Data Warehousing – IS213

Final Project
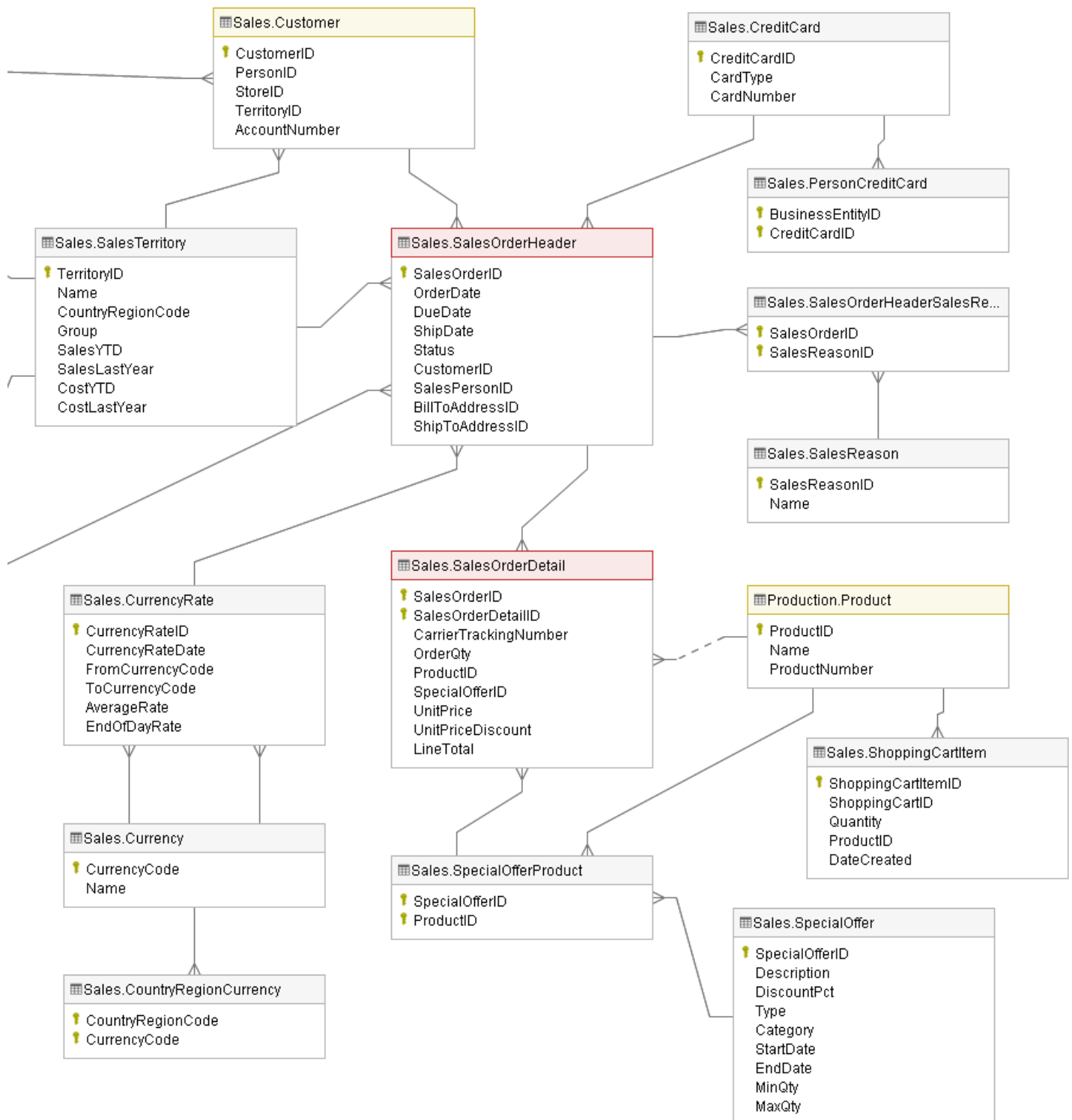
Team Details:

- Seif Yahia   [ID: 20210172]

- Shawky Ebrahim   [ID: 20210184]

- Mostafa Ali   [ID: 20210393]

- Youssef Ezzat   [ID: 20210542]

- Youssef Mahmoud   [ID: 20210486]

Supervised by: Eng. Ahmed Galal

1. Physical model of the source system

AdventureWorks Database Sales module holds information about sales orders, special offers and products.

Sales.Customer
- CustomerID
- PersonID
- StoreID
- TerritoryID
- AccountNumber

Sales.CreditCard
- CreditCardID
- CardType
- CardNumber

Sales.PersonCreditCard
- BusinessEntityID
- CreditCardID

Sales.SalesTerritory
- TerritoryID
- Name
- CountryRegionCode
- Group
- SalesYTD
- SalesLastYear
- CostYTD
- CostLastYear

Sales.SalesOrderHeader
- SalesOrderID
- OrderDate
- DueDate
- ShipDate
- Status
- CustomerID
- SalesPersonID
- BillToAddressID
- ShipToAddressID

Sales.SalesOrderHeaderSalesRe...
- SalesOrderID
- SalesReasonID

Sales.SalesReason
- SalesReasonID
- Name

Sales.CurrencyRate
- CurrencyRateID
- CurrencyRateDate
- FromCurrencyCode
- ToCurrencyCode
- AverageRate
- EndOfDayRate

Sales.SalesOrderDetail
- SalesOrderID
- SalesOrderDetailID
- CarrierTrackingNumber
- OrderQty
- ProductID
- SpecialOfferID
- UnitPrice
- UnitPriceDiscount
- LineTotal

Production.Product
- ProductID
- Name
- ProductNumber

Sales.ShoppingCartItem
- ShoppingCartItemID
- ShoppingCartID
- Quantity
- ProductID
- DateCreated

Sales.Currency
- CurrencyCode
- Name

Sales.SpecialOfferProduct
- SpecialOfferID
- ProductID

Sales.SpecialOffer
- SpecialOfferID
- Description
- DiscountPct
- Type
- Category
- StartDate
- EndDate
- MinQty
- MaxQty

Sales.CountryRegionCurrency
- CountryRegionCode
- CurrencyCode

2. Dimensional model

a. Define the business processes that you will model

1. Order Status Tracking Process: Monitor and analyze the status of orders over time by using Fact 'OrderStatus Fact'.

Business Use Case: Analyzing order processing efficiency by tracking the distribution of orders across different statuses over time.

---

2. Order Fulfillment Process: Track the lifecycle of sales orders from creation to fulfillment by using fact 'orderDates Fact'.

Business Use Case: Analyzing order lead times, monitoring delivery schedule adherence, and identifying delays.

---

3. Order Revenue Analysis: Measure and analyze product sales revenue by using the 'OrderRevenue Fact'.

Business Use Case: Analyzing sales performance by product, territory, and special promotions, monitoring revenue trends, and assessing the effectiveness of discount strategies.

b. Declare the grain of each fact table (simply, describe what a fact table record represents).

OrderStatus Fact	
TimeID	Represent time from the time dimension
StatusId	Represent status id from table statusDim
NumberOfOrders	Represent the number of orders relative to specific status in the specific time frame

OrderDates Fact	
TimeID	Represent time from the time dimension
SalesOrderID	Represent Order ID of SalesOrderHeader Dim
OrderDate	Represent the start date of order
DueDate	Represent the expected received date of order
ShipDate	Represent the Shipping date of order

OrderRevenue Fact	
TimeID	Represent time from the time dimension
ProductID	Represent product ID from product dimension
SalesOrderID	Represent Order id from table from sales OrderDetails Dim
SalesOrderDetailsID	Represent Order Detail id from table from sales OrderDetails dimension
SpecialOfferID	Represent special order ID from SpecialOrder dimension
TerritoryName	Represent territory Name from SalesOrderHeader dimension
ProductSold	Represent Quantity of products sold
UnitPrice	Represent listing price of product
DiscountPrice	Represent discount price of product
RevenueWithoutDiscount	Represent gaining without Discount
LineTotal	Represent gaining with Discount

c. Define the type of each fact table.

Fact Table	Type
OrderDates Fact	Cumulative
OrderStatus Fact	Periodic Snapshot
OrderRevenue Fact	Transaction

d. Define the dimensions and the type of each one

Dimension	Type
Time Dimension	● Static ● Role Playing ● Conformed
Status Dimension	● Static
Product Dimension	● Slowly Changing (Type 1)
SalesOrderHeader Dimension	● Slowly Changing (Type 1)
SalesOrderDetails Dimension	● Slowly Changing (Type 2)
SpecialOffer Dimension	● Slowly Changing (Type 4)

e. Define the measures that will appear in the fact tables and the type of each one.

Measure	Fact Table	Type
OrderDate	OrderDates FT	Non additive
DueDate	OrderDates FT	Non additive
ShipDate	OrderDates FT	Non additive
NumberOfOrders	OrderStatus FT	Semi additive
UnitPrice	OrderRevenue FT	Non additive
DiscountPrice	OrderRevenue FT	Non additive
TerritoryName	OrderRevenue FT	Non additive
ProductSold (OrderQty)	OrderRevenue FT	Fully additive
Revenue	OrderRevenue FT	Fully additive
RevenueWithoutDiscount (Derived Column)	OrderRevenue FT	Fully additive

f. Physical Model (Star Schema)

Cumulative

OrderDates Fact

- TimeID
- SalesOrderID
- OrderDate => Non
- DueDate => Non
- ShipDate => Non

Static

StatusDim

- StatusID
- Name

Periodic Snapshot

OrderStatus Fact

- TimeID
- StatusID
- NumberOfOrders => Semi

Static
Role Playing
Conformed

TimeDim

- TimeID
- Year
- Month
- Day
- Hour

Slowly Changing (Type 1)
Incremental Loading

SalesOrderHeaderDim

- SalesOrderID
- OrderDate
- DueDate => Changed
- ShipDate => Changed
- SalesOrderNumber
- PurchaseOrderNumber
- CustomerID
- SalesPersonID
- TerritoryID
- TerritoryName

Transactional

OrderRevenue Fact

- TimeID
- ProductID
- SalesOrderID
- SalesOrderDetailsID
- SpecialOfferID
- TerritoryName => Non
- ProductSold => Fully
- UnitPrice => Non
- DiscountPrice => Non
- RevenueWithoutDiscount => (Derived) Fully
- Revenue => Fully

Slowly Changing (Type 2)

SalesOrderDetailsDim

- SalesOrderID
- SalesOrderDetailID
- ProductID
- SpecialOfferID
- OrderQT => Changed
- UnitPrice
- UnitPriceDiscount => Changed
- LineTotal

Slowly Changing (Type 4)
Whole Extraction

SpecialOfferDim

- SpecialOfferID
- DiscountPct => Changed
- Type
- Category
- StartDate => Changed
- EndDate => Changed

Slowly Changing (Type 1)
Whole Extraction

ProductDim

- ProductID
- Name => Changed
- ListPrice
- MakeFlag
- FinishedGoodsFlag
- Weight => Changed

3. Screenshots of SSIS control and data flow tasks
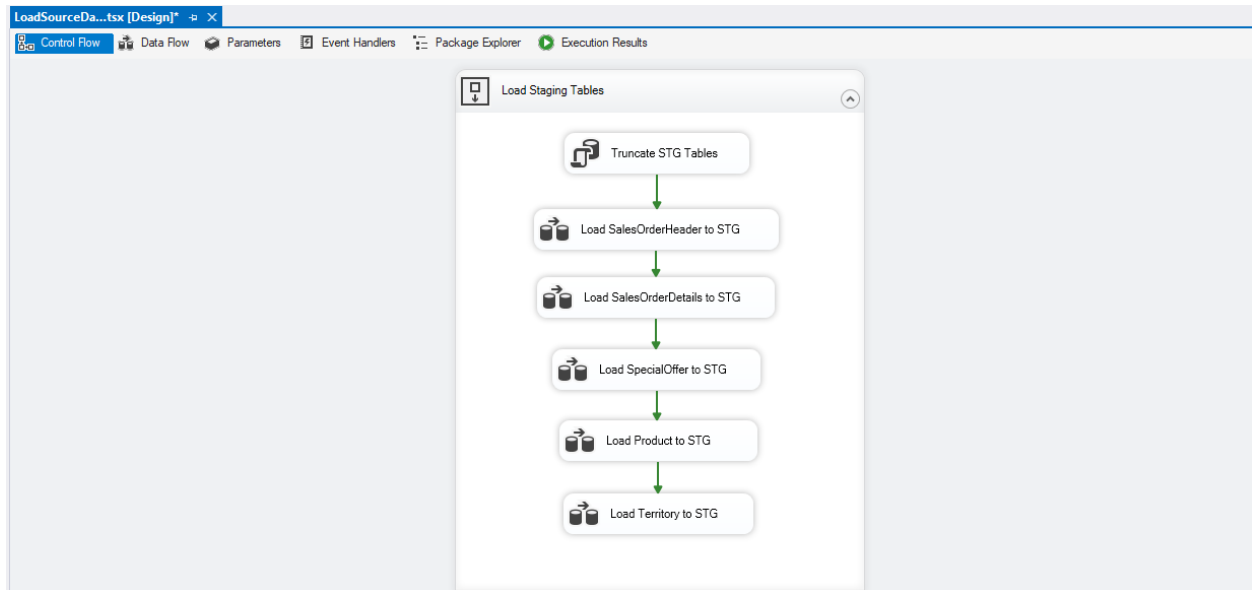
Control Flow To Load Data From <u>Source DB</u> To Staging Area

Data Flow To Load <u>SalesOrderHeader</u> Into Staging Area

Same goes inside all data flow tasks in the sequence container above
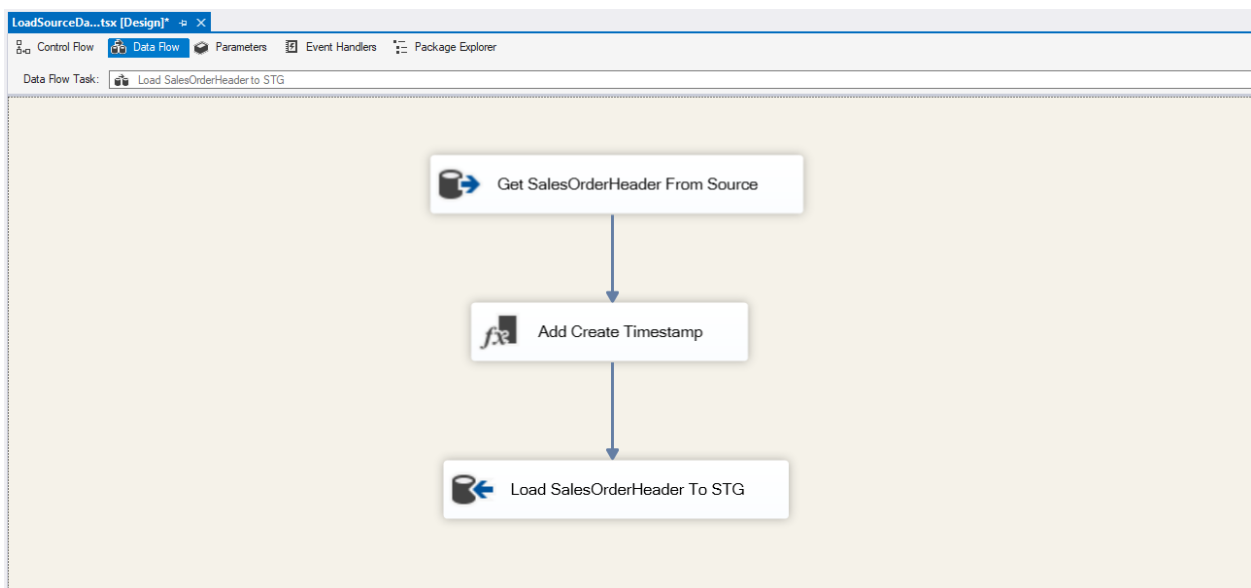
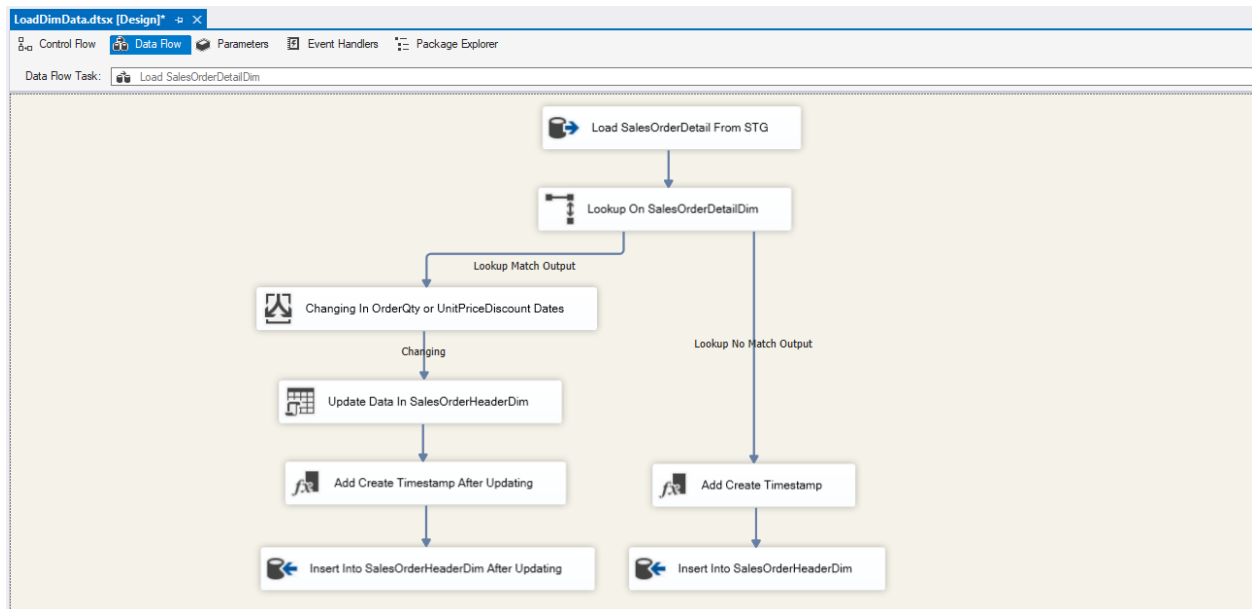# Control Flow To Load Dimension Tables From Staging Area
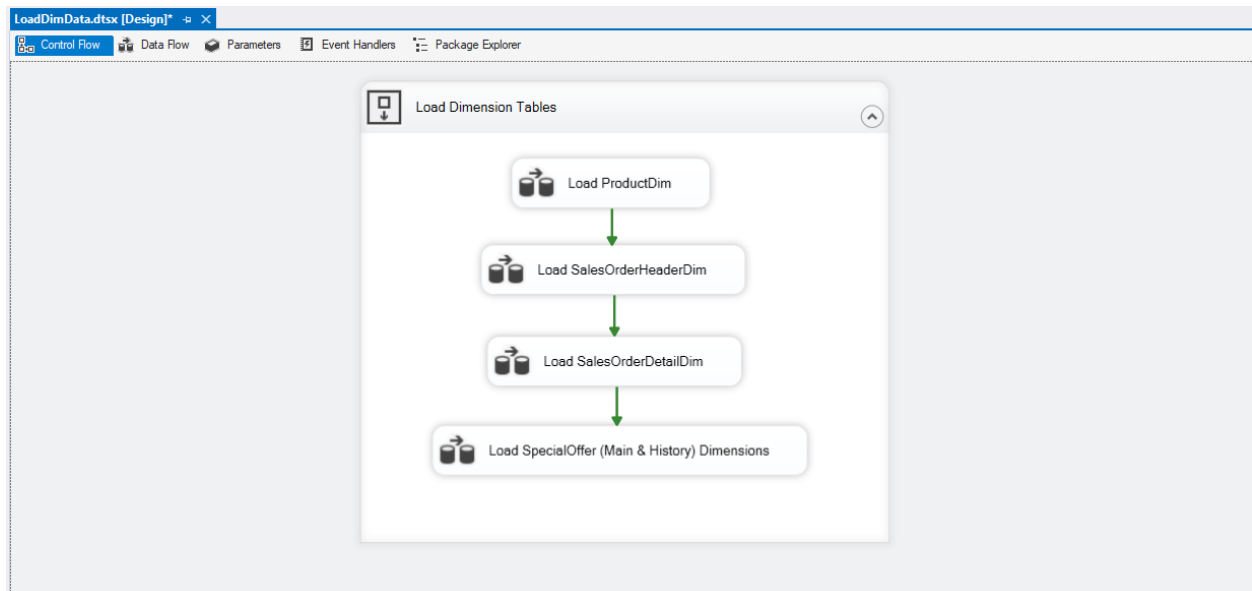
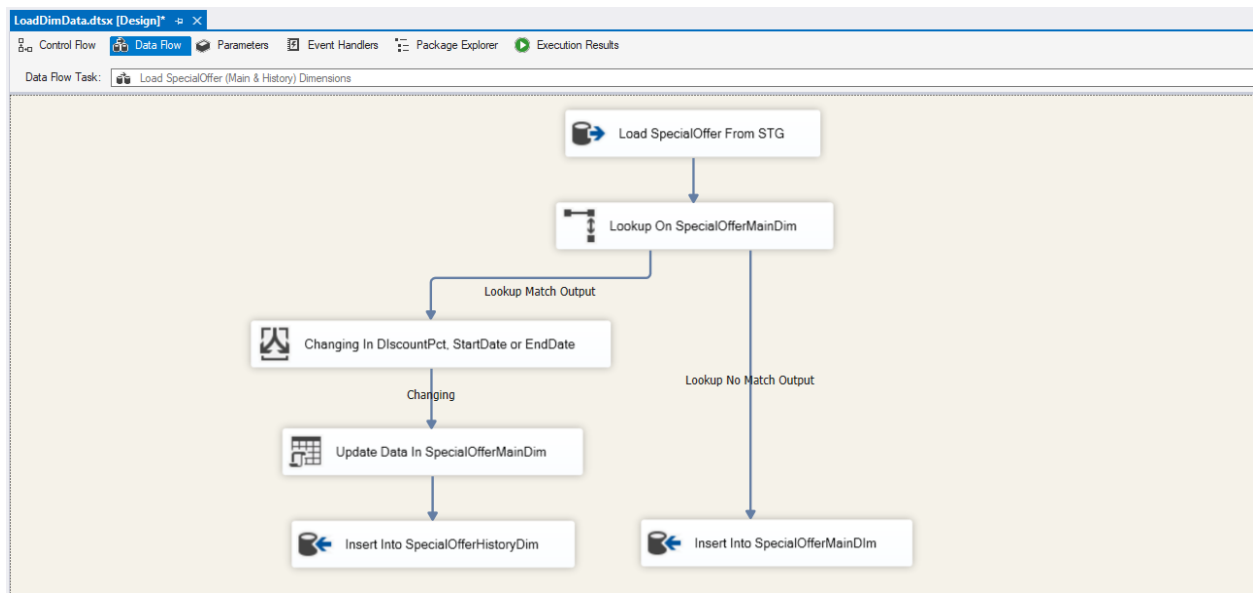# Data Flow To Load ProductDim From Staging Area
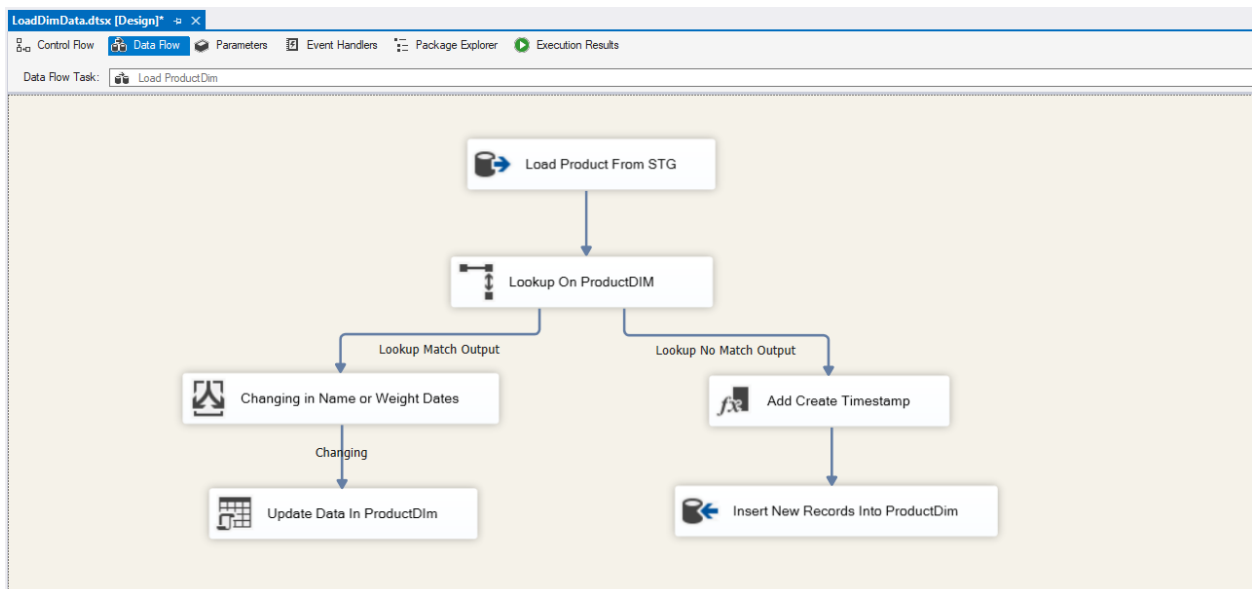
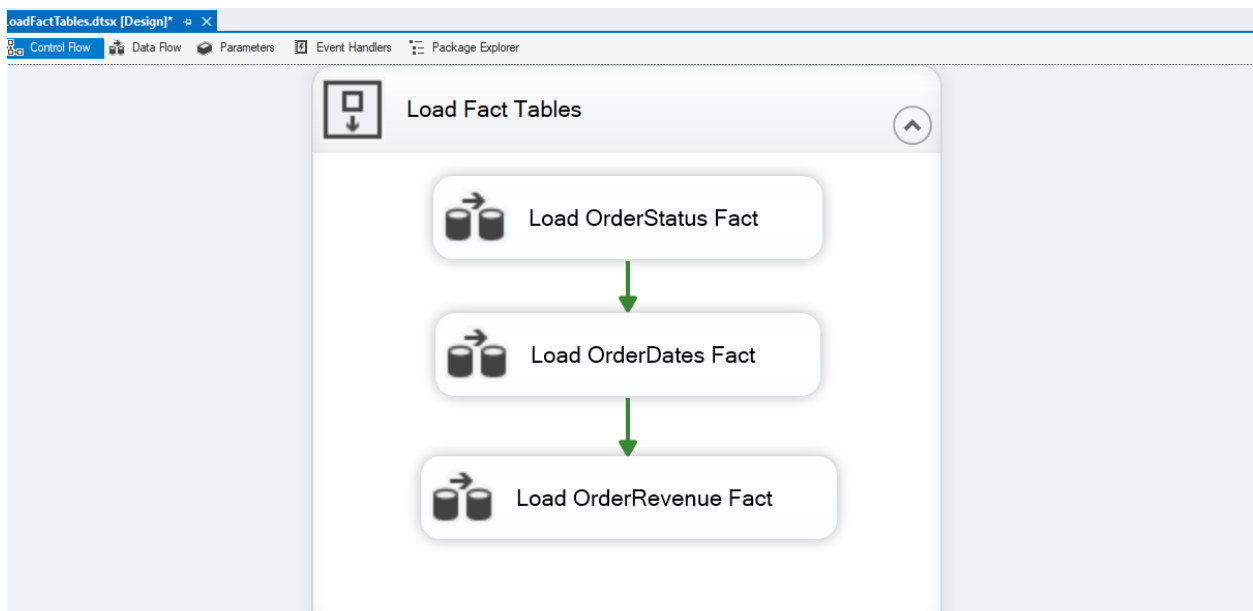# Data Flow To Load SalesOrderHeaderDim From Staging Area
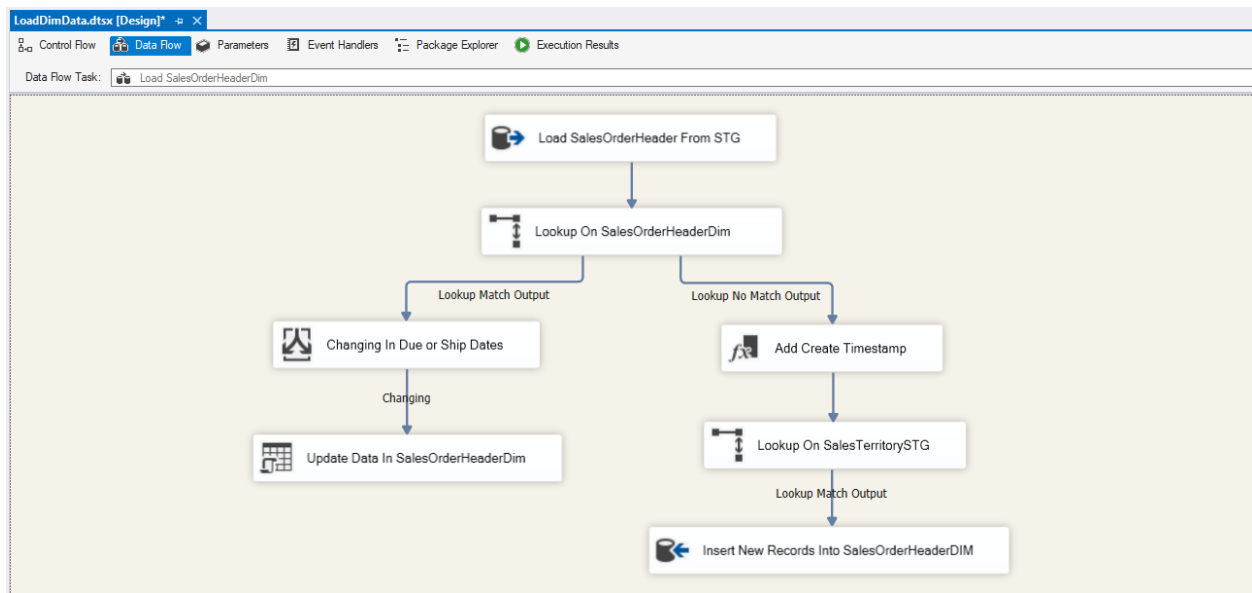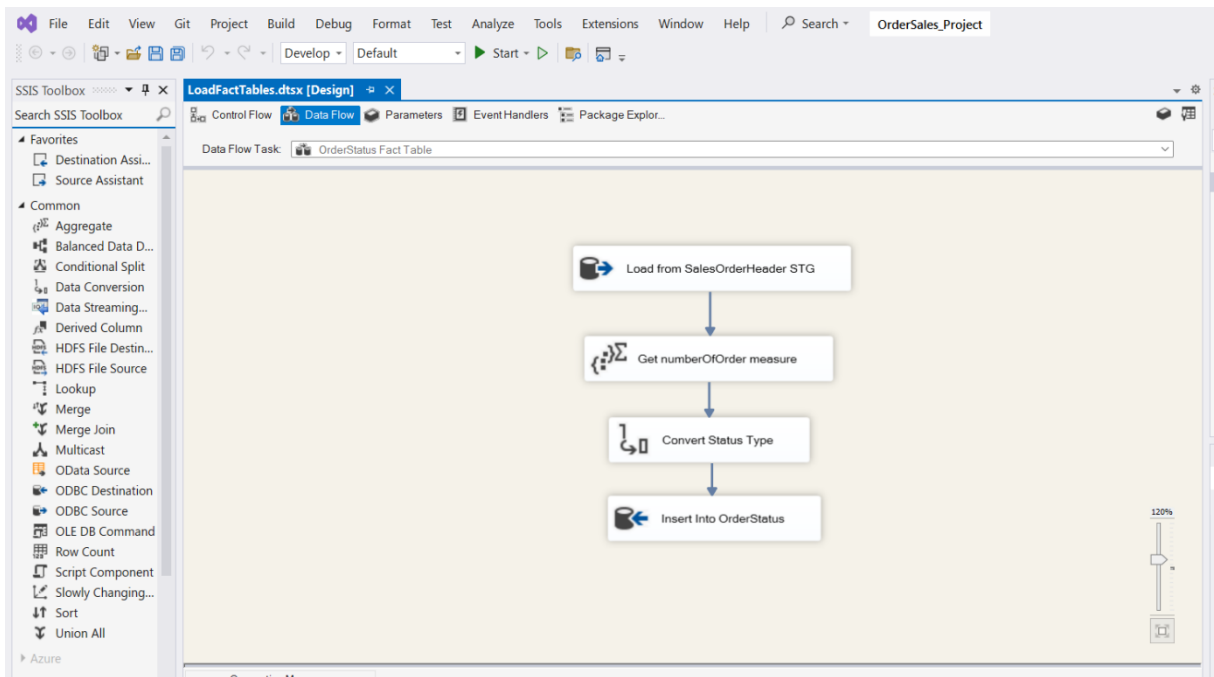

Data Flow To Load SalesOrderDetailDim From Staging Area

Data Flow To Load <u>SpecialOffer</u> (Main & History) Dimensions

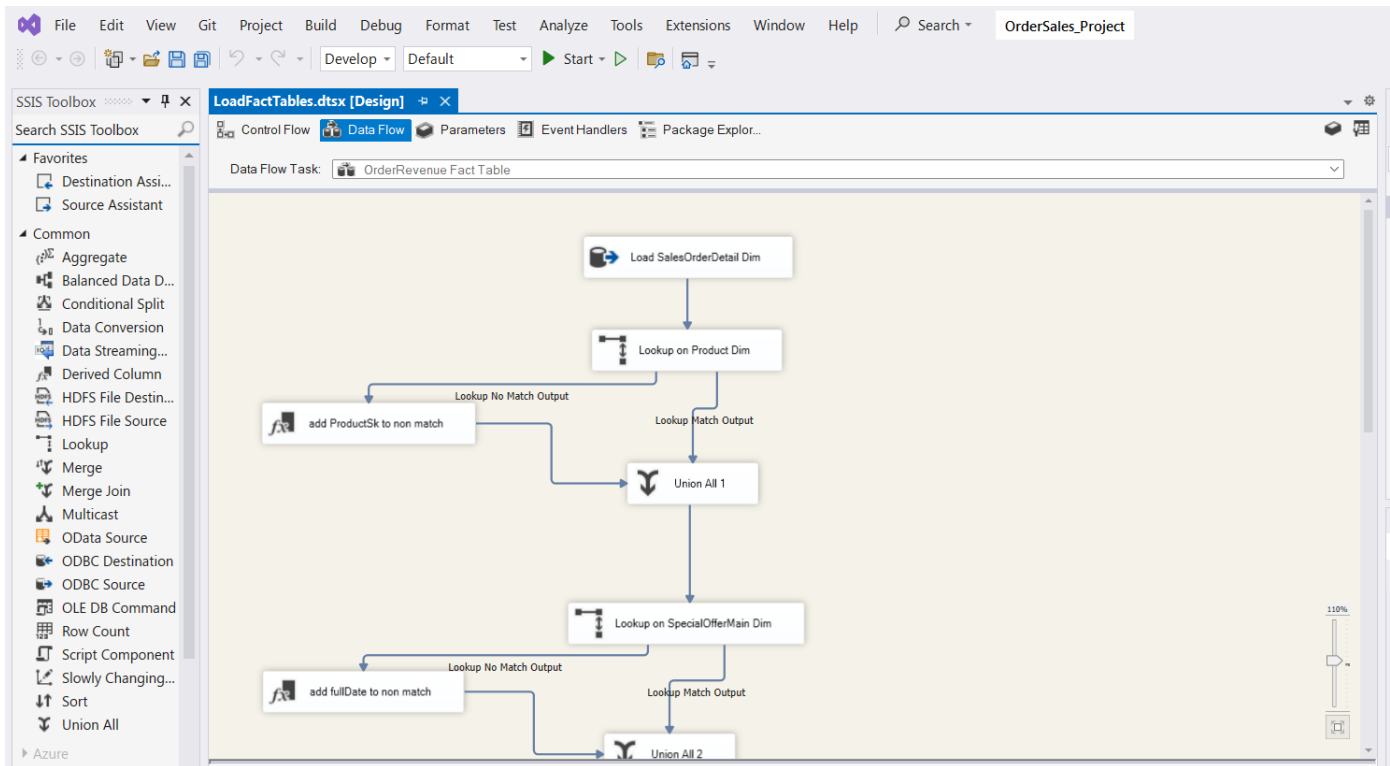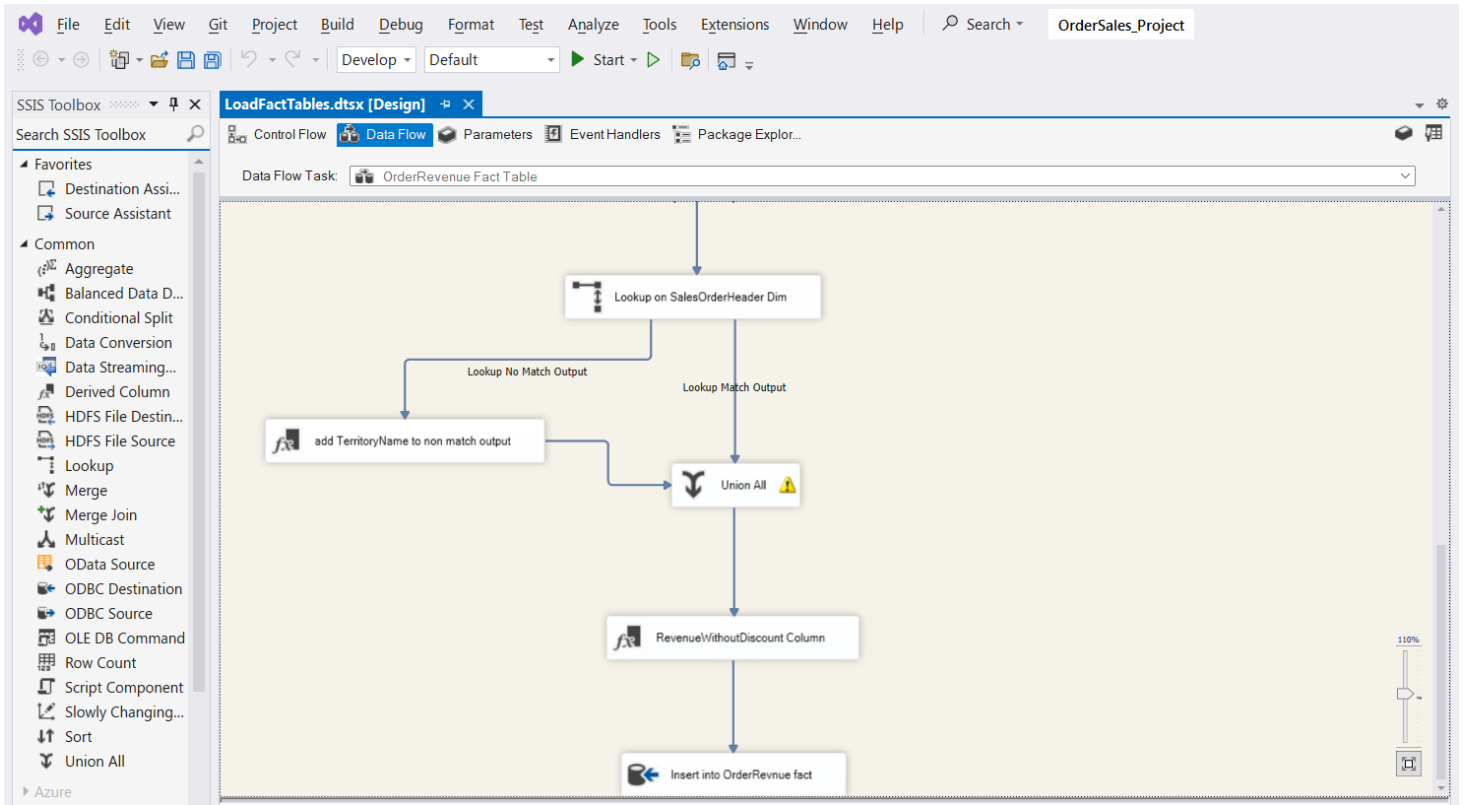# Control Flow To Load <u>Fact</u> Tables

- Fact OrderStatus

- Fact OrderDates
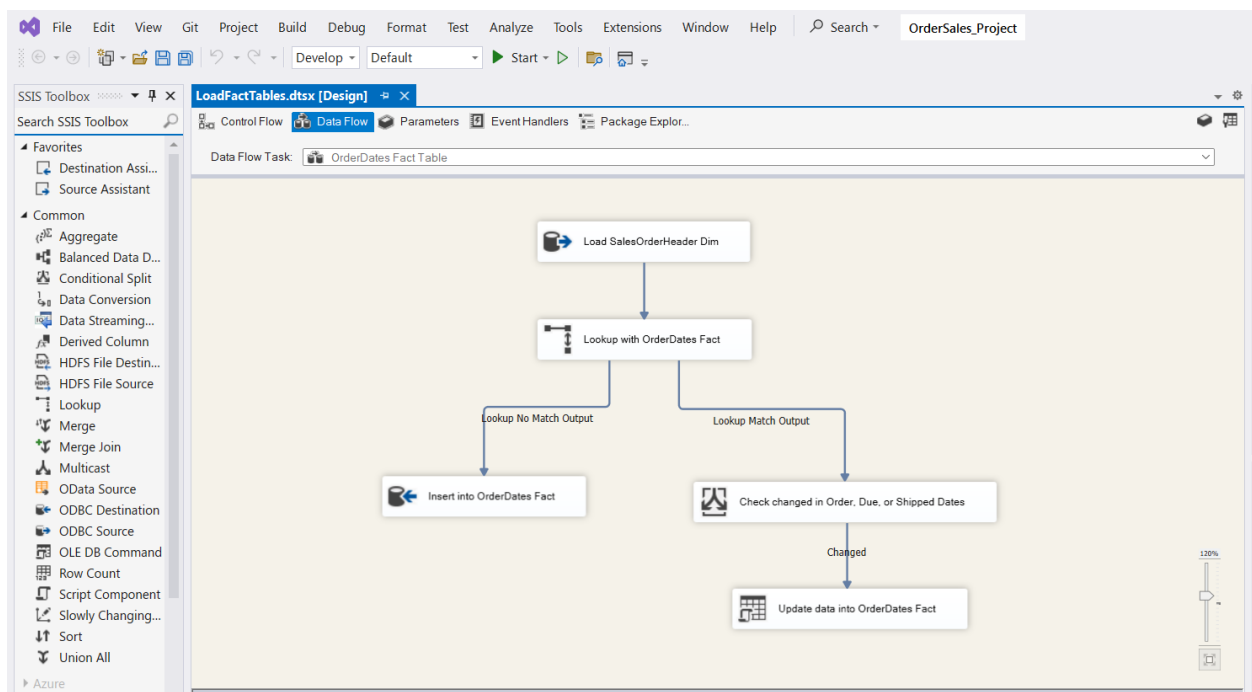
- **Fact OrderRevenue**
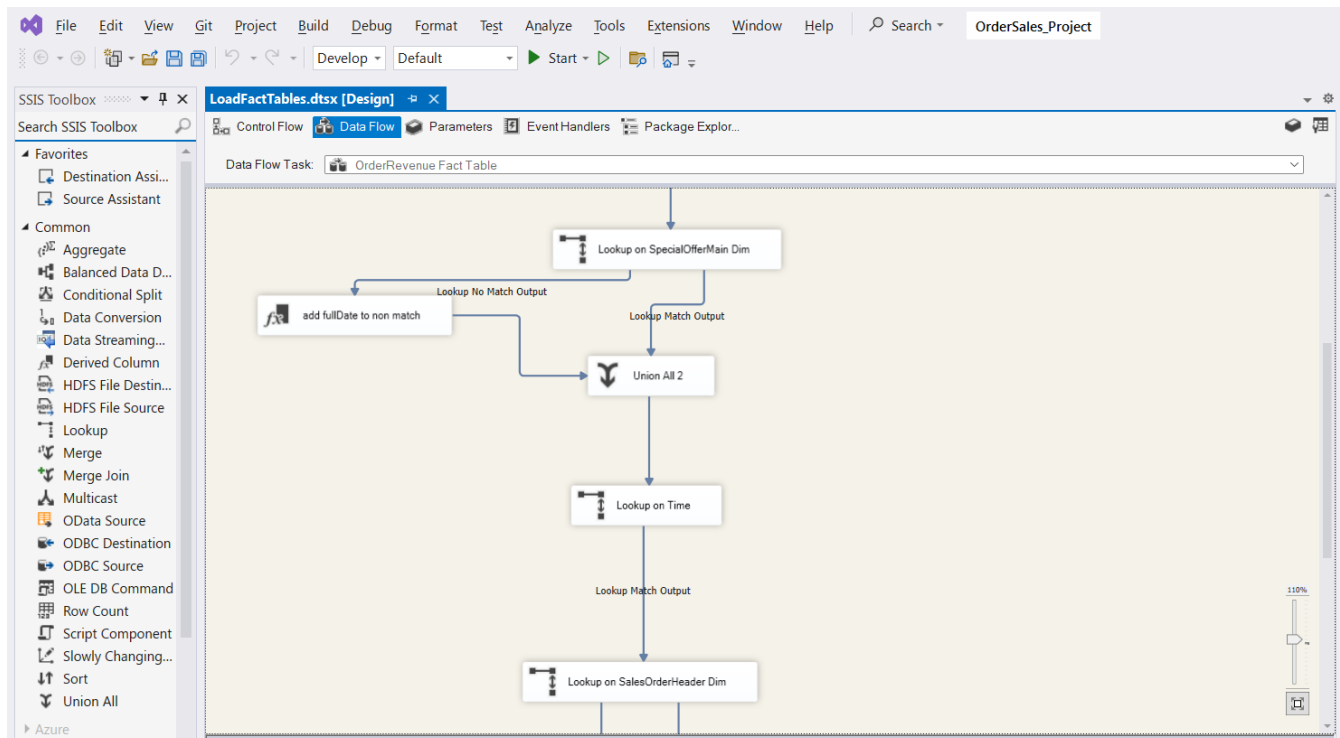
4. Queries on each fact table

a) Queries to get insights about the <u>OrderStatus</u> Fact Table

```sql
-- Number of orders over time dimension
SELECT TimeID,
       SUM(NumberOfOrders) AS No_Orders_Over_Time
FROM [DWH].F_OrderStatus
GROUP BY TimeID;
```

	TimeID	No_Orders_Over_Time
1	201106070000	43
2	201106080000	4
3	201106090000	5
4	201106100000	2
5	201106110000	5

Query executed successfully.   WIN10\SQL_SRVR (16.0 RTM)   WIN10\iSeFz (52)   DWH_Project   00:00:00   1,124 rows

```sql
-- Number of orders over status dimension
SELECT StatusID,
       SUM(NumberOfOrders) AS No_Orders_Over_Status
FROM [DWH].F_OrderStatus
GROUP BY StatusID
ORDER BY StatusID;
```

	StatusID	No_Orders_Over_Status
1	5	31465

```sql
-- Total number of orders all time
SELECT SUM(NumberOfOrders) AS Total_No_Orders
FROM [DWH].F_OrderStatus;
```

	Total_No_Orders
1	31465

```sql
-- Number of orders in each year
SELECT LEFT(TimeID, 4) AS Time_Year,
       SUM(NumberOfOrders) AS No_Orders_In_Year
FROM [DWH].F_OrderStatus
GROUP BY LEFT(TimeID, 4);
```

	Time_Year	No_Orders_In_Year
1	2011	1566
2	2012	3774
3	2013	13752
4	2014	12373

Query executed successfully.   WIN10\SQL_SRVR (16.0 RTM) WIN10\iSeFz (52) DWH_Project 00:00:00 4 rows

b) Queries to get insights about the OrderDates Fact Table

```sql
-- Total orders orders, shipped and due
SELECT SalesOrderID,
       OrderDate,
       ShipDate,
       DueDate
FROM [DWH].F_OrderDates;
```

	SalesOrderID	OrderDate	ShipDate	DueDate
1	1	2011-05-31	2011-06-07	2011-06-12
2	2	2011-05-31	2011-06-07	2011-06-12
3	3	2011-05-31	2011-06-07	2011-06-12
4	4	2011-05-31	2011-06-07	2011-06-12
5	5	2011-05-31	2011-06-07	2011-06-12

Query executed successfully.   WIN10\SQL_SRVR (16.0 RTM) WIN10\iSeFz (52) DWH_Project 00:00:01 31,465 rows

c) Queries to get insights about the <u>OrderRevenue</u> Fact Table

```sql
-- Sum of measures over all dimensions
SELECT TimeID,
       ProductID,
       SpecialOfferID,
       SUM(ProductSold) AS Sum_Product_Sales,
       SUM(RevenueWithoutDiscount) AS Sum_Revenue_Without_Discount,
       SUM(Revenue) AS Sum_Revenue
FROM [DWH].F_OrderRevenue
GROUP BY TimeID,
         ProductID,
         SpecialOfferID;
```

233 %

Results | Messages

	TimeID	ProductID	SpecialOfferID	Sum_Product_Sales	Sum_Revenue_Without_Discount	Sum_Revenue
1	201105010000	388	2	449	14060.075800	13778.874284
2	201105010000	270	3	21	9043.534500	8591.357775
3	201105010000	471	2	71	98176.002600	96212.482548
4	201104010000	214	1	743	4235.100000	4235.100000
5	201304300000	460	14	144	138752.874000	111479.113200

✅ Query executed successfully.   WIN10\SQL_SRVR (16.0 RTM) | WIN10\iSeFz (52) | DWH_Project | 00:00:01 | 484 rows

```sql
-- Sum of measures over Time and Product dimensions
SELECT TimeID,
       ProductID,
       SUM(ProductSold) AS Sum_Product_Sales,
       SUM(RevenueWithoutDiscount) AS Sum_Revenue_Without_Discount,
       SUM(Revenue) AS Sum_Revenue
FROM [DWH].F_OrderRevenue
GROUP BY TimeID,
         ProductID;
```

233 %

Results | Messages

	TimeID	ProductID	Sum_Product_Sales	Sum_Revenue_Without_Discount	Sum_Revenue
1	201104010000	504	1165	438795.874000	438795.874000
2	201104010000	218	429	21445.710000	21445.710000
3	201104010000	364	2894	46854.886300	46854.886300
4	201104010000	381	2239	176952.000000	176952.000000
5	201104010000	358	394	17727.636000	17727.636000

✅ Query executed successfully.   WIN10\SQL_SRVR (16.0 RTM) | WIN10\iSeFz (52) | DWH_Project | 00:00:00 | 412 rows

```sql
-- Sum of measures over Time and SpecialOffer dimensions
SELECT TimeID,
       SpecialOfferID,
       SUM(ProductSold) AS Sum_Product_Sales,
       SUM(RevenueWithoutDiscount) AS Sum_Revenue_Without_Discount,
       SUM(Revenue) AS Sum_Revenue
FROM [DWH].F_OrderRevenue
GROUP BY TimeID,
         SpecialOfferID;
```

	TimeID	SpecialOfferID	Sum_Product_Sales	Sum_Revenue_Without_Discount	Sum_Revenue
1	201203140000	7	456	386042.610000	250927.696500
2	201304300000	13	1581	536310.757500	458091.193875
3	201304300000	11	680	10706.940000	9100.899000
4	201104010000	1	238944	102372622.260400	102372622.260400
5	201105010000	4	2321	137942.815900	124148.534310

Query executed successfully.   WIN10\SQL_SRVR (16.0 RTM) | WIN10\iSeFz (52) | DWH_Project | 00:00:00 | 12 rows

```sql
-- Sum of measures over Product and SpecialOffer dimensions
SELECT ProductID,
       SpecialOfferID,
       SUM(ProductSold) AS Sum_Product_Sales,
       SUM(RevenueWithoutDiscount) AS Sum_Revenue_Without_Discount,
       SUM(Revenue) AS Sum_Revenue
FROM [DWH].F_OrderRevenue
GROUP BY ProductID,
         SpecialOfferID;
```

	ProductID	SpecialOfferID	Sum_Product_Sales	Sum_Revenue_Without_Discount	Sum_Revenue
1	277	2	14	27607.918800	27055.760424
2	364	1	2894	46854.886300	46854.886300
3	357	4	28	1049.860000	944.874000
4	220	2	877	25018.635000	24518.262300
5	307	1	190	16897.080000	16897.080000

Query executed successfully.   WIN10\SQL_SRVR (16.0 RTM) | WIN10\iSeFz (52) | DWH_Project | 00:00:00 | 484 rows

```sql
-- Sum of measures over Time dimension
SELECT TimeID,
       SUM(ProductSold) AS Sum_Product_Sales,
       SUM(RevenueWithoutDiscount) AS Sum_Revenue_Without_Discount,
       SUM(Revenue) AS Sum_Revenue
FROM [DWH].F_OrderRevenue
GROUP BY TimeID;
```

	TimeID	Sum_Product_Sales	Sum_Revenue_Without_Discount	Sum_Revenue
1	201104010000	238944	102372622.260400	102372622.260400
2	201403010000	382	43165.236000	25899.141600
3	201304300000	3036	1304675.143500	1079516.631675
4	201204300000	796	79685.161200	57434.907480
5	201203140000	456	386042.610000	250927.696500
6	201105010000	31300	6187698.902300	6059980.762233

Query executed successfully.   WIN10\SQL_SRVR (16.0 RTM) | WIN10\iSeFz (52) | DWH_Project | 00:00:00 | 6 rows

```sql
-- Sum of measures over Product dimension
SELECT ProductID,
       SUM(ProductSold) AS Sum_Product_Sales,
       SUM(RevenueWithoutDiscount) AS Sum_Revenue_Without_Discount,
       SUM(Revenue) AS Sum_Revenue
FROM [DWH].F_OrderRevenue
GROUP BY ProductID;
```

	ProductID	Sum_Product_Sales	Sum_Revenue_Without_Discount	Sum_Revenue
1	261	346	302678.724000	302678.724000
2	238	90	32120.820000	32120.820000
3	355	445	16017.330000	16017.330000
4	215	90	513.000000	513.000000
5	378	3865	8232.916400	8232.597632

Query executed successfully.   WIN10\SQL_SRVR (16.0 RTM)   WIN10\iSeFz (52)   DWH_Project   00:00:00   266 rows

```sql
-- Sum of measures over SpecialOffer dimension
SELECT SpecialOfferID,
       SUM(ProductSold) AS Sum_Product_Sales,
       SUM(RevenueWithoutDiscount) AS Sum_Revenue_Without_Discount,
       SUM(Revenue) AS Sum_Revenue
FROM [DWH].F_OrderRevenue
GROUP BY SpecialOfferID;
```

	SpecialOfferID	Sum_Product_Sales	Sum_Revenue_Without_Discount	Sum_Revenue
1	9	304	71408.688000	49986.081600
2	3	10713	1092256.134600	1037643.327870
3	7	456	386042.610000	250927.696500
4	1	238944	102372622.260400	102372622.260400
5	4	2321	137942.815900	124148.534310

Query executed successfully.   WIN10\SQL_SRVR (16.0 RTM)   WIN10\iSeFz (52)   DWH_Project   00:00:00   12 rows

```sql
-- Total measures without any aggregation
SELECT SUM(ProductSold) AS Total_Product_Sales,
       SUM(RevenueWithoutDiscount) AS Total_Revenue_Without_Discount,
       SUM(Revenue) AS Total_Revenue
FROM [DWH].F_OrderRevenue;
```
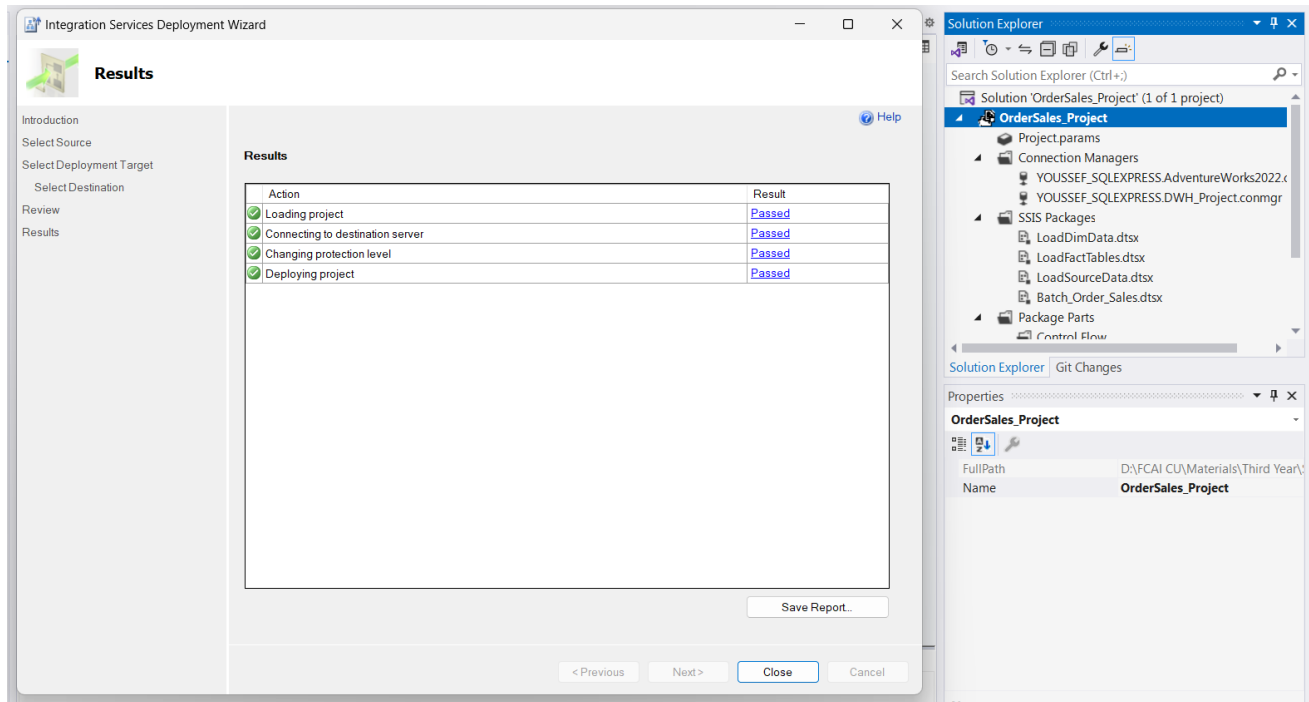
|	Total_Product_Sales	Total_Revenue_Without_Discount	Total_Revenue
1	274914	110373889.313400	109846381.399888

Query executed successfully.   WIN10\SQL_SRVR (16.0 RTM)   WIN10\iSeFz (52)   DWH_Project   00:00:00   1 rows

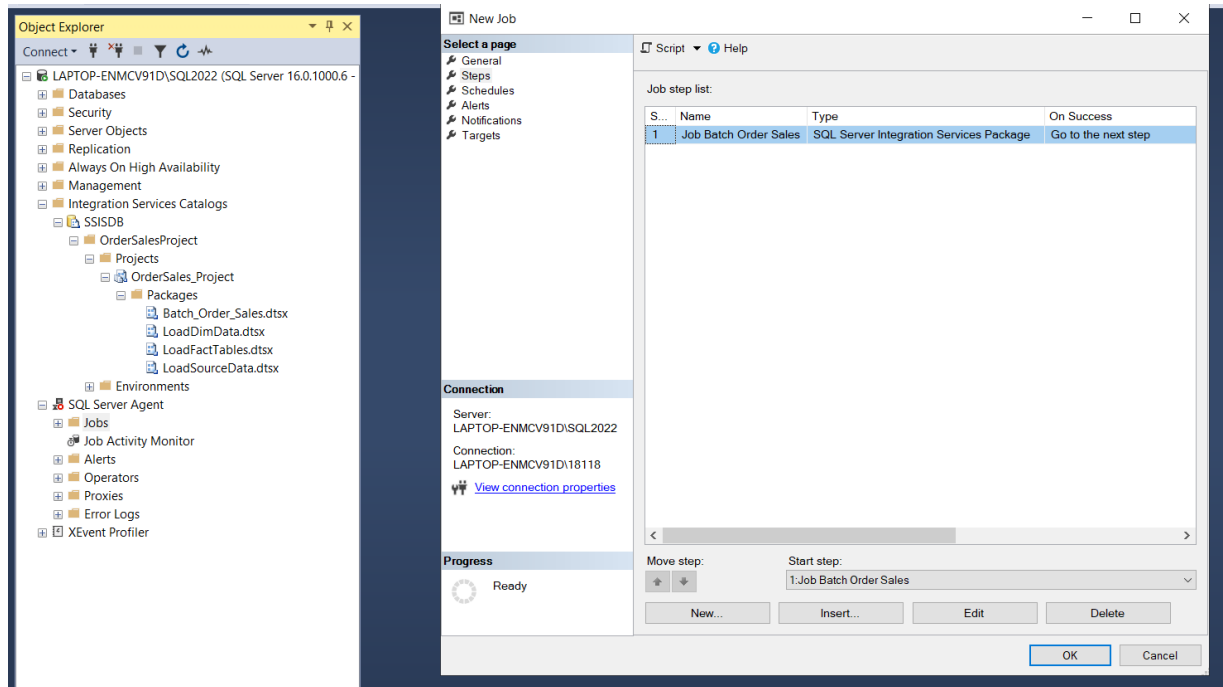# 5. Screenshots of the deployed SSIS packages with their schedule
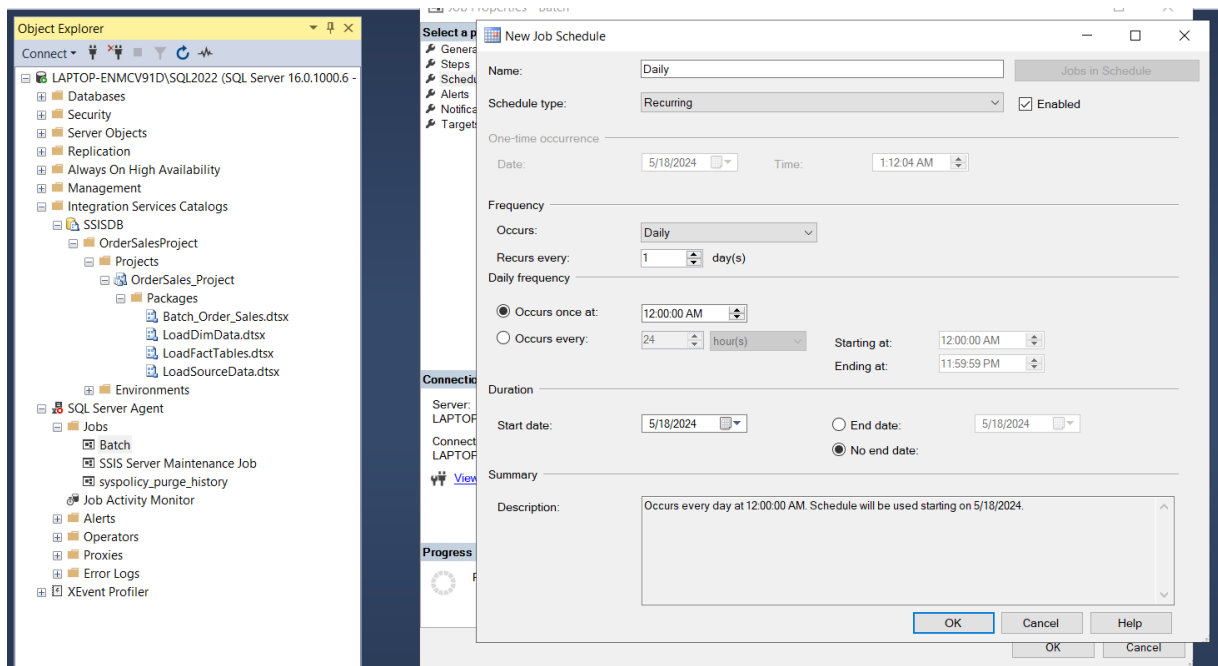
Packages Deployed Successfully From SSIS

Batch Package Executed Successfully In SSMS After Deployment
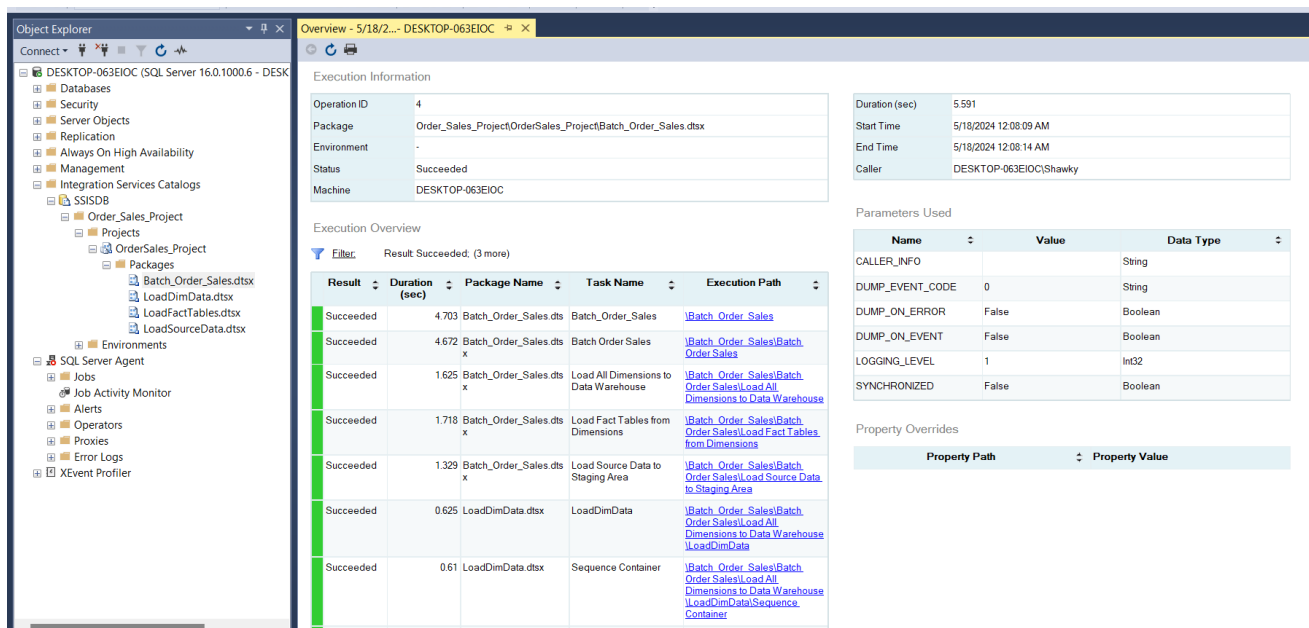
Batch Job For Scheduling


Job Scheduled For Daily Run

6. BONUS Power BI DWH Dashboard
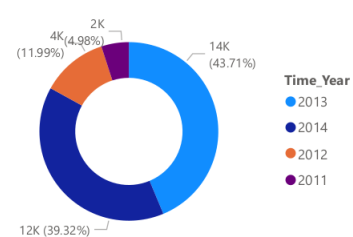
Fact OrderStatus Dash Board

TimeID	Sum of NO_Of_Orders_Each_Time ▲
201107240000	1
201109050000	1
201110010000	1
201201280000	1
201205040000	1
201210100000	1
201106100000	2
201106190000	2
201106240000	2
201108030000	2
201108040000	2
201108190000	2
201108270000	2
201109040000	2
201109100000	2
201109210000	2
201110020000	2
201110070000	2
201110290000	2
201111160000	2
201201250000	2
201202140000	2
Total	**31465**

Sum of Total_Of_Orders by StatusID

0K (0.45%)
5K (15.7%)
8K (24.53%)
5K (17.32%)
7K (21.9%)
6K (20.1%)

StatusID
- 4
- 6
- 2
- 5
- 3
- 1
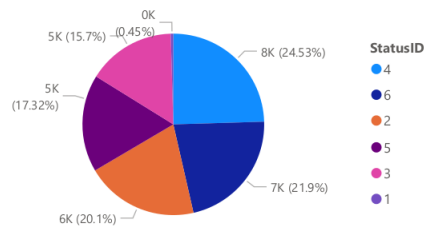
31K
Sum of Total_NO_Of_Orders

Sum of NO_Of_Orders_In_Year by Time_Year

2K
4K (4.98%)
14K (43.71%)
12K (39.32%)

Time_Year
- 2013
- 2014
- 2012
- 2011

Sum of NO_Of_Orders_In_Range and Count of TimeID

148

25   296
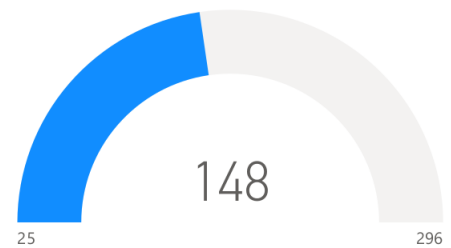
Fact OrderDates Dash Board

SalesOrderID	OrderDate	ShipDate ▲	DueDate
1	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
2	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
3	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
4	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
5	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
6	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
7	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
8	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
9	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
10	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
11	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
12	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
13	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
14	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
15	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
16	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
17	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
18	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
19	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
20	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص
21	31/05/2011 12:00:00 ص	07/06/2011 12:00:00 ص	12/06/2011 12:00:00 ص

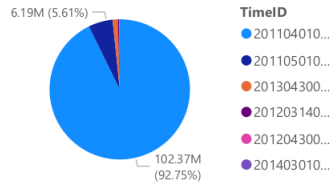**Sum of NO_Ordered_Orders_Month by Time_Month**

3153
3089
2887
2838
2733
2548
2527
2485
2401
2389
2212
2203

Time_Month: 05, 04, 11, 01, 03, 12, 02, 10, 07, 08, 06, 09
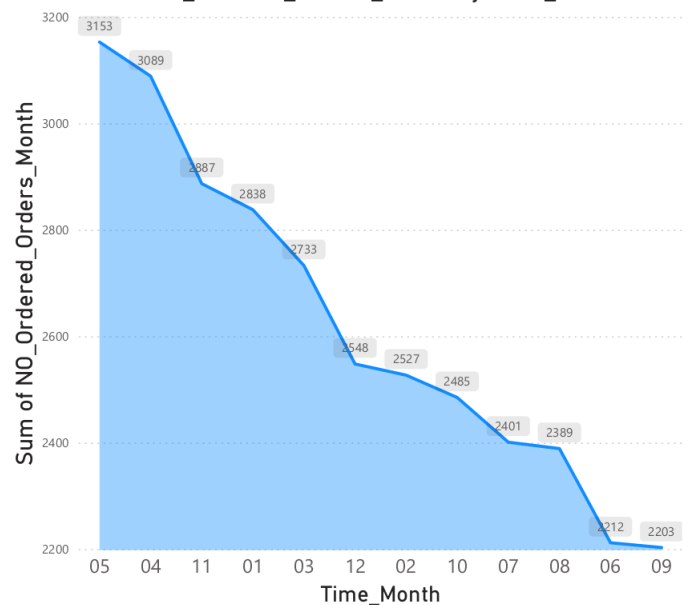
Fact OrderRevenue Dash Board

Count of TimeID by Sum_Of_Revenue

Sum_Of_Revenue
- 25899.1416
- 57434.90748
- 250927.6965
- 1079516.631...
- 6059980.762...
- 102372622.2...
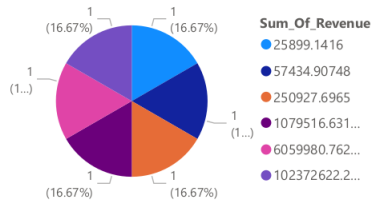
Sum of Sum_Of_Revenue_Without_Discount by TimeID

6.19M (5.61%)
102.37M (92.75%)

TimeID
- 201104010...
- 201105010...
- 201304300...
- 201203140...
- 201204300...
- 201403010...

275K
Sum of Total_Sum_Of_Product

Count of TimeID by Sum_Of_Product

Sum_Of_Product
- 382
- 456
- 796
- 3036
- 31300
- 238944

Sum of Sum_Of_Revenue_Without_Discount and Sum of Sum_Of_Revenue

109.8... (49.8...)
110.3... (50.1...)
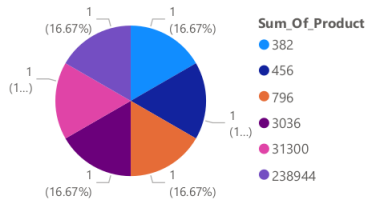
- Sum of Su...
- Sum of Su...

109.85M
Sum of Total_Sum_Of_Revenue
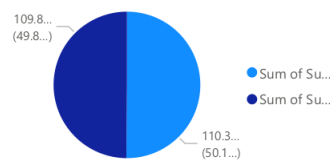
110.37M
Sum of Total_Sum_Of_Revenue_Without_Discount

Fact OrderRevenue Dash Board

Sum of Sum_Of_Revenue, Sum of Sum_Of_Revenue_Without_Discount and Sum of Sum_Of_Product by SpecialOfferID

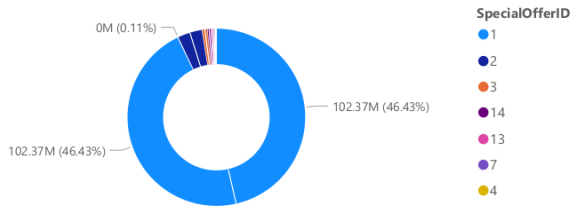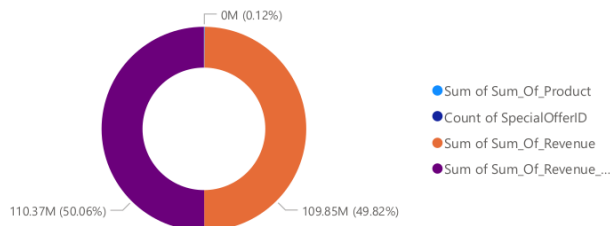0M (0.11%)
102.37M (46.43%)
102.37M (46.43%)

SpecialOfferID
- 1
- 2
- 3
- 14
- 13
- 7
- 4

Sum of Sum_Of_Product, Count of SpecialOfferID, Sum of Sum_Of_Revenue and Sum of Sum_Of_Revenue_Without_Discount

0M (0.12%)
110.37M (50.06%)
109.85M (49.82%)

- Sum of Sum_Of_Product
- Count of SpecialOfferID
- Sum of Sum_Of_Revenue
- Sum of Sum_Of_Revenue_...

ProductID	Sum of Sum_Of_Product	Sum of Sum_Of_Revenue	Sum of Sum_Of_Revenue_Without_Discount
212	6266	157,772.39	158,865.91
213	6532	160,869.52	162,076.70
214	1107	6,060.39	6,147.64
215	90	513.00	513.00
216	6743	165,406.62	166,720.91
217	8311	51,229.45	51,512.28
218	429	21,445.71	21,445.71
219	3636	115,249.21	115,509.37
220	6592	198,754.98	200,240.74
221	2980	95,611.20	95,689.15
222	485	394,255.57	394,255.57
223	486	395,182.70	395,182.70
224	109	89,872.17	89,872.17
227	940	177,635.90	177,635.90
228	129	24,844.69	24,844.69
230	991	194,692.60	194,782.57
231	671	132,125.25	132,125.25
232	100	20,104.44	20,104.44
234	996	195,933.41	195,933.41
235	697	137,213.49	137,256.51
237	250	89,224.50	89,224.50
238	90	22,120.83	22,120.83
Total	**274914**	**109,846,381.40**	**110,373,889.31**