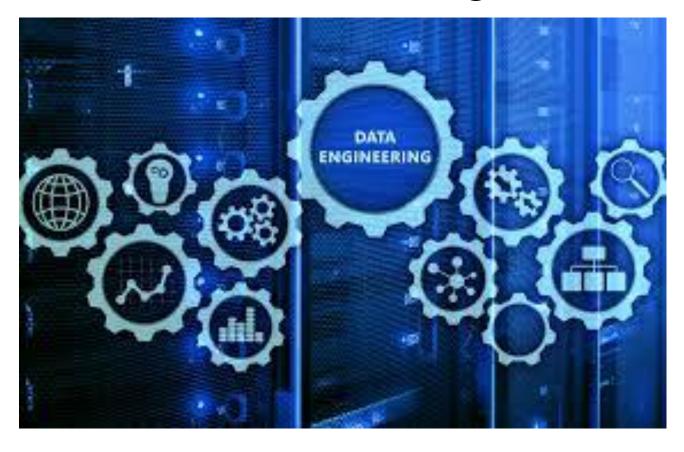
Customer 360 Data Integration



Contents

Project Overview	3
GitHub link	
Tools required	
Architecture	
Bronze/Raw Layer	
Silver Layer	
Gold Layer	12
PowerBI	17

Project Overview

The Customer 360 Data Integration project aims to create a unified view of customers for a retail business by integrating data from online transactions, in-store purchases, customer service interactions, and loyalty programs. This comprehensive view helps the business better understand customer behavior, personalize experiences, and support data-driven decisions. The project uses an end-to-end data pipeline built on the Azure ecosystem, including Azure Synapse Analytics, Azure Data Lake Storage (ADLS), Azure SQL Database, and Power BI. It follows a layered architecture—from raw data ingestion to staging, analytics modeling, and final visualization.

Outcomes

Raw Layer: Source data stored in ADLS containers.

Curated Layer (Silver): Cleaned and joined data in Synapse Analytics.

Analytics Layer (Gold): Aggregated tables and customer insights in Azure SQL Database.

Power BI Dashboard: Interactive Customer 360 dashboard visualizing purchasing behaviour, service interactions, and loyalty metrics.

GitHub link

https://github.com/DeepthiChethi/Azure_DE/upload/Madhumita/Project3.

Tools required

ADLSGEN2

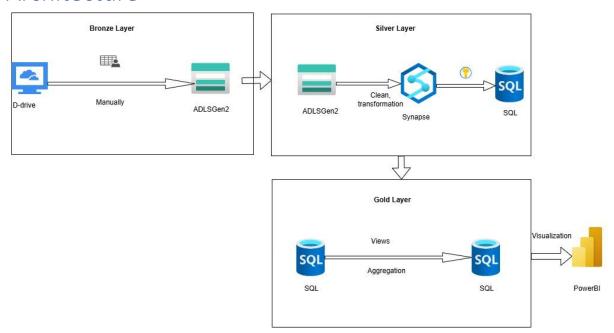
Azure Synapse Analytics

Key Vault

Azure SQL Database

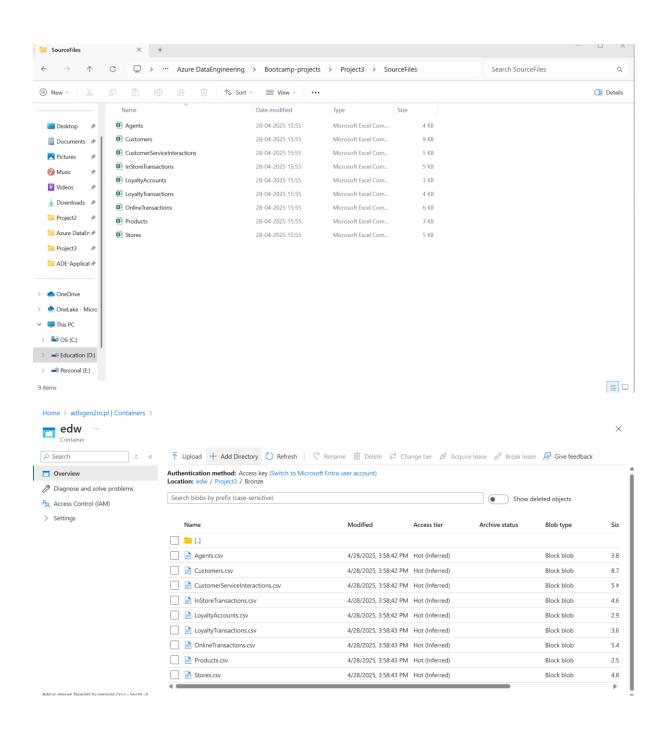
PowerBI

Architecture



Bronze/Raw Layer

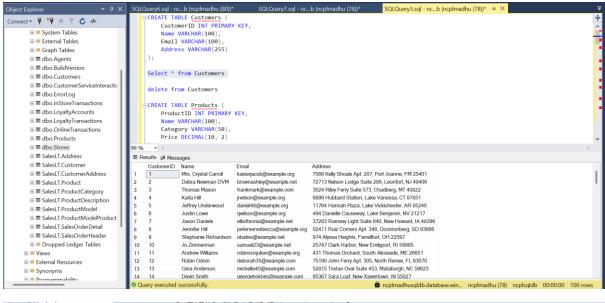
Download the 9 files from <u>Kaggle's Customer 360 Data</u> and save them to the D-drive. Then, manually upload all 9 CSV files to Azure Data Lake Storage Gen2 (ADLS Gen2).

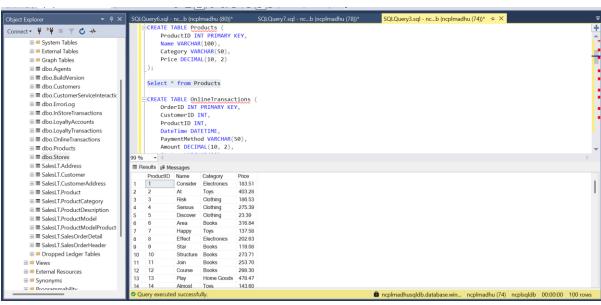


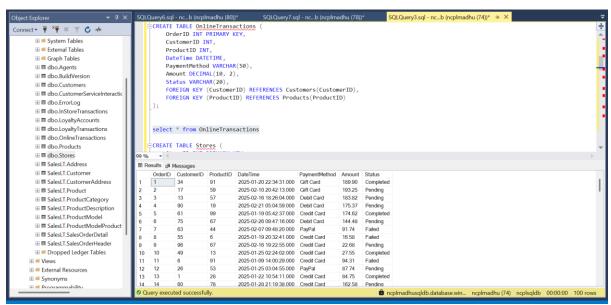
Silver Layer

Create the tables in Azure SQL tables.

Tables



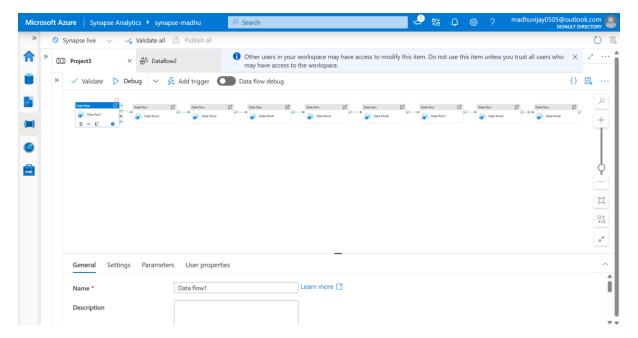








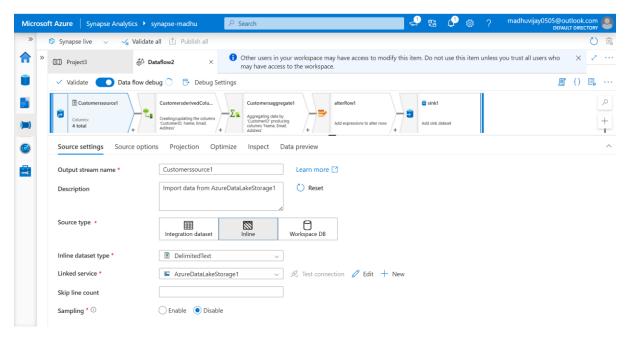
Create dataflows in Azure Synapse Analytics to get data from bronze layer and clean, transformation the data and load it into Azure SQL database.

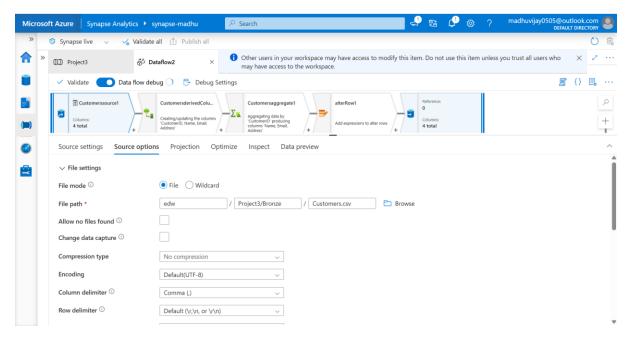


Customer File

Source

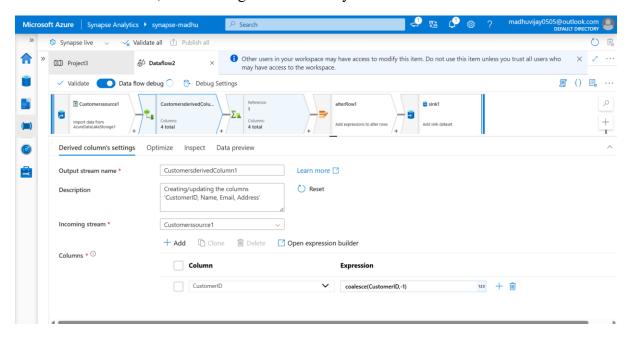
To get data from ADLSGen2, we use source activity





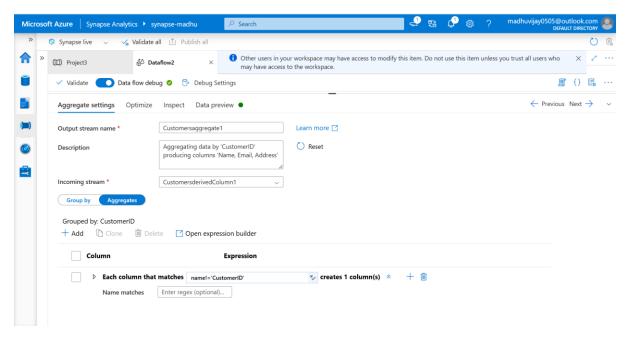
Derived Activity

To remove null values, we are using derived activity



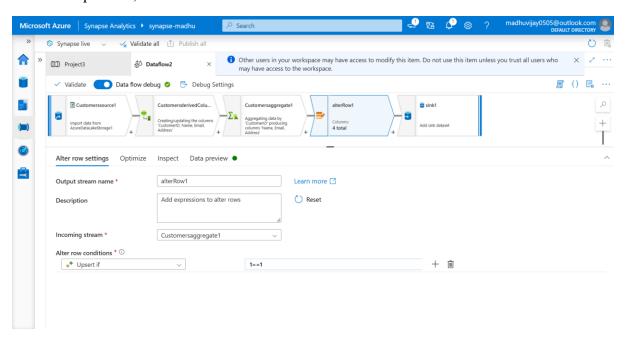
Aggregate Activity

To remove the duplicate values, we are using aggregate activity.



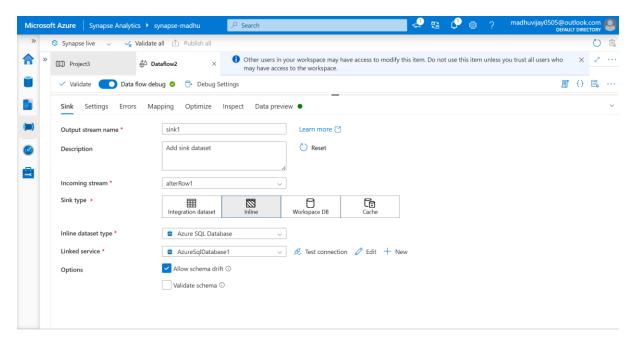
AlterRow

To set row policies, we use AlterRow



Sink

Loading data to SQL database, we use sink activity



Repeat the same steps for other files too.

Gold Layer

Views-1

```
CREATE VIEW View_AverageOrderValue AS

SELECT

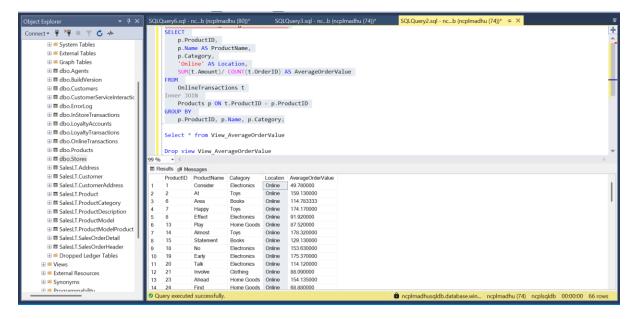
p.ProductID,
p.Name AS ProductName,
p.Category,
    'Online' AS Location,

SUM(t.Amount)/ COUNT(t.OrderID) AS AverageOrderValue

FROM
OnlineTransactions t

Inner JOIN
Products p ON t.ProductID = p.ProductID

GROUP BY
p.ProductID, p.Name, p.Category;
```



VIEW 2

SELECT

```
CREATE VIEW View_CustomerSegmentation AS
WITH CustomerSpending AS (
  SELECT
    c.CustomerID,
    c.Name,
    SUM(t.Amount) AS TotalSpend,
    COUNT(t.OrderID) AS PurchaseFrequency,
    1.TierLevel
  FROM
    dbo.Customers c
  LEFT JOIN
    OnlineTransactions t ON c.CustomerID = t.CustomerID
  LEFT JOIN
    LoyaltyAccounts 1 ON c.CustomerID = 1.CustomerID
  GROUP BY
    c.CustomerID, c.Name, 1.TierLevel
```

CustomerID,

Name,

TotalSpend,

PurchaseFrequency,

TierLevel,

CASE

WHEN TotalSpend >= (SELECT PERCENTILE_CONT(0.9) WITHIN GROUP (ORDER BY TotalSpend) OVER ()) THEN 'High-Value Customer'

WHEN PurchaseFrequency = 1 THEN 'One-Time Buyer'

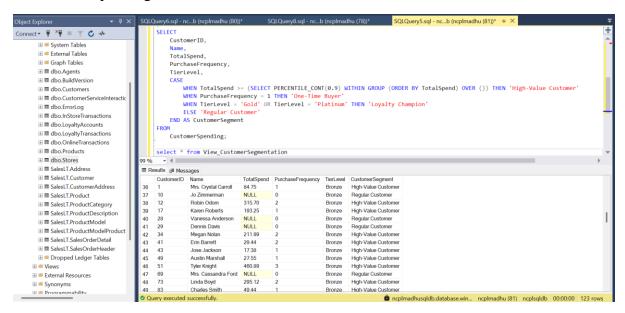
WHEN TierLevel = 'Gold' OR TierLevel = 'Platinum' THEN 'Loyalty Champion'

ELSE 'Regular Customer'

END AS CustomerSegment

FROM

CustomerSpending;



VIEW 3

CREATE VIEW View PeakTrafficAnalysis AS

SELECT

DATENAME(WEEKDAY, DateTime) AS DayOfWeek,

DATEPART(HOUR, DateTime) AS HourOfDay,

'Online' AS Channel,

COUNT(*) AS TransactionCount

FROM

OnlineTransactions

GROUP BY

DATENAME(WEEKDAY, DateTime),

DATEPART(HOUR, DateTime)

UNION ALL

SELECT

DATENAME(WEEKDAY, DateTime) AS DayOfWeek,

DATEPART(HOUR, DateTime) AS HourOfDay,

'In-Store' AS Channel,

COUNT(*) AS TransactionCount

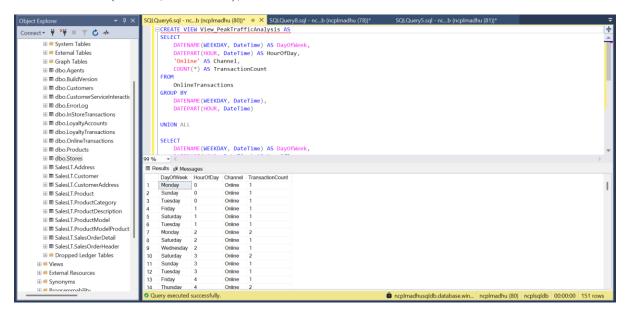
FROM

InStoreTransactions

GROUP BY

DATENAME(WEEKDAY, DateTime),

DATEPART(HOUR, DateTime);



VIEW 4

CREATE VIEW View_AgentInteractionStats AS

SELECT

- a.AgentID,
- a. Name AS AgentName,
- a.Department,
- a.Shift,

COUNT(csi.InteractionID) AS TotalInteractions,

SUM(CASE WHEN csi.ResolutionStatus = 'Resolved' THEN 1 ELSE 0 END) AS ResolvedCount,

CAST(

100.0 * SUM(CASE WHEN csi.ResolutionStatus = 'Resolved' THEN 1 ELSE 0 END)

/ NULLIF(COUNT(csi.InteractionID), 0) AS DECIMAL(5,2)

) AS ResolutionRatePercent

FROM

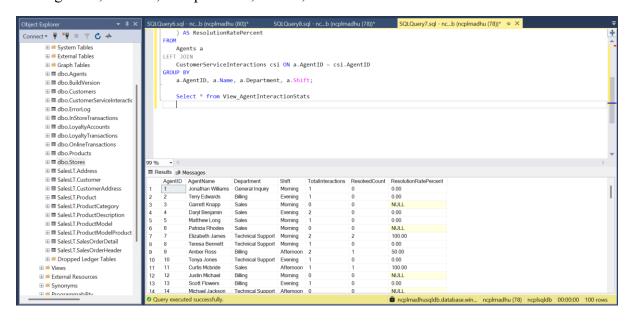
Agents a

LEFT JOIN

CustomerServiceInteractions csi ON a.AgentID = csi.AgentID

GROUP BY

a.AgentID, a.Name, a.Department, a.Shift;



PowerBI

