

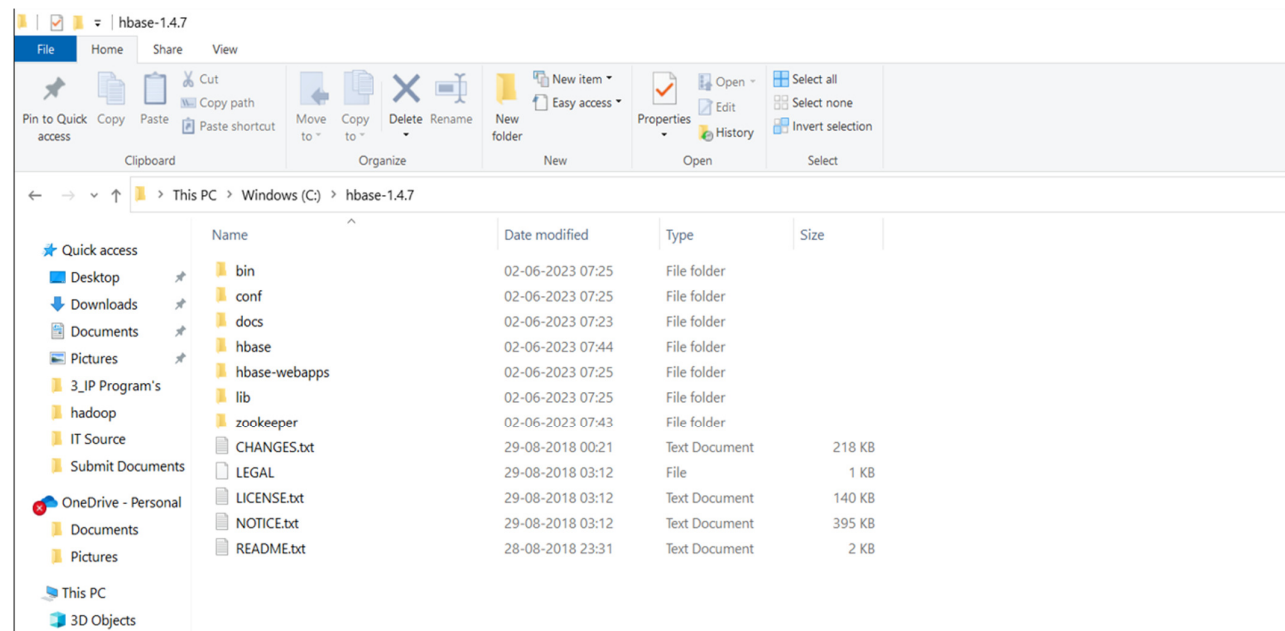
## Aim: Implement an application that store big data in Hbase/ Mongodb/ Pig using Hadoop / R.

Hbase - Standalone mode installation.

### STEP - 1: Extract the HBase file

Download from: <http://www.apache.org/dyn/closer.lua/hbase/>

Extract file hbase-1.4.7-bin.tar.gz and place under "C:\hbase-1.4.7", you can use any preferred location –

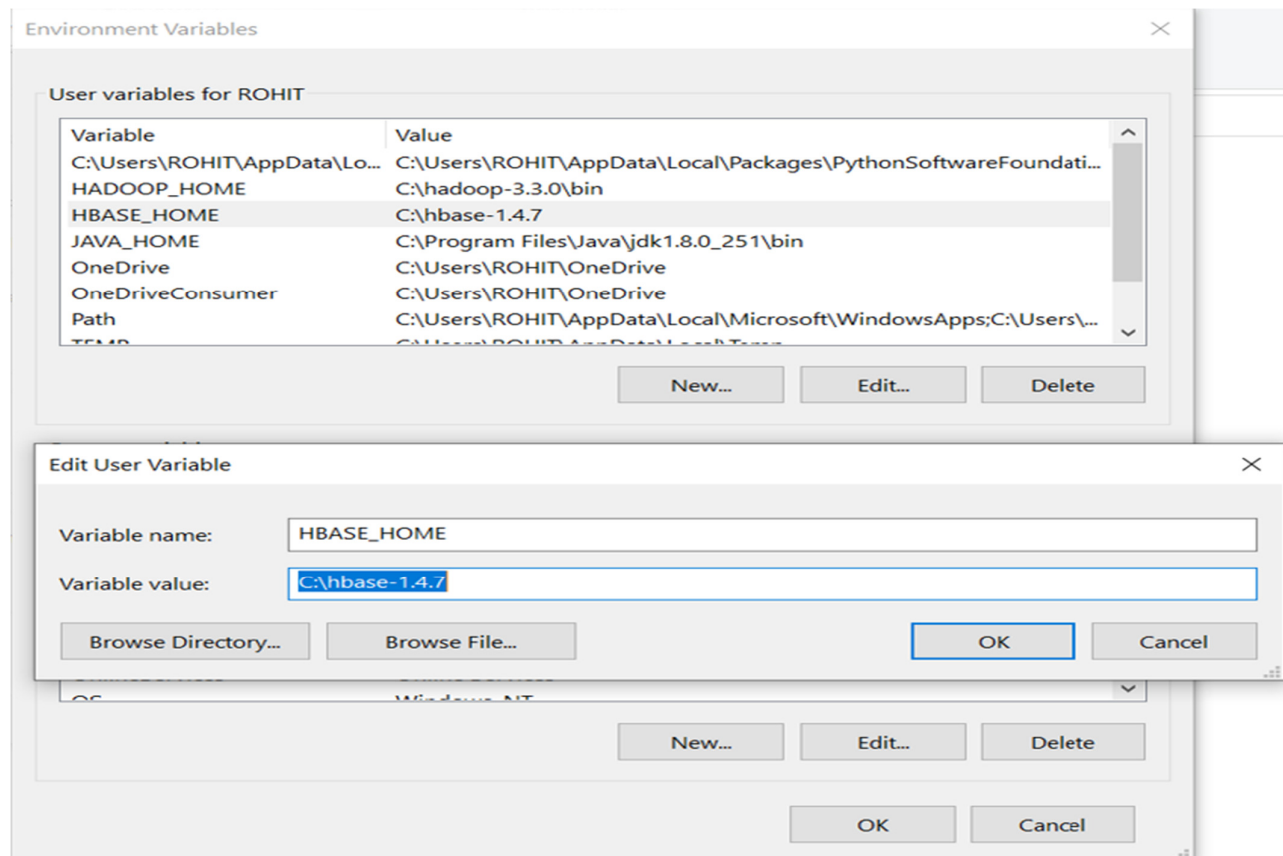
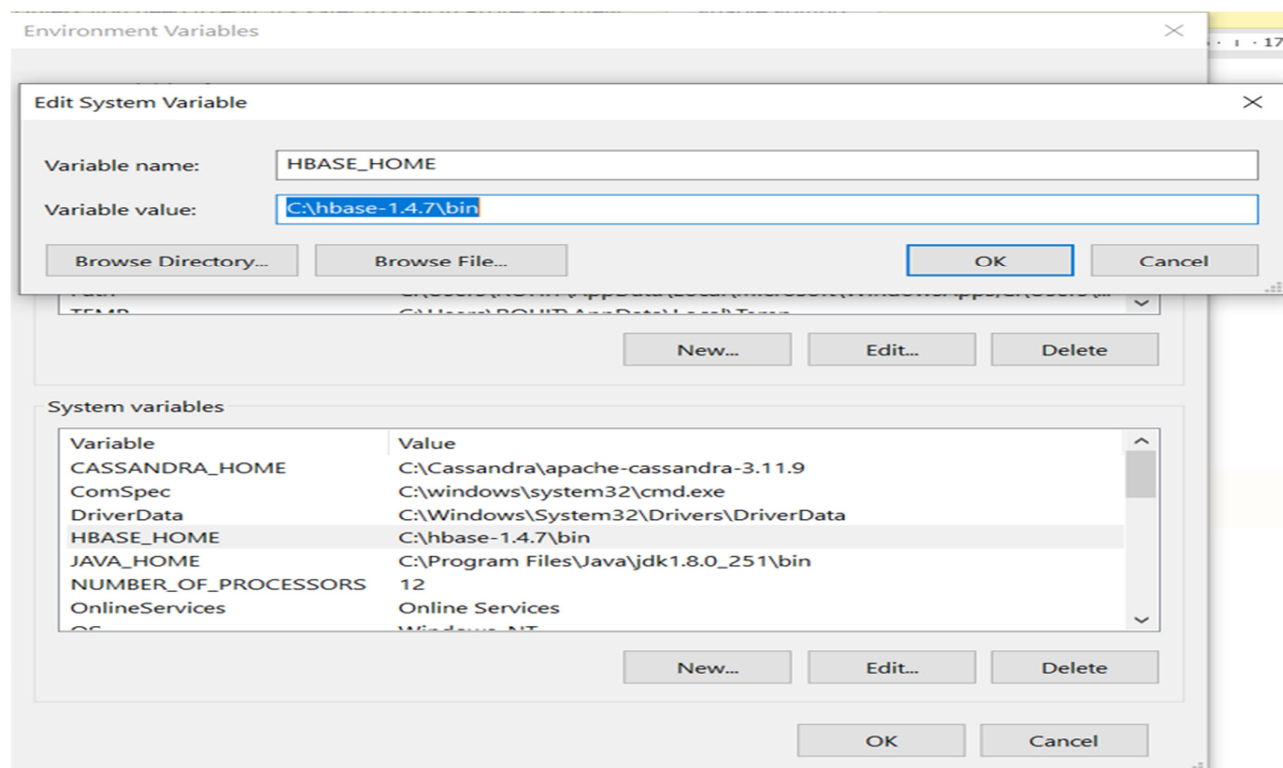


### STEP - 2: Configure Environment variable.

Set the path for the following Environment variable (User Variables) on windows 10 –

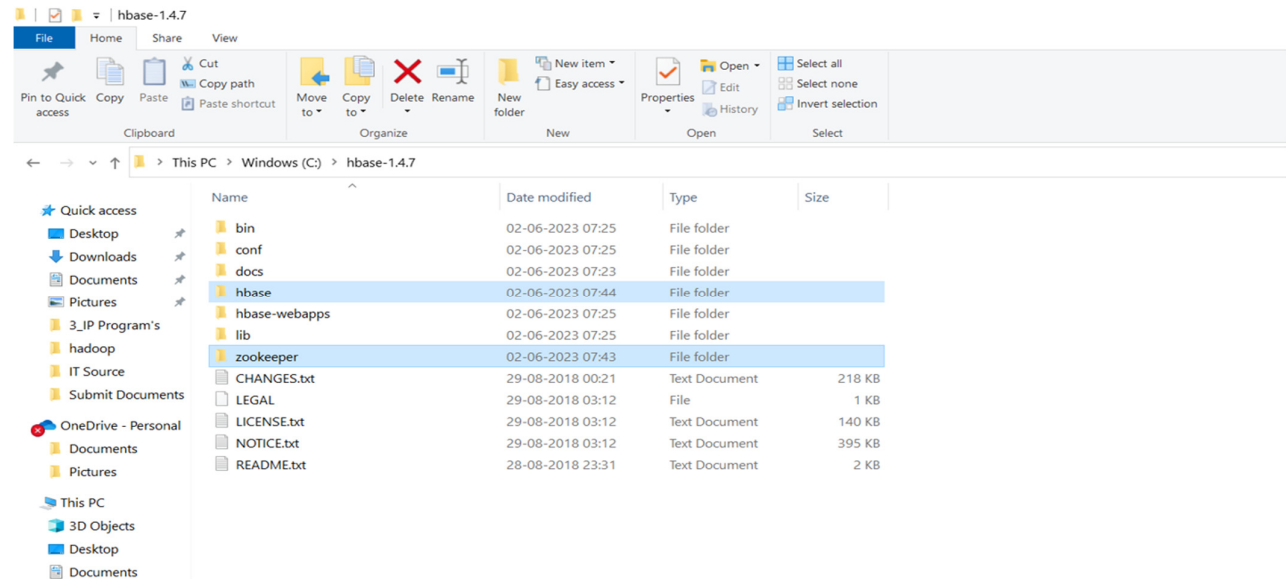
- HBASE\_HOME - C:\hbase-1.4.7

This PC -> Right Click -> Properties -> Advanced System Settings -> Advanced -> Environment Variables.

**STEP - 3: Configure System variable.**

**STEP - 4: Create required folders.**

1. Create folder "hbase" under "C:\hbase-1.4.7".
2. Create folder "zookeeper" under "C:\hbase-1.4.7".

**STEP - 5: Configured required files.**

Next, essential to configure two key files with minimal required details –

- hbase-env.cmd
- hbase-site.xml

1) Edit file "C:\hbase-1.4.7\conf\hbase-env.cmd", mention JAVA\_HOME path in the location and save this file.

```
@rem set JAVA_HOME=c:\apps\java
```

```
set JAVA_HOME=C:\PROGRA~1\Java\jdk1.8.0_251
```

```
hbase-env.cmd - Notepad
File Edit Format View Help
@rem/**
@rem * Licensed to the Apache Software Foundation (ASF) under one
@rem * or more contributor license agreements. See the NOTICE file
@rem * distributed with this work for additional information
@rem * regarding copyright ownership. The ASF licenses this file
@rem * to you under the Apache License, Version 2.0 (the
@rem * "License"); you may not use this file except in compliance
@rem * with the License. You may obtain a copy of the License at
@rem *
@rem * http://www.apache.org/licenses/LICENSE-2.0
@rem *
@rem * Unless required by applicable law or agreed to in writing, software
@rem * distributed under the License is distributed on an "AS IS" BASIS,
@rem * WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
@rem * See the License for the specific language governing permissions and
@rem * limitations under the License.
@rem */

@rem Set environment variables here.

@rem The java implementation to use. Java 1.7+ required.
@rem set JAVA_HOME=c:\apps\java
set JAVA_HOME=C:\PROGRA~1\Java\jdk1.8.0_251

@rem Extra Java CLASSPATH elements. Optional.
@rem set HBASE_CLASSPATH=
```

2) Edit file "C:\hbase-1.4.7\conf\hbase-site.xml", paste below xml paragraph and save this file.

```
<configuration>
  <property>
    <name>hbase.rootdir</name>
    <value>C:\hbase-1.4.7\hbase</value>
  </property>
  <property>
    <name>hbase.zookeeper.property.dataDir</name>
    <value>C:\hbase-1.4.7\zookeeper</value>
  </property>
  <property>
    <name> hbase.zookeeper.quorum</name>
    <value>127.0.0.1</value>
  </property>
</configuration>
```

All HMaster and ZooKeeper activities point out to this hbase-site.xml.

### STEP - 6: Start HBase

Open command prompt and change directory to "C:\hbase-1.4.7\bin" and type "start-hbase.cmd" to start HBase.

```
C:\Windows\System32\cmd.exe - hbase shell
Microsoft Windows [Version 10.0.19045.2965]
(c) Microsoft Corporation. All rights reserved.

C:\hbase-1.4.7\bin>start-hbase.cmd
```

It will open a separate instances of cmd for following tasks – HBase Master

```
HBase Distribution - C:\hbase-1.4.7\bin\hbase.cmd [cmd.exe]
Accesses=0, cachingHits=0, cachingHitsRatio=0, evictions=59, evicted=0, evictedPerRun=0.0
2023-06-02 07:53:59,212 INFO [Idle-Rpc-Conn-Sweeper-pool2-t1] ipc.AbstractRpcClient: Cleanup idle connection to LAPTOP-10M148EB/192.168.43.154:58392
2023-06-02 07:53:59,212 INFO [Idle-Rpc-Conn-Sweeper-pool2-t1] ipc.AbstractRpcClient: Cleanup idle connection to LAPTOP-10M148EB/192.168.43.154:58392
2023-06-02 07:55:11,077 WARN [RpcServer.default.FP00.Fifo.handler-29,queue=2,port=58392] regionserver.HRegion: No such column family in batch mutation.
org.apache.hadoop.hbase.regionserver.HRegionColumnFamilyException: Column family bigdata does not exist in region students,1685672447484.931620f97af94b020b6b41cf97d6bae, in table 'students'
, {NAME => 'big data', BLOOMFILTER => 'ROW', VERSIONS => '1', IN_MEMORY => 'false', KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', COMPRESSION => 'NONE', M
IN_VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0'}, {NAME => 'valia', BLOOMFILTER => 'ROW', VERSIONS => '1', IN_MEMORY => 'false', KEEP_DELETED_CELLS
=> 'FALSE', DATA_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', COMPRESSION => 'NONE', MIN_VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REPLICATION_SCOPE => '0'}
at org.apache.hadoop.hbase.regionserver.HRegion.checkFamily(HRegion.java:8299)
at org.apache.hadoop.hbase.regionserver.HRegion.checkFamilies(HRegion.java:4096)
at org.apache.hadoop.hbase.regionserver.HRegion.checkAndPrepareMutation(HRegion.java:4108)
at org.apache.hadoop.hbase.regionserver.HRegion.doMiniBatchMutation(HRegion.java:3215)
at org.apache.hadoop.hbase.regionserver.HRegion.batchMutate(HRegion.java:3108)
at org.apache.hadoop.hbase.regionserver.HRegion.batchMutate(HRegion.java:3050)
at org.apache.hadoop.hbase.regionserver.RSRpcServices.doBatchOp(RSRpcServices.java:916)
at org.apache.hadoop.hbase.regionserver.RSRpcServices.doNonAtomicRegionMutation(RSRpcServices.java:844)
at org.apache.hadoop.hbase.regionserver.RSRpcServices.multi(RSRpcServices.java:2405)
at org.apache.hadoop.hbase.protobuf.generated.ClientProtos$ClientService$2.callBlockingMethod(ClientProtos.java:36621)
```

### STEP - 7: Validate HBase.

Post successful execution of HBase, verify the installation using following commands.

- hbase –version

```
Select C:\Windows\System32\cmd.exe - hbase shell

C:\hbase-1.4.7\bin>hbase -version
java version "1.8.0_251"
Java(TM) SE Runtime Environment (build 1.8.0_251-b08)
Java HotSpot(TM) 64-Bit Server VM (build 25.251-b08, mixed mode)
```

- jps

```
Select C:\Windows\System32\cmd.exe - hbase shell

C:\hbase-1.4.7\bin>hbase -version
java version "1.8.0_251"
Java(TM) SE Runtime Environment (build 1.8.0_251-b08)
Java HotSpot(TM) 64-Bit Server VM (build 25.251-b08, mixed mode)

C:\hbase-1.4.7\bin>jps
24176 Jps
5132 HMaster
```

If we can see HMaster is in running mode, then our installation is okay.

### STEP - 8: Execute HBase Shell.

The standalone mode does not require Hadoop daemons to start. HBase can run independently. HBase shell can start by using "hbase shell" and it will enter into interactive shell mode –

```
Select C:\Windows\System32\cmd.exe - hbase shell

C:\hbase-1.4.7\bin>hbase shell
2023-06-02 07:46:06,024 ERROR [main] util.Shell: Failed to locate the winutils binary in the hadoop binary path
java.io.IOException: Could not locate executable C:\hadoop-3.3.0\bin\bin\winutils.exe in the Hadoop binaries.
    at org.apache.hadoop.util.Shell.getQualifiedBinPath(Shell.java:382)
    at org.apache.hadoop.util.Shell.getWinUtilsPath(Shell.java:397)
    at org.apache.hadoop.util.Shell.<clinit>(Shell.java:390)
    at org.apache.hadoop.util.StringUtils.<clinit>(StringUtils.java:80)
    at org.apache.hadoop.conf.Configuration.getBoolean(Configuration.java:1437)
    at org.apache.hadoop.hbase.HBaseConfiguration.checkDefaultsVersion(HBaseConfiguration.java:67)
    at org.apache.hadoop.hbase.HBaseConfiguration.addHbaseResources(HBaseConfiguration.java:81)
    at org.apache.hadoop.hbase.HBaseConfiguration.create(HBaseConfiguration.java:96)
    at org.apache.hadoop.hbase.HColumnDescriptor.<clinit>(HColumnDescriptor.java:151)
    at java.lang.Class.forName0(Native Method)
    at java.lang.Class.forName(Class.java:348)
```

### STEP-9: Some hands on activities.

#### 1. Create a simple table

create 'students', 'bigdata', 'valia'

```
Select C:\Windows\System32\cmd.exe - hbase shell

HBase Shell
Use "help" to get list of supported commands.
Use "exit" to quit this interactive shell.
Version 1.4.7, r763f27f583cf8fd7ecf79fb6f3ef57f1615dbf9b, Tue Aug 28 14:40:11 PDT 2018

hbase(main):001:0> create 'students', 'big data', 'valia'
NoMethodError: undefined method `create' for #<Object:0x46e64760>

hbase(main):002:0> create 'students', 'big data'
NoMethodError: undefined method `create' for #<Object:0x46e64760>

hbase(main):003:0> create 'students', 'big data', 'valia'
0 row(s) in 1.5490 seconds
```

#### 2. List the table has been created.

- List

```
=> Hbase::Table - students
hbase(main):004:0>
hbase(main):005:0> list
TABLE
students
1 row(s) in 0.0140 seconds

=> ["students"]
hbase(main):006:0> put 'students', 'row1', 'bigdata:hadoop', 'hadoop Tutorials'
```

#### 3. Insert some data to above created table

put 'tablename', 'rowname', 'columnvalue', 'value'

put 'students', 'row1', 'bigdata:hadoop', 'hadoop tutorial'

```
Select C:\Windows\System32\cmd.exe - hbase shell

hbase(main):007:0> put 'students', 'row1', 'valia:hadoop', 'hadoop Tutorials'
0 row(s) in 0.0070 seconds
```

#### 4. List all rows in the table

scan 'students'

```
hbase(main):008:0> scan 'students'
ROW                                COLUMN+CELL
row1                                column=valia:hadoop, timestamp=1685672770496, value=hadoop Tutorials
1 row(s) in 0.0110 seconds

hbase(main):009:0> _
```

\*\*\*\*\*END\*\*\*\*\*