A GUIDE TO EXCEL & FME

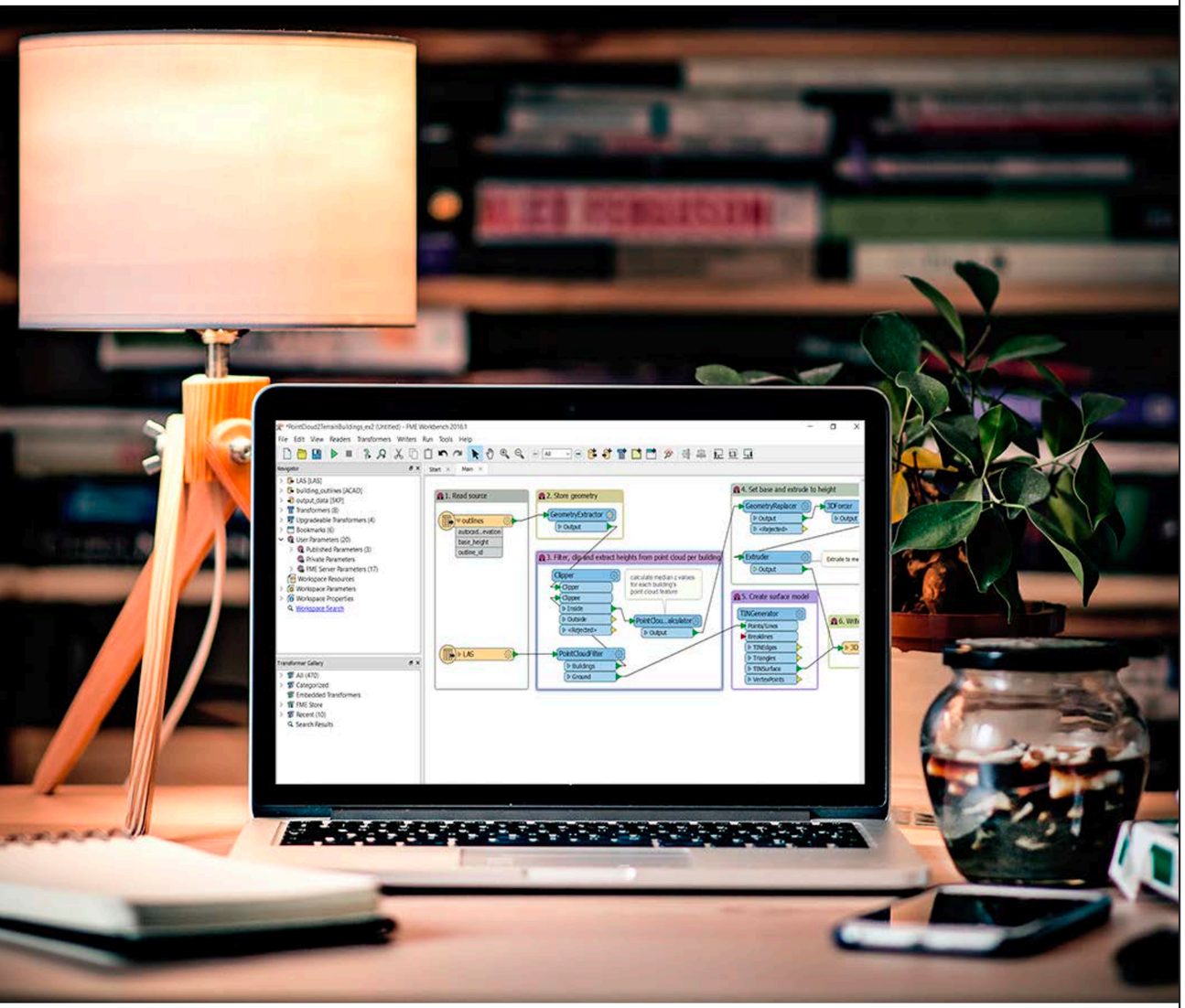
TABLE OF CONTENTS

Why Process Excel Data with FME?

Excel is everywhere in most organizations. Everyone is comfortable using it, everyone has it installed and most applications can import and export XLS data. It's relatively straightforward to use and if you're only working with a small amount of data across one or a few worksheets, there's no need to bring FME into the process.

FME is an asset in situations where you're working with **LARGE** amounts of data across multiple sheets, and many manual tasks are required. Copying, pasting and entering functions by hand carries the risk of introducing error and can be time consuming. FME's automated workflows can execute the same functions as Excel - filtering, sorting, conditional testing - to protect the data's integrity and save you time.

FME's integration platform makes it simple to connect hundreds of systems, transform data in unlimited ways, and automate workflows.



GETTING STARTED WITH EXCEL IN FME

A Note on Language in FME versus Excel

When working in FME and reading its documentation keep these equivalencies in mind:

Feature Type = Worksheet or Named Range

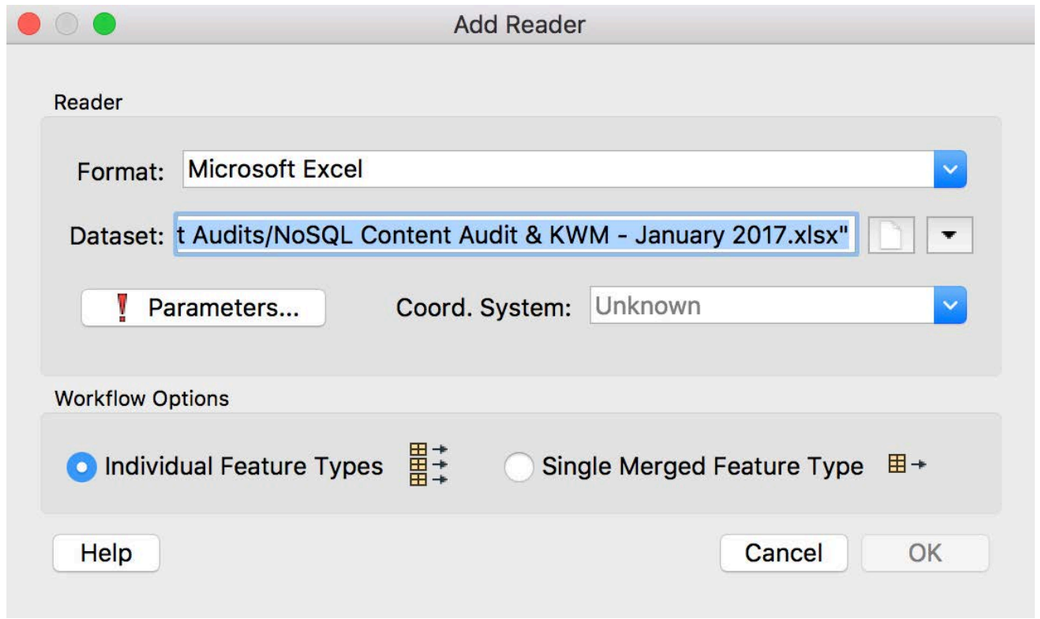
Attribute = Column

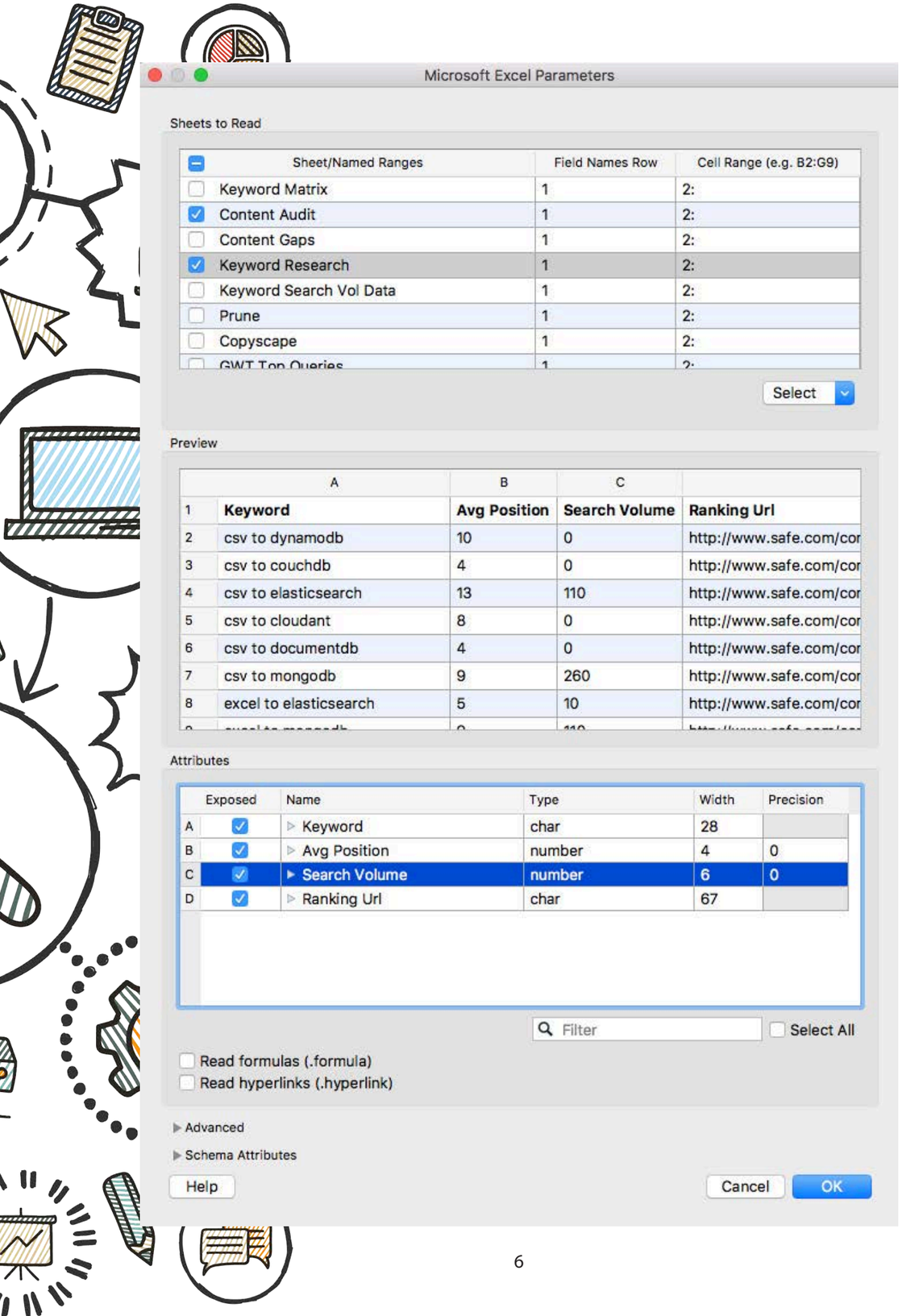
Feature = Row

Dataset = Excel File

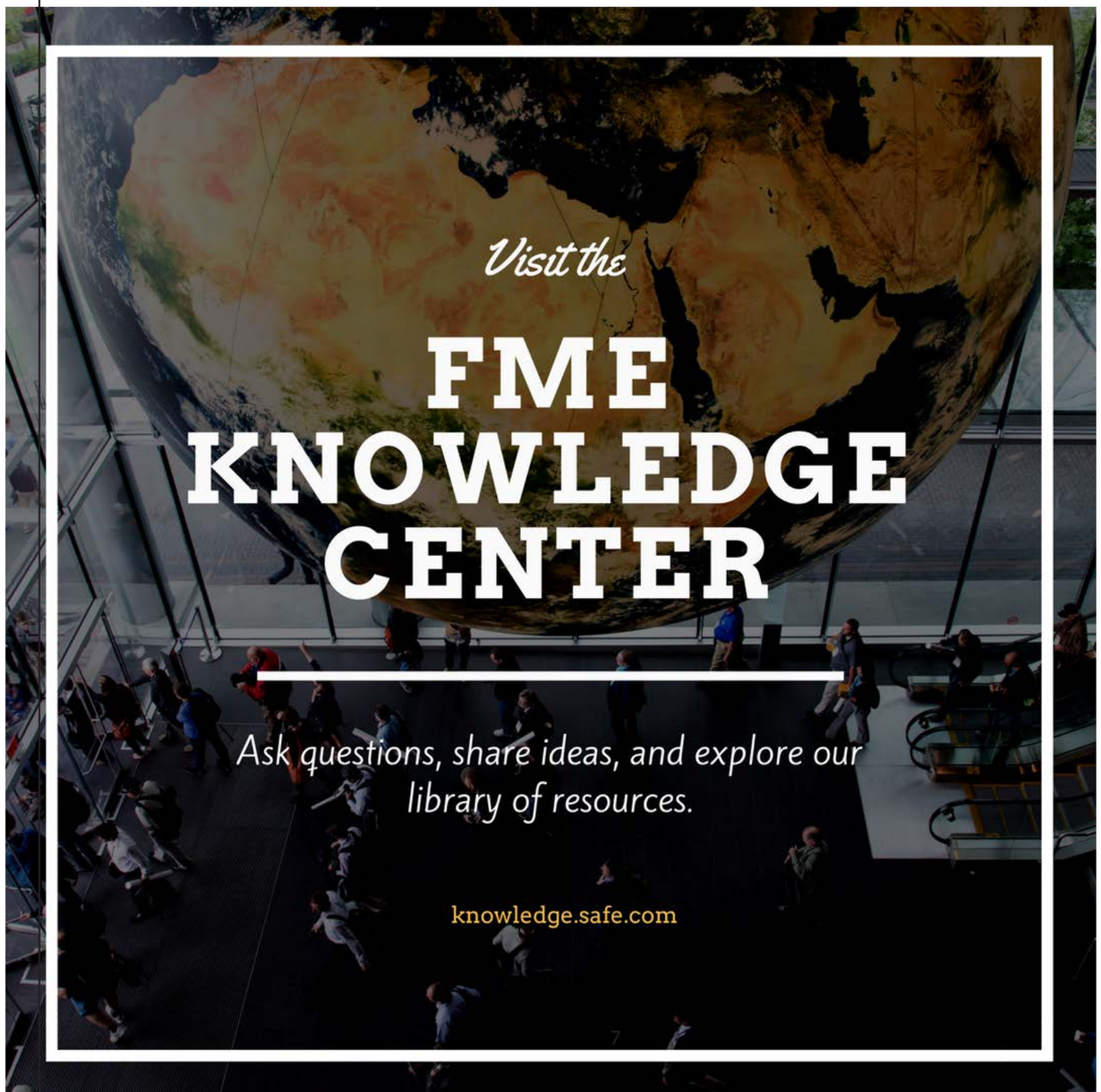
Opening Excel Files

Use the “Add Reader” tool to open Excel files with FME. In the “Add Reader” dialog, select Microsoft Excel as the format and locate the desired file. After this open the Parameters dialog to select sheets to import and to inspect the data.

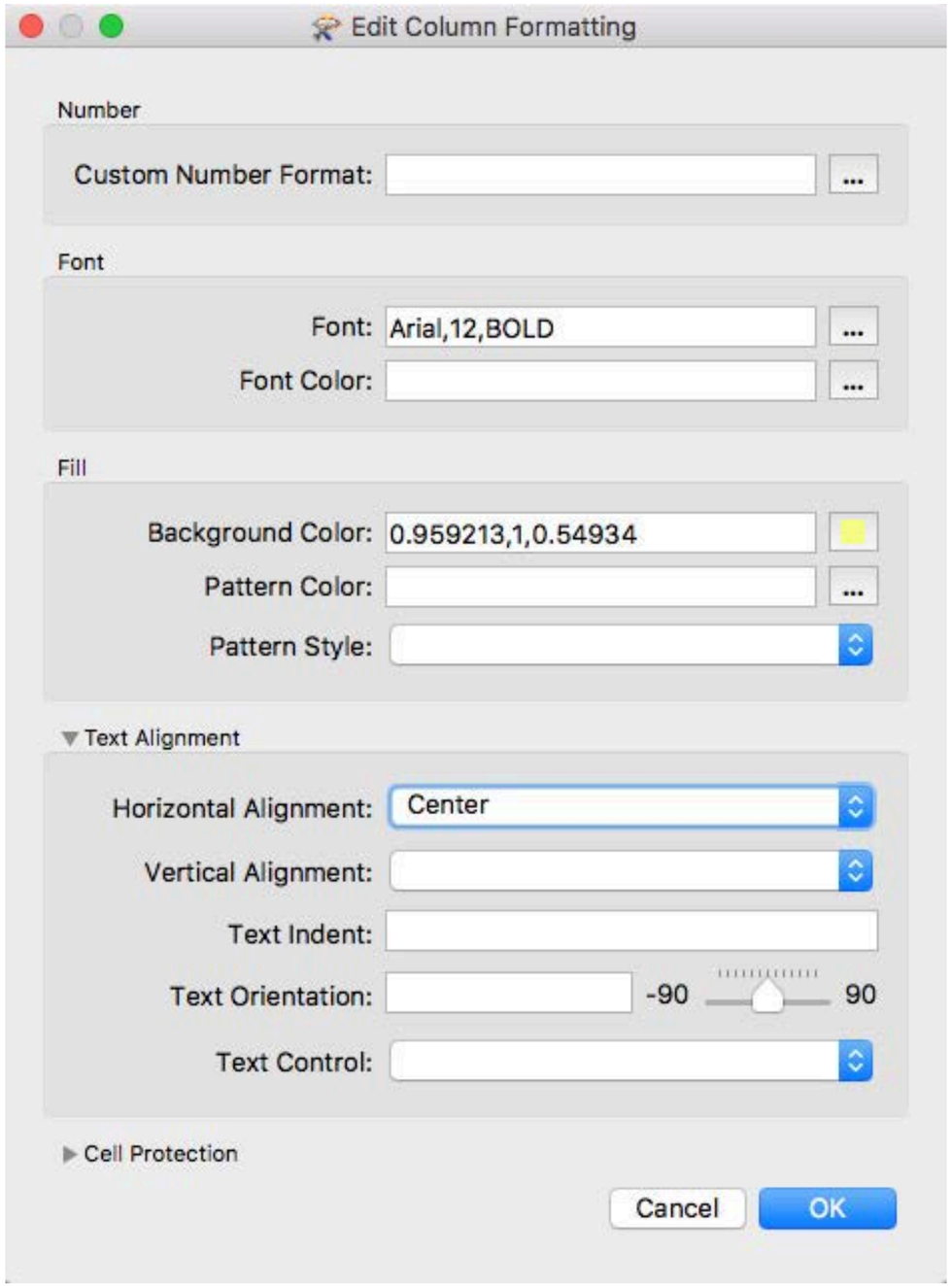




Select the worksheets to be imported in the “Sheets to Read” section. When you highlight a worksheet you get a preview of how the data looks, and can inspect and set data types in the “Atributes” section. You can also set the reader to recognize and preserve formulas and hyperlinks.



Set style elements for fonts and cells in a column by clicking the corresponding space under “Formatting”.

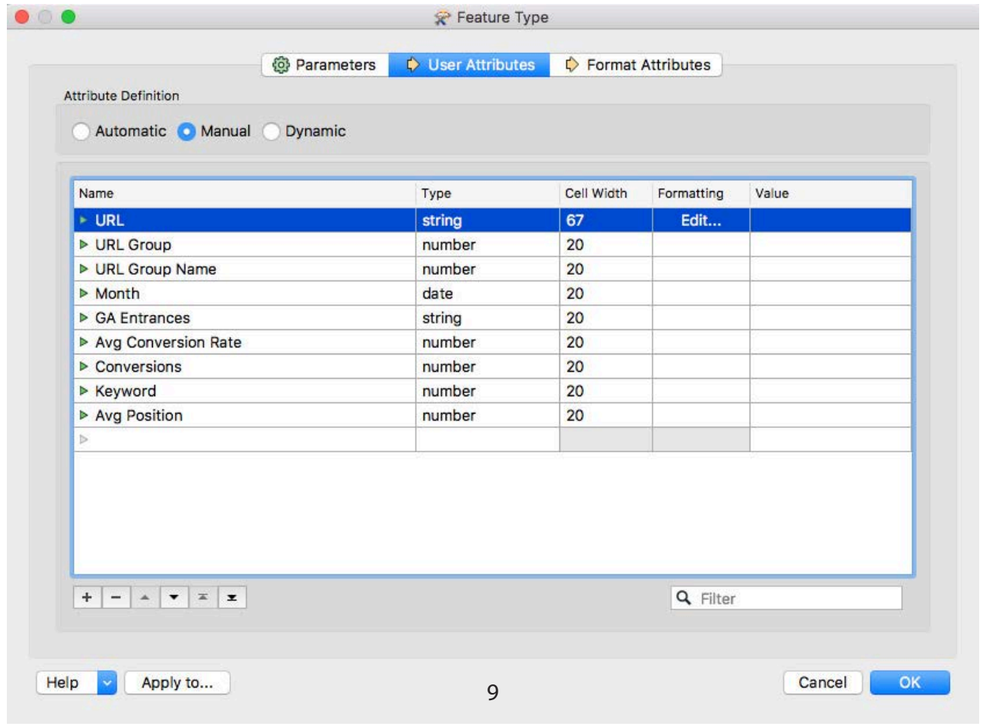


CREATING EXCEL FILES

FME lets you construct worksheets within Workbench or write data to an existing Excel template. In each case you begin by adding an Excel writer using the “Add Writer” tool. Set the format to Microsoft Excel and specify where the data will be written to and what the name of the file will be. In the case of writing data to a template, select the template file as the destination.

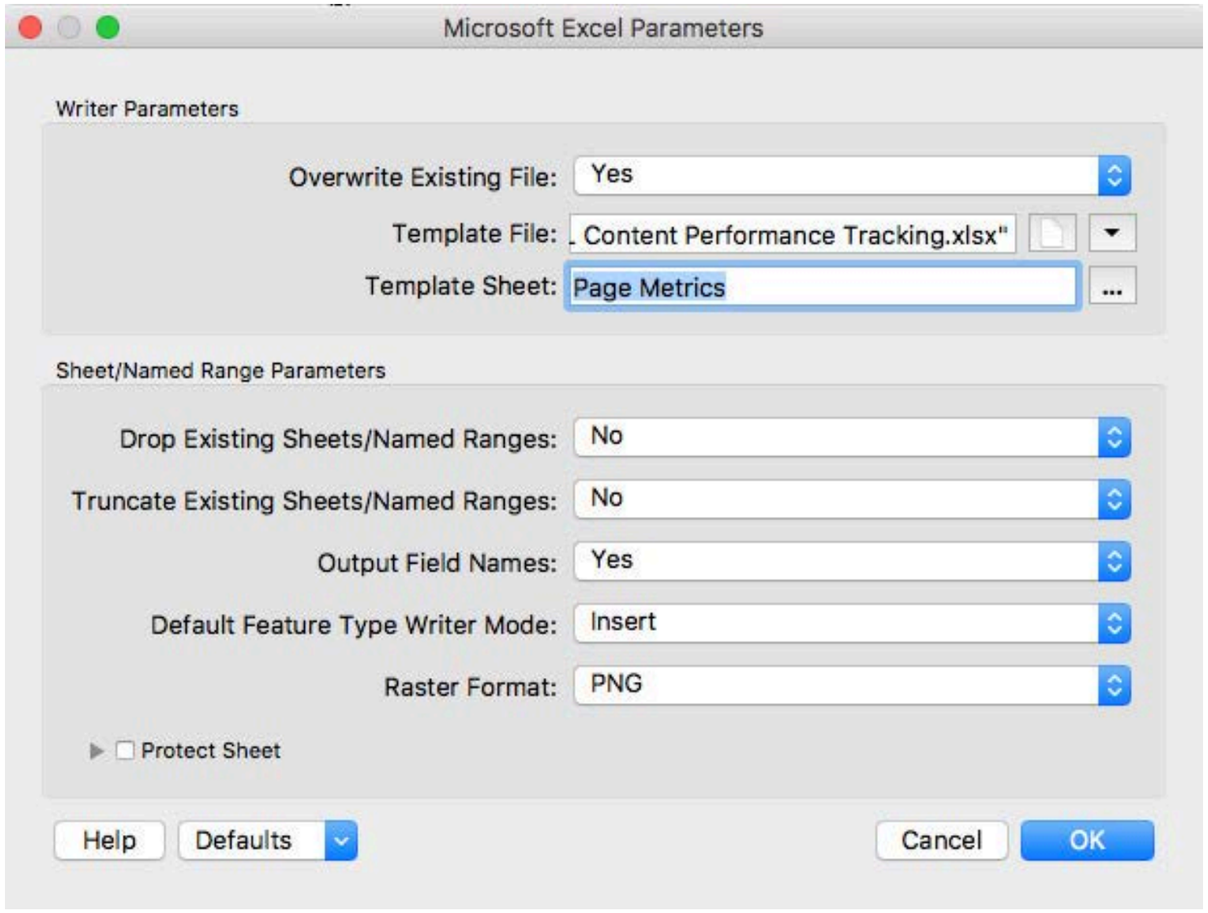
Constructing an Excel spreadsheet within FME Workbench

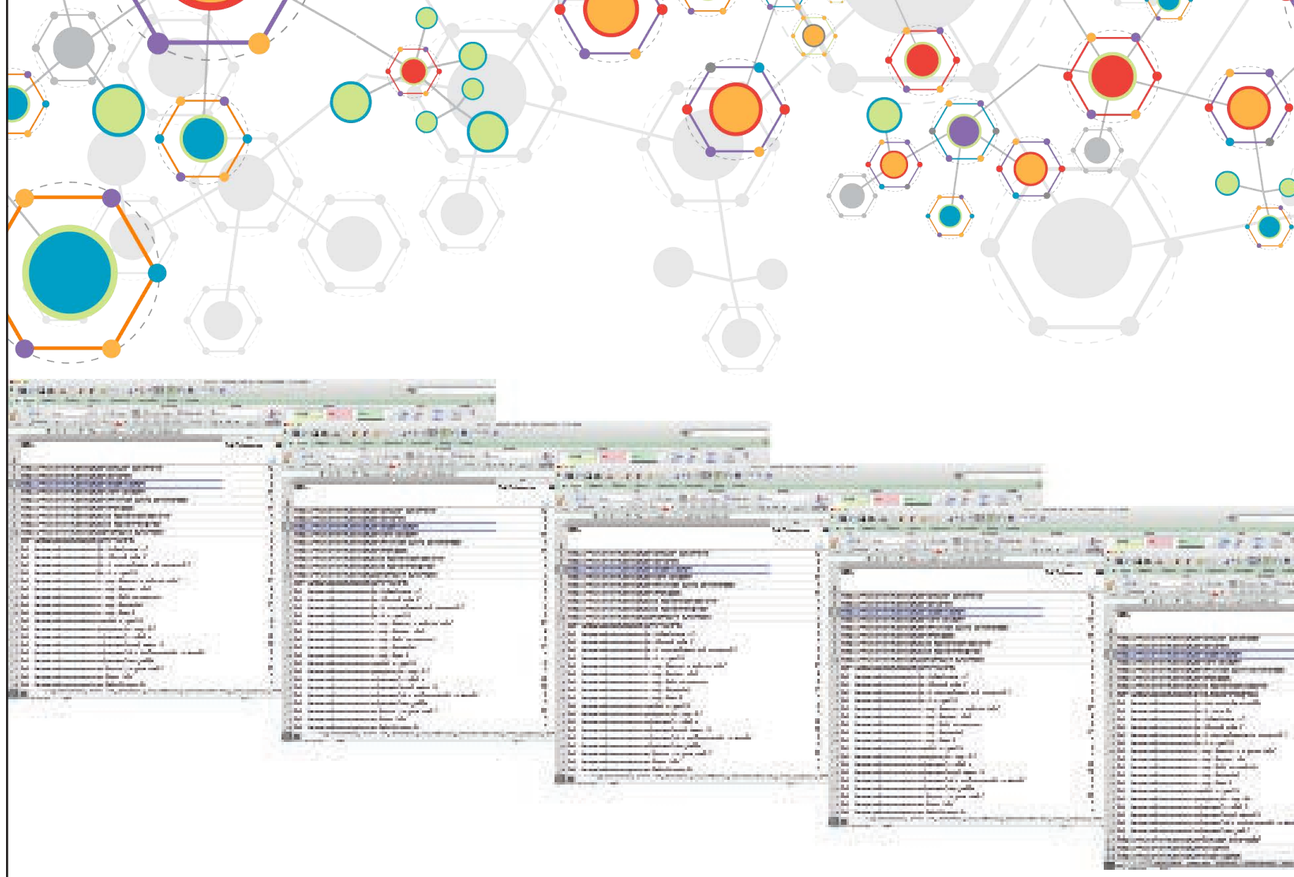
After the Writer has been added to the workspace, open up its dialog. Columns are created in the “User Attributes” section. Set “Attribute Definition” to “Manual” and enter the names of the columns as you’d like them to appear in the final spreadsheet and indicate the type of data. These columns will be populated when the workflow is run with values from Attributes with the same name. Working with Attributes (columns) will be discussed further later on.



Writing to an Excel Template

In the case of writing data to a template, you have an Excel file whose data needs to be refreshed from time to time to reflect changes. This is useful when generating reports that summarize raw data into meaningful charts and tables. The sheet that holds the definition of the layout of results is designated as the template in the FME Writer's Parameter dialog. Set "Overwrite Existing File" to "Yes" so that the data is overwritten and not appended.



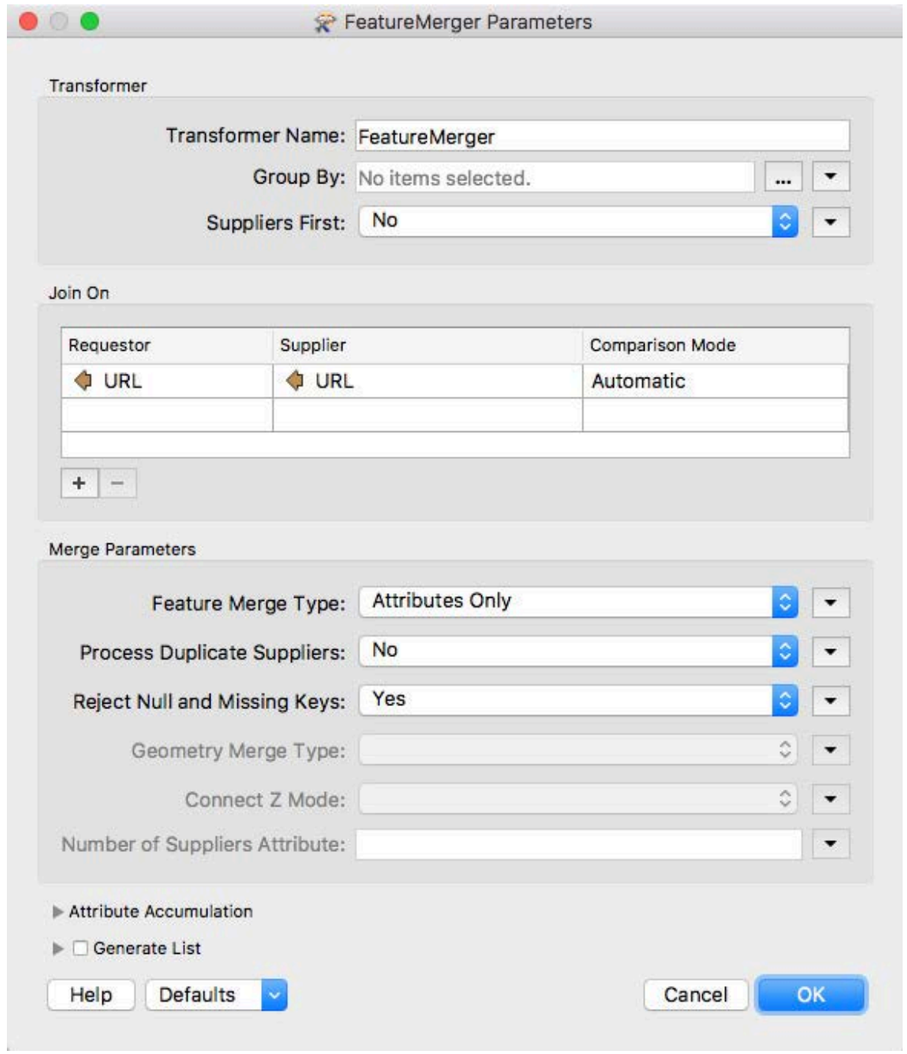


INTEGRATING DATA FROM MULTIPLE WORKSHEETS

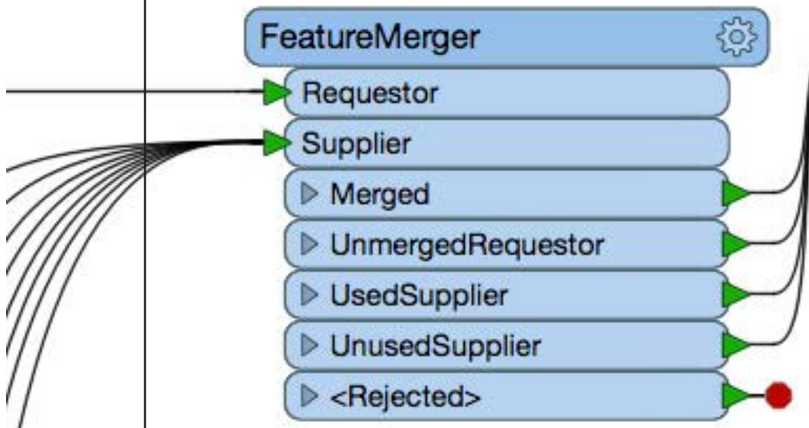
Appending Multiple Worksheets with the Same Structure

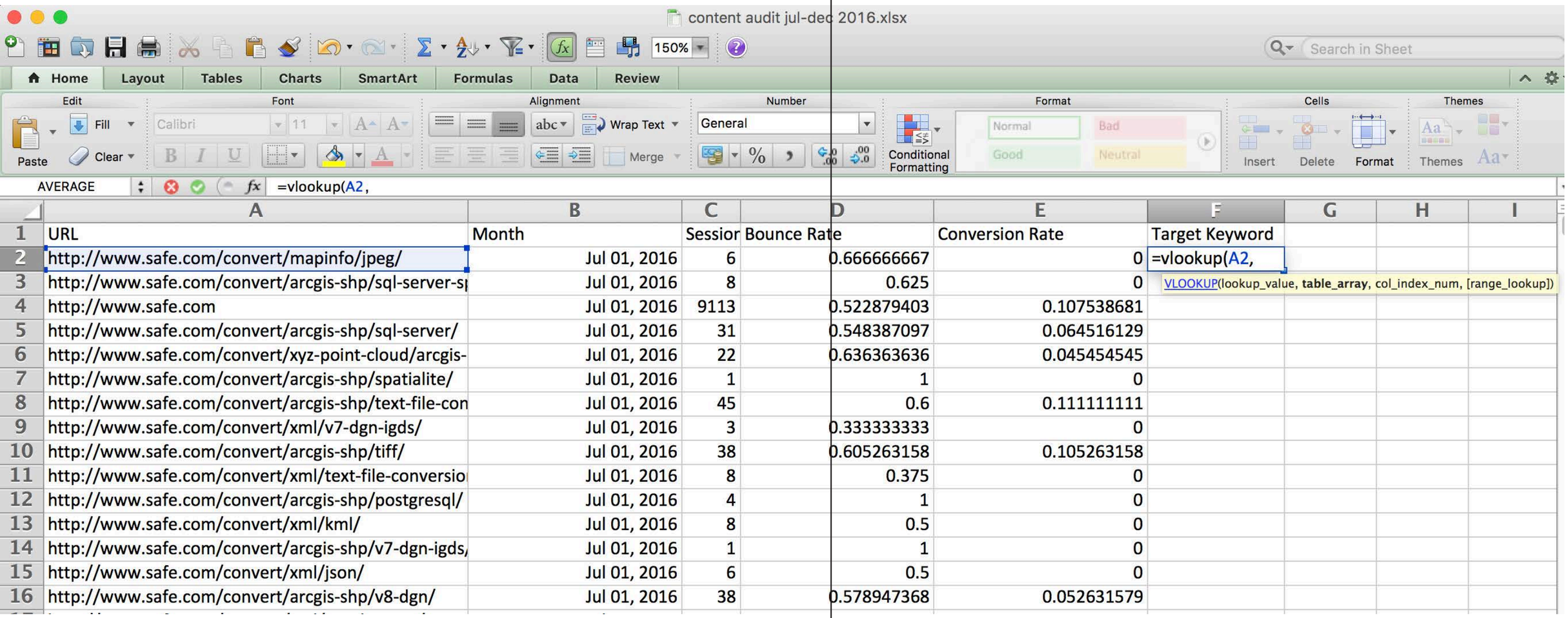
Appending worksheets is a common task with Excel data, especially for projects that involve analyzing data collected at different times, by different people, or in different places.

In the FeatureMerger dialog, indicate which field the sheets should be joined on and set “Process Duplicate Suppliers” to “Yes”.



This tool categorizes data into output ports based on whether it found a match between the Requestor and Suppliers: “Merged”, “UnmergedRequestor”, “UsedSupplier”, and “UnusedSupplier”. This comes in handy for other tasks like Vlookups, but when merging worksheets you want all the data carried over, so connect all four output ports to the Excel Writer, or whichever transformer tool comes next in the workflow.



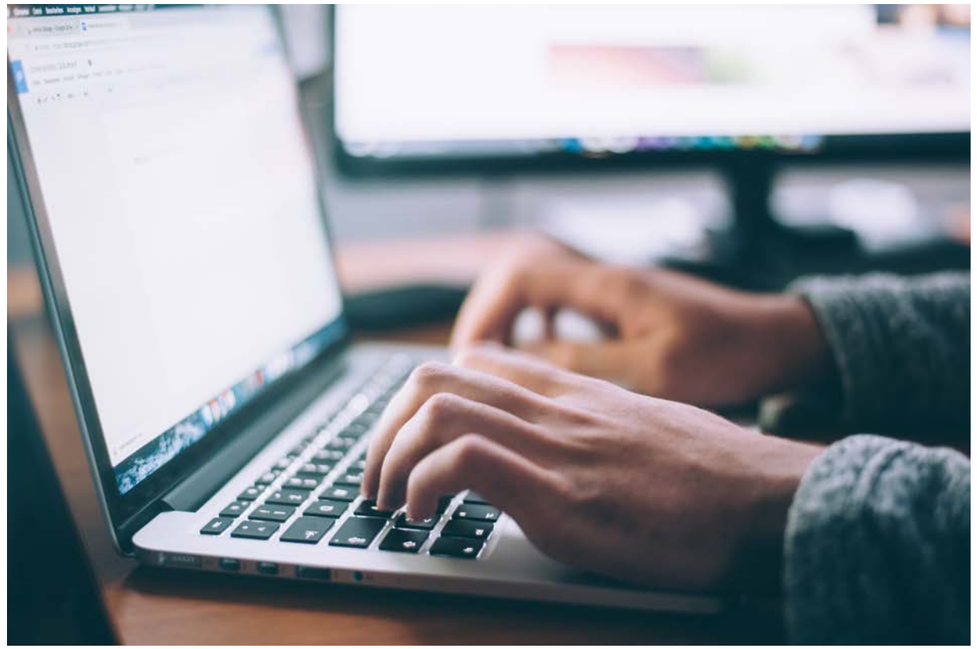


	A	B	C	D	E	F	G	H	I
1	URL	Month	Session	Bounce Rate	Conversion Rate	Target Keyword			
2	http://www.safe.com/convert/mapinfo/jpeg/	Jul 01, 2016	6	0.666666667	0	=vlookup(A2,			
3	http://www.safe.com/convert/arcgis-shp/sql-server-sp	Jul 01, 2016	8	0.625	0				
4	http://www.safe.com	Jul 01, 2016	9113	0.522879403	0.107538681				
5	http://www.safe.com/convert/arcgis-shp/sql-server/	Jul 01, 2016	31	0.548387097	0.064516129				
6	http://www.safe.com/convert/xyz-point-cloud/arcgis-	Jul 01, 2016	22	0.636363636	0.045454545				
7	http://www.safe.com/convert/arcgis-shp/spatialite/	Jul 01, 2016	1	1	0				
8	http://www.safe.com/convert/arcgis-shp/text-file-con	Jul 01, 2016	45	0.6	0.111111111				
9	http://www.safe.com/convert/xml/v7-dgn-igds/	Jul 01, 2016	3	0.333333333	0				
10	http://www.safe.com/convert/arcgis-shp/tiff/	Jul 01, 2016	38	0.605263158	0.105263158				
11	http://www.safe.com/convert/xml/text-file-conversio	Jul 01, 2016	8	0.375	0				
12	http://www.safe.com/convert/arcgis-shp/postgresql/	Jul 01, 2016	4	1	0				
13	http://www.safe.com/convert/xml/kml/	Jul 01, 2016	8	0.5	0				
14	http://www.safe.com/convert/arcgis-shp/v7-dgn-igds/	Jul 01, 2016	1	1	0				
15	http://www.safe.com/convert/xml/json/	Jul 01, 2016	6	0.5	0				
16	http://www.safe.com/convert/arcgis-shp/v8-dgn/	Jul 01, 2016	38	0.578947368	0.052631579				

Running a Vlookup with FME

Vlookup is a handy function in Excel that allows you to search a list for an item then return an associated value when a match is found. For example, in the NoSQL web page example used in this guide, target keyword data is returned by searching a separate worksheet using the common URL column. In Excel, this is accomplished by entering the vlookup function and parameters into a cell. In FME, the same result is achieved using the FeatureMerger.

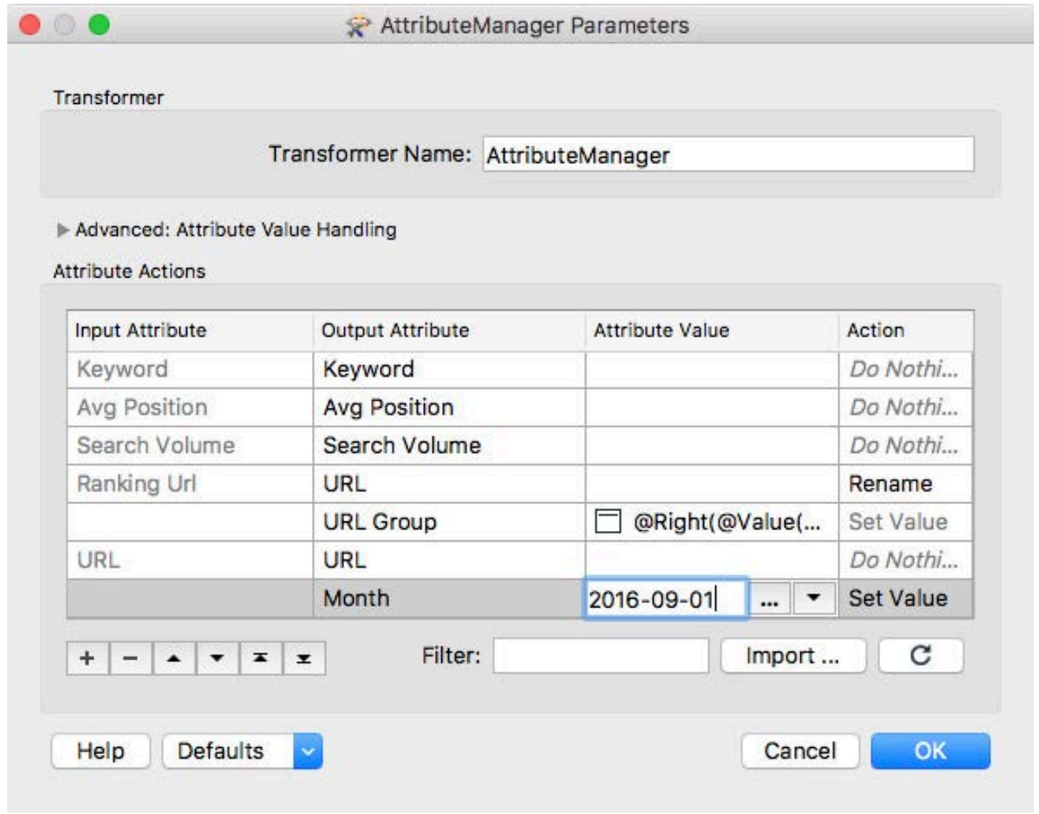
The lookup value between Requestor and Suppliers is indicated with the “Join On” parameter in the FeatureMerger dialog. In this situation you’re only interested in matches between two worksheets so only the “Merged” output port of the FeatureMerger is connected to the Excel Writer (or next transformer in the workflow). The FeatureMerger will return all associated values for a match. Specify which values are to be written when you configure the column names in the Excel Writer dialog.



MANIPULATING EXCEL DATA & STRUCTURES

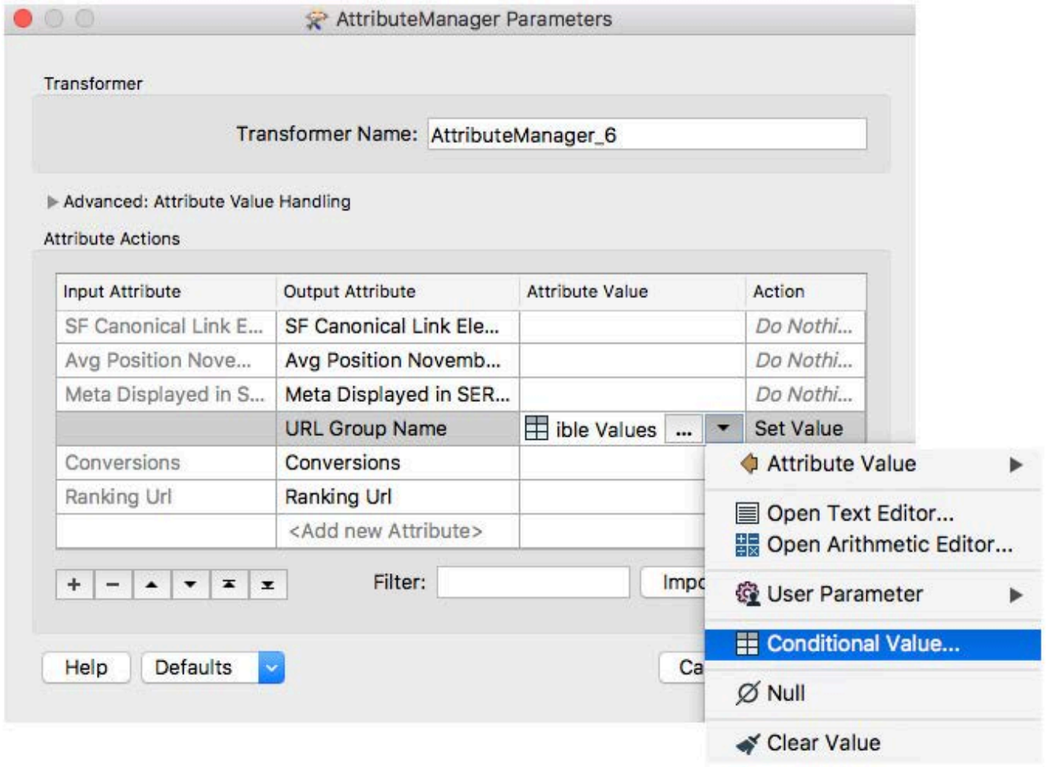
Adding Columns & Assigning Values

You may want to add columns while working with Excel data in FME. For instance, if you're merging multiple worksheets and you wish to add a unique identifier like a date or location to each, or perhaps you want to derive values based on the data in other columns. In FME this is accomplished with the AttributeManager. You can either assign a constant value to a new column or derive values using functions or conditional statements.

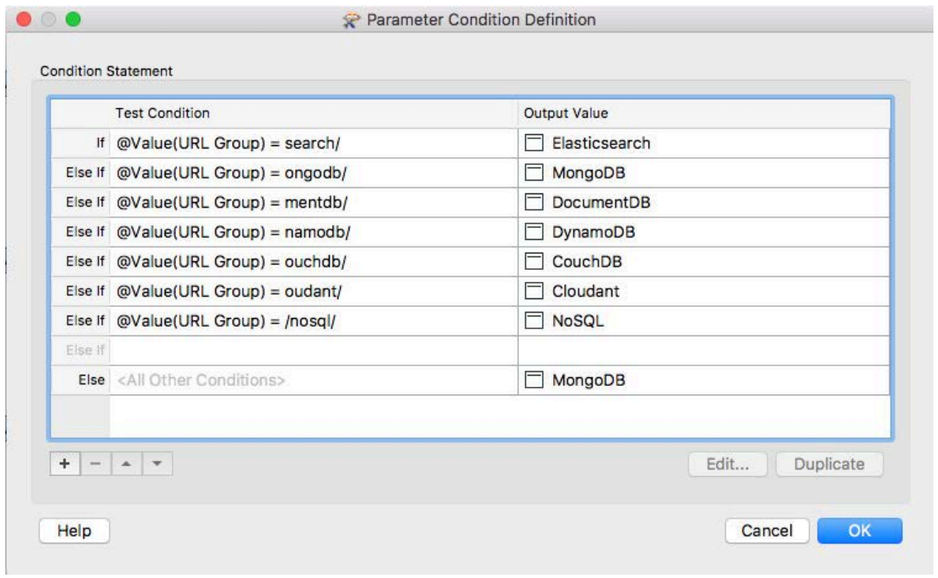


Example: Assigning a constant value

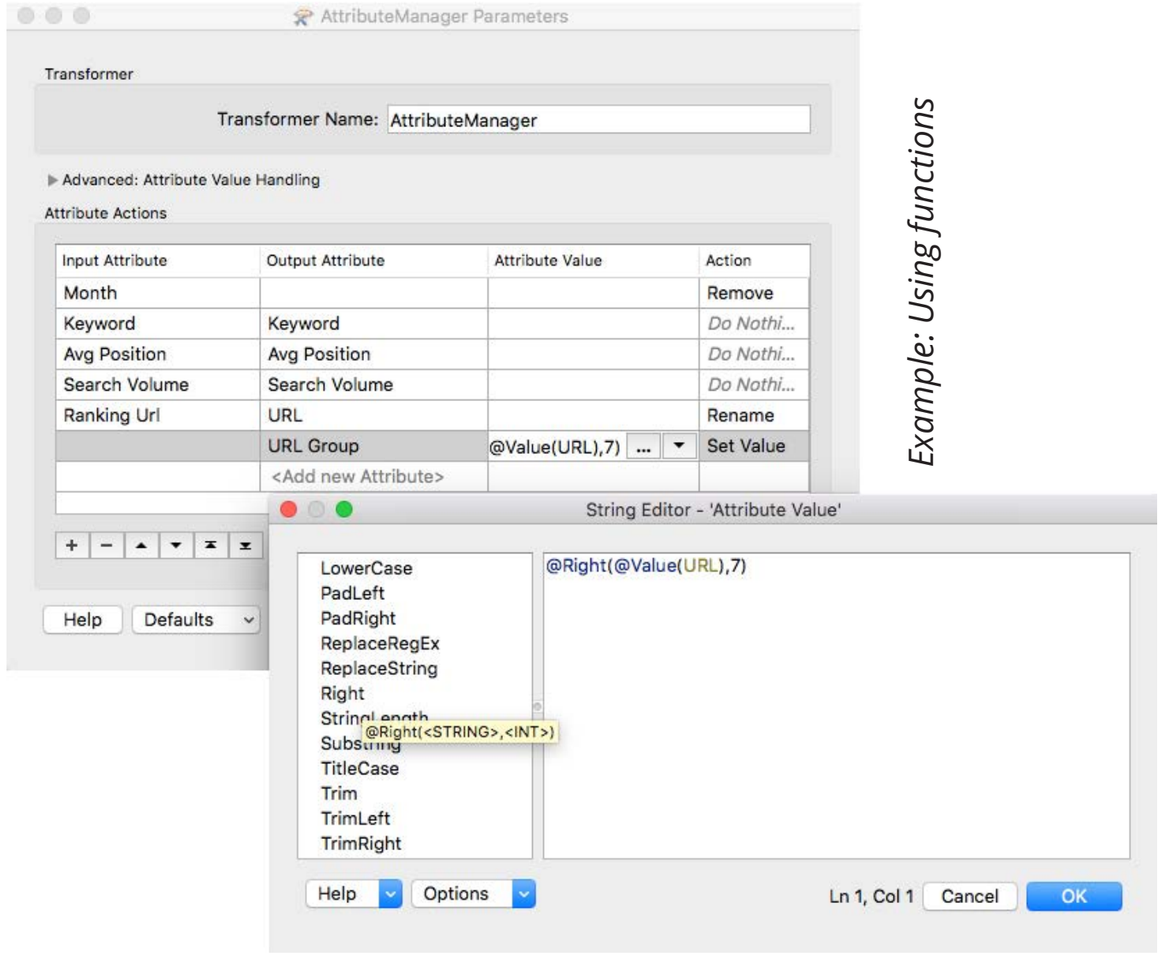
In this scenario you'll need to use a separate AttributeManager for each worksheet imported into the Workspace.



Example: Using conditional statements.

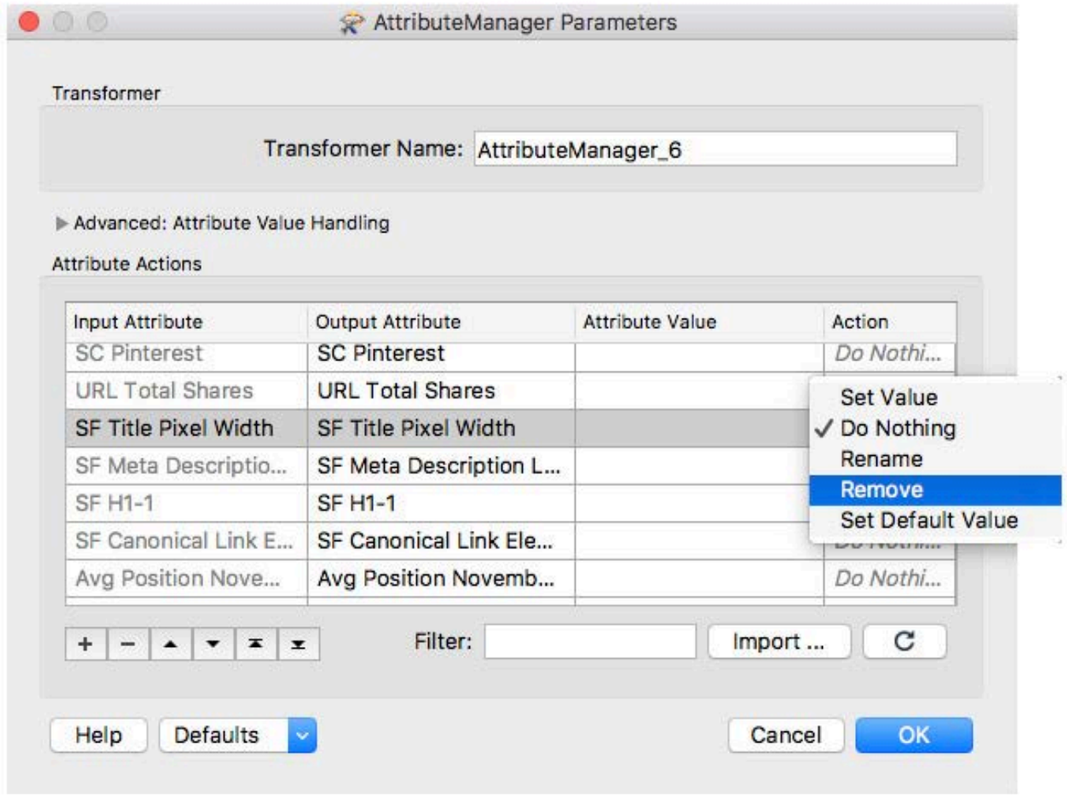


Example: Using functions

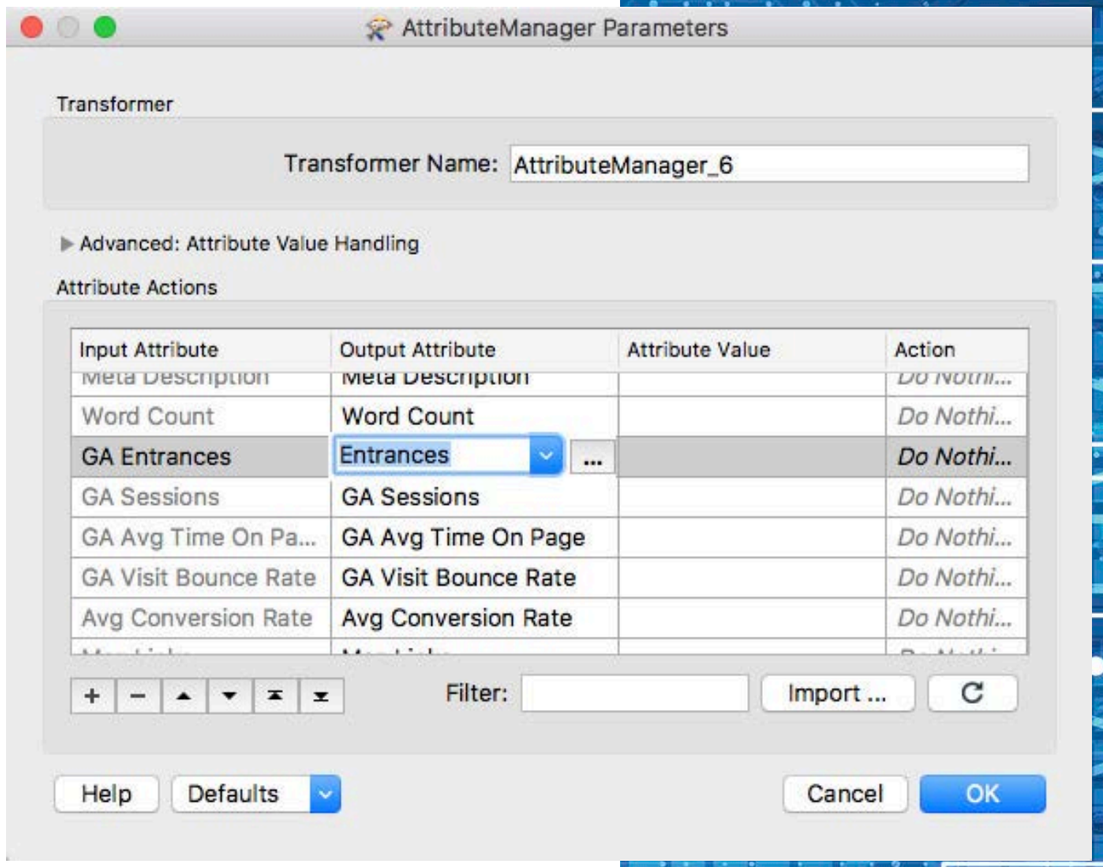


Removing, Renaming & Reordering Columns

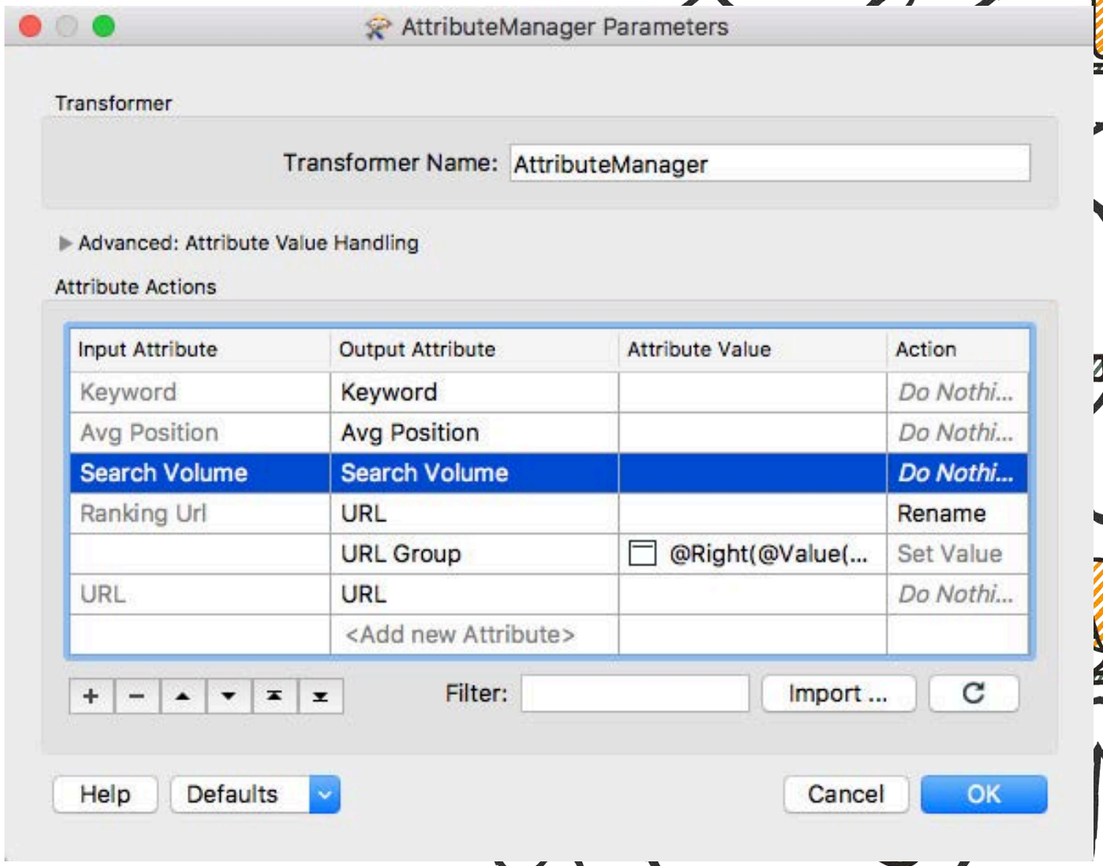
The AttributeManager can also be used to remove, rename and reorder columns.



Example: Remove



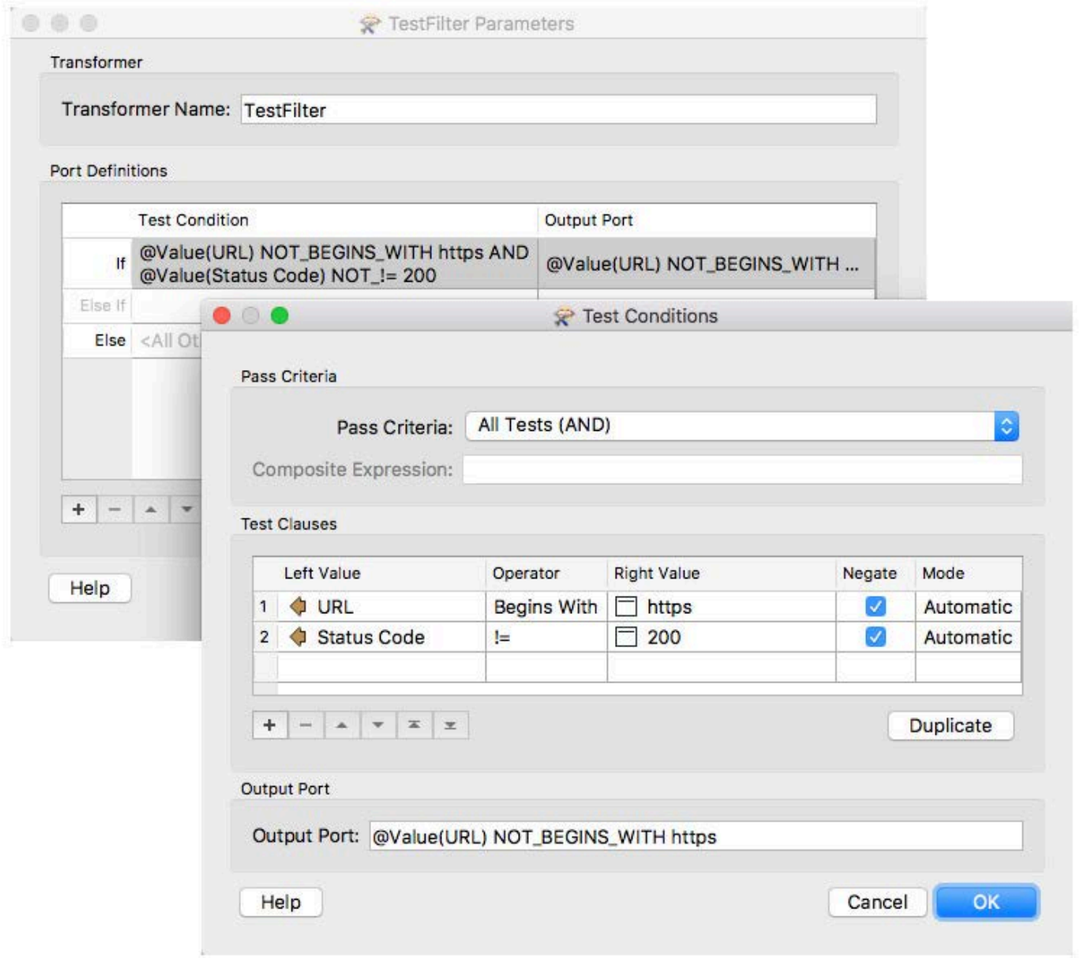
Example: Rename



Example: Reorder

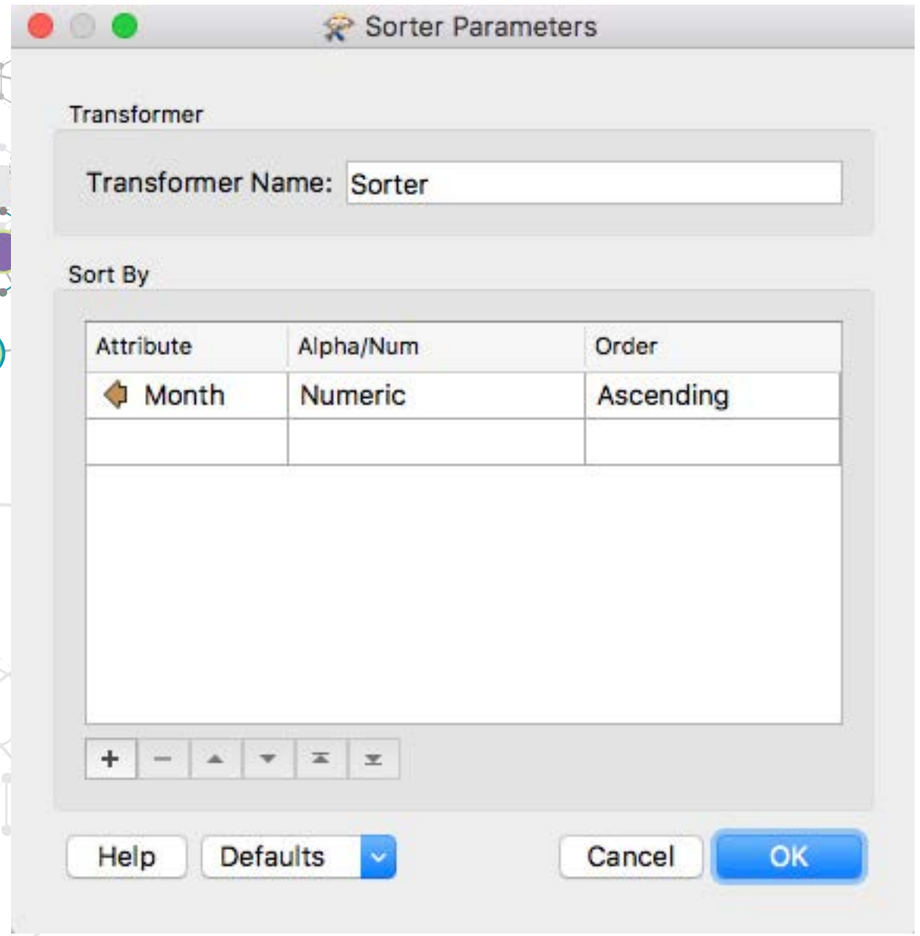
Filtering

Easy enough in Excel - just use the filter tool. But what if you're interested in a subset of data that is spread across many worksheets? FME can be used to merge these worksheets then filter down to only the data you need using the TestFilter. This tool filters rows using test conditions that direct data that pass to one or more output ports, and those that fail to another. It is worth noting that FME's Tester can accomplish a similar result using a simple pass/fail test.



Sorting

FME has a Sorter transformer that enables you to sort rows in a very similar way as Excel. Specify the column (attribute) to sort by and indicate whether the sort is alphabetical or numeric, ascending or decreasing. You can also sort by more than one column.



Restructuring Tables

Sometimes you may need to restructure how the data is presented in a table. For instance you may want to transpose the data or consolidate data from several columns into one. FME's AttributeExploder along with some of its other tools help you accomplish these types of tasks. The AttributeExploder takes a dataset and creates a row for each column / value pair, essentially breaking it down into its tiniest pieces. Once this is done you can combine the data back together in a way that fits your needs.

Example: Transposing tables

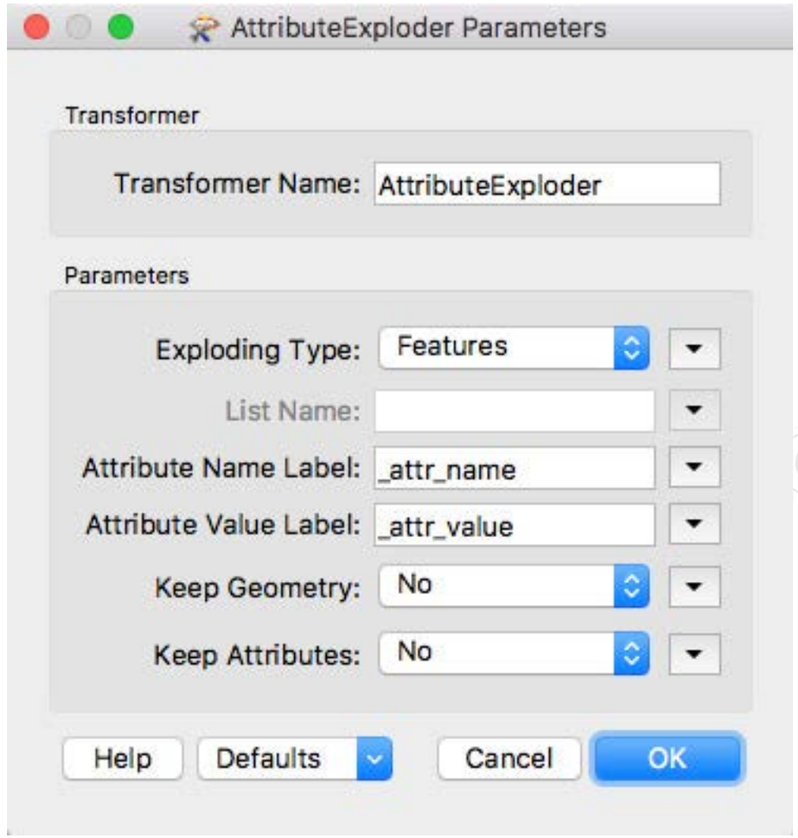
From this:

	A	B	C	D	E	F	G	H	I
1	URL Group Name	20161001	20161101	20161201	20170101	20170201	20170301	20170401	20170501
2	Cloudant	15	16	11	5	17	13	25	15
3	CouchDB	18	28	26	33	36	15	32	48
4	DocumentDB	17	25	30	50	73	75	79	83
5	DynamoDB	146	135	129	231	169	224	239	193
6	Elasticsearch	59	72	76	122	174	257	236	229
7	MongoDB	54	80	89	163	[illegible]	355	527	823
8	NoSQL	126	27	9	14	[illegible]	12	22	4

To this:

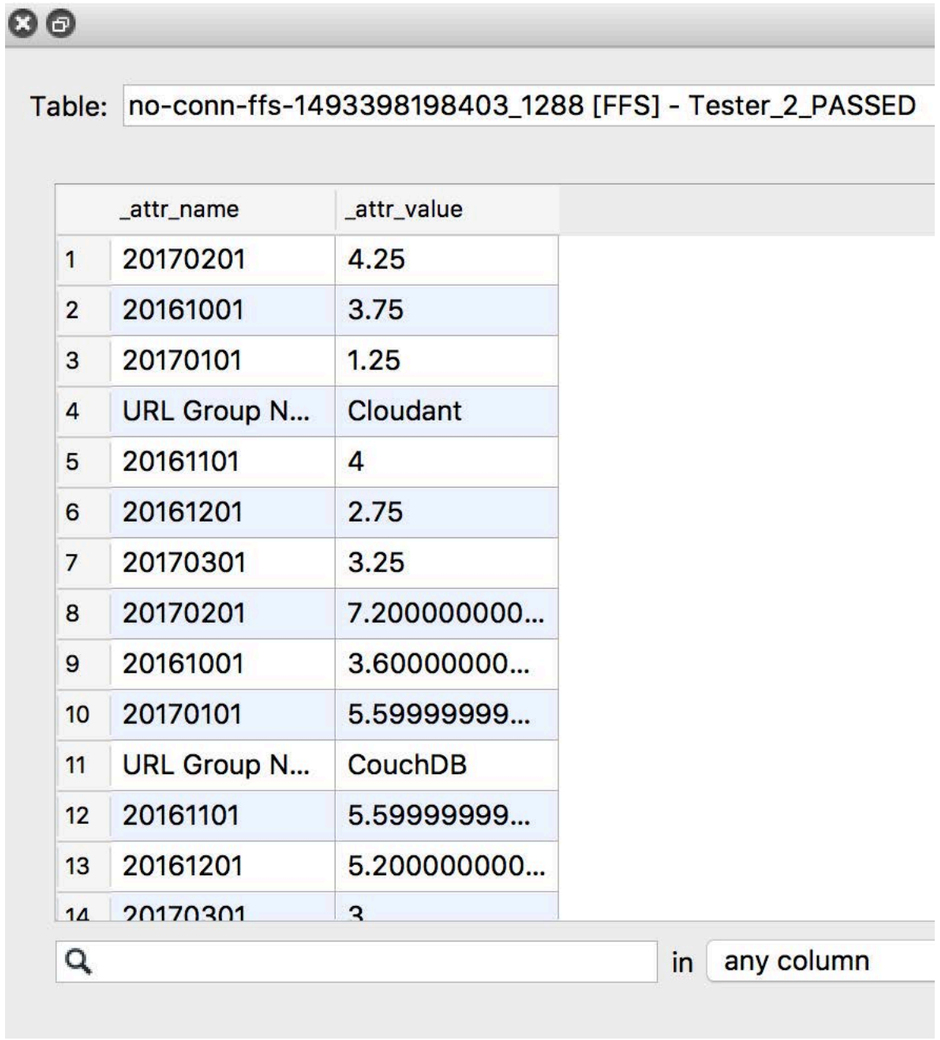
	A	B	C	D	E	F	G	H
1	URL Group Name	Cloudant	CouchDB	DocumentDB	DynamoDB	Elasticsearch	MongoDB	NoSQL
2	20161001	15	18	17	146	59	54	126
3	20161101	16	28	25	135	72	80	27
4	20161201	11	26	30	129	76	89	9
5	20170101	5	33	50	231	122	163	14
6	20170201	17	36	73	169	174	299	6
7	20170301	13	15	75	224	257	355	12

Connect your data to an AttributeExploder and leave everything as the default setting.



At this point you'll need to use a Tester to filter out FME format attributes. Refer to the Knowledge article and example workspace linked at the end of this section for full details.

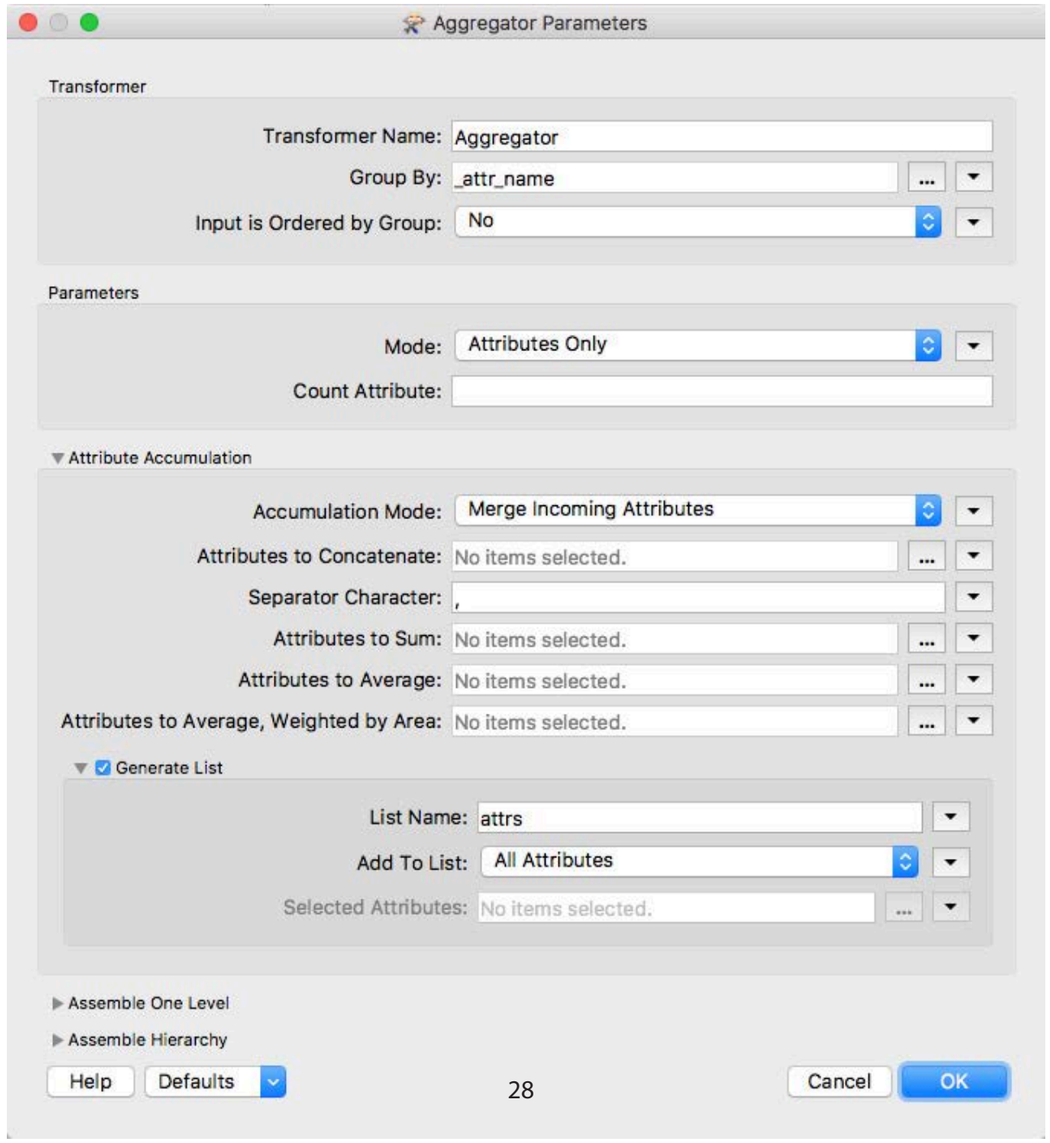
When you investigate what is coming out of the Tester in the FME Data Inspector you'll see a long table that consists of two columns: one for the attribute (column) names and one for the paired value.



	_attr_name	_attr_value
1	20170201	4.25
2	20161001	3.75
3	20170101	1.25
4	URL Group N...	Cloudant
5	20161101	4
6	20161201	2.75
7	20170301	3.25
8	20170201	7.200000000...
9	20161001	3.60000000...
10	20170101	5.59999999...
11	URL Group N...	CouchDB
12	20161101	5.59999999...
13	20161201	5.200000000...
14	20170301	3

To combine these bits of data back together into your new transposed tables, use the Aggregator. In the dialog, set “Group By” to “_attr_name”, “Mode” to “Attributes Only” and “Accumulation Mode” to “Merge Incoming Attributes”.

These crucial settings tell the transformer to create a row in a new table for each column (‘attr_name’) from the original table. It will then merge all the attributes together to form the new table. Check “Generate List” to tell transformer to create a list (column) for each value associated to a given feature. Give the list a name and set “Add To List” to “All Attributes”.



Taking a look at what is coming out of the Aggregator in the Data Inspector you’ll see a table where the former column names are now listed in rows. Values are listed in columns the transformer has created using the List Name that you indicated following this pattern: listname{0}._attr_value, listname{1}._attr_value, …

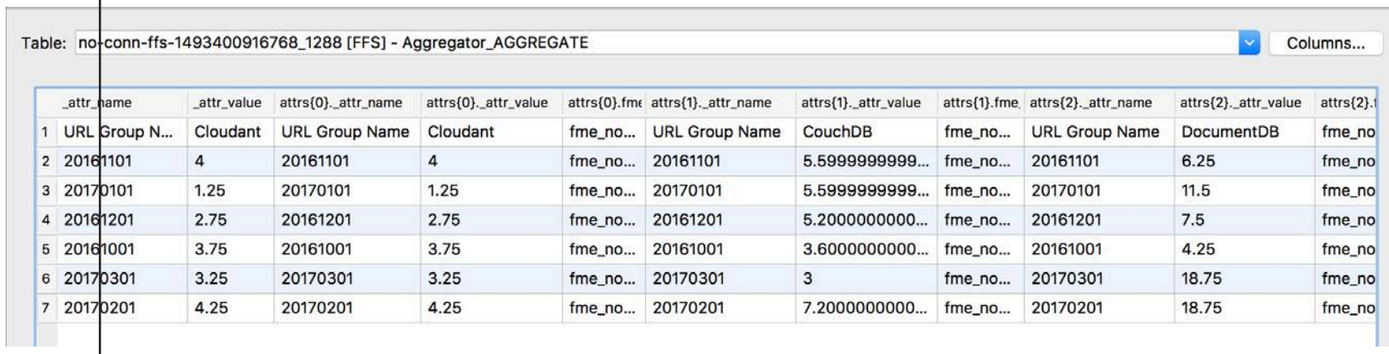
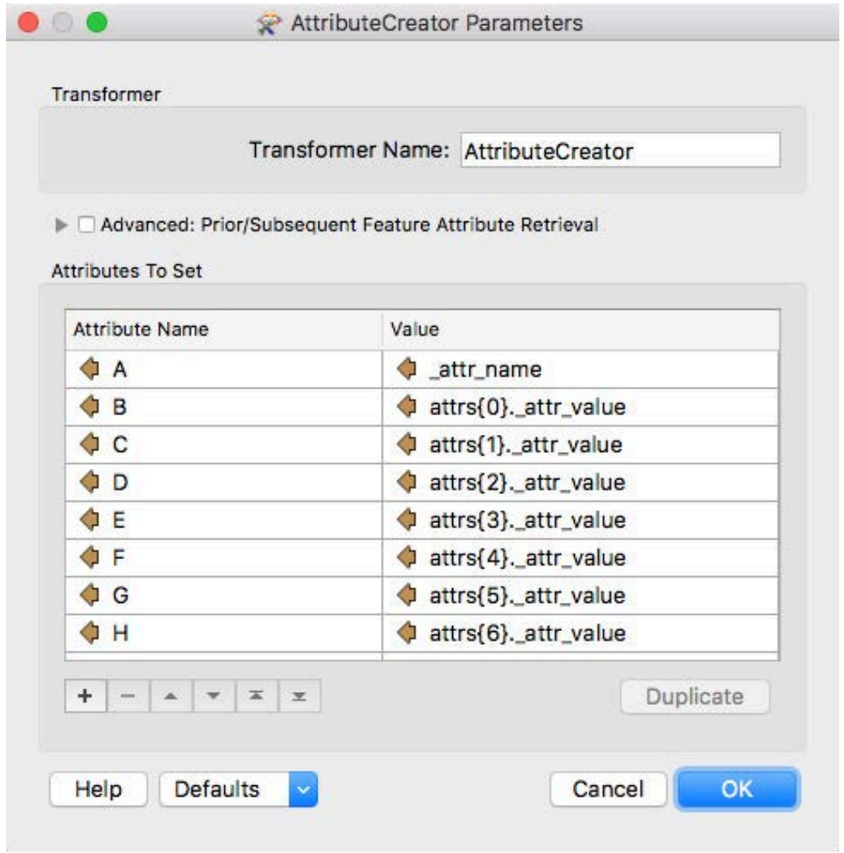


Table: no-conn-ffs-1493400916768_1288 [FFS] - Aggregator_AGGREGATE — Columns...

	_attr_name	_attr_value	attrs{0}._attr_name	attrs{0}._attr_value	attrs{0}.fme	attrs{1}._attr_name	attrs{1}._attr_value	attrs{1}.fme	attrs{2}._attr_name	attrs{2}._attr_value	attrs{2}.
1	URL Group N...	Cloudant	URL Group Name	Cloudant	fme_no...	URL Group Name	CouchDB	fme_no...	URL Group Name	DocumentDB	fme_no
2	20161101	4	20161101	4	fme_no...	20161101	5.5999999999...	fme_no...	20161101	6.25	fme_no
3	20170101	1.25	20170101	1.25	fme_no...	20170101	5.5999999999...	fme_no...	20170101	11.5	fme_no
4	20161201	2.75	20161201	2.75	fme_no...	20161201	5.2000000000...	fme_no...	20161201	7.5	fme_no
5	20161001	3.75	20161001	3.75	fme_no...	20161001	3.6000000000...	fme_no...	20161001	4.25	fme_no
6	20170301	3.25	20170301	3.25	fme_no...	20170301	3	fme_no...	20170301	18.75	fme_no
7	20170201	4.25	20170201	4.25	fme_no...	20170201	7.2000000000...	fme_no...	20170201	18.75	fme_no

To clean up the table (eliminate duplicate information or FME format jargon) use the AttributeCreator. Create a new column for “_attr_name” and each associated listname{#}._attr_value. In FME if you assign letters as attribute names when working with Excel, the writer will assume they are referring to column designations. In the writer tool manually enter the attribute names as “A, B, C …”.

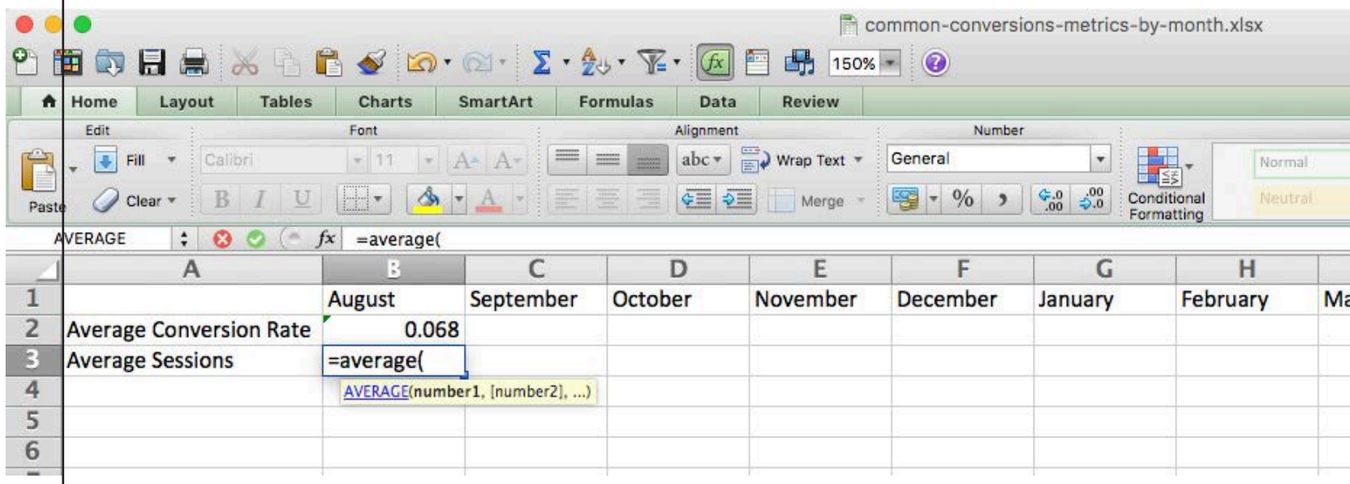


ANALYTICS & REPORTING

Excel is great for analyzing and reporting on data, but once again, there are situations where FME's built-in tools provide advantages with respect to data integrity and efficiency.

Statistics

Calculating statistics in Excel is relatively straightforward; however it involves manually constructing tables, entering functions and named ranges. And you'll have to carry out this task each time you'd like to update it with new data. This can be tedious in addition to threatening the quality of the data.



common-conversions-metrics-by-month.xlsx

Home | Layout | Tables | Charts | SmartArt | Formulas | Data | Review

Edit | Font | Alignment | Number | Format

Paste | Fill | Clear | Calibri | 11 | Wrap Text | Merge | General | Conditional Formatting | Normal | Bad | Neutral | Calculation

AVERAGE | =average(Sheet1!D2:D25

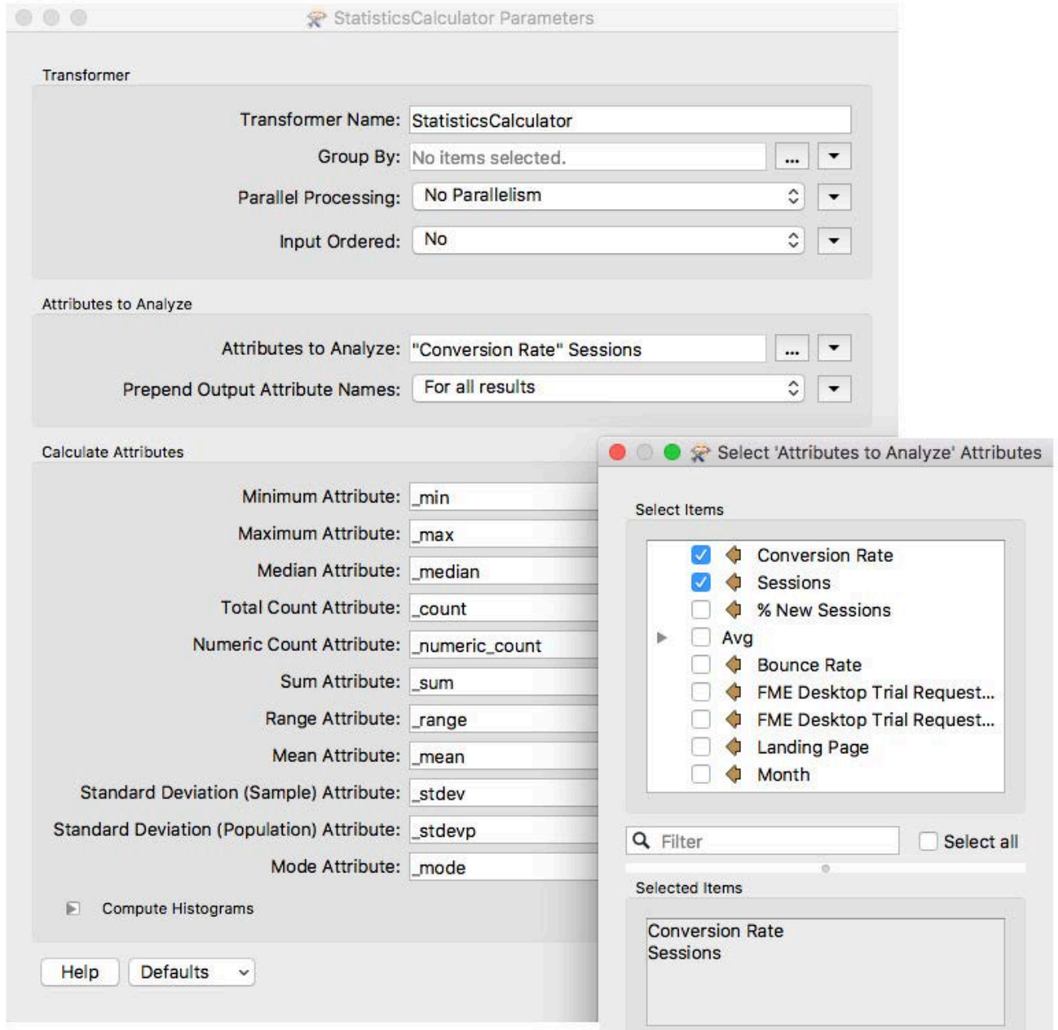
AVERAGE(**number1**, [number2], ...)

	Landing Page	Month	B	C	D
1	Landing Page	Month		Conversion Rate	Sessions
2	http://www.safe.com/convert/		Aug 01, 2016	0	2
3	http://www.safe.com/convert/3ds-max/		Aug 01, 2016	0	1
4	http://www.safe.com/convert/3ds-max/microsoft-directx-x-file/		Aug 01, 2016	0	10
5	http://www.safe.com/convert/access-mdb/		Aug 01, 2016	0	3
6	http://www.safe.com/convert/access-mdb/arcgis-shp/		Aug 01, 2016	0.086956522	23
7	http://www.safe.com/convert/access-mdb/csv/		Aug 01, 2016	0.192307692	26
8	http://www.safe.com/convert/access-mdb/dbf/		Aug 01, 2016	0.25	4
9	http://www.safe.com/convert/access-mdb/dwg/		Aug 01, 2016	0	11
10	http://www.safe.com/convert/access-mdb/file-geodatabase-api/		Aug 01, 2016	0	1
11	http://www.safe.com/convert/access-mdb/file-geodatabase/		Aug 01, 2016	0	42
12	http://www.safe.com/convert/access-mdb/gml/		Aug 01, 2016	0	2
13	http://www.safe.com/convert/access-mdb/mapinfo/		Aug 01, 2016	0.125	8
14	http://www.safe.com/convert/access-mdb/oracle/		Aug 01, 2016	0	1
15	http://www.safe.com/convert/access-mdb/postgis/		Aug 01, 2016	0	1
16	http://www.safe.com/convert/access-mdb/postgresql/		Aug 01, 2016	0	3
17	http://www.safe.com/convert/access-mdb/sql-server/		Aug 01, 2016	0.0625	16
18	http://www.safe.com/convert/access-mdb/v7-dgn-igds/		Aug 01, 2016	0	1
19	http://www.safe.com/convert/adac/		Aug 01, 2016	0	3
20	http://www.safe.com/convert/additional-military-layers-aml/		Aug 01, 2016	0	2
21	http://www.safe.com/convert/ais/		Aug 01, 2016	0	3
22	http://www.safe.com/convert/aixm-5/		Aug 01, 2016	0	1
23	http://www.safe.com/convert/aixm/		Aug 01, 2016	0	1
24	http://www.safe.com/convert/arcgis-layer/		Aug 01, 2016	0	1
25	http://www.safe.com/convert/arcgis-online-feature-service/file-geodatabase-api/		Aug 01, 2016	0	1
26	http://www.safe.com/convert/arcgis-shp/3ds-max/		Aug 01, 2016	0	
27	http://www.safe.com/convert/arcgis-shp/access-mdb/		Aug 01, 2016	0.060606061	33
28	http://www.safe.com/convert/arcgis-shp/arcgis-online-feature-service/		Aug 01, 2016	0	1
29	http://www.safe.com/convert/arcgis-shp/arcinfo-coverage/		Aug 01, 2016	0	1
30	http://www.safe.com/convert/arcgis-shp/arcsde-geodatabase/		Aug 01, 2016	0	3

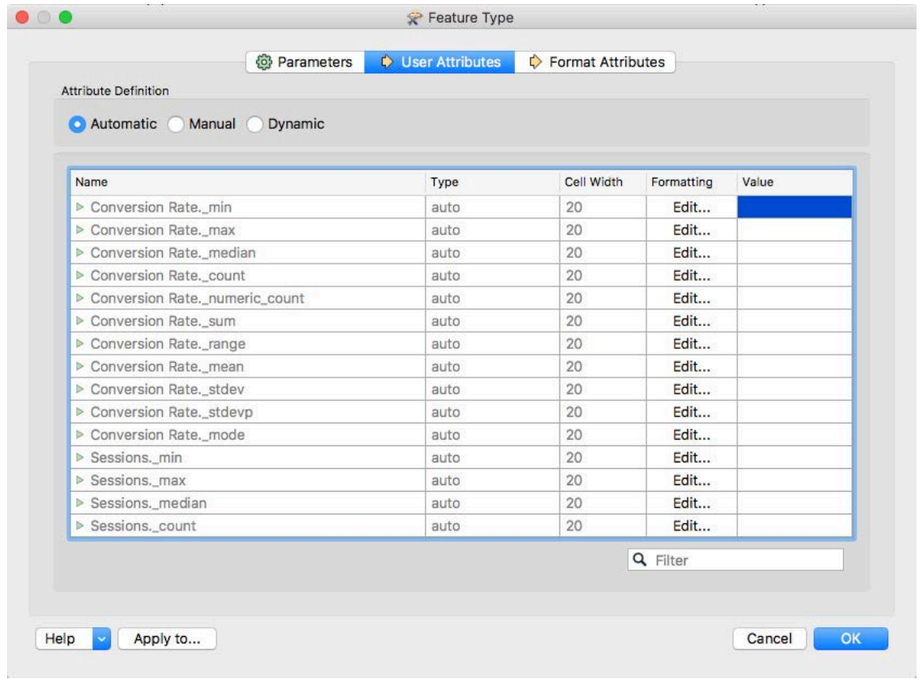
24R x 1C

In FME, the StatisticsCalculator tool generates a customized statistical report that automatically refreshes with new data each time the workflow is run.

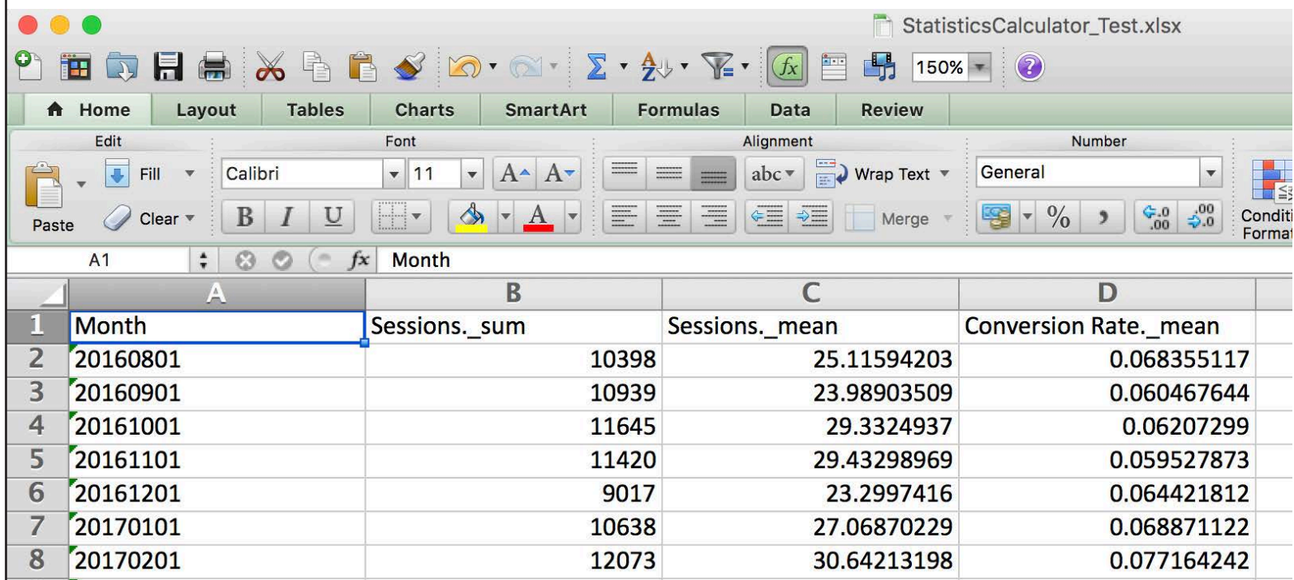
Add a StatisticsCalculator and a Writer to the workspace. Connect the output of this data to the StatisticsCalculator and the “Summary” port of the StatisticsCalculator to the Writer. In the StatisticsCalculator dialog, choose the columns to analyze with the “Attributes to Analyze” function. If you are analyzing more than one attribute, select “For All Results” for the “Prepend Output Attribute Names” parameter. This is to ensure that results for the same type of statistic calculated from different attributes can be distinguished.



In the Writer dialog in the “User Attributes” section, set “Attribute Definition” to “Automatic”. FME will automatically name your columns according to the format: “Attribute._statistic”.



Once the names of all possible attributes are listed using the Automatic setting, you can switch to manual and customize the statistics you’d like in your report by deleting the ones you don’t need.

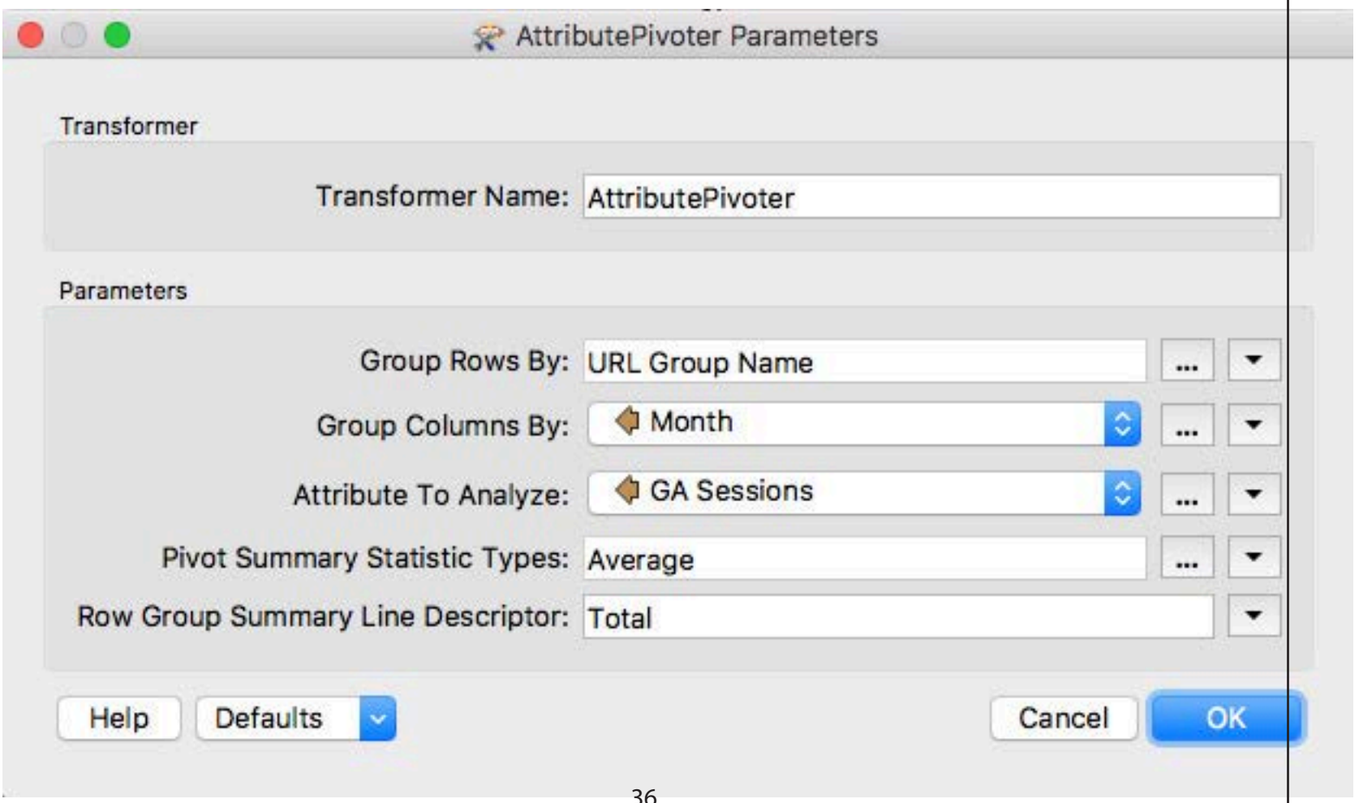


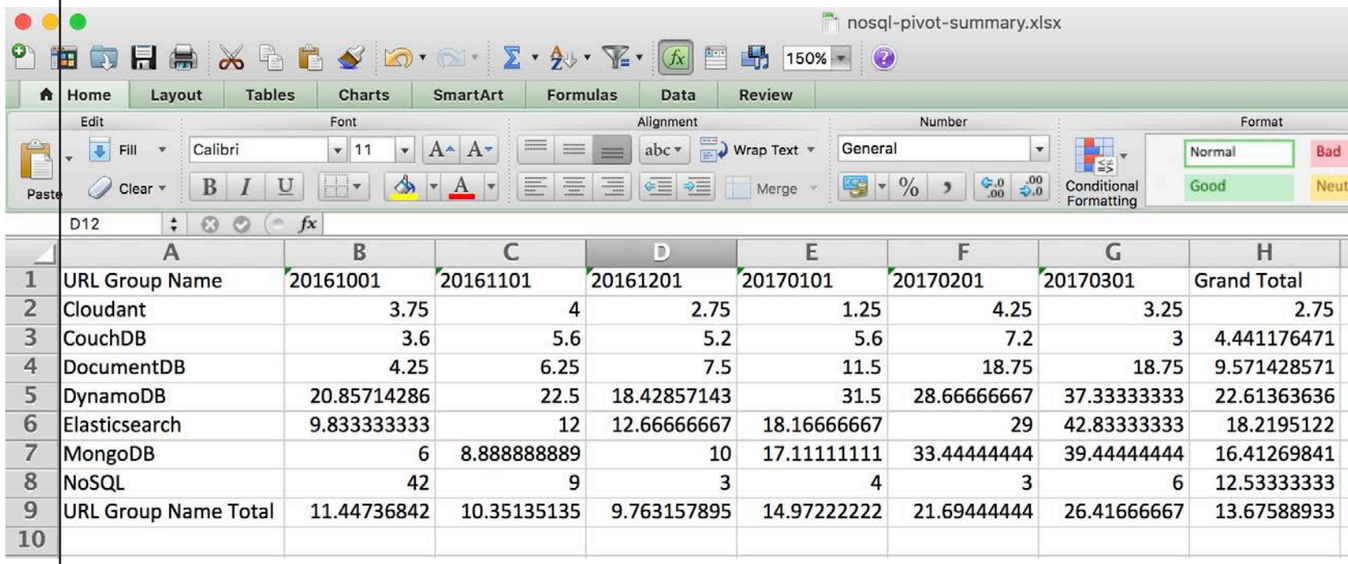
	A	B	C	D
1	Month	Sessions._sum	Sessions._mean	Conversion Rate._mean
2	20160801	10398	25.11594203	0.068355117
3	20160901	10939	23.98903509	0.060467644
4	20161001	11645	29.3324937	0.06207299
5	20161101	11420	29.43298969	0.059527873
6	20161201	9017	23.2997416	0.064421812
7	20170101	10638	27.06870229	0.068871122
8	20170201	12073	30.64213198	0.077164242

Pivot Tables

Pivot Tables are one of Excel's most useful capabilities, enabling users to analyze data by "pivoting" it around one field. In many cases it makes sense to generate pivot reports using Excel; however, if you are using FME to process Excel data, it is useful to know that the AttributePivoter generates "pivot-like" reports.

Once you've added an AttributePivoter to your workspace and connected the data to its input port, set the parameters in the dialog including what attribute (column) to analyze, and for what type of statistic. Add a writer and connect both the "Data" and "Summary" output ports from the AttributePivoter. In the writer dialog, set the format to Excel and in the "User Attributes" section set the "Attribute Definition" parameter to "Dynamic". The schema is automatically created for you based on the pivot output.





	A	B	C	D	E	F	G	H
1	URL Group Name	20161001	20161101	20161201	20170101	20170201	20170301	Grand Total
2	Cloudant	3.75	4	2.75	1.25	4.25	3.25	2.75
3	CouchDB	3.6	5.6	5.2	5.6	7.2	3	4.441176471
4	DocumentDB	4.25	6.25	7.5	11.5	18.75	18.75	9.571428571
5	DynamoDB	20.85714286	22.5	18.42857143	31.5	28.66666667	37.33333333	22.61363636
6	Elasticsearch	9.833333333	12	12.66666667	18.16666667	29	42.83333333	18.2195122
7	MongoDB	6	8.888888889	10	17.11111111	33.44444444	39.44444444	16.41269841
8	NoSQL	42	9	3	4	3	6	12.53333333
9	URL Group Name Total	11.44736842	10.35135135	9.763157895	14.97222222	21.69444444	26.41666667	13.67588933
10								

Both FME and Excel have their individual strengths. Excel is a great platform for laying out, organizing and analyzing data, and it is easily shared with others. FME allows you to process data with automated, hands-free workflows. Used in conjunction, the result is Excel data that is formatted exactly how you need, quickly, and without manual interference that can harm data's integrity.

To learn more, visit **knowledge.safe.com**