



DATA WAREHOUSE PROJECT

ADVENTUREWORKS COMPANY

Chorn Cheadalin

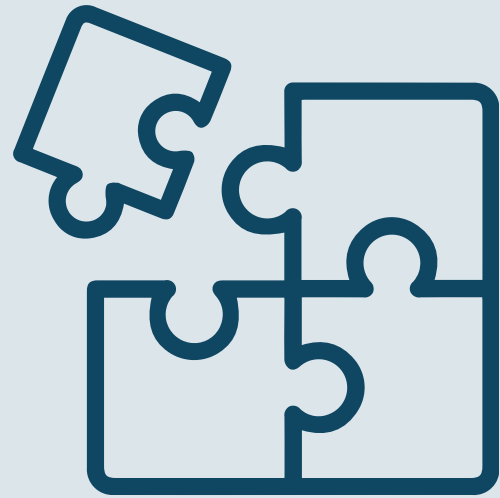


Project Overview

AdventureWorks2022 into AdventureWorks_DWH

The company faces challenges with scattered, unorganized data that slows analytical queries and leads to decisions based on partial or outdated data. To address this, it is implementing a dedicated Data Warehouse to consolidate, cleanse, and transform data from multiple sources into an integrated, analytics-ready system.

Project Objectives



Data Integration

- ETL from AdventureWorks2022 into Staging then to Data Warehouse
- Ensure data consistency and accuracy through cleansing and validation



Data Modeling

- Design a Galaxy Schema model for analytical queries
- Develop Dimension Tables (DimDate, DimCurrency, DimProduct, DimPromotion, DimCustomer, DimStore) and Fact Tables (FactStoreSales, and FactOnlineSales)



Technique Implementing

- Implement an incremental ETL process to update only new or changed records, improving performance.

Project Overview

Build a Data Warehouse (AdventureWorks_DWH) for
AdventureWorks Company



Why do we have to transform and load data to Data Warehouse?



Feature	OLTP (AdventureWorks2022)	OLAP (AdventureWorks_DWH)
Purpose	Daily operations (insert, update, delete)	Strategic analysis & reporting
Data Model	Normalized (100+ tables, many joins)	Galaxy Schema (simple, fast queries)
Query Type	Short, fast, single record	Long, complex, aggregates
Performance	Fast for 1 row	Fast for many rows
Historical Data	Current only	Full history + trends



Project Architecture

AdventureWorks_DWH uses a three-layer architecture for efficient ETL and reporting: data moves from the source through staging (extraction & transformation) to the analytical data warehouse, enabling fast, reliable insights.

Source

(AdventureWorks2022)

- OLTP: Raw transactional data (Product, SpecialOffer, SalesTerritory, Store, etc.)

Staging

(SSIS)

- Extract, clean, transform, validate; handles incremental loads

Data Warehouse

(AdventureWorks_DWH)

- OLAP: Galaxy schema
- Dimension Tables: DimCurrency, DimStore, DimProduct, etc.
- Fact Tables: FactStoreSales and FactOnlineSales



Data Flow Workflow

1 Extract: AdventureWorks2022

2 Staging: temporary storage for cleansing

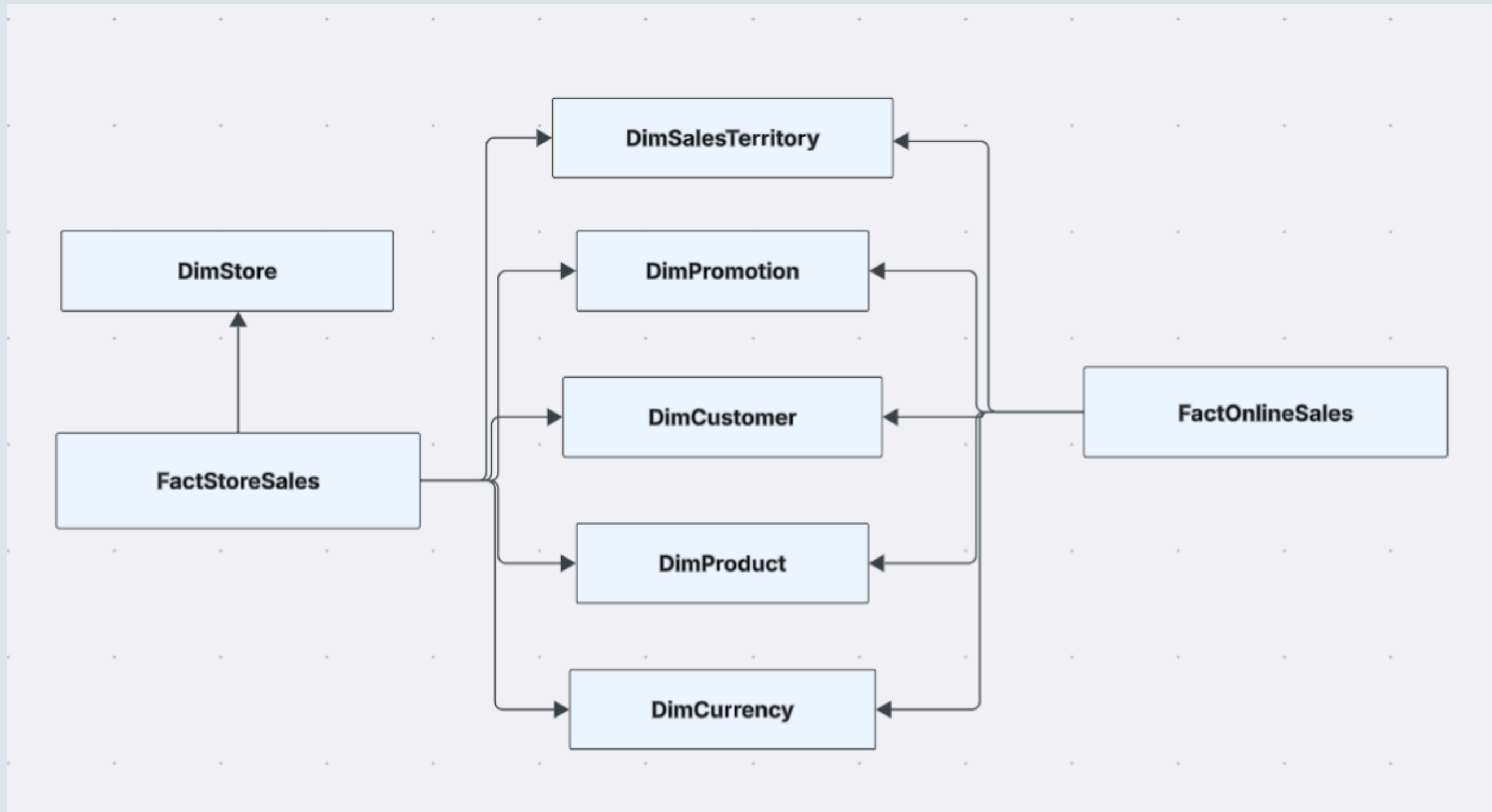
3 Transform: Clean, validate, apply rules

4 Incremental Load: Load only new/updated data

5 Load: Into AdventureWork_DWH (Dim + Fact tables)

Data Modeling Design

For the AdventureWorks_DWH, the data model is designed using a Galaxy Schema approach



Project Task Breakdown(Environment)



To set up the environment we need to have:

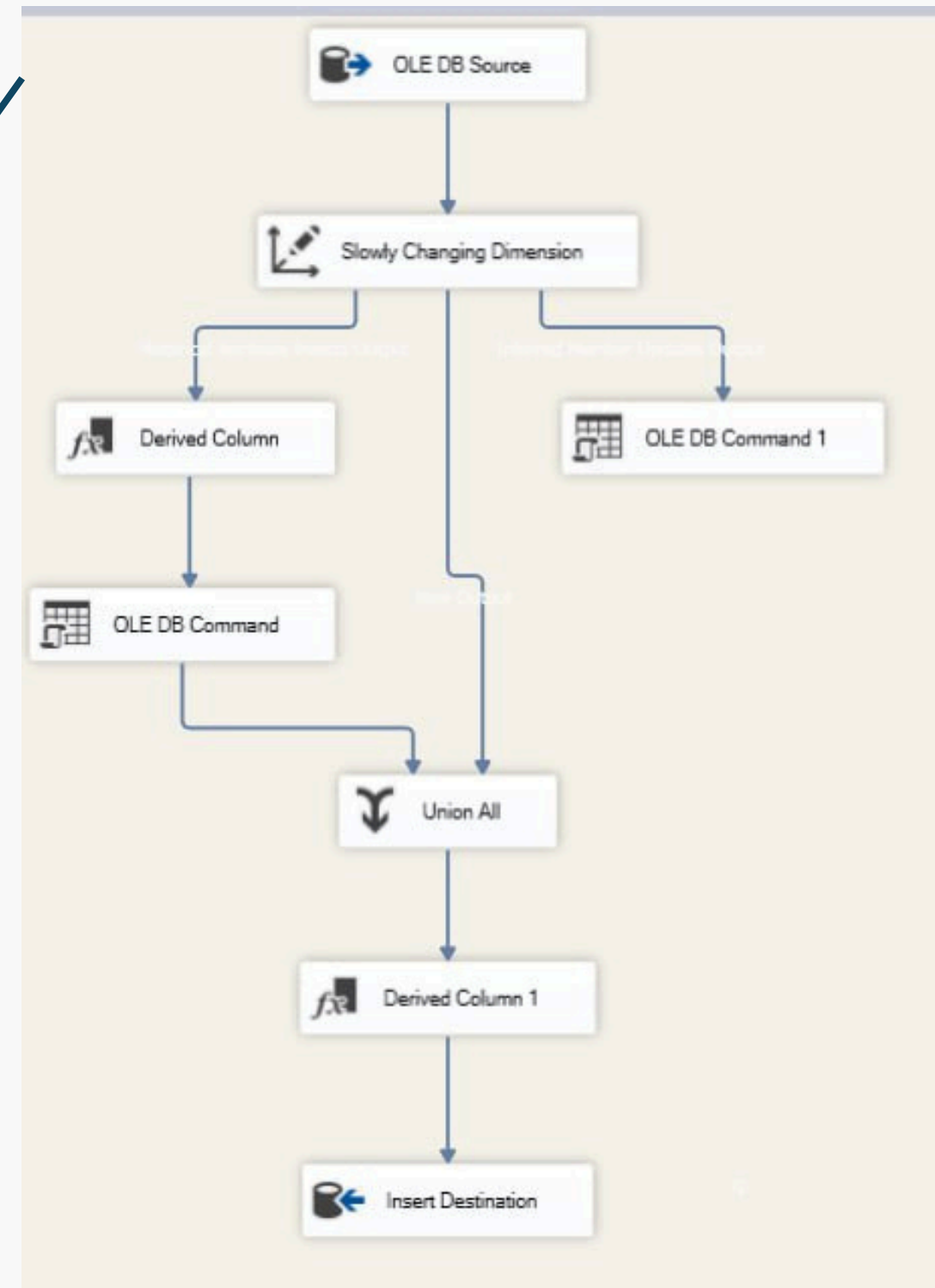
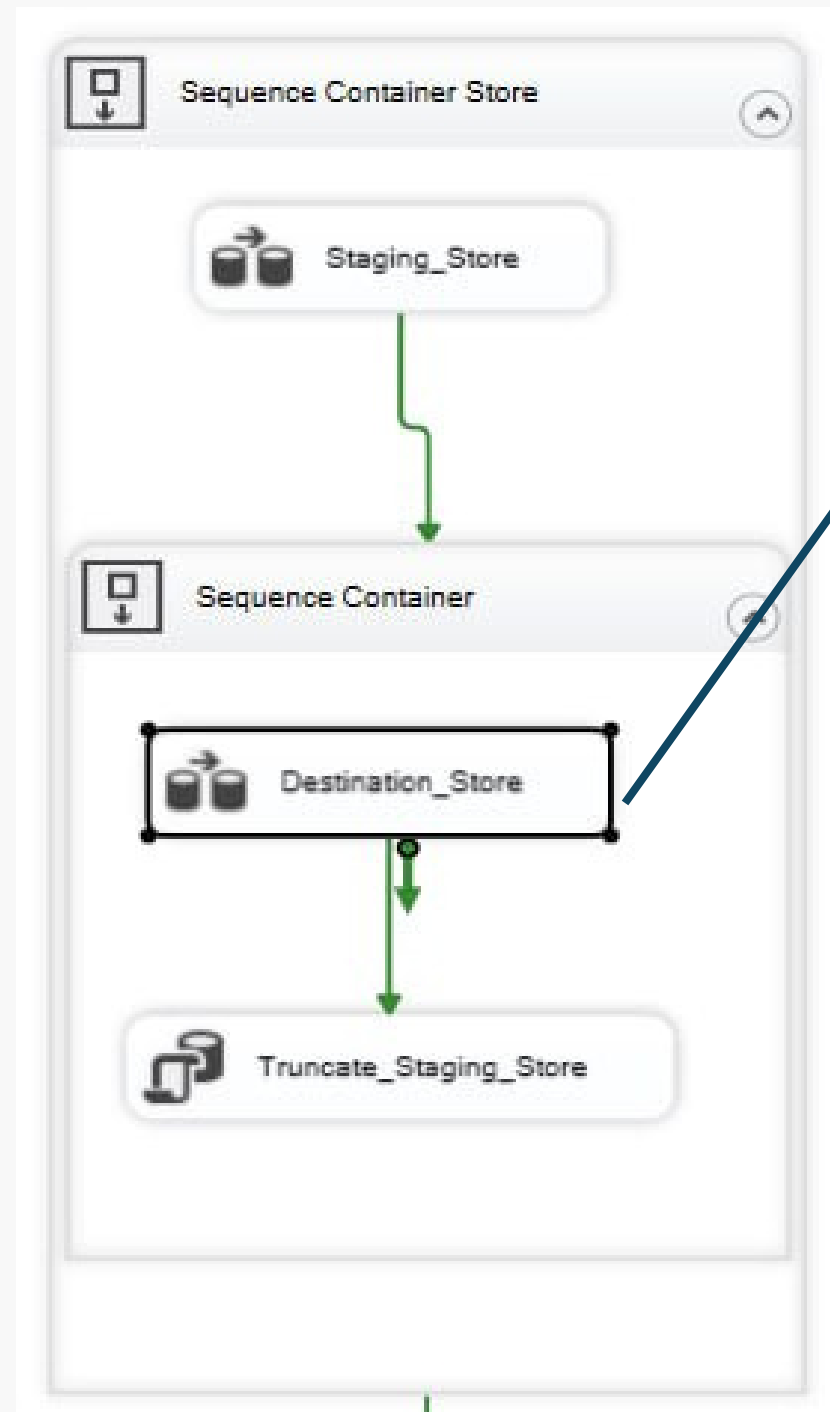
- Visual Studio
- Microsoft SQL Server

The Database needed:

- Source: AdventureWorks2022
- Staging: AdventureWorks_Staging
- Destination: AdventureWorks_DWH



Project Task Breakdown (ETL)



Project Task Breakdown (ETL)



Execute SQL Task Editor

Configure the properties required to run SQL statements and stored procedures using the selected connection.

General
Parameter Mapping
Result Set
Expressions

General

Name	Get Last Load Date Store
Description	Execute SQL Task

Options

TimeOut	0
CodePage	1252
TypeConversionMode	Allowed

Result Set

ResultSet	None
-----------	------

SQL Statement

ConnectionType	OLE DB
Connection	DESKTOP-912UJD3.AdventureWorks2022
SQLSourceType	Direct input
SQLStatement	IF NOT EXISTS (SELECT * FROM dbo.TEST_STORE_LAST_UPDATE_StoreSale)
IsQueryStoredProcedure	BEGIN
BypassPrepare	INSERT INTO dbo.TEST_STORE_LAST_UPDATE_StoreSale (DATE_LAST_UPDATE)
	SELECT DATEADD(DAY, -1, MIN(OrderDate))
	FROM dbo.vw_FactStoreSales;
	END

Name
Specifies the name of the task.

Browse... Build Query... Parse Query

OK Cancel Help

Project Task Breakdown (ETL)



Project Task Breakdown (ETL)



Execute SQL Task Editor

Configure the properties required to run SQL statements and stored procedures using the selected connection.

General
Parameter Mapping
Result Set
Expressions

General	
Name	Update Table
Description	Execute SQL Task
Options	
TimeOut	0
CodePage	1252
TypeConversionMode	Allowed
Result Set	
ResultSet	None
SQL Statement	
ConnectionType	OLE DB
Connection	DESKTOP-912UJD3.AdventureWorks2022
SQLSourceType	Direct input
SQLStatement	UPDATE [dbo].[TEST_STORE_LAST_UPDATE_StoreSale]
IsQueryStoredProcedure	SET DATE_LAST_UPDATE = (SELECT MAX(orderDate)
BypassPrepare	FROM [dbo].[vw_FactStoreSales])

Name
Specifies the name of the task.

Browse... Build Query... Parse Query

OK Cancel Help

Project Task Breakdown (Deploy and Schedule)

SQL Server Agent

Jobs

Daily ETL Report Sale

SSIS Server Maintenance Job

syspolicy_purge_history

AdventureWorks_DWH_Dalin

Job Activity Monitor



message_alert



Alert: ETL Success



Job Schedule Properties - LoadToDataWarehouse

Name: LoadToDataWarehouse Jobs in Schedule

Schedule type: Recurring Enabled

One-time occurrence

Date: 07-Nov-2025 Time: 11:42:02 PM

Frequency

Occurs: Daily

Recurs every: 1 day(s)

Daily frequency

☒ Occurs once at: 6:00:00 AM

☐ Occurs every: 1 hour(s)

Starting at: 6:00:00 AM

Ending at: 11:59:59 PM

Duration

Start date: 07-Nov-2025 ☐ End date: 07-Nov-2025

☒ No end date:

Summary

Description: Occurs every day at 6:00:00 AM. Schedule will be used starting on 07-Nov-2025.

OK Cancel Help



Thank you

