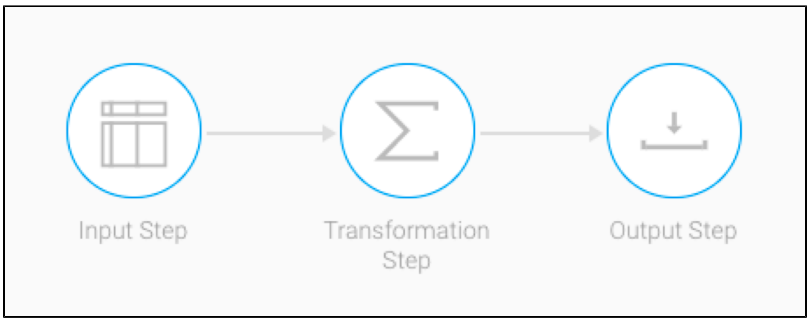
Creating a Basic Transformation Flow

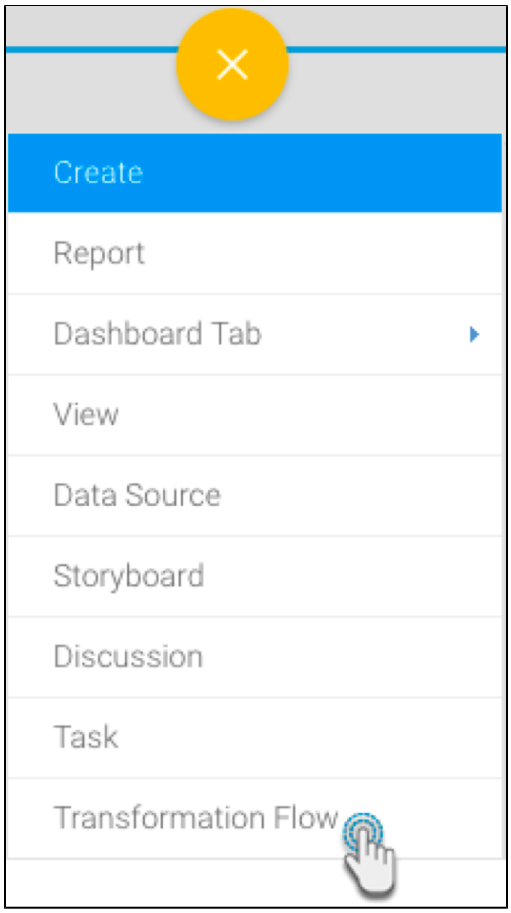
Overview

In this example, we will cover how to create a simple transformation flow, through the transformation flow builder. (Click here to learn about this builder and its components). This involves setting up an input step, adding a transformation step, and connecting the two together. Then add an output step and join the transformation step to it. Each step requires configuration when added to the flow.

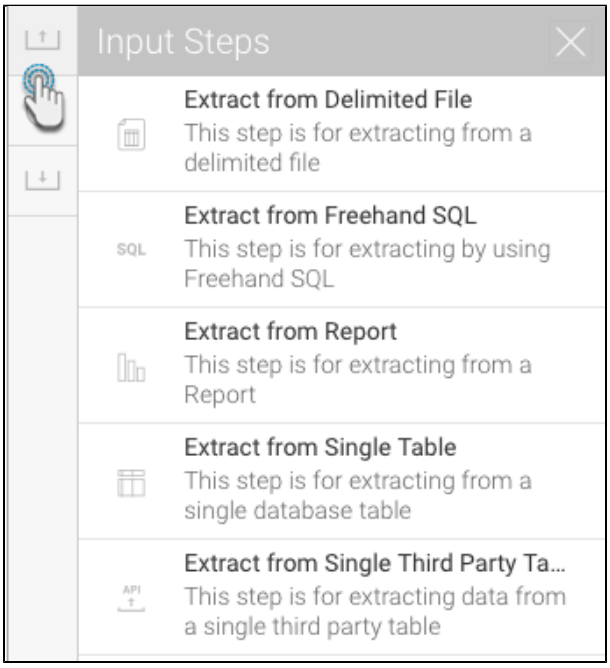


1. Click on the Create button in the top-right corner.
2. Then select the Transformation Flow option.

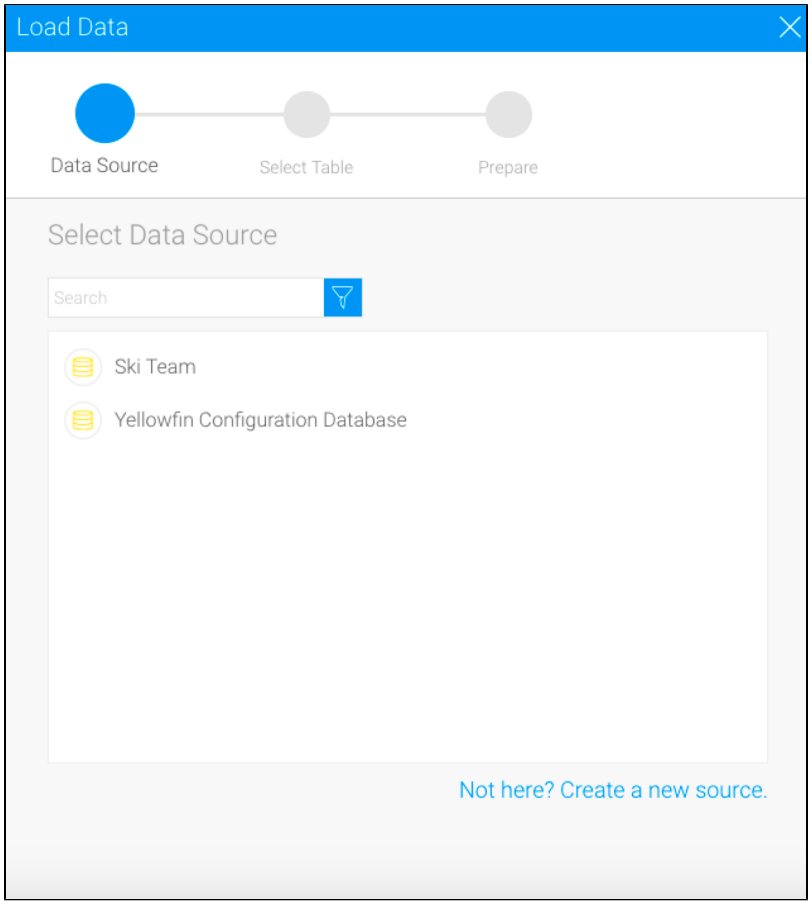
> ⚠ If you do not see this option, you may not have security access to transformation flows. Learn how to get access here.



3. You will be taken to the transformation flow builder.
4. Hover your cursor over the Input Steps button on the left side to make the panel appear.

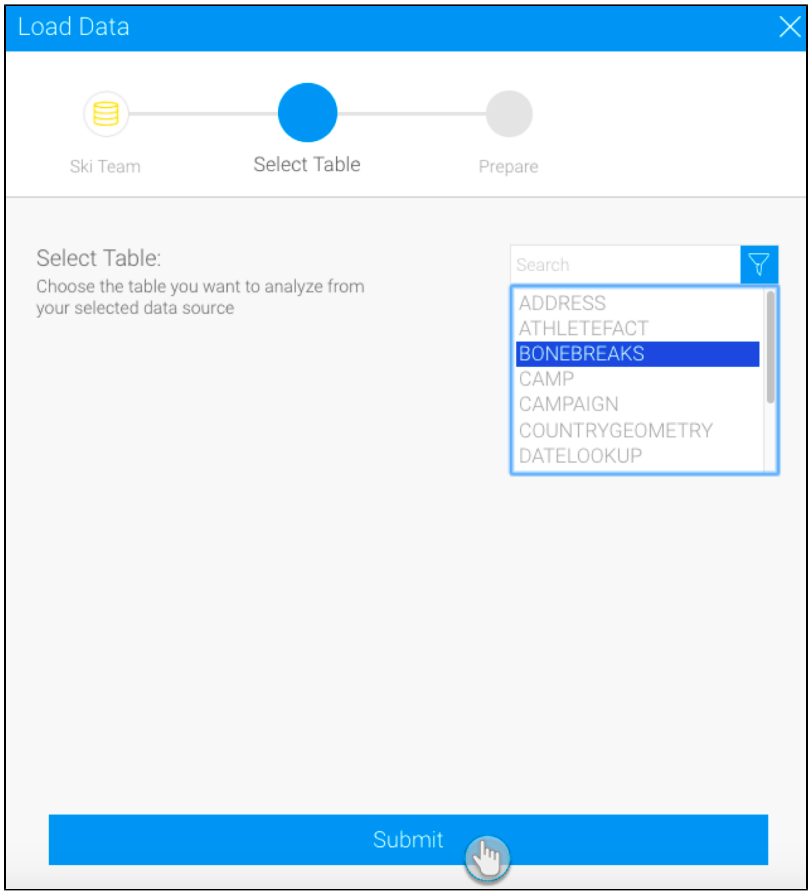


5. Drag an input step from this panel onto the canvas. (For this procedure, we will use the single database table as an example. Click here to learn about the different input steps.)
6. On doing so, a popup will appear, requiring information on the selected data source. (The steps within this popup will differ for each type of input step.)

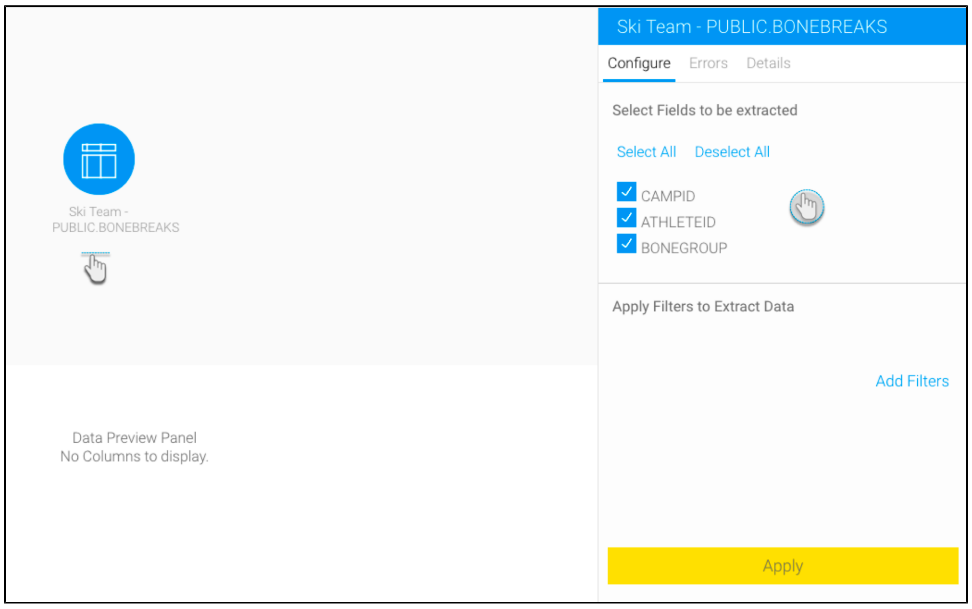


7. Select the data source, and then a single table (since our selected input was a single table).

8. Then click on the Submit button.

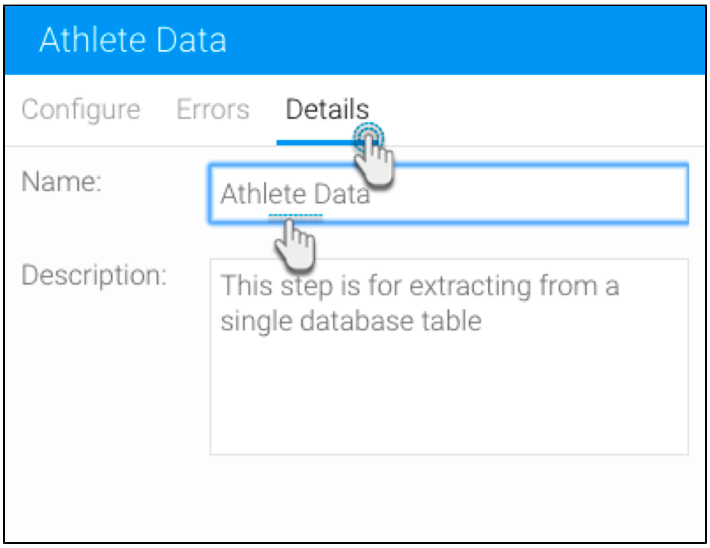


9. Next, click on the newly created input step in the canvas. Its details will appear in the configure flow panel.

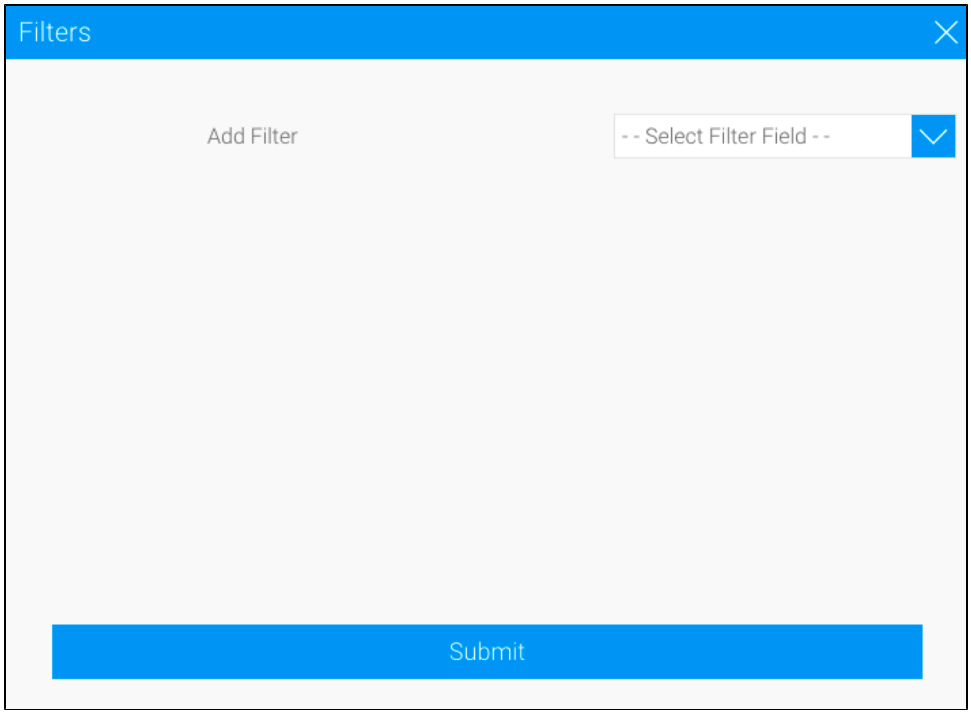


10. **Note:** Clicking on a blank space in the canvas will show details of the entire flow, but clicking on a specific step shows only that step's configurations.

11. If required, you can rename the input step. Click on the Details tab in the Configuration panel, and provide a new name, or even an updated description.

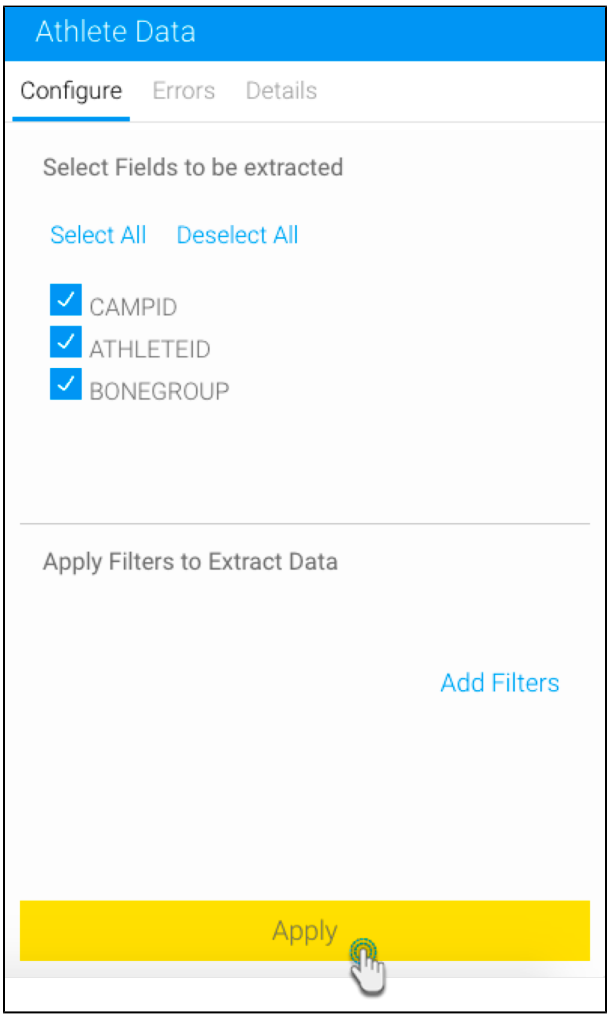


12. Click on the Configure tab and select only the fields that you want data to be extracted from.
13. Click on the Add Filter button to apply filters to control the data being extracted, when the Filter popup appears.

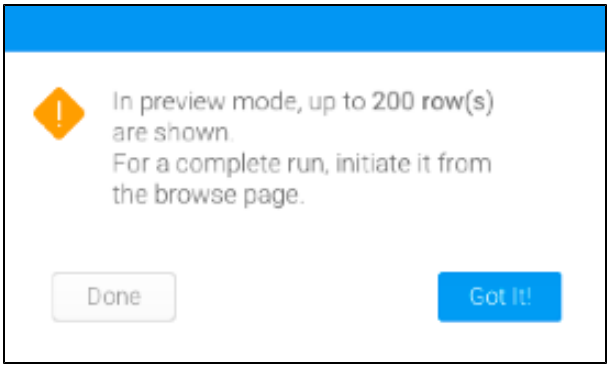


14. Once you're done with the configuration, click on the Apply button.

> You must click on the Apply button to save the configuration details. If you do not click on this button, and click somewhere else first, you will lose the configuration details.



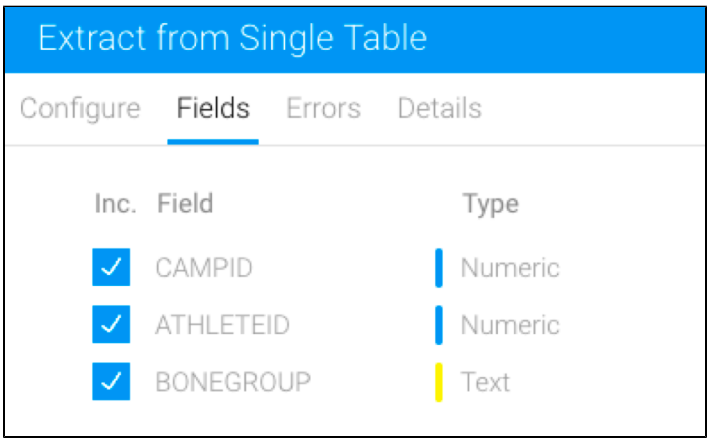
15. **Note:** Clicking on the Apply button for the first time will bring up the following pop up. Click *Done* to close this popup. (It will not show up again for the current user's session. Alternatively, if you click on *Got It!*, the pop up will never be brought up again.)



16. On doing so, the data preview panel will display the data extracted from the configured data source. A maximum row setting of 200 applies.

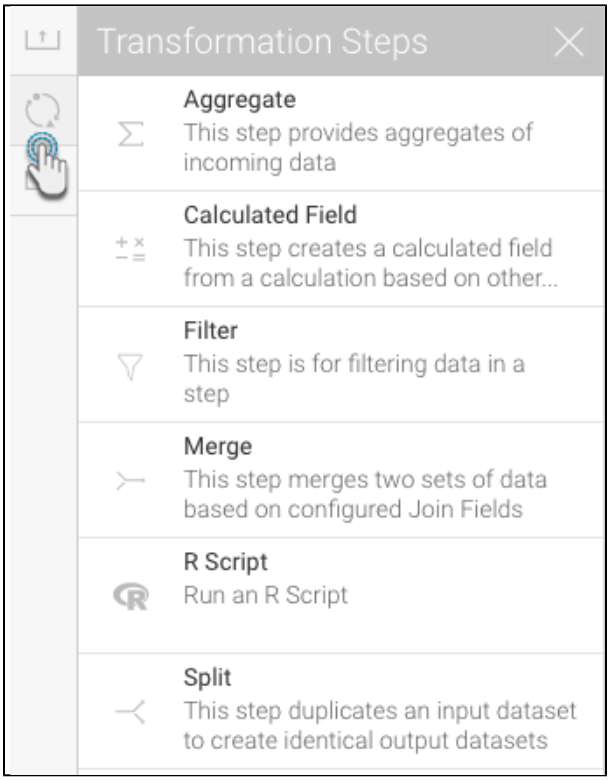
	CAMPID NUMERIC	ATHLETEID NUMERIC	BONEGRO... TEXT
1	12242	9154	Scapula
2	12301	9159	Carpus
3	12242	9199	Clavicle
4	12423	9266	Radius

17. You can apply transformations to fields directly from the data preview panel. Click here to learn more.
18. A new Fields tab will also appear in the step configuration panel. Use this to manage the data fields that are extracted from the step.

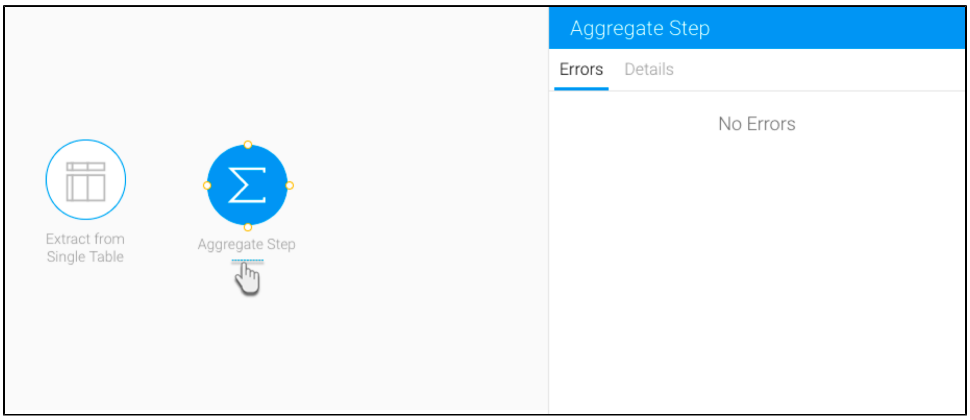


go back

1. Next, expand the transformation step panel by hovering on its icon in the step builder, and then dragging a transformation step onto the canvas.

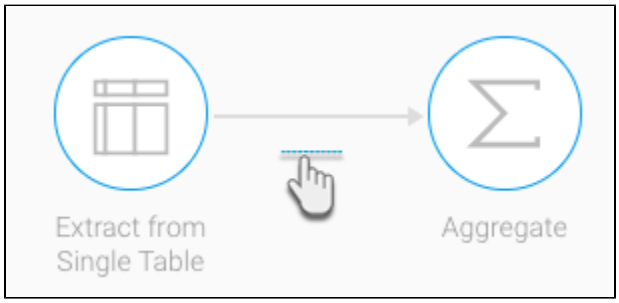


2. A transformation step will appear on the canvas.

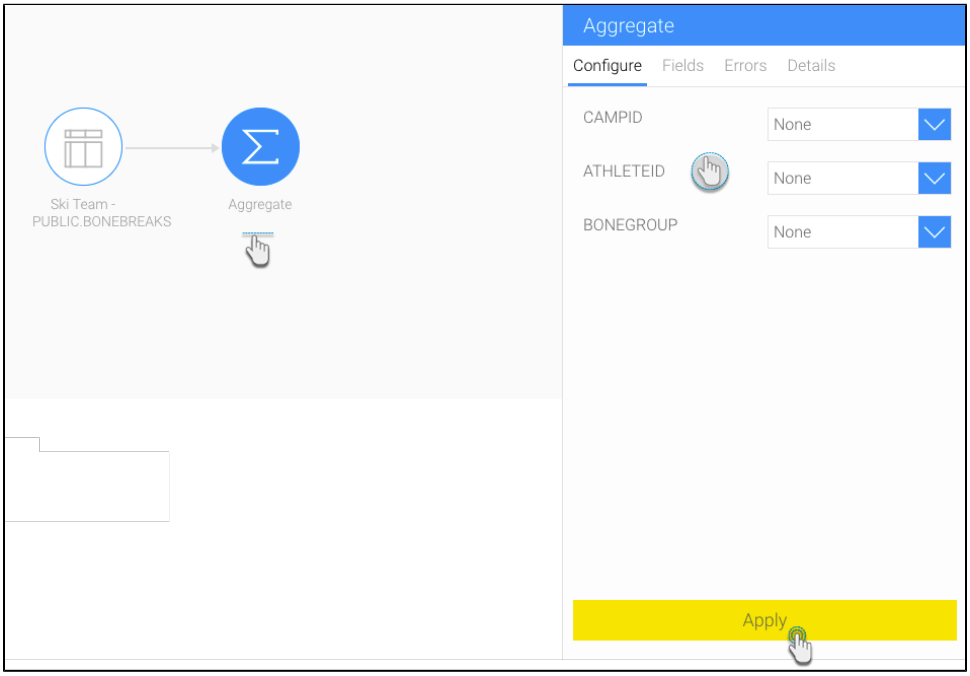


3. Next, create a connection between the input step and the transformation step. This is done by hovering over an input step to make connecting points appear.

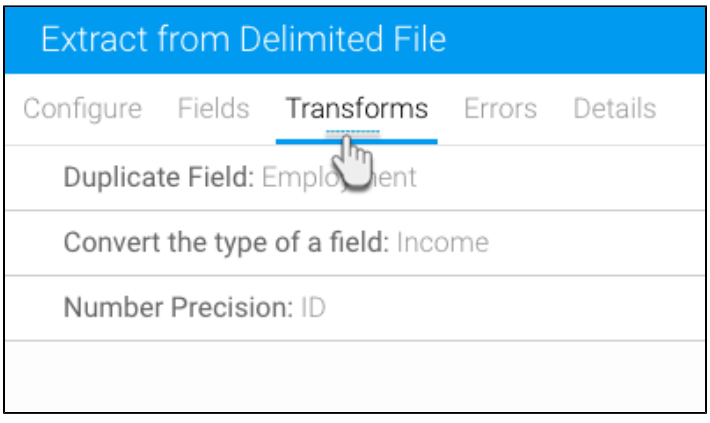
4. Drag a point and connect it to the transformation step to form a relationship between them. (Click here to learn about rules involving connections, and how to break existing connections.)



5. Now you can configure the transformation step, by clicking on its icon and using the step configuration panel. (Each transformation step type will be configured differently.)
6. Save changes using the Apply button.



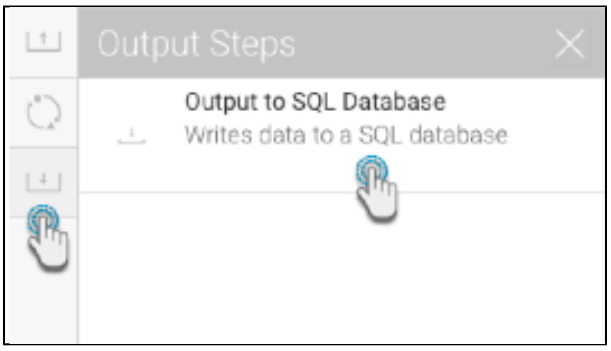
7. **Note:** When a field transformation is made to the data, a Transforms tab will appear in the step configuration panel.



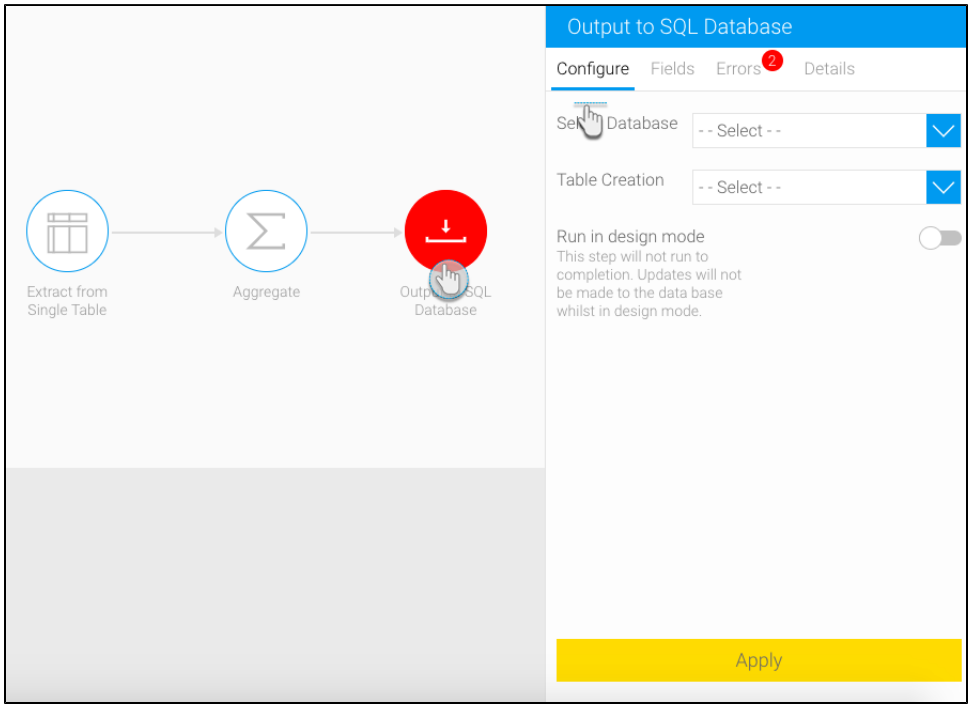
go back

Warning: Extra caution should be taken when configuring the output step so that any existing source data does not get accidentally overwritten. The output to database step has a variety of options such as dropping tables, truncating (deleting) data, adding new rows, and updating existing data. It is recommended that a test database is used when developing transformation flows. Further, the output step is disabled by default when in edit mode (see Run in Design mode option).

1. Extend the output step panel by hovering on its icon, and drag an output step onto the canvas.



2. Next, connect the transformation step to the output step as mentioned above (in part 2, steps 3 and 4).
3. And then configure the output step through the panel on the right-side. Click here to learn more about configuring this step.
   Note: By default, the output step will be highlighted as red to signify that it contains errors. This is because it has not been configured yet.



4. You can also view the flow's status in the run logs tab in the flow panel.
5. Click on the run button in the top header menu to execute the flow. Each step will process the number of rows that has been configured for the design panel (with the max. limit being 200), or until it runs out of data, or if it processes an error. By default, any output to database step will not run. This behaviour can be overridden by enabling the Run in Design mode option in the output step configuration panel. This should only be done subject to the warnings listed earlier on this page.

go back

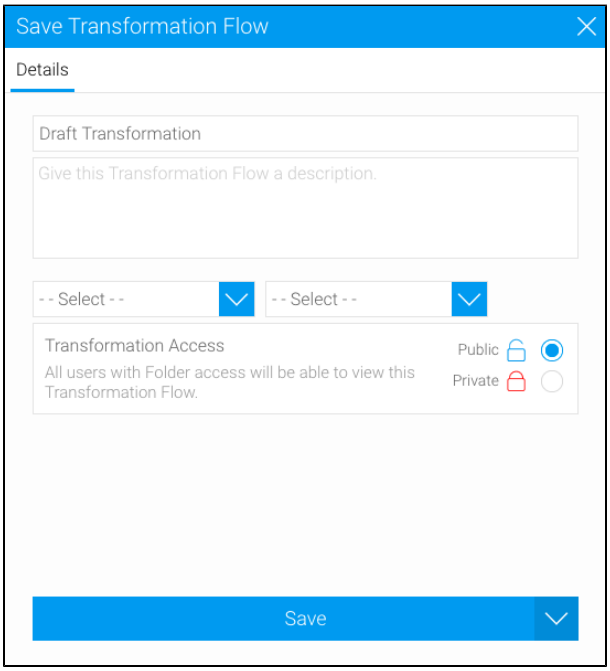
Once a flow has been created, you can save and publish it. Although you can manually execute draft flow, scheduled flows will have to be published in order to be executed.

Only valid flows may be published. A flow is valid when:

- It has at least one output step.
- All steps are configured.
- All steps have the correct number of inputs and outputs connected.

1. Click on the Publish button to do so.

2. Provide details in the popup that appears, such as giving your transformation flow a proper name and choosing the correct access rights.



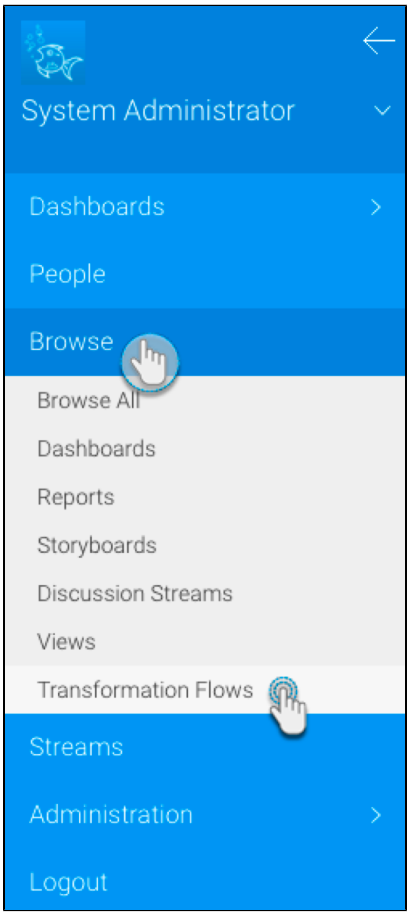
go back

Once a flow is created, you can either execute it manually or schedule its execution.

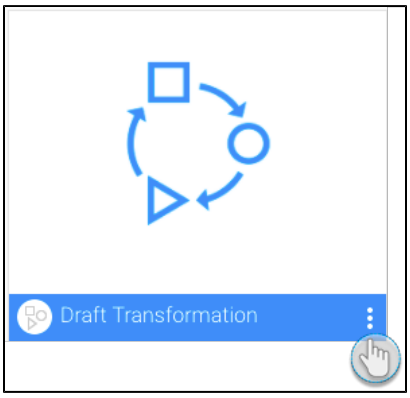
Manual Execution

This process shows how to execute the transformation flow manually:

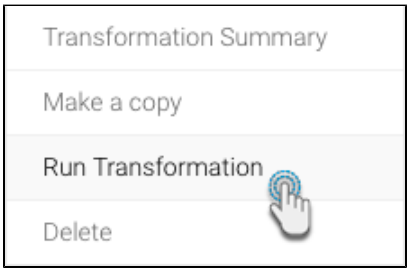
1. Navigate to left side menu > Browse > Transformation Flows.



2. Locate the flow that you want to run, and click on its menu icon. (The menu icon will appear when you hover the cursor on the bottom-right corner of the thumbnail. This menu can also be invoked by right-clicking the flow.)



3. Then click on the Run Transformation option.

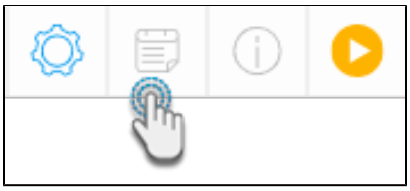


go back

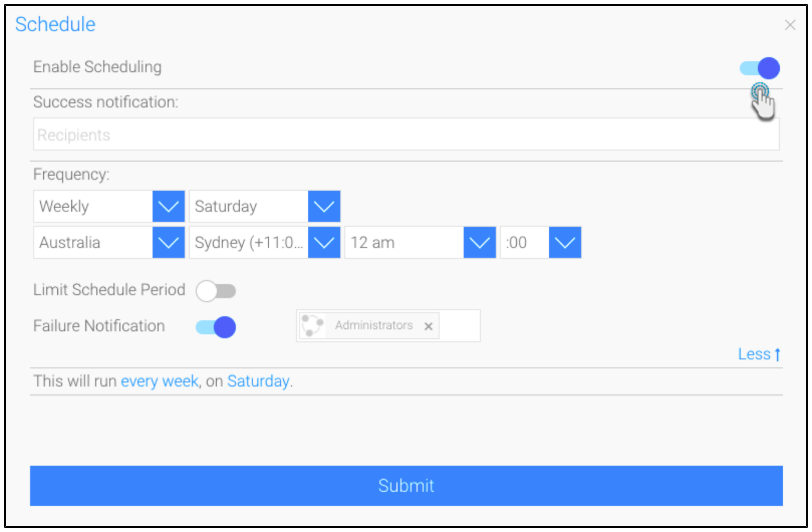
Scheduled Execution

Here's how to schedule a data transformation flow to be executed through the scheduler:

1. Click on the schedule button in the transformation builder's header menu.

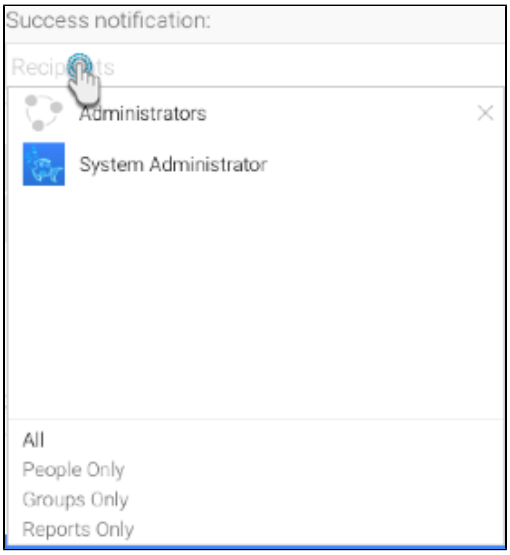


2. When the scheduling popup appears, click on the button to enable it.

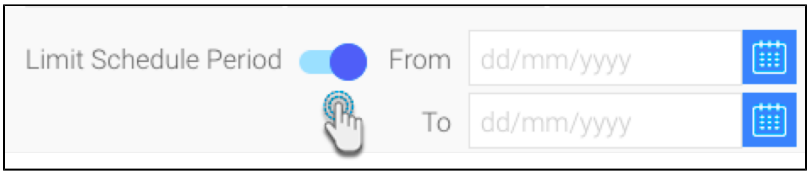


3. Provide scheduling details to run the batch at specified intervals:

a. Click on the Recipients textbox to select users to notify when the batch has executed successfully. You can add individual users and even groups.



b. Choose frequency with which to run the transformation flow.
c. Enable option to limit schedule period. This allows you to establish a time period within which the batch will be executed. Once the end period passes, the transformation will stop being executed.



d. Enable the Failure Notification button and add user groups to notify if the transformation failed to be executed.

e. Finally, click on the Submit button.

go back