

Informe Tarea 3

Integrantes:

- Nelson Sepulveda | 202004610-7
- Francisco Domínguez | 202104520-1
- Vicente Henríquez | 202051507-7

Dificultades Presentadas: Durante la realización de la tarea, primeramente, se nos dificultó el uso de Apache Airflow, debido a que la configuración de la conexión y los puertos a utilizar nos generaban problemas. Para solucionar esto, utilizamos como base el archivo de configuración airflow.cfg del workspace_airflow que se mostró de ejemplo en una de las clases. Luego otra de las dificultades y la que mas tiempo nos tomó arreglar se presentó a la hora de manejar data proveniente de un bucket con el uso de MinIO, debido a que cuando en el flujo de trabajo se llegaba a la ejecución del pipeline en Apache Beam, este fallaba. Por esto, decidimos utilizar Cursor para identificar y corregir el problema, para cual utilizamos una serie de prompts que se resumen a lo siguiente:

“Actualmente en el pipeline se presentan estos errores cuando se ejecuta la tarea run_beam_pipeline:

(Logs de la consola de Airflow)

Identifica que está ocurriendo y proponme una solución. Para esto analiza los archivos del proyecto para detectar posibles errores entre múltiples archivos”