

Destrinchando o GCP

Google Cloud Dataproc

Clone do Código

```
git clone https://github.com/marcelomarques05/destrinchando-o-gcp.git
cd destrinchando-o-gcp
WORKING_DIR=$(pwd)
# Atenção: Troque o valor abaixo para o nome do seu projeto
PROJECT_ID="SEU_ID_PROJETO"
gcloud config set project $PROJECT_ID
```

Criação do Bucket

```
gsutil mb gs://$PROJECT_ID
```

Copia dos arquivos para o bucket

```
cd dataproc/spark-bigquery
gsutil cp counts_by_subreddit.py gs://$PROJECT_ID
gsutil cp backfill* gs://$PROJECT_ID
```