

Natural Language Search

Date published: 2020-10-30

Date modified: 2024-02-29

CLOUDERA

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

Natural language search UI.....	4
Search results.....	5
Order of search results.....	6
Search bookmarks.....	6
Search history.....	8
Search types.....	8
Hit highlighting.....	8
Match field names.....	8
Matching suggested questions.....	9
Bookmarked searches.....	9
Search history.....	10
View more searches.....	10
Global search.....	11
Global search supported keywords.....	12

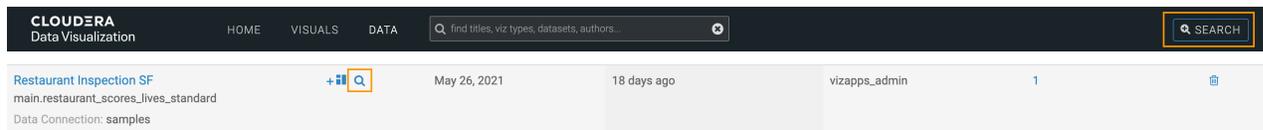
Natural language search UI

Cloudera Data Visualization enables you to find relevant information in the dataset through natural language statements, and renders the result in a graphical format. Using this type of search you can type in a sentence, using regular, spoken language, that describes the information you are looking for.

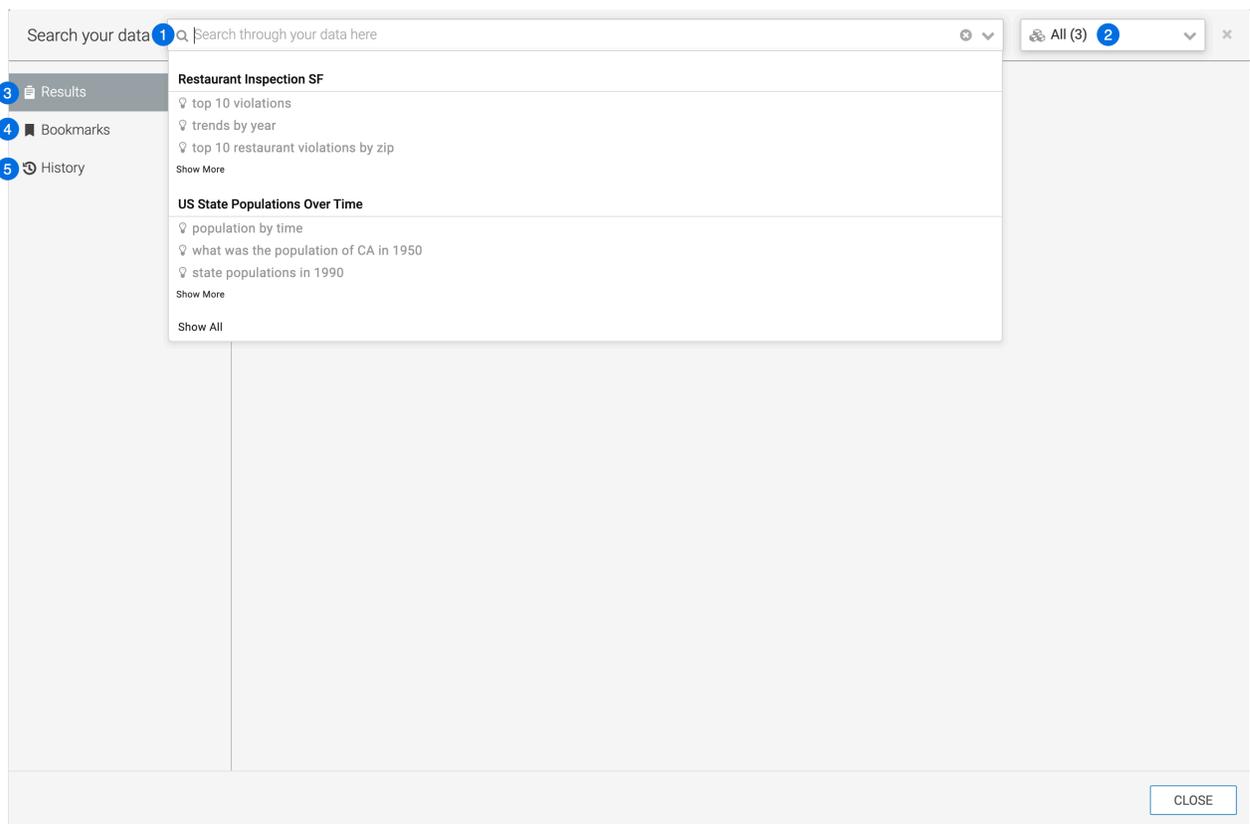
To enable Natural Language Search (NLS) modal globally, see [Enabling NLS in Site Settings](#). For the required configurations, see [Enabling dataset for NLS](#).



Note: To use NLS after enabling search in the Menu Bar and configuring datasets for searching, click **SEARCH** in the top right corner or the icon next to searchable datasets.



When you click **SEARCH**, the search modal window opens. It is shown in the figure below:



1. Search text box - You can enter search terms here. For more information on search suggestions in the drop-down list, see [Search types](#).



Note: The drop-down list initially shows a history of queries, organized by dataset. As you type, the search window changes to deliver intermediate results and shows many additional features of the search.

2. Dataset selector - By default the value is 'All' unless you access the modal from a search-enabled dataset page. You can narrow your search to a specific dataset, or a list of datasets - See [Changing the dataset search domain](#).
3. Results tab - See [Search results](#).
4. Bookmarks tab - See [Search bookmarks](#).

5. History tab - See Search history.

Related Information

[Enabling NLS in Site Settings](#)

[Enabling dataset for NLS](#)

[Search types](#)

[Changing dataset search domain](#)

[Search results](#)

[Search bookmarks](#)

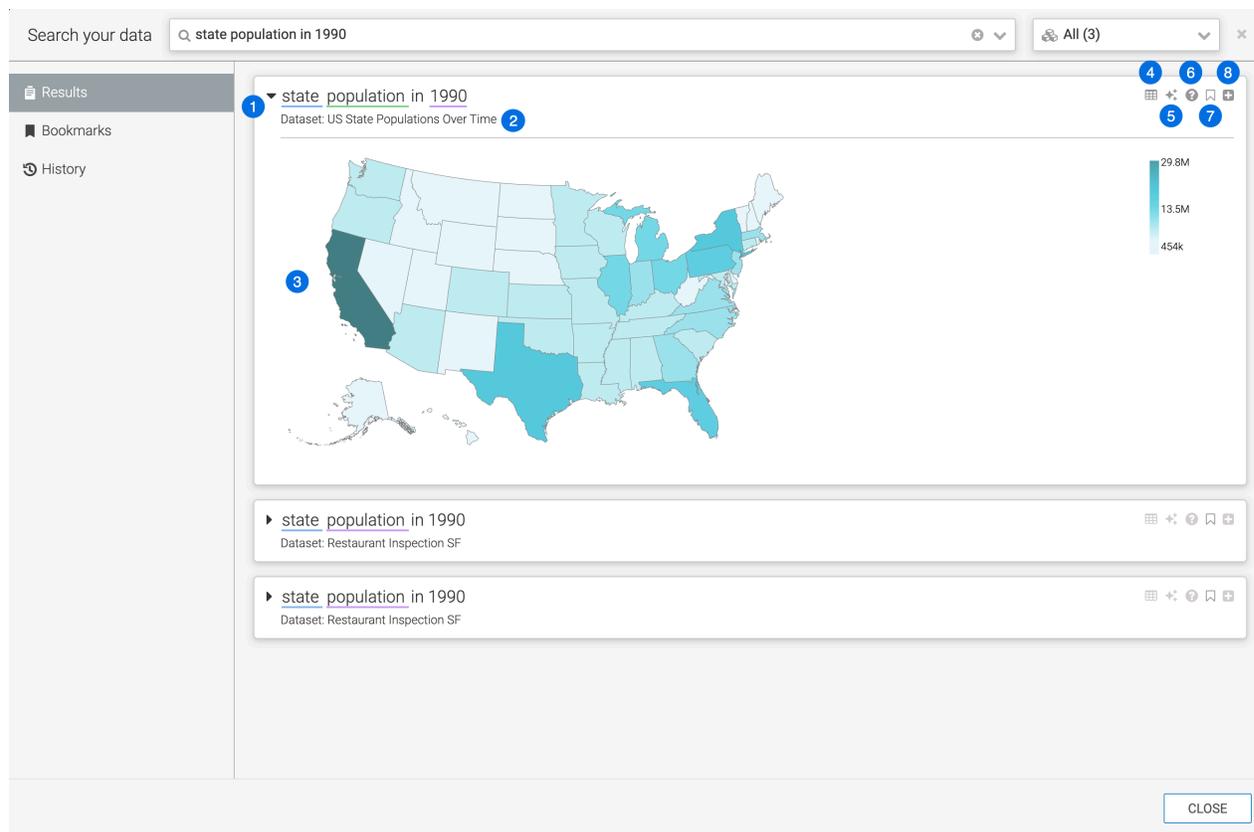
[Search history](#)

Search results

Unless the search phrase explicitly specifies the type of visual representation, Cloudera Data Visualization displays Natural Language Search results using a type that best matches the number and type of fields included in the search, for every searched dataset.

Cloudera Data Visualization displays the results of a Natural Language search directly in the search modal window.

The main features of the search results are shown in the figure below:



1. Search criteria

This is identical to your original search. The various components of the search statement show colored highlights. If you move your mouse over these parts of the statement, the tooltip displays the details of the component.

2. Result dataset

The dataset where a particular search result is found. Note that each search result may represent a different dataset.

3. Result visual

The search result itself is a visual. Note that Cloudera Data Visualization displays results in a default visual type that best represents the number and type of fields in the search phrase.

4. View and download data

This will show the data that is being graphed with an option to download as a CSV file.

5. Explore

Clicking this icon enables you to choose an alternate visual. For more information, see [Changing result visual type through explore](#).

6. Explain

Hovering over this icon shows the dataset and connection informations, the SQL query and shelf information for the visual.

7. Bookmark

Clicking this icon adds the visual to the bookmarked results. Clicking it again to removes it.

8. Dashboard actions

Clicking this icon adds the visual to a dashboard.

Related Information

[Changing result visual type through explore](#)

Order of search results

Cloudera Data Visualization automatically ranks the results of Natural Language Search.

When search produces multiple results from different datasets, you can see the most relevant visual at the top of the list. The ranking is based on the highest number and accuracy of matched search terms.

Search bookmarks

Cloudera Data Visualization enables you to bookmark the results of Natural Language Search, and to easily manage these bookmarks. You can revisit the results using bookmarks without having to locate or search each time.

The Bookmarks tab of the Search modal window has the features displayed in the figure below:

1. ADD TO DASHBOARD

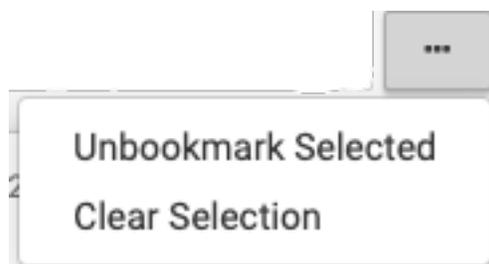
Selecting this option adds selected visuals to an existing dashboard. Note that this is a bulk operation.

2. CREATE DASHBOARD

Selecting this option creates a new dashboard that contains selected visuals. Note that this is a bulk operation.

3. Supplemental menu

Supplemental options enable you to make bulk changes to selected results.



4. Bookmarked results

The main area of the Bookmarks tab shows a list of bookmarked results. By default, they are in collapsed mode.

5. View and download data

This shows the data that is being graphed with an option to download as a .csv file.

6. Explore

Clicking this icon enables you to choose an alternate visual. See Changing result visual type through explore.

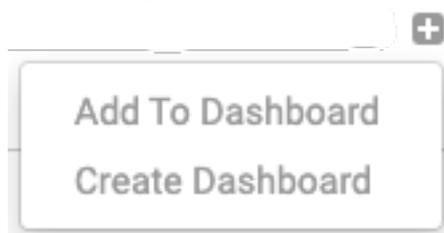
7. Explain

Hovering over this icon shows the dataset and connection information, the SQL query and shelf information for the visual.

8. Unbookmark

This enables you to remove visuals from the bookmarks.

9. Dashboard options



10. Selector

You can select or deselect visuals to make bulk actions on them.

Related Information

[Changing result visual type through explore](#)

Search history

You can browse the history of Natural Language Search in Cloudera Data Visualization, and select the results listed there.

The History tab of the Search page consists of a list of searches made previously, and the date of the search. Cloudera Data Visualization groups this list into three categories:

- Last 7 Days
- Last 30 Days
- Older Searches

Search types

For faster search results, Cloudera Data Visualization enables you to choose from different Natural Language Search types.

In illustrating these topics, the search-enabled, sample datasets Cereals and World Life Expectancy are used.

Hit highlighting

To easily differentiate the visuals in the Natural Language Search results, Cloudera Data Visualization highlights the visuals that match the text in the search text box.

When the user enters text in the search text box, the dropdown list displays all the strings that match the characters entered in the search box. To easily identify and select one of the options, the typed characters are highlighted in a different color.

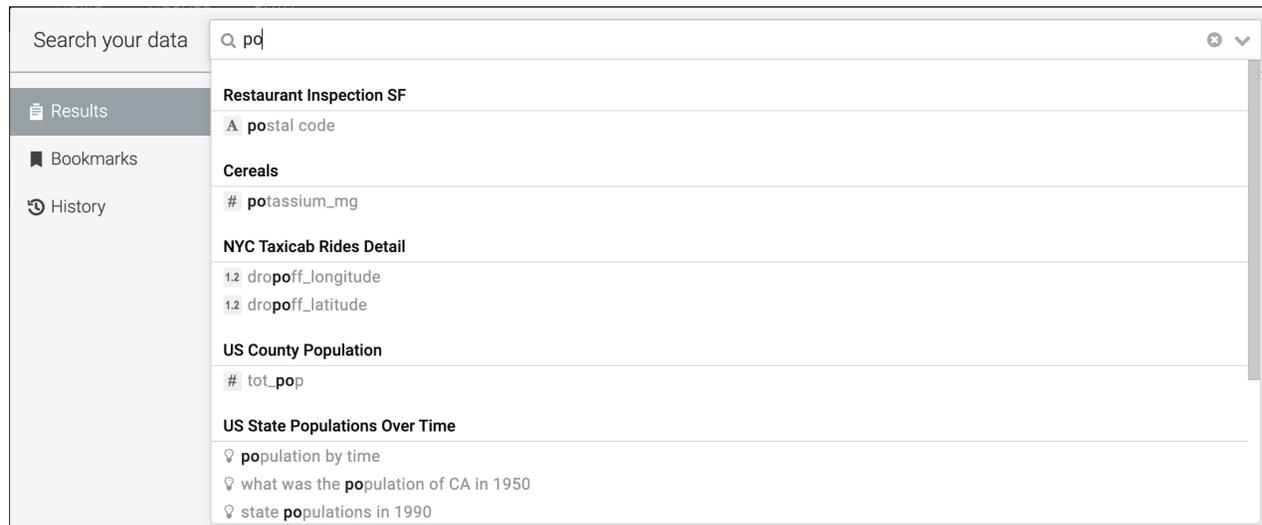
Match field names

For faster search results, Cloudera Data Visualization enables you to search visuals based on the field names entered in the Natural Language Search text box.

When you enter a field name in the search text box, the dropdown list highlights all the field names that match the characters entered in the search box. For example, if you type 'po' then the dropdown list highlights the fields

names, 'tot_pop', 'potassium_mg', and so on. The dropdown list also matches within the strings, so in our example, 'dropoff_longitude' field name is also highlighted.

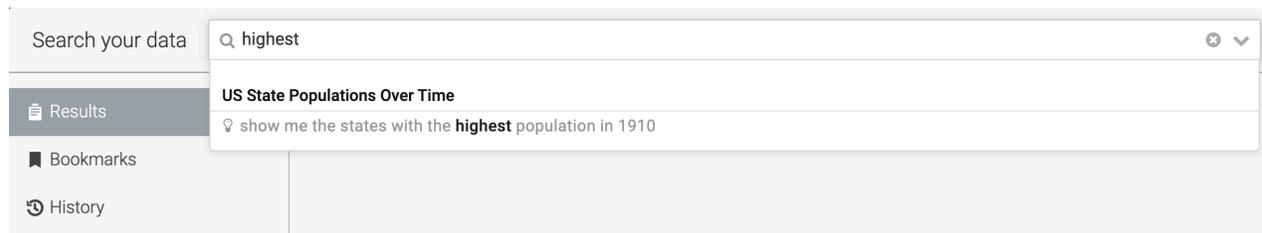
In the following example, 'dropoff_longitude' and 'dropoff_latitude' field names are also highlighted.



Matching suggested questions

For easy and faster search, Cloudera Data Visualization suggests predefined questions in the Natural Language Search drop-down list.

When you start to enter text in the search text box, the drop-down list displays predefined suggested questions that match the characters entered in the search box. In the below example, when the characters 'highest' are entered, the drop-down list highlights a suggestion with the characters 'highest' in the string.



Bookmarked searches

For easier and faster results, Cloudera Data Visualization enables you to search from a list of favorites that are saved as bookmarks.

When you first click the search box and before you start typing, your most recent bookmarked searches appear in the drop-down list.

After you start typing, if the bookmarked options are still relevant for your current search, you can select one of the bookmarked options, as they are still shown in the drop-down list.

The top screenshot shows the search interface with a search box and a dropdown menu. The dropdown menu contains suggestions like "population by time", "what was the population of CA in 1950", and "state populations in 1990". Below these suggestions are sections for "World Life Expectancy" (with sub-items like "life_expectancy", "gdp_per_capita", "country_5") and "Bookmarked Queries" (with "population by country"). A "Search History" section shows "Canada".

The bottom screenshot shows the "Bookmarks" tab selected in the sidebar. The main area displays a list of bookmarked queries, each with a title, dataset name, and date. The queries are: "top 10 restaurant violations by zip" (Dataset: Restaurant Inspection SF, Jun 15, 2021), "state populations in 1990" (Dataset: US State Populations Over Time, Jun 15, 2021), "top 10 restaurant violations by zip" (Dataset: Restaurant Inspection SF, Jun 13, 2021), and "population by country" (Dataset: World Life Expectancy, May 27, 2021).

Search history

For easier and faster search results, Cloudera Data Visualization enables you to choose from the most recent Natural Language Searches.

When you first click the search box and before you start typing, your most recent searches appear in the drop-down list.

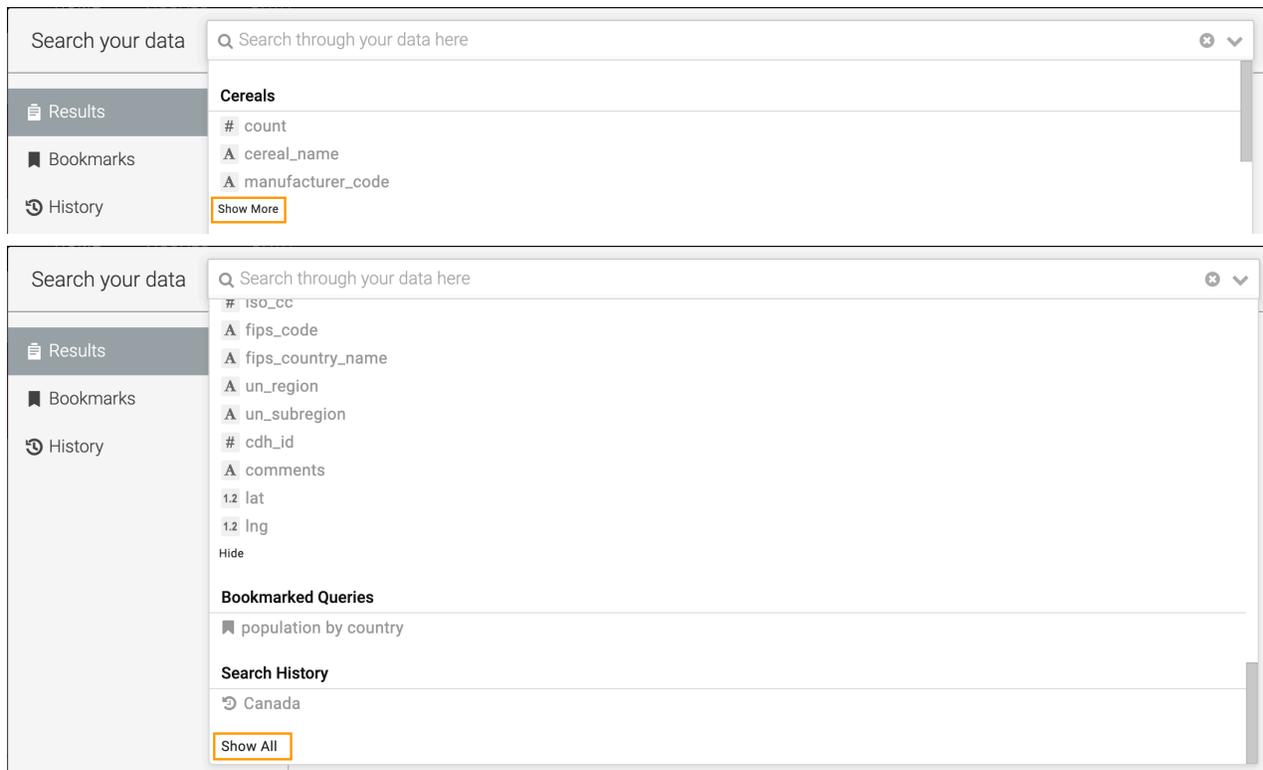
After you start typing, if your previous search is still relevant for your current search, you can select one of the previous search options, as they continue to appear in the drop-down list.

View more searches

To declutter the Natural Language Search drop-down window, Cloudera Data Visualization displays only a few matching values for any particular suggestion type.

If you click Show More in the search drop-down list under a particular dataset, all search results will be displayed for that dataset.

If you click Show All at the very end of the drop down list, all available searches will be displayed.

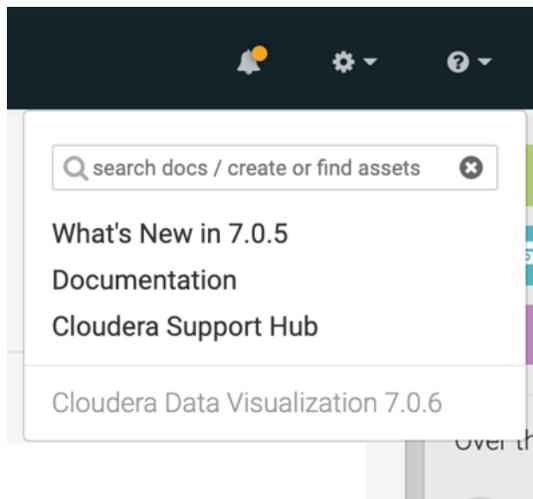


Global search

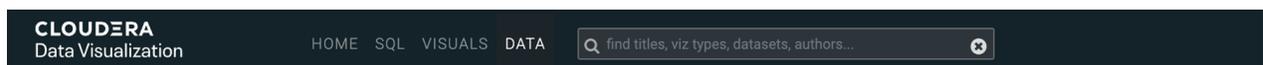
Global search can be used to quickly access certain features without clicking through the corresponding menus.

Global search can be used to quickly create or find assets, and perform certain actions, or to easily search through the documentation.

On the main of page Data Visualization (HOME interface) you can access the search text box from the Help menu.



If you switch to another interface, for example to DATA, you can find the search box on the main navigation bar.



Global search supported keywords

The keywords currently supported by global search in Cloudera Data Visualization.

Create

Create app or Create application

Creates a new, empty App and opens the App Designer page.

Create dashboard

Creates a new, empty Dashboard and opens the Dashboard Designer page.

Create data connection or Create connection

Opens the Data page and the Create New Data Connection modal.

Create dataset

Opens the New dataset modal from the last viewed data connection.

Run SQL

Opens the SQL page.

Find

Find data or Find dataset

Opens the data page and inputs remaining text as search input.

Find app or Find application

Opens the visuals page, Apps tab selected, and input remaining text as search input.

Find dashboard

Opens the visuals page, Dashboards tab selected, and inputs remaining text as search input.

None keywords

Everything not listed becomes a search term to be used for searching through the documentation. The search opens in a separate window. For and about keywords are removed from the beginning if included.