# Lab Data modeling and DAX

## SCENARIO

VanArsdel is a company that manufactures and sells sporting goods. The company has offices in the United States (US) and several other countries. Its sales comprise of US sales and International sales. VanArsdel’s sales come from its owned manufactured products, as well as other manufacturers’ products.

You have successfully brought the US sales data from the Access database and the International sales data from a collection of CSV files to Power BI Desktop. Before you can start analyzing your data, you need to manage the table relationships within your data model and create new ones if necessary. To do so, you might need to create calculated columns or calculated tables for the relationships to be based on.

Once you have all the relationships created, you can create visualizations and start to analyze the data. However, you need to create additional measures to perform more advanced analysis with your data, which includes:

* Comparing last year sales and last year YTD sales.
* Comparing sales of VanArsdel's manufactured goods to other manufacturers.

In this lab, you will create calculated columns, calculated tables, and create table relationships in your data model based on the calculated columns and tables you created. In addition, you will write several DAX expressions to create measures to be used to analyze VanArsdel’s sales data. Specifically, you will create the following measures:

* **Total Sales**: calculates the total sales.
* **LY Sales**: calculates last year sales.
* **Sales Var**: calculates sales variance between this year and last year sales.
* **Sales Var %**: calculates sales variance between this year and last year sales in percentage.
* **YTD Sales**: calculates YTD sales.
* **LY YTD Sales**: calculates last year YTD sales.
* **YTD Sales Var**: calculates sales variance between this year and last year YTD sales.
* **YTD Sales Var %**: calculates sales variance between this year and last year YTD sales in percentage.
* **Total VanArsdel Sales**: calculates sales for VanArsdel manufactured goods.
* **% Sales Market Share**: calculates the percentage of VanArsdel manufactured goods from the total sales.

Before starting this lab, you should review the **Power BI Desktop Modelling** module in this course. Then, if you have not already done so, follow the instructions in the **Set up the Lab Environment** section of this course to set up the lab environment.

## WHAT YOU’LL NEED

* A computer with the latest version of Power BI Desktop installed on it.
* The following Power BI Desktop file:
* The “Lab Data modeling and DAX” file

## Exercise 1: Manage Table Relationships

Power BI Desktop has automatically detected and created table relationships. So the first step is to ensure all the relationships are properly created, and if not, create them yourselves.

1. Start with the "[Lab 2 - Starting.pbix](https://github.com/MicrosoftLearning/Analyzing-Visualizing-Data-PowerBI/blob/master/Lab2/Lab%202%20-%20Starting.zip?raw=true)" file.
2. Open the **Relationship** view.
3. Ensure that there is a many to one relationship with both cross directional filtering from the **ProductID**column on the **Sales** table to the **ProductID**column on the **Products**table. If not, create the relationship by dragging the **ProductID** column on the **Sales** table to the**ProductID** column on the **Products** table.
4. Ensure that there is a many to one relationship with both cross directional filtering from the **ManufacturerID**column on the **Products**table to the **ManufacturerID**column on the **Manufacturers**table. If not, create the relationship.
5. Ensure that there is a many to one relationship with both cross directional filtering from the **Date** column on the **Sales** table to the **Date** column on the **Date** table. If not, create the relationship.

Now you want to create a relationship between the **Sales** table and the **Locations** table. First, you merge the **Country** and **Zip** columns in both **Sales** and **Locations** table as a new column, **CountryZip**. Then, you create a relationship on the **CountryZip**column for both tables.

1. Edit the **Locations**table in the **Data** view.
2. Add a new column named **CountryZip**by concatenating the value from the **Country** column, a comma and a space character, and the value from the **Zip** column. (Hint: the calculated column formula look as follows: CountryZip = Locations[Country] & ", " & Locations[Zip])
3. Edit the **Sales**table in the **Data** view.
4. Add a new column named **CountryZip**by concatenating the value from the **Country Name** column, a comma and a space character, and the value from the **Zip** column. (Hint: the calculated column formula look as follows: CountryZip = Sales[Country Name] & ", " & Sales[Zip])
5. Open the **Relationship** view.
6. Link the newly created **CountryZip**column on the **Sales** table to the newly created **CountryZip**column on the **Locations** table.

### **Lab Questions**

Review the relationship that you have just created on the CountryZip columns. (Hint: Double-click on the newly created relationship).

* 1. What is the cardinality of the relationship?

Many to One

One to One

One to Many

Many to Many

* 1. What is the Cross filter direction of the relationship?

Single

Double

Both

Multiple

## Exercise 2: Last Year Comparison

You want to know how much sales (revenue) in total does the VanArsdel have and to compare this with the figure from the same period last year. You need to create several calculated measures to help with this comparison. To do so, in either the **Report** view or the **Data** view, right-click the **Sales** table, click **New Measure**, and type in the corresponding DAX formulas for the measure you want to create. This will create the measures with the **Home Table** properties set to the **Sales** table.

Specifically, you will create the following measures:

* **Total Sales**: calculates the total sales. Format this measure as **Currency**. (Hint: Check out the **SUM** function).
* **Sales LY**: calculates last year sales. Format this measure as **Currency**. (Hint: You might find the **CALCULATE** and **SAMEPERIODLASTYEAR** function useful).
* **Sales Var**: calculates sales variance between this year and last year sales. Format this measure as **Currency**. (Hint: This is simply the difference between **Total Sales** and **Sales LY**).
* **Sales Var %**: calculates sales variance between this year and last year sales in percentage. Format this measure as **Percentage**. (Hint: This is simply the percentage of **Sales Var** from **Sales LY**. You might find the **DIVIDE** function useful).

### **Lab Questions**

Answer the following questions using the measures you created. (Do NOT include currency symbols or thousands separators). The fastest way to do this is to drag the measures you created to the Report view and format them as a table visualization. Do not use any level of filtering to answer the questions.

1. What is the figure for the Total Sales measure? (to two decimal places)
2. What is the figure for the Sales Var % measure? (to two decimal places) %

Exercise 3: Year to Date

Year-to-date (YTD) is a period starting from the beginning of the current year and continuing up to the present date. You want to calculate the YTD sales and compare this with the figure from the same period last year. Specifically, you will create the following measures:

* **YTD Sales**: calculates the YTD sales. Format this measure as **Currency**. (Hint: Check out the **TOTALYTD** function).
* **YTD LY Sales**: calculates last year YTD sales. Format this measure as **Currency**. (Hint: You might find the **CALCULATE** and **SAMEPERIODLASTYEAR** function useful).
* **YTD Sales Var**: calculates sales variance between this year and last year YTD sales. Format this measure as **Currency**. (Hint: This is simply the difference between **YTD Sales** and **YTD** **LY Sales**).
* **YTD Sales Var %**: calculates sales variance between this year and last year YTD sales in percentage. Format this measure as **Percentage**. (Hint: This is simply the percentage of **YTD** **Sales Var** from **YTD** **LY Sales**. You might find the **DIVIDE** function useful).

### **Lab Questions**

Answer the following questions using the measures you created. (Do NOT include currency symbols or thousands separators). The fastest way to do this is to drag the measures you created to the Report view and format them as a table visualization. Do not use any level of filtering to answer the questions.

1. What is the figure for the YTD LY Sales measure? (to two decimal places) 
2. What is the figure for the YTD Sales Var % measure? (enter the **absolute** value, to two decimal places)  %

## Exercise 4: Market Share

VanArsdel’s sales comprise of products manufactured by VanArsdel and other companies. You want to know how much of these sales are VanArsdel’s own manufactured products. You decide to show this share in numbers and %. Specifically, you will create the following measures:

* **VanArsdel Total Sales**: calculates sales where the products manufacturer is VanArsdel. Format this measure as **Currency**. (Hint: Use the **CALCULATE** function and filter by Manufacturer).
* **VanArsdel % Sales Market Share**: calculates the percentage of sales of VanArsdel manufactured products from the total sales. Format this measure as **Percentage**.

### **Lab Questions**

Answer the following questions using the measures you created. (Do NOT include currency symbols or thousands separators). The fastest way to do this is to drag the measures you created to the Report view and format them as a table visualization. Do not use any level of filtering to answer the questions.

1. What is the figure for the VanArsdel Total Sales measure? (to two decimal places)
2. What is the figure for the VanArsdel % Sales Market Share measure? (to two decimal places) %

## Exercise 5: Optimize the Data Model

Now that you have the table relationships defined and the measures created, you want to optimize the data model before you create the visualizations.

1. Open the **Data** view.
2. Ensure both the **International Sales** and **Country Population** table are hidden from the report view.
3. Hide the following fields on the **Date** table from the report view.
   * **MonthNo**
   * **MonthID**
   * **Month**
4. Sort the **MonthName** column by the **MonthNo** column.
5. Hide the **CountryZip**field on the **Locations**table from the report view.
6. Hide the **ManufacturerID** field on the **Manufacturers** table from the report view.
7. Hide the following fields on the **Products**table from the report view.
   * **ProductID**
   * **ManufacturerID**
   * **Manufacturer**
8. Hide the following fields on the **Sales**table from the report view.
   * **ProductID**

* **Date**
* **Zip**
* **Units**
* **Revenue**
* **Country Name**
* **CountryZip**