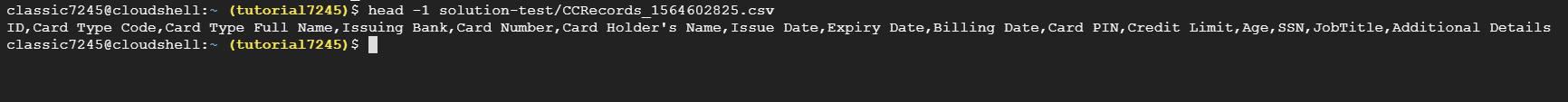
Assignment 2: Part 1: ML as a service(C)

# Migrate Sensitive Data in BigQuery Using Dataflow & Cloud DLP

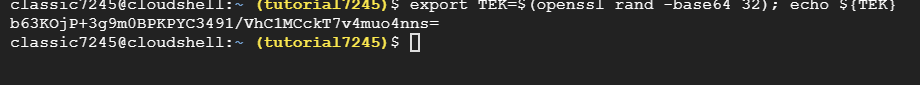
# 

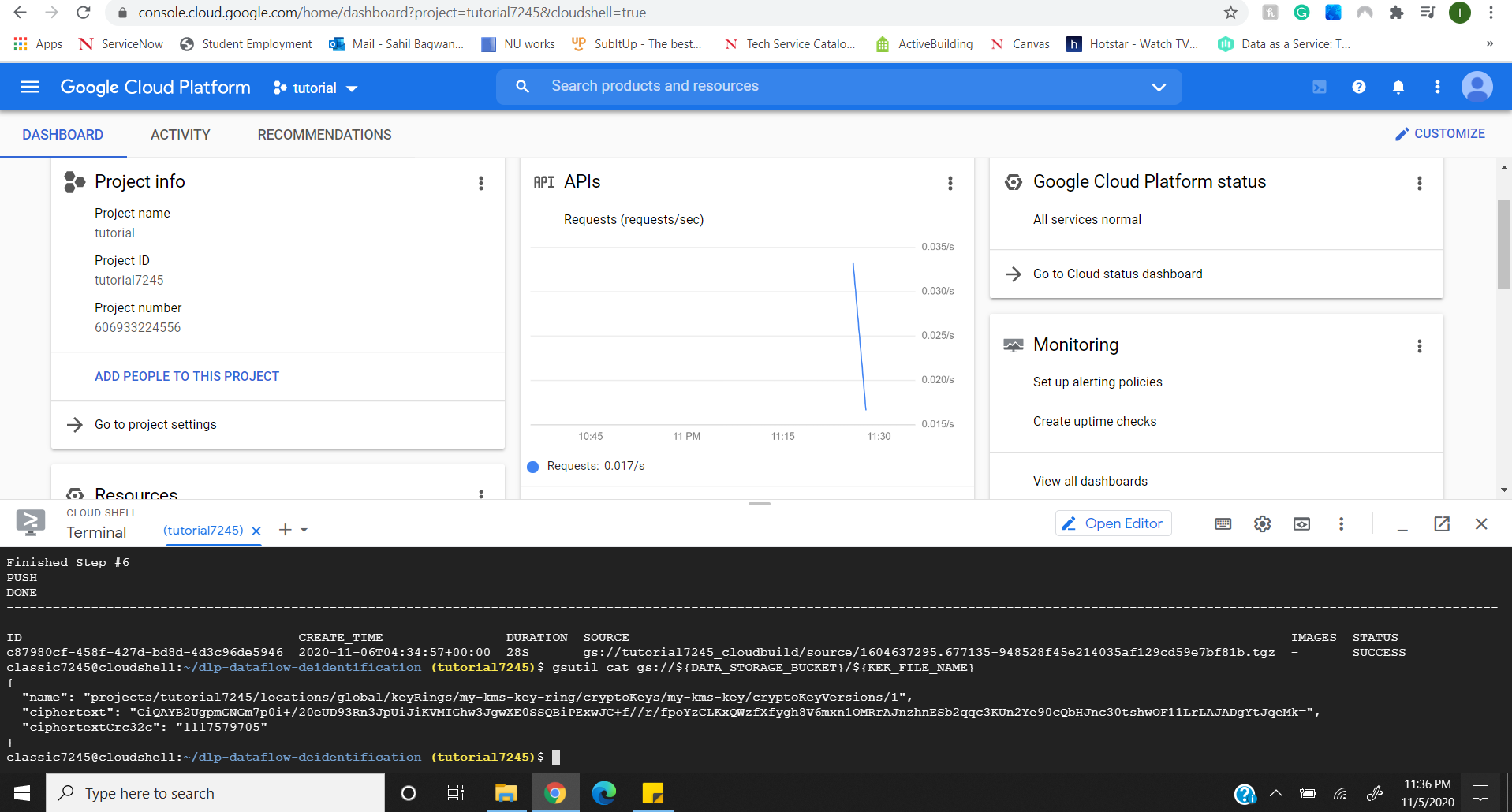
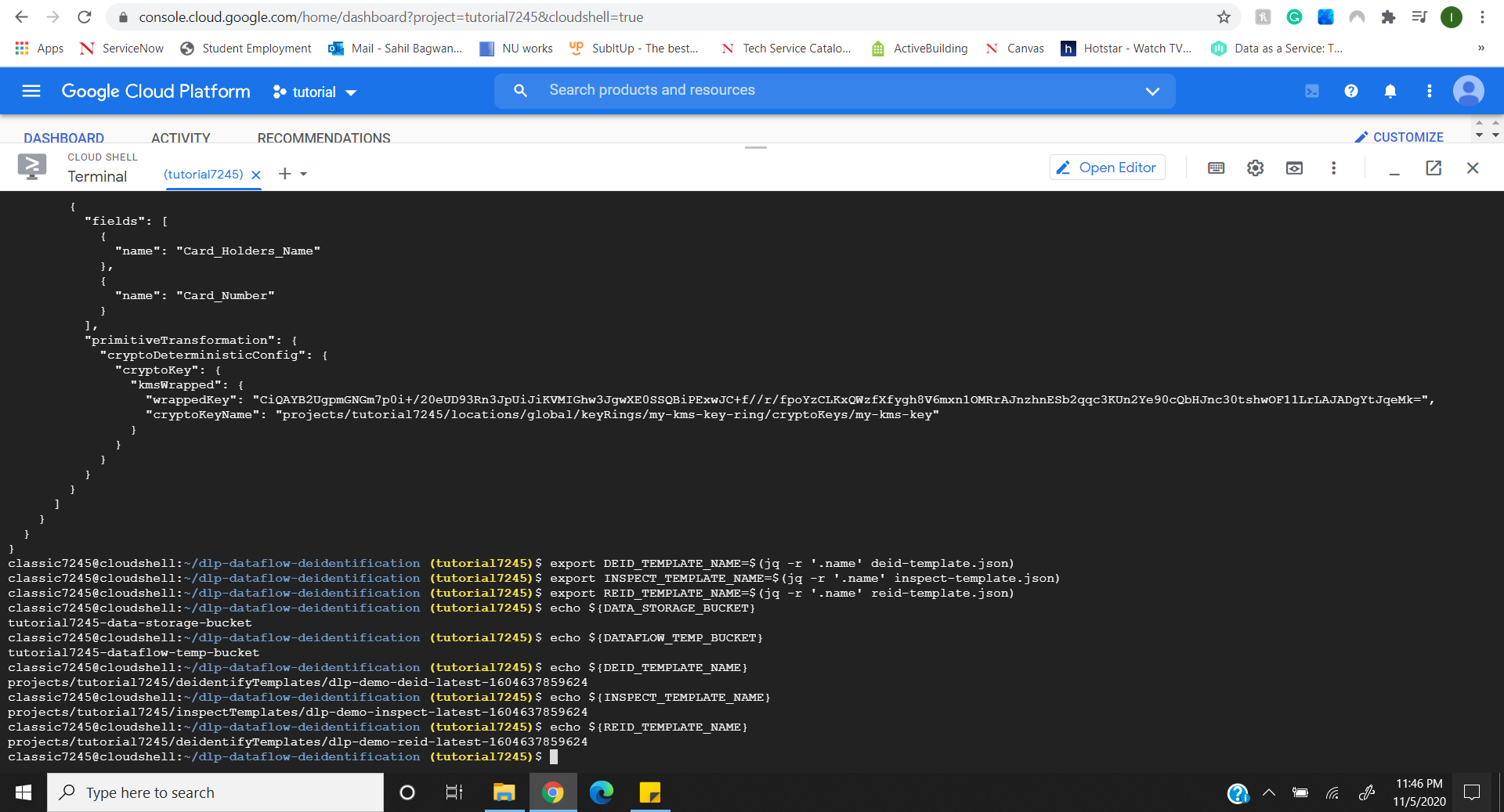
Steps:

1. Create 2 buckets,the first bucket stores the sample dataset and the second bucket stores temporary data for the automated pipeline.
2. Decompress the contents and validate the files downloaded.
3. Identify which columns might require DLP de-identification by examining the header record in the CSV file.



1. Create a BigQuery dataset and create an encryption key.



1. Export the key, key ring, and KEK file as variables and Enable the Cloud KMS admin and key encrypter roles for the Cloud Build service account.
2. Create a KEK and Validate it.
3. Activate the service account.
4. Create the identification and de-identification template, validate it and export the template names.
5. Run the automated data-flow pipeline to de-identify the dataset

