**Reference : Online Advertising**

In this document, you will describe the code and the overall steps taken to solve the project.

**Step 1**: create a EMR Cluster.

Step2 : In root(sudo -i)

install Mysql-connector (pip install mysql-connector)

install pykafka (pip install pykafka)

install FLASK (pip install -U FLASK)

install requests (pip install requests)

step3: setting up mysql as per the pdf from sqoop.( Setting up MySQL on EMR Cluster)

Step4: Create a database in maria db with the name upgrad

Step5: use database upgrad; and then create the tables as per the mysql queries pdf

Step6: download the users data from the link: **wget -P /root/** [**https://de-capstone-project1.s3.amazonaws.com/users\_500k.csv**](https://de-capstone-project1.s3.amazonaws.com/users_500k.csv)**.**

Step7: load the users data into the users table using the sql query in mariadb.

**LOAD data local infile '~/users\_500k.csv' \**

**INTO TABLE users \**

**fields terminated by '|' \**

**ignore 1 rows;**

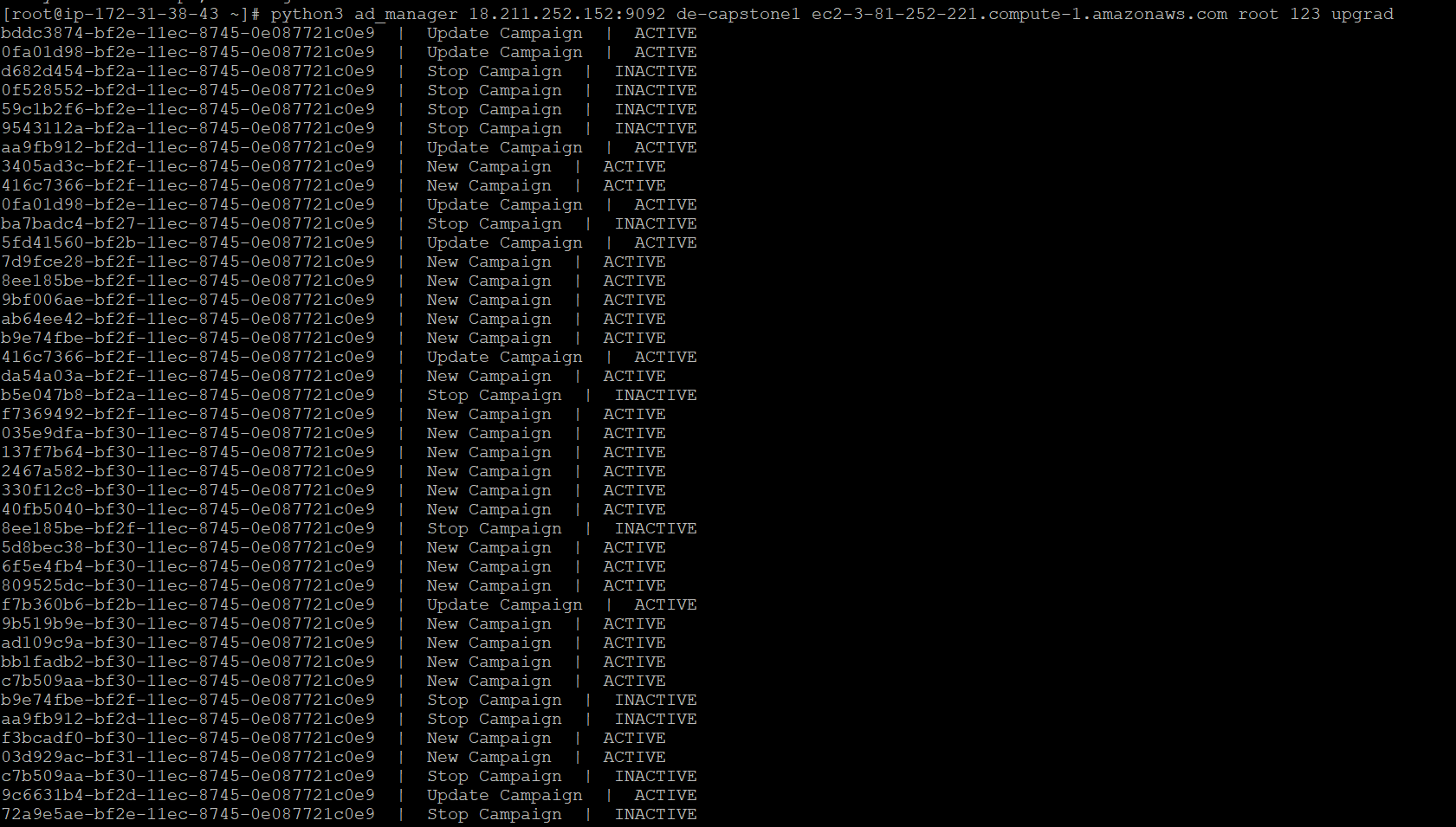
Step8: create a ad\_manager file in root (vi ad\_manager)

Step9: copy paste the code from python file into the file we created in root by pressing I for insert operation and then exit using :wq!

Step10: similarly do the following for the ad\_server and ad\_simulator.

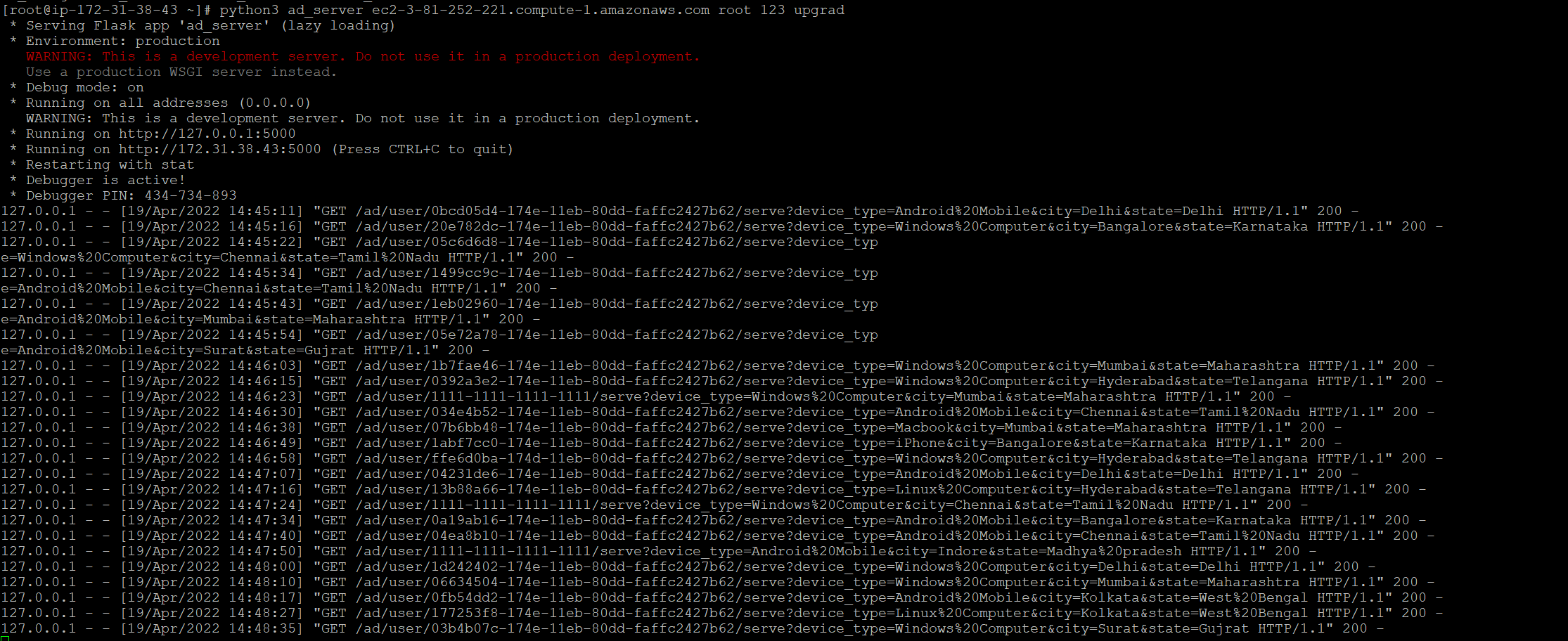
Step11: now run the ad\_manager file using : **python3 ad\_manager 18.211.252.152:9092 de-capstone1 ec2-3-81-252-221.compute-1.amazonaws.com root 123 upgrad**

**Output:**



Step 12: duplicate the session and then run the ad\_server file using the command:

**python3 ad\_server ec2-3-81-252-221.compute-1.amazonaws.com root 123 upgrad**

**Output:** 

Step13: again duplicate the session and run the user\_simulator file using the command:

**python3 user\_simulator ec2-3-81-252-221.compute-1.amazonaws.com root 123 upgrad http 0.0.0.0 5000 0.0.0.0 8000**

**we should be able to see some output in the putty.**