ICP 1:

Topic : Installing Cloudera and visualize Hadoop data with Hue

Task:

* Install Cloudera
* Load datasets into HDFS
* Append both files
* Visualize the result file with Hue
* Display first and last 5 lines of the result file
* Load new file and append data of all the 3 datasets

Features:

* Cloudera
* Hadoop
* Hue

Questions:

* Creating new directory BDP:



* Copy the files to hadoop hdfs





appendToFile:



appendToFile – copies files from local file system to a destination file system

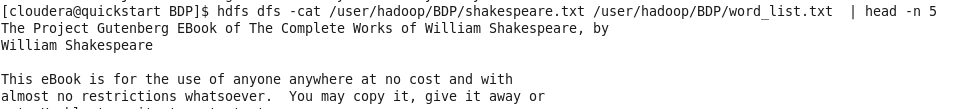
Appending the files using Cat command and storing moving the output to HDFS using put command,



View the first 5 lines of merged dataset using appropriate hdfs commands



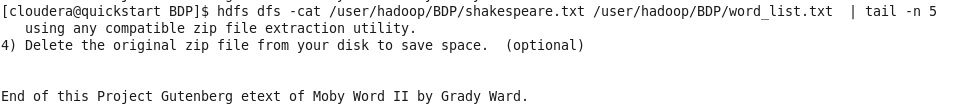
Output



View the first 5 lines of merged dataset using appropriate hdfs commands



Output:



Create a new text file and load it into hdfs and try to append all three datasets.





Visualize file with Hue

