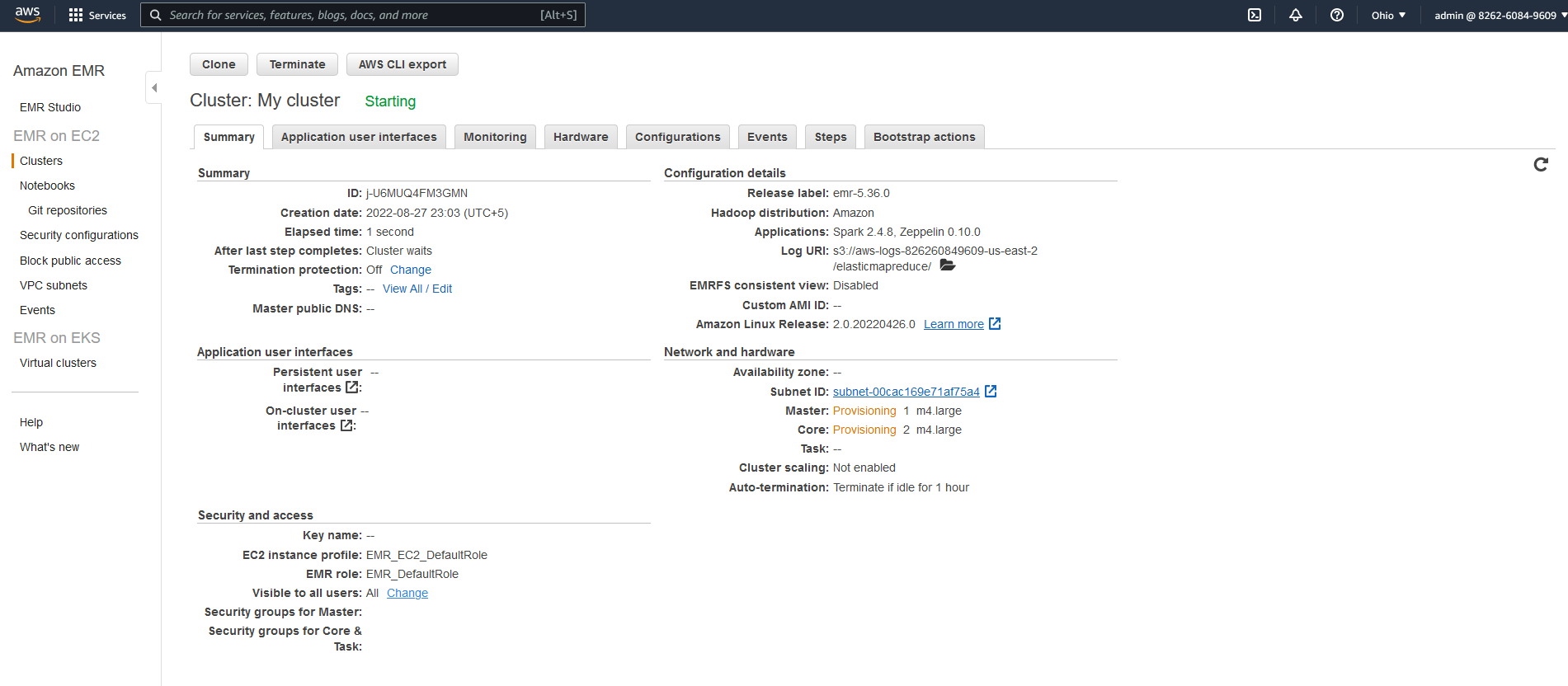
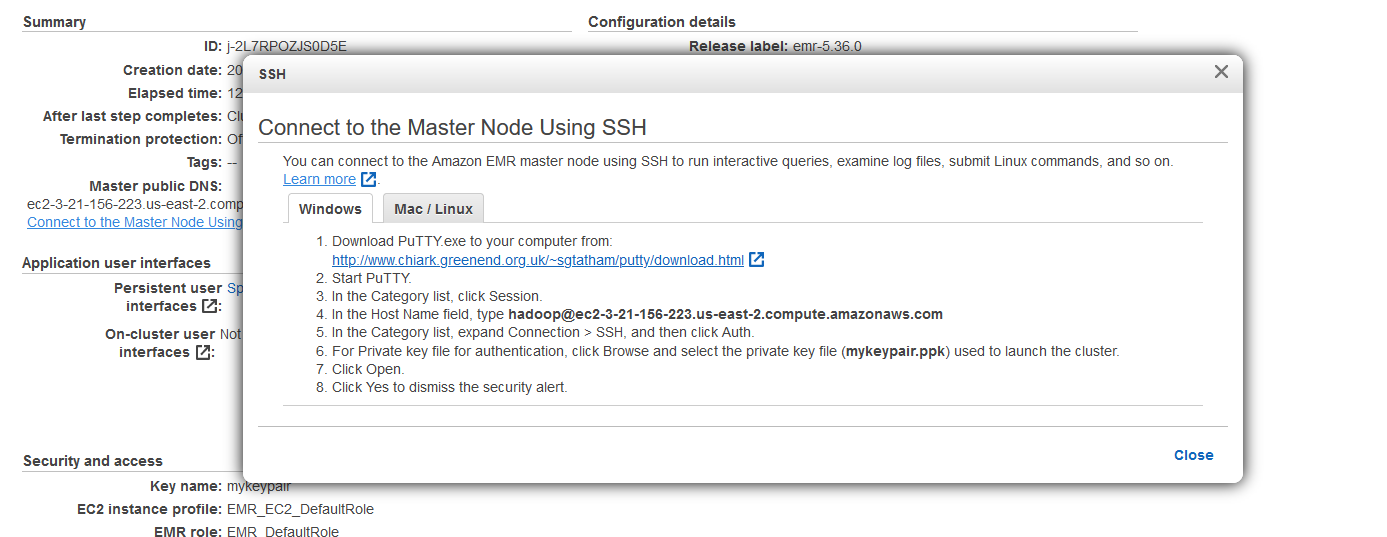
Step 1: Create the EMR Cluster

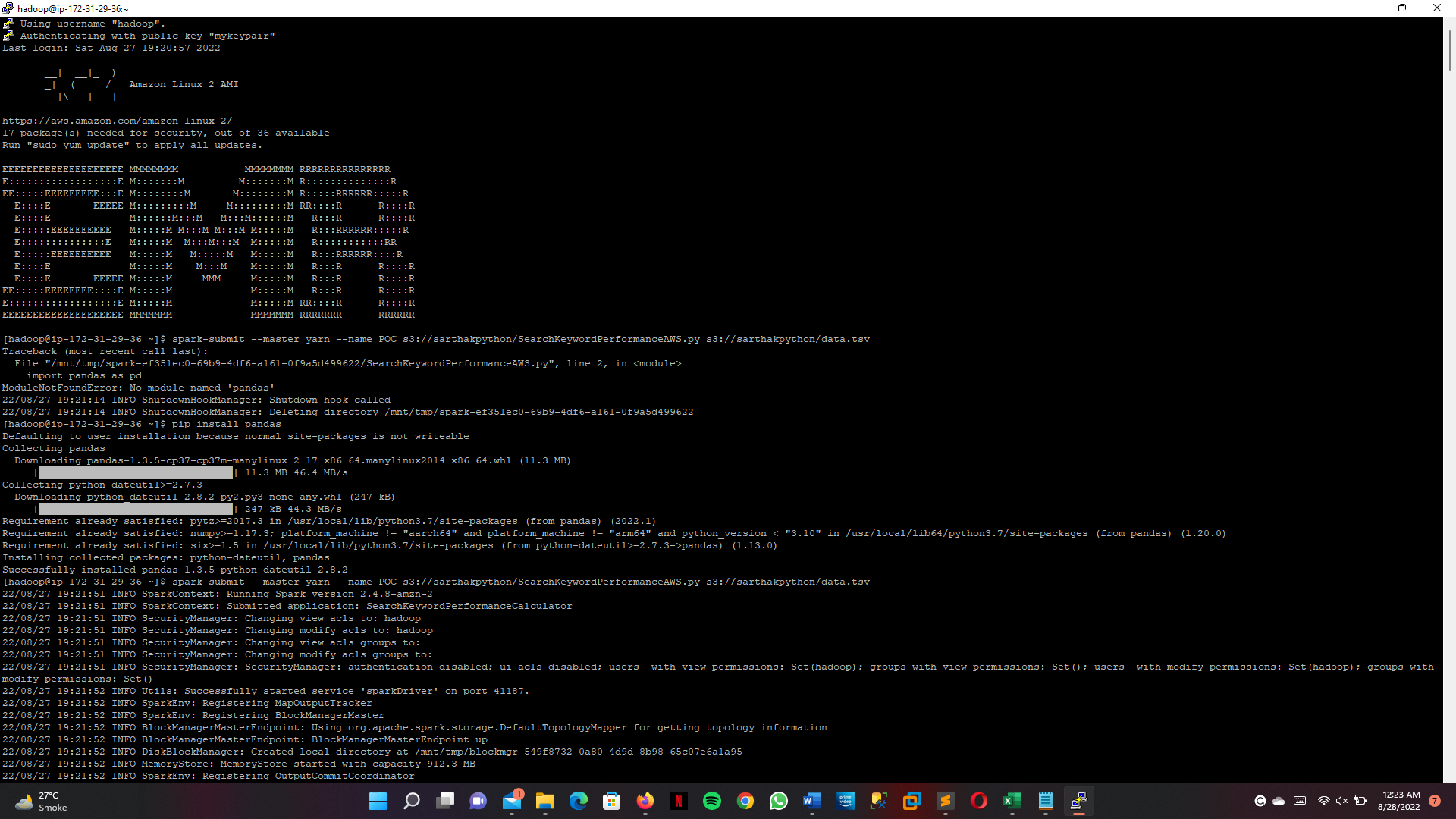


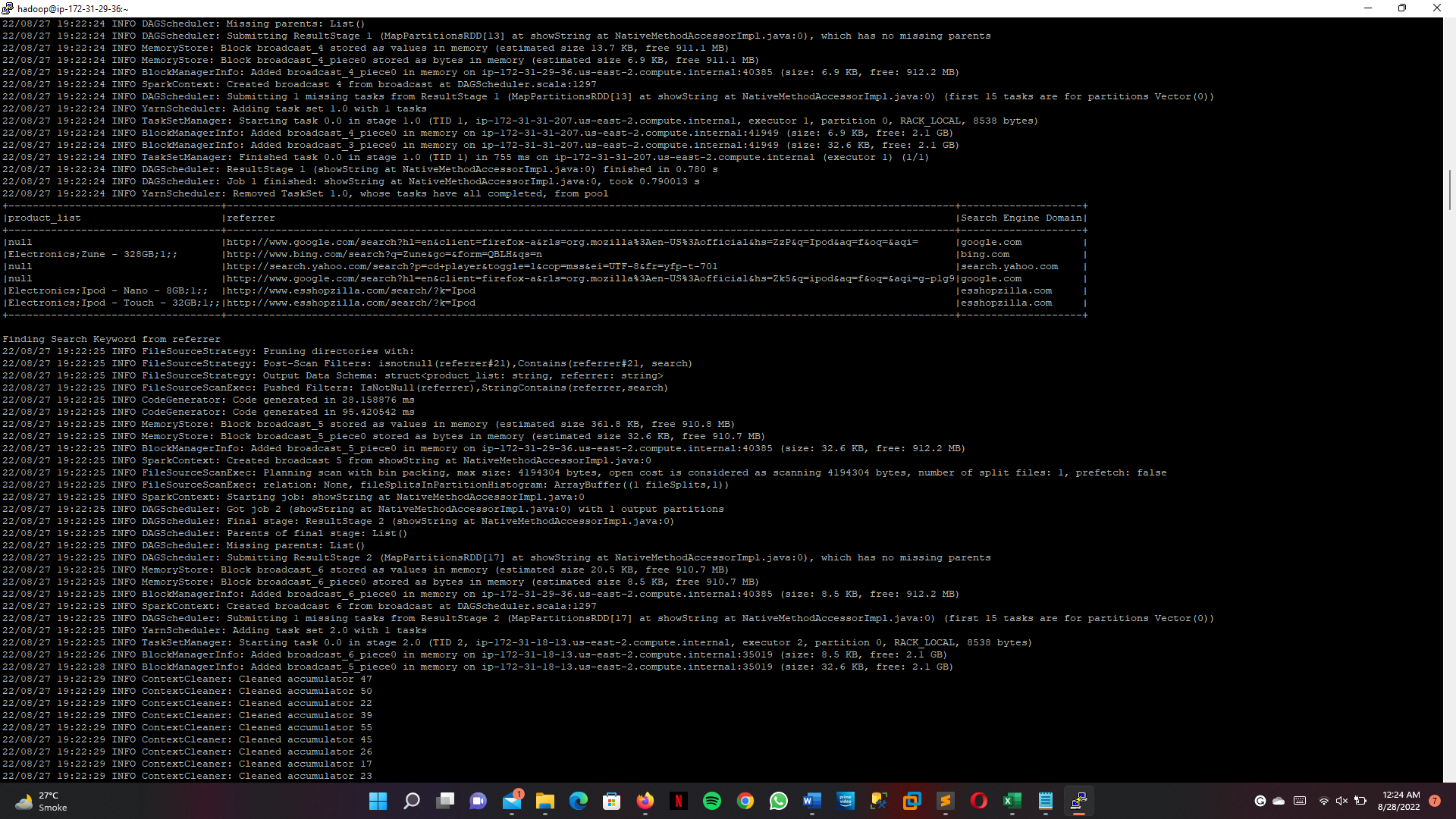
Step2: Login to the master node of the cluster

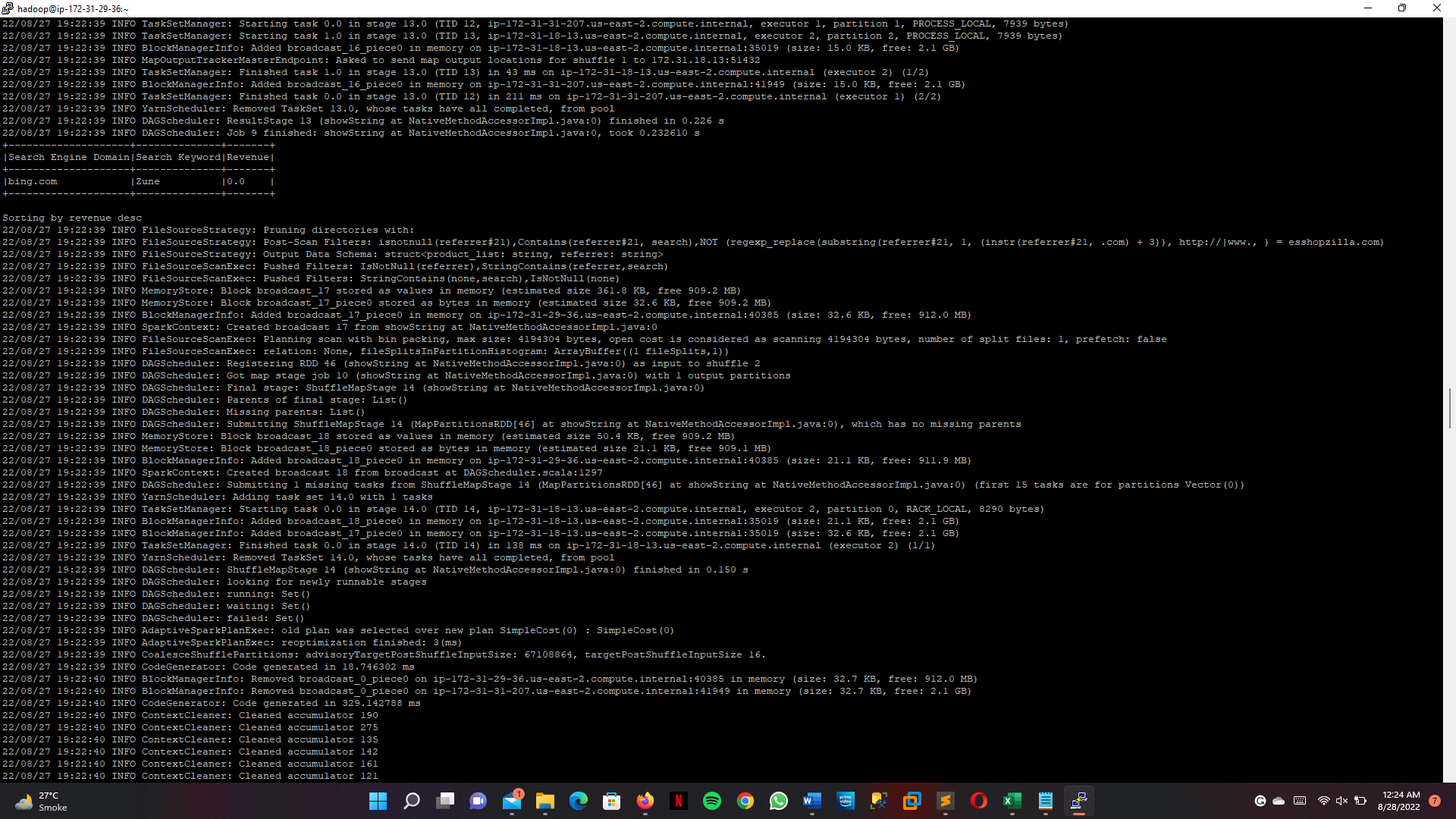


Step3: Run the spark application

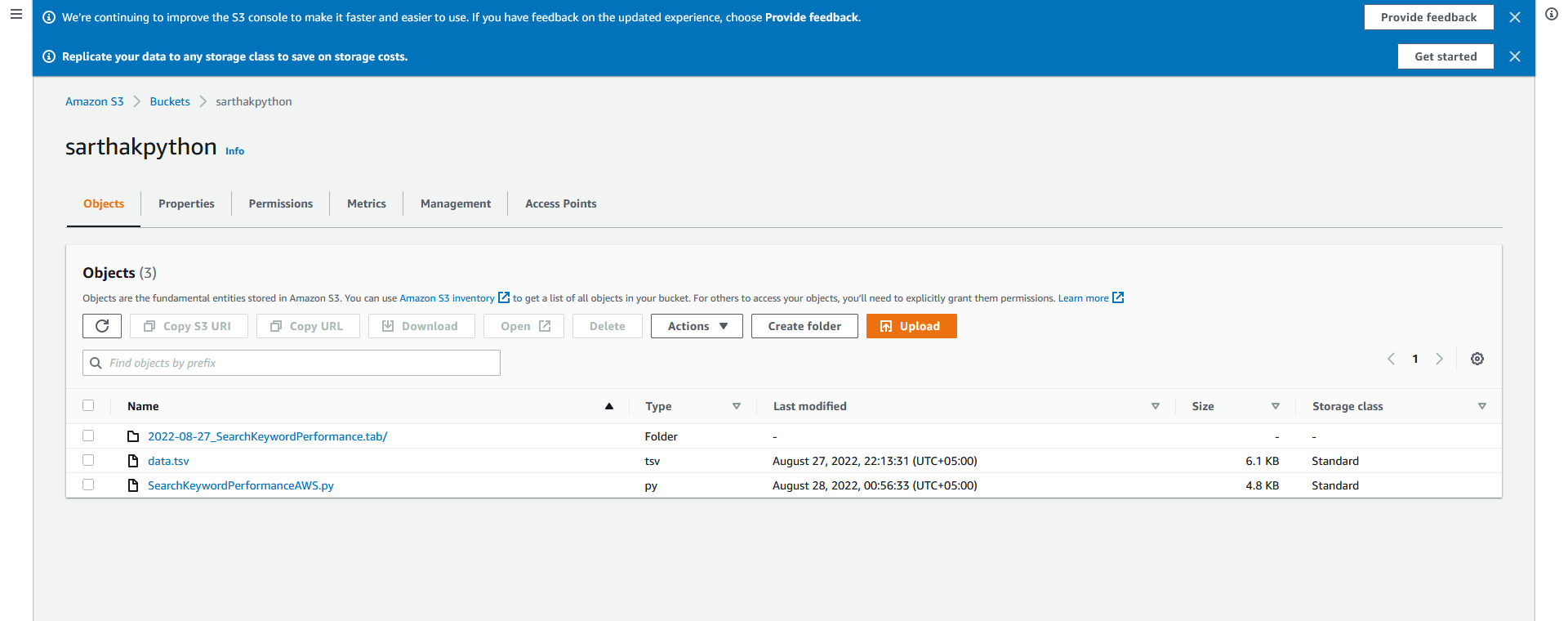
spark-submit --master yarn --name POC s3://sarthakpython/SearchKeywordPerformanceAWS.py s3://sarthakpython/data.tsv







Step 4 : Verify the result created on S3.



Step 5: Terminate the cluster

