Ejercicio Neural Works en GCP

Datos se reciben mediante webhook y sistema de mensajería Pub/Sub que envía mensaje para notificar a suscriptores, por ejemplo Bigquery, quien al recibir mensaje almacenará los datos en una tabla particionada por día.

Para realizar las transformaciones se consumen los datos mediante API de Bigquery en lenguaje python, mediante script que realizará los procesos de descarga, transformación y carga en base de datos SQL y/o tabla de Bigquery.

Para automatizar el proceso de ingesta de datos se utiliza Airflow Apache que es compatible con GCP si la máquina es parte de Google Cloud y si los archivos .py almacenados en Cloud Storage.

