**Lab – ETL Testing**

December , 2015

This document is confidential and contains proprietary information, including trade secrets of CitiusTech. Neither the document nor any of the information contained in it may be reproduced or disclosed to any unauthorized person under any circumstances without the express written permission of CitiusTech.

**Document Control**

|  |  |
| --- | --- |
| Version | 1 |
| Created by | Dimple Manwani |
| Updated by |  |
| Reviewed by | Sugandha Agarwal |
| Date | 12th Oct 2015 |
| Earlier versions | NA |

**Table of Contents**

[Prerequisites 3](#_Toc439252909)

[Getting Started 3](#_Toc439252910)

[Data Model 4](#_Toc439252911)

[Restore Test DB – BIDataMart 5](#_Toc439252912)

[Test Design 8](#_Toc439252913)

[Patient Registration 8](#_Toc439252914)

[Source to Target Mapping 8](#_Toc439252915)

[Test Execution 9](#_Toc439252916)

[Job Creation & Deployment 9](#_Toc439252917)

[Conclusion: 10](#_Toc439252918)

[TO DO: 11](#_Toc439252919)

[Clinics Information 11](#_Toc439252920)

[Patient Visit Information 11](#_Toc439252921)

[Patient Claims Information 12](#_Toc439252922)

[Hospital Staff Details 12](#_Toc439252923)

# Prerequisites

Participants undergoing this lab should be well-versed with the following,

* SQL Concepts
* Datawarehousing Concepts

# Getting Started

* **Overview**

This lab book describes how to perform testing of different ETL Processes.

* **Software Requirement**

Should have access to Microsoft SQL Server.

* **Business Requirement**

LifeCare BI Portal has reports tab which displays various reports based on authorization rights. The data which is made available for reporting purposes is extracted, transformed and loaded into the datawarehouse as per business needs.

This document covers detailed description for each of the data load or ETL process to integrate data from various clinics/hospitals associated with Lifecare Group.

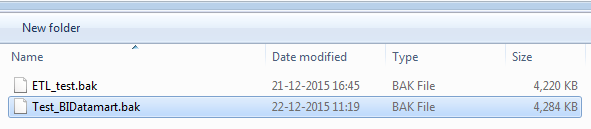
# Data Model

Given below is the data model of ‘BIDataMart’ for your reference,

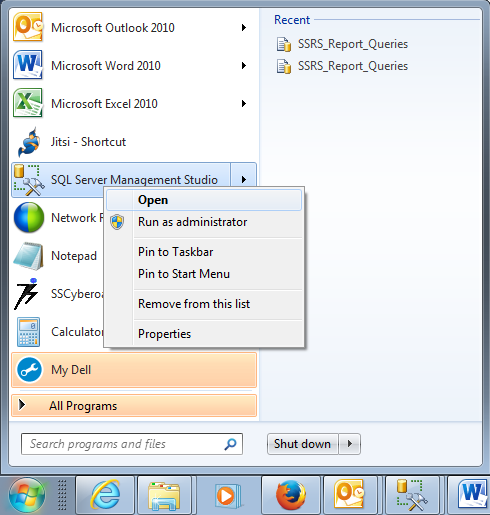


# Restore Test DB – BIDataMart

Copy the backup file Test\_BIDatamart.bak on your local machine.



Open SQL Server Management Studio from Start Menu.



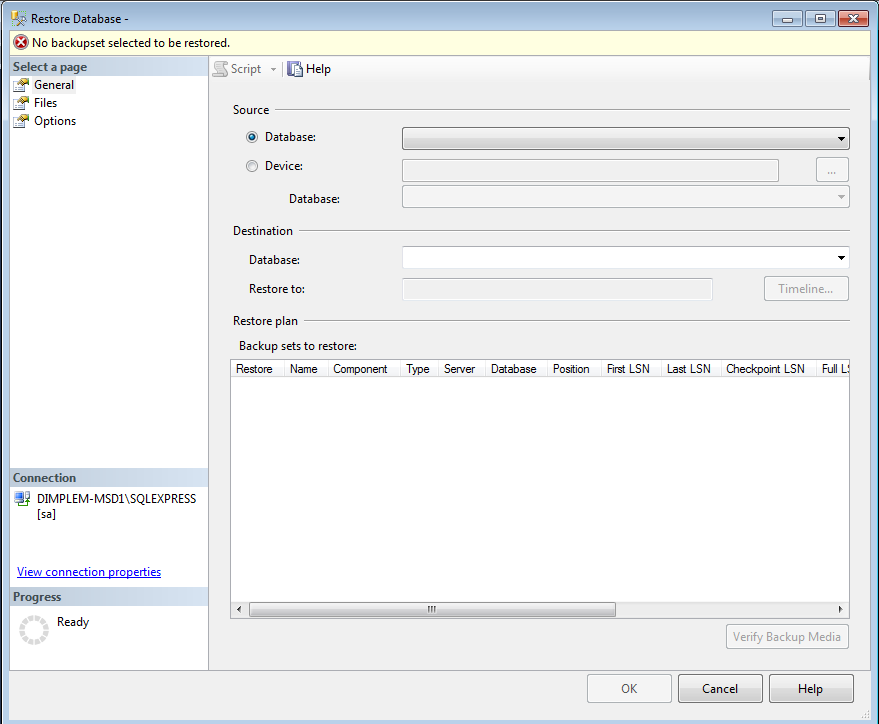
Connect to the DB Engine, using SQL Server Authentication

User- sa

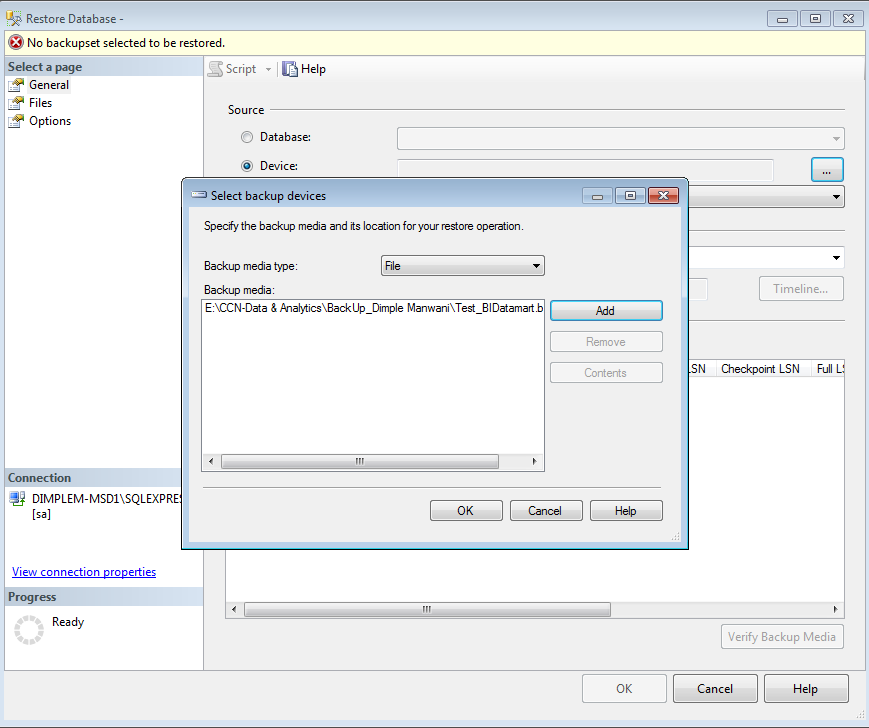
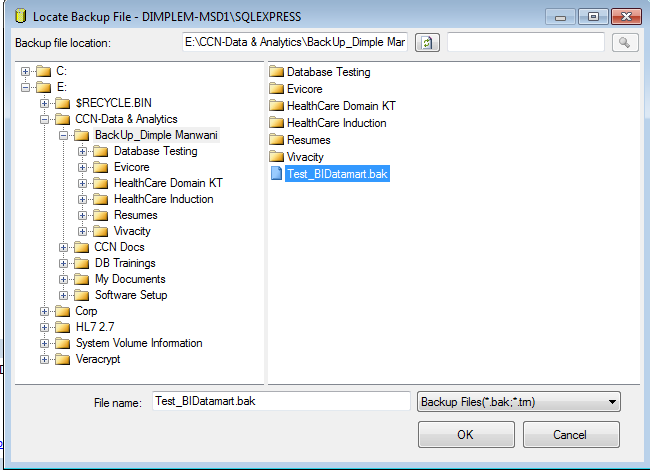
Password- password\_123



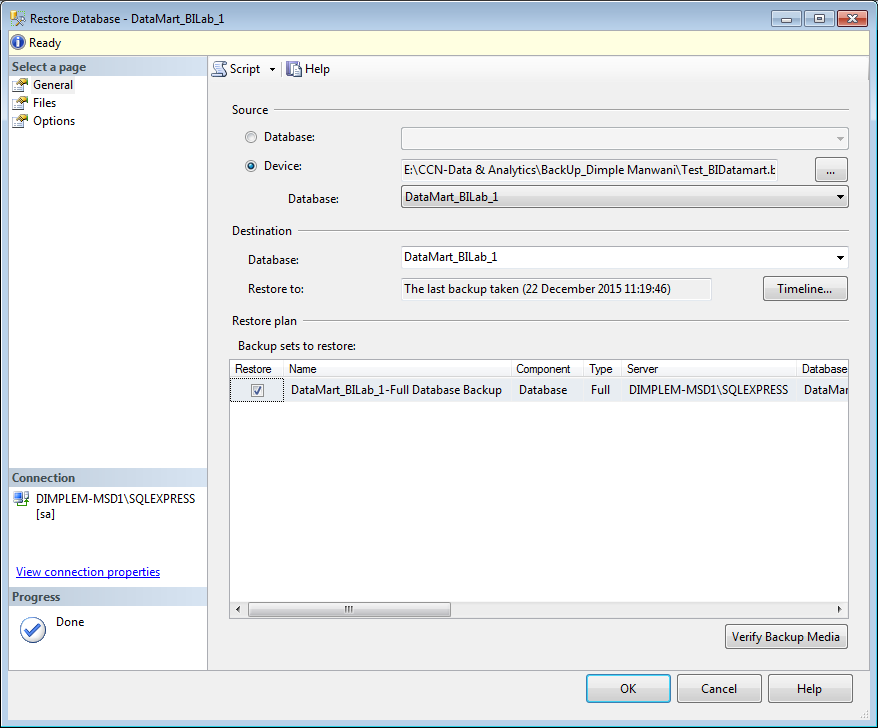
Right-Click on the Databases folder -> Restore Database



Select Device -> Browse & Add the backup file ‘Test\_BIDatamart.bak’

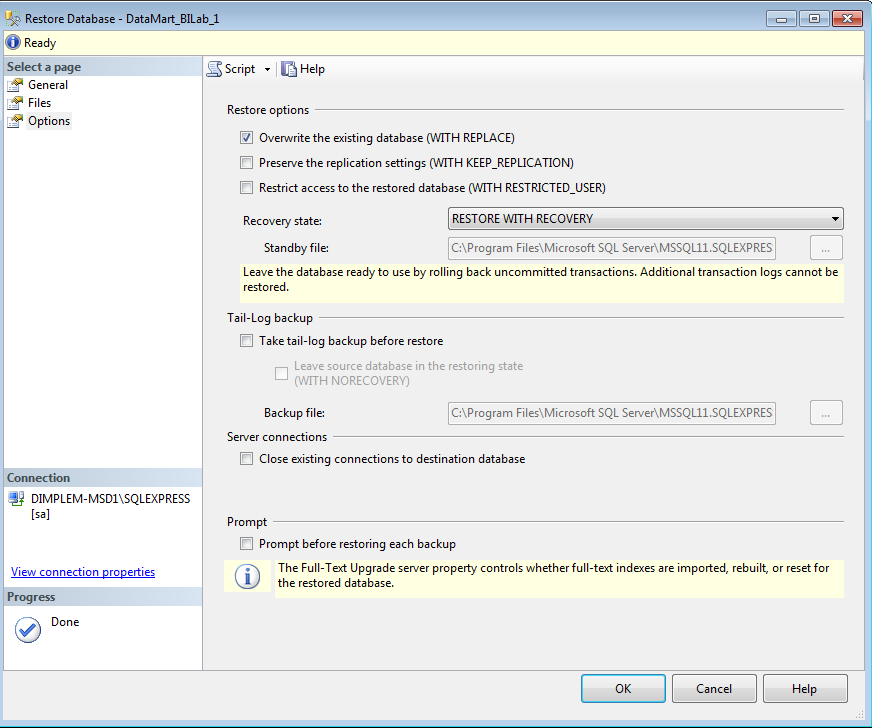
Verify the backup file name appears under Restore plan tab.



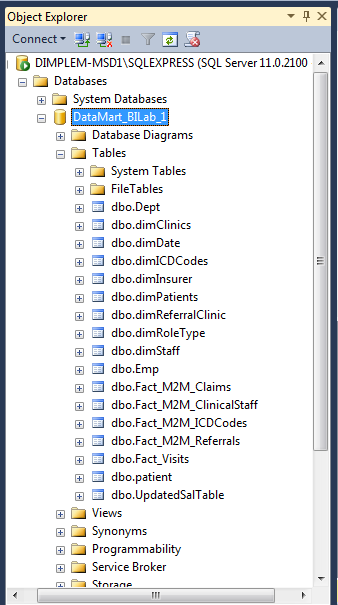
Go to the Options tab,

Check **‘Overwrite the existing database’**

Uncheck **‘Take tail-log backup before restore’**. Click OK.



Check Test DB is now restored-



# Test Design

Now, lets understand each of the ETL data flow requirements mentioned to design test scenarios & test cases for them,

## Patient Registration

Patient registration information is captured in two ways:

* via web portal as the patient is registered by the hospital staff on LifeCare Web Portal on their first visit
* via IVR call by the Patient at the hospital/clinic

#### Patient personal information will consist of-

[PatientID]

[PatientName]

[ClinicID]

[State]

[City]

[ZipCode]

[ContactNo]

[Gender]

[BirthDate]

[NextOfKin]

Once the new patient is registered, the information will be transferred to the data warehouse. Over a period of time a registered patient can have multiple visit entries which will be available in the data warehouse.

Changes in the patient information are maintained as per SCD Type-2 dimension. Patient Data is loaded on daily basis through a data loading job process via a comma separated file(.csv).

Refer detailed source to target mapping attached below. Refer attachment for test cases for the data load of Patient Registration.



## Source to Target Mapping

Column level mapping for each data loading process is specified in below attached excel.

****

# Test Execution

Let us execute the test cases for Patient Registration and capture the test evidence.

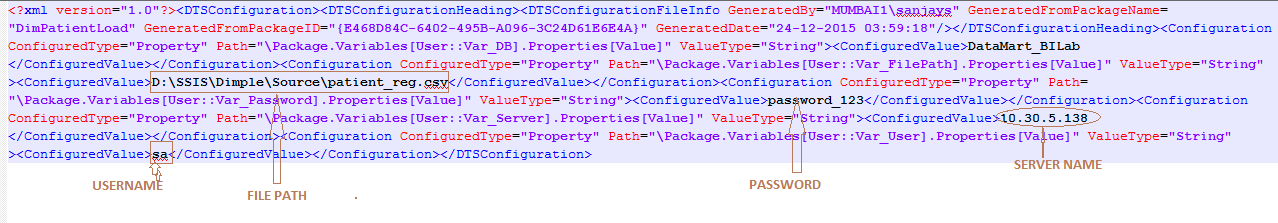
Note that we are using SSIS as an ETL tool and SQL Server as a DWH environment for this training purpose.

## Job Creation & Deployment

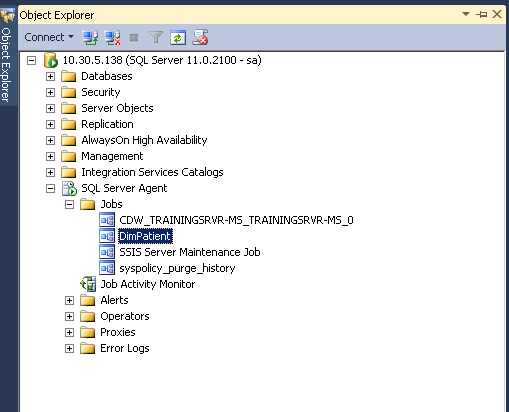
Follow the steps given below to deploy the SSIS package in your environment for test execution,

1. Copy the folder “DTSConfig” from shared path on your local machine C:\ drive.
2. Config file is located @C:\DTSConfig\config4.dtsConfig. Config file location should be static.
3. Config file (config4.dtsConfig) components that needs to be altered for deployment are listed below-

* File Path(Path of the patient\_reg.csv file)
* Server(Database Instance Name)
* Userid( Database Username mostly “sa”)
* Password



1. Alter the location of the SSIS package “DimPatientLoad.dtsx” in the job script, “Job Creation Script.sql” and run the script under SQL Server Management Studio for SQL job creation.



1. Right-Click on the job name “DimPatient” and Click on ‘Start Job at Step…’ to run job and load the required data.
2. Execute the ETL test cases as mentioned in the document above.

## Conclusion:

ETL Testing includes-

* Verify that data is transformed correctly according to various business requirements and rules.
* Make sure that all projected data is loaded into the data warehouse without any data loss and truncation.
* Make sure that ETL application appropriately rejects, replaces with default values and reports invalid data.
* Make sure that data is loaded in data warehouse within prescribed and expected time frames to confirm improved performance and scalability.

# TO DO:

Create & Execute TestScenarios and TestCases for each of the below requirements,

## Clinics Information

It stores the data of all the clinics which are associated with the LifeCare Group of Healthcare.

The attributes which needs to be stored in the warehouse schema are-

[ClinicID]

[ClinicName]

[Location]

Clinic information is a one-time load into the warehouse. Whenever a new clinic is added with the enterprise, the data is transferred to the data mart.

## Patient Visit Information

Each of the Clinics at LifeCare Group maintains Patient visit data which is required to be transferred to the data warehouse on daily basis. The Patient visit details are utilized for various reporting needs at LifeCare Group.

Each of the clinics provide following data which is stored in the data warehouse-

[VisitID]

[PatientID]

[ClinicID]

[ArrivedTime]

[PhyCheckinTime]

[PhycheckOutTime]

[IsCritical]

[OtherTranscriptions]

[Satisfied]

[VisitCharges]

[Status]

[AppointedTo]

[VisitDate]

[WaitTimeInMinutes]

Patient visit data for each clinic is loaded through a daily data loading job process.

## Patient Claims Information

A claim is generated whenever a patient/insurer undergoes medical expenses and it is submitted by the Clinic for re-imbursement

Each of the clinics provides following data which is stored in the data warehouse for claims made so far,

[VisitID]

[InsurerID]

[ClinicName]

[ClaimDate]

[ChargesPaidDate]

[TotalCharges]

[ChargesDue]

[ChargesPaid]

Patient Claims data for each Clinic is loaded through a daily data loading job process.

## Hospital Staff Details

Each of the Clinics at LifeCare Group maintains staff information which needs to be transferred to the data warehouse. Each of the clinics provide following data,

[MemberID]

[MemberName]

[RoleName]

[Gender]

[ZipCode]

[ContactNo]

[City]

[Region]

[BirthDate]

Hospital Data is loaded into warehouse by running monthly batch load process.