HDFS Task 1

Answer .

What is the Namenode's URI and which file is it configured in?

Ans: Name node is the master and management node of Hadoop system.in core-site.xml file

2. Where on a local file system will Namenode store its image and which file is it configured in?

Ans: FSImage is where Namenode stores its Imagae and it is stored in XML.config

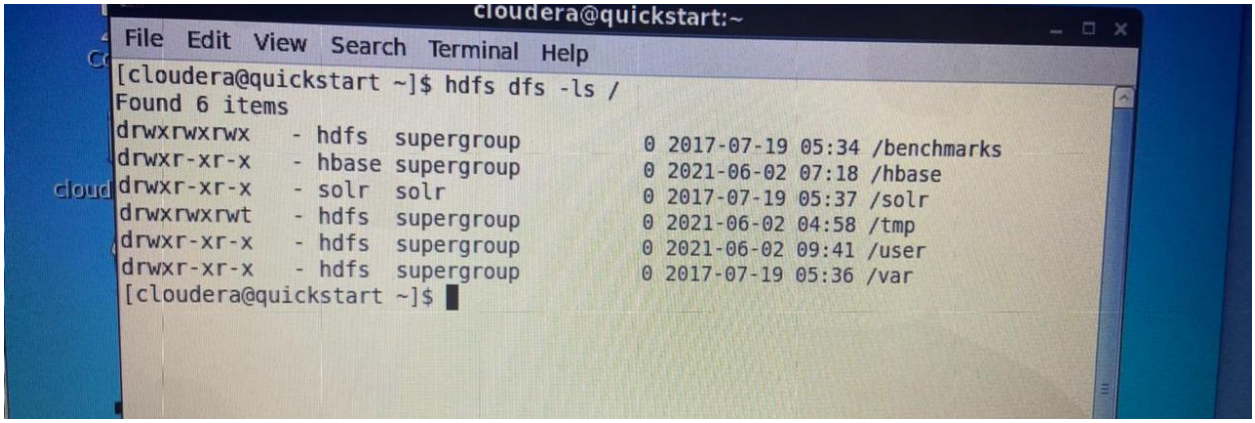
3. Where on a local file system will Datanode store its blocks and which file is it configured in?

Ans:In Hadoop home etc folder the hdfs-site.xml file is there in that file data node property is configured
4. What is the block replication and which file is it configured in?

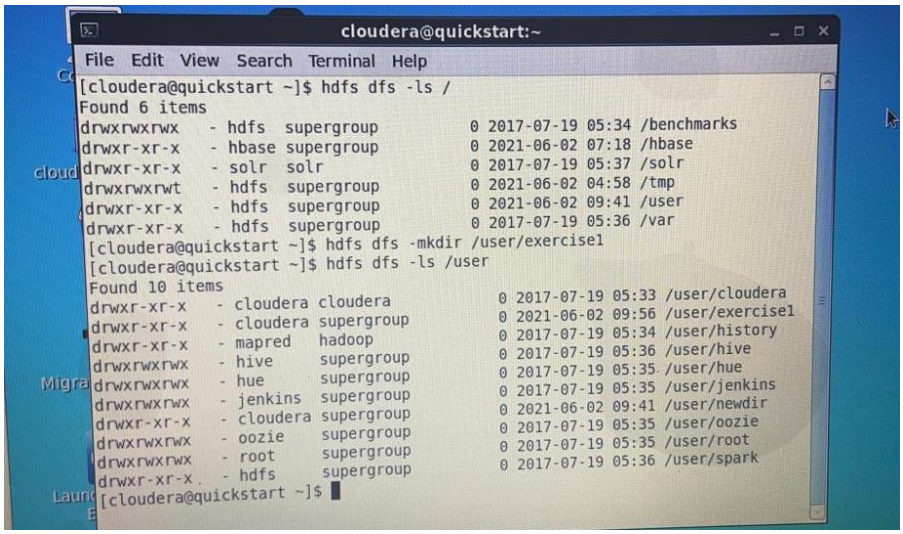
Ans: Its a copy of the data of node of rack, each data should be placed in different nodes of different rack.It is configured in hdfs-site.xml

Perform:

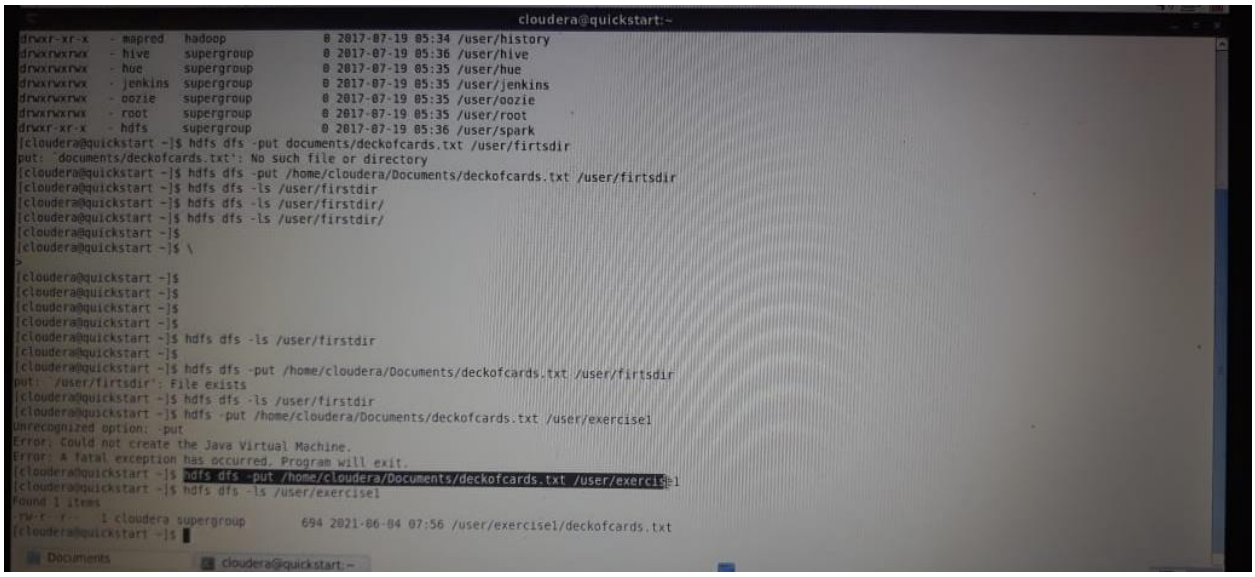
1. Start HDFS and verify that it's running ?



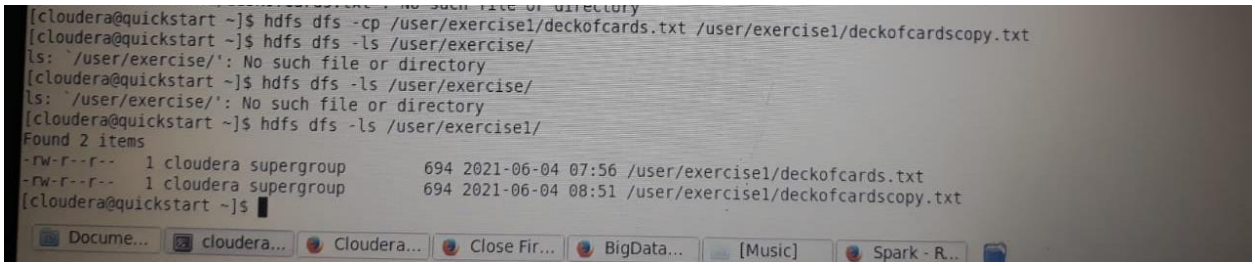
2. Create a new directory /exercise1 on HDFS ?



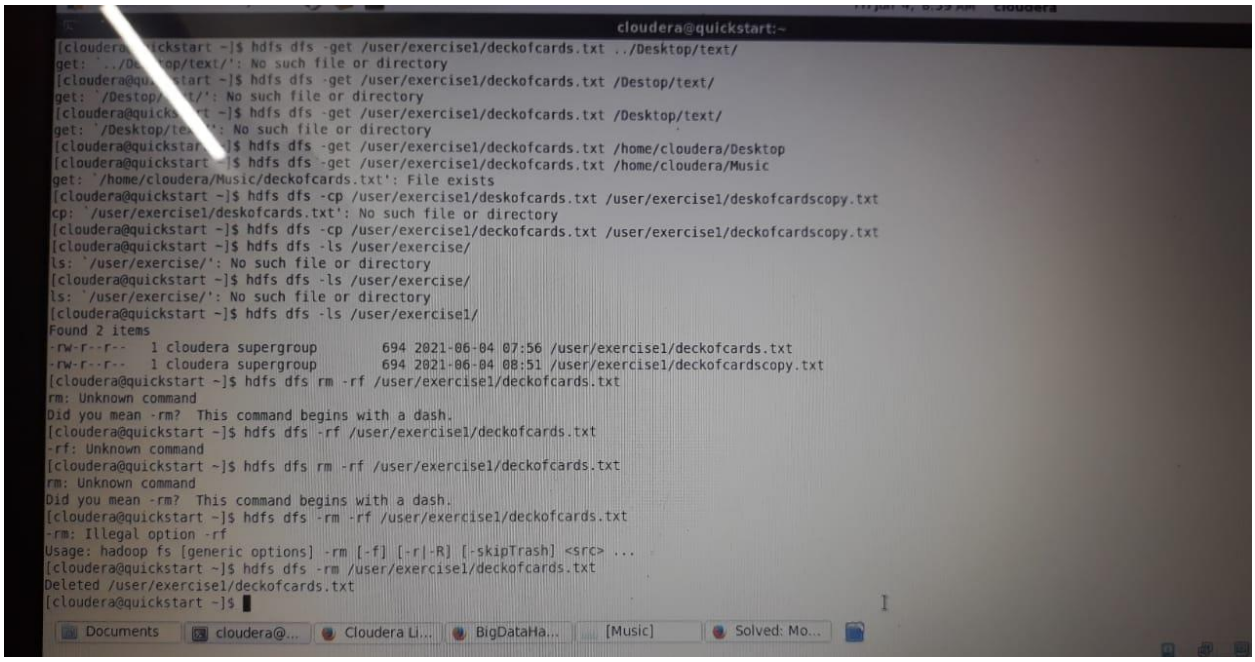
3. Upload GitHub repo sample_data/deckofcards.txt to HDFS under /exercise1 directory



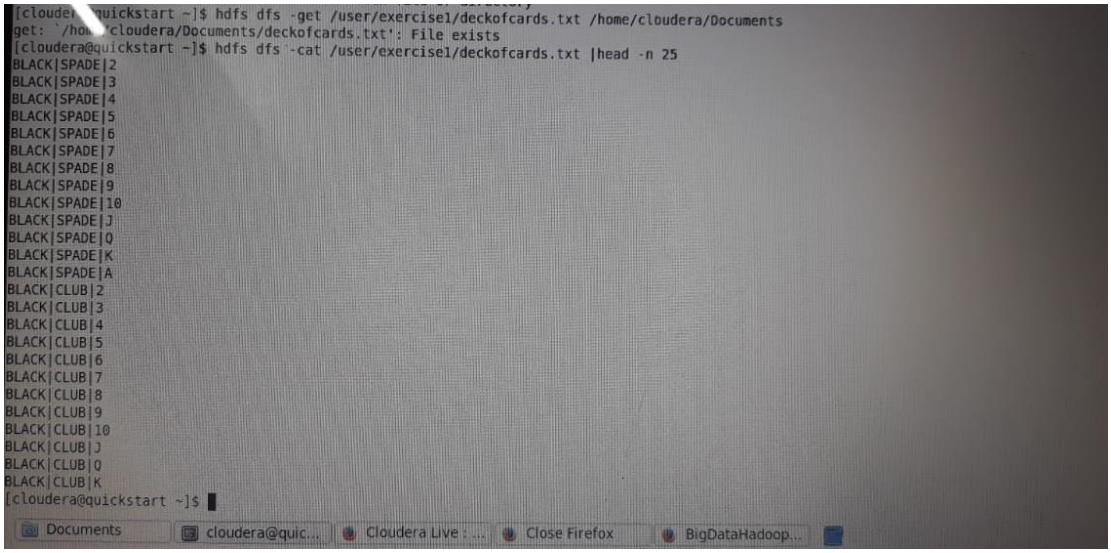
4. View the content of the /exercise1 directory



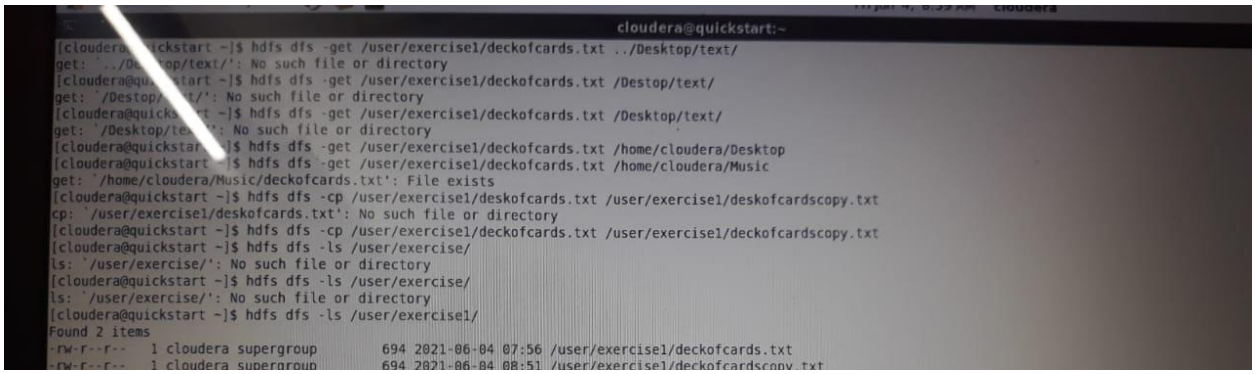
5. Determine the size of the hamlet.txt file in KB that resides on HDFS (not local directory)



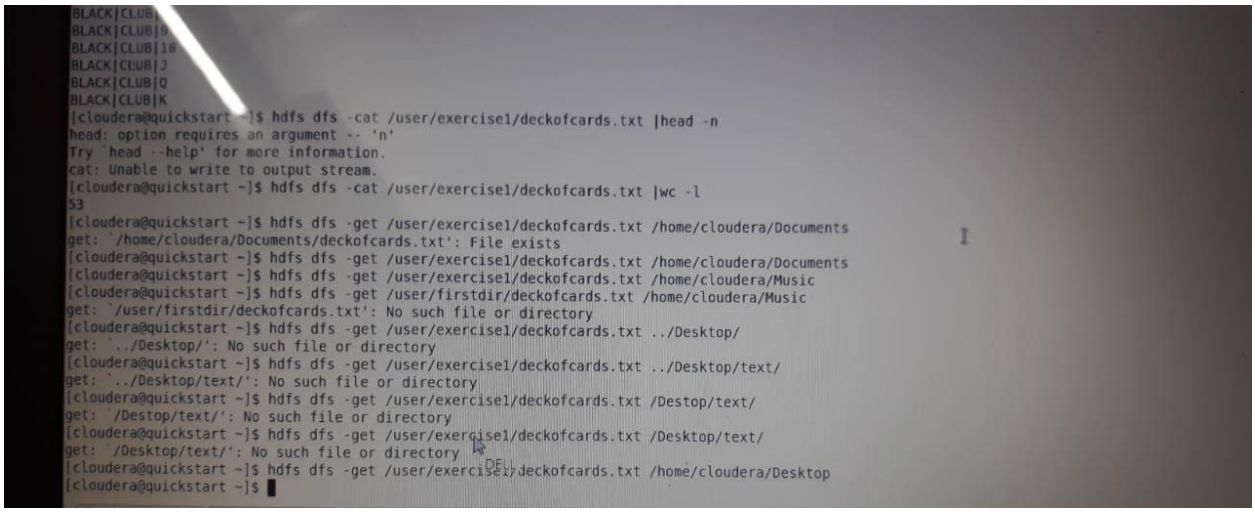
6. Print the first 25 lines to the screen from deckofcards.txt on HDFS



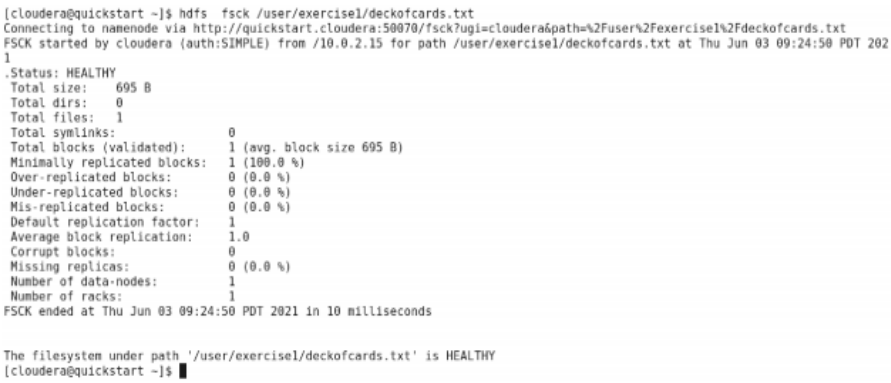
7. Copy deckofcards.txt to deckofcardsCopy.txt



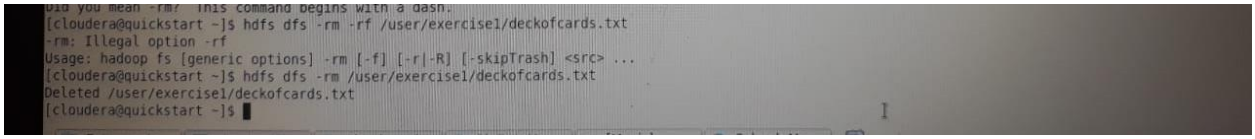
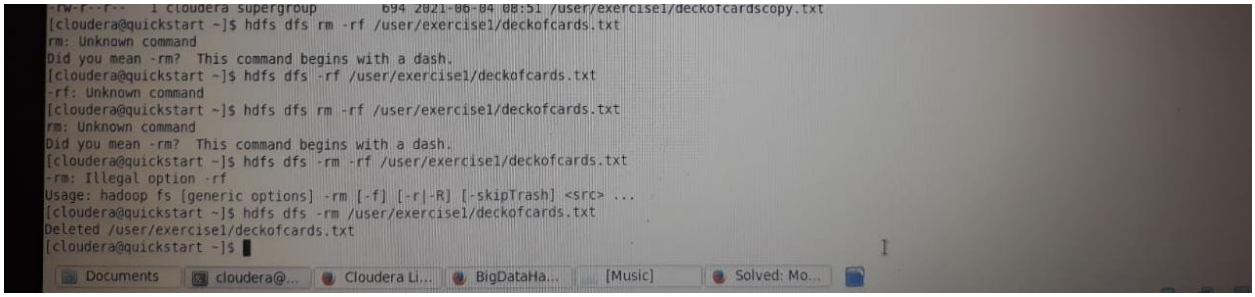
8. Copy deckofcards.txt back to local file system and name it deckofcards.copy.txt



9. Check the entire filesystem for inconsistencies/problems



10. Delete deckofcards.txt from HDFS



11. Delete the /exercise1 directory from HDFS