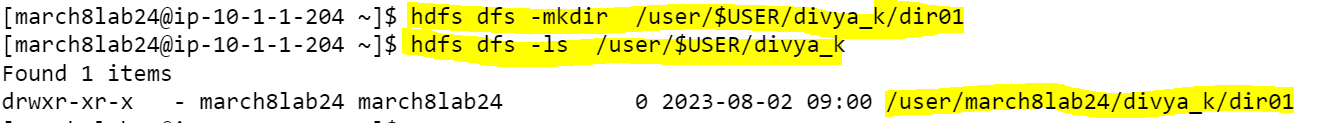
**Hadoop Commands**

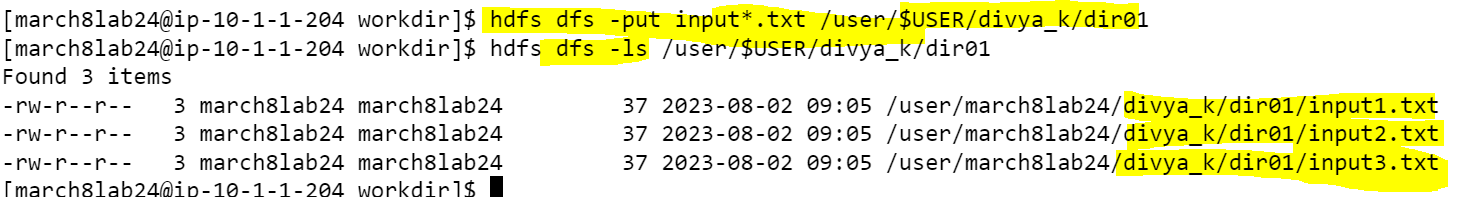
**Problem Statement 01** : **Working with HDFS Commands Download Dataset**

1. Create a folder in HDFS by name “dir01” and move input1.txt , input2.txt and input3.txt into /dir01.

Ans : hdfs dfs -mkdir /user/$USER/divya\_k/dir01



hdfs dfs -put input\*.txt /user/$USER/divya\_k/dir01



1. List only the file names present in “/dir01”

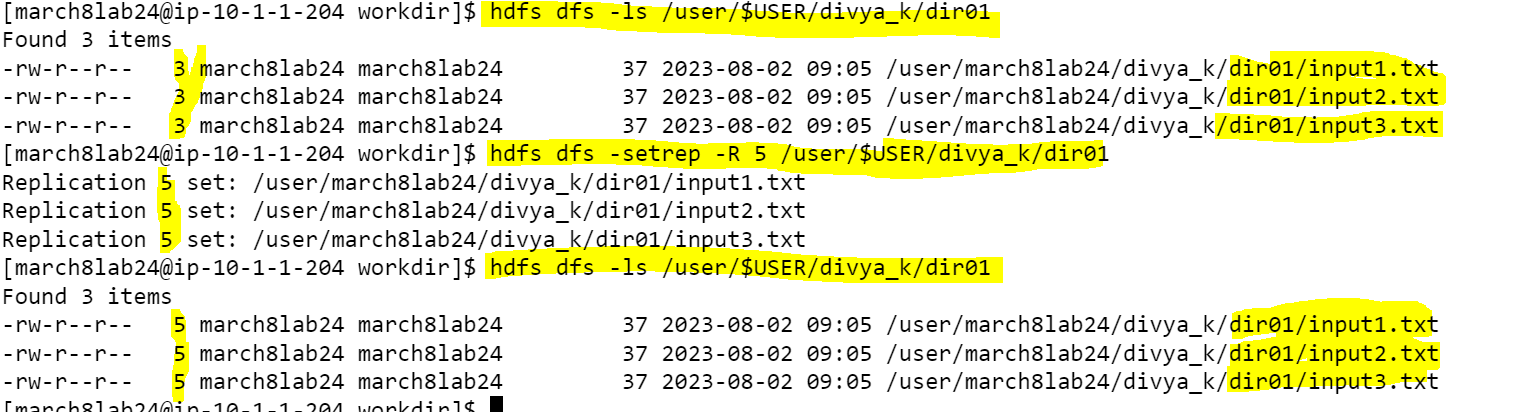
Ans: hdfs dfs -ls /user/$USER/divya\_k/dir01

A close up of numbers

Description automatically generated

3.Change the replication factor for the content present in directory “/dir01” to 5 and display the replication factor for the files present in “/dir01”.

Ans: hdfs dfs -setrep -R 5 /user/$USER/divya\_k/dir01



4.Create a folder in HDFS by name “scenario01” and create directory “level01” inside “scenario01” directory and create another directory “level02” inside directory “level01”. Once the required directories are created copy input1.txt to scenario01, input2.txt to level01 and input3.txt to level02 and finally recursively print only the file names present in directory scenario01.

Ans : hdfs dfs -mkdir /user/$USER/divya\_k/scenario01

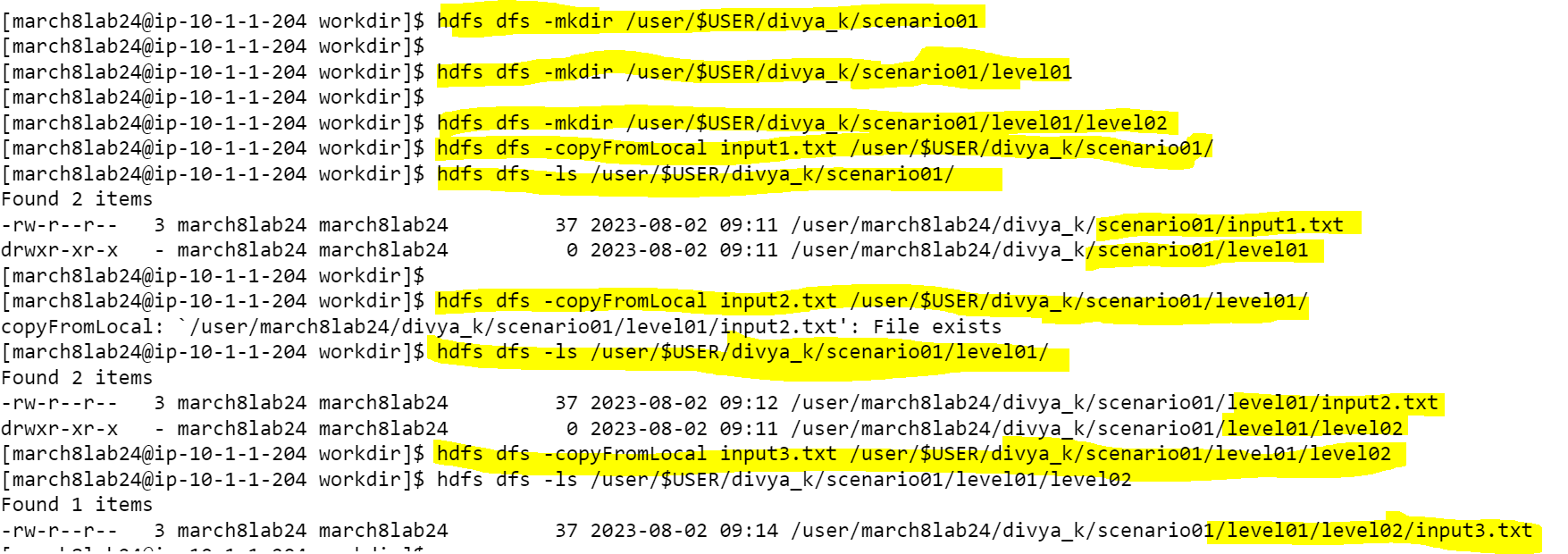
hdfs dfs -mkdir /user/$USER/divya\_k/scenario01/level01

hdfs dfs -mkdir /user/$USER/divya\_k/scenario01/level01/level02

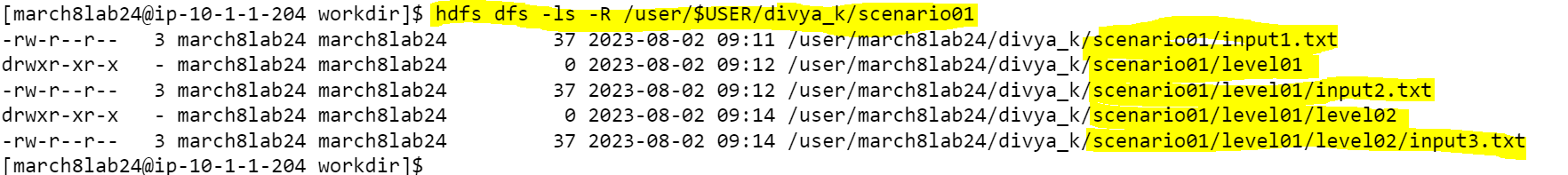
hdfs dfs -copyFromLocal input1.txt /user/$USER/divya\_k/scenario01/

hdfs dfs -copyFromLocal input2.txt /user/$USER/divya\_k/scenario01/level01/

hdfs dfs -copyFromLocal input3.txt /user/$USER/divya\_k/scenario01/level01/level02



hdfs dfs -ls -R /user/$USER/divya\_k/scenario01



**Problem Statement 02** : **Working with YARN Commands Commands**

1. Run MapReduce Program and capture the application Id of the job

Ans: cd /opt/cloudera/parcels/CDH/jars

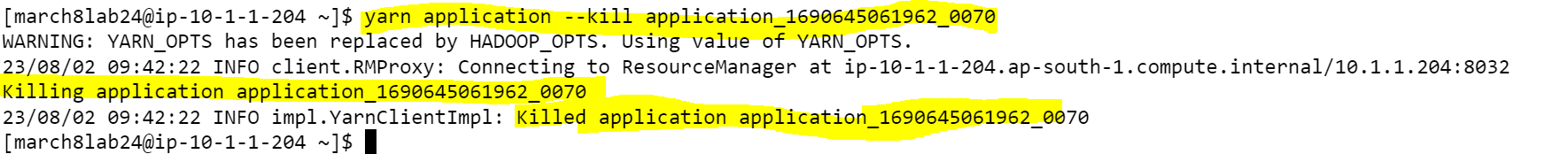
yarn jar hadoop-mapreduce-examples-3.0.0-cdh6.2.1.jar wordcount /user/$USER/divya\_k/ /user/$USER/divya\_k/output/

A screenshot of a computer code

Description automatically generated

1. Re-run the MapReduce program and kill the application using the yarn command.

Yarn application –kill application\_1690645061962\_0070



1. List all the applications which are RUNNING state

yarn application --list --appStates RUNNING

A close up of a text

Description automatically generated

4. View the logs of any of the jobs which are already completed

Ans: yarn logs --applicationId application\_1690645061962\_0071

A close-up of a document

Description automatically generated

Note: Due to more logs available for above application id so I pasted last some lines from snippet.