

# Telecom Business Intelligence Analysis & Reporting



BY

**ALI YADULLAH**

**Institute of Information technology**

**QUAID I AZAM UNIVERSITY ISLAMABAD**

**2014-2018**

*Dedicated to*

*My Parents*

*&*

*My brother Faiz Hassan*

*for their endless support, affection, trust and encouragement*



# STATEMENT OF SUBMISSION

This is to certify that Mr. Ali Yadullah Registration No. 04161413007 has successfully completed the final project as “*Telecom BI Analysis & Reporting*” Quaid-i-Azam University, Islamabad to fulfill the partial requirement of the degree “*Bachelor of Sciences in Information Technology*”.

External Examiner

Dr. Malik Ahmad Kamran  
Assistant Professor

Department Computer Sciences, COMSATS  
Islamabad.

Internal Examiner

Ms. Sidra Batool Kazmi  
Lecturer

Institute of Information Technology,  
Quaid-i-Azam University, Islamabad.

## **Acknowledgement**

First and foremost, I thank **Allah Almighty** for His countless blessing and for giving me courage, knowledge and ability to complete this work. My family's support, love and guidance were with me whatever I pursue.

I would like to express my deepest gratitude to **my parents, my father** for a lot support throughout my academic career and my mother for her love, affection, patience, encouragement and prayers.

I express my gratitude to my kind supervisor **Miss Sidra Batool Kazmi** who kept my moral high by appreciation and motivation. I am extremely fortunate to have her as my supervisor.

I am hugely grateful to my external supervisor Mr Gohar Abbas who had always been pushing me to showcase my talent to my fullest. He was always available whenever I tried to approach him. His valuable suggestions significantly improved this project and without his precious guidance I would have never been able to develop such a system. Special thanks to Samiullah Roman.

It is my bounden duty to pay tributes to my worthy teachers and staff members of Institute of Information Technology, Quaid-i-Azam University Islamabad for their help and support to develop an understanding of the project.

Last but not least I wish to avail myself of this opportunity, express a sense of gratitude and love to my beloved fellows either seniors, juniors, classmates, or university fellows for their moral, manual support, strength, and help and for everything which they done for me.

**ALI YADULLAH BANGASH**

## **Abstract**

Telecom business intelligence analysis and reporting is a data warehouse & business intelligence project. This project is developed to give Telecom Company a way of taking better business decision in order to improve their business. Application takes the telecom customers data and transform, clean and load into the data warehouse. The data in the data warehouse in being stored in facts and dimension tables following the data warehouse architecture schema in star and snowflake schema. Analysis will be carrying out on the data this lead to data visualization. Visual dashboard and report will be generated on that data. These visualized data reports then presented to the company manger, end users. From these visualized reports the company takes better business decisions. These decisions lead a company to improve their business and earn profit.

## **Preface**

Chapter 1 includes the introduction to the project, existing system, and proposed solution, scope of the system, project objectives and risk analysis. Brief description about software and hardware resources used.

Chapter 2 includes requirement analysis, functional and nonfunctional requirement. It also include system requirements of the project

Chapter 3 includes a system design, system architecture, use case diagram, activity diagram and data warehouse schemas such as star schema and snowflake schema

Chapter 4 includes a brief overview of tools and technologies used in order to implement the project

Chapter 5 includes the major component and activates to implement the project

Chapter 6 include test scenario and test cases of the system

Chapter 7 includes conclusion to this project and future work regarding this project.



## Contents

1.1 Introduction.....	11
1.2 Existing System.....	12
1.3 Proposed System.....	13
1.3.1 Advantages.....	13
1.4 Objectives.....	14
1.5 Scope.....	14
1.6 Resource Identification.....	15
1.6.1 Software Resources.....	15
1.6.2 Hardware Resources.....	15
1.7 Risk Analysis.....	16
1.7.1 When to use Risk Analysis.....	16
2 Requirement Analysis.....	17
2.1 Introduction.....	17
2.2 Requirement Definition.....	17
2.3 Requirement Elicitation.....	18
2.4 Requirement Analysis.....	18
2.5 Requirement Specification.....	18
2.6 Requirement Validation.....	18
2.7 Functional Requirement.....	19
2.8 Non Functional Requirement.....	20
2.9 System Requirement.....	21
2.9.1 User Interface Requirement.....	21
2.9.2 Software Requirement.....	21
2.9.3 Hardware Requirement.....	21
3 System Analysis and Design.....	22
3.1 Introduction.....	22
3.2 Analysis.....	22

3.3 Design .....	22
3.4 System Architecture.....	23
3.5 Use case diagrams.....	24
3.5.1 Elements of the use case diagram.....	24
3.6 Activity Diagram .....	26
3.7 Star Schema .....	29
3.8 Snowflake Schema.....	30
4 Tools and Languages .....	31
4.1 Microsoft Power BI .....	31
4.1.1 Requirement for MS Power BI.....	32
4.2 Microsoft Excel.....	32
4.2.1 Requirement for MS Excel .....	33
4.3 SQL .....	33
4.4 Microsoft Word.....	33
4.5 Microsoft Visio .....	34
4.6 VMware Workstation Pro:.....	34
4.7 Teradata Administrator .....	35
4.8 Teradata SQL Assistant .....	35
5 System Implementation .....	36
5.1 Introductions .....	36
5.2 Objective of Implementation Phase.....	36
5.3 Project Objectives .....	37
5.4 Getting Data.....	38
5.5 Data warehouse.....	39
5.5.1 How Data Warehouse Works .....	39
5.5.2 Data warehouse Schema Architecture.....	40
5.3 ETL .....	40



5.3.1 Extracting the data .....	40
5.3.2 Transforming the data.....	40
5.3.3 Loading the data .....	41
5.3.4 How ETL Works.....	41
5.4 ETL and Business Intelligence .....	42
5.5 Data Analysis .....	42
5.6 Data Visualization.....	43
5.6.1 History of Data Visualization .....	43
5.6.2 How Data Visualization Works.....	44
5.6.3 Benefits of Data Visualization.....	44
5.6.4 Tool Used for Visualization .....	46
6 System Testing.....	48
6.1 Test Scenario & Test Cases .....	48
7 Conclusive Discussion and Future Work.....	52
7.1 Conclusion .....	52
7.2 Future Work.....	52

## List of Tables and Figures

Table and Figures	Page Number
Table 1.1 Software Resources	15
Table 1.2 Hardware Resources	15
Fig 3.1 System Architecture.	23
Fig 3.2 Use Case Diagram	25
Fig 3.3 Activity Diagram Of System	27
Fig 3.4 Activity Diagram Of Data Warehouse	28
Fig 3.5 Star Schema	29
Fig 3.6 Snow Flake Schema	30
Fig 5.1 Customers Data Sheet	38
Fig 5.2 ETL Architecture	41
Fig 5.3 Visualization Report	47

# CHAPTER 1

## 1.1 Introduction

This project is basically about business intelligence analysis and reporting .Business intelligence (BI) is a technology – driven process for analyzing data and presenting actionable information to help executives, managers and other corporate end users to make informed business decisions. BI encompasses a wide variety of tools, applications and methodologies that enable organizations to collect data from internal systems and external sources prepare it for analysis, develop and run queries against that data; and create reports, dashboards and data visualizations to make the analytical results available to corporate decision- makers as well as operational works. Who make decisions? People make decisions .How? People make decisions go all in or fold hold sell or buy...People also make poor decision. Intelligence is making decisions and this is the people who make decisions. As intelligence +information= right decisions. Information helped provide answers to unlock secrets answers from answers people to make decisions.

How did business people get answers?? They had something called data .They had business data about people, products and places. They also had questions. Which products are the best? How are my sales performing? How are my people doing? Before they could get the answers they needed insight into that data .otherwise it was guess work .Data has value, it was precious. Where do you put something precious? It needed to be stored. Before the 1960 data was stored in the filing cabinet. When computer arrives it started to get stored in different spaces like floppy, hard drive. But storing something precious that way was risky business and difficult to manage. Eventually data is stored in place name database which is invented by Edger Codd in 1969 which is about the same time we landed on the moon. A database provides to store the business data about people, product and places. But it wasn't that straight forward to get the data in here required a lot of expertise. So in 1970 business application were created. Why business applications? because they provide the better why to collect data .The data access in not easy because data in coming from the multiple sources so for that data should be organized so they moved the data into dataware house.so now the data could be organized coming from multiple locations organized and able to access and now the data could be served .I will take the unorganized data from the database convert that into information .that information will be as a result of performing a lot of analysis on that

data. Further the information will be taken to the next step which is the reporting stage ... Making reports on that data will executives, managers and other corporate end users to make informed business decisions. That decisions will their product better and increase their sale helping make a lot of money and eliminate the risk and also make better decisions in the future for their healthy business.

## 1.2 Existing System

In today world most of the business applications are operational systems based on traditional transaction system or co called OLTP system (online transaction processing). An OLTP system is a common data processing system in today's enterprises. These systems have day to day transactions data which keep changes on everyday basis. In my existing telecom system database there is detailed and current data, and schema used to store transactional databases is the entity model (usually 3NF).Data is placed in *normalize form* for the sake of efficient transaction processing. Suppose the telecom company wants to expend their business and need to take better business decision in order to get more benefit and earn more profit so that business decision will not carry out on that day to day data because in traditional transaction system database there are number of table joins, performing analytical queries is very complex. Reporting is typically limited to more static, siloed needs. For example, you might generate a monthly report of the telecom company churn user. These reports are helpful— particularly for real-time reporting for bedside care—but they don't allow in-depth analysis. So in order to cover these limitations I proposed a new way of data analysis which may lead us to take better business decisions for that particular company. When the company takes better business decision surely they got benefit

## 1.3 Proposed System

The proposed system is based on business intelligence whereas in business intelligence we have to analyze the data and presenting actionable information to help executives, managers and other corporate end users to make informed business decisions. **OLAP (Online Analytical Processing)** is the technology behind many Business Intelligence (BI) applications. OLAP data comes from the various OLTP databases. We extract data from the telecom company different OLTP databases we than transform and standardized that data and the import to the OLAP data warehouse and the run different queries on that data then we take data into business intelligence application for data analysis and data visualization and then create reports, dashboards to make the analytical result based on that reports the company take better business decisions

### 1.3.1 Advantages

Here are different ways business intelligence can enhance an organization and help decision makers make more informed, effective choices.

- Derive knowledge from a sea of data.
- Understand what drives revenue for your business.
- Personalize your sales strategy and anticipate objections.
- Create a more effective business model
- Get a single, consolidated view of your customers..
- Discover the big-picture "why" and what motivates consumer behavior.

## **1.4 Objectives**

The basic objective of the project is to develop a system which takes the operational or traditional transactional data of the telecom company and transform that data to analytical data then perform data analysis and data visualization, create intelligent reports and unique dashboards for the telecom company based on that reports the company take better business decision. Once we perform data analysis and run different queries with the respect to company need and then based on these analysis daily basic reports or weekly may be monthly report will be generated Based on these reports intelligent business decision can make the company earn profit and decrease the loss generates revenue. Increase number of customer and even facilitate their churn user to rejoin.

## **1.5 Scope**

The project is purely a business intelligence project and its basic functionality is to provide the user and the telecom company a way to take better decision based on their own data this decision can increase their profit or grow their business or fill-up their dreams. The project scope is limited to reporting and data visualization, data analysis based on the historic or previous data of the telecom company. The system takes the company data as an input these might stores at different databases, flat files or even excel spreadsheets then extract and clean that data to store in the data warehouse then analysis and is performed on that data, analytical data then convert to a report or a dashboard which make a user able to understand and make good business decisions. The good business decision help the organization to increase their revenue earn more profit and get rid of business losses. So these business intelligence report help them to fill full their dreams

## 1.6 Resource Identification

Following resources will be used in order to build this particular project

### 1.6.1 Software Resources

	Operating system	Microsoft Windows 8
	Development tools	Microsoft Power BI, Microsoft Excel, VMware workstation pro, Teradata Administrator, Teradata SQL Assistant
	Documentations	Microsoft Word, Microsoft Visio
	Language	SQL
	Data warehouse	Teradata Administrator

Table 1.1

### 1.6.2 Hardware Resources

1	System	Intel(R) Core™ i5-3317U
2	Processor	1.70Ghz
3	Hard Disk	500GB
4	RAM	8.00GB
5	System type	64 bit operating system, x64-based processor

Table 1.2

## 1.7 Risk Analysis

Risk is made up of two parts: the probability of something going wrong, and the negative consequences if it does. Risk Analysis is a process that helps you identify and manage potential problems that could undermine key business initiatives or projects.

To carry out a Risk Analysis, you must first identify the possible threats that you face, and then estimate the likelihood that these threats will materialize

### 1.7.1 When to use Risk Analysis

Risk analysis is useful in many situations:

- When you're planning projects, to help you anticipate and neutralize possible problems.
- When you're deciding whether or not to move forward with a project.
- When you're improving safety and managing potential risks in the workplace.
- When you're preparing for events such as equipment or technology failure, theft, staff sickness, or natural disasters..

As far as this project is concern in my opinion following error can occur and we also use counter measures to eliminate this error in order to get better performance of a system.

- Corrupt data: there must be a situation when the user or company data is corrupt .So we cannot understand the corrupt data. To encounter this we simply need new user data
- Data duplication: there exists duplication in the data, this case analysis issue. To encounter this we have to clean the data and transform the data
- Data lose: we can face the situation where we lost the data unintentionally. We can retrieve these data by running the backup files for that data



## CHAPTER 2

### 2 Requirement Analysis

#### 2.1 Introduction

Requirements analysis, also called requirements engineering is the process of determining user expectations for a new or modified product. These features, called requirements, must be quantifiable, relevant and detailed. Requirement analysis involves frequent communication with system users to determine specific feature expectations, resolution of conflict or ambiguity in requirements as demanded by the various users or group of users, avoidance of feature creep and documentation of all aspects of the project development process from start to finish. Energy should be directed towards the final product which conforms to client needs rather than attempting to mold user expectations to fit the requirements. . If the requirements are not completely defined then there might be inconsistencies in the final product. The most errors in code are detected after development and testing and mostly half of them are in requirements and design. To prevent these errors, we analyze each requirement by requirement analysis process in this chapter the requirements of the proposed system are discussed in detail to achieve the goals and objectives of the system.

#### 2.2 Requirement Definition

Requirements definition is the most crucial part of the project. Incorrect, inaccurate, or excessive definition of requirements must necessarily result in schedule delays, wasted resources, or customer dissatisfaction. The requirements analysis should begin with business or organizational requirements and translate those into project requirements. If meeting stated requirements will be unreasonably costly, or take too long, the project requirements may have to be negotiated down, down-scoped or down-sized, in discussions with customers or sponsors. Any discussion of requirements analysis methods will quickly become specific to the type of project effort. Many industry areas have specific, proven techniques for obtaining thorough and accurate definition of requirements. Sometimes it is useful to write a draft user's manual as a way to define requirements. While the methods may differ, the principles remain the same across all types and sizes of projects.

The requirements analysis should cover the whole scope of the project. It must be comprehensive and thorough. It must consider the views and needs of all the project stakeholders. It is easy to leave scope out of a requirements analysis or to omit necessary clarity or detail thereby making the requirements definition ambiguous. The completed requirements analysis should be reviewed and approved by the customer or project sponsor before work continues. In requirement analysis first we clearly define the requirements of proposed system. There are two types of requirements. One is called functional requirements and the other is called non-functional requirements. The main functionalities that system should perform are functional requirements. Non-functional are those that the customer does not demand but the programmer and developer should keep in mind to fulfill the quality aspects of the system like usability and efficiency etc.

**2.3 Requirement Elicitation:** The elicitation activity consists of gathering information, understanding the stakeholders need, and articulating high level requirements.

**2.4 Requirement Analysis:** The analysis activity examines the high level requirements and determines if they are clear, complete, and free of contradictions, and then define the strategy to address these issues.

**2.5 Requirement Specification:** The specification activity defines the behavior of a system in development and determines the method for requirements documentation (i.e., natural-language documents, process models, business definitions, use cases, user stories, or process specification)

**2.6 Requirement Validation:** The validation activity involves sessions with users, stakeholders, and functional experts to determine mitigation and issue resolving plans for conflicting requirements before projects move into the development phase.

## 2.7 Functional Requirement

The Functional Requirements Specification documents the operations and activities that a system must be able to perform. The Functional Requirements Specification is designed to be read by a general audience. Readers should understand the system, but no particular technical knowledge should be required to understand the document.

Functional requirements may include, technical details, data manipulation and processing, calculations and other specific functionality that delineates what a system made-up to achieve.

Following are the functional requirements of above described system. Functional requirements are supported by non-space functional requirements (also known as quality requirements), which impose constraints on the design or requirements are expressed in the form "system must do requirement".

- Data Extraction: data extraction will be from different telecom databases, flat files , excel spreadsheet or from other different sources
- Data Staging area (DSA): we also call it ETL (Extract , transform and load) the data from the different data sources than propagated to the DSA where it transformed and cleaned up and converted into the standardized data before loaded to the data warehouse.
- Data warehouse: cleaned data then loaded to data warehouse, data warehouse work as repository for storing data from the different sources the data stores in the data warehouse in form of fact and dimension table.
- Data Analysis: data then processed for data analysis, different queries will be performed on that data. Meaning full analysis will be perform on the data in order to better understanding.
- Reporting: data then load to Business intelligence application where different reports and dashboards will be created from that data. These reports and dashboard give help to a company to take better business decisions.
- Report Presentation: the visualized report is the ready for presentation, we can adopt different method of presenting these report. We can present it on web application or on the excel spread sheet our even on our mobile device

## 2.8 Non Functional Requirement

The definition for a non-functional requirement is that it essentially specifies how the system should behave and that it is a constraint upon the systems behavior. One could also think of non-functional requirements as quality attributes for of a system.

Non-functional requirements cover all the remaining requirements which are not covered by the functional requirements. Non-functional requirements essentially specify how the system should behave and that it is a constraint upon the system's behavior. They specify benchmarks that judge the operation of a system, rather than specific behaviors.

Every prerequisite must be objective; there must be some quantifiable approach to evaluate whether the necessity has been met. Following are some non-functional requirements

- **Performance:** The system must perform the core functionalities successfully.
- **Reliability:** System is sufficiently reliable and it performs all its planned capacities and operations in its domain, without experiencing any failure.
- **Efficiency:** Efficiency of the system defines satisfaction of purpose without misuse of assets, for example, memory, space and processor usage, system transfer speed and time.
- **Usability:** Usability refers to convenience and practicality of use.
- **Feature Enhancement ability:** Feature Enhancement has its own importance, because with the passage of time system evolves and there should be a need for more functions, so system should allow future Enhancements. This functionality may be achieved by system design. System should be designed in such a way that it supports the enhancement
- **Multi-Platform Delivery:** The use of application is not limited to laptop or personal system user can use this BI application on his/her android and iOS devices



## **2.9 System Requirement**

Following is the system requirement to complete this project

### **2.9.1 User Interface Requirement**

- GUI(Graphical user interface ) along with meaningful frames and interface

### **2.9.2 Software Requirement**

- Operating system will be required. Windows operating system 8 and above

### **2.9.3 Hardware Requirement**

Minimum hardware requirements are:

- 1.5GHz CPU
- 2 GB RAM
- 256 HARD DISK
- 1024\*1024 of display

# CHAPTER 3

## 3 System Analysis and Design

### 3.1 Introduction

System analysis and design deal with planning the development of information systems through understanding and specifying in detail what a system should do and how the components of the system should be implemented and work together. System analysts solve business problems through analyzing the requirements of information systems and designing such systems by applying analysis and design

### 3.2 Analysis

This section addresses security considerations. Key security activities include

- Conduct the risk assessment and use the results to supplement the baseline security controls;
- Analyze security requirements;
- Perform functional and security testing
- Prepare initial documents for system certification and accreditation

Although this section presents the information security components in a sequential top-down manner, the order of completion is not necessarily fixed. Security analysis of complex systems will need to be iterated until consistency and completeness is achieved.

### 3.3 Design

The process of defining the architecture, components, and data of a system to satisfy specified requirements in order to complete a project

### 3.4 System Architecture

System architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system. A system architecture can comprise system components, the expand systems developed, that will work together to implement the overall system. System architecture conveys the informational content of the elements comprising a system, the relationships among those elements, and the rules governing those relationships. The architectural components and set of relationships between these components that an architecture description may consist of hardware, software, documentation, facilities, manual procedures, or roles played by organizations or people.

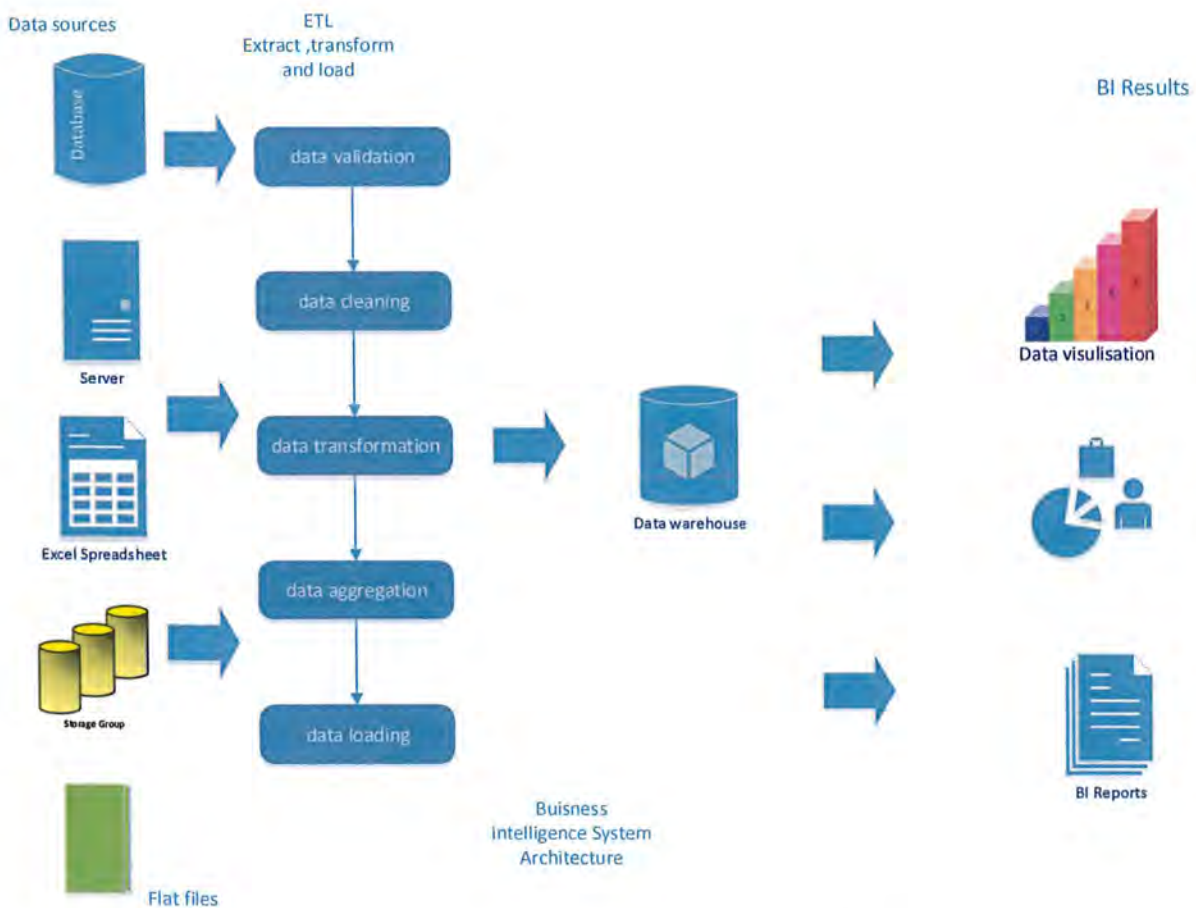


Fig 3.1 System Architecture Diagram

### 3.5 Use case diagrams

Use case diagrams are usually referred to as behavior diagrams used to describe a set of use cases that some system or systems should or can perform in collaboration with one or more external users of the system. Each use case should provide some observable and valuable result to the actors or other stakeholders of the system.

Use case diagrams are used to specify:

- External needs on a subject, mandatory usages of a system - to capture what a system under construction is meant to do
- The functionality offered by a subject – what the system can do
- The requirements, which the specified subject poses on its environment - by defining how environment should interact with the subject so that it will be able to perform its services

#### 3.5.1 Elements of the use case diagram

**Actor:** An actor portrays any entity that performs certain roles in a given system. It is, in a use case diagram interacts with a use case.

**Use case:** A use case, in a use case diagram, is a visual depiction of distinct business functionality in a system.

**System boundary:** A system boundary defines what will be the scope of system. A system cannot have infinite functionality as it defines the limits of the system.

**Relationships:** A relationship between two use cases is basically a dependency between the two use cases. The types of relationships are: Include, Extend, Generalization etc.



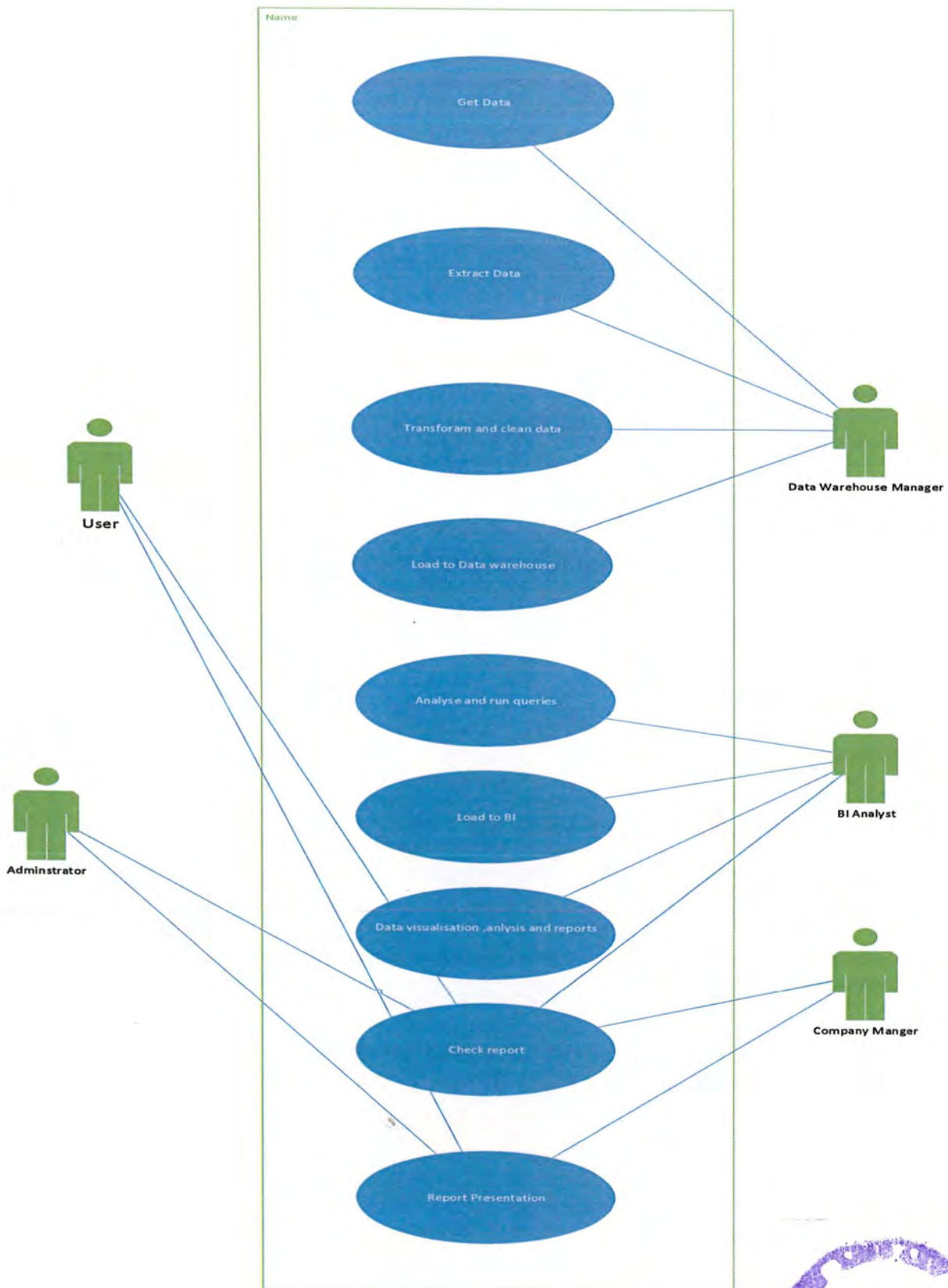


Fig 3.2 Use case diagram of system



### 3.6 Activity Diagram

Activity diagrams, which are related to program flow plans (flowcharts), are used to illustrate activities. In the external view, we use activity diagrams for the description of those business processes that describe the functionality of the business system.

Contrary to use case diagrams, in activity diagrams it is obvious whether actors can perform business use cases together or independently from one another.

Activity diagrams allow you to think functionally. Purists of the object-oriented approach probably dislike this fact. We, on the other hand, regard this fact as a great advantage, since users of object-oriented methods, as well as users of functional thinking patterns, find a common and familiar display format, which is a significant aid for business-process modeling.

Following are the notations used for activity diagrams:

- **Activity:** The rounded rectangles represent activities that occur.
- **Initial Node:** The filled in circles is the starting point of the diagram.
- **Final Node:** The filled circle with a border is the ending point.
- **Fork:** A black bar with one flow going into it and several leaving it
- **Join:** A black bar with several flows entering it and one leaving it.
- **Decision:** A diamond with one flow entering and several leaving.
- **Merge:** A diamond with several flows entering and one leaving.
- **Flow Final:** The circle with the X through it. This indicates that the process stops at this point.



Fig 3.3 Activity diagram of System

### 3.7 Star Schema

In data warehousing and business intelligence (BI), a star schema is the simplest form of a dimensional model, in which data is organized into *facts* and *dimensions*. A fact is an event that is counted or measured, such as a sale or login. A dimension contains reference information about the fact, such as date, product, or customer. A star schema is diagramed by surrounding each fact with its associated dimensions. The resulting diagram resembles a star.

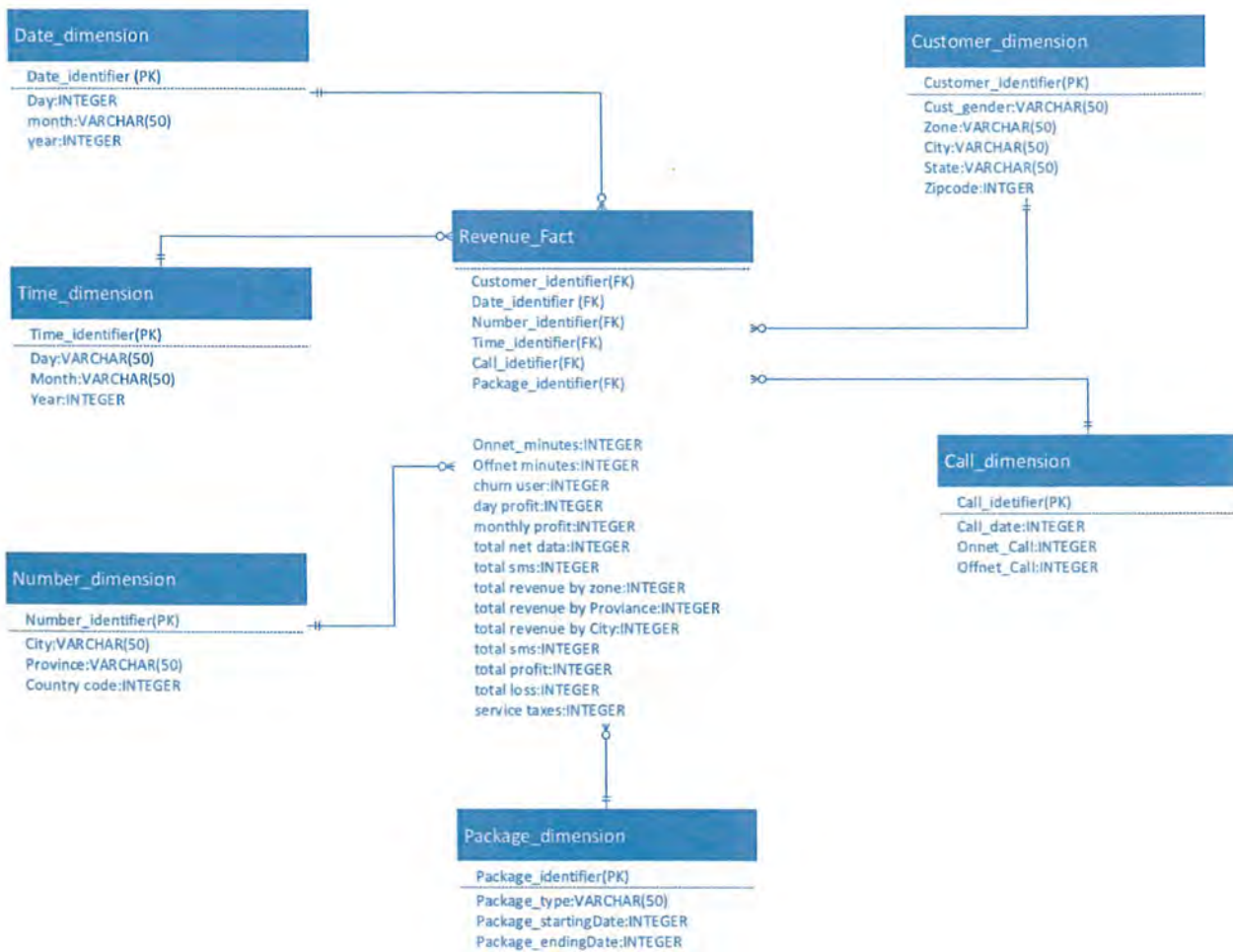


Fig 3.5 Star Schema of data warehouse

### 3.8 Snowflake Schema

This snowflake schema stores exactly the same data as the star schema. The fact table has the same dimensions as it does in the star schema example. The most important difference is that the dimension tables in the snowflake schema are normalized. Interestingly, the process of normalizing dimension tables is called snow flaking.

Once again, visually the snowflake schema reminds us of its namesake, with several layers of dimension tables creating an irregular snowflake-like shape.

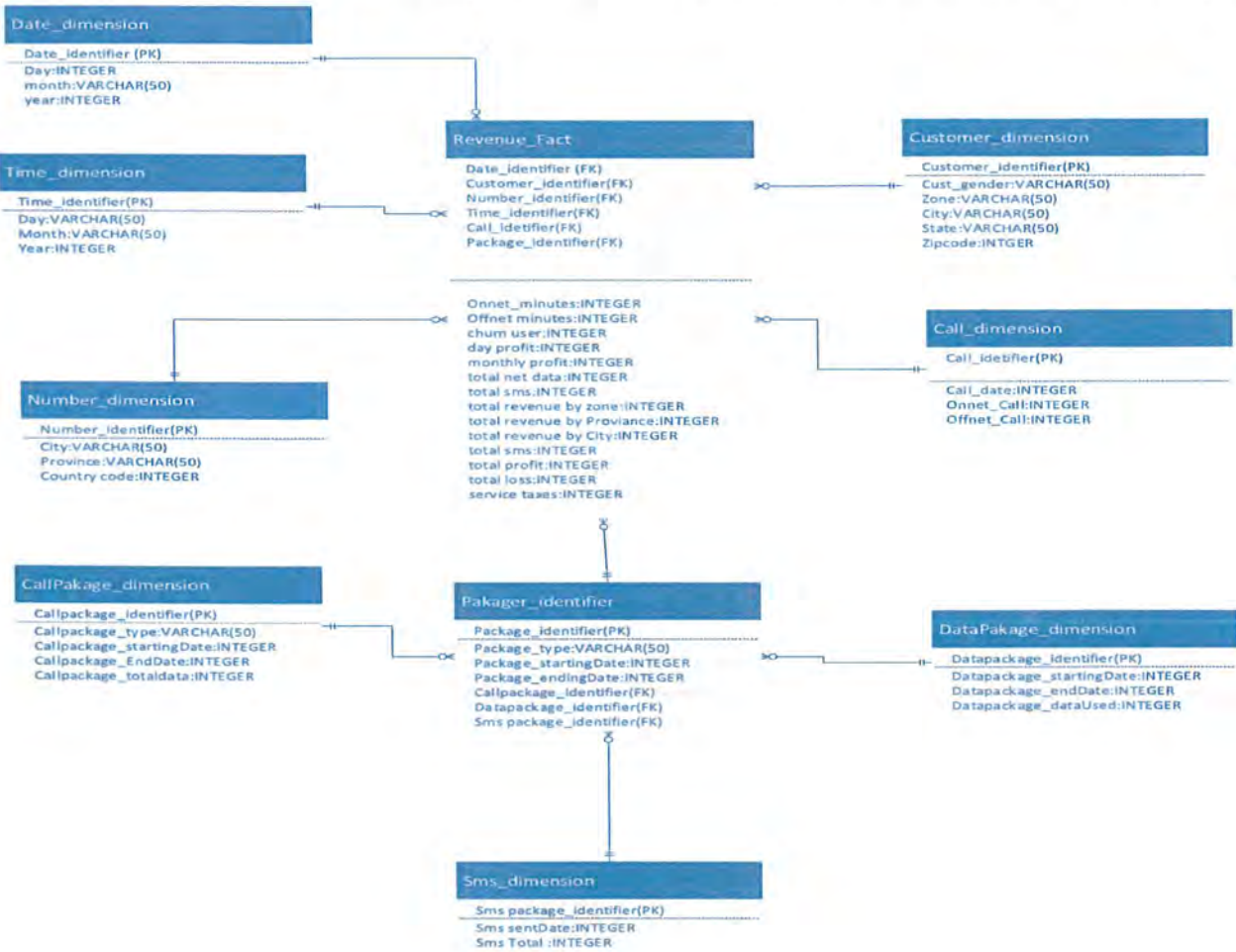


Fig 3.6 Snowflake schema of data warehouse



## CHAPTER 4

### 4 Tools and Languages

Following are the tools and technology which are being used for the development of the system

- Microsoft Power BI
- Microsoft Excel
- SQL
- Microsoft word
- MS Visio
- VMware workstation pro
- Teradata Administrator
- Teradata SQL Assistant

#### 4.1 Microsoft Power BI

Power BI is a business analytics service provided by Microsoft. It provides interactive visualizations with self-service business intelligence capabilities, where end users can create reports and dashboards by themselves, without having to depend on information technology staff or database administrators. Power BI provides cloud-based BI services, known as "Power BI Services", along with a desktop based interface, called "Power BI Desktop". It offers data warehouse capabilities including data preparation, data discovery and interactive dashboards.

### 4.1.1 Requirement for MS Power BI

The following list provides the minimum requirements to run **Power BI**

- Windows 7 / Windows Server 2008 R2, or later
- Internet Explorer 9 or later
- **Memory (RAM):** At least 1 GB available, 1.5 GB or more recommended
- **Display:** At least 1440x900 or 1600x900 (16:9) recommended. Lower resolutions such as 1024x768 or 1280x800 are not recommended, as certain controls (such as closing the startup screen) display beyond those resolutions.
- **CPU:** 1 gigahertz (GHz) or faster x86- or x64-bit processor recommended

### 4.2 Microsoft Excel

Microsoft Excel is a spreadsheet developed by Microsoft for Windows, macOS, Android and iOS. It features calculation, graphing tools, pivot tables Microsoft Excel has the basic features of all spreadsheets, using a grid of *cells* arranged in numbered *rows* and letter-named *columns* to organize data manipulations like arithmetic operations In addition, it can display data as line graphs, histograms and charts, and with a very limited three-dimensional graphical display or in general, as a design tool that asks the user questions and provides answers and reports

## 4.2.1 Requirement for MS Excel

- **CPU:** 1 gigahertz (GHz) or faster x86- or x64-bit processor with SSE2 instruction set
- **RAM:** 1 gigabyte (GB) RAM (32-bit); 2 gigabytes (GB) RAM (64-bit)
- **HARD DISK:** 3.0 gigabytes (GB) available
- **DISPLAY:** Graphics hardware acceleration requires a DirectX10 graphics card and a 1024 x 576 or higher resolution monitor

## 4.3 SQL

Structured Query Language is a domain-specific language used in programming and designed for managing data held in a relational database management system (RDBMS), or for stream processing in a relational data stream management system (RDSMS). SQL offers two main advantages: first, it introduced the concept of accessing many records with one single command; and second, it eliminates the need to specify *how* to reach a record, e.g. with or without an index. The scope of SQL includes data query, data manipulation (insert, update and delete), data definition (schema creation and modification), and data access control. Although SQL is often described as, and to a great extent is, a declarative language (4GL), it also includes procedural elements.

## 4.4 Microsoft Word

Microsoft Word is a word processor developed by Microsoft. Word contains rudimentary desktop publishing capabilities and is the most widely used word processing program on the market. Word files are commonly used as the format for sending text documents via e-mail because almost every user with a computer can read a Word document by using the Word application. All my project documentation is carried out on Microsoft word.



## 4.5 Microsoft Visio

Microsoft Visio (formerly Microsoft Office Visio) is a diagramming and vector graphics application and is part of the Microsoft Office family. Microsoft made Visio 2013 for Windows available in two editions: Standard and Professional. The Standard and Professional editions share the same interface, but the Professional edition has additional templates for more advanced diagrams and layouts, as well as capabilities intended to make it easy for users to connect their diagrams to data sources and to display their data graphically. The Professional edition features three additional diagram types, as well as intelligent rules, validation, and sub process (diagram breakdown). Visio Professional is also offered as an additional component of an Office365 subscription.

## 4.6 VMware Workstation Pro:

VMware Workstation is a hosted hypervisor that runs on x64 versions of Windows and Linux operating systems (an x86 version of earlier releases was available) it enables users to set up virtual machines (VMs) on a single physical machine, and use them simultaneously along with the actual machine. Each virtual machine can execute its own operating system, including versions of Microsoft Windows, Linux, BSD, and MS-DOS. VMware Workstation is developed and sold by VMware, Inc., a division of Dell Technologies. There is a free-of-charge version, VMware Workstation Player, for non-commercial use. An operating systems license is needed to use proprietary ones such as Windows. Ready-made Linux VMs set up for different purposes are available from several sources. VMware Workstation supports bridging existing host network adapters and sharing physical disk drives and USB devices with a virtual machine. It can simulate disk drives; an ISO image file can be mounted as a virtual optical disc drive, and virtual hard disk drives are implemented as .vmdk files. VMware Workstation Pro can save the state of a virtual machine (a "snapshot") at any instant. These snapshots can later be restored, effectively returning the virtual machine to the saved state,<sup>[5]</sup> as it was and free from any post-snapshot damage to the VM. VMware Workstation includes the ability to group multiple virtual machines in an inventory folder. The machines in such a folder can then be powered on and powered off as a single object, useful for testing complex client-server environments

## **4.7 Teradata Administrator**

Teradata Administrator is a client based graphical interface used to perform database administration tasks on Teradata and Aster Databases, as well as copying data between Aster, Teradata, and Hadoop Database Systems. Teradata Studio is built on the Eclipse® Rich Client Platform (RCP), taking advantage of the RCP framework for building and deploying native GUI applications. It extends the Eclipse Data Tools Platform (DTP) to enhance Teradata Database access. Teradata Studio can be run on Windows, Linux, and Mac OSX operating systems. The Eclipse platform is designed as an open tools platform but architected so that its components (or a subset of its components) can be used to build RCP applications. This allows Teradata Studio to benefit from the many high quality Eclipse features available while focusing on value-add for Teradata.

## **4.8 Teradata SQL Assistant**

Teradata SQL Assistant stores, retrieves, and manipulates data from Teradata Database (or any database that provides an ODBC interface). Teradata SQL Assistant for Microsoft Windows is a query and information discovery tool used to retrieve data from Teradata Database or any ODBC-compliant database server. ... .NET Data Provider for Teradata can also be used to connect to Teradata Database from Teradata SQL Assistant for Microsoft Windows.

# CHAPTER 5

## 5 System Implementation

### 5.1 Introductions

Implementation simply means carrying out the activities described in your work plan. Project implementation (or project execution) is the phase where visions and plans become reality. This is the logical conclusion, after evaluating, deciding, visioning, planning, applying for funds and finding the financial resources of a project. The implementation stage is the construction period of an application. The work exist without this phase.

We begin the development stage not long after analysis and the design is finished, we are clear about what we need to build and we know our requirements. All the thing, studies and steps, which we have been through, were the planning for the implementation stage. It is done with the help of tools, programming skills and techniques, which are pillars of our application.

### 5.2 Objective of Implementation Phase

The objectives of the implementation phase can be summarized as follow:

- Putting the action plan into operation
- Achieving tangible change and improvements
- Ensuring that new infrastructure, new institutions and new resources are sustainable in every aspect
- Ensuring that any unforeseen conflicts that might arise during this stage are resolved
- Ensuring transparency with regard to finances
- Ensuring that potential benefits are not captured by elites at the expenses of poorer social groups

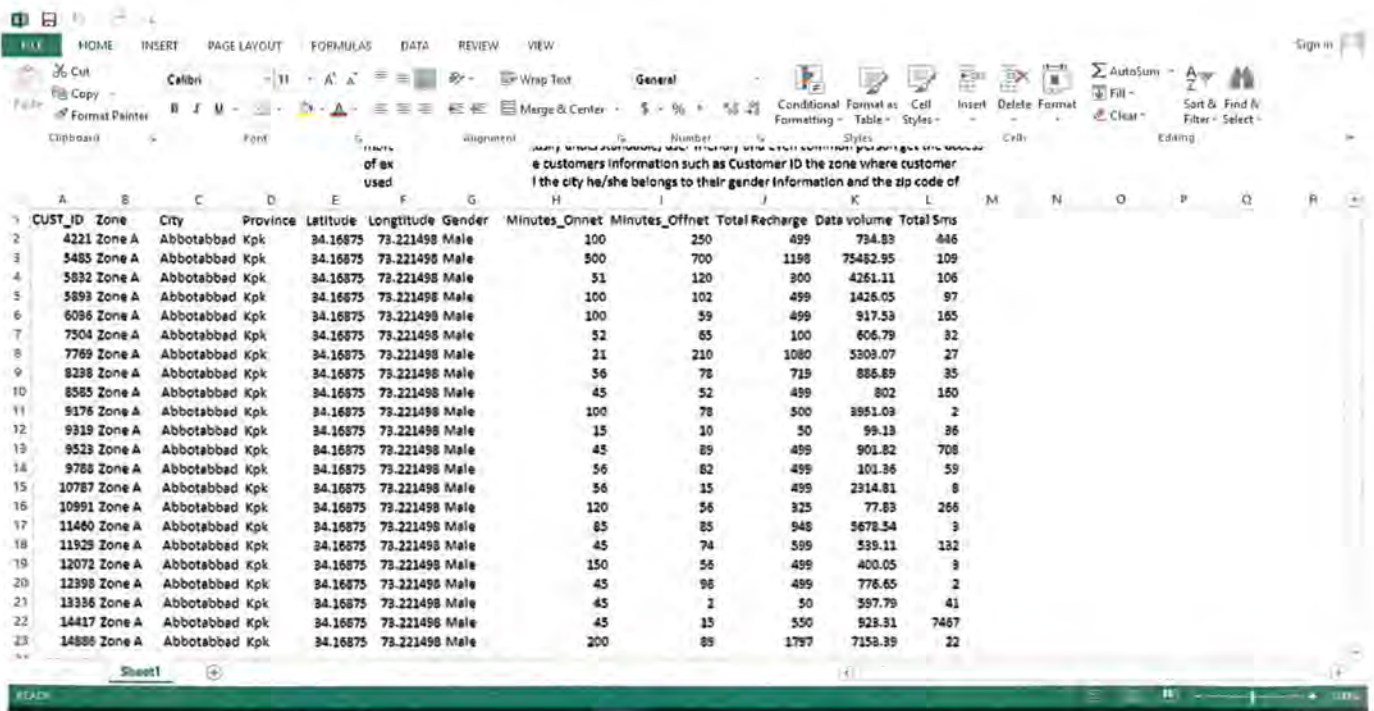
## 5.3 Project Objectives

Following are the basic objective of this phase:

- Getting the data. Data can be in different sources but we have data in excel format.
- Making a data warehouse
- Loading of data to the warehouse
- Analysis on the data
- Data visualization. Reports and dashboards of data.

## 5.4 Getting Data

The project deals with the data without the data we are not be able to implement what we are dreaming off. Data having different sources may be it comes from the operation databases, servers, storage group, flat files, excel sheet etc. We have the telecom data in the excel spreadsheet format. Basically that data is taken from the telecom company operational database and is in the excel format because I have not directly access to that database. That why we have the privilege to get the data in more likely format which is easily understandable, user friendly and even common person get the access of excel. The data contain the customers information such as Customer ID the zone where customer used to live the province and the city he/she belongs to their gender information and the zip code of his/her cities and most likely the recharge he/she makes throughout the month, the sms they send and the On net and Off net minutes they used and the mobile data they used in the whole month. These customer record and their data is in our custody. And most probably we are going to store these data in order to carry out analysis on that create visualized report in order to understand and take future decisions



CUST_ID	Zone	City	Province	Latitude	Longitude	Gender	Minutes_Onnet	Minutes_Offnet	Total Recharge	Data volume	Total Sms
4221	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	100	250	499	734.83	446
5485	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	900	700	1198	75482.95	109
5892	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	51	120	300	4261.11	106
5893	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	100	102	499	1426.05	97
6086	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	100	59	499	917.53	165
7504	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	52	65	100	606.79	32
7769	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	21	210	1080	5309.07	27
8298	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	56	78	719	886.89	35
8585	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	45	52	499	802	160
9176	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	100	78	500	3951.09	2
9319	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	15	10	50	99.13	36
9529	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	45	89	499	901.82	708
9768	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	56	82	499	101.36	59
10787	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	56	15	499	2314.81	8
10991	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	120	56	325	77.83	266
11460	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	85	85	948	5678.54	3
11929	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	45	74	599	539.11	132
12072	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	150	56	499	400.05	3
12398	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	45	98	499	776.65	2
13336	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	45	1	50	397.79	41
14417	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	45	15	550	923.31	7467
14880	Zone A	Abbotabbad	Kpk	34.16875	73.221498	Male	200	89	1797	7159.39	22

Fig 5.1 Customer Data Sheet

## **5.5 Data warehouse**

A data warehouse is a federated repository for all the data collected by an enterprise's various operational systems, be they physical or logical. Data warehousing emphasizes the capture of data from diverse sources for access and analysis rather than for transaction processing.

A data warehouse is a:

- Subject-oriented
- Integrated
- Time varying
- Non-volatile

### **5.5.1 How Data Warehouse Works**

A Data Warehouse works as a central repository where information arrives from one or more data sources. Data flows into a data warehouse from the transactional system and other relational databases.

Data may be:

- Structured
- Semi-structured
- Unstructured data

The data is processed, transformed, and ingested so that users can access the processed data in the Data Warehouse through Business Intelligence tools, SQL clients, and spreadsheets. A data warehouse merges information coming from different sources into one comprehensive database.

By merging all of this information in one place, an organization can analyze its customers more holistically. This helps to ensure that it has considered all the information available. Data warehousing makes data mining possible. Data mining is looking for patterns in the data that may lead to higher sales and profits.

## 5.5.2 Data warehouse Schema Architecture

Data Warehouse environment usually transforms the relational data model into some special

Architectures. There are many schema models designed for data warehousing but the most

Commonly used are:

- Star schema
- Snowflake schema
- Fact constellation schema

The determination of which schema model should be used for a data warehouse should be based upon the analysis of project requirements, accessible tools and project team preferences.

## 5.3 ETL

(Extract, Transform and Load) is a process in data warehousing responsible for pulling data out of the source systems and placing it into a data warehouse. During this process, data is taken (extracted) from a source system, converted (transformed) into a format that can be analyzed, and stored (loaded) into a data warehouse or other system ETL involves the following tasks:

### 5.3.1 Extracting the data

From source systems (SAP, ERP, other operational systems), data from different source systems is converted into one consolidated data warehouse format which is ready for transformation

### 5.3.2 Transforming the data

It may involve the following tasks:

- applying business rules (so-called derivations, e.g., calculating new measures and dimensions),
- Cleaning (e.g., mapping NULL to 0 or "Male" to "M" and "Female" to "F" etc.),

- Filtering (e.g., selecting only certain columns to load),
- Splitting a column into multiple columns and vice versa,
- Joining together data from multiple sources (e.g., lookup, merge),
- Transposing rows and columns,
- Applying any kind of simple or complex data validation (e.g., if the first 3 columns in a row are empty then reject the row from processing)

**5.3.3 Loading the data** Load the data into a data warehouse or data repository other reporting applications

### 5.3.4 How ETL Works

Data from one or more sources is extracted and then copied to the data warehouse. When dealing with large volumes of data and multiple source systems, the data is consolidated. ETL is used to migrate data from one database to another, and is often the specific process required to load data to and from data marts and data warehouses, but is a process that is also used to large convert (transform) databases from one format or type to another



Fig 5.2 ETL Architecture



## 5.4 ETL and Business Intelligence

ETL is an important part of today's business intelligence (bi). It is the IT process from which data from disparate sources can be put in one place to programmatically analyze and discover business insights.

## 5.5 Data Analysis

Data analysis is a primary component of data mining and Business Intelligence (BI) and is key to gaining the insight that drives business decisions. Organizations and enterprises analyze data from a multitude of sources using Big Data management solutions and customer experience management solutions that utilize data analysis to transform data into actionable insights. This is a process inspecting, cleansing, transforming, and modeling data with the goal of discovering useful information, informing conclusions, and supporting decision-making. Data analysis is a process for obtaining raw data and converting it into information useful for decision-making by users. Data is collected and analyzed to answer questions, test hypotheses or disprove theories

Data analysis involves asking questions about what happened, what is happening, Data analysis is a proven way for organizations and enterprises to gain the information they need to make better decisions, serve their customers, and increase productivity and revenue. The benefits of data analysis are almost too numerous to count, and some of the most rewarding benefits include getting the right information for your business, getting more value out of IT departments, creating more effective marketing campaigns, gaining a better understanding of customers, and so on.

But, there is so much data available today that data analysis is a challenge. Namely, handling and presenting all of the data are two of the most challenging aspects of data analysis. Traditional architectures and infrastructures are not able to handle the sheer amount of data that is being generated today, and decision makers find it takes longer than anticipated to get actionable insight from the data.

Fortunately, data management solutions and customer experience management solutions give enterprises the ability to listen to customer interactions, learn from behavior and contextual

information, create more effective actionable insights, and execute more intelligently on insights in order to optimize and engage targets and improve business practices.

## **5.6 Data Visualization**

Data visualization is a general term that describes any effort to help people understand the significance of data by placing it in a visual context. Patterns, trends and correlations that might go undetected in text-based data can be exposed and recognized easier with data visualization software. Data visualization refers to the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization is an accessible way to see and understand trends, outliers, and patterns in data. It enables decision makers to see analytics presented visually, so they can grasp difficult concepts or identify new patterns. With interactive visualization, you can take the concept a step further by using technology to drill down into charts and graphs for more detail, interactively changing what data you see and how it's processed.

Today's data visualization tools go beyond the standard charts and graphs used in Microsoft Excel spreadsheets, displaying data in more sophisticated ways such as info graphics, dials and gauges, geographic maps, spark lines, heat maps, and detailed bar, pie and fever charts. The images may include interactive capabilities, enabling users to manipulate them or drill into the data for querying and analysis. Indicators designed to alert users when data has been updated or predefined conditions occur can also be included.

### **5.6.1 History of Data Visualization**

The concept of using pictures to understand data has been around for centuries, from maps and graphs in the 17th century to the invention of the pie chart in the early 1800s. Several decades later, one of the most cited examples of statistical graphics occurred when Charles Minard mapped Napoleon's invasion of Russia. The map depicted the size of the army as well as the path of Napoleon's retreat from Moscow – and tied that information to temperature and time scales for a more in-depth understanding of the event.

It's technology, however, that truly lit the fire under data visualization. Computers made it possible to process large amounts of data at lightning-fast speeds. Today, data visualization has become a rapidly evolving blend of science and art that is certain to change the corporate landscape over the next few years

### **5.6.2 How Data Visualization Works**

Most of today's data visualization tools come with connectors to popular data sources, including the most common relational databases, data warehouses and a variety of cloud storage platforms. The visualization software pulls in data from these sources and applies a graphic type to the data.

Data visualization software allows the user to select the best way of presenting the data, but, increasingly, software automates this step. Some tools automatically interpret the shape of the data and detect correlations between certain variables and then place these discoveries into the chart type that the software determines is optimal.

Typically, data visualization software has a dashboard component that allows users to pull multiple visualizations of analyses into a single interface, generally a web portal.

### **5.6.3 Benefits of Data Visualization**

Our eyes are drawn to colors and patterns. We can quickly identify red from blue, square from circle. Our culture is visual, including everything from art and advertisements to TV and movies.

Data visualization is another form of visual art that grabs our interest and keeps our eyes on the message. When we see a chart, we quickly see trends and outliers. If we can see something, we internalize it quickly. It's storytelling with a purpose. If you've ever stared at a massive spreadsheet of data and couldn't see a trend, you know how much more effective a visualization can be.

### **Common general types of data visualization:**

- Charts
- Tables
- Graphs
- Maps
- Info graphics
- Dashboards
- **More specific examples of methods to visualize data:**
- Area Chart
- Bar Chart
- Box-and-whisker Plots
- Bubble Cloud
- Bullet Graph
- Cartogram
- Circle View
- Dot Distribution Map
- Gantt Chart
- Heat Map
- Highlight Table

- Histogram
- Matrix
- Network
- Polar Area
- Radial Tree
- Scatter Plot (2D or 3D)
- Stream graph
- Text Tables
- Timeline
- Tree map
- Wedge Stack Graph
- Word Cloud
- And any mix-and-match combination in a dashboard!

#### **5.6.4 Tool Used for Visualization**

Power BI is a business analytics service provided by Microsoft. It provides interactive visualizations with self-service business intelligence capabilities, where end users can create reports and dashboards by themselves, without having to depend on information technology staff or database administrators

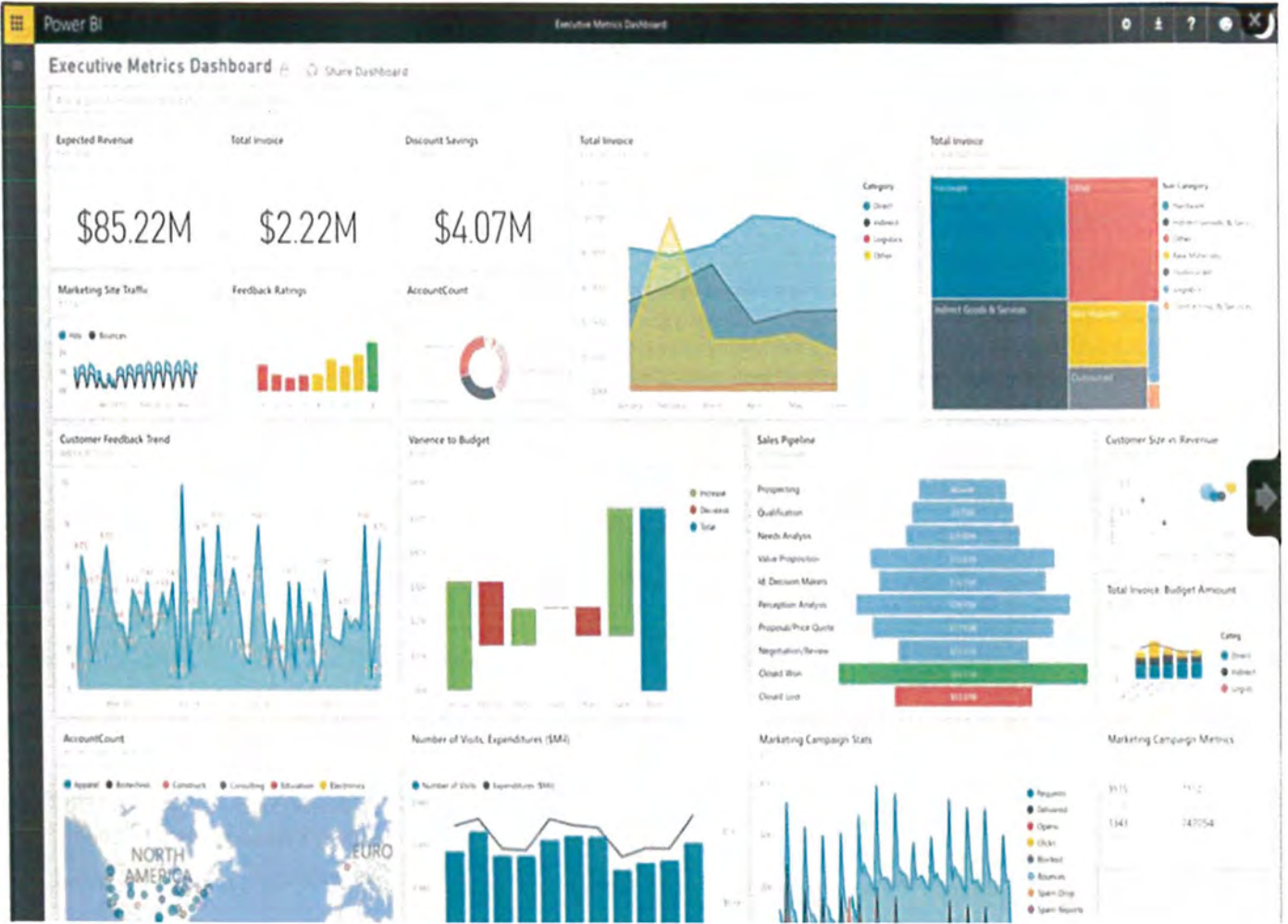


Fig 5.3 Visualized Report

## CHAPTER 6

### 6 System Testing

Testing of Data Warehouse/Business Intelligence (DW/BI) applications is a little different than testing traditional transactional applications as it requires data-centric testing approach

#### 6.1 Test Scenario & Test Cases

Following are generic test cases that need to be validated for any BI Testing Project

Test Scenarios	Test Cases
<b>ETL verification</b>	<ul style="list-style-type: none"><li>• Verify data is mapped correctly from source to target system</li><li>• Verify all tables and their fields are copied from source to target</li><li>• Verify that null fields are not populated</li><li>• Verify data is neither garbled nor truncated</li><li>• Verify data type and format in target system is as expected</li><li>• Verify there is no duplicity of data in the target system</li><li>• Verify transformations are applied correctly</li></ul>

- Verify that the precision of data in numeric fields is accurate
- Verify exception handling is robust

**Test Scenarios****Test Cases****Staging data**

- Reconciliation check- record count between the STG (staging) tables and target tables are same after applying filter rules
- Insert a record which is not loaded into target table for given key combination
- Copy records, sending same records that are already loaded into target tables-should not be loaded
- Update a record for a key when value columns changed on day\_02 loads
- Delete the records logically in the target tables
- Values loaded by process tables
- Values loaded by reference tables



---

**Test Scenarios****Test Cases**

---

**Data Loading in BI**

- Check if the target and source data base are connected well and there are no access issues.
- For a full load, check the truncate option and ensure its working fine.
- While loading the data, check for the performance of the session
- Check for non-fatal errors.
- Verify you can fail the calling parent task if the child task fails.
- Verify that the logs are updated
- Verify mapping and workflow parameters are configured accurately
- Verify the number of tables in source and target systems is the same

---

**Test Scenarios****Test Cases**

---

**BI Reports**

- Display date and time
- Decimal precision for key figures
- In a given page display the number of rows and columns
- Free characteristics in the report
- How are blank values/data displayed for both characteristics and key figures in the report
- Whether search for characteristics is based on key or key text as applicable
- Does search option on text is case sensitive- Upper, Lower or both

## CHAPTER 7

### 7 Conclusive Discussion and Future Work

#### 7.1 Conclusion

The purpose of the project is to provide end users, company manager a way of taking better business decision in order to improve their business.

Telecom BI analysis & Reporting was a small effort in a bid to accomplish the purpose

This project comprised upon two major parts, the first part is based upon designing a data warehouse, and the other part of the project includes the Business Intelligence reporting

#### 7.2 Future Work

Feature enhancement has its own importance, because with the passage of time system evolves and there is a need for more features and functionalities, so system should allow future enhancement. Purposed System is developed by keeping this in mind.

This project can be improved in the following ways:

- Instead of carry out simple analysis , python & R be used for the data analysis
- Instead of carry data visualization with Power BI which is user friendly and need no help from the professionals, visualization will be carry out with tableau.