Part 1: Single Stream ----------1

Part 2: Append the Second Stream----------38

Using Alteryx

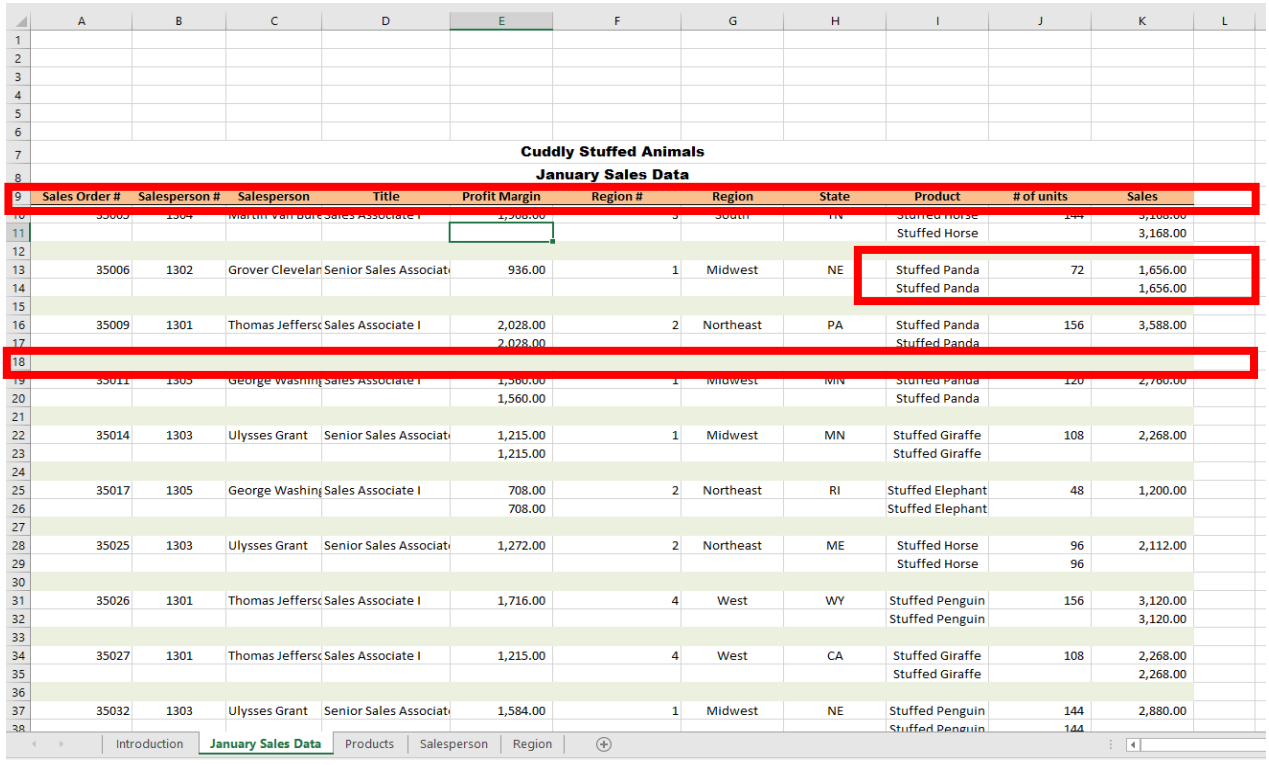
ETL software is used to automate the redundant manipulation of data. In this exercise you will use Alteryx to:

Extract the Data file(s) from Excel

Transform the data

Load the transformed data into Excel

The sales data for *Cuddly Stuffed Animals* are stored in an Excel sheet titled January Sales Data. The column headings are in the ninth row, some rows have partial duplicate data and some rows are blank.



Cuddly Stuffed Animals
January Sales Data

Sales Order #	Salesperson #	Salesperson	Title	Profit Margin	Region #	Region	State	Product	# of units	Sales
								Stuffed Horse		3,168.00
35006	1302	Grover Clevelan	Senior Sales Associat	936.00	1	Midwest	NE	Stuffed Panda	72	1,656.00
								Stuffed Panda		1,656.00
35009	1301	Thomas Jefferso	Sales Associate I	2,028.00	2	Northeast	PA	Stuffed Panda	156	3,588.00
				2,028.00				Stuffed Panda		
				1,560.00	1	Midwest		Stuffed Panda	120	2,760.00
				1,560.00				Stuffed Panda		
35014	1303	Ulysses Grant	Senior Sales Associat	1,215.00	1	Midwest	MN	Stuffed Giraffe	108	2,268.00
				1,215.00				Stuffed Giraffe		
35017	1305	George Washing	Sales Associate I	708.00	2	Northeast	RI	Stuffed Elephant	48	1,200.00
				708.00				Stuffed Elephant		
35025	1303	Ulysses Grant	Senior Sales Associat	1,272.00	2	Northeast	ME	Stuffed Horse	96	2,112.00
								Stuffed Horse	96	
35026	1301	Thomas Jefferso	Sales Associate I	1,716.00	4	West	WY	Stuffed Penguin	156	3,120.00
								Stuffed Penguin		3,120.00
35027	1301	Thomas Jefferso	Sales Associate I	1,215.00	4	West	CA	Stuffed Giraffe	108	2,268.00
								Stuffed Giraffe		2,268.00
35032	1303	Ulysses Grant	Senior Sales Associat	1,584.00	1	Midwest	NE	Stuffed Penguin	144	2,880.00
								Stuffed Penguin	144	

Introduction | January Sales Data | Products | Salesperson | Region

Sample Reports

Product	Units	Sales	Profit Margin
Stuffed Elephant	4,392	$109,800.00	$ 64,782.00
Stuffed Giraffe	8,052	$169,092.00	$ 90,585.00
Stuffed Horse	12,120	$266,640.00	$160,590.00
Stuffed Panda	7,488	$172,224.00	$ 97,344.00
Stuffed Penguin	5,016	$100,320.00	$ 55,176.00
Stuffed Unicorn	3,060	$ 73,440.00	$ 40,545.00
Total	40,128	$891,516.00	$509,022.00

Salesperson	Units	Sales	Profit Margin
1301 - Thomas Jefferson	9,528	$210,000.00	$119,466.00
1302 - Grover Cleveland	7,128	$162,276.00	$ 92,601.00
1303 - Ulysses Grant	4,116	$ 90,168.00	$ 51,294.00
1304 - Martin Van Buren	4,476	$ 98,388.00	$ 56,247.00
1305 - George Washington	10,452	$232,536.00	$133,557.00
1306 - John Tyler	4,428	$ 98,148.00	$ 55,857.00
Total	40,128	$891,516.00	$509,022.00

Start Alteryx

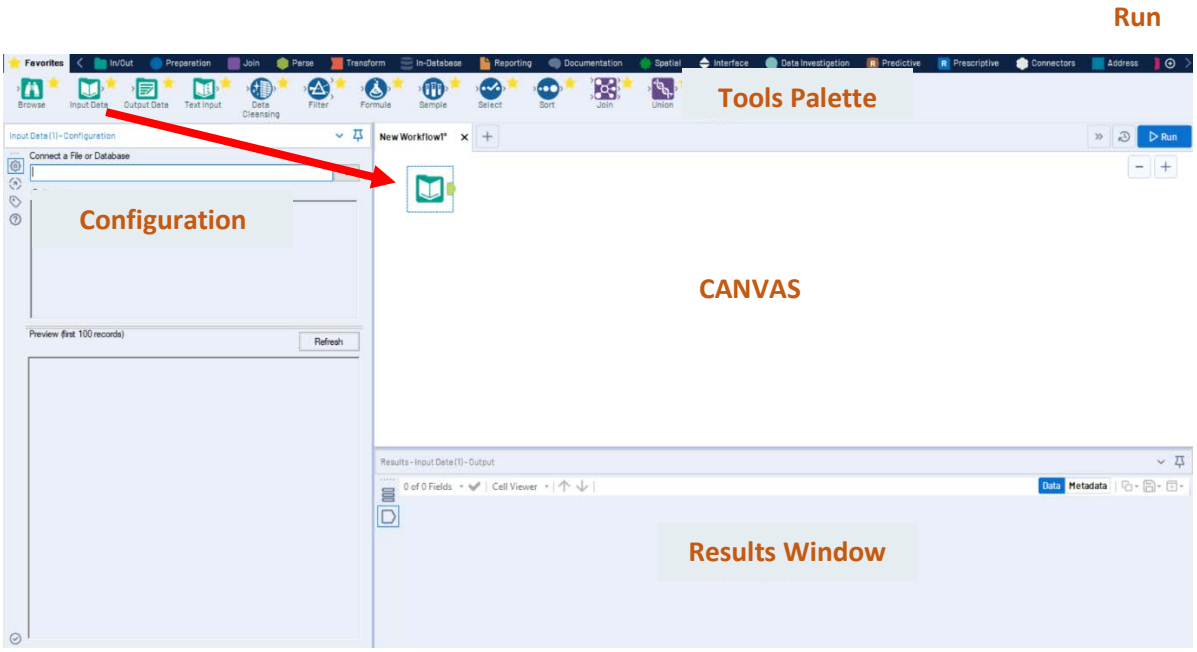
Note: You cannot extract data from an open Excel file so make sure that the Excel file is closed before starting to input the data.

There are four major parts of the Alteryx window. The first is the Canvas, and that's where we're going to create the workflow. 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 us the result after we select "Run" to apply the changes.

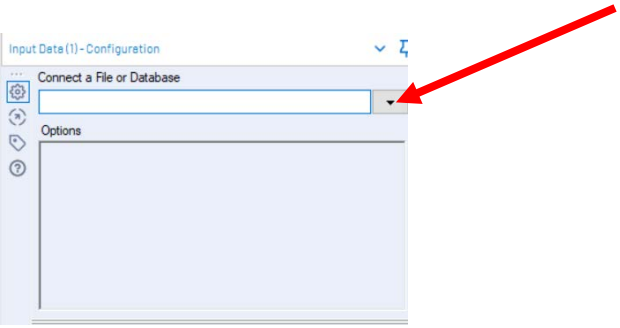
Tool 1: Input Data Tool – Input Excel file

Drag the "Input Data Tool"

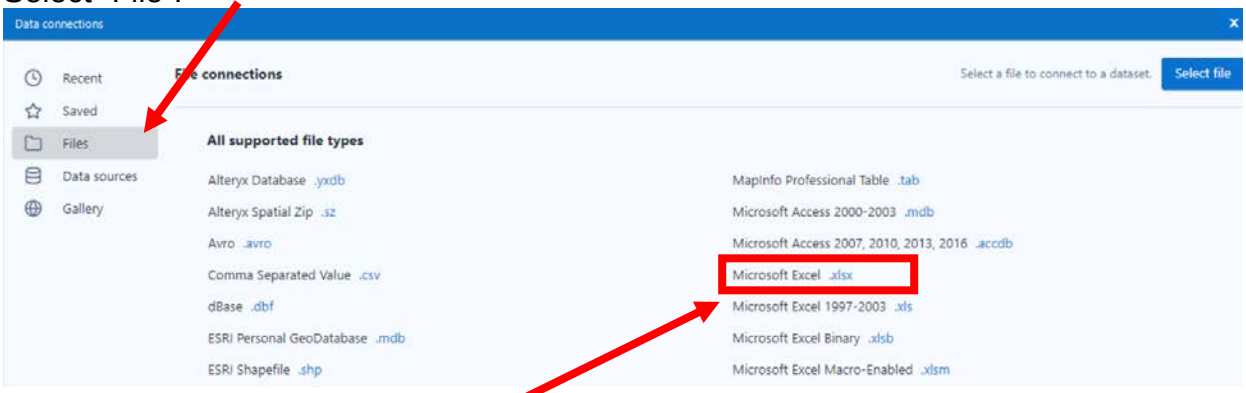
from the Tools Palette to the Canvas.



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



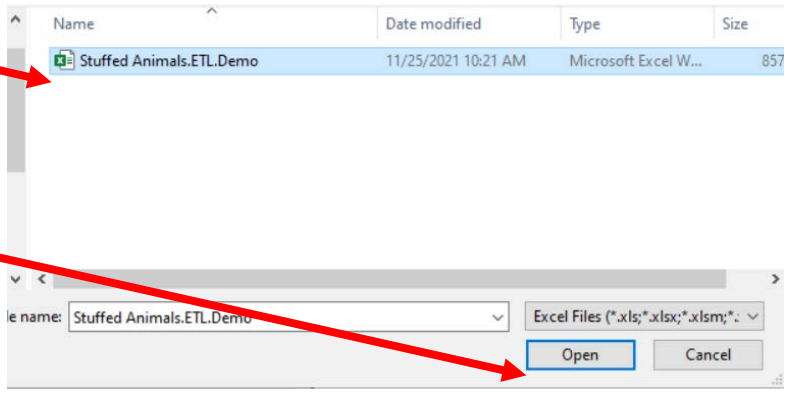
Select “File”.



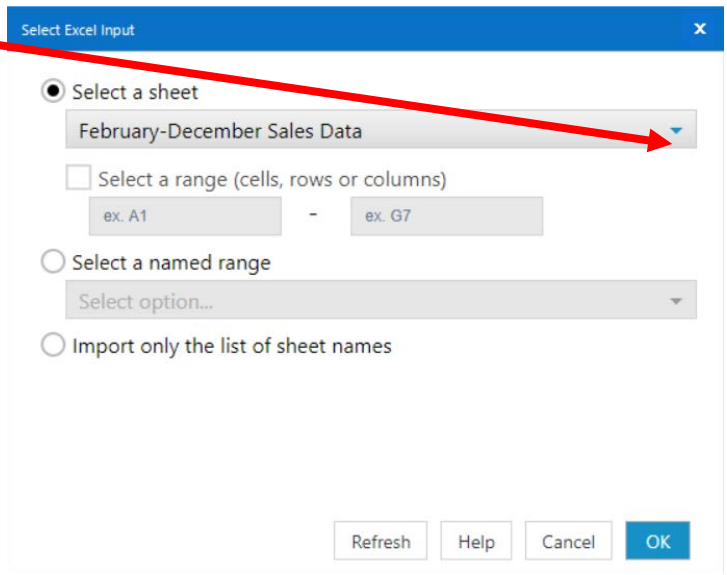
Select “Microsoft Excel .xlsx”.

Browse for the desired file.

Select “Open”.

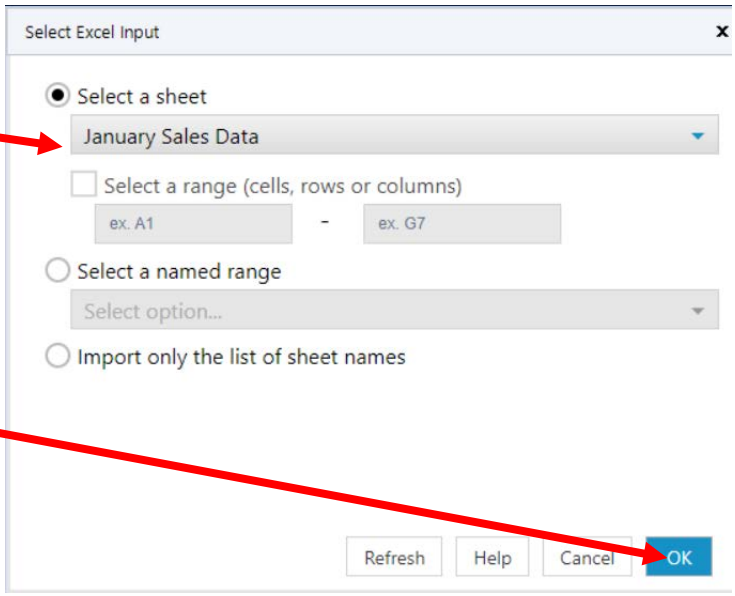


Select the drop-down.



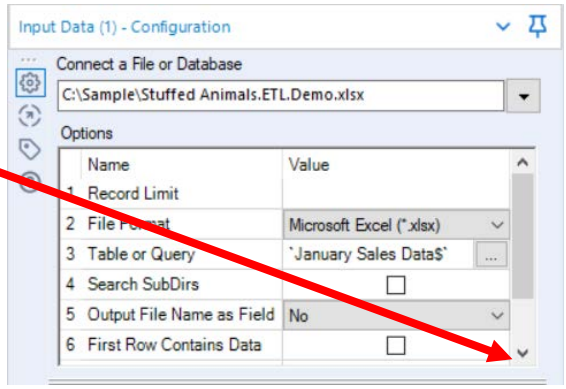
Select “January Sales Data”.

Select “OK”.

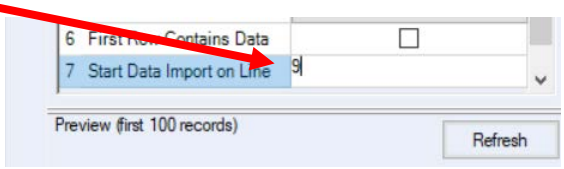


Remember that the field names are in the nineth row.

In the Configuration area, scroll down to option 7.

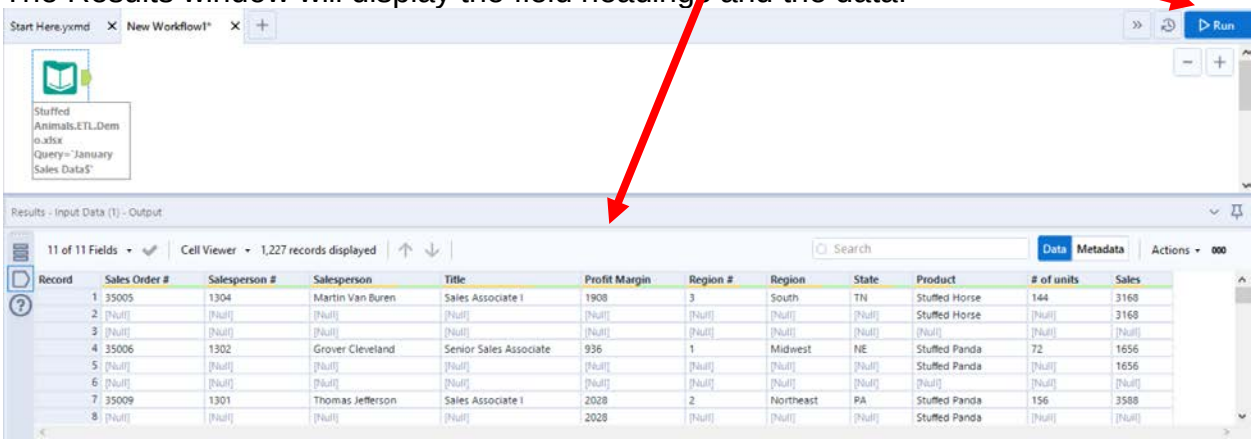


Change the number in “Start Data Import on Line” to “9”.



Select “Run”.

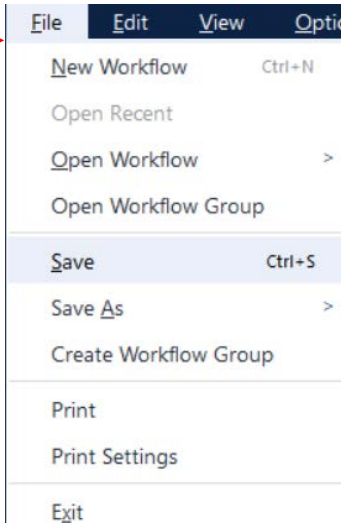
The Results window will display the field headings and the data.



The workflow must be saved.

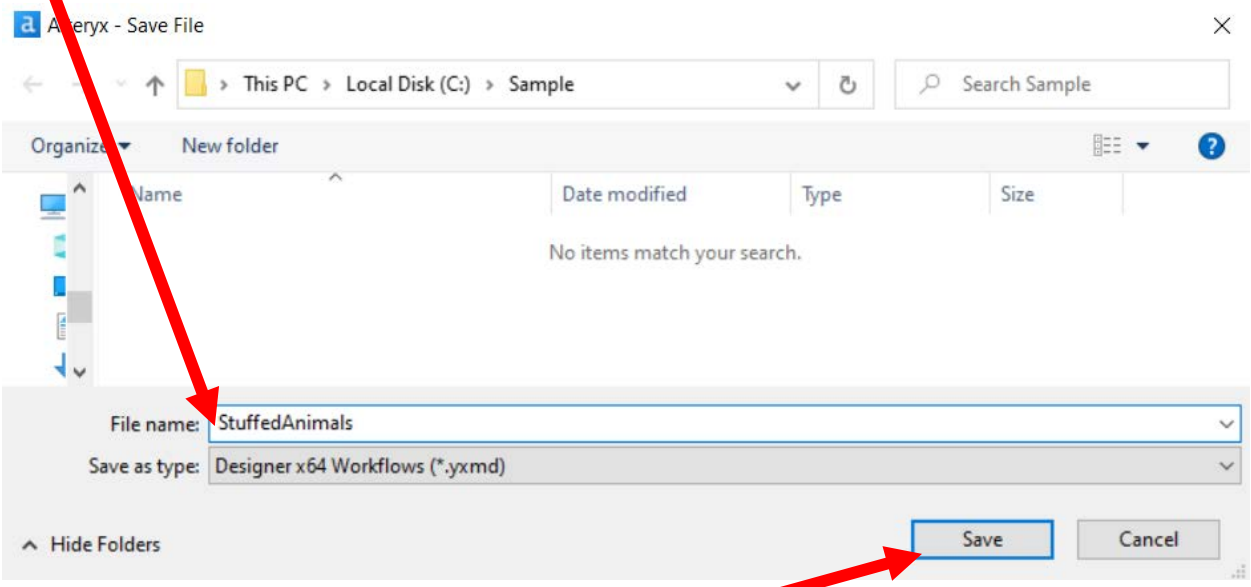
Select “File”.

Select “Save”.



Browse to the required directory.

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

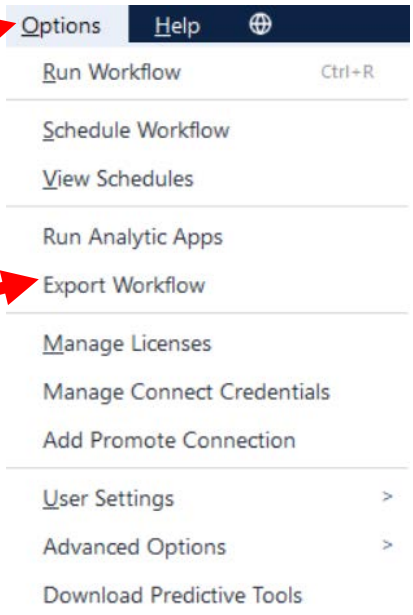


Select “Save”.

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”.

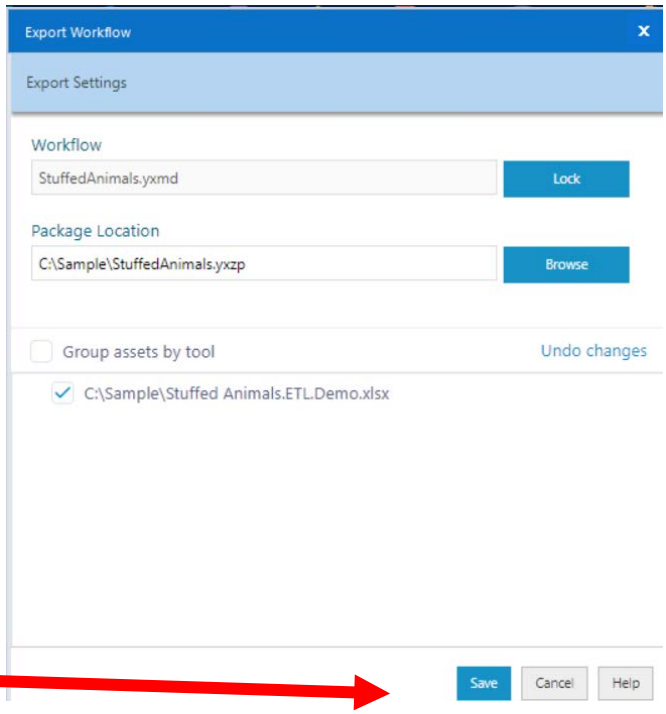
Select “Export Workflow”.



Verify the “Workflow”.

Verify the “Package Location”.

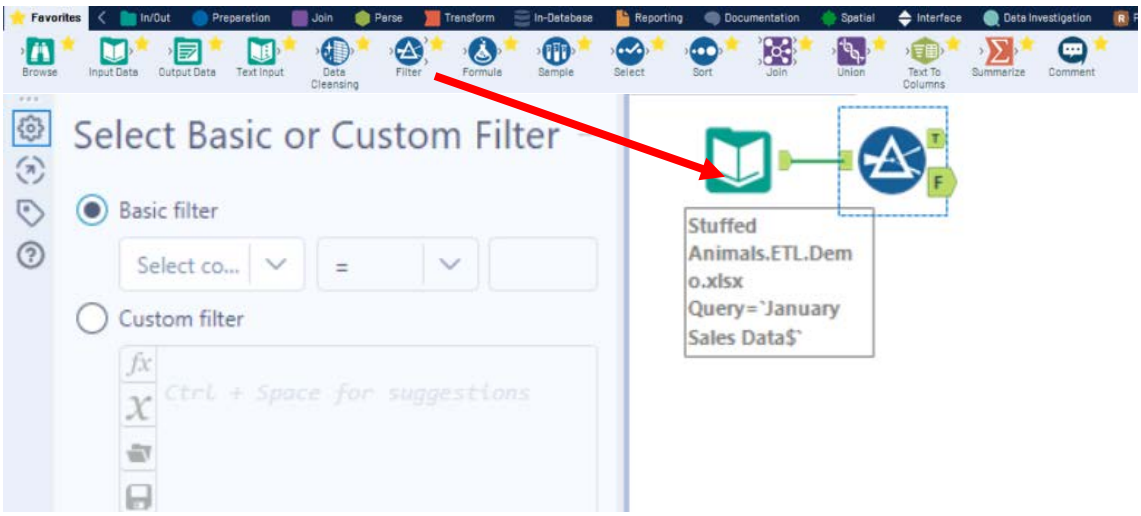
Select “Save”.



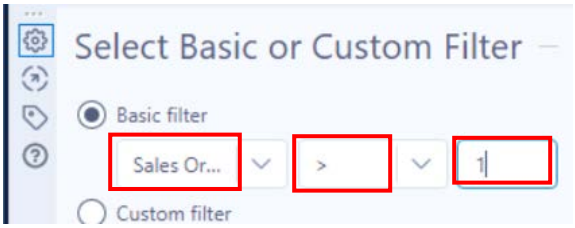
Tool 2: Filter Tool – Exclude all rows that do not contain data

There are several rows that do not contain an sales order number. We can modify the working dataset so that it includes only those rows that contain account numbers greater than one. Note that we are not deleting anything from the source data file. The working dataset will be split into two subsets. The first subset, “True”, will include all rows with account numbers greater than one. The second subset, “False”, will contain all rows with account numbers less than one.

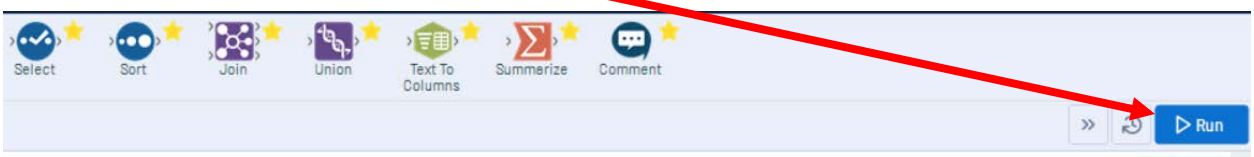
Drag the “**Filter Tool**” from the Tools Palette to the right of the “Input Data Tool” on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



With the “**Filter Tool**” selected, the “Basic” filter in the Configuration area can be modified to filter the field “Sales Order” for those accounts greater than “1’.



To apply the latest change, select “Run”.

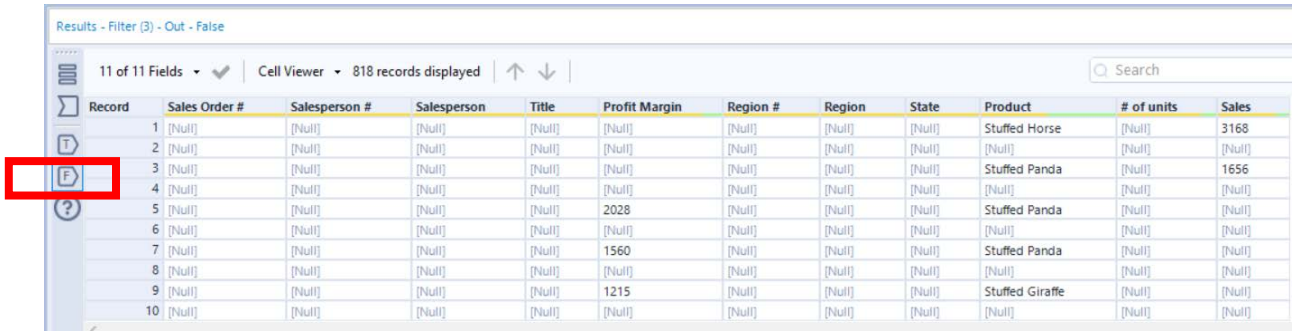


We can view three different datasets in the Results window.

The records that meet the filtered conditions, “Sales Order > 1”, are displayed as T (True).

Record	Sales Order #	Salesperson #	Salesperson	Title	Profit Margin	Region #	Region	State	Product	# of units	Sales
1	35005	1304	Martin Van Buren	Sales Associate I	1908	3	South	TN	Stuffed Horse	144	3168
2	35006	1302	Grover Cleveland	Senior Sales Associate	936	1	Midwest	NE	Stuffed Panda	72	1656
3	35009	1301	Thomas Jefferson	Sales Associate I	2028	2	Northeast	PA	Stuffed Panda	156	3588
4	35011	1305	George Washington	Sales Associate I	1560	1	Midwest	MN	Stuffed Panda	120	2760
5	35014	1303	Ulysses Grant	Senior Sales Associate	1215	1	Midwest	MN	Stuffed Giraffe	108	2268
6	35017	1305	George Washington	Sales Associate I	708	2	Northeast	RI	Stuffed Elephant	48	1200
7	35025	1303	Ulysses Grant	Senior Sales Associate	1272	2	Northeast	ME	Stuffed Horse	96	2112
8	35026	1301	Thomas Jefferson	Sales Associate I	1716	4	West	WY	Stuffed Penguin	156	3120
9	35027	1301	Thomas Jefferson	Sales Associate I	1215	4	West	CA	Stuffed Giraffe	108	2268
10	35032	1303	Ulysses Grant	Senior Sales Associate	1584	1	Midwest	NE	Stuffed Penguin	144	2880

The records that do not meet the filtered conditions are displayed as F (False).



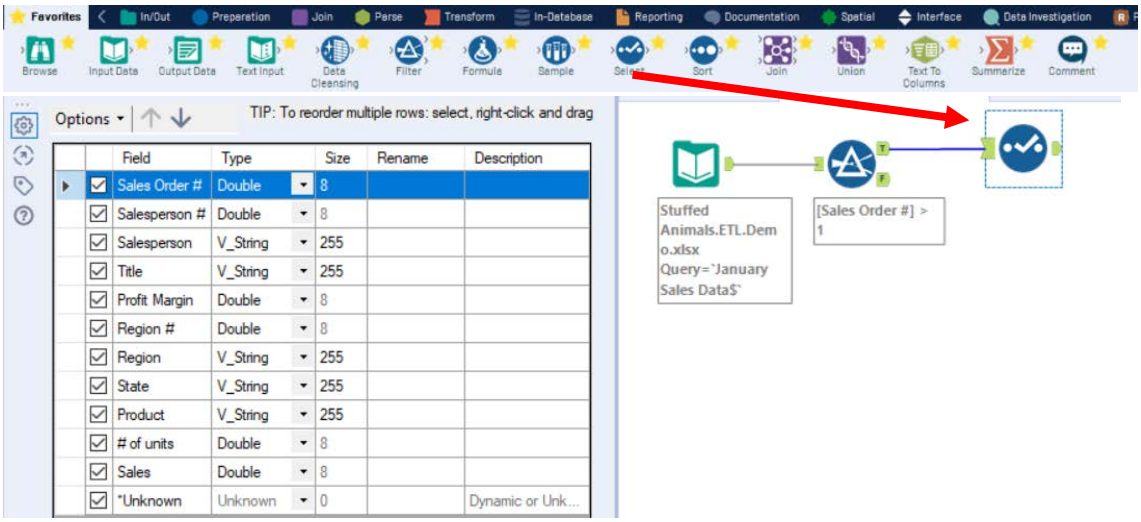
Record	Sales Order #	Salesperson #	Salesperson	Title	Profit Margin	Region #	Region	State	Product	# of units	Sales
1	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	Stuffed Horse	[Null]	3168
2	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
3	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	Stuffed Panda	[Null]	1656
4	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
5	[Null]	[Null]	[Null]	[Null]	2028	[Null]	[Null]	[Null]	Stuffed Panda	[Null]	[Null]
6	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
7	[Null]	[Null]	[Null]	[Null]	1560	[Null]	[Null]	[Null]	Stuffed Panda	[Null]	[Null]
8	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
9	[Null]	[Null]	[Null]	[Null]	1215	[Null]	[Null]	[Null]	Stuffed Giraffe	[Null]	[Null]
10	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]

The data being filtered can also be displayed.

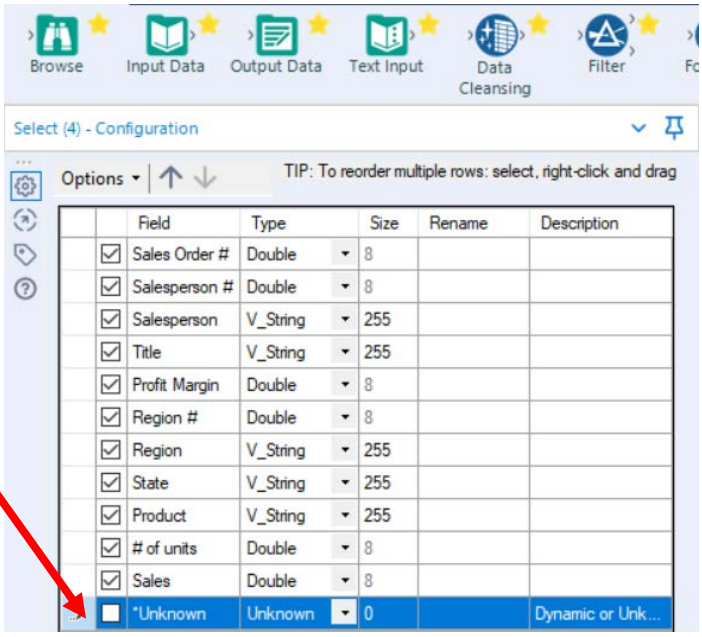
Record	Sales Order #	Salesperson #	Salesperson	Title	Profit Margin	Region #	Region	State	Product	# of units	Sales
1	35005	1304	Martin Van Buren	Sales Associate I	1908	3	South	TN	Stuffed Horse	144	3168
2	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	Stuffed Horse	[Null]	3168
3	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
4	35006	1302	Grover Cleveland	Senior Sales Associate	936	1	Midwest	NE	Stuffed Panda	72	1656
5	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	Stuffed Panda	[Null]	1656
6	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
7	35009	1301	Thomas Jefferson	Sales Associate I	2028	2	Northeast	PA	Stuffed Panda	156	3588
8	[Null]	[Null]	[Null]	[Null]	2028	[Null]	[Null]	[Null]	Stuffed Panda	[Null]	[Null]
9	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
10	35011	1305	George Washington	Sales Associate I	1560	1	Midwest	MN	Stuffed Panda	120	2760

Tool 3: Select Tool – Remove fields that contain no data, from "*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.

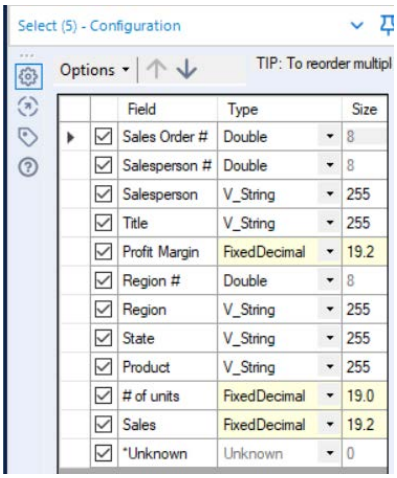


Remove any extra fields by unchecking them.

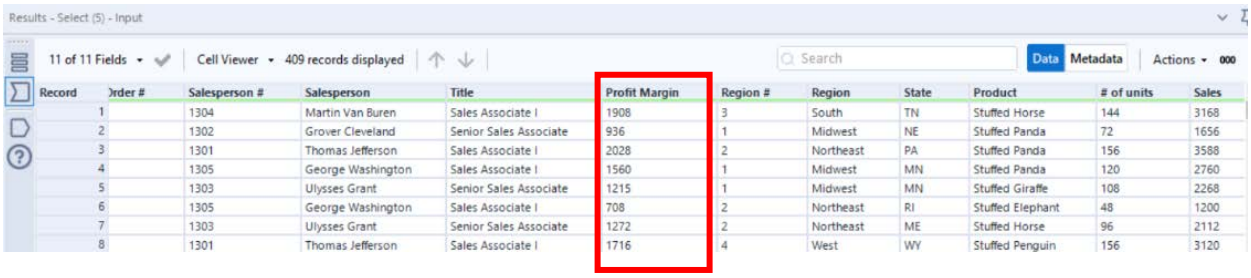


Change the type of all numeric fields that will be used in calculations to “FixedDecimal”. The “# of units should have zero decimal places and dollar values should have two decimal places.

.

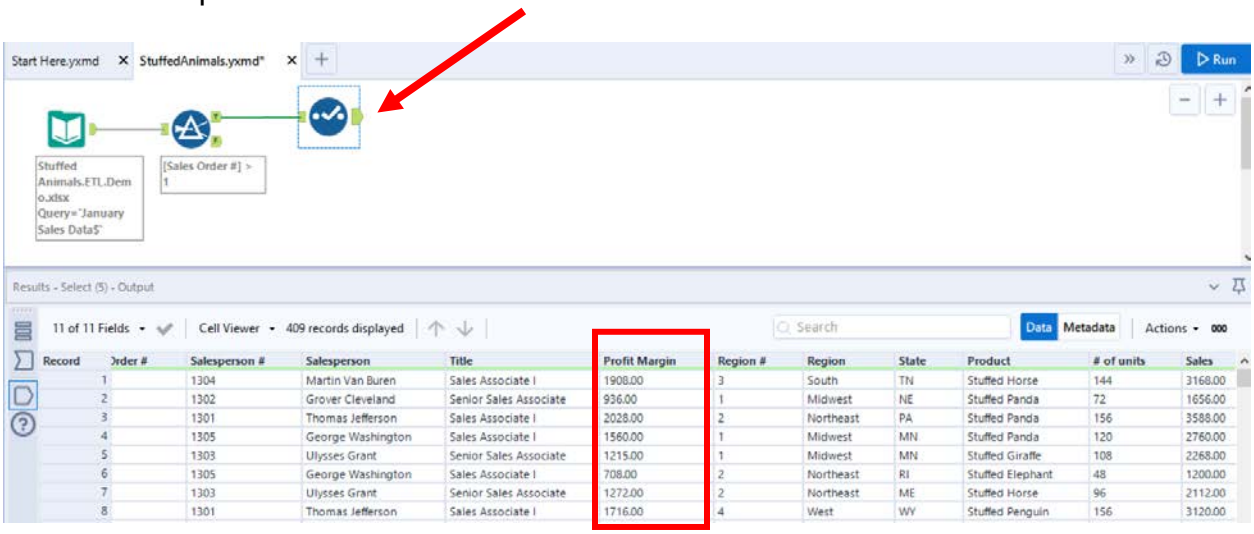


Before making changes:



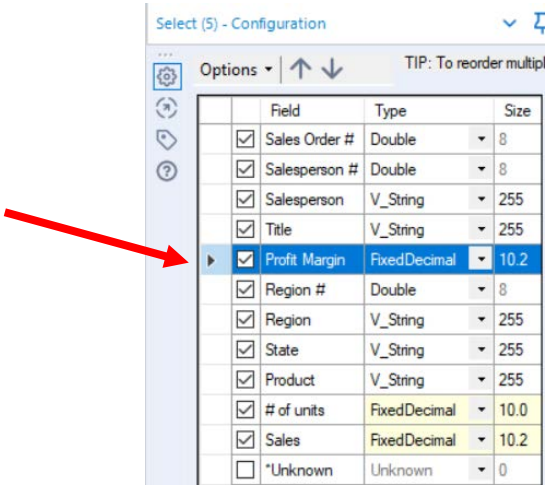
Select Run.

Select the output area of the Select tool.

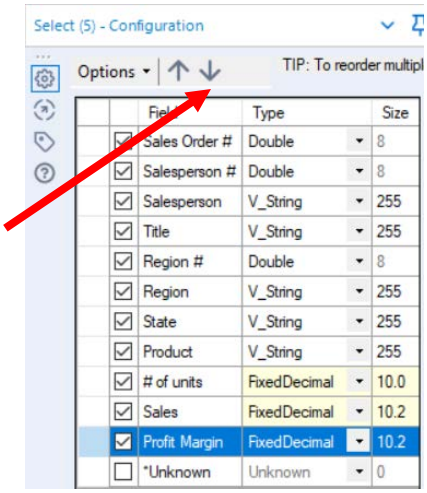


Change the order of the fields by moving “Profit Margin” to the end.

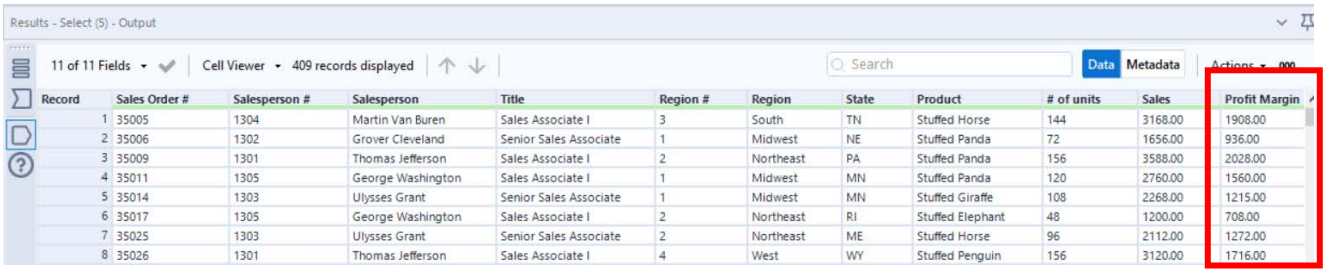
In the Configuration section select “Profit Margin”.



Use the down arrow to move “Profit Margin” to the end of the list.



Select Run.



Temporary Tool 4: Browse Tool – The Data in the Fields Must Conform to Predefined Rules.

In the original Excel file, The Cuddly Stuffed Animals company provided the only acceptable values for “Product”, “Salesperson” and “Region”. Sometime data is entered incorrectly into the data.

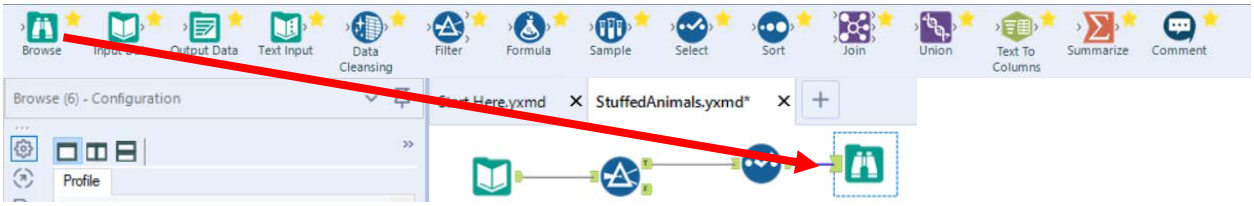
Product ID	Product
101	Stuffed Penguin
102	Stuffed Horse
103	Stuffed Unicorn
104	Stuffed Giraffe
105	Stuffed Panda
106	Stuffed Elephant

Code	Salesperson
1301	Thomas Jefferson
1302	Grover Cleveland
1303	Ulysses Grant
1304	Martin Van Buren
1305	George Washington
1306	John Tyler

Code	Region
1	Midwest
2	Northeast
3	South
4	West

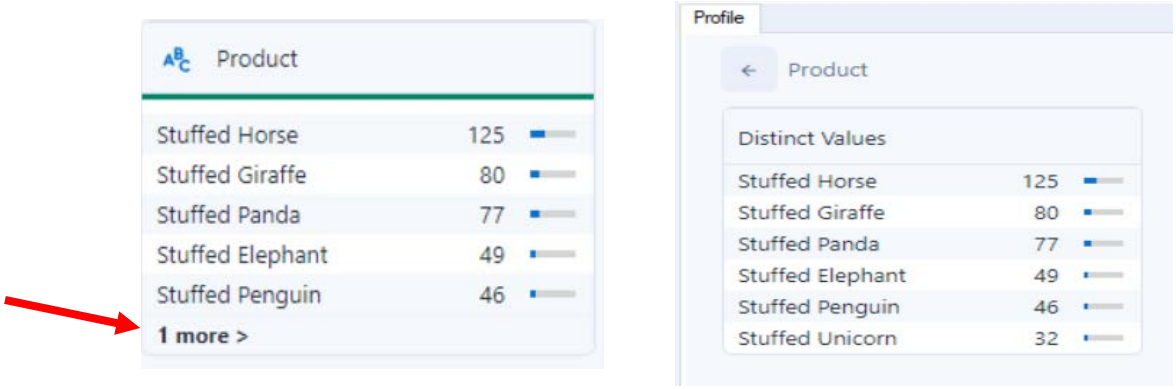
The “Product” field can only contain one of the six products. Browse the data in the “Product” field to verify that all the data is correct. Make any changes that are required.

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



Select Run.

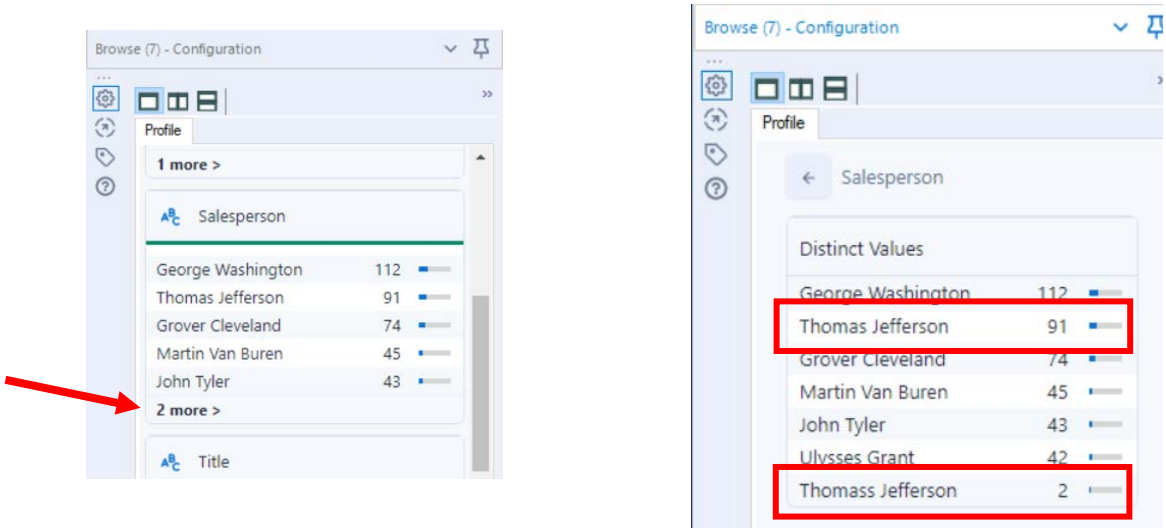
In the Configuration section scroll down to “Product” and select “1 more”.



The data is clean, there are only six products listed.

The “Salesperson” field can only contain one of the six names. Browse the data in the “Salesperson” field to verify that all the data is correct. Make any changes that are required.

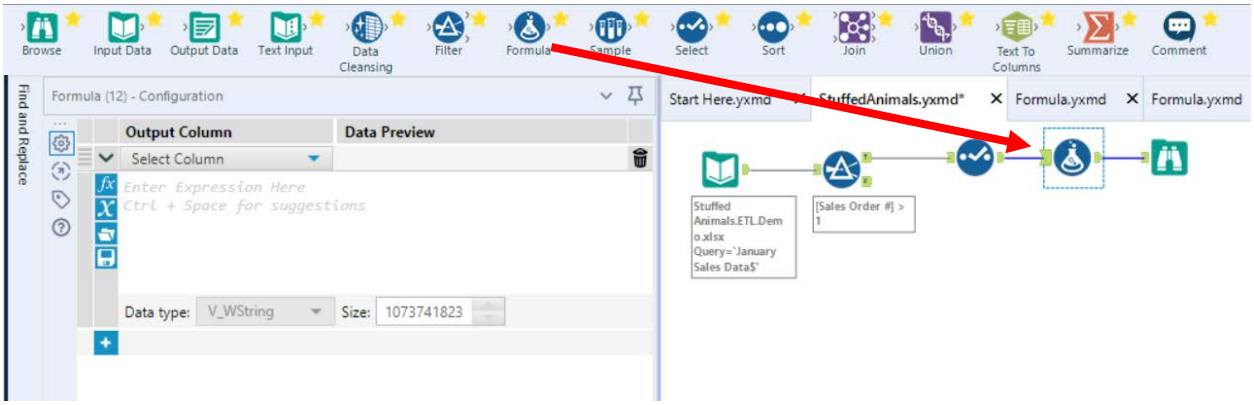
In the Configuration section scroll down to “Salesperson” and select “2 more”.



“Thomass Jefferson” must be replaced with “Thomas Jefferson” for a total of 93 distinct values.

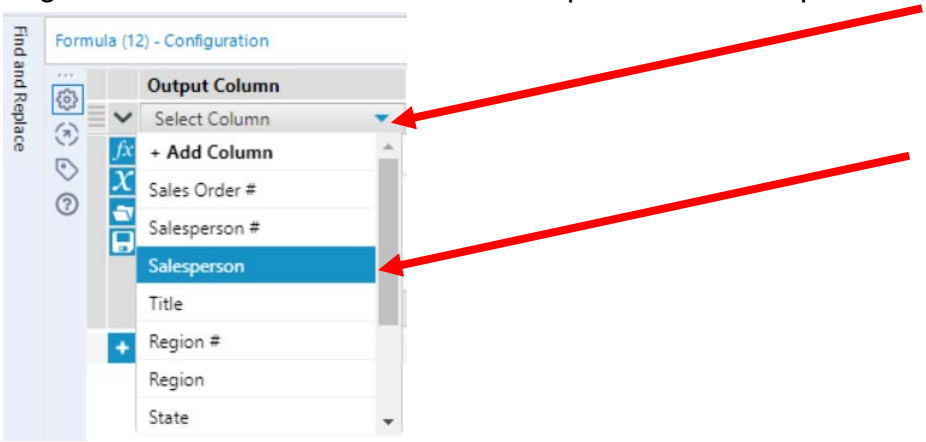
Tool 4: Formula Tool – The Data in the Fields Must Conform to Predefined Rules.

Drag the "**Formula Tool**" from the Tools Palette between the output of the "**Select Tool**" and the "Browse Tool" on the Canvas. The three will be connected automatically.

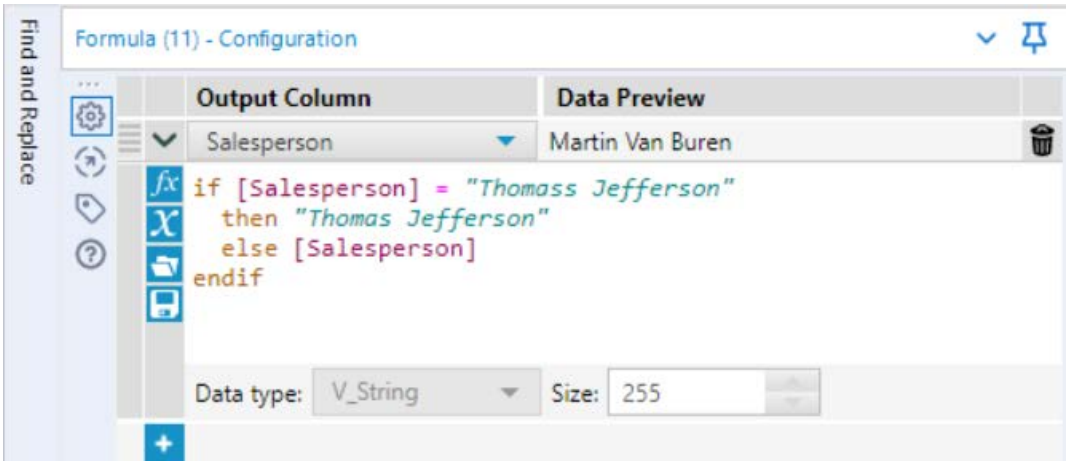


If the salesperson's name is "Thomass Jefferson" replace it with "Thomas Jefferson" otherwise leave the existing name in the field. Do not create a new field, put the output back into the "Salesperson" field.

In the configuration section select the field Salesperson from the pull-down list.

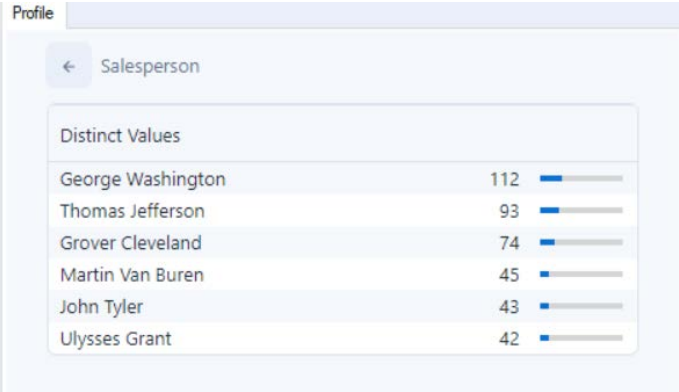


Type in the formula.



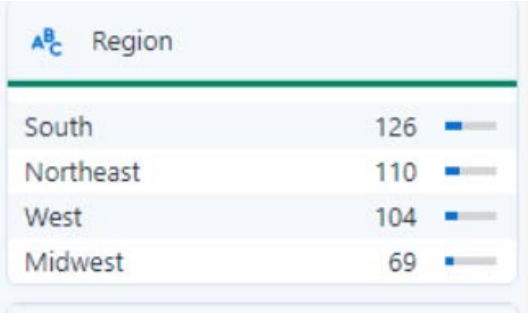
Select the Browse step.

Select Run.



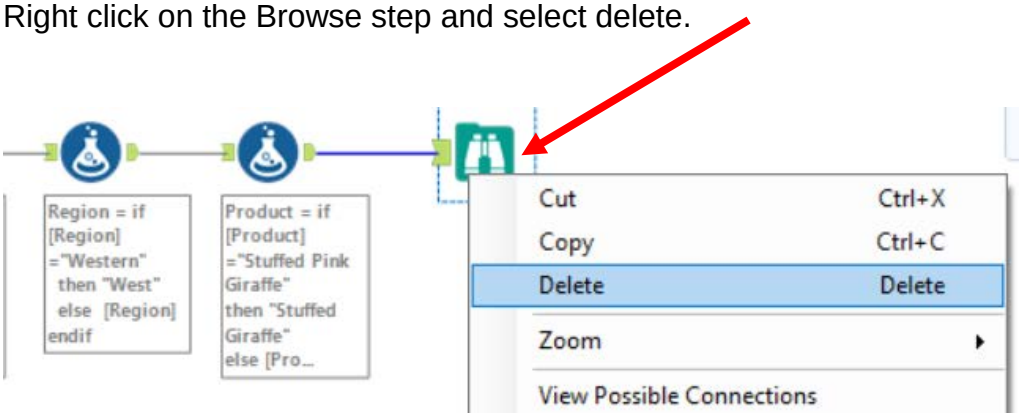
In the Configuration section scroll down to “Salesperson”. There are only six salespersons and Thomas Jefferson has 93 distinct values.

While in the Browse step, verify the field “Region”



The data is clean, only the four regions are listed.

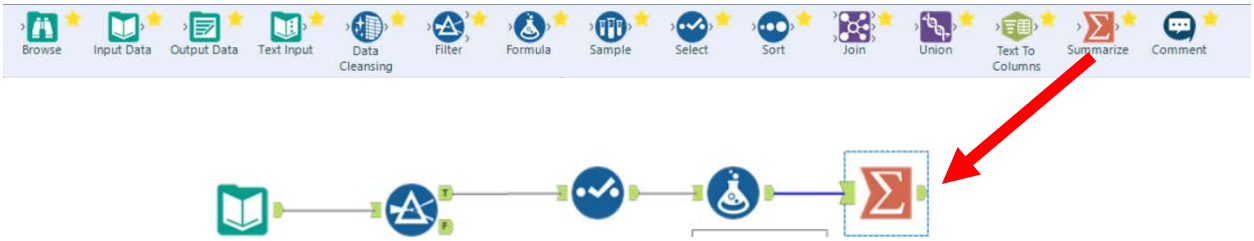
Right click on the Browse step and select delete.



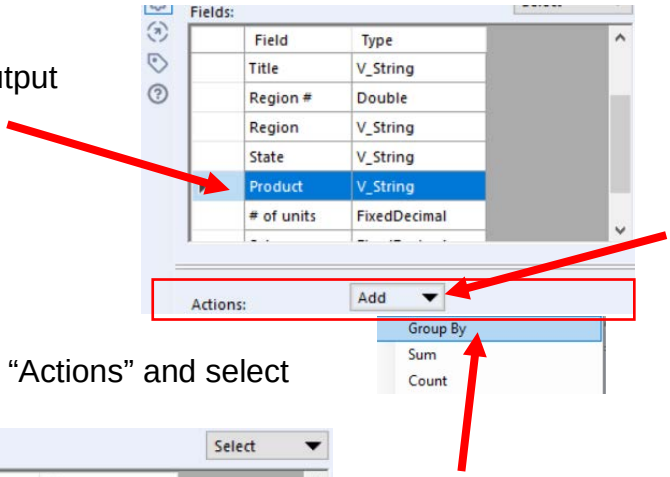
Tool 5: Select Tool Summary: We now want to create a report by product by first summarizing on "# of units" by product.

Product	# of units	Sales	Profit Margin
Stuffed Frog	1,524	$ 33,528	$20,574
Stuffed Unicorr	408	8,160	4,692
Stuffed Panda	1,140	26,220	15,675
Stuffed Koala	624	14,976	8,112
Stuffed Elephai	4,044	84,924	49,539
Stuffed Bear	2,376	59,400	35,046
Total	10,116	$227,208	$133,638

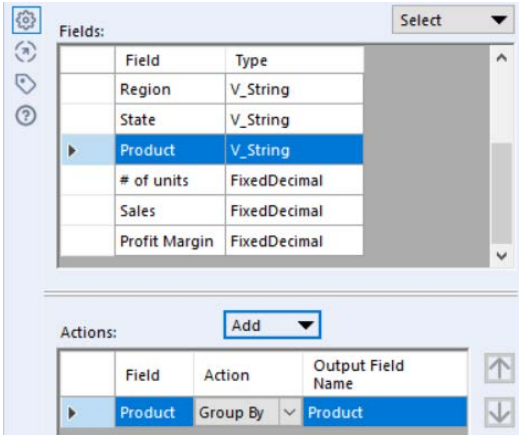
Drag a "**Summary Tool**" to the right of the "**Formula tool**" on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



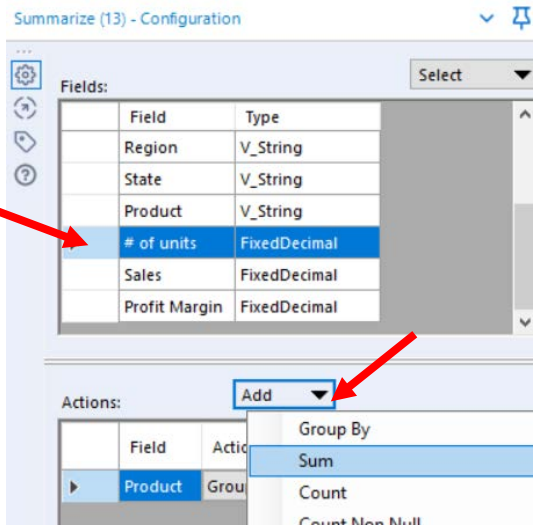
The first column will be Product. In the output Configuration area, select "Product".



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

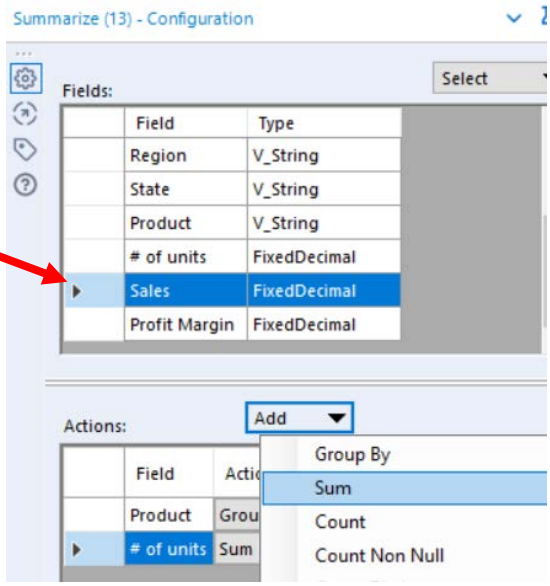


The second column will be “# of units”.



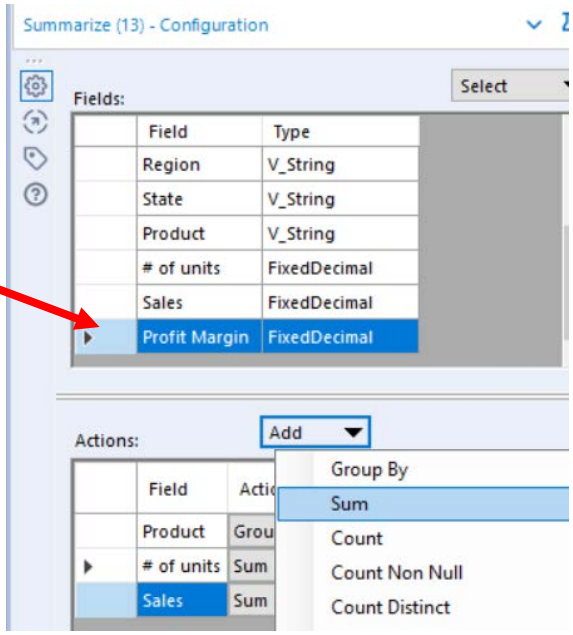
Click on the drop-down arrow on ”Add” in the “Actions” and select “Sum”.

The third column will be “Sales”.



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

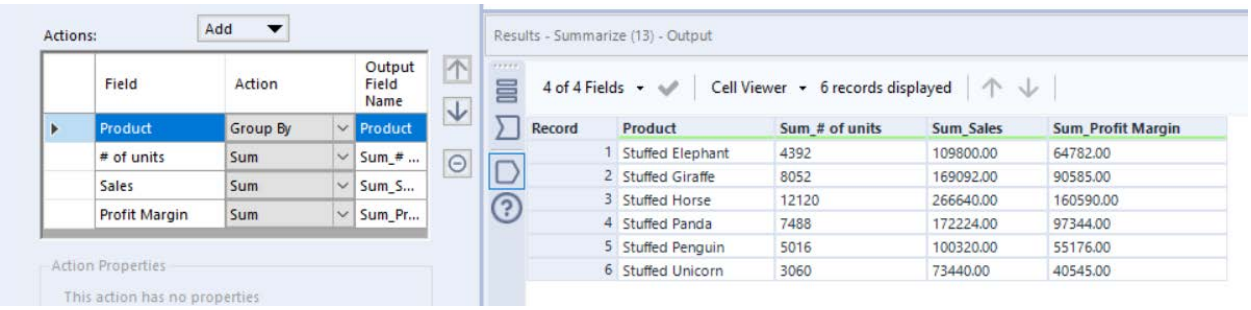
The fourth column will be “Profit Margin”.



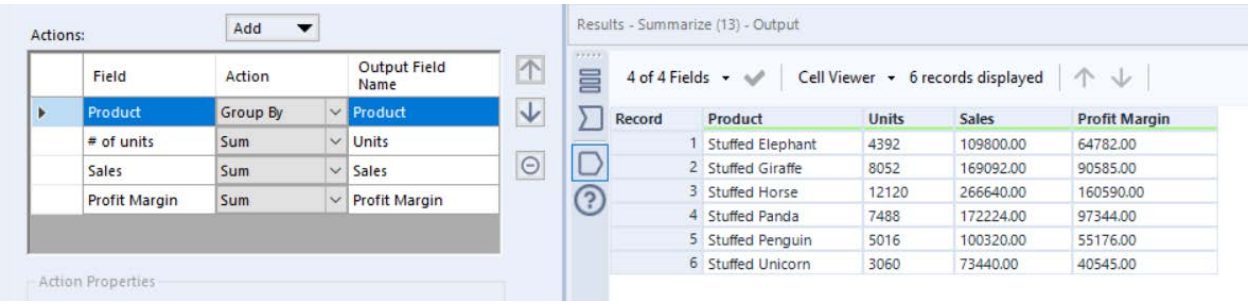
Click on the drop-down arrow on ”Add” in the “Actions” and select “Sum”.

Select “Run”.

The report is created.

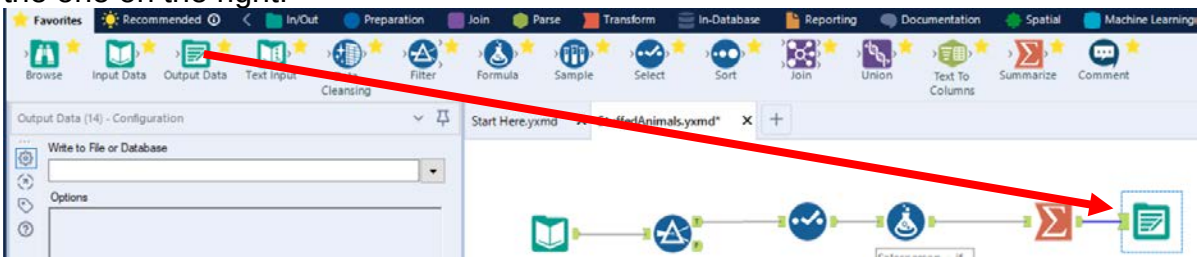


Change the “Output Field Names” and select “Run”

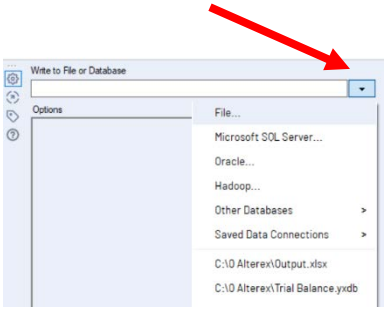


Tool 6: Select Tool Output Data Tool: We now want to export the cleaned and assembled Product Report to Excel.

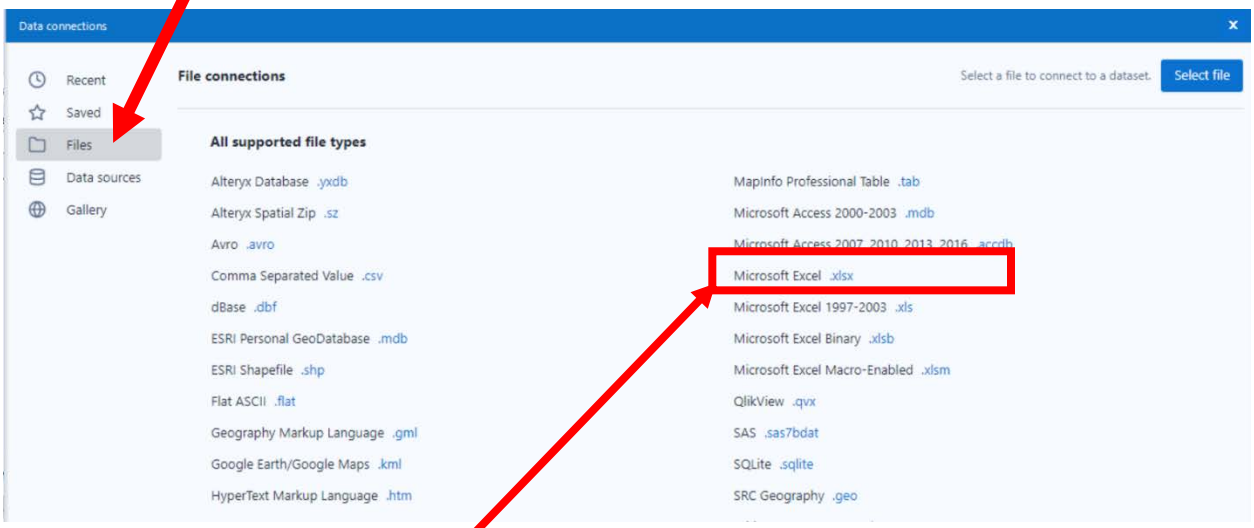
Drag an "**Output Data Tool**" to the right of the "**Summary Tool**" on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



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



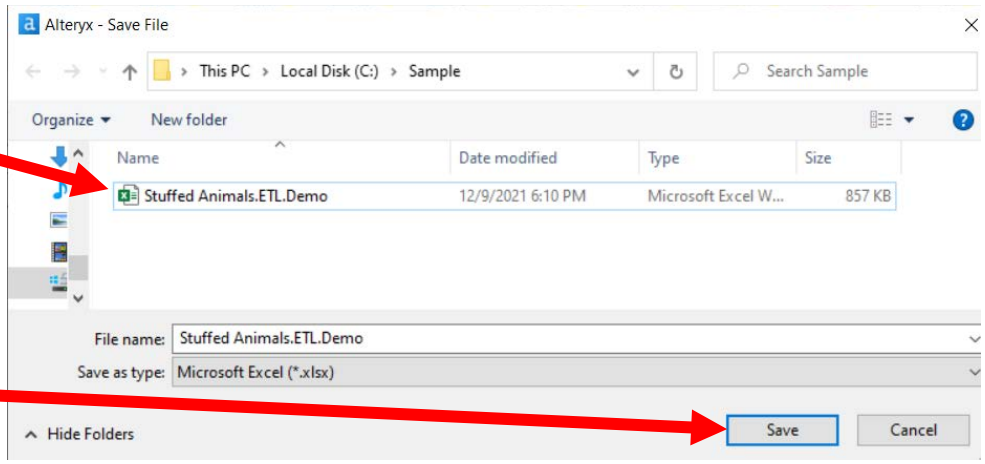
Select "File".



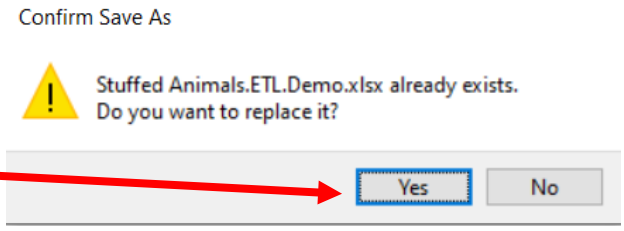
Select "Microsoft Excel .xlsx"

Select the required file name.

Select “Save”.

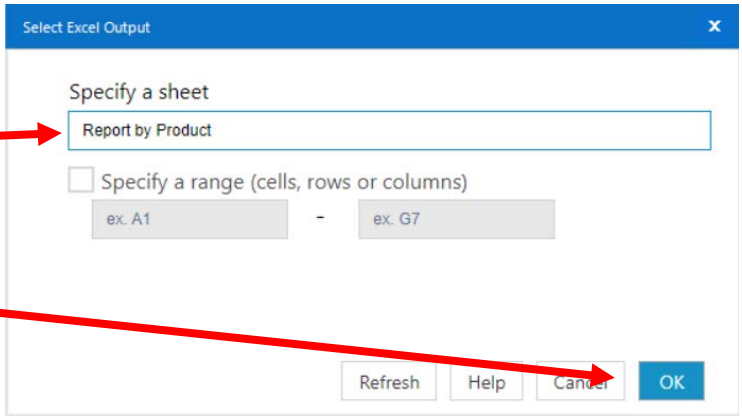


Click on “Yes” to confirm “Save”.



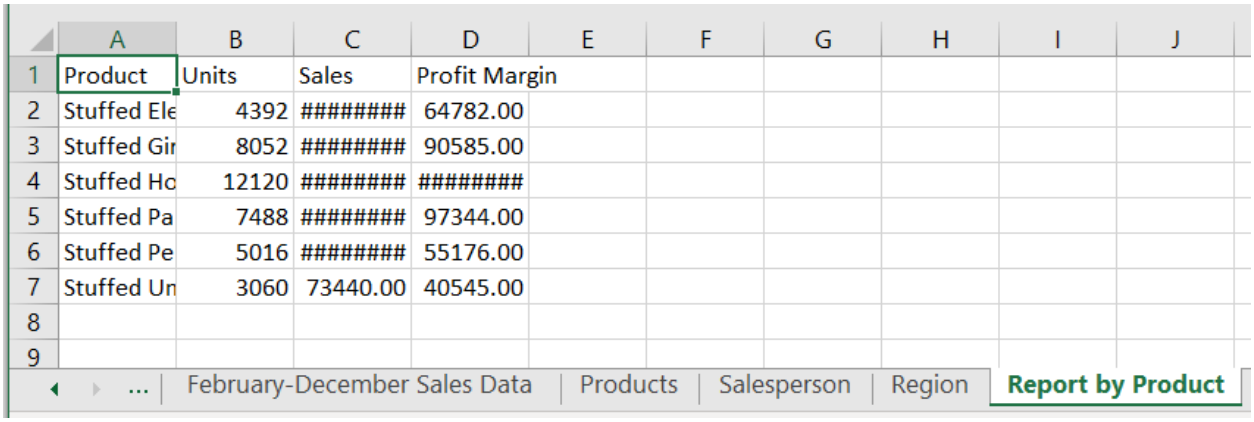
Enter the sheet name.

Select “OK”.



Select “Run”.

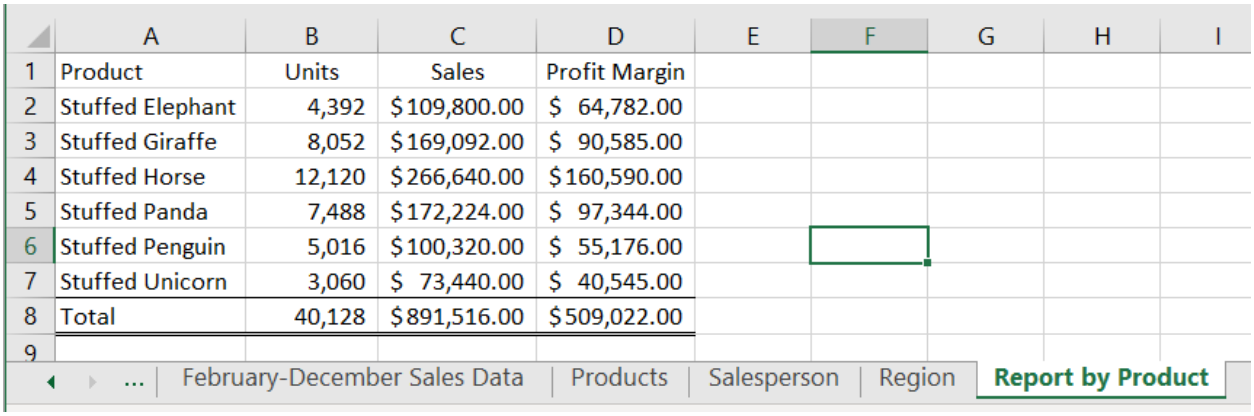
Open the file in Excel.



	A	B	C	D	E	F	G	H	I	J
1	Product	Units	Sales	Profit Margin						
2	Stuffed Ele	4392	########	64782.00						
3	Stuffed Gir	8052	########	90585.00						
4	Stuffed Ho	12120	########	########						
5	Stuffed Pa	7488	########	97344.00						
6	Stuffed Pe	5016	########	55176.00						
7	Stuffed Un	3060	73440.00	40545.00						
8										
9										

... | February-December Sales Data | Products | Salesperson | Region | **Report by Product**

Change the columns' widths and formats.

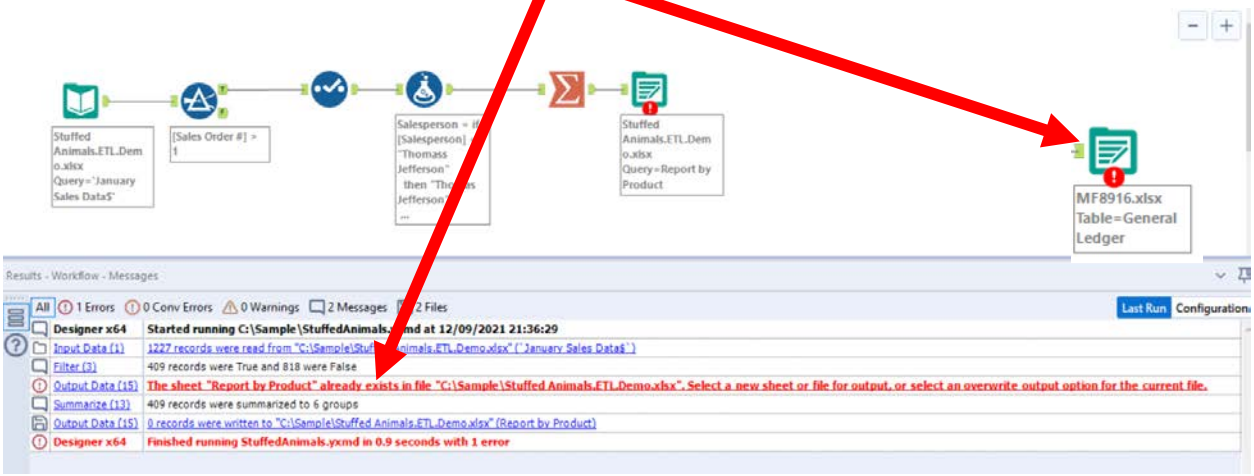


	A	B	C	D	E	F	G	H	I
1	Product	Units	Sales	Profit Margin					
2	Stuffed Elephant	4,392	$109,800.00	$ 64,782.00					
3	Stuffed Giraffe	8,052	$169,092.00	$ 90,585.00					
4	Stuffed Horse	12,120	$266,640.00	$160,590.00					
5	Stuffed Panda	7,488	$172,224.00	$ 97,344.00					
6	Stuffed Penguin	5,016	$100,320.00	$ 55,176.00					
7	Stuffed Unicorn	3,060	$ 73,440.00	$ 40,545.00					
8	Total	40,128	$891,516.00	$509,022.00					
9									

... | February-December Sales Data | Products | Salesperson | Region | **Report by Product**

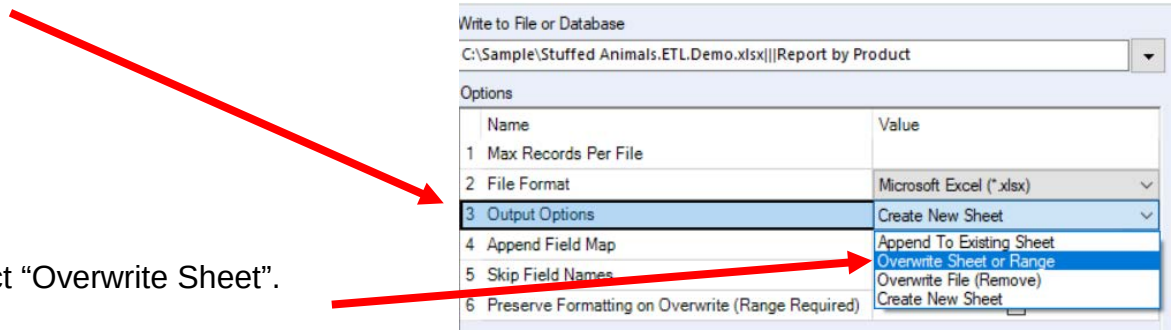
"Save" and "Close" the Excel file.
Return to Alteryx and select "Run" for the second time.

But, there is an error! The sheet already exits.



In the output Configuration area,
select “Output Options” number 3.

Select “Overwrite Sheet”.

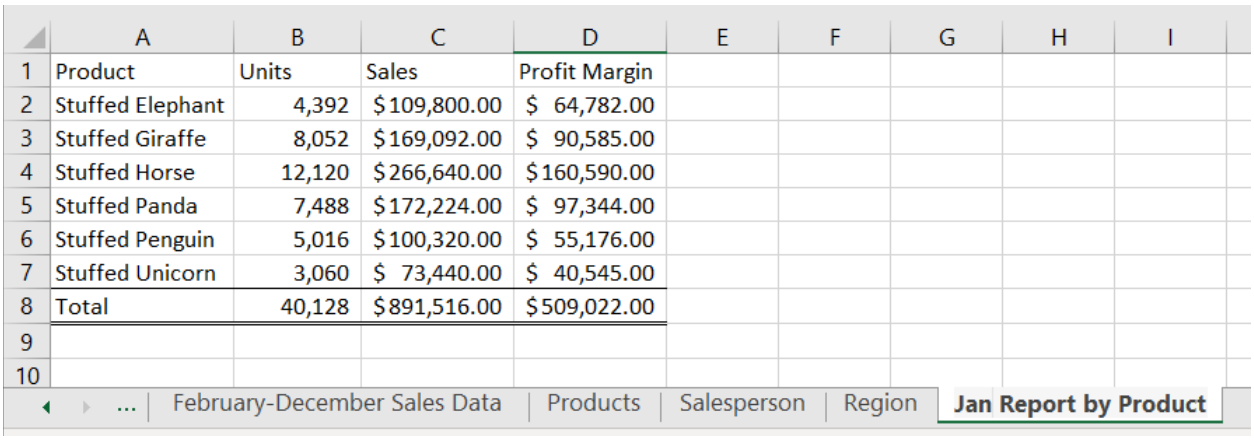


Select “Run”.

Open the Excel file.

	A	B	C	D	E	F	G	H	I	J
1	Product	Units	Sales	Profit Margin						
2	Stuffed Ele	4392	########	64782.00						
3	Stuffed Gir	8052	########	90585.00						
4	Stuffed Ho	12120	########	########						
5	Stuffed Pa	7488	########	97344.00						
6	Stuffed Pe	5016	########	55176.00						
7	Stuffed Un	3060	73440.00	40545.00						
8										
9										

... | February-December Sales Data | Products | Salesperson | Region | **Report by Product**



	A	B	C	D	E	F	G	H	I
1	Product	Units	Sales	Profit Margin					
2	Stuffed Elephant	4,392	$109,800.00	$ 64,782.00					
3	Stuffed Giraffe	8,052	$169,092.00	$ 90,585.00					
4	Stuffed Horse	12,120	$266,640.00	$160,590.00					
5	Stuffed Panda	7,488	$172,224.00	$ 97,344.00					
6	Stuffed Penguin	5,016	$100,320.00	$ 55,176.00					
7	Stuffed Unicorn	3,060	$ 73,440.00	$ 40,545.00					
8	Total	40,128	$891,516.00	$509,022.00					
9									
10									

... | February-December Sales Data | Products | Salesperson | Region | **Jan Report by Product**

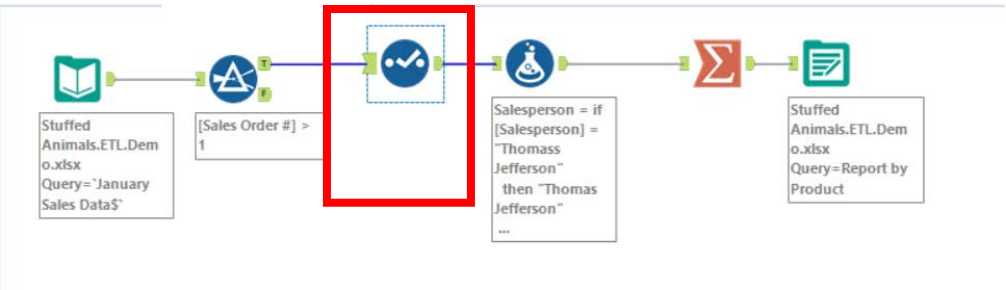
Reformat the Excel sheet and rename the formatted sheet “Jan Report by Product”. When the Alteryx flow runs it will create a new unformatted sheet “Report by Product”.

The “Salesperson” report will summarize the data by “Salesperson #” with the salesperson’s name attached for clarity. We now want to combine the fields with a space, a dash and another space between the two fields.

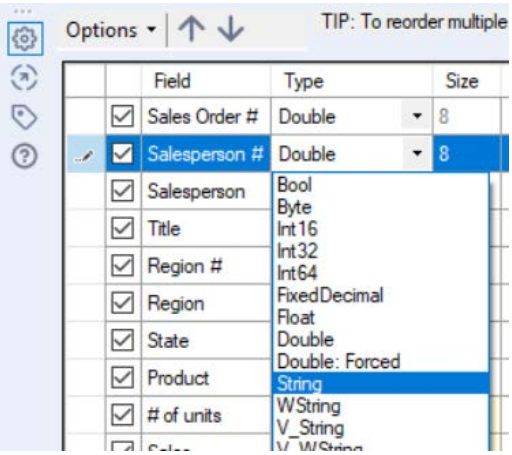
Salesperson	# of units	Sales	Profit Margin
1405 - James Polk	1,812	$41,568	$24,438
1425 - John Tyler	2,172	48,336	28,110
1541 - Ulysses Grant	756	17,112	10,062
1576 - James Madison	1,776	39,216	23,160
1588 - George Washington	1,584	35,784	21,168
1776 - James Monroe	2,016	45,192	26,700
Total	10,116	$227,208	$133,638

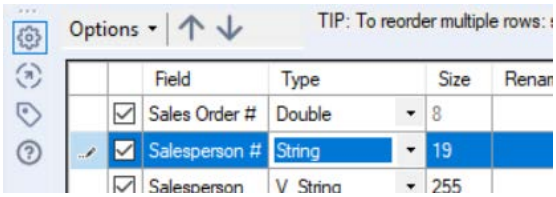
One field, “Salesperson #” is a numeric field, while the other field, “Salesperson’ is a string. It is easier to combine filed if they are both strings.

Select the third tool, the Select Tool from the flow.



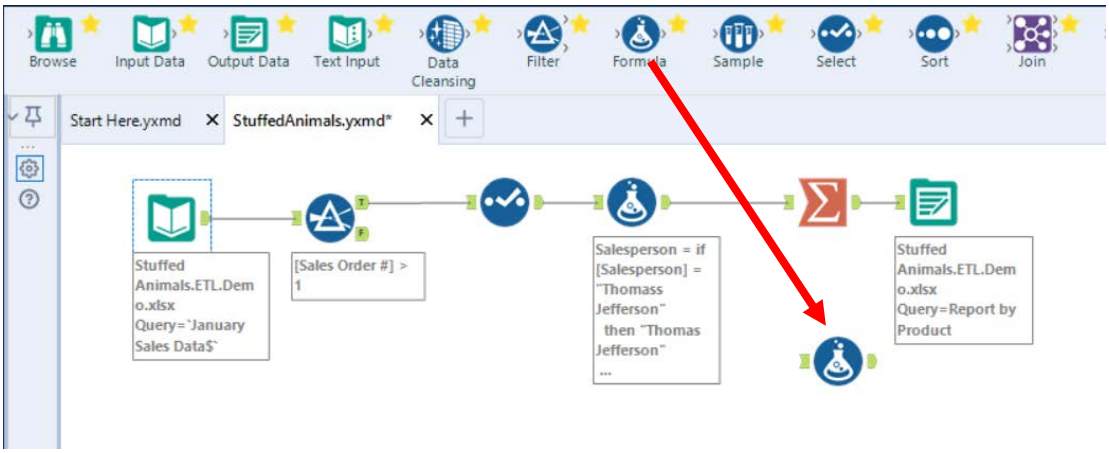
In the configuration section select the pull down in the “Type” column for the “Salesperson #” field and select “String”.



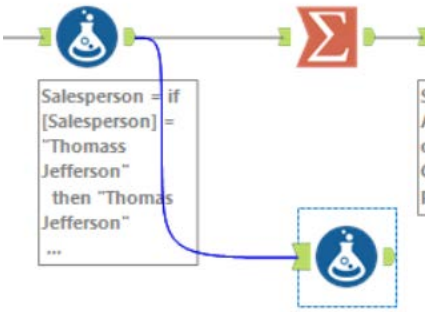


Tool 7: Select Tool Formula: We can now want to combine the fields “Salesperson #” and “Salesperson”.

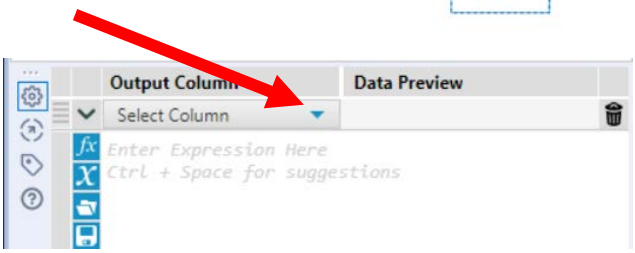
Drag a “**Formula** **Tool**” under the “**Running Balance Tool**” branching from the “**Formula Tool**” on the Canvas.



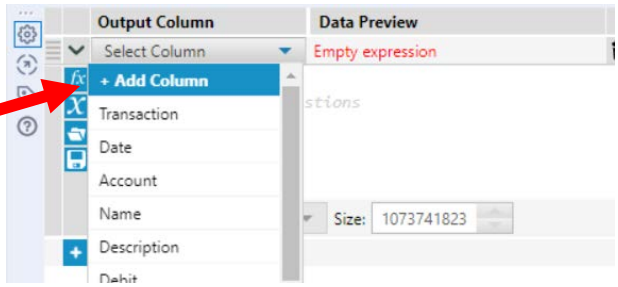
To connect the tools, drag from the outflow of the Tool 4, Formula Tool to the left inflow of the Tool 7, Formula tool.



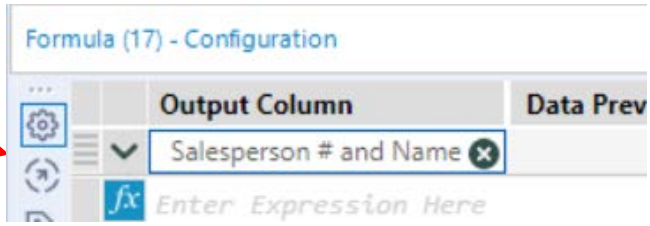
In the output Configuration area, click on the drop-down arrow of “Select Column”.



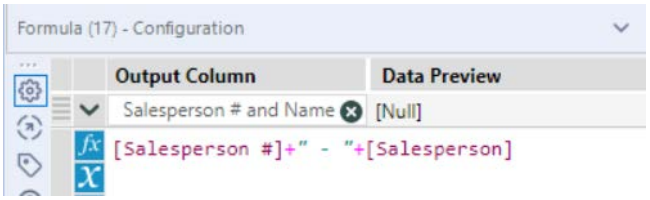
Select “Add Column”.



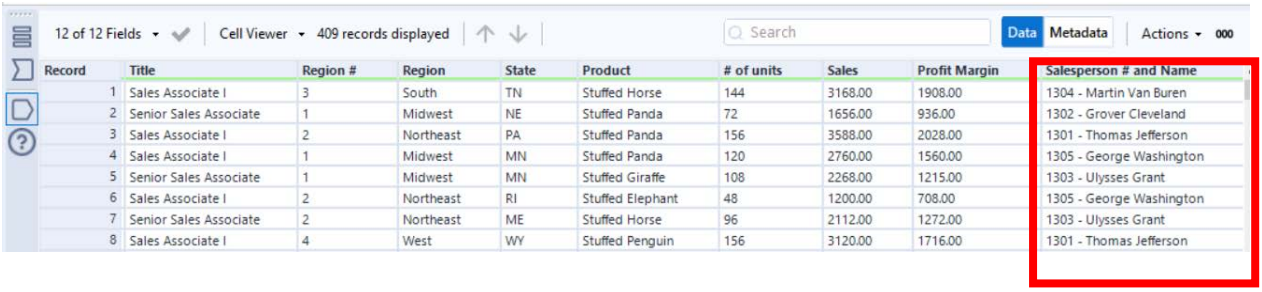
Type “Salesperson # And Name”.



Type the following formula
[Salesperson #]+" - "+[Salesperson]



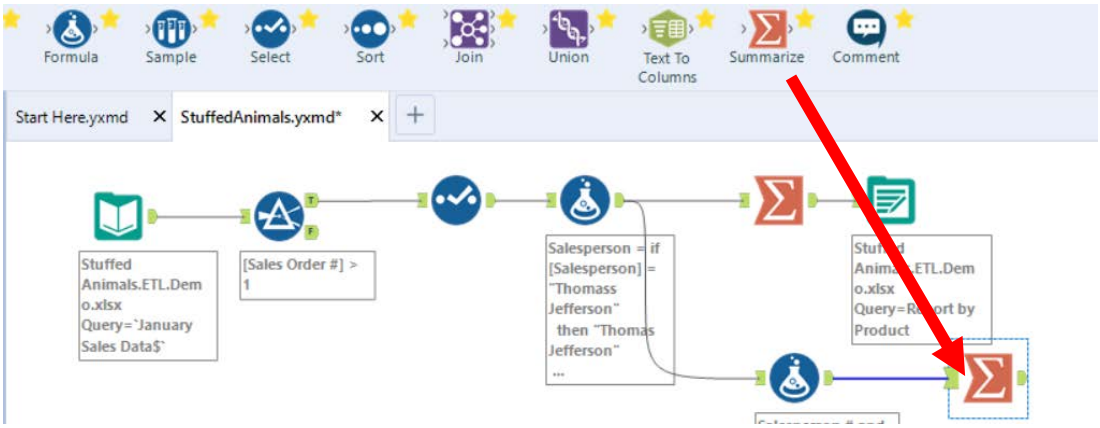
Select “Run”.



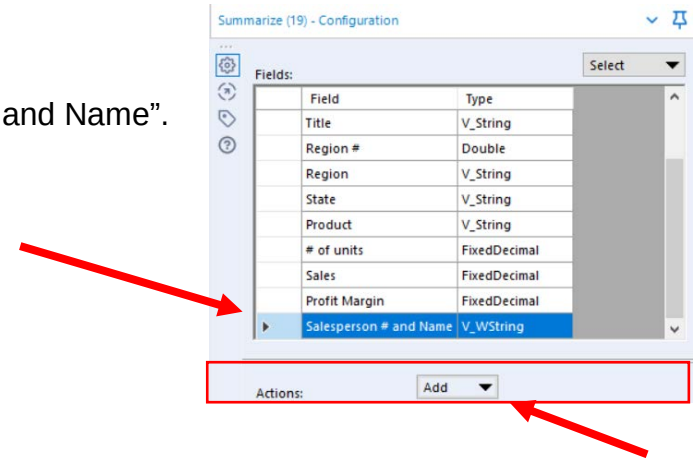
Record	Title	Region #	Region	State	Product	# of units	Sales	Profit Margin	Salesperson # and Name
1	Sales Associate I	3	South	TN	Stuffed Horse	144	3168.00	1908.00	1304 - Martin Van Buren
2	Senior Sales Associate	1	Midwest	NE	Stuffed Panda	72	1656.00	936.00	1302 - Grover Cleveland
3	Sales Associate I	2	Northeast	PA	Stuffed Panda	156	3588.00	2028.00	1301 - Thomas Jefferson
4	Sales Associate I	1	Midwest	MN	Stuffed Panda	120	2760.00	1560.00	1305 - George Washington
5	Senior Sales Associate	1	Midwest	MN	Stuffed Giraffe	108	2268.00	1215.00	1303 - Ulysses Grant
6	Sales Associate I	2	Northeast	RI	Stuffed Elephant	48	1200.00	708.00	1305 - George Washington
7	Senior Sales Associate	2	Northeast	ME	Stuffed Horse	96	2112.00	1272.00	1303 - Ulysses Grant
8	Sales Associate I	4	West	WY	Stuffed Penguin	156	3120.00	1716.00	1301 - Thomas Jefferson

Tool 8: Select Tool Summary: We now want to create a “Salesperson Report”.

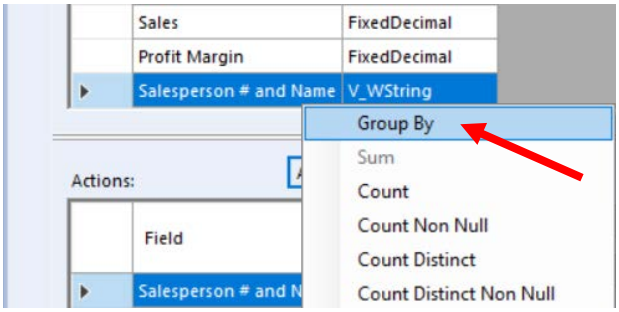
Drag a “**Summary Tool**” to the right of the “**Formula Tool**” on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



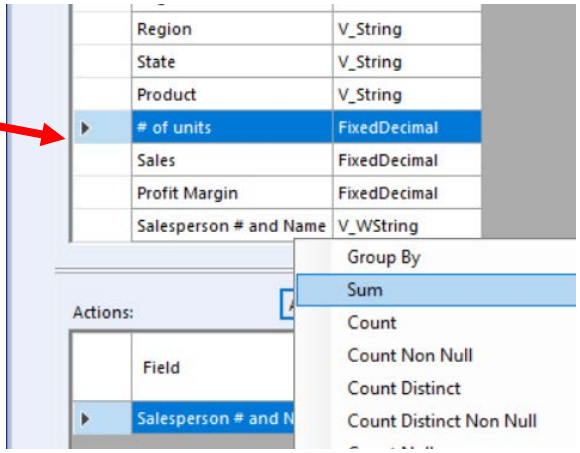
Select the first column, “Salesperson # and Name”.



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

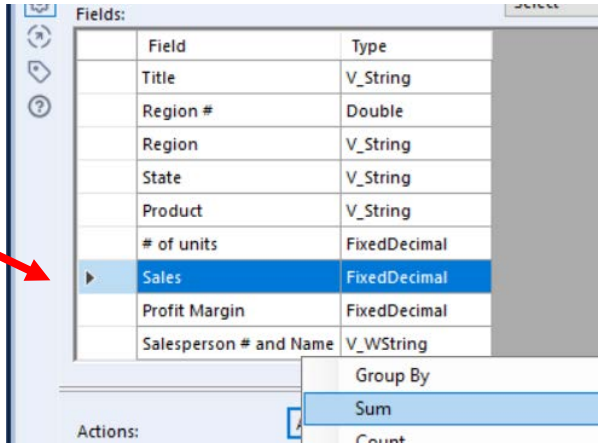


The second column will be “# of units”.



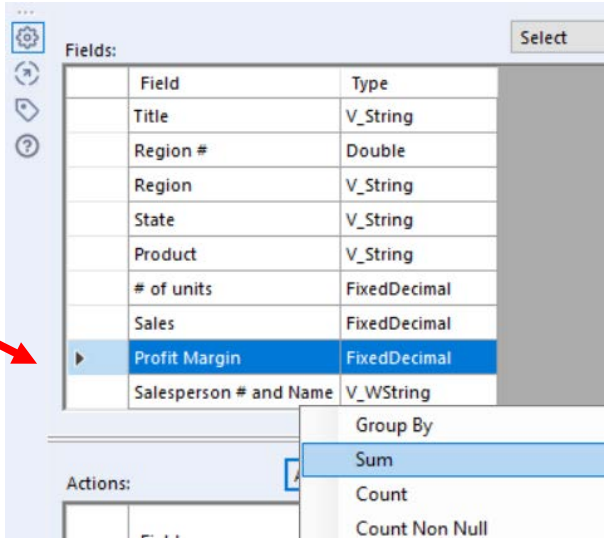
Click on the drop-down arrow on “Add” in “Actions” and select “Sum”.

The third column will be “Sales”.



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

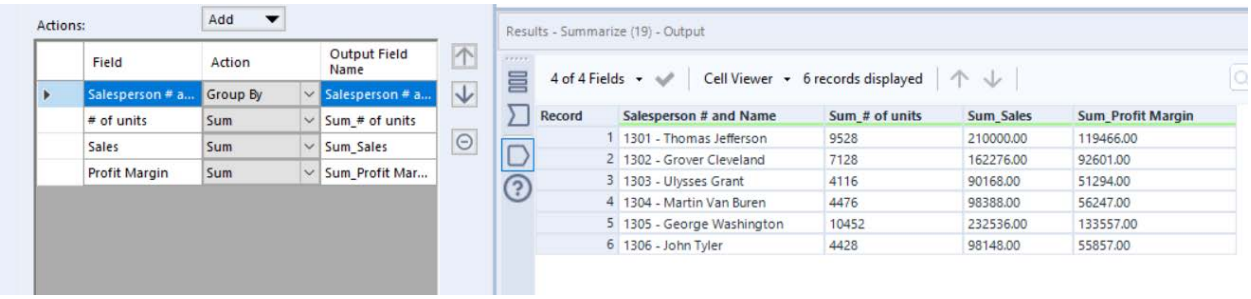
The fourth column will be “Profit Margin”.



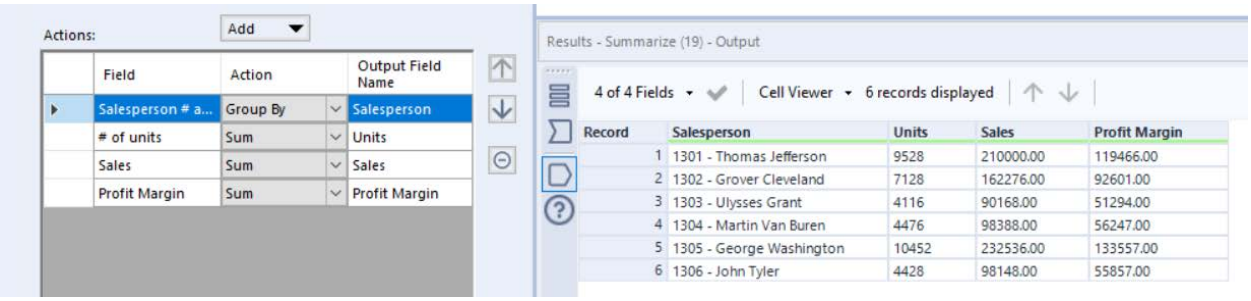
Click on the drop-down arrow on “Add” in “Actions” and select “Sum”.

Select “Run”.

And, the report is created, right click on the output arrow if nessacary.

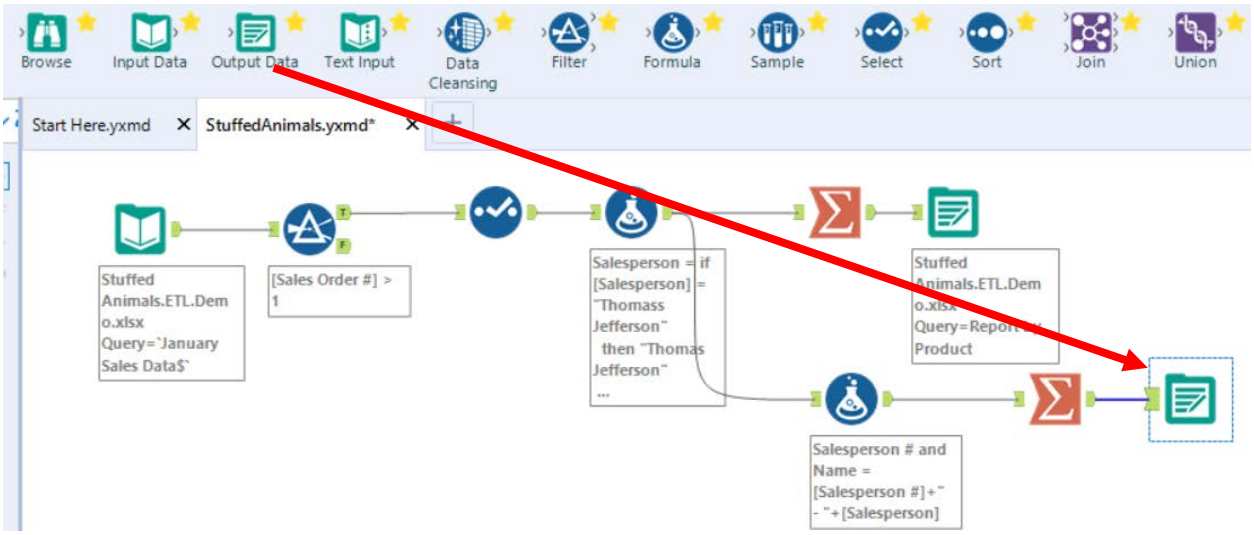


Change the “Output Field Names” and select “Run”

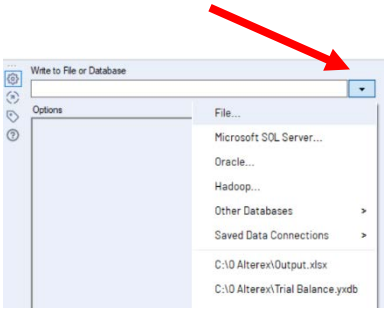


Tool 9: Select Output Data Tool: We now want to export the cleaned and assembled Salesperson Report to Excel.

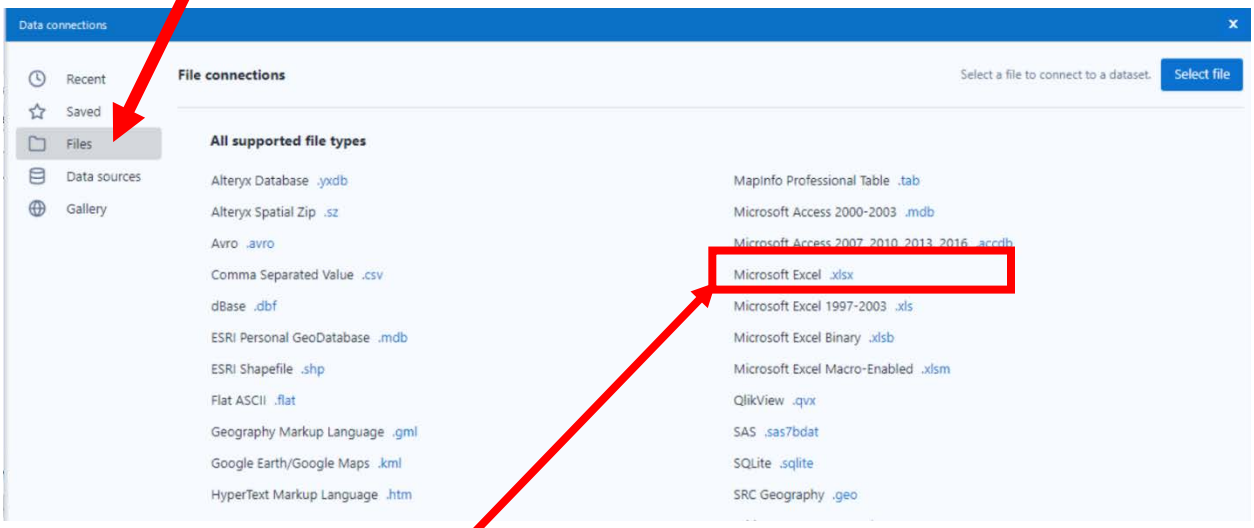
Drag an "**Output Data Tool**" to the right of the "**Summary Tool**" on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



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



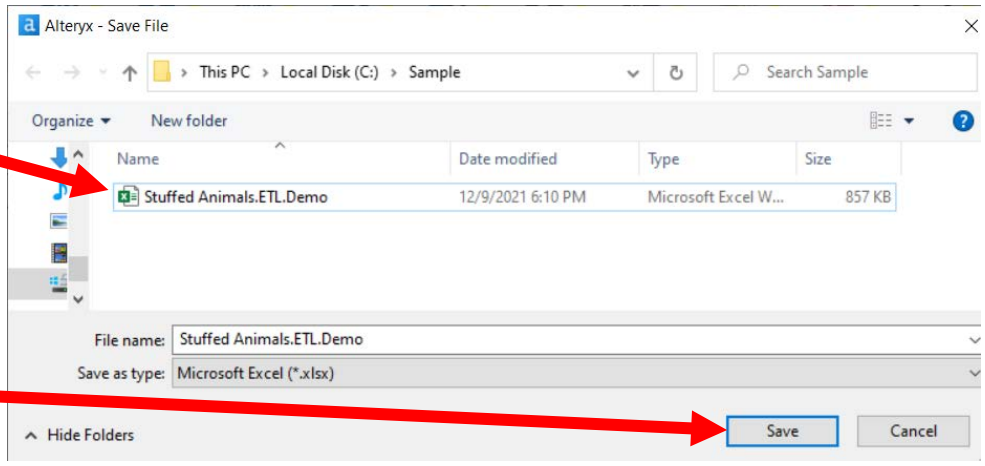
Select “File”.



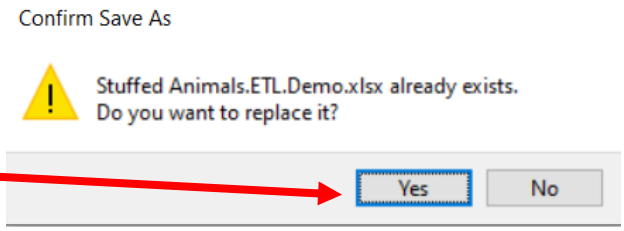
Select “Microsoft Excel .xlsx”

Select the required file name.

Select “Save”.

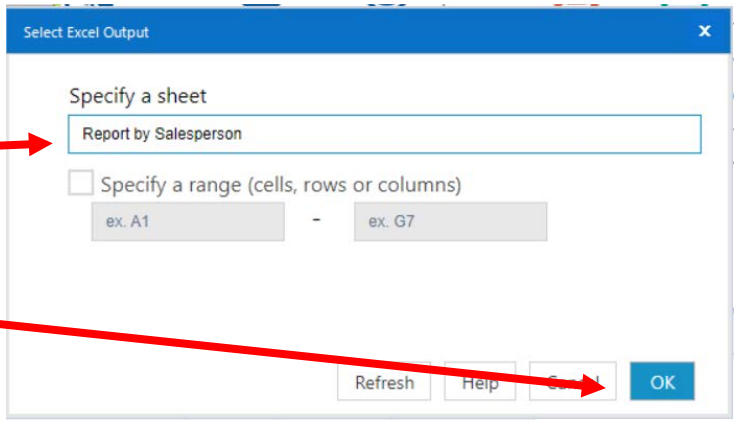


Click on “Yes” to confirm “Save”.



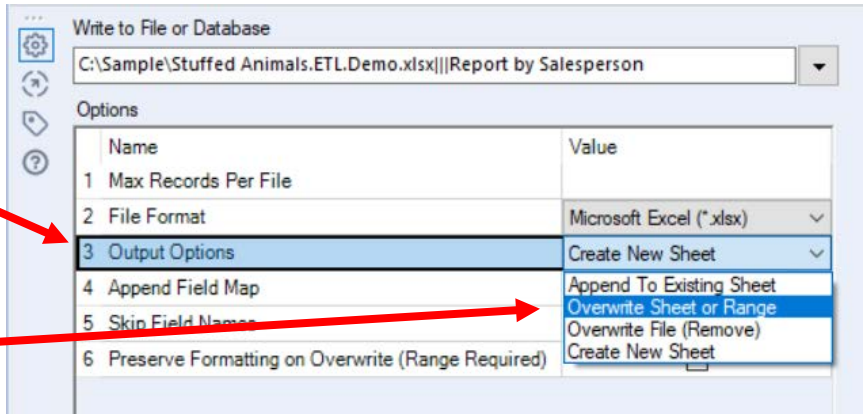
Enter the sheet name.

Select “OK”.

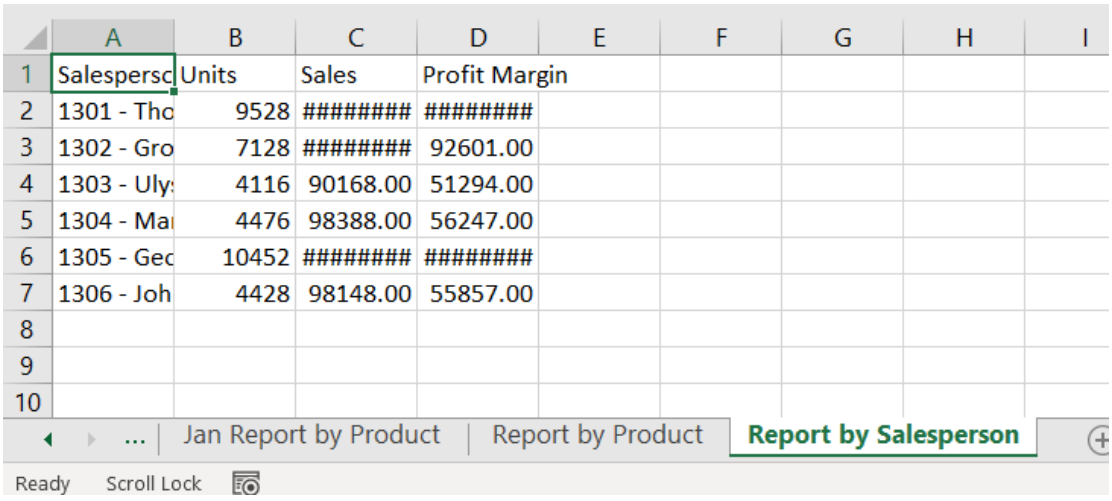


In the output Configuration area,
select “Output Options” number 3.

Select “Overwrite Sheet”
(which drops the old sheet)



Select Run
Open the Excel file



	A	B	C	D	E	F	G	H	I
1	Salespersc	Units	Sales	Profit Margin					
2	1301 - Tho	9528	########	########					
3	1302 - Gro	7128	########	92601.00					
4	1303 - Uly	4116	90168.00	51294.00					
5	1304 - Mai	4476	98388.00	56247.00					
6	1305 - Gec	10452	########	########					
7	1306 - Joh	4428	98148.00	55857.00					
8									
9									
10									

... | Jan Report by Product | Report by Product | **Report by Salesperson**

Ready Scroll Lock

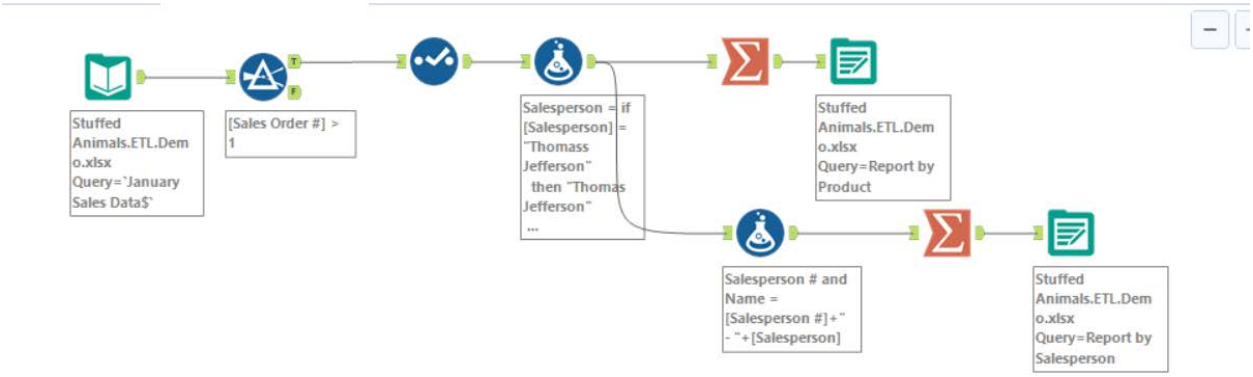
Change the format and rename the sheet so the format changes are not lost, save, close Excel and return to Alteryx.

	A	B	C	D	E	F	G
1	Salesperson	Units	Sales	Profit Margin			
2	1301 - Thomas Jefferson	9,528	$210,000.00	$119,466.00			
3	1302 - Grover Cleveland	7,128	$162,276.00	$ 92,601.00			
4	1303 - Ulysses Grant	4,116	$ 90,168.00	$ 51,294.00			
5	1304 - Martin Van Buren	4,476	$ 98,388.00	$ 56,247.00			
6	1305 - George Washington	10,452	$232,536.00	$133,557.00			
7	1306 - John Tyler	4,428	$ 98,148.00	$ 55,857.00			
8	Total	40,128	$891,516.00	$509,022.00			
9							
10							

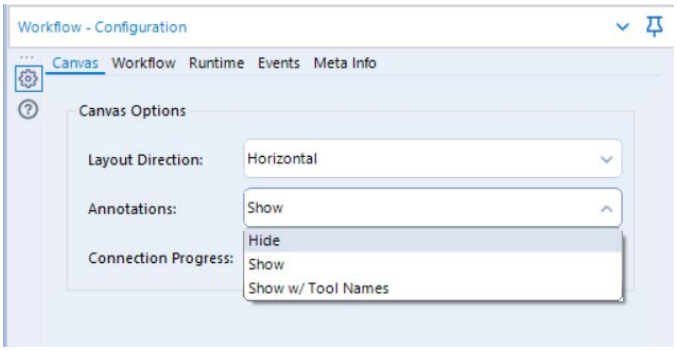
... | Jan Report by Product | Report by Product | **Jan Report by Salesperson**

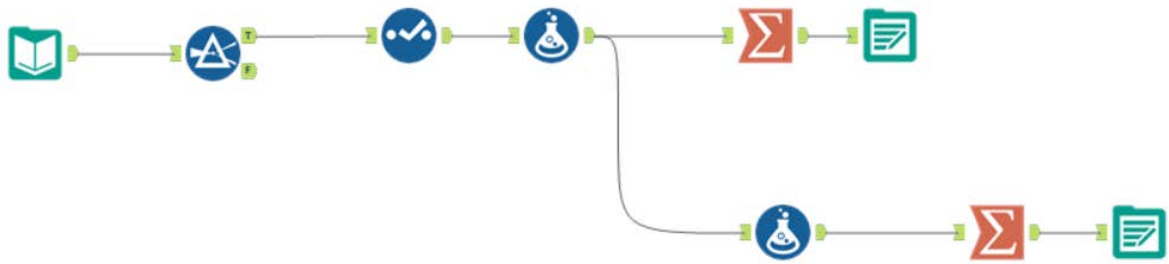
Ready Scroll Lock

The canvas is very busy looking because of the annotations.



Select any blank area on the canvas and in the Configuration section change “Annotations” to “Hide”.

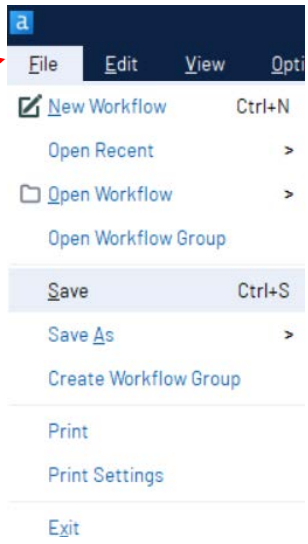




The workflow must be saved.

Select “File”.

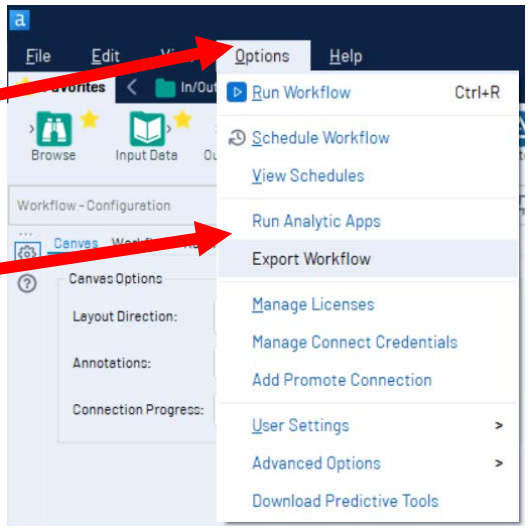
Select “Save”.



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

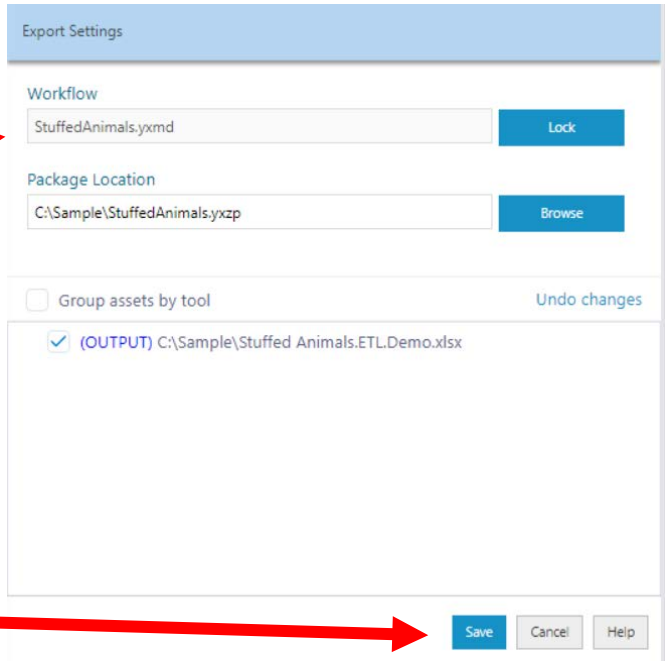
Select “Options”.

Select “Export Workflow”.

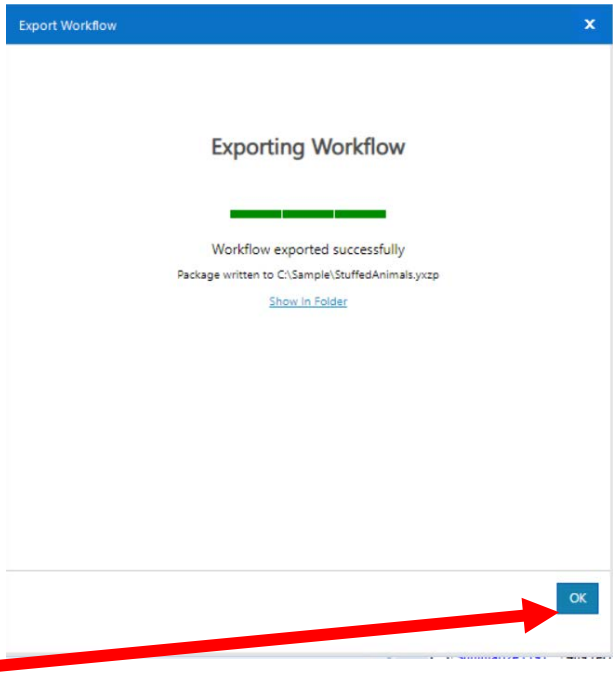


Verify the name of the “Workflow”.

Verify the “Package Location”.



Select “Save”.



Select “OK”.

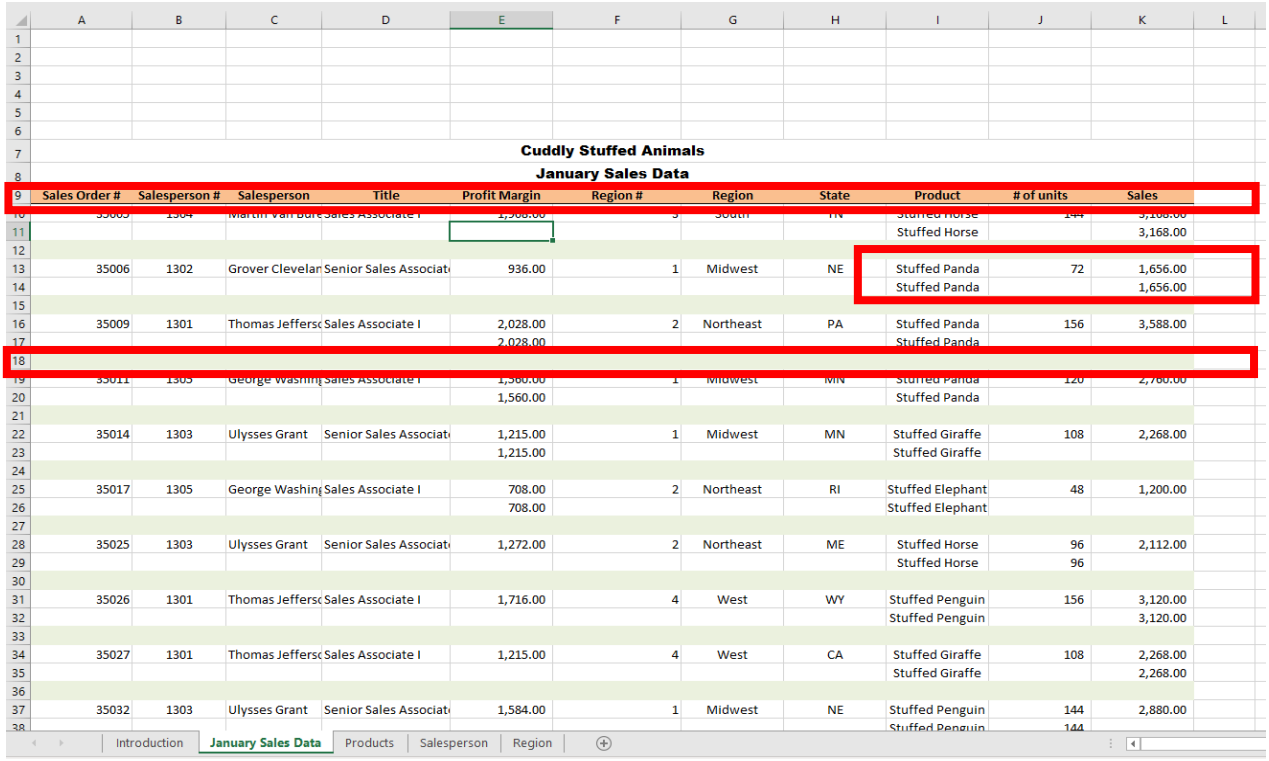
Close Alteryx

Part 2

Using Alteryx

Append a Second Stream

The sales data for *Cuddly Stuffed Animals* are stored in an Excel sheet titled January Sales Data. The column headings are in the ninth row, some rows have partial duplicate data and some rows are blank.



Cuddly Stuffed Animals
January Sales Data

Sales Order #	Salesperson #	Salesperson	Title	Profit Margin	Region #	Region	State	Product	# of units	Sales
								Stuffed Horse		3,168.00
35006	1302	Grover Clevelan	Senior Sales Associat	936.00	1	Midwest	NE	Stuffed Panda	72	1,656.00
								Stuffed Panda		1,656.00
35009	1301	Thomas Jefferso	Sales Associate I	2,028.00	2	Northeast	PA	Stuffed Panda	156	3,588.00
				2,028.00				Stuffed Panda		
				1,560.00				Stuffed Panda		
35014	1303	Ulysses Grant	Senior Sales Associat	1,215.00	1	Midwest	MN	Stuffed Giraffe	108	2,268.00
				1,215.00				Stuffed Giraffe		
35017	1305	George Washing	Sales Associate I	708.00	2	Northeast	RI	Stuffed Elephant	48	1,200.00
				708.00				Stuffed Elephant		
35025	1303	Ulysses Grant	Senior Sales Associat	1,272.00	2	Northeast	ME	Stuffed Horse	96	2,112.00
								Stuffed Horse	96	
35026	1301	Thomas Jefferso	Sales Associate I	1,716.00	4	West	WY	Stuffed Penguin	156	3,120.00
								Stuffed Penguin		3,120.00
35027	1301	Thomas Jefferso	Sales Associate I	1,215.00	4	West	CA	Stuffed Giraffe	108	2,268.00
								Stuffed Giraffe		2,268.00
35032	1303	Ulysses Grant	Senior Sales Associat	1,584.00	1	Midwest	NE	Stuffed Penguin	144	2,880.00

Introduction | January Sales Data | Products | Salesperson | Region

Sample Reports

Product	Units	Sales	Profit Margin
Stuffed Elephant	4,392	$109,800.00	$ 64,782.00
Stuffed Giraffe	8,052	$169,092.00	$ 90,585.00
Stuffed Horse	12,120	$266,640.00	$160,590.00
Stuffed Panda	7,488	$172,224.00	$ 97,344.00
Stuffed Penguin	5,016	$100,320.00	$ 55,176.00
Stuffed Unicorn	3,060	$ 73,440.00	$ 40,545.00
Total	40,128	$891,516.00	$509,022.00

Salesperson	Units	Sales	Profit Margin
1301 - Thomas Jefferson	9,528	$210,000.00	$119,466.00
1302 - Grover Cleveland	7,128	$162,276.00	$ 92,601.00
1303 - Ulysses Grant	4,116	$ 90,168.00	$ 51,294.00
1304 - Martin Van Buren	4,476	$ 98,388.00	$ 56,247.00
1305 - George Washington	10,452	$232,536.00	$133,557.00
1306 - John Tyler	4,428	$ 98,148.00	$ 55,857.00
Total	40,128	$891,516.00	$509,022.00

Start Alteryx

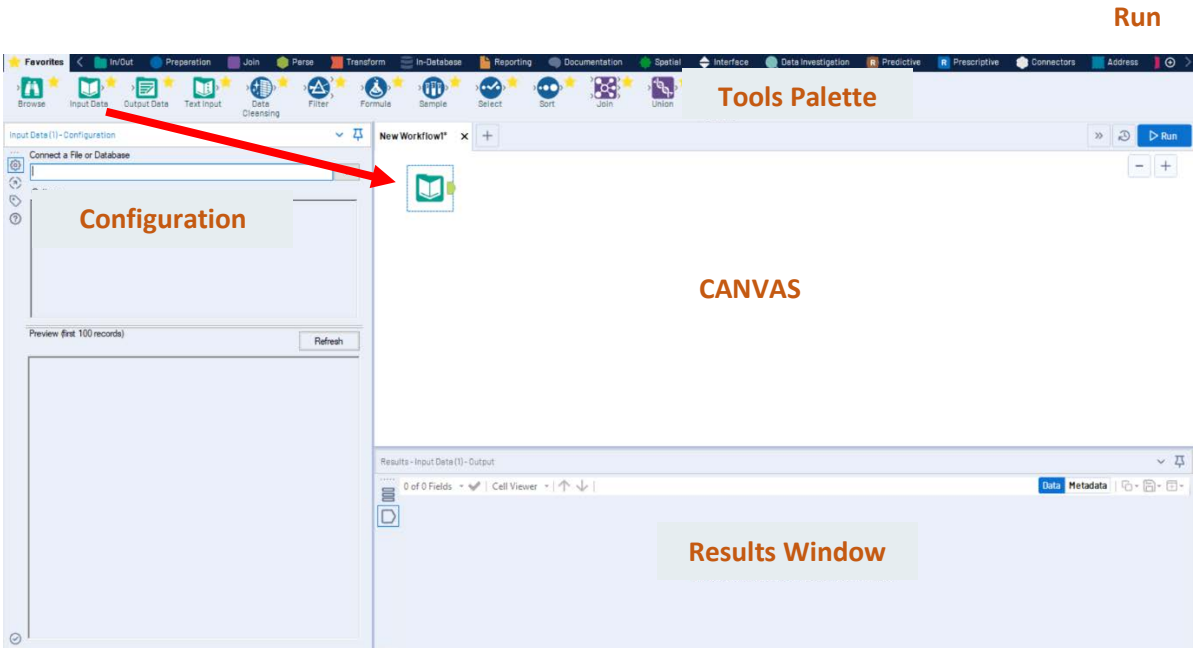
Note: You cannot extract data from an open Excel file so make sure that the Excel file is closed before starting to input the data.

There are four major parts of the Alteryx window. The first is the Canvas, and that's where we're going to create the workflow. 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 us the result after we select "Run" to apply the changes.

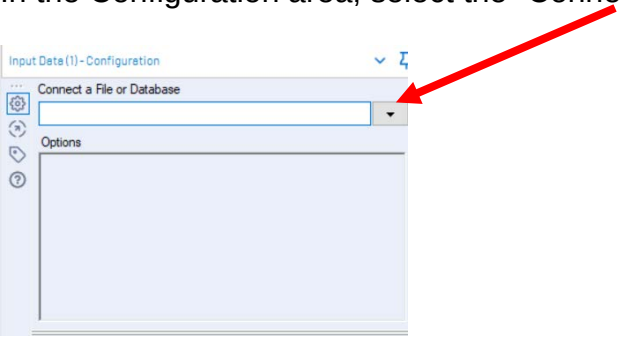
Tool 1: Input Data Tool – Input Excel file

Drag the "Input Data Tool"

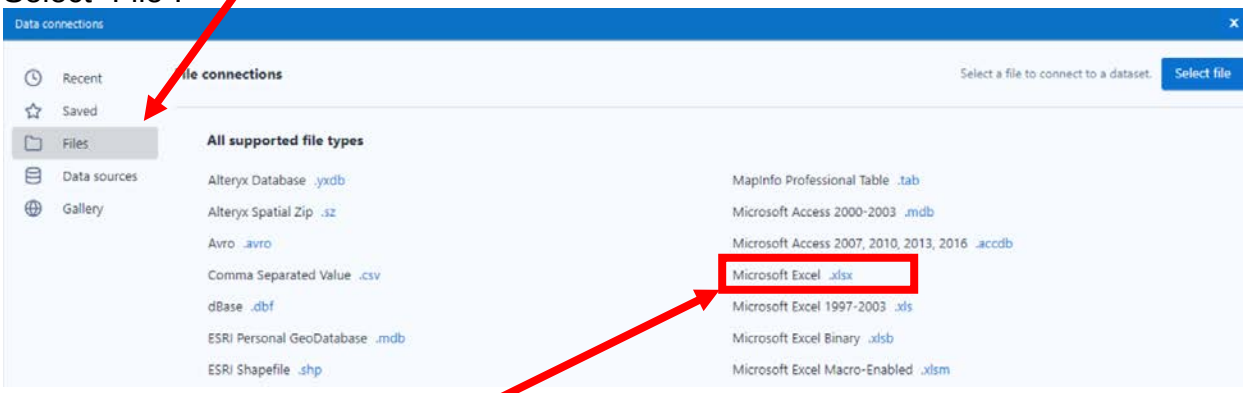
from the Tools Palette to the Canvas.



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



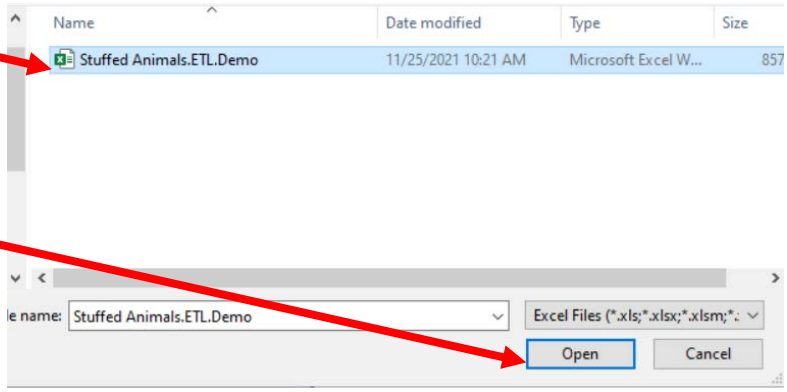
Select “File”.



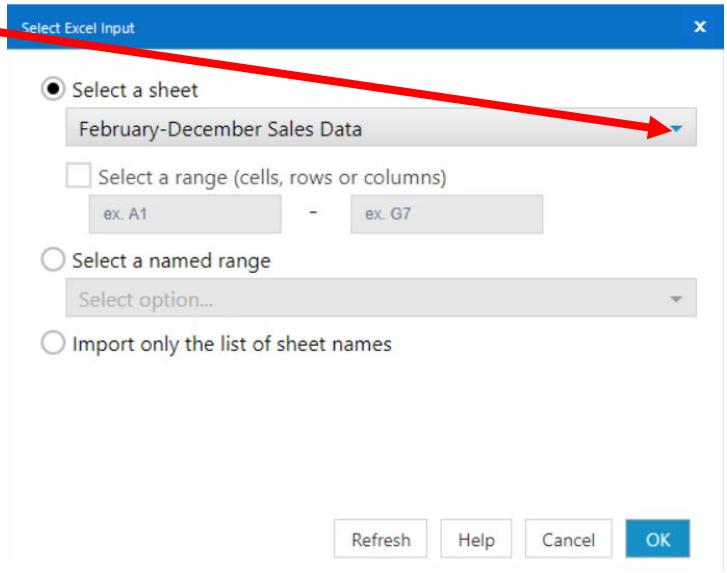
Select “Microsoft Excel .xlsx”.

Browse for the desired file.

Select “Open”.

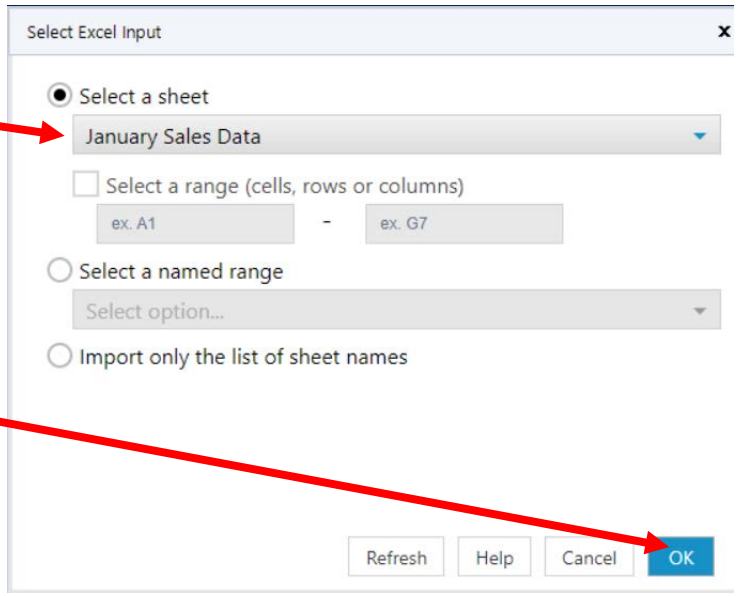


Select the drop-down.



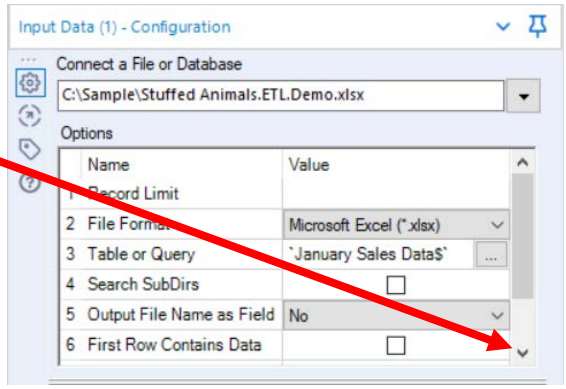
Select “January Sales Data”.

Select “OK”.

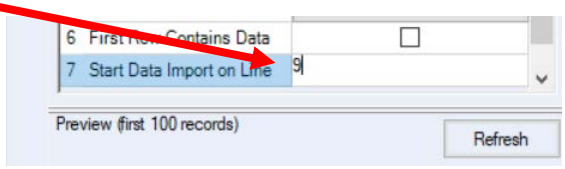


Remember that the field names are in the nineth row.

In the Configuration area, scroll down to option 7.

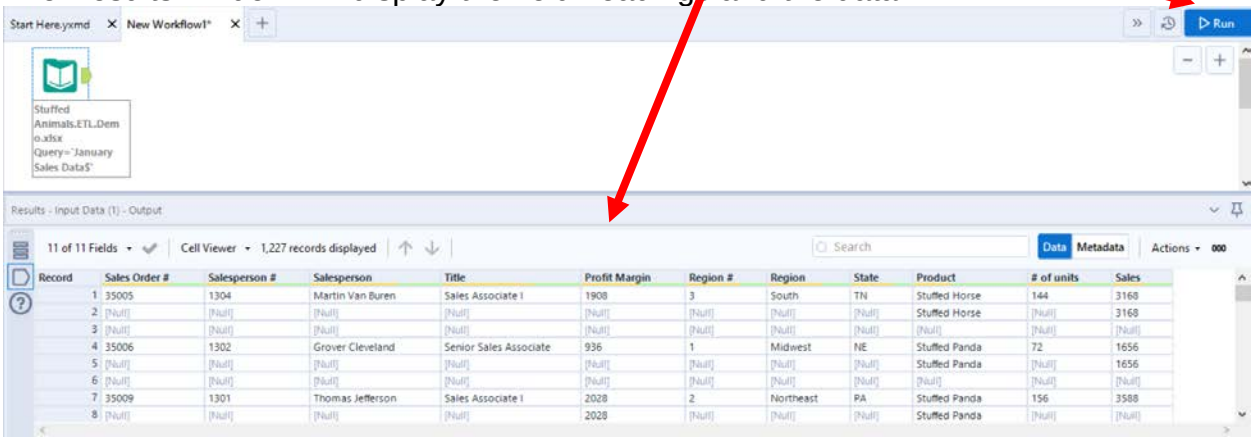


Change the number in “Start Data Import on Line” to “9”.



Select “Run”.

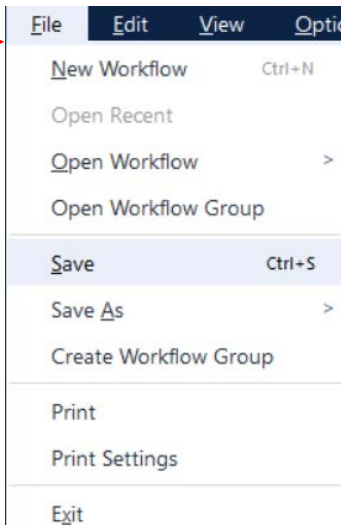
The Results window will display the field headings and the data.



The workflow must be saved.

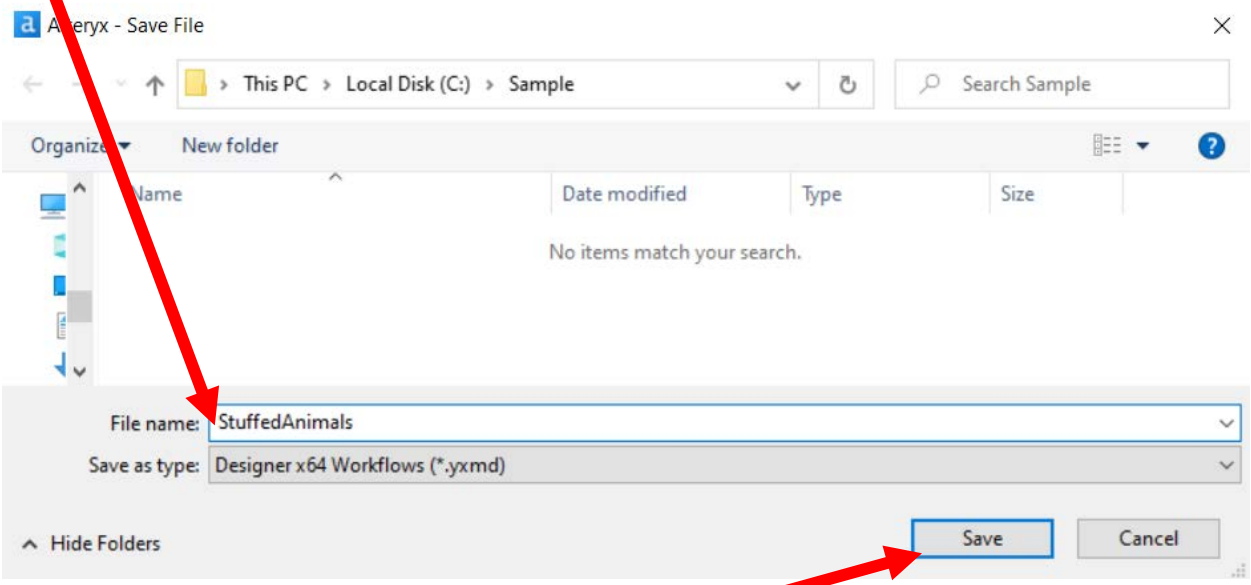
Select “File”.

Select “Save”.



Browse to the required directory.

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

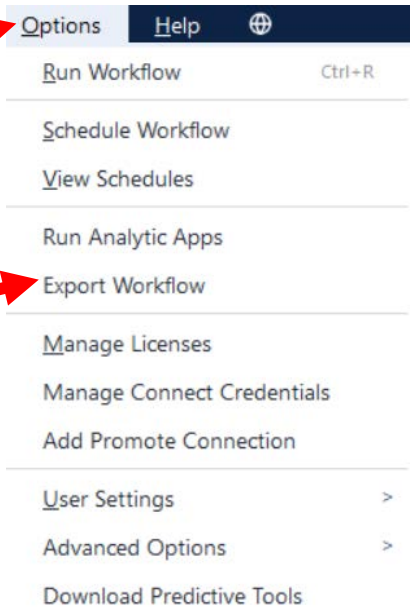


Select “Save”.

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”.

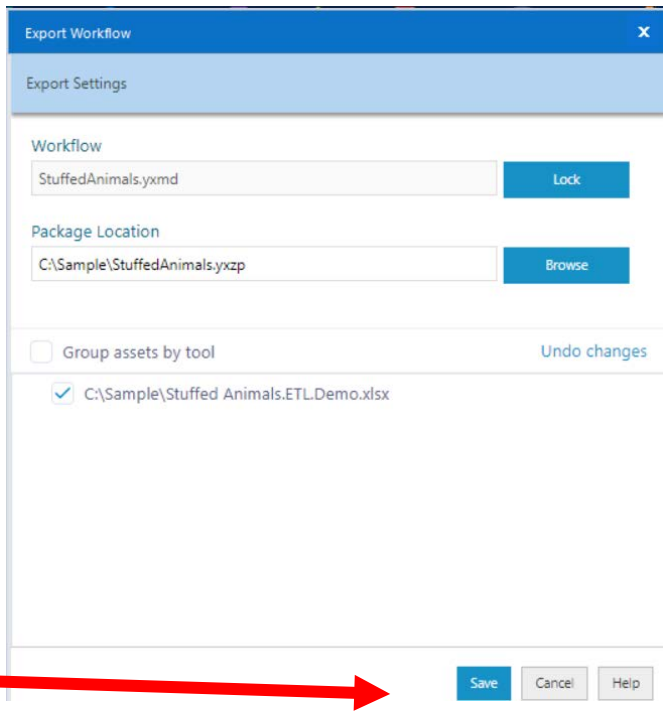
Select “Export Workflow”.



Verify the “Workflow”.

Verify the “Package Location”.

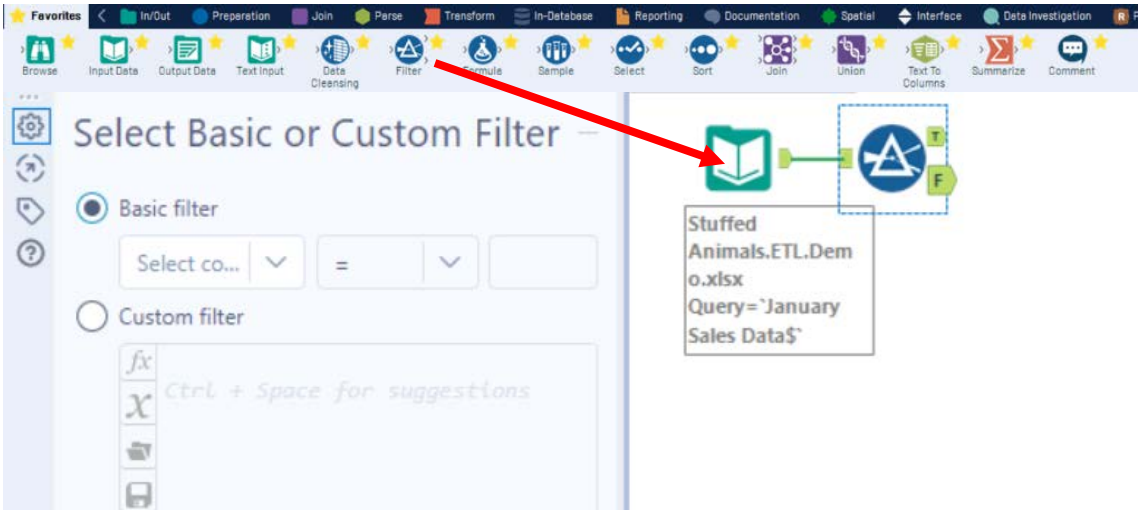
Select “Save”.



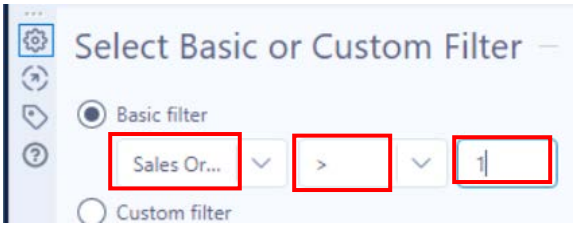
Tool 2: Filter Tool – Exclude all rows that do not contain data

There are several rows that do not contain an sales order number. We can modify the working dataset so that it includes only those rows that contain account numbers greater than one. Note that we are not deleting anything from the source data file. The working dataset will be split into two subsets. The first subset, “True”, will include all rows with account numbers greater than one. The second subset, “False”, will contain all rows with account numbers less than one.

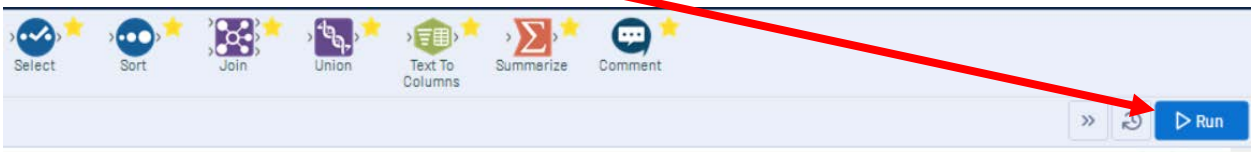
Drag the “**Filter Tool**” from the Tools Palette to the right of the “Input Data Tool” on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



With the “**Filter Tool**” selected, the “Basic” filter in the Configuration area can be modified to filter the field “Sales Order” for those accounts greater than “1’.



To apply the latest change, select “Run”.

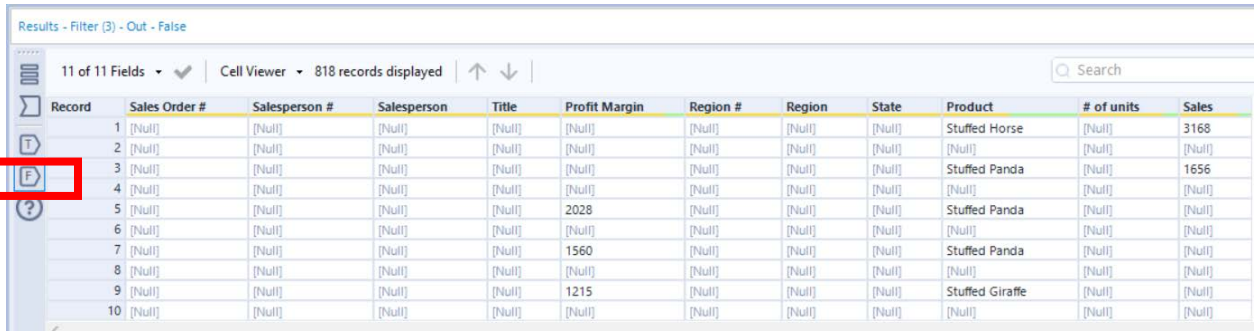


We can view three different datasets in the Results window.

The records that meet the filtered conditions, “Sales Order > 1”, are displayed as T (True).

Record	Sales Order #	Salesperson #	Salesperson	Title	Profit Margin	Region #	Region	State	Product	# of units	Sales
1	35005	1304	Martin Van Buren	Sales Associate I	1908	3	South	TN	Stuffed Horse	144	3168
2	35006	1302	Grover Cleveland	Senior Sales Associate	936	1	Midwest	NE	Stuffed Panda	72	1656
3	35009	1301	Thomas Jefferson	Sales Associate I	2028	2	Northeast	PA	Stuffed Panda	156	3588
4	35011	1305	George Washington	Sales Associate I	1560	1	Midwest	MN	Stuffed Panda	120	2760
5	35014	1303	Ulysses Grant	Senior Sales Associate	1215	1	Midwest	MN	Stuffed Giraffe	108	2268
6	35017	1305	George Washington	Sales Associate I	708	2	Northeast	RI	Stuffed Elephant	48	1200
7	35025	1303	Ulysses Grant	Senior Sales Associate	1272	2	Northeast	ME	Stuffed Horse	96	2112
8	35026	1301	Thomas Jefferson	Sales Associate I	1716	4	West	WY	Stuffed Penguin	156	3120
9	35027	1301	Thomas Jefferson	Sales Associate I	1215	4	West	CA	Stuffed Giraffe	108	2268
10	35032	1303	Ulysses Grant	Senior Sales Associate	1584	1	Midwest	NE	Stuffed Penguin	144	2880

The records that do not meet the filtered conditions are displayed as F (False).



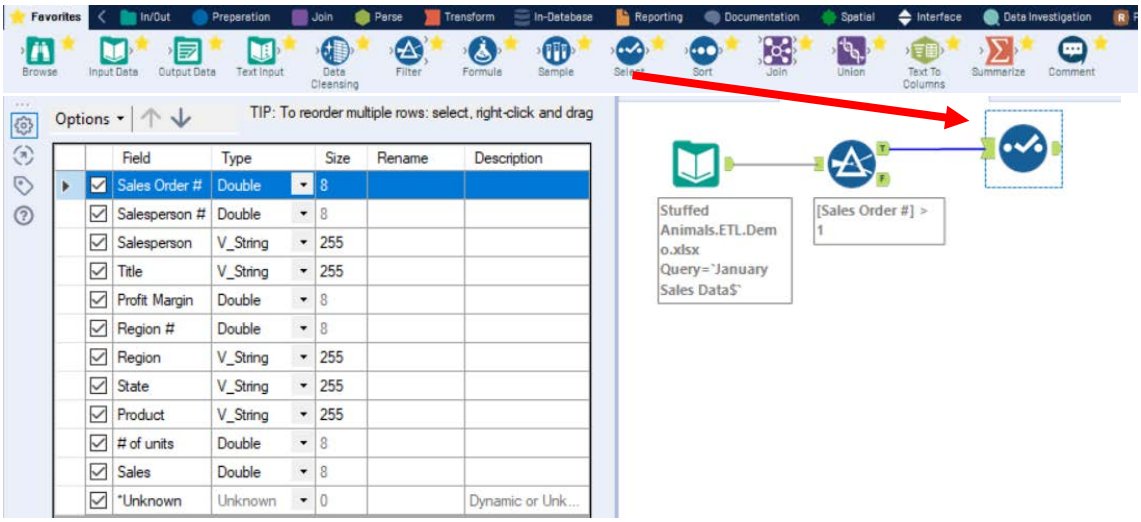
Record	Sales Order #	Salesperson #	Salesperson	Title	Profit Margin	Region #	Region	State	Product	# of units	Sales
1	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	Stuffed Horse	[Null]	3168
2	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
3	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	Stuffed Panda	[Null]	1656
4	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
5	[Null]	[Null]	[Null]	[Null]	2028	[Null]	[Null]	[Null]	Stuffed Panda	[Null]	[Null]
6	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
7	[Null]	[Null]	[Null]	[Null]	1560	[Null]	[Null]	[Null]	Stuffed Panda	[Null]	[Null]
8	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
9	[Null]	[Null]	[Null]	[Null]	1215	[Null]	[Null]	[Null]	Stuffed Giraffe	[Null]	[Null]
10	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]

The data being filtered can also be displayed.

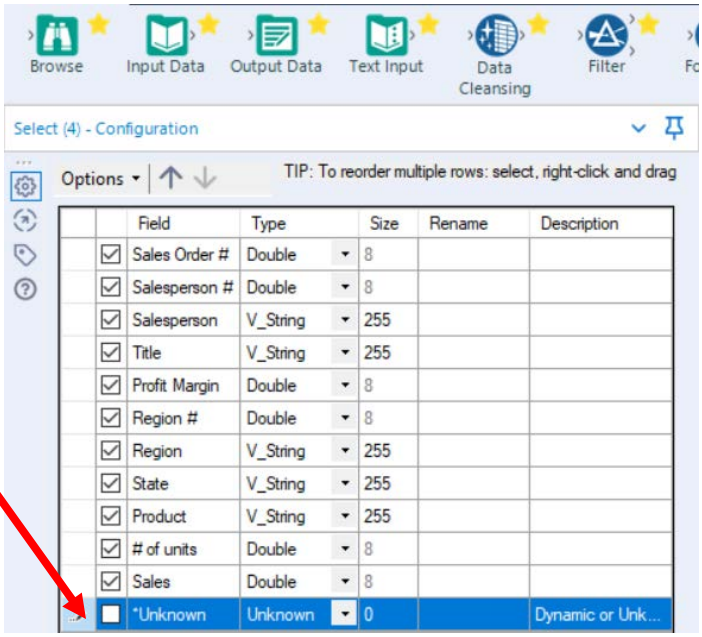
Record	Sales Order #	Salesperson #	Salesperson	Title	Profit Margin	Region #	Region	State	Product	# of units	Sales
1	35005	1304	Martin Van Buren	Sales Associate I	1908	3	South	TN	Stuffed Horse	144	3168
2	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	Stuffed Horse	[Null]	3168
3	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
4	35006	1302	Grover Cleveland	Senior Sales Associate	936	1	Midwest	NE	Stuffed Panda	72	1656
5	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	Stuffed Panda	[Null]	1656
6	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
7	35009	1301	Thomas Jefferson	Sales Associate I	2028	2	Northeast	PA	Stuffed Panda	156	3588
8	[Null]	[Null]	[Null]	[Null]	2028	[Null]	[Null]	[Null]	Stuffed Panda	[Null]	[Null]
9	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]	[Null]
10	35011	1305	George Washington	Sales Associate I	1560	1	Midwest	MN	Stuffed Panda	120	2760

Tool 3: Select Tool – Remove fields that contain no data, from "*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.

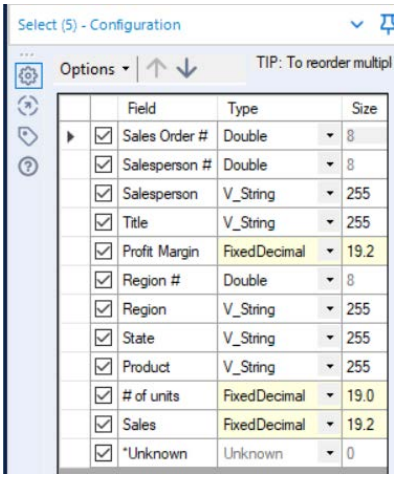


Remove any extra fields by unchecking them.

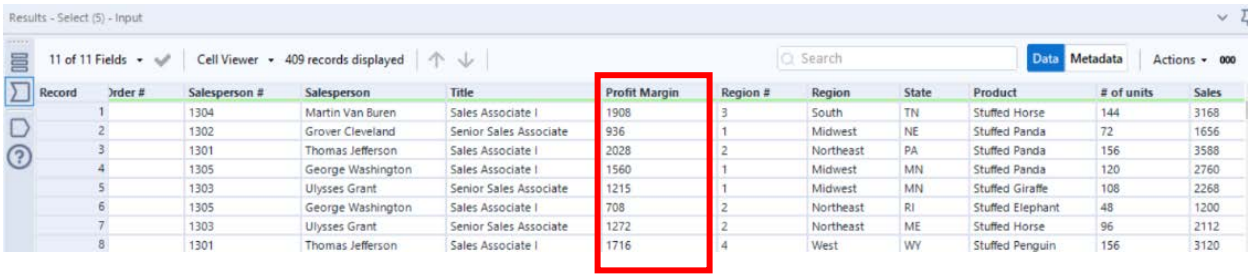


Change the type of all numeric fields that will be used in calculations to “FixedDecimal”. The “# of units should have zero decimal places and dollar values should have two decimal places.

.

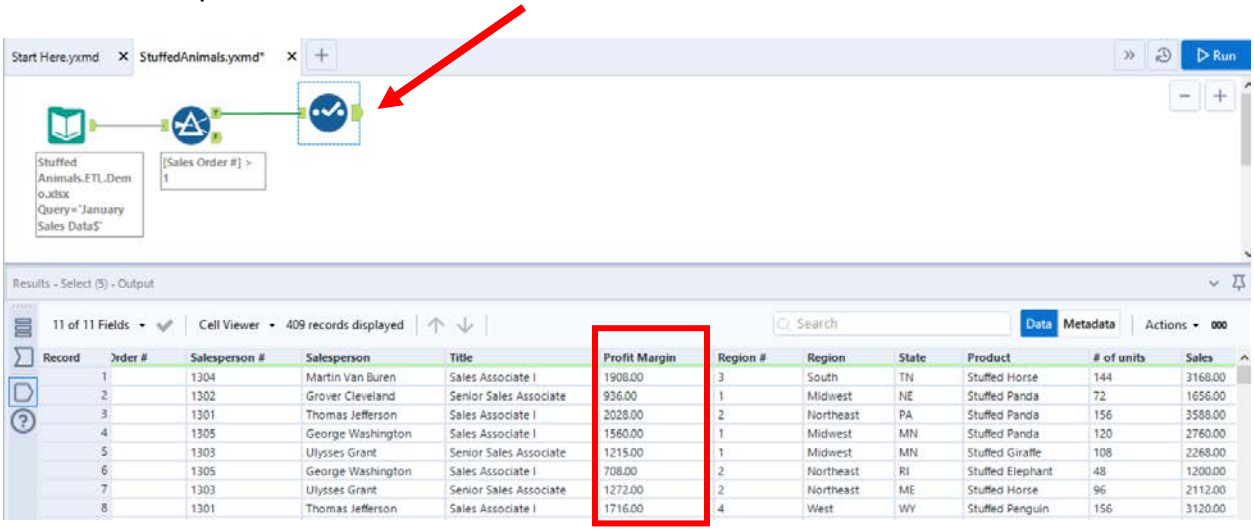


Before making changes:



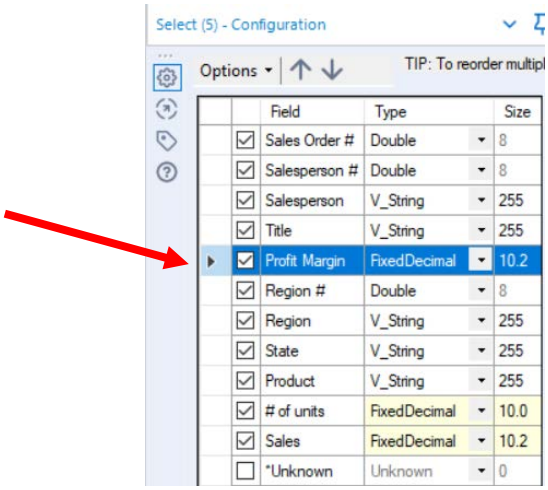
Select Run.

Select the output area of the Select tool.

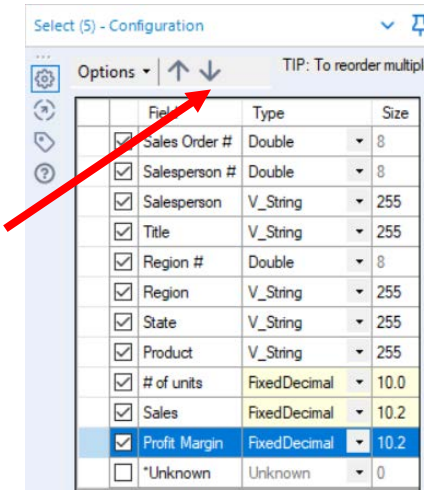


Change the order of the fields by moving “Profit Margin” to the end.

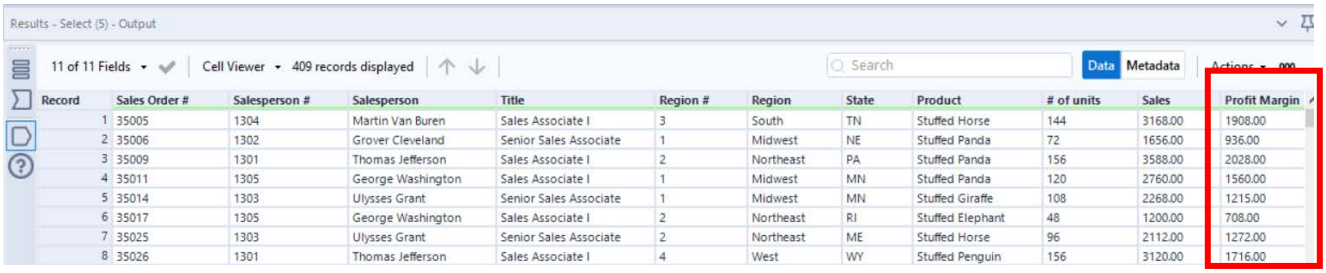
In the Configuration section select “Profit Margin”.



Use the down arrow to move “Profit Margin” to the end of the list.



Select Run.



Temporary Tool 4: Browse Tool – The Data in the Fields Must Conform to Predefined Rules.

In the original Excel file, The Cuddly Stuffed Animals company provided the only acceptable values for "Product", "Salesperson" and "Region". Sometime data is entered incorrectly into the data.

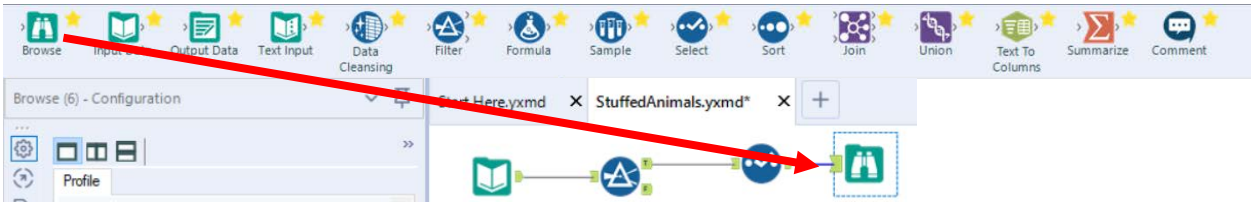
Product ID	Product
101	Stuffed Penguin
102	Stuffed Horse
103	Stuffed Unicorn
104	Stuffed Giraffe
105	Stuffed Panda
106	Stuffed Elephant

Code	Salesperson
1301	Thomas Jefferson
1302	Grover Cleveland
1303	Ulysses Grant
1304	Martin Van Buren
1305	George Washington
1306	John Tyler

Code	Region
1	Midwest
2	Northeast
3	South
4	West

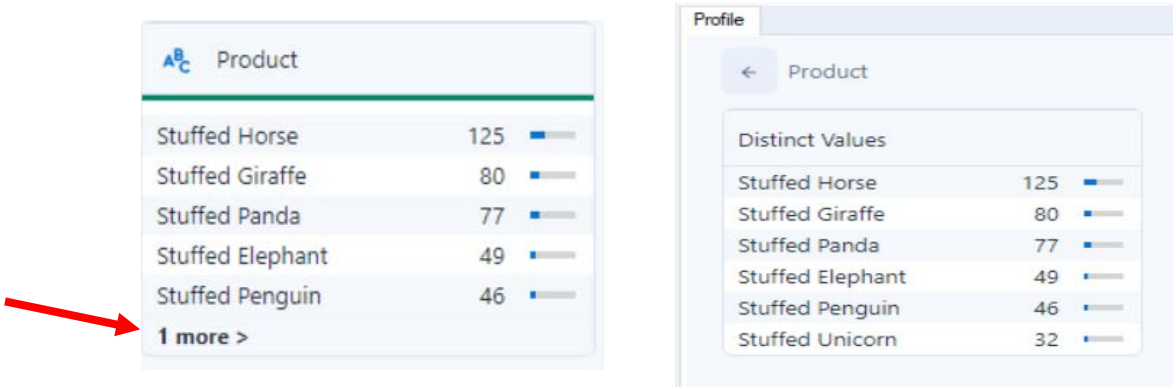
The "Product" field can only contain one of the six products. Browse the data in the "Product" field to verify that all the data is correct. Make any changes that are required.

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



Select Run.

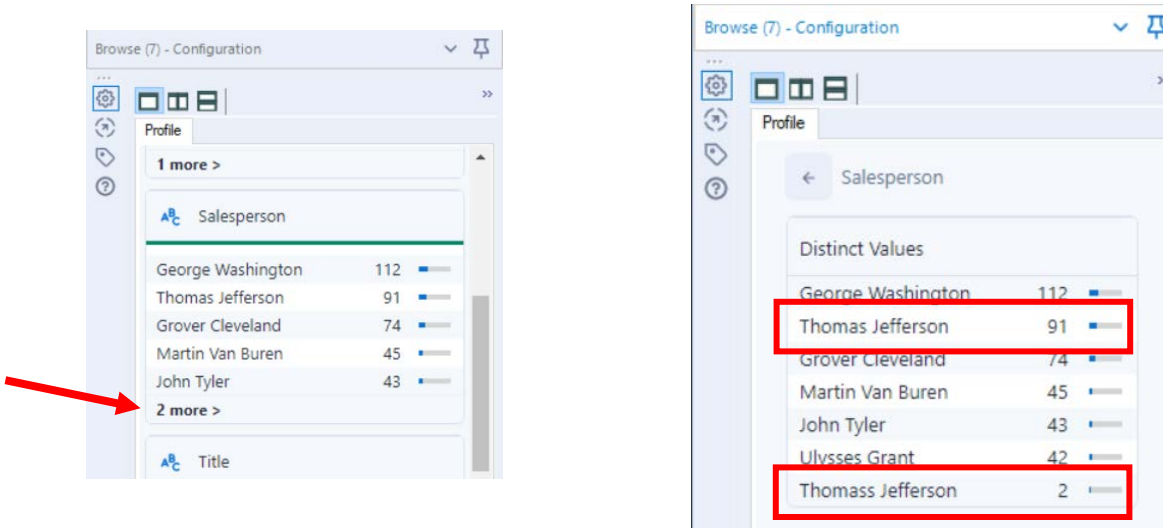
In the Configuration section scroll down to “Product” and select “1 more”.



The data is clean, there are only six products listed.

The “Salesperson” field can only contain one of the six names. Browse the data in the “Salesperson” field to verify that all the data is correct. Make any changes that are required.

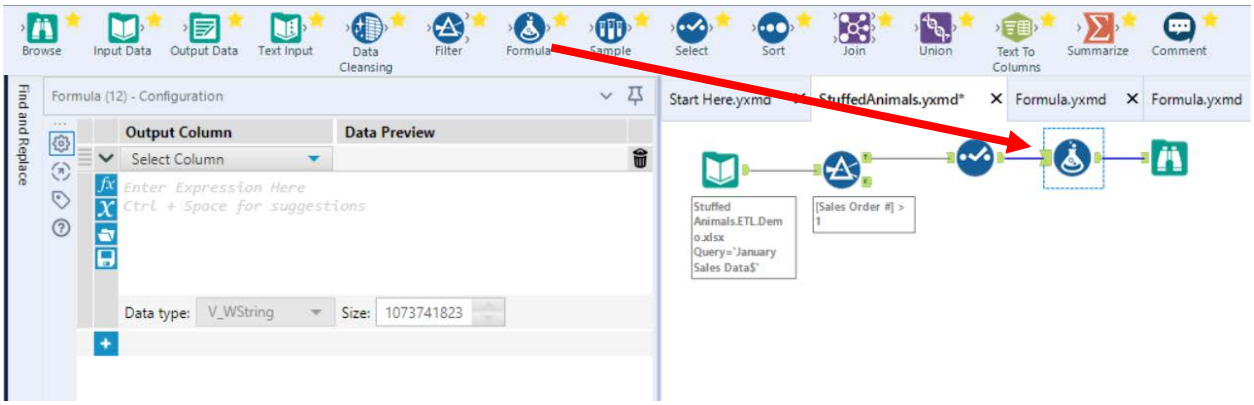
In the Configuration section scroll down to “Salesperson” and select “2 more”.



“Thomass Jefferson” must be replaced with “Thomas Jefferson” for a total of 93 distinct values.

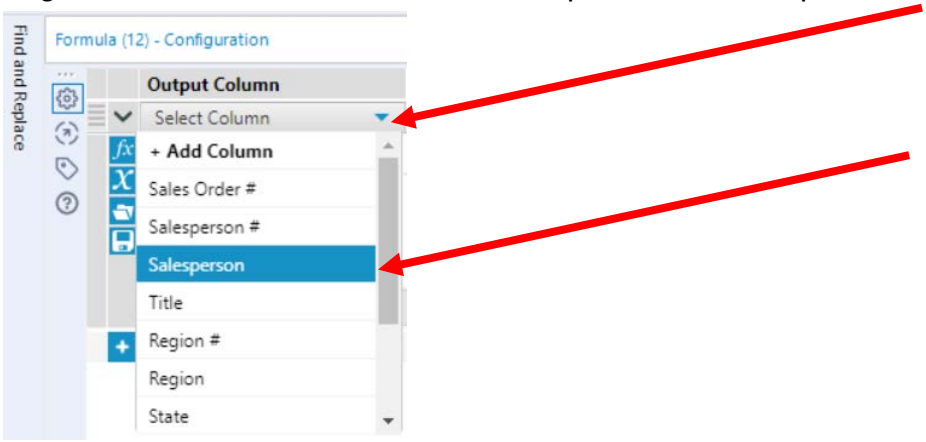
Tool 4: Formula Tool – The Data in the Fields Must Conform to Predefined Rules.

Drag the "**Formula Tool**" from the Tools Palette between the output of the "**Select Tool**" and the "Browse Tool" on the Canvas. The three will be connected automatically.

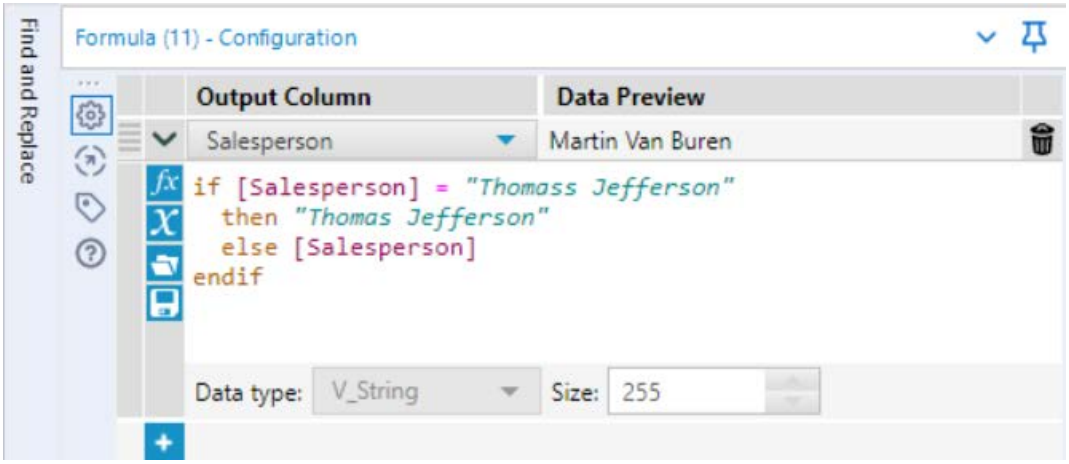


If the salesperson's name is "Thomass Jefferson" replace it with "Thomas Jefferson" otherwise leave the existing name in the field. Do not create a new field, put the output back into the "Salesperson" field.

In the configuration section select the field Salesperson from the pull-down list.

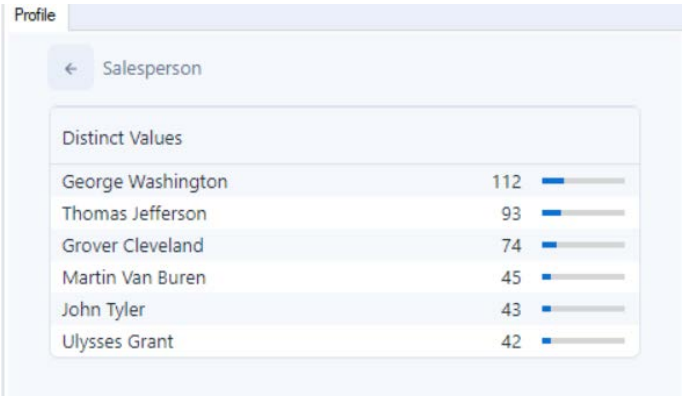


Type in the formula.



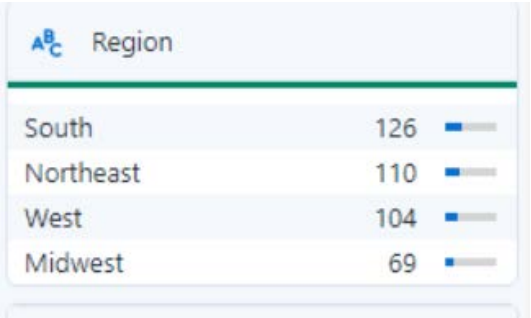
Select the Browse step.

Select Run.



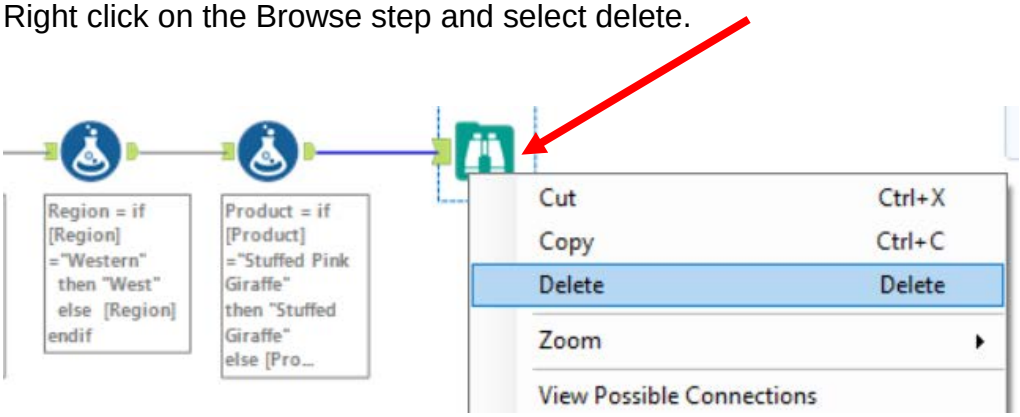
In the Configuration section scroll down to “Salesperson”. There are only six salespersons and Thomas Jefferson has 93 distinct values.

While in the Browse step, verify the field “Region”



The data is clean, only the four regions are listed.

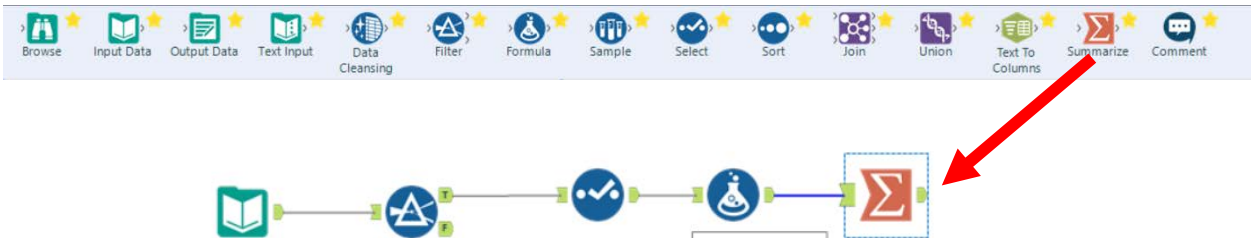
Right click on the Browse step and select delete.



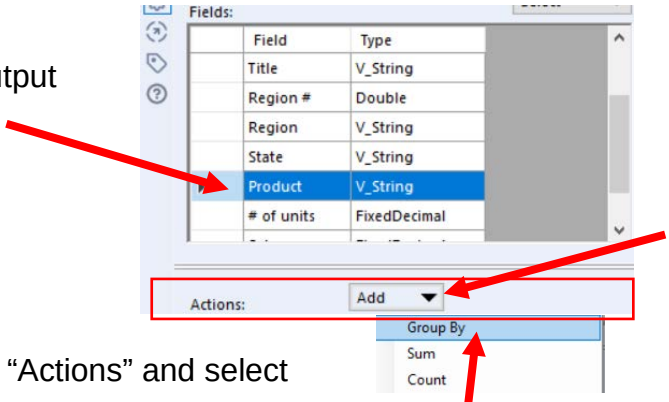
Tool 5: Select Tool Summary: We now want to create a report by product by first summarizing on "# of units" by product.

Product	# of units	Sales	Profit Margin
Stuffed Frog	1,524	$ 33,528	$20,574
Stuffed Unicorr	408	8,160	4,692
Stuffed Panda	1,140	26,220	15,675
Stuffed Koala	624	14,976	8,112
Stuffed Elephai	4,044	84,924	49,539
Stuffed Bear	2,376	59,400	35,046
Total	10,116	$227,208	$133,638

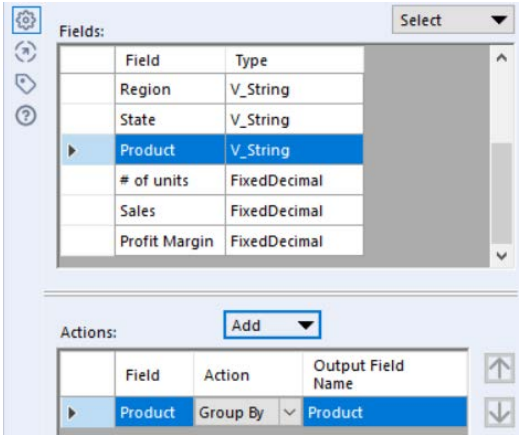
Drag a "**Summary Tool**" to the right of the "**Formula tool**" on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



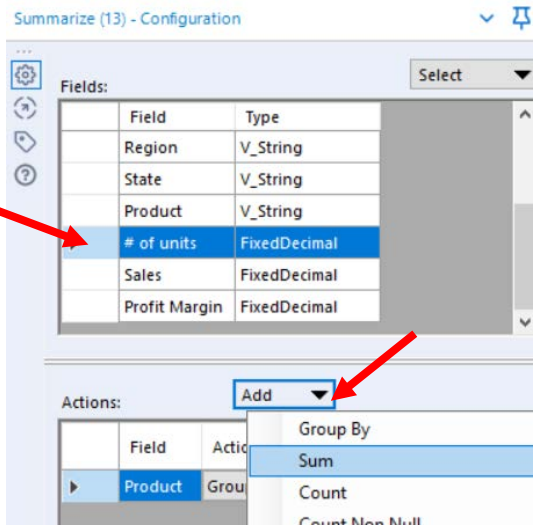
The first column will be Product. In the output Configuration area, select "Product".



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

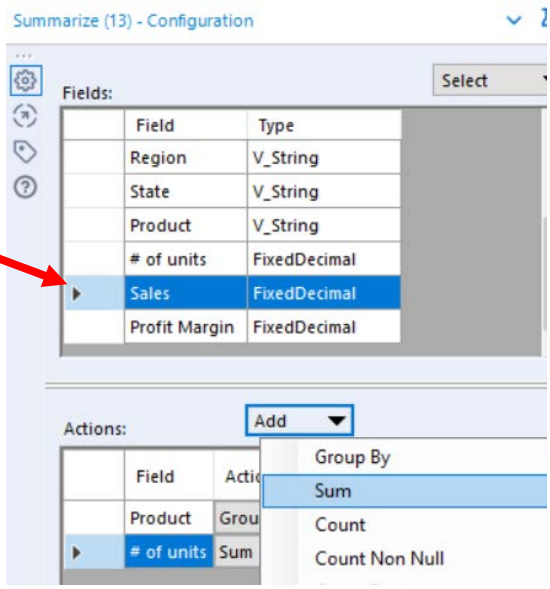


The second column will be “# of units”.



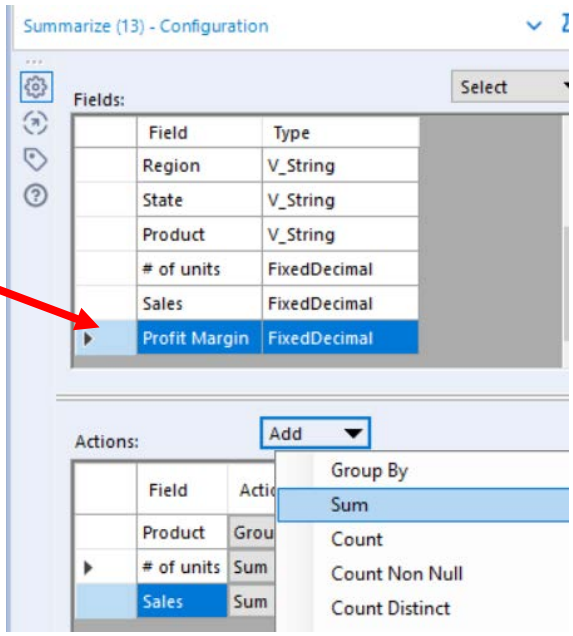
Click on the drop-down arrow on ”Add” in the “Actions” and select “Sum”.

The third column will be “Sales”.



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

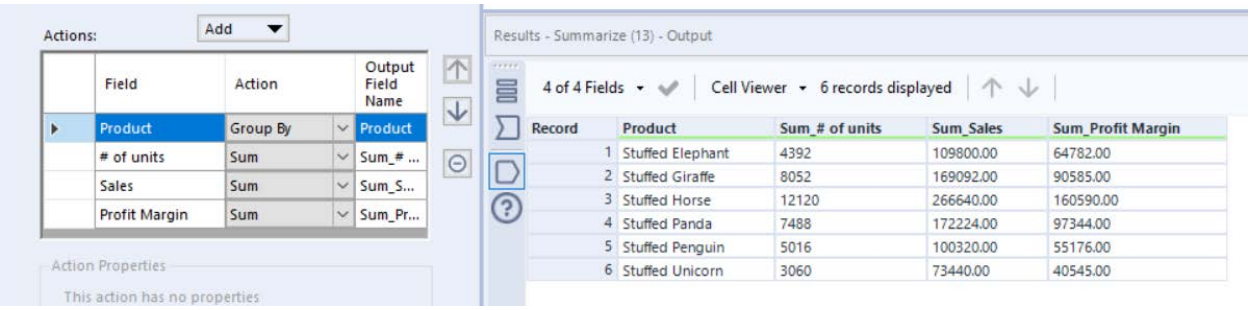
The fourth column will be “Profit Margin”.



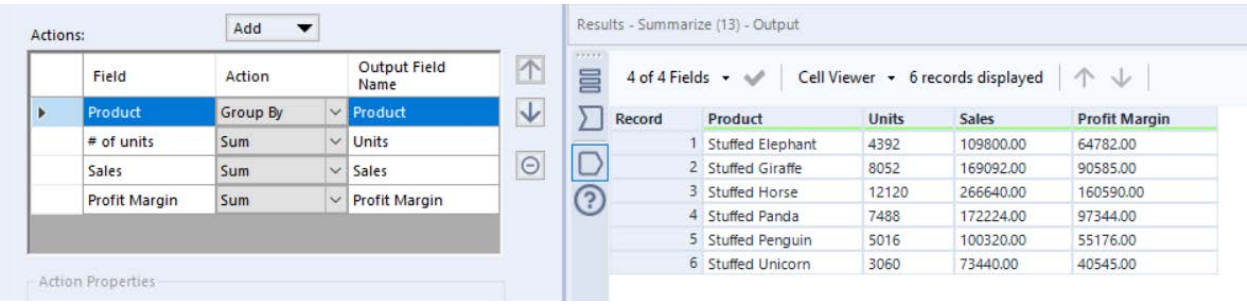
Click on the drop-down arrow on ”Add” in the “Actions” and select “Sum”.

Select “Run”.

The report is created.

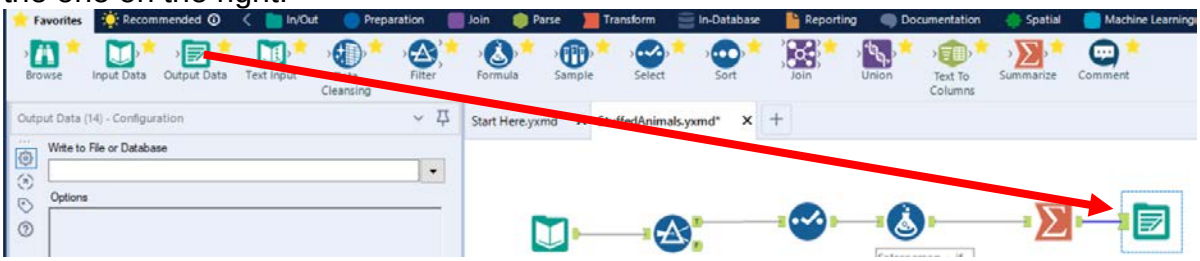


Change the “Output Field Names” and select “Run”

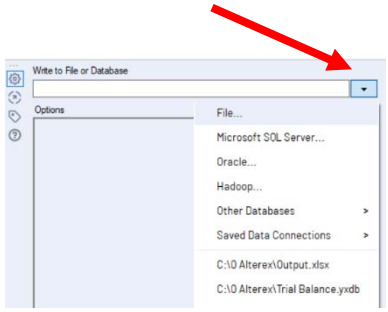


Tool 6: Select Tool Output Data Tool: We now want to export the cleaned and assembled Product Report to Excel.

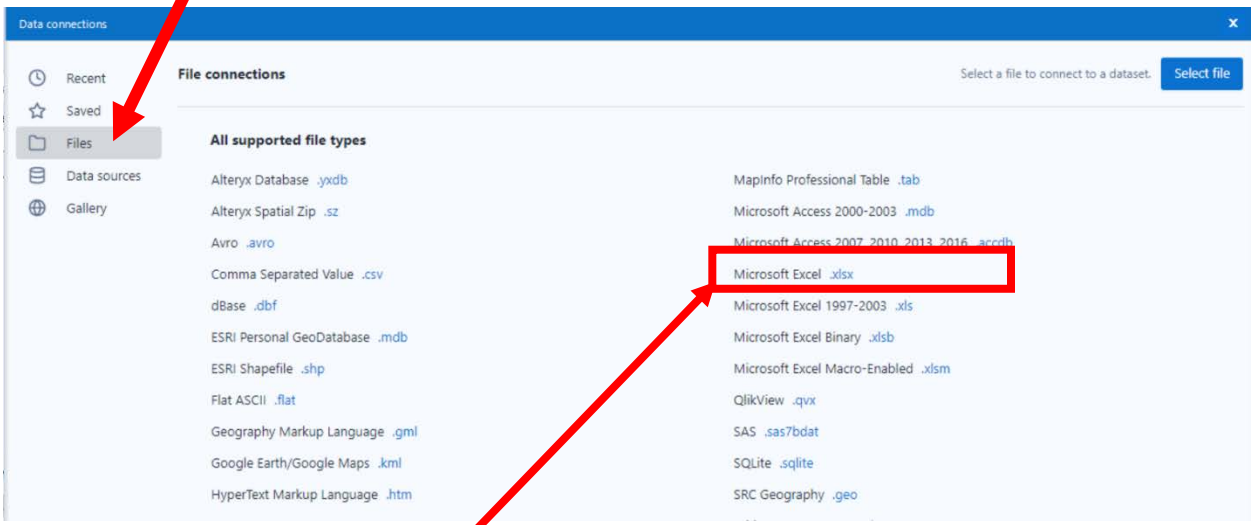
Drag an "**Output Data Tool**" to the right of the "**Summary Tool**" on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



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



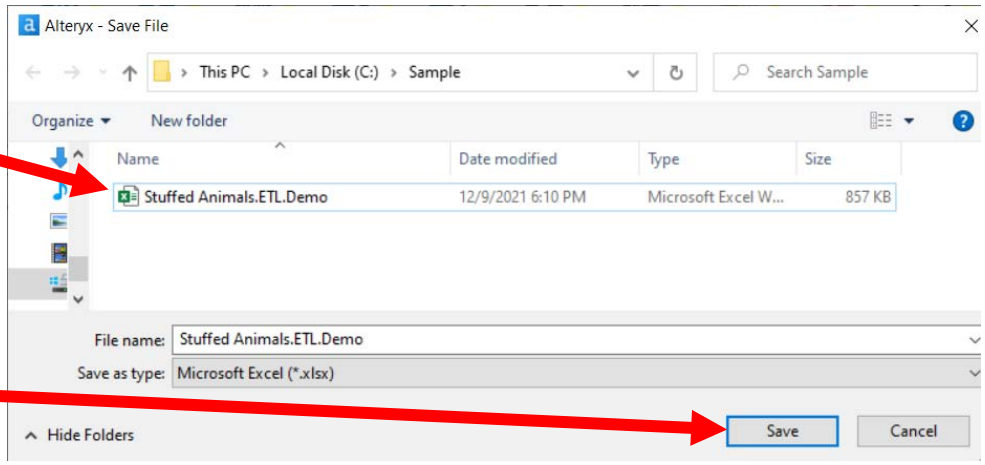
Select "File".



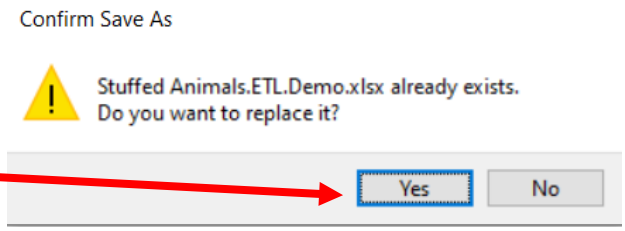
Select "Microsoft Excel .xlsx"

Select the required file name.

Select “Save”.

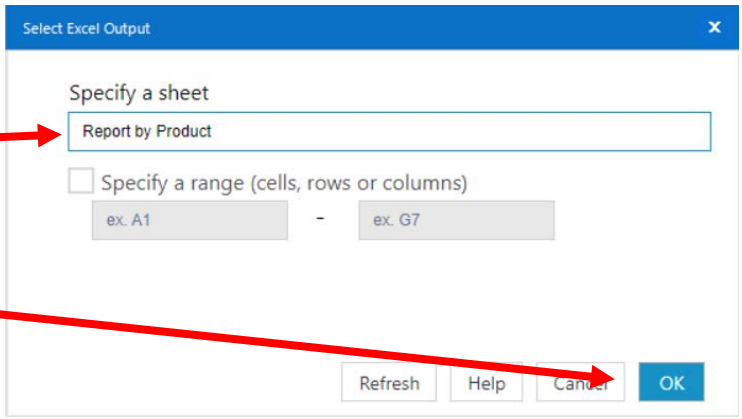


Click on “Yes” to confirm “Save”.



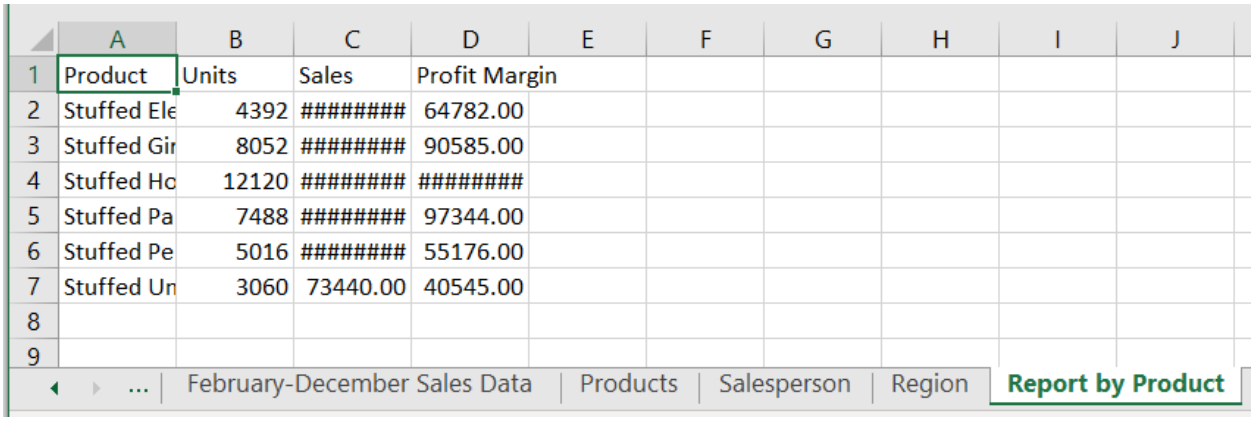
Enter the sheet name.

Select “OK”.



Select “Run”.

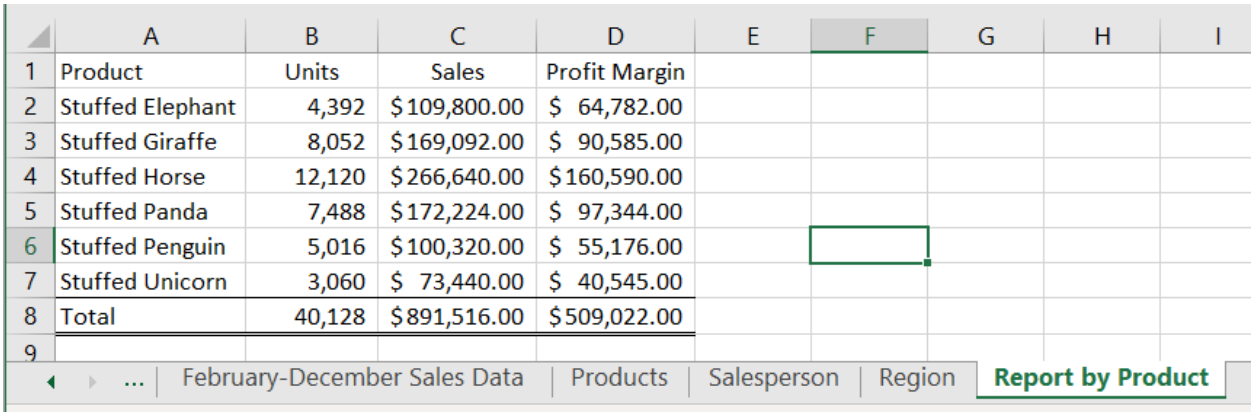
Open the file in Excel.



	A	B	C	D	E	F	G	H	I	J
1	Product	Units	Sales	Profit Margin						
2	Stuffed Ele	4392	########	64782.00						
3	Stuffed Gir	8052	########	90585.00						
4	Stuffed Ho	12120	########	########						
5	Stuffed Pa	7488	########	97344.00						
6	Stuffed Pe	5016	########	55176.00						
7	Stuffed Un	3060	73440.00	40545.00						
8										
9										

... | February-December Sales Data | Products | Salesperson | Region | **Report by Product**

Change the columns' widths and formats.

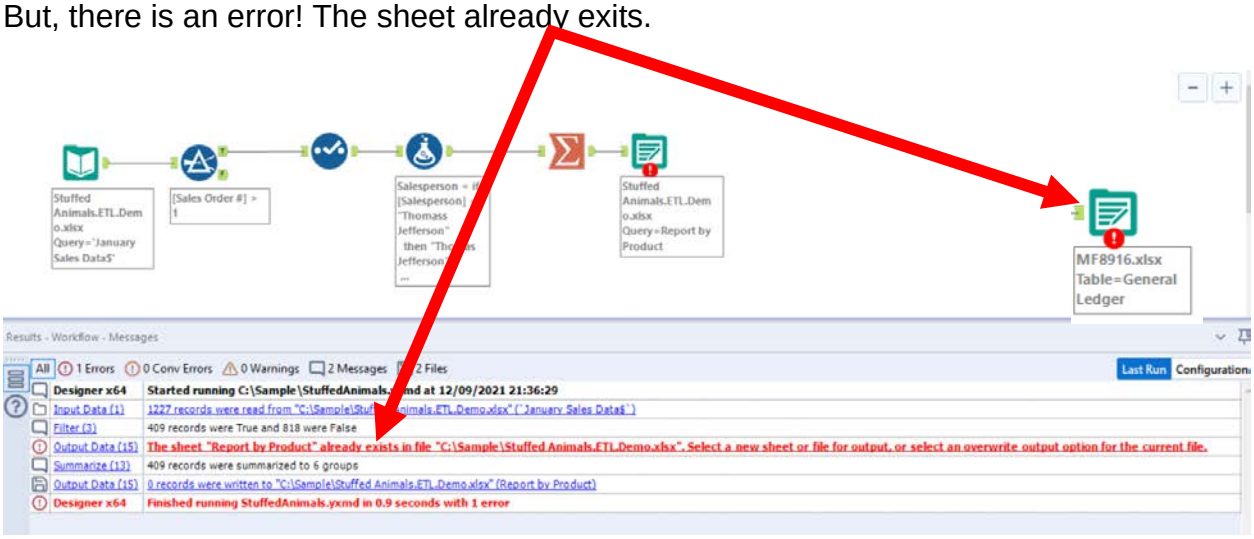


	A	B	C	D	E	F	G	H	I
1	Product	Units	Sales	Profit Margin					
2	Stuffed Elephant	4,392	$109,800.00	$ 64,782.00					
3	Stuffed Giraffe	8,052	$169,092.00	$ 90,585.00					
4	Stuffed Horse	12,120	$266,640.00	$160,590.00					
5	Stuffed Panda	7,488	$172,224.00	$ 97,344.00					
6	Stuffed Penguin	5,016	$100,320.00	$ 55,176.00					
7	Stuffed Unicorn	3,060	$ 73,440.00	$ 40,545.00					
8	Total	40,128	$891,516.00	$509,022.00					
9									

... | February-December Sales Data | Products | Salesperson | Region | **Report by Product**

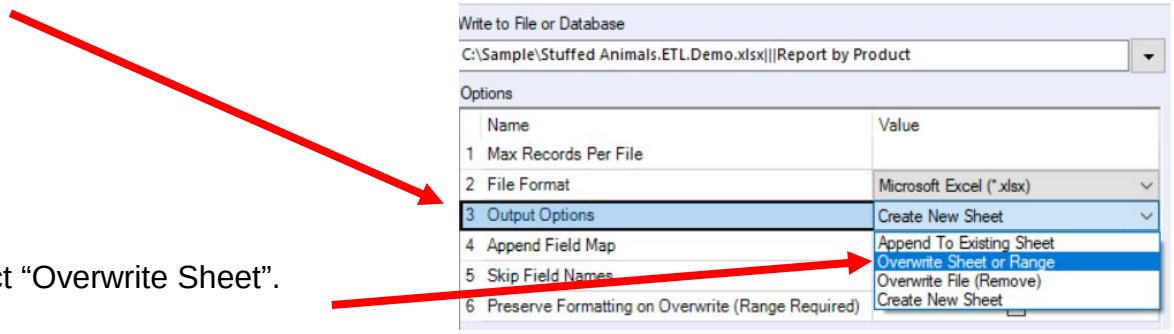
"Save" and "Close" the Excel file.
Return to Alteryx and select "Run" for the second time.

But, there is an error! The sheet already exits.



In the output Configuration area,
select “Output Options” number 3.

Select “Overwrite Sheet”.

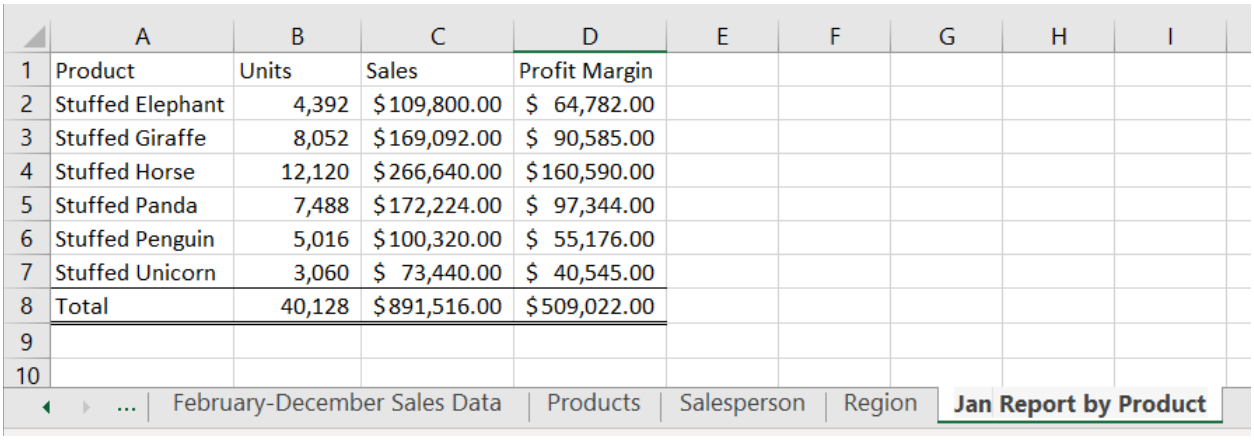


Select “Run”.

Open the Excel file.

	A	B	C	D	E	F	G	H	I	J
1	Product	Units	Sales	Profit Margin						
2	Stuffed Ele	4392	########	64782.00						
3	Stuffed Gir	8052	########	90585.00						
4	Stuffed Ho	12120	########	#########						
5	Stuffed Pa	7488	########	97344.00						
6	Stuffed Pe	5016	########	55176.00						
7	Stuffed Un	3060	73440.00	40545.00						
8										
9										

... | February-December Sales Data | Products | Salesperson | Region | **Report by Product**



	A	B	C	D	E	F	G	H	I
1	Product	Units	Sales	Profit Margin					
2	Stuffed Elephant	4,392	$109,800.00	$ 64,782.00					
3	Stuffed Giraffe	8,052	$169,092.00	$ 90,585.00					
4	Stuffed Horse	12,120	$266,640.00	$160,590.00					
5	Stuffed Panda	7,488	$172,224.00	$ 97,344.00					
6	Stuffed Penguin	5,016	$100,320.00	$ 55,176.00					
7	Stuffed Unicorn	3,060	$ 73,440.00	$ 40,545.00					
8	Total	40,128	$891,516.00	$509,022.00					
9									
10									

... | February-December Sales Data | Products | Salesperson | Region | **Jan Report by Product**

Reformat the Excel sheet and rename the formatted sheet “Jan Report by Product”. When the Alteryx flow runs it will create a new unformatted sheet “Report by Product”.

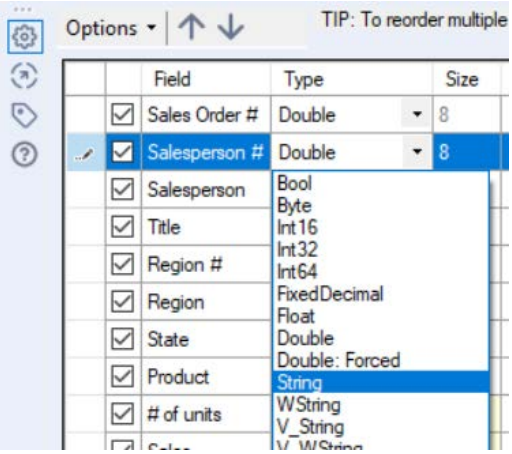
The “Salesperson” report will summarize the data by “Salesperson #” with the salesperson’s name attached for clarity. We now want to combine the fields with a space, a dash and another space between the two fields.

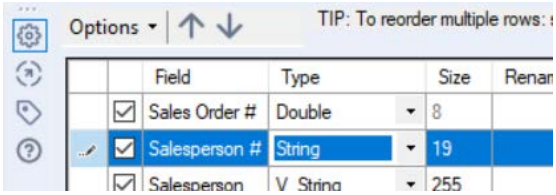
Salesperson	# of units	Sales	Profit Margin
1405 - James Polk	1,812	$41,568	$24,438
1425 - John Tyler	2,172	48,336	28,110
1541 - Ulysses Grant	756	17,112	10,062
1576 - James Madison	1,776	39,216	23,160
1588 - George Washington	1,584	35,784	21,168
1776 - James Monroe	2,016	45,192	26,700
Total	10,116	$227,208	$133,638

One field, “Salesperson #” is a numeric field, while the other field, “Salesperson’ is a string. It is easier to combine filed if they are both strings.

Select the third tool, the Select Tool from the flow.

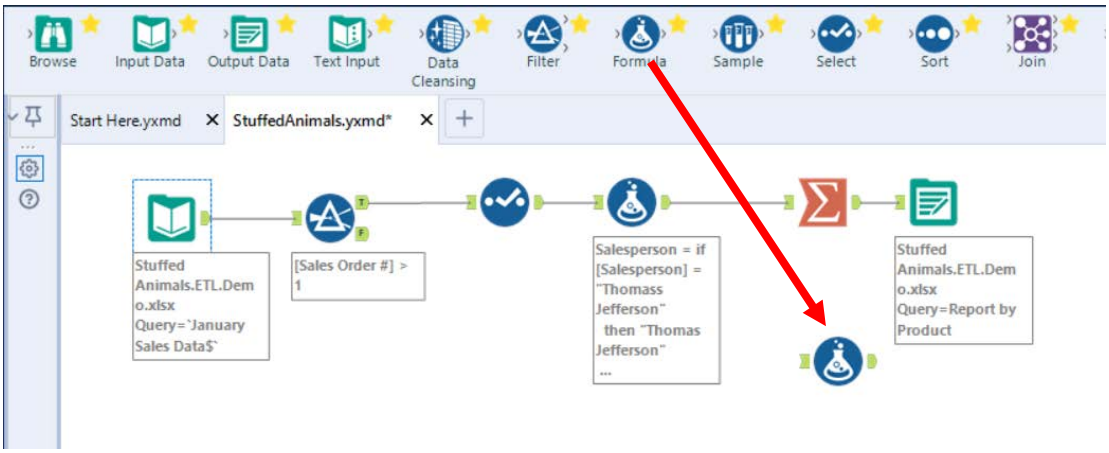
In the configuration section select the pull down in the “Type” column for the “Salesperson #” field and select “String”.



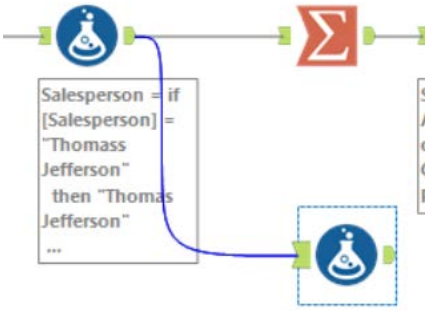


Tool 7: Select Tool Formula: We can now want to combine the fields “Salesperson #” and “Salesperson”.

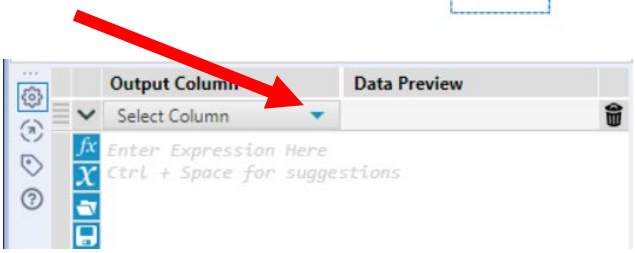
Drag a “**Formula** **Tool**” under the “**Running Balance Tool**” branching from the “**Formula Tool**” on the Canvas.



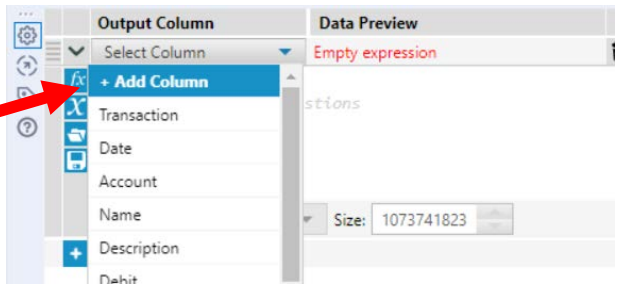
To connect the tools, drag from the outflow of the Tool 4, Formula Tool to the left inflow of the Tool 7, Formula tool.



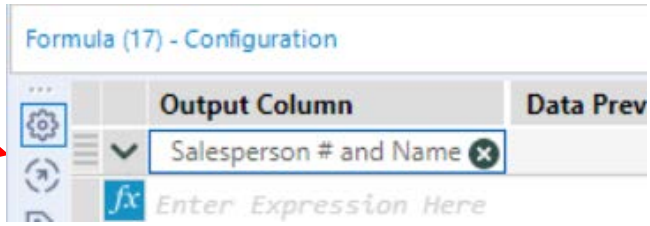
In the output Configuration area, click on the drop-down arrow of “Select Column”.



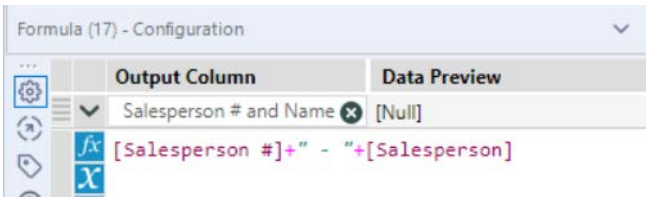
Select “Add Column”.



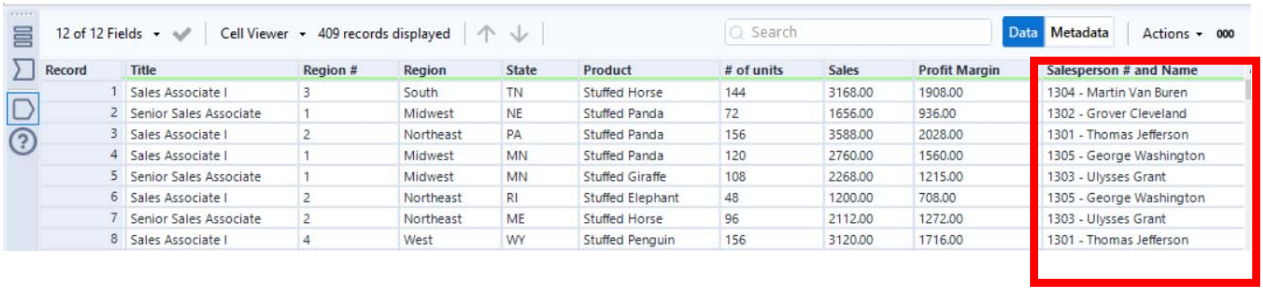
Type “Salesperson # And Name”.



Type the following formula
[Salesperson #]+" - "+[Salesperson]



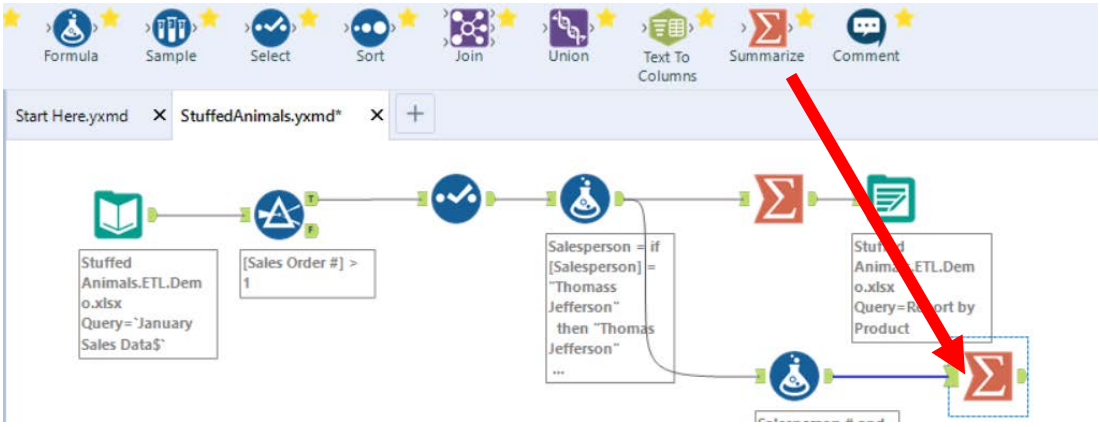
Select “Run”.



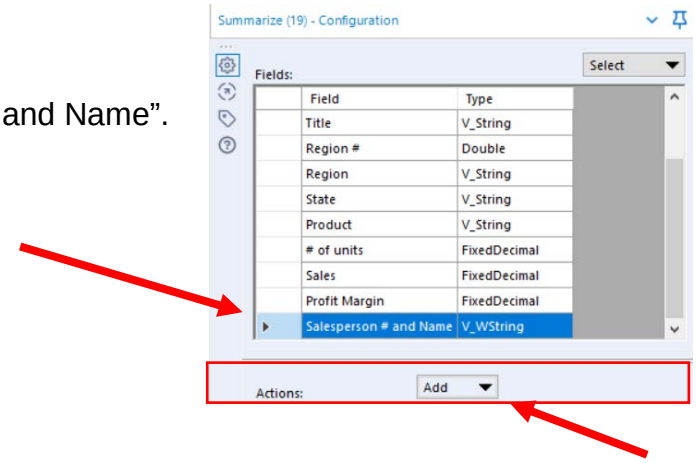
Record	Title	Region #	Region	State	Product	# of units	Sales	Profit Margin	Salesperson # and Name
1	Sales Associate I	3	South	TN	Stuffed Horse	144	3168.00	1908.00	1304 - Martin Van Buren
2	Senior Sales Associate	1	Midwest	NE	Stuffed Panda	72	1656.00	936.00	1302 - Grover Cleveland
3	Sales Associate I	2	Northeast	PA	Stuffed Panda	156	3588.00	2028.00	1301 - Thomas Jefferson
4	Sales Associate I	1	Midwest	MN	Stuffed Panda	120	2760.00	1560.00	1305 - George Washington
5	Senior Sales Associate	1	Midwest	MN	Stuffed Giraffe	108	2268.00	1215.00	1303 - Ulysses Grant
6	Sales Associate I	2	Northeast	RI	Stuffed Elephant	48	1200.00	708.00	1305 - George Washington
7	Senior Sales Associate	2	Northeast	ME	Stuffed Horse	96	2112.00	1272.00	1303 - Ulysses Grant
8	Sales Associate I	4	West	WY	Stuffed Penguin	156	3120.00	1716.00	1301 - Thomas Jefferson

Tool 8: Select Tool Summary: We now want to create a "Salesperson Report".

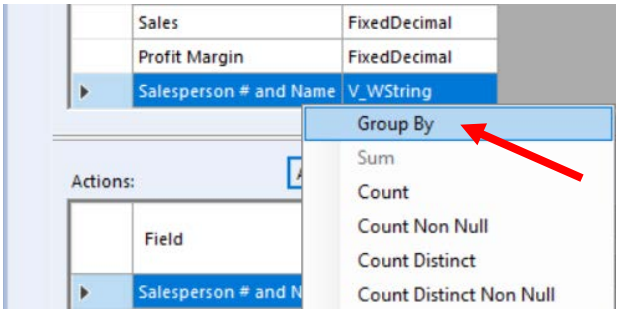
Drag a "**Summary Tool**" to the right of the "**Formula Tool**" on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



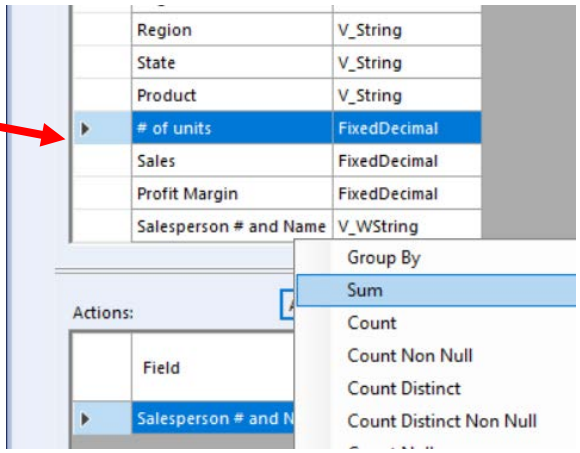
Select the first column, "Salesperson # and Name".



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

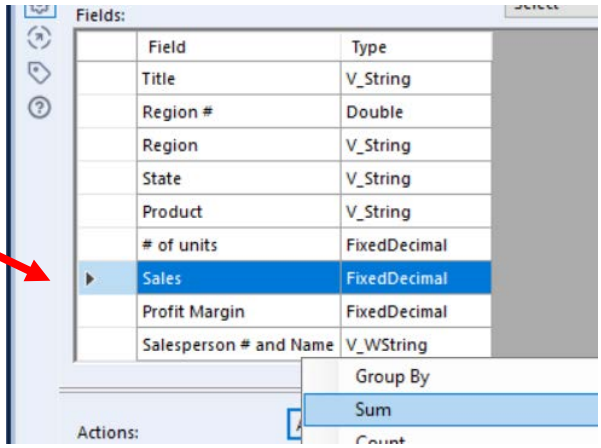


The second column will be “# of units”.



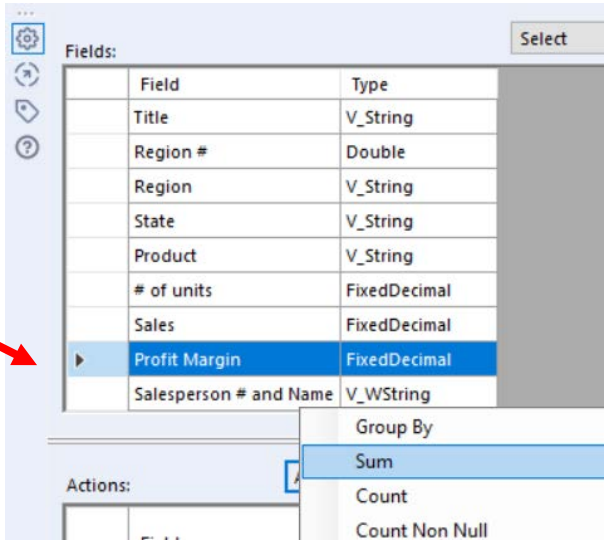
Click on the drop-down arrow on “Add” in “Actions” and select “Sum”.

The third column will be “Sales”.



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

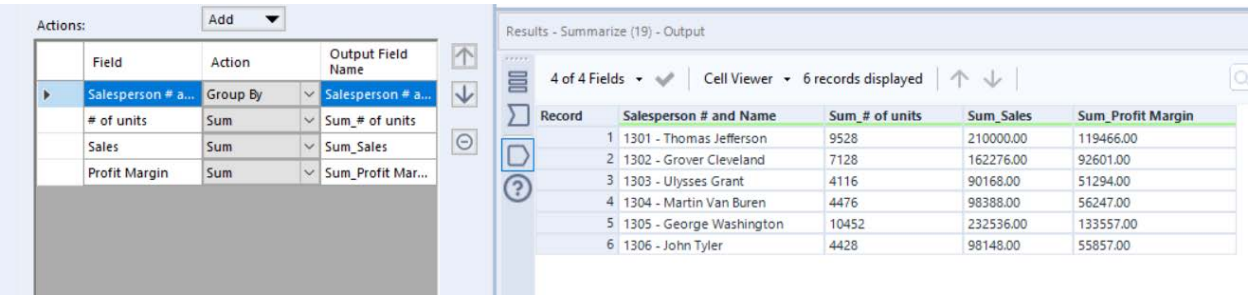
The fourth column will be “Profit Margin”.



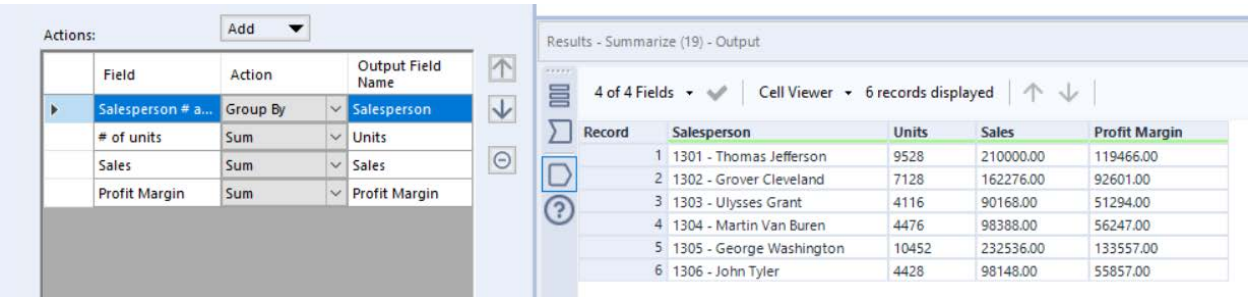
Click on the drop-down arrow on “Add” in “Actions” and select “Sum”.

Select “Run”.

And, the report is created, right click on the output arrow if nessacary.

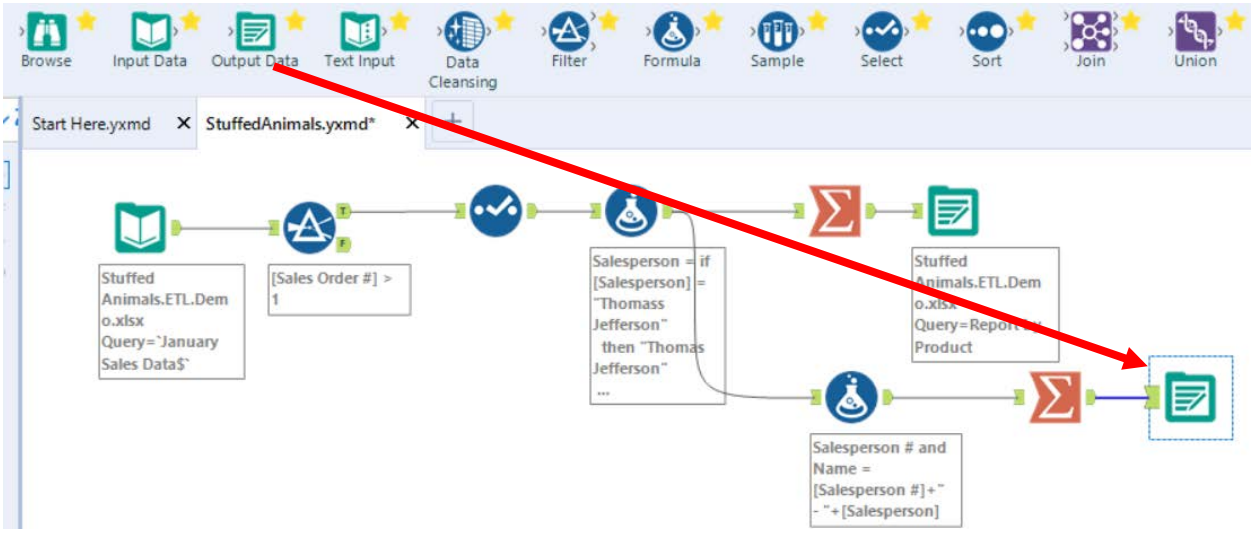


Change the “Output Field Names” and select “Run”

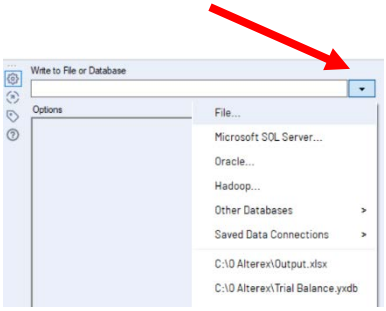


Tool 9: Select Output Data Tool: We now want to export the cleaned and assembled Salesperson Report to Excel.

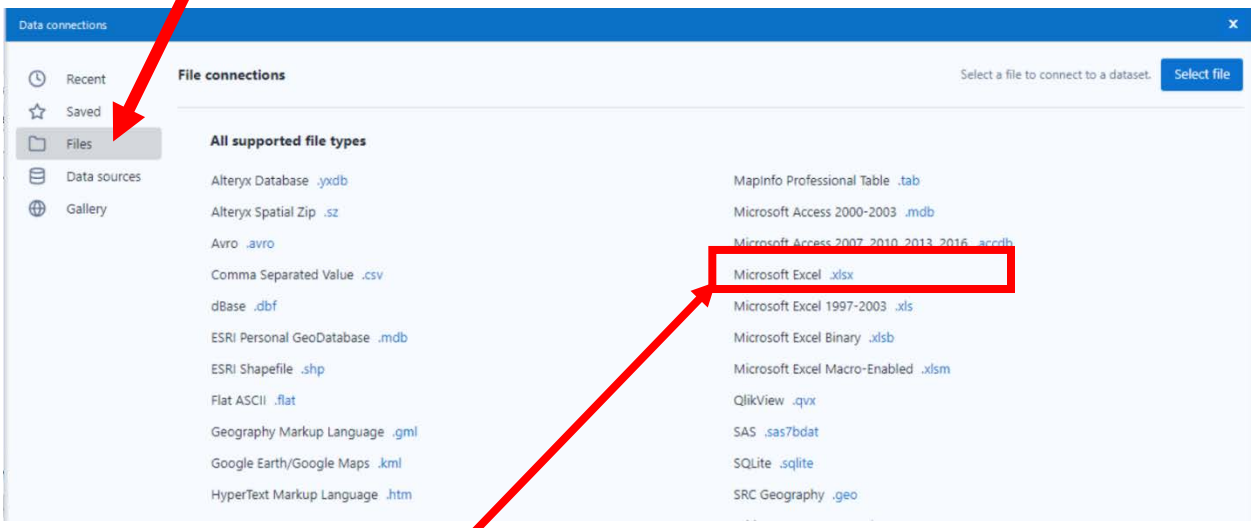
Drag an "**Output Data Tool**" to the right of the "**Summary Tool**" on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



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



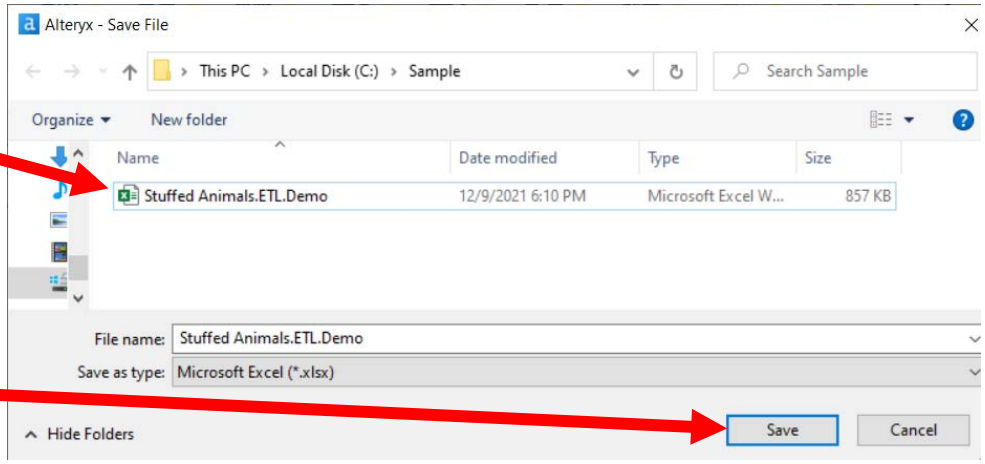
Select “File”.



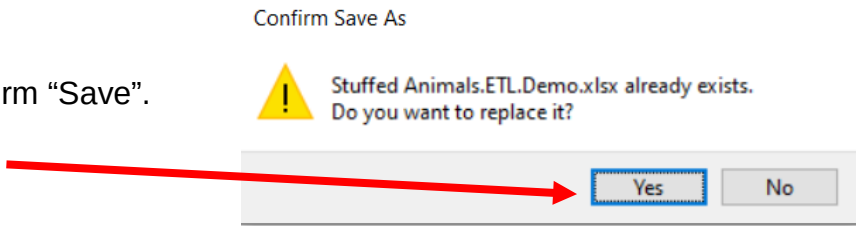
Select “Microsoft Excel .xlsx”

Select the required file name.

Select “Save”.

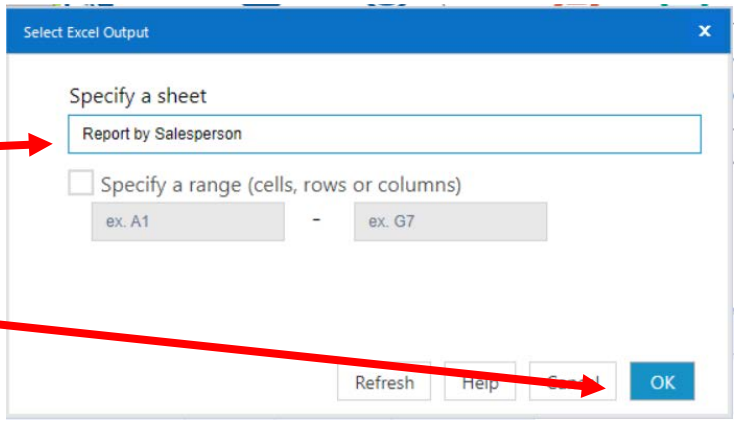


Click on “Yes” to confirm “Save”.



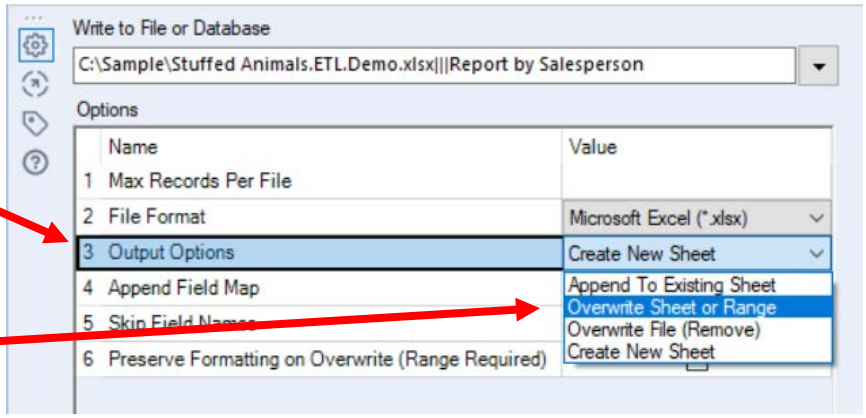
Enter the sheet name.

Select "OK".

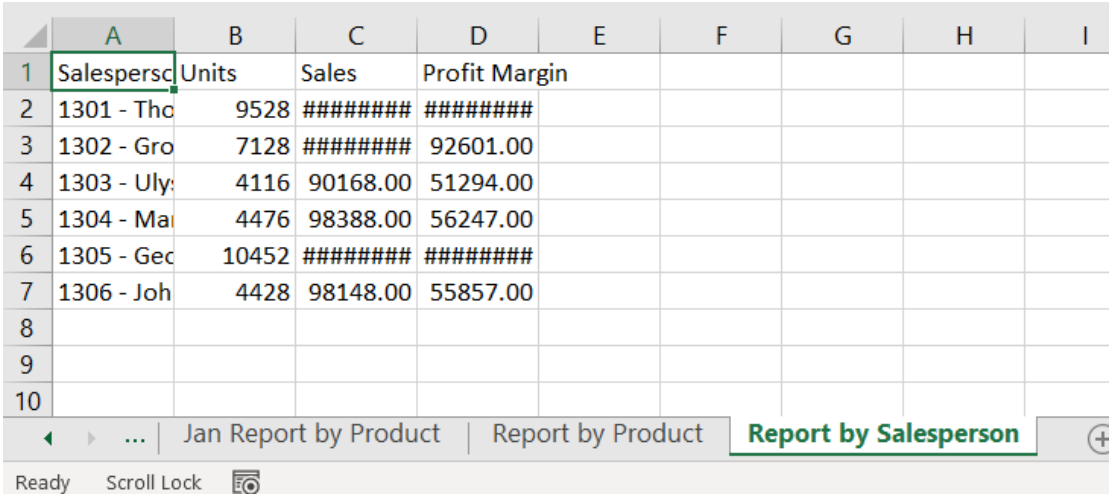


In the output Configuration area,
select "Output Options" number 3.

Select "Overwrite Sheet"
(which drops the old sheet)



Select Run
Open the Excel file



	A	B	C	D	E	F	G	H	I
1	Salespersc	Units	Sales	Profit Margin					
2	1301 - Tho	9528	########	########					
3	1302 - Gro	7128	########	92601.00					
4	1303 - Uly	4116	90168.00	51294.00					
5	1304 - Mai	4476	98388.00	56247.00					
6	1305 - Gec	10452	########	########					
7	1306 - Joh	4428	98148.00	55857.00					
8									
9									
10									

... | Jan Report by Product | Report by Product | Report by Salesperson

Ready Scroll Lock

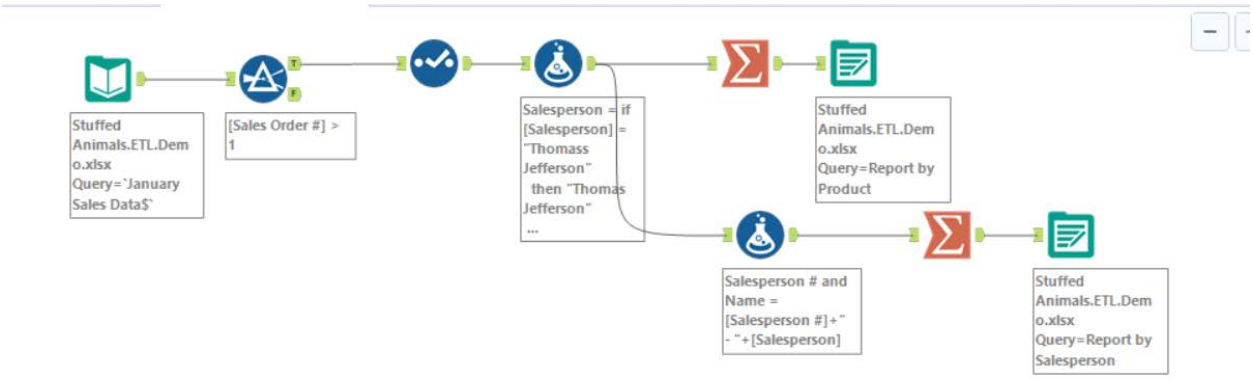
Change the format and rename the sheet so the format changes are not lost, save, close Excel and return to Alteryx.

	A	B	C	D	E	F	G
1	Salesperson	Units	Sales	Profit Margin			
2	1301 - Thomas Jefferson	9,528	$210,000.00	$119,466.00			
3	1302 - Grover Cleveland	7,128	$162,276.00	$ 92,601.00			
4	1303 - Ulysses Grant	4,116	$ 90,168.00	$ 51,294.00			
5	1304 - Martin Van Buren	4,476	$ 98,388.00	$ 56,247.00			
6	1305 - George Washington	10,452	$232,536.00	$133,557.00			
7	1306 - John Tyler	4,428	$ 98,148.00	$ 55,857.00			
8	Total	40,128	$891,516.00	$509,022.00			
9							
10							

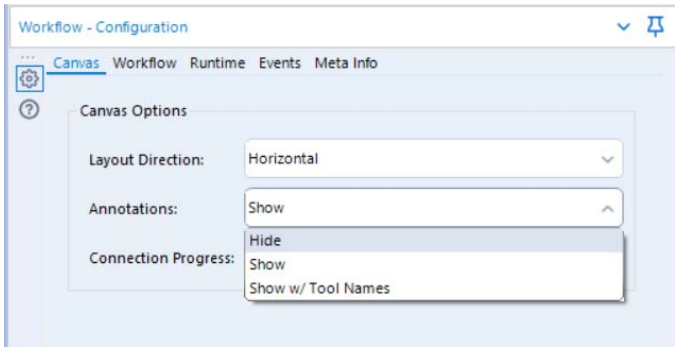
... | Jan Report by Product | Report by Product | **Jan Report by Salesperson**

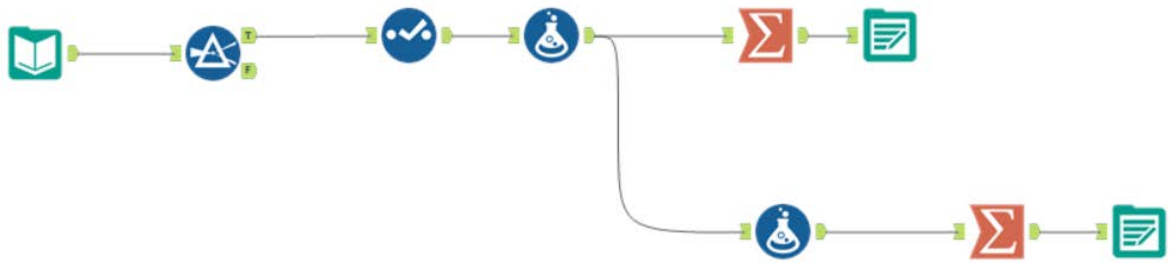
Ready Scroll Lock

The canvas is very busy looking because of the annotations.



Select any blank area on the canvas and in the Configuration section change “Annotations” to “Hide”.

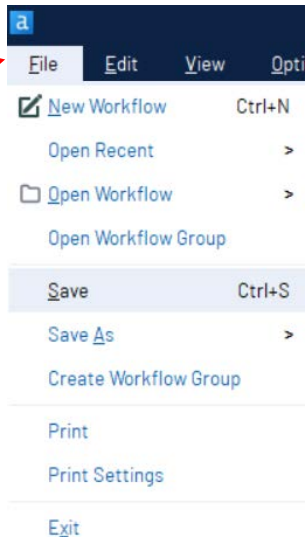




The workflow must be saved.

Select “File”.

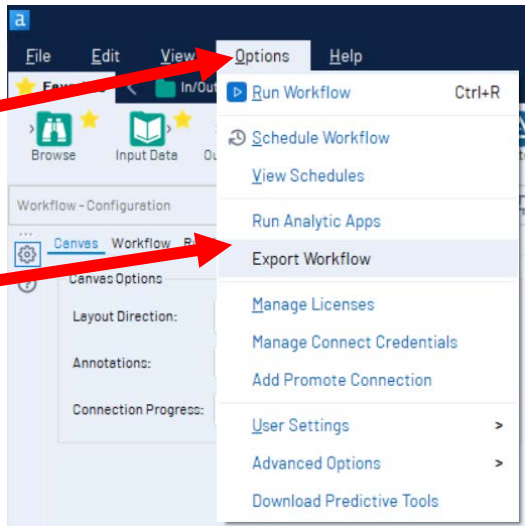
Select “Save”.



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

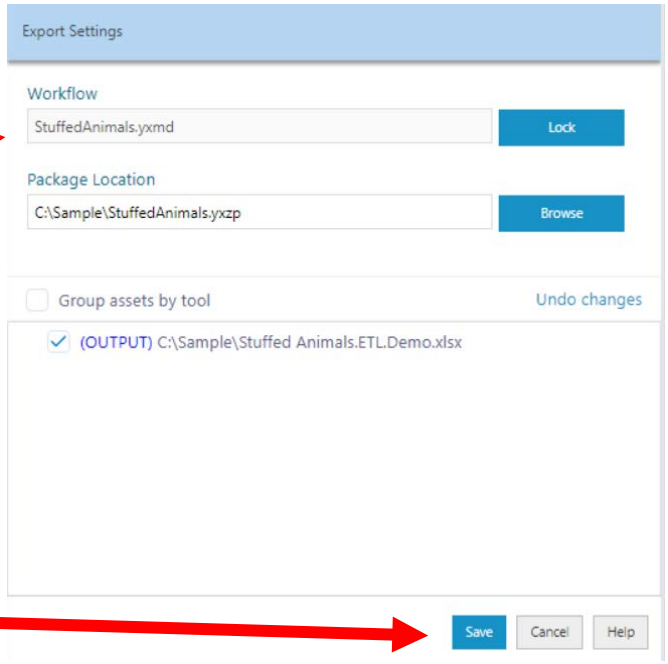
Select “Options”.

Select “Export Workflow”.

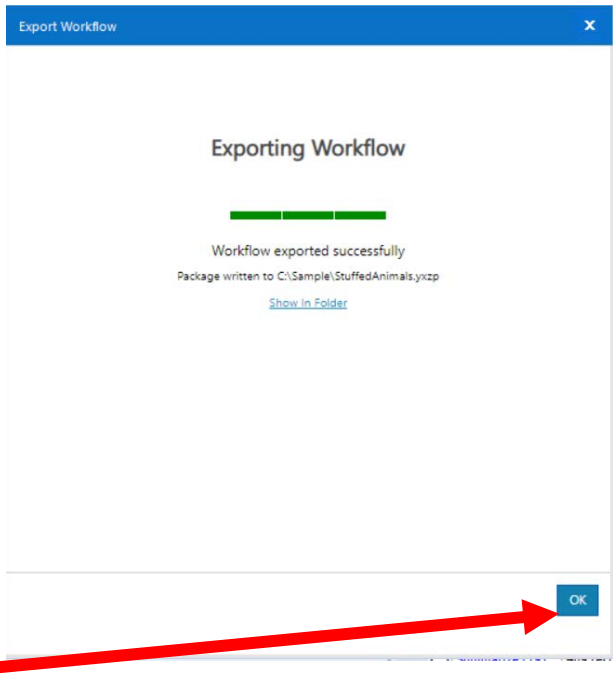


Verify the name of the “Workflow”.

Verify the “Package Location”.



Select “Save”.

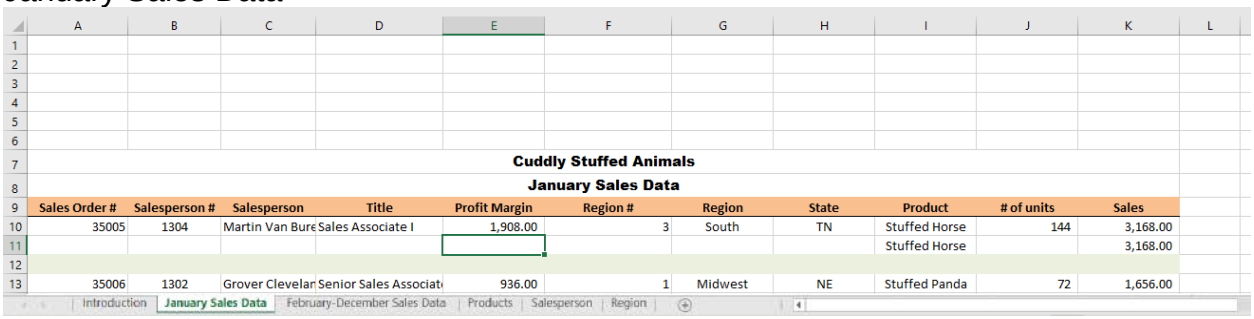


Select “OK”.

Part 2

The “February-December Sales Data” tab contains data that was created in a different system than the “January Sales Data”. You cannot append two data sets unless the field names and field types are identical. The fields however, may be ordered differently.

January Sales Data



	A	B	C	D	E	F	G	H	I	J	K	L
1												
2												
3												
4												
5												
6												
7					Cuddly Stuffed Animals							
8					January Sales Data							
9	Sales Order #	Salesperson #	Salesperson	Title	Profit Margin	Region #	Region	State	Product	# of units	Sales	
10	35005	1304	Martin Van Bure	Sales Associate I	1,908.00	3	South	TN	Stuffed Horse	144	3,168.00	
11									Stuffed Horse		3,168.00	
12												
13	35006	1302	Grover Clevelar	Senior Sales Associat	936.00	1	Midwest	NE	Stuffed Panda	72	1,656.00	

Introduction | January Sales Data | February-December Sales Data | Products | Salesperson | Region

February-December Sales Data

	A	B	C	D	E	F	G	H	I	J	K
1											
2					Cuddly Stuffed Animals						
3					February-December Sales						
4	Order #	SalesPerson	Title	Region Name	Region Code	*State*	Article	# of pieces sold	Selling Price Per Unit	Profit Margin	
5	37060	1303 : Ulysses Grant	Senior Sales Associate	South	3	TX	Stuffed Horse	120	22.00	1,590.00	
6							Stuffed Horse			1,590.00	
7											
8											
9	37064	1301 : Thomas Jefferson	Sales Associate I	Northeast	2	CT	Stuffed Horse	144	22.00	1,908.00	
10							Stuffed Horse				

Introduction | January Sales Data | February-December Sales Data | Products | Salesperson | Region

Ready

As you review the new sheet note the differences:

- The column headings are in row 4.
- The column headings are different.
- The columns are in a different order.
- Column B contains both the Salesperson’s number and name.
- Selling Price Per Units has replaced Sales.

Sample Reports

Product	# of units	Sales	Profit Margin
Stuffed Frog	1,524	$ 33,528	$20,574
Stuffed Unicorr	408	8,160	4,692
Stuffed Panda	1,140	26,220	15,675
Stuffed Koala	624	14,976	8,112
Stuffed Elephai	4,044	84,924	49,539
Stuffed Bear	2,376	59,400	35,046
Total	10,116	$227,208	$133,638

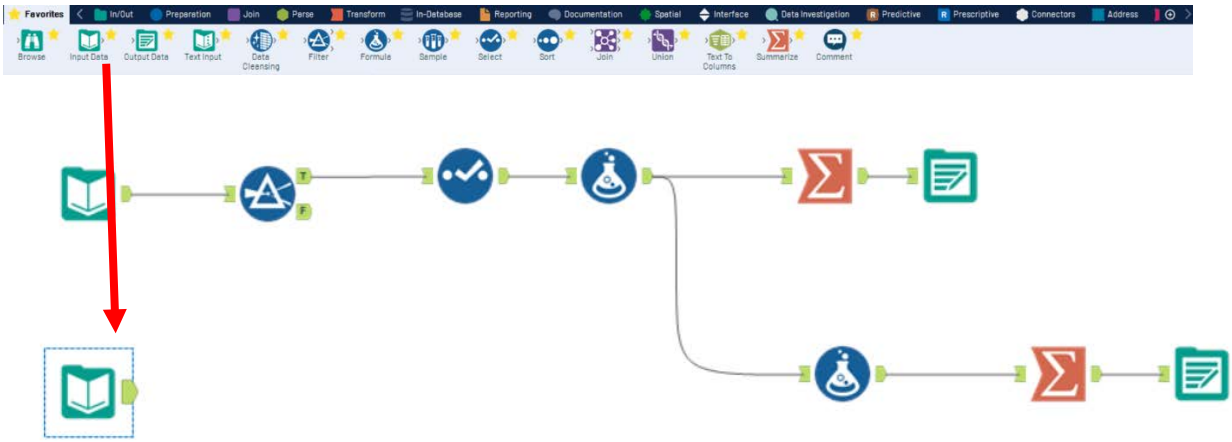
Salesperson	# of units	Sales	Profit Margin
1405 - James Polk	1,812	$41,568	$24,438
1425 - John Tyler	2,172	48,336	28,110
1541 - Ulysses Grant	756	17,112	10,062
1576 - James Madison	1,776	39,216	23,160
1588 - George Washington	1,584	35,784	21,168
1776 - James Monroe	2,016	45,192	26,700
Total	10,116	$227,208	$133,638

Tool 10: Input Data Tool – Input Excel file

Drag the "Input Data Tool"

from the Tools Palette to the Canvas.

Run

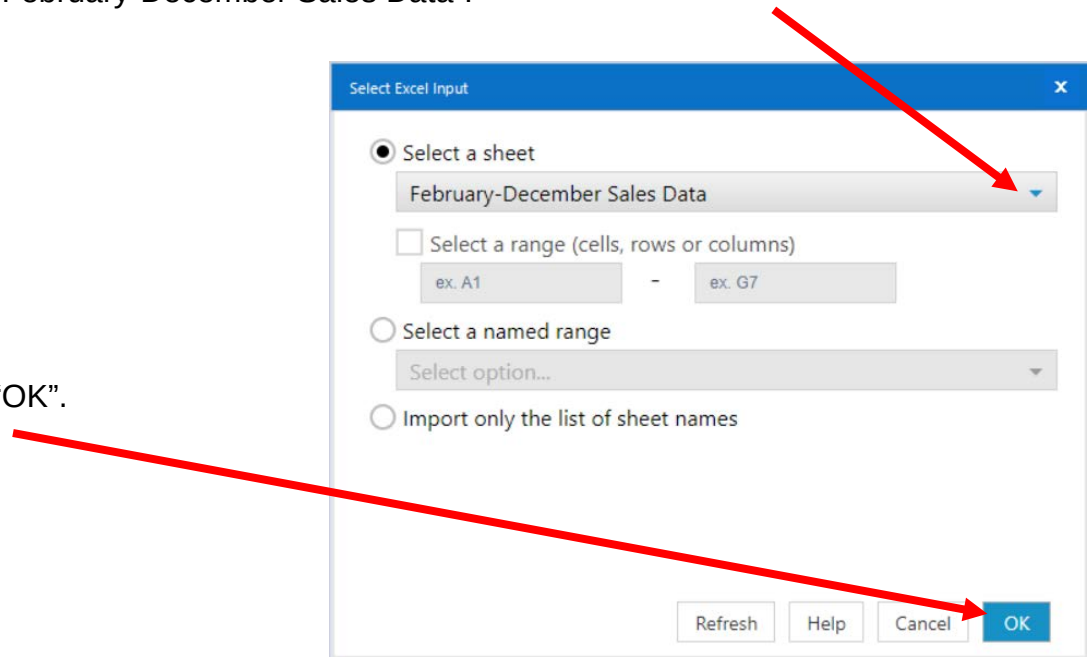


In the Configuration area, select the “Connect a File or Database” drop-down arrow.

Select “Recent connection”.

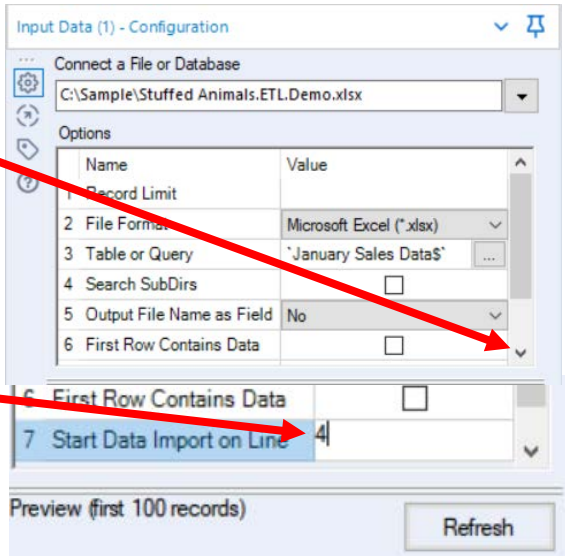
Select “February-December Sales Data”.

Select “OK”.



Remember that the field names are in the fourth row.

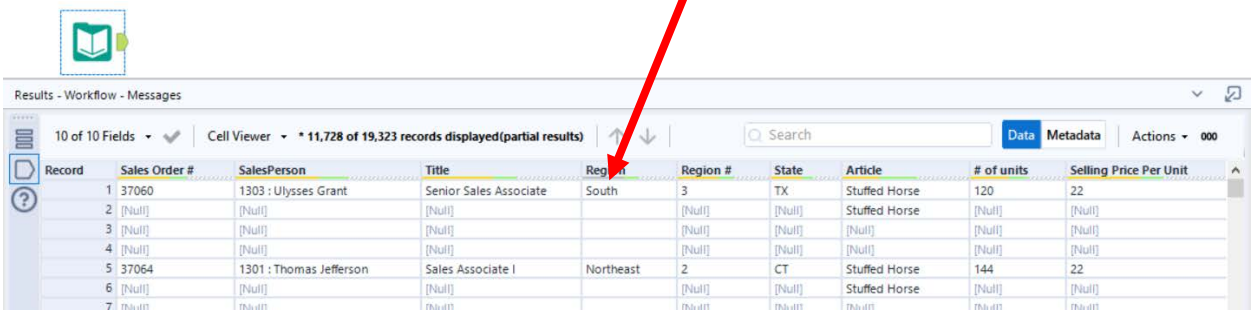
In the Configuration area, scroll down to option 7.



Change the number in “Start Data Import on Line” to “4”.

Select “Run”.

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

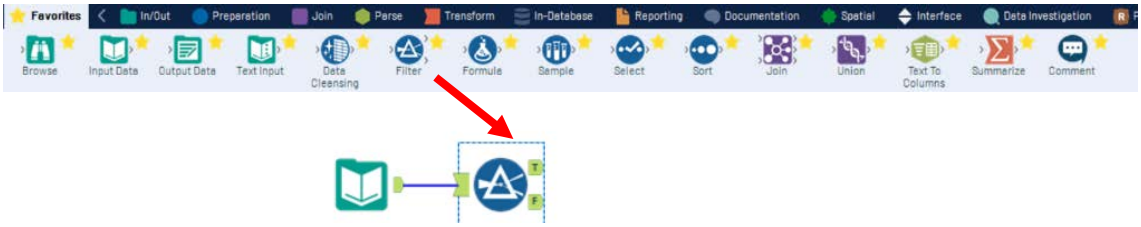


Record	Sales Order #	SalesPerson	Title	Region	Region #	State	Article	# of units	Selling Price Per Unit
1	37060	1303 : Ulysses Grant	Senior Sales Associate	South	3	TX	Stuffed Horse	120	22
2	[Null]	[Null]	[Null]		[Null]	[Null]	Stuffed Horse	[Null]	[Null]
3	[Null]	[Null]	[Null]		[Null]	[Null]	[Null]	[Null]	[Null]
4	[Null]	[Null]	[Null]		[Null]	[Null]	[Null]	[Null]	[Null]
5	37064	1301 : Thomas Jefferson	Sales Associate I	Northeast	2	CT	Stuffed Horse	144	22
6	[Null]	[Null]	[Null]		[Null]	[Null]	Stuffed Horse	[Null]	[Null]

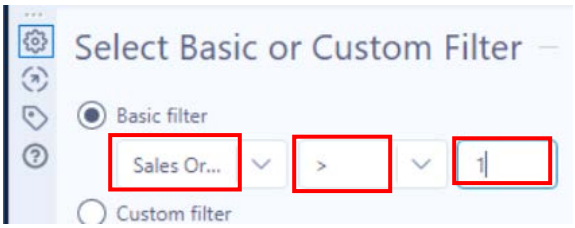
Tool 11: Filter Tool – Exclude all rows that do not contain data

There are several rows that do not contain an sales order number. We can modify the working dataset so that it includes only those rows that contain account numbers greater than one. Note that we are not deleting anything from the source data file. The working dataset will be split into two subsets. The first subset, “True”, will include all rows with account numbers greater than one. The second subset, “False”, will contain all rows with account numbers less than one.

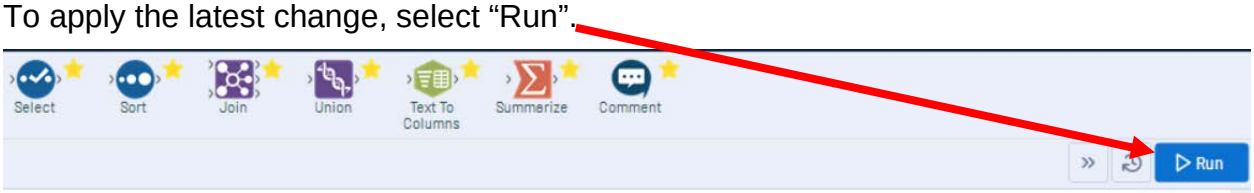
Drag the “**Filter Tool**” from the Tools Palette to the right of the “Input Data Tool” on the Canvas. The two will be connected automatically. The outflow from the tool on the left is the inflow to the one on the right.



With the “**Filter Tool**” selected, the “Basic” filter in the Configuration area can be modified to filter the field “Sales Order” for those accounts greater than “1’.

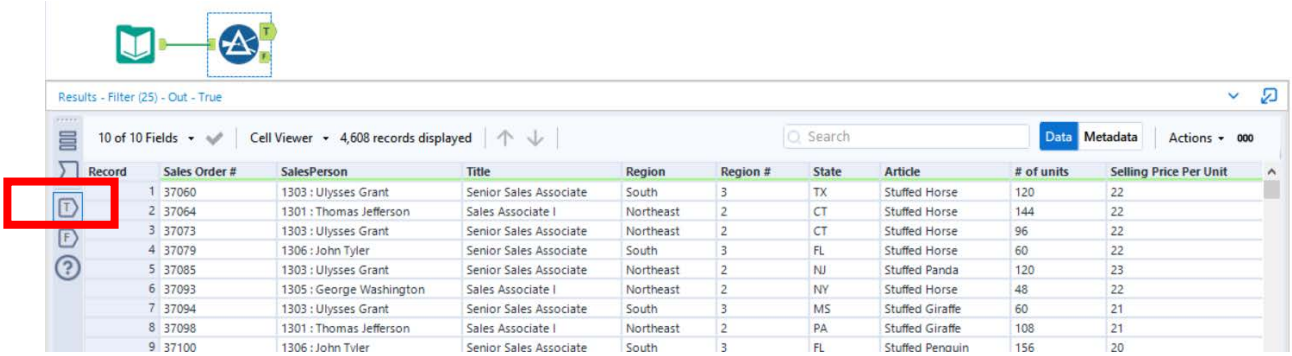


To apply the latest change, select “Run”.



We can view three different datasets in the Results window.

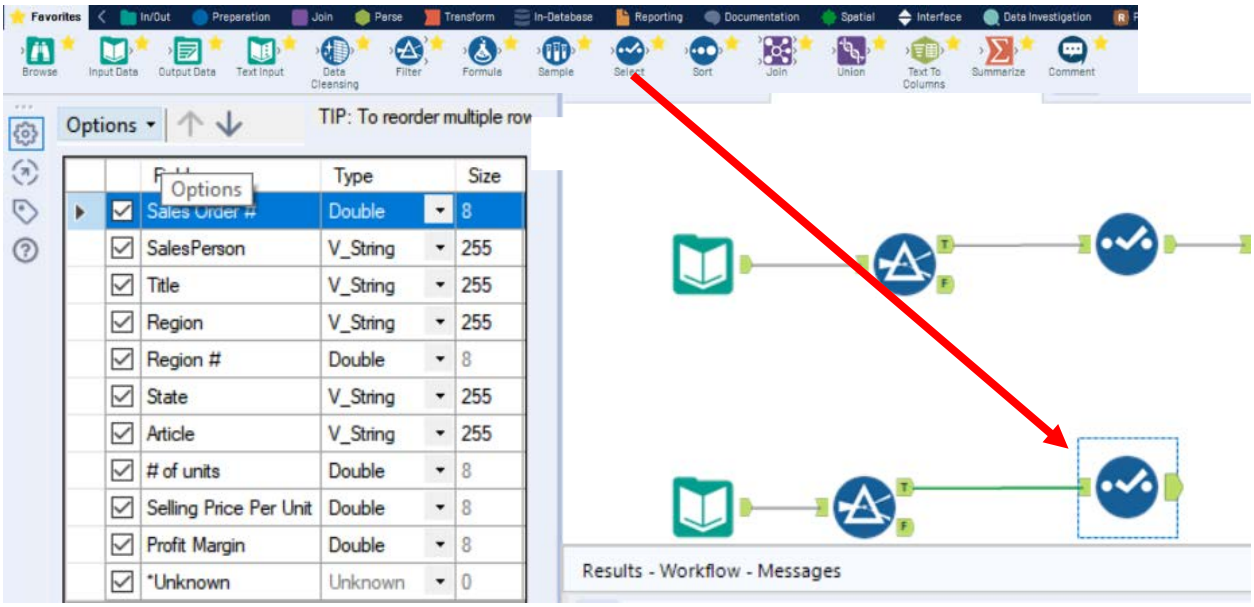
The records that meet the filtered conditions, “Sales Order > 1”, are displayed as T (True).



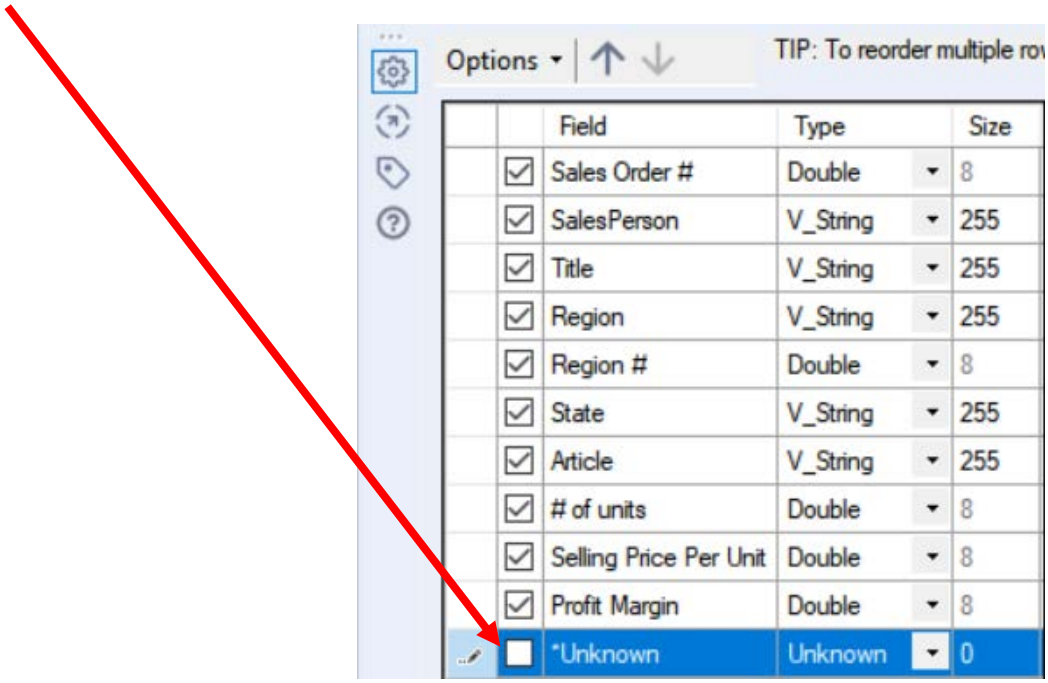
Record	Sales Order #	SalesPerson	Title	Region	Region #	State	Article	# of units	Selling Price Per Unit
1	37060	1303 : Ulysses Grant	Senior Sales Associate	South	3	TX	Stuffed Horse	120	22
2	37064	1301 : Thomas Jefferson	Sales Associate I	Northeast	2	CT	Stuffed Horse	144	22
3	37073	1303 : Ulysses Grant	Senior Sales Associate	Northeast	2	CT	Stuffed Horse	96	22
4	37079	1306 : John Tyler	Senior Sales Associate	South	3	FL	Stuffed Horse	60	22
5	37085	1303 : Ulysses Grant	Senior Sales Associate	Northeast	2	NJ	Stuffed Panda	120	23
6	37093	1305 : George Washington	Sales Associate I	Northeast	2	NY	Stuffed Horse	48	22
7	37094	1303 : Ulysses Grant	Senior Sales Associate	South	3	MS	Stuffed Giraffe	60	21
8	37098	1301 : Thomas Jefferson	Sales Associate I	Northeast	2	PA	Stuffed Giraffe	108	21
9	37100	1306 : John Tyler	Senior Sales Associate	South	3	FL	Stuffed Penguin	156	20

Tool 3: Select Tool – Remove fields that contain no data, from "*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 any extra fields by unchecking them.



To apply the latest change, select "Run.

Transforming and cleaning the data

To append the sets of data the "Field Names" and "Types" of "February-December Sales Data" must be the same as those in the data set "January Sales Data". There are ten fields in the "February-December Sales Data" stream, and we will review them one by one making all necessary changes.

January Sales Data

	Field	Type	Size	Renan
☑	Sales Order #	Double	8	
☑	SalesPerson	V_String	255	
☑	Title	V_String	255	
☑	Region	V_String	255	
☑	Region #	Double	8	
☑	State	V_String	255	
☑	Article	V_String	255	
☑	# of units	FixedDecimal	10.2	
☑	Selling Price Per Unit	FixedDecimal	10.2	
☑	Profit Margin	FixedDecimal	10.2	

February-December Sales Data

	Field	Type	Size	Rename
☑	Sales Order #	Double	8	
☑	SalesPerson	V_String	255	
☑	Title	V_String	255	
☑	Region	V_String	255	
☑	Region #	Double	8	
☑	State	V_String	255	
☑	Article	V_String	255	
☑	# of units	Double	8	
☑	Selling Price Per Unit	Double	8	
☑	Profit Margin	Double	8	

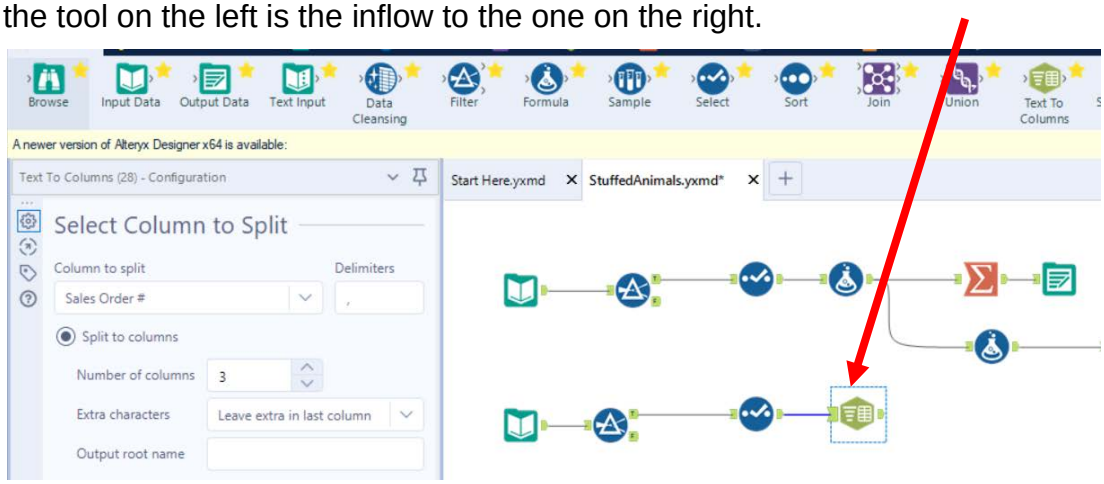
1) The field "**Sales Order #"** has no changes

2) The field "**SalesPerson**" contains two fields, "Salesperson #" (1303) and "Salesperson" (Ulysses Grant") separated by, " : ", a space, a colon and another space. The characters that separate two fields are referred to as delimiters. Common delimiters are commas, dashes, and semicolons.

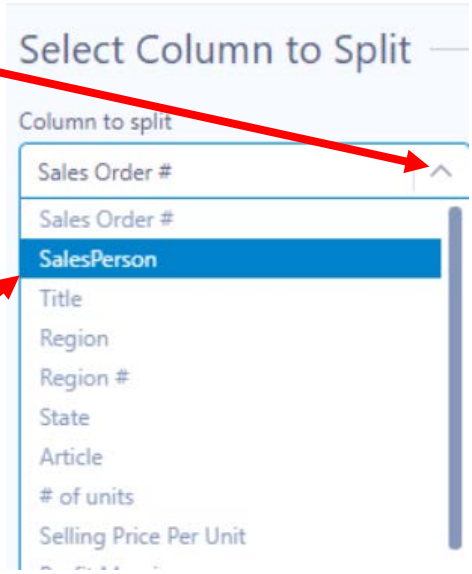
   We will convert one field into two fields. Click on the plus sign to the right of the "February-December Sales Data"".

Tool 2: Text To Columns Tool – Separate one column into many columns

Drag the "**Text To Columns 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 "**Text To Columns Tool**" selected, click the drop-down arrow on the "Column to Split" option.

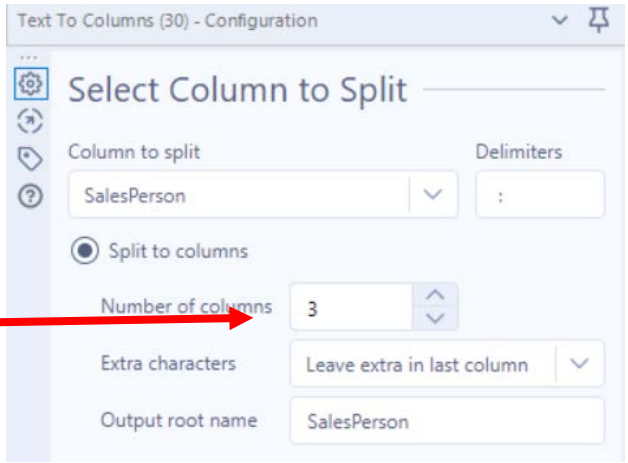


Select the field "SalesPerson".

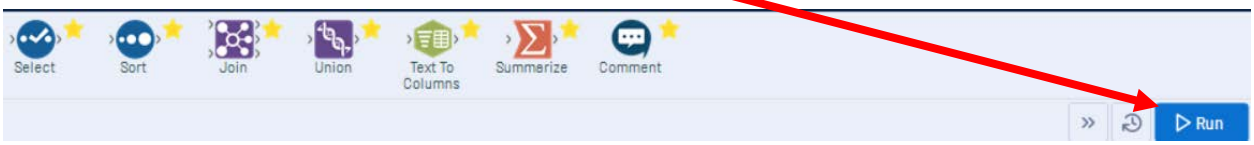
Enter "a space, a colon and another space" as the delimiter.

Enter "3" as the number of columns.

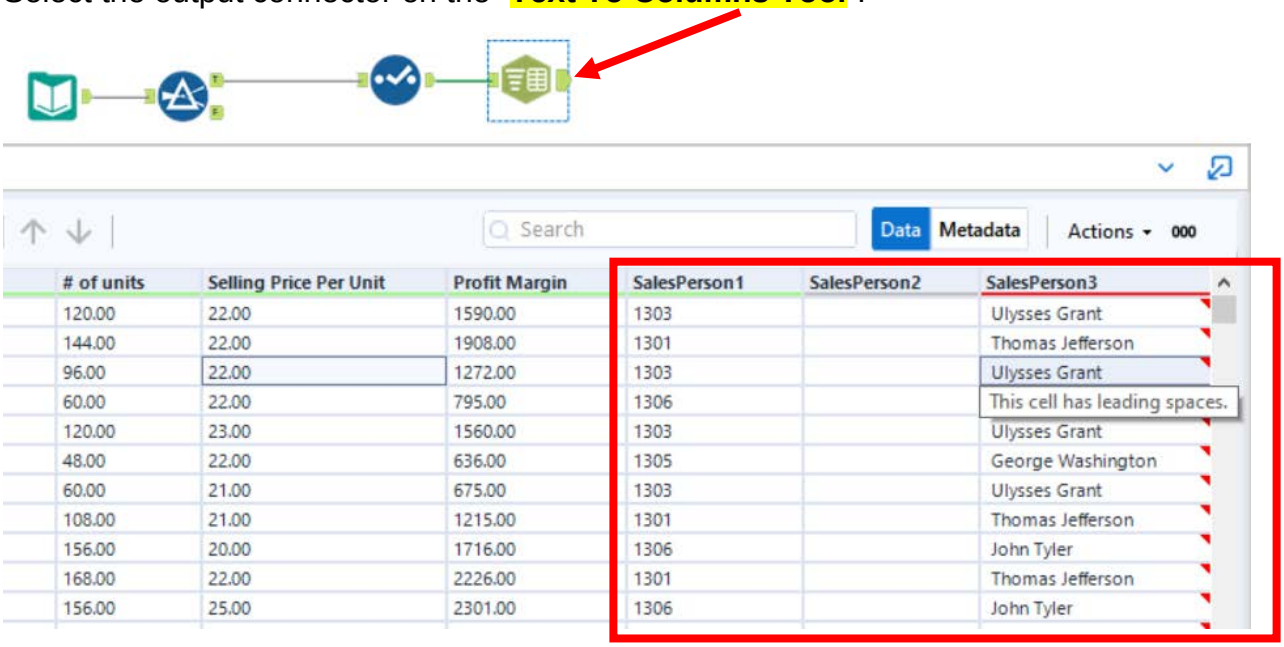
The defaults are OK.



To apply the latest change, select "Run".



Select the output connector on the "**Text To Columns Tool**".



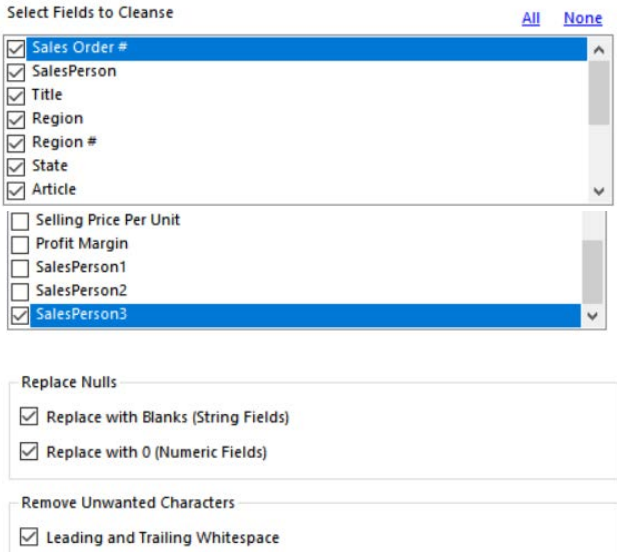
The "Results" window will display the field headings and the data.

Three fields were created, SalesPerson1, SalesPerson2 and SalesPerson3 that has a leading space.

Tool 2: Data Cleaning Tool – Remove Leading Spaces

Drag the "**Data Cleaning Tool**" from the Tools Palette to the right of the "**Text To Columns 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 configuration area, select None for the "Select Fields to Cleanse".



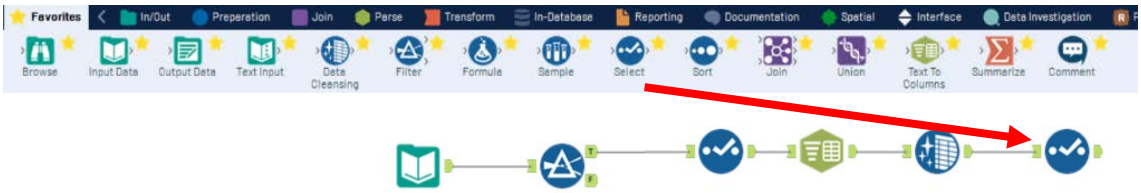
Check the field "SalesPerson3".

To apply the latest change, select "Run".

SalesPerson1	SalesPerson2	SalesPerson3
1303		Ulysses Grant
1301		Thomas Jefferson
1303		Ulysses Grant
1306		John Tyler
1303		Ulysses Grant
1305		George Washington
1303		Ulysses Grant
1301		Thomas Jefferson
1306		John Tyler
1301		Thomas Jefferson

Tool 3: Select Tool – Change the names of fields and remove fields.

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



2 The original field "**SalesPerson"** must be removed.

		Field	Type	Size	Rename
1	☑	Sales Order #	Double	8	
2	☐	SalesPerson	V_String	255	
3	☑	Title	V_String	255	
4	☑	Region	V_String	255	
5	☑	Region #	Double	8	
6	☑	State	V_String	255	
7	☑	Article	V_String	255	
8	☑	# of units	Double	8	
9	☑	Selling Price Per Unit	Double	8	
10	☑	Profit Margin	Double	8	
11	☑	SalesPerson1	V_String	255	
12	☑	SalesPerson2	V_String	255	
13	☑	SalesPerson3	V_String	255	

3 The field "**Title**" has no changes.
4 The field "**Region**" has no changes.
5 The field "**Region #**" has no changes.
6 The field "**State**" has no changes.
7 The field "**Article**" has to be renamed "**Product".**

	Article	V_String	255	Product

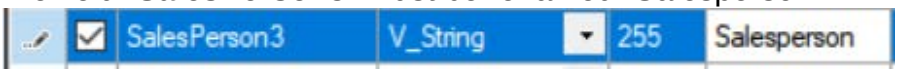
8 The fields "# of units", Selling Price Per Unit and Profit Margin must be changed.
9
10

☑	# of units	FixedDecimal	10.0
☑	Selling Price Per Unit	FixedDecimal	10.2
☑	Profit Margin	FixedDecimal	10.2

11 The field “**SalesPerson1”** has to be renamed “Salesperson #”.

12 The field “**SalesPerson2**” should be removed.

13 The field “**SalesPerson3**” must be renamed “Salesperson”.



To apply the latest change, select “Run”.

	Sales Order #	Title	Region	Region #	State	Product	# of units	Selling Price Per Unit	Profit Margin	Salesperson #	Salesperson
1	37060	Senior Sales Associate	South	3	TX	Stuffed Horse	120.00	22.00	1590.00	1303	Ulysses Grant
2	37064	Sales Associate I	Northeast	2	CT	Stuffed Horse	144.00	22.00	1908.00	1301	Thomas Jefferson
3	37073	Senior Sales Associate	Northeast	2	CT	Stuffed Horse	96.00	22.00	1272.00	1303	Ulysses Grant
4	37079	Senior Sales Associate	South	3	FL	Stuffed Horse	60.00	22.00	795.00	1306	John Tyler
5	37085	Senior Sales Associate	Northeast	2	NJ	Stuffed Panda	120.00	23.00	1560.00	1303	Ulysses Grant
6	37093	Sales Associate I	Northeast	2	NY	Stuffed Horse	48.00	22.00	636.00	1305	George Washington
7	37094	Senior Sales Associate	South	3	MS	Stuffed Giraffe	60.00	21.00	675.00	1303	Ulysses Grant
8	37098	Sales Associate I	Northeast	2	PA	Stuffed Giraffe	108.00	21.00	1215.00	1301	Thomas Jefferson
9	37100	Senior Sales Associate	South	3	FL	Stuffed Penguin	156.00	20.00	1716.00	1306	John Tyler
10	37109	Sales Associate I	Midwest	1	WI	Stuffed Horse	168.00	22.00	2226.00	1301	Thomas Jefferson
11	37117	Senior Sales Associate	West	4	WY	Stuffed Elephant	156.00	25.00	2301.00	1306	John Tyler

Temporary Tool 4: Browse Tool – The Data in the Fields Must Conform to Predefined Rules.

In the original Excel file, The Cuddly Stuffed Animals company provided the only acceptable values for “Product”, “Salesperson” and “Region”. Sometime data is entered incorrectly into the data.

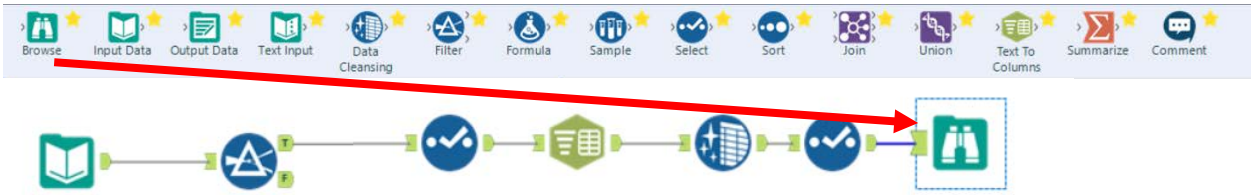
Product ID	Product
101	Stuffed Penguin
102	Stuffed Horse
103	Stuffed Unicorn
104	Stuffed Giraffe
105	Stuffed Panda
106	Stuffed Elephant

Code	Salesperson
1301	Thomas Jefferson
1302	Grover Cleveland
1303	Ulysses Grant
1304	Martin Van Buren
1305	George Washington
1306	John Tyler

Code	Region
1	Midwest
2	Northeast
3	South
4	West

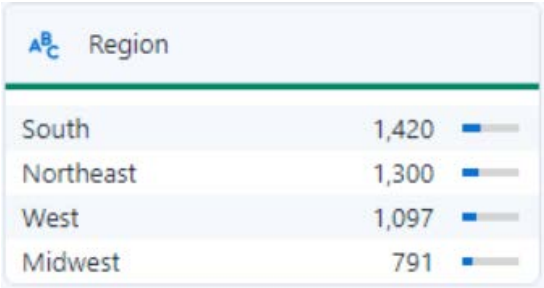
The “Product” field can only contain one of the six products. Browse the data in the “Product” field to verify that all the data is correct. Make any changes that are required.

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



Select Run.

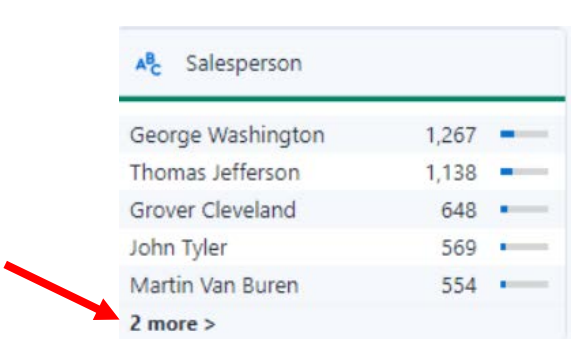
In the Configuration section scroll down to “Region”

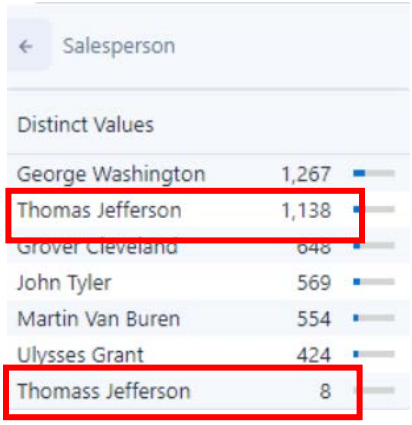


The data is clean, there are only four Regions listed.

The “Salesperson” field can only contain one of the six names. Browse the data in the “Salesperson” field to verify that all the data is correct. Make any changes that are required.

In the Configuration section scroll down to “Salesperson” and select “2 more”.

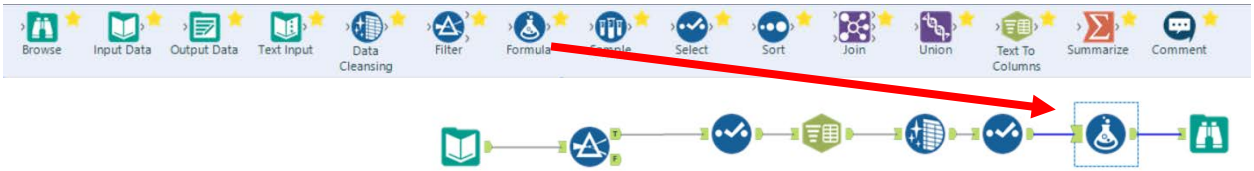




“Thomass Jefferson” must be replaced with “Thomas Jefferson” for a total of 1,146 distinct values.

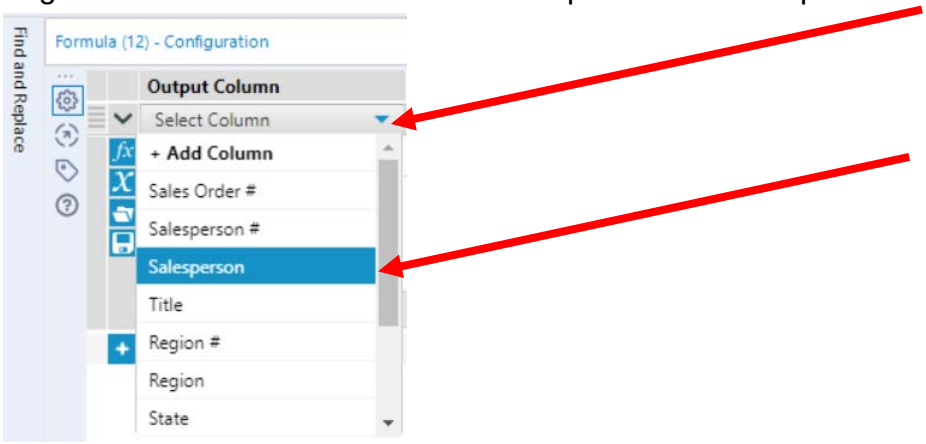
Tool 4: Formula Tool – The Data in the Fields Must Conform to Predefined Rules.

Drag the "**Formula Tool**" from the Tools Palette between the output of the "**Select Tool**" and the "Browse Tool" on the Canvas. The three will be connected automatically.

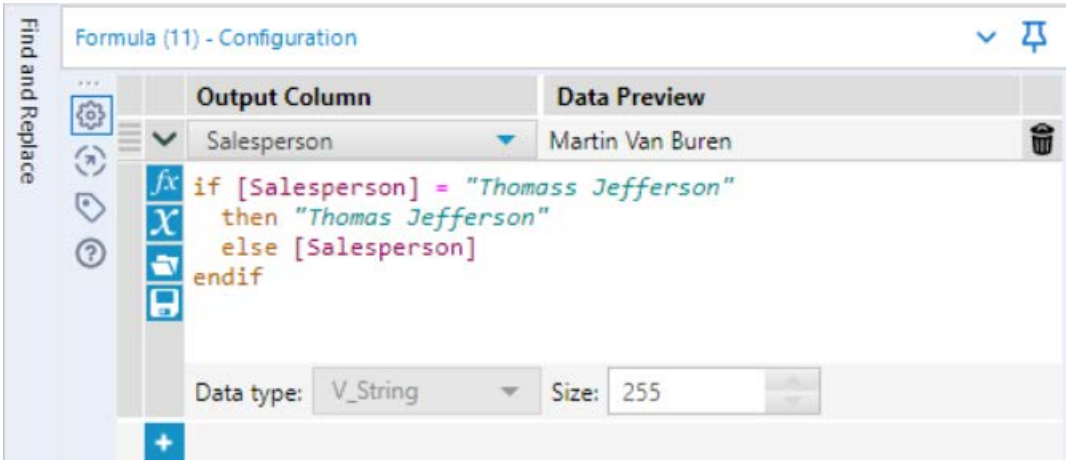


If the salesperson's name is "Thomass Jefferson" replace it with "Thomas Jefferson" otherwise leave the existing name in the field. Do not create a new field, put the output back into the "Salesperson" field.

In the configuration section select the field Salesperson from the pull-down list.

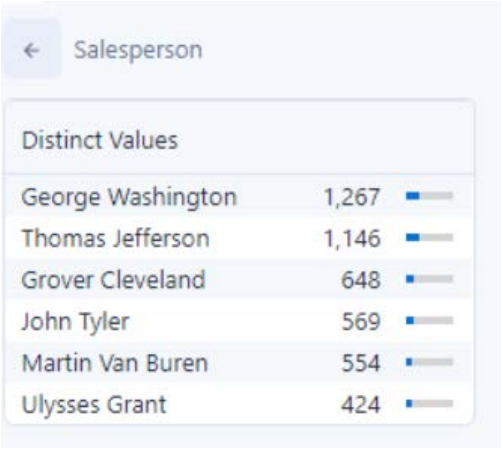


Type in the formula.



Select Run.

Select the Browse step.

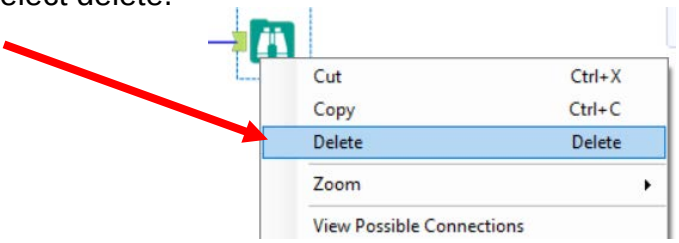


In the Configuration section scroll down to "Salesperson". There are only six salespersons and Thomas Jefferson has 1,146 distinct values.

While in the Browse step, verify the field "Product" and select "1 more".

The data is clean, there are only six Products listed.

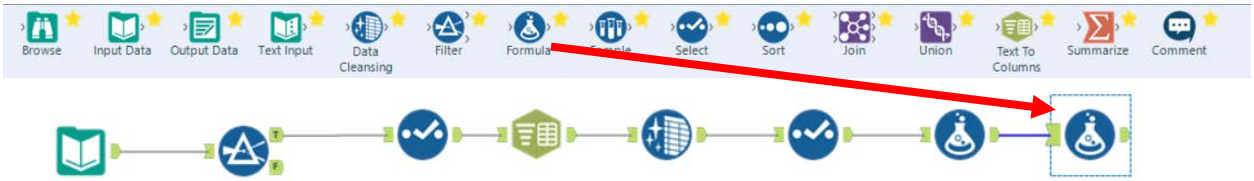
Right click on the Browse step and select delete.



The field "**Selling Price Per Unit**" does not exist in the "January Sales Data" stream of data. Instead, there is a field "Sales", the product of the "# of units" and the "Selling Price Per Unit".

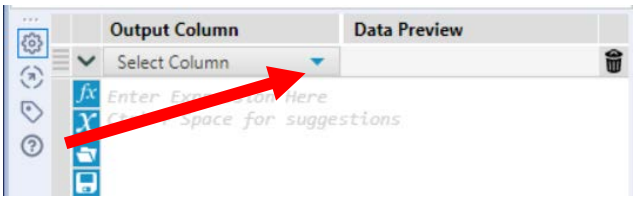
Tool 4: Formula Tool – Calculate Sales

Drag the "**Formula Tool**" from the Tools Palette between the output of the "**Select Tool**" and the "Browse Tool" on the Canvas. The three will be connected automatica

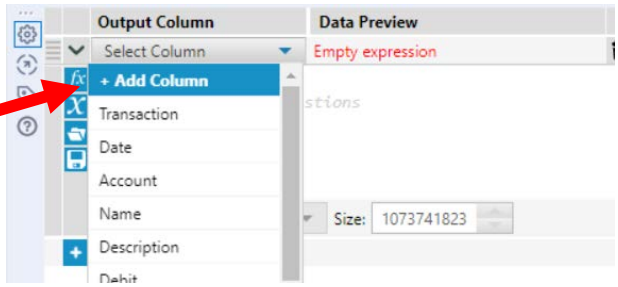


In the configuration section select the Add Column from the pull-down list.

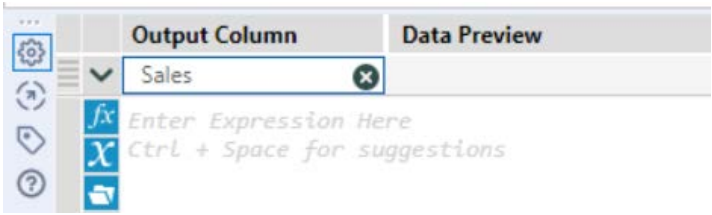
In the output Configuration area, click on the drop-down arrow of "Select Column".



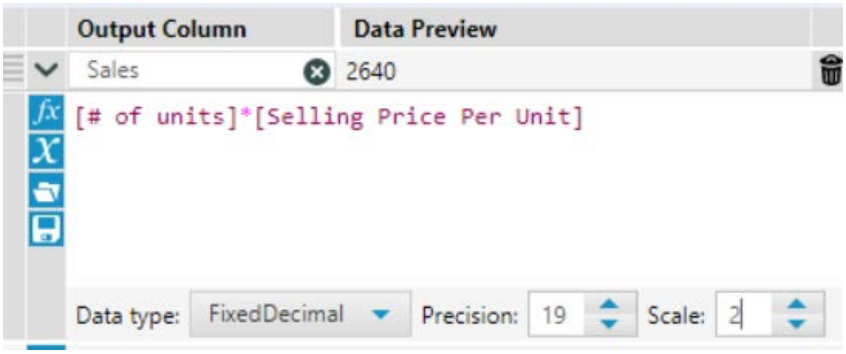
Select "Add Column".



Type "Sales".



Type the following formula [# of units]*[Selling Price Per unit].



Change the Data type to FixedDecimal with two decimal places.

Select “Run”.

# of units	Selling Price Per Unit	Profit Margin	Salesperson #	Salesperson	Sales
120.00	22.00	1590.00	1303	Ulysses Grant	2640.00
144.00	22.00	1908.00	1301	Thomas Jefferson	3168.00
96.00	22.00	1272.00	1303	Ulysses Grant	2112.00
60.00	22.00	795.00	1306	John Tyler	1320.00

The “Field Names” and “Types” of “February-December Sales Data” must be the same as those in the data set “January Sales Data”.

January Sales Data

	Field	Type	Size
▶	Sales Order #	Double	8
	Salesperson #	String	19
	Salesperson	V_String	255
	Title	V_String	255
	Region #	Double	8
	Region	V_String	255
	State	V_String	255
	Product	V_String	255
	# of units	FixedDecimal	10.0
	Sales	FixedDecimal	10.2
	Profit Margin	FixedDecimal	10.2

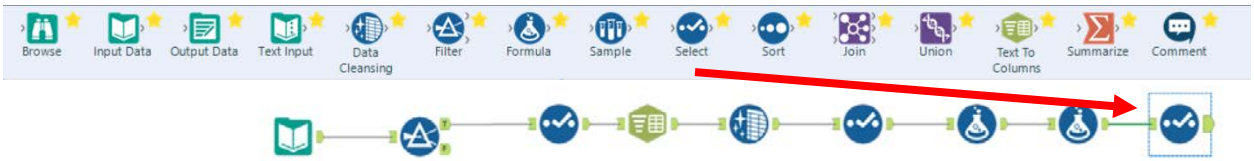
February-December Sales Data

	Field	Type	Size
▶	Sales Order #	Double	8
	Title	V_String	255
	Region	V_String	255
	Region #	Double	8
	State	V_String	255
	Product	V_String	255
	# of units	FixedDecimal	10.0
	Selling Price Per Unit	FixedDecimal	10.2
	Profit Margin	FixedDecimal	10.2
	Salesperson #	V_String	255
	Salesperson	V_String	255
	Sales	FixedDecimal	19.2
	*Unknown	Unknown	0

Rearrange the February-December Data fields to match the January Sales Data.

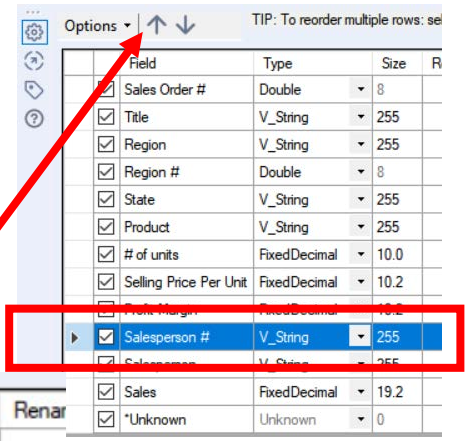
Tool 3: Select Tool – Change the order of fields and remove fields.

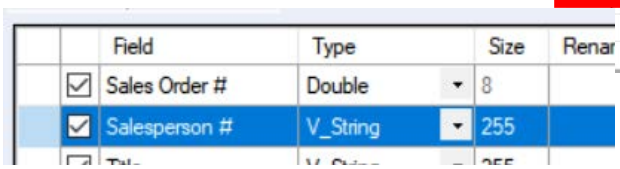
Drag the "**Select Tool**" from the Tools Palette to the True output of the "**Data Cleaning Tool**" on the Canvas. The two will be connected automatically.



Select "Salesperson #"

Press the up arrow to move "Salesperson #" to the second field.

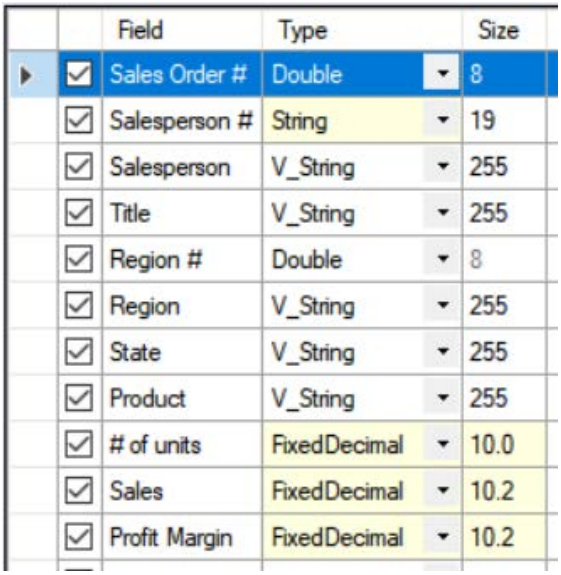




Change the type of the field "Salesperson #" to String with a size of 19.

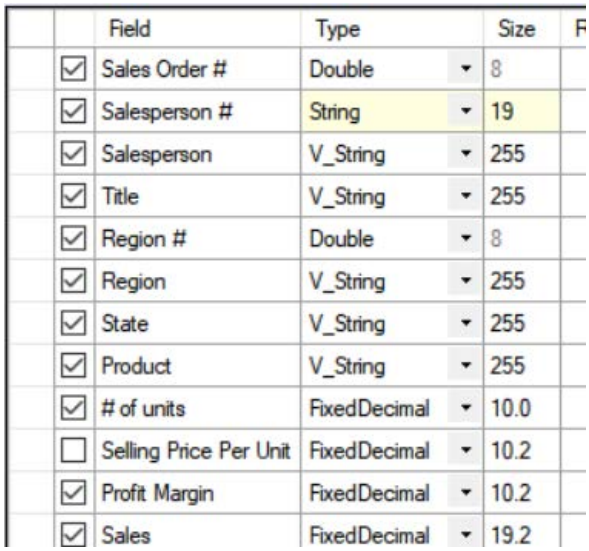
Rearrange the February-December Data fields to match the January Sales Data and remove the Selling Price Per Unit field.

January Sales Data



	Field	Type	Size
▶	Sales Order #	Double	8
	Salesperson #	String	19
	Salesperson	V_String	255
	Title	V_String	255
	Region #	Double	8
	Region	V_String	255
	State	V_String	255
	Product	V_String	255
	# of units	FixedDecimal	10.0
	Sales	FixedDecimal	10.2
	Profit Margin	FixedDecimal	10.2

February-December Sales Data



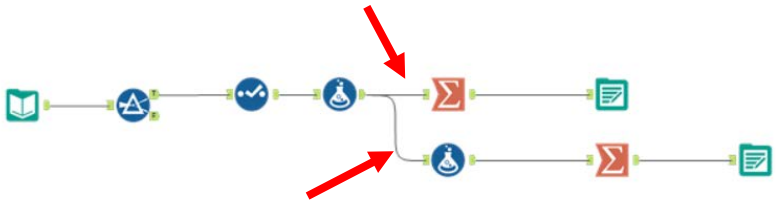
	Field	Type	Size	F
	Sales Order #	Double	8	
	Salesperson #	String	19	
	Salesperson	V_String	255	
	Title	V_String	255	
	Region #	Double	8	
	Region	V_String	255	
	State	V_String	255	
	Product	V_String	255	
	# of units	FixedDecimal	10.0	
	Selling Price Per Unit	FixedDecimal	10.2	
	Profit Margin	FixedDecimal	10.2	
	Sales	FixedDecimal	19.2	

Tool 4: Union Tool – Append Streams

Now that the data streams are the same, field names and types, we can append the second stream to the first stream. First separate the report section from stream one.

Right click between the Formula Tool and the Summarize Tool on the first stream and select “Delete Connection.

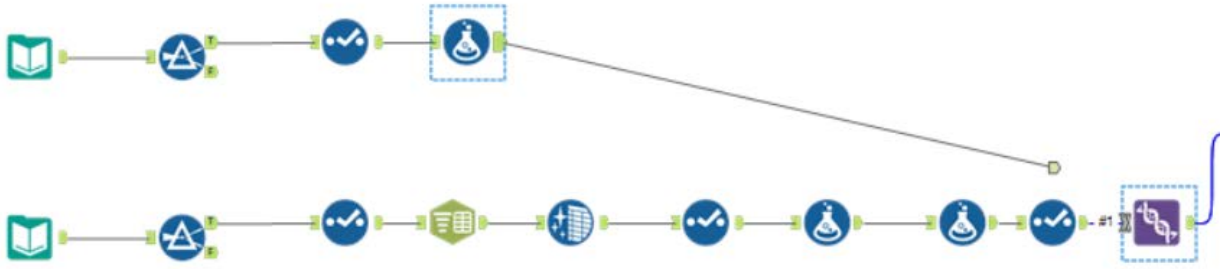
Right click between the Formula Tool and the Formula Tool on the first stream and select “Delete Connection.



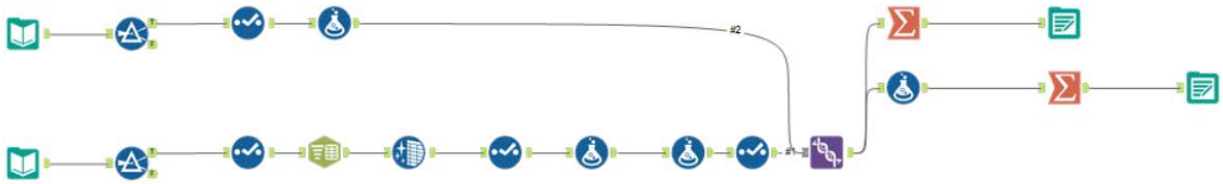
Drag the section to the right.

Drag the “**Union Tool**” from the Tools Palette to the output of the last “**Select Tool**” on the second stream Canvas and the Formula Tool from the report section. The three will be connected automatically.

Drag the output from the Formula Tool on the first stream to the Input of the Union Tool.



Drag the output from the Union Tool to the input of the first Summarize Tool in the report section.



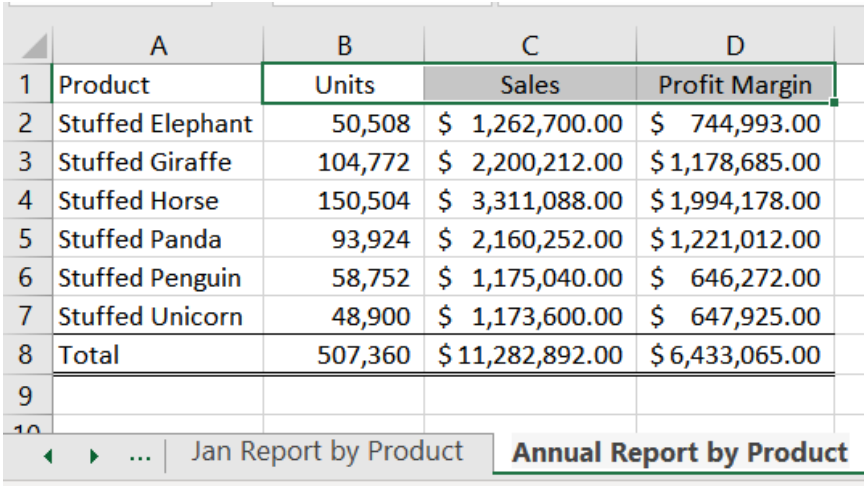
Select Run.
The combined stream has 5,017 rows.

Save the file.
Export the workflow.
Close Alteryx.
Open the Excel file to the Report by Product.

	A	B	C	D
1	Product	Units	Sales	Profit Margin
2	Stuffed Elephant	50508	1262700.00	744993.00
3	Stuffed Giraffe	104772	2200212.00	1178685.00
4	Stuffed Horse	150504	3311088.00	1994178.00
5	Stuffed Panda	93924	2160252.00	1221012.00
6	Stuffed Penguin	58752	1175040.00	646272.00
7	Stuffed Unicorn	48900	1173600.00	647925.00
8				

... Report by Product | Jan Report by Salesp ..

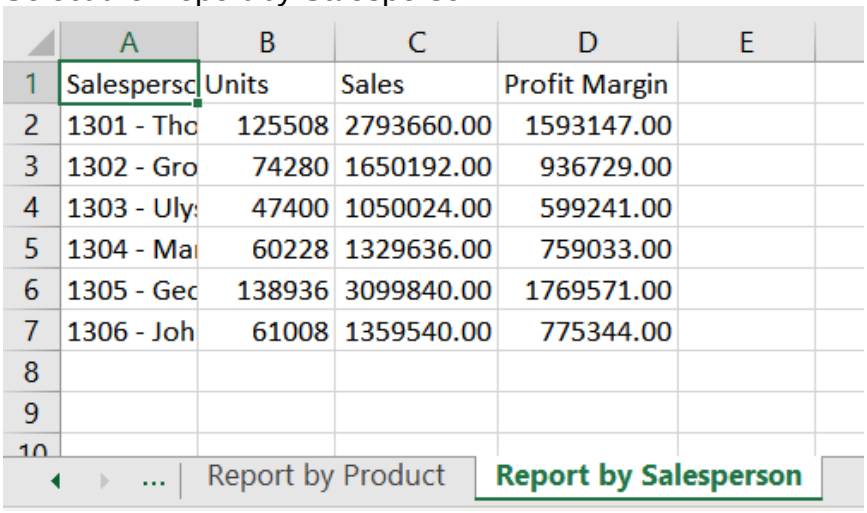
Fix the format and rename the sheet.



	A	B	C	D
1	Product	Units	Sales	Profit Margin
2	Stuffed Elephant	50,508	$ 1,262,700.00	$ 744,993.00
3	Stuffed Giraffe	104,772	$ 2,200,212.00	$ 1,178,685.00
4	Stuffed Horse	150,504	$ 3,311,088.00	$ 1,994,178.00
5	Stuffed Panda	93,924	$ 2,160,252.00	$ 1,221,012.00
6	Stuffed Penguin	58,752	$ 1,175,040.00	$ 646,272.00
7	Stuffed Unicorn	48,900	$ 1,173,600.00	$ 647,925.00
8	Total	507,360	$ 11,282,892.00	$ 6,433,065.00
9				

... | Jan Report by Product | **Annual Report by Product**

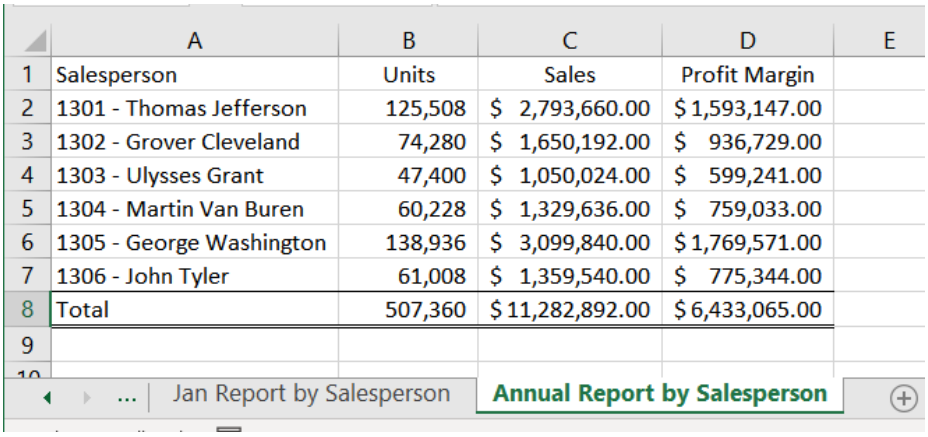
Select the Report by Salesperson.



	A	B	C	D	E
1	Salesperso	Units	Sales	Profit Margin	
2	1301 - Tho	125508	2793660.00	1593147.00	
3	1302 - Gro	74280	1650192.00	936729.00	
4	1303 - Uly	47400	1050024.00	599241.00	
5	1304 - Mar	60228	1329636.00	759033.00	
6	1305 - Geo	138936	3099840.00	1769571.00	
7	1306 - Joh	61008	1359540.00	775344.00	
8					
9					

... | Report by Product | **Report by Salesperson**

Fix the format and rename the sheet.



	A	B	C	D	E
1	Salesperson	Units	Sales	Profit Margin	
2	1301 - Thomas Jefferson	125,508	$ 2,793,660.00	$ 1,593,147.00	
3	1302 - Grover Cleveland	74,280	$ 1,650,192.00	$ 936,729.00	
4	1303 - Ulysses Grant	47,400	$ 1,050,024.00	$ 599,241.00	
5	1304 - Martin Van Buren	60,228	$ 1,329,636.00	$ 759,033.00	
6	1305 - George Washington	138,936	$ 3,099,840.00	$ 1,769,571.00	
7	1306 - John Tyler	61,008	$ 1,359,540.00	$ 775,344.00	
8	Total	507,360	$ 11,282,892.00	$ 6,433,065.00	
9					

... | Jan Report by Salesperson | **Annual Report by Salesperson**

Save and close Excel.