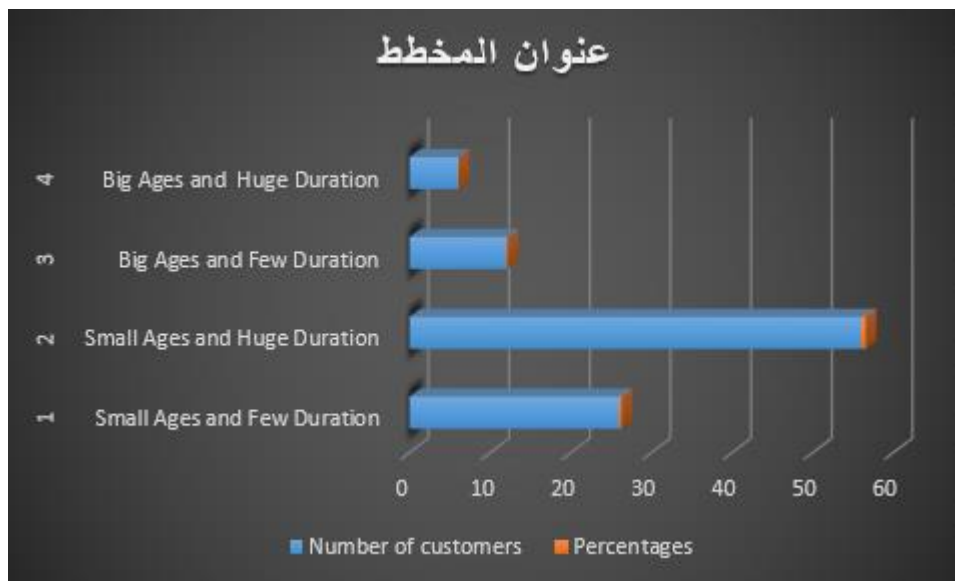
Figure No. 4.41 Evaluation of age and duration queries

4.4.5   Join between Finance dimension and Customer dimension (Finance size and customer marital status):

The following query get the customer that meet the condition, there are four conditions in this query and they are:

1- Married and Small size of loan

Figure No. 4.42 Married and Small size query

2- Married and Big size of loan

Figure No. 4.43 Married and Big size query

3- Single and Small size of loan
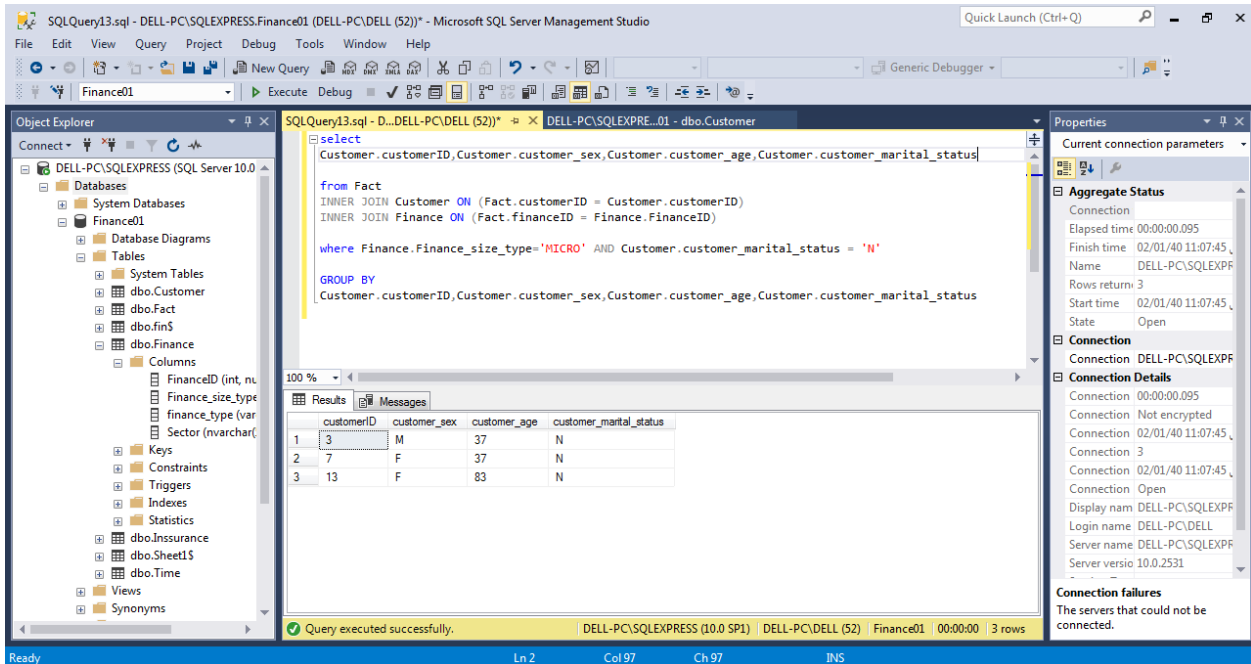
Figure No. 4.44 Single and Small size query

4- Single and Big size of loan
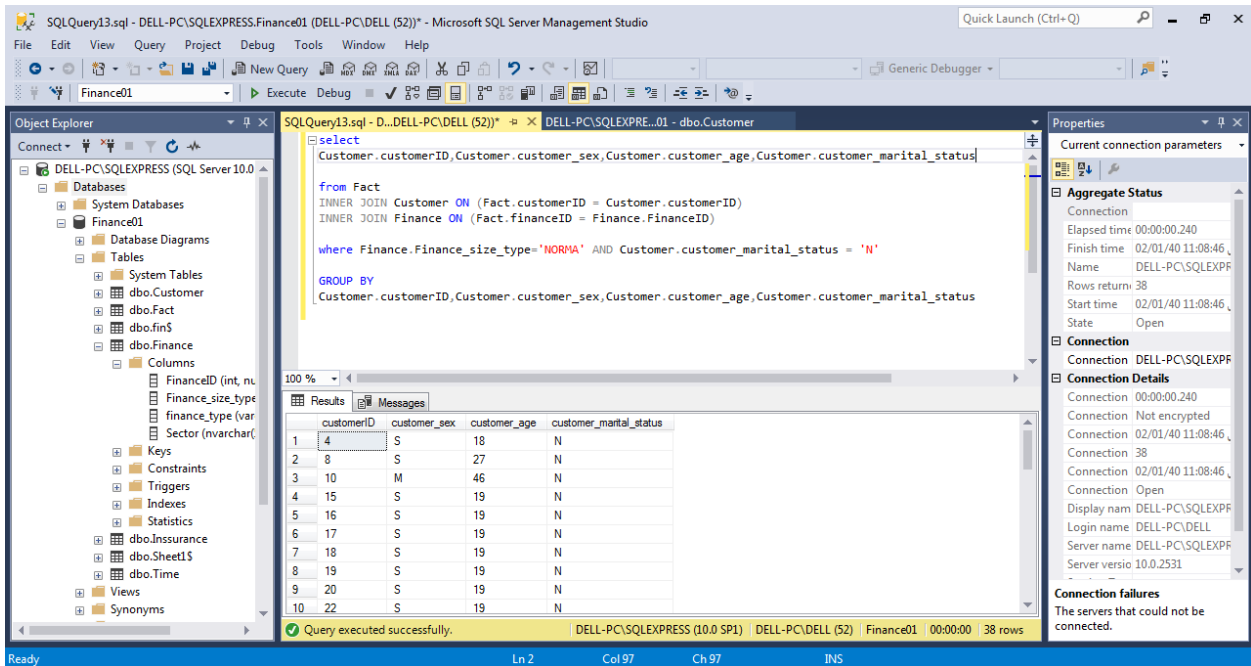
Figure No. 4.45 Single and Big size query

Evaluation of the fifth query:

- Category of results:

Number	Category	From – To
1	Very Small	1……10
2	Small	10 …. 20
3	Medium	20 …..30
4	Big	30……40

Table 4.5 Results category

Number	Marital Status and Loan Size Dimensions Values	Number of customers	Category
1	Married and Small size of loan	1	Very Small
2	Married and Big size of loan	7	Very Small
3	Single and Small size of loan	3	Very Small
4	Single and Big size of loan	38	Big

Table 4.6 Status and Loan Size evaluation
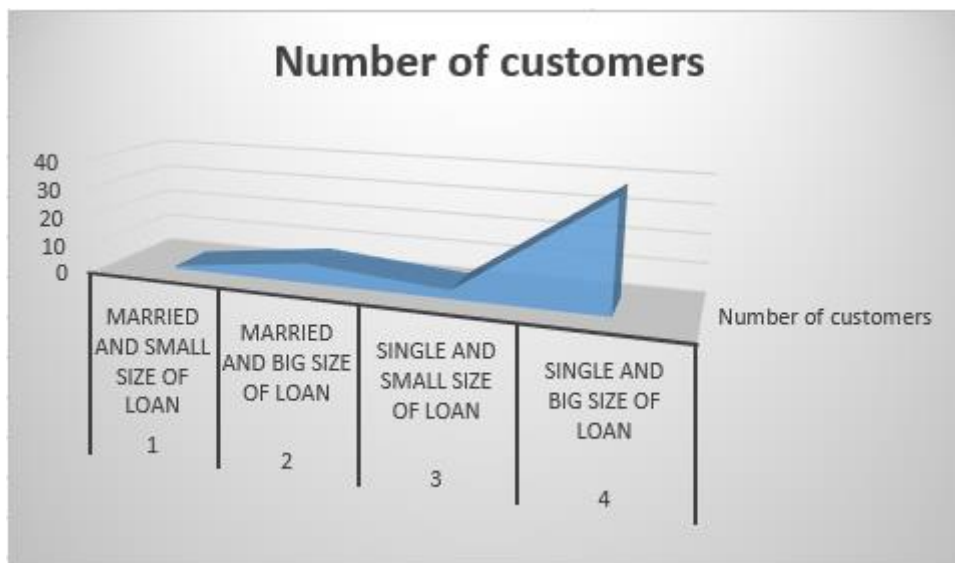
Figure No. 4.46 Evaluation of status and size queries

4.4.6 Join between Finance dimension and Customer dimension (Finance size and customer income):

The following query get the customer that meet the condition, there are six conditions in this query and they are:

1- Few Income and Micro size of loan

Figure No. 4.47 Few income and Micro size query

2- Few Income and Small size of loan

Figure No. 4.48 Few income and Small size query

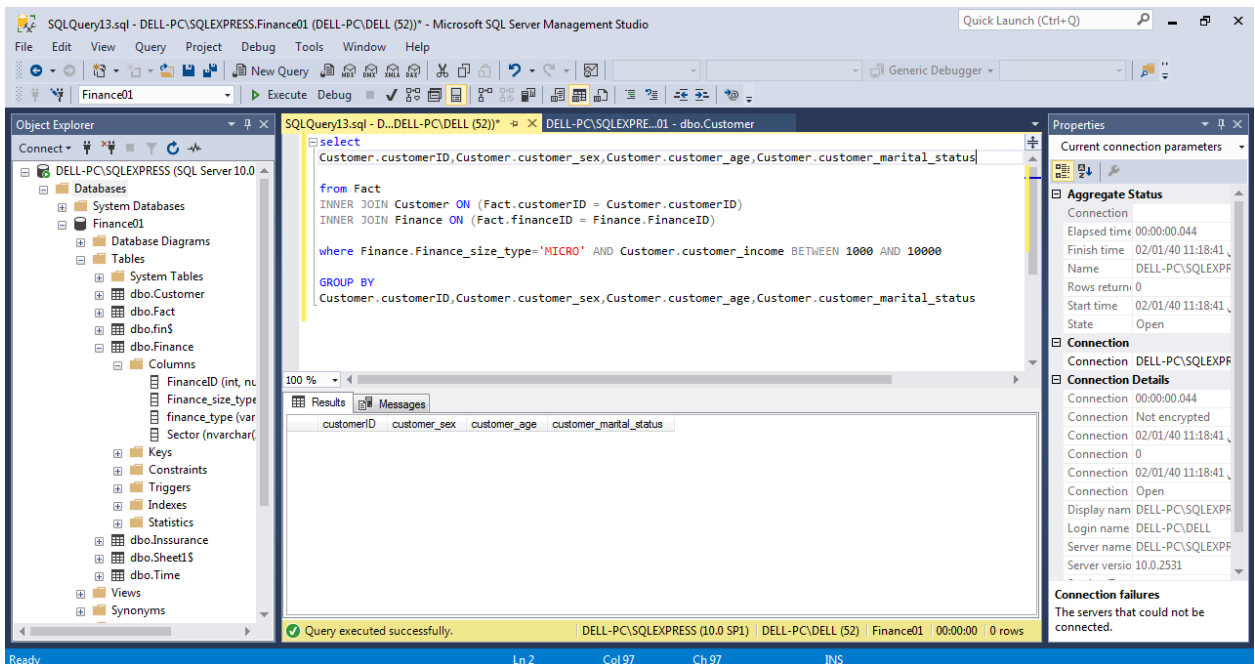3- Few Income and Normal size of loan

Figure No. 4.49 Few income and Normal size query
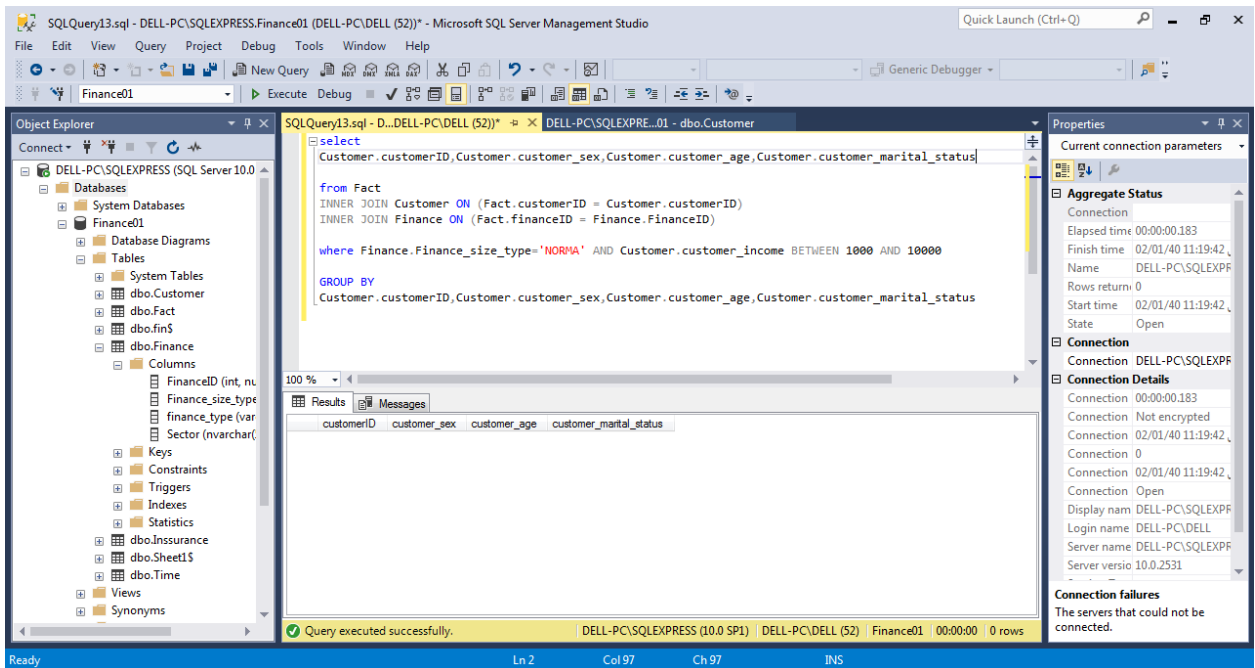
4- Huge Income and Micro size of loan

Figure No. 4.50 Huge income and Micro size query

5- Huge Income and Small size of loan

Figure No. 4.51 Huge income and Small size query

6- Huge Income and Normal size of loan

Figure No. 4.52 Huge Income and Normal size query

Evaluation of the sixth query:

Number	Customer Income and Loan Size Values	Number of customers	Percentages
1	Few Income and Micro size of loan	4	4.00%
2	Few Income and Small size of loan	50	50.00%
3	Few Income and Normal size of loan	45	45.00%
4	Huge Income and Micro size of loan	0	0.00%
5	Huge Income and Small size of loan	1	1.00%
6	Huge Income and Normal size of loan	0	0.00%

Table 4.7 Income and Loan Size evaluation

Figure No. 4.53 Evaluation of income and size queries

4.5 Measuring the customer value:

4.5.1 The Customer has been divided to three Segmentations:
- Low income customer (annual income between 500000 and 1000000)
- Medium income customer (annual income between 1000000 and 1800000)
- High income customer (annual income more than 1800000)

4.5.2 The Customer value measuring constrains :
- Date : from 24/11/2017 to 24/11/2018
- Number of transaction : 1000 transaction
- Type of data : transactional data
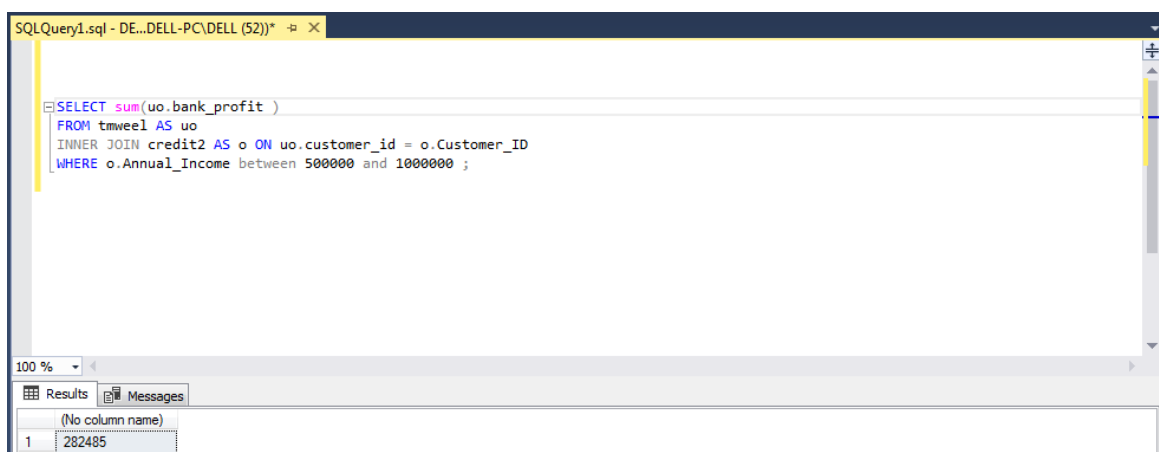
4.5.3 The bank gets the revenues from different streams and they are:

1- Funding Stream
- The attributes of funding stream:
  - customer_id
    - Data type: varchar(50)
    - Value : unique customer id
  - Amount_of_funding
    - Data type: integer
    - Value : fixed value (50000 SDG)
  - Profit_from_financing
    - Data type: integer
    - Value: different revenue according to the project
  - bank_profit
    - Data type: integer
    - Value : 17% revenue form the attribute "Profit_from_financing"

- The execution of funding stream:

```
SELECT sum(uo.bank_profit )
FROM tmweel AS uo
INNER JOIN credit2 AS o ON uo.customer_id = o.Customer_ID
WHERE o.Annual_Income between 500000 and 1000000 ;
```

(No column name)
282485

Figure No. 4.54 Funding stream

- Customer value calculated for the three segmentations as follow:
  - low income:
    - The value per year = 282485 SDG
    - The value per quarter = 70621.25 SDG
    - The percentage of low income revenue from hole revenue per quarter=25.3 %
  - medium income:
    - The value per year = 545788 SDG
    - The value per quarter = 136447 SDG
    - The percentage of low income revenue from hole revenue per quarter=48.8 %
  - high income:
    - The value per year = 287917 SDG
    - The value per quarter = 71979.25 SDG
    - The percentage of low income revenue from hole revenue per quarter=25.7 %

2- Transferring Stream:
- The attributes of transferring stream:
  - customer_id
    - Data type: varchar(50)
    - Value : unique customer id
  - Amount_to_transfer
    - Data type: integer
    - Value : from 1000 to 50000
  - Benefit
    - Data type: integer
    - Value : fixed value (400)
  - Date
    - Data type: varchar(50)
    - Value : from 24/11/2017 to 24/11/2018

- The execution of transferring stream:


```
SQLQuery1.sql - DE...DELL-PC\DELL (52))*  ⊕ ✕

SELECT sum(uo.Benefit )
FROM transfer2 AS uo
INNER JOIN credit AS o ON uo.customer_id = o.Customer_ID
WHERE o.Annual_Income between 500000 and 1000000 ;

100 %

Results   Messages
     (No column name)
1    159600
```

Figure No. 4.55 Transferring stream

- Customer value calculated for the three segmentation as follow:
    - Low income:
        - The value per quarter = 159600 SDG
        - The percentage of low income revenue from hole revenue per quarter=39.9%
    - Medium income:
        - The value per quarter = 120400 SDG
        - The percentage of low income revenue from hole revenue per quarter=30.1%
    - High income:
        - The value per quarter = 120000 SDG
        - The percentage of low income revenue from hole revenue per quarter=30%

3- ATM Stream:
- The attributes of ATM stream:
    - customer_id
        - Data type: varchar(50)
        - Value : unique customer id
    - customer_bank
        - Data type: varchar(50)
        - Value : banks such as (alkhartoum bank- Omdurman bank- faisal bank – farmer bank- alshimal bank – alneel bank – alarabi bank)
    - Amount_received
        - Data type : integer
        - Value :fixed value (4)
    - Date
        - Data type: varchar(50)
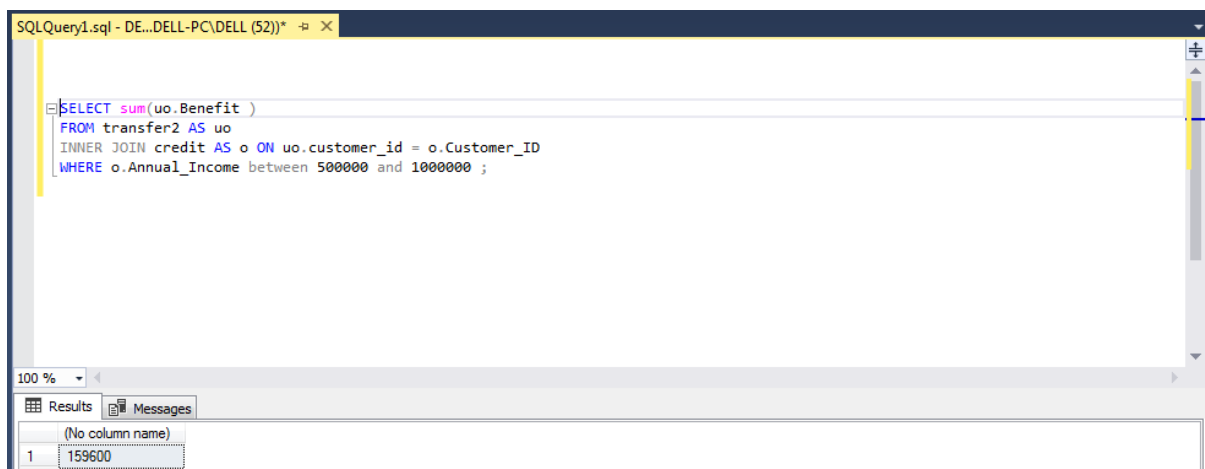   - Value: from 24/11/2017 to 24/11/2018
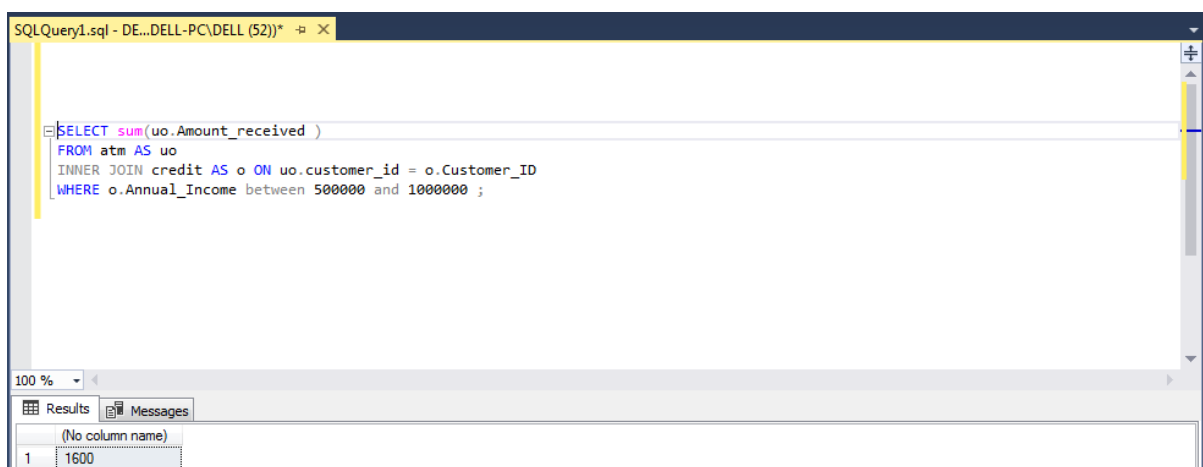
- The execution of ATM stream:

Figure No. 4.56 ATM stream

- Customer value calculated for the three segmentation as follow:
  - Low income:
    - o The value per quarter = 1600 SDG
    - o The percentage of low income revenue from hole revenue per quarter=40%
  - Medium income:
    - o The value per quarter = 1200 SDG
    - o The percentage of low income revenue from hole revenue per quarter=30%
  - High income:
    - o The value per quarter = 1200 SDG
    - o The percentage of low income revenue from hole revenue per quarter=30%

4- Customer operation Stream:
  - The attributes of customer operation stream:
    - customer_id
      - o data type : varchar(50)
      - o Value : unique customer id
    - customer_operation
      - o data type: varchar(50)
      - o value : operations such as(withdraw- Deposit- Bank check – local transfer)
    - Amount_received
      - o Data type: integer
      - o Value : values(0.5 – 1 – 2 - 6)
    - Date
      - o data type: varchar(50)
      - o Value: from 24/11/2017 to 24/11/2018

  - The execution of customer operation stream:


```
SQLQuery1.sql - DE...DELL-PC\DELL (52))*

SELECT sum(uo.Amount_received )
FROM opr AS uo
INNER JOIN credit AS o ON uo.customer_id = o.Customer_ID
WHERE o.Annual_Income between 500000 and 1000000 ;

100 %
Results   Messages
   (No column name)
1   942
```
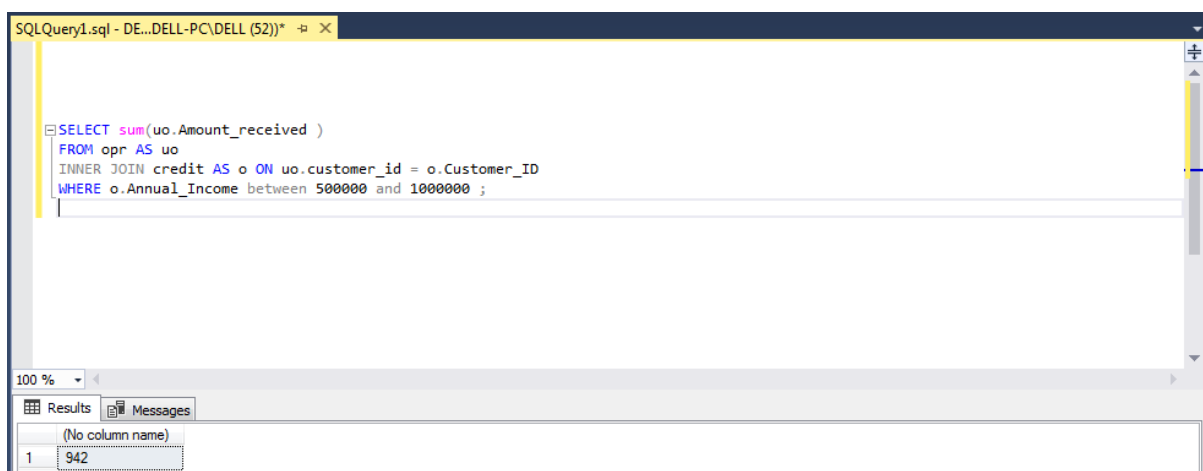
Figure No. 4.57 Customer operation stream

- Customer value calculated for the three segmentation as follow:
  - Low income:
    - o The value per quarter = 942 SDG
    - o The percentage of low income revenue from hole revenue per quarter=41.5%
  - Medium income:
    - o The value per quarter = 687 SDG
    - o The percentage of low income revenue from hole revenue per quarter=30.2%
  - High income:
    - o The value per quarter = 640 SDG
    - o The percentage of low income revenue from hole revenue per quarter=28.2%

5- Deposits Stream:
  - The attributes of deposits stream:
    - customer_id
      - o Data type : varchar(50)
      - o Value : unique customer id
    - Amount_deposited
      - o Data type : integer
      - o Value : from 50000 to 500000
    - Interest_rate
      - o Data type :real
      - o Value: rates as (4% - 5% - 8% - 10% - 13%- 15%)
    - Number_of_years
      - o Data type : integer
      - o Value : numbers as (3-4-5)
    - Date
      - o Data type : varchar(50)
      - o Value: from 24/11/2017 to 24/11/2018
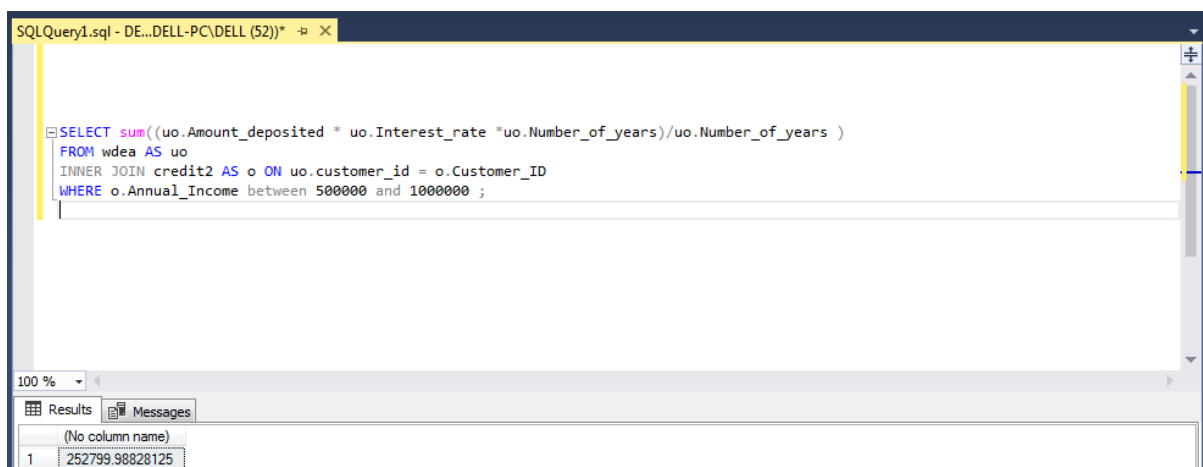 - The execution of deposit stream:

Figure No. 4.58 Deposit stream

83

- Customer value calculated for the three segmentation as follow:
  - Low income:
    - o The value per year = 252799.98 SDG
    - o The value per quarter =63199.995 SDG
    - o The percentage of low income revenue from hole revenue per quarter=16.3%
  - Medium income:
    - o The value per year = 865399.94 SDG
    - o The value per quarter =216349.985 SDG
    - o The percentage of low income revenue from hole revenue per quarter=56.1%
  - High income:
    - o The value per year = 423999.98 SDG
    - o The value per quarter =105999.995 SDG
    - o The percentage of low income revenue from hole revenue per quarter=27.4%

6- Loan Stream:
  - The attributes of loan stream:
    - customer_id
      - o Data type : varchar(50)
      - o Value : unique customer id
    - Loan_amount
      - o Data type :integer
      - o Value : from 100000 to 1000000
    - Benefit
      - o Data type :integer
      - o Value : from 20000 to 200000
    - Number_of_years
      - o Data type :integer
      - o Values : numbers as (3- 4- 5)
    - Date
      - o Data type : varchar(50)
      - o Value: from 24/11/2017 to 24/11/2018
  - The execution of loan stream:


```
SELECT sum(uo.Benefit )
  FROM loan AS uo
  INNER JOIN credit AS o ON uo.customer_id = o.Customer_ID
  WHERE o.Annual_Income between 500000 and 1000000 ;
```

Figure No. 4.59 Loan stream

84

- Customer value calculated for the three segmentation as follow:
  - Low income:
    - o The value per year = 860000 SDG
    - o The value per quarter =215000 SDG
    - o The percentage of low income revenue from hole revenue per quarter=30.4%
  - Medium income:
    - o The value per year = 1660000 SDG
    - o The value per quarter =415000 SDG
    - o The percentage of low income revenue from hole revenue per quarter=58.8%
  - High income:
    - o The value per year = 300000 SDG
    - o The value per quarter =75000 SDG
    - o The percentage of low income revenue from hole revenue per quarter=10.6%

4.6 Summary

The chapter covers the execution of the model and implementing the multidimensional model across the data that collected and measuring the customer value in different bank streams.

Chapter 5   Conclusion

5.1 Conclusion

This research discussed the growing of technology and its impact in the business process across the market. Also discussed the competition that inspired between the companies, organizations and competitors that affect the business process.

The main part of this process is the customer, so with that customer the business can success or fail. Thus, we need to focus on customer and measuring its value by looking to the customer in multidimensional view. The main limitation of Customer Relationship management (CRM) is looking to the customer in one dimension.

In this situation, business intelligence needed to measure the customer value. One of the technologies that apply the business intelligence is the data warehouse, which used in reporting and data analysis. For multidimensional view, we use OLAP technology to analysis data. In addition, we made a data mart, which built for a specific task.

The result of the research that we reached demonstrate that dealing with data in multidimensional view generates accurate reports that help in decision-making.

5.2 FUTURE WORK

The proposed framework method based on relational database in organizations. The main future works is to apply and extend the proposed framework on web based application project.