

# Einstein Analytics:

## How to build a DataSet with DataFlow

Last update: 05/03/2020

Trailheads:

- Einstein Analytics Data Integration Basics > [Extract Salesforce Data into Analytics](#)
- Superbadge: [Einstein Analytics Data Preparation Specialist](#)

Example:

Create a dataset named "TEST\_2020"

1. Options on creating:

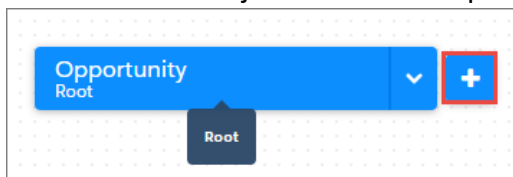
- Analytics Studio > click **Create** and select **Dataset** > Click **Salesforce Data** > give name > From the **Select dataflow** picklist, select **Default Salesforce Dataflow** > see point2 from the part of "Pick a Salesforce object to start"
- Analytics Studio > Data Manager > Dataflow & Recipes > **Default Salesforce Dataflow** > click



> give name > see part2

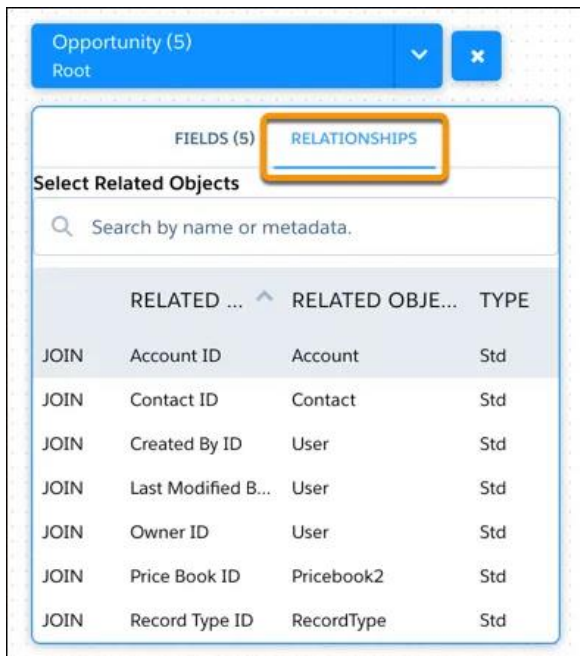
2. Following 1\_1), in the "Pick a Salesforce object to start", it will be the same for 1\_2). We need to choose a Root-object. See trailhead 1 for explanation of Root-Object. As an example here:

- In the object list, click **Opportunity** >
- Hover over the object and click the plus (+) >

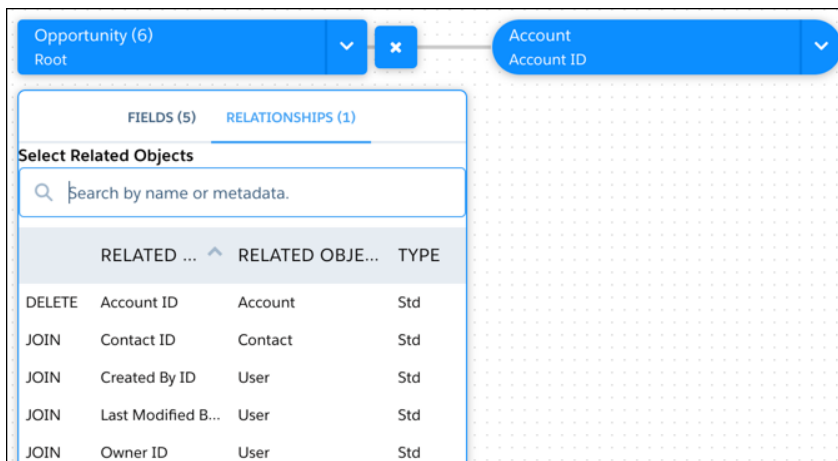


You see a list of the object's fields.

- Click to select each field that you need. If you can't see a field, start entering its name in the **Search by name or metadata** search box to filter the list. Select these fields:
  - i. Amount
  - ii. Close Date
  - iii. Created Date
  - iv. Name
  - v. Stage
- To open a list of objects related to your root object, at the top of the field list, click the **Relationships** tab. This is where you can select the fields you need from the related accounts and users.

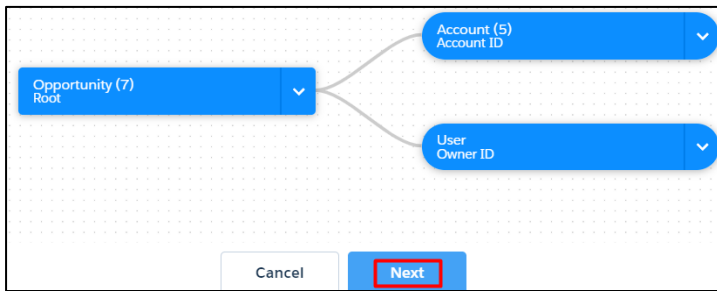


- In front of the **Account ID** relationship, click **Join**. You see the Account object on the dataset builder canvas.

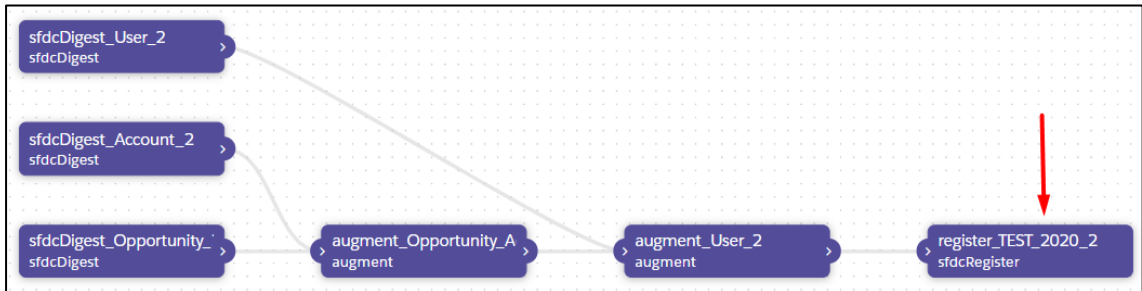


- In front of the **Owner ID** relationship, click **Join**. You see the User object on the canvas.
- Click the X next to Opportunity to close the selection window
- **Account** object - click the plus (+), select:
  - Account Name
  - Billing City
  - Billing Country
  - Industry
  - SIC Code
- **User** object - click the plus (+), select:
  - Full Name
  - Title

- Click Next



- Now back to the canvas and a new set of nodes are build there, targeting to the dataset “TEST\_2020”. So if I run the dataflow now, the selected columns will be written into this dataset



- After running the dataflow, you can find the dataset from DataManager, Apps, or in all item if it doesn't belong to any apps