

SYNTÈSE

Lorsque le programme reconnaît **un ou plusieurs** fichiers .csv :

TIMESTAMP	OP	EVENT TYPE	INFO
11:39:12,302	convert_all_csvs	STEP_START	Started execution of step "convert_all_csvs".
11:39:12,324	convert_all_csvs	INFO	Traitement de : htsdata (41 to 70).csv -> htsdata (41 to 70).xlsx
11:39:19,284	convert_all_csvs	INFO	Fichier sauvegardé : C:\Users\faceb\OneDrive - DVHE\Documents\TRAVAIL\STAGE\Dagster\dagster-csv-to-excel\csv_to_excel_project\data\outputs\htsdata
11:39:19,294	convert_all_csvs	STEP_OUTPUT	Yielded output "result" of type "Any". (Type check passed).
11:39:19,433	convert_all_csvs	ASSET_MATERIALIZE	Materialized value convert_all_csvs. asset_key convert_all_csvs [View Asset] path C:\Users\faceb\OneDrive - DVHE\Documents\TRAVAIL\STAGE\Dagster\dagster-csv-to-excel\tmp_dagster_home_nbn7xnsilstorage\convert_all_csvs
11:39:20,381	convert_all_csvs	HANDLED_OUTPUT	Handled output "result" using IO manager path C:\Users\faceb\OneDrive - DVHE\Documents\TRAVAIL\STAGE\Dagster\dagster-csv-to-excel\tmp_dagster_home_nbn7xnsilstorage\convert_all_csvs
11:39:20,392	convert_all_csvs	STEP_SUCCESS	Finished execution of step "convert_all_csvs" in 8.07s.
11:39:23,466	-	ENGINE_EVENT	Multiprocess executor: parent process exiting after 32.2s (pid: 39584) pid 39584
11:39:23,496	-	RUN_SUCCESS	Finished execution of run for "__ASSET_JOB".
11:39:23,624	-	ENGINE_EVENT	Process for run exited (pid: 39584).

```
# On loggue un message pour dire quel fichier on est en train de traiter
context.log.info(f"✅ Traitement de : {filename} -> {new_filename}")

# Lecture du CSV
df = pd.read_csv(input_path)

# Ecriture en Excel (garder le même nom de base)
df.to_excel(output_path, index=False)

# On ajoute le chemin à notre liste de succès
processed_files.append(output_path)

context.log.info(f"✅ Fichier sauvegardé : {output_path}")
```

Un bout du code assets.py

Lorsque le programme ne reconnaît **pas un ou plusieurs** fichiers .csv :

```
# si la liste est vide, on prévient qu'on n'a rien trouvé
if not processed_files:
    context.log.warning("⚠ Aucun fichier CSV trouvé dans le dossier inputs !")
```

- Pourquoi **Materialized** ?

Cela montre tout simplement que le programme est arrivé à terme. Dans le code, le programme se finit qu'il ait trouvé un fichier .csv à convertir ou non. Le programme n'a pas été fait pour tourner en boucle jusqu'à trouver un fichier à convertir.

TIMESTAMP	OP	EVENT TYPE	INFO
1:15:32,679	convert_all_csvs	STEP_START	Started execution of step "convert_all_csvs".
1:15:32,690	convert_all_csvs	WARNING	⚠ Aucun fichier CSV trouvé dans le dossier inputs !
1:15:32,758	convert_all_csvs	STEP_OUTPUT	Yielded output "result" of type "Any". (Type check passed).
1:15:32,824	convert_all_csvs	ASSET_MATERIALIZE	Materialized value convert_all_csvs. asset_key convert_all_csvs [View Asset] path C:\Users\faceb\OneDrive - DVHE\Documents\TRAVAIL\STAGE\Dagster\dagster-csv-to-excel\tmp_dagster_home_nbn7xnsilstorage\convert_all_csvs
1:15:33,430	convert_all_csvs	HANDLED_OUTPUT	Handled output "result" using IO manager path C:\Users\faceb\OneDrive - DVHE\Documents\TRAVAIL\STAGE\Dagster\dagster-csv-to-excel\tmp_dagster_home_nbn7xnsilstorage\convert_all_csvs
1:15:33,454	convert_all_csvs	STEP_SUCCESS	Finished execution of step "convert_all_csvs" in 764ms.
1:15:44,619	-	ENGINE_EVENT	Multiprocess executor: parent process exiting after 31.83s (pid: 27056) pid 27056
1:15:44,627	-	RUN_SUCCESS	Finished execution of run for "__ASSET_JOB".
1:15:44,781	-	ENGINE_EVENT	Process for run exited (pid: 27056).

Lorsque le programme n'arrive pas à lire un fichier .csv mais qu'il en lit un autre en même temps :

- Pourquoi ?

Il n'y avait pas de try et except dans mon assets.py ce qui veut dire qu'un simple fichier corrompu pouvait faire crasher l'analyse. Dans bad_data.csv, Pandas a lu l'en-tête (ligne 1) et a vu 4 colonnes. Arrivé à la ligne 4, il a trouvé 9 morceaux séparés par des virgules. Il ne sait pas faire rentrer 9 valeurs dans 4 cases ce qui crash l'analyse.

Sans try/except :

14:15:17,346	Since start of run	16,286	0,023	FAILURE
14:15:17,651				
dagster._core.errors.DagsterExecutionStepExecutionError: Error occurred while executing op "convert_all_csvs": The above exception was caused by the following exception: pandas.errors.ParserError: Error tokenizing data. C error: Expected 4 fields in line 4, saw 9				
Stack Trace: File "C:\Users\faceb\AppData\Local\Programs\Python\Python313\Lib\site-packages\dagster_core\execution\plan\utils.py", line 57, in op_execution_error_boundary yield , File "C:\Users\faceb\AppData\Local\Programs\Python\Python313\Lib\site-packages\dagster_utils__init__.py", line 394, in iterate_with_context next_output = next(iterator) raise AnException("The stack trace is too long to display") from e View full message				
14:15:24,722	-	ENGINE_EVENT	Multiprocess executor: parent process exiting after 23.36s (pid: 8292)	pid 8292
14:15:24,755	-	RUN_FAILURE	Execution of run for "__ASSET_JOB" failed. Steps failed: ['convert_all_csvs'].	
14:15:25,006	-	ENGINE_EVENT	Process for run exited (pid: 8292).	

Avec :

```
# >>> DÉBUT DE LA ZONE PROTÉGÉE <<<
try:
    # 1. On essaie de lire le CSV
    # C'est ici que 'bad_data.csv' va déclencher une alerte, mais pas un crash
    df = pd.read_csv(input_path)

    # Petite vérification supplémentaire si le fichier est vide
    if df.empty:
        raise ValueError("Le fichier est vide")

    # 2. on écrit le fichier Excel
    df.to_excel(output_path, index=False)

    # 3. si on arrive ici, c'est que tout s'est bien passé
    context.log.info(f"✅ Succès : {new_filename}")
    processed_files.append(output_path)

except Exception as e:
    # >>> ZONE DE GESTION D'ERREUR <<<
    # Si n'importe quoi se passe mal au-dessus, on atterrit ici.
    error_message = f"❌ ÉCHEC sur {filename}. Raison : {str(e)}"

    # On note L'erreur en rouge dans les logs
    context.log.error(error_message)

    # On ajoute le fichier à la liste des échecs pour le bilan
    failed_files.append(filename)

# >>> FIN DE LA ZONE PROTÉGÉE <<<
```

EVENT TYPE	INFO
LOGS_CAPTURED	Started capturing logs in process (pid: 34956). step_keys: ["convert_all_csvs"] captured_logs: View stdout / stderr
CSVS STEP_START	Started execution of step "convert_all_csvs".
CSVS INFO	✅ Tentative de traitement : bad_data.csv
CSVS ERROR	❌ ÉCHEC sur bad_data.csv. Raison : Error tokenizing data. C error: Expected 4 fields in line 4, saw 9
CSVS INFO	✅ Tentative de traitement : htsdata (41 to 70).csv
CSVS INFO	✅ Succès : htsdata (41 to 70).xlsx
CSVS WARNING	⚠ Terminé avec des erreurs. Fichiers échoués (1) : ['bad_data.csv']
CSVS STEP_OUTPUT	Yielded output "result" of type "Any". (Type check passed).
CSVS ASSET_MATERIALIZE	Materialized value convert_all_csvs. asset_key: convert_all_csvs [View Asset] path: C:\Users\faceb\OneDrive - DVHE\Documents\TRAVAIL\STAGE\Danster\dagster-csv-to-excel\two_dagster_hc
CSVS HANDLED_OUTPUT	Handled output "result" using IO manager "io_manager" path: C:\Users\faceb\OneDrive - DVHE\Documents\TRAVAIL\STAGE\Danster\dagster-csv-to-excel\two_dagster_hc
CSVS STEP_SUCCESS	Finished execution of step "convert_all_csvs" in 8.44s.
ENGINE_EVENT	Multiprocess executor: parent process exiting after 24.57s (pid: 35656) pid 35656
RUN_SUCCESS	Finished execution of run for "__ASSET_JOB".
ENGINE_EVENT	Process for run exited (pid: 35656).