DataClarity

WHAT'S NEW AND RELEASE NOTES

Release: 2022.1

DataClarity

IN THIS RELEASE

NEW FEATURES AND IMPROVEMENTS 2

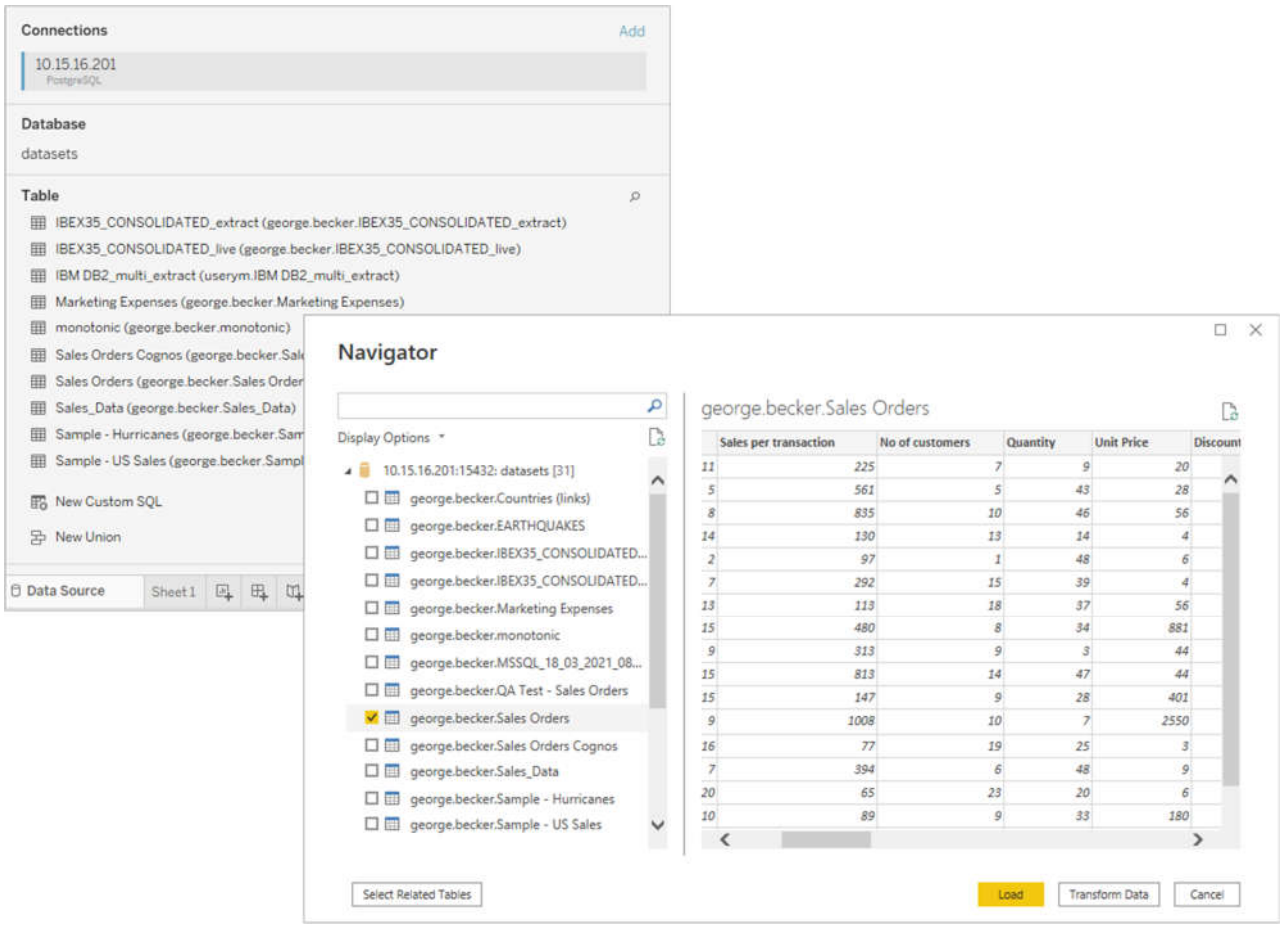
NEW FEATURES AND IMPROVEMENTS

DATA PREPARATION

Dataset names when using other BI tools

Previously, when accessing DataClarity datasets from a third-party BI tool like Tableau or Power BI, you had all private and shared datasets displayed under a single public schema. To avoid possible name duplicates, all the datasets' names had an ID automatically appended. For example, "Sales Orders" was displayed as "sales_orders_1234". The approach has been changed to reflect the original names as displayed in Data Preparation. Thus, starting with this release, you will find datasets displayed under different schemas based on the dataset ownership, for example:

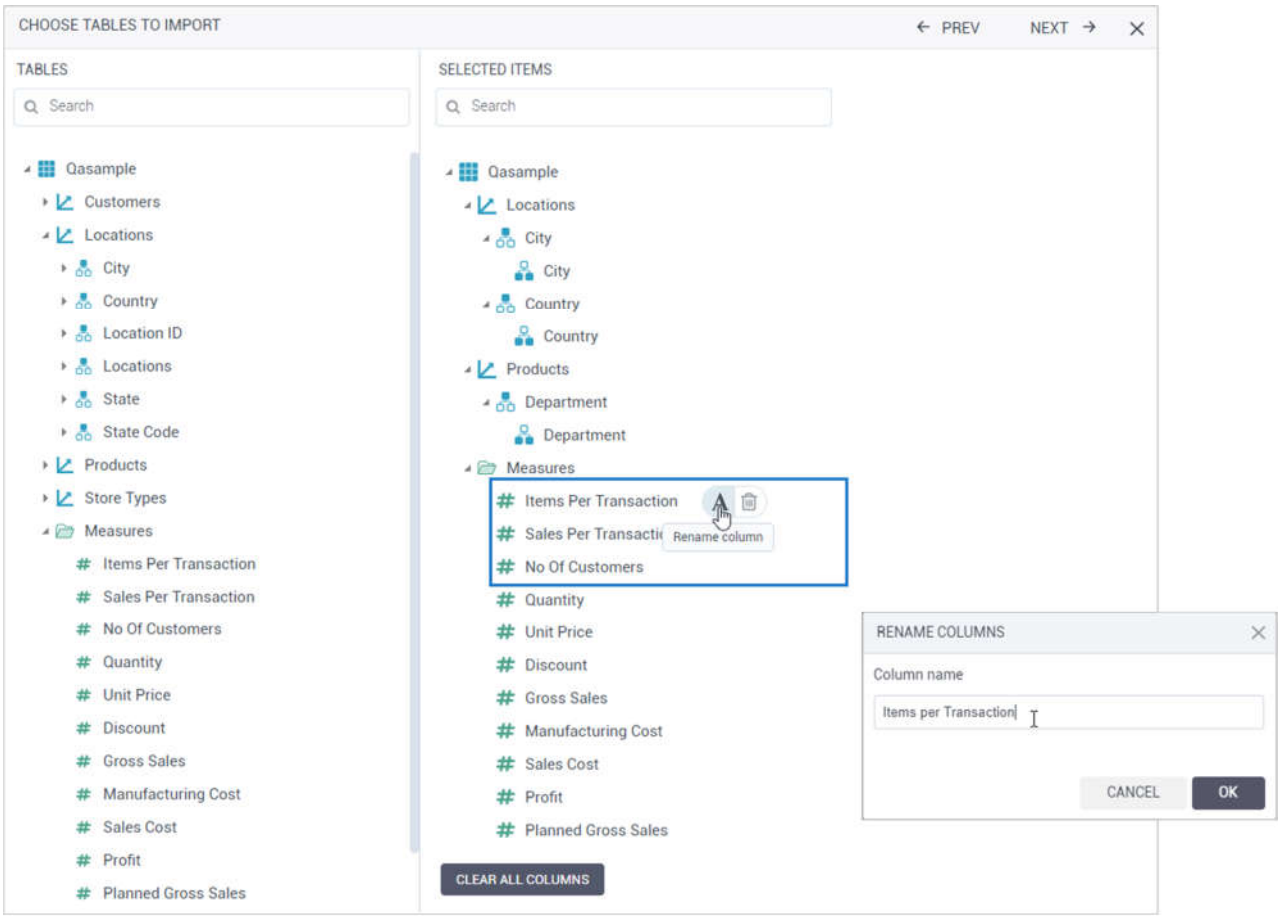
- george.becker.Sales Orders (a dataset created by your user George Becker)
- johnsmith.Sales Orders (a dataset with the same name that was shared by John Smith)



Copyright © 2022 DataClarity Corporation

Rename individual columns

In addition to renaming all the columns in bulk on the **Choose tables to import** page, you can now rename each column individually when adding a new data source.

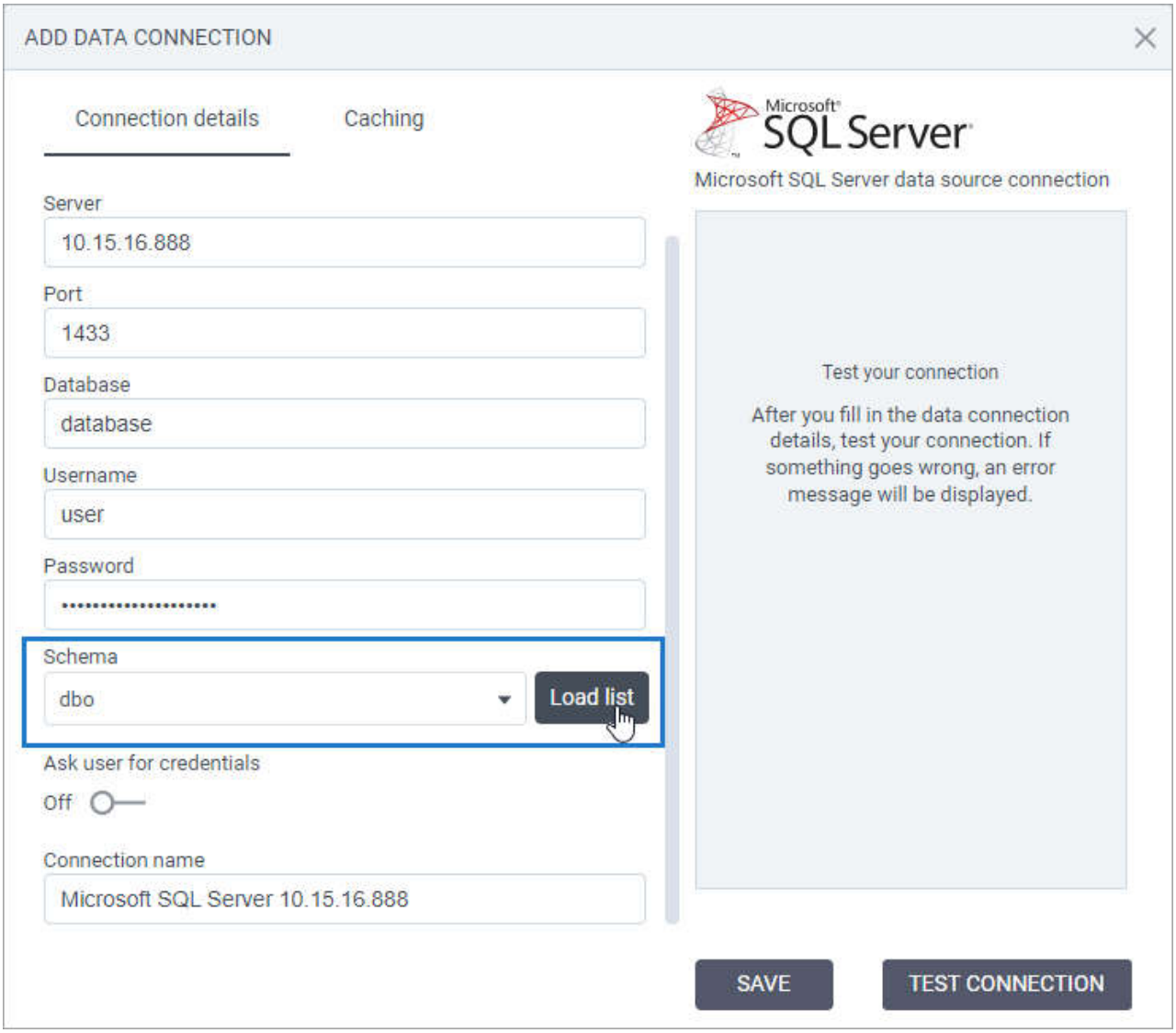


Select schemas for data sources

Starting with this release, you need to specify a database schema when adding a data connection to the following sources:

- Microsoft SQL Server
- PostgreSQL
- DB2
- Apache Derby
- Google Cloud SQL
- PostgreSQL

By selecting a schema, you narrow down the data for the data connection and improve query engine performance. After you enter server credentials, click **Load list** to view available schemas in the database and select one. If you edit a data connection created before this release, you will be prompted to select a schema.

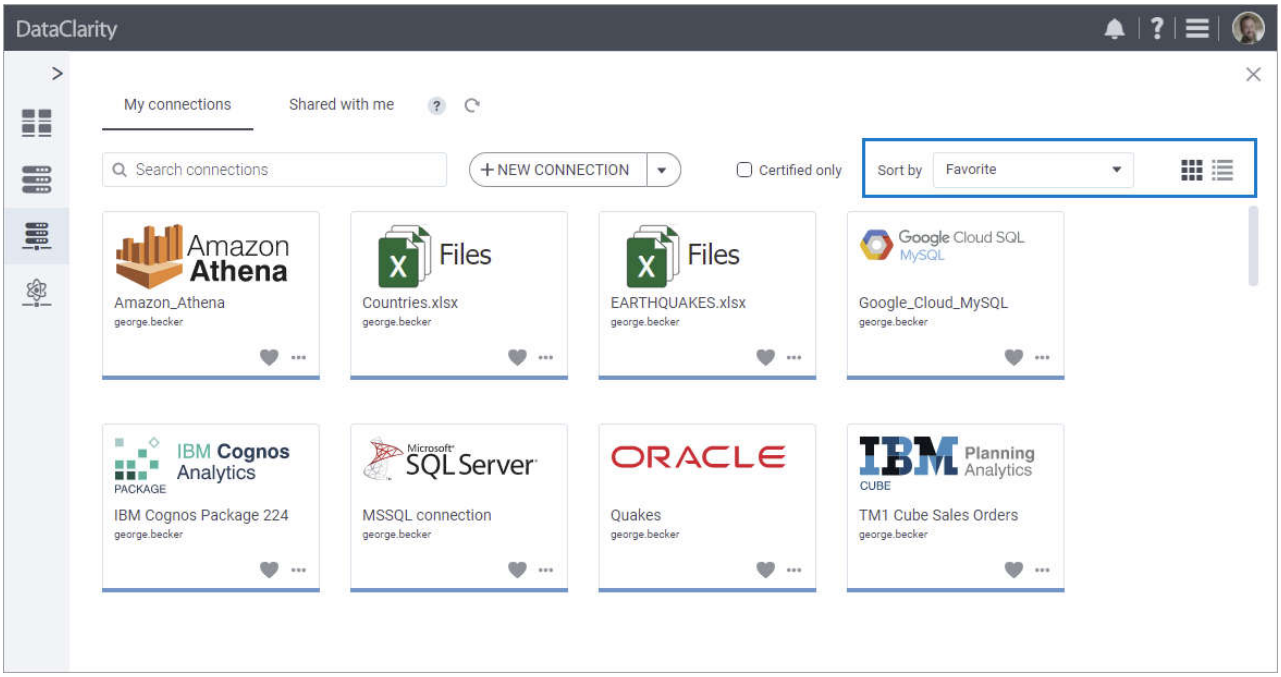


Copyright © 2022 DataClarity Corporation

Save your viewing preferences

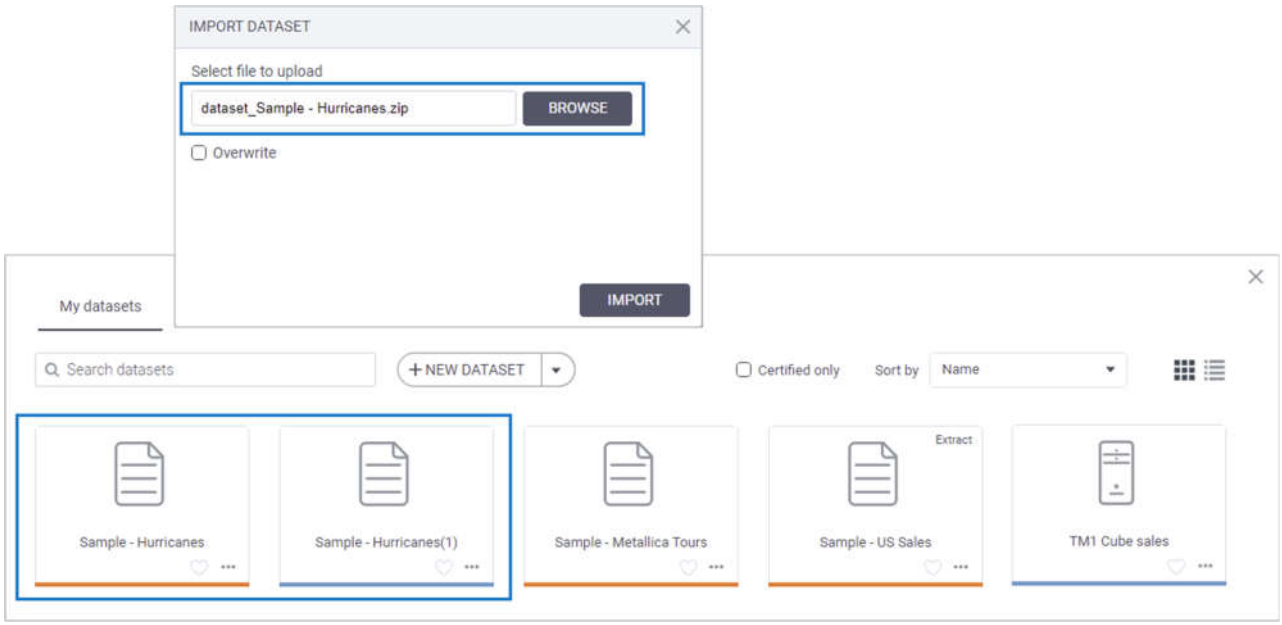
Previously, when you switched to **List view** or selected a different sorting option to view datasets, the selections were saved only within the user session. In other words, the view and sorting were restored to the default values with each subsequent login. The user experience has been improved by saving your viewing preferences using the browser's cookies. Additionally, the default soring has been changed to **Last created** to have the most recent resources listed first.

The same improvements are applied to data connections and AI connections.



Automatically resolve duplicate names for the imported datasets

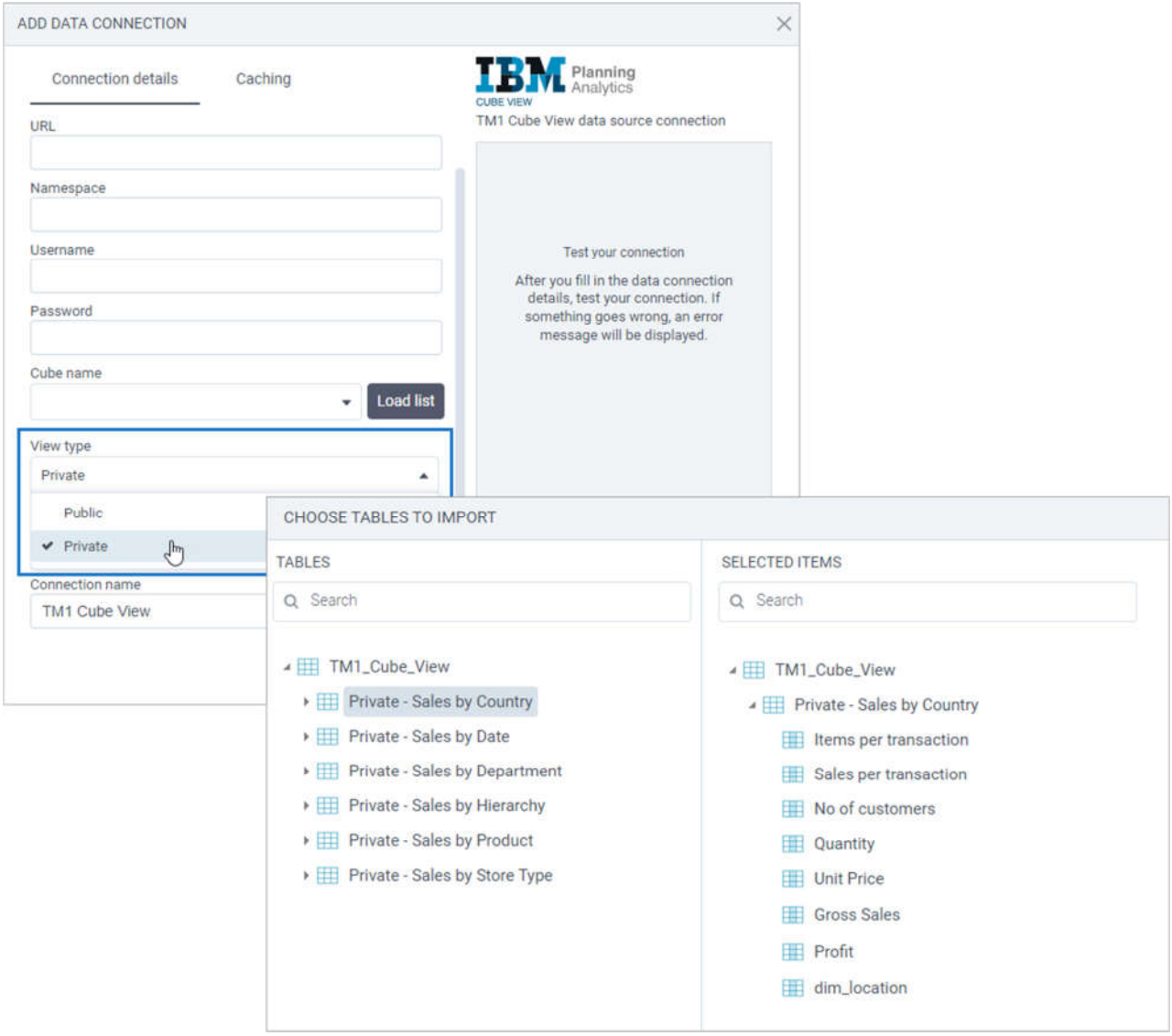
Now, if you are importing a dataset with the name that already exists in the **Datasets** pane, the name of the imported dataset will automatically include an index number in parentheses. For example, if you are importing the dataset named “Sales” and you already have a dataset with this name, it will be imported as “Sales(1).”



Copyright © 2022 DataClarity Corporation

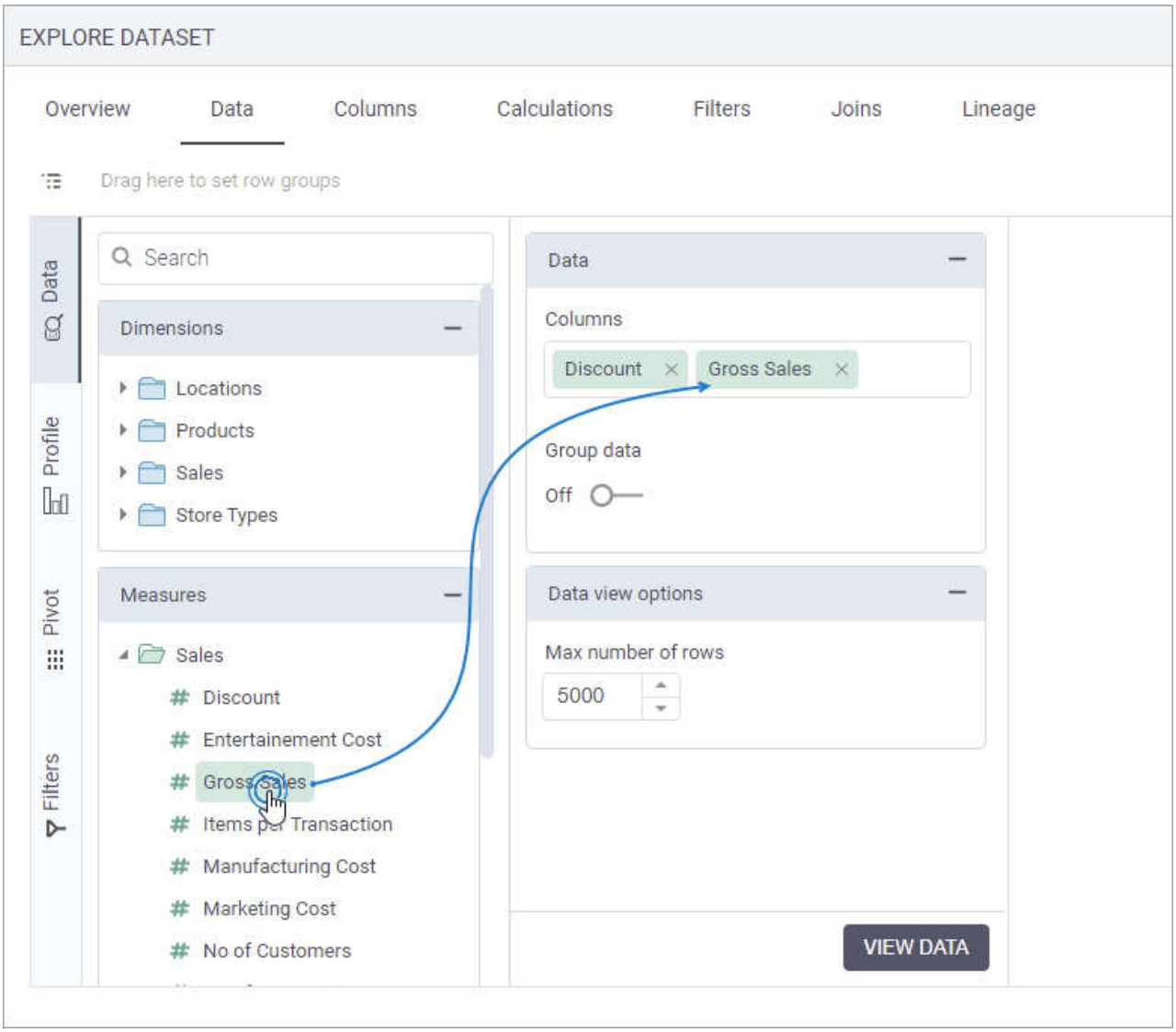
Select private views to create TM1 cube view datasets

Previously, after creating a data connection to a TM1 cube view, you could select a public view for a dataset. Now, you can create data connections with access to your private views. The new **View type** dropdown allows you to select the type of cube view: **Private** or **Public**. The **Public** option is selected by default.



Double-click to select columns for data exploration

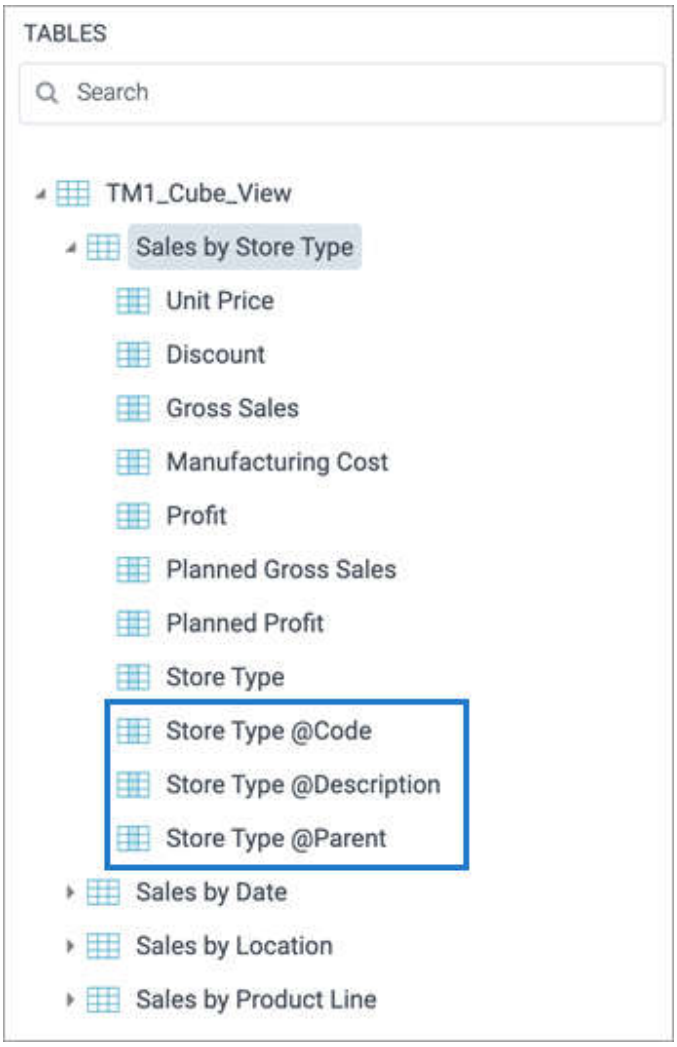
Previously, to select columns for data exploration, you needed to drag them into the **Columns** field of the **Explore dataset** pane. Now, you can quickly add columns by double-clicking them under the **Dimensions** and **Measures** sections.



Copyright © 2022 DataClarity Corporation

Create a dataset based on a TM1 MDX view

Previously, you could create datasets based on native TM1 cube views. The TM1 cube view driver has been extended to support MDX views (named MDX expressions stored inside the TM1 Server). Now, after creating an **IBM Planning Analytics / TM1 Cube View** connection, you can select MDX views for your datasets. Moreover, you can select attributes for MDX views the same way as you do for native TM1 views. The attributes are listed as columns that follow the "column_name @attribute_name" naming pattern.



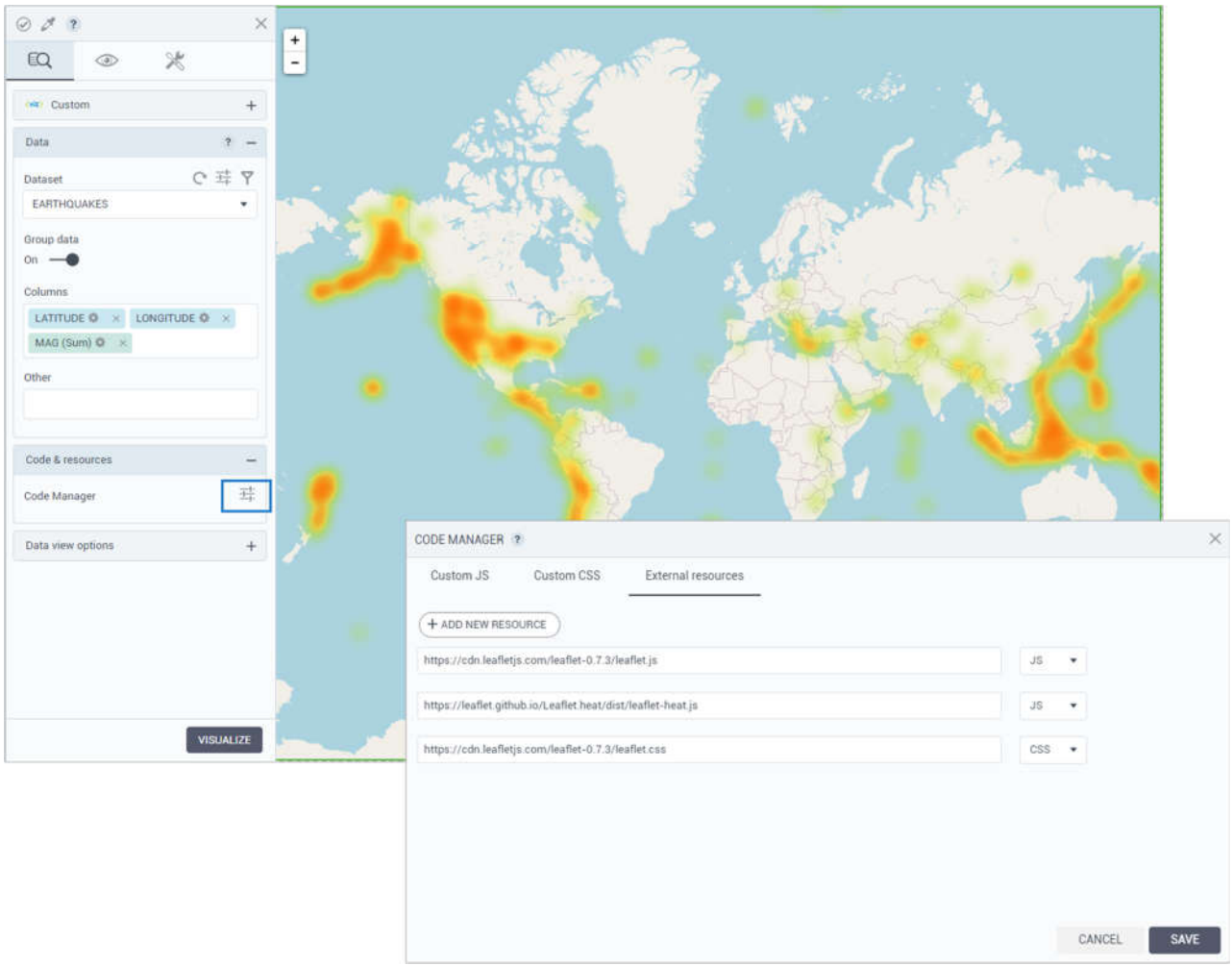
STORYBOARDS

Custom visualization widget

If you want to visualize your data in a specific way that is not possible with the visualizations included in Storyboards, you can build custom visualizations. The new **Custom viz** widget allows you to write custom JavaScript code by using popular third-party visualization libraries such as D3, amCharts, or Charts.js.

To create a custom visualization, select data columns that generate the query, and then write the code to process and visualize the data in the **Code Manager** using the following tabs:

- **Custom JS** – Add your custom JavaScript code.
- **Custom CSS** – Specify custom CSS code.
- **External resources** – Link the JavaScript and CSS resources from an external library.

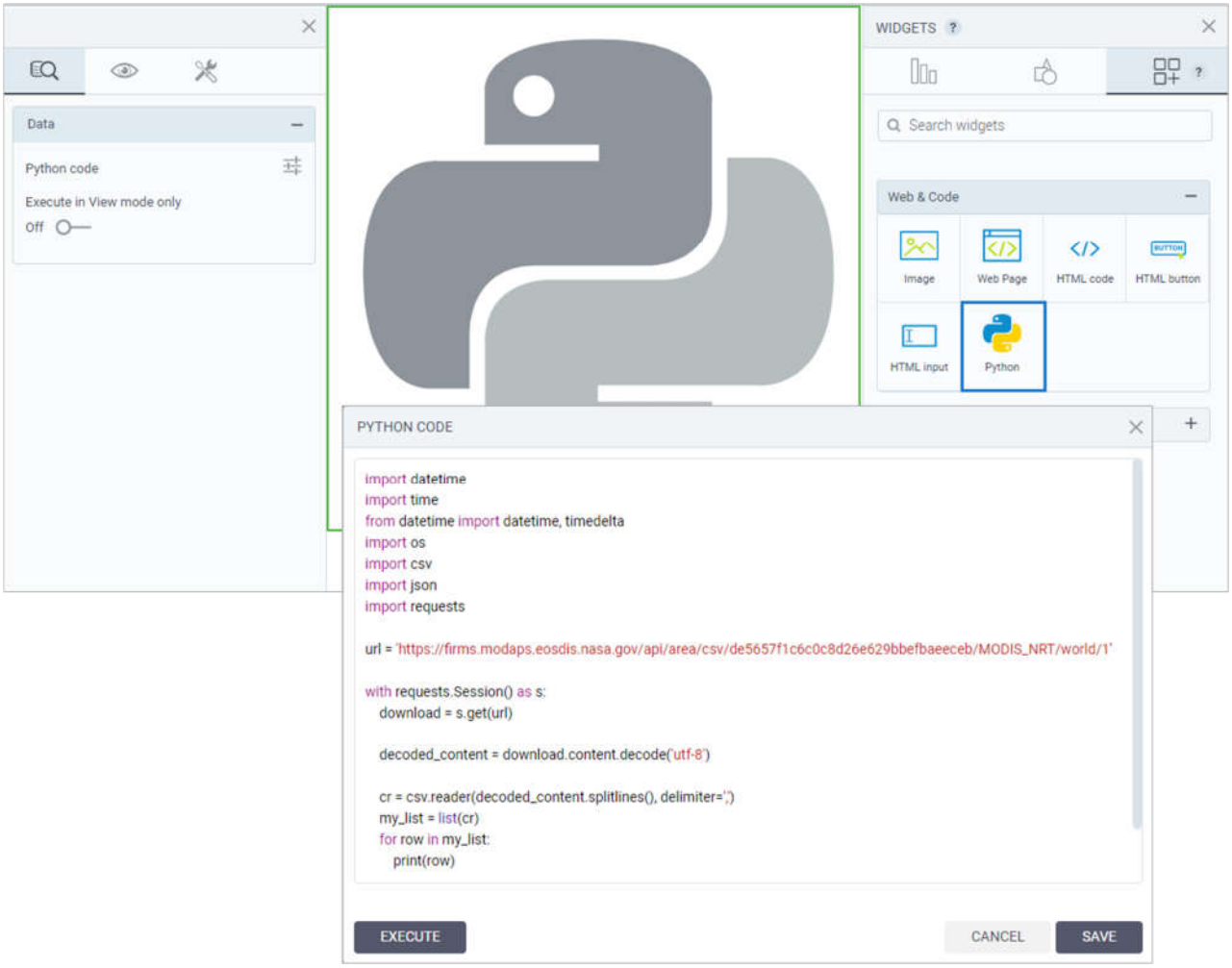


Copyright © 2022 DataClarity Corporation

Python code widget

Starting with this release, power users can execute Python code on a storyboard's page. The code is executed directly on the built-in Python server and allows using the results to feed other custom visualizations and, therefore, enhance the data analysis.

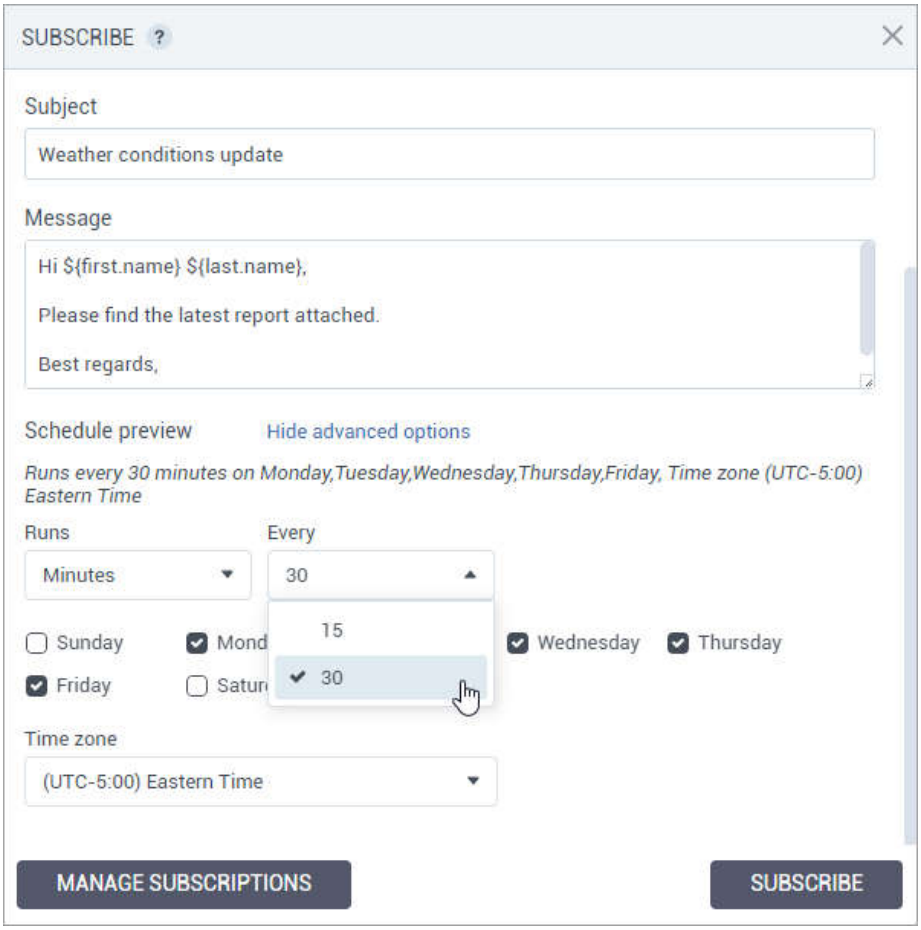
You can find the new widget in the **Widgets** pane, on the **Other widgets** tab under the **Web & Code** category. You can add your code by clicking **manage python code** on the **data** tab. To preview and verify the code results, click **Execute**. If you do not want to run the code when your storyboard is in Edit mode, turn on **Execute in View mode only**. This way, switching a storyboard to View mode will be a trigger to execute the code.



Schedule a storyboard subscription to run every 15 or 30 minutes

With a storyboard subscription, the subscribers receive emails with storyboard snapshots on a scheduled basis. Previously, you could schedule a subscription on a monthly, weekly, daily, and hourly basis. Starting with this release, you can schedule subscription emails to run every 15 or 30 minutes and select specific days of the week.

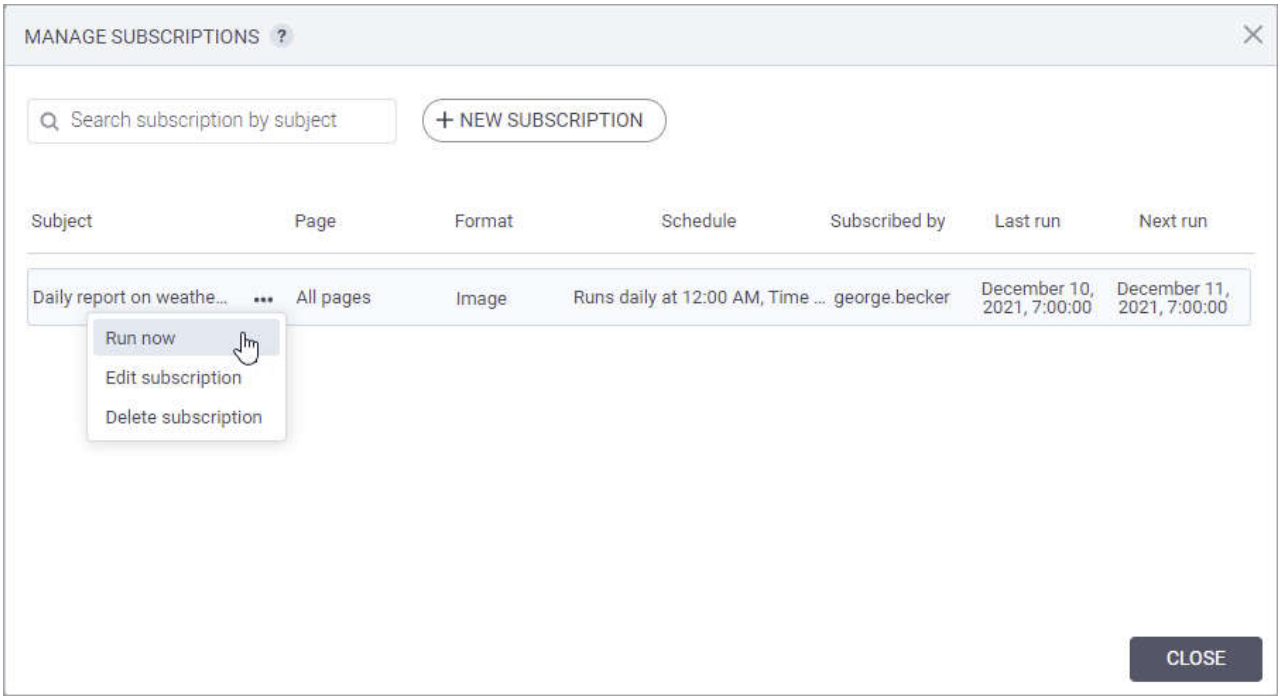
To set the new frequency, open the **Subscribe** dialog, click **Show advanced options**, and in the **Runs** dropdown, select **Minutes**.



Copyright © 2022 DataClarity Corporation

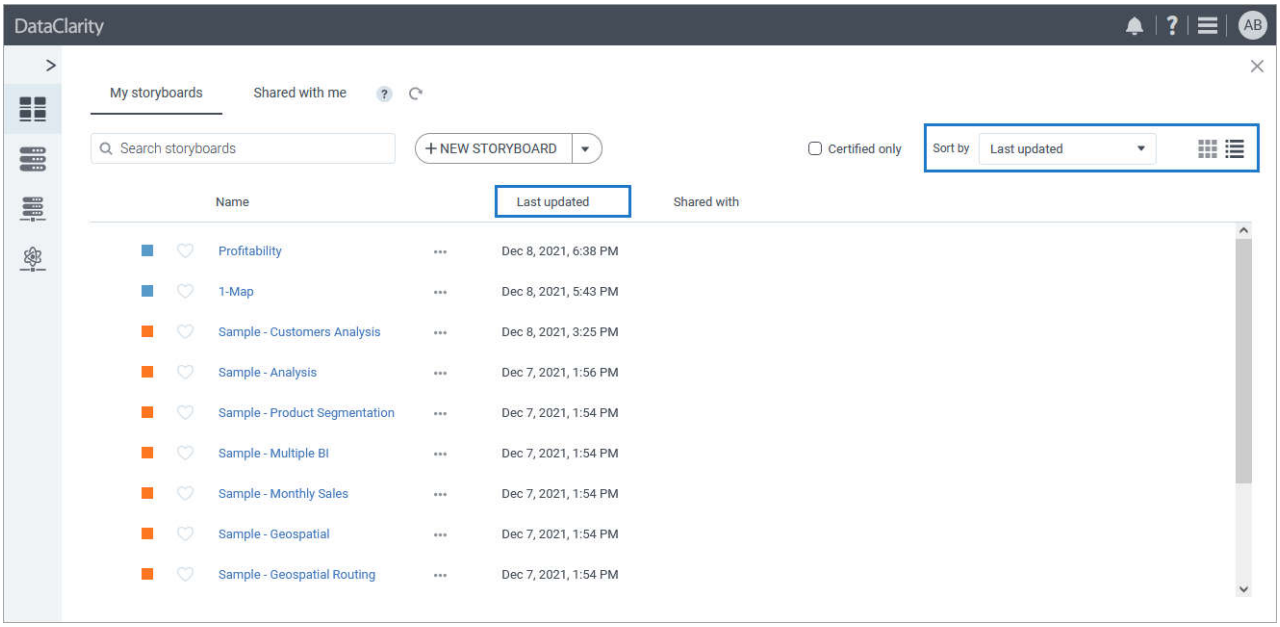
Send a storyboard subscription email on demand

Now, you can test the added subscription jobs by sending an email on demand. In the **Manage subscriptions** window, for a subscription, point to **More actions** and select the new menu option—**Run now**. After confirmation, the subscription email request is sent immediately. We recommend using this option only for testing purposes or in urgent situations.



Save your viewing preferences

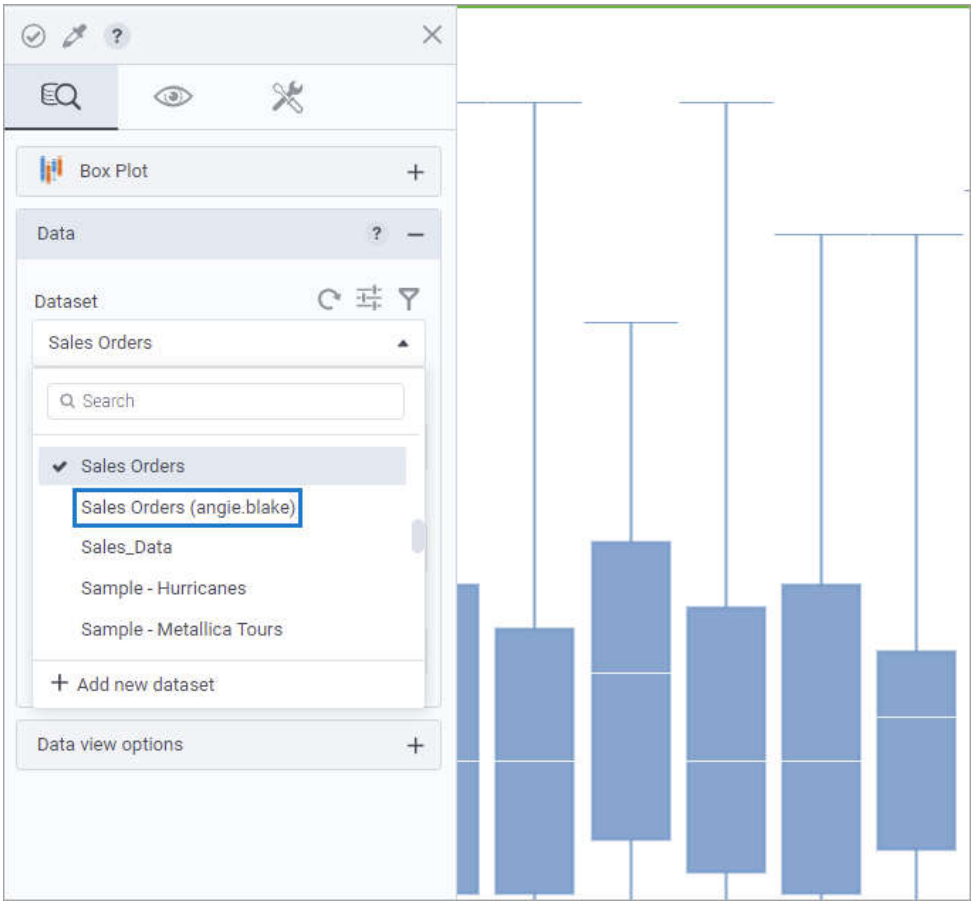
Previously, when you switched to **List view** or selected a different sorting option to view storyboards, the selections were saved only within the user session. In other words, the view and sorting were restored to the default values with each subsequent login. The user experience has been improved by saving your viewing preferences using the browser's cookies. Additionally, the default sorting has been changed to **Last created** to have the most recent storyboards listed first.



Copyright © 2022 DataClarity Corporation

Distinguish shared datasets in a visualization

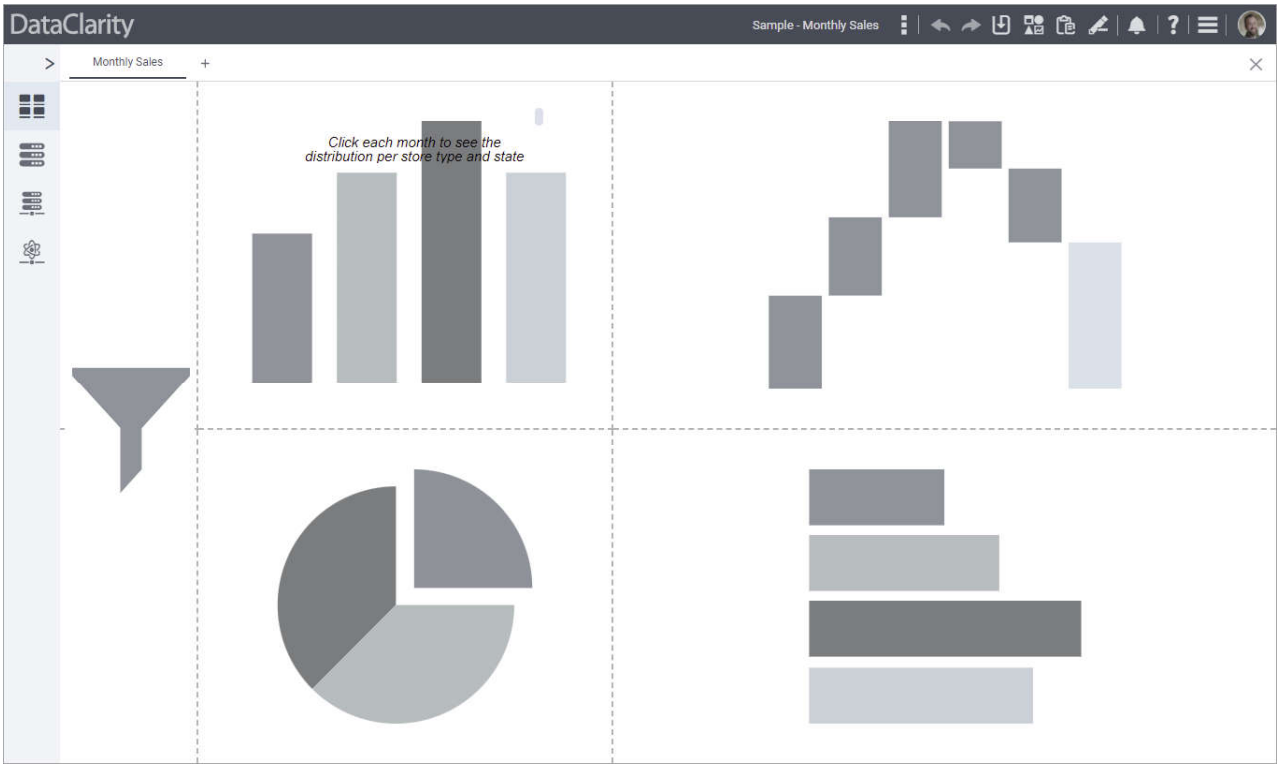
Previously, when selecting a dataset for visualization, you could not differentiate between your datasets and those shared with you. Now, if a dataset is shared with you, you can view its owner's username in parentheses next to the dataset name in the **Dataset** dropdown. For example, "Sales Orders (angie.blake)" is a dataset shared by the user Angie Blake.



Edit a storyboard without running the widgets

Previously, when opening a storyboard for editing, all its widgets ran automatically. As a result, for storyboards containing many widgets with complex data science calculations, any modification was time-consuming. The UX has been improved, and now, when you open a storyboard by selecting **More actions** > **Modify** > **Edit**, the visualizations do not run automatically. This way, you can quickly modify any widgets on a storyboard.

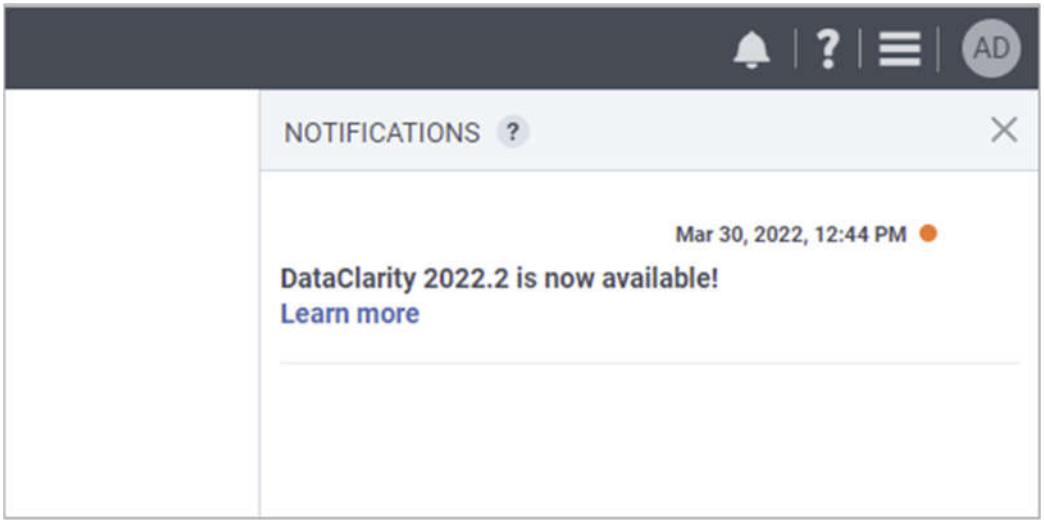
You can visualize all the widgets by switching a storyboard to View mode. You can still run each widget individually by clicking **Visualize** on the widget settings pane.



Copyright © 2022 DataClarity Corporation

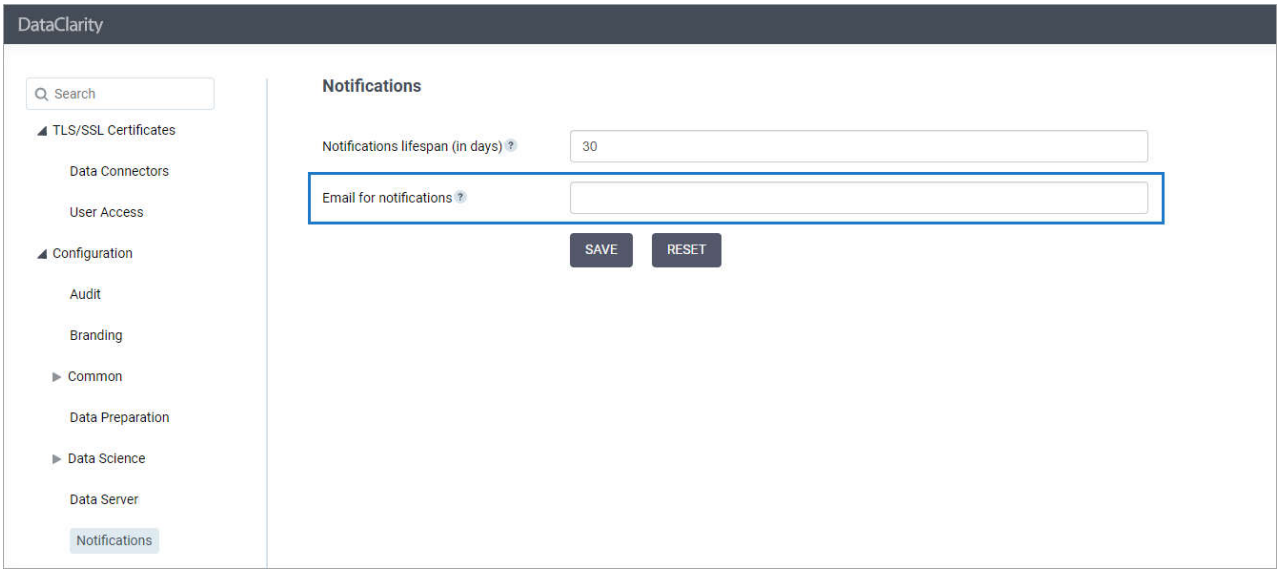
Receive notifications about new versions

Now you will receive notifications about each new version of the DataClarity Platform that is available for installation. The announcements appear in the **Notifications** pane and include a version number and the link to the *What's New and Release Notes* document. Additionally, you can specify an email for such notifications in **Configuration Manager** > **Notifications**, in the **Email for notifications** field.



Configure email for notifications

You can now specify an email for receiving notifications about each new version of the DataClarity Platform that is available for you to install. The new **Email for notifications** field has been added in **Configuration Manager** > **Notifications**.



Copyright © 2022 DataClarity Corporation

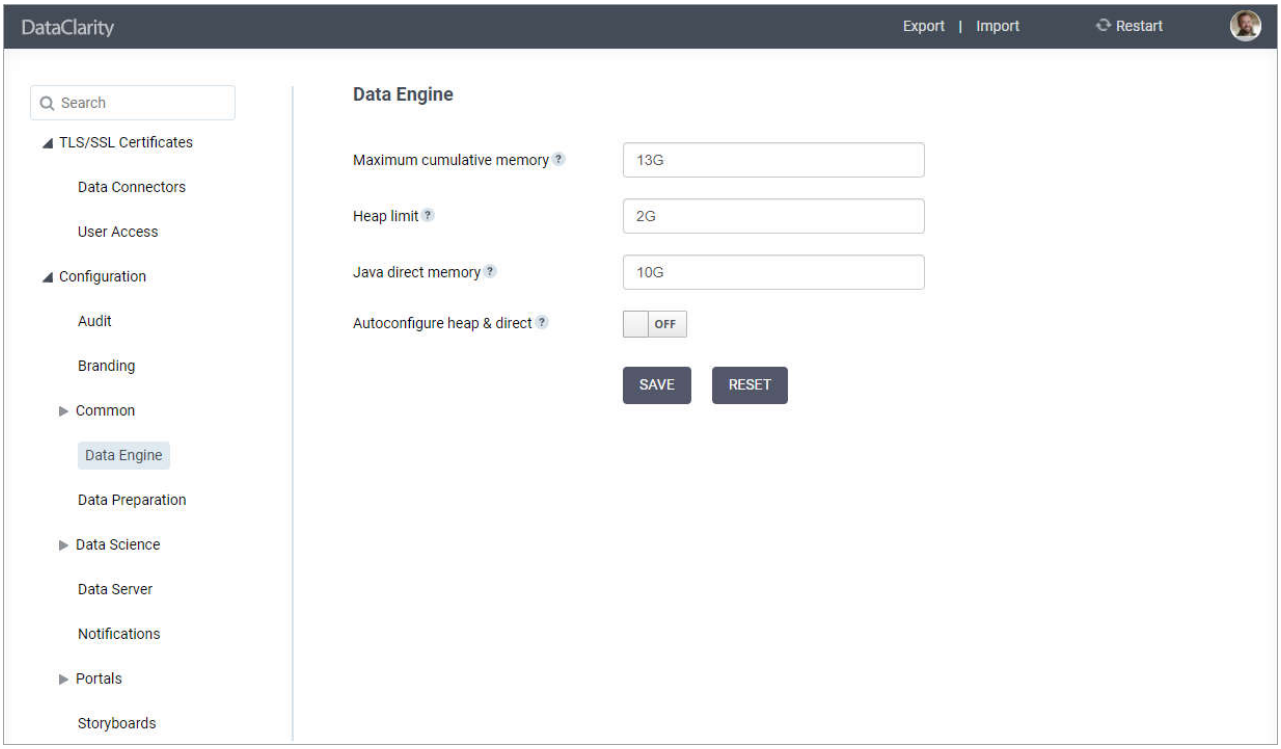
Configure Data Engine's memory settings

You can now control how much memory to allocate to Data Engine. This way, you can improve the data & query engine's speed and the platform server efficiency. In Configuration Manager, on the new **Data Engine** pane, you can find the following memory settings:

- **Maximum cumulative memory** – The maximum cumulative memory allocated to the Data Engine process during startup.
- **Heap limit** – The maximum theoretical JVM (Java virtual machine) heap limit.
- **Java direct memory** – Java direct memory allocated to query processing.
- **Autoconfigure heap & direct** – Choose how to define memory limits:
    - If **On**, Data Engine automatically determines the best allocation between heap and direct memory limits based on the specified **Maximum cumulative memory**. In this case, the values entered in the **Heap limit** and **Java direct memory** are ignored.
    - If **Off**, Data Engine uses the memory limits specified in the respective fields.

If you scale Data Engine to more pods, each pod will have the same memory limits. For example, if you have the max memory set to 16 GB, and you have two pods, then Data Engine uses 32 GB (16 x 2) as the limit.

For more information on how to allocate memory to Data Engine, refer to *Configuration Manager Help*.

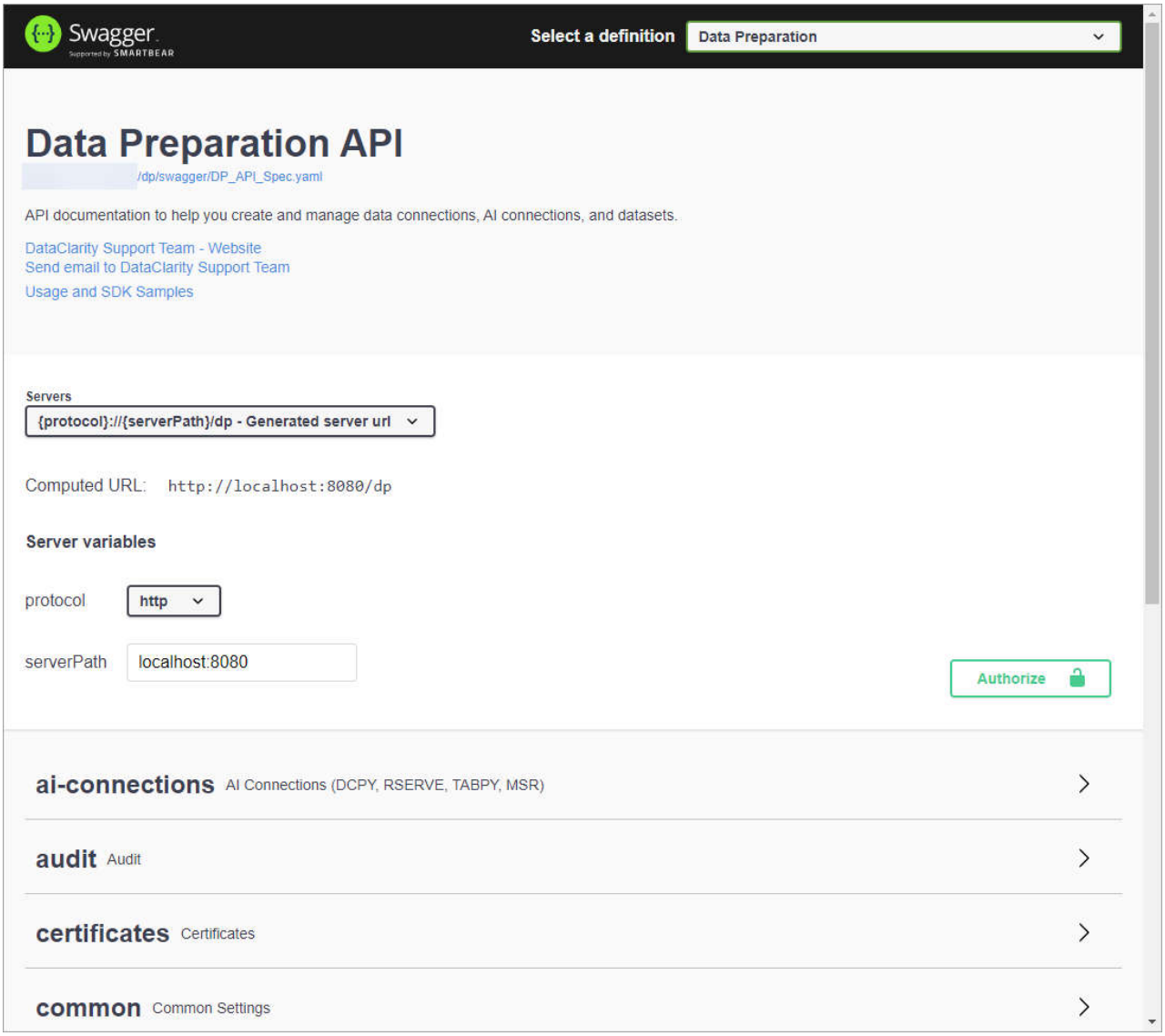


REST API

OpenAPI specification & Swagger

DataClarity provides REST APIs to let you leverage, automate, or incorporate DataClarity Platform functions into your website or application. The DataClarity's API is based on REST principles and provides standard HTTP methods for getting, creating, updating, and deleting the platform's resources.

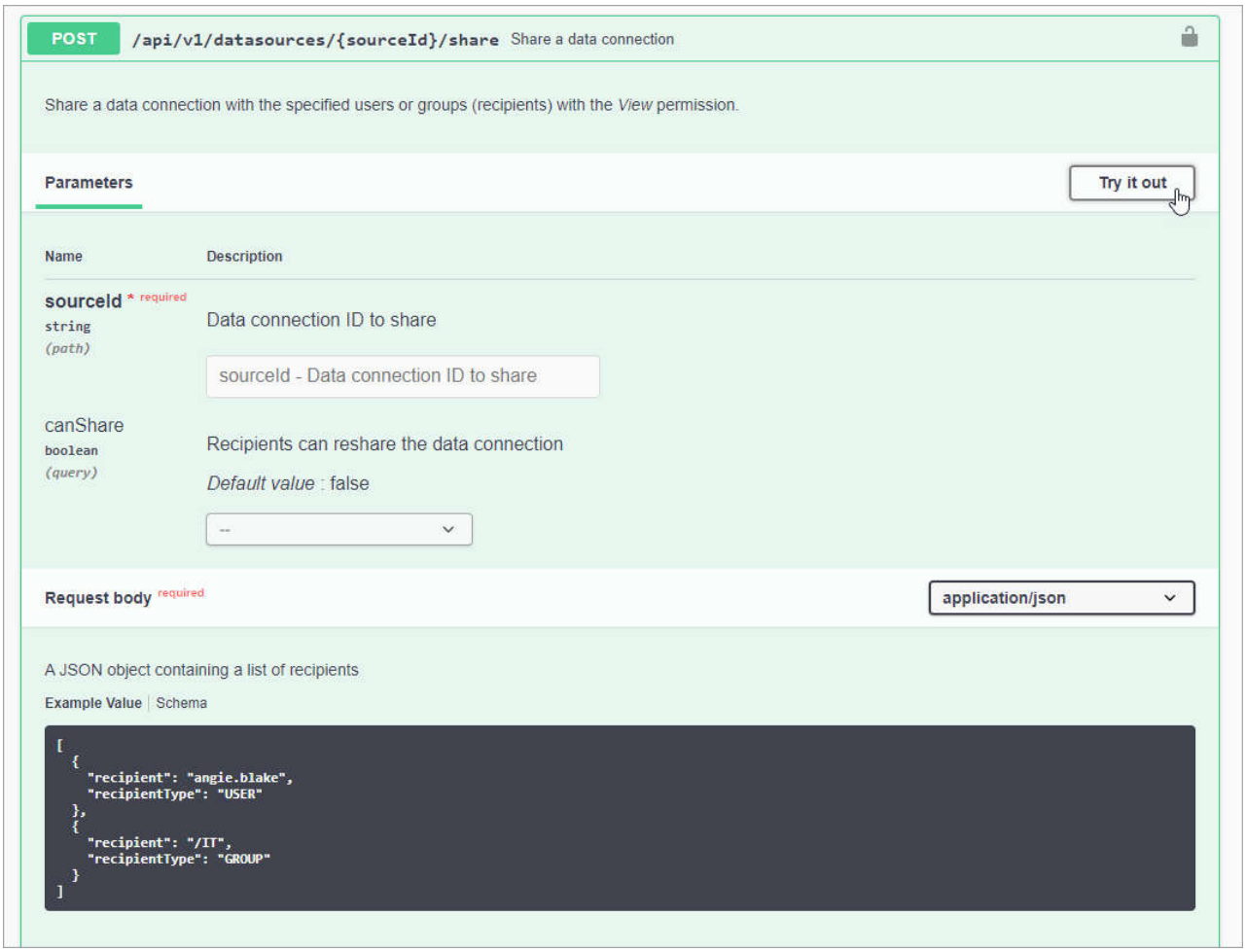
You can now benefit from the improved and restructured API reference documentation provided in Swagger, a fully interactive documentation tool that allows you to visualize and interact with API.



Copyright © 2022 DataClarity Corporation

Interact with API in Swagger

Each API request now includes a summary, description, and examples of a request body and response body where applicable. You can authenticate with the Bearer token and use the “Try it out” feature to experiment with the API before integrating it into your code.



OpenAPI

Swagger is generating the interactive API documentation based on the OpenAPI specification, namely the OpenAPI definition file version 3.0.3 in YAML format. You can download the file, view it in a text editor, or even import it in a Postman collection. The examples in the specification use sample data and credentials, so make sure you use your data for testing.

```yaml
openapi: 3.0.3
info:
  title: Data Preparation API
  description: API documentation to help you create and manage data connections, AI connections, and datasets.
  contact:
    name: DataClarity Support Team
    url: 'https://support.dataclaritycorp.com/hc/en-us'
    email: 'customercare@dataclaritycorp.com'
  version: ''
externalDocs:
  url: '/dp/'
  description: 'Usage and SDK Samples'
servers:
  - url: '{protocol}://{serverPath}/dp'
    description: Generated server url
    variables:
      protocol:
        enum:
          - 'http'
          - 'https'
        default: 'http'
      serverPath:
        # note! no enum here means it is an open value
        default: 'localhost:8080'
        description: this value is assigned by the service provider, in this example `prod.com`
tags:
  - name: ai-connections
    description: AI Connections (DCPY, RSERVE, TABPY, MSR)
  - name: audit
    description: Audit
  - name: certificates
    description: Certificates
  - name: common
    description: Common Settings
  - name: dataset-extract
    description: Dataset Extracts
  - name: datasets
    description: Datasets
  - name: datasource
    description: Data Connections
  - name: datasource-upload
    description: File Upload
  - name: demo-data
    description: Sample Content
  - name: preview
    description: Data Preview
  - name: tags
    description: Tags
  - name: user-setting
    description: User Settings
paths:
  '/api/v1/aiconnections':
    get:
      tags:
```


Copyright © 2022 DataClarity Corporation

API references

API reference documentation per each component is provided with the Platform's installation, where "localhost" is the name or IP address that was configured for the Platform.

- **Data Preparation API** – https://localhost/dp/swagger-ui/index.html
- **Storyboards API** – https://localhost/sb/swagger-ui/index.html
- **Scheduler API** – https://localhost/scheduler/swagger-ui/index.html
- **Notification API** – https://localhost/notification/swagger-ui/index.html