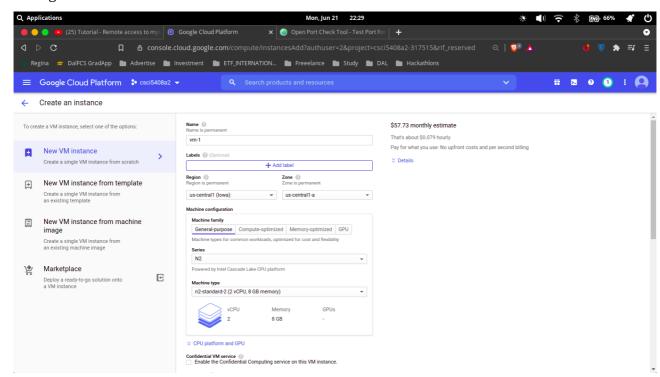
CSCI 5408 DATA MANAGEMENT AND WAREHOUSING ASSIGNMENT – 2

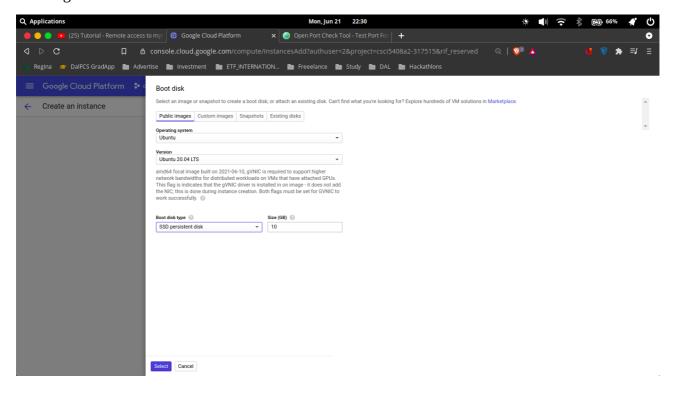
SUBMITTED TO: PROF. SAURABH DEY

SUBMITTED BY: DHRUV DOSHI B00883311

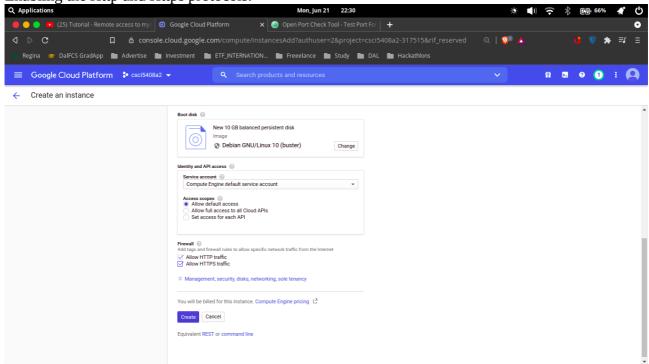
Starting the VM Instance:



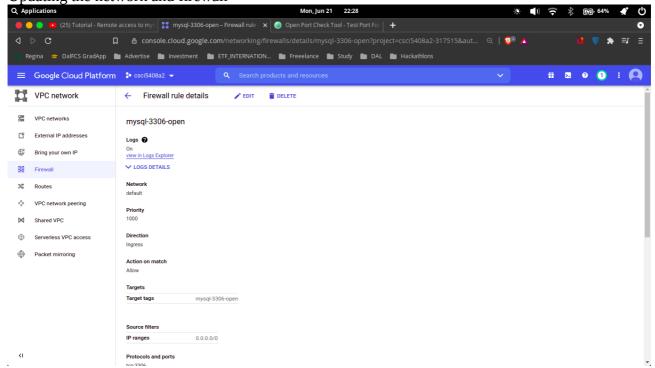
Choosing the Disk:



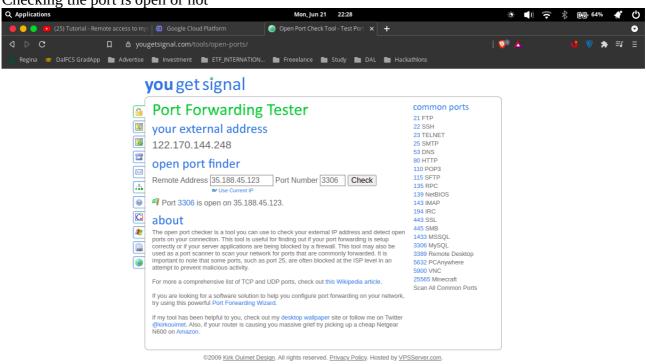
Enabling the Http and Https protocols:



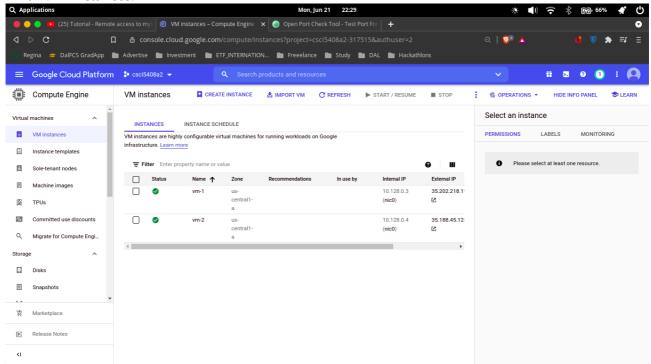
Updating the network and firewall



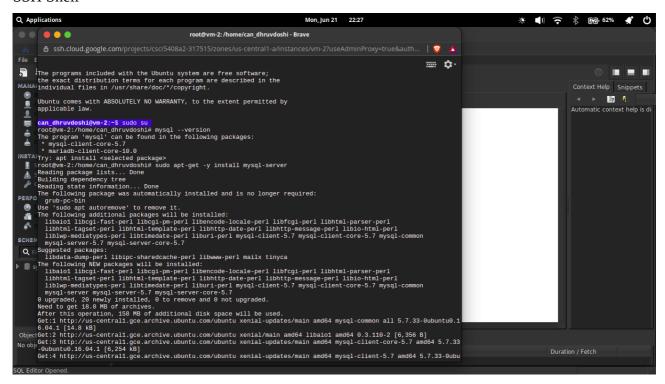
Checking the port is open or not

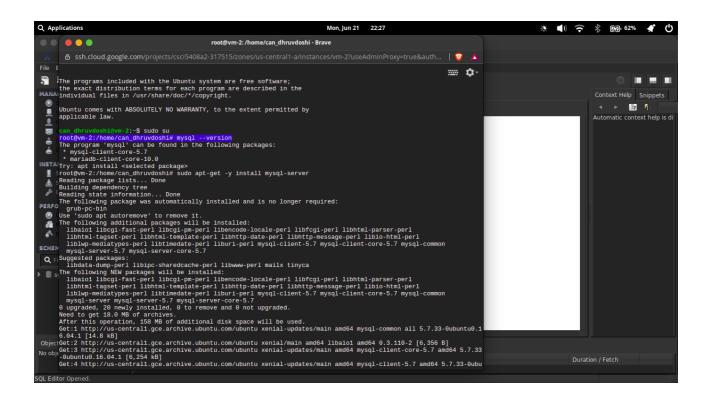


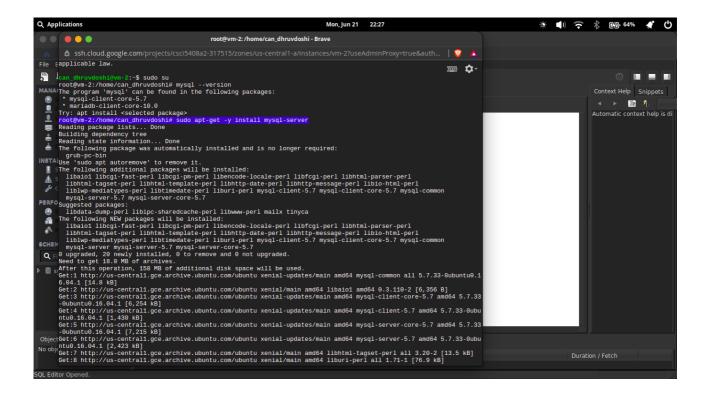
Final VM instances:

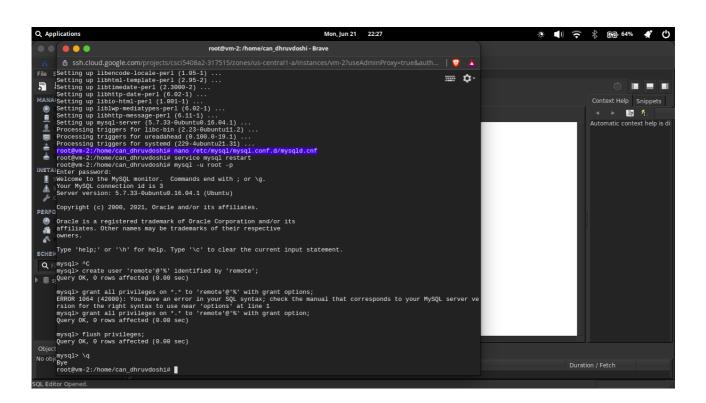


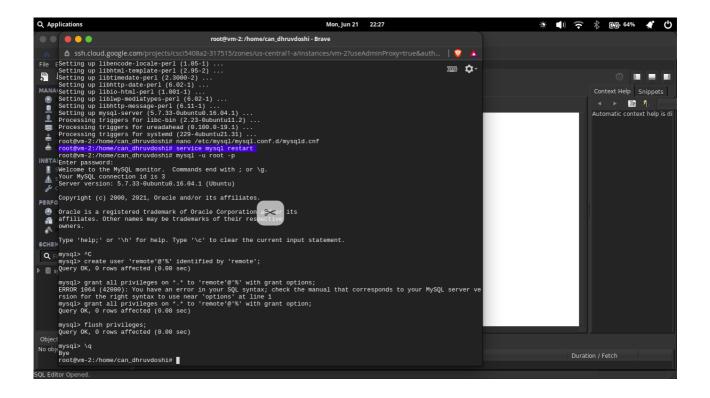
SSH Shell

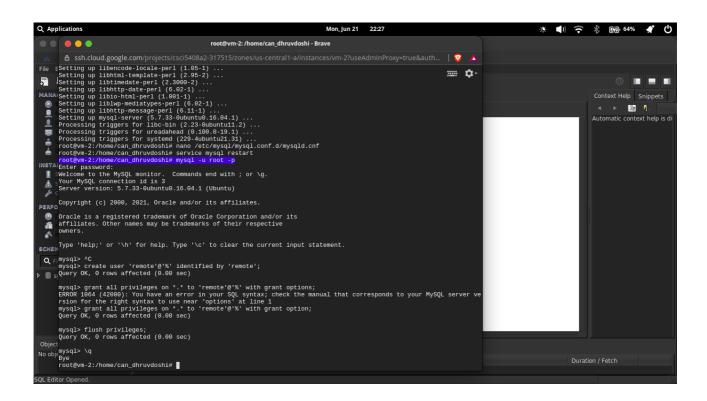


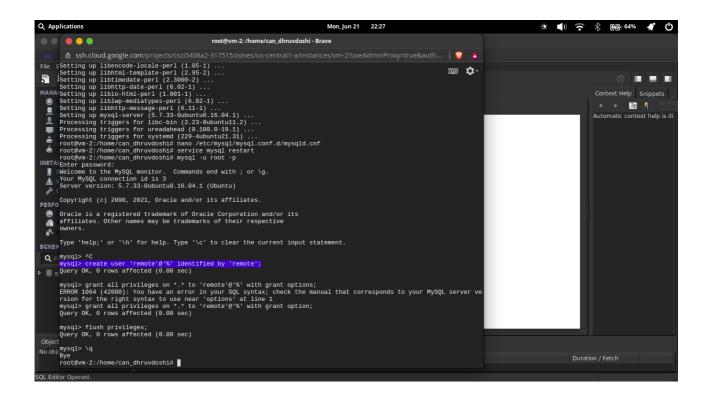


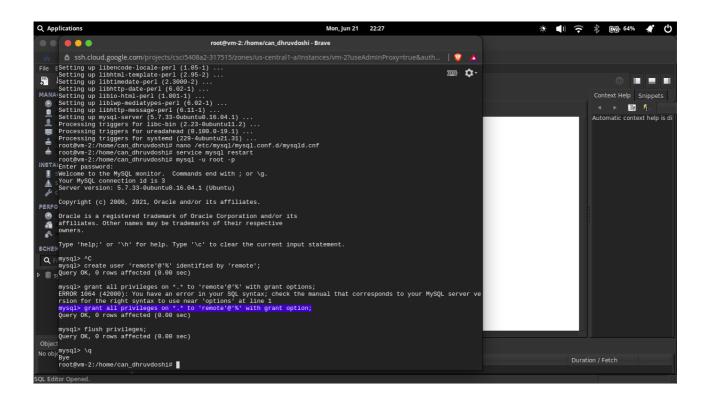


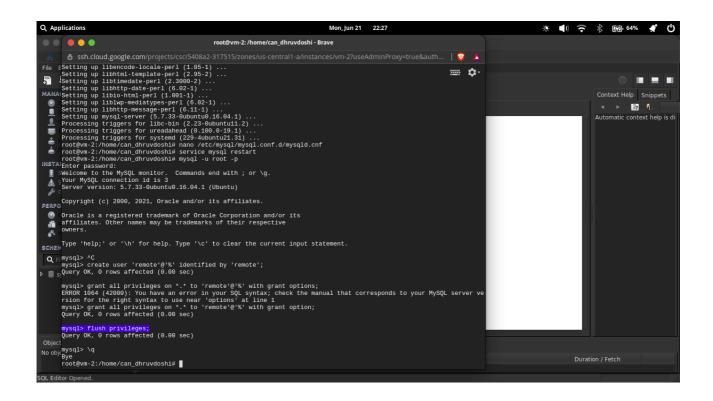


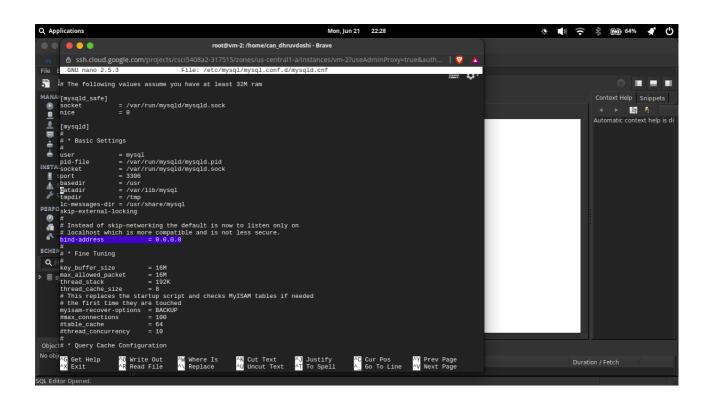




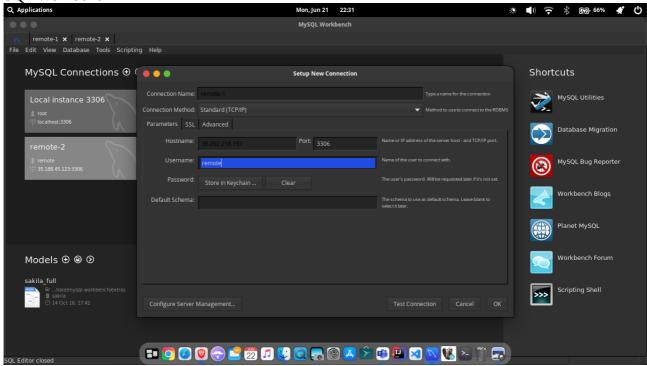


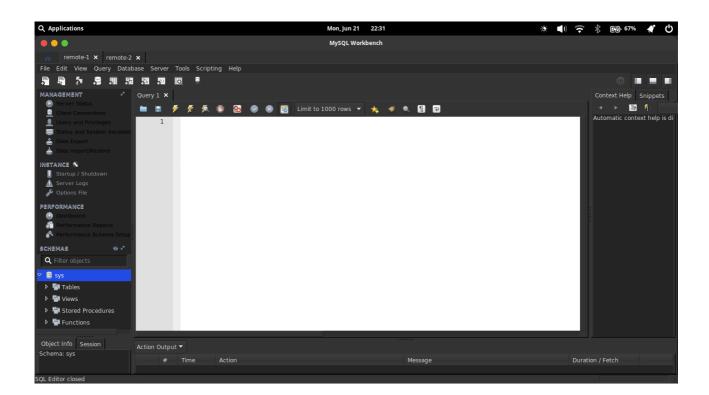


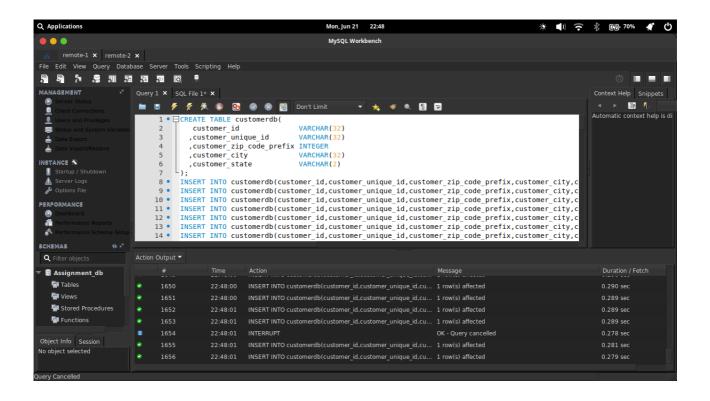


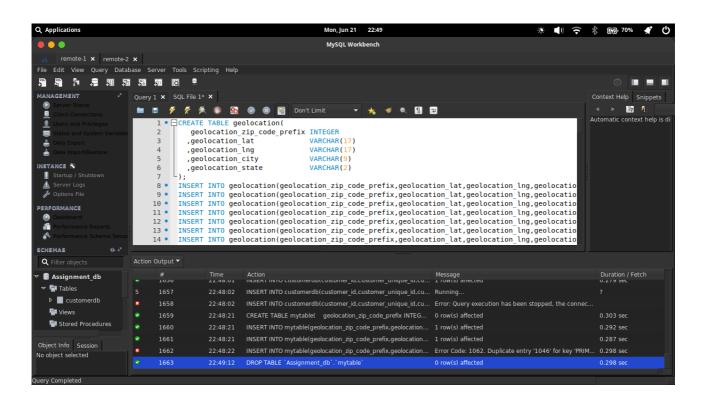


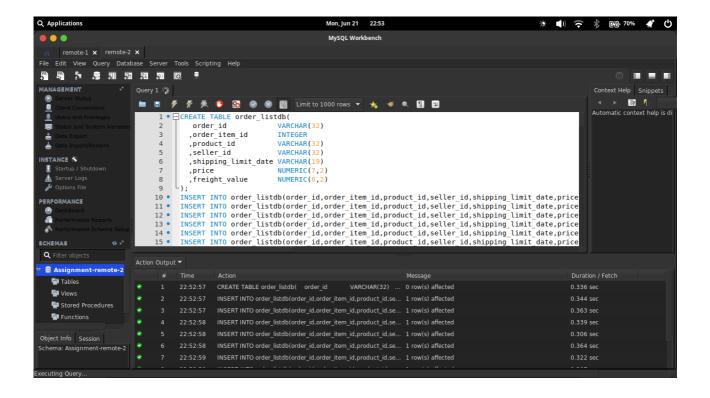
SQL Workbench

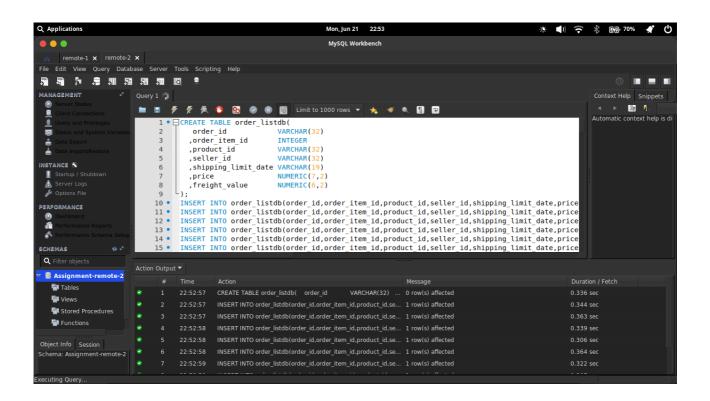


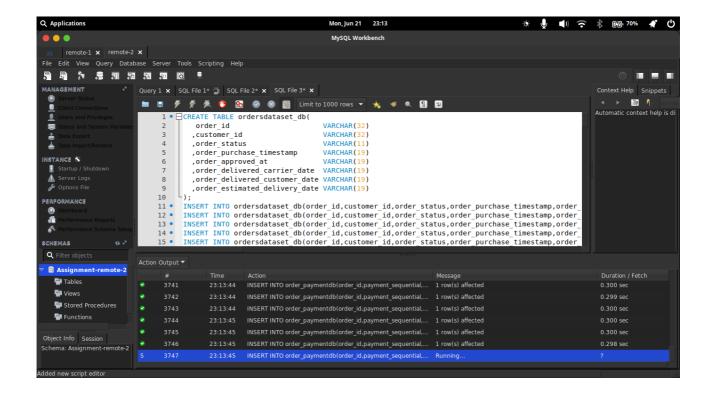












VM1

In this instance there are two tables which does not have any primary key or constraints. However there are no data conflicts which could be found as it is in the highest normal form with no data duplications. Here I have updated the constraints with new primary key in one table and a composite primary key in geological data table.

VM2

In this part also there is also there is no requirenment of cleaning of data as they are almost in the highest normal form possible. Continuing with there we could add some relationships betweens these all tables.

Taking a sceraio where we have a table for orders which would communicate with all other tables and then we could have a whole schema where there are relationships.

Here also we have added multiple primary keys across all of the tables. I used to work with reverse engineering and then make the suitable changes further on I used forward engineering to move those changes to the database table.

Here at the end of the work we have the whole systems which is interconnected.

Dump Files:

There are two initial dump files which had been created with the raw data after moving it into the sql from the csv provided files

There are another two dump files after the update.

REFERENCES:

[1]"Datalab documentation | Cloud Datalab Documentation | Google Cloud", *Google Cloud*, 2021. [Online]. Available: https://cloud.google.com/datalab/docs. [Accessed: 21- Jun- 2021]

[2]K. Ouimet, "Open Port Check Tool - Test Port Forwarding on Your Router", *Yougetsignal.com*, 2021. [Online]. Available: https://www.yougetsignal.com/tools/open-ports/. [Accessed: 21- Jun-2021]

[3]"Before you continue to YouTube", *Youtu.be*, 2021. [Online]. Available: https://youtu.be/8oF4ku_7vxM. [Accessed: 21- Jun- 2021]