**BigQuery Table Creation and Data Loading Using JSON Files**

**Student Name:** Sakina banu   
**Project:** Google Cloud – BigQuery Table Loading  
**Date:** May 7, 2025  
**Dataset Used:** AI Flights Data – varunkumari (JSON format only)

**Objective**

To demonstrate the ability to create and load both partitioned and non-partitioned BigQuery tables using **JSON files**, following the documented approach using GCP Cloud Shell, BigQuery CLI, and Cloud Storage.

**Step 1: Enable BigQuery and Cloud Storage APIs**

Enabled via the GCP Console:  
<https://console.cloud.google.com/flows/enableapi?apiid=bigquery,storage_component>

**Step 2: Download and Prepare Dataset**

* Downloaded AI Flights Data from Kaggle.
* Used the JSON file: 2025-04-28.json.

**Step 3: Upload File to Cloud Storage**

Used Cloud Shell to upload the file:

gsutil cp ~/archive/2025-04-28.json gs://flight-data-sakina-2025/raw/flight/

**Step 4: Create Schema File**

Created a schema.json file in Cloud Shell Editor with the following content:

[

{ "name": "flight\_date", "type": "DATE", "mode": "REQUIRED" },

{ "name": "airline\_code", "type": "STRING", "mode": "NULLABLE" },

{ "name": "flight\_num", "type": "INTEGER", "mode": "NULLABLE" },

{ "name": "source\_airport", "type": "STRING", "mode": "NULLABLE" },

{ "name": "destination\_airport", "type": "STRING", "mode": "NULLABLE" },

{ "name": "departure\_time", "type": "INTEGER", "mode": "NULLABLE" },

{ "name": "departure\_delay", "type": "INTEGER", "mode": "NULLABLE" },

{ "name": "arrival\_time", "type": "INTEGER", "mode": "NULLABLE" },

{ "name": "arrival\_delay", "type": "INTEGER", "mode": "NULLABLE" },

{ "name": "airtime", "type": "INTEGER", "mode": "NULLABLE" },

{ "name": "distance", "type": "INTEGER", "mode": "NULLABLE" },

{ "name": "id", "type": "INTEGER", "mode": "REQUIRED" }

]

**Step 5: Create BigQuery Dataset**

bq mk data\_analysis

**Step 6: Create Partitioned Table**

bq mk -t \

--schema=schema.json \

--time\_partitioning\_field=flight\_date \

data\_analysis.flight\_delays\_json\_partitioned

**Step 7: Load Data into Partitioned Table**

bq load --source\_format=NEWLINE\_DELIMITED\_JSON \

data\_analysis.flight\_delays\_json\_partitioned \

gs://flight-data-sakina-2025/raw/flight/2025-04-28.json \

schema.json

**Step 8: Validate with Query (Optional)**

bq query --use\_legacy\_sql=false \

'SELECT flight\_date, airline\_code, departure\_delay

FROM data\_analysis.flight\_delays\_json\_partitioned

WHERE flight\_date = "2025-04-28"

LIMIT 5'

A computer screen shot of a black screen

AI-generated content may be incorrect.

A screen shot of a computer

AI-generated content may be incorrect.