1. Download Virtual Box and Cloudera

2. Import Cloudera to Virtual Box and run it

3. Run Cloudera

   3.1. Create file in hdfs hadoop fs -mkdir /user/cloudera/icp1

   3.2. Push shakespeare.txt to hdfs hadoop fs -put shakespeare.txt /user/cloudera/icp1

   3.3. Append word_file.txt hadoop fs -appendToFile word_list.txt /user/cloudera/icp1/shakespeare.txt

   3.4. Query the first and last five lines hadoop fs -cat /user/cloudera/icp1/shakespeare.txt | head -n 5 hadoop fs -cat /user/cloudera/icp1/shakespeare.txt | tail -n 5

   3.5. Create a new text file, push it to hdfs and append it to merged dataset. hadoop fs -put random_text.txt /user/cloudera/icp1 hadoop fs -appendToFile random_text.txt /user/cloudera/icp1/shakespeare.txt