

Creating Formula Fields on Data Lake Objects (DLOs)

Data Lake Object (DLO)

A **Data Lake Object (DLO)** is the first layer in Salesforce Data Cloud where raw data from various sources (CSV uploads, APIs, connectors, etc.) is ingested and stored.

DLOs hold **unprocessed or semi-structured data** and act as the foundation for transformations into **Data Model Objects (DMOs)**.

Example:

A DLO named Purchase_Transactions_DLO might contain the following fields:

transaction_id	customer_id	purchase_date	amount	region
T1001	C001	2025-10-10	850	East
T1002	C002	2025-10-12	450	West
T1003	C001	2025-10-15	25	East

2. Formula Fields on DLOs

A **formula field** in a DLO is a **derived field** that performs a calculation or transformation on existing fields.

It is not physically stored in the ingested dataset; rather, it is computed dynamically using an expression or formula.

You can use formula fields to:

- Combine or manipulate existing field values
 - Apply conditional logic (e.g., IF statements)
 - Format text, numbers, or dates
 - Compute derived metrics (e.g., discount, tax, profit margin)
 - Standardize data formats for downstream DMOs
-

3. Why Use Formula Fields on DLOs?

Formula fields are highly valuable for **data preparation** and **validation** because they:

1. Eliminate the need to transform data externally before ingestion.
 2. Allow quick creation of computed attributes without new data sources.
 3. Ensure consistent, reusable business logic across datasets.
 4. Help simplify DMO transformations by pre-deriving key metrics.
 5. Improve accuracy and speed in data analysis.
-

4. Where You Can Create Formula Fields

Formula fields can be created in Salesforce Data Cloud:

- On **Data Lake Objects (DLOs)** – for raw data preparation.
- On **Data Model Objects (DMOs)** – for unified and cleaned data.
- On **Calculated Insight Objects (CIOs)** – for aggregated and KPI-level metrics.

However, creating them **at the DLO layer** is ideal for **basic derived attributes** before unification or identity resolution.

5. Steps to Create a Formula Field on a DLO

Step 1: Navigate to Data Cloud Setup

1. Go to **Salesforce Data Cloud** from the App Launcher.
2. Open the **Data Manager** workspace.

Step 2: Locate Your Data Lake Object

1. In the **Data Streams** or **Data Explorer**, locate the DLO (e.g., `Purchase_Transactions_DLO`).
2. Click **Schema** or **Fields** to view its metadata.

Step 3: Add a New Formula Field

1. Click **Create Field** or **Add Field**.
2. Select **Formula Field** as the field type.
3. Provide:
 - **Field Label** – e.g., “Discounted Amount”
 - **Field API Name** – e.g., `discounted_amount__c`
 - **Data Type** – Number, Text, Date, or Boolean
 - **Formula Definition** – use the expression builder

Step 4: Define the Formula Logic

Use Salesforce formula syntax, similar to what you use in CRM formulas.

For example:

```
amount * 0.9
```

(If applying a flat 10% discount.)

Step 5: Save and Validate

1. Save the formula.
 2. Return to **Data Explorer** to preview data with your new computed field.
 3. Validate that formula output appears as expected.
-

6. Example Formula Field Scenarios

Example 1: Discount Calculation

Objective: Calculate a 10% discount on the purchase amount.

Formula:

```
amount * 0.9
```

Result:

amount	discounted_amount
850	765
450	405
25	22.5

Example 2: Conditional Logic

Objective: Flag high-value transactions over \$500.

Formula:

```
IF(amount > 500, "High Value", "Regular")
```

Result:

amount	transaction_type
850	High Value
450	Regular
25	Regular

Example 3: Date Transformation

Objective: Extract the month name from purchase_date.

Formula:

TEXT(MONTH(purchase_date))

Result:

purchase_date	purchase_month
2025-10-10	10
2025-10-12	10

Example 4: Combine Fields

Objective: Create a full transaction label for reporting.

Formula:

customer_id & "-" & transaction_id

Result:

customer_id	transaction_id	transaction_label
Co01	T1001	Co01-T1001
Co02	T1002	Co02-T1002

Example 5: Data Cleansing

Objective: Replace null values in payment mode with “Unknown”.

Formula:

```
IF(ISBLANK(payment_mode), "Unknown", payment_mode)
```

Result:

payment_mode	cleaned_payment_mode
Credit Card	Credit Card
(blank)	Unknown

7. Real-World Use Case

Company Background

A financial services firm is ingesting customer transaction data into Salesforce Data Cloud.

They need to categorize transactions by region, flag high-value deals, and calculate commission percentages before unification.

Implementation Steps

Step	Description	Tool
1	Ingest transaction CSV into Transaction_DLO.	Data Stream
2	Create formula fields: commission_amount, region_category, and high_value_flag.	DLO Schema Editor
3	Validate the computed data in Data Explorer .	Data Explorer
4	Map these new fields into the Sales_Transaction_DMO for unified data.	Data Model Builder

Outcome

- Reduced external data preparation time by 30%.
 - Simplified transformation logic in DMOs.
 - Improved consistency across analytics dashboards.
-

8. Best Practices for Formula Fields in DLOs

Best Practice	Description
Keep formulas simple	Complex formulas slow down processing during ingestion or querying.
Validate field types	Ensure numeric, text, and date types are correctly defined.
Use meaningful names	Clearly name derived fields, e.g., net_amount, transaction_flag.
Avoid redundant fields	Don't recreate the same formula at multiple layers (DLO, DMO, CIO).
Reuse logic downstream	If a formula is valuable for analytics, propagate it into DMOs or CIOs.
Test before activation	Use Data Explorer to preview and validate output accuracy.

9. Summary

Concept	Description
DLO (Data Lake Object)	Raw data source layer in Salesforce Data Cloud.
Formula Field	A computed, dynamic field derived from existing fields using an expression.
Purpose	Data enrichment, cleansing, transformation, and standardization.
Created In	DLO Schema Editor or Data Explorer interface.
Common Uses	Calculations, categorization, conditional logic, concatenation, and data formatting.

10. Key Takeaway

Formula fields on DLOs let you transform, enrich, and validate raw data *inside Salesforce Data Cloud* — without requiring external preprocessing or scripts. They make data cleaner, more meaningful, and immediately usable for building DMOs, CIOs, and downstream analytics.