

BITACORA DE EJECUCIÓN

TITULO: Creación de Google cloud, la máquina virtual y el sector de almacenamiento.

OBJETIVO: Crear la máquina virtual en Google cloud, con los sectores de almacenamiento, y los códigos de Python para la ejecución de los procesos.

ARCHIVOS:

- `main.py`: Ejecuta procesos de Web scraping, los almacena en Google cloud, los procesa. Y los deja almacenados en Cloud Storage. Listos para ser consumidos por Big Query y el sector de Machine Learning.

PROCEDIMIENTO

1. Crear cuenta de GCP.
2. Habilitamos las APIs de Bucket (Cloud Storage), Big Query, Looker, Compute Engine y Vertex AI, Google Schedule.
3. En Compute Engine se crea la instancia de la maquina virtual. Se le instalan los paquetes necesarios para ejecutar un script de Python.
4. Se configura Google Schedule, para prender la Maquina Virtual en un tiempo determinado.
5. Al prender la maquina virtual se ejecuta el “`main.py`” de forma automática.
6. El script crea la carpeta en el Bucket, donde se va a almacenar los archivos del WebScraping y los archivos transformados para el consumo en Big Query y el sector de Machine Learning.
7. Una vez que se ejecuta todo lo programado, Google Schedule se encarga de apagarlo a una hora determinada por la programación prevista.