

Bigquery

09/12/2019 LINKS

- Exercise: [kaggle.com/learn/intro-to-sql](https://www.kaggle.com/learn/intro-to-sql) > "As & with" tutorial
- Docs: cloud.google.com/bigquery/docs/

SETUP

from google.cloud import bigquery

client = bigquery.Client() *- capitalized more specific than pool*

dataset_ref = client.dataset("<table-name>", project="<proj-name>")

dataset = client.get_dataset(dataset_ref) ← can be used to PEEK

"Joining Data" has great end example broader than table

LIST TABLES • get sense of tables in database

table_list = list(client.list_tables(dataset))

for table in table_list:
print(table.table_id)

PEEK

take a look @ row table as df

table = client.get_table(dataset.table("<table-name>"))

client.list_rows(table, max_results=5).to_dataframe()

SAFE QUERY

• way to query to avoid overusing limits of query capacity

query = "
<SQL QUERY> *count(1) is useful to count rows*
(Use backticks for FROM)
"

limit, in bytes

safe_config = bigquery.QueryJobConfig(maximum_bytes_billed=10**10)

query_job = client.query(query, job_config=safe_config) ← describes the job

result = query_job.to_dataframe() ← actually runs as API request, ~~then~~ and converts to DataFrame