talend

Jobs

Generated by Talend Cloud Data Management Platform

Project Name	Data-cleaning	GENERATION DATE	12/04/2022, 14:50:20
AUTHOR		Talend Cloud VERSION	8.0.1.20220317_0634-patch

Summary

Project Description
Description
Preview Picture
Settings
Context List
Component List
Components Description

Project Description

Properties	Values
Name	Data-cleaning
Language	java

Description	Data cleaning and enrichment

Description

Properties	Values
Name	Teste_CNI
Author	
Version	0.1
Purpose	Teste_CNI
Status	
Description	Teste_CNI
Creation	11/04/2022, 19:25:23
Modification	11/04/2022, 19:47:57

Preview Picture
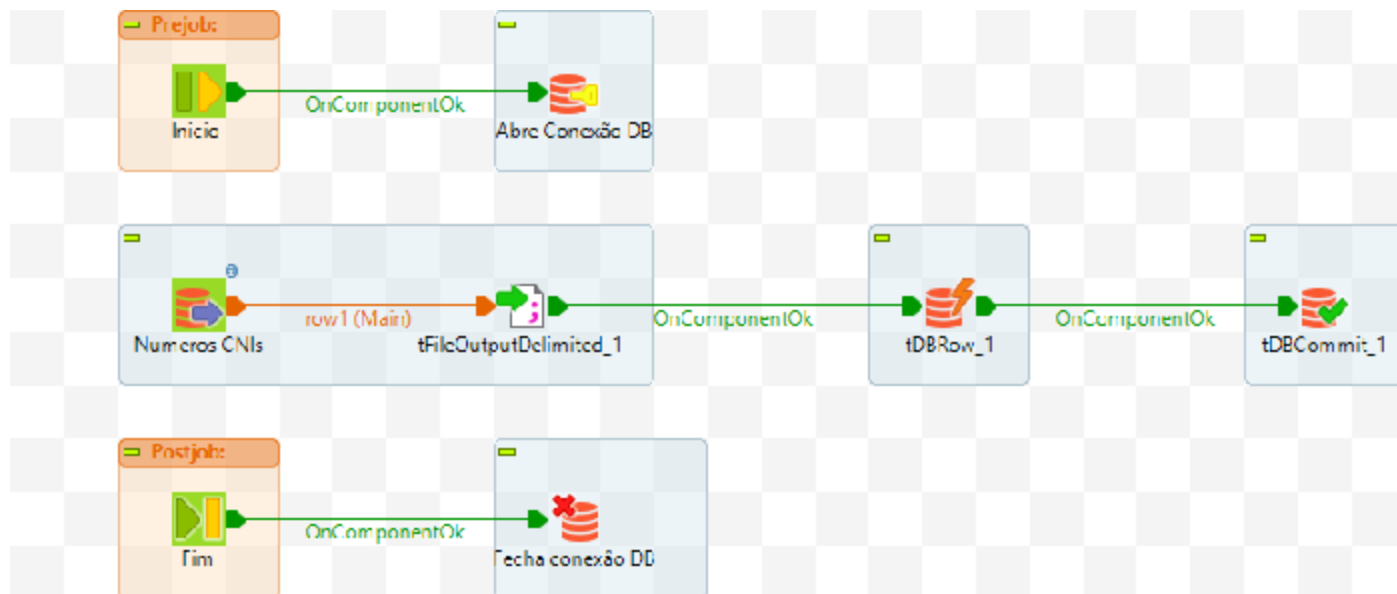
Settings

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
tDBClose_1	tMSSqlClose
tDBCommit_1	tMSSqlCommit
tDBConnection_1	tMSSqlConnection

tDBInput_1	tMSSqlInput
tDBRow_1	tMSSqlRow
tFileOutputDelimited_1	tFileOutputDelimited
tPostjob_1	tPostjob
tPrejob_1	tPrejob

Components Description

Component   tMSSqlClose

	UNIQUE NAME	tDBClose_1	INPUT(S)	tPostjob_1
	LABEL	Fecha conexão DB	OUTPUT(S)	none

Component Parameters:

Properties	Values
Database	tMSSqlClose
Unique Name	tDBClose_1
Component Name	tMSSqlClose
Version	0.102 (ALPHA)
Family	Databases/DB Specifics/MS_SQL_Server
Start	false
Startable	true
SUBTREE_START	true

END_OF_FLOW	true
Activate	true
DUMMY	false
tStatCatcher Statistics	false
Help	org.talend.help.tMSSqlClose
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
Component List	tDBConnection_1
Label format	Fecha conexão DB
Hint format	__UNIQUE_NAME__ __COMMENT__
Connection format	row
Show Information	false
Comment	
Use an existing validation rule	false
Validation Rule Type	

Original Function Parameters:
Component   tMSSqlCommit

	UNIQUE NAME	tDBCommit_1	INPUT(S)	tDBRow_1
	LABEL	__UNIQUE_NAME__	OUTPUT(S)	none

Component Parameters:

Properties	Values
Database	tMSSqlCommit
Unique Name	tDBCommit_1
Component Name	tMSSqlCommit
Version	0.102 (ALPHA)
Family	Databases/DB Specifics/MS_SQL_Server
Start	false
Startable	true
SUBTREE_START	true
END_OF_FLOW	true
Activate	true
DUMMY	false
tStatCatcher Statistics	false
Help	org.talend.help.tMSSqlCommit
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
Component List	tDBConnection_1
Close Connection	true
Schema	Built-In
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 tDBCommit_1 :

Column	Key	Type	Length	Precision	Nullable	Comment

Original Function Parameters:
Component   tMSSqlConnection

	UNIQUE NAME	tDBConnection_1	INPUT(S)	tPrejob_1
	LABEL	Abre Conexão DB	OUTPUT(S)	none

Component Parameters:

Properties	Values
Database	tMSSqlConnection
Unique Name	tDBConnection_1
Component Name	tMSSqlConnection
Version	0.102 (ALPHA)
Family	Databases/DB Specifics/MS_SQL_Server|ELT/Connections
Start	false
Startable	true
SUBTREE_START	true
END_OF_FLOW	true
Activate	true
DUMMY	false
tStatCatcher Statistics	false
Help	org.talend.help.tMSSqlConnection
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	repository: DB (MSSQL):Cadastro
JDBC Provider	JTDS
Host	"172.18.242.124"
Database Driver	MSSQL
Port	"1433"
Schema	"dbo"
Database	"cadastro"
Username	"CadSQL"
Password	**************
Encoding	"ISO-8859-15"
Additional JDBC Parameters	""
Use or register a shared DB Connection	false
Shared DB Connection Name	
Specify a data source alias	false
Data source alias	""
Authenticate using Azure Active Directory	false
Auto Commit	false
Share identity insert in multi table	false
Label format	Abre Conexão DB
Hint format	__UNIQUE_NAME__ __COMMENT__
Connection format	row
Show Information	false

Comment	
Use an existing validation rule	false
Validation Rule Type	

Original Function Parameters:
Component   tMSSqlInput

	UNIQUE NAME	tDBInput_1	INPUT(S)	none
	LABEL	Numeros CNIs	OUTPUT(S)	tFileOutputDelimited_1

Component Parameters:

Properties	Values
Database	tMSSqlInput
Unique Name	tDBInput_1
Component Name	tMSSqlInput
Version	0.102 (ALPHA)
Family	Databases/DB Specifics/MS_SQL_Server
Start	true
Startable	true
SUBTREE_START	true
END_OF_FLOW	false
Activate	true
DUMMY	false
tStatCatcher Statistics	false

Help	org.talend.help.tMSSqlInput
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
Use an existing connection	true
Component List	tDBConnection_1
Property Type	Built-In
Host	""
JDBC Provider	MSSQL_PROP
Database Driver	MSSQL
Port	"1433"
Schema	""
Database	""
Username	""
Password	**
REPOSITORY_ALLOW_AUTO_SWITCH	true
Schema	Built-In

Table Name	"SYS_DET_CLIENTES"
Query Type	Built-In
Guess Schema	""
Query	"SELECT NAME, TIPO, VALUEDOCS, tipo_docs, 'CNI' as New_Tipo from SYS_DET_CLIENTES where len(ltrim(rtrim(valuedocs))) = 13 and isnumeric(substring(ltrim(rtrim(valuedocs)), 1, 8)) = 1 and isnumeric(substring(ltrim(rtrim(valuedocs)), 9, 1)) = 0 and isnumeric(substring(ltrim(rtrim(valuedocs)), 10, 3)) = 1 and isnumeric(substring(ltrim(rtrim(valuedocs)), 13, 1)) = 0 and substring(ltrim(rtrim(valuedocs)), 5, 2) < 13 and substring(ltrim(rtrim(valuedocs)), 7, 2) < 32 and substring(ltrim(rtrim(valuedocs)), 1, 1) < 3 and tipo_docs != 'CNI' and name = 'Jocilene Medina Monteiro'"
Specify a data source alias	false
Data source alias	""
Mapping	id_MSSQL
Additional JDBC Parameters	"noDatetimeStringSync=true"
Authenticate using Azure Active Directory	false
Encoding	"ISO-8859-15"
Trim all the String/Char columns	false
Trim column	[{SCHEMA_COLUMN=NAME, TRIM=false}, {SCHEMA_COLUMN=TIPO, TRIM=false}, {SCHEMA_COLUMN=VALUEDOCS, TRIM=false}, {SCHEMA_COLUMN=tipo_docs, TRIM=false}, {SCHEMA_COLUMN=New_Tipo, TRIM=false}]
Set Query Timeout	false
Timeout (seconds)	30
Label format	Numeros CNIs
Hint format	__UNIQUE_NAME__ __COMMENT__
Connection format	row
Show Information	true

Comment	Universo de registos que vai ser tratado (numero CNI, tipo de documento diferente de CNI)	
Use an existing validation rule	false	
Validation Rule Type		

Schema for tDBInput_1 :

Column	Key	Type	Length	Precision	Nullable	Comment
NAME	false	String	2000		true	
TIPO	false	String	200		true	
VALUEDOCS	false	String	250		true	
tipo_docs	false	String	240		true	
New_Tipo	false	String	3		false	

Original Function Parameters:
Component tMSSqlRow

	UNIQUE NAME	tDBRow_1	INPUT(S)	tFileOutputDelimited_1
	LABEL	__UNIQUE_NAME__	OUTPUT(S)	tDBCommit_1

Component Parameters:

Properties	Values
Database	tMSSqlRow
Unique Name	tDBRow_1
Component Name	tMSSqlRow
Version	0.102 (ALPHA)

Family	Databases/DB Specifics/MS_SQL_Server
Start	false
Startable	true
SUBTREE_START	true
END_OF_FLOW	true
Activate	true
DUMMY	false
tStatCatcher Statistics	false
Help	org.talend.help.tMSSqlRow
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
Use an existing connection	true
Component List	tDBConnection_1
Property Type	Built-In
JDBC Provider	MSSQL_PROP
Host	""

Port	"1433"
Schema	""
Database	""
Database Driver	MSSQL
Username	""
Password	**
REPOSITORY_ALLOW_AUTO_SWITCH	true
Schema	Built-In
Reject Schema	Built-In
Table Name	""
Turn on identity insert	false
Query Type	Built-In
Query	"UPDATE SYS_DET_CLIENTES set tipo_docs = 'CNI' where len(ltrim(rtrim(valuedocs))) = 13 and isnumeric(substring(ltrim(rtrim(valuedocs)), 1, 8)) = 1 and isnumeric(substring(ltrim(rtrim(valuedocs)), 9, 1)) = 0 and isnumeric(substring(ltrim(rtrim(valuedocs)), 10, 3)) = 1 and isnumeric(substring(ltrim(rtrim(valuedocs)), 13, 1)) = 0 and substring(ltrim(rtrim(valuedocs)), 5, 2) < 13 and substring(ltrim(rtrim(valuedocs)), 7, 2) < 32 and substring(ltrim(rtrim(valuedocs)), 1, 1) < 3 and tipo_docs != 'CNI' and name = 'Jocilene Medina Monteiro'"
Specify a data source alias	false
Data source alias	""
Die on error	false
Additional JDBC Parameters	""
Authenticate using Azure Active Directory	false
Propagate QUERY's recordset	false
use column	

Use PreparedStatement	false
Set PreparedStatement Parameters	[]
Encoding	"ISO-8859-15"
Commit every	10000
Set Query Timeout	false
Timeout (seconds)	30
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 tDBRow_1 :

Column	Key	Type	Length	Precision	Nullable	Comment

Original Function Parameters:
Component   tFileOutputDelimited

	UNIQUE NAME	tFileOutputDelimited_1	INPUT(S)	tDBInput_1
	LABEL	__UNIQUE_NAME__	OUTPUT(S)	tDBRow_1

Component Parameters:

Properties	Values

Unique Name	tFileOutputDelimited_1
Component Name	tFileOutputDelimited
Version	0.101 (ALPHA)
Family	File/Output
Startable	false
SUBTREE_START	false
END_OF_FLOW	true
Activate	true
DUMMY	false
tStatCatcher Statistics	false
Help	org.talend.help.tFileOutputDelimited
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
Use Output Stream	false
Output Stream	outputStream

File Name	"C:/Users/Rui/workspace/talend/Numeros_CNI_out.csv"
Row Separator	"\n"
Use OS line separator as row separator when CSV Row Separator is set to CR,LF or CRLF.	true
CSV Row Separator	"\n"
Field Separator	";"
Append	false
Include Header	false
Compress as zip file	false
REPOSITORY_ALLOW_AUTO_SWITCH	true
Schema	Built-In
Advanced separator (for numbers)	false
Thousands separator	","
Decimal separator	"."
CSV options	false
Escape char	"""
Text enclosure	"""
Create directory if does not exist	true
Split output in several files	false
Rows in each output file	1000
Custom the flush buffer size	false
Row number	1
Output in row mode	false

Encoding	"ISO-8859-15"
Don't generate empty file	false
Throw an error if the file already exist	true
Min column number of optimize code	90
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 tFileOutputDelimited_1 :

Column	Key	Type	Length	Precision	Nullable	Comment
NAME	false	String	2000		true	
TIPO	false	String	200		true	
VALUEDOCS	false	String	250		true	
tipo_docs	false	String	240		true	
New_Tipo	false	String	3		false	

Original Function Parameters:
Component tPostjob

	UNIQUE NAME	tPostjob_1	INPUT(S)	none
	LABEL	Fim	OUTPUT(S)	tDBClose_1

Component Parameters:

Properties	Values
Unique Name	tPostjob_1
Component Name	tPostjob
Version	0.102 (ALPHA)
Family	Orchestration
Start	true
Startable	true
SUBTREE_START	true
END_OF_FLOW	true
Activate	true
DUMMY	false
tStatCatcher Statistics	false
Help	!!!HELP!!!
Update components	true
IREPORT_PATH	
JAVA_LIBRARY_PATH	C:\Program Files (x86)\Talend-Studio\studio\configuration\lib\java
Subjob color	255;220;180
Title color	230;100;0
Enable parallel execution	false

Number of parallel executions	2
PARALLELIZE_KEEP_EMPTY	false
Label format	Fim
Hint format	__UNIQUE_NAME__ __COMMENT__
Connection format	row
Show Information	false
Comment	
Use an existing validation rule	false
Validation Rule Type	

Original Function Parameters:
Component   tPrejob

	UNIQUE NAME	tPrejob_1	INPUT(S)	none
	LABEL	Inicio	OUTPUT(S)	tDBConnection_1

Component Parameters:

Properties	Values
Unique Name	tPrejob_1
Component Name	tPrejob
Version	0.102 (ALPHA)
Family	Orchestration
Start	true
Startable	true

SUBTREE_START	true
END_OF_FLOW	true
Activate	true
DUMMY	false
tStatCatcher Statistics	false
Help	org.talend.help.tPrejob
Update components	true
IREPORT_PATH	
JAVA_LIBRARY_PATH	C:\Program Files (x86)\Talend-Studio\studio\configuration\lib\java
Subjob color	255;220;180
Title color	230;100;0
Enable parallel execution	false
Number of parallel executions	2
PARALLELIZE_KEEP_EMPTY	false
Label format	Inicio
Hint format	__UNIQUE_NAME__ __COMMENT__
Connection format	row
Show Information	false
Comment	
Use an existing validation rule	false
Validation Rule Type	

Original Function Parameters: