**Reseller Sales Analysis Based on AdventureWork2017**

**Business requirements:**

Adventure Works is an American cycles company that sells bikes, accessories, clothing and components, etc. This company has around 300 employees including 29 sales representatives. The primary distribution channel for Adventure Works through the retail stores of their resellers and online website to individual customers. These resellers are located in Australia, Canada, France, Germany, the United Kingdom, and the United States, which mainly provide products to North American, European and Asian commercial markets. It wants to set up a reseller sales system to track sales to all of their resellers.

Resellers are typically stores or wholesales, they have names, locations and a specific sales representative from Adventure Works to contact with. Individual customers are person contains tile, name and email information. Each customer may have multiple email addresses. For resellers and individual customers, they all have a unique account number to be identified by Adventure Works.

Sales representatives have the same basic information as individual customers, however, as an employee, their profiles also include sales quota, bonus, commission percent, this year’s total sales, and last year’s total sales.

When a deal happened with resellers or individual customers, there is a sales order generated to contain this deal’s information. This sales order contains customer (reseller or individual customer) information, sales location information, merchandise information and other details like order number, order date, transaction status, purchase order number, tax amount, shipping fee and total due.

Sales location contains group (continent), country, region, sales this year, sales last year, cost this year and cost last year. Merchandise including product information and order details like quantity, price, discount, shipping tracking number. There are different discount types suited for different sales, like discount names, percent, category, start date, and end date.

For products that Adventure Works sells to customers, it includes detailed information like product name, number, color, size, weight, class, style, cost, price, category, start date, end date and discontinued date.

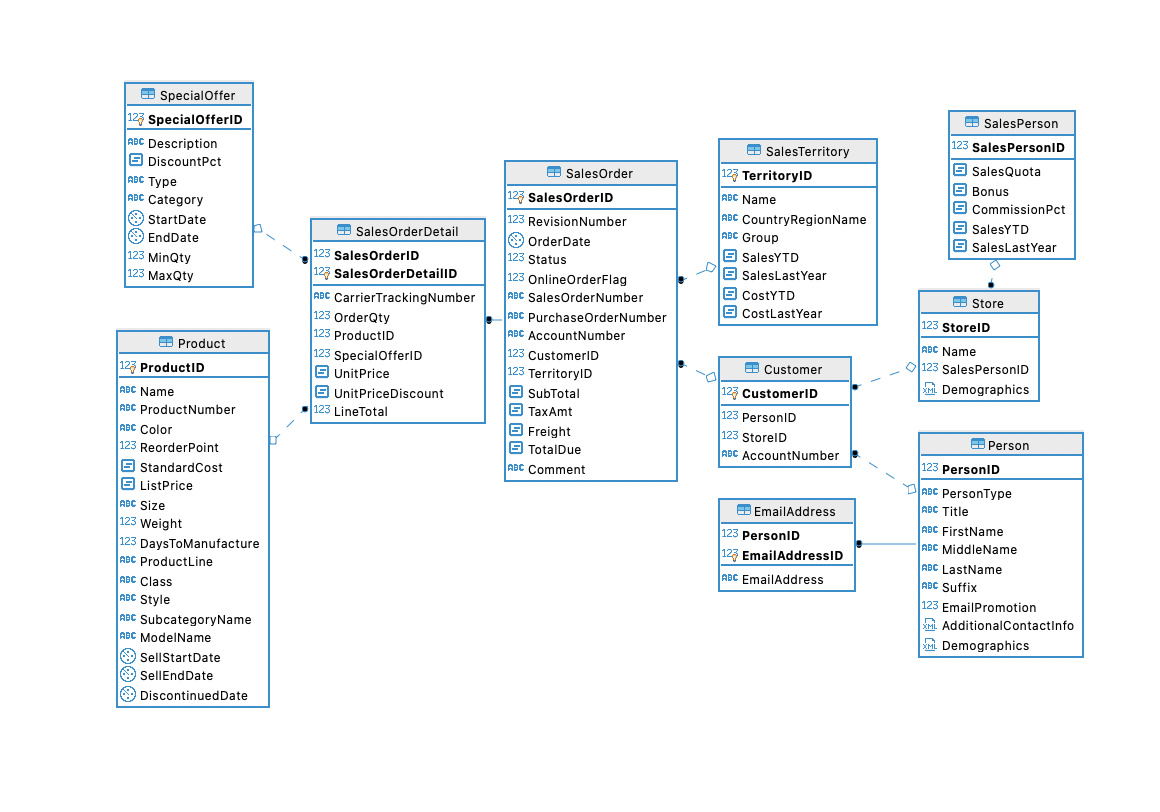
We would like to track sales to all of their resellers for Adventure Works. We will analyze which resellers are better at selling certain types of products in certain regions, and how things like products’ color, size, weight, class, category, cost, and style may impact sales to resellers. In particular, we would like to identify preferences for specific resellers so that we may more effectively predict which products they may be interested in. Similarly, we would like to be able to determine if certain locations have product preferences. Also, we will track what kind of promotion will influence sales. We would like to determine which salespeople will sell more certain things in certain cities. We need to track the date of each order to be able to track the impact of our change over time.

**Technical Requirements:**

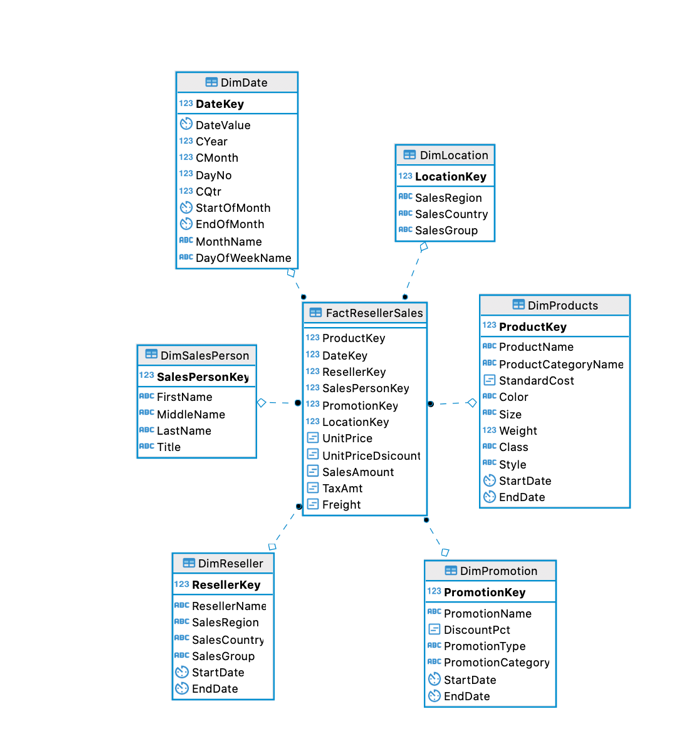
1. The schemas should be created in a new database called ‘AdventureWorks2021’.
2. All tables should in third normal form.
3. There should be an identity column on each table in the database, and created as the primary key.
4. All columns should be appropriately constrained to prevent missing data.
5. All related tables are properly constrained using foreign key.
6. All foreign key columns should have the same name as the column they reference.
7. The database created to satisfy above requirements should be properly normalized.
8. Using both T-SQL and SSIS to build tables required for designed data warehouse star schema model including appropriate constraints. The new data warehouse should be called ‘AW2021\_DW’.
9. Determine the unique/ business key in dimension table, set SCD level appropriately, and also build appropriate indexes .
10. Create stage tables and preload tables appropriately, and create proper procedures to extract, transform and load data to data warehouse.

**Business Scenario:**

Adventure Works Cycles is looking to broaden its market share by targeting their sales to their best resellers. In order to broaden its resellers’ market share, Adventure Works Cycles wants to predict potential resellers by investigating resellers’ location by selling price and shipping expenses. They also want to use resellers’ demographic attributes to forecast average selling prices and shipping expenses. They intend to optimize their new product pricing strategies and sales by similar product information.

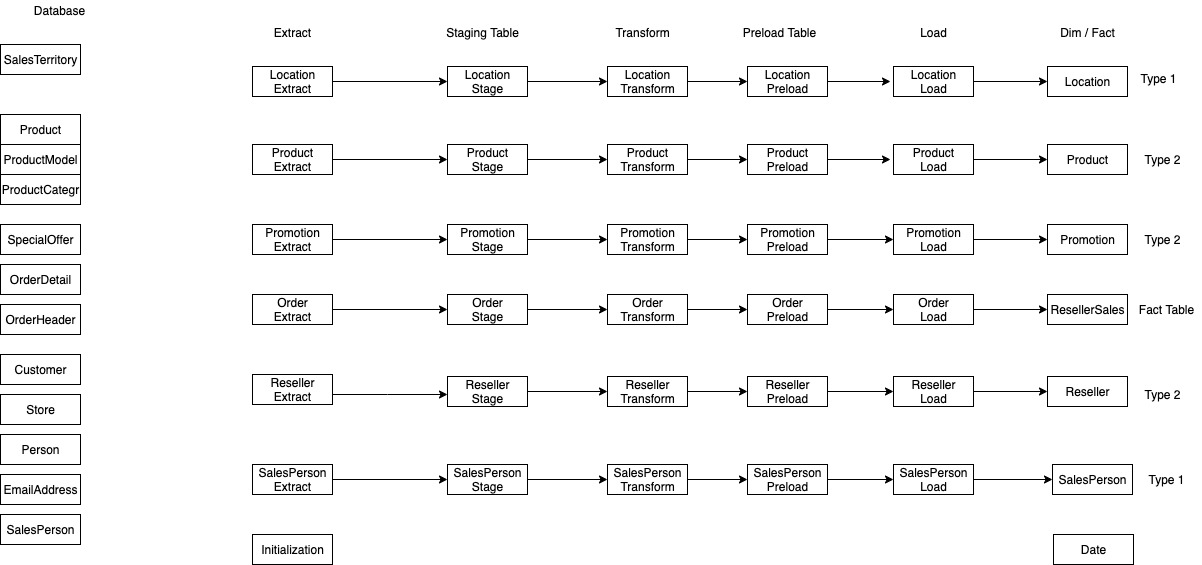
**Database Design:**

Different from AdventureWorks2017 database design, we rebuilt entities and relationships in third normal form by merging tables and reorganizing relationships to suit our business requirements.

**Data Warehouse Design:**

Here we also redesigned multidimensional model for our data mart different from AdventureWorks2017 DW to suit our business scenario.

**ETL Model Design:**



**SSIS Design:**

1. Using SCD node to complete ETL process:
2. Location Dimension Table

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

1. Reseller Dimension Table

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generatedA screenshot of a computer

Description automatically generated

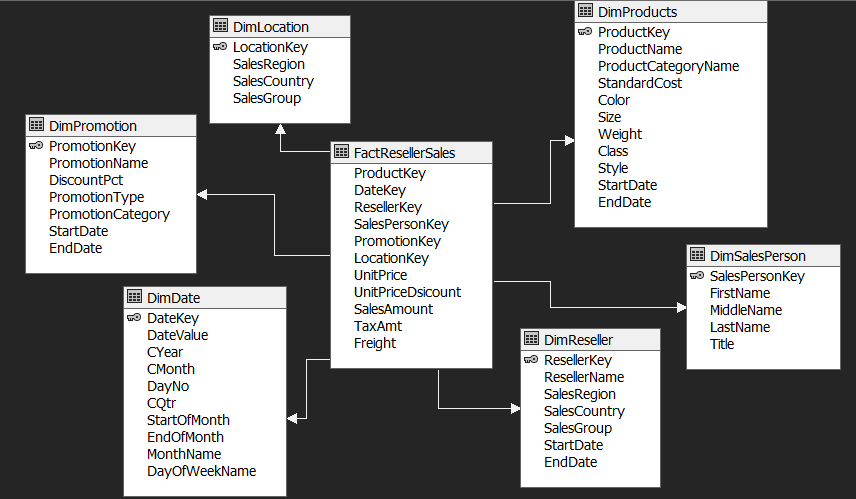
1. Using SQL command to complete ETL process:

A screenshot of a cell phone

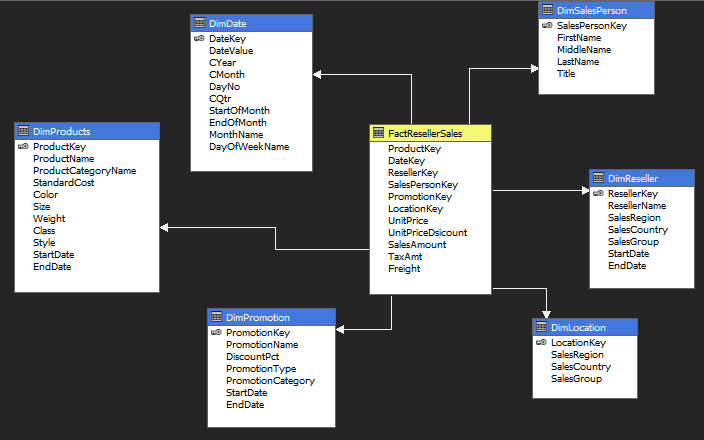
Description automatically generated

**SSAS design:**

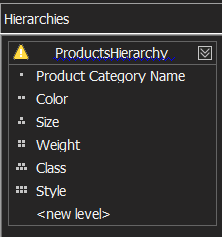
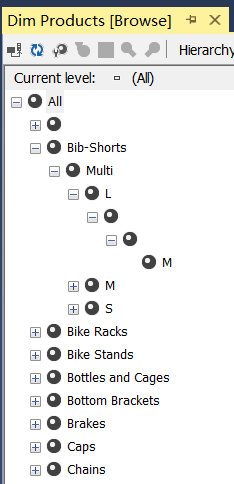
1. Data Source View for our Data Warehouse.



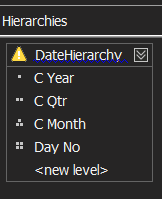
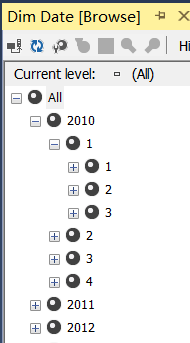
1. Cube Structure.



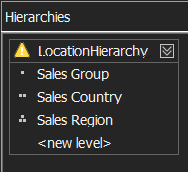
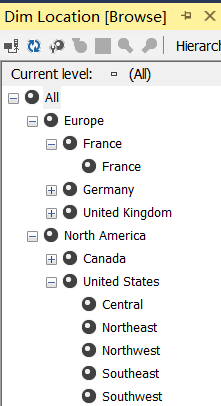
1. Product Dimension Hierarchy.

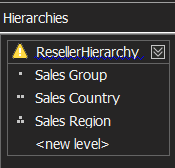
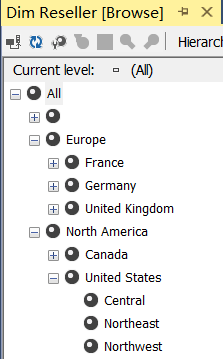
1. Date Dimension Hierarchy.

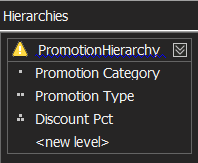
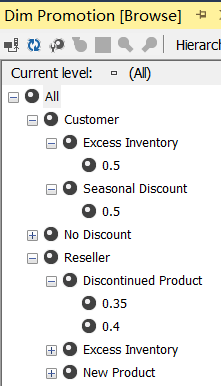
1. Location Dimension Hierarchy.

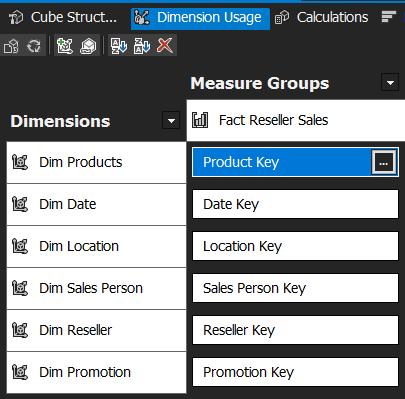
1. Reseller Dimension Hierarchy.

1. Promotion Dimension Hierarchy.

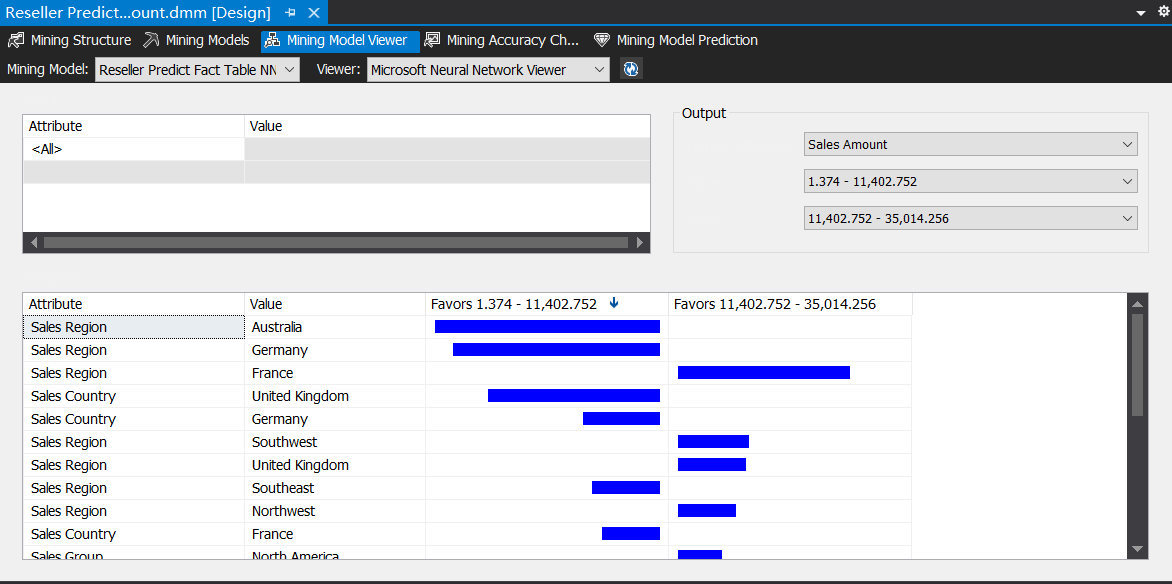
 

1. Dimension Usage.



**Data Mining Structure:**

1. Regression problem: Using Products attributes to predict Sales Amount.



1. Classification problem: Using Products attributes to predict Reseller Location.

