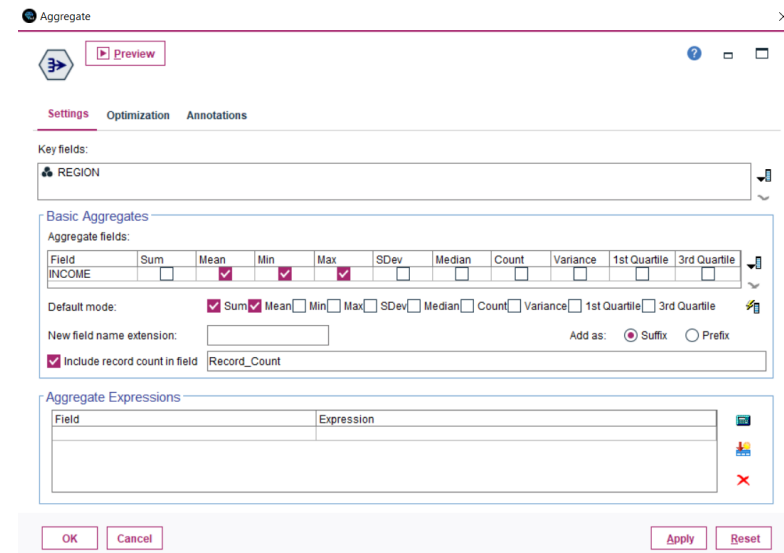
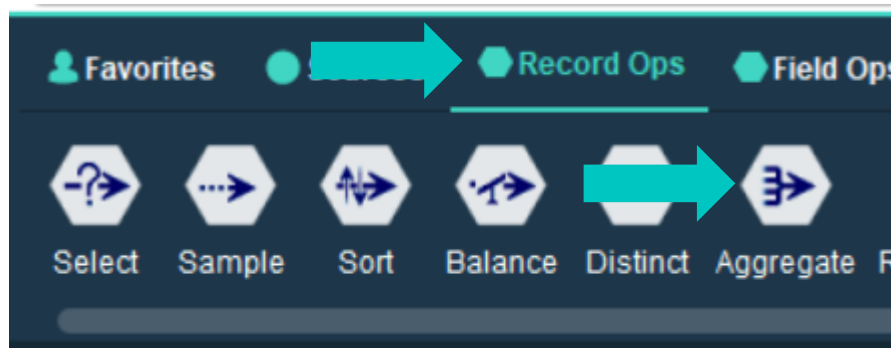


Aggregate Records

Tech Tips - IBM SPSS Modeler

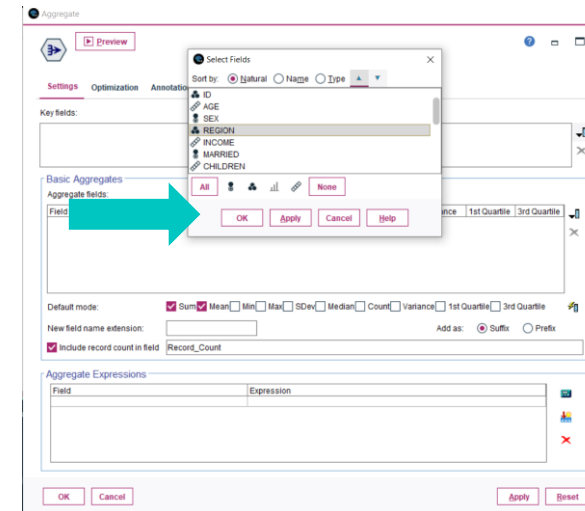
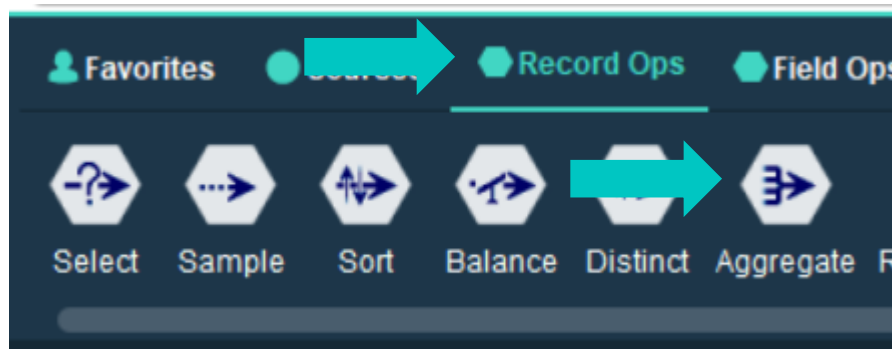
Tech Tips – Aggregate Records

- Here’s a quick tip to aggregate records in IBM SPSS Modeler.
- The **Aggregate** node makes data aggregation quick and easy and is located on the **Record Ops** palette.



Tech Tips – Aggregate Records

- To aggregate data go to the **Record Ops** palette. Select the **Aggregate** node and drag it onto the stream canvas. You can also double click the node to drop it onto the stream canvas. Once it is on the canvas you can connect it to your stream.
- Double click to open the node. We want to aggregate our data by 'REGION' and obtain average, minimum and maximum incomes for each region. Click on the **Key fields** chooser button and select the field you want to aggregate data by. Click **OK**.



Tech Tips – Aggregate Records

- Next, select **Aggregate fields** using the field chooser button. We will select 'INCOME'. Click **OK**. Select statistics. We will tick mean, the minimum and maximum. Also, **Record Count** will provide counts for each region. This can be turned on or off. Click **OK**. Now your data will be aggregated. You can preview the data by clicking in the Preview button.

The first screenshot shows the 'Aggregate' dialog box with the 'Select Fields' sub-dialog open. The 'Select Fields' dialog has 'INCOME' selected in the list. The 'Aggregate' dialog has 'REGION' in the 'Key fields' section and 'INCOME' in the 'Aggregate fields' section. The 'Basic Aggregates' section for 'INCOME' has 'Mean', 'Min', and 'Max' checked. The 'Include record count in field' checkbox is checked, and 'Record_Count' is entered in the field name extension.

The second screenshot shows the 'Aggregate' dialog box after configuration. 'REGION' is the key field. 'INCOME' is the aggregate field with 'Mean', 'Min', and 'Max' selected. 'Record_Count' is the record count field.

The third screenshot shows the 'Preview from Aggregate Node (5 fields, 4 records) #2' window. It displays a table with the following data:

	REGION	INCOME_Mean	INCOME_Min	INCOME_Max	Record_Count
1	INNER_CITY	28343.703	6531	60635	145
2	RURAL	30633.200	8005	60176	40
3	TOWN	24762.927	5960	61476	110
4	SUBURBAN	26211.254	9785	51468	63



Thank You

For more information
please visit spssanalyticspartner.com