

## Task 1

- Run juffle-shop dbt project from <https://github.com/dbt-labs/jaffle-shop>
- Use either Bigquery, or Snowflake, or Redshift as a target database
- Don't use DBT Cloud env, only use **dbt command line**
- Provide for acceptance:
  - screenshots of lineage graph
  - screenshot of tables on your environment
  - log of executed commands

## Task 2

- Check other DBT projects:  
<https://docs.getdbt.com/faqs/Project/example-projects>
- Repeat one of them in either Bigquery, or Snowflake, or Redshift.

# Task 1

## Create a new repository from template

The screenshot shows a GitHub repository page for 'dbt-lab-jaffle-shop'. The repository is public and was generated from 'dbt-labs/jaffle-shop'. It has 1 branch and 0 tags. The page includes standard GitHub navigation buttons like 'Pin', 'Unwatch', 'Go to file', 'Add file', and 'Code'.

**Let's try to work with BigQuery: create a new project in GCP and enable BigQuery API**

The screenshot shows the Google Cloud Platform API library for the BigQuery API. It is part of the 'Jaffle Shop Project'. The 'BigQuery API' page is displayed, showing it is enabled. Navigation links include 'OVERVIEW', 'DOCUMENTATION', and 'RELATED PRODUCTS'.

## Create a new service account and provide a minimum set of permissions

The screenshot shows the 'Create service account' wizard. Step 1: Service account details is completed. Step 2: Grant this service account access to project (optional) is in progress. It lists five roles: BigQuery Data Viewer, BigQuery Filtered Data Viewer, BigQuery Data Editor, BigQuery Job User, and BigQuery Read Session User. Each role has an 'IAM condition (optional)' section with a '+ ADD IAM CONDITION' button. A note says: 'Grant this service account access to Jaffle Shop Project so that it has permission to complete specific actions on the resources in your project.' Buttons at the bottom include '+ ADD ANOTHER ROLE', 'CONTINUE', 'DONE', and 'CANCEL'.

## Create a JSON key

Filter Enter property name or value

Email	Status	Name ↑	Description	Key ID	Key creation date	OAuth 2 Client ID	Actions
<a href="#">dbt-service-account@jaffle-shop-project.iam.gserviceaccount.com</a>	Enabled	dbt-service-account	Service account for dbt to access BigQuery	No keys		116841703435227037463	<ul style="list-style-type: none"><li>Manage details</li><li>Manage permissions</li><li>Manage keys</li><li>View metrics</li><li>View logs</li><li>Disable</li><li>Delete</li></ul>

Type	Status	Key	Creation date	Expiration date	
	Active	ce90cf7f2790bd7cb4abcba8db8a9994f1a997a6	Dec 9, 2024	Jan 1, 10000	

## Create a dataset

### Create dataset

Project ID \*

jaffle-shop-project

[CHANGE](#)

Dataset ID \*

jaffle\_shop

Letters, numbers, and underscores allowed

Location type [?](#)

Region

Specify a region to colocate your datasets with other Google Cloud services.

Multi-region

Allow BigQuery to select a region within a group to achieve higher quota limits.

Region \*

us-central1 (Iowa)

jaffle-shop-project

☆ ::

Create new

SQL QUERY

Recently opened

jaffle\_shop  
Tap to open in current tab  
(Cmd + Tap - open in new tab,  
Shift + Cmd + Tap - open in split tab)

Queries

Notebooks

Data canvases

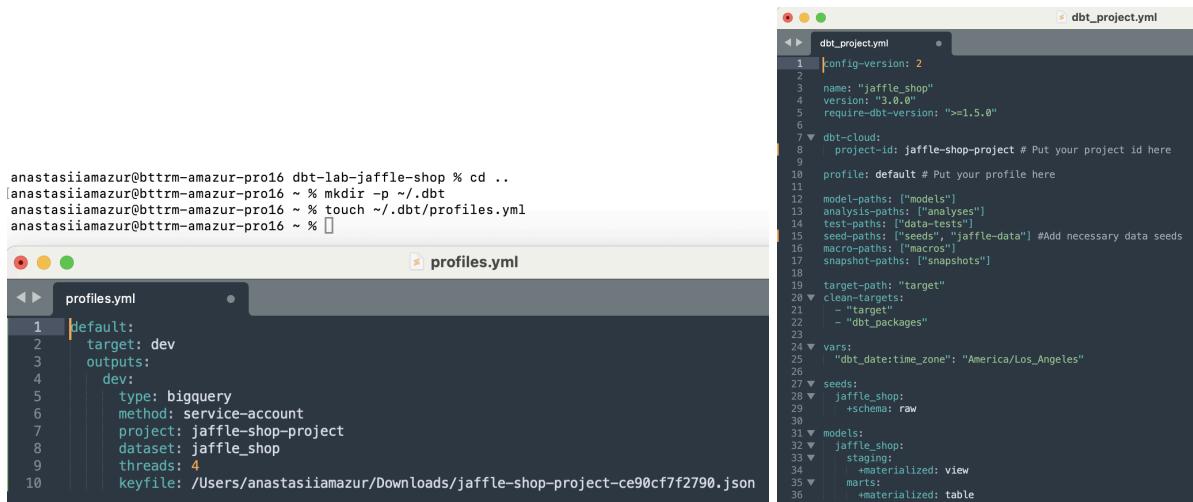
Data preparations

Workflows

External connections

jaffle\_shop

## Setting up the basic connection files



The screenshot shows a terminal window with three tabs. The top tab is titled 'dbt\_project.yml' and contains the following YAML code:

```
1 config-version: 2
2
3 name: "jaffle_shop"
4 version: "3.0.0"
5 require-dbt-version: ">=1.5.0"
6
7 dbt-cloud:
8   project-id: jaffle-shop-project # Put your project id here
9
10 profile: default # Put your profile here
11
12 model-paths: ["models"]
13 analysis-paths: ["analyses"]
14 test-paths: ["data_tests"]
15 seed-paths: ["seeds","jaffle-data"] #Add necessary data seeds
16 macro-paths: ["macros"]
17 snapshot-paths: ["snapshots"]
18
19 target-path: "target"
20 clean-targets:
21   - "target"
22   - "dbt_packages"
23
24 vars:
25   "dbt_date_time_zone": "America/Los_Angeles"
26
27 seeds:
28   jaffle_shop:
29     +schema: raw
30
31 models:
32   jaffle_shop:
33     staging:
34       +materialized: view
35   marts:
36     +materialized: table
```

The middle tab is titled 'profiles.yml' and contains the following YAML code:

```
1 default:
2   target: dev
3   outputs:
4     dev:
5       type: bigquery
6       method: service-account
7       project: jaffle-shop-project
8       dataset: jaffle_shop
9       threads: 4
10      keyfile: /Users/anastasiiamazur/Downloads/jaffle-shop-project-ce90cf7f2790.json
```

The bottom tab shows a command-line session:

```
anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % cd ..
anastasiiamazur@bttrm-amazur-pro16 ~ % mkdir -p ~/.dbt
anastasiiamazur@bttrm-amazur-pro16 ~ % touch ~/.dbt/profiles.yml
anastasiiamazur@bttrm-amazur-pro16 ~ %
```

## Installing BigQuery packages

Since my Python 3.12 environment is managed by Homebrew, I cannot install packages globally due to the externally-managed-environment restriction

```
anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % python3.12 -m pip install dbt-bigquery
error: externally-managed-environment

x This environment is externally managed
↳ To install Python packages system-wide, try brew install
xyz, where xyz is the package you are trying to
install.

If you wish to install a Python library that isn't in Homebrew,
use a virtual environment:

python3 -m venv path/to/venv
source path/to/venv/bin/activate
python3 -m pip install xyz
```

To avoid conflicts and ensure isolated package management, I'll use a virtual environment (venv)

```
anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % python3.12 -m venv venv
anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % source venv/bin/activate
(venv) anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % pip install --upgrade pip
pip install dbt-bigquery
Requirement already satisfied: pip in ./venv/lib/python3.12/site-packages (24.0)
Collecting pip
  Downloading pip-24.3.1-py3-none-any.whl.metadata (3.7 kB)
  Downloading pip-24.3.1-py3-none-any.whl (1.8 MB)
    1.8/1.8 MB 4.5 MB/s eta 0:00:00
(venv) anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % dbt --version
Core:
- installed: 1.8.9
- latest: 1.8.9 - Up to date!
Plugins:
- bigquery: 1.8.3 - Up to date!
```

## Check the connection

```
(venv) anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % dbt debug
23:46:57  Running with dbt=1.8.9
23:46:57  db version: 1.8.9
23:46:57  python version: 3.12.4
23:46:57  python path: /Users/anastasiiamazur/dbt-lab-jaffle-shop/venv/bin/python3.12
23:46:57  os info: macOS-14.4.1-arm64-arm-64bit
23:46:58  Using profiles dir at /Users/anastasiiamazur/.dbt
23:46:58  Using profiles.yml file at /Users/anastasiiamazur/.dbt/profiles.yml
23:46:58  Using dbt_project.yml file at /Users/anastasiiamazur/dbt-lab-jaffle-shop/dbt_project.yml
23:46:58  adapter type: bigquery
23:46:58  adapter version: 1.8.3
23:46:58  Configuration:
23:46:58    profiles.yml file [OK found and valid]
23:46:58    dbt_project.yml file [OK found and valid]
23:46:58  Required dependencies:
23:46:58    - git [OK found]

23:46:58  Connection:
23:46:58    method: service-account
23:46:58    database: jaffle-shop-project
23:46:58    execution_project: jaffle-shop-project
23:46:58    schema: jaffle_shop
23:46:58    location: None
23:46:58    priority: None
23:46:58    maximum_bytes_billed: None
23:46:58    impersonate_service_account: None
23:46:58    job_retry_delay_in_seconds: None
23:46:58    job_retry_limit: 1
23:46:58    job_creation_timeout_seconds: None
23:46:58    job_execution_timeout_seconds: None
23:46:58    timeout_seconds: None
23:46:58    client_id: None
23:46:58    token_uri: None
23:46:58    dataproc_region: None
23:46:58    dataproc_cluster_name: None
23:46:58    gcs_bucket: None
23:46:58    dataproc_batch: None
23:46:58  Registered adapter: bigquery=1.8.3
23:47:00  Connection test: [OK connection ok]

23:47:00  All checks passed!
```

## Dependency loading

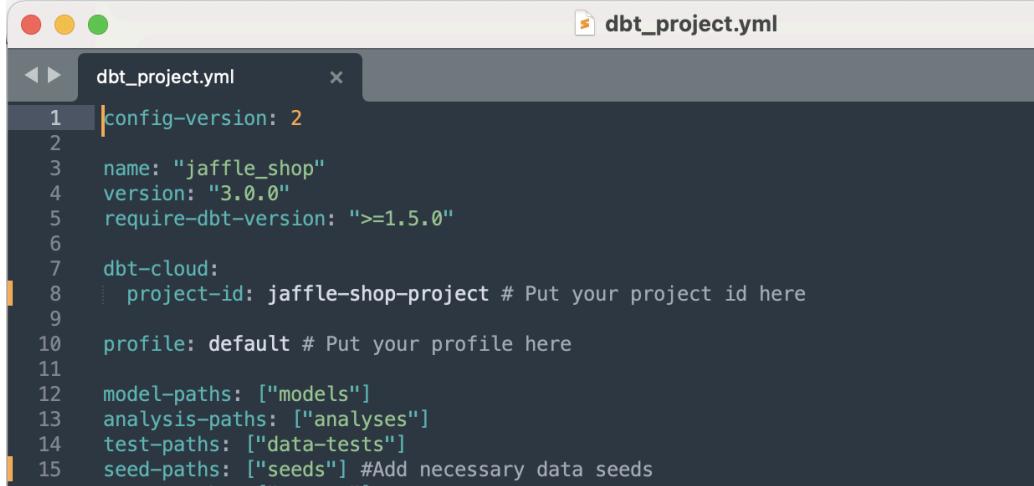
```
(venv) anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % dbt deps
23:52:39  Running with dbt=1.8.9
23:52:42  WARNING: The git package "https://github.com/dbt-labs/dbt-audit-helper.git"
is pinned to the "main" branch.
This can introduce breaking changes into your project without warning!
See https://docs.getdbt.com/docs/package-management#section-specifying-package-versions
23:52:43  WARNING: The git package "https://github.com/dbt-labs/dbt-audit-helper.git"
is pinned to the "main" branch.
This can introduce breaking changes into your project without warning!
See https://docs.getdbt.com/docs/package-management#section-specifying-package-versions
23:52:43  Updating lock file in file path: /Users/anastasiiamazur/dbt-lab-jaffle-shop/package-lock.yml
23:52:43  Installing dbt-labs/dbt_utils
23:52:44  Installed from version 1.1.1
23:52:44  Updated version available: 1.3.0
23:52:44  Installing calogica/dbt_date
23:52:44  Installed from version 0.10.0
23:52:44  Updated version available: 0.10.1
23:52:44  Installing https://github.com/dbt-labs/dbt-audit-helper.git
23:52:46  Installed from revision 74072850a5bccd90235576d67530c98e3b7437f4
23:52:46
23:52:46  Updates available for packages: ['dbt-labs/dbt_utils', 'calogica/dbt_date']
Update your versions in packages.yml, then run dbt deps
```

## Data download

```
(venv) anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % dbt seed
23:53:28  Running with dbt=1.8.9
23:53:29  Registered adapter: bigquery=1.8.3
23:53:29  Unable to do partial parsing because saved manifest not found. Starting full parse.
23:53:30  Found 13 models, 6 seeds, 27 data tests, 6 sources, 19 metrics, 784 macros, 6 semantic models, 3 saved queries, 3 unit tests
23:53:30
23:53:34  Concurrency: 4 threads (target='dev')
23:53:34
23:53:34  1 of 6 START seed file raw.raw_customers ..... [RUN]
23:53:34  2 of 6 START seed file raw.raw_items ..... [RUN]
23:53:34  3 of 6 START seed file raw.raw_orders ..... [RUN]
23:53:34  4 of 6 START seed file raw.raw_products ..... [RUN]
23:53:41  4 of 6 OK loaded seed file raw.raw_products ..... [INSERT 10 in 6.78s]
23:53:41  5 of 6 START seed file raw.raw_stores ..... [RUN]
23:53:41  1 of 6 OK loaded seed file raw.raw_customers ..... [INSERT 935 in 7.14s]
23:53:41  6 of 6 START seed file raw.raw_supplies ..... [RUN]
23:53:47  5 of 6 OK loaded seed file raw.raw_stores ..... [INSERT 6 in 6.15s]
23:53:47  6 of 6 OK loaded seed file raw.raw_supplies ..... [INSERT 65 in 6.01s]
23:53:50  2 of 6 OK loaded seed file raw.raw_items ..... [INSERT 9090 in 16.39s]
s]
23:53:56  3 of 6 OK loaded seed file raw.raw_orders ..... [INSERT 61948 in 21.68]
s]
23:53:56  Finished running 6 seeds in 0 hours 0 minutes and 25.37 seconds (25.37s).
23:53:56  Completed successfully
23:53:56
23:53:56  Done. PASS=6 WARN=0 ERROR=0 SKIP=0 TOTAL=6
```

## Cleaning phase: remove the jaffle-data folder and jaffle-data from the seed-paths list

```
(venv) anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % rm -r jaffle-data  
(venv) anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % nano dbt_project.yml
```



```
config-version: 2  
name: "jaffle_shop"  
version: "3.0.0"  
require-dbt-version: ">=1.5.0"  
  
dbt-cloud:  
  project-id: jaffle-shop-project # Put your project id here  
  
profile: default # Put your profile here  
  
model-paths: ["models"]  
analysis-paths: ["analyses"]  
test-paths: ["data-tests"]  
seed-paths: ["seeds"] #Add necessary data seeds
```

Rerun the dbt seed command: this should confirm that dbt is no longer trying to download data from the jaffle-data folder

```
(venv) anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % dbt seed  
00:04:16 Running with dbt=1.8.9  
00:04:16 Registered adapter: bigquery=1.8.3  
00:04:17 Unable to do partial parsing because a project config has changed  
00:04:18 [WARNING]: Configuration paths exist in your dbt_project.yml file which do not apply to any resources.  
There are 1 unused configuration paths:  
- seeds.jaffle_shop  
00:04:18 Found 13 models, 27 data tests, 6 sources, 19 metrics, 784 macros, 6 semantic models, 3 saved queries, 3 unit tests  
00:04:18  
00:04:18 Nothing to do. Try checking your model configs and model specification args  
00:04:18  
00:04:18 Completed successfully  
00:04:18  
00:04:18 Done. PASS=0 WARN=0 ERROR=0 SKIP=0 TOTAL=0
```

## Run a dbt build to build the project

```
(venv) anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % dbt build  
00:07:52 Running with dbt=1.8.9  
00:07:52 Registered adapter: bigquery=1.8.3  
00:07:53 [WARNING]: Configuration paths exist in your dbt_project.yml file which do not apply to any resources.  
There are 1 unused configuration paths:  
- seeds.jaffle_shop  
00:07:53 Found 13 models, 27 data tests, 6 sources, 19 metrics, 784 macros, 6 semantic models, 3 saved queries, 3 unit tests  
00:07:53  
00:07:54 Concurrency: 4 threads (target='dev')  
00:07:54 1 of 46 START sql table model jaffle_shop.metricflow_time_spine ..... [RUN]  
00:07:54 2 of 46 START sql view model jaffle_shop.stg_customers ..... [RUN]  
00:07:54 7 of 46 START unit_test stg.locations::test_does_location_opened_at_trunc_to_date [RUN]  
00:07:54 3 of 46 START sql view model jaffle_shop.stg_order_items ..... [RUN]  
00:07:55 BigQuery adapter: https://console.cloud.google.com/bigquery?project=jaffle-shop-project&j=bq:US:5c86c06b-985f-4842-880a-e8ce20eadd96&page=queryresults  
00:07:55 BigQuery adapter: https://console.cloud.google.com/bigquery?project=jaffle-shop-project&j=bq:US:002a51f4-6634-4037-9304-7fb6380674a9&page=queryresults  
00:07:55 2 of 46 ERROR creating sql view model jaffle_shop.stg_customers ..... [ERROR in 0.91s]  
00:07:55 3 of 46 ERROR creating sql view model jaffle_shop.stg_order_items ..... [ERROR in 0.91s]  
00:07:55 4 of 46 START sql view model jaffle_shop.stg_orders ..... [RUN]  
00:07:55 5 of 46 BigQuery adapter: https://console.cloud.google.com/bigquery?project=jaffle-shop-project&j=bq:US:c60e2809-1401-4d55-930d-de69404870fe&page=queryresults  
00:07:56 4 of 46 ERROR creating sql view model jaffle_shop.stg_orders ..... [ERROR in 0.74s]  
00:07:56 6 of 46 START sql view model jaffle_shop.stg_supplies ..... [RUN]  
00:07:56 BigQuery adapter: https://console.cloud.google.com/bigquery?project=jaffle-shop-project&j=bq:US:03c3e328-e687-4528-84f4-bcef8c2c5f61&page=queryresults  
00:07:56 5 of 46 ERROR creating sql view model jaffle_shop.stg_products ..... [ERROR in 0.76s]  
00:07:56 8 of 46 SKIP test not_null_stg_customers_customer_id ..... [SKIP]  
00:07:56 9 of 46 SKIP test unique_stg_customers_customer_id ..... [SKIP]  
00:07:56 10 of 46 SKIP test not_null_stg_order_items_order_id ..... [SKIP]  
00:07:56 11 of 46 SKIP test not_null_stg_order_items_order_item_id ..... [SKIP]  
00:07:56 12 of 46 SKIP test unique_stg_order_items_order_id ..... [SKIP]
```

```

00:08:03 Finished running 6 view models, 27 data tests, 7 table models, 3 unit tests, 3 saved queries in 0 hours 0 minutes and 9.93 seconds (9.93s).
00:08:03 Completed with 9 errors and 0 warnings:
00:08:03 Database Error in model stg_customers (models/staging/stg_customers.sql)
Not found: Dataset jaffle-shop-project:jaffle_shop was not found in location US
compiled code at target/run/jaffle_shop/models/staging/stg_customers.sql
00:08:03 Database Error in model stg_order_items (models/staging/stg_order_items.sql)
Not found: Dataset jaffle-shop-project:jaffle_shop was not found in location US
compiled code at target/run/jaffle_shop/models/staging/stg_order_items.sql
00:08:03 Database Error in model stg_orders (models/staging/stg_orders.sql)
Not found: Dataset jaffle-shop-project:jaffle_shop was not found in location US
compiled code at target/run/jaffle_shop/models/staging/stg_orders.sql
00:08:03 Database Error in model stg_products (models/staging/stg_products.sql)
Not found: Dataset jaffle-shop-project:jaffle_shop was not found in location US
compiled code at target/run/jaffle_shop/models/staging/stg_products.sql
00:08:03 Database Error in model stg_supplies (models/staging/stg_supplies.sql)
Not found: Dataset jaffle-shop-project:jaffle_shop was not found in location US
compiled code at target/run/jaffle_shop/models/staging/stg_supplies.sql
00:08:03 Exception on worker thread. 'SavedQuery' object has no attribute 'schema'
00:08:03 Exception on worker thread. 'SavedQuery' object has no attribute 'schema'
00:08:03 Exception on worker thread. 'SavedQuery' object has no attribute 'schema'
00:08:03 Database Error in model stg_locations (models/staging/stg_locations.sql)
Not found: Dataset jaffle-shop-project:jaffle_shop was not found in location US
compiled code at target/run/jaffle_shop/models/staging/stg_locations.sql
00:08:03 Done. PASS=2 WARN=0 ERROR=9 SKIP=35 TOTAL=46

```

This is a combined command that executes:

- dbt run: run models
- dbt test: tests
- dbt seed: loading seed data, if necessary
- dbt snapshot: create snapshot tables

And I have some mistakes both in testing and in building data. The errors indicate that dbt build cannot find some resources in BigQuery, in particular, the jaffle\_shop Dataset in the specified region.

To fix this, I need to check if the `profiles.yml` file contains the correct Project ID and Dataset name (spoiler: no, I didn't specify the region 😊)

`jaffle_shop`

---

**Dataset info**

Dataset ID	jaffle-shop-project.jaffle_shop
Created	Dec 9, 2024, 1:27:47 AM UTC+2
Default table expiration	Never
Last modified	Dec 9, 2024, 1:27:47 AM UTC+2
Data location	us-central1
Description	
Default collation	
Default rounding mode	ROUNDING_MODE_UNSPECIFIED
Time travel window	7 days
Case insensitive	false
Labels	
Tags	

**Dataset replica info** [PREVIEW](#)

Primary location	us-central1
------------------	-------------

```

profiles.yml
1 default:
2   target: dev
3   outputs:
4     dev:
5       type: bigquery
6       method: service-account
7       project: jaffle-shop-project # Exact GCP Project ID
8       dataset: jaffle_shop # Dataset name
9       location: us-central1 # Correct region
10      threads: 4
11      keyfile: /Users/anastasiiamazur/Downloads/jaffle-shop-project-ce90cf7f2790.json # Path to the JSON key

```

Now let's try again 🙌

And I got errors again due to conflicts between raw data regions and our jaffle\_shop dataset 😅

```
00:27:42 Completed with 9 errors and 0 warnings:
00:27:42
00:27:42   Database Error in model stg_order_items (models/staging/stg_order_items.sql)
00:27:42     Not found: Dataset jaffle-shop-project:raw was not found in location us-central1
00:27:42     compiled code at target/run/jaffle_shop/models/staging/stg_order_items.sql
00:27:42
00:27:42   Database Error in model stg_customers (models/staging/stg_customers.sql)
00:27:42     Not found: Dataset jaffle-shop-project:raw was not found in location us-central1
00:27:42     compiled code at target/run/jaffle_shop/models/staging/stg_customers.sql
00:27:42
00:27:42   Database Error in model stg_orders (models/staging/stg_orders.sql)
00:27:42     Not found: Dataset jaffle-shop-project:raw was not found in location us-central1
00:27:42     compiled code at target/run/jaffle_shop/models/staging/stg_orders.sql
00:27:42
00:27:42   Database Error in model stg_products (models/staging/stg_products.sql)
00:27:42     Not found: Dataset jaffle-shop-project:raw was not found in location us-central1
00:27:42     compiled code at target/run/jaffle_shop/models/staging/stg_products.sql
00:27:42
00:27:42   Database Error in model stg_supplies (models/staging/stg_supplies.sql)
00:27:42     Not found: Dataset jaffle-shop-project:raw was not found in location us-central1
00:27:42     compiled code at target/run/jaffle_shop/models/staging/stg_supplies.sql
00:27:42
00:27:42     Exception on worker thread. 'SavedQuery' object has no attribute 'schema'
00:27:42
00:27:42     Exception on worker thread. 'SavedQuery' object has no attribute 'schema'
00:27:42
00:27:42     Exception on worker thread. 'SavedQuery' object has no attribute 'schema'
00:27:42
00:27:42   Database Error in model stg_locations (models/staging/stg_locations.sql)
00:27:42     Not found: Dataset jaffle-shop-project:raw was not found in location us-central1
00:27:42     compiled code at target/run/jaffle_shop/models/staging/stg_locations.sql
00:27:42
00:27:42 Done. PASS=2 WARN=0 ERROR=9 SKIP=35 TOTAL=46
```

**Dataset info**

Dataset ID	jaffle-shop-project.raw
Created	Dec 9, 2024, 1:53:33 AM UTC+2
Default table expiration	Never
Last modified	Dec 9, 2024, 1:53:33 AM UTC+2
Data location	US
Description	
Default collation	
Default rounding mode	ROUNDING_MODE_UNSPECIFIED
Time travel window	7 days
Case insensitive	false
Labels	
Tags	

**Dataset replica info** [PREVIEW](#)

Primary location	US
------------------	----

BigQuery does not support work between different regions, so I can't solve the problem without transferring the Dataset. The least time-consuming and convenient solution is to transfer the raw to us-central1.

## Resolving the conflict of regions

1. Delete previous `raw` data schema
2. Create a new `raw` data schema in the correct region
3. Return the data to the local cloned repository and string “jaffle-data” in the `seed-paths` of the `dbt_project.yml` file
4. Run `dbt seed` and check for data availability
5. Delete the dataset and the “jaffle-data” parameter to avoid conflicts in further work
6. Rerun `dbt build`

Now everything is fine 👍

```
00:48:15 Finished running 6 view models, 27 data tests, 3 unit tests, 7 table models, 3 saved queries in 0 hours 0 minutes and 49.20 seconds (49.20s).
00:48:15
00:48:15 Completed successfully
00:48:15
00:48:15 Done. PASS=46 WARN=0 ERROR=0 SKIP=0 TOTAL=46
```

Table	Schema
raw_customers	raw
raw_items	raw
raw_orders	raw
raw_products	raw
raw_stores	raw
raw_supplies	raw

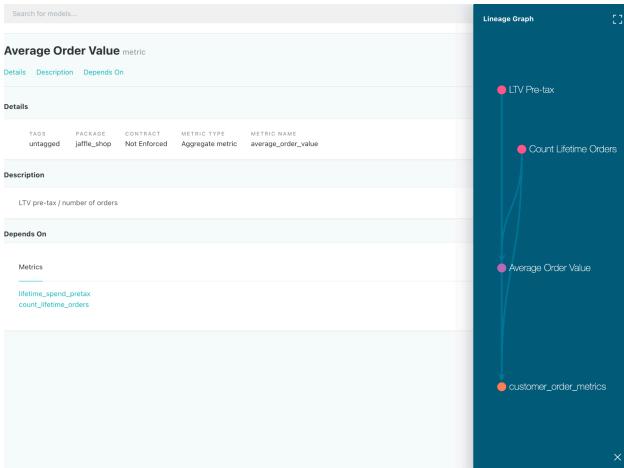
## Lineage graphs

```
(venv) anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % dbt docs generate

00:51:51  Running with dbt=1.8.9
00:51:52  Registered adapter: bigquery=1.8.3
00:51:53  [WARNING]: Configuration paths exist in your dbt_project.yml file which do not apply to any resources.
There are 1 unused configuration paths:
- seeds.jaffle_shop
00:51:53  Found 13 models, 27 data tests, 6 sources, 19 metrics, 784 macros, 6 semantic models, 3 saved queries, 3 unit tests
00:51:53
00:51:53  Concurrency: 4 threads (target='dev')
00:51:53
00:51:55  Building catalog
00:51:58  Catalog written to /Users/anastasiiamazur/dbt-lab-jaffle-shop/target/catalog.json
(venv) anastasiiamazur@bttrm-amazur-pro16 dbt-lab-jaffle-shop % dbt docs serve

00:52:14  Running with dbt=1.8.9
Serving docs at 8080
To access from your browser, navigate to: http://localhost:8080
```

**Metric: Average Order Value (just for example of simple graph)**



## Full data schema



## Task 2

Let's solve task for Google Analytics 4 project 🤝

[Source code](#), [documentation](#)

**Same initial settings were made with minor fixes for errors regarding parameter versions in files like this:**

```
02:02:57 [WARNING]: Deprecated functionality
The `source-paths` config has been renamed to `model-paths`. Please update your
`dbt_project.yml` configuration to reflect this change.
02:02:57 [WARNING]: Deprecated functionality
The `data-paths` config has been renamed to `seed-paths`. Please update your
`dbt_project.yml` configuration to reflect this change.
```

### Check previous connections and settings

```
(venv) anastasiiamazur@bttrm-amazur-pro16 stacktonic-db-example-project % dbt debug
02:06:03 Running with dbt=1.8.9
02:06:03 dbt version: 1.8.9
02:06:03 python version: 3.12.4
02:06:03 python path: /Users/anastasiiamazur/dbt-lab-jaffle-shop/venv/bin/python3.12
02:06:03 os info: macOS-14.4.1-arm64-arm-64bit
02:06:04 Using profiles dir at /Users/anastasiiamazur/.dbt/profiles.yml
02:06:04 Using dbt_project.yml file at /Users/anastasiiamazur/stacktonic-db-example-project/dbt_project.yml
02:06:04 adapter type: bigquery
02:06:04 adapter version: 1.8.3
02:06:04 Configuration:
02:06:04   profiles.yml file [OK found and valid]
02:06:04   dbt_project.yml file [OK found and valid]
02:06:04 Required dependencies:
02:06:04   - git [OK found]

02:06:04 Connection:
02:06:04   method: service-account
02:06:04   database: sublime-forest-444201-e5
02:06:04   execution_project: sublime-forest-444201-e5
02:06:04   schema: test_dataset
02:06:04   location: us-central1
02:06:04   priority: None
02:06:04   maximum_bytes_billed: None
02:06:04   impersonate_service_account: None
02:06:04   job_retry_deadline_seconds: None
02:06:04   job_retries: 1
02:06:04   job_creation_timeout_seconds: None
02:06:04   job_execution_timeout_seconds: None
02:06:04   timeout_seconds: None
02:06:04   client_id: None
02:06:04   token_uri: None
02:06:04   dataproc_region: None
02:06:04   dataproc_cluster_name: None
02:06:04   gcs_bucket: None
02:06:04   dataproc_batch: None
02:06:04 Registered adapter: bigquery=1.8.3
02:06:06 Connection test: [OK connection ok]

02:06:06 All checks passed!
```

### Get access to source data

The project "as is" is configured to use the [Google Analytics 4 example dataset](#). Therefore, I connected the entire package of available datasets to my project through the 'Explore public datasets' functionality, and then found the sample I needed separately `bigquery-public-data.ga4_obfuscated_sample_ecommerce`

[← Product details](#)



Google Analytics Sample

[Obfuscated Google Analytics 360 data](#)

Twelve months of obfuscated Google Analytics 360 data

[View dataset](#)

## Create new source dataset

### Create dataset

Project ID \*  
sublime-forest-444201-e5 [CHANGE](#)

Dataset ID \*  
ga4\_ecommerce\_custom

Letters, numbers, and underscores allowed

Location type [?](#)

Region  
Specify a region to colocate your datasets with other Google Cloud services.

Multi-region  
Allow BigQuery to select a region within a group to achieve higher quota limits.

Multi-region \*  
US (multiple regions in United States)

## Uploading data to a separate table of the newly created dataset

## Update the necessary technical files

The screenshot shows a code editor with two tabs open:

- src\_google\_analytics.yml**: A configuration file for Google Analytics data. It specifies a version of 2, sources (Google Analytics with database 'sublime-forest-444201' and dataset 'ga4\_ecommerce\_custom'), freshness rules (warn after 24 hours, error after 30 hours), and a loaded\_at\_field for timestamp parsing. It also defines a table named 'events'.
- profiles.yml**: A configuration file for a StackTonic project. It defines a target 'dev' with outputs for 'dev' and 'prod'. The 'dev' output uses BigQuery with service account authentication, targeting the 'sublime-forest-444201-e5' project and dataset 'dbt\_staging'. The 'prod' output uses BigQuery with service account authentication, targeting the 'sublime-forest-444201-e5' project and dataset 'dbt\_prod'.

## Result overview

