Cloudera Data Visualization 8.0.9

Starting Cloudera Data Visualization in Cloudera AI

Date published: 2020-10-30
Date modified: 2025-12-16

https://docs.cloudera.com/

Legal Notice

Contents

Cloudera Data Visualization deployment scenarios in Cloudera AI

You can use Cloudera Data Visualization in Cloudera AI in two ways:

- In embedded mode through the Data Discovery and Visualization application (Data tab)
- In standalone mode deployed as a separate Cloudera AI application

Both deployment options are available in cloud and on-premises environments but they serve different use cases and offer slightly different capabilities.

Data Discovery and Visualization application (Data tab)

This integrated lightweight version of Cloudera Data Visualization is designed for quick, in-context data exploration and visualization directly within your Cloudera AI project. It is ideal for data scientists and ML engineers who need to interactively explore datasets without leaving the Cloudera AI workspace.

Key characteristics:

- Can be accessed from the Data tab within a Cloudera AI project
- Operates in the context of the project's resources and permissions
- Ideal for rapid exploration and iteration on datasets related to Cloudera AI experiments and model development

User interface and admin differences:

- "Site Settings" is renamed to "Explore Data Settings" to reflect its more localized and limited scope within the Cloudera AI project.
- "Advanced Settings" are not available, broader administrative functions are managed at the Cloudera AI application level or by a standalone Cloudera Data Visualization instance.
- Dashboard emailing is not supported in this mode.

  Workaround: To email a dashboard, use a standalone Cloudera Data Visualization application to access the full email functionality.

Note: Permissions are inherited from Cloudera AI project roles, and with synced permissions, the project owner has equivalent permissions as an administrator.

Standalone Cloudera Data Visualization application

This deployment provides the full capabilities of Cloudera Data Visualization, making it suitable for enterprise-grade dashboards, collaboration, reporting and other production use cases.

Key characteristics:

- Deployed as a separate application within Cloudera AI
- Has a dedicated URL and can be accessed independently like other Cloudera AI applications
- Operates independently with its own resources, lifecycle, and user management

Key capabilities:

- Dashboard creation and sharing
- Email support and scheduled reports
- Full access to Site Settings and optionally Advanced Settings (depending on version and configuration)
- Enhanced sharing and administrative controls

Related Information

Deploying a standalone Cloudera Data Visualization application in Cloudera AI

Using Data Discovery and Visualization in Cloudera AI

Deploying a standalone 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.

If you want upgrade your Cloudera Data Visualization Runtime image to a newer version in an air-gapped deployment, you need to manually load the specific Runtime image into the cluster. For general guidance on air-gapped installation see *Updating ML Runtime images on Cloudera AI installations* in the Cloudera AI documentation.

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.

Related Information

Updating ML Runtime images on Cloudera AI installations

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.

Using Data Discovery and Visualization in Cloudera AI

Exploratory data science and visualization makes it simple for data scientists to get started on a data science project. With the Data Discovery and Visualization application, you can automatically discover the data sources available to you from within the standard Cloudera AI user interface and provide various features for exploratory data visualization and analysis.

From the Data tab in Cloudera AI you can:

- Connect to data sources that are available in your project
- Explore the data with SQL to understand its basic shape and characteristics

- Create named datasets that you can reference later
- Create visualizations of the data to understand its properties
- Create dashboards that you can share with your team

For more information on how to work with exploratory data visualization in Cloudera AI, see the Data Discovery and Visualization documentation.

Upgrading Cloudera Data Visualization in Cloudera AI

Cloudera recommends following the Cloudera Data Visualization release cadence by upgrading to the latest version as it becomes available. This practice ensures that your Cloudera Data Visualization instances benefit from the latest new features, security enhancements, and bug fixes.

About this task

This guide outlines the available options performing an upgrade and the steps of the upgrade process. If you need to downgrade, follow the Restoring Cloudera Data Visualization in Cloudera AI procedure using the backup you created earlier.

Before you begin

- Before initiating the upgrade, review the Known issues and limitations in Cloudera Data Visualization in the Release Notes for both the currently running version and the version to which you plan to upgrade to. This helps you identify potential challenges, limitations, or specific actions required for a successful upgrade.
- If you are using an external database like MySQL, MariaDB, Postgres, or Oracle, ensure you back up the database and restore it as a new database that the new Cloudera Data Visualization instance can connect to.

Procedure

1. Stop the application.

   On the Applications page in Cloudera AI, click Actions Stop to stop the selected Cloudera Data Visualization application.
2. Back up the SQLite metadata database.
   a) Navigate to your project's Overview page and click New Session.
   b) When your session is running, click >_ Terminal Access above the session log pane to open a terminal window.

      The default working directory is /home/cdsw, where all your project files reside.
   c) Navigate to /home/cdsw/.arc.

      This directory contains the Cloudera Data Visualization logs, files, settings, and the arcviz.db database.
   d) Run the following command to create a copy of your current Cloudera Data Visualization metadata.

      sqlite3 arcviz.db ".backup arcviz.db-BACKUP-*[***VERSION***]*"

      You can use this backup in case you need to revert to a previous Cloudera Data Visualization version.
3. Navigate to the Application Settings page.
4. Locate the runtime configuration option, where you can select the desired runtime for your application.
5. Select the runtime you want to set for your application from the available options.
6. Click Update Application to save the change.

   The application will restart automatically with the newly selected runtime.

   Note: Allow some time for the restart process to complete.

Restoring a Cloudera Data Visualization version in Cloudera AI

Restoring a previous version means reverting your Cloudera Data Visualization application to an earlier version. You can restore your system to a known stable state, especially when issues arise during an upgrade or if the new version introduces major problems that impact functionality, and you need to return to the previous Cloudera Data Visualization version to maintain system functionality. If you decide to downgrade, follow the steps outlined below.

Before you begin

To restore a previous version, you should have either connected your system to an external database (for example: MySQL, MariaDB, Postgres, Oracle) or created a backup of the SQLite metadata database.

Note:

If you are using an external database like MySQL, MariaDB, Postgres, or Oracle, you must back up the database before the upgrade and restore it as a new database for the Cloudera Data Visualization instance to connect to after the downgrade.

Procedure

1. Stop the application.

   On the Applications page in Cloudera AI, click Actions Stop to stop the selected Cloudera Data Visualization application.
2. Back up the SQLite metadata database.
   a) Navigate to your project's Overview page and click New Session.
   b) When your session is running, click >_ Terminal Access above the session log pane to open a terminal window.

      The default working directory is /home/cdsw, where all your project files reside.
   c) Navigate to /home/cdsw/.arc.

      This directory contains the Cloudera Data Visualization logs, files, settings, and the arcviz.db database.
   d) Run the following command to create a copy of your current Cloudera Data Visualization metadata.

      sqlite3 arcviz.db ".backup arcviz.db-BACKUP-*[***VERSION***]*"

      You can use this backup when reverting to a previous Cloudera Data Visualization version.
3. Create a new Cloudera Data Visualization application in your Cloudera AI project.
4. Copy the backed-up SQLite database to /home/cdsw/.arc/ in the new project.
5. Once the backup is restored, start the new Cloudera Data Visualization application.