ONE - Workshop

DQ Advanced – Data Export

Prepared for: v15.4.x

Prepared by: Ataccama

Dated: November 2024

Contents of the Document

Introduction

In this workshop, you will **export** the data of your existing catalog items to various destinations, enabling their use outside the ONE Platform.

Tasks

We have three tasks in this workshop through which we can practice exporting our catalog items' data to a .*csv* file, a database, as well as ONE Data.
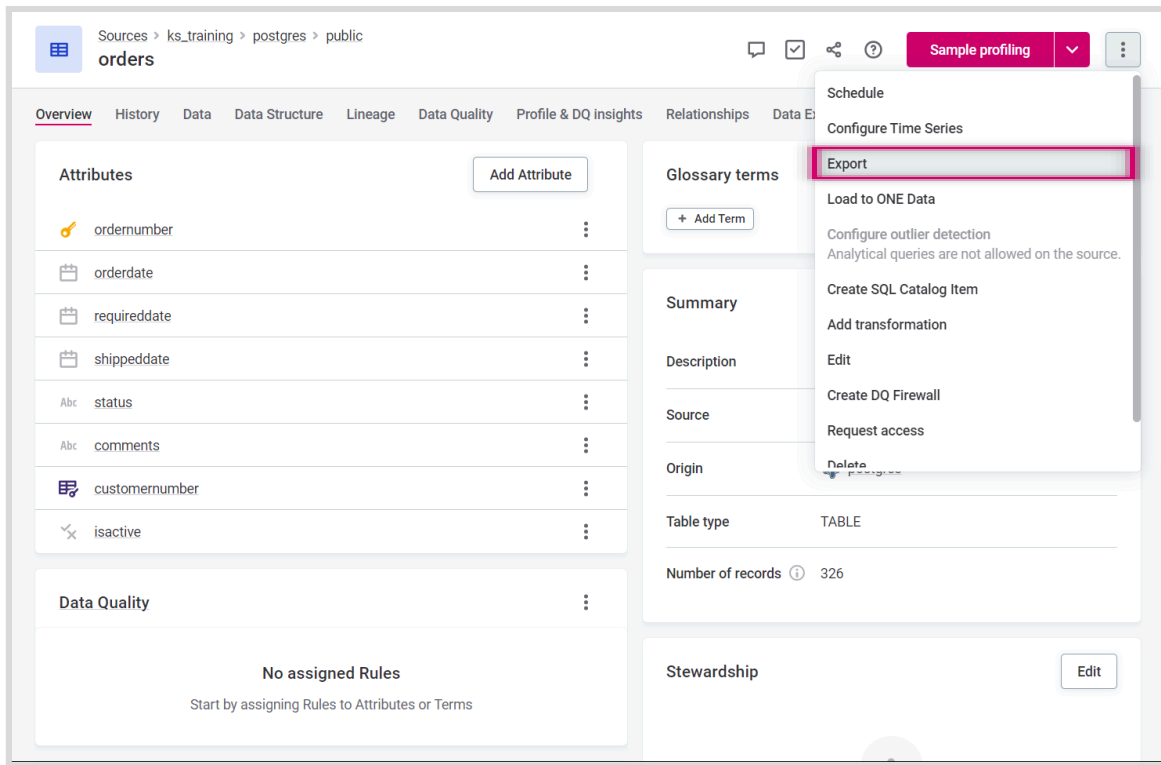
1. Export an existing catalog item to a file

In the first task, we will export an Exiting catalog Item **orders** into a *.csv file.

> › Navigate to the list of catalog items in your source: **Knowledge Catalog** >> **Sources** >> **training** >> **Catalog Items**.
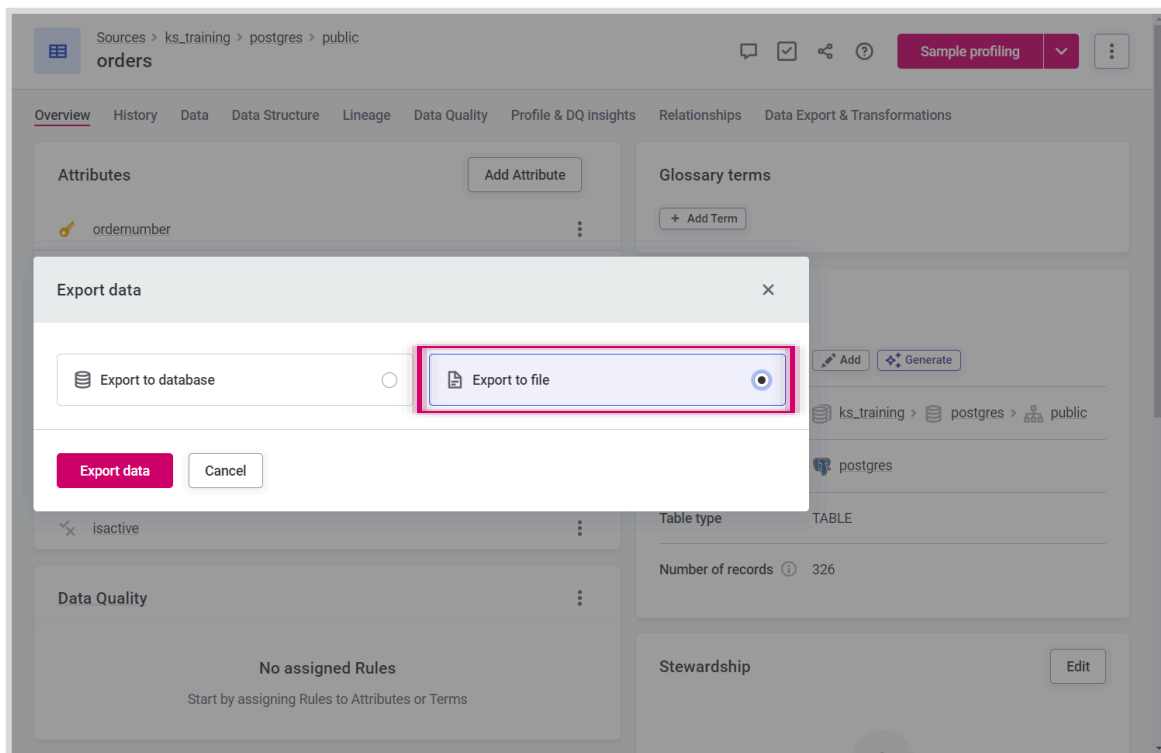

> › Click on the name of the catalog item **orders** to open it.
> › From the options menu (⋮), select **Export**.

› Select **Export to file** to export data to a file and then click on the **Export data** button.

*Alternatively, you can navigate to the **Data export project** tab of your **selected catalog item**, select **Export data**, and follow the procedure described in one of the following scenarios based on where you need the data to be exported.*

› From the list of available attributes, select **ordernumber**, **orderdate**, **status**, and **customernumber** for export from the catalog item, then select **Next**.
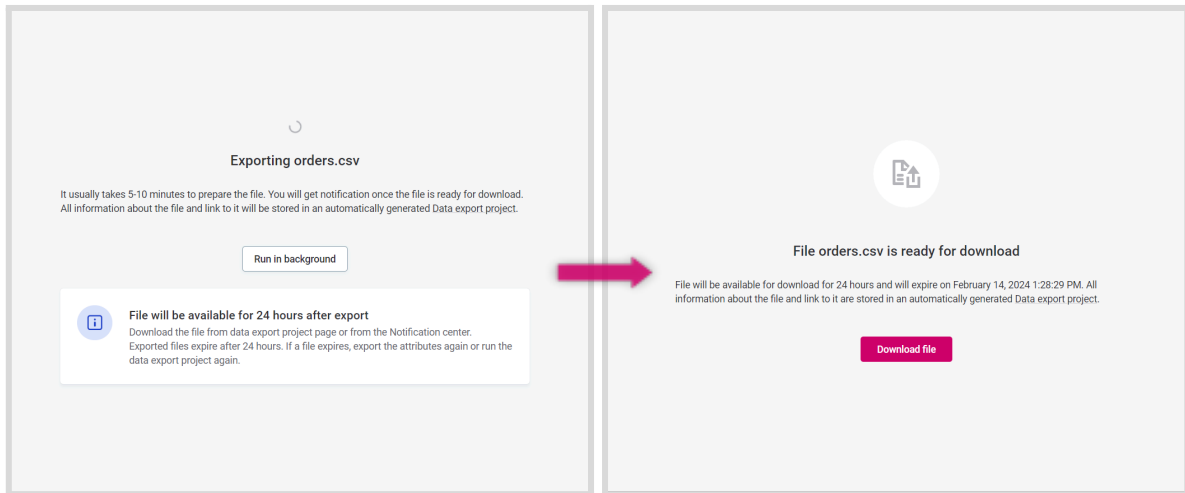

› Provide a file name for your export (e.g., "**orders_info**"), and select the file type **CSV** (Semicolon separated). Next, click the **Export** button.

› You can also monitor the progress of the running exporting process. Once the export is finished, a download button appears. Select **Download** to download your file.

> ❯ Once the items are downloaded successfully you can see them on your local machine.

NOTE

To see all data export projects created in ONE, navigate to **Data Catalog > Catalog Items > Data Export Projects.** To see the data export projects from a corresponding catalog item, navigate to Data Catalog > Catalog Items > [your catalog item] > Data export project tab.

2. Export an existing catalog items to a database

In this task, we will export the catalog item **customers** to a database. To do this, we need a database with write permissions. If you plan to use the training database, you must create a source and connect to it. Alternatively, if the source was created in previous workshops, it must be reconfigured to ensure write permissions are enabled.

> *Training Database specifications:*
>
> **Name:** *postgres*
>
> **JDBC**: *jdbc:postgresql://db-training.czhmef985xre.eu-central-1.rds.amazonaws.com:5432/postgres*
>
> **Credentials name:** *training*
>
> **Username:** *training*
>
> **Password:** *AtaccamaONE*

› Navigate to the **Knowledge Catalog** section - **Sources**, then select the source you created previously.

› On the **Connections** tab, open the options menu (⋮) for your database connection (**postgres**) and select **Show details**.
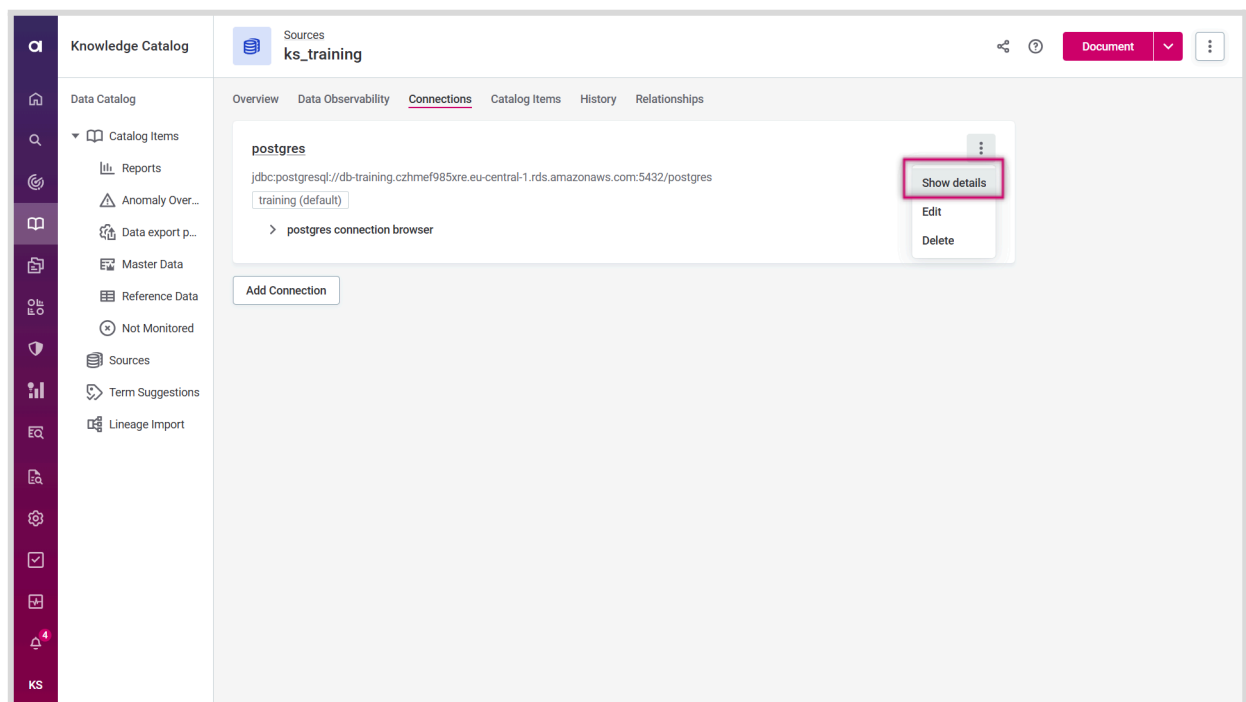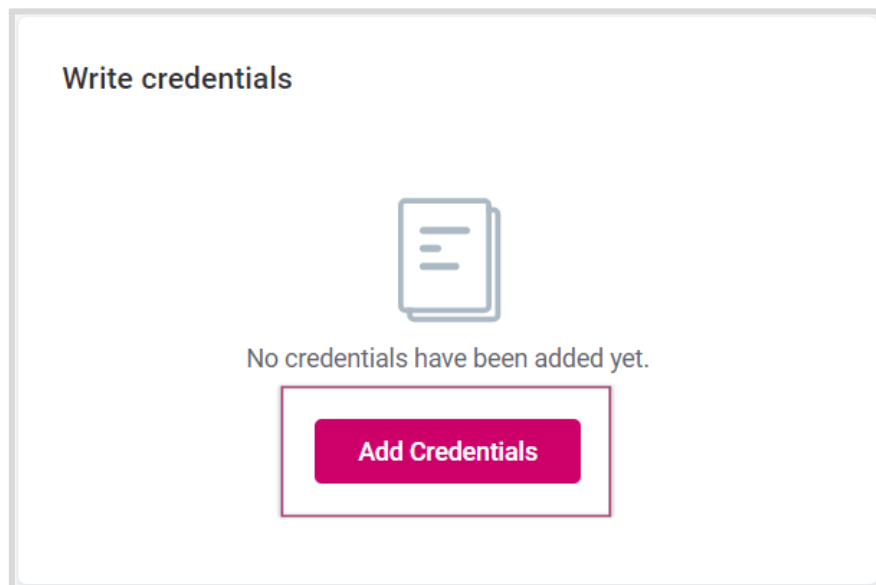

› In **Credentials,** make sure that you have set them as default. Otherwise, use the more options menu to **Set as default.**
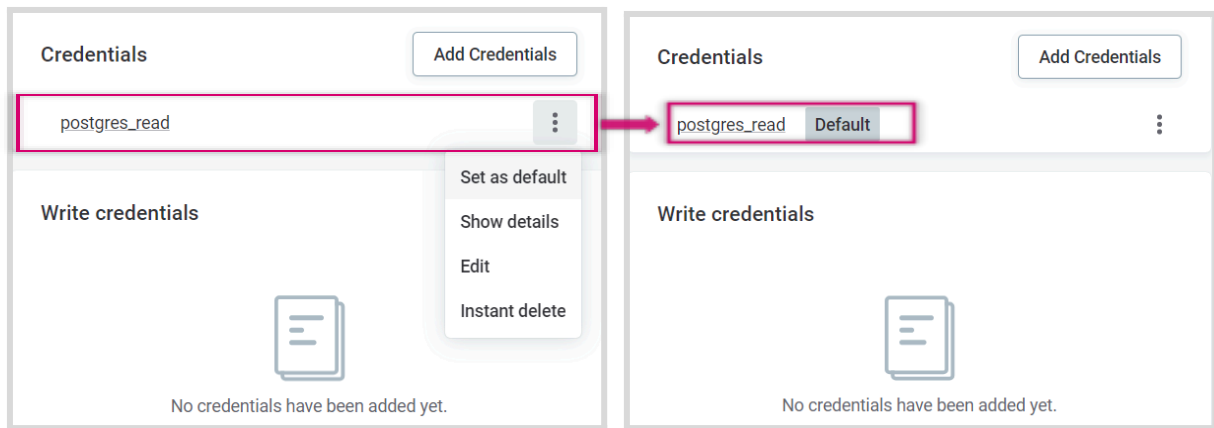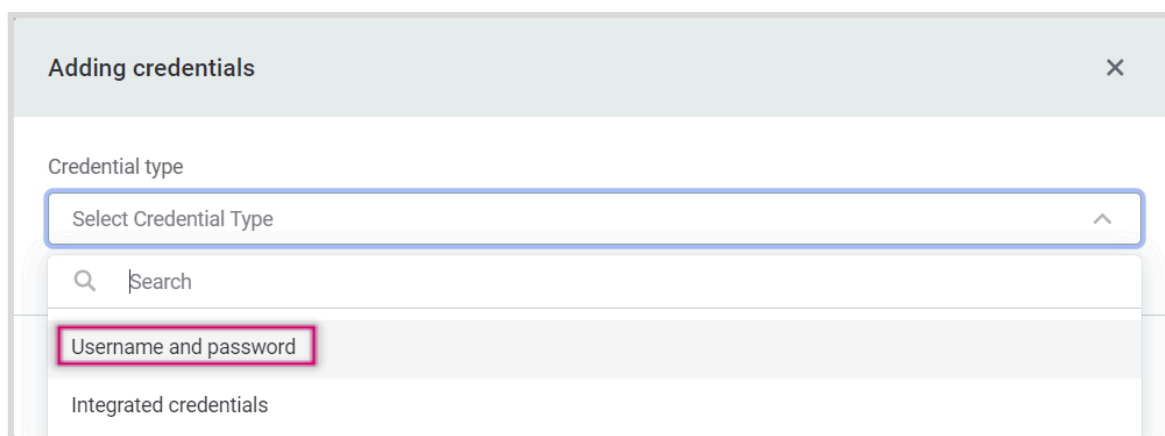
Next, we need to add **Write credentials** for the existing connection:

› Navigate to the **Write credentials** section and click the **Add Credentials** button.

› Select the ***Username and password*** as credential type.

› Use the same credentials as you have for previous credentials for reading data (Username: *training* Password: *AtaccamaONE*). Give it a name, for instance, '**postgres_write**'.

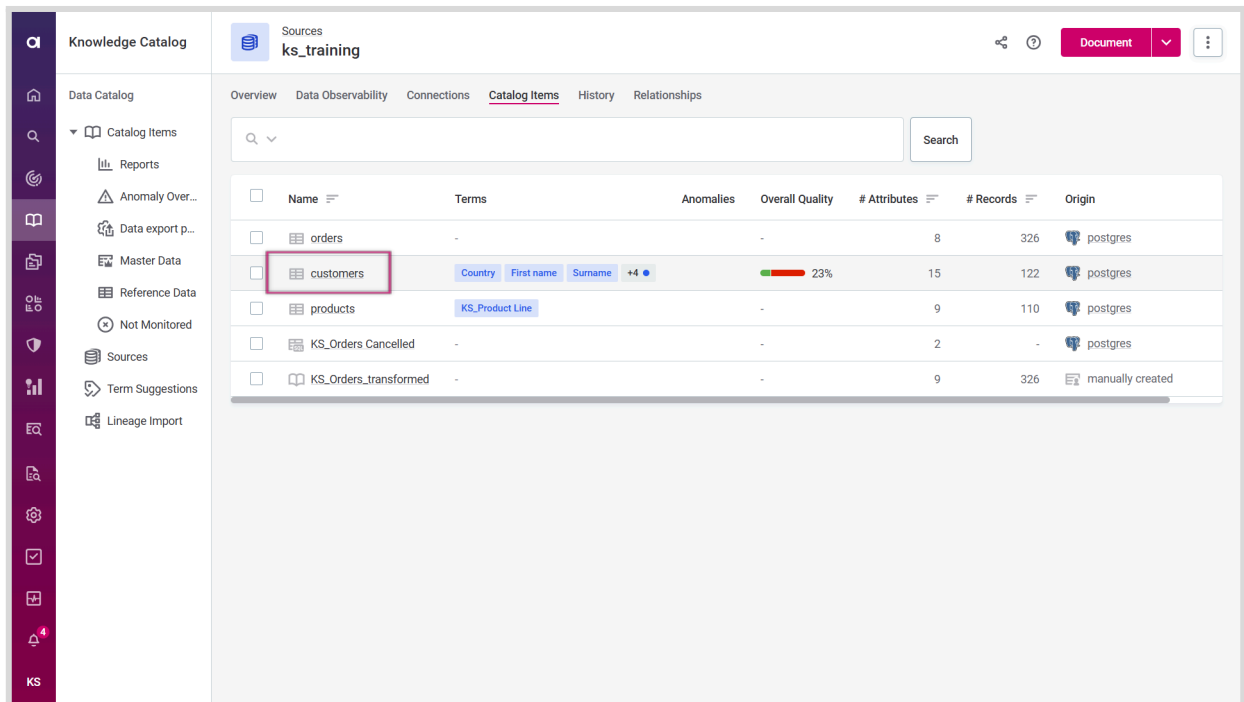› Click **Save and Publish**.

› Set credentials for writing data as **Default**.

› **Publish** all changes.

Once completed, you can begin exporting your data to this database using the newly added credentials.

› Navigate to the list of catalog items in your source: **Knowledge Catalog** >> **Sources** >> **training** >> **Catalog Items**.

› Click on the name of the catalog item **customers** to open it.

› From the more options menu (⋮), select **Export**.

› Select **Export to database** export data to a database and then click on the **Export data** button.
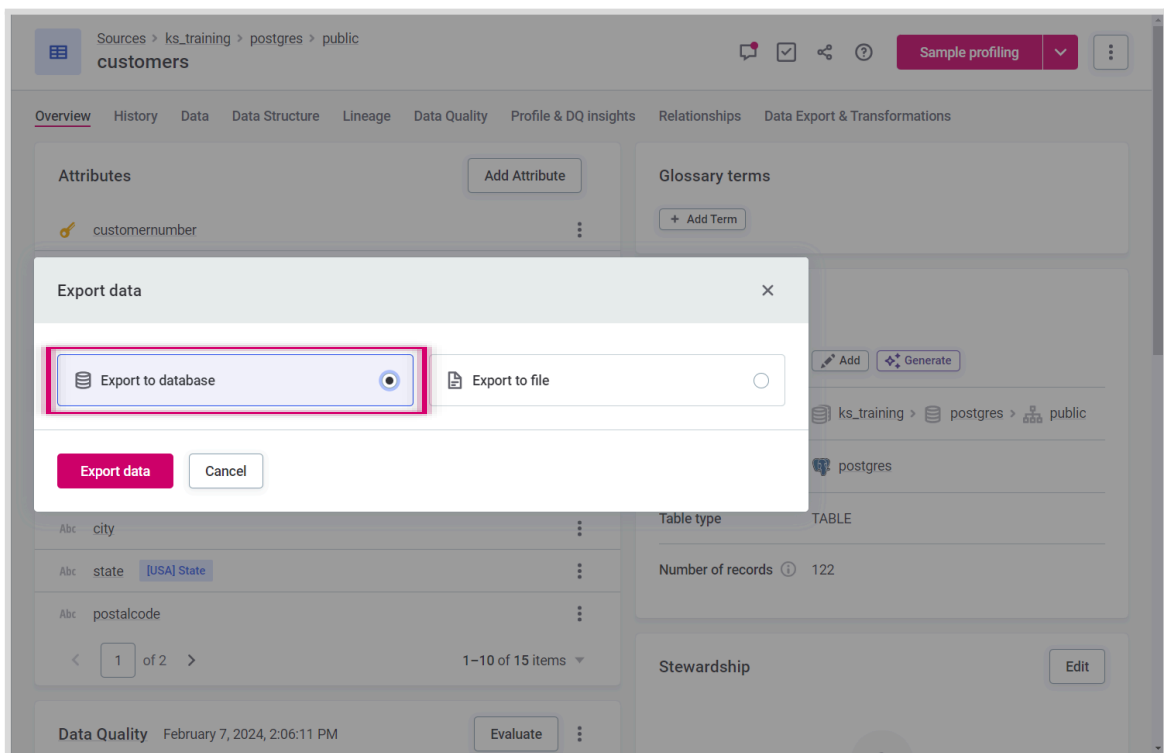
> Provide a **name** for the new export project (e.g., "**<prefix>_customers_export_db**").
> Click **Next** to continue.

> **Add** the following **attributes** to export by ticking their box selectors:

> Select the **Target Source** ("**<prefix>_training**"), then the connection in the data source ("**postgres**").

› Create a new table by selecting the schema **public**, and then clicking on the **Create table** button. Give the name for the new table (e.g., "**<prefix>_customers_phone**").

*You can save the draft of your project at any point of the configuration. To see and edit drafts, navigate to **Data Catalog > Catalog Items > Data Export Projects**.*

› Once the new table is created, it appears at the top of the list. When you are finished editing, click **Next**.

› Click **Save & run project** on the right bottom corner to save and run the Export data project.

> › Once completed, you can find the run results of the export on the **Data Export & Transformations** tab of the catalog item exported.

Now, if accessible, you can also go to the database and check your exported data table.

3. Load an existing catalog items to ONE Data

In this task, we will load the **existing catalog item** products into **ONE Data**. This action creates a table that can be managed directly in ONE Data within the Ataccama ONE Platform. The table can be used as reference data in validity and detection rules. Additionally, it allows you to quickly modify some data before exporting it back to the data source.

› Once more, navigate to the list of catalog items in your source: **Knowledge Catalog** >> **Sources** >> **training** >> **Catalog Items**.
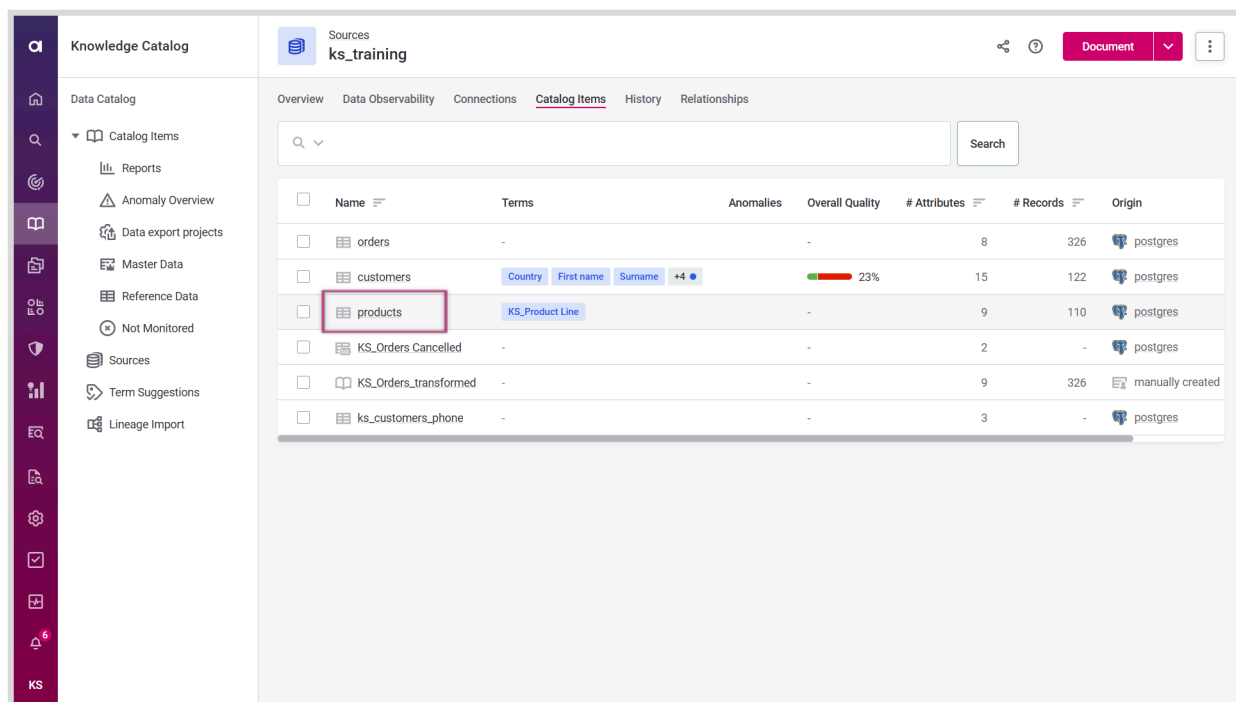
› This time, we will work with the catalog item **products**. Click its name to open it.

› Choose **Load to ONE Data** from the options menu (⋮).
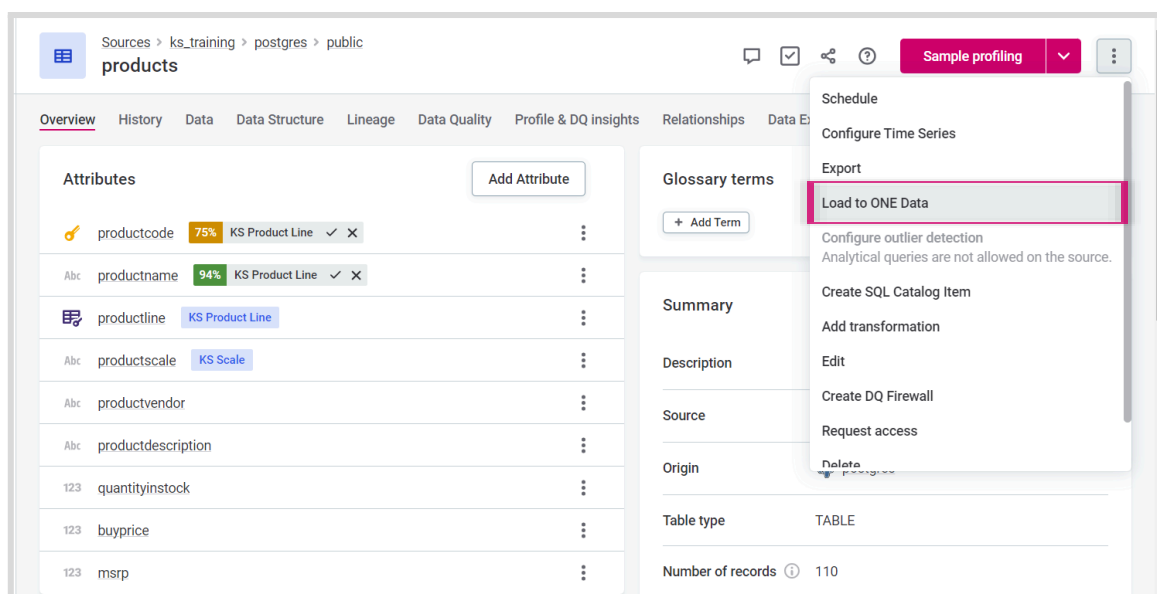
› Select the **Export type** options **Data only**, then click **Next**.

› Select the **attributes** for export, and once you have chosen the attributes, select **Next**.
  - **productcode**
  - **productname**
  - **productvendor**
  - **quantityinstock**

› Enter a unique name(**<prefix>_product_info**) for the table and optionally a description. Select **Create table.**

› Depending on the size of your catalog item, it might take a few minutes to get everything ready. To continue working with the platform in the meantime, select **Run in background**.

› Table **<prefix>_products_info** is ready to use in **ONE Data**.

Conclusion

We have now reached the end of another Ataccama ONE workshop!

In this lab, we learned how to perform the export operations to transfer the catalog item data to various destinations.