Part1 & Part2:

```
xiao@a-10-27-0-69 ~ % ssh xm2074@peel.hpc.nyu.edu
xm2074@peel.hpc.nyu.edu's password:
Last failed login: Mon Feb  7 15:39:14 EST 2022 from 10.27.0.69 on ssh:notty
There was 1 failed login attempt since the last successful login.
Last login: Mon Feb  7 15:38:20 2022 from 10.27.0.69

[xm2074@hlog-1 ~]$ hadoop fs -ls /
Found 4 items
drwxr-xr-x   - hbase hbase              0 2022-01-13 11:09 /hbase
drwxrwxr-x   - solr  solr               0 2021-01-15 18:01 /solr
drwxrwxrwt   - hdfs  supergroup         0 2022-01-21 14:10 /tmp
drwxr-xr-x   - hdfs  supergroup         0 2022-02-07 14:42 /user
[xm2074@hlog-1 ~]$ hadoop fs -ls
[xm2074@hlog-1 ~]$ hadoop fs -ls /user/xm2074
[xm2074@hlog-1 ~]$ hadoop fs -mkdir lab1
[xm2074@hlog-1 ~]$ hadoop fs -ls
Found 1 items
drwxrwx---+  - xm2074 xm2074            0 2022-02-07 15:40 lab1
[xm2074@hlog-1 ~]$
```

```
[xm2074@hlog-2 ~]$ touch /scratch/xm2074/data.txt
[xm2074@hlog-2 ~]$ hadoop fs -put /scratch/yourNetID/data.txt lab1
put: `/scratch/yourNetID/data.txt': No such file or directory
[xm2074@hlog-2 ~]$ hadoop fs -put /scratch/xm2074/data.txt lab1
[xm2074@hlog-2 ~]$ hadoop fs -ls lab1
Found 1 items
-rw-rw----+  3 xm2074 xm2074            0 2022-02-07 22:57 lab1/data.txt
[xm2074@hlog-2 ~]$ hadoop fs -cat lab1/data.txt
[xm2074@hlog-2 ~]$ hadoop fs -rm -r lab1
22/02/07 22:57:42 INFO fs.TrashPolicyDefault: Moved: 'hdfs://horton.hpc.nyu.edu:
8020/user/xm2074/lab1' to trash at: hdfs://horton.hpc.nyu.edu:8020/user/xm2074/.
Trash/Current/user/xm2074/lab1
[xm2074@hlog-2 ~]$ hadoop fs -ls
Found 3 items
drwx------+  - xm2074 xm2074            0 2022-02-07 22:57 .Trash
drwxrwx---+  - xm2074 xm2074            0 2022-02-07 21:53 .sparkStaging
-rw-rw----+  3 xm2074 xm2074            0 2022-02-07 21:40 data.txt
[xm2074@hlog-2 ~]$
```

Part3:

```
[xm2074@hlog-1 ~]$ spark-shell --deploy-mode client
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLe
Spark context Web UI available at http://hlog-1.hpc.nyu.edu:4040
Spark context available as 'sc' (master = yarn, app id = application_1640038803823_0447
Spark session available as 'spark'.
Welcome to
      ____              __
     / __/__  ___ _____/ /__
    _\ \/ _ \/ _ `/ __/  '_/
   /___/ .__/\_,_/_/ /_/\_\   version 2.4.0-cdh6.3.4
      /_/

Using Scala version 2.11.12 (Java HotSpot(TM) 64-Bit Server VM, Java 1.8.0_181)
Type in expressions to have them evaluated.
Type :help for more information.

scala> :help
All commands can be abbreviated, e.g., :he instead of :help.
:edit <id>|<line>        edit history
```

```
scala> sc.
accumulable             defaultMinPartitions    isStopped               runApproximateJob
accumulableCollection   defaultParallelism      jars                    runJob
accumulator             deployMode              killExecutor            sequenceFile
addFile                 doubleAccumulator       killExecutors           setCallSite
addJar                  emptyRDD                killTaskAttempt         setCheckpointDir
addSparkListener        files                   listFiles               setJobDescription
appName                 getAllPools             listJars                setJobGroup
applicationAttemptId    getCheckpointDir        longAccumulator         setLocalProperty
applicationId           getConf                 makeRDD                 setLogLevel
binaryFiles             getExecutorMemoryStatus master                  sparkUser
binaryRecords           getLocalProperty        newAPIHadoopFile        startTime
broadcast               getPersistentRDDs       newAPIHadoopRDD         statusTracker
cancelAllJobs           getPoolForName          objectFile              stop
cancelJob               getRDDStorageInfo       parallelize             submitJob
cancelJobGroup          getSchedulingMode       range                   textFile
cancelStage             hadoopConfiguration     register                uiWebUrl
clearCallSite           hadoopFile              removeSparkListener     union
clearJobGroup           hadoopRDD               requestExecutors        version
collectionAccumulator   isLocal                 requestTotalExecutors   wholeTextFiles

scala> sc.version
res0: String = 2.4.0-cdh6.3.4

scala> val myConstant: Int = 2437
myConstant: Int = 2437

scala> myConstant
res1: Int = 2437

scala> myConstant.
!=   <=          ceil        isInfinity      isWhole        toBinaryString   toOctalString
%    ==          compare     isNaN           longValue      toByte           toRadians
&    >           compareTo   isNegInfinity   max            toChar           toShort
*    >=          doubleValue isPosInfinity   min            toDegrees        unary_+
+    >>          floatValue  isValidByte     round          toDouble         unary_-
-    >>>         floor       isValidChar     self           toFloat          unary_~
/    ^           getClass    isValidInt      shortValue     toHexString      underlying
<    abs         intValue    isValidLong     signum         toInt            until
<<   byteValue   isInfinite  isValidShort    to             toLong           |
```

```
scala> myConstant.to
to                toByte     toDegrees   toFloat      toInt    toOctalString   toShort
toBinaryString    toChar     toDouble    toHexString  toLong   toRadians       toString

scala> myConstant.toFloat
res2: Float = 2437.0

scala> myConstant
res3: Int = 2437

scala> myConstant.toFloat.toInt
res4: Int = 2437

scala> val myString = myConstant.toString
myString: String = 2437

scala> :type myString
String
```

```
scala> :q
[xm2074@hlog-1 ~]$
```