**Problem Statement**

Scenario: Adventureworks is a Bicycle company that sells bicycles worldwide. They work with a partner to handle their internet sales. Their partner exports the sales history on a monthly basis to Azure Blob storage. The format of the files is delimited text

Adventureworks is also introducing new products to their portfolio in the upcoming months. They exported an entire catalog of their products to a shared BLOB storage location. The format of the files is delimited text.

The location for the data is here:

## Account Name: g1stornlb57tqnxgtls

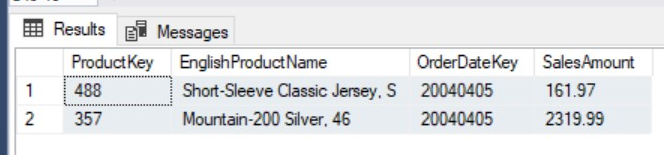
Account Access Key: Vwu6L1WgCwbh1HzCBXGzN+9ZwfHje7YUkO3mRCNXtW6kl45PTqUUQbvnAQtwhEKpPWhyV0cfSJ9R4LHaQoSd5w==

Container Name: advdata

They would like to use this data to draw meaningful insights. They approached you to build an end to end solution that addresses data ingestion, cleansing, preparation, analysis, and orchestration.

You need build a Modern Data Warehouse (MDW) solution that meets the following criteria:

1. Ingest the CSV data from Azure Storage into Azure SQL Data Warehouse in your Azure subscription.
2. Perform necessary steps to ingest and load the data into a table named ‘FactInternetSalesStg’ in your SQLDW database.
3. Load the products dump file into a table named ‘DimProductStg’ in the most efficient way possible.
4. Write a query to join both tables above to generate a report that looks similar to below:



1. Create a reporting user named ‘rptusr’ with absolute minimum privileges to execute the query above.
2. The query observed to be not performing optimally. Investigate the reasons for this performance issue and make necessary changes to the table structure so that query can perform optimally.
3. This report is expected to run multiple times, configure your Data Warehouse to make your subsequent executions of this query faster.

Bonus: Management loved your report. But when they tried to run the same report for different products and for different dates, they didn’t like the performance of the report. Can you create a generic solution that works optimally for all products and dates?

Use the following resources for guidance:

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/load-data-wideworldimportersdw>

<https://docs.microsoft.com/en-us/azure/data-factory/>

<https://docs.microsoft.com/en-us/azure/sql-data-warehouse/sql-data-warehouse-manage-monitor>

<https://docs.microsoft.com/en-us/azure/sql-database/sql-database-manage-logins?toc=/azure/sql-data-warehouse/toc.json>