



Non Graded Challenge : Basic Data Wrangling in SQL

FTDS Phase 0 - Week 2

Contents



Objective	_____	03
Data	_____	04
Task	_____	05
Submission	_____	06

Non Graded Challenge

Objective

This Non Graded Challenge was created to evaluate the concept of Data Wrangling in SQL with the aim of:

1. Able to use the SELECT, FROM, and WHERE keywords to get data from a specific column.
2. Able to Explore the table schema.

Non Graded Challenge

Data

You will use `austin_bikeshare` dataset from Google BigQuery Public Dataset. For example, you can change the following syntax:

```
from google.colab import auth
from google.cloud import bigquery
import pandas as pd

auth.authenticate_user()
project_id = "your_project_id"
client = bigquery.Client(project=project_id)

# put your query here
query = """
SELECT *
FROM `bigquery-public-data.austin_bikeshare.bikeshare_trips`
"""

# Set up the query
query_job = client.query(query).to_dataframe()
query_job
```

Non Graded Challenge

Task

Create a query to answer the following questions:

- How many tables are in the `austin_bikeshare` dataset?
- How many columns in the `bikeshare_stations` table have `STRING` data?

Create a query to filter data based on the following conditions:

- Retrieve all trips where the `bike_id` is not null and the duration is greater than 1000 minutes.
- Filter trips where the `trip_id` is "Walk Up" and the `end_station_name` is "Stolen".

Non Graded Challenge

Submission

1. Create a file to store your answers.
2. Save this assignment with filename : *h8dsft_ngc_basic_data_wrangling_in_sql.ipynb*.
3. Push your answer into your own GitHub repository.



Thank You