



ONE - Workshop

DQ Advanced – Data Export

Prepared for: v15.4.x

Prepared by: Ataccama

Dated: November 2024

Contents of the Document

Introduction	3
Tasks	3
1. Export an existing catalog item to a file	3
2. Export an existing catalog items to a database	7
3. Load an existing catalog items to ONE Data	14
Conclusion	17

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

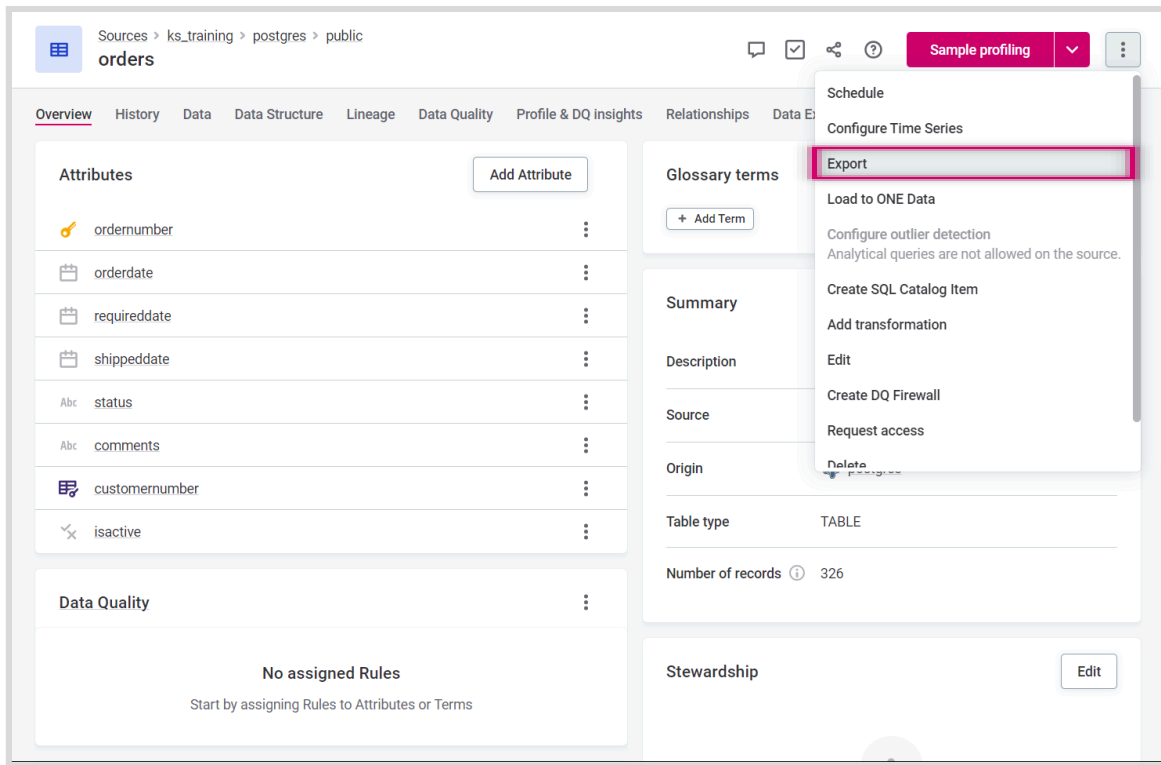
In the first task, we will export an Existing catalog Item **orders** into a *.csv file.

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

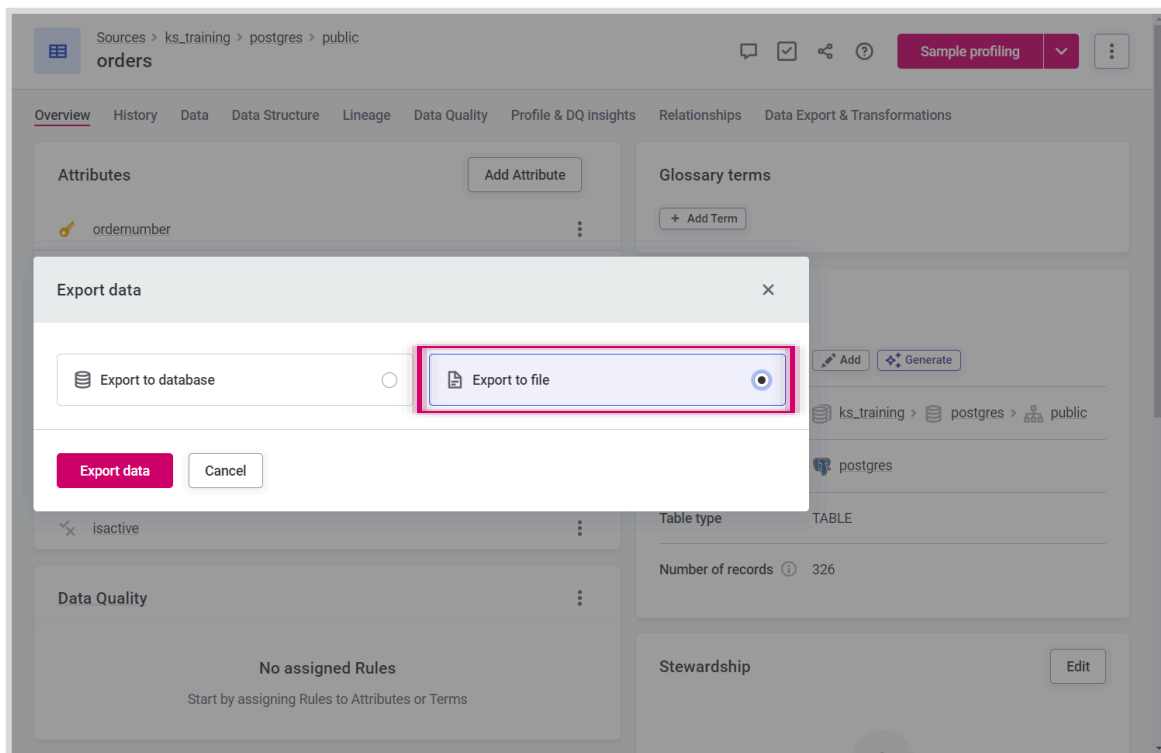
The screenshot shows the 'Knowledge Catalog' interface. On the left is a sidebar with navigation icons. The main area displays the 'Sources' section for 'ks_training'. Under the 'Catalog Items' tab, a table lists several items. The 'orders' item is highlighted with a red box. The table has columns for Name, Terms, Anomalies, Overall Quality, # Attributes, # Records, and Origin.

	Name	Terms	Anomalies	Overall Quality	# Attributes	# Records	Origin
<input type="checkbox"/>	orders	-	-	-	8	326	postgres
<input type="checkbox"/>	customers	Country First name Surname +4	-	23%	15	122	postgres
<input type="checkbox"/>	products	KS_Product Line	-	-	9	110	postgres
<input type="checkbox"/>	KS_Orders Cancelled	-	-	-	2	-	postgres
<input type="checkbox"/>	KS_Orders_transformed	-	-	-	9	326	manually created

- › 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**.

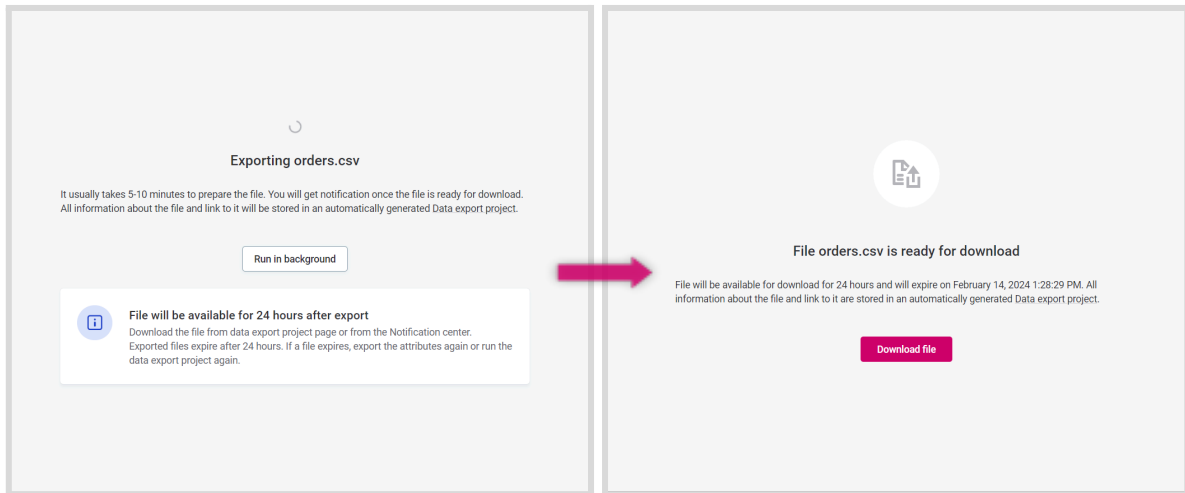
The screenshot shows the 'Export data to file' window with a progress bar at the top indicating '1. Select attributes' and '2. Export file'. A notification box states: 'File size limit is 100 MB. Data will be exported record by record starting from first record. Records that are above file size limit will not be included in to the file.' Below this, the 'Select attributes' section has a sub-header 'Select what data to export. You can export all or marked data from a list.' and three tabs: 'Attributes', 'Data Preview', and 'Profile & DQ Insights'. The 'Attributes' tab is active, showing a search bar and a table of attributes. The table has columns: 'Data type', 'Name', 'Validity', 'Terms', and 'Evaluated'. The attributes listed are: 'ordernumber' (checked), 'orderdate' (checked), 'requireddate' (unchecked), 'shippeddate' (unchecked), 'status' (checked), 'comments' (unchecked), and 'customernumber' (checked). A 'Next' button is located at the bottom right of the window.

	Data type	Name	Validity	Terms	Evaluated
<input checked="" type="checkbox"/>		ordernumber	-		-
<input checked="" type="checkbox"/>		orderdate	-		-
<input type="checkbox"/>		requireddate	-		-
<input type="checkbox"/>		shippeddate	-		-
<input checked="" type="checkbox"/>	Abc	status	-		-
<input type="checkbox"/>	Abc	comments	-		-
<input checked="" type="checkbox"/>		customernumber	-		-

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

The screenshot shows the 'Name your file' section with a 'File name' input field containing 'orders_info' and a character count '11/200'. Below this is the 'Select file type' section with the sub-header 'Export current dataset into delimited text file'. There are three radio button options: 'CSV (Comma separated)', 'CSV (Semicolon separated)' (which is selected), and 'TSV (Tab separated)'. There is also a 'TXT (Tab separated)' option below the others.

- 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.

AutoSave Off orders_info

File Home Insert Draw Page Layout Formulas Data Review View Automate Help

POSSIBLE DATA LOSS Some features might be lost if you save this workbook in the comma-delimited (.csv) format. To preserve these features, save it in an Excel file format.

C10 Shipped

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	ordernumber	orderdate	status	customernumber										
2	10100	1/6/2003	Shipped	363										
3	10101	1/9/2003	Shipped	128										
4	10102	1/10/2003	Shipped	181										



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

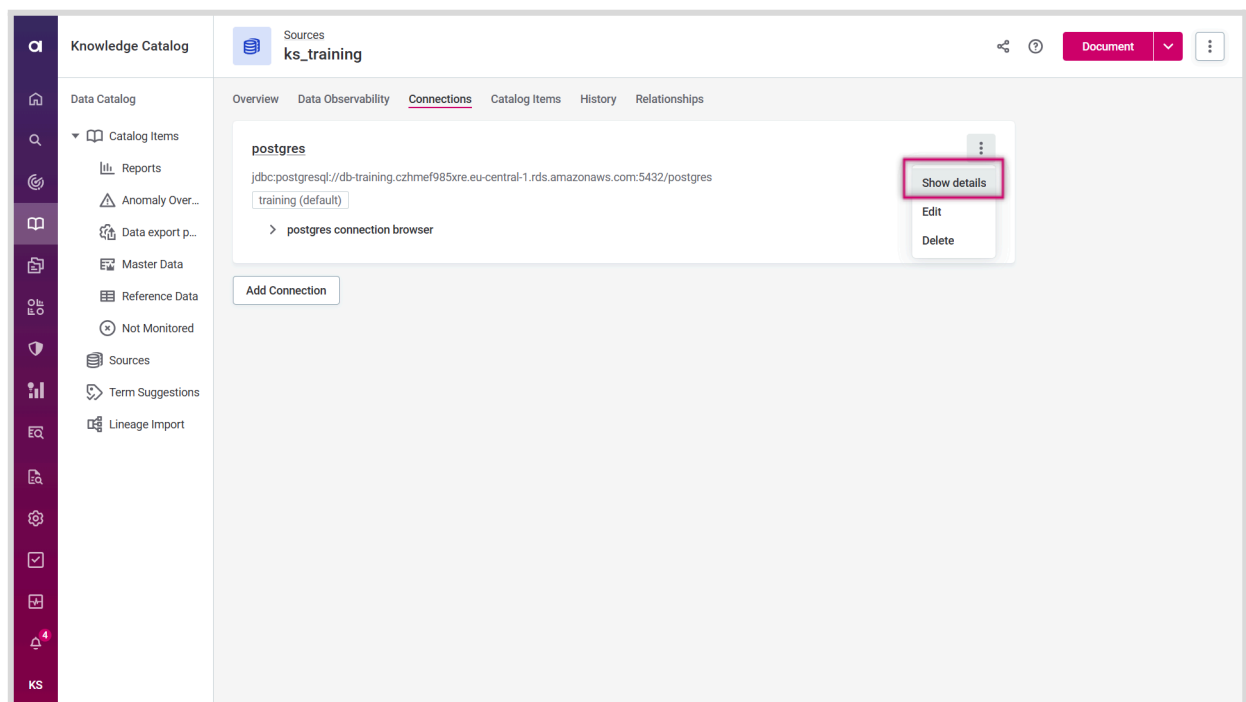
JDBC: <jdbc:postgresql://db-training.czhmef985xre.eu-central-1.rds.amazonaws.com:5432/postgres>

Credentials name: training

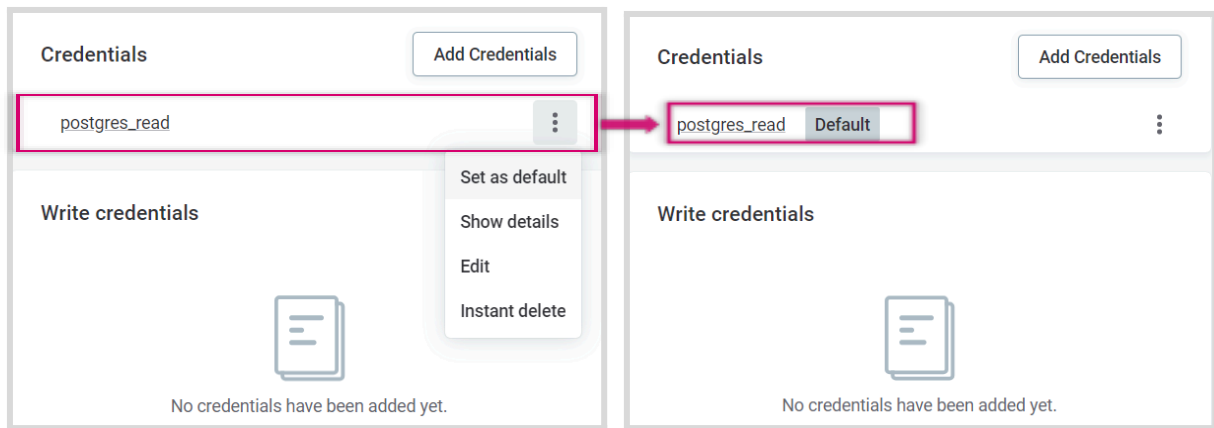
Username: training

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**.

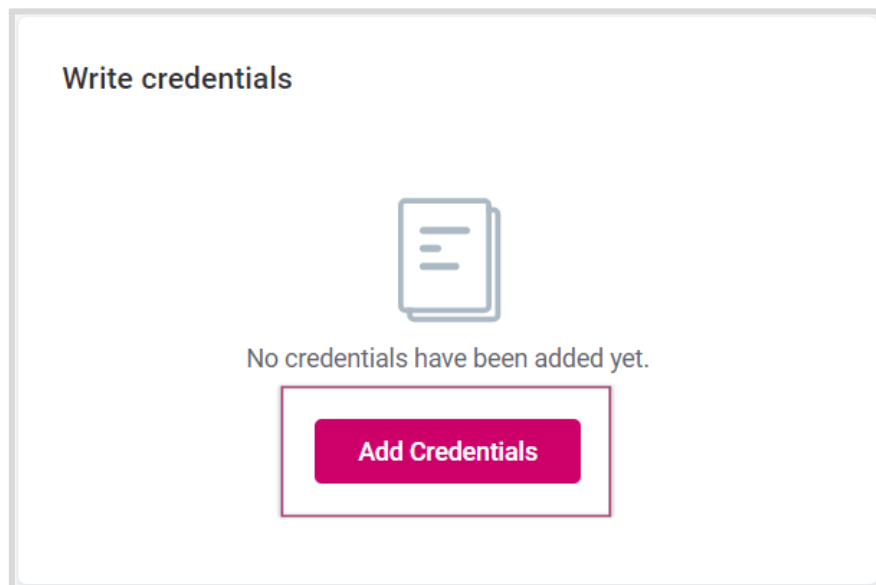


- › In **Credentials**, make sure that you have set them as default. Otherwise, use the more options menu to **Set as default**.

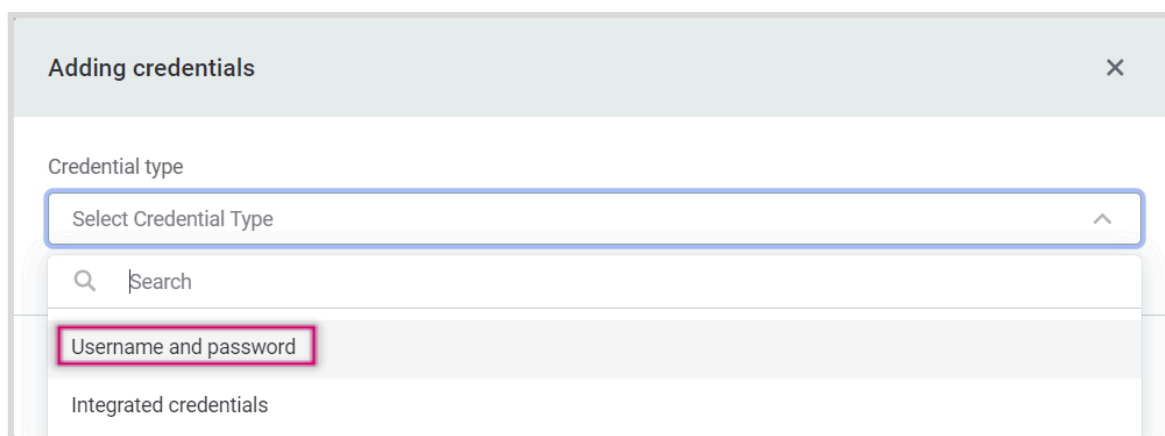


Next, we need to add **Write credentials** for the existing connection:

- › 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**.

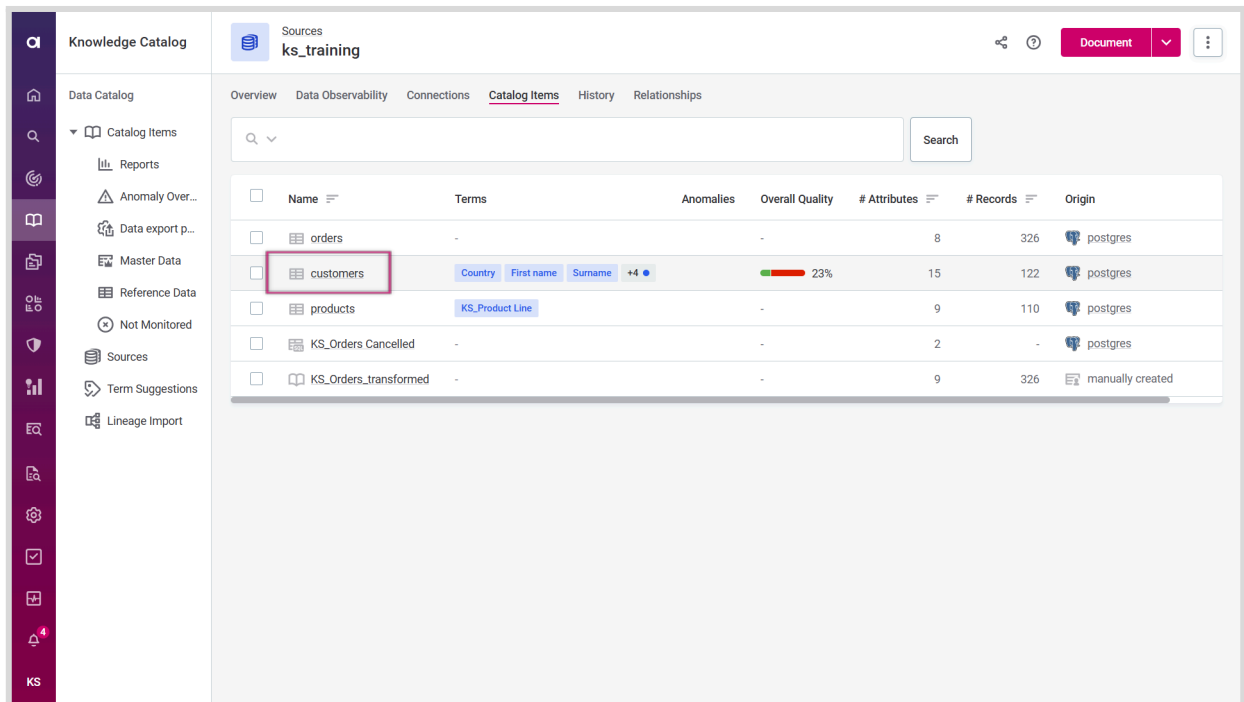
The image shows two sections of a configuration interface. The top section is titled 'Credentials' and contains a table with one row: 'postgres_read' and 'Default'. To the right of this row is a vertical ellipsis menu. Above the table is an 'Add Credentials' button. The bottom section is titled 'Write credentials' and contains a table with one row: 'postgres_write' and 'Default'. To the right of this row is a vertical ellipsis menu. Above the table is an 'Add Credentials' button.

- › **Publish** all changes.

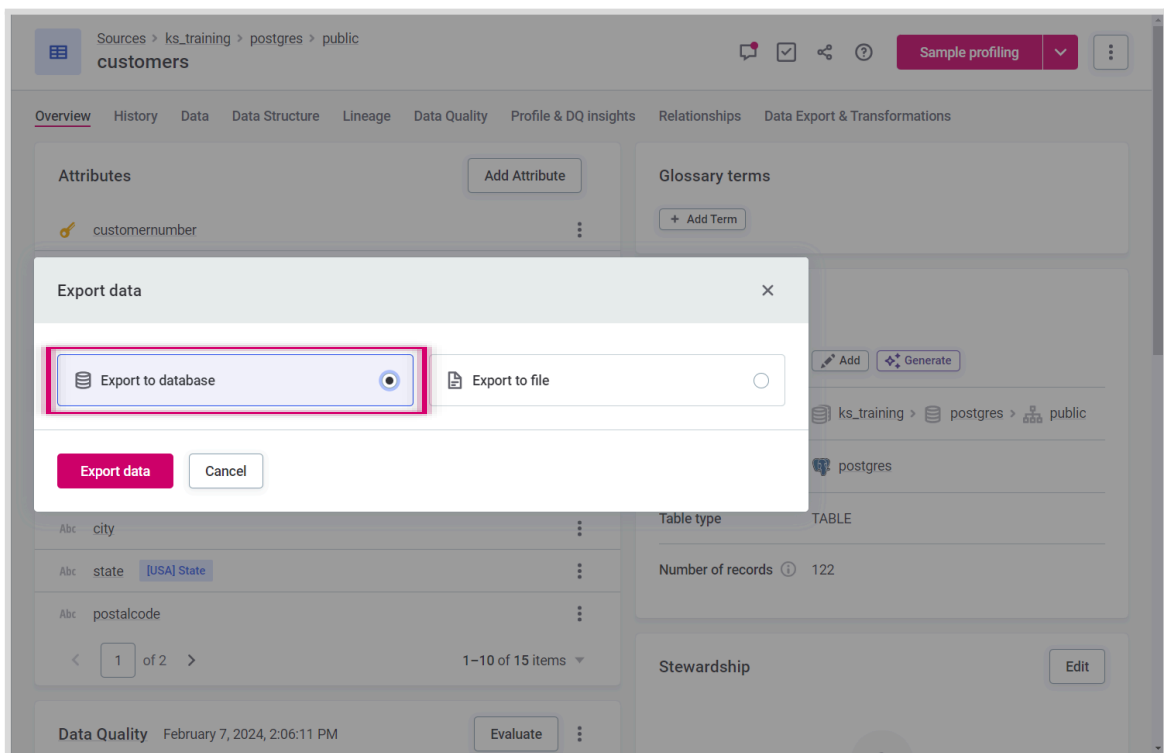
The image shows the top navigation bar of a configuration interface. It features a yellow bar with the text 'Edited' and a help icon. To the right are four buttons: 'View published version', 'View changes', 'Revert changes', and 'Publish' (which is highlighted with a red box). Below this is a white bar showing the breadcrumb 'Sources > ks_training' and the source name 'postgres'. To the right of the source name are three buttons: 'Test Connection', 'Document', and a dropdown menu with a vertical ellipsis icon.

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.





To view your data before selecting attributes, go to **Data preview tab**. In case you select **an existing table**, the entire content of this table will be replaced by the exported data. To see the results of the data quality evaluation, go to **Profile & DQ insights tab**.

- Provide a **name** for the new export project (e.g., “<prefix>_customers_export_db”). Click **Next** to continue.

The screenshot shows the 'New project' step of the 'Export data' wizard. The 'Name' field is highlighted with a red box and contains the text 'KS_customers_export_db'. The 'Description' field is empty. A 'Next' button is visible at the bottom right.

Step
1. New project
2. Select source
3. Select target
4. Overview

New project

Data export project will store your configuration. You can edit it later and run it again to refresh your exported data. All projects are listed inside Catalog item in Data Export Projects tab.

Name *

KS_customers_export_db

Description

Enter a brief description

Next

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

The screenshot shows the 'Select source' step of the 'Export data' wizard. The 'Attributes' tab is selected, and a table of attributes is displayed. The 'phone' attribute is highlighted with a red box, and its checkbox is checked. A 'Next' button is visible at the bottom right.

Attribute	Validity	Terms	Evaluated
customernumber	-		February 7, 2024, 2:06:11 PM
customername	-		February 7, 2024, 2:06:11 PM
contactlastname	84%	Surname	February 7, 2024, 2:06:11 PM
contactfirstname	91%	First name	February 7, 2024, 2:06:11 PM
phone	32%	[North America] Phone Number	February 7, 2024, 2:06:11 PM

Select source *

Click on the checkbox next to attribute(s) which you want to export data from. You will move data from the existing tables in the source database to your database.

Record limit

Set maximum number of records that will be exported per project.

Limit is not set

Attributes Data Preview Profile & DQ Insights

Standard view Hidden columns Clear sort

Back to New project **Next**

- 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”).

The screenshot shows the 'Export data' dialog with the '3. Select target' step active. A 'Create table' modal is open, allowing the user to enter a table name. The modal includes a 'Table name' input field with the text 'ks_customers_phone', a 'Create' button, and a 'Cancel' button. A message states: 'Table will be created after you run the data export project.' The background dialog shows the '1. Select target' step with 'ks_training' selected, and the '2. Select connection' step with 'postgres' selected. A 'Save draft' button is visible in the top right of the dialog.



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**.

The screenshot shows the 'Export data' dialog with the '3. Specify location' step active. The 'postgres' connection is selected. A message states: 'Data will be replaced. Data in the table will be replaced by new exported data.' The breadcrumb shows 'You are in: ks_training > postgres'. A search bar is present. Below, a table lists available tables:

Name	Type
ks_customers_phone	TABLE
rm_training	TABLE
rm_training_sample	TABLE

The 'ks_customers_phone' row is selected. A 'Create table' button is in the top right. At the bottom right, a 'Next' button is highlighted with a red box. A 'Back to Select source' button is at the bottom left.

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

Export data

1. New project — 2. Select source — 3. Select target — 4. Overview

×

Overview

Save draft

Review the information provided here to help you complete data export project. First you need to save the project and then you can run it to export data.

General information

Project name

KS_customers_export_db

Description

Add a description

State

Ready to run

Source catalog item

customers

Target catalog item

Catalog item will be created after you successfully run data export project.

Data source

ks_training

Connection

postgres

Location

public > ks_customers_phone

Row limit

Unlimited

Attributes

Name

customernumber

customername

phone

Back to Select target

Save & run project

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

Knowledge Catalog

Data Catalog

Catalog Items

Sources

Term Suggestions

Lineage Import

Sources > ks_training > postgres > public

customers

Sample profiling

⋮

Overview

History

Data

Data Structure

Lineage

Data Quality

Profile & DQ insights

Relationships

Data Export & Transformations

Data Export

Standard view

Hidden columns

Name	Last run result	Target	Available for download	Last run	Project state
KS_customers_export_db	Completed	–	–	June 3, 2024, 9:56:46 AM	Ready to run

Transformation plans

Standard view

Hidden columns

Clear sort

Create transformation plan

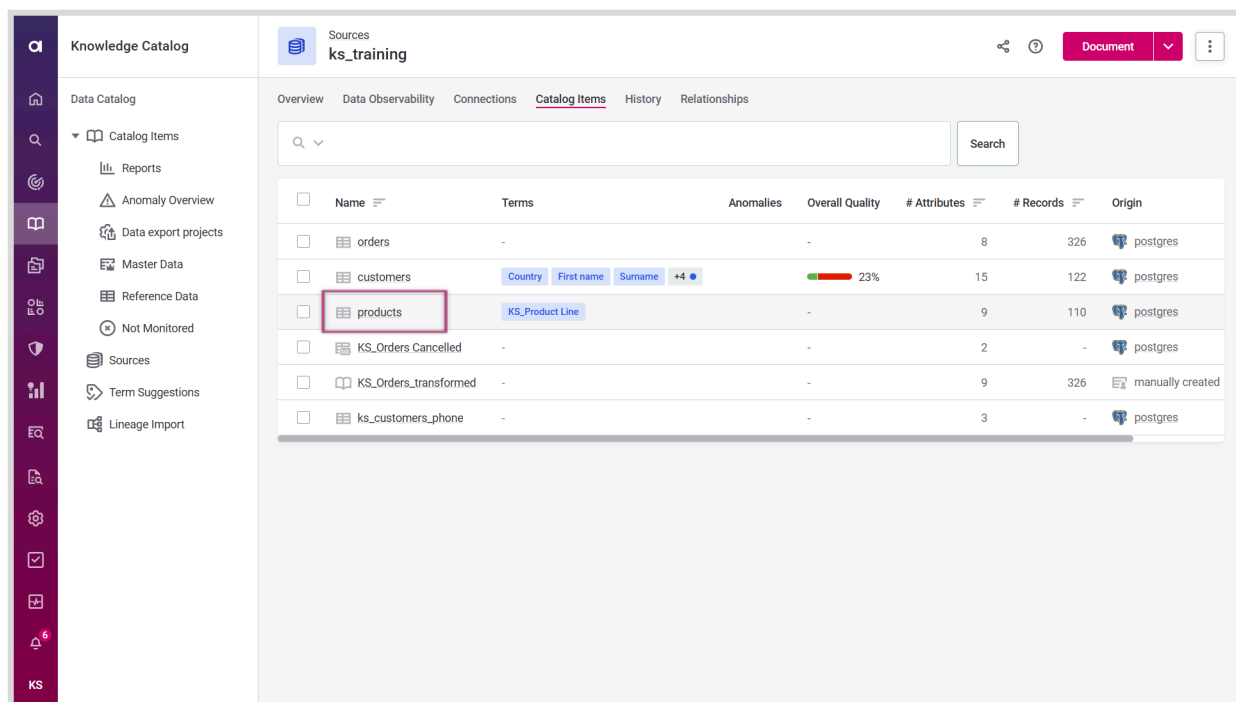
Name	Description	Status	Input type	Output type
KS_customers_transformation_plan	Transforming a Catalog Item with Data Transformation...	No issues	Catalog item input	ONE Data writer
KS_Transformation_Plan_Customer_DQ_Report		No issues	Monitoring project post processing input	File output

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

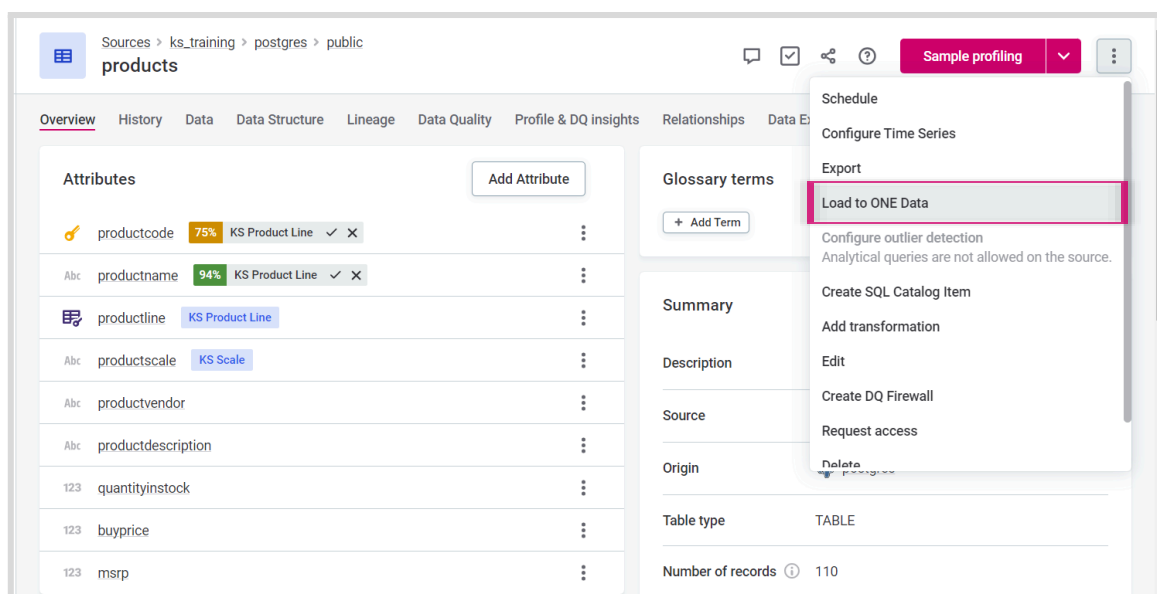
3. Load an existing catalog items to ONE Data

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**.

Load to ONE Data 1. Select data load type 2. Select attributes 3. Table options

Data load type

Full (data with DQ results) ☐ **Data only** ☒ Deduplicated data ☐

Next

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

Load to ONE Data 1. Select data load type 2. Select attributes 3. Table options

Select attributes

Attributes Data Preview Profile & DQ Insights 4 attributes selected X Clear selection

Standard view Hidden columns Clear sort

	Data type	Name	Validity	Terms	Evaluated
<input checked="" type="checkbox"/>	🔑	productcode	-		February 7, 2024, 1:59:54 PM
<input checked="" type="checkbox"/>	Abc	productname	-		February 7, 2024, 1:59:54 PM
<input type="checkbox"/>	📊	productline	100%	KS Product Line	February 7, 2024, 2:00:40 PM
<input type="checkbox"/>	Abc	productscale	100%	KS Scale	February 7, 2024, 1:59:54 PM
<input checked="" type="checkbox"/>	Abc	productvendor	-		February 7, 2024, 1:59:54 PM
<input type="checkbox"/>	Abc	productdescription	-		February 7, 2024, 1:59:54 PM
<input checked="" type="checkbox"/>	123	quantityinstock	-		February 7, 2024, 1:59:54 PM

< Back to selecting data load type Next

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

Load to ONE Data 1. Select data load type 2. Select attributes 3. Table options

Table options

Create new ONE Data table ☒ Overwrite existing ONE Data table ☐

Name
ks_product_info

Description
B I </> :≡ :≡ H1 H2 ⌕ /

Stewardship
Owner
Select stewardship group

< Back to selecting attributes **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**.

ONE Data Tables
ks_product_info

Data Overview History Lineage Data Quality Profile & DQ insights Relationships Data Export & Transformations

Filter Evaluate Data Quality

	productcode	productname	productvendor	quantityinstock	dmm_record_id
1	S10_1678	1969 Harley Davidson UL...	Min Lin Diecast	7933	c32514d4-1681-4d7b-b985...
2	S10_1949	1952 Alpine Renault 1300	Classic Metal Creations	7305	7cbb6cf1-4550-4e96-b461...
3	S10_2016	1996 Moto Guzzi 1100i	Highway 66 Mini Classics	6625	09868d37-1e4d-47f0-8c6f...
4	S10_4698	2003 Harley-Davidson Ea...	Red Start Diecast	5582	a9f34618-19a6-4fdd-8325...
5	S10_4757	1972 Alfa Romeo GTA	Motor City Art Classics	3252	80234dcd-ca59-4516-b473...
6	S10_4962	1962 LanciaA Delta 16V	Second Gear Diecast	6791	da45b8bf-638b-4129-a1d8...
7	S12_1099	1968 Ford Mustang	Autoart Studio Design	68	4986eae7-4d65-46d9-93ec...
8	S12_1108	2001 Ferrari Enzo	Second Gear Diecast	3619	e71c8f5a-5f28-4a02-9f19...
9	S12_1666	1958 Setra Bus	Welly Diecast Productio...	1579	ec6035be-3a8a-4677-a88e...
10	S12_2823	2002 Suzuki XRE0	Unimax Art Galleries	9997	028c8f45-7b9c-42b4-919d...
11	S12_3148	1969 Corvair Monza	Welly Diecast Productio...	6906	3128dbc5-c0be-40c2-bf5d...
12	S12_3380	1968 Dodge Charger	Welly Diecast Productio...	9123	8b200651-8f3f-42b5-a9ac...
13	S12_3891	1969 Ford Falcon	Second Gear Diecast	1049	f0502f18-f295-4775-98f4...
14	S12_3990	1970 Plymouth Hemi Cuda	Studio M Art Models	5663	89c269ff-e1aa-47a8-a281...
15	S12_4473	1957 Chevy Pickup	Exoto Designs	6125	6bd233d7-57c5-44e3-85d8...
16	S12_4675	1969 Dodge Charger	Welly Diecast Productio...	7323	7eh4806f-4f8c-41f5-a1eh...

Total records: 110

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.