BTECH 451

Orion Health – Product Quality and Reliability

Priyal Bangia
Pban49

Contents

ABSTRACT

This report discusses the issue of reports being manually created as well as the data stored in unreadable, unstructured form within Orion Health who provide products to many clinics and hospitals locally and globally. The process takes a lot of time and energy, it gets difficult for Complaint Administrators and Data Analytics to carry out the process to provide meaningful data to management in a fast paced environment. Many more companies follow a similar manual procedure who are also not using the best approaches carry out processes which come with data retrieval and reporting. I have come up with my own solution to reduce the effort that goes into this using SSIS and SSRS. The report also discusses the different types of software which can be used for storing data, carrying out the ETL process and reporting. It is critical to pay attention to the complaints in a quick manner which companies receive to ensure the quality of the product is maintained, customers are satisfied and also increase due to good services provided.

INTRODUCTION

This project is entirely based on Business Intelligence. BI is the transformation of raw data into meaningful and useful information to make it easier for analyse in order to create better decision making. BI can handle enormous amounts of unstructured data to help identity, develop and create new strategic opportunities and provides historical, current and predictive business operation views. This project is aimed at providing Orion Health with an improved data retrieval and reporting service using a suitable software. The project focuses on transforming the data into a structured form where data is easy to interpret and querying (data retrieval) is quicker.

People comprehend data better through pictures than by reading numbers in rows and columns. So by visualizing data, you are able to more effectively ask and answer important questions such as “Where are sales growing,” “What is driving growth” and “What are the characteristics of my customers using different services?” With the ability to quickly answer questions, your data becomes a competitive advantage instead of an underutilized asset [5]. Data is no longer just a bunch of rows and columns. It's actually something more meaningful *which can* help quickly discover patterns in your data that unveil trends and reveal unexpected insights in order to have the focus on correcting the right area as well as prioritising how much work and effort needs to be put into certain products or issues with the product. For this Orion to get a clear view of how their products are doing. It is essential to retrieve requested reports as fast as possible.

The goal is to automate data retrieval and reporting by performing an ETL process to extract data from the database and transform it in a way where it is much easier to analyse the structure further in order to discover useful information which could help keep track of the complaints/performance of individual components by observing patterns and trends which then will help aid make decisions in the first half of the project. The second relates with research of the types of reporting software that provides support for better decision making.

THE COMPANY

This begin when founder Ian McCrae and CEO of Orion Health left his role at Ernst & Young as a Business Consultant in 1992. He wanted a job which was more interesting, challenging and would be able to make a difference in the world. Internet was not used in those days. Together with people who had a similar vision, formed a consulting firm where they obtained various IT projects related to Auckland District Health Board who required computerization and to connect all hospitals. Using their ingenuity they places microwave dishes on the roof of the hospitals to facilitate data sharing. The outcome came to be a level of inter-connectedness which had not been carried out before. The benefits and savings made by interacting was huge. They started focusing on innovation where the team focus on coming up with creative ideas and the success of projects opened up pathways for many other projects locally and globally in countries such as Australia, Spain, France, USA and many more. Orion Health offices are located in 23 countries with over 1,100 employees, the company grew at a speed of approximately 30% per year and is evolving more and more in the future.

Orion Health is now a leading global provider of HIE and Healthcare integration solutions. Products and implementations which were built by this company are used by 30+ countries.

Orion health implements and provides various facilities to hospitals such as collaborative care to securely and easily exchange clinical information with other organizations, electronic health record, hospital information system and efficiency tools. Also another product called Orion health mobile for iPhone or apple devices provides complete flexibility for healthcare professionals. All these products give an outcome of increased quality, efficiency, accuracy, collaboration of data and saved costs/risks as well as improved patient safety and experience.

THE PROBLEM

Currently, the major issue with the process is that the reports are being manually generated which takes a lot of time and energy and is very tedious, the cause of this being the data is not structure and in a readable form. Every time new complaints come in, the reports are not automated accordingly. The process of updating the data being used to generate the reports to actually creating the report is manual which contain numerous steps in between. This takes a lot of time an energy and is very tedious, it becomes difficult for Complaint Administrators and Data Analytics to carry out the process to provide meaningful data to management in a fast paced environment.

Current Process

The reporting process begins with issues coming in on a support application which Orion Health uses (Support tracker). Before an investigation, Regional analysts will complete a Pre-investigation on the issue that has come in and reading through what the customer has written. The result of that questionnaire then determines whether the report is alleged to be a problem represented by 1 or 0 emphasising yes or no which is recorded in the database table. After the investigation is carried out, regional analysts then complete a post investigation questionnaire based on the previous investigation that was executed, the issue and its fix. The result of this questionnaire determines if the alleged problem is legitimate or not (for example – a bug) which is also represented by a 1 or 0 in the table as a separate row for the post questionnaire for the same reference id(redundant data), another id called documentid (for the questionnaires) is stated to provide a unique primary key with the eventid. Therefor there are two rows for every eventid representing pre and post results. If either result is 1/yes then this reference is considered as a complaint (FDA essentially defines complaints as any allegations of issues with the product).
After a report/ticket is determined as a complaint it is read through by the global consultant for that product/application. They will 'code' this complaint down to three levels:

- Problem Code, i.e. Environment, configuration, design etc.
- Problem Area, Specific a read of the application
- Problem Detail, the specific detail within that area

This enables any analysis to be done over either a broad spectrum or highly detailed. These actions are essential for one to do as in today's world, technology is not capable to do so but the following process is where the issue occurs.

Once these processes are done, the Quality Consultant and Complaints Administrator uses the support application to query the expected output, the data then needs to be filtered and manipulated then exported into an csv file then into an excel file where the consultant has to manually create graphs, tables and analytics using excel and tableau. The graphs are then pasted onto a word document which is taken to meetings or passed onto department managers.

As the data is updated on the support system, the whole process needs to be carried out again therefore the query needs to be written out again which then is manipulated and exported into a csv file then excel or where the graphs are updated or tableau where the new versions of the graphs are created again. The Quality Consultant and Complaints Administrator carries out this cycle continuously and also when the management requires up-to-date information at the point of time or even whether is it yearly, monthly or weekly.

Therefore a lot of reports are going to be generated related to different aspects. The final reports are used to:

- Display trends over time, details and arising issues at Complaint meetings with product management
- Generate Metrics for Management
- Compare Development priorities with the main complaints in support and provide a feedback loop to ensure the right issues are being tackled by development.
- Give feedback to regional and global support teams on improvement
- A wide variety of other analytics

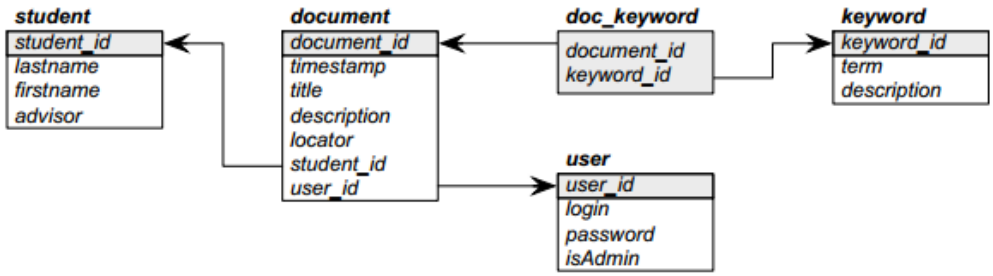
Orion health is a fast paced work environment so the problem with the process is reports are being manually generated which takes a lot of time an energy and is very tedious and to some extent, not the best approach. Time is a huge concern for large businesses. Tasks should be done as quickly as possible in order to carry on with other tasks, this becomes a concern when some tasks are to be done sequential and dependent on the other for example, the manager needs to analyse a report for a product then decides to request for a more detailed report related to the product before making any strategic or even operational decisions.

Research/Related Work -

Most commercial businesses information system are based on the relational and object oriented database models. These are called Descriptive data structures. Descriptive data structures describe the design and implementation of the information organization of non-spatial data.

Relational Data Structure

Stores information about the data as well as the way in which they are related, hence the name 'relational database'. This is the most commonly used database model used to implement a database application. Consists of a collection of tables that are linked by primary keys. The model also consists of schemas which refers to the database design or instances (snapshots of the data at a point in time). An example of a simple schema is illustrated in the image below where relations are drawn as boxes with relations attributes listed in the box. Primary Keys highlighted and foreign keys represented as arrows from the foreign keys to the referencing attributes to the primary keys. [1]



Queries can be made against this database using SQL operations and therefore due to the relations that are defined between tables, time to seek data is reduced as well as space for data storage is reduced. Normalization can be used to eliminate non-simple domains and data redundancy of data where data manipulation anomalies is prevented, this is built on the concept of normal forms such as 1NF, 2NF and 3NF.

Other benefit of using this approach:
Easy to use language – SQL plays a huge role in the success of relational models.
Conceptually Simple – features, fundamentals and concepts are easy to pick up.
Able to avoid inconsistent records.

Object Oriented Data Structure

Different O-O implementations have different data structures unlike the relational data structure where the data structure is formalized. When object oriented programming gained prominence, there was attempts to get rid of relational databases and put Object databases in place but these attempts were not successful. Now Object-relational mapping algorithms are used where Object Oriented co-exists with an underlying relational database. [1] OO consists of object structure, type constructor and object identity.

This interest was initiated by requirements in new areas of database applications. Some applications call for data modelling capabilities not supported by traditional models of databases or current implementations of database management systems (relational databases). [2]

Unlike relational databases, the data structures of objects in an object-oriented database are encapsulated (hidden) and cannot be manipulated directly by generalised procedures. This is because generalised procedures are only possible if the data structure is known and uniform over all objects, eg. the relation or table in the relational model is a known and uniform structure, allowing generalised procedures such as query operations to manipulate tabular structures independently of actual contents. Instead, each ADO presents a public set of methods that operate over its private data structures. This allows great flexibility in the design and definition of objects while at the same time allows object capabilities to be shared and re-used amongst persistent objects.

In relation to this project I personally thought Object Oriented programming can become more time consuming and quite technical as you are required to make new classes, methods, define memory in which a lot more coding is nesssary if anything changes after the project for example, new tables, relations. The Quality Consultant and Complaint Administrator is the one to be carrying out this job after and it is more beneficial to take an approach which is more simple and something that is easier to learn such as approaches traditional approaches relating to relational databases rather than to have them learn something completely new, there may be more chance of errors made.

MOMIS

Aims at creating a conceptualization of that involved local sources by means of Global virtual views.

Relevant Prototype

Based on the idea that analysing an attribute domain, we may find values which may be clustering because strongly related.
Based on the following tasks:

- Data Pre-Processing
- Similarity Computation
- Clustering technique
- Name Selection
- Decision dashboard

Informatica

The products of this software development company is based on data integration, areas such as Cloud Data Integration, Data Quality, Data virtualization, ETL, Data Replication and more. These features form a toolset for maintaining data warehouses. Informatica ETL tool 'power center" offers many features of data transformation such as lookups, sorting/ aggregating data, transforming data with functions. Using this platform the ETL process is simplified, allows you to initiate ETL Projects in a cost effective manner. Comparatively this has greatest flexibility of your ETL process with the ability to extract more enterprise data types.
Informatica allows easy training and tool availability for software companies, most of the other ETL tools lack in this aspect. Some tools are very expensive whereas some just have to face challenges such as debugging and ease of use, looking at it in this perspective, Informatica is an ideal ETL tool. Success deployment ratio is quite high, close to 100%. During a ETL tool survey 2014, Informatica gained 4 out of 5 stars over all. For a physical table storage, the rows are partitioned.

Oracle

Oracle Warehouse Builder (OWB) is another ETL tool incorporated in Oracle which allows you to build, manage and maintain data integration processes. Further it offers capabilities for data auditing, graphical environment for relational/dimensional modelling.
Oracle Business Intelligence Enterprise Edition 11g, provides all the capabilities of that of BI Publisher so I considered this thereafter. Not only just reporting this supports many capabilities including OLAP Analytics and Enterprise reporting, access, analysis and integrated web environment. Provides a common infrastructure for all these things as well [4].

Reporting based features:

- Interactive Dashboard – Provides interactive dashboards and reports with a rich variety of visualizations. Can drill, navigating, pivot and filter. The end user is working with prompts, tables, graphics and reports in a web architecture.

- Self-serve Interactive Reporting – Enables users to create new analysis from scratch or modify existing analysis without any help form IT.
- Enterprise reporting – This is delivered through BI Publisher itself. Creation of highly formatted templates, reports. Report layouts can be easily generated directly onto a web browser using BI Publisher layout in conjunction. This does not require approrientry design studio, additional costs and reduced learning curve. BI Publisher is solely focused on reporting. Data can be combined from separate data sources into a single Data Model sharable across many reports. In this case, this capability can be beneficial as I will be using Microsoft SQL Server where the data will be stored in a star schema structure. Not an ETL tool but it can manipulate logical schemas instead of physical schemas. BI Publisher is able to sort the data and format it but this is for report purposes so Microsoft SQL server is necessary for where the data is stored and BI Publisher can be used on the surface for reporting. Output formats for the document generation can be selected between interactive, HTML, Excel, Powerpoint, PDF and batch report generation [3].

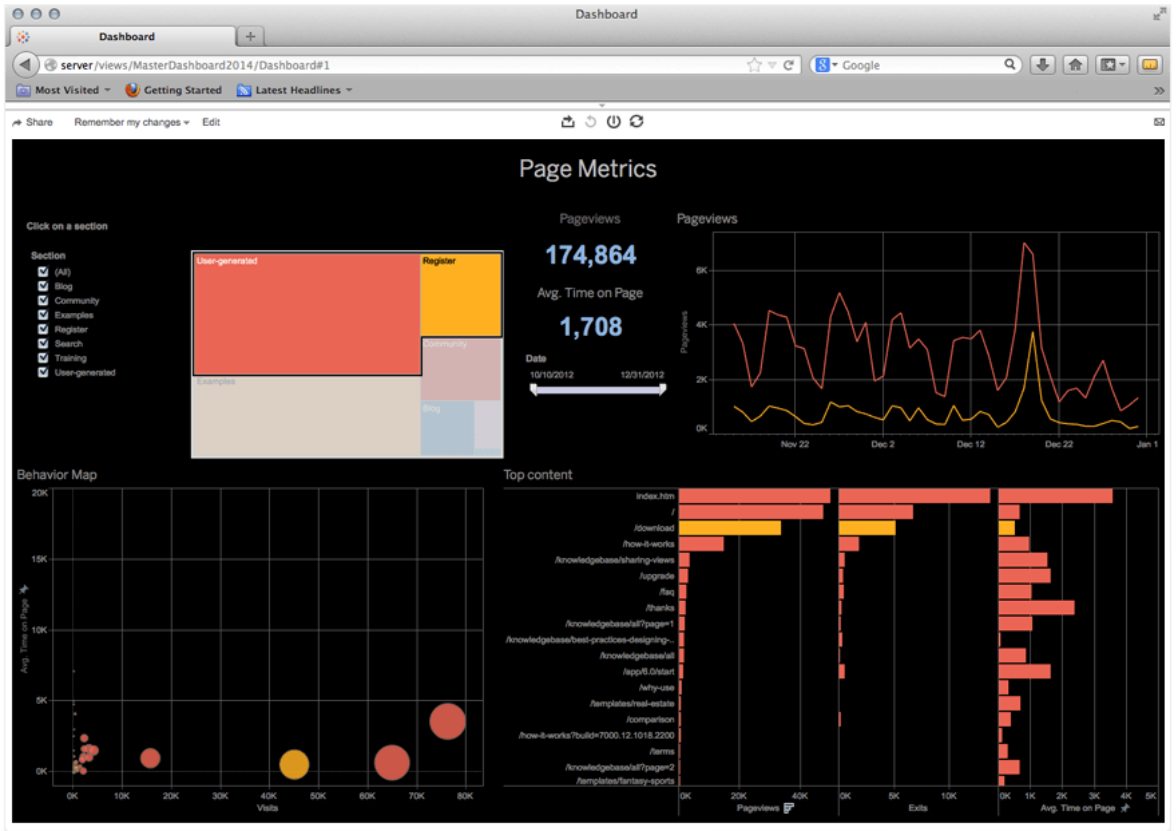
  The reports can be viewed online, sent to particular printers, fax or emails as compared to before where the Quality Consultant and Complaint Administrator will have carried out a process of first creating the graphs and tables on excel then copy and paste these onto Microsoft Word then format it then finally print/email it. Even better, reports can be scheduled. This saves time in case a product manager requests for in-depth details regarding their product and has a huge impact on time when weekly reports are already schedules to be generated, the administrator does not need to update the database first as that is going to be automated unlike the current system there is a need to update the data then create updated reports on top of that.

  The layout can support 150 languages and includes automatic internationalization of date and time which can be helpful for collaboration with overseas branches of Orion Health.

Tableau

Tableau's aim is to equip people to help them see and understand data. This is a software which helps build reports and visualizations in tableau desktop then publish to tab server for sharing. A range of capabilities relating to visual analytics in areas such as data discovery, data visualization, Business Dashboards, Mobile Business Intelligence and Time Series Analysis. Time Series Analysis plays a huge role in business decisions where you are able to perform time comparisons, find trends and analyse by time, week, month or year.

A huge advantage that comes along with Tableau is the ease of use, the software is very user friendly. It is not a must that the user has to have technical skills in order to use this effectively, people such as business users, senior management are able to use this as well. A reason for this is that you are able to create queries without having to physically code due to a drag and drop facility. Whereas SSIS is more developer centric reporting application. Similar to Oracle and unlike the current process in place updates are done automatically so you are able to get up to date data with a live connection and also on a schedule if defined. Another feature which is similar was to be able to share live on the web, using Tableau mobile devices can also be used for this purpose. It would be more interactive and professional for the administrator to create this and share the web Brower around to the department regarding their product rather then printed pieces of paper of graphs posted on a word document. Orion health uses this software as well as Excel so I thought it would be good to use Tableau and explore more features this semester when I carry out part 2 (reporting) of my project.



SOLUTION

Automating both report and data retrieval will free up a lot more resource time and effort. On this note the accuracy will also increase. To achieve this, initially I researched the different ETL techniques using different software which were discussed in the research section of this report. I have also researched about Microsoft SQL Server which I did not

discuss in the previous section but I will discuss why I chose this approach in this section as this is the software I have decided to use throughout the first half of my project which is data integration using SSIS. The second half of my project deals with reporting and using the best software and approach for reporting which also revealed more about the products and benefits management. This half of the project is to be done this semester although so far I have included about my findings so far relating to reporting.

Microsoft SQL server

This is a commercial relational database management system whose function is to store and retrieve data as requested. Access is similar but it is a database which is more of a database management system that permits you to build an interactive UI where the end user is able to enter data and report, however Microsoft SQL Server does not have a front end but allows you to store data. SQL Server manages many services such as relational databases, Integration services, Analysis Services and Reporting Services. Data is stored in a database. The SQL server upon it is able to support the different types of data types (text, numeric, float, char, Boolean). The data in the storage space is separated into numbered pages of size 8Kb. The page is marked with a header which consists of page number, page type, used and free space and other metadata about the page. Larger volumes of data can be stored including objects such as video, text and photographs, Access has the same capability but it can handle only 1Gigabyte of data as compared to Microsoft SQL server which can handle terra bytes.

Concurrent access is taken care of by SQL server in two modes, first being pessimistic concurrency and the other, optimistic concurrency. Locks are used when there is need to control access in the pessimistic mode. Optimistic concurrency control differs in the sense that it allows a new version of a row to be created when the row is updated instead of replacing or overriding the row.

Many more exist as mentioned earlier but not all are necessary to function on the database, it depends on what you are looking for. The add-on services that I will be looking at is to do with data integration and reporting. Microsoft SQL Server supports better security features such as strong authentication, efficient access control, permission management tools and better performance which is why I choose to keep with Microsoft SQL server instead of Access due to the large amount of confidential data Orion Health has to store.

Business Intelligence Development Studio

BIDS is the IDE (Integrated Development Environment) for Microsoft utilizing the Reporting, integrating services which means it is a software that provides facilities to programmers for software development. It is equipped with service specific extensions such as projects for reporting services will use this type of extension and things such as data mining structures will use extensions corresponding to Analysis services.

SSRS (SQL Server Reporting Services)

This service generates reports in a RDL file from the data collected from the SQL server database. Microsoft Visual Studio produces the same outputs as long as Business Intelligence Development Studio is installed. We are not limited to RDL files when it comes to output, RDL file can be rendered in any other format such as PDF, XML, CSV and more.

SSIS (SQL Service Integration Services)

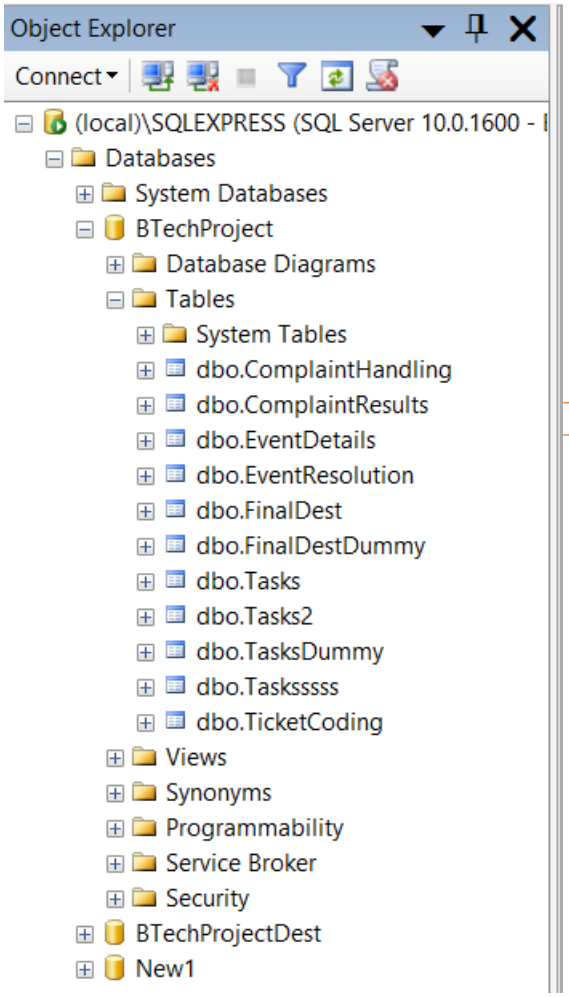
DTS was build upon to contain more transformation features and was replaced by SSIS in 2005. This allows for ETL capabilities for SQL Servers. SSIS is consists of GUI tools to create workflows such as extracting data, querying data, transforming data, exporting data to destination. Used entirely for Data integration and workflow applications. To avoid out problem of taking to much time for reporting which concerned the data stored not kept up to date, SSIS resolves this issue as it used to automate maintainace for SQL server databases and updates. Using this will keep the data up to date rather than having the Admin to update every time a report needs to be generated causing the whole cycle to take place continuously.

Because Business Intelligence Development studio is the IDE for integrated services, SSIS can be developed though this. There may be various data sources from which the data needs to be considered, this will not be an issue when using SSIS as one is permitted to use heterogeneous data sources so this is beneficial for the future when necessary. This can also be applied in my case where I was given separate tables with rows and columns in a csv format, I was able to import and preform SSIS operations effectively using Business Intelligence Development Studio without any issues.

However SSIS is more developer centric as compared to software such as Tableau but Microsoft SQL Server gives the ability to store data, preform data integration using SSIS and also reporting services using SSRS all in one place whereas if I was to use Tableau I would still require a place for storing data and then use Tableau only as a surface. Similar case with Oracle, I did not go ahead with that because Orion Health is already using Microsoft SQL Server to perform their transactional data even though if it is not in the best structure so it would be best to use the same platform and prove the system from there rather than to shift all the data into an Oracle platform therefore I decided not to use Oracle Business Intelligence Enterprise for the ETL process as this would also be a cross platform just like using Tableau with Microsoft SQL Server. So to keep everything all together in a effective and efficient manner.

Implementation

I was given 6 tables with data, using Microsoft SQL Management Studio I connected to my own server using (local)\SQL.EXPRESS and created a new database and imported those CSV files to create tables and also an empty database for which the outcome will reside.



These are the 6 tables I have been working with, tasks is the table where some transformations take place and the data gets stored within in this. FinalDest is the table the final output resides in at this stage to make sure everything is coming out as expected.

These tables contain metadata about certain things which is also shown by the following images of the tables (dimension tables – CSV formatting) which you will have to zoom into see the details.

EventDetails

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	eventId	applicationId	Application	versic	Client	RequestType	Priority	Responsible	State	Diagnosis	Arrival	synopsis	contactPersonName	clientId	SupportStaff	SupportStaffId
2	ABCDE	00:00.0	RHA	5.2.0	VIS	Issue Report	Level 2 - Production Degradation	NA Region	Closed	Other	49:28.5	Rhapsody Production Crash	Mano M	abc	Michael Stanford	michael.s
3	BCDEF	00:00.0	RHA	5.2.0	BAN	Issue Report	Level 3 - Functional Impact	NA Region	Resolved	Product Bug	27:06.1	Rhapsody 5.2 Crash 12/31/2012	Christof C	abc	Terry Pan	terry.p
4	HIJKL	00:00.0	RHA	3.2.0	BAN	Issue Report	Level 1 - Production Failure	EMEA Region	Closed	Environment Issue	57:29.6	rhapsody failure	Dasarsh D	def	Terry Pan	terry.p
5	IJKLM	00:00.0	RHA	5.2	VIS	Technical Question	Level 4 - No Production Impact	NA Region	Closed	Not Specified	32:13.5	Question on comm point icon in Management Console	Sid S	ghi	Tony D	tony
6	JKLMN	00:00.0	RHA	5.1	SAT	Technical Question	Level 4 - No Production Impact	NA Region	Closed	Config Advice	12:13.1	Want to test messages	Saad S	jkl	Michael Stanford	michael.s
7	MNOPQ	00:00.0	RHA	5.2.0	JOT	Issue Report	Level 3 - Functional Impact	EMEA Region	Closed	Environment Issue	38:21.0	Rhapsody not processing messages	Rina R	mn	Terry Pan	terry.p
8	OPQRS	00:00.0	RHA	4.0.0	SAT	Issue Report	Level 3 - Functional Impact	EMEA Region	Closed	Environment Issue	01:18.6	Interface problem?	Dave D	opq	Tony D	tony
9	CDEFG	00:00.0	RHA	5.4	JOT	Issue Report	Level 3 - Functional Impact	EMEA Region	Closed	Product Bug	10:41.8	CAnnot add new segments to HL7 with database looku	Saketh S	res	Samuel Wills	samw
10	KLMNO	00:00.0	RHA	3.4.0	CHR	Issue Report	Level 4 - No Production Impact	NA Region	Closed	Account Requests	46:34.2	License file sent not working	Mandy M	tuv	Michael Stanford	michael.s
11	PQRST	00:00.0	RHA	5.4.0	CHR	Technical Question	Level 3 - Functional Impact	EMEA Region	Closed	Release request	30:29.5	NHS Communication Point	Kunaal K	wxy	Terry Pan	terry.p
12	GHIJK	00:00.0	RHA	5.4	RIN	Technical Question	Level 4 - No Production Impact	NA Region	Closed	Product Bug	52:13.7	SNMP MIB and Documentation	Ashwin A	sdk	Samuel Wills	samw
13	FGHIJ	00:00.0	RHA	5.1.0	VAD	Technical Question	Level 1 - Production Failure	EMEA Region	Closed	Environment Issue	51:58.9	we have a production interface down and need urgen	Nizam N	kse	Michael Stanford	michael.s
14	EFGHI	00:00.0	RHA	5.2.0.(	DAV	Issue Report	Level 1 - Production Failure	EMEA Region	Resolved	Product Bug	51:02.6	Incorrect Date in Acknowledgements	Lawrence L	knf	Felisa Pan	fel
15	DEFGH	00:00.0	RHA	3.4.0	RIN	Issue Report	Level 3 - Functional Impact	EMEA Region	Closed	Environment Issue	24:53.5	rhapsody 3 not processing trak adt messages	Xu X	mwo	Terry Pan	terry.p
16	LMNOP	00:00.0	RHA	5.4.0	SIV	Issue Report	Level 3 - Functional Impact	NA Region	Closed	Other	13:39.7	JavaScript	Michael M	wee	Michael Stanford	michael.s
17	GHIJK	00:00.0	RHA	5.5.0	SAA	Issue Report	Level 4 - No Production Impact	NA Region	Suspended - Client Feedback	Not Specified	45:40.0	Unclean Shutdown Detected	Mano M	boo	Michael Stanford	michael.s
18	DEFGH	36	RHA	5.5.1	NIZ	Issue Report	Level 4 - No Production Impact	NA Region	Resolved	Product Bug	12:16.3	Server app aborts	Christof C	mei	Tony D	tony
19	RSTUV	138	LAB	8.3	NIZ	Technical Question	Level 4 - No Production Impact	Projects Admir	In Progress	Not Specified	45:34.5	Interpreted value from HL7 not update in HIS	Dasarsh D	pll	Felisa Pan	fel
20	STUVW	150	RAD	8.3.13	SAA	Issue Report	Level 2 - Production Degradation	HIS Region	Suspended - Client Feedback	Not Specified	10:25.8	Cannot send to PACS	Saketh S	bbb	Terry Pan	terry.p

ComplaintHandling

EventId	vigilanceR	fieldActior	complaintClosure
ABCDE	00:00.0	1	1
BCDEF	00:00.0	1	1
CDEFG	00:00.0	1	1
DEFGH	00:00.0	0	1
EFGHI	00:00.0	1	1
FGHIJ	00:00.0	1	1
GHIJK	00:00.0	1	1

TicketCoding – relates to when the Gobal Consultant manually codes a complain into categories such as 'Envrionment', 'Defect', 'By Design' and also states related details.

	A	B	C	D	E	F	G	H
1	EventId	problemCode	problemCodeDesc	problemArea	problemAreaDesc	problemDetail	problemDetailDesc	lastModified
2	ABCDE	DEFECT	Defect	JVM	JVM	NULL	NULL	00:00.0
3	BCDEF	UNDERINV	Under Investigation	JVM	JVM	NULL	NULL	00:00.0
4	HIJKL	ENV	Environment	NED	Not Enough Detail	NULL	NULL	00:00.0
5	MNOPQ	ENV	Environment	CPFILTER	Comm Point & Filters	NULL	NULL	00:00.0
6	CDEFG	DEFECT	Defect	CPFILTER	Comm Point & Filters	NULL	NULL	00:00.0
7	OPQRS	ENV	Environment	CPFILTER	Comm Point & Filters	NULL	NULL	00:00.0
8	PQRST	BYDES	By Design	CERT	Certificates	NULL	NULL	53:18.4

Tasks

	A	B	C	D	E
1	EventId	taskType	taskStatus	taskStatusName	documentId
2	ABCDE	Pre Investigation	4	Completed	e164a660-e18e-4b83-b910-c7526318d19b
3	ABCDE	Complaint Handling	4	Completed	f6899545-c85e-4abc-8e58-f763fdb9e24b
4	ABCDE	Post Investigation	1	New	NULL
5	BCDEF	Pre Investigation	4	Completed	9d22193c-514c-445b-a51d-c9f49a40eedc
6	BCDEF	Post Investigation	1	New	3ab13829-0ba1-469d-b36d-03425bcad505
7	BCDEF	Complaint Handling	4	Completed	e9d1f9ac-0e26-4a58-9be3-9c1f19a60d0a
8	HIJKL	Pre Investigation	4	Completed	5e1b8855-9ee2-45d2-b72b-111665b30d33
9	HIJKL	Post Investigation	4	Completed	2bfb41b0-0746-4048-ad73-c02dd4ee984d
10	IJKLM	Pre Investigation	4	Completed	5470f5a1-bacd-4f26-b045-a87897292f9e
11	IJKLM	Post Investigation	3	InProgress	77f411ed-a4db-486e-bda8-9295c85dbe46
12	JKLMN	Pre Investigation	4	Completed	8bd54caa-a879-4316-bb61-89faeced5c43
13	JKLMN	Post Investigation	1	New	NULL
14	MNOPQ	Pre Investigation	4	Completed	6417fced-2182-4799-9edb-efce47f6315d
15	MNOPQ	Post Investigation	3	InProgress	9325e321-eaca-4b96-9376-ba700aaf2fe4
16	CDEFG	Pre Investigation	4	Completed	3857bbb6-db80-4376-b53a-3c04efd010d2
17	CDEFG	Complaint Handling	4	Completed	9b9c2f63-e7e5-4a2c-bad9-975407cbdf23
18	CDEFG	Post Investigation	1	New	7fcc0876-a0e2-47d0-82e2-081f382825f8
19	KLMNO	Pre Investigation	4	Completed	552c39ba-8874-4299-ae14-e251c01d7aed
20	KLMNO	Pre Investigation	1	New	3a127ace-f3cb-4a4f-8078-71d43ea2d8f0
21	KLMNO	Post Investigation	4	Completed	77eb5038-a475-408c-a1c7-aa7e998ac999
22	KLMNO	Complaint Handling	1	New	bb106e1c-b0b4-49cb-841f-6fe8b9523b48
23	PQRST	Pre Investigation	1	New	331833a6-c3d7-49a4-8146-f3a37852ce0a
24	PQRST	Complaint Handling	1	New	cca35c87-b661-4fb4-b7e2-8d43d709686c
25	GHIJK	Pre Investigation	4	Completed	ead1a230-f1f3-4639-bb7a-e5fc675eefed
26	GHIJK	Complaint Handling	4	Completed	4138563a-e735-4b30-ade0-68950d8d1fb6
27	GHIJK	Post Investigation	1	New	cdcbb262-91db-4741-9573-5d307933821e
28	FGHIJ	Pre Investigation	4	Completed	666cf23f-9c29-457e-9635-2fdb6dbe43d2

ComplaintResult – has multiple rows for 1 eventid for a complaint to illustrate the results of pre and post investigation. Note: Eventid is left empty due to confidential purposes.

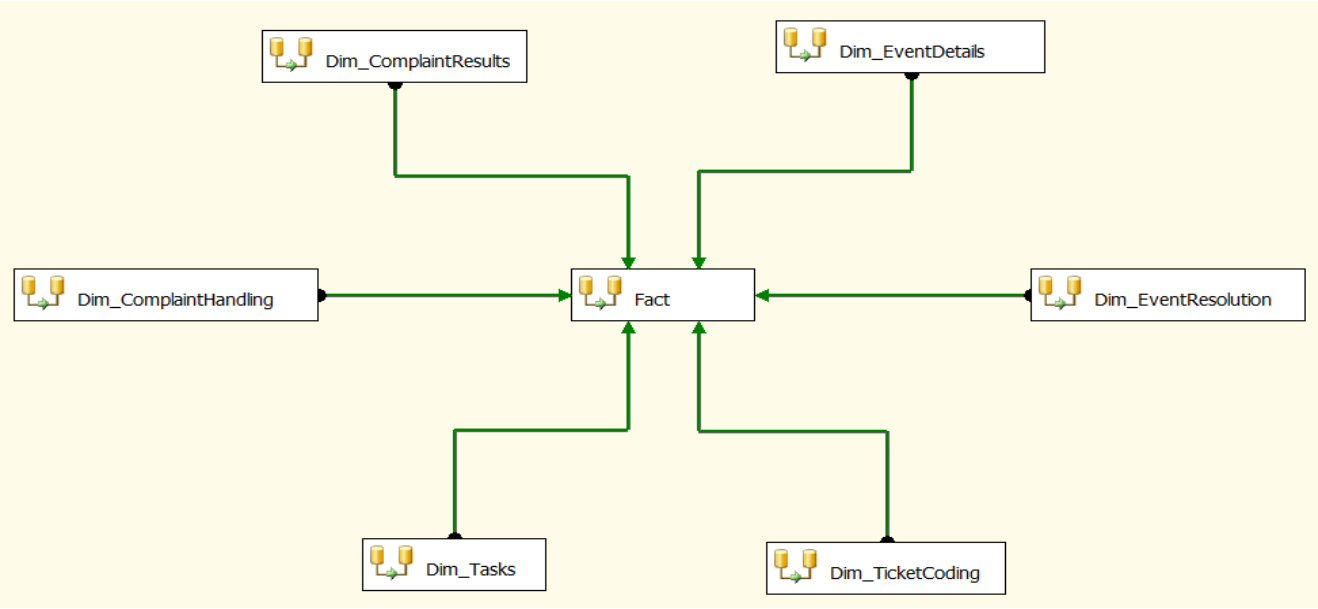
documentId	EventId	Type	question1	question2	question3	question4	question5	question6	Result	CompletedDate
2bfb41b0-0746-4048-ad73-c02dd4ee984d	00:00.0	POST	0	0	0	0	00:00.0	0	0	28:32.7
3857bbb6-db80-4376-b53a-3c04efd010d2	00:00.0	PRE	1	1	0	NULL	NULL	NULL	1	22:06.5
42b3858a-4d58-4321-8fc4-11657a5aecac	00:00.0	PRE	0	1	0	NULL	NULL	NULL	1	02:00.6
5470f5a1-bacd-4f26-b045-a87897292f9e	00:00.0	PRE	0	0	0	NULL	NULL	NULL	0	30:10.8
552c39ba-8874-4299-ae14-e251c01d7aed	00:00.0	PRE	0	0	0	NULL	NULL	NULL	0	37:18.5
5e1b8855-9ee2-45d2-b72b-111665b30d33	00:00.0	PRE	0	0	0	NULL	NULL	NULL	0	28:19.1
6417fced-2182-4799-9edb-efce47f6315d	00:00.0	PRE	0	0	0	NULL	NULL	NULL	0	50:41.8
666cf23f-9c29-457e-9635-2fdb6dbe43d2	00:00.0	PRE	0	1	0	NULL	NULL	NULL	1	00:08.7
77eb5038-a475-408c-a1c7-aa7e998ac999	00:00.0	POST	1	0	1	1	1	1	1	38:33.2
8bd54caa-a879-4316-bb61-89faeced5c43	00:00.0	PRE	0	0	0	NULL	NULL	NULL	0	58:37.8
8d6cc01c-3922-43aa-b392-6654d0ee417d	00:00.0	PRE	1	0	0	NULL	NULL	NULL	1	17:24.1
9d22193c-514c-445b-a51d-c9f49a40eedc	00:00.0	PRE	0	1	0	NULL	NULL	NULL	1	21:47.6
e164a660-e18e-4b83-b910-c7526318d19b	00:00.0	PRE	1	1	0	NULL	NULL	NULL	1	47:18.0
e30751fc-c7ab-4401-8f3c-e853a6b05213	00:00.0	PRE	0	1.00E+00	0	NULL	NULL	NULL	1	48:54.6
ead1a230-f1f3-4639-bb7a-e5fc675eefed	00:00.0	PRE	0	1	0	NULL	NULL	NULL	1	37:29.6
f4e85a02-1510-4794-a178-a1cd616e1a86	00:00.0	PRE	0	1	1	NULL	NULL	NULL	1	16:40.9
f608b0a9-430f-473b-873d-44a34ed53ba3	RSTUV	PRE	0	1	0	NULL	NULL	NULL	1	29:48.7

EventResolution – specifies when a complaint has been resolved.

	A	B
1	eventId	ResolvedDate
2	HIJKL	46:34.3
3	IJKLM	26:32.9
4	JKLMN	15:41.3
5	KLMNO	00:44.9
6	GHIJK	58:52.2
7	FGHIJ	48:48.0
8	LMNOP	15:37.8
9		

Star schema

I had decided to use a schema where the data can be seen as a blueprint of the structure of the database to keep formed integrity. I have used one fact table which refers to 6 other dimension tables (the 6 tables in the previous image) which makes it easier to query compared to a snowflake schema. The dimensions are the descriptive attributes related to fact data and the fact table holds the important information relating to complains as you will see in an output image later on in the report.

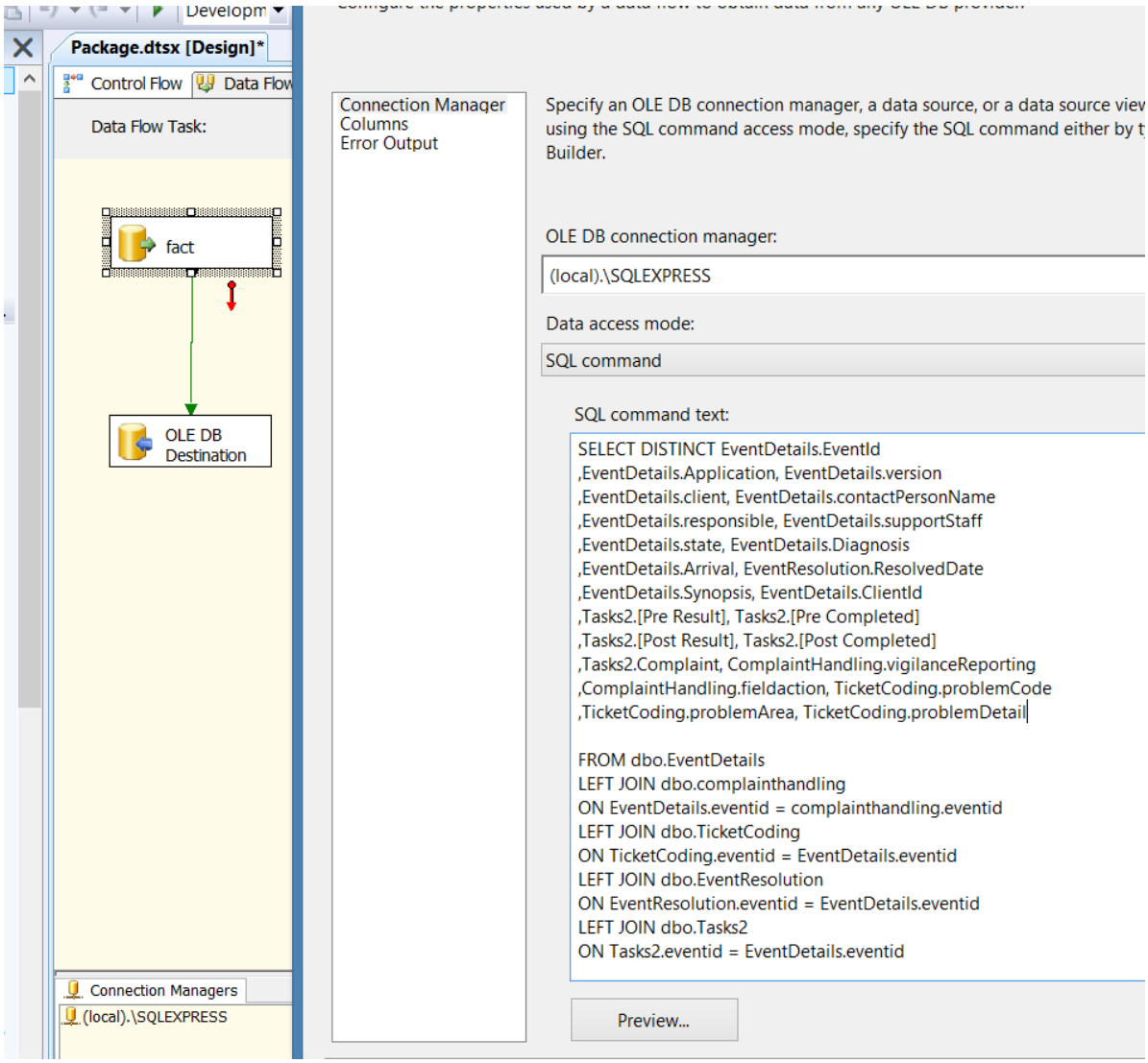


Fact table consists of low level of uniform detail whereas dimension define a range of characteristics. For example in this case the following dimensions are:

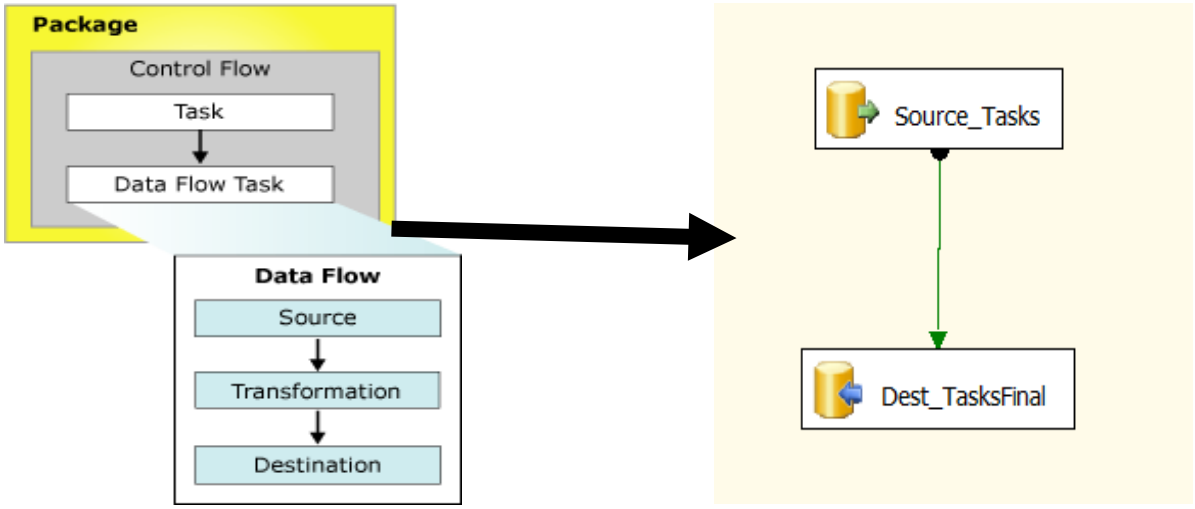
- Complaint Handing related details such as closer, field of action.
- Complaint results – consists of the pre and post results depending on the questionnaire with the results and date/time completed.
- Tasks – states the type of task (either pre/post investigation) and weather the task has been completed with a corresponding documented.
- Event Details – this means attributes related to an complaint that has come in, includes details such as application it concerns, client, request type (technical/report), priority of the complaint from 1-4 (Production failure to No production impact), diagnosis, synopisis, staff responsible for this issue and their id's.
- Event Resolution – the date a problem has been resolved of that eventid.
- And lastly, Ticket Coding which expands more on the problem code (DEFECT, BYDES etc) for the complaint. Expresses what it stand for and the problem details of a complaint.

Attributes which were required for the daily reports that are being generated consists within the fact table and therefore queries required for reporting will be used against this information rather than the individual tables themselves. Here is the structure of the output (not the real output due to confidential data). My supervisor carried out this process using the current system and I carried this out using SSIS and both of our outcomes matched out with the right amount of rows and transformed columns. There is a huge benefit to my approach now, as complaints come in, the database is automatically updated which is going

to be working towards the goal of automated data retrieval and reporting rather than the manual process each time. The following images shows the sql query for retrieving the valid information within the fact table using Joins on the 6 table. The outcome however is not accurate as it is a structure of what I retrieve due to confidential purposes and you may have to zoom-in in order to see details for the outcome.



Control flow contains workflow of tasks to be executed. However, there a need for transformation between the Extract and Load steps with are done in the Data flow for a table. Data flow defines the flow of data from source to destination where data flows can also flow between your selected entities, source, transformation, and destination.



Looking at the following images for the complaintResult table(shown in figure 1) that specifies the type, pre investigation, post investigation (different rows for the same complaint), the outcome for each question, the result (if it an actual issue) done by the Regional Analyst. Because the data is redundant and structured in a very inefficient manner. I took the data from the table and transformed it into what I will be displaying in the final outcome and stored it into the tasks table as there is also a inner join to obtain the eventid between the two tables as they are not stated in the complaintResult table. This transformation is used to ensure ease when using queries for the final output. The transformation that took place was to have pre and post columns for one eventid (shown in figure 2) with the data entry of yes or no according to the result for example the pre column will contain a yes if the result from the questionnaires was 1 or no if 0. To do this I had to split the complaintResult table into two tables, one with eventid and Pre investigation and the completed date, same for the post investigation.

	EventId	Pre Investigation	CompletedDate
1	ABCDE	1	47:18.0
2	BCDEF	1	21:47.6
3	CDEFG	1	22:06.5
4	DEFGH	1	17:24.1
5	EFGHI	1	16:40.9
6	FGHIJ	1	00:08.7
7	GHIJK	1	37:29.6
8	HIJKL	0	28:19.1
9	IJKLM	0	30:10.8
10	JKLMN	0	58:37.8
11	KLMNO	0	37:18.5
12	LMNOP	1	02:00.6
13	MNOPQ	0	50:41.8
14	RSTUV	1	29:48.7

Process –

Obtain a table with only Pre Investigation from

the source table using this SQL query:

```
SELECT DISTINCT t.EventId, cr.Result, cr.CompletedDate
FROM  dbo.Tasks AS t INNER JOIN
          dbo.ComplaintResults AS cr ON cr.documentId = t.documentId
WHERE (cr.Type = 'PRE')
```

	EventId	Post Investigation	CompletedDate
1	HIJKL	0	28:32.7
2	KLMNO	1	38:33.2

Obtain a table with only Pre Investigation from the source table using this SQL query:

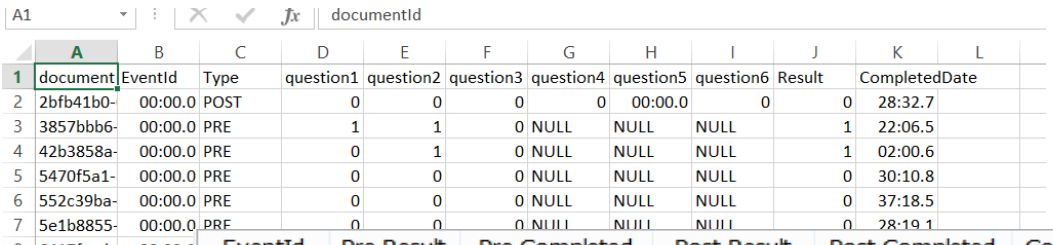
```
SELECT DISTINCT t.EventId, cr.Result, cr.CompletedDate
FROM  dbo.Tasks AS t INNER JOIN
          dbo.ComplaintResults AS cr ON cr.documentId = t.documentId
WHERE (cr.Type = 'POST')
```

	EventId	PRE	Pre Complet...	POST	Post Res...	Complaint
1	ABCDE	Yes	47:18.0	No	NULL	Yes
2	BCDEF	Yes	21:47.6	No	NULL	Yes
3	CDEFG	Yes	22:06.5	No	NULL	Yes
4	DEFGH	Yes	17:24.1	No	NULL	Yes
5	EFGHI	Yes	16:40.9	No	NULL	Yes
6	FGHIJ	Yes	00:08.7	No	NULL	Yes
7	GHIJK	Yes	37:29.6	No	NULL	Yes
8	HIJKL	No	28:19.1	No	28:32.7	Not Started
9	IJKLM	No	30:10.8	No	NULL	Not Started
10	JKLMN	No	58:37.8	No	NULL	Not Started
11	KLMNO	No	37:18.5	Yes	38:33.2	Yes
12	LMNOP	Yes	02:00.6	No	NULL	Yes
13	MNOPQ	No	50:41.8	No	NULL	Not Started
14	RSTUV	Yes	29:48.7	No	NULL	Yes

This is now the new Tasks table formed using the SQL Query:

```
SELECT DISTINCT
        pre.EventId, (CASE pre.[Result] WHEN 1 THEN 'Yes' WHEN 0 THEN 'No' ELSE 'No'
END) AS PRE, pre.CompletedDate AS [Pre Completed],
        (CASE post.[Result] WHEN 1 THEN 'Yes' WHEN 0 THEN 'No' ELSE 'No' END) AS POST,
post.CompletedDate AS [Post Result], (CASE WHEN post.[Result] = 1 OR
        pre.[Result] = 1 THEN 'Yes' ELSE 'Not Started' END) AS Complaint
FROM  dbo.Pre AS pre LEFT OUTER JOIN
        dbo.Post AS post ON post.EventId = pre.EventId
```

Original table Vs transformed table.



	A	B	C	D	E	F	G	H	I	J	K	L
1	document	EventId	Type	question1	question2	question3	question4	question5	question6	Result	CompletedDate	
2	2bfb41b0-	00:00.0	POST	0	0	0	0	00:00.0	0	0	28:32.7	
3	3857bbb6-	00:00.0	PRE	1	1	0	NULL	NULL	NULL	1	22:06.5	
4	42b3858a-	00:00.0	PRE	0	1	0	NULL	NULL	NULL	1	02:00.6	
5	5470f5a1-	00:00.0	PRE	0	0	0	NULL	NULL	NULL	0	30:10.8	
6	552c39ba-	00:00.0	PRE	0	0	0	NULL	NULL	NULL	0	37:18.5	
7	5e1b8855-	00:00.0	PRE	0	0	0	NULL	NULL	NULL	0	28:19.1	
8	6417fced-	00:00.0										
9	666cf23f-9	00:00.0										
10	77eb5038-	00:00.0										
11	8bd54caa-	00:00.0										
12	8d6cc01c-	00:00.0										
13	9d22193c-	00:00.0										
14	e164a660-	00:00.0										
15	e30751fc-(	00:00.0										
16	ead1a230-	00:00.0										
17	f4e85a02-	00:00.0										
18	f608b0a9-	RSTUV										
19												
20												

Figure1

Figure 2

EventId	Pre Result	Pre Completed	Post Result	Post Completed	Complaint
ABCDE	Yes	47:18.0	Not Done	NULL	Yes
BCDEF	Yes	21:47.6	Not Done	NULL	Yes
CDEFG	Yes	22:06.5	Not Done	NULL	Yes
DEFGH	Yes	17:24.1	Not Done	NULL	Yes
EFGHI	Yes	16:40.9	Not Done	NULL	Yes
FGHIJ	Yes	00:08.7	Not Done	NULL	Yes
GHIJK	Yes	37:29.6	Not Done	NULL	Yes
HIJKL	No	28:19.1	No	28:32.7	Not Sta...
IJKLM	No	30:10.8	Not Done	NULL	Not Sta...
JKLMN	No	58:37.8	Not Done	NULL	Not Sta...
KLMNO	No	37:18.5	Yes	38:33.2	Yes
LMNOP	Yes	02:00.6	Not Done	NULL	Yes
MNOPQ	No	50:41.8	Not Done	NULL	Not Sta...
RSTUV	Yes	29:48.7	Not Done	NULL	Yes

CONCLUSION

Reports are generated on a weekly/monthly basis to take into meetings where complaints are discussed with managers. For these the data is firstly updated manually then queried and transformed to import into a CSV file and finally into Excel or Tableau to create graphs

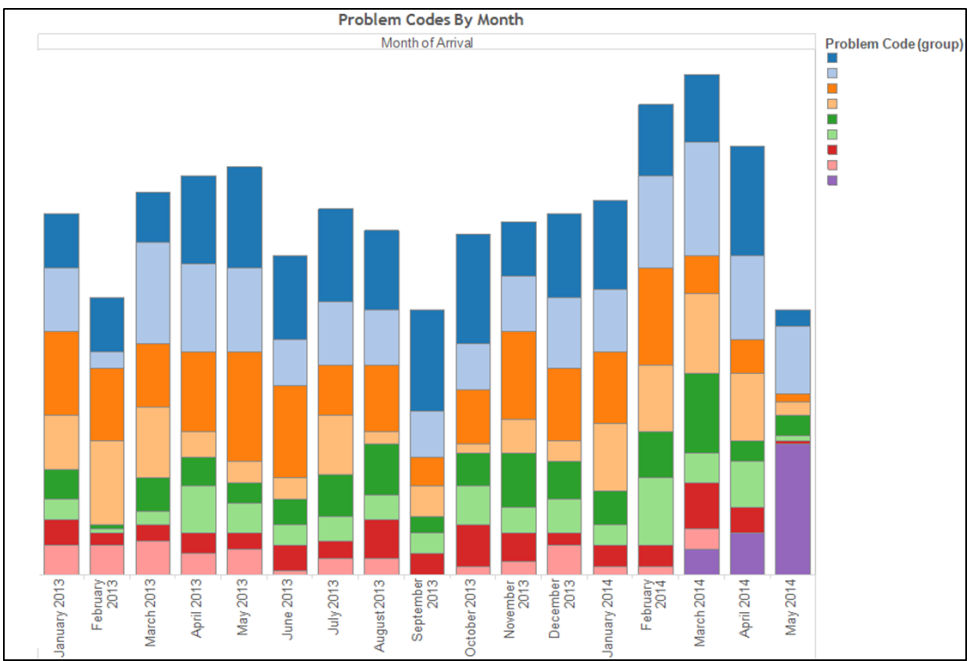
and tables. These graphs are then placed in word for analysis, meetings and to be passed on to managers on request. This process very time consuming as it is and also when this cycle has to be carried out by the Quality Consultant and Complaint Administrator on a regular basis.

The Quality Consultant and Complaint Administrator gets held up carrying out these steps when this can be made more quicker and efficient so resources. By avoiding these unnecessary processes can give the Quality Consultant and Complaint Administrator to proceed with other important tasks that have to be completed to improve products and services. Orion Health is a huge company so there are many tasks to be done, this inefficient process can be dealt with so many software platforms which are available.

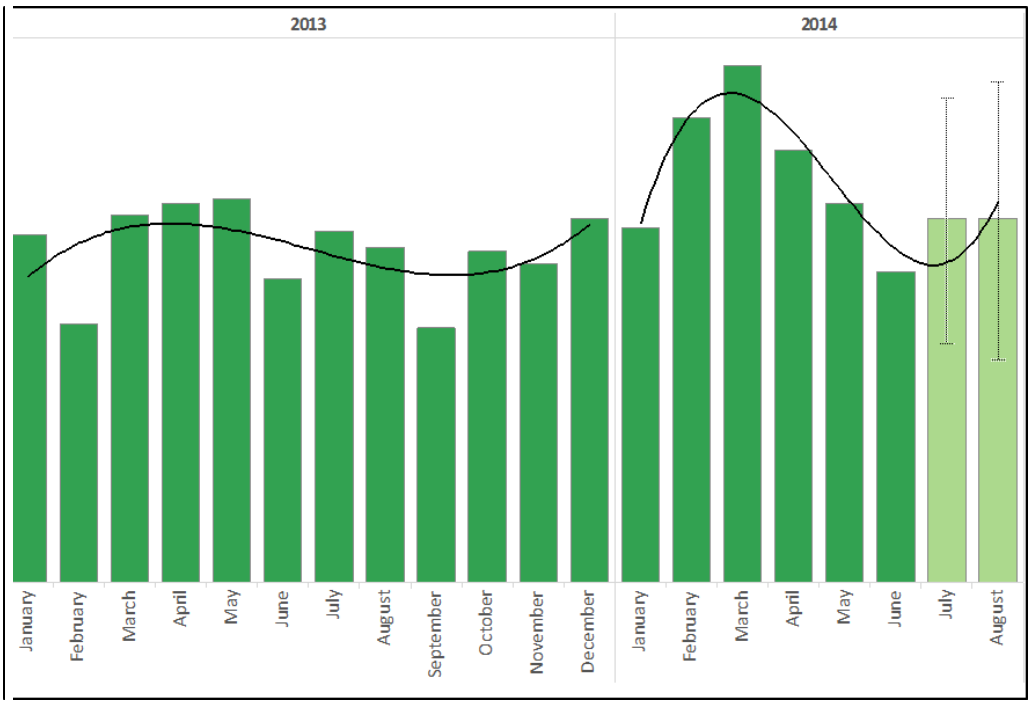
Having a fast process for data retrieval and reporting will help Orion Health in decrease defects and improving they're products. The generation of these types of analytics, everybody is able to look at the same data in the same way and they can always drill down and see what they want to look at, decision-making happens much faster [6]. Product managers are able to ask for more detailed reporting based on the product they are responsible therefore having the data structured using a star schema makes data retrieval easier and faster. Even just by interpreting the data itself may show significant changes during a time period for each of the products, components or services that Orion Health provides. Using data visualizations for reporting helps uncover deeper factors and shed light on the service the product has been providing over a time frame e.g. if the problem is in fact something other than what the product manager has interpreted to be but complaints are still arising due to the fact that the real concern is another factor. Time and effort may have been wasted on improving aspects which do not affect the product.

During this semester I will be looking at more reporting tools and software as I have done with SSRS, Tableau and Oracle Publisher. I am looking to do more research on Visual Studio, Test Track and Assembla. Out of these I will use one software which will be best suited for the requirements of Orion Health based on their daily reporting use and also products as well. For example, I will look to generate simple graphs which are regularly used such as:

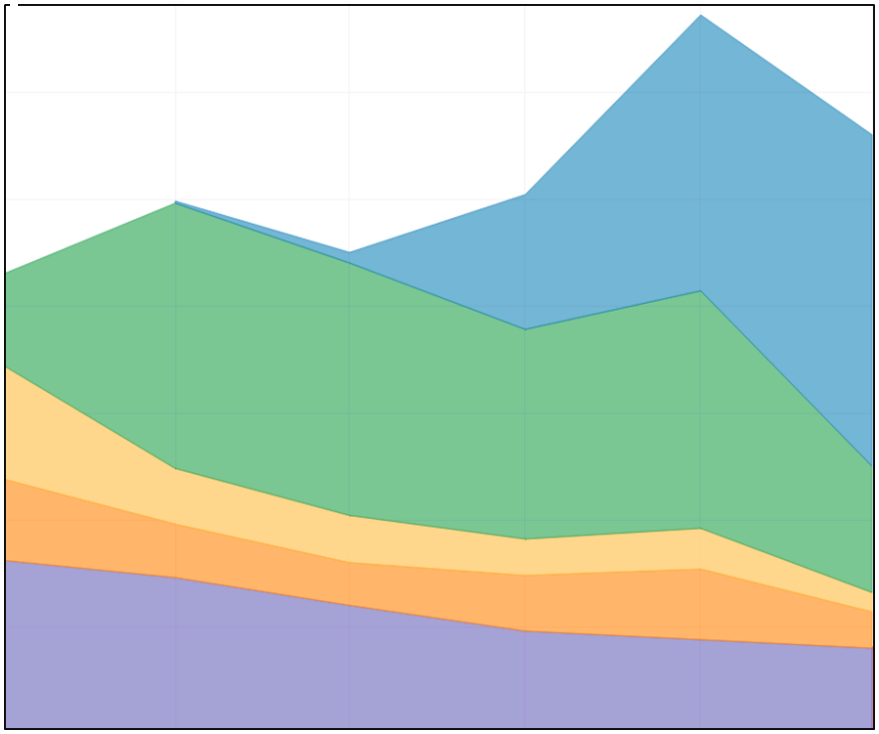
- Problem Codes during a period of time: Looks over the overall complaints over a given time grouped by problem codes (Defect, environment, configuration)



- Complaints Over Time – illustrates the total volume of complaints grouped by products over time with trend lines and forecasting for the future time to come.



- Complaint Versions Over Time: Versions are released for a products, this looks at the versions of an application that the complaints are raised.



Bibliography

1. Relational Database Model. (2014). [online] Available at: http://myweb.lmu.edu/dondi/share/db/relational1.pdf [Accessed 14 Aug. 2014].
2. Oriented Object Modelling. (2014). [online] Available at: http://coronet.iicm.edu/is/scripts/lesson04.pdf [Accessed 14 Aug. 2014].
3. BI Public. (2014). [online] Available at: http://www.oracle.com/technetwork/middleware/bi-publisher/overview/bipublisherdatasheet-129370.pdf?ssSourceSiteId=ocomen [Accessed 14 Aug. 2014].
4. BI Enterprise Edition. (2014). [online] Available at: http://www.oracle.com/us/bi-enterprise-edition-plus-ds-078848.pdf [Accessed 14 Aug. 2014].
5. Data Visualization. (2014). [online] Available at: http://www.tableausoftware.com/solutions/data-visualization [Accessed 14 Aug. 2014].
6. Tableau. (2014). [online] Available at: http://www.tableausoftware.com/learn/stories/seagate-innovates-putting-data-hands-business-users [Accessed 14 Aug. 2014].
7. Tableau. (2014). [online] Available at: http://www.tableausoftware.com/learn/stories/seagate-innovates-putting-data-hands-business-users [Accessed 14 Aug. 2014].
8. Tableau VS Microsoft. (2014). [online] Available at: http://viz-analyis.blogspot.co.nz/2011/10/tableau-vs-microsoft-bi.html [Accessed 14 Aug. 2014].

9. Oracle. (2014). [online] Available at: http://www.oracle.com/us/bi-enterprise-edition-plus-ds-078848.pdf [Accessed 14 Aug. 2014].