

# **Sri Lanka Institute of Information Technology**



## **Data Warehousing & Business Intelligence**

### **Assignment 01**

**Submitted by:**

**Rathnasooriya P.U (IT20156206)**

## Contents

1.Data Set Selection.....	3
01.    Background: - .....	3
02.    Content: - .....	3
03.    ER Diagram.....	4
2.    Preparation of Data Sources.....	5
3.    Solution Architecture.....	6
4.    Data warehouse design & development.....	7
5.    ETL Development .....	8
5.1 ETL –Source to Staging.....	8
5.1.1 Load Data Department Staging Table .....	8
5.1.2 Load Data Doctor Staging Table.....	9
5.1.3 Load Data Location Staging Table .....	9
5.1.4 Load Data Hospital Staging Table.....	10
5.1.5 Load Data patient Staging Table .....	10
5.1.6 Load Data Branch Staging Table .....	11
6.    Staging To Data Warehouse.....	12
6.1 Load Department Data.....	12
6.2 Load Doctor Data .....	13
6.3 Load Patient Data.....	13
6.4 Load Location Data.....	14
6.5 Load FactHospital Data .....	15
7.    Data Warehouse Update SISS Package .....	16
7.1 Load accm_txn_complete_time Data .....	16
7.2 Load_txn_process_time_houres Data .....	17
8.    Accumulated Fact Table.....	18

# 1.Data Set Selection.

## 01.Background: -

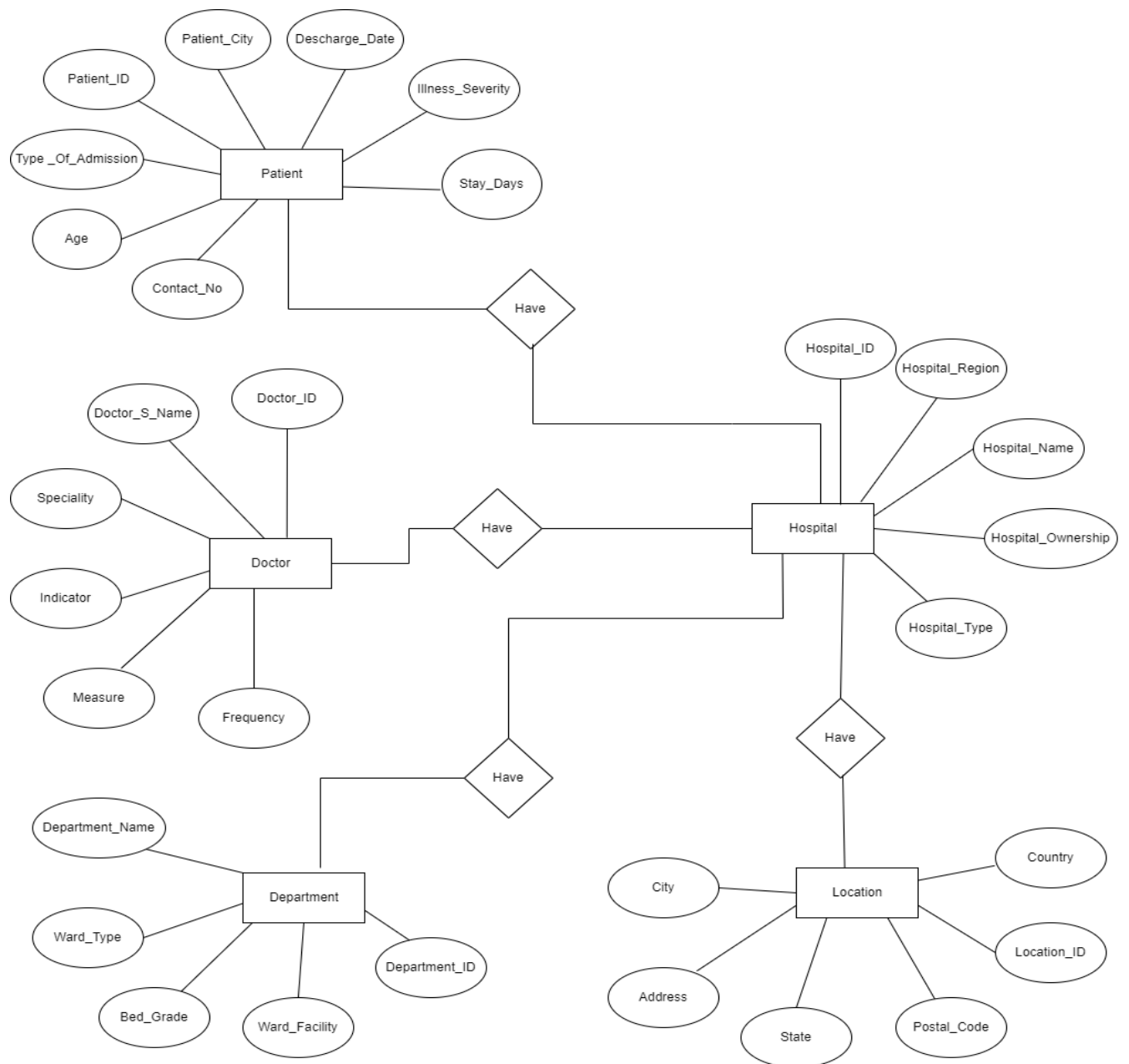
With rapidly growing demand for healthcare in hospitals and intensive care units (ICUs) throughout the world, the COVID-19 epidemic has put unprecedented strain on health systems. As the epidemic spreads, evaluating the related healthcare resource requirements (beds, personnel, and equipment) has become a top issue for many countries. Estimates of how long patients with COVID-19 require various levels of hospital care are needed to forecast future demand.

## 02.Content: -

The data set was downloaded from -:

[COVID-19 Hospitals Treatment Plan | Kaggle](#)

### 03. ER Diagram



## 2. Preparation of Data Sources.

In order to data extraction need to prepare the data sources. From my main data source, I have extracted to type of data sources.

1. Text file (.txt)

2. Excel files (.xlsx)

- Hospital Excel Data File.
- Department Excel Data File.
- Doctor Excel Data File.
- Customer Excel Data File.
- Patient Excel File.
- Accm\_txn\_complete\_time Excel File.

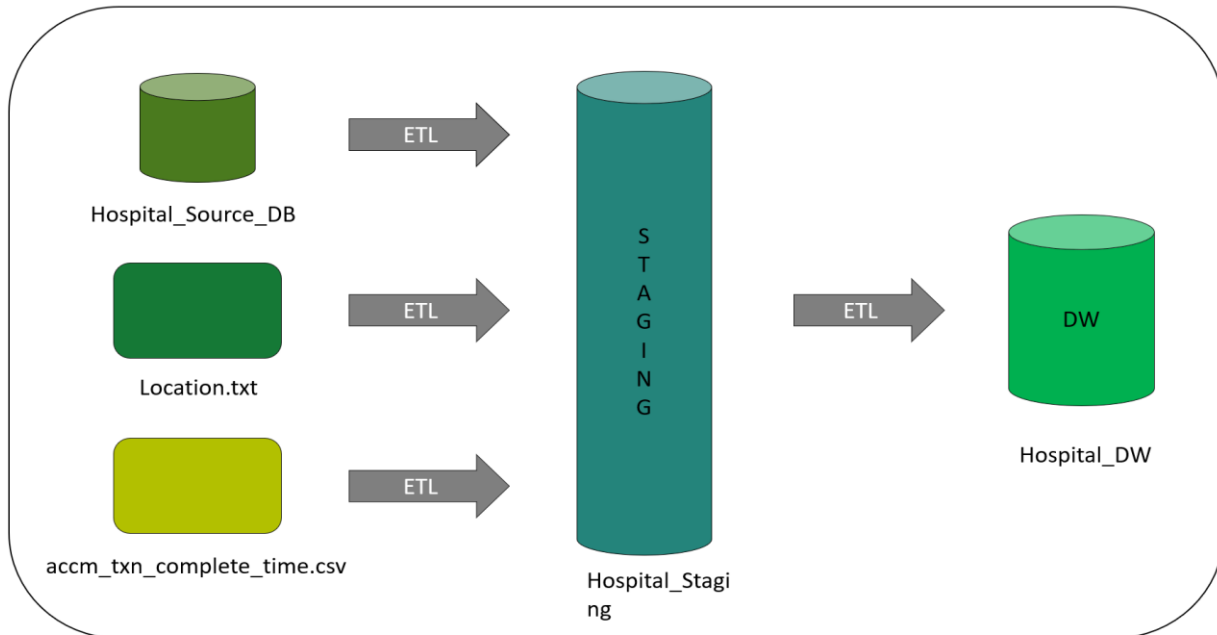
From these dataset, I import above excel files to SQL SERVER MANAGEMENT STUDIO and create the database called **Hospital\_sourceDB** Database.

**Hospital\_sourceDB** have following tables -:

- Hospital table.
- Department Table.
- Doctor Table.
- Customer Table.
- Patient Table.

**Text file** –This text file include all the hospital address details including address, city, state, postal code and the country.

### 3. Solution Architecture.



#### Hospital stagingDB

- Hospital\_Staging
- Location\_Staging
- Department\_Staging
- Doctor\_Staging
- Location\_Staging
- Patient\_Staging
- accm\_txn\_complete\_time\_Staging

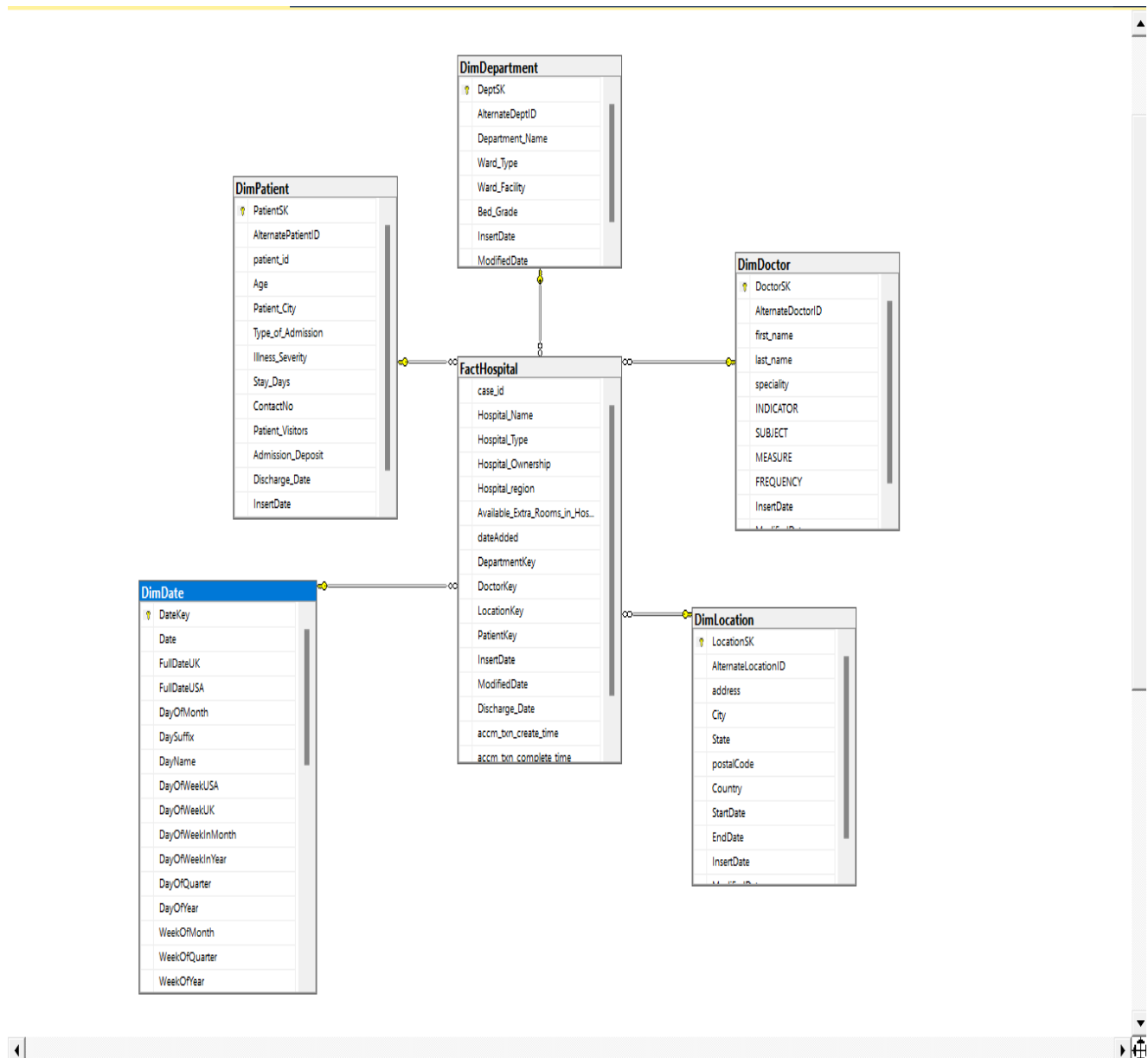
#### Hospital DW

- DimLocation
- DimDepartmen

- DimDoctor
- DimLocation
- FactHospital

## 4. Data warehouse design & development.

### Relational Diagram – Star Schema.

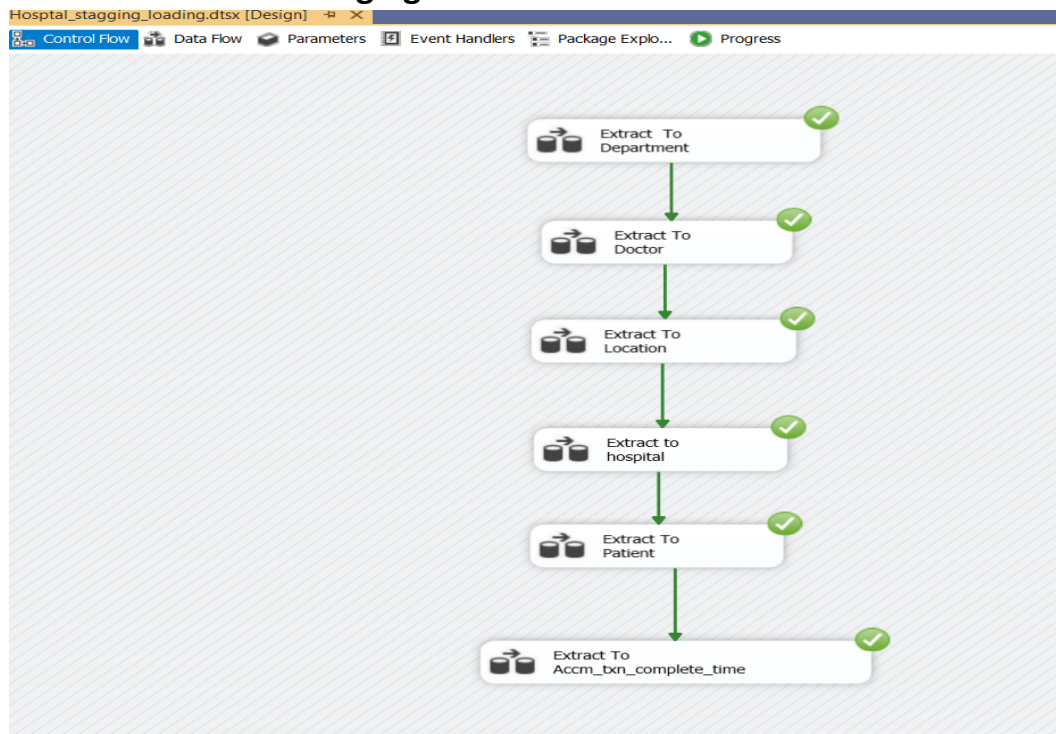


DimLocation is slowly changing dimension. Hospital Address may be changed in future. Therefore, I get it as slowly changing attribute.

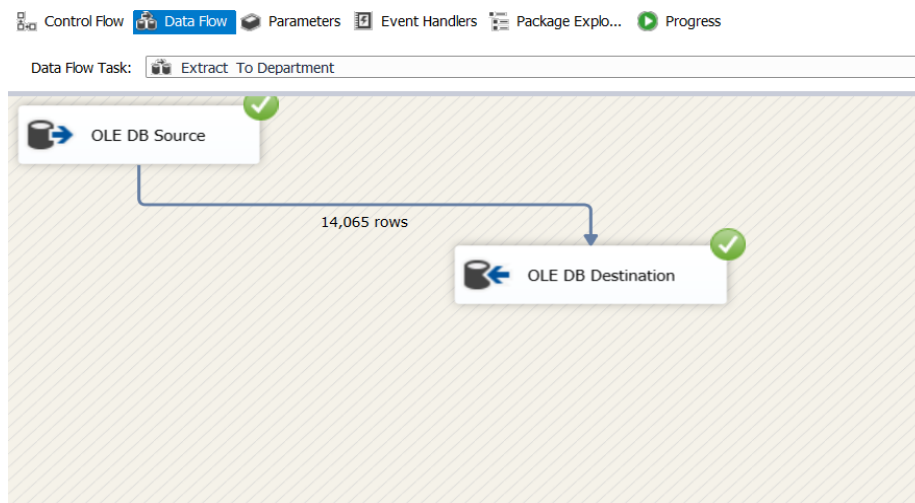
Country-> City-> Postal Code-> State-> Address = This is the Hierarchies.

## 5. ETL Development

### 5.1 ETL –Source to Staging

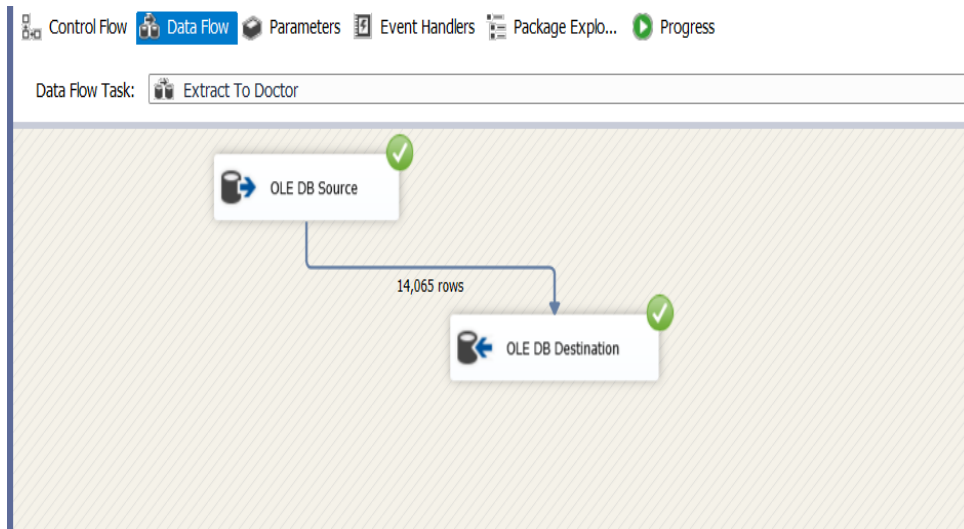


#### 5.1.1 Load Data Department Staging Table

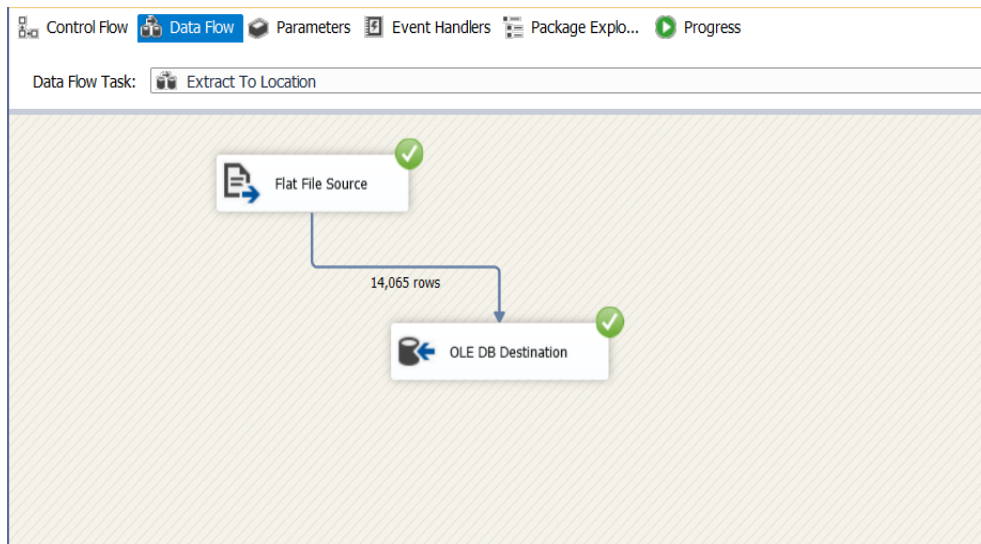




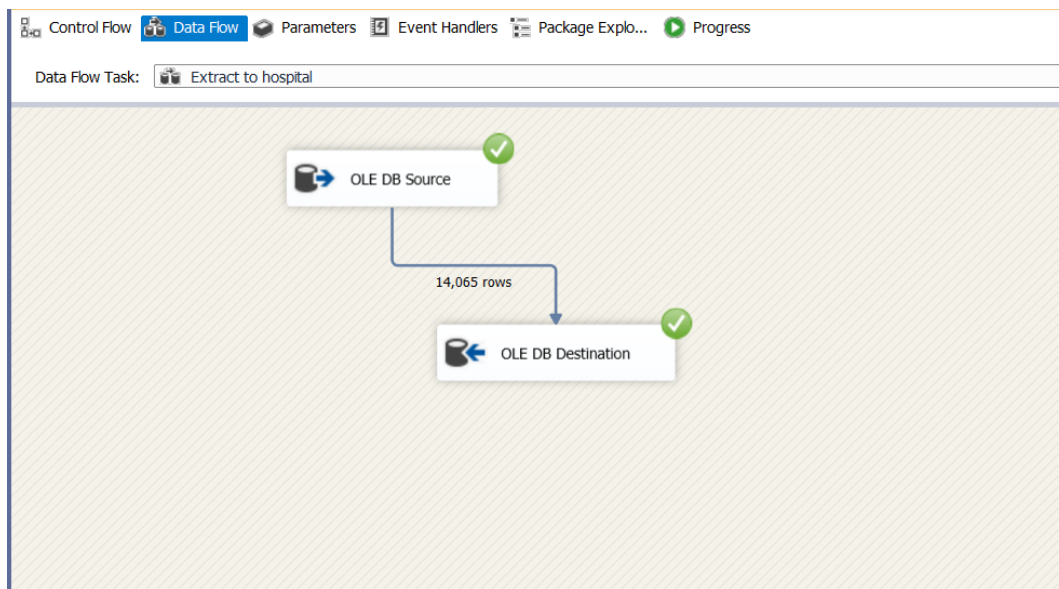
### 5.1.2 Load Data Doctor Staging Table



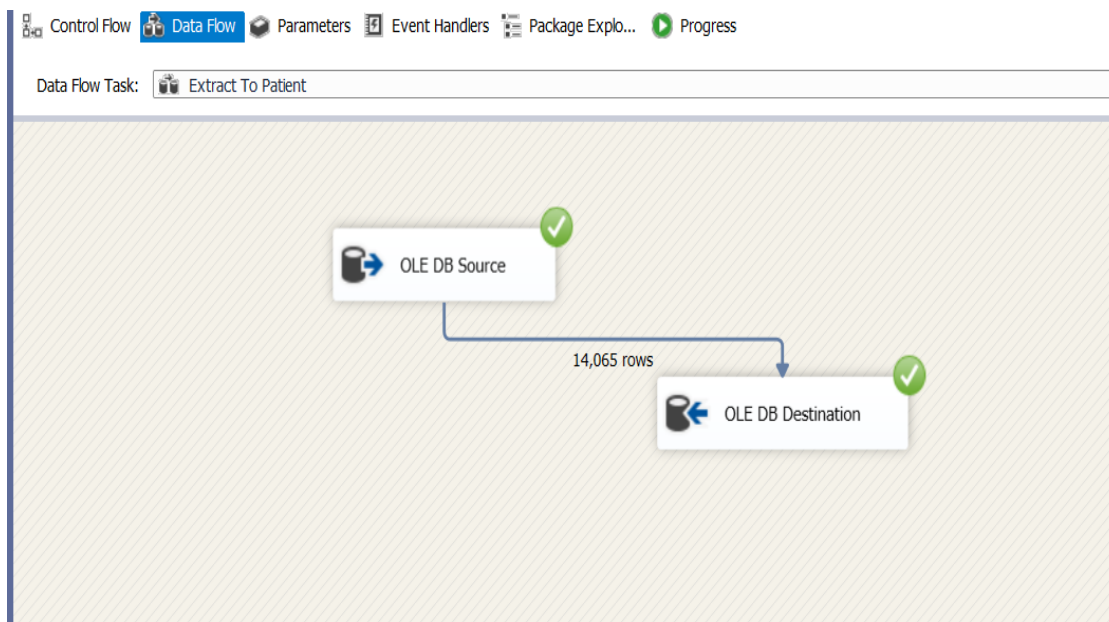
### 5.1.3 Load Data Location Staging Table



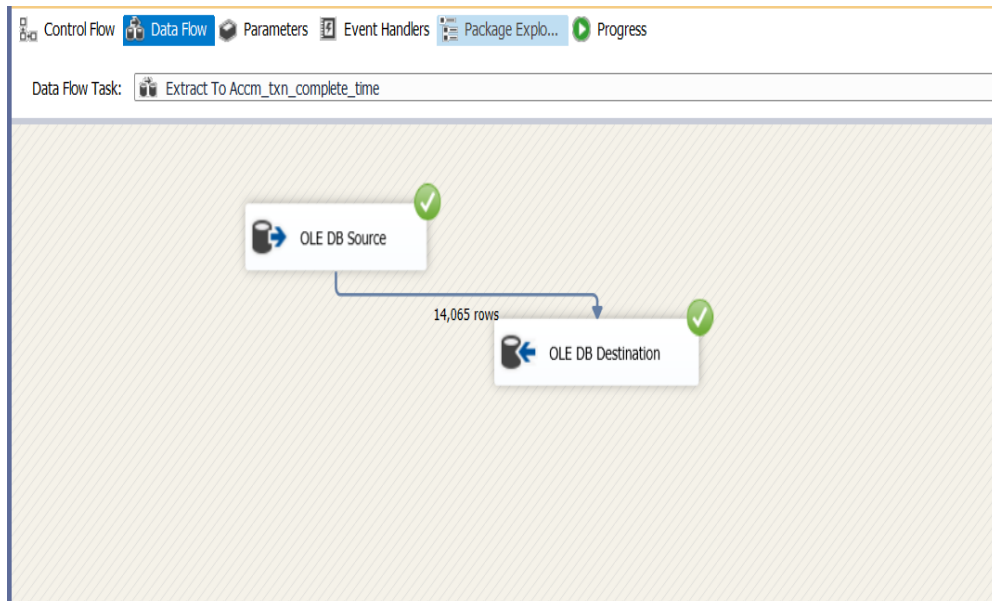
### 5.1.4 Load Data Hospital Staging Table



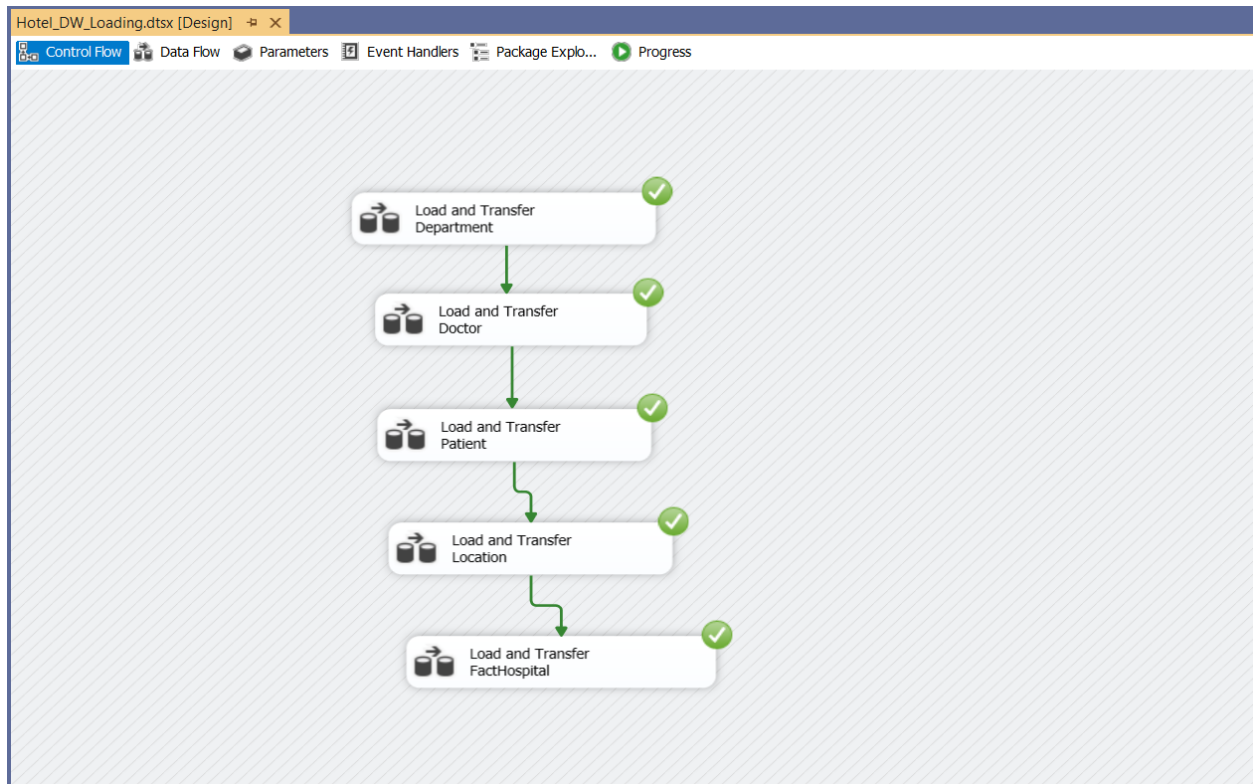
### 5.1.5 Load Data patient Staging Table



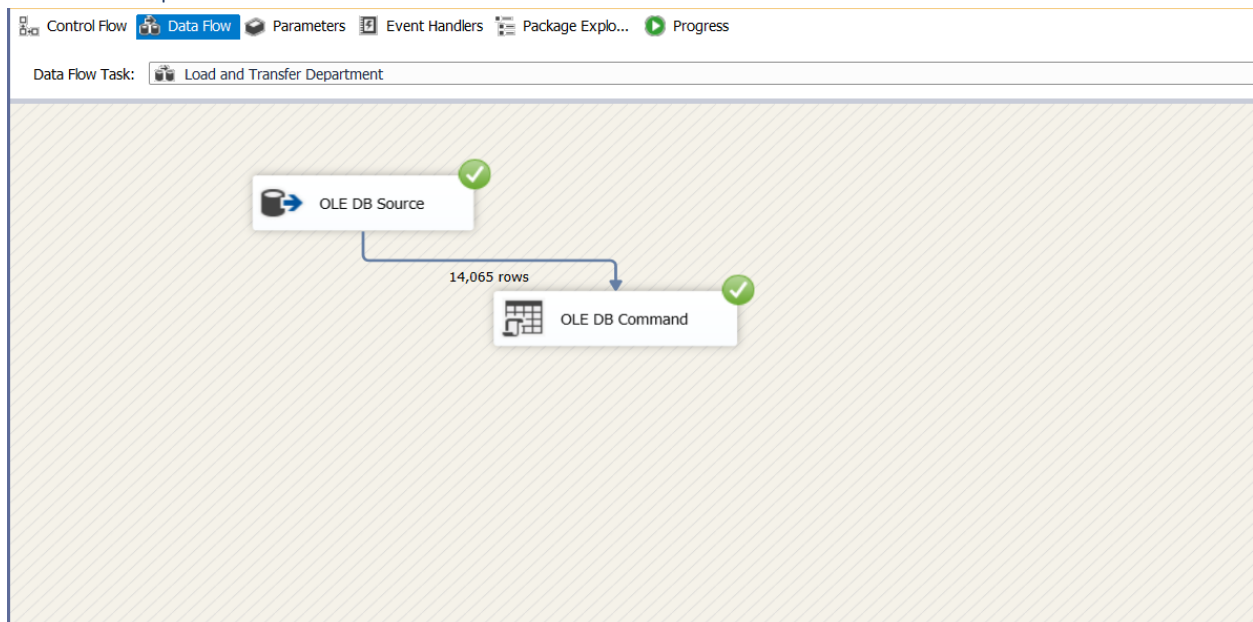
### 5.1.6 Load Data Branch Staging Table



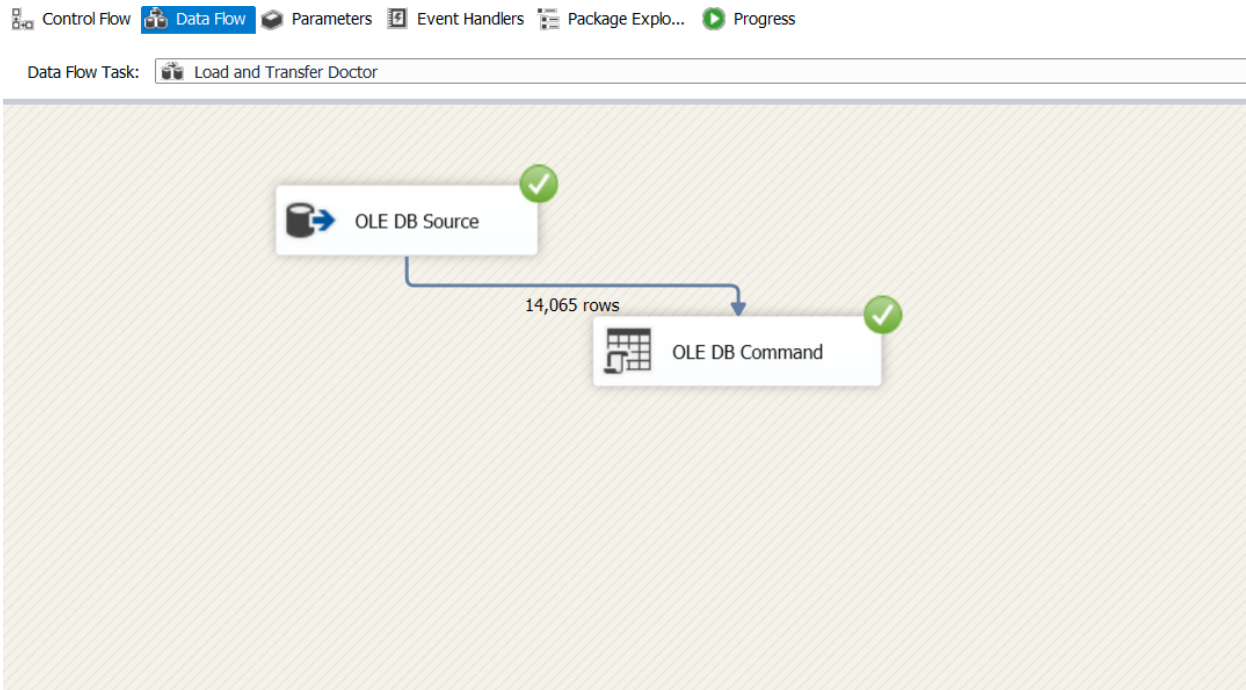
## 6. Staging To Data Warehouse



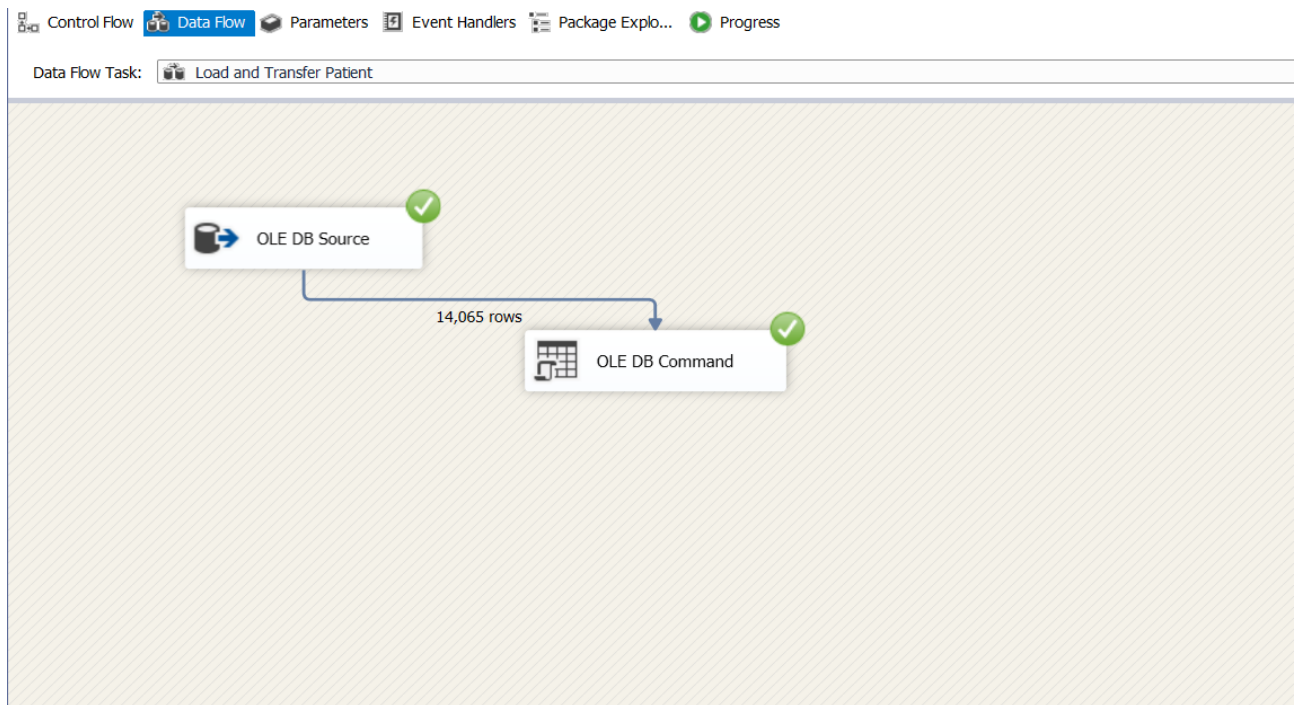
### 6.1 Load Department Data



## 6.2 Load Doctor Data



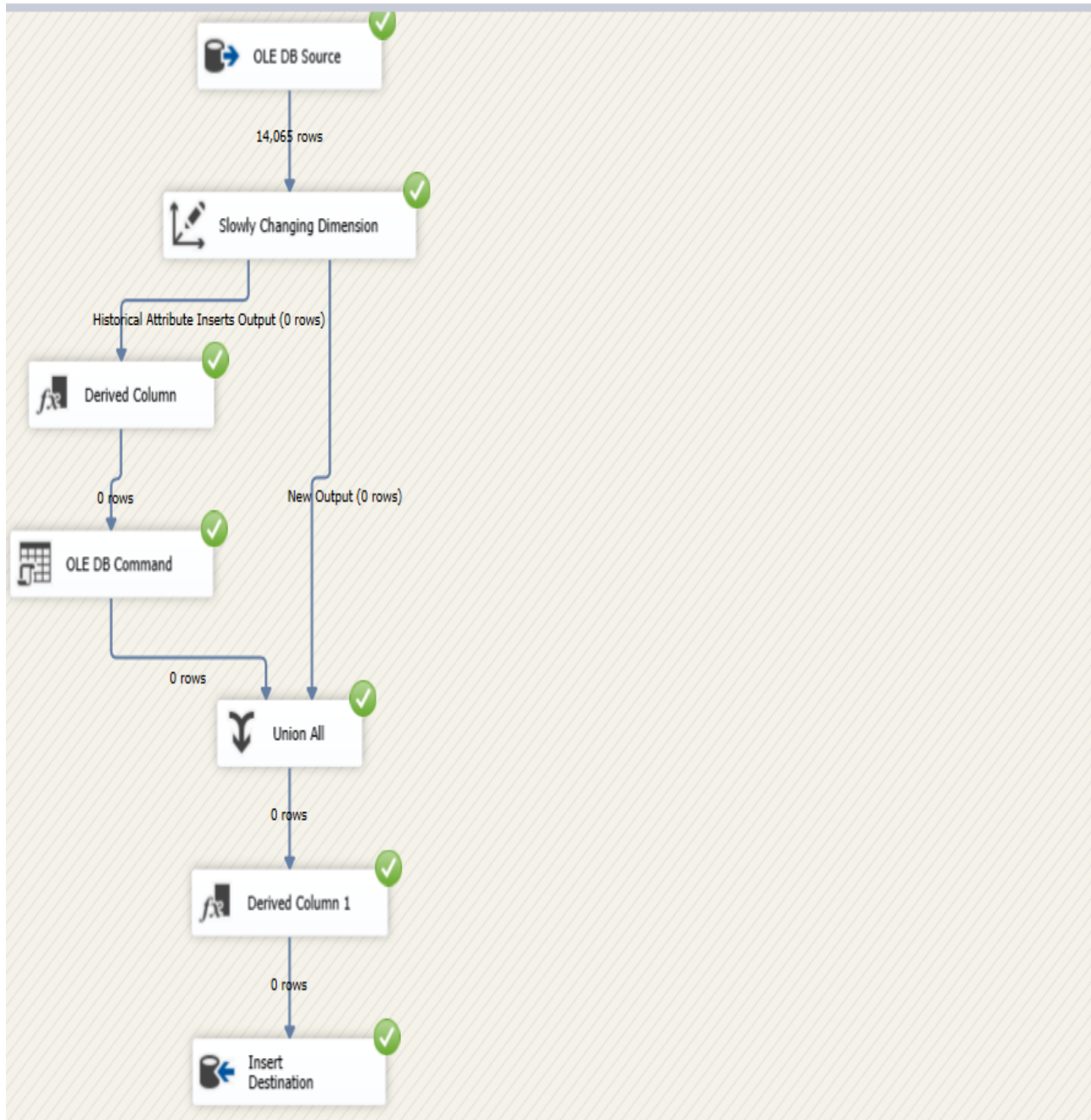
## 6.3 Load Patient Data



## 6.4 Load Location Data

Control Flow Data Flow Parameters Event Handlers Package Explo... Progress

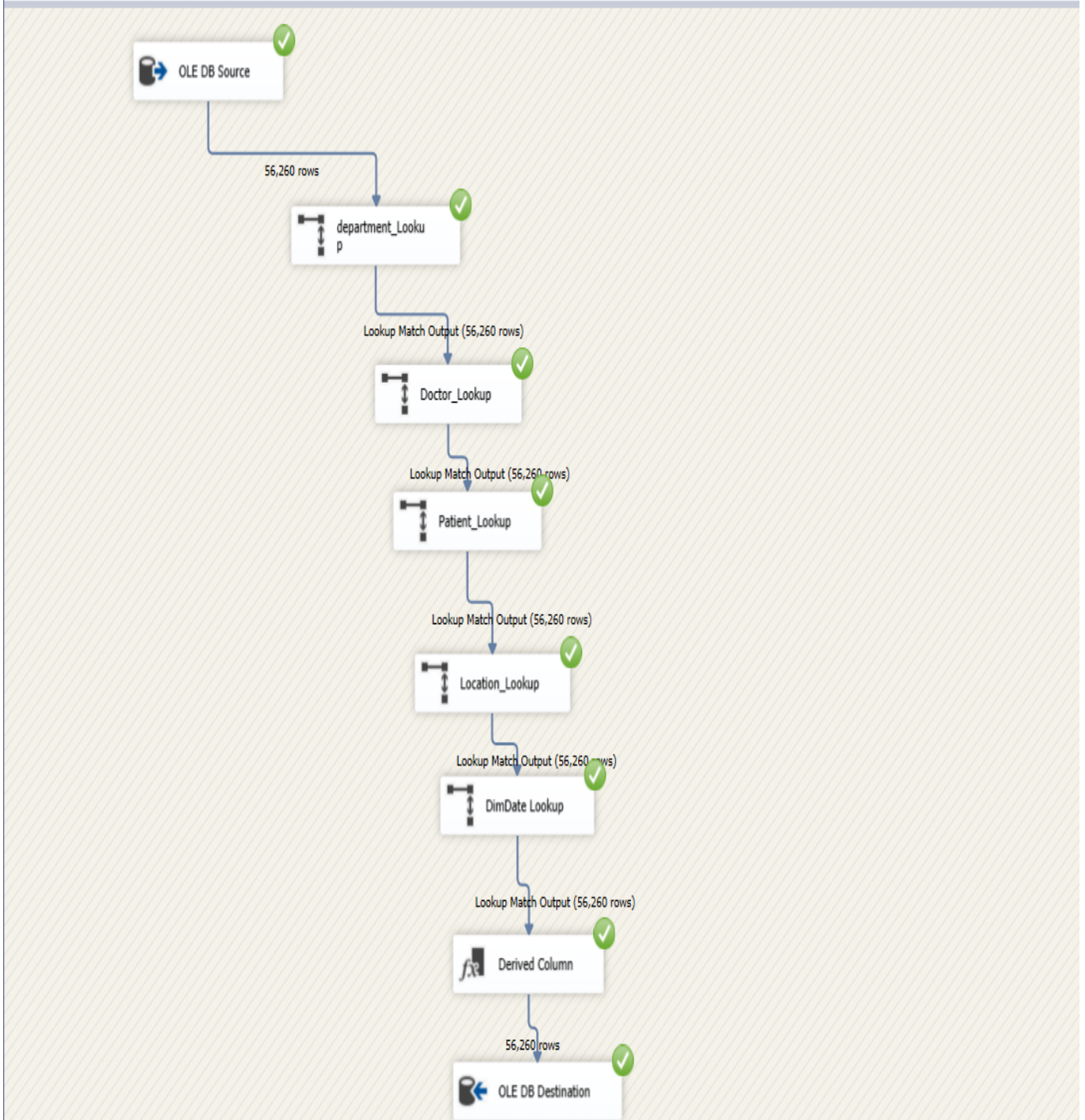
Data Flow Task: Load and Transfer Location



## 6.5 Load FactHospital Data

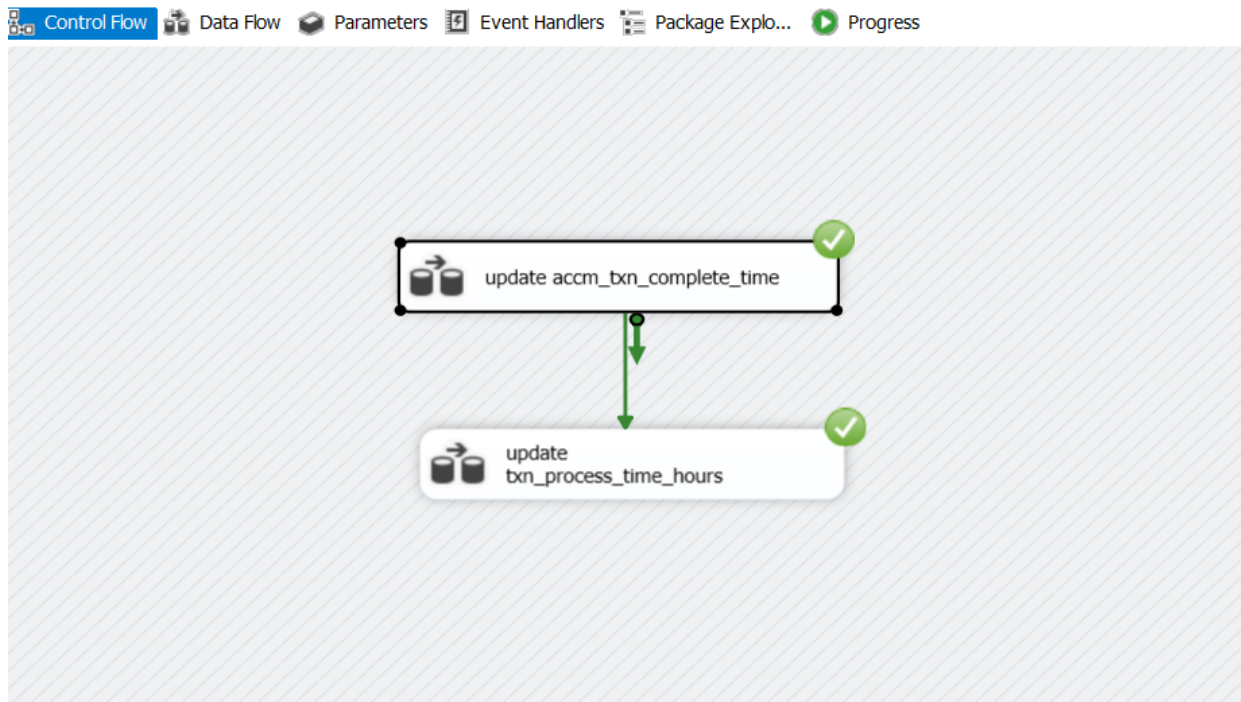
Control Flow Data Flow Parameters Event Handlers Package Explo... Progress

Data Flow Task: Load and Transfer FactHospital

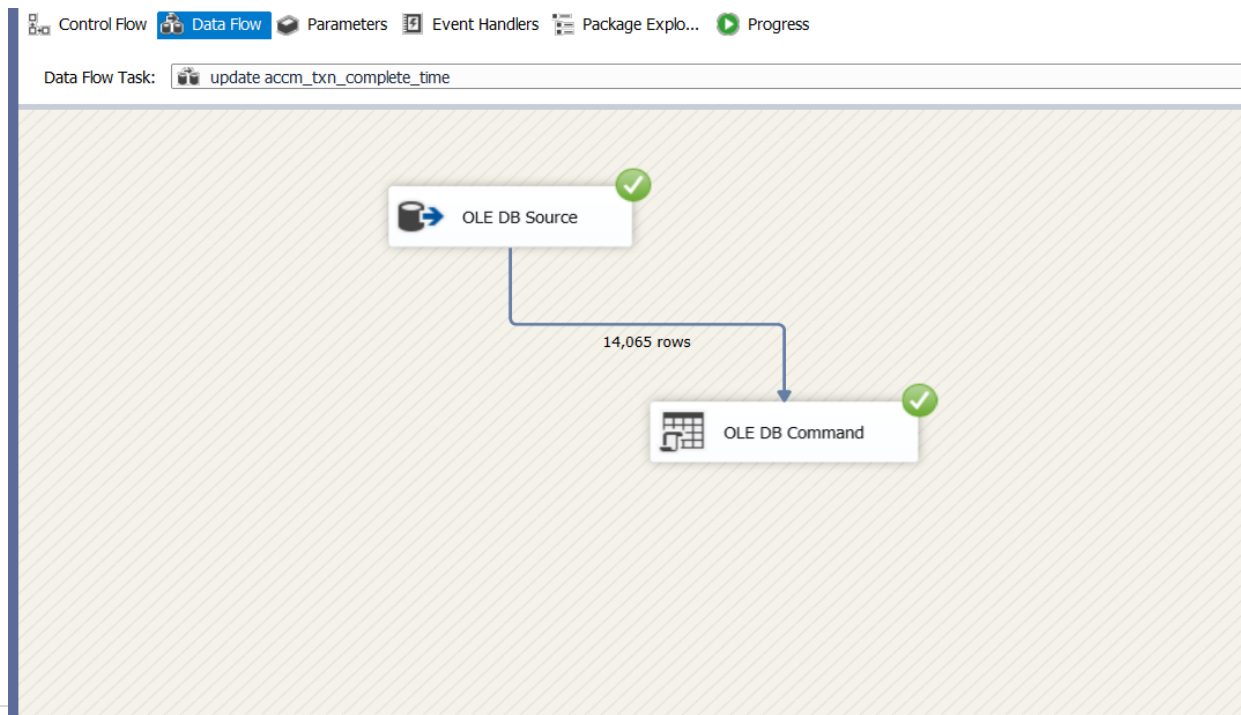




## 7. Data Warehouse Update SISS Package

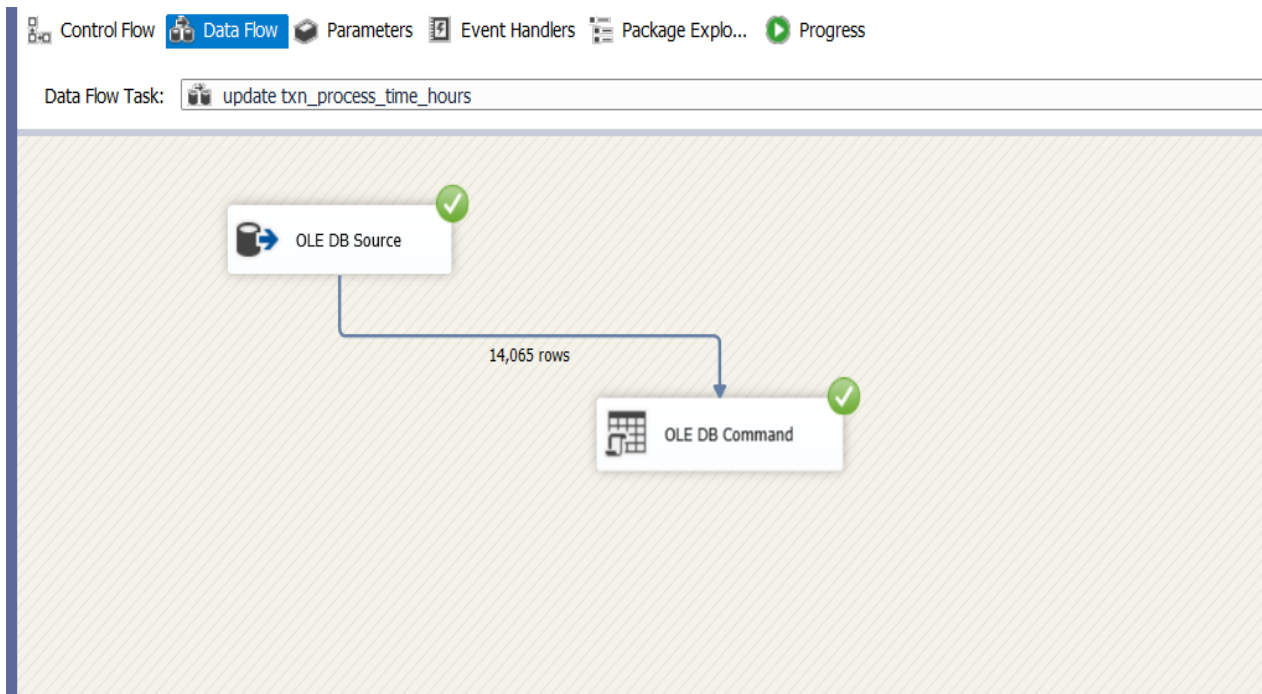


### 7.1 Load accm\_txn\_complete\_time Data





## 7.2 Load\_txn\_process\_time\_houres Data



# 8. Accumulated Fact Table

SQLQuery1.sql - DESKTOP-1K8MV8\MSSQLSERVER1\Hospital\_DW (DESKTOP-1K8MV8\DELL (67)) - Microsoft SQL Server Management Studio

Quick Launch (Ctrl+Q)

File Edit View Query Project Tools Window Help

Execute

Object Explorer

Connect

SQLQuery1.sql - D:\K8MV8\DELL (67) \*

\*\*\*\*\* Script for SelectTopRows command from SSMS \*\*\*\*\*

```
SELECT TOP (1000) [case_id]
```

Results Messages

case_id	Hospital_Name	Hospital_Type	Hospital_Ownership	Hospital_region	Available_Extra_Rooms_in_Hospital	dateAdded	DepartmentKey	DoctorKey	LocationKey	PatientKey	InsertDate
1	5643	NORTH FLORIDA REGIONAL MEDICAL CENTER	Acute Care Hospitals	Voluntary non-profit - Church	0	4	20210630	5643	13038	5643	2022-06-03 23:53:50.357
2	5644	MEMORIAL HOSPITAL OF TAMPA	Acute Care Hospitals	Government - Local	0	2	20210630	5644	13039	5644	2022-06-03 23:53:50.357
3	5645	KENDALL REGIONAL MEDICAL CENTER	Acute Care Hospitals	Government - Local	2	1	20210630	5645	13040	5645	2022-06-03 23:53:50.357
4	5646	BAYFRONT HEALTH DADE CITY	Acute Care Hospitals	Voluntary non-profit - Other	2	1	20210630	5646	13041	5646	2022-06-03 23:53:50.357
5	5647	OCALA REGIONAL MEDICAL CENTER	Acute Care Hospitals	Voluntary non-profit - Private	1	2	20210630	5647	13042	5647	2022-06-03 23:53:50.357
6	5648	BLAKE MEDICAL CENTER	Acute Care Hospitals	Government - Local	1	5	20210630	5648	13043	5648	2022-06-03 23:53:50.357
7	5649	SEBASTIAN RIVER MEDICAL CENTER	Acute Care Hospitals	Voluntary non-profit - Private	1	5	20210630	5649	13044	5649	2022-06-03 23:53:50.357
8	5650	GULF COAST MEDICAL CENTER LEE MEM HEALTH SYSTEM	Acute Care Hospitals	Voluntary non-profit - Other	0	2	20210630	5650	13045	5650	2022-06-03 23:53:50.357
9	5651	FORT WALTON BEACH MEDICAL CENTER	Acute Care Hospitals	Government - Local	1	4	20210630	5651	13046	5651	2022-06-03 23:53:50.357
10	5652	UNIVERSITY HOSPITAL AND MEDICAL CENTER	Acute Care Hospitals	Voluntary non-profit - Private	1	4	20210630	5652	13047	5652	2022-06-03 23:53:50.357
11	5653	ORANGE PARK MEDICAL CENTER	Acute Care Hospitals	Government - Local	1	4	20210630	5653	13048	5653	2022-06-03 23:53:50.357
12	5654	WESTSIDE REGIONAL MEDICAL CENTER	Acute Care Hospitals	Government - Local	1	2	20210630	5654	13049	5654	2022-06-03 23:53:50.357
13	5655	MEMORIAL HOSPITAL PEMBROKE	Acute Care Hospitals	Government - Local	0	3	20210630	5655	13050	5655	2022-06-03 23:53:50.357
14	5656	WEST FLORIDA HOSPITAL	Acute Care Hospitals	Voluntary non-profit - Private	0	5	20210630	5656	13051	5656	2022-06-03 23:53:50.357
15	5657	PUTNAM COMMUNITY MEDICAL CENTER	Acute Care Hospitals	Government - Local	0	3	20210630	5657	13052	5657	2022-06-03 23:53:50.357
16	5658	FAWCETT MEMORIAL HOSPITAL	Acute Care Hospitals	Voluntary non-profit - Other	0	3	20210630	5658	13053	5658	2022-06-03 23:53:50.357
17	5659	NORTHSIDE HOSPITAL	Acute Care Hospitals	Voluntary non-profit - Other	0	5	20210701	5659	13054	5659	2022-06-03 23:53:50.357
18	5660	ANNE BATES LEACH EYE HOSPITAL	Acute Care Hospitals	Voluntary non-profit - Other	1	4	20210701	5660	13055	5660	2022-06-03 23:53:50.357
19	5661	GULF COAST REGIONAL MEDICAL CENTER	Acute Care Hospitals	Government - Local	2	4	20210701	5661	13056	5661	2022-06-03 23:53:50.357
20	5662	BRANDON REGIONAL HOSPITAL	Acute Care Hospitals	Voluntary non-profit - Private	2	2	20210701	5662	13057	5662	2022-06-03 23:53:50.357
21	5663	CAPE CORAL HOSPITAL	Acute Care Hospitals	Voluntary non-profit - Private	2	2	20210701	5663	13058	5663	2022-06-03 23:53:50.357
22	5664	LAWNWOOD REGIONAL MEDICAL CENTER & HEART INST.	Acute Care Hospitals	Voluntary non-profit - Private	1	3	20210701	5664	13059	5664	2022-06-03 23:53:50.357
23	5665	LARGO MEDICAL CENTER	Acute Care Hospitals	Voluntary non-profit - Other	2	2	20210701	5665	13060	5665	2022-06-03 23:53:50.357
24	5666	SEVEN RIVERS REGIONAL MEDICAL CENTER	Acute Care Hospitals	Voluntary non-profit - Other	1	2	20210701	5666	13061	5666	2022-06-03 23:53:50.357
25	5667	RAILROAD HOSPITAL	Acute Care Hospitals	Proprietary	1	3	20210701	5667	13062	5667	2022-06-03 23:53:50.357
26	5668	JURTER MEDICAL CENTER	Acute Care Hospitals	Government - Local	1	4	20210702	5668	13063	5668	2022-06-03 23:53:50.357
27	5669	CAPITAL REGIONAL MEDICAL CENTER	Acute Care Hospitals	Government - Local	2	4	20210702	5669	13064	5669	2022-06-03 23:53:50.357
28	5670	TAMPA COMMUNITY HOSPITAL	Acute Care Hospitals	Government - Local	1	4	20210702	5670	13065	5670	2022-06-03 23:53:50.357
29	5671	REGIONAL MEDICAL CENTER BAYONET POINT	Acute Care Hospitals	Voluntary non-profit - Other	0	2	20210702	5671	13066	5671	2022-06-03 23:53:50.357
30	5672	DELRAY MEDICAL CENTER	Acute Care Hospitals	Government - Local	0	2	20210702	5672	13067	5672	2022-06-03 23:53:50.357
31	5673	SOUTH BAY HOSPITAL	Acute Care Hospitals	Government - Local	1	2	20210702	5673	13068	5673	2022-06-03 23:53:50.357
32	5674	ST LUCIE MEDICAL CENTER	Acute Care Hospitals	Voluntary non-profit - Church	1	5	20210702	5674	13069	5674	2022-06-03 23:53:50.357

Query executed successfully.

DESKTOP-1K8MV8\MSSQLSERVER... DESKTOP-1K8MV8\DELL (67) Hospital\_DW 00:00:00 1,000 rows

SQLQuery1.sql - DESKTOP-1K8MV8\MSSQLSERVER1\Hospital\_DW (DESKTOP-1K8MV8\DELL (67)) - Microsoft SQL Server Management Studio

Quick Launch (Ctrl+Q)

File Edit View Query Project Tools Window Help

Execute

Object Explorer

Connect

SQLQuery1.sql - D:\K8MV8\DELL (67) \*

\*\*\*\*\* Script for SelectTopRows command from SSMS \*\*\*\*\*

```
SELECT TOP (1000) [case_id]
```

Results Messages

case_id	Hospital_region	Available_Extra_Rooms_in_Hospital	dateAdded	DepartmentKey	DoctorKey	LocationKey	PatientKey	InsertDate	ModifiedDate	Discharge_Date	acorn_tn_create_time	acorn_tn_complete_time	tn_process_time_hours
1	0	4	20210630	5643	13038	5643	5643	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-01 00:00:00.000	-71
2	0	2	20210630	5644	13039	5644	5644	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-02 00:00:00.000	-47
3	2	1	20210630	5645	13040	5645	5645	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-03 00:00:00.000	-23
4	2	1	20210630	5646	13041	5646	5646	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-04 00:00:00.000	1
5	1	2	20210630	5647	13042	5647	5647	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-05 00:00:00.000	25
6	1	5	20210630	5648	13043	5648	5648	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-06 00:00:00.000	49
7	1	5	20210630	5649	13044	5649	5649	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-07 00:00:00.000	73
8	0	2	20210630	5650	13045	5650	5650	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-08 00:00:00.000	97
9	1	4	20210630	5651	13046	5651	5651	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-09 00:00:00.000	121
10	1	4	20210630	5652	13047	5652	5652	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-10 00:00:00.000	145
11	1	4	20210630	5653	13048	5653	5653	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-11 00:00:00.000	169
12	1	2	20210630	5654	13049	5654	5654	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-12 00:00:00.000	193
13	0	3	20210630	5655	13050	5655	5655	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-13 00:00:00.000	217
14	0	5	20210630	5656	13051	5656	5656	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-14 00:00:00.000	241
15	0	3	20210630	5657	13052	5657	5657	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-15 00:00:00.000	-479
16	0	5	20210630	5658	13053	5658	5658	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-16 00:00:00.000	-455
17	0	5	20210701	5659	13054	5659	5659	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-17 00:00:00.000	-431
18	1	4	20210701	5660	13055	5660	5660	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-18 00:00:00.000	-407
19	2	4	20210701	5661	13056	5661	5661	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-19 00:00:00.000	-383
20	2	2	20210701	5662	13057	5662	5662	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-20 00:00:00.000	-359
21	2	2	20210701	5663	13058	5663	5663	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-21 00:00:00.000	-335
22	1	3	20210701	5664	13059	5664	5664	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-22 00:00:00.000	-311
23	2	2	20210701	5665	13060	5665	5665	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-23 00:00:00.000	-287
24	1	2	20210701	5666	13061	5666	5666	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-24 00:00:00.000	-263
25	1	3	20210701	5667	13062	5667	5667	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-25 00:00:00.000	-239
26	1	4	20210702	5668	13063	5668	5668	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-26 00:00:00.000	-215
27	2	4	20210702	5669	13064	5669	5669	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-27 00:00:00.000	-191
28	1	4	20210702	5670	13065	5670	5670	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-28 00:00:00.000	-167
29	0	2	20210702	5671	13066	5671	5671	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-29 00:00:00.000	-143
30	0	2	20210702	5672	13067	5672	5672	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-30 00:00:00.000	-119
31	1	2	20210702	5673	13068	5673	5673	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-31 00:00:00.000	-95
32	1	5	20210702	5674	13069	5674	5674	2022-06-03 23:53:50.357	2022-06-03 23:53:50.357	NULL	2022-06-03 23:53:50.357	2022-06-01 00:00:00.000	-71

Query executed successfully.

DESKTOP-1K8MV8\MSSQLSERVER... DESKTOP-1K8MV8\DELL (67) Hospital\_DW 00:00:00 1,000 rows

