



## WHAT'S NEW AND RELEASE NOTES

Release: 2020.5

## IN THIS RESLEASE

**NEW FEATURES AND IMPROVEMENTS..... 2**

<b>DATA PREPARATION.....</b>	<b>2</b>
Certify datasets.....	2
Datasets export and import.....	4
Datasets tags .....	6
Data cleaning.....	7
Split a column into multiple columns .....	10
Bulk rename columns in a data source.....	11
Resolve duplicate column names .....	12
String function for splitting columns .....	13
Scalar script function for data science.....	14
Support BLOB data type in datasets.....	15
Load more datasets into view .....	16
AI connections export and import.....	17
Built-in DataClarity Python server .....	18
Encrypted data connections credentials .....	18
<b>STORYBOARDS.....</b>	<b>19</b>
Certify storyboards .....	19
Enhanced sorting in widgets.....	21
Background image for the Table widget .....	22
Visualize image columns in the Table widget.....	23
Margins for geospatial widgets .....	24
Customize lines and direction arrows in the Path map .....	24
<b>GENERAL.....</b>	<b>25</b>
Change content ownership when deleting a user .....	25
Audit user activities .....	26
Configure notifications lifespan.....	27
Message broker configuration .....	27

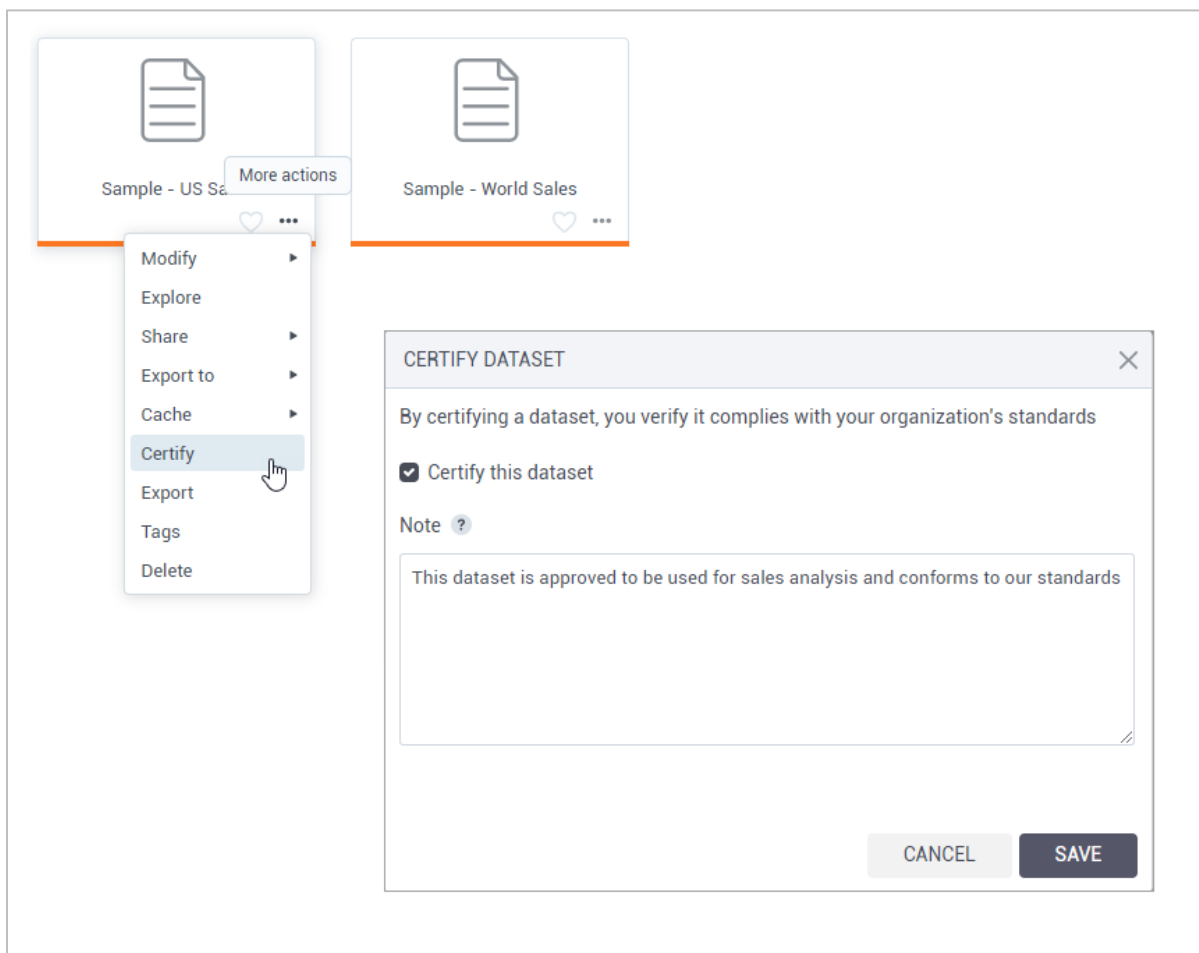
## NEW FEATURES AND IMPROVEMENTS

### DATA PREPARATION

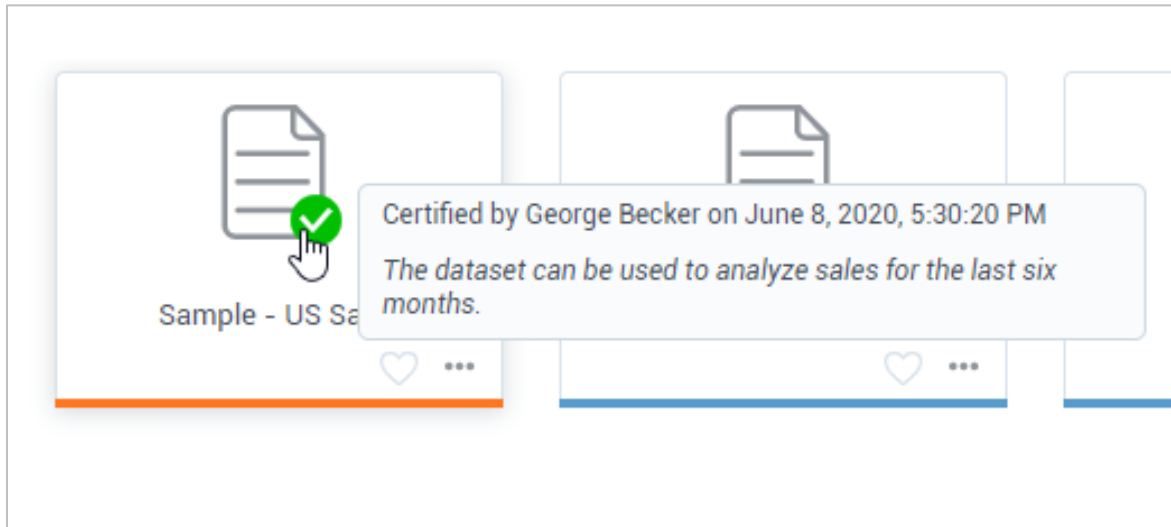
#### *Certify datasets*

To help data analysts find datasets that are trusted and recommended for their analysis, you can certify the datasets. For best search results, only certify the datasets that are valuable and considered the official source of enterprise information. Datasets certification ensures the data is reliable and can be used across the organization.

Before performing this task, confirm that users have the dedicated permission to certify datasets, granted in Access Manager. If you can certify datasets, you can find the **Certify** option under the **More actions** menu for the dataset. Users can mark datasets as certified by selecting the **Certify this dataset** check box. In addition, you can add a note about the certification status with descriptive information.



Certified datasets appear with a green check mark over the dataset icon. Hovering over the certification icon will reveal a tooltip with the certification information. The tooltip also shows who certified the dataset and the date and time of the certification.

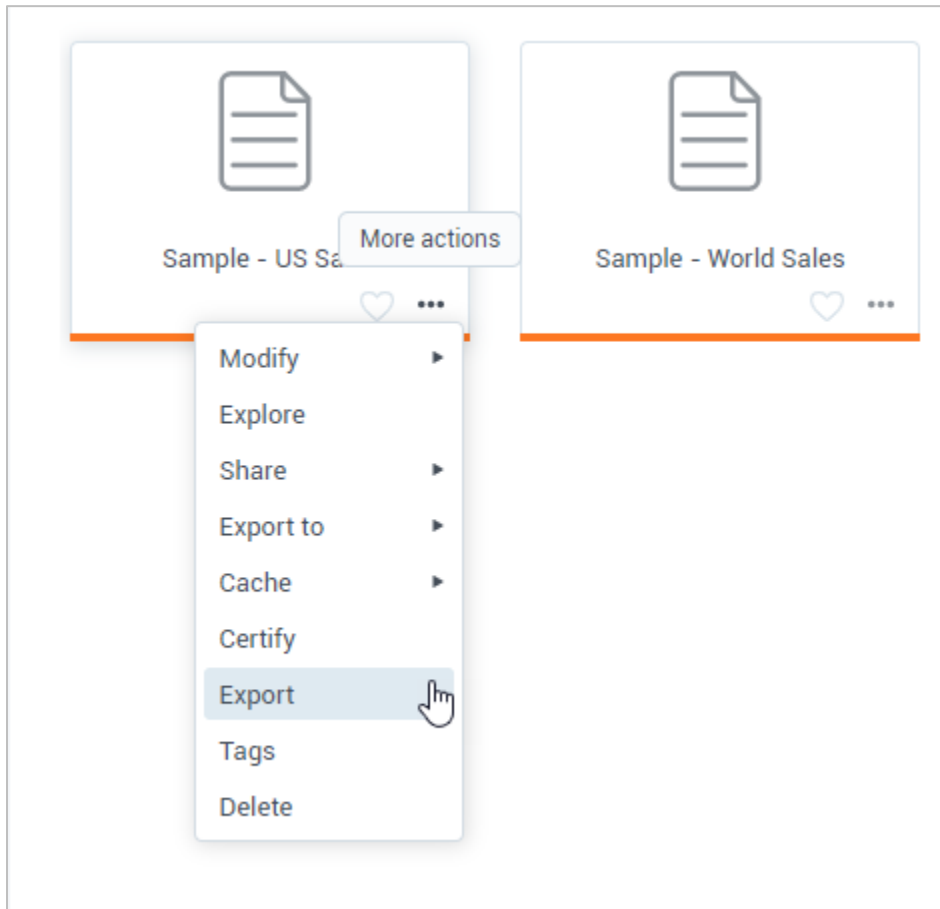


Each change to the previously certified dataset removes its certification, and as a result, will require the dataset's recertification. If multiple users are certifying a dataset at the same time, the last saved changes are applied.

### *Datasets export and import*

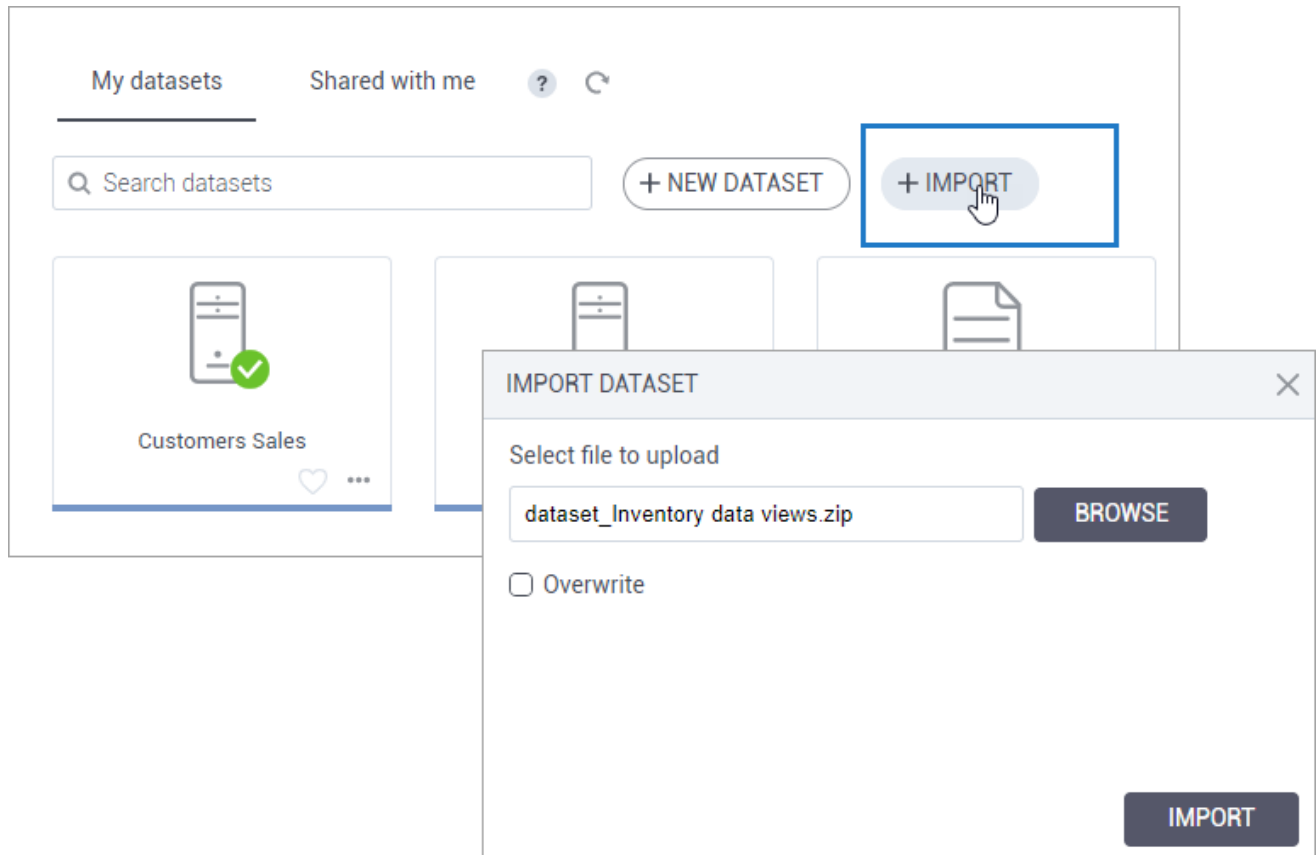
In this release, dataset management capabilities include the options for dataset export and import. This way, users can save time and reuse the datasets in different environments.

To export a dataset, in the **Datasets** pane, select the dataset, and under the **More actions** menu, click **Export**. The dataset with all underlying data sources, connections, and AI connections will be then saved into a ZIP file.



To import a dataset, you will need to upload a corresponding ZIP file. In the **Datasets** pane, clicking the **Import** button will open the dialog for the file import. You can specify whether to overwrite the dataset if the same element has already been imported previously. If the same dataset already exists, then the dataset will not be imported.

SS



### Datasets tags

Tags are keywords that you can add to your datasets to help users better categorize and find datasets. When searching for datasets, users can now use tags for their search criteria.

Datasets can include two types of tags:

- **Public tags** — Automatically visible within a tenant so that other users may reuse the same tags for their datasets. For a shared dataset, public tags are available to all recipients, in view mode.
- **Private tags** — Visible only for the user who created them.

**TAGS MSSQL DATASET** [X]

**Public tags**

mall × sales ×

**Private tags**

blended

blended (New Tag)

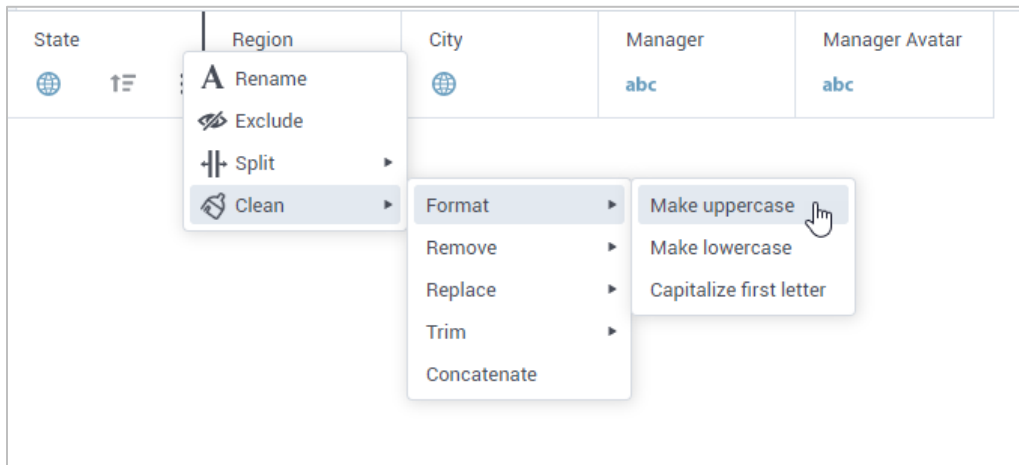
CANCEL OK

## Data cleaning

A range of data cleaning functions is now included in the column menu to help dataset modelers ensure data accuracy and consistency. The **Clean** column menu contains the following data cleaning options: **Format**, **Remove**, **Replace**, **Trim**, and **Concatenate**.

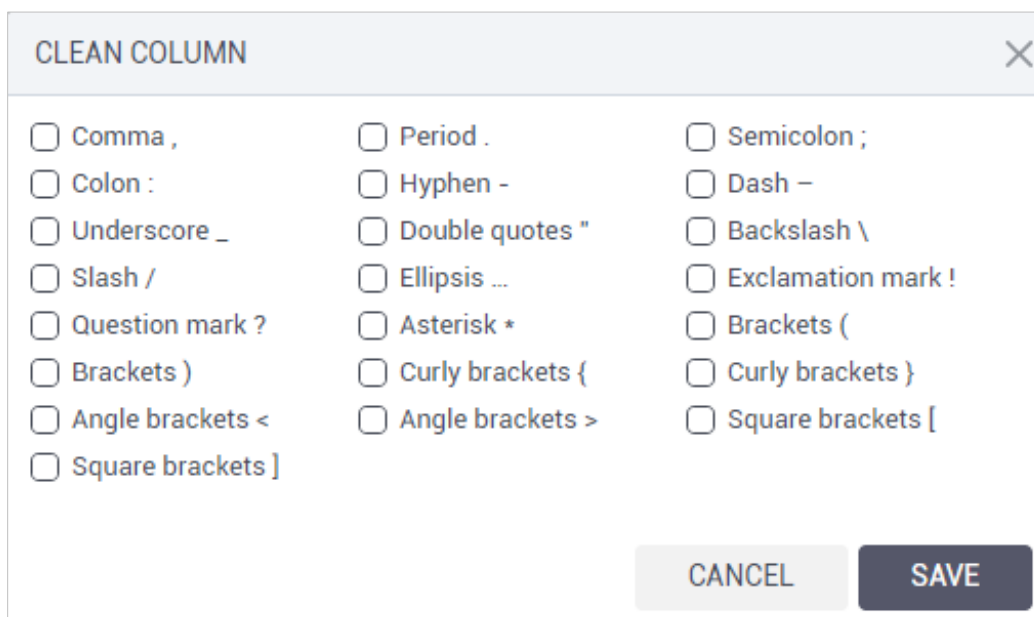
### Format values

The **Format** menu allows you to convert all the column values to uppercase, lowercase, or capitalize the first letter. Formatting can be a great time-saver to ensure consistency in values naming and capitalization.



### Remove unnecessary values

You can quickly remove NULLs, blank values, and punctuation marks of your choice by using the **Remove** option for the string column types.





For the numeric type of data, the available options allow you to remove NULLs, zero values, or negative values.

No of Transactions	No of Customers	Quantity	Unit Price
#		#	#
4		15	12
5			
7	30		
10	12	48	30

A Rename

Exclude

$\Sigma$  Aggregation

Clean

Remove

Concatenate

NULLs

Zero values

Negative values

### Replace NULLs or any specified value

In the new **Replace** column option, you can replace all the occurrences of the NULL or any specified value with another value. This option can be also used to correct spelling errors.

CLEAN COLUMN

Value

United States of America

Replace with

USA

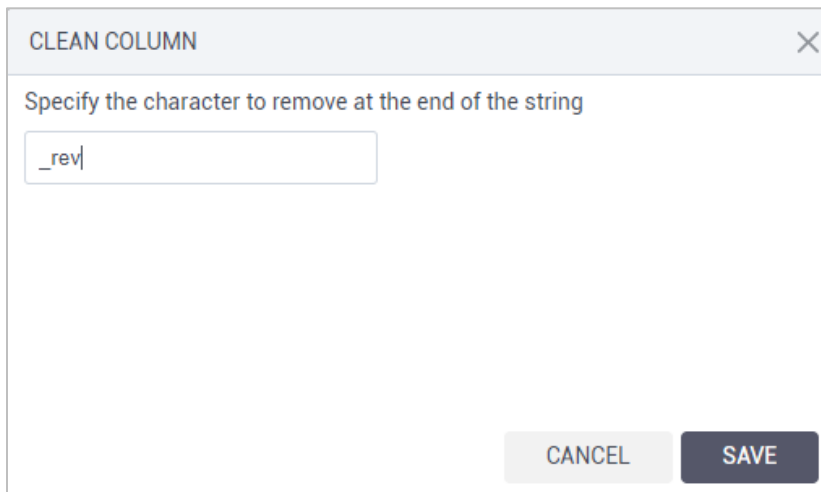
CANCEL

SAVE

## Trim spaces and characters

Get rid of unnecessary spaces or any leading/trailing characters of your choice by using the **Trim** options:

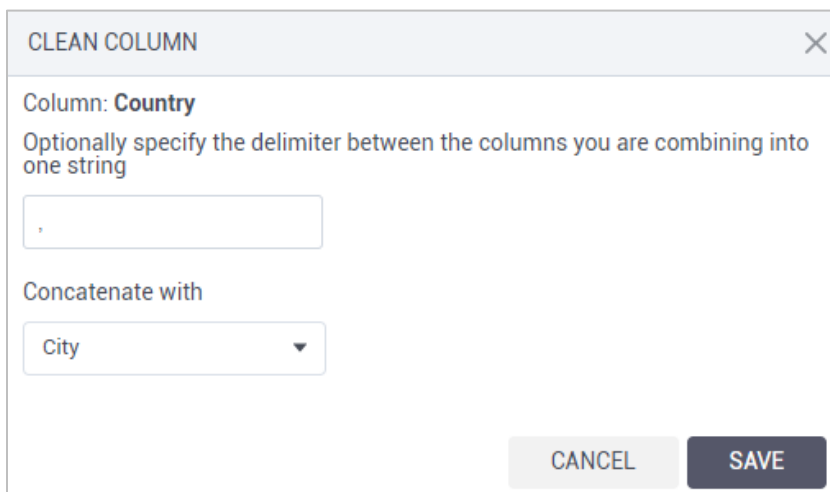
- **Leading & Trailing spaces** – To remove spaces at the beginning and the end of the string.
- **Leading spaces** – To remove spaces at the beginning of the string.
- **Trailing spaces** – To remove spaces at the end of the string.
- **All spaces** – To remove all the spaces.
- **Leading character** – To remove the specified character at the beginning of the string.
- **Trailing character** – To remove the specified character at the end of the string.



A screenshot of a dialog box titled "CLEAN COLUMN" with a close button (X) in the top right corner. The main text inside the dialog says "Specify the character to remove at the end of the string". Below this text is a text input field containing the string "\_rev". At the bottom of the dialog, there are two buttons: "CANCEL" and "SAVE".

## Concatenate columns

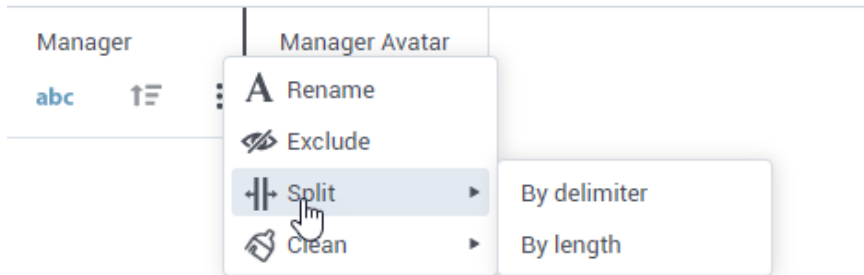
Using the new **Concatenate** column option, you can join the selected string column with another one. Moreover, you can specify a separator in between, if needed. In the separator field, one space is set as the default option.



A screenshot of a dialog box titled "CLEAN COLUMN" with a close button (X) in the top right corner. The main text inside the dialog says "Column: **Country**" and "Optionally specify the delimiter between the columns you are combining into one string". Below this text is a text input field containing a comma (','), which is the default separator. Further down, there is a label "Concatenate with" followed by a dropdown menu showing "City". At the bottom of the dialog, there are two buttons: "CANCEL" and "SAVE".

### Split a column into multiple columns

In this release, with the new **Split** column menu, you can split one column into up to 3 columns based on a selected delimiter or by characters' length.



To split a column by a delimiter, select a separator from the list or specify a custom character at which to split the column.

A screenshot of a dialog box titled 'SPLIT COLUMN BY DELIMITER'. Inside the dialog, there is a text prompt: 'Select or enter the separator to split the text column'. Below this, there are two dropdown menus. The first is labeled 'Delimiter' and has 'Comma (,)' selected. The second is labeled 'Number of columns' and has '2 columns' selected. At the bottom right of the dialog, there are two buttons: 'CANCEL' and 'SAVE'.

If you opt for splitting by length, specify the number of characters to repeatedly split the text column. For example, you can split the selected column into two columns every five characters. This option can be helpful if you have IDs of a certain length at the beginning – you can add one split that equals the ID length.

SPLIT COLUMN BY LENGTH

×

Specify the number of characters to repeatedly split the text column

Number of characters

4

Number of columns

2 columns

CANCEL

SAVE

### *Bulk rename columns in a data source*

To keep dataset column names consistent, you can benefit from the bulk columns renaming within a data source. Next to a data source name, click **More options** > **Rename** > **Rename all columns**, and specify a pattern on how to rename all the columns:

- **Add prefix** – Add a data source name or any custom text, at the beginning of the column name.
- **Add suffix** – Add a data source name or any custom text, at the end of the column name.

The example at the bottom of the dialog box will reflect all the selections based on the first column name. The example is convenient to preview the pattern results that you are applying to all the columns.

RENAME ALL COLUMNS

×

Add prefix

Source name

Add suffix

Example: Sales.Sales Date

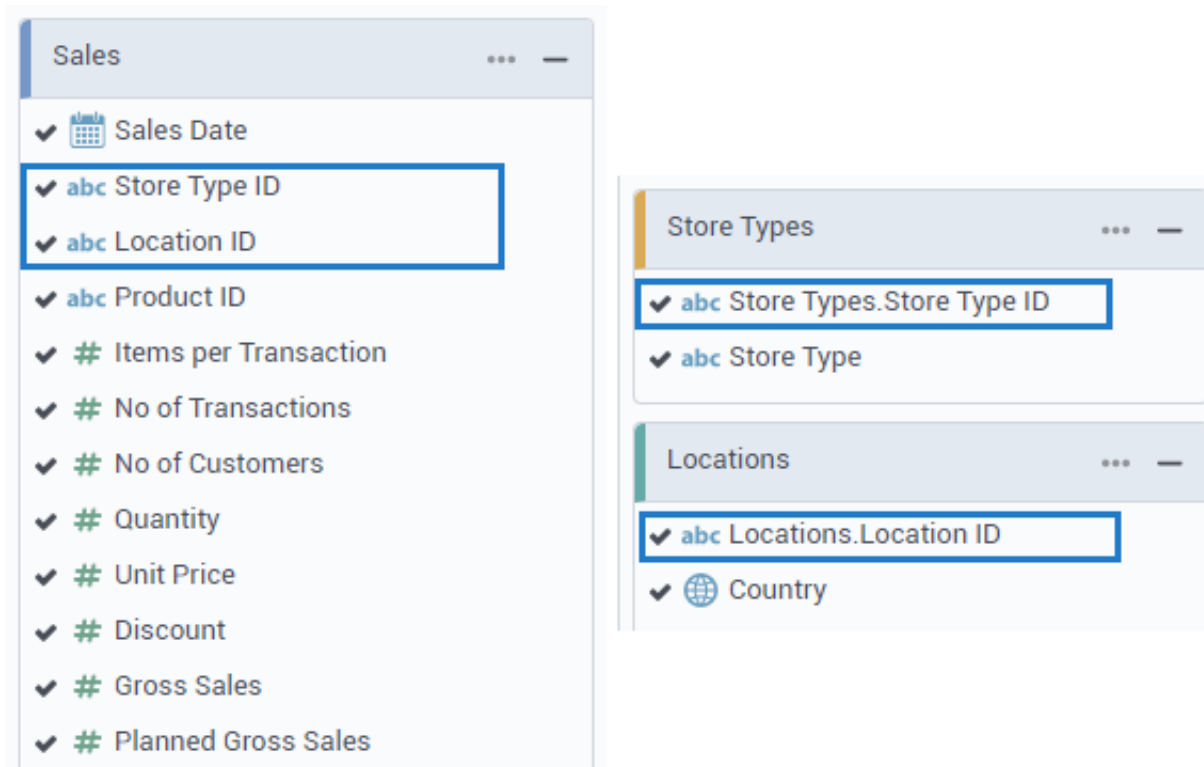
CANCEL

OK

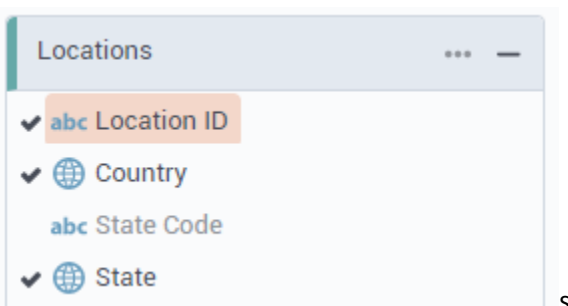
### Resolve duplicate column names

This release offers a way to resolve duplicate column names. When clicking **Finish** in the last step of the dataset creation process, you will see the list of duplicate columns, and there you have three options on how to proceed:

- **Hide duplicates** – The first occurrence of a duplicate column remains in the dataset, the rest of duplicate columns become hidden.
- **Rename duplicates** – The system automatically renames all duplicate columns by adding the source name in the beginning as follows.



- **OK** – Click to resolve all the duplicates manually. The duplicates are highlighted in red to help users find the columns if they decide to rename them one by one manually.



### String function for splitting columns

The new split function now appears in the list of string functions in the **Calculations** pane. This function allows you to obtain a specific part of the string based on a separator of your choice.

CALCULATIONS ?

Search

Dimensions

City

Country

abc Location ID

abc Manager

abc Manager Avatar

Region

State

abc State Code

Measures

× Manager Last Name

Dimension

SPLIT\_PART([Manager], ',', 2)

SAVE

Search functions

String

SPLIT\_PART

STUFF

SUBSTRING (str, start, len)

SUBSTRING (str, len)

**SPLIT\_PART(str, delimiter, part)**  
Splits a string (str) based on the specified delimiter and returns the part at the specified position.  
The part number to return starts with 1.  
**Example:**  
SPLIT\_PART('John,Smith', ',', 2) = 'Smith'

### Scalar script function for data science

Prior this release, you could use only the **SCRIPT** function that sent a set of values as a table to the server for processing and received an array of rows for a new calculated column. To cover the use case where each row needs to be executed separately, you can now use a **SCALAR** function in the **Calculations** pane.

The screenshot shows the 'CALCULATIONS' pane in DataClarity. On the left, there are two lists: 'Dimensions' (Location ID, Product ID, Sales Date, Store Type ID) and 'Measures' (Discount, Entertainment Cost, Gross Sales, Items per Transaction, Manufacturing Cost, Marketing Cost, No of Customers, No of Transactions, Planned Gross Sales, Planned Profit, Profit, Quantity, Sales Cost). The main area on the right is for configuring a calculation. It has a 'Calculation name' field, a 'Measure' dropdown, and a large text area for the script. Below the script area are 'EDIT SCRIPT', 'CLEAR', and 'SAVE' buttons. There is also a 'Search functions' field and a 'Data Science' dropdown. A list of functions is shown: SCALAR, SCRIPT, MLLIB.CLUSTER, and MLLIB.BICLUSTER. The 'SCALAR' function is selected, and its details are shown below: 'SCALAR(connector\_id, data\_type, code, columns)' and a description: 'Executes a scalar script by using the specified connector. To define a scalar script calculation: 1. Drag the function, and then click the Edit Script button that appears. 2. Select AI connector, language, and the format of the returned value. 3. Enter the code and select the columns. 4. Click Apply. The formula is automatically updated according to the data you specified. 5. Save your changes.'

After clicking the **Edit script** button, the calculation type is set to **Scalar** for such a function. You can quickly switch to the **Script** function by selecting **Vector** in the drop-down list.

SCRIPT

Connector

Language

Calculation

Result

DataClarityPy (built-in)

Python

Scalar

String

Type script

Click to add columns

APPLY

### *Support BLOB data type in datasets*

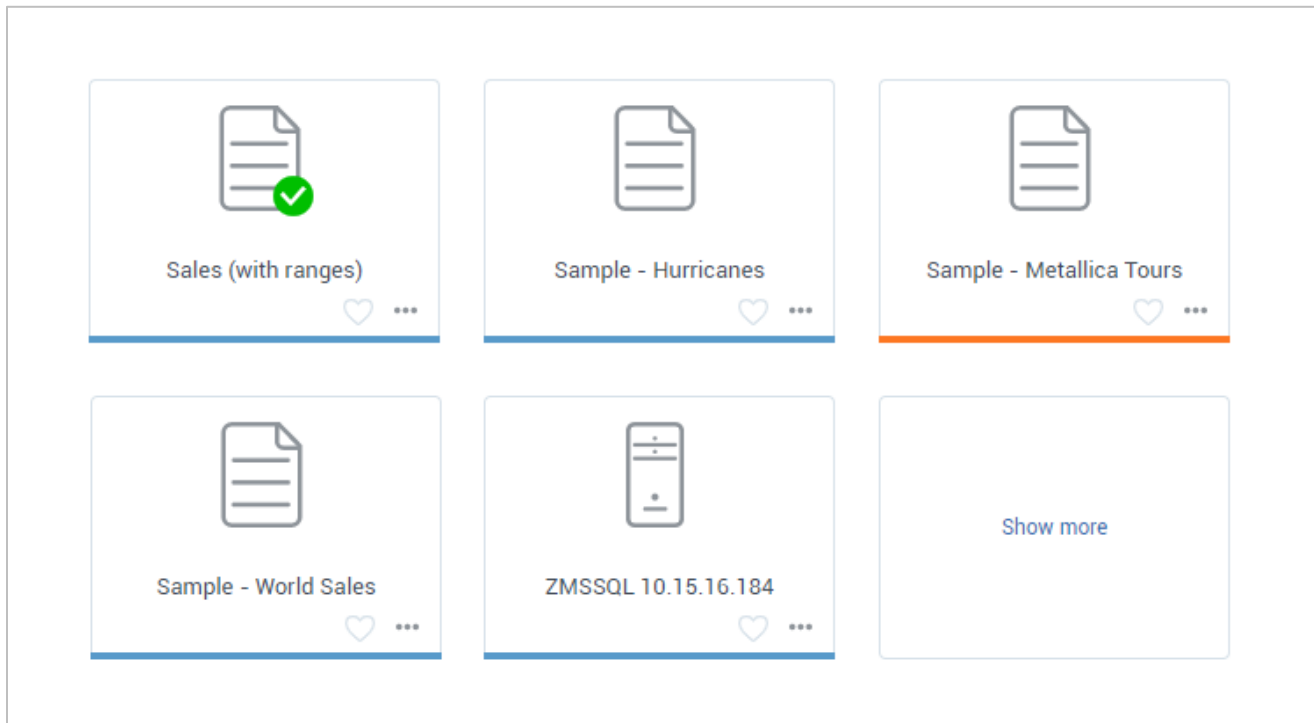
Now, you can use the BLOB type of columns with image binary data in the datasets. To visualize image type columns, you can use the **Table** widget. For details, see *Image columns in Table widget* section.



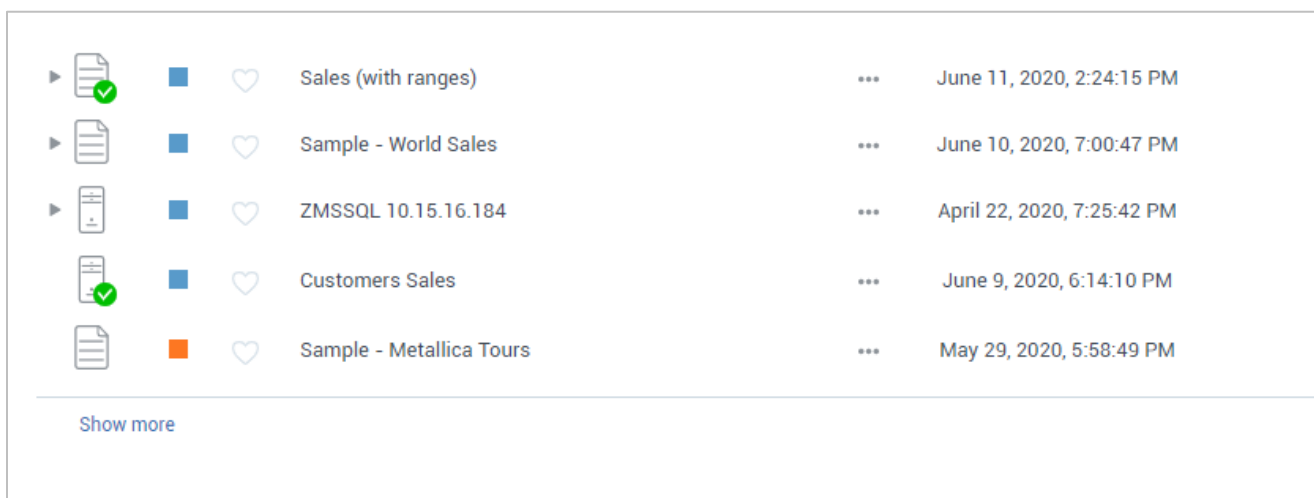
### Load more datasets into view

The new pagination technique has improved the performance of Data Preparation. When you open the **Datasets** pane, it loads the first set of 20 datasets. If the number of user datasets exceeds 20, the new **Show more** link appears at the end of the dataset list. Click this link to load another set of 20 datasets into the view, and so on.

In **Tile** view, the last tile appears with the **Show more** link on it.



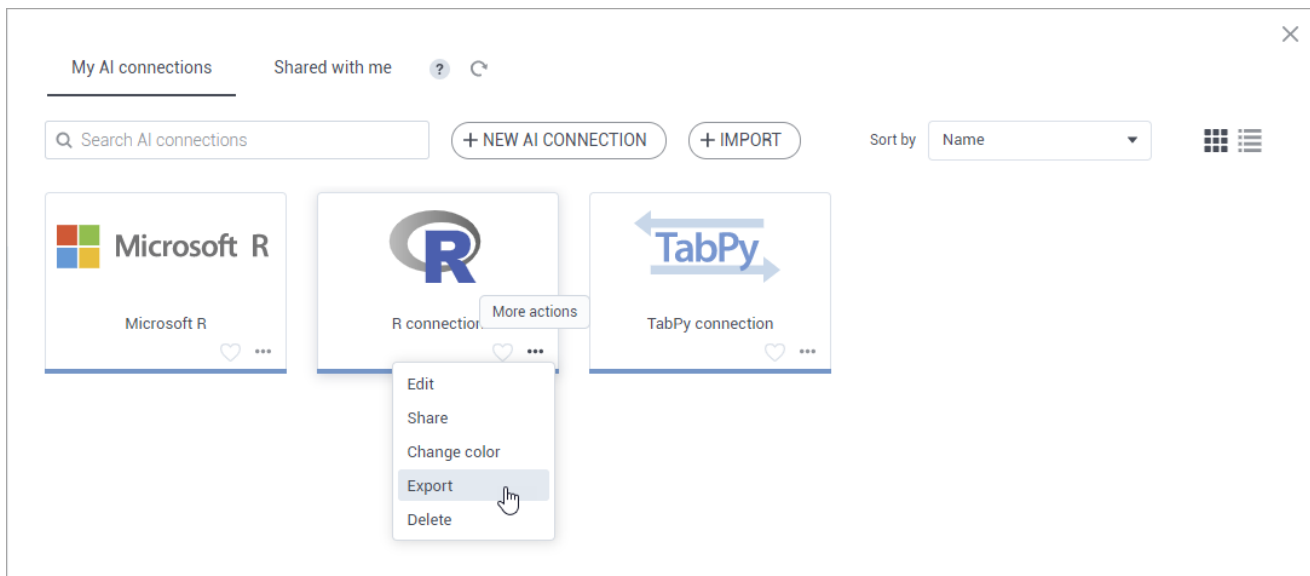
In **List** view, the **Show more** link is placed right under the last dataset in the list.



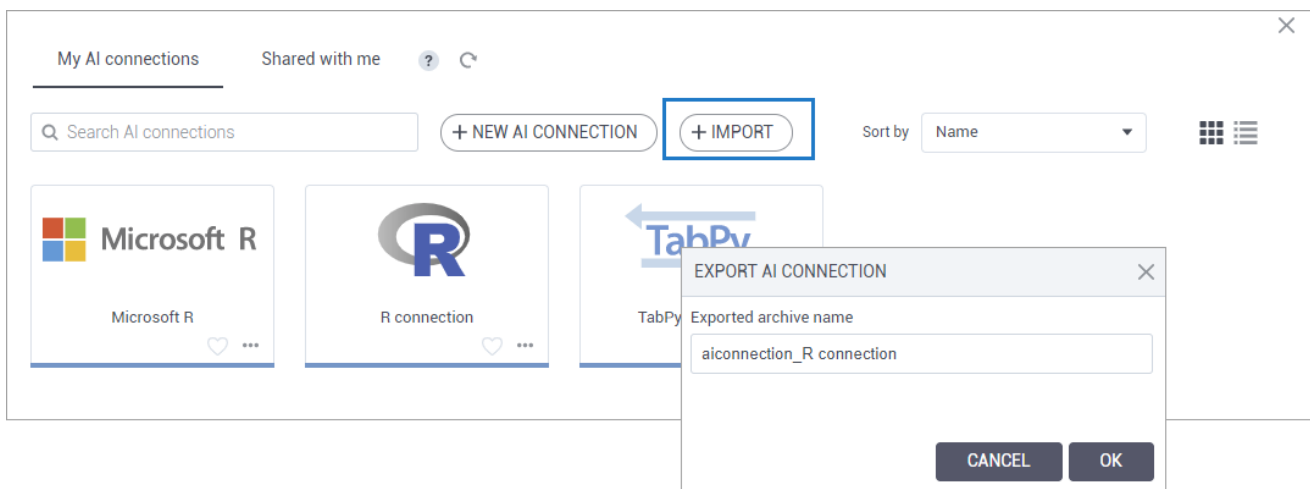
### AI connections export and import

When managing AI connections, you can now benefit from the new capabilities like connection export and import. AI connections can be easily reused in different environments saving your time spent for manually recreating new ones.

To export any of your AI connections, in the **AI connections** pane, find the needed item, and under the **More actions** menu, click the new **Export** option. The AI connection details that you are exporting will be saved into a ZIP file.

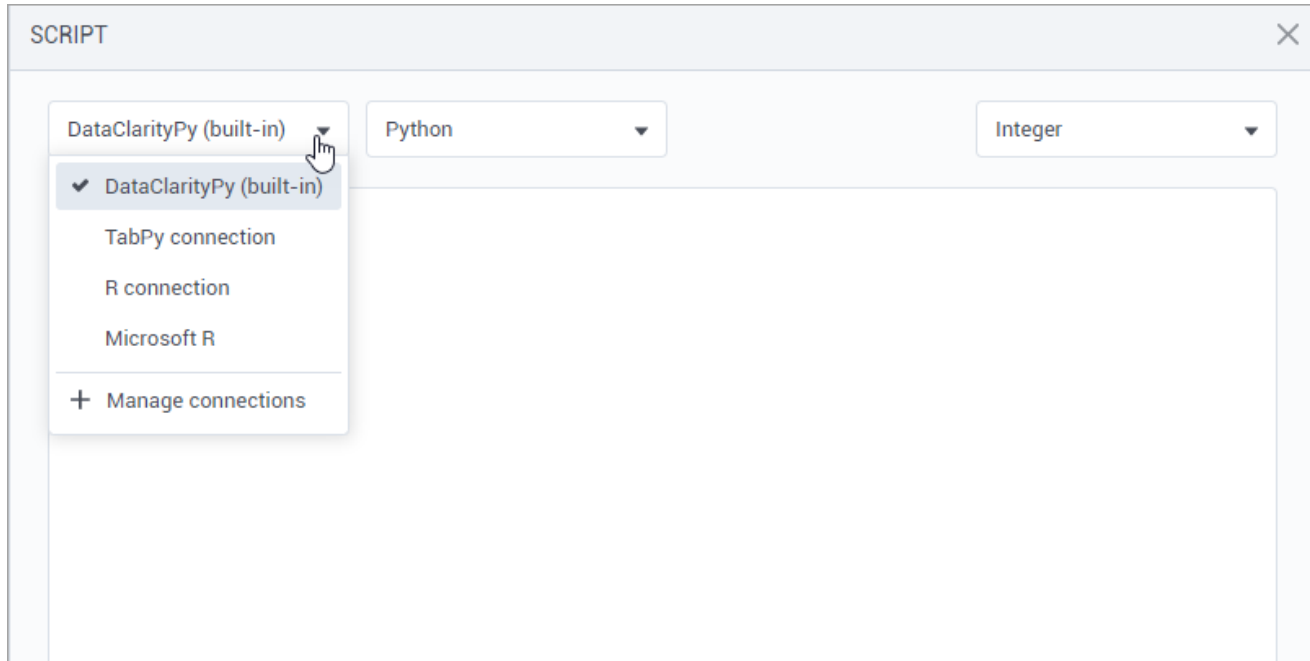


To import an AI connection, upload a corresponding ZIP file with the connection. In the **AI connections** pane, clicking the **Import** button will open the dialog for the file import. You can specify whether to overwrite the connection if the same data connection has already been imported. If the same connection already exists, then the data connection is not imported.



### *Built-in DataClarity Python server*

Now you don't need to create an AI connection to use the built-in DataClarity Python server. It is available right away when working with a script function.



### *Encrypted data connections credentials*

The data connection credentials are now stored in an encrypted format to ensure the security of data sources.

## STORYBOARDS

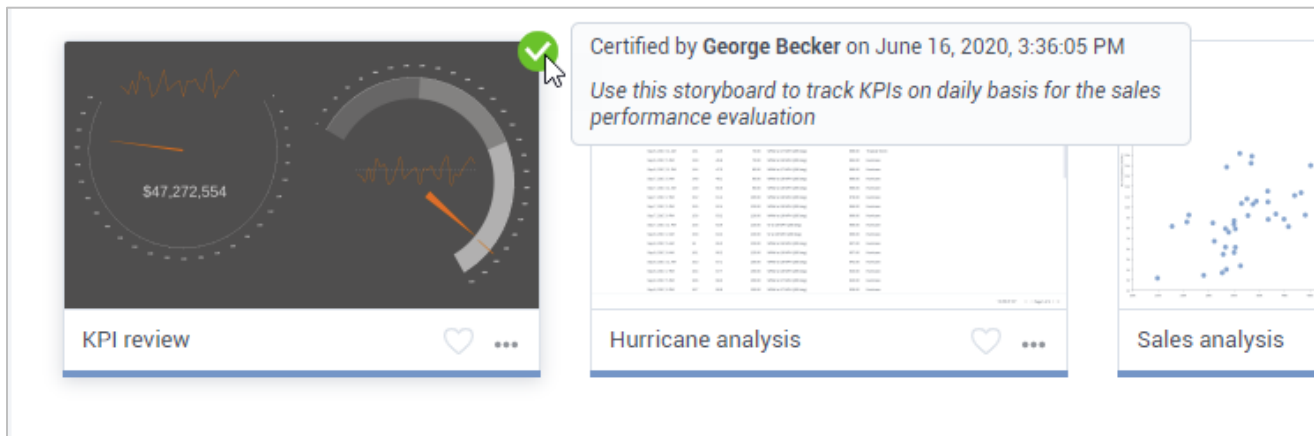
### *Certify storyboards*

To help users find storyboards that are trusted and recommended for their analysis, you can certify the storyboards. For best search results, only certify the storyboards that are valuable and considered the official source of enterprise information. Storyboard certification ensures the data is reliable and can be used across the organization.










Before performing this task, confirm that users have the dedicated permission to certify storyboards, granted in Access Manager. If you can certify storyboards, you can find the **Certify** option under the **More actions** menu for the storyboard. Users can mark storyboards as certified by selecting the **Certify this storyboard** check box. In addition, you can add a note about the certification status with descriptive information.

The screenshot displays the DataClarity interface with two storyboards: "KPI review" and "Sales analysis". The "KPI review" storyboard features four circular gauges with values like \$47,272,554. The "Sales analysis" storyboard shows a scatter plot of data points. A "More actions" menu is open for the "KPI review" storyboard, listing options: Modify, Subscribe, Share, Publish to web, Duplicate, Download, **Certify** (highlighted with a hand cursor), and Delete. A "CERTIFY STORYBOARD" dialog box is open in the foreground. It contains the text: "By certifying a storyboard, you verify it complies with your organization's standards". Below this is a checked checkbox labeled "Certify this storyboard". There is a "Note" field with a question mark icon, containing the text: "Use this storyboard to evaluate all KPIs for the daily sales performance". At the bottom of the dialog are "CANCEL" and "SAVE" buttons. A green checkmark icon is visible in the top right corner of the "Sales analysis" storyboard.

The storyboards have a green check mark in the upper-right corner of the tile. Pointing to the certification icon will reveal additional certification information.

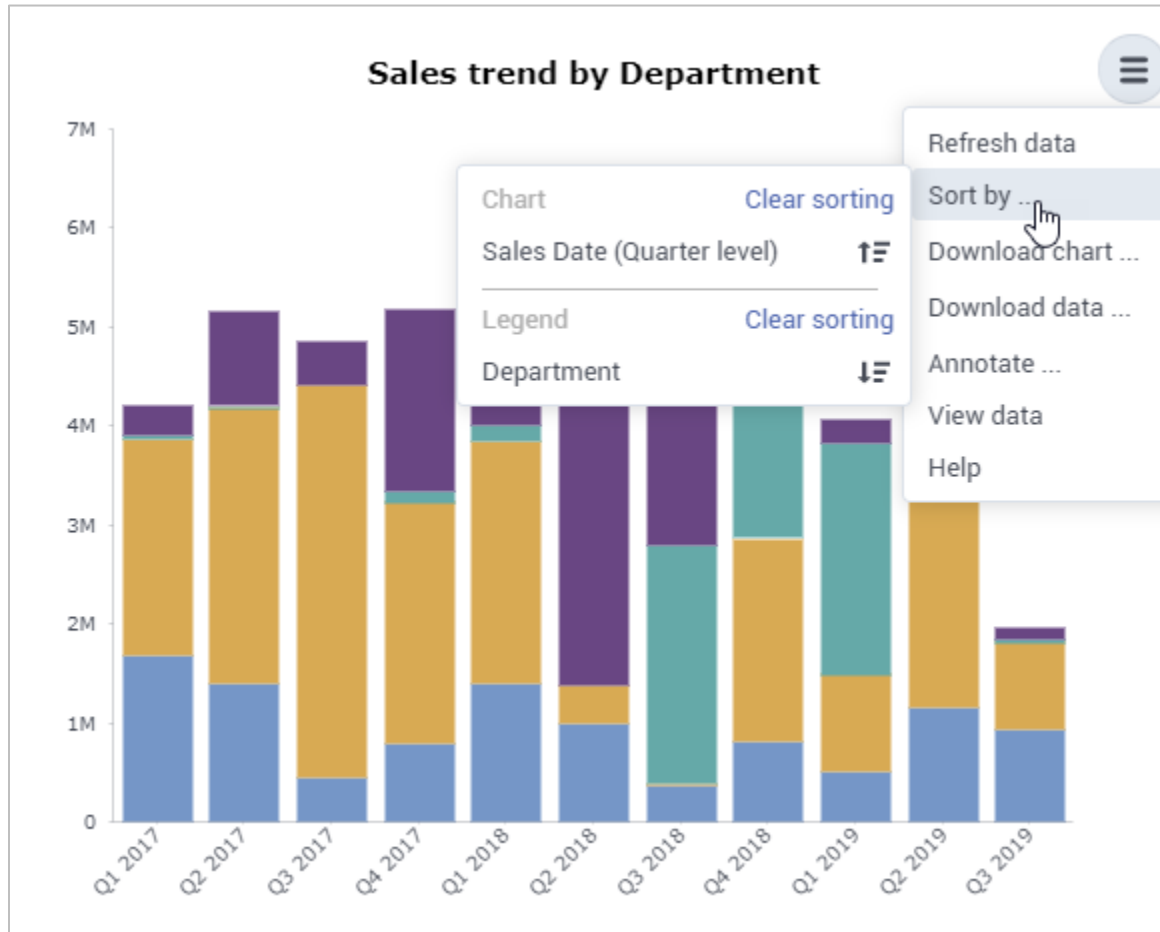


In **List** view, the certification icon appears in the first column in front of the certified storyboard.

Name				Last updated	Subscribers
		Calculations	...	June 19, 2020, 3:45:23 PM	
		KPI review	...	June 18, 2020, 4:06:00 PM	
		Sales analysis	...	June 2, 2020, 2:23:28 PM	

### Enhanced sorting in widgets

When sorting data within the widget, the columns you sort are now shown under dedicated sections. Click the data to change the sorting order. Additionally, you can clear sorting that was previously applied to the widget.



## Background image for the Table widget

In this release, users can further style the **Table** widget by setting a background image.

The screenshot displays the DataClarity Table widget interface. The left sidebar contains configuration options for the table, and the main area shows a table with 5 columns and 20 rows. The background image is a stack of business cards.

**Configuration Options (Left Sidebar):**

- Palette and style (+)
- Header (+)
- Rows (+)
- Sidebar (+)
- Toolbar (+)
- Pagination (-)
  - Show pagination: On (toggle)
  - Background color: (dropdown)
  - Font color: (dropdown)
  - Font size: 14 (input)
- Background (+)
- Title (+)
- Margins (+)
- Number precision (+)

**Table Data:**

Items per Transaction	Gross Sales	No of Customers	Profit	Manufacturing Cost
33.00	167.40	20.00	147.40	3.60
33.00	1.98k	15.00	1.82k	40.80
37.00	1.02k	30.00	935.66	21.50
17.00	1.32k	12.00	1.21k	28.80
39.00	323.00	21.00	291.14	6.80
41.00	1.43k	12.00	1.31k	30.00
16.00	1.92k	10.00	1.77k	40.00
46.00	176.40	12.00	155.90	3.60
24.00	736.96	10.00	673.00	15.68
11.00	27.65k	9.00	25.52k	576.00
22.00	810.00	8.00	739.72	18.00
26.00	1.19k	9.00	1.09k	25.00
23.00	450.80	12.00	408.78	9.80
44.00	436.50	15.00	396.05	9.00
6.00	167.40	12.00	147.45	3.60
46.00	593.75	10.00	540.92	12.50
24.00	3.84k	12.00	3.54k	83.52
16.00	323.00	12.00	291.08	6.80
33.00	1.43k	14.00	1.31k	30.00
10.00	1.92k	9.00	1.77k	40.00
35.00	176.40	8.00	155.91	3.60
48.00	736.96	11.00	673.08	15.68
49.00	864.00	21.00	790.58	18.00
28.00	810.00	11.00	739.76	18.00

**Footer:** 1 to 50 of 4,275 | Page 1 of 86












### Visualize image columns in the Table widget

In this release, using the **Table** widget, you can visualize images provided as blobs or as URLs pointing to the respective files.

When selecting such columns in the **data** tab, specify one of the available aggregation options to ensure the image is interpreted accordingly:

- **Image as Base64**
- **Image as Hex**
- **Image as URL**

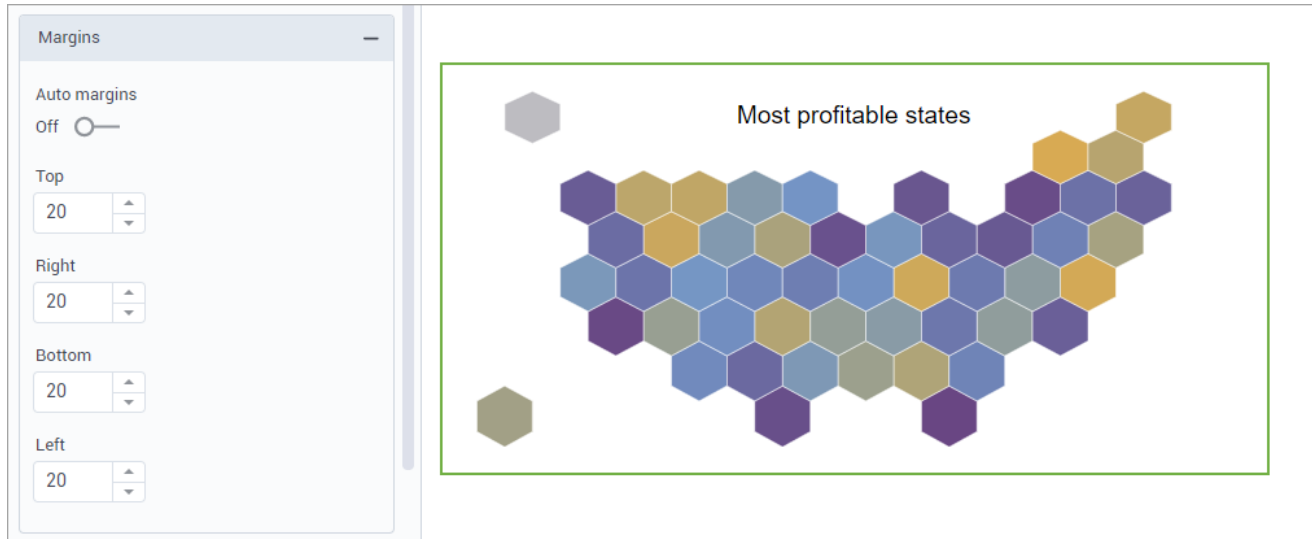
The screenshot shows the DataClarity Table widget interface. On the left, the 'Data' tab is active, showing the 'Sample - US Sales' dataset. The 'Columns' section lists 'Manager', 'Manager Avatar (url)', 'Country', and 'No of Cust'. A dropdown menu is open for 'Manager Avatar (url)', showing options: 'NONE', 'Image as Base64', 'Image as Hex', and 'Image as URL'. The 'Image as URL' option is selected. The 'Table' widget is also visible in the top left corner. On the right, the table displays the following data:

Manager	Manager Avatar	Country	No of Customers
Baron Meyer		United States	12.88k
Brodie Patel		United States	8.12k
Bryant Hanna		United States	10.78k
Gregory Williams		United States	11.86k
Jaylee Fry		United States	4.34k
Beckham Freeman		United States	8.82k
Adyson Heath		United States	9.16k
Rylee Mckee		United States	8.33k
Alexandria Vaughn		United States	8.54k
Liana Whitehead		United States	7.92k
Edward Gentry		United States	6.7k



### *Margins for geospatial widgets*

In addition to the default margins, you can now define any margins needed to display your geospatial widgets. Just enter the value for the top, right, left, and bottom margins.



### *Customize lines and direction arrows in the Path map*

To better visualize map routes, the following customization capabilities have been introduced:

- **Scale line & arrow** – To apply an automatic scale to the path lines and arrows. This way, the size of the lines and arrows does not change when you zoom a map.
- **Arrow width** – To set the width of the direction arrow.
- **Arrow height** – To set the height of the direction arrow.

## GENERAL

*Change content ownership when deleting a user*

In this release, you have more flexibility with managing content that was created by the platform users. If a content owner is deleted in Access Manager, administrators can now choose whether to delete the content permanently or reassign it to another user. This feature may be especially useful if you need to comply with the GDPR requirement on personal data deletion from the system.

**DELETE USER** [X]

Are you sure you want to permanently delete the user John Flemming?

Select owner ?

george.becker

You can select a new owner to reassign the content of John Flemming. In this case, ensure all the respective permissions are granted to the owner. If you don't select a new owner, all the content of John Flemming will be permanently deleted.

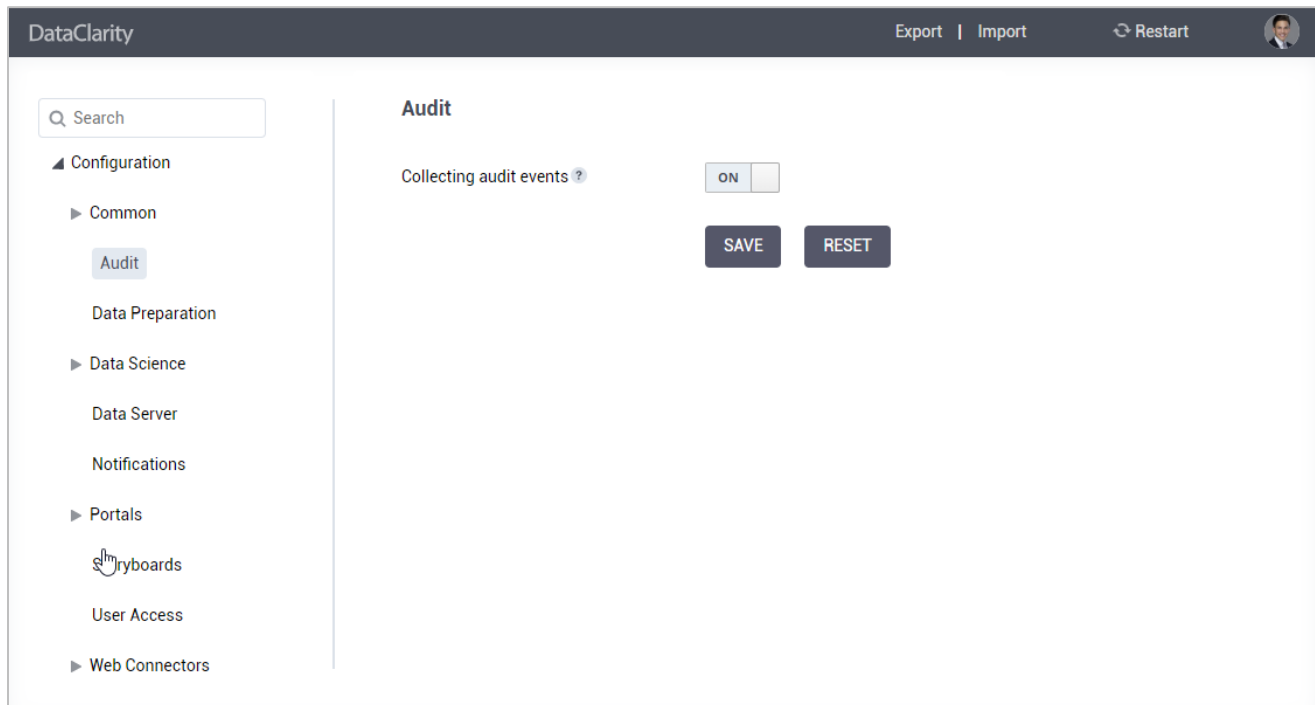
[X] [v]

**CANCEL** **DELETE**

If the content that is being transferred requires new user authentication, the new owner will receive an email asking to get authenticated in the application.

### Audit user activities

Administrators can now use the auditing service to track various user activities in Storyboards and Data Preparation. The new configuration setting, **Collecting audit events**, has been added to the Configuration Manager to allow administrators to activate or deactivate the audit logging.



### Configure notifications lifespan

DataClarity administrators can now configure for how long to retain the notifications and storyboard images linked to them. By default, the notifications and the storyboard images are cleaned regularly to prevent storage overflow. The default notification lifespan is set to 30 days since the notification creation date.

The screenshot shows the DataClarity web interface. The top header bar contains the 'DataClarity' logo, 'Export | Import' links, a 'Restart' button with a circular arrow icon, and a user profile icon. On the left, a sidebar menu includes a search bar and a list of configuration categories: Configuration (expanded), Common, Audit, Data Preparation, Data Science, Data Server, Notifications (highlighted), Portals, Storyboards, User Access, and Web Connectors. The main content area is titled 'Notifications'. It features a label 'Notifications lifespan (in days) ?' next to a text input field containing the value '30'. Below the input field are two buttons: 'SAVE' and 'RESET'.

### Message broker configuration

The screenshot shows the DataClarity web interface for the 'Message Bus' configuration. The top header bar is identical to the previous screenshot. The sidebar menu on the left shows the 'Configuration' section expanded, with 'Common' also expanded. Under 'Common', the items 'Logging', 'Database', and 'Message Bus' (highlighted) are listed. The main content area is titled 'Message Bus'. It contains two configuration fields: 'Message broker username ?' with a text input field containing 'artemis\_user', and 'Message broker password ?' with a password input field showing masked characters '.....'. Below these fields are 'SAVE' and 'RESET' buttons.