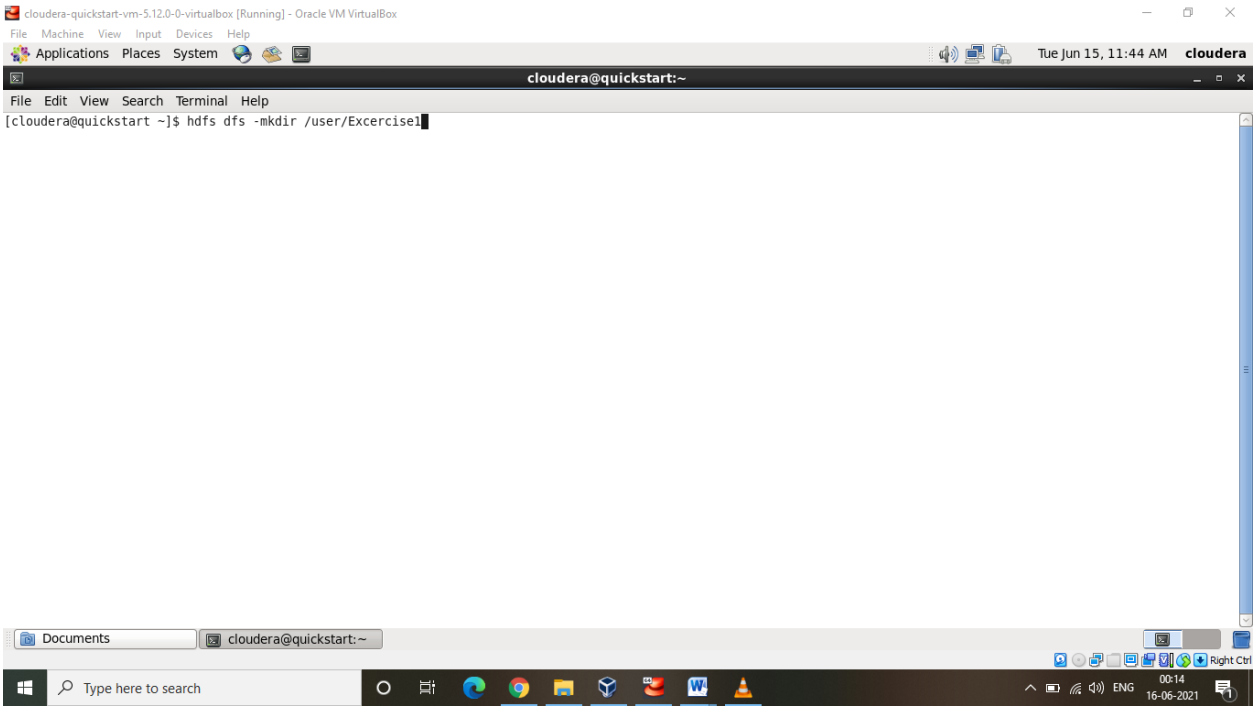
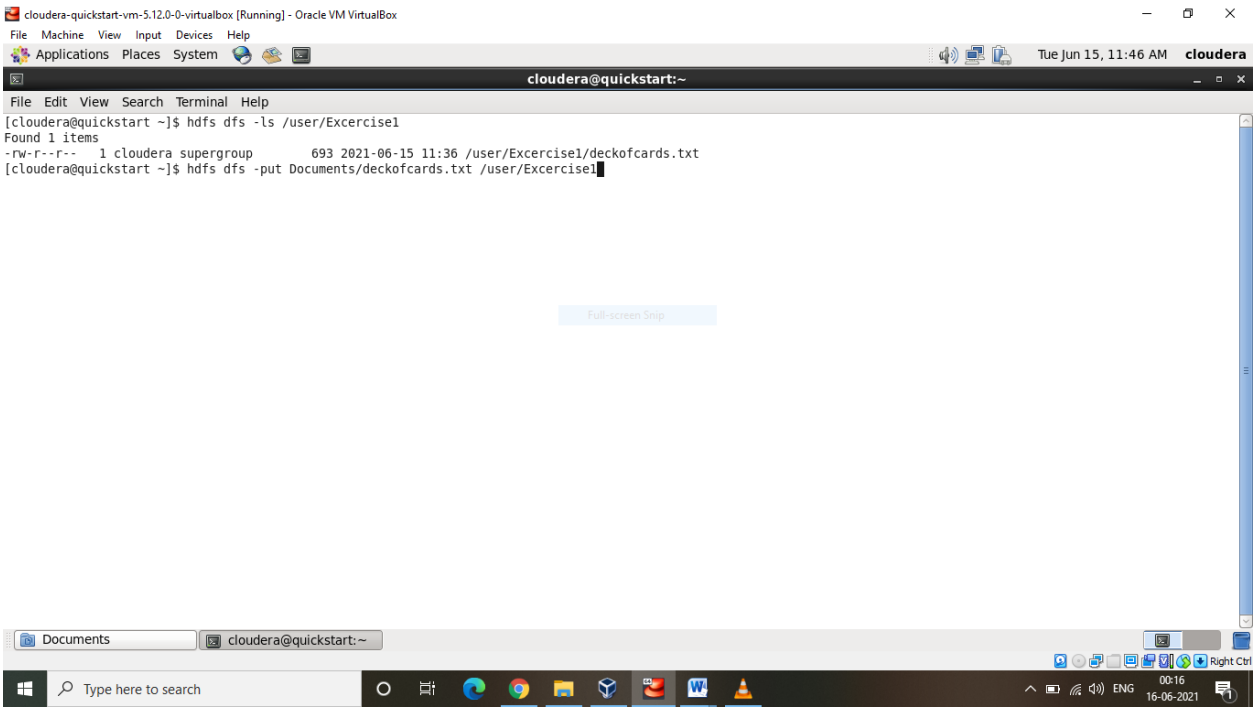
HDFS Task1

Perform

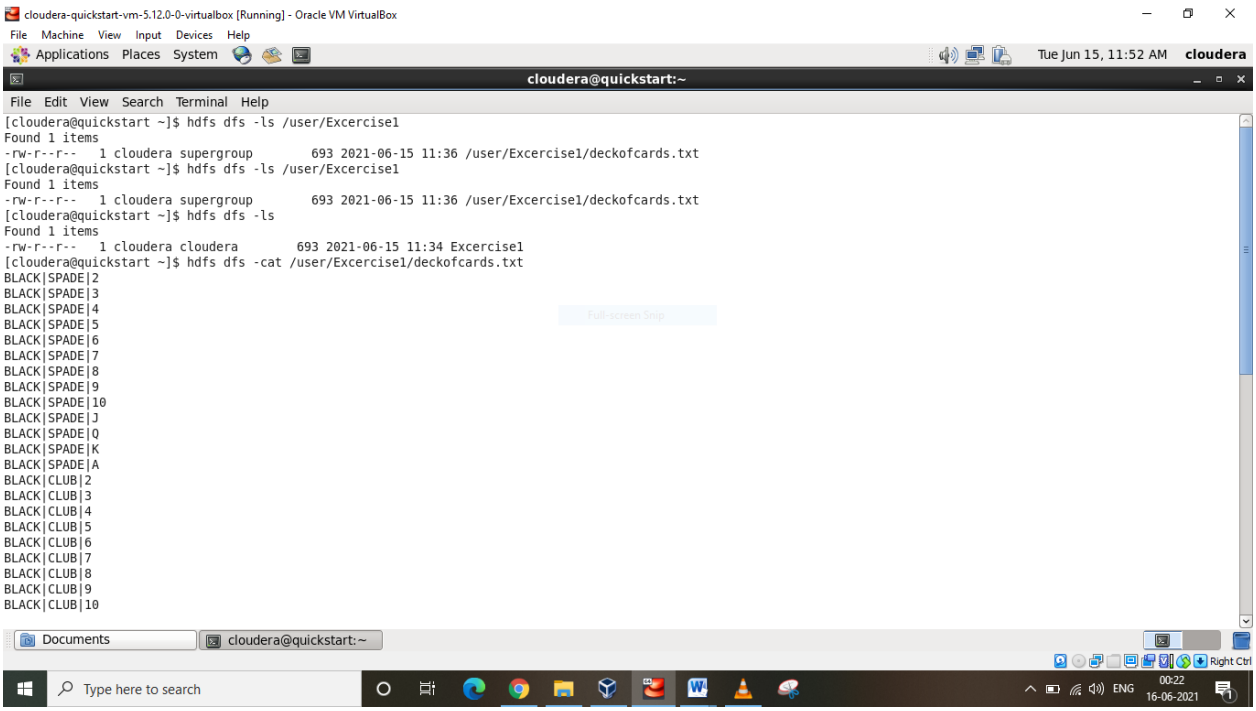
1. Start HDFS and verify that it's running
2. Create a new directory /exercise1 on HDFS



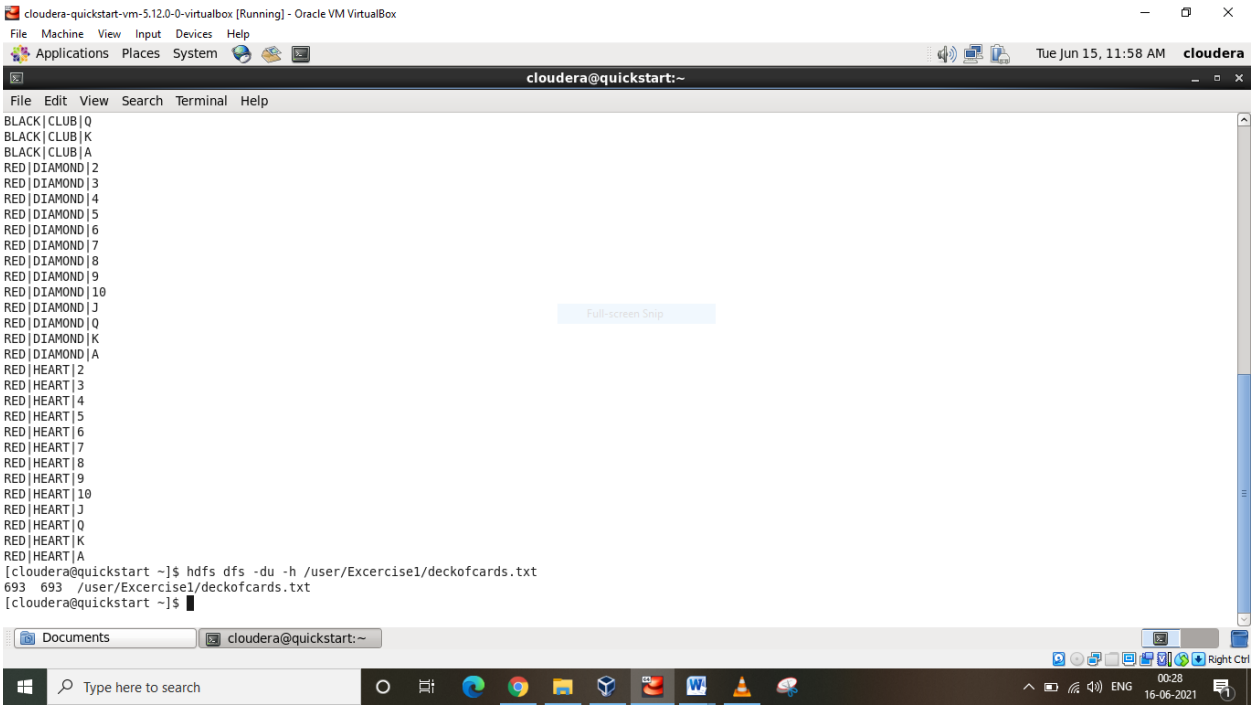
3. Upload GitHub repo sample_data/deckofcards.txt to HDFS under /exercise1 directory



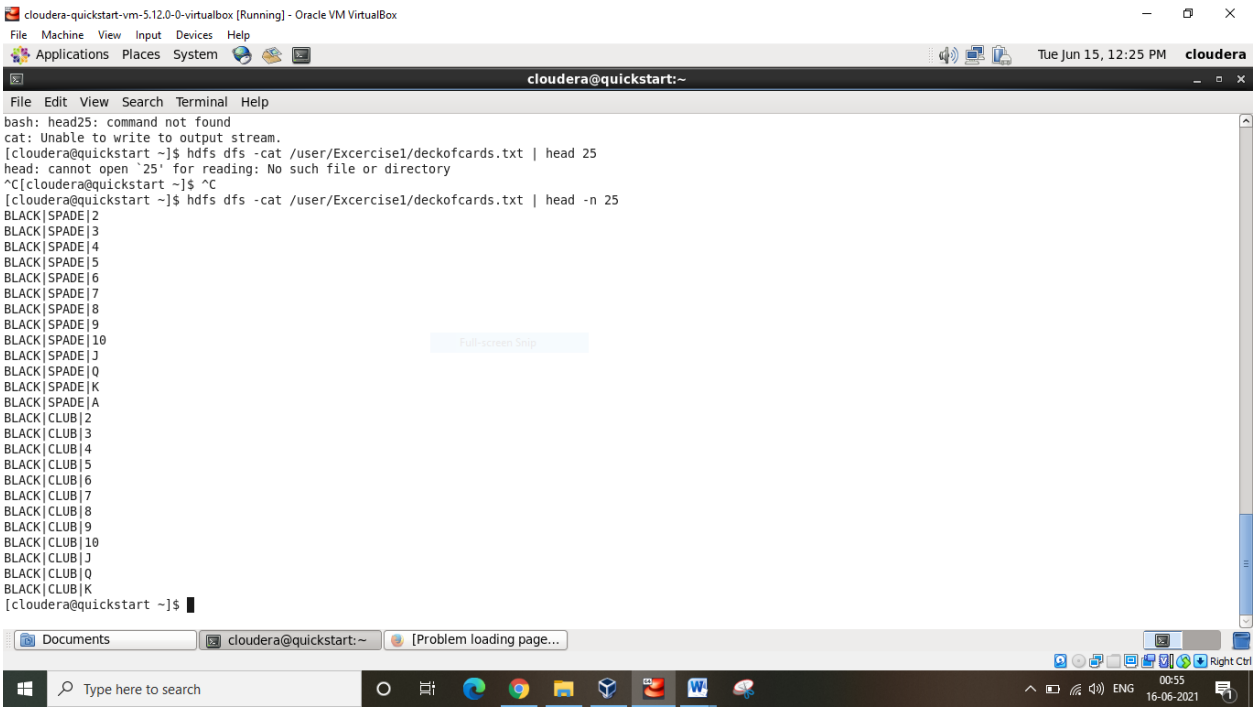
4. View the content of the /exercise1 directory



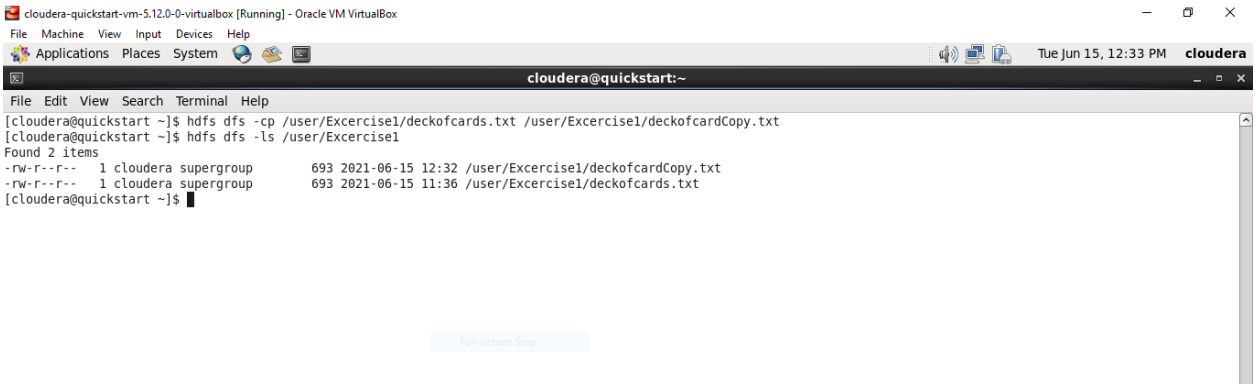
5. Determine the size of the hamlet.txt file in KB that resides on HDFS (not local directory)



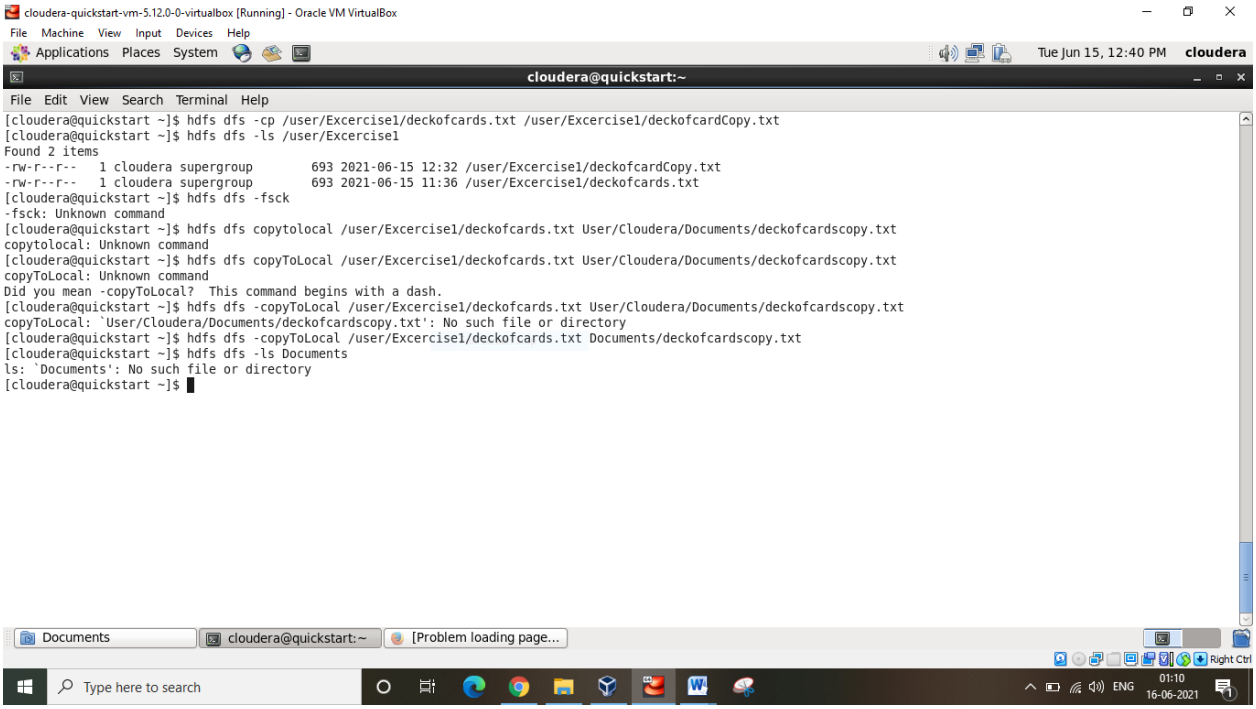
6. Print the first 25 lines to the screen from deckofcards.txt on HDFS



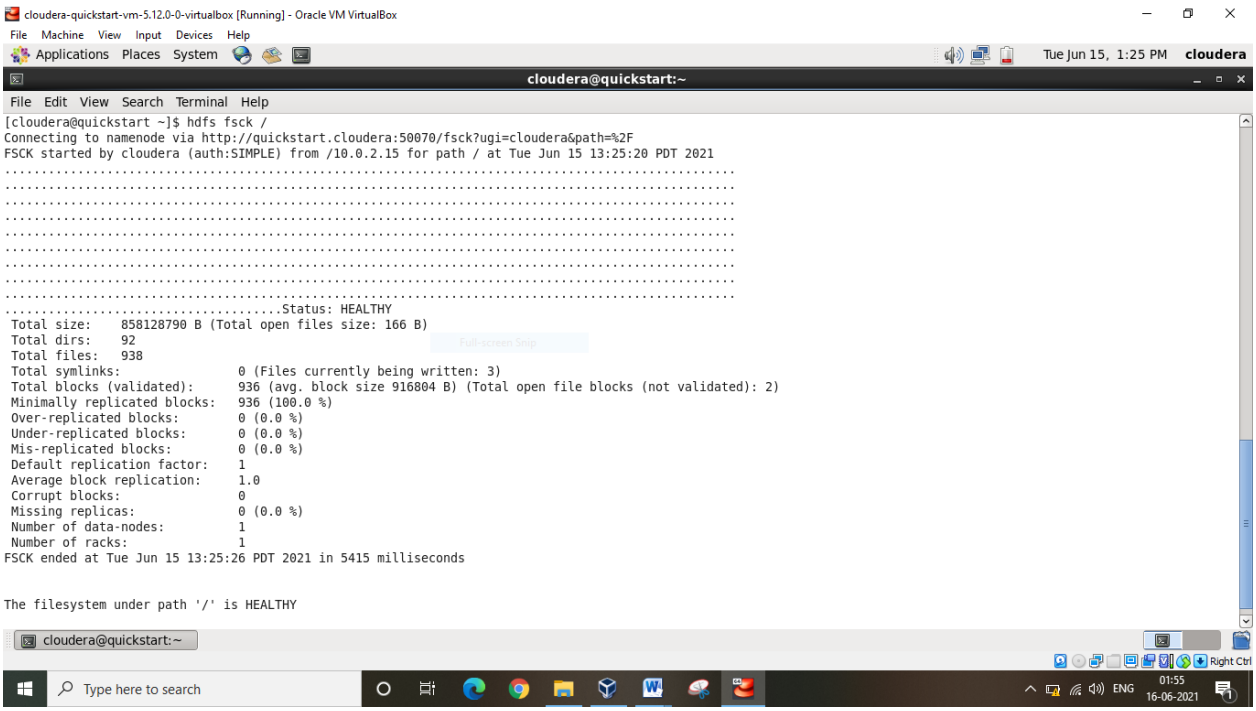
7. Copy deckofcards.txt to deckofcardsCopy.txt



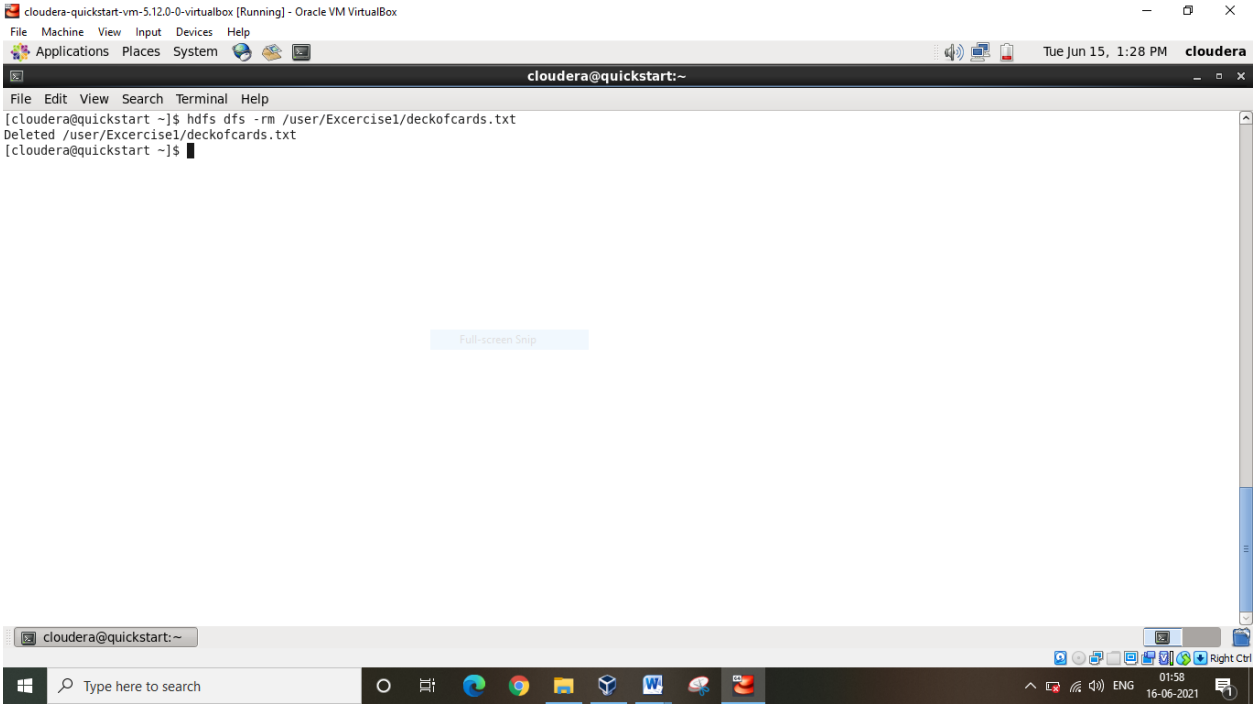
8. Copy deckofcards.txt back to local file system and name it deckofcards.copy.txt



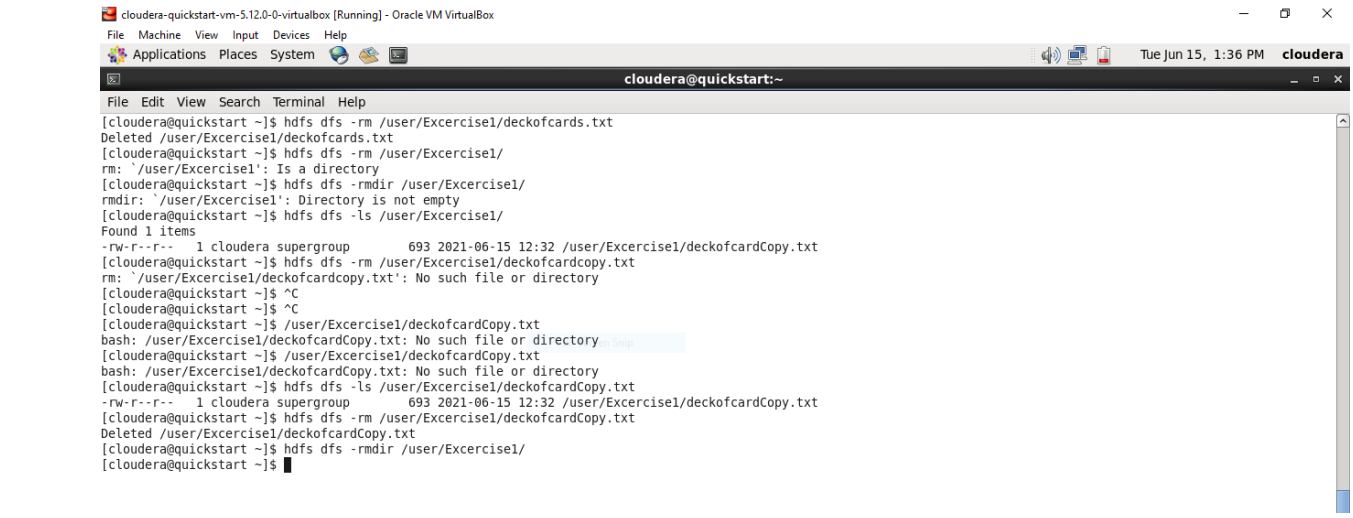
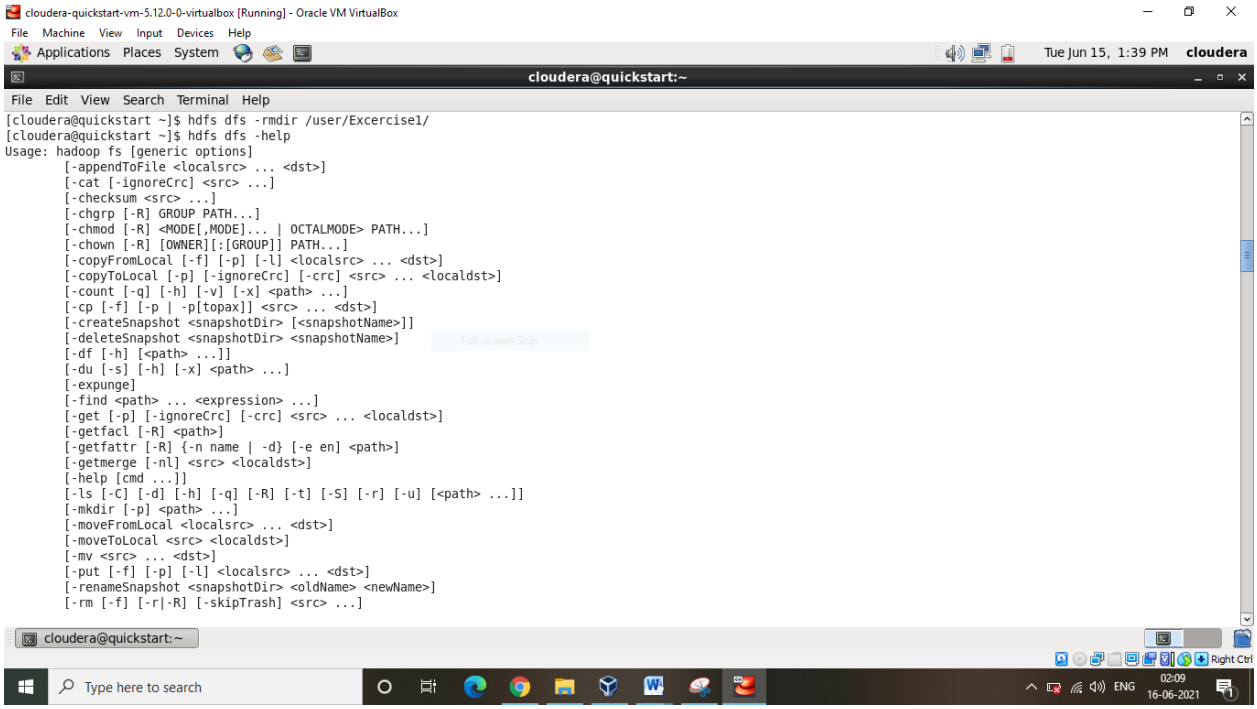
9. Check the entire filesystem for inconsistencies/problems



10. Delete deckofcards.txt from HDFS



11. Delete the /exercise1 directory from HDFS



12. Take a second to look at other available shell options