



Jobs

Generated by Talend Cloud Data
Management Platform

Project Name	talendcloud	GENERATION DATE	31 Jan, 2019 2:59:20 PM
AUTHOR	Joshi, Nishad	Talend Cloud VERSION	7.1.1.20181026_1147

Summary

[Project Description](#)

[Description](#)

[Preview Picture](#)

[Settings](#)

[Context List](#)

[Component List](#)

[Components Description](#)

Project Description

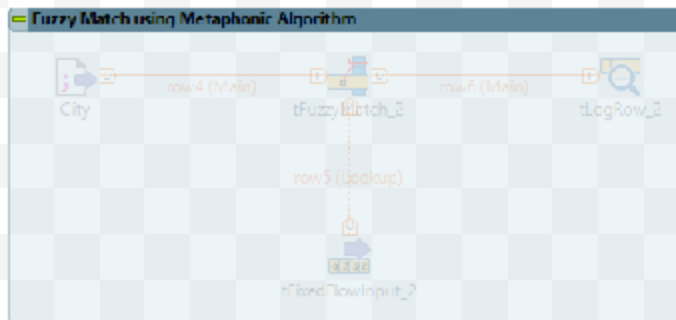
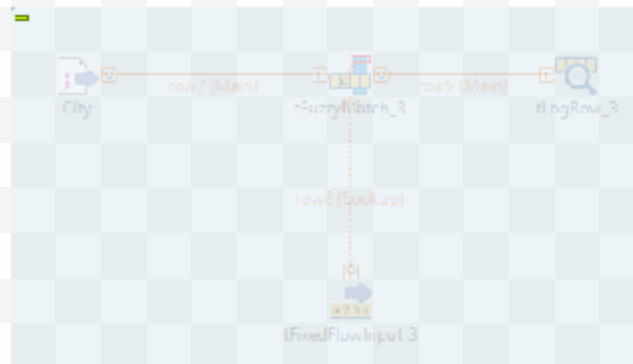
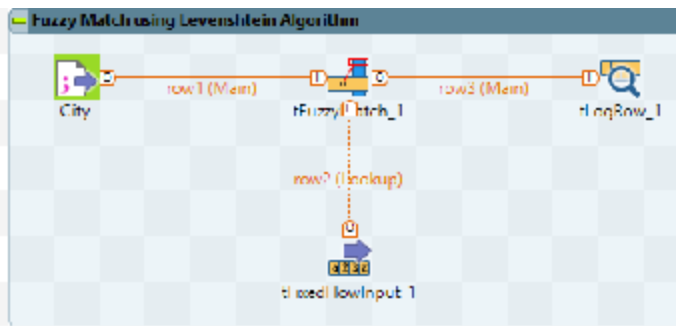
Properties	Values
Name	talendcloud
Language	java

Description	
-------------	--

Description

Properties	Values
Name	FuzzyMatch
Author	Joshi, Nishad
Version	1.0
Purpose	
Status	
Description	
Creation	27 Jul, 2016 1:58:23 AM
Modification	7 Aug, 2018 3:27:28 PM

Preview Picture



Settings

Extra settings

Name	Value
COMP_DEFAULT_FILE_DIR	
Multi thread execution	false

Implicit tContextLoad	false
-----------------------	-------

Status & Logs

Name	Value
Use statistics (tStatCatcher)	false
Use logs (tLogCatcher)	false
Use volumetrics (tFlowMeterCatcher)	false
On Console	false
On Files	false
On Databases	false
Catch components statistics	false
Catch runtime errors	true
Catch user errors	true
Catch user warnings	true

Context List

ContextDefault


Name	Prompt	Need Prompt?	Type	Value	Source
------	--------	--------------	------	-------	--------

Component List

Component Name	Component Type
tFileInputDelimited_1	tFileInputDelimited
tFixedFlowInput_1	tFixedFlowInput
tFuzzyMatch_1	tFuzzyMatch
tLogRow_1	tLogRow

Components Description

Component tFileInputDelimited

	UNIQUE NAME	tFileInputDelimited_1	INPUT(S)	none
	LABEL	City	OUTPUT(S)	tFuzzyMatch_1

Component Parameters:

Properties	Values
Unique Name	tFileInputDelimited_1
Component Name	tFileInputDelimited
Version	0.102 (ALPHA)
Family	File/Input
Start	true
Startable	true

SUBTREE_START	true
END_OF_FLOW	false
Activate	true
DUMMY	false
tStatCatcher Statistics	false
Help	org.talend.help.tFileInputDelimited
Update components	true
IREPORT_PATH	
JAVA_LIBRARY_PATH	C:\Talend-Studio-20181026_1147-V7.1.1\configuration\lib\java
Subjob color	
Title color	
Enable parallel execution	false
Number of parallel executions	2
PARALLELIZE_KEEP_EMPTY	false
Property Type	Built-In
REPOSITORY_ALLOW_AUTO_SWITCH	true
Schema	repository: DELIM:City - metadata
File name/Stream	"F:/Trainings/CITI/SampleFiles/City_Delimited.txt"
CSV options	false
Row Separator	"\n"
CSV Row Separator	"\n"
Field Separator	"", ,


Escape char	""
Text enclosure	""
Header	1
Footer	0
Limit	40
Skip empty rows	false
Uncompress as zip file	false
Die on error	false
Schema	repository: DELIM:City - metadata
!!!TEMP_DIR.NAME!!!	"D:/Talend-Tools-Studio-20151214_1327-V6.1.1/workspace"
Advanced separator (for numbers)	false
Thousands separator	" "
Decimal separator	"."
Extract lines at random	false
Number of lines	10
Trim all columns	false
Check column to trim	[[{SCHEMA_COLUMN=ID, TRIM=false}, {SCHEMA_COLUMN=Name, TRIM=false}, {SCHEMA_COLUMN=CountryCode, TRIM=false}, {SCHEMA_COLUMN=District, TRIM=false}, {SCHEMA_COLUMN=Population, TRIM=false}]]
Check each row structure against schema	false
Check date	false
Encoding	"UTF-8"
Split row before field	false

Permit hexadecimal (0xNNN) or octal (0NNNN) for numeric types	false
Decode table	[[{SCHEMA_COLUMN=ID, DECODE=false}, {SCHEMA_COLUMN=Name, DECODE=false}, {SCHEMA_COLUMN=CountryCode, DECODE=false}, {SCHEMA_COLUMN=District, DECODE=false}, {SCHEMA_COLUMN=Population, DECODE=false}]]
!!!DESTINATION.NAME!!!	
Min column number of optimize code	100
Label format	City
Hint format	__UNIQUE_NAME__ __COMMENT__
Connection format	row
Show Information	false
Comment	
Use an existing validation rule	false
Validation Rule Type	

Schema for metadata :

Column	Key	Type	Length	Precision	Nullable	Comment
ID	false	Integer	2		true	
Name	false	String	19		true	
CountryCode	false	String	3		true	
District	false	String	15		true	
Population	false	Integer	7		true	

Original Function Parameters:
Component tFixedFlowInput

	UNIQUE NAME	tFixedFlowInput_1	INPUT(S)	none
	LABEL	__UNIQUE_NAME__	OUTPUT(S)	tFuzzyMatch_1

Component Parameters:


Properties	Values
Unique Name	tFixedFlowInput_1
Component Name	tFixedFlowInput
Version	0.102 (ALPHA)
Family	Misc
Start	false
Startable	true
SUBTREE_START	false
END_OF_FLOW	false
Activate	true
DUMMY	false
tStatCatcher Statistics	false
Help	org.talend.help.tFixedFlowInput
Update components	true
IREPORT_PATH	
JAVA_LIBRARY_PATH	C:\Talend-Studio-20181026_1147-V7.1.1\configuration\lib\java
Subjob color	
Title color	

Enable parallel execution	false
Number of parallel executions	2
PARALLELIZE_KEEP_EMPTY	false
REPOSITORY_ALLOW_AUTO_SWITCH	true
Schema	Built-In
Number of rows	1
Use Single Table	false
Values	[[SCHEMA_COLUMN=CityName, VALUE=]]
Use Inline Table	false
Inline Table	[[CityName="newLine"], {CityName="newLine"}, {CityName="newLine"}, {CityName="newLine"}]
Use Inline Content(delimited file)	true
Row Separator	"\n"
Field Separator	"", ,
Content	Kabul Qandahar Kandahar Herat Mazar-e-Sharif Mazare-Sharif Amsterdam Rotterdam Roterdaam Haag Utrecht Breda Apeldoorn Nijmegen
Label format	__UNIQUE_NAME__
Hint format	__UNIQUE_NAME__ __COMMENT__
Connection format	row
Show Information	false
Comment	
Use an existing validation rule	false
Validation Rule Type	

Schema for tFixedFlowInput_1 :

Column	Key	Type	Length	Precision	Nullable	Comment
CityName	false	String	50		true	

Original Function Parameters:
Component tFuzzyMatch

	UNIQUE NAME	tFuzzyMatch_1	INPUT(S)	tFuzzyMatch_1 , tFixedFlowInput_1
	LABEL	__UNIQUE_NAME__	OUTPUT(S)	tLogRow_1

Component Parameters:

Properties	Values
Unique Name	tFuzzyMatch_1
Component Name	tFuzzyMatch
Version	0.103 (ALPHA)
Family	Data_Quality/Matching
Startable	false
SUBTREE_START	false
END_OF_FLOW	false
Activate	true
DUMMY	false
tStatCatcher Statistics	false
Help	org.talend.help.tFuzzyMatch
Update components	true
IREPORT_PATH	


JAVA_LIBRARY_PATH	C:\Talend-Studio-20181026_1147-V7.1.1\configuration\lib\java
Subjob color	
Title color	
Enable parallel execution	false
Number of parallel executions	2
PARALLELIZE_KEEP_EMPTY	false
REPOSITORY_ALLOW_AUTO_SWITCH	TRUE
Schema	Built-In
Matching type	matchDoubleMetaphone
Case Sensitive	false
Look Up	CityName
Min. distance	0
Max. distance	3
Matching column	Name
Unique matching	false
Matching item separator	","
Label format	__UNIQUE_NAME__
Hint format	__UNIQUE_NAME__ __COMMENT__
Connection format	row
Show Information	false
Comment	
Use an existing validation rule	false

Validation Rule Type	
----------------------	--

Schema for metadata :

Column	Key	Type	Length	Precision	Nullable	Comment
ID	false	Integer	2		true	
Name	false	String	19		true	
CountryCode	false	String	3		true	
District	false	String	15		true	
Population	false	Integer	7		true	
VALUE	false	String	255		true	
MATCHING	false	String	255		true	

Original Function Parameters:
Component tLogRow

	UNIQUE NAME	tLogRow_1	INPUT(S)	tFuzzyMatch_1, tFixedFlowInput_1
	LABEL	__UNIQUE_NAME__	OUTPUT(S)	none

Component Parameters:

Properties	Values
Unique Name	tLogRow_1
Component Name	tLogRow
Version	0.101 (ALPHA)

Family	Logs_Errors
Startable	false
SUBTREE_START	false
END_OF_FLOW	true
Activate	true
DUMMY	false
tStatCatcher Statistics	false
Help	org.talend.help.tLogRow
Update components	true
IREPORT_PATH	
JAVA_LIBRARY_PATH	C:\Talend-Studio-20181026_1147-V7.1.1\configuration\lib\java
Subjob color	
Title color	
Enable parallel execution	false
Number of parallel executions	2
PARALLELIZE_KEEP_EMPTY	false
REPOSITORY_ALLOW_AUTO_SWITCH	true
Schema	Built-In
Basic	false
Table (print values in cells of a table)	true
Vertical (each row is a key/value list)	false
Print unique name	true

Print label	false
Print unique name and label	false
Field Separator	" "
Print header	true
Print component unique name in front of each output row	false
Print schema column name in front of each value	false
Use fixed length for values	false
Lengths	[{SCHEMA_COLUMN=ID, LENGTH=10}, {SCHEMA_COLUMN=Name, LENGTH=10}, {SCHEMA_COLUMN=CountryCode, LENGTH=10}, {SCHEMA_COLUMN=District, LENGTH=10}, {SCHEMA_COLUMN=Population, LENGTH=10}, {SCHEMA_COLUMN=VALUE, LENGTH=10}, {SCHEMA_COLUMN=MATCHING, LENGTH=10}]
Print content with log4j	true
Min column number of optimize code	100
Label format	__UNIQUE_NAME__
Hint format	__UNIQUE_NAME__ __COMMENT__
Connection format	row
Show Information	false
Comment	
Use an existing validation rule	false
Validation Rule Type	

Schema for metadata :

Column	Key	Type	Length	Precision	Nullable	Comment
ID	false	Integer	2		true	

Name	false	String	19		true	
CountryCode	false	String	3		true	
District	false	String	15		true	
Population	false	Integer	7		true	
VALUE	false	String	255		true	
MATCHING	false	String	255		true	

Original Function Parameters: