**DOCUMENT TO INSTALL AND RUN HIVE ON MAC**

1. Download and install Java (JDK) 8 from <https://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html>

Accept the licence and download the version for mac(Create an oracle account if needed. BTW, I don’t like this step). Double click the downloaded dmg file and Install Java.

A screenshot of a cell phone

Description automatically generated

1. Download and extract hive and Hadoop in your desired folder, say ‘Assignment’.
2. <http://apachemirror.wuchna.com/hadoop/common/hadoop-2.10.0/hadoop-2.10.0.tar.gz>
3. <http://mirrors.estointernet.in/apache/hive/hive-2.3.6/apache-hive-2.3.6-bin.tar.gz>

A screenshot of a social media post

Description automatically generated

1. Now, be nice and tell everything to your MAC about it ☺, so that MAC will be nice to let you execute hive command from anywhere in the terminal.
2. You should set JAVA\_HOME, HADOOP\_HOME and HIVE\_HOME variables in ~/.bash\_profile (Caution! The weird symbols before the file name are important in the order). You can use the ‘nano’ editor to do this… Just type the following in terminal

nano ~/.bash\_profile

Please look how I have ‘set’ the variables and ‘included’ them in the ‘PATH’. JAVA\_HOME is kind of set separately because I already had Java. So, you can set as like I have set the other variables here (between the red lines). Press control+O to write and press enter to save the file and come back to terminal (See the commands for the editor at the bottom).

A screenshot of a cell phone

Description automatically generated

1. To the changes to take effect, run the following in the terminal

source ~/.bash\_profile

1. Check if you have installed Java and Hadoop correctly,

A screenshot of a cell phone

Description automatically generated

1. We are not going to use Hadoop but for hive to work, looks like it’s needed. We will be using out local filesystem only. So, now navigate into the hive installation folder and create a tmp folder.

A screenshot of a cell phone

Description automatically generated

1. Now let’s tell hive to use this. Navigate to ‘conf’ folder, duplicate the hive-default.xml.template file and name it as hive-site.xml. Do the following changes in this file. I have uploaded my file to GitHub. You can directly pull it inside your conf folder if you are lazy to do the changes(<https://github.com/karthi190/utils/blob/master/hive-site.xml>)
   1. Copy and paste the following into the <configuration> tag on top of the file

<property>

<name>system:java.io.tmpdir</name>

<value>/Users/Karthi/ISB-AMPBA2020/Term-2/BDM2/Assignment/apache-hive-2.3.6-bin/tmp</value>

<description>

Local directory path

</description>

</property>

<property>

<name>system:user.name</name>

<value>user.name</value>

<description>

User name

</description>

</property>

* 1. Modify the values in the file as follows (highlighted in yellow)

<property>

<name>hive.metastore.schema.verification</name>

<value>false</value>

</property>

<property>

<!-- this should eventually be deprecated since the metastore should supply this -->

<name>hive.metastore.warehouse.dir</name>

<value>file:///tmp</value>

<description></description>

</property>

1. Run the following in terminal.

schematool -dbType derby -initSchema

1. After it runs successfully, Run hive in terminal to see the hive> prompt. Try executing some queries and have fun. It might be a lil slow, but that’s ok.

A screenshot of a social media post

Description automatically generated