Step1:

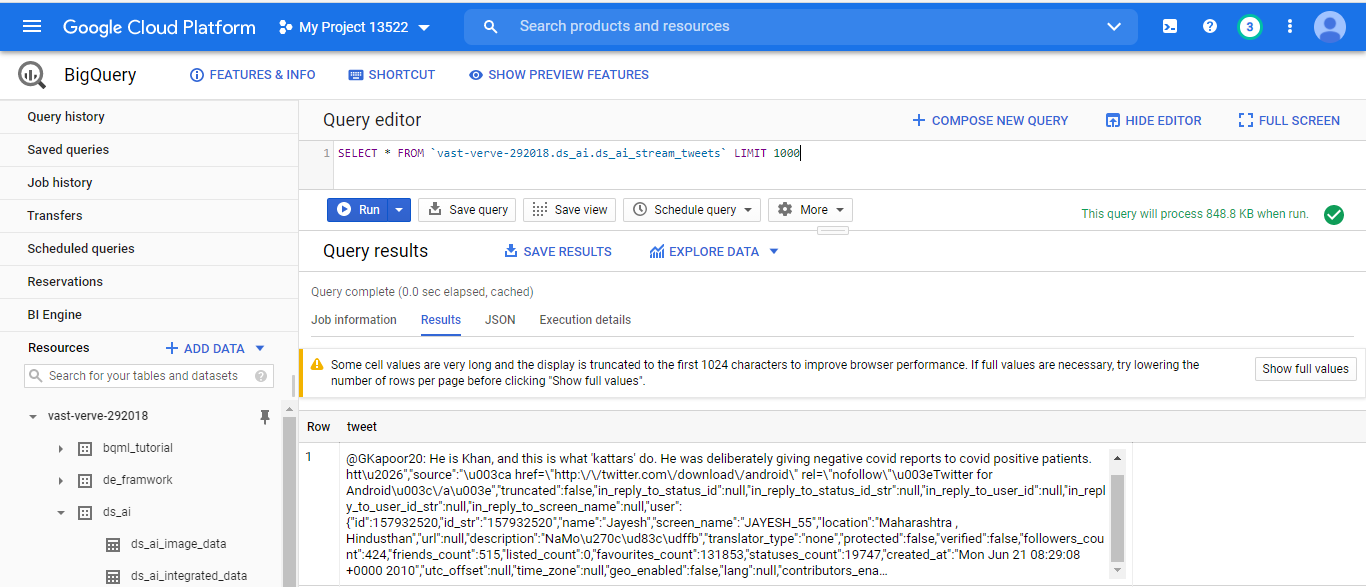
Loading Streaming tweets(with some search term) into Bigquery table .

Requirements:

Library – tweepy

Tokens – Twitter api keys

Storage – Bigquery Table

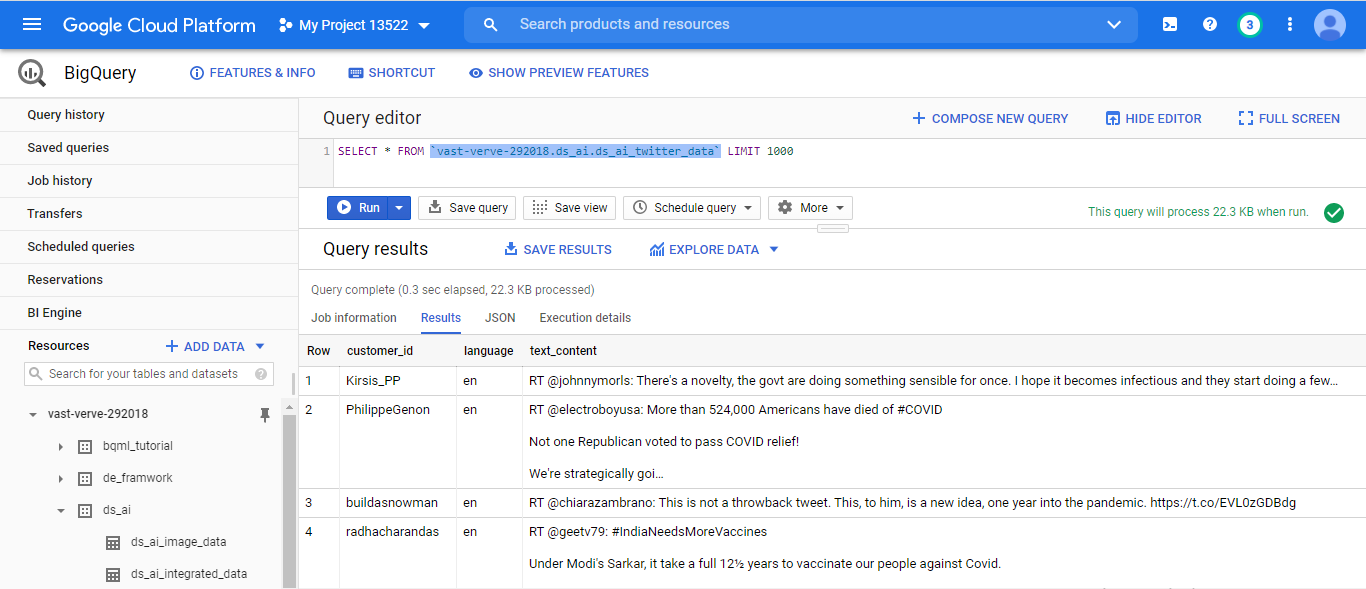


Step2:

Parsing tweets from above table and getting the required fields from tweets and load it into another Bigquery table.

Requirements:

Library – opencv used for age,gender detection from profile image

Storage – Bigquery Table

Step3:

Retrieving Youtube data using Youtube data API from GCP Service and Loading the video meta-data into Bigquery table.

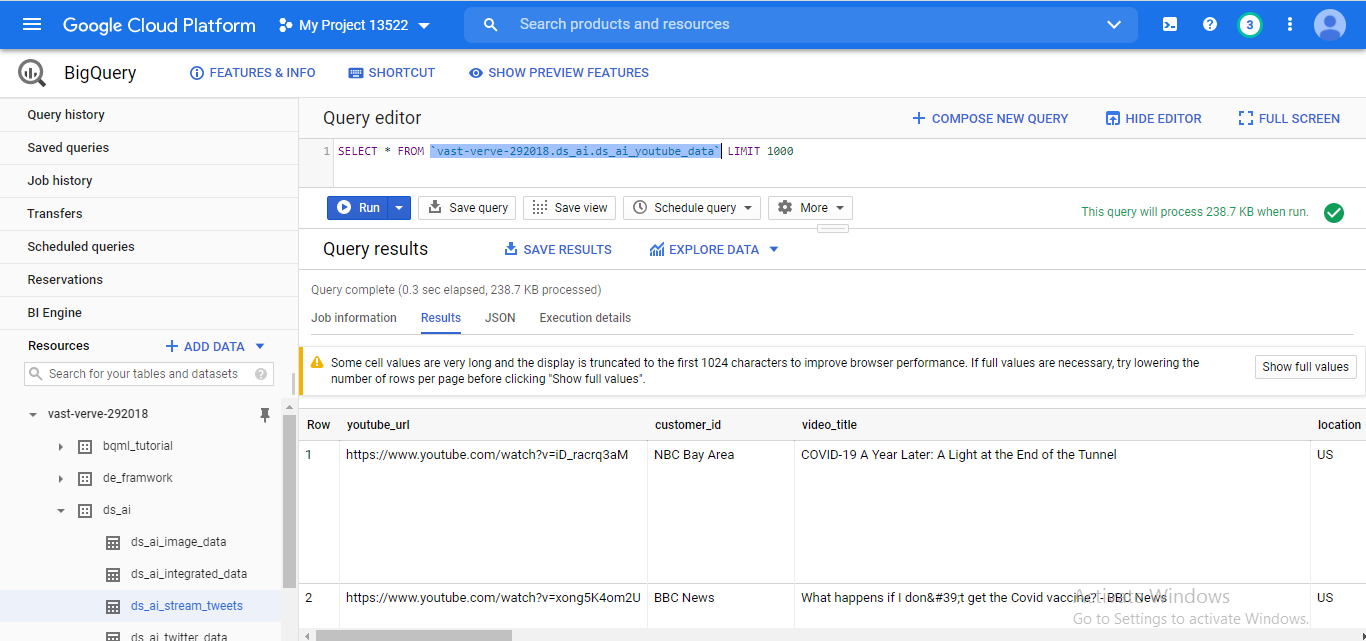
Requirements:

API’s:

– Youtube data API for youtube data retrieval

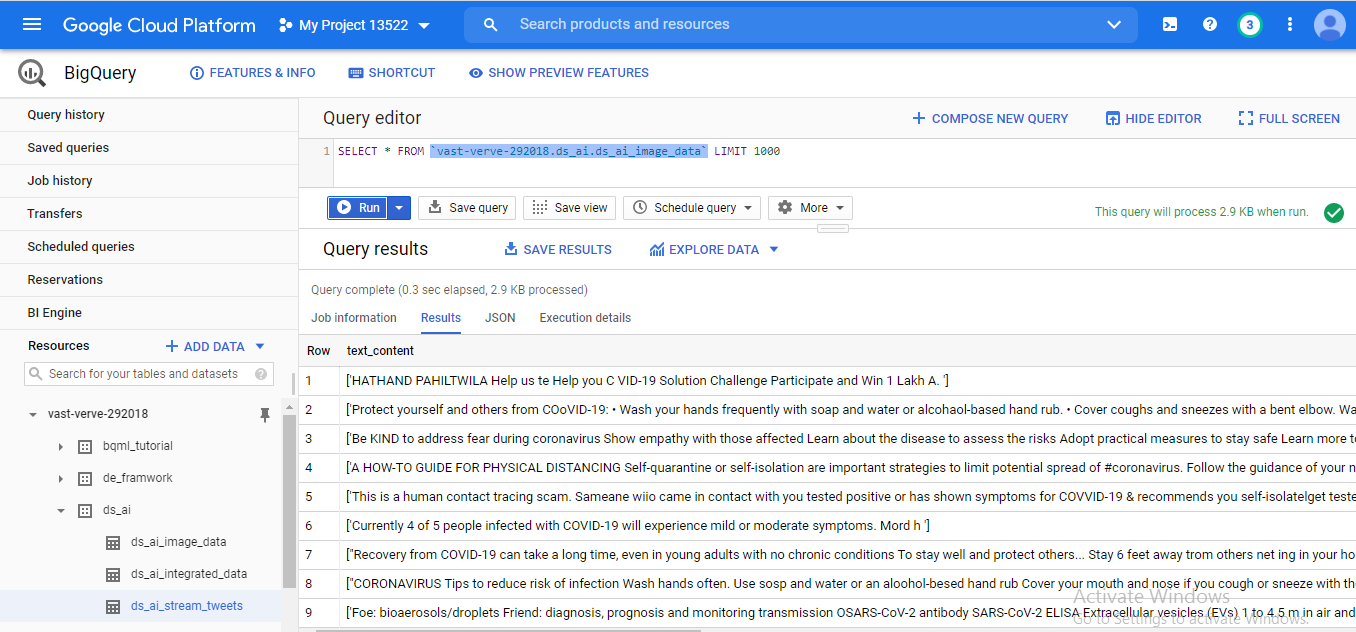
-- Video Intelligence API for speech transcription, label detection

Storage : Bigquery Table for meta data, Google Storage for actual video content



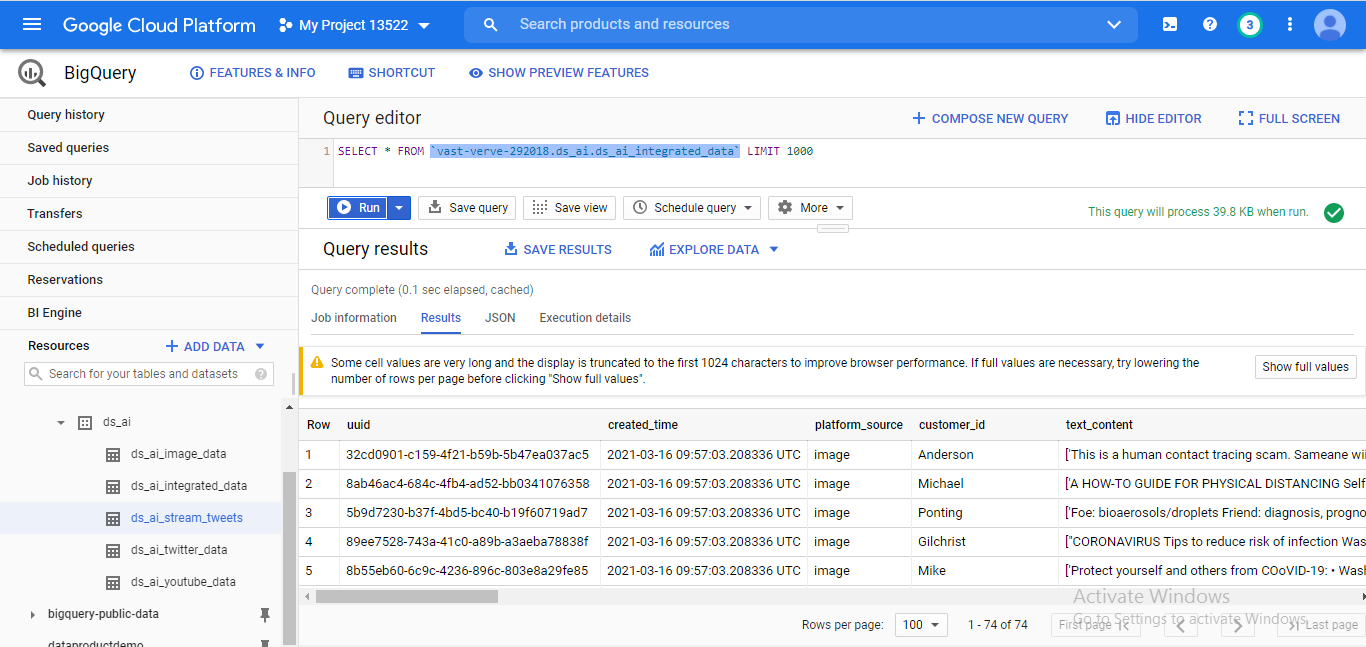
Step4:

Loading Image data into cloud storage. And extracting text from it and simulate customer,age,gender for the image data. Finally, loading this data into another Bigquery table. Here, Cloud Vision API Library used for text extraction from image file.



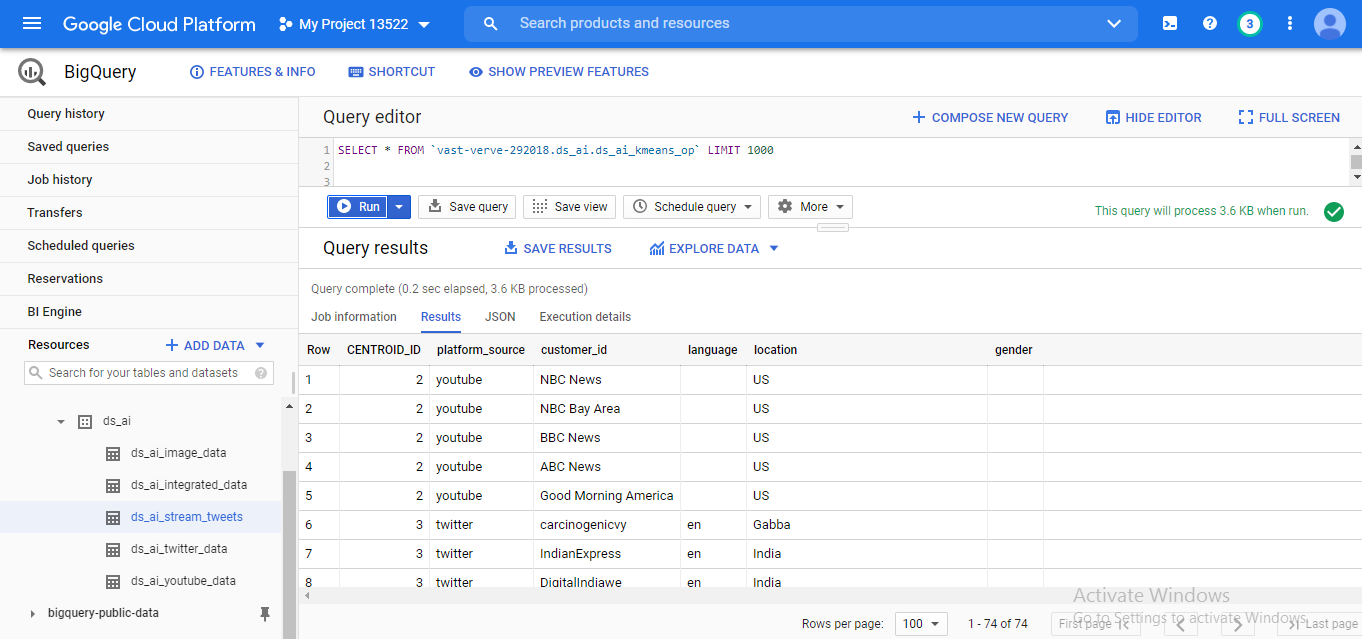
Step5:

Integrating all tweet,youtube,image data into single bigquery table.



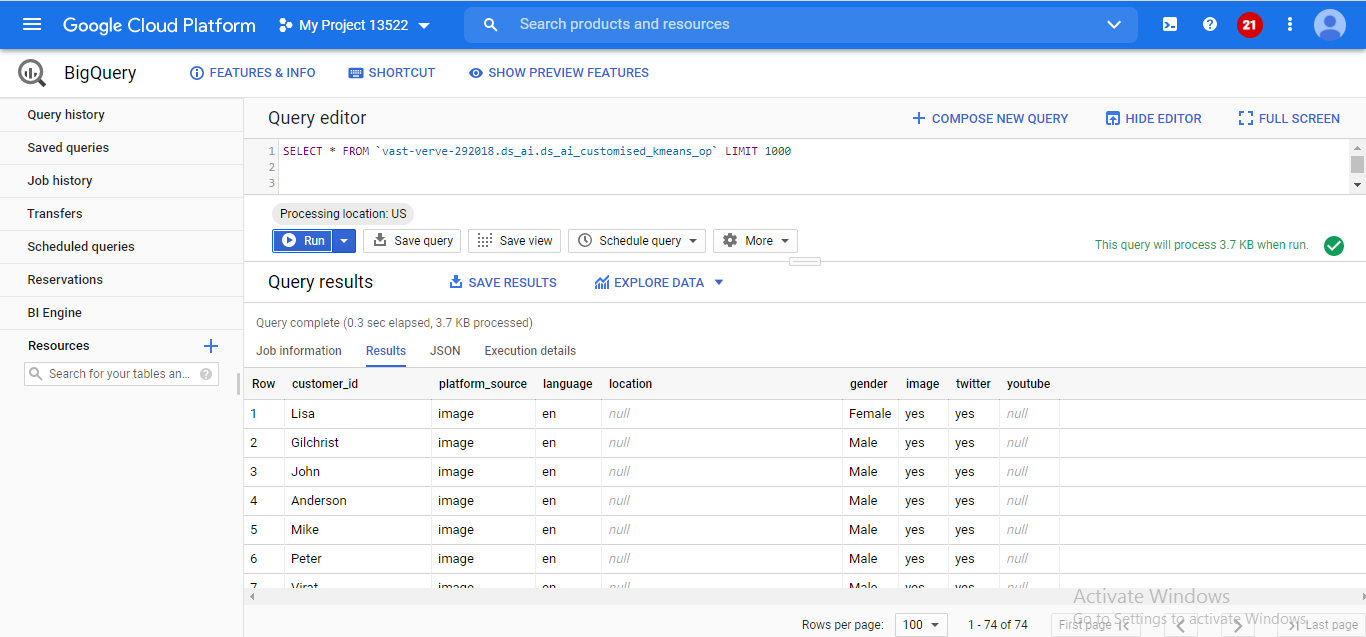
Step6:

Applying Kmeans Model for customer grouping with the help of Bigquery ML. Then the model predicts the customer group for given features.



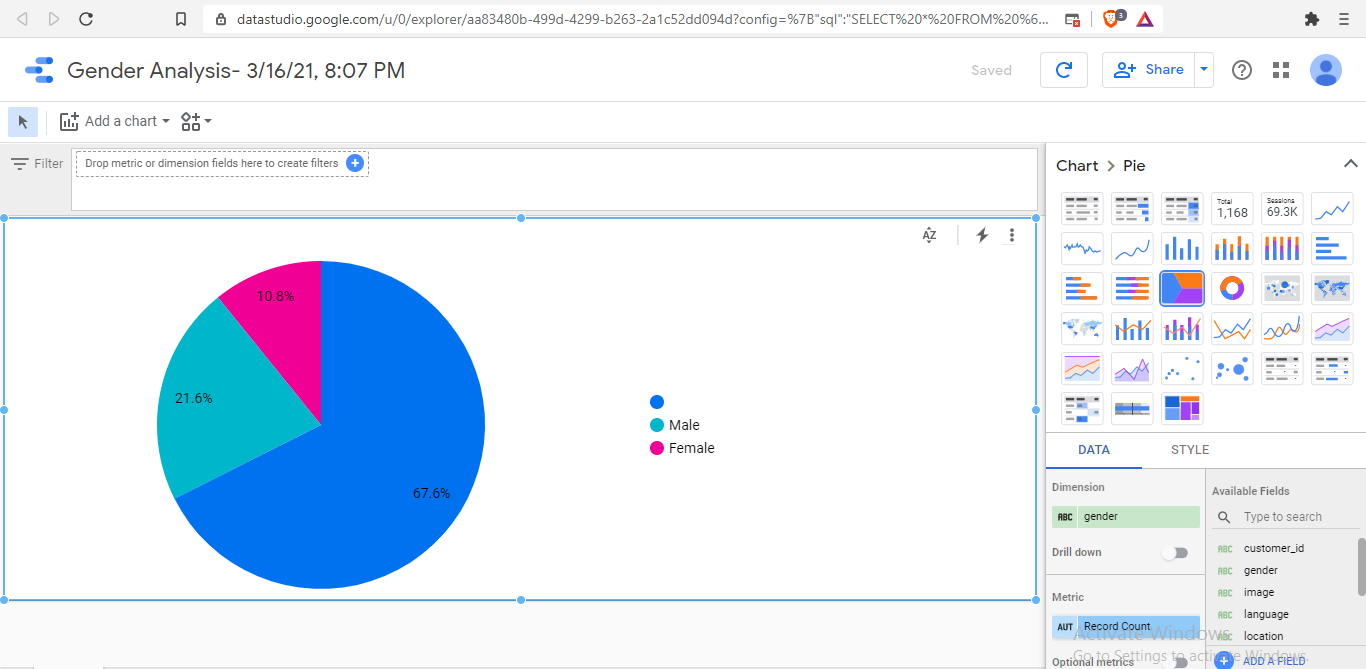
Step7:

Finally, we are customizing the kmeans prediction, based on the centroid value.



Step8:

Some analysis with Google Data Studio – Gender analysis:



Step9:

Common User Profile Across Social Media: User Access Trends

