Northeastern University

College of Engineering

INFO 7290: Data Warehousing and Business Intelligence

Sales Data Warehouse

Group 8

JayaChandra Chimakurthi (001065324)

Vyshnavi Bhyravajosula (001085050)

Kirtikumar Waykos (001029199)

Pradeep Poonati (001316491)

Index

Revision History:

Date	Version	Changes	Author
03/27/2021	1.0	Initial Document	All members
04/06/2021	1.1	Added pipelines for managing SCD in the tables. Created facts and dimensions for all the tables. Built Dimension Model Chart. Added Professor’s comment as appendix	All members
04/18/2021	1.2	Added how the data is getting populated to fact tables, expanded the business scenarios. Explained how the data is analyzed and insights were drawn to business users using OLAP cubes.	All members
04/20/2021	1.3	Created and populated data into dimensions. Used required transformations in SSIS process.	All members
04/21/2021	1.4	Added SCD’s and populated fact tables with data.	All members
04/22/2021	1.5	Created OLAP cubes and extracted data into Excel.	All members
04/26/2021	1.6	Analysis in PowerBI	All members
04/27/2021	1.7	Documentation and Presentation	All members

1. Objective

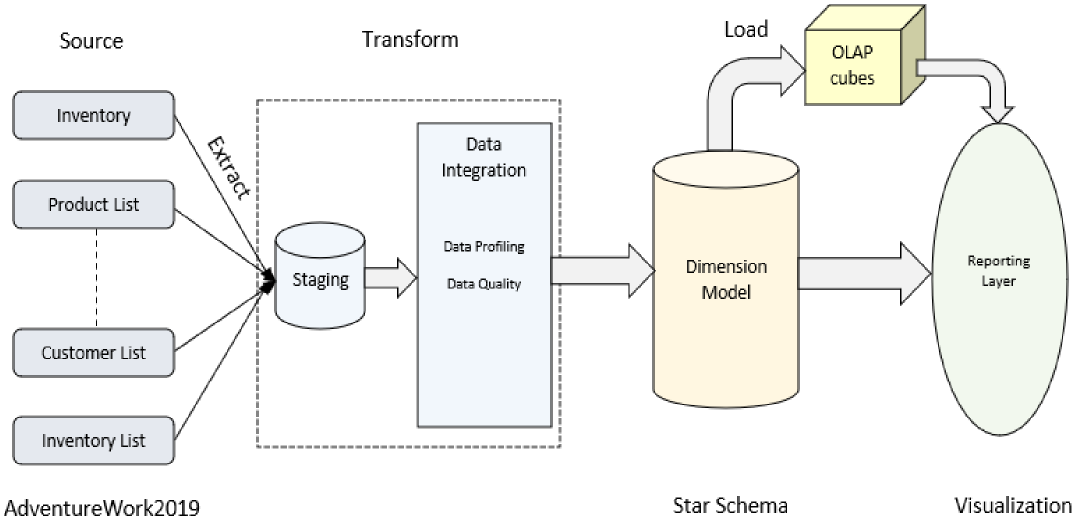
We are using AdventureWorks19 database to understand and analyze the sales by monitoring costs, discounts and selling prices. We are identifying factors that influence sales and analyze the Sales based on different quarters and periods. We understand the sales of products based on their geographical location and analyze the list price and standard cost based on the product category.

2. Data Description

Adventure Work is a fictional company that manufactures the bicycles and other related accessories products such as Helmet, Water bottle, Clothing, Frames, and other premium services. It distributes its products primarily through retail stores located across multiple countries and directly to individual customers through the internet.

AdventureWorks database is meant to be a one-stop solution for all data-related needs of the organization, and with time, more data will be added to the Database. Our process for warehousing and analysis will concern only with the Products, Sales, Customers, along with geographical data.

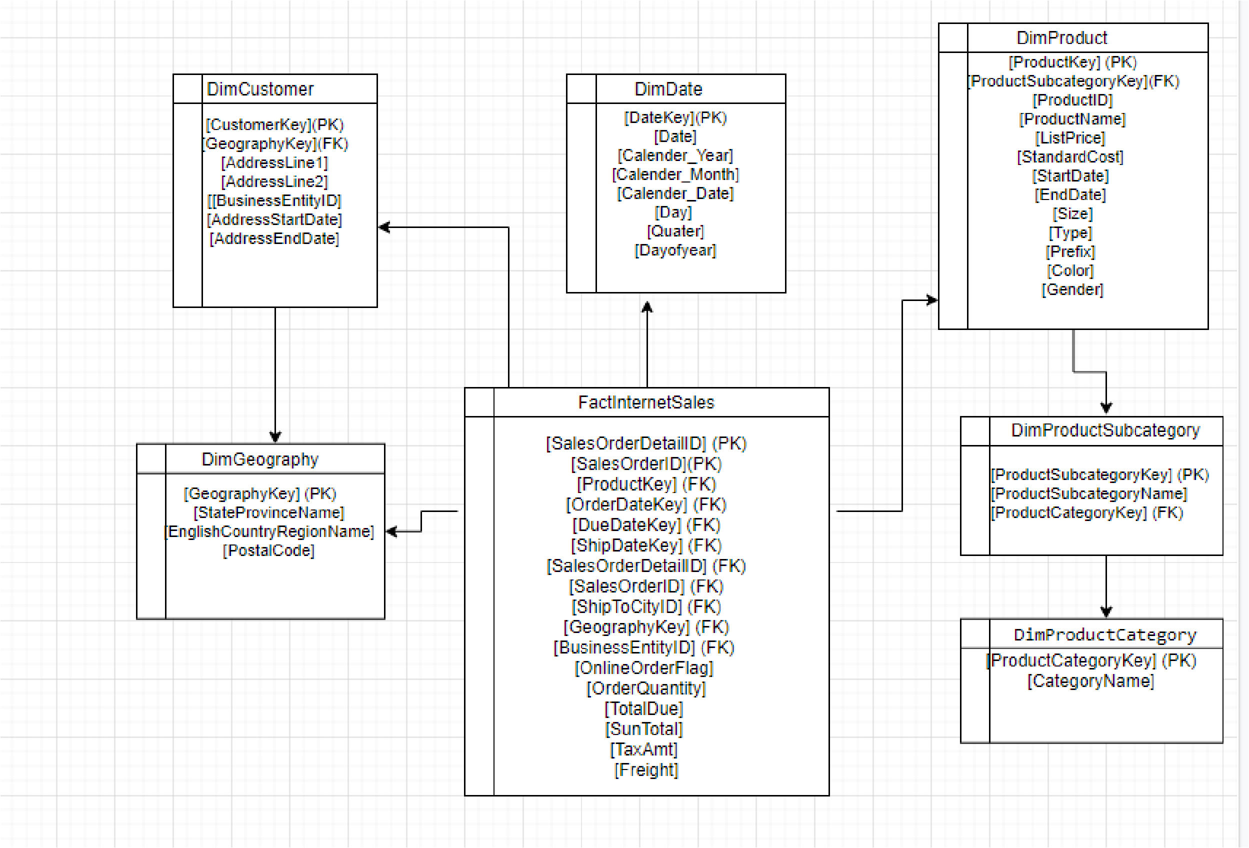
3. Business Intelligence Dataflow Architecture



4. Workflow:

- Data is extracted from the AdventureWork2019 database. We extracted tables according to business purpose and load the data into the centralized platform, i.e., staging area. We do not make any changes to data while loading into the Staging area.
- We apply transformations where required; we also perform integration tasks by applying business rules.
- After the integration process, data is then loaded into the dimensions and facts.
- With SQL Server Analysis Services, we build OLAP cubes and create hierarchies. We deploy the cubes and extract the data into excel files by creating offline cubes to analyze the relations between Product, Sales, Customers by different geographic territories and other factors.
- The final dimensional model is visualized using PowerBI to justify the use of DataWarehouse and answer various questions through visual techniques.

5. Dimension Model:



We built a snowflake schema, which is used to represent multidimensional data.

5.1 Dimensional Model Description

We have worked on one Fact table, i.e, FactInternetSales for analysis and the other fact table (FactInternetSalesReason) did not meet our business scenarios.

- FactInternetSales
  Fact internetsales table consists of 15 columns with one primary key: SalesOrderDetailID. The FactInternetSales table has many to one relation with the dimension tables.
- FactInternetSalesReason:
  FactInternetSalesReason table consists of 3 columns with three primary keys columns; SalesOrderNumber, SalesOrderLineNumber and SalesReasonKey.

In our dimensional model, the dimensions include:

DimProduct, DimDate, DimCustomer, DimGeography, DimProductSubCategory, DimProductCategory

- DimProduct:
  This has 13 columns with one primary key (ProductKey), and DimProductSubCategory reference the DimProduct with the help of ProductSubCategoryKey.
- DimDate:
  This has 8 columns with primary key (DateKey), it deals with the different dates that are present in various tables i.e, Ship Date, Order Date and Due Date.
- DimCustomer: This has 7 columns with one primary key (CustomerKey) and this table has complete details of the customer and has the reference to DimGeography using GeographyKey
- DimGeography has six columns with Geographykey as primary key this section tells about the geographical categories like city ,state, region zip code etc.
- DimProductCategory has two rows with productcategorykey as primary key, it gives information on product categories like Bikes,Components etc.
- Dimproductsubcategory has three columns with productsubcategorykey as primary key and productCategory key as foreign key, this section gives details about subcategories of the product.

5.2 First time load

Until loading files, it's critical to understand what kind of cleaning is needed. We need to double-check the rows after they've been loaded. After the data has been loaded, we would search to see how the rows in the original data source match the rows in the destination.

Staging:

• Raw data is loaded into the staging area from different tables, errors are handled, and we set referential integrity constraints as required.

• Each data flow task undergoes a truncation stage to avoid data loading duplication, even though the data flow task is run several times.

• Current data is loaded followed by an incremental load data from the AdventureWorks Database.

5.3 How do Dimensions and Facts get populated with data?

Dimension tables are used as a collection of reference information about a business. In a snowflake schema model, dimension tables offer descriptive characteristics of the fact table with the help of their attributes.

Fact tables are loaded from the transaction tables such as Production.Product by mapping the fact columns.

The columns of the transaction tables become measures and the primary keys on the transaction table such as Productkey, Customerkey becomes the dimensional key columns of the fact table when loading the data from OLTP transaction table to Fact table.

For our analysis we are storing the details about the Products (Product, ProductCategory, ProductSubcategory) along with the Sales details with respect to the products. We are also creating and modifying the Date Columns according to our business scenarios.

The dimensions are populated with the following source tables from AdventureWork2019

a. DimProduct
   [Production].[Product]
   [Production].[ProductCategory],
   [Production].[ProductSubcategory],
   [Production].[ProductModel]
b. DimCustomer
   [Person].[Person]
   [Person].[PersonPhone]
   [Person].[EmailAddress]
   [Person].[Address]
c. DimGeography
   [Person].[Address]
   [Person].[StateProvince]
   [Person].[CountryRegion]
   [person].[AddressType]
d. DimDate
   [sales].[SalesOrderHeader]
e. DimProductCategory
   [production].[ProductCategory]
f. DimProductSubCategory
   [production].[ProductSubcategory]

6. Slowly Changing Dimensions (SCD)

Slowly Changing Dimensions are used to capture the changing data within the dimension over time. In this project we are implementing Type 2 SCD for Customer and Product dimension. We are adding a new record when the change occurs on the dimension and updating the old record as inactive with the help of enddate attribute.

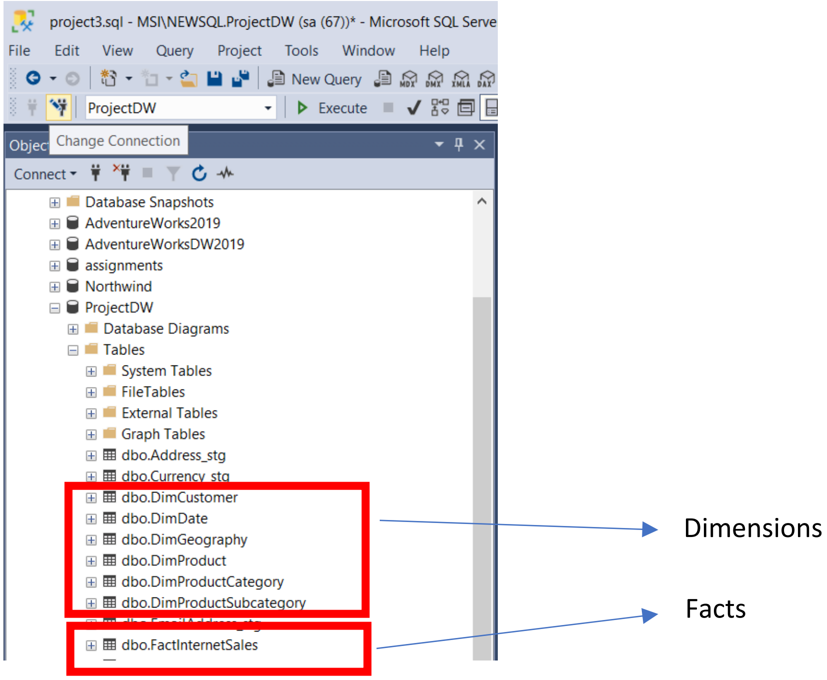
This helps the business to do the historical reporting purposes as the change is captured in the new record.

- AddressLine1 attribute on the DimCustomer dimension.

If Customer updates Address, then we track changes with AddressStartDate and AddressEndDate. (SCD Type2).

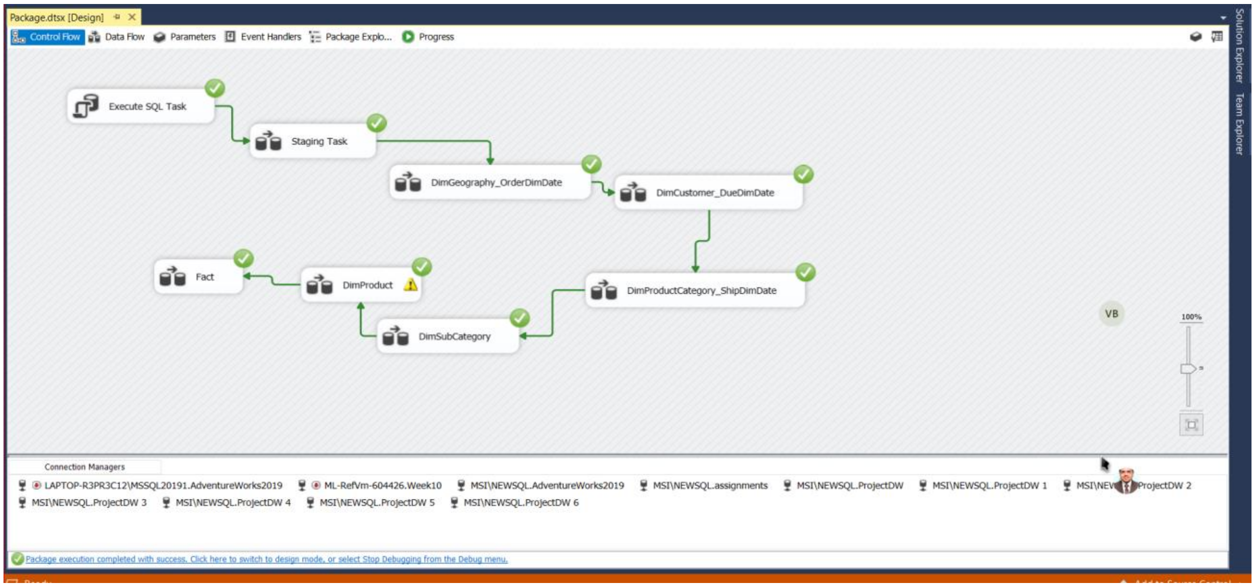
- StandardCost attribute on DimProduct
  If StandardCost is updated, then we track changes with CostStartDate and CostEndDate. (SCD Type2).

7. Fact Tables and Dimension Tables Design- SSIS



Database Snapshots
AdventureWorks2019
AdventureWorksDW2019
assignments
Northwind
ProjectDW
Database Diagrams
Tables
System Tables
FileTables
External Tables
Graph Tables
dbo.Address_stg
dbo.Currency_stg
dbo.DimCustomer
dbo.DimDate
dbo.DimGeography
dbo.DimProduct
dbo.DimProductCategory
dbo.DimProductSubcategory
dbo.EmailAddress_stg
dbo.FactInternetSales

Control Flow:



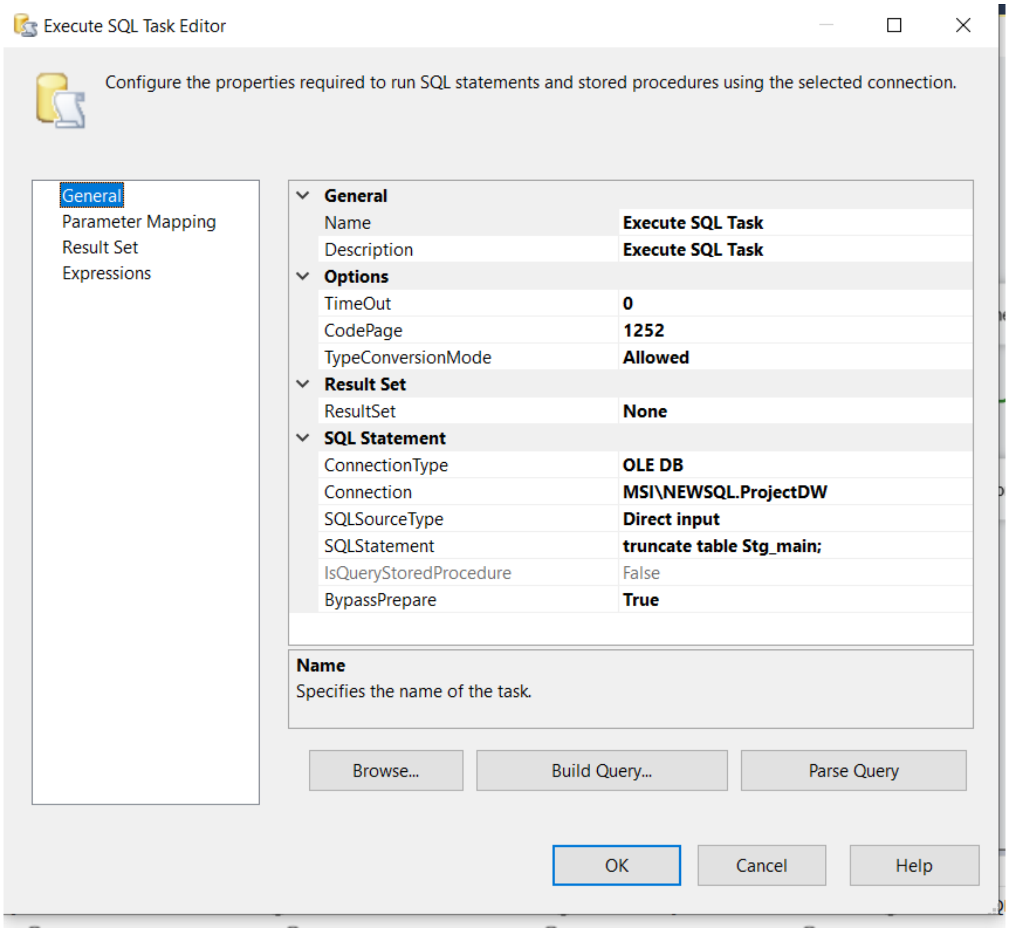
Connection Managers

LAPTOP-R3PR3C12\MSSQL20191.AdventureWorks2019 · ML-RefVm-604426.Week10 · MSI\NEWSQL.AdventureWorks2019 · MSI\NEWSQL.assignments · MSI\NEWSQL.ProjectDW · MSI\NEWSQL.ProjectDW 1 · MSI\NEW ProjectDW 2 · MSI\NEWSQL.ProjectDW 3 · MSI\NEWSQL.ProjectDW 4 · MSI\NEWSQL.ProjectDW 5 · MSI\NEWSQL.ProjectDW 6

Package execution completed with success. Click here to switch to design mode, or select Stop Debugging from the Debug menu.

Execute SQL Task:

We are cleaning our staging area for every process.



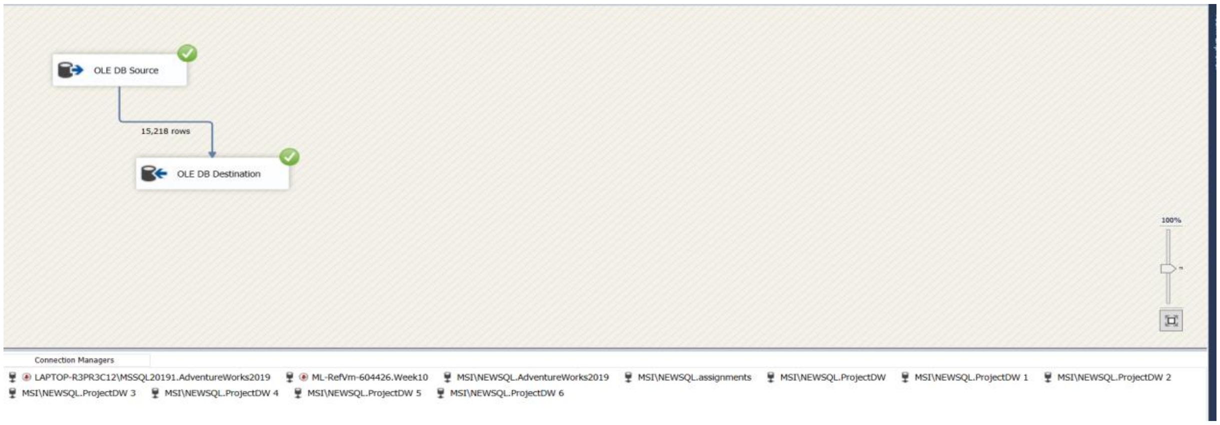
Configure the properties required to run SQL statements and stored procedures using the selected connection.

General	
Name	**Execute SQL Task**
Description	**Execute SQL Task**
Options	
TimeOut	**0**
CodePage	**1252**
TypeConversionMode	**Allowed**
Result Set	
ResultSet	**None**
SQL Statement	
ConnectionType	**OLE DB**
Connection	**MSI\NEWSQL.ProjectDW**
SQLSourceType	**Direct input**
SQLStatement	**truncate table Stg_main;**
IsQueryStoredProcedure	False
BypassPrepare	**True**

Name
Specifies the name of the task.

Staging Area:

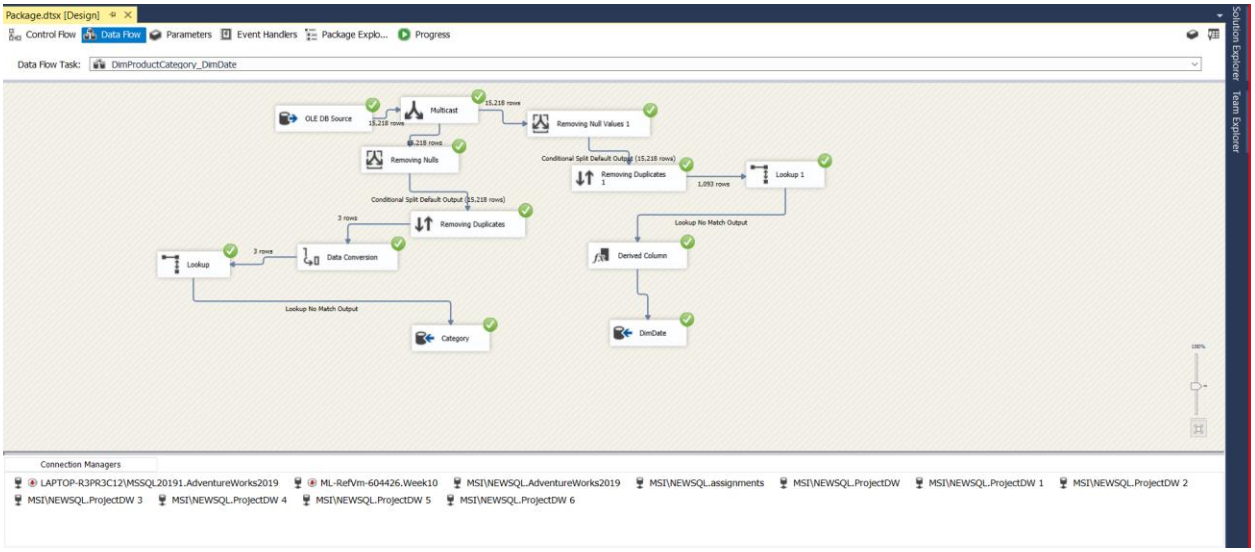
We are extracting data from Source table from AdventureWorks2019 and loading into the staging area



Following workflows shows tables being populated after performing cleansing and transformations:

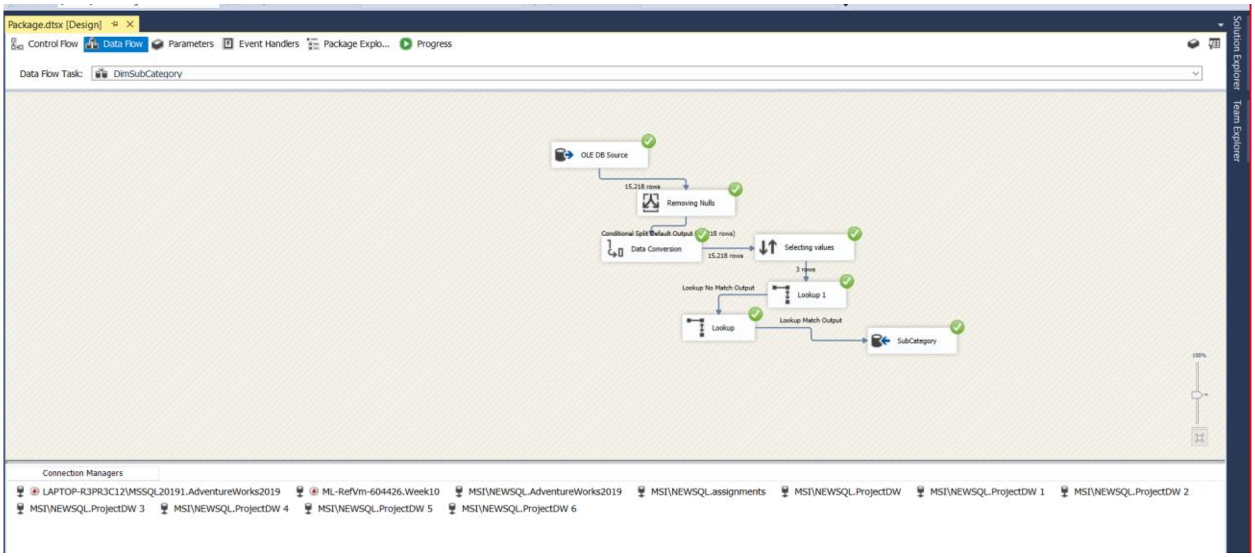
1. **DimProduct Category and DimDate**

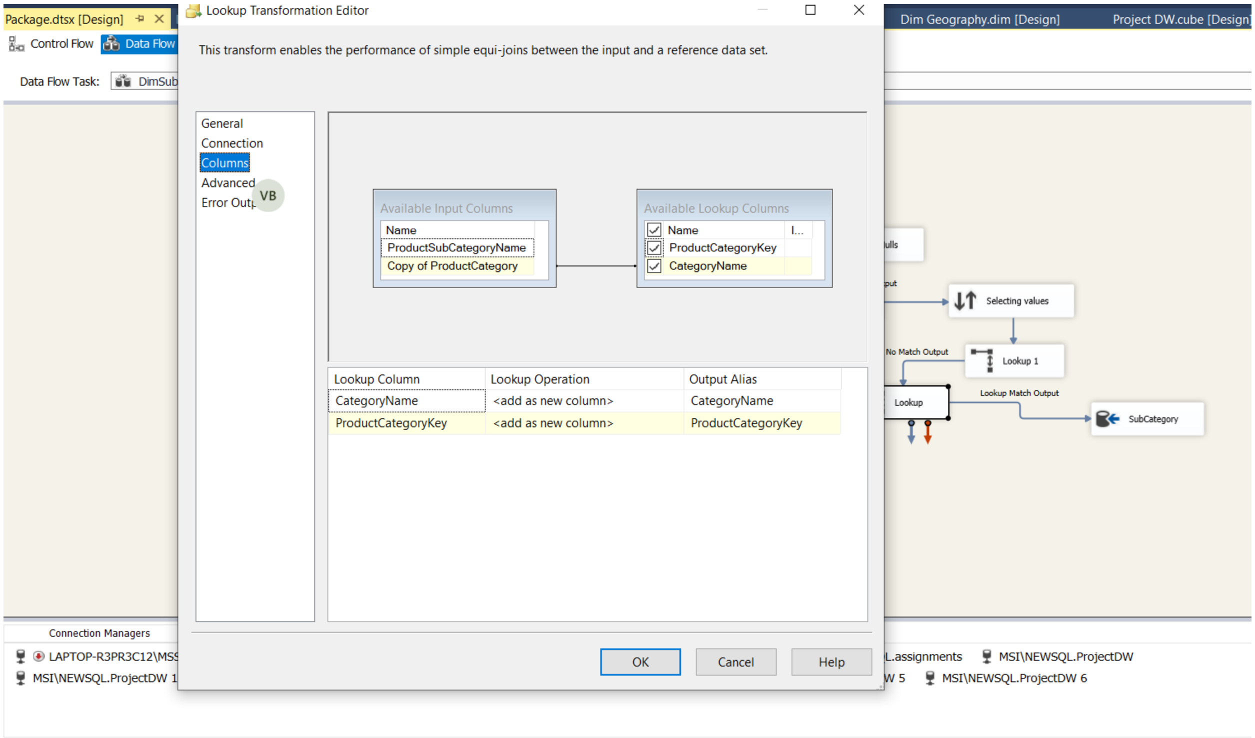
   We are populating DimProductCategory and DimDate(for ShipDate) by using a multicast, checking the nulls and removing nulls. We then further remove the duplicates and do data conversions as required and add lookups to get no match outputs only.



2. DimSubCategory

We are populating DimSubCategory from the staging area. We are performing the required transformations and adding a lookup to get the foreign keys of 'ProductCategoryKey' from the DimProductCategory table.



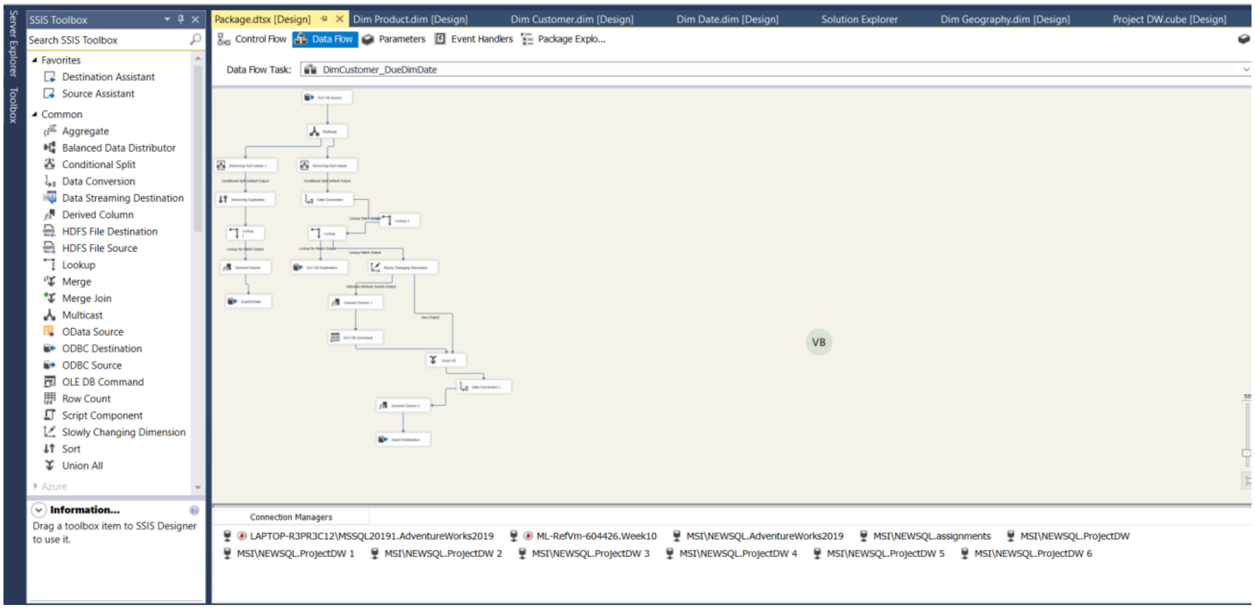


This transform enables the performance of simple equi-joins between the input and a reference data set.

Lookup Column	Lookup Operation	Output Alias
CategoryName	<add as new column>	CategoryName
ProductCategoryKey	<add as new column>	ProductCategoryKey

3. DimCustomer and DimDate:

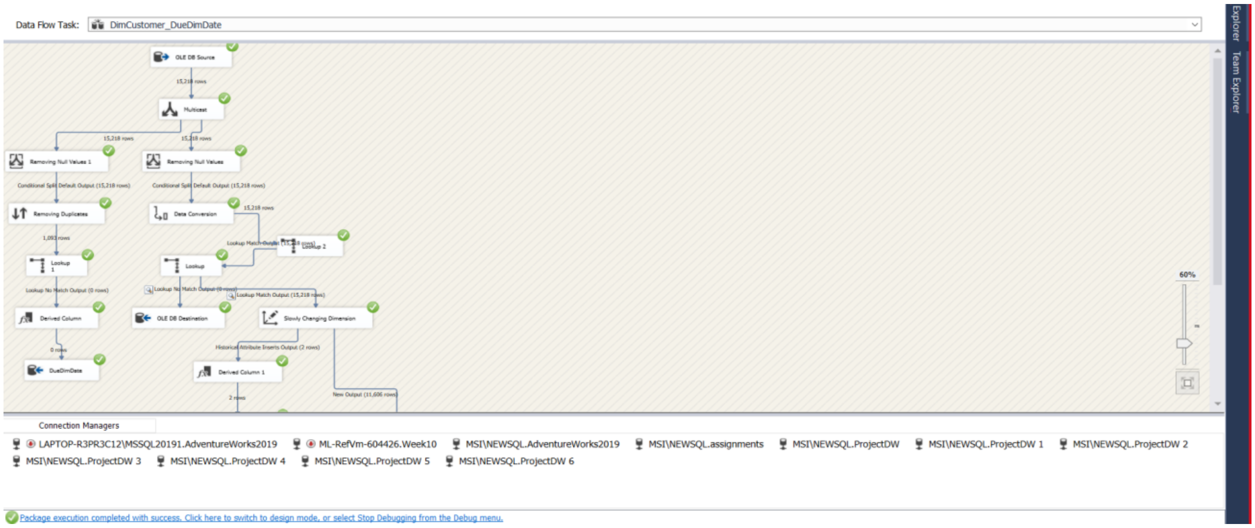
We have populated DimCustomer and DimDate(for DueDate) from the staging. We used a multicast, checking the nulls and removing nulls. We then further remove the duplicates and do data conversions as required and add lookups to get no match outputs only. We have used SCD on AddressLine1 to track changes when a update occurs.



Drag a toolbox item to SSIS Designer to use it.

Data Flow Task: DimCustomer_DueDimDate

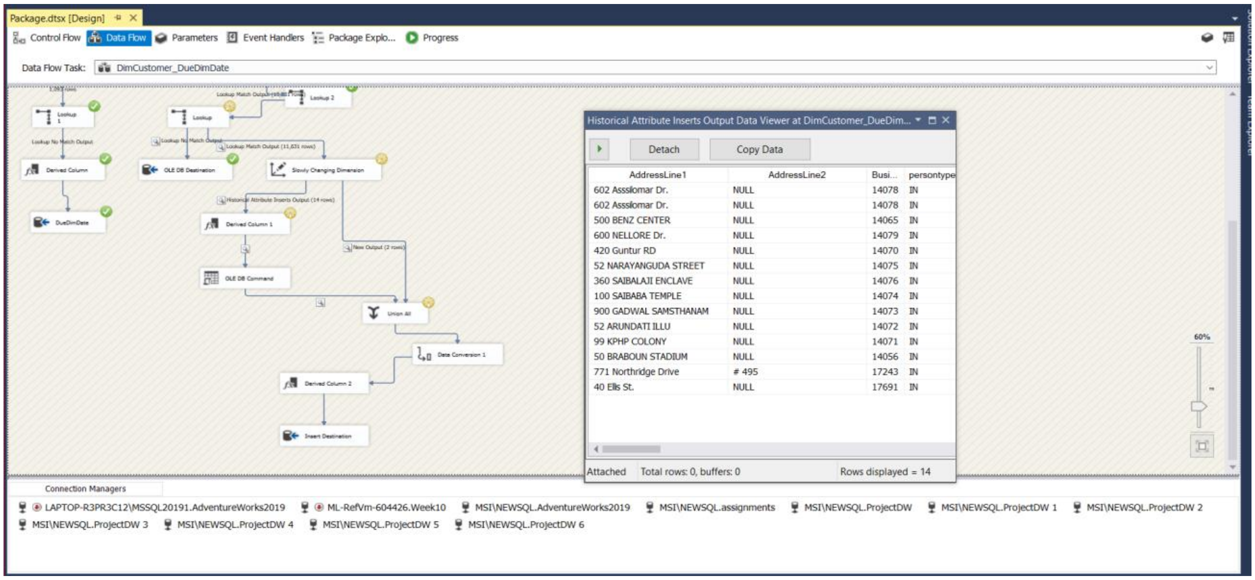
LAPTOP-R3PR3C12\MSSQL20191.AdventureWorks2019 ML-RefVm-604426.Week10 MSI\NEWSQL.AdventureWorks2019 MSI\NEWSQL.assignments MSI\NEWSQL.ProjectDW
MSI\NEWSQL.ProjectDW 1 MSI\NEWSQL.ProjectDW 2 MSI\NEWSQL.ProjectDW 3 MSI\NEWSQL.ProjectDW 4 MSI\NEWSQL.ProjectDW 5 MSI\NEWSQL.ProjectDW 6



Data Flow Task: DimCustomer_DueDimDate

LAPTOP-R3PR3C12\MSSQL20191.AdventureWorks2019 ML-RefVm-604426.Week10 MSI\NEWSQL.AdventureWorks2019 MSI\NEWSQL.assignments MSI\NEWSQL.ProjectDW MSI\NEWSQL.ProjectDW 1 MSI\NEWSQL.ProjectDW 2
MSI\NEWSQL.ProjectDW 3 MSI\NEWSQL.ProjectDW 4 MSI\NEWSQL.ProjectDW 5 MSI\NEWSQL.ProjectDW 6

Package execution completed with success. Click here to switch to design mode, or select Stop Debugging from the Debug menu.

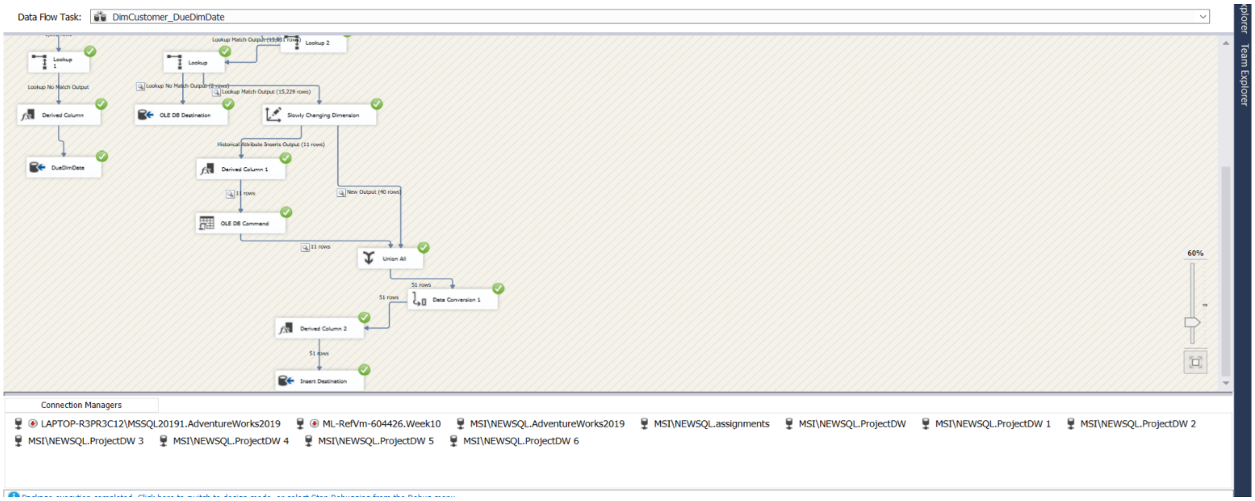


Historical Attribute Inserts Output Data Viewer at DimCustomer_DueDim...

Detach | Copy Data

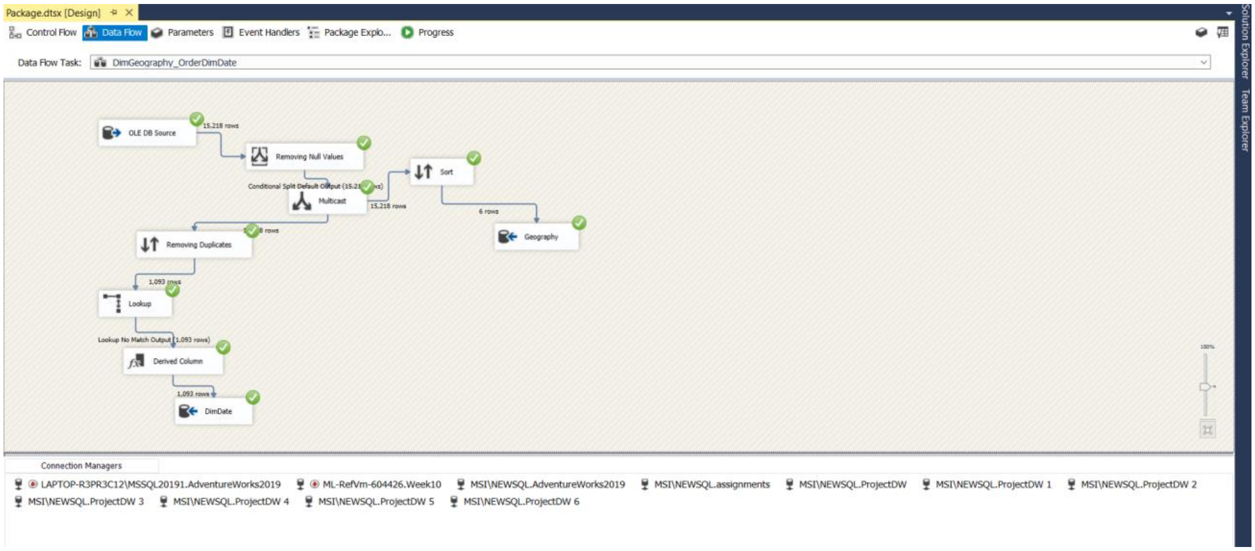
AddressLine1	AddressLine2	Busi...	persontype
602 Asssiomar Dr.	NULL	14078	IN
602 Asssiomar Dr.	NULL	14078	IN
500 BENZ CENTER	NULL	14065	IN
600 NELLORE Dr.	NULL	14079	IN
420 Guntur RD	NULL	14070	IN
52 NARAYANGUDA STREET	NULL	14075	IN
360 SAIBALAJI ENCLAVE	NULL	14076	IN
100 SAIBABA TEMPLE	NULL	14074	IN
900 GADWAL SAMSTHANAM	NULL	14073	IN
52 ARUNDATI ILLU	NULL	14072	IN
99 KPHP COLONY	NULL	14071	IN
50 BRABOUN STADIUM	NULL	14056	IN
771 Northridge Drive	# 495	17243	IN
40 Ellis St.	NULL	17691	IN

Attached | Total rows: 0, buffers: 0 | Rows displayed = 14



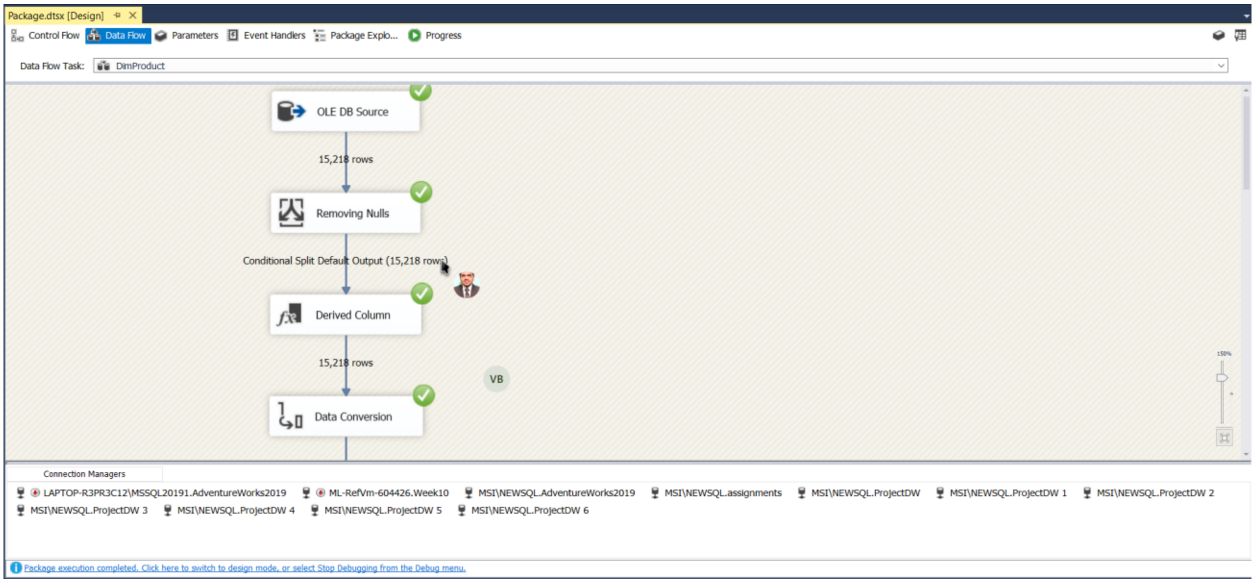
4. Geography and DimDate

We are populating DimProductCategory and DimDate(for OrderDate) by using a multicast, checking the nulls and removing nulls. We then further remove the duplicates and do data conversions as required and add lookups to get no match outputs only.

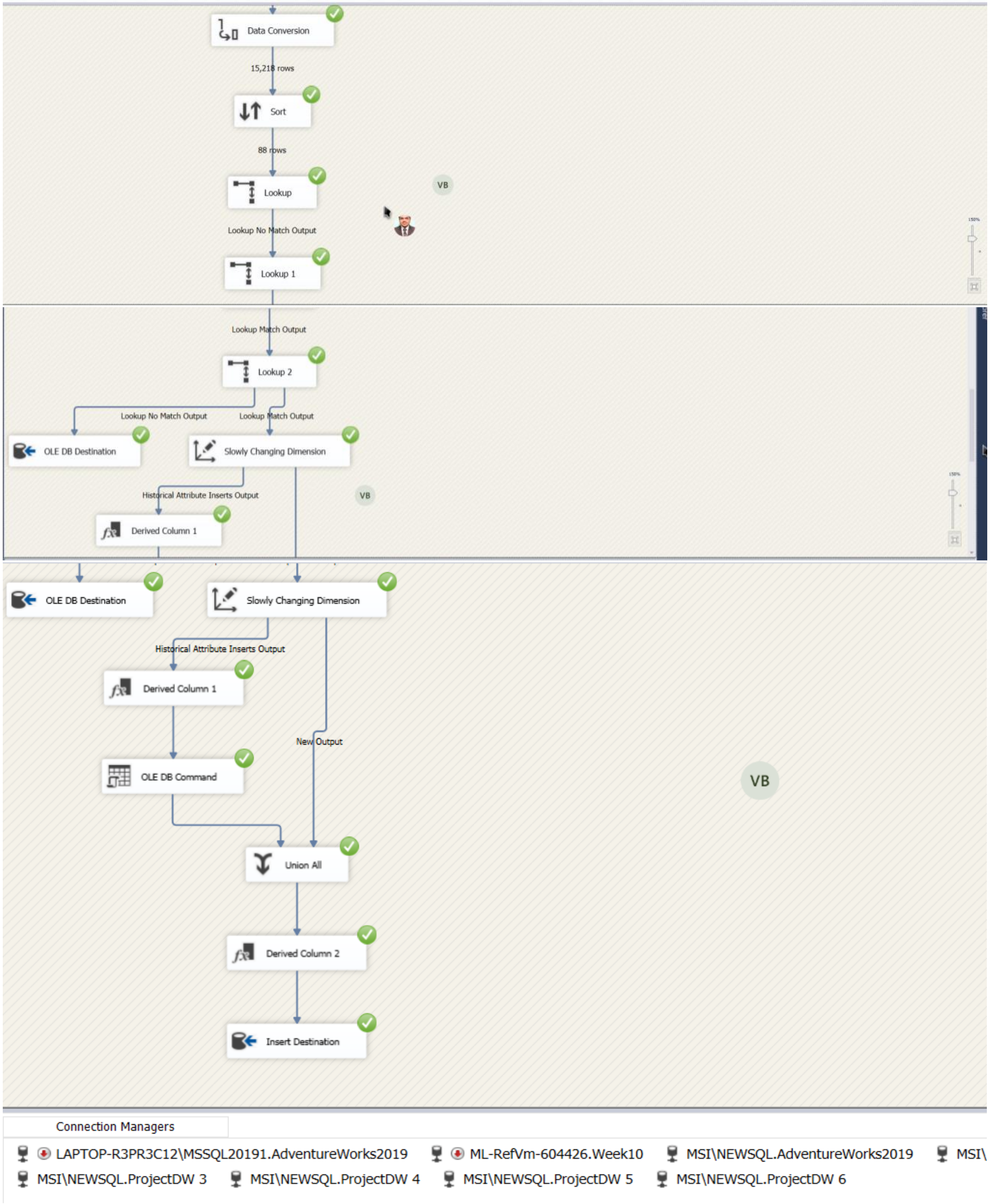


5. DimProduct:

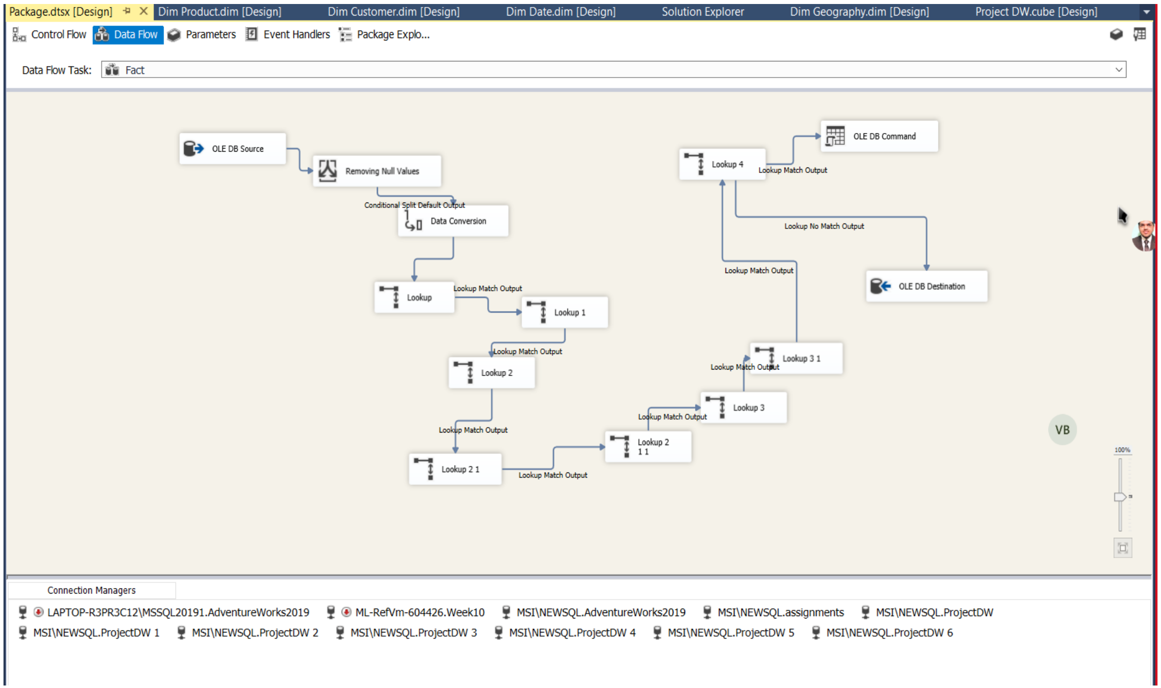
We are populating DimProduct from the staging. We are performing the required transformations and adding a lookup to get the foreign keys of ‘ProductSubCategoryKey’ from the DimProductSubCategory table. We are adding an SCD to monitor any changes occurring in StandardCost.

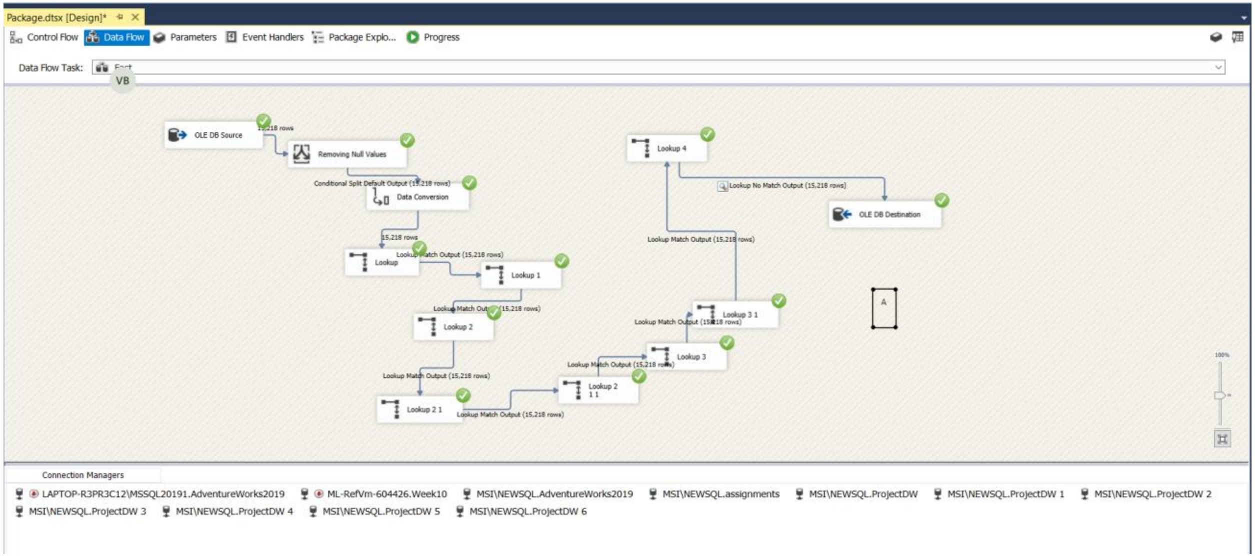


Package execution completed. Click here to switch to design mode, or select Stop Debugging from the Debug menu.


Data Conversion
15,218 rows
Sort
88 rows
Lookup
VB
Lookup No Match Output
Lookup 1
Lookup Match Output
Lookup 2
Lookup No Match Output
Lookup Match Output
OLE DB Destination
Slowly Changing Dimension
Historical Attribute Inserts Output
VB
Derived Column 1
OLE DB Destination
Slowly Changing Dimension
Historical Attribute Inserts Output
Derived Column 1
New Output
OLE DB Command
VB
Union All
Derived Column 2
Insert Destination
Connection Managers
LAPTOP-R3PR3C12\MSSQL20191.AdventureWorks2019
ML-RefVm-604426.Week10
MSI\NEWSQL.AdventureWorks2019
MSI\
MSI\NEWSQL.ProjectDW 3
MSI\NEWSQL.ProjectDW 4
MSI\NEWSQL.ProjectDW 5
MSI\NEWSQL.ProjectDW 6

1. Fact table:

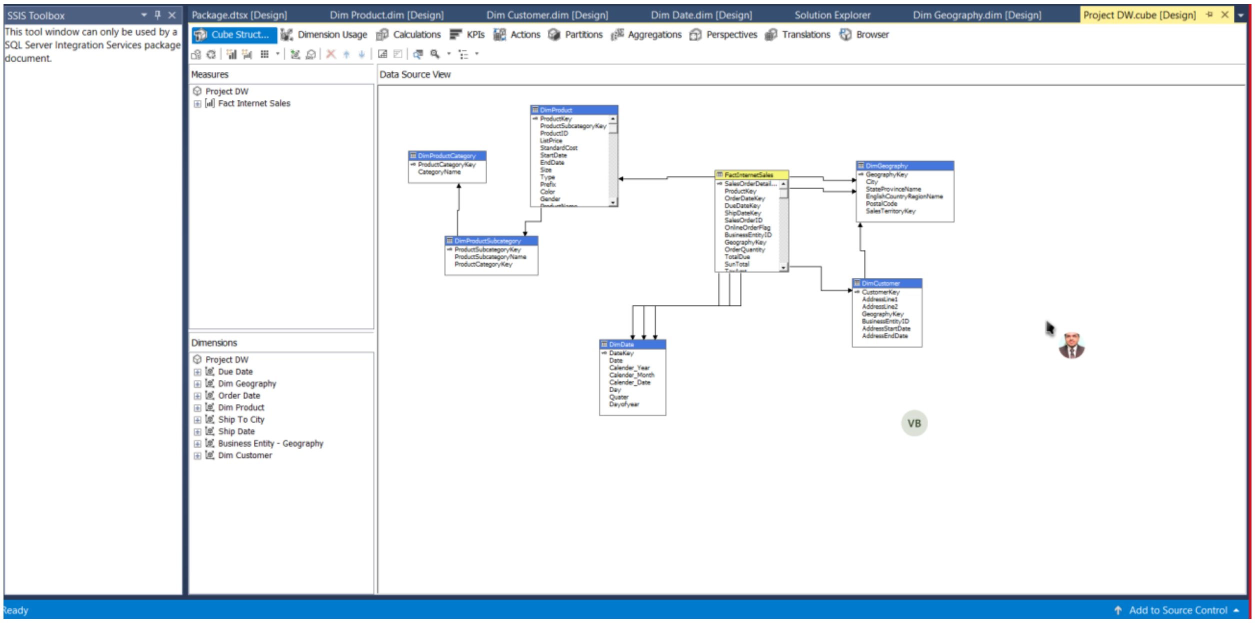

Package.dtsx [Design]
Dim Product.dim [Design]
Dim Customer.dim [Design]
Dim Date.dim [Design]
Solution Explorer
Dim Geography.dim [Design]
Project DW.cube [Design]
Control Flow
Data Flow
Parameters
Event Handlers
Package Explo...
Data Flow Task:
Fact
OLE DB Source
Removing Null Values
Conditional Split Default Output
Data Conversion
Lookup
Lookup Match Output
Lookup 1
Lookup Match Output
Lookup 2
Lookup Match Output
Lookup 2 1
Lookup Match Output
Lookup 2 1 1
Lookup Match Output
Lookup 3
Lookup Match Output
Lookup 3 1
Lookup Match Output
Lookup 4
Lookup Match Output
OLE DB Command
Lookup No Match Output
OLE DB Destination
VB
100%
Connection Managers
LAPTOP-R3PR3C12\MSSQL20191.AdventureWorks2019
ML-RefVm-604426.Week10
MSI\NEWSQL.AdventureWorks2019
MSI\NEWSQL.assignments
MSI\NEWSQL.ProjectDW
MSI\NEWSQL.ProjectDW 1
MSI\NEWSQL.ProjectDW 2
MSI\NEWSQL.ProjectDW 3
MSI\NEWSQL.ProjectDW 4
MSI\NEWSQL.ProjectDW 5
MSI\NEWSQL.ProjectDW 6

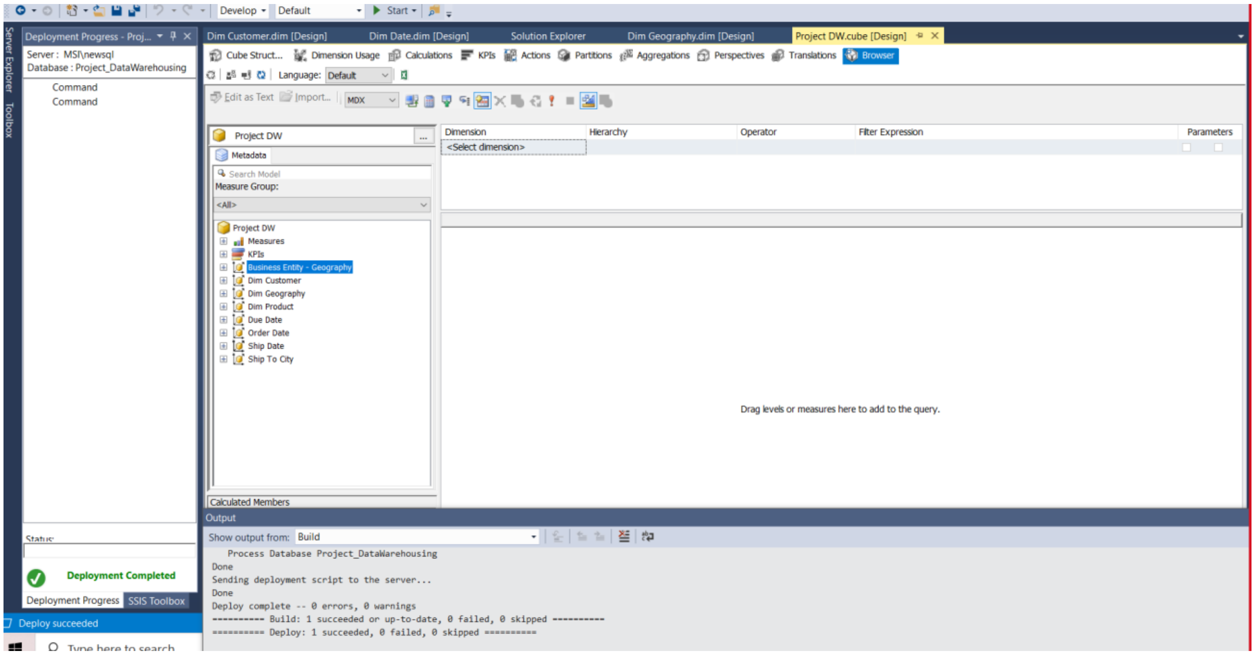

Package.dtsx [Design]*
Control Flow
Data Flow
Parameters
Event Handlers
Package Explo...
Progress
Data Flow Task:
Fact
VB
OLE DB Source
15,218 rows
Removing Null Values
Conditional Split Default Output (15,218 rows)
Data Conversion
15,218 rows
Lookup
Lookup Match Output (15,218 rows)
Lookup 1
Lookup Match Output (15,218 rows)
Lookup 2
Lookup Match Output (15,218 rows)
Lookup 2 1
Lookup Match Output (15,218 rows)
Lookup 2 1 1
Lookup Match Output (15,218 rows)
Lookup 3
Lookup Match Output (15,218 rows)
Lookup 3 1
Lookup Match Output (15,218 rows)
Lookup 4
Lookup No Match Output (15,218 rows)
OLE DB Destination
A
100%
Connection Managers
LAPTOP-R3PR3C12\MSSQL20191.AdventureWorks2019
ML-RefVm-604426.Week10
MSI\NEWSQL.AdventureWorks2019
MSI\NEWSQL.assignments
MSI\NEWSQL.ProjectDW
MSI\NEWSQL.ProjectDW 1
MSI\NEWSQL.ProjectDW 2
MSI\NEWSQL.ProjectDW 3
MSI\NEWSQL.ProjectDW 4
MSI\NEWSQL.ProjectDW 5
MSI\NEWSQL.ProjectDW 6

8. Cube Design-SSAS:

Online analytical processing cubes are used to perform multidimensional analysis at high speeds on large volumes of data from a data warehouse. We created two hierarchies for Dimgeography and Dimdate to perform analysis and extract insights at a higher granular level.

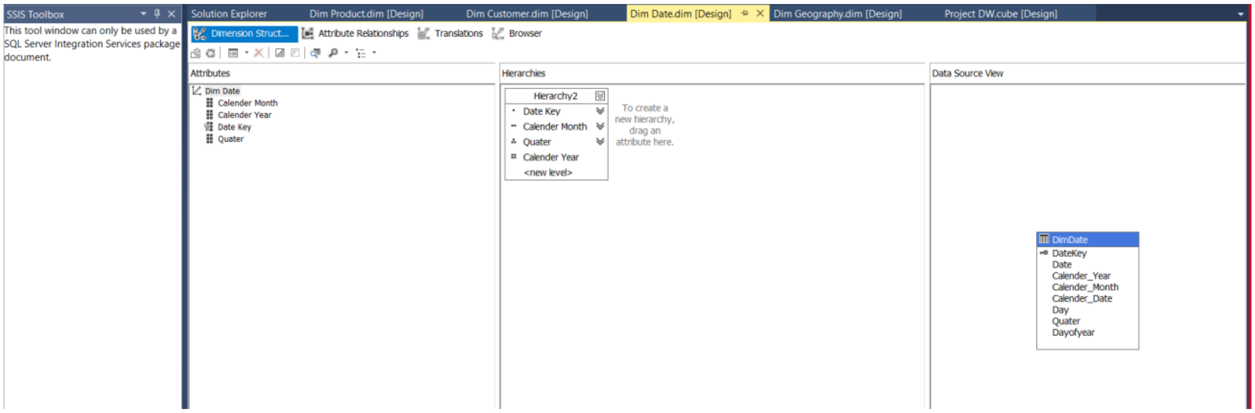
OLAP CUBE:





```
    Process Database Project_DataWarehousing
Done
Sending deployment script to the server...
Done
Deploy complete -- 0 errors, 0 warnings
========== Build: 1 succeeded or up-to-date, 0 failed, 0 skipped ==========
========== Deploy: 1 succeeded, 0 failed, 0 skipped ==========
```

Hierarchy of DimDate:



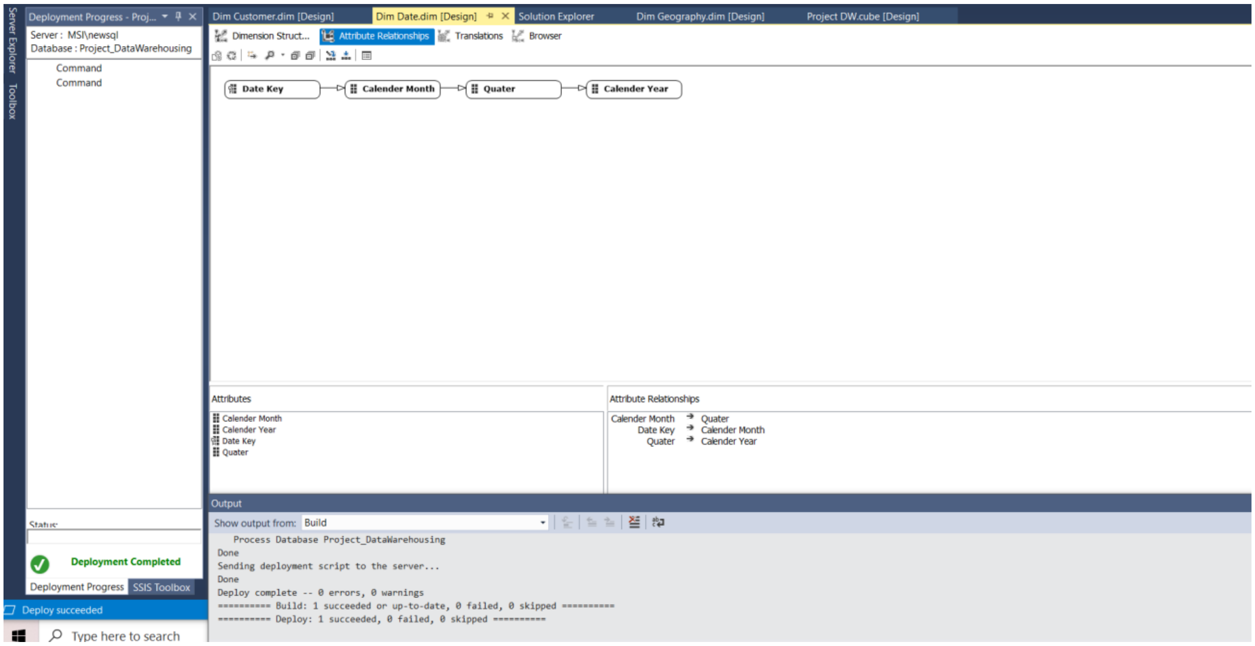
This tool window can only be used by a SQL Server Integration Services package document.

Dim Date
- Calender Month
- Calender Year
- Date Key
- Quater

Hierarchy2
- Date Key
- Calender Month
- Quater
- Calender Year
- <new level>

To create a new hierarchy, drag an attribute here.

DimDate
- DateKey
- Date
- Calender_Year
- Calender_Month
- Calender_Date
- Day
- Quater
- Dayofyear



Server : MSI\newsql
Database : Project_DataWarehousing

Command
Command

Status:

Deployment Completed

Deploy succeeded

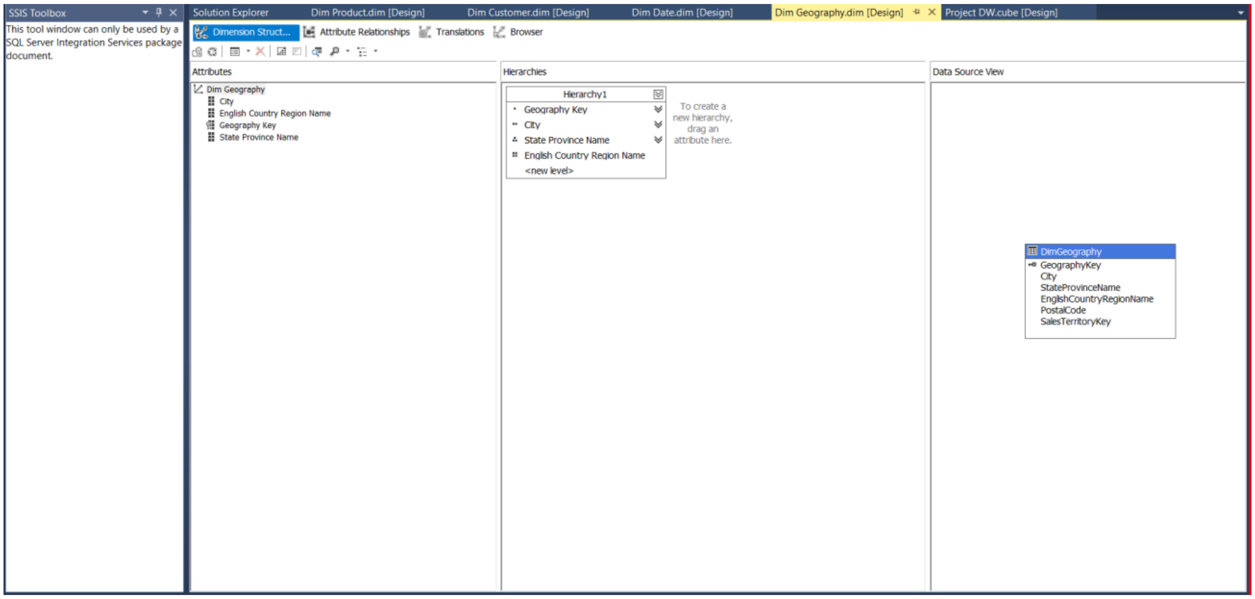
Calender Month
Calender Year
Date Key
Quater

Calender Month	→	Quater
Date Key	→	Calender Month
Quater	→	Calender Year

Show output from: Build

```
   Process Database Project_DataWarehousing
Done
Sending deployment script to the server...
Done
Deploy complete -- 0 errors, 0 warnings
========== Build: 1 succeeded or up-to-date, 0 failed, 0 skipped ==========
========== Deploy: 1 succeeded, 0 failed, 0 skipped ==========
```

Hierarchy of DimGeography



This tool window can only be used by a SQL Server Integration Services package document.

Attributes:

- Dim Geography
  - City
  - English Country Region Name
  - Geography Key
  - State Province Name

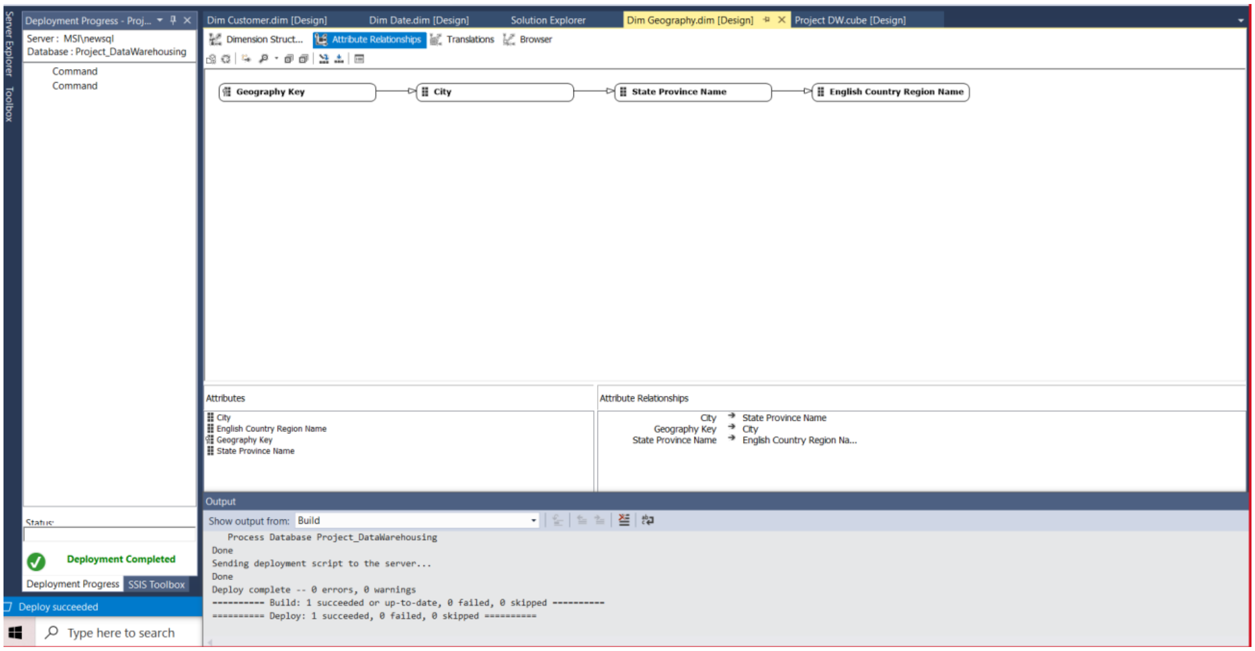
Hierarchy1:

- Geography Key
- City
- State Province Name
- English Country Region Name
- <new level>

To create a new hierarchy, drag an attribute here.

DimGeography:

- GeographyKey
- City
- StateProvinceName
- EnglishCountryRegionName
- PostalCode
- SalesTerritoryKey



Server : MSI\newsql
Database : Project_DataWarehousing

Command
Command

Status:

Deployment Completed

Deploy succeeded

Attributes:

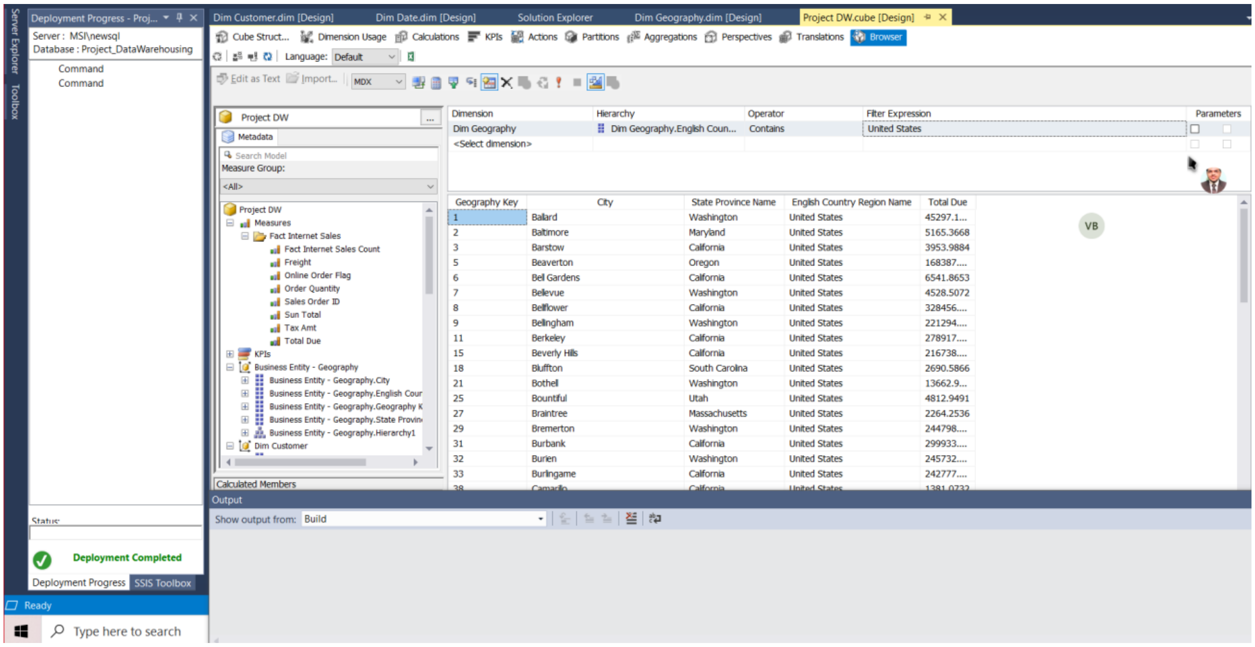
- City
- English Country Region Name
- Geography Key
- State Province Name

Attribute Relationships:

City	→	State Province Name
Geography Key	→	City
State Province Name	→	English Country Region Na...

Show output from: Build

```
   Process Database Project_DataWarehousing
Done
Sending deployment script to the server...
Done
Deploy complete -- 0 errors, 0 warnings
========== Build: 1 succeeded or up-to-date, 0 failed, 0 skipped ==========
========== Deploy: 1 succeeded, 0 failed, 0 skipped ==========
```

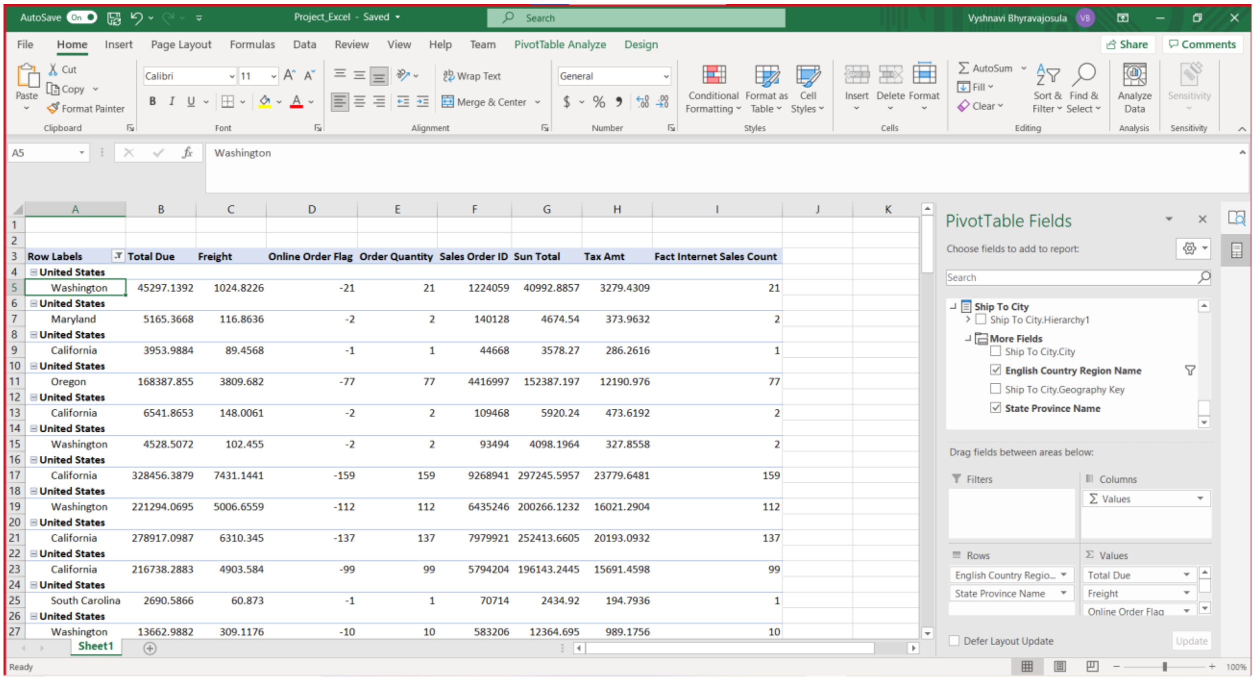


Dimension	Hierarchy	Operator	Filter Expression
Dim Geography	Dim Geography.English Coun...	Contains	United States
<Select dimension>			

Geography Key	City	State Province Name	English Country Region Name	Total Due
1	Ballard	Washington	United States	45297.1...
2	Baltimore	Maryland	United States	5165.3668
3	Barstow	California	United States	3953.9884
5	Beaverton	Oregon	United States	168387....
6	Bell Gardens	California	United States	6541.8653
7	Bellevue	Washington	United States	4528.5072
8	Bellflower	California	United States	328456....
9	Bellingham	Washington	United States	221294....
11	Berkeley	California	United States	278917....
15	Beverly Hills	California	United States	216738....
18	Bluffton	South Carolina	United States	2690.5866
21	Bothell	Washington	United States	13662.9...
25	Bountiful	Utah	United States	4812.9491
27	Braintree	Massachusetts	United States	2264.2536
29	Bremerton	Washington	United States	244798....
31	Burbank	California	United States	299933....
32	Burien	Washington	United States	245732....
33	Burlingame	California	United States	242777....

Deployment Completed

Creating Offline files using Excel:



A5 Washington

	A	B	C	D	E	F	G	H	I
3	Row Labels	Total Due	Freight	Online Order Flag	Order Quantity	Sales Order ID	Sun Total	Tax Amt	Fact Internet Sales Count
4	⊟ United States								
5	Washington	45297.1392	1024.8226	-21	21	1224059	40992.8857	3279.4309	21
6	⊟ United States								
7	Maryland	5165.3668	116.8636	-2	2	140128	4674.54	373.9632	2
8	⊟ United States								
9	California	3953.9884	89.4568	-1	1	44668	3578.27	286.2616	1
10	⊟ United States								
11	Oregon	168387.855	3809.682	-77	77	4416997	152387.197	12190.976	77
12	⊟ United States								
13	California	6541.8653	148.0061	-2	2	109468	5920.24	473.6192	2
14	⊟ United States								
15	Washington	4528.5072	102.455	-2	2	93494	4098.1964	327.8558	2
16	⊟ United States								
17	California	328456.3879	7431.1441	-159	159	9268941	297245.5957	23779.6481	159
18	⊟ United States								
19	Washington	221294.0695	5006.6559	-112	112	6435246	200266.1232	16021.2904	112
20	⊟ United States								
21	California	278917.0987	6310.345	-137	137	7979921	252413.6605	20193.0932	137
22	⊟ United States								
23	California	216738.2883	4903.584	-99	99	5794204	196143.2445	15691.4598	99
24	⊟ United States								
25	South Carolina	2690.5866	60.873	-1	1	70714	2434.92	194.7936	1
26	⊟ United States								
27	Washington	13662.9882	309.1176	-10	10	583206	12364.695	989.1756	10

PivotTable Fields

Choose fields to add to report:

- Ship To City
  - □ Ship To City.Hierarchy1
  - More Fields
    - □ Ship To City.City
    - ☑ English Country Region Name
    - □ Ship To City.Geography Key
    - ☑ State Province Name

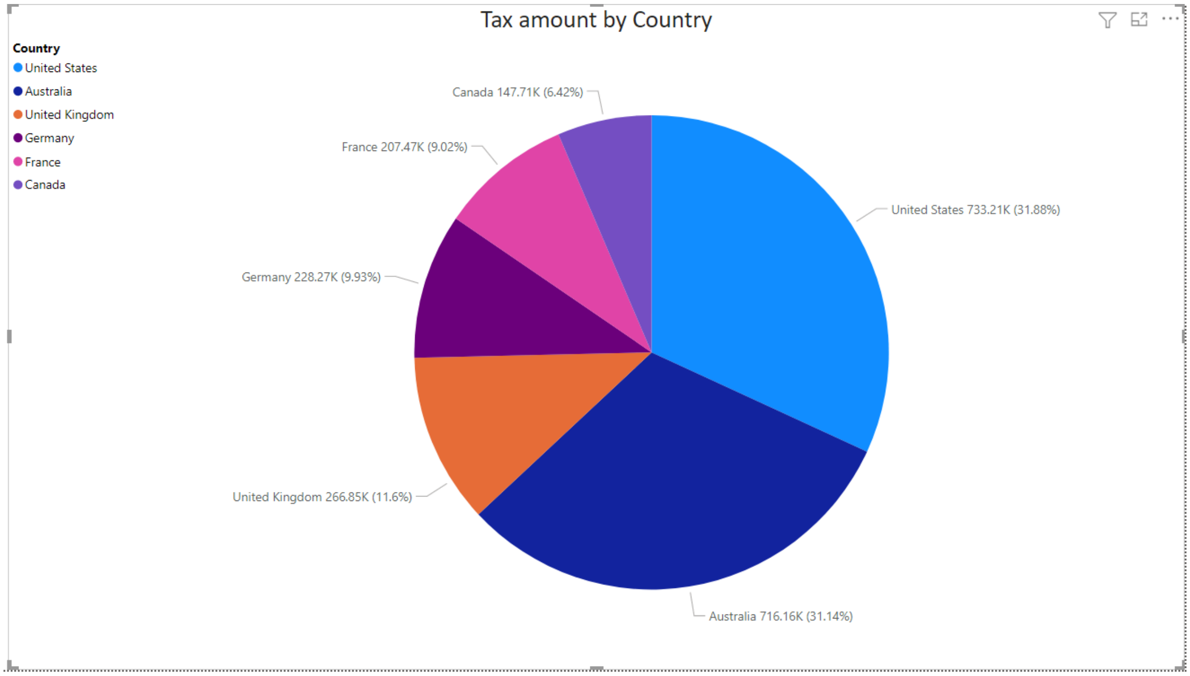
Drag fields between areas below:

- Filters
- Columns: Σ Values
- Rows: English Country Regio..., State Province Name
- Values: Total Due, Freight, Online Order Flag

□ Defer Layout Update

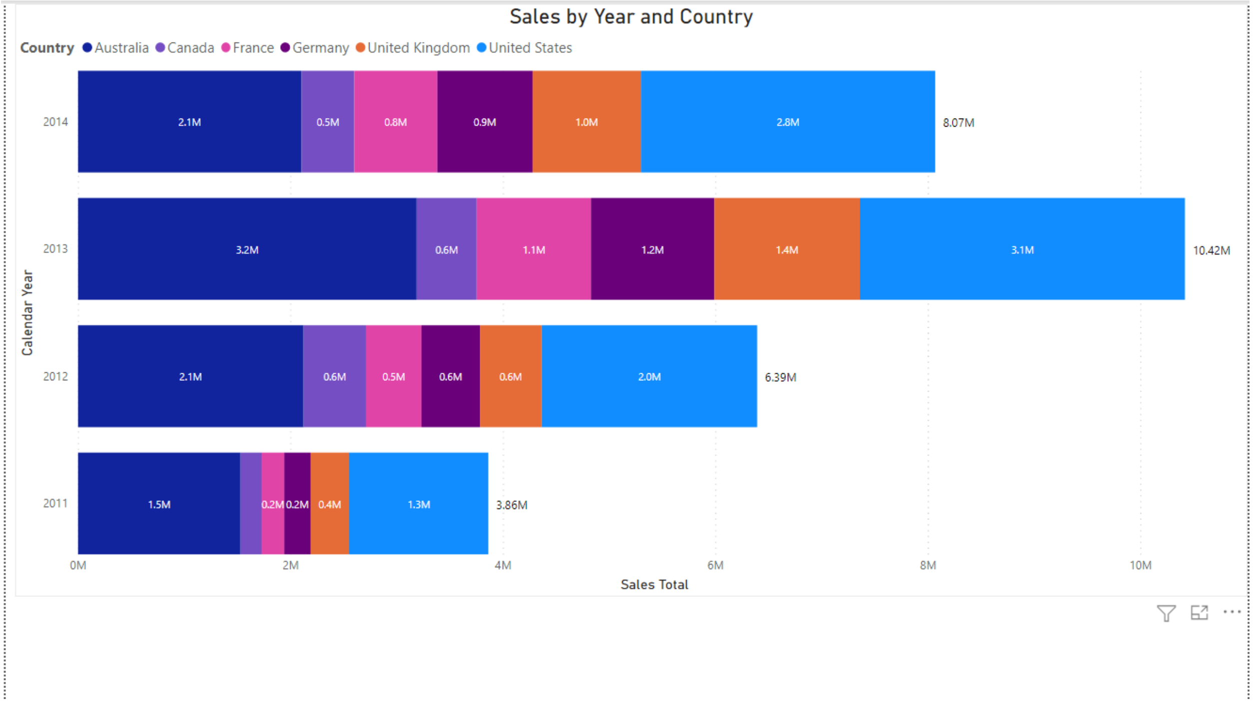
9. Analysis and Visualizations Using PowerBI:

1. Tax Amount by Country:



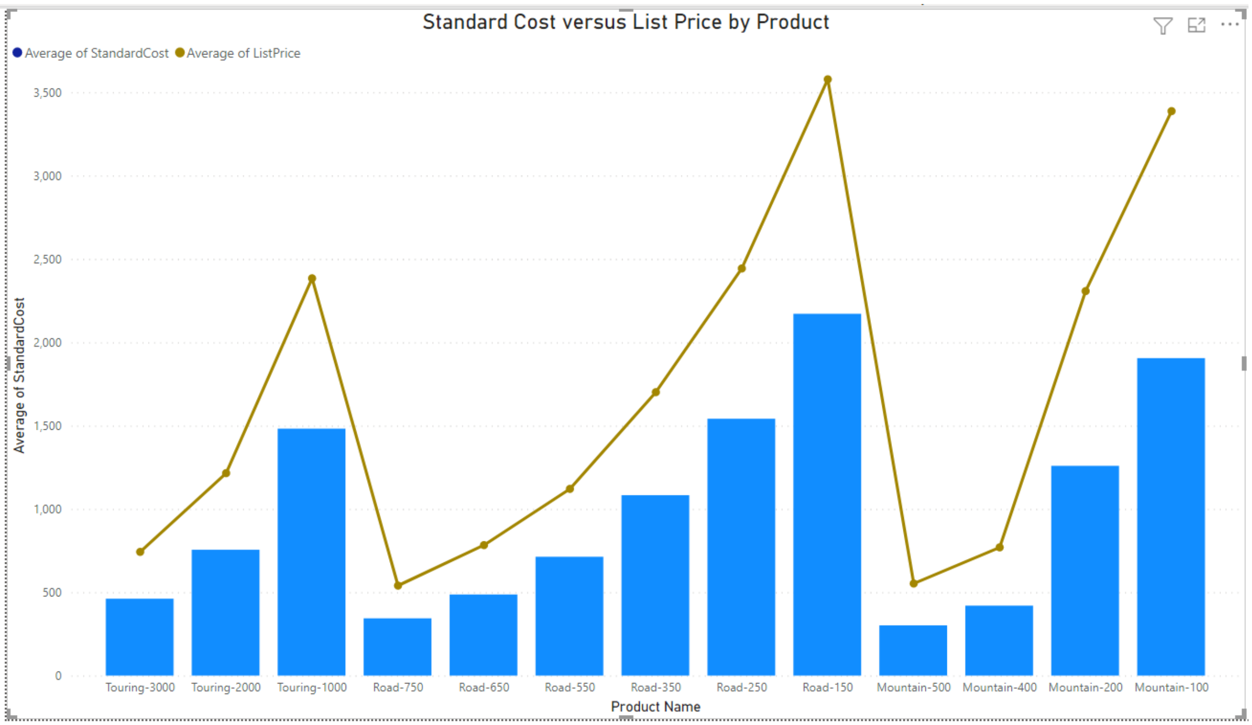
The above pie chart Company has paid highest amount of tax in USA and Australia. Both sharing total 62% of company's total tax payment in between. While other countries account for less than 10% of tax.

2. Sales by Year and Country:



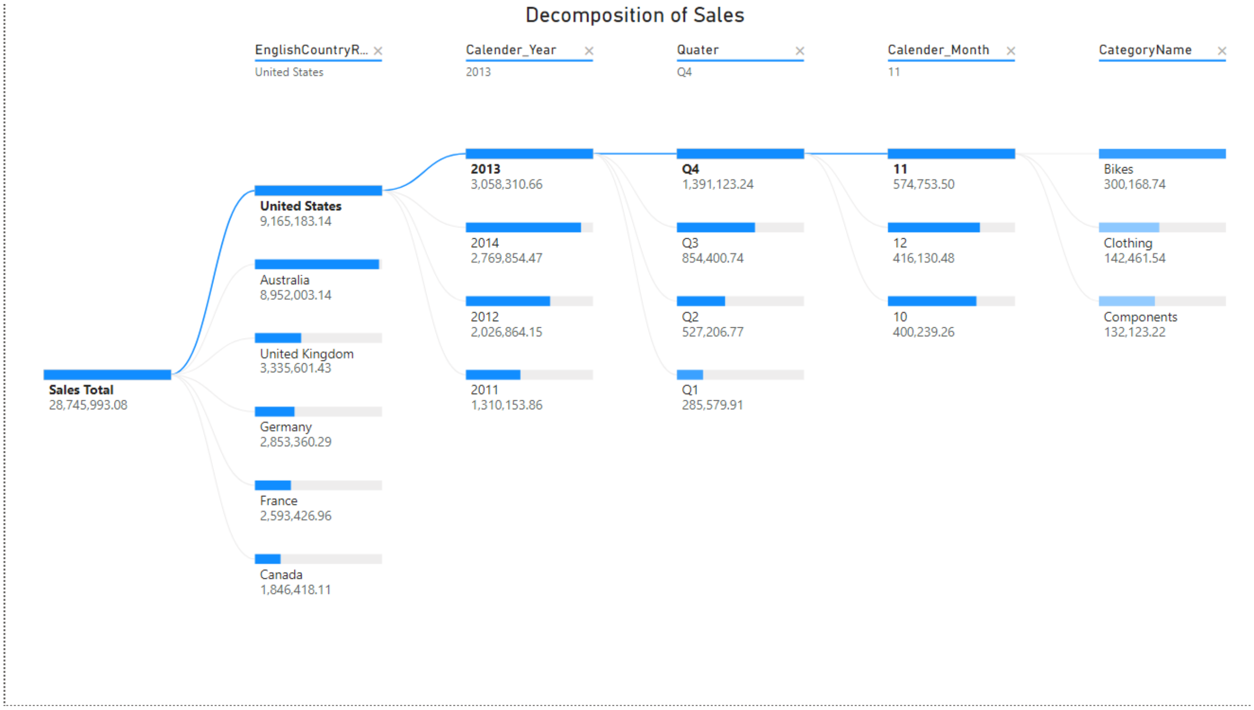
The highest sales are seen by United States and Australia. Company sales are less in European countries. Hence, it is suggested that company should maintain market in these countries. However, Company also needs to focus on increasing it's sales in European Countries.

3. Standard cost versus List price by Product



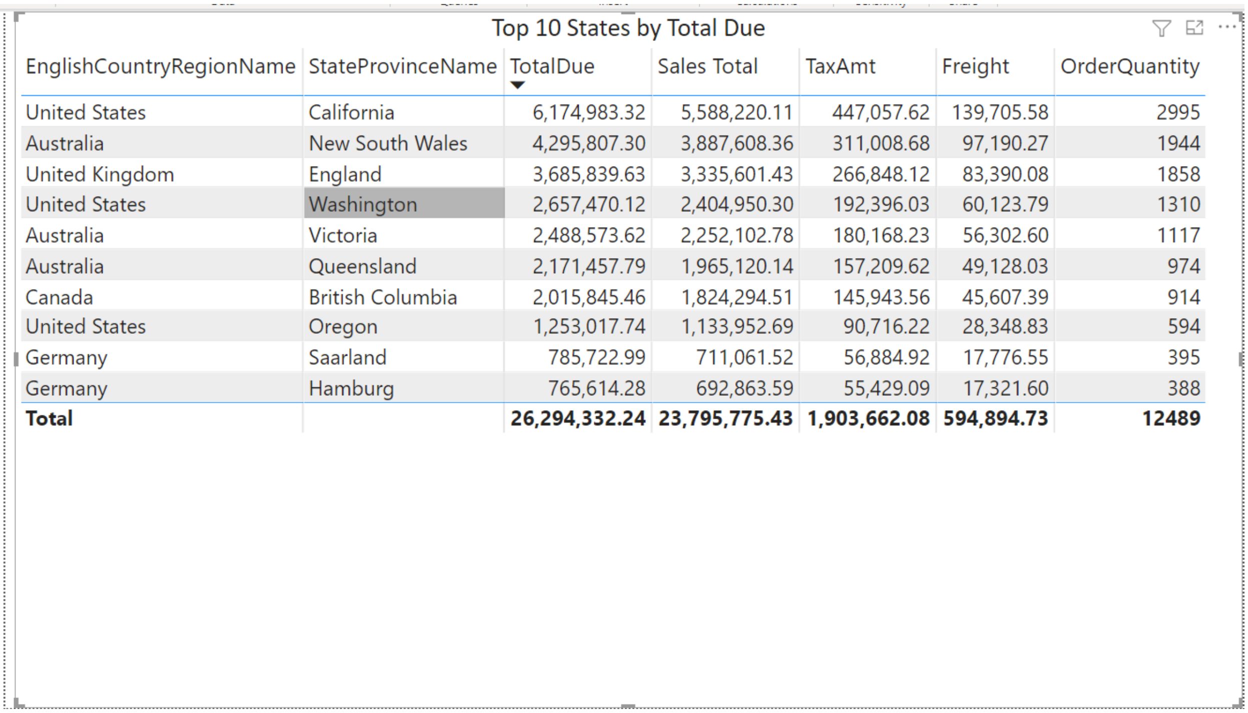
While some Product category has higher than average Standard cost, Mountain Bikes can be sold for higher Margin. Road bike - 250 and Road bike-150 have higher Margin as Mountain Bike.

4. Decomposition of Sales



In USA, Q4 of fiscal year 2013 has highest company had highest sales in November. Possibly because of the Holiday season occurring in November and December.

5. Top 10 States by Total due



Top 10 States by Total Due

EnglishCountryRegionName	StateProvinceName	TotalDue	Sales Total	TaxAmt	Freight	OrderQuantity
United States	California	6,174,983.32	5,588,220.11	447,057.62	139,705.58	2995
Australia	New South Wales	4,295,807.30	3,887,608.36	311,008.68	97,190.27	1944
United Kingdom	England	3,685,839.63	3,335,601.43	266,848.12	83,390.08	1858
United States	Washington	2,657,470.12	2,404,950.30	192,396.03	60,123.79	1310
Australia	Victoria	2,488,573.62	2,252,102.78	180,168.23	56,302.60	1117
Australia	Queensland	2,171,457.79	1,965,120.14	157,209.62	49,128.03	974
Canada	British Columbia	2,015,845.46	1,824,294.51	145,943.56	45,607.39	914
United States	Oregon	1,253,017.74	1,133,952.69	90,716.22	28,348.83	594
Germany	Saarland	785,722.99	711,061.52	56,884.92	17,776.55	395
Germany	Hamburg	765,614.28	692,863.59	55,429.09	17,321.60	388
Total		**26,294,332.24**	**23,795,775.43**	**1,903,662.08**	**594,894.73**	**12489**

Company must receive the above amount from the respective Countries. It also shows the Tax amount and Order Quantity in those countries. We can see highest selling regions in each country.

10. Software and Tools Used:

- Microsoft SQL Server Management Studio
- ER/Studio
- Microsoft Visual Studio 2017(SSDT)
- PowerBI
- Excel

References

- https://www.mssqltips.com/
- https://www.sqlshack.com/

Professors Comments (03/28/2021):

- Intro paragraph telling me what the project / document is for. You have some of it later in the document but you s
- As the document gets large you would want to consider a table of contents.
- I don't understand what the Timeline section is for
- Data pipeline looks good. I'd like a full list of tables…
- You only show me one fact table and you are strictly reverse engineering the DW dbase, you can make your life easier but trimming it back.
- You will need a real model not just a reversed engineered one. They have a number of columns. I need to know the full list -- Analyse the business queries for our model and find redundant columns.
- The business scenarios are good to list
- Softwares typically isn't used as a plural-done
- I see no mention of SCDs

-

Professors Comments (04/11/2021):

- Pipeline is a good level of detail
- Explain how the sources will also populate the facts
- Telling me the number of columns doesn't tell me anything… tell me what you are storing
- SCD descriptions are good
- Add a modification history to the document
- Expand the business scenarios some more
- Explain your olap model
- Describe the visualizations