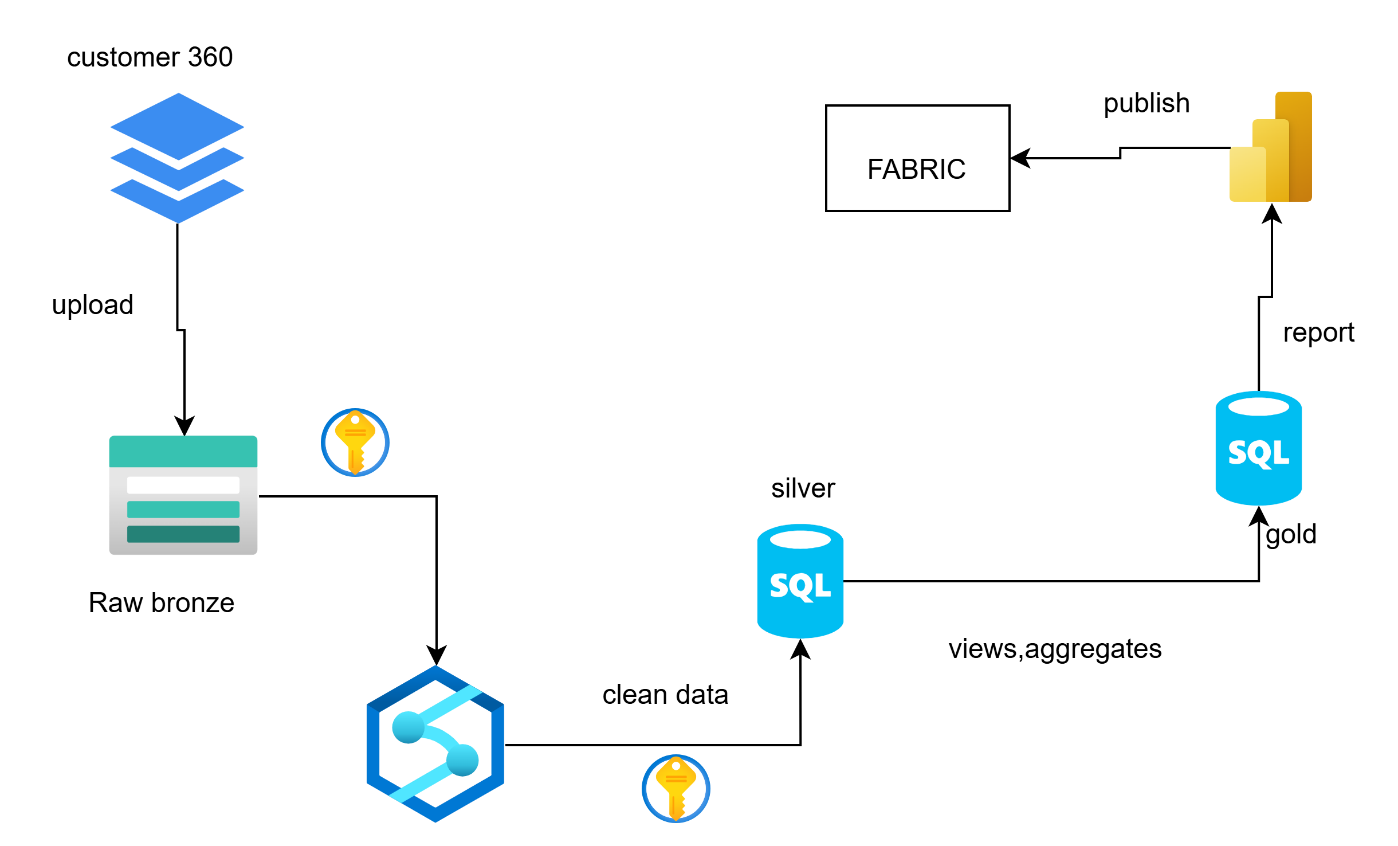
***CUSTOMER 360***

***DATA INTEGRATION***

**Overview**

A retail business wants to build a unified Customer 360 view by integrating data from multiple sources, including online transactions, in-store purchases, customer service interactions, and loyalty programs. This project uses a mix of fact and dimension tables to ensure a clean, scalable structure.

ARCHITECTURE



### **Tools and Technologies**

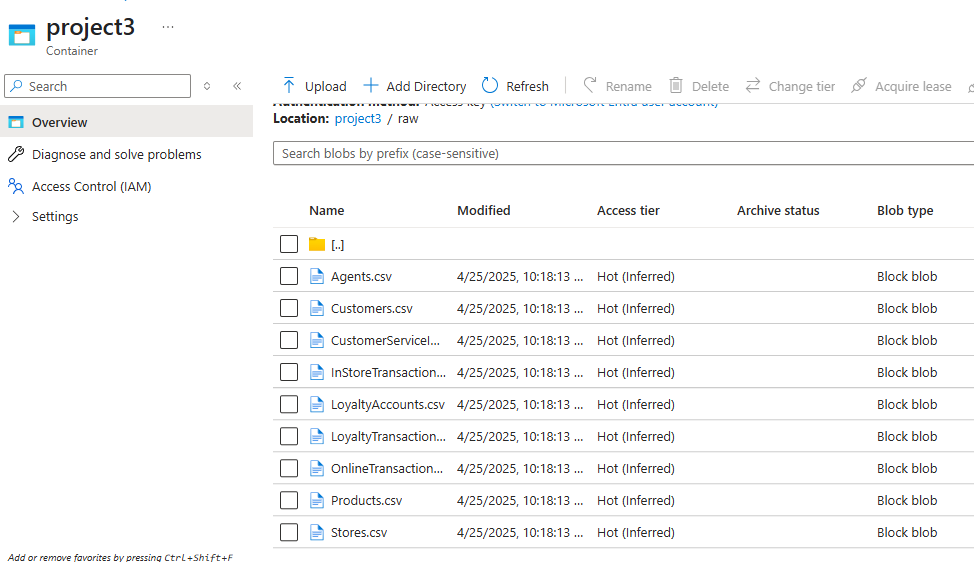
* Azure Synapse Analytics
* Azure Data Lake Storage (ADLS)
* Azure SQL Database
* Power BI

**STEP-1: Ingest Data**

**Reference Dataset** For realistic sample data, refer to [Kaggle’s Customer 360 Data](https://www.kaggle.com/datasets/varunkumari/customer-360-data).

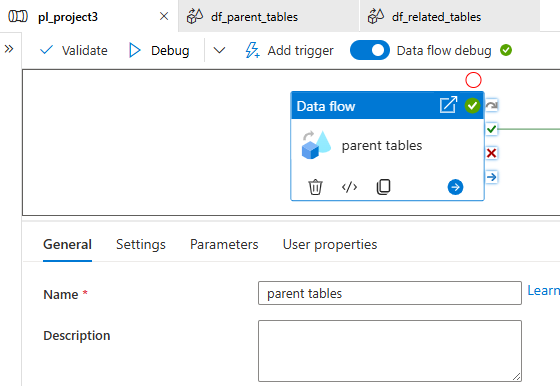
Download the above customer 360 dataset from the provided link to local system.

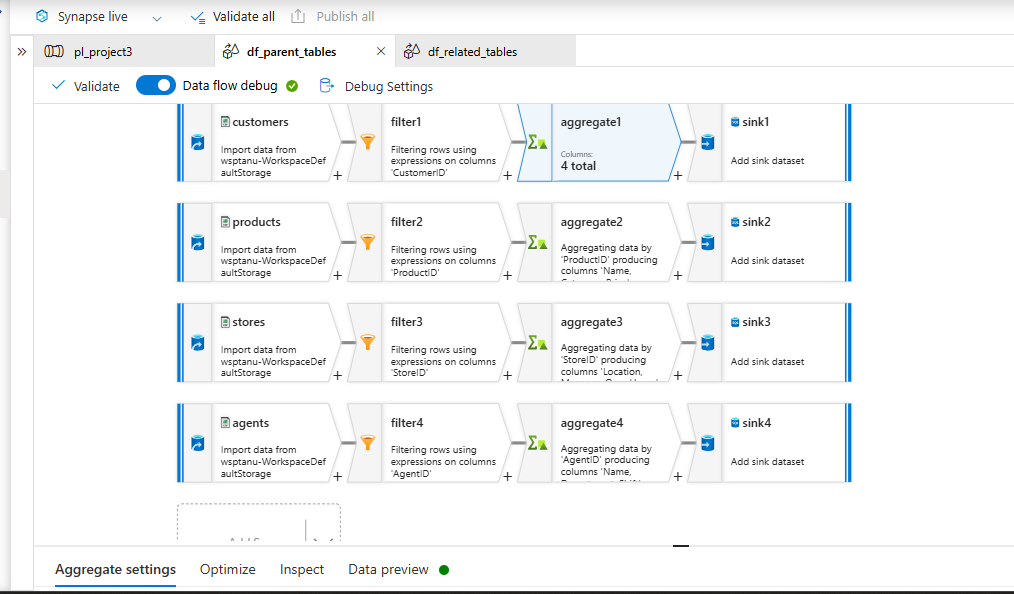
Now upload all the files in the dataset to project3 container raw folder in ADLS GEN 2 storage account.



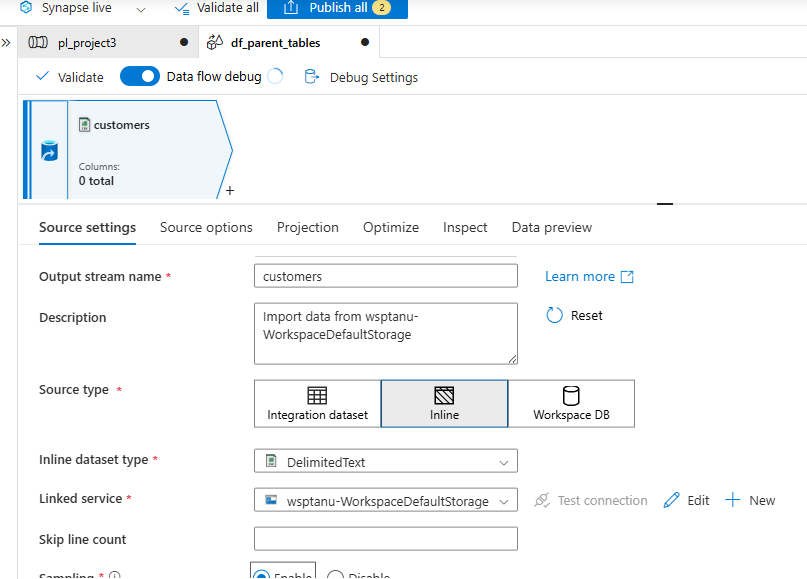
**STEP-2: Define Staging (curated) schema**

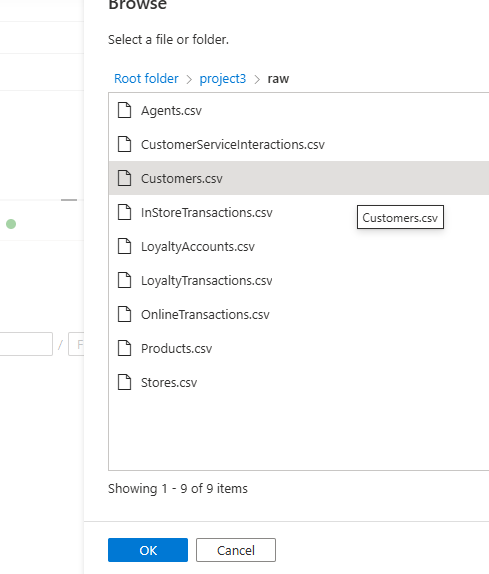
To implement cleaning, let’s define the dataflows in a pipeline as follows

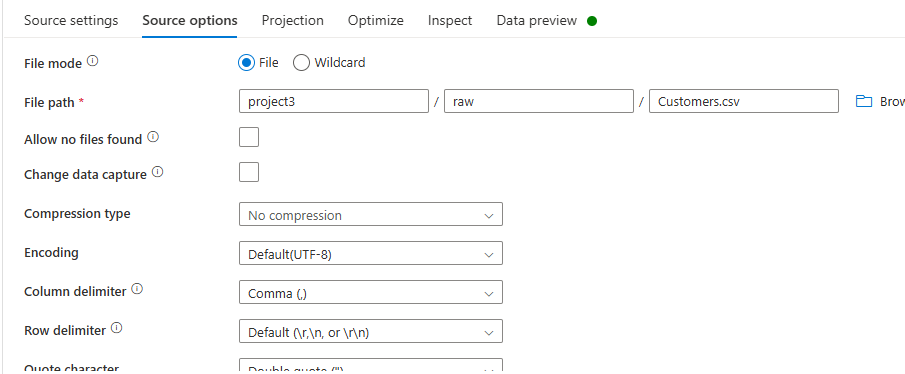




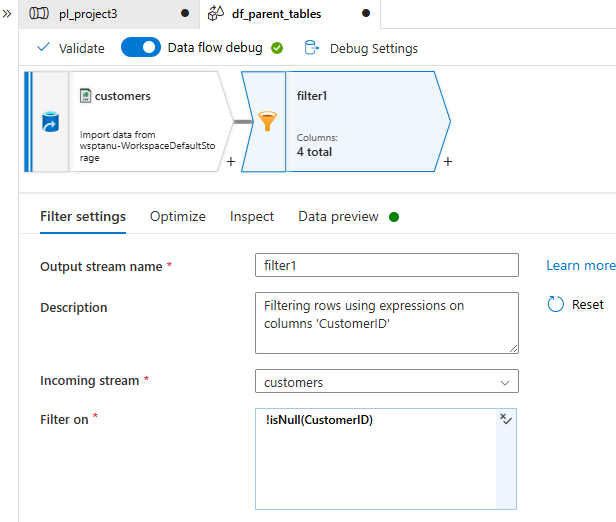
Let us take first source as customers in ADLS GEN 2



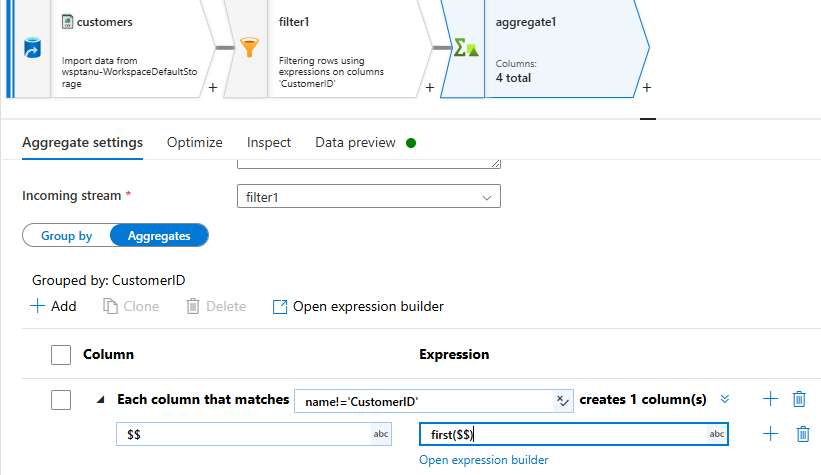


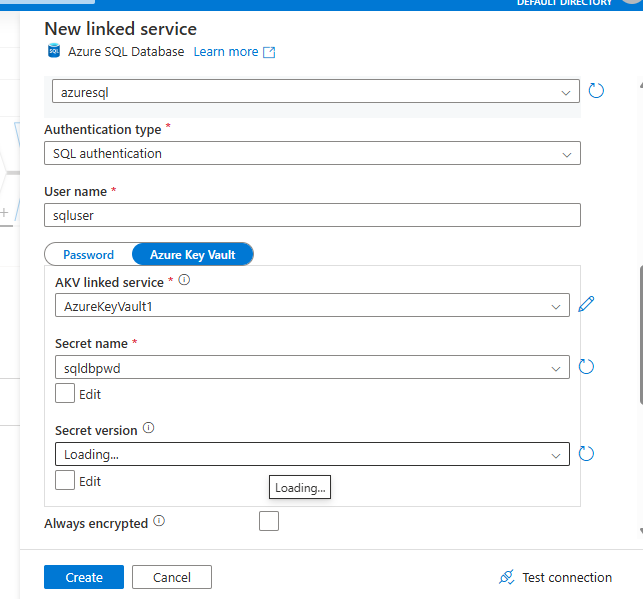


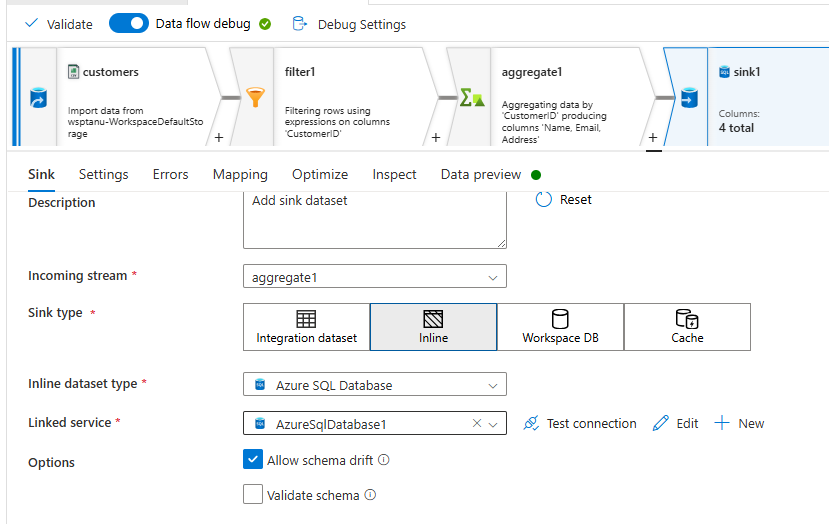
Add filter to remove nulls

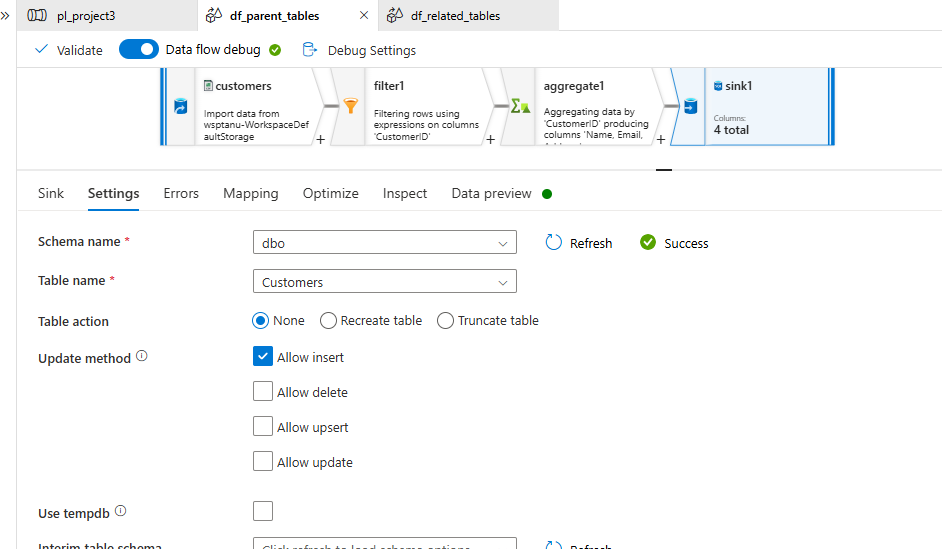


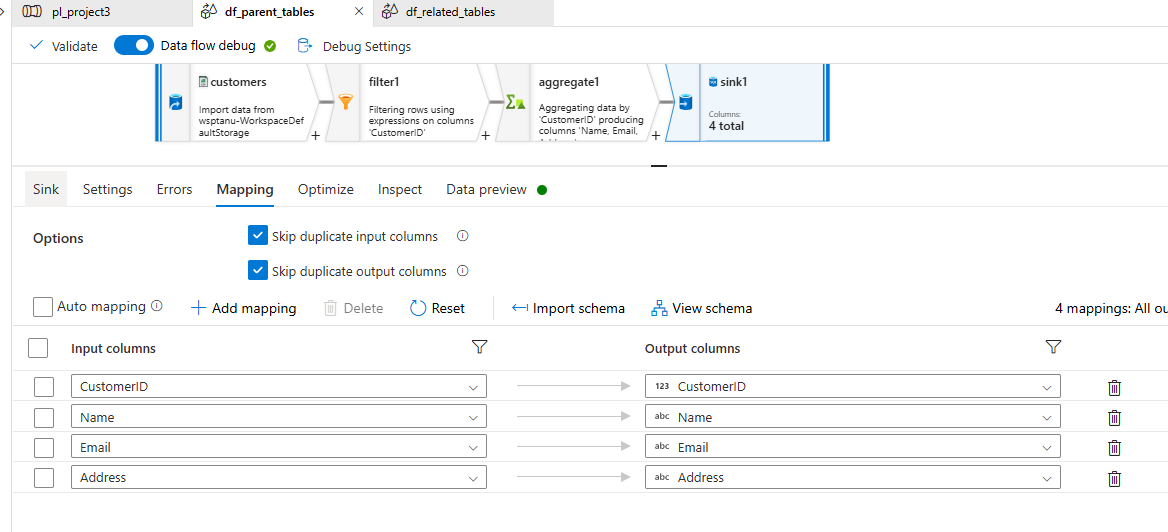
Now add aggregate to remove any duplicates



Now add sink here as follows

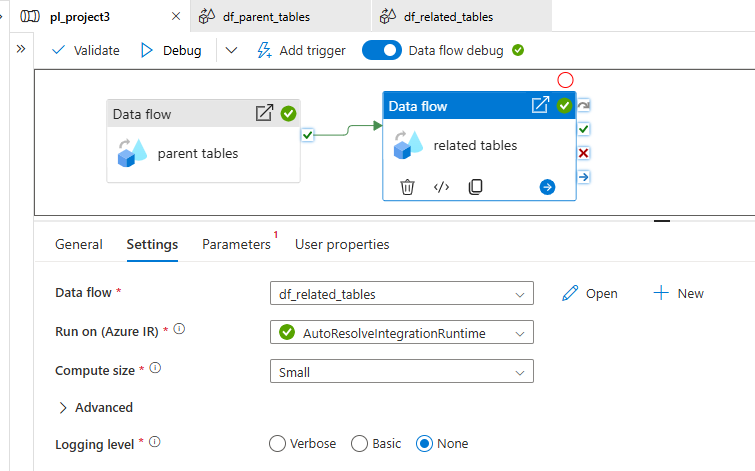


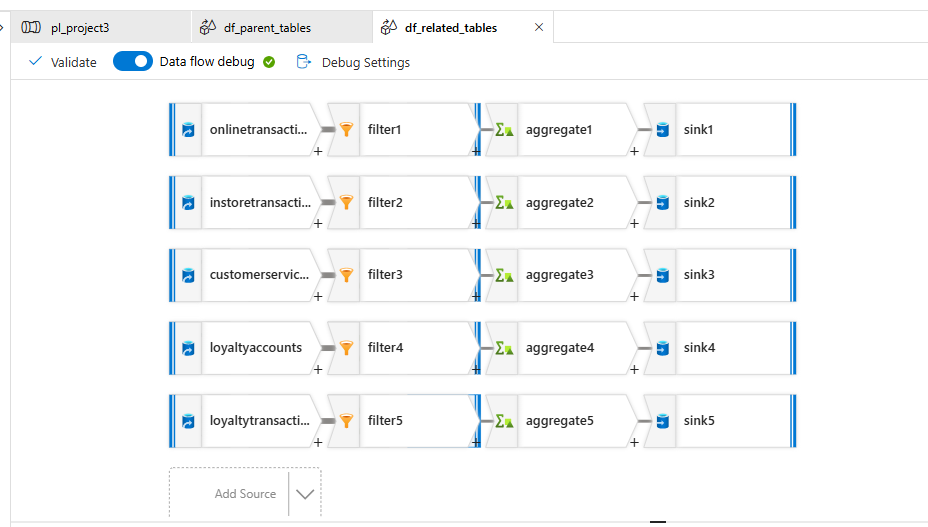




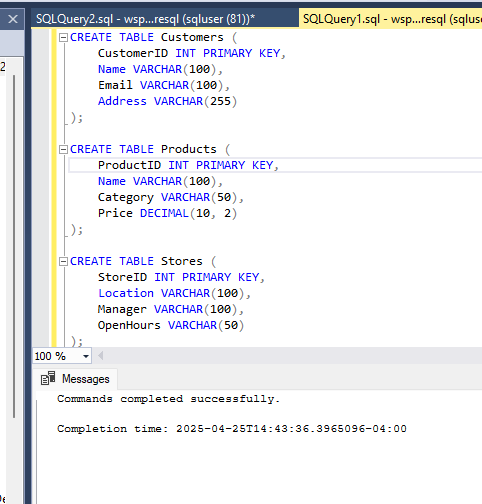
In the same way we add products, stores, agents to the same dataflow and add filter, aggregate and sink to their respective tables.

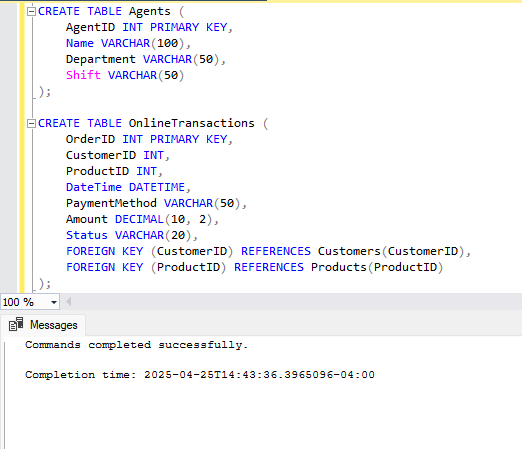
Here again we take another dataflow as related tables to parent tables as follows

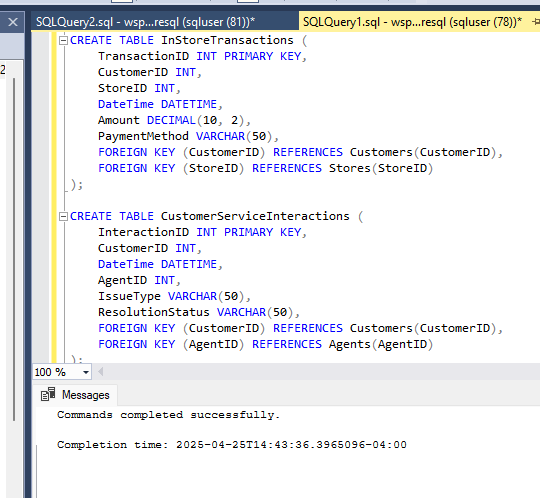


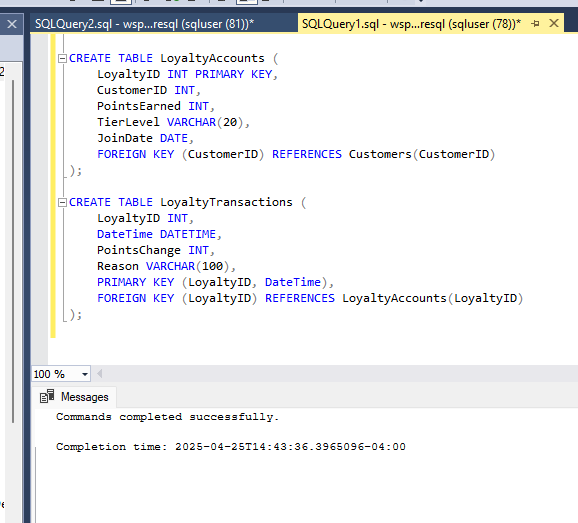


Here we have added remaining data sources (online transactions, instore transactions, customer service interactions, loyalty accounts, loyalty transactions) and cleaned then later given the respective sink tables.







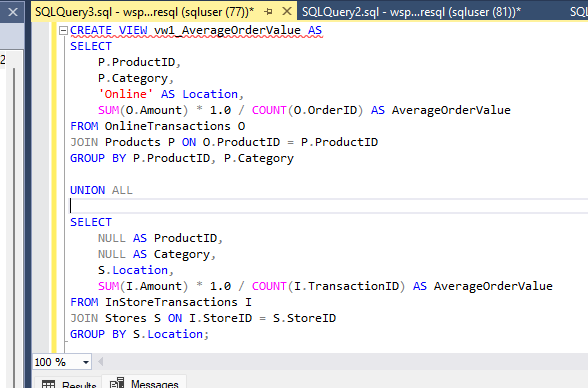


The above tables are defined and then provided as sink.

**STEP-3: Define Analytics (gold) Schema**

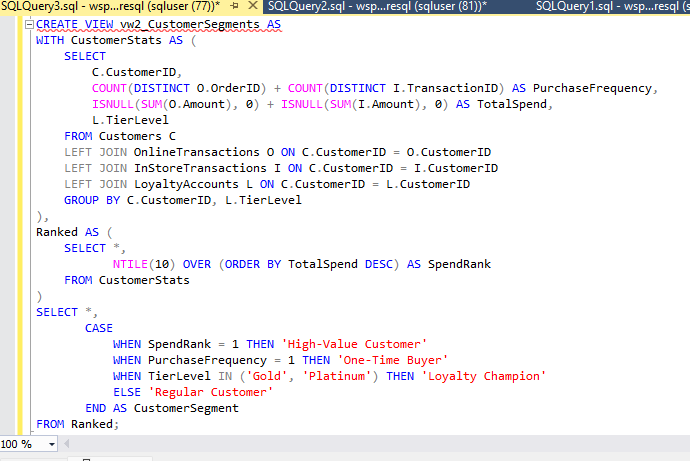
View 1 - for Average Order Value (AOV)

SUM(Amount) / COUNT(OrderID) per product, category, and location.

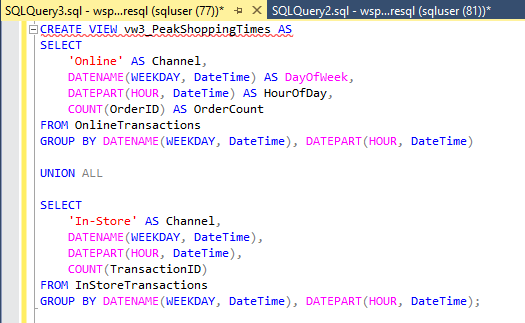


View 2 - for Segment customers based on total spend, purchase frequency, and loyalty tier (**LoyaltyAccounts.TierLevel**).

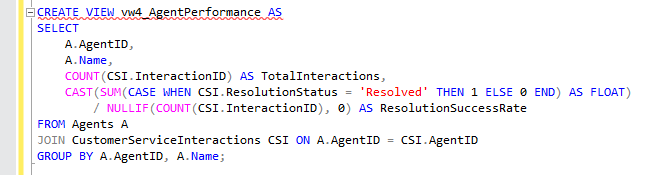
Example: "High-Value Customers" (Top 10% spenders), "One-Time Buyers," "Loyalty Champions."



View 3 - for Analyze **DateTime** to find peak days and times in-store vs. online.

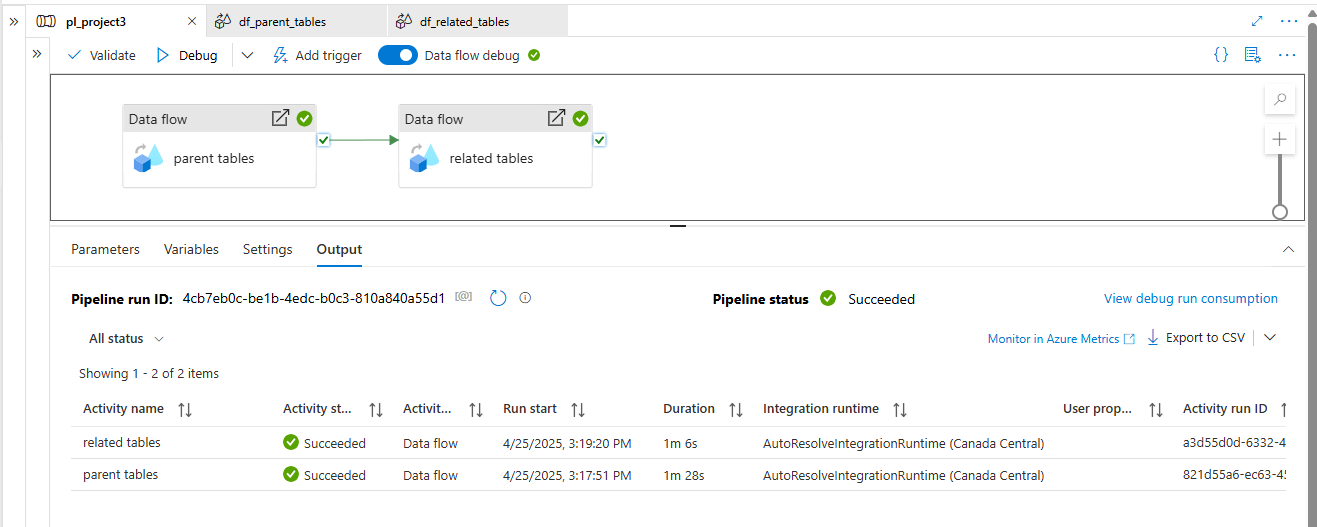


View 4 - for Number of interactions and resolution success rates per agent (**ResolutionStatus**).

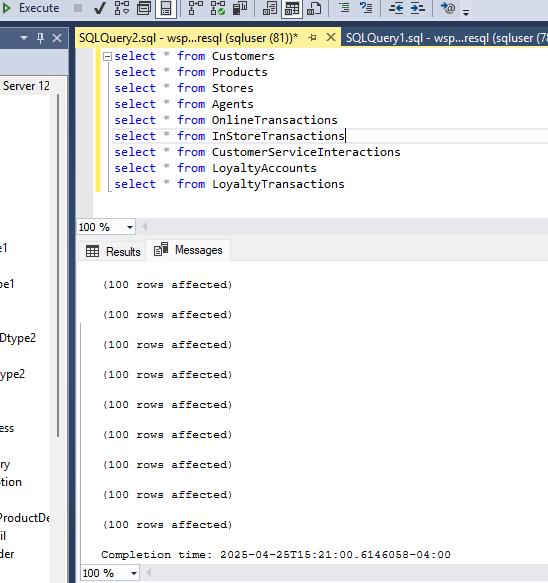


**STEP-4: Load Curated Data**

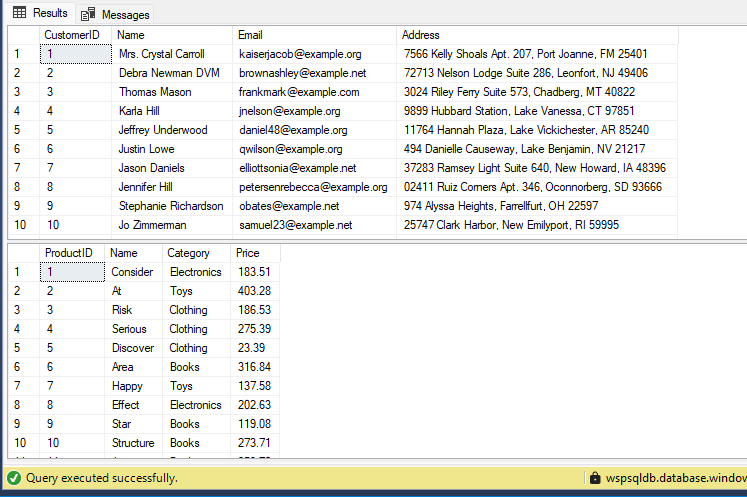
To load data to staging database tables, here we debug the pipeline which we created earlier.

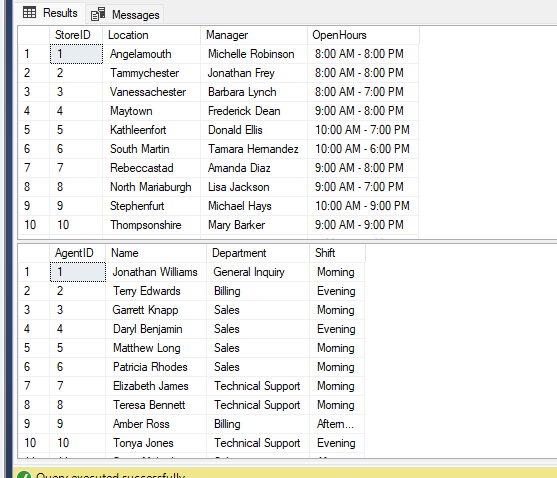


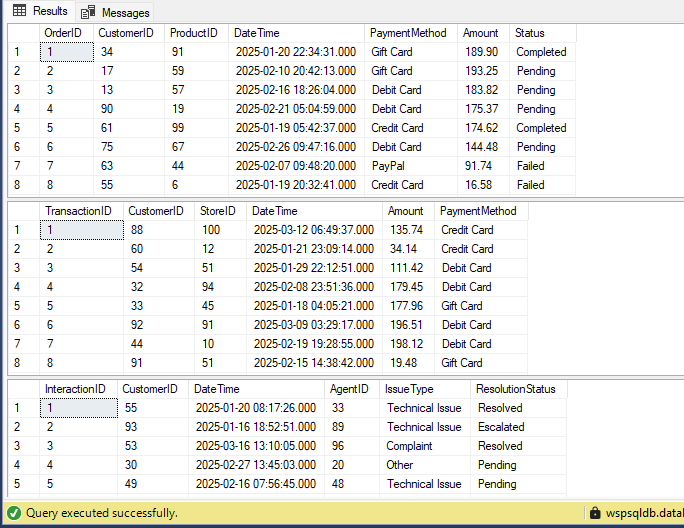
Once the pipeline execution was successful, we check for data in SQL Database.

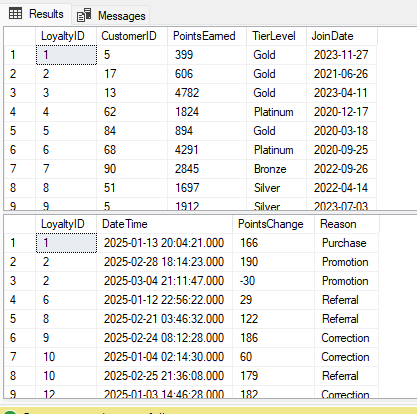


We can see all the tables are loaded with data



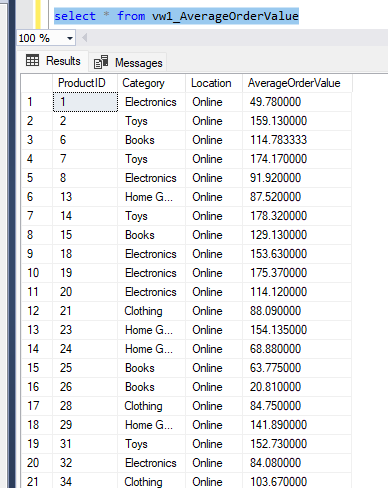


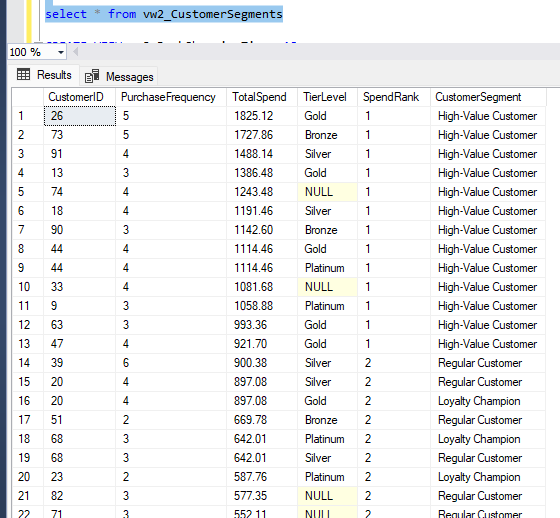


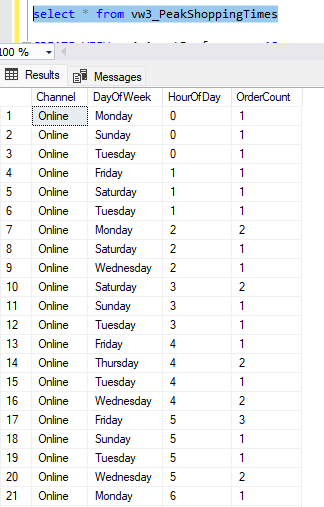


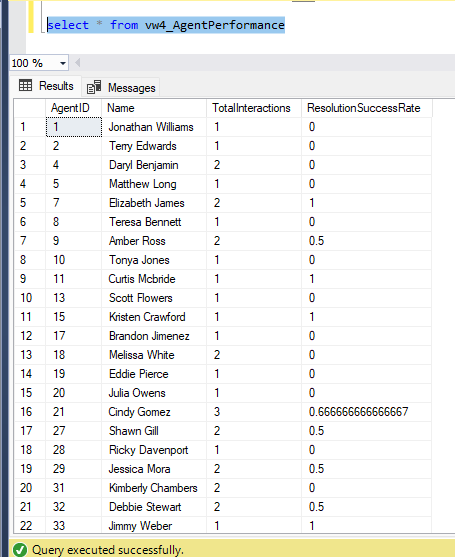
**STEP-5: Load Analytics Data**

To load the Analytics data we execute the views which we created in step 2



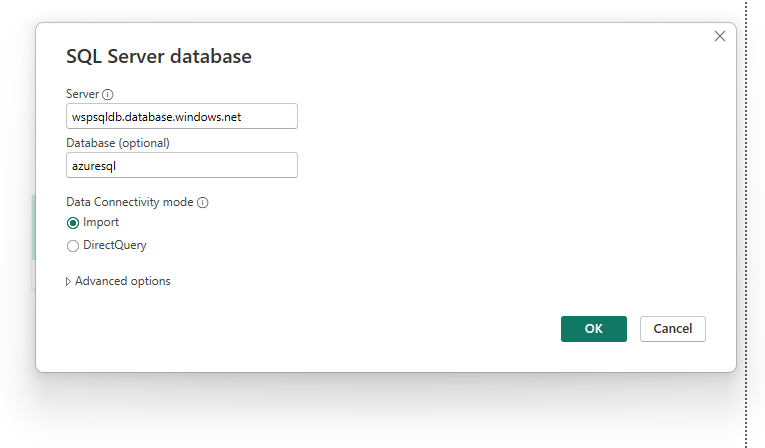




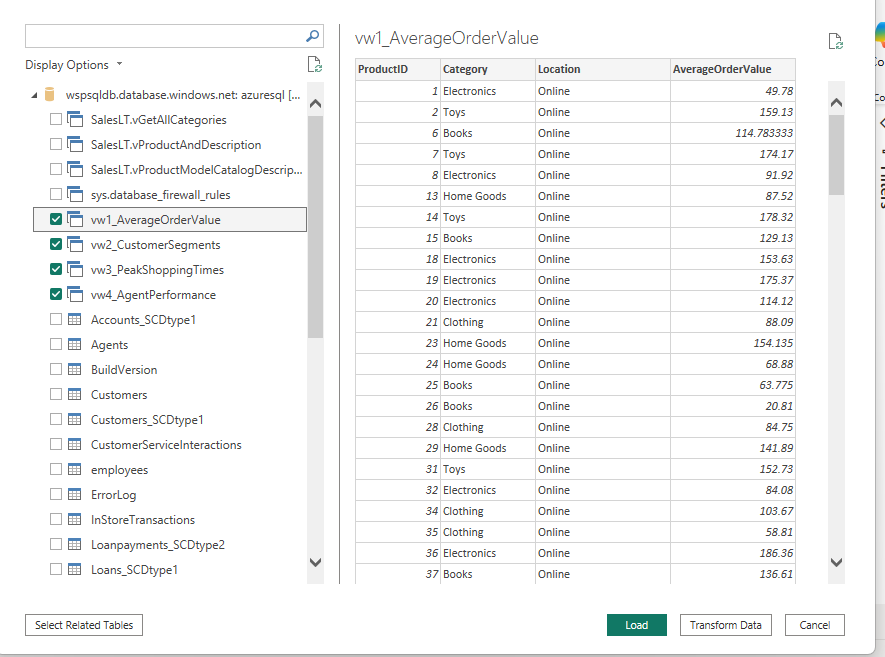


**STEP-6: Build a Power BI Dashboard**

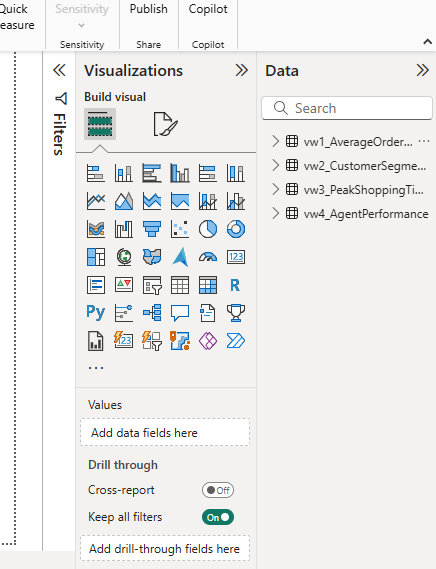
Go to power bi desktop 🡪home🡪sql server



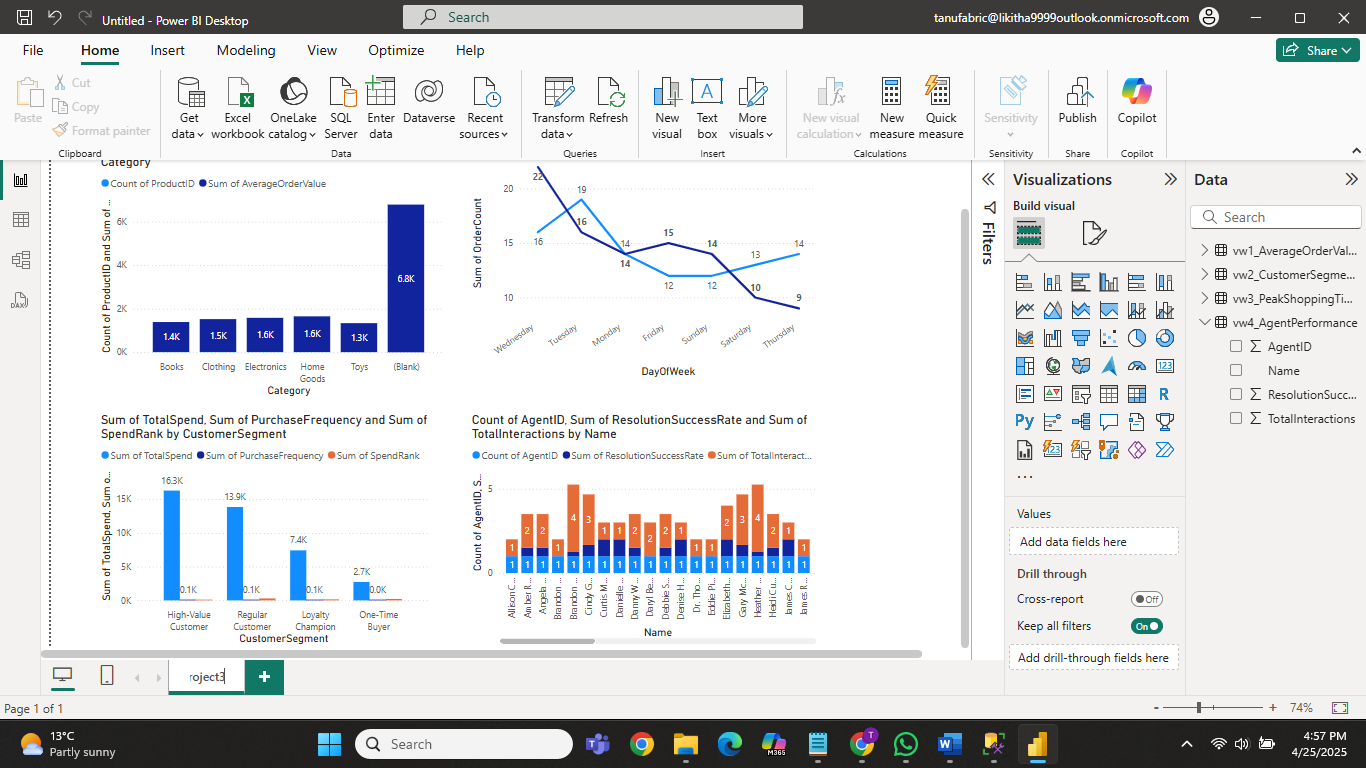
Select views in sql database and load



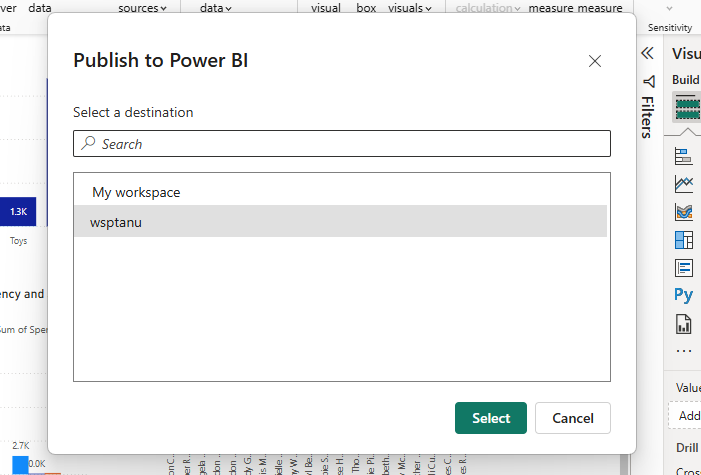
Data appears as follows

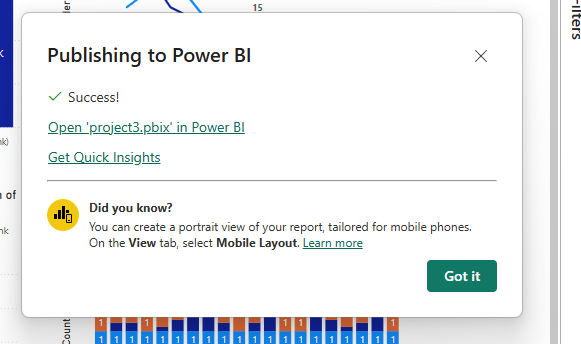


Expand views and generate report



Now save and publish this in fabric workspace



  
open fabric and see the report

