

COPY FILES INTO HADOOP

Learning Goals

By the end of this activity, you will be able to:

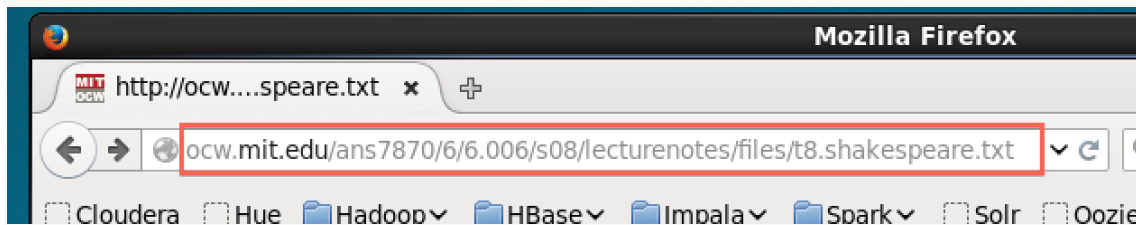
- Interact with Hadoop using the command-line application.
- Copy files into and out of the Hadoop Distributed File System (HDFS).

Instructions

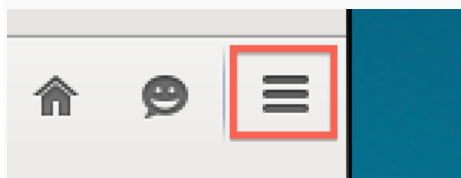
1. **Open a browser.** Open the browser by click on the browser icon on the top left of the screen.



2. **Download the Shakespeare.** We are going to download a text file to copy into HDFS. Enter the following link in the browser: <http://ocw.mit.edu/ans7870/6/6.006/s08/lecturenotes/files/t8.shakespeare.txt>



Once the page is loaded, click on the Open menu button.

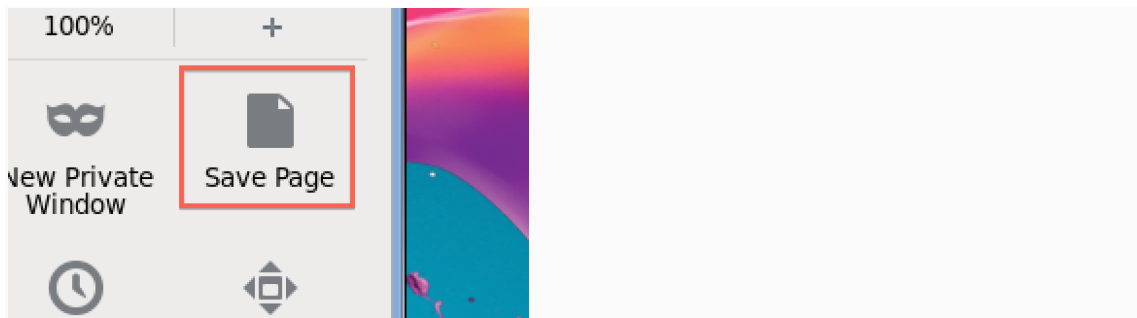


Click on Save Page

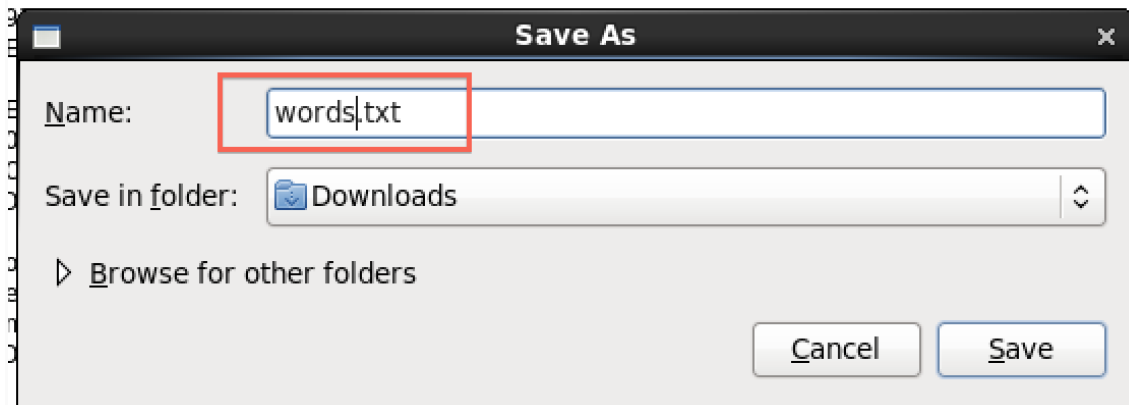
CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

Cartagena99



Change the output to words.txt and click Save.



2. **Open a terminal shell.** Open a terminal shell by clicking on the square black box on the top left of the screen.



Run `cd Downloads` to change to the Downloads directory.

```
[cloudera@quickstart ~]$ cd Downloads/  
[cloudera@quickstart Downloads]$
```

Run `ls` to see that words.txt was saved.

```
[cloudera@quickstart Downloads]$ ls  
words.txt
```

3. **Copy file to HDFS.** Run `hadoop fs -copyFromLocal words.txt` to copy the text file to HDFS.

Cartagena99

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

5. **Copy a file within HDFS.** You can make a copy of a file in HDFS. Run `hadoop fs -cp words.txt words2.txt` to make a copy of words.txt called words2.txt

```
[cloudera@quickstart Downloads]$ hadoop fs -cp words.txt words2.txt  
[cloudera@quickstart Downloads]$ █
```

We can see the new file by running `hadoop fs -ls`

```
[cloudera@quickstart Downloads]$ hadoop fs -ls  
Found 2 items  
-rw-r--r-- 1 cloudera cloudera 5458199 2016-02-12 15:14 words.txt  
-rw-r--r-- 1 cloudera cloudera 5458199 2016-02-12 15:15 words2.txt  
[cloudera@quickstart Downloads]$ █
```

6. **Copy a file from HDFS.** We can also copy a file from HDFS to the local file system. Run `hadoop fs -copyToLocal words2.txt` to copy words2.txt to the local directory.

```
[cloudera@quickstart Downloads]$ hadoop fs -copyToLocal words2.txt  
[cloudera@quickstart Downloads]$ █
```

Let's run `ls` to see that the file was copied to see that words2.txt is there.

```
[cloudera@quickstart Downloads]$ ls  
words2.txt words.txt  
[cloudera@quickstart Downloads]$ █
```

7. **Delete a file in HDFS.** Let's delete words2.txt in HDFS. Run `hadoop fs -rm words2.txt`

```
[cloudera@quickstart Downloads]$ hadoop fs -rm words2.txt  
16/02/12 15:17:01 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Empty interval = 0 minutes.  
Deleted words2.txt  
[cloudera@quickstart Downloads]$ █
```

Run `hadoop fs -ls` to see that the file is gone.

```
[cloudera@quickstart Downloads]$ hadoop fs -ls  
Found 1 items  
-rw-r--r-- 1 cloudera cloudera 5458199 2016-02-12 15:14 words.txt  
[cloudera@quickstart Downloads]$ █
```

The logo for Cartagena99 features the text 'Cartagena99' in a stylized, blue, serif font. The '99' is significantly larger and more prominent than the rest of the text. The logo is set against a light blue background with a white starburst effect behind the text.

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70