



# Adobe Experience Platform Bootcamp Deep Dive Edition

DOCUMENTATION AND LABS



## Contents

### Lab 1 - Lookup Stores

1. Lab Overview	4
2. Learning Objectives	4
3. Lab Task – Creation of the Dataflow	5
3.1. Select source data	5
3.2. Define the target dataset	7
3.3. Data Prep / Transformation	8
3.4. Schedule	10
3.5. Check scheduled execution	13

### Lab 2 Customer Accounts

1. Lab Overview	17
2. Learning Objectives	17
3. Lab Tasks – Ingest Customer Account data	18
3.1. Select Source data	18
3.2. Define the Target Dataset	21
3.3. Data Prep / Transformation	22
3.4. Preview the mapping output	23
3.5. Re-map the attributes	25
3.6. Calculated Fields	26
3.7. Schedule	32
3.8. Check execution status and error messages	34
3.9. Correct the Dataflows	41
3.10. Re-schedule the failed Dataflow	48

### Lab 3 - Historical

1. Lab Overview	56
2. Learning Objectives	56
3. Lab Tasks – Orders – Historical data	56
3.1. Select Source data	57
3.2. Define the Target dataset	58
3.3. Data Prep / Transformation	59
3.4. Remove the following mappings	59
3.5. Create Calculated Fields	60
3.6. Add new pass-through mappings	63
3.7. Product List Items	63
3.8. Final Mapping Set	64

3.9. Preview the data	66
3.10. Schedule	68
3.11. Check scheduled execution	69
3.12. Verify the data	71
<b>Lab 4 - Streaming</b>	<b>74</b>
1. Lab Overview	56
2. Learning Objectives	56
3. Lab Tasks A – Customer Accounts – Live data (Streaming)	75
3.1. Creating a Streaming Inlet	76
3.2. Upload sample data	78
3.3. Import the Customer Accounts Mapping	81
3.4. Streaming Endpoint	85
3.5. Place a REST API call	87
3.6. Verify the data in Profile	90
3.7. Learnings	91
3.8. Additional exercises	91
4. Lab Tasks B – Orders – Live data (Streaming)	94
4.1. Create a Streaming Inlet	95
4.2. Upload sample data	97
4.3. Define the Target dataset	98
4.4. Data Prep / Transformation	99
4.5. Ingest Live Data	105
4.6. Place a REST API call	107
4.7. Verify data in Profile	112
4.8. Learnings	116
4.9. Additional exercises	116



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

Adobe Experience Platform Bootcamp Deep Dive Edition

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

## 1. 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**

## 2. 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

### 3. Lab Task – Creation of the Dataflow

#### 3.1. 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.

The screenshot shows the Adobe Experience Platform interface. The left sidebar has a tree view with 'Sources' selected. Under 'Sources', 'Cloud storage' is expanded, showing options like 'Azure File Storage', 'Data Landing Zone', 'FTP (Beta)', 'Google Cloud Storage', 'Google PubSub', and 'Oracle Object Storage'. Each card has a 'Set up' button. The 'Data Landing Zone' card's 'Set up' button is highlighted with a red box.

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

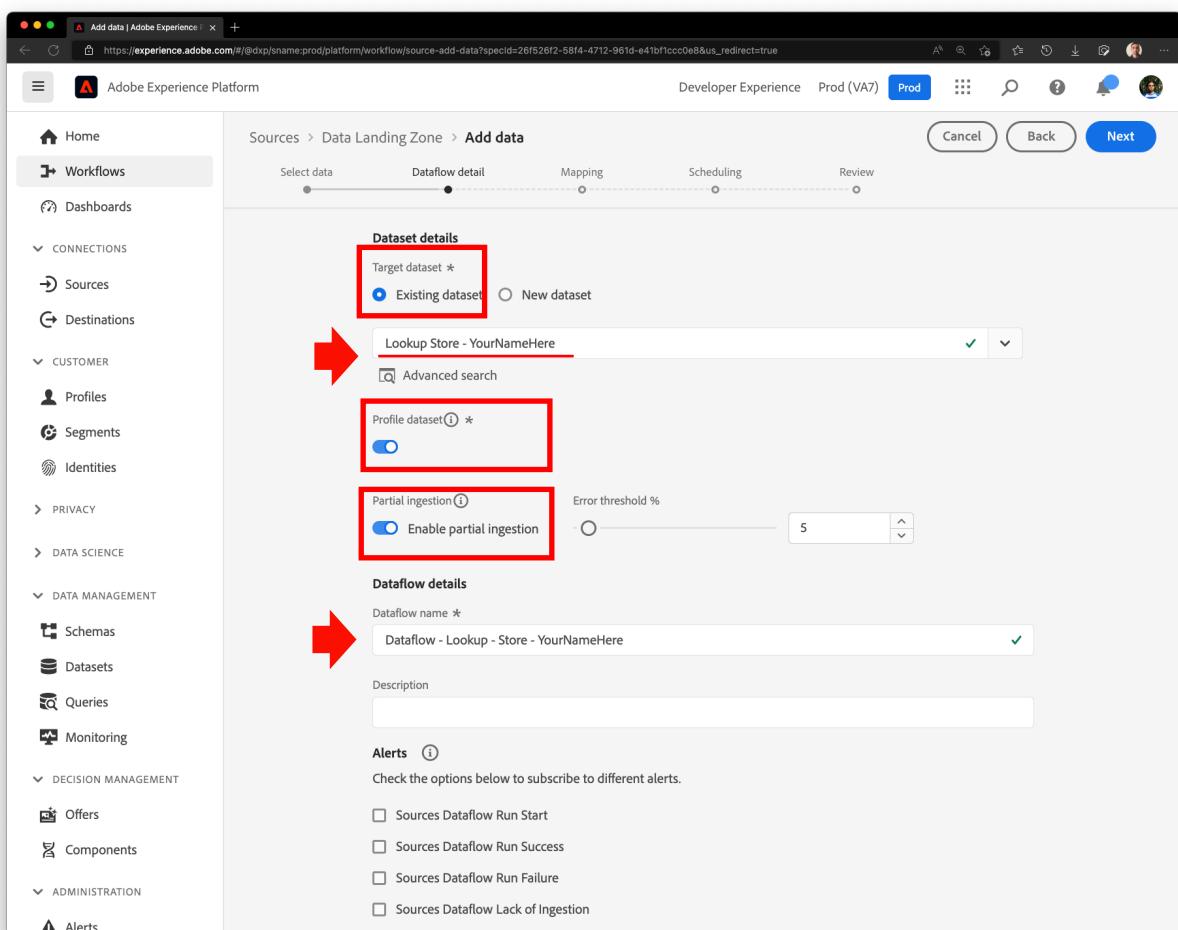
The screenshot shows the 'Add data' interface in Adobe Experience Platform. The left sidebar navigation includes Home, Workflows (selected), Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY (Policies, Requests, Audits), DATA SCIENCE (Services), and DATA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

	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

### 3.2.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.



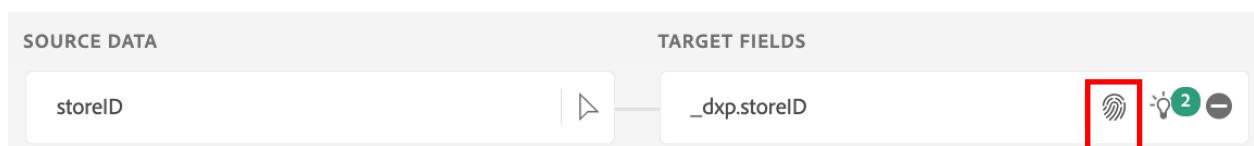
### 3.3.Data Prep / Transformation

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 data is from a CSV file named 'Lookup\_Store\_CLIENT.csv'. The target fields belong to a schema named 'dep: Lookup Store'. All 9 source fields are mapped to target fields, indicated by green progress bars. A red error icon in the top right corner shows 0 errors. The sidebar on the left shows various platform navigation links.

## 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 for the 'storeID' source field. It lists three options: 'SELECT ALL FIELDS' (unchecked), '\_dpx.storeID' (checked), and '\_dpx.address.\_id' (unchecked). A button 'Select manually' is also present. To the right, there are four small cards with icons and numbers (2, 1, 1, 1) and a search/filter icon.

### 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.

### Final Mapping Set

The final mapping set should look like this

Source Column	XDM Property
storeID	_devbc.storeID
storeName	_devbc.name
storeInagurationDate	_devbc.inagationDate
storeStreet	_devbc.address.street1
storeCity	_devbc.address.city
storeState	_devbc.address.state
storeZipCode	_devbc.address.postalCode
storeOpenTime	_devbc.openTime
storeCloseTime	_devbc.closeTime

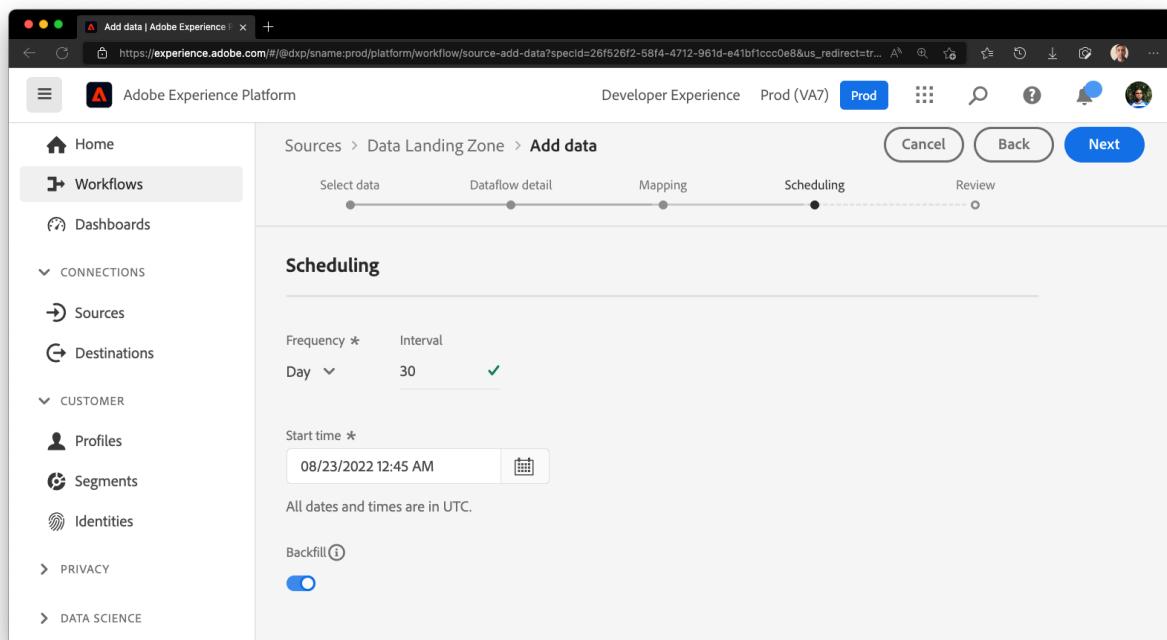
Click **Next** to proceed to schedule screen.



### 3.4.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

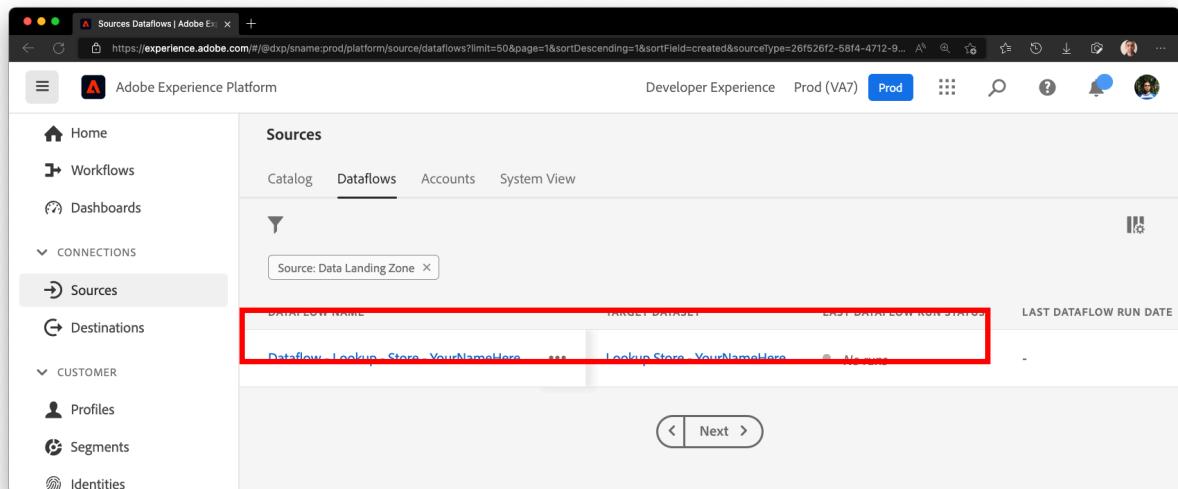


When the **Review** screen appears, click **Finish**.

The screenshot shows the 'Add data' workflow in the Adobe Experience Platform. The left sidebar navigation includes Home, Workflows (selected), Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets, Queries). The main content 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 (the last 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). The 'Finish' button is highlighted in blue at the top right.

## 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.



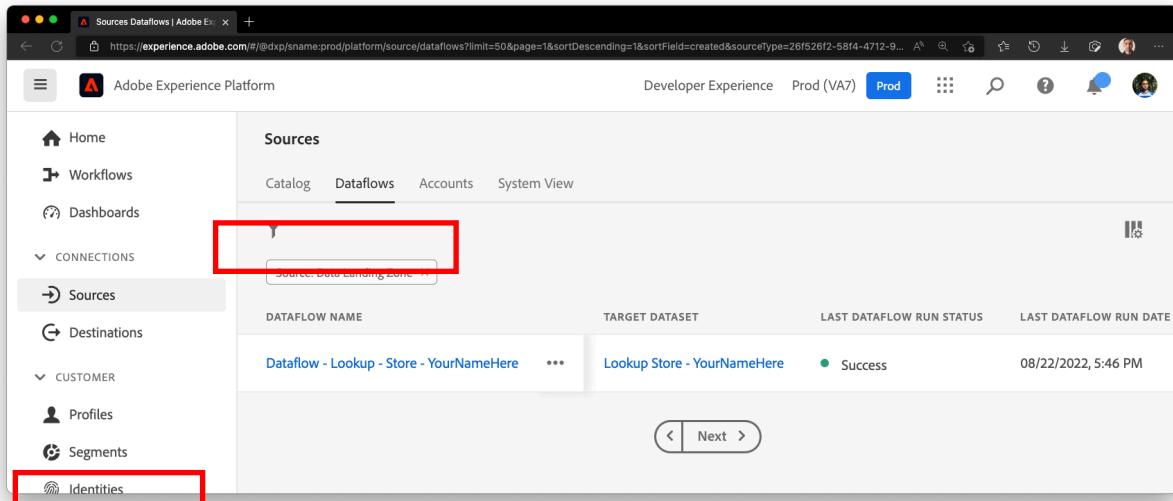
The screenshot shows the Adobe Experience Platform interface for managing dataflows. The left sidebar has sections for Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), and CUSTOMER (Profiles, Segments, Identities). The main area is titled 'Sources' with tabs for Catalog, Dataflows, Accounts,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

DATAFLOW NAME	TARGET TARGET	LAST DATAFLOW RUN STATUS	LAST DATAFLOW RUN DATE
Dataflow_Lookup_Store_YourNameHere	Lookup Store_YourNameHere	No runs	-

A red box highlights the 'Dataflow\_Lookup\_Store\_YourNameHere' row. At the bottom right of the table area is a navigation bar with '<' and 'Next >'. There is also a large red box covering the bottom portion of the page content.

### 3.5.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.

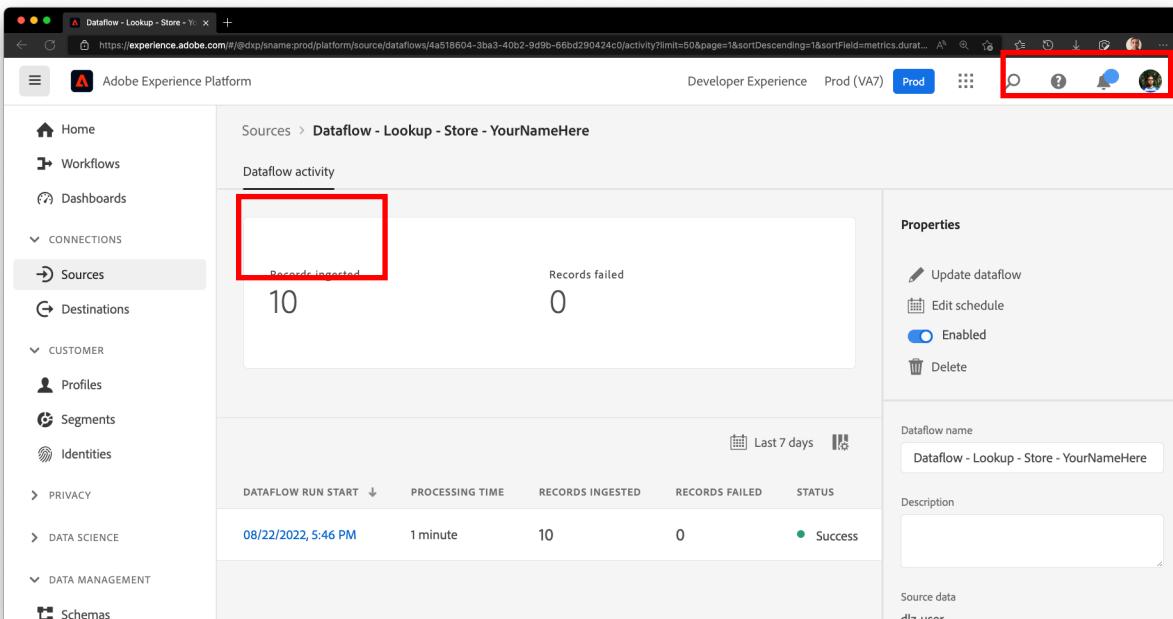


The screenshot shows the Adobe Experience Platform interface for managing dataflows. The left sidebar includes options like Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments), and Identities (highlighted with a red box). The main area is titled 'Sources' and shows the 'Dataflows' tab selected. A search bar at the top has 'Source: Data Landing Zone' typed into it. Below the search bar, a table lists a single dataflow entry:

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

At the bottom right of the table are navigation buttons for 'Next' and 'Previous'.

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



The screenshot shows the details of the 'Dataflow - Lookup - Store - YourNameHere' activity. The left sidebar remains the same. The main area is titled 'Sources > Dataflow - Lookup - Store - YourNameHere' and shows the 'Dataflow activity' section. This section displays the number of records ingested (10) and records failed (0). To the right is a 'Properties' panel with options to update the dataflow, edit the schedule, enable or disable it, and delete it. It also shows the dataflow name ('Dataflow - Lookup - Store - YourNameHere') and a description field. At the bottom, a table provides a summary of the last run: start time (08/22/2022, 5:46 PM), processing time (1 minute), records ingested (10), records failed (0), and status (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 Adobe Experience Platform Datasets Browse interface. On the left, the navigation bar includes Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations, CUSTOMER, PRIVACY, DATA SCIENCE), DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring). The Datasets section is selected. The main area displays a table of datasets:

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

A summary card on the right indicates there are 33 Datasets and lists the most recently updated datasets.

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

The screenshot shows the Adobe Experience Platform Dataset activity interface for the 'Lookup Store - YourNameHere' dataset. The left navigation bar is identical to the previous screenshot. The main area displays dataset activity metrics for the last 7 days:

Total records in previous month	Ingested records in the last 7 days	Ingested batches today	Ingested batches in the last 7 days
10	0	1	0

Below this, it shows the size of data in the previous month: 4.42 kB. The right side of the screen shows the dataset details for 'Lookup Store - YourNameHere':

- Name: Lookup Store - YourNameHere
- Description: (empty)
- Dataset ID: 63041d154655631c077d6133
- Table name: lookup\_store\_yournamehere
- Profile: (switch is off)
- 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 `_dxd → 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).



# Adobe Experience Platform

## Lab 2 - Data Pipeline Customer Accounts

Adobe Experience Platform Bootcamp Deep Dive Edition

Name \_\_\_\_\_

Sandbox \_\_\_\_\_

The screenshot shows the Adobe Experience Platform interface for managing datasets. 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' section is selected.

The main area displays a dataset named 'Lookup Store - YourNameHere'. A search bar at the top says 'Search Experience Cloud (⌘+/)'.

The dataset view has two sections:

- Schema Tree:** On the left, it shows a hierarchical schema structure:

```
dep:Lookup Store
  +-- _dpx | Object
    +-- address | Object
      +-- city | String
      +-- postalCode | String
      +-- state | String
```
- Dataset Preview:** On the right, it shows a table with columns: '\_DXP.ADDRESS.CITY', '\_DXP.CLOSETIME', and '\_DXP.ADDRESS.STREET'. The data rows are:

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



## 1. 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**

## 2. 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

### 3. 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.

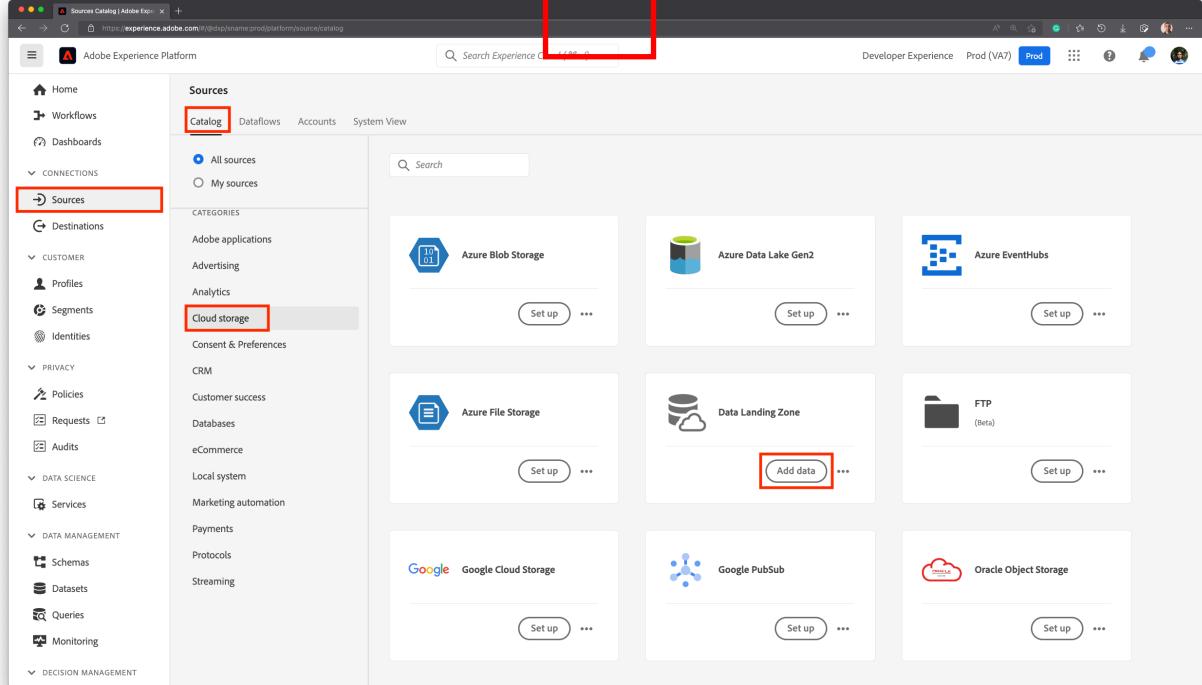
#### Pre-requisites

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

#### 3.1. 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.



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**.

The screenshot shows the 'Add data' interface in Adobe Experience Platform. The left sidebar has 'Workflows' selected. The main area is titled 'Sources > Data Landing Zone > Add data' and is at the 'Select data' step. The 'Selected data' section shows a tree view: CUSTOMER > / > dlz-user-container > project > PIPELINE > CUSTOMER. Under CUSTOMER, 'Lab\_Customer\_001.csv' is selected. To the right, there's a 'Preview' section with a table showing sample data:

	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 a 'Selected data' section for 'CUSTOMER' data, which is part of a 'PIPELINE'. Below this, there is a table with columns: SMS\_OPTIN, CUSTOMER\_ID, SHIPPING\_STREET\_ADDRESS, SHIPPING\_CITY, and SHIPPING\_STATE. The table contains 10 rows of data. The first row has a null value for 'SMS\_OPTIN'. The last row has a null value for 'SMS\_OPTIN' and a proper date format for 'ACCOUNT\_END\_DATE'.

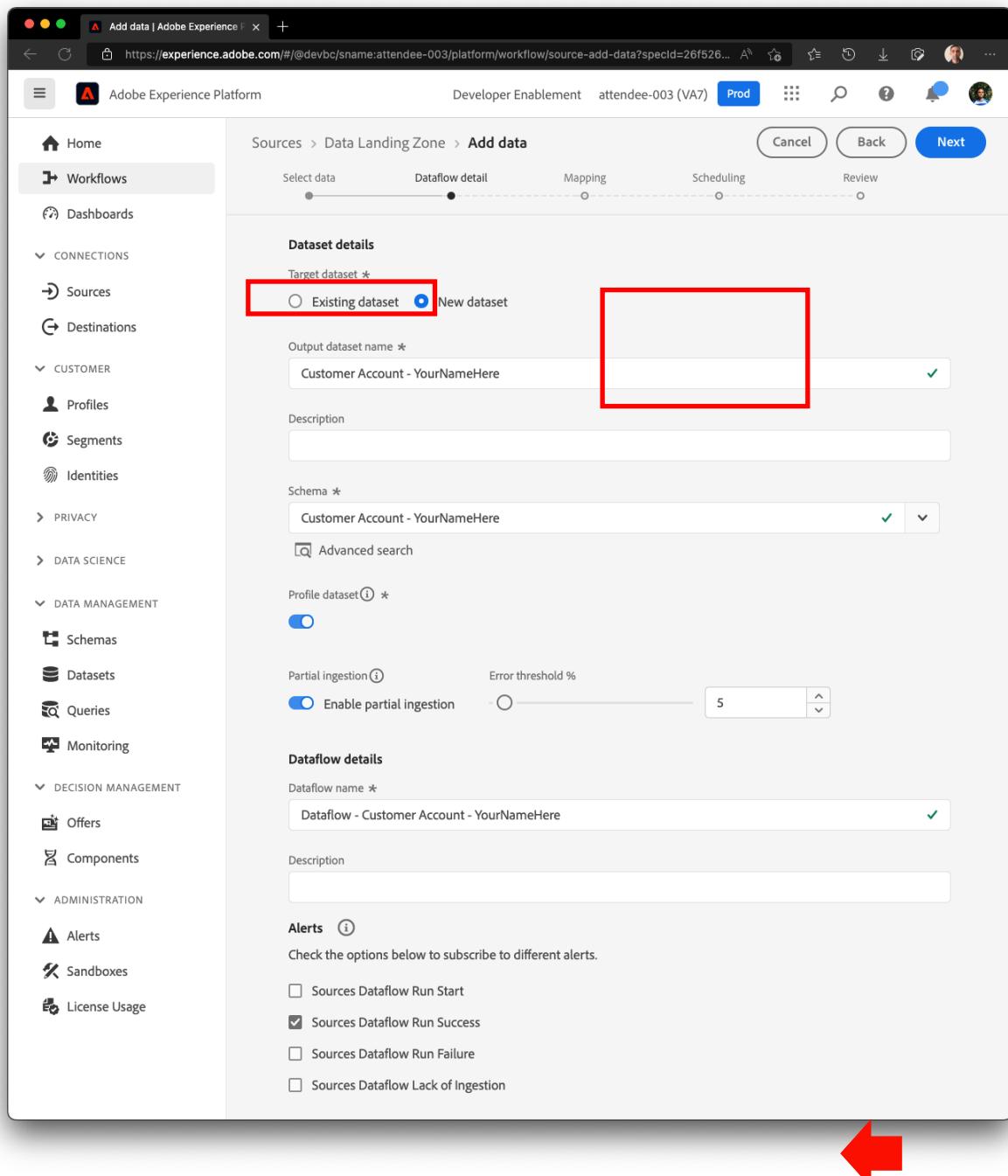
SMS_OPTIN	CUSTOMER_ID	SHIPPING_STREET_ADDRESS	SHIPPING_CITY	SHIPPING_STATE
-	736072406	706 Sloan Lane	Fresno	CA
y	495305838	985 Elmside Court	Austin	TX
n	442237547	9 Longview Center	Kansas City	MO
-	462757264	0 Lyons Crossing	Roanoke	VA
y	764162251	82972 Karstens Junction	Worcester	MA
y	983010429	6 Di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

This screenshot shows a similar 'Add data' workflow step, but for 'Lab\_Customer.csv' data. The table has different columns: BILLING\_ZIP\_CODE, PLAN\_ID, PLAN\_NAME, ACCOUNT\_CREATE\_DATE, ACCOUNT\_END\_DATE, and SOURCE. The last two rows show proper date formats for 'ACCOUNT\_END\_DATE'.

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

### 3.2.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.



### 3.3.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

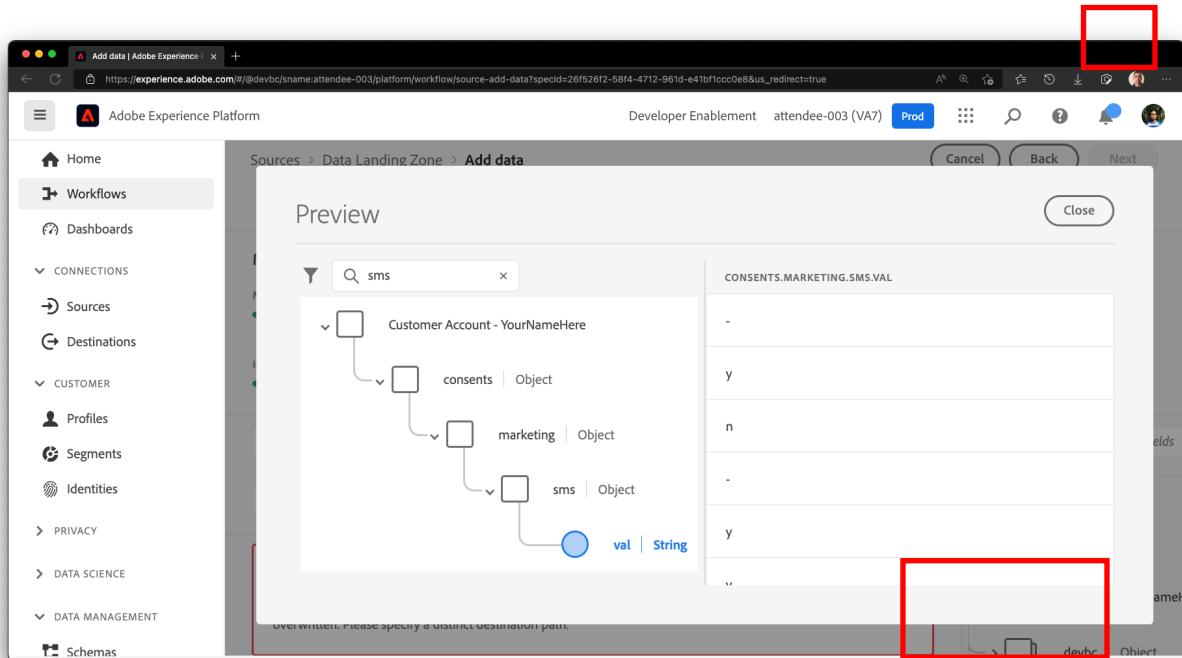
**Warning** Due to the ML Recommendations, your screen may look different than the screenshot below.

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. It displays a progress bar for mapping source fields to target dataset fields, showing 23 of 23 mapped fields, 1 of 1 required fields, and 1 of 2 identity fields. A red warning box at the bottom states: 'There was error(s) preparing mappings.' followed by the message: '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.' Below this, the 'SOURCE DATA' and 'TARGET FIELDS' sections show mappings for createDate, modifyDate, firstName, and lastName. The 'lastName' mapping is highlighted with a red border, indicating the error mentioned in the warning box.

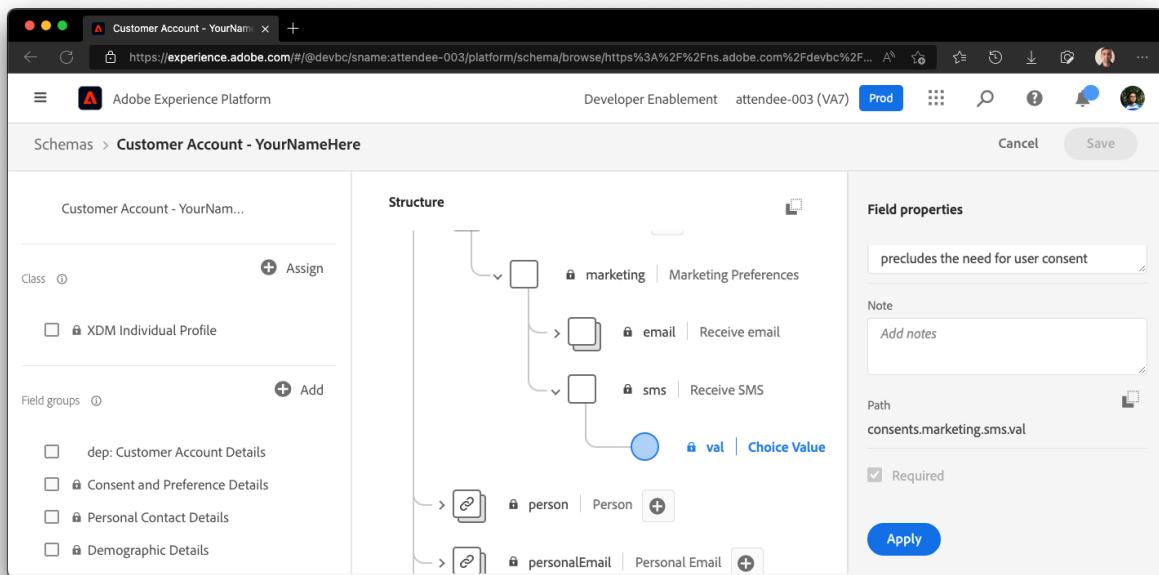
### 3.4. Preview the mapping output

#### 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 interface. On the left, the navigation sidebar includes 'Workflows' (selected), 'Sources', 'Segments', 'Identities', 'Privacy', 'Data Science', and 'Schemas'. The main area displays a tree structure under 'Customer Account - YourNameHere' with nodes for 'consents' (Object), 'marketing' (Object), and 'sms' (Object), which maps to the 'val' (String) attribute. To the right, a table titled 'CONSENTS.MARKETING.SMS.VAL' shows four rows with values: y, n, -, and y. A red box highlights the 'val' column in the table, and another red box highlights the 'marketing.sms.val' attribute in the schema structure.



The screenshot shows the 'Customer Account - YourNameHere' schema structure. The 'Structure' panel on the right shows a tree with 'marketing' (Marketing Preferences), 'email' (Receive email), 'sms' (Receive SMS), 'person' (Person), and 'personalEmail' (Personal Email). The 'Field properties' panel on the right shows the 'Path' as 'consents.marketing.sms.val', 'Required' checked, and an 'Apply' button. A red box highlights the 'Path' field in the properties panel.

## 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.

### 3.5. 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 left sidebar navigation includes Home, Workflows (selected), Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets). The main workspace displays a 'Sources > Data Landing Zone > Add data' screen. The 'Mapping' tab is active, showing a list of mappings between source fields and destination fields. One mapping for 'plan\_name' is highlighted, showing it is currently mapped to '\_devbc.plan.name'. A 'Mapping recommendations' panel on the right lists 'SOURCE FIELD' as 'plan\_name' and suggests three options: 'SELECT ALL FIELDS' (unchecked), 'person.name.lastName' (unchecked), and '\_devbc.plan.name' (checked). A 'Select manually' button is also present in this panel.

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	

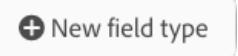
### 3.6.Calculated Fields

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

#### SMS Opt In

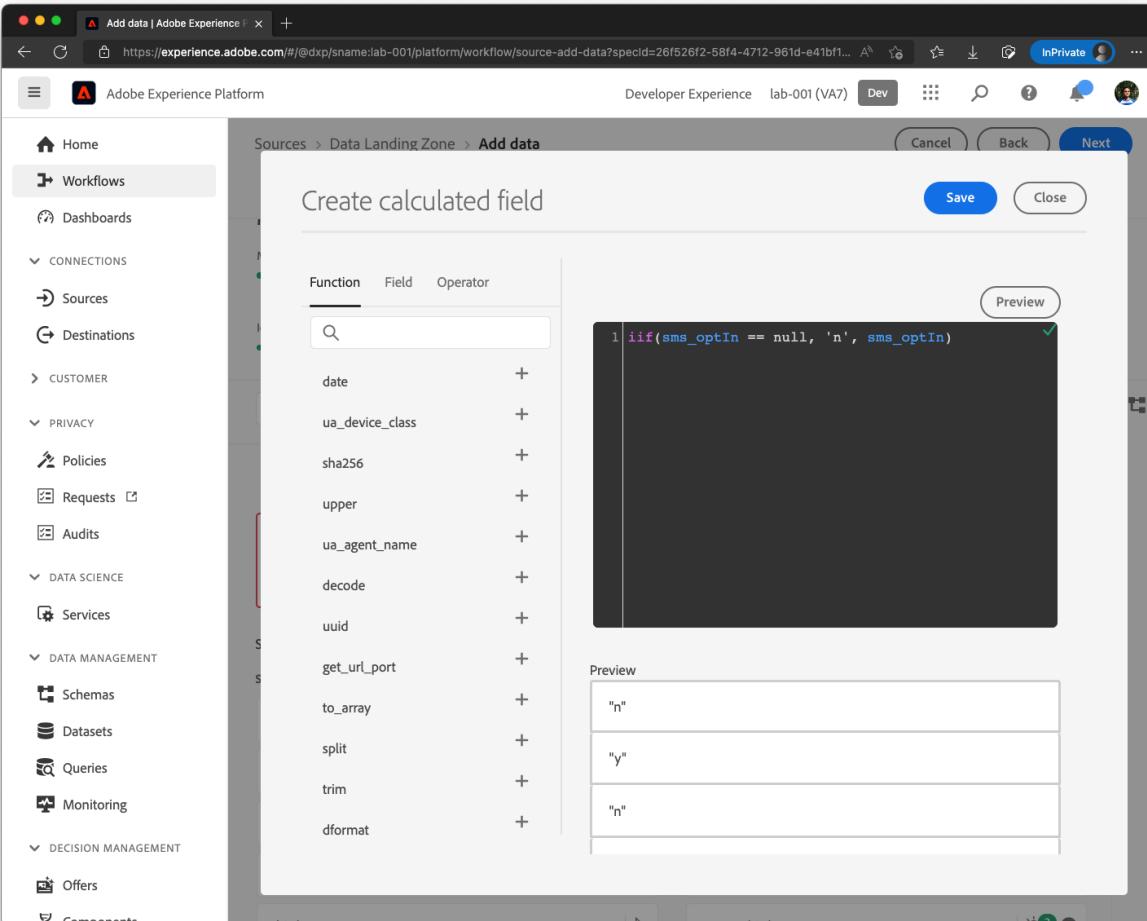
Scroll down to the mapping for XDM attribute **consents.marketing.sms.val**, and remove the mapping.



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**.

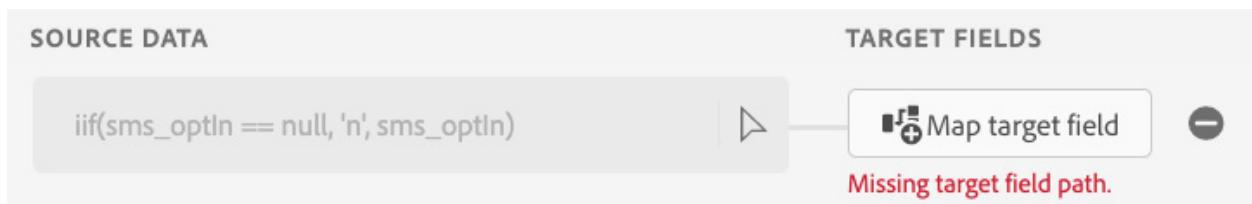


The screenshot shows the 'Add data' interface in the Adobe Experience Platform. On the left, there's a sidebar with various navigation options like Home, Workflows, Sources, Destinations, etc. The main area is titled 'Sources > Data Landing Zone > Add data'. A modal window is open titled 'Create calculated field'. Inside, 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, dformat. To the right, there's a 'Preview' section with a code editor containing the expression 'iif(sms\_optIn == null, 'n', sms\_optIn)'. Below it is a table labeled 'Preview' with three rows: 'n', 'y', and 'n'. At the bottom of the modal are 'Save' and 'Close' buttons.

**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" and "Close" 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. The left sidebar contains navigation links like Home, Workflows, Dashboards, Connections, Customer, Privacy, Data Science, Data Management, Decision Management, Offers, and Components. The main area is titled 'Sources > Data Landing Zone > Add data'. A red box highlights the 'Create calculated field' button. The right side shows a code editor with the following SQL-like 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.

**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

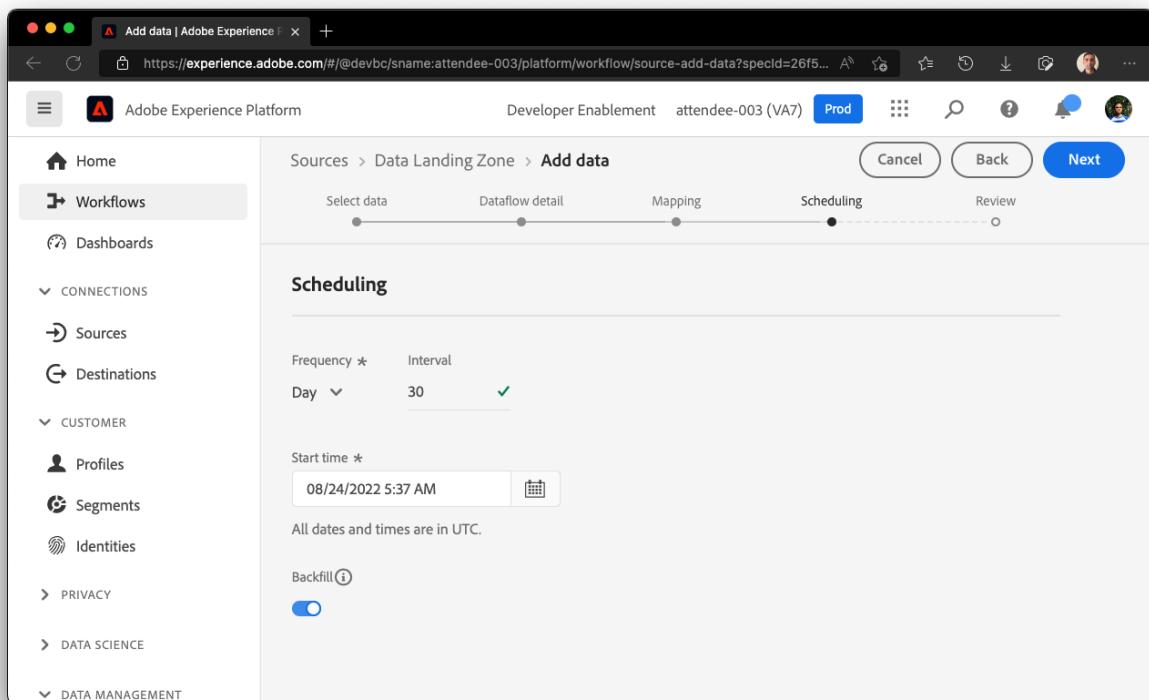
The final mapping set looks like below

#	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	

### 3.7. Schedule

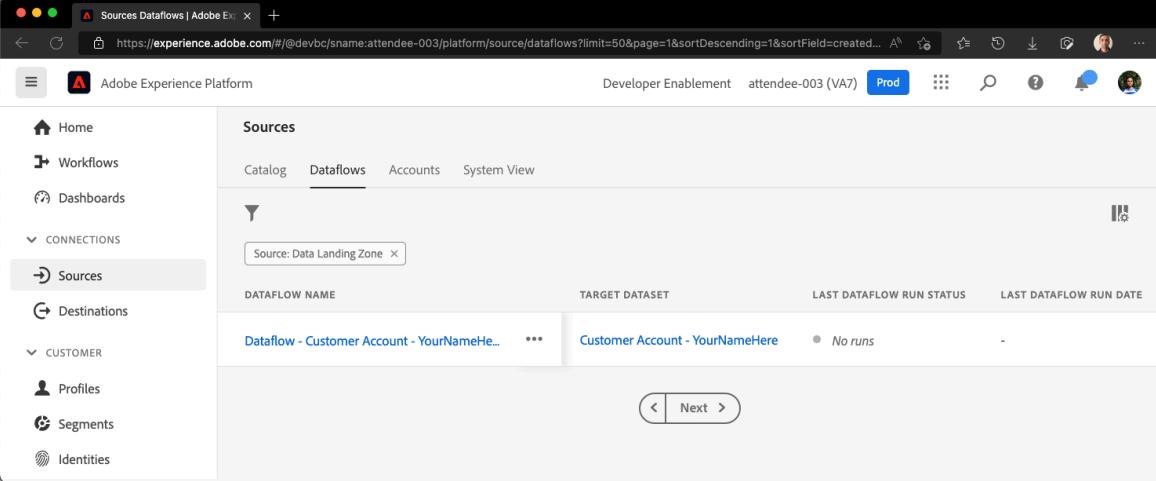
In the **Scheduling** step, set the **Frequency** to **Day** and **Interval** to **30** and leave the **Backfill** turned on. This will schedule the Dataflow to run every 30 days



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 completed, indicated by a green checkmark icon. The 'Connection' section shows an account name 'Customer Account - YourNameHere' and a path 'Lab\_Customer.csv'. The 'Assign dataset and map fields' section shows a target dataset 'Customer Account - YourNameHere' and a schema mapping 'Customer Account - YourNameHere'. The 'Scheduling' section shows a start time '08/24/2022 5:37 AM', frequency 'Day', and interval '30'. Buttons at the bottom right include 'Cancel', 'Back', and a large blue 'Finish' button.

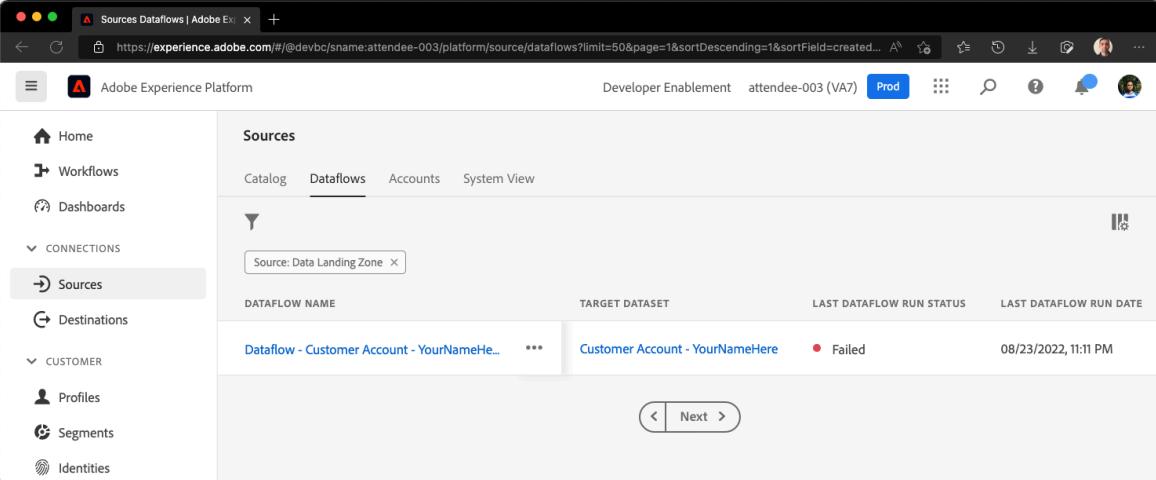
It takes few minutes to create the Dataflow. Once Dataflow is created, you will see the following screenshot. Notice that Last Dataflow Run Status indicates **No runs**. First run will kick off approximately in 15 minutes.



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 (selected). The main area is titled 'Sources' with tabs for Catalog, Dataflows (selected), Accounts, and System View. A search bar at the top says 'Source: Data Landing Zone'. Below it is a table with columns: DATAFLOW NAME, TARGET DATASET, LAST DATAFLOW RUN STATUS, and LAST DATAFLOW RUN DATE. One row is visible: 'Dataflow - Customer Account - YourNameHere...' with target dataset 'Customer Account - YourNameHere', status 'No runs', and date '-'. Navigation buttons '<' and 'Next >' are at the bottom of the table.

### 3.8. 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 shows the same interface after 15 minutes. The dataflow entry now has a red circle icon next to 'Failed' in the 'LAST DATAFLOW RUN STATUS' column, indicating an error. The 'LAST DATAFLOW RUN DATE' shows '08/23/2022, 11:11 PM'. The rest of the table and sidebar are identical to the previous screenshot.

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 Adobe Experience Platform Dataflow interface. On the left, a sidebar navigation includes Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, and DATA SCIENCE. The main area displays 'Dataflow activity' for 'Dataflow - Customer Account - YourNameHere'. It shows 'Records ingested' and 'Records failed' both at 0. Below this is a table with columns: DATAFLOW RUN START, PROCESSING TIME, RECORDS INGESTED, RECORDS FAILED, and STATUS. A single row shows '08/23/2022, 11:11 PM', '1 minute', '-' for both records, and a red dot indicating 'Failed'. To the right is a 'Properties' panel with options to Update dataflow, Edit schedule, enable or disable it, and Delete. The Dataflow name is 'Dataflow - Customer Account - YourNam...'. A large red rectangle highlights the 'STATUS' column header and the 'Failed' entry.

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 a summary of the Dataflow run, including:

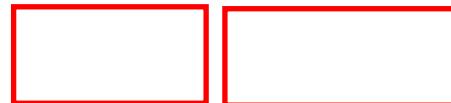
- Records ingested: -
- Records failed: -
- Total files: 1
- Size of data: 5.09 kB
- Status: Failed (red dot)
- Dataflow run start: 08/23/2022, 11:11 PM
- Dataflow run end: 08/23/2022, 11:12 PM
- Partial ingestion: Enabled: 5% Error threshold
- Error diagnostics: Enabled
- Error summary: CONNECTOR-2001-500: Error in processing (parsing, validation or transformation) the copied data.
- Dataflow run IMS org ID: 213A33B362F288B50A495C3B  
ID: 2dba063a-4333-465b-a572-61d52433c00a
- Dataset: Customer Account - YourNameHere

Below this, the "Dataflow run errors" section shows:

- Records failed (radio button selected)
- Records skipped
- Parsing errors occurred when converting or validating the data. Successfully ingested 0 rows, review the associated diagnostic files for additional details.
- Table of errors:

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

At the bottom right of the errors table are "Preview error diagnostics" and "Download" buttons.



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

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 a summary of the Dataflow run, including:

- Records ingested: -
- Records failed: -
- Total files: 1
- Size of data: 5.09 kB
- Status: Failed (red dot)
- Dataflow run start: 08/23/2022, 11:11 PM
- Dataflow run end: 08/23/2022, 11:12 PM
- Partial ingestion: Enabled: 5% Error threshold
- Error diagnostics: Enabled
- Error summary: CONNECTOR-2001-500: Error in processing (parsing, validation or transformation) the copied data.

Below this, there are tabs for Dataflow run errors (selected) and Files. Under Dataflow run errors, there is a link to Preview error diagnostics, which is highlighted with a red box. The table below shows error details:

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

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 Dataflow Run interface. The left sidebar has a tree view with categories like Home, Workflows, Dashboards, CONNECTIONS (Sources selected), DESTINATIONS, CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring), DECISION MANAGEMENT (Offers, Components), and ADMINISTRATION. The main content area shows a "Dataflow Run 2dba063a-4333-465b-a572-61d52433c00a" page under Sources > Dataflow - Customer Account - YourNameHere. A modal window titled "Error diagnostics preview" is open, displaying a table of errors. The table has columns: FILE NAME, ERROR CODE, COLUMN NAME, and DESCRIPTION. There are 10 rows, each corresponding to a file named "Lab\_Customer.csv" with an INGEST-1517-400 error code, a person.birthDayAndMonth column name, and a description stating "The value does not conform to the specified regex pattern." At the bottom of the modal, it says "INGEST-1517-400" and "The value does not conform to the specified regex pattern."

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	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.

## 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 sections like 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 'Sources' section is currently selected. In the center, a modal window titled 'Error diagnostics preview' is open, displaying a table of errors. The table has columns: FILE NAME, ERROR CODE, COLUMN NAME, and DESCRIPTION. There are 10 entries, all from 'Lab\_Customer.csv'. The first 9 entries have an 'ERROR CODE' of 'INGEST-1517-400' and the last one has 'MAPPER-3700-199'. The 'COLUMN NAME' is 'person.birthDayAndMonth' for the first 9 and 'createDate' for the last one. The 'DESCRIPTION' for the first 9 entries is 'The value does not conform to the specified regex pattern'. The last entry's description starts with 'Error transforming data for destination path \_devbc.account.createDate. Details: Unable to convert Created on 2022-04-22T19:34:17Z to schema type DATE...'. A red rectangle highlights the top of the modal window.

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...

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. On the left, a sidebar navigation menu 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 a "Sources > Dataflow - Customer Account - YourNameHere > Dataflow Run 2dba063a-4333-465b-a572-61d52433c0a" page titled "Error diagnostics preview". It lists several errors related to date conversion issues. To the right is a table of data with columns "Org" and "Value". A red box highlights the "billing\_zip\_code" row, which has a value of "93794".

Org	Value
billing_city	Fresno
billing_state	CA
billing_zip_code	93794
plan_id	m3
plan_name	pro
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.

### 3.9. Correct the Dataflows

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

The screenshot shows the Dataflow run overview page for a specific run. 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). The main content area displays the following details:

Records ingested	Records failed	Total files	Size of data	Dataflow run ID
-	-	1	5.09 kB	2dba063a-4333-465b-a572-61d52433c00a

Status: Failed (Dataflow run start: 08/23/2022, 11:11 PM; Dataflow run end: 08/23/2022, 11:12 PM)

Partial ingestion: Enabled: 5% Error threshold

Error diagnostics: Enabled

Error summary: CONNECTOR-2001-500: Error in processing (parsing, validation or transformation) the copied data.

Below this, there are tabs for Dataflow run errors and Files. Under Dataflow run errors, there are two radio buttons: Records failed (selected) and Records skipped.

Click on the Update dataflow

Records ingested      Records failed

0      0

Last 7 days

DATAFLOW RUN START	PROCESSING TIME	RECORDS INGESTED	RECORDS FAILED
08/23/2022, 11:11 PM	1 minute	-	-

Properties

- Update dataflow
- Edit schedule
- Enabled
- Delete

Dataflow name  
Dataflow - Customer Account - YourNameHere

Description

Source data  
dlz-user-container/project/PIPELINE/Lab\_Customer.csv

Target dataset  
Customer Account - YourNameHere

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

Select data

Preview:

/ > dlz-user-container > project > PIPELINE > Lab\_Customer.csv

Data format \*

Delimited

Delimiter

Sample data

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 the current step is 'Dataflow detail'. The 'Dataset details' section includes a 'Target dataset' dropdown set to 'Customer Account - YourNameHere', an 'Advanced search' button, and a 'Partial ingestion' toggle which is enabled. An 'Error threshold %' slider is set to 5. The 'Dataflow details' section shows a 'Dataflow name' input field containing 'Dataflow - Customer Account - YourNameHere', a 'Description' input field, and an 'Alerts' section with a note to check options for 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 links for Home, Workflows (selected), Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets, Queries). The main area is titled 'Sources > Data Landing Zone > Add data' and is at the 'Mapping' step. It displays a progress bar for 'Mapped fields' (24 of 24) and 'Required fields' (1 of 1). Below this, 'Identity fields' (0 of 2) are shown. An 'Errors: 0' message is present. A search bar and various mapping tools are available. The mapping table lists four source fields with their target XDM fields:

SOURCE DATA	TARGET FIELDS
concat( date_part("month", date(birth_Date))...)	person.birthDayAndMonth
date_part("year", date(birth_Date))	person.birthYear
iif(sms_optin == null, 'n', sms_optin)	consents.marketing.sms.val
lastName	person.name.lastName

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 contains tabs for Function, Field, and Operator. The Function tab is active, showing 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, dformat, and lpad. A code editor window displays 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"))
```

A preview window below shows the results of the expression for four different dates:

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

Click save. The mapping will look like this

The screenshot shows the 'Mapping schema' section of the Dataflow Mapping interface. On the left, under 'SOURCE DATA', there is a code snippet: `concat( lpad(date_part("month", date(birth_Date)).toString(), 2, '0') , birthYear )`. An arrow points from this source to the 'TARGET FIELDS' section on the right, where the target field is `person.birthDayAndMonth`.

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

The screenshot shows the 'Scheduling' step in the Dataflow process. The top navigation bar has tabs: 'Select data', 'Dataflow detail', 'Mapping', 'Scheduling' (which is highlighted with a red box), and 'Review'. The 'Frequency' dropdown is set to 'Day' and 'Interval' is set to '30'. The 'Start time' is set to '08/24/2022 9:46 PM'. A note says 'All dates and times are in UTC.' Below that is a 'Backfill' toggle switch, which is turned off.

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 shows the 'Sources > Data Landing Zone > Add data' path. A progress bar at the top indicates steps: Select data (done), Dataflow detail (done), Mapping (done), Scheduling (done), and Review (in progress). The 'Review' section is divided into three cards:

- Connection:** Account name: Customer Account - YourNameHere, Source platform: Data Landing Zone, Path: -, Columns: 23. Status: Connected.
- Assign dataset and map fields:** Target dataset: Customer Account - YourNameHere. Status: Dataset assigned.
- Scheduling:** Start time: 08/24/2022 6:11 AM, Frequency: Minute, Interval: 15. Status: Scheduled.

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

**NOTE**

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

### 3.10. 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 interface for managing dataflows. On the left, there's a sidebar with various navigation options like Home, Workflows, Dashboards, and Connections (Sources, Destinations). The main area is titled 'Sources' and has tabs for Catalog, Dataflows, Accounts, and System View. The 'Dataflows' tab is currently selected and highlighted with a red box. Below this, there's a search bar with 'Source: Data Landing Zone'. Underneath, a table lists a single dataflow entry:

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

On the right, a detailed properties panel is open for the selected dataflow. It includes fields for Dataflow name (set to 'Dataflow - Customer Account - YourNam...'), Description (empty), and Source data (set to 'dlz-user-container/project/PIPELINE/CUSTOMER/'). There are also buttons for Update dataflow, Edit schedule (which is the target of the red arrow in the image), Enabled (which is checked), and Delete.

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 various navigation options like Home, Workflows, Dashboards, and Sources. The Sources section is currently selected. In the main area, there's a list of dataflows under a 'Source: Data Landing Zone' filter. One dataflow named 'Dataflow - Customer Account' is visible. A modal window titled 'Edit schedule' is open over the list. Inside the modal, there are fields for 'Frequency \*' (set to 'Minute') and 'Interval' (set to '15'). Below these, there's a 'Start time \*' field set to '08/24/2022 9:46 PM'. A note says 'All dates and times are in UTC.' There's also a 'Backfill' toggle switch which is turned off. At the bottom of the modal are 'Cancel' and 'Save' buttons. A red box highlights the 'Frequency' and 'Interval' fields.



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 Sources Dataflows interface. The left sidebar has a 'Sources' section selected. The main area displays a dataflow named 'Dataflow - Customer Account - YourNameHere...' with a target dataset 'Customer Account - Y...'. A red box highlights the 'Description' field in the properties panel, which contains the placeholder 'Dataflow - Customer Account - YourNam...'. The properties panel also includes options for updating the dataflow, editing the schedule, enabling it, and deleting it.

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 for a specific dataflow named "Dataflow - Customer Account - YourNameHere". The left sidebar is collapsed, and the main area displays the "Dataflow activity" section. A red box highlights the "Records ingested" section, which is currently empty. Below it, a table shows four dataflow runs over the last 7 days. The first three runs are marked as "Success" with 0 records ingested and 0 records failed. The fourth run, from 08/24/2022, 2:46 PM, is marked as "Failed" with 0 records ingested and 0 records failed. The properties panel on the right shows the dataflow is enabled and has a description and source target defined.

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 interface. On the left, a sidebar navigation includes 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 displays a Dataflow activity for 'Dataflow - Customer Account - YourNameHere'. It shows 'Records ingested' as 20 and 'Records failed' as 0. Below this is a table of dataflow runs over the last 7 days:

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

The right side of the interface contains a Properties panel with options to Update dataflow, Edit schedule, 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), and Status (Enabled).

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



## **Adobe Experience Platform**

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

Adobe Experience Platform Bootcamp Deep Dive Edition

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

The screenshot shows the Dataflow run overview page for a specific run. The left sidebar contains navigation links for Home, Workflows, Dashboards, Connections (Sources selected), Customer (Profiles, Segments, Identities), Privacy, Data Science, Data Management (Schemas, Datasets, Queries, Monitoring), and Decision Management (Offers, Components). The main content area displays the following details:

Records ingested	Records failed	Total files	Size of data
20	0	1	5.09 kB

Status: Success. Dataflow run start: 08/24/2022, 3:46 PM. Dataflow run end: 08/24/2022, 3:47 PM.

Partial ingestion: Enabled: 5% Error threshold. Error diagnostics: Enabled. Error summary: CONNECTOR-2001-500: Error in processing (parsing, validation or transformation) the copied data.

Dataflow run errors tab (selected):

ERROR CODE	RECORD COUNT	DESCRIPTION
MAPPER-3700-199	0	Field transformation warning

Files tab (disabled):

FILE NAME	FILE PATH
Lab_Customer_002.csv	-

Download button: Download

IMS org ID: 213A338362F288B50A  
Dataflow run ID: 97859a00-7be9-488c-9530-b55699e4f452  
Dataset: Customer Account - YourNameHere

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

The screenshot shows the Dataflow run overview page for a specific run. The left sidebar contains navigation links for Home, Workflows, Dashboards, Connections (Sources selected), Customer (Profiles, Segments, Identities), Privacy, Data Science, Data Management (Schemas, Datasets, Queries, Monitoring), and Decision Management (Offers, Components). The main content area displays the following details:

Records ingested	Records failed	Total files	Size of data
20	0	1	5.09 kB

Status: Success. Dataflow run start: 08/24/2022, 3:46 PM. Dataflow run end: 08/24/2022, 3:47 PM.

Partial ingestion: Enabled: 5% Error threshold. Error diagnostics: Enabled. Error summary: CONNECTOR-2001-500: Error in processing (parsing, validation or transformation) the copied data.

Dataflow run errors tab (disabled):

ERROR CODE	RECORD COUNT	DESCRIPTION
MAPPER-3700-199	0	Field transformation warning

Files tab (selected):

FILE NAME	FILE PATH
Lab_Customer_002.csv	-

Download button: Download

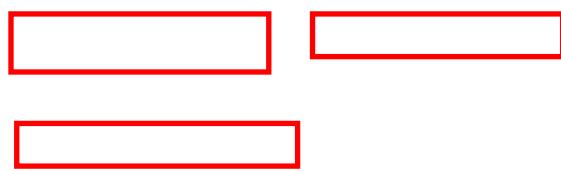
IMS org ID: 213A338362F288B50A  
Dataflow run ID: 97859a00-7be9-488c-9530-b55699e4f452  
Dataset: Customer Account - YourNameHere

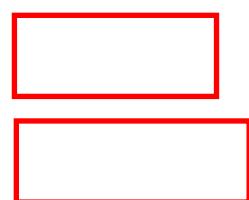
The lab is now complete.

## 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.







## 1. 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**

## 2. 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

## 3. 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.

The screenshot shows the Adobe Experience Platform Sources Catalog interface. The left sidebar navigation includes Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY (Policies, Requests, Audits), DATA SCIENCE (Services, Schemas, Datasets, Queries, Monitoring), and DECISION MANAGEMENT. The main area is titled 'Sources' and has tabs for Catalog, Dataflows, Accounts, and System View. Under 'CATEGORIES', 'Cloud storage' is selected and highlighted with a red box. The 'Data Landing Zone' source is listed, and its 'Add data' button is also highlighted with a red box.

### 3.1. 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**.

The screenshot shows the 'Add data' screen in Adobe Experience Platform. The left sidebar navigation is identical to the previous screenshot. The main area is titled 'Sources > Data Landing Zone > Add data'. The 'Select data' tab is active. In the 'Selected data:' section, 'Lab\_Historical\_Orders.json' is selected. In the 'Preview:' section, the path is shown as / > dlz-user-container > project > PIPELINE > Lab\_Historical\_Orders.json. The 'Data format' is set to JSON, 'XDM compliant' is set to No, and 'Compression type' is set to None. On the right, there is a detailed JSON schema tree and a preview table of data. The table has columns: SHIPPINGCITY, PRODUCTS.QUANTITY, and PRODUCTS.PROD. The data rows are:

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

## 3.2. 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**.

The screenshot shows the 'Add data' 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, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring), DECISION MANAGEMENT (Offers, Components), and ADMINISTRATION (Alerts). The main panel is titled 'Sources > Data Landing Zone > Add data' and shows a progress bar with steps: Select data (done), Dataflow detail (done), Mapping (not started), Scheduling (not started), and Review (not started). The 'Dataset details' section includes fields for 'Target dataset' (radio buttons for Existing dataset and New dataset, with New dataset selected), 'Output dataset name' (text input: 'Orders - YourNameHere'), 'Description' (empty text input), 'Schema' (text input: 'dep: Orders'), and 'Profile dataset' (checkbox checked). The 'Dataflow details' section includes fields for 'Dataflow name' (text input: 'Orders - Backfill - YourNameHere'), 'Description' (empty text input), and 'Alerts' (checkbox). A red box highlights the 'Profile dataset' checkbox and the 'Dataflow name' field.

**Tip**

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

### 3.3. 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 'Mapping' step is currently active. A progress bar indicates 29 of 29 fields are mapped, 1 of 3 required fields are missing, and 2 of 2 identity fields are mapped. An error message box states: "There was error(s) preparing mappings. The following unmapped paths are required: \_id, timestamp. The following unmapped paths are required: productListItems.SKU." The source data is from 'Lab\_Historical\_Orders.json' and the target schema is 'Orders'.

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

### 3.4. Remove the following mappings

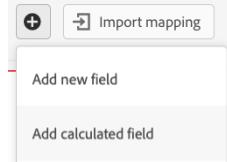
- Mapping to XDM attribute `billing.address.lastVerifiedDate`



## 3.5. 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.

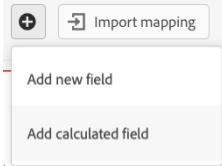
The screenshot shows the 'Create calculated field' interface. On the left is a sidebar with various platform components like Workflows, Sources, Destinations, and Data Science. The main area has tabs for Function, Field, and Operator. The Function tab is active, showing a search bar with 'concat' and a list of functions: date, ua\_device\_class, sha256, upper, ua\_agent\_name, decode, uuid, get\_url\_port, to\_array, split, trim, dformat. Below the search bar is a preview window showing the expression and its results. The preview results are:  
1 concat(orderID, "-", lastOrderStatusUpdate)  
"bfca4a37-d58b-4061-89e9-0906b71d0a45-2022-06-13T16:27:03Z"  
"471fb07f-3ee7-48bb-bf11-208aab2e6180-2022-06-06T13:22:47Z"  
"c0441e60-4c04-45c1-a621-1d4b6fe6d038-2022-06-11T10:54:31Z"  
"d2d490c7-6640-476f-a52b-d75d650c916-2022-06-14T04:29:07Z"

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 'Add data' workflow in the Adobe Experience Platform. The left sidebar includes sections like 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 panel shows a progress bar with steps: Select data, Dataflow detail, Mapping (selected), Scheduling, and Review. Below the progress bar, it says 'Mapped fields 31 of 31' and 'Required fields 2 of 3'. Under 'Identity fields', it shows '2 of 2'. A red warning box at the top right indicates 'Errors: 2' with the 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 lists several mappings from source fields to target fields, including 'order\_devbc.acqSource' which is highlighted with a red border.

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 'order.\_devbc.acqSource' entry in the target fields list, with a red message below stating 'Duplicated target fields found. Map one row per target field.'

### 3.6. Add new pass-through mappings

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**

### 3.7. Product List Items

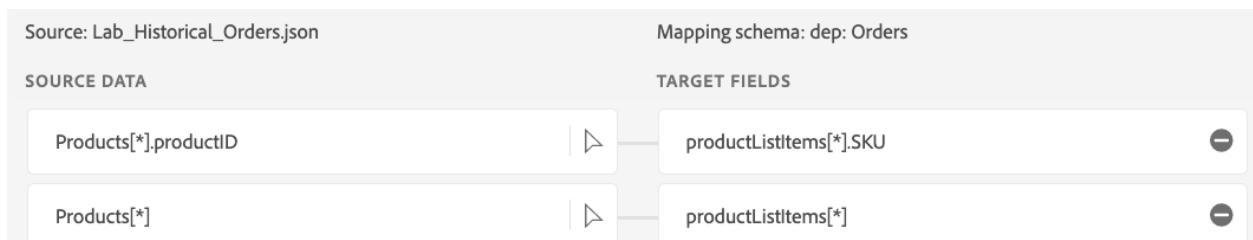
We will remove the mappings to XDM attributes **productListItems[\*].quantity** and **productListItems[\*].currencyCode**. Change the other mappings to **productListItems[\*]**

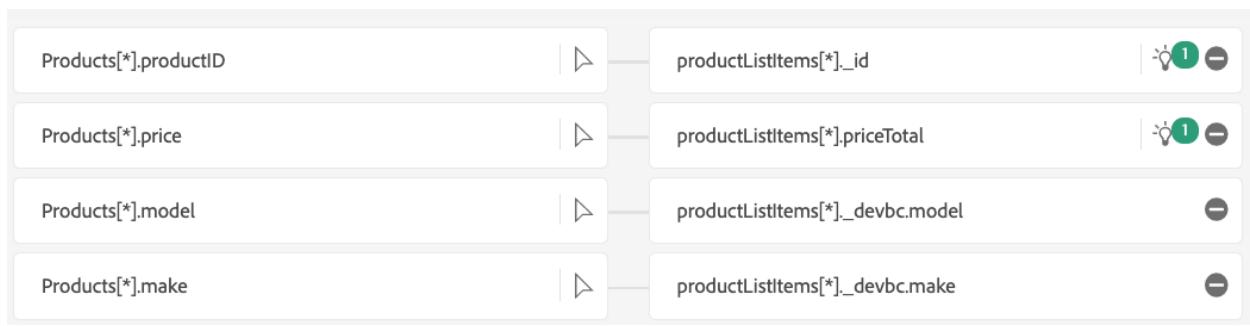
**NOTE**

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.





### 3.8. 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	

## 3.9. 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

The screenshot shows the 'Add data' preview screen in the Adobe Experience Platform. On the left, a navigation sidebar lists various platform modules like Home, Workflows, Sources, Destinations, Customer, Data Science, and Administration. The main area is titled 'Preview' and contains a search bar and a tree view of the data structure. The tree view shows a 'dep: Orders' node expanded, revealing sub-nodes for 'dpx' (Object), 'billing' (Object), 'order' (Object), 'productListItems' (Object[]), 'shipping' (Object), 'store' (Object), '\_id' (String), 'eventMerged' (String), 'eventType' (String), 'identityMap' (Type), 'producedBy' (String), and 'timestamp' (DateTime). To the right of the tree view is a table displaying data rows. The columns are labeled '\_DXP.CUSTOMERID', '\_DXP.PERSONALEMAIL', 'PRODUCTLISTITEMS.\_DXP.MAKE', and 'PRODUCTLISTITEMS'. The table contains 12 rows of data, each with a unique ID, email, make, and version number (v1, v2, v3, v4).

_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
47156525	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 'Add data' interface in the Adobe Experience Platform. The left sidebar contains a navigation menu with sections such as Home, Workflows, Sources, Destinations, Customer, Data Science, Data Management, Decision Management, and Administration. The main area is titled 'Sources > Data Landing Zone > Add data'. A 'Preview' tab is selected, showing a tree view 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	-

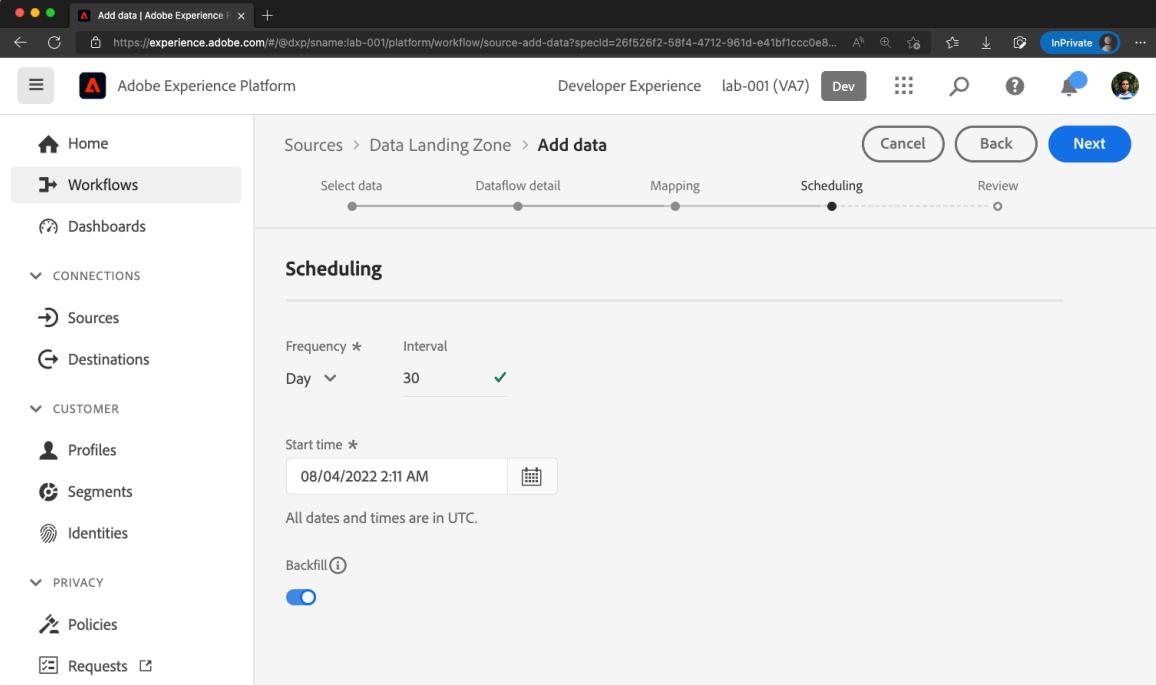
Completed mapping will look similar to the following screenshot

The screenshot shows the 'Add data' interface in the Adobe Experience Platform, specifically the 'Mapping' step. The left sidebar contains a navigation menu with sections such as Home, Workflows, Sources, Destinations, Customer, Data Science, Data Management, Decision Management, and Administration. The main area is titled 'Sources > Data Landing Zone > Add data'. The 'Mapping' tab is active. A summary bar at the top indicates 'Mapped fields: 33 of 33', 'Required fields: 3 of 3', and 'Identity fields: 0 of 2'. An error message 'Errors: 0' is displayed. Below this, a table lists source fields from 'dep: Orders' and their corresponding target fields in the 'productListItems' dataset. The table includes columns for 'SOURCE DATA' and 'TARGET FIELDS'.

SOURCE DATA	TARGET FIELDS
"inStore"	order_dxp.acqSource
Products	productListItems
billingState	billing.address.state
orderStatus	eventType
concat(orderID, " ", lastOrderStatusUpdate)	_id
lastOrderStatusUpdate	timestamp
Products[*].productID	productListItems[*].SKU

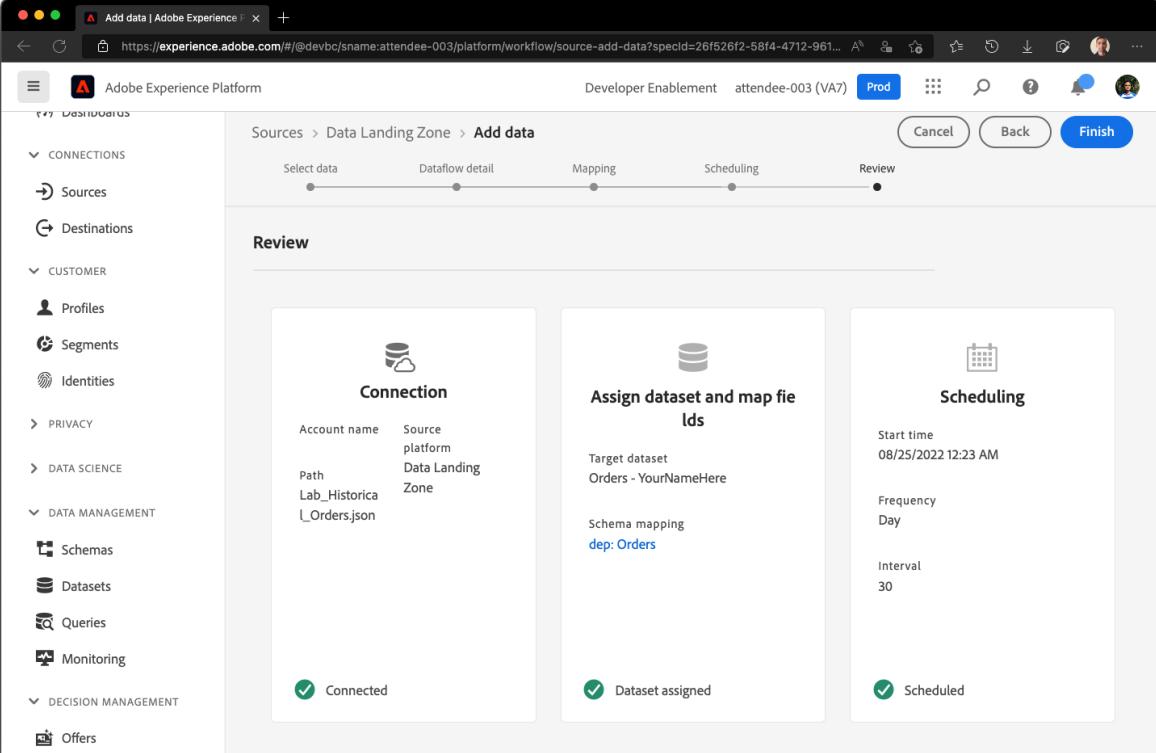
## 3.10. Schedule

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 in the Adobe Experience Platform. The left sidebar shows various platform services like Home, Workflows, Dashboards, and Sources. The main panel is titled 'Sources > Data Landing Zone > Add data' and is at the 'Scheduling' step. A progress bar indicates the steps: Select data, Dataflow detail, Mapping, Scheduling, and Review. The 'Scheduling' step is active. The configuration shows a frequency of 'Day' and an interval of '30'. The start time is set to '08/04/2022 2:11 AM'. There is a 'Backfill' toggle switch turned off. The 'Review' button is visible at the bottom right.

Review the flow and click Finish



The screenshot shows the 'Add data' workflow in the Adobe Experience Platform, now at the 'Review' step. The left sidebar is identical to the previous screenshot. The main panel shows the completed steps: 'Connection' (Account name: 'Developer Enablement', Source platform: 'Data Landing Zone', Path: 'Lab\_HistoricAll\_Orders.json'), 'Assign dataset and map files' (Target dataset: 'Orders - YourNameHere', Schema mapping: 'dep: Orders'), and 'Scheduling' (Start time: '08/25/2022 12:23 AM', Frequency: 'Day', Interval: '30'). Each step has a green checkmark indicating it is successful. The 'Finish' button is visible at the bottom right.

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".



# **Adobe Experience Platform**

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

Adobe Experience Platform Bootcamp Deep Dive Edition

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

The screenshot shows the Adobe Experience Platform Dataflows interface. On the left, a sidebar navigation includes 'CONNECTIONS' (Sources, Destinations), 'CUSTOMER' (Profiles, Segments, Identities), 'DATA SCIENCE', 'DATA MANAGEMENT', and 'Schemas'. The main area is titled 'Sources' and has tabs for Catalog, Dataflows, Accounts, and System View. Under 'Dataflows', there is a search bar 'Source: Data Landing Zone X'. A table lists two dataflows:

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

At the bottom right of the table are navigation buttons: '<' and '>'. The status 'Success' is indicated by a green dot.

### 3.11. Check scheduled execution

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

The screenshot shows the Adobe Experience Platform Dataflows interface. The sidebar navigation is identical to the previous screenshot. The main area is titled 'Sources' and has tabs for Catalog, Dataflows, Accounts, and System View. Under 'Dataflows', there is a search bar 'Source: Data Landing Zone X'. A table lists three dataflows, all of which have run successfully:

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

At the bottom right of the table are navigation buttons: '<' and '>'. The status 'Success' is indicated by a green dot.

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

Partial ingestion: Enabled: 5% Error threshold

Error diagnostics: Enabled

Error summary: -

Dataflow run ID: 94354d08-682f-4fb5-be6f-0a2ea5232d4f

Dataset: dep: Orders learner

Dataflow run errors

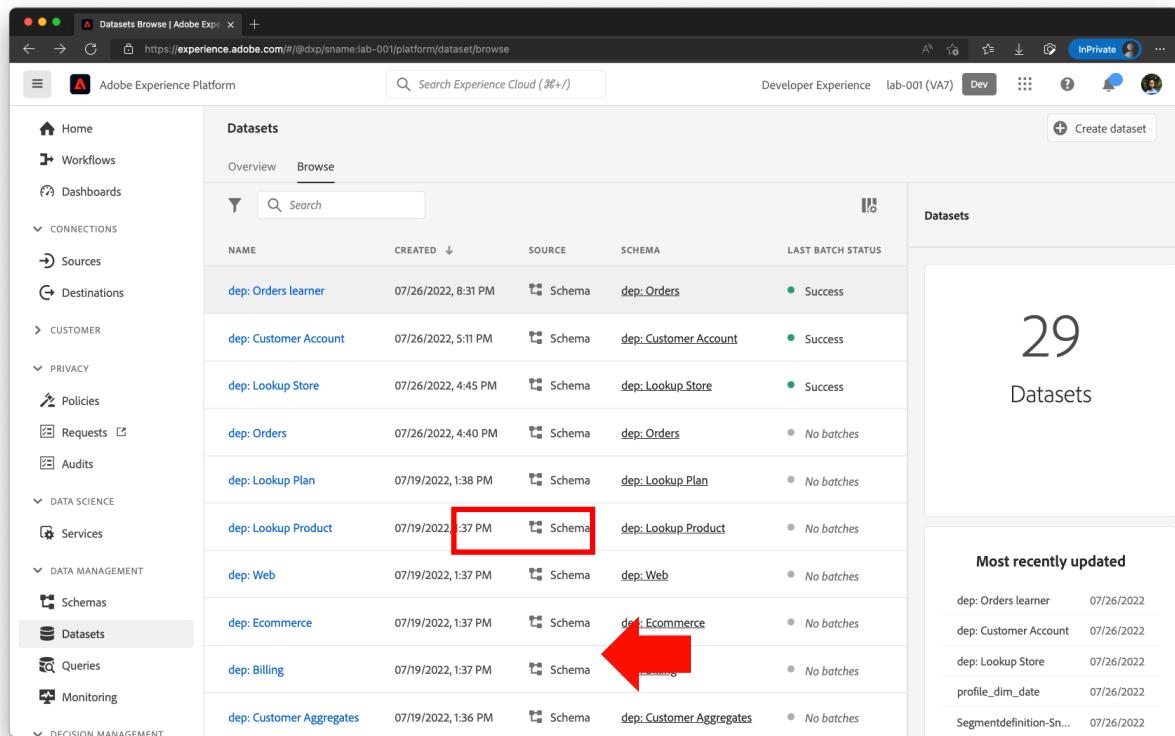
Records failed (selected)

Records skipped

ERROR CODE	DESCRIPTION

## 3.12. Verify the data

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



Datasets Browse | Adobe Experience Platform

https://experience.adobe.com/#/d@xp/sname:lab-001/platform/dataset/browse

Adobe Experience Platform InPrivate Dev

Search Experience Cloud (⌘+F)

Developer Experience lab-001 (VA7)

Create dataset

Home Workflows Dashboards CONNECTIONS Sources Destinations CUSTOMER PRIVACY Policies Requests Audits DATA SCIENCE Services DATA MANAGEMENT Schemas Datasets Queries Monitoring DECISION MANAGEMENT

Datasets Overview Browse

Search

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

29 Datasets

Most recently updated

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, including total records, ingested records, batches, and failed batches. It also shows the size of data in the previous month (58.75 kB). A chart titled 'Ingested records' tracks daily ingestions from July 20 to July 26. The right panel provides detailed information about the dataset, including its name ('dep: Orders learner'), description, dataset ID ('62e0b18f7303841c072d31ae'), table name ('dep\_orders\_learner'), profile ('Profile 1'), schema ('dep: Orders'), source ('Schema'), and creation date ('07/26/2022, 8:31 PM').

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

The screenshot shows the 'Preview dataset' modal for the 'dep: Orders learner' dataset. The modal has a search bar at the top and a tree view on the left showing the dataset structure: 'dep: Orders' containing '\_dpx', 'billing', 'order', 'productListItems', 'shipping', 'store', and '\_id'. To the right is a table of data rows with columns for 'ORDER.PAYMENT.CURRENCYCODE' and 'ORDER.ORDERID'. The table lists 10 rows of USD currency with various order IDs. At the bottom of the modal, the date '07/26/2022, 8:31 PM' is visible. A large red arrow points to the title bar of the modal window.

ORDER.PAYMENT.CURRENCYCODE	ORDER.ORDERID
USD	b438b27f-6cf5-4388-a147-4b9c14ac2aa
USD	724d1bf3-cccd-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



## 1. 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**

## 2. 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

### 3. Lab Tasks A – 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

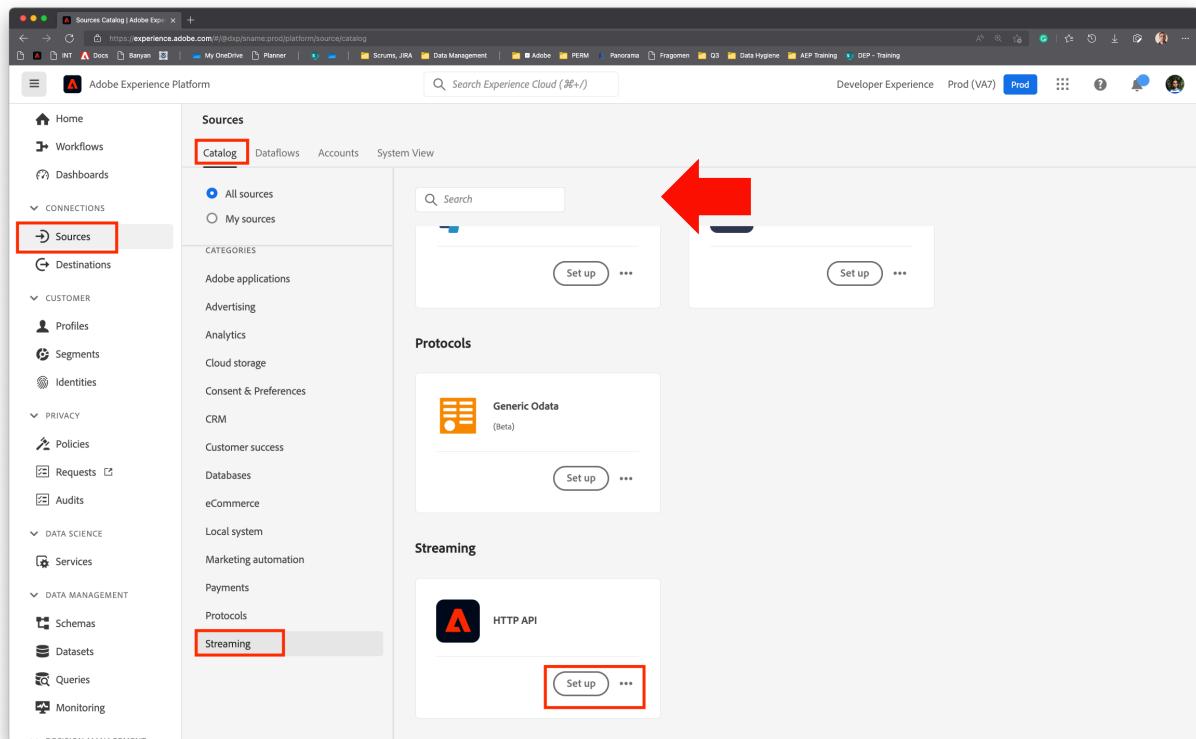
1. Lab 2 is successfully completed

#### Steps

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

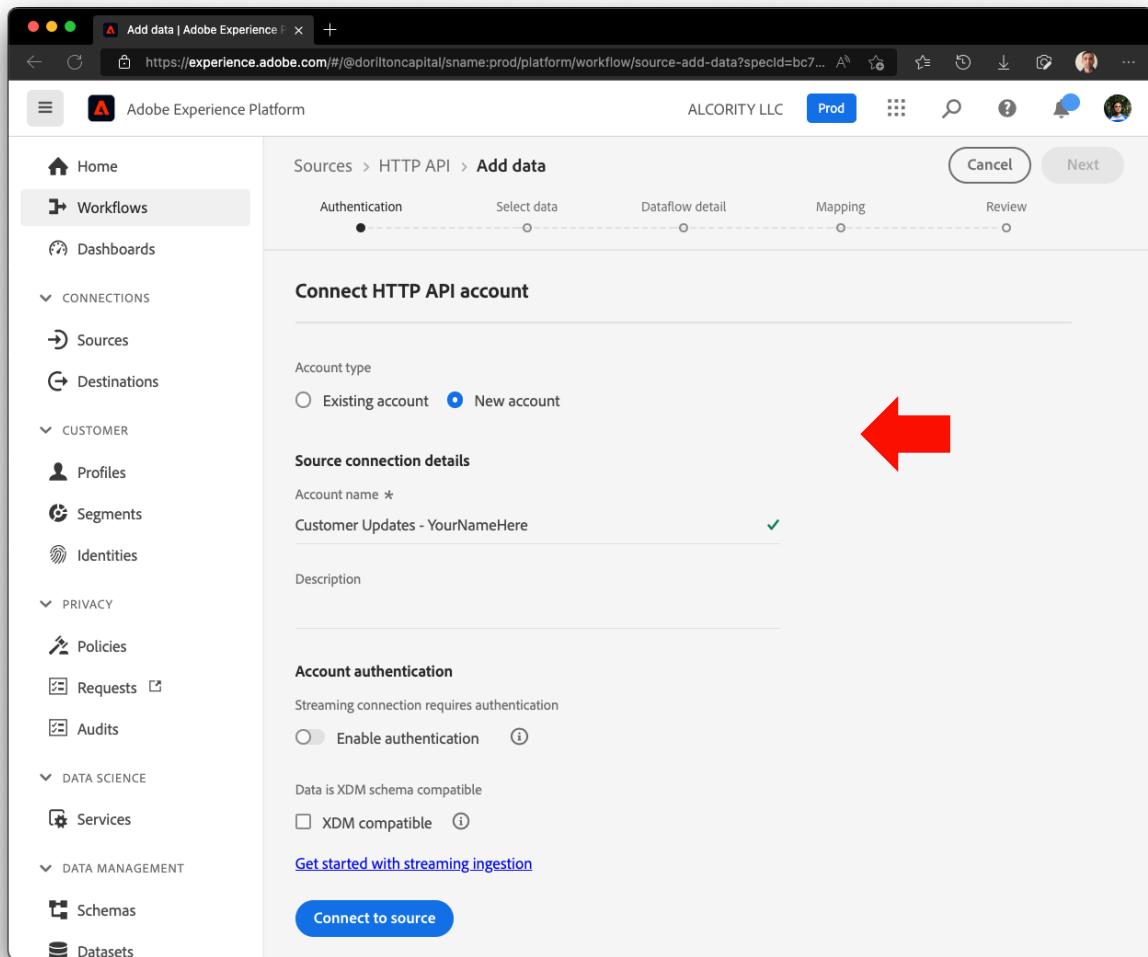
##### Tip

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.



### 3.1. 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.

The screenshot shows the Adobe Experience Platform interface for adding data via an HTTP API. The left sidebar is collapsed, and the main area shows the 'Sources > HTTP API > Add data' workflow. The 'Authentication' step is active, indicated by a solid dot under it. The 'Select data', 'Dataflow detail', 'Mapping', and 'Review' steps are shown as dashed circles. On the right, there are 'Cancel' and 'Next' buttons. A large red arrow points upwards towards the 'Next' button. The 'Source connection details' section includes fields for 'Account name' (Customer Updates - YourNameHere) and 'Description' (which is highlighted with a red box). The 'Account authentication' section has an 'Enable authentication' option selected. At the bottom, there are links for 'Get started with streaming ingestion' (with a 'Connected' status) and 'Sign in with different credentials'.

## 3.2. 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

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 content area is titled 'Add data' under 'HTTP API'. The current step is 'Select data', indicated by a solid dot on the progress bar. The 'Source data schema' section contains instructions: 'Upload sample JSON file to define source schema. Acceptable file size is up to 1GB.' Below this is a 'Drag and drop files' area with a dashed border, featuring icons for a photo, a document, and a folder. To the right, a 'Preview sample data:' section shows a small open box icon and the message 'No data file has been uploaded'. At the bottom, it says 'Upload sample JSON file to define source schema.' The top right of the page shows 'Developer Enablement attendee-003 (VA7) Prod'.

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'. A progress bar at the top indicates 'Authentication' (done), 'Select data' (done), 'Dataflow detail' (not started), 'Mapping' (not started), and 'Review' (not started). The 'Source data schema' section shows a file named 'Lab\_Single\_Customer.json' (JSON) has been uploaded. The 'Preview sample data:' section displays the following JSON structure:

```
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
```

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**.

The screenshot shows the 'Add data' workflow configuration screen in 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 progress bar shows steps: Authentication (done), Select data (done), Dataflow detail (done), Mapping (not started), and Review (not started). The 'Dataset details' section has 'Target dataset' set to 'Existing dataset' (selected) and 'Customer Account - YourNameHere' chosen from a dropdown. The 'Profile dataset' toggle is turned on. The 'Dataflow details' section has 'Dataflow name' set to 'Live Customer Updates - YourNameHere'. The 'Alerts' section contains a checkbox for 'Sources Dataflow Lack of Ingestion'.

### 3.3. 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 Adobe Experience Platform. The current step is 'Mapping'. The interface displays a progress bar for 'Mapped fields' (23 of 23) and 'Required fields' (1 of 1). Below this, an 'Identity fields' section shows 2 of 2. A warning message indicates 'Errors: 2'. A red box highlights the 'Import mapping' button. Another red box highlights an error message box containing the text: '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 'TARGET FIELDS' section shows four 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 status indicator (green with a '1') and a delete icon.

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.

The screenshot shows the Adobe Experience Platform interface with the URL <https://experience.adobe.com/#/@devbc/sname:attendee-003/platform/workflow/source-add-data?specId=bc7b00d6-623a-4dfc-9...>. The left sidebar is open, showing categories like Home, Workflows (which is selected), Dashboards, CONNECTIONS, SOURCES, DESTINATIONS, CUSTOMER, PRIVACY, DATA SCIENCE, and DATA MANAGEMENT. The 'Workflows' section contains several items: 'Dataflow - Customer Account - YourNameHere...', 'Web Streaming flow', 'Customer Streaming flow', 'Orders Streaming flow', and 'Web Feed AWS S3 flow'. The 'Dataflow - Customer Account - YourNameHere...' item is highlighted with a red box. The main content area displays the 'Import mapping' dialog. This dialog has a search bar at the top labeled 'Search dataflow'. Below it is a table with columns: 'DATAFLOW NAME', 'TARGET DATASET', and 'LAST DATAFLO...'. The table rows are:

DATAFLOW NAME	TARGET DATASET	LAST DATAFLO...
Dataflow - Customer Account - YourNameHe...	Customer Account - YourNameHere	Success
Web Streaming flow	dep: Web	Processing
Customer Streaming flow	dep: Customer Account	Processing
Orders Streaming flow	dep: Orders	Processing
Web Feed AWS S3 flow	dep: Web	Success

At the bottom of the dialog, there are two buttons: 'Select' (highlighted with a blue box) and 'Cancel'. The status bar at the bottom of the interface shows 'lastName' and 'person.name.lastName'.

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 sections for Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets), and more. The main area is titled 'Sources > HTTP API > Add data' and is currently at the 'Mapping' step. A progress bar indicates 24 of 24 Mapped fields and 1 of 1 Required fields. Below this, 'Identity fields' show 2 of 2. An 'Errors: 0' message is present. The 'SOURCE DATA' section contains three entries: 'concat( lpad(date\_part("month", date(birth\_Date))...', 'date\_part("year", date(birth\_Date))', and 'iif(sms\_optin == null, 'n', sms\_optin)'. These map to 'TARGET FIELDS': 'person.birthDayAndMonth', 'person.birthYear', and 'consents.marketing.sms.val'. The 'Mapping schema' is set to '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 current step is 'Review'. The interface includes a navigation sidebar on the left with sections like Home, Workflows, Dashboards, CONNECTIONS, Sources, Destinations, CUSTOMER, SEGMENTS, IDENTITIES, PRIVACY, DATA SCIENCE, and DATA MANAGEMENT. The main area displays two panels: 'Connection' (Account name: Customer Updates - YourNameHere; Source platform: HTTP API; Source name: Lab\_Single\_Customer.json) and 'Assign dataset and map fields' (Target dataset: Customer Account - YourNameHere; Schema mapping). A red box highlights the 'Connected' status in the Connection panel, and another red box highlights the 'Dataset assigned' status in the Assign dataset panel. A large red arrow points upwards from the bottom right towards the 'Finish' button in the top right corner.

### 3.4. 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, there is a navigation sidebar with various categories like Home, Workflows, Dashboards, CONNECTIONS (with Sources selected), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring). The main content area shows a source named "Customer Updates - YourNameHere" which is an "HTTP API". It has a summary: Account Name: Customer Updates - YourNameHere, 1 dataset, Last updated August 24, 2022. Below this is a table with columns: DATAFLOW NAME, TARGET DATASET, and LAST DATAFLOW RUN. One row shows "Live Customer Updates - YourNameHere" pointing to "Customer Account - YourNameHere" with "No runs". To the right of the table is a detailed view of the source, including its ID (b653911b-1350-4f2f-8703-13a79807954b), Source name (Customer Updates - YourNameHere), Description (empty), Category (Streaming), and Streaming endpoint (a long URL starting with https://dcs.adobedc.net/collection/518c27...). Red arrows highlight the "Sources" button in the sidebar, the "HTTP API" card, and the "Streaming endpoint" URL.

## 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, there's a navigation sidebar with sections like Home, Workflows, Dashboards, CONNECTIONS (Sources is selected), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, and DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring). The main area displays a dataflow named "Customer Updates - YourNameHere". The dataflow card shows the account name, number of datasets (1 dataset), and last update date (Last updated Aug 24, 2022). Below the card is a table with columns: DATAFLOW NAME, TARGET DATASET, and LAST DATAFLOW RUN. The dataflow name is "Live Customer Updates - YourNameHere", the target dataset is "Customer Account - YourNameHere", and the last run status is "No runs". A "Next" button is visible at the bottom of this table. To the right of the table is a "Properties" panel. At the bottom right of the main area is a "API Usage" section. A red box highlights the "API Usage" section in the properties panel, which contains the following information:

API Usage	①
Dataflow ID	00782e58-ddce-43a0-b6e9-99fe66ac9892
Dataset ID	6305c0b7cb26b31c0753fb20
Mapping set ID	e1ae59ef5d814ef7ba6e6b35a73a7af9
Mapping version	0

### 3.5. Place a REST API call

Switch to the Postman UI, and set the following

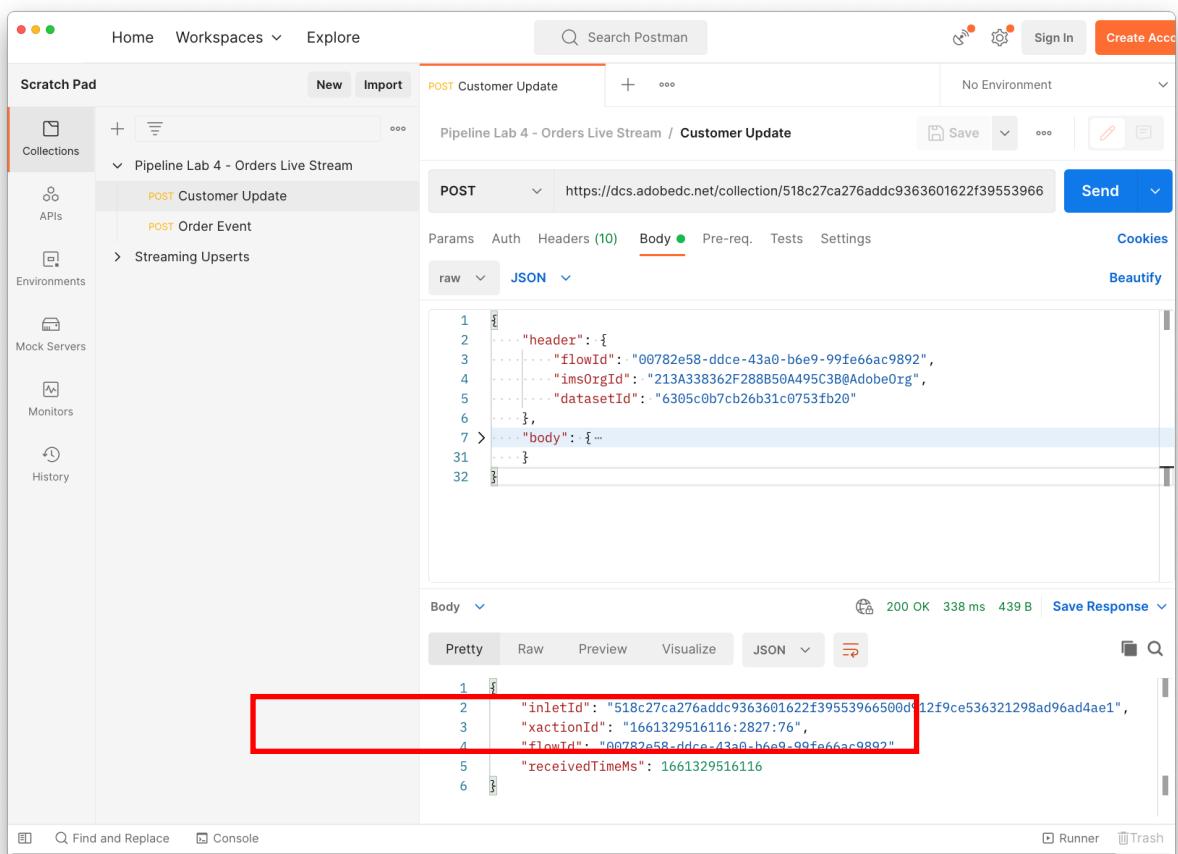
Type: POST

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 interface with a POST request to "Customer Update". The request body is a JSON object with several fields, some of which are highlighted with a red box. The response status is 200 OK, and a large red arrow points to the response body area.

```
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    "birthDate": "1940-01-15",  
15  }  
16}
```

Body 200 OK 338 ms 439 B

Pretty Raw Preview Visualize JSON ↻

1 {  
2 "inletId": "518c27ca276addc9363601622f39553966500d912f9ce536321298ad96ad4ae1",  
3 "xactionId": "1661329516116:2827:76",  
4 "flowId": "00782e58-ddce-43a0-b6e9-99fe66ac9892",  
5 "receivedTimeMs": 1661329516116  
6 }

### 3.6. 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, the navigation bar has 'Profiles' selected under 'Customer'. The main area has 'Browse' selected in the top navigation. A search bar at the top allows filtering by merge policy (set to 'Default Timebased') and identity namespace (set to 'Email'). Below the search is a large red arrow pointing to the message 'No profiles found.' which is displayed in front of a table header row containing columns for PROFILE ID, FIRST NAME, LAST NAME, and PERSONAL EMAIL. To the right of the table, a summary box displays '426 Profiles'.

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 'Identity namespace' input field in the search bar. It contains the value 'Email' with a green checkmark and a delete button. The 'Merge policy' and 'Identity value' fields are also visible above it.

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	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, the navigation sidebar is visible with sections like Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring). The main content area shows a profile named 'Profile BVrqzwVv7sqLqxJmopfmnaG3v3KJgA'. The 'Detail' tab is selected. In the 'Basic attributes' section, the 'Birth Day And Month' field is populated with '01-15'. Other fields include Address (yourEmail@adobe.com), Gender (-), Number (-), Street 1 (-), Customerid (202208240125), Email (yourEmail@adobe.com), and Channel preferences (Direct mail, Phone, SMS, Email, Facebook news f..., Twitter feed). A 'View identity graph' link is also present.

## 3.7.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.

## 3.8. 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 <small>i</small> *	Identity namespace <small>i</small>	Identity value <small>i</small>
Default Timebased	Email	yourEmail@adobe.com

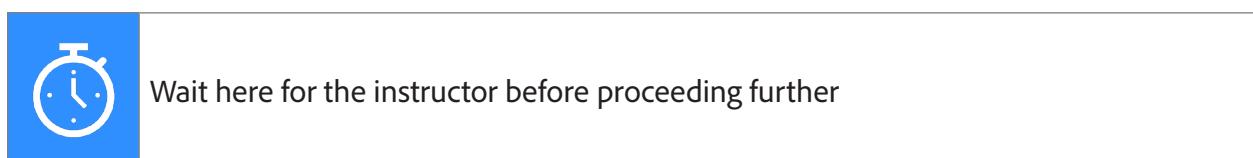
**View**    Clear

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"

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
firstName	YourFirstName	person.name.firstName
id	yourEmail@adobe.com	identityMap.email.0.id
id	202208240125	identityMap.customerid.0.id



## 4. Lab Tasks B – Orders – Live data (Streaming)

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

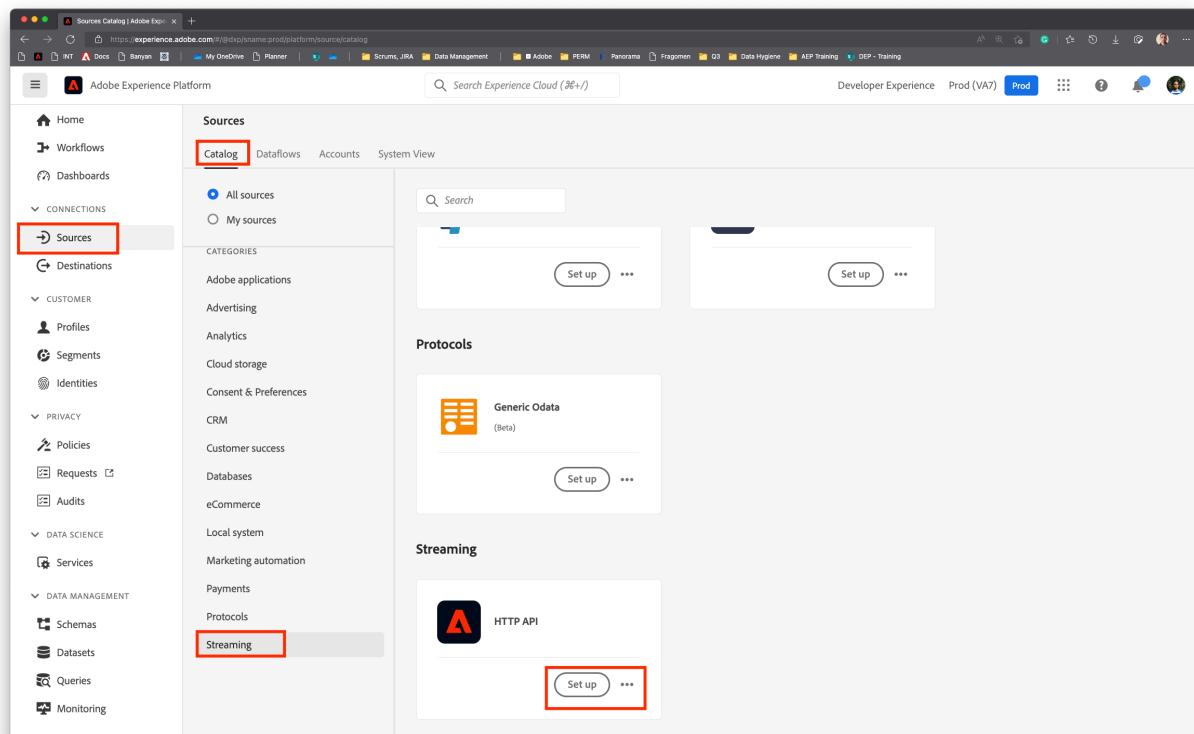
### Pre-requisites

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

### Steps

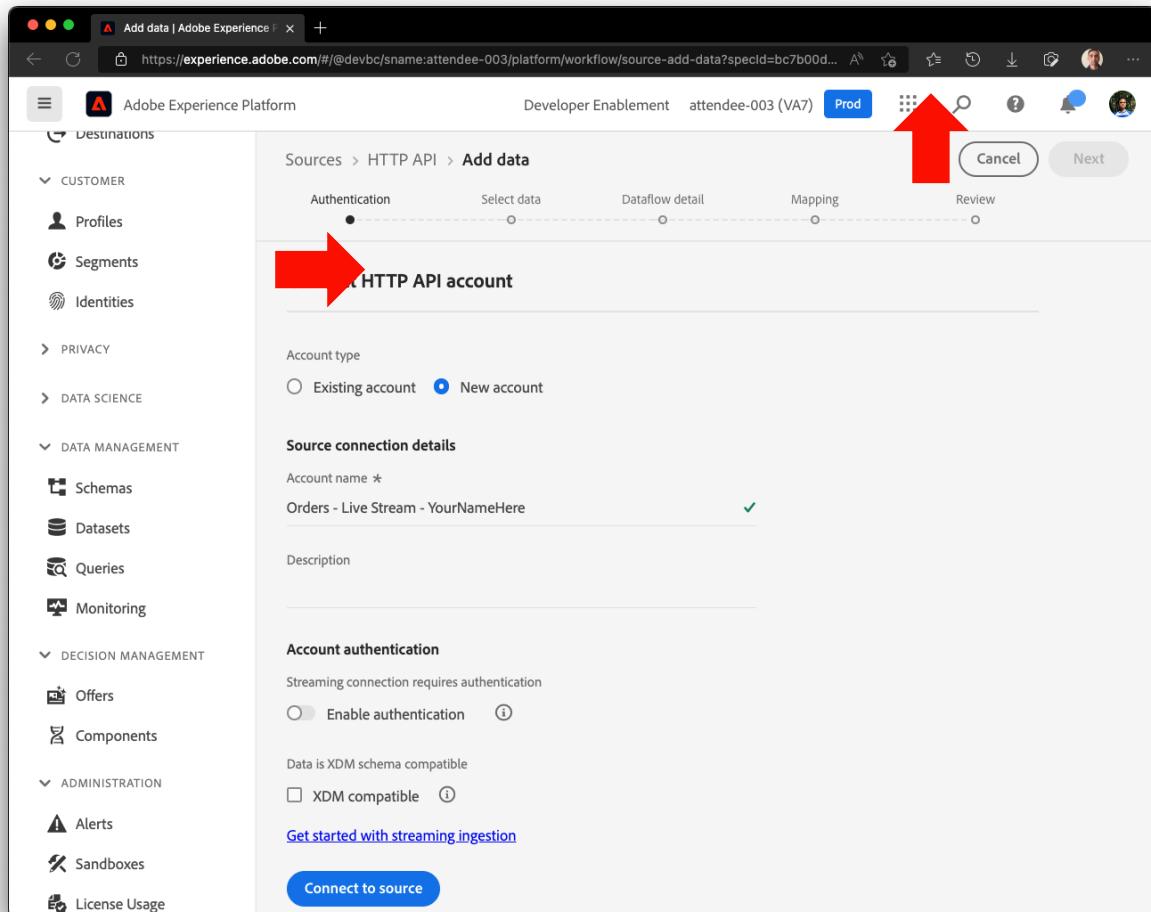
Go to Adobe Experience Platform → Sources → Catalog → Streaming. Click on **Setup / Add Data** for the HTTP API.

**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.



## 4.1.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 interface for adding data via the HTTP API. 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), and Help.

The main panel displays the 'Sources > HTTP API > Add data' workflow. The current step is 'Authentication'. The process consists of five steps: Authentication, Select data, Dataflow detail, Mapping, and Review.

**Connect HTTP API account**

**Account type:** Existing account (radio button) is selected, while New account is also available.

**Source connection details:** Account name is set to "Orders - Live Stream - YourNameHere".

**Description:** A text input field for describing the connection.

**Account authentication:** Streaming connection requires authentication. The "Enable authentication" checkbox is checked.

**Data is XDM schema compatible:** The "XDM compatible" checkbox is unchecked.

**Get started with streaming ingestion:** A link with a green checkmark icon indicates the connection is connected.

**Sign in with different credentials:** A link for signing in with alternative credentials.

At the top right, there are 'Cancel' and 'Next' buttons. The 'Next' button is highlighted in blue.

## 4.2. 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 (Profiles, Segments, Identities), PRIVACY (Policies, Requests, Audits), DATA SCIENCE (Services), and DATA MANAGEMENT (Schemas, Datasets, Queries). The main panel is titled 'Sources > HTTP API > Add data' and is currently on the 'Select data' step. It displays the 'Source data schema' which has been defined by uploading a JSON file named 'Lab\_Single\_Order.json'. The schema preview shows a hierarchical structure of fields: 'SampleSchema:a5158f9307db472bbc53b163c9dffd55' (root object) containing 'Products' (an array of objects) and 'plan' (an object). Below 'plan' are fields for 'billingCity' (String, Value: Salt Lake City), 'billingState' (String, Value: Utah), 'billingStreetAddress' (String, Value: 756 Northland Court), 'billingZip' (String, Value: 84125), 'customerID' (String, Value: 748451685), and 'lastOrderStatusUpdate' (String, Value: 2022-06-08T22:06:56Z).

## 4.3. 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.

The screenshot shows the 'Add data' interface in the Adobe Experience Platform. 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), and ADMINISTRATION (Alerts, Sandboxes, License Usage). The main panel is titled 'Sources > HTTP API > Add data' and is on the 'Dataflow detail' step of a five-step process. The 'Dataset details' section includes fields for 'Target dataset' (set to 'Existing dataset' with value 'Orders - YourNameHere') and 'Profile dataset' (checkbox checked). The 'Dataflow details' section includes fields for 'Dataflow name' (set to 'Orders - Live - YourNameHere') and 'Description'. The 'Alerts' section has a checkbox for 'Sources Dataflow Lack of Ingestion' which is unchecked. At the top right, there are 'Cancel', 'Back', and 'Next' buttons, with 'Next' being highlighted in blue.

## 4.4. Data Prep / Transformation

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

The screenshot shows the 'Add data' workflow in Adobe Experience Platform. The left sidebar includes sections for Home, Workflows (selected), Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY (Policies, Requests, Audits), DATA SCIENCE (Services), and DECISION MANAGEMENT. The main area is titled 'Sources > HTTP API > Add data' and is at the 'Mapping' step. It displays progress bars for 'Mapped fields' (29 of 29), 'Required fields' (1 of 3), and 'Identity fields' (2 of 2). A red box highlights the 'Import mapping' button. Below it, a warning message states: 'There was error(s) preparing mappings. 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. 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. The following unmapped paths are required: \_id, timestamp. The following unmapped paths are required: productListItems.SKU.' The bottom section shows a mapping table with source data from 'Orders\_Streaming\_CLIENT.json' and target fields in the 'dep: Orders' schema.

SOURCE DATA	TARGET FIELDS
shippingCity	shipping.address.city
Products[*].quantity	productListItems[*].quantity
Products[*].productId	productListItems[*]._id
Products[*].price	productListItems[*].priceTotal
Products[*].model	productListItems[*].product

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

The screenshot shows the Adobe Experience Platform interface for adding data via the HTTP API. A modal window titled "Import mapping" is open, listing various dataflows and their target datasets. The "Orders - Backfill - YourNameHere" dataflow is selected, indicated by a blue highlight. Other listed dataflows include "Dataflow - Customer Account - YourNameHere", "Live Customer Updates - YourNameHere", "Web Streaming flow", "Customer Streaming flow", "Orders Streaming flow", "Web Feed AWS S3 flow", and "Aggregates Feed AWS S3 flow". The target datasets for these flows are "Orders - YourNameHere", "Customer Account - YourNameHere", "Customer Account - YourNameHere", "dep: Web", "dep: Customer Account", "dep: Orders", "dep: Web", and "dep: Customer Aggregates" respectively.

DATAFLOW NAME	TARGET DATASET
Orders - Backfill - YourNameHere	Orders - YourNameHere
Dataflow - Customer Account - YourNameHere	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 'Errors: 0' message is shown. Below this, there's a search bar ('Search source fields') and a dropdown ('All fields'). A red box highlights this search/filter area. To the right is a 'Import mapping' button and other UI elements. The main content area shows a table mapping source data to target fields:

SOURCE DATA	MAPPING	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

A red box highlights the entire table area. At the bottom right of the main area is a 'Next' button.

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" (with value 962105342), "\_DXP.PERSONALEMAIL" (with value mchicchetto@msn.com), and "PRODUCTLISTITEM" (with value Google). The mapping tree starts with "dep: Orders" and branches into "dpx", "billing", "order", "productListItems", "shipping", and "store". Each of these objects has its own sub-mappings, such as "dpx" having "id" and "eventMergeId", and "productListItems" having "eventType" and "identityMap". 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 is a search bar ('Search source fields') and a dropdown menu ('All fields'). A 'Import mapping' button is also present. The main content area displays a mapping schema between 'Source: Lab\_Single\_Order.json' and 'Mapping schema: dep: Orders'. The 'SOURCE DATA' column lists fields like 'Products[\*].productId', 'Products[\*]', 'lastOrderStatusUpdate', 'inStore', 'concat(orderId, ":", lastOrderStatusUpdate)', 'Products[\*].model', 'Products[\*].make', 'orderStatus', and 'shippingCity'. The 'TARGET FIELDS' column lists corresponding fields like 'productListItems[\*].SKU', 'productListItems[\*]', 'timestamp', 'order.\_devbc.acqSource', '\_id', 'productListItems[\*]\_devbc.model', 'productListItems[\*]\_devbc.make', 'eventType', and 'shipping.address.city'. Each mapping entry has a delete icon on the right.

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.

## 4.5. 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 Sources interface. On the left, there's a sidebar with various categories like Dashboards, Connections, Sources (which is currently selected), Destinations, Customer, Profiles, Segments, Identities, Privacy, Data Science, Data Management (with Schemas, Datasets, Queries, Monitoring), and Decision Management (with Offers). The main content area is titled 'Sources > Orders - Live Stream - YourNameHere'. It displays a summary card for an 'HTTP API' source. The card includes the account name 'Orders - Live Stream - YourNameHere', the number of datasets (1), and the last update date (August 24, 2022). Below this, a table lists the dataflow details: DATAFLOW NAME (Orders - Live - YourNameHere), TARGET DATASET (Orders - YourNameHere), LAST DATAFLOW RUN STATUS (No runs), and LAST DATAFLOW RUN (empty). On the right, a panel titled 'Orders - Live Stream - YourNameHere' provides more details: ID (de1e6f49-d9ea-4029-b7a5-009c7c1847cd), Source name (Orders - Live Stream - YourNameHere), Description (-), Source (HTTP API), Category (Streaming), and API Usage (Streaming endpoint URL: https://dcs.adobedc.net/collection/65faa7...). A 'Next' button is visible at the bottom of the table.

### 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 Developers Bootcamp interface. The left sidebar contains navigation links for 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 is titled "Sources > Orders - Live Stream - YourNameHere". It displays an "HTTP API" entry with the account name "Orders - Live Stream - YourNameHere", 1 dataset, and last updated on August 24, 2022. Below this is a table with columns: DATAFLOW NAME, TARGET DATASET, LAST DATAFLOW RUN STATUS, and LAST DATAFLOW RUN. One row is shown: "Orders - Live - YourNameHere" (with three dots) maps to "Orders - YourNameHere" with a status of "No runs" and a run time of "-". A "Properties" panel on the right shows the status as "Processing", partial ingestion as "Not enabled", error diagnostics as "Not enabled", and creation date as "08/24/2022, 5:54 PM". The API Usage section shows a dataflow ID of "fe36eb0b-2550-4dc8-8a40-8cbd4b6d13d4", dataset ID "6306c0cac1897a1c08f7e7d4", mapping set ID "1b3643e3b4b046b4b6285982178c1864", and mapping version "0".

## 4.6. 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, the sidebar includes 'Scratch Pad' (selected), 'Collections', 'APIs', 'Environments', 'Mock Servers', 'Monitors', and 'History'. The main workspace displays a 'Pipeline Lab 4 - Orders Live Stream / Order Event' collection. A 'POST Order Event' request is selected. The 'Body' tab is active, showing the following JSON payload:

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

Below the code editor is a 'Response' section featuring a cartoon rocket and the text '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": "Glinojeck",  
    "billingZip": "06-450",  
    "plan": {  
        "ID": "m3"  
    },  
    "Products": [  
        {  
            "productID": "PRODUCT-6",  
            "quantity": 1,  
            "price": 150.23,  
            "currencyCode": "USD",  
            "model": "v2",  
            "make": "Google"  
        },  
    ],  
    "orderLineItems": [  
        {  
            "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,
    "billing": {
      "city": "Glinojeck",
      "streetAddress": "497 Kensington Drive",
      "postalCode": "06-450"
    },
    "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
}
```

The screenshot shows a Postman interface with a JSON response. The response body is a single object with four properties: `inletId`, `xactionId`, `flowId`, and `receivedTimeMs`. The `inletId` value is a long string of characters. The `xactionId` value is `1661389779753:1943:76`. The `flowId` value is a UUID. The `receivedTimeMs` value is a timestamp. The Postman interface includes tabs for 'Pretty' (selected), 'Raw', 'Preview', 'Visualize', and 'JSON'. It also shows status information: 200 OK, 350 ms, 439 B, and a 'Save Response' button.

```

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

```

## 4.7. 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 Adobe Experience Platform interface. The left sidebar has a tree structure with sections like Home, Workflows, Dashboards, CONNECTIONS, Sources, Destinations, CUSTOMER (expanded), SEGMENTS, IDENTITIES, PRIVACY (Policies, Requests, Audits), DATA SCIENCE (Services), and DATA MANAGEMENT (Schemas). Under CUSTOMER, 'Profiles' is selected and highlighted with a red box. In the main content area, the 'Profiles' section is active, indicated by a red box around the 'Browse' tab in the top navigation bar. Below it, there's a search/filter section for merge policy and identity namespace, and a table listing profiles with columns: PROFILE ID, FIRST NAME, LAST NAME, and PERSONAL EMAIL. The table contains several rows of profile IDs. To the right, there are two summary boxes: one showing '10 Profiles' and another for 'Profiles by namespace' with a customerID count of 10. The top right corner shows 'Developer Experience lab-001 (VA7) Dev' and some user icons.

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**.

Merge policy (i) \*      Identity namespace (i)      Identity value (i)  
 Default Timebased      Email      yourEmail@adobe.com  
 View      Clear

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.

The screenshot shows the Adobe Experience Platform Profile Detail page for a specific profile. The left sidebar navigation includes Home, Workflows, Dashboards, CONNECTIONS (Sources, Destinations), CUSTOMER (Profiles, Segments, Identities), PRIVACY, DATA SCIENCE, DATA MANAGEMENT (Schemas, Datasets, Queries, Monitoring). The main content area displays the following details:

- Customer profile:** Placeholder image, Profile ID: BVrqzwVv7sqLqxJmopfmnaG3v3KJgA.
- Basic attributes:**
  - Address: yourEmail@adobe.com
  - Gender: -
  - Birth Day And Month: 01-15
  - Number: -
  - Street1: -
- Linked identities:** CustomerId: 202208240125, Email: yourEmail@adobe.com
- Channel preferences:**
  - Direct mail
  - Phone
  - SMS
  - Email
  - Facebook news feed
  - Twitter feed