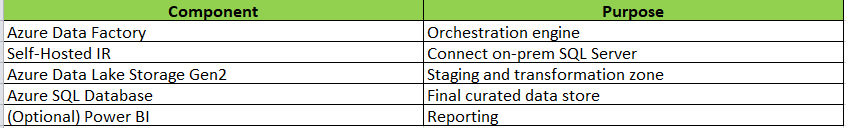
**Architecture Overview**

**Complete Azure Data Factory (ADF) project** using your **on-prim AdventureWorksLT2022 database and a Self-Hosted Integration Runtime (SHIR)** to move data into **Azure Data Lake Storage Gen2 (ADLS) and Azure SQL Database.**

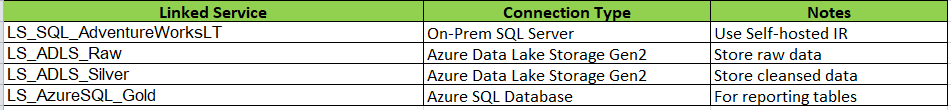
1. **Overall Architecture Overview**

****

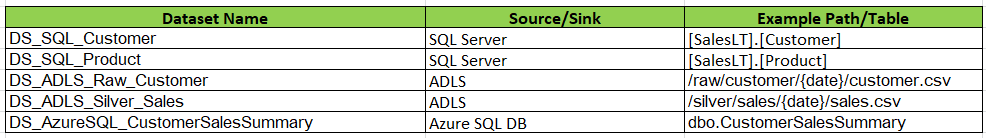
1. **Required Azure Components**



1. **Create Linked Services**



1. **Create Datasets**



1. **Create Metadata Table (In SQL Server)**

****

1. **Create Pipelines**
2. PL\_Master\_AdventureWorksETL

Main orchestrator pipeline.

Activities:

Lookup: Read table list from ADF\_Metadata\_Config

ForEach: Loop through each table

Inside ForEach → Execute PL\_Bronze\_Load

Execute Pipeline: PL\_Silver\_Transform

Execute Pipeline: PL\_Gold\_Load

Web Activity: Notify Success (Teams or Email)

1. PL\_Bronze\_Load

Extract data from SQL → ADLS Raw

Activities:

1. Copy Activity:

* Source: SQL Server Table
* Sink: ADLS /raw/{tablename}/{date}/data.csv

1. **Stored Procedure:** Insert audit record into ADF\_Audit\_Log



1. PL\_Silver\_Transform

Clean & standardize data using **Mapping Data Flow**.

Example Transformations:

* Join SalesOrderHeader + SalesOrderDetail
* Derive TotalAmount = UnitPrice \* Quantity
* Remove null customers
* Sink to /silver/sales/

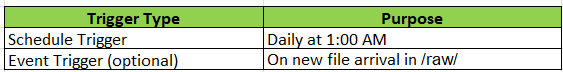
1. PL\_Gold\_Load

Aggregate & load curated data to Azure SQL DB.

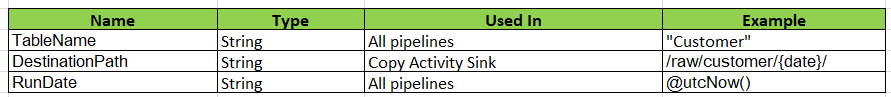
Mapping Data Flow:

* Read /silver/sales/
* Group By: CustomerID
* Derive: TotalSales, AverageOrderValue
* Sink: Azure SQL DB table dbo.CustomerSalesSummary

1. **Add Triggers**



1. **Add Parameters and Variables**



1. **Monitoring and Logging**

* Use ADF monitor tab for run history.
* Store audit logs in SQL via Stored Procedure.
* Send notifications using Web Activity (e.g., Teams or Logic App).

**Optional Enhancements**

* Row Count Validation Activity – Compare SQL vs ADLS row counts
* Delete Activity – Clean ADLS before each run
* If Condition – Skip table if no data returned
* Wait/Until – Wait for dependent files
* Parameterized Data Flows – Reuse transformations dynamically

**Result**

* We will have 3 layers (Raw → Silver → Gold)
* End-to-end orchestration using ADF
* Metadata-driven dynamic pipeline
* All ADF components covered
* Logs and notifications automated
* Ready for Power BI dashboard creation