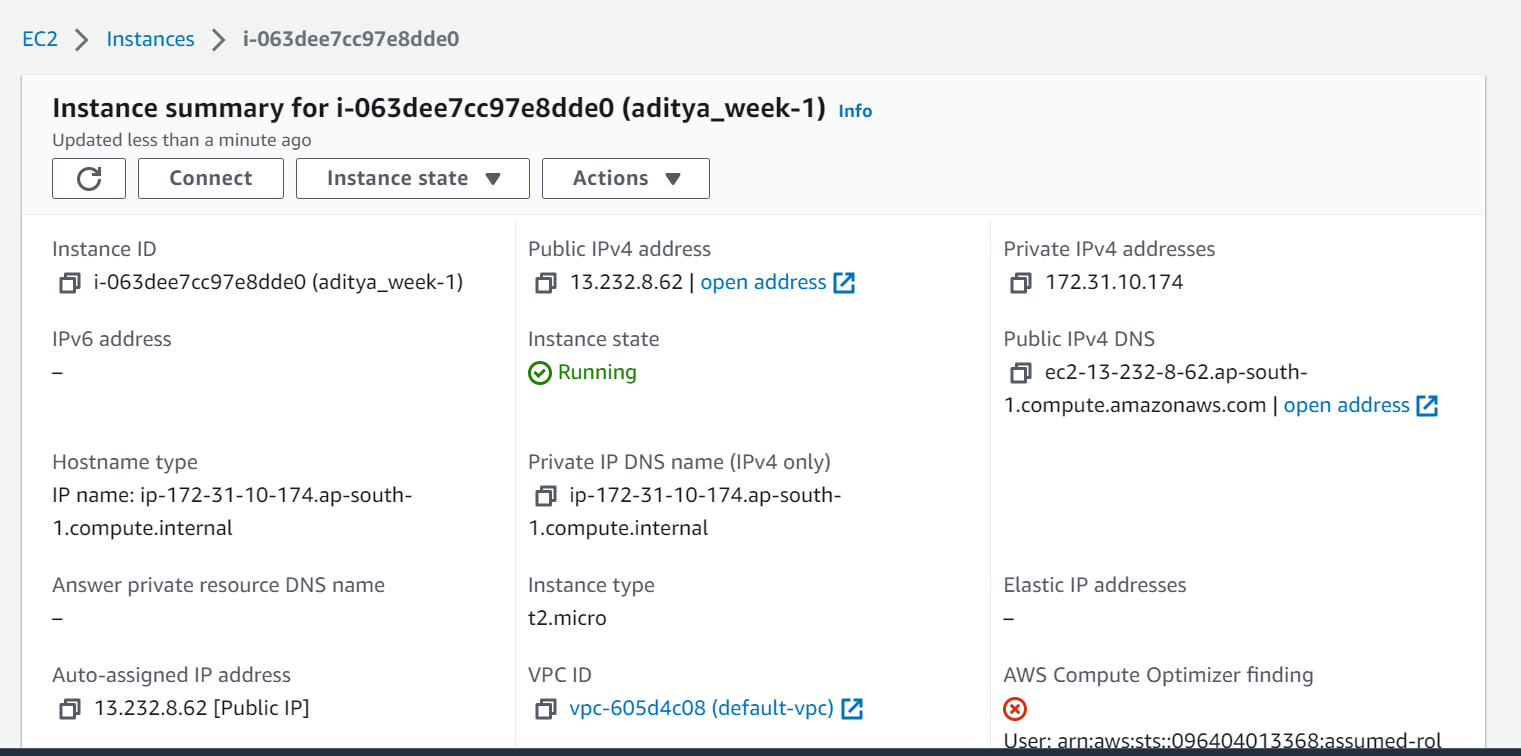
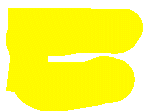
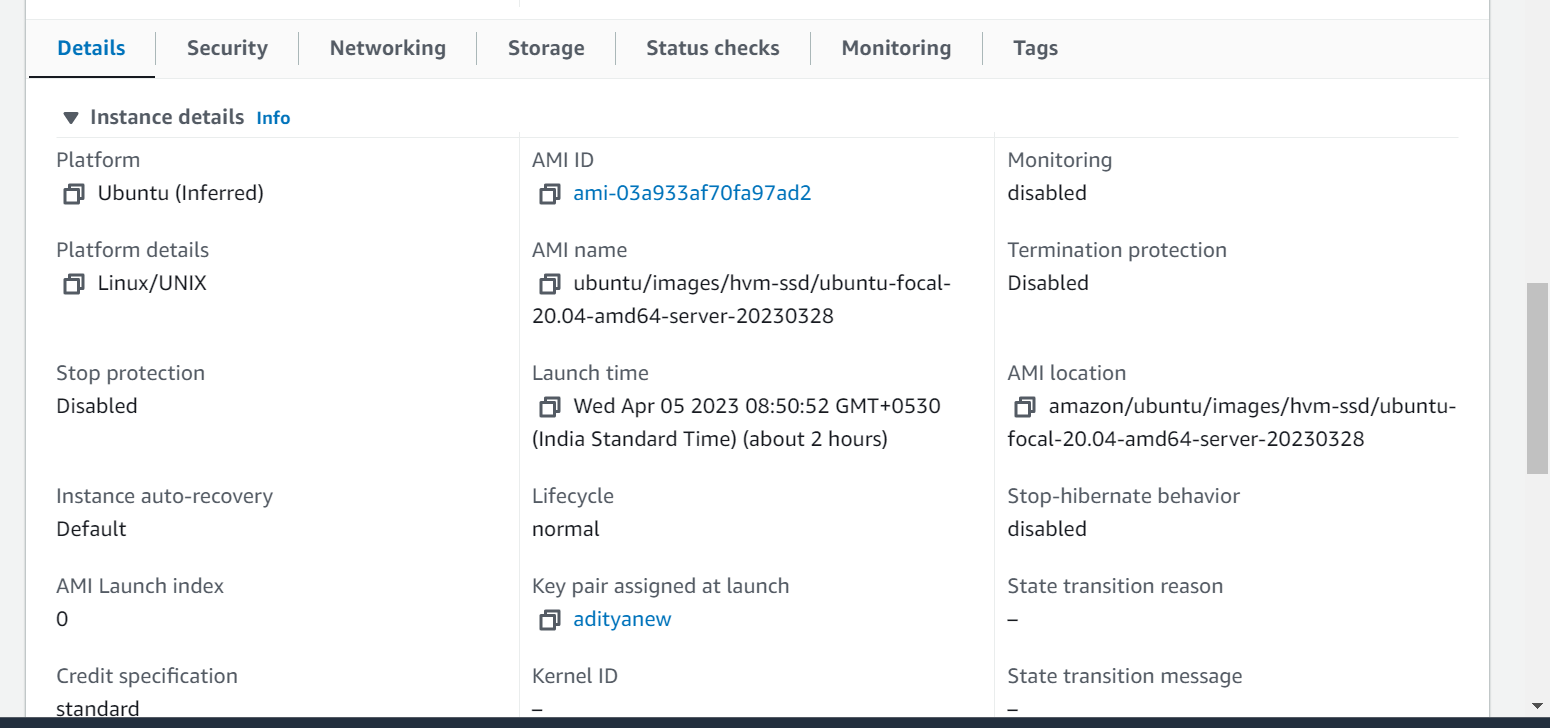
# Aditya Kumar Week -1 Task

# Task 1: Launch an EC2 instance (Amazon Linux/ Ubuntu)

Step 1: Launched EC2 Ubuntu t2.micro instance.



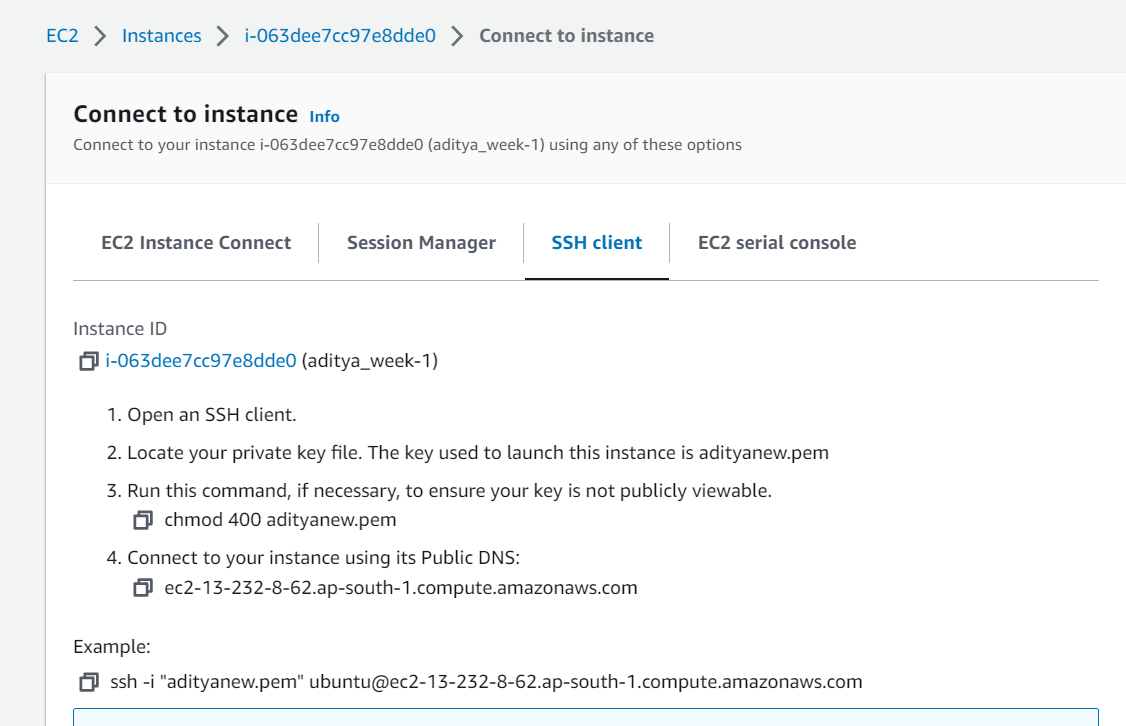


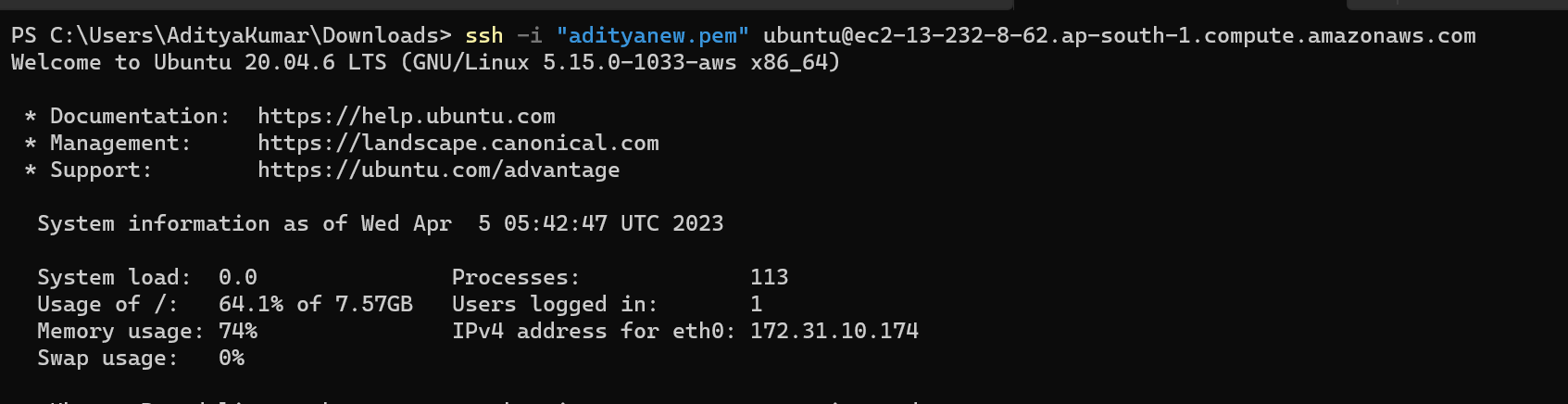


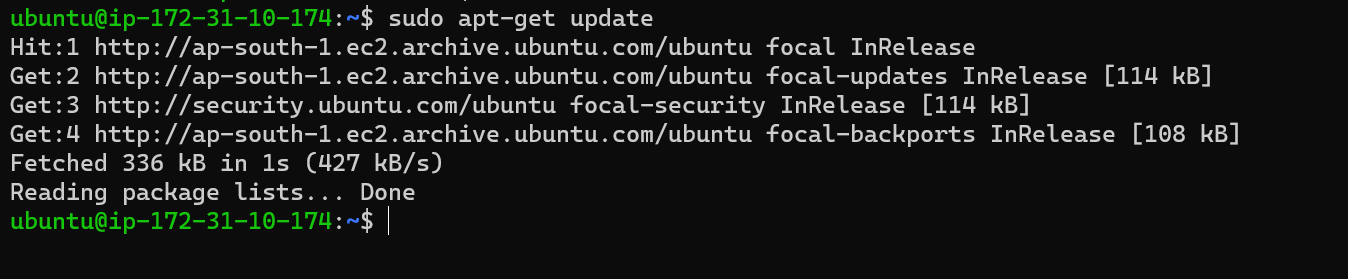


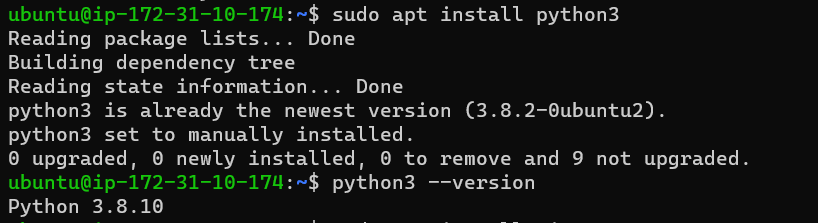
# Task 2: Install Python and MySQL.

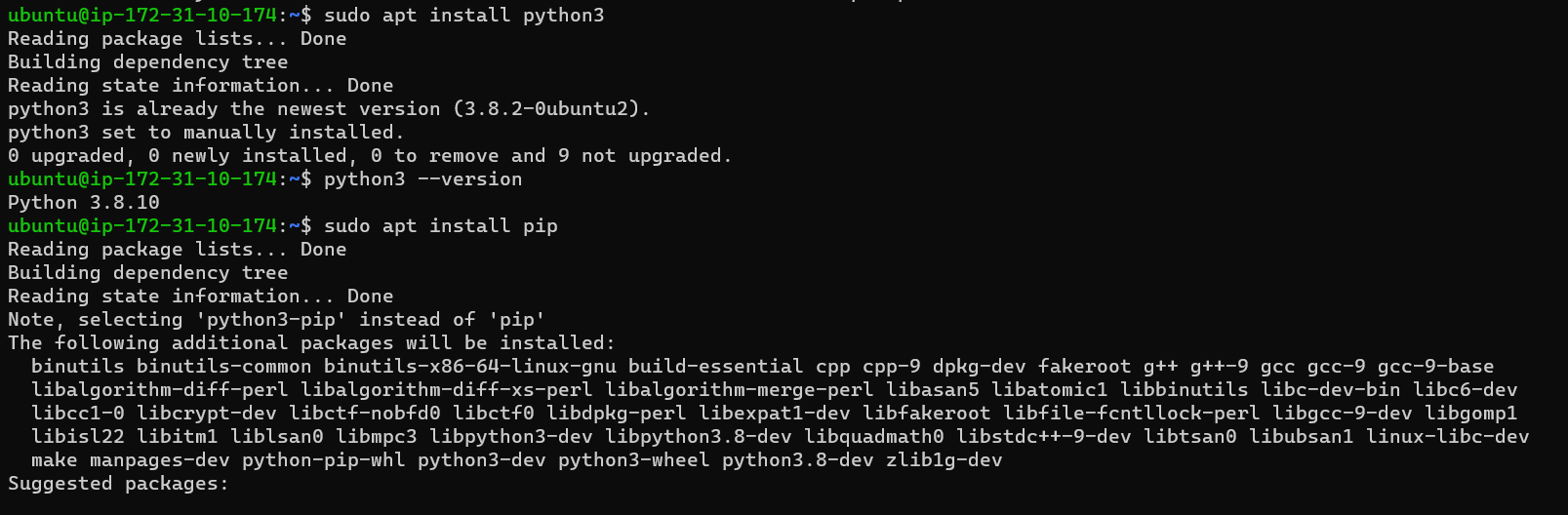
Step 1: login that instance using the ssh connection.

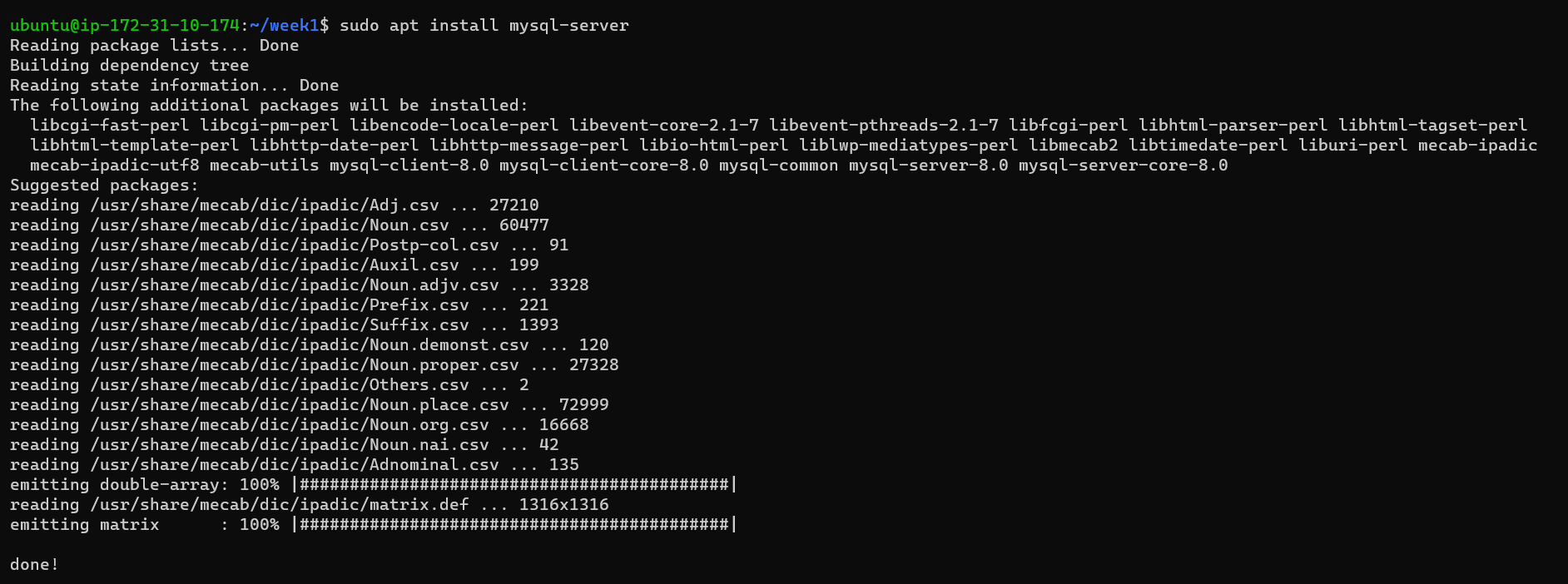
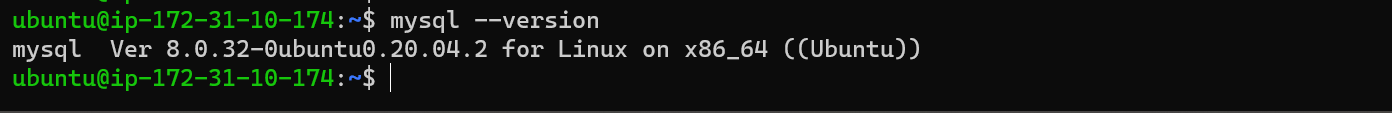
Copy the SSH client URL and paste into the windows cmd and hit the enter button. 

We have successfully login that instance.

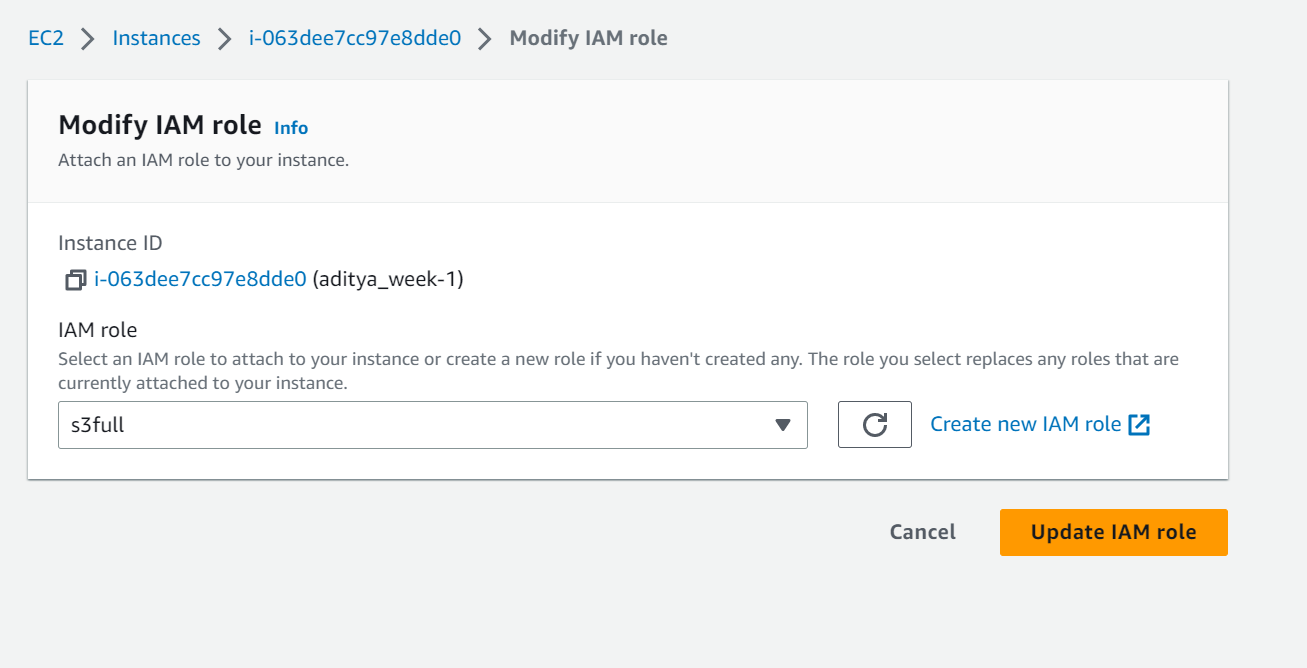
Step2: Update the EC2 instance and install the Python and MySQL. 

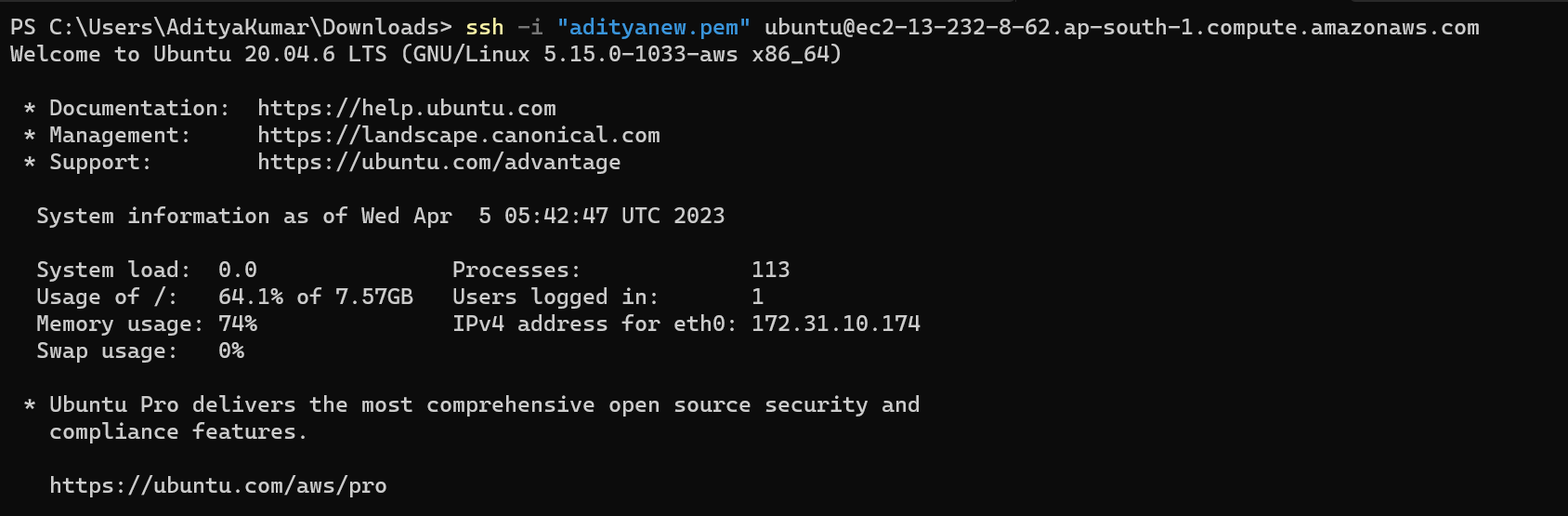
Installing Python and also checking their version.

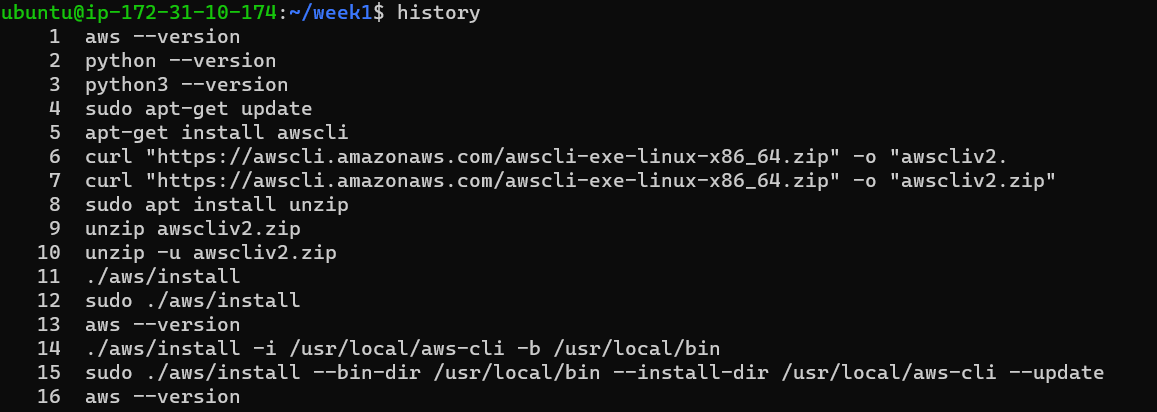
Also installing ‘pip’. 

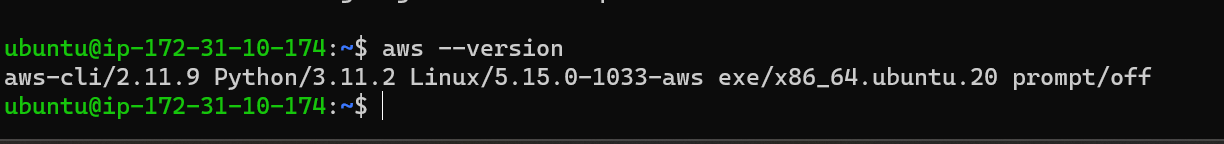
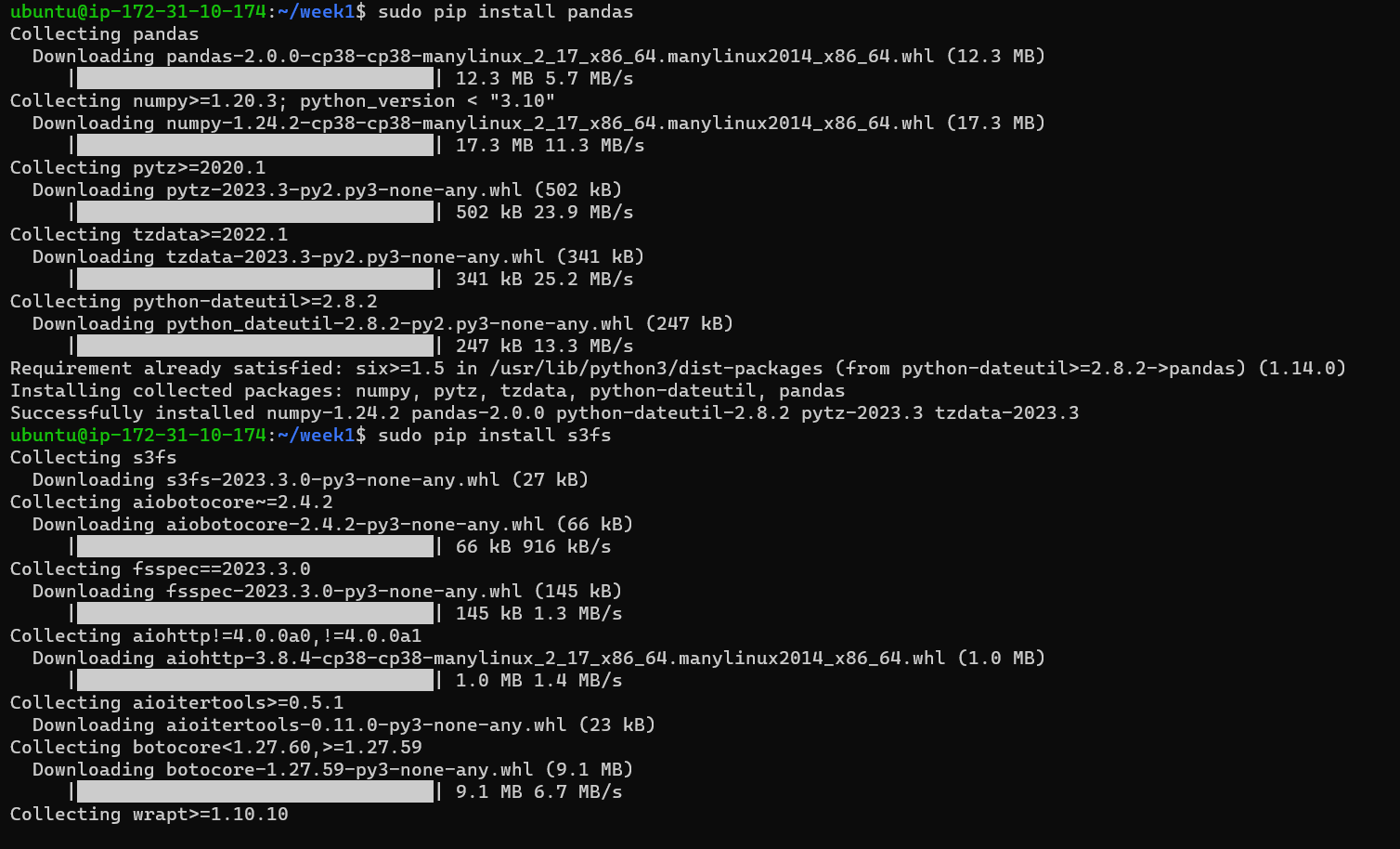
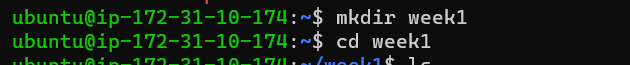
Installing MySQL:checking MySQL version:

# Task 3: Read the Students-Table.csv file from S3 using python.

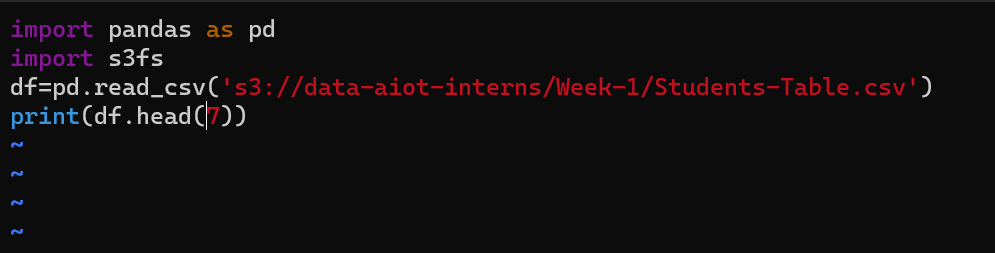
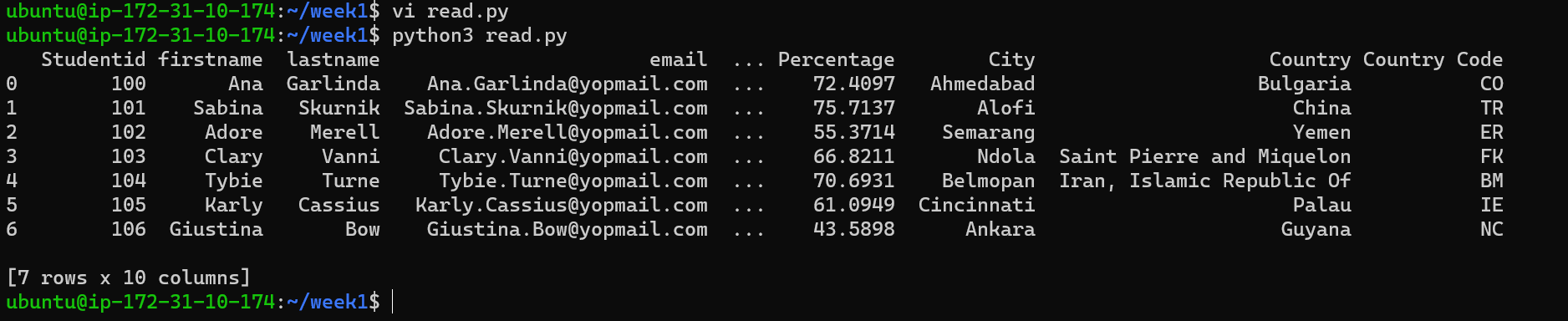
Step 1: Attach IAM role to that instance. That role having S3 full access. So that I will read data from S3 bucket

Step 2: Login that instance using SSH connection. 

Step 3: Install AWS Cli and some python library so that it will read data from s3 bucket.

Checking AWS Cli version. Installing Pandas and S3fs library using pip.Step 4: Created a empty directory to perform all the task in that directory.

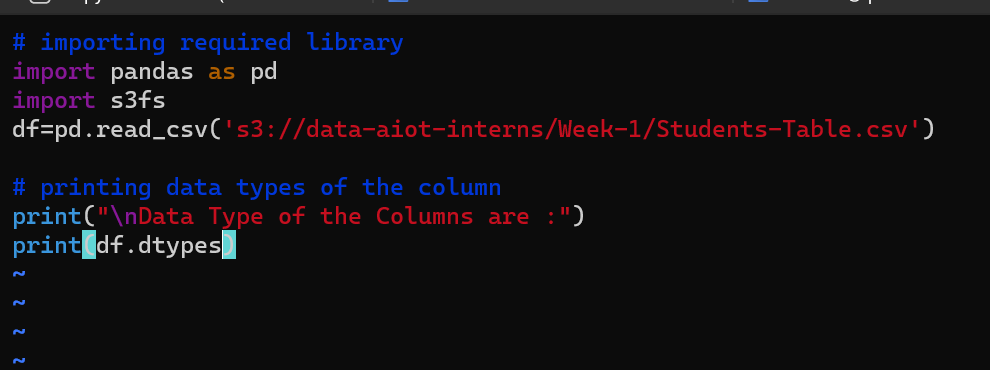
Also created python script file to read data from S3 bucket.

Read.py file: Run the pthon scrpit.

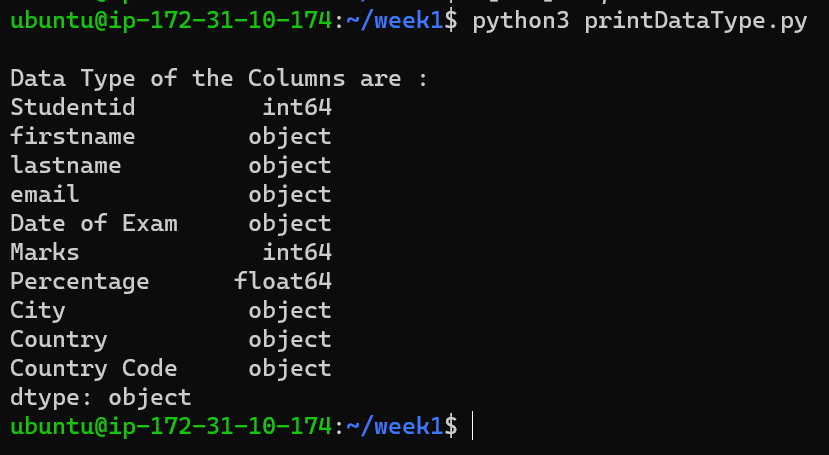
# Task 4: Print data types of all columns.

Step 1: As per till now previous state we are printing the data type of the column. By using the python script.

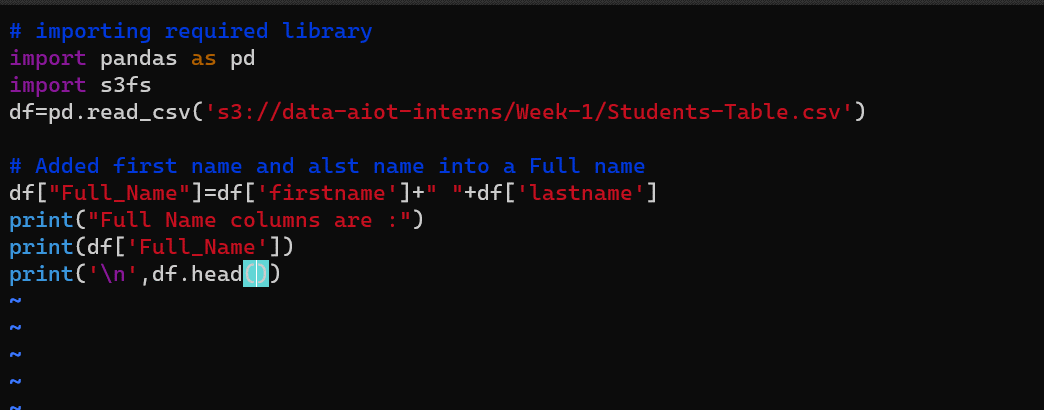
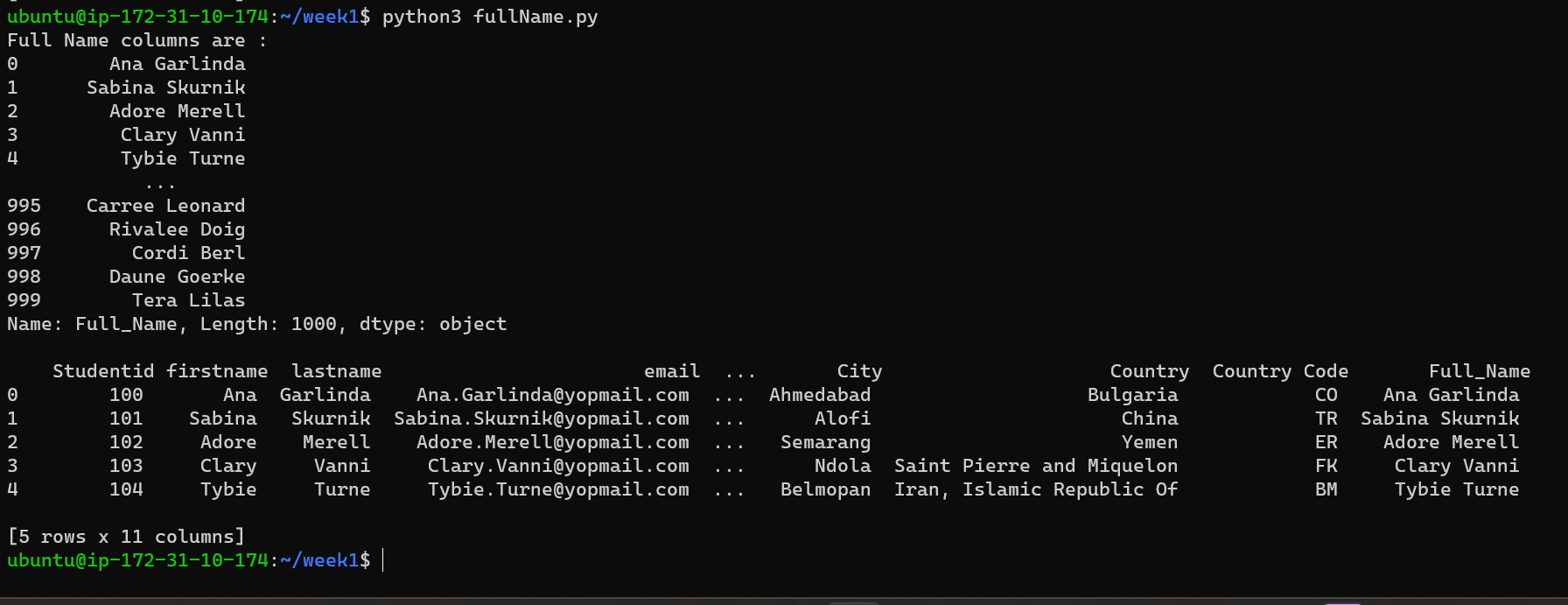
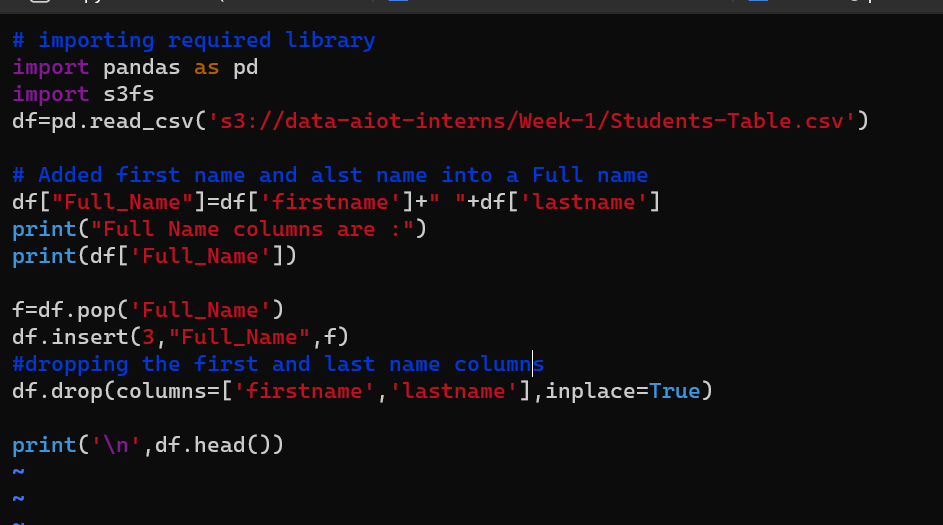
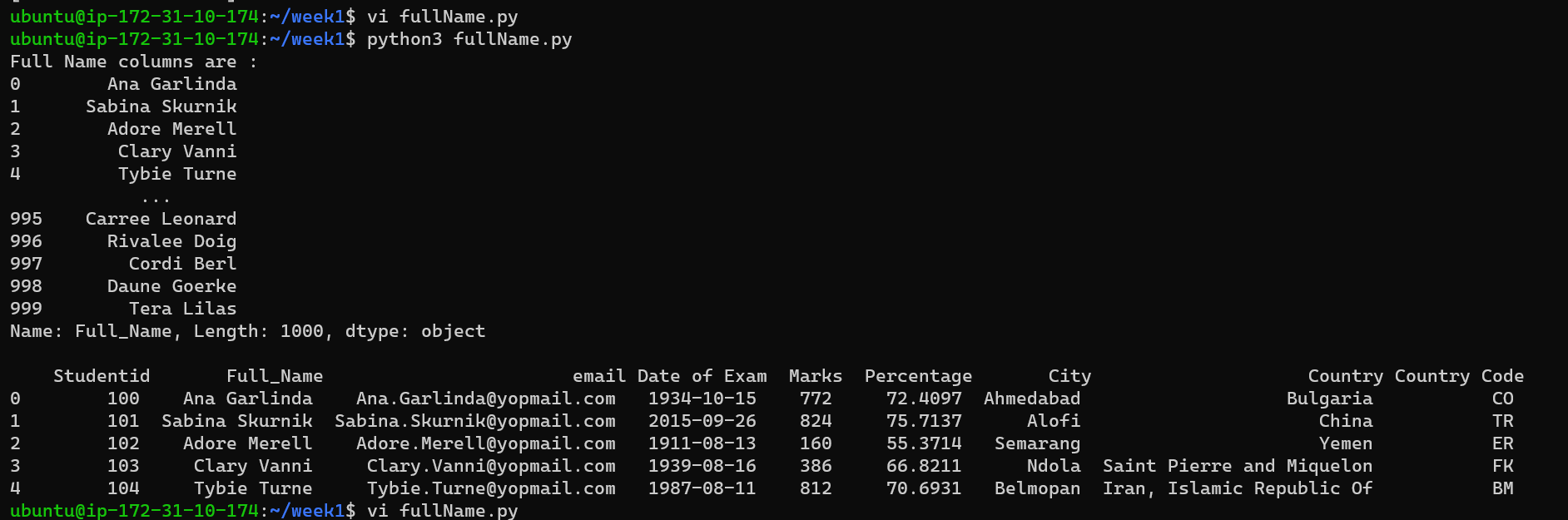
“printDatatype.py” python script file



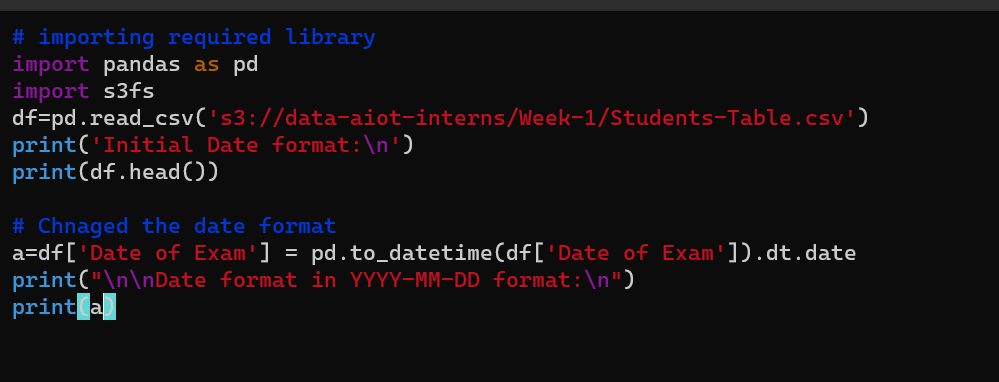
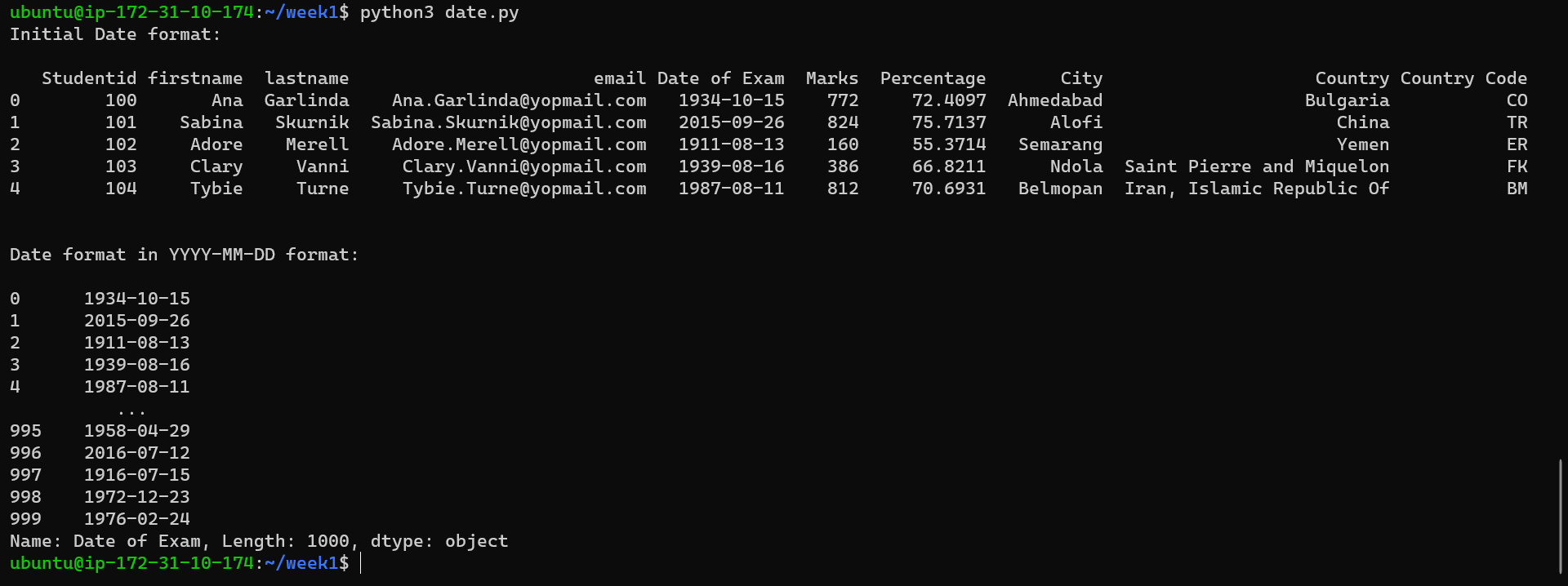
Running the script:



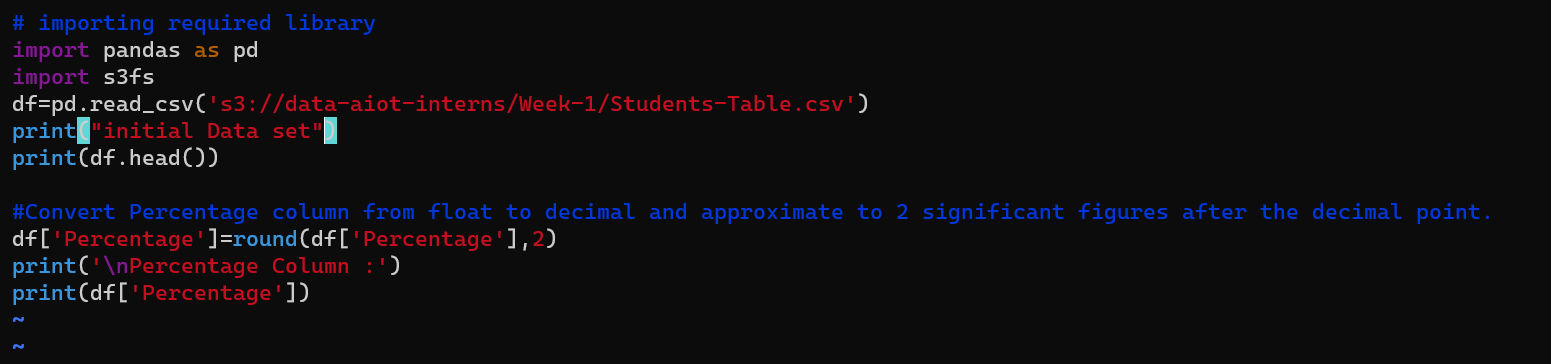
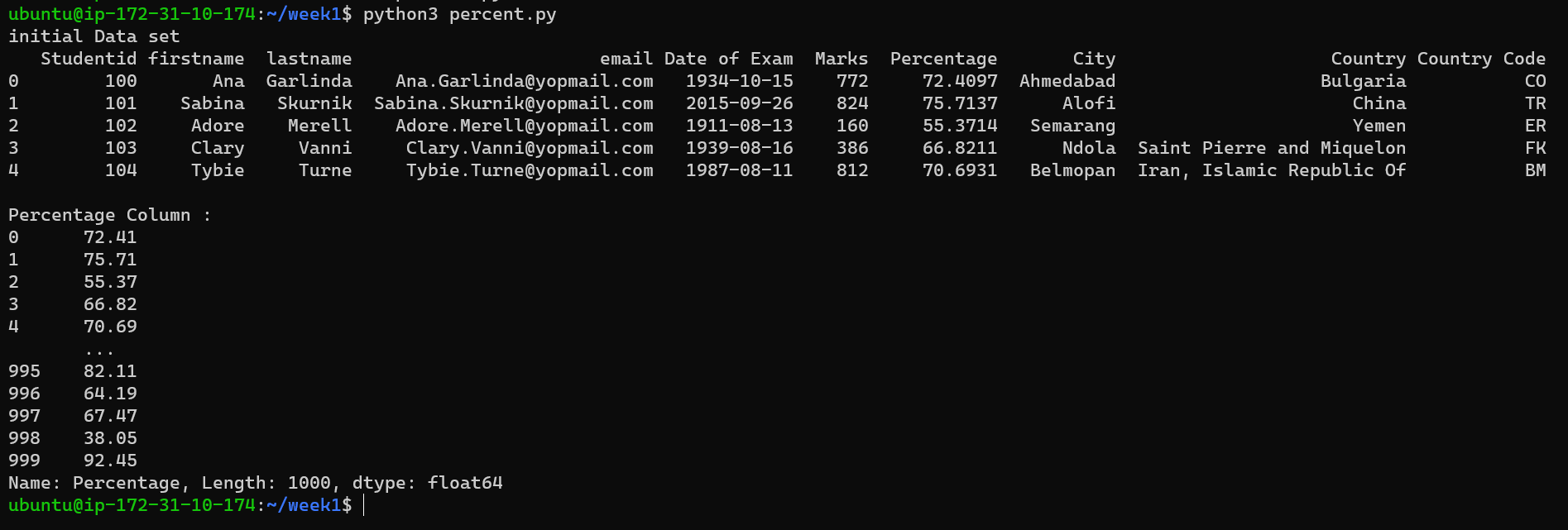
# Task 5: Combine the First Name and Last Name columns to single new column called Full Name.

Python Script file to combine the First Name and Last Name columns into a single new column called Full Name. Rnning python Scrpit file.Also inserting Full Name column at index 3 dropping the first and last name column. 

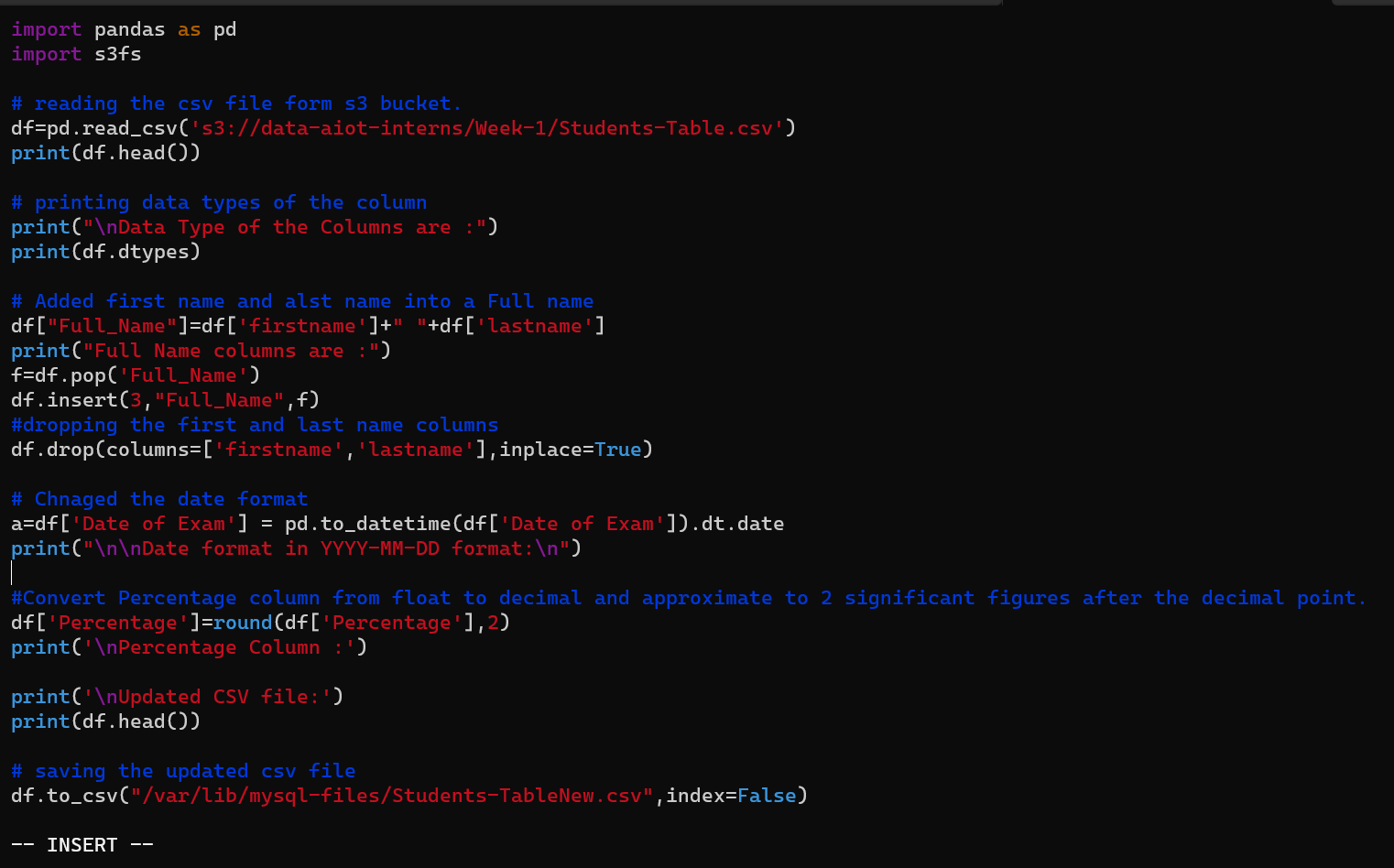
# Task 6: Convert date time format to YYYY-MM-DD.

Step 1: created a python script “date.py’ to convert date into appropriate format.Step 2: Running the script.

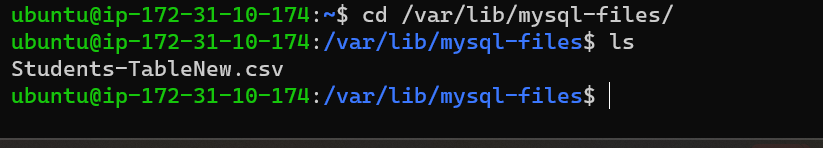
# Task 7: Convert Percentage column from float to decimal and approximate to 2 significant figures after the decimal point.

Step 1: Created a python Script file “percent.py”.Step 2: Running the Script.

Also saving the update file. For that I have combine all the of above script into a single script and saving the updated csv file.



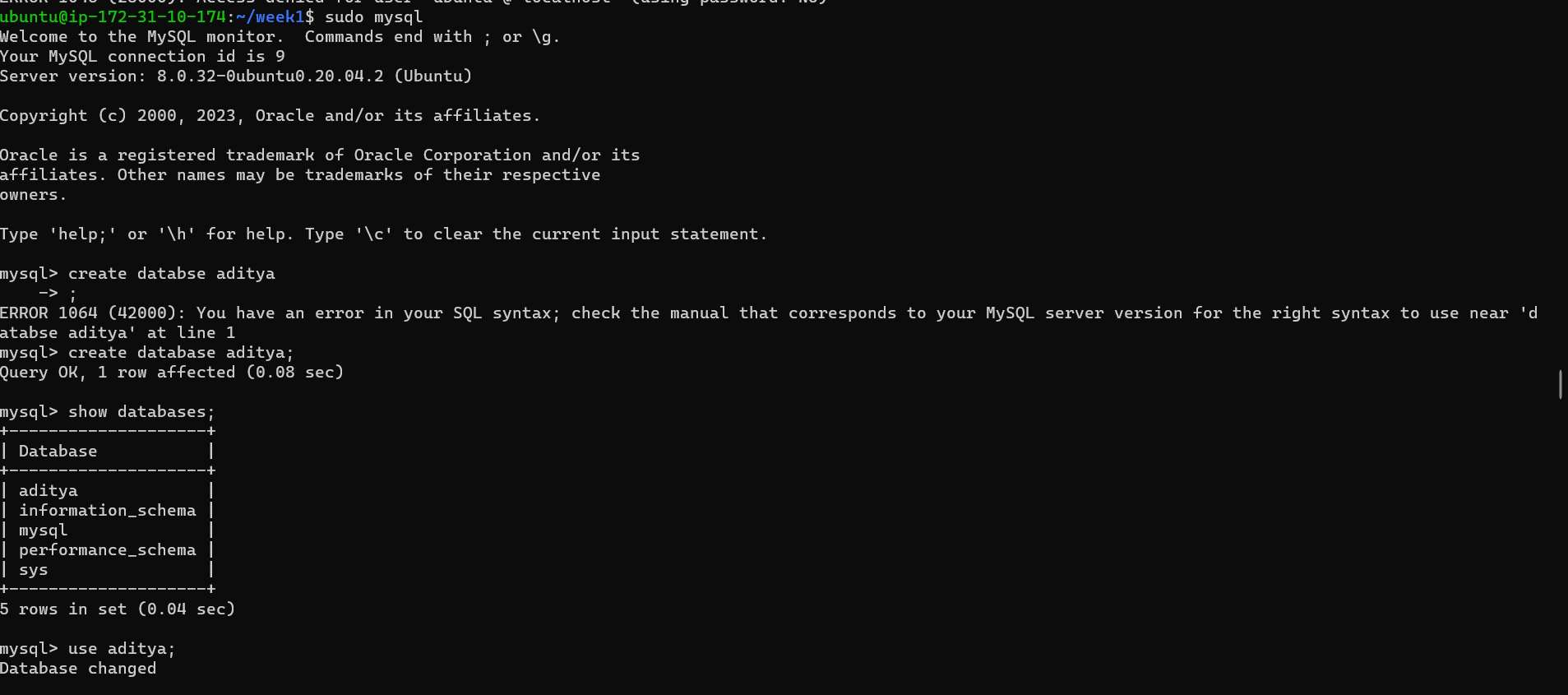
We have run this script file has been saved at specified location.



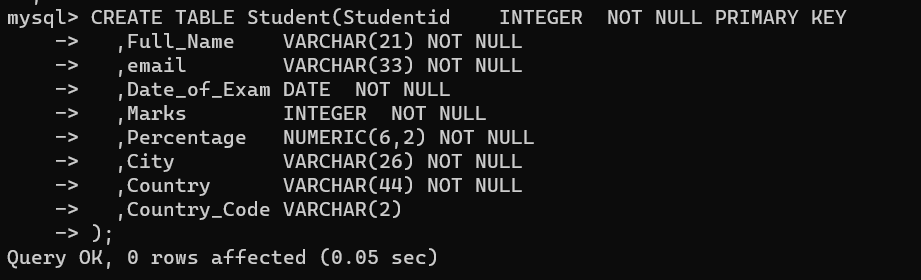


# Task 8: Create a Database in MySQL and create a table with the file modified till previous step.

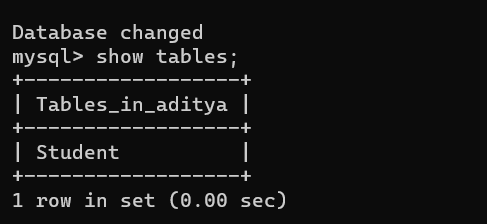
Step 1: Login to MySQL which is installed earlier and created a database and using that database to create a table. Commands are highlited in the yellow box.

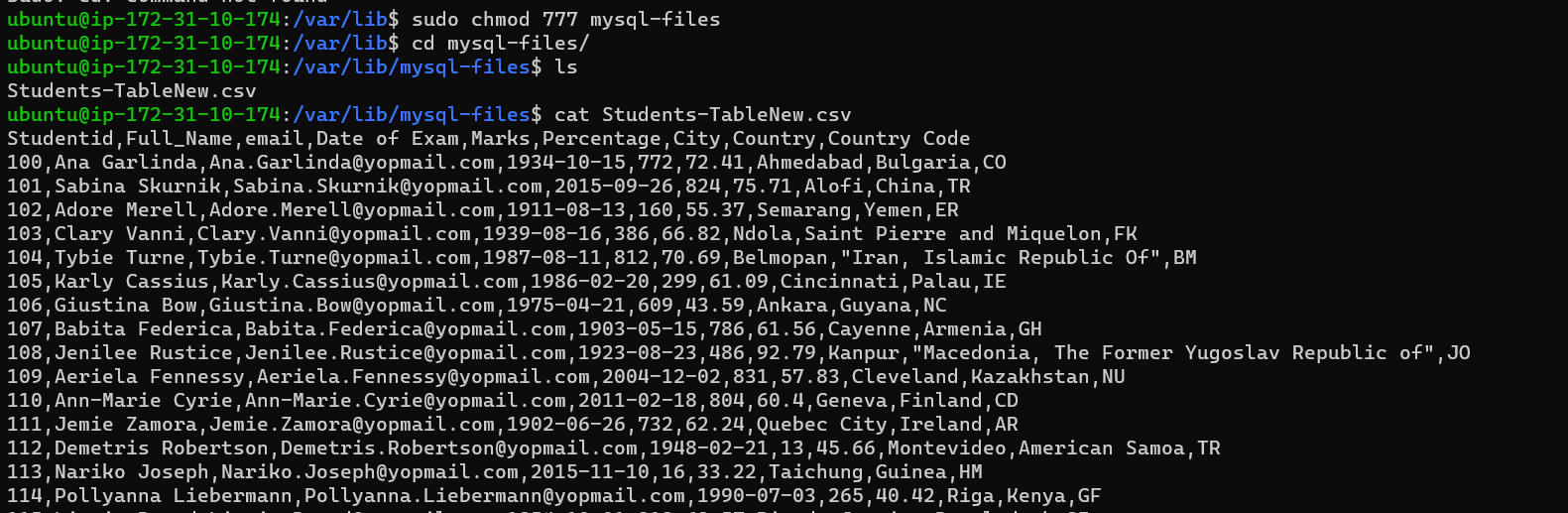


Creating a table.



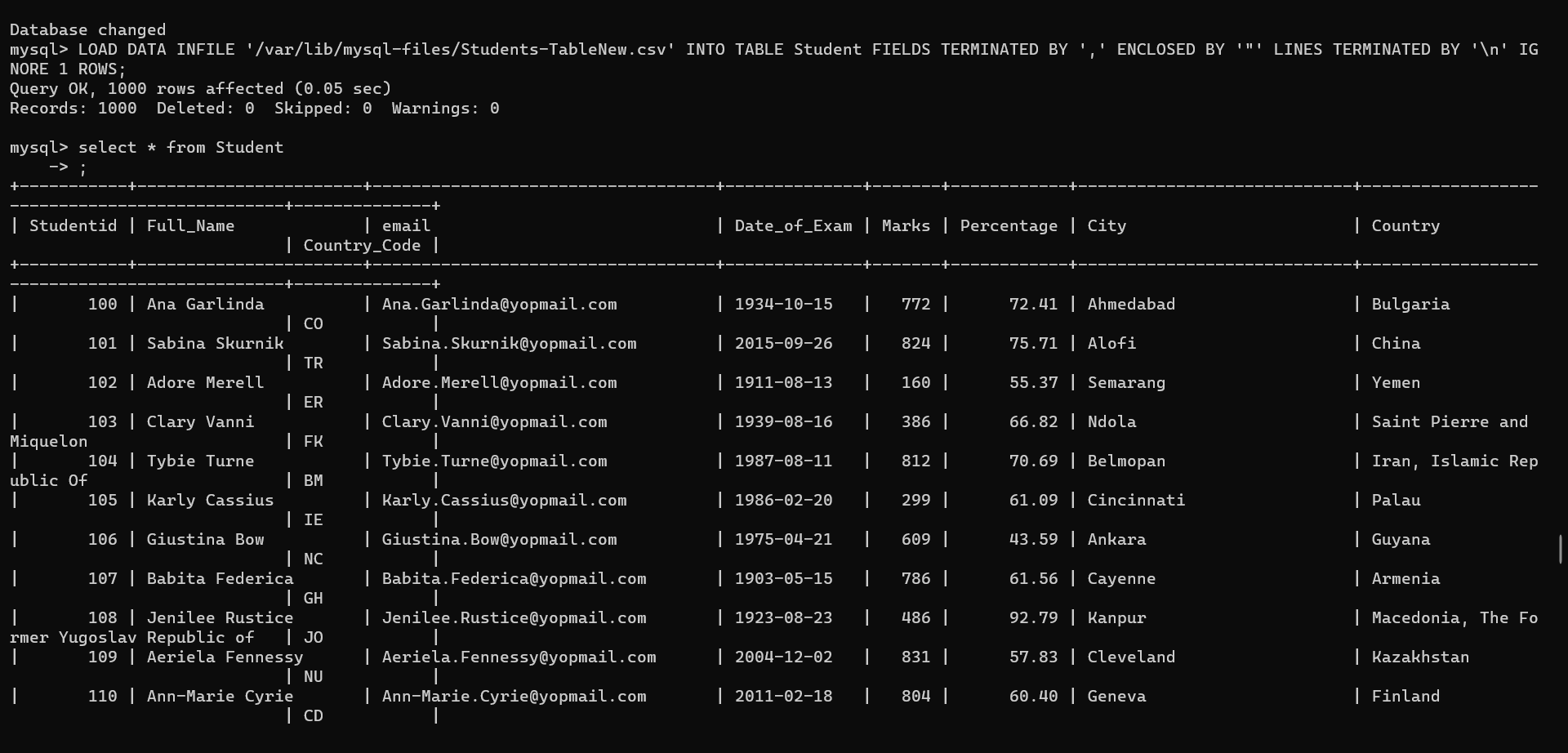
Step 2: Viewing the table.



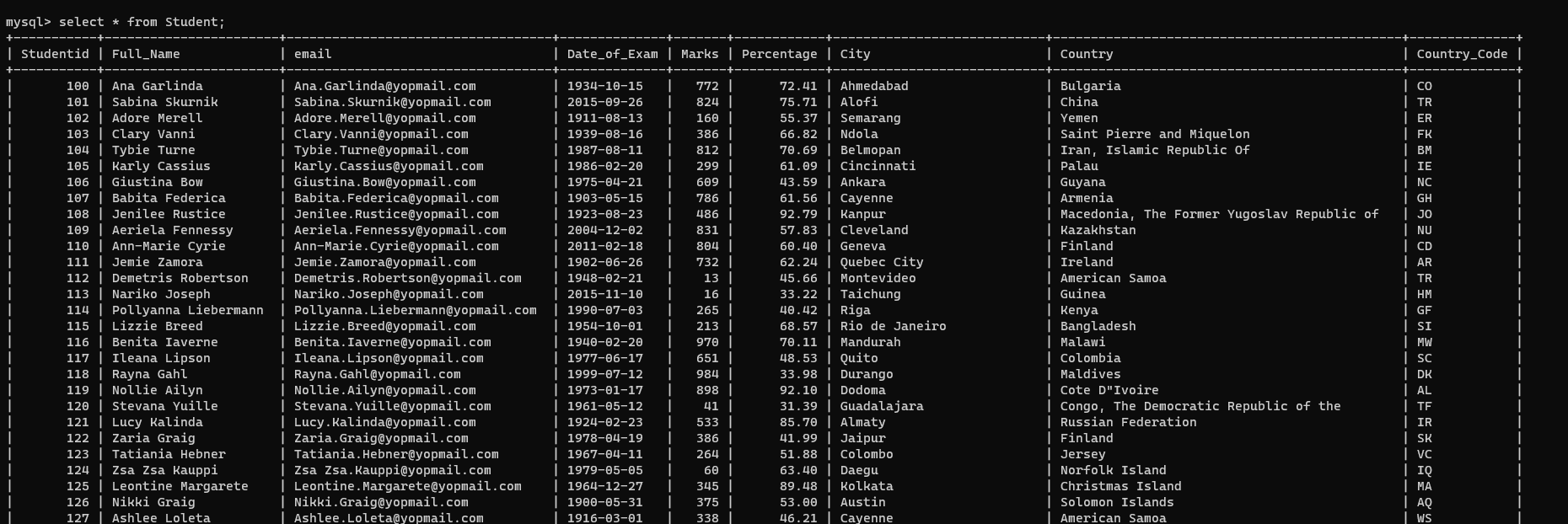
Now Changing the file permission of that folder where modified CSV file is stored in previous tasks. 

Step 3: Inserting data into table from modified csv data. And viewign the data form table. Commands are highlited in the yellow box.

LOAD DATA INFILE '/var/lib/mysql-files/Students-TableNew.csv' INTO TABLE Student FIELDS TERMINATED BY ',' ENCLOSED BY '"' LINES TERMINATED BY '\n' IGNORE 1 ROWS;



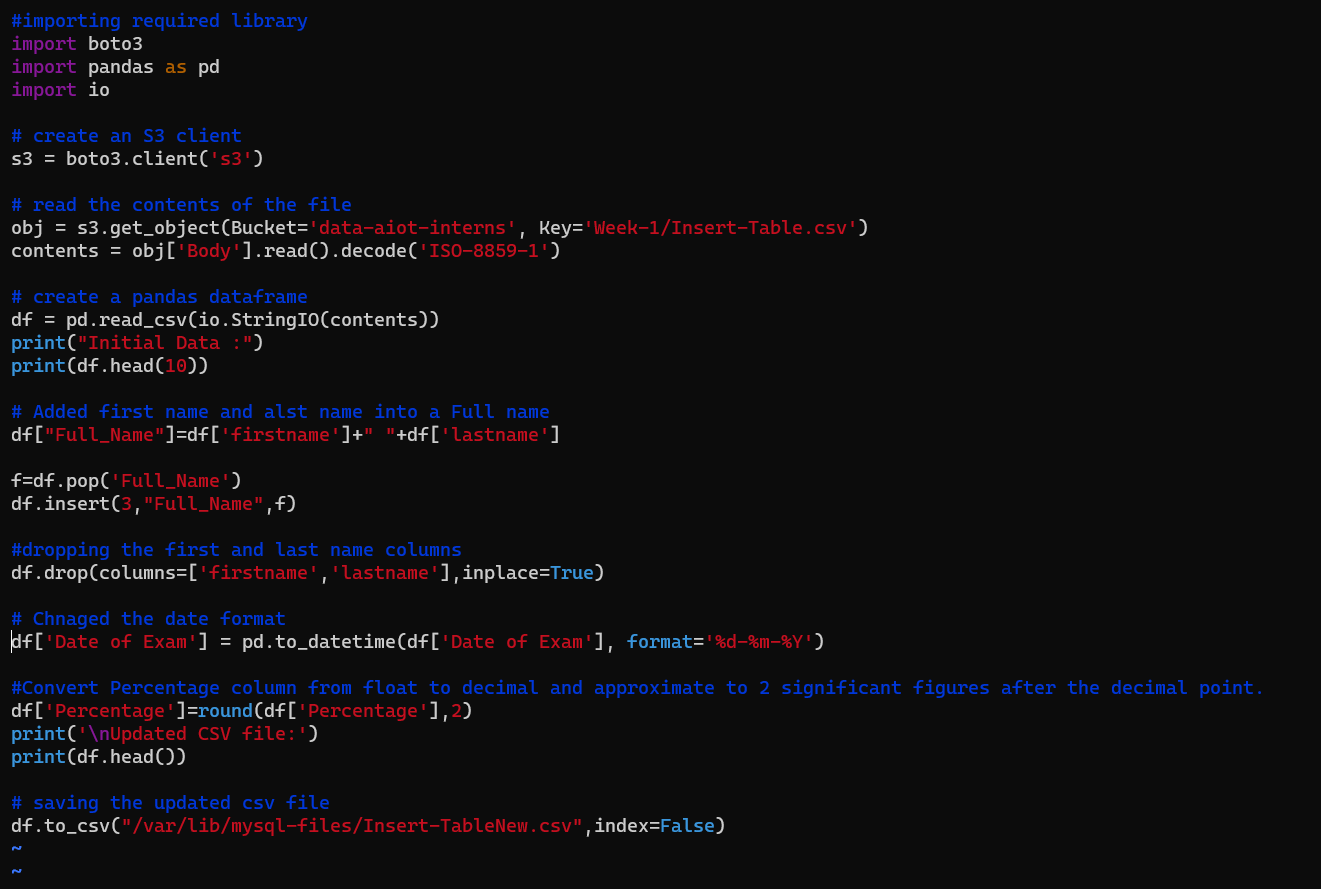
Select \* from Student;

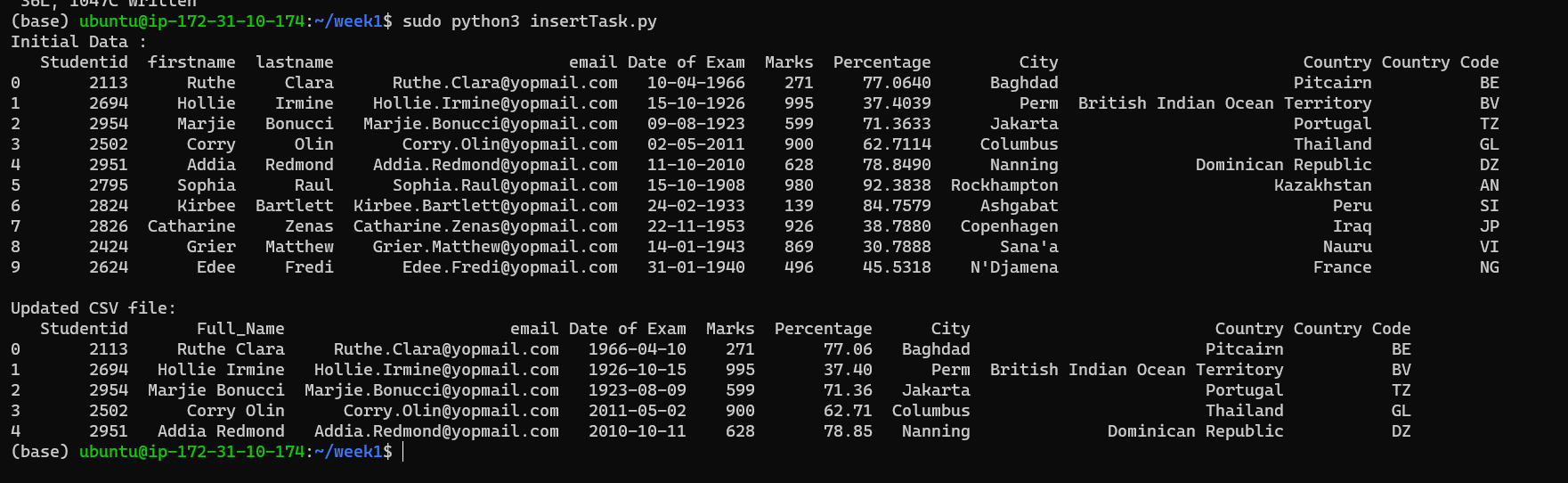


# Task 9: Insert new columns using the Insert-Table.csv in s3.

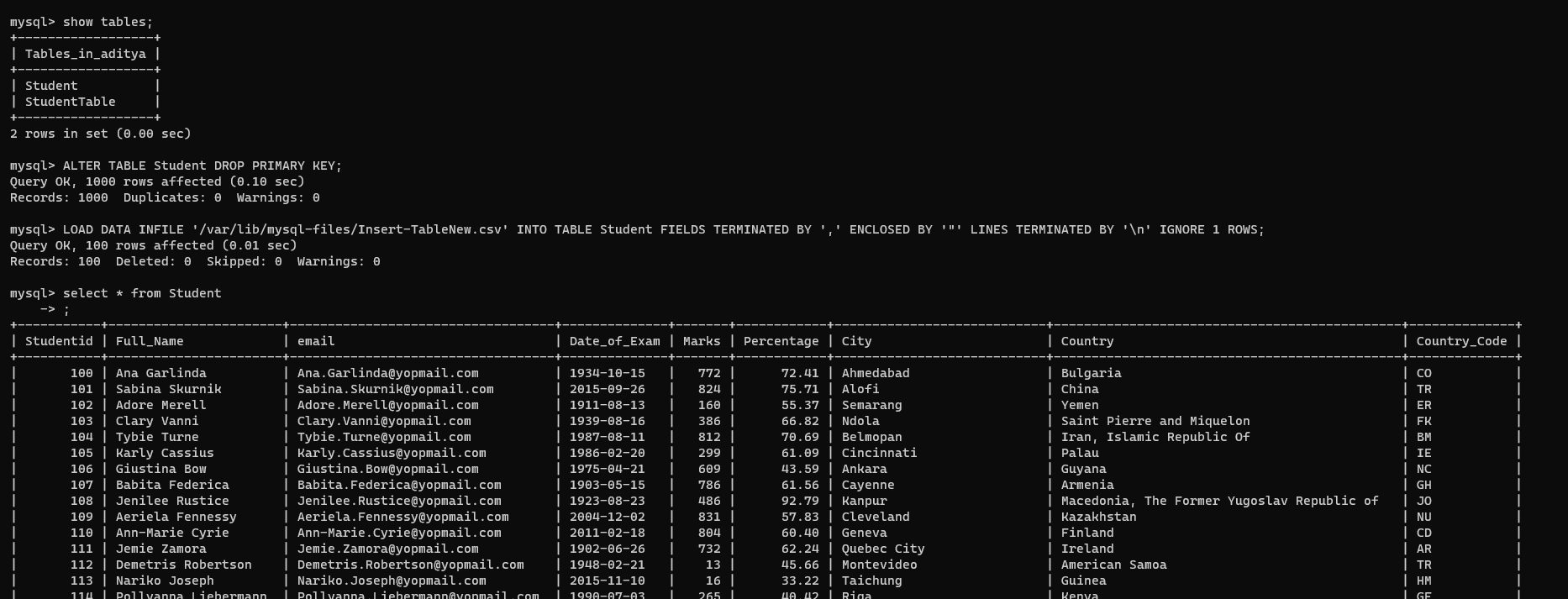
Step 1: Created a python script to read data from s3 and save it. Also make some changes like convert percentage column to upto 2 decimal point and combined first name and last name as a full name column.

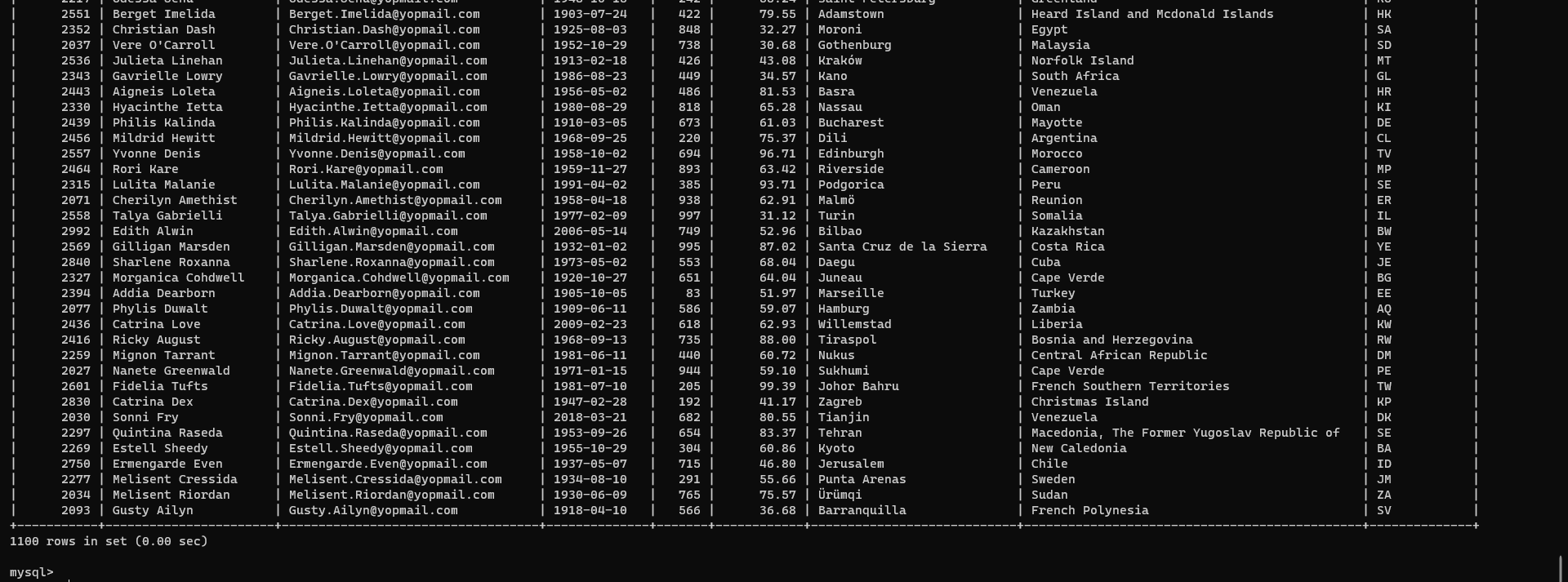
Saved file will be used to insert new rows in the table.



Step 2: Running the script to makes some changes according to requirement and save that csv file. 

Step 3: Inserting data into student table which is created earlier and student data is stored eralier. And Viewing data suing select command.

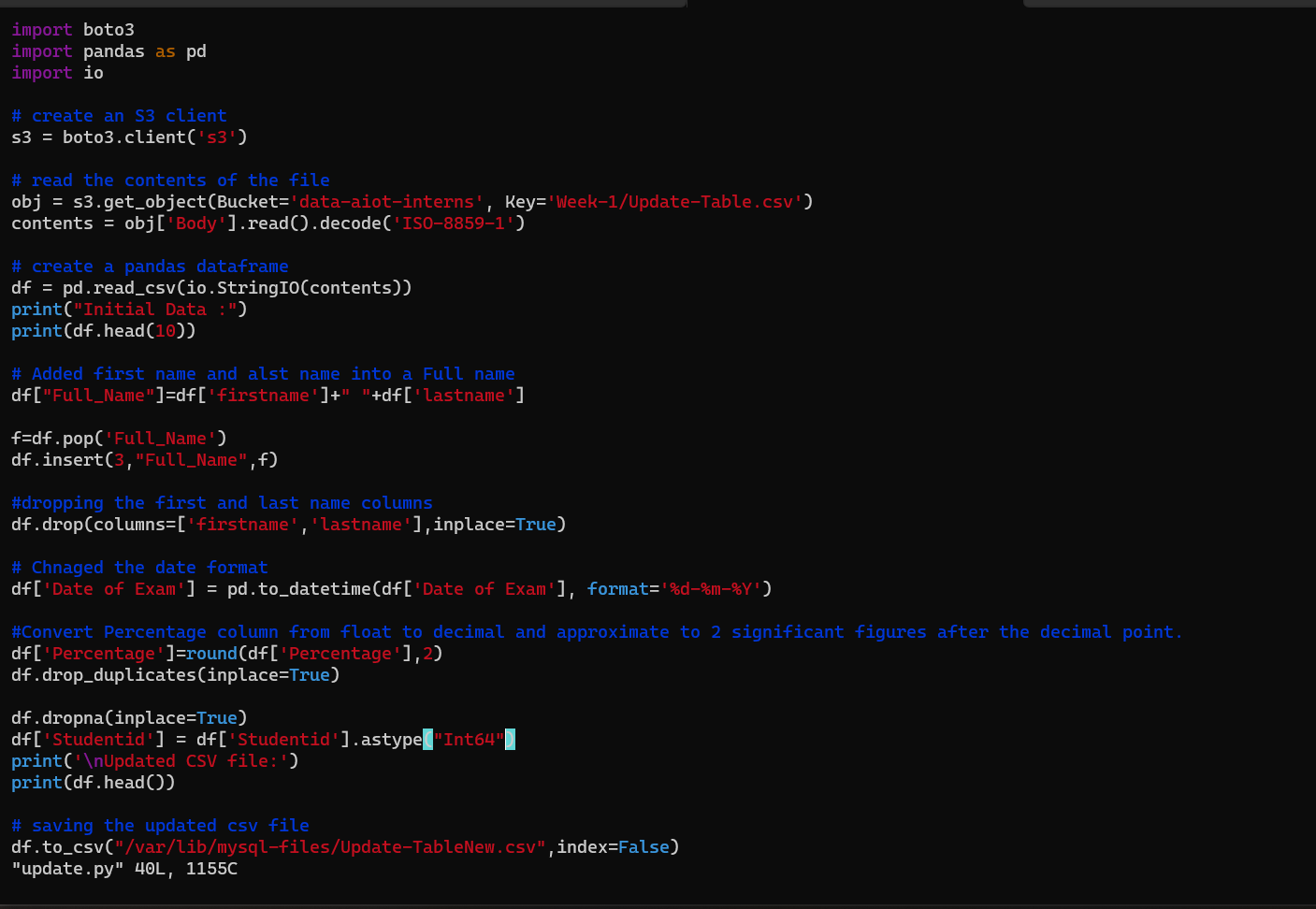
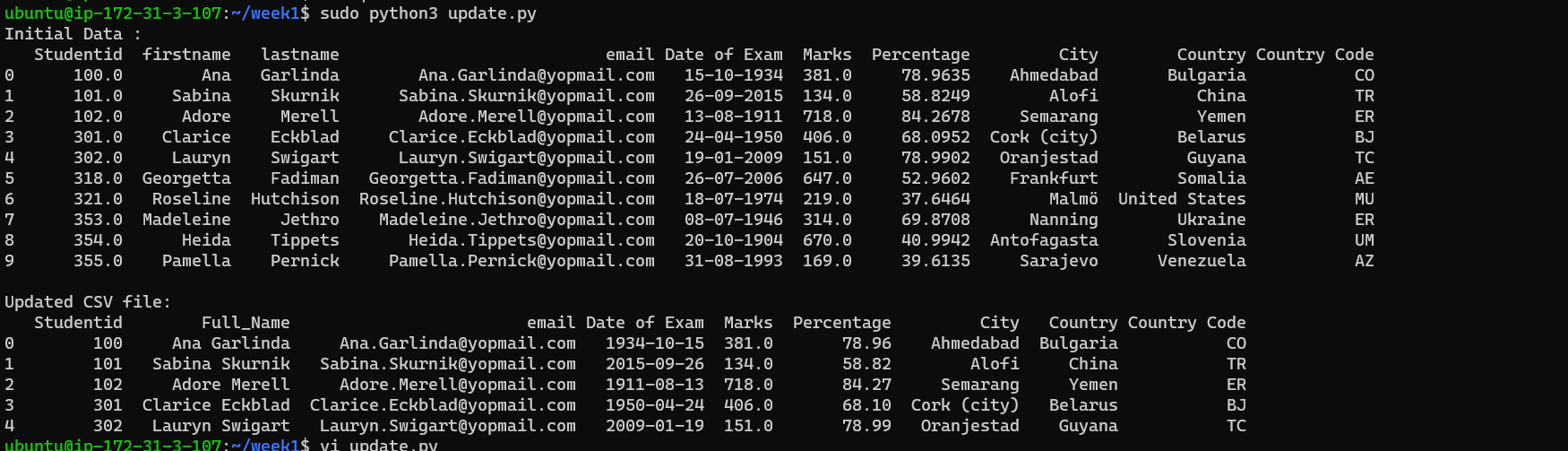
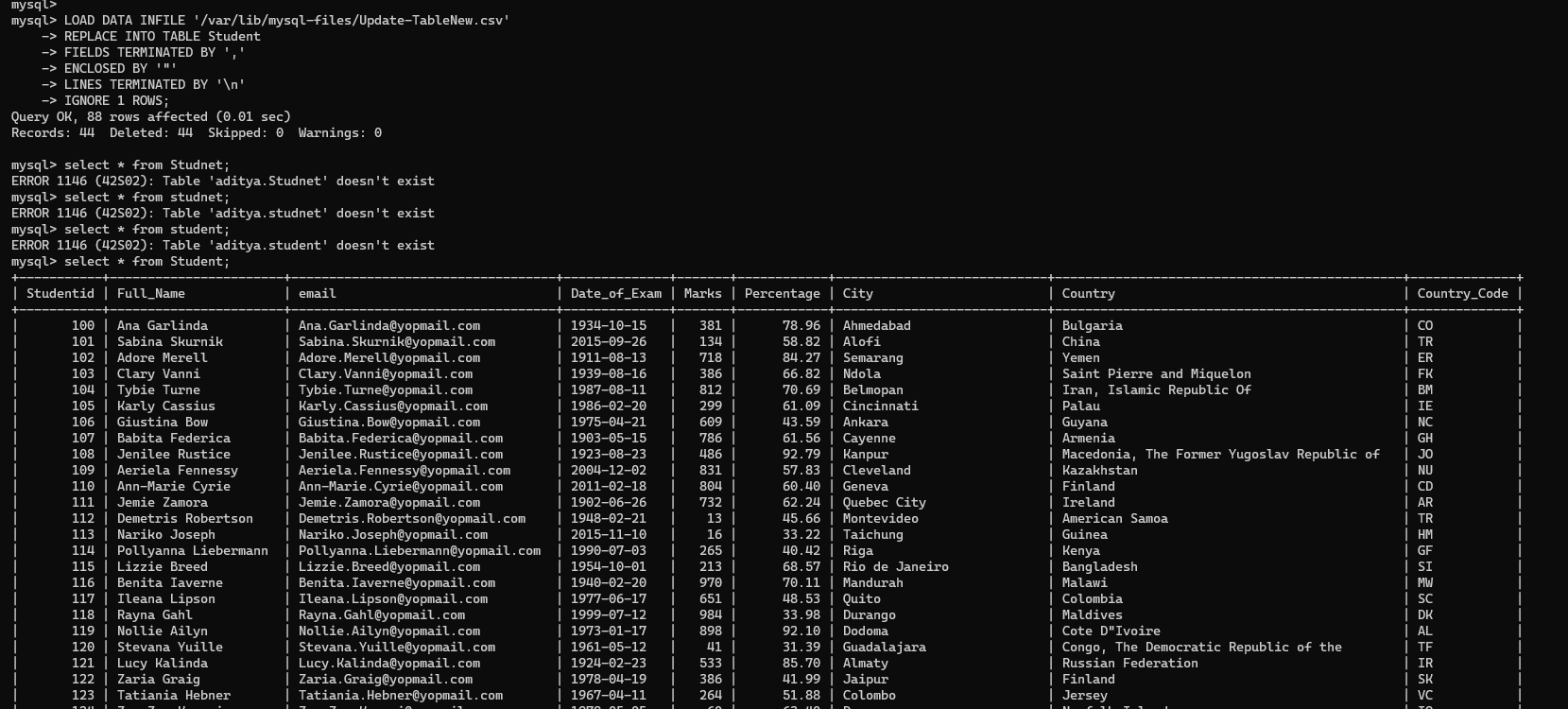




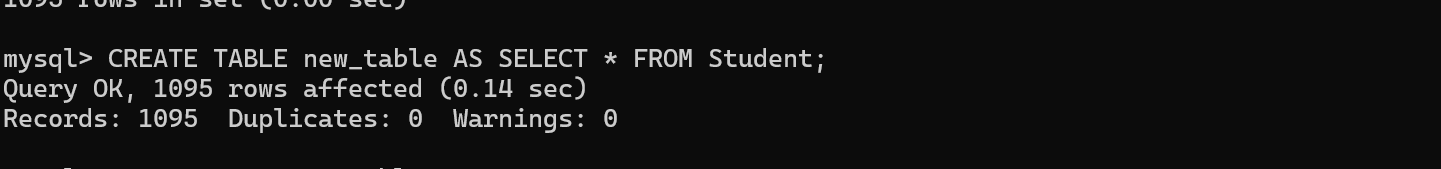
# Task 10: Update the columns mentioned in Update-Table.csv in s3.

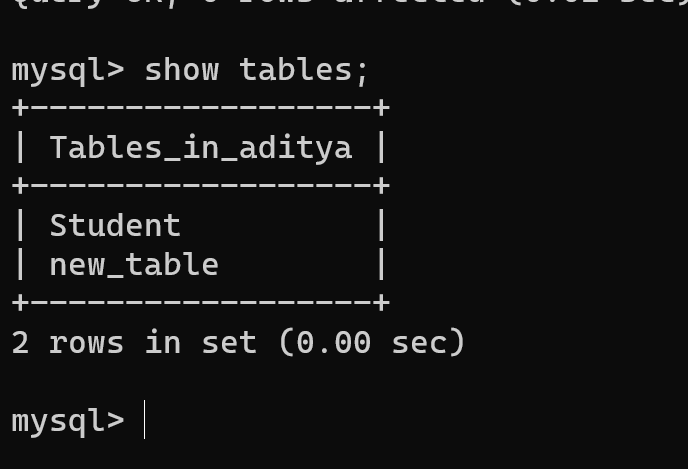
Step 1: Created a python script to read data from s3 and save it. Also make some changes like convert percentage column to upto 2 decimal point and combined first name and last name as a full name column.

Saved file will be used to update rows in the student table.

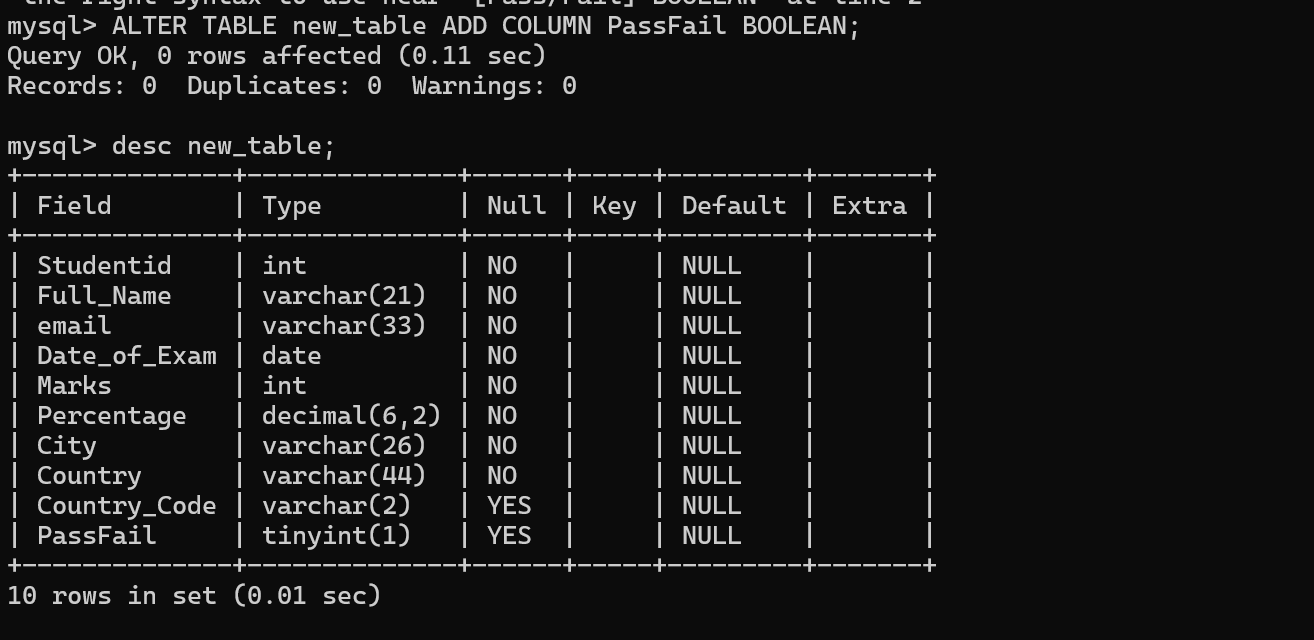
Step 2: Running the script to makes some changes according to requirement and save that csv file.Step 3: Updating table using update.csv file.

# Task 11: Create a new table and copy contents from students-table.

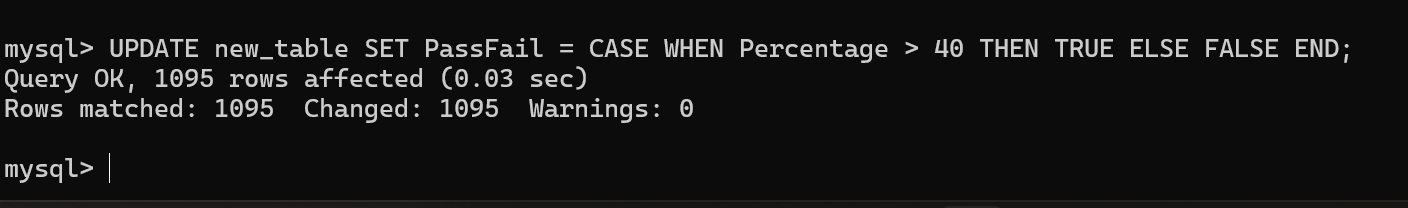
Created a new table and copy contents from student table.Viewing the new table.



# Task 12: Alter table schema and add a new column.

Alter table and adding new column name as PassFail. And describi

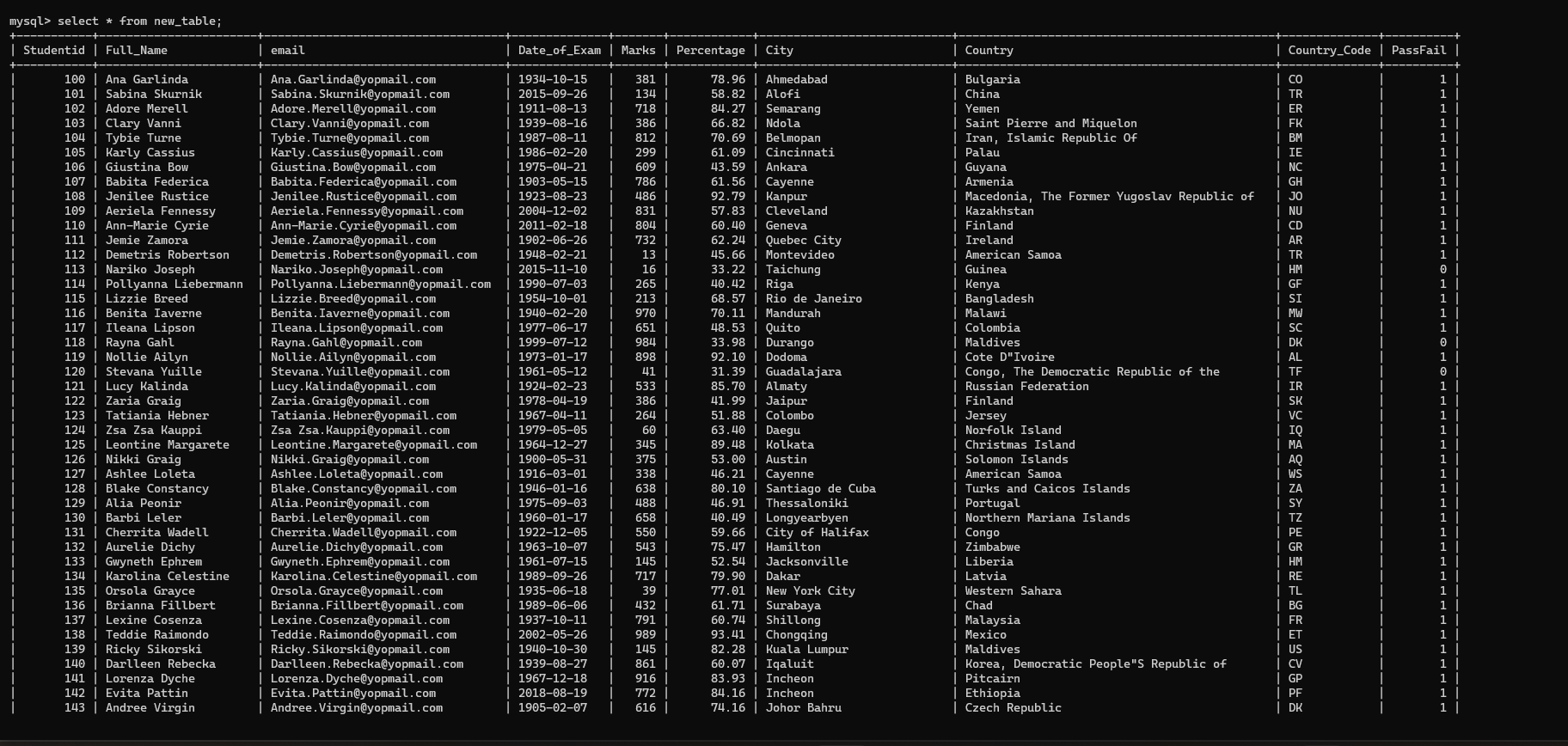
# Task 13: The column name should be Pass/Fail with Boolean data type. Percentage > 40 True else False.

Updating according to requirement.

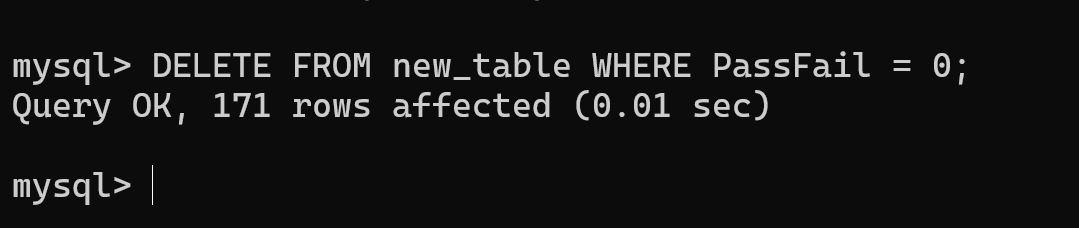
# Task 14: Show contents of new table.

Viewing the content form new table.

mysql> select \* from new\_table;

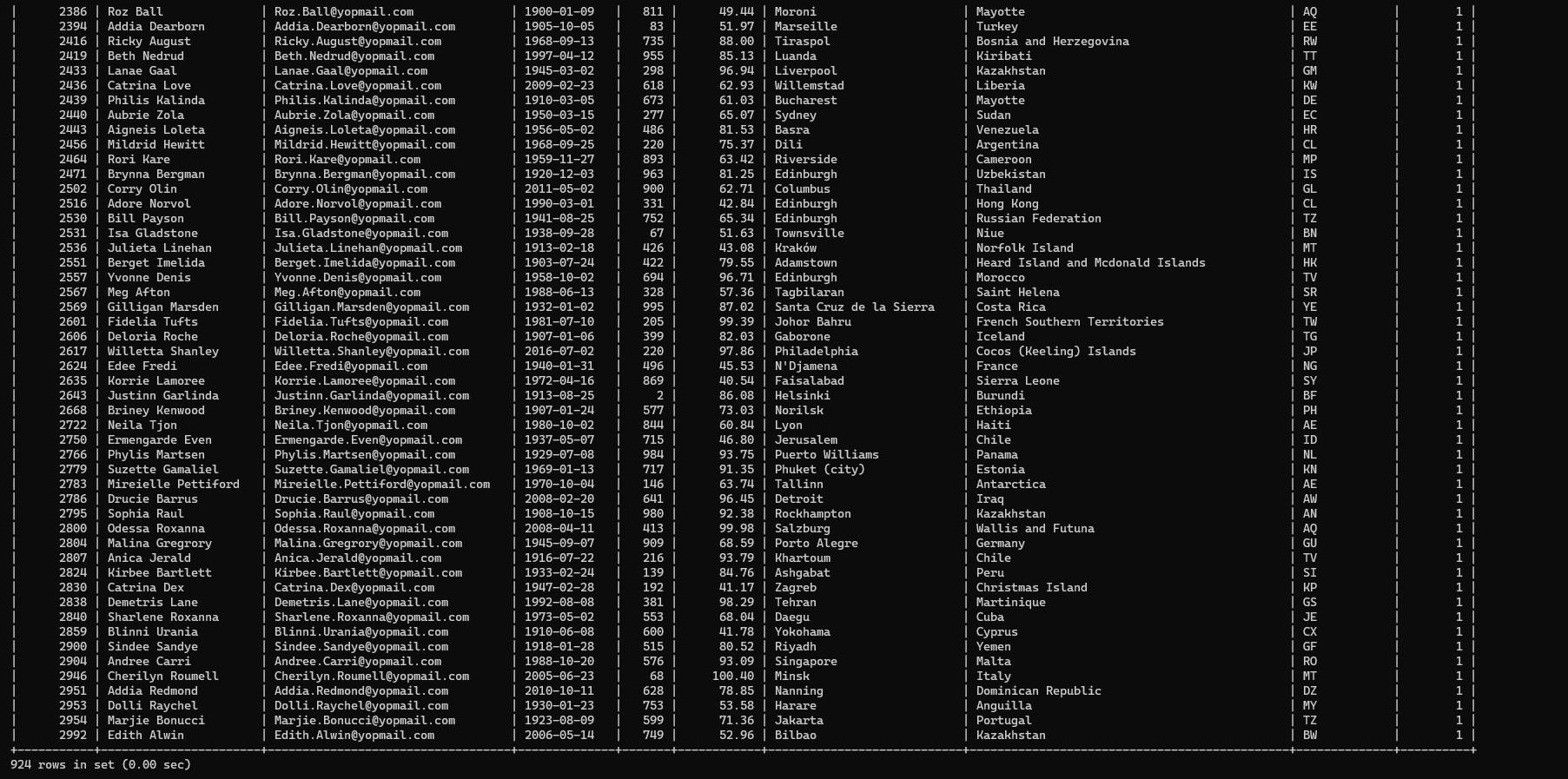


# Task 15: Delete records who failed.



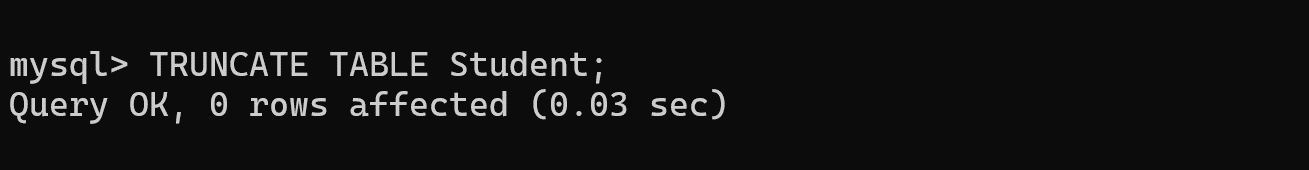
# Task 16: Show contents of new table.

mysql> select \* from new\_table;

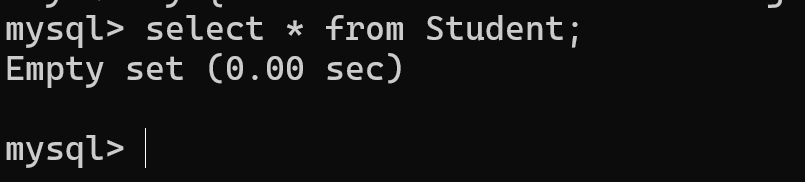
# Task 17: Truncate the students-table.

mysql> TRUNCATE TABLE Student;



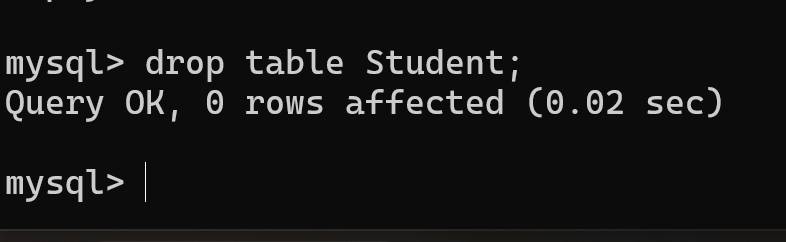
# Task 18: Display its contents.

mysql> select \* from Student;



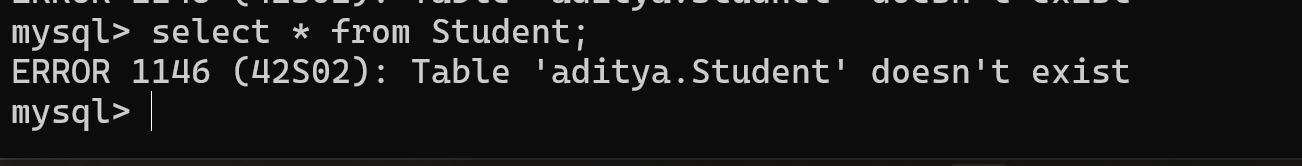
# Task 19: Drop Students-table.

mysql> drop table Student;



# Task 20: Display its contents.

mysql> select \* from Student;

Table does not exit because it was dropped in previous stage. Verifying the “student-table”via listing all the available table.