



# Adobe Experience Platform Bootcamp Deep Dive Edition

DOCUMENTATION AND LABS



# Contents

Lab Overview	4
Learning Objectives	4
Lab Task – Creation of the Dataflow	5
1. Select source data	5
2. Define the target dataset	7
3. Data Prep / Transformation	8
4. Schedule	10
5. Check scheduled execution	13
Lab Overview	17
Learning Objectives	17
Lab Tasks – Ingest Customer Account data	18
6. Select Source data	18
7. Define the Target Dataset	21
8. Data Prep / Transformation	22
9. Preview the mapping output	23
10. Re-map the attributes	25
11. Calculated Fields	26
12. Schedule	32
13. Check execution status and error messages	34
14. Correct the Dataflows	41
15. Re-schedule the failed Dataflow	48
Lab Overview	56
Learning Objectives	56
Lab Tasks – Orders – Historical data	56
3. Select Source data	57
4. Define the Target dataset	58
5. Data Prep / Transformation	59
6. Remove the following mappings	59
7. Create Calculated Fields	60
8. Add new pass-through mappings	63
9. Product List Items	63
10. Final Mapping Set	64
11. Preview the data	66
12. Schedule	68
13. Check scheduled execution	69
14. Verify the data	71

Lab Overview	74
Lab Tasks – Customer Accounts – Live data (Streaming)	75
2. Creating a Streaming Inlet	76
3. Upload sample data	78
4. Import the Customer Accounts Mapping	81
5. Streaming Endpoint	85
6. Place a REST API call	87
7. Verify the data in Profile	90
8. Learnings	91
9. Additional exercises	91
Lab Tasks – Orders – Live data (Streaming)	94
10. Create a Streaming Inlet	95
11. Upload sample data	97
12. Define the Target dataset	98
13. Data Prep / Transformation	99
14. Ingest Live Data	105
15. Place a REST API call	107
16. Verify data in Profile	112
17. Learnings	116
18. Additional exercises	116



# **Adobe Experience Platform - Data Pipeline Stores**

Adobe Experience Platform Bootcamp Deep Dive Edition

Name \_\_\_\_\_  
Sandbox \_\_\_\_\_

## Lab Overview

Ingest Stores data into AEP in batch mode using Data Landing Zone. This Lab will leverage delimited data on Landing Zone and ML Recommendations to map most of the source data. This job will be scheduled to run every 30 days.

This lab will introduce you to the basic data ingestion process.

**Expected time: 20 minutes**

## Learning Objectives

What should you walk away with after taking this Lab?

- Understand the basic data ingestion process
- Use Data Landing Zone as a source
- Use Data Prep to Map the non-XDM data to XDM
- Scheduling batch workflows
- Preview the data in dataset

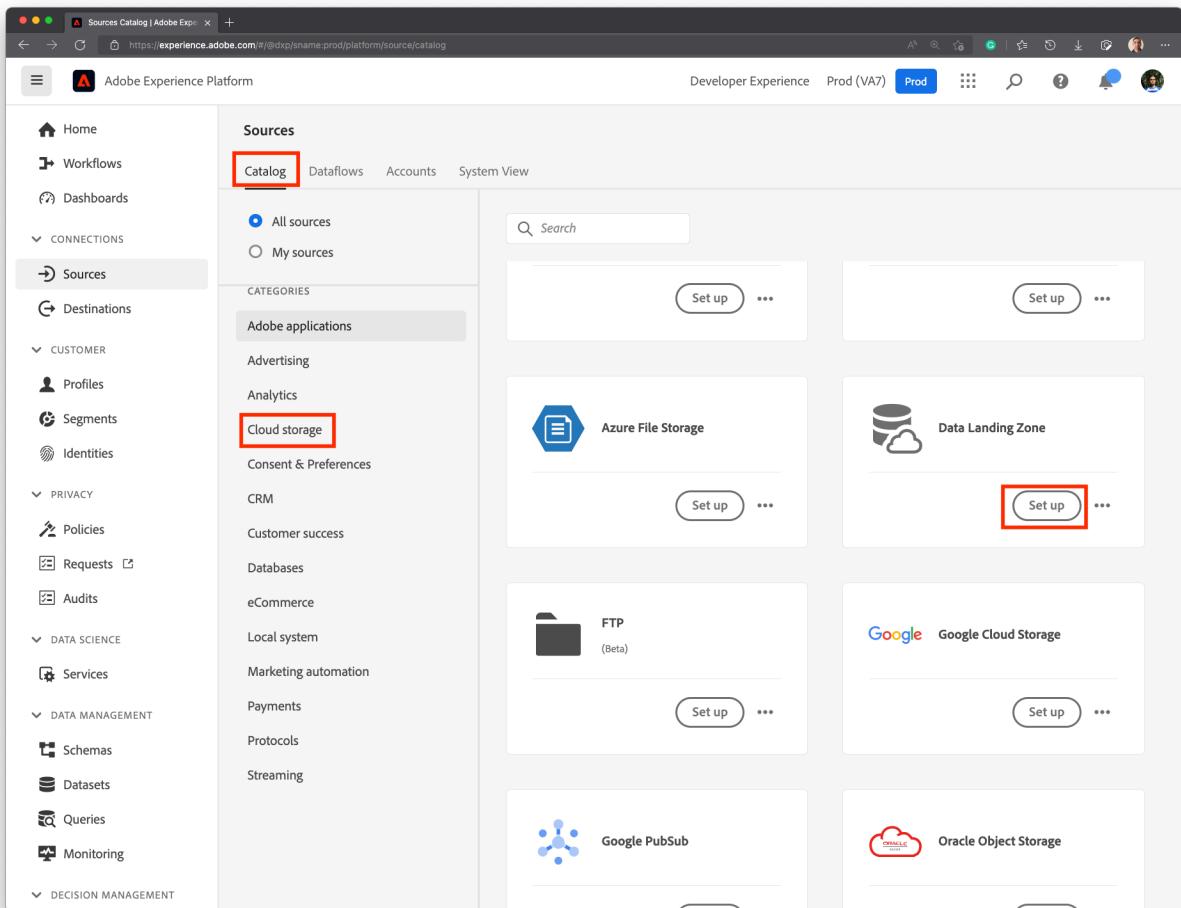
## Lab Task – Creation of the Dataflow

### Select source data

Go to Adobe Experience Platform → Sources → Catalog. In the Cloud Storage connectors, click Setup / Add Data in the Data Landing Zone card.

#### Tip

If at least one connection exists for that source, you will see “Add data” as the default action. If no connections exist for that source, you will see “Setup” as the default action.



In the next screen, navigate to **project → PIPELINE** and select the **Lab\_Lookup\_Store.csv**. On the right hand side, Data format should already be set to **Delimited**. Preview should load automatically. Click **Next**.

**Tip**

Check the list of delimiters and compression types supported by clicking on the corresponding drop downs

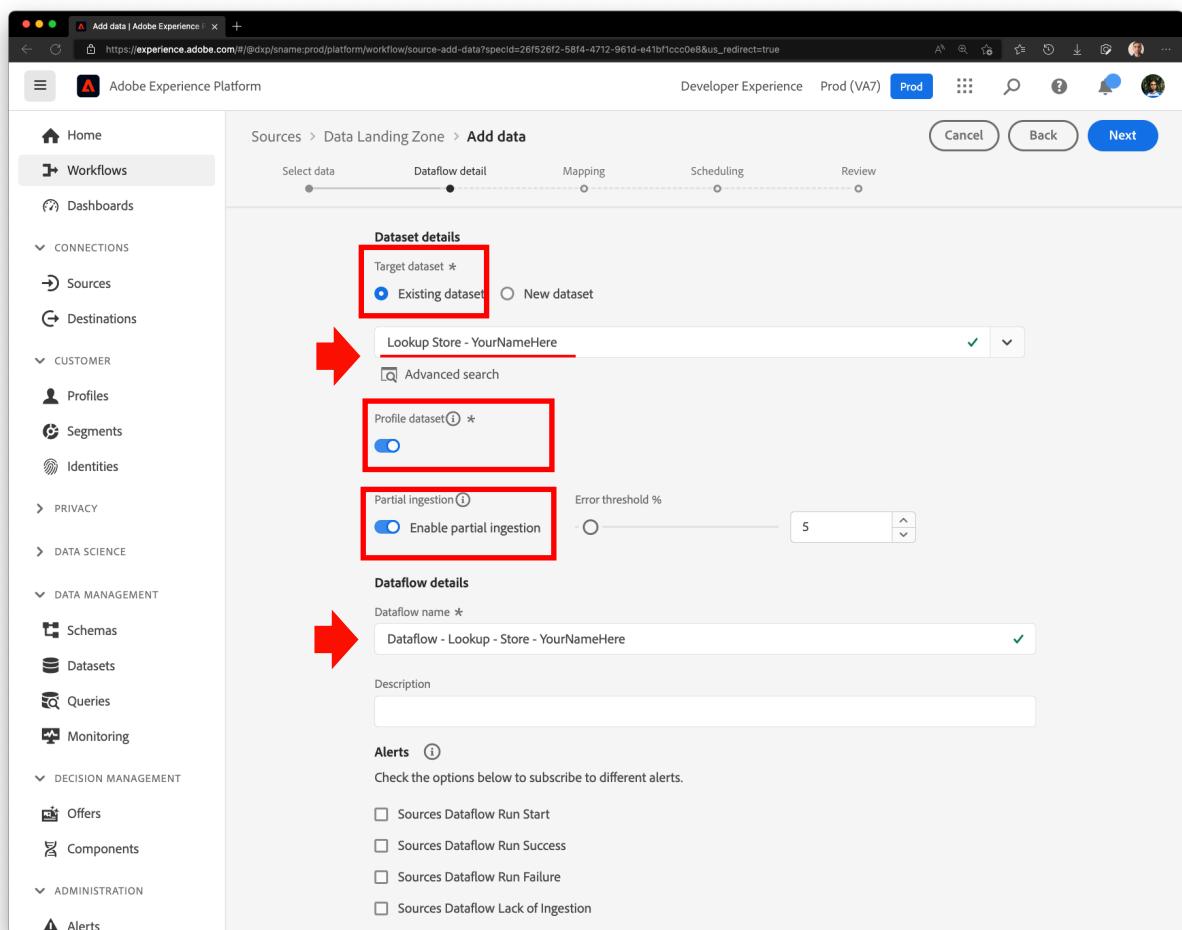
The screenshot shows the 'Add data' interface in Adobe Experience Platform. The left sidebar navigation includes Home, Workflows (selected), Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY (Policies, Requests, Audits), DATA SCIENCE (Services), and DATA MANAGEMENT. The main workspace shows 'Sources > Data Landing Zone > Add data'. The 'Selected data' panel lists 'Lab\_Lookup\_Store.csv'. The 'Preview' panel shows the full path: '/ > dlz-user-container > project > PIPELINE > Lab\_Lookup\_Store.csv'. The 'Data format' dropdown is set to 'Delimited' with a comma delimiter. A red box highlights the 'PIPELINE' path in the preview. Another red box highlights the 'Lab\_Lookup\_Store.csv' file in the 'Selected data' list. The 'Sample data' table below shows five rows of store information.

	STOREID	STORENAME	STOREINAGURATIONDATE	STOREOPENTIME
0	STORE-1	Cool Place	2021-01	open at 9am / close at 9pm
1	STORE-2	Awesome Town	2021-01	open at 9am / close at 9pm
2	STORE-3	Awesome Town	2022-04	open at 9am / close at 9pm
3	STORE-4	Best Shop	2022-01	open at 9am / close at 9pm
4	STORE-5	Cool Place	2022-01	open at 9am / close at 9pm

## Define the target dataset

In the **Dataflow detail** screen, choose the **Existing dataset** **Lookup Store - YourNameHere**. **Error Diagnostics** is already turned ON. Now, turn ON **Enable Partial Ingestion**. Set the Data flow name as **Dataflow – Lookup – Store - YourNameHere** and click **Next**.

**Tip** When Partial Ingestion is enabled, Error diagnostics are automatically enabled and hence the toggle box will disappear.



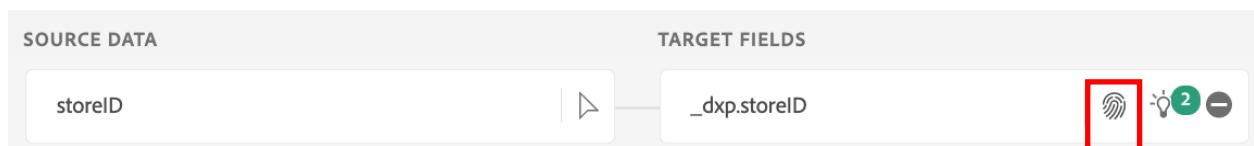
## Data Prep / Transformation

Data Prep (Mapping) step will load. The ML Recommendation Service will pre-populate the mappings for you. Please review and ensure the mappings are as follows. If all mappings look good, click **Next**.

The screenshot shows the 'Add data' workflow in the Adobe Experience Platform. The current step is 'Mapping'. The interface displays a mapping table between 'SOURCE DATA' and 'TARGET FIELDS'. The source is 'Source: Lookup\_Store\_CLIENT.csv' and the target schema is 'Mapping schema: dep: Lookup Store'. The table lists fields like storeID, storeName, storeInagurationDate, etc., with their corresponding target fields like \_dpx.storeID, \_dpx.name, \_dpx.inagurationDate, etc. A red error icon in the top right corner indicates 0 errors.

## Learnings

Identity attributes are represented with a fingerprint icon next to them



If the bulb icon next to the XDM attribute has a number greater than 1, it means there are other recommendations by the system that you can browse through and select.

The screenshot shows the 'Mapping recommendations' section of the Data Modeler. On the left, under 'SOURCE DATA', there are four fields: 'storeID', 'storeName', 'storeInagurationDate', and 'storeOpenTime'. Each field has a right-pointing arrow indicating it can be mapped. On the right, under 'SOURCE FIELD', there is a list: 'storeID', 'SELECT ALL FIELDS', '\_dpx.storeID' (with a checked checkbox), and '\_dpx.address.\_id' (with an unchecked checkbox). Below this is a button labeled 'Select manually'. To the right of the list are four small cards, each with a lightbulb icon and a number (2, 1, 1, 1) and a minus sign. The card with '2' is highlighted with a red box.

**Tip**

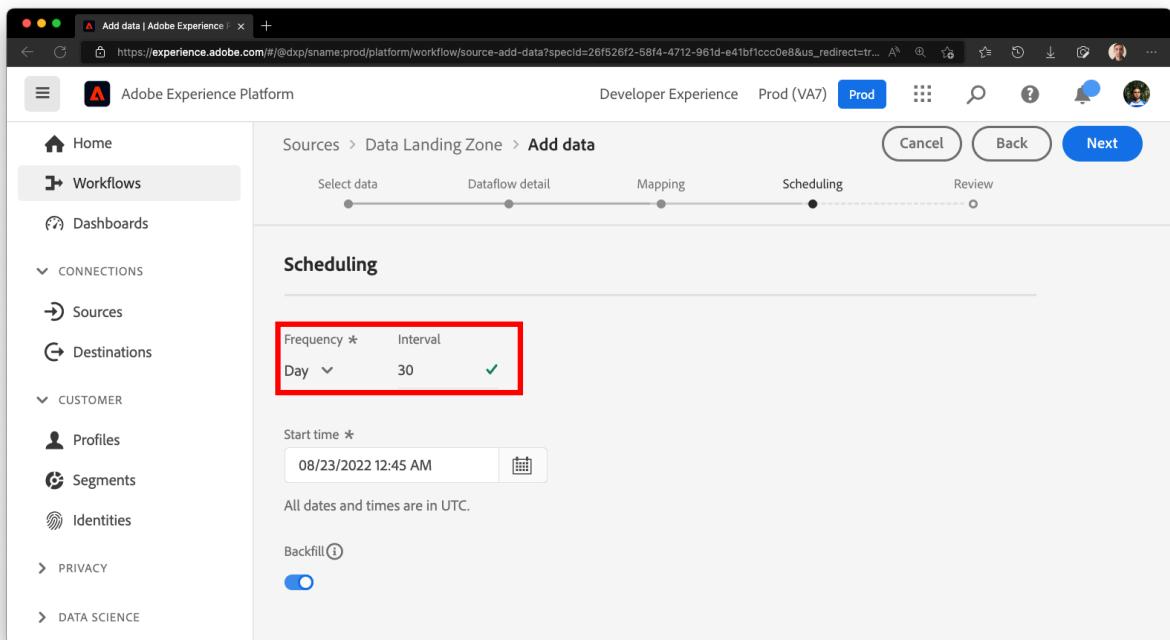
If you would like to change the recommendation, always check the new recommendation before unchecking the existing recommendation. If you uncheck all recommendations (with none checked), system treats it as an indication from the user that it got all recommendations for that source attribute wrong.

Click **Next** to proceed to schedule screen.

## Schedule

The **Scheduling** screen will load. Change the **Frequency** to **Day** and set the **Interval** as **30**. Leave the **Backfill** toggle turned ON. This will schedule the Dataflow to run once every 30 days.

<b>WARNING</b>	Do not set the Frequency to Once. Currently One-time ingestion flows cannot be edited.
<b>WARNING</b>	Schedule times are in UTC. Not your local time. By default, the Start time is set to current time + 1 minute

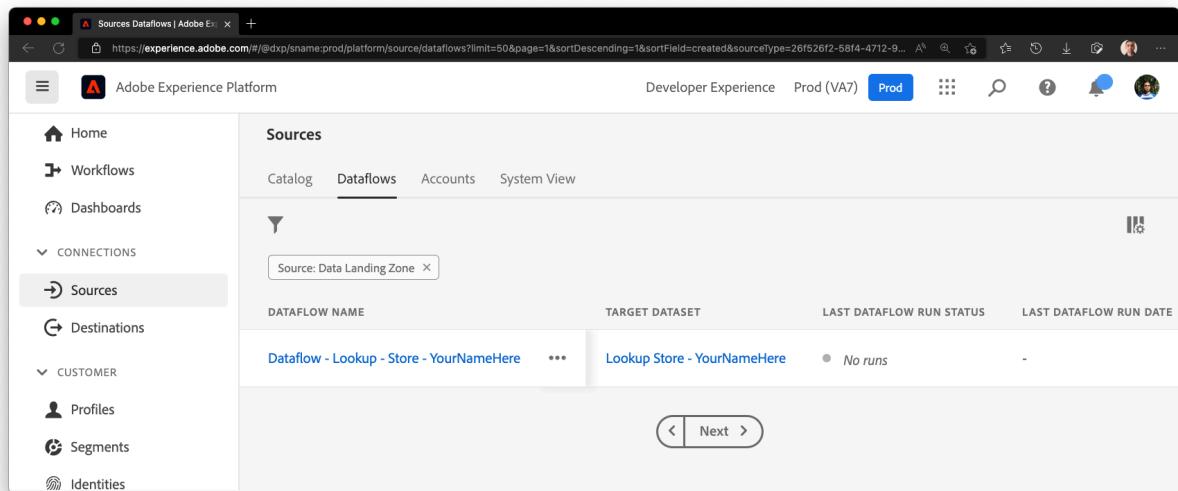


When the **Review** screen appears, click **Finish**.

The screenshot shows the 'Add data' workflow in the Adobe Experience Platform. The left sidebar shows navigation options like Home, Workflows (which is selected), Dashboards, CONNECTIONS, Sources, Destinations, CUSTOMER, PRIVACY, DATA SCIENCE, and DATA MANAGEMENT. The main area shows the 'Sources > Data Landing Zone > Add data' path. A progress bar at the top indicates steps: Select data, Dataflow detail, Mapping, Scheduling, and Review, with Review being the current step. The 'Review' section contains three cards: 'Connection' (Account name: Lab\_Lookup, Path: \_Store.csv, Columns: 9, status: Connected), 'Assign dataset and map fields' (Target dataset: Lookup Store - YourNameHere, Schema mapping status: Dataset assigned), and 'Scheduling' (Start time: 08/23/2022 12:45 AM, Frequency: Day, Interval: 30, status: Scheduled). At the bottom right are 'Cancel', 'Back', and a large blue 'Finish' button.

## Lab Task – Monitoring the Dataflow

It takes few minutes to create the Dataflow. Once Dataflow is created, you will see the following screenshot. Notice that Last Dataflow Run Status indicates **No runs**. First run will kick off approximately in 5 minutes.



The screenshot shows the Adobe Experience Platform web interface for managing dataflows. The left sidebar has a 'Sources' section selected. The main area displays a table of dataflows under the 'Dataflows' tab. One dataflow is listed:

DATAFLOW NAME	TARGET DATASET	LAST DATAFLOW RUN STATUS	LAST DATAFLOW RUN DATE
Dataflow - Lookup - Store - YourNameHere ***	Lookup Store - YourNameHere	<input checked="" type="radio"/> No runs	-

A search bar at the top says 'Source: Data Landing Zone'. Navigation buttons at the bottom include '<' and '>'. The top right shows 'Developer Experience Prod (VA7) Prod' and other user interface elements.

## Check scheduled execution

After 5 minutes, refresh the page and the following will appear. Notice the **Last Dataflow Run Status** and **Last Dataflow Run Date**. Click on the Dataflow name that appears as blue link.

DATAFLOW NAME	TARGET DATASET	LAST DATAFLOW RUN STATUS	LAST DATAFLOW RUN DATE
Dataflow - Lookup - Store - YourNameHere	Lookup Store - YourNameHere	Success	08/22/2022, 5:46 PM

Click on the Dataflow name to get a list of Dataflow Runs. 10 Records should be ingested.

DATAFLOW RUN START	PROCESSING TIME	RECORDS INGESTED	RECORDS FAILED	STATUS
08/22/2022, 5:46 PM	1 minute	10	0	Success

We will learn more about monitoring in the subsequent labs. For now, we will check that the data is loaded successfully.

## Lab Task – Verify the data

In the Left Nav bar, Go to **Datasets** and on the right panel, click on **Lookup Store – YourNameHere** that appears as a blue link.

The screenshot shows the 'Datasets' page in the Adobe Experience Platform. The left sidebar has a 'Datasets' menu item highlighted with a red box. The main area displays a table of datasets with columns: NAME, CREATED, SOURCE, and SCHEMA. One row, 'Lookup Store - YourNameHere', is highlighted with a red box. To the right, there's a summary card for 'Datasets' showing '33 Datasets' and a section for 'Most recently updated'.

NAME	CREATED	SOURCE	SCHEMA
Lookup Store - YourNameHere	08/22/2022, 5:19 PM	Schema	dep: Lookup Store
Customer Account - YourNameHere	08/22/2022, 5:18 PM	Schema	Customer Account - YourNameHere
dep: Lookup Plan	08/14/2022, 7:58 PM	Schema	dep: Lookup Plan
dep: Lookup Product	08/14/2022, 7:58 PM	Schema	dep: Lookup Product
dep: Lookup Store	08/14/2022, 7:58 PM	Schema	dep: Lookup Store
dep: Web	08/14/2022, 7:57 PM	Schema	dep: Web
dep: Ecommerce	08/14/2022, 7:57 PM	Schema	dep: Ecommerce

You will notice there are 10 records ingested. Click on the Preview Dataset

The screenshot shows the 'Dataset activity' page for 'Lookup Store - YourNameHere'. The left sidebar has a 'Datasets' menu item highlighted with a red box. The main area displays dataset activity metrics: 'Total records in previous month' (10), 'Ingested records in the last 7 days' (0), 'Ingested batches today' (1), and 'Failed batches today' (0). The right panel shows detailed dataset information: Name (Lookup Store - YourNameHere), Description (empty), Dataset ID (63041d154655631c077d6133), Table name (lookup\_store\_yournamehere), Profile (disabled), and Schema (dep: Lookup Store).

Data is loaded and displayed. Expand the XDM tree on the right to see child attributes.

**TIP**

When you click on a node such as `_dpx → address`, the right hand side will update to show only attributes within the selected parent node (`address` in this example). To see all attributes, click on the master node (`Lookup Store - YourNameHere` in this example).

The screenshot shows the Adobe Experience Platform interface. On the left, a sidebar navigation includes Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER, PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring), and DECISION MANAGEMENT. The 'Datasets' option is selected. The main area displays a dataset named 'Lookup Store - YourNameHere'. The XDM tree on the left shows a hierarchy starting with 'dep: Lookup Store', which has a child 'dpx | Object', which in turn has a child 'address | Object'. This 'address' node has three child attributes: 'city | String', 'postalCode | String', and 'state | String'. To the right of the tree, a table lists six rows of data:

	<code>_DXP.ADDRESS.CITY</code>	<code>_DXP.CLOSETIME</code>	<code>_DXP.ADDRESS.STREET</code>
1	Longhe	open at 9am / close at 9pm	68676 Lotheville Point
2	Zhenqian	open at 9am / close at 9pm	770 Fordem Circle
3	Gangou	open at 9am / close at 9pm	0 Buena Vista Drive
4	Wangjing	open at 9am / close at 9pm	7255 Dunning Junction
5	Xuebu	open at 9am / close at 9pm	706 Chinook Circle
6	Shah Alam	open at 9am / close at 9pm	35296 Hayes Street



# **Adobe Experience Platform**

## **Lab 1 - Data Pipeline Customer Accounts**

Adobe Experience Platform Bootcamp Deep Dive Edition

Name \_\_\_\_\_

Sandbox \_\_\_\_\_

## Lab Overview

Ingest Customer Accounts data into AEP in batch mode using Data Landing Zone. This Lab will leverage delimited data on Landing Zone and ML Recommendations to map most of the source data. This job will be scheduled to run every 7 days.

This lab will introduce you to customizing the ML recommendations in Data Prep and updating manual mappings.

**Expected time: 30 minutes**

## Learning Objectives

What should you walk away with after taking this Lab?

- Adding pass through mappings
- Using preview to check any data quality issues
- Address basic data quality issues using Calculated Fields

# Lab Tasks – Ingest Customer Account data

In this exercise, we will load the Customer Account data from Data Landing Zone to AEP Data Lake and Profile.

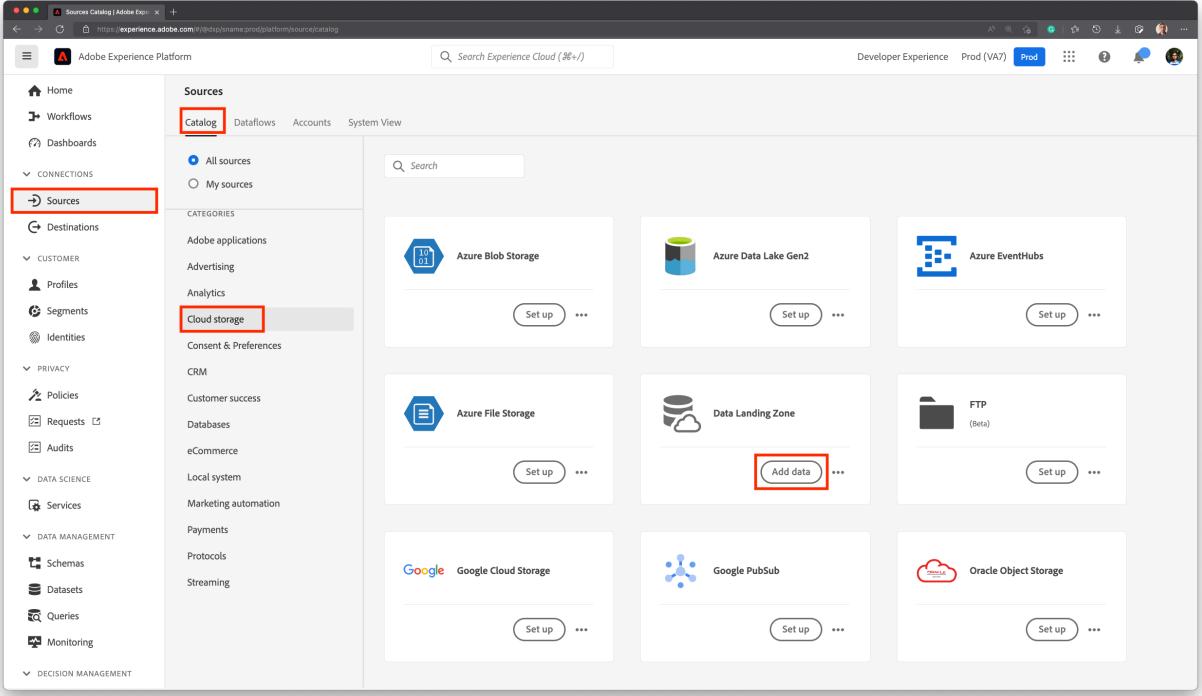
## Pre-requisites

1. Customer Account JSON file uploaded in the Azure ADLS Directory

## Select Source data

Go to Adobe Experience Platform à **Sources** à **Catalog** à **Cloud storage**. Click on **Setup / Add Data** for the Data Landing Zone.

**Tip** If at least one connection exists for that source, you will see “Add data” as the default action. If no connections exist for that source, you will see “Setup” as the default action.



The screenshot shows the Adobe Experience Platform interface. On the left, there's a sidebar with various navigation options like Home, Workflows, Dashboards, Connections, and more. Under 'Connections', the 'Sources' option is selected and highlighted with a red box. In the main content area, under 'Catalog', the 'Cloud storage' category is selected and highlighted with a red box. A grid of data sources is displayed, including Azure Blob Storage, Azure Data Lake Gen2, Azure EventHubs, Azure File Storage, Data Landing Zone (which has its 'Add data' button highlighted with a red box), FTP, Google Cloud Storage, Google PubSub, and Oracle Object Storage. Each source has a 'Set up' button and a three-dot menu icon.

In the **Select data** screen, navigate to `/dlz-user-container/project/PIPELINE`. Select the radio button next to **CUSTOMER** folder. Do NOT click on the folder name. On the righthand side, **Delimited** is automatically selected as Data format. Notice on the right side that the first file in the folder is automatically selected **Lab\_Customer\_001.csv**.

	CREATEDATE	MODIFYDATE	FIRSTNAME	LASTNAME	BIRTH_DATE
0	1660096901	2022-08-09T22:01:41Z	Larina	Loveredge	1941-04-27
1	1660096901	2022-08-09T22:01:41Z	Sabina	Heindle	1970-07-29
2	1660096901	2022-08-09T22:01:41Z	Danika	Ruffey	1946-12-09
3	1660096901	2022-08-09T22:01:41Z	Mellicent	Fernyhough	1946-08-06
4	1660096901	2022-08-09T22:01:41Z	Rosemonde	Cann	1940-12-11
5	1660096901	2022-08-09T22:01:41Z	Genevra	Pentony	1954-04-13

In the preview here, notice the following attributes:

- **sms\_optIn** has null values (shown in preview as -)
- **account\_create\_date** has string values along with date and time values. **account\_end\_date** has proper date format.

The screenshot shows the 'Add data' workflow step in Adobe Experience Platform. The left sidebar is the navigation menu. The main area shows the 'Selected data' section for a 'CUSTOMER' pipeline. A red box highlights the first column of the preview table, labeled 'SMS\_OPTIN', which contains several '-' characters indicating null values.

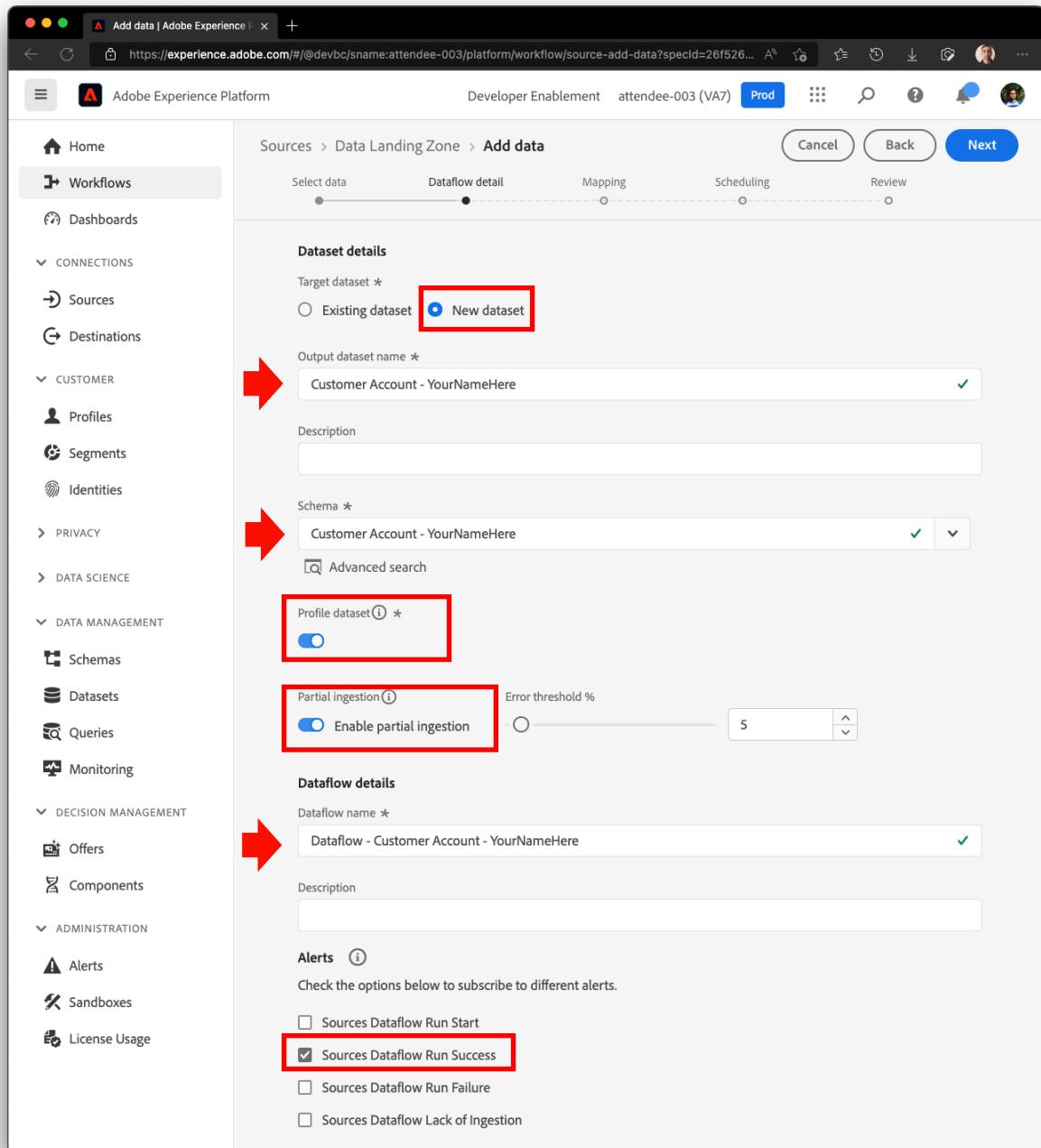
SMS_OPTIN	CUSTOMER_ID	SHIPPING_STREET_ADDRESS	SHIPPING_CITY	SHIPPING_STATE
-	736072406	706 Sloan Lane	Fresno	CA
y	495305838	985 Elmside Court	Austin	TX
n	442237547	9 Longview Center	Kansas City	MO
-	462757264	0 Lyons Crossing	Roanoke	VA
y	764162251	82972 Karstens Junction	Worcester	MA
y	983010429	6 Di Loreto Center	Concord	CA
-	318629820	2081 Commercial Parkway	Houston	TX
-	503815514	638 Helena Hill	Oakland	CA

The screenshot shows the 'Add data' workflow step in Adobe Experience Platform. The left sidebar is the navigation menu. The main area shows the 'Selected data' section for a 'Lab\_Customer.csv' pipeline. A red box highlights the last two columns of the preview table, labeled 'ACCOUNT\_CREATE\_DATE' and 'ACCOUNT\_END\_DATE', which contain string values such as 'Created on 2022-04-22T19:34:17Z' and '2022-04-03T05:13:06Z'.

BILLING_ZIP_CODE	PLAN_ID	PLAN_NAME	ACCOUNT_CREATE_DATE	ACCOUNT_END_DATE	SOURCE
93794	m3	pro	Created on 2022-04-22T19:34:17Z	2022-04-03T05:13:06Z	web
78703	m1	basic	Created on 2022-05-11T16:25:15Z	2022-03-30T05:58:31Z	web
64142	m1	basic	Created on 2022-04-16T13:36:34Z	2022-05-15T03:34:11Z	web
24009	m1	basic	Created on 2022-06-08T19:23:07Z	2022-05-09T07:51:30Z	inStore
01605	m2	ultimate	Created on 2022-05-01T07:16:46Z	2022-02-23T03:06:40Z	web
94522	m1	basic	Created on 2022-04-18T14:43:51Z	2022-02-02T11:32:51Z	inStore
77255	m1	basic	Created on 2022-05-20T22:46:16Z	2022-05-15T12:09:52Z	web
94627	m1	basic	Created on 2022-04-19T08:06:50Z	2022-01-14T16:19:08Z	web

## Define the Target Dataset

In the Dataflow detail screen, choose **New dataset**. Name the output dataset as **Customer Account - YourNameHere**. Select the schema name **Customer Account – YourNameHere**. Turn ON the **Profile dataset** toggle box. Turn ON the **Enable partial ingestion**. Set the Dataflow name as **Dataflow - Customer Account – YourNameHere**. Turn ON the alert **Sources Dataflow Run Success**. Click **Next** to continue.



## Data Prep / Transformation

In the **Mapping** screen, Recommendation engine will map several attributes automatically. However, there will be some errors that need to be addressed before moving on to next step. The initial screen will look like this

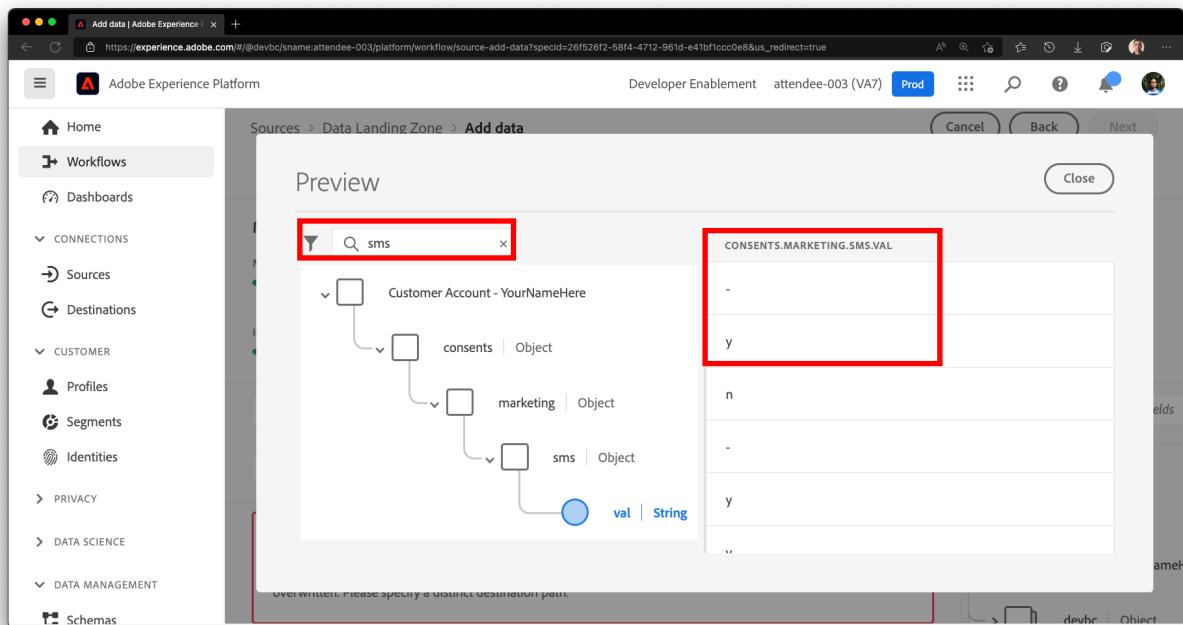
**Warning** Due to the ML Recommendations, your screen may look different than the screenshot below.

The screenshot shows the 'Add data' interface in Adobe Experience Platform. The left sidebar includes sections for Home, Workflows, Dashboards, Connections (Sources, Destinations), Customer (Profiles, Segments, Identities), Privacy (Policies, Requests, Audits), Data Science (Services), and Data Management (Schemas, Datasets). The main area is titled 'Sources > Data Landing Zone > Add data' and is currently on the 'Mapping' step. A progress bar indicates 23 of 23 Mapped fields, 1 of 1 Required fields, and 1 of 2 Identity fields. A prominent red error message box states: 'There was error(s) preparing mappings. There is a duplicate mapping for the target path person.name.lastName. The data at the XDM path will be overwritten. Please specify a distinct destination path.' The mapping table lists four source fields: createDate, modifyDate, firstName, and lastName, which are mapped to target fields: \_repo.createDate, \_repo.modifyDate, person.name.firstName, and person.name.lastName respectively. The 'lastName' mapping is highlighted with a red border.

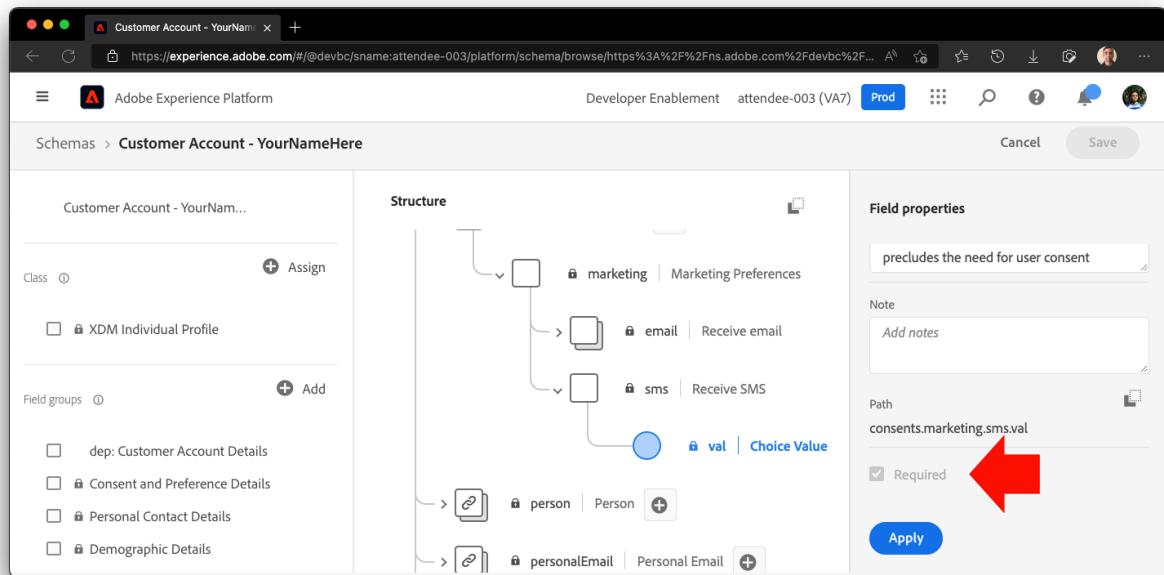
## PREVIEW THE MAPPING OUTPUT

### SMS Opt In

Preview the mapping output by clicking  In Preview screen, navigate to the consents marketing.sms.val attribute and you will notice that the mapping output has NULL values shown as hyphens (-). This is consistent with what we saw in the Preview of source in source object selection screen. However, this attribute is marked as a required attribute as seen in the schema screenshot below.



The screenshot shows the 'Add data' preview screen in Adobe Experience Platform. On the left, the navigation menu is visible with 'Workflows' selected. The main area shows a tree structure under 'Preview' with a search bar for 'sms'. The 'consents.marketing.sms.val' attribute is highlighted with a red box, showing its value as a list of hyphens (-) and a single 'y'.



The screenshot shows the 'Customer Account - YourNameHere' schema screen in Adobe Experience Platform. The 'Structure' panel on the left shows the field hierarchy. The 'Field properties' panel on the right shows the path 'consents.marketing.sms.val' and the 'Required' checkbox, which is checked and highlighted with a red arrow.

## Learnings

At runtime, the rows without a SMS value will be rejected. If Partial ingestion is not enabled, this will fail the entire batch. If partial ingestion is enabled, the rows will NULL value will be rejected and other rows will be ingested

## Account Create Date

In Preview screen, navigate to the `_devbc.account.createDate`.

ACCOUNT.ENDDATE	_DEVBC.ACCOUNT.ACQSOURCE	_DEVBC.ACCOUNT.
i-03T05:13:06Z	web	-
i-30T05:58:31Z	web	-
i-15T03:34:11Z	web	-
i-09T07:51:30Z	inStore	-
i-23T03:06:40Z	web	-
i-02T11:32:51Z	inStore	-
i-15T12:09:52Z	web	-
i-14T16:19:08Z	web	-

## Learnings

It is showing as NULL even though the source has value. This is because the source data contains string values such as "**Created on 2022-06-04T19:2409Z**". This is mapped to `_devbc.account.createDate` which is a timestamp data type. Since the input cannot be automatically converted into a timestamp, Data Prep now raises a warning (seen in Monitoring console when Dataflow runs) and NULLifies the result.

We will NOT fix this error as we will use this to understand how Data Prep / Mapper errors are reported later.

**TIP**

Note that the input file is CSV and all attributes are Strings. Data Prep automatically recognized the date attributes such as `account_end_date` and converted them into date. Similarly, Zip Code is automatically converted into number.

## Re-map the attributes

When you scroll down, you will notice **plan\_name** mapped to **person.name.lastName** (or some other invalid attribute). Click on the bulb icon (💡 2) next to it and choose the **\_devbc.plan.name**. The uncheck the **person.name.lastName**. If it is already mapped correctly, continue forward.

### WARNING

Do not uncheck the **person.name.lastName** first. If you uncheck first, the ML recommendations will reset and you will have to select the **\_devbc.plan.name** manually from the XDM schema on the right

The screenshot shows the 'Add data' workflow in the Adobe Experience Platform. The 'Mapping' step is active. In the main area, a mapping rule is shown: 'plan\_name' is mapped to '\_devbc.plan.name'. To the right, a 'Mapping recommendations' panel is open, listing three suggestions for the 'plan\_name' source field. The third suggestion, '\_devbc.plan.name', is checked and highlighted with a red box. A second red box highlights the lightbulb icon with the number '2' next to '\_devbc.plan.name', indicating that two recommendations are available for this field.

Similarly, Change the mappings as necessary to match the table below. The table is sorted in Alphabetical order of XDM fields. The rows marked with ★ are incorrect and will be corrected in the next step.

#	Source Column	XDM fields ↓	✓
1	source	_devbc.account.acqSource	
2	account_create_date	_devbc.account.createDate	
3	account_end_date	_devbc.account.endDate	
4	customer_id	_devbc.customerID	
5	plan_name	_devbc.plan.name	
6	plan_id	_devbc.plan.planID	
7	create_date	_repo.createDate	
8	modifyDate	_repo.modifyDate	
9	billing_city	billingAddress.city	
10	billing_zip_code	billingAddress.postalCode	
11	billing_state	billingAddress.state	
12	billing_street_address	billingAddress.street1	
13	email_optIn	consents.marketing.email.val	
14	sms_optIn	consents.marketing.sms.val	★
15	mobile_phone	mobilePhone.number	
16	birth_Date	person.birthDayAndMonth	★
17	first_name	person.name.firstName	
18	last_name	person.name.lastName	
19	email	personalEmail.address	
20	shipping_city	shippingAddress.city	
21	shipping_zip_code	shippingAddress.postalCode	
22	shipping_state	shippingAddress.state	
23	shipping_street_address	shippingAddress.street1	

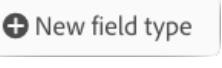
## CALCULATED FIELDS

We will now correct the two rows that were marked ★ by using Calculated fields to calculate the values.

### SMS Opt In

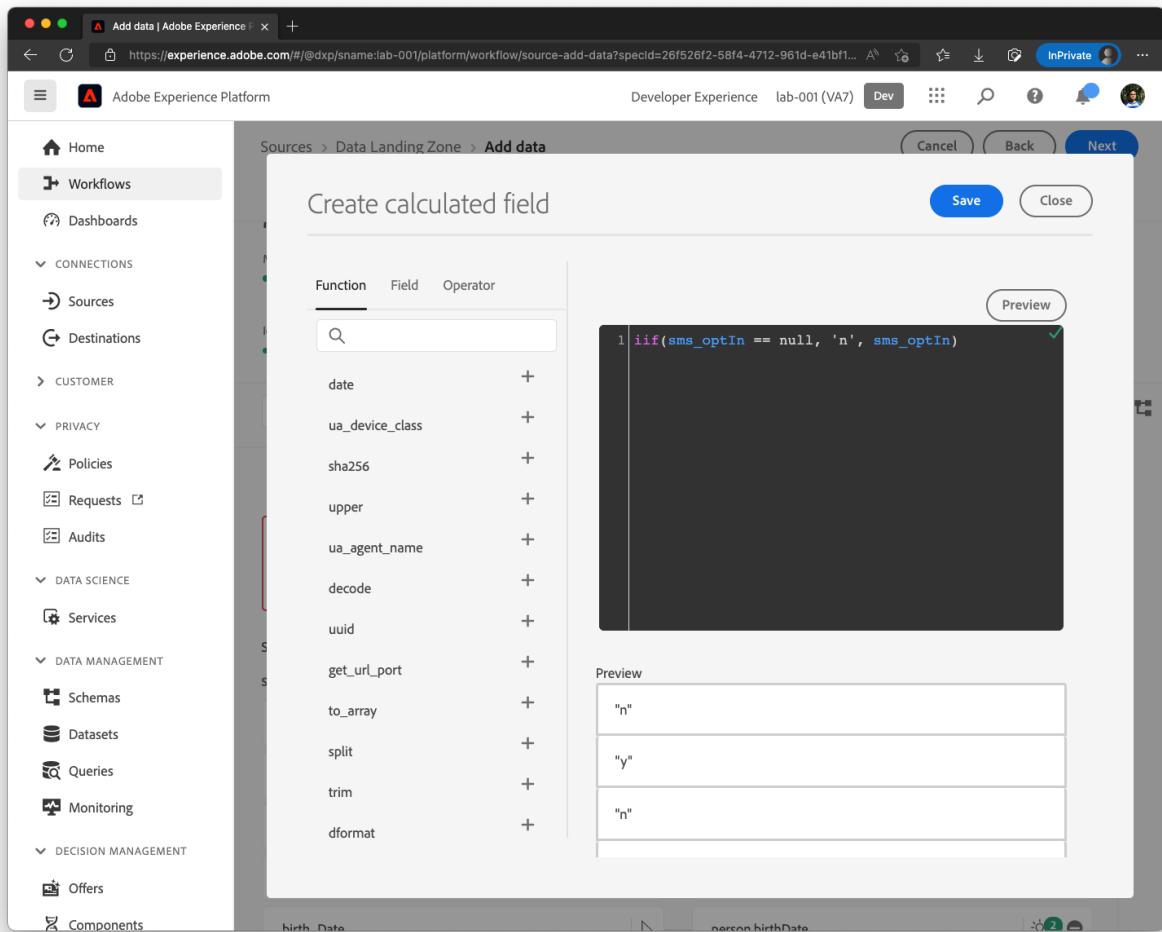
Scroll down to the mapping for XDM attribute **consents.marketing.sms.val**, and remove the mapping.



Now create a calculated field by clicking  icon(sometimes shown as  ) > **Add Calculated Field**. In the Create Calculated field dialog box write the following expression and hit **Preview**.

```
iif(sms_optIn == null, 'n', sms_optIn)
```

You should see a green checkmark in the top right corner of the black box indicating the validity of the expression and the data Preview should only show "n" or "y" as values. Then click **Save**.

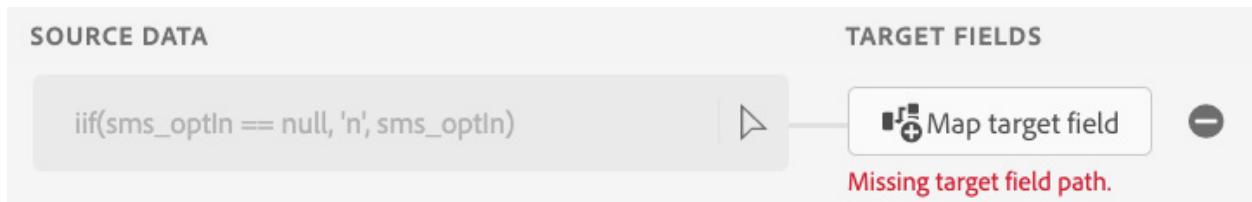


The screenshot shows the 'Add data' interface in Adobe Experience Platform. The left sidebar has 'Workflows' selected. The main area shows the 'Create calculated field' dialog. The expression 'iif(sms\_optIn == null, 'n', sms\_optIn)' is entered in the preview area, and a green checkmark indicates it's valid. The preview table shows three rows: 'n', 'y', and 'n'.

**NOTE**

For the demonstration purposes of this lab, we are assuming lack of consent as an explicit No represented by "n". Different customers may have different way of handling lack of consent

The mapping will now look like this.



Click on Map target field and choose **consents.marketing.sms.val** attribute. Final mapping will look like this:



## Learnings

Data Prep conditional operators can be used to cleanse input fields. Similar logic can be applied to email consent (**consents.marketing.email.val**)

## Birth Dates

Find the mapping associated with the XDM attribute **person.birthDayAndMonth** and remove that row by clicking on the (-) icon next to it.



If you find a mapping associated with **person.birthYear**, remove that row as well.

Add a Calculated Field by clicking the (+) icon and **Add Calculated Field**. Type in the following expression into the calculated field

```
date_part("year", date(birth_Date))
```

The screenshot shows the Adobe Experience Platform Data Prep Mapping Functions interface. On the left, there's a sidebar with navigation links like Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY (Policies, Requests, Audits), DATA SCIENCE (Services), and DATA MANAGEMENT (Schemas, Datasets). The main area is titled "Create calculated field". It has tabs for Function, Field, and Operator, with "Function" selected. A search bar is at the top of the list. Below it is a list of functions: date, ua\_device\_class, sha256, upper, ua\_agent\_name, decode, uuid, get\_url\_port, to\_array, and split. To the right of the list is a code editor window with a "Preview" button. The code in the preview is:

```
1 date_part("year", date(birth_date))
```

Below the code editor is a "Preview" section showing two rows of data: 1972 and 1940. At the bottom of the screen, there are "Save", "Cancel", "Back", and "Next" buttons.

Click Save to save this expression and map it to **person.birthYear**



Add another calculated field with the following expression and map it to **person.birthDayAndMonth**

```
concat(  
    date_part("month", date(birth_Date)).toString(),  
    "-",  
    date_part("day", date(birth_Date)).toString()  
)
```

The screenshot shows the 'Add data' interface in Adobe Experience Platform. On the left, there's a sidebar with navigation links like Home, Workflows, Dashboards, Connections (Sources, Destinations), Customer (Profiles, Segments, Identities), Privacy, Data Science, Data Management (Schemas, Datasets, Queries, Monitoring), and Decision Management (Offers, Components). The main area is titled 'Create calculated field' and has tabs for Function, Field, and Operator. A search bar is at the top of the list of functions. Below are listed functions: date, ua\_device\_class, sha256, upper, ua\_agent\_name, decode, uuid, get\_url\_port, to\_array, split, trim, and dformat. To the right, there's a code editor with the following expression:

```
concat(  
    date_part("month", date(birth_Date)).toString(),  
    "-",  
    date_part("day", date(birth_Date)).toString()  
)
```

Below the code editor is a 'Preview' section showing three rows of output: "4-27", "7-29", and "12-9". There are 'Save' and 'Close' buttons at the top right of the dialog.

**NOTE**

Notice that the month for some rows is a single digit. For example, 4-27 instead of 04-27. This will cause the data to be failed. We will use that to check errors in the monitoring console.

Now click on the **Map target field** button and assign it to **person.birthDayAndMonth**. The mapping will now look as below.



## Final Mapping Set

The final mapping set looks like below

#	SourceColumn	XDM field ↓	✓
1	Source	_devbc.account.acqSource	
2	account_create_date	_devbc.account.createDate	
3	account_end_date	_devbc.account.endDate	
4	customer_id	_devbc.customerID	
5	plan_name	_devbc.plan.name	
6	plan_id	_devbc.plan.planID	
7	billing_city	billingAddress.city	
8	billing_zip_code	billingAddress.postalCode	
9	billing_state	billingAddress.state	
10	billing_street_address	billingAddress.street1	
11	email_optIn	consents.marketing.email.val	
12	mobile_phone	mobilePhone.number	
13	first_name	person.name.firstName	
14	last_name	person.name.lastName	
15	email	personalEmail.address	
16	create_date	repo.createDate	
17	last_modified	repo.modifyDate	
18	shipping_city	shippingAddress.city	
19	shipping_zip_code	shippingAddress.postalCode	
20	shipping_state	shippingAddress.state	
21	shipping_street_address	shippingAddress.street1	

Calculated fields		
iif(sms_optIn == null, 'n', sms_optIn)	consents.marketing.sms.val	✓
concat( date_part("month", date(birth_Date)). toString(), " -", date_part("day", date(birth_Date)).to- String() )	person.birthDayAndMonth	
date_part("year", date(birth_Date))	person.birthYear	

## Schedule

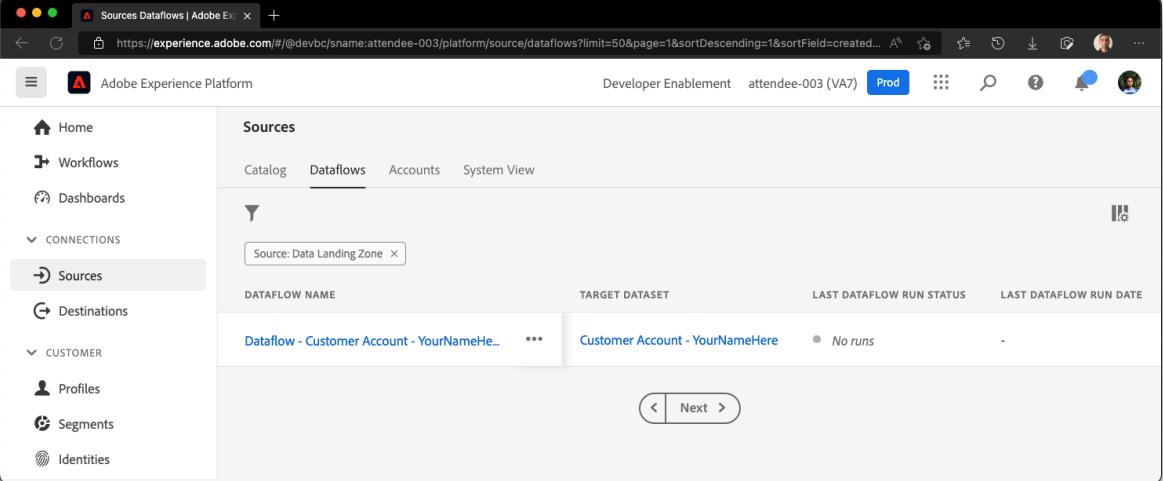
In the **Scheduling** step, set the **Frequency** to **Day** and **Interval** to **30** and leave the **Backfill** turned on. This will schedule the Dataflow to run every 30 days

The screenshot shows the 'Add data' workflow configuration interface in the Adobe Experience Platform. The left sidebar contains navigation links for Home, Workflows (selected), Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT. The main panel shows the 'Sources > Data Landing Zone > Add data' path. The top navigation bar includes 'Developer Enablement attendee-003 (VA7)', 'Prod', and other user icons. Below the path, a progress bar indicates steps: 'Select data' (done), 'Dataflow detail' (done), 'Mapping' (done), 'Scheduling' (done), and 'Review' (not yet started). The 'Scheduling' section is highlighted with a red box. It shows 'Frequency \*' set to 'Day' and 'Interval' set to '30'. A green checkmark is next to the interval input. Below this, the 'Start time \*' is set to '08/24/2022 5:37 AM' with a calendar icon. A note says 'All dates and times are in UTC.' The 'Backfill' toggle switch is turned on. At the bottom right of the scheduling section are 'Cancel', 'Back', and 'Next' buttons.

Review the data flow and click **Finish**.

The screenshot shows the 'Add data' workflow in the Adobe Experience Platform. The left sidebar navigation includes Home, Workflows (selected), Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring). The main content area is titled 'Sources > Data Landing Zone > Add data' and shows the 'Review' step of the workflow. The review steps are: Select data, Dataflow detail, Mapping, Scheduling, and Review. The 'Review' step is currently active. The 'Connection' section shows an account name 'Customer Account - YourNameHere', source platform 'Data Landing Zone', path 'Lab\_Customer.csv', and columns '23'. A green checkmark indicates it is 'Connected'. The 'Assign dataset and map fields' section shows the target dataset 'Customer Account - YourNameHere' and schema mapping 'Customer Account - YourNameHere'. A green checkmark indicates it is 'Dataset assigned'. The 'Scheduling' section shows a start time '08/24/2022 5:37 AM', frequency 'Day', and interval '30'. A green checkmark indicates it is 'Scheduled'. At the top right are 'Cancel', 'Back', and 'Finish' buttons.

It takes few minutes to create the Dataflow. Once Dataflow is created, you will see the following screenshot. Notice that Last Dataflow Run Status indicates **No runs**. First run will kick off approximately in 15 minutes.



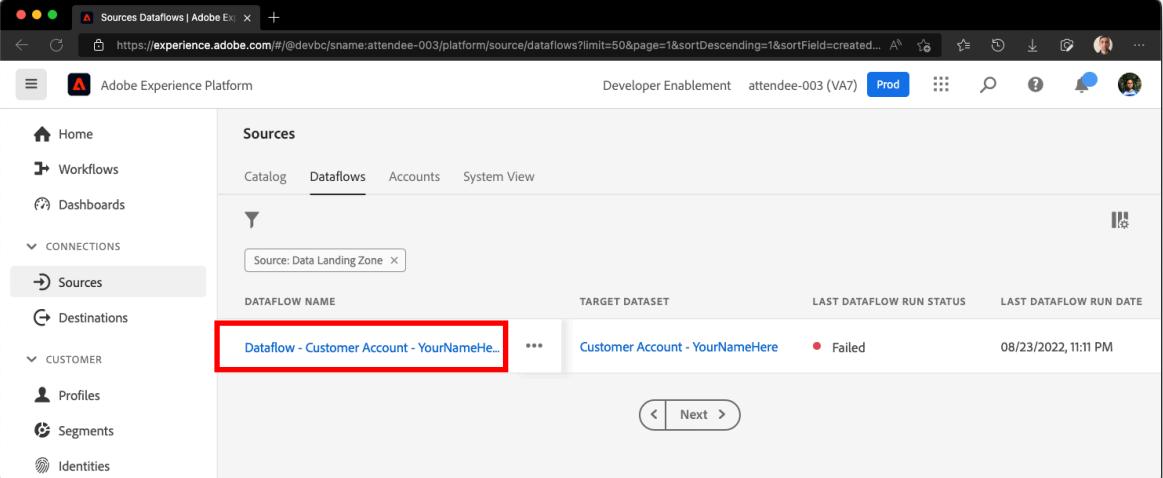
The screenshot shows the Adobe Experience Platform interface for managing dataflows. The left sidebar has sections for Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), and Sources. The Sources section is selected. The main area is titled 'Sources' with tabs for Catalog, Dataflows (which is selected), Accounts, and System View. A search bar at the top says 'Source: Data Landing Zone'. Below it, a table lists one dataflow:

DATAFLOW NAME	TARGET DATASET	LAST DATAFLOW RUN STATUS	LAST DATAFLOW RUN DATE
Dataflow - Customer Account - YourNameHe...	Customer Account - YourNameHere	No runs	-

At the bottom right of the table is a navigation bar with '<' and '>' buttons and a 'Next' button.

## Check execution status and error messages

After 15 minutes, the following will appear. Notice the **Last Dataflow Run Status** and **Last Dataflow Run Date**. Click on the Dataflow name that appears as a blue link



This screenshot is similar to the previous one, but the dataflow entry has changed. The 'Dataflow Name' column now contains a blue link 'Dataflow - Customer Account - YourNameHe...', which is highlighted with a red rectangle. The 'Last Dataflow Run Status' column shows a red circle with a white dot, indicating 'Failed'. The 'Last Dataflow Run Date' column shows the date and time '08/23/2022, 11:11 PM'.

We will notice that there is one Dataflow run and in Failed status. Click on the Dataflow run with timestamp that is listed as a blue link.

The screenshot shows the Dataflow - Customer Account page in the Adobe Experience Platform. The left sidebar has 'Sources' selected. The main area displays 'Dataflow activity' with 'Records ingested' at 0 and 'Records failed' at 0. A table lists a single run: 'DATAFLOW RUN START' is 08/23/2022, 11:11 PM; 'PROCESSING TIME' is 1 minute; 'RECORDS INGESTED' and 'RECORDS FAILED' are both '-'; 'STATUS' is 'Failed'. The 'Failed' status is highlighted with a red box. On the right, the 'Properties' panel shows the dataflow is 'Enabled' and has a name 'Dataflow - Customer Account - YourNameHere'. A 'Logs' tab is also visible.

We will notice that **INGEST-1517-400** error code is listed.

ERROR CODE	RECORD COUNT	DESCRIPTION
MAPPER-3700-199	-	Field transformation warning
INGEST-1517-400	-	The value does not conform to the specified regex pattern.

The screenshot shows the Adobe Experience Platform Dataflow Run overview page. The left sidebar contains navigation links for Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring), DECISION MANAGEMENT (Offers, Components), and ADMINISTRATION. The main content area displays the Dataflow run overview for run ID 2dba063a-4333-465b-a572-61d52433c00a. It shows 1 record ingested, 0 records failed, 1 total file, and 5.09 kB size. The status is Failed, with a start time of 08/23/2022, 11:11 PM and an end time of 08/23/2022, 11:12 PM. The error summary indicates a CONNECTOR-2001-500: Error in processing (parsing, validation or transformation) the copied data. The Dataflow run errors section shows one entry for INGEST-1517-400 with the message: "The value does not conform to the specified regex pattern." This entry is highlighted with a red border.

Click on the Preview error diagnostics to preview the error messages in detail.

The screenshot shows the Adobe Experience Platform interface for a Dataflow run. The left sidebar has a 'Sources' section selected. The main content area displays a summary of the dataflow run, including records ingested (1), records failed (0), total files (1), and size of data (5.09 kB). It also shows the status as Failed, with start and end times of 08/23/2022, 11:11 PM and 08/23/2022, 11:12 PM respectively. Below this, there's a partial ingestion summary and an error summary stating 'CONNECTOR-2001-500: Error in processing (parsing, validation or transformation) the copied data.' At the bottom, there are tabs for 'Dataflow run errors' (selected) and 'Files'. Under 'Dataflow run errors', there are two radio buttons: 'Records failed' (selected) and 'Records skipped'. A note says 'Parsing errors occurred when converting or validating the data. Successfully ingested 0 rows, review the associated diagnostic files for additional details.' A table lists error codes and descriptions:

ERROR CODE	RECORD COUNT	DESCRIPTION
MAPPER-3700-199	-	Field transformation warning
INGEST-1517-400	-	The value does not conform to the specified regex pattern.

A red box highlights the 'Preview error diagnostics' link next to the 'Dataflow run errors' tab.

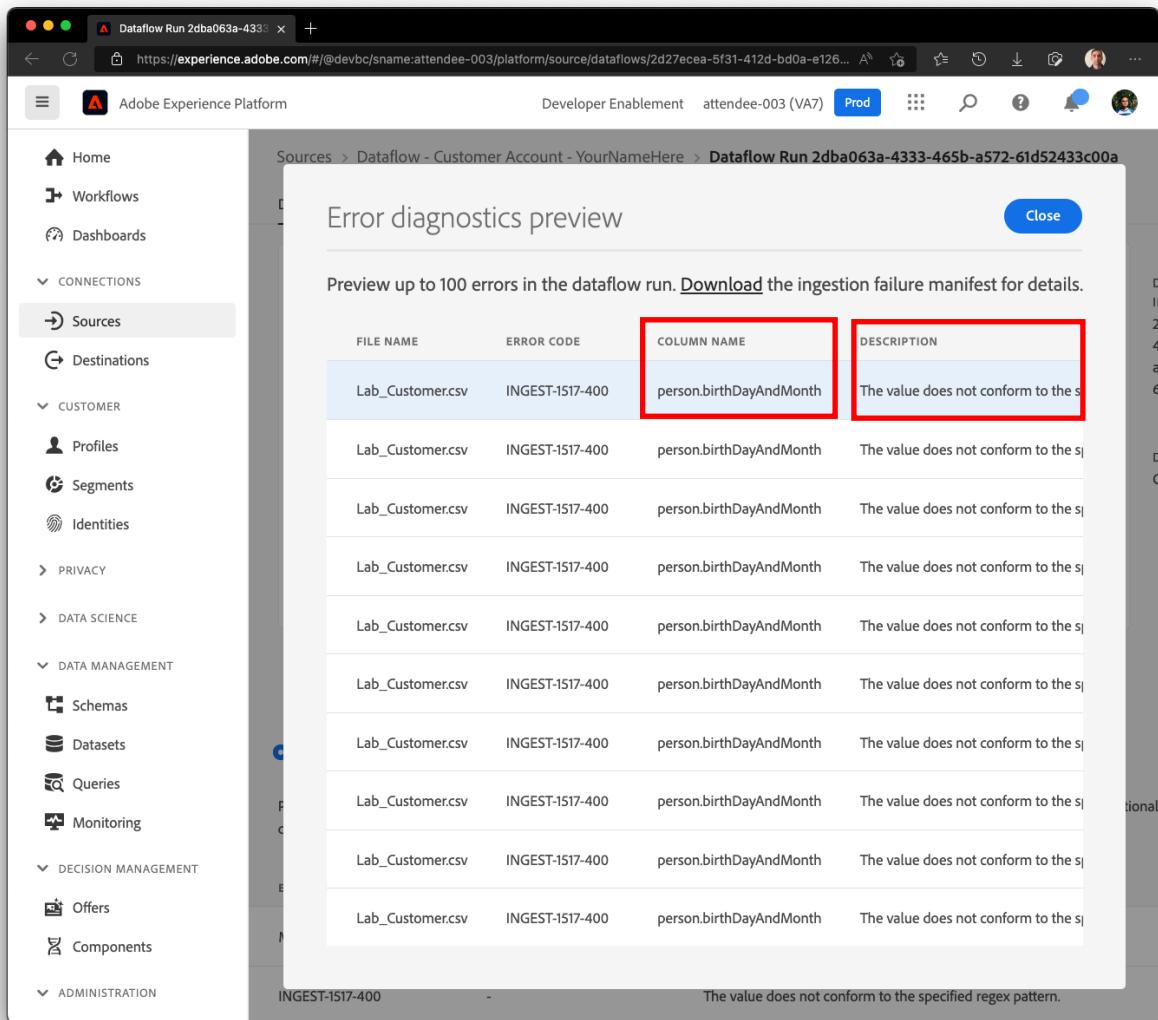
Error diagnostics preview will load. We can see the XDM attribute where the error occurred (**COLUMN NAME**) and the actual error message (**DESCRIPTION**).

## Error 1

The description will have the following text

**The value does not conform to the specified regex pattern: [0-1][0-9]-[0-9][0-9] in field: person.birthDayAndMonth of type: String.**

This error occurs because **person.birthDayAndMonth** is expected in the format of two digit month followed by a hyphen followed by two digit day. For example, the values generated are 4-27. Whereas, the values should be 04-27.



The screenshot shows the Adobe Experience Platform interface with a modal window titled "Error diagnostics preview". The modal displays a table of errors from a dataflow run. The columns are FILE NAME, ERROR CODE, COLUMN NAME, and DESCRIPTION. The first error row is highlighted with red boxes around the "COLUMN NAME" and "DESCRIPTION" columns. The "COLUMN NAME" cell contains "person.birthDayAndMonth" and the "DESCRIPTION" cell contains "The value does not conform to the specified regex pattern: [0-1][0-9]-[0-9][0-9] in field: person.birthDayAndMonth of type: String.". Below the table, a message states "Preview up to 100 errors in the dataflow run. Download the ingestion failure manifest for details." and "The value does not conform to the specified regex pattern." at the bottom.

FILE NAME	ERROR CODE	COLUMN NAME	DESCRIPTION
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern: [0-1][0-9]-[0-9][0-9] in field: person.birthDayAndMonth of type: String.
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern: [0-1][0-9]-[0-9][0-9] in field: person.birthDayAndMonth of type: String.
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern: [0-1][0-9]-[0-9][0-9] in field: person.birthDayAndMonth of type: String.
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern: [0-1][0-9]-[0-9][0-9] in field: person.birthDayAndMonth of type: String.
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern: [0-1][0-9]-[0-9][0-9] in field: person.birthDayAndMonth of type: String.
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern: [0-1][0-9]-[0-9][0-9] in field: person.birthDayAndMonth of type: String.
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern: [0-1][0-9]-[0-9][0-9] in field: person.birthDayAndMonth of type: String.
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern: [0-1][0-9]-[0-9][0-9] in field: person.birthDayAndMonth of type: String.
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern: [0-1][0-9]-[0-9][0-9] in field: person.birthDayAndMonth of type: String.
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern: [0-1][0-9]-[0-9][0-9] in field: person.birthDayAndMonth of type: String.

## Error 2

Scroll down to see the second error message associated with **createDate**. The description will have the following text

**Error transforming data for destination path \_devbc.account.createDate. Details: Unable to convert Created on 2022-04-22T19:34:17Z to schema type DATE...**

The screenshot shows the Adobe Experience Platform Dataflow Run interface. On the left, there's a navigation sidebar with options like Home, Workflows, Dashboards, Connections, Sources (which is selected), Destinations, Customer, Privacy, Data Science, Data Management, Decision Management, and Administration. The main area displays a "Dataflow Run 2dba063a-4333-465b-a572-61d52433c00a". Below it, a modal window titled "Error diagnostics preview" lists errors from a file named "Lab\_Customer.csv". The errors are categorized by file name, error code, column name, and description. One specific error is highlighted with a red box: "createDate" with the description "Error transforming data for destination path \_devbc.account.createDate. Details: Unable to convert Created on 2022-04-22T19:34:17Z to schema type DATE...".

FILE NAME	ERROR CODE	COLUMN NAME	DESCRIPTION
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern.
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern.
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern.
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern.
Lab_Customer.csv	INGEST-1517-400	person.birthDayAndMonth	The value does not conform to the specified regex pattern.
Lab_Customer.csv	MAPPER-3700-199	createDate	Error transforming data for destination path _devbc.account.createDate. Details: Unable to convert Created on 2022-04-22T19:34:17Z to schema type DATE...
Lab_Customer.csv	MAPPER-3700-199	createDate	Error transforming data for destination path _devbc.account.createDate. Details: Unable to convert Created on 2022-04-22T19:34:17Z to schema type DATE...
Lab_Customer.csv	MAPPER-3700-199	createDate	Error transforming data for destination path _devbc.account.createDate. Details: Unable to convert Created on 2022-04-22T19:34:17Z to schema type DATE...
Lab_Customer.csv	MAPPER-3700-199	createDate	Error transforming data for destination path _devbc.account.createDate. Details: Unable to convert Created on 2022-04-22T19:34:17Z to schema type DATE...
Lab_Customer.csv	MAPPER-3700-199	createDate	Error transforming data for destination path _devbc.account.createDate. Details: Unable to convert Created on 2022-04-22T19:34:17Z to schema type DATE...

This error occurs because the source data has strings such as "Created on 2022-04-22T19:34:17Z". This value cannot be converted into a Date automatically. A calculated field must be used to cleanse the data. Scroll to the right side to see the source data associated with this error code.

The screenshot shows the Adobe Experience Platform Dataflow Run interface. The left sidebar navigation includes Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring), DECISION MANAGEMENT (Offers, Components), and ADMINISTRATION. The main content area displays the 'Error diagnostics preview' for a specific run. The errors listed are related to date conversion issues, such as 'Created on 2022-04-22T19:34:17Z' being converted to a string type. A red box highlights the first error entry:

account_create_date	Created on 2022-04-22T19:34:17Z
account_end_date	2022-04-03T05:13:06Z
source	web

We will now go back and fix these errors.

## Correct the Dataflows

Go back to the Dataflows page by clicking on the Dataflow name.

The screenshot shows the Adobe Experience Platform Dataflows interface. On the left, a sidebar navigation includes Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets, Queries). The Sources section is currently selected. The main content area displays a Dataflow run overview for a run named "Dataflow Run 2dba063a-4333-465b-a572-61d52433c00a". The run status is "Failed". The error summary indicates a parsing or validation error in the connector. At the bottom, there are tabs for "Dataflow run errors" (selected) and "Files", along with links to "Preview error diagnostics" and "Download".

Click on the Update dataflow

The screenshot shows the Dataflow - Customer Account interface. On the left, a sidebar navigation includes Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets, Queries). The main area displays 'Dataflow activity' with 'Records ingested' and 'Records failed' both at 0. Below this is a table showing 'DATAFLOW RUN START' as 08/23/2022, 11:11 PM, 'PROCESSING TIME' as 1 minute, and 'RECORDS INGESTED' and 'RECORDS FAILED' both as -. To the right is a 'Properties' panel with a red box around the 'Update dataflow' button. Other properties include 'Edit schedule', 'Enabled' (switched on), and 'Delete'. Below the properties are sections for 'Dataflow name' (Dataflow - Customer Account - YourNameHere), 'Description' (empty), 'Source data' (dlz-user-container/project/PIPELINE/Lab\_Customer.csv), and 'Target dataset' (Customer Account - YourNameHere).

Data Preview screen will load. Click Next to go to Dataflow detail page.

The screenshot shows the 'Add data' screen for a Workflow. The sidebar navigation is identical to the previous screenshot. The main area shows a 'Sources > Data Landing Zone > Add data' step. It has tabs for 'Select data' (selected), 'Dataflow detail', 'Mapping', 'Scheduling', and 'Review'. The 'Select data' tab shows a preview of the 'Lab\_Customer.csv' file under 'dlz-user-container > project > PIPELINE'. The 'Data format' is set to 'Delimited' with a comma as the 'Delimiter'. A 'Sample data' table is shown below, containing four rows of customer information:

	CREATEDATE	MODIFYDATE	FIRSTNAME	LASTNAME	BIRTH_DATE
0	1660096901	2022-08-09T22:01:41Z	Larina	Loveredge	1941-04-27
1	1660096901	2022-08-09T22:01:41Z	Sabina	Heindle	1970-07-29
2	1660096901	2022-08-09T22:01:41Z	Danika	Ruffey	1946-12-09
3	1660096901	2022-08-09T22:01:41Z	Mellicent	Fernyough	1946-08-04

Click Next to proceed to Data Prep Mapper step

The screenshot shows the 'Add data' workflow in Adobe Experience Platform. The left sidebar contains navigation links for Home, Workflows (selected), Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets, Queries). The main panel title is 'Sources > Data Landing Zone > Add data'. A progress bar at the top indicates steps: Select data (done), Dataflow detail (done), Mapping (not started), Scheduling (not started), and Review (not started). The 'Dataflow detail' section includes fields for 'Target dataset \*' (Existing dataset selected, showing 'Customer Account - YourNameHere'), 'Partial ingestion' (Enable partial ingestion checked, error threshold set to 5%), and 'Dataflow name \*' (Dataflow - Customer Account - YourNameHere). The 'Description' field is empty. The 'Alerts' section has a note: 'Check the options below to subscribe to different alerts.'

Data Prep Mapper step will load. Click on the arrow icon next to the calculated field populating the **person.birthDayAndMonth** XDM attribute.

The screenshot shows the 'Add data' workflow in the Adobe Experience Platform. The left sidebar includes 'Workflows' (selected), 'Sources', 'Destinations', 'Profiles', 'Segments', 'Identities', 'Schemas', 'Datasets', and 'Queries'. The main area is titled 'Sources > Data Landing Zone > Add data' and is at the 'Mapping' step. It displays 'Mapped fields: 24 of 24' and 'Required fields: 1 of 1'. An 'Identity fields' section shows '0 of 2'. A warning message 'Errors: 0' is present. The 'SOURCE DATA' section lists four fields: 'concat( date\_part(\"month\", date(birth\_Date))...', 'date\_part(\"year\", date(birth\_Date))', 'iif(sms\_optin == null, 'n', sms\_optin)', and 'lastName'. The 'TARGET FIELDS' section maps these to 'person.birthDayAndMonth', 'person.birthYear', 'consents.marketing.sms.val', and 'person.name.lastName'. A red box highlights the arrow icon next to the first source field.

Type in the following expression in the calculated field and click Preview. Data will appear as 2 digit month and 2 digit day

```
concat(  
    lpad(date_part("month", date(birth_Date)).toString(), 2, "0"),  
    "-",  
    lpad(date_part("day", date(birth_Date)).toString(), 2, "0"))
```

The screenshot shows the 'Add data' interface in the Adobe Experience Platform. The left sidebar navigation includes Home, Workflows (selected), Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring), DECISION MANAGEMENT (Offers, Components), and ADMINISTRATION (Alerts). The main area is titled 'Create calculated field' and shows a code editor with the following SQL-like expression:

```
concat(  
    lpad(date_part("month", date(birth_Date)).toString(), 2, "0"),  
    "-",  
    lpad(date_part("day", date(birth_Date)).toString(), 2, "0"))
```

Below the code editor is a 'Preview' section displaying five rows of output:

"04-27"
"07-29"
"12-09"
"08-06"
"11-11"

Click save. The mapping will look like this

The screenshot shows the 'Mapping schema' section of the Dataflow interface. On the left, under 'Source:', there is a code editor containing the query: `concat( lpad(date_part("month", date(birth_Date)).toString(), 2, '0') , date_part("year", birth_Date).toString() ) AS birthDayAndMonth`. To the right, under 'Mapping schema: Customer Account - YourNameHere', there is a 'TARGET FIELDS' section with a single entry: `person.birthDayAndMonth`.

Leave the schedule as 30 days and click next to continue.

The screenshot shows the 'Scheduling' step in the Dataflow workflow. The navigation bar at the top indicates the current step is 'Scheduling'. The configuration includes:

- Frequency: Day, Interval: 30
- Start time: 08/24/2022 9:46 PM
- Backfill: Off

At the bottom right of the scheduling form, there are 'Cancel', 'Back', and 'Next' buttons. The 'Next' button is highlighted in blue.

Click Next to review the dataflow and then finally save the dataflow by clicking Finish

The screenshot shows the 'Add data' workflow in the Adobe Experience Platform. The left sidebar contains navigation links for Home, Workflows, Dashboards, CONNECTIONS, Sources, Destinations, CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring). The main area is titled 'Sources > Data Landing Zone > Add data' and shows a progress bar with steps: Select data, Dataflow detail, Mapping, Scheduling, and Review. The 'Review' step is currently active. It displays three cards: 'Connection' (Account name: Customer Account - YourNameHere, Path: -, Columns: 23, Status: Connected), 'Assign dataset and map fields' (Target dataset: Customer Account - YourNameHere, Schema mapping: Dataset assigned), and 'Scheduling' (Start time: 08/24/2022 6:11 AM, Frequency: Minute, Interval: 15, Status: Scheduled). At the top right are 'Cancel', 'Back', and 'Finish' buttons.

**NOTE**

The changes we made to Data Prep step will be applicable only for the future runs. It will NOT affect the previous runs.

## Re-schedule the failed Dataflow

In the Dataflow page, click on the row to select the dataflow, when the right panel opens up, click **Edit schedule**.

The screenshot shows the Adobe Experience Platform Dataflows interface. On the left, there's a sidebar with navigation links like Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets). The main area is titled 'Sources' and has tabs for Catalog, Dataflows, Accounts, and System View. A filter bar says 'Source: Data Landing Zone'. Below it, a dataflow card is selected, highlighted with a red box. The card shows 'DATAFLOW NAME: Dataflow - Customer Account - YourNameHere...' and 'TARGET DATASET: Customer Account - Y...'. To the right of the card is a 'Properties' panel with options: 'Update dataflow' (with a pencil icon), 'Edit schedule' (which is also highlighted with a red box), 'Enabled' (with a toggle switch set to 'On'), and 'Delete' (with a trash can icon). Below the properties are fields for 'Dataflow name' (set to 'Dataflow - Customer Account - YourNam...') and 'Description'. At the bottom of the properties panel, it says 'Source data: dlz-user-container/project/PIPELINE/CUSTOMER/'.

Change the schedule such that Frequency is set to **Minute** and Interval is set to **15** as shown below

The screenshot shows the Adobe Experience Platform Dataflows interface. On the left, there's a sidebar with navigation links like Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets). The main area is titled 'Sources' and shows a catalog with a single entry: 'Dataflow - Customer Account'. A modal window titled 'Edit schedule' is open over the catalog. Inside the modal, under the 'Frequency' dropdown, 'Minute' is selected and 'Interval' is set to '15'. Below this, the 'Start time' is set to '08/24/2022 9:46 PM'. There's also a 'Backfill' toggle switch which is turned off. At the bottom of the modal are 'Cancel' and 'Save' buttons.



Let the instructor know you are done so far. Instructor will upload a new data file to your environment that will be automatically picked up by this dataflow in next run (after 15 minutes).

Click on the Dataflow name on the screen to open Dataflow runs

The screenshot shows the Adobe Experience Platform Dataflows interface. On the left, a sidebar navigation includes Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets). The Sources tab is selected under CONNECTIONS. The main area displays a 'Sources' catalog with tabs for Catalog, Dataflows, Accounts, and System View. A search bar at the top of the catalog section shows 'Source: Data Landing Zone'. Below it, a table lists a single dataflow entry:

DATAFLOW NAME	TARGET DATASET
Dataflow - Customer Account - YourNameHe...	Customer Account - Y...

A red box highlights the 'Dataflow - Customer Account - YourNameHe...' entry. To the right, a 'Properties' panel is open with the following settings:

- Update dataflow
- Edit schedule
- Enabled (checkbox checked)
- Delete

Below the properties, there are fields for Dataflow name (set to 'Dataflow - Customer Account - YourNam...'), Description (empty), and Source data (set to 'dlz-user-container/project/PIPELINE/CUSTOMER/'). Navigation arrows at the bottom of the catalog table allow for 'Next' and 'Previous' steps.

You will see some Dataflow runs with Success status and NO records ingested as shown below.

The screenshot shows the Adobe Experience Platform Dataflow activity page. The left sidebar has a 'Sources' section selected. The main area displays data flow activity with a table of runs. The table columns are: DATAFLOW RUN START, PROCESSING TIME, RECORDS INGESTED, RECORDS FAILED, and STATUS. The STATUS column uses green dots for Success and red dots for Failed. The first three rows are highlighted with a red box and show 'Success' status. The fourth row shows 'Failed' status. The properties panel on the right shows the data flow name as 'Dataflow - Customer Account - YourNameHere' and the target dataset as 'Customer Account - YourNameHere'. The source data is listed as 'diz-user-container/project/PIPELINE/CUSTOMER/'. The status bar at the bottom indicates 'Last 7 days'.

DATAFLOW RUN START	PROCESSING TIME	RECORDS INGESTED	RECORDS FAILED	STATUS
08/24/2022, 3:33 PM	1 minute	-	-	Success
08/24/2022, 3:33 PM	1 minute	-	-	Success
08/24/2022, 3:33 PM	1 minute	-	-	Success
08/24/2022, 2:46 PM	3 minutes	-	-	Failed

After the instructor uploads a new datafile, you will notice a successful run that ingested the data. You should notice 20 records ingested. Click on the date timestamp for this Dataflow run

The screenshot shows the Adobe Experience Platform Dataflow activity page. On the left, there's a sidebar with navigation links like Home, Workflows, Dashboards, CONNECTIONS (Sources selected), Destinations, CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring), and DECISION MANAGEMENT (Offers, Components). The main content area displays a summary card with 'Records ingested' (20) and 'Records failed' (0). Below this is a table of dataflow runs:

DATAFLOW RUN START	PROCESSING TIME	RECORDS INGESTED	RECORDS FAILED	STATUS
08/24/2022, 3:46 PM	1 minute	20	0	Success
08/24/2022, 3:33 PM	1 minute	-	-	Success
08/24/2022, 3:33 PM	1 minute	-	-	Success
08/24/2022, 3:33 PM	1 minute	-	-	Success
08/24/2022, 2:46 PM	3 minutes	-	-	Failed

A red box highlights the first run (08/24/2022, 3:46 PM). To the right is a 'Properties' panel with options like Update dataflow, Edit schedule, Enabled (checkbox), and Delete. It also shows the Dataflow name (Dataflow - Customer Account - YourNameHere), Description, Source data (dlz-user-container/project/PIPELINE/CUSTOMER/), Target dataset (Customer Account - YourNameHere), Source (Data Landing Zone icon), and Status (Enabled checkbox).

We will notice that 20 records are now ingested into the Dataset.

The screenshot shows the Dataflow Run overview page. On the left sidebar, under 'Sources', the 'Records ingested' count is highlighted with a red box and contains the value '20'. The main panel displays various metrics: Records failed (0), Total files (1), Size of data (5.09 kB). It also shows the status as 'Success' with a green dot, and the run times from start to end. Below this, the 'Dataflow run errors' section is shown, with a red arrow pointing to the 'Files' tab. Under 'Files', the file 'Lab\_Customer\_002.csv' is listed.

Click on the Files tab and confirm the Lab\_Customer\_002.csv got loaded.

The screenshot shows the same Dataflow Run overview page as before, but with a red arrow pointing down to the 'Files' tab in the 'Dataflow run errors' section. The 'Files' tab is selected, and the file 'Lab\_Customer\_002.csv' is listed in the table below.

The lab is now complete.

## Learnings

In the second part of the lab, we learned:

1. Editing an existing Dataflow to make corrections to transformation logic
2. Change the schedule of an existing Dataflow
3. Check the files loaded in a given Dataflow run using Monitoring console.



# **Adobe Experience Platform**

## **Lab 3 - Data Pipeline Historical**

Adobe Experience Platform Bootcamp Deep Dive Edition

Name \_\_\_\_\_  
Sandbox \_\_\_\_\_

## Lab Overview

Ingest Orders historical data into AEP in batch mode using Data Landing Zone (this lab) and stream orders live data using Streaming Source HTTP API (next lab).

This lab will introduce some complex data transformations and JSON data processing

**Expected time: 60 minutes**

## Learning Objectives

What should you walk away with after taking this Lab?

- Understand ingestion of JSON data
- Use Data Landing Zone as a source
- Use Data Prep to Map the non-XDM data to XDM
- Scheduling batch workflows

## Lab Tasks – Orders – Historical data

In this exercise, we will load the Orders data from Data Landing Zone to AEP Data Lake and Profile.

### Pre-requisites

1. Orders JSON file uploaded in the Azure ADLS Directory
2. Orders Schema and Dataset are already created

### Steps

Go to Adobe Experience Platform à **Sources** à **Catalog** à **Cloud storage**. Click on **Setup / Add Data** for the Data Landing Zone.

**Tip**

If at least one connection exists for that source, you will see “**Add data**” as the default action. If no connections exist for that source, you will see “**Setup**” as the default action.

## Select Source data

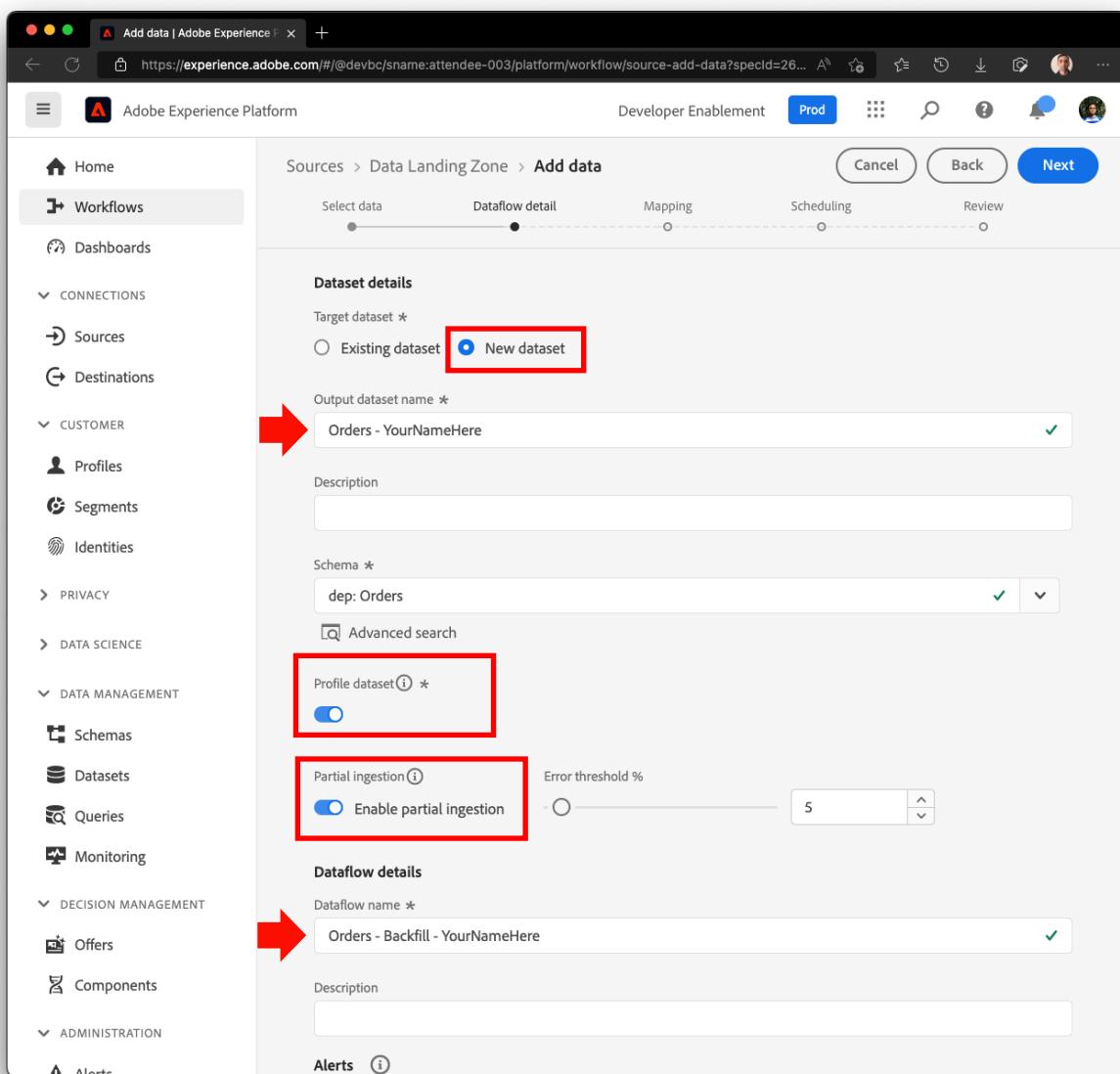
In the **Select data** screen, navigate to **dlz-user-container** → **project** → **Pipeline** and choose **Lab\_Historical\_Orders.json**. On the right-hand side, choose data format as **JSON**. Preview of the selected file is automatically displayed. Click **Next**.

SHIPPINGCITY	PRODUCTS.QUANTITY	PRODUCTS.PROD
Saint Petersburg	1	PRODUCT-17
Pittsburgh	1	PRODUCT-17
Jacksonville	1	PRODUCT-5
Norfolk	1	PRODUCT-13
Kansas City	1	PRODUCT-5
Boca Raton	1	PRODUCT-10

## Define the Target dataset

In the **Dataset details** step, choose **New dataset** option. Name the output dataset as **Orders - YourNameHere**. Set the schemas as **dep: Orders**.

Turn ON the **Profile dataset** option. Enable **Partial ingestion**. Name the Dataflow as **Dataflow - Orders – Backfill**.



### Tip

When Partial Ingestion is enabled, Error diagnostics are automatically enabled and hence the toggle box will disappear.

## Data Prep / Transformation

In the Data Prep (Mapping) step, ML Recommendations will automatically map most attributes. However, you will also see several errors. The initial screen will look similar to below. `_id` and `timestamp` are never recommended or mapped by default for Experience Events.

The screenshot shows the 'Add data' workflow in the Adobe Experience Platform. The left sidebar shows 'Workflows' selected. The main area is titled 'Sources > Data Landing Zone > Add data'. The workflow steps are: 'Select data' (done), 'Dataflow detail' (done), 'Mapping' (in progress), 'Scheduling' (not started), and 'Review' (not started). The 'Mapping' step is active, showing a progress bar for 'Mapped fields' at 29 of 29, 'Required fields' at 1 of 3, and 'Identity fields' at 2 of 2. An error message box is displayed, stating: 'There was error(s) preparing mappings.' with a red exclamation mark icon. It lists required unmapped paths: '\_id, timestamp' and 'productListItems.SKU'. Below this, a table maps source data to target fields:

Source: Lab_Historical_Orders.json	Mapping schema: dep: Orders
SOURCE DATA	TARGET FIELDS
shippingCity	shipping.address.city
Products[*].quantity	productListItems[*].quantity
Products[*].productId	productListItems[*].id

Please ensure the mappings are accurate. The initial recommendations loaded may have the following **invalid** mappings

### 3. Remove the following mappings

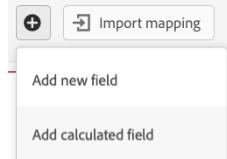
- Mapping to XDM attribute `billing.address.lastVerifiedDate`



## Create Calculated Fields

### Mapping to \_id

Create the following calculated fields by clicking on and click on Add Calculated Field.



Write the following expression and click Preview

```
concat(orderID, "-", lastOrderStatusUpdate)
```

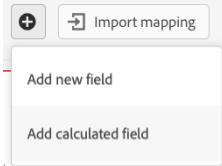
Calculated field will look similar to this. Click Save to save the calculated field.

Map the calculated field to \_id



### Mapping to order.\_devbc.acqSource

Create the following calculated field by clicking on and click on Add Calculated Field.



Write the following expression and click Preview. NOTE that this value is case sensitive and must be written exactly this way.

```
"inStore"
```

Calculated field will look similar to this. Click Save to save the calculated field.

A screenshot of the "Create calculated field" dialog in the Adobe Experience Platform. The left sidebar shows various platform components like Workflows, Sources, and Data Science. The main area has tabs for Function, Field, and Operator. Under the Function tab, there's a search bar and a list of functions: date, ua\_device\_class, sha256, upper, ua\_agent\_name, decode, uuid, get\_url\_port, to\_array, split, trim, and a partially visible function. To the right of the list is a "Preview" button, which is highlighted with a green checkmark. Below the preview is a "Preview" section showing three rows of the expression "inStore". At the bottom right of the dialog are "Save" and "Close" buttons.

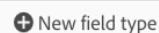
Map the calculated field to **order.\_devbc.acqSource**. The Data Prep screen will complain there is a duplicate mapping.

The screenshot shows the Adobe Experience Platform Workflows interface. On the left, there's a sidebar with various navigation options like Home, Workflows (which is selected), Dashboards, CONNECTIONS, Sources, Destinations, CUSTOMER, DATA SCIENCE, DATA MANAGEMENT, DECISION MANAGEMENT, and ADMINISTRATION. The main area is titled "Sources > Data Landing Zone > Add data". It's currently at the "Mapping" step of a five-step process. The progress bar shows "Mapped fields" (31 of 31) and "Required fields" (2 of 3). Below the progress bar, there's a search bar and a "Import mapping" button. A red-bordered box highlights an error message: "There was error(s) preparing mappings. The following unmapped paths are required: \_id, timestamp. The following unmapped paths are required: productListItems.SKU." The "TARGET FIELDS" section shows several mappings from source fields to target fields, including "inStore" to "order\_devbc.acqSource", "concat(orderID, \"-\", lastOrderStatusUpdate)" to "\_id", "shippingCity" to "shipping.address.city", "Products[\*].quantity" to "productListItems[\*].quantity", "Products[\*].productID" to "productListItems[\*]\_id", and "Products[\*].price" to "productListItems[\*].priceTotal".

Remove the other duplicate mapping for `order._devbc.acqSource` by clicking on the (-) symbol next to the row.

This screenshot shows a simplified view of the mapping interface. It has "SOURCE DATA" (containing "orderStatus") and "TARGET FIELDS" (containing "order.\_devbc.acqSource"). A red box highlights the minus (-) icon next to the "order.\_devbc.acqSource" entry in the target fields list, indicating where to click to remove it. A red warning message at the bottom states: "Duplicated target fields found. Map one row per target field."

## Add new pass-through mappings

 New field type

Add the following passthrough mappings by clicking New field type ( ) and Add new field for each row here. Some may already be present due to ML Recommendations.

Source Column	XDM Column
orderStatus	eventType
lastOrderStatusUpdate	timestamp
Products[*]	productListItems[*]
Products[*].productId	productListItems[*].SKU

NOTE that **Products[\*].productId** here is mapped to **productListItems[\*].SKU** in addition to **productListItems[\*].\_id**

## Product List Items

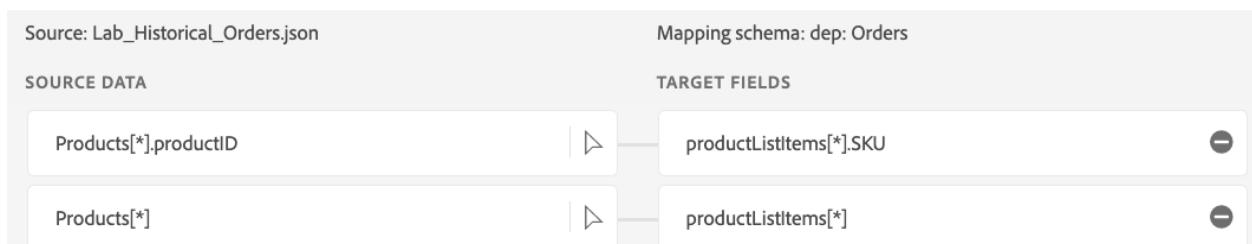
We will remove the mappings to XDM attributes **productListItems[\*].quantity** and **productListItems[\*].currencyCode**. Change the other mappings to **productListItems[\*]**

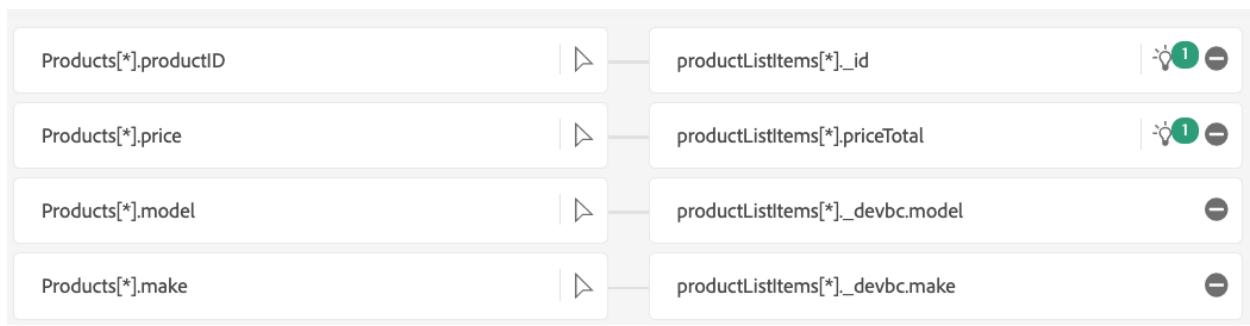
### NOTE

You will have to manually type in [\*] in the target attributes after making a selection for **Products** source object array or **productListItems** XDM attribute.

Source Column	XDM Column	Action
products[*]	productListItems[*]	Add
products[*].productId	productListItems[*].SKU	Add
products[*].productId	productListItems[*]._id	No change
products[*].make	productListItems[*]._devbc.make	Change
products[*].model	productListItems[*]._devbc.model	Change
products[*].price	productListItems[*].priceTotal	No change
products[*].quantity	productListItems[*].quantity	Remove
products[*].currencyCode	productListItems[*].currencyCode	Remove

The resultant mappings for **ProductListItems[\*]** should look like this.





## Final Mapping Set

The final mapping set should look like this:

#	Source Column	XDM Column	✓
1	orderStatus	eventType	
2	lastOrderStatusUpdate	timestamp	
3	orderID	order.orderID	
4	orderDate	order.orderDate	
5	orderTotal	order.priceTotal	
6	paymentType	order.payment.paymentType	
7	paymentAmount	order.payment.paymentAmount	
8	paymentCurrencyCode	order.payment.currencyCode	
9	paymentTransactionID	order.payment.transactionID	
10	plan.ID	order._devbc.plan.planID	
11	customerID	_devbc.customerID	
12	personalEmail	_devbc.personalEmail	
13	storeID	store.storeID	
14	shippingStreetAddress	shipping.address.street1	
15	shippingCity	shipping.address.city	
16	shippingState	shipping.address.state	
17	shippingZip	shipping.address.postalCode	
18	shippingMethod	shipping.shippingMethod	
19	shippingAmount	shipping.shippingAmount	
20	shippingDestination	shipping.shippingDestination	
21	billingStreetAddress	billing.address.street1	
22	billingCity	billing.address.city	

#	Source Column	XDM Column	✓
23	billingState	billing.address.state	
24	billingZip	billing.address.postalCode	
25	products[*]	productListItems[*]	
26	products[*].productID	<ul style="list-style-type: none"> <li>• productListItems[*]._id</li> <li>• productListItems[*].SKU</li> </ul>	
27	products[*].make	productListItems[*]._devbc.make	
28	products[*].model	productListItems[*]._devbc.model	
29	products[*].price	productListItems[*].priceTotal	
<b>Calculated Fields</b>			
30	concat(orderID, "-", lastOrderStatusUpdate)	_id	
31	"inStore"	order._devbc.acqSource	

## Preview the data

Preview the mapping output. Scroll through all the attributes to ensure there is no red exclamation next to any of the attributes on the right hand side. Preview will look similar to the below screenshot

_DXP.CUSTOMERID	_DXP.PERSONALEMAIL	PRODUCTLISTITEMS._DXP.MAKE	PRODUCTLISTITEMS
344633258	ibathow0@goo.ne.jp	Samsung	v1
892407845	fclapperton1@shutterfly.com	Samsung	v4
890611893	trentoll2@feedburner.com	Samsung	v2
729324930	fdarrow3@hud.gov	Google	v3
373432334	bdansie4@cbsnews.com	Apple	v4
687757641	mgumme5@mac.com	Google	v4
787455287	glaws6@psu.edu	Samsung	v3
565817034	dmcquillan7@reddit.com	Samsung	v1
975166720	kathersmith8@symantec.com	Google	v1
417156525	tmetterick9@a8.net	Google	v3

On the left hand side navigation of the Preview, select **productListItems** object array. Right hand side will update to show only the attributes in that object array. Notice that **productListItems.currencyCode** and **productListItems.quantity** are automatically populated (even after removing the mappings). This happens because **productListItems** as a parent object is mapped.

The screenshot shows the 'Preview' step in the 'Add data' workflow. On the left, a sidebar navigation includes 'Sources', 'Destinations', 'Customer', 'Data Science', 'Data Management', 'Decision Management', and 'Administration'. The main area displays a hierarchical tree of source fields under 'dep: Orders' and a table of target fields with their values.

PRODUCTLISTITEMS.CURRENCYCODE	PRODUCTLISTITEMS.PRICETOTAL	PRODUCTLISTITEMS.QUANTITY	PI
USD	402.23	1	-
USD	150.23	1	-
USD	150.23	1	-
USD	402.23	1	-
USD	150.23	1	-
USD	402.23	1	-
USD	402.23	1	-
USD	150.23	1	-
USD	402.23	1	-
USD	402.23	1	-

Completed mapping will look similar to the following screenshot

The screenshot shows the 'Mapping' step in the 'Add data' workflow. The interface includes tabs for 'Select data', 'Dataflow detail', 'Mapping', 'Scheduling', and 'Review'. The 'Mapping' tab is active. It displays a progress bar for 'Mapped fields' (33 of 33) and 'Required fields' (3 of 3). An error message indicates 'Errors: 0'. The 'SOURCE DATA' section lists fields like 'inStore', 'Products', 'billingState', 'orderStatus', 'concat(orderID, "...", lastOrderStatusUpdate)', 'lastOrderStatusUpdate', and 'Products[\*].productID'. The 'TARGET FIELDS' section lists corresponding fields like 'order\_dxp.acqSource', 'productListItems', 'billing.address.state', 'eventType', '\_id', 'timestamp', and 'productListItems[\*].SKU'. A 'Mapping schema: dep: Orders' label is present.

## Schedule

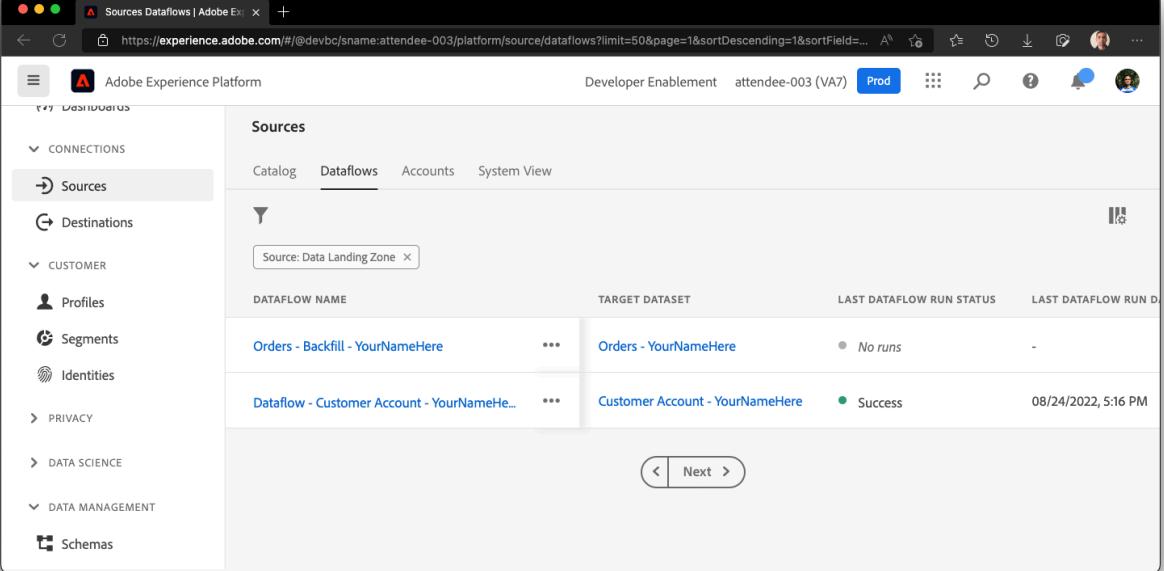
Set the schedule to run every 30 days by setting the Frequency as **Day** and Interval as **30**.

The screenshot shows the 'Add data' workflow interface in Adobe Experience Platform. The left sidebar contains navigation links for Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY (Policies), and Requests. The main panel title is 'Sources > Data Landing Zone > Add data'. A progress bar at the top indicates steps: Select data, Dataflow detail, Mapping, Scheduling, and Review. The 'Scheduling' step is active. The configuration shows 'Frequency \*' set to 'Day' and 'Interval' set to '30'. The 'Start time \*' is set to '08/04/2022 2:11 AM'. A note says 'All dates and times are in UTC.' There is also a 'Backfill' toggle switch.

Review the flow and click Finish

The screenshot shows the 'Add data' workflow interface in Adobe Experience Platform, specifically the 'Review' step. The left sidebar is identical to the previous screenshot. The main panel title is 'Sources > Data Landing Zone > Add data'. The progress bar shows the 'Review' step is completed. Three review cards are displayed: 'Connection' (Account name: Lab\_HistoricA\_L\_Orders.json, Source platform: Data Landing Zone, Path: Lab\_HistoricA\_L\_Orders.json) with status 'Connected'; 'Assign dataset and map files' (Target dataset: Orders - YourNameHere, Schema mapping: dep: Orders) with status 'Dataset assigned'; and 'Scheduling' (Start time: 08/25/2022 12:23 AM, Frequency: Day, Interval: 30) with status 'Scheduled'. Each card has a green checkmark icon indicating success.

Backfill / Historical load Dataflow will be created. Dataflow execution will not start immediately and will take few minutes. So, last Dataflow Run Status will be set to "No runs".

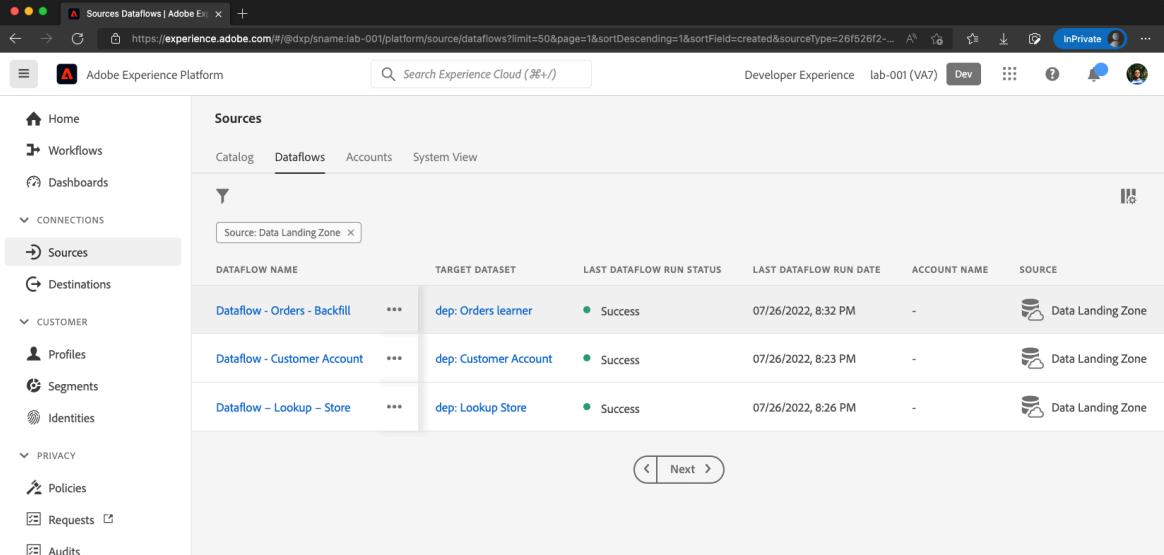


The screenshot shows the Adobe Experience Platform interface for managing dataflows. The left sidebar is titled 'Sources' and includes sections for CONNECTIONS, CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT. The main content area is titled 'Sources' and has tabs for Catalog, Dataflows, Accounts, and System View. A search bar at the top right says 'Source: Data Landing Zone'. Below it, there's a table with columns: DATAFLOW NAME, TARGET DATASET, LAST DATAFLOW RUN STATUS, and LAST DATAFLOW RUN DATE. Two rows are visible:

DATAFLOW NAME	TARGET DATASET	LAST DATAFLOW RUN STATUS	LAST DATAFLOW RUN DATE
Orders - Backfill - YourNameHere	Orders - YourNameHere	No runs	-
Dataflow - Customer Account - YourNameHe...	Customer Account - YourNameHere	Success	08/24/2022, 5:16 PM

## Check scheduled execution

After few minutes, the Dataflow will succeed. Notice the **Last Dataflow Run Status** and **Last Dataflow Run Date**.



This screenshot shows the same interface after some time. The table now displays three successful dataflows:

DATAFLOW NAME	TARGET DATASET	LAST DATAFLOW RUN STATUS	LAST DATAFLOW RUN DATE	ACCOUNT NAME	SOURCE
Dataflow - Orders - Backfill	dep: Orders learner	Success	07/26/2022, 8:32 PM	-	Data Landing Zone
Dataflow - Customer Account	dep: Customer Account	Success	07/26/2022, 8:23 PM	-	Data Landing Zone
Dataflow - Lookup - Store	dep: Lookup Store	Success	07/26/2022, 8:26 PM	-	Data Landing Zone

Click on the Dataflow name to get a list of Dataflow Runs. 10 Records should be ingested.

Records ingested: 10

Records failed: 0

DATAFLOW RUN START	PROCESSING TIME	RECORDS INGESTED	RECORDS FAILED	STATUS
07/26/2022, 8:32 PM	2 minutes	10	0	Success

Dataflow name: Dataflow - Orders - Backfill

Description:

Source data: dlz-user-container/project/JSON-

Click on the Dataflow Run Start time to see error diagnostic details.

Records ingested: 10

Records failed: 0

Total files: 1

Size of data: 10.39 kB

Status: Success

Dataflow run start: 07/26/2022, 8:32 PM

Dataflow run end: 07/26/2022, 8:34 PM

Dataflow run IDMS org ID: 37E0399C61687C4E0A495E06@A94354d08-682f-4fb5-be6f-0a2ea5232d4f

Dataset: dep: Orders learner

## Verify the data

In the Left Nav bar, Go to Datasets in Platform and click on Orders - YourNameHere

The screenshot shows the Adobe Experience Platform Datasets Browse interface. On the left, there is a navigation sidebar with various categories like Home, Workflows, Dashboards, Sources, Destinations, Customer, Privacy, Policies, Requests, Audits, Data Science, Services, Data Management, Schemas, and Datasets. The 'Datasets' option is selected. The main area has tabs for 'Overview' and 'Browse'. Under 'Browse', there is a search bar and a table listing datasets. The columns in the table are NAME, CREATED, SOURCE, SCHEMA, and LAST BATCH STATUS. The table contains the following data:

NAME	CREATED	SOURCE	SCHEMA	LAST BATCH STATUS
dep: Orders learner	07/26/2022, 8:31 PM	Schema	dep: Orders	Success
dep: Customer Account	07/26/2022, 5:11 PM	Schema	dep: Customer Account	Success
dep: Lookup Store	07/26/2022, 4:45 PM	Schema	dep: Lookup Store	Success
dep: Orders	07/26/2022, 4:40 PM	Schema	dep: Orders	No batches
dep: Lookup Plan	07/19/2022, 1:38 PM	Schema	dep: Lookup Plan	No batches
dep: Lookup Product	07/19/2022, 1:37 PM	Schema	dep: Lookup Product	No batches
dep: Web	07/19/2022, 1:37 PM	Schema	dep: Web	No batches
dep: Ecommerce	07/19/2022, 1:37 PM	Schema	dep: Ecommerce	No batches
dep: Billing	07/19/2022, 1:37 PM	Schema	dep: Billing	No batches
dep: Customer Aggregates	07/19/2022, 1:36 PM	Schema	dep: Customer Aggregates	No batches

To the right of the table, there is a summary section with the number '29' and the word 'Datasets'. Below this, there is a section titled 'Most recently updated' with the following list:

- dep: Orders learner 07/26/2022
- dep: Customer Account 07/26/2022
- dep: Lookup Store 07/26/2022
- profile\_dim\_date 07/26/2022
- Segmentdefinition-Sn... 07/26/2022

## Click on the Preview Dataset.

The screenshot shows the 'Dataset activity' page for the dataset 'dep: Orders learner'. The left sidebar includes sections for Home, Workflows, Dashboards, Connections, Sources, Destinations, Customer, Privacy, Policies, Requests, Audits, Data Science, Services, Data Management, Schemas, Datasets (selected), Queries, and Monitoring. The main content area displays dataset activity metrics for the last 7 days. Key statistics include: Total records in previous month (10), Ingested records in the last 7 days (0), Ingested batches today (1), Failed batches today (0), Size of data in previous month (58.75 kB), and Failed batches in the last 7 days (0). A chart titled 'Ingested records' shows a single data point for July 20, 2022. A legend indicates blue for 'Ingested batches' and red for 'Failed batches'. On the right, the dataset details are shown: Name (dep: Orders learner), Description (empty), Dataset ID (62e0b18f7303841c072d31ae), Table name (dep\_orders\_learner), Profile (empty), Schema (dep: Orders), Source (empty), and Created (07/26/2022, 8:31 PM).

Notice that `productListItems.currencyCode` and `productListItems.quantity` are auto populated.

The screenshot shows the 'Preview dataset' interface for the 'dep: Orders learner' dataset. The left sidebar is identical to the previous screenshot. The main area features a search bar and a tree view of the dataset schema under the heading 'dep: Orders'. The schema structure is as follows: dep: Orders > \_dpx (Object) > billing (Object) > order (Object) > productListItems (Object[]) > shipping (Object) > store (Object) > \_id (String) > eventTime (String). To the right, a table lists data rows for the 'productListItems' field. The columns are ORDER.PAYMENT.CURRENCYCODE and ORDER.ORDERID. The data includes:

ORDER.PAYMENT.CURRENCYCODE	ORDER.ORDERID
USD	b438b27f-6cf5-4388-a147-4b9c14ac2aa
USD	724d1bf3-ccd5-4942-b149-1e49fa71015b
USD	478d328b-0e24-4b34-b25b-7649af797fc
USD	d6b889f8-26c4-47ea-936c-230acb469778
USD	bfb366a-a4f1-4099-80d3-da12ddf034cf
USD	d427fa96-46b2-46fc-a4c7-5a211343b1bd
USD	76cb8abe-e2f6-46b4-a6f3-d23e4d6ab7ee
USD	4ee8c8db-20d6-46a4-af26-94cffc83b30
USD	536e0e5a-d2fb-4fe4-adf8-7cdae4ced484

At the bottom right, the timestamp is 07/26/2022, 8:31 PM.



# **Adobe Experience Platform**

## **Lab 4 - Data Pipeline Streaming**

Adobe Experience Platform Bootcamp Deep Dive Edition

Name \_\_\_\_\_  
Sandbox \_\_\_\_\_

## Lab Overview

In this lab, we will use Import Mapping capability to import Customer Accounts and Orders historical data mapping into streaming ingestion flow. Data will be ingested using Streaming HTTP API.

**Expected time: 30 minutes**

### Learning Objectives

What should you walk away with after taking this Lab?

- Creating a Streaming Inlet
- Importing mapping set from another Dataflow
- Get Dataflow ID, and Dataset ID from the UI
- Using REST API to ingest an event

# Lab Tasks – Customer Accounts – Live data (Streaming)

In this exercise, we will load the Customer Account data from Streaming source to AEP Data Lake and Profile.

## Pre-requisites

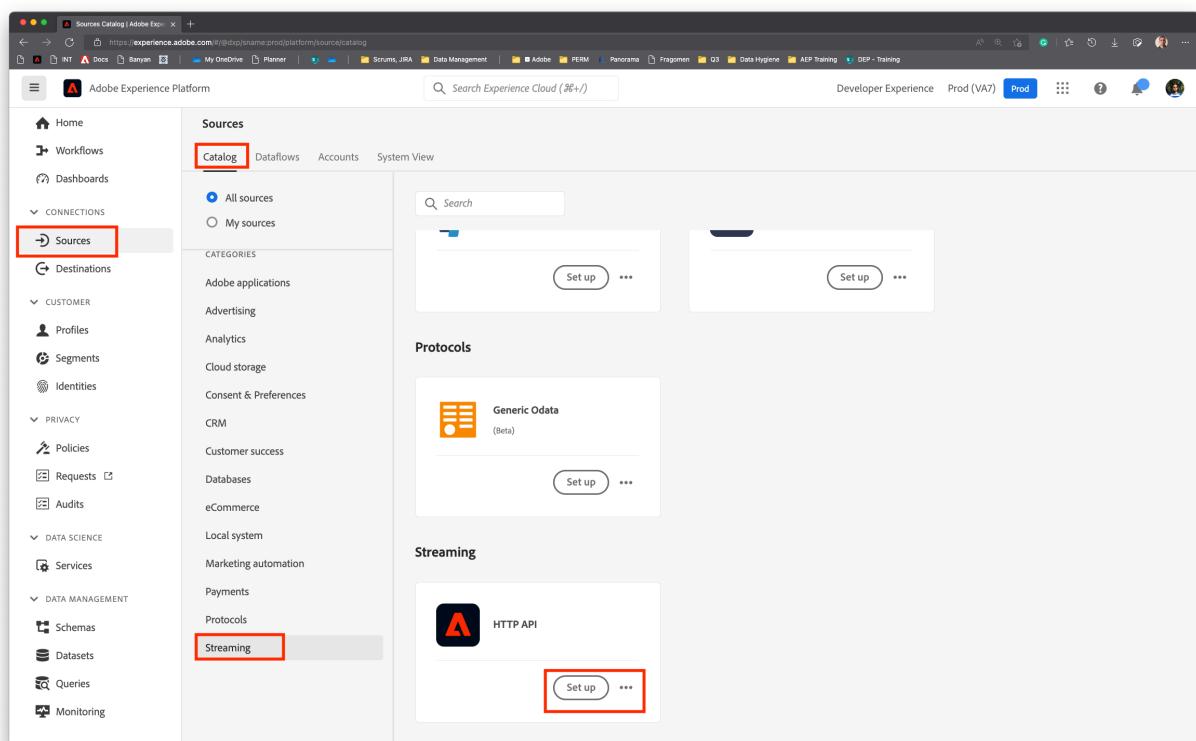
1. Lab 2 is successfully completed

## Steps

Go to Adobe Experience Platform → Sources → Catalog → Streaming. Click on **Setup / Add Data** for the HTTP API.

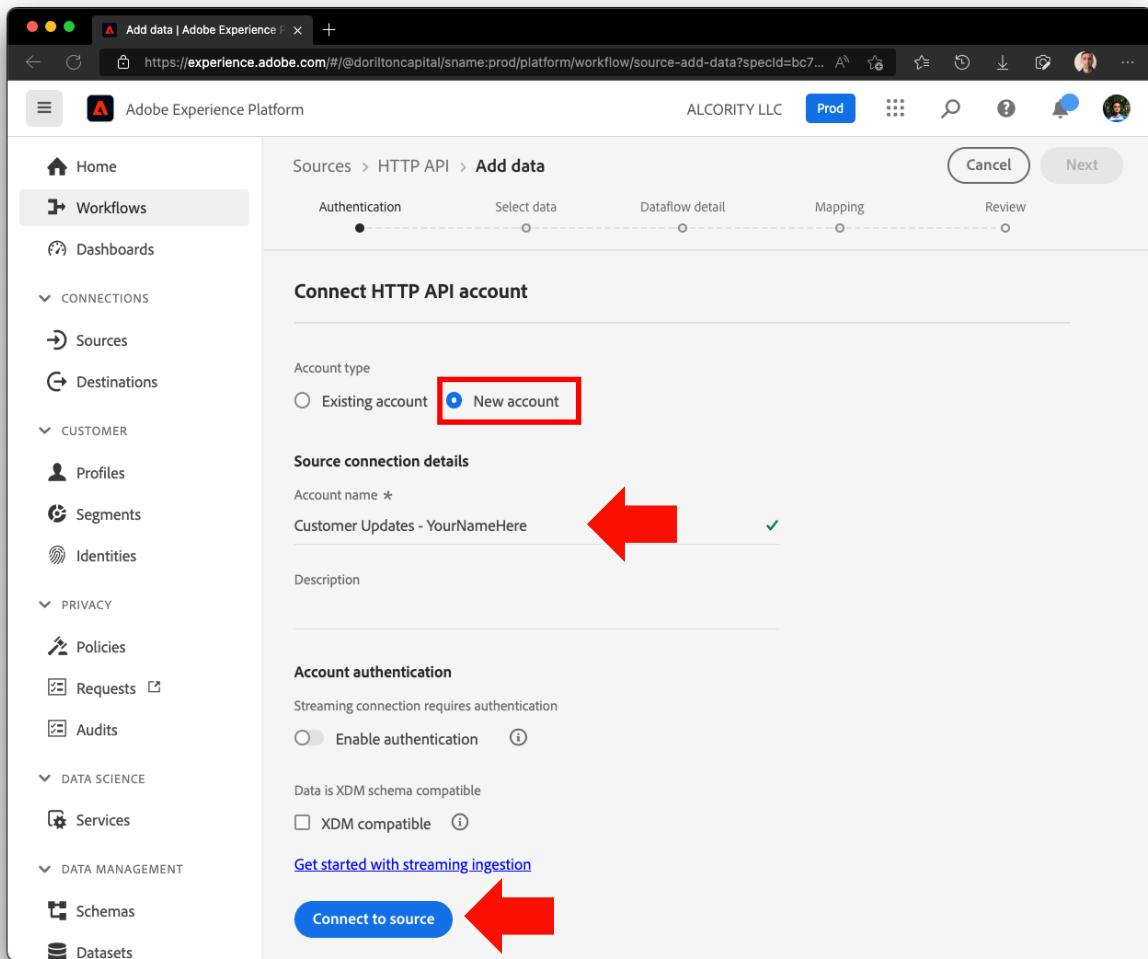
### Tip

If at least one connection exists for that source, you will see “Add data” as the default action. If no connections exist for that source, you will see “Setup” as the default action.

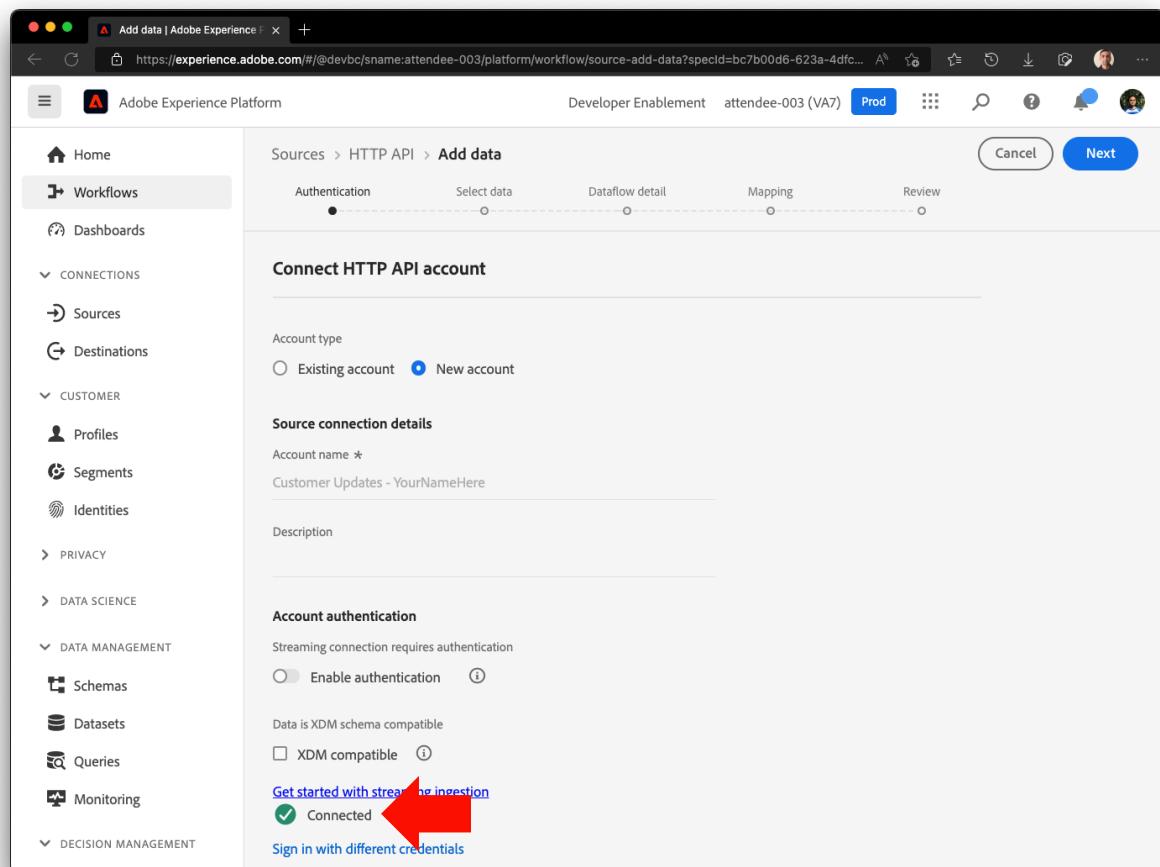


## Creating a Streaming Inlet

In the Connect HTTP API Account screen, Choose **New account** and name the account as **Customer Updates – YourNameHere**. Enable authentication is turned OFF by default. XDM compatible checkbox is OFF too. We will leave these defaults as-is. Then click **Connect to source**.

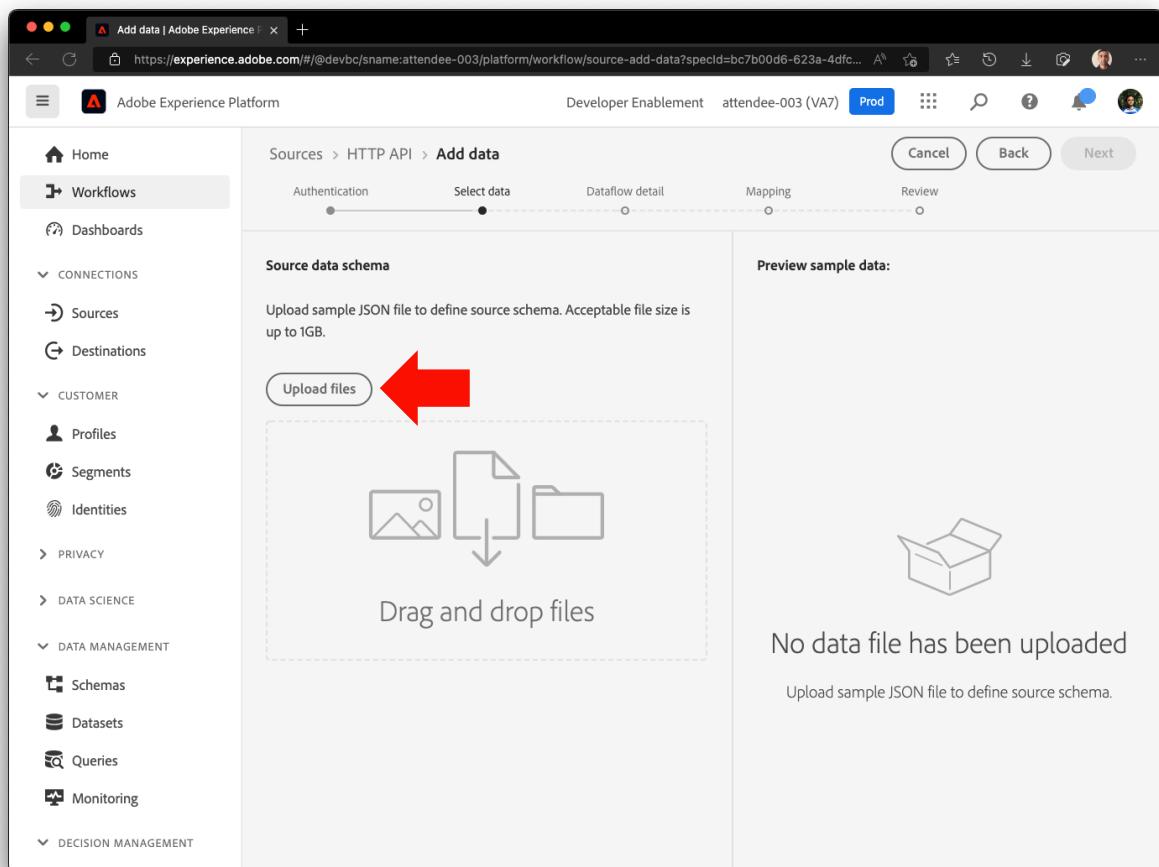


A Streaming account / endpoint is created for you, and you will see message called “Connected” on the screen. Click next to continue.



## UPLOAD SAMPLE DATA

In the Source data schema screen, upload the JSON file **Lab\_Single\_Customer.json** from your local file system to the AEP



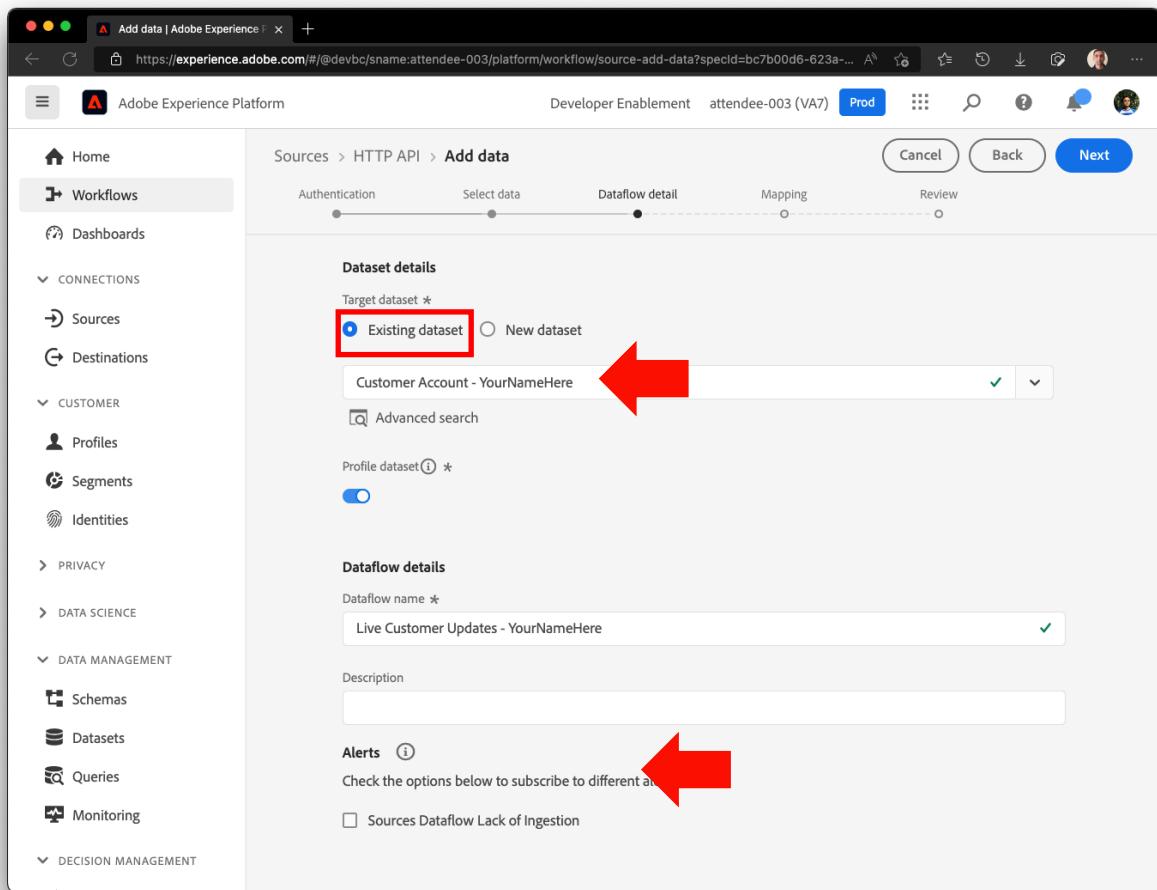
Once the file is uploaded, preview will appear as follows. Click Next to continue.

The screenshot shows the 'Add data' interface in the Adobe Experience Platform. The left sidebar navigation includes Home, Workflows (selected), Dashboards, CONNECTIONS, Sources, Destinations, CUSTOMER, Profiles, Segments, Identities, PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring), and DECISION MANAGEMENT. The main panel title is 'Sources > HTTP API > Add data'. The workflow steps are Authentication (done), Select data (done), Dataflow detail (not started), Mapping (not started), and Review (not started). A 'Source data schema' section prompts to upload a sample JSON file (maximum 1GB). A preview window titled 'Preview sample data:' shows the contents of 'Lab\_Single\_Customer.json'. The schema is defined as:

```
SampleSchema:a7bb255d430c401c921cc34da2b61774
account_create_date | String | 2022-04-20T22:19:03Z
account_end_date | String | 2022-01-20T13:15:32Z
billing_city | String | Carol Stream | Value
billing_state | String | IL | Value
billing_street_address | String | 71424 Messerschmidt C
billing_zip_code | String | 60351 | Value
birth_Date | String | 1940-01-15 | Value
createDate | String | 1660096899 | Value
```

Buttons at the top right include Cancel, Back, and Next (highlighted in blue).

In the Dataset details screen, choose **Existing dataset** and select the dataset you created in Lab 2. It was named **Customer Account – YourNameHere**. Profile dataset is toggled ON by default. Name the Dataflow as Live **Customer Updates – YourNameHere**.



## Import the Customer Accounts Mapping

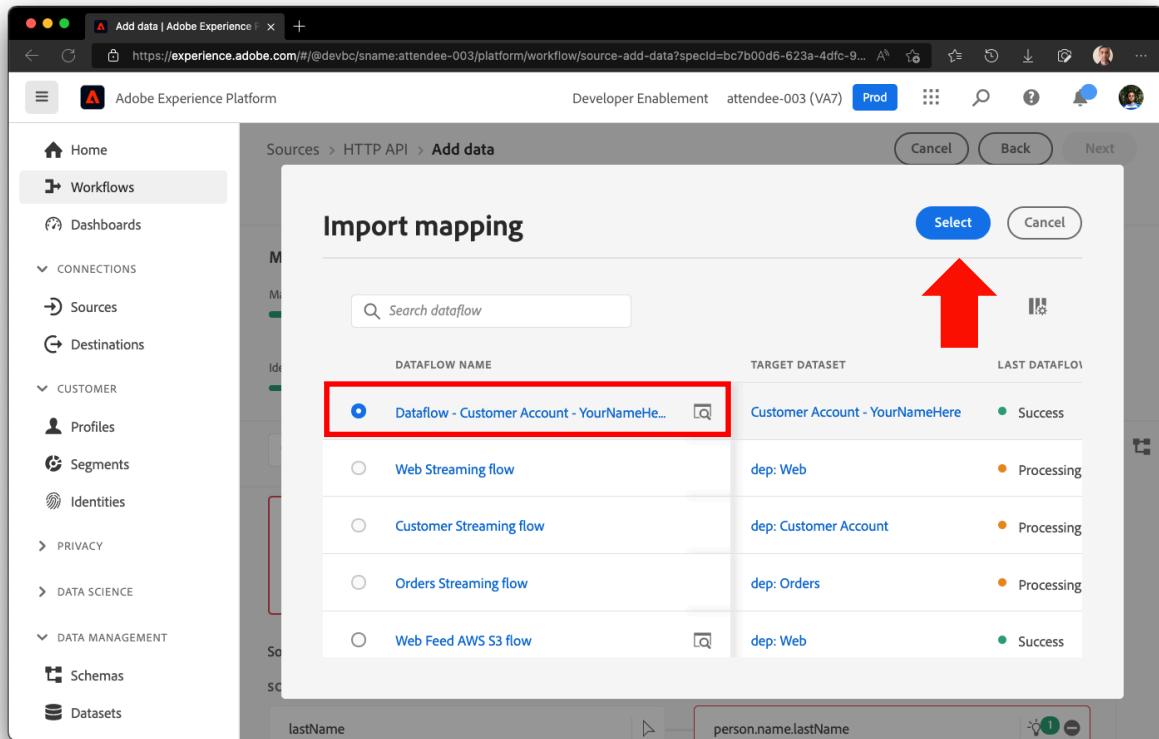
Mapping screen will load. ML recommendations will pre-populate some attributes. Some errors may also appear. Click on the Import Mapping button ( Import mapping) to import the mapping we created in Lab 2 for batch customer accounts ingestion.

The screenshot shows the 'Add data' workflow in the Adobe Experience Platform. The current step is 'Mapping'. A red arrow points to the 'Import mapping' button, which is highlighted with a red border. Below the button, a red box contains an error message: 'There was error(s) preparing mappings.' followed by 'There is a duplicate mapping for the target path person.name.lastName. The data at the XDM path will be overwritten. Please specify a distinct destination path.' The 'SOURCE DATA' section lists fields: 'lastName', 'email\_optin', 'billing\_street\_address', and 'modifyDate'. The 'TARGET FIELDS' section shows mappings: 'lastName' to 'person.name.lastName', 'email\_optin' to 'consents.marketing.email.val', 'billing\_street\_address' to 'billingAddress.street1', and 'modifyDate' to '\_repo.modifyDate'. Each mapping has a red exclamation mark icon next to it.

Choose the Dataflow created in the Lab 2. It was named **Dataflow – Customer Account – YourNameHere**.

**Warning**

If you have not completed the Lab 2 successfully, you will NOT be able to proceed further. Additionally, if you have not updated the Lab 2 dataflow to change the transformation logic for **person.birthDayAndMonth** XDM attribute, this lab will fail.



Mapping is now imported. Ensure the calculated field mapped to **person.birthDayAndMonth** has "lpad" string in it. Click next to continue.

The screenshot shows the 'Add data' workflow interface in Adobe Experience Platform. The left sidebar includes 'Workflows' under 'Sources'. The main area displays a progress bar at the top: Authentication (green), Select data (green), Dataflow detail (green), Mapping (black dot), and Review (grey). Below the progress bar, the title is 'Map desired source fields to target dataset fields'. It shows 'Mapped fields: 24 of 24' and 'Required fields: 1 of 1' both in green. Under 'Identity fields', it shows '2 of 2' in green. A red warning icon indicates 'Errors: 0'. The 'SOURCE DATA' section contains three items: 'concat( lpad(date\_part("month", date(birth\_Date))...)', 'date\_part("year", date(birth\_Date))', and 'iif(sms\_optin == null, 'n', sms\_optin)'. The 'TARGET FIELDS' section maps these to 'person.birthDayAndMonth', 'person.birthYear', and 'consents.marketing.sms.val'. The 'Mapping schema' is listed as 'Customer Account - YourNameHere'.

Review the details and click Finish to save the Dataflow.

The screenshot shows the 'Add data' workflow in the Adobe Experience Platform. The left sidebar contains navigation links for Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring). The main area shows the 'Sources > HTTP API > Add data' flow, with the 'Review' step selected. A progress bar at the top indicates steps: Authentication (done), Select data (done), Dataflow detail (done), Mapping (done), and Review (in progress). The 'Review' section contains two panels: 'Connection' (Account name: Customer Updates - YourNameHere, Source platform: HTTP API) and 'Assign dataset and map fields' (Target dataset: Customer Account - YourNameHere, Schema mapping: Lab\_Single\_Customer.json). Both panels show a green checkmark indicating they are connected or assigned. At the bottom right of the review section are 'Cancel', 'Back', and 'Finish' buttons.

## Streaming Endpoint

In the screen that appears, gather the streaming endpoint, we will use it in the API.

### Tip

If you do not see this information, most likely the Dataflow is selected on the left. Click on the Dataflow row (not the blue link) on the left again to unselect it.

The screenshot shows the Adobe Experience Platform Sources interface. On the left sidebar, under 'CONNECTIONS', 'Sources' is selected. In the main content area, there is a card for an 'HTTP API' source named 'Customer Updates - YourNameHere'. The card displays the account name, dataset count, and last update date. Below this, a table lists a single dataflow named 'Live Customer Updates - YourNameHere' with a target dataset of 'Customer Account - YourNameHere'. A red box highlights the 'Streaming endpoint' section on the right, which contains the URL: <https://dc.s.adobedc.net/collection/518c27ca276addc9363601622f39553966500d912f9ce536321298ad96ad4ae1>.

## Get Dataflow ID and Dataset ID

In the current screen, select the dataflow name, scroll down the right hand panel and take a note of the Dataflow ID and Dataset ID from the **API Usage** section.

The screenshot shows the Adobe Experience Platform Foundations Sources interface. On the left sidebar, under the 'Connections' section, 'Sources' is selected. In the main content area, there is a card for an 'HTTP API' named 'Customer Updates - YourNameHere'. The card displays the account name, dataset count (1 dataset), and last update date (August 24, 2022). Below the card is a table with three columns: 'DATAFLOW NAME', 'TARGET DATASET', and 'LAST DATAFLOW RUN'. The first row shows 'Live Customer Updates - YourNameHere' as the dataflow name, 'Customer Account - YourNameHere' as the target dataset, and 'No runs' as the last dataflow run. A red box highlights this row. To the right of the table is a 'Properties' panel. At the bottom of the page is a navigation bar with 'Next >' and '< Previous' buttons.

DATAFLOW NAME	TARGET DATASET	LAST DATAFLOW RUN
Live Customer Updates - YourNameHere	Customer Account - YourNameHere	No runs

**Properties**

Status: Processing  
Partial ingestion: Not enabled  
Error diagnostics: Not enabled  
Created: 08/24/2022, 1:15 AM

**API Usage**

Dataflow ID: 00782e58-ddce-43a0-b6e9-99fe66ac9892  
Dataset ID: 6305c0b7cb26b31c0753fb20  
Mapping set ID: e1ae59ef5d814ef7ba6e6b35a73a7af9  
Mapping version: 0

## Place a REST API call

Switch to the Postman UI, and set the following

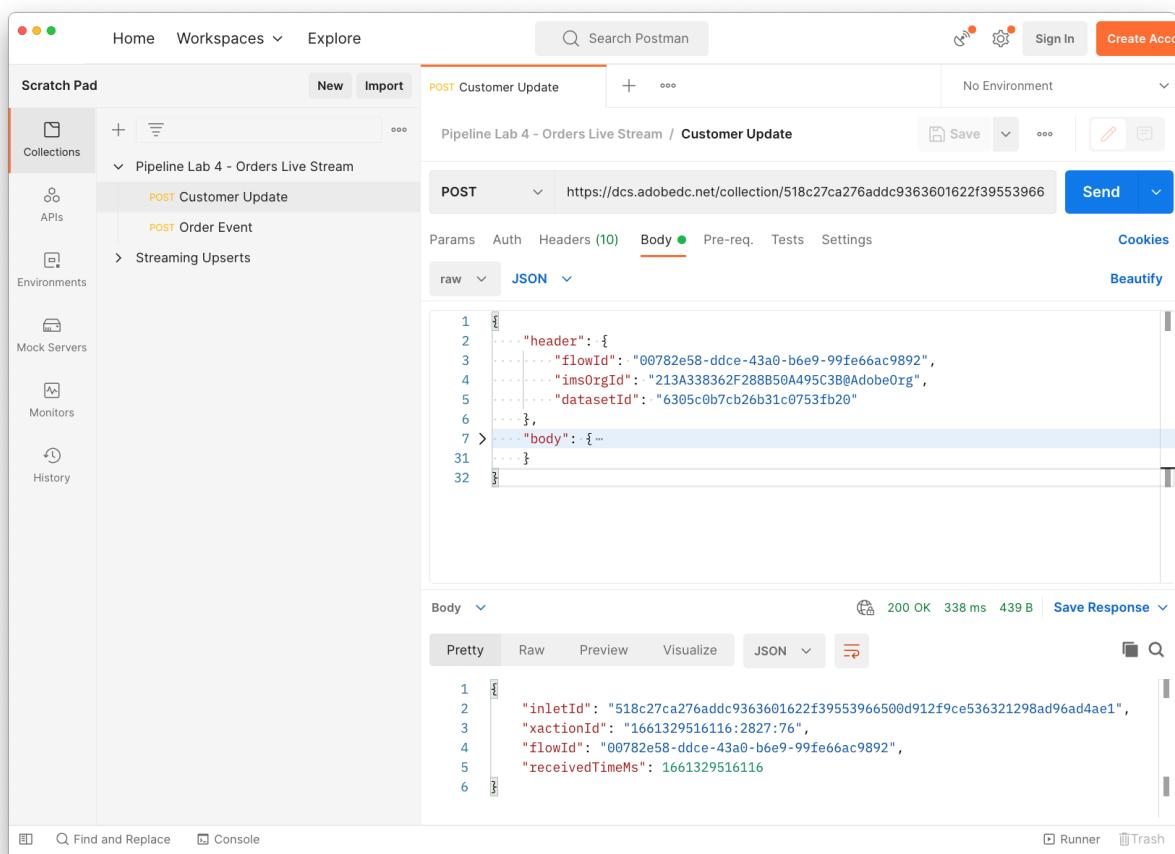
Type: **POST**

URL: « Streaming Endpoint » you captured in the previous step

Body: The below structure.

Replace your flowId, and datasetId.

```
{  
  "header": {  
    "flowId": "00782e58-ddce-43a0-b6e9-99fe66ac9892",  
    "imsOrgId": "213A338362F288B50A495C3B@AdobeOrg",  
    "datasetId": "6305c0b7cb26b31c0753fb20"  
  },  
  "body": {}  
}
```



Add the JSON request body as follows.

```

"body": {
    "customer_id": "202208240125",
    "firstName": "YourFirstName",
    "lastName": "YourLastName",
    "email": "yourEmail@adobe.com",
    "createDate": "1660096899",
    "modifyDate": "2022-08-09T22:01:39Z",
    "birth_Date": "1940-01-15",
    "mobile_phone": "309-115-7572",
    "email_optIn": "y",
    "sms_optIn": "n",
    "shipping_street_address": "71424 Messerschmidt Circle",
    "shipping_city": "Carol Stream",
    "shipping_state": "IL",
    "shipping_zip_code": "60351",
    "billing_street_address": "71424 Messerschmidt Circle",
    "billing_city": "Carol Stream",
    "billing_state": "IL",
    "billing_zip_code": "60351",
    "plan_id": "m1",
    "plan_name": "basic",
    "account_create_date": "Created on 2022-04-20T22:19:03Z",
    "account_end_date": "2022-01-20T13:15:32Z",
    "source": "inStore"
}

```

Change the following attributes in the JSON to make it unique:

Attribute Name	Type	Suggested values
customer_id	Number	Use your attendee number with date/timestamp to make it unique
firstName	String	Your First name
lastName	String	Your last name / sur name / initial
Email	Email	Must be a valid email in with proper email format.

**WARNING**

Make sure these attributes are unique. If the row is not unique, you may be overriding someone else's data or vice versa.

Click Send to ingest the data. You should receive a **200 OK** from the Platform.

The screenshot shows the Postman application interface. On the left, the 'Scratch Pad' sidebar lists collections, APIs, environments, mock servers, monitors, and history. In the center, a 'Customer Update' collection is selected under 'Pipeline Lab 4 - Orders Live Stream'. A POST request titled 'Customer Update' is shown with the URL <https://dcs.adobedc.net/collection/518c27ca276addc9363601622f39553966>. The 'Body' tab is selected, showing a JSON payload with various fields like header, body, and dataset ID. The 'Send' button is highlighted with a red arrow. Below the request, the response section shows a 200 OK status with a response time of 338 ms and a response size of 439 B. The response body is highlighted with a red box and contains the following JSON:

```
1 {  
2   "header": {  
3     "flowId": "00782e58-ddce-43a0-b6e9-99fe66ac9892",  
4     "imsOrgId": "213A338362F288B50A495C3B@AdobeOrg",  
5     "datasetId": "6305c0b7cb26b31c0753fb20"  
6   },  
7   "body": {  
8     "customer_id": "202208240125",  
9     "firstName": "YourFirstName",  
10    "lastName": "YourLastName",  
11    "email": "yourEmail@adobe.com",  
12    "createDate": "1660096899",  
13    "modifyDate": "2022-08-09T22:01:39Z",  
14    "birth_Date": "1940-01-15",  
15  }  
16 }
```

## 2. Verify the data in Profile

Go to **Profiles** in the Left Navbar and click on **Browse** tab.

The screenshot shows the Adobe Experience Platform Profiles Browse interface. On the left, there's a sidebar with various navigation options like Home, Workflows, Dashboards, and Customer profiles. Under Customer, 'Profiles' is selected and highlighted with a red arrow. At the top, the 'Browse' tab is active, also highlighted with a red arrow. The main area shows a search interface for profiles, including fields for 'Merge policy' (set to 'Default Timebased') and 'Identity namespace' (set to 'Email'). Below this, a large '426 Profiles' count is displayed. A central table is shown with columns: PROFILE ID, FIRST NAME, LAST NAME, and PERSONAL EMAIL. A placeholder message 'No profiles found.' is visible. To the right, there's a section titled 'Profiles by namespace'.

In the Browse tab, Leave the Merge policy to be the **Default Timebased** and select identity namespace as **Email**. Identity value text box appears. Type in your email (the one you used in the REST API Payload such as **yourEmail@adobe.com**) and press **View**.

A close-up view of the search filters at the top of the Profiles Browse page. The 'Merge policy' dropdown is set to 'Default Timebased'. The 'Identity namespace' dropdown is set to 'Email'. The 'Identity value' input field contains 'yourEmail@adobe.com'. A red arrow points to the 'View' button, which is highlighted in blue.

A profile row will appear below. The profile row will have the first name, last name and email you supplied in the REST API call. Click on the Profile ID to open the Profile.

The screenshot shows the results of the profile search. A table has four columns: PROFILE ID, FIRST NAME, LAST NAME, and PERSONAL EMAIL. The first row shows a profile with the ID 'BVrqzwVv7sqLqxJmopfmnaG3v3KJgA', first name 'YourFirstName', last name 'YourLastName', and email 'yourEmail@adobe.com'. A red arrow points to the 'BVrqzwVv7sqLqxJmopfmnaG3v3KJgA' profile ID.

PROFILE ID	FIRST NAME	LAST NAME	PERSONAL EMAIL
BVrqzwVv7sqLqxJmopfmnaG3v3KJgA	YourFirstName	YourLastName	yourEmail@adobe.com

Notice the **person.birthDayAndMonth** populated in the Profile.

The screenshot shows the Adobe Experience Platform Profile Detail interface. On the left, there's a navigation sidebar with sections like Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring). The main area shows a profile named 'Profile BVrqzwVv7sqLqxJmopfmnaG3v3KJgA'. The 'Basic attributes' section contains fields for Address, Gender, Birth Day And Month (which is highlighted with a red box and shows '01-15'), Number, and Street 1. Other sections include Customer profile, Linked identities, and Channel preferences (Direct mail, Phone, SMS, Email, Facebook news f..., Twitter feed).

## Learnings

1. We can always turn a batch ingestion flow in to Streaming flow by creating a new streaming connection and reusing the Mapping step from the batch flow
2. The fix we applied for **person.birthDayAndMonth** in Lab 2 is successful
3. When an update or new customer row arrived in the Streaming ingestion flow, it was immediately transformed into XDM and passed on to Profile. This demonstrates the Real-time nature of profile.

## Additional exercises

### Exercise #1

Try executing the same REST API call with a slight name to the first name / last name. You will notice that the changes reflect in Profile in matter of minutes. Do NOT change the **Customer ID** or **email** attributes as they are identities and profile relies on them. It may take up to 15 minutes for the Profile values to be updated

Overview    **Browse**    Merge Policies    Union Schema

Browse profiles by merge policy and filter by an identity value (optional).

Merge policy (i) \*    Identity namespace (i)    Identity value (i)

Default Timebased    X    Email    X    yourEmail@adobe.com    View    Clear

PROFILE ID	FIRST NAME	LAST NAME	PERSONAL EMAIL
BVrqzwVv7sqLqxJmopfmnaG3v3KJgA	YourFirstName	Updated Last Name	yourEmail@adobe.com

## Exercise #2

In the Attributes tab of the Profile, you will notice that only account end date is populated. But account create date is NOT populated. This is because we never fixed the issue with account create date in Data Prep Mapper step. Remember, that the account create date contained values such as "**Created on 2022-04-20T22:19:03Z**" whereas accepted values are similar to "**2022-04-20T22:19:03Z**"

Profile Attributes | Adobe Experience Platform

https://experience.adobe.com/#/devbc/sname:attendee-003/platform/profile/browse/BVrqzwVv7sqLqxJmopfmnaG3v3KJgA...

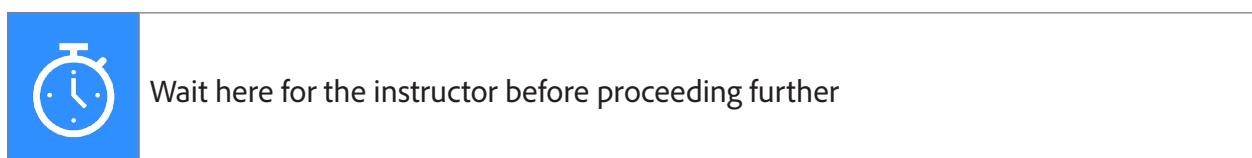
Adobe Experience Platform    Developer Enablement attendee-003 (VA7) Prod

Profiles > Profile BVrqzwVv7sqLqxJmopfmnaG3v3KJgA

Detail    **Attributes**    Events    Segment membership

Profile using merge policy Default Timebased

ATTRIBUTE ↑	VALUE	PATH
code	planID	_experience.profile.identityNamespaces._/devbc/plan/planID.namespace.code
createDate	1970-01-20T05:08:16Z	_repo.createDate
customerID	202208240125	_devbc.customerID
endDate	2022-01-20T13:15:32Z	_devbc.account.endDate
<b>firstName</b>	<b>YourFirstName</b>	<b>person.name.firstName</b>
id	yourEmail@adobe.com	identityMap.email.0.id
id	202208240125	identityMap.customerid.0.id



## Lab Tasks – Orders – Live data (Streaming)

In this exercise, we will load the Orders data from Streaming source to AEP Data Lake and Profile.

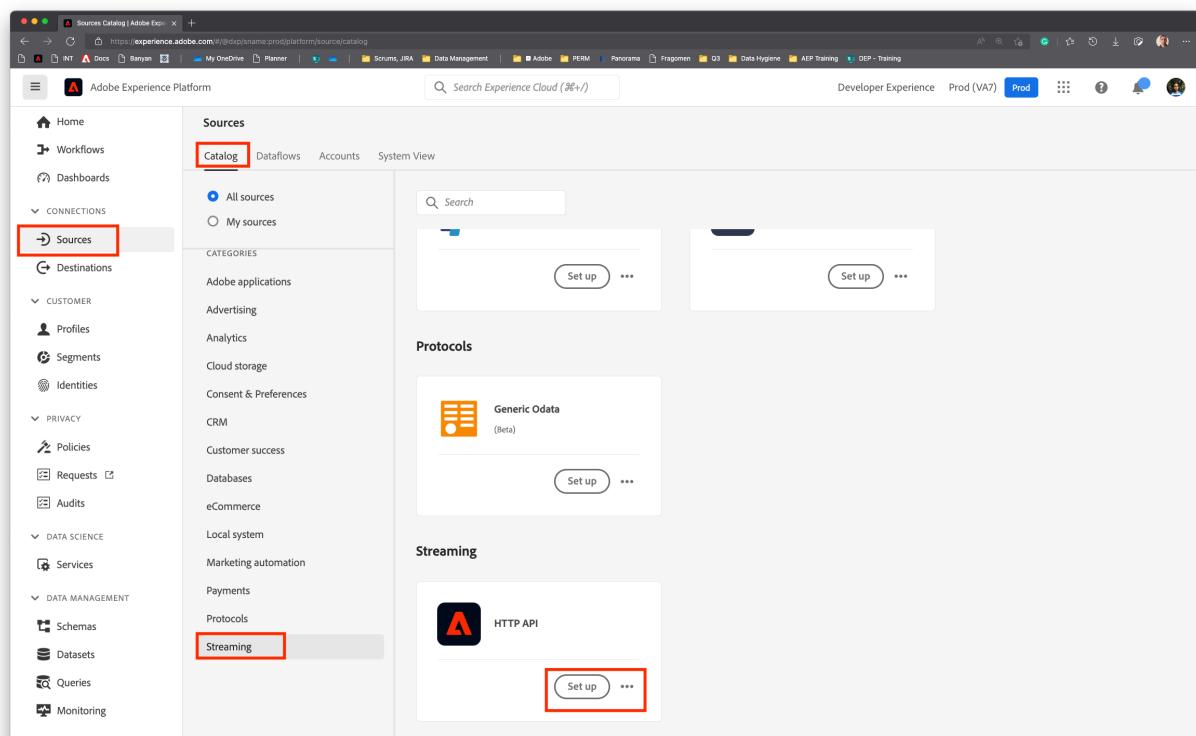
### Pre-requisites

1. Orders JSON file is downloaded from Azure ADLS and available on your local system
2. Orders Schema and Dataset are already created

### Steps

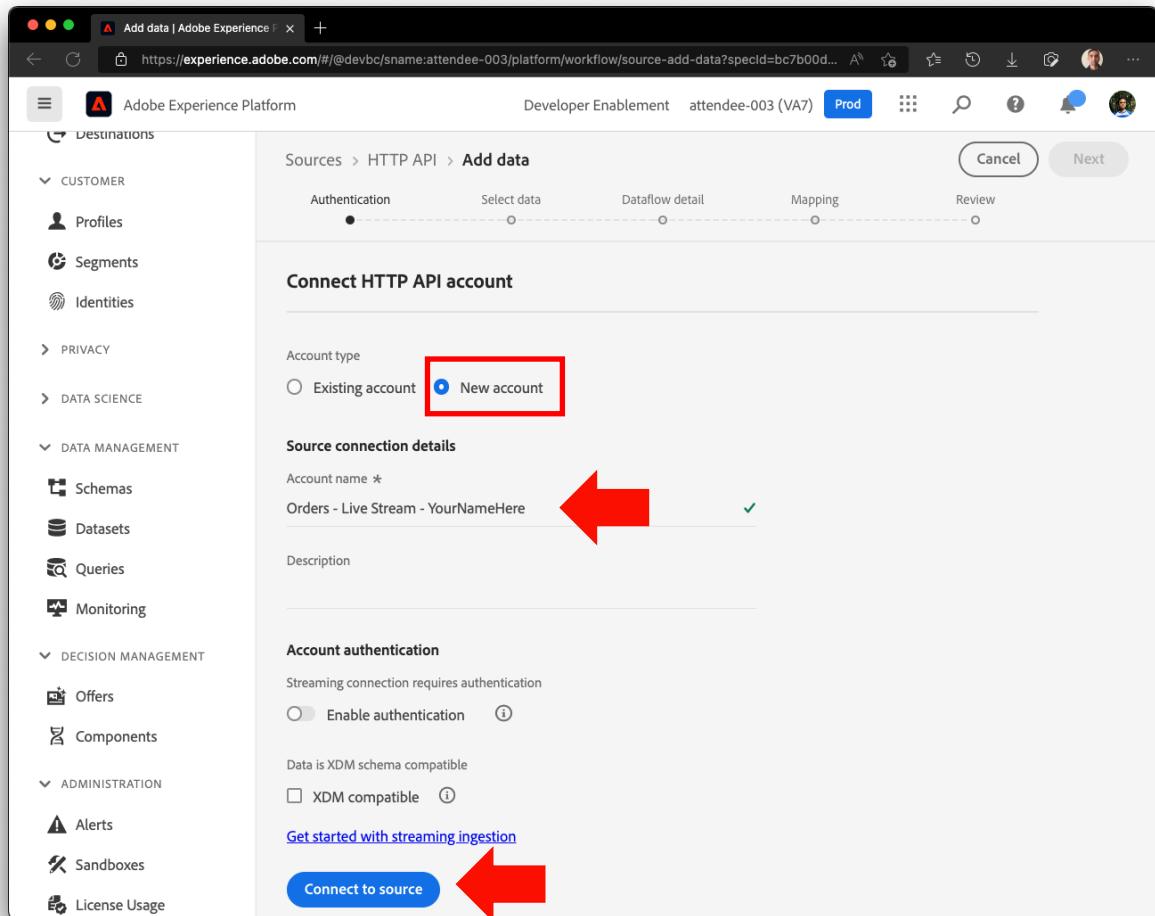
Go to Adobe Experience Platform → Sources → Catalog → Streaming. Click on **Setup / Add Data** for the HTTP API.

**Tip** If at least one connection exists for that source, you will see “**Add data**” as the default action. If no connections exist for that source, you will see “**Setup**” as the default action.



## Create a Streaming Inlet

In the Streaming Source, create a **New account** as shown in the screenshot below. For the new account, type the name as **Orders Live Stream**. Click **Connect to source**. Enable authentication must be OFF (default). XDM Compatible must be unchecked (default).



Once account is created, **Connected** message is displayed at the bottom. Click **Next** on the top right to continue.

The screenshot shows the Adobe Experience Platform 'Add data' workflow interface. The left sidebar contains navigation links for Sources, Destinations, CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring), DECISION MANAGEMENT (Offers, Components), ADMINISTRATION (Alerts, Sandboxes, License Usage). The main panel shows the 'Sources > HTTP API > Add data' workflow. The current step is 'Authentication'. The 'Source connection details' section includes an 'Account name' field with 'Orders - Live Stream - YourNameHere' and a checked 'Description' field. The 'Account authentication' section has an 'Enable authentication' checkbox checked. Below it, there's a note about XDM schema compatibility and a link to 'Get started with streaming ingestion'. A red arrow points to the 'Connected' status indicator, which is checked. Other options include 'Sign in with different credentials'.

## Upload sample data

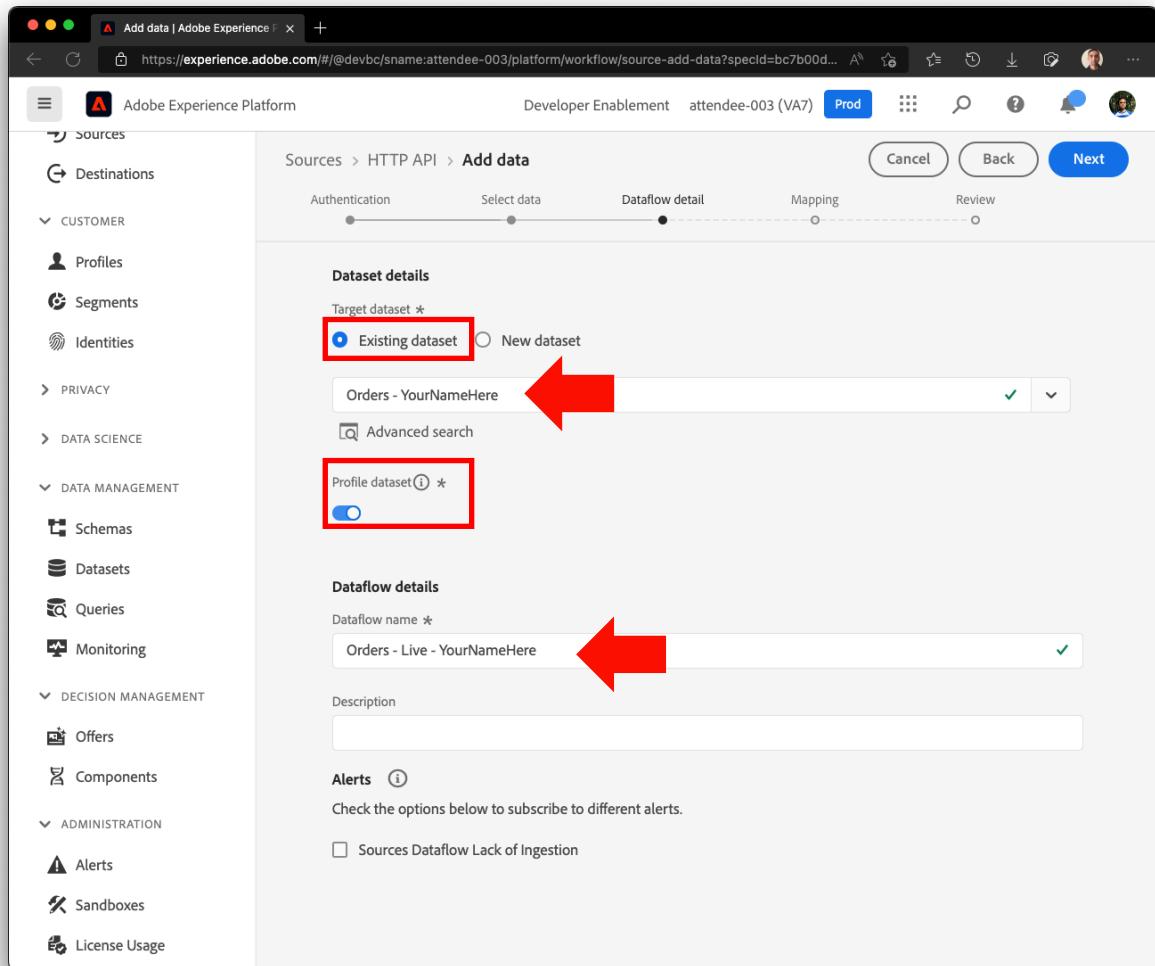
Click **Upload files** and select the Orders JSON file (**Lab\_Single\_Order.json**) from local file system. Data Preview loads automatically. Click **Next**.

The screenshot shows the 'Add data' interface in the Adobe Experience Platform. The left sidebar contains navigation links for Workflows, Dashboards, CONNECTIONS, Sources, Destinations, CUSTOMER, PRIVACY, DATA SCIENCE, and DATA MANAGEMENT. The main panel is titled 'Sources > HTTP API > Add data' and is currently on the 'Select data' step. A progress bar indicates the steps: Authentication (done), Select data (done), Dataflow detail (not started), Mapping (not started), and Review (not started). On the left, there's a 'Source data schema' section with a note: 'Upload sample JSON file to define source schema. Acceptable file size is up to 1GB.' Below it is a file upload input field containing 'Lab\_Single\_Order.json'. To the right, under 'Preview sample data:', a hierarchical tree view shows the JSON structure:

- SampleSchema:a515bf9307db472bbc53b163c9df1d55
  - Products | Object[]
  - plan | Object
    - billingCity | String ————— Salt Lake City | Value
    - billingState | String ————— Utah | Value
    - billingStreetAddress | String ————— 756 Northland Court | Value
    - billingZip | String ————— 84125 | Value
    - customerID | String ————— 748451685 | Value
    - lastOrderStatusUpdate | String ————— 2022-06-08T22:06:56Z | Value

## Define the Target dataset

In the screen, select existing dataset and choose **Orders learner - YourNameHere**. Set the Dataflow name as **Orders - Live Stream - YourNameHere**. Enable the Dataset for Profile. By toggling the **Profile dataset** box.



## Data Prep / Transformation

In the Data Prep (mapping) screen, Click on Import mapping button

The screenshot shows the 'Add data' workflow in the Adobe Experience Platform. The left sidebar contains navigation links for Home, Workflows, Dashboards, Connections, Sources, Destinations, Customer, Privacy, Policies, Requests, Audits, Data Science, and Decision Management. The main area is titled 'Sources > HTTP API > Add data' and is currently on the 'Mapping' step. A progress bar at the top indicates 'Mapped fields: 29 of 29' and 'Required fields: 1 of 3'. Below this, there are sections for 'Identity fields' (2 of 2) and an error message: 'There was error(s) preparing mappings.' The error message states: 'There is a duplicate mapping for the target path order.orderDate. The data at the XDM path will be overwritten. Please specify a distinct destination path.' It also mentions 'Errors: 3'. At the bottom, there is a table comparing 'SOURCE DATA' (e.g., shippingCity, Products[\*].quantity, Products[\*].productID, Products[\*].price, Products[\*].model) with 'TARGET FIELDS' (e.g., shipping.address.city, productListItems[\*].quantity, productListItems[\*].\_id, productListItems[\*].priceTotal, productListItems[\*].product). The 'Import mapping' button is highlighted with a red box.

Select **Orders – Backfill - YourNameHere** and click **Select**.

The screenshot shows the 'Sources' section of the Adobe Experience Platform interface. A modal window titled 'Import mapping' is open. The left sidebar lists various categories like CUSTOMER, DATA SCIENCE, and ADMINISTRATION. The main area shows a table with columns 'DATAFLOW NAME' and 'TARGET DATASET'. One row is selected, highlighted in blue, with the value 'Orders - Backfill - YourNameHere' in both columns. At the top right of the modal are 'Select' and 'Cancel' buttons, with a red arrow pointing to 'Select'. Below the table is a search bar labeled 'Search dataflow'.

DATAFLOW NAME	TARGET DATASET
Orders - Backfill - YourNameHere	Orders - YourNameHere
Dataflow - Customer Account - YourNameHe...	Customer Account - YourNameHere
Live Customer Updates - YourNameHere	Customer Account - YourNameHere
Web Streaming flow	dep: Web
Customer Streaming flow	dep: Customer Account
Orders Streaming flow	dep: Orders
Web Feed AWS S3 flow	dep: Web
Aggregates Feed AWS S3 flow	dep: Customer Aggregates

Mapping is now imported. All the previously existing mappings are Removed. Click Next.

The screenshot shows the 'Add data' interface in Adobe Experience Platform. The left sidebar includes sections for Sources, Destinations, CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring). The main area is titled 'Sources > HTTP API > Add data' and is at the 'Mapping' step. It displays progress bars for 'Mapped fields' (32 of 32), 'Required fields' (3 of 3), and 'Identity fields' (2 of 2). An error message indicates 'Errors: 0'. Below this, a mapping table compares 'SOURCE DATA' from 'Source: Lab\_Single\_Order.json' with 'TARGET FIELDS' in 'Mapping schema: dep: Orders'. The table lists the following mappings:

SOURCE DATA	TARGET FIELDS
Products["*"].productId	productListItems["*"].SKU
Products["*"]	productListItems["*"]
lastOrderStatusUpdate	timestamp
"inStore"	order_devbc.acqSource
concat(orderId, "", lastOrderStatusUpdate)	_id
Products["*"].model	productListItems["*"]_devbc.model
Products["*"].make	productListItems["*"]_devbc.make

Click on the icon to Preview data button to view the mapping output. Ensure there are no red exclamation marks next to any attributes

The screenshot shows the Adobe Experience Platform Workflows interface. On the left, a sidebar lists various categories: Home, Workflows (which is selected), Dashboards, CONNECTIONS, Sources, Destinations, CUSTOMER, Profiles, Segments, Identities, PRIVACY, Policies, Requests, Audits, DATA SCIENCE, Services, DATA MANAGEMENT, Schemas, Datasets, Queries, Monitoring, and DECISION MANAGEMENT. The main area is titled "Sources > HTTP API > Add data". A "Preview" section displays a hierarchical mapping structure. At the top right of the preview are buttons for "Cancel", "Back", and "Next". Below the preview, there are three columns: "\_DXP.CUSTOMERID" (962105342), "\_DXP.PERSONALEMAIL" (mchicchetto@msn.com), and "PRODUCTLISTIT" (Google). The mapping tree on the left starts with "dep: Orders" and branches into "dpx" (Object), "billing" (Object), "order" (Object), "productListItems" (Object[]), "shipping" (Object), "store" (Object), "\_id" (String), "eventMergeId" (String), "eventType" (String), and "identityMap" (Type). At the bottom of the preview, there are two input fields: "shippingZip" and "shipping.address.postalCode".

The final mapping set will look similar to this

The screenshot shows the 'Add data' workflow in Adobe Experience Platform. The left sidebar contains navigation links for Dashboards, CONNECTIONS, Sources, Destinations, CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring), DECISION MANAGEMENT (Offers, Components), and ADMINISTRATION (Alerts, Sandboxes, License Usage). The main area is titled 'Sources > HTTP API > Add data' and is currently on the 'Mapping' step. A progress bar at the top indicates '32 of 32 Mapped fields' and '3 of 3 Required fields'. Below the progress bar, there are sections for 'Identity fields' (2 of 2) and 'Errors: 0'. A search bar and filter dropdown ('All fields') are available. The mapping table compares 'SOURCE DATA' from 'Source: Lab\_Single\_Order.json' with 'TARGET FIELDS' in 'Mapping schema: dep: Orders'. The table lists the following mappings:

SOURCE DATA	TARGET FIELDS
Products[*].productId	productListItems[*].SKU
Products[*]	productListItems[*]
lastOrderStatusUpdate	timestamp
"inStore"	order._devbc.acqSource
concat(orderId, "-", lastOrderStatusUpdate)	_id
Products[*].model	productListItems[*]_devbc.model
Products[*].make	productListItems[*]_devbc.make
orderStatus	eventType
shippingCity	shipping.address.city

Click Next to continue, review the Dataflow and save it by clicking Finish.

The Dataflow is now created. Ensure the status of the Dataflow is **Enabled** so that it is ready to receive events.

## Ingest Live Data

We will ingest the live data using a REST API call to our HTTP API Endpoint. To do this, we will need the following:

1. Streaming Endpoint
2. Dataflow ID
3. IMS Org ID
4. Dataset ID

## Get the Streaming Endpoint

In the current screen, right hand side, note the Streaming endpoint.

**Tip** If you do not see this information, most likely the Dataflow is selected on the left. Click on the Dataflow row (not the blue link) on the left again to unselect it.

The screenshot shows the Adobe Experience Platform interface with the following details:

- Left Sidebar:** Contains sections for Dashboards, CONNECTIONS (Sources selected), Destinations, CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring), and DECISION MANAGEMENT.
- Top Bar:** Shows the URL <https://experience.adobe.com/#/@devbc/sname:attendee-003/platorm/source/accounts/de1ef49-d9ea-4029-b7a5-009c7c1847cd/activity?limit=50&page=1&sortD...>, the account name "Developer Enablement attendee-003 (VA7)", and a "Prod" button.
- Middle Content Area:** Displays the "Sources > Orders - Live Stream - YourNameHere" page. It shows a summary card for the "HTTP API" source, including the account name, dataset count (1), and last update date (August 24, 2022). Buttons for "Add data" and "Delete" are present.
- Right Panel:** Titled "Orders - Live Stream - YourNameHere". It contains fields for "ID" (de1ef49-d9ea-4029-b7a5-009c7c1847cd), "Source name" (Orders - Live Stream - YourNameHere), "Description" (empty), and "Source" (HTTP API). A "Category" field is set to "Streaming". A red box highlights the "Streaming endpoint" section, which contains the URL <https://dcs.adobedc.net/collection/65faa7288321153cb2061b85536a00b02095152e637c575a1a0ae9a5ddffff4>.

## Get Dataflow ID and Dataset ID

In the current screen, select the dataflow name, scroll down the right hand panel and take a note of the Dataflow ID and Dataset ID from the **API Usage** section.

The screenshot shows the Adobe Experience Platform Foundations Sources interface. On the left, a sidebar navigation includes Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring), and DECISION MANAGEMENT (Offers). The main content area displays an 'HTTP API' source named 'Orders - Live Stream - YourNameHere'. The source details show an account name, 1 dataset, and a last update date of August 24, 2022. Below this, a table lists a single dataflow named 'Orders - Live - YourNameHere' with target dataset 'Orders - YourNameHere', status 'No runs', and a last run date of 08/24/2022, 5:54 PM. A red box highlights this row. To the right, a 'Properties' panel shows the status as 'Processing', partial ingestion as 'Not enabled', error diagnostics as 'Not enabled', and a creation date of 08/24/2022, 5:54 PM. Another red box highlights the 'API Usage' section, which displays dataflow ID 'fe36eb0b-2550-4dc8-8a40-8cbd4b6d13d4', dataset ID '6306c0cac1897a1c08f7e7d4', mapping set ID '1b3643e3b4b046b4b6285982178c1864', and mapping version '0'. Navigation arrows at the bottom indicate the ability to view more dataflows.

### 3. Place a REST API call

In the Postman UI, set the following

Type: **POST**

URL: « Streaming Endpoint » you captured in the previous step.

Body: The below structure.

Replace your **flowId**, and **datasetId**.

```
{  
  "header": {  
    "flowId": "fe36eb0b-2550-4dcb-8a40-8cbd4b6d13d4",  
    "imsOrgId": "213A338362F288B50A495C3B@AdobeOrg",  
    "datasetId": "6306c0cac1897a1c08f7e7d4"  
  },  
  "body": {}  
}
```

The screenshot shows the Postman application interface. On the left, there's a sidebar with sections for Collections, APIs, Environments, Mock Servers, Monitors, and History. The main area shows a 'Scratch Pad' with a 'Customer Update' collection. A 'Pipeline Lab 4 - Orders Live Stream / Order Event' item is selected. The request details show a 'POST' method to 'https://dcs.adobedc.net/collection/65faa7288321153cb2061b85536a00b02095...'. The 'Body' tab is selected, displaying the JSON structure from the code block above. The JSON is:

```
1  {  
2   "header": {  
3     "flowId": "fe36eb0b-2550-4dcb-8a40-8cbd4b6d13d4",  
4     "imsOrgId": "213A338362F288B50A495C3B@AdobeOrg",  
5     "datasetId": "6306c0cac1897a1c08f7e7d4"  
6   },  
7   "body": {}  
51 }  
52 }
```

Below the JSON, there's a 'Response' section with a cartoon character holding a rocket. At the bottom, it says 'Click Send to get a response'.

Add the JSON request body as follows.

```
"body": {  
    "orderStatus": "placed",  
    "lastOrderStatusUpdate": "2022-08-11T17:43:01Z",  
    "customerID": "202208240125",  
    "personalEmail": "yourEmail@adobe.com",  
    "orderID": "unique-yourEmail-order-001",  
    "orderDate": "2022-08-11T17:43:01Z",  
    "orderTotal": 291.02,  
    "paymentType": "credit",  
    "paymentAmount": 349.35,  
    "paymentCurrencyCode": "USD",  
    "paymentTransactionID": "24596662-1a9b-4fc1-a98f-b5ca687c523f",  
    "storeID": "STORE-35",  
    "shippingStreetAddress": "30133 Golf Parkway",  
    "shippingCity": "Fullerton",  
    "shippingState": "California",  
    "shippingZip": "92640",  
    "shippingMethod": "Standard",  
    "shippingAmount": 0,  
    "shippingDestination": "store",  
    "billingStreetAddress": "497 Kensington Drive",  
    "billingCity": "Glinojec",  
    "billingZip": "06-450",  
    "plan": {  
        "ID": "m3"  
    },  
    "Products": [  
        {  
            "productID": "PRODUCT-6",  
            "quantity": 1,  
            "price": 150.23,  
            "currencyCode": "USD",  
            "model": "v2",  
            "make": "Google"  
        },  
    ]  
}
```

```
{
  "productID": "PRODUCT-11",
  "quantity": 1,
  "price": 150.23,
  "currencyCode": "USD",
  "model": "v1",
  "make": "Google"
}
]
}
```

Change the following attributes to make it unique:

Attribute Name	Type	Suggested values
customerID	Number	<ul style="list-style-type: none"> <li>1. Must be valid number</li> <li>2. MUST Match what you used in Customer Account REST API call</li> </ul>
personalEmail	Email	<ul style="list-style-type: none"> <li>1. Must be valid email with proper format.</li> <li>2. MUST match what you used in Customer Account REST API call</li> </ul>
orderID	String	<ul style="list-style-type: none"> <li>1. MUST be unique for every REST API call you make.</li> <li>2. Suggested to have your email suffixed by a sequence number</li> </ul>
last Order StatusUpdate	Timestamp	<ul style="list-style-type: none"> <li>1. Must be a valid timestamp</li> <li>2. Example format is 2022-08-11T17:43:01Z</li> <li>3. MUST be current or past date</li> </ul>

**WARNING**

Make sure these attributes are unique. If the row is not unique, system will not warn and the event will be ignored (not reported as warning / error).

Full REST call will look like this:

```
{
  "header": {
    "flowId": "fe36eb0b-2550-4dcb-8a40-8cbd4b6d13d4",
    "imsOrgId": "213A338362F288B50A495C3B@AdobeOrg",
    "datasetId": "6306c0cac1897a1c08f7e7d4"
  },
  "body": {
    "orderStatus": "placed",
    "lastOrderStatusUpdate": "2022-08-11T17:43:01Z",
    "customerID": "202208240125",
    "personalEmail": "yourEmail@adobe.com",
    "orderID": "unique-yourEmail-order-001",
    "orderDate": "2022-08-24T17:43:01Z",
    "orderTotal": 291.02,
    "paymentType": "credit",
    "paymentAmount": 349.35,
    "paymentCurrencyCode": "USD",
    "paymentTransactionID": "24596662-1a9b-4fc1-a98f-b5ca687c523f",
    "storeID": "STORE-35",
    "shippingStreetAddress": "30133 Golf Parkway",
    "shippingCity": "Fullerton",
    "shippingState": "California",
    "shippingZip": "92640",
    "shippingMethod": "Standard",
    "shippingAmount": 0,
    "shippingDestination": "store",
    "billingStreetAddress": "497 Kensington Drive",
    "billingCity": "Glinojeck",
    "billingZip": "06-450",
    "plan": {
      "ID": "m3"
    },
    "Products": [
      {
        "productID": "PRODUCT-6",
      }
    ]
  }
}
```

```

    "quantity": 1,
    "price": 150.23,
    "currencyCode": "USD",
    "model": "v2",
    "make": "Google"
},
{
    "productId": "PRODUCT-11",
    "quantity": 1,
    "price": 150.23,
    "currencyCode": "USD",
    "model": "v1",
    "make": "Google"
}
]
}
}

```

Click Send and the event will be posted to the Streaming Endpoint. The response body will be similar to this:

```
{
    "inletId": "65faa7288321153cb2061b85536a00b02095152e637c575a1a0ae9a5ddffff4",
    "xactionId": "1661389779753:1943:76",
    "flowId": "fe36eb0b-2550-4dcb-8a40-8cbd4b6d13d4",
    "receivedTimeMs": 1661389779754
}
```

Body ▼

200 OK 350 ms 439 B | Save Response ▼

Pretty Raw Preview Visualize JSON ▼ ≡ □ 🔍

```

1  {
2      "inletId": "65faa7288321153cb2061b85536a00b02095152e637c575a1a0ae9a5ddffff4",
3      "xactionId": "1661389779753:1943:76",
4      "flowId": "fe36eb0b-2550-4dcb-8a40-8cbd4b6d13d4",
5      "receivedTimeMs": 1661389779754
6  }

```

## Verify data in Profile

Go to Adobe Experience Platform → **Profiles** in the left Nav bar and click on the **Browse** on the top.

The screenshot shows the 'Profiles' section of the Adobe Experience Platform. The left sidebar has a red box around the 'Profiles' item under the 'CUSTOMER' section. The main area has a red box around the 'Browse' tab at the top. Below it, there's a search bar for 'Merge policy' set to 'Default Timebased' and an 'Identity namespace' input field. A 'View' button is next to the namespace input. To the right, a summary box says '10 Profiles'. A table below lists 10 profile IDs: BVRQZWV003W6LLip4IWpuvqaLLXq6LrP3KJg, BVRQZWV1IIK2p3Xq7KJ9\_rHmrbZXrYpnP3KJg, BVRQZWV31pK4r3p86J\_PnGq4KHkJZxnLYr3v3KJg, BVRQZWVryrl3mpePonKv3KJg, BVRQZWVsY2ra3sHnpPfnnf4KLw, and BVRQZWVtzJnK6MjXsu\_rXnf3KJg. The right side also shows a 'Profiles by namespace' section with a customerID of 10.

PROFILE ID	FIRST NAME	LAST NAME	PERSONAL EMAIL
BVRQZWV003W6LLip4IWpuvqaLLXq6LrP3KJg	-	-	-
BVRQZWV1IIK2p3Xq7KJ9_rHmrbZXrYpnP3KJg	-	-	-
BVRQZWV31pK4r3p86J_PnGq4KHkJZxnLYr3v3KJg	-	-	-
BVRQZWVryrl3mpePonKv3KJg	-	-	-
BVRQZWVsY2ra3sHnpPfnnf4KLw	-	-	-
BVRQZWVtzJnK6MjXsu_rXnf3KJg	-	-	-

In the Browse tab, Leave the Merge policy to be the **Default Timebased** and select identity namespace as **Email**. Identity value text box appears. Type in your email (the one you used in the REST API Payload such as **yourEmail@adobe.com**) and press **View**.

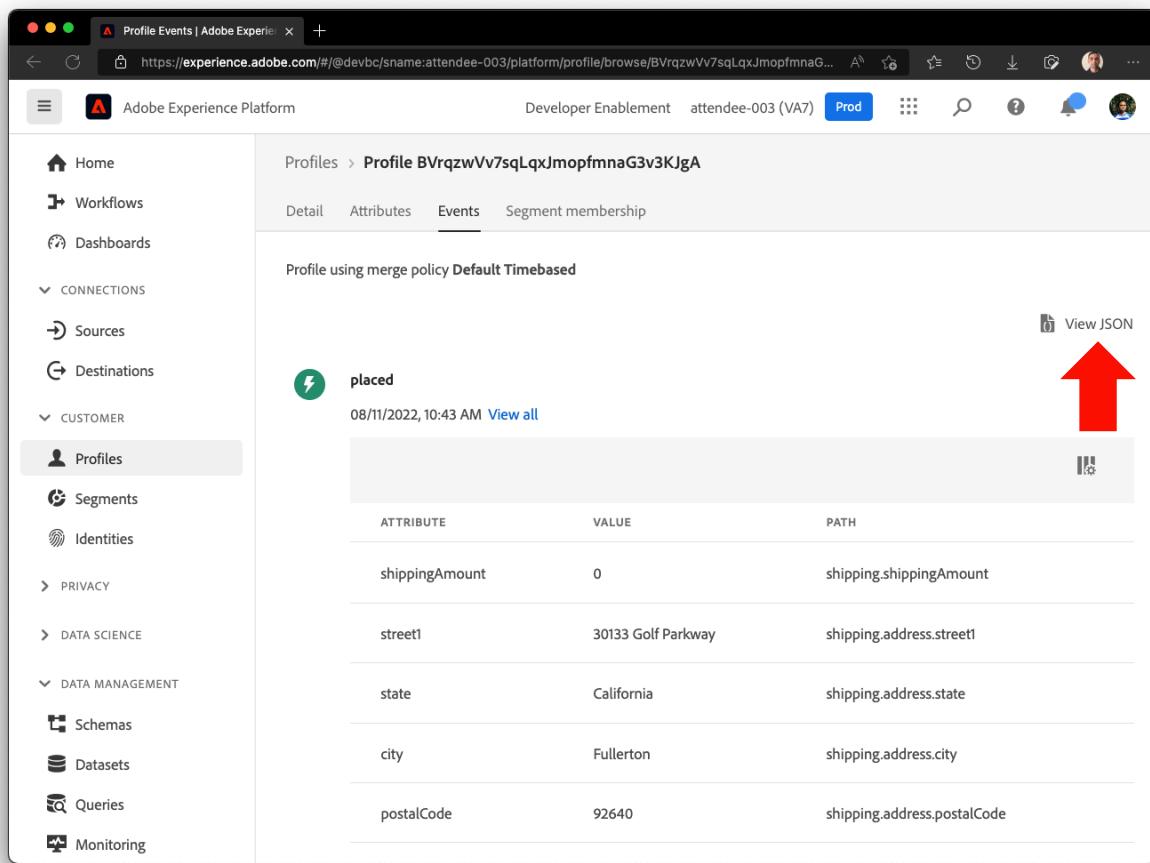


A profile row will appear below. The profile row will have the first name, last name and email you supplied in the REST API call. Click on the Profile ID to open the Profile.

PROFILE ID	FIRST NAME	LAST NAME	PERSONAL EMAIL
BVrqzwVv7sqLqxJmopfmnaG3v3KJgA	YourFirstName	Updated Last Name	yourEmail@adobe.com

Profile gets loaded. This profile was originally loaded via Customer Accounts Lab (Lab 2) and updated in the previous Customer Accounts REST call in this lab. Click on the Events tab.

Events will appear as follows. Click on **View JSON** to see full JSON of the event.



The screenshot shows the Adobe Experience Platform Profile Events interface. On the left is a sidebar with navigation links: Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring). The main area shows a profile detail for 'Profile BVrqzwVv7sqLqxJmopfmnaG3v3KJgA'. The 'Events' tab is selected, showing an event titled 'placed' from 08/11/2022, 10:43 AM. Below the event is a table with columns ATTRIBUTE, VALUE, and PATH, listing shipping information: shippingAmount (0, shipping.shippingAmount), street1 (30133 Golf Parkway, shipping.address.street1), state (California, shipping.address.state), city (Fullerton, shipping.address.city), and postalCode (92640, shipping.address.postalCode).

ATTRIBUTE	VALUE	PATH
shippingAmount	0	shipping.shippingAmount
street1	30133 Golf Parkway	shipping.address.street1
state	California	shipping.address.state
city	Fullerton	shipping.address.city
postalCode	92640	shipping.address.postalCode

Full JSON appears as below. Look for the **OrderID** attribute and verify that the value matches with what you have provided in the REST API call.

Profile Events | Adobe Experience Platform

Developer Enablement attendee-003 (VA7) Prod

Home Workflows Dashboards CONNECTIONS Sources Destinations CUSTOMER Profiles Segments Identities PRIVACY DATA SCIENCE DATA MANAGEMENT Schemas Datasets Queries Monitoring

Profiles > Profile BVrqzwVv7saLqxJmopfmnaG3v3KJgA

View JSON

```

    "transactionID": "24596662-1a9b-4fc1-a98f-b5ca687c523f",
    "paymentType": "credit"
},
"_devbc": {
    "plan": {
        "planID": "m3"
    },
    "acqSource": "inStore"
},
"orderID": "unique-yourEmail-order-001",
"priceTotal": 291.02
},
"billing": {
    "address": {
        "city": "Glinojeck",
        "street1": "497 Kensington Drive",
        "postalCode": "06-450"
    }
},

```

Close

raw JSON

```

"transactionID": "24596662-1a9b-4fc1-a98f-b5c
    "paymentType": "credit"
},
"_devbc": {
    "plan": {
        "planID": "m3"
    },
    "acqSource": "inStore"
},
"orderID": "unique-yourEmail-order-001",
"priceTotal": 291.02
},
"billing": {
    "address": {
        "city": "Glinojeck",
        "street1": "497 Kensington Drive",
        "postalCode": "06-450"
    }
},

```

	body:
7	... "orderStatus": "placed",
8	... "lastOrderStatusUpdate": "2022-08-11T17:43:01Z",
9	... "customerID": "202208240125",
10	... "personalEmail": "yourEmail@adobe.com",
11	... "orderID": "unique-yourEmail-order-001",
12	... "orderDate": "2022-08-24T17:43:01Z",
13	... "orderTotal": 291.02,
14	... "paymentType": "credit",
15	... "paymentAmount": 349.35,
16	... "paymentCurrencyCode": "USD",
17	... "paymentTransactionID": "24596662-1a9b-4fc1-a98f-b5ca687c523f",
18	... "storeID": "STORE-35",
19	... "shippingStreetAddress": "30133 Golf Parkway",
20	

## Learnings

1. Event data can be ingested via HTTP API or other Streaming connectors
2. Data often reflects in Profile in few minutes
3. Event data gets attached to the Profile data based on the Primary Identity provided inside the event (**personalEmail** in this case)

## Additional exercises

### Exercise #1

Try executing the same REST API call with a different Order ID and Order Status. You will notice that the Profile has multiple events now.

```
7 ... "body": {  
8 ...   "orderStatus": "cancelled",  
9 ...   "lastOrderStatusUpdate": "2022-08-11T17:43:01Z",  
10 ...   "customerID": "202208240125",  
11 ...   "personalEmail": "yourEmail@adobe.com",  
12 ...   "orderID": "unique-yourEmail-order-002",  
13 ...   "orderDate": "2022-08-24T17:43:01Z",  
14 ...   "orderTotal": 291.02,  
15 ...   "paymentType": "credit",  
16 ...   "paymentAmount": 349.35,  
17 ...   "paymentCurrencyCode": "USD",  
18 ...   "paymentTransactionID": "24596662-1a9b-4fc1-a98f-b5ca687c523f",  
19 ...   "storeID": "STORE-35",  
20 ...   "shippingStreetAddress": "30133 Golf Parkway",
```

Profile will show multiple events now.

The screenshot shows the Adobe Experience Platform interface for a specific profile. The left sidebar contains navigation links for Home, Workflows, Dashboards, Connections, Sources, Destinations, Customer (Profiles, Segments, Identities), Privacy, Data Science, Data Management (Schemas, Datasets, Queries, Monitoring), Decision Management (Offers, Components), and Administration (Alerts, Sandboxes, License Usage). The main content area displays the profile details for 'Profile BVrqzwVv7sqLqxJmopfmnaG3v3KJgA'. The 'Events' tab is selected, showing two entries:

- placed** (red arrow points here) - Occurred on 08/11/2022, 10:43 AM. The event table shows:

ATTRIBUTE	VALUE	PATH
shippingAmount	0	shipping.shippingAmount
street1	30133 Golf Parkway	shipping.address.street1
state	California	shipping.address.state
city	Fullerton	shipping.address.city
postalCode	92640	shipping.address.postalCode
- cancelled** - Occurred on 08/11/2022, 10:43 AM. The event table shows:

ATTRIBUTE	VALUE	PATH
shippingAmount	0	shipping.shippingAmount
street1	30133 Golf Parkway	shipping.address.street1
state	California	shipping.address.state
city	Fullerton	shipping.address.city

