BUILDING A DATA WAREHOUSE TO ANALYZE COMPUTER MAINTENANCE REQUESTS

Project report submitted in partial fulfillment of the requirement for the degree of Bachelor of Technology

in

Computer Science and Engineering/Information Technology

By

Shanu Kumar (141244)
Niharika (141377)
Akanksha Singh(143202)
Manoj Kumar(141247)
Aaron Quadros(141252)
Aastha Singh(141280)

Under the supervision of

Mr. Abhijith V C

to

Department of Computer Science & Engineering and Information Technology
Jaypee University of Information Technology, Waknaghat, Solan-173234, Himachal Pradesh

Candidate's Declaration

I hereby declare that the work presented in this report entitled " Building a Data Warehouse To Analyze Computer Maintenance Requests" in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering/Information Technology** submitted in the department of Computer Science & Engineering and Information Technology**,** Jaypee University of Information Technology, Waknaghat is an authentic record of my own work carried out over a period from April 2018 to May 2018 under the supervision of **Mr. Abhijith V C**(Systems Engineer, ETA).

The matter embodied in the report has not been submitted for the award of any other degree or diploma.

Shanu Kumar (141244) Manoj Kumar(141247)

Niharika (141377) Aaron Maurice Quadros(141252)

Akanksha Singh(143202) Aastha Singh(141280)

This is to certify that the above statement made by the candidate is true to the best of my knowledge.

(Supervisor Signature)
Mr. Abhijith V C (Systems Engineer)
Education Training and Assessment Department
Dated:

ACKNOWLEDGEMENT

We are grateful and indebted to Mr. Abhijith V C, Education Training and Assessment Department for his help and advice in completion of this project report. We also express our deep sense of gratitude and appreciation to our guide for his constant supervision, inspiration and encouragement right from the beginning of this project. We also want to thank our parents and friends for their immense support and confidence upon us. We deem it a pleasant duty to place on record our sincere and heartfelt gratitude to our project guide for his long sightedness, wisdom and co-operation which helped us in tackling crucial aspects of the project in a very logical and practical way.

Shanu Kumar (141244)

Niharika (141377)

Akanksha Singh(143202)

Manoj Kumar(141247)

Aaron Quadros(141252)

Aastha Singh(141280)

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

In large IT service environments, with the amount of systems that are placed for their respective employees many issues can crop up, these issues are logged as tickets that are handled in a ticket framework set up that enables individuals to report different issues. On the off chance that you experience difficulty with your system or any program on it, you can report your issue to the company's assistance work area and send an inconvenience ticket.

A ticket framework, is particularly intended to sort out issues as they are accounted for and monitors all fixes made, going about as an issue following framework also.

With the accessibility of sensibly huge ticket datasets, it gives data related to the kind of maintenance activities that is preferred in any timestamp. Along these lines, arranging and examining tickets turns into a basic assignment in dealing with the tasks of the administration.

In this project,we initially recognize the difficulties in real-world ticket analysis and build up a coordinated structure to effectively deal with those difficulties.

.

We identified various KPI's (Key Performance Indicators) that could be integrated with our dashboard so as to provide short and accurate analysis to help in the decision making process.

The proposed system is broadly assessed with a vast true dataset.

1. INTRODUCTION

1.1) Introduction

Business Intelligence is defined as a set of concepts and methodologies used to aid the decision making process by using of facts and figures. It enables knowledge workers to make accurate and fast decisions. It allows data to flow from operational systems to systems that support decision making. Data stored in warehouse is usually summarized or aggregated so as to facilitate decision making.

1.1.1) Data Quality

The data that is stored in data warehouse should be accurate, complete, consistent and timely. Accuracy is the degree to which the data correctly reflects the real world entity. Consistency is required so as to maintain the same version of truth across the entire enterprise. Completeness is the degree to which the required data is provided. Timeliness is when the right data reaches in right person in time.

1.1.2) Types of Enterprise Data

IT systems are divided into two processes, transactional and analytical. On-line Analytical Processing is supported by data warehousing and it is very different from on-line transaction processing applications which is supported by operational databases. OLTP applications automate day to day operations which are the source for all the data required for and contributed by various operations of the organization. This data is usually in megabytes to gigabytes in size. Maximizing the transaction throughput is a key performance metric. These applications are mainly associated with entry, storage and retrieval of data. OLTP data is very detailed to be used for decision making.

OLAP requires data to be stored in dimensional form opposed to relational form. OLAP tools are based on multi-dimensional data model. Multi-dimensional data model views data in the form of a data cube. A cube is a data structure that helps in analysis based on the dimensions required by the business problem.

1.1.3) Data Warehouse

A data warehouse is a "subject-oriented, integrated, time varying, non-volatile collection of data that is used primarily in organizational decision making." [3]

- Subject Oriented: The data warehouse provides subject oriented information rather than giving the overall information of the company's ongoing information. The focus of data warehouse is not the ongoing operation of an organization but rather the modelling and analysis of data for better decision making process.

- Integrated: Data warehouses are constructed using heterogeneous sources such as flat files, relational databases, etc. This allows the collection of data from different OLTP sources as well as data which is stored in different formats.

- Time Variant: Data warehouses contain historic data rather than day-to-day transactional data. OLTP databases are regularly backed up onto data warehouses. It is not like operational systems which store the most recent data.

- Non-Volatile: Previous records are not erased on addition of new data to the data warehouse. Data that has once been stored in the warehouse is not subjected to further modifications. This allows data to be non-volatile.

Data warehouses contain enterprise data that are sized between hundreds of gigabytes to hundreds of terabytes of data as they contain data consolidated into one storage from many operational databases over a large time period. The queries are mostly ad hoc and complex and access records in millions and perform many joins and aggregates as the data is stored in renormalized forms.

To encourage complex investigations and perception, the information in a distribution center is ordinarily displayed in a multidimensional format. Frequently, these measurements are hierarchal; time is day-month-quarter-year progressive system. Common OLAP activities incorporate rollup (expanding the level of conglomeration) and drill down (diminishing the level of conglomeration or expanding subtle element) along at least one measurement, 'slice_and_dice' (determination and projection) along with pivot (re-situating the multidimensional view)

The process includes extracting data from several data sources and external sources, then cleaning and transforming then integrating this data to load into the data warehouse. This data is periodically refreshed to reflect updates the updates that occur at the source turning the warehouse into archive storage.

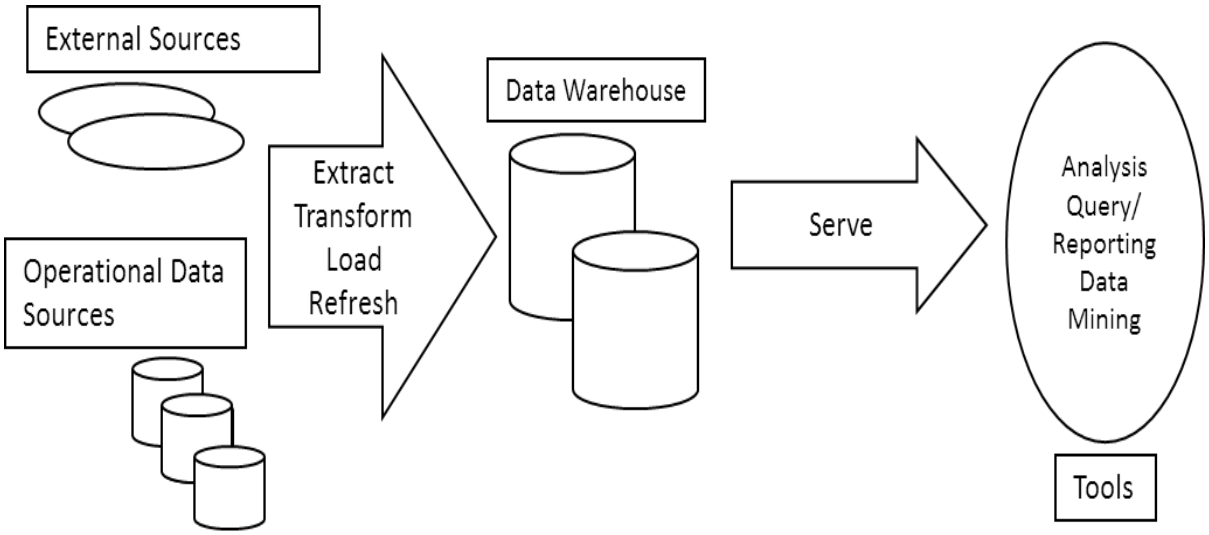


Figure 1: Data Warehousing Architecture

1.1.4) ETL

ETL processes allow the organizations to gather data that is stored in various locations in different format s to be unified and stored into one location. Thus, allowing a single version of the data stored extracted from various data sources. The data can be modified according to the needs of the analysist making it suitable for analysis.

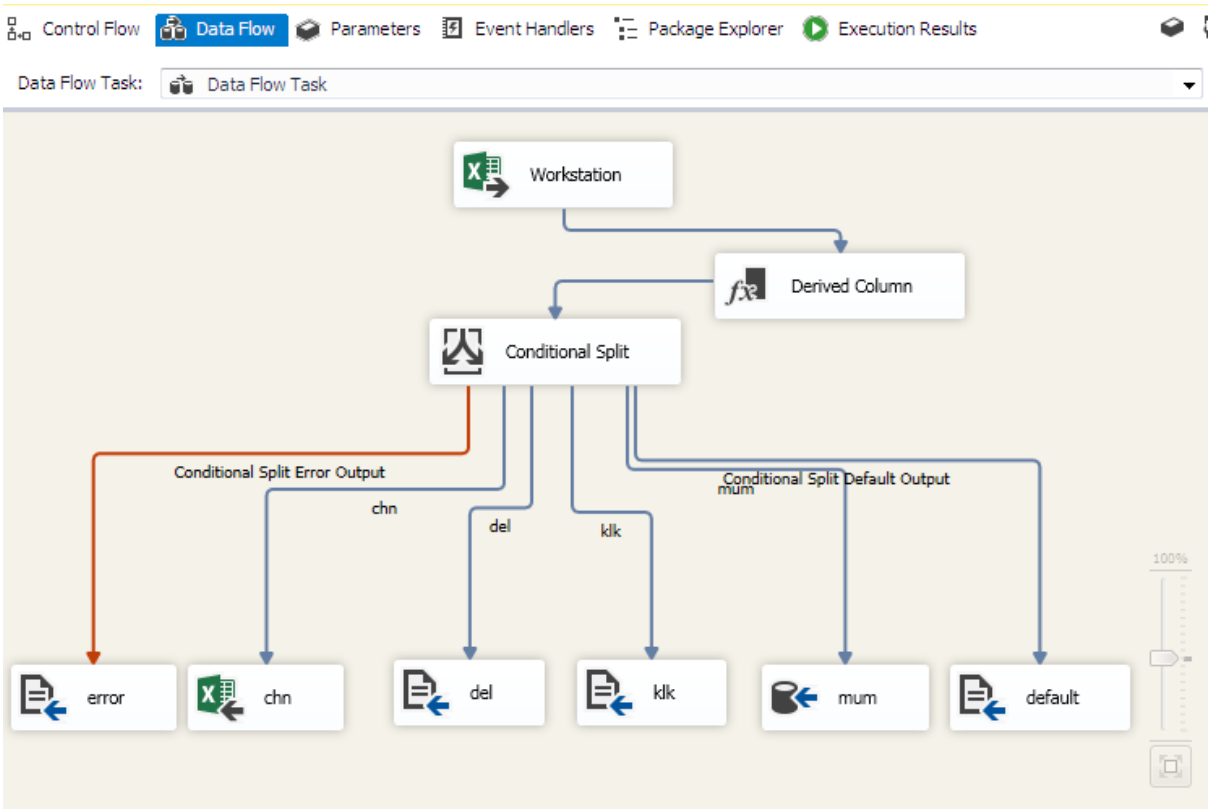


Figure 2: ETL Operation

Extraction:

The initial segment of an ETL procedure includes extracting the information from the source system(s). Much of the time, this is the most imperative part of ETL, since separating information effectively sets the phase for the achievement of ensuing procedures. Most data warehousing ventures consolidate information from various source frameworks. Each different framework may likewise utilize an alternate data model or structure. Common information source positions incorporate structured databases, XML, JSON and excel documents, however may likewise incorporate non-structured database structures. When

all is said in done, the extraction stage plans to change over the information into a solitary organization suitable for handling transformation.

Transformation:

The subsequent stage in the ETL procedure is transformation. After extraction of data, it must be physically transported to the final location, and changed over into the suitable arrangement. This data change may incorporate several activities, for example, cleaning, joining, and approving information or creating ascertained information in view of existing qualities.

Regardless of whether the change happens in the information distribution center or previously, there are both normal and propelled change composes that get ready information for examination.

Loading:

The last step in the ETL procedure includes stacking the changed information into the goal target. This objective might be a database or an information stockroom. There are two essential techniques for stacking information into a database: full load and incremental load. The full load technique includes a whole information dump that happens in the first run through the source which is stacked into the distribution center. The incremental load, then again, happens at general interims. These interims can be spilling additions (better for littler information volumes) or clump increases (better for bigger information volumes).

While one can plan and keep up their own ETL process, it is normally viewed as a standout amongst the most difficult and asset concentrated parts of the data warehouse venture, requiring a great deal of time and work. Numerous associations use ETL tools that help with the procedure, giving capacities and preferences inaccessible if you somehow happened to finish it all alone. These tools don’t just help with the extraction, change and stacking process, yet they can likewise help in planning the information distribution center and dealing with the information stream.

1.1.5) Schema:

STAR Schema:

STAR Schema is one of the simplest and most widely used schema which is used to build large and complex data warehouse. In STAR Schema, we have a fact table which is present at the center of the schema whereas the dimension tables are placed around it. The fact table consists of foreign keys referencing to the keys present in dimension table. Moreover, STAR Schema is the simplest version of SNOWFLAKE Schema which is used to handle small queries.

In STAR Schema both the fact table and the dimension tables are present in de-normalized form. This schema is basically used by all OLAP systems having a single dimension table for each dimension.

The following example shows the design of a STAR Schema:

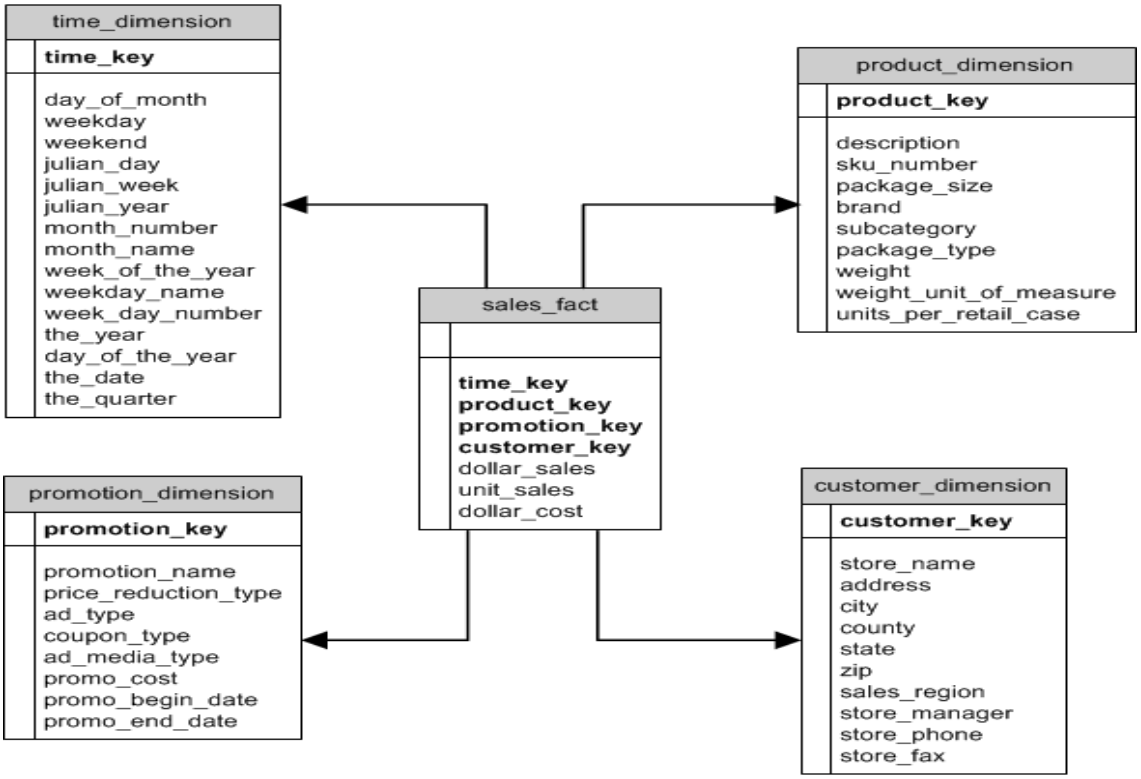


Figure 3: Star Schema

The above example consists of a fact table “sales_fact” having time_key, product_key, promotion_key and customer_key as foreign keys referencing to the primary keys present in respective dimension tables i.e. “time_dimension”,” product_dimension”,” promotion_dimension” and “customer_dimension”.

SNOWFLAKE Schema:

SNOWFLAKE Schema consists of a centralized fact table which is connected to a number of dimension tables which are further connected to sub-dimensional tables. In SNOWFLAKE Schema both the fact table and the dimension tables are present in normalized form which are used to perform complex ETL operations. As compared to STAR Schema it is difficult to perform operations in SNOWFLAKE Schema as the number of dimensions are comparatively large having a large number of joins associated with them.

The following example consists of a fact table “sales” which is associated with dimension tables but unlike STAR Schema the dimension tables are further split into sub-dimension tables like “item” and “supplier” dimension tables.

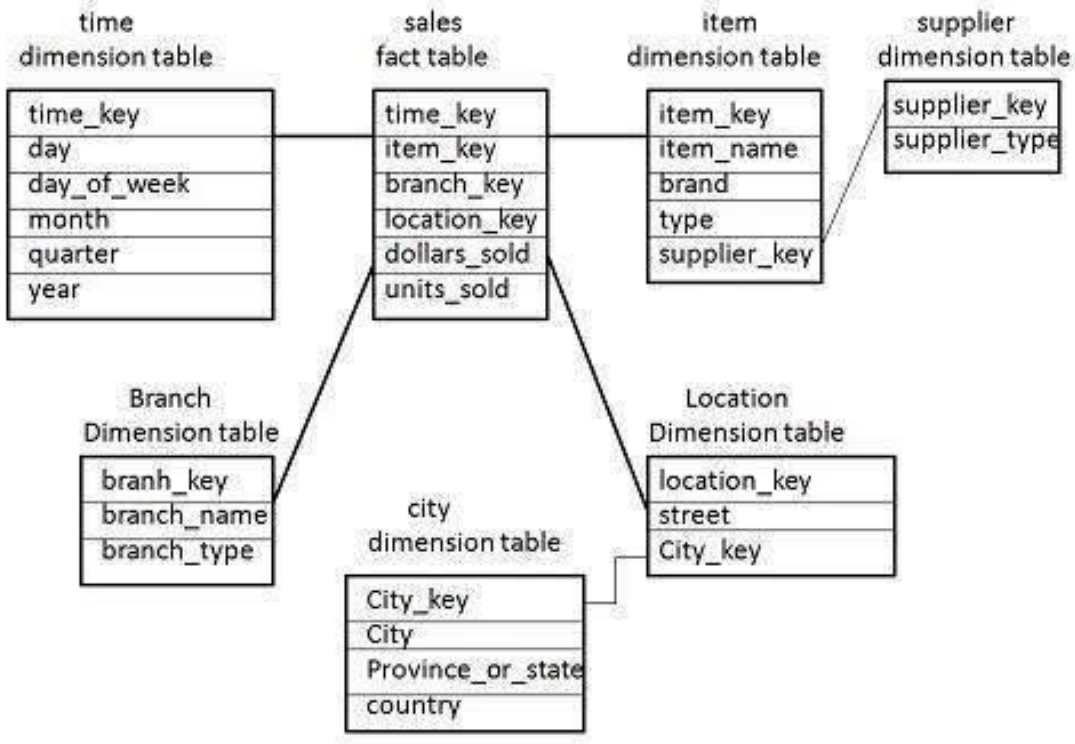


Figure 4: SnowFlake Schema

Moreover, in SNOWFLAKE Schema the query execution time is comparatively high as compared with STAR Schema as there are more number of foreign keys and high number of joins in SNOWFLAKE Schema.

1.2) PROBLEM STATEMENT

There is a large amount of tickets generated in any organization. The cost of maintenance and addressing of these tickets are often high and varies according to the complexity of the tickets.

The main purpose of an organization's maintenance activities is to support the production process providing maximal levels of accessibility, reliability and operability at acceptable cost. If this objective is followed the results will be clearly visible in the form of increased production capacity and profit of the organization. As high profit is one of the major goals of any organization, maintenance should be considered in a very serious way.

Analysis of these tickets data can provide insights which suggests actions to be taken which can reduce the cost and increase the overall performance of the organization.

Before analyzing large datasets, it has to be stored somewhere which further creates the necessity of creating a data warehouse.

Reports created from the data stored in the data warehouse must be free from anomalies and should provide single version of truth. These reports must provide meaningful insights and information which can be used by decision makers to make the necessary decisions for the organization’s benefit.

1.3 OBJECTIVE

To build a data warehouse to analyze computer maintenance requests.

To create reports using the data stored in the data warehouse to gain meaningful insights.

To create meaningful reports which help in understanding the forces that affects the business and allow decision makers to make better decisions.

To provide effective analysis based on the data collected which helps to monitor the performance and take appropriate actions.

1.4 METHODOLOGY

This chapter contains intricate explanation of different methodologies that are being used to develop the project "**Building a data warehouse to analyze computer maintenance requests**". The model for software development which is being followed is agile-scrum method. The benefit of using agile method for software buildout is unlike waterfall method requirements can be added anytime while the software is under development and modifications are incorporated anytime if needed.

The reason behind use of agile method for software development is many a times clients are not aware of the proper requirements and hence they add certain features while the software is in the development phase in such cases waterfall method are not useful as changes are not appreciated while following waterfall model. And a situation most often faced by the developers is evolution of requirement i.e. clients demands some specific task to be performed by the software which the also is a problematic situation or a case like software developed using some specific specification and after sometime of deployment the number of users increases or a change of platform is required.

Like all software development model coincides with the phases of SDLC (Software Development Life Cycle). The various phases of SDLC are: -

1) Planning
2) Analysis
3) Design
4) Development
5) Testing
6) Deployment

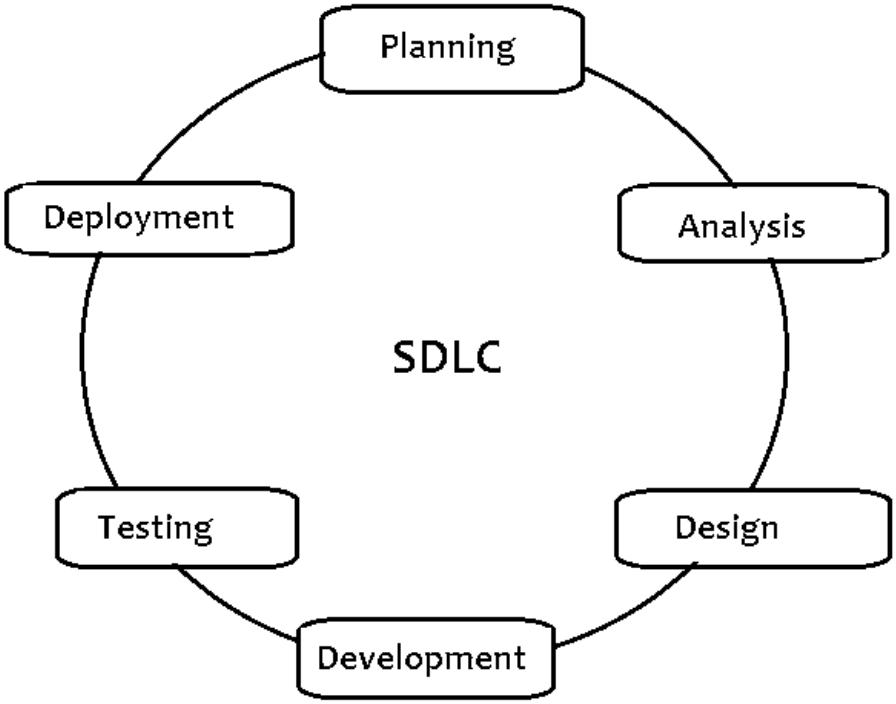


Figure 5: Software Development Life Cycle

This development of project used three major steps to implement project starting from planning (software and hardware requirements, designing schema, building databases etc.), implementing (Cognos, Power BI) and testing.

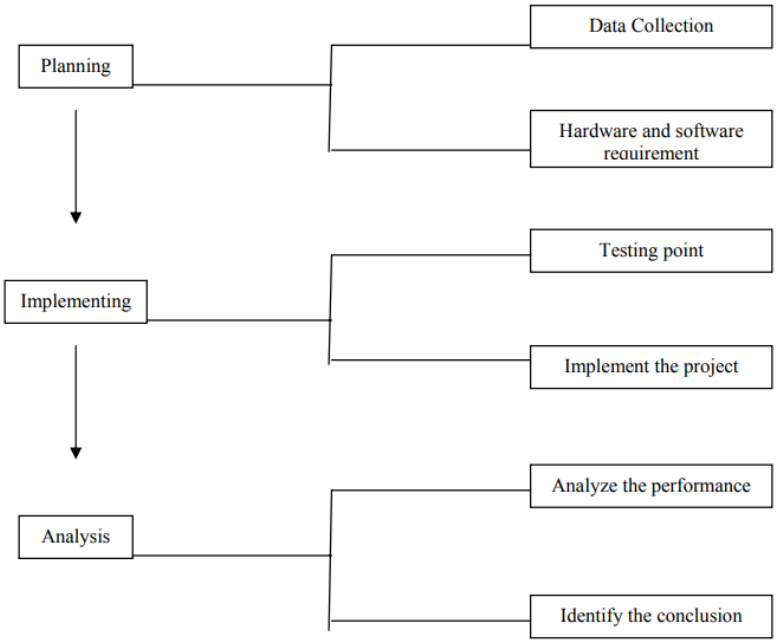


Figure 6: Steps of Methodology

Planning:

Planning phase is basically divided into two parts in the first part we gather all the requirements whereas in the second we identified all the hardware and software required for the project.

The following list provides the minimum requirements to run **SQL Server Data Tool**:

SSIS 2005, 2008, 2012, 2014, or 2016
Business Intelligence Development Studio 2005-06 or SSIS 2012, 2014, or 2016
Windows XP or later
Windows Server 2000 or later

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

- Windows 7 / Windows Server 2008 R2, or later
- .NET 4.5
- Internet Explorer version 9 or later
- **Memory (RAM):** 1 GB available, 1.5 GB or more recommended.
- **Display:** At least 1440x900 or 1600x900
- **CPU:** 1 gigahertz (GHz) or faster x86- or x64-bit processor recommended.

Data collection:

In order to check the proper working of the project datasets were needed which was created manually. Again for this requirement analysis was done to verify the database's tables needed for the project.

Fact table contains 5000 records.

DimUsers contains 2200 records.

DimGeneratedDate and DimResolvedDate contain 1300 records.

Other databases used are requester, service-area, workstations, locations etc.

Implementation: -

The project was implemented with the help of SQL Server Data Tools and Power BI. SSIS was used for creating datasets and cleansing of data. Power BI was used to develop the reporting system.

1.5 ORGANIZATION

The project focuses on building a data warehouse to analyze computer maintenance ticket requests. The report consists of table of contents including a list of figures and tables used in the report along with various screenshots. The report contains an introduction to the project followed by details about the various ETL operations involved in gathering the data and therefore cleansing it using SSIS tool which is further used in building up of reports. The report also focuses on developing the system. The organization of the project report is as follows:

Chapter 1 gives a detailed introduction about the project. It consists of introduction about the system, problem statement, objectives, various ETL (Extraction, Transformation, Loading) operations of the project and the methodology involved in the system. It also contains details about data warehouse and various anomalies associated with the datasets.

Chapter 2 consists of System development which explains about the various experimental methods and functional requirements involved in the implementation of project. It explains about various tools used while building up of project such as-SSMS (SQL Server Management Studio), SSIS (SQL Server Integration Services), Power BI Desktop and the schema used in building data warehouse. It also tells about the various key modules involved in the implementation.

Chapter 3 is Performance Analysis. It consists of various reports which are build to provide meaningful insights and information that can be used by decision makers to make the necessary decisions for the organization's benefit.

Chapter 4 is Conclusion of the project which gives the outcome of the project work carried out and also brings out the limitations of the project work and gives an enhancement about future scope.

2. LITERATURE SURVEY

2.1) RESEARCH PAPER I

Adam Wilson in “Bring your data to life with Power BI” [1] says that the idea of business intelligence has been around in different structures for over a century and a half. At its center, BI is tied in with understanding the realities - and the connection between actualities - in a way that guides basic leadership and activity.

From an innovation point of view, BI is an arrangement of methods and instruments for changing crude information into important business bits of knowledge. IT experts have reliably assumed a key part in opening an incentive from information by making and keeping up information stockrooms, assembling and uncovering complex information models, or report creation.

At the same time, end clients have been depending to a great extent on the help from IT to meet their BI needs, including the advancement of reports.

Microsoft Excel made ready for self-benefit BI, as it was the initial phase in giving business investigators the capacity to pick up bits of knowledge from information. Microsoft made self-benefit BI a reality with Power Pivot, expanding capacities to Excel that were beforehand discovered just in databases.

These developments empowered experts to convey bits of knowledge from more noteworthy volumes of information and a wide exhibit of information sources.

Moreover, BI keeps on developing and Microsoft is driving the path by bringing another age of BI to associations, with arrangements that will expand and expand on, as opposed to supplant, existing examination stages what's more, instruments. With Power BI, a business examination benefit for picturing and breaking down the greater part of your information in one place.

Power BI is a SaaS offering that empowers anybody and everybody to effortlessly associate with any of their information, make live dashboards and reports, and investigate information through intelligent perceptions whenever.

With Power BI, one can make the majority of information perceptible in a solitary area, paying little heed to where the information lives, empowering a combined perspective of business tasks.

Power BI incorporates two sidekick applications. The first is Power BI Desktop, a visual information investigation furthermore, announcing apparatus. The second is an arrangement of local, intelligent versatile applications for Windows, iOS, and Android gadgets, giving secure access to live Power BI dashboards and reports from any gadget.

What's more, Power BI can be reached out with an arrangement of REST APIs which empower engineers to incorporate customer and web arrangements with Power BI or to fabricate custom perceptions.

Power BI Desktop also gives progressed investigation capacities that allow user to structure complex information, make connections, characterize an assortment of estimations, find connections, feature exemptions and figure business results.

At the point when information is transported in to Power BI Desktop, an information display is naturally made. Power BI Desktop identifies connections consequently, orders information and applies default rundown.

In particular, Power BI Desktop's AutoDetect recognizes connections over the greater part of datasets - counting amongst cloud and on-premises information sources - to quicken examination. We you can likewise refine the model as required - for instance, by physically making connections or altering the kind of relationship. Power BI Desktop underpins coordinated, one-to-many, many-to-one, and many-to-numerous connections. We can also apply single-heading or bidirectional channels to cross-channel information for extra views.

2.2) RESEARCH PAPER II

Mr. Dishek Mankad and Mr. Preyash Dholakia in “The Study on Data Warehouse Design and Usage” [2] say that having a data warehouse firstly may give an upper hand by introducing important data from which to gauge execution and make basic changes in accordance which helps in prevailing upon contenders. Secondly, a data warehouse can upgrade business profitability since it can rapidly and productive accumulate data that precisely depicts the association. Thirdly, it encourages client relationship administration since it gives a steady perspective of clients and things over all lines of business, all divisions, and all business sectors. At long last, it may achieve cost lessening by following patterns, examples and special cases over long periods.

In the event that you needed to do plan compelling data warehouse you must know the business needs and build a business investigation system. The development of a substantial and complex data framework can be seen as the development of a huge and complex working, for which the proprietor planner and manufacturer have extraordinary views.

This view is consolidated to shape an unpredictable system that speaks to the best down, business-driven, or proprietor's point of view, and in addition the base up, manufacturer driven, or practitioner's perspective of the data framework. Four, extraordinary sees with respect to an information distribution center outline must be viewed as: the top-down view, the information source view, the information distribution center view, of the data framework.

They also add that building and utilizing a data warehouse is a perplexing undertaking since it requires business aptitude innovation abilities, and program administration aptitudes. With respect to aptitudes, constructing a data warehouse includes seeing how frameworks store and deal with their information, how to construct extractors that exchange information from the operational framework to the information product house, and how to construct warehouse revive process that keeps the data warehouse sensibly fully informed regarding the operational framework's information.

They also throw light on the various approaches used and steps involved in them. A data warehouse can be made utilizing a top-down approach, a bottom up approach or a blend of both.

The top–down approach begins with general outline and arranging. It is valuable in situations where the innovation is surely understood, and where the business issues that must be illuminated are clear and well comprehended. The bottom-up approach begins with examinations and models. This is valuable in the early phase of business demonstrating and innovation advancement. What's more, it additionally permitted an association to advance at impressive less costs and assesses the mechanical points of interest before making noteworthy duties.

The event in which we are thinking in from the product building perspective, the plan and development of an information examination, stockroom plan, information mix and testing and lastly arrangement of the data warehouse, substantial programming frameworks can be created by utilizing one of the two models. The Waterfall technique and The spiral method.

3. SYSTEM DEVELOPMENT

3.1) SOFTWARE REQUIREMENTS:

- **Microsoft SQL Server Data Tools 2015**
- **Microsoft Power BI**

3.2) HARDWARE REQUIREMENTS:

- **System Requirements for Microsoft SQL Server Data Tools 2015:**
  - ✓ 64-Bit Windows XP or newer.
  - ✓ Microsoft SQL Server 2008 or newer.
  - ✓ 6 GB hard-disk space.
  - ✓ 4 GB RAM.
- **System Requirements for Microsoft Power BI:**
  - ✓ Windows 7 / Windows Server 2008 R2, or later.
  - ✓ .NET 4.5.
  - ✓ Internet Explorer 9 or later.
  - ✓ Memory (RAM): At least 1 GB available, 1.5 GB or more recommended.

3.3) SYSTEM DESIGN

Schema Used: Snowflake Schema

It is an extended version of star schema. In star approach, each point represents a dimension where as in snowflake every point of star is divided into two points each.

The reason for preferring the snowflake composition is the change in inquiry execution because of minimized storage necessities and joining littler query tables. The fundamental disservice of the snowflake blueprint is the extra upkeep endeavors required because of the expansion number of query tables.

Agile approach is followed to develop this project. Agile approach is likewise a sort of Incremental model. Software is produced in incremental, quick cycles. Due to this, it results in small incremental release where every release depends on past usefulness. Every release altogether tries to guarantee best programming quality. Extreme Programming (XP) is as of now a stand out amongst the most understood light-footed improvement life cycle demonstrate.

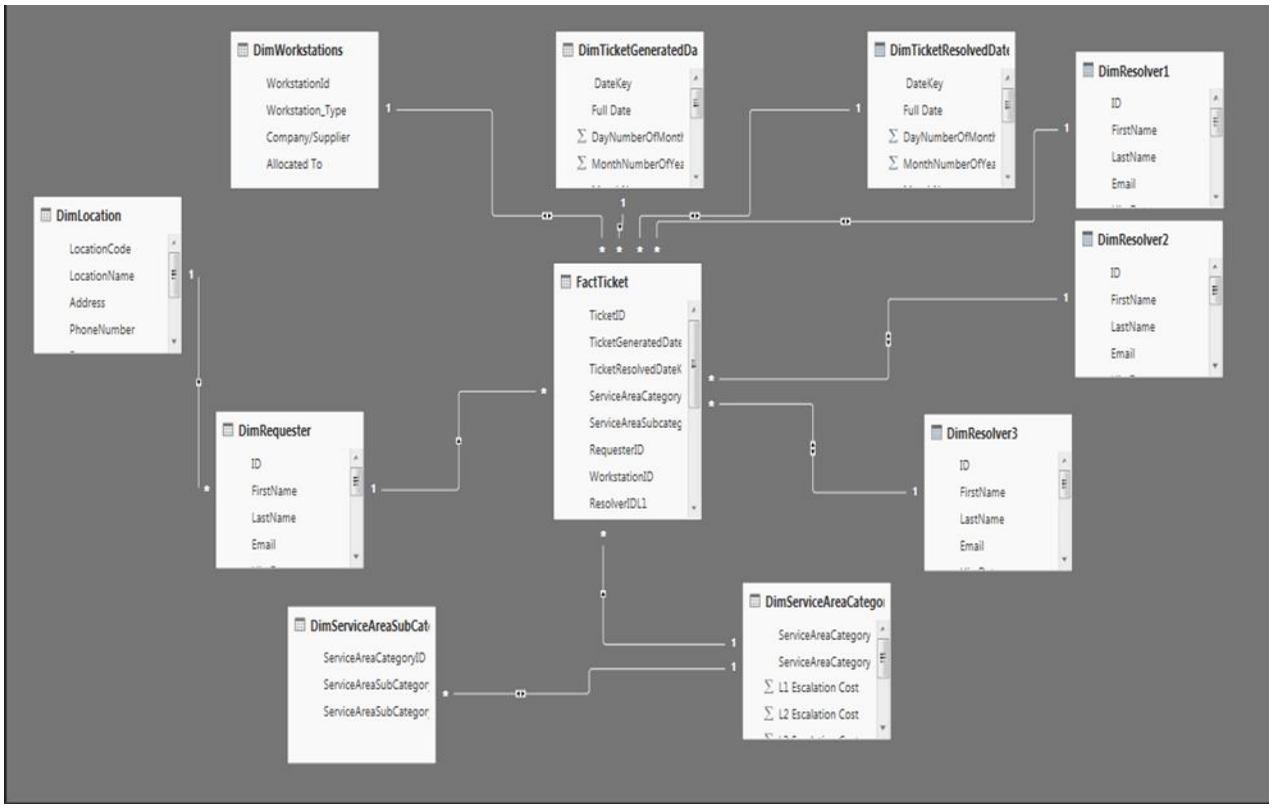


Figure 7: Schema

Tools Used

SSIS

SSIS is a product developed by Microsoft. It is a tool which is used in ETL (Extraction, Transformation and Loading) operations. It provides features like cleansing data and data movement.

The information can be gathered from various areas and different types of file formats like flat files, excel files and databases and can be put into a single destination like Data Warehouse.

SSIS Architecture

SSIS comprises of mainly two components:

SSIS Runtime Engine:

This manages the control flow task of the package. This helps in running the package and consists of precedence constraints, breakpoints and connections.

This runtime engine follows a series of sequence to implement the tasks inside a package. During execution, if the engine comes across a data flow task, it creates a data flow pipeline for that data flow task.

Some of the components of control flow tasks (apart from data flow task) are mentioned below:

- Execute SQL task: This task is used to extract, delete or change data directly from a database table.
- Send mail task: This task is used to send e-mail to a group of people at once.

SSIS Data Flow Engine/Pipeline:

SSIS Data flow pipeline (also known as transformation pipeline) handles the flow of data from various data sources along with the transformation applied on them.

Working:

The data flow engine extracts data from the source and stores it in some buffer and then applies the required transformation in that buffer. In this way, the processing of data is much faster as compared with physically copying the data at each step of integration.

The data flow task consists of various components like:

Transforms

SSIS consists of various transformations like:

- Lookup: Looks whether the data is present in the other source or not.
- Merge: Used to merge two file.
- Merge Join: Used to merge two file with some common data.
- Aggregate: Used to find the aggregate like max, min or count.
- Conditional Split: Splits the data of a source based on the condition specified.
- Multicast: Used to transfer the source file into multiple types of destinations at once.
- Data Conversion: Used to convert one data type into other.
- Derived Column: Used to add new column or change the contents of existing column by applying various operations.
- Sort: Used to sort data according to specified column.
- Union all: Used to merge two or more files

Sources

Sources are the files from which the data is being extracted to perform operations. This consists of different types of file sources like:

- Excel Source
- Flat File Source
- OLE DB Source
- XML Source

Destinations

Destinations are the file in which the data is being stored after performing all the required ETL operations. This consists of destinations like:

- Excel Destination
- Flat File Destination
- OLE DB Destination
- XML Destination

Report Studio

Power BI is a service provided by Microsoft which helps in business analytics. Its highlights incorporate intelligent representation which helps clients in making reports and dashboards with no dependency.

It also provides cloud based business analytics service. Any type of report can be created using this tool like:

- Invoice Report
- Statement Report
- Inventory Report

It is basically used when the reports:

- are to be delivered to mass audience
- require maintenance at the time of data changes

It supports various types of report features such as:

- Cross-tab
- List Report: It is the most basic type of report organized in multi-row, multi-column format.
- Quick Measure: These are the values which are calculated at the time of report creation such as total cost.
- Sorting: This is used to sort elements according to specified order.
- Matrix Report: It is used to display grouped data. It does not include detail data rather it includes only summaries.
- Value Prompt: It filters the data according to the value selected.
- Apply Filter on Report: It is used to extract the data only for some mentioned condition.
- Conditional Style: It is a formatting applied on text or background if and only if the condition specified is satisfied.
- Chart Reports: Power BI Desktop support various types of chart reports such as pie chart, column chart, line chart, bar chart etc.
- Drill through: It is used to navigate from one report to another.

Development Model: Here experimental method is used to develop model. Experimental approach is to construct programming items and administrations by ceaselessly sending new version to clients. Instead of depending on pre-characterized prerequisites or assessment based suspicions, functionalities, and highlights is approved in their genuine commercial center by leading a consistent arrangement of tests.

While building a data warehouse, following steps were followed:

- Raw data is collected from various data sources and stored at a single place.
- Data is cleansed using SSIS and all the anomalies (if present) are removed from the data such as data replication and the data is converted to a single format.
- The raw data is then converted to tables and tables are converted to .dbo format using SSMS.
- The data from these tables is then used to make reports using Power BI.
- Once the reports are made, they are published using Power BI.

4. PERFORMANCE ANALYSIS

User Stories:

4.1) As a CEO, I want to know the overall resolution rate so that I can analyze my company's efficiency.

Figure 8: Overall Resolution Rate

4.2) As a Manager, I want to know the cost per ticket so that I can view the expenses in computer maintenance.

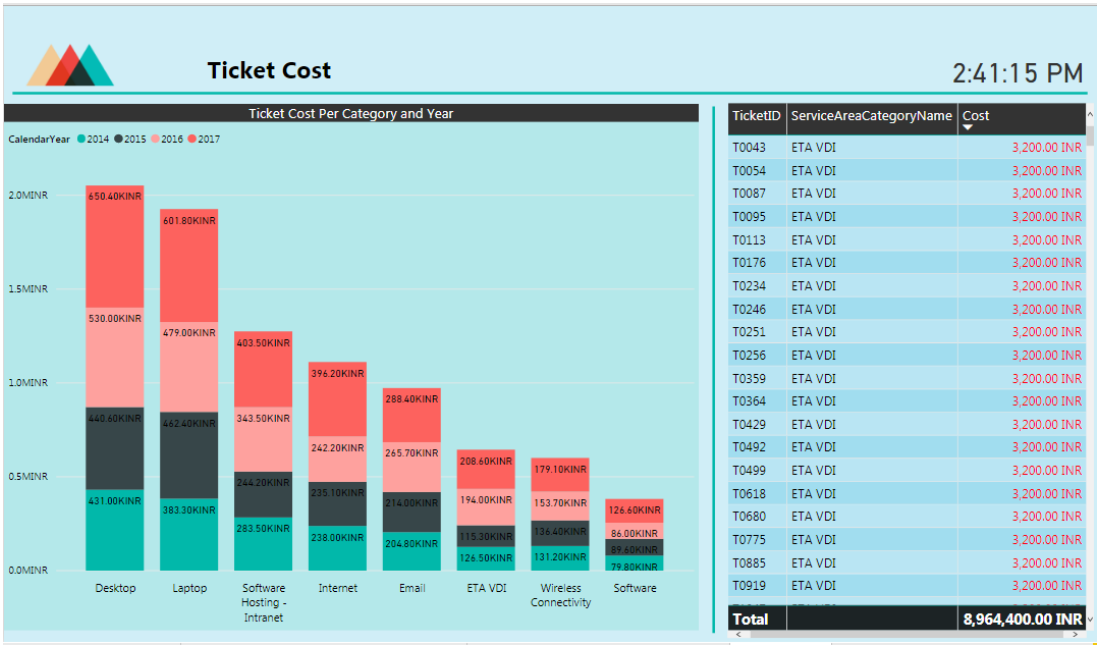


Figure 9: Cost Per Ticket

4.3) As a Manager, I want to know the number of tickets raised per department so that I can know in which department most issues are raised.

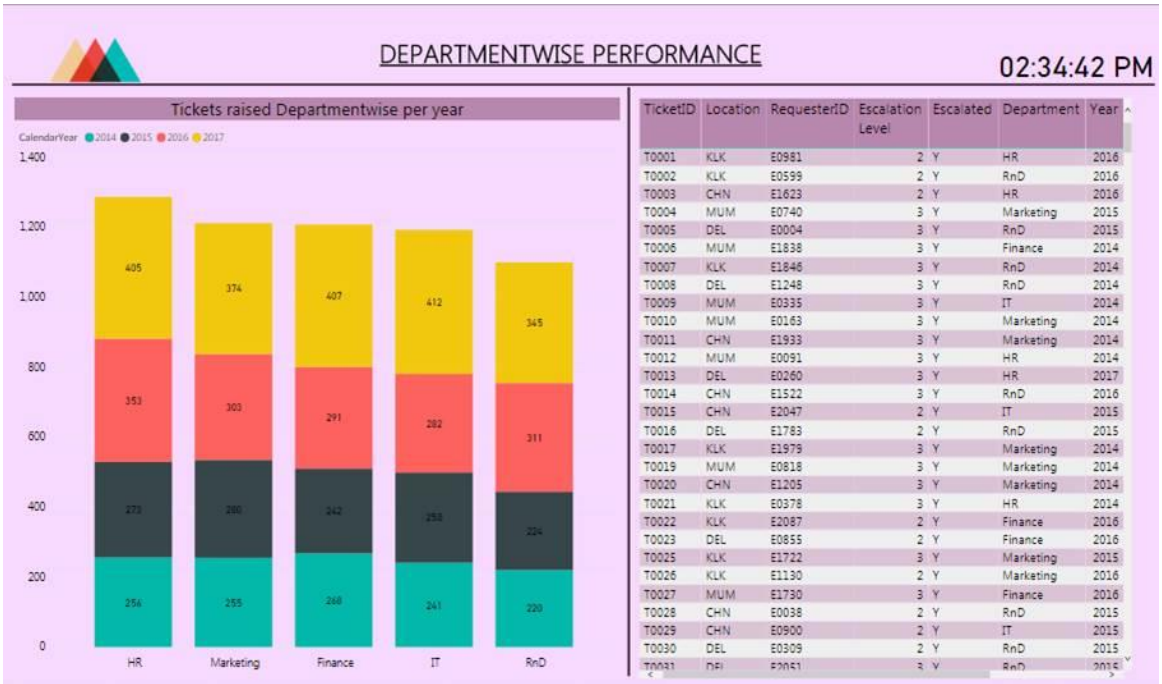


Figure 10: Department Wise Performance

4.4) As a Manager, I want to know the percent of failure for each supplier so that I can know the reliability of each supplier.

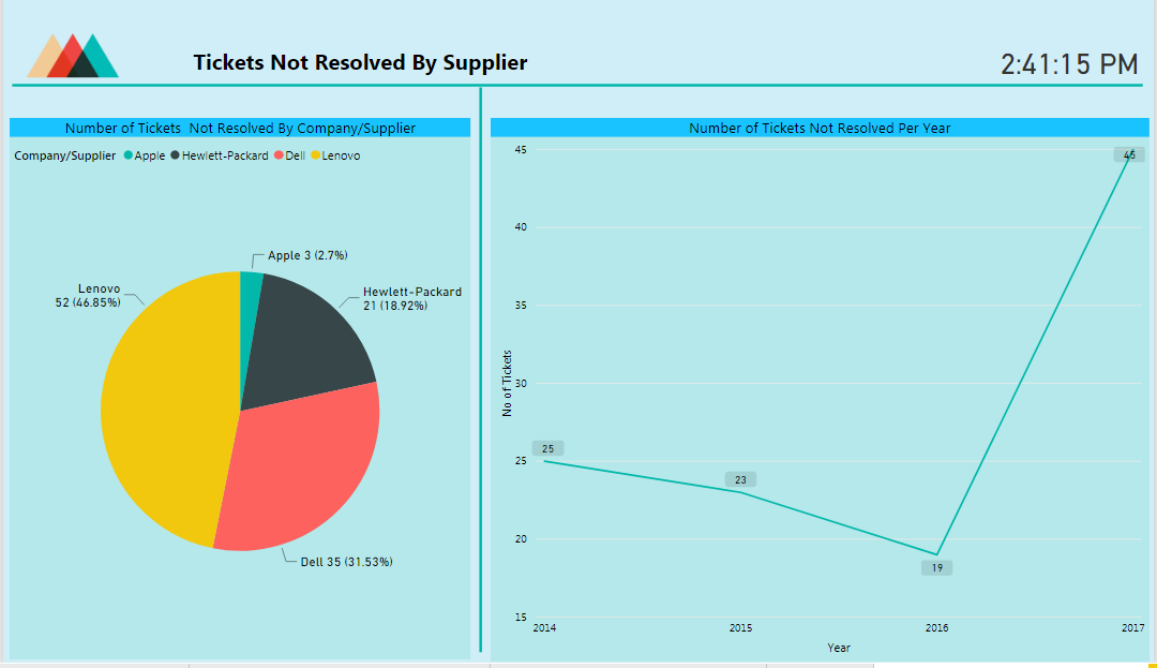


Figure 11: Percent Failure for Each Supplier

4.5) As a Manager, I want to know the efficiency rate of each employee so that I can judge their performance.

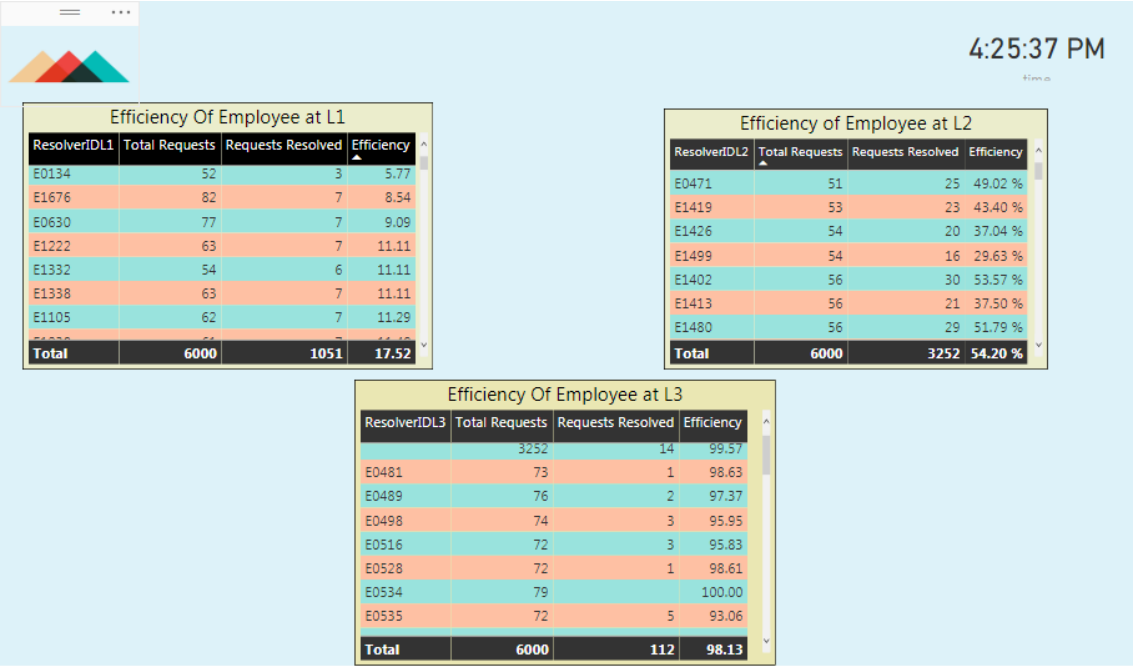


Figure 12: Efficiency Rate of Each Employee

4.6) As a CEO, I want to know the number of tickets generated and resolved per year/quarter/month so that I can analyze the company's performance.

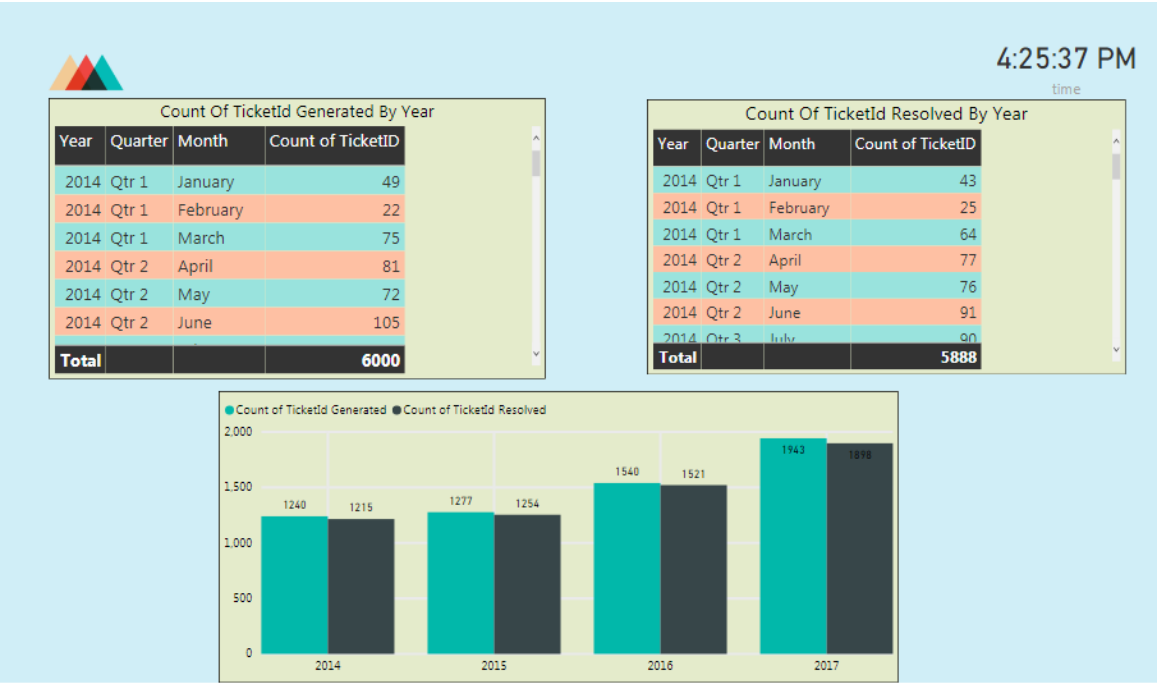


Figure 13: Ticket Generated and Resolved

4.7) As a CEO, I want to know the number of tickets raised per DC so that I can analyze the company's performance.

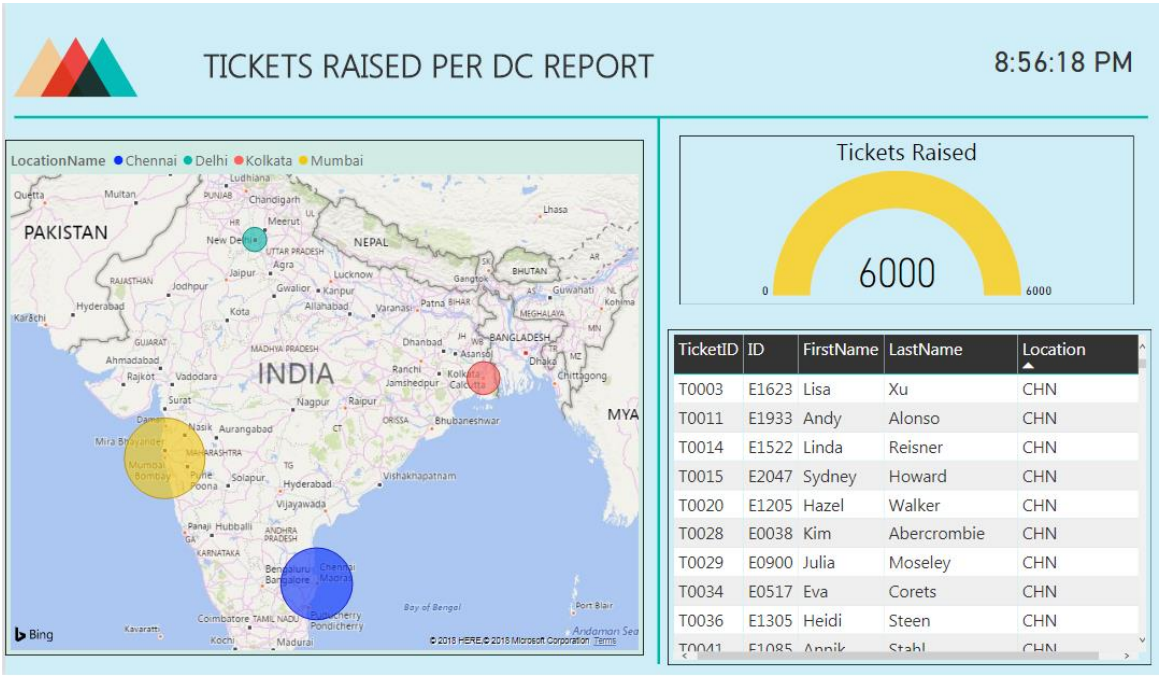


Figure 14: Number of Tickets Raised per DC

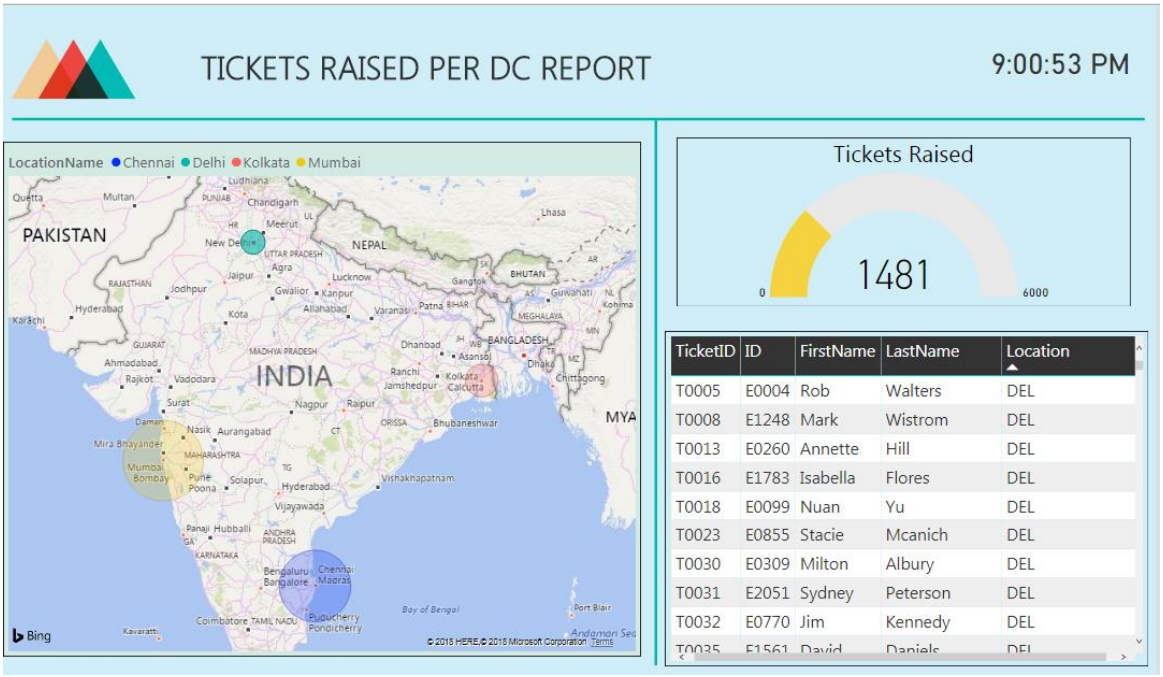


Figure 15: Details of Tickets Raised Per DC

4.8) As a Manager, I want to know the average resolution rate per request area so that I can know the time required to solve issues.

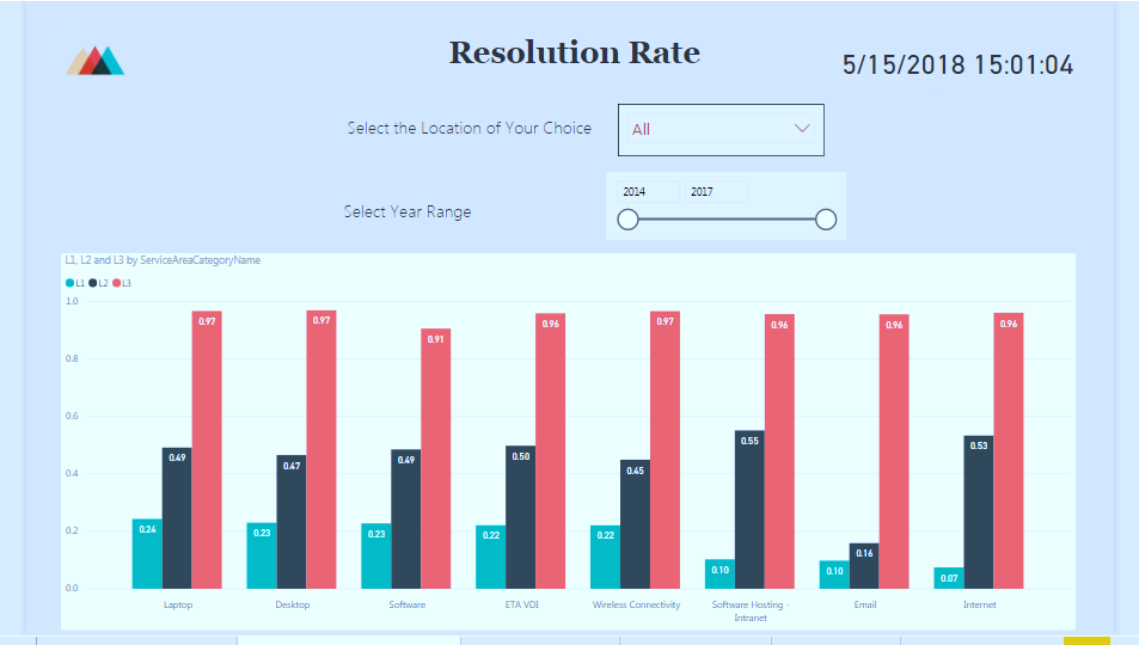


Figure 16: Resolution Rate for Each Request Area

4.9) As a Manager, I want to know how many requests are pending so that I can review the performance.

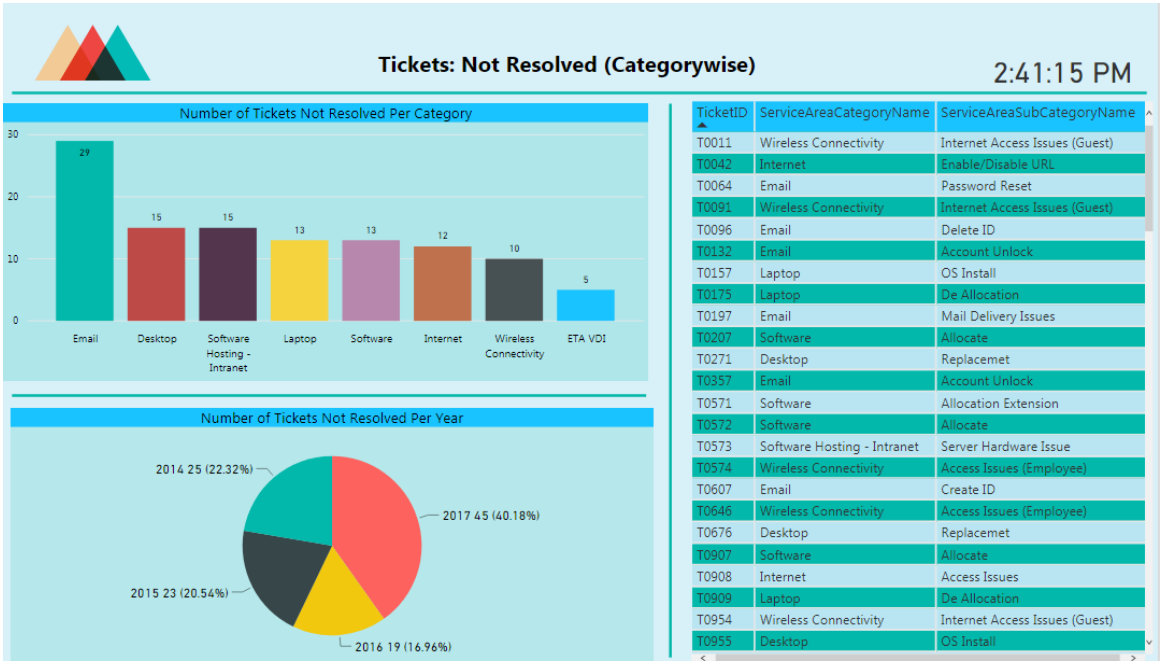


Figure 17: Details of Pending Requests

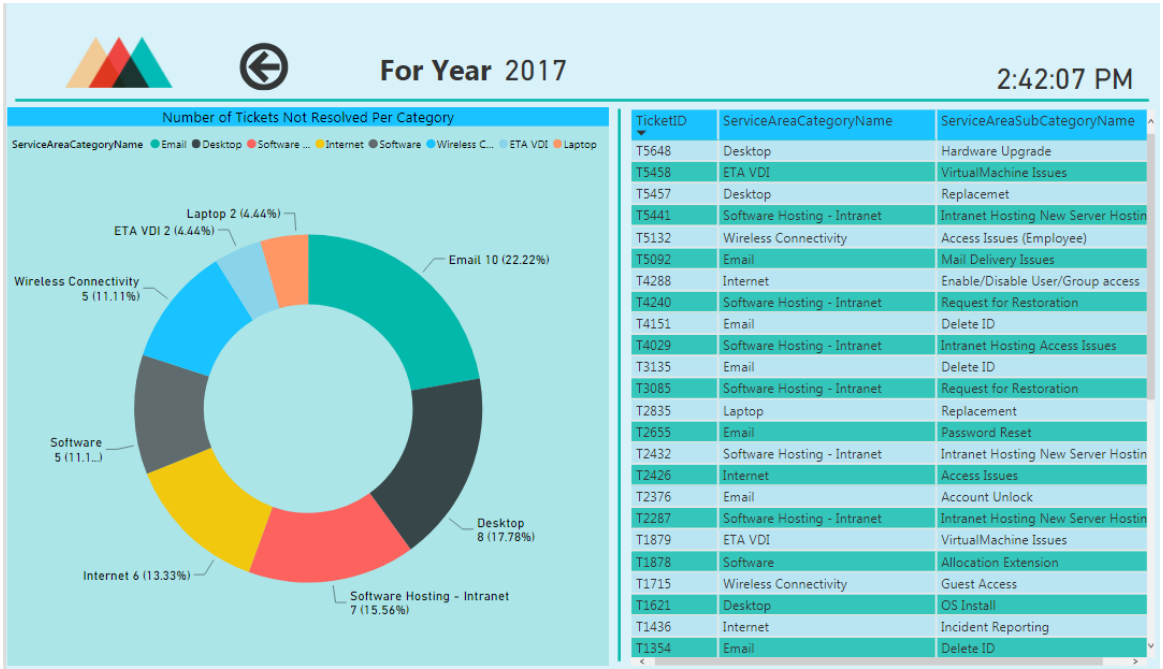


Figure 18: Details of Pending Request for a Given Year

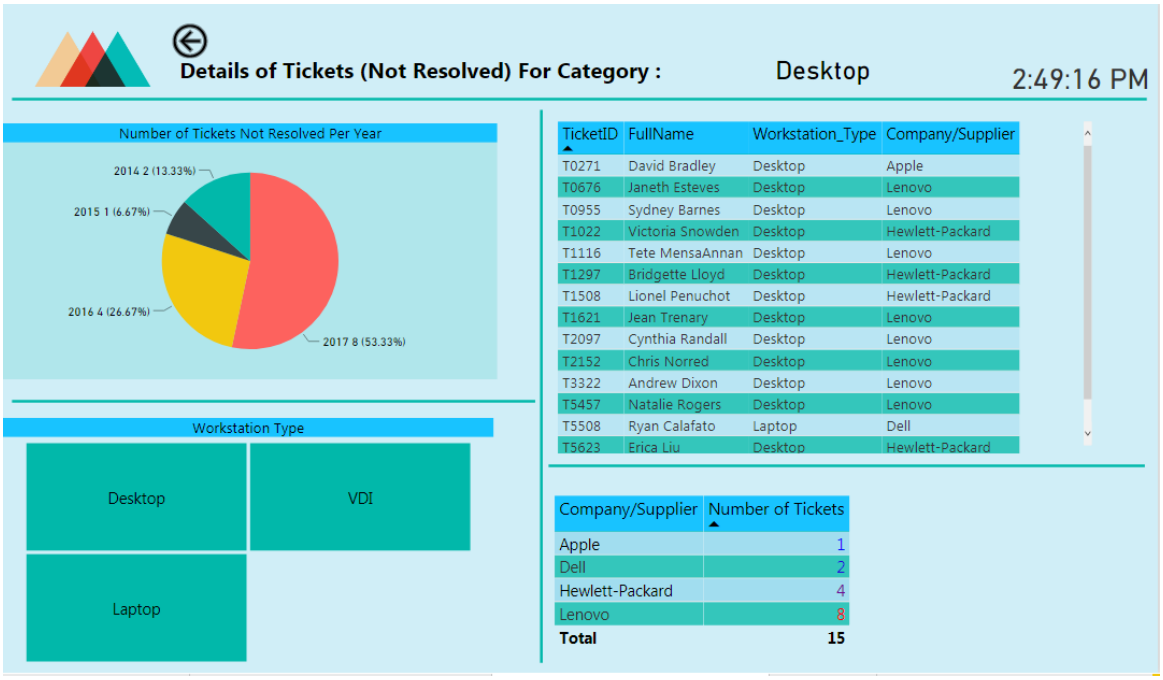


Figure 19: Details of Tickets raised for a Given Category

4.10) As a Manager, I want to know the escalation rate so that I can know the areas where employees need more training.

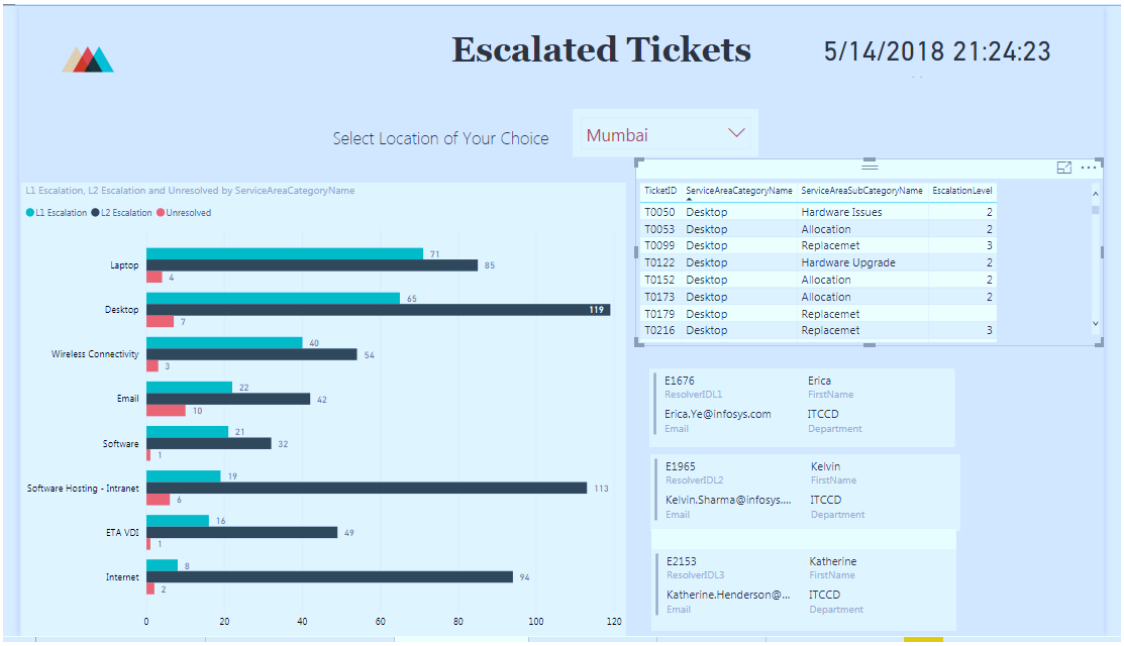
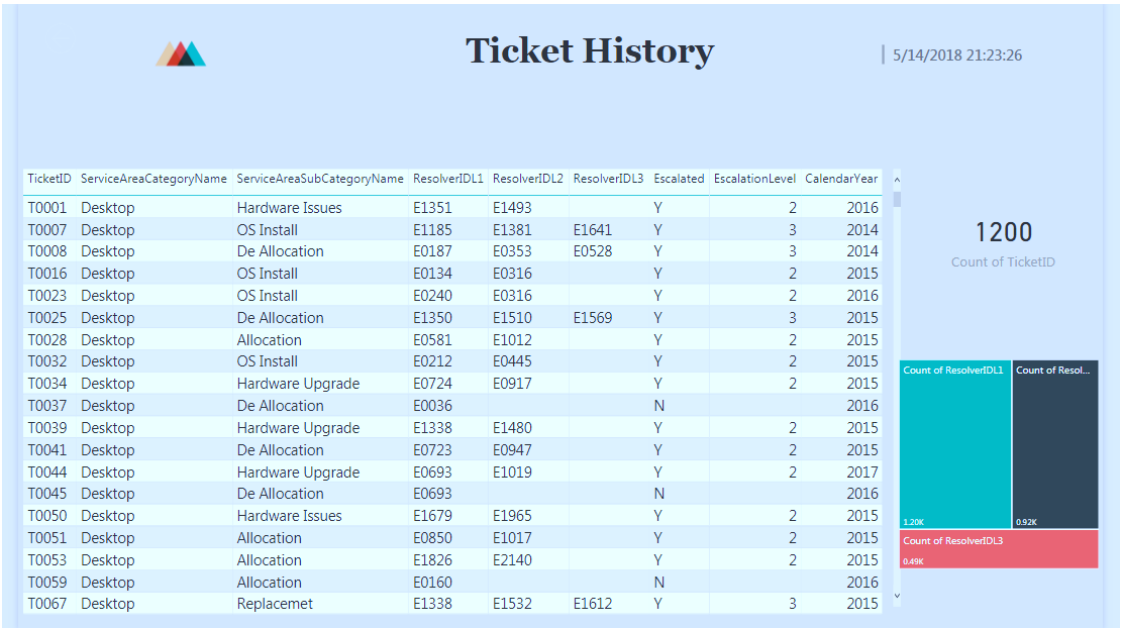


Figure 20: Escalation Rate of Tickets for a Given DC



TicketID	ServiceAreaCategoryName	ServiceAreaSubCategoryName	ResolverIDL1	ResolverIDL2	ResolverIDL3	Escalated	EscalationLevel	CalendarYear
T0001	Desktop	Hardware Issues	E1351	E1493		Y	2	2016
T0007	Desktop	OS Install	E1185	E1381	E1641	Y	3	2014
T0008	Desktop	De Allocation	E0187	E0353	E0528	Y	3	2014
T0016	Desktop	OS Install	E0134	E0316		Y	2	2015
T0023	Desktop	OS Install	E0240	E0316		Y	2	2016
T0025	Desktop	De Allocation	E1350	E1510	E1569	Y	3	2015
T0028	Desktop	Allocation	E0581	E1012		Y	2	2015
T0032	Desktop	OS Install	E0212	E0445		Y	2	2015
T0034	Desktop	Hardware Upgrade	E0724	E0917		Y	2	2015
T0037	Desktop	De Allocation	E0036			N		2016
T0039	Desktop	Hardware Upgrade	E1338	E1480		Y	2	2015
T0041	Desktop	De Allocation	E0723	E0947		Y	2	2015
T0044	Desktop	Hardware Upgrade	E0693	E1019		Y	2	2017
T0045	Desktop	De Allocation	E0693			N		2016
T0050	Desktop	Hardware Issues	E1679	E1965		Y	2	2015
T0051	Desktop	Allocation	E0850	E1017		Y	2	2015
T0053	Desktop	Allocation	E1826	E2140		Y	2	2015
T0059	Desktop	Allocation	E0160			N		2016
T0067	Desktop	Replacemet	E1338	E1532	E1612	Y	3	2015

Figure 21: Escalated Ticket History

4.11) As a Manager, I want to know the average user satisfaction so that I can analyze the department's performance.

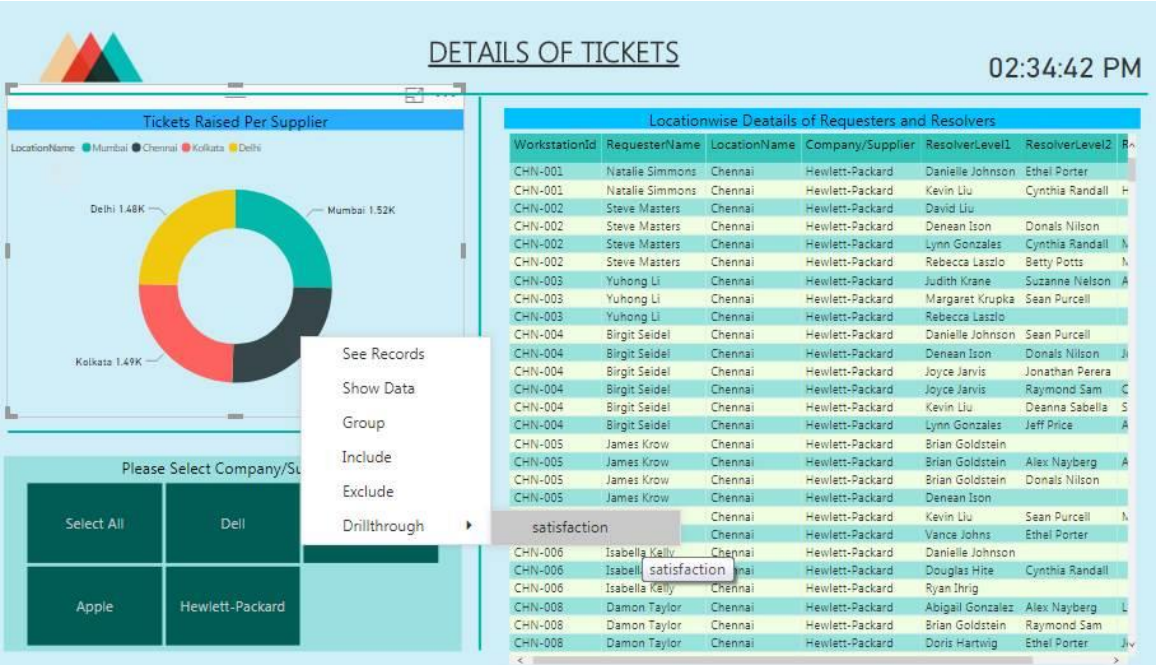


Figure 22: Overall User Satisfaction

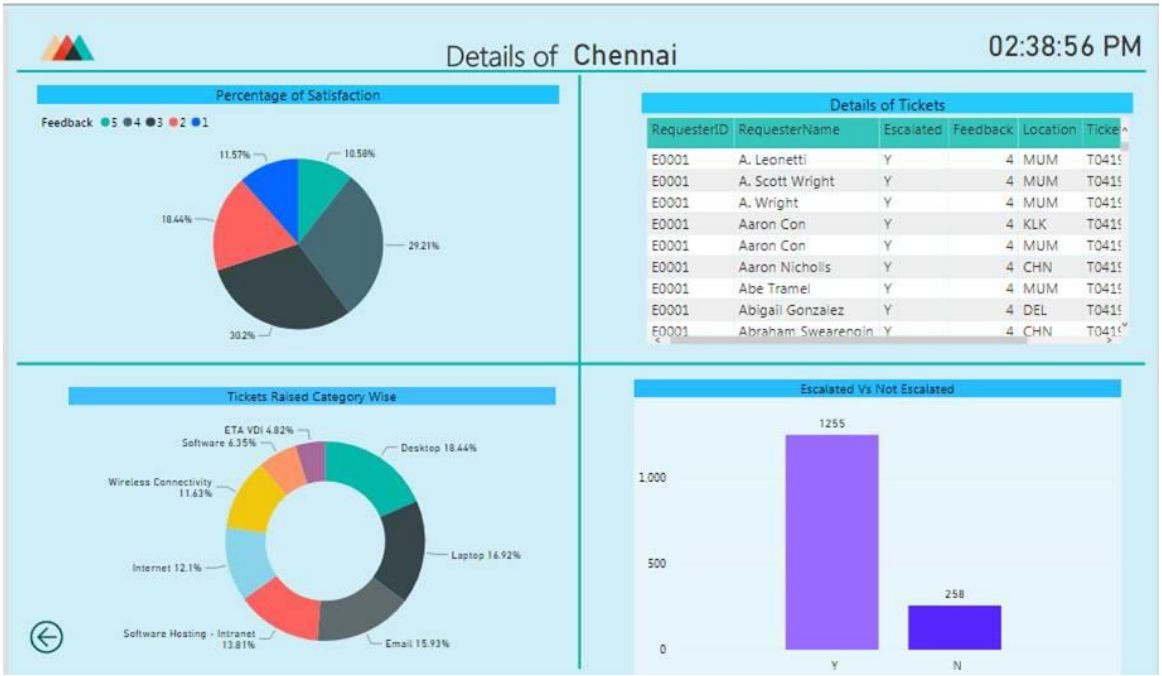


Figure 23: User Satisfaction at a Particular

4.12) As an Employee, I want to see my performance so that I can improve.

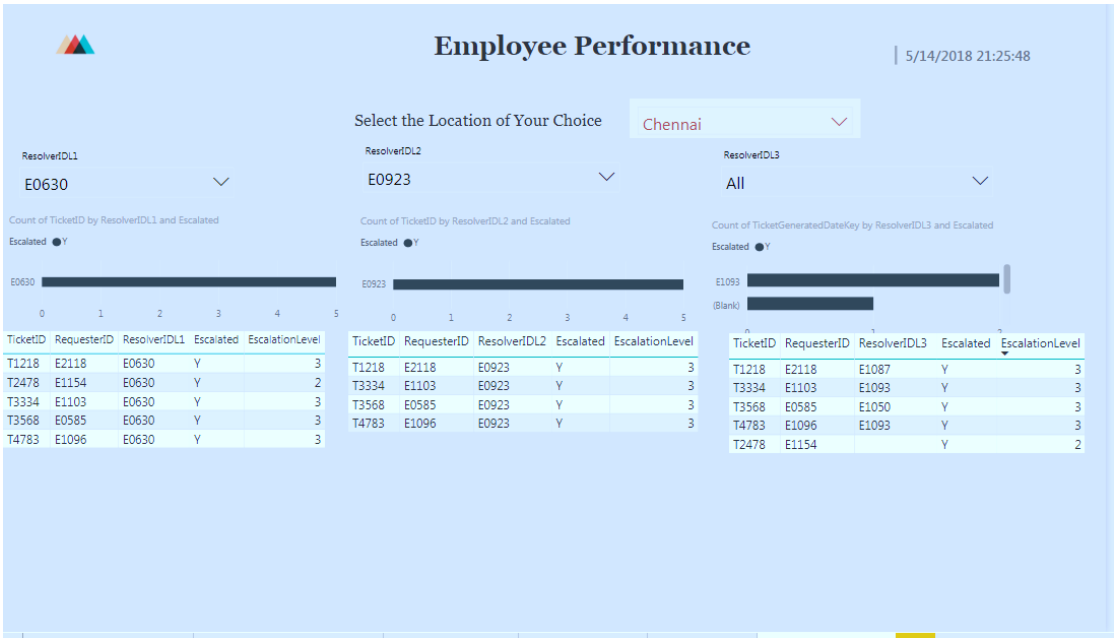


Figure 24: Employee Performance

4.13) As a User, I want to know details of the tickets I have raised previously so that I can see my request history.

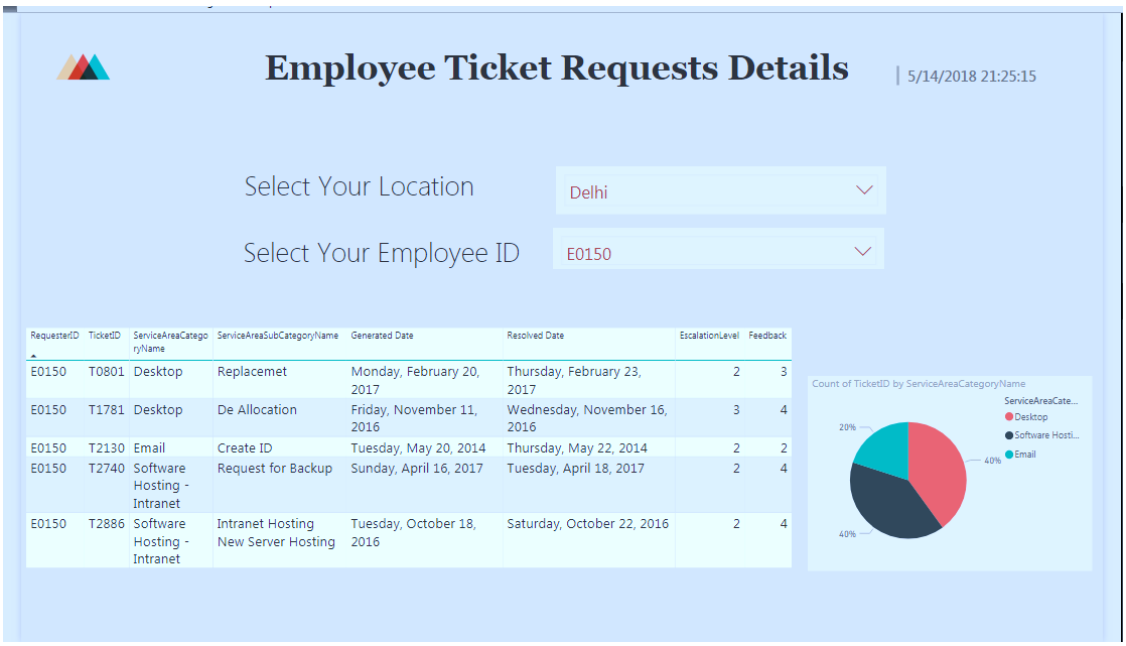


Figure 25: Employee Request History

4.14) As the CEO, I want to know the DC wise resolution so that I can analyze their performance.

Figure 26: DC Wise Resolution Rate

4.15) As the CEO, I want to know the requests raised in each request area in every DC so that I can analyze their performance.

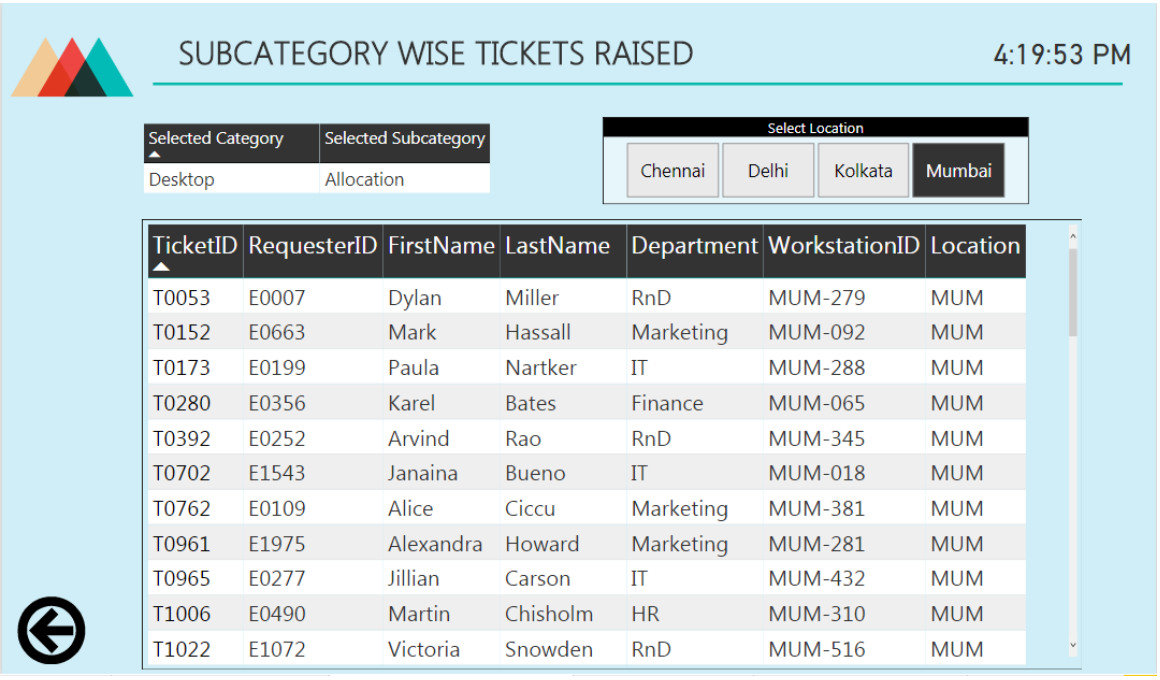


Figure 27: Details of Tickets Raised in Each Category

4.16) As the CEO, I want to know the average cost per ticket so that I can analyze expenditure per DC.

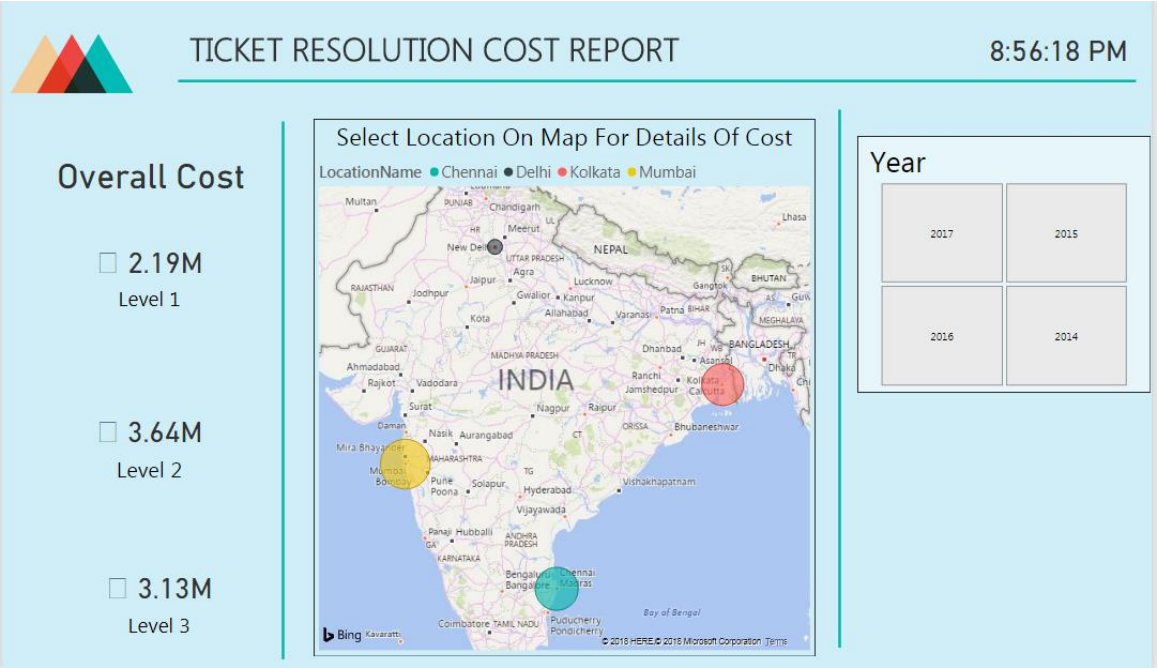


Figure 28: Cost Per Ticket Per Year

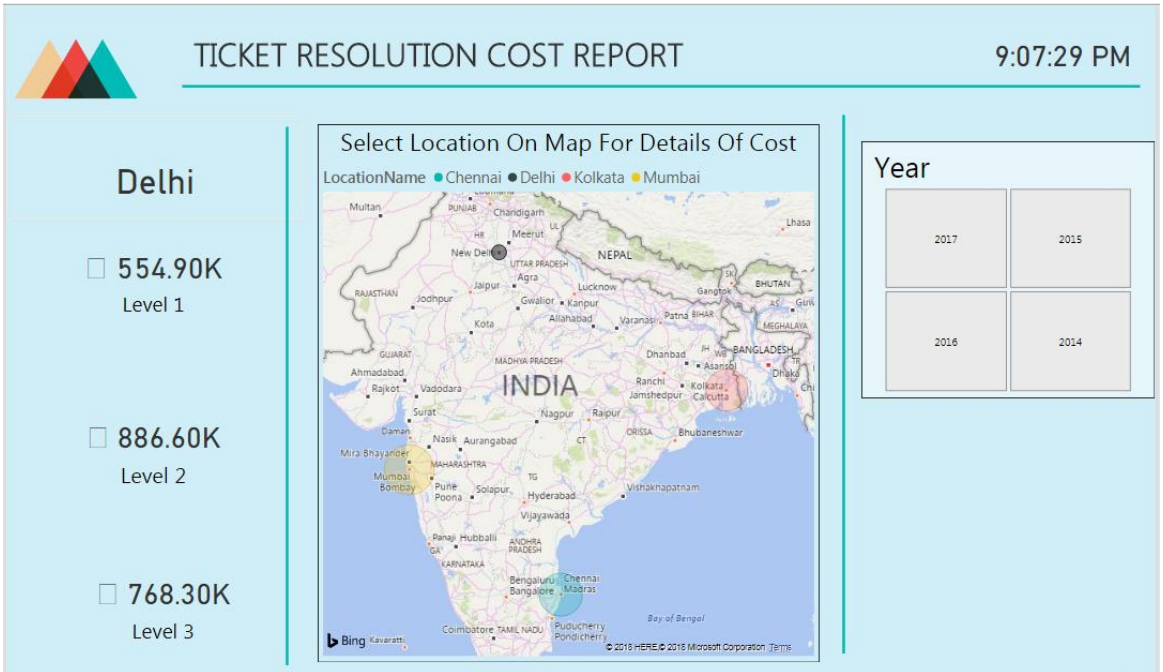


Figure 29: Cost Per Ticket Per DC Per Year

4.17) As the CEO, I want to know the number of tickets raised due to a particular supplier so that I can instruct on system improvisation.

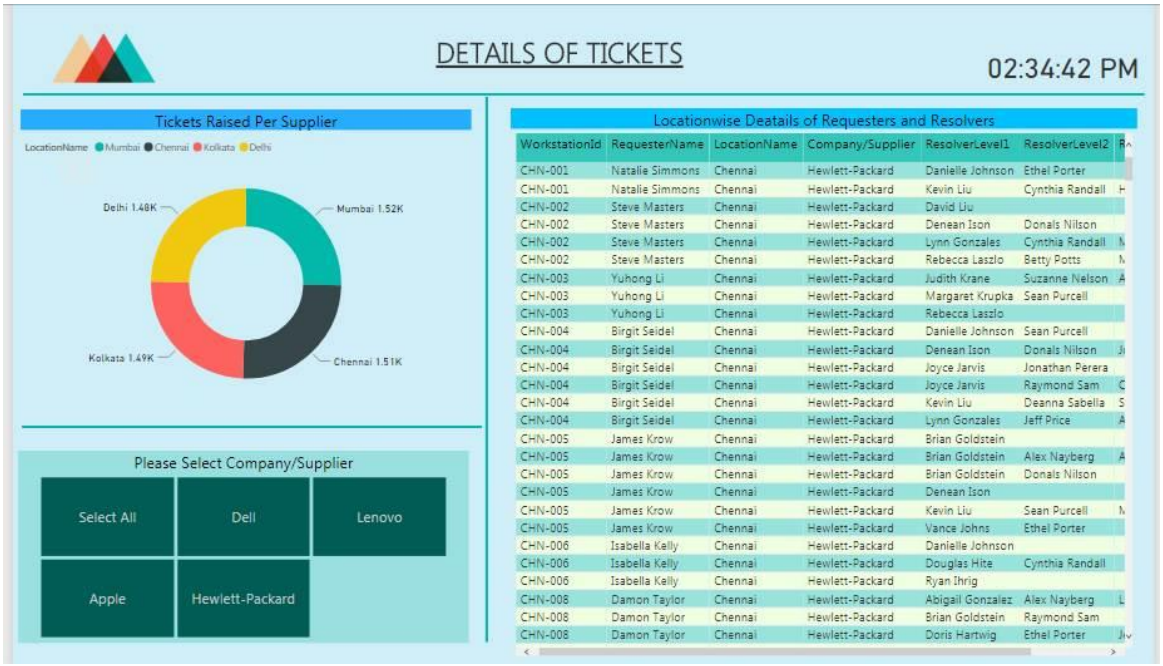


Figure 30: Tickets Raised for each Supplier

4.18) As the CEO, I want to know the number of tickets raised in every category and subcategory I can analyze which service area needs more attention.

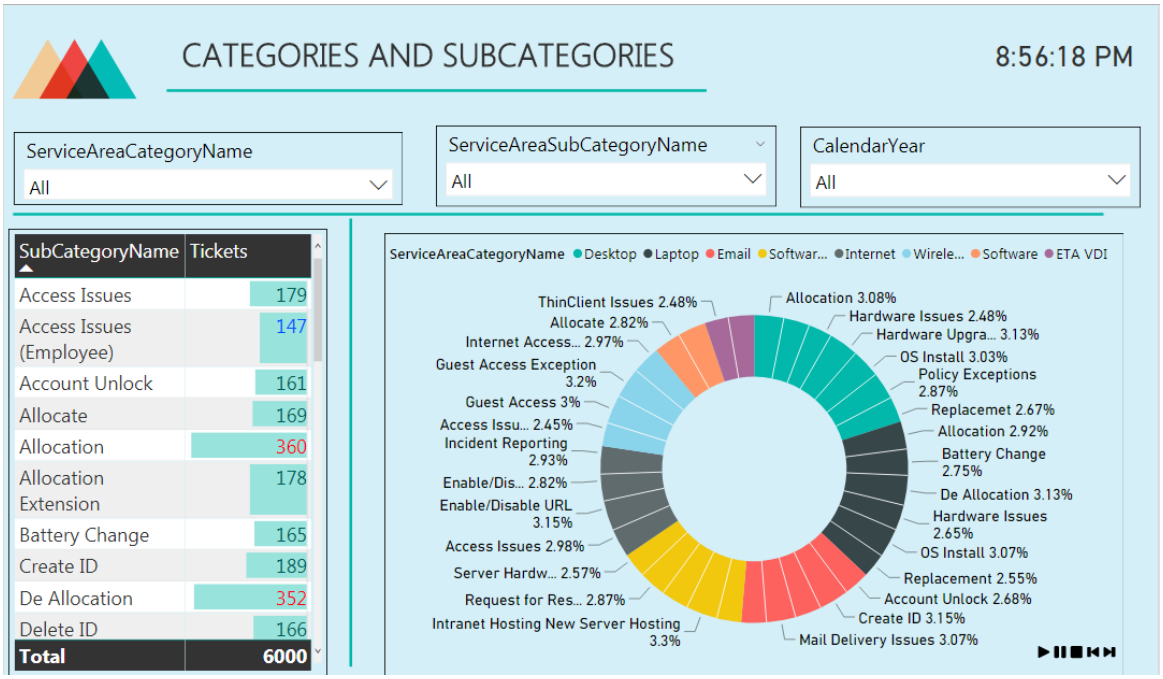


Figure 31: Tickets Raised in Each Category and Sub-Category

4.19) As a BI analyst I want to implementing security for various designations.

Figure 32: Implementing Security

4.20) As the CEO, I want to know the DC wise efficiency.

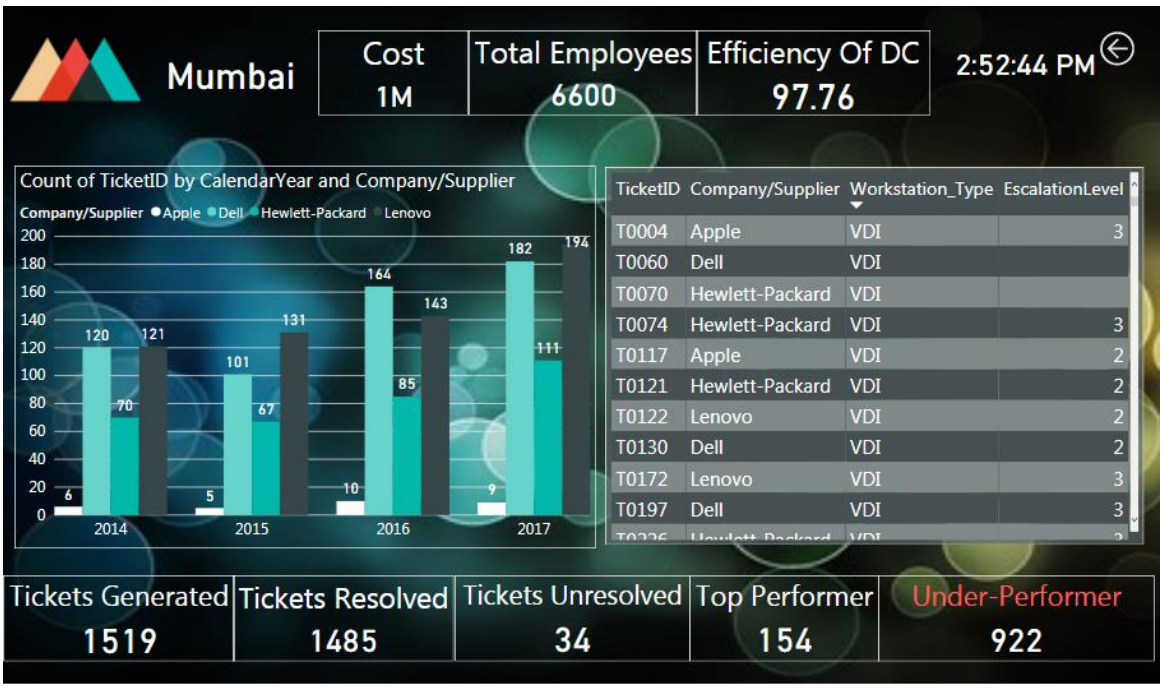


Figure 33: DC Wise Efficiency

5. CONCLUSION

This project focuses on building a data warehouse for analyzing computer maintenance tickets. Computer Maintenance takes a major role in any organization. The cost incurred in maintenance alone is usually high and it varies with the complexity of the problems raised. Without proper maintenance and support the production process of any organization is badly affected which further reduces the performance. Maximizing profit and reducing cost especially in maintenance activities is the main objective of any organization.

Analysis using proper tools helps in understanding the key factors which affects the organization in a good or a bad way. Analysis if done properly can give incredible results.

In this project we used Power BI which is a business analytics service developed by Microsoft Corporation. We created various reports which give meaningful insights and suggests actions that can improve the performance and helps in reducing cost incurred in maintenance activities. These reports allow the viewer to monitor the performance of the employees from the maintenance department which helps to know the top, average and poor performers. The reports also provide the resolution rate, cost per ticket, average cost, number of tickets raised per department, number of requests not resolved and many other views which help in making better decisions.

The viewer after analyzing these reports knows the areas where performance is low and therefore proper training is required. After knowing how many requests are not resolved in each service area and in every location the viewer now can take appropriate actions to improve the performance.

All the decisions that the viewer can make after viewing the reports is possible only after proper analysis of the maintenance tickets using the right tools.

REFERENCES

[1] Adam Wilson, Alberto Ferrari, Ariel Netz, Carl Perry, Christoph Schuler, Cindy Song, Faisal Mohamood, Jennifer Underwood, John Doyle, Josh Caplan, Kay Unkroth, Lance Delano, Marco Russo, Mey Meenakshisundaram, Michael Tejedor, Miguel Llopis, Mini Nair, Patrick Baumgartner, Sandy Rivas, Sanjay Soni, Siva Harinath, "Bring your data to life with Power BI", 2015, pp. 6-7.

[2] Mr. Dishek Mankad, Mr. Preyash Dholakia, "The Study on Data Warehouse Design and Usage", International Journal of Scientific and Research Publications, Volume 3, Issue 3, March 2013 1 ISSN 2250-3153.

[3] Surajit Chaudhuri, Umeshwar Dayal, "An Overview of Data Warehousing and OLAP Technology", ACM SIGMOD Volume 26, Issue 1, March 1997.

[4] Luna Rajbhandari, Jason Schober, "Data Warehouse and Analytics Discussion Paper",2013.

[5] Ali Radhi Al Essa, Bach, Christian, "Data Mining and Warehousing", ASEE 2014 Zone I Conference, April 3-5, 2014, University of Bridgeport, Bridgpeort, CT, USA