Demand Forecast | Data Workbench

Introduction

- This Hands-on Exercise is a part of **Building A Demand Forecast Skill**.
- This is the first exercise you will perform as part of the course.. The Steps for accessing the sandbox are common across all the exercises.
- You will create a single Workspace as part of the skill-building process.
- All the Hands-on exercises need to be executed in this Workspace.
- The data in some of the Screenshots does not match the instructions and are illustrative; in such cases, you are requested to follow the instructions.
- Use **<username>** as an identifier of all objects that you create in the Sandbox during the course.

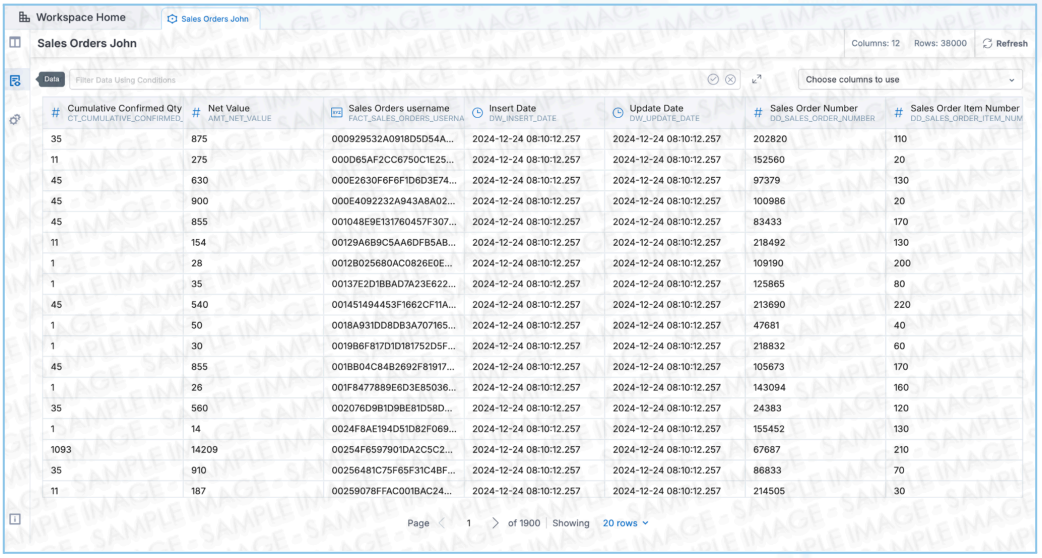
Key Takeaway

In this hands-on, you will learn to:

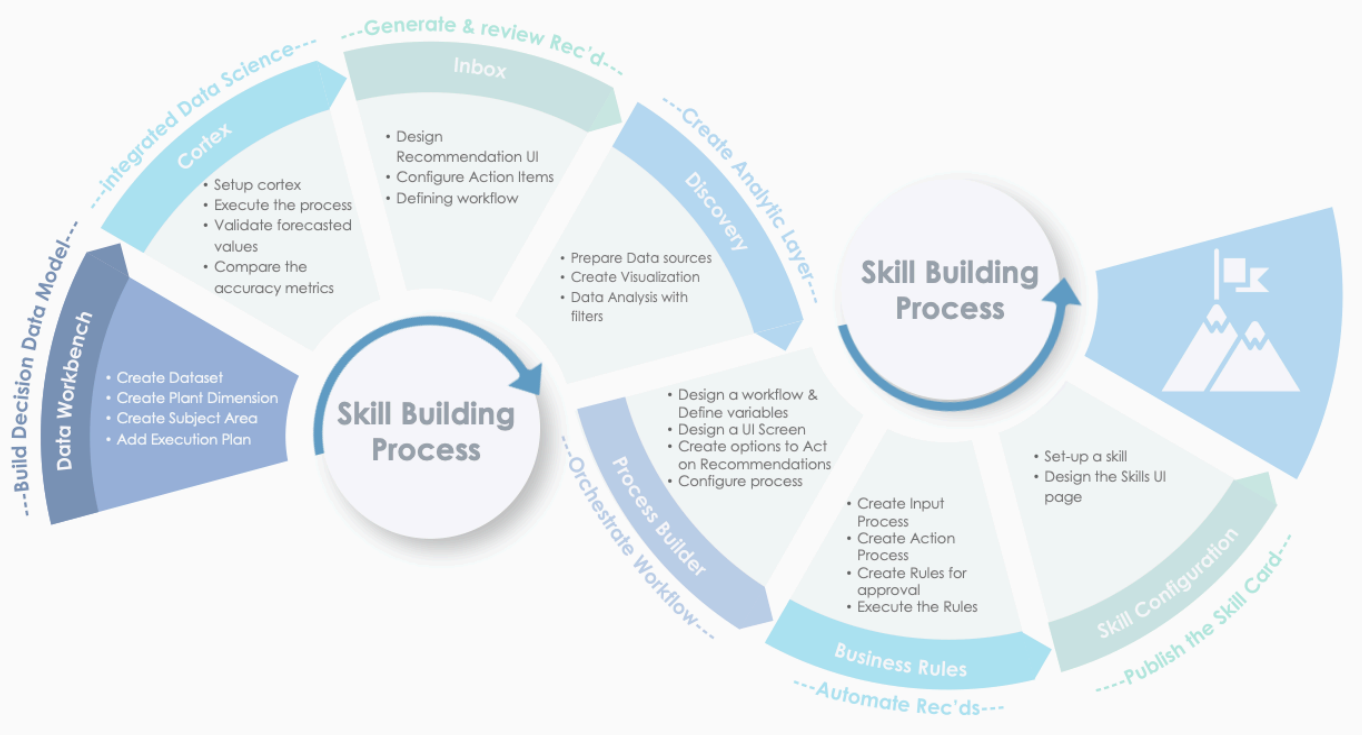
- **Create Dataset:** Extract data from a flat file using the Crawler.
- **Create Plant Dimension:** Create and populate Dimensions using both SQL queries and Insert/Update steps.
- **Create Subject Area:** Construct a Subject Area from scratch and populate it with data from multiple sources using both SQL and Insert/Update methods.
- **Add Execution Plan:** Configure the Execution Plan to orchestrate the process of building the Decision Data Model.
- Monitor job status and review execution history within the Control Room.

Expected Results

Build a **Sales Order_<Username>** Subject Area



Skill Building Process



Typical Errors

Data is not populated:

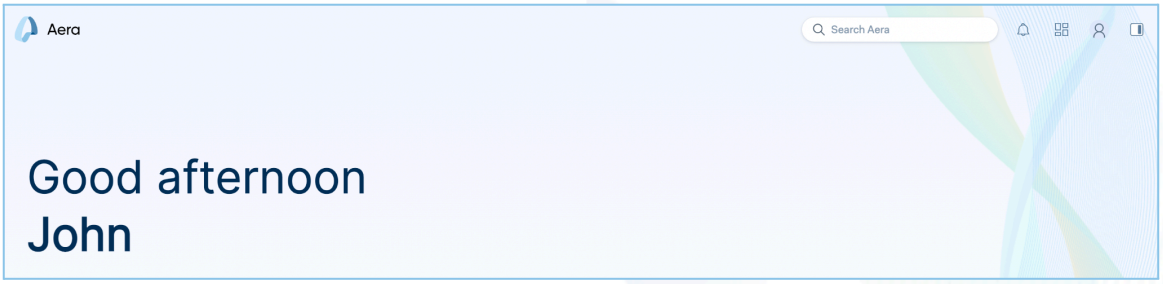
- Please review the length of the columns in your table.
- Ensure the query you write uses the correct physical names of columns.
- View detailed SQL errors to identify and fix issues quickly.
- Ensure you use the specific target table name you created to populate the objects.
- Review selection and mapping of parameters.

Tasks and Steps

Getting Started with Data Workbench

Log in to the Aera Decision Cloud

1. Log in to the Aera sandbox using your credentials. The URL is shared via email with notifications@aeratechnology.com. After a successful login, the Aera Decision Cloud homepage will be displayed.



In this exercise, you will build a Decision Data Model by orchestrating data from the source (on-premise system or cloud) to the destination (Aera Decision Data Model).

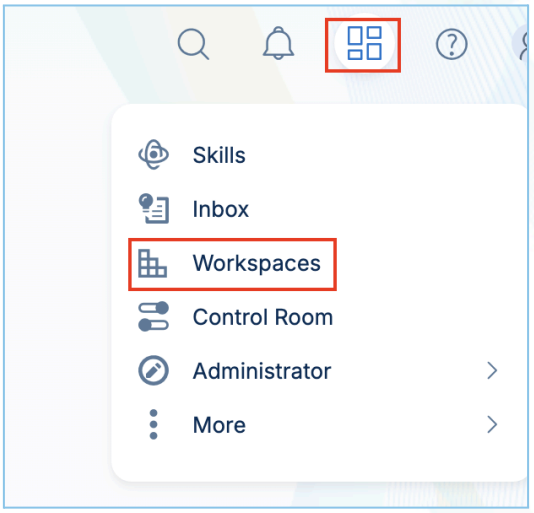
As this is a Sandbox environment and we are not connected to the source system, we will use an already **uploaded flat file** as a data source.

You will perform the following actions to build the Decision Data Model:

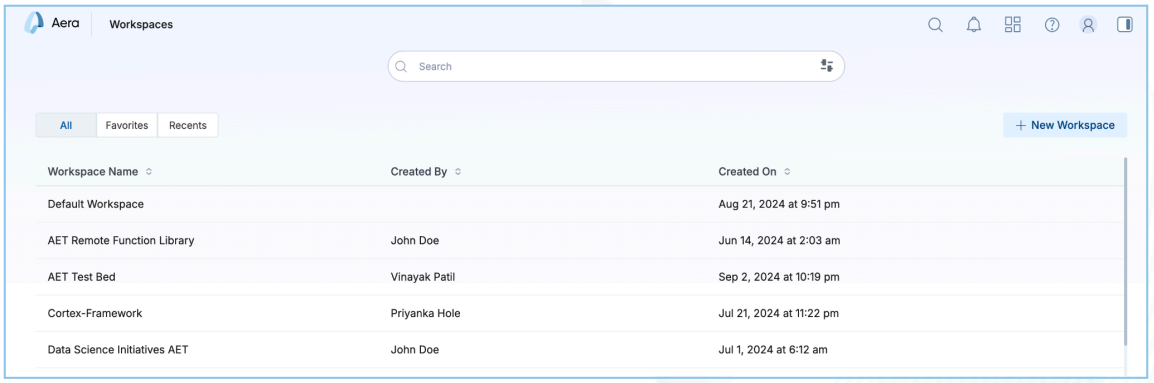
- Set up a Workspace
- Create Dataset
- Create Dimension
- Populate the Dimension
- Create a new Subject Area (Sales Order)
- Populate the Subject Area
- Add an Execution Plan

Set up a Workspace

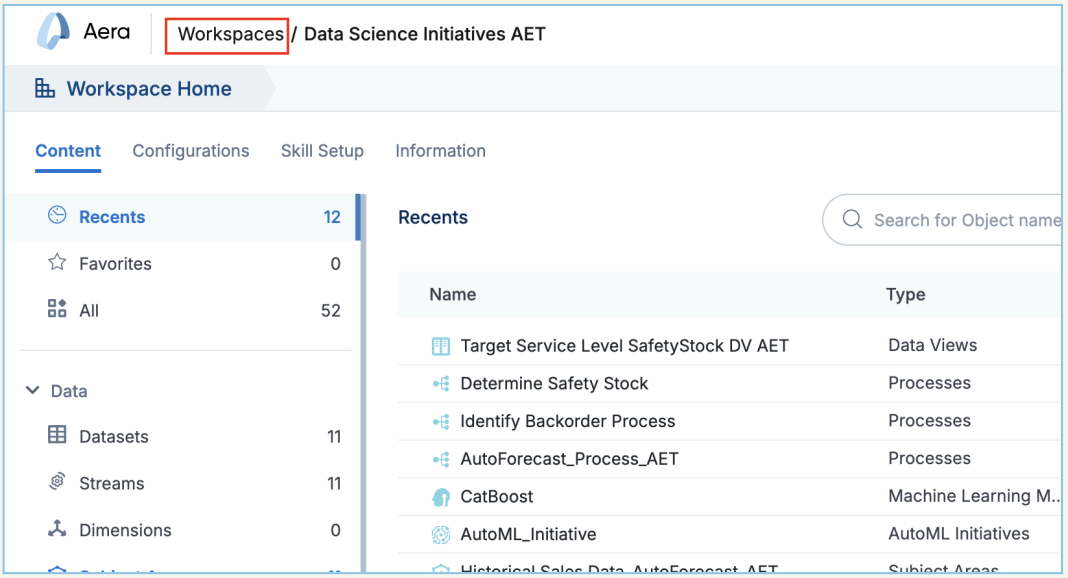
1. Click the **Aera Menu** icon on the Aera Decision Cloud home page and then click **Workspaces**.



2. After you click on **Workspaces**, you'll be taken to the Workspaces home.

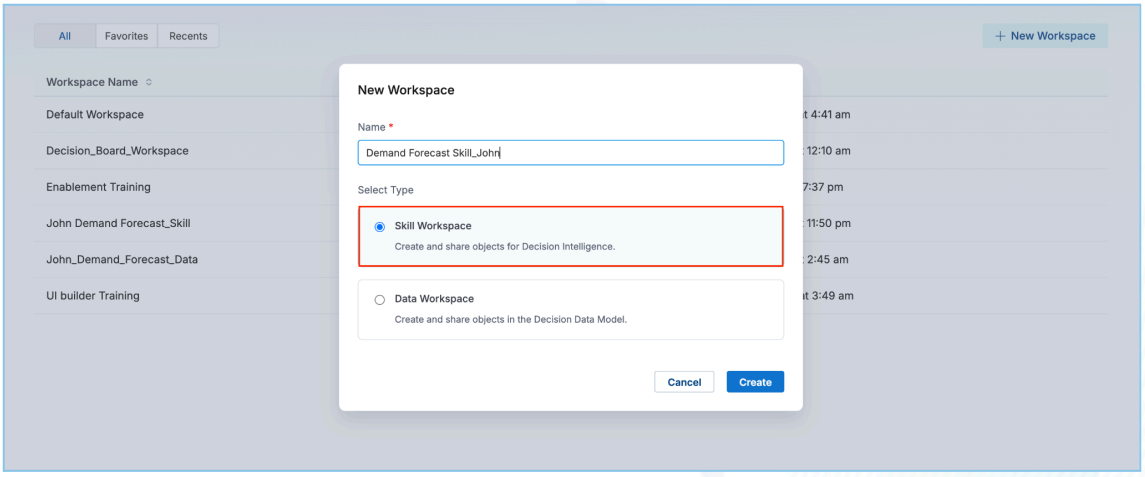


Note: If you're currently in a Workspace, click on Workspaces in the top-left corner to return to the Workspaces home.

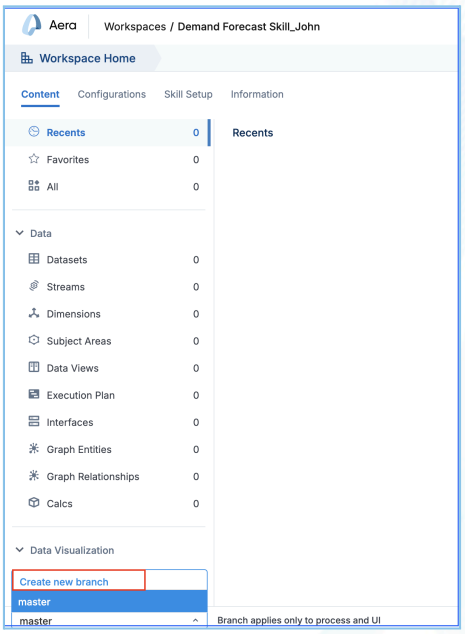


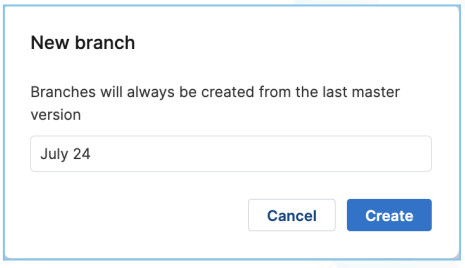
3. From the Aera Workspaces page, create a new **Skill-**type Workspace named **Demand Forecast Skill_<Username>** and click **Create**.

There are three types of Workspaces: Default Workspace, Data Workspace and Skill Workspace.
The **Default Workspace** contains reference or archived objects. It can be used to build test skills and create POCs.
The **Data Workspace** is used to build and organise a single Decision Data Model that represents a Customer's entire business. Here we can build common KPIs, Datasets, Dimensions and Subject Areas that can be reused across different Skills. It is generally a global Workspace.
Skill Workspaces are used to build and manage all the objects and configurations related to a particular skill.



4. Create a new branch named **<Month YY>** and click **Create**.



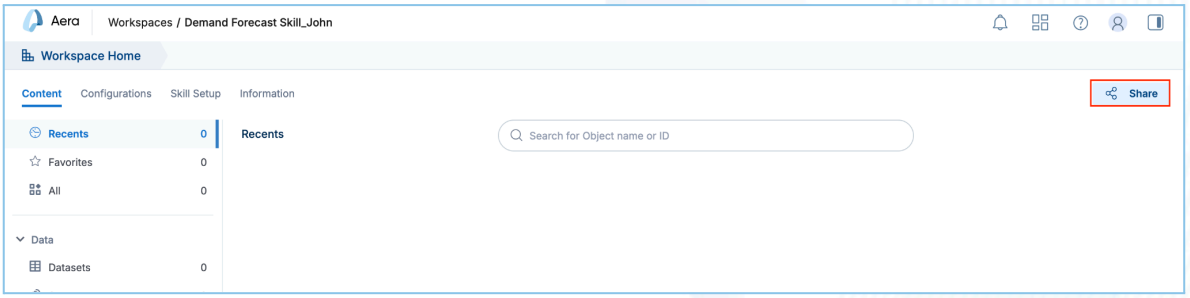


5. The newly created Workspace and its branch are now displayed.

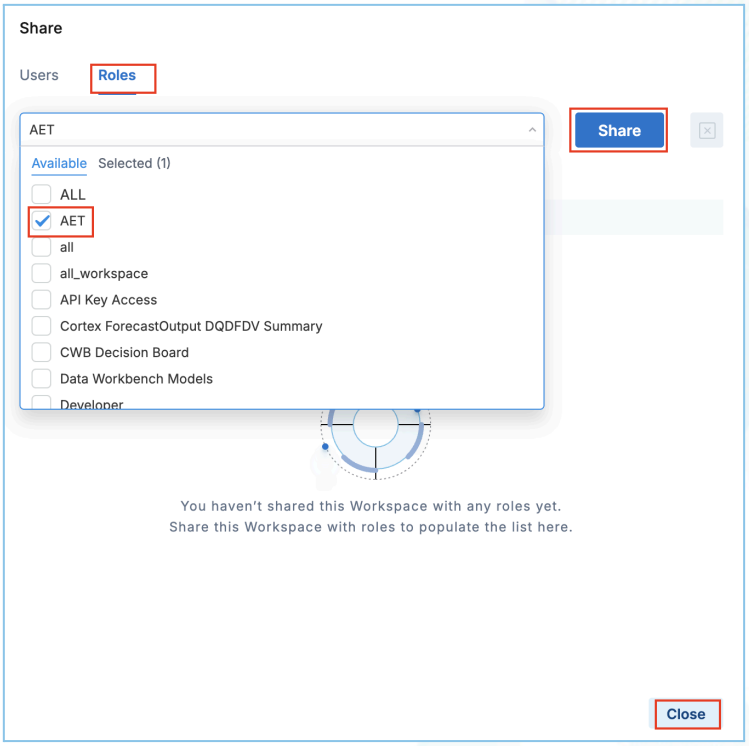
Notice there are different categories of objects in the content panel, such as Data, Data Visualization, Data Science and Automation.

Note: After creating your Workspace with the branch, share it with **AET** Roles.

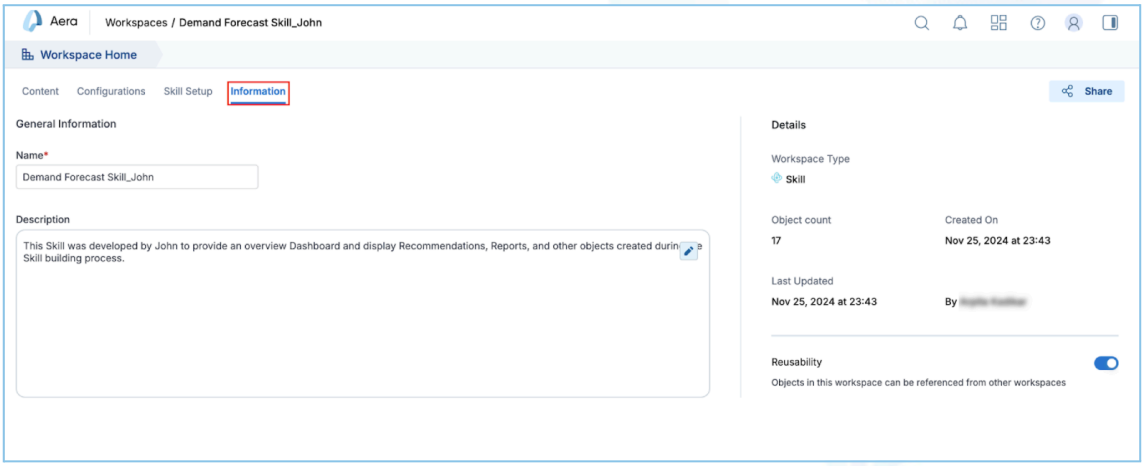
6. Click on the **Share** option to share the Workspace with the Enablement Team. (This will help the Enablement Team to review your Workspace and assist you with troubleshooting).



7. On the Share window, navigate to the **Roles** tab and Type or select the **AET** role. Once you've selected the **AET** role, click **Share** and **Close.**



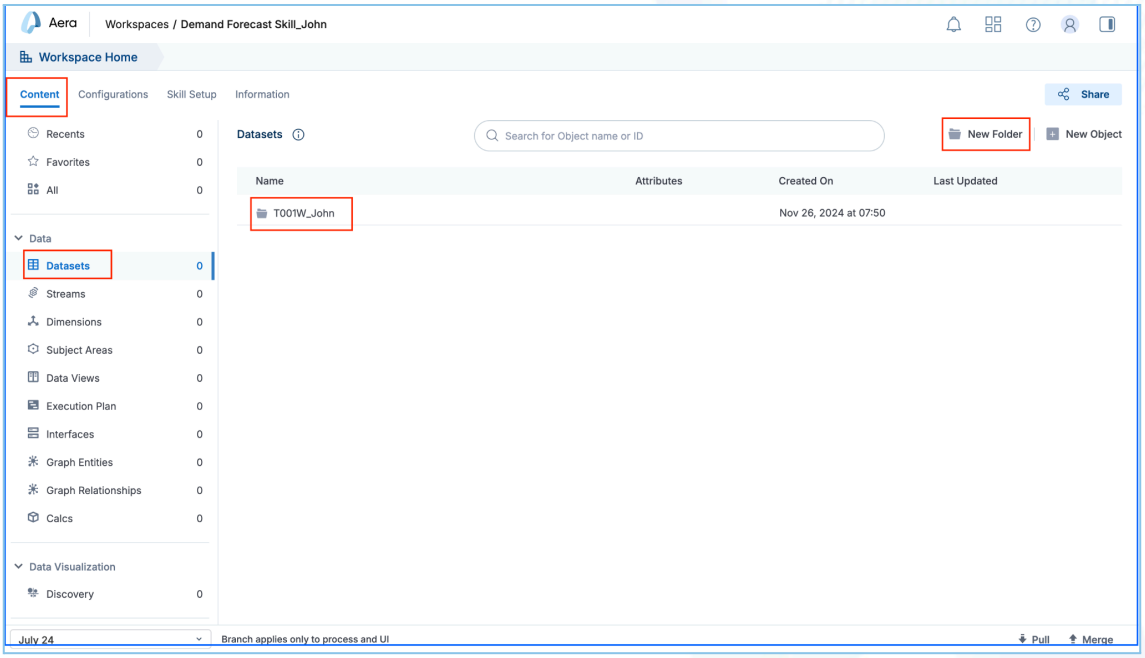
8. Click on the **Information tab** to view the essential information for a Workspace.



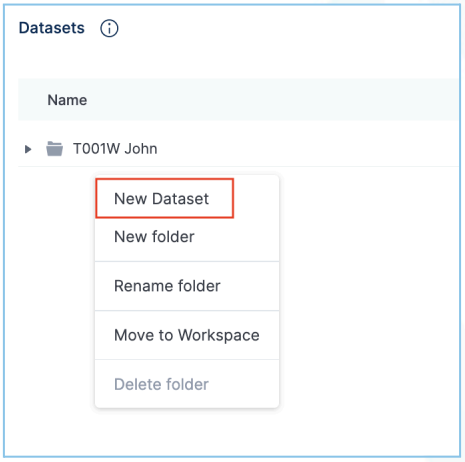
The **Information tab** provides important details about your Workspace, such as its name, type, description, the number of objects within, and the last update timestamp. From here, you specify a reusable Workspace attribute.
Once enabled, reusable objects in this Workspace can be reused in other Workspaces.

Define and Configure Data Extraction

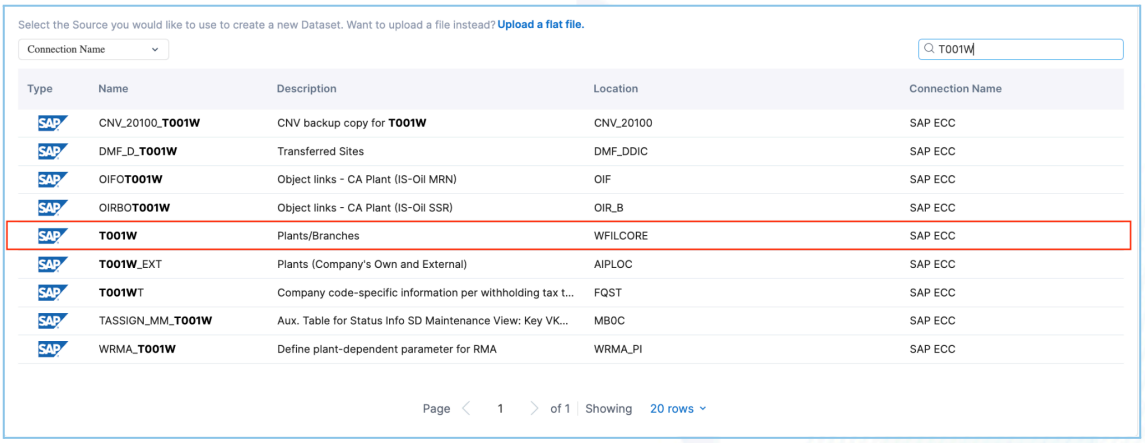
1. Navigate to **Datasets** within the Data category on the **Content** tab and click the '**New Folder**'. Name it **T001W<Username>**



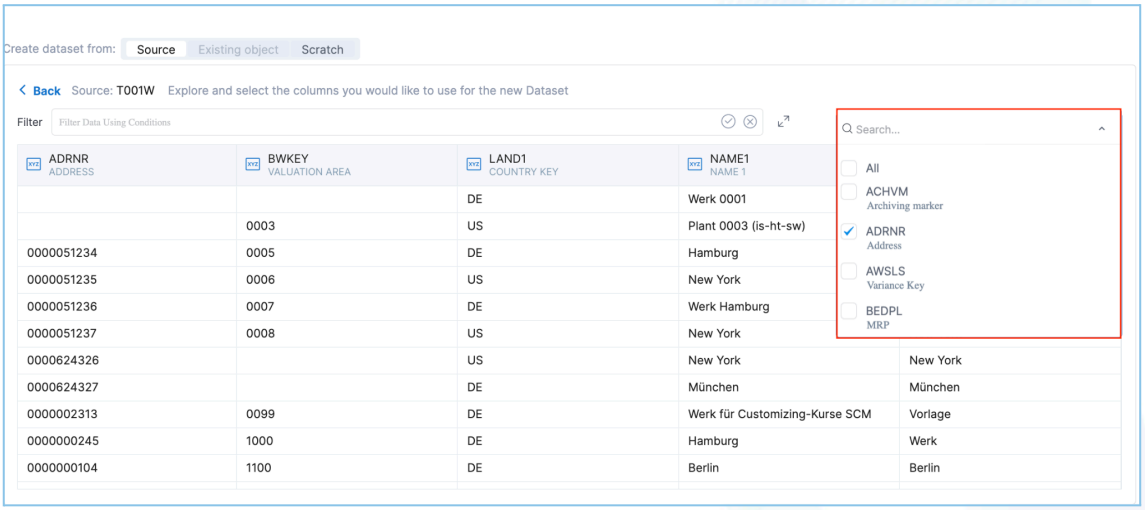
2. Next, right-click on the folder to create a dataset, and the Create Dataset page will be displayed.



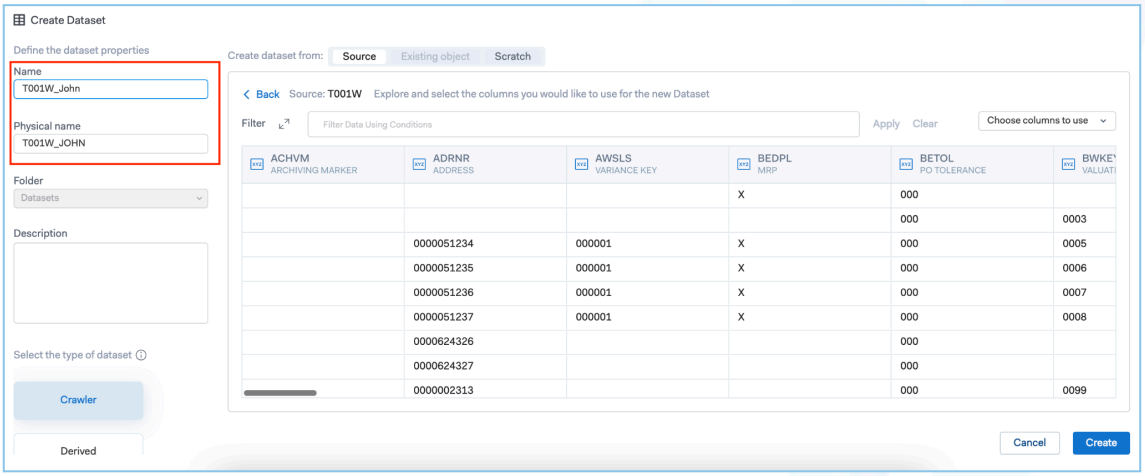
3. Navigate to the **"Source"** tab, type **"T001W"** in the search box to choose the file, and then **select the** columns listed below.



4. From the **Choose the columns to use** dropdown, deselect **ALL** and select the following columns: **ADRNR, BWKEY, LAND1, NAME1, NAME2, ORT01, PSTLZ, VKORG, WERKS.**

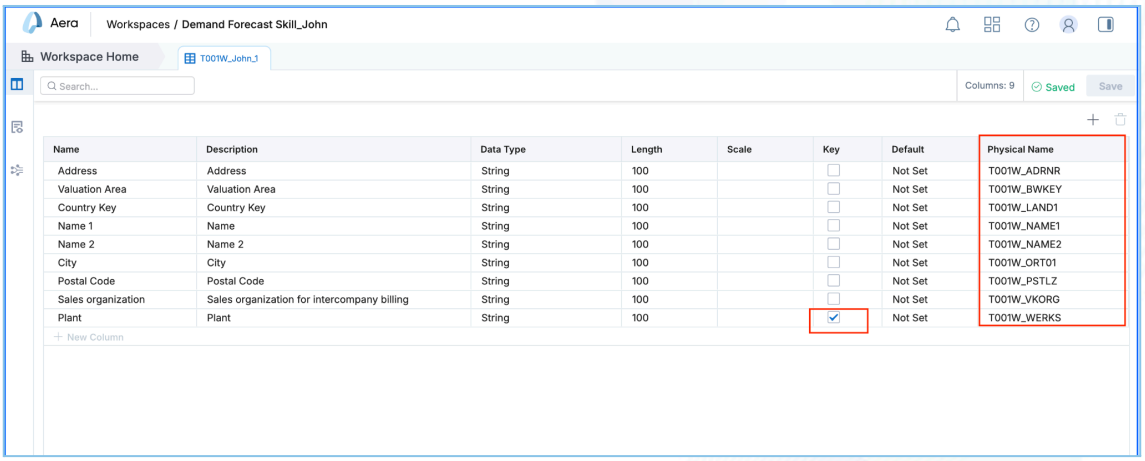


5. Name the dataset **T001W_<Username>**, then click **Create**.



6. Click the **Physical name** column to rename the fields to include the prefix **T001W_** and select the **Key** Option for the **Plant** field.

The naming convention here identifies the column names with the source file for easy reference. The physical names are then used in the queries constructed later in the exercise.

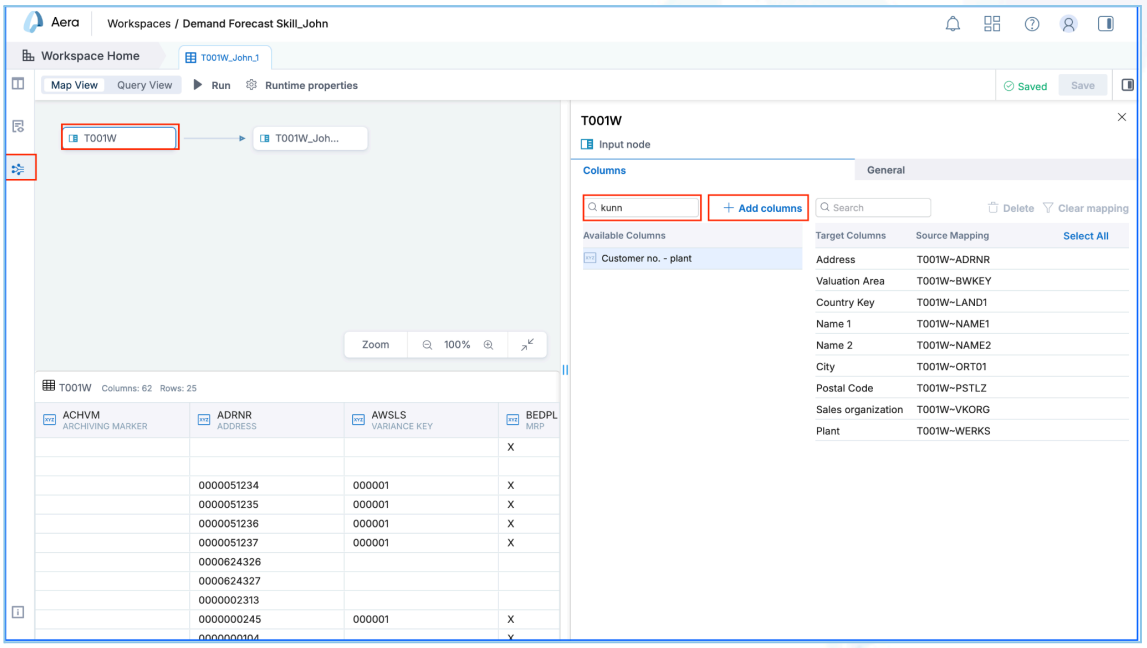


7. Click **Save**.

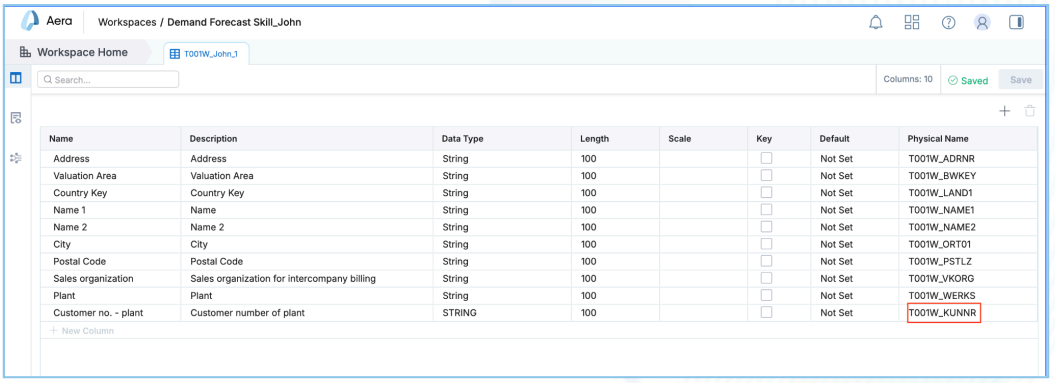
You can now specify one or more columns in a data object as a Natural key. The system uses this information to auto-generate the join conditions when the data object is used.

Modify the Crawler Configuration

1. Click the **Crawler** button to display the **Map view,** allowing you to add additional columns whenever needed.
2. Click the **T001W** Source table, find and select the "**KUNNR**" column, then click "**Add Column**" to include it and click **Save.**



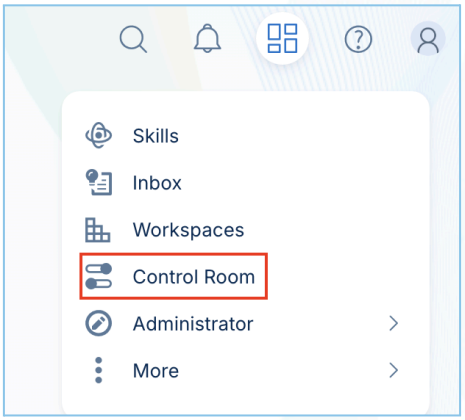
3. Navigate to the Columns tab, change the column's physical name to **T001W_KUNNR** and make sure to click "**Save**".



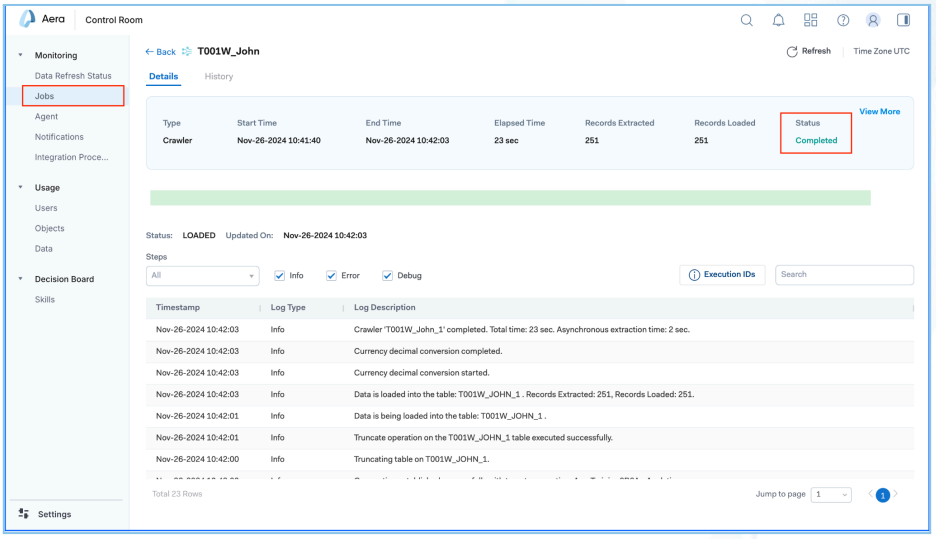
Populate the Dataset

1. From the **crawler** tab, click **Run** to populate the Data into your target table.
2. From the **Aera Menu**, navigate to **Control Room**, then click on **Jobs** under **Monitoring** in the top left corner

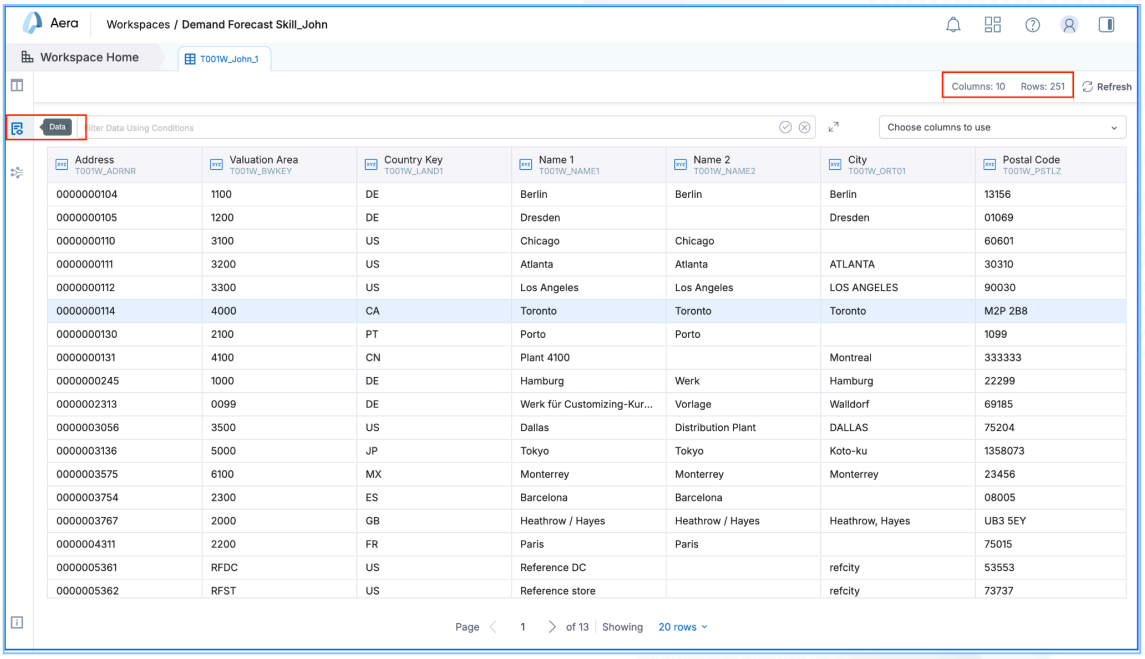
> It takes a few minutes to update the status. Try refreshing the console after 5-10 min.



3. A list of jobs is displayed with their respective status, such as completed, failed, or running.



4. Once the job is **completed**, navigate to **"Workspaces"** > **"Your Dataset"** and select the **"Data"** tab to examine the populated data. Ensure all required fields are present.



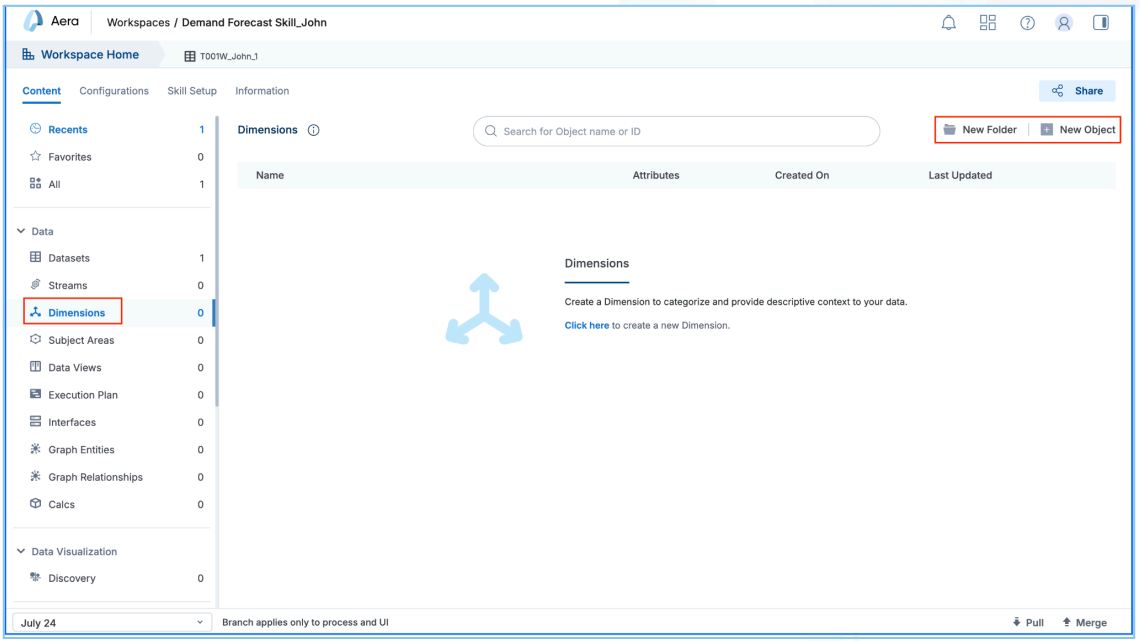
Address T001W_ADRNR	Valuation Area T001W_BWKEY	Country Key T001W_LAND1	Name 1 T001W_NAME1	Name 2 T001W_NAME2	City T001W_ORT01	Postal Code T001W_PSTLZ
0000000104	1100	DE	Berlin	Berlin	Berlin	13156
0000000105	1200	DE	Dresden		Dresden	01069
0000000110	3100	US	Chicago	Chicago		60601
0000000111	3200	US	Atlanta	Atlanta	ATLANTA	30310
0000000112	3300	US	Los Angeles	Los Angeles	LOS ANGELES	90030
0000000114	4000	CA	Toronto	Toronto	Toronto	M2P 2B8
0000000130	2100	PT	Porto	Porto		1099
0000000131	4100	CN	Plant 4100		Montreal	333333
0000000245	1000	DE	Hamburg	Werk	Hamburg	22299
0000002313	0099	DE	Werk für Customizing-Kur...	Vorlage	Walldorf	69185
0000003056	3500	US	Dallas	Distribution Plant	DALLAS	75204
0000003136	5000	JP	Tokyo	Tokyo	Koto-ku	1358073
0000003575	6100	MX	Monterrey	Monterrey	Monterrey	23456
0000003754	2300	ES	Barcelona	Barcelona		08005
0000003767	2000	GB	Heathrow / Hayes	Heathrow / Hayes	Heathrow, Hayes	UB3 5EY
0000004311	2200	FR	Paris	Paris		75015
0000005361	RFDC	US	Reference DC		refcity	53553
0000005362	RFST	US	Reference store		refcity	73737

Page 1 of 13 | Showing 20 rows

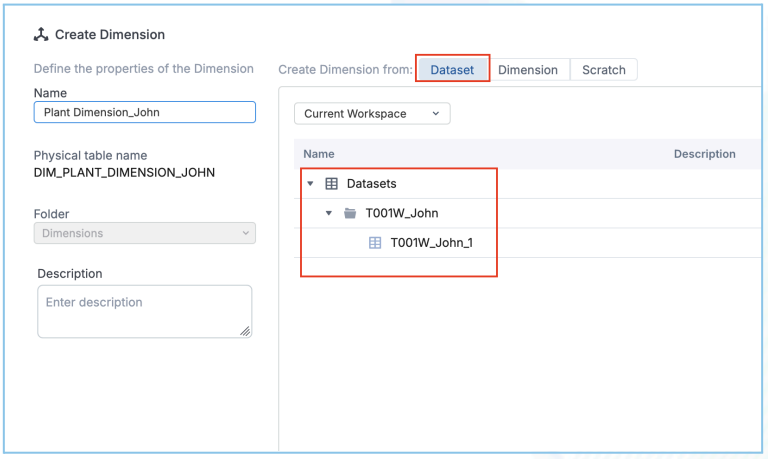
Create Dimensions

In this section
A **dimension** in a data model refers to a structure that categorizes and provides context to numerical measures, typically representing business entities like time, product, customer, location, etc.
When a dimension is created from a dataset, typically an ERP (Enterprise Resource Planning) system, it contains attributes that allow for more detailed analysis and slicing of the data.

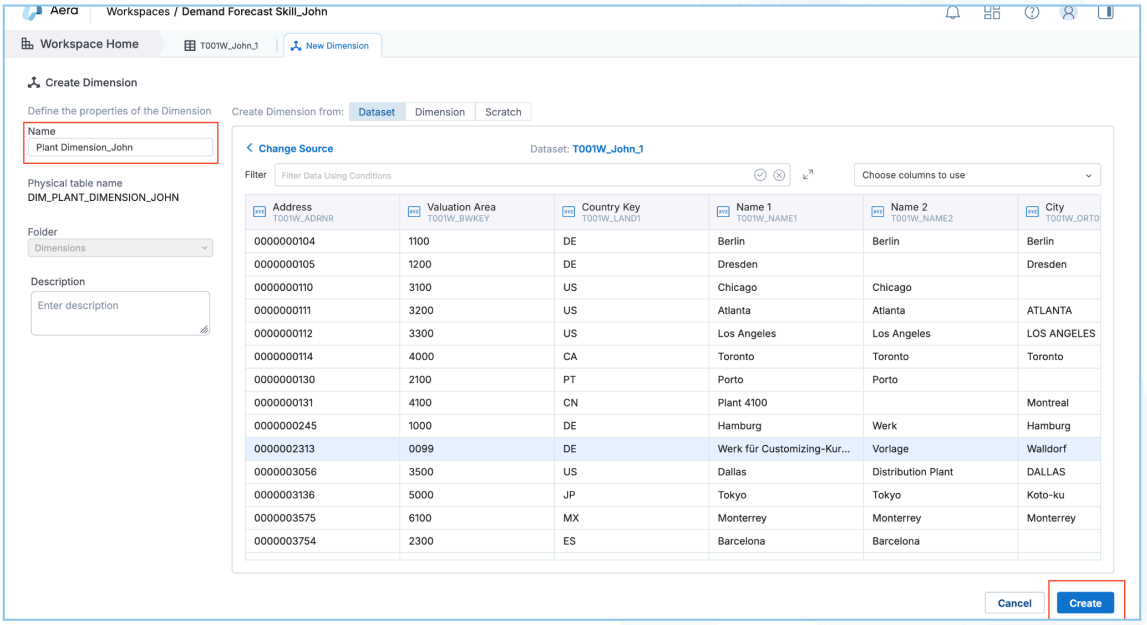
1. Click the "**New Object**" located at the top right to create a **Dimension**.



2. Select the Dataset **T001W_<Username>** and provide a name **Plant Dimension <username>**



3. Once the records are displayed, click **Create**. The Plant dimension table opens in a new tab.



4. Update the **Physical names** of the columns as shown below.

T001W_WERKS -> PLANT

T001W_NAME1 -> PLANT_NAME

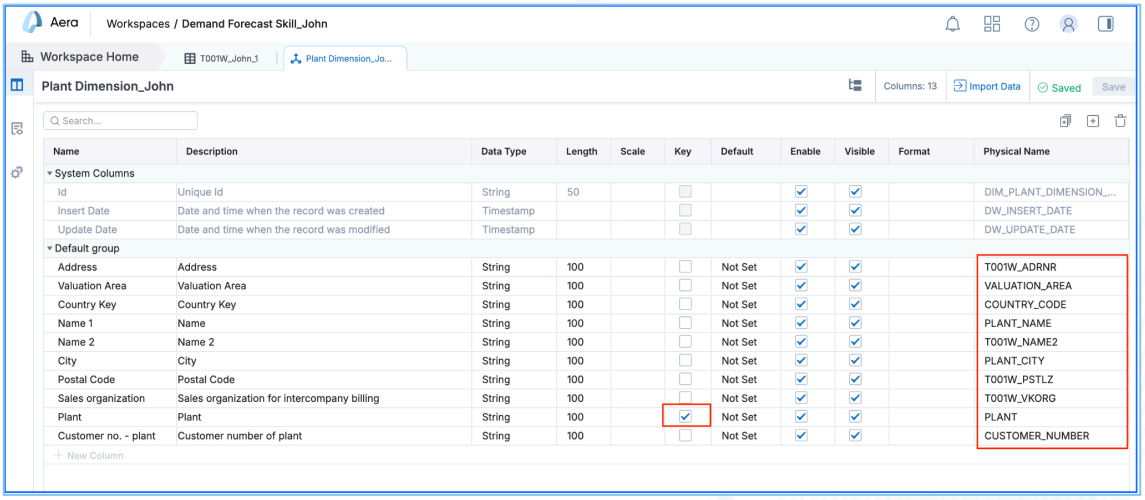
T001W_ORT01 -> PLANT_CITY

T001W_KUNNR -> CUSTOMER_NUMBER

T001W_LAND1 -> COUNTRY_CODE

T001W_BWKEY -> VALUATION_AREA

6. Enable the **Key** option for **PLANT** and click **Save**



Add Stream to Populate Dimension

Streams are a component in the platform that defines and executes the data processing steps and populates the data in the Decision Data Model.

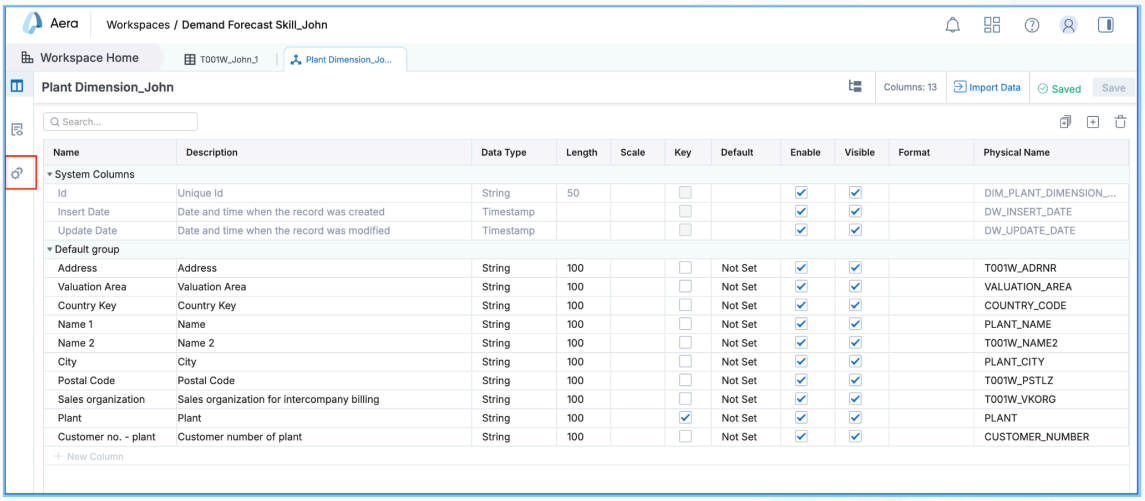
Streams support SQL queries, Formula steps and predefined Functions.

There are two ways of adding the Stream.

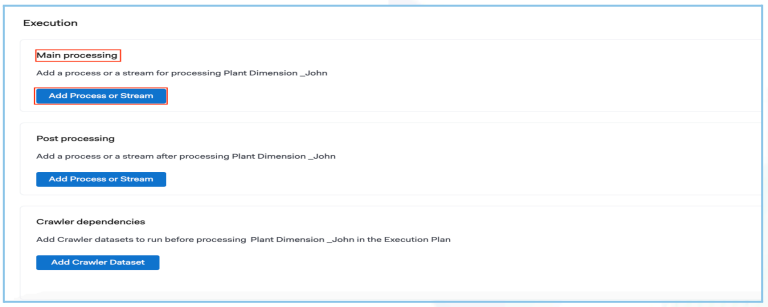
1. Click the + icon beside the Streams tab on the Workspace home.
2. Create Stream from the execution tab of the Dimensions window.

We will use the second method.

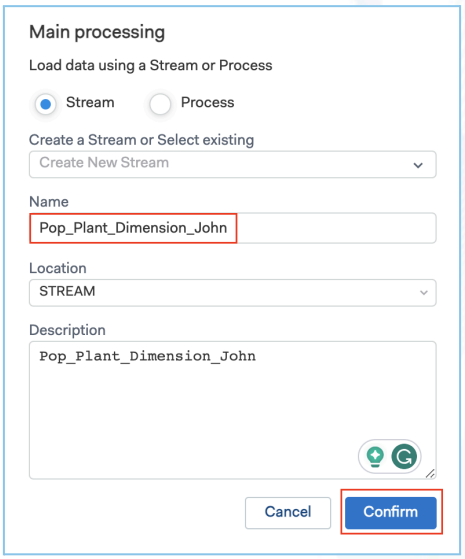
1. Click the **Executions** tab on the Plant Dimensions window.



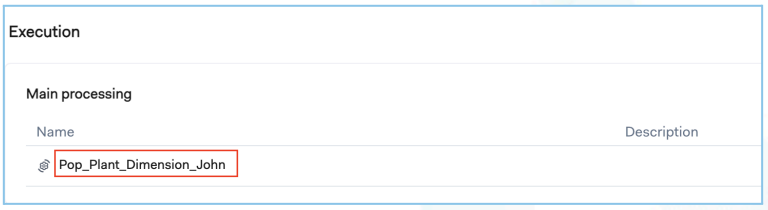
2. To add a process or stream, click on "Add Process or Stream" in the Main Processing section.



3. To load the data, select a **Stream** from the dropdown, choose **"Create New Stream,"** name it **"Pop_Plant_Dimension_<Username>**," and click **"Confirm**."



4. The Stream gets added, click on the Stream, **Pop_Plant_Dimension_<Username>**



5. A new tab is where we define the Stream steps to populate the Dimension.

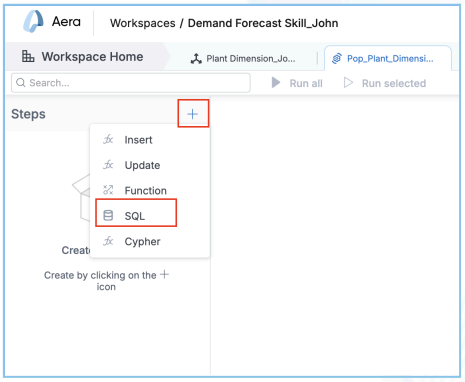
We can populate the Dimension by either using the **SQL Step** or by using the **Insert/Update** formula steps.

When using SQL queries to populate a Dimension or Subject Area, you have direct access to the underlying database and can leverage SQL's full power and flexibility. This approach is generally preferred for complex data transformations.
Insert/Update steps provide a more visual and streamlined approach to populating a Subject Area. This method is generally suitable for simpler data-loading scenarios where complex SQL queries are not required.

We will first populate the Dimension using SQL queries and later populate one more Dimension using Insert/Update Steps.

Populate Dimension Using SQL

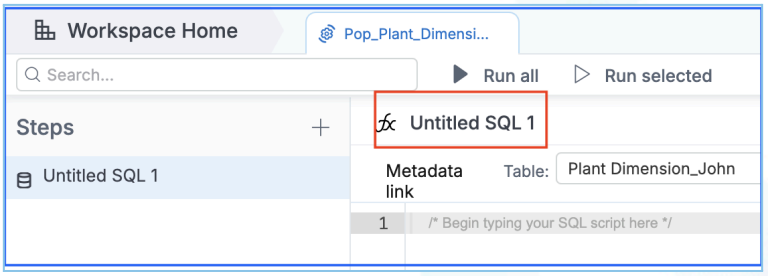
1. On the Stream Steps tab, click the "**+**" icon and select **SQL**.



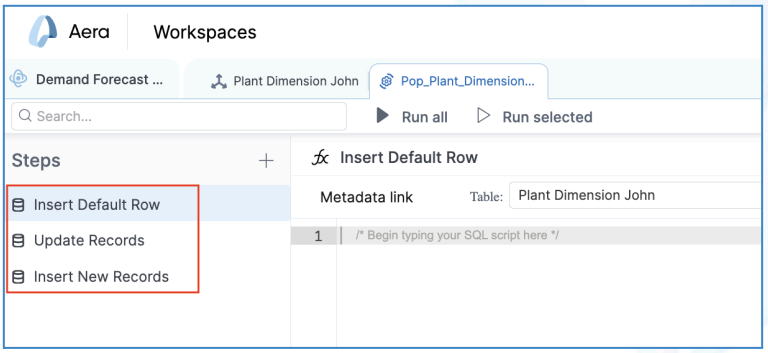
You will create three individual steps for each significant operation:

- **Insert Default Row:** We add a default row consisting of a 40-digit hash value that serves as the Dimension ID and timestamps, and we set the other fields to 'Not Set'. This step ensures that the Dimension table is not empty and we don't get a NULL table whenever a Join is performed on this Dimension.
- **Update Records:** This Step checks whether any records have changed since the last import. If so, we will update the records to reflect the new entries.
- **Insert New Records:** This Step is used to insert any new records that previously did not exist

2. Rename the step as **Insert Default Row**.



3. Add two more **SQL** steps, name them **Update Records,** and **insert new records**.

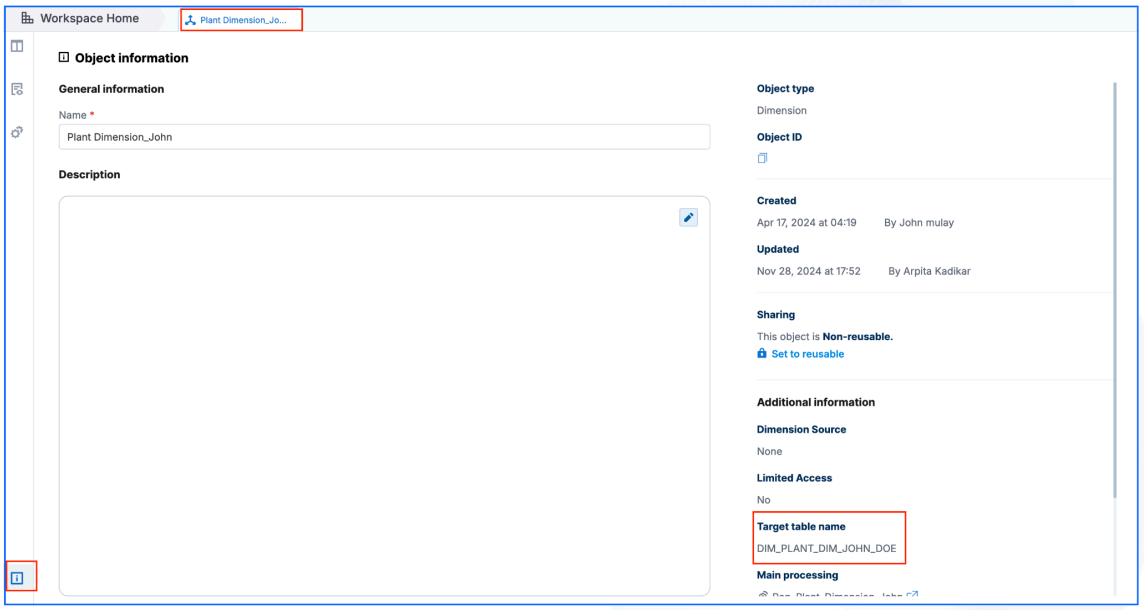


Read the query and understand the purpose of each query before you run the steps.

4. Click the first step, **Insert Default Row**, and copy-paste the below query in the right panel.

```
insert into your_dimension_table (your_dimension_tableID, DW_INSERT_DATE, DW_UPDATE_DATE,
PLANT, PLANT_NAME, PLANT_CITY, CUSTOMER_NUMBER, COUNTRY_CODE,
VALUATION_AREA)
Select
'00000000000000000000000000000000000000000', current_timestamp, current_timestamp, 'Not
Set', 'Not Set', 'Not Set', 0 ,'Not Set', 'Not Set'
from (select 1) as src
where
not exists (select 1 from your_dimension_table where your_dimension_tableID =
'00000000000000000000000000000000000000000');
```

5. To find the names of your tables, navigate to the **"Plant Dimension_<Username>"** and **"T001W_<Username>"** tabs, then click the **"About"** option (identified by an **"i"** icon).



6. In the streams tab, click "**Update Records**" in the second step and paste the following query.

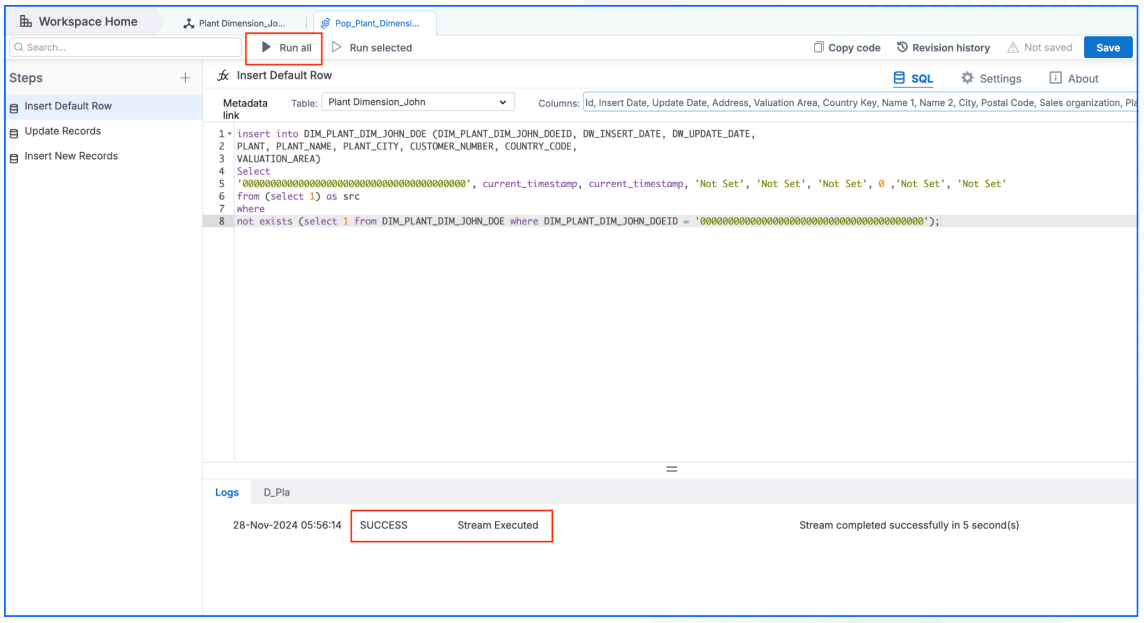
```
UPDATE your_dimension_table dp
SET COUNTRY_CODE = ifnull(T001W_LAND1, 'Not Set'),
     dp.dw_update_date = current_timestamp
from your_dataset_table mk, your_dimension_table dp
WHERE mk.T001W_WERKS = dp.PLANT
and COUNTRY_CODE <> ifnull(T001W_LAND1, 'Not Set');

UPDATE your_dimension_table dp
SET PLANT_NAME = ifnull(T001W_NAME1, 'Not Set'),
     dp.dw_update_date = current_timestamp
from your_dataset_table mk, your_dimension_table dp
WHERE mk.T001W_WERKS = dp.PLANT
```

```
and PLANT_NAME <> ifnull(T001W_NAME1, 'Not Set');

UPDATE your_dimension_table dp
SET PLANT_CITY = ifnull(T001W_ORT01, 'Not Set'),
     dp.dw_update_date = current_timestamp
from your_dataset_table mk, your_dimension_table dp
WHERE mk.T001W_WERKS = dp.PLANT
and PLANT_CITY <> ifnull(T001W_ORT01, 'Not Set');

UPDATE your_dimension_table dp
SET CUSTOMER_NUMBER = ifnull(T001W_KUNNR, 0),
     dp.dw_update_date = current_timestamp
from your_dataset_table mk, your_dimension_table dp
WHERE mk.T001W_WERKS = dp.PLANT
and CUSTOMER_NUMBER <> ifnull(T001W_KUNNR, 0);

UPDATE your_dimension_table dp
SET COUNTRY_CODE = ifnull(T001W_LAND1, 'Not Set'),
     dp.dw_update_date = current_timestamp
from your_dataset_table mk, your_dimension_table dp
WHERE mk.T001W_WERKS = dp.PLANT
and COUNTRY_CODE <> ifnull(T001W_LAND1, 'Not Set');

UPDATE your_dimension_table dp
SET VALUATION_AREA = ifnull(T001W_BWKEY, 'Not Set'),
     dp.dw_update_date = current_timestamp
from your_dataset_table mk, your_dimension_table dp
WHERE mk.T001W_WERKS = dp.PLANT
and VALUATION_AREA <> ifnull(T001W_BWKEY, 'Not Set');
```

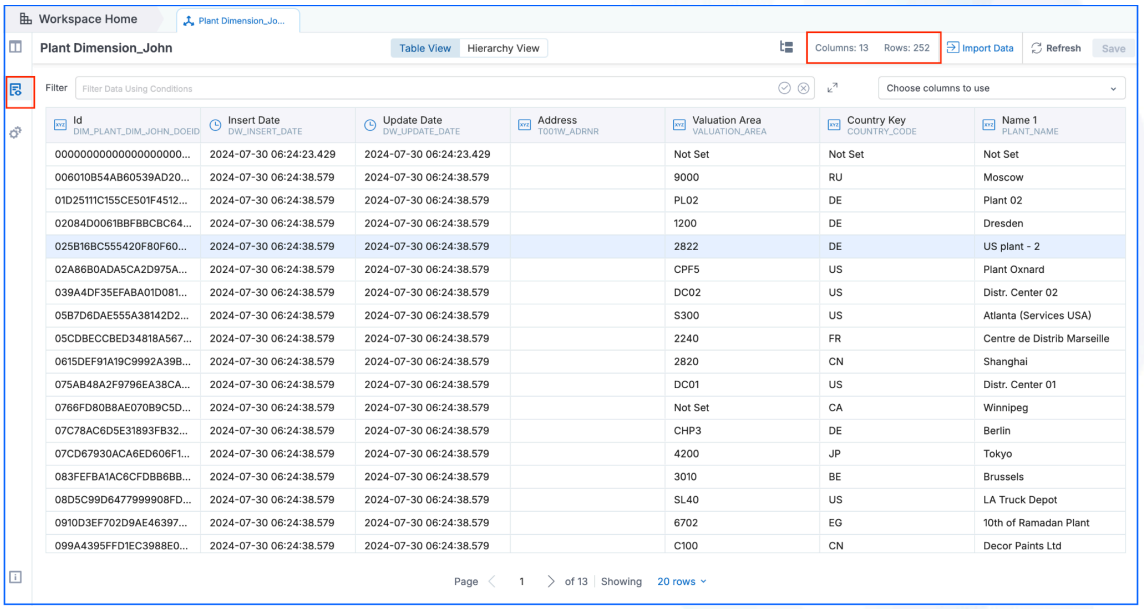
8. Click the third step, **insert new records**, and copy and paste the query below.

```
insert into your_dimension_table (your_dimension_tableID, DW_INSERT_DATE, DW_UPDATE_DATE,
PLANT, PLANT_NAME, PLANT_CITY, CUSTOMER_NUMBER, COUNTRY_CODE, VALUATION_AREA)
SELECT distinct
        ifnull(CAST(UPPER(HASH_SHA1(UPPER(IFNULL(T001W_WERKS, ' ') || '|^|' ||
                                          'SAP ECC EURO'))) AS VARCHAR(40)),
               '00000000000000000000000000000000000000000') as ID,
current_timestamp as DW_INSERT_DATE,
current_timestamp as DW_UPDATE_DATE,
ifnull(T001W_WERKS,'Not Set') as PLANT,
ifnull(T001W_NAME1,'Not Set') as PLANT_NAME,
ifnull(T001W_ORT01,'Not Set') as PLANT_CITY,
ifnull(T001W_KUNNR,0) as CUSTOMER_NUMBER,
ifnull(T001W_LAND1,'Not Set') as COUNTRY_CODE,
ifnull(T001W_BWKEY,'Not Set') as VALUATION_AREA
FROM your_dataset_table T001W
where not exists (
select 1 from your_dimension_table a where a.PLANT = T001W.T001W_WERKS);
```

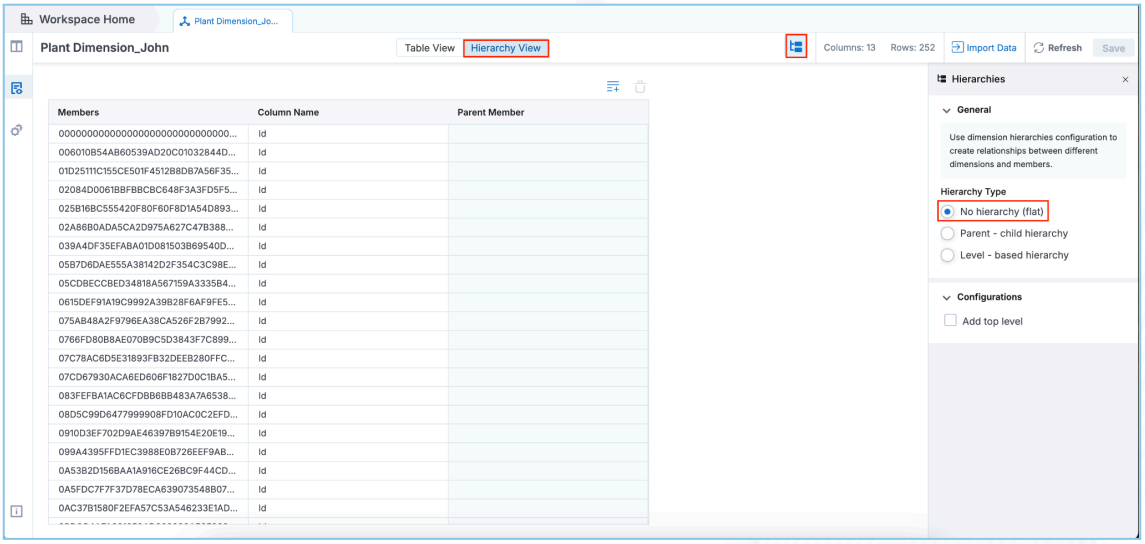
9. Click **"Save"** and then **"Run All"** to execute all three steps simultaneously; when the execution is successful, the status will display **"Completed."**



10. Go to the Dimension, **Plant Dimension_<Username>** and check the data populated by default under the **Table view** of the data tab with the column names you specified in the above steps.



11. Additionally, you can view the data in the **Hierarchy View** tab. By default, all dimensions created are treated as flat dimensions, meaning all the members are treated as being on the same level without any parent-child relationships or inherent hierarchy.



12. To see all revisions of your stream and choose an earlier version, go back to the stream **Pop_Plant_Dimension_<Username>** and click **Revision History**.

Note: You can quickly compare Stream revisions to identify changes at each step. You can select two revisions and view the differences at both the stream and step levels.

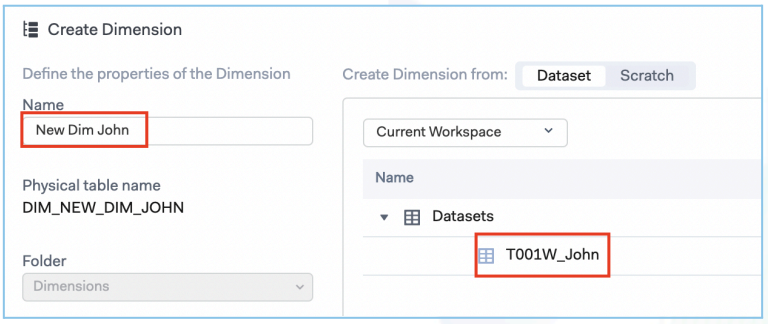
Populate Dimension Using Insert/Update Formula

In this section, you will create a New Dimension and populate it using the Insert and Update Steps. We will use three steps similar to those we did earlier for populating the Dimension using the SQL Steps.

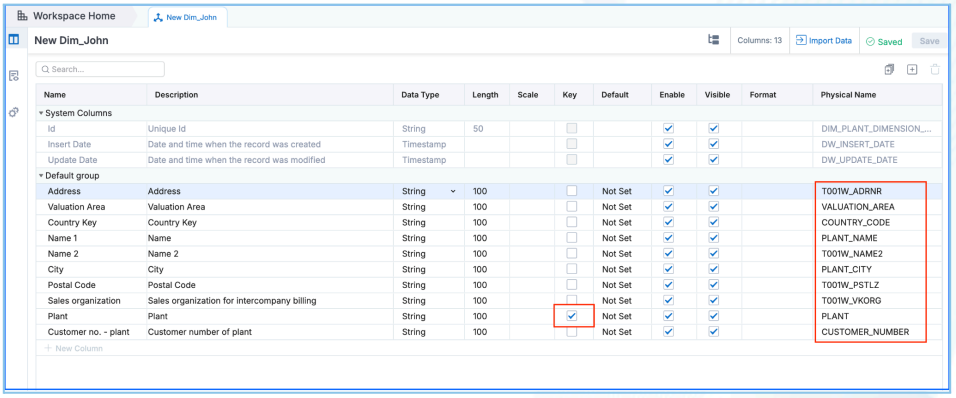
The steps for creating a Dimension remain the same. Please refer to the above steps when creating a new dimension.

Create a New Dimension

1. Create a new dimension, **"New Dim_<Username>**," using the **T001W_<Username>** dataset.

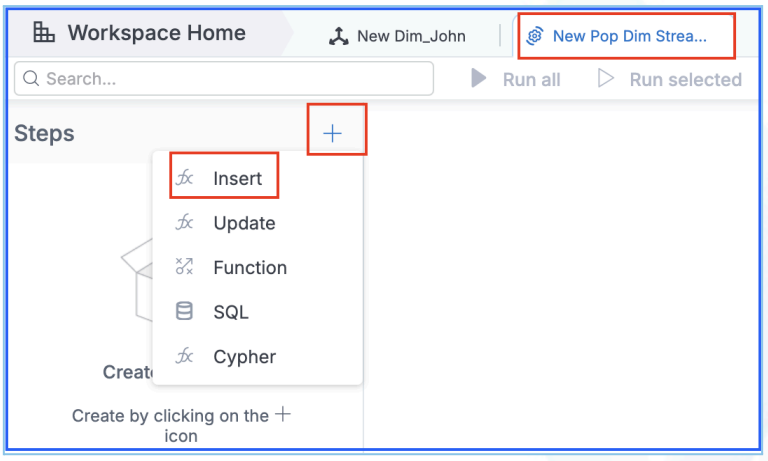


2. Rename the columns as done previously, then enable the **Key** option for **PLANT** and click **Save**.

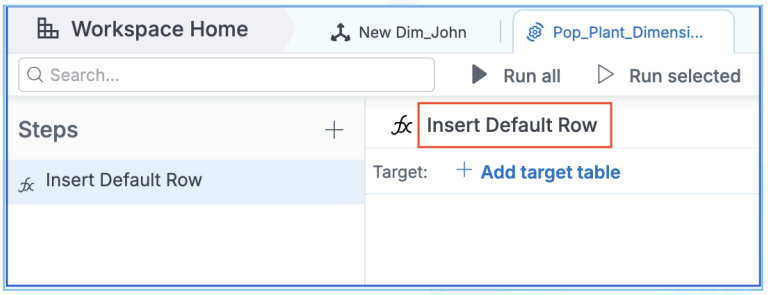


Create a New Stream to Populate the New Dimension

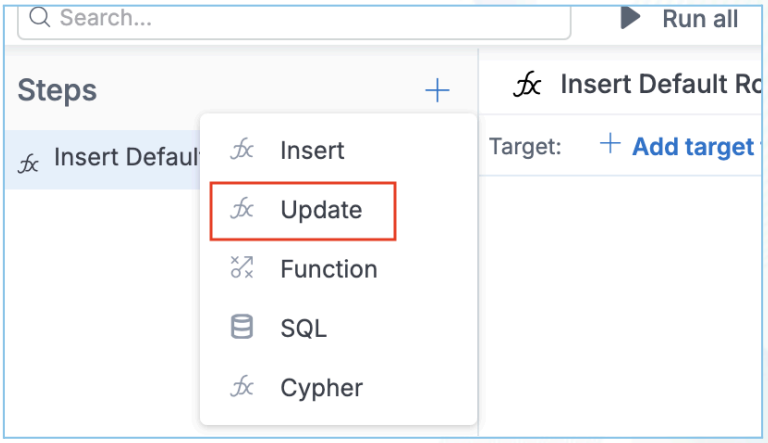
1. Follow the earlier steps and **create a New Stream** to populate the Dimension. Provide the name – **New Pop Dim Stream <Username>**
2. Open the stream and click on the + icon, and select **Insert** Step.



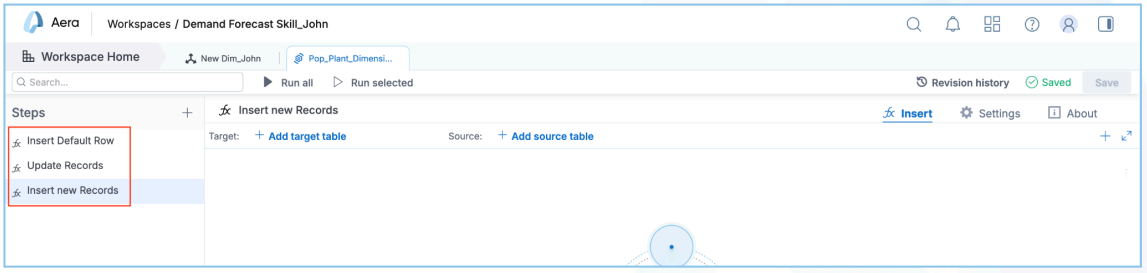
3. Rename the step as **Insert Default Row**.



4. Next, add an **Update step** and rename the step as **Update Records.**

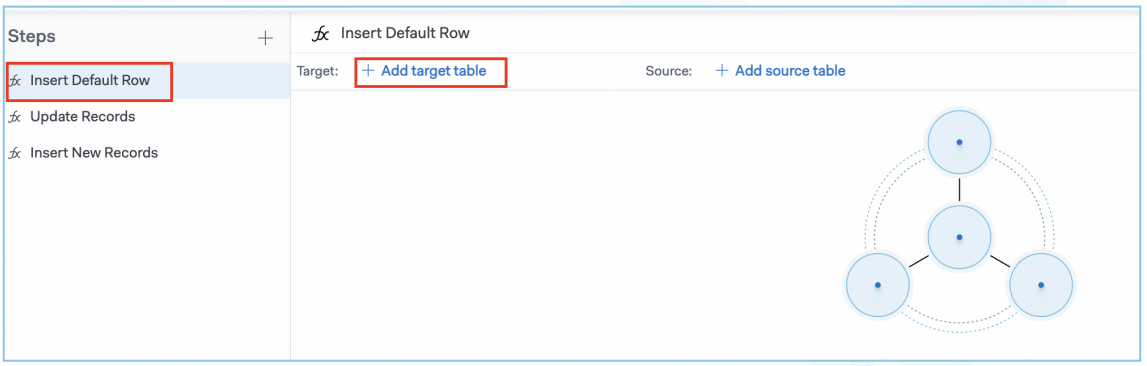


5. Add and name a new **Insert Step** as "**Insert New Records,**" then **save** the Stream.

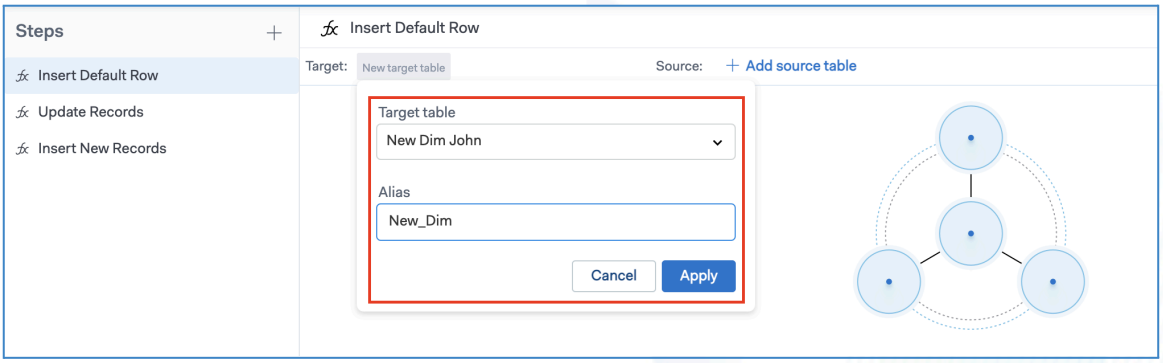


Insert Step

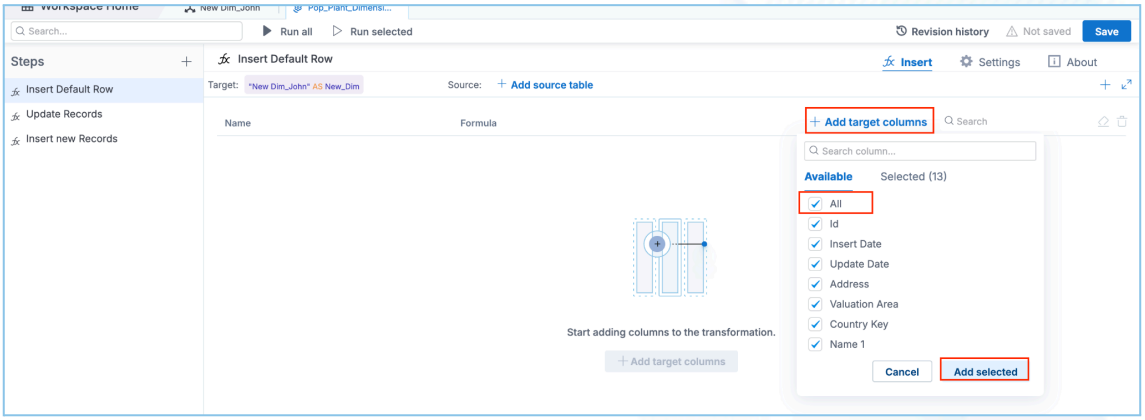
1. Navigate to the initial step, **Insert Default Row**, and select **"Add Target Table."**



2. Select the **New Dim <Username>** from the dropdown and change the Alias to **New_Dim**. Click **Apply**



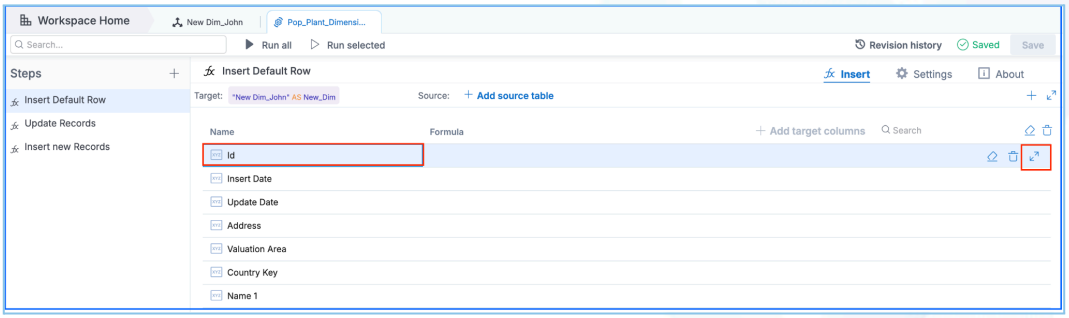
3. Next, click "**Add Target Columns**," choose "**Select All**" from the new window, and then click "**Add Selected**".



In this step, we insert a default row. Therefore, you will add the default values manually.

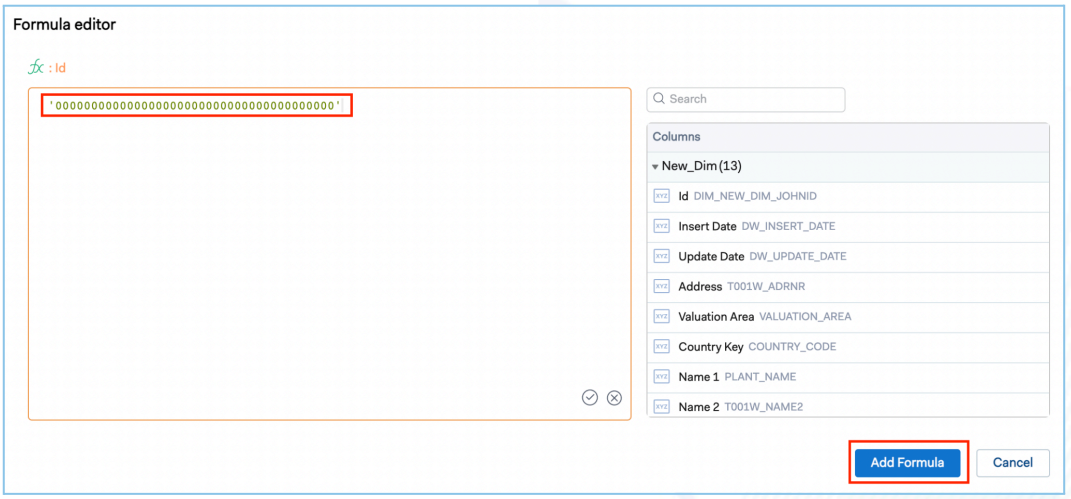
In the Formula Editor, every constant value defined should be enclosed within '**Single Quotes**' and should NOT be validated.

4. Click on the **Formula** option for the **ID** column and then click on **Expand** Icon.



5. Provide a 40 Digit Default value for **ID** column **0000000000000000000000000000000000000000**.

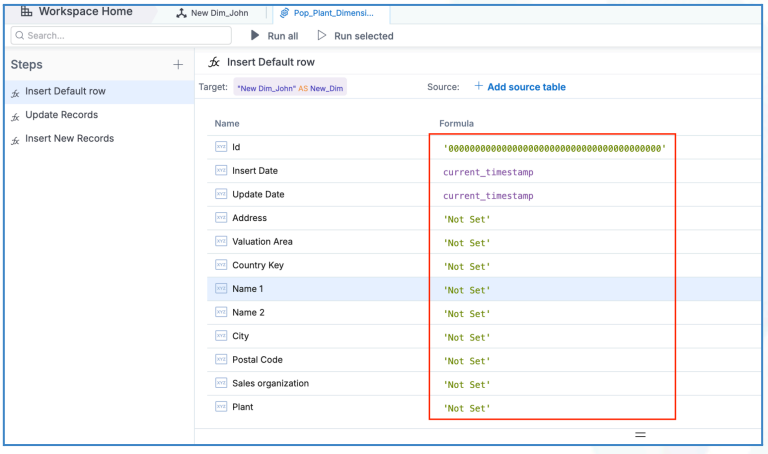
The ID field usually consists of a 40-digit hash value generated from a unique grain.



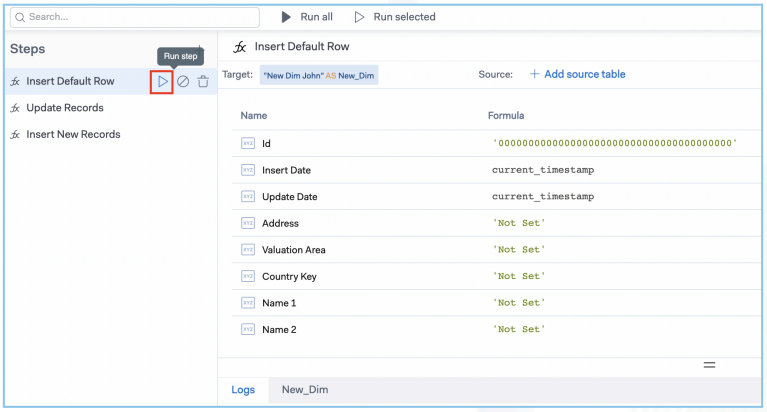
6. Click **Add Formula**. Here, we do not validate the formula as we are setting default constant values.

7. Similarly, manually enter the default values for all the remaining columns, as shown below, then click **Save.**

```
Insert Date -> current_timestamp
Update Date -> current_timestamp
Address -> 'Not Set'
Valuation Area -> 'Not Set'
Name 2 -> 'Not Set'
Name 1 -> 'Not Set'
City -> 'Not Set'
Country Key -> 'Not Set'
Postal Code -> 'Not Set'
Sales Organization -> 'Not Set'
Plant -> 'Not Set'
Customer no -> '0'
```

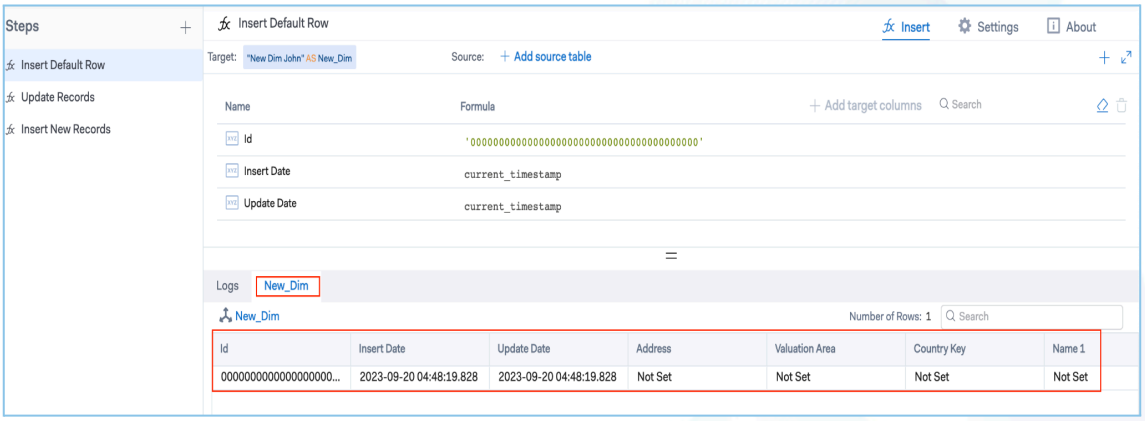


9. Hover the cursor over the **Insert Default Row** step and click on **Run Step.**



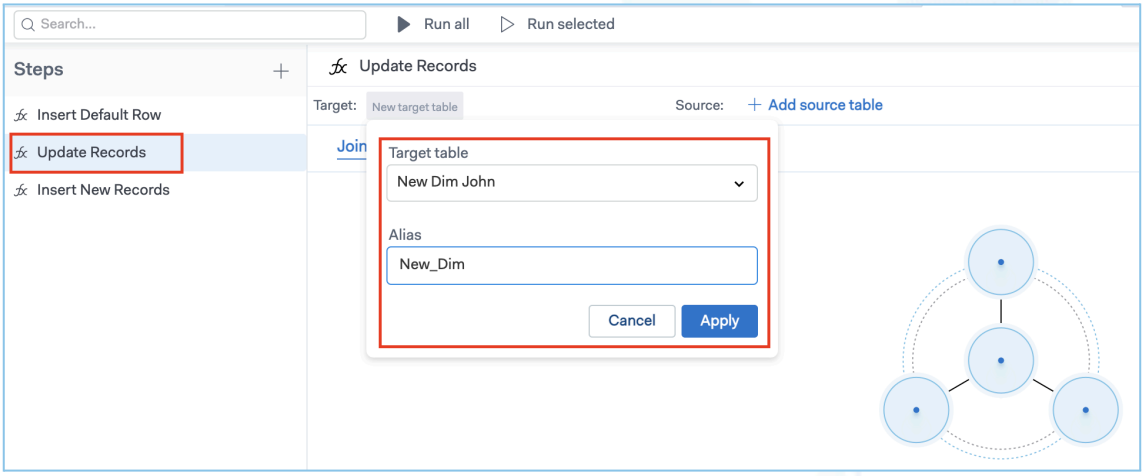
11. Make Sure the Step has run successfully.

12. Click on the **New_Dim** tab to see the default row added.

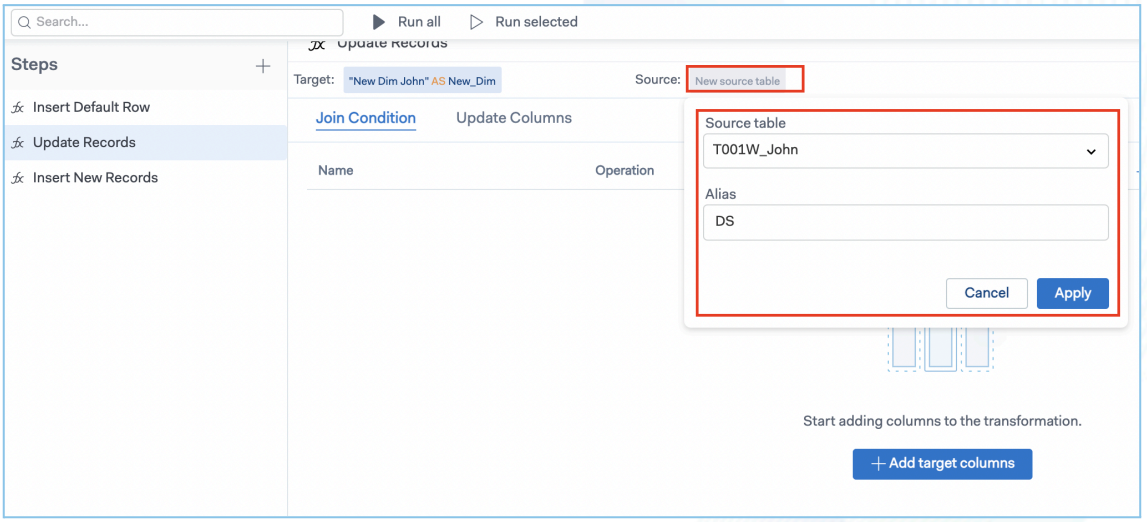


Update Step

1. Click the **Update Records** step, add the target table **New Dim <Username>**, and set the Alias to **New_Dim.**

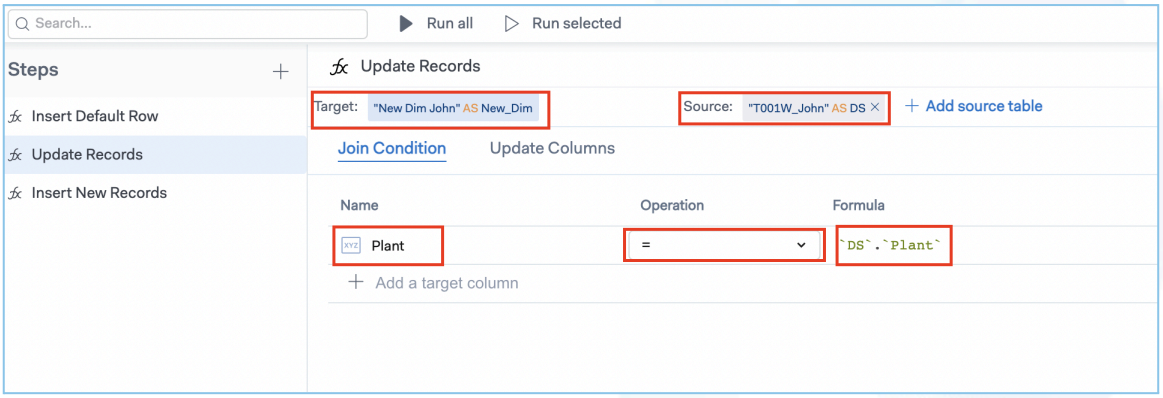


2. Next, click on **Add source table**. In the pop-up, select the source as **T001W_<Username>** and change Alias to **DS.**

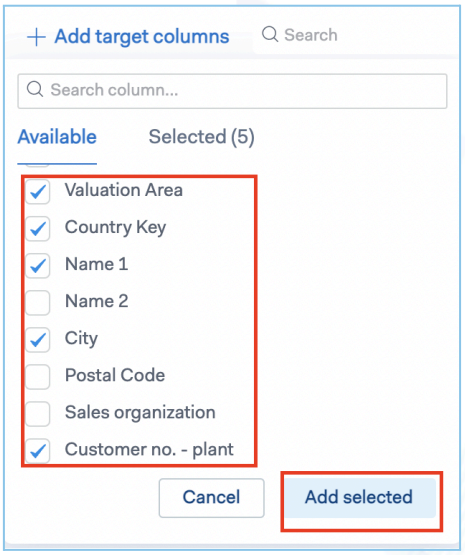


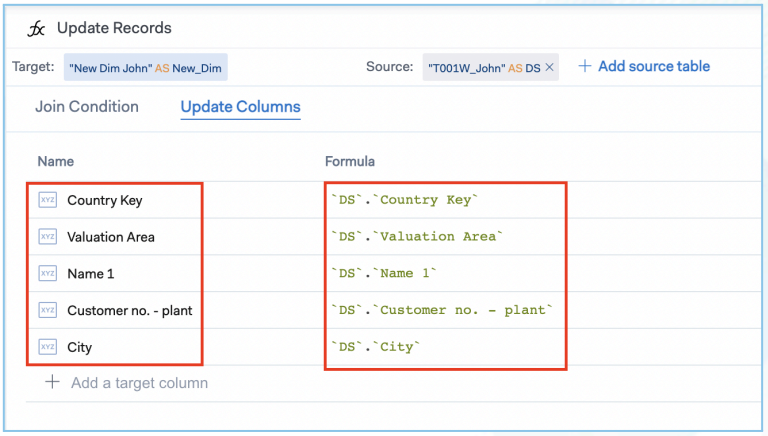
You will see that a Join condition is already recommended based on the **Natural Keys** you selected earlier while creating the Dataset and the Dimension. You had enabled the Key option for the **Plant** Field.

This is just a recommendation for a Join clause. You can change it based on your business requirements. In this exercise, we will use the same join clause.

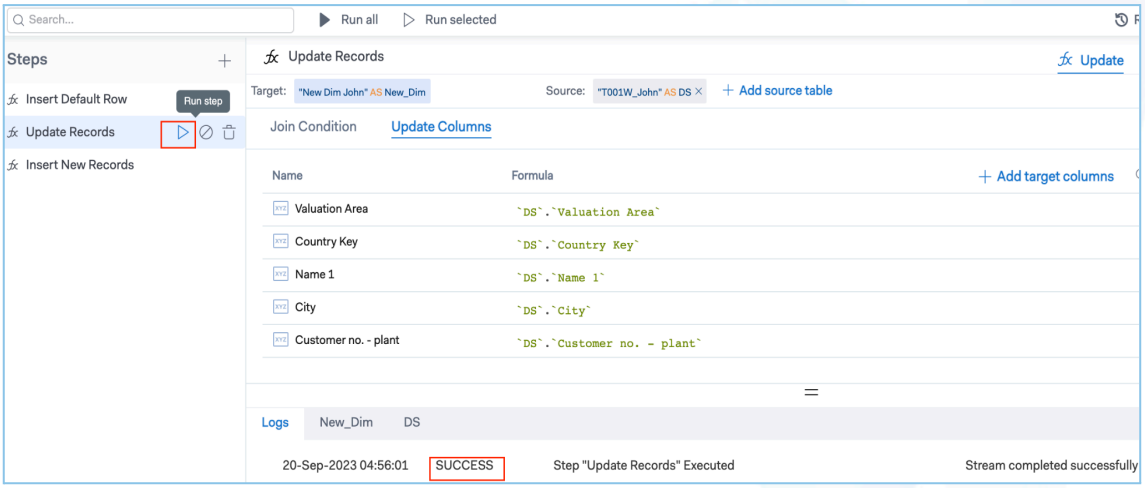


3. Next, click on the **Update columns** tab and then click **Add target columns.**
4. Select the **5 fields** we would like to update in our Plant Dimension whenever the source file changes, and then click **Add Selected.**





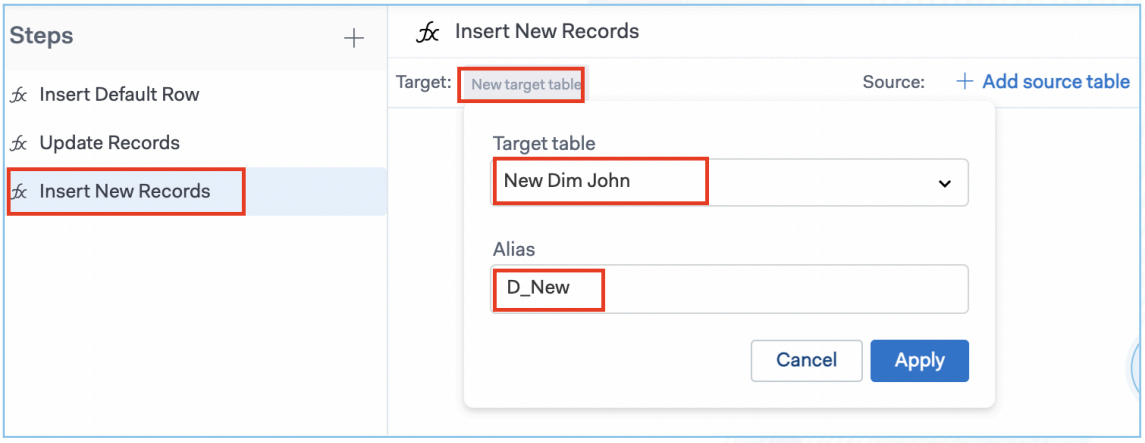
5. Click **Save** and run the **Update Records** Step.



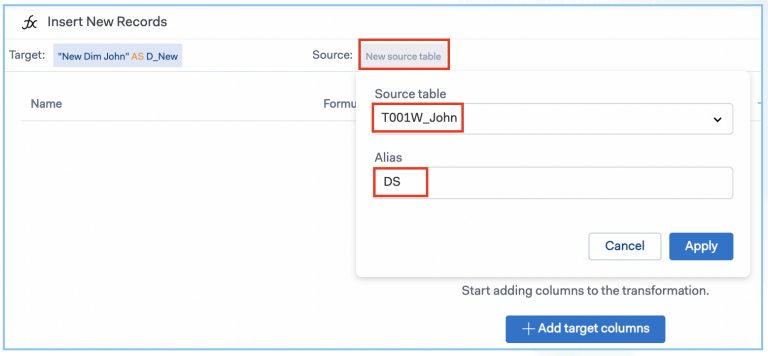
This **Update step** will update the above-mentioned columns in the New Dimension table only if corresponding changes are made to the source file. In this exercise, we have only added a single default row so far.
Next, we will add new records to this New Dimension table.

Insert New Records

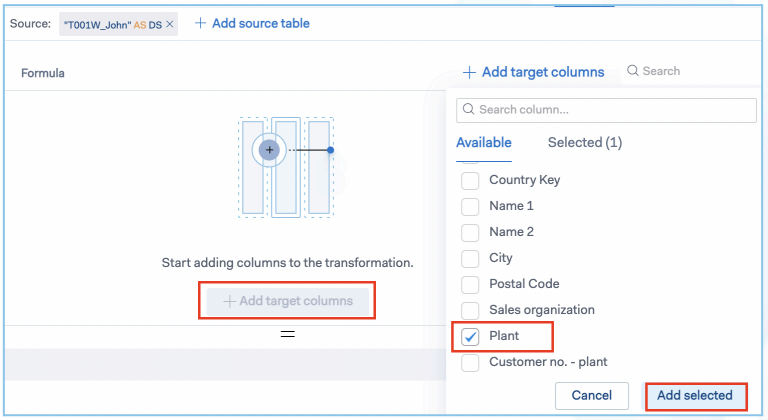
1. Follow the same step as outlined in the Update Records step and set the Alias as **D_New.**



2. Now add the source table by selecting the dataset **T001W_<Username>** and setting the Alias to **DS,** then click **Apply**.

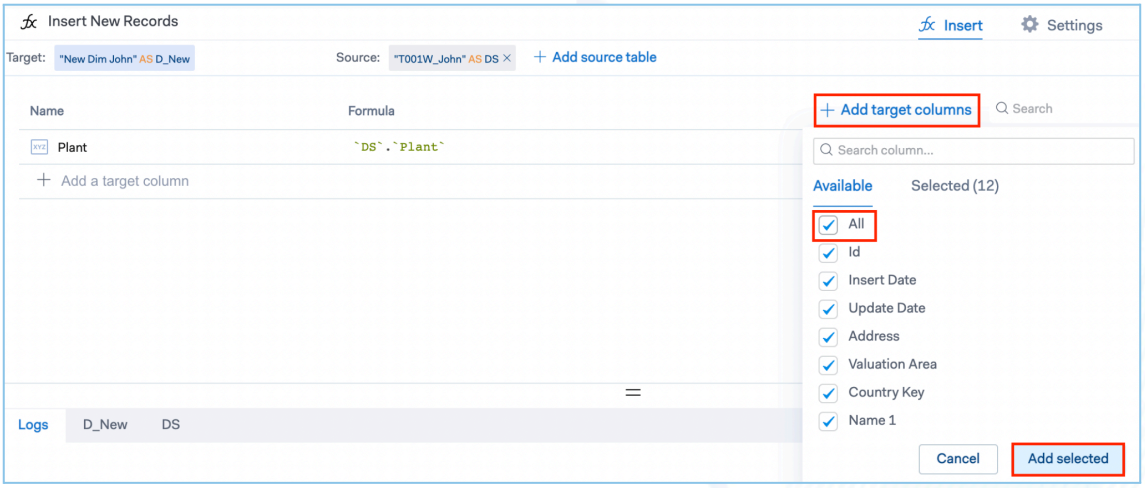


3. Click on **Add target columns**, select only the **Plant** field and click on **Add selected**.



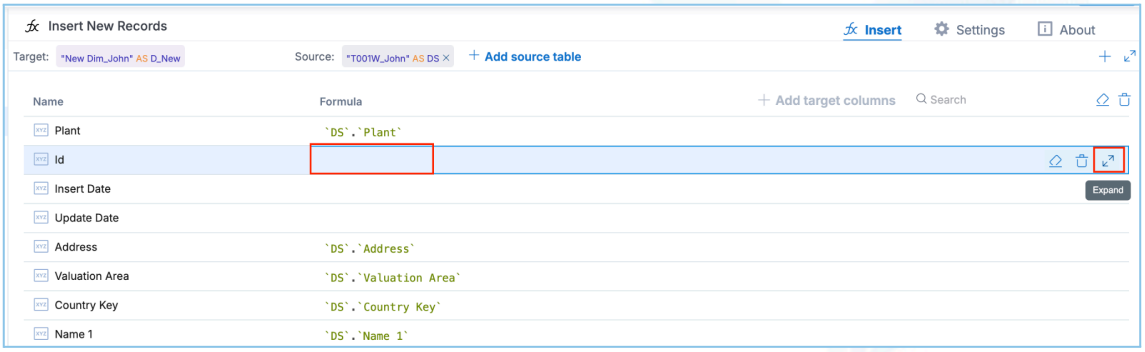
When creating a Plant Dimension, you would first need to create distinct plant entries and then populate the rest of the fields corresponding to these distinct Plant entries.

4. Next, click on **Add target columns** again, select **All**, and click **Add Selected.**

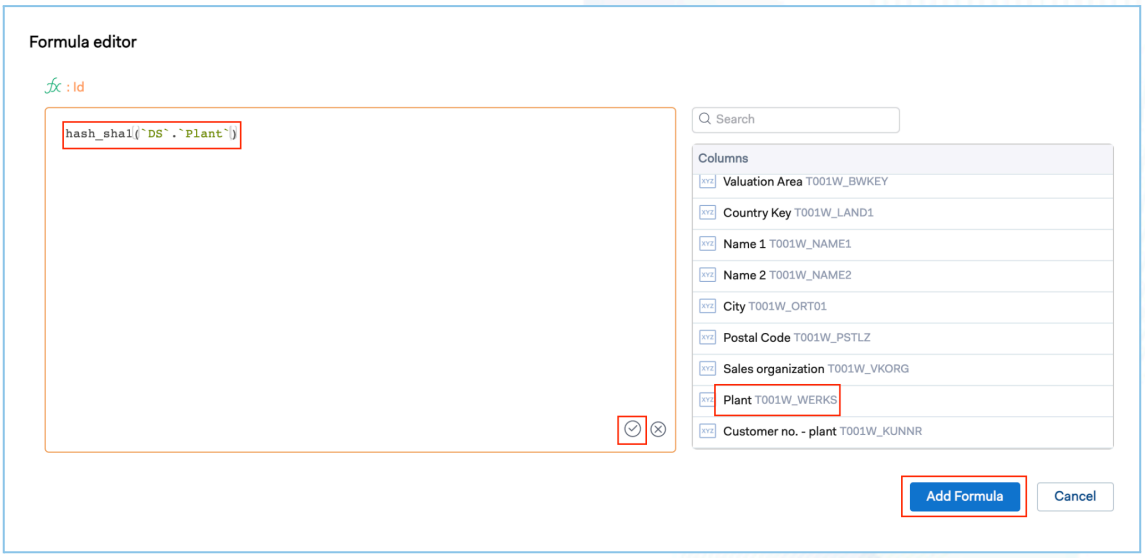


The target columns are auto-mapped to the source columns wherever a match is found. For any new columns in the Dimension that do not have a corresponding match in the source table, the values are to be provided manually, or a formula needs to be added to derive one.

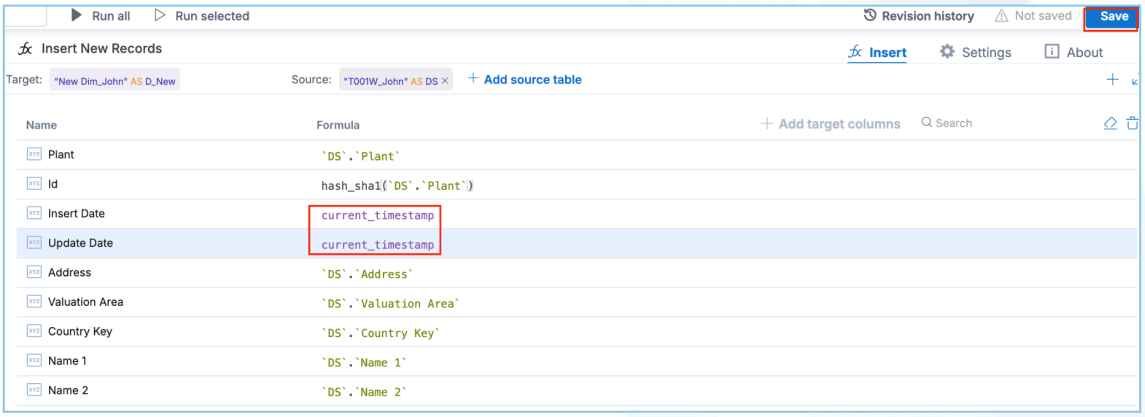
5. Next, click on the formula option next to the Id field and then click on **Expand.**



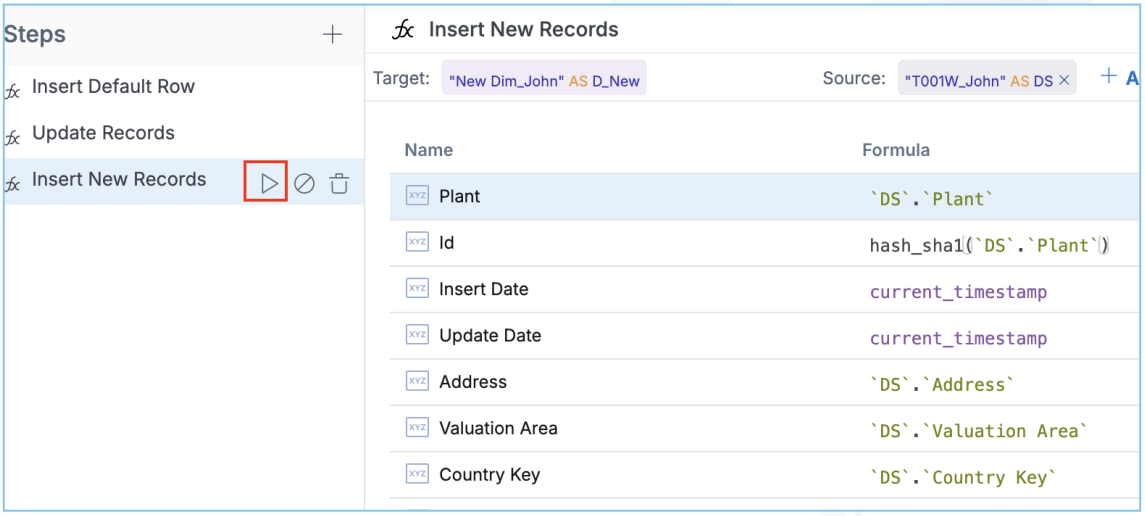
6. Add the formula `hash_sha1(`DS`.`Plant`)` in the ID field to create unique 40-digit Plant IDs, selecting the **Plant** field from the **DS** table.



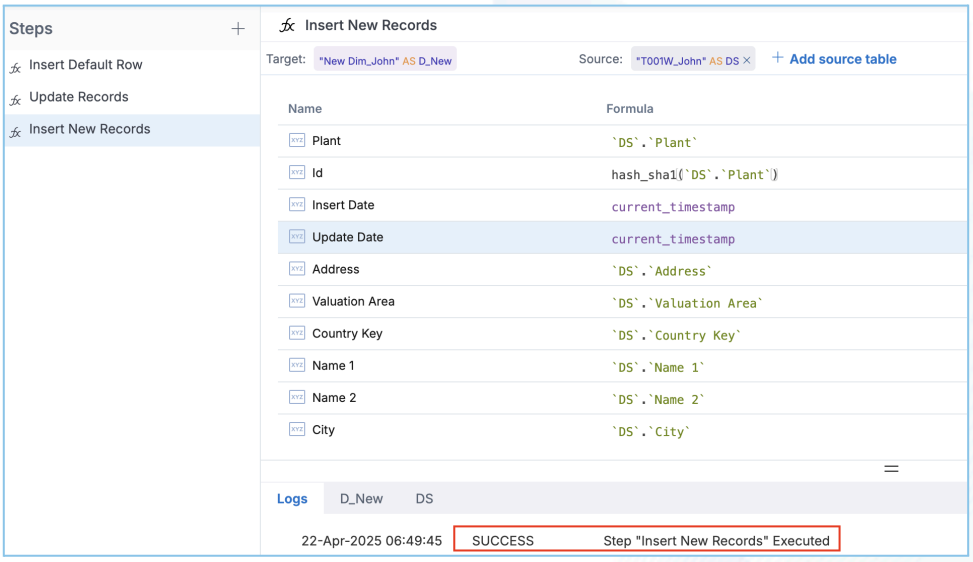
7. Click the check mark to **validate** the formula, then click **Add Formula.**
8. Next, add the formula **current_timestamp** for the **Insert Date** and the **Update Date** fields as shown.



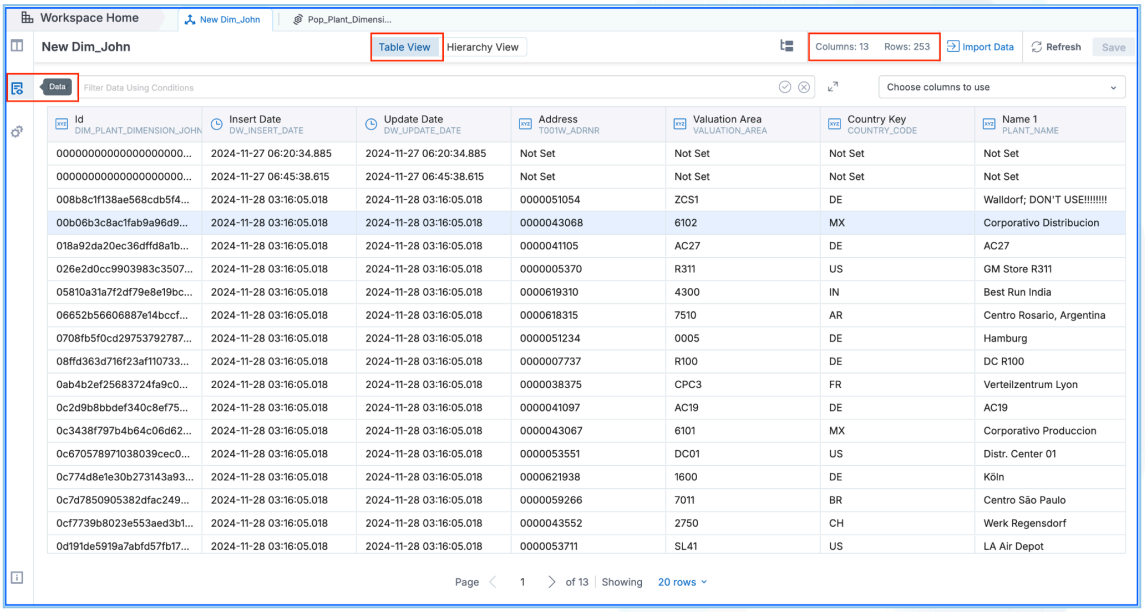
9. Ensure the **Plant** field is the first column, click **Save**, and run the **Insert New Records** step.



10. The Step gets completed successfully.



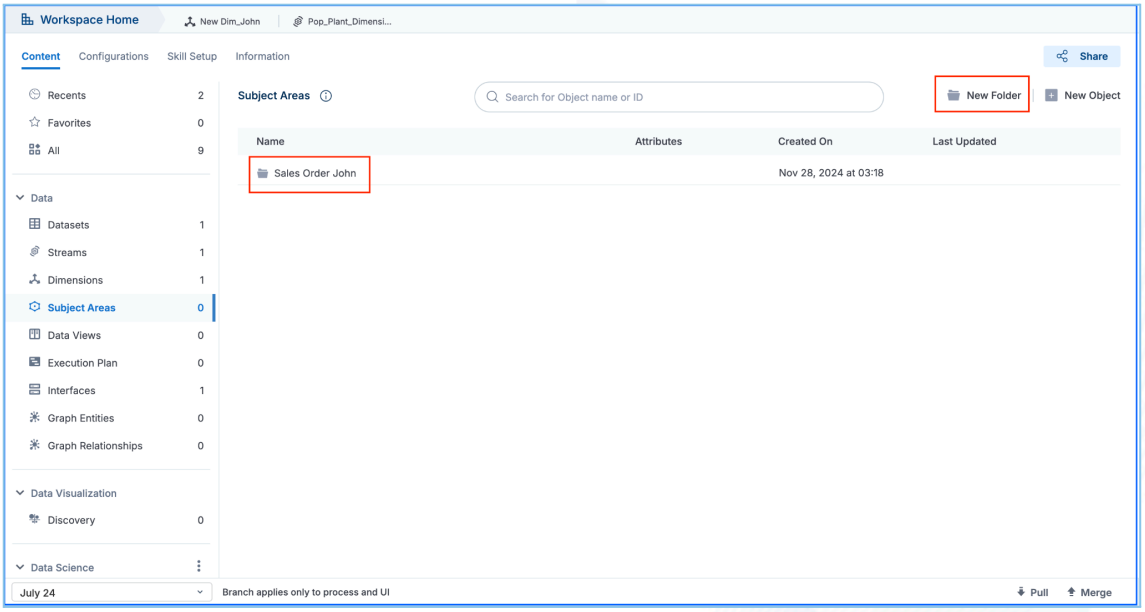
11. You can review the Data populated by going to your Dimension **New Dim <Username>** and clicking on the **Table view** under the **Data** Tab.



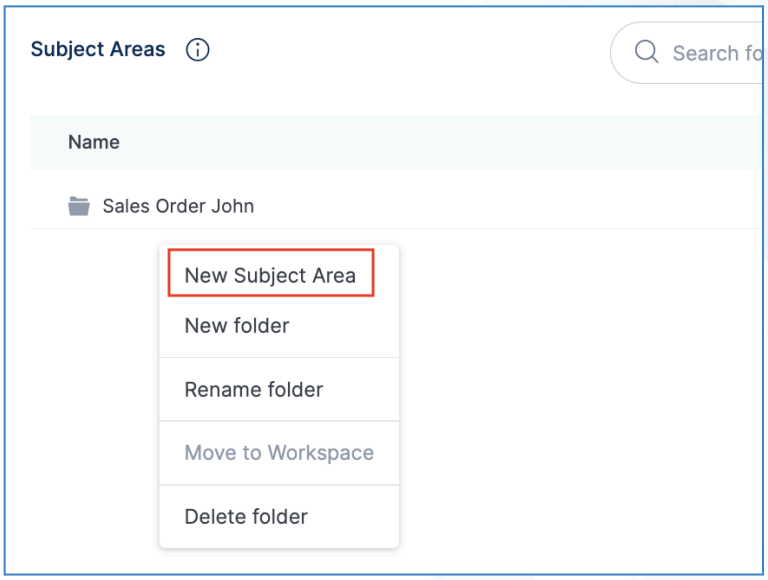
12. Close all the open tabs and go to the **Workspace home.**

Create Subject Area

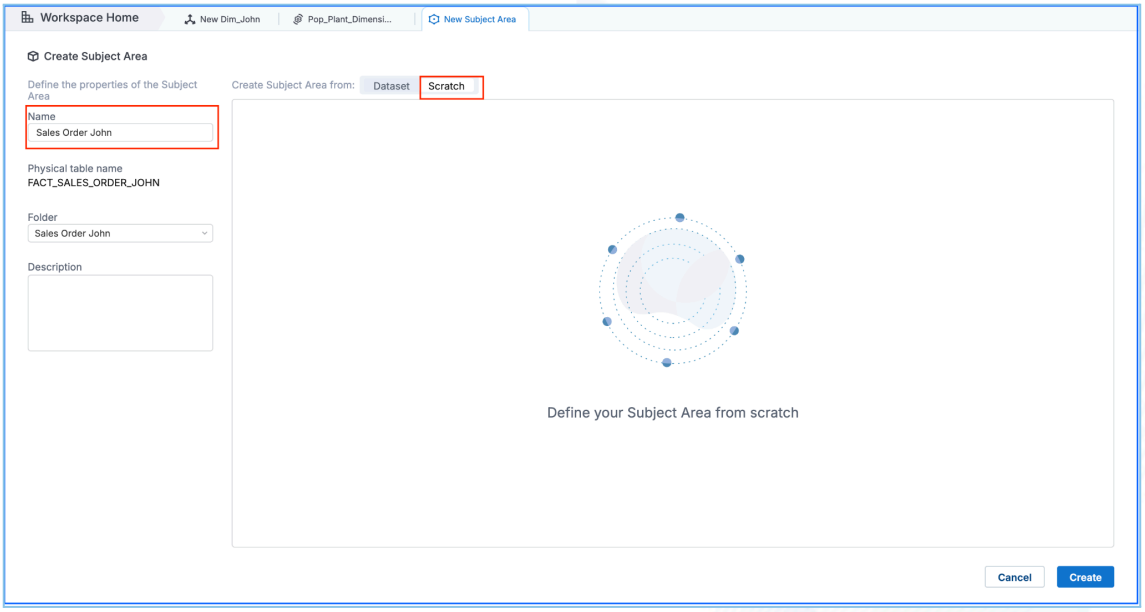
1. Navigate to the **Subject Areas** section under **Content** and tap the '**New Folder**' icon to create a new folder.
2. Provide the name **Sales Order <Username>**.



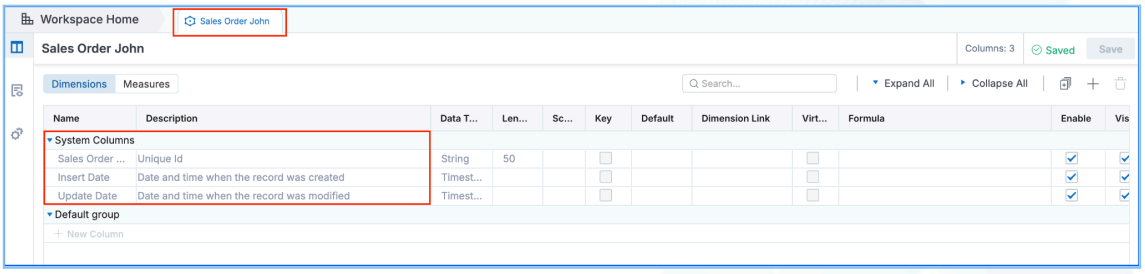
3. Next, right-click the folder to create a **Subject Area.**



4. In the new window, enter "**Sales Order <Username>** " as the name for the Subject Area, navigate to the **Scratch** tab, and click **"Create**."



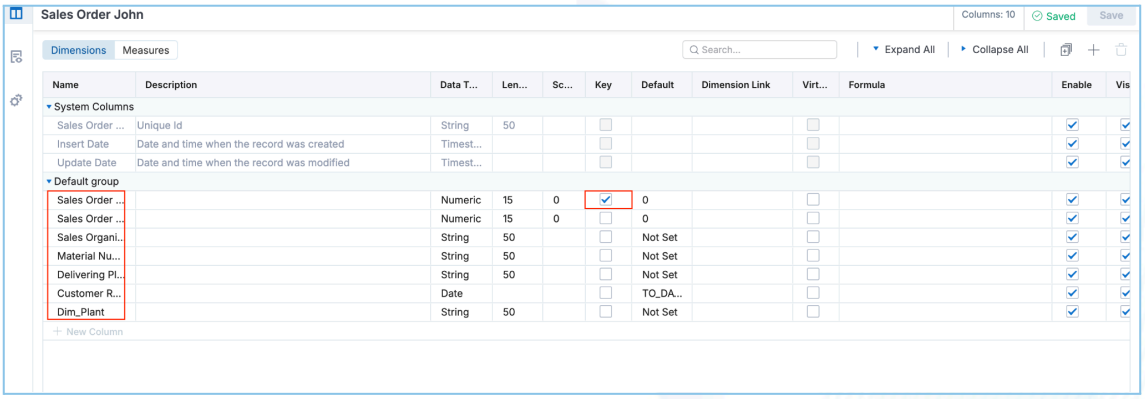
5. The subject area opens in a new tab and contains default system columns in the Dimension tab.



6. Add the following columns as **Dimension** in the **Default group** by clicking on the **+ New Column**

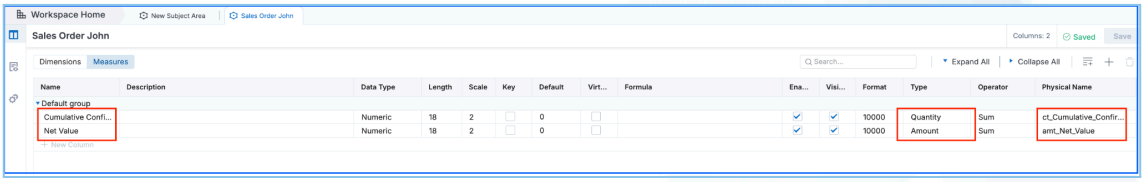
Column Name	Data Type	Length	Physical name
Sales Order Number	Numeric	15	dd_Sales_Order_Number
Sales Order Item Number	Numeric	15	dd_Sales_Order_Item_Number
Sales Organization	String	50	dd_Sales_Organization
Material Number	String	50	dd_Material_Number
Delivering Plant	String	50	dd_Delivering_Plant
Customer Requested Date	Date		dd_Customer_Requested_Date
Dim_Plant	String	50	dd_Dim_Plant

5. Enable the **Key** option for **Sales Order Number** and click **Save.**



6. Add the following columns as **Measures** and click **Save.**

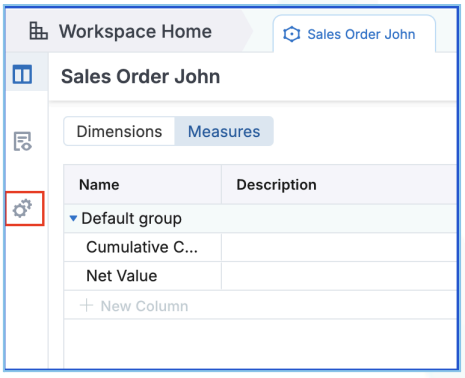
Column Name	Data Type	Length	Operator	Physical name	Type
Cumulative Confirmed Qty UoM	Numeric	18	Sum	CT_CUMULATIVE_CONFIRMED_QTY_UOM	Quantity
Net Value	Numeric	18	Sum	AMT_NET_VALUE	Amount



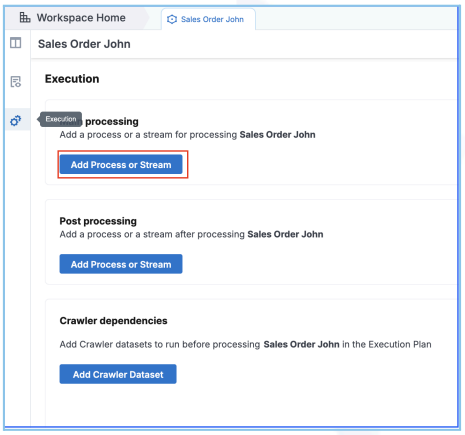
Add Stream to Populate Subject Area

Populate Subject Area using SQL Step

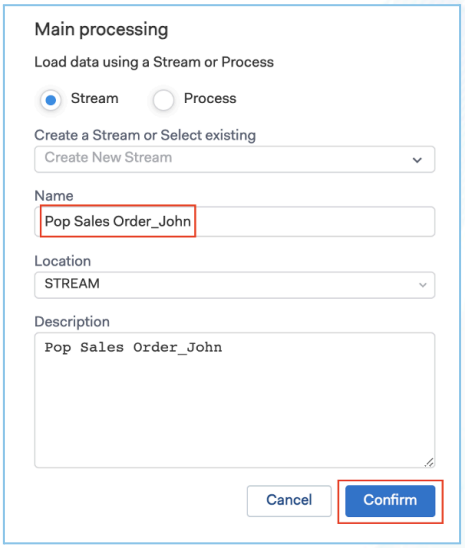
1. Click the Executions tab on the **Sales Order <Username>** window.



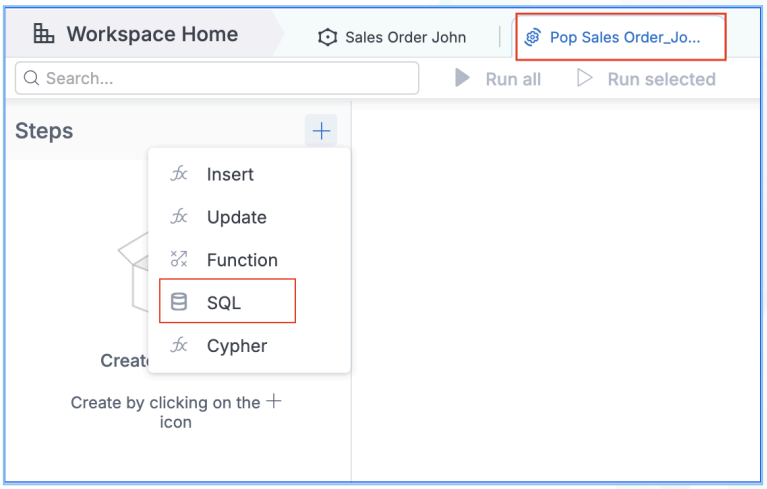
2. Click the **Add Process or Stream** in the Main Processing section.



3. To load the data, select a **Stream** from the dropdown, choose **"Create New Stream,"** name it **"Pop Sales Order_<Username>**," and click **"Confirm**."



4. On the new **Pop Sales Order_<Username>** tab, click the **+** icon and select **SQL**.

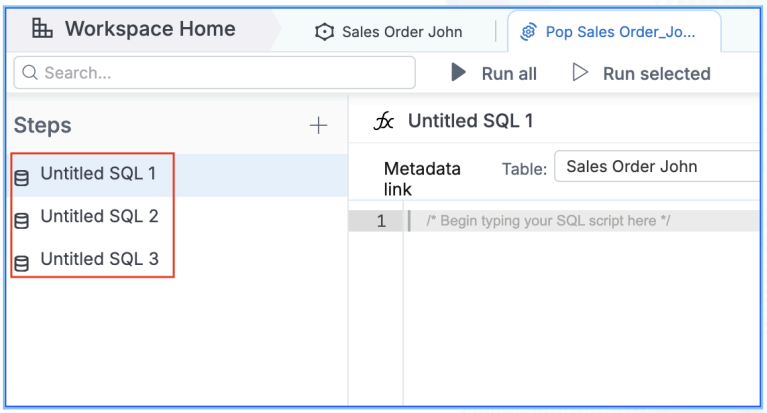


Create three individual SQL steps for each significant operation:

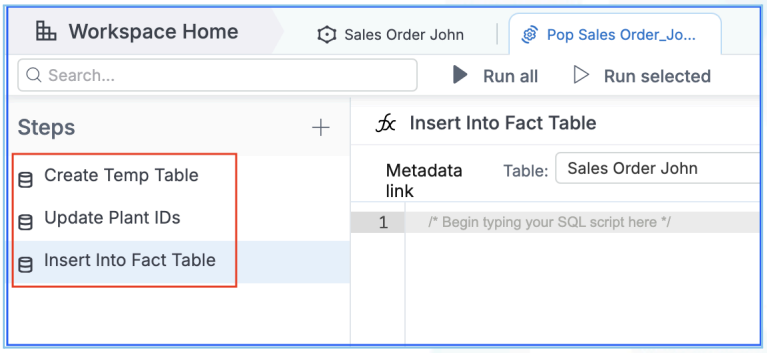
- **Create Temp table:** we create a temporary table and populate it with required fields

from VBAK – Sales Header data and VBAP- Sales Item data for every matching Sales order no for a given period.

- **Update Plant IDs:** In this step, we link and update the Plant IDs of our Subject Area with those of an external Dimension Plant Demo wherever the Plant code matches.
- **Insert into Fact table:** Once all the New records are inserted and updated, we move the data from the temporary table to the Fact Table or Subject Area.



6. Double-click to rename it as "**Create Temp Table**," "**Update Plant IDs,**" and "**Insert into Fact Table,**" respectively.

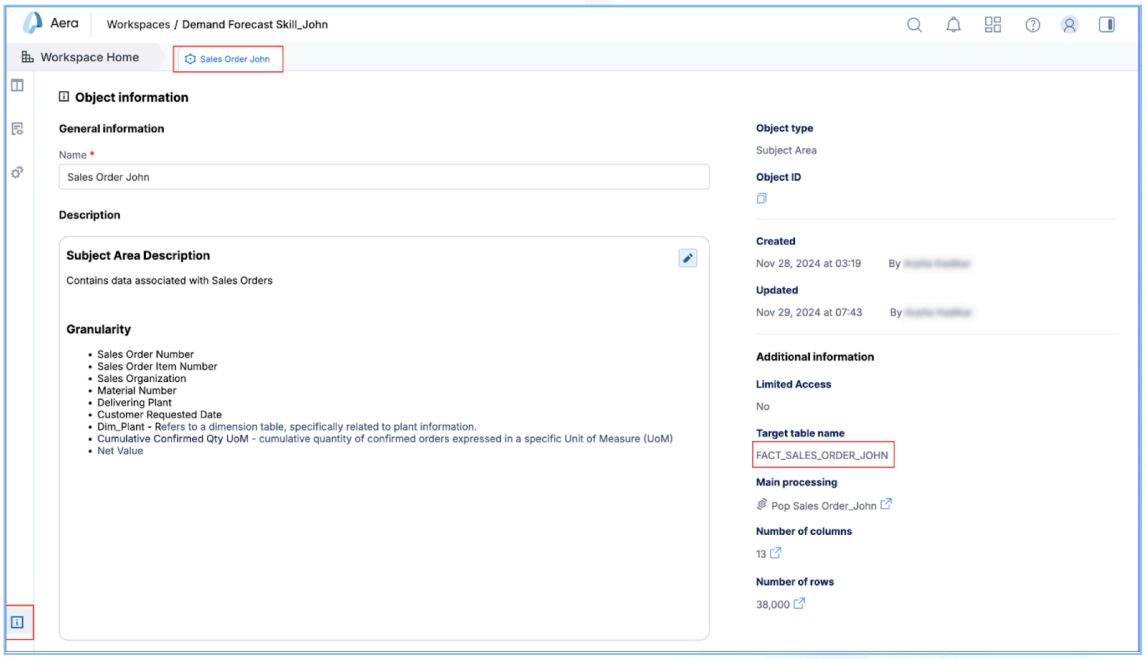


7. Click the "**Create Temp Table**" step, then copy and paste the query into the right panel.

```
DROP TABLE IF EXISTS TMP_your_subjarea_table;
CREATE TABLE TMP_your_subjarea_table LIKE your_subjarea_table;
MERGE INTO TMP_your_subjarea_table f
USING (select distinct
ifnull(CAST( UPPER( HASH_SHA1( UPPER( TRIM( LEADING '0' FROM IFNULL(VBAK_VBELN, ' ')) ||
'|^|' || IFNULL(VBAP_POSNR,' ') || '|^|' || 'SAP ECC EURO' ))) AS VARCHAR(40) ),
'00000000000000000000000000000000000000000') as ID,
current_timestamp as DW_INSERT_DATE,
current_timestamp as DW_UPDATE_DATE,
ifnull(VBAK_VBELN,0) as dd_Sales_Order_Number,
ifnull(VBAP_POSNR,0) as dd_Sales_Order_Item_Number,
ifnull(VBAK_VKORG,'Not Set') as dd_Sales_Organization,
ifnull(VBAK_VTWEG,'Not Set') as dd_Distribution_Channel,
ifnull(VBAP_MATNR,'Not Set') as dd_Material_Number,
```

```
ifnull(VBAP_WERKS,'Not Set') as dd_Delivering_Plant,
ifnull(VBAP_VSTEL,'Not Set') as dd_Shipping_Point,
ifnull(to_date(VBAK_VDATU,'MM/DD/YYYY'),'1001-01-01') as dd_Customer_Requested_Date,
ifnull(VBAP_KBMENG,0) as ct_CUMULATIVE_CONFIRMED_QTY_UOM,
ifnull(VBAP_NETWR,0) as amt_NET_VALUE,
'00000000000000000000000000000000000000' AS DD_dim_Plant
from VBAK_CSV1 vbak inner join VBAP_CSV1 vbap on vbak.VBAK_VBELN = VBAP.VBAP_VBELN
WHERE ifnull(to_date(VBAK_VDATU,'MM/DD/YYYY'),'1001-01-01') <= CURRENT_DATE) b
on b.ID=f.your_subjarea_tableID
WHEN MATCHED THEN
UPDATE set f.dd_Sales_Order_Number = b.dd_Sales_Order_Number,
f.dd_Sales_Order_Item_Number = b.dd_Sales_Order_Item_Number,
f.dd_Sales_Organization = b.dd_Sales_Organization,
f.dd_Material_Number = b.dd_Material_Number,
f.dd_Delivering_Plant = b.dd_Delivering_Plant,
f.dd_Customer_Requested_Date = b.dd_Customer_Requested_Date,
f.ct_CUMULATIVE_CONFIRMED_QTY_UOM = b.ct_CUMULATIVE_CONFIRMED_QTY_UOM,
f.amt_NET_VALUE = b.amt_NET_VALUE,
f.DW_INSERT_DATE = b.DW_INSERT_DATE,
f.DW_UPDATE_DATE = b.DW_UPDATE_DATE
WHEN NOT MATCHED THEN
INSERT (your_subjarea_tableID,
DW_INSERT_DATE,
DW_UPDATE_DATE,
DD_SALES_ORDER_NUMBER,
DD_SALES_ORDER_ITEM_NUMBER,
DD_SALES_ORGANIZATION,
DD_MATERIAL_NUMBER,
DD_DELIVERING_PLANT,
DD_CUSTOMER_REQUESTED_DATE,
CT_CUMULATIVE_CONFIRMED_QTY_UOM,
AMT_NET_VALUE,
DD_DIM_PLANT)
values (b.ID,
b.DW_INSERT_DATE,
b.DW_UPDATE_DATE,
b.DD_SALES_ORDER_NUMBER,
b.DD_SALES_ORDER_ITEM_NUMBER,
b.DD_SALES_ORGANIZATION,
b.DD_MATERIAL_NUMBER,
b.DD_DELIVERING_PLANT,
b.DD_CUSTOMER_REQUESTED_DATE,
b.CT_CUMULATIVE_CONFIRMED_QTY_UOM,
b.AMT_NET_VALUE,
b.DD_DIM_PLANT);
```

8. Replace the above-highlighted text in red with the Subject Area table name.
9. To find the names of your tables, navigate to the **"Sales Order <username>**" and "**T001W_<Username>**" tabs, then click the **"About"** option (identified by an **"i"** icon).



A description is mandatory to render the subject areas "Gen AI" ready. So that the AI Agents understand the context

10. Similarly, add the following query for the **Update Plant IDs** step.

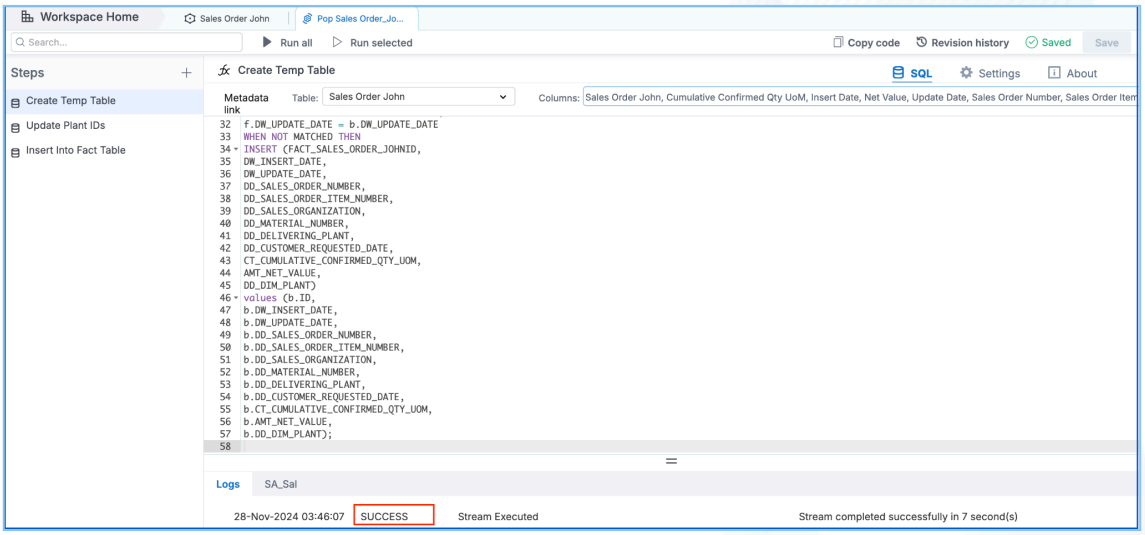
```
update TMP_your_subjarea_table f_s
set f_s.DD_dim_Plant  = d.dim_plant_demoid
from TMP_your_subjarea_table f_s
inner join DIM_PLANT_DEMO d
  on f_s.DD_DELIVERING_PLANT = d.plant_code;
```

11. Similarly, add the below query for the **Insert into Fact Table** step and **Save** the Stream.

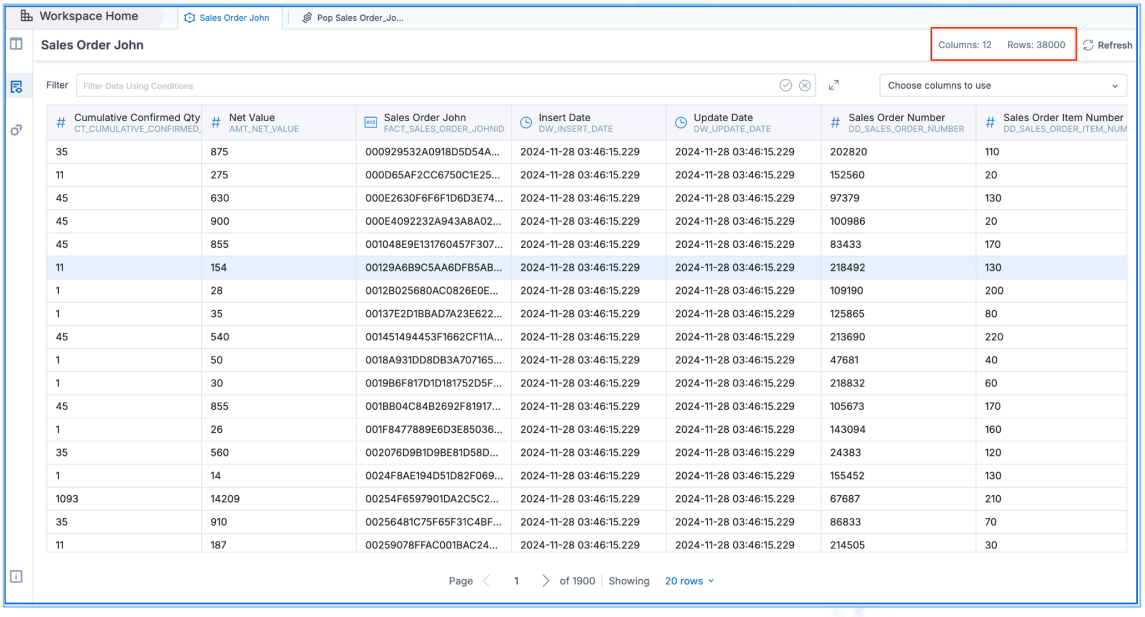
```
truncate table your_subjarea_table;
insert into your_subjarea_table
(your_subjarea_tableID,
DW_INSERT_DATE,
DW_UPDATE_DATE,
DD_SALES_ORDER_NUMBER,
DD_SALES_ORDER_ITEM_NUMBER,
DD_SALES_ORGANIZATION,
DD_MATERIAL_NUMBER,
DD_DELIVERING_PLANT,
DD_CUSTOMER_REQUESTED_DATE,
CT_CUMULATIVE_CONFIRMED_QTY_UOM,
AMT_NET_VALUE,
DD_DIM_PLANT)
select b.your_subjarea_tableID,
b.DW_INSERT_DATE,
b.DW_UPDATE_DATE,
b.DD_SALES_ORDER_NUMBER,
```

```
b.DD_SALES_ORDER_ITEM_NUMBER,
b.DD_SALES_ORGANIZATION,
b.DD_MATERIAL_NUMBER,
b.DD_DELIVERING_PLANT,
b.DD_CUSTOMER_REQUESTED_DATE,
b.CT_CUMULATIVE_CONFIRMED_QTY_UOM,
b.AMT_NET_VALUE,
b.DD_DIM_PLANT
from TMP_your_subjarea_table b;

DROP TABLE IF EXISTS TMP_your_subjarea_table;
```

12. Click **"Run All"** to execute all steps at once, changing the Stream status to **SUCCESS** upon completion.

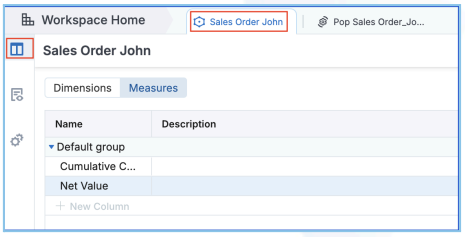


14. Click the Data tab of the **Sales Order Username** to review the populated data.

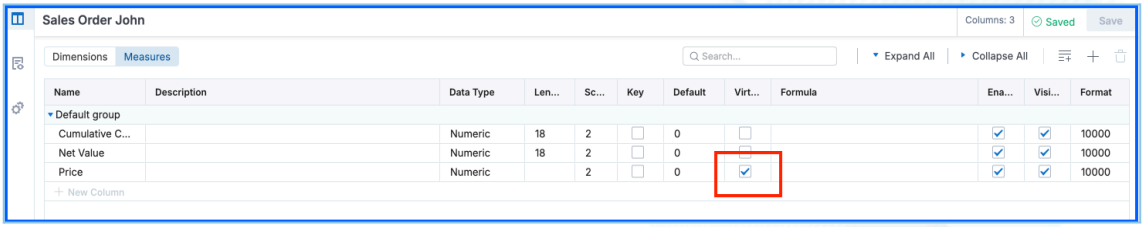


Add a Virtual Measure

1. Navigate to the subject area **Sales Order <username>** and click **Measures**.

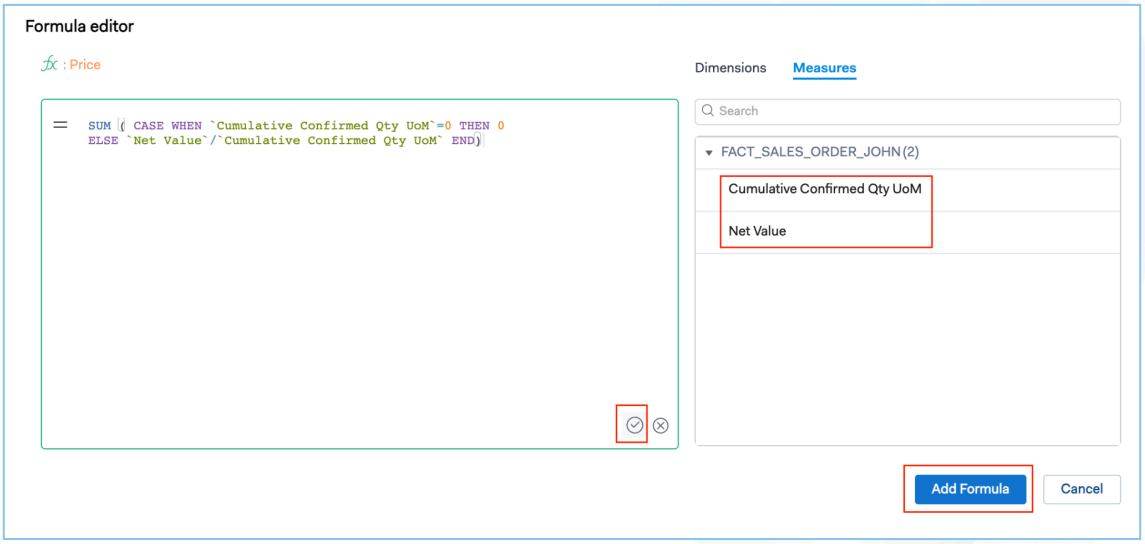


2. Add a new column named **Price** and select it as a virtual column to enable the formula field. Open the formula window by clicking the ||| icon.

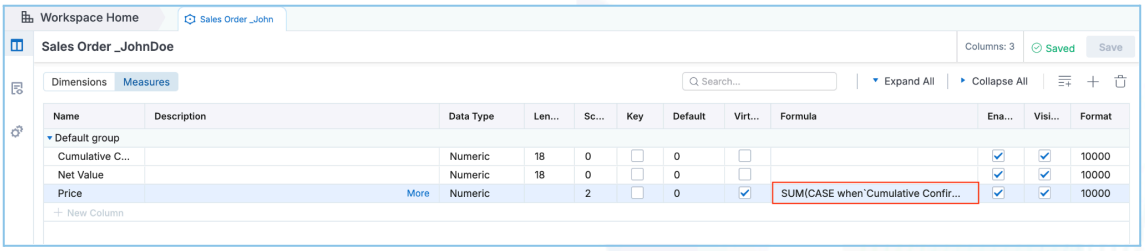


3. Enter the provided formula in the new formula window, **validate** it, and click **Add Formula**.

```
SUM(CASE when`Cumulative Confirmed Qty UoM` =0 Then 0
ELSE `Net Value ` / `Cumulative Confirmed Qty UoM` END)
```

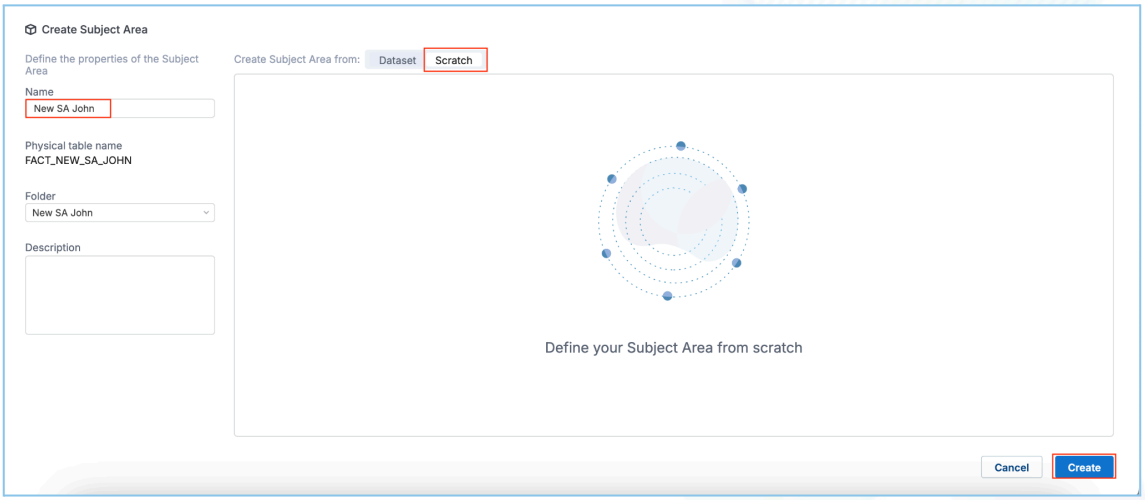


5. The formula is added as shown and saved. The Virtual Measure calculates and appears only during run time.

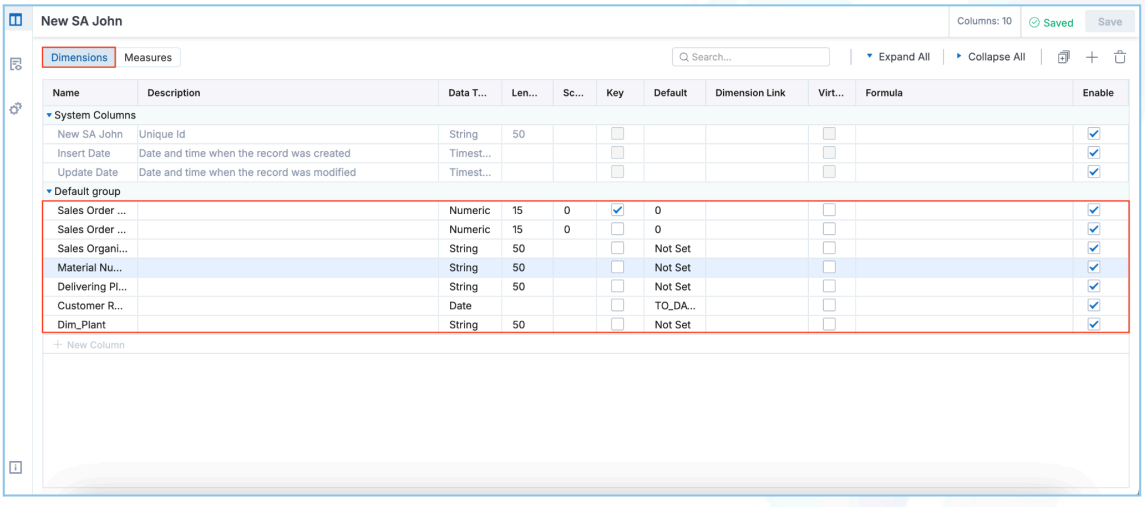


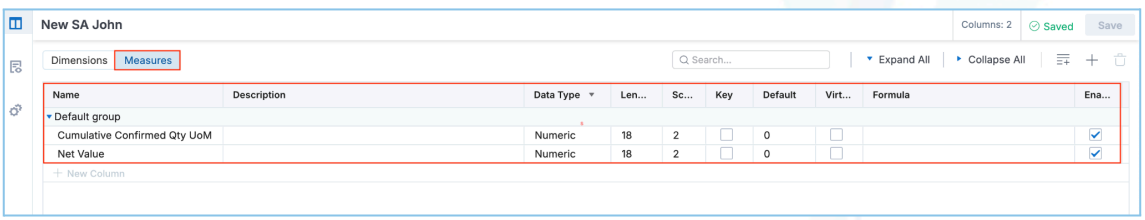
Create a New Subject Area and Populate using the Insert/Update step.

1. Follow the same steps outlined in the **'Create Subject Area'** to create a new folder **named "New SA <Username>** " and a subject area titled **"New SA <Username>**" with the same Dimensions and Measures as created above.

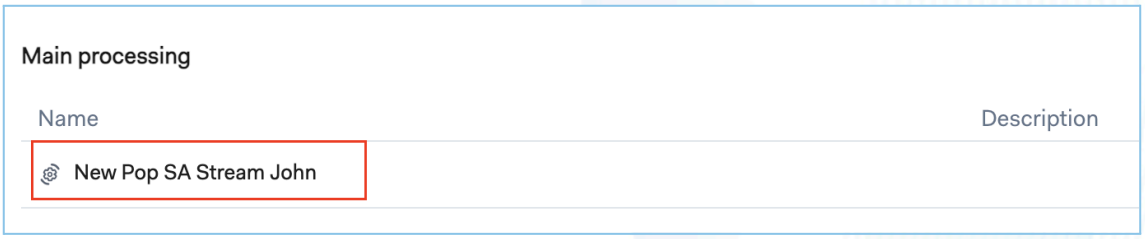


2. Enable the **Key** option for **Sales Order Number.**

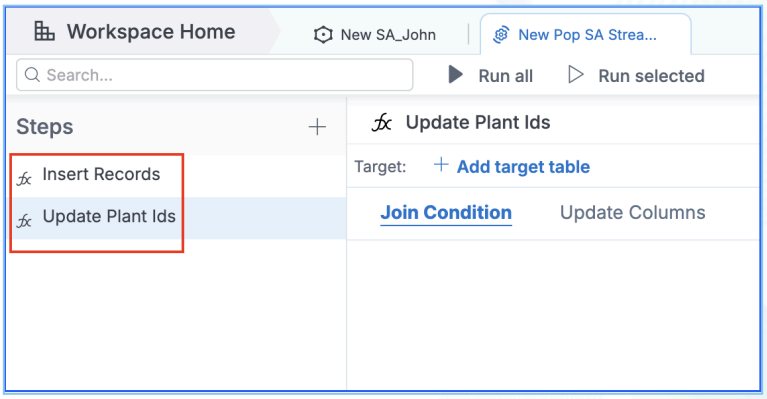




3. Add a Stream to the Main Processing as done earlier and name it **New Pop SA Stream <Username>.**



4. Open the Stream, add the insert step '**Insert Records**' and update step '**Update Plant Ids**.'

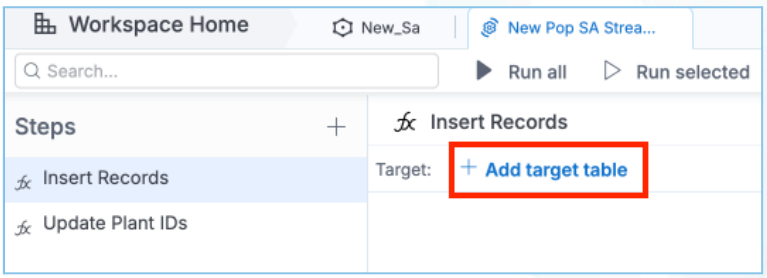


We will use only two steps to populate the New Subject Area.

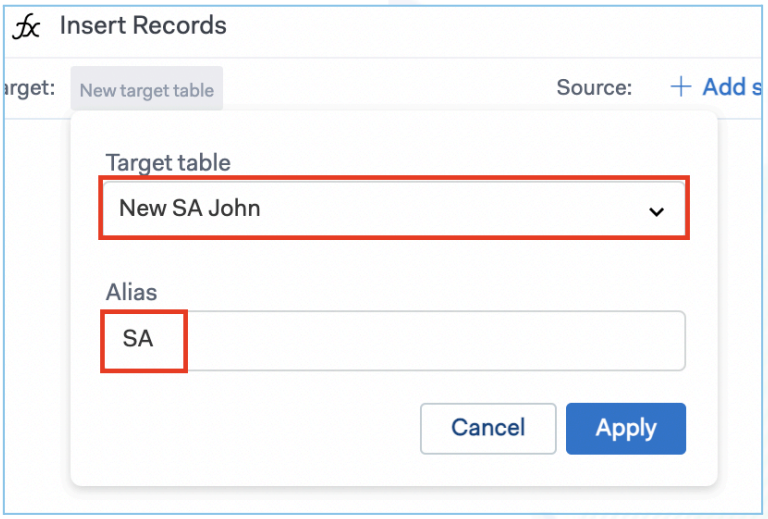
In the **first step**, we will populate the Subject Area fields by joining data from VBAK and VBAP datasets. These two datasets have already been created and are available. These datasets were also used as part of the data population using SQL.
In the **Second Step**, we will link and update the Plant IDs as we did earlier.

Insert Records

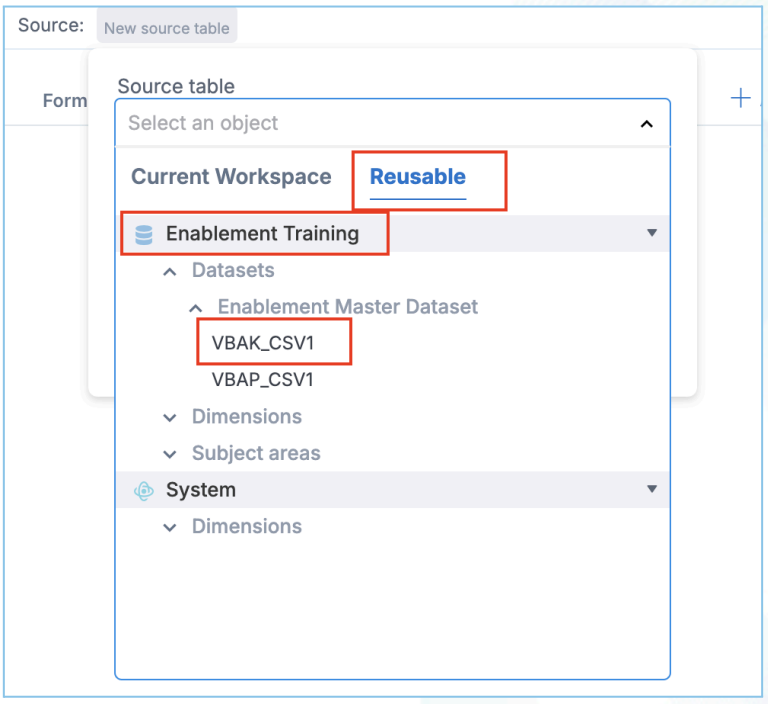
1. Click on the **Insert Records** Step and **Add target table.**



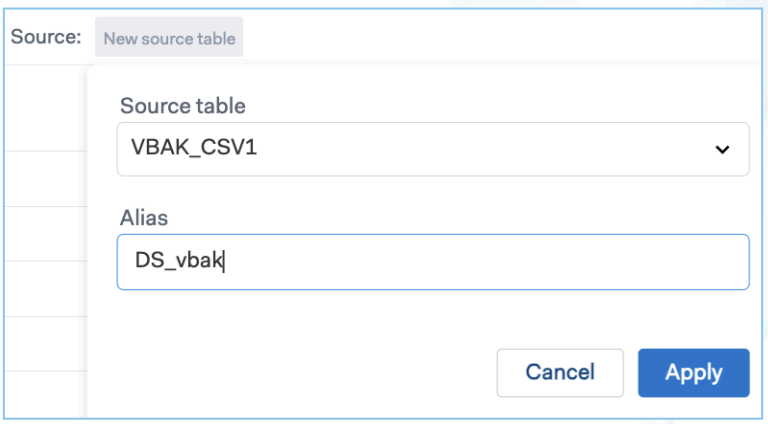
2. Select the **New SA <Username>** from the dropdown and change the Alias to **SA** and click **Apply.**



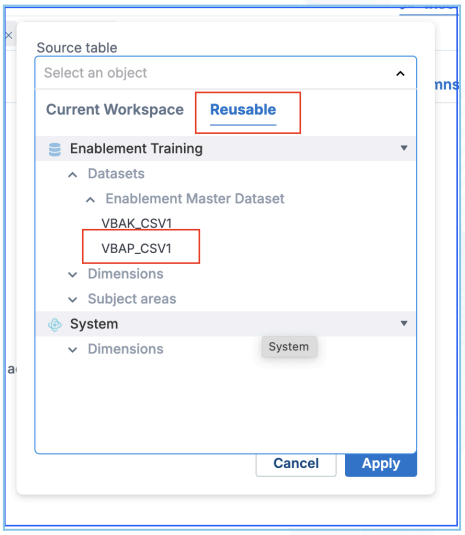
3. Click "**Add Source Table,**" select "**Reusable**" from the drop-down, and navigate to the **VBAK_CSV1** dataset.



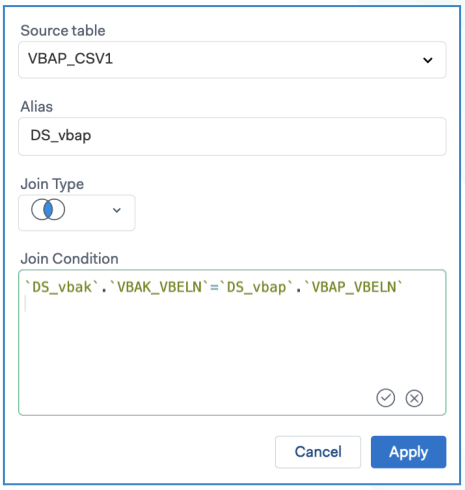
4. Set the Alias as **DS_vbak** and click **Apply.**



5. Select **"Add source table"** again, and choose the **VBAP_CSV1** dataset from the drop-down.

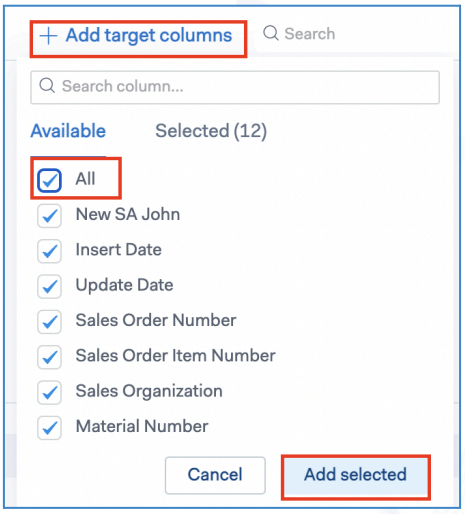


6. Set alias **DS_vbap**, use **Inner Join** with the condition **`DS_vbak`.`VBAK_VBELN`=`DS_vbap`.`VBAP_VBELN`**, validate, then click **Apply**.

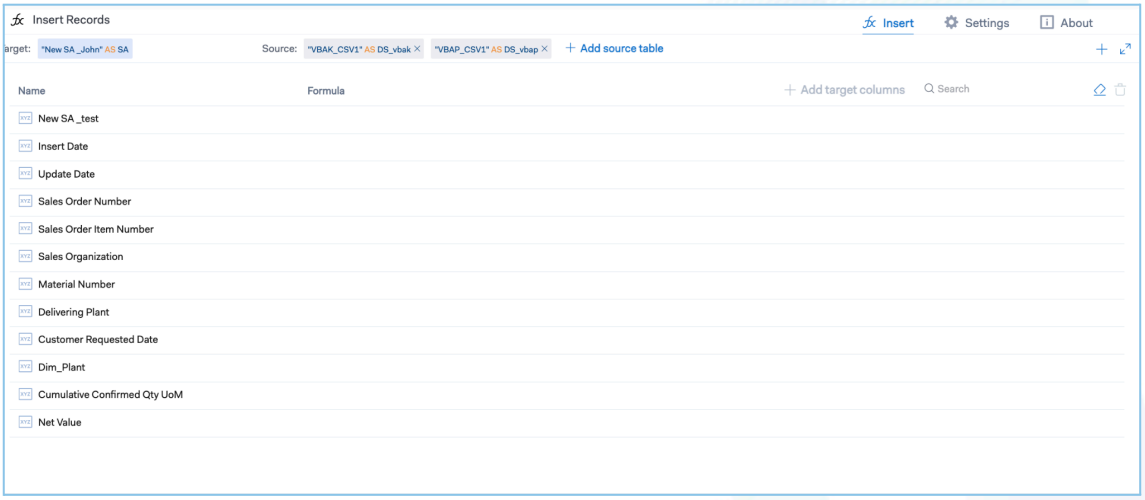


We will bring all the fields from the Sales Header and Sales Item dataset for all the matching Sales Order Numbers.

10. Click **Add target columns**, select **ALL**, and click **Add Selected**

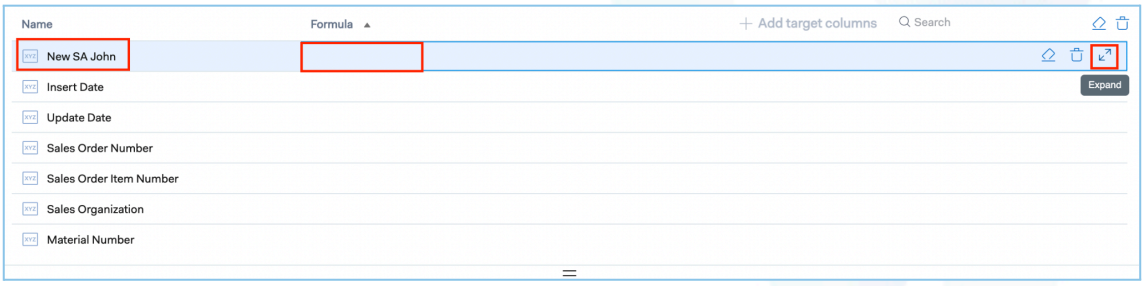


11. We see all the target columns have been added.



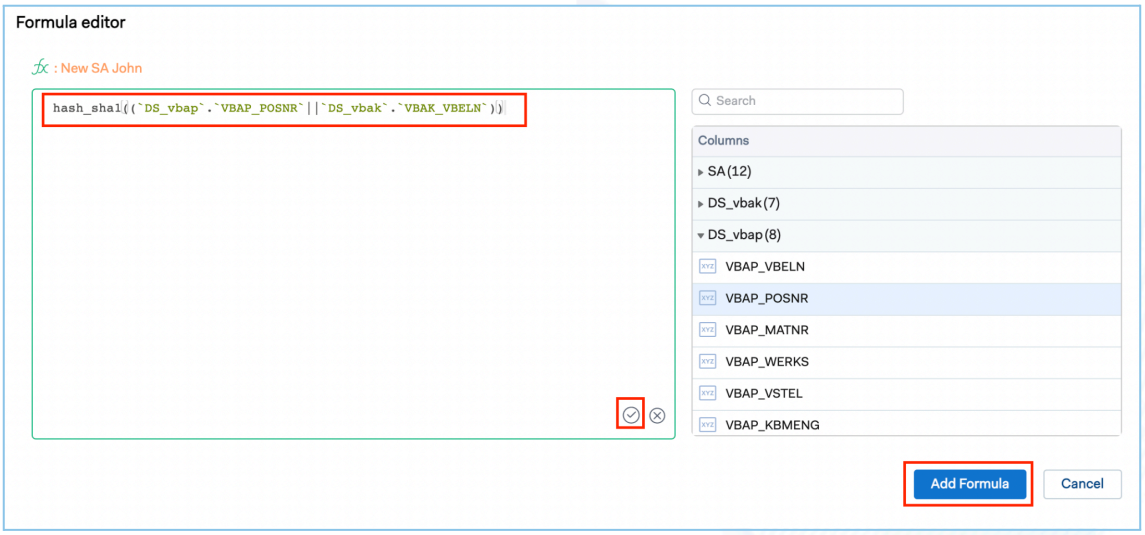
Map Target columns to Source columns

1. Click on the **formula** option next to the **New SA John** field and then click on **Expand** to add the formula.



2. Enter the below formula:

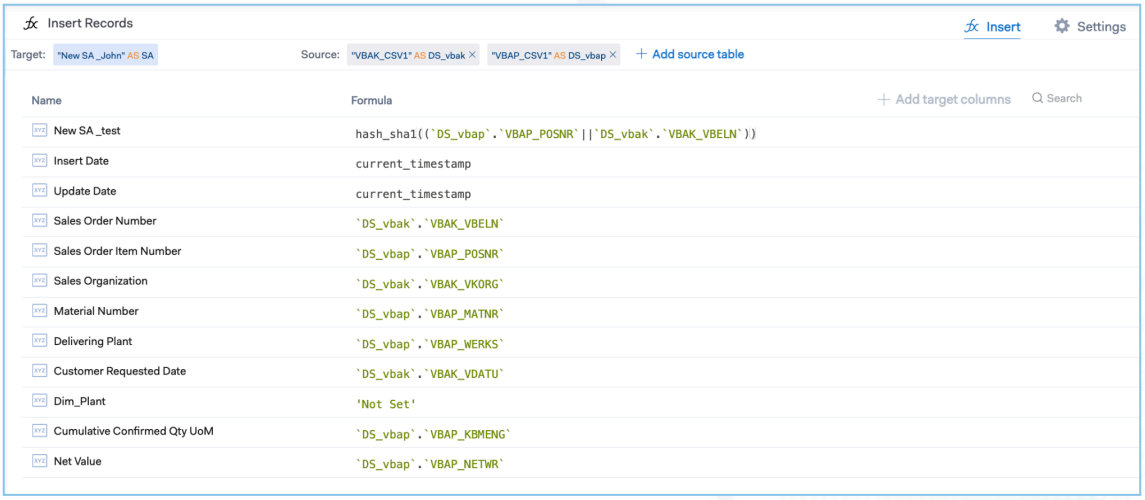
   hash_sha1((`DS_vbap`.`VBAP_POSNR`||`DS_vbak`.`VBAK_VBELN`)) then **Validate** and **Add Formula.**



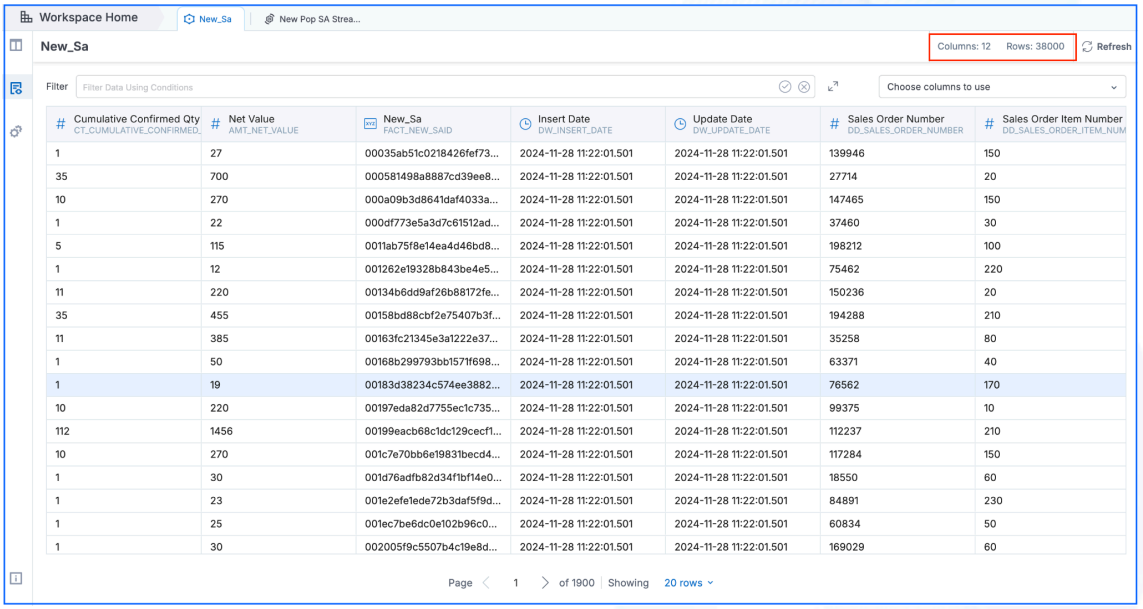
In the field **New SA, John** acts as a unique **ID**. We are creating a 40-digit hash value by combining **Sales Order No** and **Sales Order Item No**

3. Similarly, add the formula for the remaining fields as given below:

Sl No.	Target Column	Formula
1	Insert Date	current_timestamp
2	Update Date	current_timestamp
3	Sales Order Number	`DS_vbak`.`VBAK_VBELN`
4	Sales Order Item Number	`DS_vbap`.`VBAP_POSNR`
5	Sales Organization	`DS_vbak`.`VBAK_VKORG`
6	Material Number	`DS_vbap`.`VBAP_MATNR`
7	Delivering Plant	`DS_vbap`.`VBAP_WERKS`
8	Customer Requested Date	`DS_vbak`.`VBAK_VDATU`
9	Dim_Plant	'Not Set'
10	Cumulative Confirmed Qty UoM	`DS_vbap`.`VBAP_KBMENG`
11	Net Value	`DS_vbap`.`VBAP_NETWR`



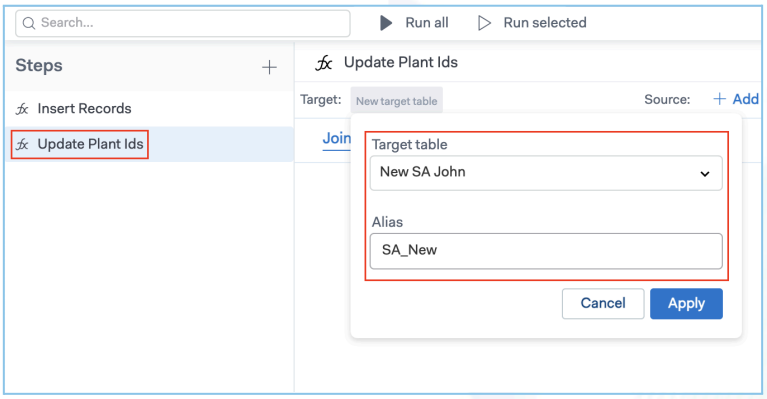
5. Click **Save** and **Run** the Step, then check the **data** under the **New SA Username**'s Data tab after successful execution.



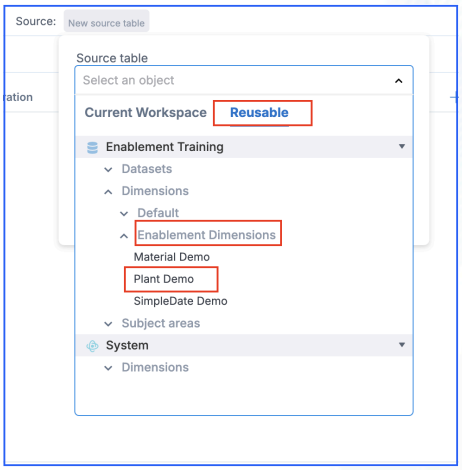
6. Also, notice that the **Dim_Plant** field has **'Not Set'** as the value. We will update them in the next step.

Update Plant Ids

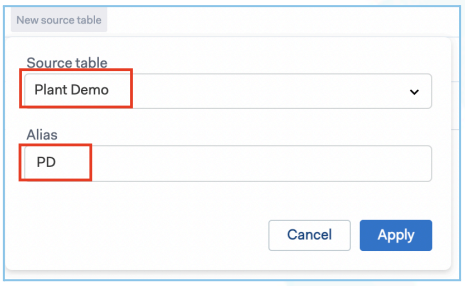
1. Follow the same step as outlined in the Insert Records step and set the Alias as **SA_New**.



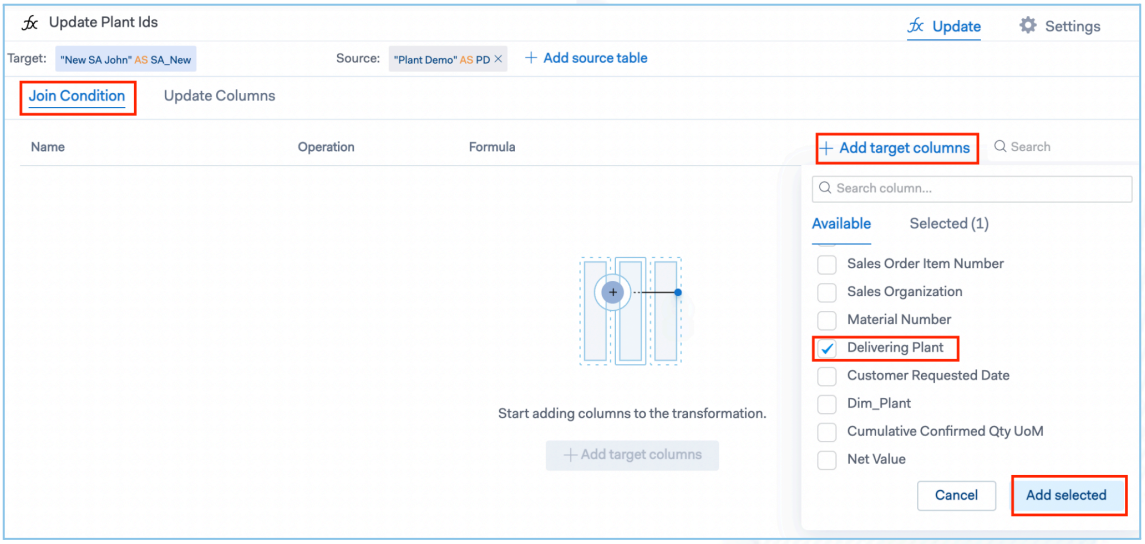
2. Click **Add Source Table**, select "**Reusable**" from the drop-down, and navigate to the **Plant Demo** dimension.



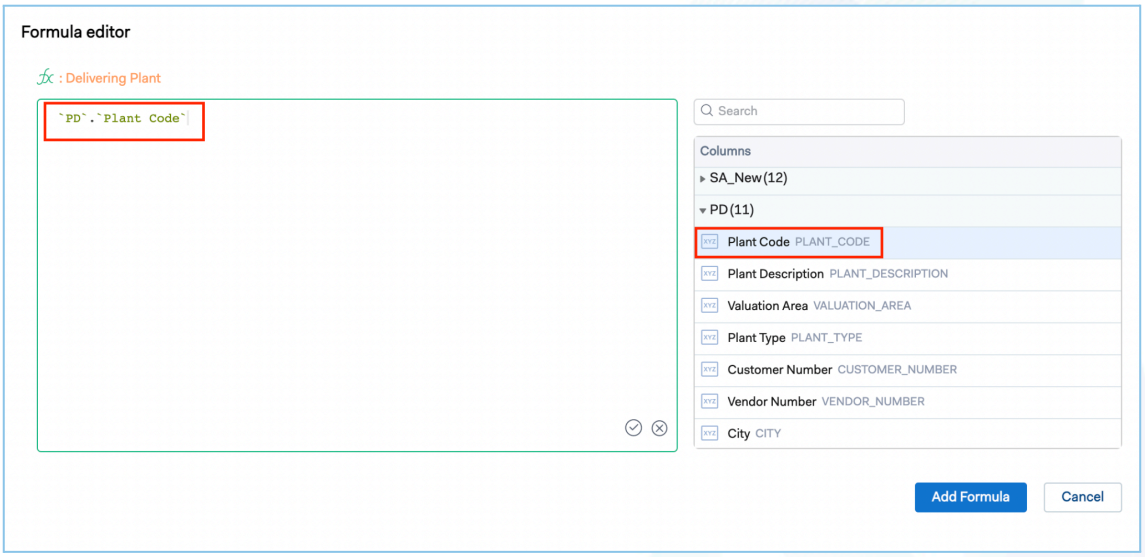
3. Set Alias as **PD** and click **Apply.**

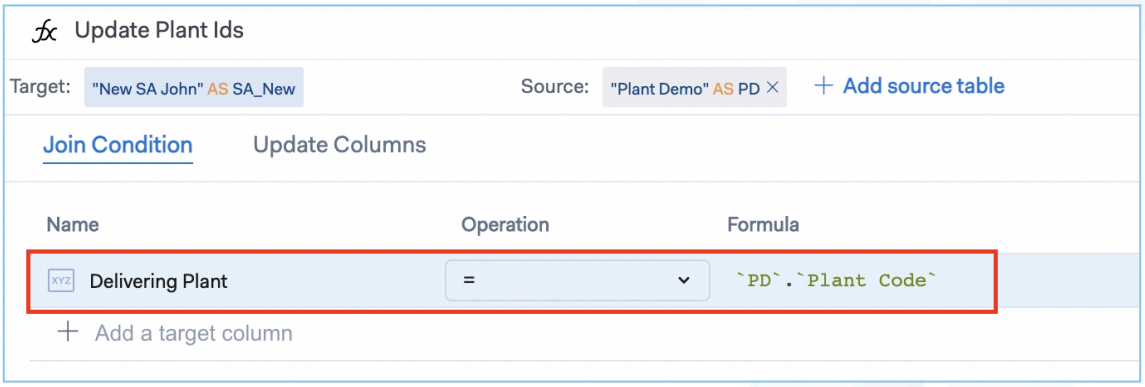


4. Under the **Join Condition** tab, click on **Add target columns**, and select **Delivering Plant** and click on **Add Selected.**

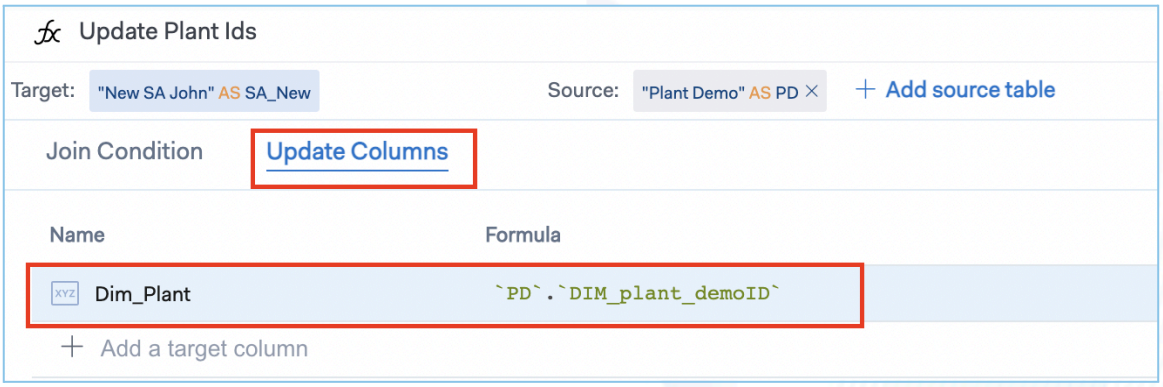


5. Choose the Operation = and provide the formula as shown below and click **Save.**



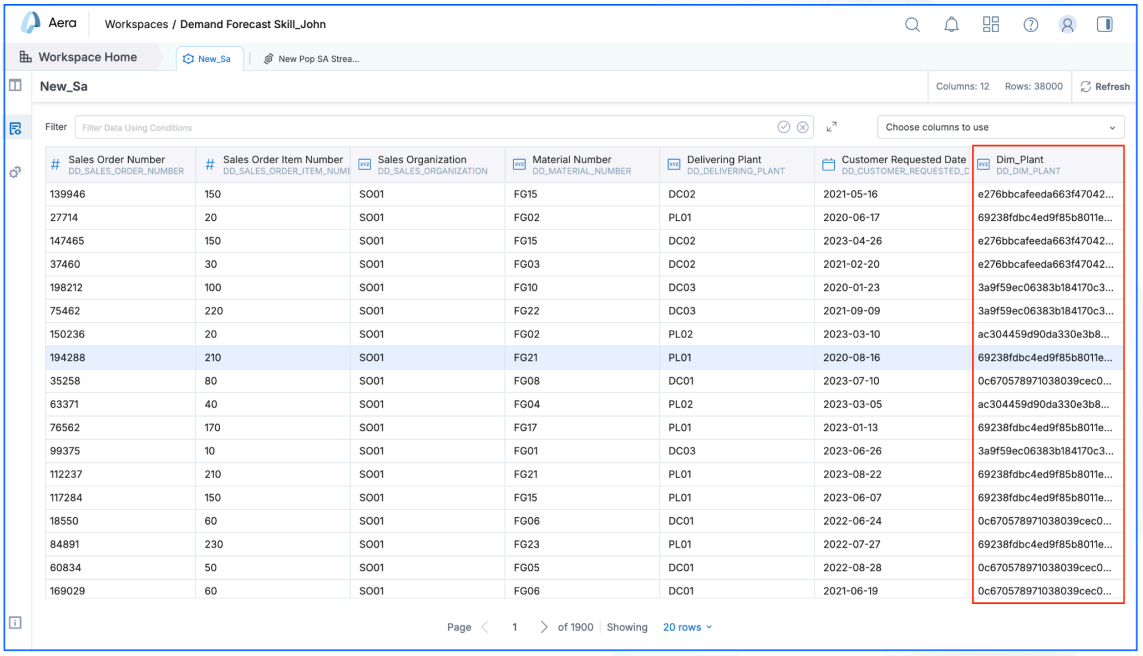


6. Navigate to the **Update Columns** tab and enter the **target column** and **formula** as shown below.



For all the matching **Delivering Plants**,we will update the **Plant IDs** by mapping them to the **Plant IDs** from the external Dimension **Plant Demo**

7. Click **Save** and **Run** the step.
8. After completing the step, go to the **New SA Username** Subject Area's Data tab, click Refresh, and check the updated **Dim_Plant** IDs.



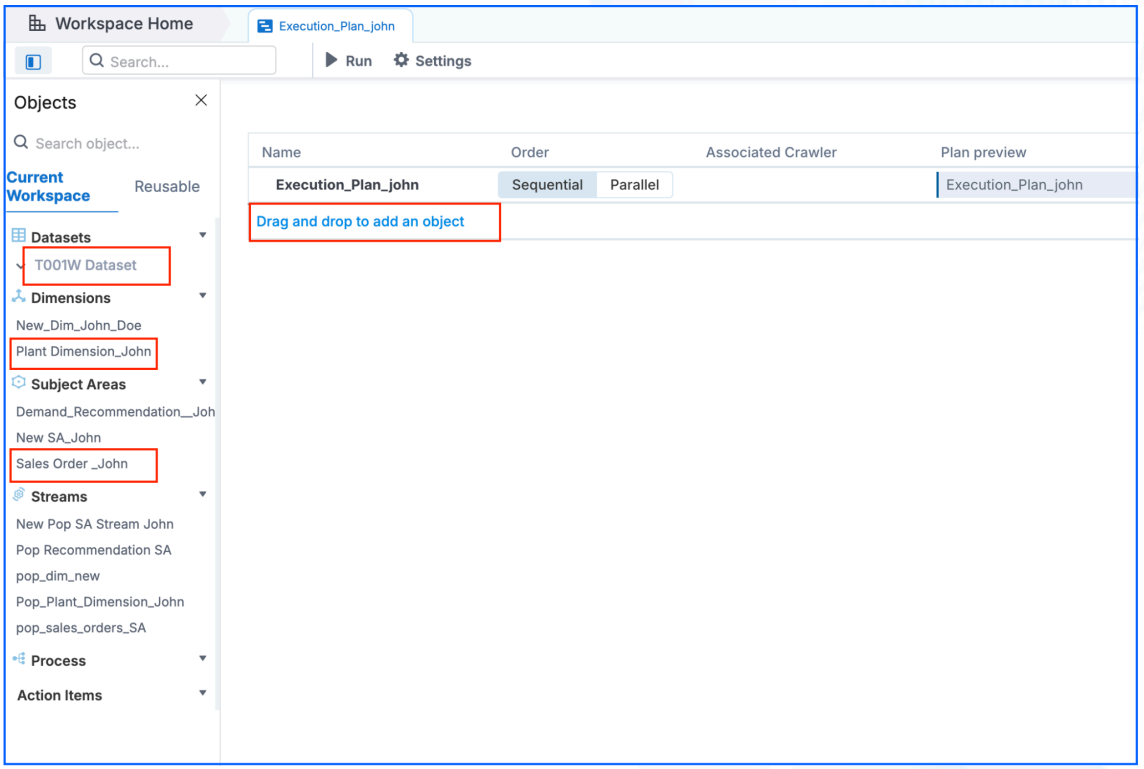
9. You have the Subject Area populated using both **SQL** and the **Insert/Update** Steps

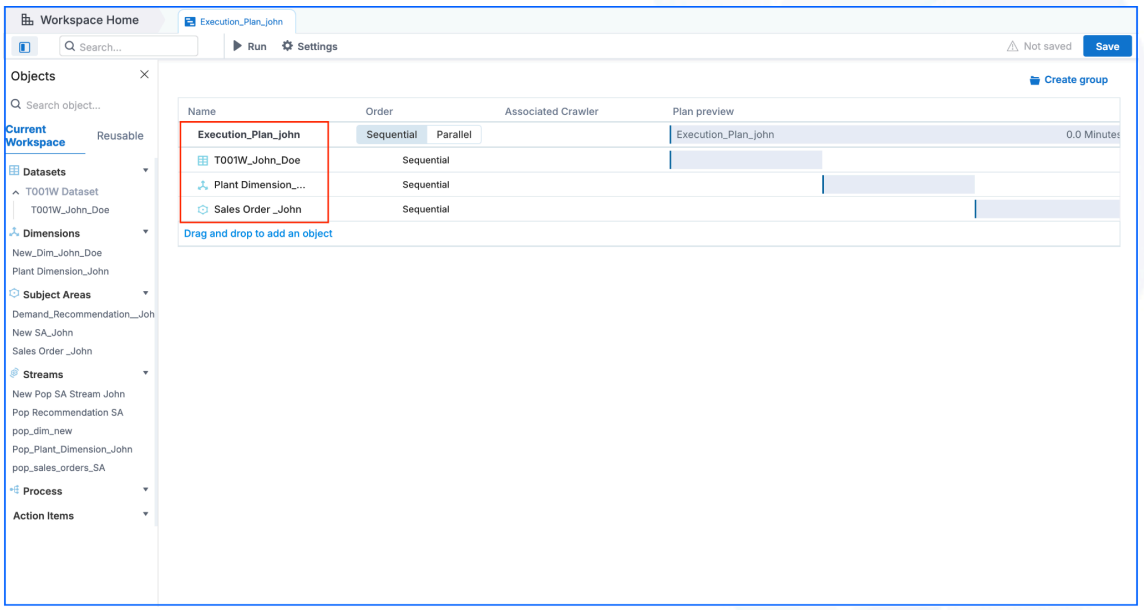
Add an Execution Plan

Transactions are updated regularly in the source system and need to be refreshed/updated in Aera DDM. Thisis done by configuring an Execution plan and adding it to a scheduler.
Create a new Execution Plan and add the objects that need to be executed in the below order:

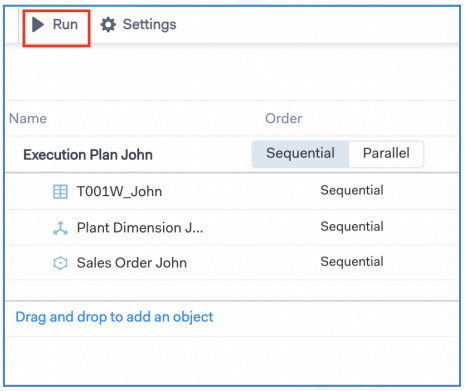
- Dataset
- Dimension
- Subject Area

1. Navigate to the **Execution Plan** section under **Content** and tap the '**+**' icon to create a new Execution Plan and provide the name as **Execution Plan <username>**.
2. Drag and drop your **Dataset**, **Dimension**, and **Subject Area** as shown.

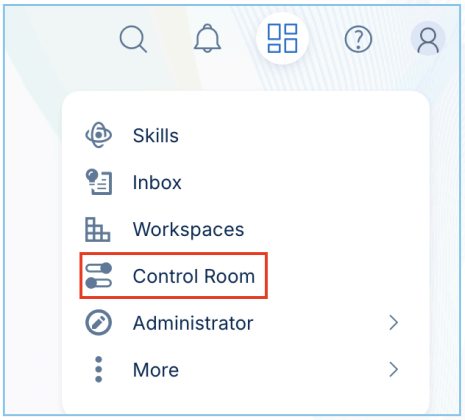




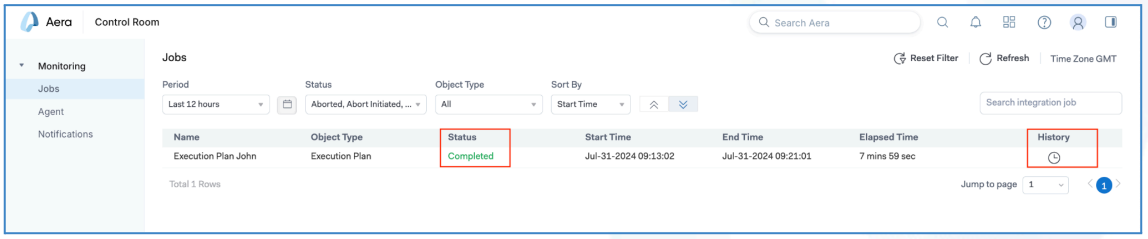
3. Ensure **sequential** order, **save** the execution plan, and click **"Run."**



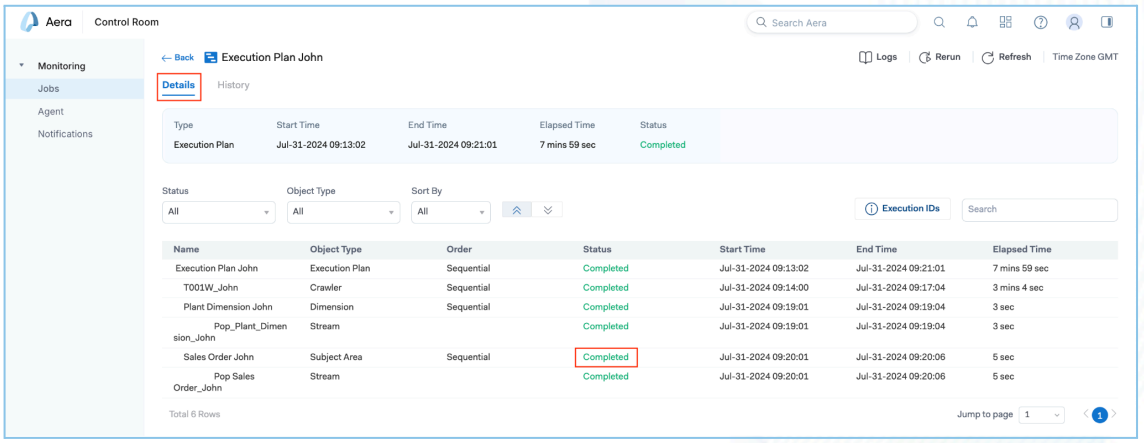
4. You can monitor the progress of your execution plan by navigating to the **Monitoring**.
5. From the **Aera Menu**, navigate to **Control Room**, then click on **Monitoring** in the top left corner, and finally go to Jobs.



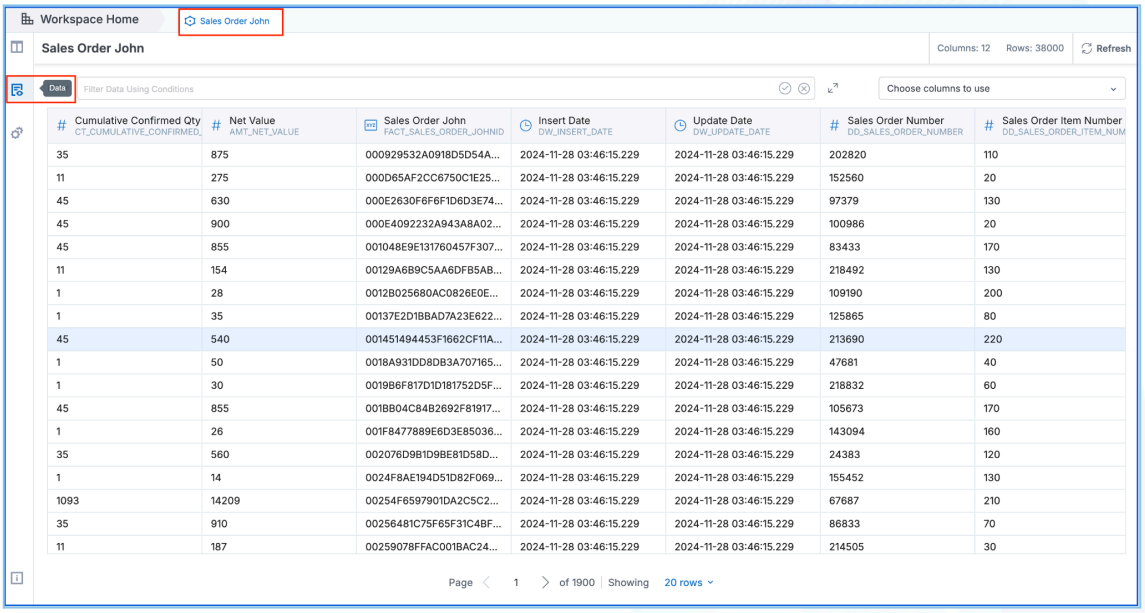
6. You can see the status of the Execution Plan in the **Jobs.** The execution plan will take a few minutes to complete. Refresh at regular intervals to see the progress.



7. Once the Execution Plan's status is displayed as **Completed**, click on **History** to view its history. Then, go to **Details** to obtain detailed information about your Execution Plan.



8. The successful execution confirms the population of your Subject Area.



SUMMARY

In this hands-on exercise, you have:

- Successful setup of a Dataset ("**T001W_<Username>**") containing relevant plant data.
- Creation of two Dimensions ("**Plant Dimension_<Username>**", "**New Dim_<Username>**") representing plant-related information.
- Configure and populate two Subject Areas ("**Sales Order_<Username>**", "**New SA_<Username>**") capturing sales order data linked to plant dimensions.
- Configuration and successful run of an Execution Plan to automate the data loading and refresh process.

Note: This exercise is limited to creating and populating subject areas.

In the upcoming modules, the subject area and dimensions will be created here, but they will use a pre-created Subject Area and Dimensions.