

### ETL TP : 3

1. Créer une nouvelle Transformation.
2. Dans le groupe d'étapes Input, faites glisser sur le canevas l'étape "Generate Rows", et configurez-la comme indiqué :

Generate Rows

Step name:

Limit:

Never stop generating rows: ☐

Interval in ms (delay):

Current row time field name:

Previous row time field name:

Fields:

	Name	Type	Format	Length	Precision	Currency	Decimal	Group	Value	Set empty
1	start_date	Date	yyyy-MM-dd						2021-01-20	N
2	end_date	Date	yyyy-MM-dd						2021-02-15	N

Buttons: ? Help, OK, Preview, Cancel

3. Dans la catégorie d'étapes Transformer, ajoutez l'étape "Calculator", et créez un hop qui va de l'étape "Generate Rows" à celle-ci.
4. Double-cliquez sur l'étape "Calculator" et ajoutez le champ nommé diff\_dates comme différence entre end\_date et start\_date. exactement comme vous l'avez fait dans le TP2.
5. Exécutez un aperçu. Vous devriez voir une seule ligne avec trois champs : la date de début, la date de fin, et un champ avec le nombre de jours entre les deux.

Examine preview data

Rows of step: Calculator (1 rows)

	start_date	end_date	diff_date
1	2021-01-20	2021-02-15	26

Close

6. Ajoutez maintenant l'étape "Clone row". Vous la trouverez à l'intérieur du groupe d'étapes "Utility group" et configurez-la comme indiqué :

**Clone row**

Step name:

Nr clones:

Nr clone in field? ☒

Nr Clone field:

Output fields

Add clone flag to output? ☐

Clone flag field:

Add clone num to output? ☒

Clone num field:

[? Help](#) [OK](#) [Cancel](#)

7. Créez un saut de l'étape de "Calculator" vers cette nouvelle étape et configurez-la comme indiqué :

**Calculator**

Step name:

Fields:

	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove	Con
1	a_single_date	Date A + B Days	start_date	delta					N	Con

[? Help](#) [OK](#) [Cancel](#)

8. Lancez un aperçu final. Vous devriez voir ceci :



Examine preview data

Rows of step: Calculator 2 (27 rows)

	start_date	end_date	diff_dates	delta	a_single_date
1	2021-01-20	2021-02-15	26	0	2021/01/20 00:00:00.000
2	2021-01-20	2021-02-15	26	1	2021/01/21 00:00:00.000
3	2021-01-20	2021-02-15	26	2	2021/01/22 00:00:00.000
4	2021-01-20	2021-02-15	26	3	2021/01/23 00:00:00.000
5	2021-01-20	2021-02-15	26	4	2021/01/24 00:00:00.000
6	2021-01-20	2021-02-15	26	5	2021/01/25 00:00:00.000
7	2021-01-20	2021-02-15	26	6	2021/01/26 00:00:00.000
8	2021-01-20	2021-02-15	26	7	2021/01/27 00:00:00.000

Close

Maintenant vous allez exécuter la Transformation et inspecter les données pendant que la Transformation est en cours d'exécution. Avant de le faire, nous allons apporter quelques modifications à la Transformation afin qu'elle s'exécute lentement, ce qui nous permettra de voir en détail ce qui se passe.

1. Modifiez l'étape "Generate Rows" et changez la plage de dates. En tant que end\_date, tapez 2023-12-31.
2. Dans le groupe "Utility group ", faites glisser vers la zone de travail l'étape "Delay row" . Avec cette étape, nous allons délibérément retarder chaque ligne de données.
3. Faites glisser l'étape vers le hop entre l'étape "Clone row" et la deuxième étape de "Calculator", jusqu'à ce que la largeur change.



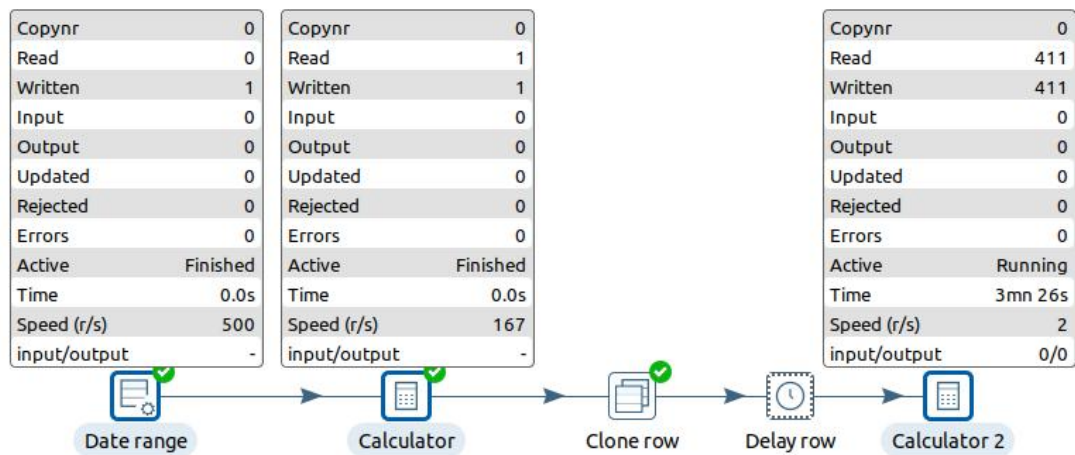
4. Double-cliquez sur l'étape "Delay row" et configurez-la en utilisant les informations suivantes : comme Timeout, tapez 500, et dans la liste déroulante, sélectionnez Millisecondes
5. Sauvegardez la transformation et exécutez-la. Vous verrez qu'elle s'exécute à un rythme lent.

#### Execution Results

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)
1	Date range	0	0	1	0	0	0	0	0	Finished	0.0s	500
2	Calculator	0	1	1	0	0	0	0	0	Finished	0.0s	167
3	Clone row	0	1	1076	0	0	0	0	0	Finished	0.0s	107,600
4	Delay row	0	96	95	0	0	0	0	0	Running	47.8s	2
5	Calculator 2	0	95	95	0	0	0	0	0	Running	47.8s	2

6. Sans arrêter l'exécution, cliquez sur la deuxième étape Calculator. Une fenêtre popup apparaîtra, décrivant les résultats de l'exécution de cette étape en temps réel. temps réel. Ctrl-cliquez sur deux autres étapes : l'étape Generate Rows et l'étape Clone row et l'étape Clone row. Pour chaque étape sélectionnée, vous verrez les métriques de l'étape en cours d'exécution :



7. Cliquez avec le bouton droit de la souris sur la deuxième étape de la Calculator et naviguez jusqu'à Sniff Test During Execution | Sniff test output rows :

The screenshot shows the data pipeline interface with the 'Calculator' step selected. A right-click context menu is open, displaying various options. The 'Sniff Test During Execution' option is highlighted, and its sub-menu is visible, showing 'Sniff test input rows', 'Sniff test output rows', and 'Sniff test error handling'.

Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active
0	0	1	0	0	0	0	0	Finished
0	1	1	0	0	0	0	0	Finished
0	1	1076	0	0	0	0	0	Finished
0	742	741	0	0	0	0	0	Running
0	741	741	0	0	0	0	0	Running

8. Pendant l'exécution de la transformation, vous avez expérimenté la fonction de sniffer les lignes de sortie. De la même manière, vous auriez pu sélectionner l'option Sniff test input rows pour voir les lignes de données entrantes.

Examine preview data

Rows of step: Calculator 2 (0 rows)

	start_date	end_date	diff_dates	delta	a_single_date
14	2021-01-20	2023-12-31	1075	251	2021/09/28 00:00:00.000
15	2021-01-20	2023-12-31	1075	252	2021/09/29 00:00:00.000
16	2021-01-20	2023-12-31	1075	253	2021/09/30 00:00:00.000
17	2021-01-20	2023-12-31	1075	254	2021/10/01 00:00:00.000
18	2021-01-20	2023-12-31	1075	255	2021/10/02 00:00:00.000
19	2021-01-20	2023-12-31	1075	256	2021/10/03 00:00:00.000
20	2021-01-20	2023-12-31	1075	257	2021/10/04 00:00:00.000
21	2021-01-20	2023-12-31	1075	258	2021/10/05 00:00:00.000

Close