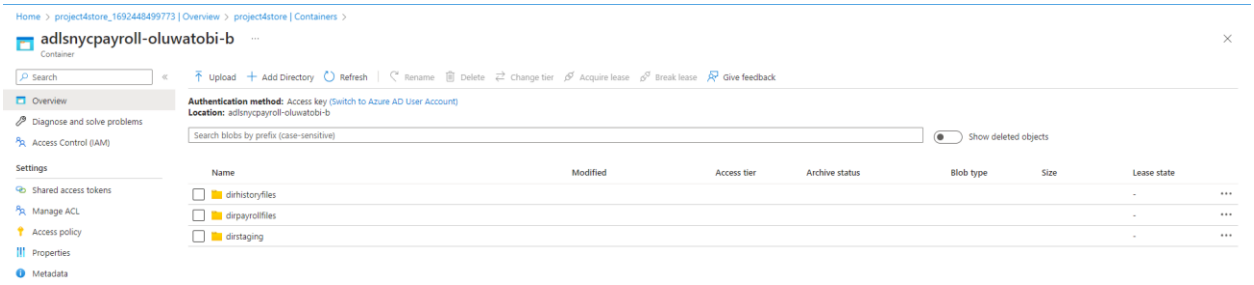
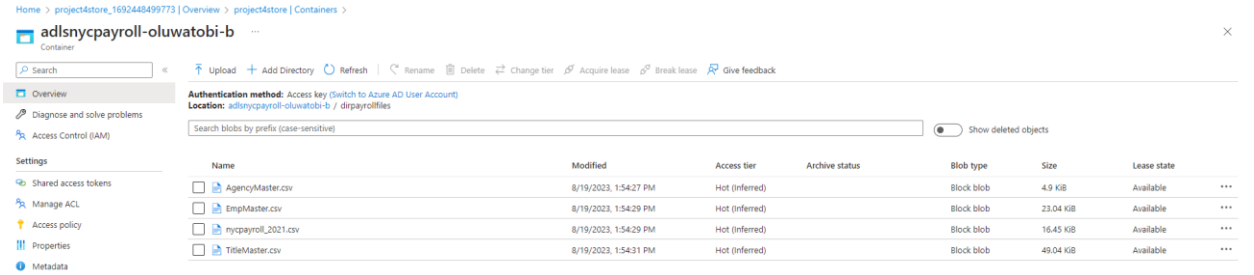
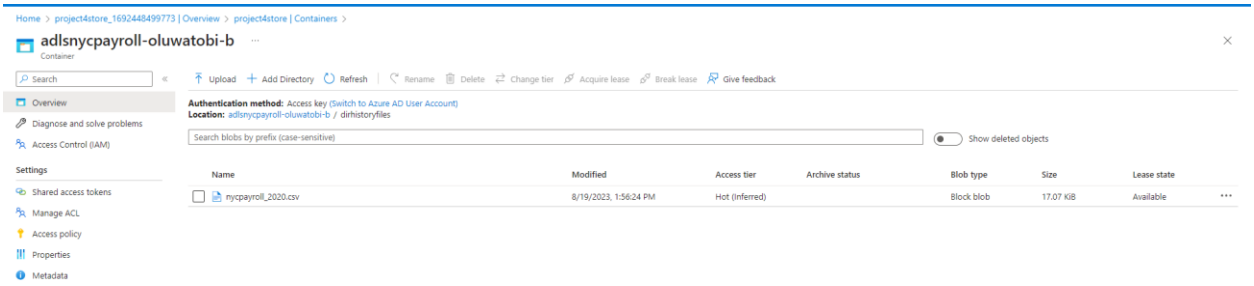
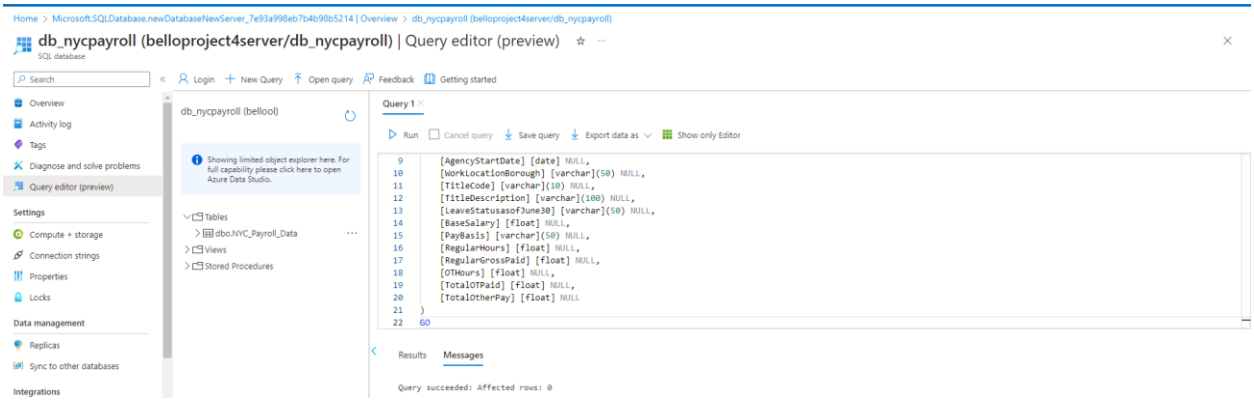
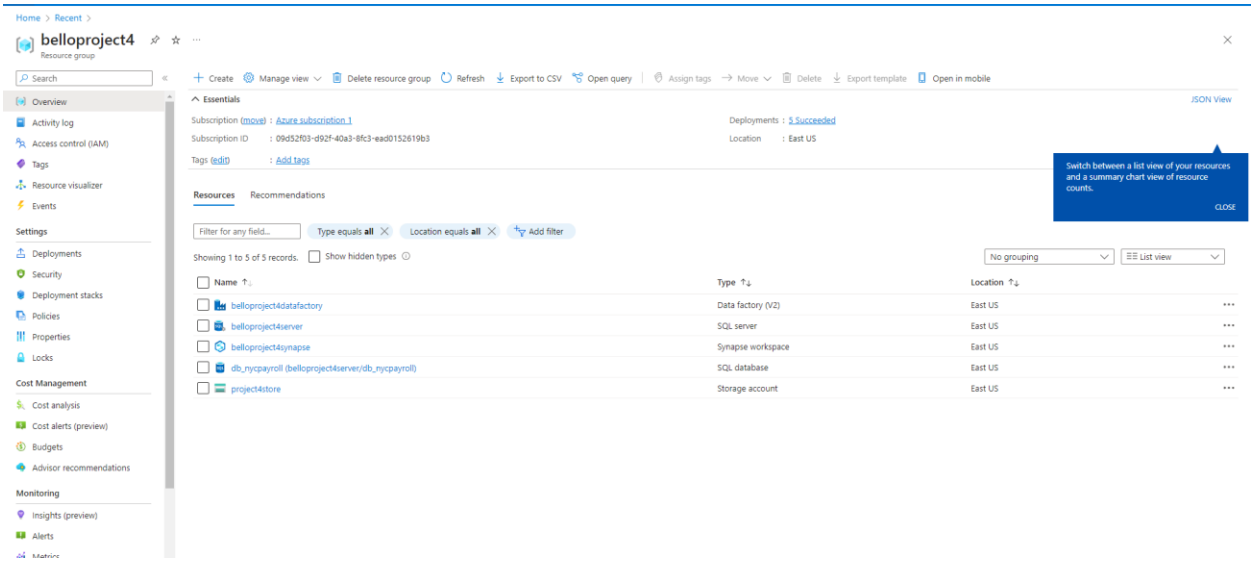
Step 1: Prepare the Data Infrastructure

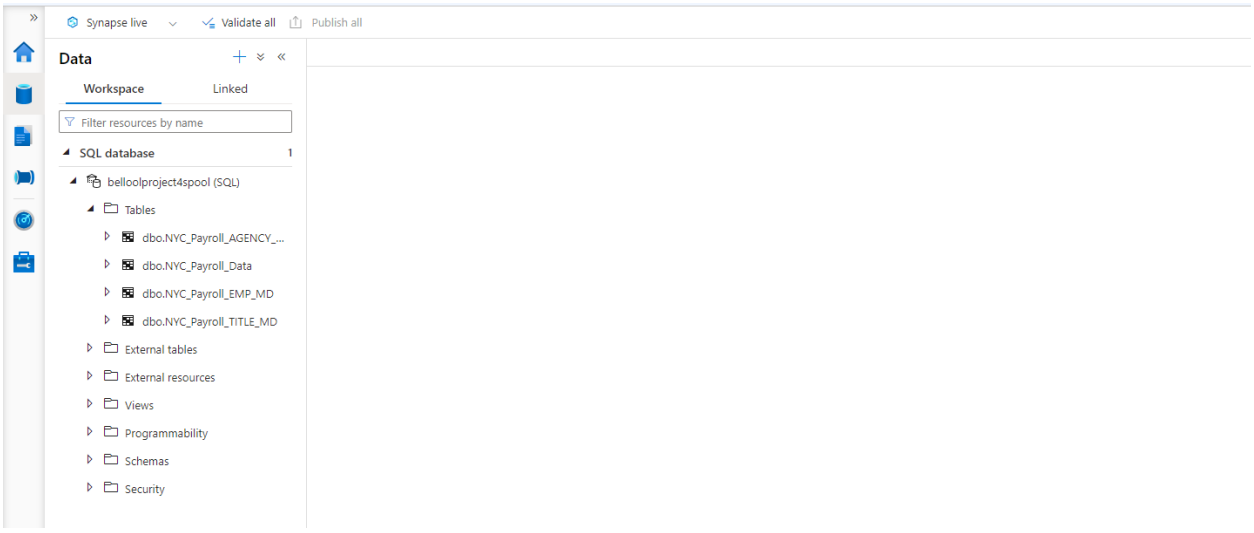




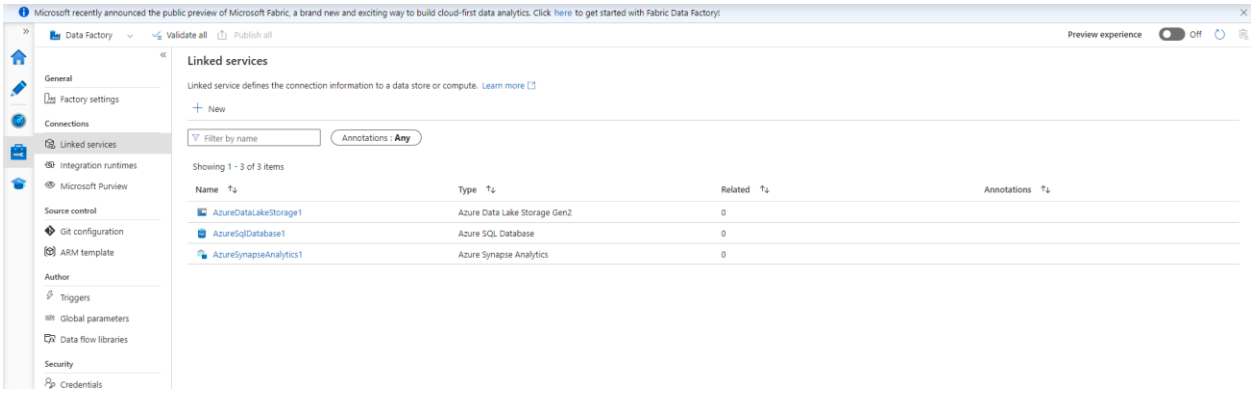




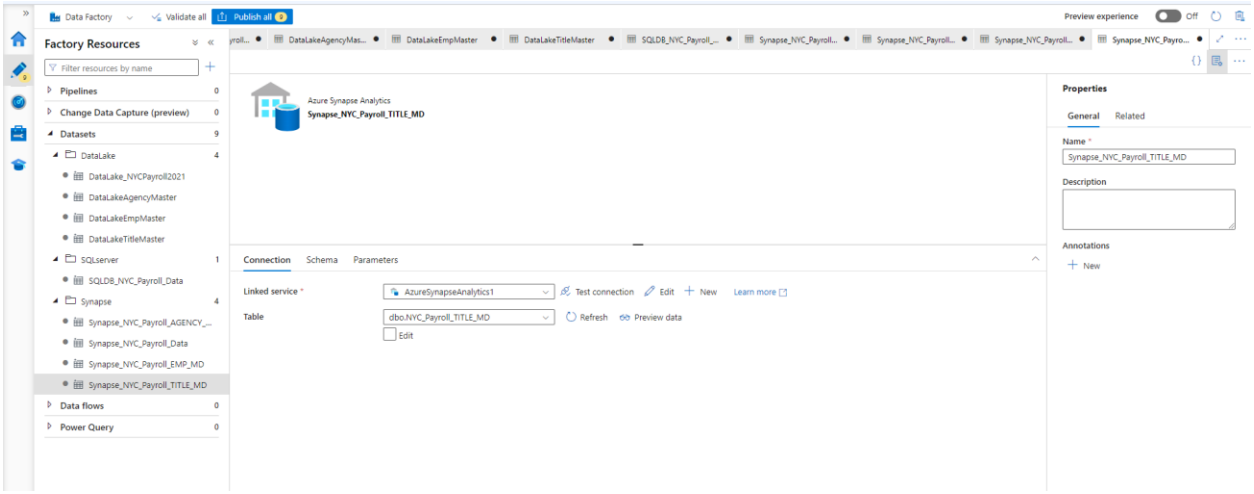


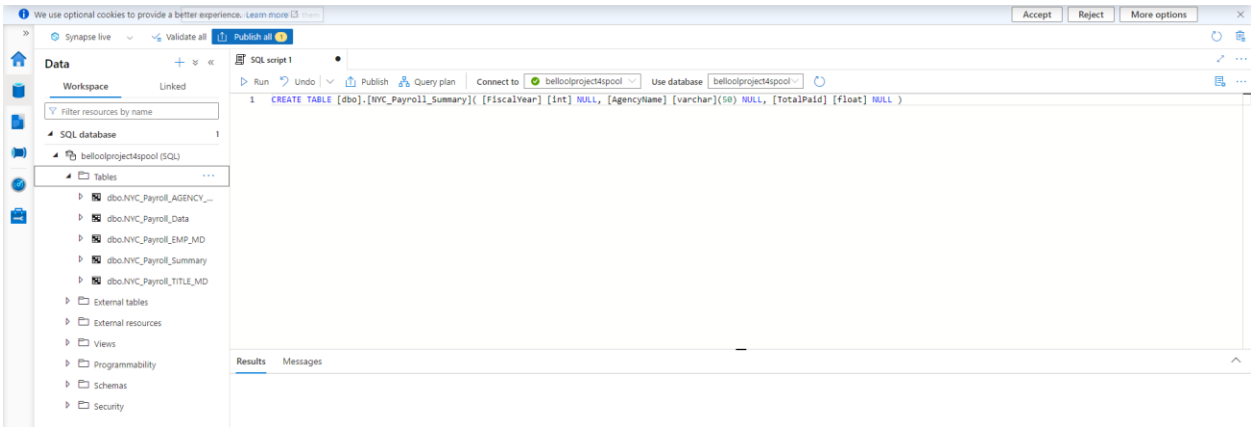


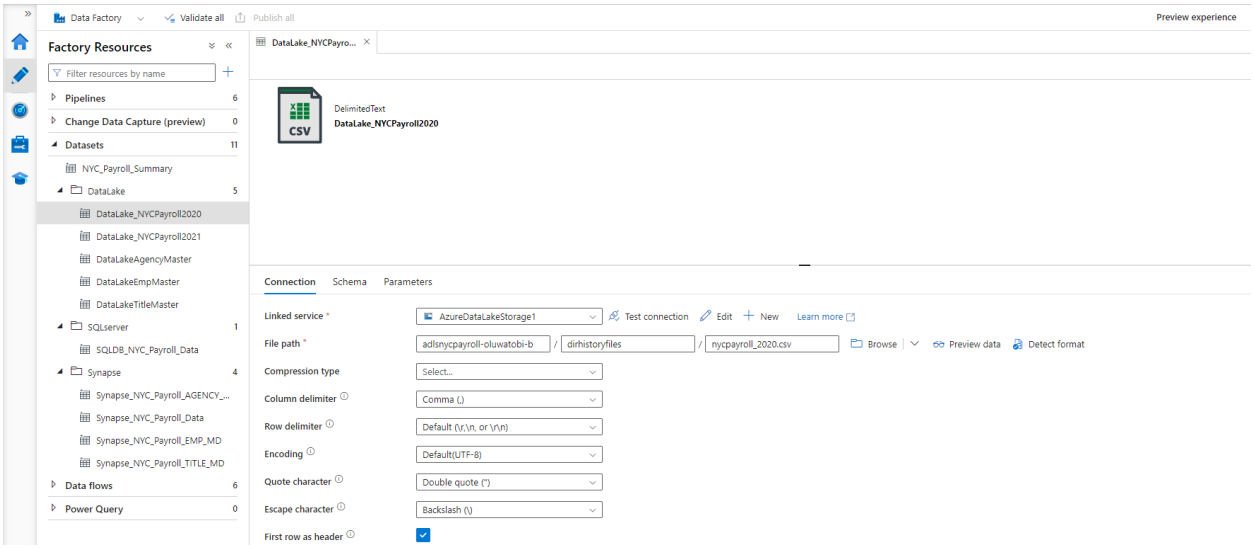
Step 2: Create Linked Services



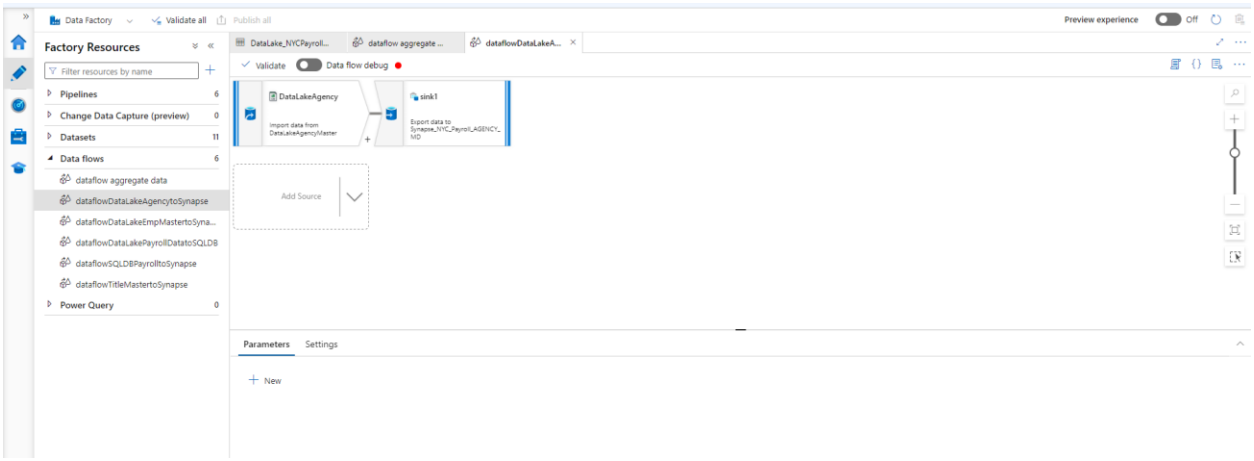
Step 3: Create Datasets in Azure Data Factory



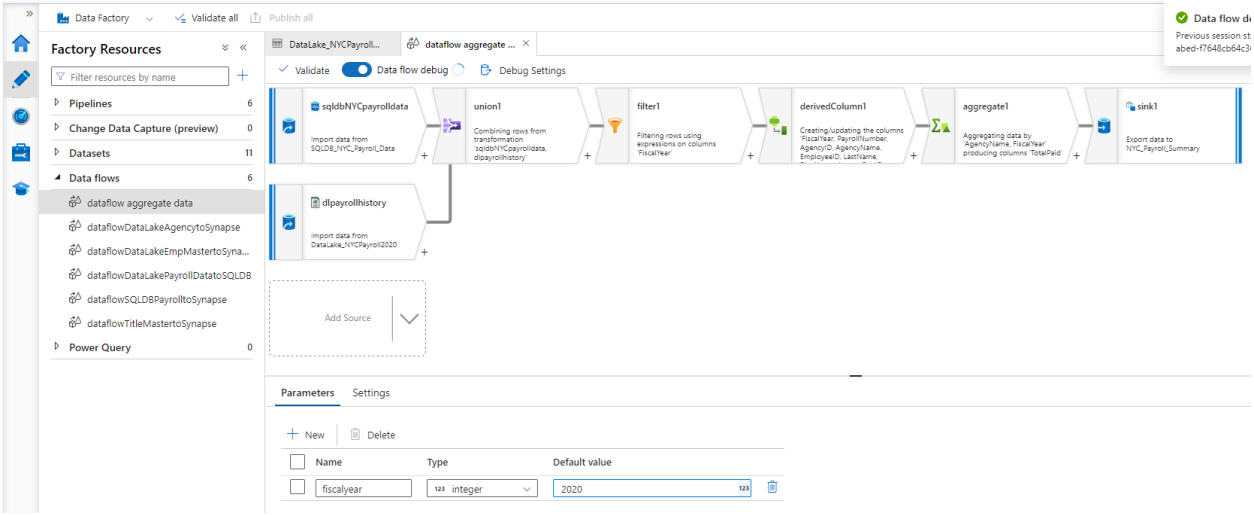


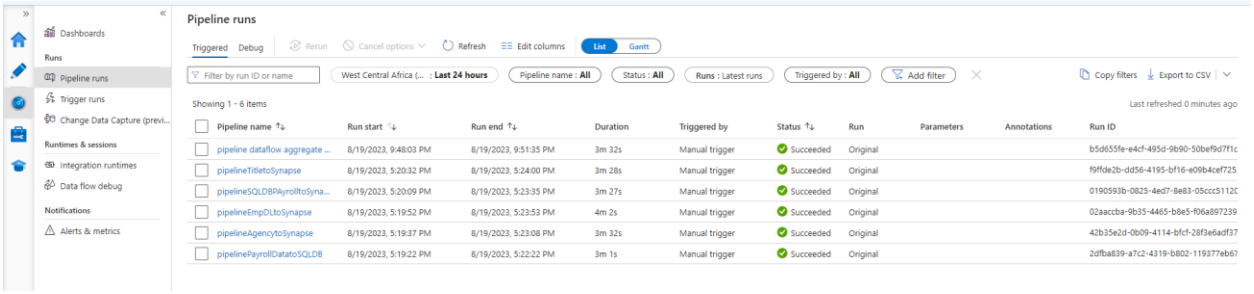


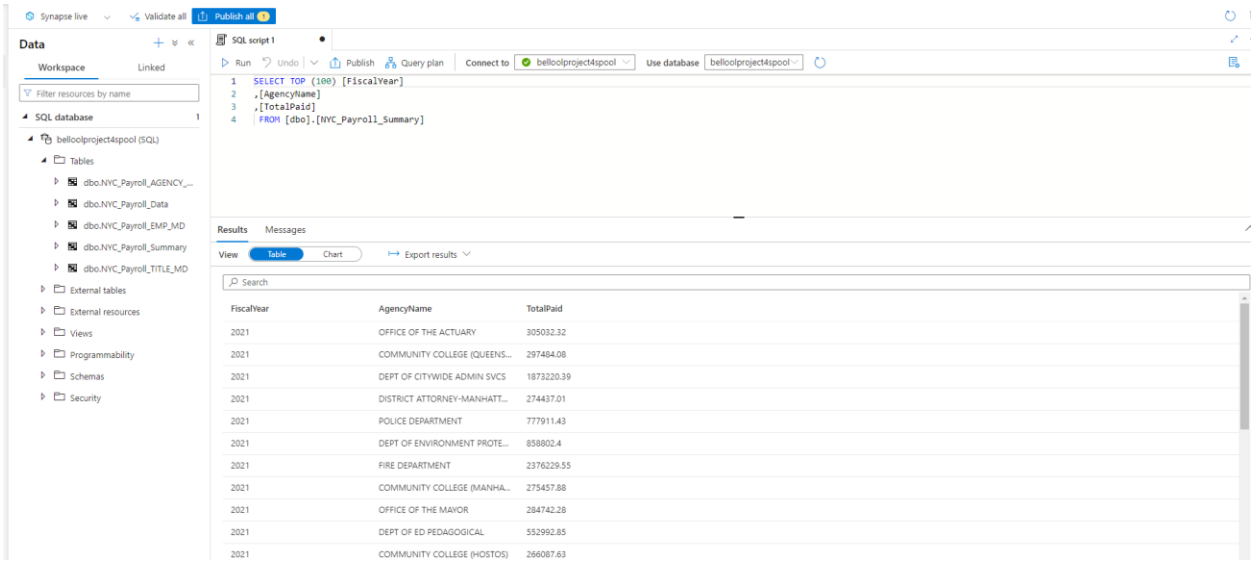
Step 4: Create Data Flows



Step 5: Data Aggregation and Parameterization


Data Factory
Validate all
Publish all
Factory Resources
Filter resources by name
Pipelines 6
Change Data Capture (preview) 0
Datasets 11
Data flows 6
dataflow aggregate data
dataflowDataLakeAgencytoSynapse
dataflowDataLakeEmpMastertoSyna...
dataflowDataLakePayrollDatatoSQLDB
dataflowSQLDBPayrolltoSynapse
dataflowTitleMastertoSynapse
Power Query 0
DataLake_NYCPayroll...
dataflow aggregate ...
Validate
Data flow debug
Debug Settings
sqldbNYCpayrolldata
Import data from SQLDB_NYC_Payroll_Data
union1
Combining rows from transformation sqldbNYCpayrolldata, dlpayrollhistory
filter1
Filtering rows using expressions on columns 'FiscalYear'
derivedColumn1
Creating/updating the columns 'FiscalYear, PayrollNumber, AgencyID, AgencyName, EmployeeID, LastName,
aggregate1
Aggregating data by 'AgencyName, FiscalYear' producing columns 'TotalPaid'
sink1
Export data to NYC_Payroll_Summary
dlpayrollhistory
Import data from DataLake_NYCPayroll2020
Add Source
Parameters
Settings
New
Delete
Name
Type
Default value
fiscalyear
integer
2020


Dashboards
Runs
Pipeline runs
Trigger runs
Change Data Capture (previ...
Runtimes & sessions
Integration runtimes
Data flow debug
Notifications
Alerts & metrics
Pipeline runs
Triggered
Debug
Rerun
Cancel options
Refresh
Edit columns
List
Gantt
Filter by run ID or name
West Central Africa (... : Last 24 hours
Pipeline name : All
Status : All
Runs : Latest runs
Triggered by : All
Add filter
Copy filters
Export to CSV
Showing 1 - 6 items
Last refreshed 0 minutes ago
Pipeline name
Run start
Run end
Duration
Triggered by
Status
Run
Parameters
Annotations
Run ID
pipeline dataflow aggregate ... 8/19/2023, 9:48:03 PM 8/19/2023, 9:51:35 PM 3m 32s Manual trigger Succeeded Original b5d655fe-e4cf-495d-9b90-50bef9d7f1c
pipelineTitletoSynapse 8/19/2023, 5:20:32 PM 8/19/2023, 5:24:00 PM 3m 28s Manual trigger Succeeded Original f9ffde2b-dd56-4195-bf16-e09b4cef725
pipelineSQLDBPayrolltoSyna... 8/19/2023, 5:20:09 PM 8/19/2023, 5:23:35 PM 3m 27s Manual trigger Succeeded Original 0190593b-0825-4ed7-8e83-05ccc5112
pipelineEmpDLtoSynapse 8/19/2023, 5:19:52 PM 8/19/2023, 5:23:53 PM 4m 2s Manual trigger Succeeded Original 02aaccba-9b35-4465-b8e5-f06a897239
pipelineAgencytoSynapse 8/19/2023, 5:19:37 PM 8/19/2023, 5:23:08 PM 3m 32s Manual trigger Succeeded Original 42b35e2d-0b09-4114-bfcf-28f3e6adf37
pipelinePayrollDatatoSQLDB 8/19/2023, 5:19:22 PM 8/19/2023, 5:22:22 PM 3m 1s Manual trigger Succeeded Original 2dfba839-a7c2-4319-b802-119377eb67


Synapse live
Validate all
Publish all
Data
Workspace
Linked
Filter resources by name
SQL database
belloolproject4spool (SQL)
Tables
dbo.NYC_Payroll_AGENCY_...
dbo.NYC_Payroll_Data
dbo.NYC_Payroll_EMP_MD
dbo.NYC_Payroll_Summary
dbo.NYC_Payroll_TITLE_MD
External tables
External resources
Views
Programmability
Schemas
Security
SQL script 1
Run
Undo
Publish
Query plan
Connect to
belloolproject4spool
Use database
belloolproject4spool
1 SELECT TOP (100) [FiscalYear]
2 ,[AgencyName]
3 ,[TotalPaid]
4 FROM [dbo].[NYC_Payroll_Summary]
Results
Messages
View
Table
Chart
Export results
Search
FiscalYear AgencyName TotalPaid
2021 OFFICE OF THE ACTUARY 305032.32
2021 COMMUNITY COLLEGE (QUEENS... 297484.08
2021 DEPT OF CITYWIDE ADMIN SVCS 1873220.39
2021 DISTRICT ATTORNEY-MANHATT... 274437.01
2021 POLICE DEPARTMENT 777911.43
2021 DEPT OF ENVIRONMENT PROTE... 858802.4
2021 FIRE DEPARTMENT 2376229.55
2021 COMMUNITY COLLEGE (MANHA... 275457.88
2021 OFFICE OF THE MAYOR 284742.28
2021 DEPT OF ED PEDAGOGICAL 552992.85
2021 COMMUNITY COLLEGE (HOSTOS) 266087.63