

TP1 - CSV File Reader

1. créer un nouveau job nommé "TP2".
2. Créer une nouvelle métadonnée de type "File delimited"

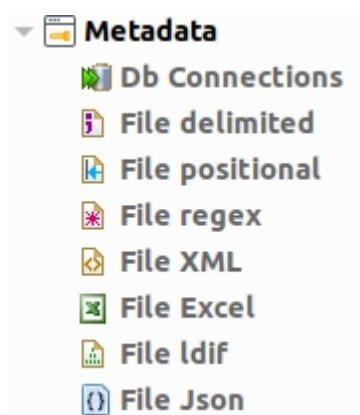


Image 1

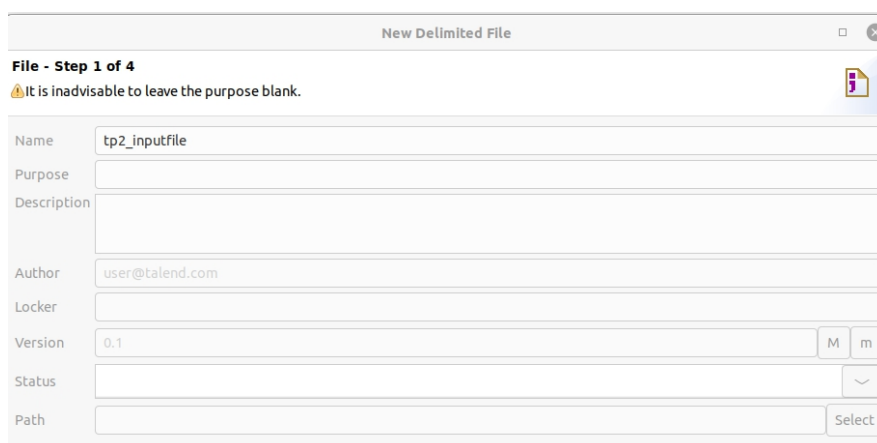
A screenshot of the 'New Delimited File' dialog box in Talend. The dialog is titled 'New Delimited File' and shows 'File - Step 1 of 4'. A warning icon and text state: 'It is inadvisable to leave the purpose blank.' The form contains the following fields: 'Name' (tp2_inputfile), 'Purpose' (empty), 'Description' (empty), 'Author' (user@talend.com), 'Locker' (empty), 'Version' (0.1) with 'M' and 'm' radio buttons, 'Status' (empty) with a dropdown arrow, and 'Path' (empty) with a 'Select' button.

Image 2

- Sélectionnez le fichier csv
- Ajouter la configuration suivante pour lire le fichier :

File - Step 3 of 4
Add a Metadata File on repository
Define the setting of the parse job

File Settings
 Encoding: **US-ASCII**
 Field Separator: **Comma** Corresponding Character: **","**
 Row Separator: **Standard EOL** Corresponding Character: **"\n"**

Rows To Skip
 If any rows must be ignored, specify the following parameters
 Header: ☒ **1**
 Footer: ☐
☐ Skip empty row

Escape Char Settings
☐ CSV ☒ **Delimited**
 Escape Char: **Empty**
 Text Enclosure: **Empty**
☐ Split row before field

Limit Of Rows
 If the number of lines must be limited, specify this number
 Limit:

Preview/Output
☒ Set heading row as column names Refresh Preview

```

Transaction_date,Product,Price,Payment_Type,Name,City,State,Country,Account_Created,Last_Login,Latitude,Longitude
1/2/09 6:17 ,Product1,1200,Mastercard,carolina,Basildon,England,United Kingdom,1/2/09 6:00,1/2/09 6:08,51.5,-1.1166667
1/2/09 4:53,Product1,1200,Visa,Betina,Parkville ,MO,United States,1/2/09 4:42,1/2/09 7:49,39.195,-94.68194
1/2/09 13:08,Product1,1200,Mastercard,Federica e Andrea,Astoria ,OR,United States,1/1/09 16:21,1/3/09 12:32,46.18806,-123.83
  
```

Export as context Revert Context

< Back Next > Cancel Finish

Image 3

- Ajoutez le format de date correct:

Description of the Schema

Column	Key	Type	<input checked="" type="checkbox"/> Null	Date Pattern (Ctrl+Space available)	Length	Precision	Default	Comment
Transaction_date	<input type="checkbox"/>	Date	<input checked="" type="checkbox"/>	"MM/dd/yyyy HH:mm"	8	0		
Product	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		4	0		
Price	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		10	0		
Payment_Type	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		17	0		
Name	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		28	0		
City	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		16	0		
State	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		14	0		
Country	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>					
Account_Created	<input type="checkbox"/>	Date	<input checked="" type="checkbox"/>	"MM/dd/yyyy HH:mm"		0		

Image 4

- Ajouter un nouveau Context group:

Create / Edit a context group

Step 1 of 2
⚠ It is inadvisable to leave the purpose blank.

Name: **tp2_context**

Purpose:

Description:

Author: **user@talend.com**

Locker:

Version: **0.1** M m

Status: ^

Path: Select

< Back Next > Cancel Finish

Image 5

7. Ajouter une nouvelle variable de type répertoire avec le chemin du dossier “input”

Step 2 of 2
Define the contexts, variables and values

	Name	Type	Comment	Default Value	
1	input_dir	Directory		/tp2/input/	<input type="checkbox"/>

+ -

Default context environment: Default

< Back Next > Cancel Finish

Image 6

8. Ajouter un composant de type “tfileinputDelimited”
9. Dans la vue contextuelle, ajoutez le groupe contextuel créé, en cliquant sur le bouton indiqué sur l'image.

Job(TP2 0.1) Contexts(TP2) Component Run (Job TP2)

	Name	Type	Comment	Default Value	
1	tp2_context (from repository cont				
2	input_dir	Directory			<input type="checkbox"/>

+ - ↑ ↓ [Document Icon]

Default context environment: Default

Image 7

10. Sur la vue des composants, ajoutez la configuration suivante:

Property Type Built-In ⓘ

Schema Repository DELIM:tp2_inputfile - metadata * ☐ Edit schema ☐

"When the input source is a stream or a zip file, footer and random shouldn't be bigger than 0."

File name/Stream context.input_dir+"inputfile.csv" * ...

Row Separator "\n" Field Separator "," *

☐ CSV options

Header 1 Footer 0 Limit

☒ Skip empty rows ☐ Uncompress as zip file ☐ Die on error

Image 8

Notez que :

- ◆ Le schéma a été choisi à partir du métadonnée déjà créées "tp1_inputFile".
- ◆ Nous utilisons la variable contextuelle "input_dir" pour spécifier le chemin du fichier

11. Ajouter un composant de type tLogRow et Tester le job

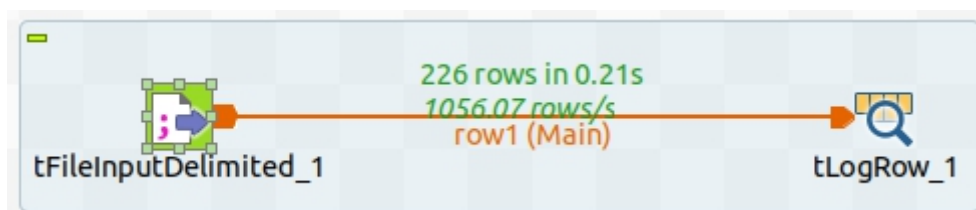


Image 9

Execution

Run

Kill

Clear

Starting job TP2 at 11:06 05/03/2022.

[statistics] connecting to socket on port 3577

[statistics] connected

Transaction_date	Product	Price	Payment_Type	Name	City
01/02/0009 06:17	Product1	1200	Mastercard	carolina	Basildon
01/02/0009 04:53	Product1	1200	Visa	Betina	Parkville
01/02/0009 13:08	Product1	1200	Mastercard	Federica e Andrea	Astoria
01/03/0009 14:44	Product1	1200	Visa	Gouya	Echuca
01/04/0009 12:56	Product2	3600	Visa	Gerd W	Cahaba Heigh
01/04/0009 13:19	Product1	1200	Visa	LAURENCE	Mickleton
01/04/0009 20:11	Product1	1200	Mastercard	Fleur	Peoria
01/02/0009 20:09	Product1	1200	Mastercard	adam	Martin
01/04/0009 13:17	Product1	1200	Mastercard	Renee Elisabeth	Tel Aviv
01/04/0009 14:11	Product1	1200	Visa	Aidan	Chatou
01/05/0009 02:42	Product1	1200	Diners	Stacy	New York
01/05/0009 05:39	Product1	1200	Amex	Heidi	Eindhoven
01/02/0009 09:16	Product1	1200	Mastercard	Sean	Shavano Park
01/05/0009 10:08	Product1	1200	Visa	Georgia	Eagle
01/02/0009 14:18	Product1	1200	Visa	Richard	Riverside
01/04/0009 01:05	Product1	1200	Diners	Leanne	Julianstown
01/05/0009 11:37	Product1	1200	Visa	Janet	Ottawa
01/06/0009 05:02	Product1	1200	Diners	barbara	Hyderabad
01/06/0009 07:45	Product2	3600	Visa	Sabine	London

☐ Line limit

10000

☐ Wrap

Image 10