**Assignment 6: BigQuery on GCP**

**Prerequisites:**

* Should have google account with access to GCP console

**Task 1: Creating dataset**

1. Since we are using BigQuery for the first we need to enable BigQuery API from GCP marketplace

Graphical user interface, text, application, email

Description automatically generated

**Graphical user interface, text, application, email

Description automatically generated**

1. Now we can create dataset in BigQuery. First, click on hamburger icon on the top left side of the screen, scroll to the option BigQuery and click on it so that you can navigate to BigQuery service on GCP

Graphical user interface, application

Description automatically generated

1. Under Explorer section you find project name, just click on 3 dot icon to find create dataset option

Graphical user interface, application

Description automatically generated

1. Give the dataset ID(I gave as babynames), select location as US and click on CREATE DATASETGraphical user interface, application, Teams

   Description automatically generated

You can see dataset info here

Graphical user interface, text, application, email

Description automatically generated

**Task 2: Creating table and loading data into it**

1. At first, we need data to load into the table. To do so, we can download the data file from US social security administration website - https://www.ssa.gov/oact/babynames/limits.html.
2. The file downloaded is of compressed format hence we extract it and select any of the text file as data to be loaded into table. I selected **yob2021.txt**

Graphical user interface, application, table

Description automatically generated

1. On dataset information screen you will find CREATE TABLE icon towards right, click on it to create new table under babynames dataset

Graphical user interface, text, application, email

Description automatically generated

1. Give the table configurations which are required and keep other settings as is. We are now going to create table with above selected data file. Here, we upload data while creating the table. Once you set-up all settings, click on CREATE TABLE

Graphical user interface, application

Description automatically generated

1. After few moments, table creation and data load will complete you can see job details like below

Graphical user interface, application

Description automatically generated

**Task 3: Preview the table and data rows**

1. On the explorer window, you can click on **names\_2021** table then click on PREVIEW button

**Graphical user interface, application, Teams

Description automatically generated**

1. Table preview will be shown as below. Can you see total 50 rows present in the data

Graphical user interface, application

Description automatically generated

1. You can also see schema of the table by clicking on SCHEMA button

Graphical user interface, application

Description automatically generated

1. You can also query data by clicking on QUERY button and write SQL query to fetch data from table( I have written the query: SELECT \* FROM `adta-5240-sri-charan.babynames.names\_2021 LIMIT 4). Here I have put limit as 4 as an example. Running your queries with limit clause will save few bucks

**Graphical user interface, application

Description automatically generated**

**Note:** Make sure you delete dataset once activity is done to avoid unnecessary charges.