from google.cloud import bigquery

# Replace 'your-project-id' and 'your-dataset-id' with your actual project and dataset IDs

project\_id = 'your-project-id'

dataset\_id = 'your-dataset-id'

# Initialize a BigQuery client

client = bigquery.Client(project=project\_id)

# Example: Create a new dataset if it doesn't exist

dataset\_ref = client.dataset(dataset\_id)

dataset = bigquery.Dataset(dataset\_ref)

try:

client.get\_dataset(dataset)

print(f'Dataset {dataset\_id} already exists.')

except Exception as e:

print(f'Dataset {dataset\_id} does not exist. Creating...')

client.create\_dataset(dataset)

print(f'Dataset {dataset\_id} created successfully.')

# Example: Load data into a BigQuery table

table\_id = 'your-table-id'

table\_ref = dataset\_ref.table(table\_id)

table = bigquery.Table(table\_ref)

# Replace 'your-data.csv' with the path to your actual data file

data\_file\_path = 'your-data.csv'

# Specify job configuration (load job)

job\_config = bigquery.LoadJobConfig(

source\_format=bigquery.SourceFormat.CSV,

skip\_leading\_rows=1,

autodetect=True,

)

with open(data\_file\_path, 'rb') as source\_file:

job = client.load\_table\_from\_file(source\_file, table, job\_config=job\_config)

# Wait for the job to complete

job.result()

print(f'Data loaded into table {table\_id} successfully.')