TIETOEVRY – DATABASE – FINAL ASSESSMENT

SWEEKAR K T/EI13047

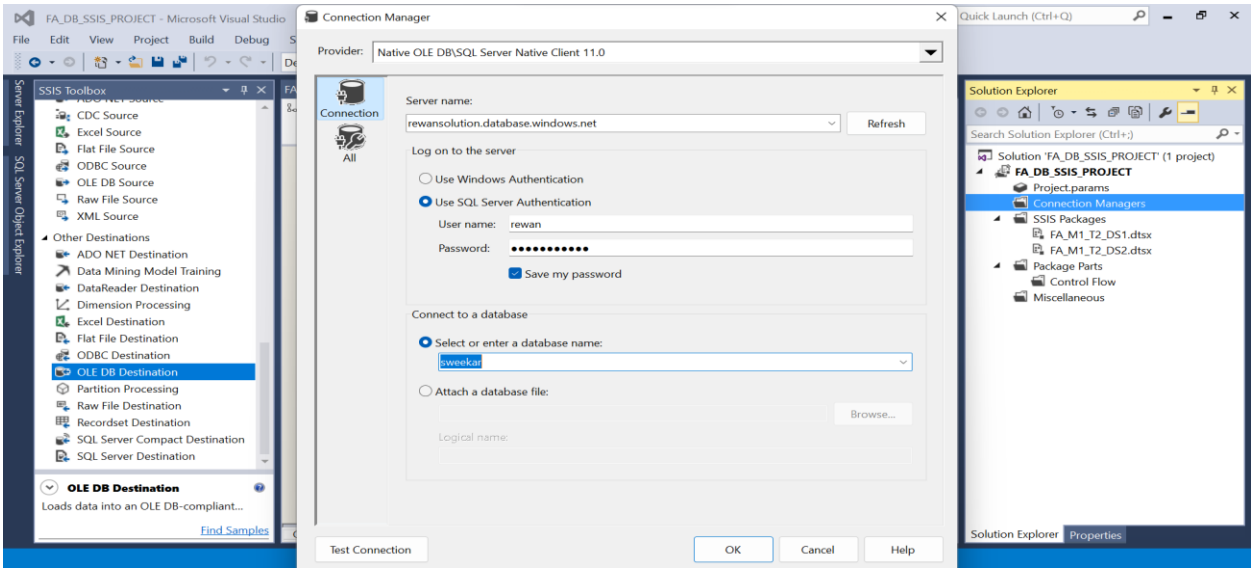
Module 2 SSIS: 20 Points

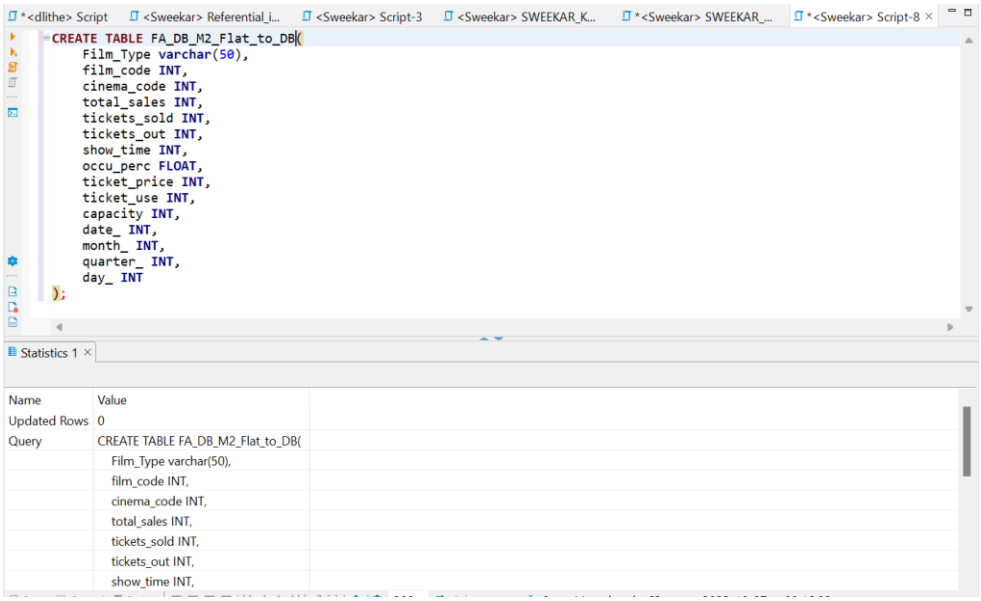
Entry criteria: Transactions tables created in Module 1, task 2

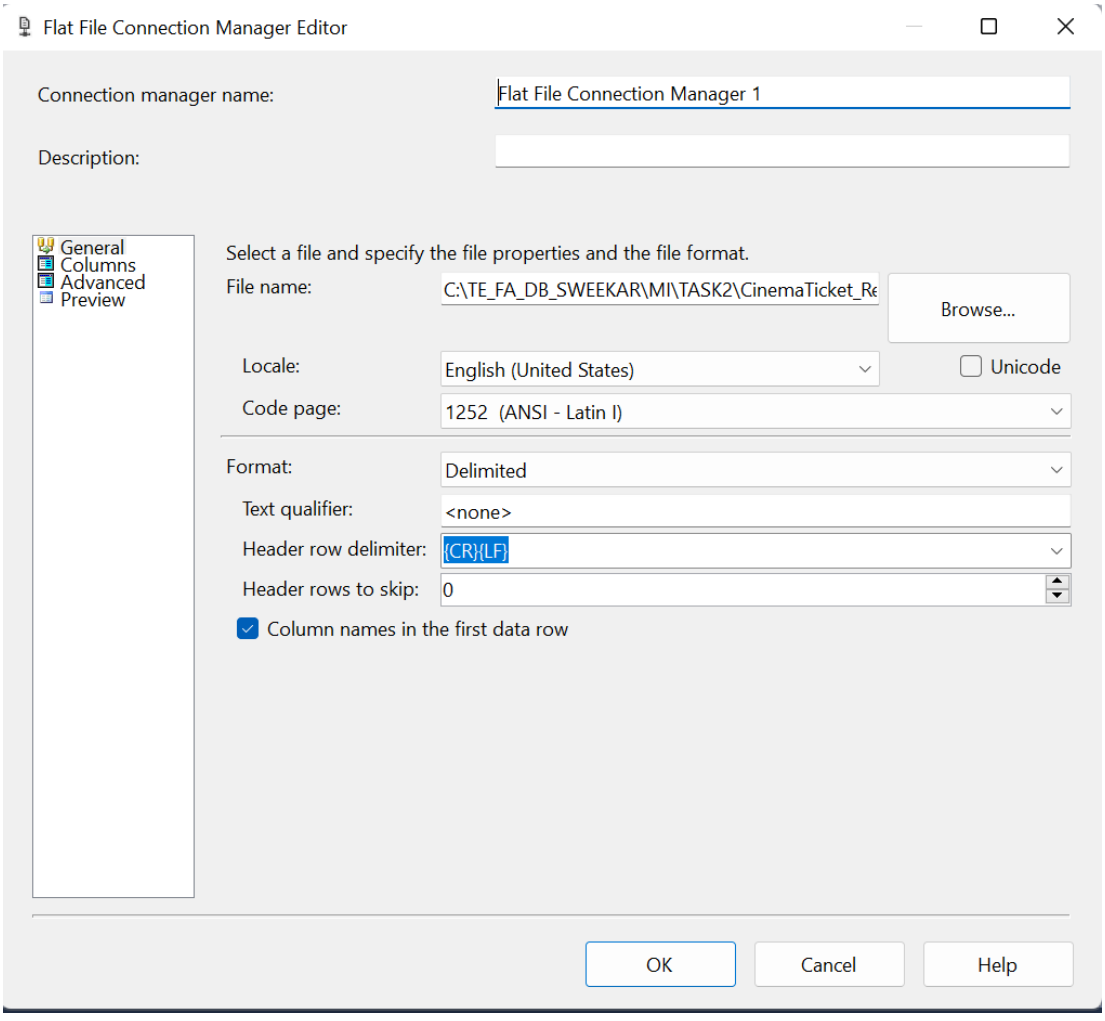
1. **Make connection to different data source (Oracle DB, Flat file, SQL Server)**

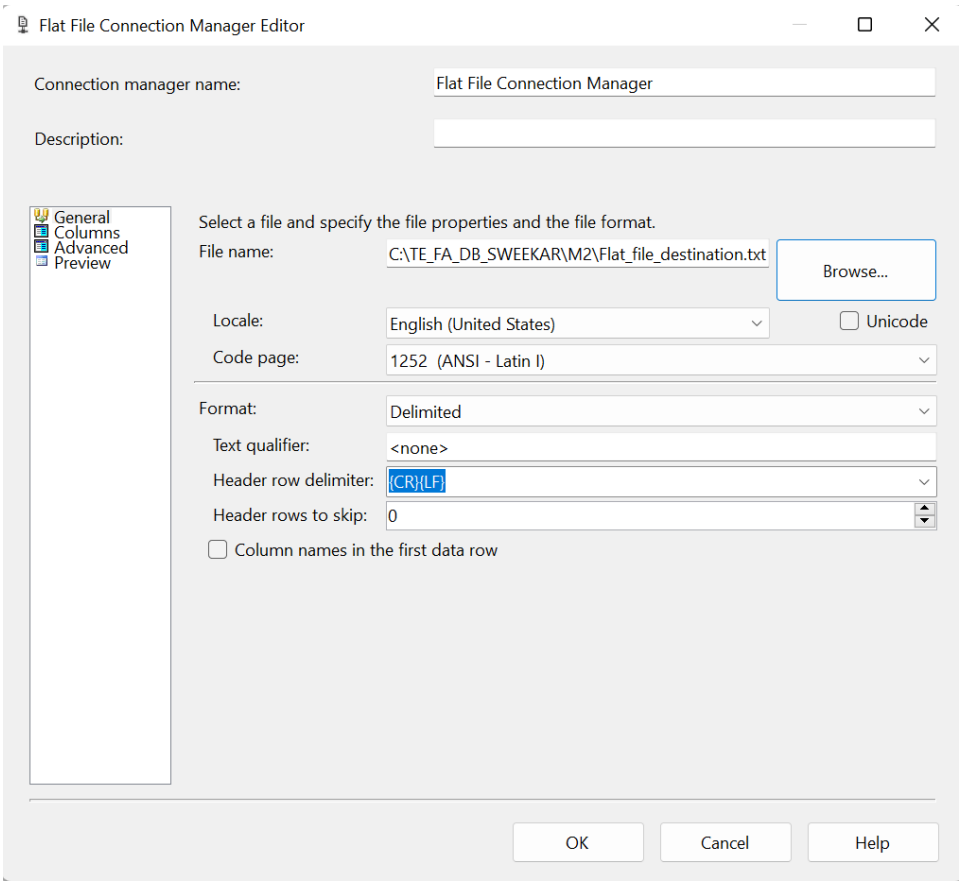
Data Sources: Car Sales and Film

First adding SQL server to oledb connections in SSIS

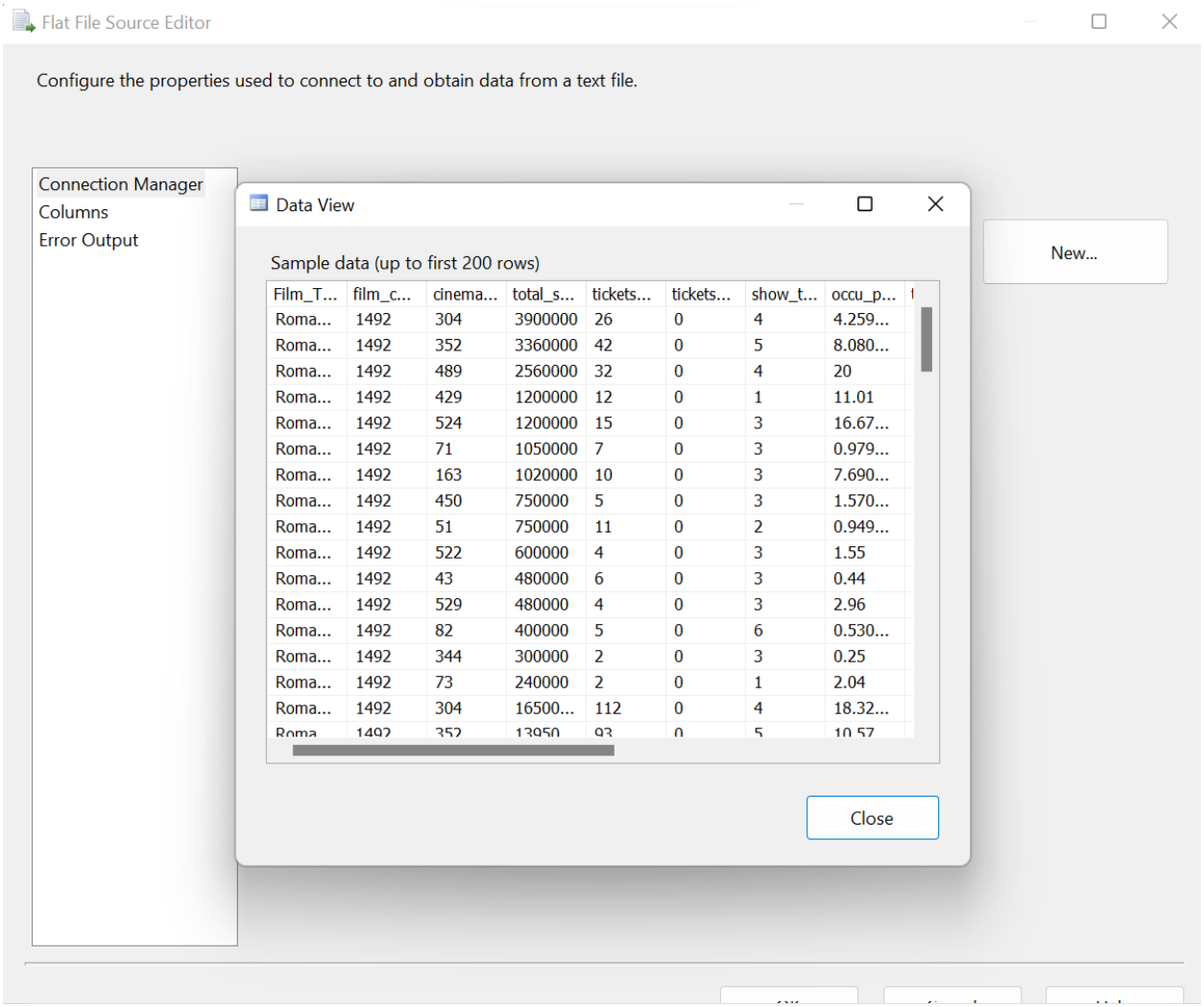



*<dlithe> Script
<Sweekar> Referential_i...
<Sweekar> Script-3
<Sweekar> SWEEKAR_K...
*<Sweekar> SWEEKAR_...
*<Sweekar> Script-8
CREATE TABLE FA_DB_M2_Flat_to_DB(
Film_Type varchar(50),
film_code INT,
cinema_code INT,
total_sales INT,
tickets_sold INT,
tickets_out INT,
show_time INT,
occu_perc FLOAT,
ticket_price INT,
ticket_use INT,
capacity INT,
date_ INT,
month_ INT,
quarter_ INT,
day_ INT
);
Statistics 1
Name
Value
Updated Rows
0
Query
CREATE TABLE FA_DB_M2_Flat_to_DB(
Film_Type varchar(50),
film_code INT,
cinema_code INT,
total_sales INT,
tickets_sold INT,
tickets_out INT,
show_time INT,

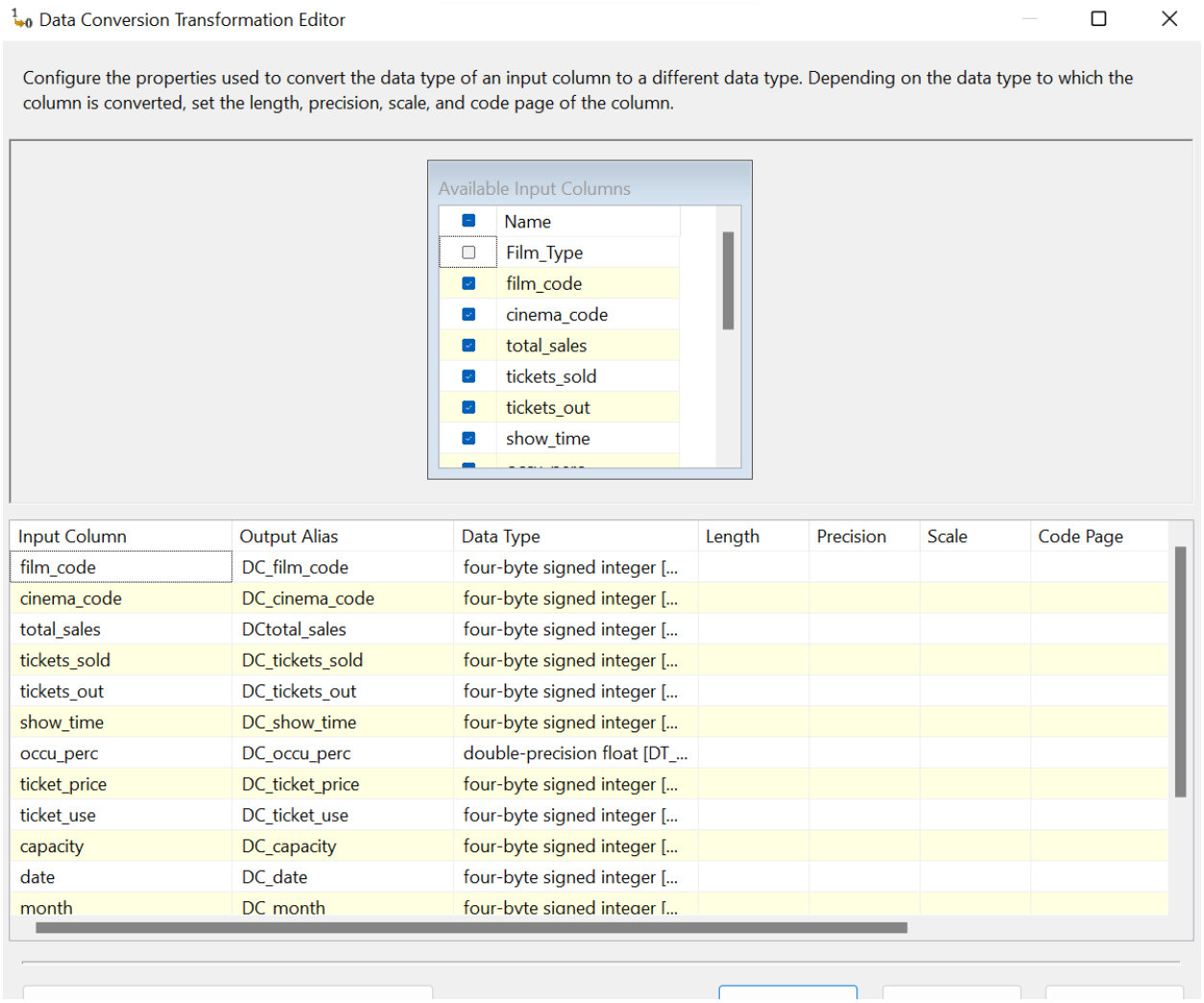

Flat File Connection Manager Editor
Connection manager name:
Flat File Connection Manager 1
Description:
General
Columns
Advanced
Preview
Select a file and specify the file properties and the file format.
File name:
C:\TE_FA_DB_SWEEKAR\MI\TASK2\CinemaTicket_Re
Browse...
Locale:
English (United States)
Unicode
Code page:
1252 (ANSI - Latin I)
Format:
Delimited
Text qualifier:
<none>
Header row delimiter:
{CR}{LF}
Header rows to skip:
0
Column names in the first data row
OK
Cancel
Help


Flat File Connection Manager Editor
Connection manager name:
Flat File Connection Manager
Description:
General
Columns
Advanced
Preview
Select a file and specify the file properties and the file format.
File name:
C:\TE_FA_DB_SWEEKAR\M2\Flat_file_destination.txt
Browse...
Locale:
English (United States)
Unicode
Code page:
1252 (ANSI - Latin I)
Format:
Delimited
Text qualifier:
<none>
Header row delimiter:
{CR}{LF}
Header rows to skip:
0
Column names in the first data row
OK
Cancel
Help

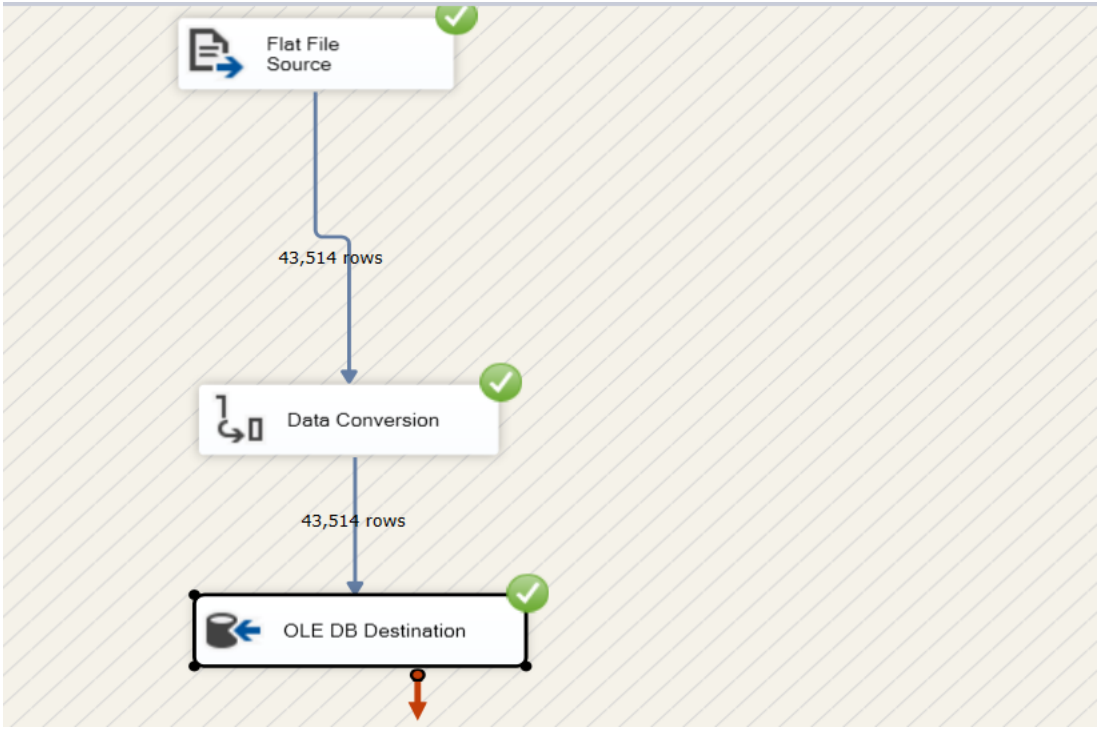
2. Extract


Flat File Source Editor
Configure the properties used to connect to and obtain data from a text file.
Connection Manager
Columns
Error Output
Data View
Sample data (up to first 200 rows)
Film_T... film_c... cinema... total_s... tickets... tickets... show_t... occu_p...
Roma... 1492 304 3900000 26 0 4 4.259...
Roma... 1492 352 3360000 42 0 5 8.080...
Roma... 1492 489 2560000 32 0 4 20
Roma... 1492 429 1200000 12 0 1 11.01
Roma... 1492 524 1200000 15 0 3 16.67...
Roma... 1492 71 1050000 7 0 3 0.979...
Roma... 1492 163 1020000 10 0 3 7.690...
Roma... 1492 450 750000 5 0 3 1.570...
Roma... 1492 51 750000 11 0 2 0.949...
Roma... 1492 522 600000 4 0 3 1.55
Roma... 1492 43 480000 6 0 3 0.44
Roma... 1492 529 480000 4 0 3 2.96
Roma... 1492 82 400000 5 0 6 0.530...
Roma... 1492 344 300000 2 0 3 0.25
Roma... 1492 73 240000 2 0 1 2.04
Roma... 1492 304 16500... 112 0 4 18.32...
Roma 1492 352 13950 93 0 5 10.57
Close
New...

3. Transform


Data Conversion Transformation Editor
Configure the properties used to convert the data type of an input column to a different data type. Depending on the data type to which the column is converted, set the length, precision, scale, and code page of the column.
Available Input Columns
Name
Film_Type
film_code
cinema_code
total_sales
tickets_sold
tickets_out
show_time
Input Column
Output Alias
Data Type
Length
Precision
Scale
Code Page
film_code
DC_film_code
four-byte signed integer [...
cinema_code
DC_cinema_code
four-byte signed integer [...
total_sales
DCtotal_sales
four-byte signed integer [...
tickets_sold
DC_tickets_sold
four-byte signed integer [...
tickets_out
DC_tickets_out
four-byte signed integer [...
show_time
DC_show_time
four-byte signed integer [...
occu_perc
DC_occu_perc
double-precision float [DT_...
ticket_price
DC_ticket_price
four-byte signed integer [...
ticket_use
DC_ticket_use
four-byte signed integer [...
capacity
DC_capacity
four-byte signed integer [...
date
DC_date
four-byte signed integer [...
month
DC_month
four-byte signed integer [...

4. Load to new target system - SQL Server



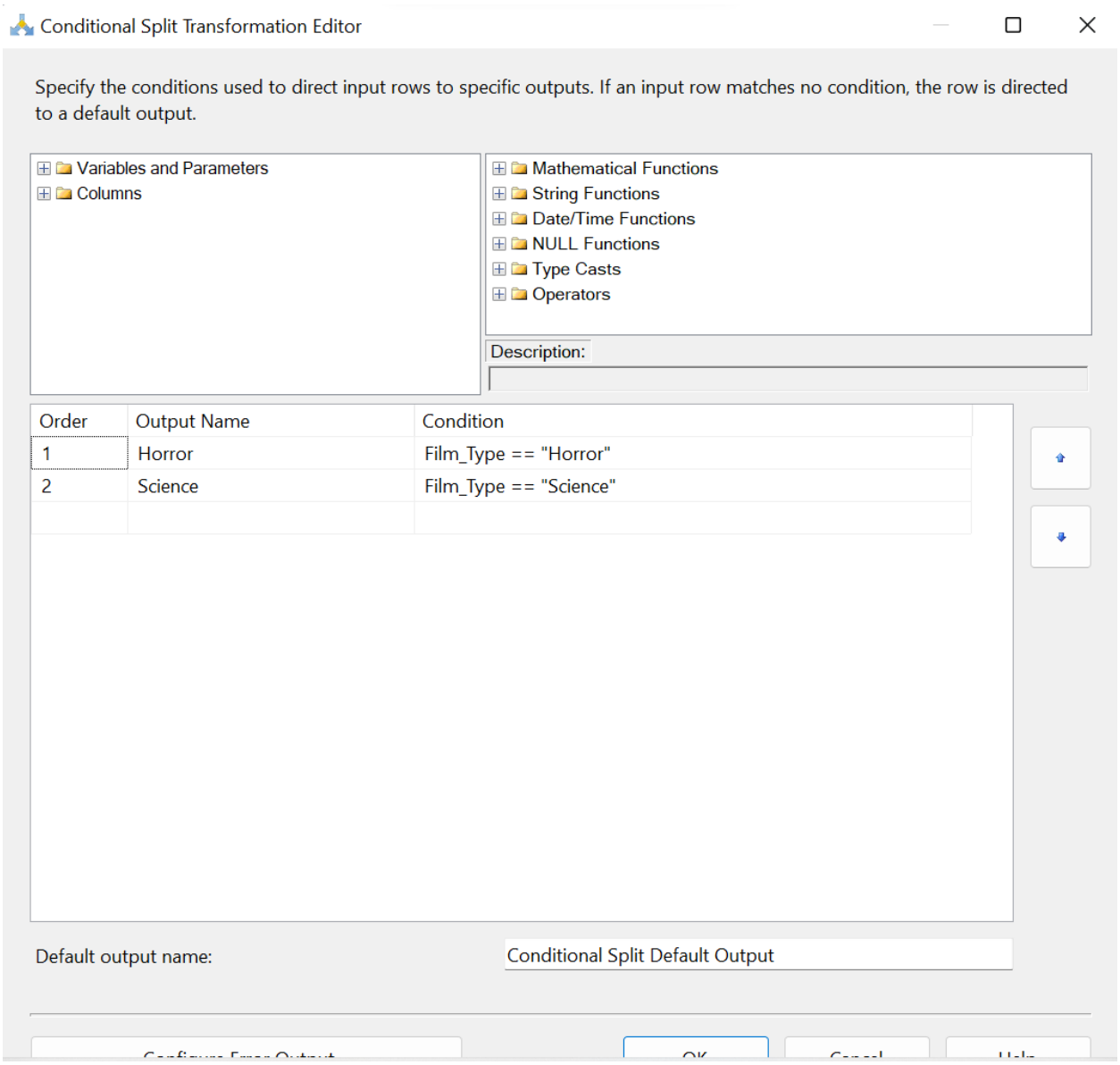

```
SELECT * FROM FA_DB_M2_Flat_to_DB;
```

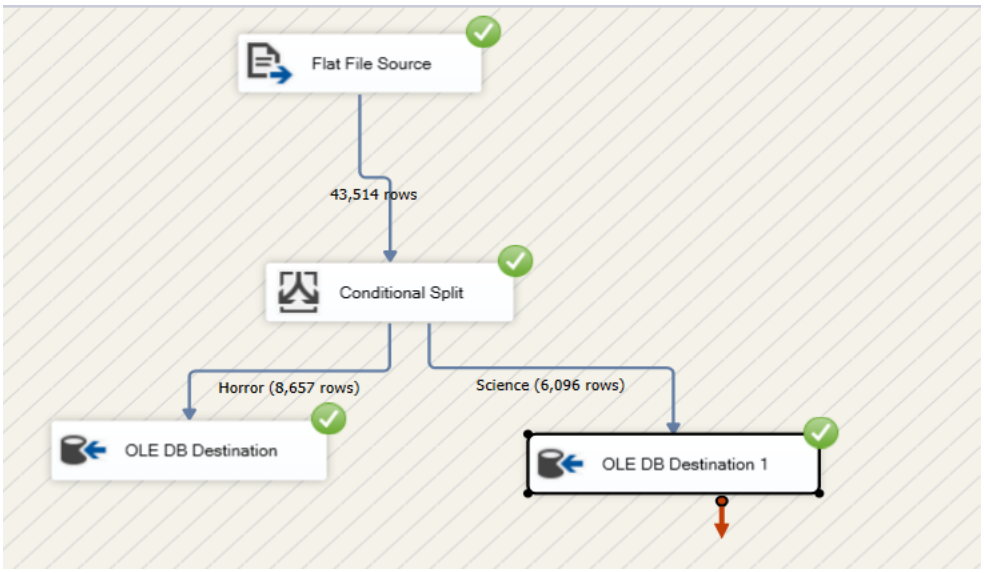
Results 1

SELECT * FROM FA_DB_M2_Flat_to_DB | Enter a SQL expression to filter results (use Ctrl+Space)

	Film_Type	film_code	cinema_code	total_sales	tickets_sold	tickets_out	show_time	occu_perc	ticket_price
1	Romance	1,492	304	3,900,000	26	0	4	4.2599999999999998	150,000
2	Romance	1,492	352	3,360,000	42	0	5	8.0800000000000001	80,000
3	Romance	1,492	489	2,560,000	32	0	4	20	80,000
4	Romance	1,492	429	1,200,000	12	0	1	11.01	100,000
5	Romance	1,492	524	1,200,000	15	0	3	16.670000000000002	80,000
6	Romance	1,492	71	1,050,000	7	0	3	0.97999999999999998	150,000
7	Romance	1,492	163	1,020,000	10	0	3	7.6900000000000004	102,000
8	Romance	1,492	450	750,000	5	0	3	1.5700000000000001	150,000
9	Romance	1,492	51	750,000	11	0	2	0.94999999999999996	68,181
10	Romance	1,492	522	600,000	4	0	3	1.55	150,000

Here I am creating seperate tables based on Film type column using conditional split, considering that different teams are reviewing the different film types.






```
SELECT * FROM FA_AGG_TABLE_HORROR;
SELECT * FROM FA_AGG_TABLE_SCIENCE;
```

Results 1 | Results 1 (2)

SELECT * FROM FA_AGG_TABLE_HORROR; SI — Data filter is not supported

	Film_Type	film_code	cinema_code	total_sales	tickets_sold	tickets_out	show_time	occu_perc	ticket_price
1	Horror	1484	448	121350000	814	0	10	30.969999999999999	149078.62409999999
2	Horror	1484	304	49920000	336	0	10	15.539999999999999	148571.42860000001
3	Horror	1484	186	47440000	593	0	7	16.739999999999998	80000
4	Horror	1484	163	46440000	487	0	7	40.920000000000002	95359.342919999996
5	Horror	1484	518	41220000	565	4	10	53.100000000000001	72955.752210000006
6	Horror	1484	183	38160000	477	0	7	15.140000000000001	80000
7	Horror	1484	141	31650000	221	0	6	20.969999999999999	143212.6697
8	Horror	1484	34	30200000	302	0	5	45.07	100000
9	Horror	1484	51	26450000	305	0	12	9.3599999999999994	86721.311480000004
10	Horror	1484	450	25200000	176	0	5	27.079999999999998	143181.81820000001

```
SELECT * FROM FA_AGG_TABLE_HORROR;
SELECT * FROM FA_AGG_TABLE_SCIENCE;
```

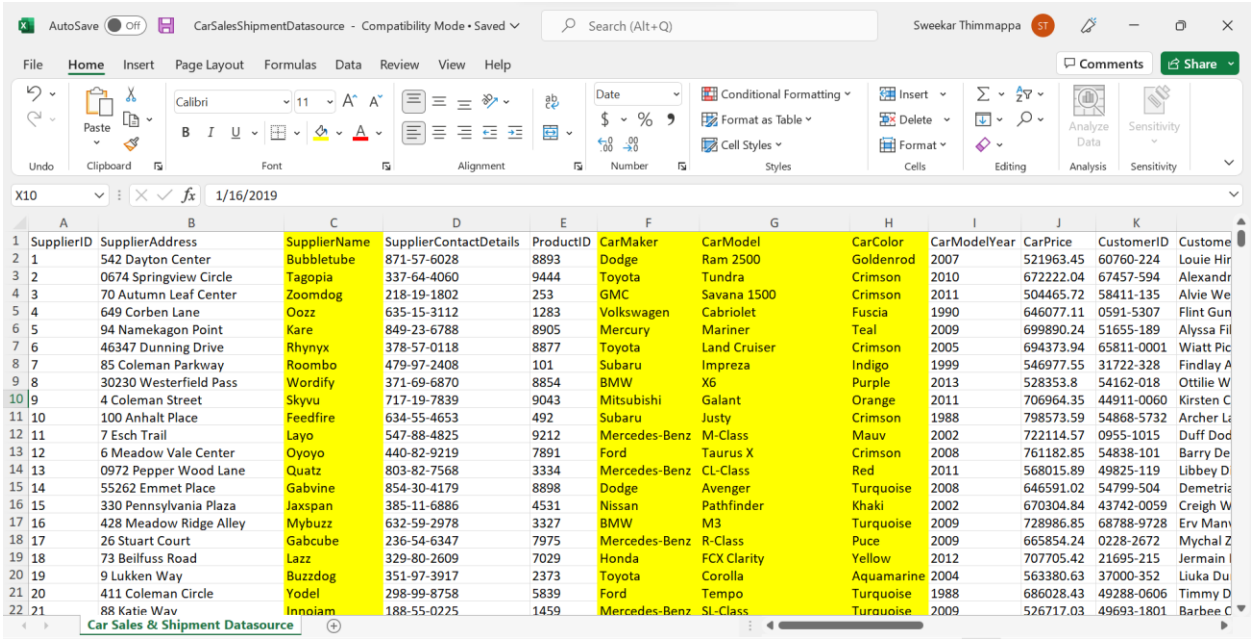
Results 1 | Results 1 (2)

SELECT * FROM FA_AGG_TABLE_HORROR; SI — Data filter is not supported

	Film_Type	film_code	cinema_code	total_sales	tickets_sold	tickets_out	show_time	occu_perc	ticket_price
1	Science	1485	448	132449998	1656	0	10	87.159999999999997	79981.882849999995
2	Science	1485	304	58730000	707	0	7	61.799999999999997	83069.306930000006
3	Science	1485	352	57360000	724	0	5	70.159999999999997	79226.519339999999
4	Science	1485	71	51280000	643	1	3	66.359999999999999	79751.166410000005
5	Science	1485	344	50480000	636	0	5	48.07	79371.069180000006
6	Science	1485	489	47840000	608	0	16	64.170000000000002	78684.210529999997
7	Science	1485	72	46340000	775	0	15	56.710000000000001	59793.548390000004
8	Science	1485	39	43050000	865	0	3	96.650000000000006	49768.78613
9	Science	1485	452	42800000	535	0	5	73.790000000000006	80000
10	Science	1485	513	41820000	697	0	3	93.810000000000002	60000

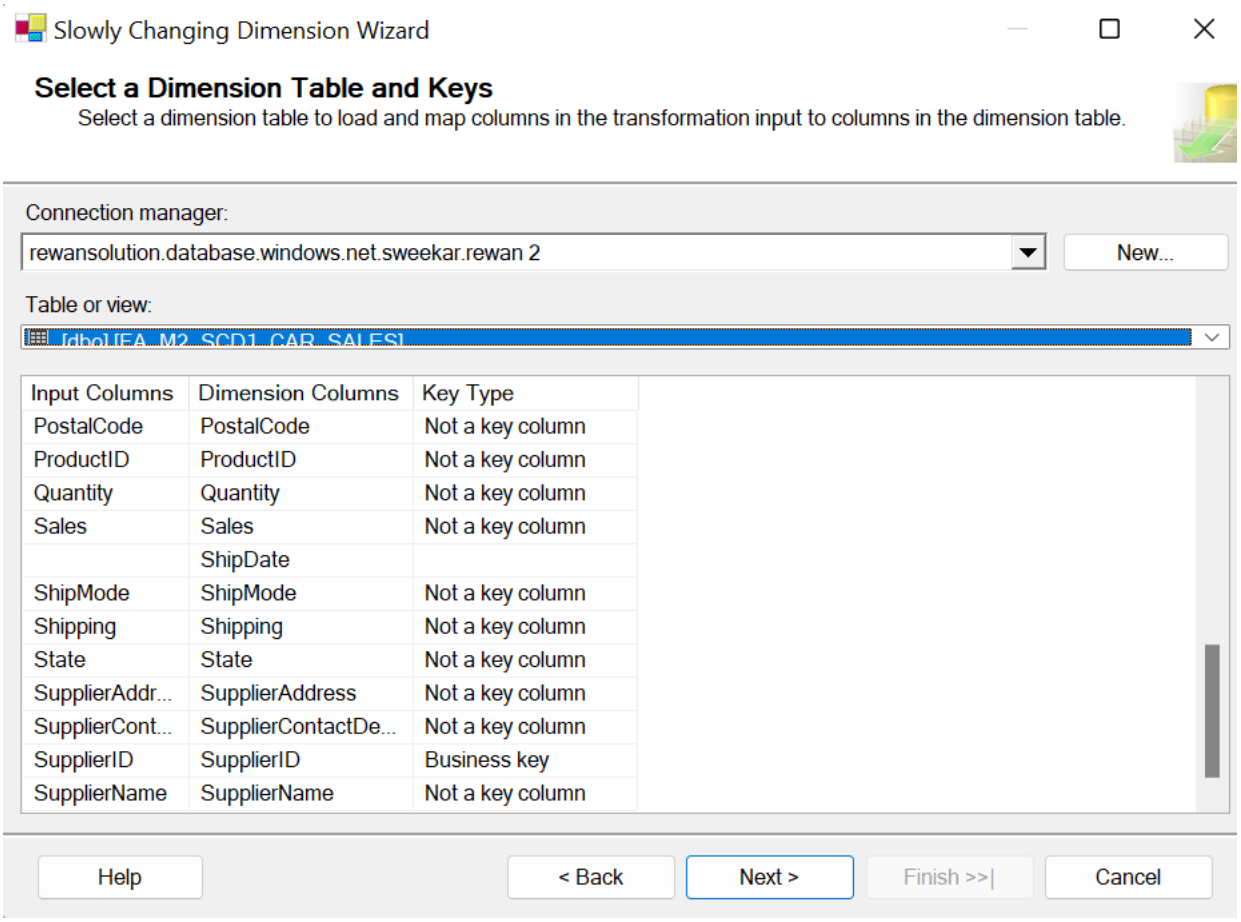
5. Perform SCD 1 & SCD2 dimension table modelling

SCD TYPE 1

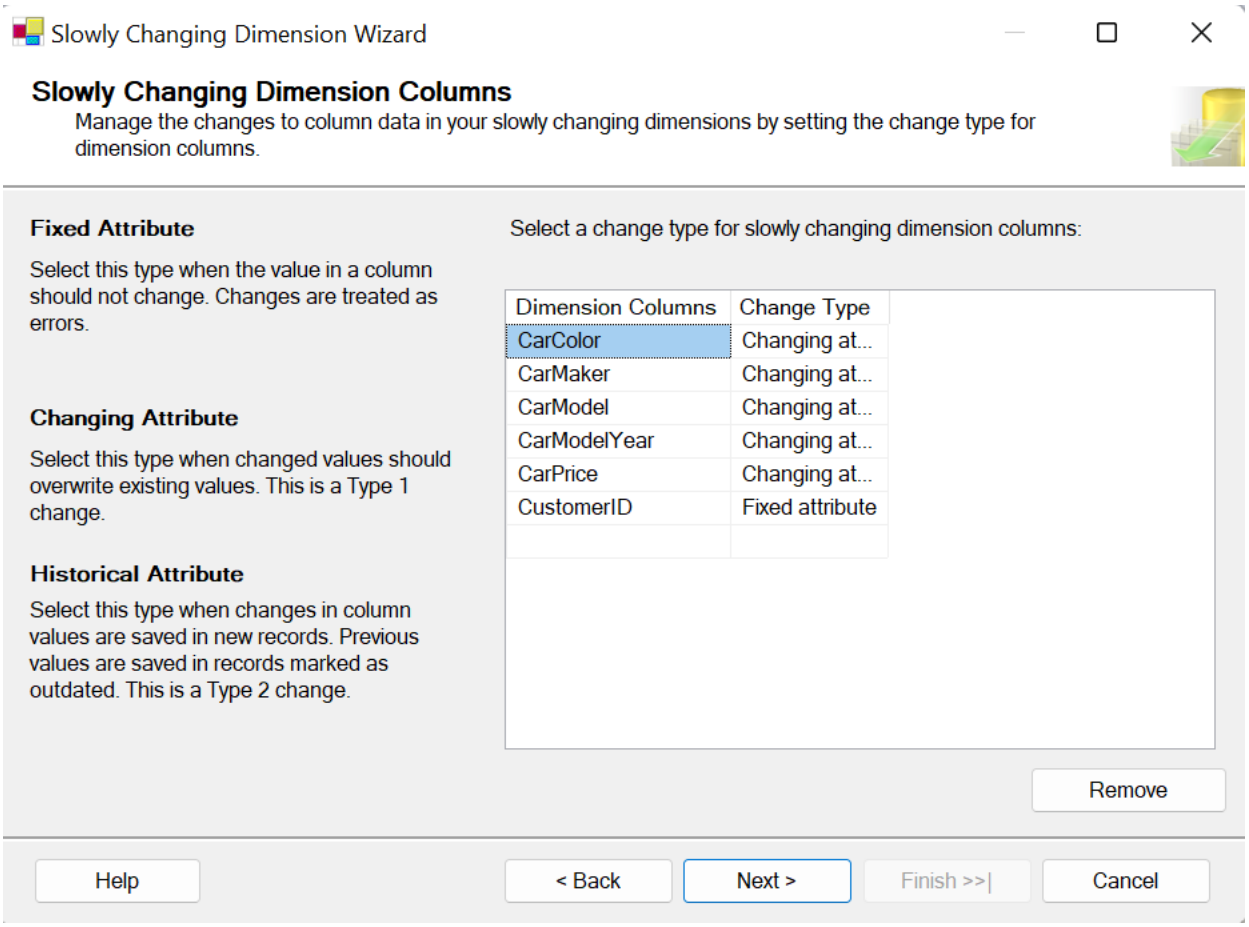


	A	B	C	D	E	F	G	H	I	J	K	
1	SupplierID	SupplierAddress	SupplierName	SupplierContactDetails	ProductID	CarMaker	CarModel	CarColor	CarModelYear	CarPrice	CustomerID	Custome
2	1	542 Dayton Center	Bubbletube	871-57-6028	8893	Dodge	Ram 2500	Goldenrod	2007	521963.45	60760-224	Louie Hir
3	2	0674 Springview Circle	Tagopia	337-64-4060	9444	Toyota	Tundra	Crimson	2010	672222.04	67457-594	Alexandr
4	3	70 Autumn Leaf Center	Zoomdog	218-19-1802	253	GMC	Savana 1500	Crimson	2011	504465.72	58411-135	Alvie We
5	4	649 Corben Lane	Oozz	635-15-3112	1283	Volkswagen	Cabriolet	Fuscia	1990	646077.11	0591-5307	Flint Gun
6	5	94 Namekagon Point	Kare	849-23-6788	8905	Mercury	Mariner	Teal	2009	699890.24	51655-189	Alyssa Fil
7	6	46347 Dunning Drive	Rhynyx	378-57-0118	8877	Toyota	Land Cruiser	Crimson	2005	694373.94	65811-0001	Wiatt Pic
8	7	85 Coleman Parkway	Roombo	479-97-2408	101	Subaru	Impreza	Indigo	1999	546977.55	31722-328	Findlay A
9	8	30230 Westerfield Pass	Wordify	371-69-6870	8854	BMW	X6	Purple	2013	528353.8	54162-018	Ottilie W
10	9	4 Coleman Street	Skyvu	717-19-7839	9043	Mitsubishi	Galant	Orange	2011	706964.35	44911-0060	Kirsten C
11	10	100 Anhalt Place	Feedfire	634-55-4653	492	Subaru	Justy	Crimson	1988	798573.59	54868-5732	Archer La
12	11	7 Esch Trail	Layo	547-88-4825	9212	Mercedes-Benz	M-Class	Mauv	2002	722114.57	0955-1015	Duff Dod
13	12	6 Meadow Vale Center	Oyoyo	440-82-9219	7891	Ford	Taurus X	Crimson	2008	761182.85	54838-101	Barry De
14	13	0972 Pepper Wood Lane	Quatz	803-82-7568	3334	Mercedes-Benz	CL-Class	Red	2011	568015.89	49825-119	Libbey D
15	14	55262 Emmet Place	Gabvine	854-30-4179	8898	Dodge	Avenger	Turquoise	2008	646591.02	54799-504	Demetria
16	15	330 Pennsylvania Plaza	Jaxspan	385-11-6886	4531	Nissan	Pathfinder	Khaki	2002	670304.84	43742-0059	Creigh W
17	16	428 Meadow Ridge Alley	Mybuzz	632-59-2978	3327	BMW	M3	Turquoise	2009	728986.85	68788-9728	Erv Manv
18	17	26 Stuart Court	Gabcube	236-54-6347	7975	Mercedes-Benz	R-Class	Puce	2009	665854.24	0228-2672	Mychal Z
19	18	73 Beilfuss Road	Lazz	329-80-2609	7029	Honda	FCX Clarity	Yellow	2012	707705.42	21695-215	Jermain I
20	19	9 Lukken Way	Buzzdog	351-97-3917	2373	Toyota	Corolla	Aquamarine	2004	563380.63	37000-352	Liuka Du
21	20	411 Coleman Circle	Yodel	298-99-8758	5839	Ford	Tempo	Turquoise	1988	686028.43	49288-0606	Timmy D
22	21	88 Katie Way	Innojam	188-55-0225	1459	Mercedes-Benz	SL-Class	Turquoise	2009	526717.03	49693-1801	Barbee C

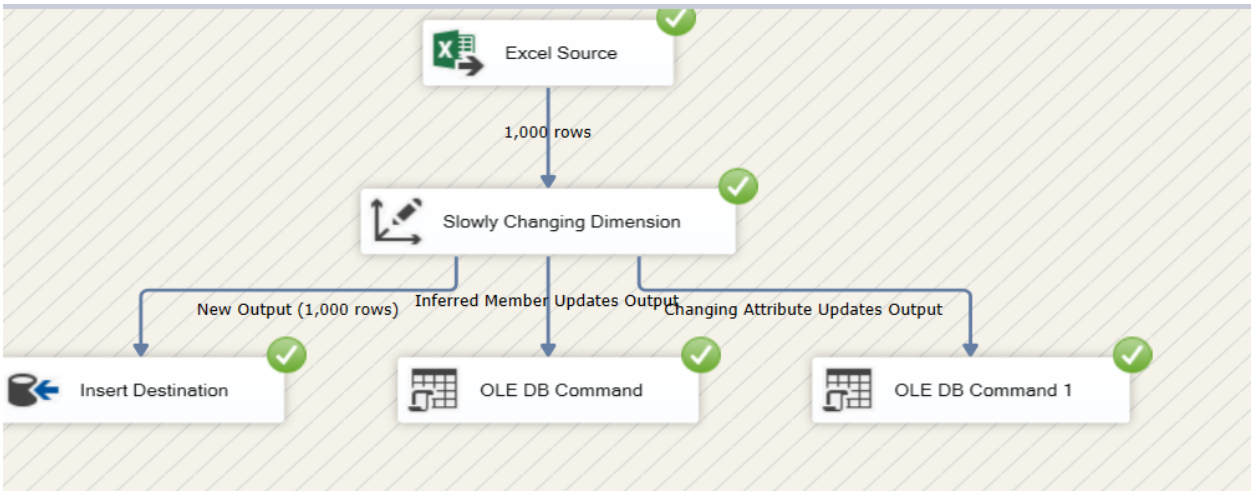
I have specified Supplier ID as the business key. Therefore, if the supplier ID changes then that row gets inserted into the destination.

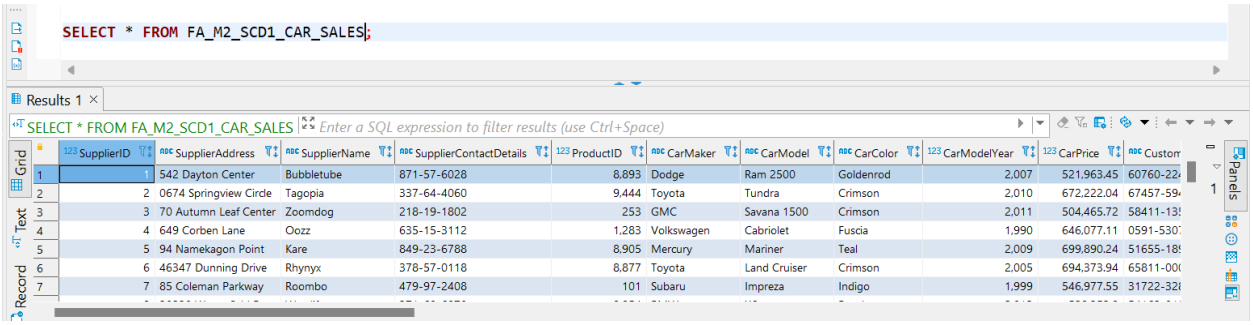


I have given the columns in the below picture as changing attributed and if any of those values change for the existing supplier ID then the change gets updated.

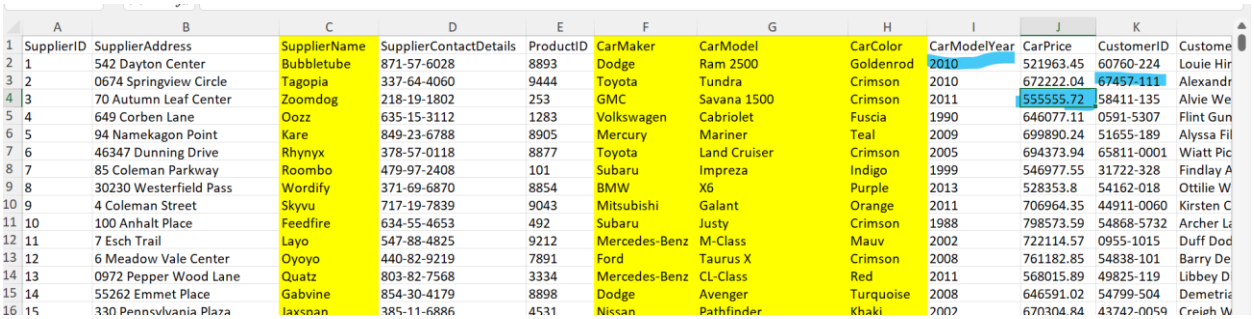


As we can see below all the 1000 rows are transferred to the destination

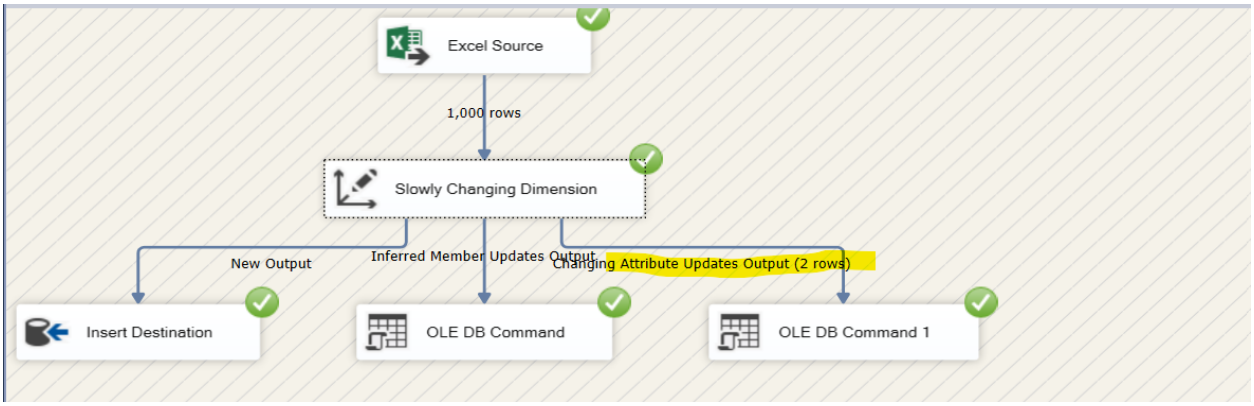


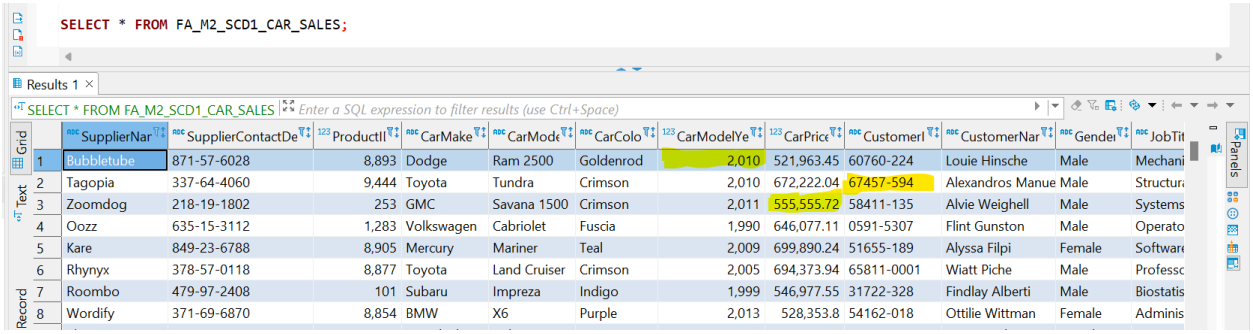


Now I'll make some changes in the source

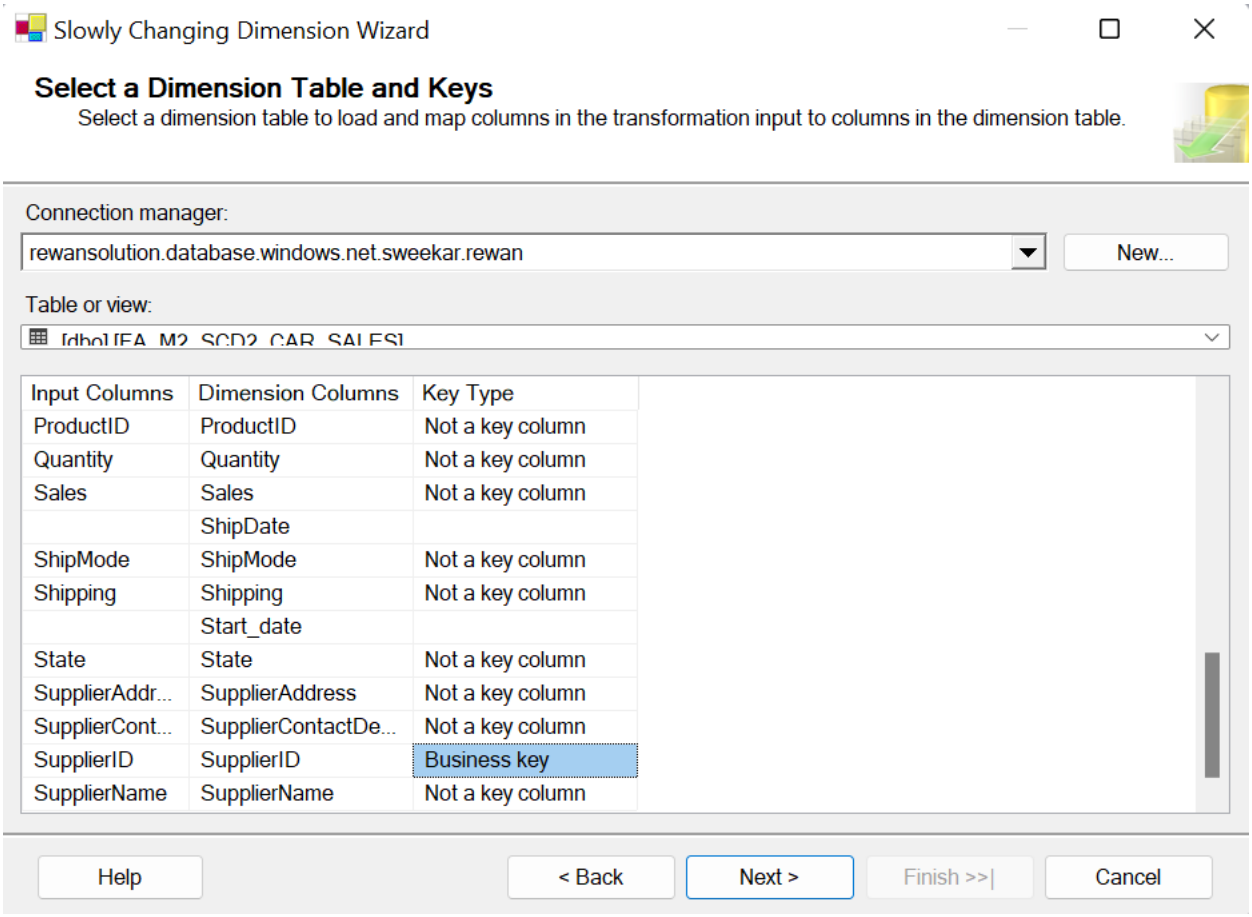


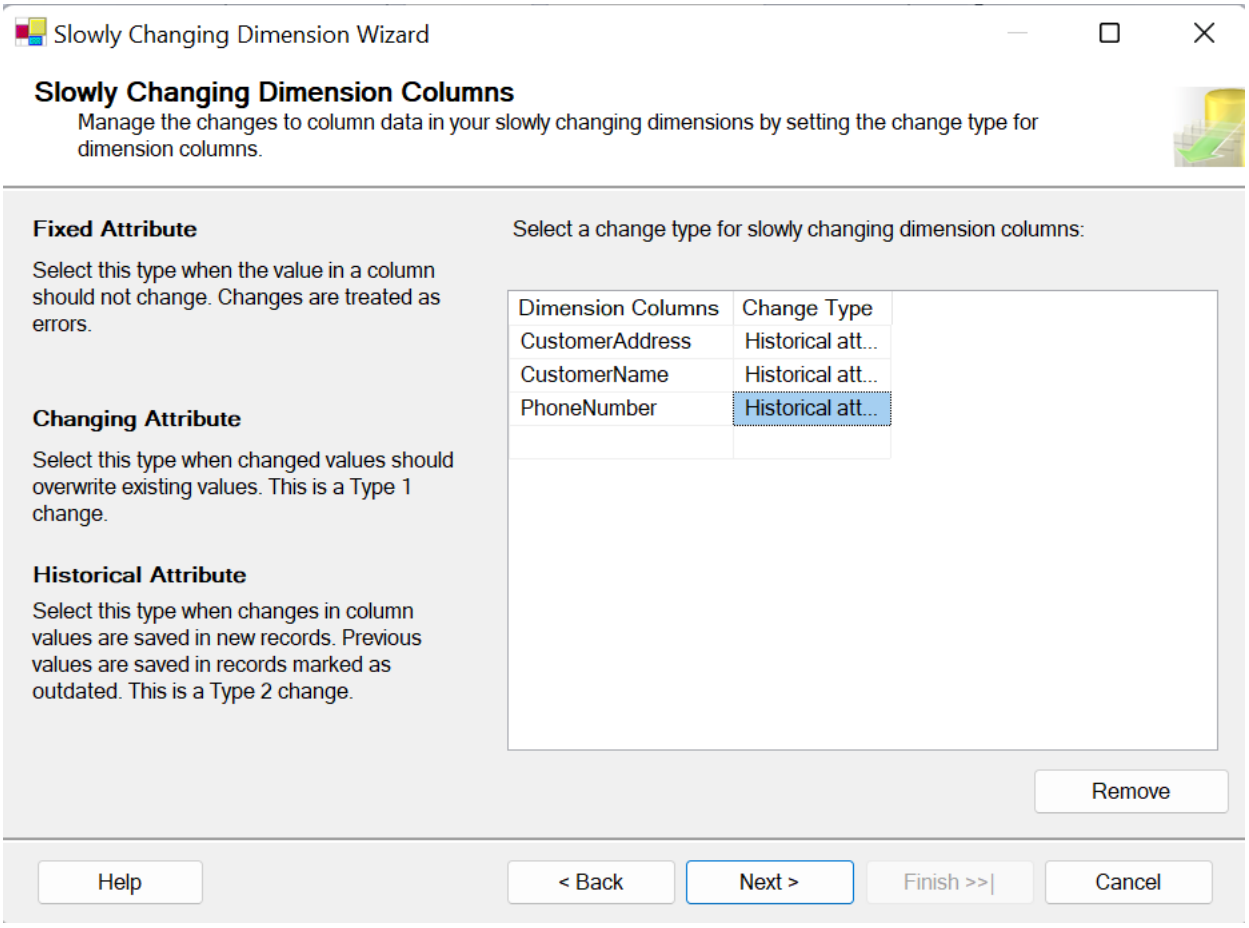
I had made changes in 3 rows but only 2 rows got updated because I had specified that Customer ID as fixed attribute, so it won't get updated.



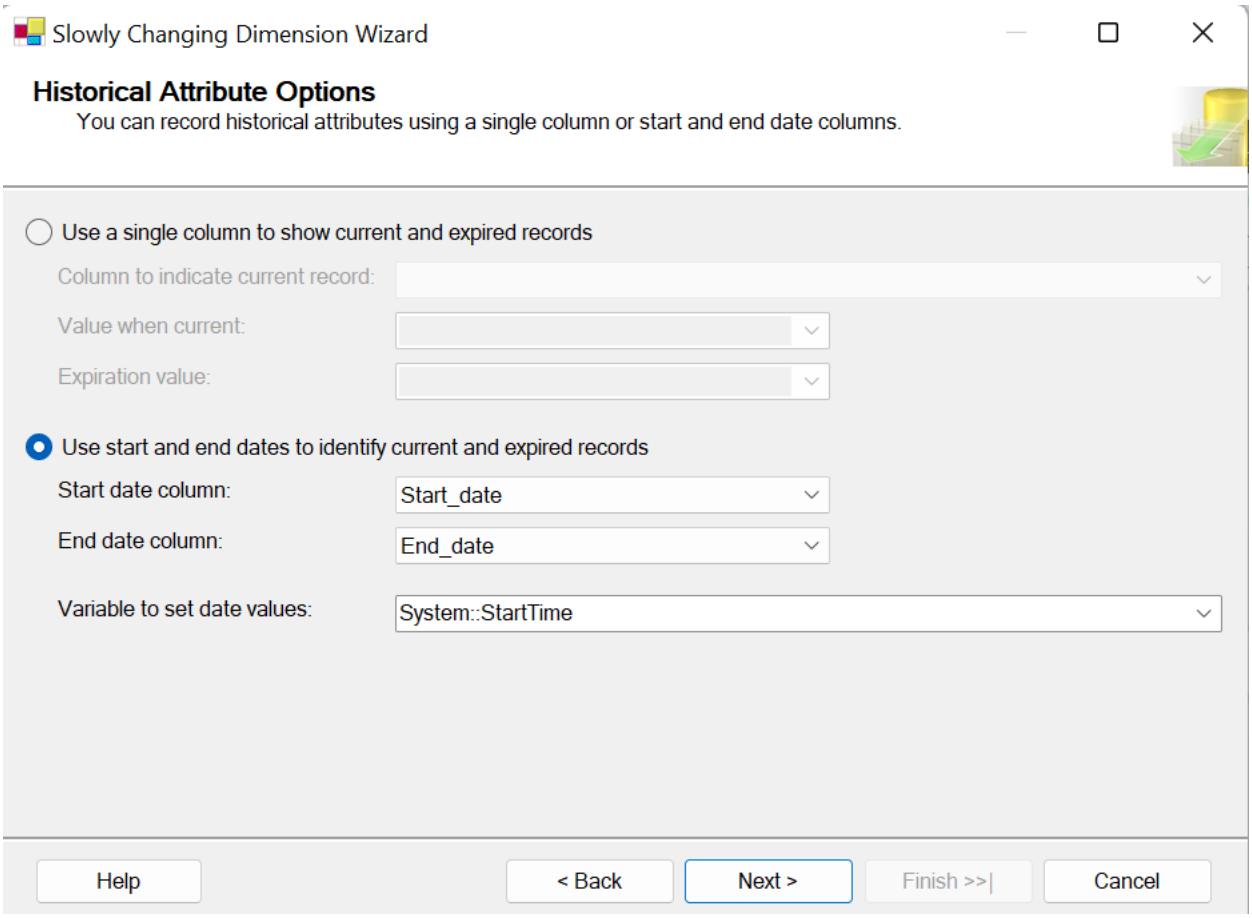


SCD TYPE 2

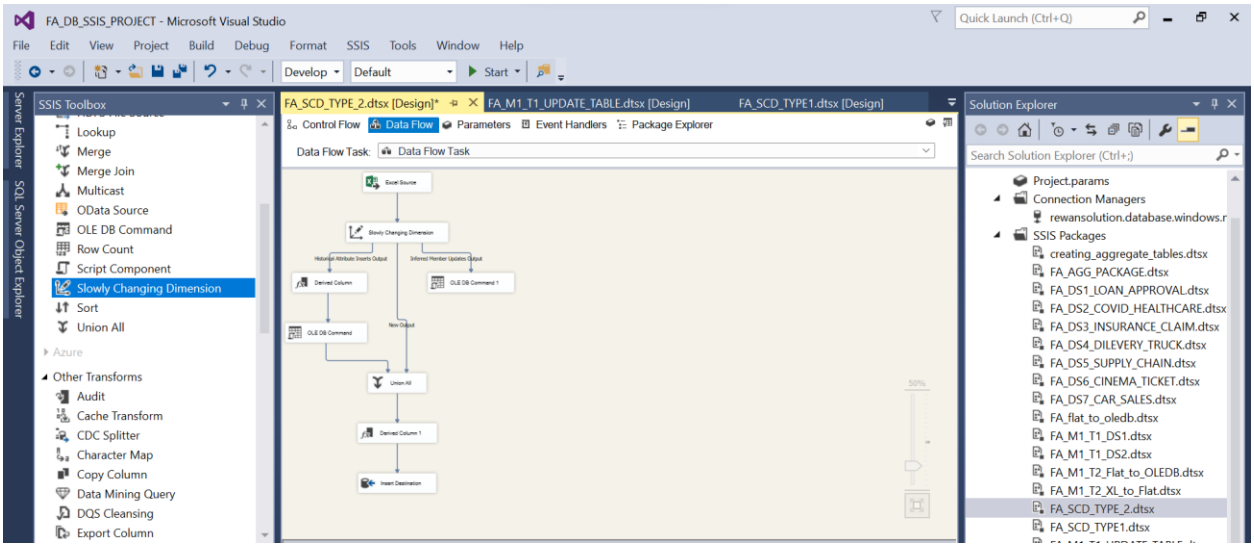

Slowly Changing Dimension Wizard
Select a Dimension Table and Keys
Select a dimension table to load and map columns in the transformation input to columns in the dimension table.
Connection manager:
rewansolution.database.windows.net.sweekar.rewan
New...
Table or view:
[dbo].[FA_M2_SCD2_CAR_SALES]
Input Columns
Dimension Columns
Key Type
ProductID
ProductID
Not a key column
Quantity
Quantity
Not a key column
Sales
Sales
Not a key column
ShipDate
ShipMode
ShipMode
Not a key column
Shipping
Shipping
Not a key column
Start_date
State
State
Not a key column
SupplierAddr...
SupplierAddress
Not a key column
SupplierCont...
SupplierContactDe...
Not a key column
SupplierID
SupplierID
Business key
SupplierName
SupplierName
Not a key column
Help
< Back
Next >
Finish >>|
Cancel

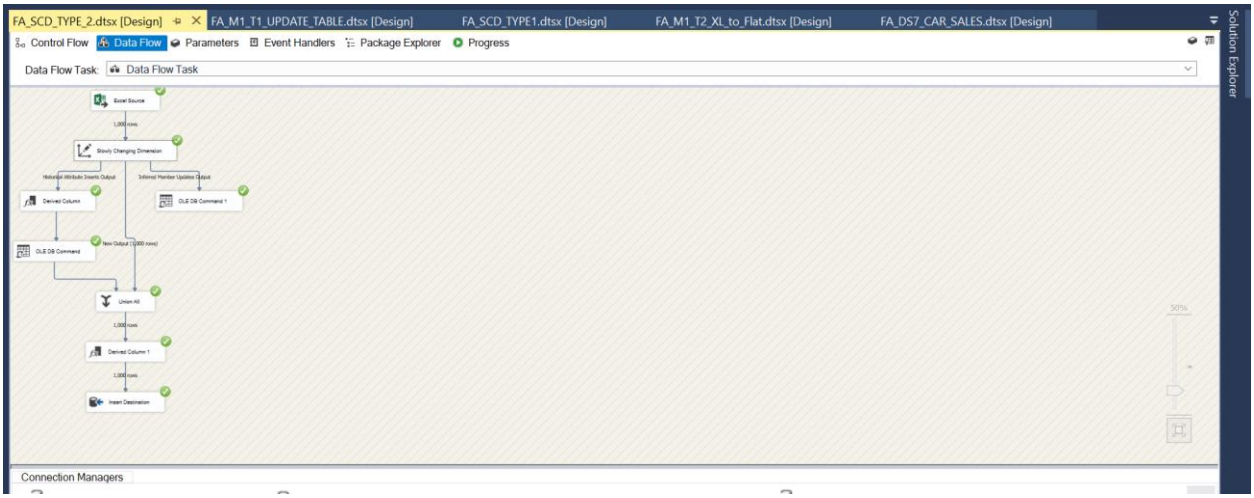

Slowly Changing Dimension Wizard
Slowly Changing Dimension Columns
Manage the changes to column data in your slowly changing dimensions by setting the change type for dimension columns.
Fixed Attribute
Select this type when the value in a column should not change. Changes are treated as errors.
Changing Attribute
Select this type when changed values should overwrite existing values. This is a Type 1 change.
Historical Attribute
Select this type when changes in column values are saved in new records. Previous values are saved in records marked as outdated. This is a Type 2 change.
Select a change type for slowly changing dimension columns:
Dimension Columns
Change Type
CustomerAddress
Historical att...
CustomerName
Historical att...
PhoneNumber
Historical att...
Remove
Help
< Back
Next >
Finish >>|
Cancel

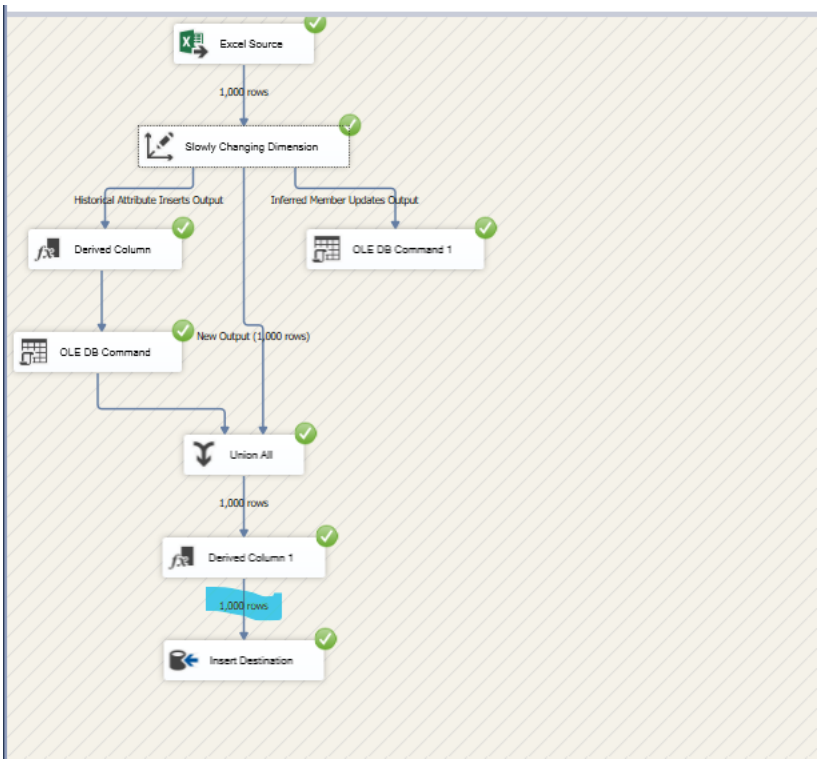
I am using start date and time and end date and time to identify the current and expired records, if the end date is null then the row is active.



The slowly changing dimension package builds itself according to our given inputs







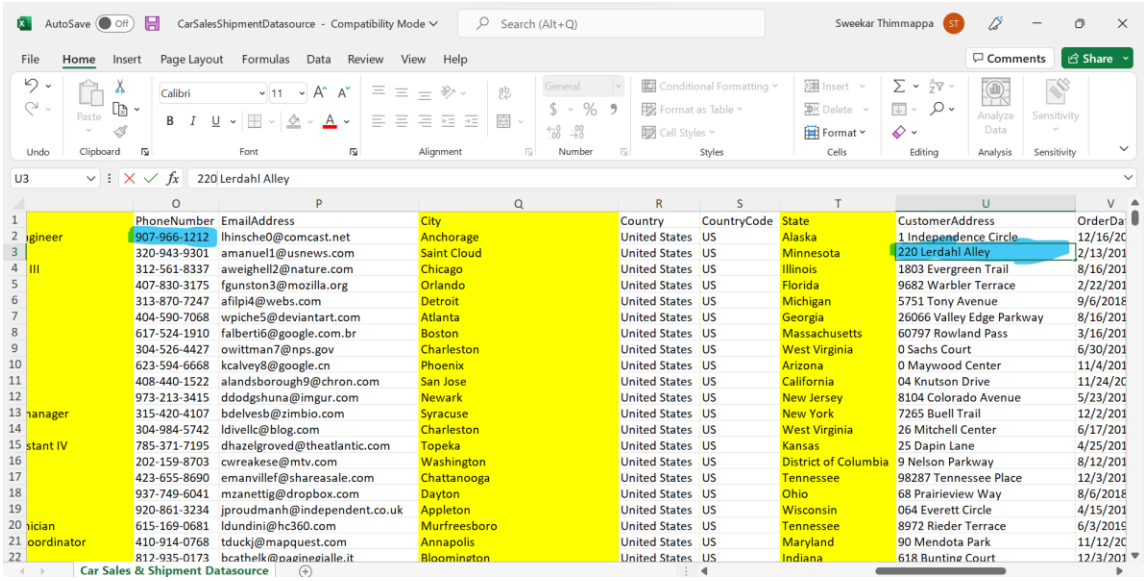

```
);

SELECT * FROM FA_M2_SCD2_CAR_SALES;
```

Results 1

SELECT * FROM FA_M2_SCD2_CAR_SALES — Enter a SQL expression to filter results (use Ctrl+Space)

	Shipping	PostalCode	Sales	Quantity	Discount	CreditCardType	CreditCard	CustomerFeedback	Start_date	End_date
1	uck	99,522	744,796.41	1	0.83	diners-club-carte-blan	408,000,000,000	Bad	2022-10-10 08:05:46.000	[NULL]
2	uck	56,398	794,773.17	1	0.79	jcb	220,000,000,000	Good	2022-10-10 08:05:46.000	[NULL]
3	r	60,674	968,244.9	1	0.28	jcb	160,000,000,000	Okay	2022-10-10 08:05:46.000	[NULL]
4	uck	32,885	942,213.82	2	0.76	jcb	910,000,000,000	Very Bad	2022-10-10 08:05:46.000	[NULL]
5	r	48,232	879,519.57	1	0.5	china-unionpay	240,000,000,000	Bad	2022-10-10 08:05:46.000	[NULL]
6	r	31,136	947,785.4	2	0.39	laser	250,000,000,000	Good	2022-10-10 08:05:46.000	[NULL]
7	uck	2,208	883,836.81	2	0.79	diners-club-carte-blan	528,100,000,000	Very Bad	2022-10-10 08:05:46.000	[NULL]
8	uck	25,336	779,626.29	1	0.9	jcb	140,000,000,000	Bad	2022-10-10 08:05:46.000	[NULL]



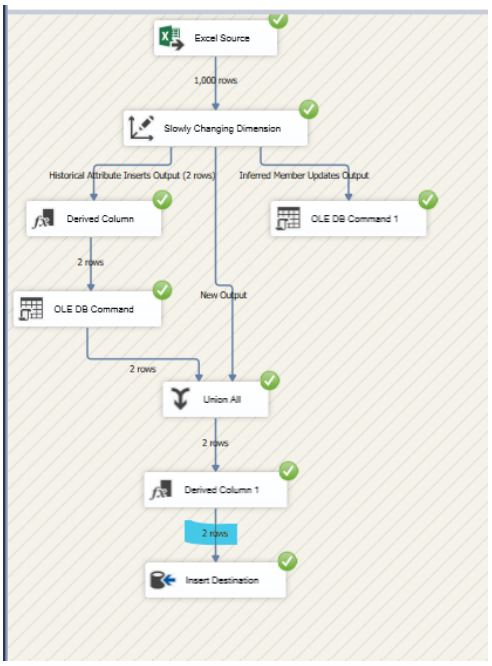
AutoSave Off | CarSalesShipmentDatasource - Compatibility Mode | Search (Alt+Q) | Sweekar Thimmappa

File Home Insert Page Layout Formulas Data Review View Help | Comments | Share

U3 | 220 Lerdahl Alley

	O	P	Q	R	S	T	U	V
1	PhoneNumber	EmailAddress	City	Country	CountryCode	State	CustomerAddress	OrderDa
2	907-966-1212	lhinsche0@comcast.net	Anchorage	United States	US	Alaska	1 Independence Circle	12/16/20
3	320-943-9301	amanuel1@usnews.com	Saint Cloud	United States	US	Minnesota	220 Lerdahl Alley	2/13/201
4	312-561-8337	aweighell2@nature.com	Chicago	United States	US	Illinois	1803 Evergreen Trail	8/16/201
5	407-830-3175	fgunston3@mozilla.org	Orlando	United States	US	Florida	9682 Warbler Terrace	2/22/201
6	313-870-7247	afilpi4@webs.com	Detroit	United States	US	Michigan	5751 Tony Avenue	9/6/2018
7	404-590-7068	wpiche5@deviantart.com	Atlanta	United States	US	Georgia	26066 Valley Edge Parkway	8/16/201
8	617-524-1910	falberti6@google.com.br	Boston	United States	US	Massachusetts	60797 Rowland Pass	3/16/201
9	304-526-4427	owittman7@nps.gov	Charleston	United States	US	West Virginia	0 Sachs Court	6/30/201
10	623-594-6668	kcalvey8@google.cn	Phoenix	United States	US	Arizona	0 Maywood Center	11/4/201
11	408-440-1522	alandsborough9@chron.com	San Jose	United States	US	California	04 Knutson Drive	11/24/20
12	973-213-3415	ddodgshuna@imgur.com	Newark	United States	US	New Jersey	8104 Colorado Avenue	5/23/201
13	315-420-4107	bdelvesb@zimbio.com	Syracuse	United States	US	New York	7265 Buell Trail	12/2/201
14	304-984-5742	ldivellc@blog.com	Charleston	United States	US	West Virginia	26 Mitchell Center	6/17/201
15	785-371-7195	dhazelgroved@theatlantic.com	Topeka	United States	US	Kansas	25 Dapin Lane	4/25/201
16	202-159-8703	cwreakese@mtv.com	Washington	United States	US	District of Columbia	9 Nelson Parkway	8/12/201
17	423-655-8690	emanvillef@shareasale.com	Chattanooga	United States	US	Tennessee	98287 Tennessee Place	12/3/201
18	937-749-6041	mzanettig@dropbox.com	Dayton	United States	US	Ohio	68 Prairieview Way	8/6/2018
19	920-861-3234	jproudmanh@independent.co.uk	Appleton	United States	US	Wisconsin	064 Everett Circle	4/15/201
20	615-169-0681	ldundini@hc360.com	Murfreesboro	United States	US	Tennessee	8972 Rieder Terrace	6/3/2019
21	410-914-0768	tduckj@mapquest.com	Annapolis	United States	US	Maryland	90 Mendota Park	11/12/20
22	812-935-0173	bcathelk@paginegialle.it	Bloomington	United States	US	Indiana	618 Bunting Court	12/3/201

Car Sales & Shipment Datasource




```
SELECT * FROM FA_M2_SCD2_CAR_SALES;
```

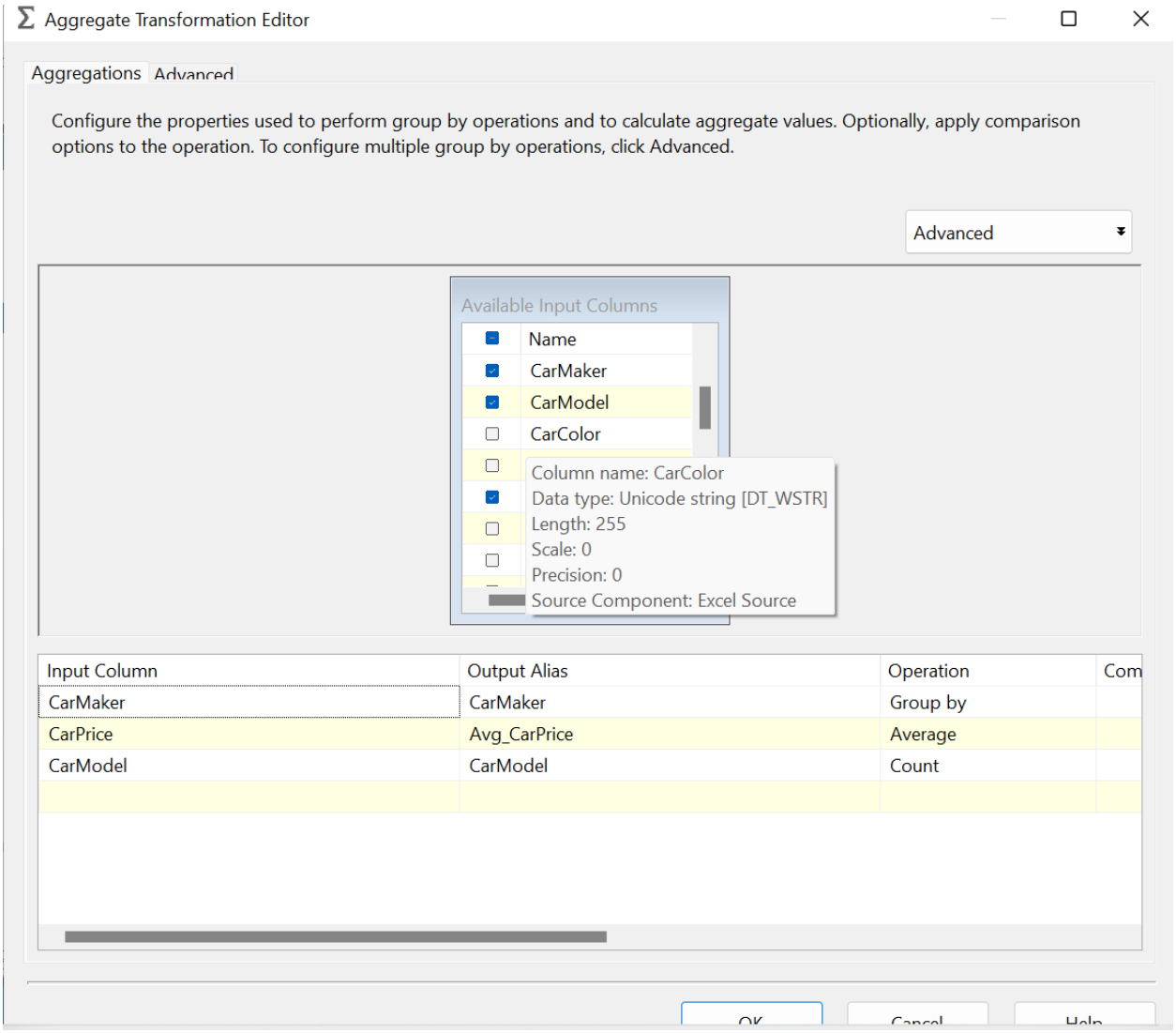
Results 1

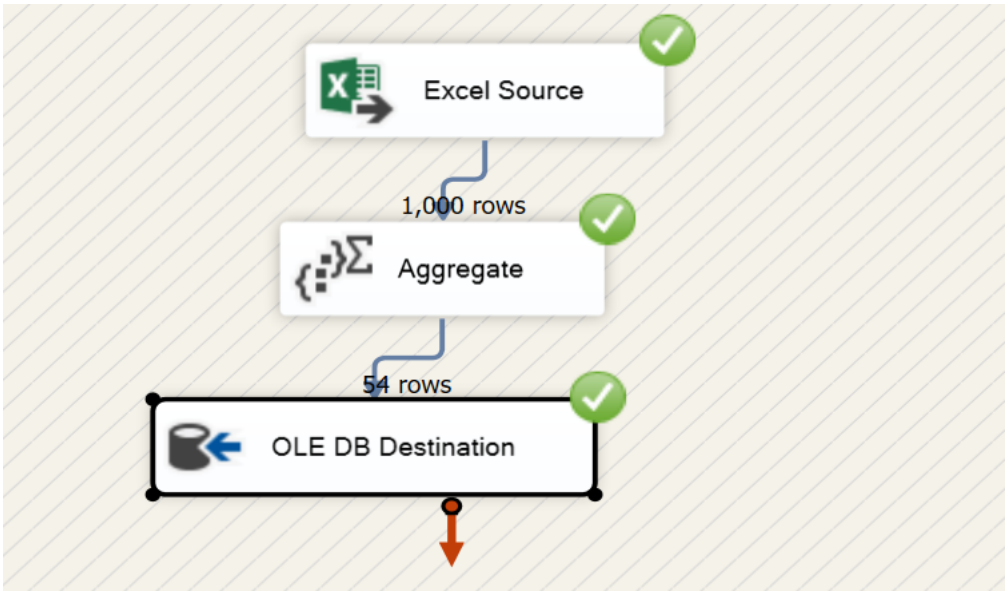
SELECT * FROM FA_M2_SCD2_CAR_SALES | Enter a SQL expression to filter results (use Ctrl+Space)

	Shipping	PostalCode	Sales	Quantity	Discount	CreditCardType	CreditCard	Customer	Start_date	End_date
1	ck	99,522	744,796.41	1	0.83	diners-club-carte-blan	408,000,000,000	Bad	2022-10-10 08:05:46.000	2022-10-10 08:12:39.000
2	ck	56,398	794,773.17	1	0.79	jcb	220,000,000,000	Good	2022-10-10 08:05:46.000	2022-10-10 08:12:39.000
3		60,674	968,244.9	1	0.28	jcb	160,000,000,000	Okay	2022-10-10 08:05:46.000	[NULL]
4	ck	32,885	942,213.82	2	0.76	jcb	910,000,000,000	Very Bad	2022-10-10 08:05:46.000	[NULL]
5		48,232	879,519.57	1	0.5	china-unionpay	240,000,000,000	Bad	2022-10-10 08:05:46.000	[NULL]
6		31,136	947,785.4	2	0.39	laser	250,000,000,000	Good	2022-10-10 08:05:46.000	[NULL]
7	ck	2,208	883,836.81	2	0.79	diners-club-carte-blan	528,100,000,000	Very Bad	2022-10-10 08:05:46.000	[NULL]
8	ck	25,336	779,626.29	1	0.9	jcb	140,000,000,000	Bad	2022-10-10 08:05:46.000	[NULL]

Save | Cancel | Script | 200 | 200+ | Rows: 1 | 200 row(s) fetched - 162ms, on 2022-10-10 at 08:13

6. Create aggregate table based on the particular column (ex: Country code). Refer to the different data source from target system






```
SELECT * FROM FA_AGG_CARMAKER;
```

Results 1 | Statistics 2

SELECT * FROM FA_AGG_CARMAKER — Enter a SQL expression to filter results (use Ctrl+Space)

	CarMaker	Avg_CarPrice	CarModel
1	Mazda	653,824.2518604654	40
2	Honda	666,820.732105263	19
3	Isuzu	670,861.4777777778	18
4	Chevrolet	638,869.9008433735	77
5	Eagle	649,742.5775	4
6	Lexus	614,979.4188888888	18
7	Scion	612,516.7466666667	3
8	Acura	666,380.7090909091	11
9	Saturn	551,125.0900000001	2

Save | Cancel | Script | 200 | 54 | Rows: 1 | 54 row(s) fetched - 165ms, on 2022-10-07 at 15:24:5

IST | en | Writable | Smart Insert | 103 : 31 : 2543 | Sel: 0 | 0