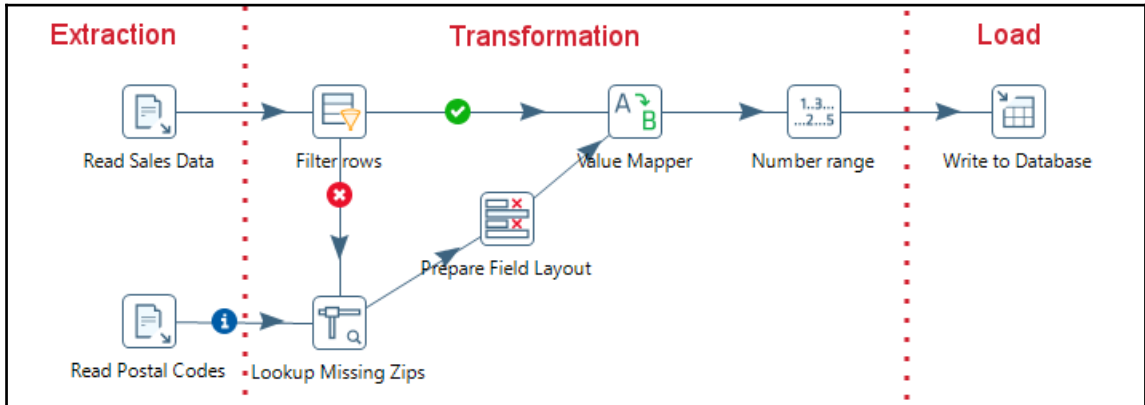


Chapter 01: Getting Started with Pentaho Data Integration

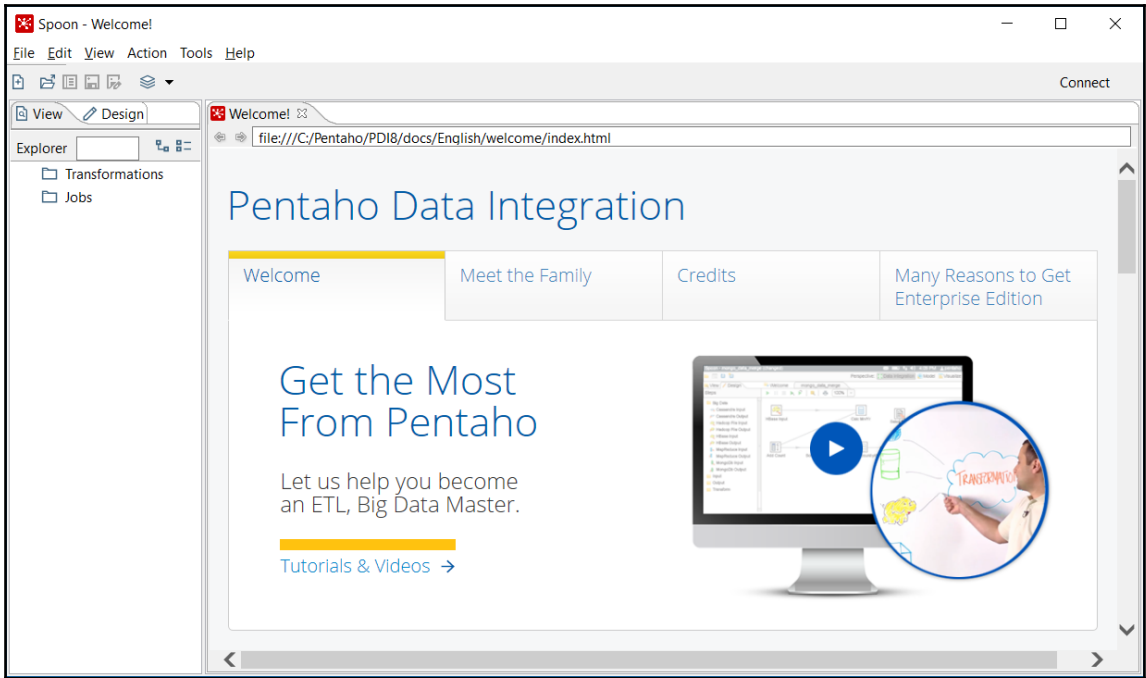


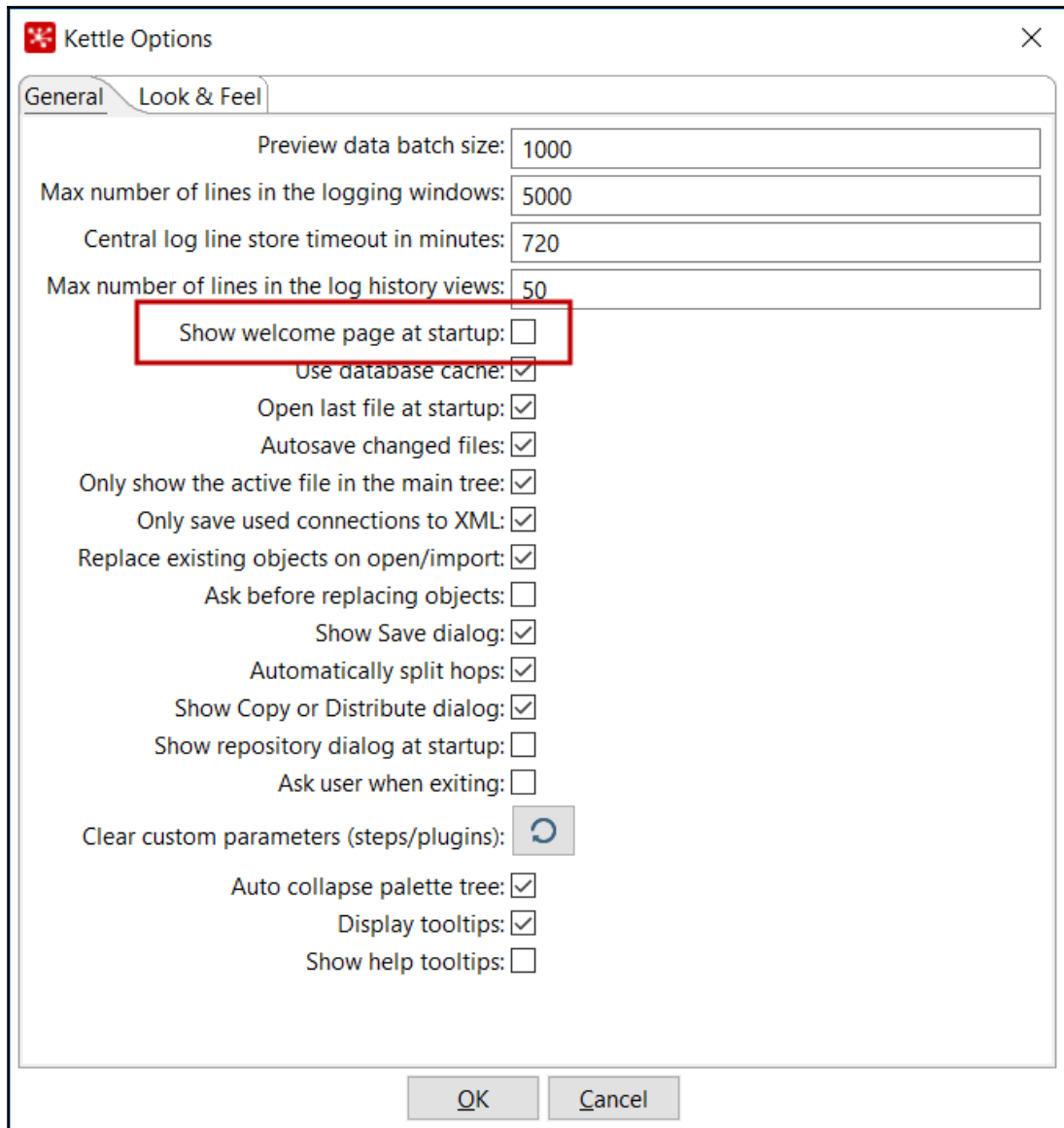
Summary | [Files](#) | [Reviews](#) | [Support](#) | [Wiki](#) | [News](#) | [Donate](#)

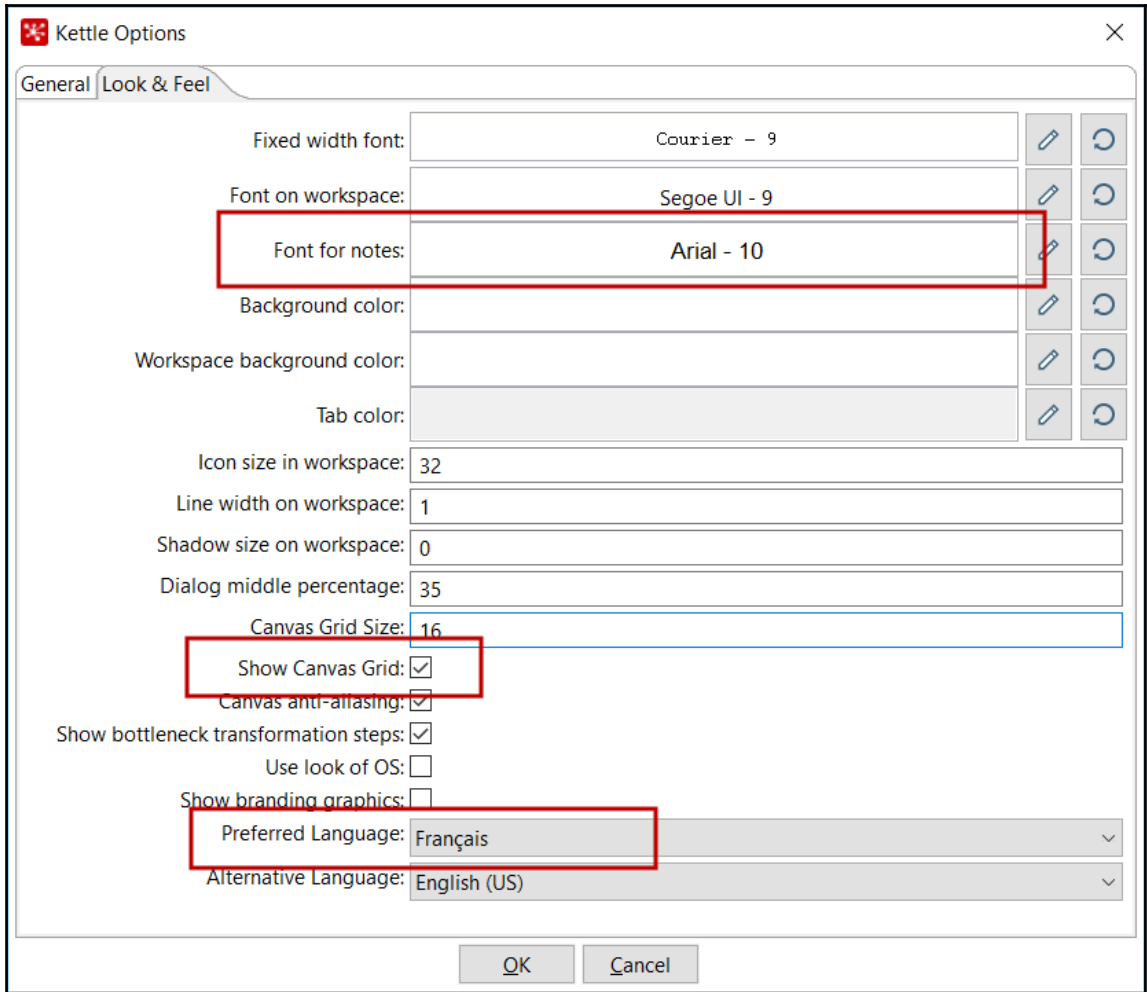
Looking for the latest version? [Download pdi-ce-8.0.0.0-28.zip \(979.8 MB\)](#)

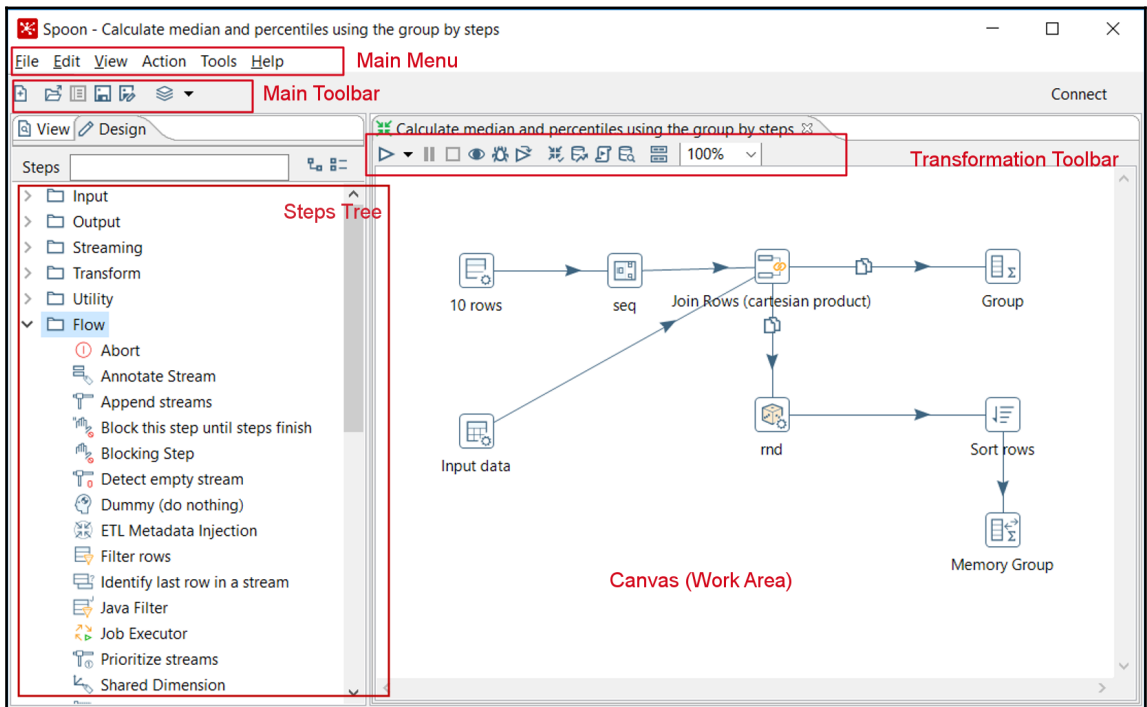
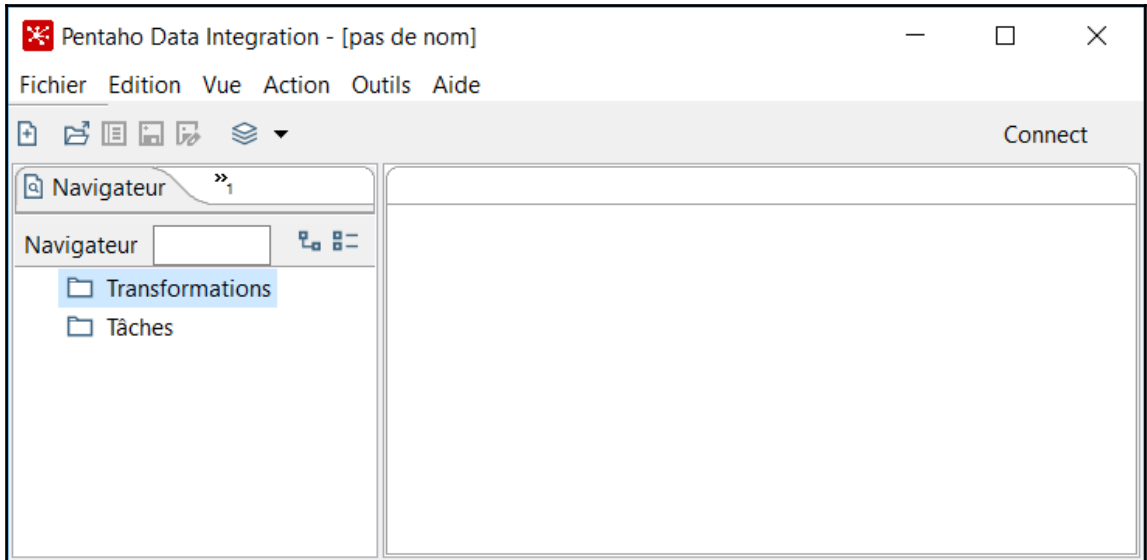
Home / Data Integration

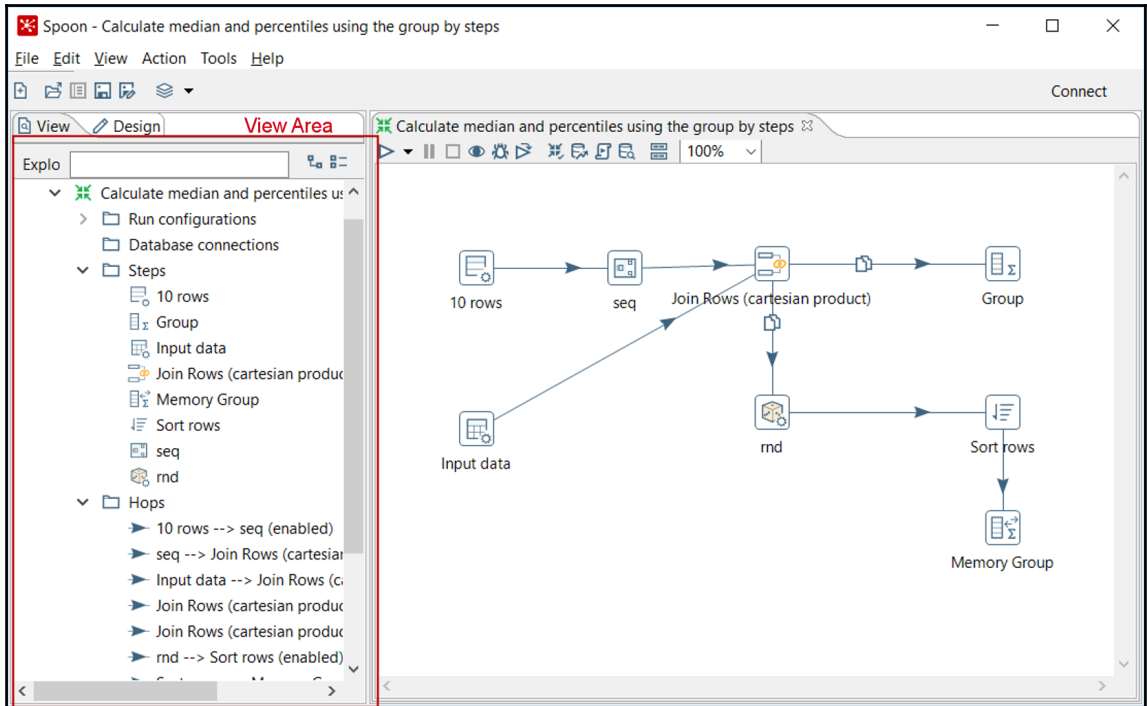
Name ↕	Modified ↕	Size ↕	Downloads / Week ↕
↑ Parent folder			
7.1	2017-05-22		6,233
7.0	2016-11-09		199
6.1	2016-04-13		307
6.0	2015-12-10		254
5.4	2015-06-15		77











Marketplace

Available Installed

Type: All Stage: All Search What are stages?

PDI MySQL Plugin Pentaho	3	Available TRUNK-SNAPSHOT (...)	Install
CPython Script Executor Mark Hall	2	Available 1.1 (Trunk)	Install
PDI MQTT Steps Mark Hall	2	Available 1 (Trunk)	Install
PDI NuoDB Plugin NuoDB	4	Available 1.0-SNAPSHOT (Stable)	Install
Apple Push Notification Joel Latino	3	Available 1.0.1 (Stable)	Install
Android Push Notification Joel Latino	3	Available 1.0.1 (Stable)	Install

✕

Vertica Bulk Loader

Pentaho

Version: 5.2.0.1-216 (Stab... ▼

Install ↕

Uninstall ↕

Information

Branch: Stable

Version: 5.2.0.1-216

Developer: Pentaho

Description

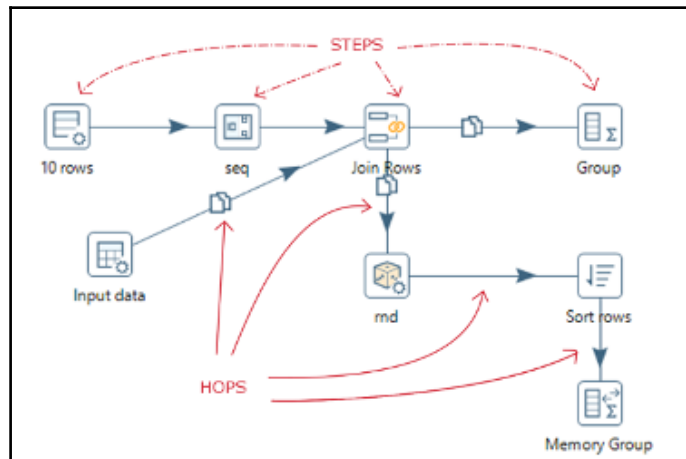
This is a plugin for performing bulk data loads into a Vertica Analytic Database using a COPY FROM STDIN statement through JDBC

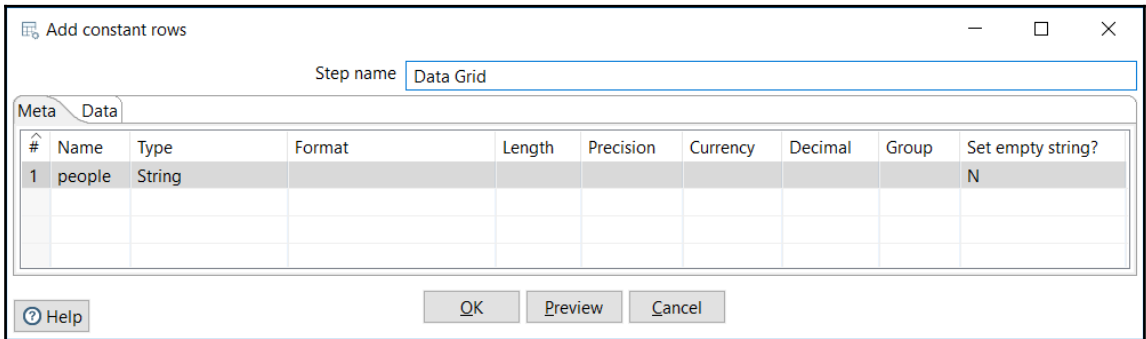
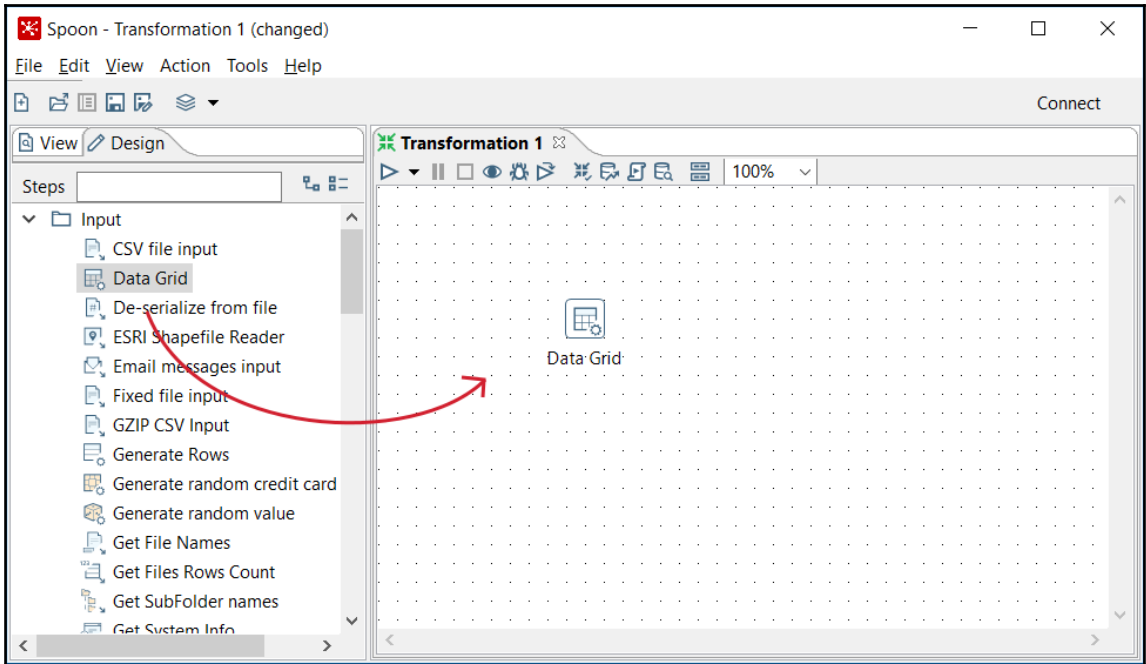
4

Mature Release

Indicates a successfully adopted project in a mature state.

[Submit your feedback](#)





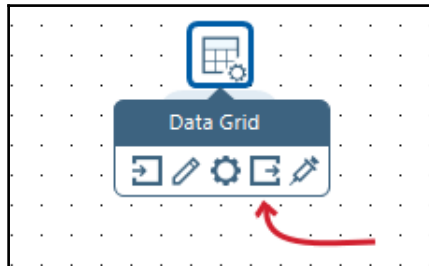
Add constant rows

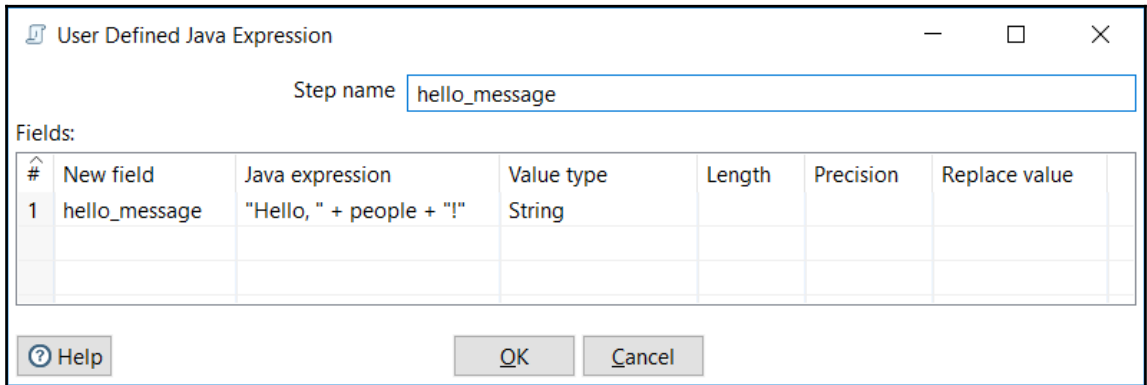
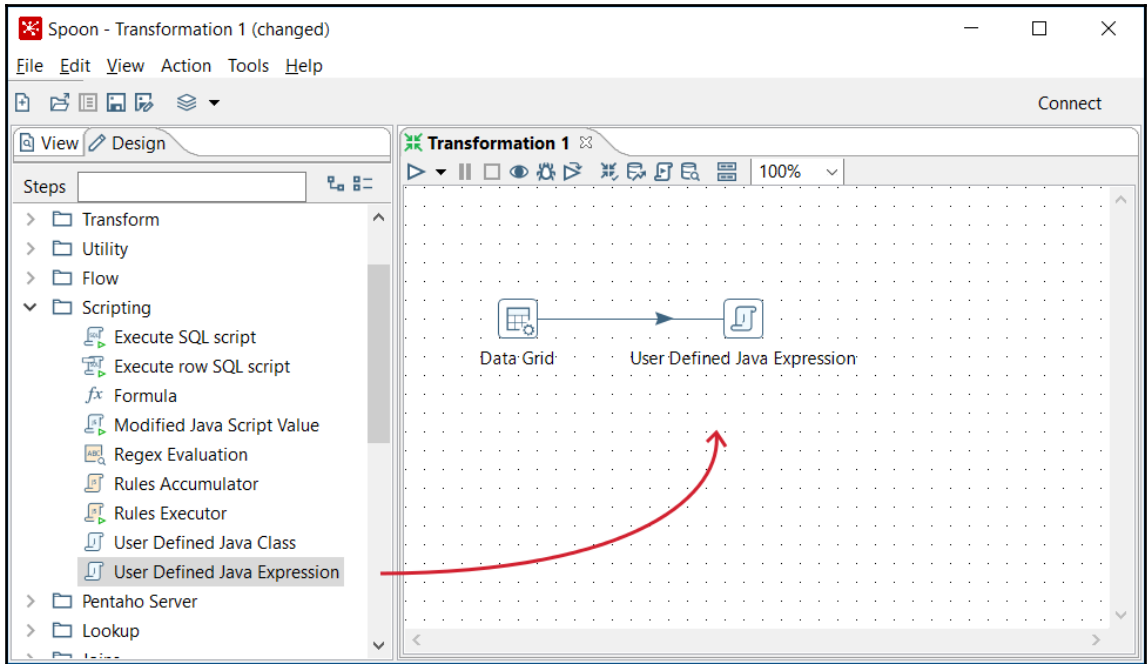
Step name

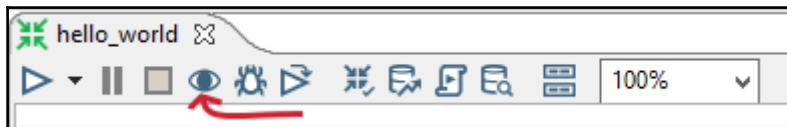
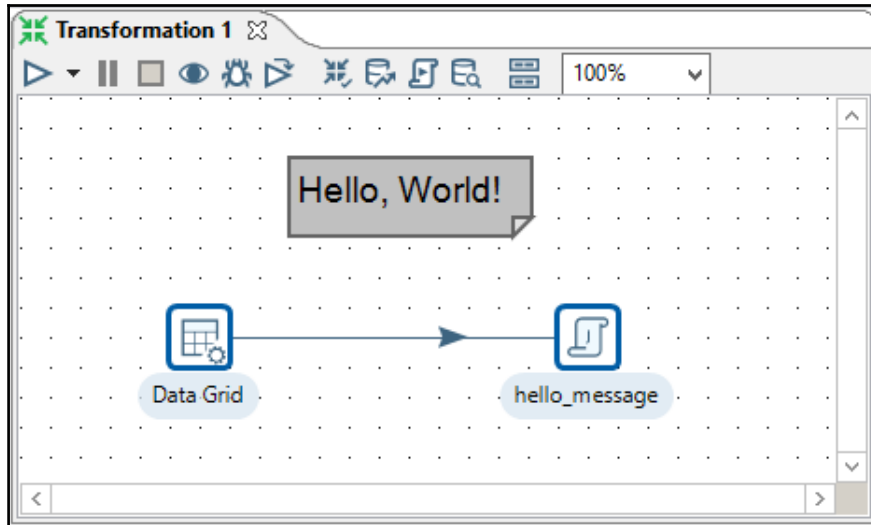
Meta Data

#	people
1	John
2	Mary
3	Ammy
4	Erik
5	Laura

Help OK Preview Cancel







Spoon - hello_world

File Edit View Action Tools Help

Connect

View Design

Steps

- > Input
- > Output
- > Streaming
- > Transform
- > Utility
- > Flow
- > Scripting
- > Pentaho Server
- > Lookup
- > Joins
- > Data Warehouse
- > Validation
- > Statistics
- > Big Data
- > Agile
- > Cryptography
- > Palo
- > Open ERP

hello_world

100%

Hello, World!

Data Grid

hello_message

Examine preview data

Rows of step: hello_message (5 rows)

#	people	hello_message
1	John	Hello, John!
2	Mary	Hello, Mary!
3	Ammy	Hello, Ammy!
4	Erik	Hello, Erik!
5	Laura	Hello, Laura!

Close

Execution Results

Logging Execution History Step Metrics

#	Stepname	Copynr	Read	Write
1	Data Grid	0	0	5
2	hello_message	0	5	5

hello_world

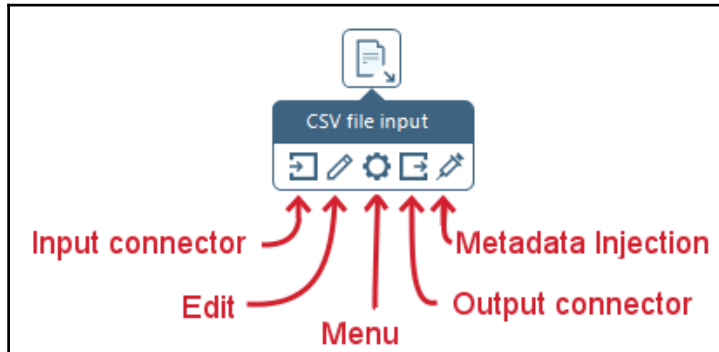
100%

The screenshot displays the Apache Spark IDE (Spoon) interface. The main workspace shows a transformation named 'Hello, World!' with two steps: 'Data Grid' and 'hello_message'. The 'Execution Results' panel is open, showing a log of the execution process.

Execution Results

- 2017/11/16 19:08:15 - Spoon - Spoon
- 2017/11/16 19:43:02 - Spoon - Spoon
- 2017/11/16 19:43:22 - Spoon - Spoon
- 2017/11/16 20:13:50 - hello_world - Dispatching started for transformation [hello_world]
- 2017/11/16 20:13:50 - Data Grid.0 - Finished processing (I=0, O=0, R=0, W=5, U=0, E=0)
- 2017/11/16 20:13:50 - hello_message.0 - Finished processing (I=0, O=0, R=5, W=5, U=0, E=0)
- 2017/11/16 20:13:50 - Spoon - The transformation has finished!!

Chapter 02: Getting Started with Transformations



#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1	project_name	String							none
2	start_date	Date	yyyy-MM-dd						none
3	end_date	Date	yyyy-MM-dd						none

Examine preview data

Rows of step: CSV file input (6 rows)

#	project_name	start_date	end_date
1	Project A	2016-01-10	2016-01-25
2	Project B	2016-04-03	2016-07-21
3	Project C	2017-01-15	<null>
4	Project D	2015-09-03	2015-12-20
5	Project E	2016-05-11	2016-05-31
6	Project F	2011-12-01	2013-11-30

< >

Close Show Log

Calculator

Step name:

Fields:

#	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove
1	diff_dates	Date A - Date B (in days)	end_date	start_date		Integer			N

Number ranges

Step name:

Input field:

Output field:

Default value(if no range matches):

Ranges (min <= x < max):

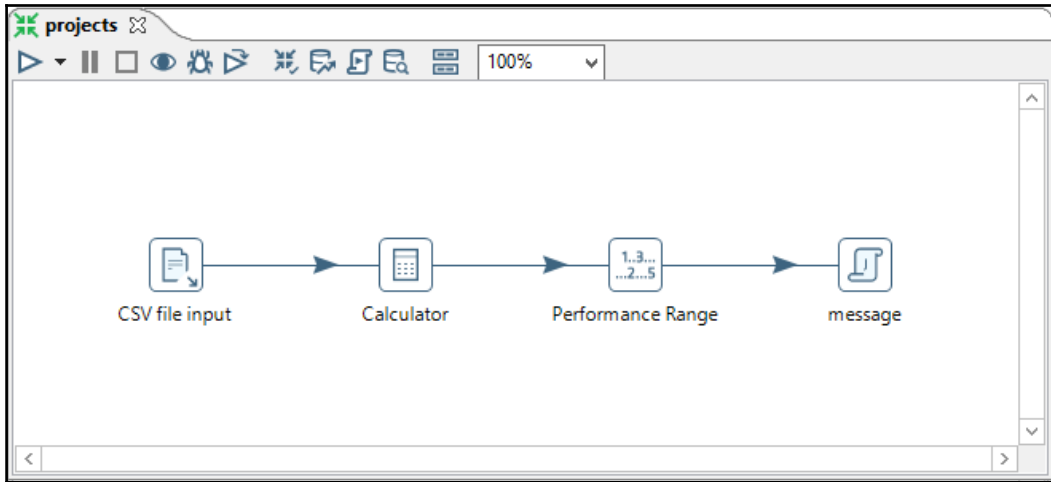
#	Lower Bound	Upper Bound	Value
1		30.0	excellent
2	30.0	80.0	very good
3	80.0	160.0	good
4	160.0		poor

User Defined Java Expression

Step name:

Fields:

#	New field	Java expression	Value type	Length	Precision	Replace value
1	duration	(diff_dates == null)?"unknown":diff_dates + " days"				
2	message	"The performance was " + performance				



The screenshot shows the same workflow as above, but with an error dialog box overlaid on the 'Calculator' step. The dialog box contains the following text:

No preview rows found

Sorry, during preview there weren't any rows to display for this step.

OK

Below the dialog box, the 'Execution Results' section is visible, showing a table with the following data:

#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active
1	Calculator	0	3	2	0	0	0	0	1	Stopped
2	Performance Range	0	1	0	0	0	0	0	0	Stopped
3	message	0	0	0	0	0	0	0	0	Stopped
4	CSV file input	0	0	6	7	0	0	0	0	Stopped

Execution Results

#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active
1	Calculator	0	3	2	0	0	0	0	1	Stopped
2	Performance Range	0	2	2	0	0	0	0	0	Stopped
3	message	0	1	0	0	0	0	0	0	Stopped
4	CSV file input	0	0	6	7	0	0	0	0	Stopped

Examine preview data

Rows of step: Calculator (5 rows)

#	project_name	start_date	end_date	diff_dates
1	Project A	2016-01-10	2016-01-25	15
2	Project B	2016-04-03	2016-07-21	109
3	Project D	2015-09-03	2015-12-20	108
4	Project E	2016-05-11	2016-05-31	20
5	Project F	2011-12-01	2013-11-30	730

Close

Examine preview data

Rows of step: message (5 rows)

#	project_name	start_date	end_date	diff_dates	performance	duration	message
1	Project A	2016-01-10	2016-01-25	15	excellent	<null>	<null>
2	Project B	2016-04-03	2016-07-21	109	good	<null>	<null>
3	Project D	2015-09-03	2015-12-20	108	good	<null>	<null>
4	Project E	2016-05-11	2016-05-31	20	excellent	<null>	<null>
5	Project F	2011-12-01	2013-11-30	730	poor	<null>	<null>

Close

Examine preview data

Rows of step: message (5 rows)

#	project_name	start_date	end_date	diff_dates	performance	duration	message
1	Project A	2016-01-10	2016-01-25	15	excellent	15 days	The performance was excellent
2	Project B	2016-04-03	2016-07-21	109	good	109 days	The performance was good
3	Project D	2015-09-03	2015-12-20	108	good	108 days	The performance was good
4	Project E	2016-05-11	2016-05-31	20	excellent	20 days	The performance was excellent
5	Project F	2011-12-01	2013-11-30	730	poor	730 days	The performance was poor

Close

Generate Rows

Step name: date range

Limit: 1

Never stop generating rows:

Interval in ms (delay): 5000

Current row time field name: now

Previous row time field name: FiveSecondsAgo

Fields:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Value	Set empty string?
1	start_date	Date	yyyy-MM-dd						2017-01-01	N
2	end_date	Date	yyyy-MM-dd						2017-06-30	N

Help OK Preview Cancel

Examine preview data

Rows of step: Clone row (181 rows)

#	start_date	end_date	diff_dates	delta
1	2017-01-01	2017-06-30	180	0
2	2017-01-01	2017-06-30	180	1
3	2017-01-01	2017-06-30	180	2
4	2017-01-01	2017-06-30	180	3
5	2017-01-01	2017-06-30	180	4
6	2017-01-01	2017-06-30	180	5
7	2017-01-01	2017-06-30	180	6
8	2017-01-01	2017-06-30	180	7
9	2017-01-01	2017-06-30	180	8
10	2017-01-01	2017-06-30	180	9

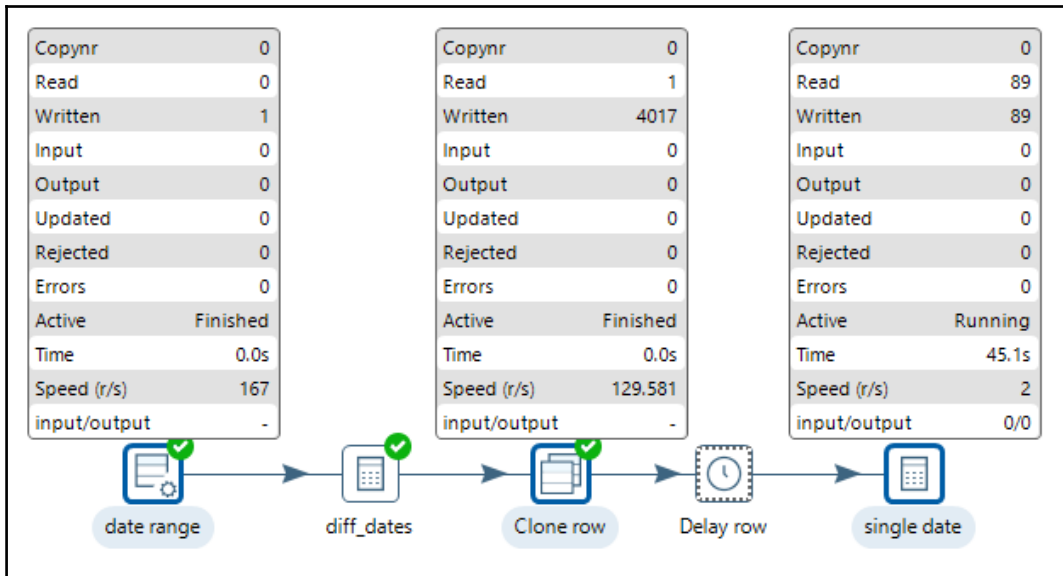
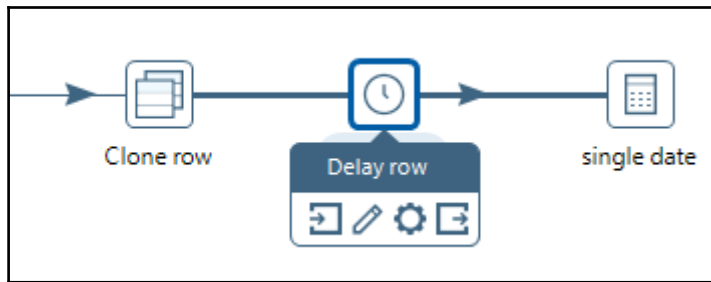
Close

Examine preview data

Rows of step: single date (181 rows)

#	start_date	end_date	diff_dates	delta	a_single_date
1	2017-01-01	2017-06-30	180	0	2017-01-01
2	2017-01-01	2017-06-30	180	1	2017-01-02
3	2017-01-01	2017-06-30	180	2	2017-01-03
4	2017-01-01	2017-06-30	180	3	2017-01-04
5	2017-01-01	2017-06-30	180	4	2017-01-05
6	2017-01-01	2017-06-30	180	5	2017-01-06
7	2017-01-01	2017-06-30	180	6	2017-01-07
8	2017-01-01	2017-06-30	180	7	2017-01-08
9	2017-01-01	2017-06-30	180	8	2017-01-09
10	2017-01-01	2017-06-30	180	9	2017-01-10

Close



fields

#	project_name	start_date	end_date	diff_dates
1	Project A	2016-01-10	2016-01-25	15
2	Project B	2016-04-03	2016-07-21	109
3	Project D	2015-09-03	2015-12-20	108
4	Project E	2016-05-11	2016-05-31	20
5	Project F	2011-12-01	2013-11-30	730

rowset

Examine preview data

Rows of step: single date (181 rows)

#	start_date	end_date	diff_dates	delta	a_single_date
1	2017-01-01	2017-06-30	180	0	2017-01-01
2	2017-01-01	2017-06-30	180	1	2017-01-02
3	2017-01-01	2017-06-30	180	2	2017-01-03
4	2017-01-01	2017-06-30	180	3	2017-01-04
5	2017-01-01	2017-06-30	180	4	2017-01-05
6	2017-01-01	2017-06-30	180	5	2017-01-06
7	2017-01-01	2017-06-30	180	6	2017-01-07
8	2017-01-01	2017-06-30	180	7	2017-01-08
9	2017-01-01	2017-06-30	180	8	2017-01-09
1..	2017-01-01	2017-06-30	180	9	2017-01-10

Close

Step fields and their origin

Step name: Calculator

Fields:

#	Fieldname	Type	Length	Precision	Step origin	Storage	Mask	Currency	Decimal	Group	Trim	Comments
1	start_date	Date	-	-	date range	normal	yyyy-MM-dd				none	
2	end_date	Date	-	-	date range	normal	yyyy-MM-dd				none	
3	diff_dates	Integer	-	0	Calculator	normal					none	DATE_DIFF

Edit origin step Cancel

Select / Rename values

Step name

Select & Alter Remove Meta-data


Fields to alter the meta-data for :

#	Fieldname	Rename to	Type	Length	Precision	Binary to Normal?	Format
1	start_date		Date			N	yyyy-MM-dd
2	end_date		Date			N	yyyy-MM-dd

Get fields to change

Help OK Cancel

Warning

 Step 'Select values' has '2' destination steps to which it sends rows of data.

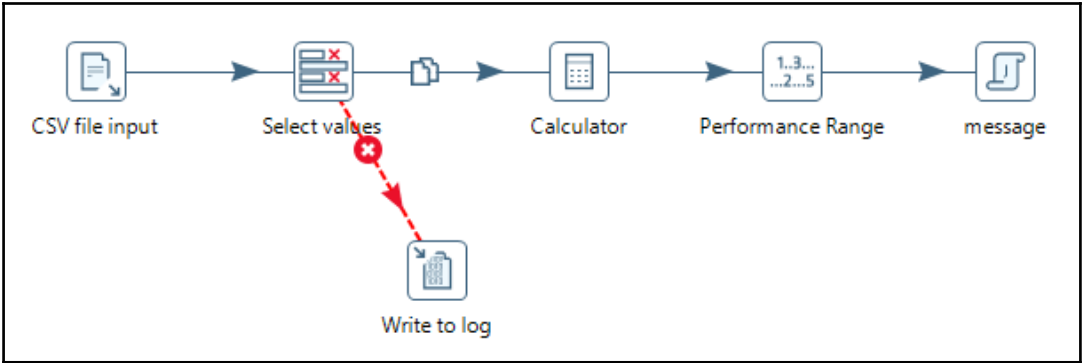
Besides explicit partitioning, there are different ways of sending rows to destination steps:

- Distribute rows: destination steps receive rows in turns
- Custom Distribution: for example, Load Balancing
- Copy rows: all rows are sent to all destination steps.

Which method would you like to use? (The default is 'Distribute rows')

Please, don't show this warning anymore (also available in the options dialog).

Distribute Copy



Examine preview data

Rows of step: Write to log (1 rows)

#	project_name	start_date	end_date	error_desc
1	Project C	2017-01-15	???	end_date String<binary-string> : couldn't convert string [???] to a date using format [yyy]

Close

Step error handling settings

Error handling stepname: Select values

Target step: Write to log

Enable the error handling?

Nr of errors fieldname:

Error descriptions fieldname:

Error fields fieldname:

Error codes fieldname:

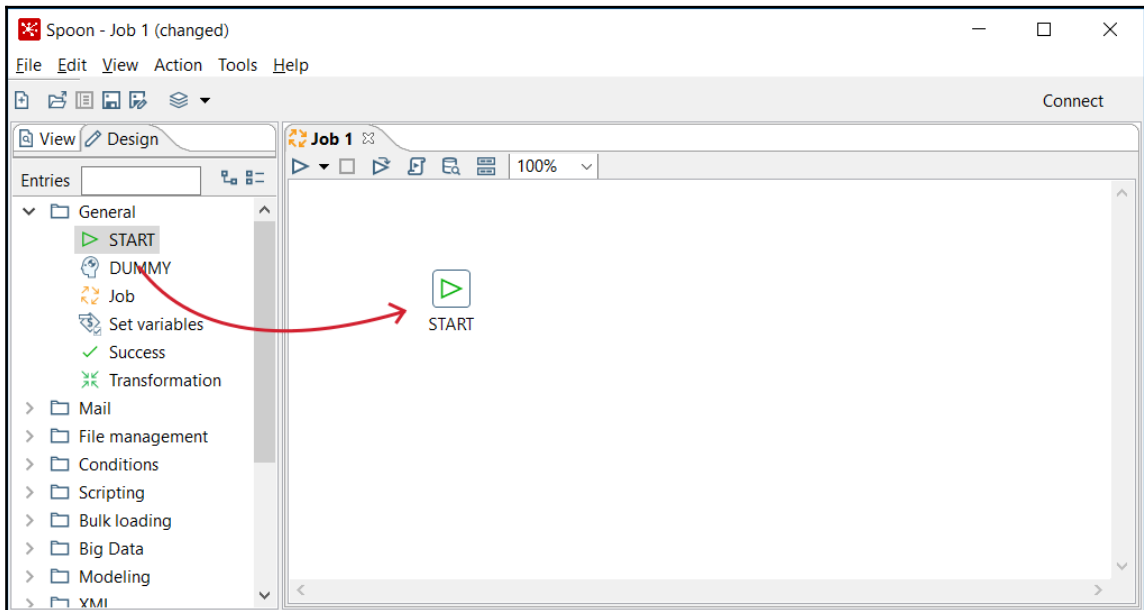
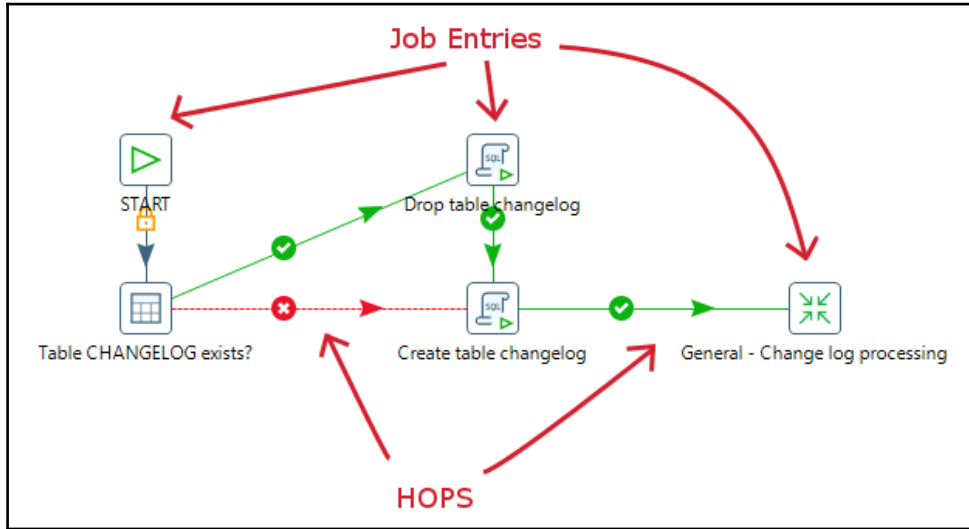
Max nr errors allowed:

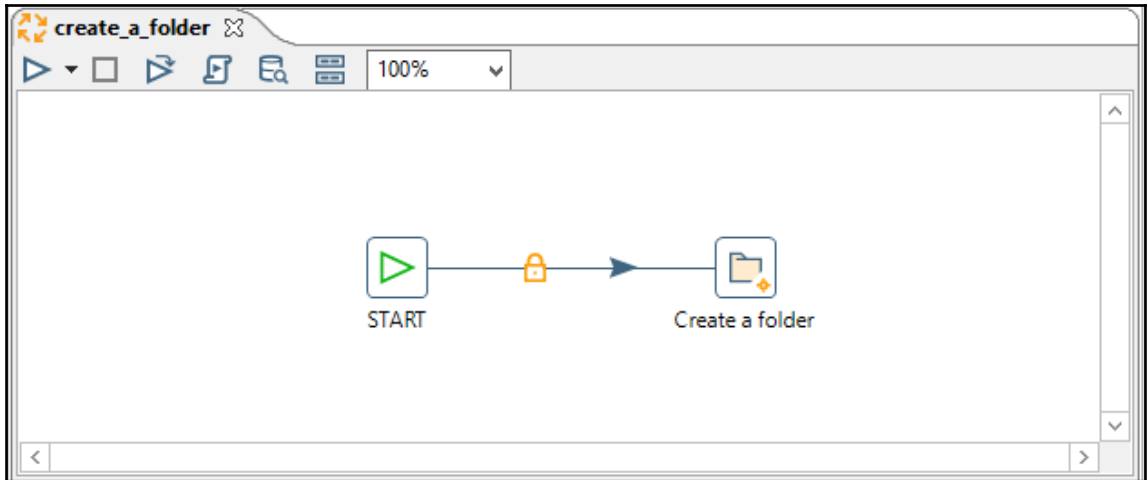
Max % errors allowed (empty==100%):

Min nr of rows to read before doing % evaluation:

OK Cancel

Chapter 03: Creating Basic Task Flows





Execution Results

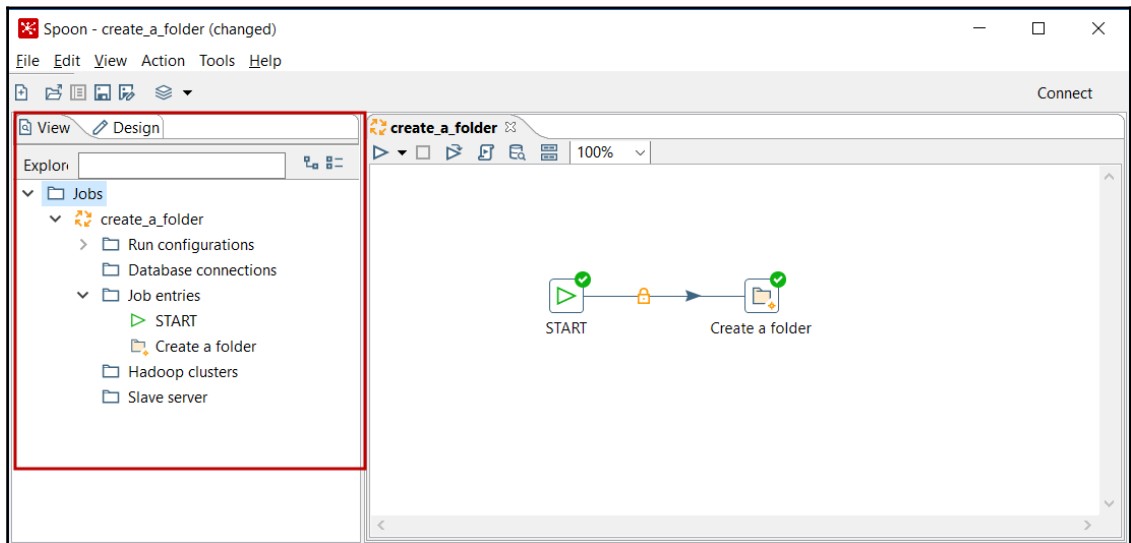
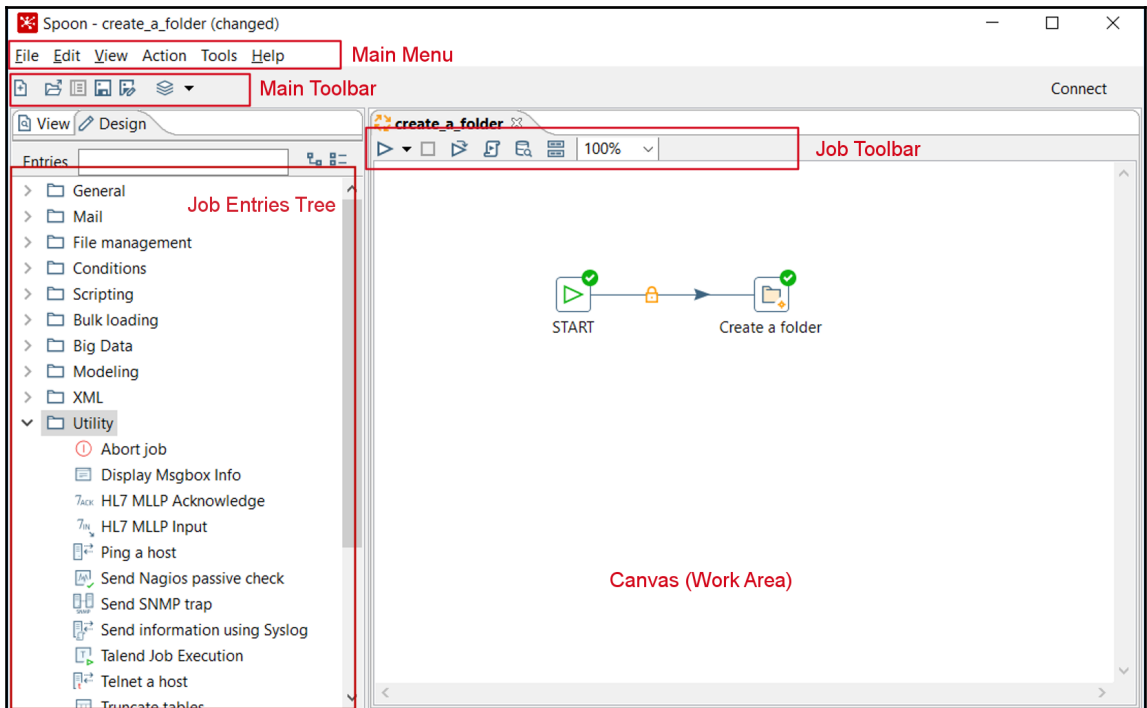
Logging | History | Job metrics | Metrics

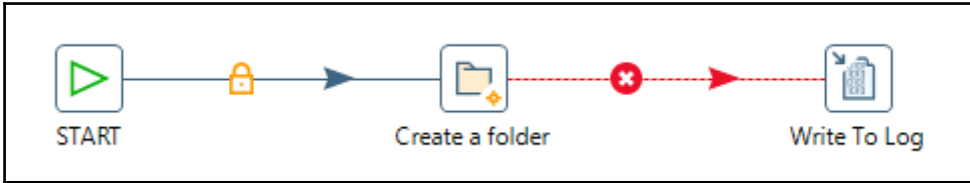
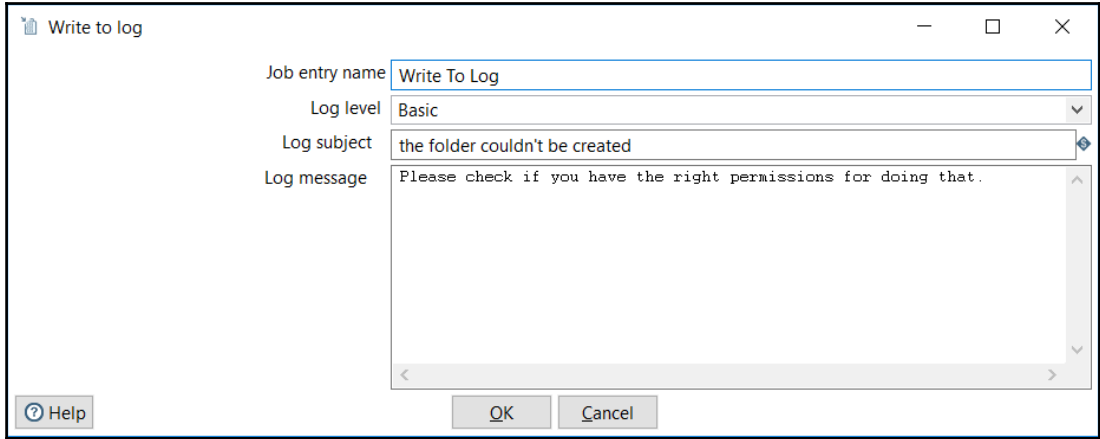
Job / Job Entry	Comment	Result	Reason	Filename
create_a_folder				
Job: create_a_folder	Start of job execution		start	
START	Start of job execution		start	
START	Job execution finished	Success		
Create a folder	Start of job execution		Followed unconditional link	
Create a folder	Job execution finished	Success		
Job: create_a_folder	Job execution finished	Success	finished	

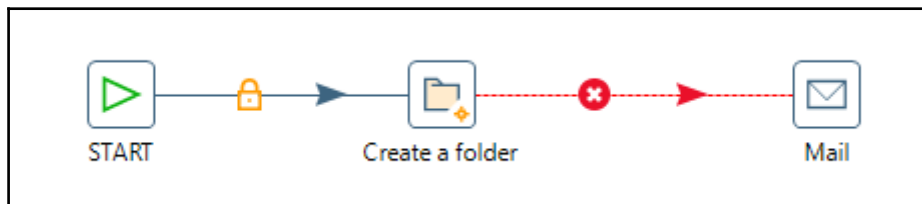
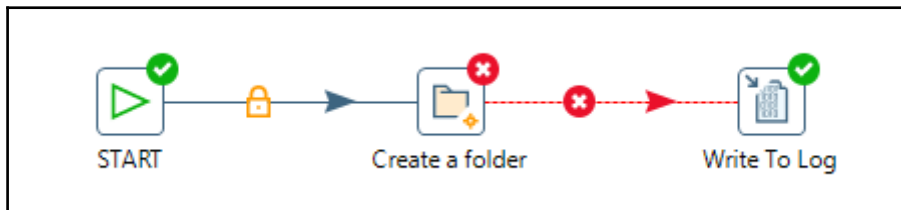
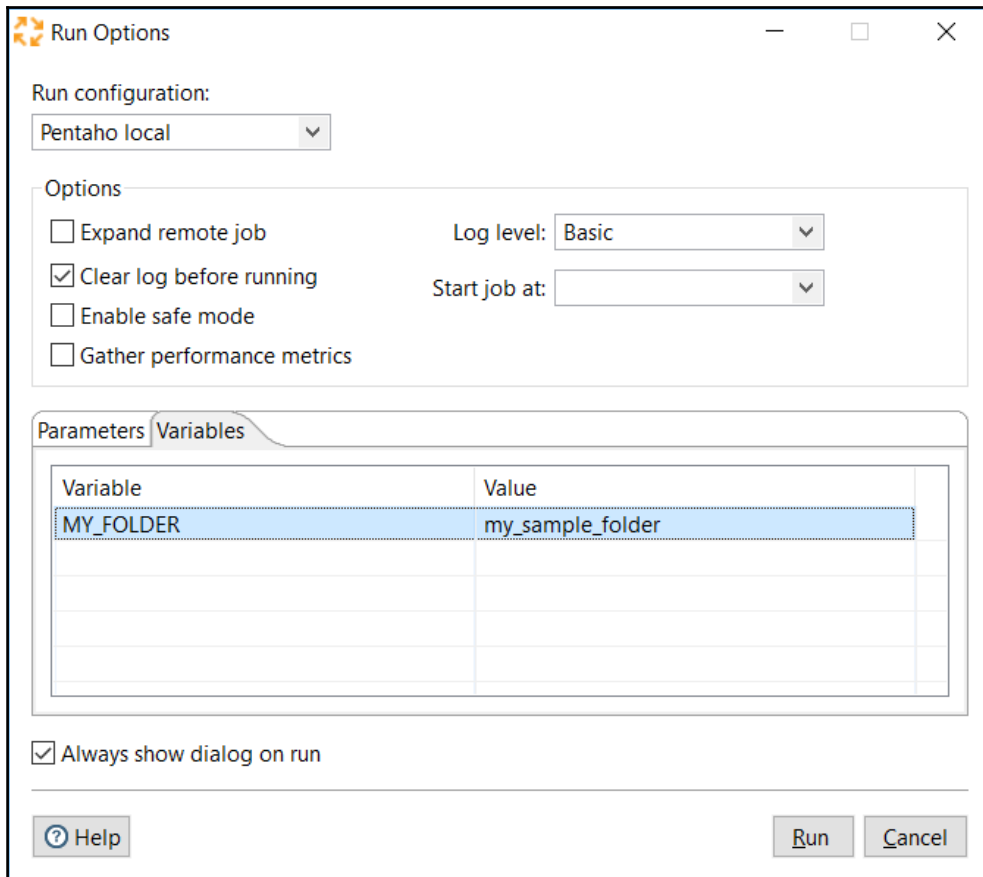
Execution Results

Logging | History | Job metrics | Metrics

2017/11/17 08:33:51 - Spoon - Save file as...
 2017/11/17 08:35:22 - Spoon - Spoon
 2017/11/17 08:35:33 - Spoon - Starting job...
 2017/11/17 08:35:33 - create_a_folder - Start of job execution
 2017/11/17 08:35:33 - create_a_folder - Starting entry [Create a folder]
 2017/11/17 08:35:33 - create_a_folder - Finished job entry [Create a folder] (result=[true])
 2017/11/17 08:35:33 - create_a_folder - Job execution finished
 2017/11/17 08:35:33 - Spoon - Job has ended.







Job mail details

Name of mail job entry: Mail reporting the error

Addresses Server EMail Message Attached Files

SMTP Server

SMTP Server: \${SMTP_SERVER}

Port: \${SMTP_PORT}

Authentication

Use authentication?

Authentication user: \${AUTH_USER}

Authentication password: ●●●●●●●●●●●●●●●●

Use secure authentication?

Secure connection type: SSL

Help OK Cancel

Job mail details

Name of mail job entry: Mail reporting the error

Addresses Server EMail Message Attached Files

Message Settings

Include date in message?

Only send comment in mail body?

Use HTML format in mail body?

Encoding UTF-8

Manage priority

Priority Normal

Importance Normal

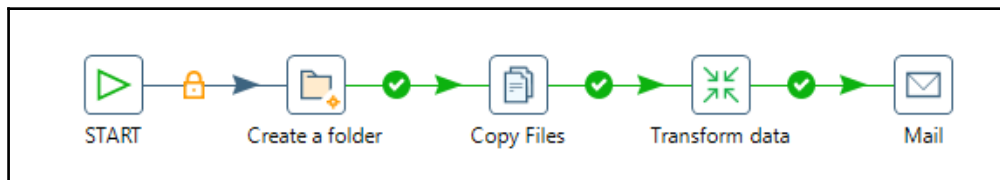
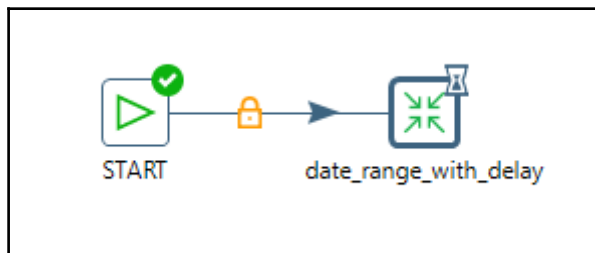
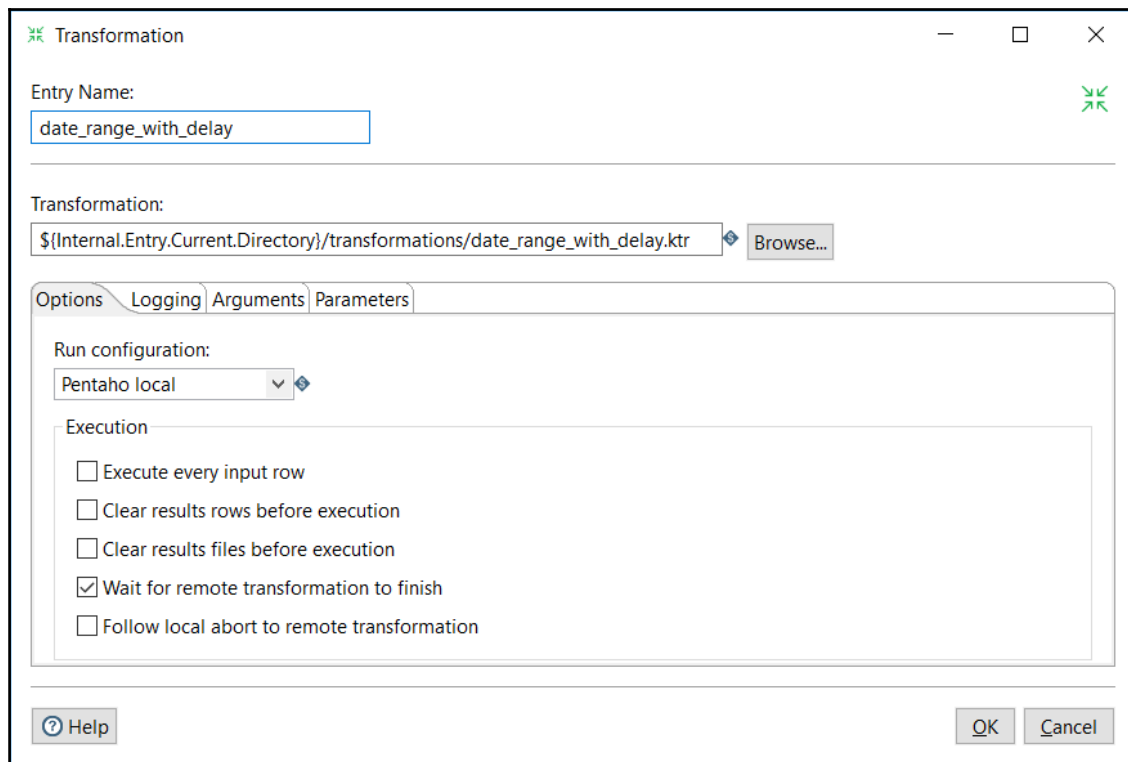
Sensitivity Normal




Message

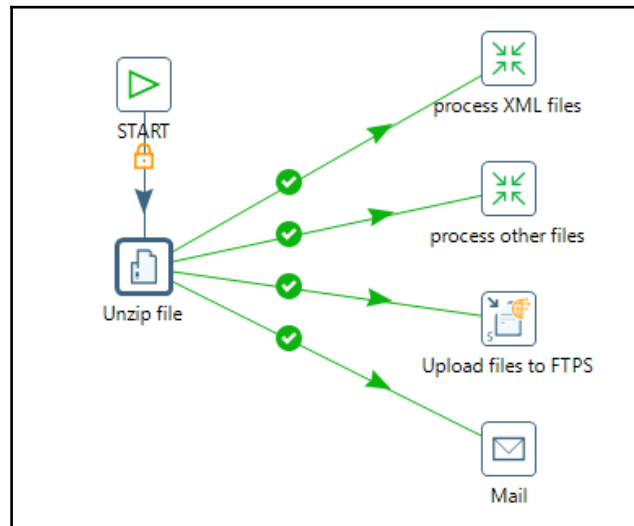
Subject: Pentaho Report

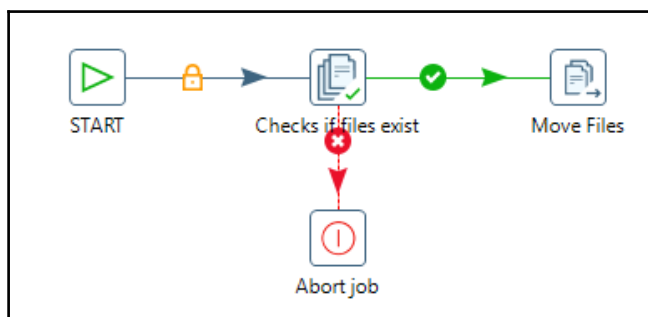
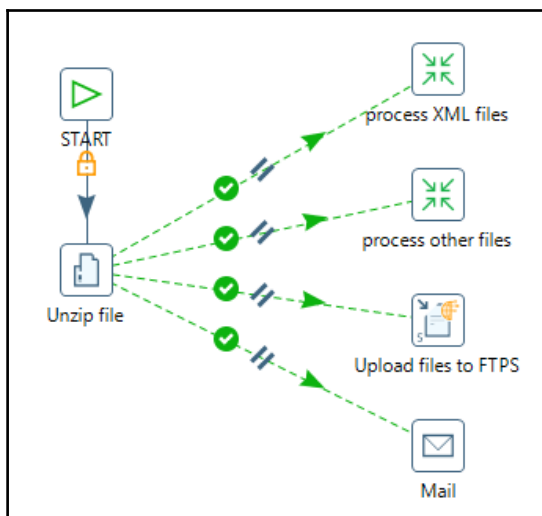
Comment: An error occurred while creating the folder \${MY_FOLDER}

Help OK Cancel



A hop that looks...	And had the following icon...	Represents...	Which means that...
Dark gray		Unconditional execution	The destination entry executes no matter what the result of the previous entry is.
Green		Execution upon success	The destination entry executes only if the previous job entry was successful.
Red		Execution upon failure	The destination entry executes only if the previous job entry failed.





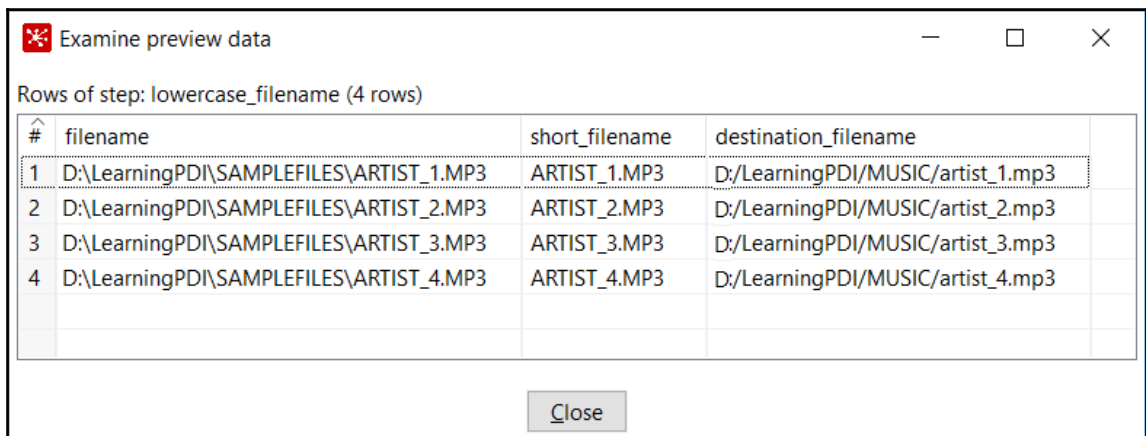
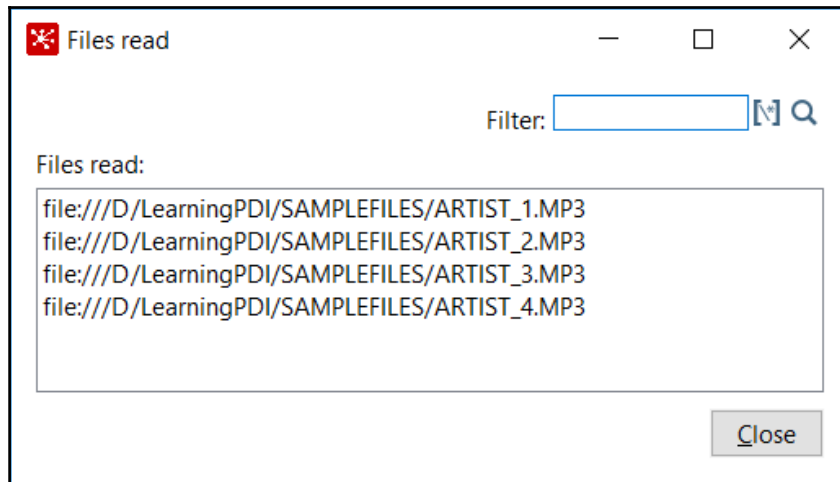
File/Folder source

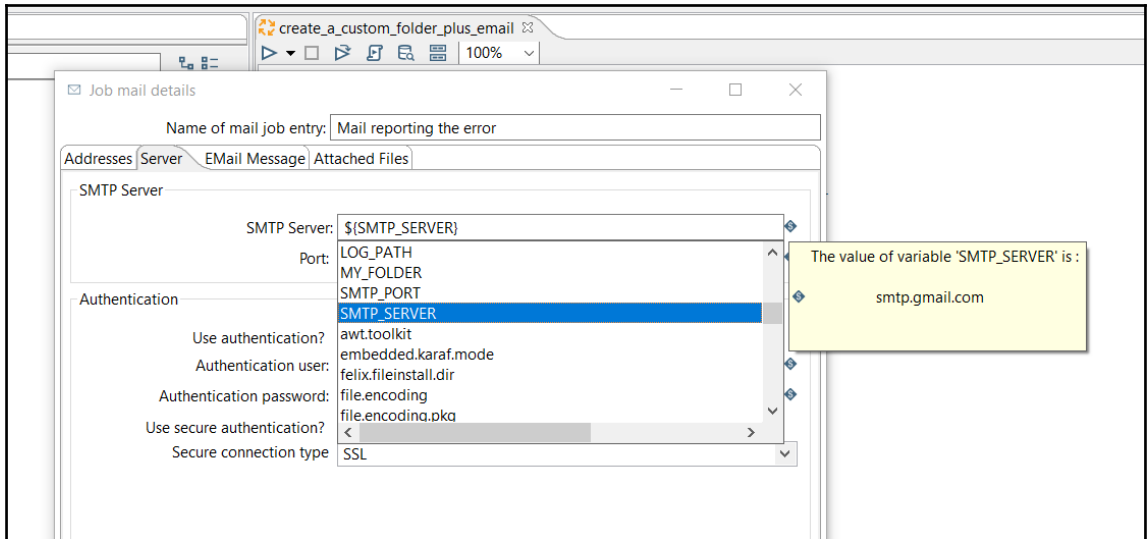
File/Folder destination

Wildcard (RegExp)

Files/Folders:

	File/Folder source	File/Folder destination	Wildcard (RegExp)	
1	D:\LearningPDI\SAMPLEFILES	D:\LearningPDI\MUSIC	ARTIST_*.MP3	<input type="button" value="Delete"/>
				<input type="button" value="Edit"/>





Chapter 04: Reading and Writing Files

File or directory

Regular Expression

Exclude Regular Expression

Selected files:

#	File/Directory	Wildcard (RegExp)	Exclude wildcard	Required	Include subfol
1	D:\LearningPDI\SAMPLEFILES\sales_data.csv			N	N

< III >

Text file input

Step name

File Content Error Handling Filters Fields Additional output fields

Filetype

Separator

Enclosure

Allow breaks in enclosed fields?

Escape

Header Number of header lines

Footer Number of footer lines

Wrapped lines? Number of times wrapped

Paged layout (printout)? Number of lines per page

Document header lines

Compression

No empty rows

Include filename in output? Filename fieldname

Rownum in output? Rownum fieldname

Rownum by file?

Format

Encoding

Length

Limit

Be lenient when parsing dates?

The date format Locale

Result filenames

Add filenames to result

Examine preview data

Rows of step: Read Sales Data (200 rows)

#	ORDERDATE	ORDERNU...	ORDERLI...	PRODUCTCODE	PRODUCTLINE	QUANTI...	PRICEEACH	SALES
1	2/20/2004 0:00	10223	10	S24_4278	Planes	23	74.62	1716.26
2	11/21/2004 0:...	10337	3	S18_4027	Classic Cars	36	100	5679.36
3	6/16/2003 0:00	10131	2	S700_4002	Planes	26	85.13	2213.38
4	7/6/2004 0:00	10266	5	S18_1984	Classic Cars	49	100	6203.4
5	10/16/2004 0:...	10310	4	S24_2972	Classic Cars	33	41.91	1383.03
6	12/4/2004 0:00	10353	4	S700_2834	Planes	48	68.8	3302.4
7	1/20/2005 0:00	10370	8	S12_1666	Trucks and Buses	49	100	8470.14
8	3/11/2004 0:00	10229	6	S24_2300	Trucks and Buses	48	100	5704.32
9	7/19/2004 0:00	10270	6	S12_1666	Trucks and Buses	28	100	4094.72
10	8/25/2003 0:00	10145	4	S32_4485	Motorcycles	27	100	3251.34

Close Show Log

Text file input

Step name: Read Sales Data

#	Name	Type	Format	Posit...	Length	Pr...	Curre...	Decimal	Group	Null if	De...	Trim type	Repeat
1	ORDERDATE	Date	MM/dd/yyyy		15		\$.	,	-		both	N
2	ORDERNUMBER	Integer	#		15		\$.	,	-		both	N
3	ORDERLINENUMBER	Integer	#		15		\$.	,	-		both	N
4	PRODUCTCODE	String			9		\$.	,	-		both	N
5	PRODUCTLINE	String			16		\$.	,	-		both	N
6	QUANTITYORDERED	Integer	#		15		\$.	,	-		both	N
7	PRICEEACH	Number	###		5		\$.	,	-		both	N
8	SALES	Number	###		7		\$.	,	-		both	N

Get Fields Minimal width

Help OK Preview rows Cancel

File or directory Add Browse...

Regular Expression

Exclude Regular Expression

Selected files:

#	File/Directory	Wildc...	Ex...	Required	Include subfolders
1	D:\LearningPDI\SAMPLEFILES\sales_data_APAC.csv			N	N
2	D:\LearningPDI\SAMPLEFILES\sales_data_EMEA.csv			N	N
3	D:\LearningPDI\SAMPLEFILES\sales_data_Japan.csv			N	N

Delete Edit

Examine preview data — □ ×

Rows of step: Read several Sales Data (613 rows)

#	ORDERDATE	ORDERNU...	ORD...	PRODUCTCODE	PRODUCTLINE	QUAN...	PRICEEACH	SALES	file_name
77	11/04/2003	10169	9	S32_2206	Motorcycles	26	39.83	1035.58	sales_data_APAC.csv
78	11/24/2004	10342	8	S18_3782	Motorcycles	26	55.95	1454.7	sales_data_APAC.csv
79	07/16/2003	10139	2	S18_3136	Vintage Cars	20	90.06	1801.2	sales_data_APAC.csv
80	11/21/2003	10193	2	S18_3136	Vintage Cars	23	100	2769.89	sales_data_APAC.csv
81	07/19/2004	10270	11	S18_3136	Vintage Cars	38	100	4775.08	sales_data_APAC.csv
82	09/11/2003	10148	4	S24_2766	Classic Cars	21	73.6	1545.6	sales_data_APAC.csv
83	11/01/2004	10316	8	S24_3949	Planes	30	77.79	2333.7	sales_data_EMEA.csv
84	05/07/2003	10121	5	S10_1678	Motorcycles	34	81.35	2765.9	sales_data_EMEA.csv
85	05/26/2004	10252	5	S18_3782	Motorcycles	31	52.84	1638.04	sales_data_EMEA.csv
86	11/11/2003	10180	9	S10_1678	Motorcycles	29	86.13	2497.77	sales_data_EMEA.csv
87	09/30/2004	10299	1	S24_4620	Classic Cars	32	80.84	2586.88	sales_data_EMEA.csv

Close Show Log

Regular Expression

Exclude Regular Expression

Selected files:

#	File/Directory	Wildcard (RegExp)	Exclude wildcard	Required	Include subfolders
1	D:\LearningPDI\SAMPLEFILES\	sales_data_+\.csv		N	N

Files read — □ ×

Filter: 🔍

Files read:

```
file:///D:/LearningPDI/SAMPLEFILES/sales_data_APAC.csv
file:///D:/LearningPDI/SAMPLEFILES/sales_data_EMEA.csv
file:///D:/LearningPDI/SAMPLEFILES/sales_data_Japan.csv
```

Close

Get data from XML

Step name: Get data from XML

File Content Fields Additional output fields

#	Name	XPath	Element	Result type	Type	Format	Length
1	filename	.	Node	Value of	None		

Examine preview data

Rows of step: Get data from XML (1 rows)

#	filename
1	sales_data_Japan

Close Show Log

Help OK Preview rows Cancel

User Defined Java Expression

Step name: filename

Fields:

#	New field	Java expression	Value type	Length	Precision	Replace value
1	filename	"D:/LearningPDI/SAMPLEFILES/" + filename + ".csv"	String			filename

Help OK Cancel

Accept filenames from previous steps

Accept filenames from previous step

Pass through fields from previous step

Step to read filenames from

Field in the input to use as filename

Calculator

Step name

Fields:

#	New field	Calculation	Field A	Field B	F...	Value type	L...	Pre...	Remove	Conversion mask
1	a_single_date	Date A + B Days	start_date	delta		Date			N	yyyy-MM-dd
2	year	Year of date A	a_single_date			Integer			N	
3	month	Month of date A	a_single_date			Integer			N	
4	day	Day of month of date A	a_single_date			Integer			N	
5	week_of_year	Week of year of date A	a_single_date			Integer			N	
6	quarter	Quarter of date A	a_single_date			Integer			N	

< >

Text file output

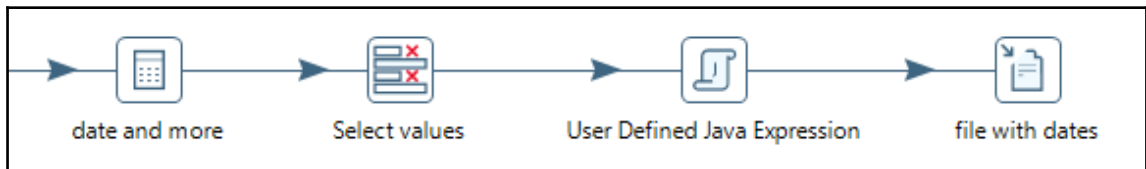
Step name: file with dates

File Content Fields

#	Name	Type	Format	Length	Precision	Curre
1	a_single_date	Date	yyyy-MM-dd			
2	year	Integer	#			
3	month	Integer	#			
4	day	Integer	#			
5	week_of_year	Integer	#			
6	quarter	Integer	#			

Get Fields Minimal width

Help OK Cancel



Select & Alter Remove Meta-data

Fields to alter the meta-data for:

#	Fieldname	Rename to	Type	Length	Precisi...	Binary to N...	Format
1	start_date		String			N	yyyyMMdd
2	end_date		String			N	yyyyMMdd

Get fields to change

Examine preview data

Rows of step: filename (200 rows)

#	ORDERDATE	ORDERNU...	ORD...	PRODUCTCODE	PRODUCTLINE	QUA...	PRICEEACH	SALES	filename
1	02/20/2004	10223	10	S24_4278	Planes	23	74.62	1716.26	D:/LearningPDI/files/sales_Planes
2	11/21/2004	10337	3	S18_4027	Classic Cars	36	100	5679.36	D:/LearningPDI/files/sales_Classic Cars
3	06/16/2003	10131	2	S700_4002	Planes	26	85.13	2213.38	D:/LearningPDI/files/sales_Planes
4	07/06/2004	10266	5	S18_1984	Classic Cars	49	100	6203.4	D:/LearningPDI/files/sales_Classic Cars
5	10/16/2004	10310	4	S24_2972	Classic Cars	33	41.91	1383.03	D:/LearningPDI/files/sales_Classic Cars
6	12/04/2004	10353	4	S700_2834	Planes	48	68.8	3302.4	D:/LearningPDI/files/sales_Planes
7	01/20/2005	10370	8	S12_1666	Trucks and Buses	49	100	8470.14	D:/LearningPDI/files/sales_Trucks and Buses
8	03/11/2004	10229	6	S24_2300	Trucks and Buses	48	100	5704.32	D:/LearningPDI/files/sales_Trucks and Buses
9	07/19/2004	10270	6	S12_1666	Trucks and Buses	28	100	4094.72	D:/LearningPDI/files/sales_Trucks and Buses
10	08/25/2003	10145	4	S32_4485	Motorcycles	27	100	3251.34	D:/LearningPDI/files/sales_Motorcycles

Close

Text file output

Step name: files by product line

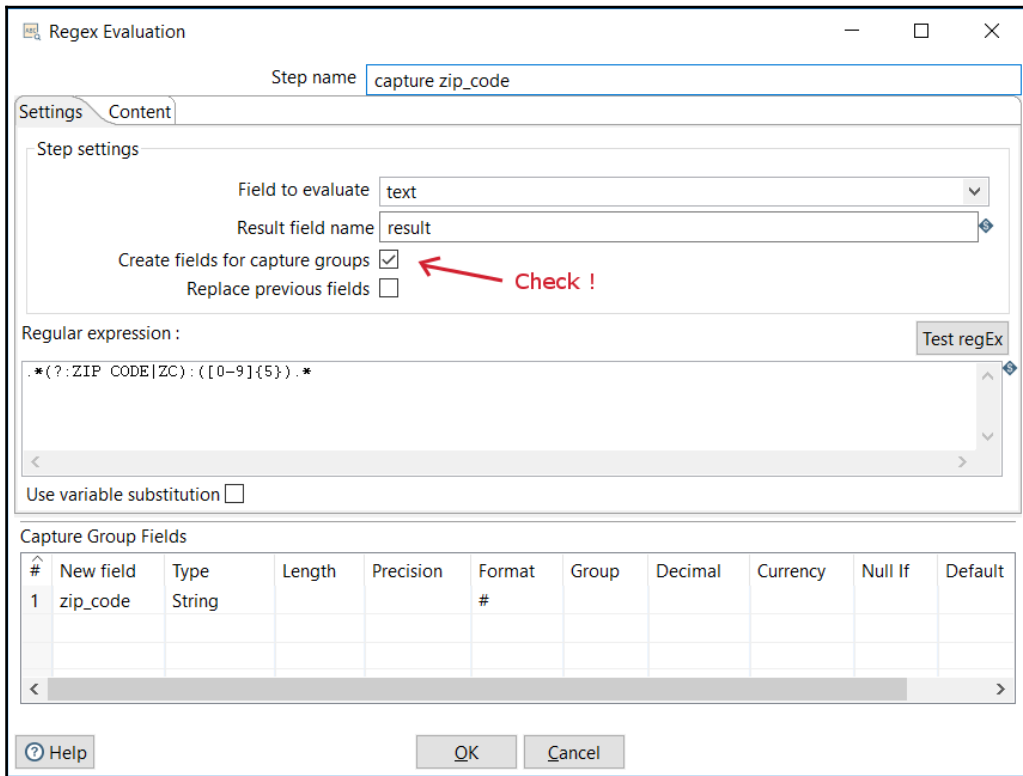
File Content Fields

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim Type	Null
1	PRODUCTLINE	String		16					both	
2	PRODUCTCODE	String		9					both	
3	QUANTITYORDERED	Integer	#	15	0	\$.	,	both	
4	PRICEEACH	String	\$.00	5		\$.	,	both	
5	SALES	Number	\$.00	7		\$.	,	both	

Get Fields Minimal width

Help OK Cancel

Chapter 05: Manipulating PDI Data and Metadata



Examine preview data

Rows of step: capture zip_code (6 rows)

#	text	result	zip_code
1	asdfg	N	<null>
2	AAAA ZC:39391 ----	Y	39391
3	//// ZIP CODE:92000	Y	92000
4	ZC: 1234	N	<null>
5	ZIP CODE UNAVAILABLE	N	<null>
6	101010	N	<null>

Close

Examine preview data

Rows of step: ZIP CODE >> POSTAL CODE (6 rows)

#	text	text2
1	asdfg	asdfg
2	AAAA ZC:39391 ----	AAAA ZC:39391 ----
3	//// ZIP CODE:92000	//// POSTAL CODE:92000
4	ZC: 1234	ZC: 1234
5	ZIP CODE UNAVAILABLE	POSTAL CODE UNAVAILABLE
6	101010	101010

Close

Examine preview data

Rows of step: ZIP CODE | ZC >> POSTAL CODE (6 rows)

#	text	text2
1	asdfg	asdfg
2	AAAA ZC:39391 ----	AAAA POSTAL CODE:39391 ----
3	//// ZIP CODE:92000	//// POSTAL CODE:92000
4	ZC: 1234	POSTAL CODE: 1234
5	ZIP CODE UNAVAILABLE	POSTAL CODE UNAVAILABLE
6	101010	101010

Close

Examine preview data

Rows of step: ZIP CODE | ZC >> POSTAL CODE (uses group reference) (6 rows)

#	text	text2
1	asdfg	asdfg
2	AAAA ZC:39391 ----	AAAA POSTAL CODE:39391 ----
3	//// ZIP CODE:92000	//// POSTAL CODE:92000
4	ZC: 1234	ZC: 1234
5	ZIP CODE UNAVAILABLE	ZIP CODE UNAVAILABLE
6	101010	101010

Close



Calculator

Step name: week_before

Fields:

#	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove	Conversion mask
1	minus7	Set field to constant value A	-7			Integer			N	
2	week_before	Date A + B Days	sample_date	minus7		Date			N	yyyy-MM-dd

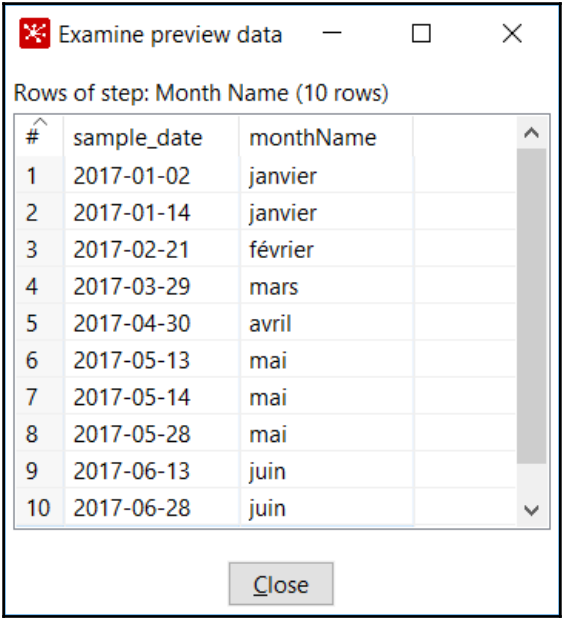
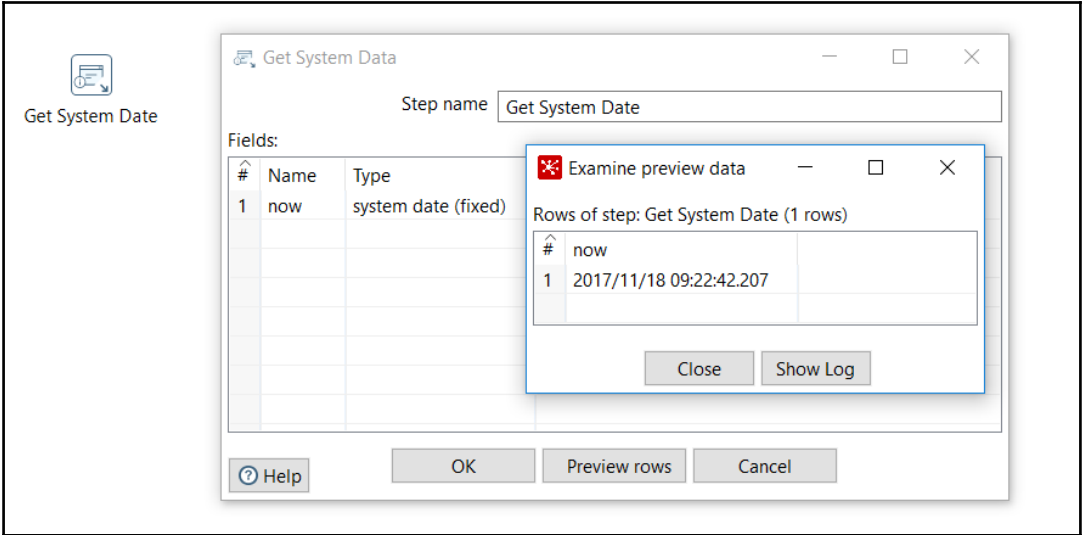
Help OK Cancel

Examine preview data

Rows of step: week_before (10 rows)

#	sample_date	minus7	week_before
1	2017-01-02	-7	2016-12-26
2	2017-01-14	-7	2017-01-07
3	2017-02-21	-7	2017-02-14
4	2017-03-29	-7	2017-03-22
5	2017-04-30	-7	2017-04-23
6	2017-05-13	-7	2017-05-06
7	2017-05-14	-7	2017-05-07
8	2017-05-28	-7	2017-05-21
9	2017-06-13	-7	2017-06-06
10	2017-06-28	-7	2017-06-21

Close

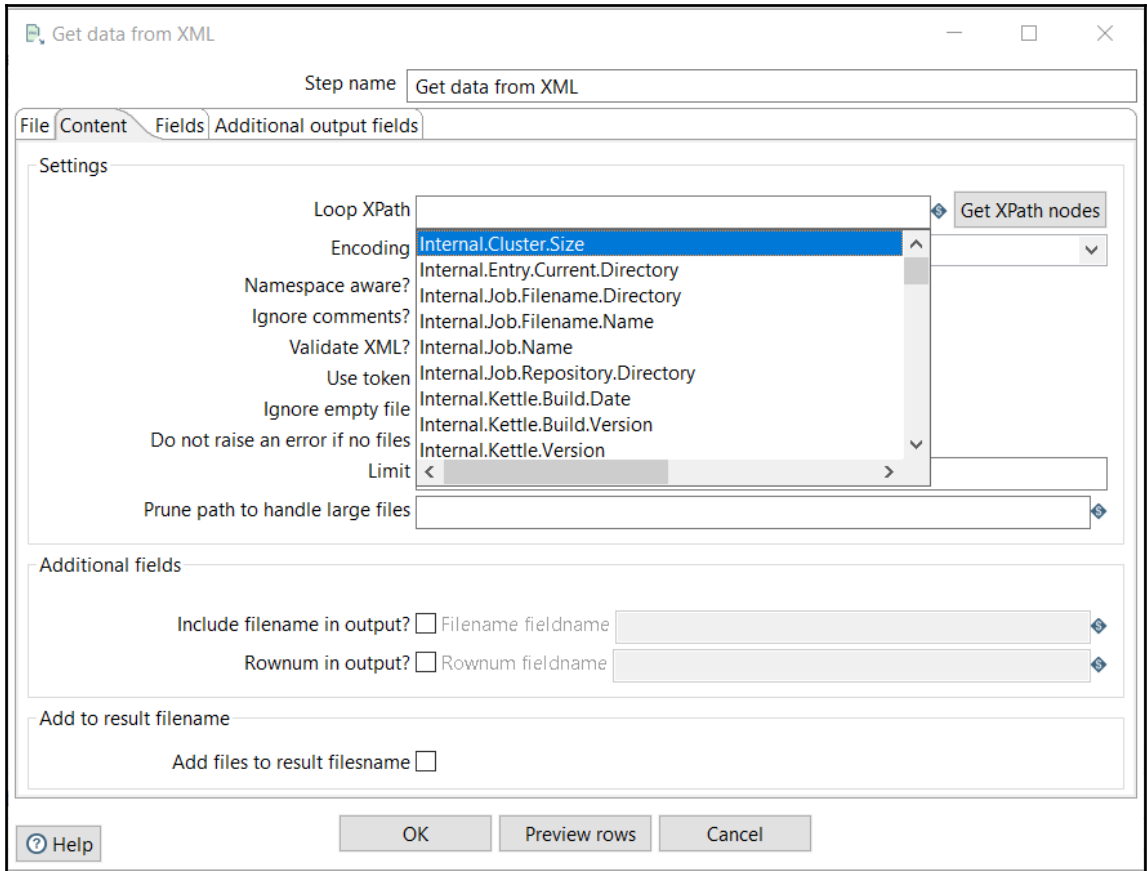


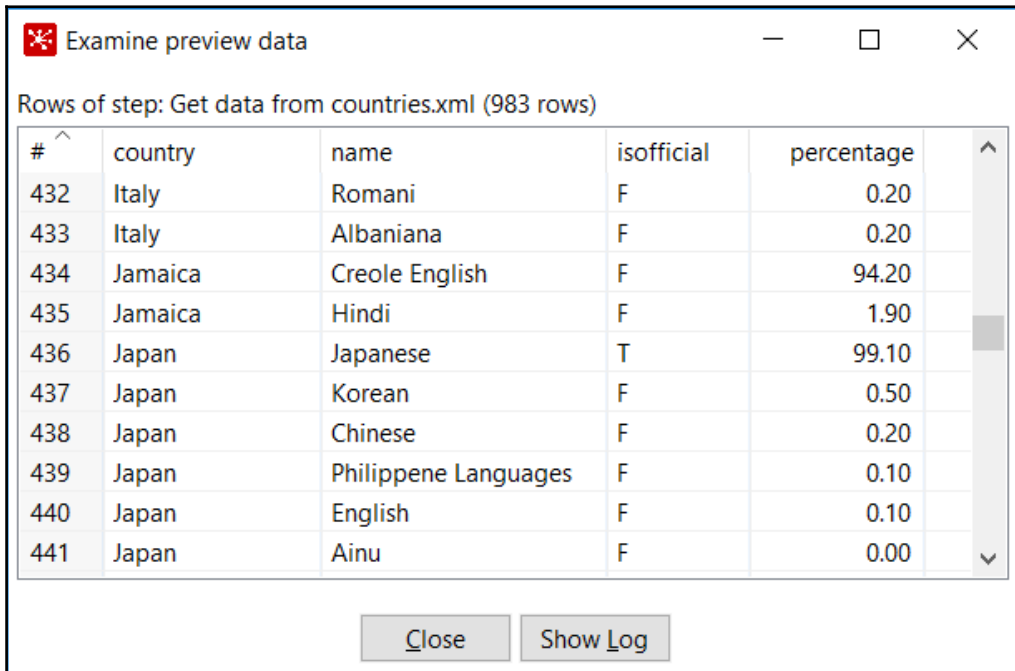
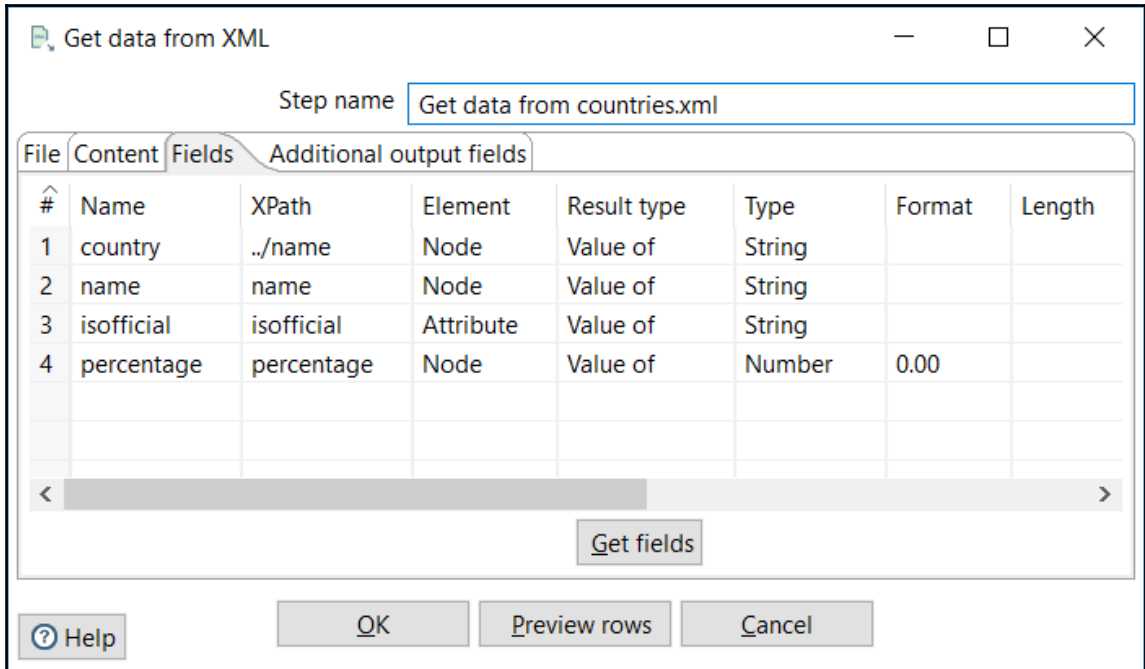
Examine preview data

Rows of step: Select values (181 rows)

#	sample_date
1	2017-01-01
2	2017-01-02
3	2017-01-03
4	2017-01-04
5	2017-01-05
6	2017-01-06
7	2017-01-07
8	2017-01-08
9	2017-01-09
10	2017-01-10

Close





Examine preview data

Rows of step: colors (20 rows)

#	palette_name	color
1	BlackAndWhite	#ffffff
2	BlackAndWhite	#b9b9b9
3	BlackAndWhite	#8e8e8e
4	BlackAndWhite	#727272
5	BlackAndWhite	#000000
6	primary	#f40707
7	primary	#1b00c7
8	primary	#e7e035
9	primary	#17d640
10	primary	#f6f3f3

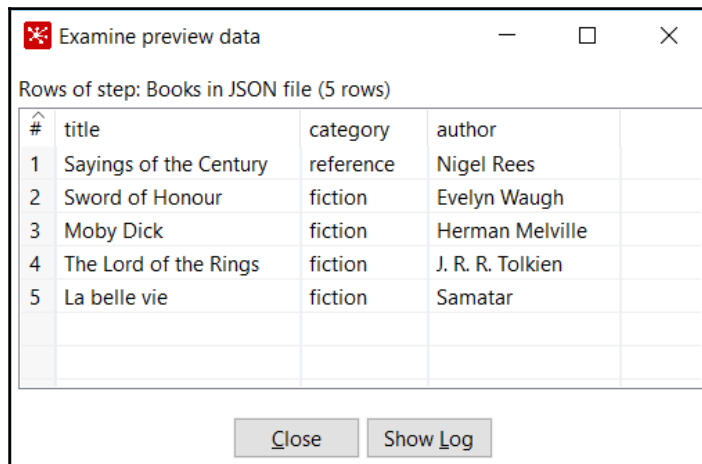
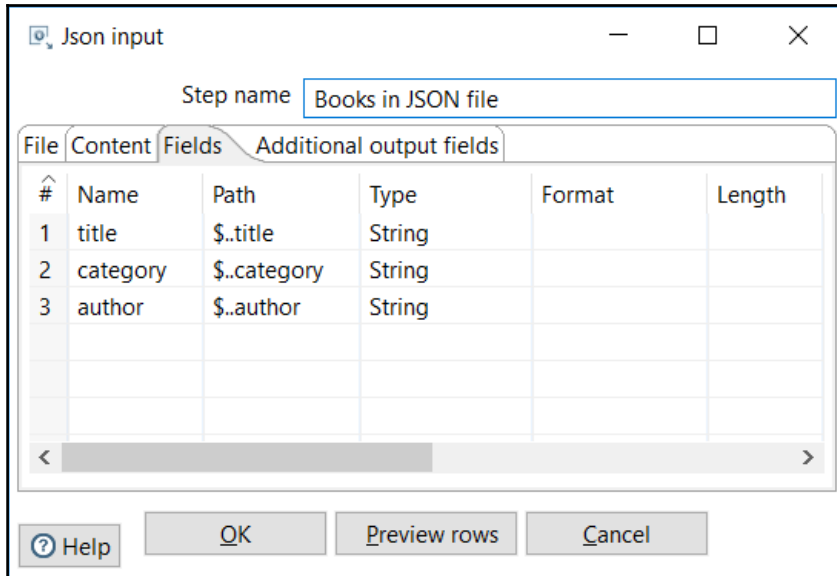
Close

Plugin browser

Plugin type: Step

#	Plugin type	ID	Name	Description
9.	Step	Abort	Abort	Abort a transformation
8.	Step	AccessInput	Microsoft Access Input	Read data from a Microsoft Access file
1.	Step	AccessOutput	Microsoft Access Output	Stores records into an MS-Access database table.
1.	Step	AddXML	Add XML	Encode several fields into an XML fragment
1.	Step	AnalyticQuery	Analytic Query	Execute analytic queries over a sorted dataset (LEAD/LAG/FIRST/LAS
4.	Step	Append	Append streams	Append 2 streams in an ordered way
4.	Step	AutoDoc	Automatic Documentation Outp...	This step automatically generates documentation based on input in t
6.	Step	AvroInputNew	Avro Input	Reads data from Avro file
1.	Step	AvroOutput	Avro Output	Writes data to an Avro file according to a mapping
3.	Step	BlockUntilStepsFinish	Block this step until steps finish	Block this step until selected steps finish.

Close



Chapter 06: Controlling the Flow of Data

Examine preview data

Rows of step: Split field to rows (1000 rows)

#	line	n	word
1	deposit, complicated dislocations of.--Relations between ancient orifices	1	deposit,
2	deposit, complicated dislocations of.--Relations between ancient orifices	1	complicated
3	deposit, complicated dislocations of.--Relations between ancient orifices	1	dislocations
4	deposit, complicated dislocations of.--Relations between ancient orifices	1	of.--Relations
5	deposit, complicated dislocations of.--Relations between ancient orifices	1	between
6	deposit, complicated dislocations of.--Relations between ancient orifices	1	ancient
7	deposit, complicated dislocations of.--Relations between ancient orifices	1	orifices
8	of eruption and subsequent axes of injection.--Iquique, Peru, fossils of,	2	of
9	of eruption and subsequent axes of injection.--Iquique, Peru, fossils of,	2	eruption
1..	of eruption and subsequent axes of injection.--Iquique, Peru, fossils of,	2	and

Close Stop Get more rows

Examine preview data

Rows of step: word_length (1000 rows)

#	n	word	word_length
2	1	complicated	11
3	1	dislocations	12
4	1	ofRelations	11
5	1	between	7
6	1	ancient	7
7	1	orifices	8
8	2	of	2
9	2	eruption	8
1..	2	and	3

Close Stop Get more rows

Filter rows

Step name: length > 4

Send 'true' data to step: [dropdown]

Send 'false' data to step: [dropdown]

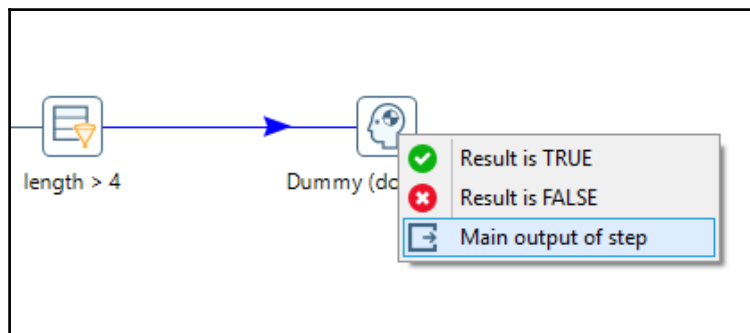
The condition:

[input] +

word_length > [input] [input] (Integer)

[input] 4

[input] Help [input] OK [input] Cancel



Examine preview data

Rows of step: PREVIEW (1000 rows)

#	n	word	word_length
1	1	deposit	7
2	1	complicated	11
3	1	dislocations	12
4	1	ofRelations	11
5	1	between	7
6	1	ancient	7
7	1	orifices	8
8	2	eruption	8
9	2	subsequent	10
1..	2	injectionlquique	16

Close Stop Get more rows

Examine preview data

Rows of step: combined condition (393 rows)

#	n	word	word_length
26	161	geographical	12
27	164	geological	10
28	168	geological	10
29	173	rocks	5
30	176	rocks	5
31	177	geological	10
32	179	geologists	10
33	182	rocks	5
34	191	geologistsand	13
35	192	rocks	5

Close

HITACHI Inspire the Next | Dashboards | Projects | Issues | Agile | Search | Log In

Hitachi Vantara | Pentaho Jira Case Tracking

Search [Save as] [Export] [Tools]

Pentaho Data... | New Feature | Open, In Pro... | Assignee: All | Contains text | More | Advanced

Order by | Pentaho Data Integration - Kettle / PDI-16772 | 1 of 592

update the Java Class PDI step, to use Java 1.8x (instead of Java 1.5)

Details

- Type: New Feature
- Status: OPEN
- Severity: Unknown
- Resolution: Unresolved
- Affects Version/s: None
- Fix Version/s: None
- Component/s: None
- Labels: None
- Story Points: 0
- Notice: When an issue is open.

People

- Assignee: Jens Bleuel
- Reporter: Mark Monroe II
- Votes: Vote for this issue
- Watchers: Start watching this issue

Microsoft Excel input

Step name: Pentaho Platform Tracking.xls

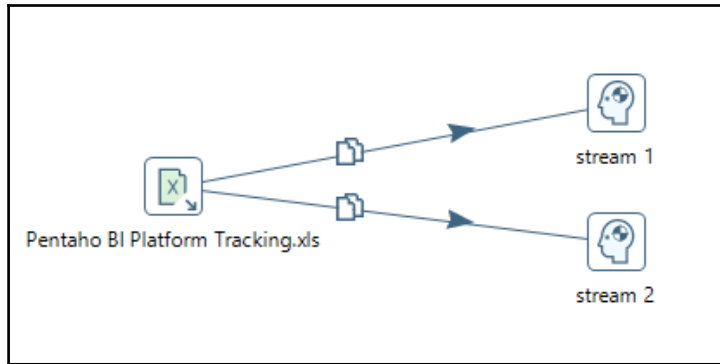
Files | Sheets | Content | Error Handling | Fields | Additional output fields

List of sheets to read

#	Sheet name	Start row	Start column
1		3	0

Get sheetname(s)...

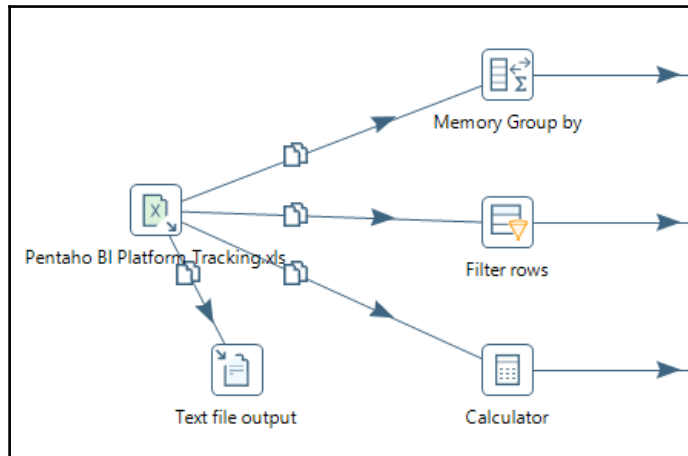
Help | OK | Preview rows | Cancel

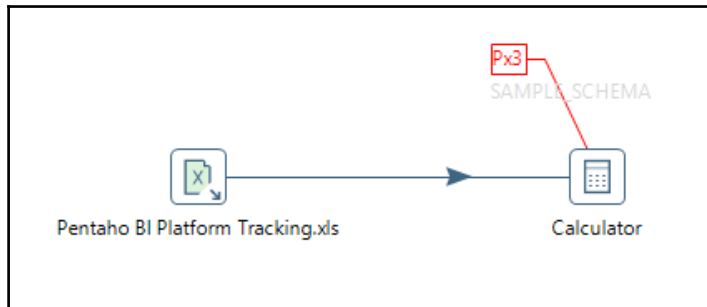
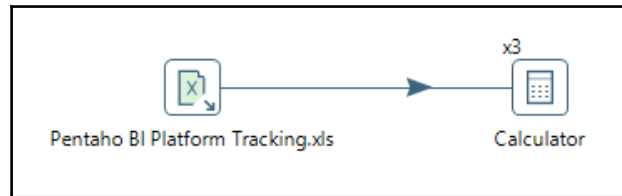
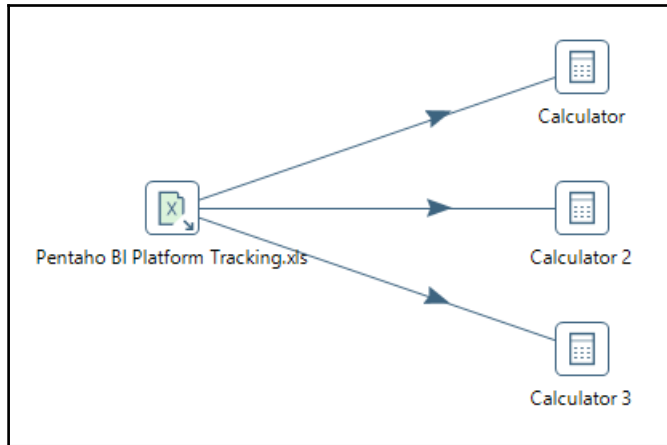


Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

#	Stepname	Copynr	Read	Written	Input	Output	Updated
1	Pentaho BI Platform Tracking.xls	0	0	1186	593	0	0
2	stream 1	0	593	593	0	0	0
3	stream 2	0	593	593	0	0	0

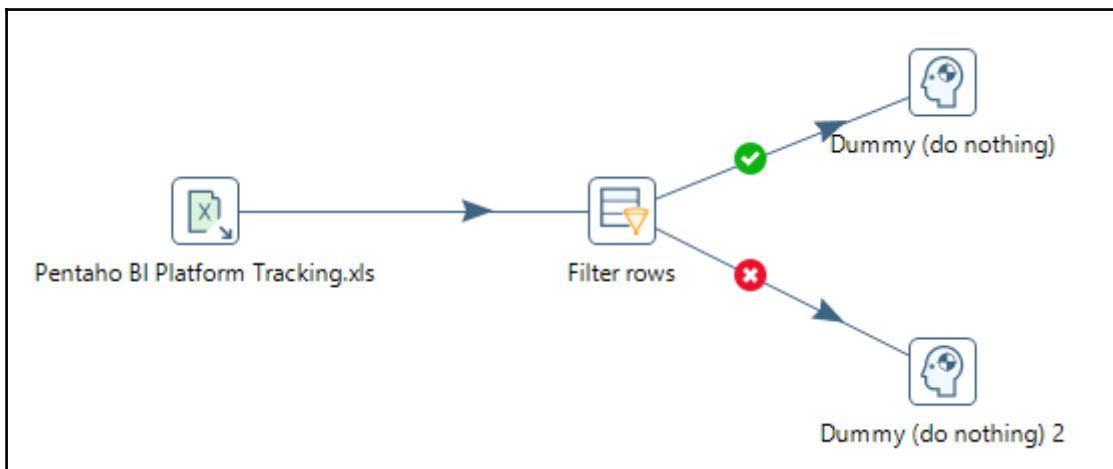




Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

#	Stepname	Copynr	Read	Written	Input	Output	Updated
1	Pentaho BI Platform Tracking.xls	0	0	593	593	0	0
2	Calculator	0	147	147	0	0	0
3	Calculator	1	375	375	0	0	0
4	Calculator	2	71	71	0	0	0



Examine preview data

Rows of step: High and Urgent Issues (113 rows)

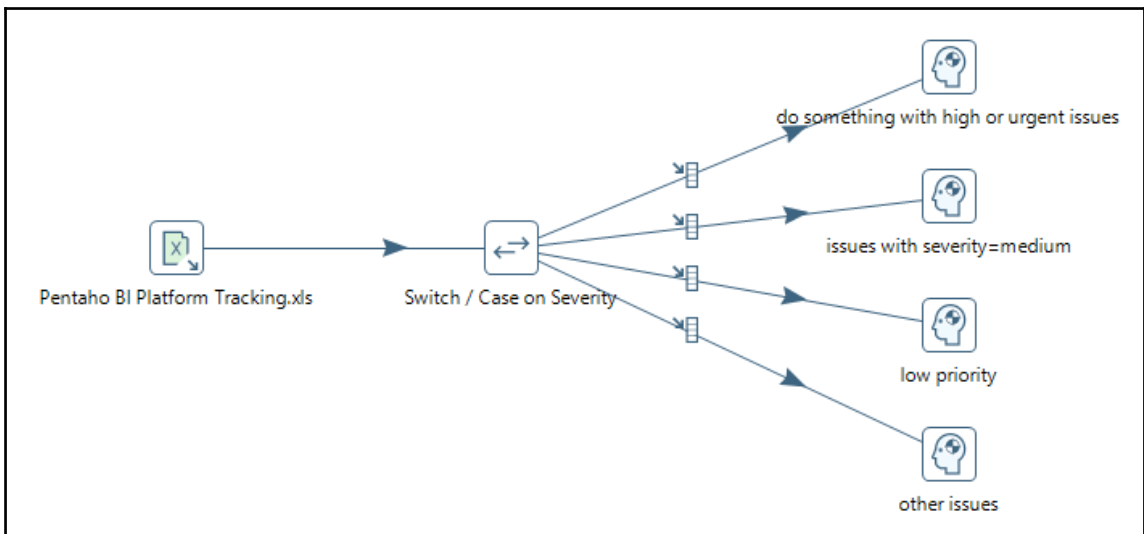
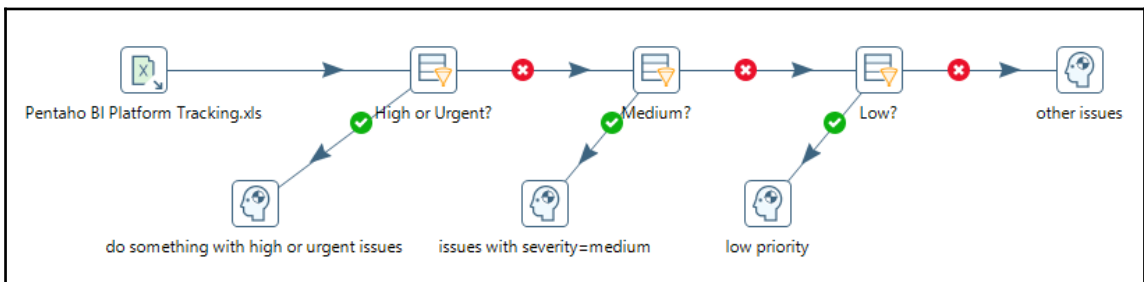
#	Issue Type	Summary	Story ...	Assignee	Sub...	Severity	Status	Resolution
1	New Feature	HBase Output step does not work with Spark AEL	0.0	Unassigned	<n...	Urgent	Open	Unresolved
2	New Feature	XML Output step incorrectly encodes 'less-than' character	0.0	Unassigned	<n...	Urgent	Open	Unresolved
3	New Feature	Provide an option/ability to read hyperlink URL's from an Excel...	0.0	Unassigned	<n...	High	Open	Unresolved
4	New Feature	Add Metadata Injection (MDI) support to the Database Looku...	0.0	Unassigned	<n...	High	Open	Unresolved
5	New Feature	Data Services can't handle SQL "CASE" in ORDER BY	0.0	Unassigned	<n...	High	Open	Unresolved
6	New Feature	Data Services - Can't do operations between fields (example: ...	0.0	Unassigned	<n...	High	Open	Unresolved
7	New Feature	Include Microsoft JDBC driver by default	0.0	Unassigned	<n...	High	Open	Unresolved
8	New Feature	Metadata injection on metadata injection should support dyna...	0.0	Unassigned	<n...	High	Open	Unresolved
9	New Feature	When we output the same field type (date) multiple times onl...	0.0	Unassigned	<n...	High	Open	Unresolved
1..	New Feature	Expand Remote Job via API call	0.0	Unassigned	<n...	High	Open	Unresolved

Close

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

#	Stepname	Copynr	Read	Written	Input	Output	Updated
1	Pentaho BI Platform Tracking.xls	0	0	593	593	0	0
2	High, Urgent?	0	593	593	0	0	0
3	High and Urgent Issues	0	113	113	0	0	0
4	other issues	0	480	480	0	0	0



↔ Switch / case

Step name: Switch / Case on Severity

Field name to switch: Severity

Use string contains comparison:

Case value data type: String

Case value conversion mask:

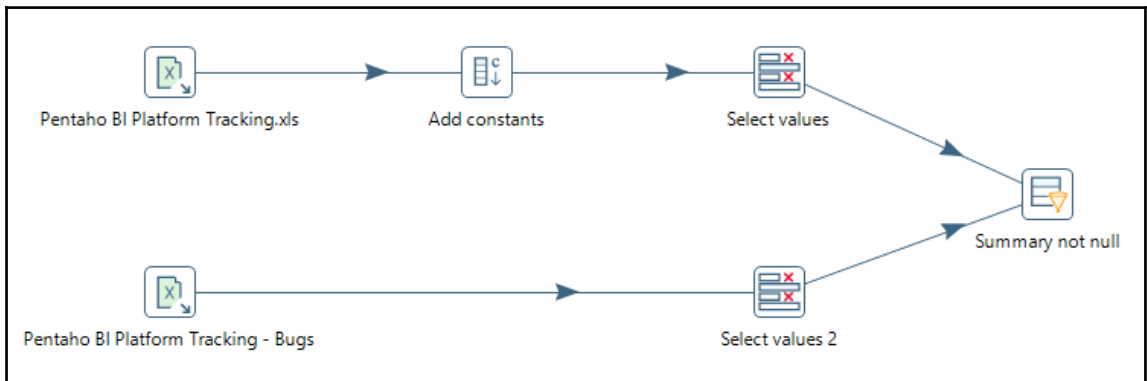
Case value decimal symbol:

Case value grouping symbol:

Case values	#	Value	Target step
	1	High	do something with high or urgent issues
	2	Urgent	do something with high or urgent issues
	3	Medium	issues with severity=medium
	4	Low	low priority

Default target step: other issues

Help OK Cancel



Examine preview data

Rows of step: Summary not null (1000 rows)

#	Issue Type	Summary	Severity	Status	Affects Version/s
5..	New Feature	As an ETL Designer, I want the ability to easily navigate to the transformation logs ass...	High	Open	N/A
5..	New Feature	As an ETL Designer/Administrator, I want a Data Integration Environment (D.I.E)	Medium	Open	N/A
5..	New Feature	As an ETL Designer, I want support for REF CURSOR in Oracle (DB PROC step)	Low	Open	N/A
5..	New Feature	As an ETL designer, I want the ability to define formulas based on values from the stre...	Low	Open	N/A
5..	New Feature	As an ETL Designer, I want manageable 'Favorites' for Steps and Job Entries	Low	Open	N/A
5..	New Feature	As an ETL Designer, I want a Cobol input file step	High	Open	N/A
5..	Bug	java.lang.ClassCastException: org.eclipse.swt.layout.FormData cannot be cast to org.ec...	Unknown	Open	8.0.0 GA
5..	Bug	Typo on SQL popup	Unknown	Open	7.0.0 GA, 7.0.0.4 GA, 8.0.0 GA, Master
5..	Bug	Transformation runs twice when specifying Pentaho Server Run Configuration	Urgent	Open	8.0.0 GA
5..	Bug	Miscalculation with big number data type in calculator step	High	Open	7.1.0.1 GA
5..	Bug	\${InternalJob. [...]} variables are not marked as deprecated in the Spoon UI	Medium	Open	7.1.0 GA, 8.0.0 GA

Close Stop Get more rows

Pentaho BI Platform Tracking.xls → Add constants → Summary not null

Warning!

This hop causes the target step to receive rows with mixed layout!

We detected rows with varying number of fields, this is not allowed in a transformation. The first row cont

OK Details

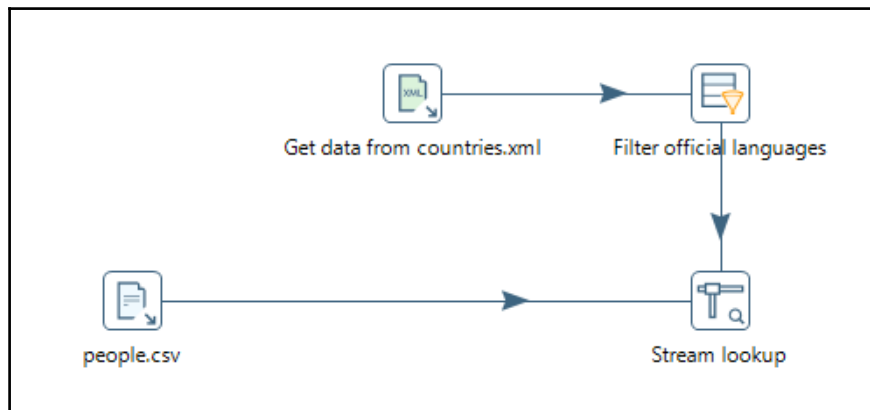
Execution Results

Examine preview data

Rows of step: Filter official languages (237 rows)

#	country	language	isofficial	percentage
1	Afghanistan	Pashto	T	52.40
2	Afghanistan	Dari	T	32.10
3	Albania	Albaniana	T	97.90
4	Algeria	Arabic	T	86.00
5	American Samoa	Samoan	T	90.60
6	American Samoa	English	T	3.10
7	Andorra	Catalan	T	32.30
8	Anguilla	English	T	0.00
9	Antigua and Barbuda	English	T	0.00
1..	Argentina	Spanish	T	96.80

Close



Stream Value Lookup

Step name

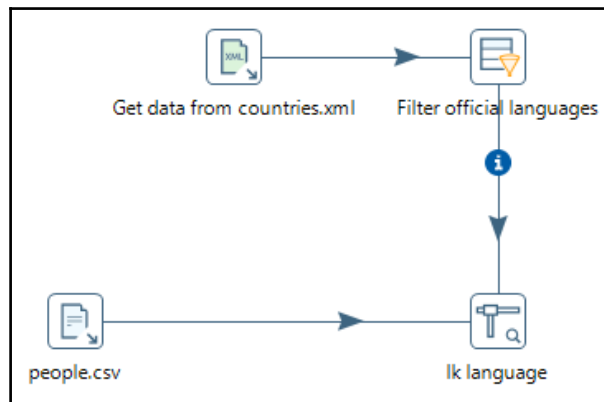
Lookup step

The key(s) to look up the value(s):

#	Field	LookupField
1	Country_Name	country

Specify the fields to retrieve :

#	Field	New name	Default	Type
1	language			String



Examine preview data

Rows of step: lk language (14 rows)

#	ID	Country_Name	Name	language
1	1	Russia	Mikhail Davydova	<null>
2	2	Russia	Anastasia Davydova	<null>
3	3	Spain	Carmen Rodriguez	Spanish
4	4	Spain	Francisco Delgado	Spanish
5	5	Japan	Natsuki Harada	Japanese
6	6	Japan	Emiko Suzuki	Japanese
7	7	China	Lin Jiang	Chinese
8	8	China	Wei Chiu	Chinese
9	9	United States	Chelsea Thompson	English
1..	10	United States	Cassandra Sullivan	English

Close

Chapter 07: Cleansing, Validating, and Fixing Data

Value Mapper

Step name:

Fieldname to use:

Target field name (empty=overwrite):

Default upon non-matching:

Field values:

#	Source value	Target value
1	USA	United States
2	U.S.A.	United States
3	Russia	Russian Federation

User Defined Java Expression

Step name:

Fields:

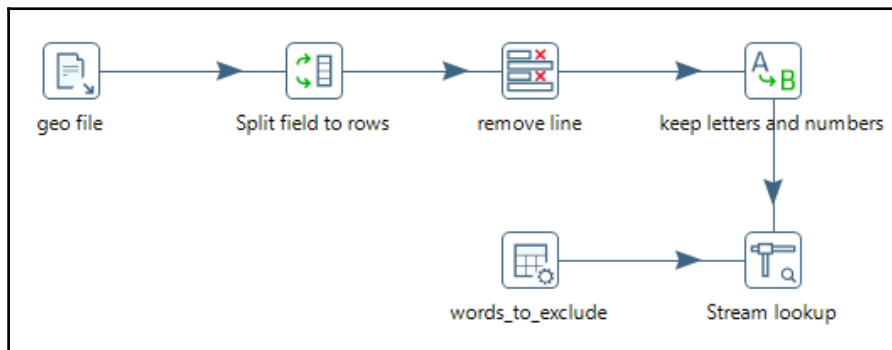
#	New field	Java expression	Value type	Length	Precision	Replace value
1	Country_Name	(new_name == null)?Country_Name:new_name	String			Country_Name

Examine preview data

Rows of step: keep letters and numbers (1000 rows)

#	n	word
2..	4	subsidence
2..	4	with
2..	4	partial
3..	5	elevations
3..	5	during
3..	5	the
3..	5	cretaceoolitic
3..	5	periodOn
3..	5	the
3..	5	elevation
3..	5	and
3..	6	structure
3..	6	of
4..	6	the

Close Stop Get more rows



Stream Value Lookup

Step name: look for word

Lookup step: words_to_exclude

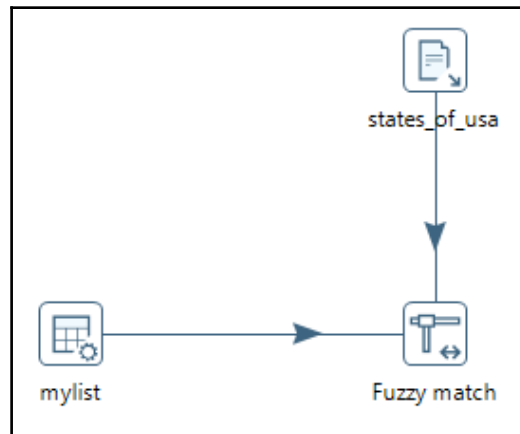
The key(s) to look up the value(s):

#	Field	LookupField
1	word	words_to_exclude

Specify the fields to retrieve :

#	Field	New name	Default	Type
1	words_to_exclude	found		String

Preserve memory (costs CPU)
 Key and value are exactly one integer
 Use sorted list (i.s.o. hashtable)



Fuzzy string search

Step name

General Fields

Lookup stream (source)

Lookup step

Lookup field

Main stream

Main stream field

Settings

Algorithm

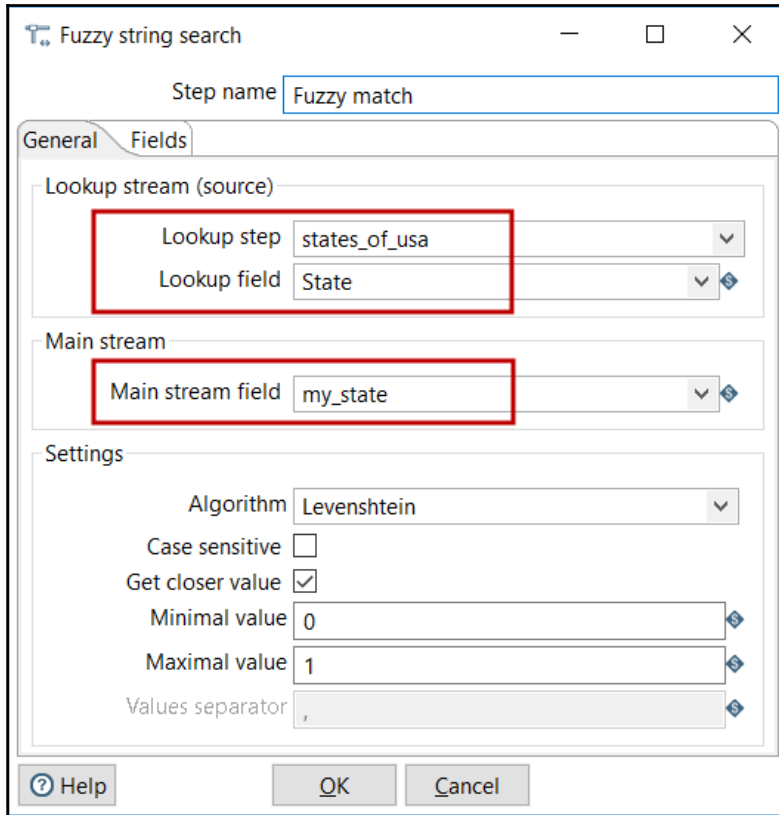
Case sensitive

Get closer value

Minimal value

Maximal value

Values separator



Examine preview data

Rows of step: Fuzzy match (17 rows)

#	my_state	match	distance
1	California	<null>	<null>
2	Colorado	Colorado	1
3	Washington	Washington	0
4	Masachusetts	Massachusetts	1
5	Alsaka	<null>	<null>
6	Conneticut	Connecticut	1
7	Road Island	<null>	<null>
8	Hawai	Hawaii	1
9	Ohio	Ohio	0
1..	Kentuky	Kentucky	1
1..	Pensylvania	Pennsylvania	1
1..	Louisiana	Louisiana	0
1..	Arizonia	Arizona	1
1..	Hawaii	Hawaii	0
1..	Mississippi	Mississippi	1
1..	California	California	0
1..	Howaii	Hawaii	1

Close

Examine preview data

Rows of step: Fuzzy match (17 rows)

#	my_state	match ^	distance
1	California	<null>	<null>
2	Alsaka	<null>	<null>
3	Road Island	<null>	<null>
4	Arizona	Arizona	1
5	California	California	0
6	Calorado	Colorado	1
7	Conneticut	Connecticut	1
8	Hawai	Hawaii	1
9	Hawaii	Hawaii	0
1..	Howaii	Hawaii	1
1..	Kentuky	Kentucky	1
1..	Louisiana	Louisiana	0
1..	Masachusetts	Massachusetts	1
1..	Mississippi	Mississippi	1
1..	Ohio	Ohio	0
1..	Pensylvania	Pennsylvania	1
1..	Washington	Washington	0

Close

User Defined Java Expression

Step name:

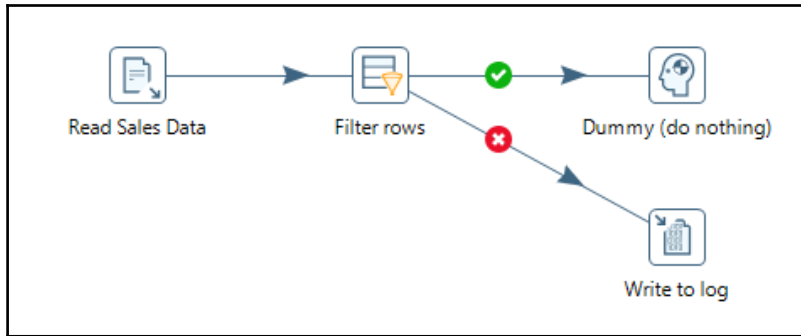
Fields:

#	New field	Java expression	Value type	Length	Precision	Replace value
1	my_state	(match== null)?my_state:match	String			my_state

Examine preview data

Rows of step: Sort rows (14 rows)

#	my_state
1	Alaska
2	Arizona
3	California
4	Colorado
5	Connecticut
6	Hawaii
7	Kentucky
8	Louisiana
9	Massachusetts
1..	Mississippi
1..	Ohio
1..	Pennsylvania
1..	Road Island
1..	Washington



The condition:

(String)

Write to log

Step name

Log level

Print header

Limit rows?

Nr of rows to print

Write to log

Fields

#	Field
1	ORDERNUMBER
2	ORDERLINENUMBER
3	PRODUCTCODE

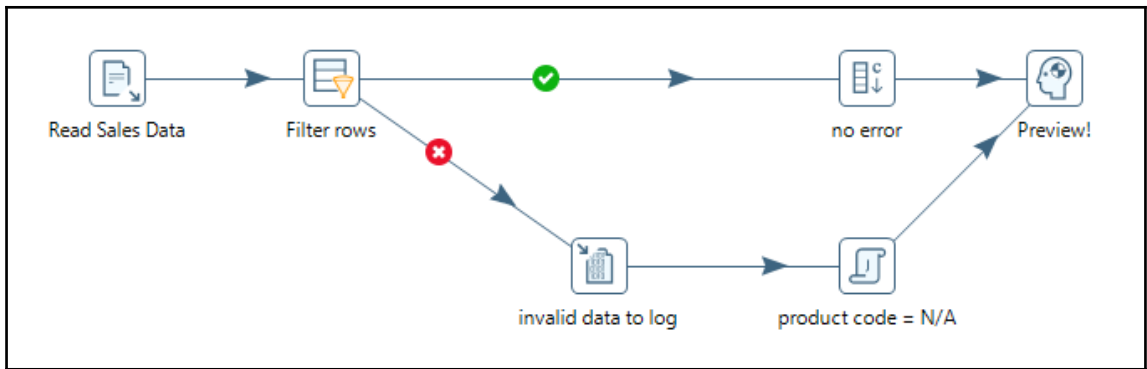
User Defined Java Expression

Step name:

Fields:

#	New field	Java expression	Value type	Length	Precision	Replace value
1	PRODUCTCODE	"N/A"	String			PRODUCTCODE
2	ERR_FLAG	true	Boolean			
3	ERR_DESC	"Invalid product code"	String			

Help OK Cancel



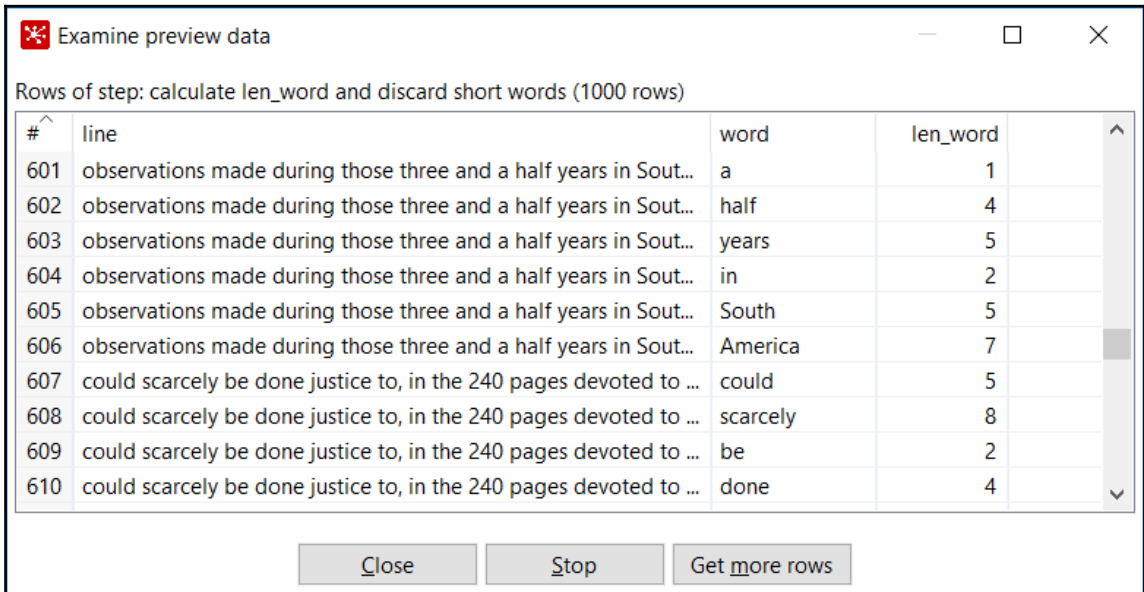
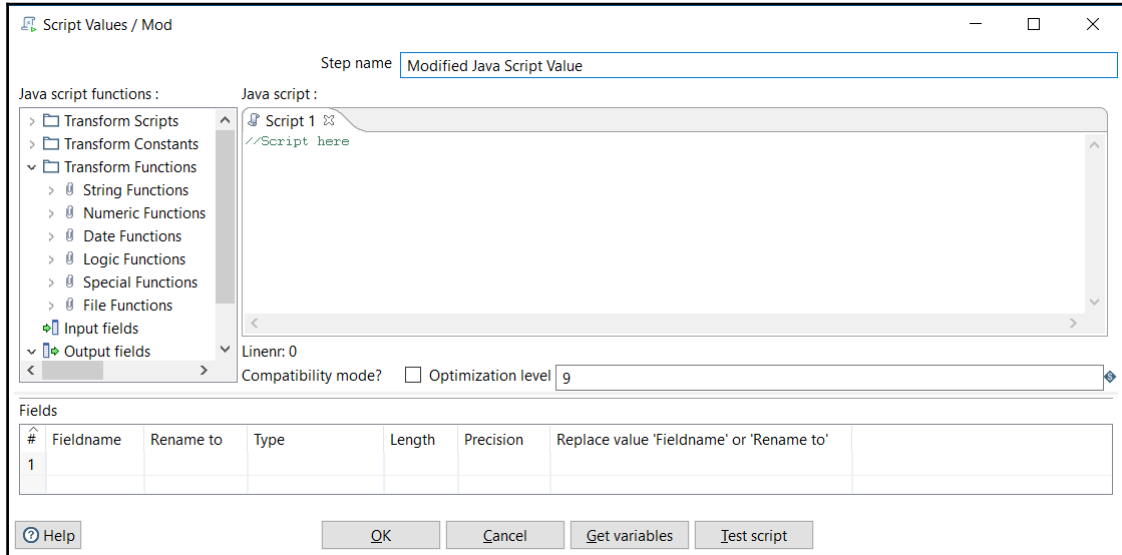
Examine preview data

Rows of step: Preview! (200 rows)

#	ORDERDATE	ORDERN...	ORDE...	PRODUCTCODE	PRODUCTLINE	QUA...	PRICEEACH	SALES	ERR_FLAG	ERR_DESC
1..	08/21/2004	10284	1	S18_2581	Planes	31	71.81	2226.11	N	
1..	05/09/2005	10415	3	S72_1253	Planes	42	57.61	2419.62	N	
1..	02/17/2005	10381	2	S18_1097	Trucks and Buses	48	98	4704	N	
1..	03/19/2004	10231	2	S12_1108	Classic Cars	42	100	8378.58	N	
1..	09/09/2004	10293	6	S18_3259	Trains	22	100	2418.24	N	
1..	06/16/2003	10131	2	N/A	Planes	26	85.13	2213.38	Y	Invalid product code
1..	12/04/2004	10353	4	N/A	Planes	48	68.8	3302.4	Y	Invalid product code
1..	03/30/2005	10398	7	N/A	Ships	36	100	3910.32	Y	Invalid product code
1..	10/06/2003	10155	3	N/A	Planes	44	85.87	3778.28	Y	Invalid product code
1..	06/15/2004	10258	1	N/A	Classic Cars	45	80.92	3641.4	Y	Invalid product code
1..	12/17/2004	10361	11	N/A	Planes	35	100	4277.35	Y	Invalid product code
1..	12/04/2004	10353	9	N/A	Planes	39	100	5043.87	Y	Invalid product code

Close

Chapter 08: Manipulating Data by Coding



#	Fieldname	Rename to	Type	Length	Precision	Replace value 'Fieldname' or 'Rename to'
1	len_word		Integer		0	N
2	u_word	word	String			Y

Examine preview data

Rows of step: calculate len_word and discard short words (1000 rows)

#	line	word	len_word
601	observations made during those three and a half years in South America	A	1
602	observations made during those three and a half years in South America	HALF	4
603	observations made during those three and a half years in South America	YEARS	5
604	observations made during those three and a half years in South America	IN	2
605	observations made during those three and a half years in South America	SOUTH	5
606	observations made during those three and a half years in South America	AMERICA	7
607	could scarcely be done justice to, in the 240 pages devoted to their	COULD	5
608	could scarcely be done justice to, in the 240 pages devoted to their	SCARCELY	8
609	could scarcely be done justice to, in the 240 pages devoted to their	BE	2
610	could scarcely be done justice to, in the 240 pages devoted to their	DONE	4

Close Stop Get more rows

Java script:

```
Script 1
var len_word = word.length;
var u_word = upper(word);
```

- Add new
- Add copy
- Set Transform Script
- Set Start Script
- Set End Script
- Remove Script Type

Java script functions :

- Transform Scripts
 - Script 1
 - Item_0
- Transform Constants
- Transform Functions
- Input fields
 - line
 - word
- Output fields
 - Please use the 'Re

Java script :

Script 1 [X] Item_0

```
var len_word = word.length;
var u_word = upper(word);
```

new script tab

Examine preview data

Rows of step: calculate len_word and discard short words (1000 rows)

#	line	word	len_word
601	been written by him. But apart from those geological questions, which have	APART	5
602	been written by him. But apart from those geological questions, which have	FROM	4
603	been written by him. But apart from those geological questions, which have	THOSE	5
604	been written by him. But apart from those geological questions, which have	GEOLOGICAL	10
605	been written by him. But apart from those geological questions, which have	QUESTIONS,	10
606	been written by him. But apart from those geological questions, which have	WHICH	5
607	been written by him. But apart from those geological questions, which have	HAVE	4
608	an important bearing on biological thought and speculation, such as the	IMPORTANT	9
609	an important bearing on biological thought and speculation, such as the	BEARING	7
610	an important bearing on biological thought and speculation, such as the	BIOLOGICAL	10

Close Stop Get more rows

Generate Rows

Step name: ## TEST DATA ##

Limit: 10

Never stop generating rows:

Interval in ms (delay):

Current row time field name:

Previous row time field name:

Fields:

#	Name	Type	Format	Length	Precision	Currency	Decim...	Group	Value	Set empty string?
1	line	String							test value test value	N
2	word	String					.	,	Pentaho	N

Buttons: Help, OK, Preview, Cancel

Examine preview data

Rows of step: calculate len_word and discard short words (10 rows)

#	line	word	len_word
1	test value test value	PENTAHO	7
2	test value test value	PENTAHO	7
3	test value test value	PENTAHO	7
4	test value test value	PENTAHO	7
5	test value test value	PENTAHO	7
6	test value test value	PENTAHO	7
7	test value test value	PENTAHO	7
8	test value test value	PENTAHO	7
9	test value test value	PENTAHO	7
10	test value test value	PENTAHO	7

Buttons: Close, Show Log

#	text	prop_code
9	Kitchen	MCX-001
10	Basement	MCX-001
11	Bathroom on main floor	MCX-001
12	2 car garage	MCX-001
13	Attached parking	MCX-001
14	More Features: Eat-In Kitchen Area, Kitchen Pantry, Deck, Fence	MCX-001
15	Property Code: MCX-002	MCX-002
16	5 bedrooms	MCX-002
17	5 baths	MCX-002
18	Style: Colonial	MCX-002

User Defined Java Class

Step name: User Defined Java Class

Classes and code fragments: Class code

- Classes
 - Code Snippets
 - Common use
 - Step status
 - Step logging
 - Step/Row listeners
 - Row manipulation
 - Uncommon use
 - Input fields
 - Info fields
 - Output fields

Processor

Line #: 0

Fields Parameters Info steps Target steps

Fields Clear the result fields?

#	Fieldname	Type	Length	Precision
1				

Help OK Cancel Test class

Fields Parameters Info steps Target steps

Fields Clear the result fields?

#	Fieldname	Type	Length	Precision
1	len_word	Integer		

User Defined Java Class

Step name: calculate len_word

Classes and code fragments: Class code

- Classes
 - Code Snippets
 - Input fields
 - Info fields
 - Output fields
 - line
 - word
 - len_word

```

Processor
public boolean processRow(StepMetaInterface smi, StepDataInterface sdi) throws KettleExcept
{
    Object[] r = getRow();
    if (r == null) {
        setOutputDone();
        return false;
    }

    if (first)
    {
        first = false;
    }
}
  
```

Line #: 0

Fields Parameters Info steps Target steps

Fields Clear the result fields?

#	Fieldname	Type	Length	Precision
---	-----------	------	--------	-----------

Help OK Cancel Test class

Examine preview data

Rows of step: calculate len_word and discard short words (1000 rows)

#	line	word	len_word
2..	least attention, up to the present time is that which treats of the geology	time	4
2..	least attention, up to the present time is that which treats of the geology	is	2
2..	least attention, up to the present time is that which treats of the geology	that	4
2..	least attention, up to the present time is that which treats of the geology	which	5
2..	least attention, up to the present time is that which treats of the geology	treats	6
2..	least attention, up to the present time is that which treats of the geology	of	2
2..	least attention, up to the present time is that which treats of the geology	the	3
2..	least attention, up to the present time is that which treats of the geology	geology	7
2..	of South America. The actual writing of this book appears to have occu...	of	2
2..	of South America. The actual writing of this book appears to have occu...	South	5

Close Stop Get more rows

Examine preview data

Rows of step: calculate len_word and discard short words (1000 rows)

#	line	word	len_word
1..	of South America. The actual writing of this book appears to have occupied	WRITING	7
1..	of South America. The actual writing of this book appears to have occupied	THIS	4
1..	of South America. The actual writing of this book appears to have occupied	BOOK	4
2..	of South America. The actual writing of this book appears to have occupied	APPEARS	7
2..	of South America. The actual writing of this book appears to have occupied	HAVE	4
2..	of South America. The actual writing of this book appears to have occupied	OCCUPIED	8
2..	Darwin a shorter period than either of the other volumes of the series; his	DARWIN	6
2..	Darwin a shorter period than either of the other volumes of the series; his	SHORTER	7
2..	Darwin a shorter period than either of the other volumes of the series; his	PERIOD	6
2..	Darwin a shorter period than either of the other volumes of the series; his	THAN	4

Close Stop Get more rows

Generate Rows

Step name:

Limit:

Never stop generating rows:

Interval in ms (delay):

Current row time field name:

Previous row time field name:

Fields:

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Value	Set empty
1	line	String							test value test value	N
2	word	String					.	,	geological	N

Help OK Preview Cancel

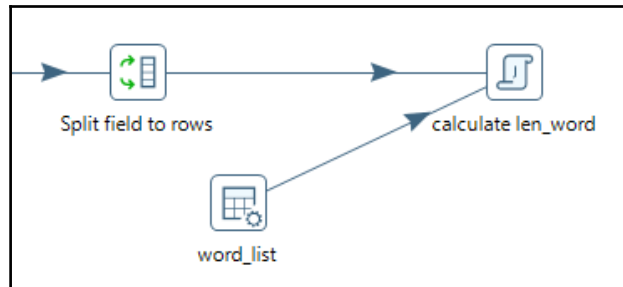
Examine preview data

Rows of step: calculate len_word and discard short words (10 rows)

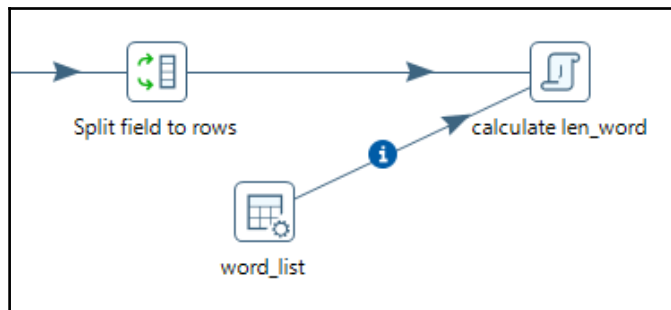
#	line	word	len_word
1	test value test value	GEOLOGICAL	10
2	test value test value	GEOLOGICAL	10
3	test value test value	GEOLOGICAL	10
4	test value test value	GEOLOGICAL	10
5	test value test value	GEOLOGICAL	10
6	test value test value	GEOLOGICAL	10
7	test value test value	GEOLOGICAL	10
8	test value test value	GEOLOGICAL	10
9	test value test value	GEOLOGICAL	10
10	test value test value	GEOLOGICAL	10

Close Show Log

Fields				Parameters				Info steps				Target steps			
Parameters:															
#	Tag	Value	Description												
1	CODE	\${SAMPLE_CODE}													
2	THRESHOLD	3													



Fields				Parameters				Info steps				Target steps			
Info steps:															
#	Tag	Step	Description												
1	MY_LIST	word_list													



Chapter 09: Transforming the Dataset

Sort rows

Step name: Sort rows

Sort directory: %%java.io.tmpdir%%

TMP-file prefix: out

Sort size (rows in memory): 1000000

Free memory threshold (in %):

Compress TMP Files?

Only pass unique rows? (verifies keys only)

Fields:

#	Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted?
1	PRODUCTLINE	Y	N	N	0	N
2	SALES	N	N	N	0	N

Examine preview data

Rows of step: Sort rows (200 rows)

#	ORDERDATE	PRODUCTLINE	PRODUCTCODE	QUANTIT...	SALES
6..	11/17/2004	Classic Cars	S24_1046	23	1307.32
6..	01/12/2005	Classic Cars	S24_1046	21	1267.77
6..	09/21/2003	Classic Cars	S24_2840	30	1209.30
6..	09/08/2004	Classic Cars	S24_2840	39	1172.34
6..	12/02/2003	Classic Cars	S24_2972	21	777.00
6..	11/24/2004	Motorcycles	S10_1678	41	7737.93
6..	11/11/2003	Motorcycles	S24_1578	48	5355.36
6..	01/07/2005	Motorcycles	S50_4713	44	4984.32
6..	10/15/2004	Motorcycles	S10_4698	20	4570.40
7..	09/30/2004	Motorcycles	S32_4485	38	4382.16

Group By

Step name:

Include all rows?

Temporary files directory:

TMP-file prefix:

Add line number, restart in each group

Line number field name:

Always give back a result row

The fields that make up the group:

#	Group field
1	PRODUCTLINE
2	PRODUCTCODE

Aggregates:

#	Name	Subject	Type
1	qty_distinct_dates	ORDERDATE	Number of Distinct Values (N)
2	avg_sales	SALES	Average (Mean)
3	total_sales	SALES	Sum

Select the preview step:

Step name:

Examine preview data

Rows of step: Sort by Product Line and Product Code (200 rows)

#	ORDERDATE	PRODUCTLINE	PRODUCTCODE	QUANTITY...	SALES
6..	11/04/2003	Classic Cars	S24_3856	34	3819.56
6..	05/03/2005	Classic Cars	S24_4048	31	4253.20
6..	05/05/2004	Classic Cars	S24_4048	46	6311.20
6..	06/15/2004	Classic Cars	S700_2824	45	3641.40
6..	11/05/2004	Classic Cars	S700_2824	34	4248.30
6..	01/29/2003	Classic Cars	S700_2824	42	4460.82
6..	11/24/2004	Motorcycles	S10_1678	41	7737.93
6..	10/15/2004	Motorcycles	S10_4698	20	4570.40
6..	01/05/2005	Motorcycles	S10_4698	22	3664.10
6..	01/05/2005	Motorcycles	S12_2823	22	3877.06
7..	07/23/2004	Motorcycles	S18_2625	37	2353.20

Close

Examine preview data

Rows of step: Group by PRODUCTCODE (98 rows)

#	PRODUCTLINE	PRODUCTCODE	qty_distinct_dates	avg_sales	total_sales
30	Classic Cars	S24_3856	2	3363.265	6726.53
31	Classic Cars	S24_4048	2	5282.2	10564.4
32	Classic Cars	S700_2824	3	4116.84	12350.52
33	Motorcycles	S10_1678	1	7737.93	7737.93
34	Motorcycles	S10_4698	2	4117.25	8234.5
35	Motorcycles	S12_2823	1	3877.06	3877.06
36	Motorcycles	S18_2625	2	2024.58	4049.16
37	Motorcycles	S18_3782	4	2115.59	8462.36
38	Motorcycles	S24_1578	2	3827.18	7654.36
39	Motorcycles	S24_2000	1	2096.32	2096.32
40	Motorcycles	S24_2360	1	2632.89	2632.89

Close

Examine preview data

Rows of step: Group everything (1 rows)

#	qty_distinct_dates	avg_sales	total_sales
1	127	3561.94355	712388.71

Close

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors
1	Read Sales Data	0	0	200	201	0	1	0	0
2	Select values	0	200	200	0	0	0	0	0
3	Sort by Product Line and Product Code	0	200	200	0	0	0	0	0
4	Group by PRODUCTCODE	0	200	98	0	0	0	0	0

Examine preview data

Rows of step: Group by PRODUCTCODE (200 rows)

#	ORDERDATE	PRODUCTLINE	PRODUCTCODE	QUANTITYORDERED	SALES	qty_distinct_dates	avg_sales	total_sales
60	11/04/2003	Classic Cars	S24_3856	34	3819.56	2	3363.265	6726.53
61	05/03/2005	Classic Cars	S24_4048	31	4253.20	2	5282.2	10564.4
62	05/05/2004	Classic Cars	S24_4048	46	6311.20	2	5282.2	10564.4
63	06/15/2004	Classic Cars	S700_2824	45	3641.40	3	4116.84	12350.52
64	11/05/2004	Classic Cars	S700_2824	34	4248.30	3	4116.84	12350.52
65	01/29/2003	Classic Cars	S700_2824	42	4460.82	3	4116.84	12350.52
66	11/24/2004	Motorcycles	S10_1678	41	7737.93	1	7737.93	7737.93
67	10/15/2004	Motorcycles	S10_4698	20	4570.40	2	4117.25	8234.5
68	01/05/2005	Motorcycles	S10_4698	22	3664.10	2	4117.25	8234.5
69	01/05/2005	Motorcycles	S12_2823	22	3877.06	1	3877.06	3877.06
70	07/23/2004	Motorcycles	S18_2625	37	2353.20	2	2024.58	4049.16

Close

#	film	Year	Genres	Director	Actors
1	Persepolis	2007	Animation Comedy Drama History	Marjane Satrapi	Chiara Mastroianni, Catherine Deneuve, Danielle D
2	Trois couleurs - Rouge	1994	Drama	Krzysztof Kiesowski	Irène Jacob, Jean-Louis Trintignant, Frédérique Fed
3	Les Misérables	1933	Drama History	Raymond Bernard	
4	Au revoir, les enfants	1987	Drama	Louis Malle	
5	La France	2007	Drama Musical Romance War	Serge Bozon	Sylvie Testud, Pascal Greggory, Guillaume Verdier
6	L'Atalante	1934	Drama Romance	n/a	Michel Simon, Dita Parlo, Jean Dasté
7	La môme	2007	Biography Drama Music	Olivier Dahan	Marion Cotillard, Sylvie Testud, Pascal Greggory, Er
8	MR 73	2008	Crime Drama	n/a	Daniel Auteuil, Olivia Bonamy, Catherine Marchal,
9	Manon Des Sources		Drama Romance	Claude Berri	Yves Montand, Daniel Auteuil, Emmanuelle Béart
10	Un Coeur en Hiver	1992	Drama Romance Music	Claude Sautet	Daniel Auteuil, Emmanuelle Béart, André Dussollier

 Examine preview data - □ ×

Rows of step: Filter null rows (371 rows)

#	feature	description	film
1	Year	2007	Persepolis
2	Genre	Animation Comedy Drama History	Persepolis
3	Directed by	Vincent Paronnaud	Persepolis
4	Directed by	Marjane Satrapi	Persepolis
5	Script	Vincent Paronnaud, Marjane Satrapi	Persepolis
6	Music	Olivier Bernet	Persepolis
7	Cast	Chiara Mastroianni, Catherine Deneuve, Danielle Darrieux	Persepolis
8	Year	1994	Trois couleurs - Rouge
9	Genre	Drama	Trois couleurs - Rouge
1..	Directed by	Krzysztof Kiesowski	Trois couleurs - Rouge
1..	Cast	Irène Jacob, Jean-Louis Trintignant, Frédérique Feder, Jean-Pierre Lorit, Samue...	Trois couleurs - Rouge
1..	Year	1933	Les Misérables
1..	Genre	Drama History	Les Misérables

Close

Denormaliser

Step name: 1 row x film

The key field: feature

The fields that make up the grouping:

#	Group field
1	film

Get Fields

Target fields:

#	Target fieldname	Value fieldname	Key value	Type
1	Year	description	Year	String
2	Genres	description	Genre	String
3	Director	description	Directed by	String
4	Actors	description	Cast	String

Get lookup fields

Help OK Cancel

Examine preview data

Rows of step: 1 row x film (97 rows)

#	film	Year	Genres	Director	Actors
1	Persepolis	2007	Animation Comedy Drama History	Marjane Satrapi	Chiara Mastroianni, Catherine Deneuve, [
2	Trois couleurs - Rouge	1994	Drama	Krzysztof Kieslowski	Irène Jacob, Jean-Louis Trintignant, Frédé
3	Les Misérables	1933	Drama History	Raymond Bernard	<null>
4	Au revoir, les enfants	1987	Drama	Louis Malle	<null>
5	La France	2007	Drama Musical Romance War	Serge Bozon	Sylvie Testud, Pascal Greggory, Guillaume
6	L'Atalante	1934	Drama Romance	<null>	Michel Simon, Dita Parlo, Jean Dasté
7	La môme	2007	Biography Drama Music	Olivier Dahan	Marion Cotillard, Sylvie Testud, Pascal Gr
8	MR 73	2008	Crime Drama	<null>	Daniel Auteuil, Olivia Bonamy, Catherine
9	Manon Des Sources	<null>	Drama Romance	Claude Berri	Yves Montand, Daniel Auteuil, Emmanue
1..	Un Coeur en Hiver	1992	Drama Romance Music	Claude Sautet	Daniel Auteuil, Emmanuelle Béart, André

Close

Examine preview data

Rows of step: Filter null rows (371 rows)

#	feature	description	film
32	Directed by	Claude Berri	Manon Des Sources
33	Produced by	Pierre Grunstein	Manon Des Sources
34	Genre	Drama Romance	Manon Des Sources
35	Cast	Yves Montand, Daniel Auteuil, Emmanuelle Béart	Manon Des Sources

Close

Target fields:

#	Target fieldname	Value fieldname	Key value	Type
1	Year	description	Year	String
2	Genres	description	Genre	String
3	Director	description	Directed by	String
4	Actors	description	Cast	String

Examine preview data

Rows of step: 1 row x film (1 rows)

#	film	Year	Genres	Director	Actors
1	Manon Des Sources	<null>		Claude Berri	

Close

Examine preview data

Rows of step: 1 row x film (1 rows)

#	film	Year	Genres	Director	Actors
1	Manon Des Sources	<null>	Drama Romance	Claude Berri	Yves Montand, Daniel Auteuil, Emmanuelle Béart

Close

Target fields:

#	Target fieldname	Value fieldname	Key value	Type	F...	L...	P...	C..	D...	G...	N...	Aggregation
1	Year	description	Year	String								-
2	Genres	description	Genre	String								-
3	Director	description	Directed by	String								Concatenate strings separated by ,
4	qty_directors	description	Directed by	String								Number of Values
5	Actors	description	Cast	String								-

Examine preview data

Rows of step: 1 row x film (97 rows)

#	film	Year	Genres	Director	qty_directors	Actors
1	Persepolis	2007	Animation Comedy Drama ...	Vincent Paronnaud, Marjane Satrapi	2	Chiara Mastroianni, Catherine Deneuve, Da
2	Trois couleurs - Rouge	1994	Drama	Krzysztof Kieslowski	1	Irène Jacob, Jean-Louis Trintignant, Frédéri
3	Les Misérables	1933	Drama History	Raymond Bernard	1	<null>
4	Au revoir, les enfants	1987	Drama	Louis Malle	1	<null>
5	La France	2007	Drama Musical Romance ...	Serge Bozon	1	Sylvie Testud, Pascal Greggory, Guillaume \
6	L'Atalante	1934	Drama Romance	<null>	0	Michel Simon, Dita Parlo, Jean Dasté

Close

Examine preview data

Rows of step: Sales (3 rows)

#	YEAR	Classic Cars (qty)	Classic Cars (\$)	Planes (qty)	Planes (\$)	Ships (qty)	Ships (\$)	Trains (qty)	Trains (\$)	Vintage Cars (qty)	Vintage Cars (\$)
1	2015	2320.0	2645939.0	703.0	666778.0	1122.0	994561.0	607.0	510842.0	156.0	12611.0
2	2016	1500.0	17107364.0	547.0	5188158.0	955.0	8465292.0	598.0	5032677.0	566.0	4575529.0
3	2017	854.0	9739793.0	199.0	1887465.0	998.0	8846452.0	1005.0	8457928.0	144.0	1164092.0

Close Show Log

#	YEAR	product_line	qty	amount
1	2015	Cars	2320,0	264593,9
2	2015	Planes	703,0	66677,8
3	2015	Ships	1122,0	99456,1
4	2015	Trains	607,0	51084,2
5	2016	Cars	1500,0	171073,64
6	2016	Planes	547,0	51881,58
7	2016	Ships	955,0	84652,92
8	2016	Trains	598,0	50326,77
9	2017	Cars	854,0	97397,93
10	2017	Planes	199,0	18874,65
11	2017	Ships	998,0	88464,52
12	2017	Trains	1005,0	84579,28

Row Normaliser
— □ ×

Step name

Type field

Fields

#	Fieldname	Type	new field
1			

Help
OK
Cancel
Get Fields

Row Normaliser

Step name: Row Normaliser

Type field: product_line

Fields

#	Fieldname	Type	new field
1	Classic Cars (qty)	Cars	qty
2	Classic Cars (\$)	Cars	amount
3	Planes (qty)	Planes	qty
4	Planes (\$)	Planes	amount
5	Ships (qty)	Ships	qty
6	Ships (\$)	Ships	amount
7	Trains (qty)	Trains	qty
8	Trains (\$)	Trains	amount

Help OK Cancel Get Fields

Examine preview data

Rows of step: Select values (12 rows)

#	YEAR	product_line	qty	amount
1	2015	Cars	2320.0	2645939.0
2	2015	Planes	703.0	666778.0
3	2015	Ships	1122.0	994561.0
4	2015	Trains	607.0	510842.0
5	2016	Cars	1500.0	17107364.0
6	2016	Planes	547.0	5188158.0
7	2016	Ships	955.0	8465292.0
8	2016	Trains	598.0	5032677.0

Close

Examine preview data

Rows of step: Select values (200 rows)

#	PRODUCTLINE	MONTH	SALES
1	Planes	2004-02	1716.26
2	Classic Cars	2004-11	5679.36
3	Planes	2003-06	2213.38
4	Classic Cars	2004-07	6203.40
5	Classic Cars	2004-10	1383.03
6	Planes	2004-12	3302.40
7	Trucks and Buses	2005-01	8470.14
8	Trucks and Buses	2004-03	5704.32
9	Trucks and Buses	2004-07	4094.72
1..	Motorcycles	2003-08	3251.34

Close

The fields that make up the group:

#	Group field
1	PRODUCTLINE
2	MONTH

Get Fields

Aggregates :

#	Name	Subject	Type	Value
1	SALES	SALES	Sum	

Get lookup fields

Examine preview data

Rows of step: Group by (107 rows)

#	PRODUCTLINE	MONTH	SALES
1	Classic Cars	2003-01	4460.82
2	Classic Cars	2003-03	3485.82
3	Classic Cars	2003-05	10947.64
4	Classic Cars	2003-07	4380.2
5	Classic Cars	2003-09	9509.91
6	Classic Cars	2003-10	12214.02
7	Classic Cars	2003-11	30383.93
8	Classic Cars	2003-12	4946.88
9	Classic Cars	2004-01	10114.56
1..	Classic Cars	2004-02	12095.68

Close

Analytic Query

Step name:

The fields that make up the group:

#	Group field
1	PRODUCTLINE

Get Fields

Analytic Functions :

#	New Field Name	Subject	Type	N
1	MONTH_PREV	MONTH	LAG "N" rows BACKWARD in get Subject	1
2	SALES_PREV	SALES	LAG "N" rows BACKWARD in get Subject	1

Get lookup fields

Help OK Cancel

Examine preview data

Rows of step: Analytic Query (107 rows)

#	PRODUCTLINE	MONTH	SALES	MONTH_PREV	SALES_PREV
20	Classic Cars	2004-12	3042.41	2004-11	17190.23
21	Classic Cars	2005-01	11007.31	2004-12	3042.41
22	Classic Cars	2005-02	14424.2	2005-01	11007.31
23	Classic Cars	2005-03	11013.37	2005-02	14424.2
24	Classic Cars	2005-04	8374.69	2005-03	11013.37
25	Classic Cars	2005-05	15336.58	2005-04	8374.69
26	Motorcycles	2003-03	1173.15	<null>	<null>
27	Motorcycles	2003-08	3251.34	2003-03	1173.15
28	Motorcycles	2003-10	4483.36	2003-08	3251.34
29	Motorcycles	2003-11	5355.36	2003-10	4483.36

Close

Calculator

Step name: Calculator

Fields:

#	New field	Calculation	Field A	Field B	F..	Value type	Length	Precision	Remove	Conversion mask
1	dif	A - B	SALES	SALES_PREV		Number			Y	
2	perc	100 * A / B	dif	SALES_PREV		Number			N	0.00

Help OK Cancel

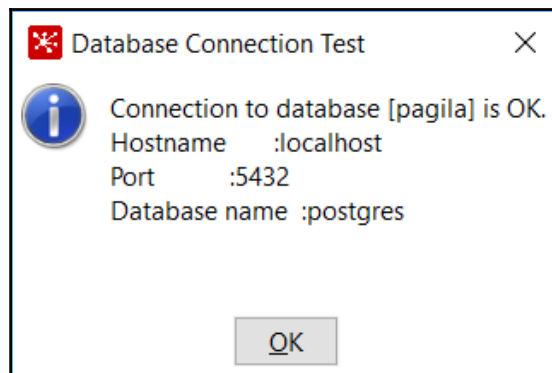
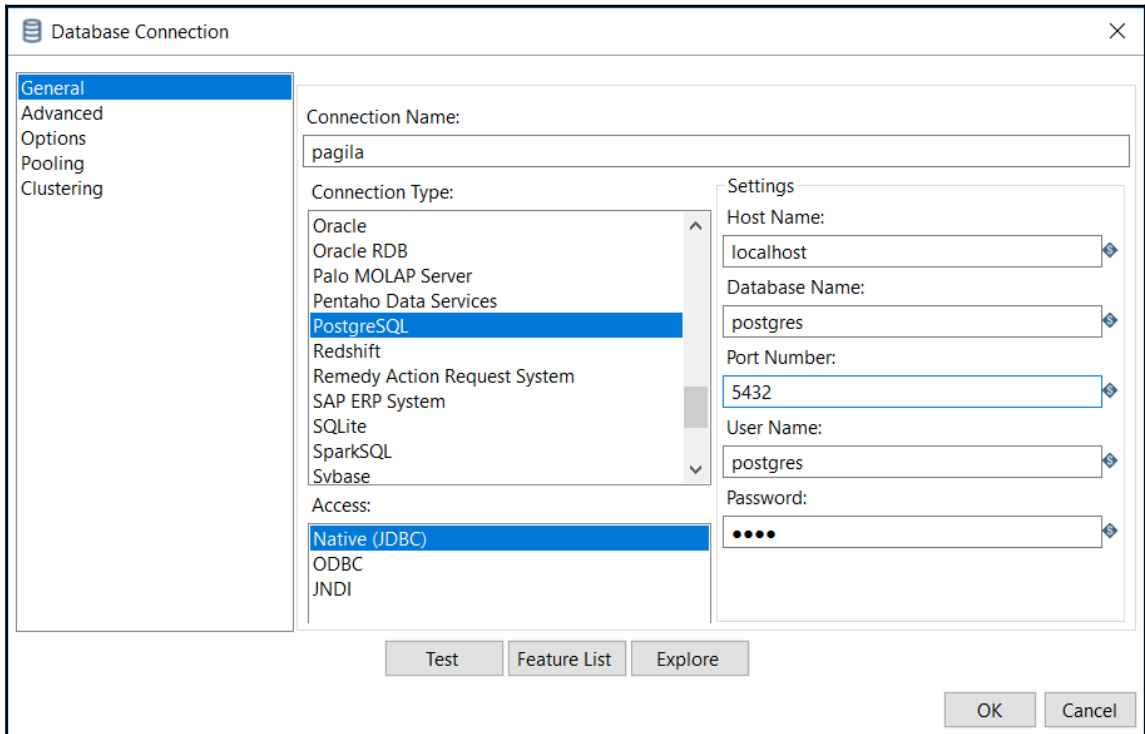
Examine preview data

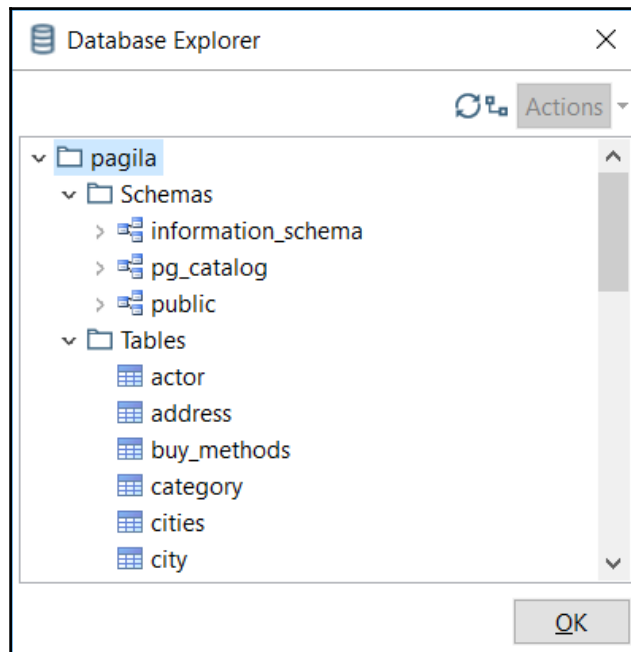
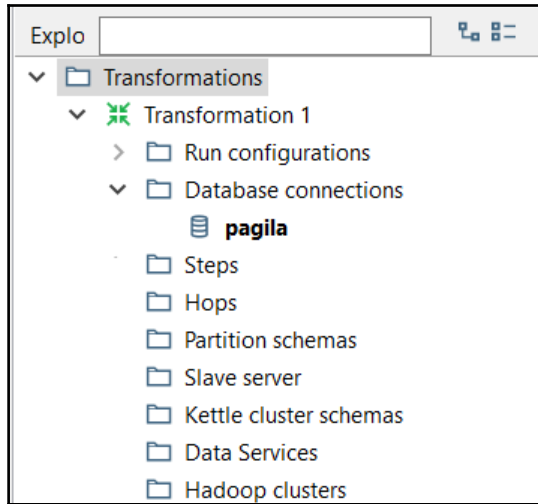
Rows of step: Calculator (107 rows)

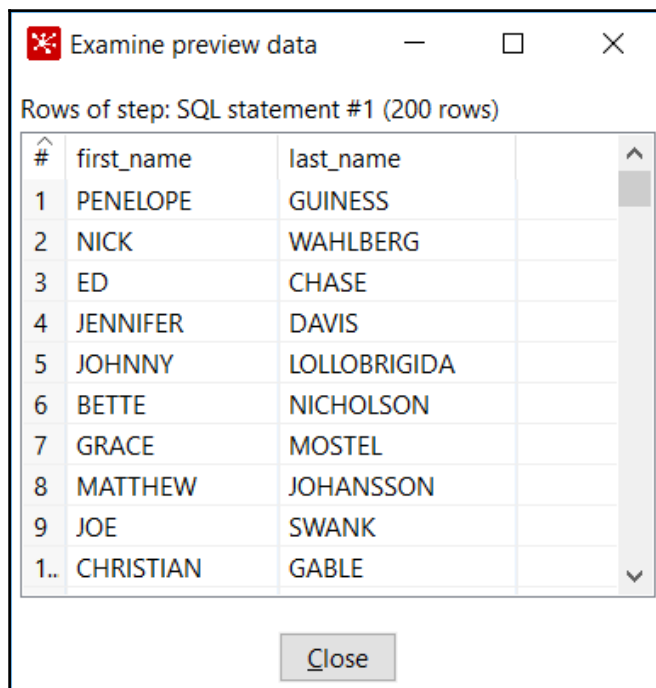
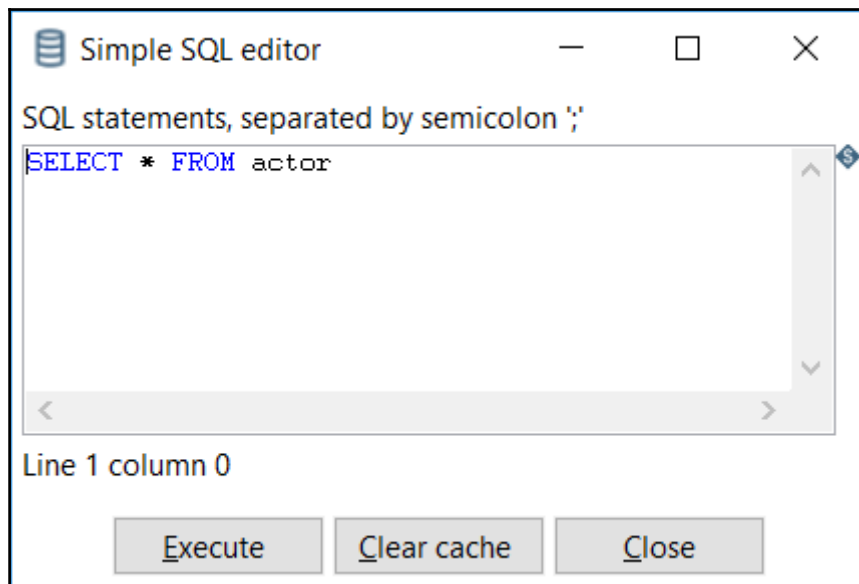
#	PRODUCTLINE	MONTH	SALES	MONTH_PREV	SALES_PREV	perc
20	Classic Cars	2004-12	3042.41	2004-11	17190.23	-82.30
21	Classic Cars	2005-01	11007.31	2004-12	3042.41	261.80
22	Classic Cars	2005-02	14424.2	2005-01	11007.31	31.04
23	Classic Cars	2005-03	11013.37	2005-02	14424.2	-23.65
24	Classic Cars	2005-04	8374.69	2005-03	11013.37	-23.96
25	Classic Cars	2005-05	15336.58	2005-04	8374.69	83.13
26	Motorcycles	2003-03	1173.15	<null>	<null>	<null>
27	Motorcycles	2003-08	3251.34	2003-03	1173.15	177.15
28	Motorcycles	2003-10	4483.36	2003-08	3251.34	37.89
29	Motorcycles	2003-11	5355.36	2003-10	4483.36	19.45

Close

Chapter 10: Loading Data Marts with PDI





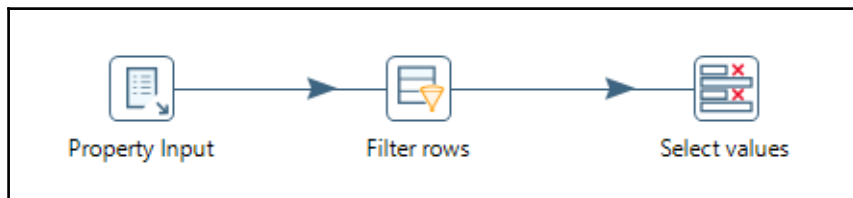


Examine preview data

Rows of step: sample cities (600 rows)

#	city_id	city	country_id	last_update
1	1	A Corua (La Corua)	87	2017/02/15 09:45:25.000000000
2	2	Abha	82	2017/02/15 09:45:25.000000000
3	3	Abu Dhabi	101	2017/02/15 09:45:25.000000000
4	4	Acua	60	2017/02/15 09:45:25.000000000
5	5	Adana	97	2017/02/15 09:45:25.000000000
6	6	Addis Abeba	31	2017/02/15 09:45:25.000000000
7	7	Aden	107	2017/02/15 09:45:25.000000000
8	8	Adoni	44	2017/02/15 09:45:25.000000000
9	9	Ahmadnagar	44	2017/02/15 09:45:25.000000000
1..	10	Akishima	50	2017/02/15 09:45:25.000000000

Close Show Log



Examine preview data

Rows of step: Select value (1 rows)

#	Value
1	Comedy

Close

Examine preview data

Rows of step: films (58 rows)

#	title
1	AIRPLANE SIERRA
2	ANTHEM LUKE
3	BRINGING HYSTERICAL
4	CAPER MOTIONS
5	CAT CONEHEADS
6	CLOSER BANG
7	CONNECTION MICROCOSMOS
8	CONTROL ANTHEM
9	CRAZY HOME
1..	DADDY PITTSBURGH

Close

Parameters Variables

Variable	Value
CATEGORY	Comedy
Internal.Entry.Current.Directory	file:///C:/Users/user/Dropbox/projects/Learning P...
Internal.Job.Filename.Directory	Parent Job File Directory
Internal.Job.Filename.Name	Parent Job Filename
Internal.Job.Name	Parent Job Name
Internal.Job.Repository.Directory	Parent Job Repository Directory

Table output

Step name

Connection

Target schema

Target table

Commit size

Truncate table

Ignore insert errors

Specify database fields

Main options Database fields

Fields to insert:

#	Table field	Stream field
1		

Insert / Update

Step name: Insert / Update

Connection: pagila [Edit... New... Wizard...]

Target schema: [Browse...]

Target table: lookup table [Browse...]

Commit size: 100

Don't perform any updates:

The key(s) to look up the value(s):

#	Table field	Comparator	Stream field1	Stream field2
1				

[Get fields]

Update fields:

#	Table field	Stream field	Update
1			

[Get update fields]

[Edit mapping]

[Help] [OK] [Cancel] [SQL]

Delete

Step name: Delete

Connection: pagila [Edit... New... Wizard...]

Target schema: [Browse...]

Target table: lookup table [Browse...]

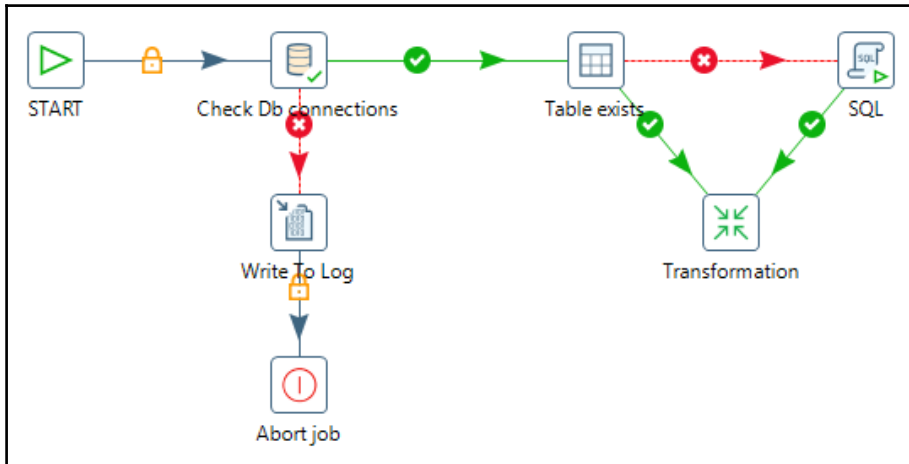
Commit size: 100

The key(s) to look up the value(s):

#	Table field	Comparator	Stream field1	Stream field2
1				

[Get fields]

[Help] [OK] [Cancel]



Database Value Lookup

Step name: Database lookup

Connection: pagila [Edit... New... Wizard...]

Lookup schema: [Browse...]

Lookup table: lookup table [Browse...]

Enable cache?

Cache size in rows (0=cache): 0

Load all data from table

The key(s) to look up the value(s):

#	Table field	Comparator	Field1	Field2
1				

Values to return from the lookup table :

#	Field	New name	Default	Type
1				

Do not pass the row if the lookup fails

Fail on multiple results?

Order by: []

[? Help] [OK] [Cancel] [Get Fields] [Get lookup fields]

The key(s) to look up the value(s):

#	Table field	Comparator	Field1	Field2
1	first_name	=	name	

Values to return from the lookup table :

#	Field	New name	Default	Type
1	first_name		Not found	String
2	last_name		Not found	String

Examine preview data

Rows of step: look for Actors (3 rows)

#	name	first_name	last_name
1	DANY	Not found	Not found
2	MORGAN	MORGAN	WILLIAMS
3	RALPH	RALPH	CRUZ

Close

Add constant rows

Step name: sample words

#	word1	word2
1	hunter	shark
2	robot	crocodile
3	dog	scientist

Help OK Preview Cancel

The parameters to use:

#	Parameter fieldname	Parameter Type
1	word1	String
2	word2	String

Examine preview data

Rows of step: look for films (9 rows)

#	word1	word2	title	description
1	hunter	shark	ATLANTIS CAUSE	A Thrilling Yarn of a Feminist And a Hunter who mus
2	hunter	shark	CELEBRITY HORN	A Amazing Documentary of a Secret Agent And a A
3	hunter	shark	FALCON VOLUME	A Fateful Saga of a Sumo Wrestler And a Hunter wh
4	robot	crocodile	BONNIE HOLOCAUST	A Fast-Paced Story of a Crocodile And a Robot who
5	robot	crocodile	CAMPUS REMEMBER	A Astounding Drama of a Crocodile And a Mad Cov
6	robot	crocodile	CASPER DRAGONFLY	A Intrepid Documentary of a Boat And a Crocodile
7	dog	scientist	DANGEROUS UPTOWN	A Unbelievable Story of a Mad Scientist And a Wo

Close

Examine preview data

Rows of step: STATEMENT (3 rows)

#	word1	word2	STATEMENT
1	hunter	shark	SELECT title, description FROM public.film WHERE upper(description) like '%HUNTER%' AND upper(description) like '%SHARK%' LIMIT 3
2	robot	crocodile	SELECT title, description FROM public.film WHERE upper(description) like '%ROBOT%' AND upper(description) like '%CROCODILE%' LIMIT 3
3	dog	scientist	SELECT title, description FROM public.film WHERE upper(description) like '%DOG%' AND upper(description) like '%SCIENTIST%' LIMIT 3

Close

Dynamic SQL join

Step name: look for films

Connection: pagila [Edit... New... Wizard...]

SQL field name: STATEMENT

Number of rows to return: 0

Outer join?

Replace variables

Query only on parameters change

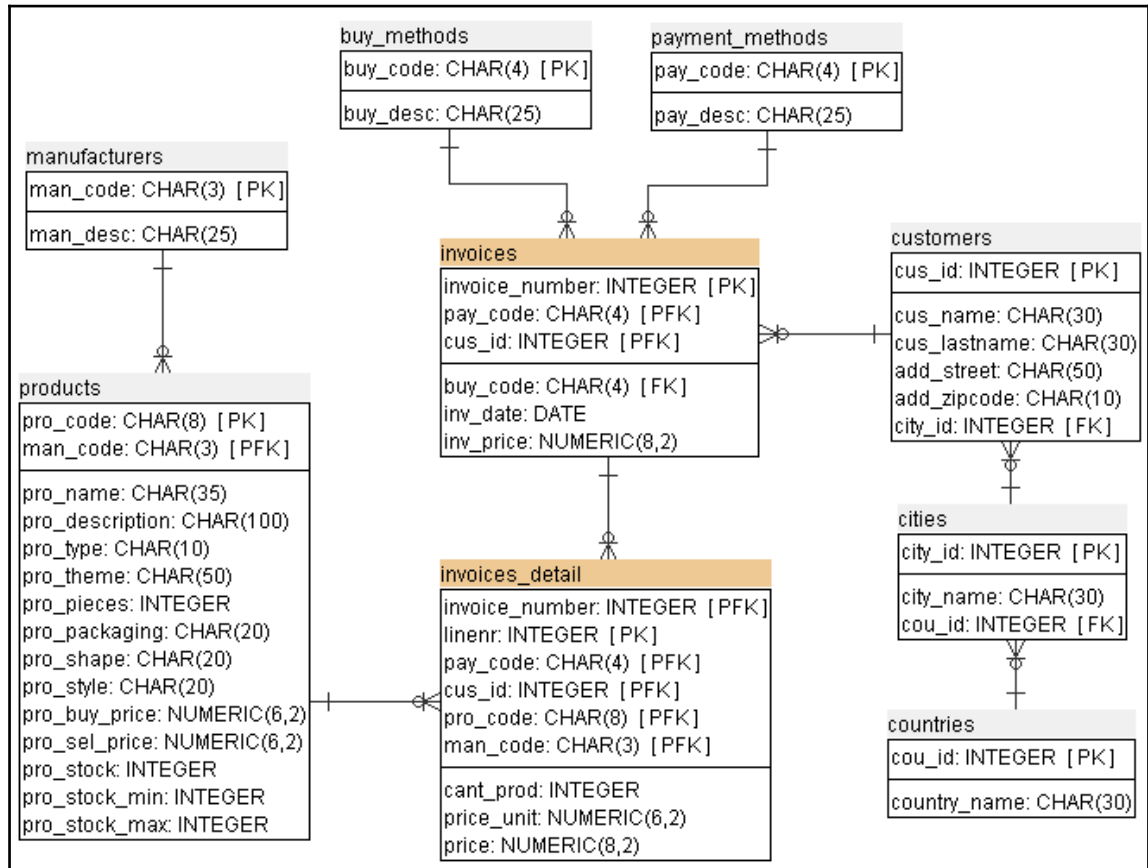
Template SQL (to retrieve Meta data)

```
SELECT 'sample title' as TITLE, 'sample desc.' as DESCRIPTION;
```

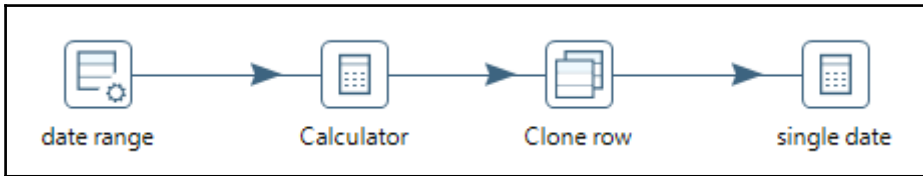
Line: 1 Column: 0

[Help] [OK] [Cancel]

Chapter 11: Loading Data Marts with PDI



surrogate key			business key
id	name	theme	id_js_prod
488	Fire!	Comedy/Cartoon	FAL428G
489	Tour de France	Comedy/Cartoon	FAL45Q7
490	Princesses and Horses - Cinderella	Horses	FAL46IO
491	Christmas Carriage	Christmas	FAL4ATQ
492	At The Beach	Comedy/Cartoon	FAL4DJ4
493	Christmas Spirits	Christmas	FAL4FGA
494	Pick a Friend	Dogs	FAL4H54
495	German Beer Festival	Comedy/Cartoon	FAL4HMU
496	Mickey and Friends	Disney	FAL4LC9
497	Hidden Lake Chateau	Castles	FAL4MME
498	The Airshow	Comedy/Cartoon	FAL4MVE



Calculator										
Step name: date attributes										
#	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove	Conversion mask
1	a_single_date	Date A + B Days	start_date	delta		Date			N	yyyyMMdd
2	year	Year of date A	a_single_date			Integer			N	
3	month	Month of date A	a_single_date			Integer			N	
4	day	Day of month of date A	a_single_date			Integer			N	
5	week_day	Day of week of date A	a_single_date			Integer			N	

Examine preview data

Rows of step: date attributes (1000 rows)

#	start_date	end_date	diff_dates	delta	a_single_date	year	month	day	week_day
2	2000-01-01	2020-12-31	7670	1	20000102	2000	1	2	1
3	2000-01-01	2020-12-31	7670	2	20000103	2000	1	3	2
4	2000-01-01	2020-12-31	7670	3	20000104	2000	1	4	3
5	2000-01-01	2020-12-31	7670	4	20000105	2000	1	5	4
6	2000-01-01	2020-12-31	7670	5	20000106	2000	1	6	5
7	2000-01-01	2020-12-31	7670	6	20000107	2000	1	7	6
8	2000-01-01	2020-12-31	7670	7	20000108	2000	1	8	7
9	2000-01-01	2020-12-31	7670	8	20000109	2000	1	9	1
1..	2000-01-01	2020-12-31	7670	9	20000110	2000	1	10	2

Close Stop Get more rows

Main options Database fields

Fields to insert:

#	Table field	Stream field
1	dateid	a_single_date
2	year	year
3	month	month
4	day	day
5	week_day	week_day

Get fields

Enter field mapping

Key fields (to look up row in table):

#	Dimension field	Field in stream
1	id_js	city_id

Examine preview data

Rows of step: lk_regions (322 rows)

#	id	city	country	region	id_js	lastupdate
2..	279	London	United Kingdom	N/A	181001	2017/11/20
2..	280	Austin	United States	N/A	182001	2017/11/20
2..	281	Baltimore	United States	N/A	182002	2017/11/20
2..	282	Boston	United States	N/A	182003	2017/11/20
2..	283	Charlotte	United States	N/A	182004	2017/11/20
2..	284	Chicago	United States	N/A	182005	2017/11/20
2..	285	Columbus	United States	N/A	182006	2017/11/20
2..	286	Dallas	United States	N/A	182007	2017/11/20

Close

Keys Fields

Key fields (to look up row in dimension):

#	Dimension field	Field in stream
1	id_js_prod	pro_code
2	id_js_man	man_code

Technical key field New name

Creation of technical key

Use table maximum + 1

Use sequence

Use auto increment field

Version field

Stream Datefield

Date range start field Min. year

Use an alternative start date?

Table date range end Max. year

Keys Fields

Lookup/Update fields

#	Dimension field	Stream field to compare with	Type of dimension update
1	name	pro_name	Insert
2	theme	pro_theme	Insert
3	current		Last version (without stream field as source)

< >

Examine preview data

Rows of step: lk_puzzles (1000 rows)

#	id	name	theme	id_js_prod	id_js_man	start_date	end_date	version	current
3..	365	Candelabra	Fantasy	ED13_11	EDU	1900/01/01	2199/12/31	1	Y
3..	366	Capri Cove	Beaches	ED13_12	EDU	1900/01/01	2199/12/31	1	Y
3..	367	Cars Double Pack	Childrens	ED13_14	EDU	1900/01/01	2199/12/31	1	Y
3..	368	Cars	Disney	ED13_15	EDU	1900/01/01	2199/12/31	1	Y
3..	369	Celebration	Sealife	ED13_17	EDU	1900/01/01	2199/12/31	1	Y
3..	370	Clown School	Sealife	ED13_18	EDU	1900/01/01	2199/12/31	1	Y
3..	371	Collage by E.B.Leighton	Miscellany	ED13_19	EDU	1900/01/01	2199/12/31	1	Y
3..	372	Amboseli Park	Wild	ED13_2	EDU	1900/01/01	2199/12/31	1	Y
3..	373	Cradle of Life	Big cats	ED13_20	EDU	1900/01/01	2199/12/31	1	Y
3..	374	Da Vinci's World	Da Vinci	ED13_21	EDU	1900/01/01	2199/12/31	1	Y

Close

Examine preview data

Rows of step: SQL statement #1 (59 rows)

#	id	name	theme	id_js_prod	id_js_man	start_date	end_date	version	current
1	658	A Gondola Ride in Venice	Famous Landmarks	JUMBO101	JUM	1900/01/01	2017/10/01	1	N
2	1029	A Gondola Ride in Venice	Caribbean	JUMBO101	JUM	2017/10/01	2199/12/31	2	Y
3	659	Afternoon Tea	Dogs	JUMBO102	JUM	1900/01/01	2199/12/31	1	Y
4	660	Ascending and Descending	Escher	JUMBO103	JUM	1900/01/01	2199/12/31	1	Y
5	661	Bond of Union	Artists	JUMBO104	JUM	1900/01/01	2017/10/01	1	N
6	1030	Bond of Union	Fantasy	JUMBO104	JUM	2017/10/01	2199/12/31	2	Y
7	662	Cars	Disney	JUMBO105	JUM	1900/01/01	2199/12/31	1	Y
8	663	Cat in a Basket	Cats	JUMBO106	JUM	1900/01/01	2199/12/31	1	Y
9	664	Cinderellas Grand Arrival	Castles	JUMBO107	JUM	1900/01/01	2017/10/01	1	N
1..	1031	Cinderellas Grand Arrival	Disney	JUMBO107	JUM	2017/10/01	2199/12/31	2	Y

Close

Database Value Lookup

Step name: DB LK Region

Connection: js_dw [Edit... New... Wizard...]

Lookup schema: [Browse...]

Lookup table: lk_regions [Browse...]

Enable cache?

Cache size in rows (0=cache): 0

Load all data from table

The key(s) to look up the value(s):

#	Table field	Comparator	Field1	Field2
1	id_js	=	city_id	

Values to return from the lookup table :

#	Field	New name	Default	Type
1	id	id_region		Integer

Do not pass the row if the lookup fails

Fail on multiple results?

Order by: []

[Help] [OK] [Cancel] [Get Fields] [Get lookup fields]

Keys Fields

Key fields (to look up row in dimension):

#	Dimension field	Field in stream
1	id_js_prod	product_id
2	id_js_man	manuf_id

Technical key field: New name:

Creation of technical key:

Use table maximum + 1

Use sequence

Use auto increment field

Version field:

Stream Datefield:

Date range start field: Min. year:

Use an alternative start date?

Table date range end: Max. year:

id	name	theme	id_js_prod	id_js_man	start_date	end_date	version	current
658	A Gondola Ride in Venice	Famous Landmarks	JUMBO101	JUM	1900/01/01 ...	2017/10/01 ...	1	N
1029	A Gondola Ride in Venice	Caribbean	JUMBO101	JUM	2017/10/01 ...	2199/12/31 ...	2	Y

Combination Lookup / Update

Step name: Combination LK/Update buy & payment method

Connection: js_dw [Edit...] [New...] [Wizard...]

Target schema: [Browse...]

Target table: lk_junk_sales [Browse...]

Commit size: 100 Cache size: 9999

Pre-load the cache?

Key fields (to look up row in table):

#	Dimension field	Field in stream
1	buy_method	buy_desc
2	payment_method	pay_desc

Technical key field: id

Creation of technical key:

- Use table maximum + 1
- Use sequence []
- Use auto increment field

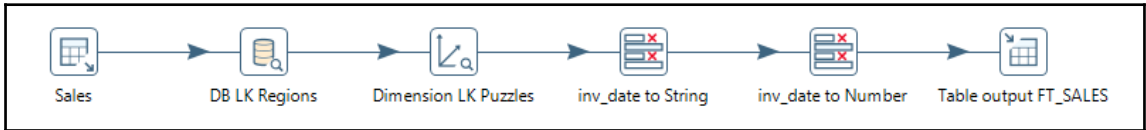
Remove lookup fields?

Use hashcode?

Hashcode field in table: []

Date of last update field (optional): []

[Help] [OK] [Cancel] [Get Fields] [SQL]



Chapter 12: Creating Portable and Reusable Transformations

Job entry name: Set variables

Properties file

Name of properties file

Variable scope: Valid in the Java Virtual Machine

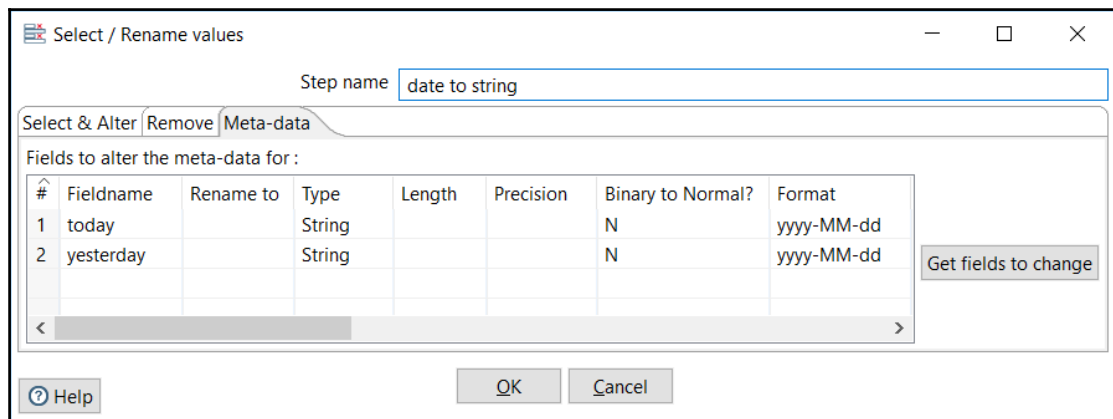
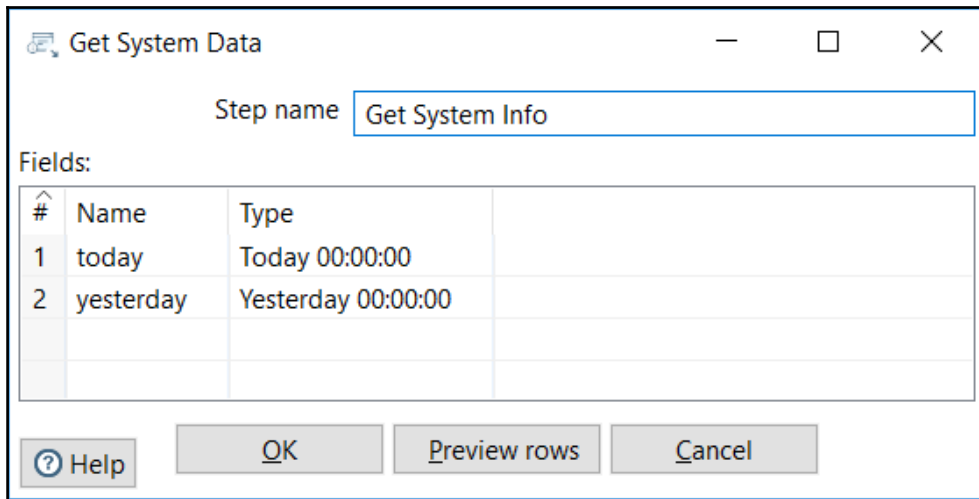
Settings

Variable substitution?

Variables :

#	Variable name	Value	Variable scope type
1	SUCCESS	1	Valid in the Java Virtual Machine
2	WARNING	2	Valid in the current job
3	ERROR	9	Valid in the parent job
4	PATH	/home/pentaho	Valid in the root job

Help OK Cancel



Examine preview data

Rows of step: Get System Info (1 rows)

#	today	yesterday	
1	2017/11/19 00:00:00.000	2017/11/18 00:00:00.000	

Close Show Log

Set Environment Variables

Step name : Set Variables


Apply formatting

Field values:

#	Field name	Variable name	Variable scope type	Default value
1	today	CURRENT_DATE	Valid in the root job	
2	yesterday	PREVIOUS_DATE	Valid in the root job	

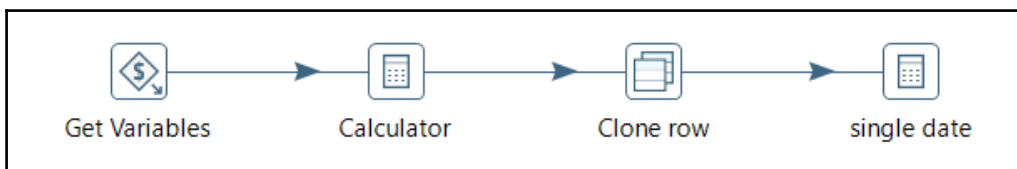
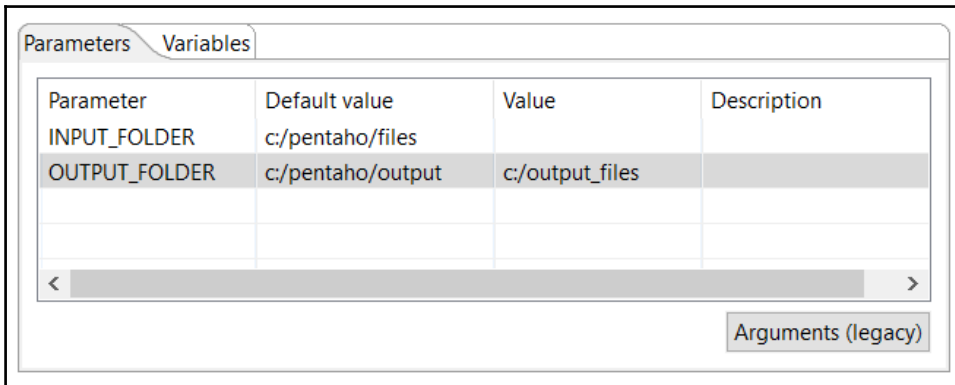
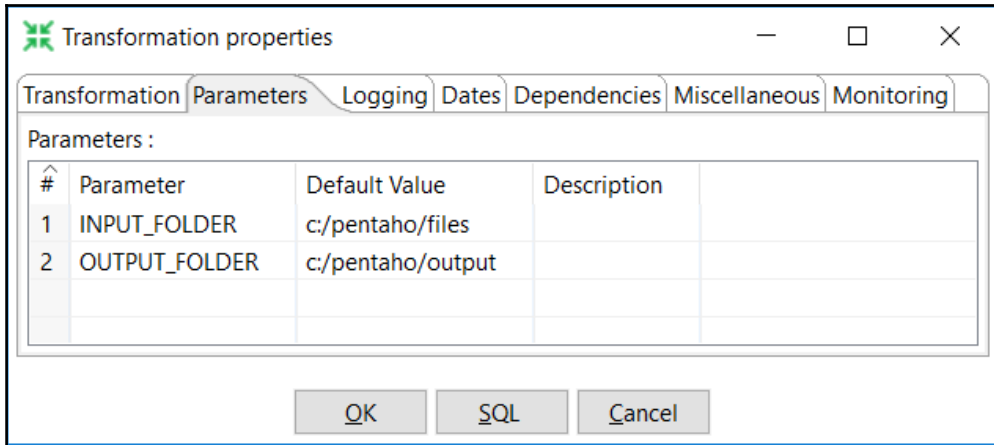
Help OK Cancel Get Fields

Notice

 Please remember that the variables you define with this step can't be used in this transformation. That is simply because all steps in a transformation run in parallel without a certain order of execution. As alternative correct usage, you can set the variables you want to use in the first transformation of a job.

Don't show this message again.

Close



Get Variable

Step name:

Fields:

#	Name	Variable	Type	Format	Length
1	start_date	\${DATE_FROM}	Date	yyyy-MM-dd	
2	end_date	\${DATE_TO}	Date	yyyy-MM-dd	

Examine preview data

Rows of step: single date (1000 rows)

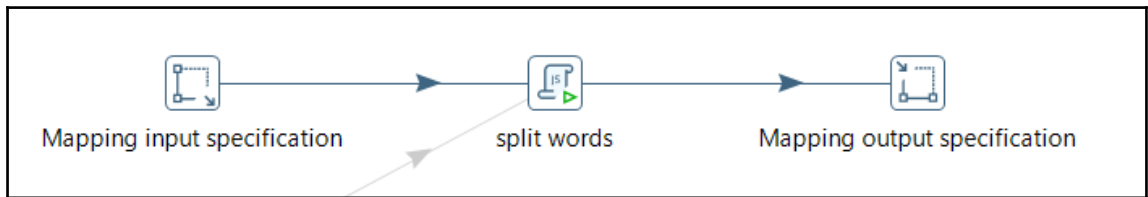
#	start_date	end_date	diff_dates	delta	a_single_date
1	2010-01-01	2020-12-31	4017	0	2010-01-01
2	2010-01-01	2020-12-31	4017	1	2010-01-02
3	2010-01-01	2020-12-31	4017	2	2010-01-03
4	2010-01-01	2020-12-31	4017	3	2010-01-04
5	2010-01-01	2020-12-31	4017	4	2010-01-05
6	2010-01-01	2020-12-31	4017	5	2010-01-06
7	2010-01-01	2020-12-31	4017	6	2010-01-07
8	2010-01-01	2020-12-31	4017	7	2010-01-08
9	2010-01-01	2020-12-31	4017	8	2010-01-09
1..	2010-01-01	2020-12-31	4017	9	2010-01-10

Examine preview data

Rows of step: split words (5 rows)

#	word	first_word	other_words
1	dog	Dog	
2	cat	Cat	
3	bird	Bird	
4	african elephant	African	ELEPHANT
5	australian sea lion	Australian	SEA LION

Close



Mapping Input Specification

Step name

The required input fields for this mapping (sub-transformation) :

#	Name	Type	Length	Precision
1	word	String		

Include unspecified fields, ordered by name

Help OK Cancel

Simple Mapping (sub-transformation)

Step Name:

Transformation:

Parameters Input Output

Fieldname from source step	Fieldname to mapping input step
name	word

Update mapped fieldnames downstream

Parameters Input Output

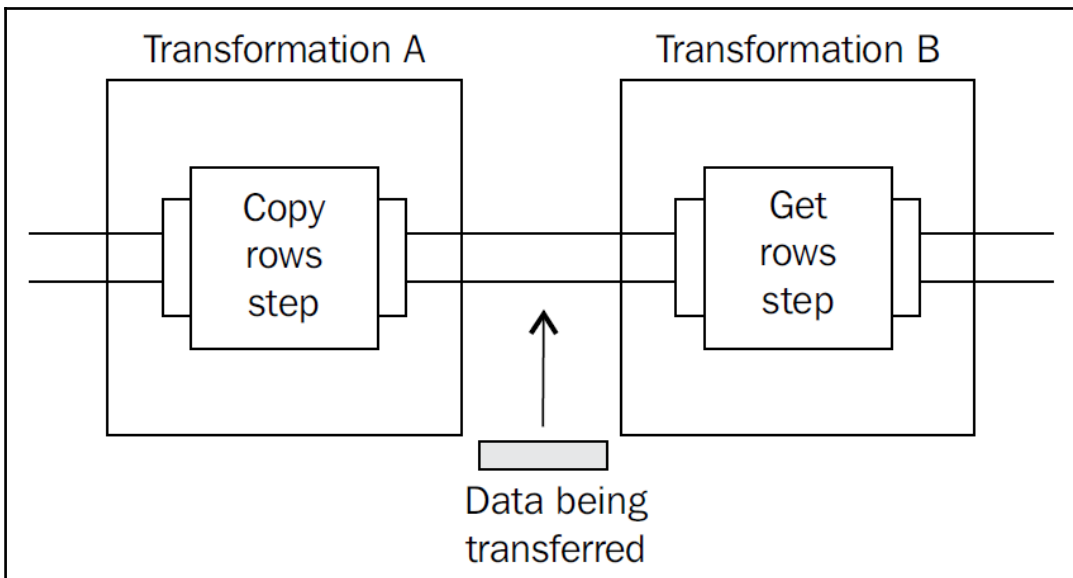
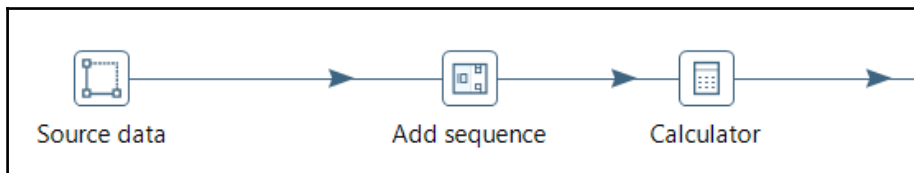
Fieldname from mapping step	Fieldname to target step
first_word	firstname
other_words	lastname

Examine preview data

Rows of step: preview (9 rows)

#	student_code	name	writing	reading	speaking	listening	firstname	lastname
1	80711-85	William Miller	81	83	80	90	William	MILLER
2	20362-34	Jennifer Martin	87	76	70	80	Jennifer	MARTIN
3	75283-17	Margaret Wilson	99	94	90	80	Margaret	WILSON
4	83714-28	Helen Thomas	89	97	80	80	Helen	THOMAS
5	61666-55	Maria Thomas	88	77	70	80	Maria	THOMAS
6	00647-35	David Collins	88	95	90	90	David	COLLINS

Close



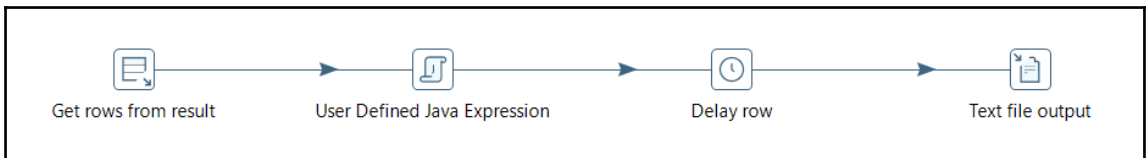
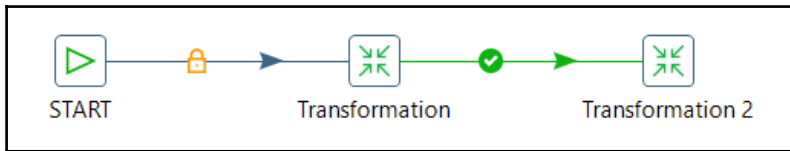
Get rows from previous result

Step name:

Fields :

#	Fieldname	Type	Length	Precision
1	student_co...	String		
2	name	String		
3	writing	Number		
4	reading	Number		
5	speaking	Number		
6	listening	Number		

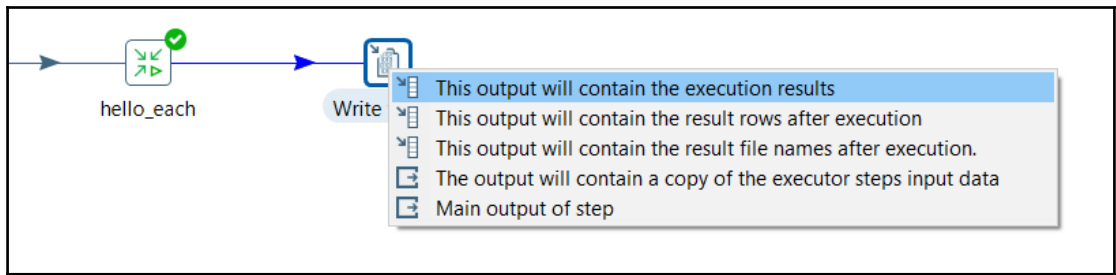
Help OK Cancel



Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active
1	Examinations	0	0	83	108	0	5	0	0	Finished
2	writing below 60	0	83	4	0	0	0	0	0	Finished
3	hello_each	0	4	0	0	0	0	0	0	Finished



Parameters Execution results Row grouping Result rows Result files

Target step for execution results:
 preview results

Field description	Field name
Execution time (ms)	
Execution result	
Number of errors	
Number of rows read	
Number of rows written	
Number of rows input	
Number of rows output	
Number of rows rejected	
Number of rows updated	
Number of rows deleted	
Number of files retrieved	
Exit status	exitstatus
Execution logging text	ExecutionLogText
Log channel ID	

Examine preview data

Rows of step: preview execution results (4 rows)

#	exitstatus	ExecutionLogText
1	0	2017/11/19 09:36:00 - hello_each - Dispatching started for transformation [hello_each]2017/11/19 09:...
2	0	2017/11/19 09:36:00 - hello_each - Dispatching started for transformation [hello_each]2017/11/19 09:...
3	0	2017/11/19 09:36:00 - hello_each - Dispatching started for transformation [hello_each]2017/11/19 09:...
4	0	2017/11/19 09:36:00 - hello_each - Dispatching started for transformation [hello_each]2017/11/19 09:...

Close

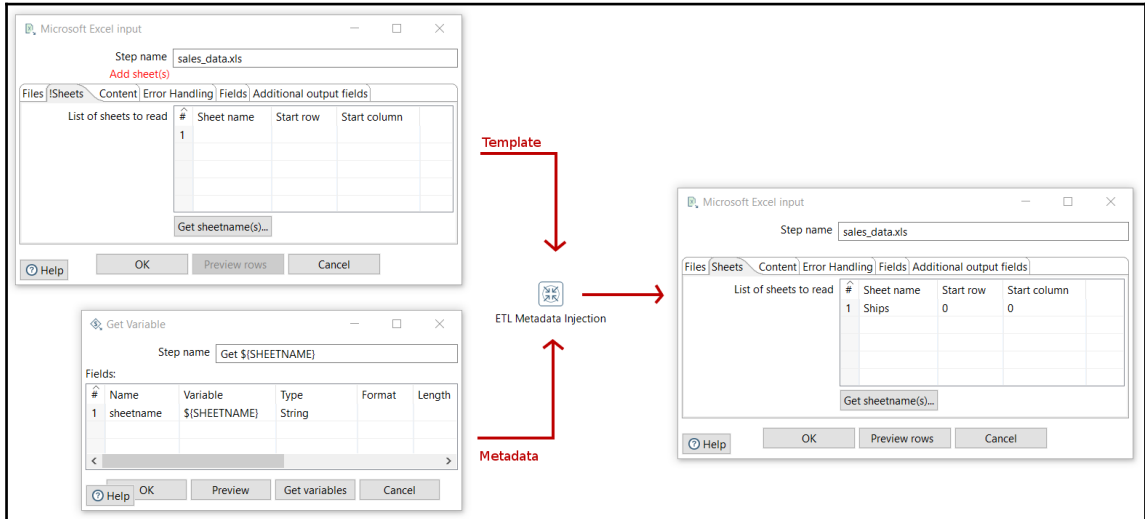
Parameters Execution results Row grouping Result rows Result files

Target step for result rows:
 preview results

Expected layout for result rows:

Field name	Data type	Length	Precision
student_code	String		
name	String		
message	String		

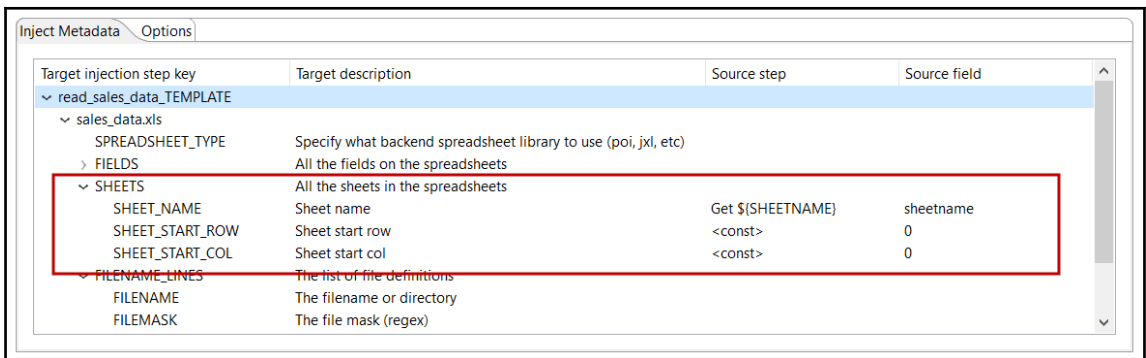
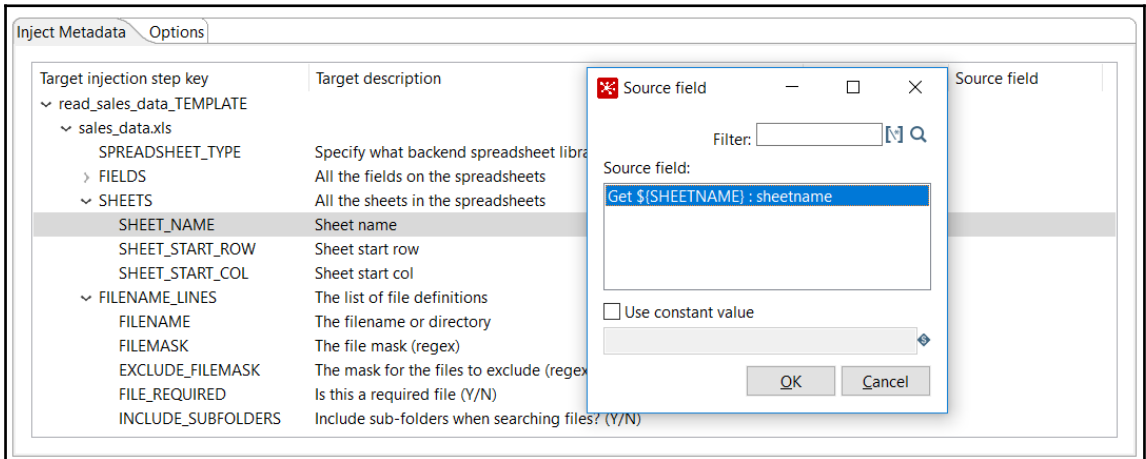
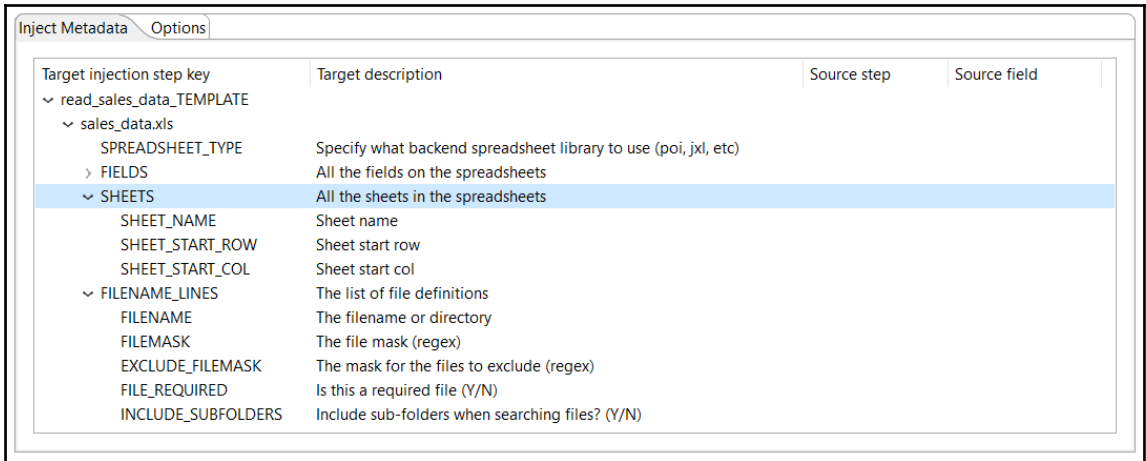
Chapter 13: Implementing Metadata Injection

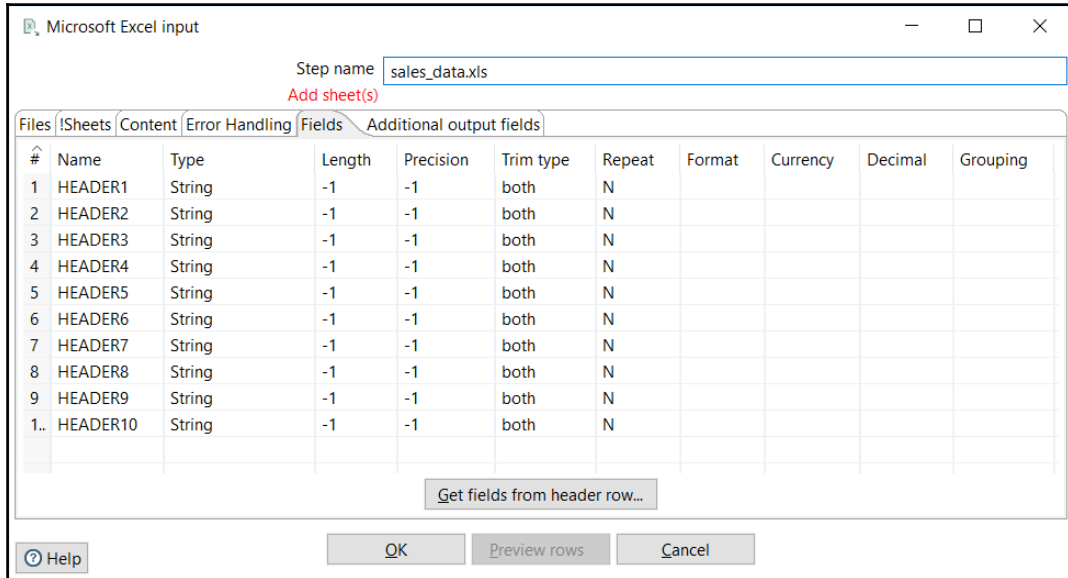
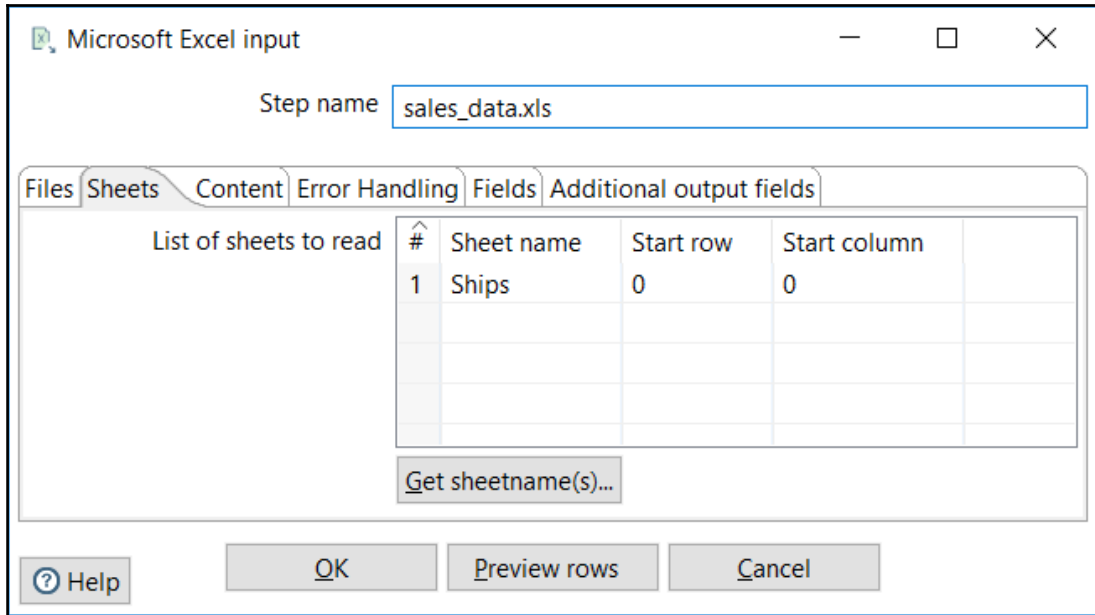


	A	B	C	D	E	F	G
1	ORDERDATE	ORDERNUMBER	PRODUCTCODE	QUANTITYORDERED	PRICEEACH	SALES	
2	3/30/2005	10398	S700_2047	36	100	3910.32	
3	1/9/2004	10209	S18_3029	28	100	2817.92	
4	9/28/2003	10153	S700_1138	43	64.67	2780.81	
5	10/14/2004	10306	S700_2610	43	75.17	3232.31	
6	2/17/2003	10106	S18_3029	41	83.44	3421.04	
7	3/20/2004	10232	S700_3962	35	82.43	2885.05	
8	2/23/2005	10384	S700_1938	49	100	6397.44	
9	11/21/2004	10337	S700_1938	36	70.3	2530.8	
10	6/16/2003	10130	S18_3029	40	96.34	3853.6	
11	5/6/2005	10414	S700_1138	37	71.34	2639.58	
12	7/21/2004	10273	S72_3212	37	45.86	1696.82	
13	11/1/2004	10316	S700_2610	48	74.45	3573.6	
14	2/17/2003	10106	S700_2047	30	100	3177.3	
15	3/30/2005	10398	S18_3029	28	72.26	2023.28	

intro Ships

Sheet 2 / 2 PageStyle_Ships Sheet 2 | STD | Sum=0





Examine preview data

Rows of step: sales_data.xls (1 rows)

#	HEADER1	HEADER2	HEADER3	HEADER4	HEADERS5	HEADER6	HEADER7	HEADER8	HEADER9	HEADER10
1	ORDERDATE	ORDERNUMBER	ORDERLINENUMBER	PRODUCTCODE	PRODUCTLINE	QUANTITYORDERED	PRICEEACH	SALES	<null>	<null>

Close Show Log

Row Normaliser

Step name: Row Normaliser

Type field: HEADER

Fields

#	Fieldname	Type	new field
1	HEADER1	HEADER1	fieldname
2	HEADER2	HEADER2	fieldname
3	HEADER3	HEADER3	fieldname
4	HEADER4	HEADER4	fieldname
5	HEADER5	HEADERS5	fieldname
6	HEADER6	HEADER6	fieldname
7	HEADER7	HEADER7	fieldname
8	HEADER8	HEADER8	fieldname
9	HEADER9	HEADER9	fieldname
1..	HEADER10	HEADER10	fieldname

Help OK Cancel Get Fields

Examine preview data

Rows of step: Row Normaliser (10 rows)

#	HEADER	fieldname
1	HEADER1	ORDERDATE
2	HEADER2	ORDERNUMBER
3	HEADER3	ORDERLINENUMBER
4	HEADER4	PRODUCTCODE
5	HEADERS5	PRODUCTLINE
6	HEADER6	QUANTITYORDERED
7	HEADER7	PRICEEACH
8	HEADER8	SALES
9	HEADER9	<null>
1..	HEADER10	<null>

Close

Inject Metadata Options

Target injection step key	Target description	Source step	Source field
<ul style="list-style-type: none"> ▼ read_sales_data_TEMPLATE <ul style="list-style-type: none"> > sales_data.xls ▼ select product info <ul style="list-style-type: none"> SELECT_UNSPECIFIED Include unspecified fields, ordered by name <ul style="list-style-type: none"> ▼ FIELDS Selected fields <ul style="list-style-type: none"> FIELD_NAME Fieldname FIELD_RENAME Rename to FIELD_LENGTH Length FIELD_PRECISION Precision ▼ REMOVES Removed fields <ul style="list-style-type: none"> REMOVE_NAME Fieldname > METAS List of fields to change metadata for 			
		fieldname in list	HEADER
		fieldname in list	fieldname

Select / Rename values

Step name:

Select & Alter Remove Meta-data

Fields :

#	Fieldname	Rename to	Length	Precision
1	HEADER4	PRODUCTCODE		
2	HEADER5	PRODUCTLINE		
3	HEADER6	QUANTITYORDERED		

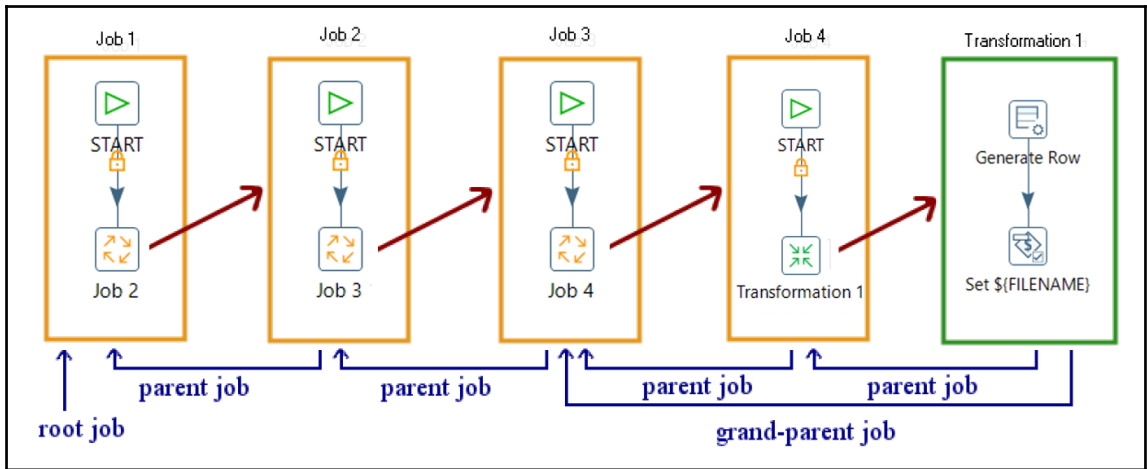
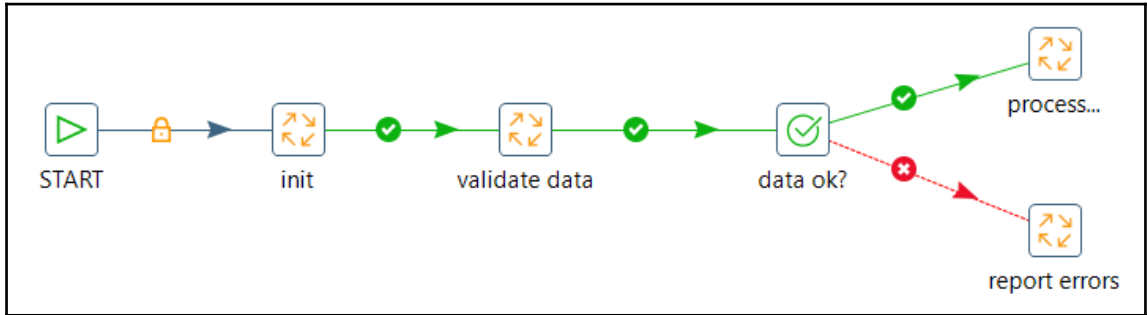
Get fields to select

Edit Mapping

Include unspecified fields, ordered by

Help OK Cancel

Chapter 14: Creating Advanced Jobs



Options | Logging | Arguments | Parameters

Copy results to parameters

Pass parameter values to sub transformation

Parameter	Stream Column Name	Value
OUTPUT_FOLDER		c:/pentaho/files2017

Get Parameters

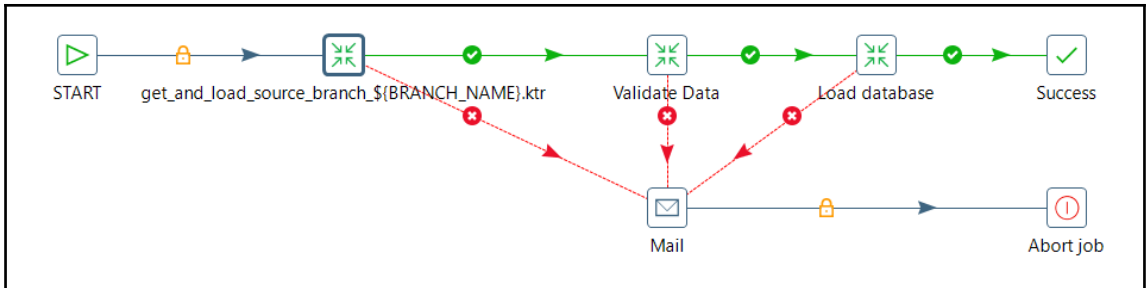
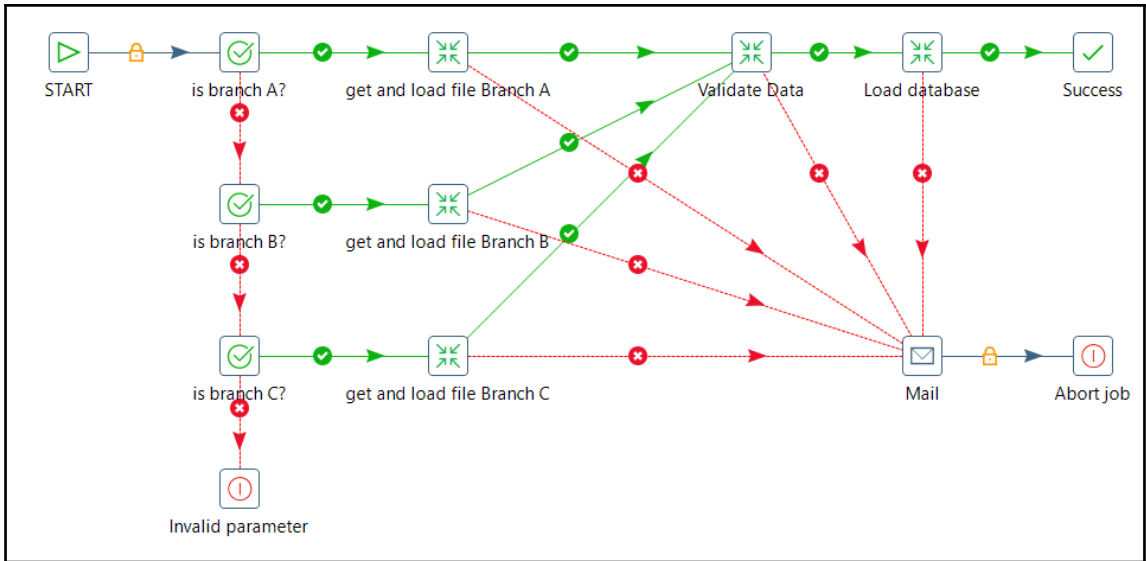
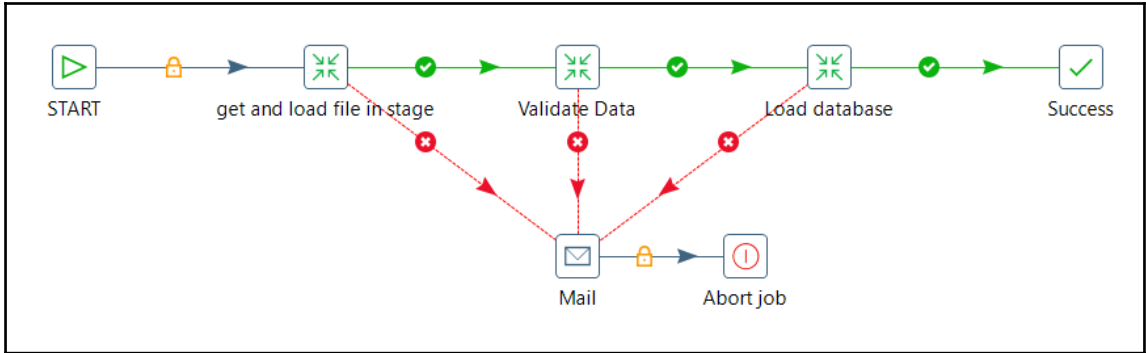
Options | Logging | Arguments | Parameters

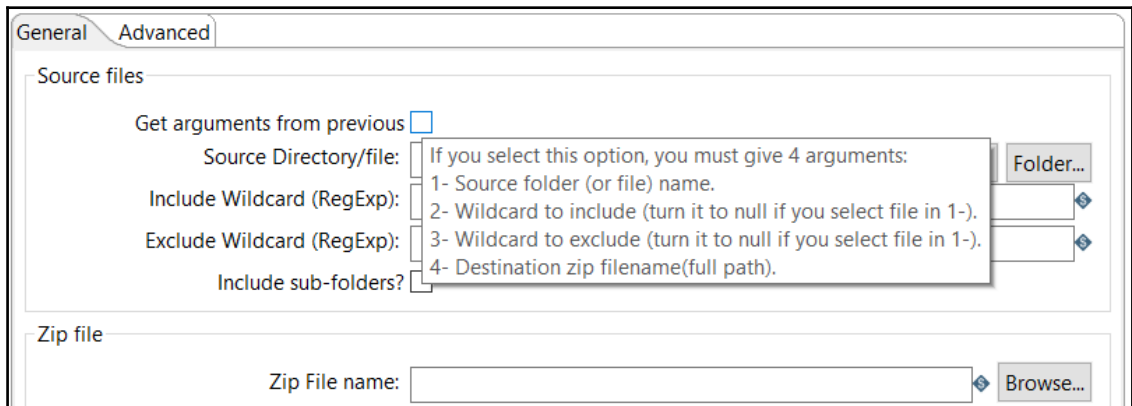
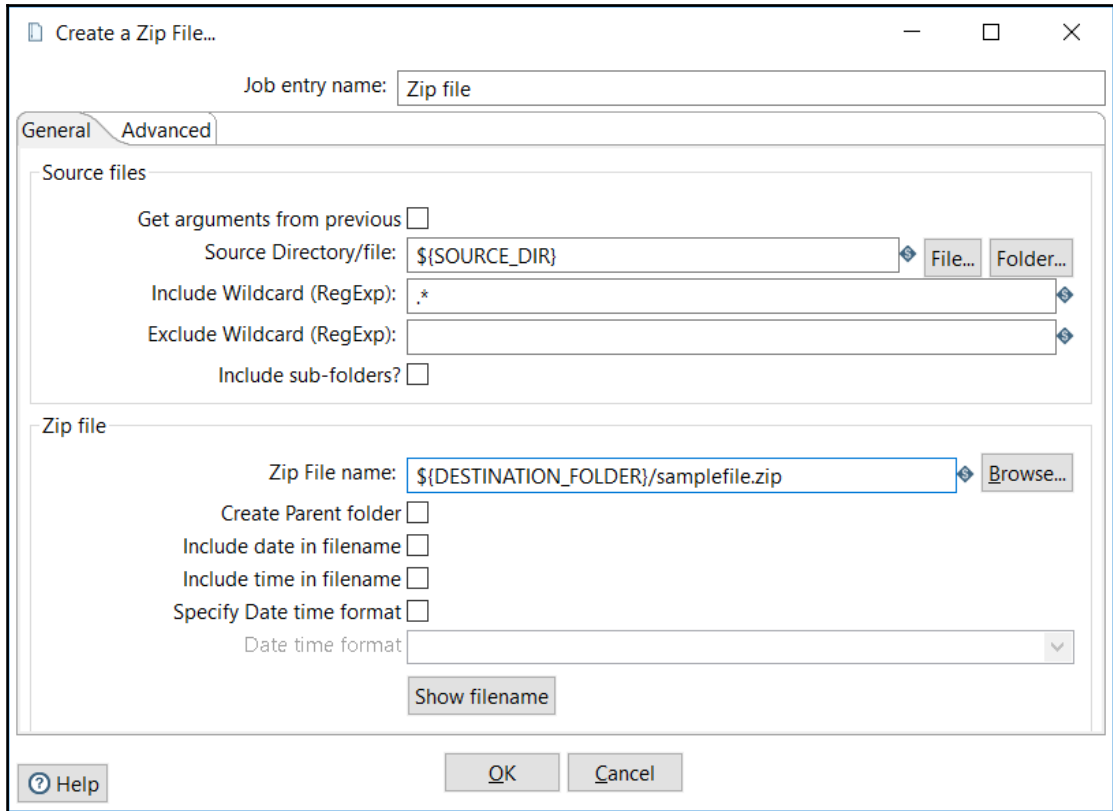
Copy results to parameters

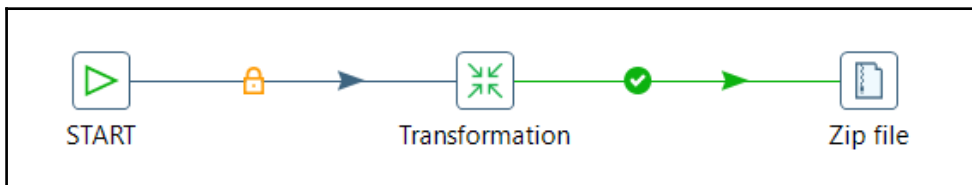
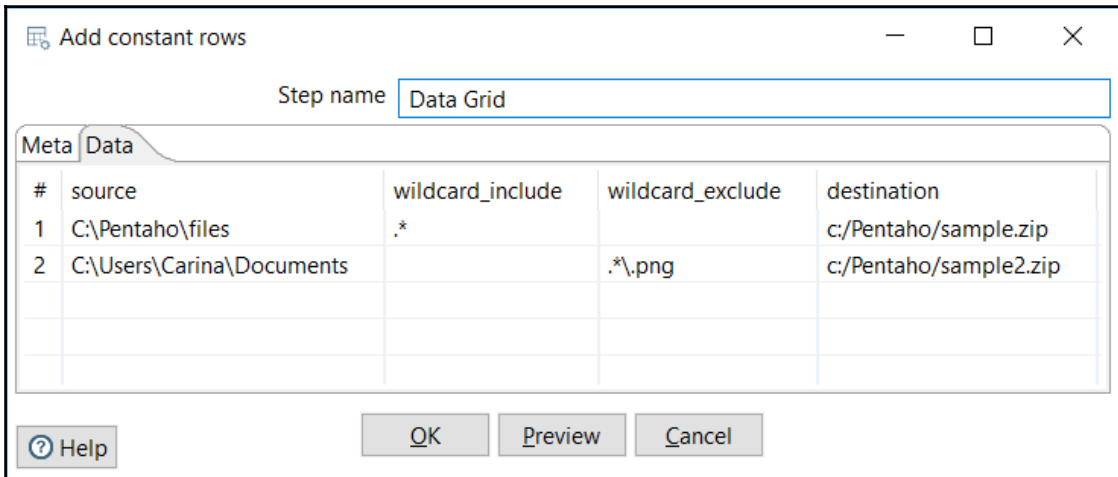
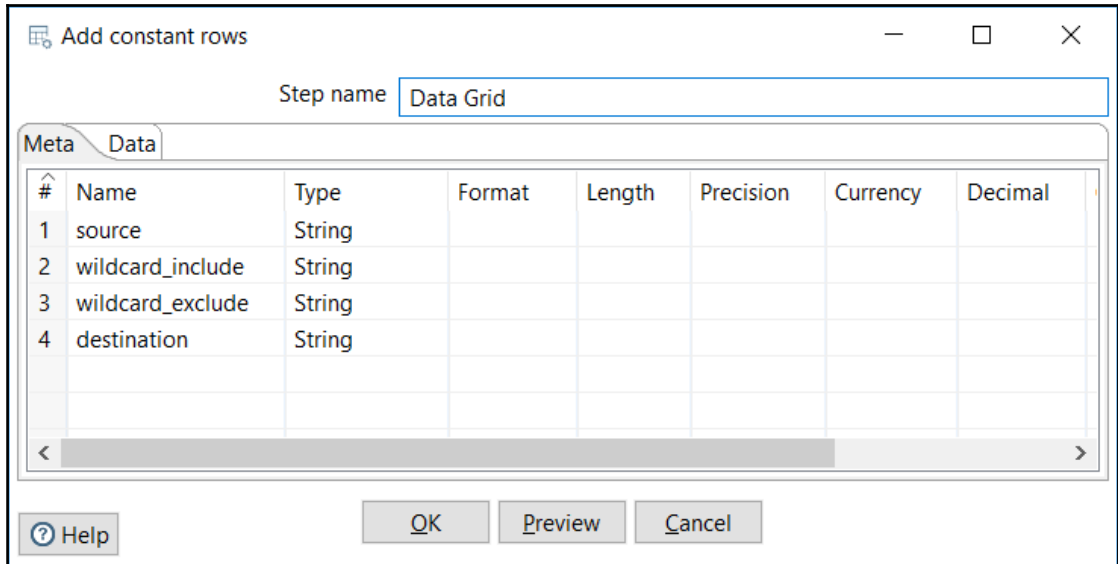
Pass parameter values to sub transformation

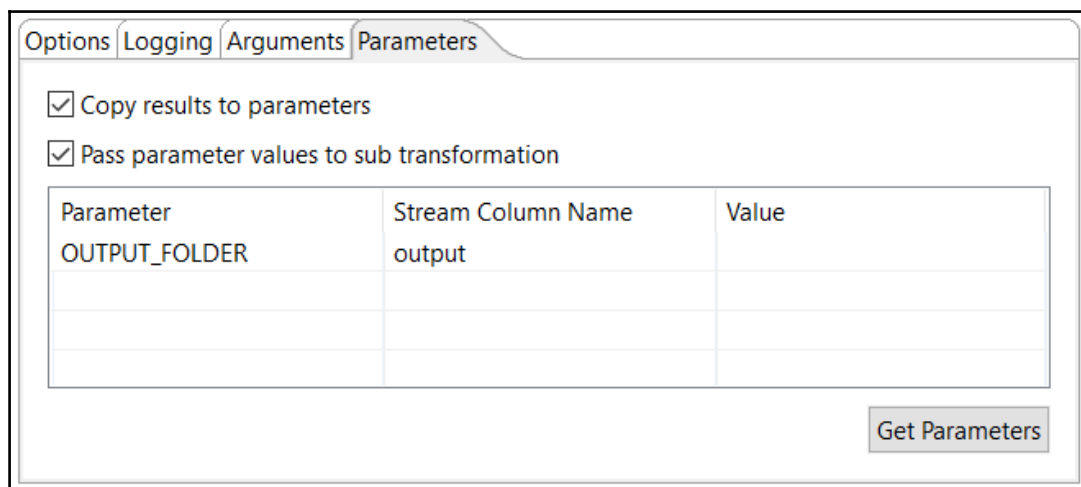
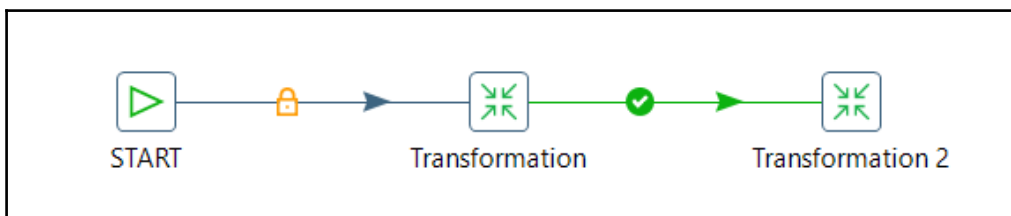
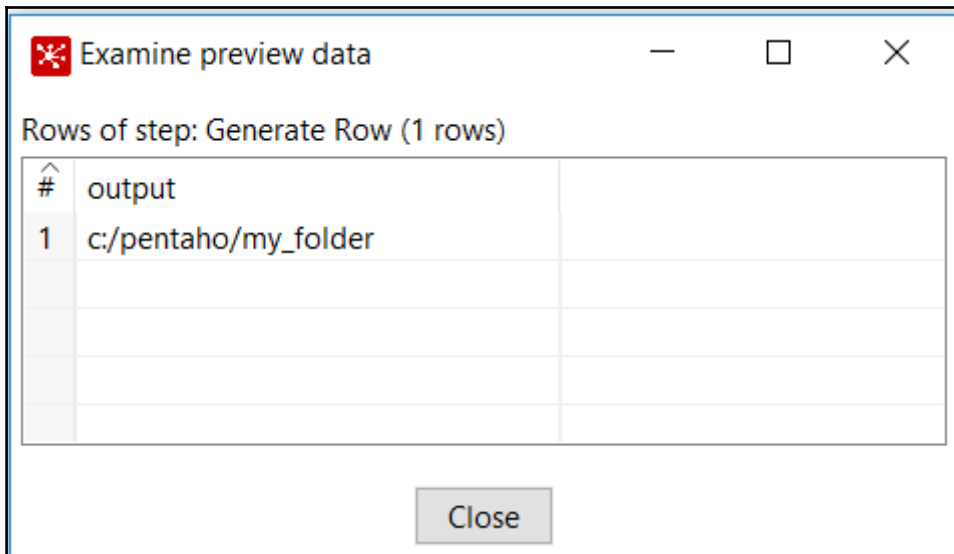
Parameter	Stream Column Name	Value
OUTPUT_FOLDER		\${OUTPUT_FOLDER}

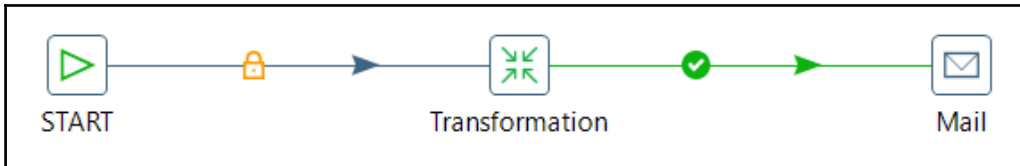
Get Parameters











Job mail details

Name of mail job entry: Mail with attachment

Addresses | Server | EMail Message | Attached Files

Files added in result filename

Attach file(s) to message?

Select file type: **General**
Log
Error line
Error
Warning

Zip files to single archive?

Name of zip archive:

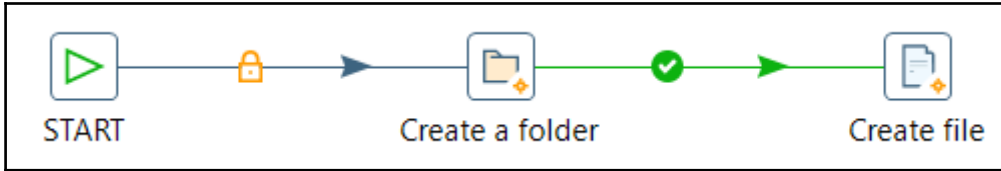
Embedded images

Filename Add Browse files ...

Content ID

#	Image	Content ID	
1			Delete Edit

Help OK Cancel



Add constant rows

Step name:

Meta Data

#	foldername
1	c:/pentaho/files/folder1
2	c:/pentaho/files/folder2
3	c:/pentaho/files/folder3
4	c:/pentaho/files/folder4
5	c:/pentaho/files/folder5

Buttons: Help, OK, Preview, Cancel

folders → Generate random value → filename

Generate random value dialog:

Step name:

#	Name	Type
1	random_value	Random string

Buttons: Help, OK, Cancel

User Defined Java Expression dialog:

Step name:

#	New field	Java expression	Value type	Length	P
1	filename	"sample_" + random_value + ".tmp"	String		

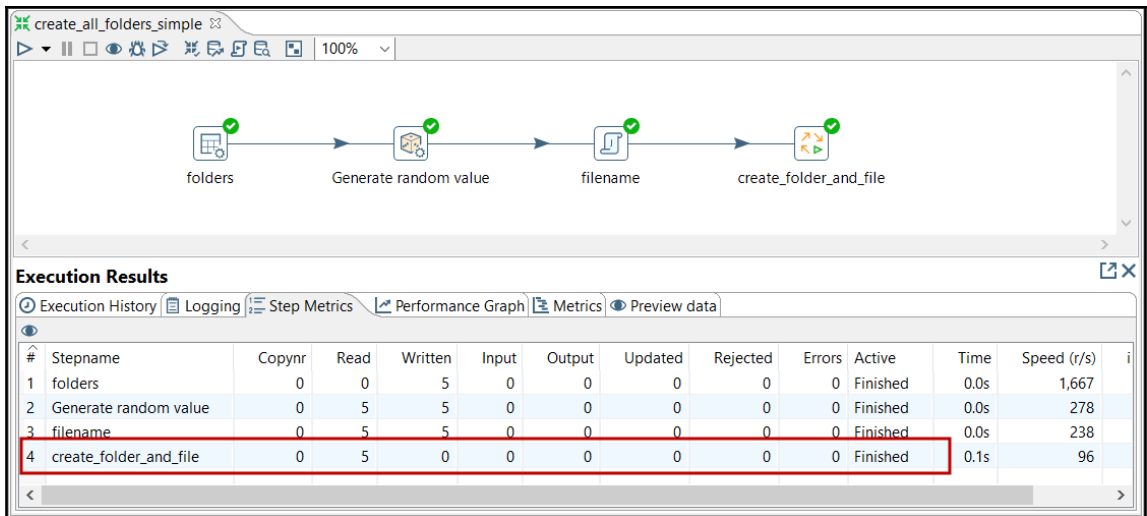
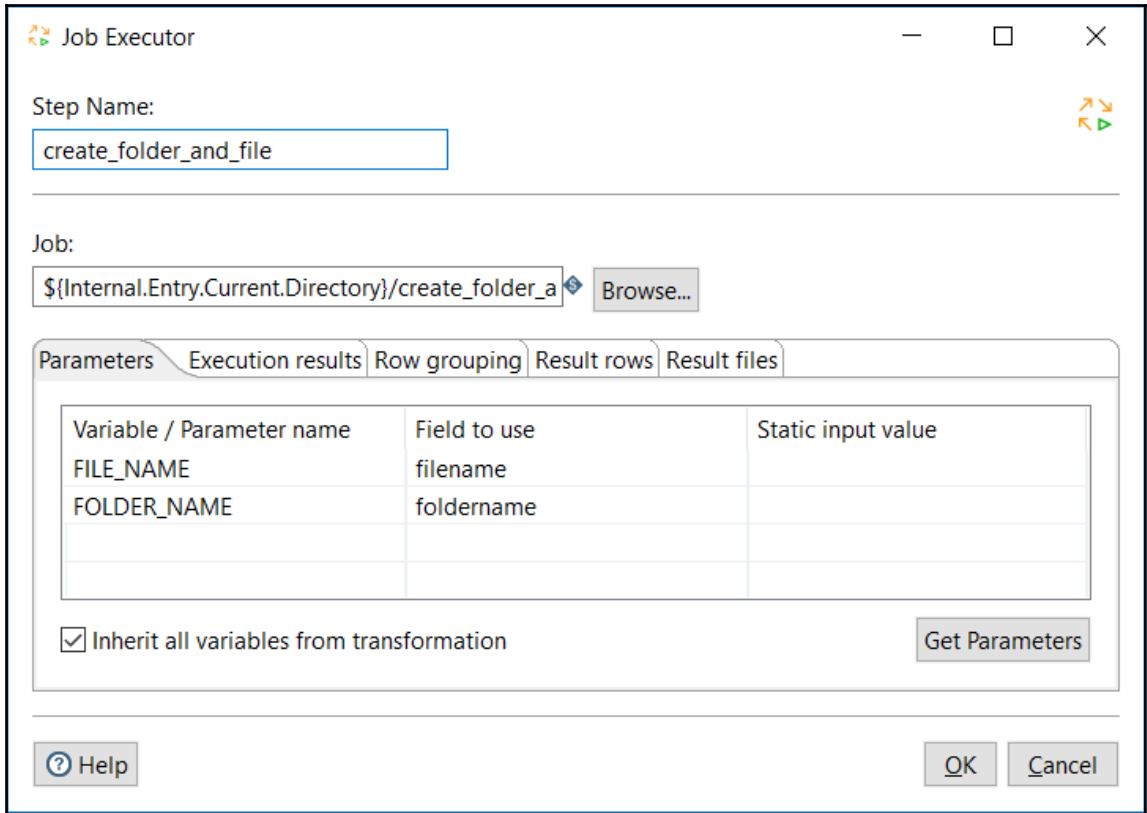
Buttons: Help, OK, Cancel

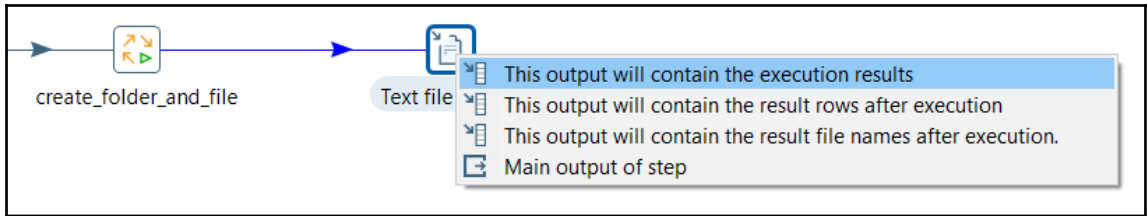
Examine preview data

Rows of step: filename (5 rows)

#	foldername	random_value	filename
1	c:/pentaho/files/folder1	4932seoa1v0h1	sample_4932seoa1v0h1.tmp
2	c:/pentaho/files/folder2	3sf4v5ri5v7iq	sample_3sf4v5ri5v7iq.tmp
3	c:/pentaho/files/folder3	787ligpm72otn	sample_787ligpm72otn.tmp
4	c:/pentaho/files/folder4	3anvam0la18ud	sample_3anvam0la18ud.tmp
5	c:/pentaho/files/folder5	654l09ihfrkr6	sample_654l09ihfrkr6.tmp

Close





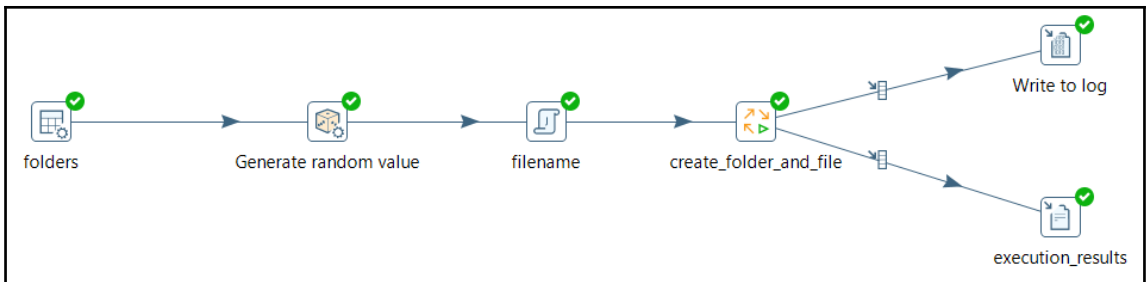
Parameters		Execution results	Row grouping	Result rows	Result files
Target step for execution results:					
execution_results					
Field description	Field name				
Execution time (ms)					
Execution result					
Number of errors					
Number of rows read					
Number of rows written					
Number of rows input					
Number of rows output					
Number of rows rejected					
Number of rows updated					
Number of rows deleted					
Number of files retrieved					
Exit status					
Execution logging text	ExecutionLogText				
Log channel ID					

Examine preview data

Rows of step: execution_results (5 rows)

#	ExecutionLogText
1	2017/10/26 23:45:53 - create_folder_and_file - Starting entry [Create a folder]2017/10/26 23:45:53...
2	2017/10/26 23:45:53 - create_folder_and_file - Starting entry [Create a folder]2017/10/26 23:45:53...
3	2017/10/26 23:45:53 - create_folder_and_file - Starting entry [Create a folder]2017/10/26 23:45:53...
4	2017/10/26 23:45:53 - create_folder_and_file - Starting entry [Create a folder]2017/10/26 23:45:53...
5	2017/10/26 23:45:53 - create_folder_and_file - Starting entry [Create a folder]2017/10/26 23:45:53...

Close

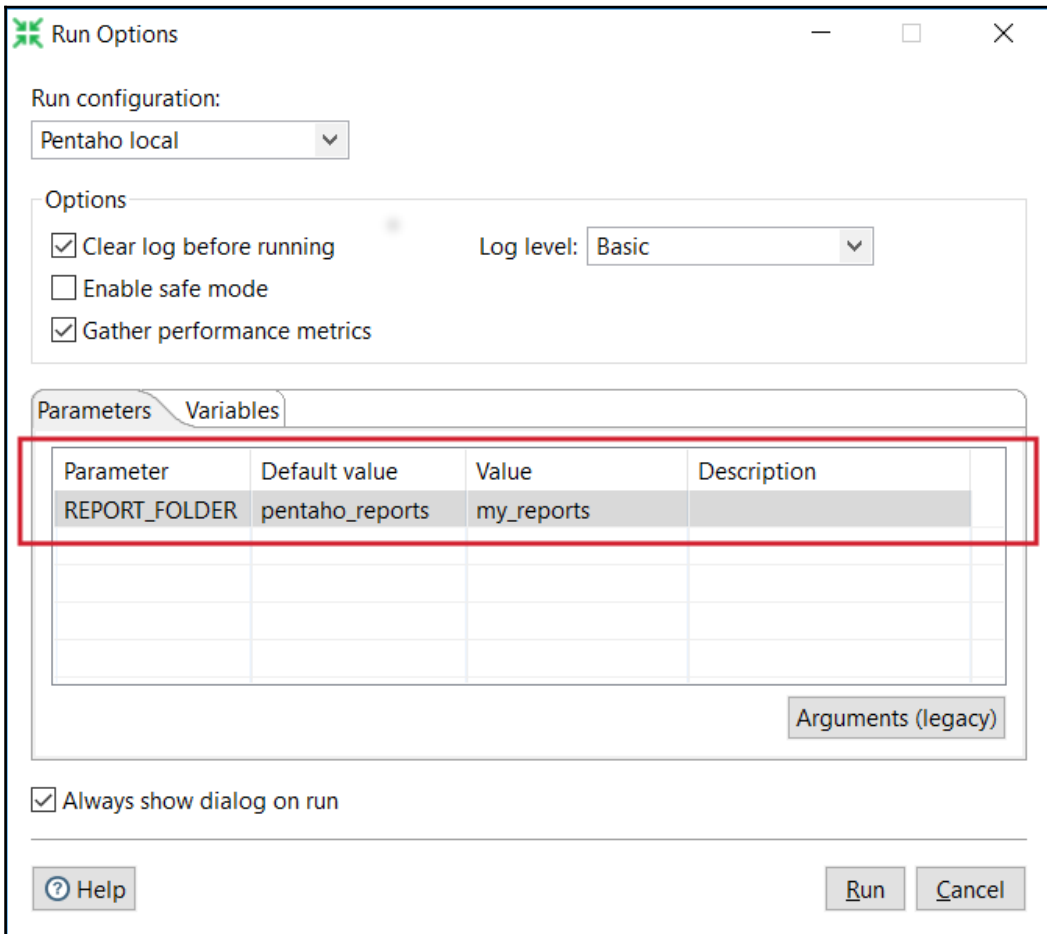


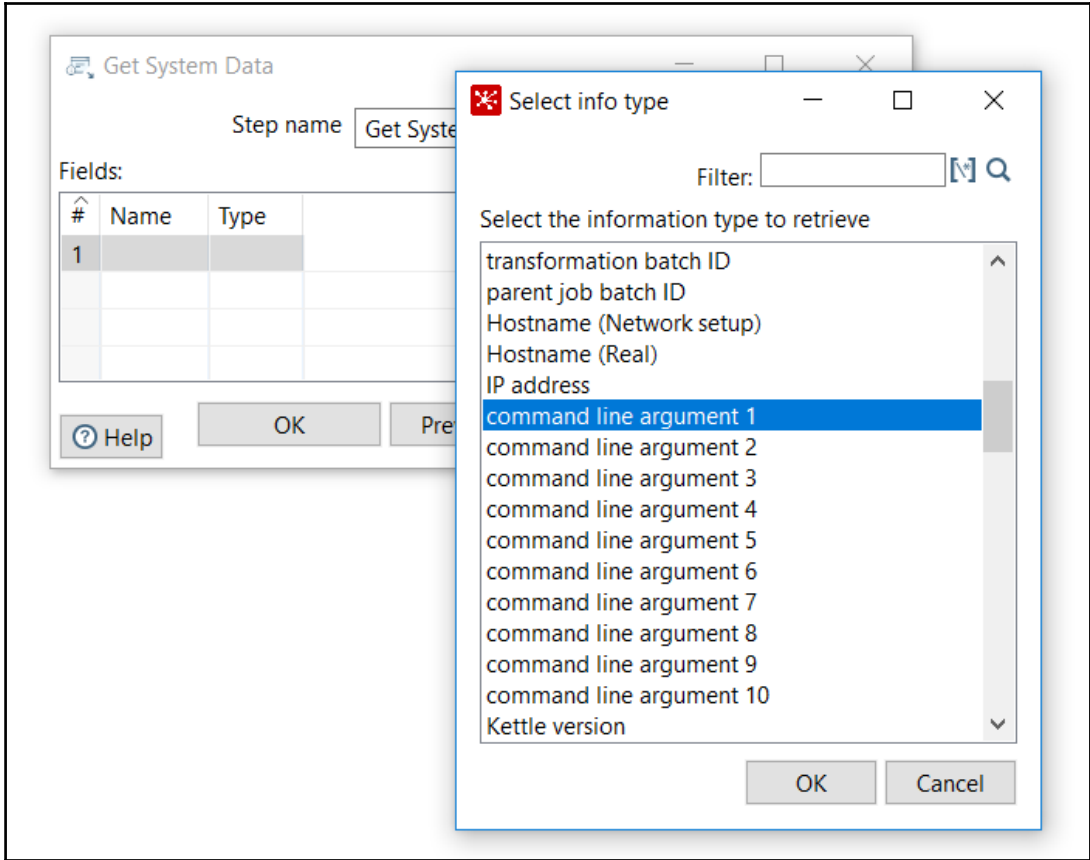
Parameters | Execution results | Row grouping | Result rows | Result files

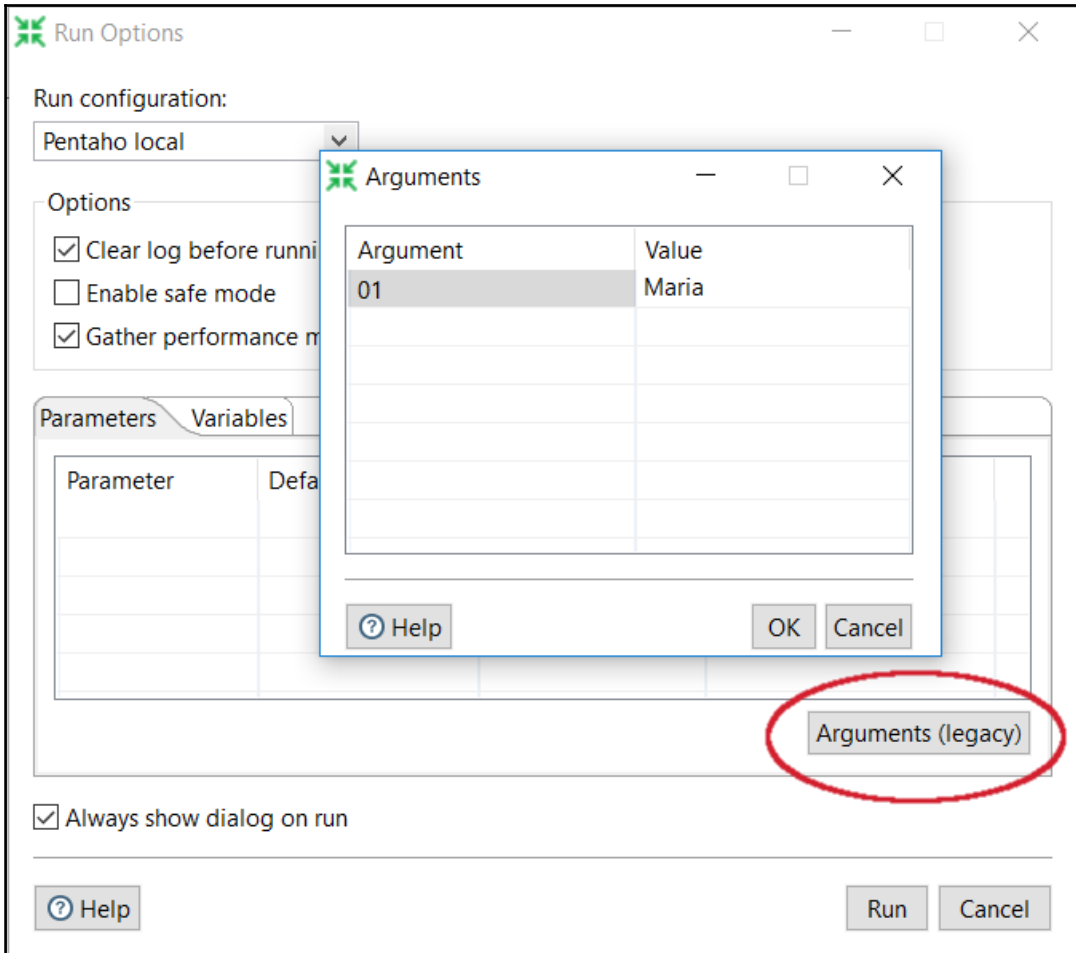
Target step for result files information:

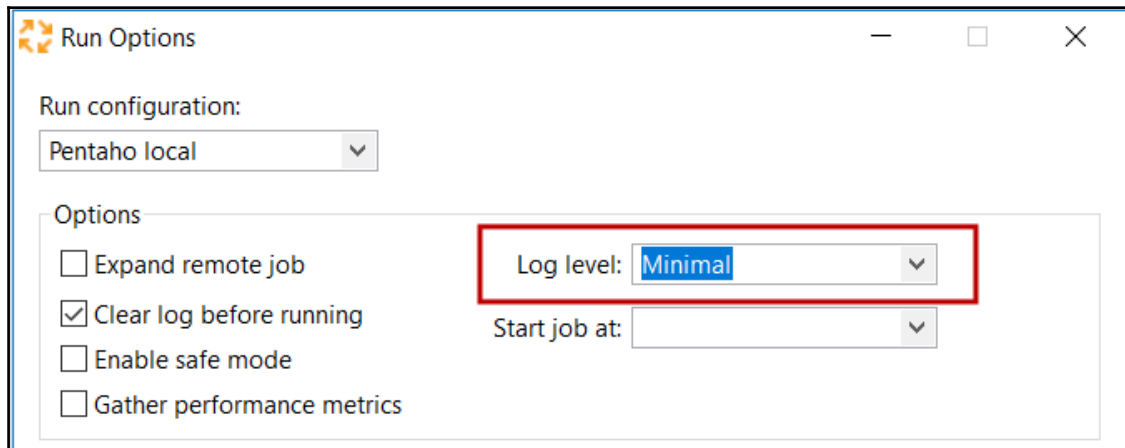
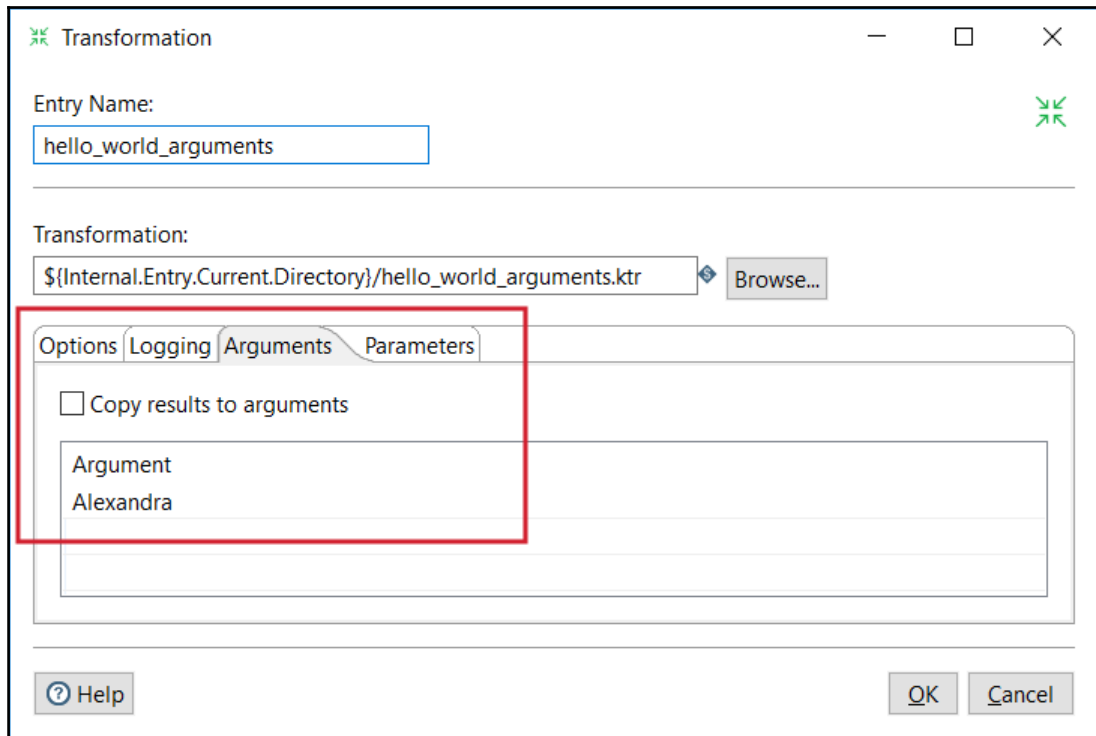
Result file name field:

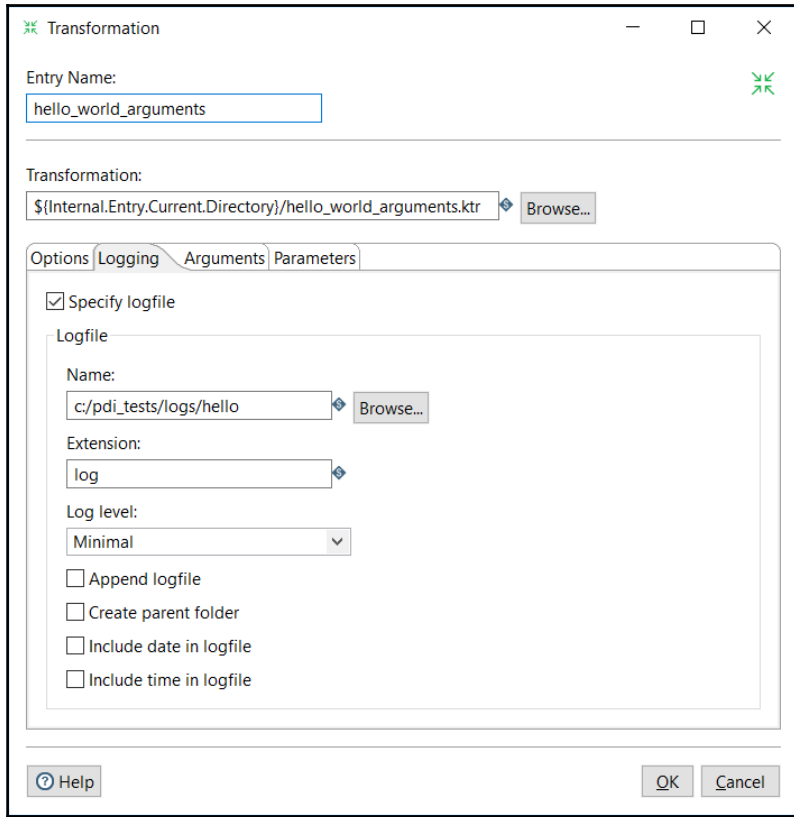
Chapter 15: Launching Transformations and Jobs from the Command Line



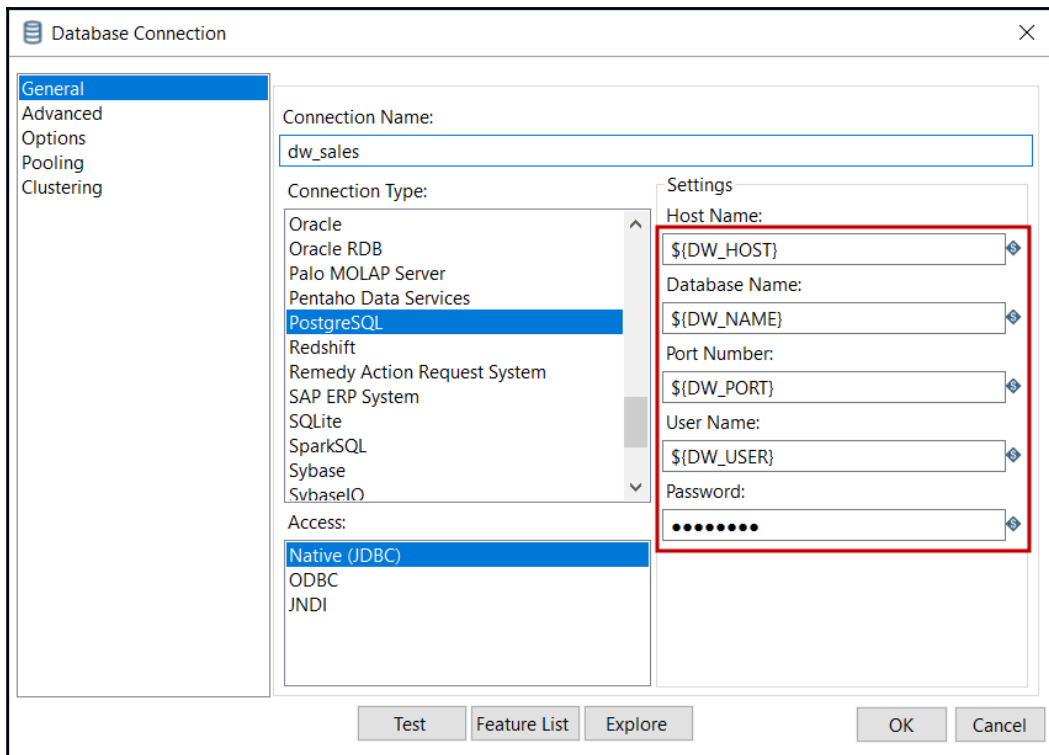
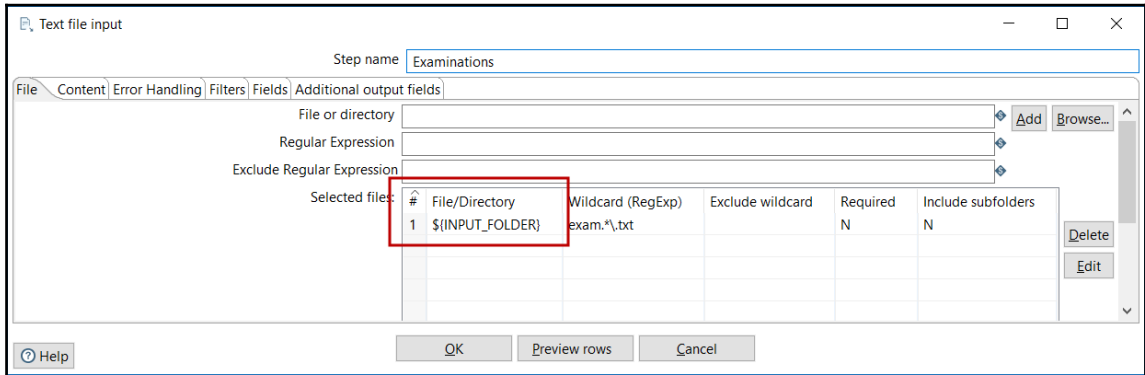


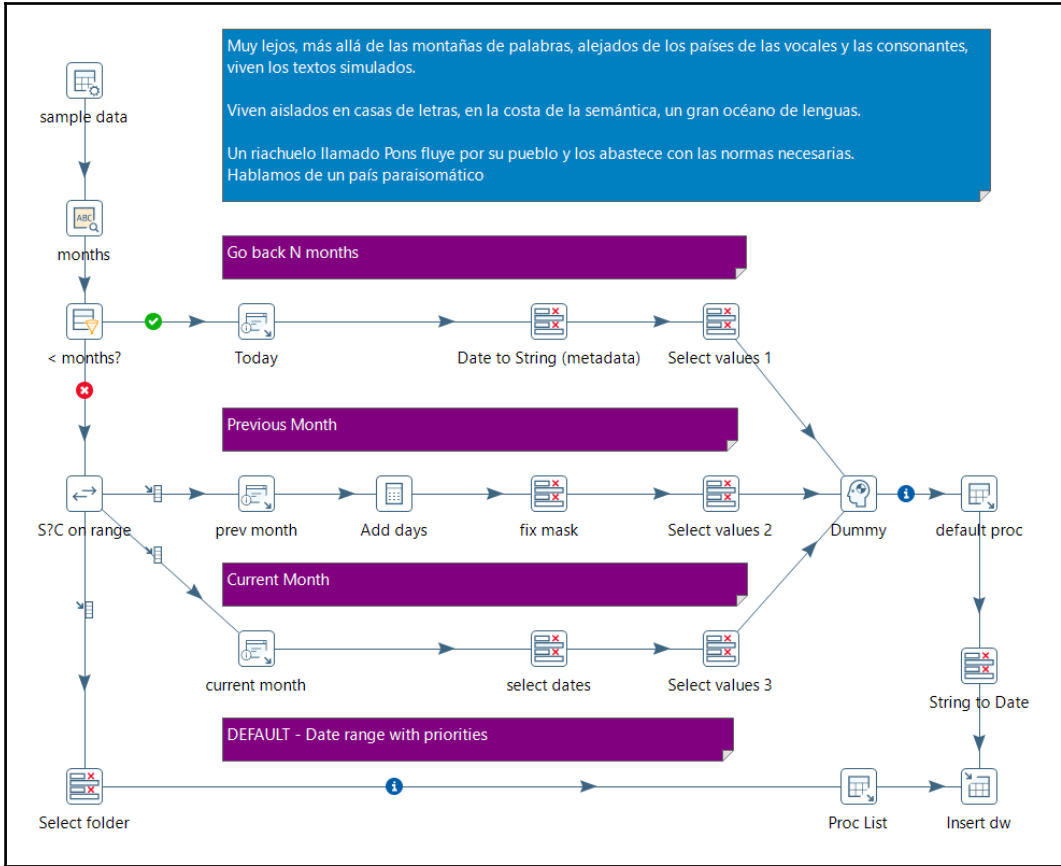


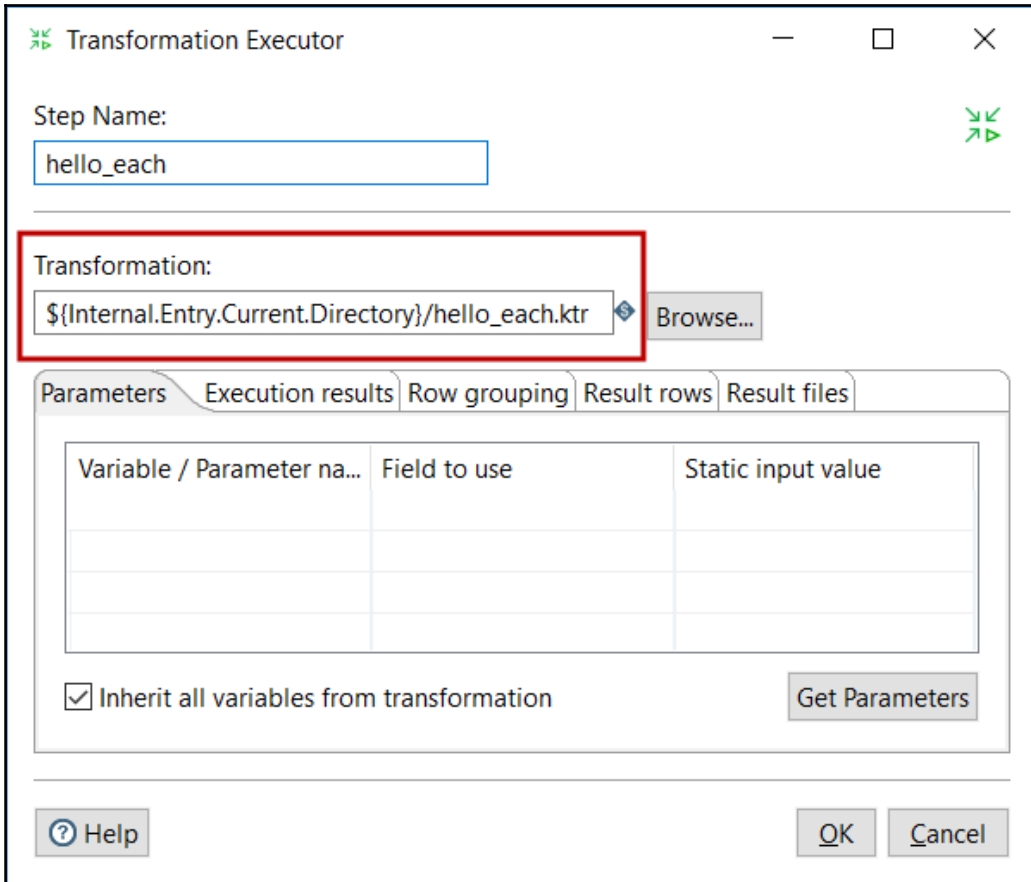




Chapter 16: Best Practices for Designing and Deploying a PDI Project



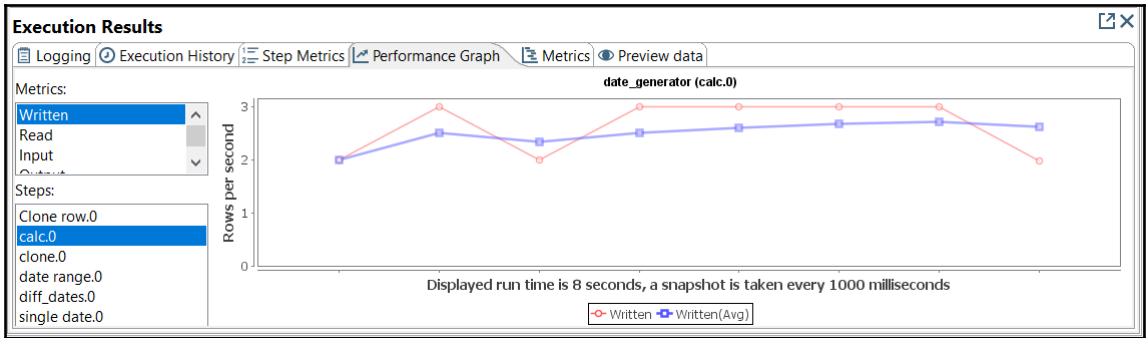




Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)	input/output
1	date range	0	0	1	0	0	0	0	0	Finished	0.0s	59	-
2	diff_dates	0	1	1	0	0	0	0	0	Finished	0.0s	26	-
3	Clone row	0	1	4017	0	0	0	0	0	Finished	0.1s	58,217	-
4	clone	0	24	23	0	0	0	0	0	Running	8.6s	3	3993/0
5	calc	0	23	23	0	0	0	0	0	Running	8.6s	3	0/0
6	single date	0	23	23	0	0	0	0	0	Running	8.6s	3	0/0



New System Variable

Variable name:

Variable value: