

Creating Master Segments

You define your data model in **Data Workbench > Master Segments**.

In Treasure Data Workflows (**Data Workbench > Workflows**) a series of workflows are automatically generated to create and refresh the master segment data model. Data in the data model is refreshed according to the schedule that you define for the master segment.

Creating and Defining Master Segments

1. From the TD Console, select Data Workbench > Master Segments.
2. Select **New Master Segment**.
3. Specify Master Table , Attribute Tables , and Behavior Tables.

The screenshot shows the configuration form for a Master Segment. It is divided into several sections:

- Schedule & Timezone:** Includes a 'Schedule' dropdown set to 'Monthly (d,hh:mm:ss format)' and a 'Run at' field set to '1,00:00:00'. Below it is a 'Timezone' dropdown set to 'US/Pacific'.
- Master Table:** A separate box containing a 'Database' dropdown set to 'sample_db' and a 'Table' dropdown set to 'sample_table'.
- Resource Pool:** Features two radio buttons: 'Presto and Hive' (which is selected) and 'Hive only'.
- Presto Pool Name:** A text input field containing 'Presto resource pool name'.
- Hive Pool Name:** A text input field containing 'Hive resource pool name'.
- Attribute Table (0) (+)** and **Behavior Table (0) (+)** are shown at the bottom, indicating no additional tables are currently defined.

4. Specify the schedule by which the master segment data is refreshed. We recommend that you run a refresh on a daily basis, to renew the dataset. The master table needs to be updated as well, before master segment creation begins.
5. Optionally, specify the default timezone.
6. Specify the Master database and table name. For example:

This is a close-up view of the 'Master Table' configuration section. It shows two dropdown menus side-by-side:

- The 'Database' dropdown is set to 'sample_datasets'.
- The 'Table' dropdown is set to 'nasdaq'.


7. Optionally, specify the resource pool. This allows you to define how your allotted Treasure Data resources are used to process the data.

We generally recommend Hive only, because it's designed for large-scale joins and known for robustness, while Presto and Hive might perform faster for smaller data sets.





8. Optionally, specify the attribute tables. The Foreign Key and Join Key is used to join Master Table with Attribute Tables. Include a td_ column here for automatic data enrichment.


Attribute Table (2) 


Foreign Key


close 

Database Table Join Key


 sample_dat  nasdaq  close 

Columns (1) 





 Column Name


close  close


Foreign Key


volume 

Database Table Join Key

 sample_dat  nasdaq  volume 

Columns (1) 

 Column Name

volume  volume

9. Optionally, specify tables from which behaviors can be derived. Use the Name field to give a user-friendly name to the specific column. Include a td_ column here for automatic data enrichment.

Behavior Table (1) +

Name

Foreign Key

⊖

Database	Table	Join Key
<input type="text" value="_mobile_sdk"/>	<input type="text" value="android"/>	<input type="text" value="activity_date"/>

Columns (0) + All Columns

10. Save, name, and run your master segment.
11. Optionally, view the progress of the workflow that is created by defining the master segment. From the main navigation, select Data Workbench > Workflows and select the Master Segment filter.

– TYPE

- User-Defined
- Master Segment
- Predictive Scoring
- Activation

The screenshot shows the Data Workbench interface. The left sidebar contains navigation options: Queries, Workflows (selected), Databases, and Master Segments. The main content area displays the workflow 'audience' with a 'SUCCESS' status. The workflow details include: Session (Today 4:54 pm), Project (cdp_audience_3234), Workflow (audience), Status (SUCCESS), Last Attempt (Today 4:54 pm), Duration (21 secs), Session ID (169659), Revision (508618d0-48da-4f97-b3a9-5075f96db581), and Session UUID (ca52d086-1766-4ed0-b7de-eeb55a1a5a39). Below the details is a 'Timeline' section with a table showing task execution progress.

Task	Duration	Execution	Logs
+get_queries	3 secs	<div style="width: 100%; background-color: #28a745;"></div>	WF D E
+create_database	0 ms	<div style="width: 100%; background-color: #28a745;"></div>	WF D E
+customers	7 secs	<div style="width: 100%; background-color: #28a745;"></div>	WF O E
+behaviors	1 sec	<div style="width: 100%; background-color: #28a745;"></div>	WF D E
+enrichments	2 secs	<div style="width: 100%; background-color: #28a745;"></div>	WF D E
+predictive_scoring	1 sec	<div style="width: 100%; background-color: #28a745;"></div>	WF D E

12. To verify that the process ran and created your master segment, navigate to Profiles.



Audience Studio



MASTER SEGMENT

m-j_sample_db



Dashboard



Segments



Profiles



Predictive Scoring



Profiles API Tokens