

Starting Data Visualization in CDSW

Date published: 2020-10-30

Date modified: 2024-06-28

The Cloudera logo is displayed in a bold, orange, sans-serif font. The word "CLOUDERA" is written in all caps, with a stylized 'E' that has a horizontal bar extending to the right.

Legal Notice

© Cloudera Inc. 2024. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

Deploying a CDV application in CDSW.....	4
Creating a CDSW project with CDV Runtime.....	4
Creating Data Visualization application in CDSW.....	4
Starting Data Visualization in CDSW.....	5
Using CDV in air-gapped CDSW deployment.....	6

Deploying a CDV application in CDSW

Learn how to deploy Cloudera Data Visualization (CDV) in Cloudera Data Science Warehouse (CDSW).

Creating a CDSW project with CDV Runtime

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

Procedure

1. Click Projects on the left sidebar of CDSW.
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 Data Visualization application in CDSW

Learn how to create a Data Visualization application in Cloudera Data Science Workbench (CDSW).

Procedure

1. Go to your CDSW project's Overview page.
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. Make sure it contains URL friendly characters.
- Description – Enter 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 in the location: `/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 CDSW project.
 - Edition – Select the edition of the Runtime variant you want to use for your application.



Note: For Data Visualization, the selected edition automatically defines the version of the Runtime variant.

All available runtimes are listed in the CDSW 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

In a few minutes, you should see the application status change from Starting to Running on the Applications page. Your application has a hosted, fully-functional Data Visualization platform.

You can restart, stop, or delete a CDSW application from the supplemental menu of the application. If you want to make changes to the application, click Application Details and go to the Settings tab.

What to do next

Start Data Visualization in CDSW.

Related Information

[Application Limitations](#)

Starting Data Visualization in CDSW

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

Procedure

1. On the Applications page, click the name of your Data Visualization application to access the login interface.

2. Log in to 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 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 CDSW 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.

Related Information

[User interface overview](#)

[Data Visualization quickstart](#)

Using CDV in air-gapped CDSW deployment

New Cloudera Data Visualization (CDV) Runtime releases are automatically added to your deployment when an internet connection is available. However, in air-gapped Cloudera Data Science Workbench (CDSW) deployments, if you want to use CDV, you need to manually load the CDV Visualization Runtime in the cluster.

Before you begin

Ensure you have the following:

- CDSW 1.9.0 or higher to support runtimes
- Root access to all cluster nodes
- CDSW installed as a parcel
- Admin access on the CDSW cluster
- Proficiency in Docker, SQL, and Linux

Procedure

1. Download the repo-assembly.json file from the [‘artifacts’ directory of the latest CDV version](#).
2. Download the Docker image to an Internet-connected node or your local machine.

```
image_identifier=$(jq -r '.runtimedataviz[0].image_identifier' repo-assembly.json)
docker pull "${image_identifier}"
```

3. Load the Docker image to all cluster nodes (master and all workers) using the `docker save` and `docker load` commands.

```
docker save -o runtimedataviz.tar <image_name>
docker load -i runtimedataviz:image_identifier
```

4. Verify that the Docker image is available on all nodes and has its original name and tag using the `docker images` command.

You will get a summary view of the Docker images, showing details such as the repository name, tag, image ID, creation date, and size.

5. Add the Cloudera Data Visualization image as a custom runtime, using the original Docker image name.

For example: `docker.repository.cloudera.com/cloudera/cdv/runtimedataviz:7.1.2-b53`

For detailed instructions, see [Adding New ML Runtimes](#).