

Defining Aggregates Yourself

You can define your own aggregates for use cases that fall outside of those covered by system-defined aggregates.

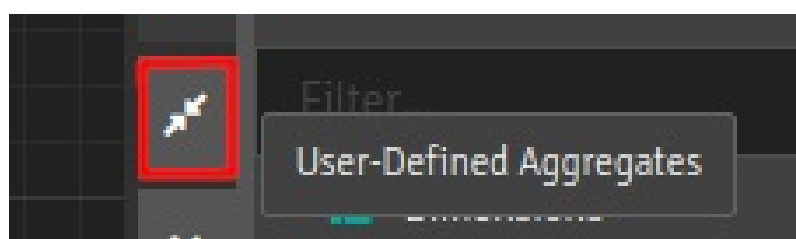
Before You Begin

- ▶ Read about [When to Define Your Own Aggregate Tables](#) and [Partitioned User-Defined Aggregates](#).
- ▶ If your aggregate definition will use one or more joins to dimensional datasets and the aggregate is allowed to be built with incremental builds, follow these steps before creating the aggregate:
 1. Read the section Incremental rebuilds of aggregates that use joins in [About Incremental Rebuilds](#) to ensure that you are aware of the special considerations for such aggregates.
 2. On the Cube Designer canvas, edit the fact dataset by clicking the gear icon in its toolbar and selecting **Allow Aggregates** and **Allow incremental builds**. Refer to [About Incremental Rebuilds](#) for info about selecting an incremental indicator and setting a grace period.
 3. On the Cube Designer canvas, double-click a dimension that will be joined to your aggregate. Then, edit the dimension dataset by clicking the gear icon in its toolbar and selecting both **Allow Aggregates** and **Safe when joined to incremental**. Follow this step for each dimension that will be joined to your aggregate.
- ▶ If you want instances of your aggregate definition to be partitioned:
 - ▶ Ask a super user to verify that the engine configuration settings TABLES.CREATE.PARTITIONS.ENABLED and AGGREGATES.CREATE.PARTITION.USERDEFINEDAGGREGATE.ENABLED are both set to True, as described in [Engine Settings for User-Defined Aggregates Only](#).
 - ▶ Follow the steps [here](#) for setting a partition key on a UDA from within the Cube Designer Canvas.

Procedure

1. Open the cube that you want to define an aggregate for.
2. In the Cube Designer, click the **User-Defined Aggregates** (double arrow) icon in the tool panel on the right to open the Aggregate panel.

Figure 1. The User-Defined Aggregates icon, highlighted by an orange box in this image, in the tool panel. Hover over this icon to see the tooltip that displays the icon's name.



3. In the Aggregates panel, click the menu icon and choose **Add Aggregate**.

Figure 2. The Create User-Defined Aggregate dialog.

4. In the **Create User-Defined Aggregate** dialog, enter the name for the user-defined aggregate.



Note: Aggregate table names used by the query engine are system-generated, but they will include the first 14 characters of the user-supplied name at the end of the internal ID name. This name can help you identify when a user-defined aggregate is used in a query. For example: `as_agg_internal-id_my-uda-name`

5. (Optional) Click **Add Notes** on the right to enter a description of why you are creating the aggregate--for example, for testing or for permanent use.
6. Use the **Dimensions** tab on the left to select the dimension attributes to include in the aggregate definition.

These are the data values that will be used to group the summarized measure data in the resulting aggregate table.



Note: User-defined aggregate definitions are fixed. They do not include all levels of a hierarchy unless explicitly selected.

7. If you want the AtScale engine to partition instances of this aggregate:

1. Choose the **Partitions** tab on the right.
2. Open the menu for a dimension, and specify whether the partition should be defined on the key column, name column, or both.

When the engine builds an instance of this aggregate, the engine will create a partition for each combination of values in the dimensional attributes. The number of partitions depends on the left-to-right order of the attributes, as well as the number of values there are for each attribute. You can drag the attributes to change their left-to-right order.

Essentially, the partitioning key functions as a GROUP BY column. Queries against the aggregate must use this dimensional attribute in a WHERE clause. A good candidate for a partitioning key is a set of dimensional attributes that together have a few hundred to under one thousand value combinations. (The default maximum number of partitions that the AtScale engine allows is 800. Super users can increase the maximum, though AtScale recommends that it be set no higher than 1000.)

8. If you need to specify distribution keys, you can do this in the **Distributions** tab on the right.

If your aggregate data warehouse supports distribution keys, then AtScale will use the specified keys when creating the aggregate table. For more information, see [Settings for large table optimization](#).

9. Use the **Measures** tab on the left to select the measures to include in the aggregate definition.

This is the data that will be summarized in the resulting aggregate table.



Note: By default, AtScale does not create system-defined aggregate tables that aggregate on DISTINCT COUNT (estimated only) measures. If you want to configure your system to do so, you can set the `aggregates.Create.Allowexactdistinctcountmeasures.Enabled` engine setting to `True`; for more information, see [Other Settings for System-Defined Aggregates Only](#).

Additionally, you can define your own aggregate tables to aggregate on such measures. However, the AtScale engine only uses those tables when queries match their dimensions exactly. If any other dimensions are added (sort, filter, and so on), the match is invalidated. If this occurs, AtScale bypasses the aggregate and calculates against the fact table(s).

10. Click **Save** when you are done.

What To Do Next

The user-defined aggregate table will be created the next time the cube is published. You can check the status of the aggregate creation on the Aggregates page.

1. At the top of the Design Center, click **Aggregates**.
2. On the left, check **Instances** and **Build History**.

If the AtScale engine fails to build the first instance of your aggregate and you are using a partitioning key, there are three different actions that you can try for a successful build:

- ▲ Use a partitioning key that has a cardinality that is lower than the configured maximum allowed number of partitions.
- ▲ Increase the maximum allowed number of partitions. (However, AtScale recommends that you set the maximum no higher than 1000.)
- ▲ Edit the user-defined aggregate by unselecting the partitioning key.