



Jobs

Generated by Talend Real-time Big Data Platform

Project Name	DataPreparation	GENERATION DATE	27 Jul 2023, 1:13:33 PM
AUTHOR	Dael, Glenn Marvin	Talend Real-time VERSION	8.0.1.20211109_1610

Summary

[Project Description](#)

[Description](#)

[Preview Picture](#)

[Settings](#)

[Context List](#)

[Component List](#)

[Components Description](#)

Project Description

Properties	Values
Name	DataPreparation
Language	java

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

Description

Properties	Values
Name	Data_Stewardship
Author	Dael, Glenn Marvin
Version	0.1
Purpose	
Status	
Description	
Creation	27 Jul 2023, 7:11:25 AM
Modification	27 Jul 2023, 11:40:10 AM

Preview Picture

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
tDatasetOutput_1	tDatasetOutput
tFileInputExcel_1	tFileInputExcel

tFileInputExcel_2	tFileInputExcel
tHashInput_1	tHashInput
tHashOutput_1	tHashOutput
tMap_1	tMap

Components Description

Component tDatasetOutput

	UNIQUE NAME	tDatasetOutput_1	INPUT(S)	tHashInput_1
	LABEL	__UNIQUE_NAME__	OUTPUT(S)	none

Component Parameters:

Properties	Values
Unique Name	tDatasetOutput_1
Component Name	tDatasetOutput
Family	Talend Data Preparation
Startable	false
SUBTREE_START	false
END_OF_FLOW	true
Activate	true
Help	HELP
Update components	true


IREPORT_PATH	
Property Type	Built-In
Query Type	Built-In
tStatCatcher Statistics	false
Enable parallel execution	false
Number of parallel executions	2
PARALLELIZE_KEEP_EMPTY	false
Schema	{"type":"record","name":"tDatasetOutput_1","fields":[{"name":"UserId","type":["string","null"],"di.table
Url	"http://10.10.10.109:9999"
Email	"g.dael@mtc.local"
Password	"Justdoitnow123!"
Mode	CreateOrUpdate
Dataset Name	"Personal-Test"
Dataset Name	""
Fetch Schema	
Limit	10000
Dataset ID	""
REPOSITORY_ALLOW_AUTO_SWITCH	True
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 tDatasetOutput_1 :

Column	Key	Type	Length	Precision	Nullable	Comment
UserId	false	String			true	
FirstName	false	String			true	
LastName	false	String			true	
Sex	false	String			true	
Phone	false	String			true	
DateOfBirth	false	String			true	
JobTitle	false	String			true	
Email	false	String			true	

Original Function Parameters:
Component tFileInputExcel

	UNIQUE NAME	tFileInputExcel_1	INPUT(S)	none
	LABEL	Email	OUTPUT(S)	tMap_1

Component Parameters:

Properties	Values
------------	--------

Unique Name	tFileInputExcel_1
Component Name	tFileInputExcel
Version	0.102 (ALPHA)
Family	File/Input
Start	false
Startable	true
SUBTREE_START	false
END_OF_FLOW	false
Activate	true
DUMMY	false
tStatCatcher Statistics	false
Help	org.talend.help.tFileInputExcel
Update components	true
IREPORT_PATH	
JAVA_LIBRARY_PATH	C:\Program Files (x86)\Talend-Studio\studio\configuration\lib\java
Subjob color	
Title color	
Enable parallel execution	false
Number of parallel executions	2
PARALLELIZE_KEEP_EMPTY	false
Property Type	Built-In
Read excel2007 file format (xlsx)	true


File name/Stream	"D:/Data Integration/Talend/Training Materials/Sample Data/Sample.xlsx"
Password	**
All sheets	false
Sheet list	[[SHEETNAME="Email", USE_REGEX=false]]
Header	1
Footer	0
Limit	
Affect each sheet(header&&footer)	false
First column	1
Last column	
Die on error	false
REPOSITORY_ALLOW_AUTO_SWITCH	True
Schema	Built-In
Schema	Built-In
Advanced separator (for numbers)	false
Thousands separator	" "
Decimal separator	"."
Trim all columns	false
Check column to trim	[[SCHEMA_COLUMN=UserId, TRIM=false], {SCHEMA_COLUMN=Email, TRIM=false}]
Convert date column to string	false
Check need convert date column	[[SCHEMA_COLUMN=UserId, CONVERTDATE=false, PATTERN="MM-dd-yyyy"], {SCHEMA_COLUMN=Email, CONVERTDATE=false, PATTERN="MM-dd-yyyy"}]
Encoding	"ISO-8859-15"

Read real values for numbers	false
Stop reading on encountering empty rows	false
Don't validate the cells	false
Ignore the warning	false
Generation mode	USER_MODE
Include phonetic runs	true
Set minimum inflation rate	false
INFLATION_RATIO	
Label format	Email
Hint format	__UNIQUE_NAME__ __COMMENT__
Connection format	row
Show Information	false
Comment	
Use an existing validation rule	false
Validation Rule Type	

Schema for tFileInputExcel_1 :

Column	Key	Type	Length	Precision	Nullable	Comment
UserId	false	String			true	
Email	false	String			true	

Original Function Parameters:
Component tFileInputExcel

	UNIQUE NAME	tFileInputExcel_2	INPUT(S)	none
	LABEL	People	OUTPUT(S)	tMap_1 , tHashInput_1

Component Parameters:

Properties	Values
Unique Name	tFileInputExcel_2
Component Name	tFileInputExcel
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.tFileInputExcel
Update components	true
IREPORT_PATH	
JAVA_LIBRARY_PATH	C:\Program Files (x86)\Talend-Studio\studio\configuration\lib\java
Subjob color	
Title color	

Enable parallel execution	false
Number of parallel executions	2
PARALLELIZE_KEEP_EMPTY	false
Property Type	Built-In
Read excel2007 file format (xlsx)	true
File name/Stream	"D:/Data Integration/Talend/Training Materials/Sample Data/Sample.xlsx"
Password	**
All sheets	false
Sheet list	[{SHEETNAME="People", USE_REGEX=false}]
Header	1
Footer	0
Limit	
Affect each sheet(header&&footer)	false
First column	1
Last column	
Die on error	false
REPOSITORY_ALLOW_AUTO_SWITCH	true
Schema	Built-In
Schema	Built-In
Advanced separator (for numbers)	false
Thousands separator	" " ,
Decimal separator	" " .


Trim all columns	false
Check column to trim	[[{SCHEMA_COLUMN=Index, TRIM=false}, {SCHEMA_COLUMN=UserId, TRIM=false}, {SCHEMA_COLUMN=FirstName, TRIM=false}, {SCHEMA_COLUMN=LastName, TRIM=false}, {SCHEMA_COLUMN=Sex, TRIM=false}, {SCHEMA_COLUMN=Phone, TRIM=false}, {SCHEMA_COLUMN=DateOfBirth, TRIM=false}, {SCHEMA_COLUMN=JobTitle, TRIM=false}]]
Convert date column to string	false
Check need convert date column	[[{SCHEMA_COLUMN=Index, CONVERTDATE=false, PATTERN="MM-dd-yyyy"}, {SCHEMA_COLUMN=UserId, CONVERTDATE=false, PATTERN="MM-dd-yyyy"}, {SCHEMA_COLUMN=FirstName, CONVERTDATE=false, PATTERN="MM-dd-yyyy"}, {SCHEMA_COLUMN=LastName, CONVERTDATE=false, PATTERN="MM-dd-yyyy"}, {SCHEMA_COLUMN=Sex, CONVERTDATE=false, PATTERN="MM-dd-yyyy"}, {SCHEMA_COLUMN=Phone, CONVERTDATE=false, PATTERN="MM-dd-yyyy"}, {SCHEMA_COLUMN=DateOfBirth, CONVERTDATE=false, PATTERN="MM-dd-yyyy"}, {SCHEMA_COLUMN=JobTitle, CONVERTDATE=false, PATTERN="MM-dd-yyyy"}]]
Encoding	"ISO-8859-15"
Read real values for numbers	false
Stop reading on encountering empty rows	false
Don't validate the cells	false
Ignore the warning	false
Generation mode	USER_MODE
Include phonetic runs	true
Set minimum inflation rate	false
INFLATION_RATIO	
Label format	People
Hint format	__UNIQUE_NAME__ __COMMENT__
Connection format	row
Show Information	false
Comment	

Use an existing validation rule	false
Validation Rule Type	

Schema for tFileInputExcel_2 :

Column	Key	Type	Length	Precision	Nullable	Comment
Index	false	String			true	
UserId	false	String			true	
FirstName	false	String			true	
LastName	false	String			true	
Sex	false	String			true	
Phone	false	String			true	
DateOfBirth	false	String			true	
JobTitle	false	String			true	

Original Function Parameters:
Component tHashInput

	UNIQUE NAME	tHashInput_1	INPUT(S)	tFileInputExcel_2
	LABEL	__UNIQUE_NAME__	OUTPUT(S)	tDatasetOutput_1

Component Parameters:

Properties	Values
Unique Name	tHashInput_1

Component Name	tHashInput
Version	0.102 (ALPHA)
Family	Technical
Start	false
Startable	true
SUBTREE_START	true
END_OF_FLOW	false
Activate	true
DUMMY	false
tStatCatcher Statistics	false
Help	org.talend.help.Hash
Update components	true
IREPORT_PATH	
JAVA_LIBRARY_PATH	C:\Program Files (x86)\Talend-Studio\studio\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
Link with a tHashOutput	true


Component list	tHashOutput_1
Data from file	"C:/Users/Glenn/workspace/in.hash"
Clear cache after reading	false
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 tHashInput_1 :

Column	Key	Type	Length	Precision	Nullable	Comment
UserId	false	String			true	
FirstName	false	String			true	
LastName	false	String			true	
Sex	false	String			true	
Phone	false	String			true	
DateOfBirth	false	String			true	
JobTitle	false	String			true	
Email	false	String			true	

Original Function Parameters:

Component tHashOutput

	UNIQUE NAME	tHashOutput_1	INPUT(S)	tMap_1
	LABEL	__UNIQUE_NAME__	OUTPUT(S)	none

Component Parameters:

Properties	Values
Unique Name	tHashOutput_1
Component Name	tHashOutput
Version	0.101 (ALPHA)
Family	Technical
Startable	false
SUBTREE_START	false
END_OF_FLOW	true
Activate	true
DUMMY	false
tStatCatcher Statistics	false
Help	org.talend.help.Hash
Update components	true
IREPORT_PATH	
JAVA_LIBRARY_PATH	C:\Program Files (x86)\Talend-Studio\studio\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
Link with a tHashOutput	false
Component list	
Data write model	MEMORY
File to persist data	"C:/Users/Glenn/workspace/in.hash"
Max memory heap size(MB)	2
Keys management	KEEP_ALL
Append	true
hash key from input connector	false
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 tHashOutput_1 :

Column	Key	Type	Length	Precision	Nullable	Comment
--------	-----	------	--------	-----------	----------	---------

UserId	false	String			true	
FirstName	false	String			true	
LastName	false	String			true	
Sex	false	String			true	
Phone	false	String			true	
DateOfBirth	false	String			true	
JobTitle	false	String			true	
Email	false	String			true	

Original Function Parameters:
Component tMap

	UNIQUE NAME	tMap_1	INPUT(S)	tFileInputExcel_2 , tFileInputExcel_1
	LABEL	__UNIQUE_NAME__	OUTPUT(S)	tHashOutput_1

Component Parameters:

Properties	Values
tStatCatcher Statistics	false
Mapping links display as:	AUTO
Temp data directory path:	

Max buffer size (nb of rows):	2000000
Ignore trailing zeros for BigDecimal	true
Show Information	false
Comment	
Use an existing validation rule	false

Mapper table for tMap_1 (input):
Mapper table Properties(row1):

Properties	Values
Name	row1
Matching-mode	UNIQUE_MATCH
isMinimized	false
isReject	false
isRejectInnerJoin	false
isInnerJoin	false
expressionFilter	null

Metadata Table Entries(row1):

Name	Type	Expression	isNullable
Index	String		true
UserId	String		true
FirstName	String		true
LastName	String		true
Sex	String		true
Phone	String		true
DateOfBirth	String		true
JobTitle	String		true

Constraint Table Entries(row1):

Name	Type	Expression	isNullable
------	------	------------	------------

Mapper table Properties(row2):

Properties	Values
Name	row2
Matching-mode	UNIQUE_MATCH
isMinimized	false
isReject	false
isRejectInnerJoin	false
isInnerJoin	false
expressionFilter	null

Metadata Table Entries(row2):

Name	Type	Expression	isNullable
UserId	String	row1.UserId	true
Email	String		true

Constraint Table Entries(row2):

Name	Type	Expression	isNullable
------	------	------------	------------

Mapper table for tMap_1 (output):**Mapper table Properties(People_Email):**

Properties	Values
Name	People_Email
Matching-mode	
isMinimized	false
isReject	false
isRejectInnerJoin	false
isInnerJoin	false
expressionFilter	null

Metadata Table Entries(People_Email):

Name	Type	Expression	isNullable
UserId	String	row1.FirstName	true
FirstName	String	row1.FirstName	true
LastName	String	row1.LastName	true
Sex	String	row1.Sex	true
Phone	String	row1.Phone	true
DateOfBirth	String	row1.DateOfBirth	true
JobTitle	String	row1.JobTitle	true
Email	String	row2.Email	true

Constraint Table Entries(People_Email):

Name	Type	Expression	isNullable
------	------	------------	------------

Mapper table for tMap_1 (var):

Mapper table Properties(Var):

Properties	Values
Name	Var
Matching-mode	
isMinimized	true
isReject	false
isRejectInnerJoin	false
isInnerJoin	false
expressionFilter	null

Metadata Table Entries(Var):

Name	Type	Expression	isNullable
------	------	------------	------------

Constraint Table Entries(Var):

Name	Type	Expression	isNullable
------	------	------------	------------