Project 1

PetStore Enterprise Data Warehouse

by

Ying Hu

## Project 1

## PetStore Enterprise Data Warehouse

Overview

PetStore is a nationwide seller of pets and merchandize. To separately monitor their inventory, the purchasing department created two databases – Animal Supply and Merchandise Supply. The sales department maintains the customer order fulfillment information in the Sales database. These are standalone databases and mostly perform functions for their own departments with delayed inter-departmental communications. PetStore wants to make all the data accessible to all users to increase visibility across different functions and business processes. You are asked to integrate these databases to create an Enterprise Data Warehouse for PetStore.

## Data Warehouse Business Requirements

As part of the company’s Business Intelligence initiative, PetStore would like to create data marts with the following goals:

***Sales Information***. Senior management would like to be able to track sales of animals and merchandise by customer, employee, and supplier, with the goal of establishing which animals and merchandise are the top sellers which employees place the most orders, and who are the best suppliers.

***Procurement Process***. There is a need to analyze the procurement process to see if the time between when the order is placed and when it is received can be improved. The analysis needs to be done for both animal and merchandise purchases by supplier.

***Product Flow Analysis***. Management requires a means to track each animal from purchase to sold date by supplier, breed, and category. This is essential for the upkeep of the animals with minimum stay in enclosures.

***Sales Coverage Analysis***. Management wants to expand the business by analyzing state wise sales coverage across customer and supplier. Management also want to analyze the sales territories covered by the employees.

***Profitability***. Most importantly, the senior managers want to know the overall profit by animal, merchandise, customer, and employee.

1. **Create an Enterprise Bus Matrix to support the above business requirements**. **(5 points)**

Please see excel file.

Find out the common/conformed dimensions and fact table—(fact table for calculation only) or the whole thing you want to see:

1. **Using SQL Server Management Studio create the enterprise data warehouse by the name PetStoreEDW**. **Identify dimensions, fact tables, and conformed dimensions. Are there any factless fact tables and outriggers? Submit the database diagram of PetStoreEDW**. **(7 points).**

Dimensions tables: Employee, Merchandise, Supplier, Animal, Category, City, Breed, Customer.

Fact tables: AnimalOrderItem, AnimalOrder, OrderItem, MerchadiseOrder, Customer Acct, SaleItem, SaleAnimal, Sale

Conformed dimensions: Merchandise, Supplier, Employee, Animal, Date

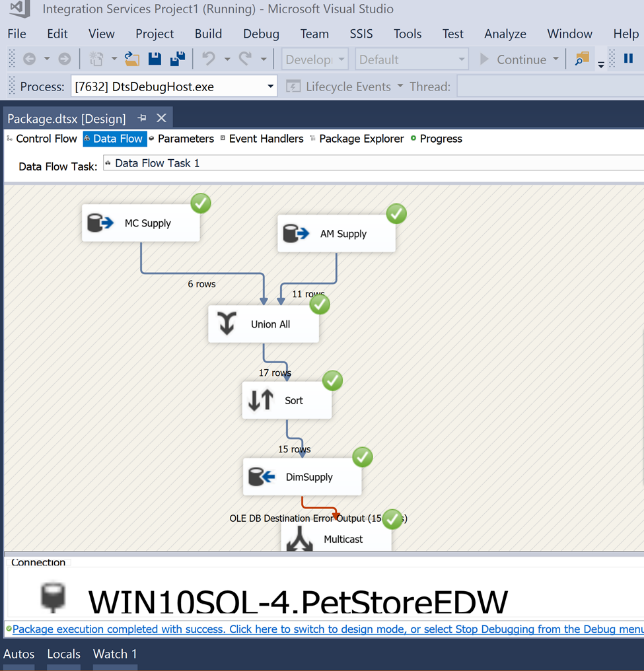
Factless fact table: Breed

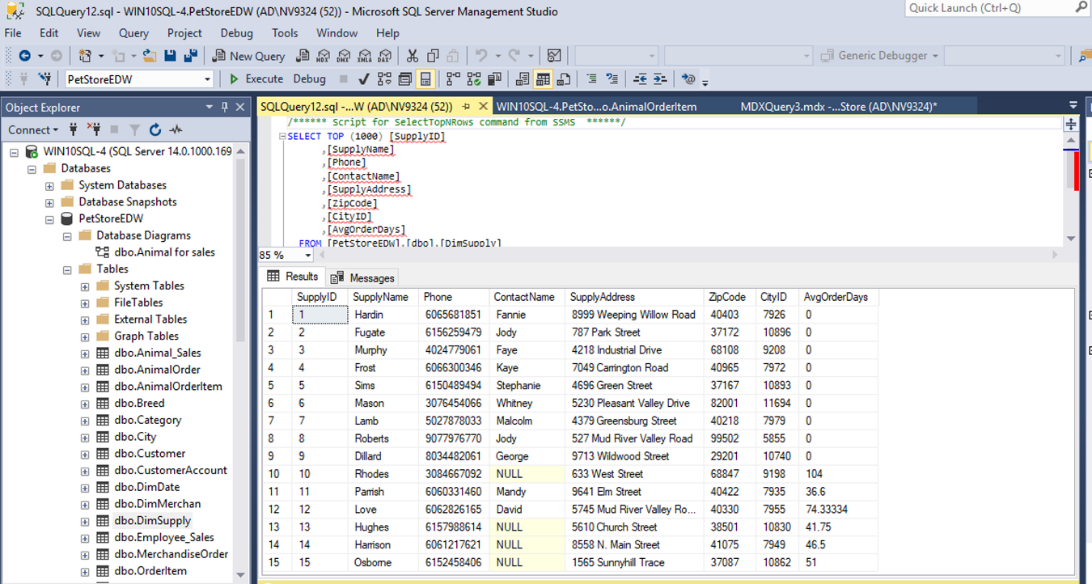
Outriggers: City, Category

**PetStoreEdw diagram:**

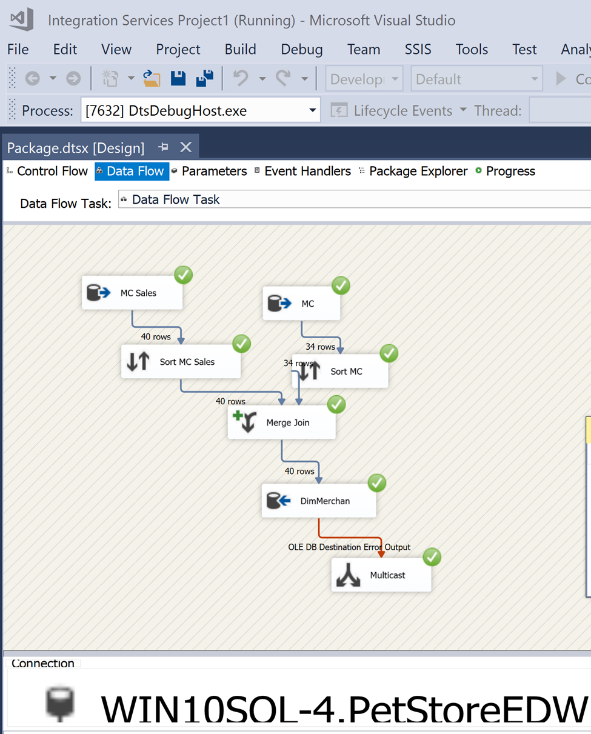


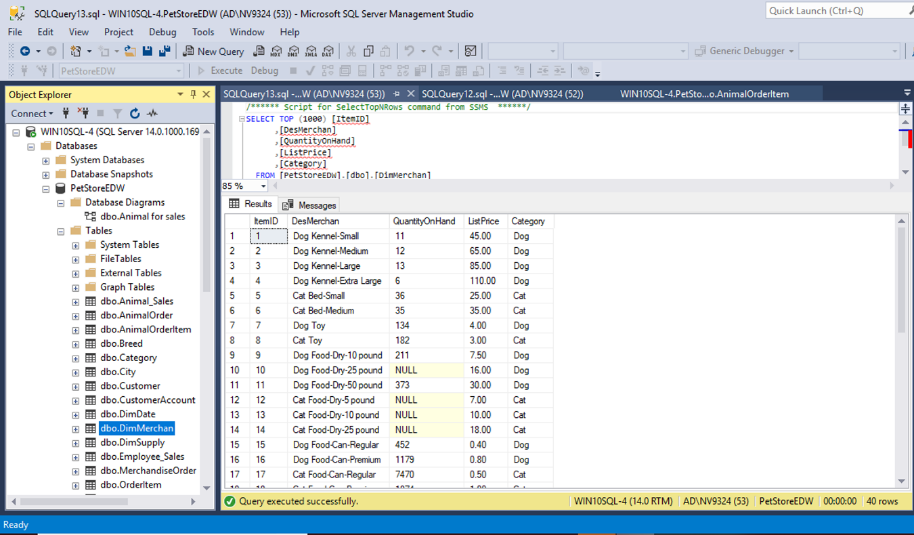
**DimSupply: (Conformed dimension: Supplier):**

****

****

**DimMerchan: (Conformed dimension: Merchandise)**



****

**3. Create separate OLAP cubes using only the relevant dimensions and fact tables required for the following queries. Take screenshots of the cube structures and dimensional usages after successfully deploying the cube. While creating the cube, also create following new calculated members if required:**

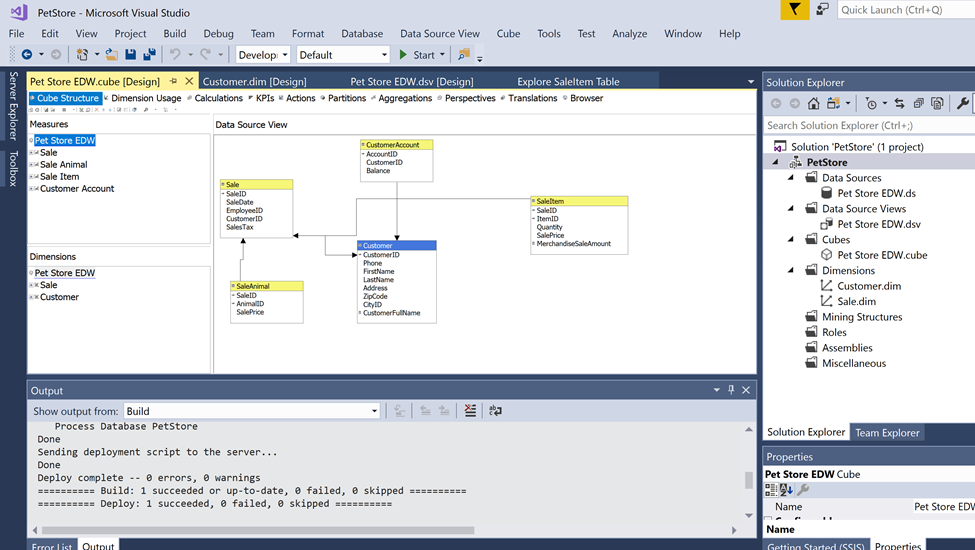
* **AnimalOrderToReceiveTime in the AnimalOrder table**
* **MerchandiseOrderToReceiveTime in the MerchandiseOrder table**
* **CustomerFullName in Customer table**
* **EmployeeFullName in DimEmployee table**
* **MerchandiseCostAmount i.e. Quantity times Cost in OrderItem table**
* **MerchandiseSaleAmount i.e. Quantity time SalePrice in SaleItem table**

**Create following reports using MDX query to drill across fact tables:**

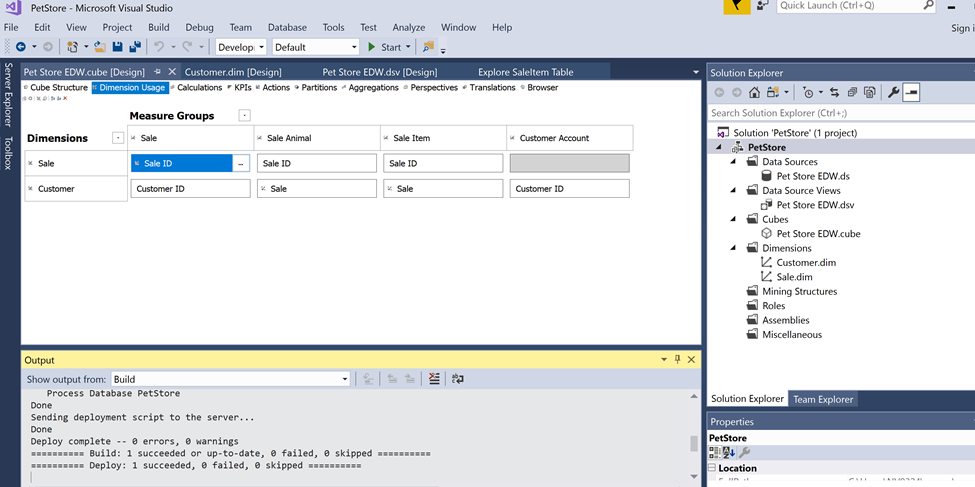
1. **Sales By Customer (1 point):** Display the animal sale price, sales tax, merchandise sale amount, and customer balance for each customer. Will you use Sale table as fact or dimension or both in the cube? Explain.

Answer: Sale table for both, because we need sale table connect and match customer table to find out the CityID for tax. And we also need sale table as a measure to calculate sale tax from sale price.

cube structures:

****

dimensional usages:

****

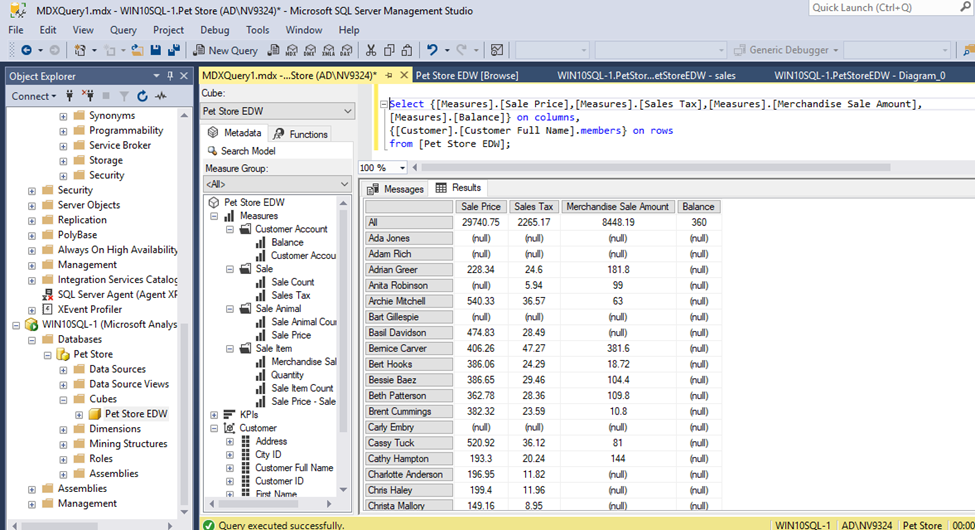
MDX query:

#1

select{[Measures].[Sale Price],[Measures].[Sale Tax],[Measures].[Merchandise Sale Amount],[Measures].[Balance]}

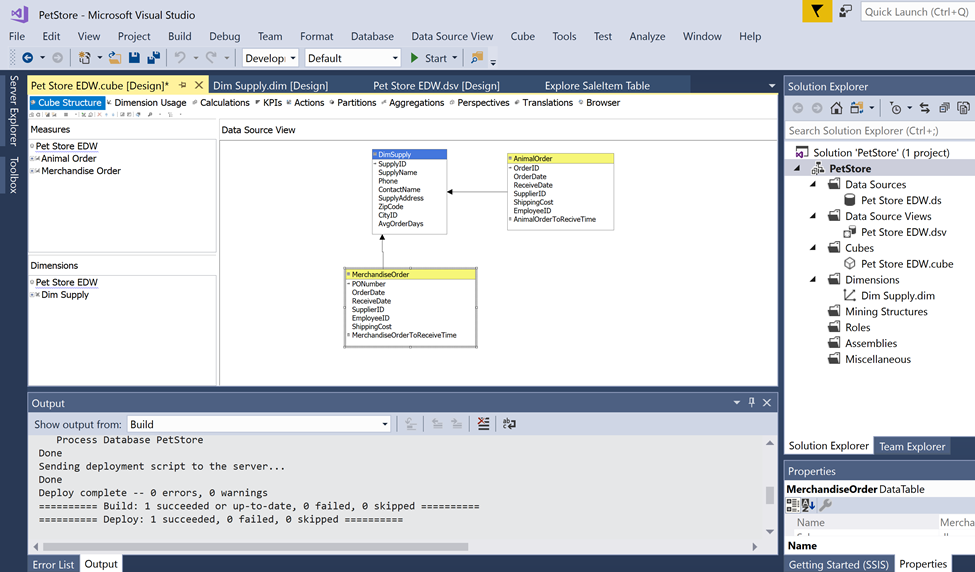
on columns,

{[Customer].[Customer Full Name].members} on rows from [Pet Store EDW];

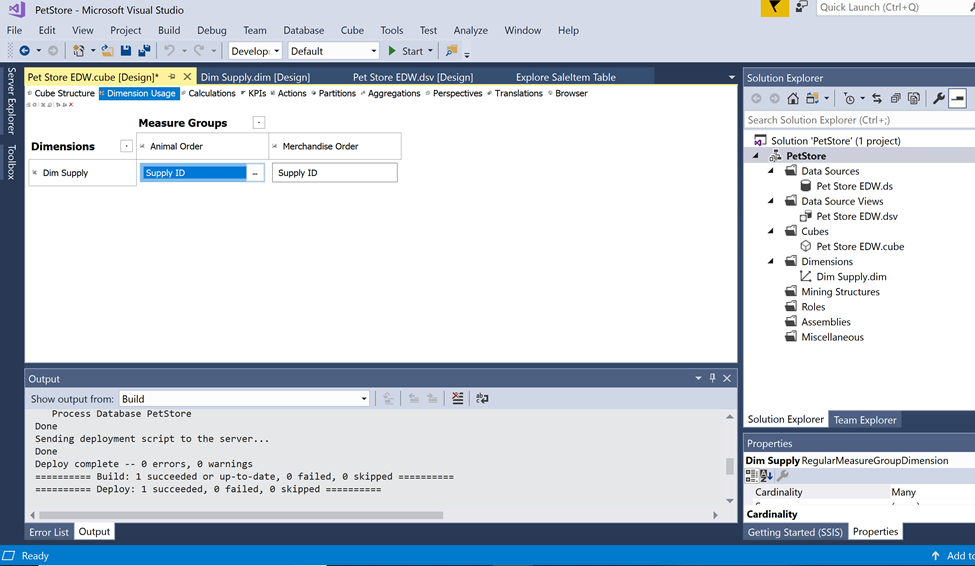
****

1. **Shipping costs By Supplier (1 point):** Display the shipping cost for animals and merchandise for each supplier.

cube structures:

****

dimensional usages:

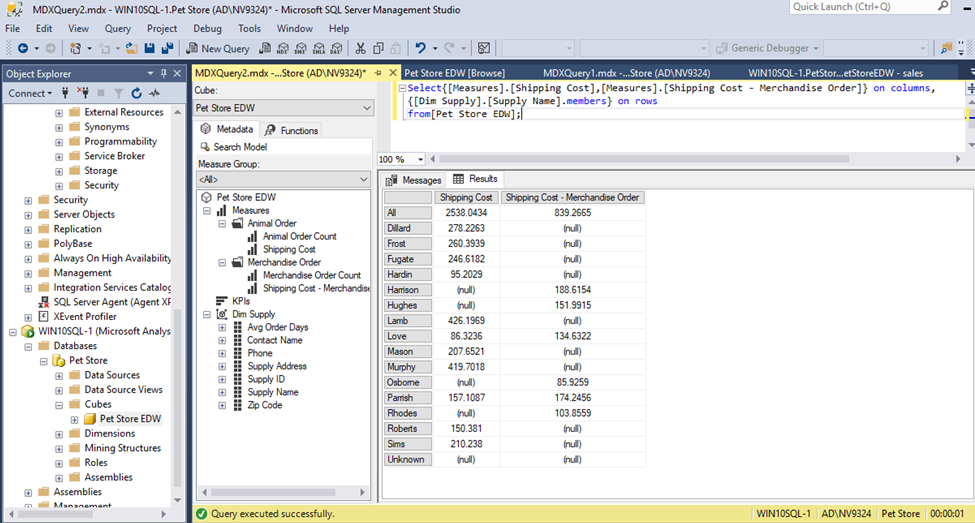
****

MDX query:

#2

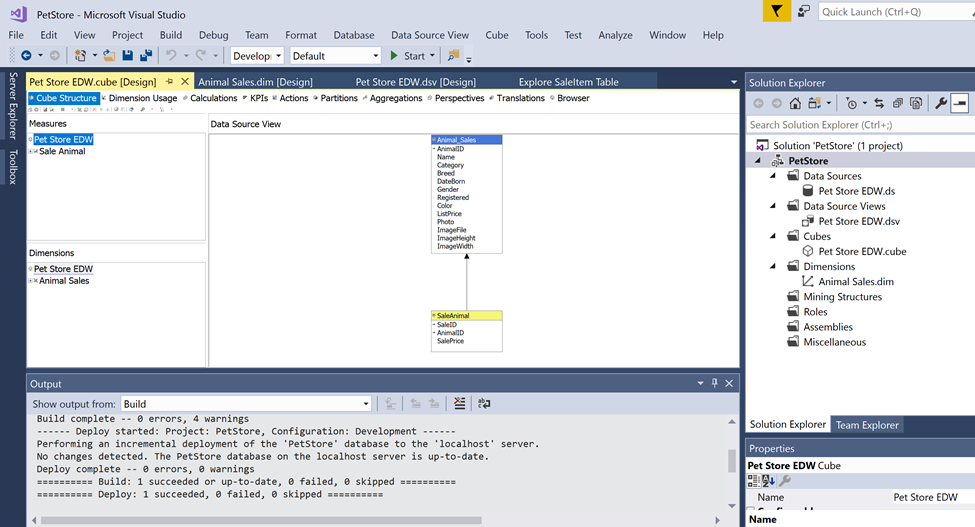
select{[Measures].[Shipping Cost],[Measures].[Shipping Cost – Merchandise Order]} on columns,

{[Dim Supply].[Supply Name].members} on rows from [Pet Store EDW];

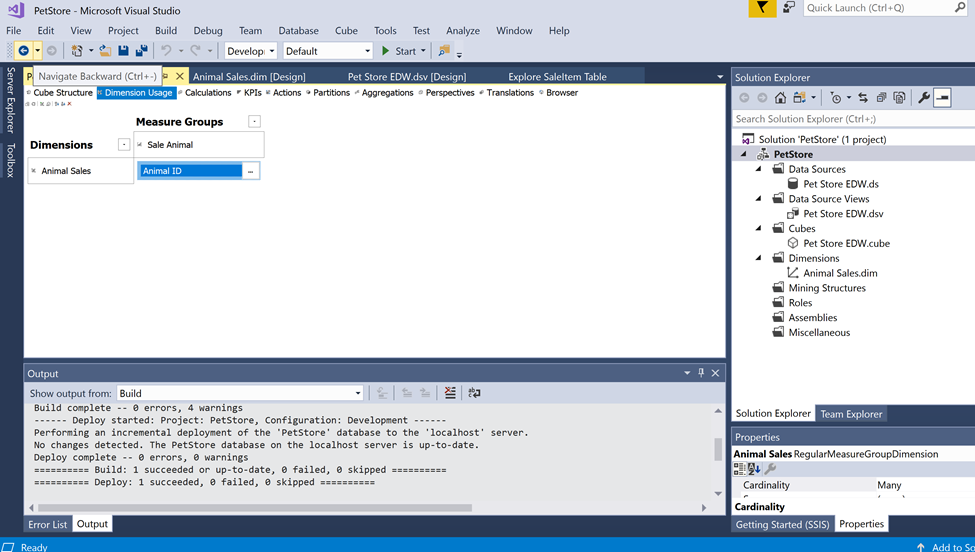
****

1. **Difference between list and sale prices By AnimalID (1 point):** Display the list price, sale price, and the difference of list and sale prices for each Animal ID.

cube structures:

****

dimensional usages:



MDX query:

#3

with member[Measures].[List Price]

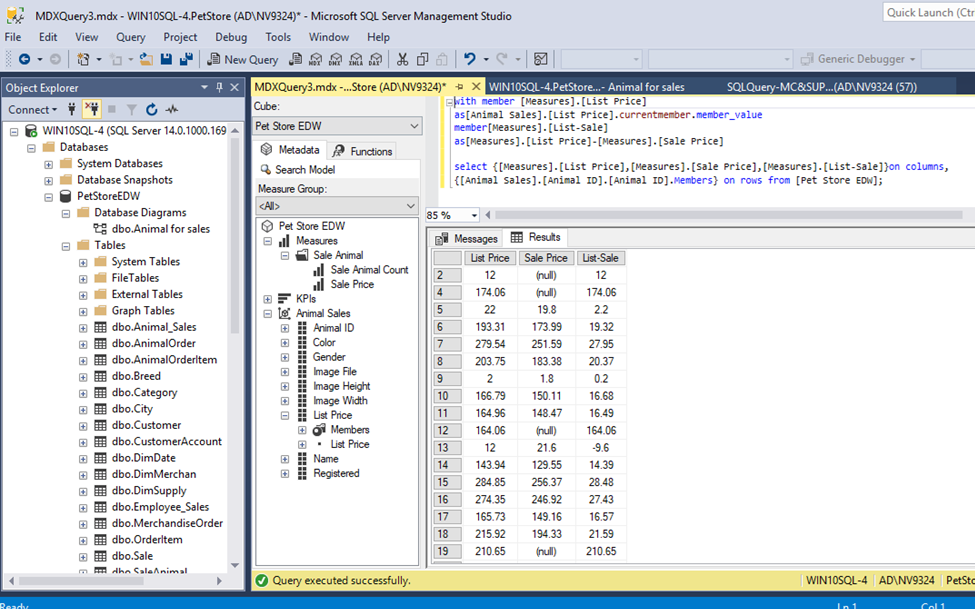
as[Animal Sales].[List Price].currentmember.member\_value

member[Measures].[List-Sale]

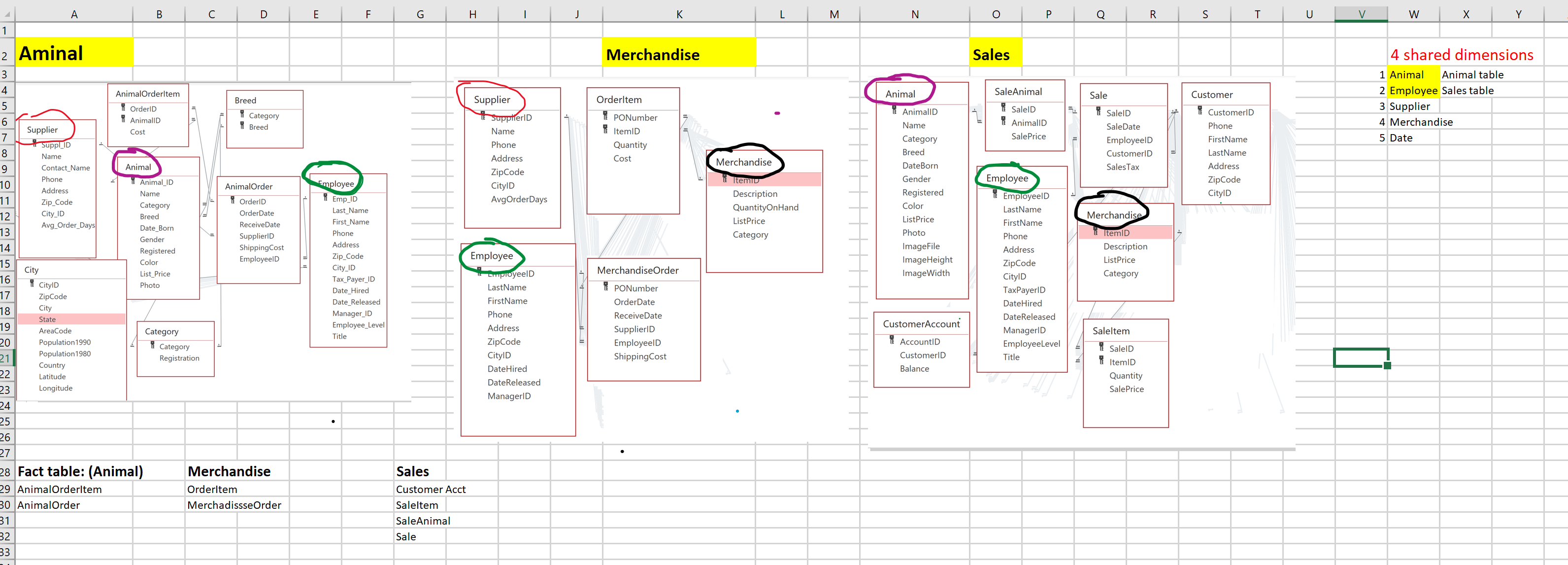
as[Measures].[List Price]-[Measures].[Sale Price]

select{[Measures].[List Price],[Measures].[Sale Price],[Measures].[List-Sale]} on columns,

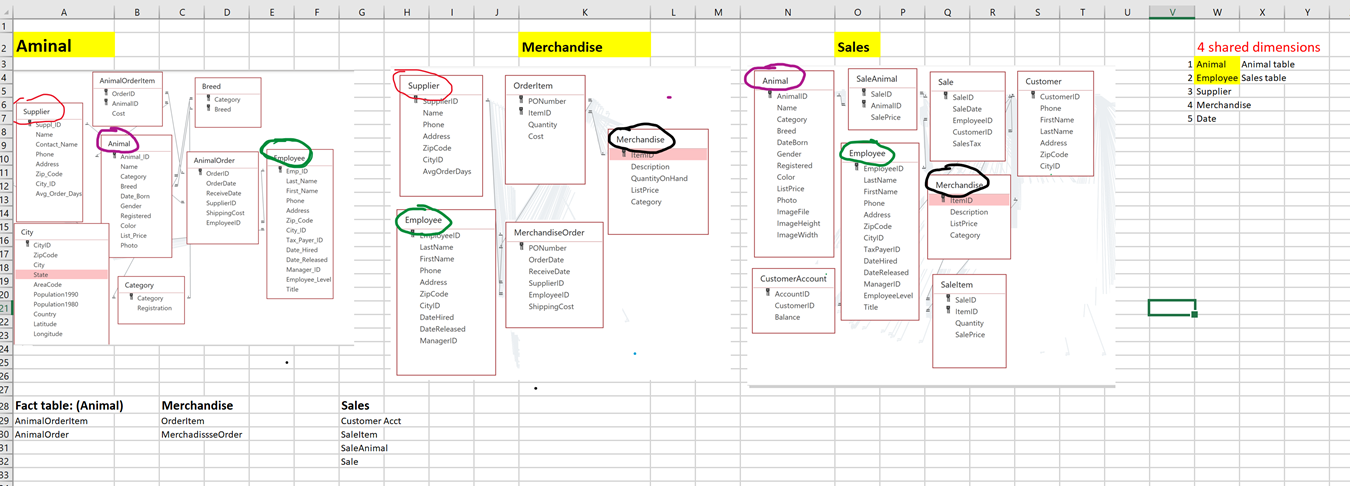
{[Animal Sales].[Animal ID].[Animal ID].Members} on rows from [Pet Store EDW];



1. Steps:
2. Create business matrix ( please see excel sheet project1)
3. Found out the common/superset dimensions (please see excel sheet Dimensions)



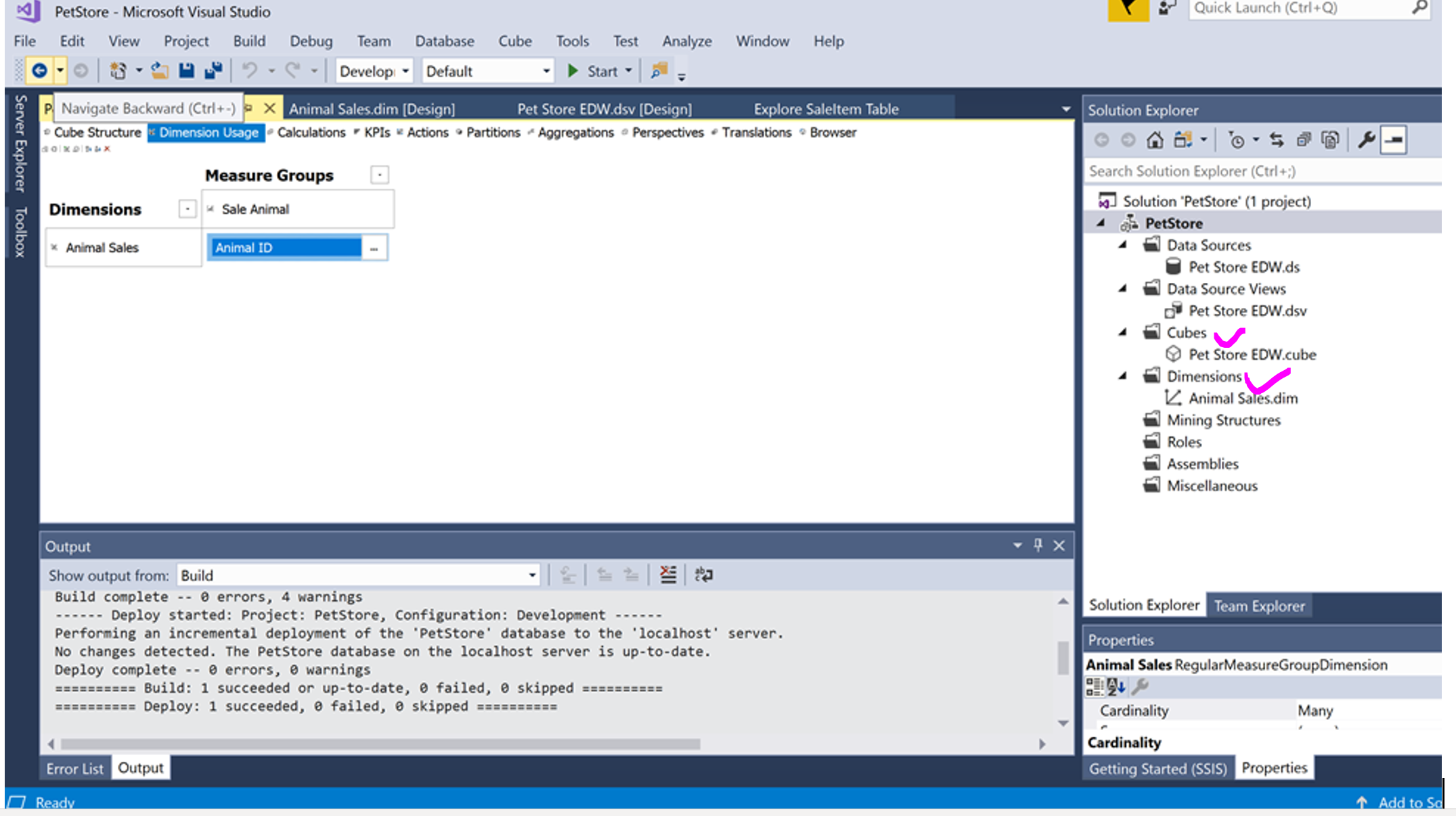
1. Because animal and employee are the super set. No need to do merge. Choose the largest tables. And delete other two duplicated tables.
2. Create queries/schemas for supplier and merchandiser in database SSMS.
3. Supplier and merchandiser in different table are totally have different columns, we need to use SSIS merge them.
4. After done merging, delete the old duplicated tables. Keep both merged tables.
5. Generate the database diagram and connect the relation and find out the pk and fk.



1. Then go to visual 2017-> SSAS OLAP analyst:

Fact table🡪 go to cube(because it is for measure)

Dimension table🡪 go to dimension (because it is not for measurement)



1. Then deploy and go to DMX query.