

Installation Instructions

Java

```
brew cask install java
```

Hadoop

```
brew install hadoop
```

SSH Localhost

```
ssh localhost
```

Text Data

Download from free [Project Gutenberg](#)

Executable Python Scripts

```
chmod +x mapper.py
```

```
chmod +x reduce.py
```

To test locally:

```
echo "foo foo quux labs foo bar quux" | ./mapper.py | sort  
| ./reduce.py
```

Format HDFS

```
hadoop namenode -format
```

Start Hadoop

```
$ cd /usr/local/Cellar/hadoop/2.7.3/libexec/sbin  
$ ./start-dfs.sh  
$ ./start-yarn.sh
```

Hadoop Cluster

```
cd /usr/local/Cellar/hadoop/2.7.3  
bin/hadoop dfs -copyFromLocal /tmp/gutenberg gutenberg  
bin/hadoop jar /usr/local/Cellar/hadoop/2.7.3/libexec/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.7.3.jar  
wordcount gutenberg gutenberg-output  
bin/hadoop dfs -cat gutenberg-output/part-r-00000 >> result.txt
```

View Result

```
open result.txt
```

```
1. ssh
× bash %1 × ssh %2 × bash %3
thing, 48
thing," 5
thing,--out 1
thing--though 1
thing. 14
thing." 6
thing." 2
thing: 5
thing; 11
thing;" 1
thing? 2
thing?" 1
thing_ 1
thinge 3
thinges 7
things 314
things! 5
things' 1
things, 63
things," 5
things. 14
things." 4
things." 1
things: 3
things; 11
things? 1
think 581
think! 1
think, 79
think," 6
think," 2
think. 6
think." 5
think." 4
think: 3
```