Data Warehousing

Data Warehouse Design

Final project

By
Kantapong Horaraung
Pasin Sukumalchan
Norapath Arjanurak
Tanyaton Oranrigsupak

Objective

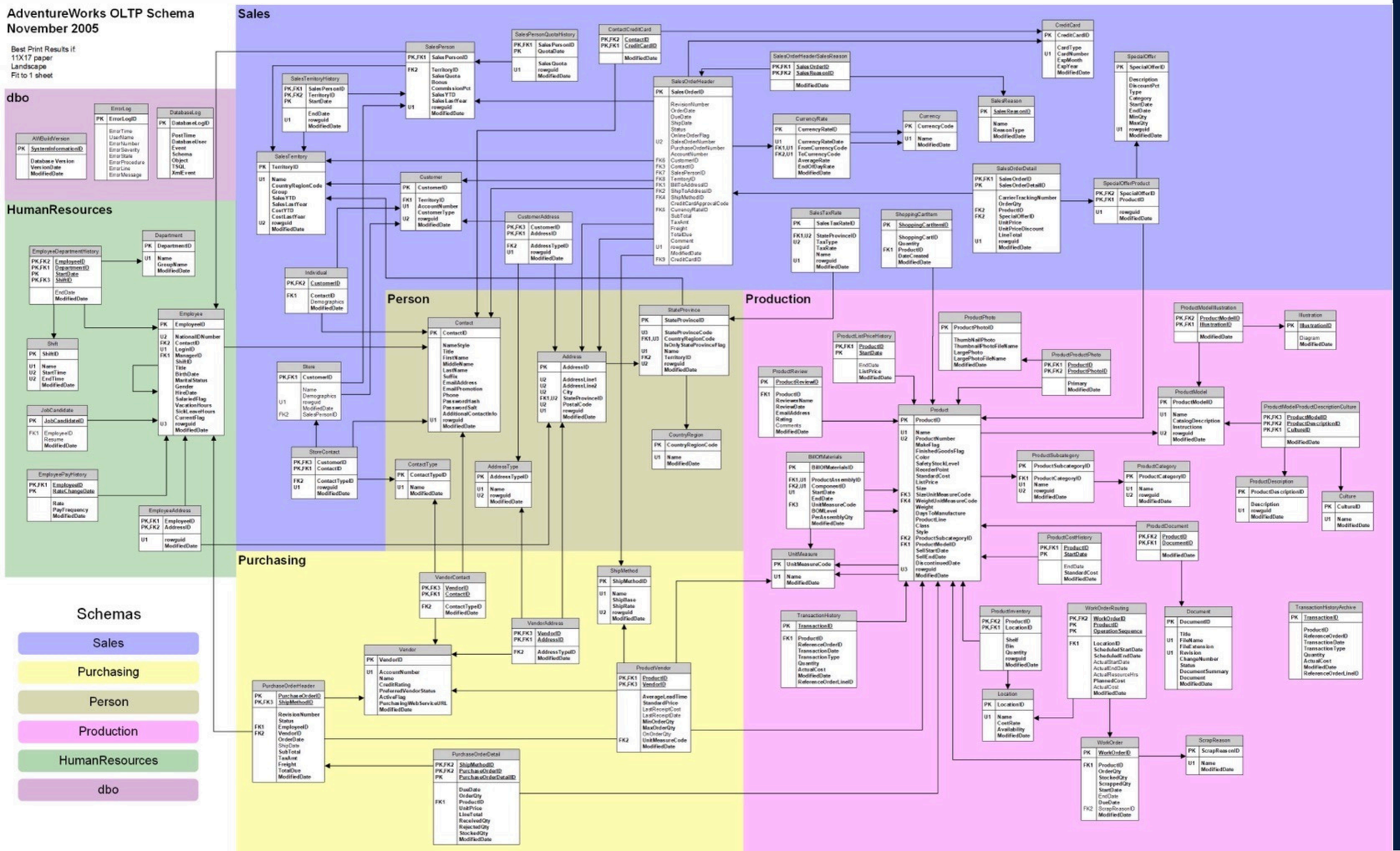
To design and implement a data warehouse based on the online transactional database to create a visualization and find business insight.

Tools & Technologies Used:

Microsoft SQL Server - SQL Server Management Studio (SSMS)
AWS Relational Database System (RDS)
AWS Glue

Source System – AdventureWorks

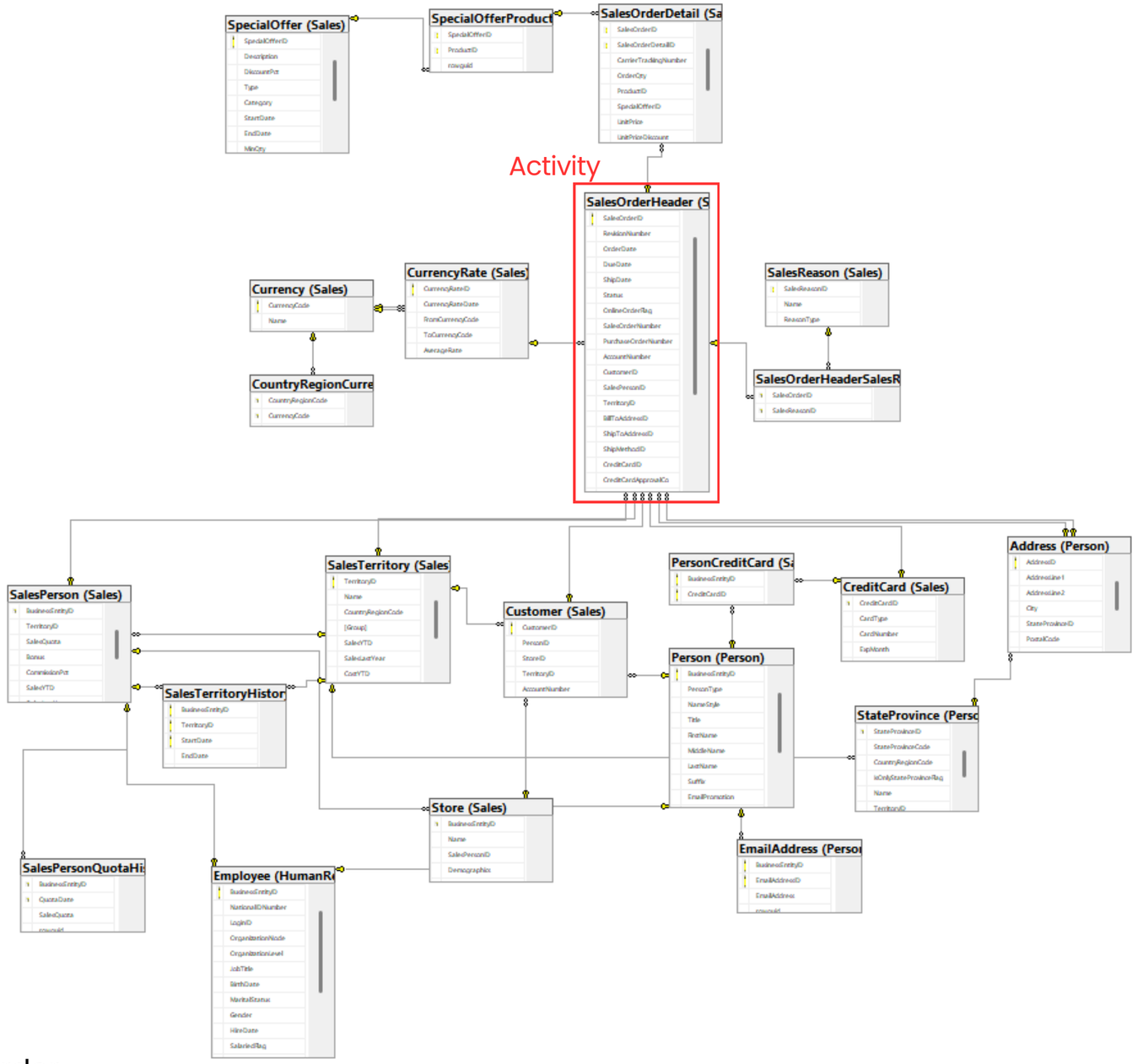
AdventureWorks is a sample database from Microsoft that simulates a fictional company selling bikes and cycling gear, designed to show how online transaction processing works.



Selecting Activities and Related Components -

Sales Order activity

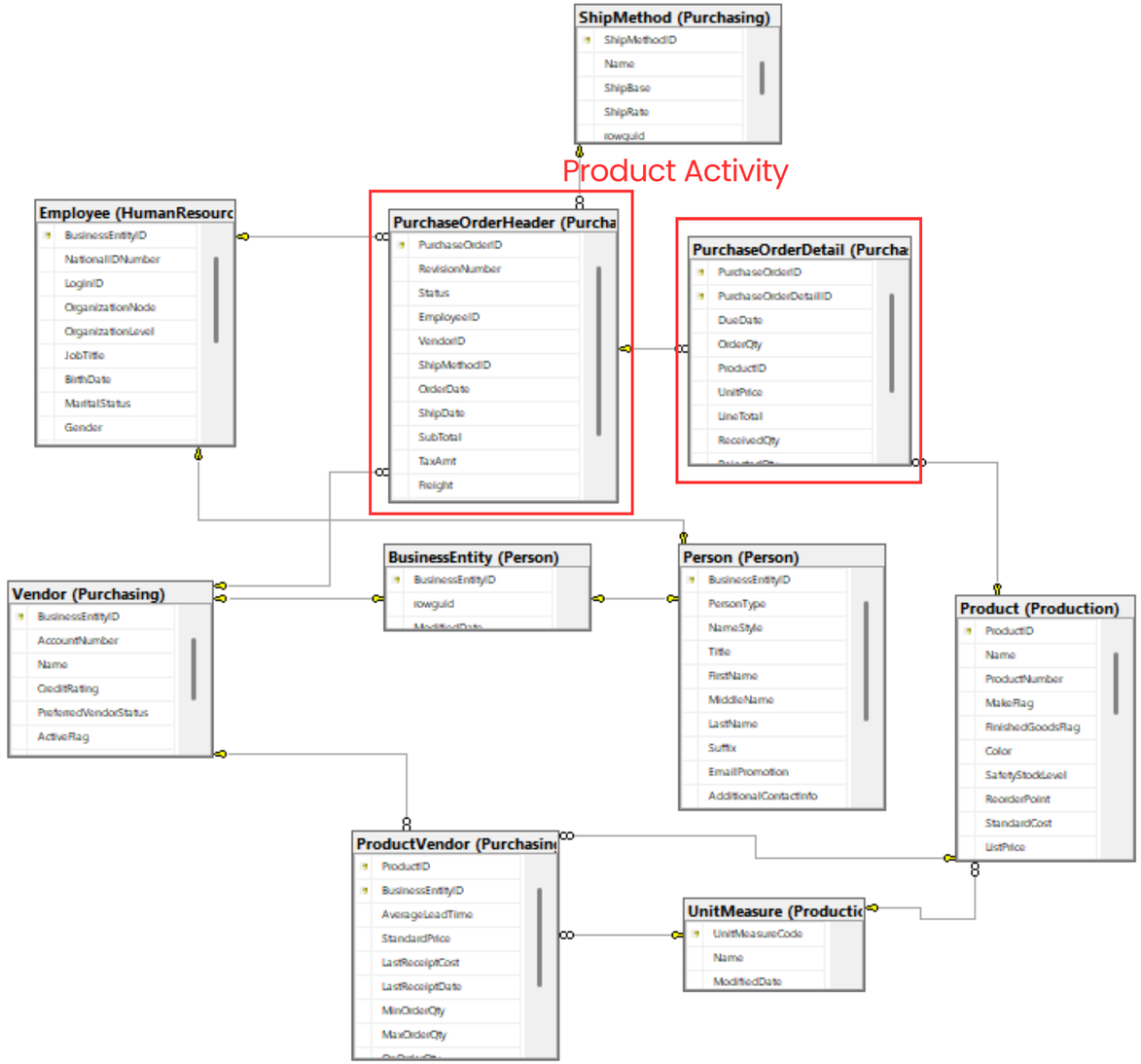
- **Step 1**: Sales Order Process
- **Step 2**:
  - Level 1: Sales per Order
  - Level 2: Sales per Product
- **Step 3:**
  - Date_dim
  - Customer_dim
  - Employee_dim, SalesEmployee_dim
  - CurrencyRate_dim
  - Creditcard_dim
  - Shipment_dim
  - Address_dim, Store_dim
  - **DD**: SalesOrderID
- **Step 4:**
  - SubTotal MONEY = SalesOrderHeader.SubTotal
  - TaxAmt MONEY = SalesOrderHeader.TaxAmt
  - Freight MONEY = SalesOrderHeader.Freight
  - TotalDue MONEY = SalesOrderHeader.TotalDue
  - Total_discount MONEY = OrderPromotion.Total_discount
  - Total_price MONEY = OrderPromotion.Total_price
  - NetPrice MONEY = OrderPromotion.NetPrice
  - Total_product_order INT = OrderPromotion.Total_product_order



Selecting Activities and Related Components -

Purchasing Product activity

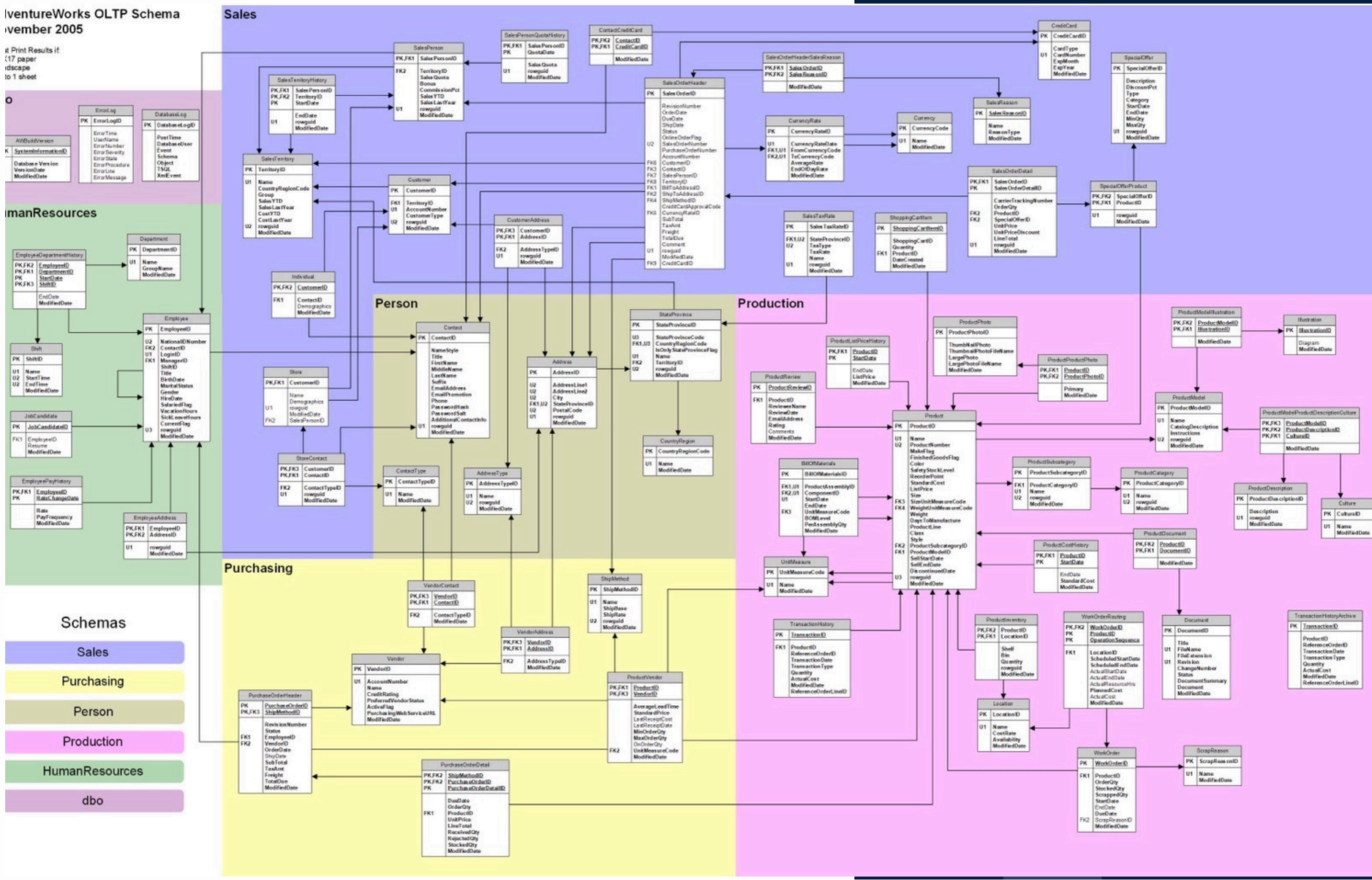
- **Step 1**: Purchasing Product Activity
- **Step 2**:
  - Level 1: Purchase per Order
  - Level 2: Purchase per Product
- **Step 3:**
  - Date_dim
  - Product_dim
  - Employee_dim
  - Shipment_dim
  - Vendor_dim
  - **DD**: PurchaseOrderID, PurchaseOrderDetailID
- **Step 4:**
  - OrderQty SMALLINT = PurchaseOrderDetail.OrderQty
  - UnitPrice MONEY = PurchaseOrderDetail.UnitPrice
  - LineTotal MONEY = PurchaseOrderDetail.LineTotal
  - ReceivedQty DECIMAL(8, 2) = PurchaseOrderDetail.ReceivedQty
  - RejectedQty DECIMAL(8, 2) = PurchaseOrderDetail.RejectedQty
  - StockedQty DECIMAL(8, 2) = (ReceivedQty - RejectedQty)
  - SubTotal MONEY = PurchaseOrderHeader.SubTotal
  - TaxAmt MONEY = PurchaseOrderHeader.TaxAmt
  - Freight MONEY = PurchaseOrderHeader.Freight
  - TotalDue MONEY = PurchaseOrderHeader.TotalDue



List of all Selected Sources

Sources:

- Sales.SalesOrderDetail
- Sales.SalesOrderHeader
- Sales.Customer
- Person.Person
- HumanResources.Employee
- Sales.SalesPerson
- Sales.CurrencyRate
- Sales.CreditCard
- Purchasing.ShipMethod
- Person.Address
- Person.StateProvince
- Production.Product
- Purchasing.Vendor
- Sales.SalesOrderDetail



Data Preparation and Data Cleaning 1

Removing ‘Sales.Customer’ rows

- Some rows can contain PersonID = null or StoreID = null or Both

	CustomerID	PersonID	StoreID	TerritoryID	AccountNumber	rowguid	ModifiedDate
698	698	*NULL*	640	1	AW00000698	6bc4ff5a-0696-4c2a-88f9-08d1dba91b74	2014-09-12 11:15:07.263
699	699	*NULL*	842	5	AW00000699	38a528aa-0402-4ed9-abbb-983bceb37d40	2014-09-12 11:15:07.263
700	700	*NULL*	1030	5	AW00000700	540fb57f-f81a-4794-8e9a-c071c4f0bc18	2014-09-12 11:15:07.263
701	701	*NULL*	844	6	AW00000701	61ae9625-8bd8-48b7-b171-8a90cbba494c	2014-09-12 11:15:07.263
702	11000	13531	*NULL*	9	AW00011000	477586b3-2977-4e54-b1a8-569ab2c7c4d4	2014-09-12 11:15:07.263
703	11001	5454	*NULL*	9	AW00011001	c32a8084-9077-4f13-9738-1e2da7c1dcd9	2014-09-12 11:15:07.263
704	11002	11269	*NULL*	9	AW00011002	45715dd8-2f57-4a39-beb4-6a8f99d59794	2014-09-12 11:15:07.263
705	11003	11358	*NULL*	9	AW00011003	7e240efc-7ee6-4814-93a8-269821157e18	2014-09-12 11:15:07.263

- Remove the rows(From 19,185 rows to 635 rows)

	CustomerID	PersonID	StoreID	TerritoryID	AccountNumber	rowguid	ModifiedDate
1	29484	291	292	5	AW00029484	a24c1240-1a8d-427b-91e9-6197f8cf3cad	2014-09-12 11:15:07.263
2	29485	293	294	4	AW00029485	392ae773-d7ec-48ac-b8d0-6e65b770285c	2014-09-12 11:15:07.263
3	29486	295	296	3	AW00029486	2a5aba2e-8db0-4856-a773-21d185f1679e	2014-09-12 11:15:07.263
4	29487	297	298	2	AW00029487	676811a9-9e7f-49af-baac-f6b0b053246e	2014-09-12 11:15:07.263
5	29488	299	300	9	AW00029488	eeee9a74-eaee-40e7-b4f7-b45e87ccb5a6	2014-09-12 11:15:07.263
6	29489	301	302	4	AW00029489	c353fe38-6147-40a3-944d-3736f6297b8c	2014-09-12 11:15:07.263
7	29490	303	304	1	AW00029490	7f2d6183-6aee-4ad1-973d-f45a19b70bf7	2014-09-12 11:15:07.263
8	29491	305	306	5	AW00029491	62a7b13e-8ac5-4612-b97c-b16a90856bdd	2014-09-12 11:15:07.263

Data Preparation and Data Cleaning 2

Filtering ‘Person.Person’

- Contain a lot of unused rows in the person.person

	RowsInPerson
1	19972

- Filter on the row that contains "Employee” and “Customer” (2000 rows)

	BusinessEntityID	PersonType	NameStyle	Title	FirstName	MiddleName	LastName	Suffix	EmailPr
1	1	EM	0	NULL	Ken	J	Sánchez	NULL	0
2	2	EM	0	NULL	Terri	Lee	Duffy	NULL	1
3	3	EM	0	NULL	Roberto	NULL	Tamburello	NULL	0
4	4	EM	0	NULL	Rob	NULL	Walters	NULL	0
5	5	EM	0	Ms.	Gail	A	Erickson	NULL	0
6	6	EM	0	Mr.	Jossef	H	Goldberg	NULL	0
7	7	EM	0	NULL	Dylan	A	Miller	NULL	2
8	8	EM	0	NULL	Diane	L	Margheim	NULL	0
9	9	EM	0	NULL	Gigi	N	Matthew	NULL	0
10	10	EM	0	NULL	Michael	NULL	Raheem	NULL	2
11	11	EM	0	NULL	Ovidiu	V	Cracium	NULL	0
12	12	EM	0	NULL	Thierry	B	D'Hers	NULL	2
13	13	EM	0	Ms.	Janice	M	Galvin	NULL	2
14	14	EM	0	NULL	Michael	I	Sullivan	NULL	2
15	15	EM	0	NULL	Sharon	B	Salavaria	NULL	2
16	16	EM	0	NULL	David	M	Bradley	NULL	1
17	17	EM	0	NULL	Kevin	F	Brown	NULL	2
18	18	EM	0	NULL	John	L	Wood	NULL	2
19	19	EM	0	NULL	Mary	A	Dempsey	NULL	1

Data Preparation and Data Cleaning 3

Removing ‘Sales.Person’ rows

- Remove the row with SaleQuota == null

	BusinessEntityID	TerritoryID	SalesQuota	Bonus	CommissionPct	SalesYTD	SalesLastYear	rowguid	ModifiedDat
1	274	*NULL*	*NULL*	0.00	0.00	559697.5639	0.00	48754992-9ee0-4c0e-8c94-9451604e3e02	2010-12-28
2	285	*NULL*	*NULL*	0.00	0.00	172524.4512	0.00	cfdbef27-b1f7-4a56-a878-0221c73bae67	2013-03-07
3	287	*NULL*	*NULL*	0.00	0.00	519905.932	0.00	1dd1f689-df74-4149-8600-59555eef154b	2012-04-09

- Remove three rows of SaleQuota == null

Results Messages

	BusinessEntityID	TerritoryID	SalesQuota	Bonus	CommissionPct	SalesYTD	SalesLastYear	rowguid	ModifiedDate
1	275	2	300000.00	4100.00	0.012	3763178.1787	1750406.4785	1e0a7274-3064-4f58-88ee-4c6586c87169	2011-05-24 0
2	276	4	250000.00	2000.00	0.015	4251368.5497	1439156.0291	4dd9eee4-8e81-4f8c-af97-683394c1f7c0	2011-05-24 0
3	277	3	250000.00	2500.00	0.015	3189418.3662	1997186.2037	39012928-bfec-4242-874d-423162c3f567	2011-05-24 0
4	278	6	250000.00	500.00	0.01	1453719.4653	1620276.8966	7a0ae1ab-b283-40f9-91d1-167abf06d720	2011-05-24 0
5	279	5	300000.00	6700.00	0.01	2315185.611	1849640.9418	52a5179d-3239-4157-ae29-17e868296dc0	2011-05-24 0
6	280	1	250000.00	5000.00	0.01	1352577.1325	1927059.178	be941a4a-fb50-4947-bda4-bb8972365b08	2011-05-24 0
7	281	4	250000.00	3550.00	0.01	2458535.6169	2073505.9999	35326ddb-7278-4fef-b3ba-ea137b69094e	2011-05-24 0
8	282	6	250000.00	5000.00	0.015	2604540.7172	2038234.6549	31fd7fc1-dc84-4f05-b9a0-762519eacacc	2011-05-24 0

Data Preparation and Data Cleaning 4

Perform aggregate on ‘SalesOrderDetail’

- One SalesOrderID can have many SalesOrderDetails; Thus, we perform aggregate and create OrderPromotionDim

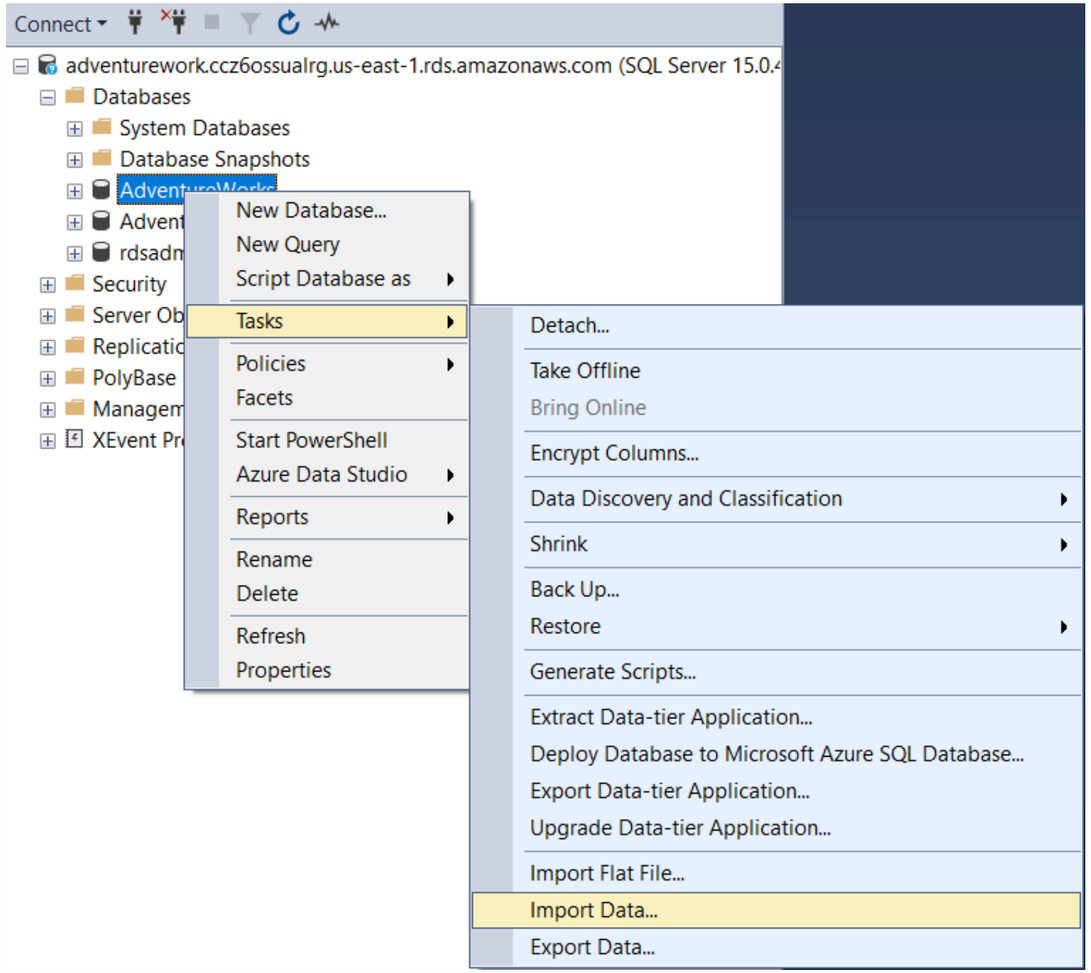
	SalesOrderID	SalesOrderDetailID	OrderQty	UnitPrice	UnitPriceDiscount	LineTotal
1	43659	1	1	2024.994	0.00	2024.994000
2	43659	2	3	2024.994	0.00	6074.982000
3	43659	3	1	2024.994	0.00	2024.994000
4	43659	4	1	2039.994	0.00	2039.994000
5	43659	5	1	2039.994	0.00	2039.994000
6	43659	6	2	2039.994	0.00	4079.988000
7	43659	7	1	2039.994	0.00	2039.994000
8	43659	8	3	28.8404	0.00	86.521200
9	43659	9	1	28.8404	0.00	28.840400
10	43659	10	6	5.70	0.00	34.200000
11	43659	11	2	5.1865	0.00	10.373000
12	43659	12	4	20.1865	0.00	80.746000

- The aggregate tables

	SalesOrderID	Total_discount	Total_price	NetPrice	Total_product_order	ModifiedDate
1	43659	0.000000	20565.6206	20565.620600	26	2025-05-06
2	43660	0.000000	1294.2529	1294.252900	2	2025-05-06
3	43661	0.000000	32726.4786	32726.478600	38	2025-05-06
4	43662	0.000000	28832.5289	28832.528900	54	2025-05-06

Load Back to Source database

- Directly load the clear table to the source database



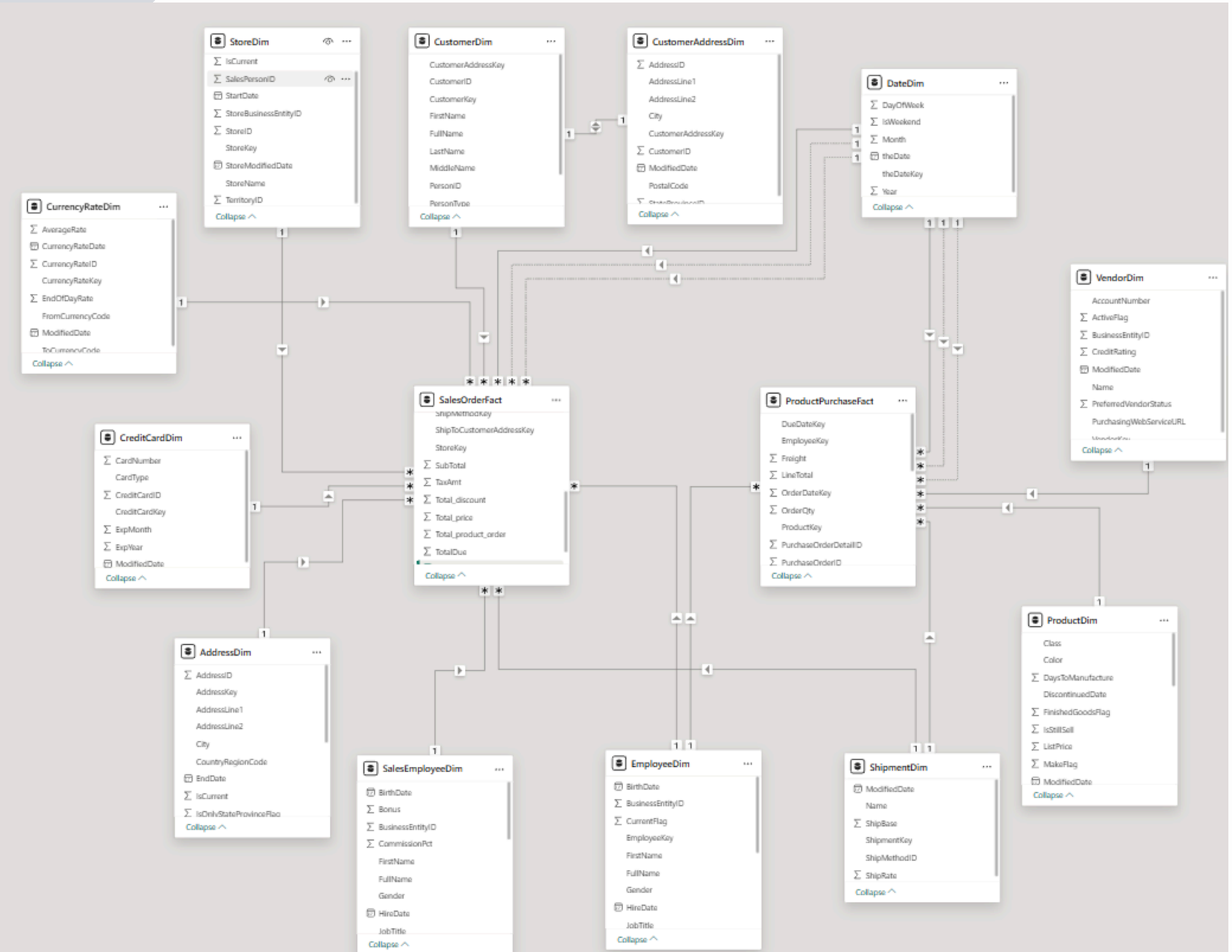
*Note: We try S3 before but it consume a lot of resources

SCD Reasons of Dimension

Dimension Name	SCD Type	Reason
Date_dim	0	Predefined and No change
Customer_Dim/ CustomerAddress_dim	2 and Outrigger	Track change and Saved space in long-run
Address_dim	2	Track change and Collect important data
Employee_dim/ SalesEmployee_dim	2 and 1 (SCD 7)	Track the current sales and Track the change
CurrencyRate_dim	1	Daily rate changes and can be consume space in lon
Credit_card_dim	1	Tracks expiry/date info changes
Shipment_dim	1	Shipping info changes slowly and Not track change
Store_dim	2	Store name and type changes tracked
Product_dim	2	Tracks product detail changes
Vendor_dim	1	Tracks vendor info and Not track change

Planning A Schema

Entity-relationship diagran from the Power Bi

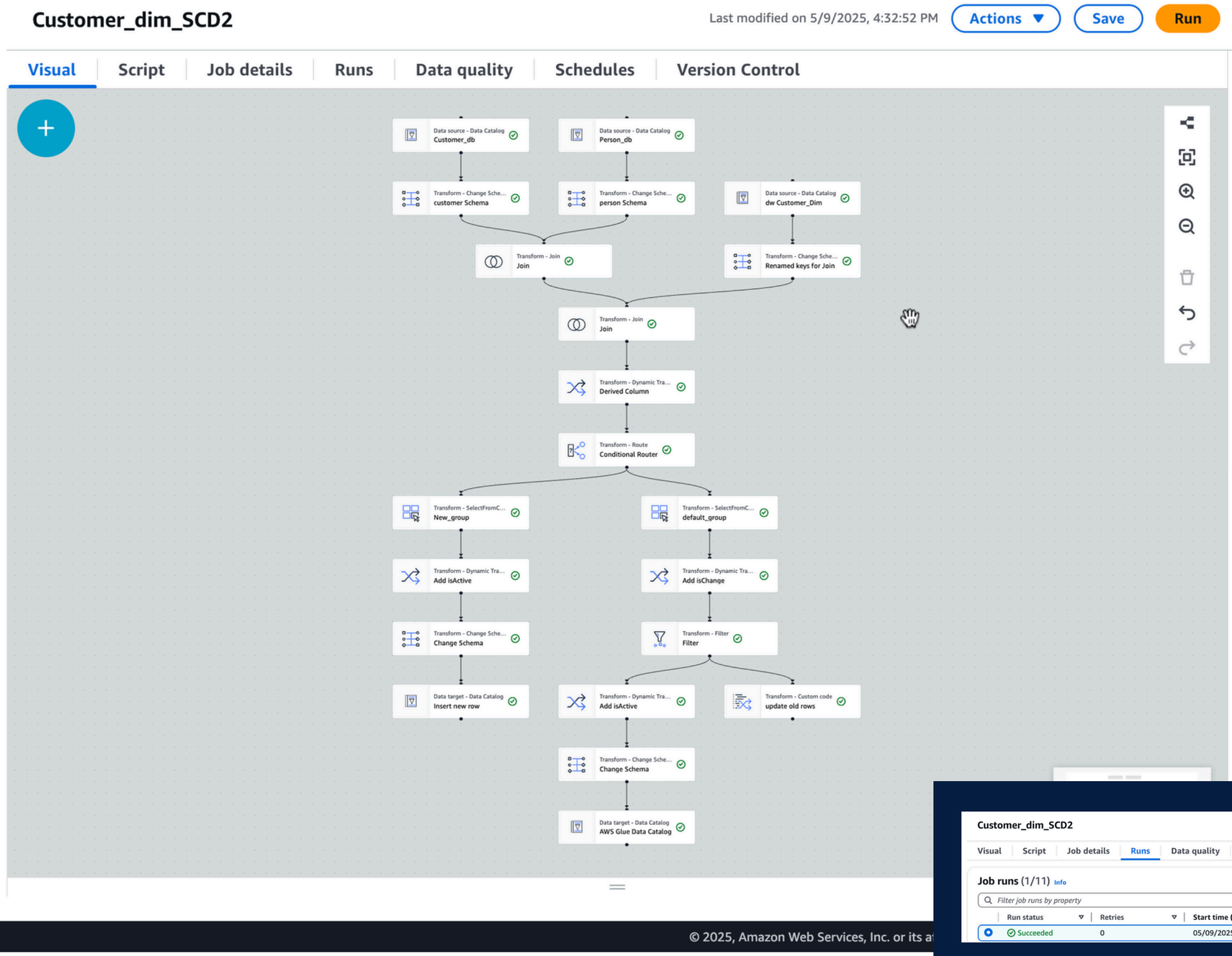


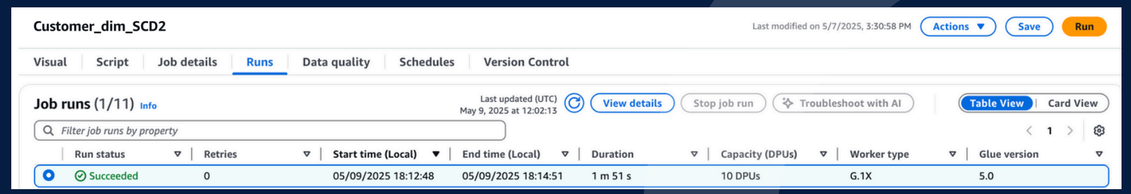
Conformed Bus Matrix

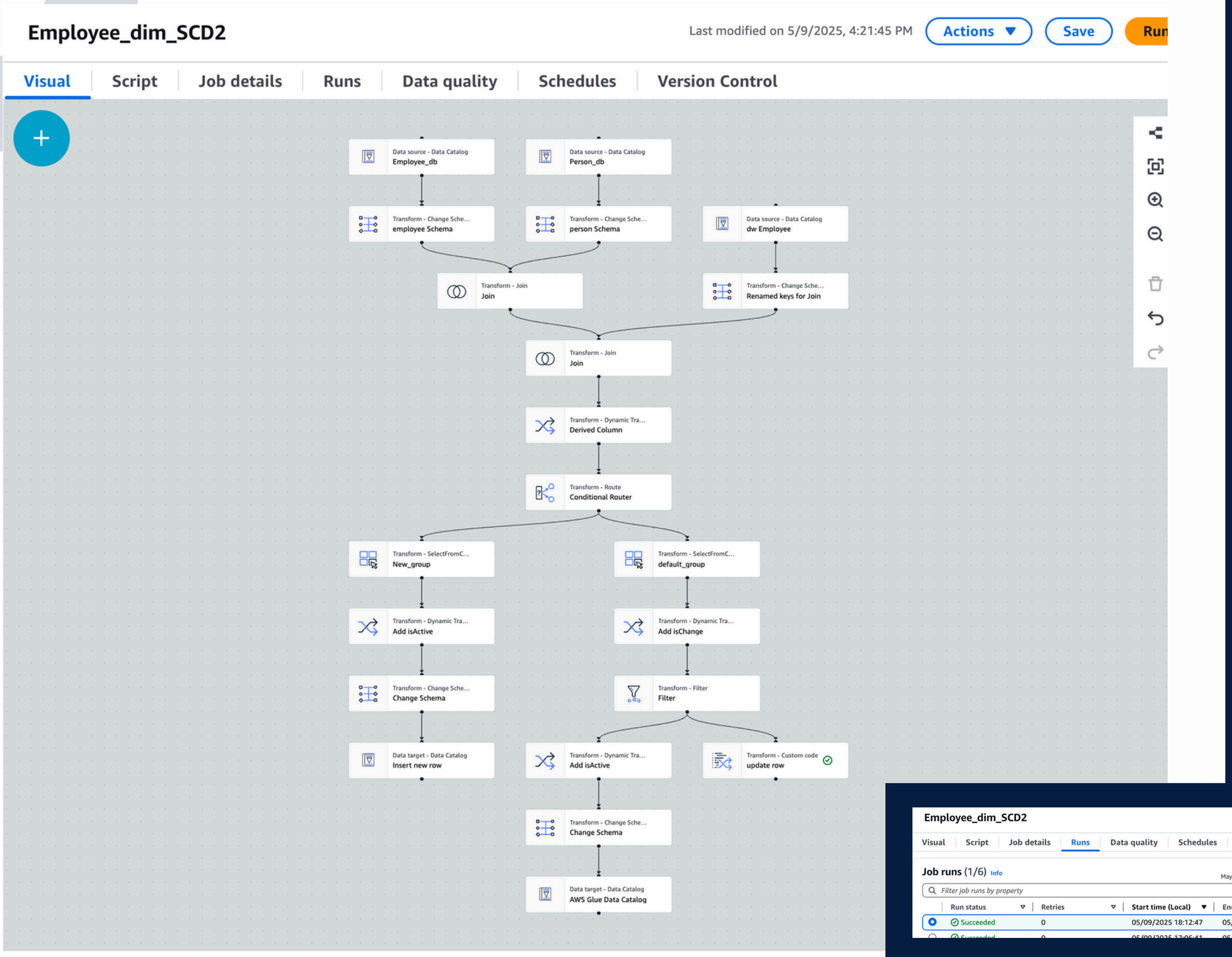
Dimension Name	FactSalesOrder	FactPurchasingProduct
Date_Dim	✅	✅
Customer_Dim	✅	
Sales_Employee_Dim	✅	
Employee_Dim	✅	✅
CustomerAddress_Dim	✅	
Shipment_Dim	✅	✅
CreditCard_Dim	✅	
CurrencyRate_Dim	✅	
Store_Dim	✅	
Product_Dim		✅
Vendor_Dim		✅

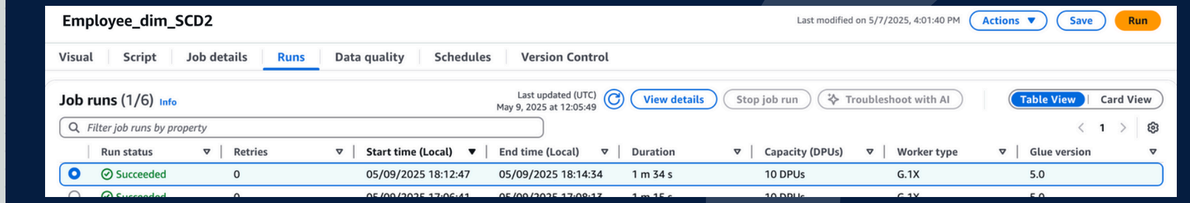
Start ETL process!

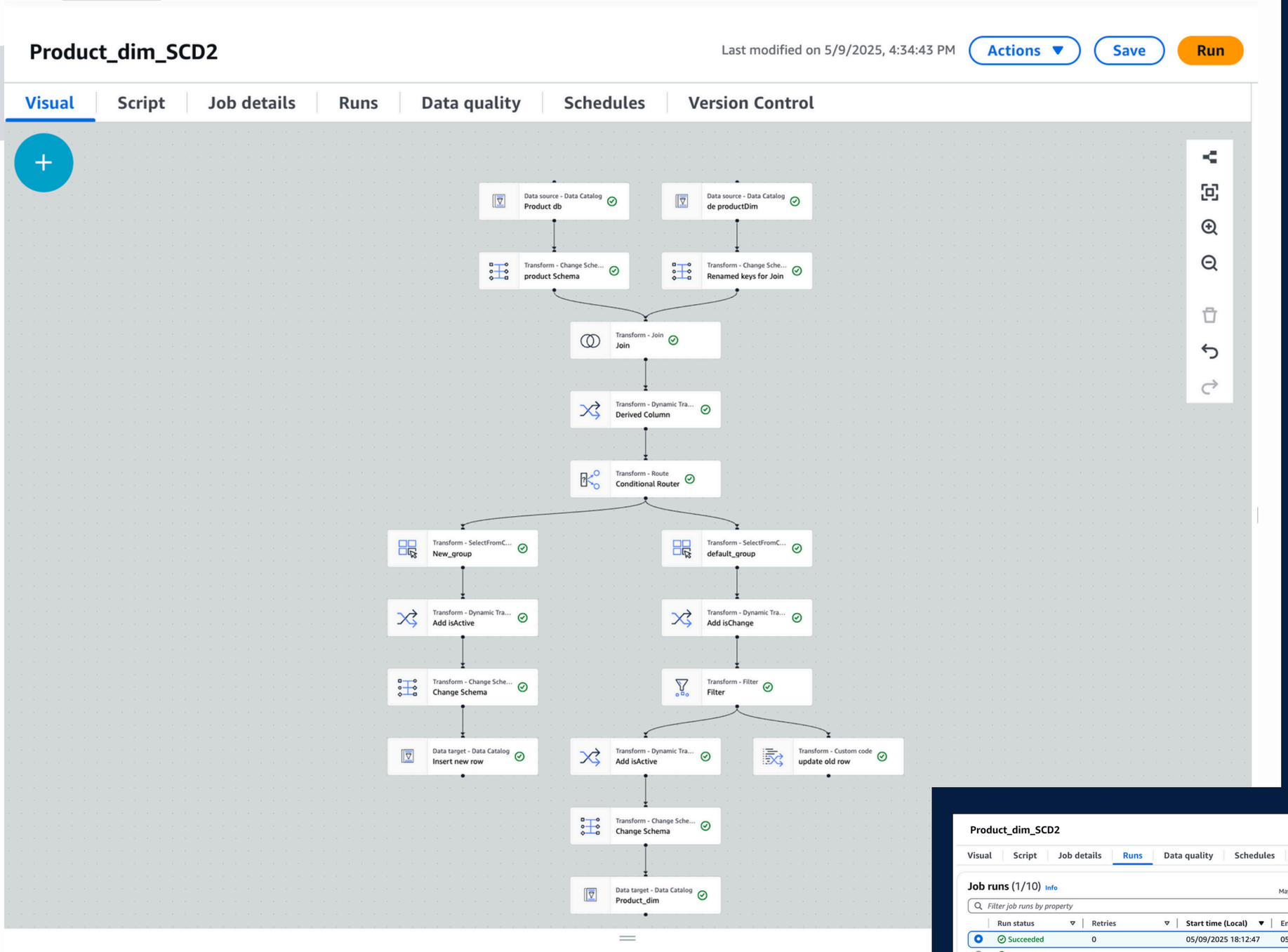
Using AWS Glue

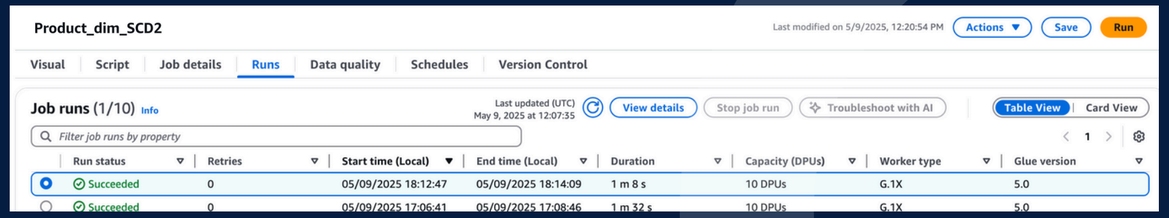


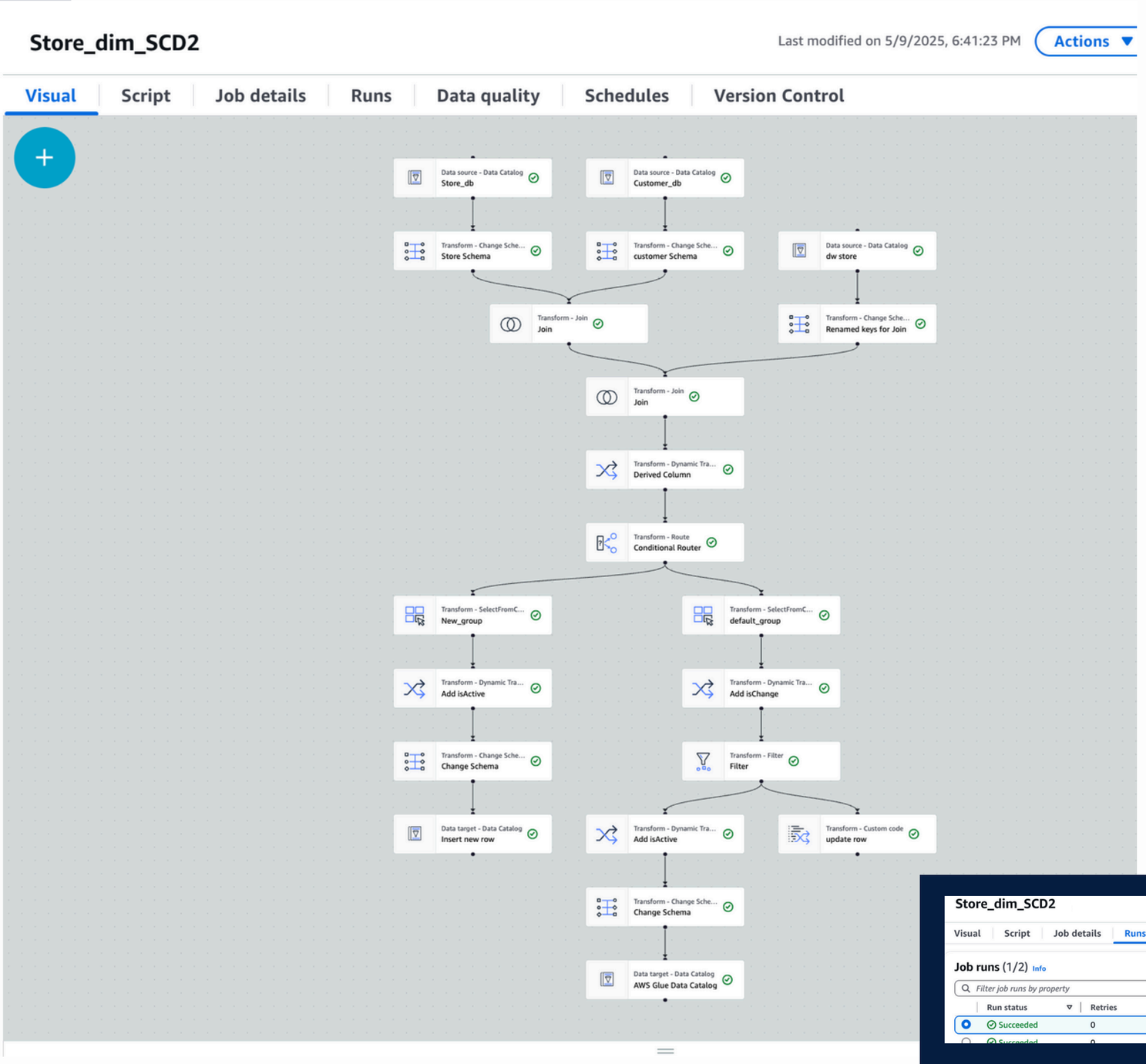


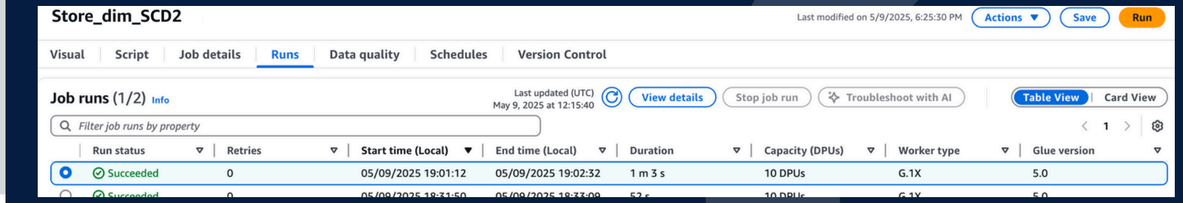

Employee_dim_SCD2
Last modified on 5/9/2025, 4:21:45 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Data source - Data Catalog
Employee_db
Data source - Data Catalog
Person_db
Transform - Change Sche...
employee Schema
Transform - Change Sche...
person Schema
Data source - Data Catalog
dw Employee
Transform - Join
Join
Transform - Change Sche...
Renamed keys for Join
Transform - Join
Join
Transform - Dynamic Tra...
Derived Column
Transform - Route
Conditional Router
Transform - SelectFromC...
New_group
Transform - SelectFromC...
default_group
Transform - Dynamic Tra...
Add isActive
Transform - Dynamic Tra...
Add isChange
Transform - Change Sche...
Change Schema
Transform - Filter
Filter
Data target - Data Catalog
Insert new row
Transform - Dynamic Tra...
Add isActive
Transform - Custom code
update row
Transform - Change Sche...
Change Schema
Data target - Data Catalog
AWS Glue Data Catalog

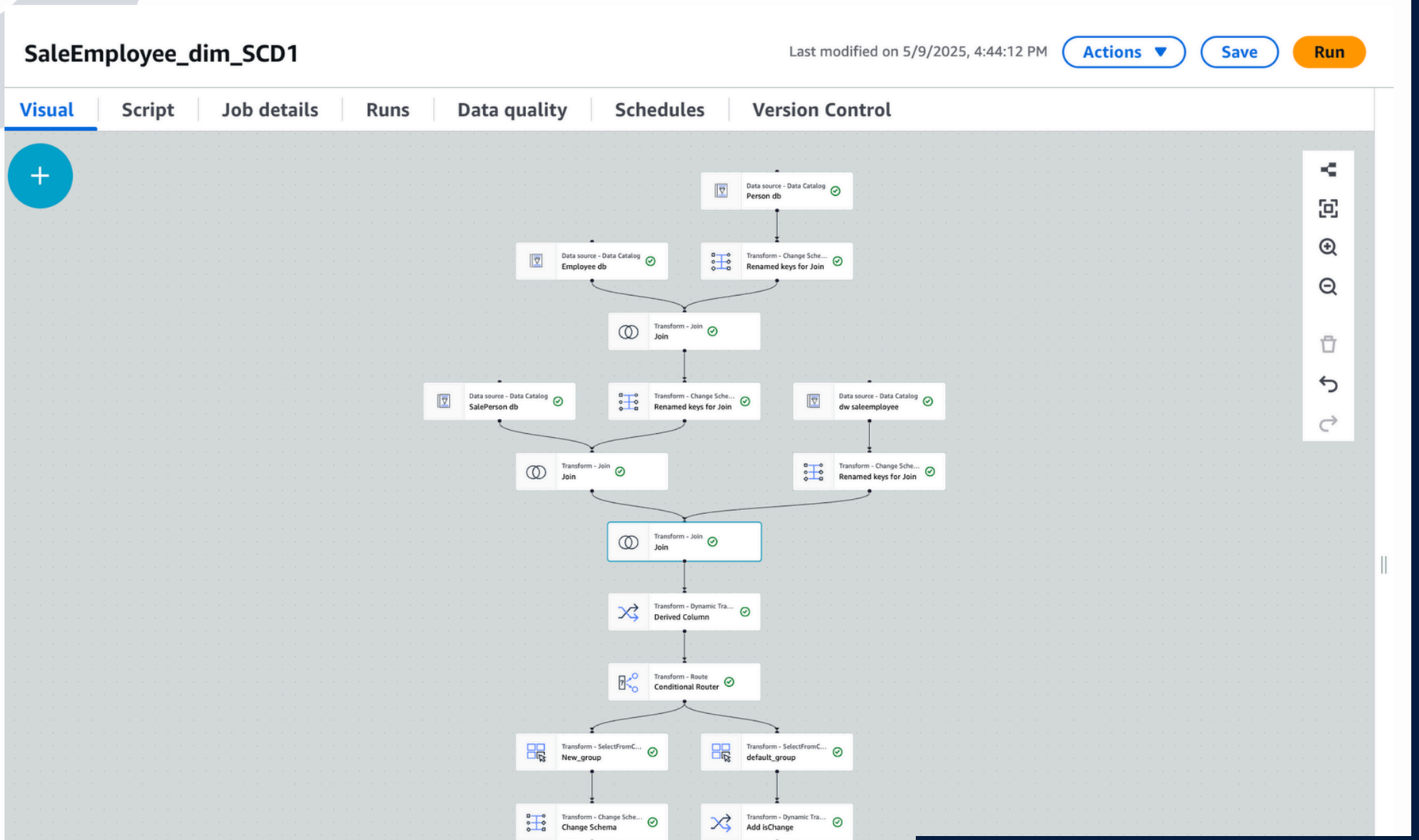

Employee_dim_SCD2
Last modified on 5/7/2025, 4:01:40 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Job runs (1/6) Info
Last updated (UTC)
May 9, 2025 at 12:05:49
View details
Stop job run
Troubleshoot with AI
Table View
Card View
Filter job runs by property
Run status
Retries
Start time (Local)
End time (Local)
Duration
Capacity (DPUs)
Worker type
Glue version
Succeeded
0
05/09/2025 18:12:47
05/09/2025 18:14:34
1 m 34 s
10 DPUs
G.1X
5.0

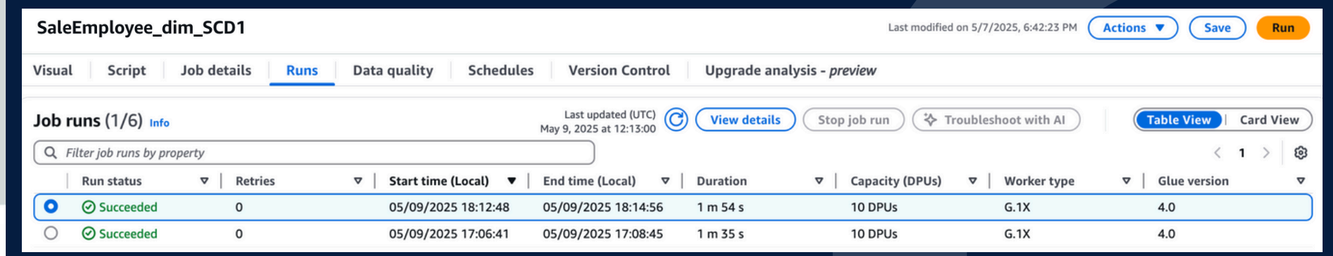

Product_dim_SCD2
Last modified on 5/9/2025, 4:34:43 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Data source - Data Catalog
Product db
Data source - Data Catalog
de productDim
Transform - Change Sche...
product Schema
Transform - Change Sche...
Renamed keys for Join
Transform - Join
Join
Transform - Dynamic Tra...
Derived Column
Transform - Route
Conditional Router
Transform - SelectFromC...
New_group
Transform - SelectFromC...
default_group
Transform - Dynamic Tra...
Add isActive
Transform - Dynamic Tra...
Add isChange
Transform - Change Sche...
Change Schema
Transform - Filter
Filter
Data target - Data Catalog
Insert new row
Transform - Dynamic Tra...
Add isActive
Transform - Custom code
update old row
Transform - Change Sche...
Change Schema
Data target - Data Catalog
Product_dim

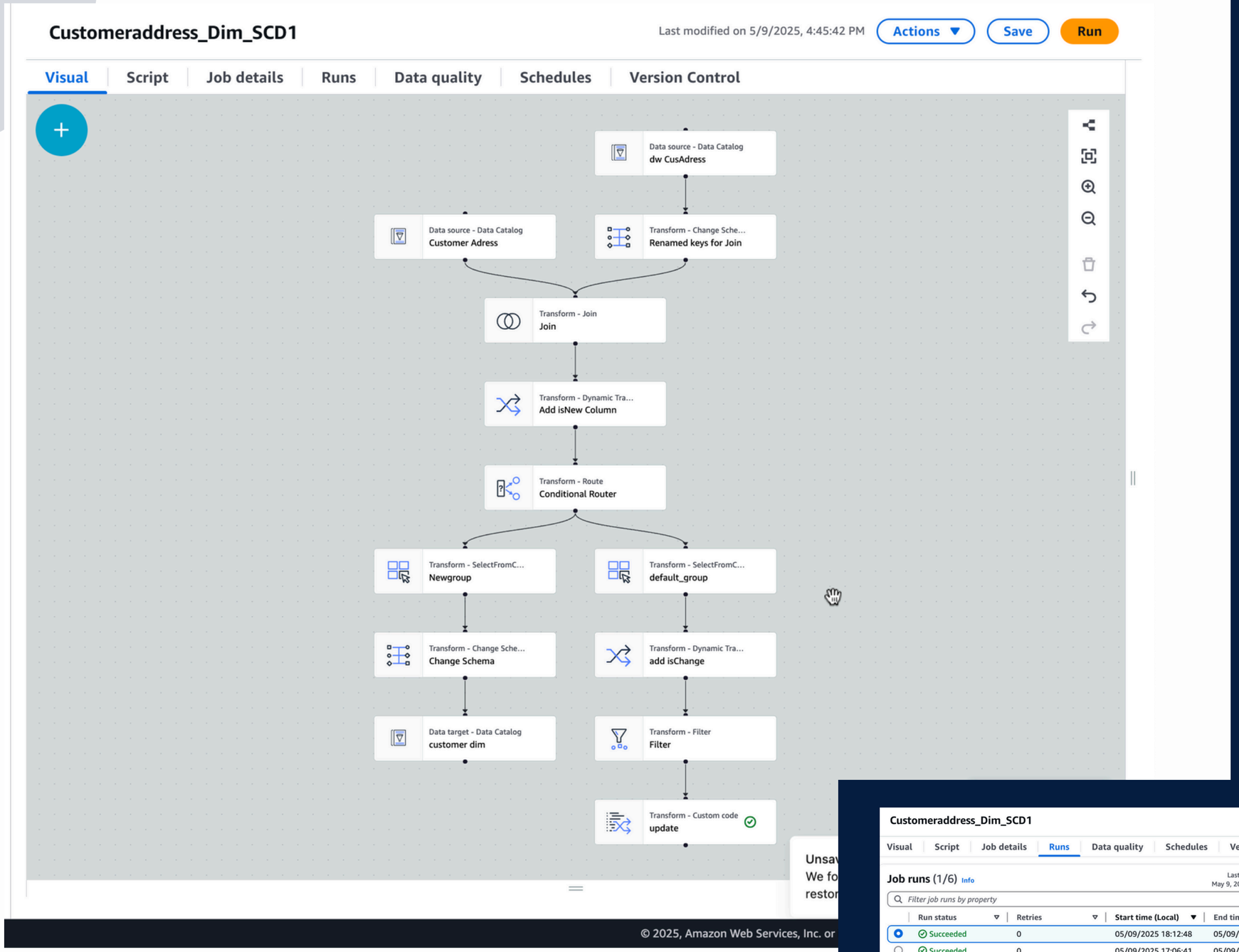

Product_dim_SCD2
Last modified on 5/9/2025, 12:20:54 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Job runs (1/10) Info
Last updated (UTC)
May 9, 2025 at 12:07:35
View details
Stop job run
Troubleshoot with AI
Table View
Card View
Filter job runs by property
Run status
Retries
Start time (Local)
End time (Local)
Duration
Capacity (DPUs)
Worker type
Glue version
Succeeded
0
05/09/2025 18:12:47
05/09/2025 18:14:09
1 m 8 s
10 DPUs
G.1X
5.0
Succeeded
0
05/09/2025 17:06:41
05/09/2025 17:08:46
1 m 32 s
10 DPUs
G.1X
5.0

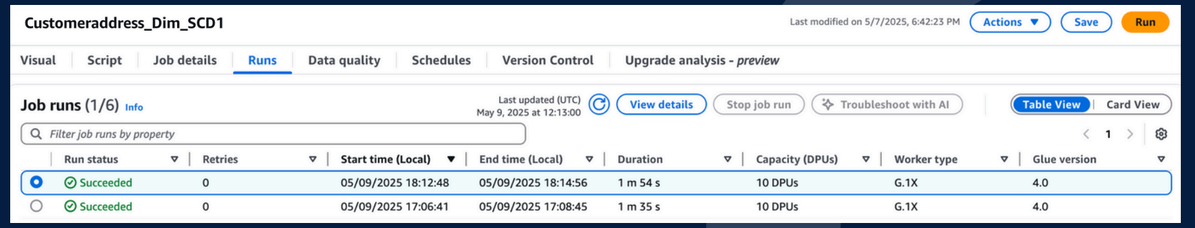

Store_dim_SCD2
Last modified on 5/9/2025, 6:41:23 PM
Actions
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Data source - Data Catalog
Store_db
Data source - Data Catalog
Customer_db
Transform - Change Sche...
Store Schema
Transform - Change Sche...
customer Schema
Data source - Data Catalog
dw store
Transform - Join
Join
Transform - Change Sche...
Renamed keys for Join
Transform - Join
Join
Transform - Dynamic Tra...
Derived Column
Transform - Route
Conditional Router
Transform - SelectFromC...
New_group
Transform - SelectFromC...
default_group
Transform - Dynamic Tra...
Add isActive
Transform - Dynamic Tra...
Add isChange
Transform - Change Sche...
Change Schema
Transform - Filter
Filter
Data target - Data Catalog
Insert new row
Transform - Dynamic Tra...
Add isActive
Transform - Custom code
update row
Transform - Change Sche...
Change Schema
Data target - Data Catalog
AWS Glue Data Catalog

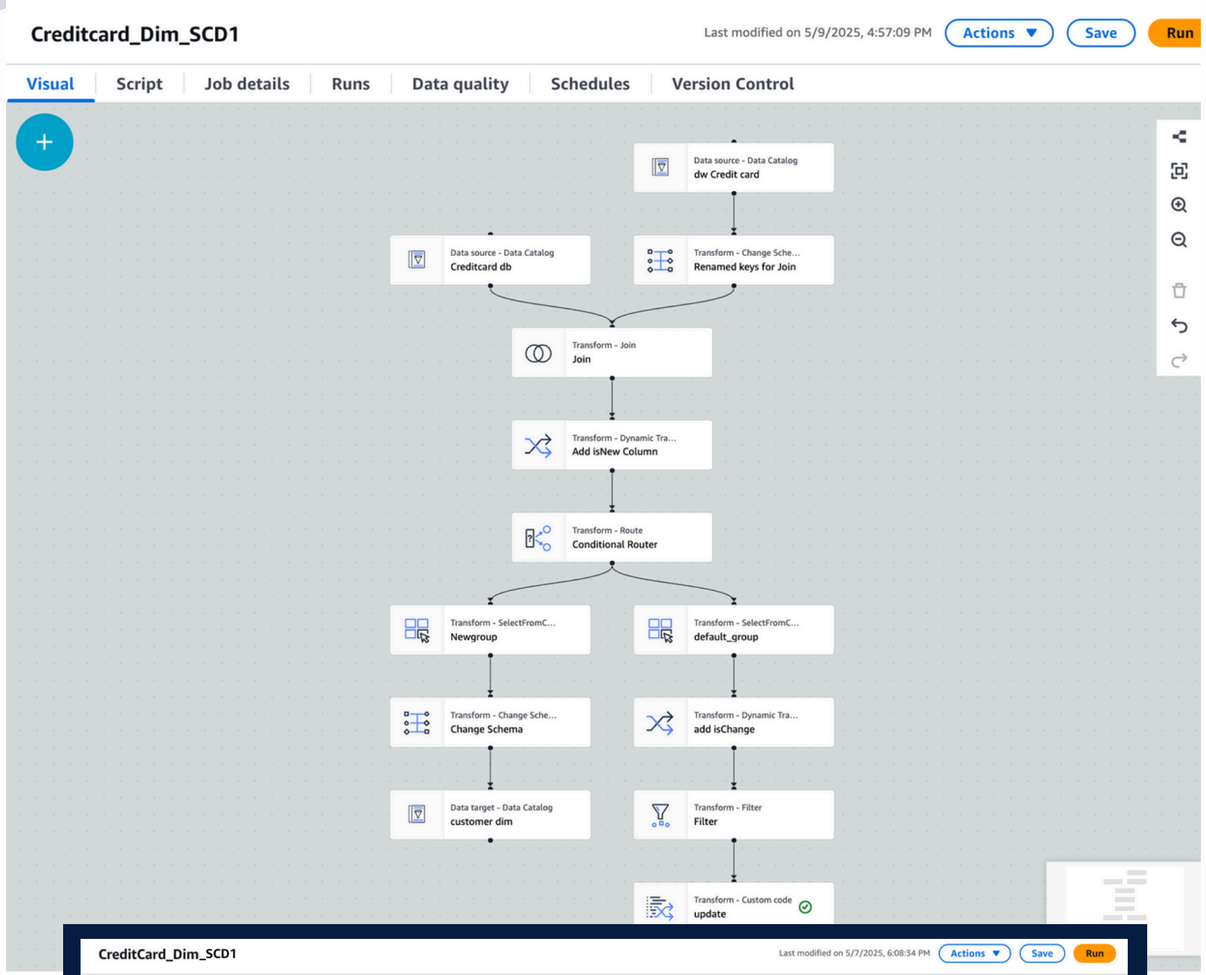

Store_dim_SCD2
Last modified on 5/9/2025, 6:25:30 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Job runs (1/2)
Info
Last updated (UTC)
May 9, 2025 at 12:15:40
View details
Stop job run
Troubleshoot with AI
Table View
Card View
Filter job runs by property
Run status
Retries
Start time (Local)
End time (Local)
Duration
Capacity (DPUs)
Worker type
Glue version
Succeeded
0
05/09/2025 19:01:12
05/09/2025 19:02:32
1 m 3 s
10 DPUs
G.1X
5.0

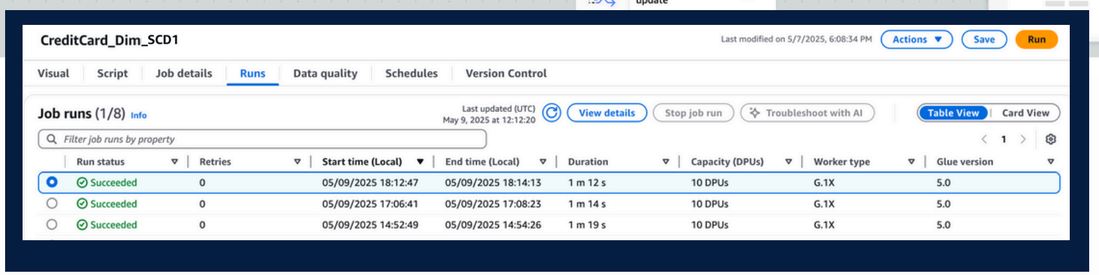

SaleEmployee_dim_SCD1
Last modified on 5/9/2025, 4:44:12 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Data source - Data Catalog
Person db
Data source - Data Catalog
Employee db
Transform - Change Sche...
Renamed keys for Join
Transform - Join
Join
Data source - Data Catalog
SalePerson db
Transform - Change Sche...
Renamed keys for Join
Data source - Data Catalog
dw saleemployee
Transform - Join
Join
Transform - Change Sche...
Renamed keys for Join
Transform - Join
Join
Transform - Dynamic Tra...
Derived Column
Transform - Route
Conditional Router
Transform - SelectFromC...
New_group
Transform - SelectFromC...
default_group
Transform - Change Sche...
Change Schema
Transform - Dynamic Tra...
Add isChange
Data target - Data Catalog
Insert new row
Transform - Filter
Filter
Transform - Custom code
Custom Transform

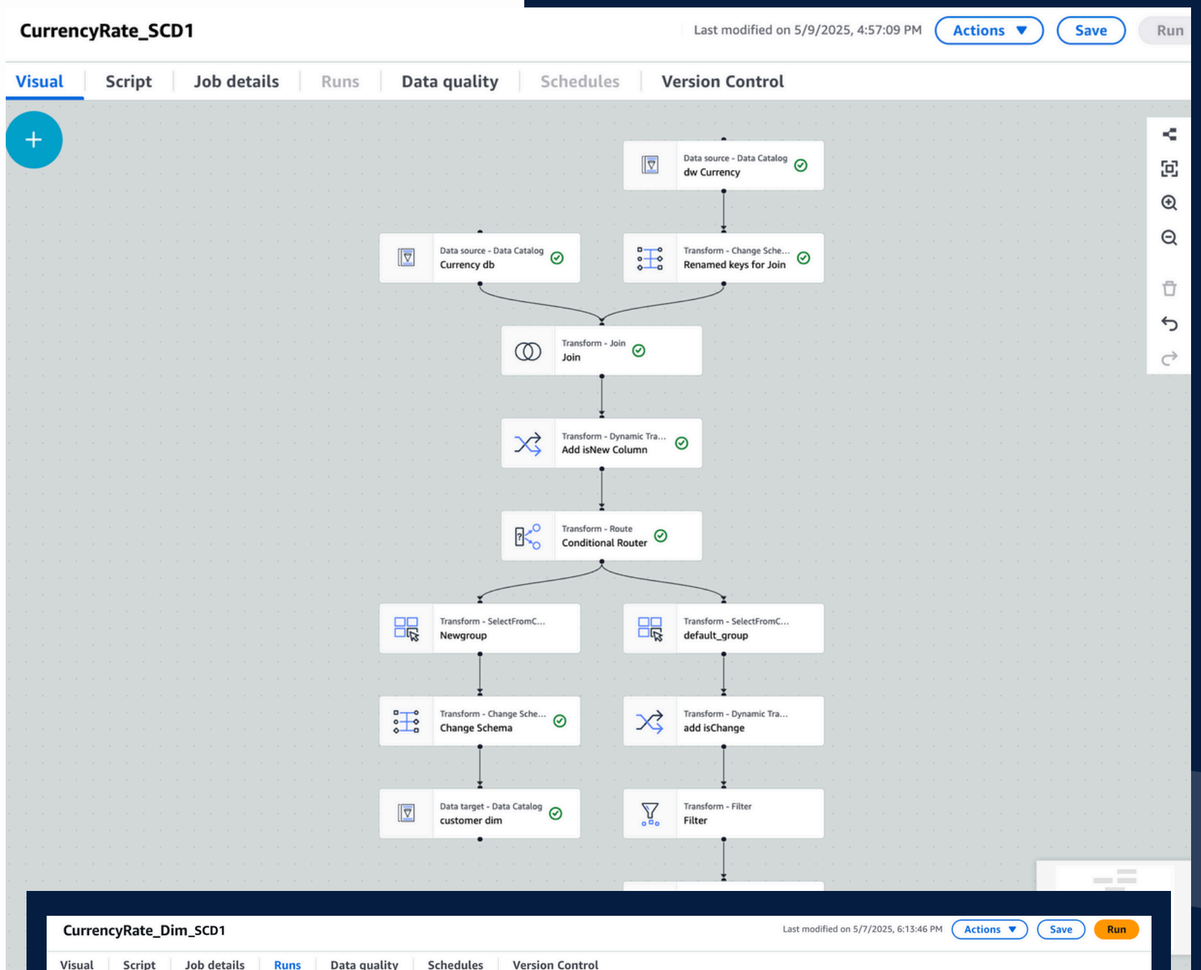

SaleEmployee_dim_SCD1
Last modified on 5/7/2025, 6:42:23 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Upgrade analysis - preview
Job runs (1/6) Info
Last updated (UTC)
May 9, 2025 at 12:13:00
View details
Stop job run
Troubleshoot with AI
Table View
Card View
Filter job runs by property
Run status
Retries
Start time (Local)
End time (Local)
Duration
Capacity (DPUs)
Worker type
Glue version
Succeeded
0
05/09/2025 18:12:48
05/09/2025 18:14:56
1 m 54 s
10 DPUs
G.1X
4.0
Succeeded
0
05/09/2025 17:06:41
05/09/2025 17:08:45
1 m 35 s
10 DPUs
G.1X
4.0

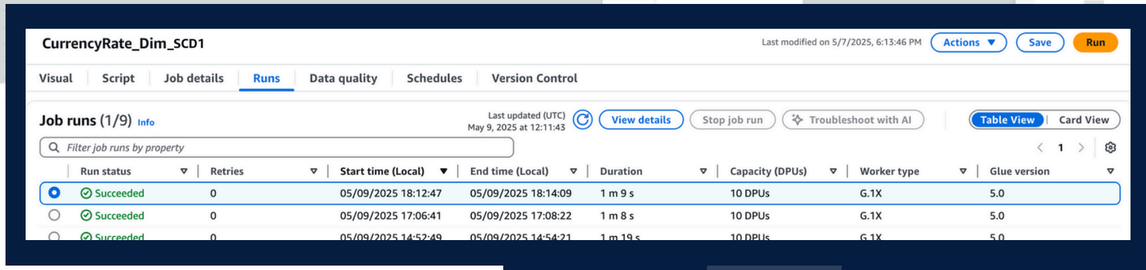

Customeraddress_Dim_SCD1
Last modified on 5/9/2025, 4:45:42 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Data source - Data Catalog
dw CusAdress
Data source - Data Catalog
Customer Adress
Transform - Change Sche...
Renamed keys for Join
Transform - Join
Join
Transform - Dynamic Tra...
Add isNew Column
Transform - Route
Conditional Router
Transform - SelectFromC...
Newgroup
Transform - SelectFromC...
default_group
Transform - Change Sche...
Change Schema
Transform - Dynamic Tra...
add isChange
Data target - Data Catalog
customer dim
Transform - Filter
Filter
Transform - Custom code
update
© 2025, Amazon Web Services, Inc. or

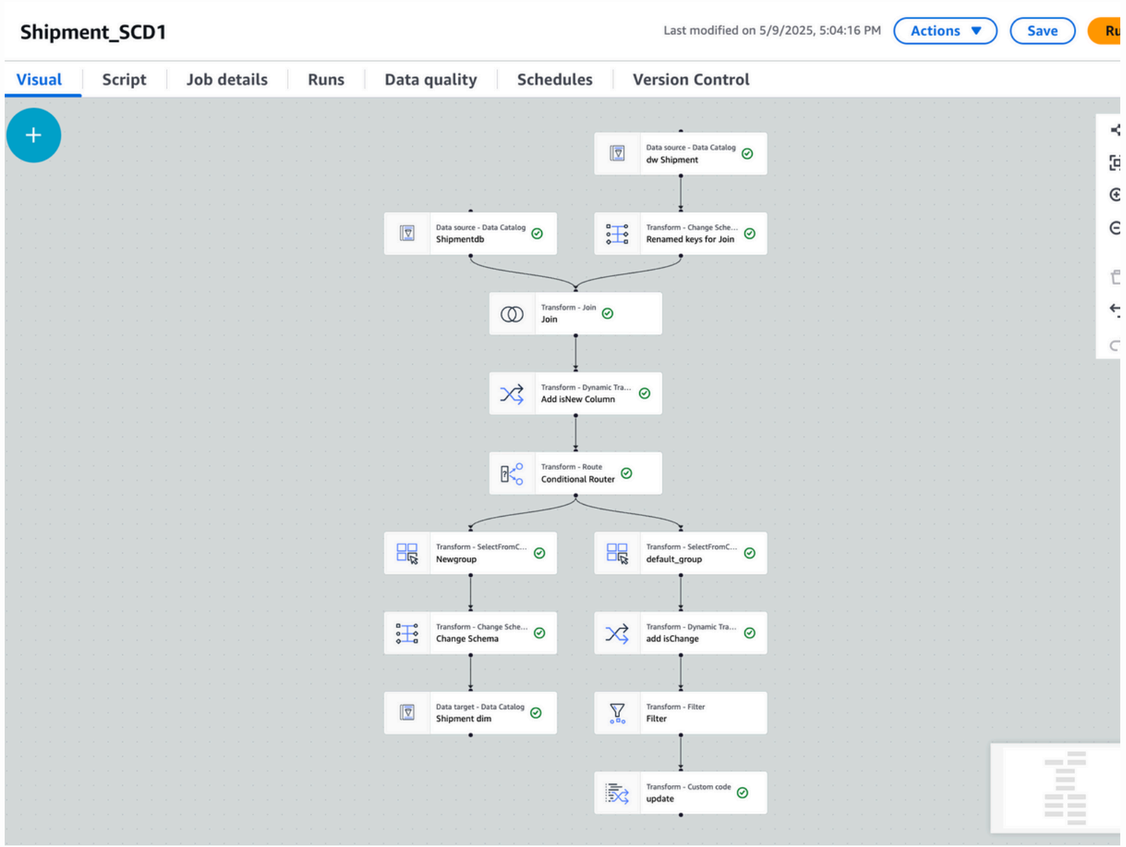

Customeraddress_Dim_SCD1
Last modified on 5/7/2025, 6:42:23 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Upgrade analysis - preview
Job runs (1/6) Info
Last updated (UTC)
May 9, 2025 at 12:13:00
View details
Stop job run
Troubleshoot with AI
Table View
Card View
Filter job runs by property
Run status
Retries
Start time (Local)
End time (Local)
Duration
Capacity (DPUs)
Worker type
Glue version
Succeeded
0
05/09/2025 18:12:48
05/09/2025 18:14:56
1 m 54 s
10 DPUs
G.1X
4.0
Succeeded
0
05/09/2025 17:06:41
05/09/2025 17:08:45
1 m 35 s
10 DPUs
G.1X
4.0

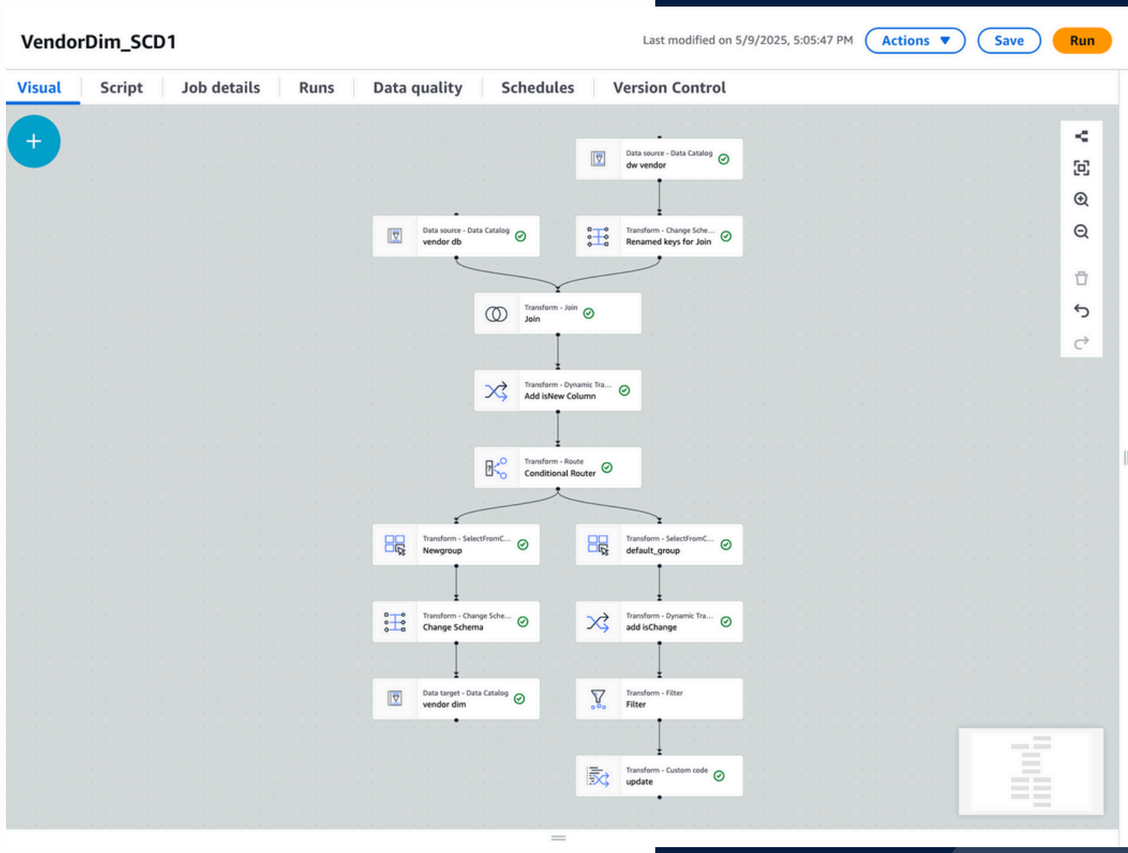

Creditcard_Dim_SCD1
Last modified on 5/9/2025, 4:57:09 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Data source - Data Catalog
dw Credit card
Data source - Data Catalog
Creditcard db
Transform - Change Sche...
Renamed keys for Join
Transform - Join
Join
Transform - Dynamic Tra...
Add isNew Column
Transform - Route
Conditional Router
Transform - SelectFromC...
Newgroup
Transform - SelectFromC...
default_group
Transform - Change Sche...
Change Schema
Transform - Dynamic Tra...
add isChange
Data target - Data Catalog
customer dim
Transform - Filter
Filter
Transform - Custom code
update

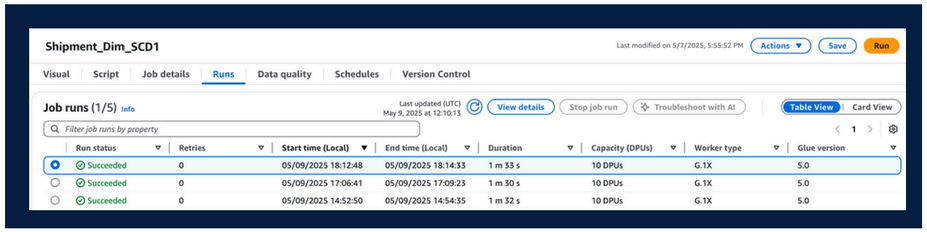

CreditCard_Dim_SCD1
Last modified on 5/7/2025, 6:08:34 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Job runs (1/8) Info
Last updated (UTC)
May 9, 2025 at 12:12:20
View details
Stop job run
Troubleshoot with AI
Table View
Card View
Filter job runs by property
Run status
Retries
Start time (Local)
End time (Local)
Duration
Capacity (DPUs)
Worker type
Glue version
Succeeded 0 05/09/2025 18:12:47 05/09/2025 18:14:13 1 m 12 s 10 DPUs G.1X 5.0
Succeeded 0 05/09/2025 17:06:41 05/09/2025 17:08:23 1 m 14 s 10 DPUs G.1X 5.0
Succeeded 0 05/09/2025 14:52:49 05/09/2025 14:54:26 1 m 19 s 10 DPUs G.1X 5.0

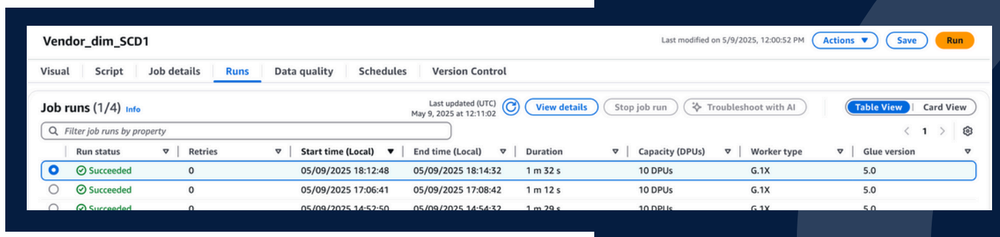

CurrencyRate_SCD1
Last modified on 5/9/2025, 4:57:09 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Data source - Data Catalog
dw Currency
Data source - Data Catalog
Currency db
Transform - Change Sche...
Renamed keys for Join
Transform - Join
Join
Transform - Dynamic Tra...
Add isNew Column
Transform - Route
Conditional Router
Transform - SelectFromC...
Newgroup
Transform - SelectFromC...
default_group
Transform - Change Sche...
Change Schema
Transform - Dynamic Tra...
add isChange
Data target - Data Catalog
customer dim
Transform - Filter
Filter


CurrencyRate_Dim_SCD1
Last modified on 5/7/2025, 6:13:46 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Job runs (1/9) Info
Last updated (UTC)
May 9, 2025 at 12:11:43
View details
Stop job run
Troubleshoot with AI
Table View
Card View
Filter job runs by property
Run status
Retries
Start time (Local)
End time (Local)
Duration
Capacity (DPUs)
Worker type
Glue version
Succeeded 0 05/09/2025 18:12:47 05/09/2025 18:14:09 1 m 9 s 10 DPUs G.1X 5.0
Succeeded 0 05/09/2025 17:06:41 05/09/2025 17:08:22 1 m 8 s 10 DPUs G.1X 5.0

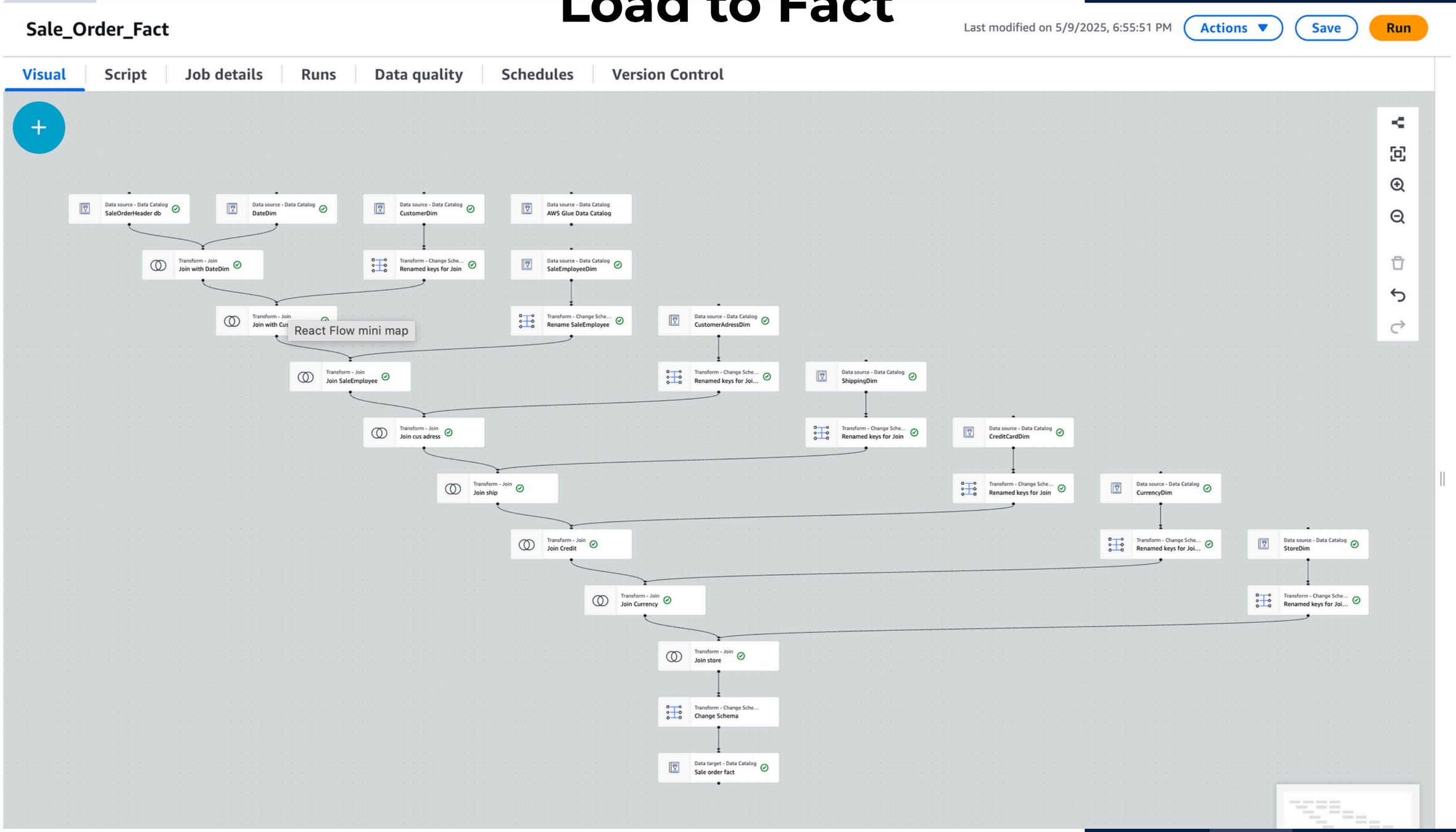

Shipment_SCD1
Last modified on 5/9/2025, 5:04:16 PM
Actions
Save
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Data source - Data Catalog
dw Shipment
Data source - Data Catalog
Shipmentdb
Transform - Change Sche...
Renamed keys for Join
Transform - Join
Join
Transform - Dynamic Tra...
Add isNew Column
Transform - Route
Conditional Router
Transform - SelectFromC...
Newgroup
Transform - SelectFromC...
default_group
Transform - Change Sche...
Change Schema
Transform - Dynamic Tra...
add isChange
Data target - Data Catalog
Shipment dim
Transform - Filter
Filter
Transform - Custom code
update


VendorDim_SCD1
Last modified on 5/9/2025, 5:05:47 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Data source - Data Catalog
dw vendor
Data source - Data Catalog
vendor db
Transform - Change Sche...
Renamed keys for Join
Transform - Join
Join
Transform - Dynamic Tra...
Add isNew Column
Transform - Route
Conditional Router
Transform - SelectFromC...
Newgroup
Transform - SelectFromC...
default_group
Transform - Change Sche...
Change Schema
Transform - Dynamic Tra...
add isChange
Data target - Data Catalog
vendor dim
Transform - Filter
Filter
Transform - Custom code
update

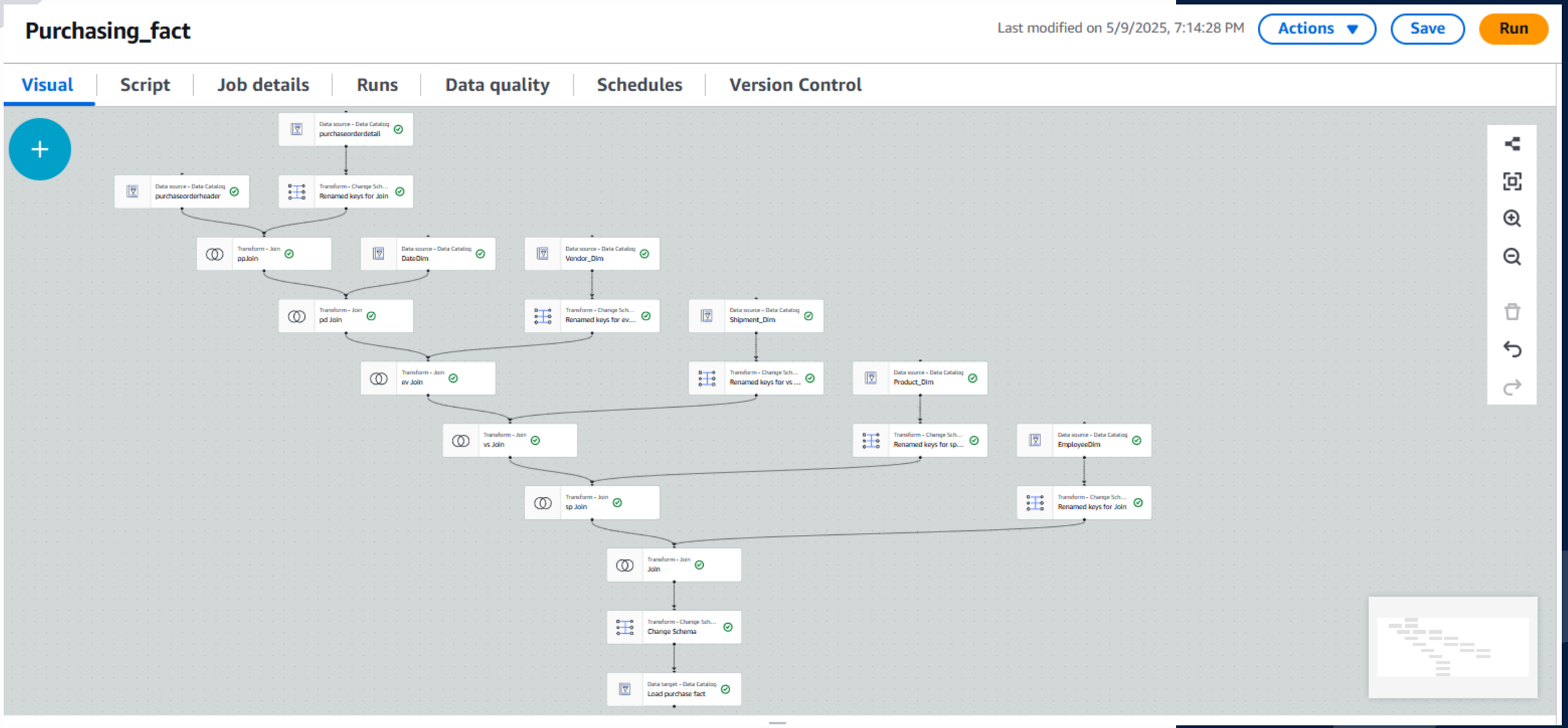

Shipment_Dim_SCD1
Last modified on 5/7/2025, 5:55:52 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Job runs (1/5) Info
Last updated (UTC) May 9, 2025 at 12:10:13
View details
Stop job run
Troubleshoot with AI
Table View
Card View
Filter job runs by property
Run status
Retries
Start time (Local)
End time (Local)
Duration
Capacity (DPUs)
Worker type
Glue version
Succeeded 0 05/09/2025 18:12:48 05/09/2025 18:14:33 1 m 33 s 10 DPUs G.1X 5.0
Succeeded 0 05/09/2025 17:06:41 05/09/2025 17:09:23 1 m 30 s 10 DPUs G.1X 5.0
Succeeded 0 05/09/2025 14:52:50 05/09/2025 14:54:35 1 m 32 s 10 DPUs G.1X 5.0


Vendor_dim_SCD1
Last modified on 5/9/2025, 12:00:52 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Job runs (1/4) Info
Last updated (UTC) May 9, 2025 at 12:11:02
View details
Stop job run
Troubleshoot with AI
Table View
Card View
Filter job runs by property
Run status
Retries
Start time (Local)
End time (Local)
Duration
Capacity (DPUs)
Worker type
Glue version
Succeeded 0 05/09/2025 18:12:48 05/09/2025 18:14:32 1 m 32 s 10 DPUs G.1X 5.0
Succeeded 0 05/09/2025 17:06:41 05/09/2025 17:08:42 1 m 12 s 10 DPUs G.1X 5.0

Load to Fact

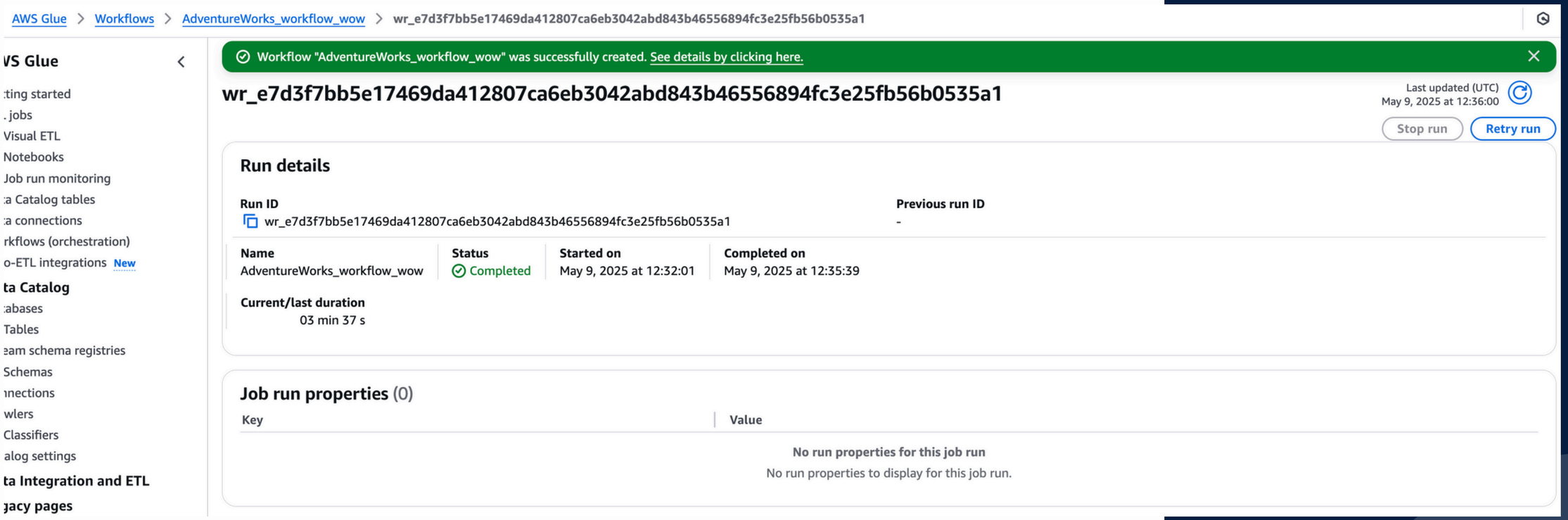

Sale_Order_Fact
Last modified on 5/9/2025, 6:55:51 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
React Flow mini map
Data source - Data Catalog SaleOrderHeader db
Data source - Data Catalog DateDim
Data source - Data Catalog CustomerDim
Data source - Data Catalog AWS Glue Data Catalog
Transform - Join Join with DateDim
Transform - Change Sche... Renamed keys for Join
Data source - Data Catalog SaleEmployeeDim
Transform - Join Join with Cus
Transform - Change Sche... Rename SaleEmployee
Data source - Data Catalog CustomerAdressDim
Transform - Join Join SaleEmployee
Transform - Change Sche... Renamed keys for Joi...
Data source - Data Catalog ShippingDim
Transform - Join Join cus adress
Transform - Change Sche... Renamed keys for Join
Data source - Data Catalog CreditCardDim
Transform - Join Join ship
Transform - Change Sche... Renamed keys for Join
Data source - Data Catalog CurrencyDim
Transform - Join Join Credit
Transform - Change Sche... Renamed keys for Joi...
Data source - Data Catalog StoreDim
Transform - Join Join Currency
Transform - Change Sche... Renamed keys for Joi...
Transform - Join Join store
Transform - Change Sche... Change Schema
Data target - Data Catalog Sale order fact

Load to Fact


Purchasing_fact
Last modified on 5/9/2025, 7:14:28 PM
Actions
Save
Run
Visual
Script
Job details
Runs
Data quality
Schedules
Version Control
Data source - Data Catalog
purchaseorderdetail
Data source - Data Catalog
purchaseorderheader
Transform - Change Sch...
Renamed keys for Join
Transform - Join
pqJoin
Data source - Data Catalog
DateDim
Data source - Data Catalog
Vendor_Dim
Transform - Join
pd Join
Transform - Change Sch...
Renamed keys for ev...
Data source - Data Catalog
Shipment_Dim
Transform - Join
ev Join
Transform - Change Sch...
Renamed keys for vs ...
Data source - Data Catalog
Product_Dim
Transform - Join
vs Join
Transform - Change Sch...
Renamed keys for sp...
Data source - Data Catalog
EmployeeDim
Transform - Join
sp Join
Transform - Change Sch...
Renamed keys for Join
Transform - Join
Join
Transform - Change Sch...
Change Schema
Data target - Data Catalog
Load purchase fact

Workflow (Orchestration)

Workflow (Orchestration)



THANK YOU!