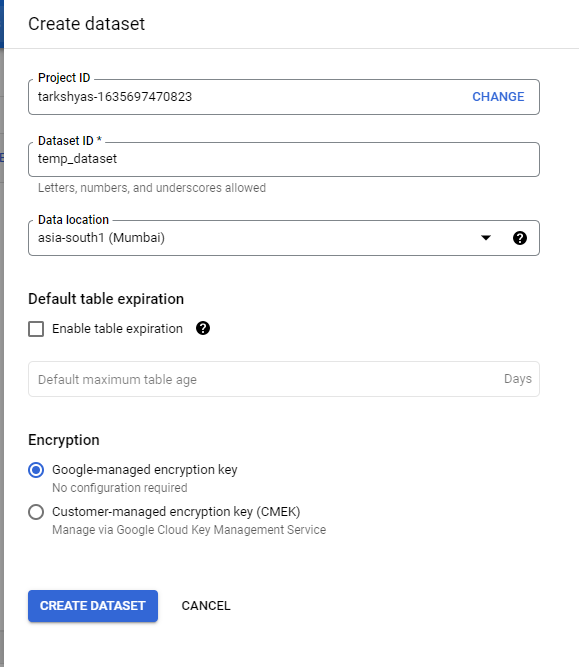
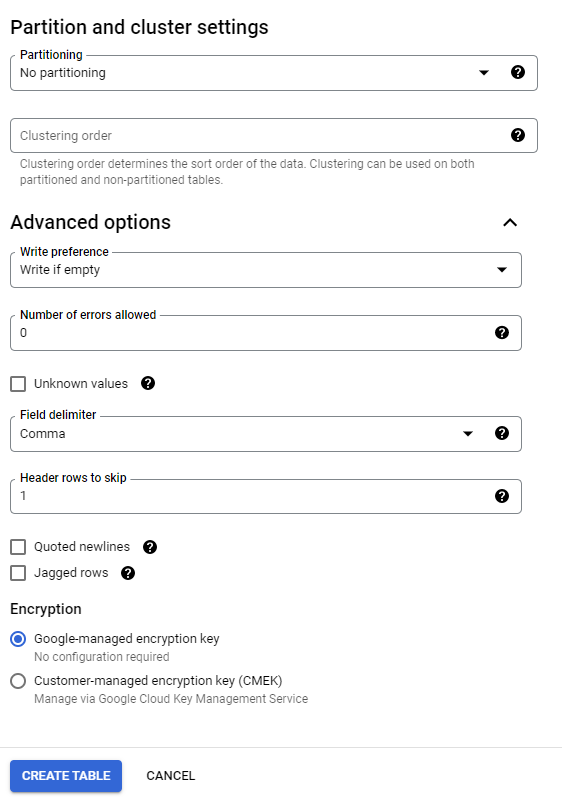
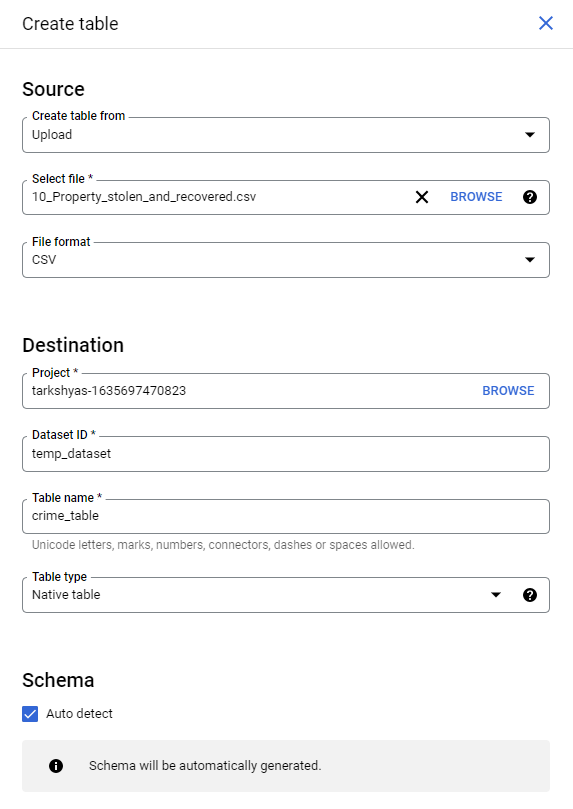
**12) BigQuery concepts**

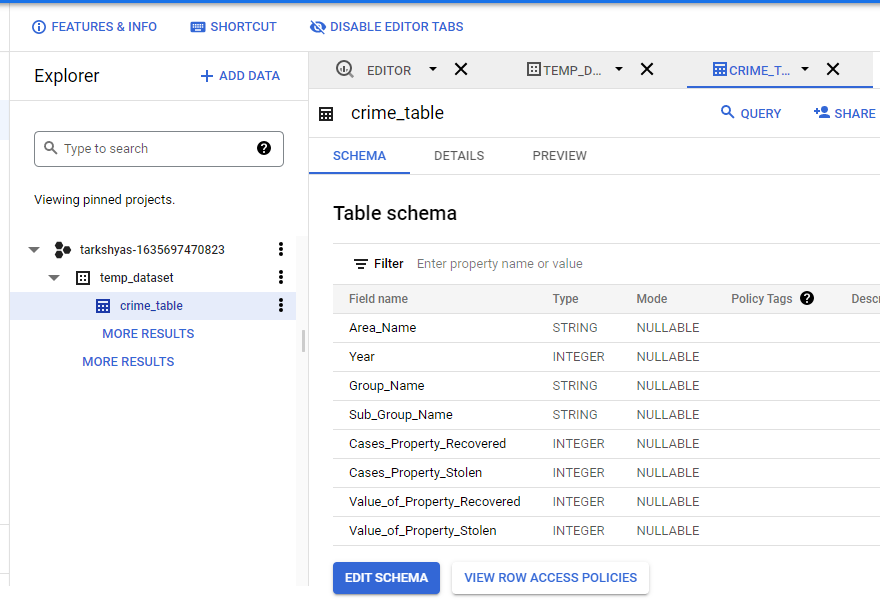
Step 1 : Create a dataset inside project first



Step 2 : Create a table inside the dataset and upload csv file which contains data. Also we can upload files from cloud storage too.

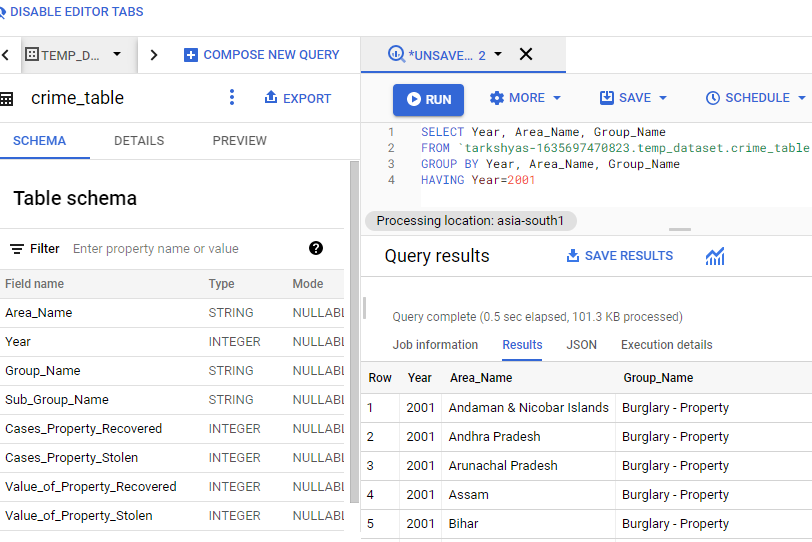


Step 3 : Dataset and table created are,

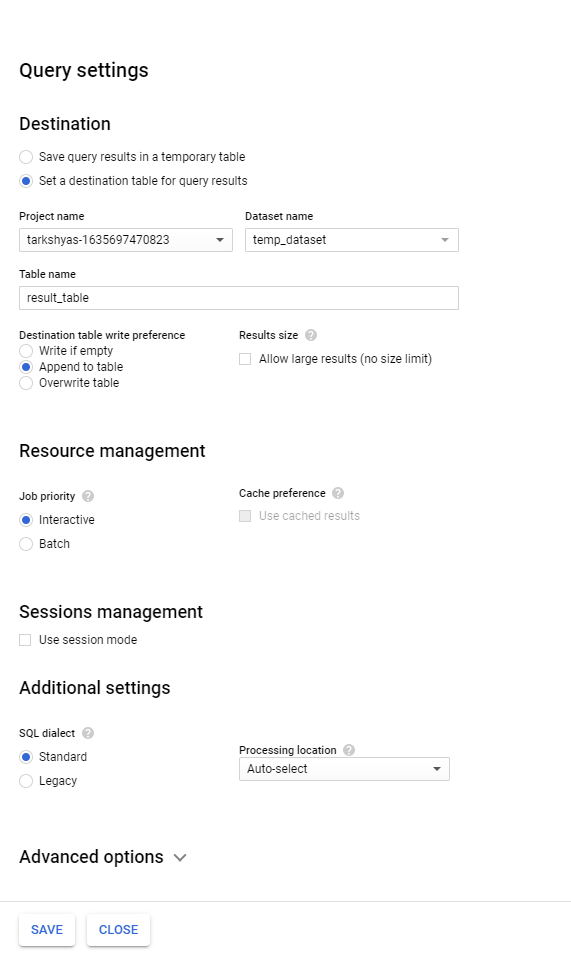


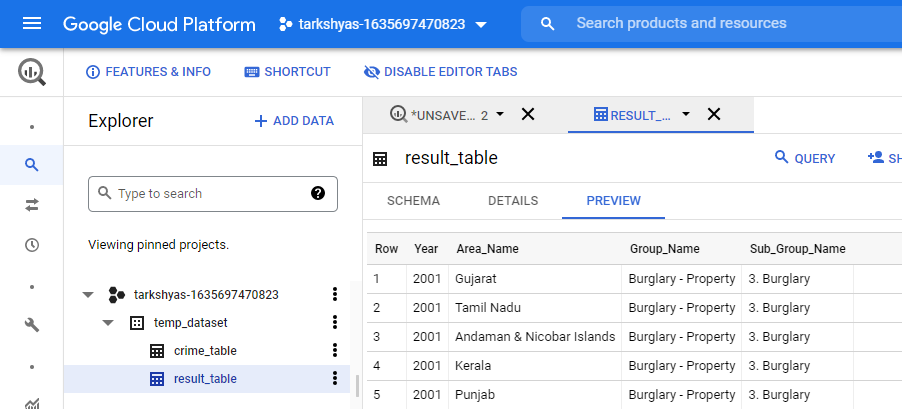
**-- GROUP BY & HAVING --**

Step 4 : Run the query to group data by year,area\_name and group\_name in the year 2001

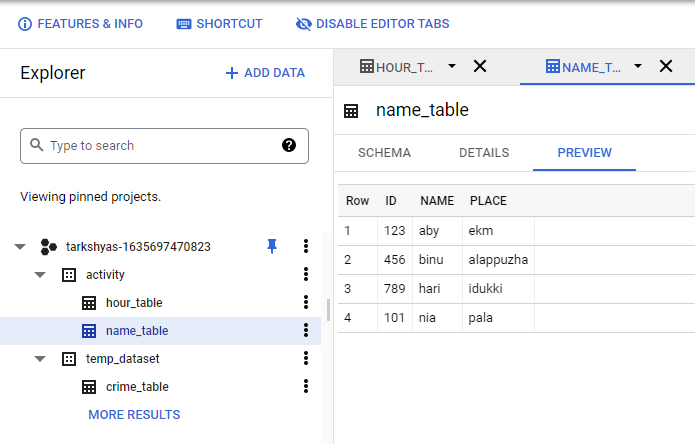


Step 5 : This query result can be stored in a new table ie, result\_table

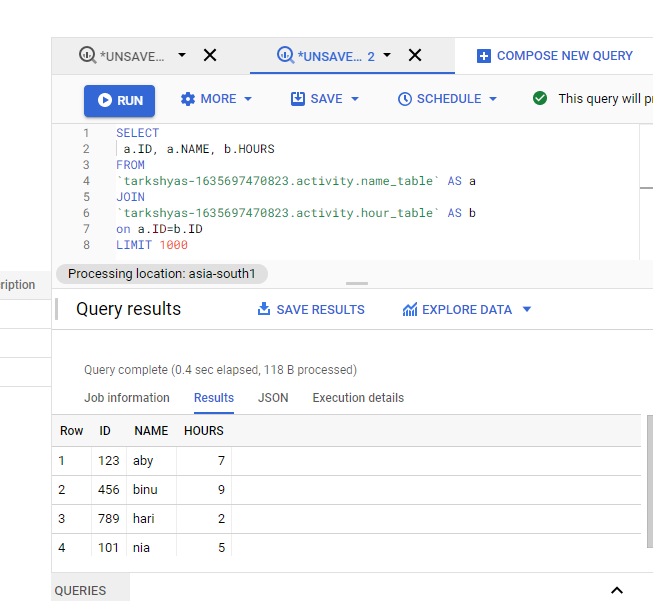




**-- JOIN QUERY --**

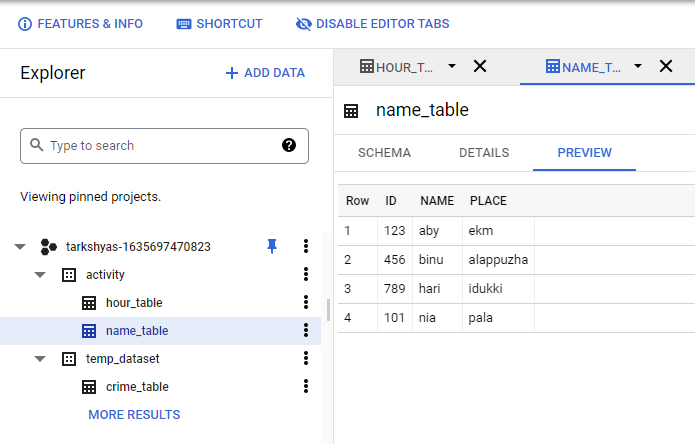
Step 6 : Create name table and hour table and join both tables

Step 7 : Resultant table after joining

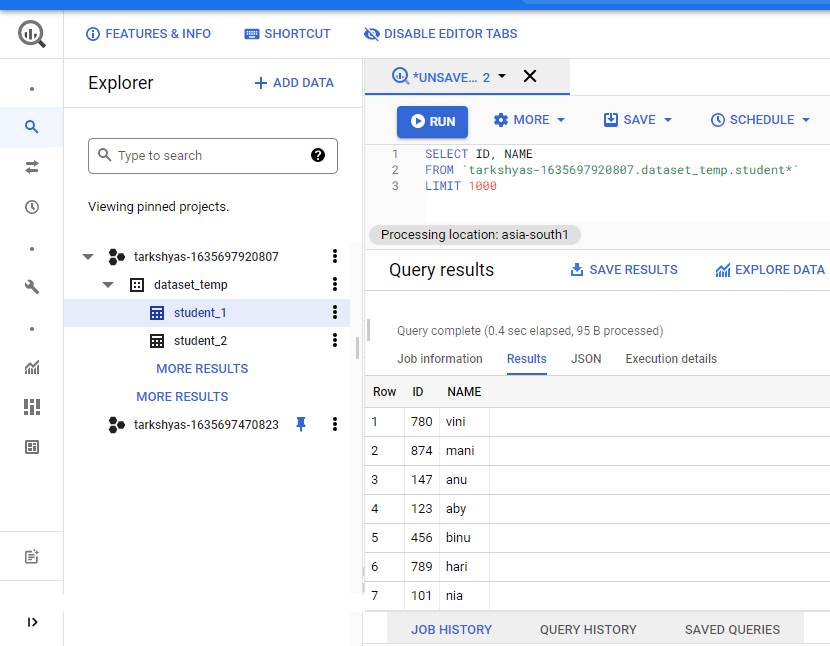


**-- WILD CARD QUERY –**

Step 8 : Create two tables with common names such as student\_1 and student\_2

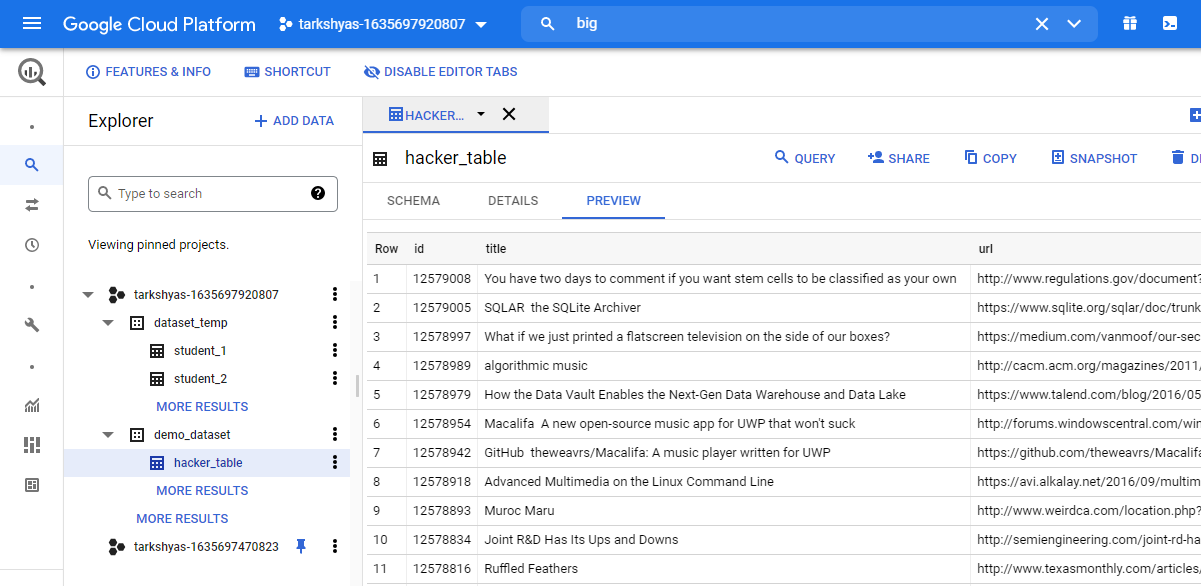


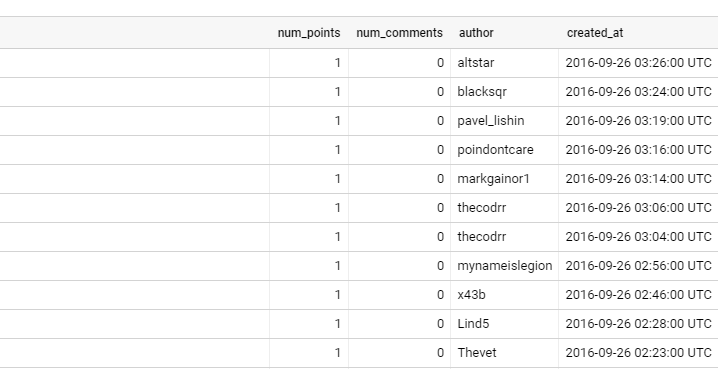
Step 9 : After executing the query the resultant table will displays all the contents in both tables.



**-- PARTITIONING & CLUSTERING –**

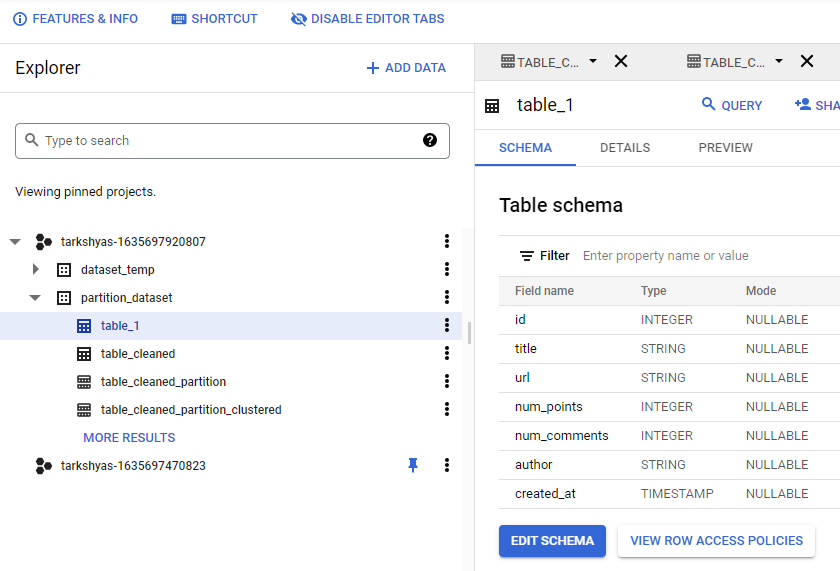
Step 10 : Create dataset and table , then upload file to be partitioned . Here the hacker table contents are :

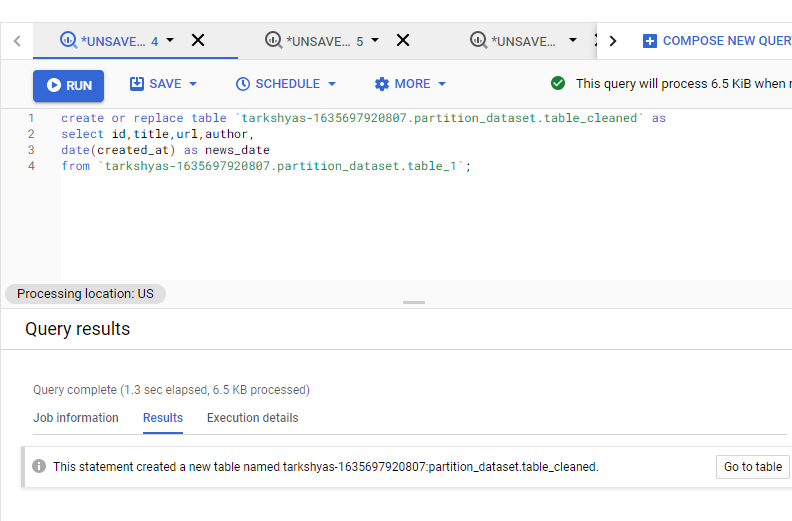




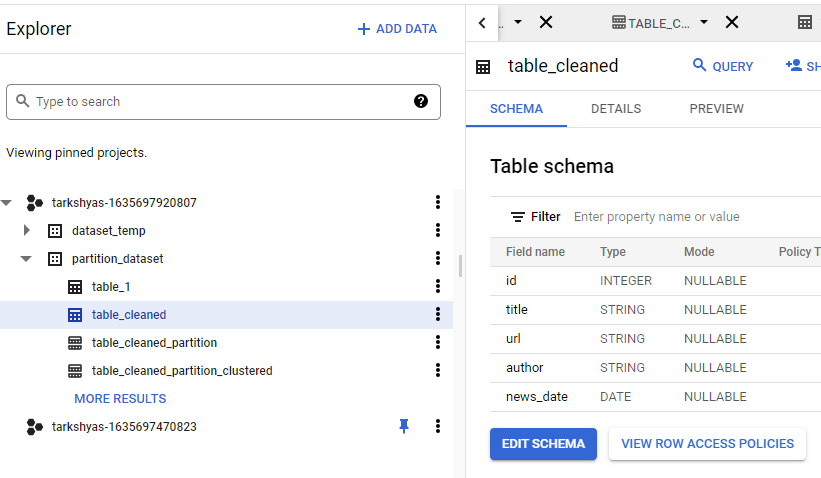
Step 11 : Clean that table so that we need to avoid the time in the row created\_at and filter the date only

***Table\_1 before cleansing***

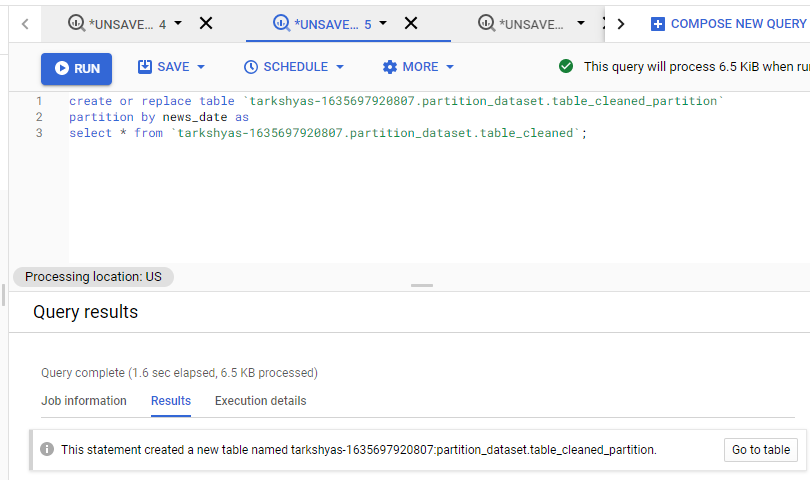




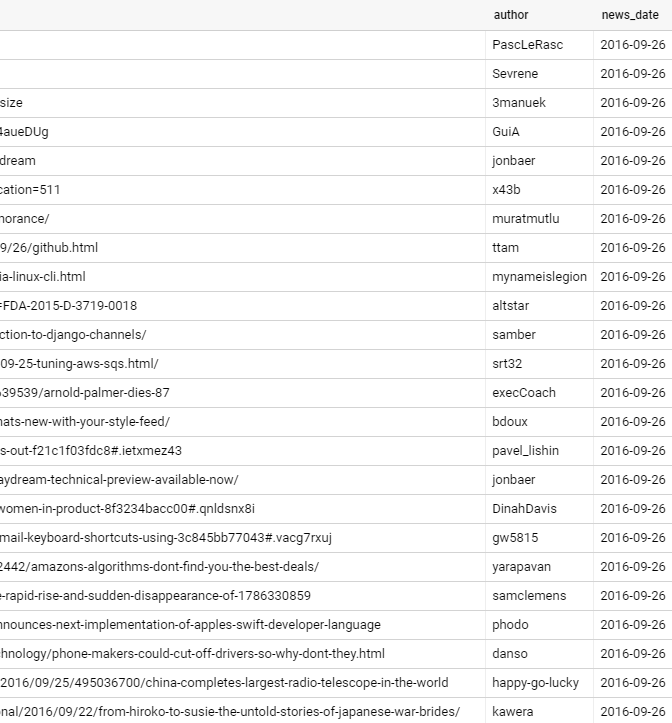
***Table\_cleaned***

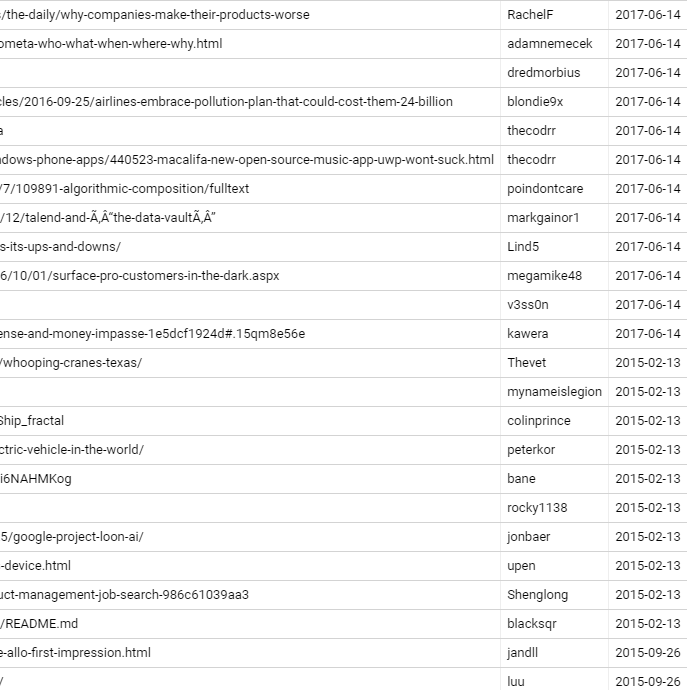


Step 12 : Partition the table based on their news\_date so that tables are divided into different groups with similar date

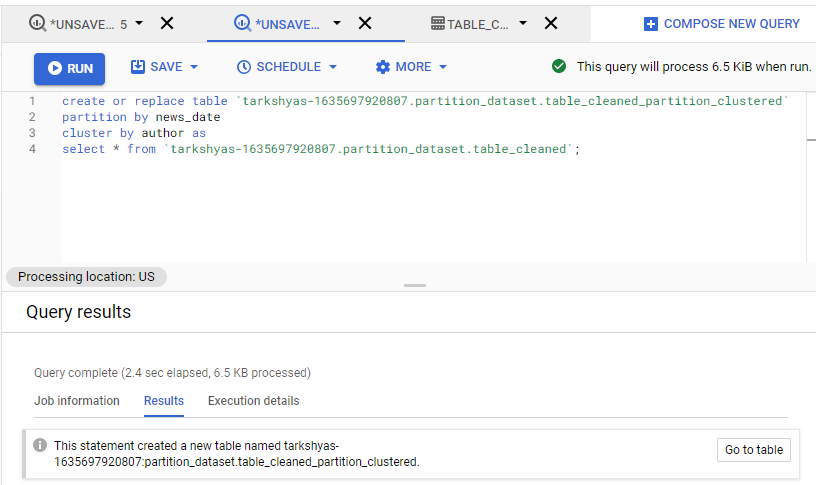


***Partitioned\_table***

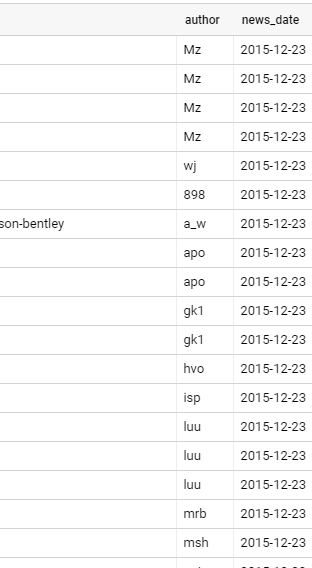
******

******

Step 13 : Group together the data with similar author in the partitioned table using cluster query.



***Clustered\_table***

******