Mark Friedman
University of Miami

Using Alteryx to Join Data Streams

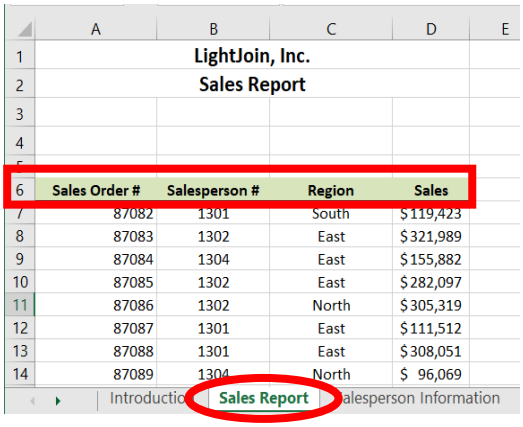
Alteryx is used by several Big Four accounting firms to automate the ETL process. In this exercise, you will:

Extract the "*Sales Report"*, "Salesperson Information" and "Comm SalesPerson by Region" data from Excel.

Transform the data into a flat data set in which all the data for a specific sales order is on one row. Aggregate the commission in dollars for each salesperson by region.

Load the transformed data into Excel.

There are three streams of data for *LightJoin, Inc.* stored in an Excel workbook. The first stream, sales information is stored in a sheet titled "Sales Report". Notice that the column headings are in the sixth row.



	A	B	C	D	E
1		LightJoin, Inc.			
2		Sales Report			
3					
4					
6	Sales Order #	Salesperson #	Region	Sales	
7	87082	1301	South	$119,423	
8	87083	1302	East	$321,989	
9	87084	1304	East	$155,882	
10	87085	1302	East	$282,097	
11	87086	1302	North	$305,319	
12	87087	1301	East	$111,512	
13	87088	1301	East	$308,051	
14	87089	1304	North	$ 96,069	

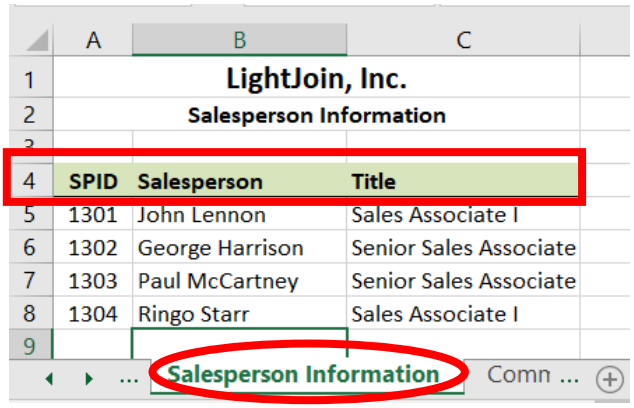
Introductio | Sales Report | alesperson Information

Notice that the last "Sales Order #" is 87175, and there are 94 rows of data (row 100 – row 6 = 94 rows).

6	Sales Order #	Salesperson #	Region	Sales
97	87172	1302	West	$154,798
98	87173	1304	North	$353,819
99	87174	1304	East	$256,322
100	87175	1303	North	$ 78,635

Mark Friedman
University of Miami

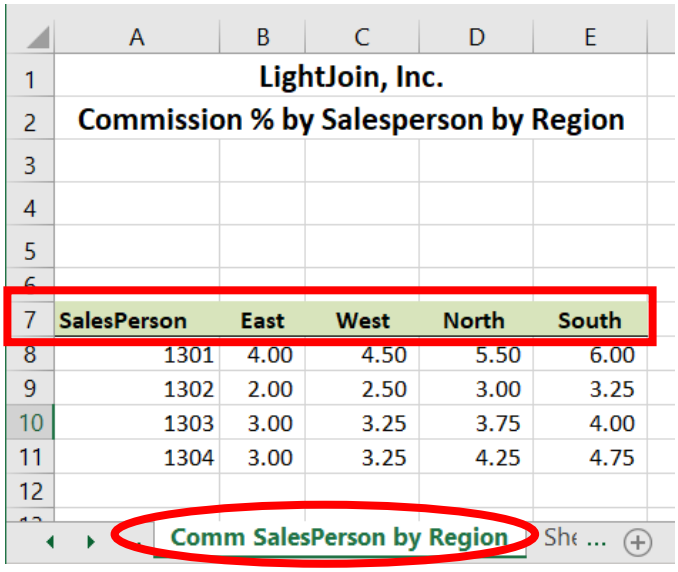
The second stream contains data regarding the Salesperson's name and title are stored in a sheet titled "Salesperson Information". Notice that the column headings are in the fourth row.



	A	B	C
1		LightJoin, Inc.	
2		Salesperson Information	
4		SPID	Salesperson	Title
5		1301	John Lennon	Sales Associate I
6		1302	George Harrison	Senior Sales Associate
7		1303	Paul McCartney	Senior Sales Associate
8		1304	Ringo Starr	Sales Associate I

Salesperson Information | Comm ...

The third stream contains the commission percentage earned by the salesperson by region in which the sales were made. Data regarding the Salesperson's commission rate is stored in a sheet titled "Comm SalesPerson by Region". Notice that the column headings are in the seventh row.



	A	B	C	D	E
1	LightJoin, Inc.				
2	Commission % by Salesperson by Region				
7	SalesPerson	East	West	North	South
8	1301	4.00	4.50	5.50	6.00
9	1302	2.00	2.50	3.00	3.25
10	1303	3.00	3.25	3.75	4.00
11	1304	3.00	3.25	4.25	4.75

Comm SalesPerson by Region | She ...

The deliverable is a commission in dollars report bysalesperson, loaded back into Excel.

	A	B
1	Salesperson	Commission Expense
2	George Harrison	$90,935.34
3	John Lennon	$226,149.84
4	Paul McCartney	$127,193.52
5	Ringo Starr	$274,189.61

Mark Friedman
University of Miami

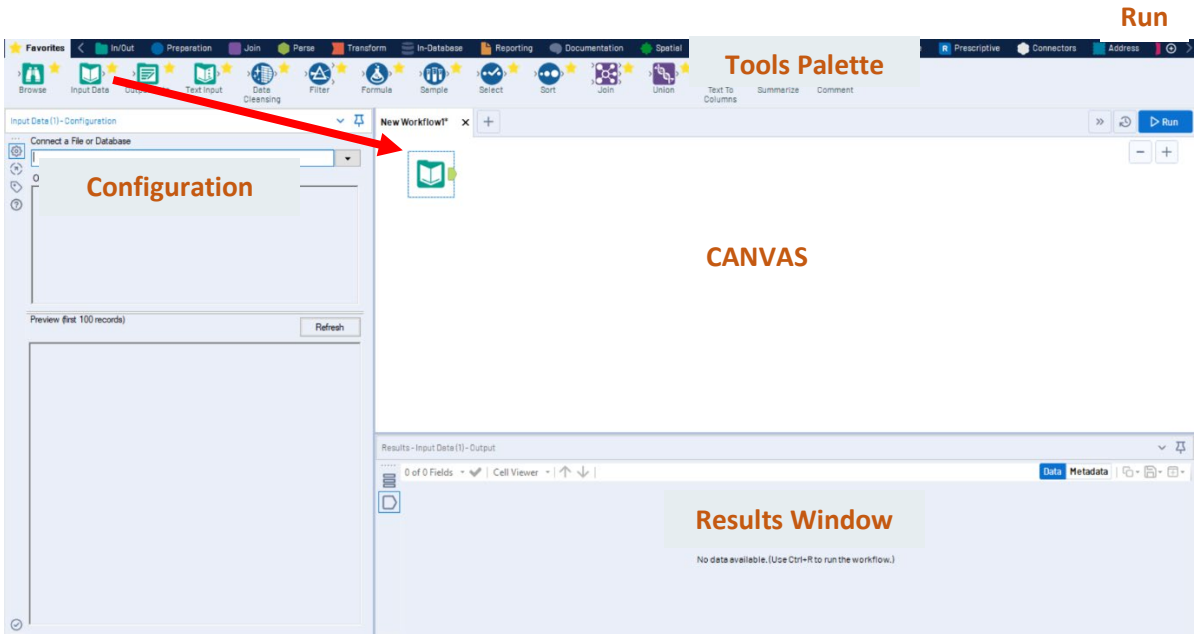
Start Alteryx

Note: You cannot extract data from an open Excel file, so make sure that the Excel file is closed before starting to input the data.

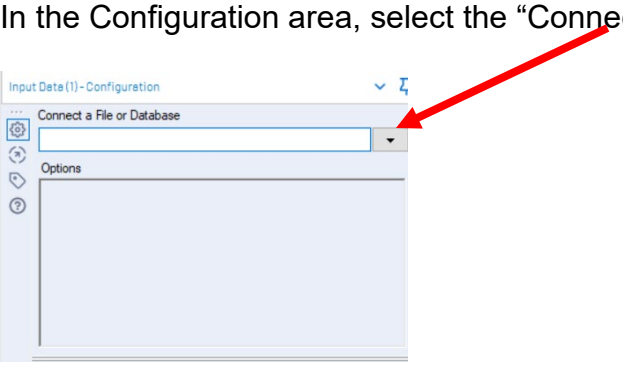
There are four major parts of the Alteryx window. The first is the Canvas, and that's where the workflow will be created. Second, the Tools Palette is where all the tools are stored. Third, the Configuration area is where the tool settings are modified. Lastly, the Results window shows the outcome after "Run" is selected to apply the changes.

Tool 1: Input Data Tool – Input Excel file

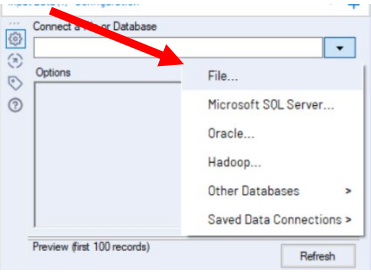
Drag the "Input Data Tool"  from the Tools Palette to the Canvas.



In the Configuration area, select the “Connect a File or Database” drop-down arrow.

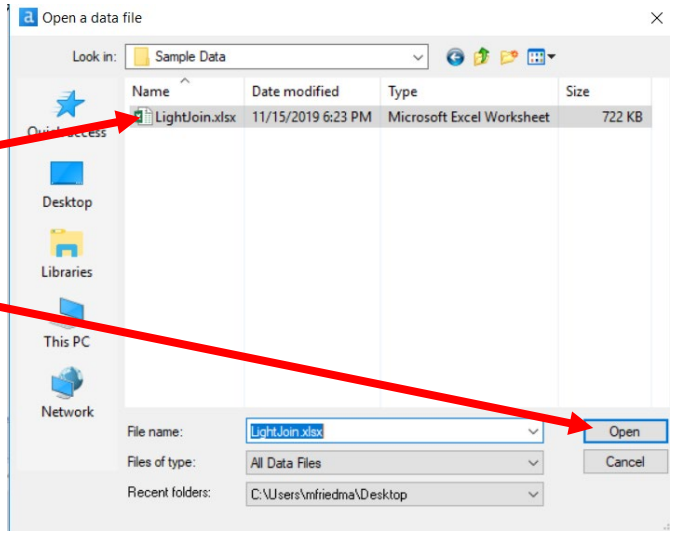


Select “File”.



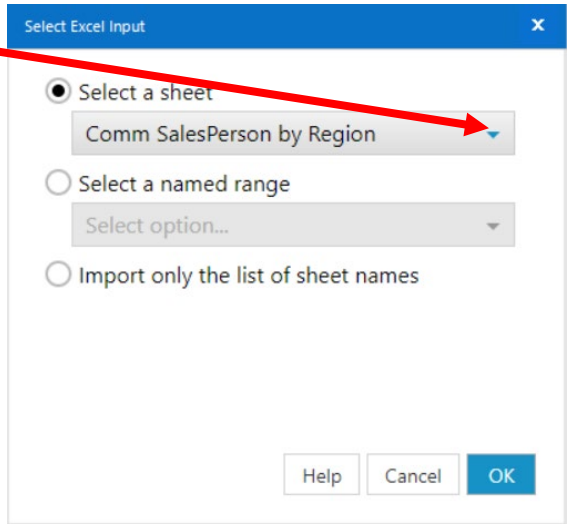
Browse for the desired file and select the version with the .xlsx file extension.

Select “Open”.



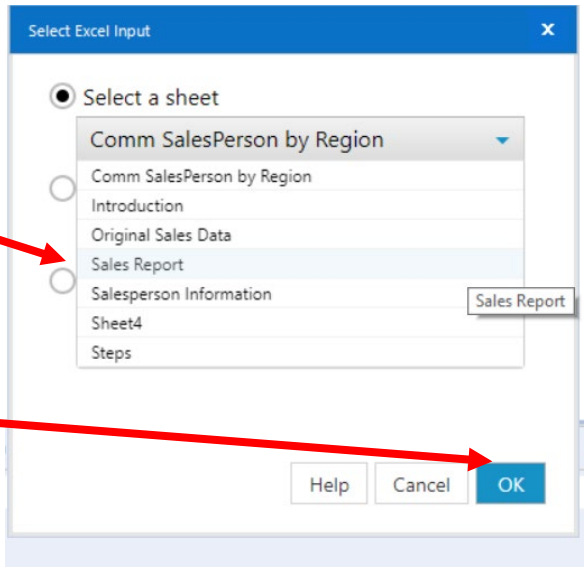
Mark Friedman
University of Miami

Select the drop-down arrow to the right of “Comm SalesPerson by Region”.

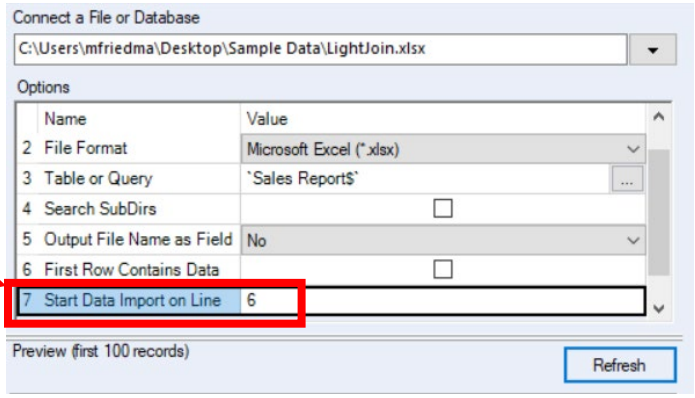


Select “Sales Report”.

Select “OK”.

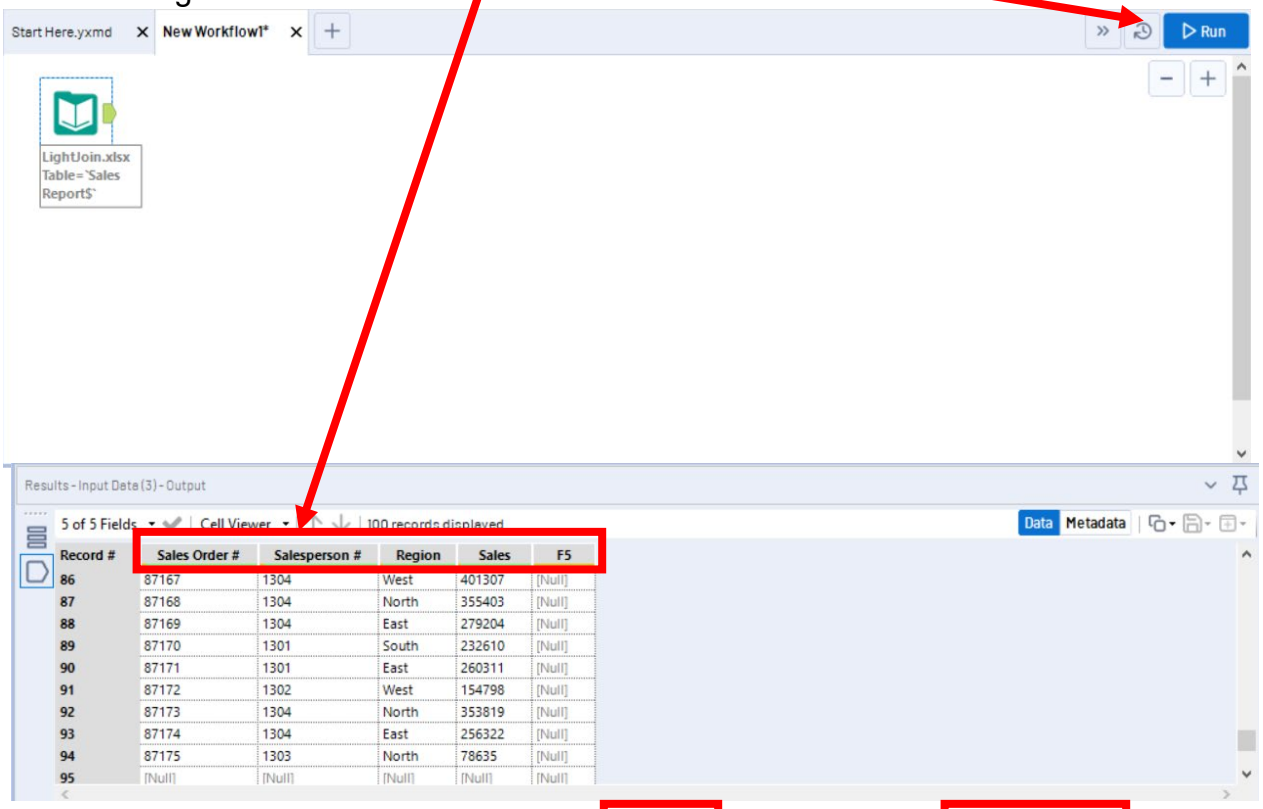


Remember that the field names are in the sixth row.
In the Configuration area, change the number in “Start Data Import on Line” to “6”.

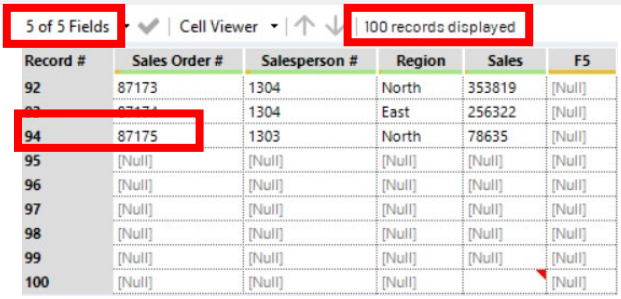


Select “Run”.

The Results window will display the field headings and the data.



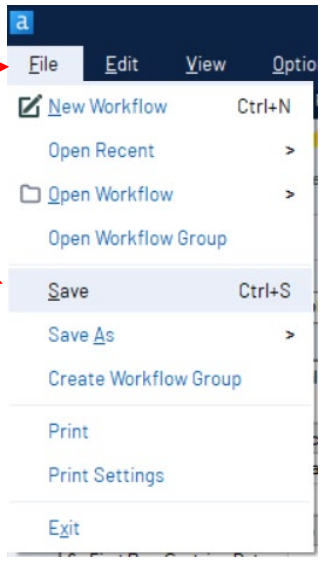
Notice that five fields were extracted while we need only four fields and that there are 100 records. The extra rows that contain “Null” and the extra fields must be removed. There are 94 rows of data and the last one is “Sales Order #” 87175.



The workflow must be saved.

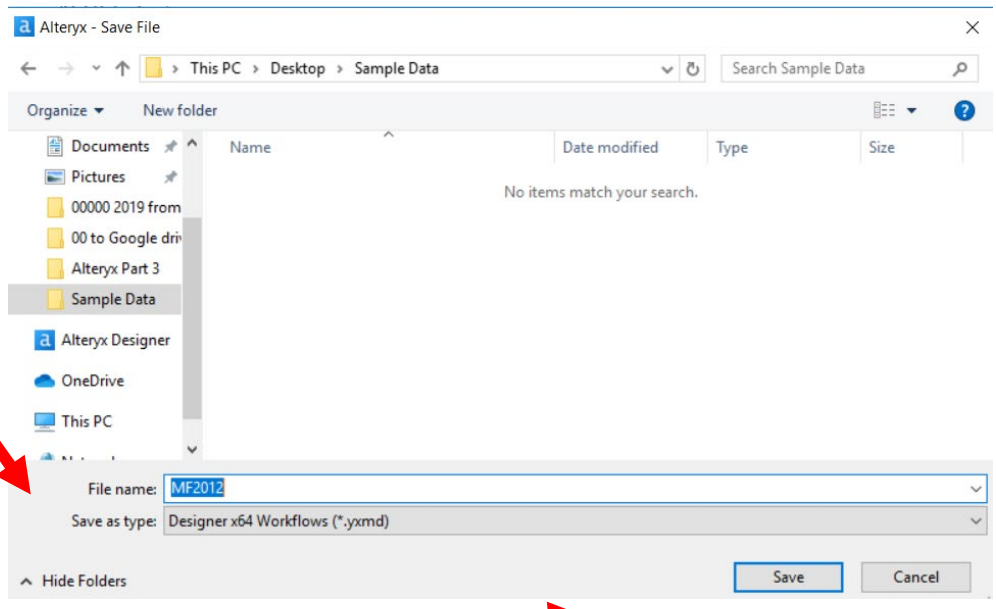
Select “File”.

Select “Save”.



Browse to the required directory.

Type in the name of the file that will be uploaded.

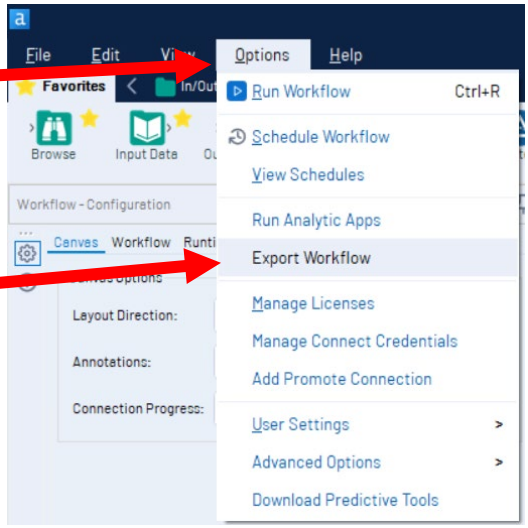


Select “Save”.

Export the workflow and the related input files so that the project can be completed on another computer or uploaded to the web.

Select “Options”.

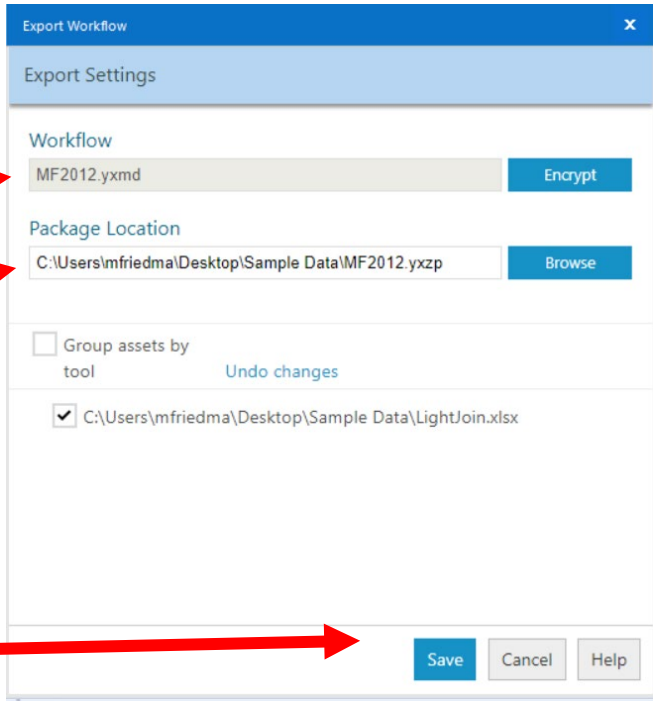
Select “Export Workflow”.



Verify the “Workflow”.

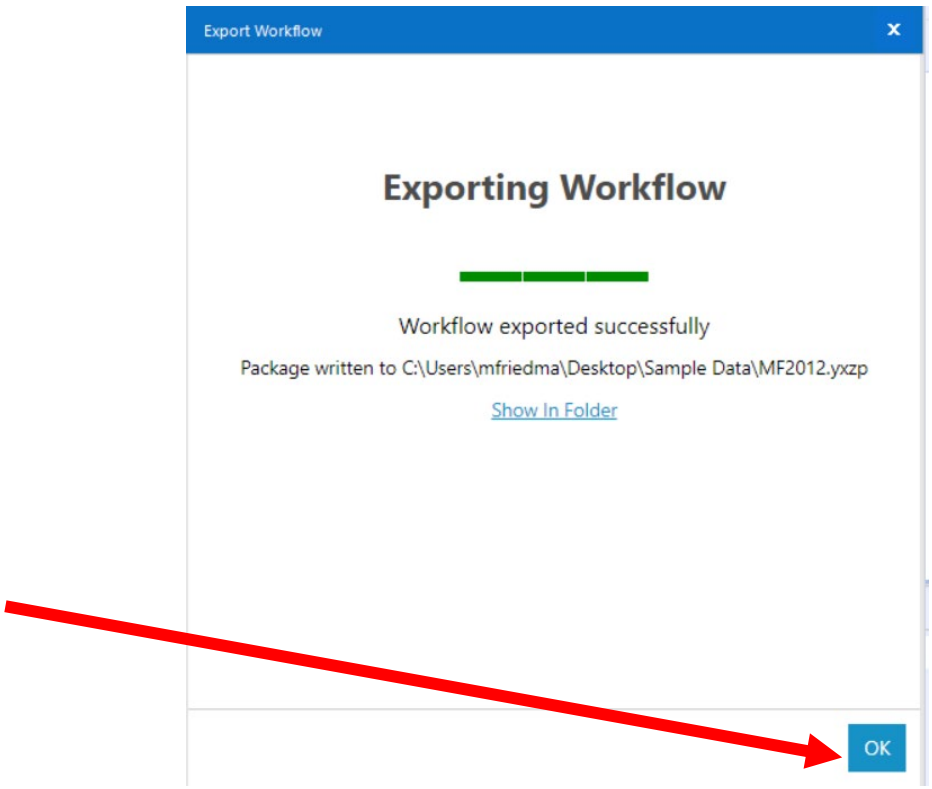
Verify the “Package Location”.

Select “Save”.



Mark Friedman
University of Miami

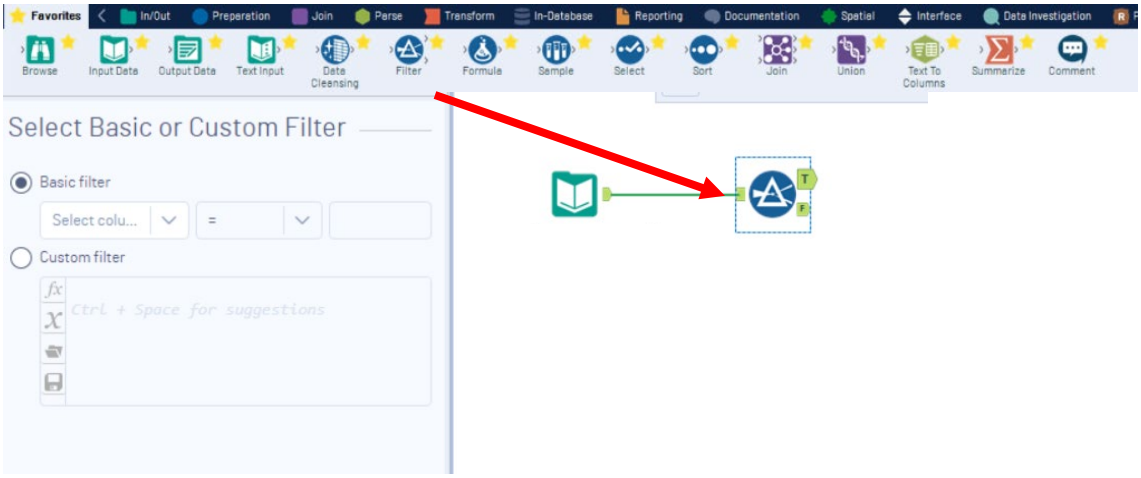
Select “OK”



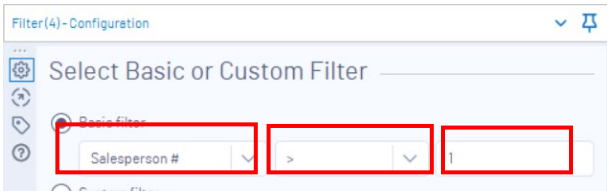
Tool 2: Filter Tool – Exclude all rows that do not contain data

There are several rows that do not contain a sales order number. We can modify the working dataset so that it includes only those rows that contain sales order numbers greater than one. Note that nothing is being deleted from the source data file. The working dataset will be split into two subsets. The first subset, "True", will include all rows with "Sales Order #" greater than one. The second subset, "False", will contain all rows with "Sales Order #" less than one.

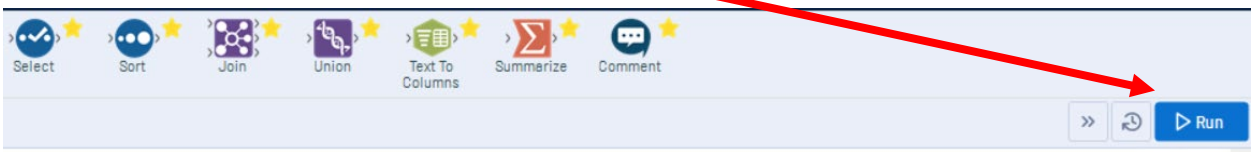
Drag the "**Filter Tool**"  from the Tools Palette to the right of the "Input Data Tool" on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



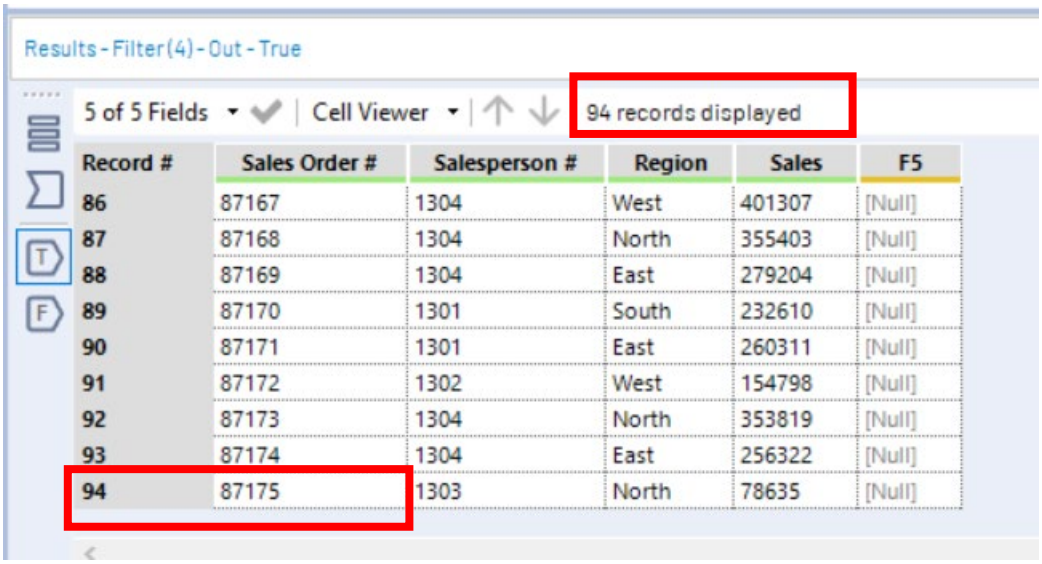
With the "**Filter Tool**" selected, the "Basic" filter in the Configuration area can be modified to filter the field "Salesperson #" for those accounts greater than "1'.



To apply the latest change, select "Run".



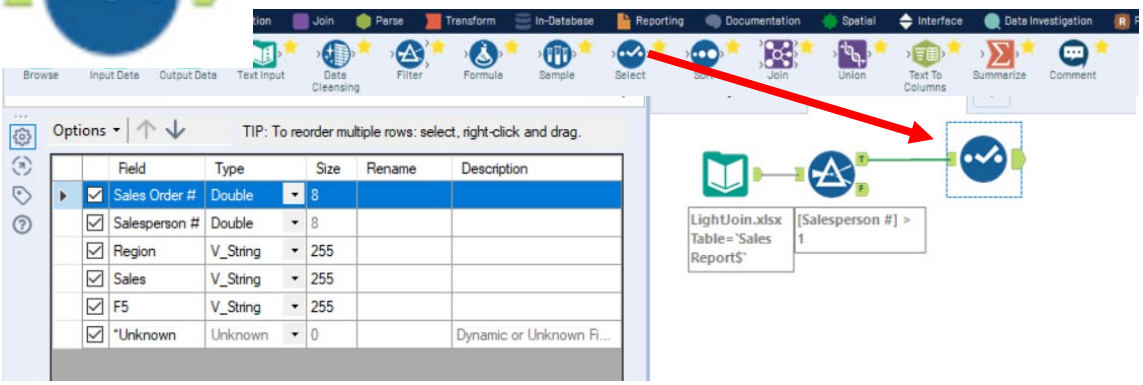
The records that meet the filtered conditions, “Salesperson #” > 1”, are displayed as T (True). There are 94 records displayed and the last one is “Sales Order #” 87175.



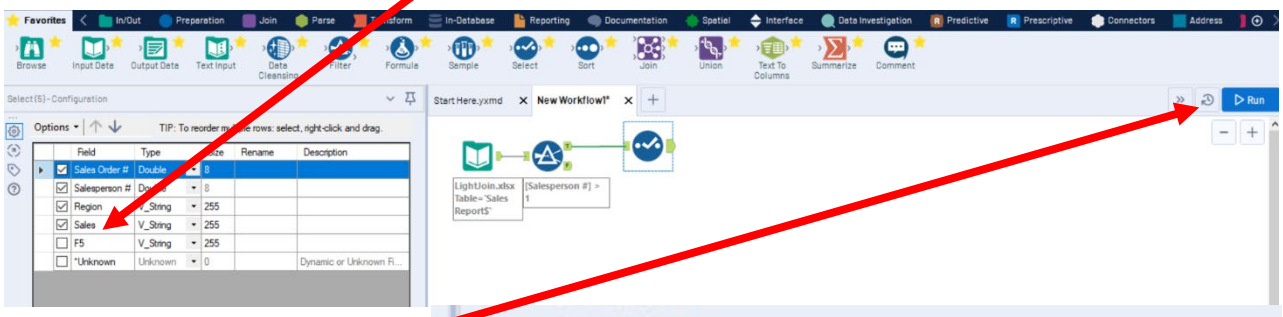
Record #	Sales Order #	Salesperson #	Region	Sales	F5
86	87167	1304	West	401307	[Null]
87	87168	1304	North	355403	[Null]
88	87169	1304	East	279204	[Null]
89	87170	1301	South	232610	[Null]
90	87171	1301	East	260311	[Null]
91	87172	1302	West	154798	[Null]
92	87173	1304	North	353819	[Null]
93	87174	1304	East	256322	[Null]
94	87175	1303	North	78635	[Null]

Tool 3: Select Tool – Remove fields that contain no data, F5 and Unknown

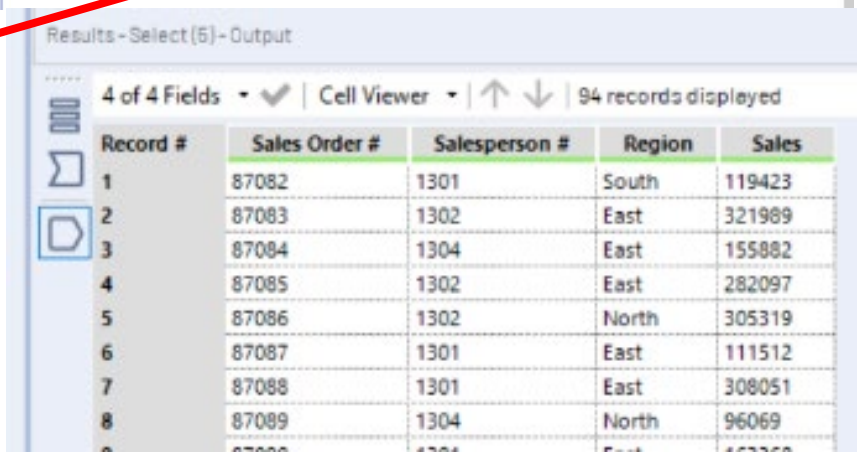
Drag the "**Select Tool**" from the Tools Palette to the True output of the "**Filter Tool**" on the Canvas. The two will be connected automatically.



Remove the extra fields by unchecking them.

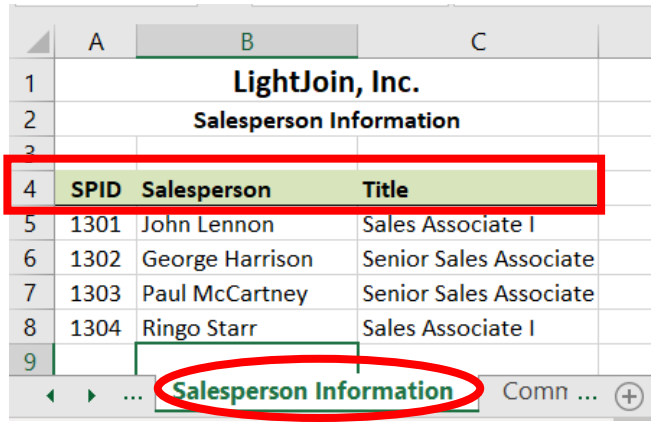


Select "Run".



Repeat Tools 1- 3 and extract the data from the “Salesperson Information” sheet.

Data regarding the Salesperson’s name and title are stored in an Excel sheet titled “Salesperson Information”. Notice that the column headings are in the fourth row.



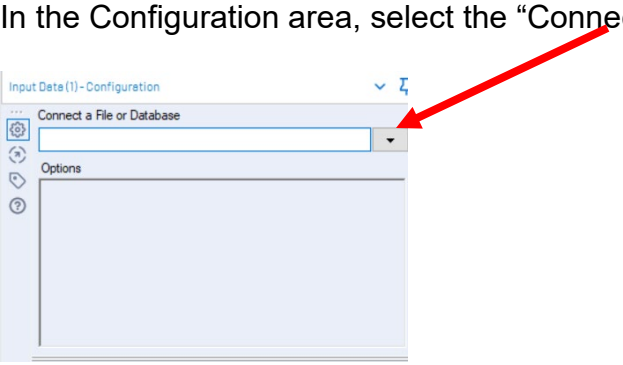
	A	B	C
1		LightJoin, Inc.	
2		Salesperson Information	
4	SPID	Salesperson	Title
5	1301	John Lennon	Sales Associate I
6	1302	George Harrison	Senior Sales Associate
7	1303	Paul McCartney	Senior Sales Associate
8	1304	Ringo Starr	Sales Associate I
9			

Salesperson Information Comn ...

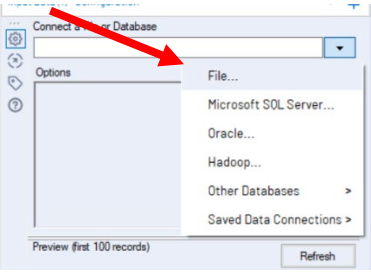
Tool 4: Input Data Tool – Input Excel file

Drag the “Input Data Tool” from the Tools Palette to the Canvas.

In the Configuration area, select the “Connect a File or Database” drop-down arrow.

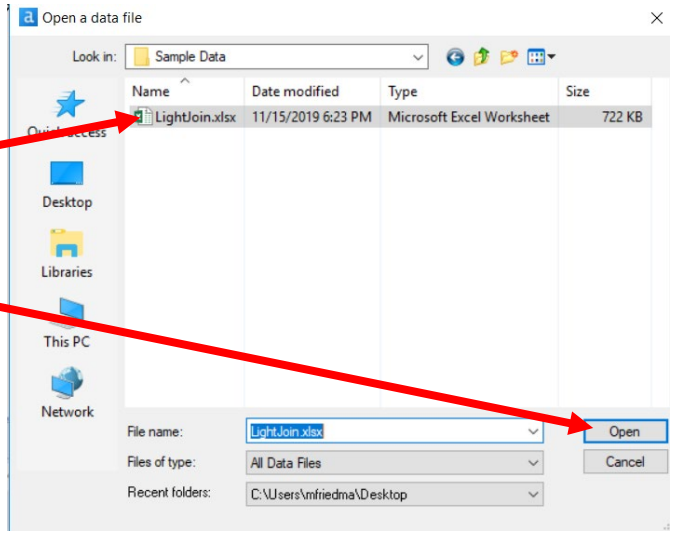


Select “File”.

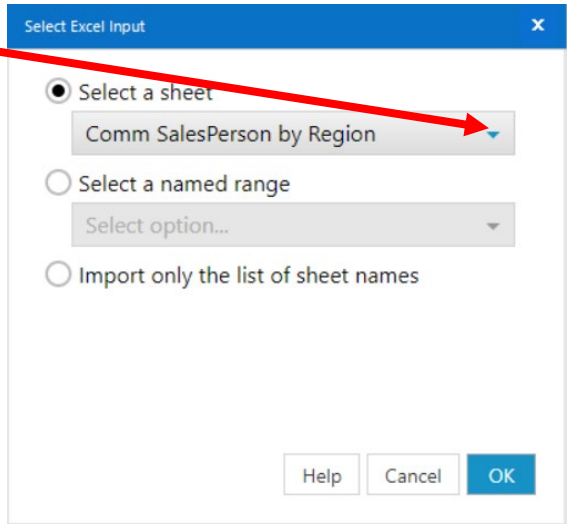


Browse for the desired file and select the version with the .xlsx file extension.

Select “Open”.

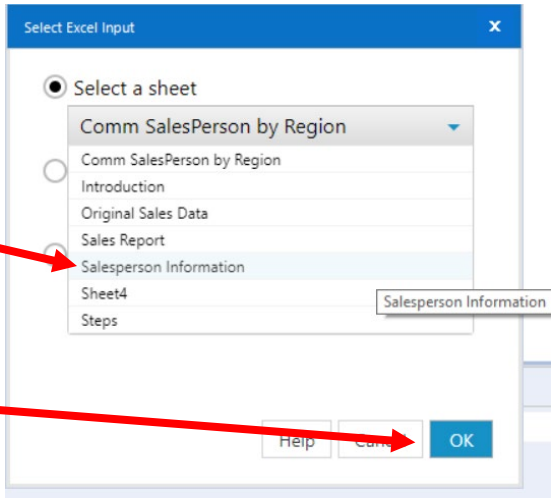


Select the drop-down arrow to the right of “Comm SalesPerson by Region”.

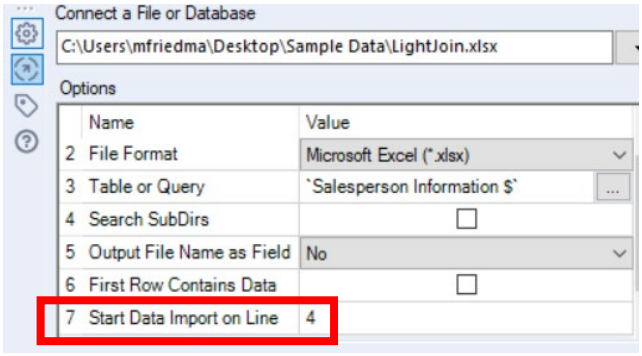


Select “Salesperson Information”.

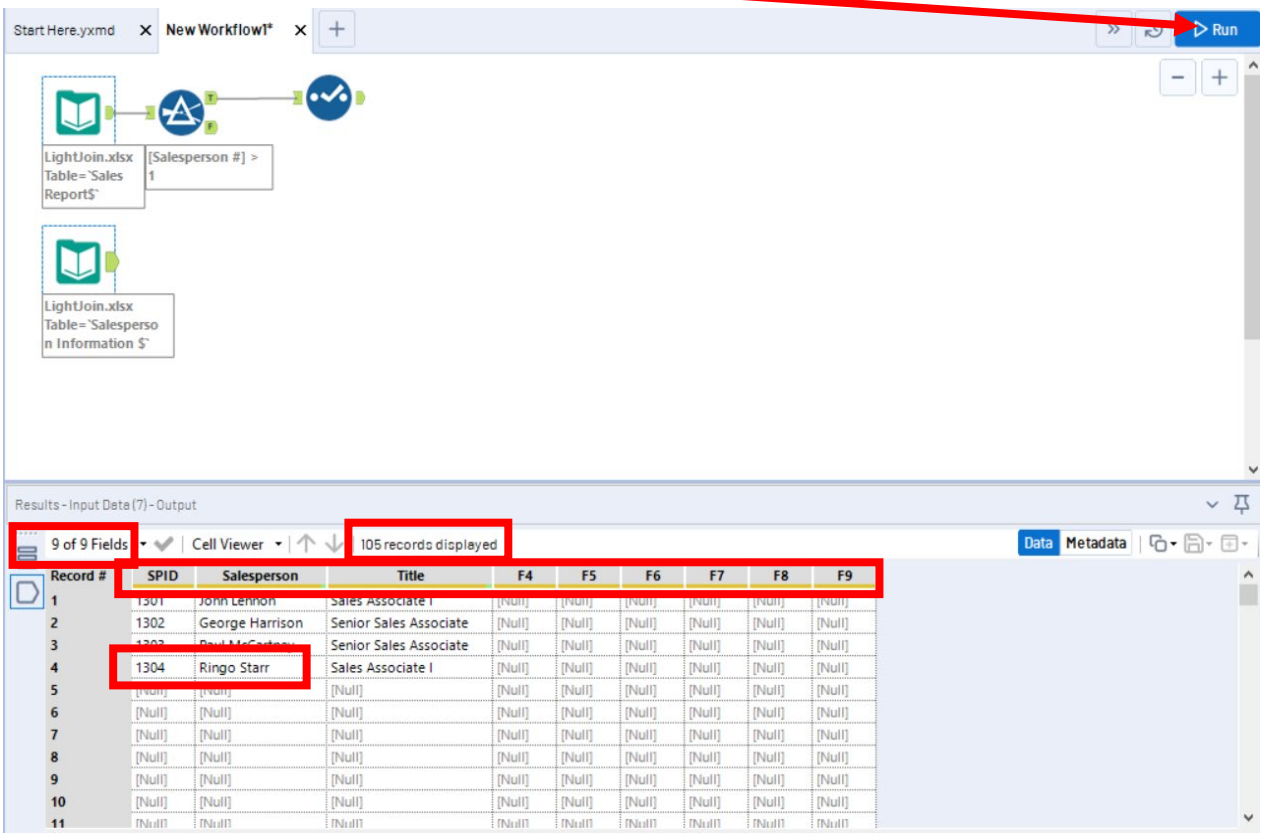
Select “OK”.



Remember that the field names are in the fourth row.
In the Configuration area, change the number in “Start Data Import on Line” to “4”.



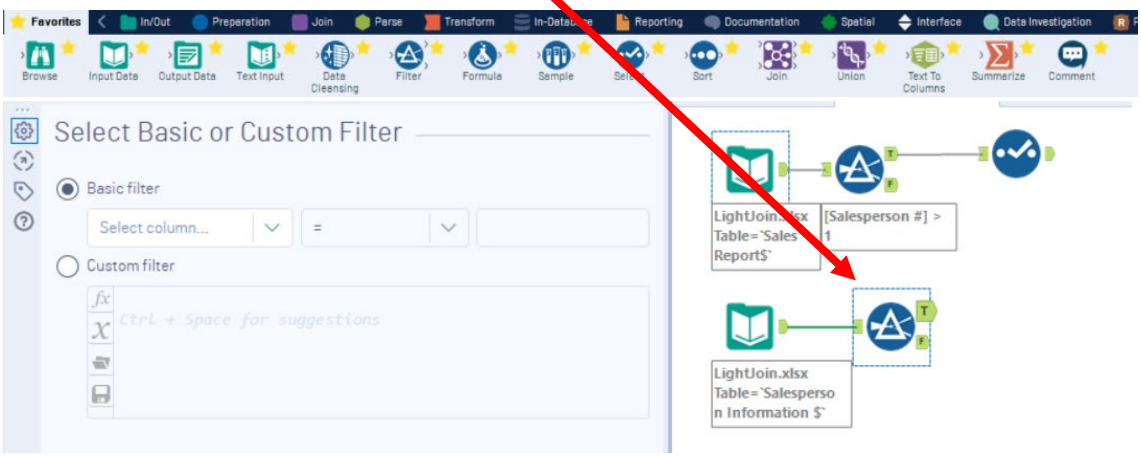
Select “Run”.



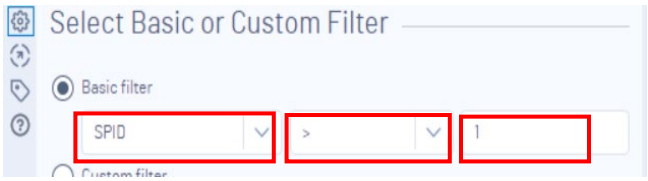
The Results window will display the field headings and the data. Notice that nine fields were extracted while only three fields are needed and that there are 105 records. The extra rows that contain [Null] and the extra fields must be removed. There are four rows of data and the last Salesperson ID, “SPID” is 1304.

Tool 5: Filter Tool – Exclude all rows that do not contain data

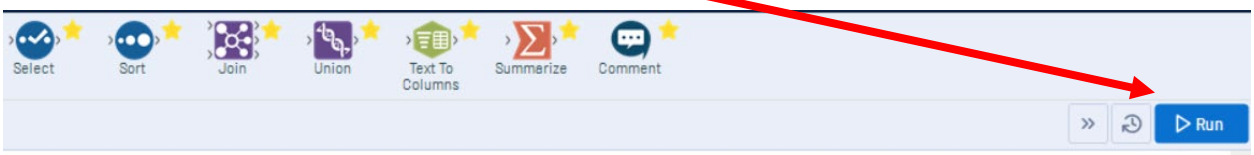
Drag the "**Filter Tool**"  from the Tools Palette to the right of the "Input Data Tool" on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



With the "**Filter Tool**" selected, the "Basic" filter in the Configuration area can be modified to filter the field "SPID" for those accounts greater than "1'.



To apply the latest change, select "Run".

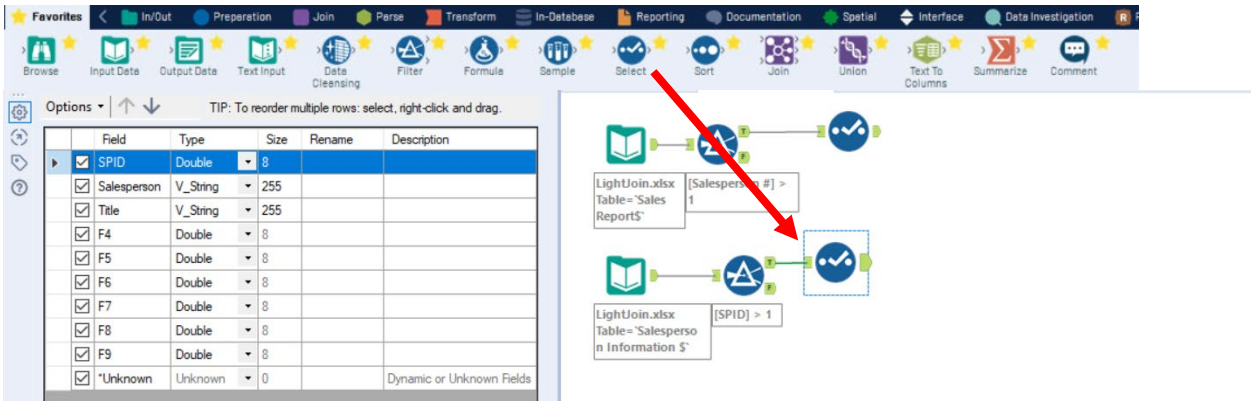


The records that meet the filtered conditions, "SPID > 1", are displayed as T (True). There are four records displayed and the last one is "SPID" 1304.

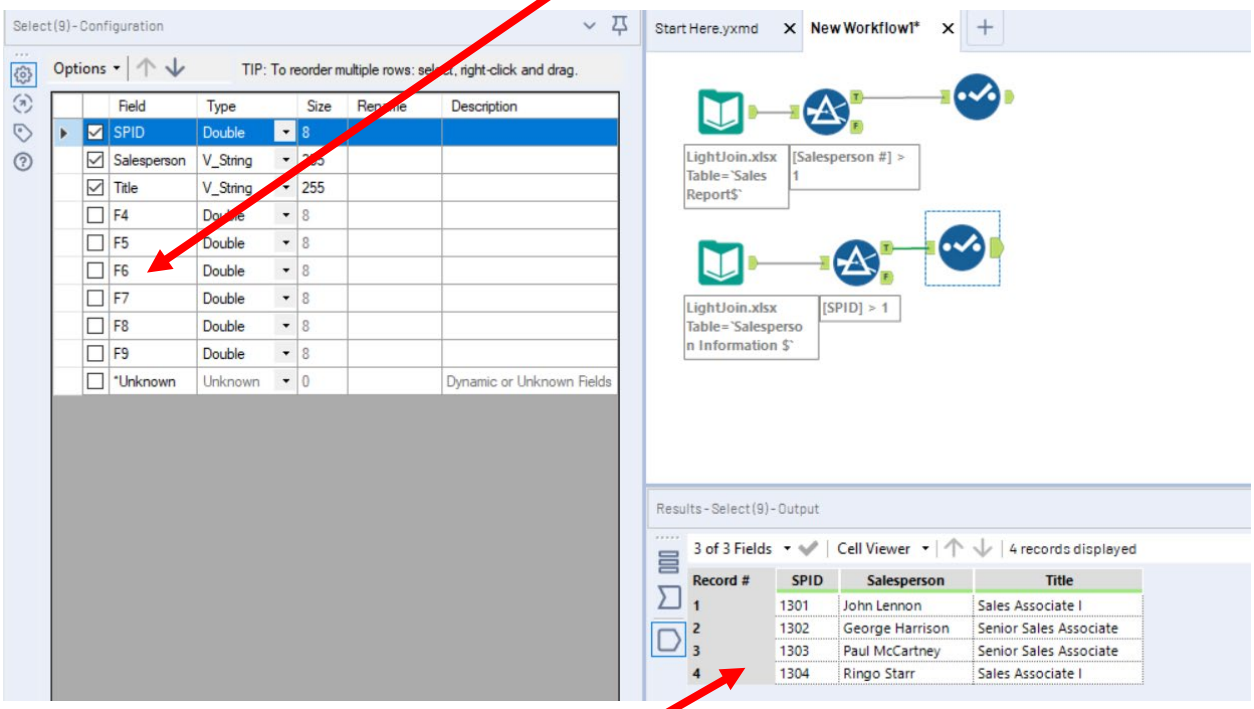
Record #	SPID	Salesperson	Title	F4	F5	F6	F7	F8	F9
1	1301	John Lennon	Sales Associate I	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
2	1302	George Harrison	Senior Sales Associate	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
3	1303	Paul McCartney	Senior Sales Associate	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
4	1304	Ringo Starr	Sales Associate I	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]

Tool 6: Select Tool – Remove Fields That Contain No Data

Drag the "**Select Tool**" from the Tools Palette to the True output of the "**Filter Tool**" on the Canvas. The two will be connected automatically.



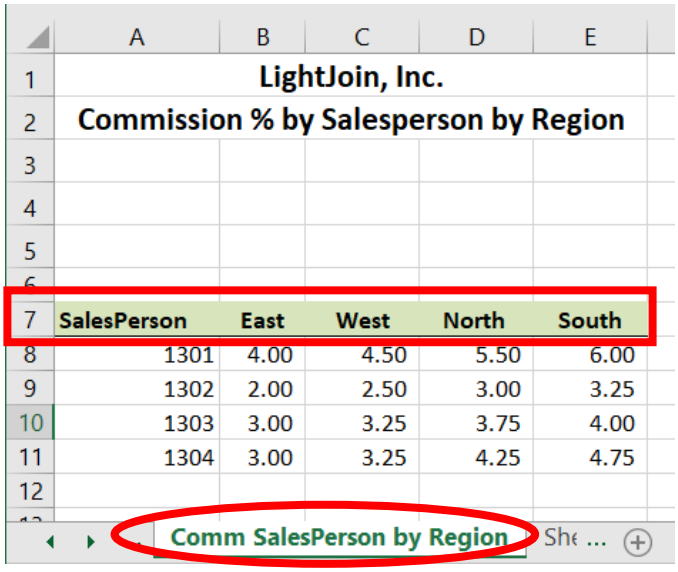
Remove the extra fields by unchecking them.



Select "Run".

Review the data.

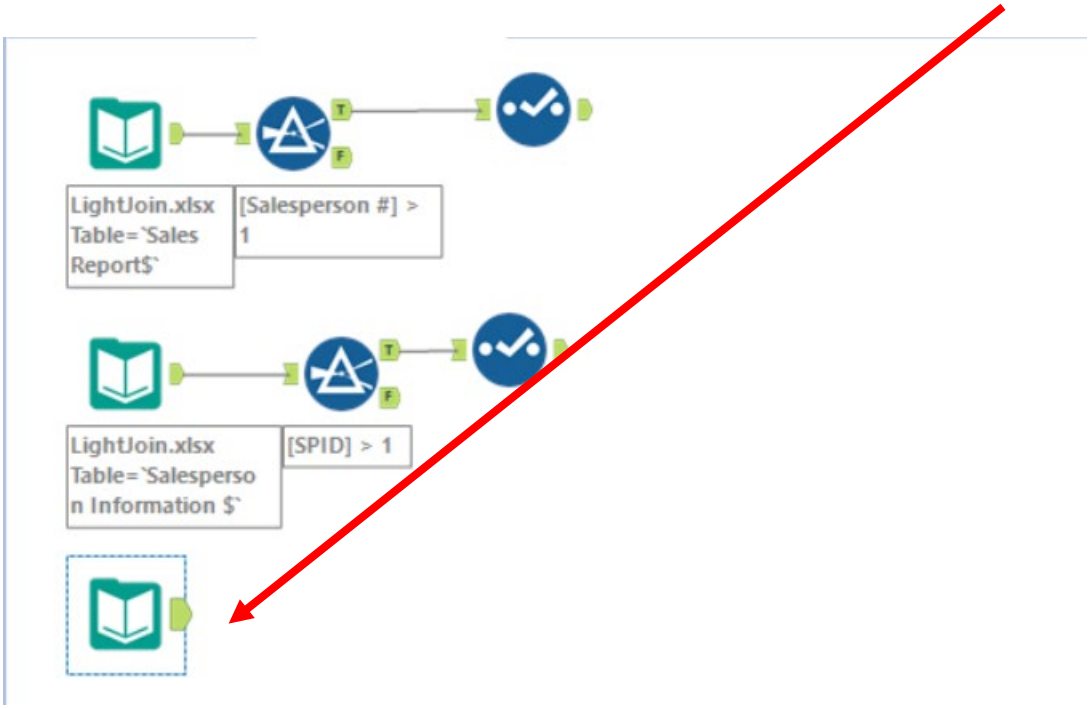
The commission percentage earned by the salesperson is a function of the region in which the sales was made. Data regarding the Salesperson's commission rate are stored in an Excel sheet titled "Comm SalesPerson by Region". Notice that the column headings are in the seventh row.



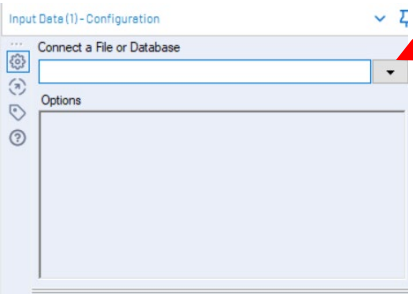
	A	B	C	D	E
1	LightJoin, Inc.				
2	Commission % by Salesperson by Region				
3					
4					
5					
7	SalesPerson	East	West	North	South
8	1301	4.00	4.50	5.50	6.00
9	1302	2.00	2.50	3.00	3.25
10	1303	3.00	3.25	3.75	4.00
11	1304	3.00	3.25	4.25	4.75
12					

Tool 7: Input Data Tool – Input Excel File

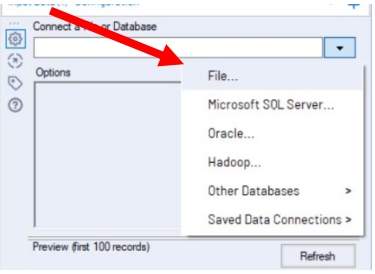
Drag the "Input Data Tool" from the Tools Palette to the Canvas.



In the Configuration area, select the “Connect a File or Database” drop-down arrow.

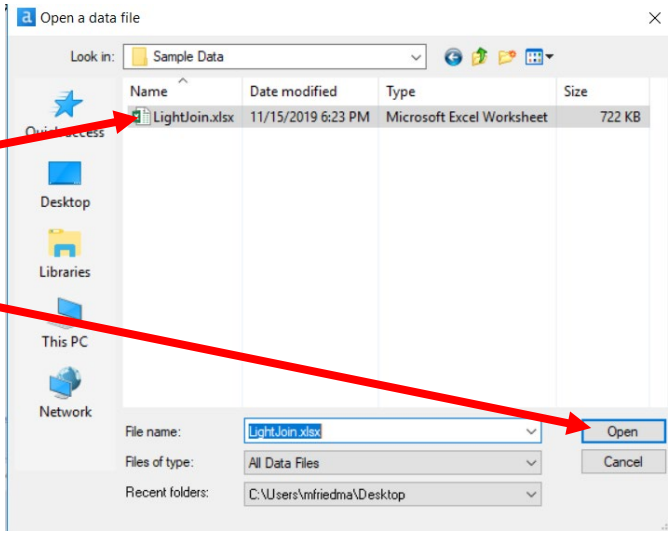


Select “File”.



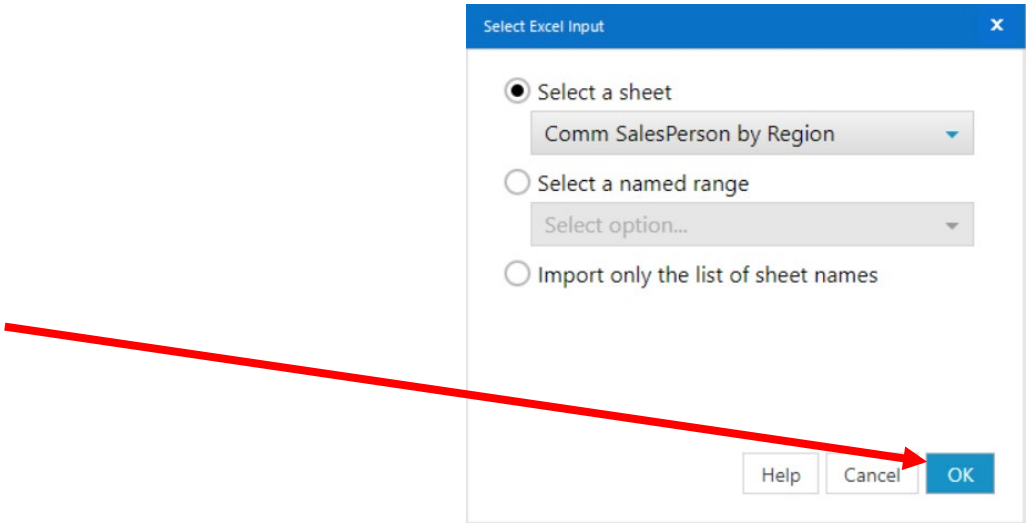
Browse for the desired file and select the version with the .xlsx file extension.

Select “Open”.

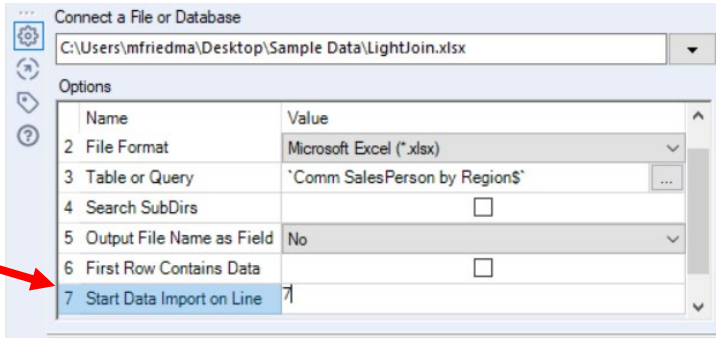


The “Comm SalesPerson by Region” sheet is the default sheet.

Select “OK”.

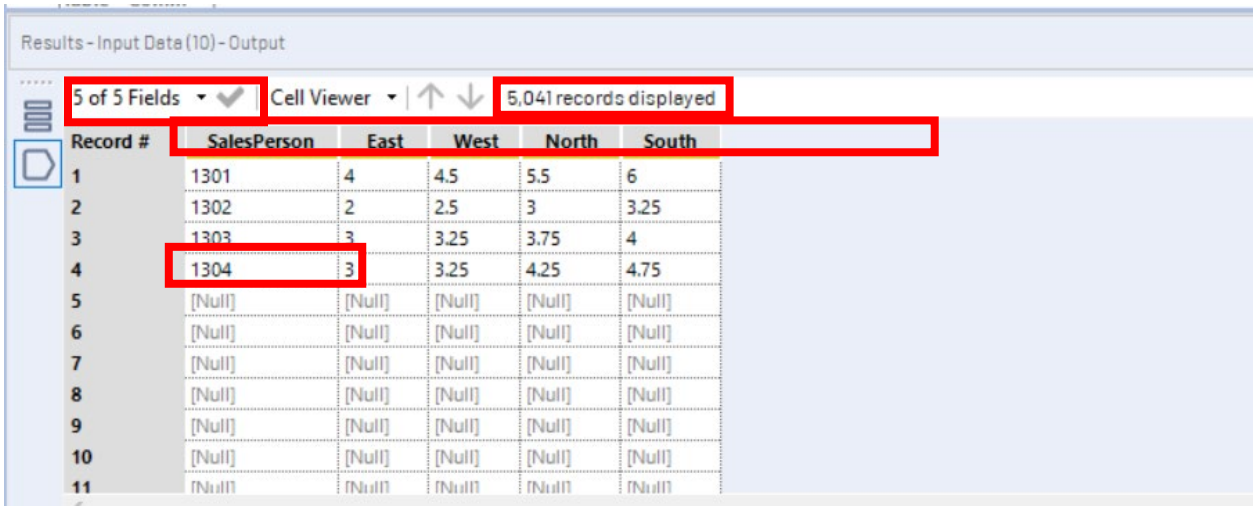


Remember that the field names are in the seventh row.
In the Configuration area, change the number in “Start Data Import on Line” to “7”.



Select “Run”.

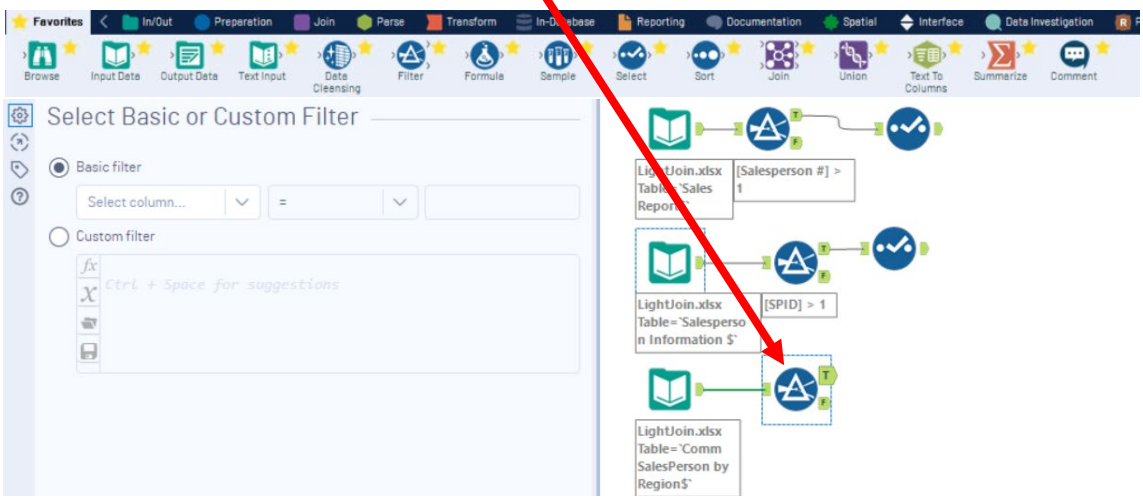
The Results window will display the field headings and the data. Notice that five fields were extracted and that there are 5041 records. The extra rows that contain [Null] and must be removed. There are four rows of data and the last “Salesperson” is 1304.



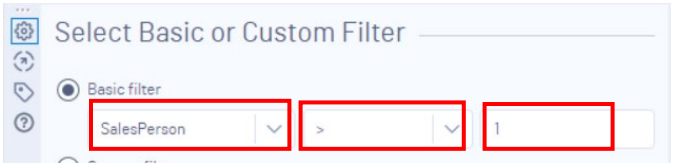
Record #	SalesPerson	East	West	North	South
1	1301	4	4.5	5.5	6
2	1302	2	2.5	3	3.25
3	1303	3	3.25	3.75	4
4	1304	3	3.25	4.25	4.75
5	[Null]	[Null]	[Null]	[Null]	[Null]
6	[Null]	[Null]	[Null]	[Null]	[Null]
7	[Null]	[Null]	[Null]	[Null]	[Null]
8	[Null]	[Null]	[Null]	[Null]	[Null]
9	[Null]	[Null]	[Null]	[Null]	[Null]
10	[Null]	[Null]	[Null]	[Null]	[Null]
11	[Null]	[Null]	[Null]	[Null]	[Null]

Tool 8: Filter Tool – Exclude All Rows That Do Not Contain Data

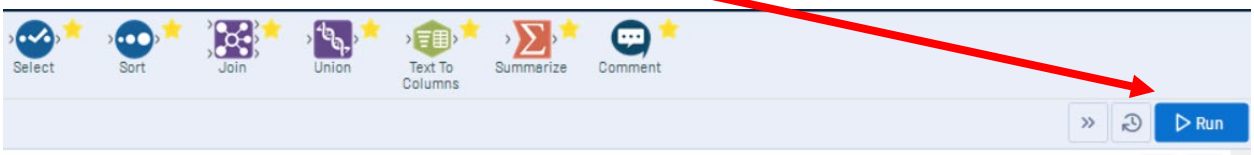
Drag the "**Filter Tool**"  from the Tools Palette to the right of the "Input Data Tool" on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



With the "**Filter Tool**" selected, the "Basic" filter in the Configuration area can be modified to filter the field "SPID" for those accounts greater than "1'.



To apply the latest change, select "Run".



The records that meet the filtered conditions, "SalesPerson> 1", are displayed as T (True). There are four records displayed and the last one is "SPID" 1304.

Results - Filter (11) - Out - True

5 of 5 Fields | Cell Viewer | 4 records displayed

Record #	SalesPerson	East	West	North	South
1	1301	4	4.5	5.5	6
2	1302	2	2.5	3	3.25
3	1303	3	3.25	3.75	4
4	1304	3	3.25	4.25	4.75

Review the streams of data:

Stream 1, Sales Report" file

4 of 4 Fields | Cell Viewer | 94 records displayed

Record #	Sales Order #	Salesperson #	Region	Sales
1	87082	1301	South	119423
2	87083	1302	East	321989
3	87084	1304	East	155882
4	87085	1302	East	282097
5	87086	1302	North	305319
6	87087	1301	East	111512
7	87088	1301	East	308051
8	87089	1304	North	96069
9	87090	1301	East	163360
10	87091	1304	North	269853
11	87092	1303	North	75460

Stream 2, Salesperson Information" file

3 of 3 Fields | Cell Viewer | 4 records displayed

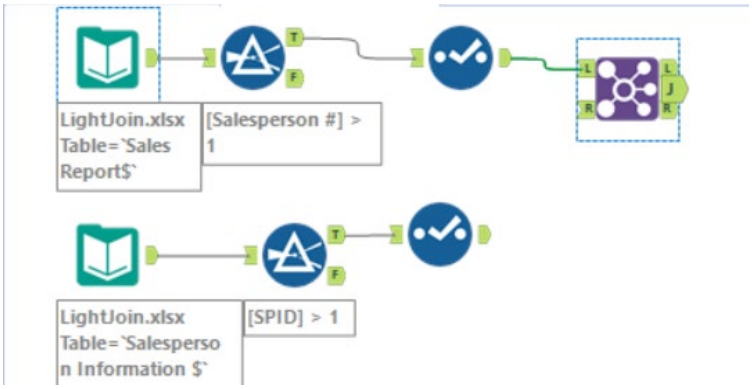
Record #	SPID	Salesperson	Title
1	1301	John Lennon	Sales Associate I
2	1302	George Harrison	Senior Sales Associate
3	1303	Paul McCartney	Senior Sales Associate
4	1304	Ringo Starr	Sales Associate I

The "Salesperson #" field in the "Sales Report" file appears many times while the "SPID" field is the primary key of the "Salesperson Information" file. Join the two streams.

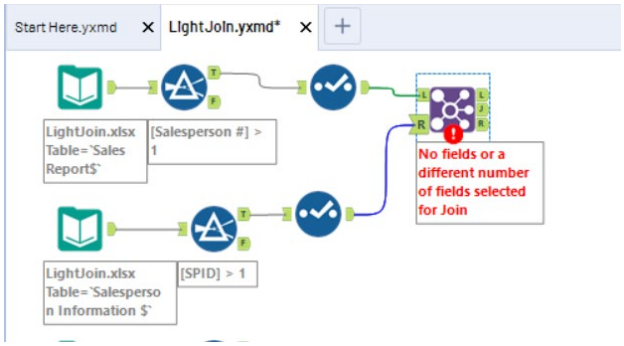
Tool 9: Join Tool – Combine Data From Two Streams On A Common Field

Drag the "**Join Tool**" 

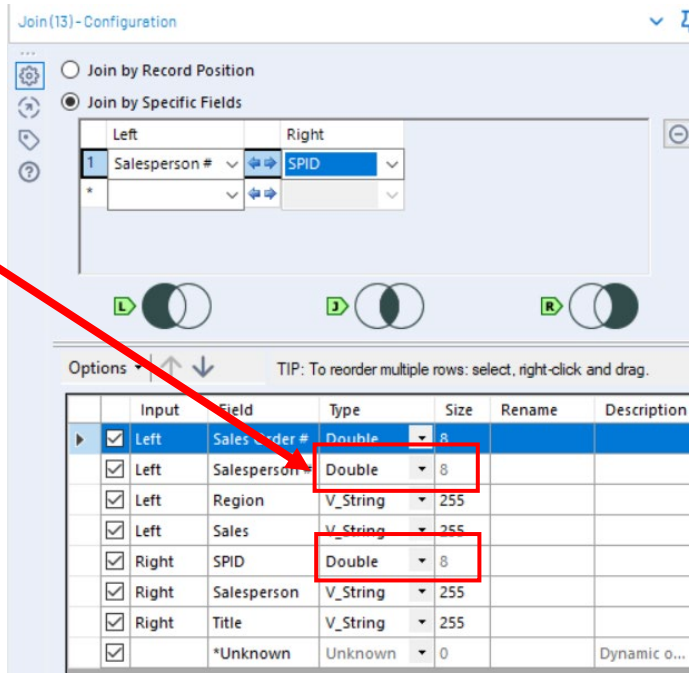
from the Tools Palette to the right of the "Select Tool" from the first data streams, on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



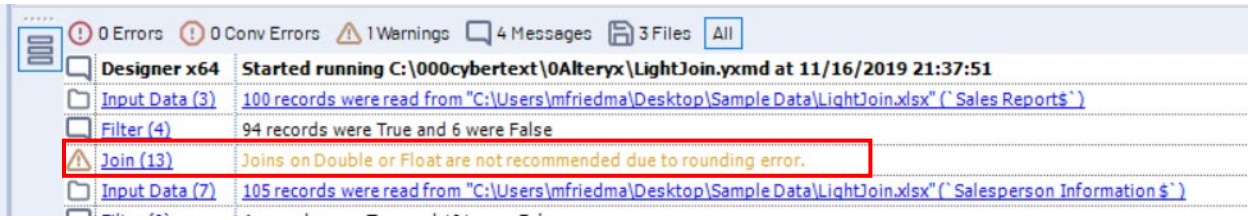
Connect the output of the "Select Tool" from the second data stream to the R,(Right) inflow of the "Join Tool".



With the "**Join Tool**" selected, in the Configuration area select "Salesperson #" on the left and "SPID" on the right. The two fields must be the same type; in this case "Double", and the same size, in this case 8.

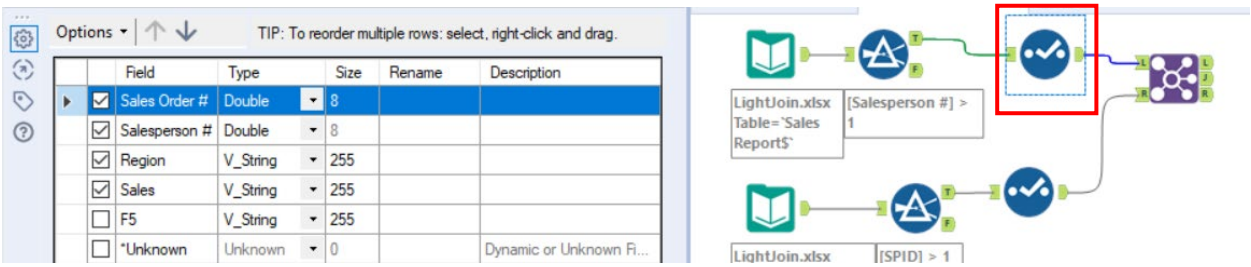


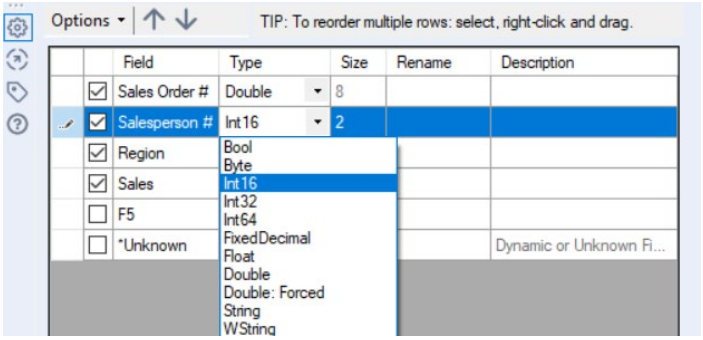
To apply the latest change, select "Run".



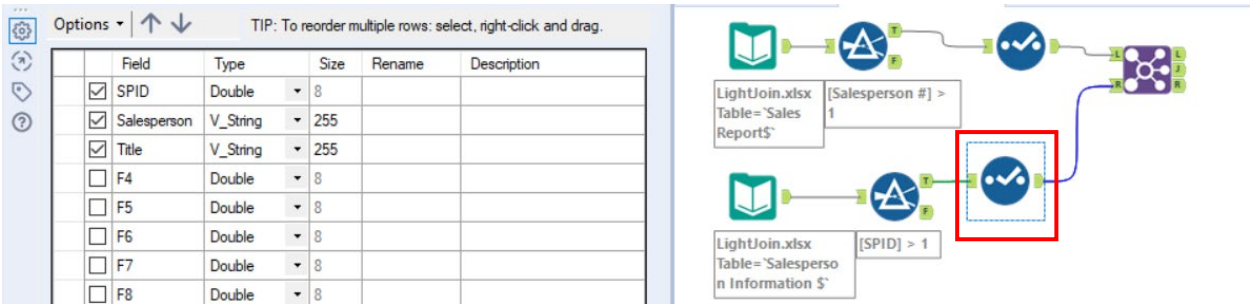
There is a possible error because the Join used "Double" fields.

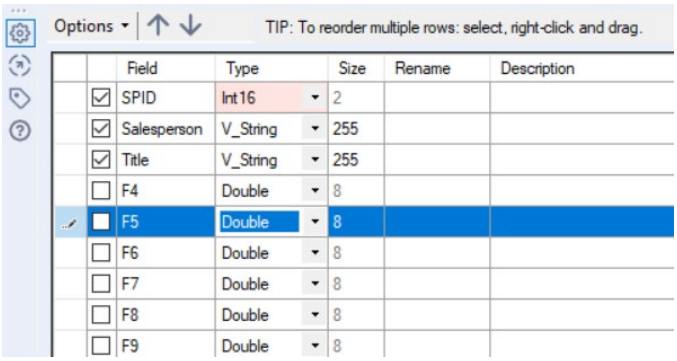
Highlight the Select Tool in the first stream and change the field type of the "Salesperson #" from Double to Integer, Int 16.





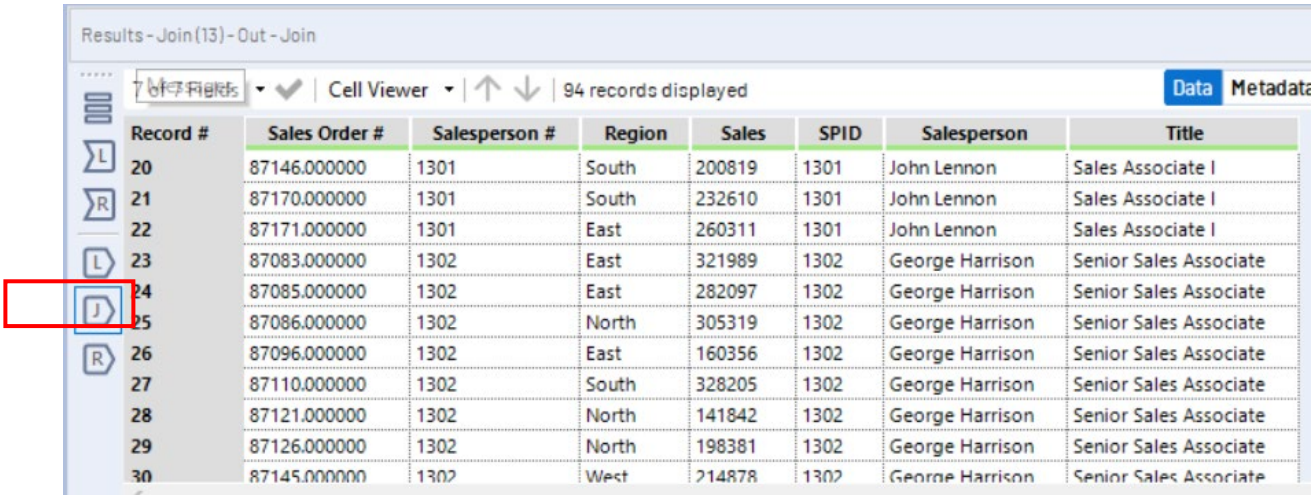
Highlight the Select Tool in the second stream and change the field type of the "SPID" from Double to Fixed Decimal.





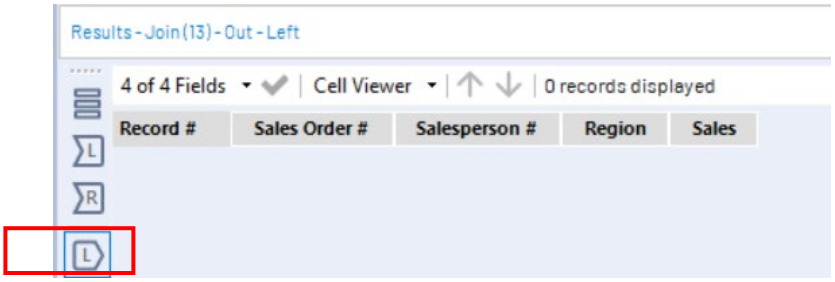
Select Run to apply.

In the result area, select J, Join, and review the output.



Record #	Sales Order #	Salesperson #	Region	Sales	SPID	Salesperson	Title
20	87146.000000	1301	South	200819	1301	John Lennon	Sales Associate I
21	87170.000000	1301	South	232610	1301	John Lennon	Sales Associate I
22	87171.000000	1301	East	260311	1301	John Lennon	Sales Associate I
23	87083.000000	1302	East	321989	1302	George Harrison	Senior Sales Associate
24	87085.000000	1302	East	282097	1302	George Harrison	Senior Sales Associate
25	87086.000000	1302	North	305319	1302	George Harrison	Senior Sales Associate
26	87096.000000	1302	East	160356	1302	George Harrison	Senior Sales Associate
27	87110.000000	1302	South	328205	1302	George Harrison	Senior Sales Associate
28	87121.000000	1302	North	141842	1302	George Harrison	Senior Sales Associate
29	87126.000000	1302	North	198381	1302	George Harrison	Senior Sales Associate
30	87145.000000	1302	West	214878	1302	George Harrison	Senior Sales Associate

Verify that nothing was unmatched from L, Left data set. Sales orders with an invalid salesperson's number would be unmatched.



Record #	Sales Order #	Salesperson #	Region	Sales

Look at R, Right data set. If anything is not matched, there would be salespeople that were not involved in any of the recorded sales.

Record #	SPID	Salesperson	Title

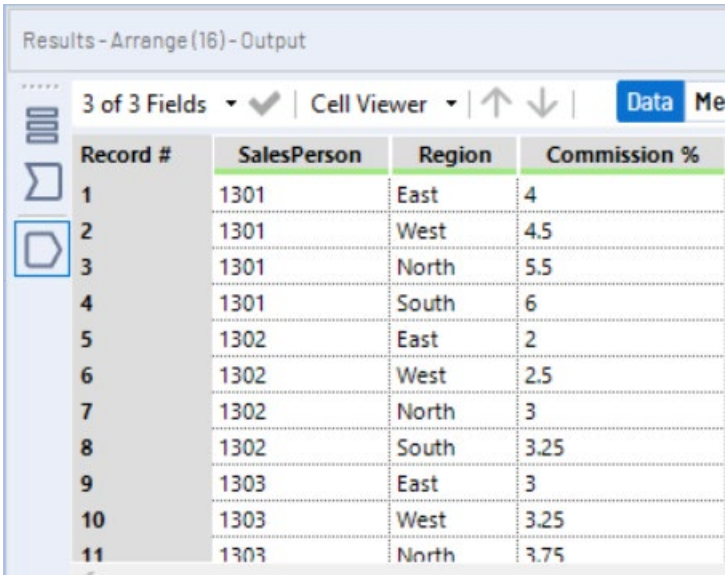
The third stream must be transformed before combining it with the already combined first and second stream of data. The third stream of data must be arranged into a data set with each commission rate on a separate row.

.

Third stream as extracted and transformed:

Record #	SalesPerson	East	West	North	South
1	1301	4	4.5	5.5	6
2	1302	2	2.5	3	3.25
3	1303	3	3.25	3.75	4
4	1304	3	3.25	4.25	4.75

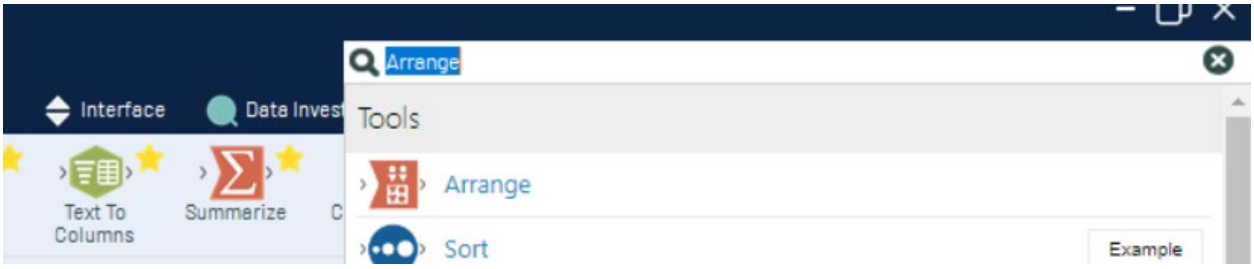
Third stream arranged into a data set with each commission rate on a separate row:



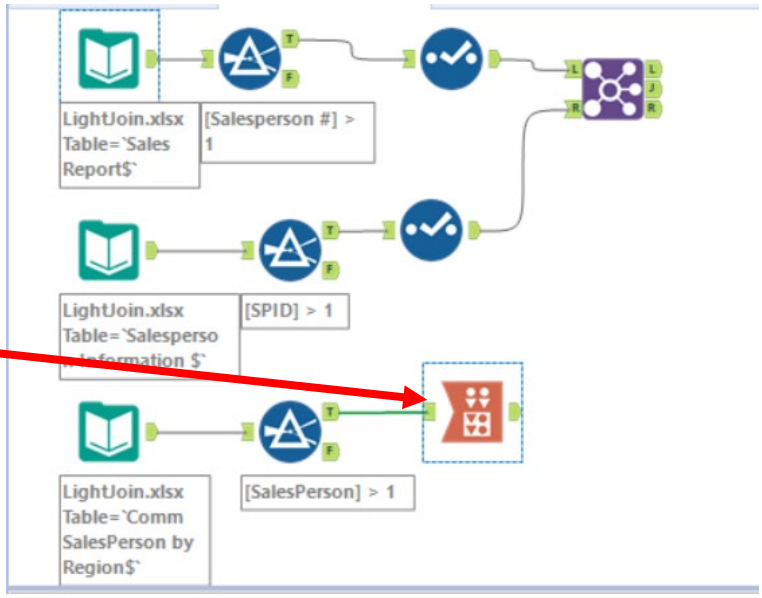
Record #	SalesPerson	Region	Commission %
1	1301	East	4
2	1301	West	4.5
3	1301	North	5.5
4	1301	South	6
5	1302	East	2
6	1302	West	2.5
7	1302	North	3
8	1302	South	3.25
9	1303	East	3
10	1303	West	3.25
11	1303	North	3.75

Tool 10: Arrange Tool – Each Record Is Turned Into Multiple Records

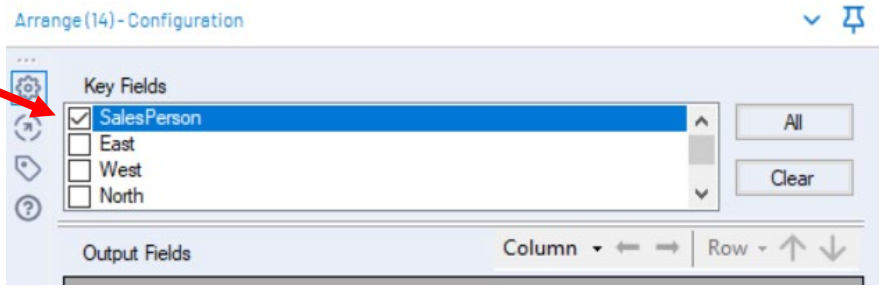
In the "Search for Tools Box" type "Arrange".



Drag the "Arrange Tool" from the Tools Palette to the right of the "Filter Tool" from the third data streams, on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.

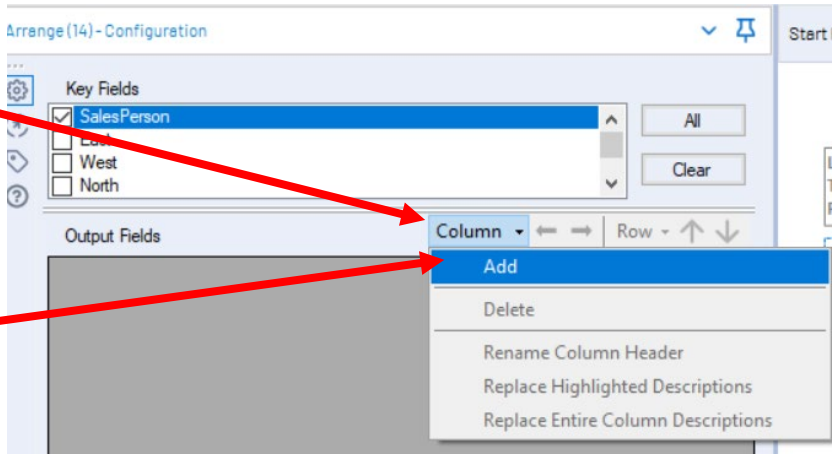


With the "Arrange Tool" selected, in the Configuration Area, under "Key Fields", check SalesPerson.



Select column.

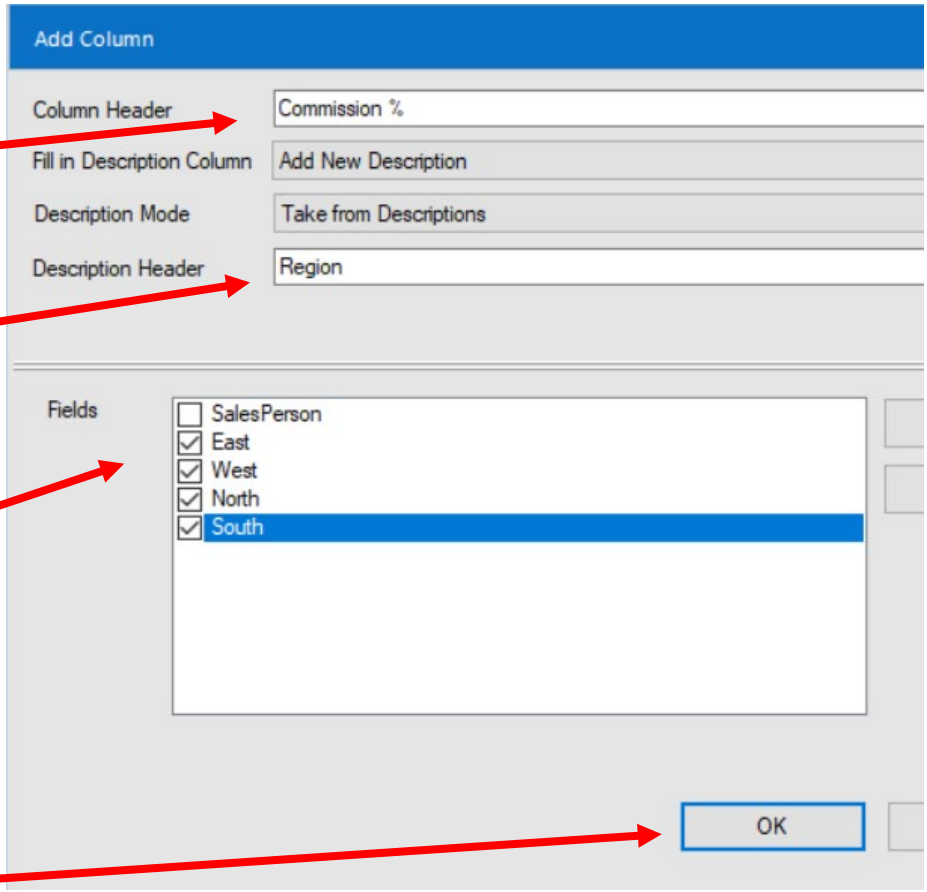
Select Add.



In the "Column Header" type: Commission %.

In the "Description Header" type: Region.

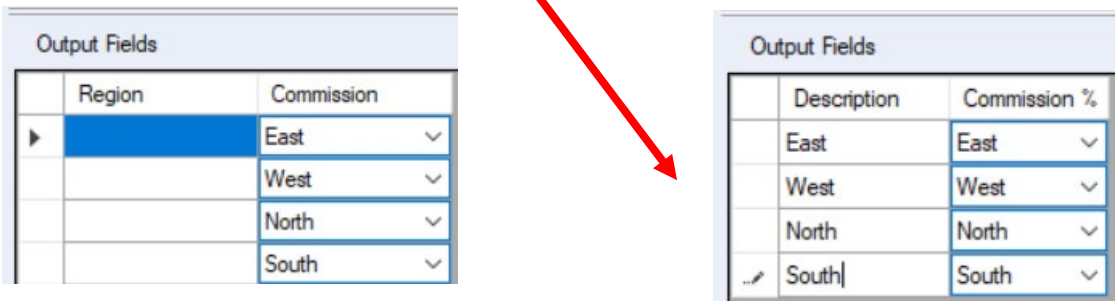
In the "Fields" check: East, West, North and South.



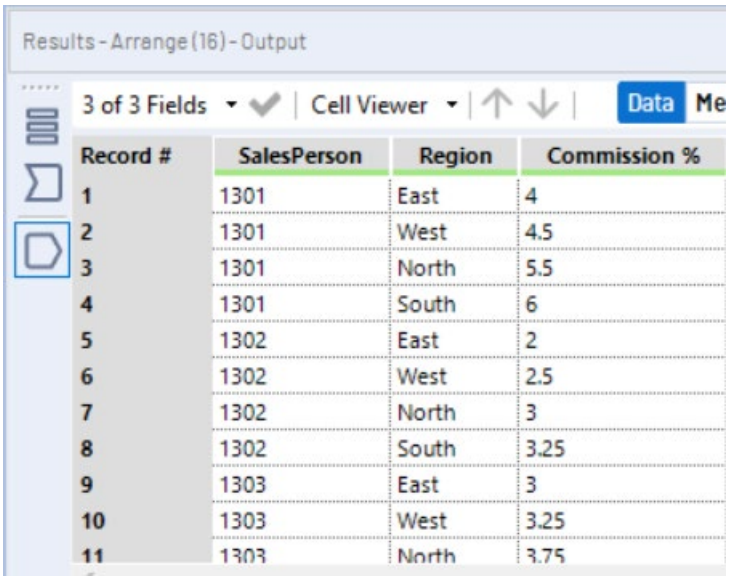
Select "OK".

Select Run to apply.

In the "Output Fields Description" type, East, West, North South



Select Run to apply.



Record #	SalesPerson	Region	Commission %
1	1301	East	4
2	1301	West	4.5
3	1301	North	5.5
4	1301	South	6
5	1302	East	2
6	1302	West	2.5
7	1302	North	3
8	1302	South	3.25
9	1303	East	3
10	1303	West	3.25
11	1303	North	3.75

Compare to original data.

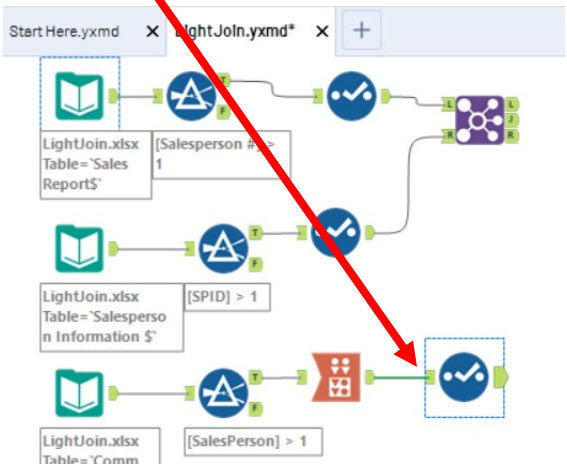
Record #	SalesPerson	East	West	North	South
1	1301	4	4.5	5.5	6
2	1302	2	2.5	3	3.25
3	1303	3	3.25	3.75	4
4	1304	3	3.25	4.25	4.75

To combine two streams in a one to many relationship there must be one field that is unique for each row of the data. That unique field is referred to as the primary key. The primary field will be created by combining the "Salesperson" field with a "." and the "Region" field. The primary key for the first record will be "1301.East".

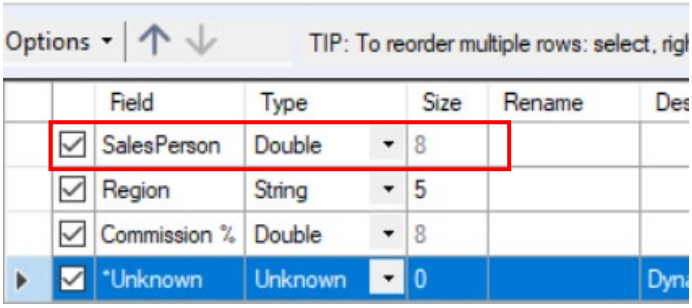
Tool 11 Select Tool: Change The Field Type

The fields “SalesPerson” and “Region” must be combined. “Region” is a string, but “SalesPerson” could be a string or a numeric.

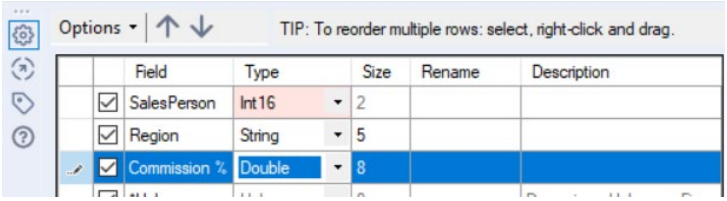
Drag the “**Select Tool**” from the Tools Palette to the output of the “**Arrange Tool**” on the third stream. The two will be connected automatically.



With the “**Select Tool**” selected, in the Configuration area note that “Salesperson #” is a numeric field with the type “Double”.

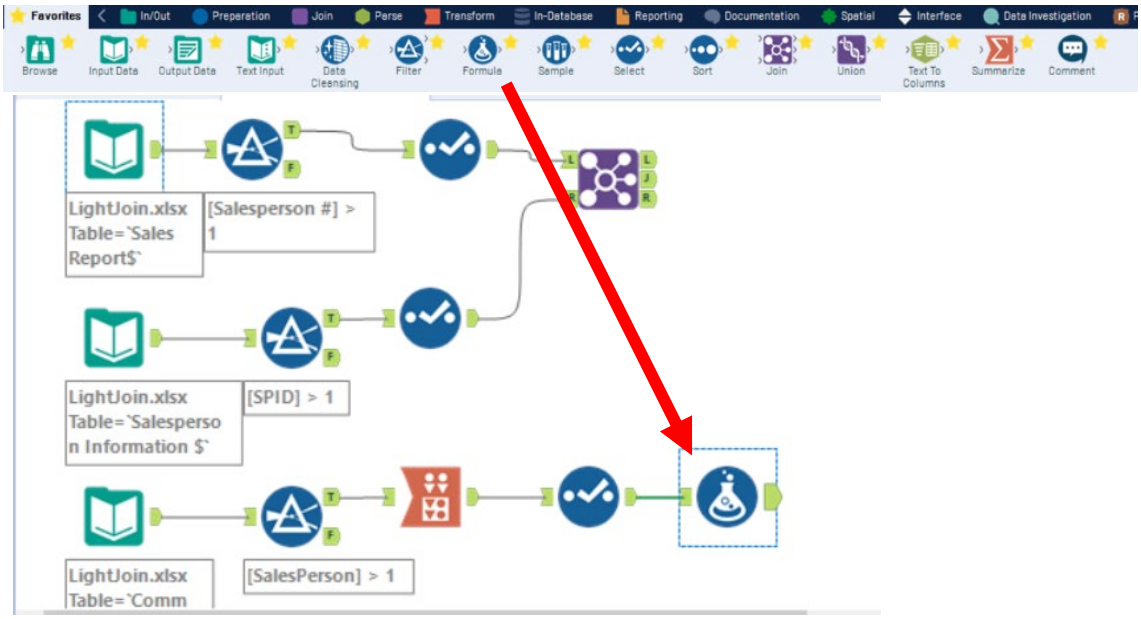


Change the type to Integer, “Int 16” to match the field definition that was used in stream 1 and stream 2.

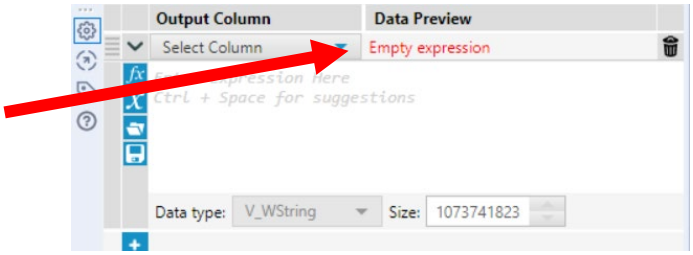


Tool 12: Formula Tool - Create A Primary Key Field

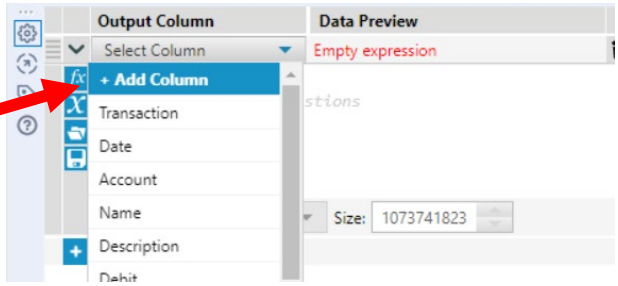
Drag a "**Formula Tool**"  from the Tools Palette to the output of the "**Select Tool**" on the third stream. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



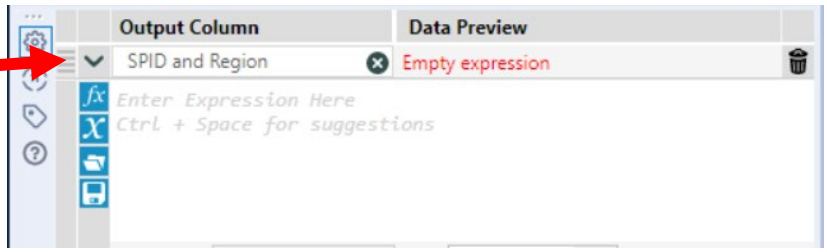
In the output Configuration area, click on the drop-down arrow of "Select Column".



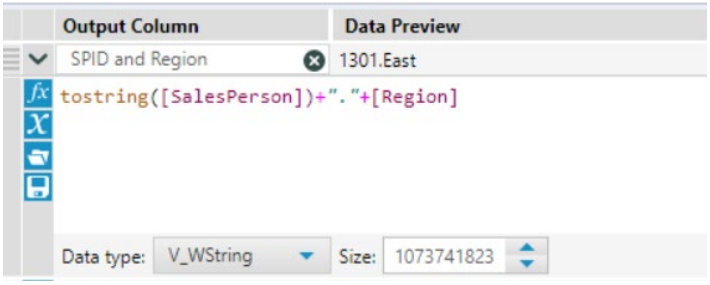
Select "Add Column".



Type “SPID and Region”.



Type the following formula
tostring([Salesperson])+"."+[Region]



Select “Run”.

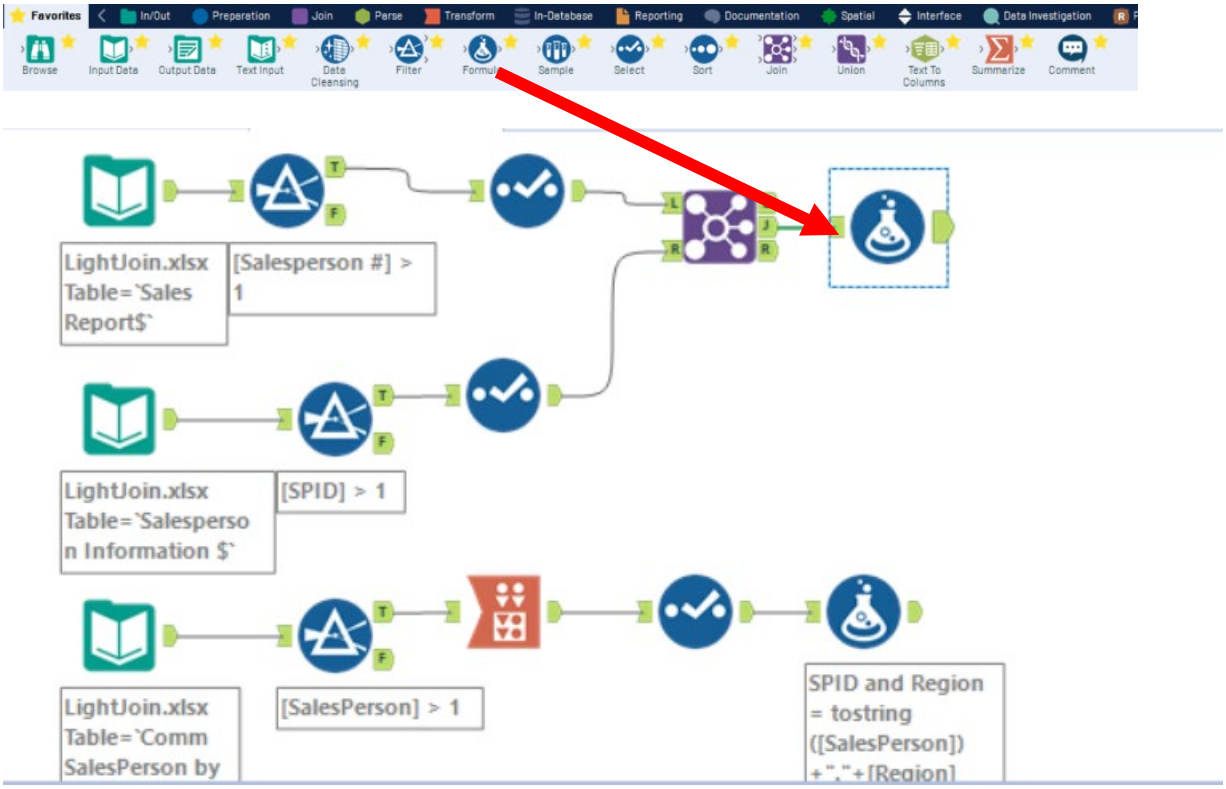
4 of 4 Fields | Cell Viewer | 16 records displayed | Data | Me

Record #	SalesPerson	Region	Commission %	SPID and Region
1	1301	East	4	1301.East
2	1301	West	4.5	1301.West
3	1301	North	5.5	1301.North
4	1301	South	6	1301.South
5	1302	East	2	1302.East
6	1302	West	2.5	1302.West
7	1302	North	3	1302.North
8	1302	South	3.25	1302.South
9	1303	East	3	1303.East
10	1303	West	3.25	1303.West
11	1303	North	3.75	1303.North

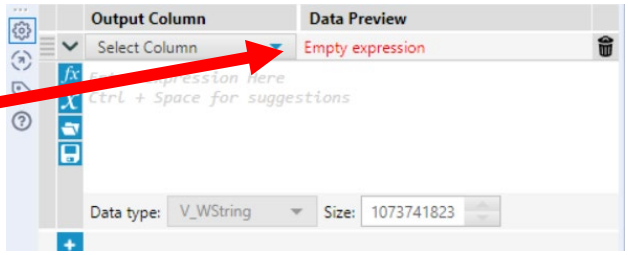
Now add the “SPID and Region” field to the combined first and second stream of data. The field is referred to as a foreign key. It will permit us to link to the data in the third stream.

Tool 13: Formula Tool - Create A Foreign Key Field

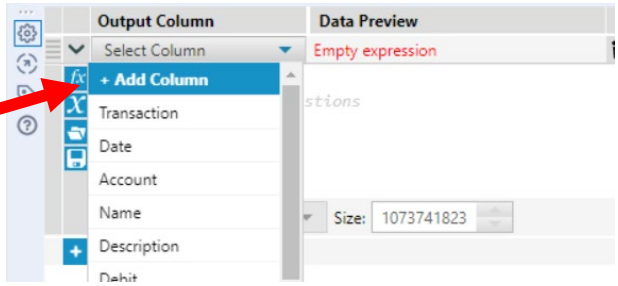
Drag a “**Formula Tool**”  from the Tools Palette to the J output of the “**Join Tool**”. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



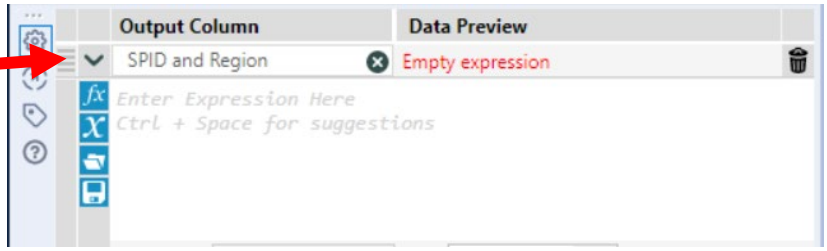
In the output Configuration area, click on the drop-down arrow of “Select Column”.



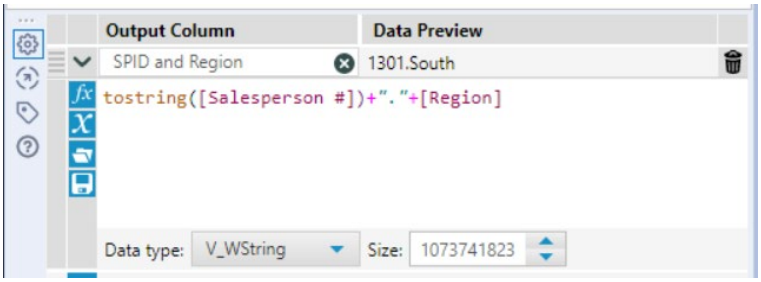
Select “Add Column”.



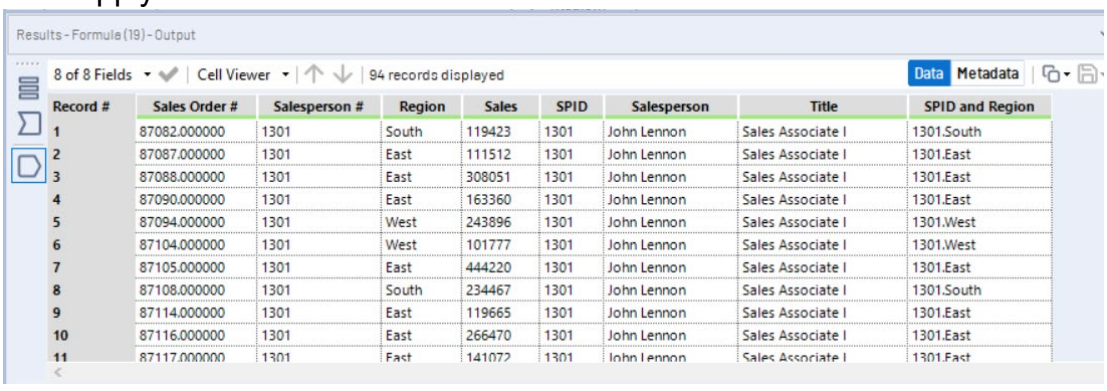
Type “SPID and Region”.



Type the following formula
tostring([Salesperson])+" "+[Region]



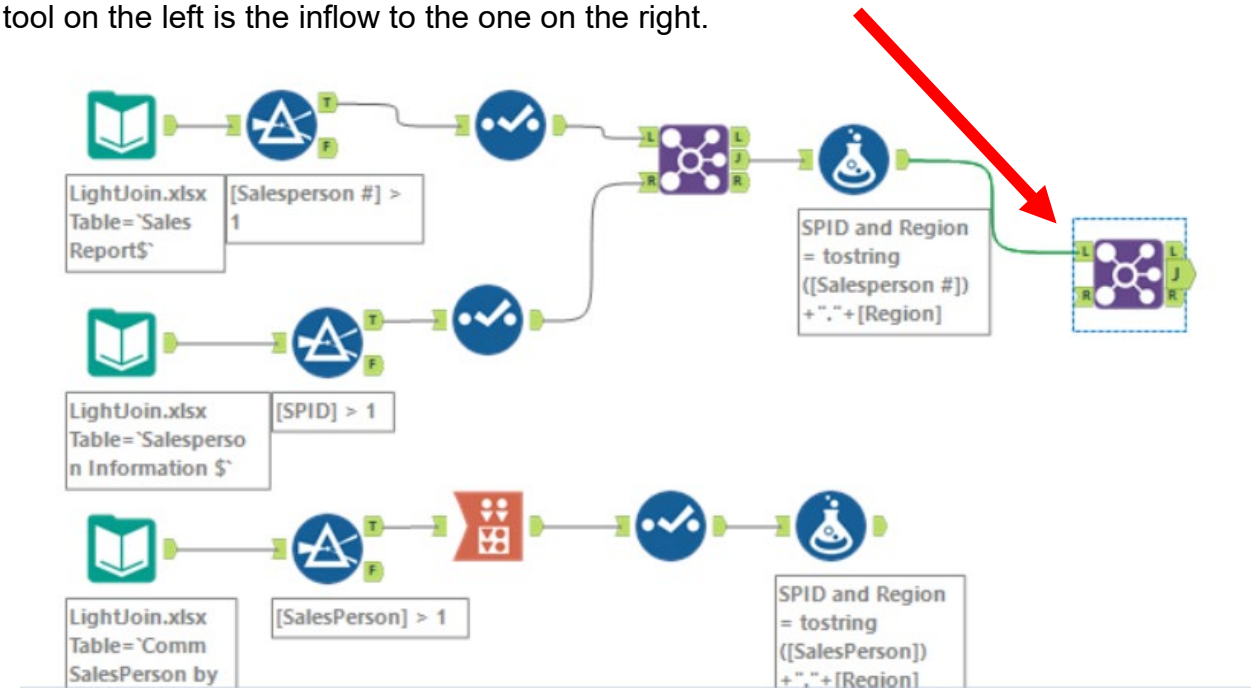
Select “Run” to apply.



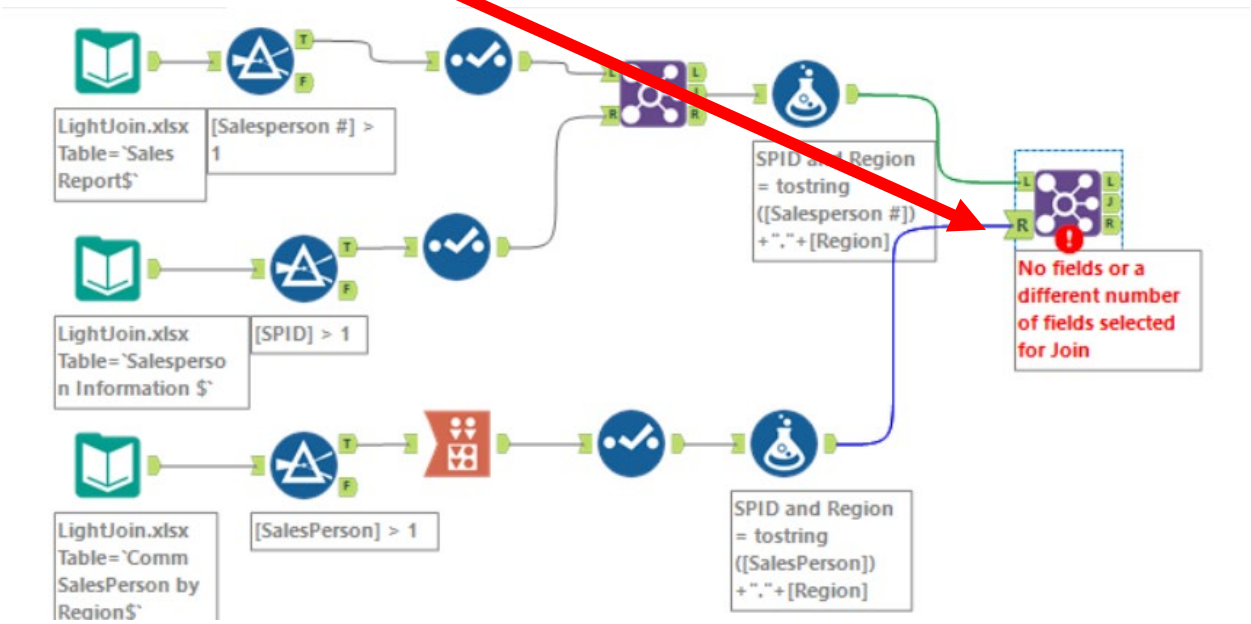
Record #	Sales Order #	Salesperson #	Region	Sales	SPID	Salesperson	Title	SPID and Region
1	87082.000000	1301	South	119423	1301	John Lennon	Sales Associate I	1301.South
2	87087.000000	1301	East	111512	1301	John Lennon	Sales Associate I	1301.East
3	87088.000000	1301	East	308051	1301	John Lennon	Sales Associate I	1301.East
4	87090.000000	1301	East	163360	1301	John Lennon	Sales Associate I	1301.East
5	87094.000000	1301	West	243896	1301	John Lennon	Sales Associate I	1301.West
6	87104.000000	1301	West	101777	1301	John Lennon	Sales Associate I	1301.West
7	87105.000000	1301	East	444220	1301	John Lennon	Sales Associate I	1301.East
8	87108.000000	1301	South	234467	1301	John Lennon	Sales Associate I	1301.South
9	87114.000000	1301	East	119665	1301	John Lennon	Sales Associate I	1301.East
10	87116.000000	1301	East	266470	1301	John Lennon	Sales Associate I	1301.East
11	87117.000000	1301	East	141072	1301	John Lennon	Sales Associate I	1301.East

Tool 14: Join Tool - Combine First And Second Stream Of Data To The Third Stream Of Data

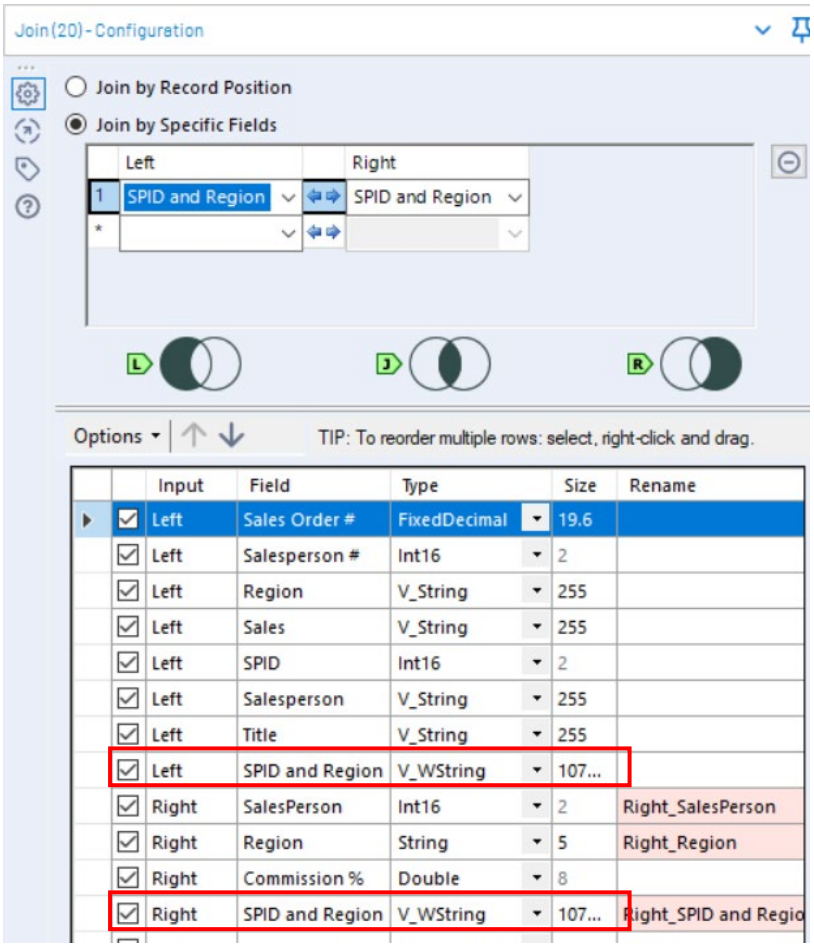
Drag the "**Join Tool**" from the Tools Palette to the right of the "Formula Tool" from the first two data streams. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



Connect the output of the "Formula Tool" from the third data stream to the R (Right) inflow of the "Join Tool".

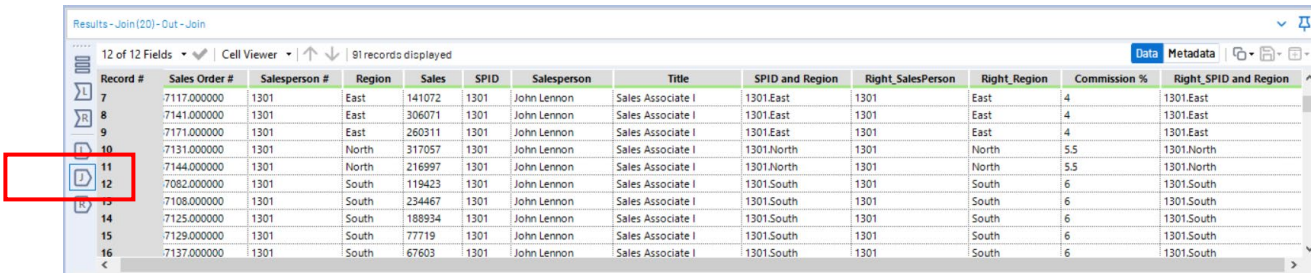


With the "**Join Tool**" selected, in the Configuration area, from the drop-down choose "SPID and Region" on the left and "SPID and Region" on the right. The two fields must be the same type, in this case V_WString, and the same size, in this case 107.



Select Run to apply.

In the result area select J, Join and review the output.

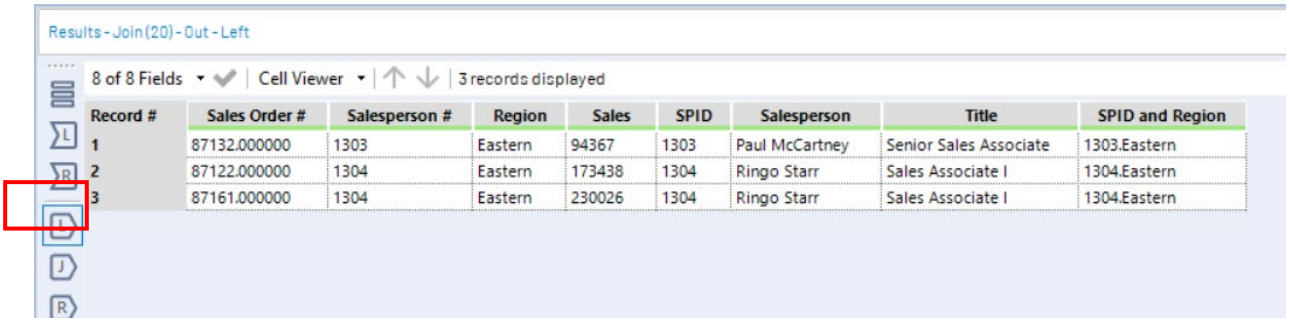


Verify to the original data stream.

Record #	SalesPerson	East	West	North	South
1	1301	4	4.5	5.5	6
2	1302	2	2.5	3	3.25
3	1303	3	3.25	3.75	4
4	1304	3	3.25	4.25	4.75

Commission %	Right_SPID and Region
4	1301.East
4	1301.East
4	1301.East
5.5	1301.North
5.5	1301.North
6	1301.South
6	1301.South

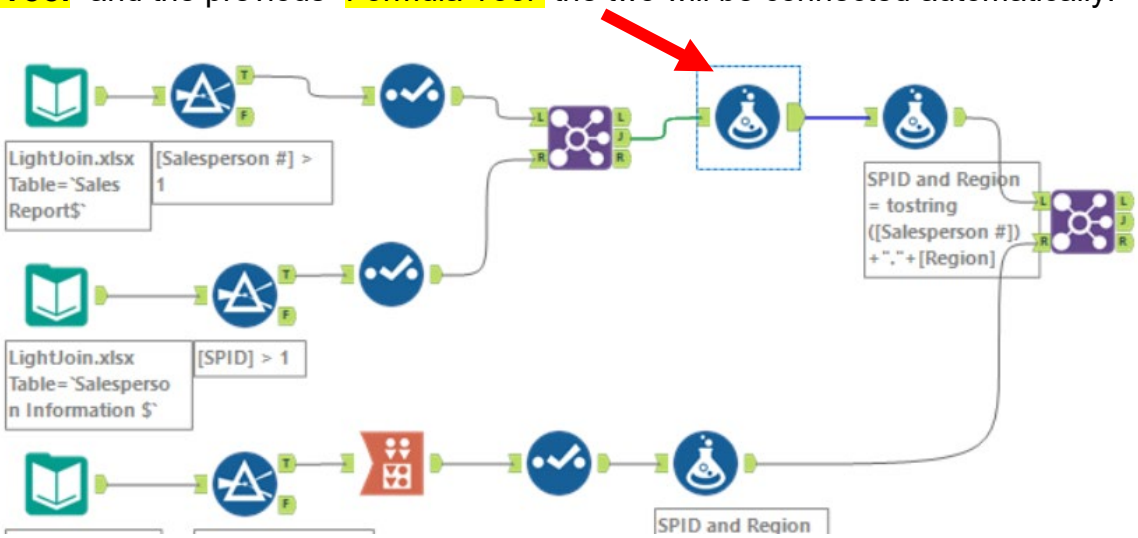
Verify that nothing was unmatched from L, Left data set. There are three records that were not matched. Either the “Salesperson #” or the “Region” is not correct. The In Sales Order # 87132, region was entered as “Eastern” when it should have been “East”.



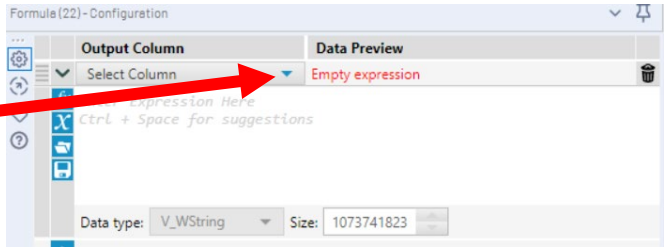
Record #	Sales Order #	Salesperson #	Region	Sales	SPID	Salesperson	Title	SPID and Region
1	87132.000000	1303	Eastern	94367	1303	Paul McCartney	Senior Sales Associate	1303.Eastern
2	87122.000000	1304	Eastern	173438	1304	Ringo Starr	Sales Associate I	1304.Eastern
3	87161.000000	1304	Eastern	230026	1304	Ringo Starr	Sales Associate I	1304.Eastern

Tool 15: “Formula Tool” - Correct Data Errors

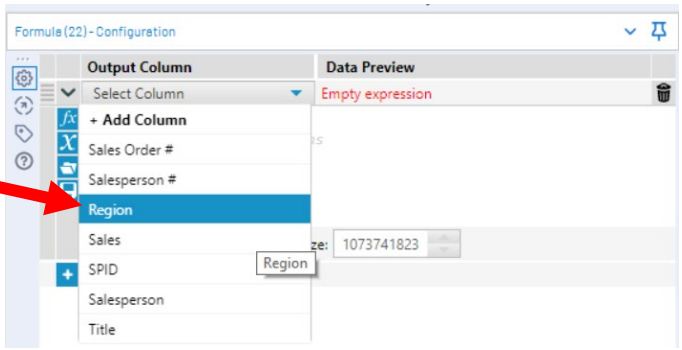
Drag a “**Formula Tool**”  from the Tools Palette between the output of the first “**Join Tool**” and the previous “Formula Tool” the two will be connected automatically.



With the "**Formula Tool**" selected, in the Configuration area, from the drop-down, "Select Column".

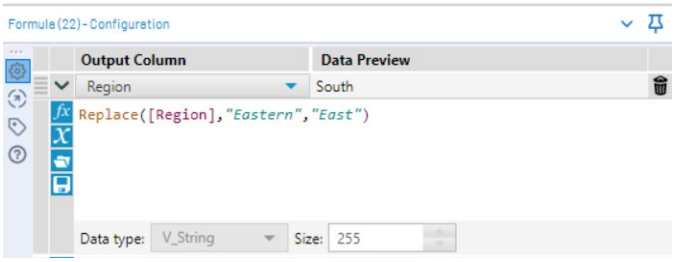


Select "Region".

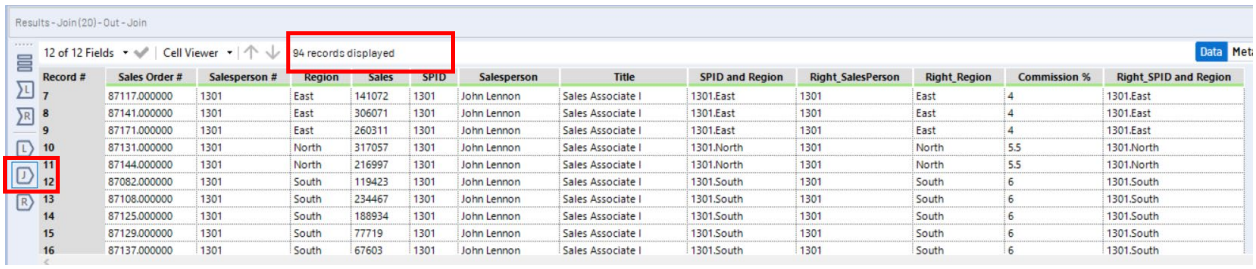


Type the following formula:

Replace([Region],"Eastern","East")



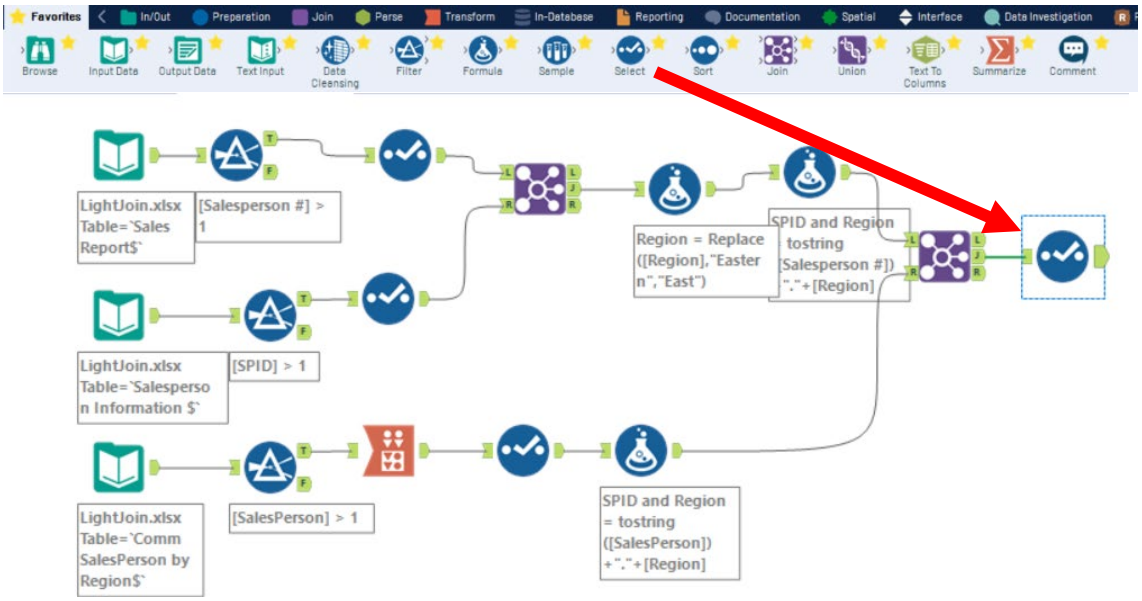
Select "Run" to apply and select the J output from the last Join.



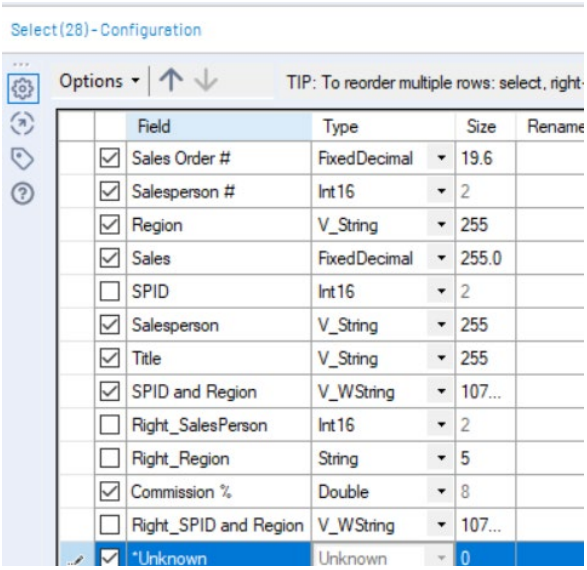
Record #	Sales Order #	Salesperson #	Region	Sales	SPID	Salesperson	Title	SPID and Region	Right_SalesPerson	Right_Region	Commission %	Right_SPID and Region
7	87117.000000	1301	East	141072	1301	John Lennon	Sales Associate I	1301.East	1301	East	4	1301.East
8	87141.000000	1301	East	306071	1301	John Lennon	Sales Associate I	1301.East	1301	East	4	1301.East
9	87171.000000	1301	East	260311	1301	John Lennon	Sales Associate I	1301.East	1301	East	4	1301.East
10	87131.000000	1301	North	317057	1301	John Lennon	Sales Associate I	1301.North	1301	North	5.5	1301.North
11	87144.000000	1301	North	216997	1301	John Lennon	Sales Associate I	1301.North	1301	North	5.5	1301.North
12	87082.000000	1301	South	119423	1301	John Lennon	Sales Associate I	1301.South	1301	South	6	1301.South
13	87108.000000	1301	South	234467	1301	John Lennon	Sales Associate I	1301.South	1301	South	6	1301.South
14	87125.000000	1301	South	188934	1301	John Lennon	Sales Associate I	1301.South	1301	South	6	1301.South
15	87129.000000	1301	South	77719	1301	John Lennon	Sales Associate I	1301.South	1301	South	6	1301.South
16	87137.000000	1301	South	67603	1301	John Lennon	Sales Associate I	1301.South	1301	South	6	1301.South

Tool 16: Select Tool – Remove Fields That Contain Duplicate Data

Drag the "**Select Tool**" from the Tools Palette to the Joined output of the "**Join Tool**" on the Canvas. The two will be connected automatically.



Remove the extra fields by unchecking them.



Select “Run” to apply.

Results - Select (28) - Output

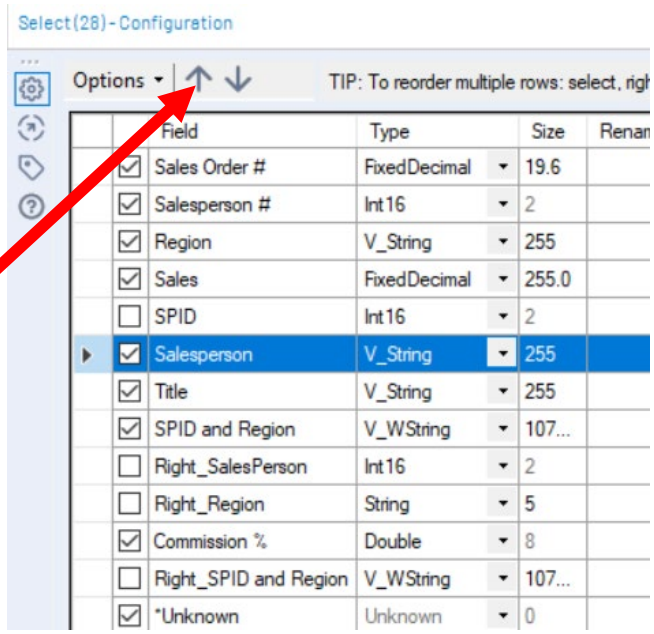
8 of 8 Fields | Cell Viewer | 94 records displayed

Record #	Sales Order #	Salesperson #	Region	Sales	Salesperson	Title	SPID and Region	Commission %
1	87087.000000	1301	East	111512	John Lennon	Sales Associate I	1301.East	4
2	87088.000000	1301	East	308051	John Lennon	Sales Associate I	1301.East	4
3	87090.000000	1301	East	163360	John Lennon	Sales Associate I	1301.East	4
4	87105.000000	1301	East	444220	John Lennon	Sales Associate I	1301.East	4
5	87114.000000	1301	East	119665	John Lennon	Sales Associate I	1301.East	4
6	87116.000000	1301	East	266470	John Lennon	Sales Associate I	1301.East	4
7	87117.000000	1301	East	141072	John Lennon	Sales Associate I	1301.East	4
8	87141.000000	1301	East	306071	John Lennon	Sales Associate I	1301.East	4
9	87171.000000	1301	East	260311	John Lennon	Sales Associate I	1301.East	4
10	87131.000000	1301	North	317057	John Lennon	Sales Associate I	1301.North	5.5

Select the field “Salesperson”.

Move it to before the field “Region” by selecting the up arrow three times.

Move the other fields as desired.



Select “Run” to apply.

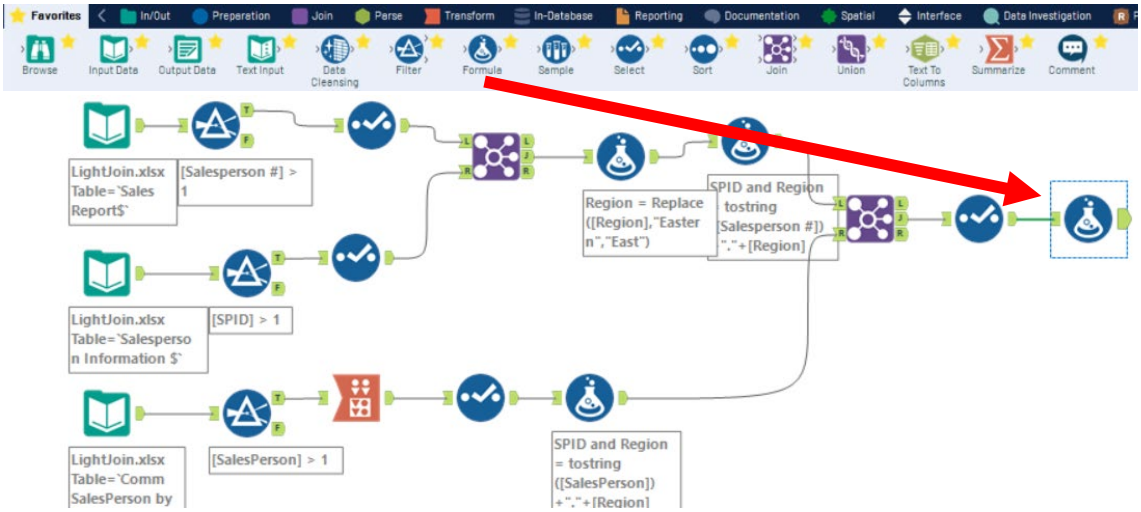
Results - Select (28) - Output

8 of 8 Fields | Cell Viewer | 94 records displayed

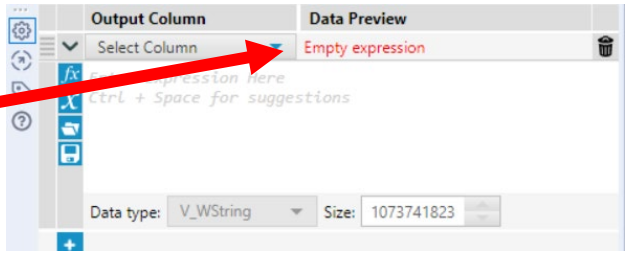
Record #	Sales Order #	Salesperson #	Salesperson	Title	Region	SPID and Region	Sales	Commission %
1	87087.000000	1301	John Lennon	Sales Associate I	East	1301.East	111512	4
2	87088.000000	1301	John Lennon	Sales Associate I	East	1301.East	308051	4
3	87090.000000	1301	John Lennon	Sales Associate I	East	1301.East	163360	4
4	87105.000000	1301	John Lennon	Sales Associate I	East	1301.East	444220	4
5	87114.000000	1301	John Lennon	Sales Associate I	East	1301.East	119665	4
6	87116.000000	1301	John Lennon	Sales Associate I	East	1301.East	266470	4
7	87117.000000	1301	John Lennon	Sales Associate I	East	1301.East	141072	4
8	87141.000000	1301	John Lennon	Sales Associate I	East	1301.East	306071	4
9	87171.000000	1301	John Lennon	Sales Associate I	East	1301.East	260311	4
10	87131.000000	1301	John Lennon	Sales Associate I	North	1301.North	317057	5.5

Tool 17: Formula Tool: Calculate The Commission On Every Sales Order

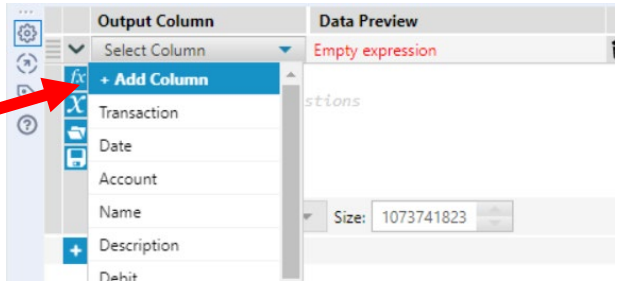
Drag a "**Formula Tool**" from the Tools Palette to the output of the "**Select Tool**" on the third stream. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



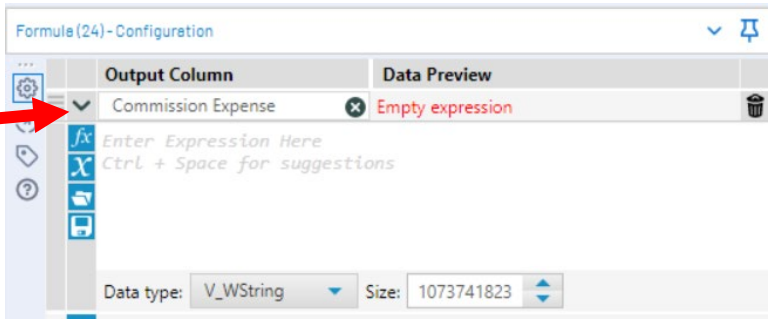
In the output Configuration area, click on the drop-down arrow of "Select Column".



Select "Add Column".

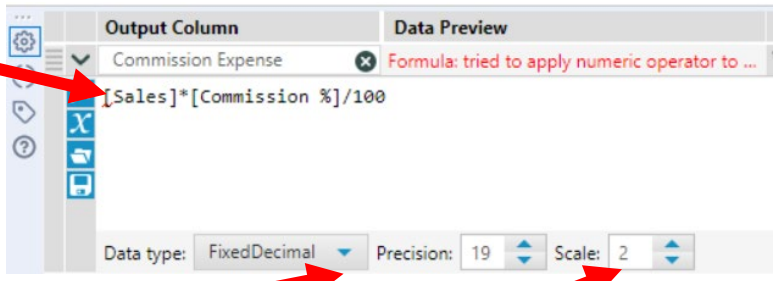


Type “Commission Expense”



Type the following formula:

[Sales]*[Commission %]/100

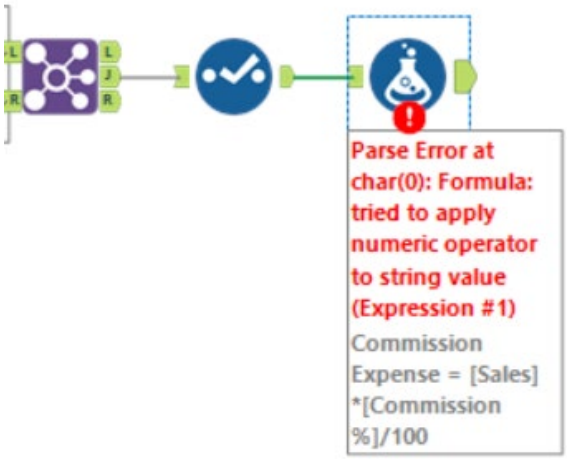


Change the “Data type:” to FixedDecimal.

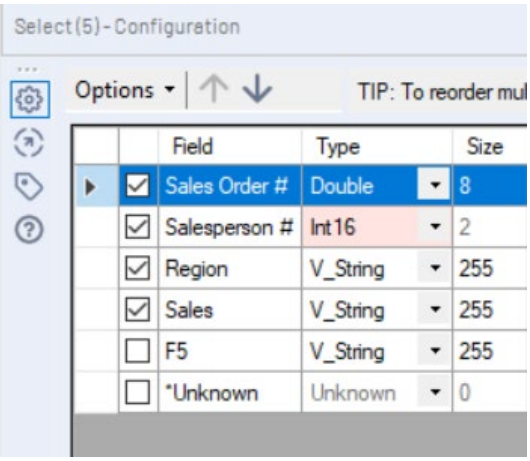
Change the Scale (number of decimal places) to two.

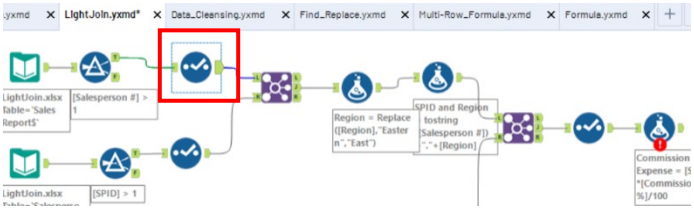
Select “Run” to apply.

There is an error. Somewhere in the first stream, the “Sales” field is defined as a string.

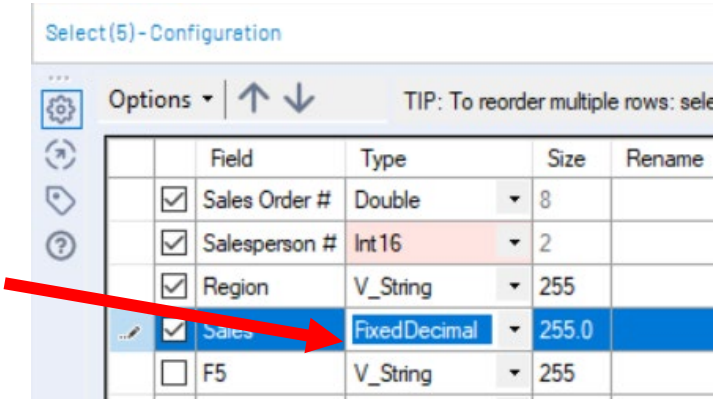


Click on the first “Select Tool”

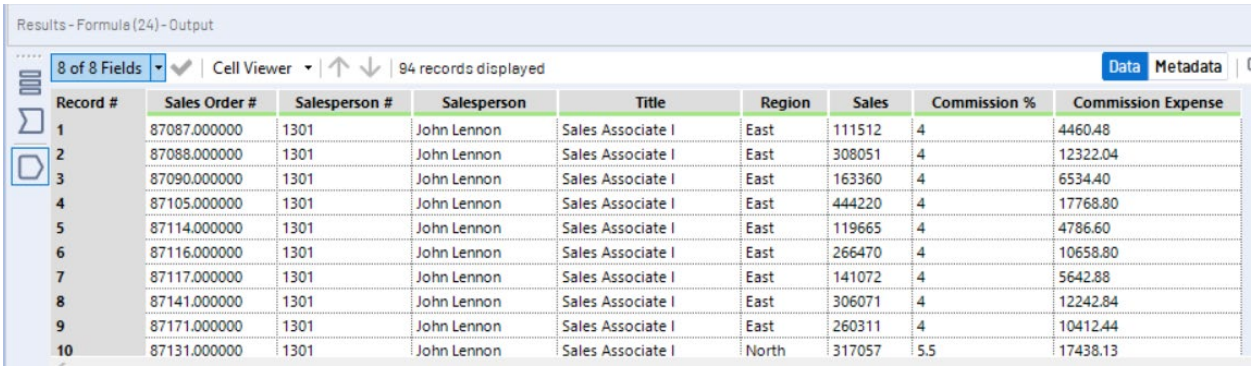




Change the “Sales” type from V_String to a FixedDecimal.



Select “Run” to apply. Select the outflow from the Commission Expense “Formula Tool” if necessary.

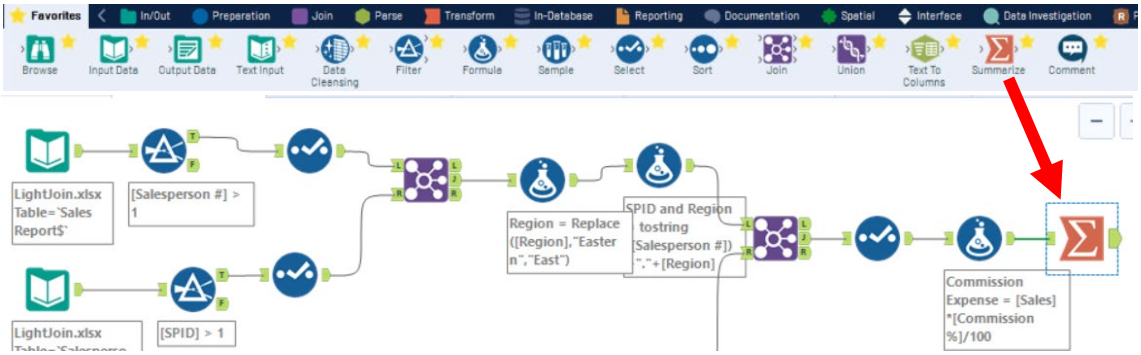


Record #	Sales Order #	Salesperson #	Salesperson	Title	Region	Sales	Commission %	Commission Expense
1	87087.000000	1301	John Lennon	Sales Associate I	East	111512	4	4460.48
2	87088.000000	1301	John Lennon	Sales Associate I	East	308051	4	12322.04
3	87090.000000	1301	John Lennon	Sales Associate I	East	163360	4	6534.40
4	87105.000000	1301	John Lennon	Sales Associate I	East	444220	4	17768.80
5	87114.000000	1301	John Lennon	Sales Associate I	East	119665	4	4786.60
6	87116.000000	1301	John Lennon	Sales Associate I	East	266470	4	10658.80
7	87117.000000	1301	John Lennon	Sales Associate I	East	141072	4	5642.88
8	87141.000000	1301	John Lennon	Sales Associate I	East	306071	4	12242.84
9	87171.000000	1301	John Lennon	Sales Associate I	East	260311	4	10412.44
10	87131.000000	1301	John Lennon	Sales Associate I	North	317057	5.5	17438.13

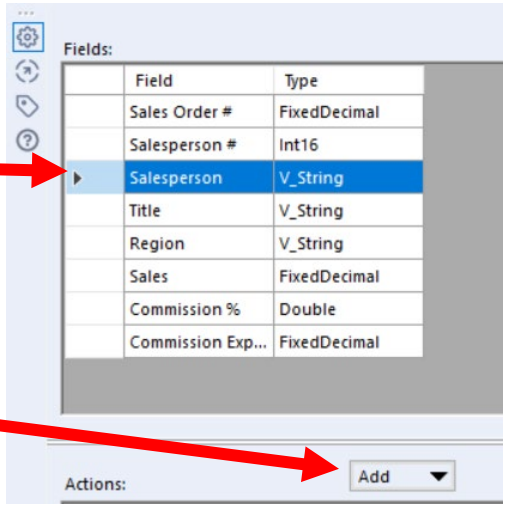
Check the math, $111,512 x 4% = $4,460.48.

Tool 18: Summarize Tool: Create A "Commission Expense Report" By Salesperson.

Drag a "**Summarize Tool**"  to the right of the "**Formula tool**" on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.

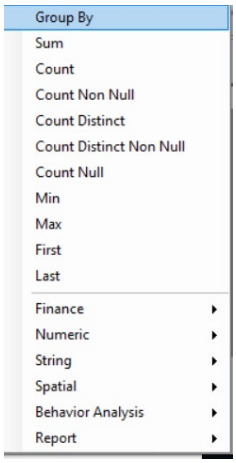


In the output Configuration area, select "Salesperson".



Click on the drop-down arrow on "Actions" and select "Add".

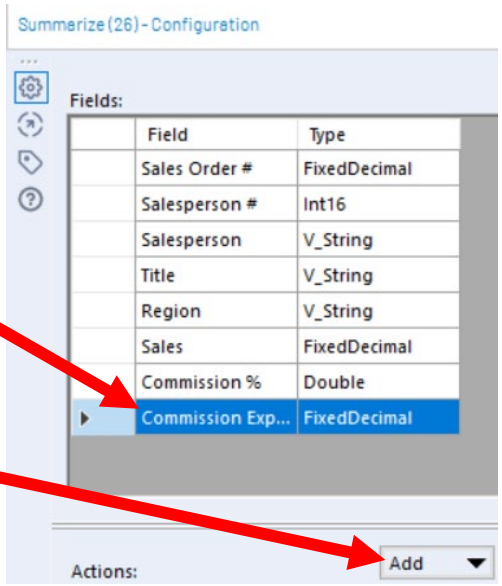
Select "Group By".

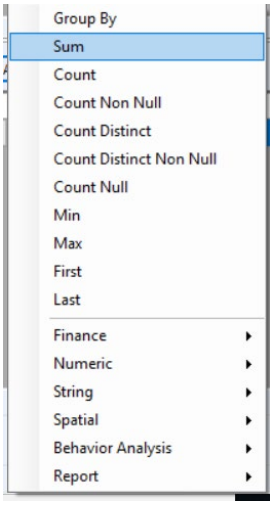


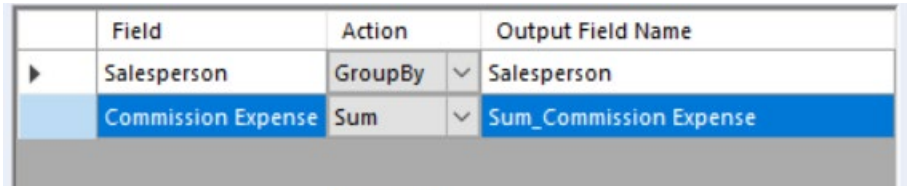
Select “Commission Expense”.

Click on the drop-down arrow on “Actions” and select “Add”.

Select “Sum”.

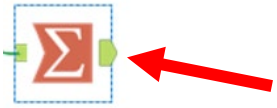






Select “Run”.

Select the output of the “Summarize Tool”.

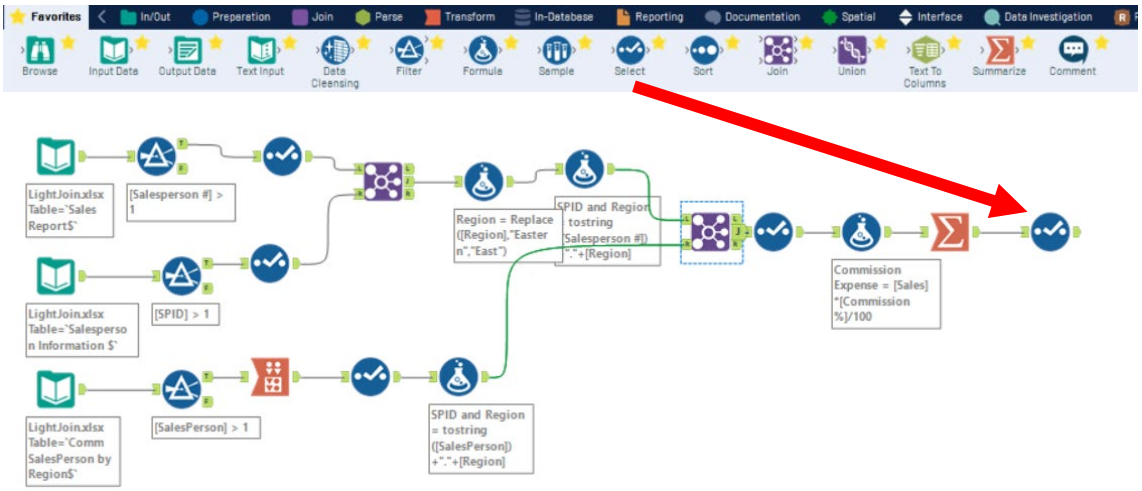


The “Commission Expense Report” is created.

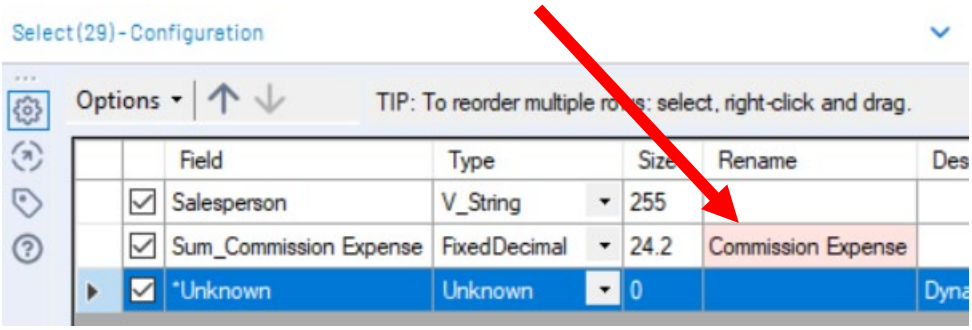
Record #	Salesperson	Sum_Commission Expense
1	George Harrison	90935.34
2	John Lennon	226149.84
3	Paul McCartney	127193.52
4	Ringo Starr	274189.61

Tool 19: Select Tool – Rename A Field

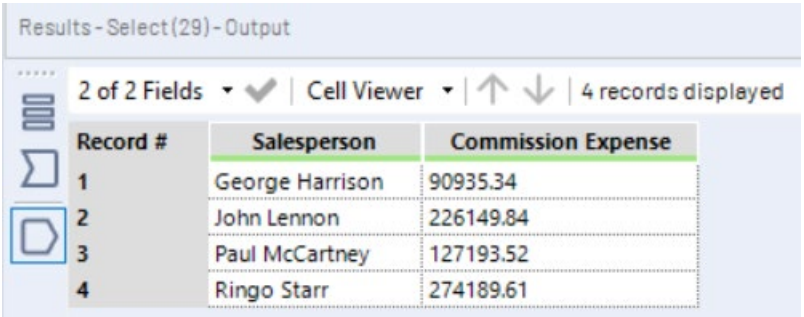
Drag the "**Select Tool**" from the Tools Palette to the output of the "**Summarize Tool**" on the Canvas. The two will be connected automatically.



Rename Sum_Commission Expense to Commission Expense.

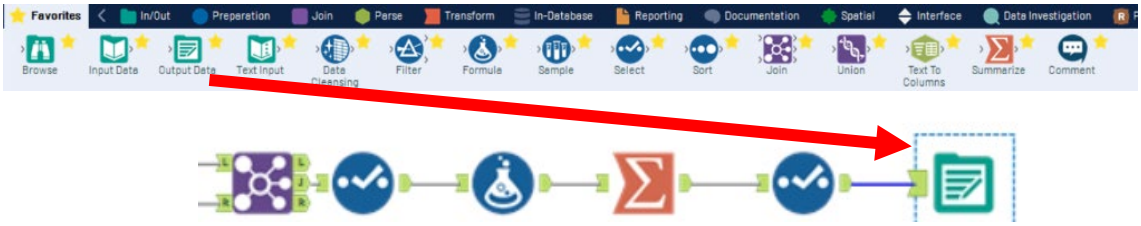


Select "Run" to apply.



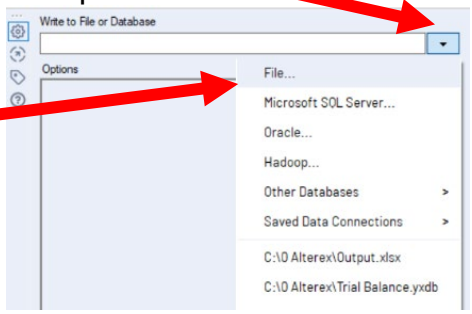
Tool 20: Output Data Tool: Export Report to Excel

Drag an "**Output Data Tool**" to the right of the "**Select Tool**" on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



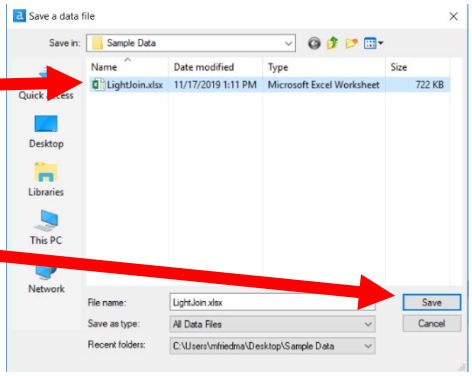
In the output Configuration area, select the drop-down arrow.

Select "File".

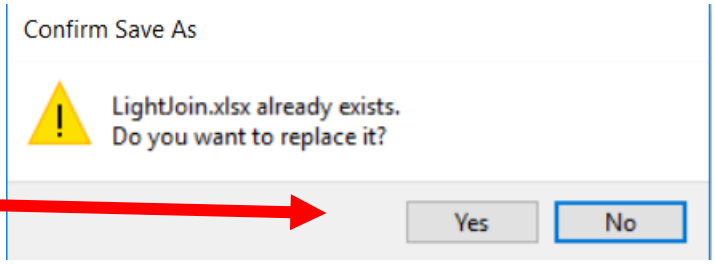


Select the required file name.

Select "Save".

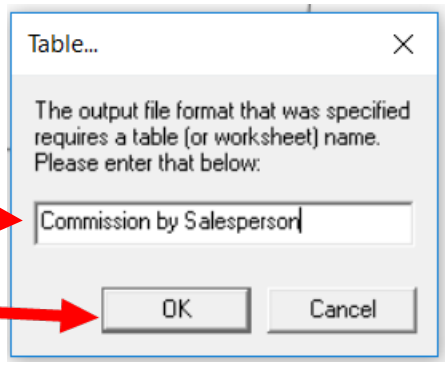


Click on “Yes” to confirm “Save”.

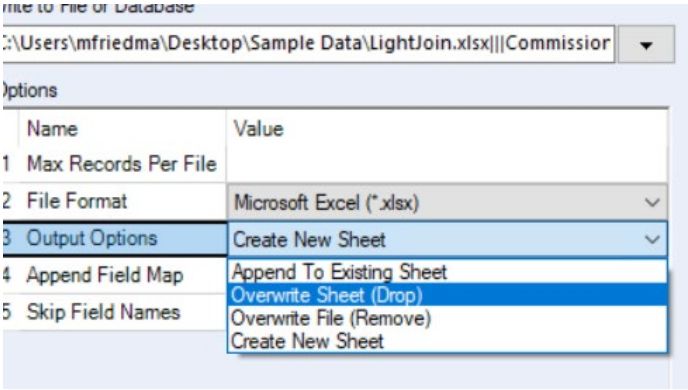


Enter the “Sheet Name”.

Select “OK”.



Change “3 Output Options” to Overwrite Sheet (Drop) to avoid producing an error the second time you save the sheet to the file.



Select “Run” to apply.

Open the file in Excel.

	A	B	C	D
1	Salespersc	Commission Expense		
2	George Ha	90935.34		
3	John Lenn	########		
4	Paul McCa	########		
5	Ringo Star	########		
6				

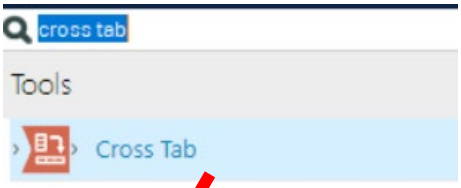
Change the columns' widths and formats.

	A	B
1	Salesperson	Commission Expense
2	George Harrison	$90,935.34
3	John Lennon	$226,149.84
4	Paul McCartney	$127,193.52
5	Ringo Starr	$274,189.61
6		

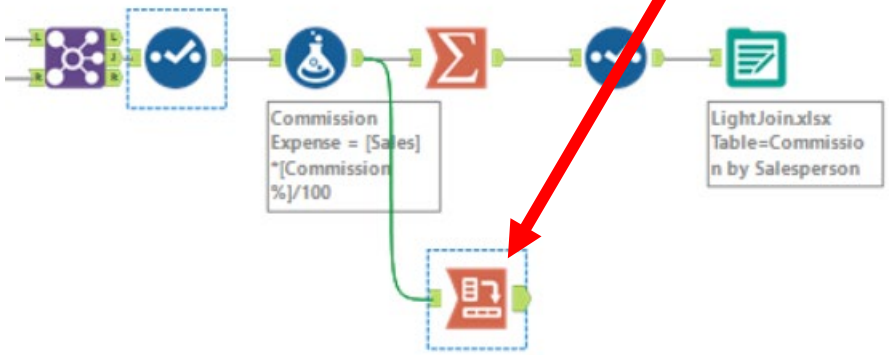
"Save" and "Close" the Excel file.
Return to Alteryx.

Tool 21: Cross Tab Tool: Change Vertical Data to Multiple Horizontal Columns

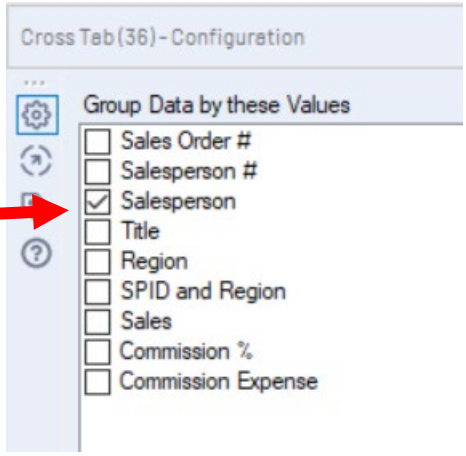
In the “Search for Tools” box type, “Cross Tab”.



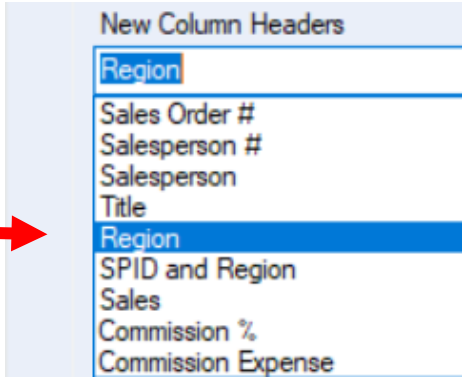
Drag the “**Cross Tab Tool**” from the Tools Palette to the right of the “Formula Tool” and below the **“Summarize Tool”** on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



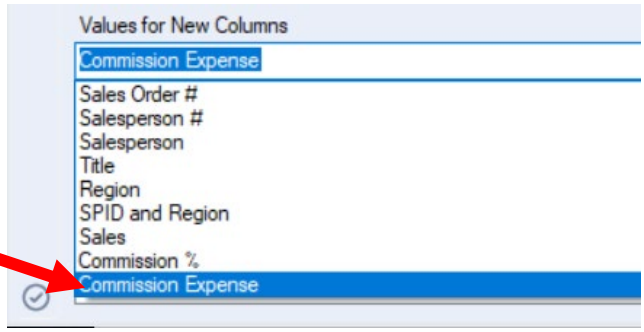
The leftmost column of the cross tab will be the “Salesperson”. In the output Configuration area, “Group Data by these Values”, select “Salesperson”.



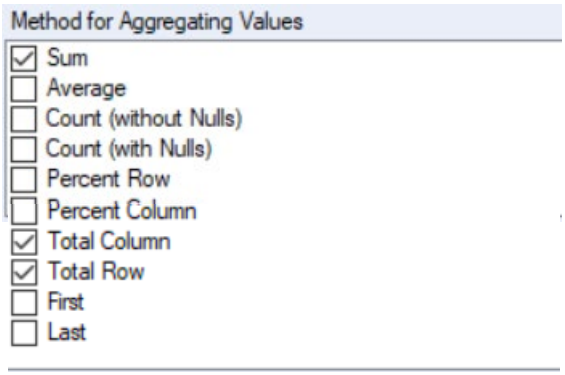
The column headings of the cross tab will be the “Region”. In the output Configuration area, “New Column Headers” select “Region”.



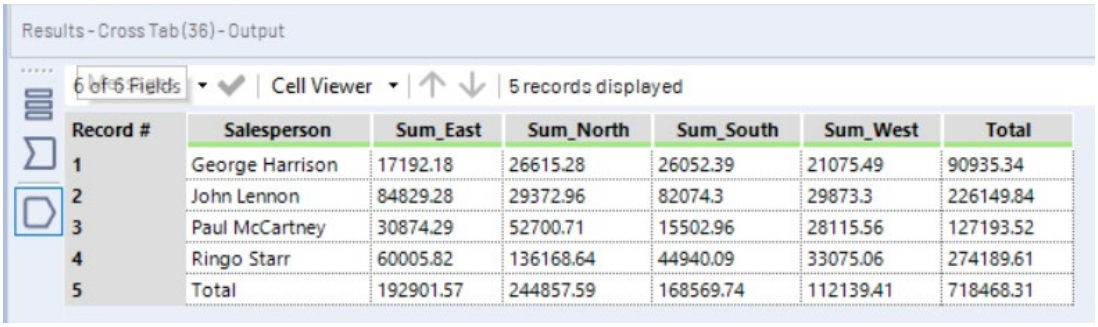
In the output Configuration area, “Values for New Column Headers” select “Commission Expense”.



In the output Configuration area, “Method for Aggregating Values” select: “Sum”, “Total Column” and “Total Row”.



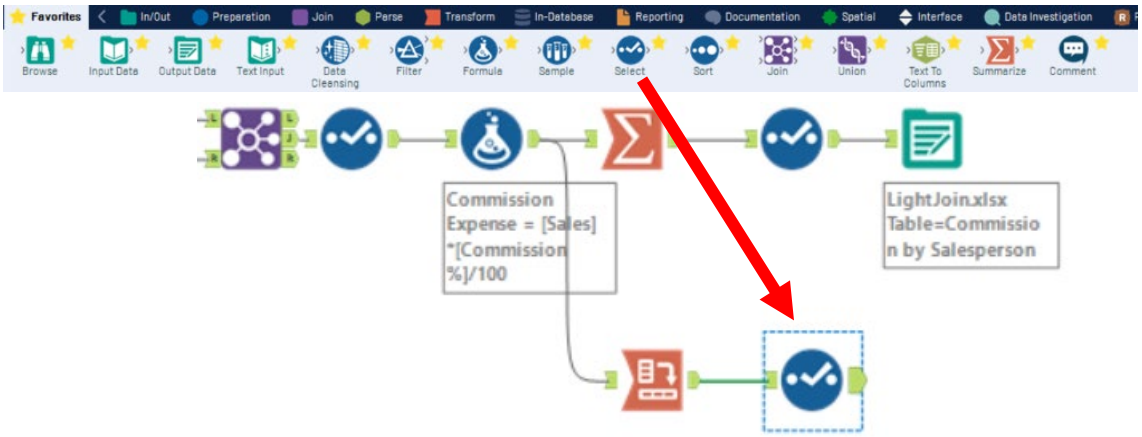
Select “Run” to apply.



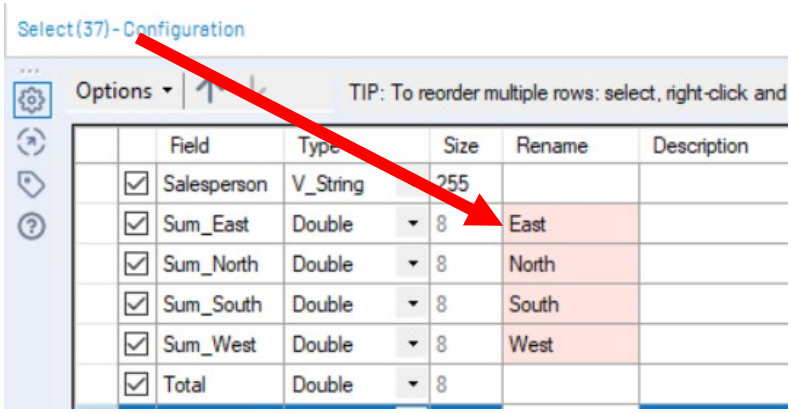
Record #	Salesperson	Sum_East	Sum_North	Sum_South	Sum_West	Total
1	George Harrison	17192.18	26615.28	26052.39	21075.49	90935.34
2	John Lennon	84829.28	29372.96	82074.3	29873.3	226149.84
3	Paul McCartney	30874.29	52700.71	15502.96	28115.56	127193.52
4	Ringo Starr	60005.82	136168.64	44940.09	33075.06	274189.61
5	Total	192901.57	244857.59	168569.74	112139.41	718468.31

Tool 22: Select Tool – Rename Fields

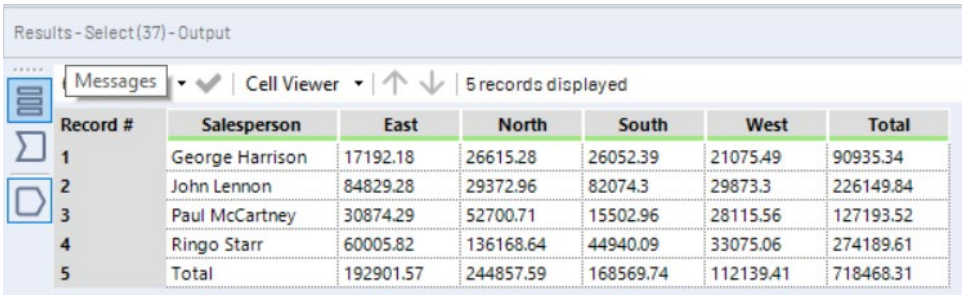
Drag the "**Select Tool**" from the Tools Palette to the output of the "**Cross Tab Tool**" on the Canvas. The two will be connected automatically.



Rename the column headers.



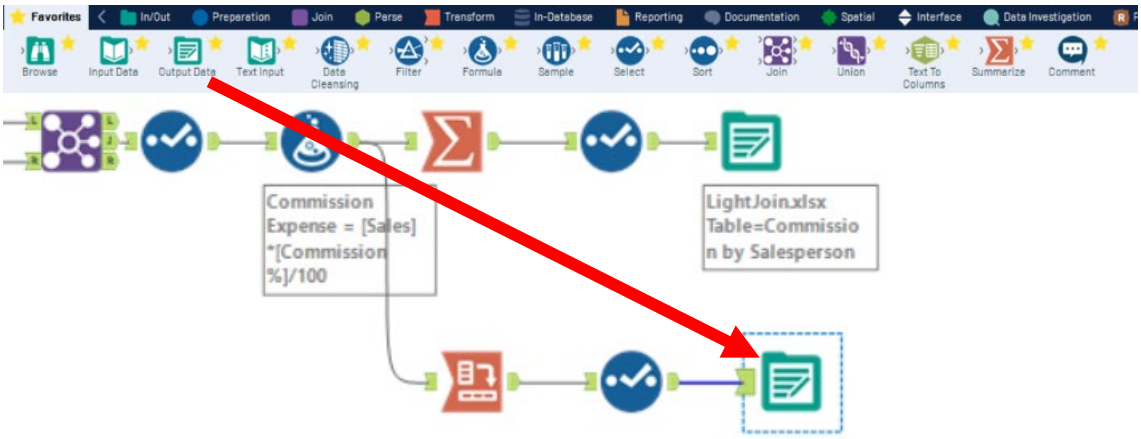
Select "Run" to apply.



Record #	Salesperson	East	North	South	West	Total
1	George Harrison	17192.18	26615.28	26052.39	21075.49	90935.34
2	John Lennon	84829.28	29372.96	82074.3	29873.3	226149.84
3	Paul McCartney	30874.29	52700.71	15502.96	28115.56	127193.52
4	Ringo Starr	60005.82	136168.64	44940.09	33075.06	274189.61
5	Total	192901.57	244857.59	168569.74	112139.41	718468.31

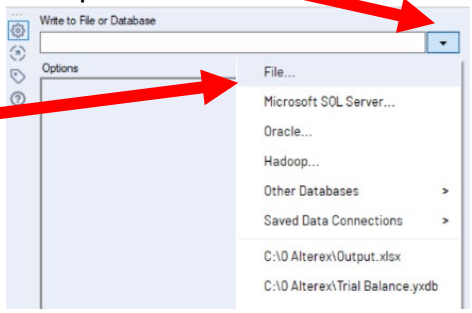
Tool 23: Output Data Tool: Export the cleaned and assembled "Comm by Salesperson by Region" report.

Drag an "**Output Data Tool**" to the right of the "**Select Tool**" on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



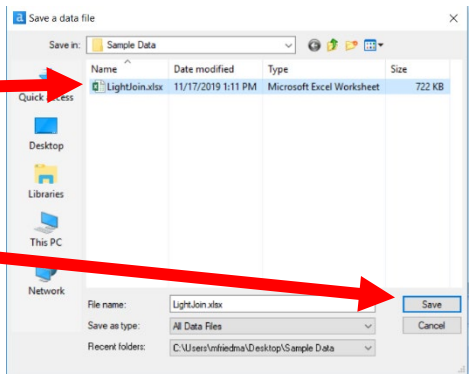
In the output Configuration area, select the drop-down arrow.

Select "File".

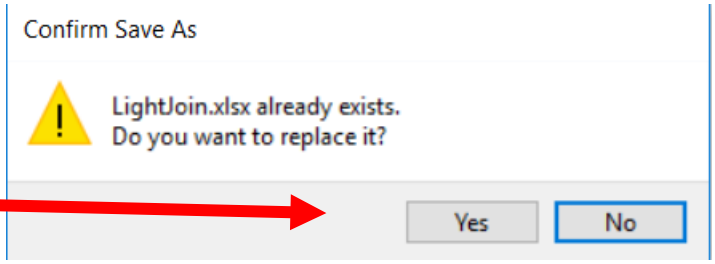


Select the required file name.

Select "Save".

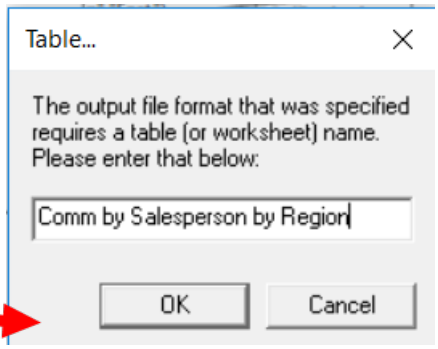


Click on “Yes” to confirm “Save”.

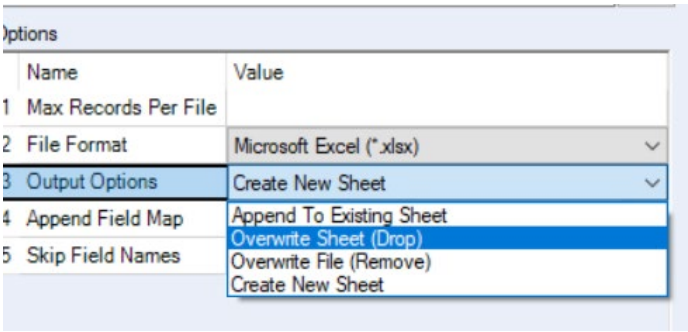


Enter the “Sheet Name”.

Select “OK”.



Change “3 Output Options” to Overwrite Sheet (Drop) to avoid producing an error the second time you save the sheet to the file.



Select “Run” to apply.

Open the file in Excel.

	A	B	C	D	E	F	G
1	Salespersc	East	North	South	West	Total	
2	George Ha	17192.18	26615.28	26052.39	21075.49	90935.34	
3	John Lenn	84829.28	29372.96	82074.3	29873.3	226149.8	
4	Paul McCa	30874.29	52700.71	15502.96	28115.56	127193.5	
5	Ringo Star	60005.82	136168.6	44940.09	33075.06	274189.6	
6	Total	192901.6	244857.6	168569.7	112139.4	718468.3	
7							

Change the columns' widths and formats.

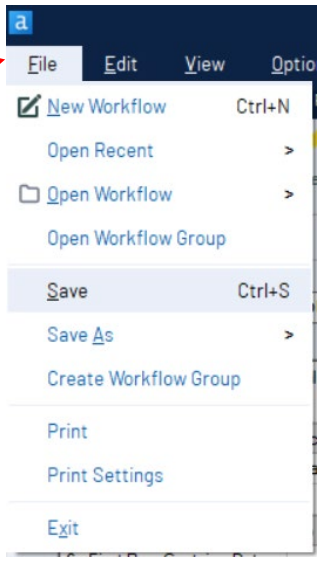
	A	B	C	D	E	F
1	Salesperson	East	North	South	West	Total
2	George Harrison	$ 17,192.18	$ 26,615.28	$ 26,052.39	$ 21,075.49	$ 90,935.34
3	John Lennon	84,829.28	29,372.96	82,074.30	29,873.30	226,149.84
4	Paul McCartney	30,874.29	52,700.71	15,502.96	28,115.56	127,193.52
5	Ringo Starr	60,005.82	136,168.64	44,940.09	33,075.06	274,189.61
6	Total	$192,901.57	$244,857.59	$168,569.74	$112,139.41	$718,468.31
7						
8						

"Save" and "Close" the Excel file.
Return to Alteryx.

The workflow must be saved.

Select “File”.

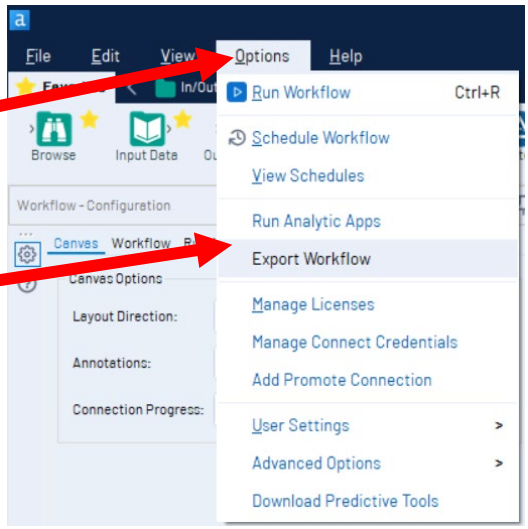
Select “Save”.



Export the workflow and the related input files so that the project can be completed on another computer or uploaded to the web.

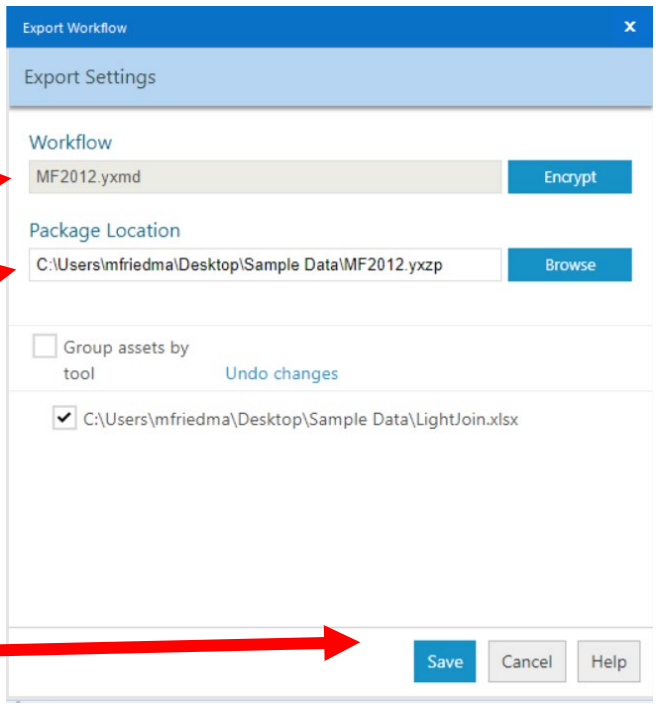
Select “Options”.

Select “Export Workflow”.

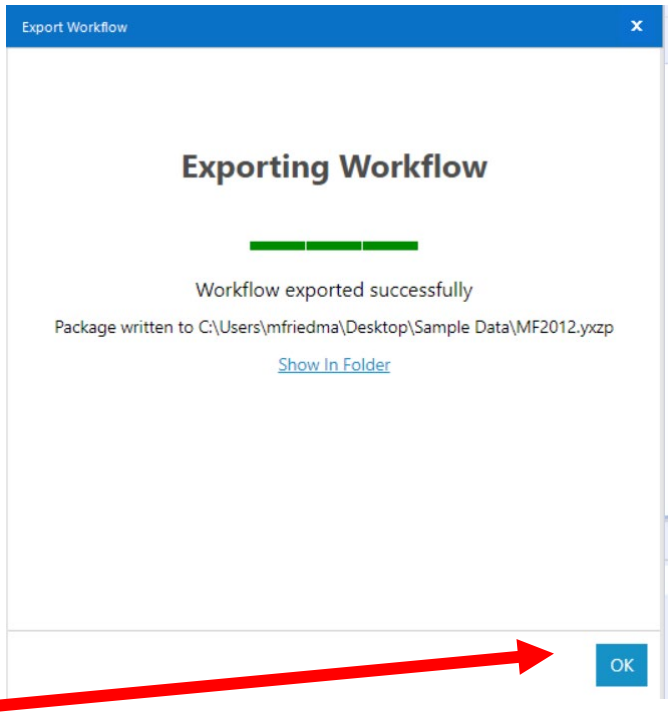


Verify the name of the “Workflow”

Verify the “Package Location”.



Select “Save”.



Select “OK”.

Exit Alteryx.