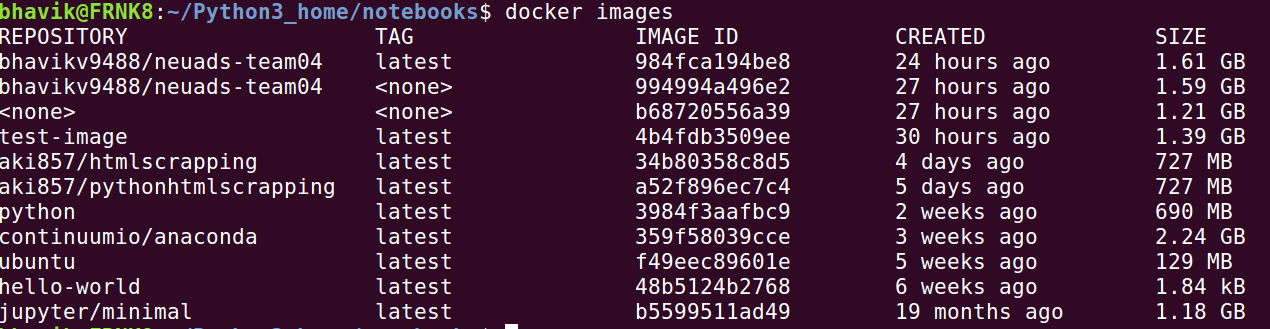
To start with the docker execution :

1. List the docker images available on the system

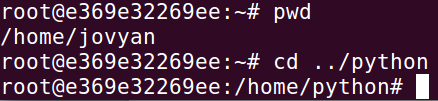


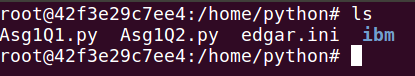
2. Run the docker image you need to execute with its Image ID available from the above step.



3. Now you have logged on to the Docker image. Change the working directory to /home/python. This the directory where both the source scripts and the related files are available

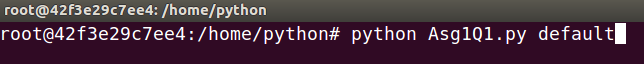
(PS : Default directory is /home/jovyan as created by the maintainer of the image)





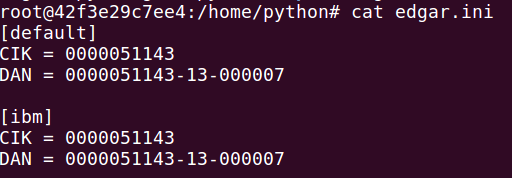
1. To execute the code run the script as shown below with the argument as required by the script

**Execute Asg1Q1.py** - here default is the argument which runs for the IBM’s CIK and DAN number. (This is the section in the .ini file under which IBM’s details are set)



To execute for another CIK and DAN number, the edgar.ini file can be modified and the required CIK and DAN number under a new section can be added. Execute the script with the argument as the new section name created in the .ini file

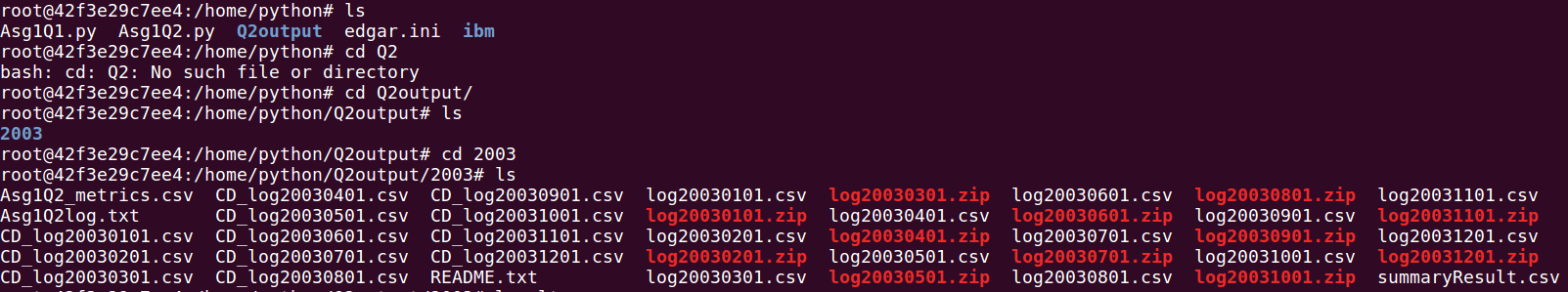
(As seen below default and ibm are the section numbers)



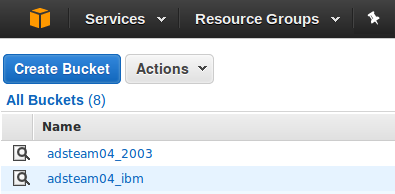
**Execute Asg1Q2.py** - Here 2003 is the year to be passed for the script to extract data for.



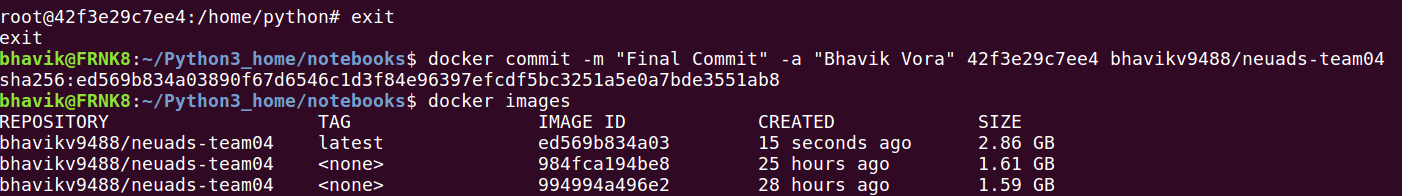
Seen below is the output directory and its contents post execution of the script



Also, there is a bucket created on the S3 with the name adsteam04\_<section> which is the output for part1 and bucket adsteam04\_<YEAR>, output for part2.



This directory will have the cleaned and combined output file, a log file of the script and the output file generated by the summary metrics script which is embeded in the same Q2 script file (Asg1Q2.py)



The docker was successfully committed, stopped and pushed to the docker repository.

