A blue and orange logo

AI-generated content may be incorrect.

**Sri Lanka Institute of Information Technology**

DATA WAREHOUSING AND BUSINESS INTELLIGNECE(IT3021)

Year 03 Semester 01 – 2025

**ASSIGNMENT 1**

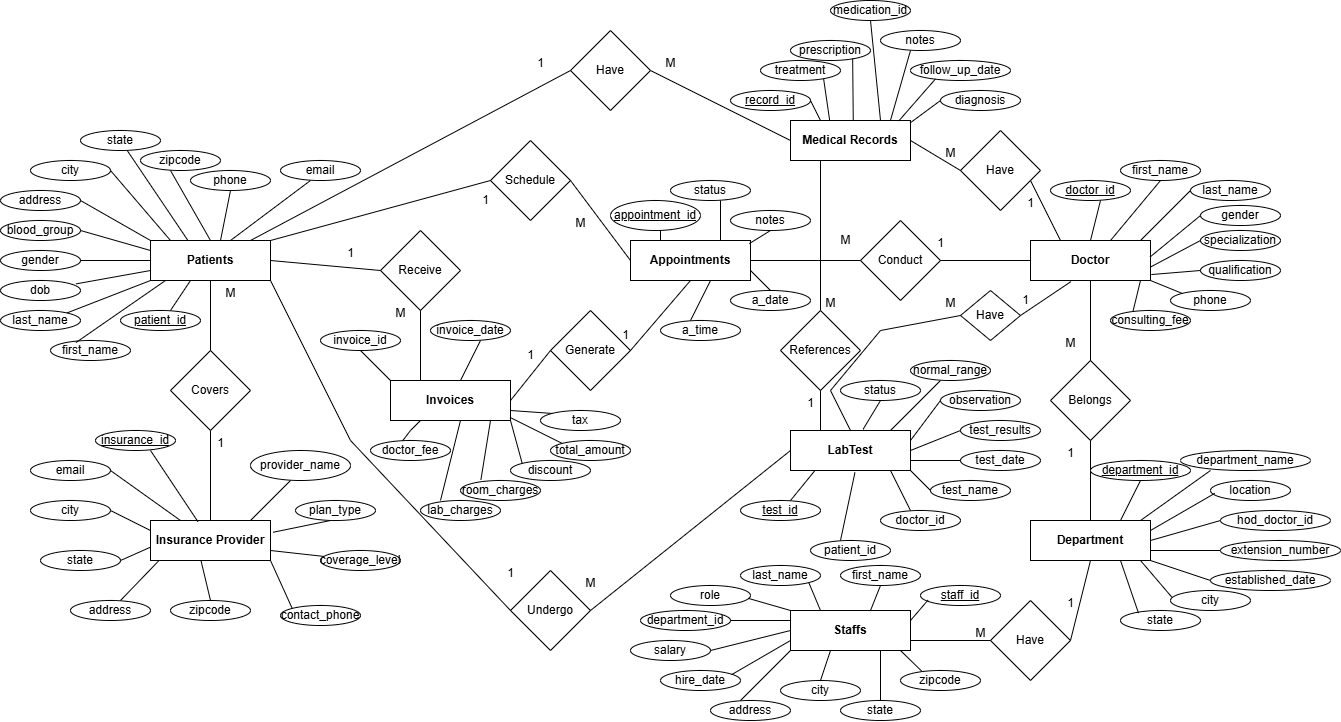
**BY M.Y IQRA [IT22071620]**

Step 1

* 1. Description of the DataSet

For this project, I designed and developed a simulated Healthcare operational system. Multiple files and database tables were prepared to handle healthcare transactions. The system included data for patient management, doctors, appointments, medications, lab tests, departments, and insurance providers.

1.2 ER DIAGRAM



STEP 2

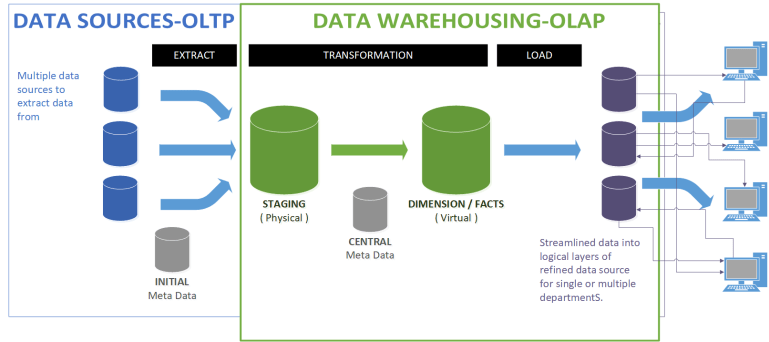
2.1 Description of the DataSources

The original raw data was available in two formats:

* CSV [Comma-seperated Values] files
* Patients.csv
* Doctors.csv
* Appointments.csv
* Medications.csv
* InsuranceProviders.csv
* LabTests.csv
* Medical.Records.csv
* Invoices.csv
* SQLdatabase tables [For Direct Extraction)
* TxnUpdate

STEP 3

3.1 High Level DW & BI Solution architecture



3.2 SourceDB [HealthCare]

* dbo.Appointment
* dbo.Department
* dbo.Doctors
* dbo.InsuranceProviders
* dbo.Invoice
* dbo.LabTest
* dbo.Medical
* dbo.Patients
* dbo.Staffs

StageDB

* dbo.StgAppointmnet
* dbo.StgDepartment
* dbo.StgDoctors
* dbo.StgInsuranceProviders
* dbo.StgInvoice
* dbo.StgMedical
* dbo.StgMedication
* dbo.StgPatients
* dbo.StgStaffs

DataWareHouseDB[HospitalDW]

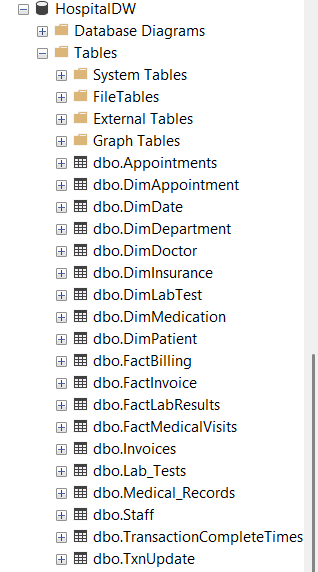
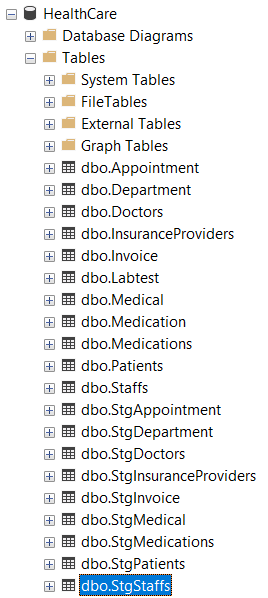
* dbo.DimAppoinmnet
* dbo.DimDate
* dbo.DimDepartment
* dbo.DimDoctor
* dbo.DimInsurance
* dbo.DimLabTest
* dbo.DimMedication
* dbo.DimPatient
* dbo.FactMedicalVisits
* dbo.FactInvoice
* dbo.TxmUpdate

ETL Process

Extract –> Read data from csv files and SQL tables

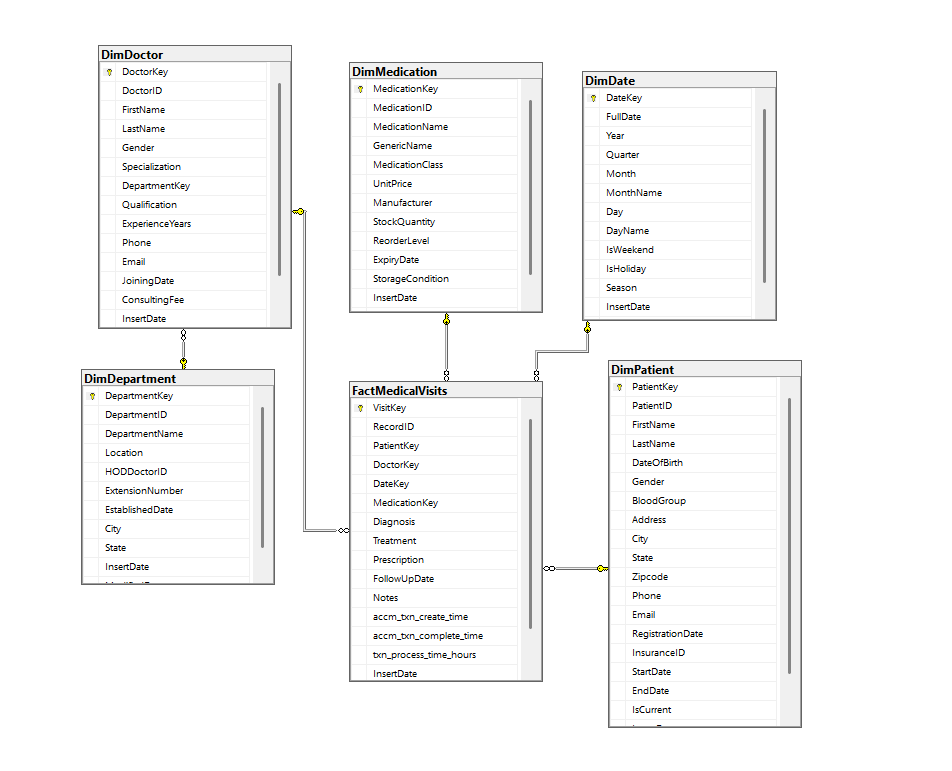
Transform -> Cleanse,formate and enrich the data (eg , used data conversion ,Dimension key,lookups and Derived columns),

Load -> Insert processed data into the staging Area and from there into Fact and Dimension tables of the DW

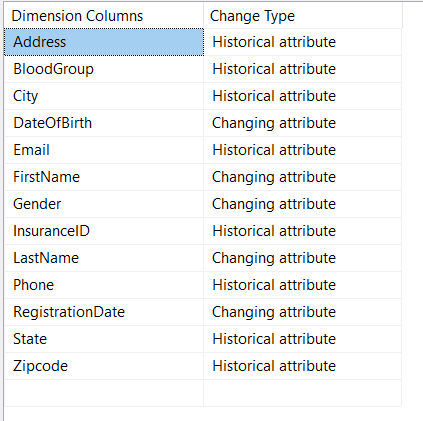


STEP 4

4.1 Relational Diagram StarSchema

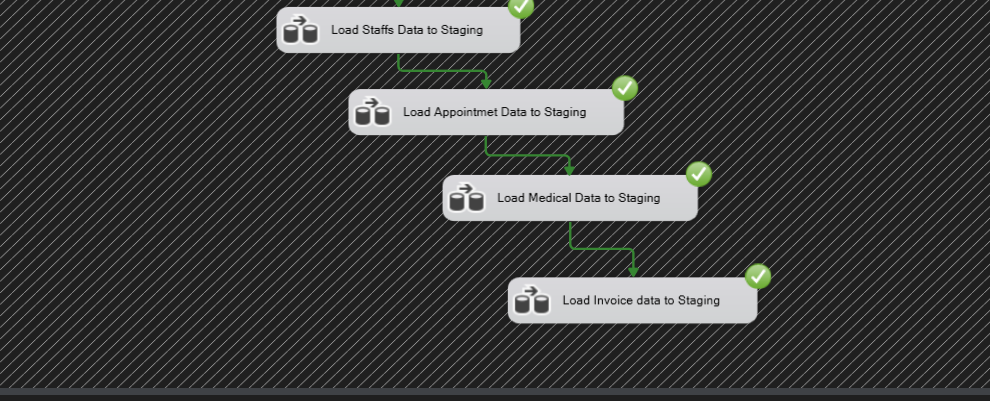
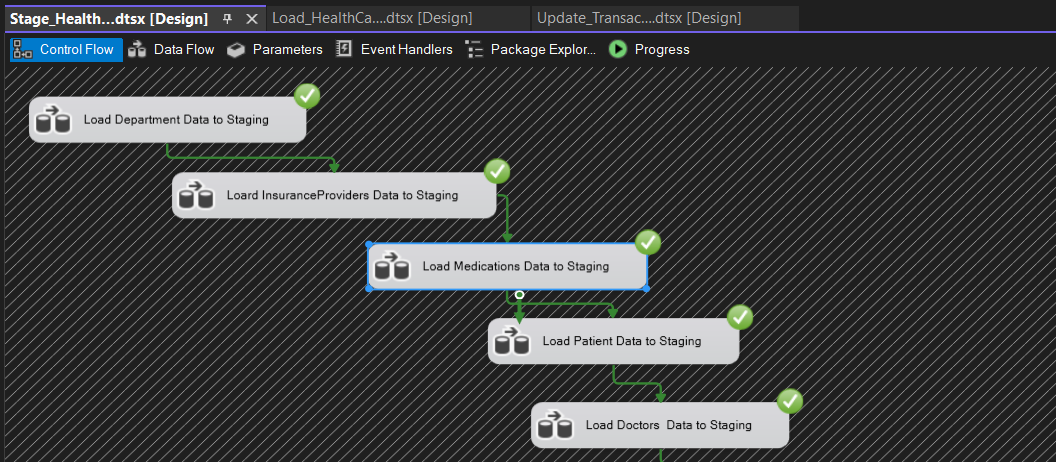


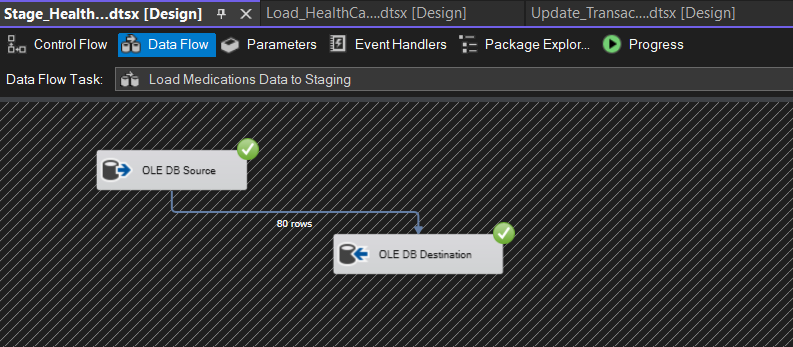
DimPatient was Implemented as a Slowly Changing Dimension.This table is designed to track changes in key patient attributes . So when a new record is inserted with updated values while the previous record is marked as inactive .This ensures that no data loss across the time

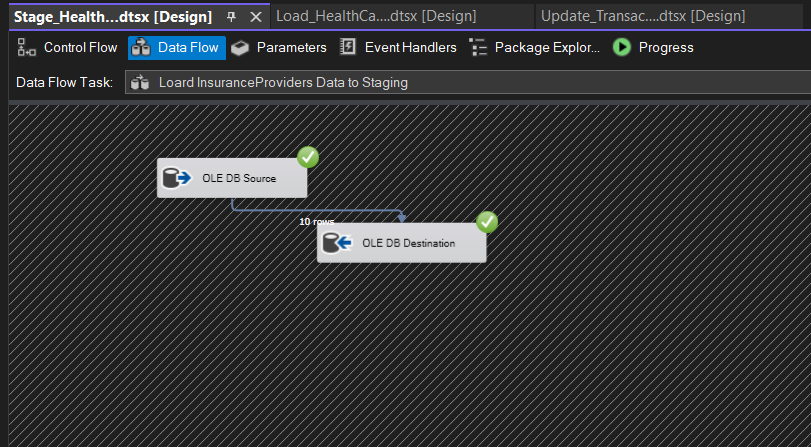


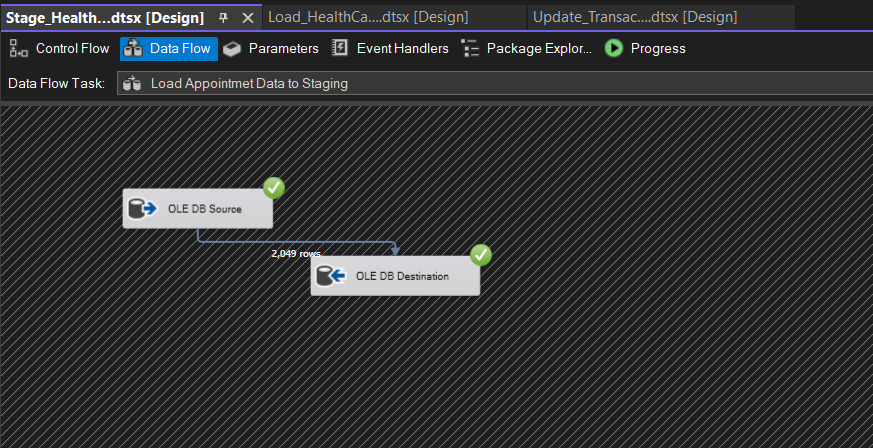
Step 5

5.1 Load the Data Source to Staging Area







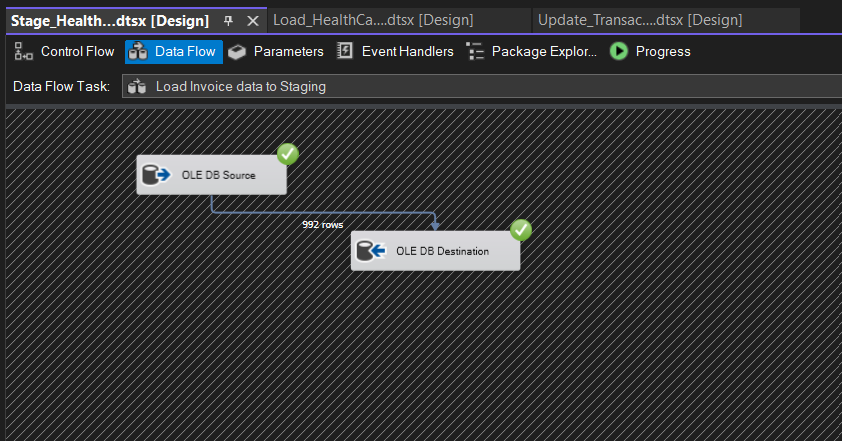


A screenshot of a computer

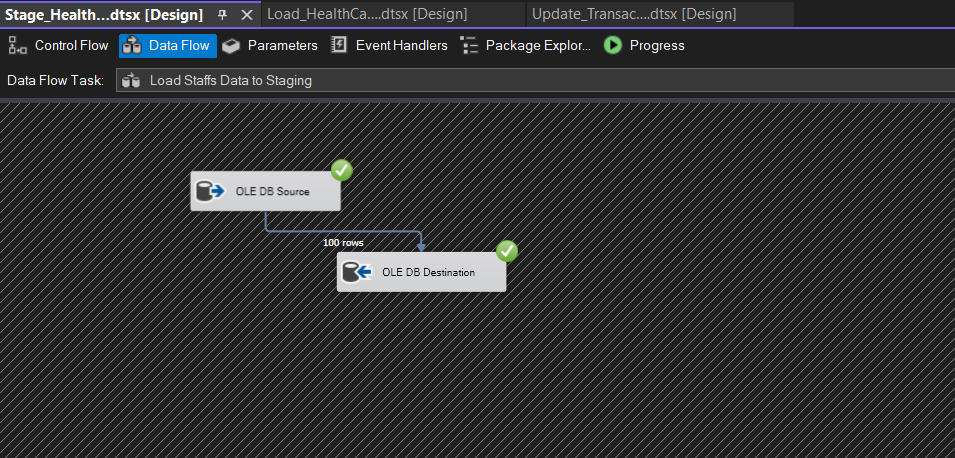
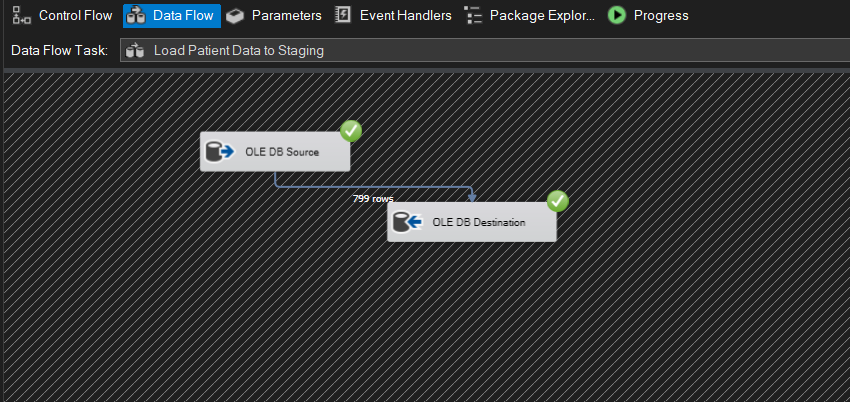
AI-generated content may be incorrect.

A screenshot of a computer

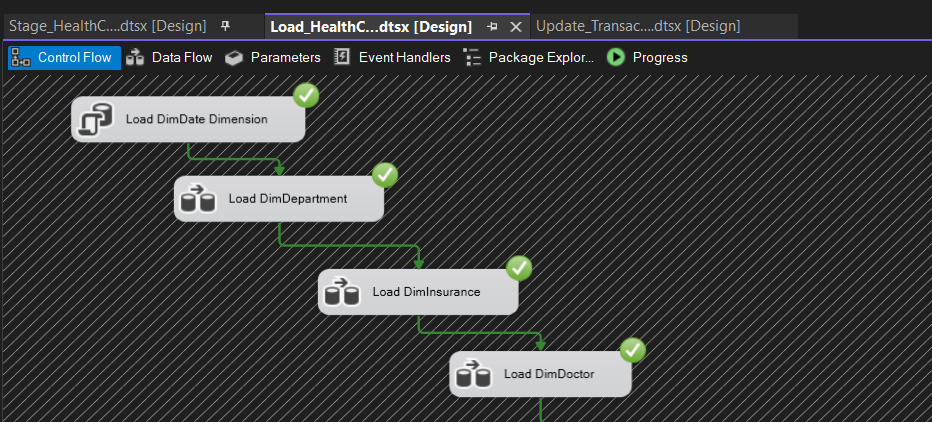
AI-generated content may be incorrect.

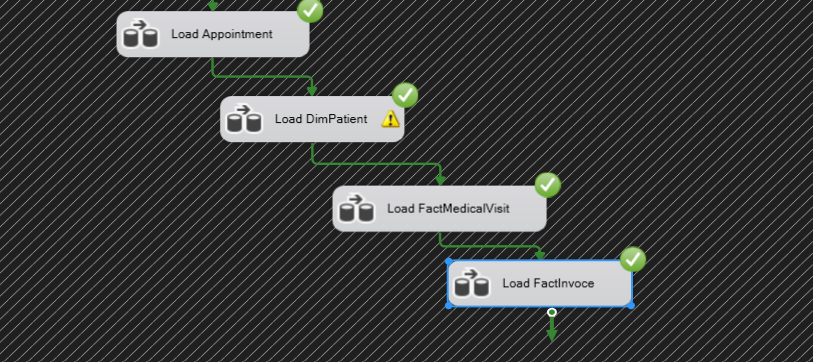


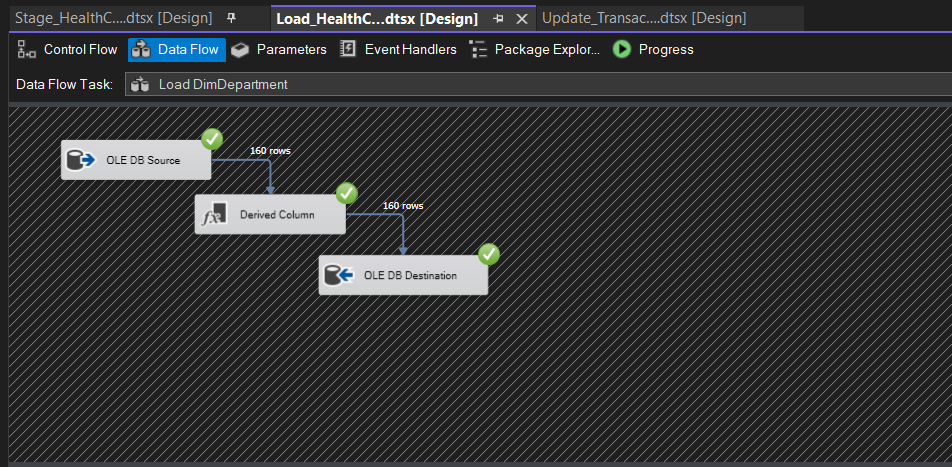
A screenshot of a computer

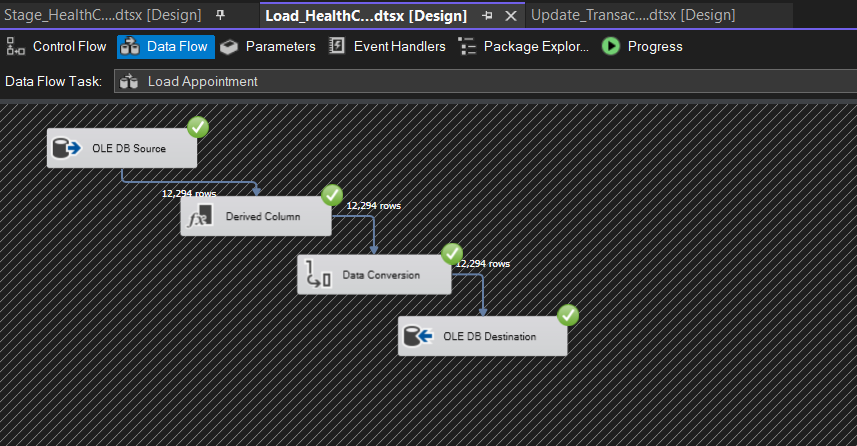
AI-generated content may be incorrect.

5.2 Load Staging Area to DataWarehouse

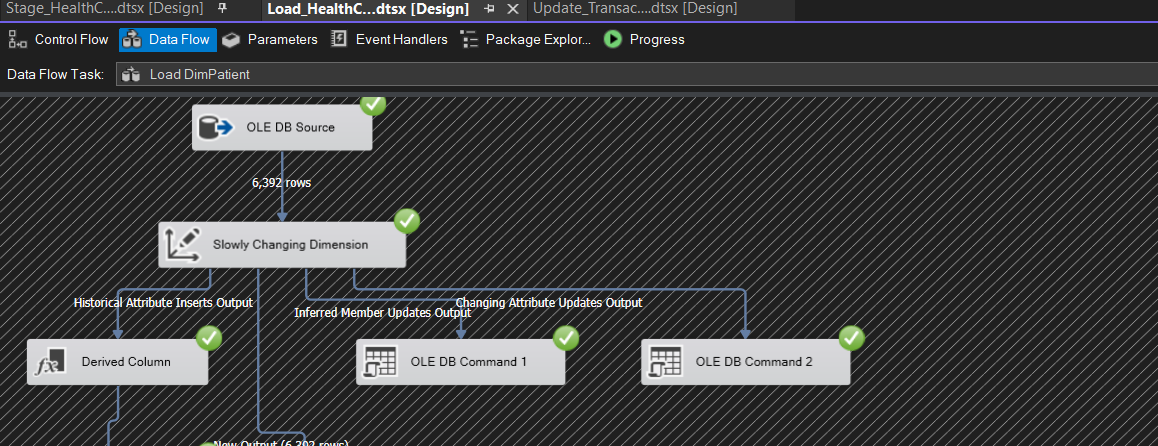






A screenshot of a computer

AI-generated content may be incorrect.A screenshot of a computer

AI-generated content may be incorrect.A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer program

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.A screenshot of a computer

AI-generated content may be incorrect.A screenshot of a computer

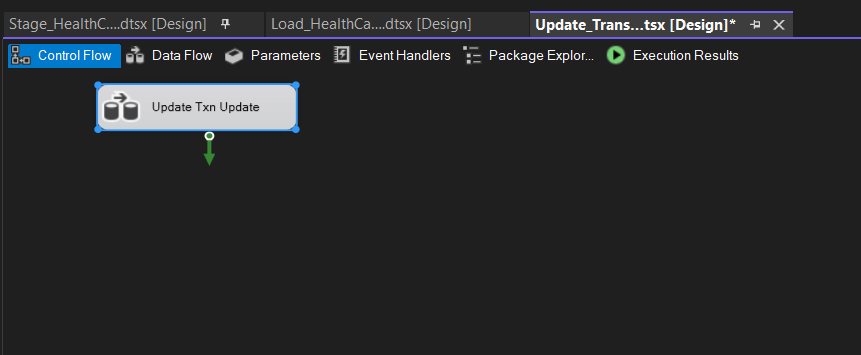
AI-generated content may be incorrect.A screenshot of a computer

AI-generated content may be incorrect.A screenshot of a computer

AI-generated content may be incorrect.

Step 6

Accumalting Fact Tables



A diagram of a data flow

AI-generated content may be incorrect.

