

HW 07 Staging

DWBI

Shivam Verma

Data is first loaded into a staging table then using data conversion, transformation is done. Finally, that data is loaded in destination table.

Create Table Script

For Staging Table

```
CREATE TABLE [dbo].[covid_data_staging](
[iso_code] [varchar](255) NULL,
[continent] [varchar](255) NULL,
[location] [varchar](255) NULL,
[date] [varchar](255) NULL,
[total_cases] [varchar](255) NULL,
[new_cases] [varchar](255) NULL,
[new_cases_smoothed] [varchar](255) NULL,
[total_deaths] [varchar](255) NULL,
[new_deaths] [varchar](255) NULL,
[new_deaths_smoothed] [varchar](255) NULL,
[total_cases_per_million] [varchar](255) NULL,
[new_cases_per_million] [varchar](255) NULL,
[new_cases_smoothed_per_million] [varchar](255) NULL,
[total_deaths_per_million] [varchar](255) NULL,
[new_deaths_per_million] [varchar](255) NULL,
[new_deaths_smoothed_per_million] [varchar](255) NULL,
[total_tests] [varchar](255) NULL,
[new_tests] [varchar](255) NULL,
[new_tests_smoothed] [varchar](255) NULL,
[total_tests_per_thousand] [varchar](255) NULL,
[new_tests_per_thousand] [varchar](255) NULL,
[new_tests_smoothed_per_thousand] [varchar](255) NULL,
[tests_per_case] [varchar](255) NULL,
[positive_rate] [varchar](255) NULL,
[tests_units] [varchar](255) NULL,
[stringency_index] [varchar](255) NULL,
[population] [varchar](255) NULL,
[population_density] [varchar](255) NULL,
[median_age] [varchar](255) NULL,
[aged_65_older] [varchar](255) NULL,
[aged_70_older] [varchar](255) NULL,
[gdp_per_capita] [varchar](255) NULL,
[extreme_poverty] [varchar](255) NULL,
[cardiovasc_death_rate] [varchar](255) NULL,
[diabetes_prevalence] [varchar](255) NULL,
[female_smokers] [varchar](255) NULL,
[male_smokers] [varchar](255) NULL,
[handwashing_facilities] [varchar](255) NULL,
[hospital_beds_per_thousand] [varchar](255) NULL,
```

```

[life_expectancy] [varchar](255) NULL,
[human_development_index] [varchar](255) NULL
) ON [PRIMARY]
GO

```

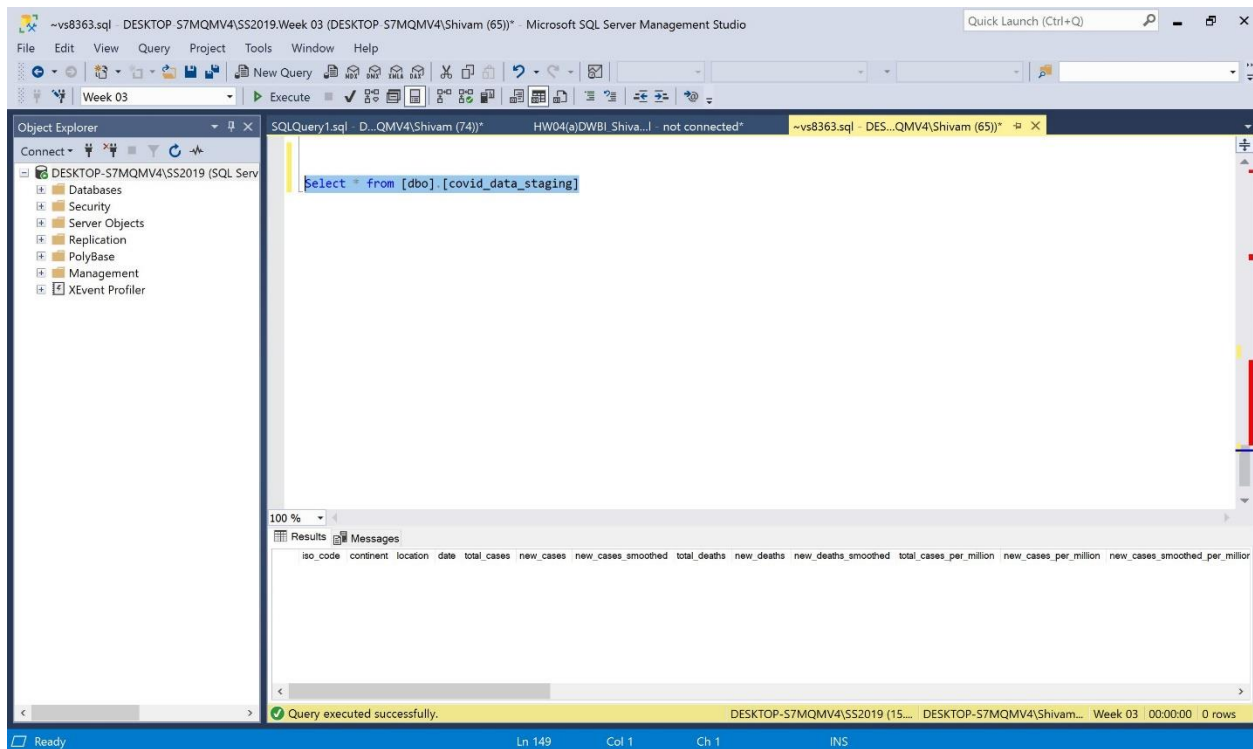
For Destination Table

```

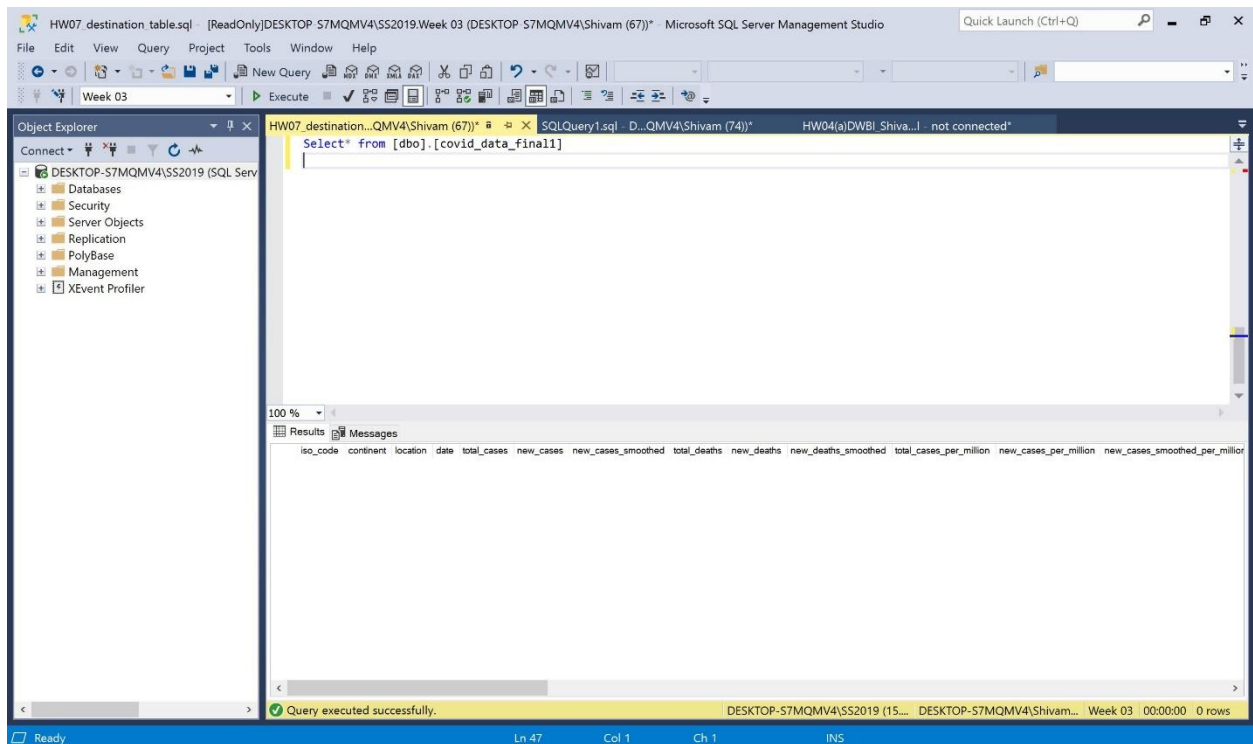
CREATE TABLE [dbo].[covid_data_final](
[iso_code] [varchar](255) NULL,
[continent] [varchar](255) NULL,
[location] [varchar](255) NULL,
[date] [date] NULL,
[total_cases] [int] NULL,
[new_cases] [int] NULL,
[new_cases_smoothed] [float] NULL,
[total_deaths] [int] NULL,
[new_deaths] [int] NULL,
[new_deaths_smoothed] [float] NULL,
[total_cases_per_million] [float] NULL,
[new_cases_per_million] [float] NULL,
[new_cases_smoothed_per_million] [float] NULL,
[total_deaths_per_million] [float] NULL,
[new_deaths_per_million] [float] NULL,
[new_deaths_smoothed_per_million] [float] NULL,
[total_tests] [float] NULL,
[new_tests] [int] NULL,
[new_tests_smoothed] [float] NULL,
[total_tests_per_thousand] [float] NULL,
[new_tests_per_thousand] [float] NULL,
[new_tests_smoothed_per_thousand] [float] NULL,
[tests_per_case] [float] NULL,
[positive_rate] [float] NULL,
[tests_units] [varchar](255) NULL,
[stringency_index] [float] NULL,
[population] [bigint] NULL,
[population_density] [float] NULL,
[median_age] [float] NULL,
[aged_65_older] [float] NULL,
[aged_70_older] [float] NULL,
[gdp_per_capita] [float] NULL,
[extreme_poverty] [float] NULL,
[cardiovasc_death_rate] [float] NULL,
[diabetes_prevalence] [float] NULL,
[female_smokers] [float] NULL,
[male_smokers] [float] NULL,
[handwashing_facilities] [float] NULL,
[hospital_beds_per_thousand] [float] NULL,
[life_expectancy] [float] NULL,
[human_development_index] [float] NULL
) ON [PRIMARY]
GO

```

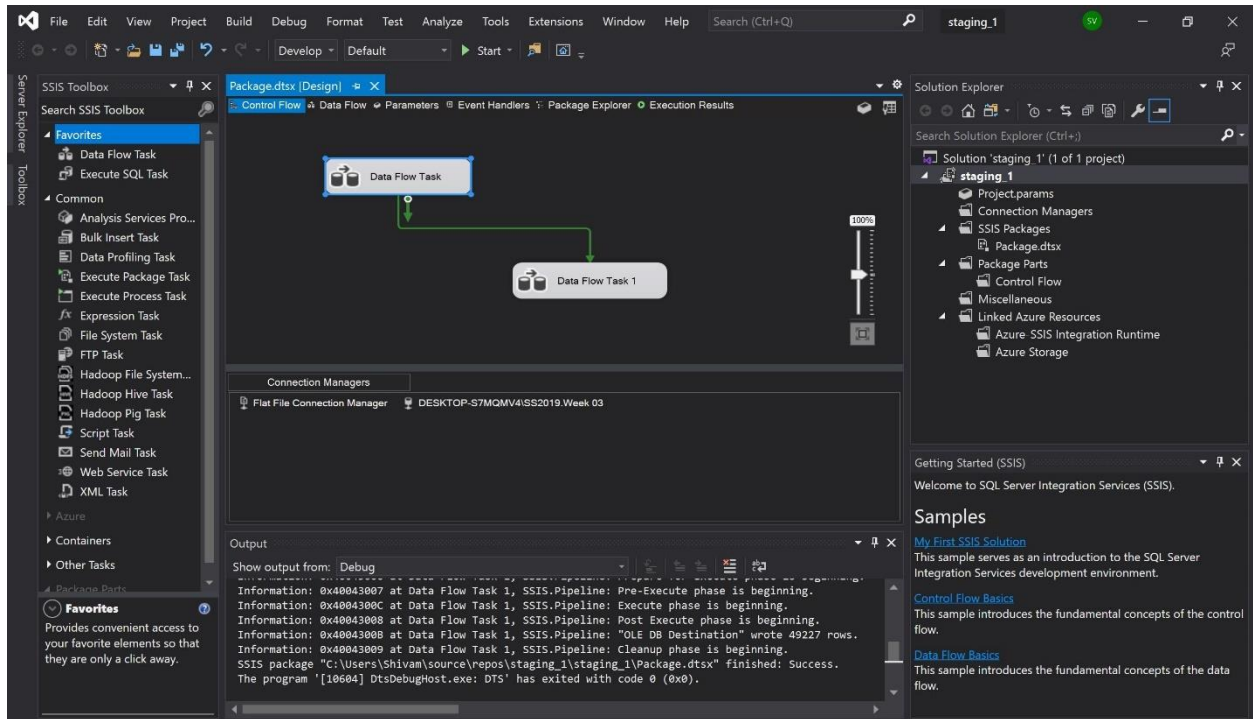
Staging table before data population



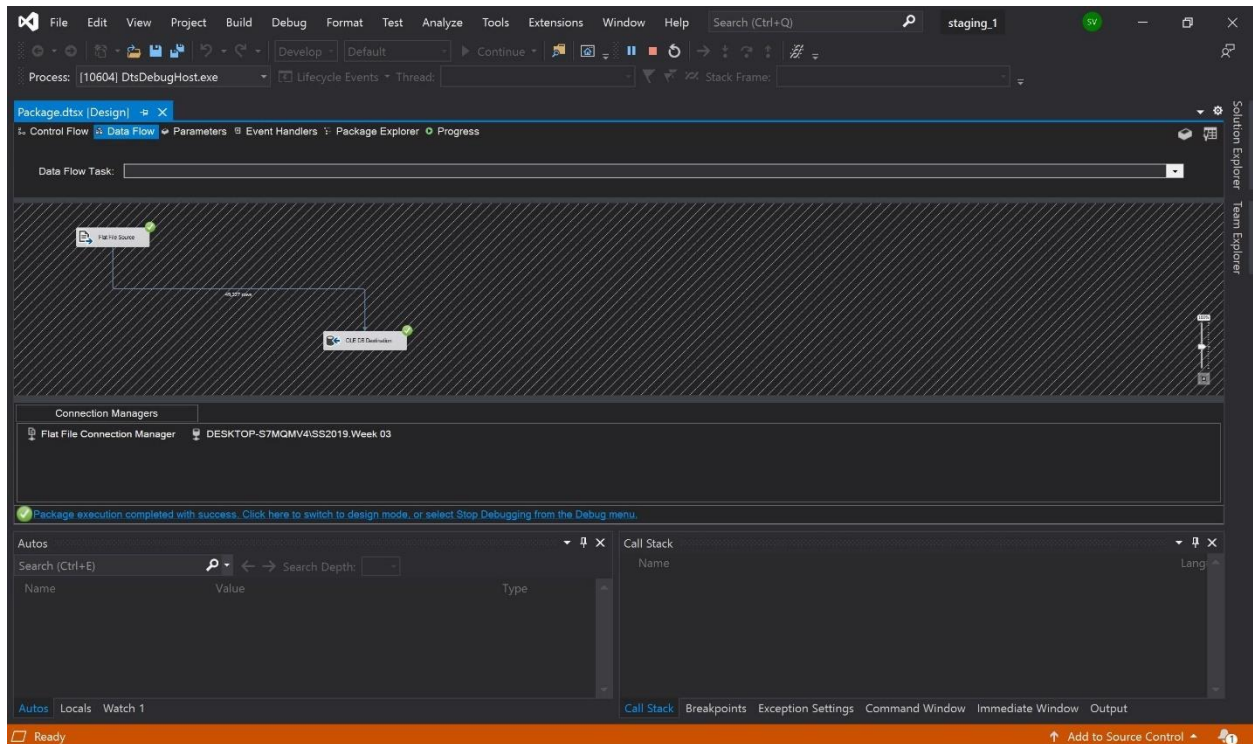
Final table before data population



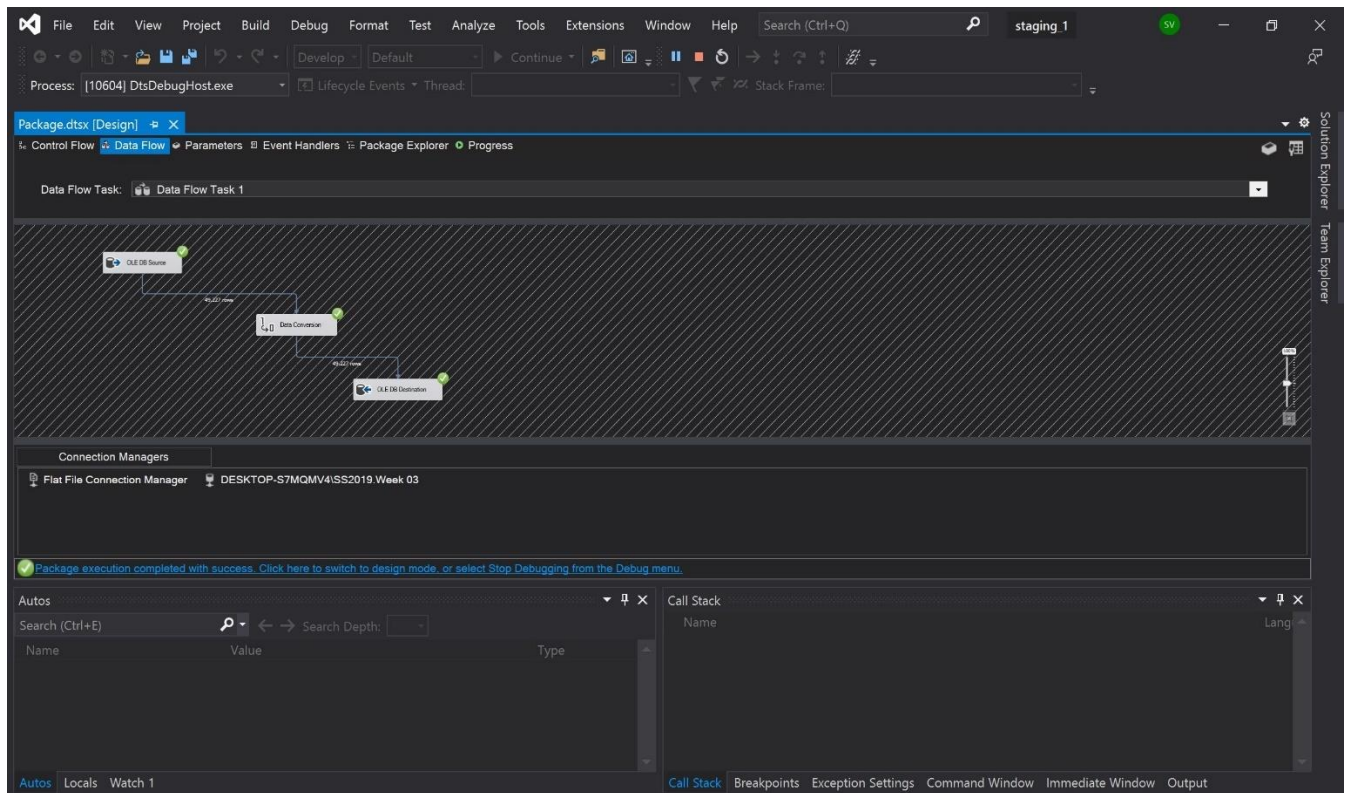
Screenshot of Control Flow



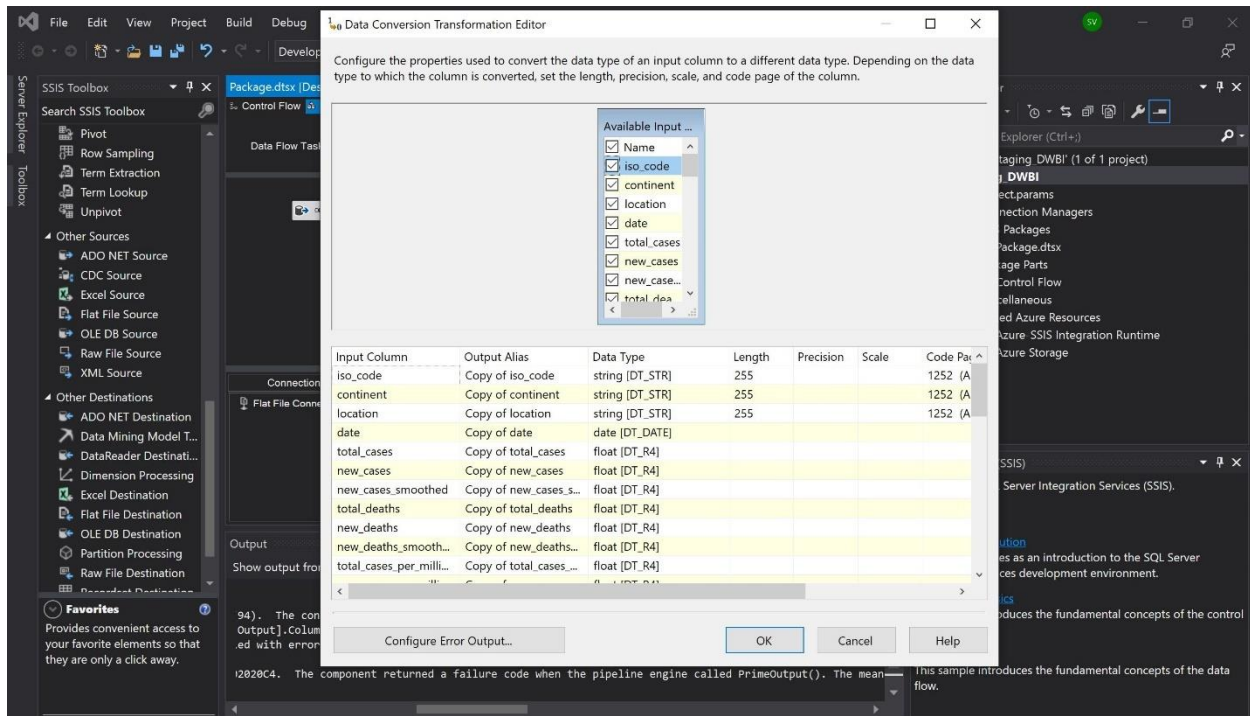
Screenshot of Data flow Task 1 (populating data from flat file to staging table)



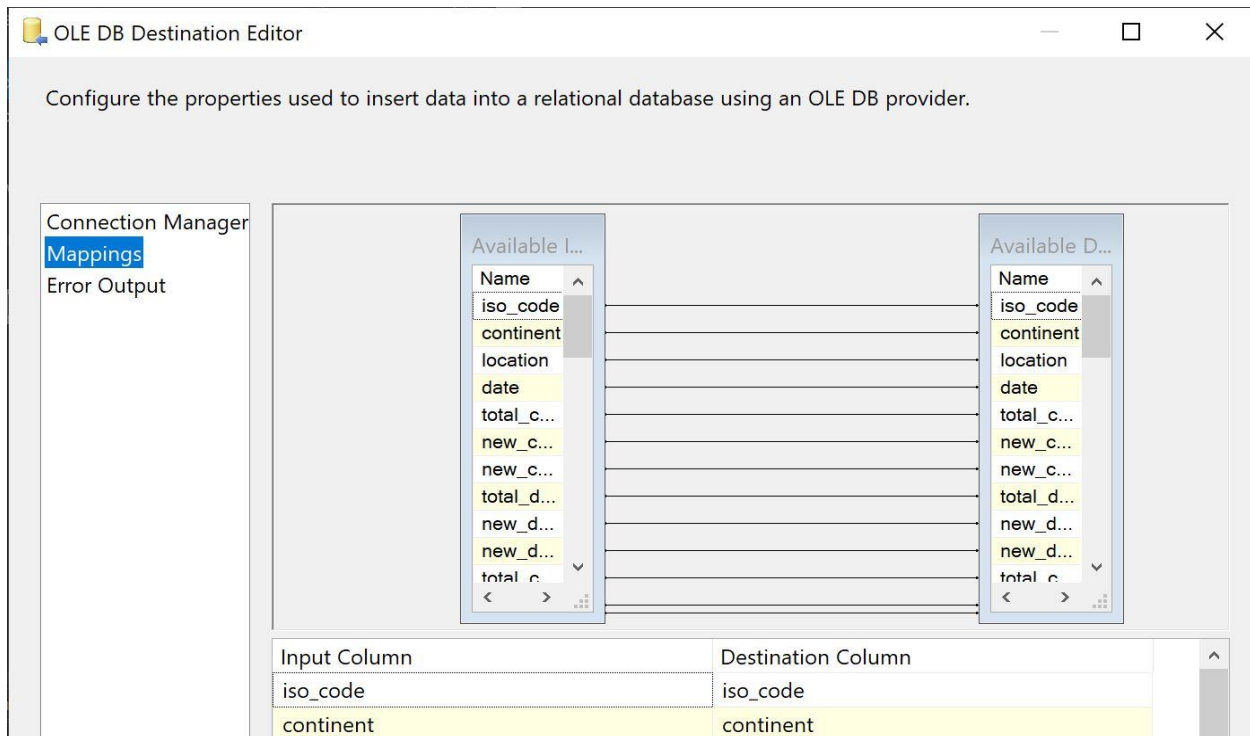
Screenshot of Data flow Task 2 (populating data from staging table to destination table)



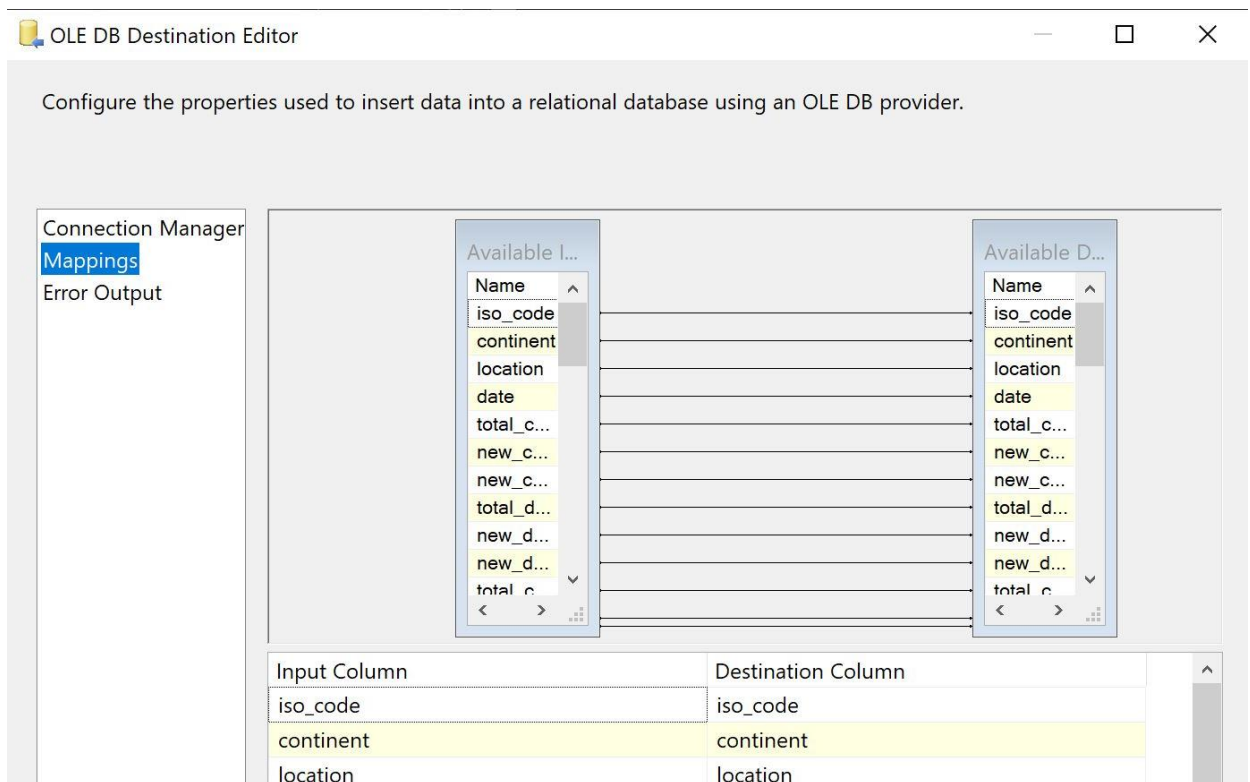
Screenshot of Data Conversion Task



Screenshot of staging table mappings



Screenshot of Destination table mappings



Screenshot of data populated into staging table

The screenshot shows the Microsoft SQL Server Management Studio interface. The top menu bar includes File, Edit, View, Query, Project, Tools, Window, and Help. Below the menu is a toolbar with various icons for file operations and database management. The main window displays a query executed against the 'DESKTOP-S7MQMV4\S2019' database. The query is 'Select * from [dbo].[covid_data_staging]'. The results are shown in a grid format with columns: iso_code, continent, location, date, total_cases, new_cases, new_cases_smoothed, total_deaths, new_deaths, new_deaths_smoothed, total_cases_per_million, new_cases_per_million, and new_cases_smoothed. The data rows show information for Afghanistan (AFG) across various dates from 2019-12-31 to 2020-02-01. A status bar at the bottom indicates 'Query executed successfully.' and provides details about the execution: 'DESKTOP-S7MQMV4\S2019 (15...)' and 'DESKTOP-S7MQMV4\Shivam...' with a duration of 'Week 03 00:00:03' and '49,227 rows'.

Screenshot of data populated into destination table

The screenshot displays the Microsoft SQL Server Management Studio interface. The query window shows the following SQL statement:

```
Select* from [dbo].[covid_data_final1]
```

The results pane displays a table with 12 columns and 31 rows of data. The columns are: iso_code, continent, location, date, total_cases, new_cases, new_cases_smoothed, total_deaths, new_deaths, new_deaths_smoothed, total_cases_per_million, new_cases_per_million, and new_cases_smoothed. The data represents COVID-19 statistics for Afghanistan from December 31, 2019, to January 30, 2020.

	iso_code	continent	location	date	total_cases	new_cases	new_cases_smoothed	total_deaths	new_deaths	new_deaths_smoothed	total_cases_per_million	new_cases_per_million	new_cases_smoothed
1	AFG	Asia	Afghanistan	2019-12-31	0.0	0.0		0.0	0.0		0.0	0.0	
2	AFG	Asia	Afghanistan	2020-01-01	0.0	0.0		0.0	0.0		0.0	0.0	
3	AFG	Asia	Afghanistan	2020-01-02	0.0	0.0		0.0	0.0		0.0	0.0	
4	AFG	Asia	Afghanistan	2020-01-03	0.0	0.0		0.0	0.0		0.0	0.0	
5	AFG	Asia	Afghanistan	2020-01-04	0.0	0.0		0.0	0.0		0.0	0.0	
6	AFG	Asia	Afghanistan	2020-01-05	0.0	0.0		0.0	0.0		0.0	0.0	
7	AFG	Asia	Afghanistan	2020-01-06	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	AFG	Asia	Afghanistan	2020-01-07	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	AFG	Asia	Afghanistan	2020-01-08	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10	AFG	Asia	Afghanistan	2020-01-09	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
11	AFG	Asia	Afghanistan	2020-01-10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
12	AFG	Asia	Afghanistan	2020-01-11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
13	AFG	Asia	Afghanistan	2020-01-12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
14	AFG	Asia	Afghanistan	2020-01-13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
15	AFG	Asia	Afghanistan	2020-01-14	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
16	AFG	Asia	Afghanistan	2020-01-15	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
17	AFG	Asia	Afghanistan	2020-01-16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
18	AFG	Asia	Afghanistan	2020-01-17	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
19	AFG	Asia	Afghanistan	2020-01-18	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	AFG	Asia	Afghanistan	2020-01-19	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
21	AFG	Asia	Afghanistan	2020-01-20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
22	AFG	Asia	Afghanistan	2020-01-21	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
23	AFG	Asia	Afghanistan	2020-01-22	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
24	AFG	Asia	Afghanistan	2020-01-23	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
25	AFG	Asia	Afghanistan	2020-01-24	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
26	AFG	Asia	Afghanistan	2020-01-25	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
27	AFG	Asia	Afghanistan	2020-01-26	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
28	AFG	Asia	Afghanistan	2020-01-27	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
29	AFG	Asia	Afghanistan	2020-01-28	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
30	AFG	Asia	Afghanistan	2020-01-29	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
31	AFG	Asia	Afghanistan	2020-01-30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

The status bar at the bottom indicates: Query executed successfully. DESKTOP-S7MQMV4\SS2019 (15... DESKTOP-S7MQMV4\Shivam... Week 03 00:00:01 49,227 rows