Hi All,!

Requirements:

The copy of the dataset in the data warehouse should reflect exactly the current state of the data at the source. For the sake of this exercise, assume that the dataset is updated daily at the source.

The business intelligence team needs to run queries that aggregate these incidentes along the following dimensions: time period, district, and battalion

**Context Executed: I downloaded a csv file upload on GCP Cloud using BIG QUERY resource and others GCP components to help with it, schedule and extracted Data. Resources like topics, Cloud Functions, Trigger e Scheduler were part of the strategy.**

We had some memories limits so we needed to reduce process power and size. Results like report file extract was written and stored in a cloud bucket

Below it is the evidence of the test and it will be packed together with files result in the github

**First we created our structure and test environment in GCP**

Interface gráfica do usuário, Texto, Aplicativo, Email

Descrição gerada automaticamente

**Second we created our GCP Topic**

Tela de computador com texto preto sobre fundo branco

Descrição gerada automaticamente

**Third we created our GCP Cloud Function**

Tela de computador com texto preto sobre fundo branco

Descrição gerada automaticamente

**Then we configure**

Interface gráfica do usuário, Texto, Aplicativo, Email

Descrição gerada automaticamente

**Then we scheduled a daily JOB**

Interface gráfica do usuário, Texto, Aplicativo, Email

Descrição gerada automaticamente

**Here is the folder with files and the report**

Interface gráfica do usuário, Texto, Aplicativo, Email

Descrição gerada automaticamente

Interface gráfica do usuário, Texto, Aplicativo, Email

Descrição gerada automaticamente

Memory Limit

Interface gráfica do usuário, Texto, Aplicativo, Email

Descrição gerada automaticamente

**Here is the log file register for Loaded 681467 rows.**

Interface gráfica do usuário, Texto, Aplicativo, Email

Descrição gerada automaticamente

<https://github.com/oseiascampos1963/cloud_file_upload.git>

main.py file is the Cloud file to update the table and generate report main2.py is a local file to upload files from local