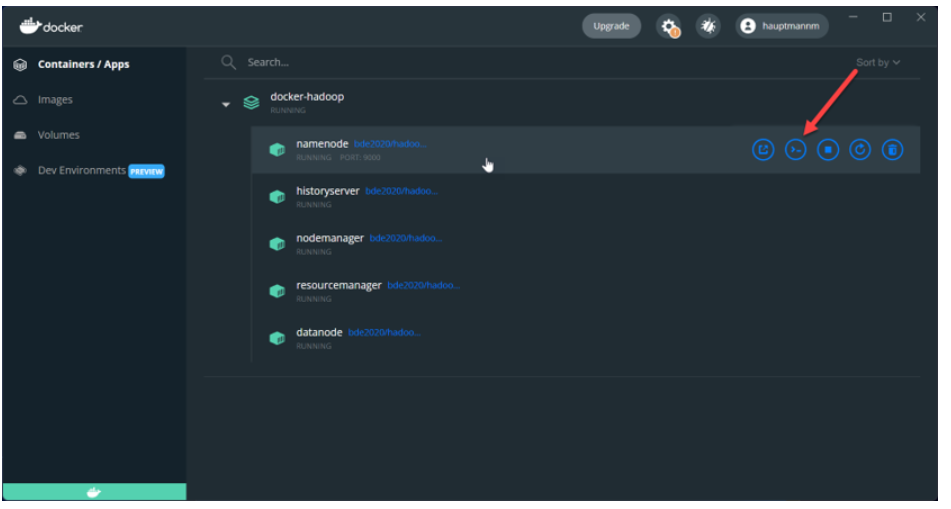
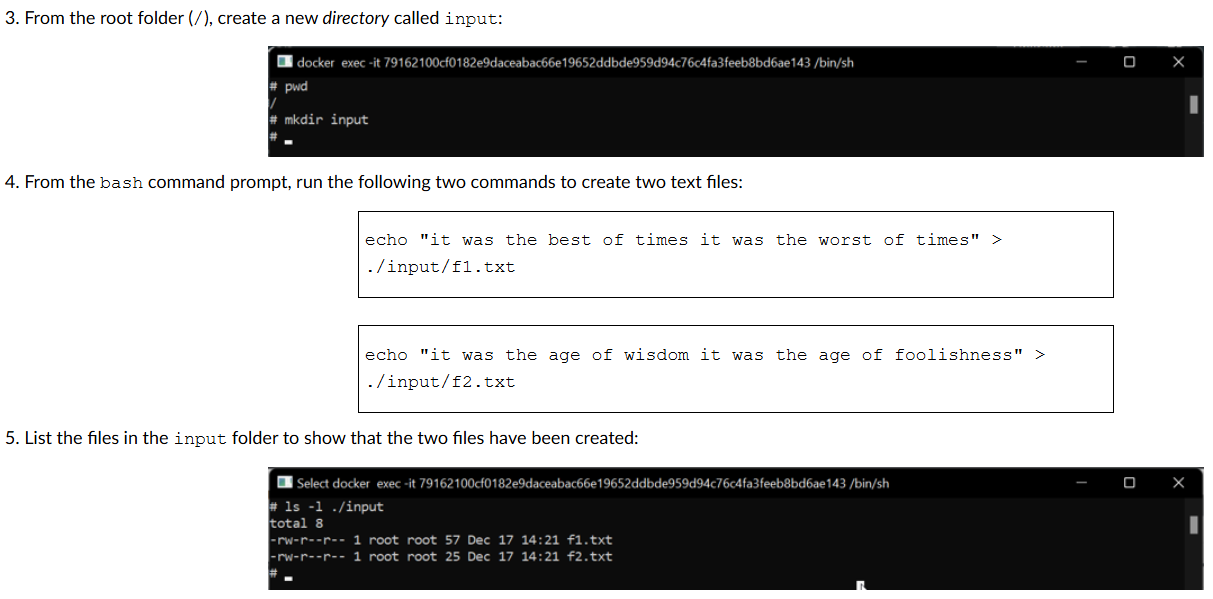
Counting words in a text file is a great example to illustrate how MapReduce works in Hadoop. The image below shows how the text provided by Dr. Sanchez is broken down (Mapped) and then (Reduced) to provide the word count (output):



Follow the steps below to perform the simple word count program that Dr. Sanchez demonstrates in the videos in this section, [Using Hadoop to Handle Big Data](https://classroom.emeritus.org/courses/10605/pages/using-hadoop-to-handle-big-data-06-05):

1. Make sure that you installed Hadoop in Docker and that all of the *containers* are running.
2. From the Docker desktop, select <CLI> on the namenode *container*:





6. To allow Hadoop to access the text files to perform the word count, you will need to copy the files from the *container*’s file system to the HDFS. From the bash command prompt, run the following command to create a folder called input in the HDFS:

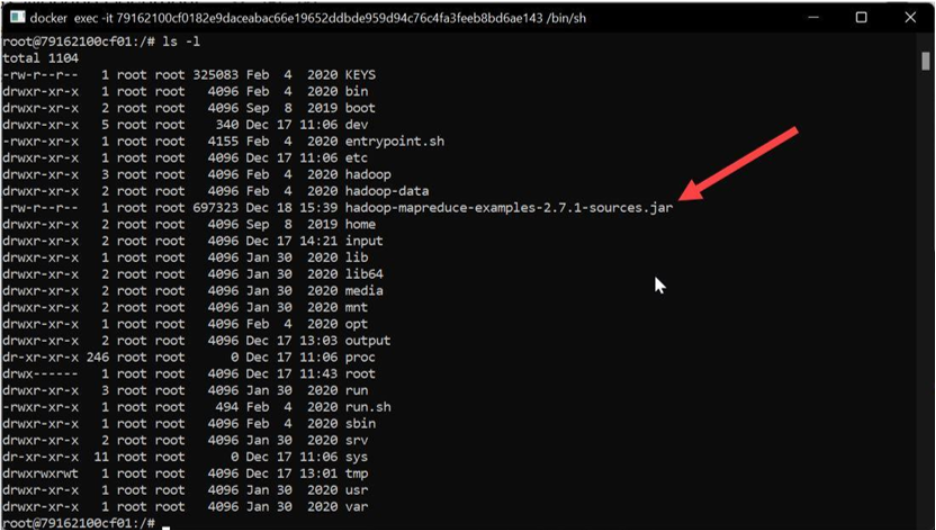
| hadoop fs -mkdir -p input |
| --- |

7. Next, you need to copy the two text files that you created to the HDFS. From the bash command prompt, run the following command to copy the files from the input folder to the HDFS:

| hdfs dfs -put ./input/\* input |
| --- |

8. You need to download the jar file containing the word count program. You can use curl to download the jar file and save it to the namenode *container*. From the Docker desktop, locate the namenode *container* and select <CLI> to open a bash window. Run the following command to download the hadoop-mapreduce-examples-2.7.1-sources.jar file:

| curl -L https://repo1.maven.org/maven2/org/apache/hadoop/hadoop-mapreduce-examples/2.7.1/hadoop-mapreduce-examples-2.7.1-sources.jar --output hadoop-mapreduce-examples-2.7.1-sources.jar |
| --- |

9. List the files in the *container* to show that the jar file has been downloaded to the namenode *container*:

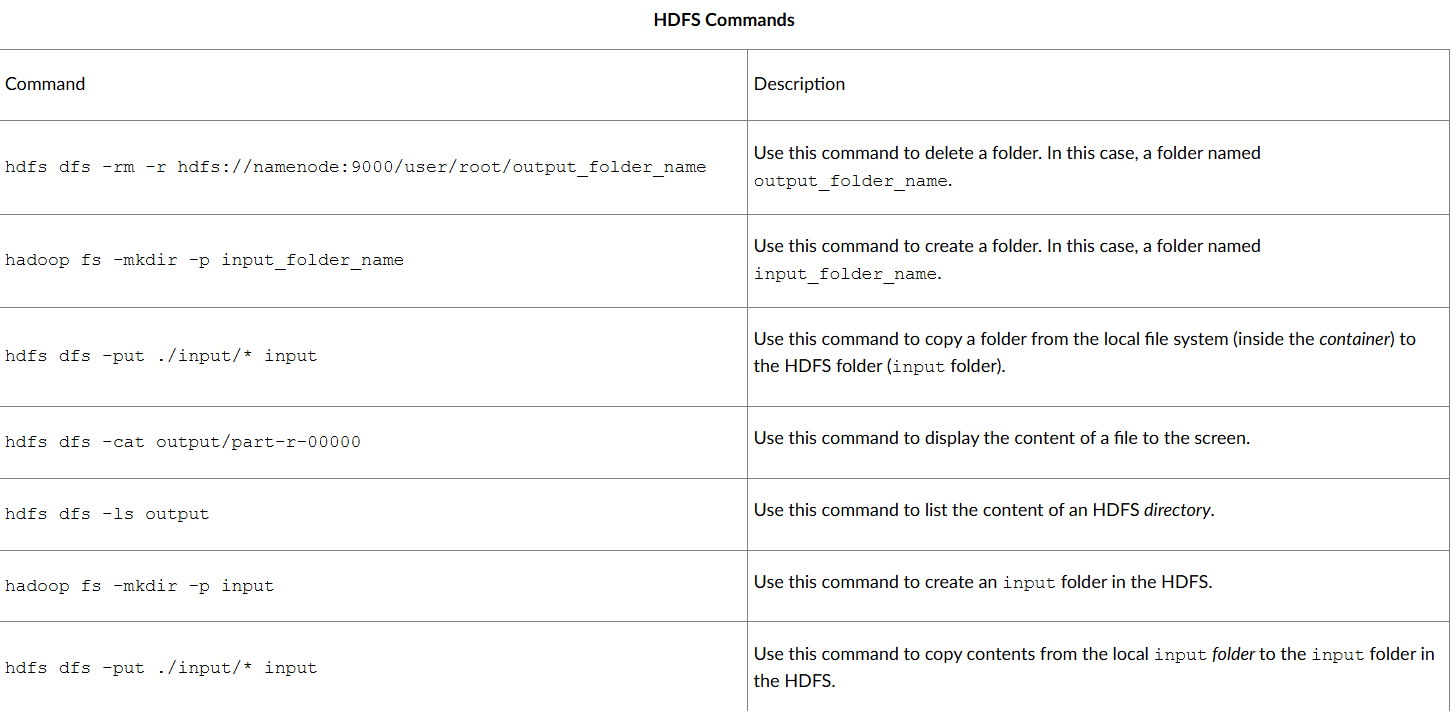
10. Now you are ready to run the word count program. From the namenode Docker *container* bash window, enter the following command:

| hadoop jar hadoop-mapreduce-examples-2.7.1-sources.jar org.apache.hadoop.examples.WordCount input output |
| --- |

11. To look at the results of the program, you will need to look at the output folder in the HDFS. From the namenode Docker *container* bash window, enter the following HDFS command to look at the results from running the word count program:

| hdfs dfs -cat output/part-r-00000 |
| --- |

Here are some useful commands that you can use to inspect and manipulate the HDFS:

In this mini-lesson, you learned how to run a simple word count program in Hadoop to illustrate the MapReduce architecture. You also learned to manipulate the HDFS to copy files from the local file system to the HDFS and view the content of files in the HDFS.