

Cloudera Data Catalog Release Notes

Date published: 2023-12-16

Date modified: 2024-12-11

The Cloudera logo is displayed in a bold, orange, sans-serif font. The word "CLOUDERA" is written in all caps, with a stylized horizontal line through the letter 'E'.

Legal Notice

© Cloudera Inc. 2025. 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

What's new.....	4
Known issues.....	4
Fixed issues.....	5
Cumulative hotfixes: Cloudera Data Catalog Private Cloud.....	5
Cloudera Data Catalog Private Cloud 1.5.4-CHF1.....	5
Fixed issues in Cloudera Data Catalog (1.5.4-CHF1).....	5
Known issues in Cloudera Data Catalog (1.5.4-CHF1).....	8
Cloudera Data Catalog Private Cloud 1.5.4-CHF3.....	8
Known issues in Cloudera Data Catalog (1.5.4-CHF3).....	8
Service Packs: Cloudera Data Catalog on Private Cloud.....	9
Cloudera Data Catalog 1.5.4 SP1.....	9
What's new in 1.5.4 SP1.....	9
Fixed issues in 1.5.4 SP1.....	9
Known issues in 1.5.4 SP1.....	10
Cloudera Data Catalog 1.5.4 SP2.....	11
What's new in 1.5.4 SP2.....	11
Fixed issues in 1.5.4 SP2.....	12

What's new

This section lists major features and updates for the Cloudera Data Catalog service.

This release of Cloudera Data Catalog on CDP Private Cloud Base 1.5.4 includes the following features and enhancements:

- Bug fixes and other performance improvements

Known issues

Learn about the known issues in Cloudera Data Catalog, the impact or changes to the functionality, and the workaround.

CDPDSS-2706: Data Steward user role is not reflecting

Data Steward role is not reflecting in Control Plane.

None

CDPDSS-3051: Profiler Jobs are not in the sorted order

Profiler jobs are incorrectly ordered when the list cannot fit into a single page under Profiler Jobs . The latest job is displayed in the last page.

CDPDSS-3047: The profiled data is not shown for Ranger Audit Profiler

The profiled data is not shown for Ranger Audit Profiler if the active NameNode is switched over to standby.

Switch the NameNode in standby to active state. For more information on restarting nodes, see [Starting, Stopping, and Restarting Role Instances](#) or [Administrative commands](#).

CDPDSS-3057: Failed profiler job because of missing columns in log entries.

When the logs to be profiled by the Ranger Audit Profiler have missing columns, the profiling job fails. The ranger audit profiler pod log contains the following message: Ranger Audit Profiler failed: Failed to run Ranger Audit Profiler. Error: "[*COLUMN1*, '*COLUMN2*', ...] not in index".

COLUMN1, *COLUMN2* show the missing column entries in the logs.

Each Ranger audit profiler job has 3 sub-jobs (identified by the same job ID prefix), any of which could fail because of the above reason. It is possible that one or even two sub-jobs pass till the above scenario is encountered.

None

CDPDSS-3353: Profiler Configuration page is giving error, if browser's preferred language is not UK English

The Profilers Configs is may be incorrectly rendered if the web browser's preferred language is not set to en-gb.

CDPDSS-2913: Asset Details view does not provide details of the asset - exception

Entities without lineage information may display the get lineage request failed Error from Atlas Not Found : Received 404..

CDPDSS-3472: Profiler jobs are showing up twice when two data lakes are present

Cloudera Data Catalog does not support more than one data lakes (or base clusters). Connecting two data lakes will result in all profiler jobs showing as duplicated in the **Profilers** page.

Fixed issues

This section lists the issues that have been fixed since the previous version.

CDPDSS-2707: Custom Sensitivity Profiler (CSP) regex logic fails intermittently for complex regex

The issue was fixed by changing the regex validation flow. The previously used complex regex had a different type of syntax, which was directly passed on to the code. The profilers failed when these complex regular expressions were passed directly. Therefore, the complex regex strings are encrypted now and they are now stored in the database. Once the profiler execution starts and fetches this complex regex, it first decodes the string and then uses the regex mechanism.

CDPDSS-2830: Cloudera Data Catalog running on JDK-17 in an Private Cloud environment causes ClassNotFoundException and further failures

Cloudera Data Catalog running on JDK-17 in Private Cloud causes a ClassNotFoundException and further failures which leads to the service being unusable. The service is restored by upgrading the Google Guice dependency manager to 6.0.0.

CDPDSS-2839: Hadoop version upgraded from 7.2.17.200-29 to 7.2.17.301-3

The Hadoop version is upgraded from 7.2.17.200-29 to 7.2.17.301-3 to remediate CVE-2023-44487 and other CVE fixes are implemented as well.

CDPDSS-2843: While upgrading to Data Services 1.5.3, an unnamed module @0x4e38d975 exception error is thrown

The issue causing exceptions from the unnamed module, which previously necessitated a restart of the Cloudera Data Catalog pod, has been addressed.

Cumulative hotfixes: Cloudera Data Catalog Private Cloud

Review the cumulative hotfixes that have been shipped for Cloudera Data Catalog 1.5.4.

Cloudera Data Catalog Private Cloud 1.5.4-CHF1

Review the features, fixes, and known issues in the Cloudera Data Catalog 1.5.4 Cumulative hotfix 1 release.

Fixed issues in Cloudera Data Catalog (1.5.4-CHF1)

This section lists issues fixed in this release of the Cloudera Data Catalog service.

CDPDSS-3037: Asset Details page does not highlight link for Datasets

The **Asset Details** now correctly highlights links to Datasets.

Datasets / Details / Asset Details

item

Properties

Type: HIVE TABLE # of Columns: ?? Data Lake: [Redacted] Datasets: 8 Owner: hive Created On: Mon Jun 17 2024 10:01:51 GMT+0200 (Cent... Last Access Time: Mon Jun 17 2024 10:01:51 GMT+020... Table Type: EXTERNAL_TABLE Database: test_dss_db DB Catalog: cm Parent: test_dss_db	Qualified Name test_dss_db.item@cm Comment + Add Comment Description + Add Description
---	--

CDPDSS-2958: Schema rows are not ordered as in their database

Rows under the **Schema** tab in **Asset Details** are ordered based on their position in their parent database.

Datasets / Details / Asset Details

Overview Schema Metadata Audits Policy Access Audits

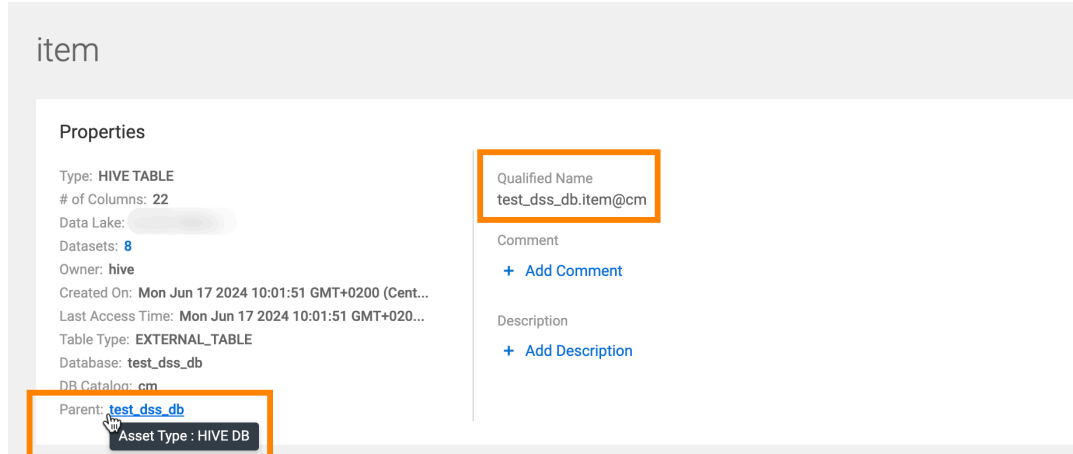
[Edit](#)

Chart Type	Name	Type	Unique Values *	Null Values	Max	Min	Mean	Comment	Classifications	Terms
▼	l_item_sk	int	17852	0	18,000	1	9,000.5			
▼	l_item_id	char(16)	8713	0						
▼	l_rec_start_date	date	4	18000						
▼	l_rec_end_date	date	3	18000						
▼	l_item_desc	varchar(200)	12819	0						
▼	l_current_price	decimal(7,2)	2917	45	99.99	0.09	9.52			
▼	l_wholesale_cost	decimal(7,2)	2137	46	87.36	0.02	5.75			
▼	l_brand_id	int	952	44	10,016,017	1,001,001	5,502,225.27			
▼	l_brand	char(50)	734	0						
▼	l_class_id	int	15	35	16	1	5.45			
▼	l_class	char(50)	93	0						
▼	l_category_id	int	10	42	10	1	5.49			
▼	l_category	char(50)	11	0						
▼	l_manufact_id	int	1000	38	1,000	1	334.57			
▼	l_manufact	char(50)	971	0						
▼	l_size	char(20)	8	0						

CDPDSS-2959: Asset Details does not provide a link to the parent database

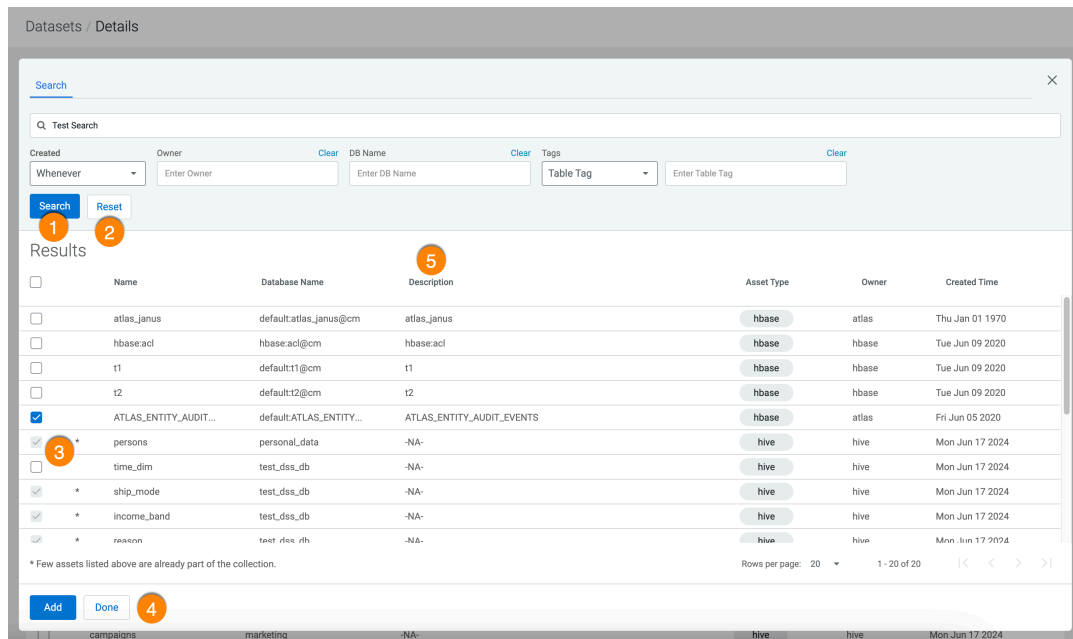
- Now, the parent database can be accessed via a link in **Asset Details**.
- The Qualified Name is no longer trimmed. Longer names are displayed in multiple rows.

Datasets / Details / Asset Details



CDPDSS-2923: Improvements for the Dataset Detail page when adding assets to a Dataset

1. The Search button is already active after the search terms are entered but before the Enter key is hit.
2. Now, clicking Reset clears all the fields.
3. Already added assets are clearly marked with a check mark as added and non-editable.
4. When adding new assets to the Dataset, the Add and Done buttons are no longer hidden while the page is waiting for the API response.
5. The Description column is added to include the Datasets Details menu. The value of the Description field is the same as in Asset Details.



CDPDSS-3042: Edit classification is failing with guid undefined error

Trying to edit the classifications in the **Asset Details** screen no longer results in the "List properties for entity request failed Error from Atlas. Not Found : Received 404. Body {"errorCode": "ATLAS-404-00-005","errorMessage": "Given instance guid undefined is invalid/not found"}" message.

CDPDSS-3049: Expanded graph is displayed on every page of the Schema tab

Previously, in the **Asset Details** page under the **Schema** tab, the bar graph of profiled assets might stayed visible for a selected asset even when moving to the next page in the list. Now, changing the page correctly closes the previously selected asset, and only displays the assets in the active page.

CDPDSS-3050: Username is not visible in Cloudera Data Catalog for the Local Administrator user

The username for local administrator users is no longer hidden. It is shown in the bottom left corner of the navigation sidebar.

CDPDSS-3048: Filtering based on Job status Finished is not working

The profiler job filter Finished is working correctly now, displaying the relevant jobs.

CDPDSS-3047: Not able to see the profiled data of Ranger Audit Profiler if the Active NameNode is switched over to standby

The profiled data is shown correctly now after the Ranger Audit Profiler is run.

CDPDSS-3051: Profiler Jobs are not in the sorted order

Profiler jobs are correctly ordered from newest to oldest when the list cannot fit into a single page under Profiler Jobs .

CDPDSS-3077:

Datasets containing deleted assets no longer fail to render the assets list with the following error: pages.dataset.labels.err-undefined-msg(Error Code - undefined)

Known issues in Cloudera Data Catalog (1.5.4-CHF1)

This section lists issues fixed in this release of the Cloudera Data Catalog service.

CDPDSS-3353: Profiler Configuration page is giving error, if browser's preferred language is not UK English

The Profilers Configs is may be incorrectly rendered if the web browser's preferred language is not set to en-gb.

CDPDSS-2913: Asset Details view does not provide details of the asset - exception

Entities without lineage information may display the get lineage request failed Error from Atlas Not Found : Received 404..

Cloudera Data Catalog Private Cloud 1.5.4-CHF3

Review the features, fixes, and known issues in the Cloudera Data Catalog 1.5.4 Cumulative hotfix 3 release.

Known issues in Cloudera Data Catalog (1.5.4-CHF3)

This section lists issues fixed in this release of the Cloudera Data Catalog service.

CDPDSS-3346: Profiler pods are stuck in init state

The CDP_ACCESS_KEY_ID and CDP_PRIVATE_KEY keys do not receive values when upgrading from Cloudera Data Catalog 1.5.2 to 1.5.4.

1. Use the following command to restart the DSS App: `kubectl rollout restart deployment/cdp-release-dssapp -n <namespace>`
2. Delete and relaunch the profilers after the upgrade.

CDPDSS-3353: Profiler Configuration page is giving error, if browser's preferred language is not UK English

The Profilers Configs is may be incorrectly rendered if the web browser's preferred language is not set to en-gb.

CDPDSS-2913: Asset Details view does not provide details of the asset - exception

Entities without lineage information may display the get lineage request failed Error from Atlas Not Found : Received 404.

Service Packs: Cloudera Data Catalog on Private Cloud

Review the Service pack features and issues for Cloudera Data Catalog 1.5.4.

Cloudera Data Catalog 1.5.4 SP1

Review the features, fixes, and known issues in the Cloudera Data Catalog 1.5.4 Service Pack 1 release.

What's new in 1.5.4 SP1

This section lists major features and updates for the Cloudera Data Catalog service.

This release of Cloudera Data Catalog on CDP Private Cloud Base 1.5.4 SP1 includes the following features and enhancements:

- Bug fixes
- Datasets can only be deleted by their creators following this release.
- You can hover over the **Comment** field for individual schema entries in **Asset Details** to preview longer comments without opening them.
- The security of Cloudera Data Catalog is improved. Now our services run with the following security context:

```
'containers[].securityContext.allowPrivilegeEscalation' to 'false'.  
'containers[].securityContext.runAsUser' to an integer > 1000000.  
'containers[].securityContext.readOnlyRootFilesystem' to 'true'
```

Fixed issues in 1.5.4 SP1

This section lists the issues that have been fixed since the previous version.

CDPDSS-3395: Incorrect results are being displayed after navigation, when "owners" filter is selected

In the **Search** page, returning to the list of results after opening the **Asset Details** of one the results and repeating the search with the same filter, other entities belonging to owners different than the selected no longer appear. Only entities belonging to the selected owners are displayed.

CDPDSS-3081: Datasets: Tenant Schema is getting passed as workload username while impersonating a user on ATLAS and Ranger API calls

Before the update, while making API calls to Apache Atlas and Apache Ranger, there was a potential bug in code wherein a wrong value for workloadUserName was getting passed which caused APIs to perform an auth check with a wrong parameter, causing unwanted issues.

CDPDSS-3169: Asset Details Page - Optimized API to avoid 504 Error

An **Asset Details** page with only a few assets (Hive table) was unable to open the Hive table assets which had around 2000 columns or more. The number of columns was causing the timeout. Optimizing the API calls related to **Asset Details** led to improved performance and avoiding potential 504 errors when opening assets.

CDPDSS-3080: Deleted Entities in Atlas cannot be removed from Datasets

Before the update, entities deleted in Apache Atlas could not be removed from Cloudera Data Catalog datasets.

CDPDSS-3366: Asset Details is distorted with very long description

After the update, the very long strings in the **Description** field are trimmed and no longer distort **Asset Details** by being displayed in full length.

CDPDSS-3351: Unable to see comments on the Hive columns in Asset Details

Full comments can be displayed in **Asset Details** while hovering over the **Comment** field compared to the earlier limited view.

CDPDSS-2688: Hive Column On-demand Profiler is failing but the Asset Details page is not updated with 'Failed State'

A failed Hive Column On-Demand Profiler job is correctly displayed as failed in **Asset Details Profilers** .

CDPDSS-3353: Profiler Configuration page is giving error, if browser's preferred language is not UK English

The **Profilers Configs** is correctly rendered even if the web browser's preferred language is not set to en-gb.

CDPDSS-2913: Asset Details view does not provide details of the asset - exception

Entities without lineage information no longer display the get lineage request failed Error from Atlas Not Found : Received 404..

CDPDSS-3381: The Ranger Audit Profiler is showing status as FINISHED/SUCCESS even though there is error with the profilers

After the update, the Ranger Audit Profiler shows the correct status when the NameNode is stopped, and it is no longer stuck in FINISHED.

Known issues in 1.5.4 SP1

Learn about the known issues in Cloudera Data Catalog, the impact or changes to the functionality, and the workaround.

CDPDSS-3057: Failed profiler job because of missing columns in log entries.

When the logs to be profiled by the Ranger Audit Profiler have missing columns, the profiling job fails. The ranger audit profiler pod log contains the following message: Ranger Audit Profiler failed: Failed to run Ranger Audit Profiler. Error: "[*COLUMN1*, '*COLUMN2*', ...] not in index".

COLUMN1, *COLUMN2* show the missing column entries in the logs.

Each Ranger audit profiler job has 3 sub-jobs (identified by the same job ID prefix), any of which could fail because of the above reason. It is possible that one or even two sub-jobs pass till the above scenario is encountered.

None

CDPDSS-3472: Profiler jobs are showing up twice when two data lakes are present

Cloudera Data Catalog does not support more than one data lakes (or base clusters). Connecting two data lakes will result in all profiler jobs showing as duplicated in the **Profilers** page.

None

CDPDSS-3665: download-search-result API failing in a Cloudera Data Catalog on premises deployment

In the Search page, clicking the Download CSV button can result in a HTTP 500: Read-only file system.

1. Edit the deployment configuration for cdp-release-dssapp in the specified namespace:

```
kubectl edit deployment cdp-release-dssapp -n ***NAMESPACE***
```

2. Set allowPrivilegeEscalation to false within securityContext for the container:

```
securityContext:
  allowPrivilegeEscalation: false
```

3. Save and exit your editor.

4. Verify the changes have been applied by describing the deployment:

```
kubectl get deployment cdp-release-dssapp -n ***NAMESPACE*** -o yaml | grep -A 5 securityContext
```

5. Restart the pods to apply the changes:

```
kubectl rollout restart deployment cdp-release-dssapp -n ***NAMESPACE***
```

Cloudera Data Catalog 1.5.4 SP2

Review the features, fixes, and known issues in the Cloudera Data Catalog 1.5.4 Service Pack 2 release.

What's new in 1.5.4 SP2

This section lists major features and updates for the Cloudera Data Catalog service.

This release of Cloudera Data Catalog on CDP Private Cloud Base 1.5.4 SP2 includes the following enhancement:

- Bug fixes

- For Hive tables, the **Schema** tab is updated in **Asset Details** which can be reached from **Search**:
 - Multiple rows of the Hive Column Profiler results can be expanded at the same time.
 - Rows can display the applicable bar and pie chart.

Asset Details

us_customers Atlas

Properties

Type: HIVE TABLE
 # of Columns: 39
 Data Lake: e6127y-env-datalake
 Datasets: 0
 Owner: hive
 Created On: Mon Feb 24 2025 05:01:55 GMT+0100 (Cent...
 Last Access Time: Mon Feb 24 2025 05:01:55 GMT+010...
 Table Type: EXTERNAL_TABLE
 Database: hortoniabank
 DB Catalog: cm
 Parent: hortoniabank

Qualified Name
 hortoniabank.us_customers@cm

Comment
[+ Add Comment](#)

Description
[+ Add Description](#)

Profilers | 2

Cluster Sensitivity Profiler
 Last run: 2/24/2025, 12:14:02 PM (UTC) | Status: SUCCESS [Run](#)
 Next Schedule Run: 2/25/2025, 12:00:00 AM (UTC)

Hive Column Profiler
 Last run: 2/24/2025, 12:13:51 PM (UTC) | Status: SUCCESS [Run](#)
 Next Schedule Run: 2/25/2025, 12:00:00 AM (UTC)

Classifications Managed System Propagated

[+ Add Classification](#)

Terms

[+ Add Terms](#)

Overview Schema Metadata Audits Policy Access Audits

Search column Edit

Chart Type	Name	Type	Unique Values *	Null Values	Max	Min	Mean	Comment	Classifications	Terms
📊	gender	string	2	0						
📊	givenname	string	438	0						
📊	city	string	491	0						

📊	country	string	1	0						
---	---------	--------	---	---	--	--	--	--	--	--

📊	countryfull	string	1	0						
📊	emailaddress	string	781	0						dp_e...
📊	birthday	string	772	0						
📊	age	int	67	0	85	19	51.23			
📊	cctype	string	2	0						
📊	ccnumber	string	781	0						dp_credR...
📊	cur?	string	536	0						

Fixed issues in 1.5.4 SP2

This section lists the issues that have been fixed in 1.5.4 SP2.

CDPDSS-3665: download-search-result API failing in a Cloudera Data Catalog on premises deployment

In the Search page, the Download CSV button correctly downloads the search results after the update.

CDPDSS-3567: 504 Error on clicking Atlas Tags when data lake access is missing

After the update, the **Atlas Tag** menu can be successfully opened even without selecting a data lake.