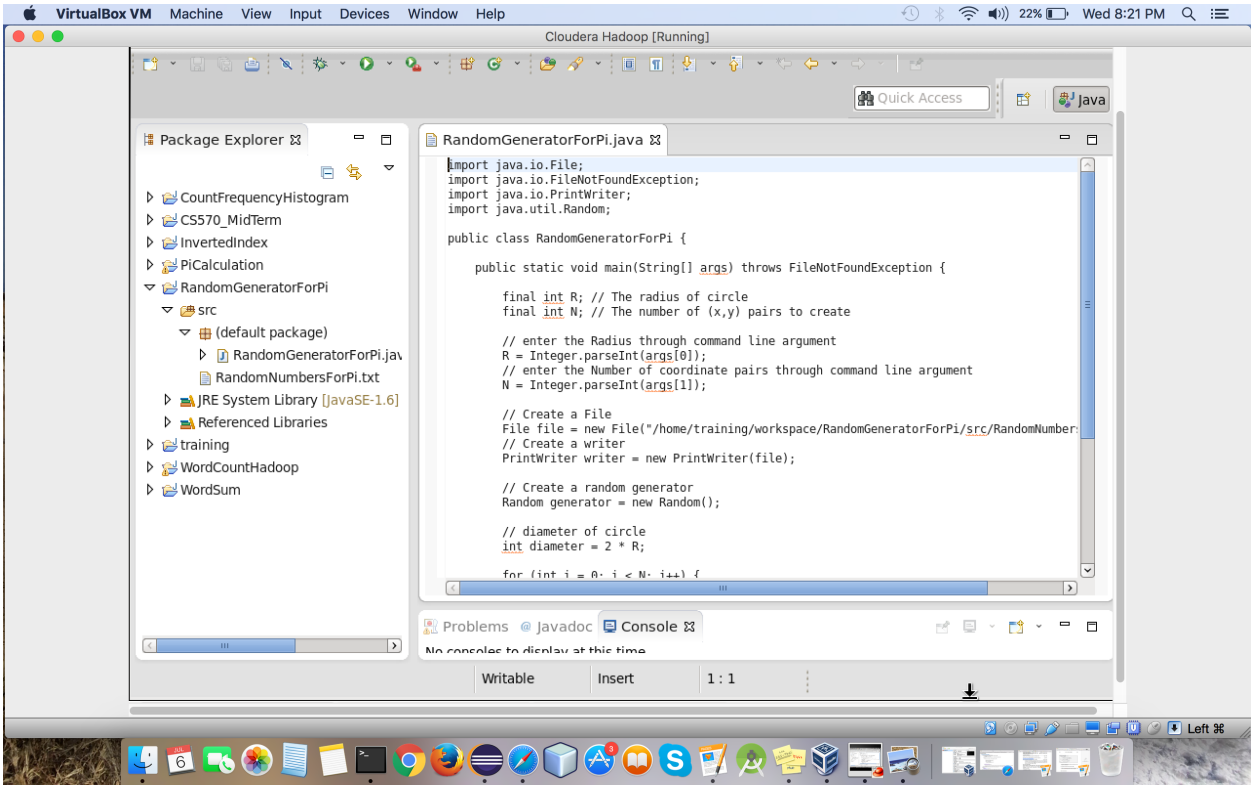
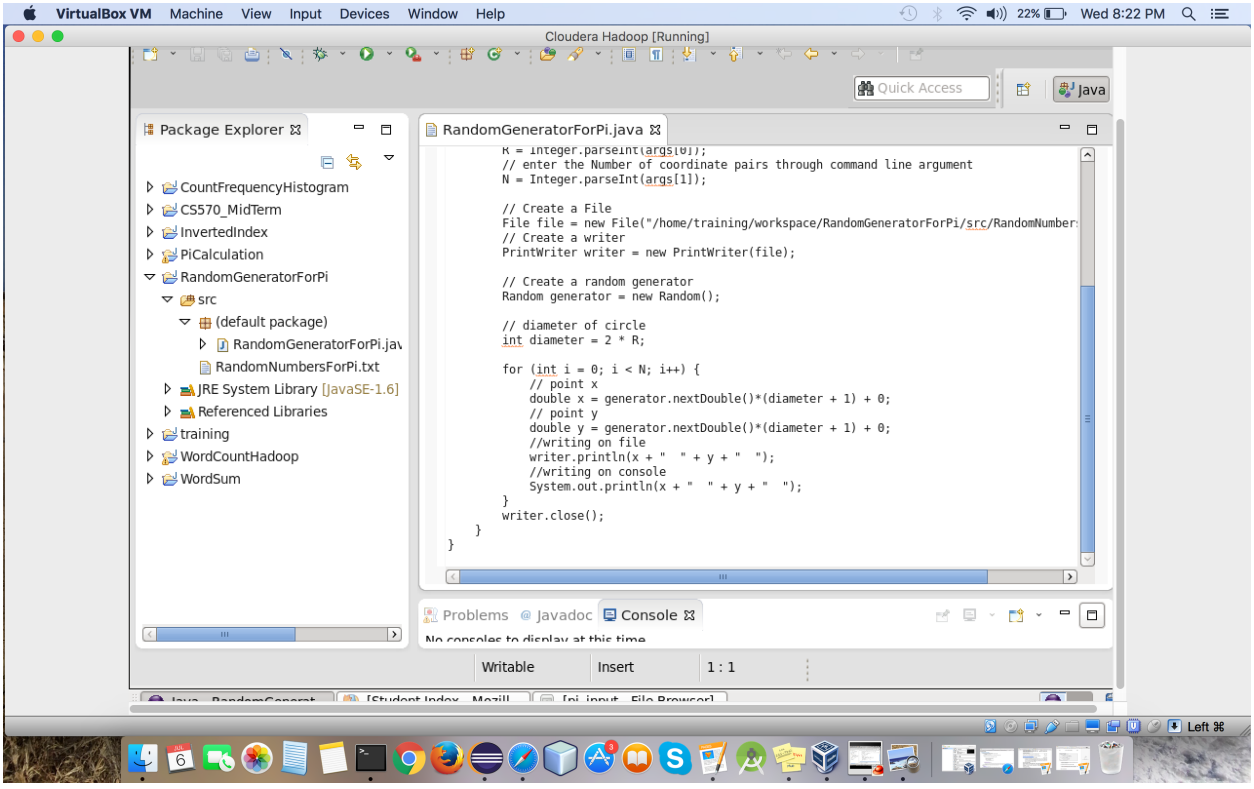
Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com

Generating a PI MapReduce program to calculate the PI value on Virtual Machine- Cloudera and then running it on Amazon Web Service:

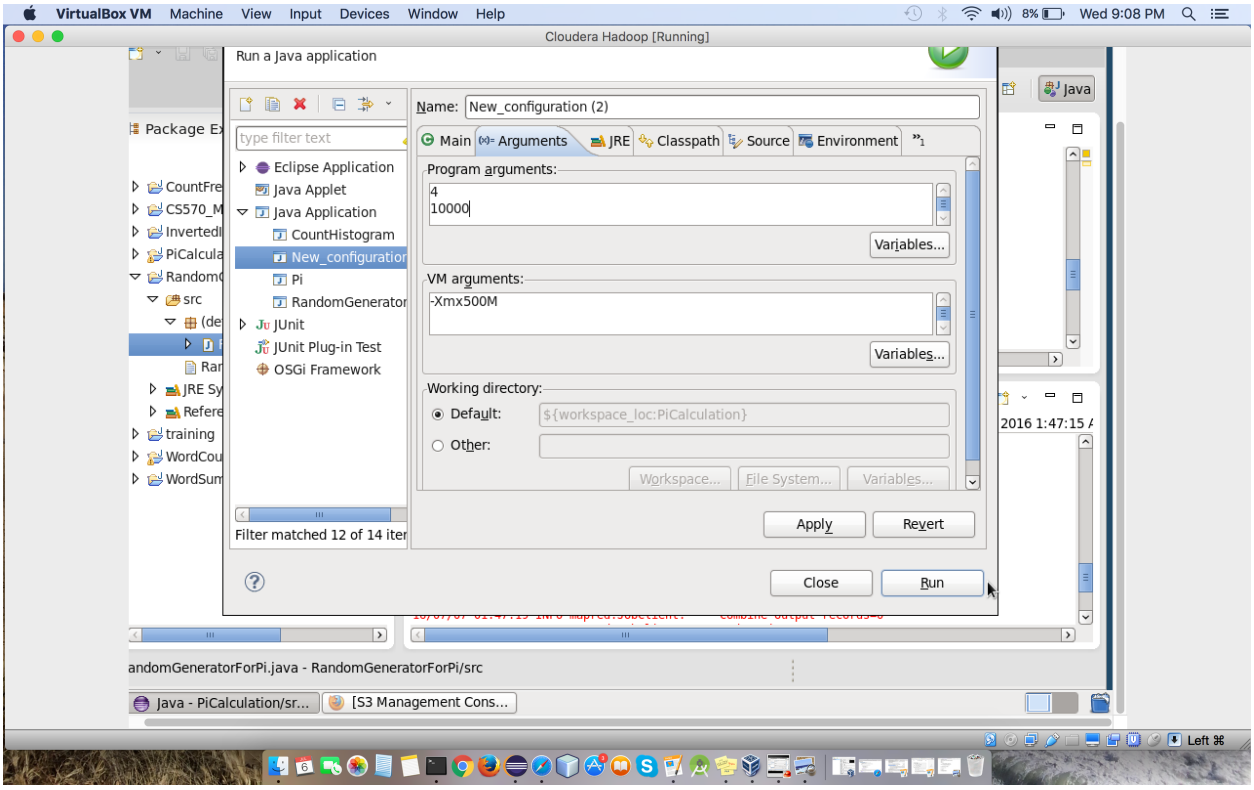
Create RandomGeneratorForPi.java class



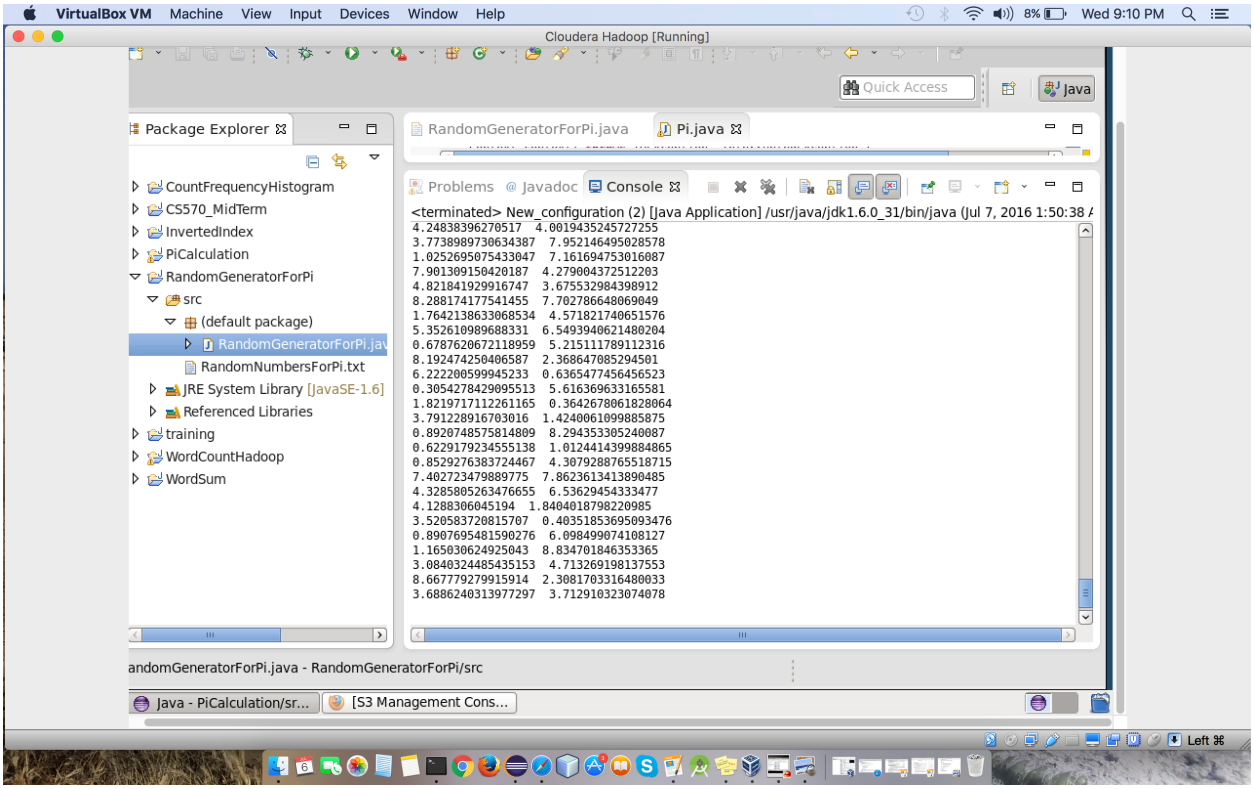


Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com

Enter Arguments as Radius=4 and Number of pairs to generate=10000:



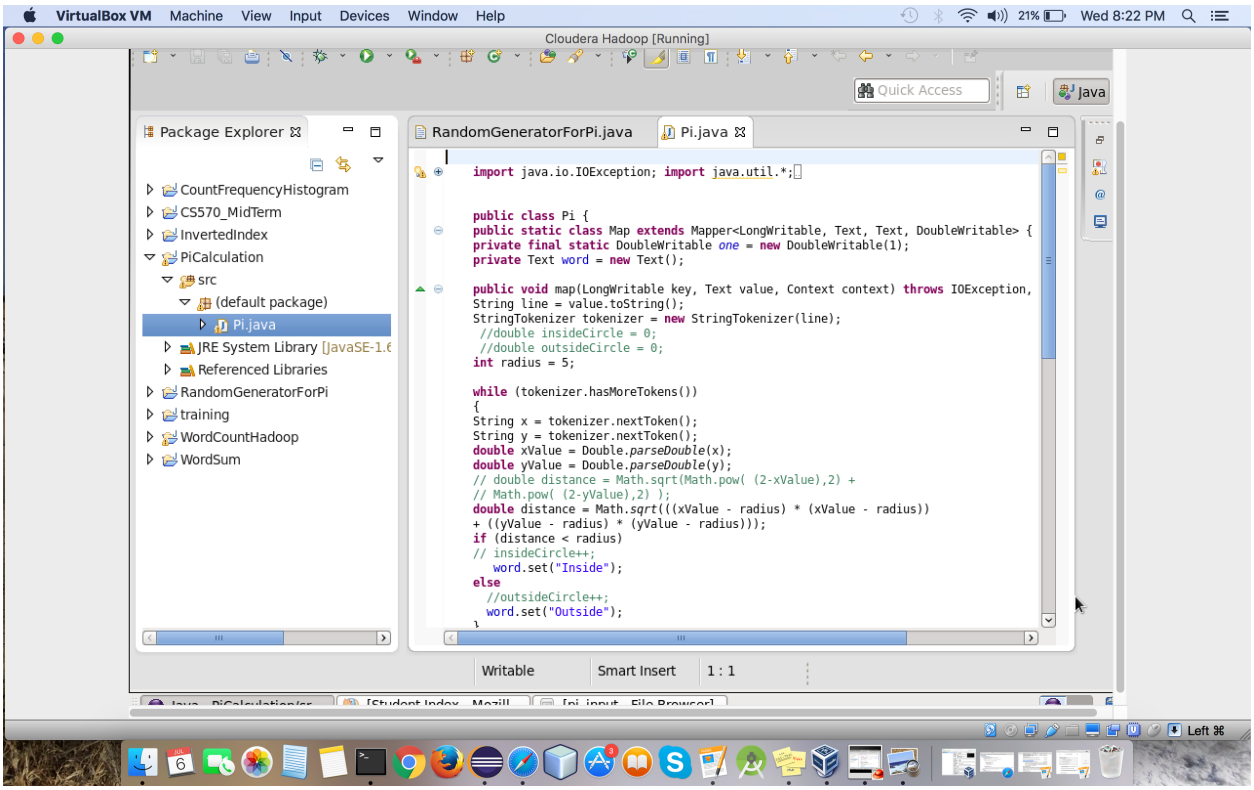
Run the RandomGeneratorForPi.java class in command line and check the output:

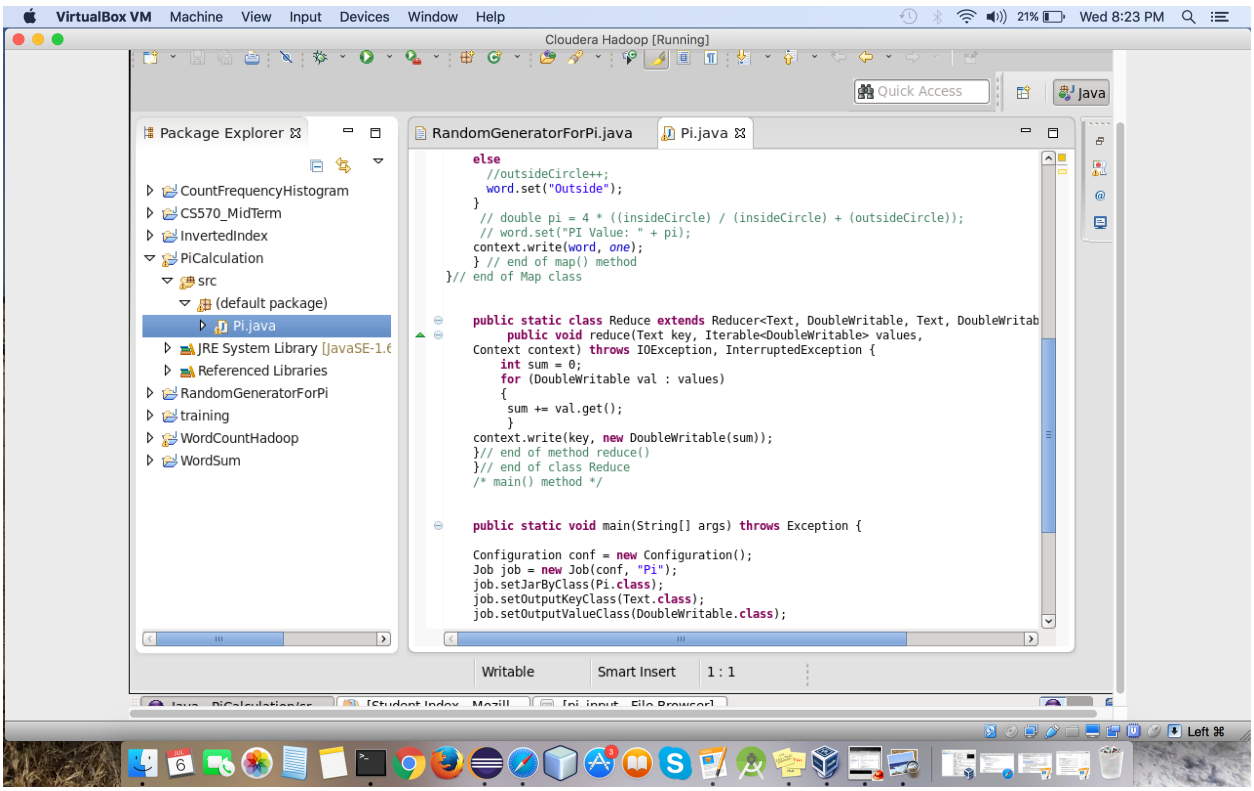


Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com

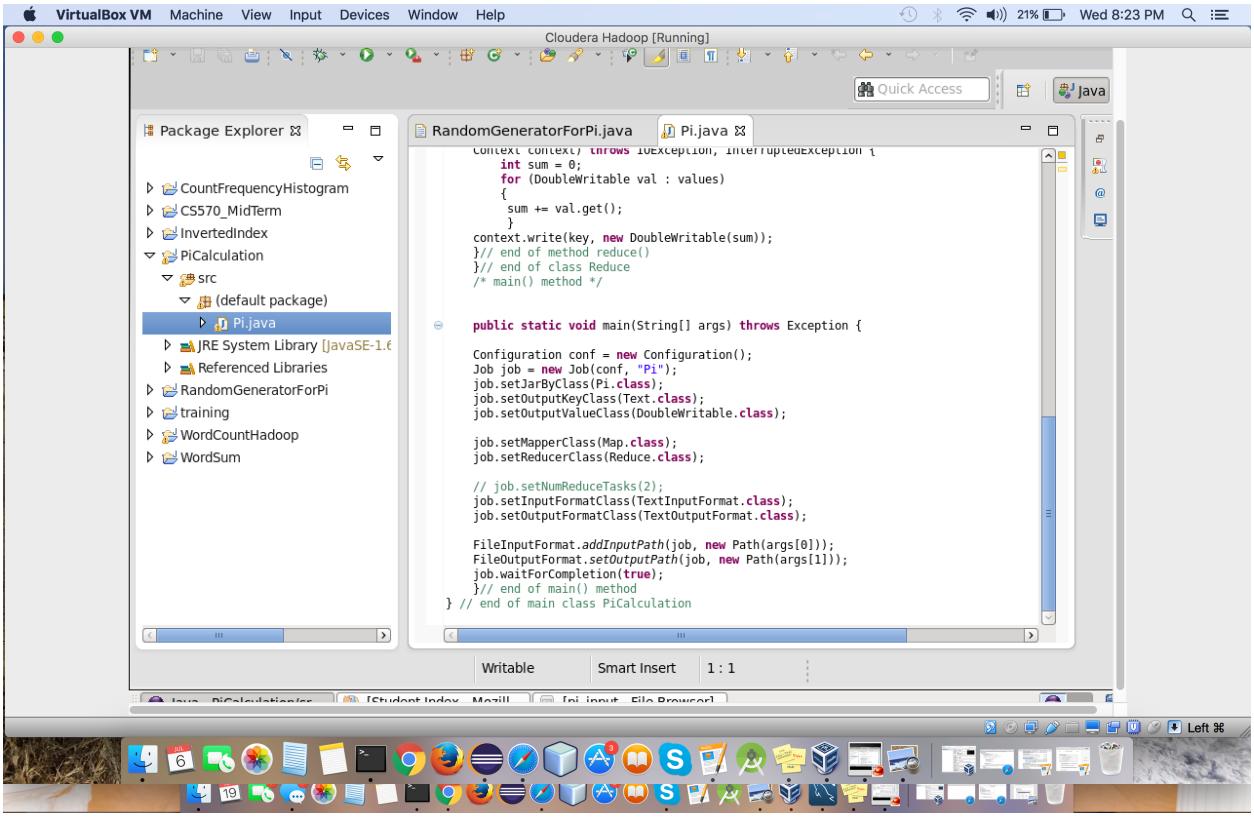
The above output generated is save in RandomGeneratorForPi.txt file.

Create a Pi.java class:

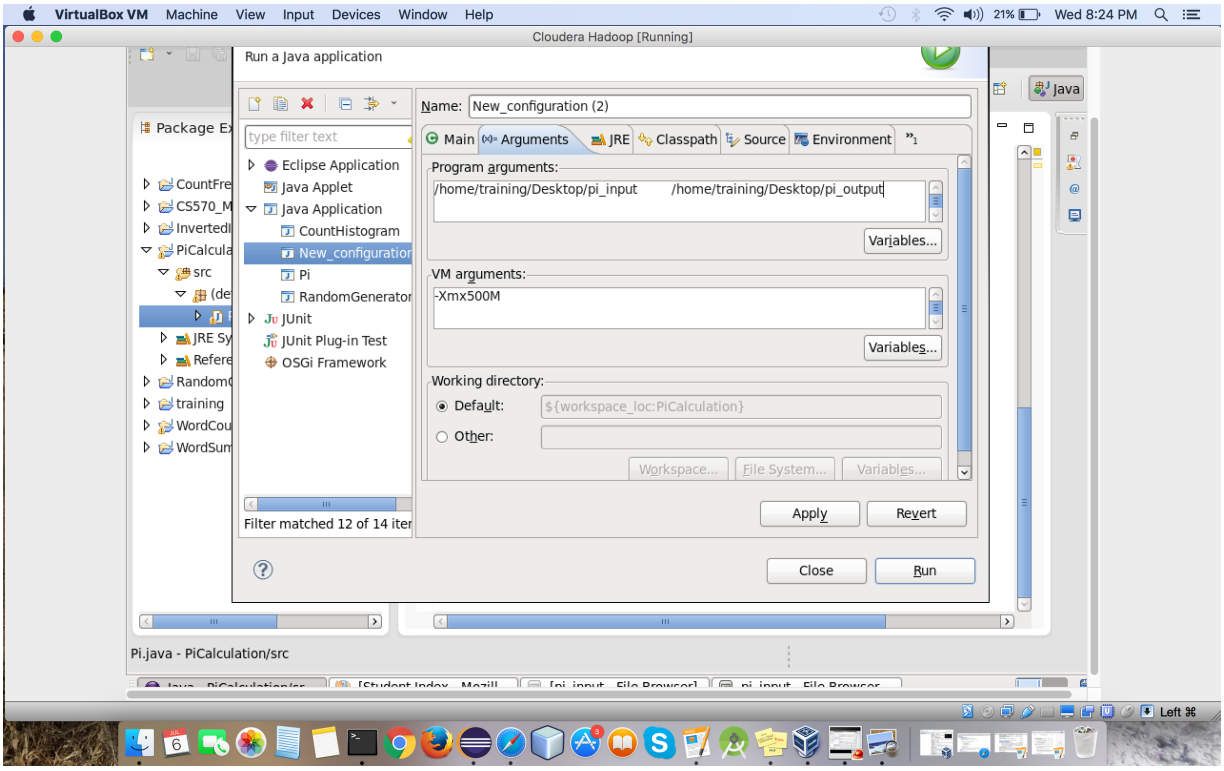




Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com

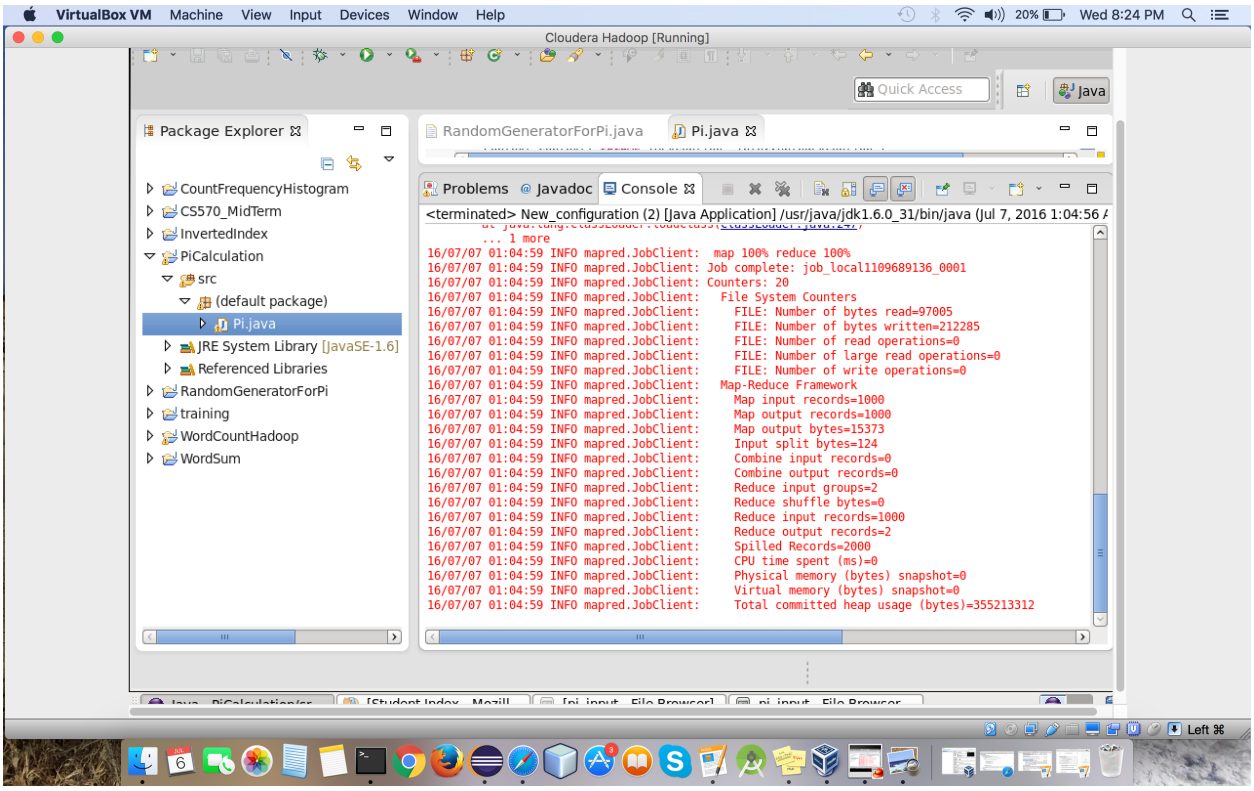


In the argument, provide the input file path where RandomGeneratorForPi.txt is saved and the path where output file will be automatically generated:

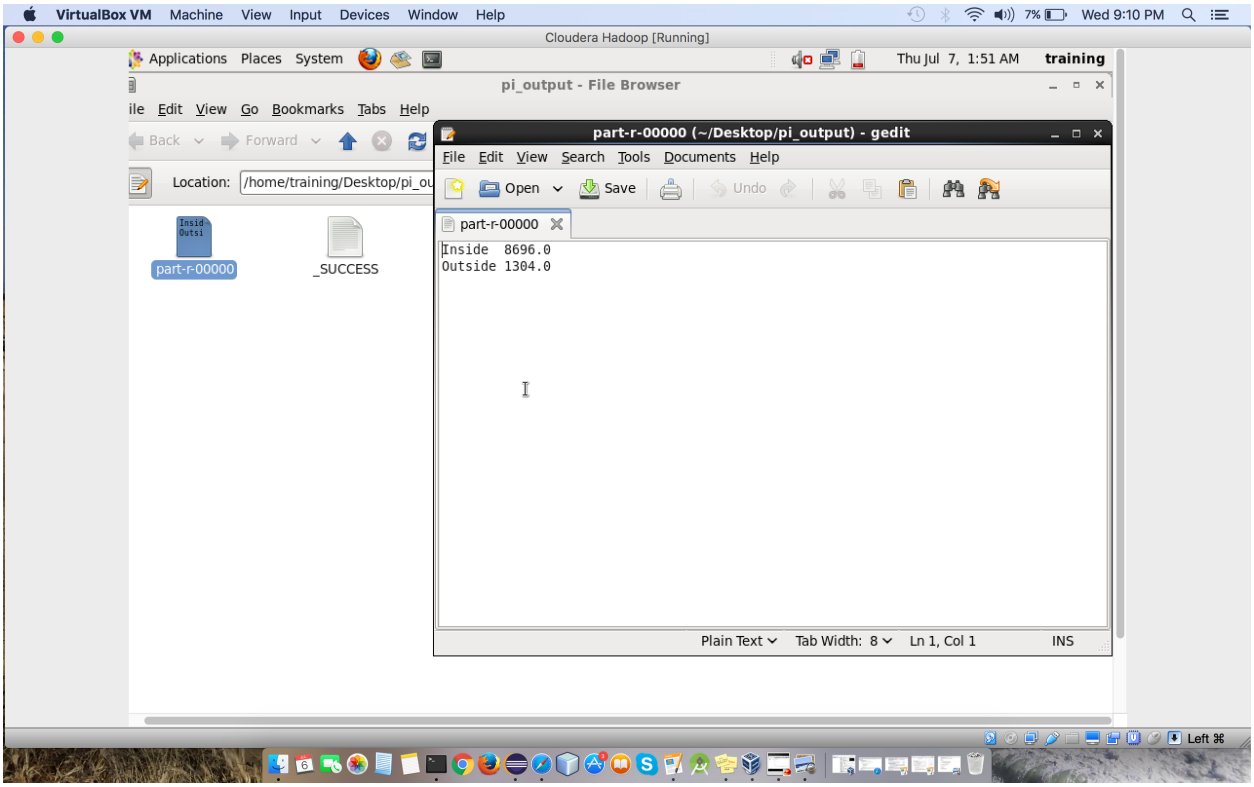


Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com

Run the Pi.java class:

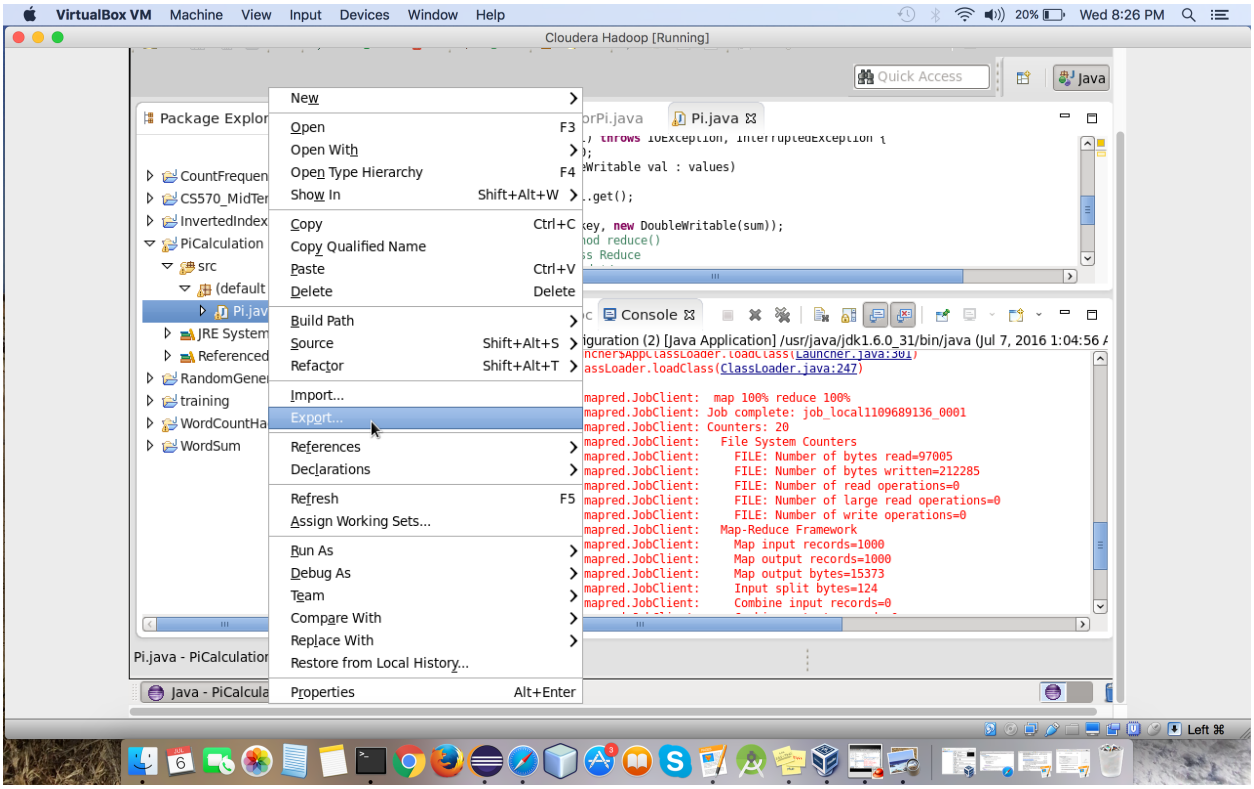


The following output is generated in Pi_output file:
Number of darts inside:8696.0
Number of darts outside: 1304.0

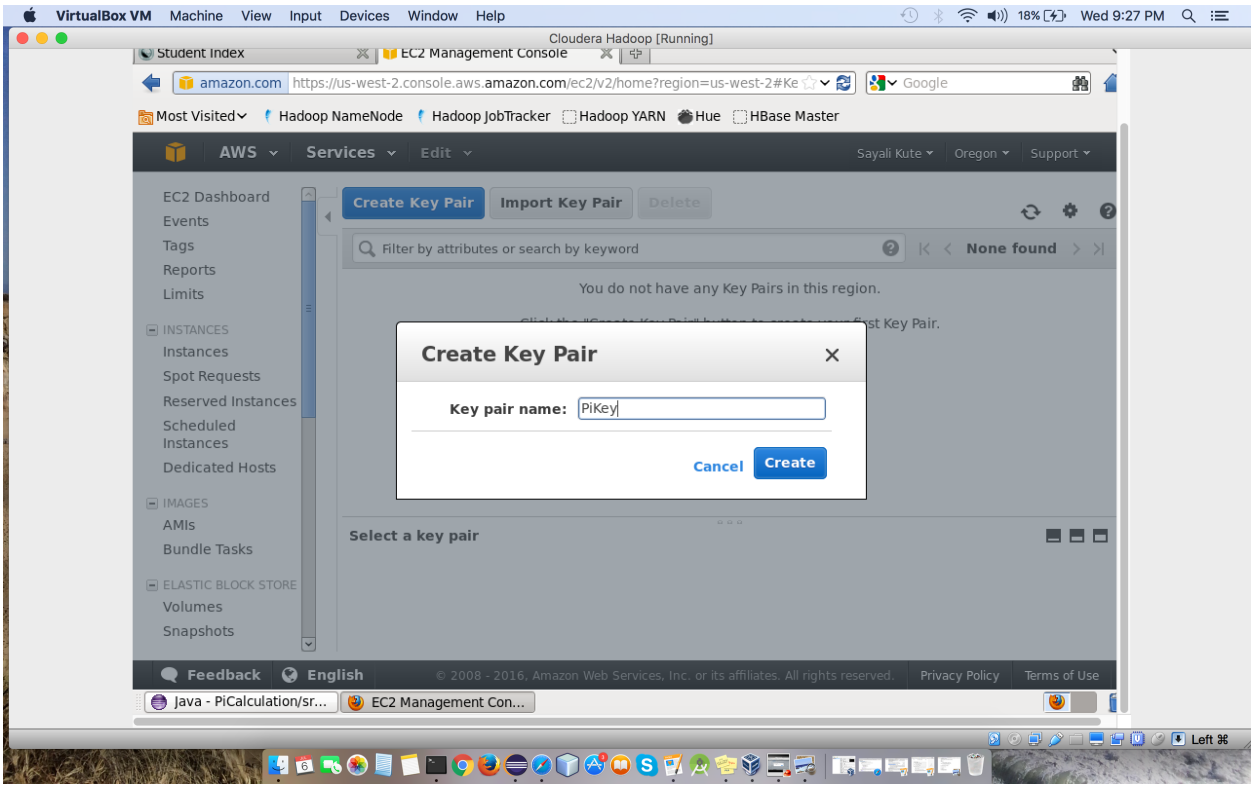


Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com

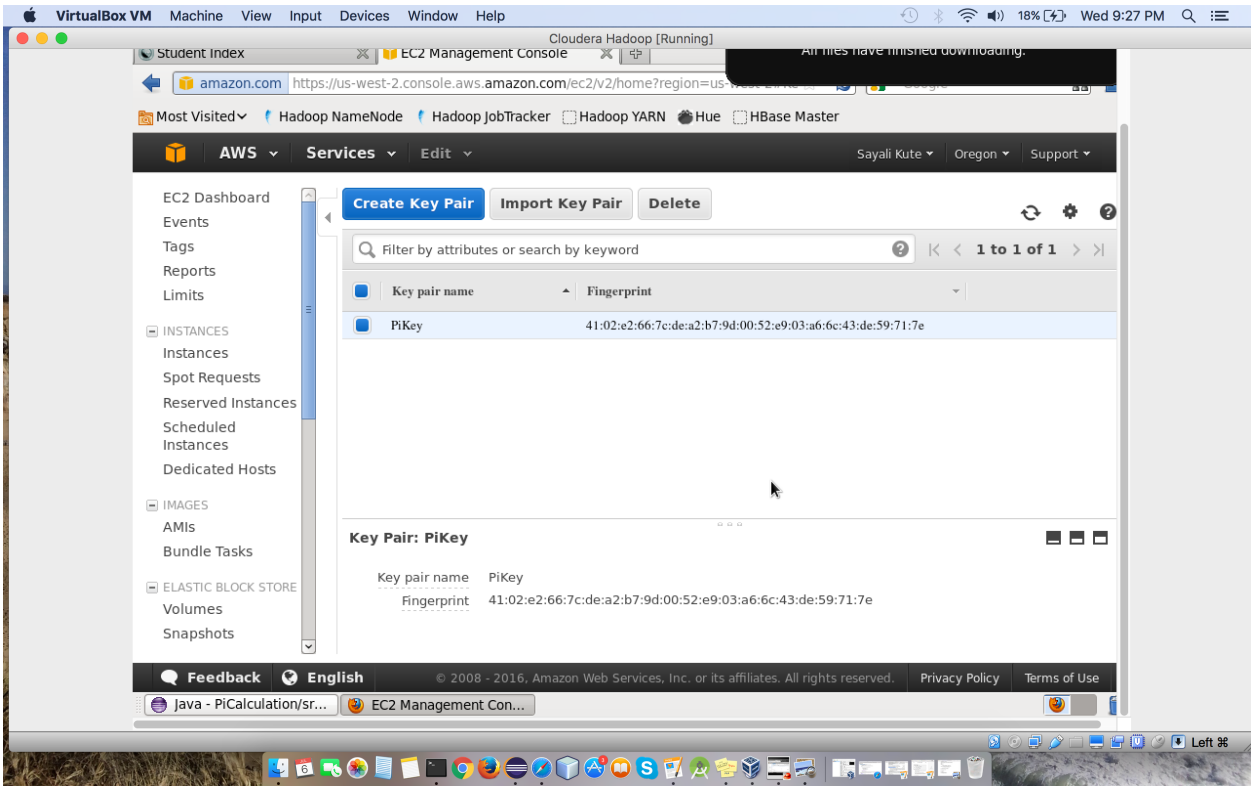
Export the Pi.java class and save it with .jar extension:



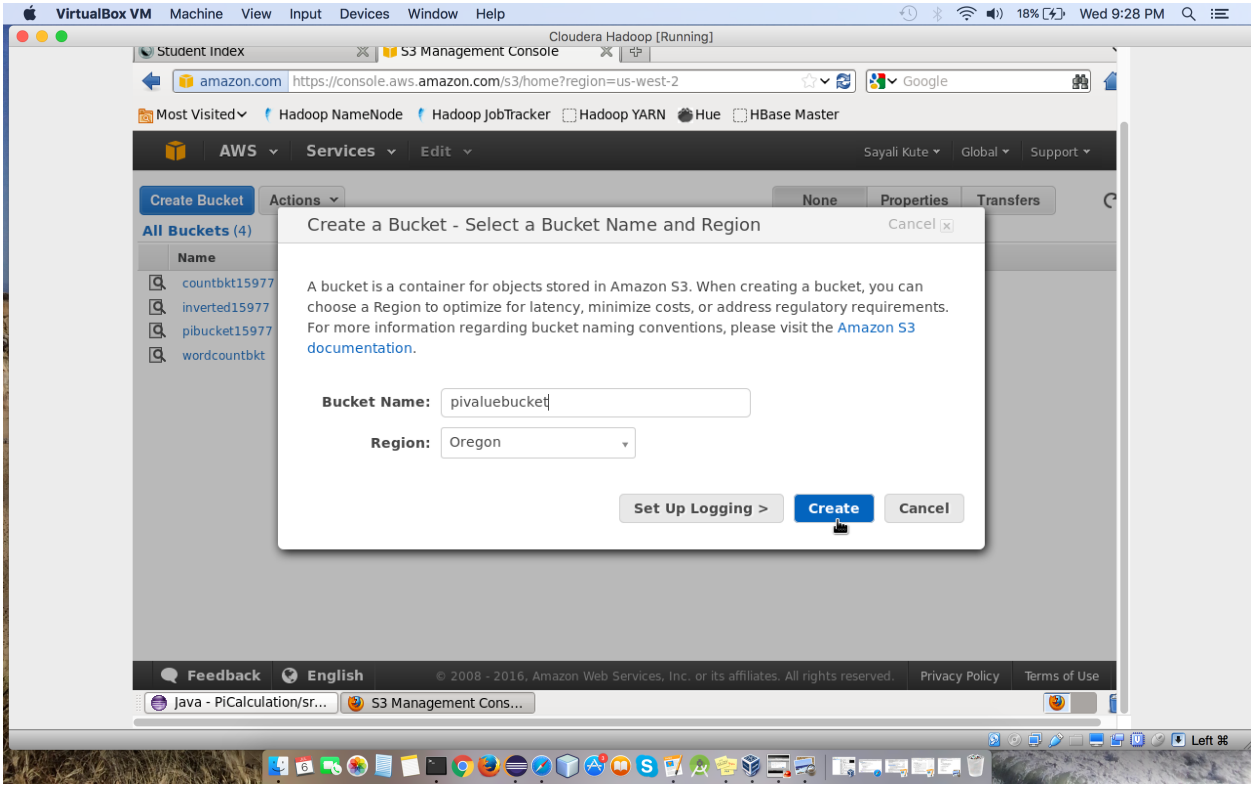
In AWS, create key pair through EC2 service:



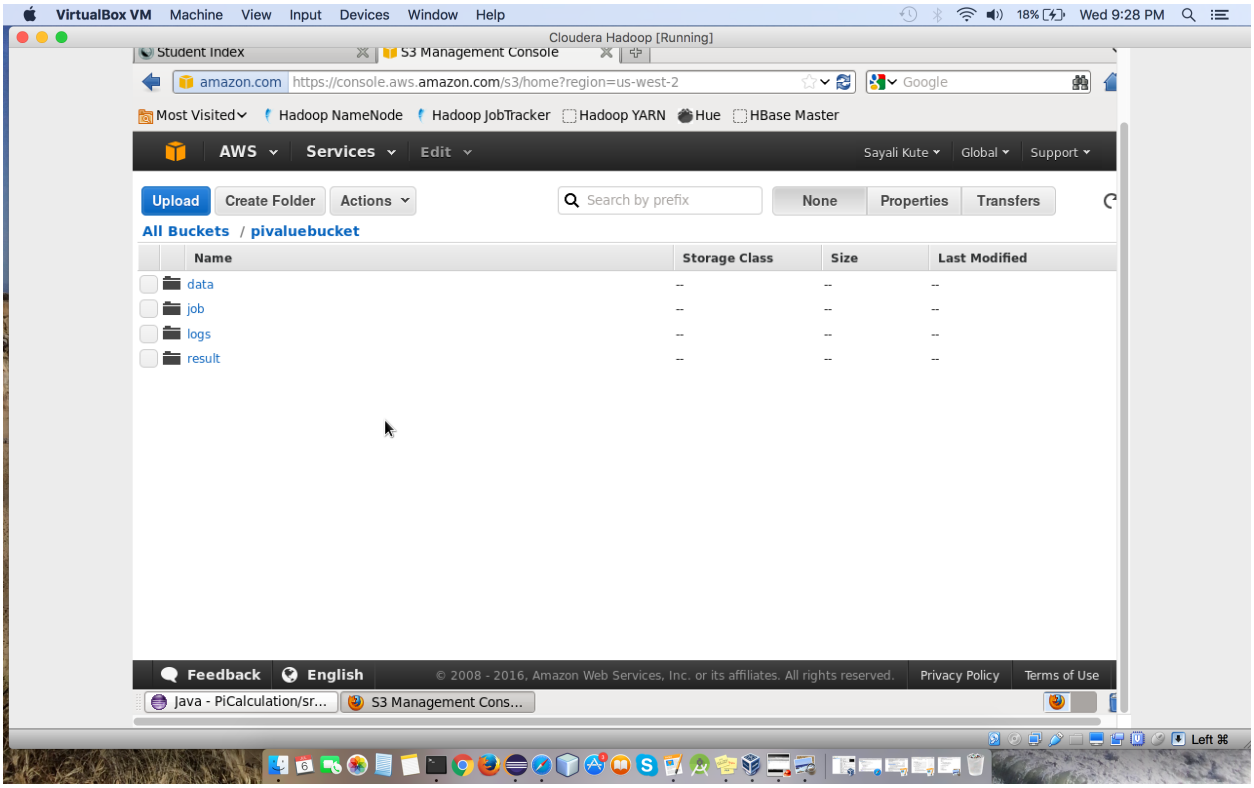
Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com



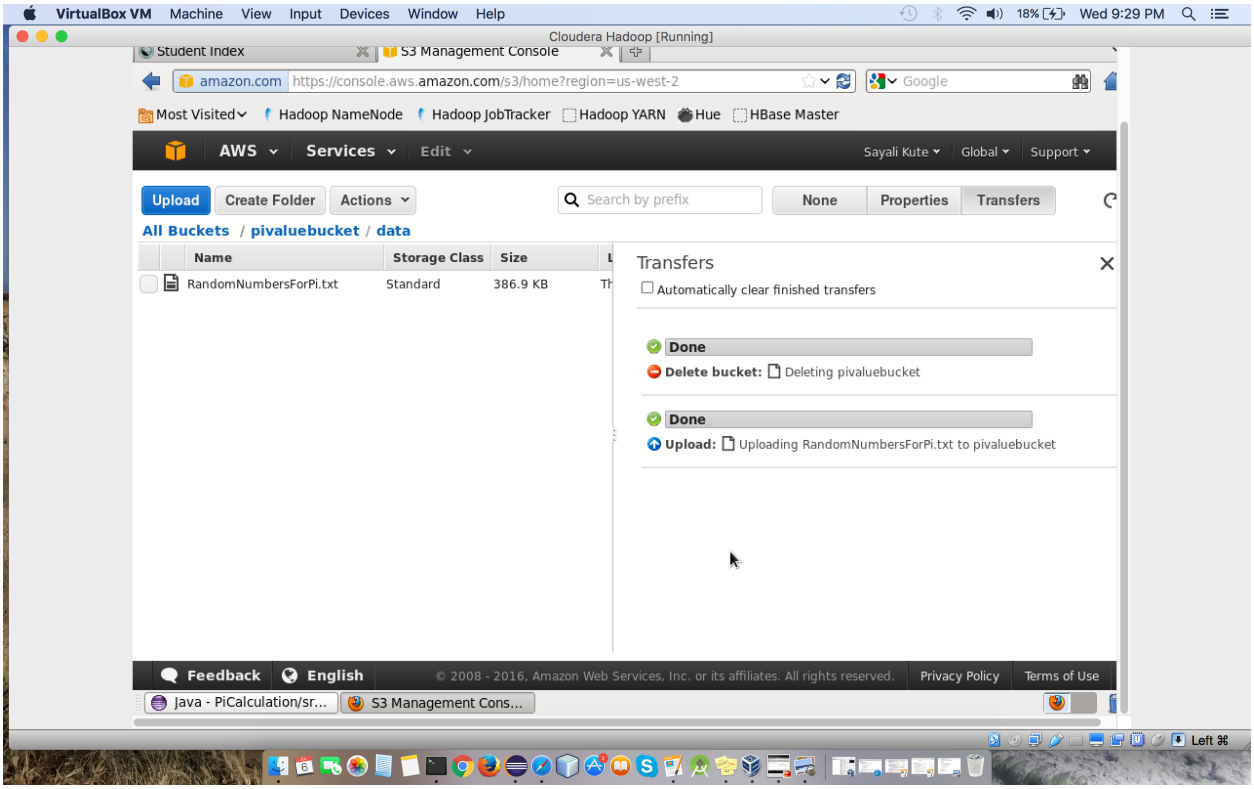
Create Bucket through S3 service:



Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com

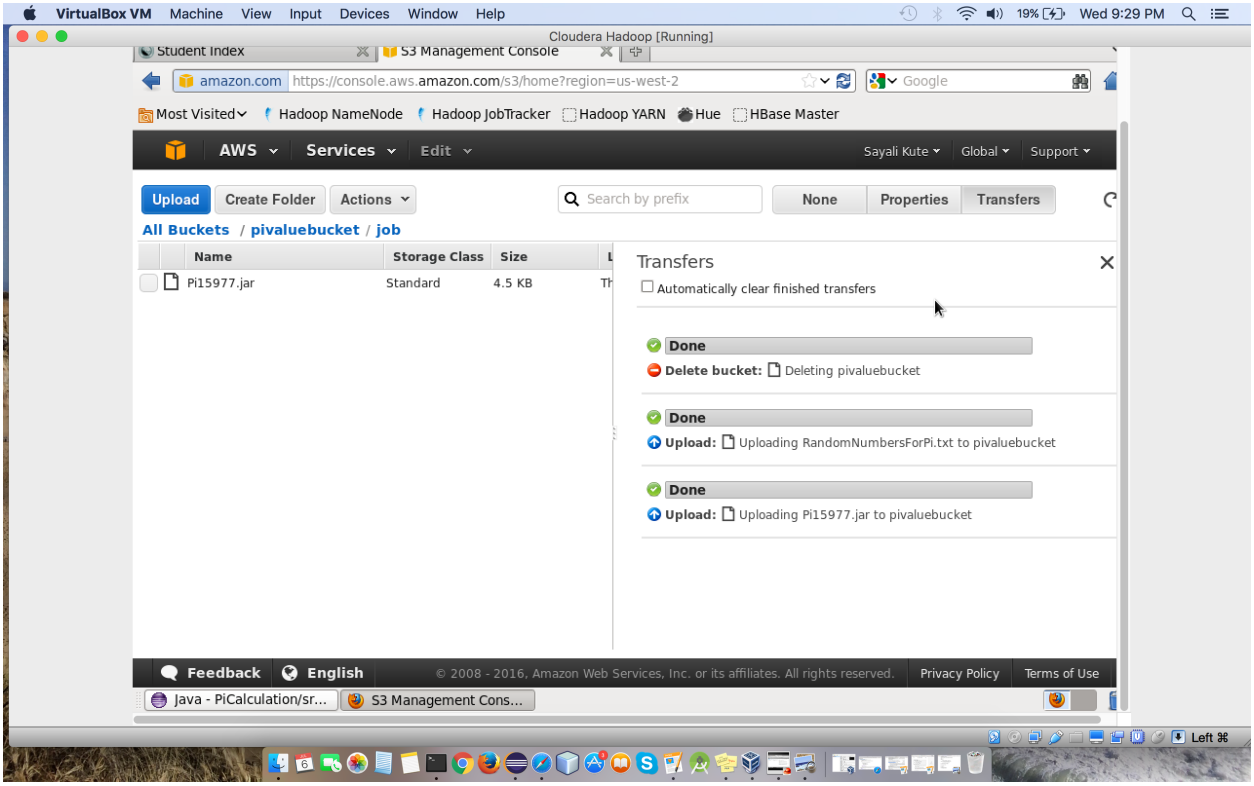


Upload the RandomGeneratorForPi.txt file in ‘data’ folder and exported Pi15977.jar file in the ‘job’ folder.

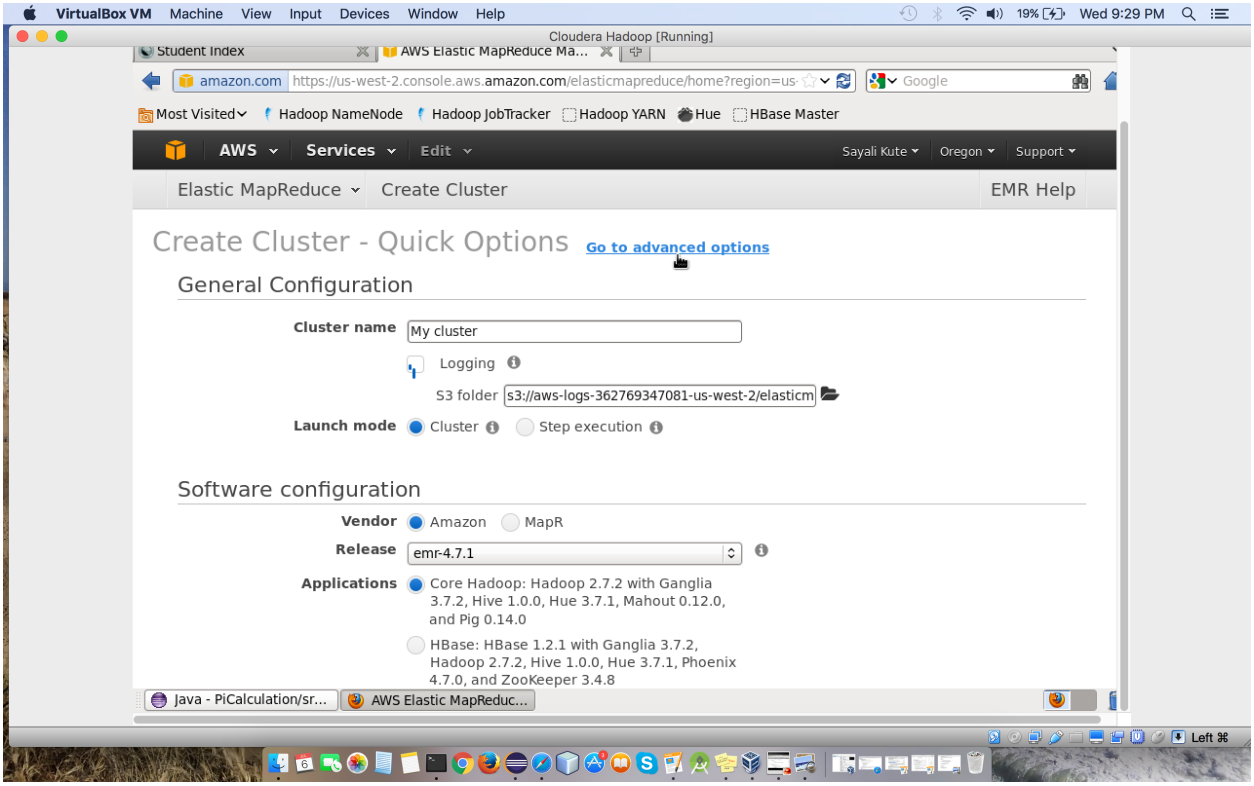


Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com

Create a cluster through EMR service and select all the required

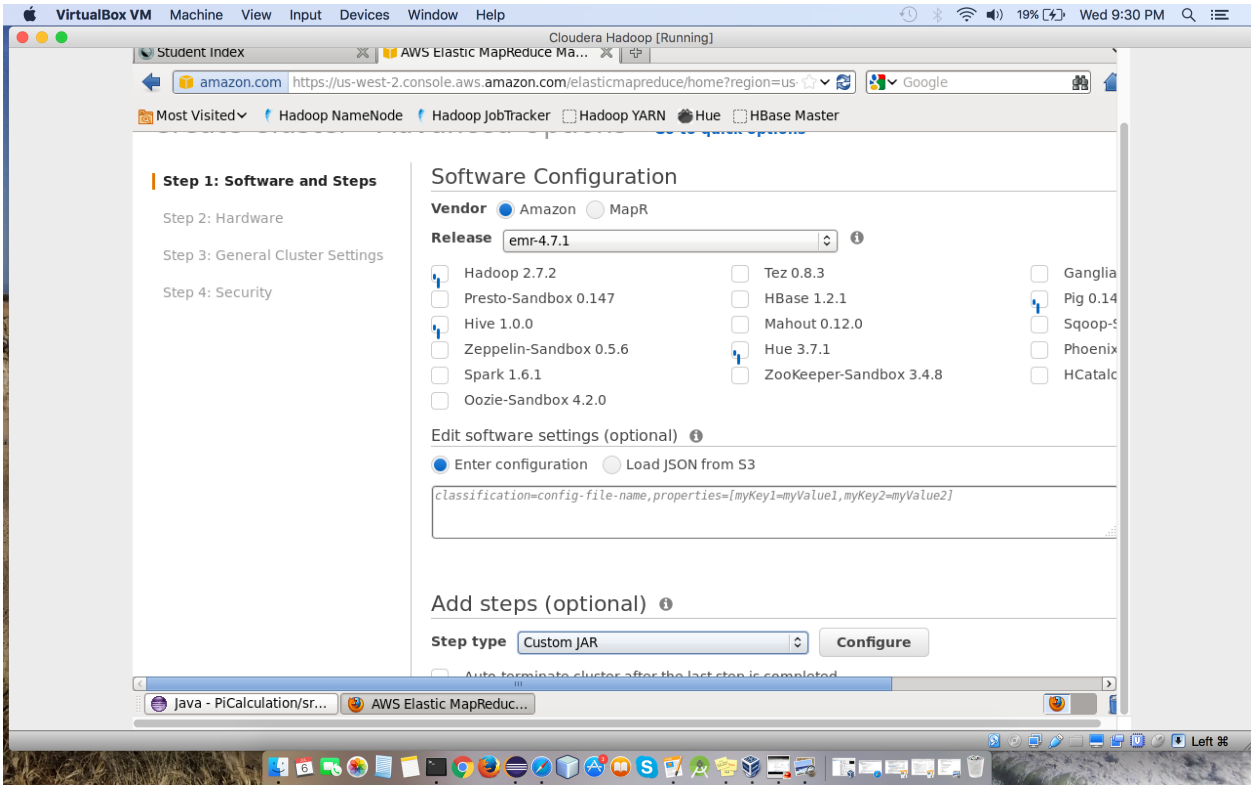


advance choices:

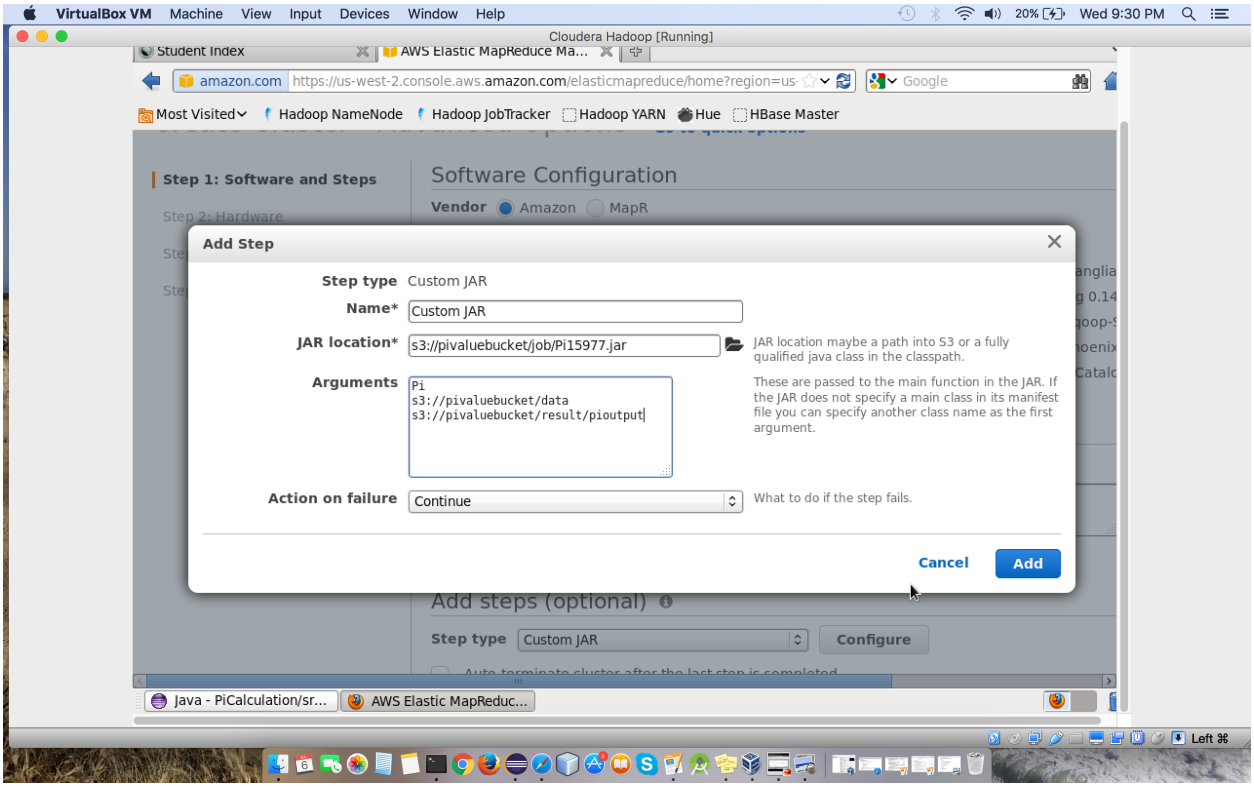


Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com

Select ‘step type’ as ‘Custom JAR’ and configure other settings as follows:

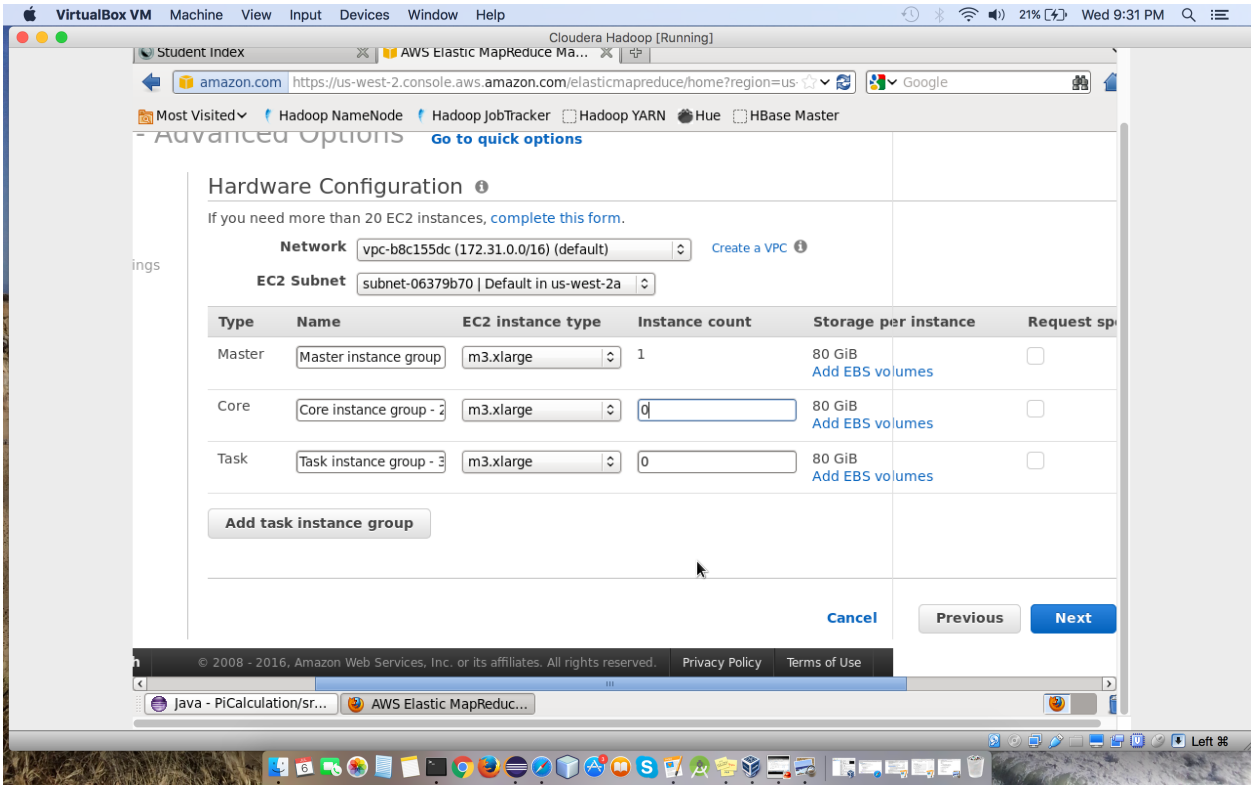


Provide JAR file location path and arguments:

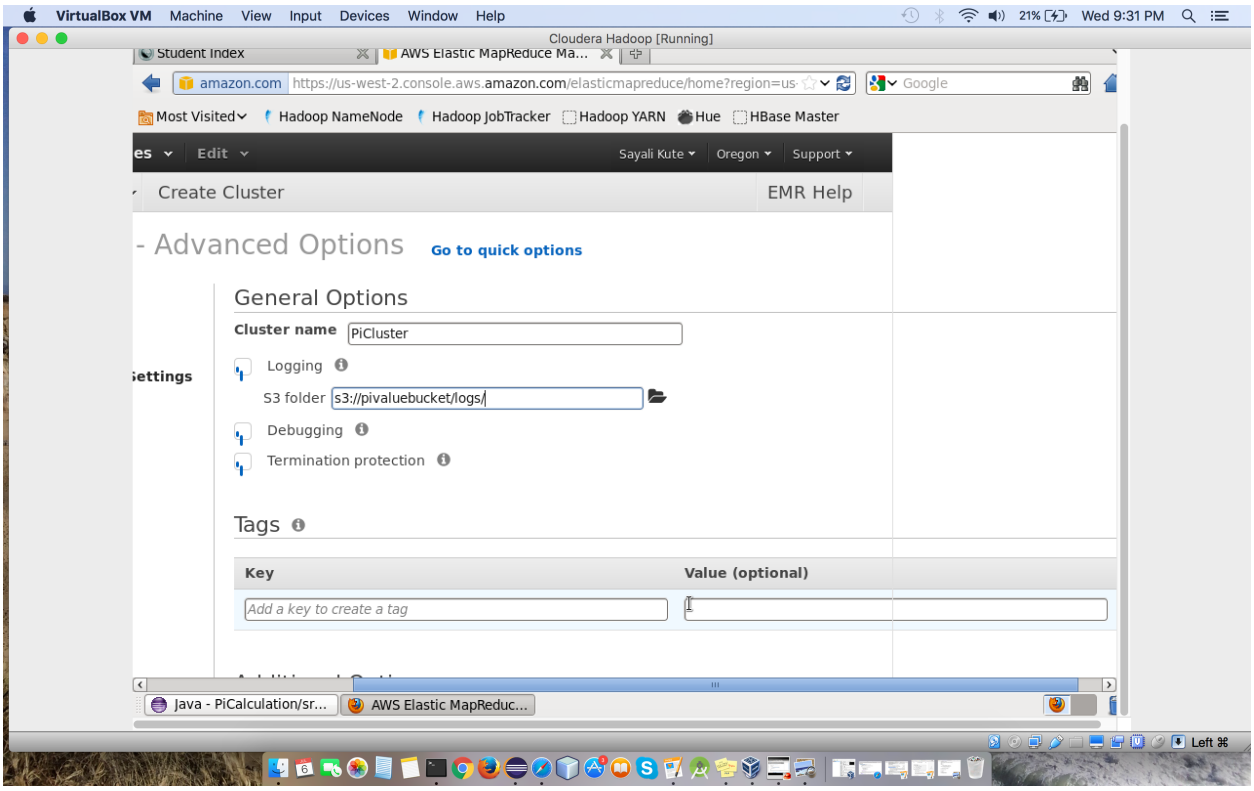


Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com

Make instance count of ‘core’ as zero:

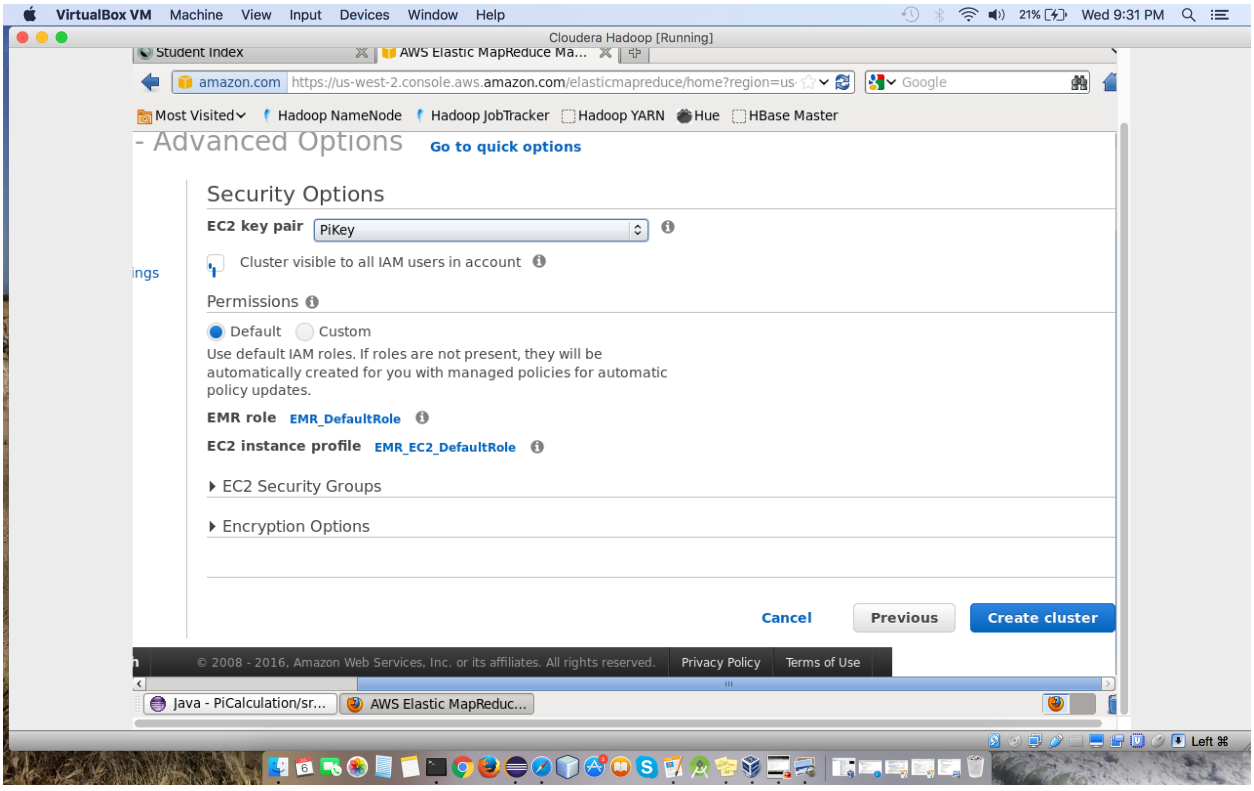


Provide the log folder path present in the S3 Service:

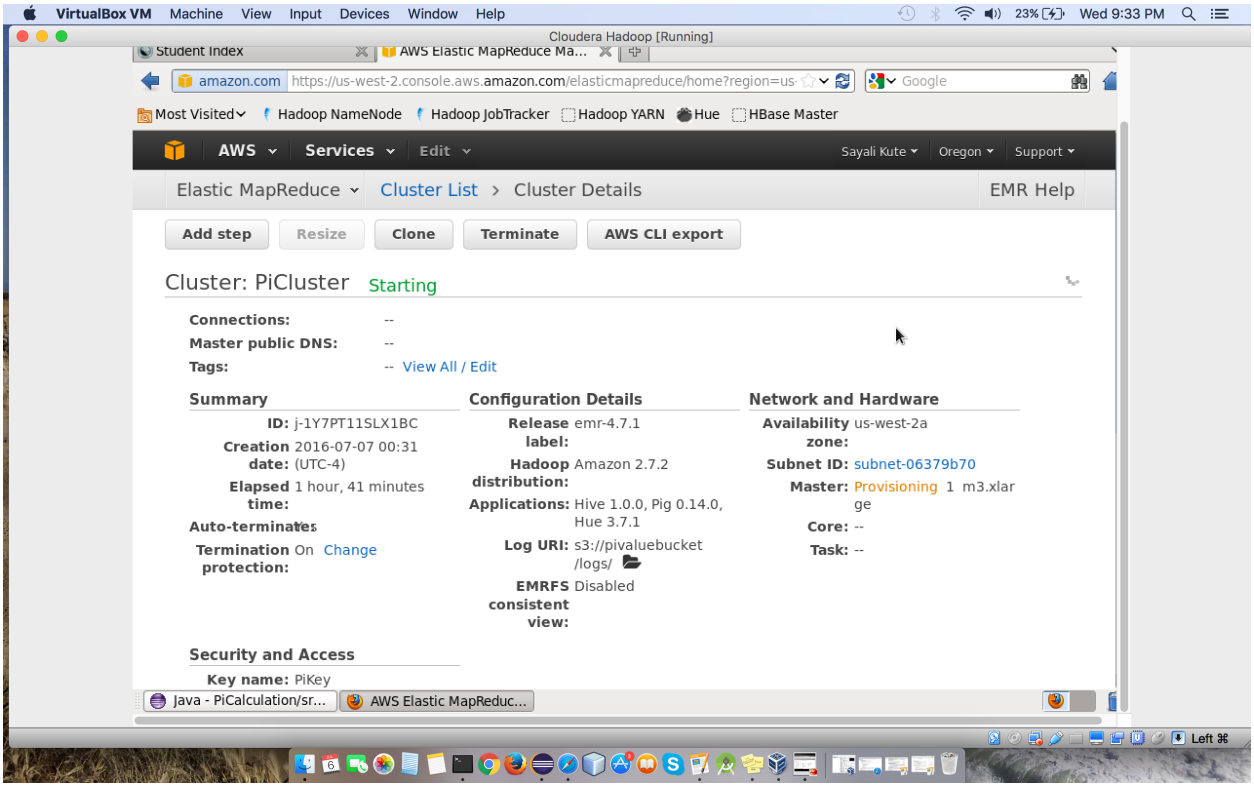


Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com

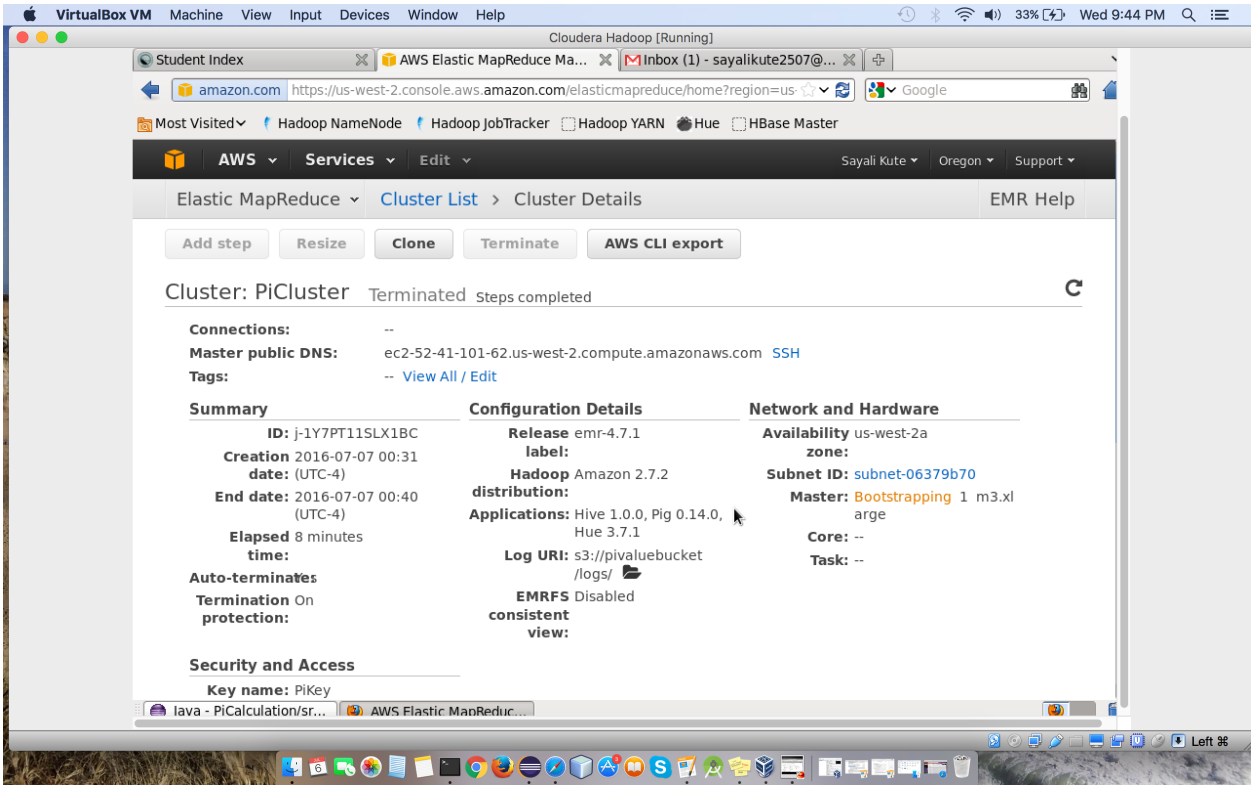
Select the appropriate EC2 key pair:



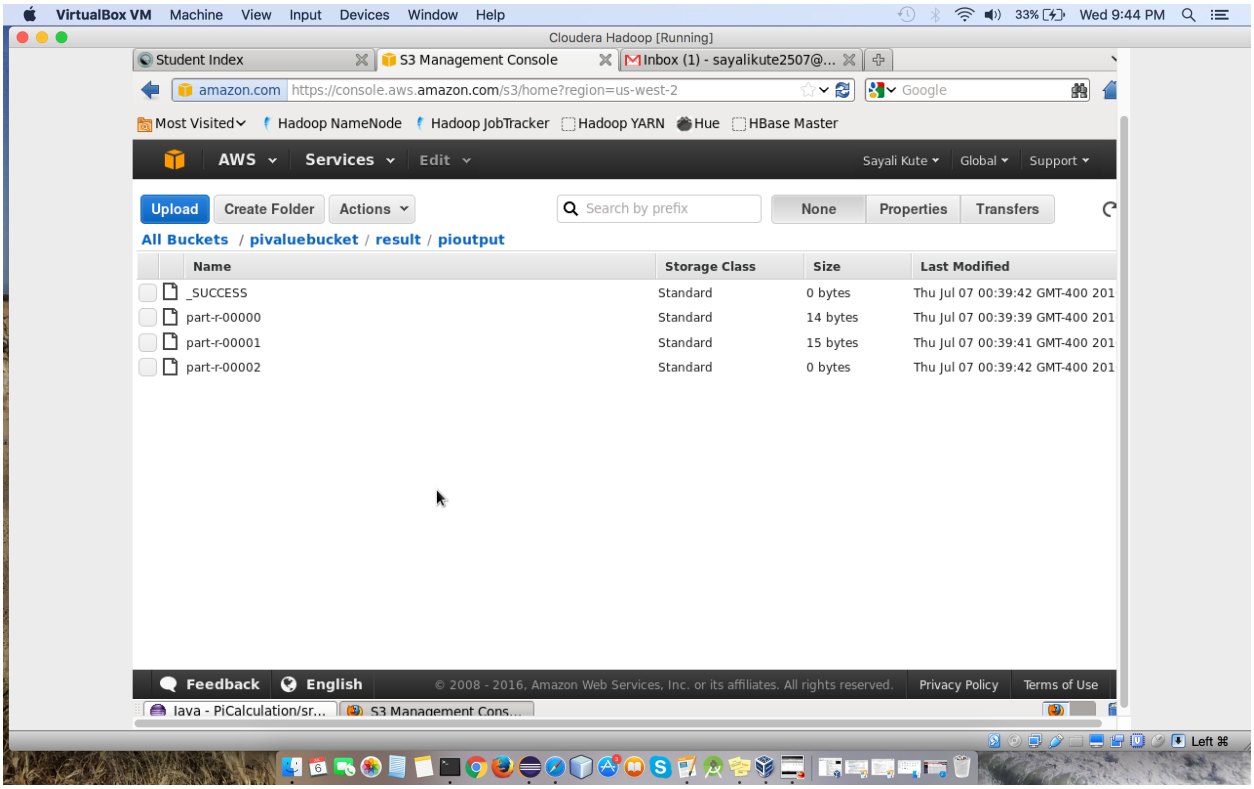
Cluster has started working and runs successfully :



Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com

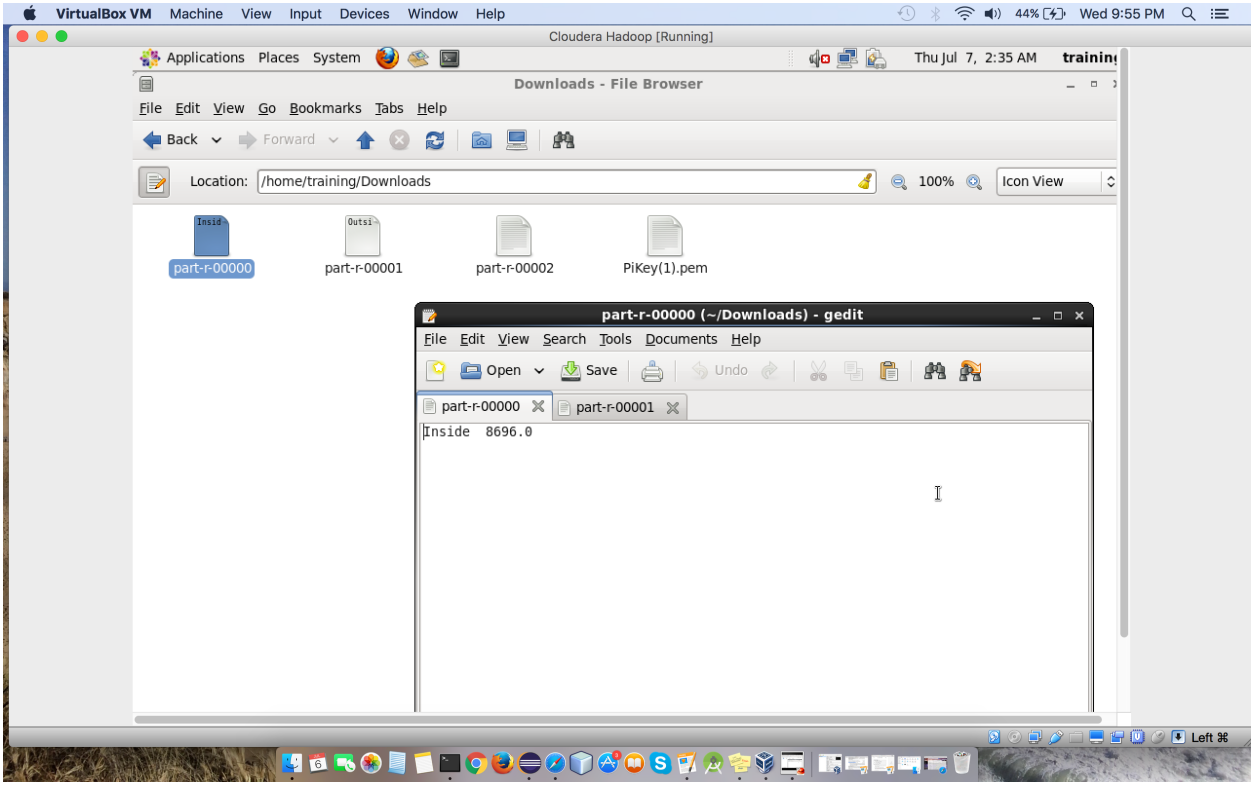


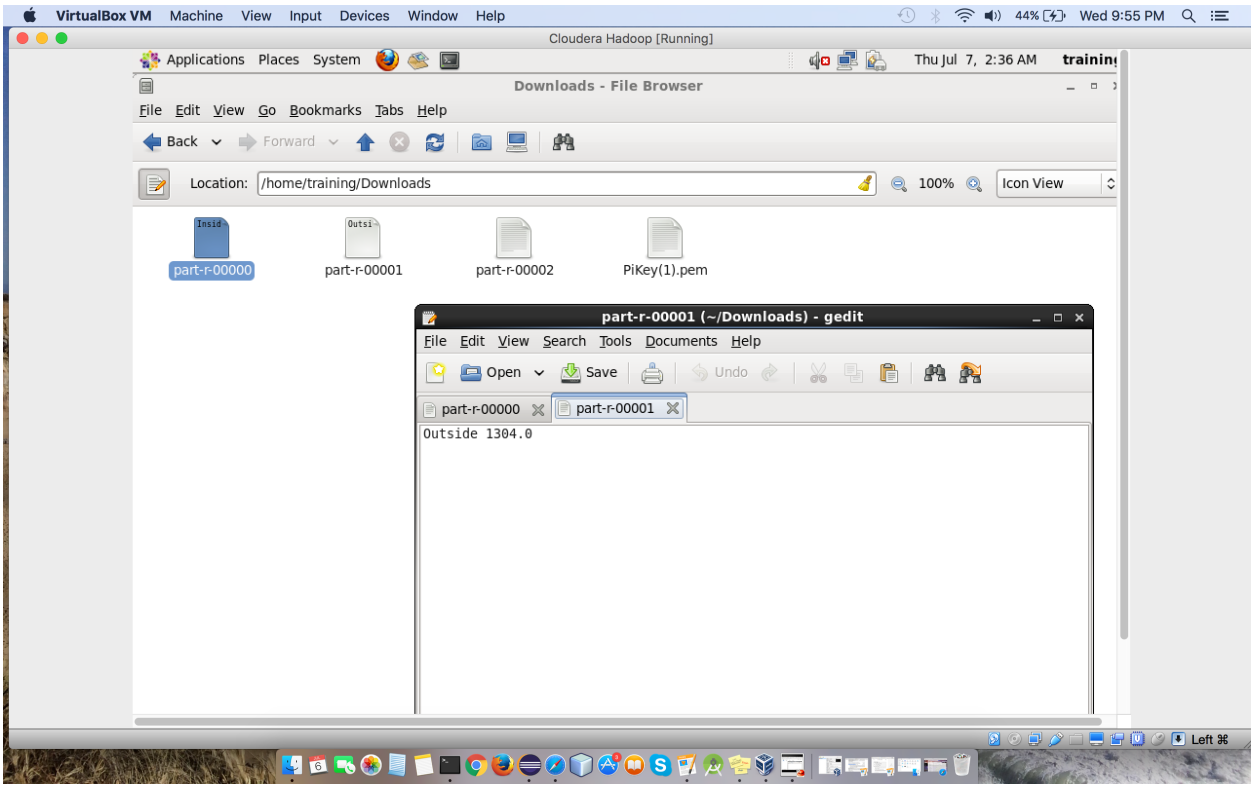
The output is generated in the 'results' folder of S3 Service:



Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com

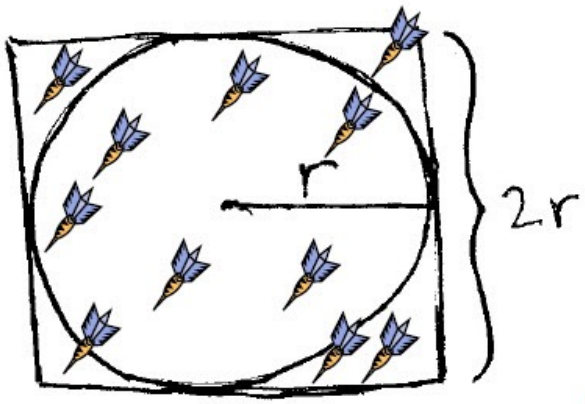
Calculation of PI value:





Name: Sayali Vishnu Kute Email ID: sayalikute2507@gmail.com

- Throw N darts on the board. Each dart lands at a random position (x,y) on the board.
  - Note if each dart landed inside the circle or not
    - Check if $x^2+y^2<r$
  - Take the total number of darts that landed in the circle as S



$$4\left(\frac{S}{N}\right)=\pi$$

Output File (pioutput): This file contains value listed as:

Inside : 8696.0
Outside : 1304.0

Formula to calculate pi : 4*(S/N) = $\boldsymbol{\pi}$
where, S = Number of darts hit inside the circle
N = Total number of darts thrown

Pi = 4*(Inside /(Inside + Outside))

Pi = 4* (8696.0 / (8696.0 + 1304.0))

Pi = 4*(8696.0/10000)

Pi= 4*0.8696

Pi= 3.4784 (approximately)

——END——