**Environment Replication**

**Components**

* Input file to simulate proprietary DB of the data source
  + This represents the final point of the client-provided Dataflow and is the fraud detection can be inserted.
  + Input file will contain the originally expected information of the client.
  + This means the customer id, item id, price, time, and process time.
  + The file will be a simple text file or CSV file for simplicity.
* Big Query table with the additional column to contain fraud score
  + This represents the destination table once fraud scores have been factored in
  + This was created manually on GCP.
* VM instance to launch dataflow job from
  + Has same specifications as what is mentioned by the client

**Solution Documentation**

Solution is provided in 2 parts: Modified client code and Demo Code

* Modified Client Code
  + The modified client code contains and additional DoFn class as well as the additonal pipeline step to add in the fraud score.
  + It is based on the demo code
* Demo Code
  + The demo code reads data from a csv file, adds a random float from 0 to 1 as the fraud score and pushes the data onto a BigQuery table.
  + This is to demonstrate the functionality of adding the fraud score onto existing data similar to that of the provided client code as well as to provide early test cases