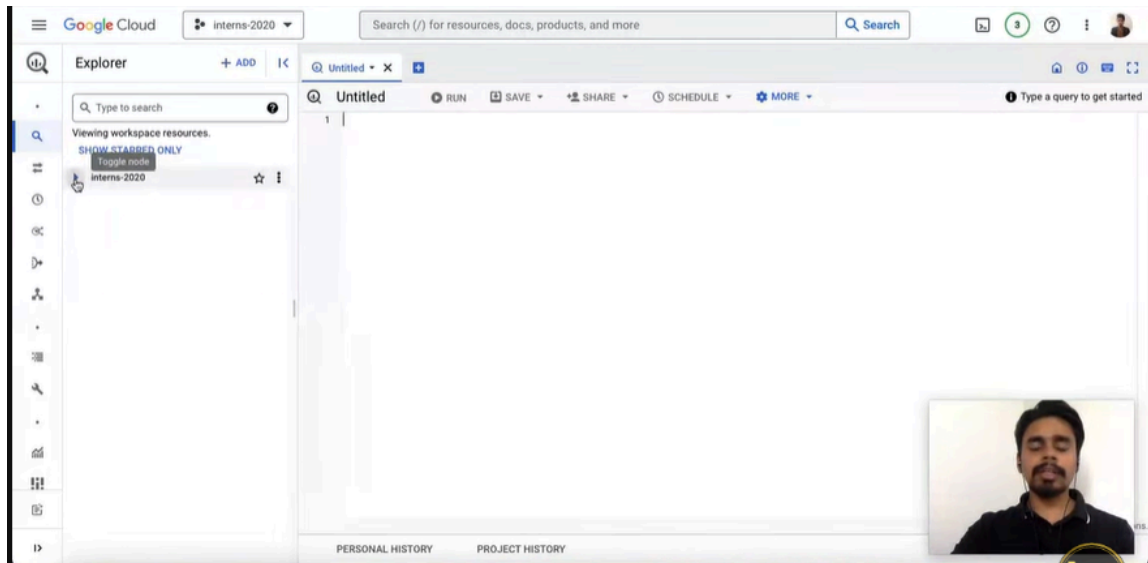


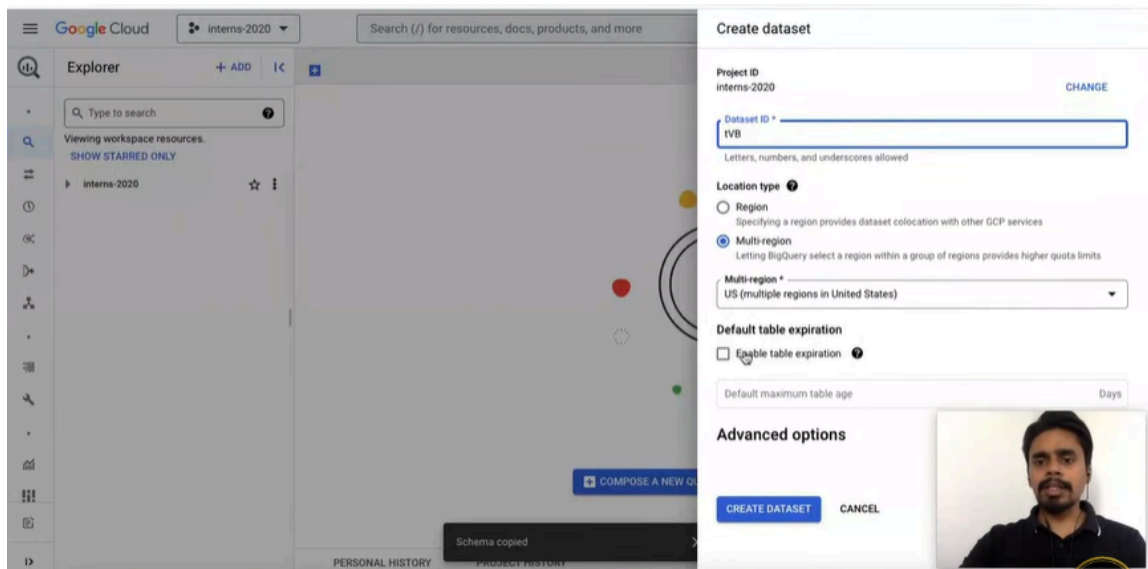
# BigQuery

## Dataset and table

1. First we have to create a project..and open big query
2. Here we have project id('interns-2020')..and inside that we can create datasets and table



3. Inside our project first..we will create the dataset ..give the details and create



4. Here we can see the dataset info

The screenshot shows the Databricks Explorer interface. On the left, the 'Explorer' sidebar lists various datasets, with 'interns-2020.tvb' selected. The main panel displays the 'Dataset info' for this dataset. The information includes:

- Dataset ID: interns-2020.tvb
- Created: May 14, 2023, 14:45 PM UTC+5:30
- Default table expiration: Never
- Last modified: May 14, 2023, 3:14:45 PM UTC+5:30
- Data location: US
- Description: (empty)
- Default collation: (empty)
- Default rounding mode: ROUNDING\_MODE\_UNSPECIFIED
- Case insensitive: false
- Labels: (empty)
- Tags: (empty)

At the bottom right, there is a video feed of a person with a yellow 'V' logo overlay.

5. Inside datasets...we will create tables

6. So to **create table follow this syntax**

```
1 CREATE TABLE 'interns-2020.tvb.employee' (empID int64, empName string);
```

'projectname.datatsetname.tablename'

7. If we click on table we can see the details of tables
8. 2nd method to create table is to use UI ..select dataset and right click and create table by giving col names...we also select partitions if our data is very largen

The screenshot shows the 'Create table' dialog box in Databricks. The 'Dataset' field is set to 'tvb'. The 'Table' field is set to 'customer'. The 'Table type' is set to 'Native table'. The 'Schema' section shows a field named 'id' with a type of 'INTEGER' and a mode of 'NULLABLE'. The 'Partition and cluster settings' section shows 'No partitioning' selected. At the bottom, there are 'CREATE TABLE' and 'CANCEL' buttons. A video feed of a person is visible in the bottom right corner.

- 3rd method is...if we have a json file..then while creating table..click schema and add json code

Create table

Destination:

Project \*  
interns-2020

Dataset \*  
tvb

Table \*  
cars

Unicode letters, marks, numbers, connectors, dashes or spaces allowed.

Table type  
Native table

Schema

☒ Edit as text

Press Alt+F1 for Accessibility Options.

```
6      "description": null,  
7      "fields": []  
8    },  
9    {  
10     "name": "model",  
11     "mode": "",  
12     "type": "STRING",  
13     "description": null,  
14     "fields": []  
15   },  
16   {  
17     "name": "brand",
```

- 4th method..is if we have csv files..then we can directly upload the file
- If we have table and data in google drive..then we can create a table inside bigquery using google drive too

## Save results

- If we want to save our query results..then we can save those in

Google Cloud | interns-2020

Search (/) for resources, docs, products, and more

Explorer

productCategory x

Untitled

```
1 SELECT * FROM `interns-2020.tvb.productCategory`  
2 WHERE brand = "Apple";
```

Query results

Row	product_category	brand	product_name	price
1	Phone	Apple	iPhone 12	1000
2	Phone	Apple	iPhone 12 Pro Max	1300
3	Phone	Apple	iPhone 12 Pro	1100
4	Laptop	Apple	MacBook Air	1200
5	Laptop	Apple	MacBook Pro 13	2000
6	Earphone	Apple	AirPods Pro	280
7	Headphone	Apple	AirPods Max	550

SAVE RESULTS

- CSV (Google Drive)  
Save up to 1GB as CSV to Google Drive.
- CSV (local file)  
Save up to 10MB as CSV locally.
- JSON (local file)  
Save up to 10MB as JSON locally.
- JSONL (newline delimited)  
Save up to 1GB as newline delimited JSON to Google Drive.
- BigQuery table  
Save results as a BigQuery table.
- Google Sheets  
Save up to 10MB to Google Sheets.
- Copy to Clipboard  
Copy up to 1MB to the clipboard.

2. We can also save our query..and execute them

The screenshot shows the Google Cloud BigQuery console. On the left, the 'Explorer' pane displays a list of workspace resources under the 'interns-2020' project, including 'productCategory'. The main editor shows a SQL query: `SELECT * FROM `interns-2020`.`productCategory` WHERE brand = 'Apple';`. A context menu is open over the query, with options 'Save query', 'Save view', and 'Save as...'. The 'Query results' pane at the bottom shows a table with 6 rows and 3 columns: 'Row', 'brand', and 'totalCount'. The results are: Row 1 (Apple, 9), Row 2 (Google, 1), Row 3 (OnePlus, 3), Row 4 (Samsung, 7), Row 5 (Dell, 3), and Row 6 (Microsoft, 2). A small video inset of a person is visible in the bottom right corner.

Row	brand	totalCount
1	Apple	9
2	Google	1
3	OnePlus	3
4	Samsung	7
5	Dell	3
6	Microsoft	2

3. And we'll be having a tab..where we can get saved queries for future use

This screenshot shows the 'Explorer' pane in the Google Cloud BigQuery console. It highlights the 'Saved queries (9)' folder under the 'interns-2020' project. A 'Saved queries' tab is visible on the right side of the Explorer pane.

4. We can also save views...views are mainly used for security purpose..here by using except method..we hidden price column

The screenshot shows the Google Cloud BigQuery console with a query named 'productQuery'. The query is: `SELECT * EXCEPT(price) FROM `interns-2020`.`productCategory` WHERE brand = 'Apple';`. The 'Query results' pane shows a table with 6 rows and 4 columns: 'Row', 'product\_category', 'brand', and 'product\_name'. The results are: Row 1 (Phone, Apple, iPhone 12), Row 2 (Phone, Apple, iPhone 12 Pro Max), Row 3 (Phone, Apple, iPhone 12 Pro), Row 4 (Laptop, Apple, MacBook Air), Row 5 (Laptop, Apple, MacBook Pro 13), and Row 6 (Earphone, Apple, AirPods Pro).

Row	product_category	brand	product_name
1	Phone	Apple	iPhone 12
2	Phone	Apple	iPhone 12 Pro Max
3	Phone	Apple	iPhone 12 Pro
4	Laptop	Apple	MacBook Air
5	Laptop	Apple	MacBook Pro 13
6	Earphone	Apple	AirPods Pro

