Hadoop Lab Exercise for Day 1

Task 1

- Delete the virtual disk that we have created in the morning.

- Import another HDP vm.

- Start the registration.

- Browse to Ambari.

Task 2

- Login as HDFS user.

- Create an *input* folder

- Upload the following files into *input* folder
    - https://raw.githubusercontent.com/cherhan/hadooptraining/master/alice.txt

    - https://raw.githubusercontent.com/cherhan/hadooptraining/master/pride.txt

    - https://raw.githubusercontent.com/cherhan/hadooptraining/master/romeo.txt

- Use MapReduce to process each files individually.

- Using MapReduce to process the entire folder

- View the results and make sure all the results are correct.