**Instructions to Install the VM**

1. Download VirtualBox from: <https://www.virtualbox.org/wiki/Downloads>
   1. Pick the appropriate binary for your operating system.
   2. Then follow the prompts for installing VirtualBox.
2. Once the installation is complete, download the latest Cloudera Quickstart VM from Cloudera’s main Web Site:

<http://www.cloudera.com/downloads/quickstart_vms/5-5.html>

You may need to download 7-zip from [www.7-zip.org](http://www.7-zip.org) to extract the Cloudera VM. After uncompressing, you will get a file called “cloudera-quickstart-vm-5.X.X-0-virtualbox.ovf.” Move it to location of your choice in File System.

1. Once you have extracted the VM, we will load the VM into VirtualBox.
   1. Open Virtual Box and click on **File -> Import Appliance...**
   2. From the file dialog open **cloudera-quickstart-vm-5.4.2-0-.clairvoyant-spark-workshop.ova** located in the decompressed (or unzipped) Cloudera VM download.
2. Update Virtual Box to the latest version if you are not able to add as described in step 4. The menu might be present in an older version, but may throw an exception/error message when you attempt to add.
3. Start the VM.
4. Once the VM starts up, you should see the Desktop within VirtualBox. This is your sandbox to play with Hadoop.
5. Test to ensure the VM was setup successfully. Open Terminal and run the following commands:
   1. $ hadoop version
      1. If the setup is successful, it will print the current version.
   2. $ spark-shell
      1. In the shell that opens run command:

scala> **sc**

1. If the setup is successful, it will print a SparkContext object