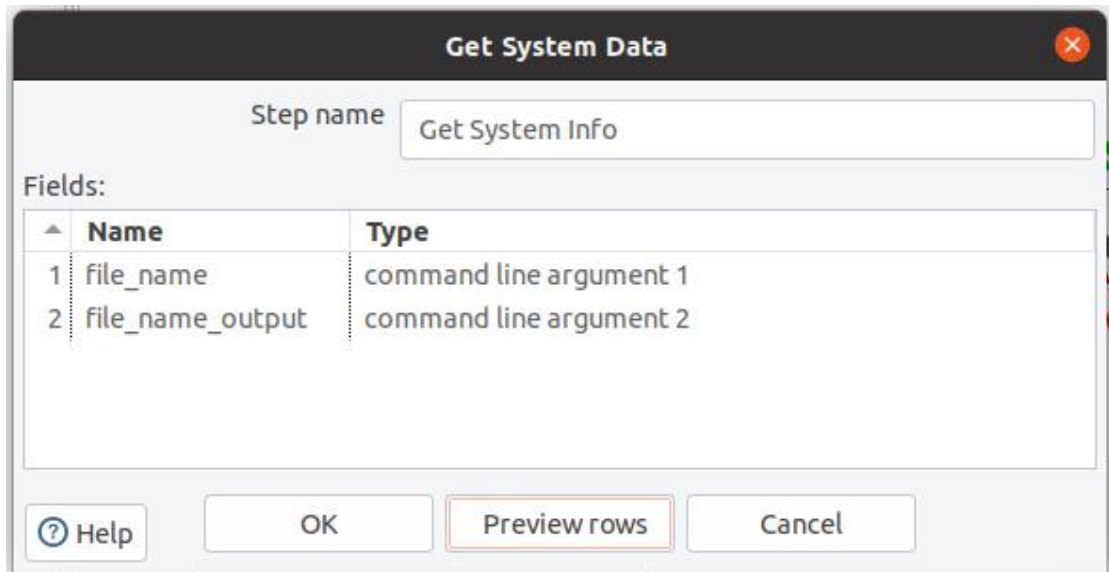


ETL TP : 5 Traitement des erreurs

Le but de ce Tp est de mettre en place une transformation avec des variables et des paramètres et de l'utiliser à partir d'un job.

1. ouvrir la transformation du tp 4
2. ajouter une étape «Get System Info» avec la configuration suivante :



Get System Data

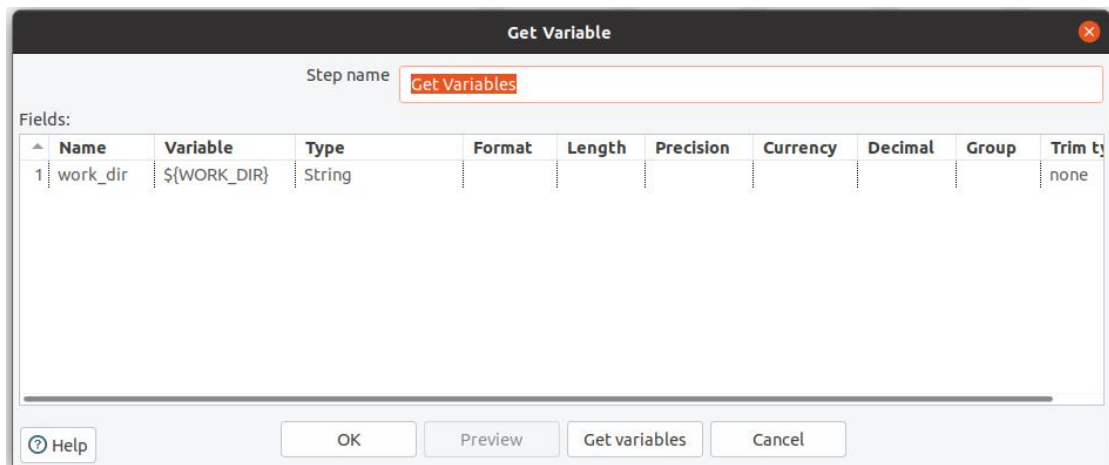
Step name: Get System Info

Fields:

	Name	Type
1	file_name	command line argument 1
2	file_name_output	command line argument 2

Buttons: ? Help, OK, Preview rows, Cancel

3. ajouter une étape «Get Variables» avec la configuration suivante :



Get Variable

Step name: Get Variables

Fields:

	Name	Variable	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1	work_dir	\${WORK_DIR}	String							none

Buttons: ? Help, OK, Preview, Get variables, Cancel

4. ajouter une étape «User Defined Java Expression» avec la configuration suivante :

User Defined Java Expression

Step name:

Fields:

	New field	Java expression	Value type	Length	Precision	Replace va
1	file_path	work_dir+"/"+file_name	String			

5. modifier les paramètres de l'étape «CSV file input» :

CSV Input

Step name:

The filename field (data from previous steps):

Include the filename in the output? ☐

Delimiter:

Enclosure:

NIO buffer size:

Lazy conversion? ☒

Header row present? ☒

Add filename to result? ☐

The row number field name (optional):

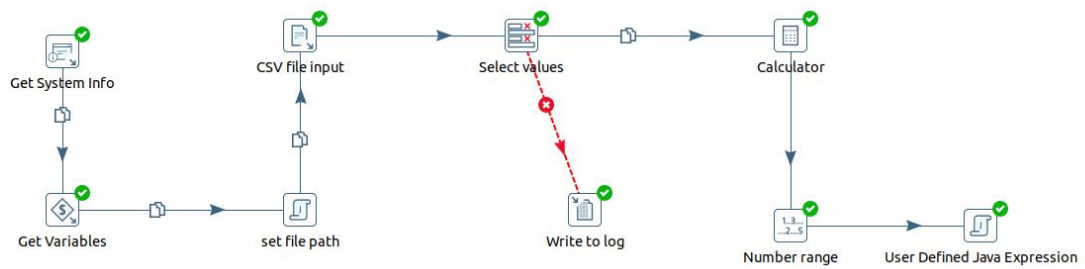
Running in parallel? ☐

New line possible in fields? ☐

File encoding:

	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1	project_name	String							none
2	start_date	String							none
3	end_date	String							none

6. votre configuration doit ressembler à ceci



7. pour les tests, nous devons définir WORK_DIR comme variable et file_name file_name_output comme arguments
8. Lorsque nous lançons une exécution, l'écran suivant s'affiche :

Run Options

Run configuration:

Options

☒ Clear log before running

Log level: Basic

☐ Enable safe mode

☒ Gather performance metrics

Parameters Variables

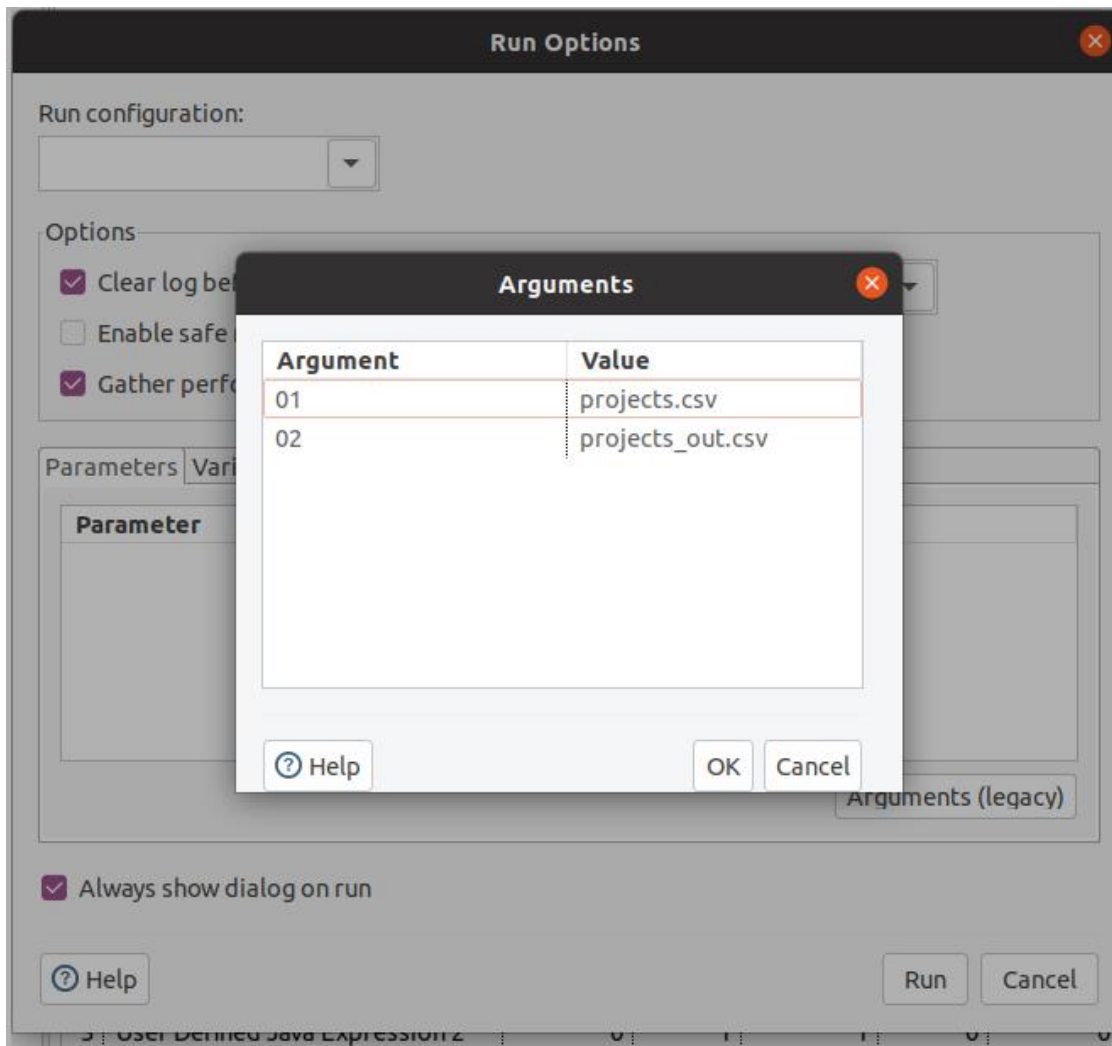
Parameter	Default value	Value	Description

Arguments (legacy)

☒ Always show dialog on run

Help

Run Cancel



9. nous ajoutons WORK_DIR dans l'onglet des variables et les deux arguments après avoir cliqué sur les arguments(legacy)
10. ajouter deux étapes «Get System Info» et «Get Variables» après «Number range» avec la configuration suivante :

Get System Data

Step name:

Fields:

	Name	Type
1	file_name	command line argument 1
2	file_name_output	command line argument 2

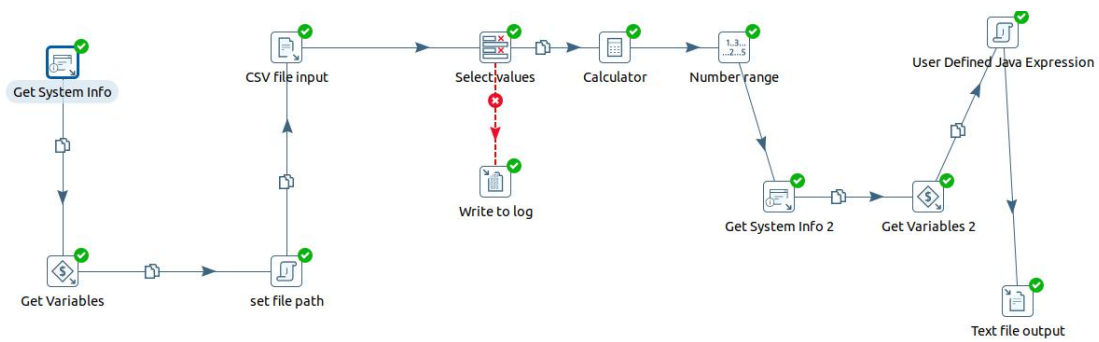
Get Variable

Step name:

Fields:

	Name	Variable	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1	work_dir	\${WORK_DIR}	String							none

11. Changer la config de «User Defined Java Expression» pour ajouter le path de fichier out.
12. Ajouter une etape «Text file output» pour générer un fichier csv avec ces colonnes: project_name,duration,message et performance:
13. corriger la colonne de performance
14. votre configuration doit ressembler à ceci



création de job

Besoin :

Créer un job qui passe plusieurs fichiers à la transformation modifiée de Tp4 et créer un fichier zip contenant tous les fichiers de sortie.

Votre configuration final doit ressembler à ceci :

