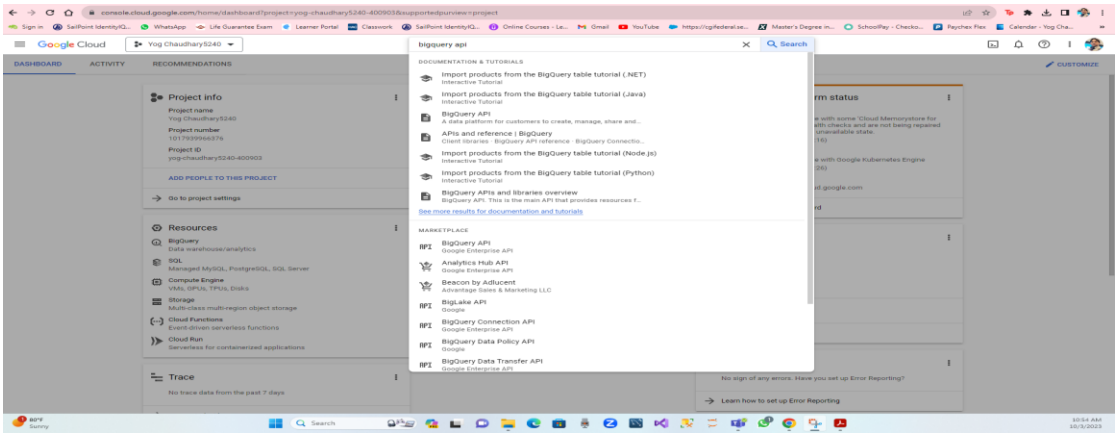
Yog Chaudhary
11727095
ADTA 5240 Week 7'th (harvesting, Storing, And Retrieving Data)
Professor: Dr. Zeynep Orhan
University Of North Texas

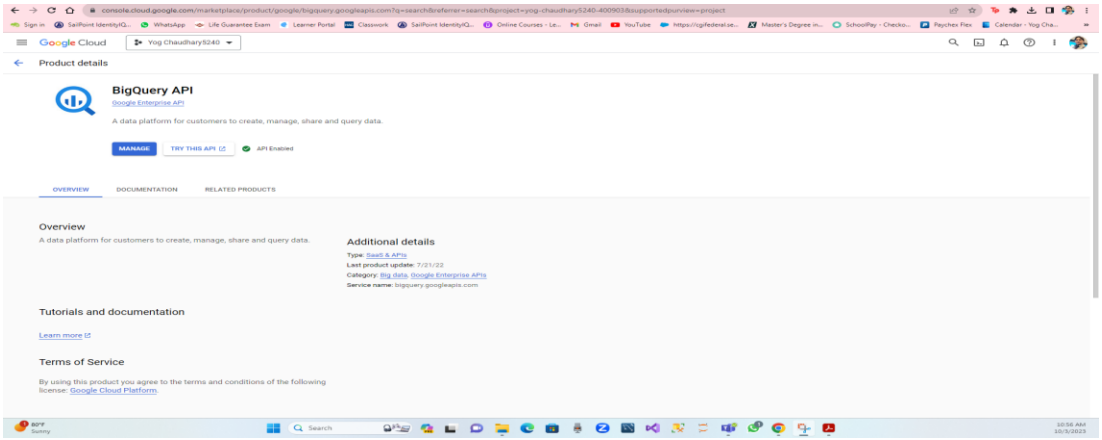
Oct 03, 2023

Creating a dataset and table, then loading data into the table in Big Query PDF.

- ❖ **Big Query in GCP:**
- For doing this we will use Google Cloud Console.
- While we are creating dataset and table and loading data into the table in Big Query, we must enable Big Query API.
- For that I opened google cloud console and checked for whether the Big Query API is enabled or not.
- I typed Big Query API in the search bar and clicked on Big Query API.
- Big Query was enabled.



- By clicking on Big Query API, I go through the below page, where it shows Big Query API was enabled.

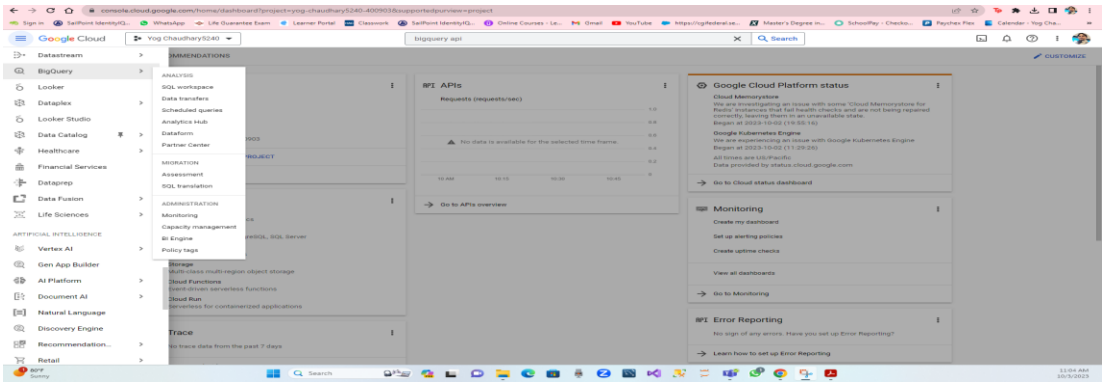


1. Download the data:

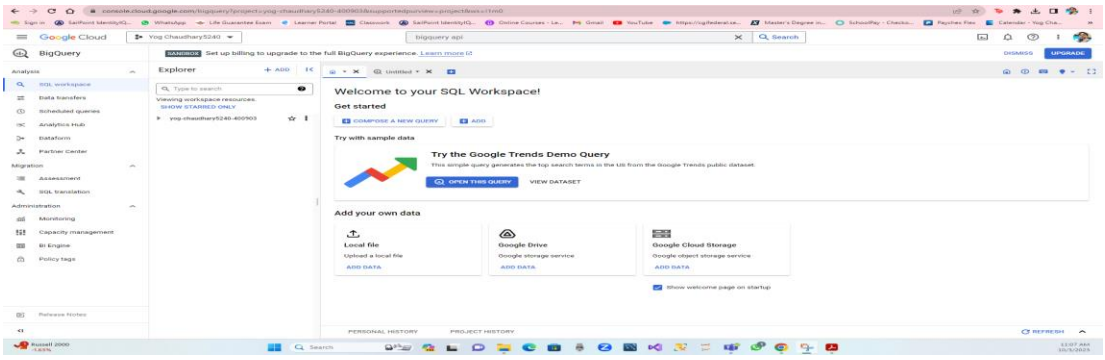
- I have downloaded the names.zip file.
- After that I extracted the file.
- In this zip file, we have a note file, which describes the dataset.
- In this file, we have a name, sex(m/f), and number of children with that name.

2. Creating a Dataset:

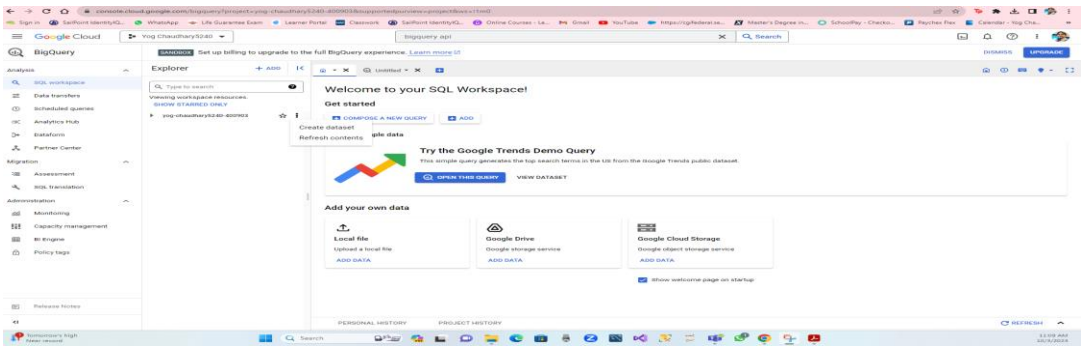
- First, to create a dataset in the cloud console for storing the data, I went through the navigation panel and clicked on Big Query.



- After clicking on a big query, I was taken to this page.

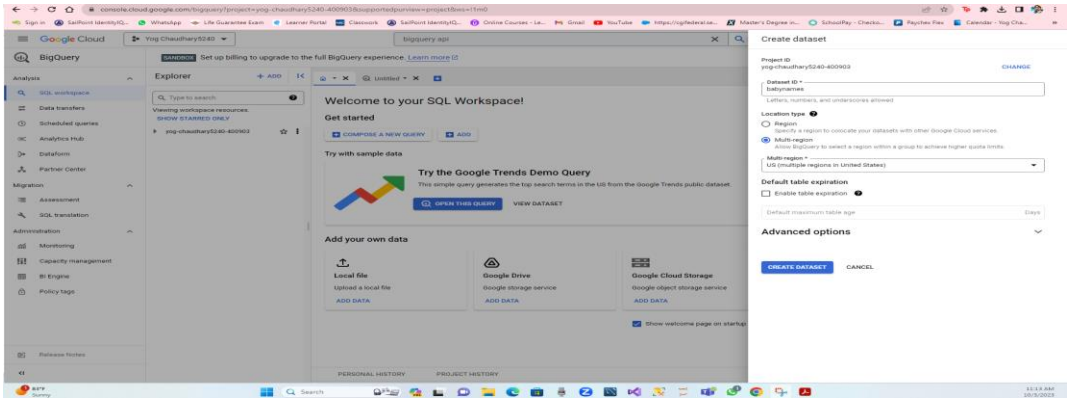


- In explorer panel I clicked on my project name (Yog Chaudhary5240).
- clicked on 3 dots.
- Then I clicked on create dataset.

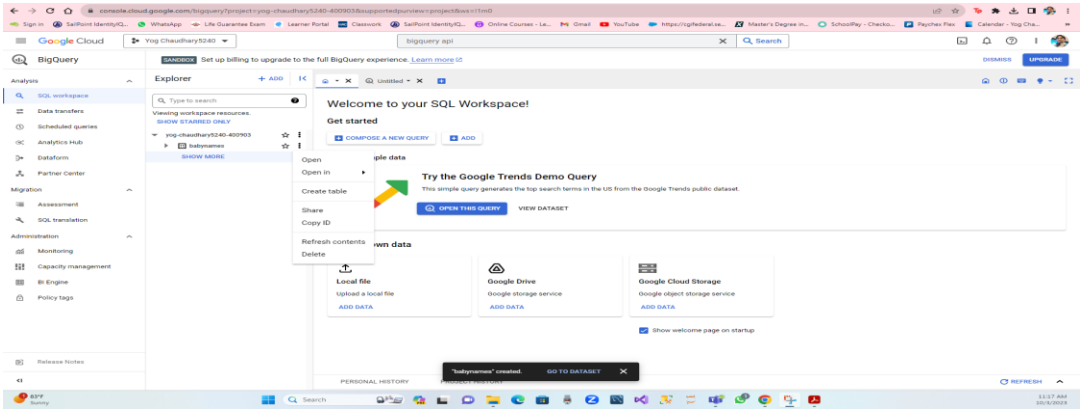


- By clicking on this I was redirected to this page.

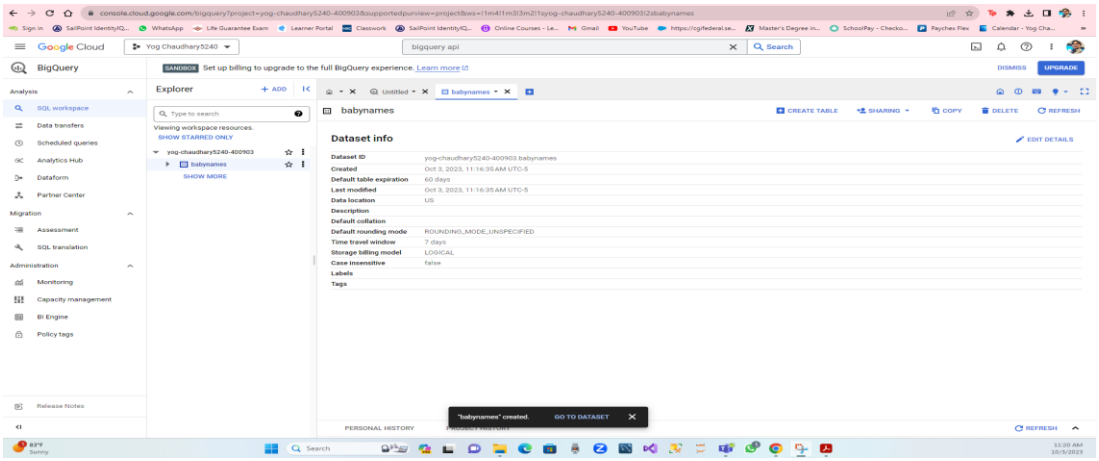
- Here I entered babynames for dataset id.
- Data location as United states (US).
- Then clicked on create data set.



- Then I was taken to this page.

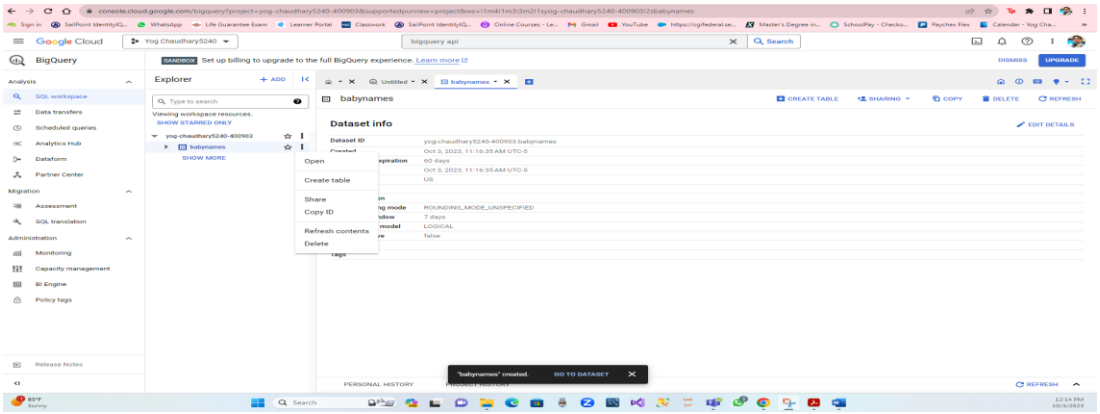


- Then, I Click Open.

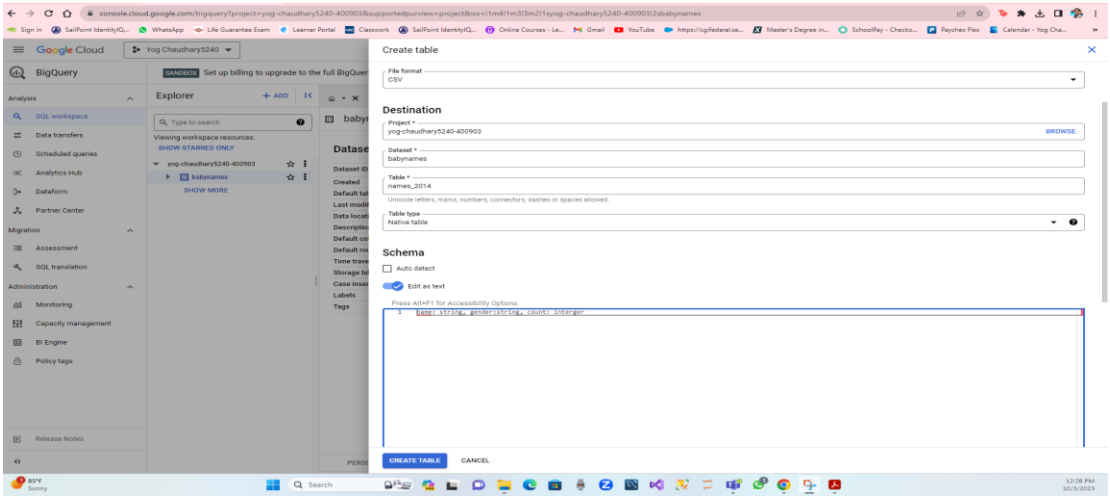


3. Loading the data into a new table:

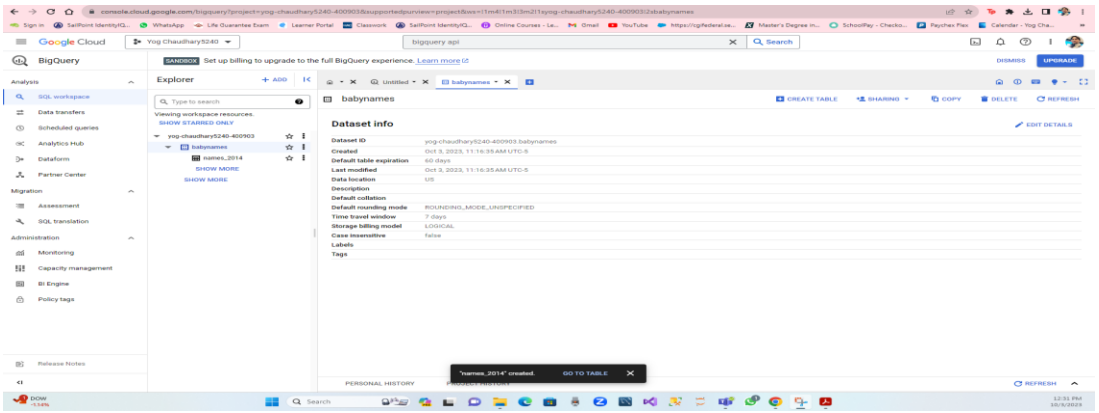
- I must load the data into a new table.
- For this, in explorer panel I clicked on babynames data set that I have created.
- I clicked on 3 dots and clicked on open.
- I was taken to the below page.
- Then I clicked on the create table.



- Hare screenshot



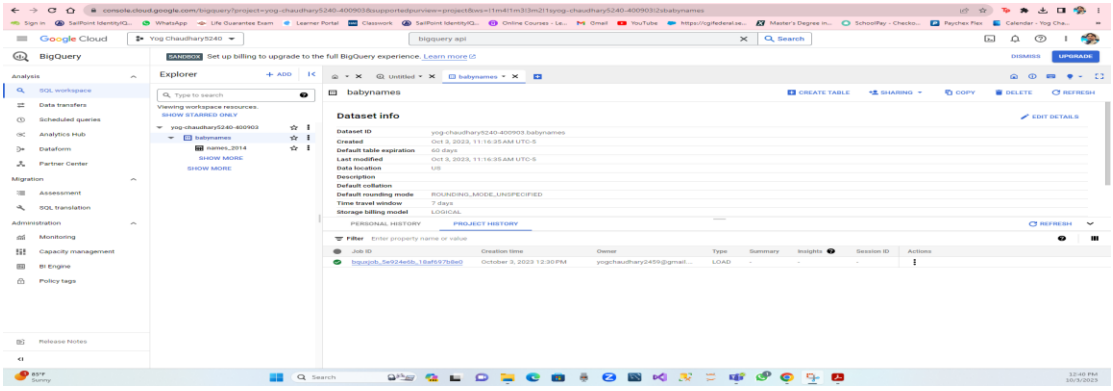
- After clicking on the create I was taken to this page.
- Here I have entered the source section.
- Create table from: Upload.
- Select file: yob2014.txt.
- File format: CSV.
- Destination section:
- Table name: names_2014.
- Schema section:
- In the schema section I clicked edit as text toggle and pasted this schema (name: string, genders: string, count: integer) in the boxes.



- Then I clicked on the create table.

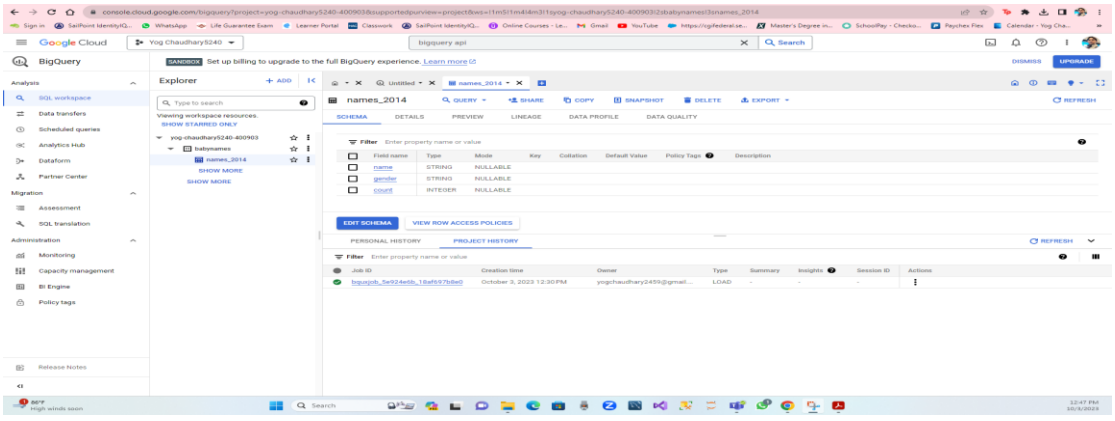
4. I waited for Big Query to create the table and load the data.

- After finishing loading the data, a check mark appeared on the job history panel.
- Here is the screenshot of that.

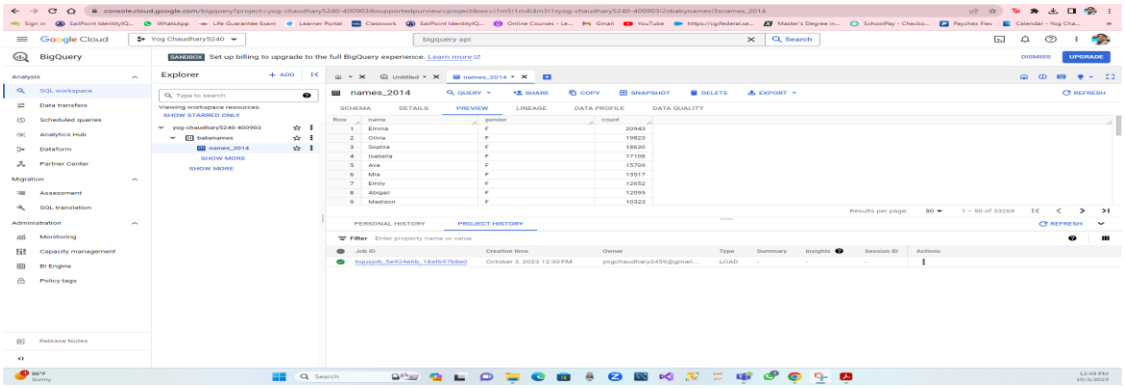


❖ **Previewing the table:**

- For previewing the first few rows of the data, I have followed these steps.
- In the explorer panel, I expanded the baby names and selected the names_2014.
- In the details panel I click on schema, the screen was shown like this.

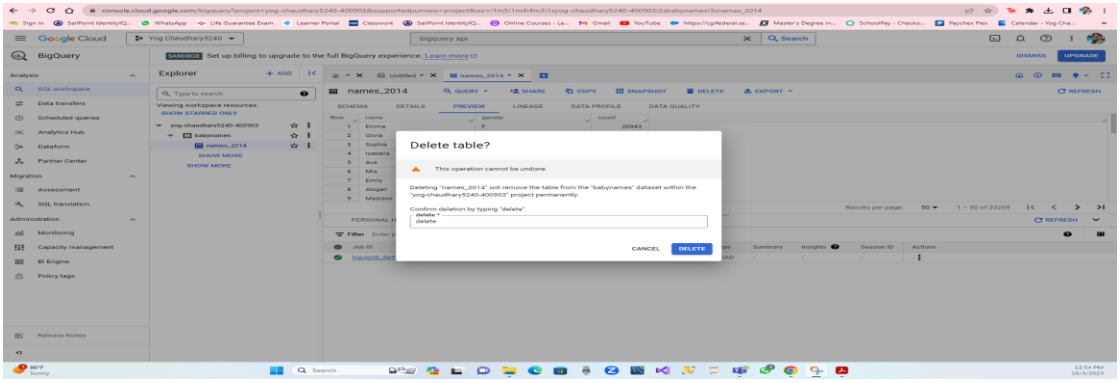


- When I click on the preview in details panel, this Big Query displays a few rows of the table.

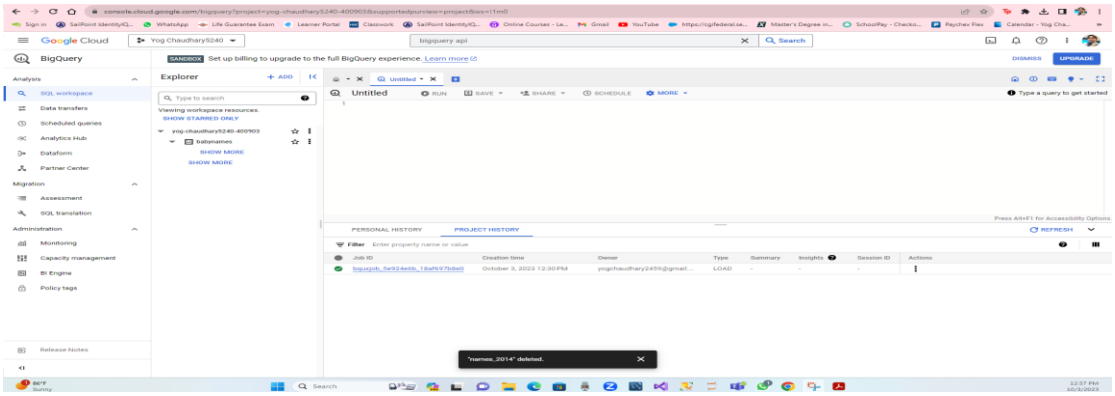


❖ **Cleanup:**

- For avoiding incurring charges of me google cloud account I must do cleanup.
- For this in explorer panel I clicked on **babynames** dataset that I have created.
- When I clicked on **babynames** dataset, I have seen “open” and “delete.”
- I clicked on delete.
- In the dialog box I entered delete and click on it.



- After this dataset, the table, and all the data has been deleted.



- Finally, delete dataset schema names_2014