Cloudera Data Visualization 7.2.9

CDP Data Visualization Top Tasks

Date published: 2020-10-30
Date modified: 2024-12-18

https://docs.cloudera.com/

Legal Notice

Contents

Deploying Cloudera Data Visualization

The first step to start using Cloudera Data Visualization is to deploy it on your chosen Cloudera platform. Cloudera Data Visualization provides consistent visualization capabilities across various Cloudera deployments and services. Once your Cloudera Data Visualization instance is up and running, you can start creating visuals, build customized dashboards, and develop analytical applications to enable users to explore data across the whole data lifecycle.

You can log into Cloudera Data Visualization using your Cloudera credentials and start analyzing data that you have access to. In Cloudera Public Cloud, Cloudera Data Visualization is included as a core capability in Cloudera Data Warehouse and Cloudera AI both for Azure and AWS. Cloudera Data Visualization is also available in Cloudera Private Cloud for Cloudera Data Warehouse and Cloudera AI, and it can be used in Cloudera Data Science Workbench, certified with CDH, HDP and Cloudera Private Cloud Base.

Deploying Cloudera Data Visualization in Cloudera Data Warehouse

Learn how to create and launch a Cloudera Data Visualization instance within the Cloudera Data Warehouse data service. You can use Cloudera Data Visualization with Cloudera Data Warehouse in both Cloudera Public Cloud and Cloudera Private Cloud to explore and visualize data stored in database catalogs providing data-driven insights throughout the entire data lifecycle.

About this task

If you want to create visuals based data stored in Cloudera Data Warehouse, you have to create a Cloudera Data Visualization instance and connect it to Hive or Impala Virtual Warehouse(s).

Cloudera Data Visualization is not tied to a particular Virtual Warehouse (VW). You can access data for your visualizations from multiple Data Catalogs using multiple Hive or Impala Virtual Warehouses in various environments. With multiple Cloudera Data Visualization instances attached to an environment, you can connect to different data sets, create dashboards for different groups and share your visualizations with different users.

Note: When you delete a Virtual Warehouse, your visual artifacts remain intact as long as the Cloudera Data Visualization instance is not deleted.

Before you begin

- You are logged into the Cloudera web interface and you have opened the Cloudera Data Warehouse service.
- You have DWAdmin role in Cloudera Data Warehouse.
- You have activated your environment.
- You have a Hive/Impala warehouse in running state.
- If you are using Cloudera Data Visualization with Cloudera Data Warehouse in Cloudera Private Cloud:

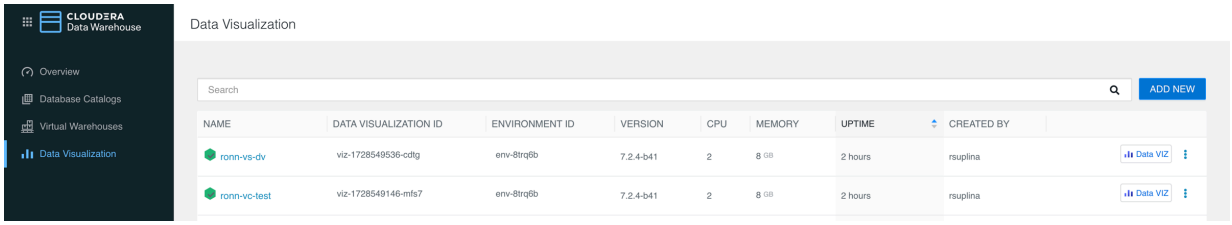
  You have an admin group created in Management Console User Management . To log in using LDAP, select the Sync Groups on Login option from Management Console Administration Authentication , so that the associated groups can be imported when you log in to Cloudera Data Visualization.

Procedure

1. In Cloudera Data Warehouse, click Data Visualization in the left navigation panel.

   A list of existing Cloudera Data Visualization instances appears, if there are any.

2. Click ADD NEW to create a new instance.
   If you are creating the first Cloudera Data Visualization instance in Cloudera Data Warehouse, click CREATE.



3. Provide the following information in the **New Data Visualization** modal window:

   Mandatory fields are marked with an asterisk.

   a) Name* – Specify a name for the instance.
   b) Environments* – Select the environment that you want Cloudera Data Visualization to be connected to.
   c) User Groups – Add user groups to allow user access to Cloudera Data Visualization for selected groups. If no group is added, all Cloudera users will have non-admin access.
   d) Admin Groups* – Add admin groups to allow configuration access to Cloudera Data Visualization for selected groups.

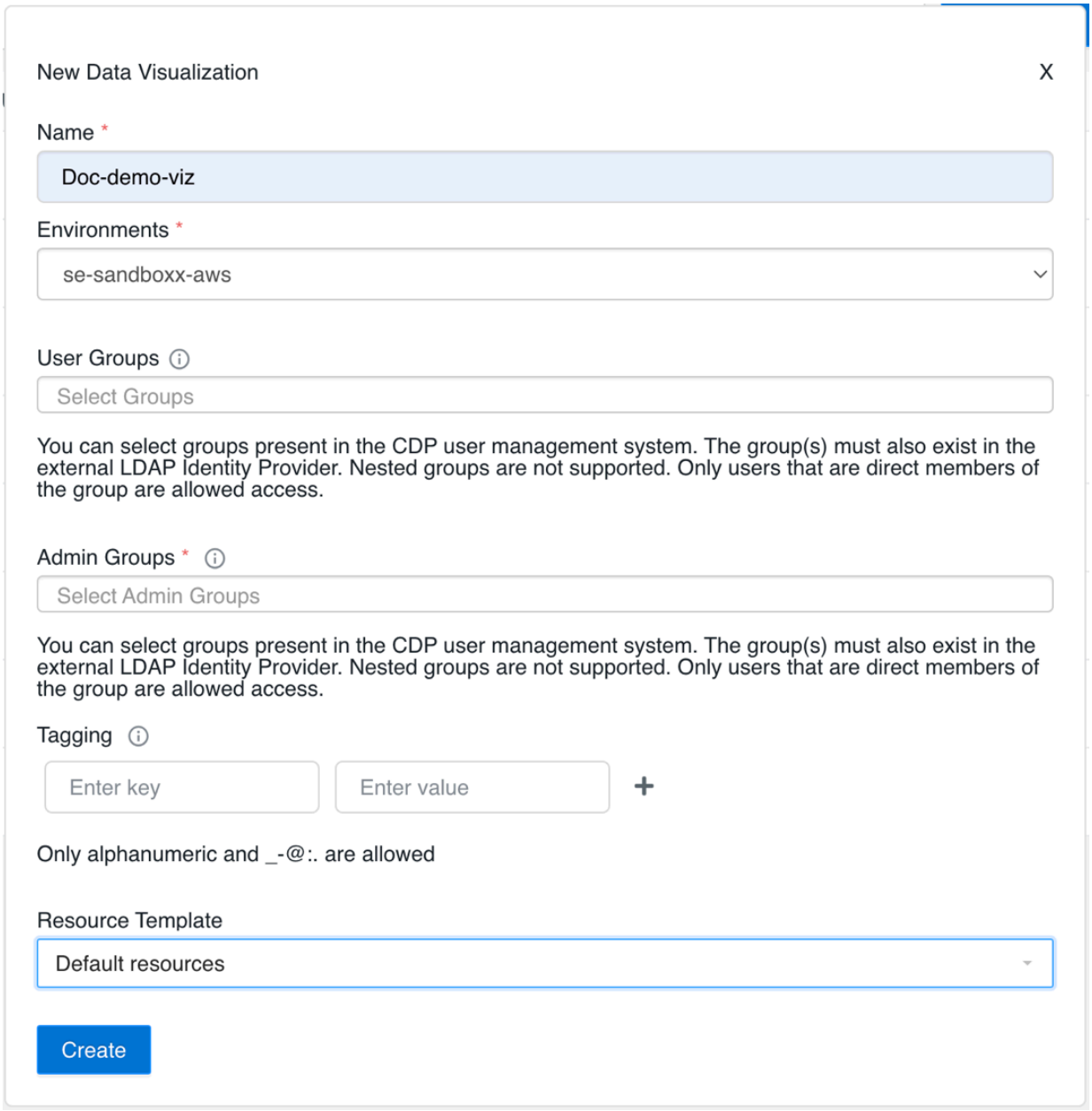
      For more information on Cloudera Data Visualization permission granularity, see the *Security model*.
   e) Tagging – Enter keys and values to apply tags to your resources for organizing them into a taxonomy.

      Note: This field is not available in Cloudera Data Warehouse on Cloudera Private Cloud.

   f) Resource Template – Select the Cloudera Data Warehouse resource template from the drop-down menu.

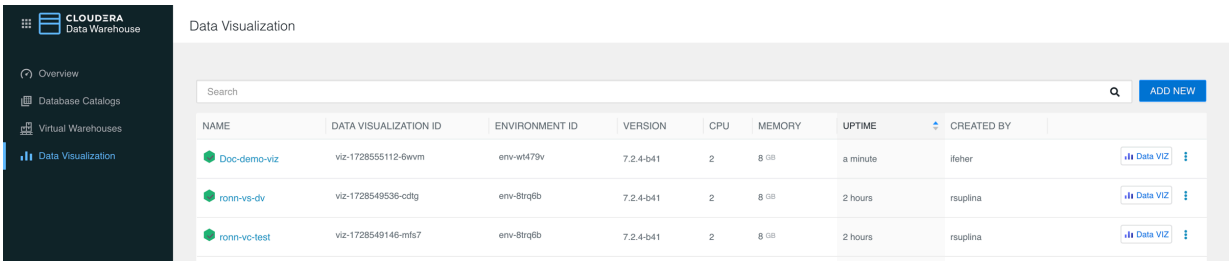
      - Default resources
      - Medium resources
      - Large resources



4. Click CREATE.

   Instance creation starts. Wait until the Cloudera Data Visualization instance is in running state.

5. You can find the list of Cloudera Data Visualization instances and environments appears under the Data Visualization menu that you can open from the left navigation panel.

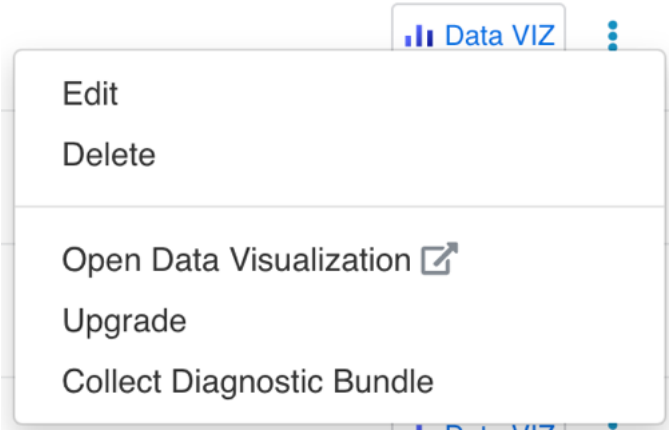


6.

Select one from the list of running Cloudera Data Visualization instances and click 

to start Cloudera Data Visualization.

Alternatively, you can click the launch/options menu (three dots) at the end of the row and click Open Data Visualization in the pop-up menu.



SSO authentication is enabled by default both in Cloudera Public Cloud and Cloudera Private Cloud. If you are logged in to the Control Plane, you will be logged into Cloudera Data Visualization automatically. If not, log into the Control Plane using your LDAP credentials. Cloudera Data Visualization opens in a new browser tab and you land on the Cloudera Data Visualization homepage, where you can explore sample dashboards and access the in-tool *Get Started guide* for help.

New users logging in to Cloudera Data Visualization are automatically added to the viz_guest_group group. You can assign the System Admin role to this group to grant new users the permission to create new connections. For more information, see *Assigning roles to users*.

Deploying a Cloudera Data Visualization application in Cloudera AI

Learn how to deploy Cloudera Data Visualization in Cloudera AI.

Creating a Cloudera AI project with Cloudera Data Visualization Runtime

Learn how to create a Cloudera AI project with Cloudera Data Visualization Runtime as the default runtime.

About this task

If you know that your project will to be running Cloudera Data Visualization, you can add the Cloudera Data Visualization Runtime when setting up the project.

Procedure

1. Click Projects on the left sidebar of Cloudera AI Workbench.
2. Click New Project.
3. Enter a Project Name.
4. You can add a description for your project.
5. Set the visibility of the project.
   - Private (default) – Only you can access the project.
   - Public – All authenticated users can view the project.

6. Under Initial Setup, select how you want to create your project.

   You can either create a blank project, or select a source for your project files.

   - Blank – Start with an empty project (no templates, files, or Git sources).
   - Templates – Pre-built example projects in R, Python, PySpark, or Scala to help you get started.
   - APMs – Use Accelerators for ML Projects to to include jobs, models, and experiments.
   - Local Files – Upload an existing project from a compressed file or folder.
   - Git – Clone a Git repository for version control and collaboration.

7. Configure which Runtime(s) will be available for this particular project.

   Projects are configured with the latest Python and R ML Runtimes. You can change this configuration in the Advanced Options view, where you can add ML Runtimes based on more detailed Editor, Kernel, Edition, and Version criteria.

   a) Enable Advanced Options to customize.
   b) For Cloudera Data Visualization 8.0.0 and higher:
      1. Set the Editor to PBJ Workbench.
      2. Select Cloudera Data Visualization from the Kernel dropdown.
      3. Edition and Version will be automatically set for Cloudera Data Visualization.
   c) For Cloudera Data Visualization versions lower than 8.0.0, set the Editor to Workbench.

      Kernel, Edition, and Version will be automatically set for Cloudera Data Visualization.
   d) Click Add Runtime.

8. Click Create Project.

What to do next

After the project is created, you can start creating your application. If you added Cloudera Data Visualization as the only Runtime during setup, it will be the default Runtime when creating applications.

Adding Cloudera Data Visualization Runtime to an exisiting Cloudera AI project

Learn how to add a Cloudera Data Visualization Runtime to an existing Cloudera AI project.

About this task

You need to manually add a Cloudera Data Visualization Runtime to your project if the workspace of your project is not set to use the Cloudera Data Visualization Runtime by default.

Procedure

1. Click Projects on the left sidebar of Cloudera AI Workbench.
2. Select the project where you want to add Cloudera Data Visualization Runtime.
3. Open Project Settings from the left navigation bar.
4. Click the Runtime tab.
5. Click Add Runtime.

   The Add new runtime to project modal window opens.

   a) For Cloudera Data Visualization 8.0.0 and higher:
      1. Set the Editor to PBJ Workbench.
      2. Select Cloudera Data Visualization from the Kernel dropdown.
      3. Edition and Version will be automatically set for Cloudera Data Visualization.
   b) For Cloudera Data Visualization versions lower than 8.0.0, set the Editor to Workbench.

      Kernel, Edition, and Version will be automatically set for Cloudera Data Visualization.

6. Click Submit.

Results

Now that Cloudera Data Visualization Runtime is added, you can select it when creating a Cloudera Data Visualization application.

What to do next

To proceed, create a Cloudera Data Visualization application.

Creating a Cloudera Data Visualization application in Cloudera AI

Learn how to create a Cloudera Data Visualization application in Cloudera AI to help you visualize and interact with your data insights. This integration allows for seamless visualization of ML Model outputs, data exploration, and reporting within the same platform.

About this task

Important:

Each Cloudera AI project can host only one standalone Cloudera Data Visualization application. Since applications share the same metadata and logs, any changes made by one application may overwrite those made by another.

Before you begin

Ensure that a Cloudera Data Visualization Runtime is available in the Cloudera AI project where you plan to create the Cloudera Data Visualization application.

For more information about ML Runtimes, see Managing ML Runtimes and Using Runtime Catalog.

Procedure

1. Navigate to the Overview page of your Cloudera AI project.
2. On the left sidebar, click Applications.
3. Click New Application.
4. Provide the following details for your new application:
   - Name – Enter a name for the application.
   - Run Application as – If the application is to run in a service account, select Service Account and pick an account from the dropdown.
   - Subdomain – Enter a subdomain that will be used to construct the URL for the web application. Use only URL-friendly characters.
   - Description – Add a description of the application.
   - Script – Use the script located at: /opt/vizapps/tools/arcviz/startup_app.py
   - Runtime – If only one Runtime is available in your project, the fields will be prepopulated. If multiple Runtimes are available, you can select which Runtime to use.
     For Cloudera Data Visualization 8.0.0 and higher
     - Editor – Select PBJ Workbench.
     - Kernel – Select Cloudera Data Visualization.
     - Edition and Version will autopopulate.

     For Cloudera Data Visualization versions lower than 8.0.0
     - Editor – Select Workbench.
     - Kernel, Edition, and Version will autopopulate.
5. Click Create Application.

Results

After a few minutes, the application status will change from Starting to Running on the Applications page. Your Cloudera Data Visualization application is now ready for use.

You can restart, stop, or delete the application using the options in the supplemental menu. If you want to make changes to the application, navigate to Application Details Settings .

What to do next

Start Cloudera Data Visualization.

Deploying a Cloudera Data Visualization application in Cloudera Data Science Workbench

Learn how to deploy Cloudera Data Visualization in Cloudera Data Science Workbench.

Creating a Cloudera Data Science Workbench project with Cloudera Data Visualization Runtime

Learn how to create a Cloudera Data Science Workbench project with a Cloudera Data Visualization Runtime as the default runtime.

Procedure

1. Click Projects on the left sidebar of Cloudera Data Science Workbench.
2. Click New Project on the Projects page.
3. Enter a project name.
4. Select the visibility of the project.
5. Under Initial Setup, you can either create a blank project, or select a source for your project files.
6. Click Create Project.

   The created project uses Cloudera ML Runtimes that are managed in the Runtime Catalog.

What to do next

After the project is created, you can start creating an application. You can select which runtime edition you want to use for the application.

Creating a Cloudera Data Visualization application in Cloudera Data Science Workbench

Learn how to create a Cloudera Data Visualization application in Cloudera Data Science Workbench to help you visualize and interact with your data insights. This integration allows for seamless data visualization, exploration, and reporting within the same platform.

About this task

The following steps will guide you on how to create a Cloudera Data Visualization application within Cloudera Data Science Workbench.

Important:

Each Cloudera Data Visualization project can host only one standalone Cloudera Data Visualization application. Since applications share the same metadata and logs, any changes made by one application can overwrite those made by another.

Procedure

1. Navigate to the Overview page of your Cloudera Data Science Workbench project.
2. On the left sidebar, click Applications.
3. Click New Application.
4. Provide the following details for your new application:
   - Name – Enter a name for the application.
   - Subdomain – Enter a subdomain that will be used to construct the URL for the web application. Esure it only contains URL-friendly characters.
   - Description – Add a description for the application.
   - Enable Unauthenticated Access – Mark the checkbox if you want to allow unauthenticated access to your application. You can also update this setting later on the Settings page of the application.

     Note: To create public applications on an ML workspace, an Admin user needs to turn on the feature flag in Admin Security by selecting Allow applications to be configured with unauthenticated access. For more information on working with public applications, see *Application Limitations* in the Cloudera Data Science Workbench documentation.
   - Script – Select the path to the startup script.

     Note: Use the script located at: /opt/vizapps/tools/arcviz/startup_app.py
   - Runtime
     - Editor – Workbench
     - Kernel – Select Cloudera Data Visualization for the kernel supported by the Runtime variant of the Cloudera Data Science Workbench project.
     - Edition – Select the edition of the Runtime variant you want to use for your application.

       Note: The selected edition determines the version of the Runtime variant.

     All available runtimes are listed in the Cloudera Data Science Workbench Runtime Catalog. For more information about ML Runtimes, see Managing ML Runtimes and Using Runtime Catalog.

     Note: New ML Runtime releases are automatically added to the deployment, if internet connection is available.
5. Click Create Application.

Results

After a few minutes, the application status will change from Starting to Running on the Applications page. Your CDV application is now fully operational.

You can restart, stop, or delete a Cloudera Data Science Workbench application from the supplemental menu of the application. If you want to make changes to the application, navigate to Application Details Settings .

What to do next

Start Cloudera Data Visualization in Cloudera Data Science Workbench.

Starting Cloudera Data Visualization in Cloudera Data Science Workbench

Learn how to start the Cloudera Data Visualization application you have created in Cloudera Data Science Workbench.

Procedure

1. On the Applications page, click the name of your Cloudera Data Visualization application to access the login interface.
2. Log in to Cloudera Data Visualization by entering your username and password. Use your workload credentials.

   If you want to log in as an administrator, you can use the following default credentials:

   - username: vizapps_admin
   - password: vizapps_admin

   When using the default credentials to log in to Cloudera Data Visualization, you are prompted to change your password at the first login.

   Important: If you use the default credentials, security issues may arise. Cloudera recommends that you change the default username and password.

   SSO authentication is disabled by default. See Authentication for information on how to permit user authentication with the Cloudera Data Science Workbench login credentials and log users in automatically.

   After logging in, you land on the homepage view of Cloudera Data Visualization. Here you can explore some sample dashboards or access the in-tool Get Started guide for help.

Connecting to a data source in Cloudera Data Visualization

Cloudera Data Visualization allows you to create connections to many types of external data sources, enhancing your data analysis and visualization capabilities.

Cloudera Data Visualization currently supports the following connection types:

- Hive
- Impala
- MariaDB
- MySQL
- PostgreSQL
- Solr [Technical Preview]
- Spark SQL
- SQLite (not supported in Cloudera Data Warehouse)
- Snowflake [Technical Preview]

This range of supported connection types provides flexibility and versatility in connecting to various data sources for comprehensive data analysis.

When using Cloudera Data Visualization with Cloudera Data Warehouse, the data connection is automatically set up, but you can connect to various other data sources as well to suit your specific requirements.

In Cloudera AI, you can connect to an Impala or a Hive data warehouse, or tie in data from predictive Cloudera AI models.

Cloudera Data Visualization also includes a samples SQLite connection that contains example datasets. This connection is read-only by design, so sample datasets cannot be modified.

Creating a Cloudera AI data connection to Impala

Learn how to connect natively to data stored in Impala when using Cloudera Data Visualization in Cloudera AI.

About this task

Before you start using data modeling and visualization functions, you must connect to your data. The following steps show you how to create a new Cloudera AI data connection in Cloudera Data Visualization to an Impala data warehouse.

Note:

You must have the Manage data connections privilege or be an admin to be able to create and manage connections in Cloudera Data Visualization.

Setting user privileges requires administrator-level access. You can log in as an administrator using the default admin account with the following credentials:

- Username: vizapps_admin
- Password: vizapps_admin

When you create a connection, you automatically get privileges to create and manage the associated datasets. You can also build dashboards and visuals within these datasets.

- For more information on the Manage data connections privilege, see *RBAC permissions*.
- For instructions on how to define privileges for a specific role, see *Setting role privileges*.
- For instructions on how to assign the administrator role to a user, see *Promoting a user to administrator*.

Before you begin

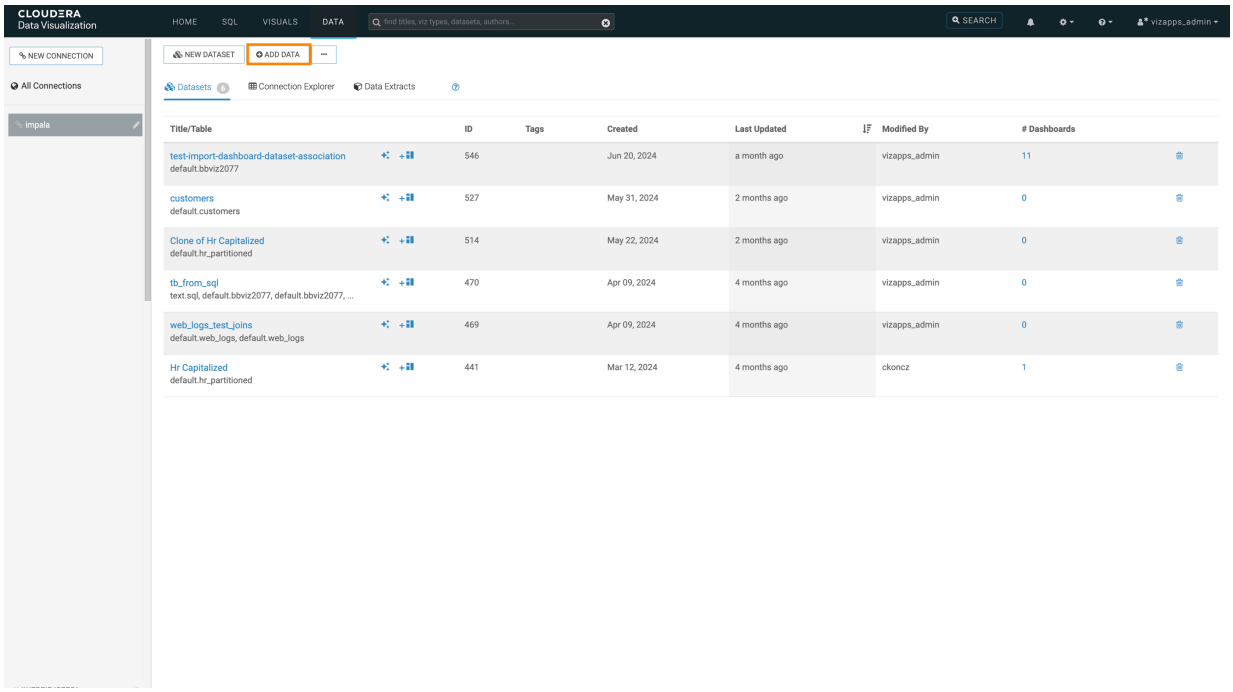
If you are using a Cloudera Private Cloud Base cluster running Impala with Kerberos for authentication, make sure that Kerberos credentials are configured in Cloudera AI before creating a Cloudera AI data connection to the Impala data warehouse. This ensures seamless integration and authentication between Cloudera Data Visualization and the Impala cluster. If you add Kerberos credentials after launching the Cloudera Data Visualization app, you need to restart the app for the changes to take effect.

For more information on using Kerberos for authentication in Cloudera AI, see *Hadoop Authentication for AI Workspaces*.

Procedure

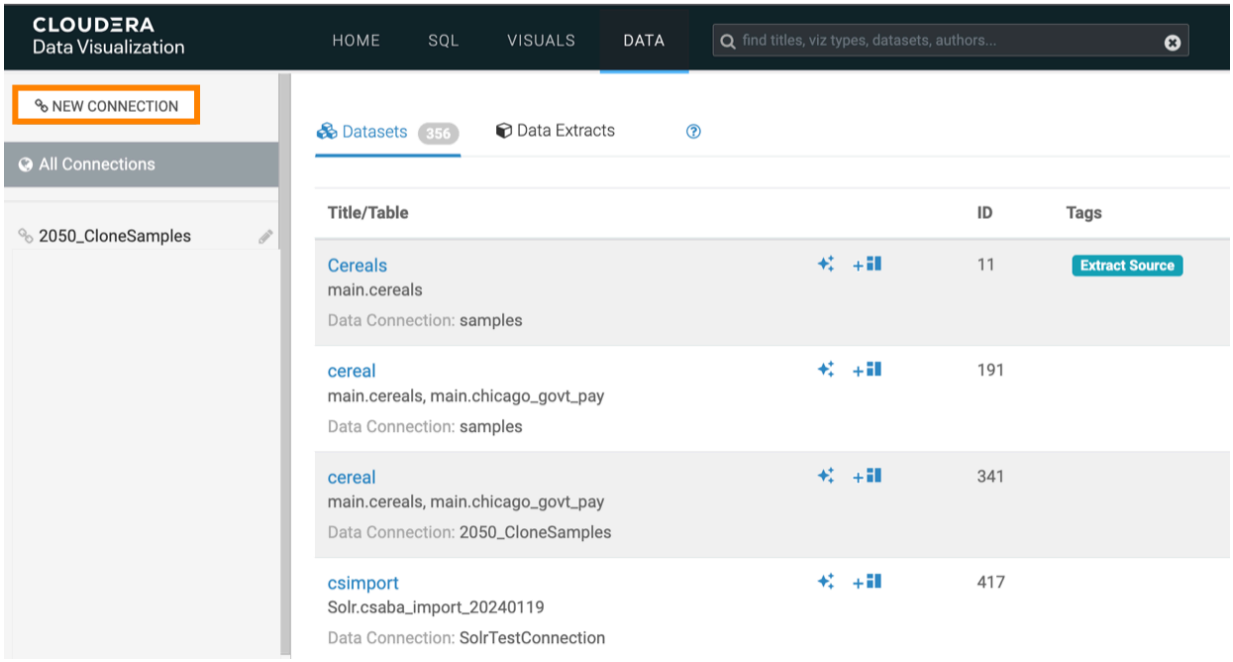
1. On the main navigation bar, click DATA.

   The DATA interface opens, displaying the Datasets tab.



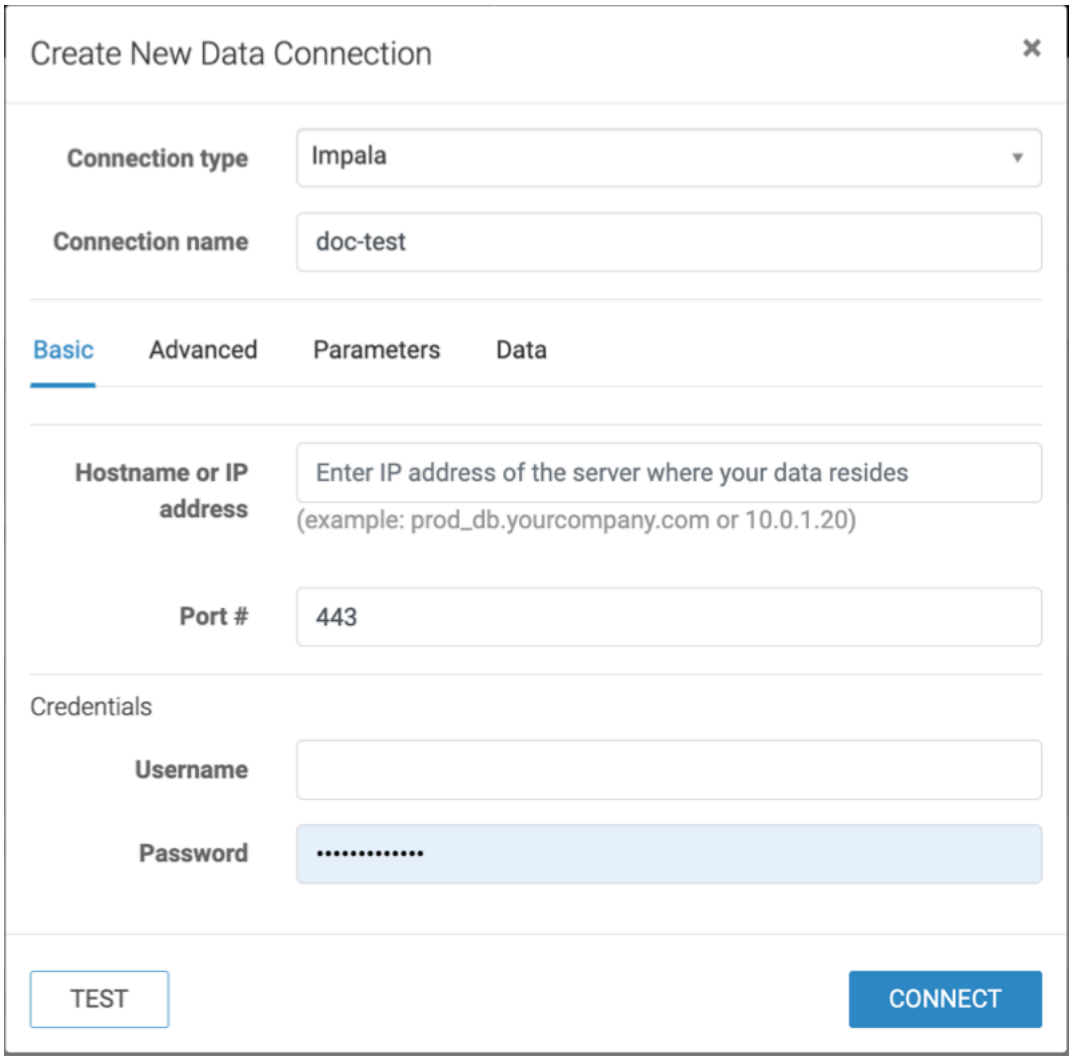
2. On the side menu bar, click NEW CONNECTION.

   Note: Only users with Manage data connections privilege or administrators can access the NEW CONNECTION button.



   The Create New Data Connection modal window appears.

3. Choose Impala from the Connection type drop-down list and assign a name to your connection.



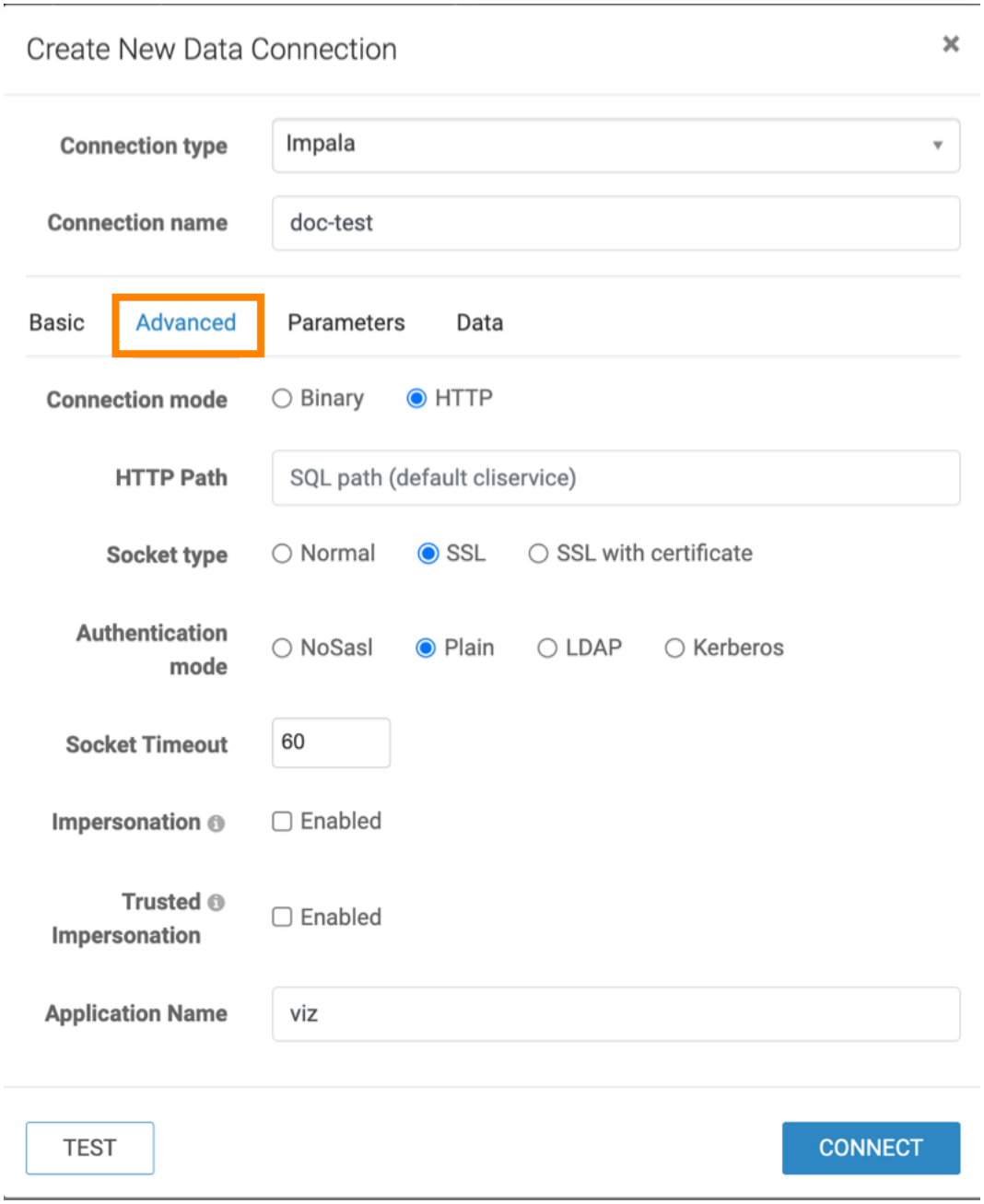
In this example, the Impala connection is made through Knox. Knox always uses TLS encryption and port 443 is the default HTTPS port.

4. Enter the hostname or IP address of the running coordinator.

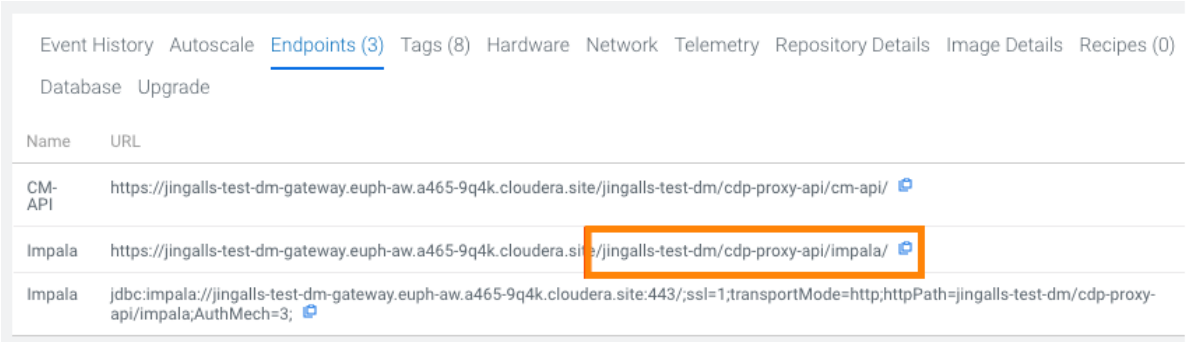
   You can retrieve this information from the JDBC URL of the Impala DW.

5. Add 443 in the Port # field.
6. Enter your workload username and password as credentials.

7. Click the Advanced tab to configure additional details.



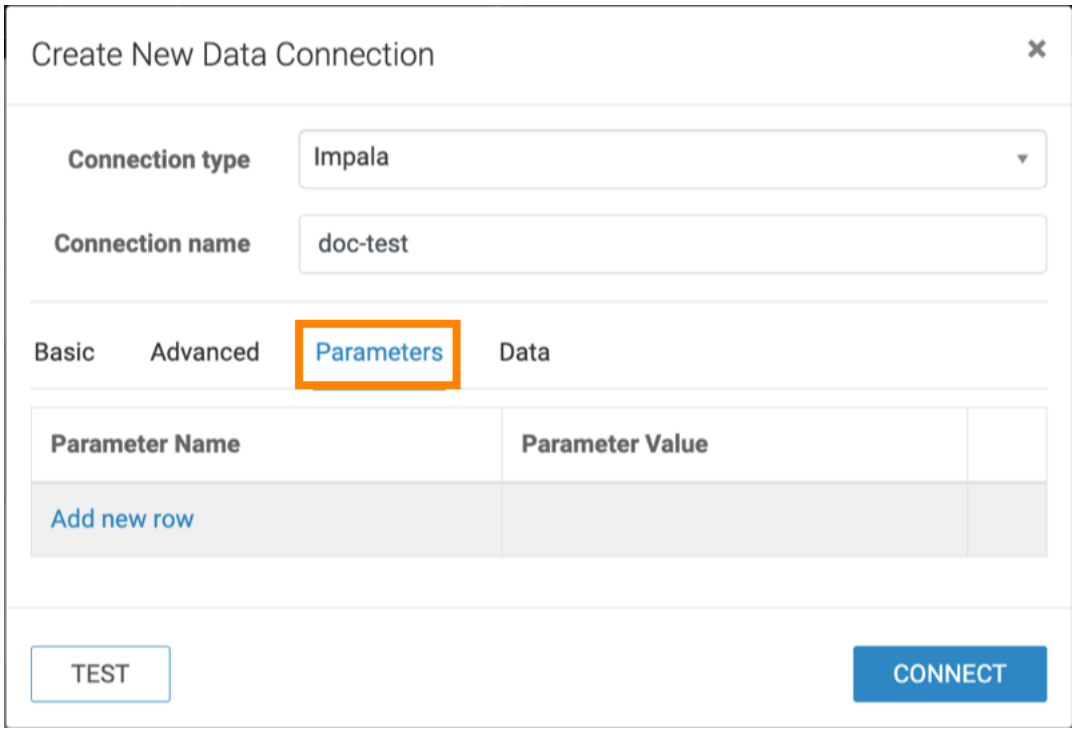
a) For HTTP connection mode, locate the Impala Endpoint for the Data Hub.

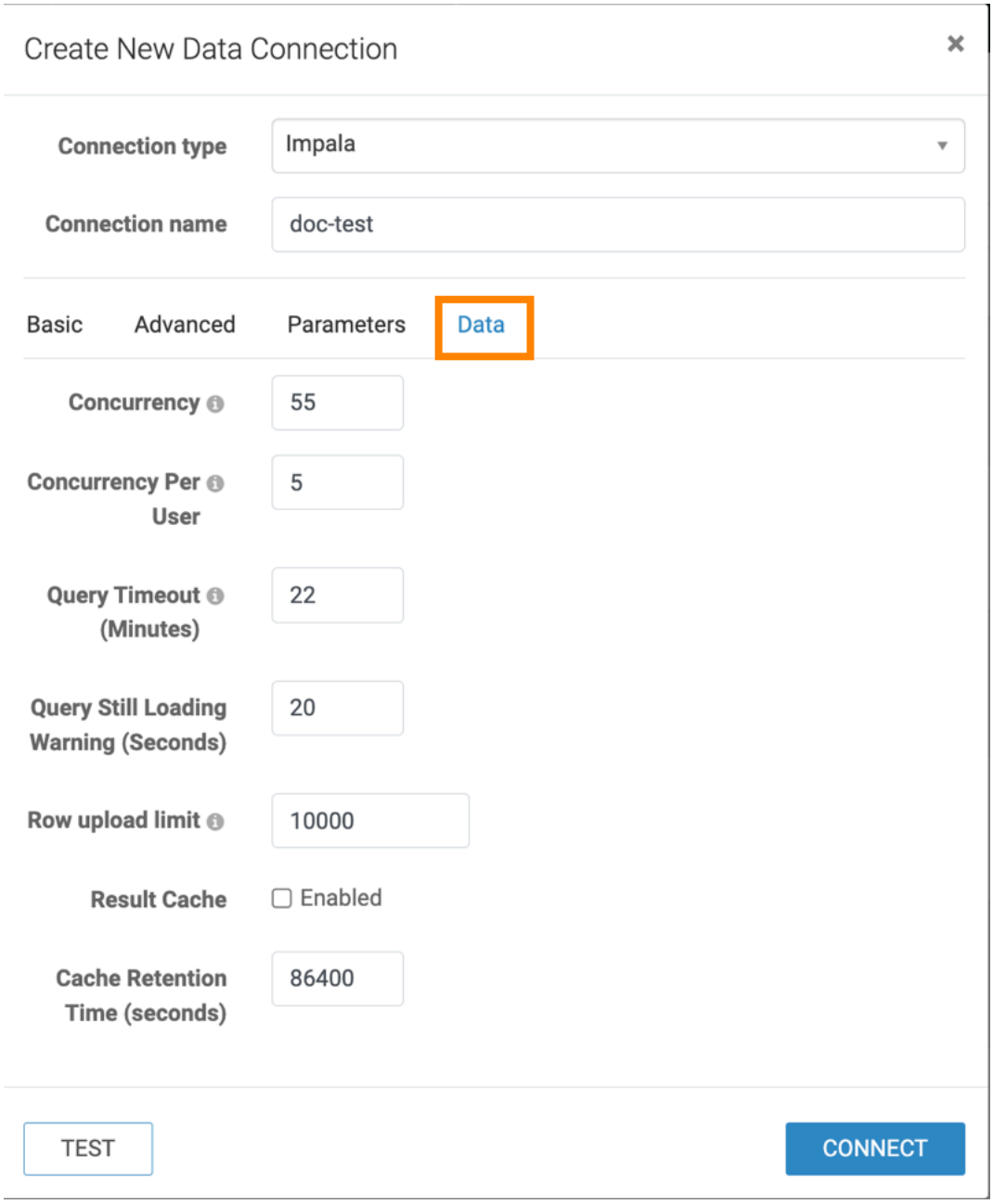


b) Copy and paste it into the HTTP Path field.

c) Set any additional details as required.

8. Check the Parameters and Data tabs for more configuration options.





9. Once you finish configuring the settings, click TEST to check the connection.
10. Click CONNECT to establish the connection.

Results

You have successfully set up a connection to a running Impala DW.

Related Information

RBAC permissions
Setting role privileges
Promoting a user to administrator

Creating a Cloudera AI data connection to a Hive data warehouse

Learn how to connect natively to data stored in Hive when using Cloudera Data Visualization in Cloudera AI.

About this task

Before you start using data modeling and visualization functions, you must connect to your data. The following steps show you how to create a new Cloudera AI data connection in Cloudera Data Visualization to a Hive data warehouse.

Note:

You must have the Manage data connections privilege or be an admin to be able to create and manage connections in Cloudera Data Visualization.

Setting user privileges requires administrator-level access. You can log in as an administrator, using the default admin account with the following credentials:

- Username: vizapps_admin
- Password: vizapps_admin

When you create a connection, you automatically get privileges to create and manage the associated datasets. You can also build dashboards and visuals within these datasets.

- For more information on the Manage data connections privilege, see *RBAC permissions*.
- For instructions on how to define privileges for a specific role, see *Setting role privileges*.
- For instructions on how to assign the administrator role to a user, see *Promoting a user to administrator*.

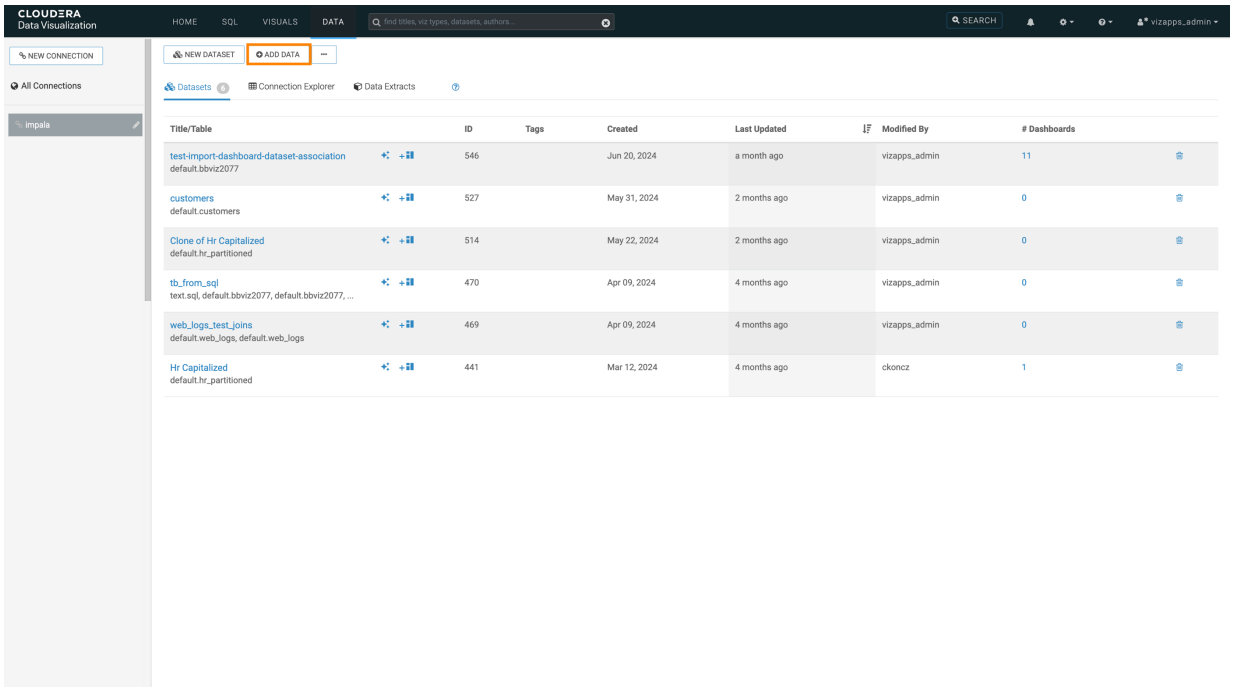
Before you begin

If you are using a Cloudera Private Cloud Base cluster running Hive with Kerberos for authentication, make sure that Kerberos credentials are configured in Cloudera AI before creating a Cloudera AI data connection to the Hive data warehouse. This ensures seamless integration and authentication between Cloudera Data Visualization and the Hive cluster. If you add Kerberos credentials after launching the Cloudera Data Visualization app, you need to restart the app for the changes to take effect.

Procedure

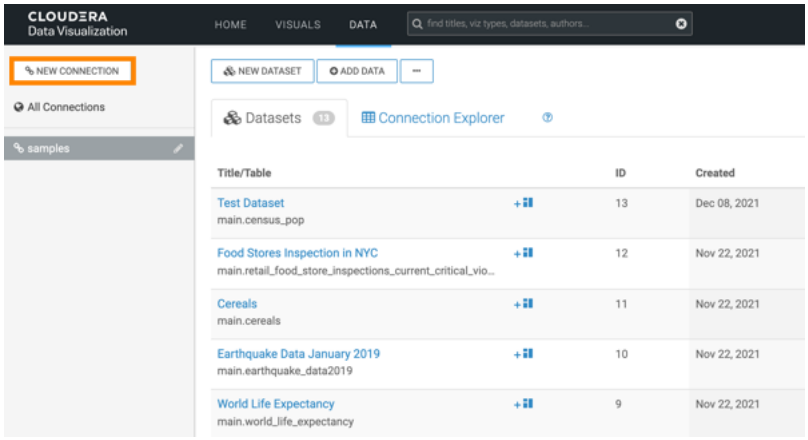
1. On the main navigation bar, click DATA.

   The DATA opens, displaying the Datasets tab.



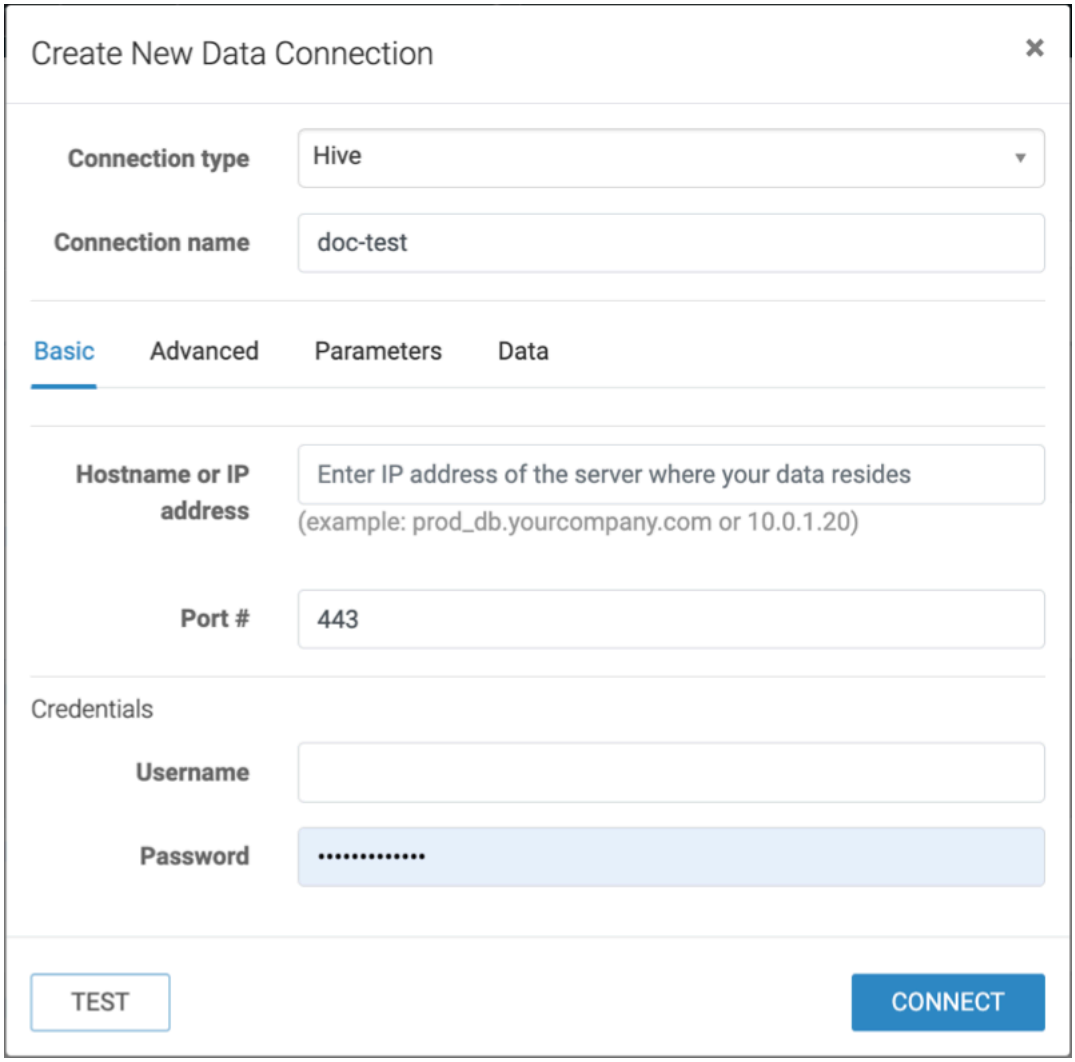
2. On the side menu bar, click NEW CONNECTION.

Note: The NEW CONNECTION button is only accessible to users assigned to roles with the Manage data connections privilege and to administrators.

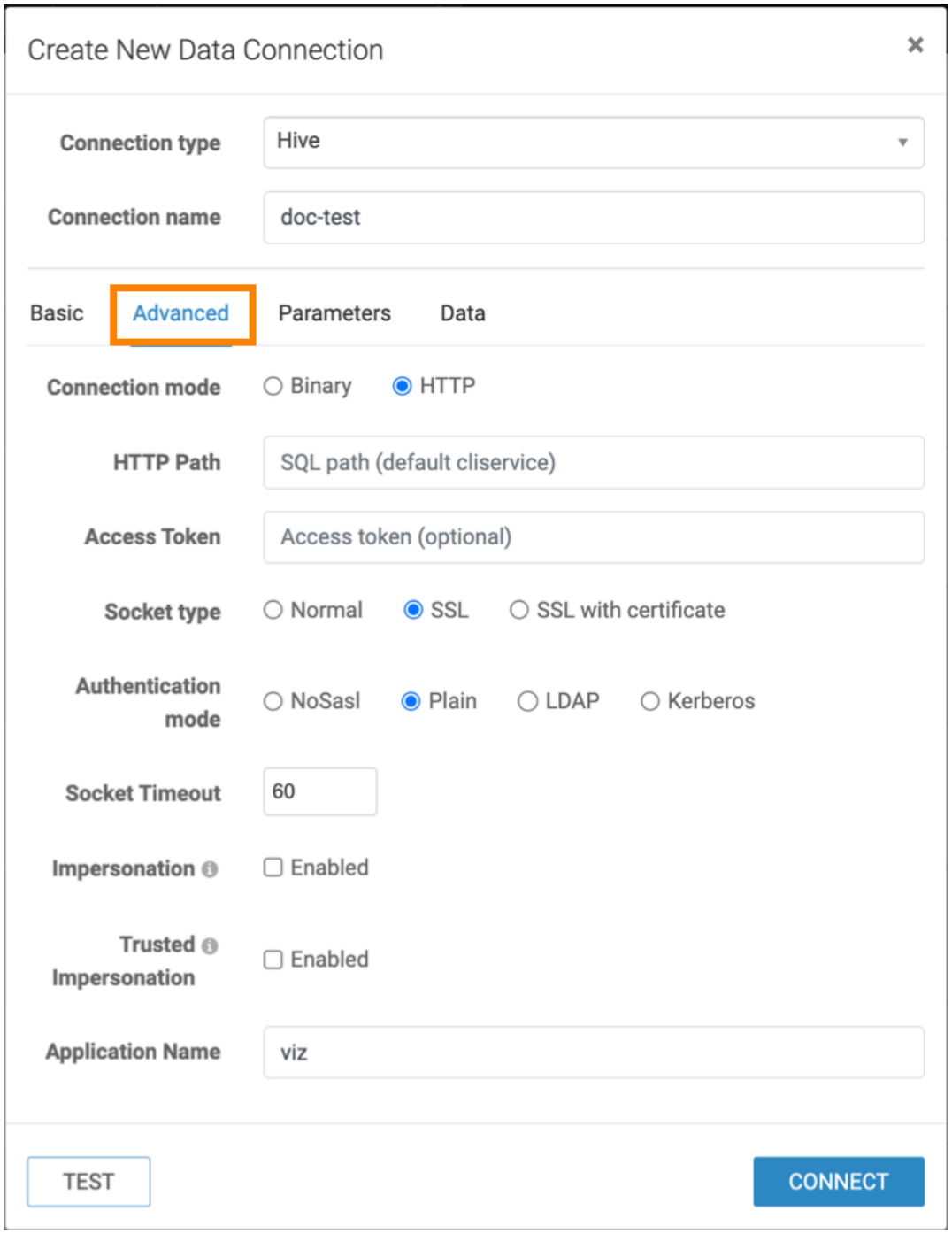


The Create New Data Connection modal window appears.

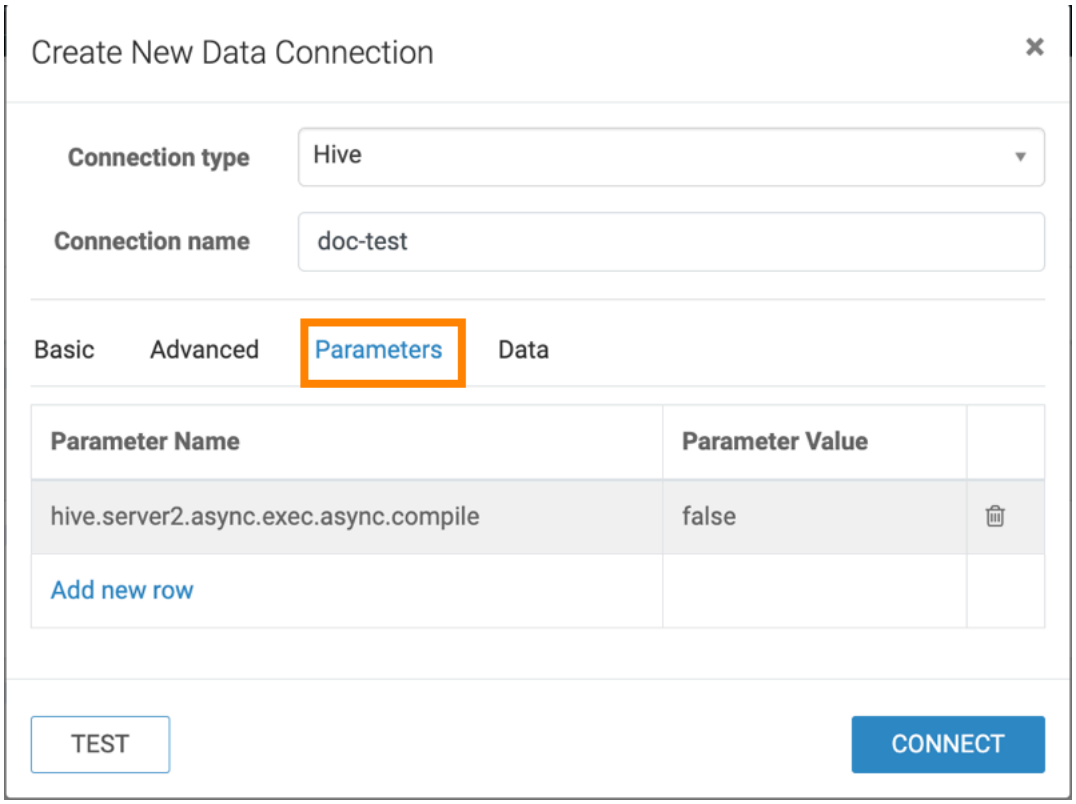
3. Choose Hive from the Connection type drop-down list and assign a name to your connection.



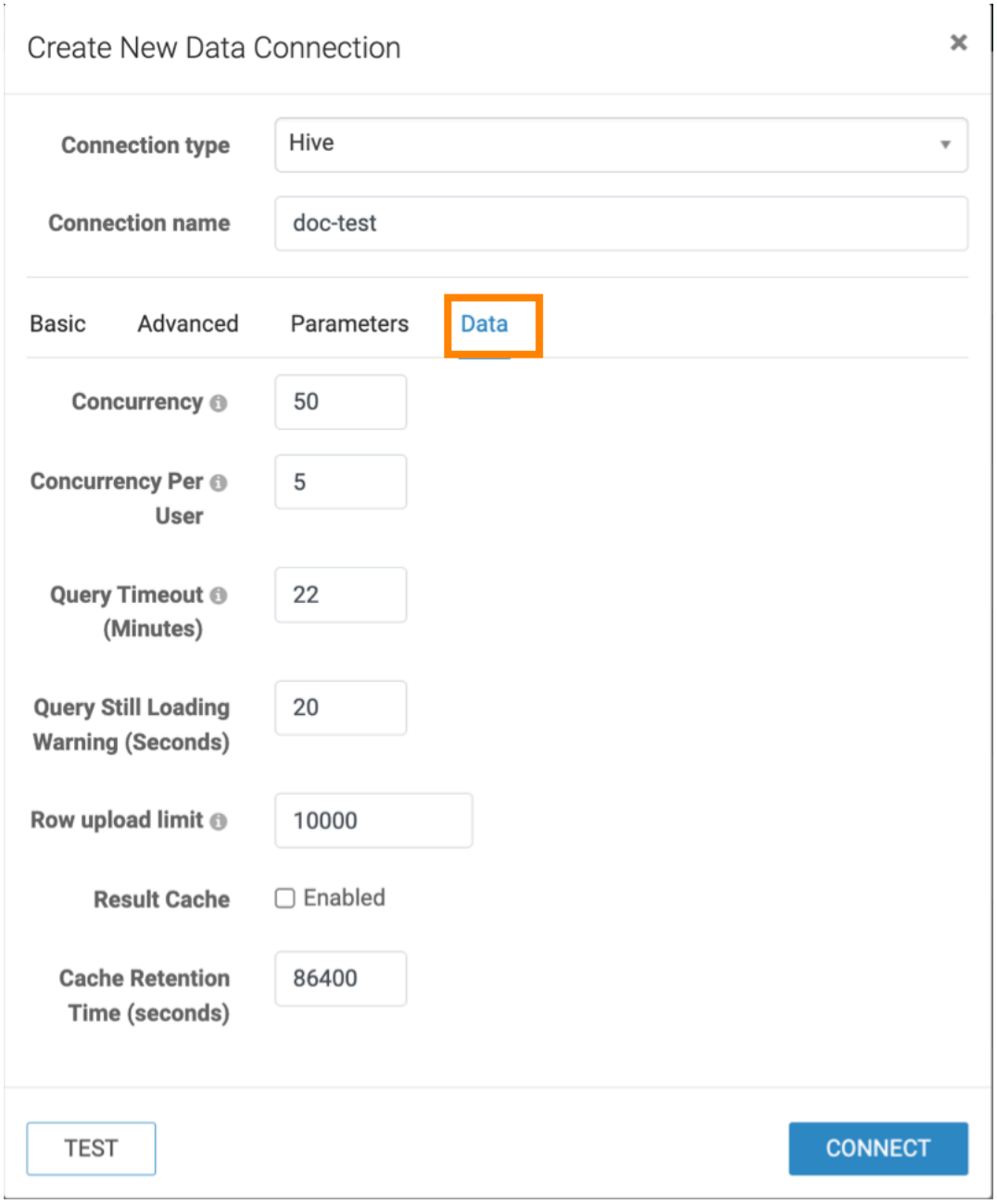
4. Enter the hostname or IP address of the running coordinator.
   You can get the coordinator hostname from the JDBC URL of the Hive DW.
5. Use port 443.
6. Enter your workload username and password as credentials.
7. Click the Advanced tab to configure the additional details.



8. Click the Parameters tab and set the hive.server2.async.exec.async.compile parameter to false.



9. Check the Data tab for more configuration options.



10. Once you finish configuring the settings, click TEST to test the connection.
11. Click CONNECT to establish the connection.

Results

You have successfully set up a connection to a running Hive DW.

Related Information

RBAC permissions
Setting role privileges
Promoting a user to administrator

Creating a Cloudera Data Warehouse data connection in Cloudera Data Visualization

Learn how to connect to data when using Cloudera Data Visualization in Cloudera Data Warehouse data service. You can connect Cloudera Data Visualization to a Virtual Warehouse to visualize your data. Similar to using a BI client, you can configure and connect to Virtual Warehouses from different clusters.

About this task

You must connect to your data prior to using the data modeling and visualization functions. You make the connection to the Virtual Warehouse when you select your warehouse in the steps below. The Cloudera Data Warehouse URL has the same compute instance ID as your Virtual Warehouse.

Note:

To create and manage connections in Cloudera Data Visualization, you must have the Manage data connections privilege or hold administrative privileges. In Cloudera Data Warehouse, these are the members of the Admin Groups associated with the Cloudera Data Visualization instance.

When you create a connection, you automatically gain privileges to create and manage datasets associated with this connection, and to build dashboards and visuals within these datasets.

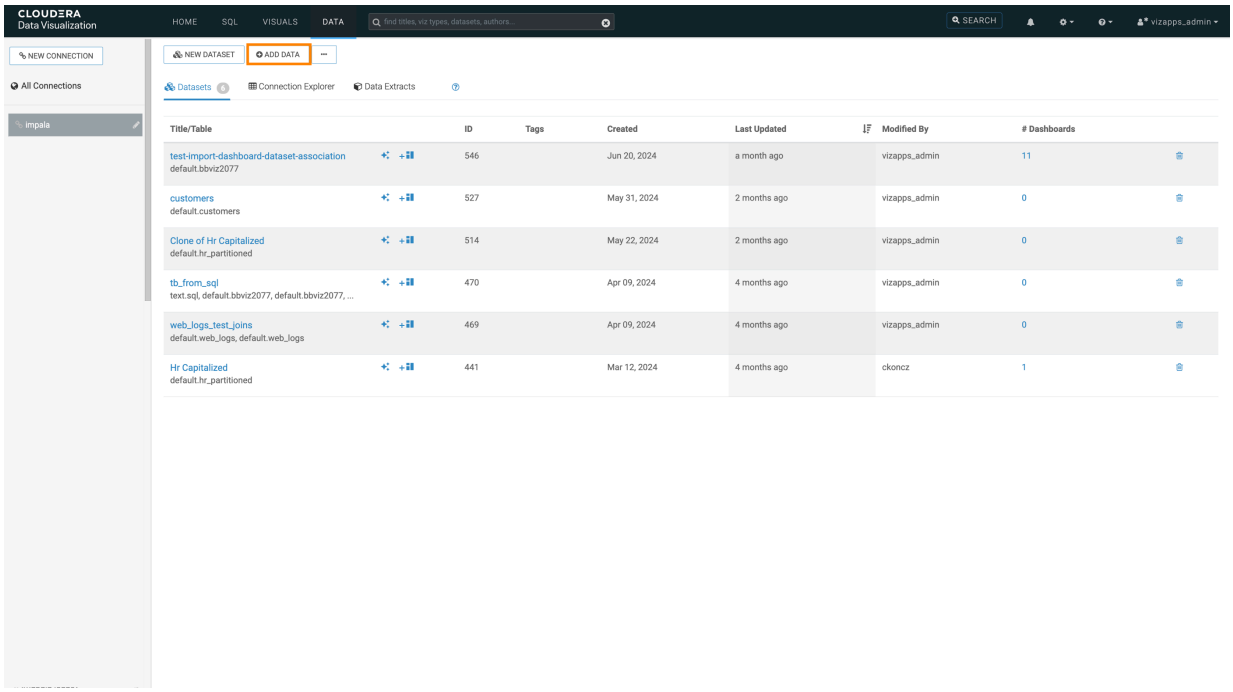
- For more information on the Manage data connections privilege, see *RBAC permissions*.
- For instructions on how to define privileges for a specific role, see *Setting role privileges*.
- For instructions on how to assign the administrator role to a user, see *Promoting a user to administrator*.

When you are creating a Hive or Impala data connection within the same cluster, the connection is considered secure and trusted, and the connection details can be auto-populated with a default authentication user.

Procedure

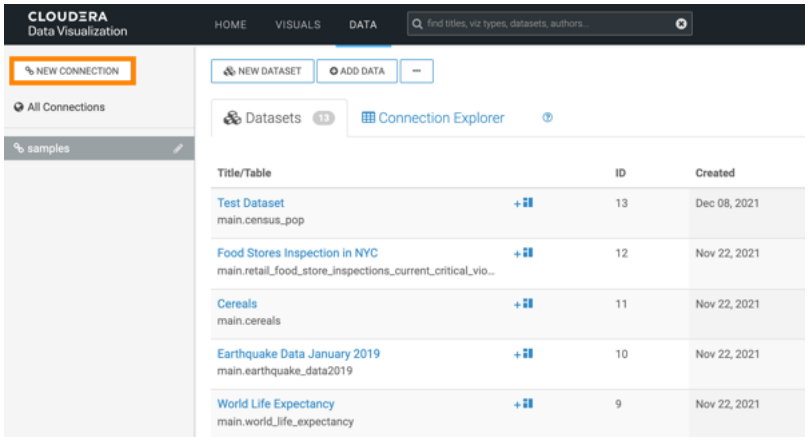
1. Start Cloudera Data Visualization from the left navigation panel in Cloudera Data Warehouse.
2. On the main navigation bar, click DATA.

   The DATA interface appears, open on the Datasets tab.



3. In the side menu bar of DATA, click NEW CONNECTION.

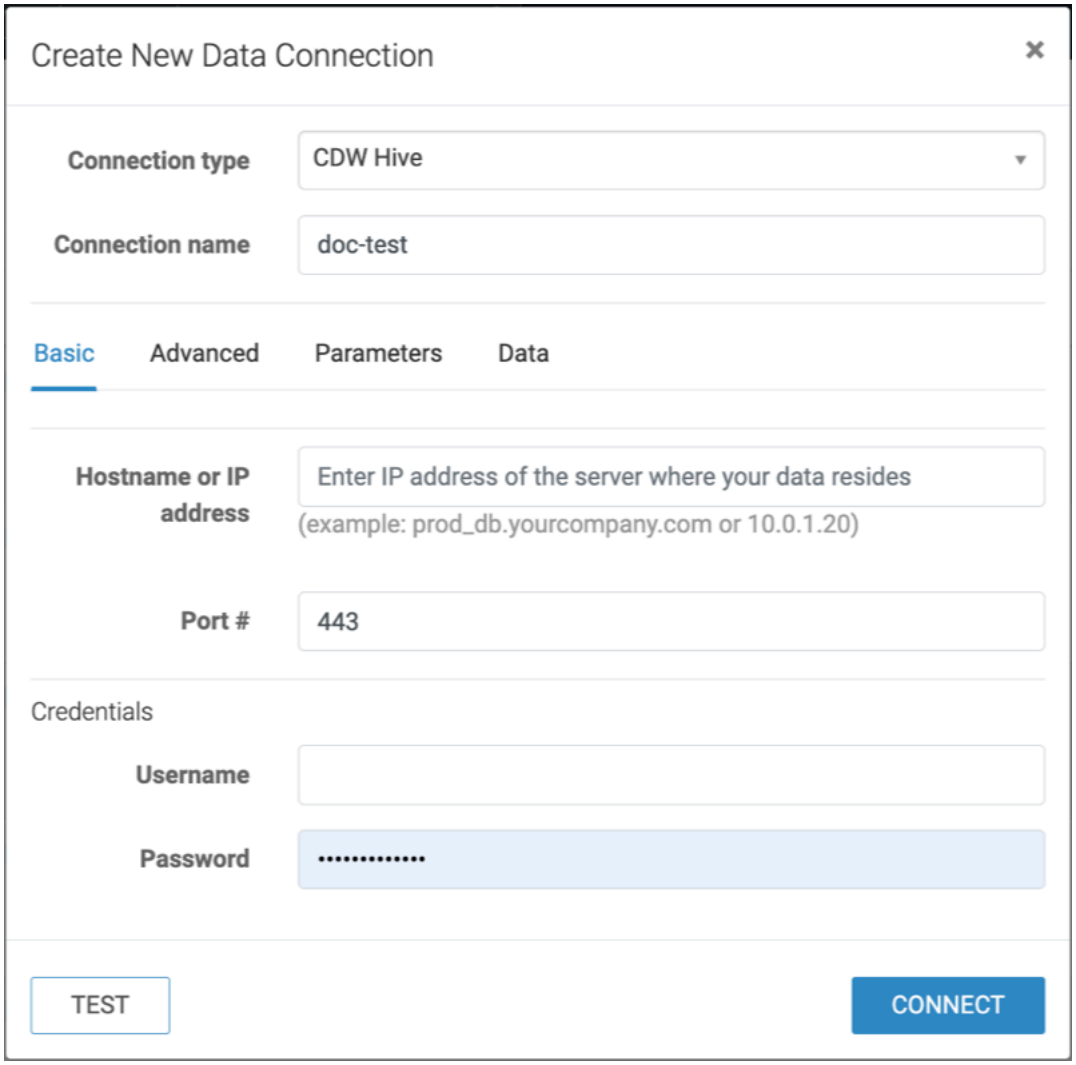
Note: The NEW CONNECTION button is only accessible to users assigned to roles with the Manage data connections privilege and to administrators.



The **Create New Data Connection** modal window appears.

4. In Connection type, select CDW Hive or CDW Impala, and provide a name for your connection.

   Note: SQLite connection is not supported in Cloudera Data Warehouse, and the option is disabled in the connection list.



5. Select a Cloudera Data Warehouse warehouse to connect to.

   For Data connection within the same cluster

   The following fields are auto-populated:

   - Hostname or IP address
   - Port #
   - Username

   For Data connection outside the cluster

   Enter the following information:

   - Hostname or IP address
   - Port #
   - Username
   - Password

6. Click the Advanced tab and configure the additional details.

   Important: Depending on the type of connection you are creating, there can be additional tabs in the Create New Data Connection modal window where you have to adjust further settings.

7. Click TEST.

   If the connection is valid, the system returns a Connection Verified message.

8. Click CONNECT.

What to do next

You can create a data set, and then start creating visuals, dashboards, and applications. For more information, see *Creating datasets* and *Creating a visual*.

Related Information

RBAC permissions

Setting role privileges

Promoting a user to administrator

Creating datasets

Creating a visual

Creating a Cloudera Data Science Workbench data connection to a data warehouse

Learn how to connect natively to data stored in a data warehouse when using Cloudera Data Visualization in Cloudera Data Science Workbench.

About this task

You must connect to your data prior to using the data modeling and visualization functionalities. The following steps show you how to create a new Cloudera Data Science Workbench data connection to a running Impala system.

Note:

To create and manage connections in Cloudera Data Visualization, you must have the Manage data connections privilege or hold administrative privileges.

Setting user privileges requires administrator-level access. To log in as an administrator, you can use the default admin account with the following credentials:

- Username: vizapps_admin
- Password: vizapps_admin

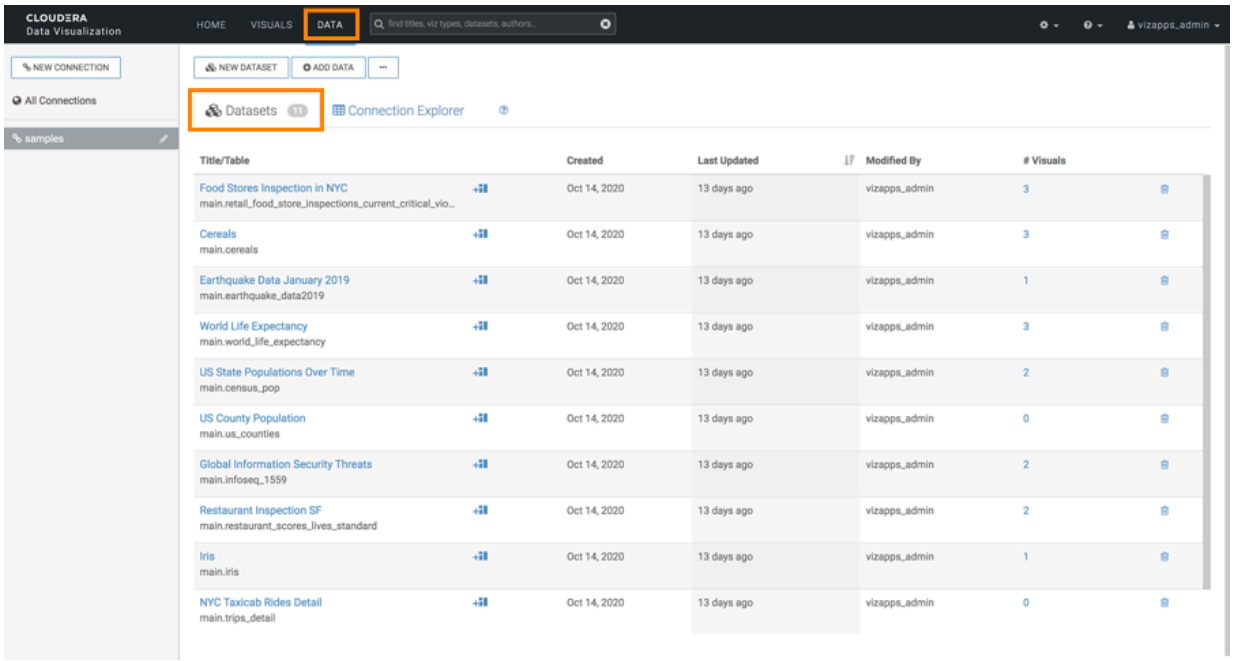
When you create a connection, you automatically gain privileges to create and manage datasets associated with this connection, and to build dashboards and visuals within these datasets.

- For more information on the Manage data connections privilege, see *RBAC permissions*.
- For instructions on how to define privileges for a specific role, see *Setting role privileges*.
- For instructions on how to assign the administrator role to a user, see *Promoting a user to administrator*.

Procedure

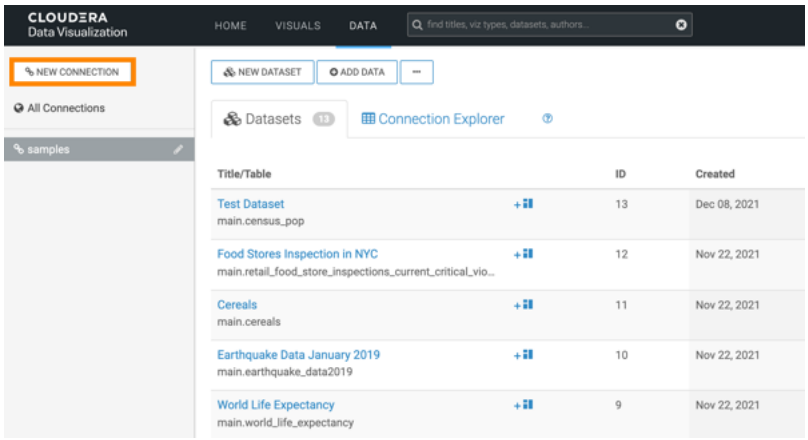
1. On the main navigation bar, click DATA.

   The DATA interface appears, open on the Datasets tab.



2. In the Data side menu bar, click NEW CONNECTION.

   Note: The NEW CONNECTION button is only accessible to users assigned to roles with the Manage data connections privilege and to administrators.



   The Create New Data Connection modal window appears.

3. Select a Connection type from the drop-down list, and provide a name for the connection.
4. Enter the hostname or IP address of the running coordinator.
5. Under Port #, enter the port number.
6. Use your workload username and password as credentials.
7. Click the Advanced tab and make the appropriate selections.

   Important: Depending on the type of connection you are creating, there can be additional tabs in the Create New Data Connection modal window where you have to adjust further settings.

8. Click TEST.

   If the connection is valid, the system returns a Connection Verified message.

9. Click CONNECT.

Results
You have set up a connection to a running data warehouse.

Related Information
RBAC permissions
Setting role privileges
Promoting a user to administrator

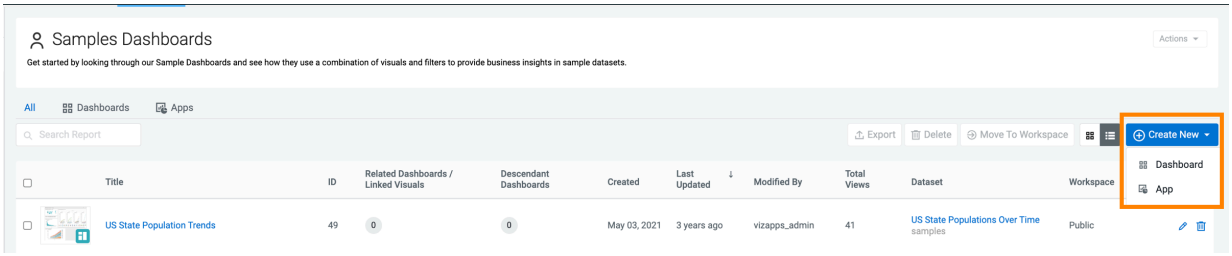
Creating a visual

You can easily create new visual representations for your data in Cloudera Data Visualization.

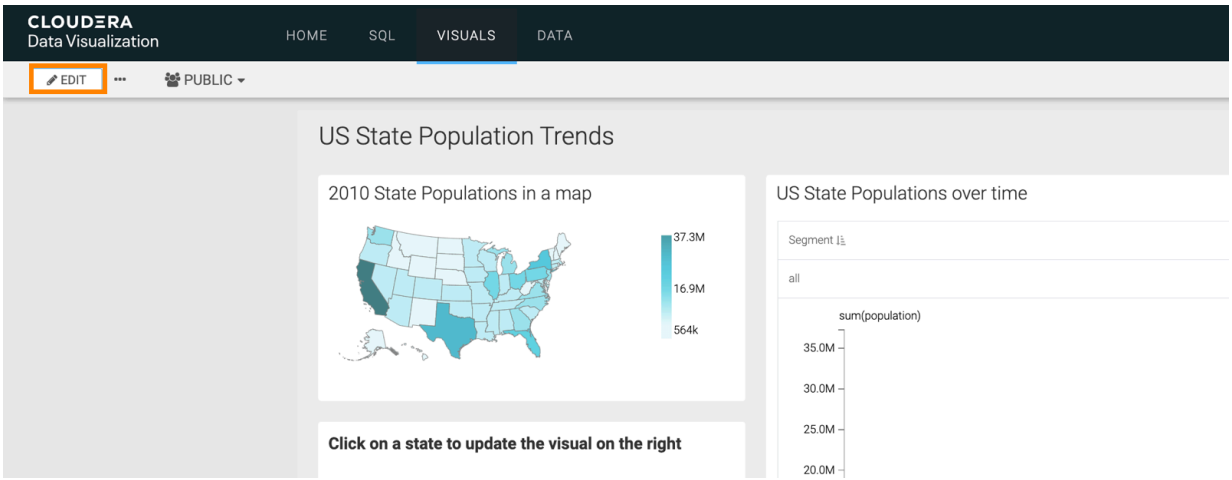
About this task
The following steps demonstrate how to create a new visual.

Procedure

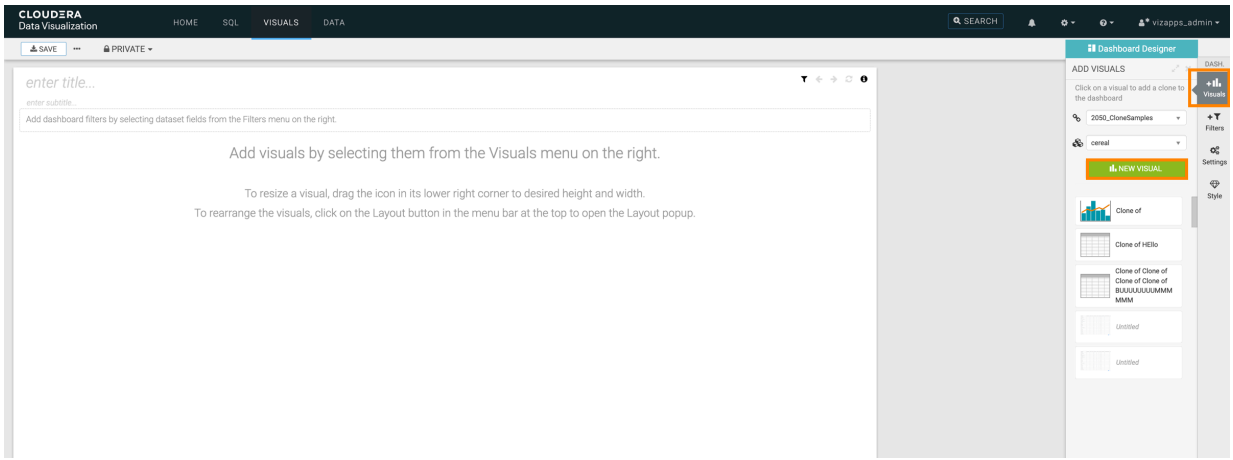
1. On the main navigation bar, click VISUALS.
2. Click Create New Dashboard to open a new, blank dashboard.



If you already have an existing dashboard and you just want to add a new visual, open the dashboard and click EDIT to make changes to it.



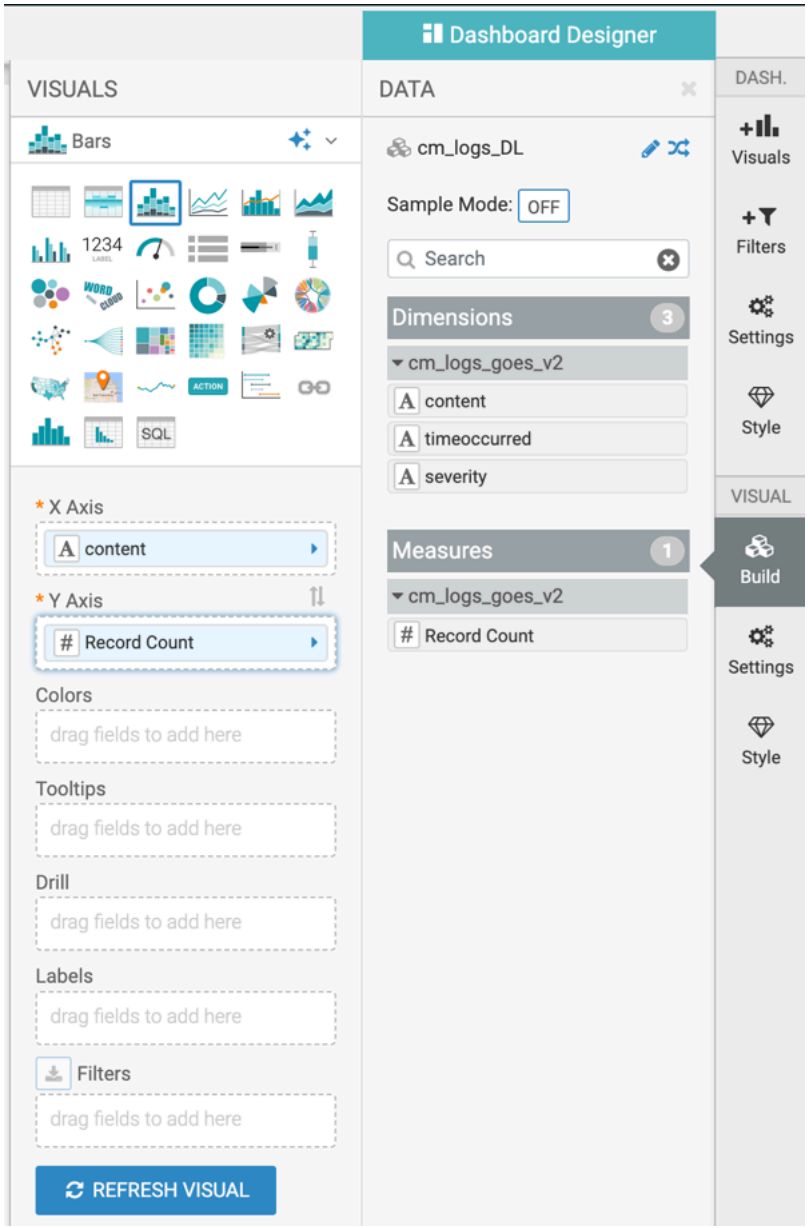
3. In the Dashboard Designer interface, open the Visuals menu from the side menu bar and click NEW VISUAL.



4. The Visual Builder appears.

   By default, a table visual is selected. For more information, see the *Visual Builder* overview.

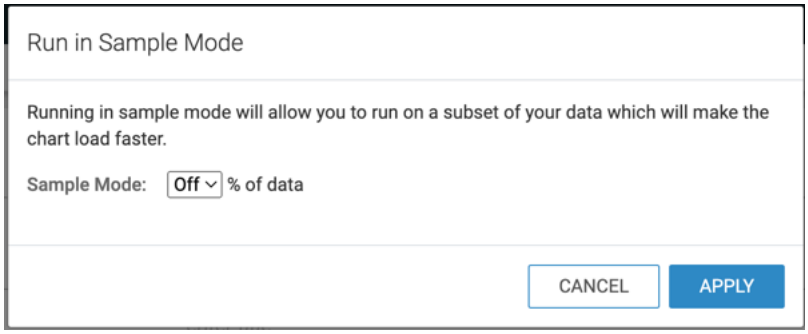
5. Choose the visual type you want to build.
   In this example, the Bars visual has been selected.



6. You can switch on Sample Mode for the visual you are building.

   With sample mode, you create the visual on a subset of your data to test it. You can define this subset as a percentage of the original data.

   Note: This feature is only available on Impala, Hive, and Spark connections.



7. Populate the shelves of the visual from the Dimensions and Measures fields:

   In this example, the following fields have been added:

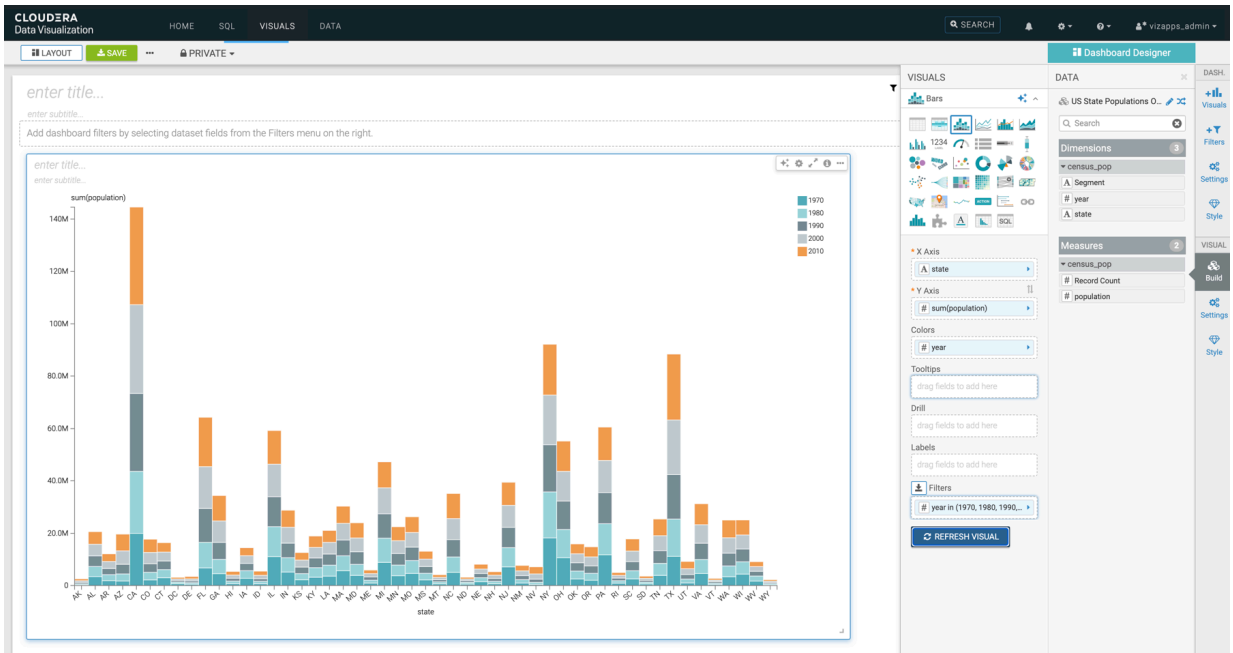
   - state dimension on the X Axis shelf
   - population measure on the Y Axis shelf

     Note: The population measure appears with its default aggregation, as sum(population).

   - year measure on the Colors shelf
   - year measure on the Filters shelf

     Note: The Filter for year module opens automatically. Years 1970, 1980, 1990, 2000, and 2010 have been selected.

8. Click REFRESH VISUAL.



9. Click the enter title... field and enter a name for the visual.
    - In this example the title 'Recent US Census Results by State' has been added.
    - You can also click the enter subtitle... field below the title of the visualization to add a brief description of the visual.
10. Click SAVE at the top left corner of the Dashboard Designer.