

## Section 1: API

In the context of a DB migration with 3 different tables (departments, jobs, employees) , create

a local REST API that must:

1. Receive historical data from CSV files

I have selected the GCP Rest APIs

The screenshot shows the Google Cloud Storage console for a bucket named 'my\_bucket\_bh'. The bucket is located in 'us-central1 (Iowa)' and uses 'Standard' storage class. It has 'No public access' and 'Delete without versioning' protection. The 'OBJECTS' tab is selected, showing a list of files:

Nombre	Tamaño	Tipo	Fecha de creación	Clase de almacenamiento	Última modificación	Acceso público	Historial
<a href="#">dag_migracion_jobs.py</a>	1.2 KB	text/x-python	26 dic 2024 02:13:11	Standard	26 dic 2024 02:13:11	No público	—
<a href="#">departments.csv</a>	189 B	application/vnd.ms-excel	26 dic 2024 05:10:03	Standard	26 dic 2024 05:10:03	No público	—
<a href="#">hired_employees.csv</a>	90.8 KB	application/vnd.ms-excel	26 dic 2024 05:30:23	Standard	26 dic 2024 05:30:23	No público	—
<a href="#">jobs.csv</a>	4.3 KB	text/csv	23 dic 2024 19:30:17	Standard	23 dic 2024 19:30:17	No público	—

2. Upload these files to the new DB
3. Be able to insert batch transactions (1 up to 1000 rows) with one request

The screenshot shows the Airflow web interface for a DAG named 'gcs\_globant\_job\_dag'. The DAG is titled 'Carga datos desde GCS a BigQuery'. The 'load\_hired\_employees' task is highlighted, showing its execution details. The task is a 'GCSToBigQueryOperator' and has completed successfully. The DAG graph shows three tasks: 'load\_jobs', 'load\_departments', and 'load\_hired\_employees', all of which are successful. The DAG is scheduled to run at 09:22:03 UTC on 2024-12-29.

Google Cloud My Project Data Engineer compo Buscar

Explorador + AGREGAR

Buscar recursos de BigQuery

Mostrar solo los destacados

dev-splice-445621-u2

- Consultas
- Notebooks
- Lienzos de datos
- Preparaciones de datos
- Cargas de trabajo
- Conexiones externas
- globant\_prueba
  - DEPARTMENTS
  - HIRED\_EMPLOYEES
  - JOBS**

JOBS CONSULTA ABRIR EN COMPARTIR COPIAR INSTANTÁNEA BORRAR

ESQUEMA DETALLES VISTA PREVIA EXPLORADOR DE TABLAS VISTA PREVIA ESTADÍSTICAS

Filtro Ingresar el nombre o el valor de la propiedad

Nombre del campo	Tipo	Modo	Clave	Intercalación	Valor predeterminado	Etiquetas de política
id	INTEGER	REQUIRED	-	-	-	-
job	STRING	REQUIRED	-	-	-	-

EDITAR ESQUEMA VER POLÍTICAS DE ACCESO DE FILA

You need to publish your code in GitHub. It will be taken into account if frequent updates are made to the repository that allow analyzing the development process. Ideally, create a markdown file for the Readme.md

Importar marcadores... Primeros pasos Algoritmos de ordena...

Dashboard

Search: Type to search

bernarditahernandez

Top repositories

Find a repository...

bernarditahernandez/GCP

bernarditahernandez/Globant\_Ejercicios

Ask Copilot

<> Rails authentication endpoint Create a profile README for me

Home

Filter

bernarditahernandez / Globant\_Ejercicios

Search: Type to search

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

Globant\_Ejercicios Public

main 1 Branch 0 Tags

Go to file Add file <> Code About

bernarditahernandez Add files via upload e062b53 - 5 minutes ago 4 Commits

Globant's Data Engineering Coding Challenge...	Add files via upload	3 hours ago
Querys_Globant_DE.sql	Add files via upload	5 minutes ago
Querys_Prueba_Globant.sql	Add files via upload	3 hours ago
README.md	Create README.md	3 hours ago
departments.csv	Add files via upload	3 hours ago
gcs_globant_job_dag.py	Add files via upload	31 minutes ago
hired_employees.csv	Add files via upload	3 hours ago
jobs.csv	Add files via upload	3 hours ago

README

About

Globant

- Readme
- Activity
- 0 stars
- 1 watching
- 0 forks

Releases

No releases published

Create a new release

Packages

No packages published

Publish your first package

Languages

## Section 2: SQL

You need to explore the data that was inserted in the previous section. The stakeholders ask for some specific metrics they need. You should create an end-point for each requirement.

### Requirements

- Number of employees hired for each job and department in 2021 divided by quarter. The table must be ordered alphabetically by department and job.

The screenshot shows the Google Cloud BigQuery interface. On the left is the 'Explorador' (Explorer) pane with a search bar and a tree view of resources. The main area displays a SQL query titled 'Consulta sin título' (Untitled query) with the following code:

```
1 SELECT t2.string_field_1 AS departments,
2       t3.string_field_1 AS jobs,
3       SUM(CASE WHEN EXTRACT(QUARTER FROM t1.timestamp_field_2) = 1 THEN 1 ELSE 0 END) AS Q1,
4       SUM(CASE WHEN EXTRACT(QUARTER FROM t1.timestamp_field_2) = 2 THEN 1 ELSE 0 END) AS Q2,
5       SUM(CASE WHEN EXTRACT(QUARTER FROM t1.timestamp_field_2) = 3 THEN 1 ELSE 0 END) AS Q3,
6       SUM(CASE WHEN EXTRACT(QUARTER FROM t1.timestamp_field_2) = 4 THEN 1 ELSE 0 END) AS Q4
7 FROM `dev-splice-445621-u2.globant_prueba.HIRED_EMPLOYEES` t1 INNER JOIN
8     `dev-splice-445621-u2.globant_prueba.DEPARTMENTS` t2 ON
9     (t1.int64_field_3=t2.int64_field_0) INNER JOIN
10    `dev-splice-445621-u2.globant_prueba.JOBS` t3 ON
11    (t1.int64_field_4=t3.int64_field_0)
12 WHERE EXTRACT(YEAR FROM t1.timestamp_field_2)=2021
```

Below the query, the 'Resultados de la consulta' (Query results) section is visible, showing a table with columns for 'INFORMACIÓN DEL TRABAJO' (Job Information) and 'DETALLES DE LA EJECUCIÓN' (Execution Details). The table is sorted by department and job, and the results are displayed for the year 2021.

INFORMACIÓN DEL TRABAJO	RESULTADOS	GRÁFICO	JSON	DETALLES DE LA EJECUCIÓN	GRÁFICO DE EJECUCIÓN	
Fila	departments	jobs	Q1	Q2	Q3	Q4
155	Support	Web Developer II	0	0	0	1
156	Support	Editor	0	0	1	1
157	Support	Web Developer III	0	1	0	0
158	Support	Assistant Media Planner	0	0	0	2
159	Support	Statistician I	0	2	0	0

At the bottom right, it indicates 'Resultados por página: 50' (Results per page: 50) and '151 - 200 de 899' (151 - 200 of 899).

employees than the mean of employees hired in 2021 for all the departments, ordered by the number of employees hired (descending).

Google Cloud

My Project Data Engineer

Buscar (/) recursos, documentos, productos y más

Explorador

+ AGREGAR

<

Buscar recursos de BigQuery

Mostrar solo los destacados

dev-splice-445621-u2

Consultas

Notebooks

Lienzos de datos

Preparaciones de datos

Cargas de trabajo

Conexiones externas

globant\_prueba

RESUMEN

Actualmente no se seleccionó nada

Consulta sin título

EJECUTAR

GUARDAR

DESCARGAR

+ C

```
10 AND IFNULL(t1.int64_field_0,0)<>0
11 GROUP BY t1.int64_field_3,
12 t2.string_field_1)
13
14 SELECT * FROM HIRED_EMPLOYEES_AVG
15 WHERE count_hired>( SELECT AVG(hired) FROM (--OBTENGO EL PROMEDIO
16 SELECT t1.int64_field_3 AS id_departments,
17 COUNT(t1.int64_field_0) AS hired
18 FROM `dev-splice-445621-u2.globant_prueba.HIRED_EMPLOYEES` t1 INNER JOIN
19 `dev-splice-445621-u2.globant_prueba.DEPARTMENTS` t2 ON
20 (t1.int64_field_3=t2.int64_field_0) INNER JOIN
```

Resultados de la consulta

INFORMACIÓN DEL TRABAJO		RESULTADOS	GRÁFICO	JSON	DETALLES D
Fila	int64_field_3	string_field_1	count_hired		
1	8	Support	217		
2	5	Engineering	207		
3	6	Human Resources	203		
4	7	Services	202		
5	4	Business Development	187		
6	3	Research and Development	149		

Developed by:

Bernardita Hernández

2024-12-29