

Para este proyecto, se crearon únicamente los modelos `staging_bank_marketing` y `kpi_bankmarketing`, conforme a lo indicado en la consigna.

En el **README** del repositorio se incluye una sección con la explicación de los campos del modelo. Además, la descripción detallada de los campos y las pruebas se encuentra en el archivo `schema.yml` dentro de la carpeta `models`. Por otro lado, en las carpetas `macros` y `tests` se han añadido algunos tests básicos para aumentar la complejidad del proyecto.

Además, se emplea el archivo `sources.yml` dentro de `models` para una mejor organización del proyecto.

También se utiliza el archivo `packages.yml` para instalar la librería `dbt_utils`, lo que podría facilitar la creación de una clave primaria, dado que el dataset original no tiene una.

El dataset se cargó directamente en Google Cloud, dado que el ejercicio no especificaba el uso de un seed en dbt.

Estado de prueba gratuita: crédito por \$300.00 y 84 días restantes. Activa tu cuenta completa para obtener acceso ilimitado a todas las funciones de Google Cloud. Usa los créditos restantes y paga solo por lo que usas.

Google Cloud

bank marketing project

Buscar (/) recursos, documentos, productos y más

BigQuery

Explorador

+ AGREGAR

K

Studio

Canalizaciones e integraci...

Transferencias de datos

Canalizaciones (Dataform)

Consultas programadas

Programación ...

Administración

Analytics Hub

Etiquetas de política

Administración

Supervisión

Explorador de trabajos

Administración de capaci...

BI Engine

Recuperación ante desast...

Centro de socios

Configuración

VISTA PREVIA

Notas de versión

Buscar recursos de BigQuery

Mostrar solo los destacados

bank-marketing-project-448216

Consultas

Notebooks

Lienzos de datos

Preparaciones de datos

Cargas de trabajo

Conexiones externas

bank_marketing

raw_bank_marketing

dbt_gabriel

RESUMEN

raw_bank_marketing

bank-marketing-project-448216.bank_marketing

Última modificación

18 ene 2025, 1:34:56 p.m. UTC-3

Ubicación de los datos

northamerica-northeast1

Descripción

raw_bank_marketing

CONSULTA

COMPARTIR

COPIAR

INSTANTÁNEA

ESQUEMA

DETALLES

VISTA PREVIA

EXPLORADOR DE TABLAS

VISTA PREVIA

ESTADÍSTIC

Filtro

Ingresar el nombre o el valor de la propiedad

| | Nombre del campo | Tipo | Modo | Clave | Intercalación | Valor predeterminado | Eti |
|--------------------------|------------------|---------|----------|-------|---------------|----------------------|-----|
| <input type="checkbox"/> | age | INTEGER | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | job | STRING | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | marital | STRING | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | education | STRING | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | default | STRING | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | housing | STRING | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | loan | STRING | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | contact | STRING | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | month | STRING | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | day_of_week | STRING | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | duration | INTEGER | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | campaign | INTEGER | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | pdays | INTEGER | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | previous | INTEGER | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | poutcome | STRING | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | emp_var_rate | FLOAT | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | cons_price_idx | FLOAT | NULLABLE | - | - | - | - |
| <input type="checkbox"/> | cons_conf_idx | FLOAT | NULLABLE | - | - | - | - |

EDITAR ESQUEMA

VER POLÍTICAS DE ACCESO DE FILA

Algunas queries de prueba sobre la tabla raw para verificar la carga de la tabla raw_bank_marketing:

Explorador

+ AGREGAR

<

Buscar recursos de BigQuery

?

Mostrar solo los destacados

bank-marketing-project-448216

☆

:

Consultas

:

Consultas compartidas

:

COUNT

:

Notebooks

:

Lienzos de datos

:

Preparaciones de datos

:

Cargas de trabajo

:

Conexiones externas

:

bank_marketing

☆

:

raw_bank_marketing

☆

:

bank-ma... 216

X

*COUNT

X

+

▼

COUNT

EJECUTAR

GUARDAR CONSULTA

▼

DESCARGAR

COMPARTIR

▼

1

SELECT COUNT(*) FROM `bank-marketing-project-448216.bank_marketing.raw_bank_marketing`

Resultados de la consulta

INFORMACIÓN DEL TRABAJO

RESULTADOS

GRÁFICO

JSON

DETALLES DE LA EJEC

| Fila | f0_ |
|------|-------|
| 1 | 41188 |

```
1 SELECT * FROM `bank-marketing-project-448216.bank_marketing.raw_bank_marketing`
```

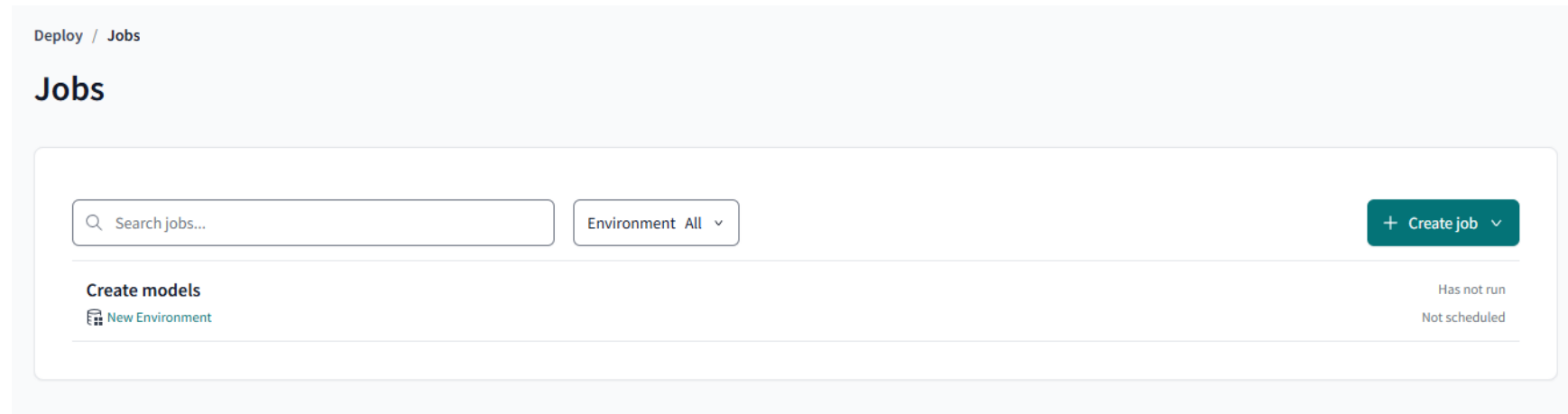
Presiona Alt+F1 para ver las opciones de accesibilidad

Resultados de la consulta

GUARDAR LOS RESULTADOS ABRIR EN

| INFORMACIÓN DEL TRABAJO | | | RESULTADOS | GRÁFICO | JSON | DETALLES DE LA EJECUCIÓN | | GRÁFICO DE EJECUCIÓN | |
|-------------------------|-----|--------|------------|---------------------|---------|--------------------------|---------|----------------------|--|
| Fila | age | job | marital | education | default | housing | loan | | |
| 1 | 38 | admin. | single | professional.course | no | no | no | | |
| 2 | 53 | admin. | single | professional.course | no | no | no | | |
| 3 | 35 | admin. | single | high.school | no | yes | no | | |
| 4 | 42 | admin. | single | university.degree | unknown | no | no | | |
| 5 | 51 | admin. | single | basic.6y | no | no | no | | |
| 6 | 35 | admin. | single | university.degree | no | no | no | | |
| 7 | 32 | admin. | single | high.school | no | no | no | | |
| 8 | 37 | admin. | single | high.school | no | yes | yes | | |
| 9 | 27 | admin. | single | university.degree | no | no | no | | |
| 10 | 34 | admin. | single | basic.9y | unknown | no | no | | |
| 11 | 53 | admin. | single | basic.6y | no | yes | yes | | |
| 12 | 26 | admin. | single | university.degree | no | yes | no | | |
| 13 | 46 | admin. | single | high.school | no | yes | no | | |
| 14 | 38 | admin. | single | university.degree | no | no | no | | |
| 15 | 44 | admin. | single | basic.9y | unknown | no | no | | |
| 16 | 32 | admin. | single | university.degree | unknown | no | no | | |
| 17 | 32 | admin. | single | university.degree | unknown | unknown | unknown | | |
| 18 | 47 | admin. | single | unknown | unknown | yes | no | | |
| 19 | 35 | admin. | single | high.school | no | yes | yes | | |
| 20 | 38 | admin. | single | basic.6y | unknown | no | no | | |
| 21 | 34 | admin. | single | university.degree | no | no | no | | |

En dbt cloud se ejecuta el job:



También se puede ejecutar en un entorno local (como una VM con Ubuntu). Para esto, es necesario que el archivo profiles.yml esté ubicado en la ruta ~/.dbt/profiles.yml. Además, se deben cumplir los siguientes requisitos: tener Python instalado junto con dbt configurado correctamente.

profiles.yml

default:

outputs:

default:

type: bigquery

method: service-account

project: bank-marketing-project-448216

dataset: dbt_gabriel

keyfile: "<RUTA DEL JSON DE GCP>"

threads: 4

target: default

Run #70471827937778

✓ Success

↺ Rerun now

Triggered 0m 40s ago

Trigger
👤 Triggered manually

Commit SHA
🔗 #016cb90

Run duration
🕒 34s

Run summary

Lineage

Model timing

Artifacts

🕒 Time in queue
Prep time 0m

- > ✓ Clone git repository0s
- > ✓ Create profile from connection BigQuery0s
- > ✓ Invoke dbt deps2s
- > ✓ Invoke dbt build19s

Finished 0m 4s ago

>  **Invoke** dbt deps

2s


▼  **Invoke** dbt build


19s

Console Logs

Debug Logs

Search logs...

 Download logs

 27 Successes ^ ▼

```
23:21:33 25 of 27 START test not_null_kpi_bank_marketing_segment ..... [RUN]
23:21:33 26 of 27 START test not_null_kpi_bank_marketing_successful_contacts ..... [RUN]
23:21:33 27 of 27 START test not_null_kpi_bank_marketing_total_contacts ..... [RUN]
23:21:33 26 of 27 PASS not_null_kpi_bank_marketing_successful_contacts ..... [PASS in 0.76s]
23:21:33 27 of 27 PASS not_null_kpi_bank_marketing_total_contacts ..... [PASS in 0.77s]
23:21:33 25 of 27 PASS not_null_kpi_bank_marketing_segment ..... [PASS in 0.79s]
23:21:33 24 of 27 PASS not_null_kpi_bank_marketing_conversion_rate_percentage ..... [PASS in 0.79s]
23:21:34
23:21:34
23:21:34 Finished running 2 table models, 25 data tests in 0 hours 0 minutes and 10.04 seconds (10.04s).
23:21:34
23:21:34
23:21:34 Completed successfully
23:21:34
23:21:34
23:21:34 Done. PASS=27 WARN=0 ERROR=0 SKIP=0 NO-OP=0 TOTAL=27
```

Queries de prueba en la BD de los modelos:

Google Cloud

bank marketing project

Buscar (/) recursos, documentos, productos y más

Buscar

Explorador

+ AGREGAR

Buscar recursos de BigQuery

Mostrar solo los destacados

bank-marketing-project-448216

Consultas

Consultas compartidas

COUNT

Notebooks

Lienzos de datos

Preparaciones de datos

Cargas de trabajo

Conexiones externas

bank_marketing

dbt_gabriel

kpi_bank_marketing

staging_bank_marketing

RESUMEN

Actualmente no se seleccionó nada

Consulta sin título

EJECUTAR

GUARDAR

DESCARGAR

COMPARTIR

PROGRAMACIÓN

ABRIR EN

MÁS

1 SELECT * FROM `bank-marketing-project-448216.dbt_gabriel.staging_bank_marketing` LIMIT 1000

Resultados de la consulta

GUARDAR

INFORMACIÓN DEL TRABAJO

RESULTADOS

GRÁFICO

JSON

DETALLES DE LA EJECUCIÓN

GRÁFICO DE EJECUCIÓN

| Fila | client_id | age | job | marital | education | default | housing |
|------|-----------|-----|---------------|----------|---------------------|---------|---------|
| 1 | 59 | 32 | management | single | university.degree | no | no |
| 2 | 199 | 32 | technician | divorced | professional.course | no | yes |
| 3 | 236 | 39 | admin. | single | high.school | no | yes |
| 4 | 318 | 38 | blue-collar | single | high.school | no | no |
| 5 | 447 | 68 | retired | married | basic.4y | no | no |
| 6 | 452 | 95 | retired | divorced | basic.6y | no | no |
| 7 | 518 | 38 | technician | married | professional.course | no | yes |
| 8 | 539 | 65 | housemaid | divorced | basic.4y | unknown | no |
| 9 | 607 | 28 | self-employed | single | university.degree | no | yes |
| 10 | 610 | 29 | admin. | single | university.degree | no | yes |
| 11 | 646 | 31 | admin. | single | university.degree | no | yes |
| 12 | 690 | 32 | technician | married | university.degree | no | no |
| 13 | 736 | 34 | blue-collar | single | basic.9y | unknown | yes |
| 14 | 768 | 35 | admin. | single | basic.9y | no | yes |
| 15 | 848 | 37 | technician | married | professional.course | no | yes |
| 16 | 1025 | 44 | blue-collar | married | basic.4y | no | no |
| 17 | 1069 | 46 | blue-collar | married | basic.9y | unknown | yes |
| 18 | 1072 | 46 | blue-collar | married | basic.9y | unknown | yes |

Cargar más

Resultados por página: 50

Historial de trabajos

Explorador

+ AGREGAR



Buscar recursos de BigQuery

☐ Mostrar solo los destacados

bank-marketing-project-448216

Consultas

Consultas compartidas

COUNT

Notebooks

Lienzos de datos

Preparaciones de datos

Cargas de trabajo

Conexiones externas

bank_marketing

dbt_gabriel

kpi_bank_marketing

staging_bank_marketing

RESUMEN

Actualmente no se seleccionó nada

Consulta sin título

EJECUTAR

GUARDAR

DESCARGAR

COMPARTIR

PRO

1 SELECT * FROM `bank-marketing-project-448216.dbt_gabriel.kpi_bank_marketing` LIMIT 1000

Resultados de la consulta

| INFORMACIÓN DEL TRABAJO | | RESULTADOS | GRÁFICO | JSON | DETALLES DE LA EJECUCIÓN | G |
|-------------------------|---------------|----------------|---------------------|---------------------|--------------------------|---|
| Fila | segment | total_contacts | successful_contacts | conversion_rate_per | | |
| 1 | 50+ | 7178 | 1082 | 15.07383672332... | | |
| 2 | 18-25 | 1661 | 347 | 20.891029500301 | | |
| 3 | 26-35 | 14846 | 1740 | 11.72032870806... | | |
| 4 | 36-50 | 17494 | 1469 | 8.397164742197... | | |
| 5 | admin. | 10421 | 1352 | 12.97380289799... | | |
| 6 | Overall | 41184 | 4640 | 11.26651126651... | | |
| 7 | retired | 1720 | 434 | 25.23255813953... | | |
| 8 | student | 875 | 275 | 31.42857142857... | | |
| 9 | unknown | 330 | 37 | 11.21212121212... | | |
| 10 | unknown | 5 | 2 | 40.0 | | |
| 11 | services | 3969 | 323 | 8.138070042831... | | |
| 12 | housemaid | 1060 | 106 | 9.999999999999... | | |
| 13 | management | 2923 | 328 | 11.22134793020... | | |
| 14 | technician | 6743 | 730 | 10.82604182114... | | |
| 15 | unemployed | 1014 | 144 | 14.20118343195... | | |
| 16 | blue-collar | 9252 | 638 | 6.895806312148... | | |
| 17 | entrepreneur | 1456 | 124 | 8.516483516483... | | |
| 18 | self-employed | 1421 | 149 | 10.48557353976... | | |

El job se ejecuta correctamente en Github Actions:

```
✓ Run DBT tests
37 21:30:42 14 of 25 START test not_null_staging_bank_marketing_age_group ..... [RUN]
38 21:30:42 15 of 25 START test not_null_staging_bank_marketing_client_id ..... [RUN]
39 21:30:42 11 of 25 PASS not_null_kpi_bank_marketing_successful_contacts ..... [PASS in 0.93s]
40 21:30:42 16 of 25 START test not_null_staging_bank_marketing_contact_duration_minutes ... [RUN]
41 21:30:43 13 of 25 PASS not_null_staging_bank_marketing_age ..... [PASS in 0.84s]
42 21:30:43 17 of 25 START test not_null_staging_bank_marketing_contact_frequency ..... [RUN]
43 21:30:43 14 of 25 PASS not_null_staging_bank_marketing_age_group ..... [PASS in 0.88s]
44 21:30:43 18 of 25 START test not_null_staging_bank_marketing_education_normalized ..... [RUN]
45 21:30:43 16 of 25 PASS not_null_staging_bank_marketing_contact_duration_minutes ..... [PASS in 0.84s]
46 21:30:43 19 of 25 START test not_null_staging_bank_marketing_has_previous_contact ..... [RUN]
47 21:30:43 15 of 25 PASS not_null_staging_bank_marketing_client_id ..... [PASS in 0.88s]
48 21:30:43 20 of 25 START test not_null_staging_bank_marketing_job_normalized ..... [RUN]
49 21:30:44 17 of 25 PASS not_null_staging_bank_marketing_contact_frequency ..... [PASS in 0.80s]
50 21:30:44 21 of 25 START test not_null_staging_bank_marketing_last_contact_date ..... [RUN]
51 21:30:44 18 of 25 PASS not_null_staging_bank_marketing_education_normalized ..... [PASS in 0.86s]
52 21:30:44 22 of 25 START test not_null_staging_bank_marketing_marital_normalized ..... [RUN]
53 21:30:44 20 of 25 PASS not_null_staging_bank_marketing_job_normalized ..... [PASS in 0.85s]
54 21:30:44 23 of 25 START test not_null_staging_bank_marketing_previous_campaign_success .. [RUN]
55 21:30:44 19 of 25 PASS not_null_staging_bank_marketing_has_previous_contact ..... [PASS in 0.91s]
56 21:30:44 24 of 25 START test positive_value_staging_bank_marketing_contact_duration_minutes [RUN]
57 21:30:45 21 of 25 PASS not_null_staging_bank_marketing_last_contact_date ..... [PASS in 0.81s]
58 21:30:45 25 of 25 START test unique_staging_bank_marketing_client_id ..... [RUN]
59 21:30:45 22 of 25 PASS not_null_staging_bank_marketing_marital_normalized ..... [PASS in 0.78s]
60 21:30:45 24 of 25 PASS positive_value_staging_bank_marketing_contact_duration_minutes ... [PASS in 0.77s]
61 21:30:45 23 of 25 PASS not_null_staging_bank_marketing_previous_campaign_success ..... [PASS in 0.82s]
62 21:30:46 25 of 25 PASS unique_staging_bank_marketing_client_id ..... [PASS in 0.80s]
63 21:30:46
64 21:30:46 Finished running 25 data tests in 0 hours 0 minutes and 6.53 seconds (6.53s).
65 21:30:46
66 21:30:46 Completed successfully
67 21:30:46
68 21:30:46 Done. PASS=25 WARN=0 ERROR=0 SKIP=0 TOTAL=25

✓ Generate DBT docs
1 ▶ Run dbt docs generate --profiles-dir /home/runner/.dbt --target default
7
7 21:30:47 Running with dbt=1.9.1
8 21:30:48 Registered adapter: bigquery=1.9.1
9 21:30:49 Found 2 models, 25 data tests, 1 source, 605 macros
10 21:30:49
```

El archivo `dbt_pipeline.yml` dentro de `.github/workflows` configura el pipeline automatizado en GitHub Actions para ejecutar los modelos DBT en BigQuery. Este pipeline se activa con un `push` o `pull_request` en el branch `main`, lo que asegura que los cambios en el código de producción se procesen automáticamente. Además, se encarga de instalar las dependencias, configurar el entorno y ejecutar los modelos DBT, pruebas y documentación.

El pipeline comienza obteniendo el código del repositorio con `actions/checkout@v2` y configurando Python 3.9 mediante `actions/setup-python@v2`, lo que permite ejecutar DBT. Luego, instala DBT y sus dependencias usando `pip`, y las dependencias del proyecto con `dbt deps`. Se crea un archivo JSON con las credenciales de Google Cloud desde los secretos de GitHub Actions, y se verifica su contenido para asegurar la autenticación correcta. A continuación, se genera el archivo `profiles.yml` con la configuración necesaria para conectar DBT a BigQuery y se verifica su creación. El pipeline ejecuta los modelos DBT, corre las pruebas para verificar la consistencia de los datos y, finalmente, genera la documentación de los modelos DBT.