Visualizing Data in Cloudera Data Warehouse Public Cloud (Preview)

Date published: 2021-08-27
Date modified: 2021-08-27

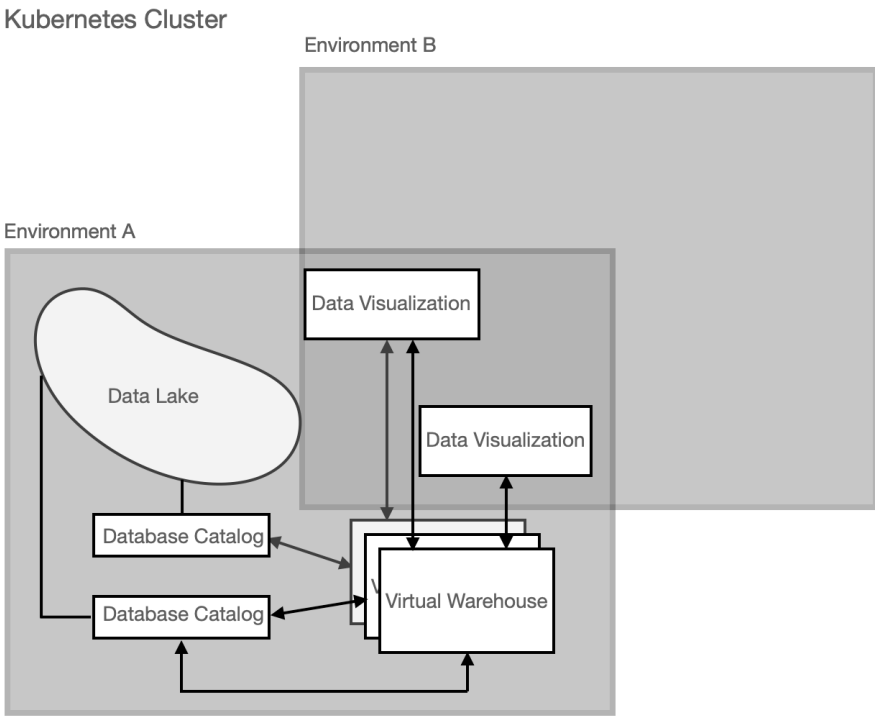
Legal Notice

Contents

Introduction

CDW integrates Data Visualization for building graphic representations of data, dashboards, and visual applications based on CDW data, or other data sources you connect to. You, and authorized users, can explore data across the whole CDP data lifecycle using graphics, such as pie charts and histograms. You arrange visuals on a dashboard for collaborative analysis.

In CDW, BI analysts who can access your environment can use these features. Data Visualization is not tied to a particular Virtual Warehouse (VW). You can access data for your visualization from multiple Data Catalogs using multiple Hive or Impala Virtual Warehouses and multiple environments.



Having multiple Data Visualization instances attached to an environment, you can create dashboards for different groups. For example, Marketing and Sales can have their own private dashboards. When you delete a Virtual Warehouse, your visuals remain intact.

Secure access to visuals

The visuals that represent your data are as important to secure as your data. You need to understand how CDP security protects your visuals.

Database catalogs and virtual data warehouses inherit the security restrictions that apply to your CDP environment. You do not need to set up security for each Database Catalog or Virtual Warehouse.

If you cannot see the data in a connected Database Catalog, check and adjust permissions for your environment and Data Warehouse.

Related Information: Creating a visual

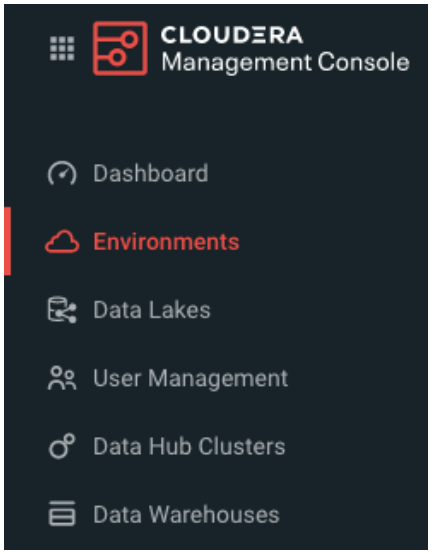
Creating Data Visualization in CDW

You need to create a Data Visualization instance if Data Visualization is not running in your environment. You can create multiple instances of Data Visualization to connect to different data sets and to share with different users. You connect an instance to a Database Catalog to ultimately create visuals of data in Hive or Impala Virtual Warehouses.

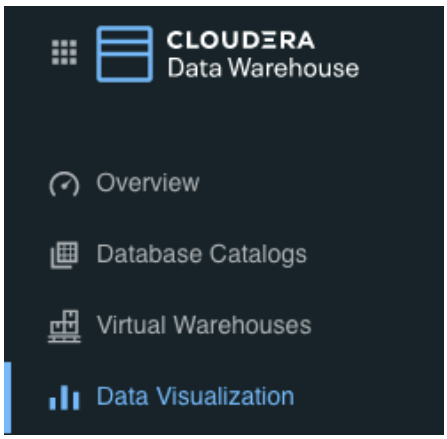
Before you begin: You have obtained a Data Visualization account.

1. Log into the CDP web interface and click Data Warehouse.

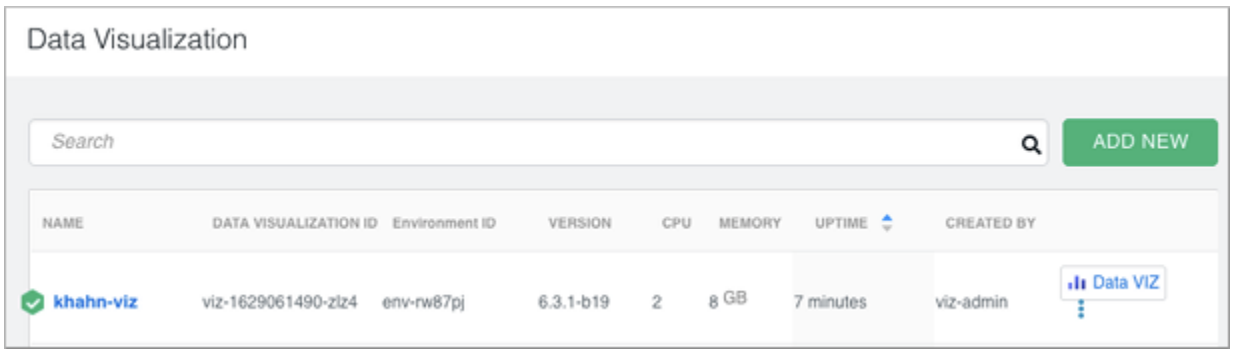
   Alternatively, if you are already logged in, in Cloudera Management Console, click Data Warehouses:



2. In Cloudera Data Warehouse, click **Data Visualization**.



A list of existing Data Visualization instances and environments appears.



3. In Data Visualization, click **Add New** to create a new instance.

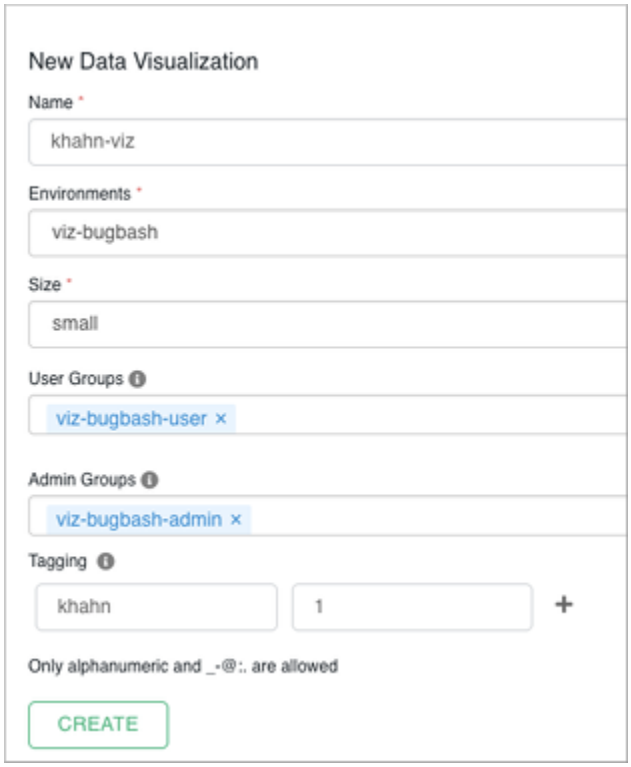
4. In New Data Visualization, specify a **Name** for the instance and select your environment.

   For example:

5. Select the **Size** of the compute instance for Data Visualization and corresponding auto-scaling thresholds in CPU | GB.

6. In **User Groups,** select who can access visuals.
7. In **Tagging,** enter keys and values to identify Data Visualization usage for billing and metering.

8. In **Admin Groups**, select who can configure Data Visualization.



9. Click **Create**.
10. From the list of Data Visualization instances and environments, click Data VIZ.

    Data Visualization opens.

    Next steps: Connect to a Virtual Warehouse to create visuals, dashboards, and applications.

Related Information: Data Visualization Documentation

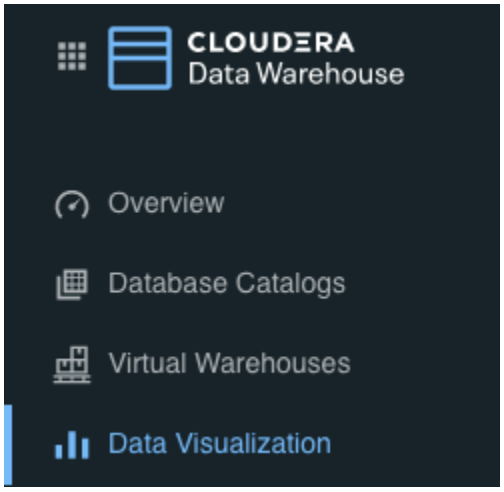
Opening Data Visualization

Opening Data Visualization to connect to data and create visuals is intuitive. You will discover more than one way to do so in CDW Data Visualization.

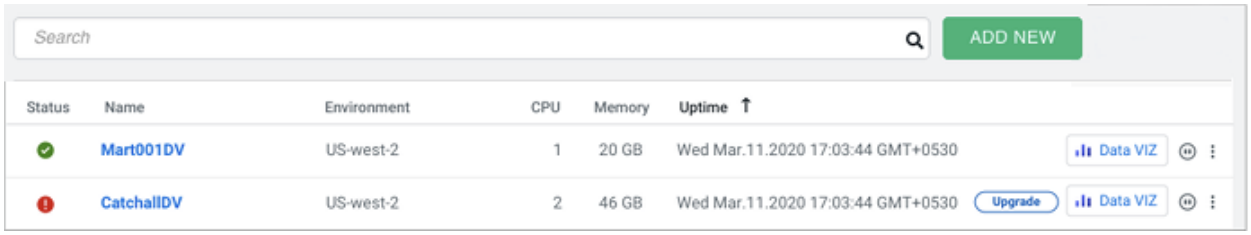
Before you begin:

- You have obtained a Data Visualization account.
- One or more Data Visualization instances are running in CDW.

1. Log into the CDP web interface and click Data Warehouse.
2. In Cloudera Data Warehouse, click **Data Visualization**.



A list of Data Visualization instances and environments appears.



3. From the list of running Data Visualization instances and environments, click Data VIZ.

Data Visualization opens.

Next steps: Connect to a Virtual Warehouse to create visuals, dashboards, and applications.

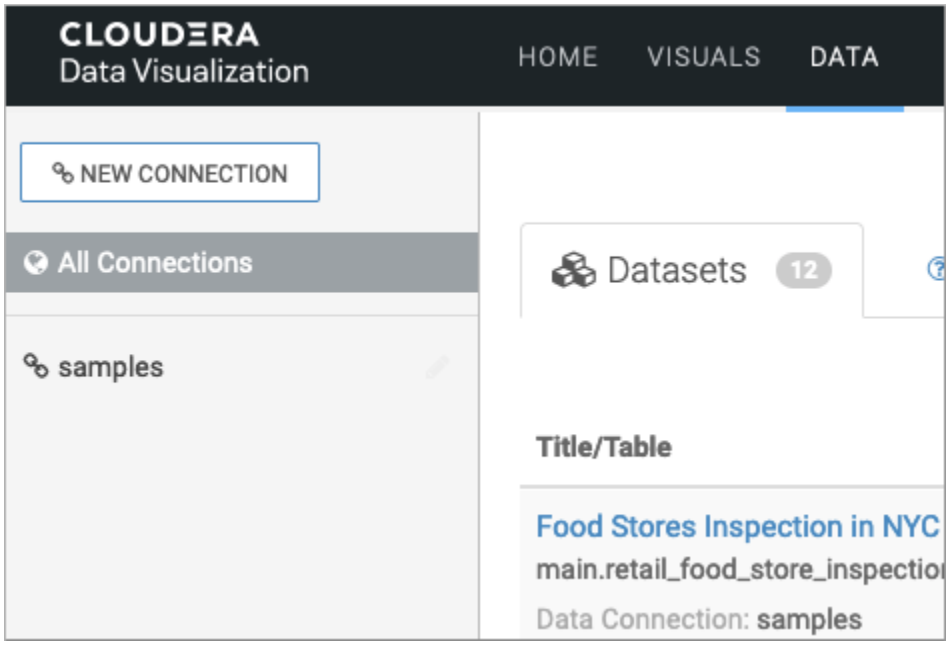
Connecting Data Visualization to data

As a BI Analyst, you can connect 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. You configure the connection in a familiar way, providing an IP address or host name.

About this task:

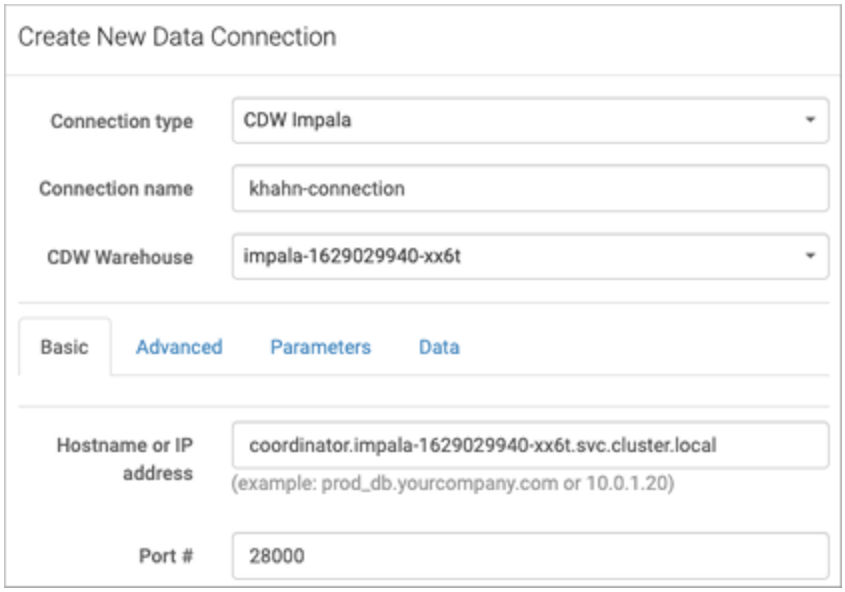
You make the connection to the Virtual Warehouse when you select your CDW Warehouse in the steps below. The CDW Warehouse URL has the same compute instance ID as your Virtual Warehouse.

1. Open Data Visualization.
2. Click **Data**.



3. Click **New Connection**.
4. In **Connection type**, select CDW Hive.
5. In **Connection Name**, enter an arbitrary name for the connection.
6. Select a CDW Warehouse to connect to.

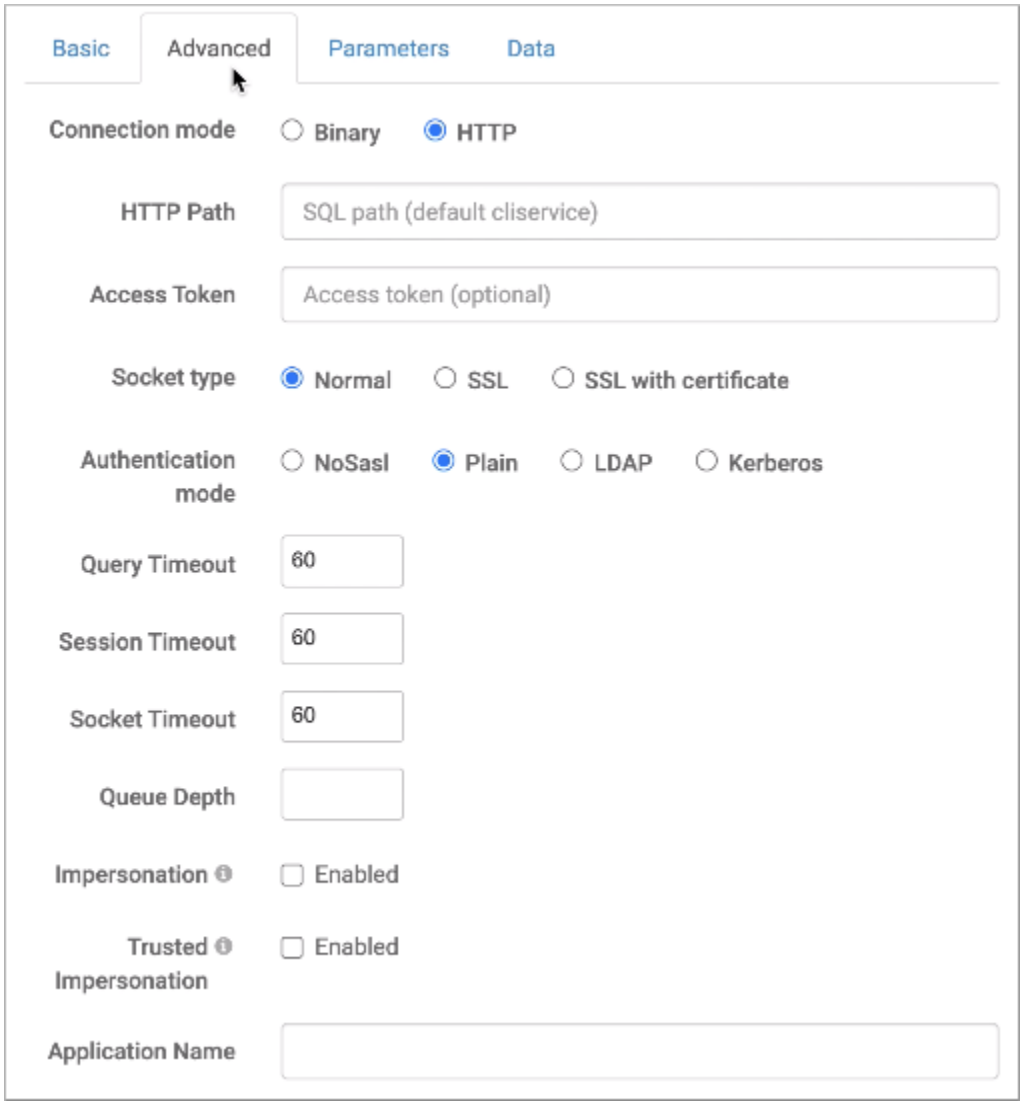
   The Hostname or IP address of the CDW Warehouse and the port number appear in the Basic tab.



7. In **Credentials**, complete the following entries:

    - Under **Username**, enter the username for establishing the connection.
    - Under **Password**, enter the password for establishing the connection.

8. On the Advanced Tab, set up the connection mode, authentication, and timeouts.

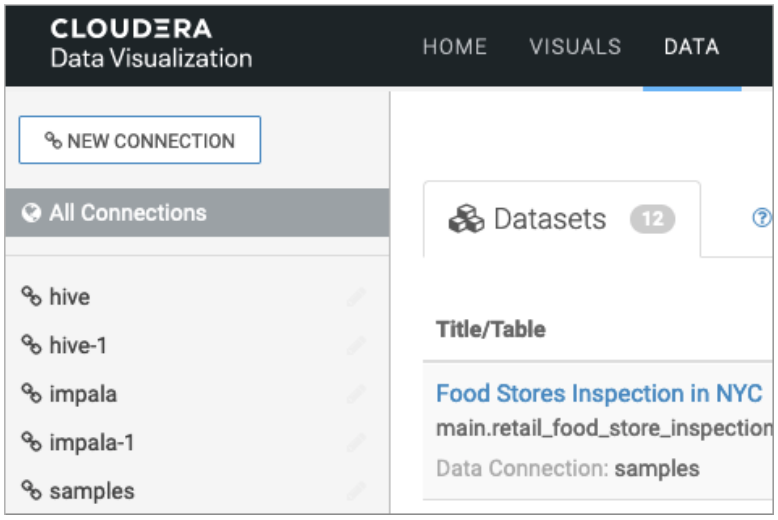


9. Click **Test**.

   If the connection is valid, the system returns a verified connection message.

10. Click **Connect**.

    The CDW data sources you are connected to appear.



Next steps: You create a Data Set, click **Visuals**, and start creating visuals, dashboards, and applications as described in the Data Visualization documentation.

Related Information:

- Creating a data set
- Creating a visual

Editing and deleting Data Visualization instances

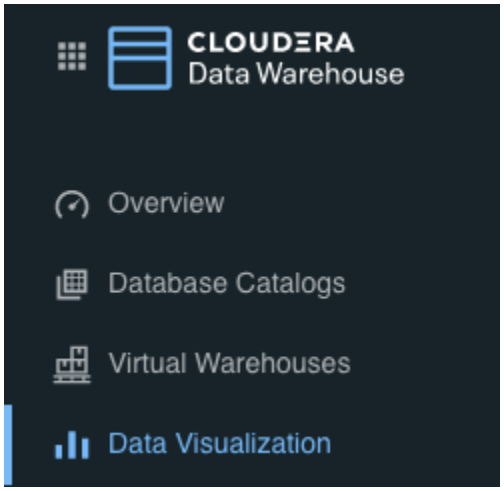
You see how to edit and delete Data Visualization instances.

Before you begin:

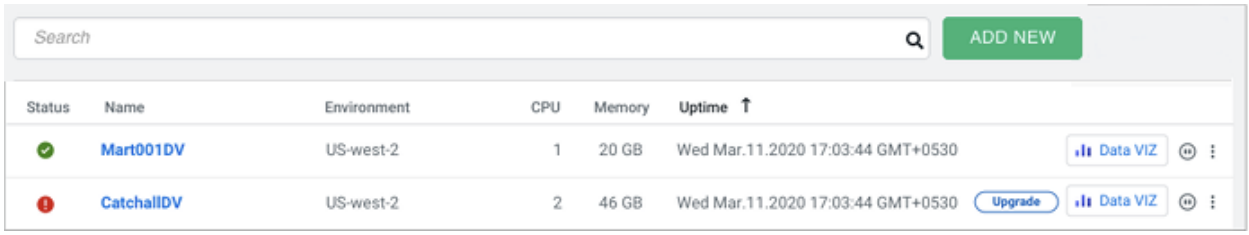
You have created the instance you want to edit or delete, or you have Admin permissions to access the instance.

1. Log into the CDP web interface and click Data Warehouse.

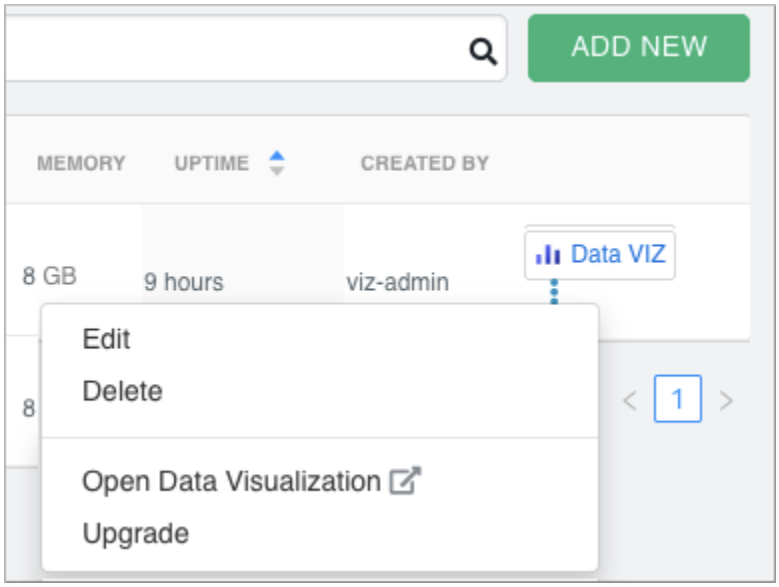
2. In Cloudera Data Warehouse, click **Data Visualization**.



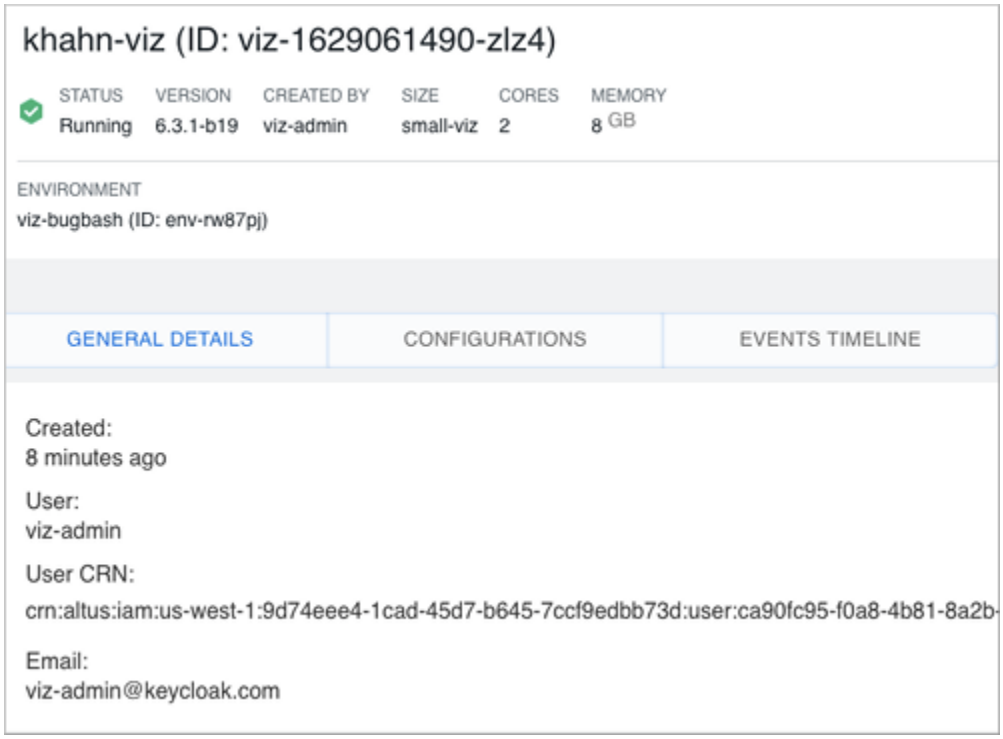
A list of Data Visualization instances and environments appears.



3. In one of the listed Data Visualization instance rows, click options .



4. Select the option to edit a Data Visualization instance.



5. Make changes to general details and other properties.

6. Click **Apply**.

7. To delete a Data Visualization instance, click options and select **Delete**.

   The visuals, dashboards, and applications you created are not deleted when you delete the instance.

Upgrading Data Visualization

You can easily upgrade Data Visualization if a later version is available in Cloudera Data Warehouse (CDW) Public Cloud.

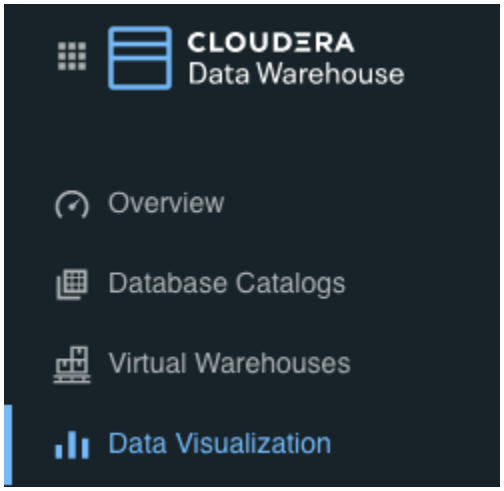
Data Visualization is independent and upgrading is easy. Upgrading Data Visualization involves no choices. Cloudera makes the latest compatible Data Visualization available on S3, and you can choose to use it.

Before you begin:

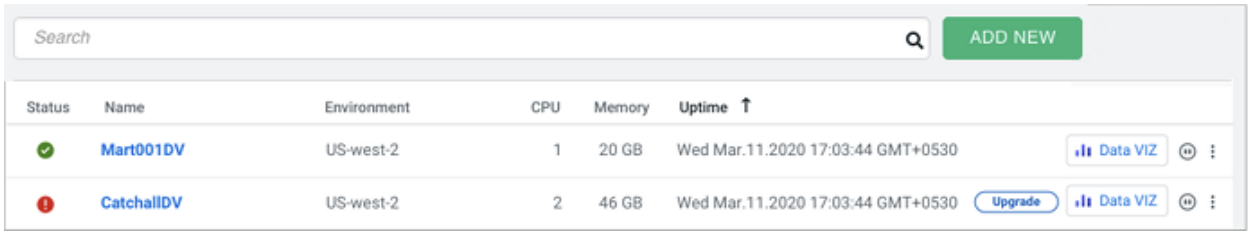
- You have obtained a Data Visualization account.
- You are logged into CDP and can access the environment of the Data Visualization to upgrade.

Required role: DWAdmin

1. In Cloudera Data Warehouse, click **Data Visualization**.



A list of Data Visualization instances appears. If Cloudera makes a later version available for any listed instance, you can upgrade that instance.



2. Click **Upgrade.**