

Big Data Analytics Lab

PMDS507P

Name: **Tufan Kundu** Registration number: **24MDT0184**

Slot: L29+L30

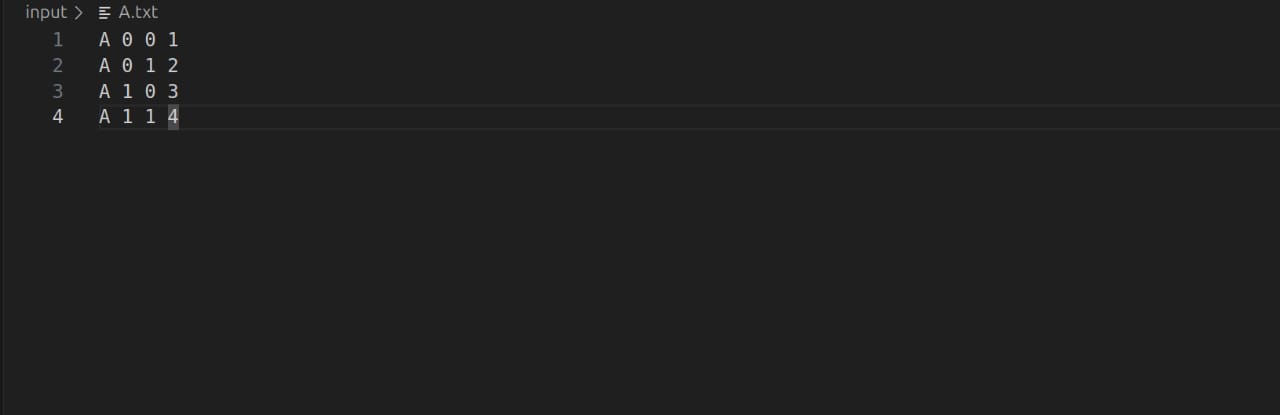
Digital Assignment 5

# 

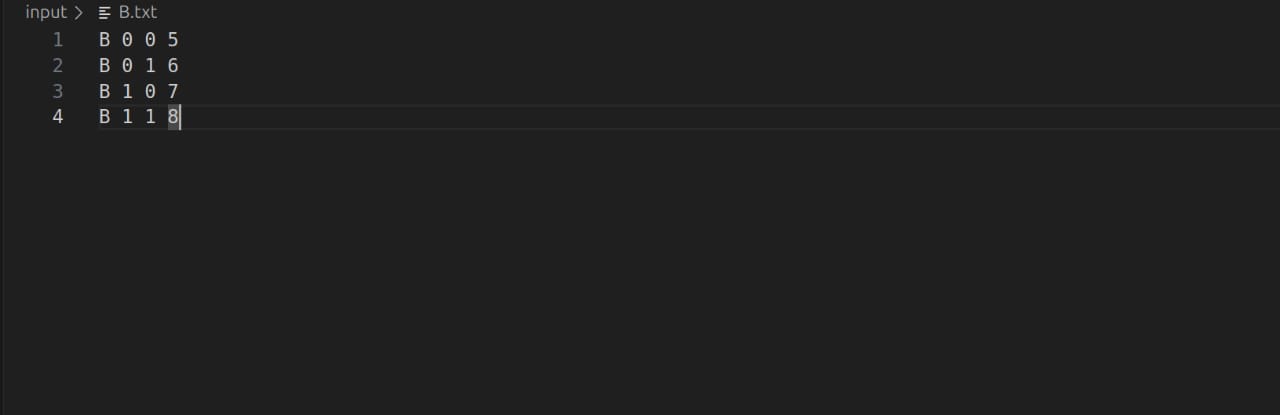
## 

## Explain the steps involved in performing a Matrix multiplication in Hadoop

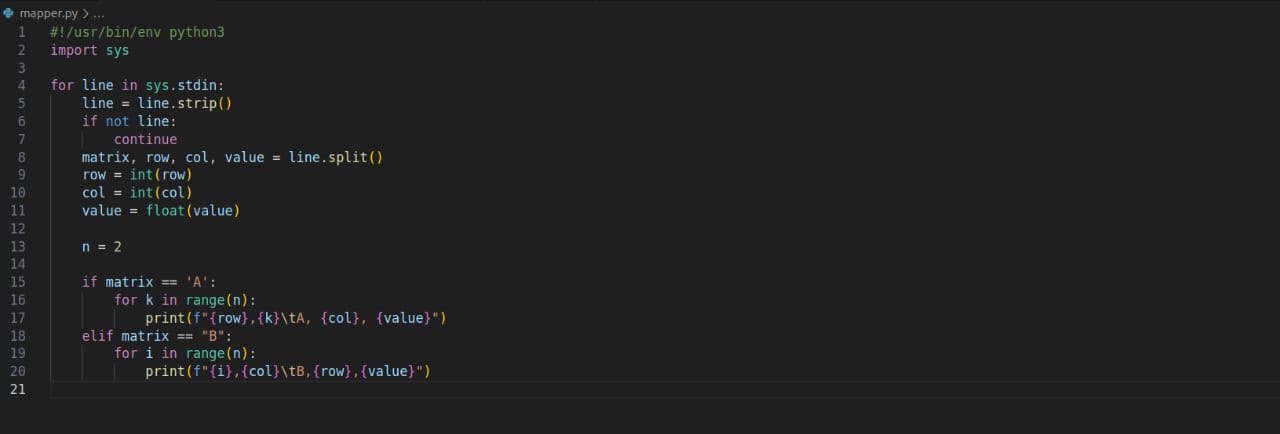
**Step 1:** Preparing the A matrix in A.txt



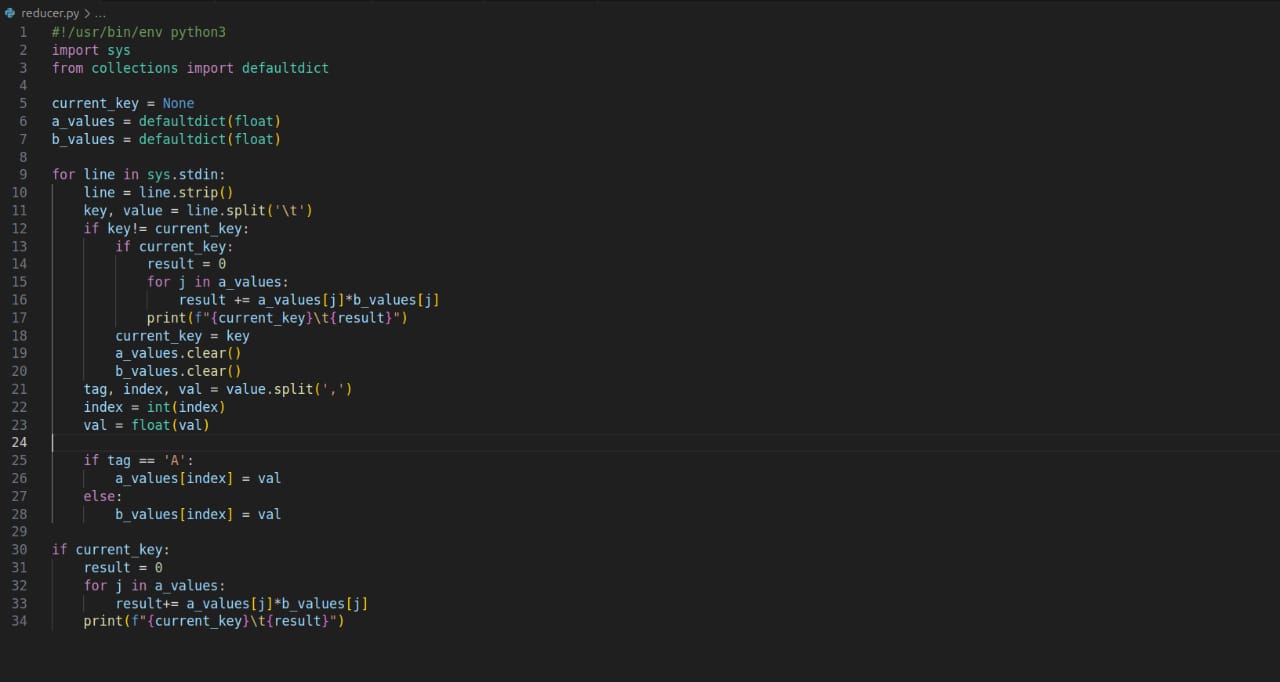
**Step 2:** Preparing the B matrix B.txt



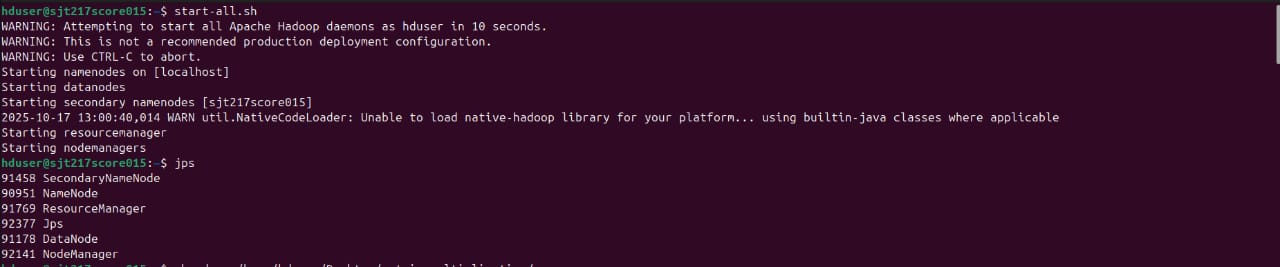
Step 3: Mapper[.py](http://reducer.py)



Step 4: [Reducer.py](http://reducer.py)



Step 5: Start HDFS and YARN and verify with jps



Step 6: Making Scripts executable

***chmod +x /home/hduser/Desktop/operationhadoop/***[***mapper.py***](http://mapper.py)

***chmod +x /home/hduser/Desktop/operationhadoop/reducer.py***



Step 7: Creating directory

***hdfs dfs -mkdir /matrix\_input***

***hdfs dfs -put /home/hduser/Desktop/matrix\_multiplication/input/\* /matrix\_input/***



Step 8: Running the mapper and reducer

***hadoop jar /home/hduser/hadoop/share/hadoop/tools/lib/hadoop-streaming-3.3.1.jar -input /matrix\_input -output /matrix\_output -mapper /home/hduser/Desktop/matrix\_multiplication/***[***mapper.py***](http://mapper.py) ***-reducer /home/hduser/Desktop/matrix\_multiplication/***[***reducer.py***](http://reducer.py) ***-file /home/hduser/Desktop/matrix\_multiplication/***[***mapper.py***](http://mapper.py) ***-file /home/hduser/Desktop/matrix\_multiplication/reducer.py***



Step 9: Displaying the result

***hdfs dfs -cat /matrix\_output/part-00000***

