# foodpanda BI Exercise

For this technical test, please execute the tasks to the best of your understanding. Leave comments where appropriate if you had to make assumptions or to explain your reasoning and understanding of the task set.

## Part 1 - SQL

To work on these questions two .csv files have been provided. Please ensure you have these files before proceeding:

1) orders.csv
2) vendors.csv

You may wish to study the data in both datasets carefully before attempting the exercise. We recommend loading the datasets on a BigQuery sandbox (https://cloud.google.com/bigquery/docs/sandbox) as an environment for you to write and test your queries.

Please attempt to write SQL queries that answer the following problems.
Retain the queries for Part 2 - Submitting the Data Model.

Expected answers are provided for reference.

### 1 - Find the Total GMV by country

Expected answer:

| Row | country_name | total_gmv |
|---|---|---|
| 1 | Taiwan | 17889.23 |
| 2 | Singapore | 3424.75 |
| 3 | Thailand | 1658.26 |

### 2 - Calculate the GMV of vendors in Taiwan and order the result their customer count

Expected answer:

| Row | vendor_name | customer_c... | total_gmv |
|---|---|---|---|
| 1 | Sweet Chinese Desserts | 14 | 13491.41 |
| 2 | I LOVE PIZZAAAAAAAAAAA | 6 | 4397.82 |

## 3 - Find the top active vendor by GMV in each country

Expected answer:

| Row | country_name | vendor_name | total_gmv |
|---|---|---|---|
| 1 | Singapore | IC House | 1924.96 |
| 2 | Taiwan | Sweet Chinese Desserts | 13491.41 |
| 3 | Thailand | 9th Cafe House | 1658.26 |

## 4 - Find the top 2 vendors per country, in each year available in the dataset

Expected answer:

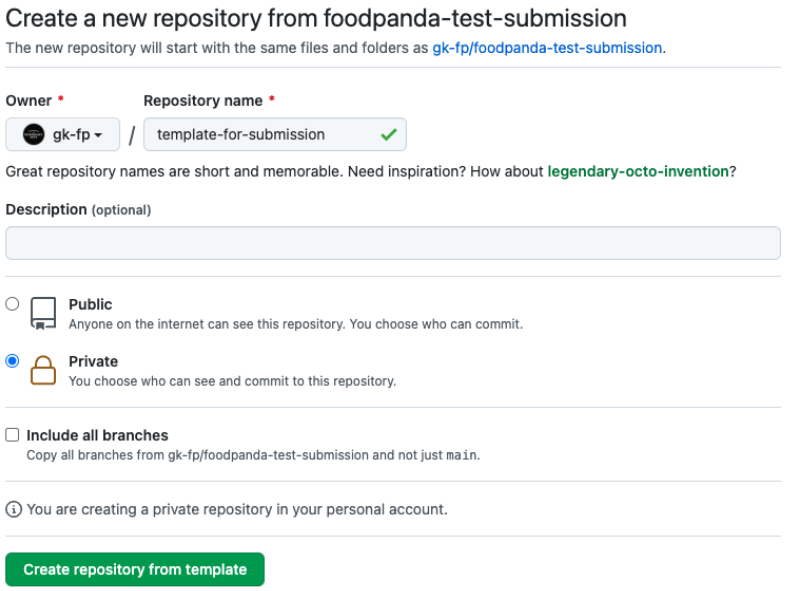
| Row | year | country_name | vendor_name | total_gmv |
|---|---|---|---|---|
| 1 | 2012-01-01T00:00:00 | Singapore | IC House | 119.34 |
| 2 | 2012-01-01T00:00:00 | Singapore | Canada Paste | 89.64 |
| 3 | 2012-01-01T00:00:00 | Taiwan | I LOVE PIZZAAAAAAAAAAA | 1613.2 |
| 4 | 2012-01-01T00:00:00 | Taiwan | Sweet Chinese Desserts | 596.1 |
| 5 | 2012-01-01T00:00:00 | Thailand | 9th Cafe House | 1658.26 |
| 6 | 2014-01-01T00:00:00 | Singapore | IC House | 1805.62 |
| 7 | 2014-01-01T00:00:00 | Singapore | WS (BM) | 768.83 |
| 8 | 2014-01-01T00:00:00 | Taiwan | Sweet Chinese Desserts | 12895.31 |
| 9 | 2014-01-01T00:00:00 | Taiwan | I LOVE PIZZAAAAAAAAAAA | 2784.62 |

# Part 2 - Submitting the Data Model

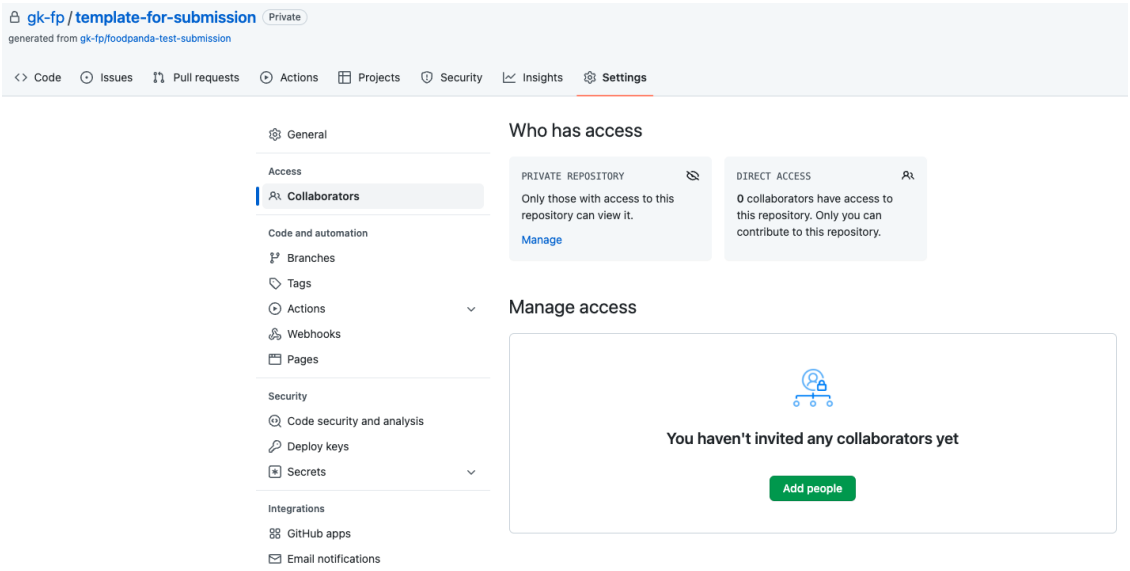
## 2.1 - Preparing a GitHub Repository

A. Ensure that you have a personal Github account.

B. Click on “Use this template” on the repository at https://github.com/gk-fp/foodpanda-test-submission.
Follow the instructions to create a new private repository on your own Github account, example shown below.

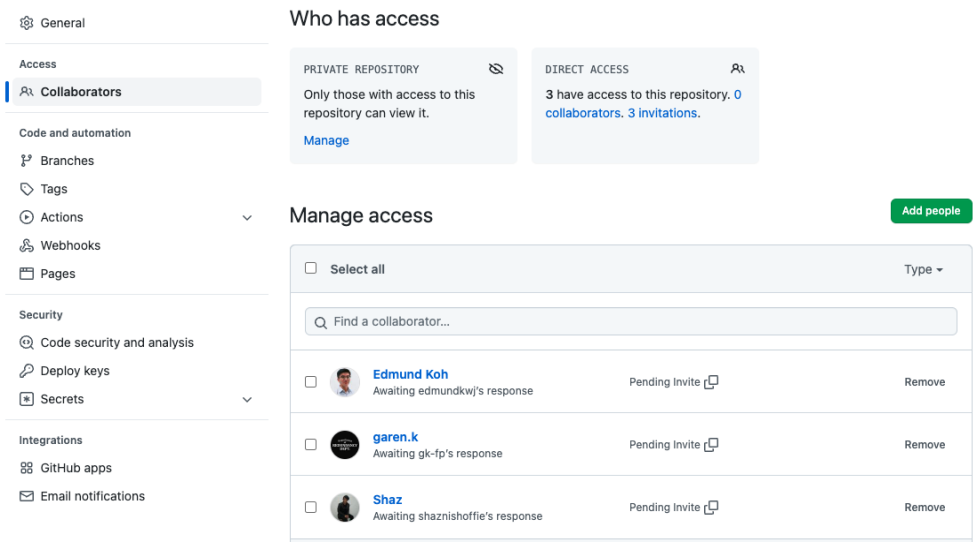



C. Click on “Settings > Collaborators > Add people”




and add the following collaborators: “`shaznishoffie`”, “`lusianadjie-fp`”, “`sweeteng-ng-fp`”

Your screen should look like below when completed.




## 2.2 - Building a DBT Model

In this section, we would like you  to create files for a dbt data model. Tools like dbt are fast becoming important in making the data analyst workflow much simpler.

*Note that you are NOT expected to create the entire dbt environment on your local machine nor do you need to do a dbt run on the files created.*

Do your best to understand the documentation for the tool and prepare the files accordingly. Here are examples of the required files.

```
{{
  config(
    materialized = 'view'
  )
}}

WITH customers AS (
  SELECT
    customers_id,
    ...
)
...
```

```
models:
  - name: q1
    description: >
      Answer to Q1 of take home test
    meta:
      owner: candidate@foodpanda.com
    columns:
      - country_name
        description: Name of country.
      - total_gmv
        description: GMV calculated for each country.
```

A. Branch the repository and give it a name “ADT-XXX-<*yourname*>” replacing the placeholder <*yourname*> with your first and last name or other unique identifier. Switch to this branch before making any further changes.
B. Please review the documentation at https://docs.getdbt.com/guides/getting-started on the dbt tool. For the purpose of this assignment, you can focus on the section “Building your first project > Build your first models”, however it is recommended that you review the other

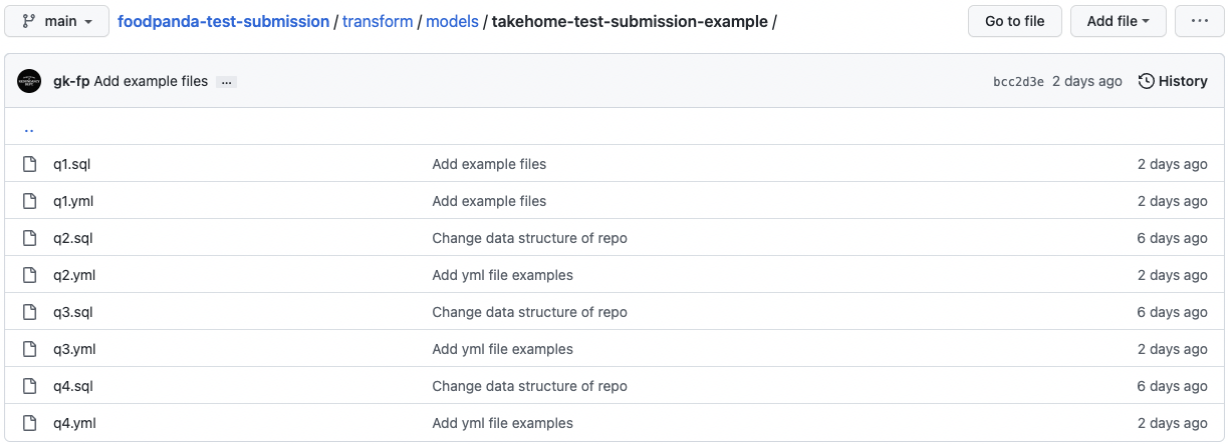
sections as well to ensure that you have a good understanding of the tool.
Please note the following points carefully:

a. It bears repeating that the purpose of this section is to test your understanding of the dbt tool through reading of documentation. As we will not use dbt Cloud in this test, commands such as dbt debug/dbt run WILL NOT work on your local machine and this is the expected behavior. However, feel free to explore dbt Cloud if it helps in your understanding of dbt.
b. It is recommended that you use a code editor such as [Sublime, VS Code, VIM, Notepad++, any other of your liking] to write these files, however this is not strictly necessary and changes can be made directly in the branch on Github if you choose to do so.
c. If using a code editor, remember to clone your repository to your local machine.
d. Once again, you are reminded that you DO NOT NEED to set up dbt on your local machine.

C. Create a new folder to store the files that you are creating for this submission, place this new folder in "transform/models".
*DO NOT overwrite the folder "transform/models/takehome-test-submission-example/".*
D. Please write each file with the assumption that the dbt tool will run the code automatically. You will be assessed on the relevance of configurations chosen for this problem.
E. Finally, make a new file that references at least one of the tables you created as part of steps C & D. The query can be as complicated or as simple as you want but should showcase that you have understood how DBT works when using the ref function.
Reference: https://docs.getdbt.com/reference/dbt-jinja-functions/ref

## 2.3 - Submitting a Pull Request

A. Please update your branch with your code for the 4 questions above in q1.sql, q2.sql, q3.sql and q4.sql in the new files created. An empty example is provided in the repository under "transform/models/takehome-test-submission-example".




B. Be sure to include all files created in Steps C and E of Section 2.2.

C. Look for the empty files order.csv and vendors.csv in the repository and replace them with the data files included in this test.
D. Submit a pull request for your branch with
    a. A clear and concise title
    b. A description that explains the work done
E. Merge the changes.
F. Include the following in your submission to the **greenhouse link** provided in the email.
    a. Github username
    b. Name of your Github repository
    c. Link to your Github repository

Part 3 - Data visualization

- Explore the Chicago Taxi Trip public Dataset from BigQuery (https://console.cloud.google.com/bigquery?p=bigquery-public-data&d=chicago_taxi_trips&page=dataset)
- Download Tableau Desktop (https://public.tableau.com/en-us/s/) or utilize Google Data Studio for this visualization exercise
- Create a story/dashboard(s) using Tableau or Data Studio (https://help.tableau.com/current/pro/desktop/en-us/stories.htm)
- Attach the final workbook (tableau public or data studio public link) along with a PDF export in the greenhouse link provided and be prepared to do a 15 minutes presentation on that workbook onsite/via call
- Take note that there is a daily quota limit for BigQuery in the free tier, see https://cloud.google.com/free/docs/free-cloud-features#free-tier-usage-limits

Note: If you're using data studio, please make sure the dashboard view settings is set to public