**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.

VanArsdel's US office stores the sales data on an Access database. VanArsdel International sales transactions are available as comma separated (CSV) files. They could be generated daily, either manually by someone, or automatically by an automated process. They are available in a dedicated folder. These CSV files have the same column structure as the sales table for the US sales that comes from the SQL Database.

You want to perform analysis on VanArsdel's worldwide sales data for the year 2000 to 2015. You need to bring all these data into Power BI Desktop before you can perform any analysis. Finally, you want to compare VanArsdel's country sales with the country population. You need to import the country population data from a less structured Excel report to Power BI.

**LAB OVERVIEW**

This lab comprises of three exercises:

1. In the first exercise, you will import data to Power BI Desktop from an Access database file.
2. In the second exercise, you will import data from CSV files which resides in a file folder. You will append this new data to the corresponding existing data that comes from the Access Database.
3. In the third exercise, you will import data to Power BI Desktop from an Excel file that is less structured.

**WHAT YOU'LL NEED**

* A computer with the latest version of Power BI Desktop installed on it.
* A copy of the Access Database containing VanArsdel's US sales data.
* 4 CSV files, containing VanArsdel’s international sales data:
  + CA Sales.csv
  + FR Sales.csv
  + DE Sales.csv
  + MX Sales.csv
* An Excel file containing country population data.

**Import from Access DB**

**Activities**

1. Import the following tables to the power bi desktop file: **bi\_date**, **bi\_geo**, **bi\_manufacturer**, **bi\_product**, and **bi\_salesFact**
2. Filter the rows on the **bi\_salesFact** query to include dates from January 1st, 2000
3. Filter the rows on the **bi\_date** query to include dates from January 1st, 2000.
4. Rename the queries as follows:
   * 1. bi\_date: **Date**
     2. bi\_geo: **Locations**
     3. bi\_manufacturer: **Manufacturers**
     4. bi\_product: **Products**
     5. bi\_salesFact: **Sales**

**Questions**

1. How many rows were imported from the Date query?
2. How many rows were imported from the Locations query?
3. How many rows were imported from the Manufacturer query?
4. How many rows were imported from the Products query?
5. How many rows were imported from the Sales table?

**Import from CSV files**

**Activities**

Perform the following steps:

1. Name the query **International Sales**.
2. Select to combine (combined binaries) the content of those 4 files.
3. Filter the rows that come from the header of the CSV files. (Hint: One way to do this is to filter out the **Country**column from records containing “Country”).
4. Filter the rows that are after **December 31st, 1999**.
5. Edit the **Sales** Query from the US Sales.
6. Append the **International Sales** Query to the **Sales** query from the US Sales.
7. In the **Sales** query, add a custom column named **Country Name** which takes the value of the **Country** column when it is not null and the value of "USA" when the **Country** column is null.
8. Remove the **Country** column.

**Questions**

1. How many rows were imported from the International Sales query?
2. How many rows were imported from the Sales query?

**Import from Excel**

**Activities**

1. Edit the query and perform the following steps:
   * 1. Name the Query **Country Population**.
     2. Remove the first four rows of the table.
     3. Promote the row that represents the title of the column to the table headers.
     4. Transform the yearly columns to rows and name the resulting columns appropriately.
     5. Rename the resulting columns to **Year** and **Population** respectively.
     6. Filter out the year **1999**.
     7. Set the **Data Type** of both the **Year** and **Population** columns to **Whole Number**

**Questions**

1. How many rows were imported from the Country Population query?
2. How many distinct values of the Year column were imported from the Country Population query?