IT16032316

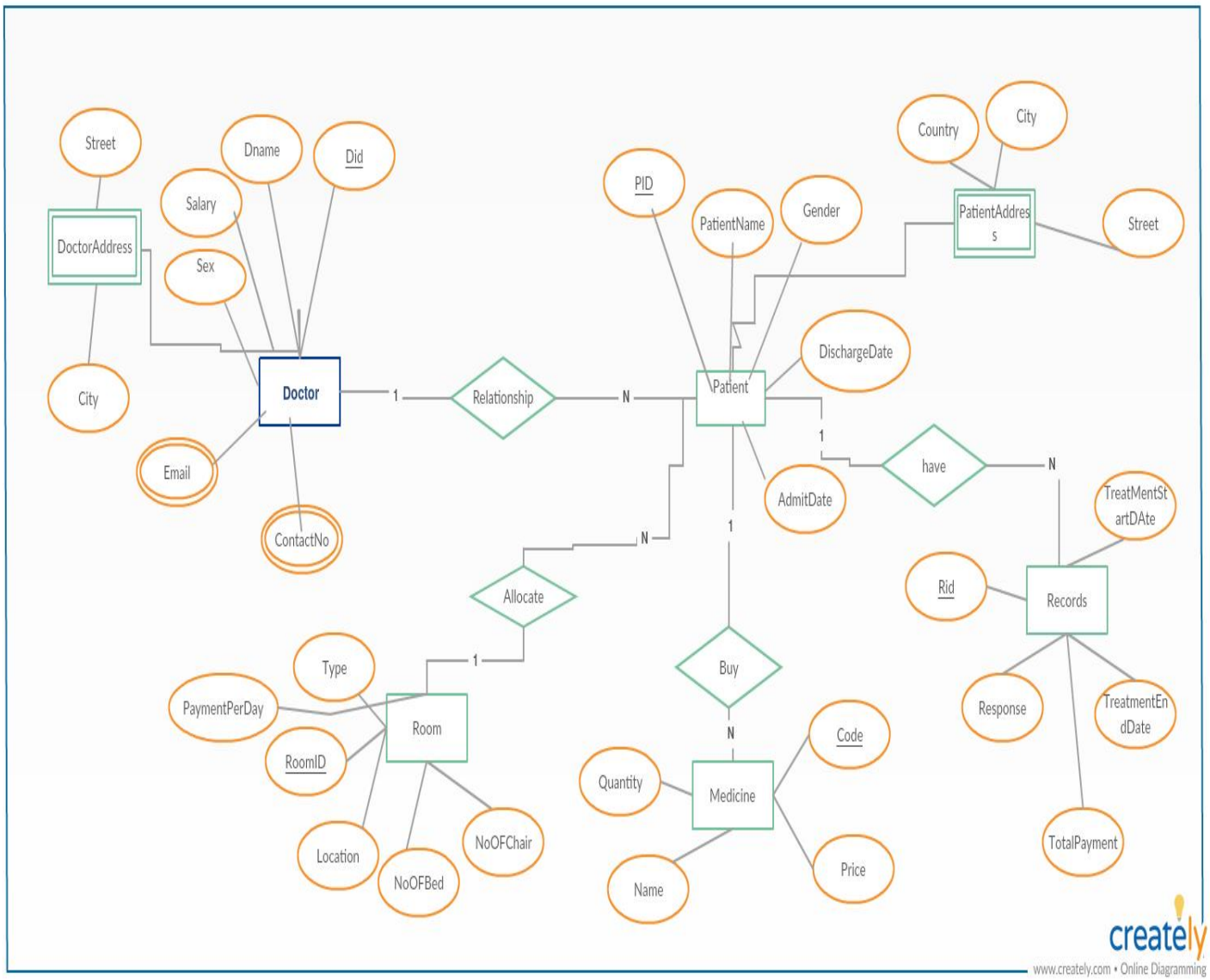
DWBI ASSIGNMENT-1

2018.05.04

The Dataset is about Hospital Management System Mostly Focused on Patient Management. So The Main Tables Took for this process are

- Patients Records Table
- Patients Details Table
- Doctor Details Table
- Medicine Details Table
- Rooms Table
- Doctors Address Table
- Patient Address Table
- Records Table

These Tables are mapped according to this scenario. This scenario is in the form of ER DIAGRAM Below



These Files are in xls and txt Format as Given Below:

Patients =>Patients.xls

Patients Address=>PatientsAddress.xls

Doctors=>Doctor.xls

Medicine =>Medicine.xls

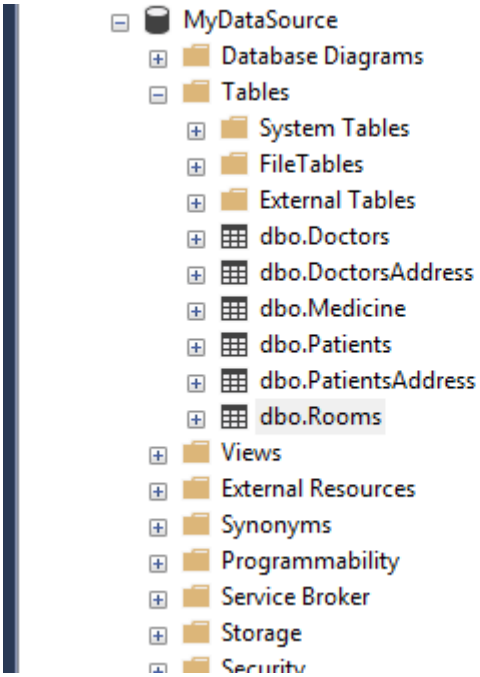
Rooms=>Rooms.xls

Doctors Address =>DoctorsAddress.xls

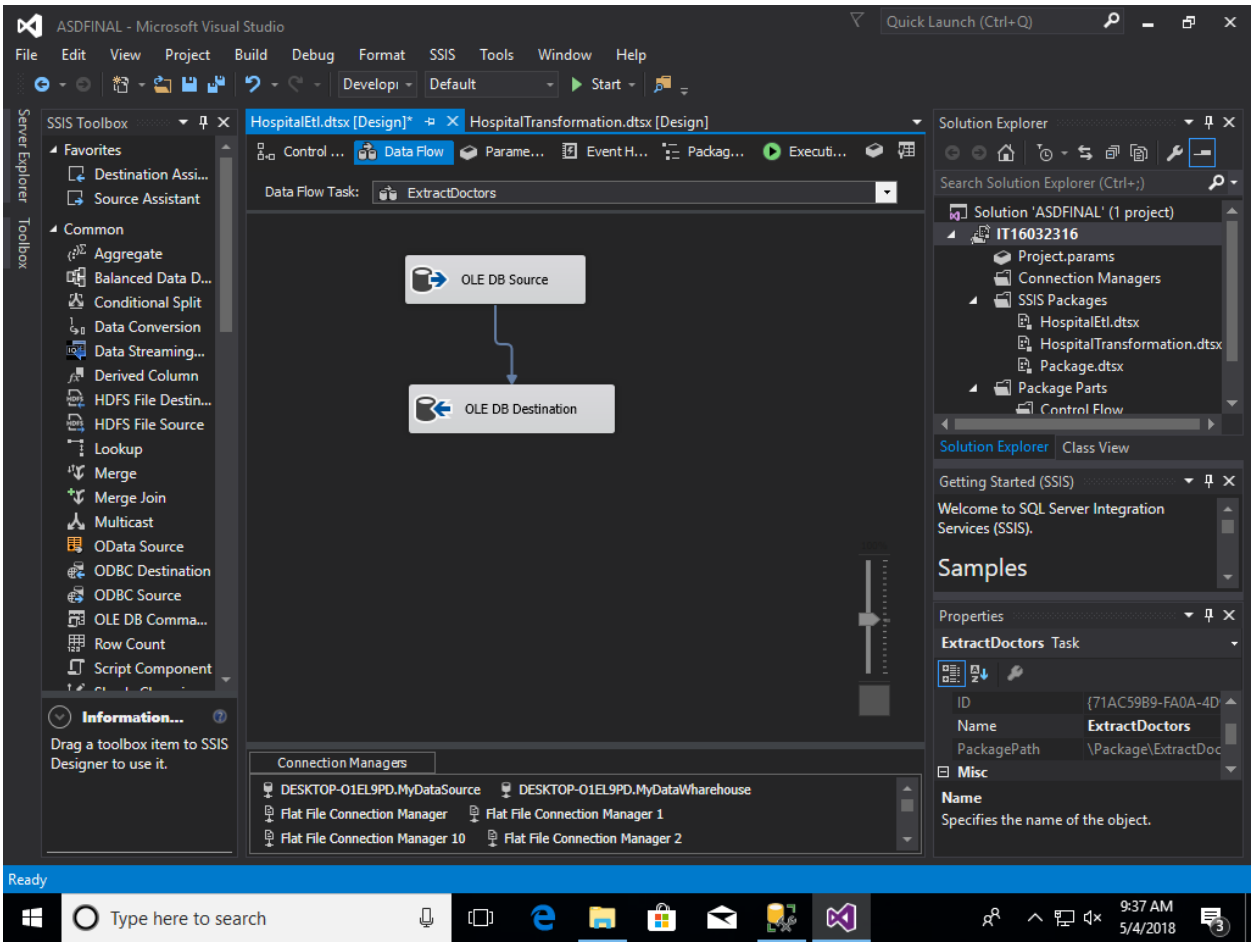
Patient Address =>PatiensAddress.xls

Records =>Records.txt

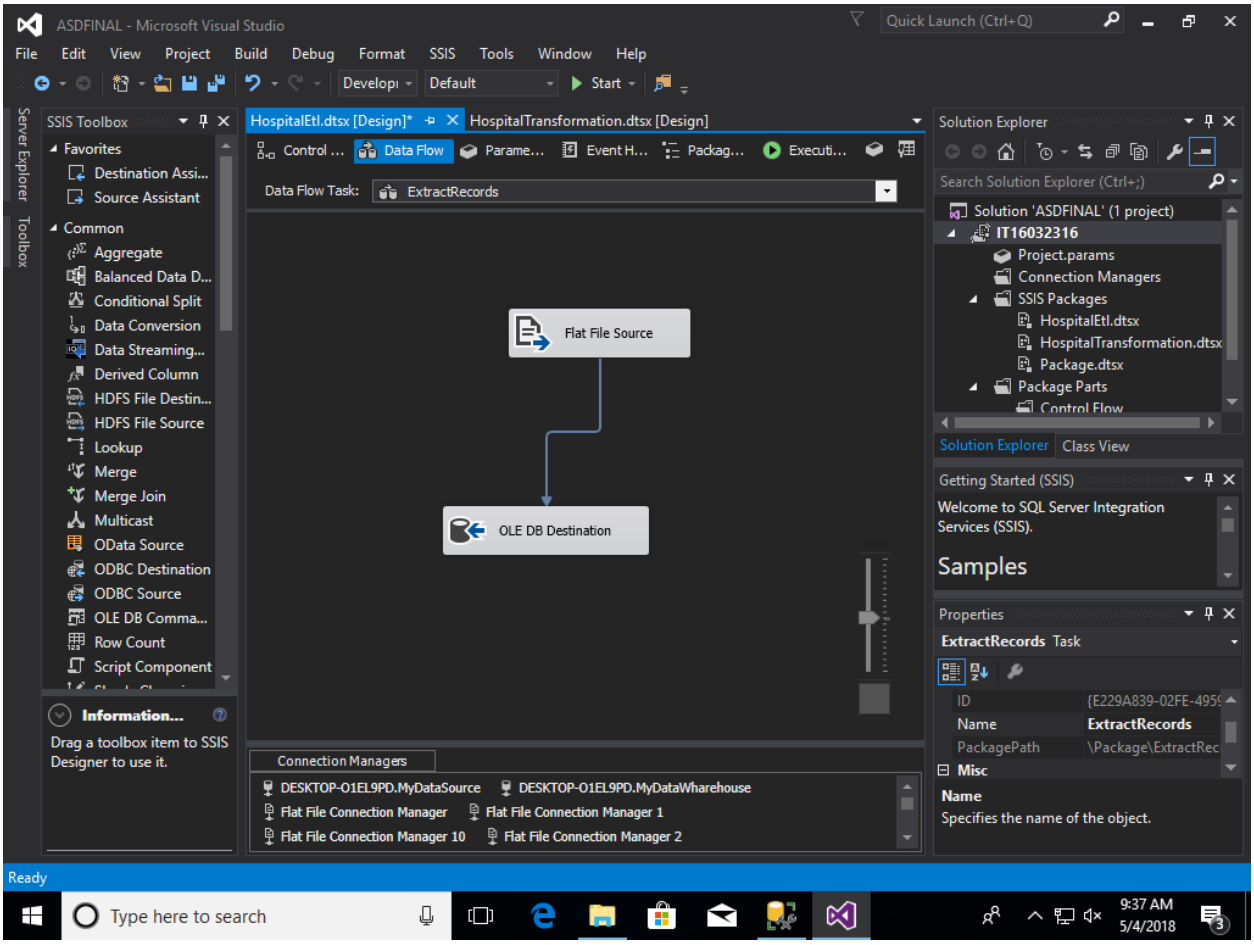
After this the xls Data was Loaded to the sql server by creating a database called DataSource.



Then by using the visual studio (ssis) All the data Loaded to the sql was took by Oldb Source and Put into Staging Area



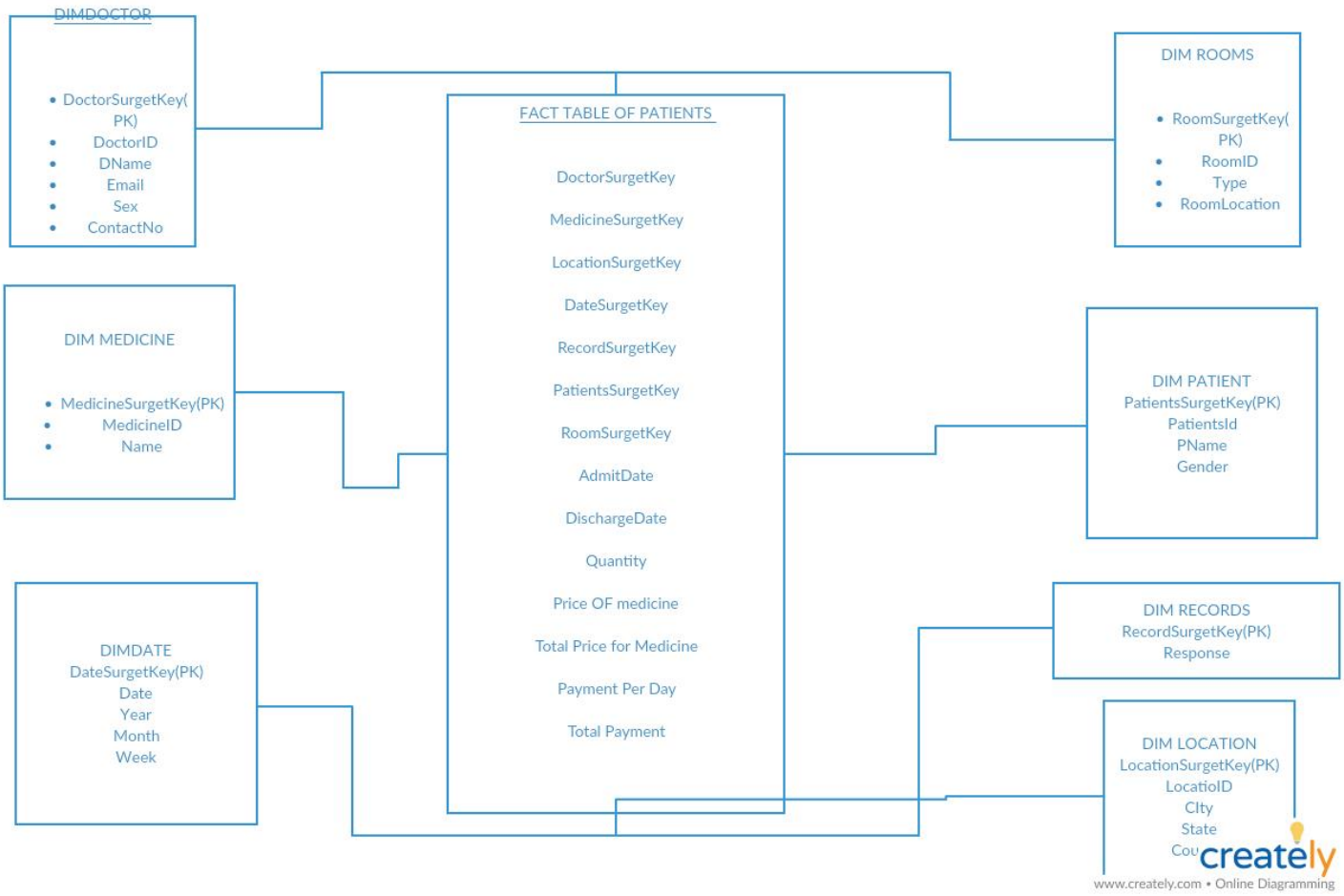
Then The Flat File Data was took directly from the records.txt file and get loaded to the Records Staging Area



After this All the Data from DataSource are went to the Staging are of “Datawharehouse” Database.

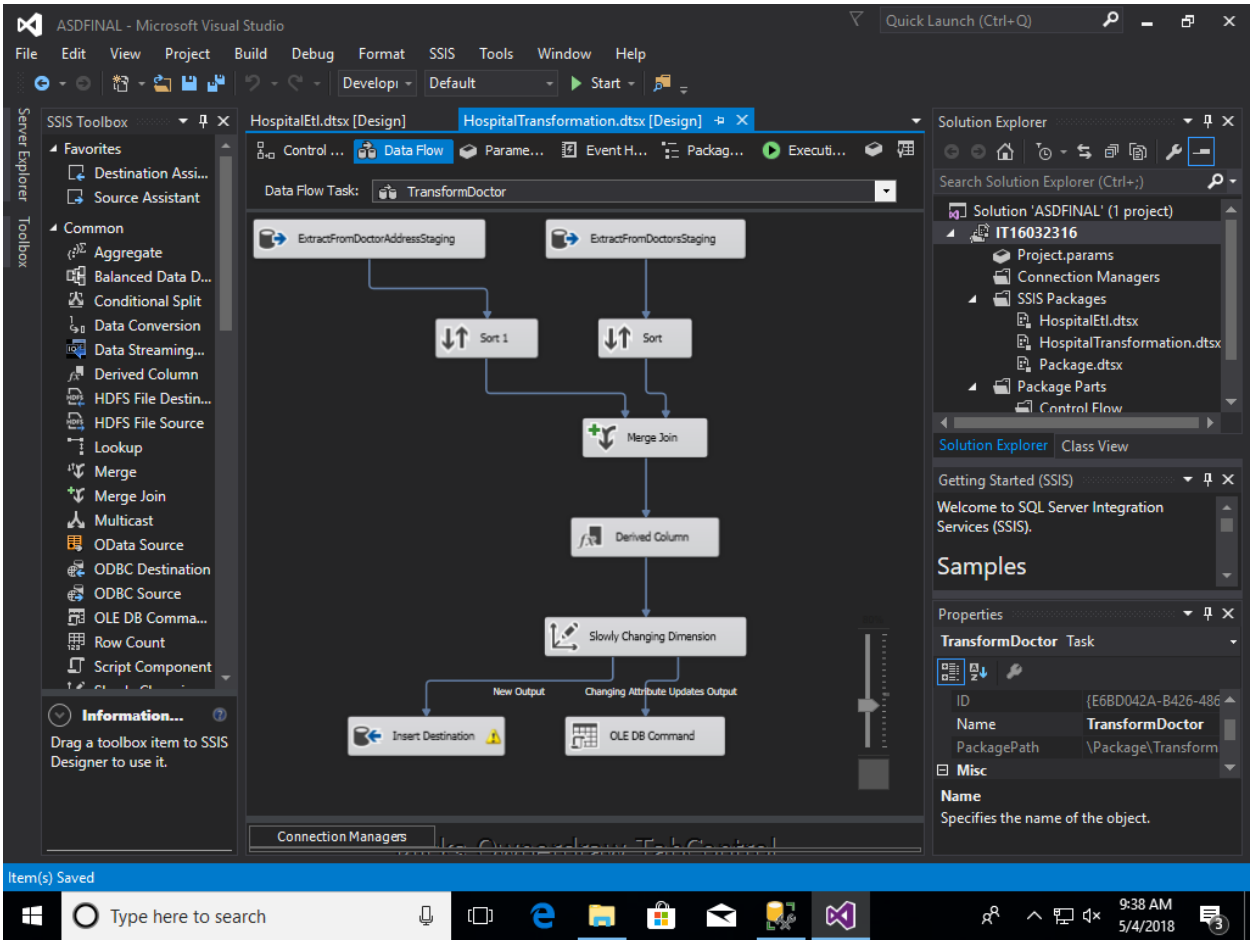
Then the next step was to Load the Data to dimension Table.

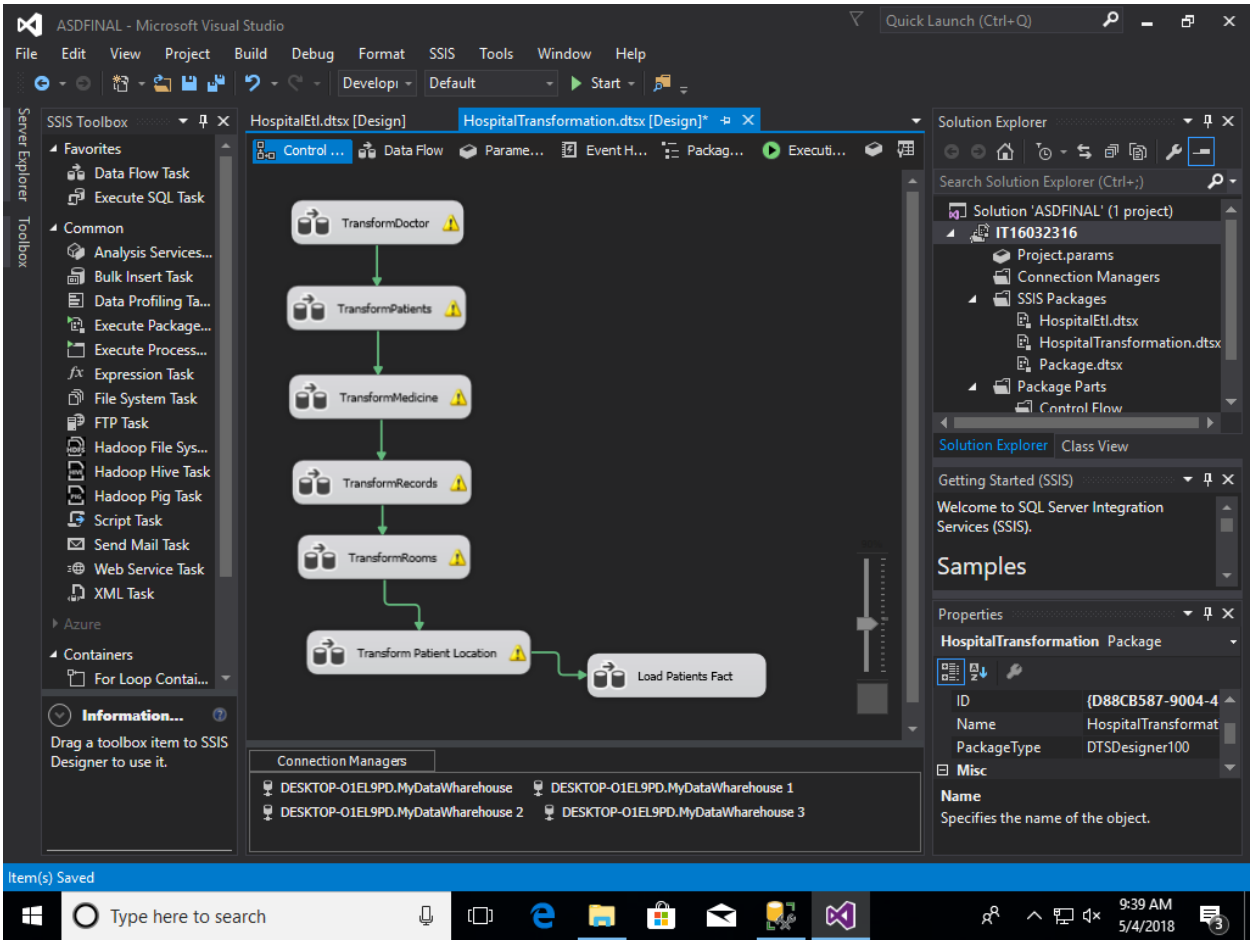
Given Below is the schema I used to map fact table and the dimension Table



So From The Staging Area the data was moved to the dimension table using slowly changing dimension task in the dataflow.

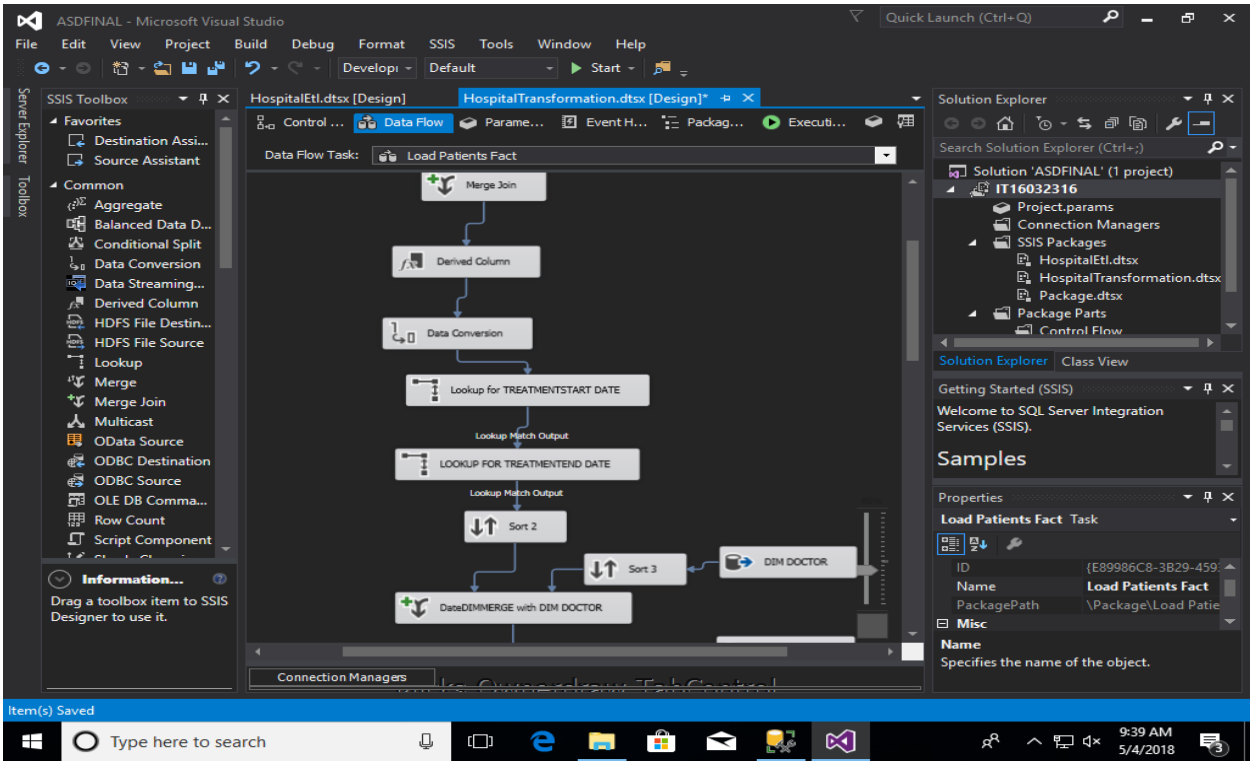
Given Below is an example the data took from DoctorsStagingArea and DoctorAddressStaging Area joined and loaded into DimDoctor.All the following procedure of Transforming data was done by the way which I have indicated now..

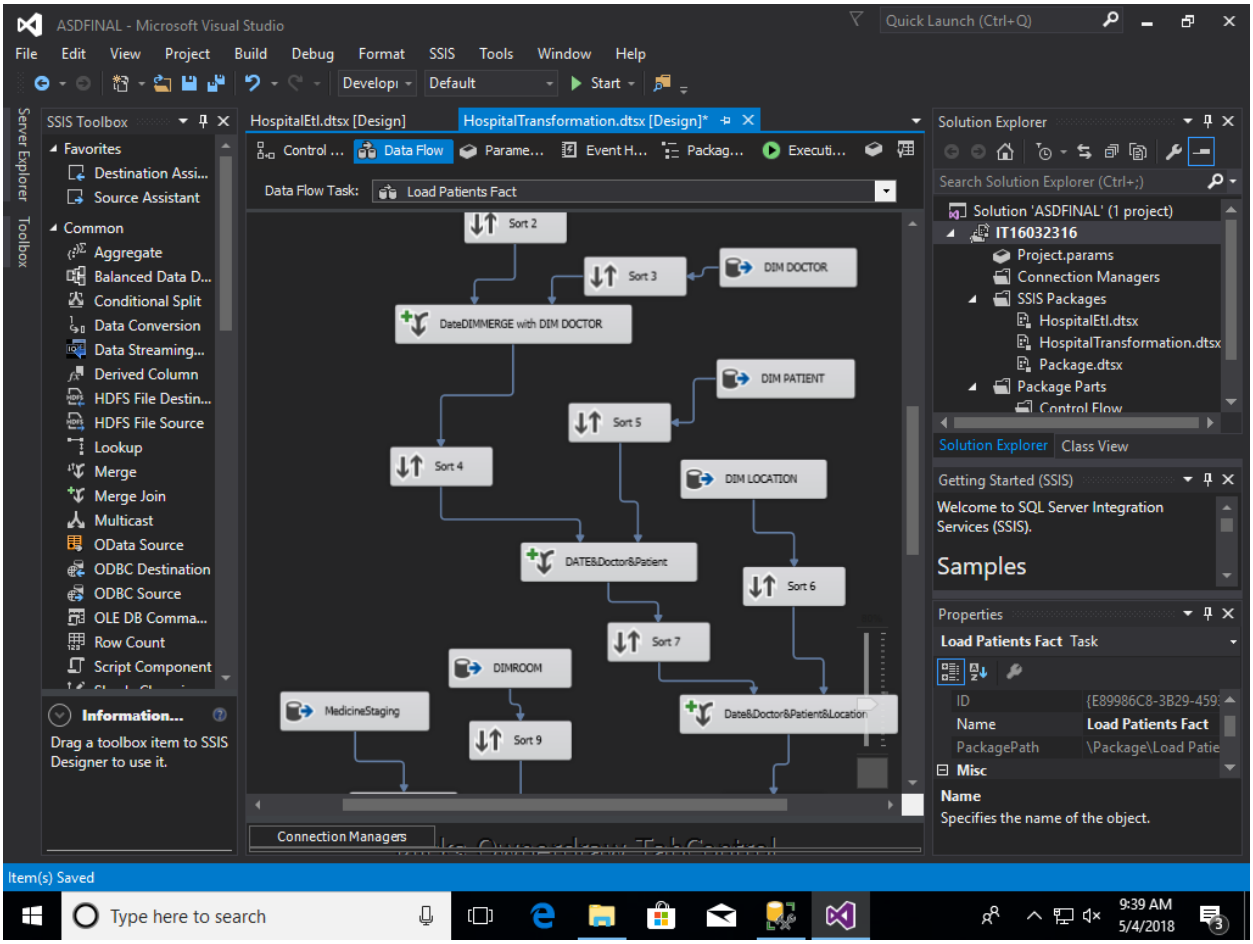



ASDFINAL - Microsoft Visual Studio
Quick Launch (Ctrl+Q)
File Edit View Project Build Debug Format SSIS Tools Window Help
Start
SSIS Toolbox
Favorites
Data Flow Task
Execute SQL Task
Common
Analysis Services...
Bulk Insert Task
Data Profiling Ta...
Execute Package...
Execute Process...
Expression Task
File System Task
FTP Task
Hadoop File Sys...
Hadoop Hive Task
Hadoop Pig Task
Script Task
Send Mail Task
Web Service Task
XML Task
Azure
Containers
For Loop Contai...
Information...
Drag a toolbox item to SSIS Designer to use it.
HospitalEtl.dtsx [Design]
HospitalTransformation.dtsx [Design]*
Control ...
Data Flow
Parame...
Event H...
Packag...
Executi...
TransformDoctor
TransformPatients
TransformMedicine
TransformRecords
TransformRooms
Transform Patient Location
Load Patients Fact
Connection Managers
DESKTOP-O1EL9PD.MyDataWharehouse
DESKTOP-O1EL9PD.MyDataWharehouse 1
DESKTOP-O1EL9PD.MyDataWharehouse 2
DESKTOP-O1EL9PD.MyDataWharehouse 3
Solution Explorer
Search Solution Explorer (Ctrl+;)
Solution 'ASDFINAL' (1 project)
IT16032316
Project.params
Connection Managers
SSIS Packages
HospitalEtl.dtsx
HospitalTransformation.dtsx
Package.dtsx
Package Parts
Control Flow
Solution Explorer
Class View
Getting Started (SSIS)
Welcome to SQL Server Integration Services (SSIS).
Samples
Properties
HospitalTransformation Package
ID
{D88CB587-9004-4
Name
HospitalTransformat
PackageType
DTSDesigner100
Misc
Name
Specifies the name of the object.
Item(s) Saved
Type here to search
9:39 AM
5/4/2018

After Loading All the dimension by this way finally Fact table was Created.

Given Below are the screenshots of creating the fact table





Here in order to get the date dimension I made the TreatentStartDate and TreatmentEndDate

In the derived column to convert it to get the format of yyyymmdd I used these formulas

(DT_WSTR,4)YEAR(birthday) + RIGHT("0" + (DT_WSTR,2)MONTH(birthday),2) + RIGHT("0" + (DT_WSTR,2)DAY(birthday),2)

SUBSTRING(TREATMENT_END_DATE,7,4) + SUBSTRING(TREATMENT_END_DATE,1,2) + SUBSTRING(TREATMENT_END_DATE,4,2)