

Hive installation:

1. Download hive

```
wget <URL>
```

2. Unzip

```
tar -xzf apache-hive-2.3.3-bin.tar.gz
```

3. Modify bashrc file

```
#HIVE VARIABLES START
export HIVE_HOME=$HADOOP_INSTALL/hive
export HIVE_CONF_DIR=$HIVE_HOME/conf
export PATH=$PATH:$HIVE_HOME/bin
#HIVE VARIABLES END
```

4. HDFS setup

```
hdfs dfs -mkdir -p /user/hive/warehouse
hdfs dfs -mkdir /tmp
hdfs dfs -chmod g+w /user/hive/warehouse
hdfs dfs -chmod g+w /tmp
```

5. Modify hive-env.sh

```
export HADOOP_HOME=/usr/local/hadoop
```

6. Modify hive-site.xml

```
<configuration>

<property>
<name>javax.jdo.option.ConnectionURL</name>
<value>jdbc:derby;;databaseName=$HIVE_HOME/metastore_db;create=true</value>
<description>
JDBC connect string for a JDBC metastore.
To use SSL to encrypt/authenticate the connection, provide database-specific SSL flag
in the connection URL.
For example, jdbc:postgresql://myhost/db?ssl=true for postgres database.
</description>
</property>

<property>
<name>hive.metastore.warehouse.dir</name>
<value>/user/hive/warehouse</value>
<description>location of default database for the warehouse</description>
</property>
```

```
<property>
<name>javax.jdo.option.ConnectionDriverName</name>
<value>org.apache.derby.jdbc.EmbeddedDriver</value>
<description>Driver class name for a JDBC metastore</description>
</property>

<property>
<name>javax.jdo.PersistenceManagerFactoryClass</name>
<value>org.datanucleus.api.jdo.JDOPersistenceManagerFactory</value>
<description>class implementing the jdo persistence</description>
</property>

</configuration>
```

7. Initialize Derbi database within hive directory

bin/schematool -initSchema -dbType derby

8. Run hive CLI to test it

Reference:

<https://www.edureka.co/blog/apache-hive-installation-on-ubuntu>

http://www.bogotobogo.com/Hadoop/BigData_hadoop_Hive_Install_On_Ubuntu_16_04.php