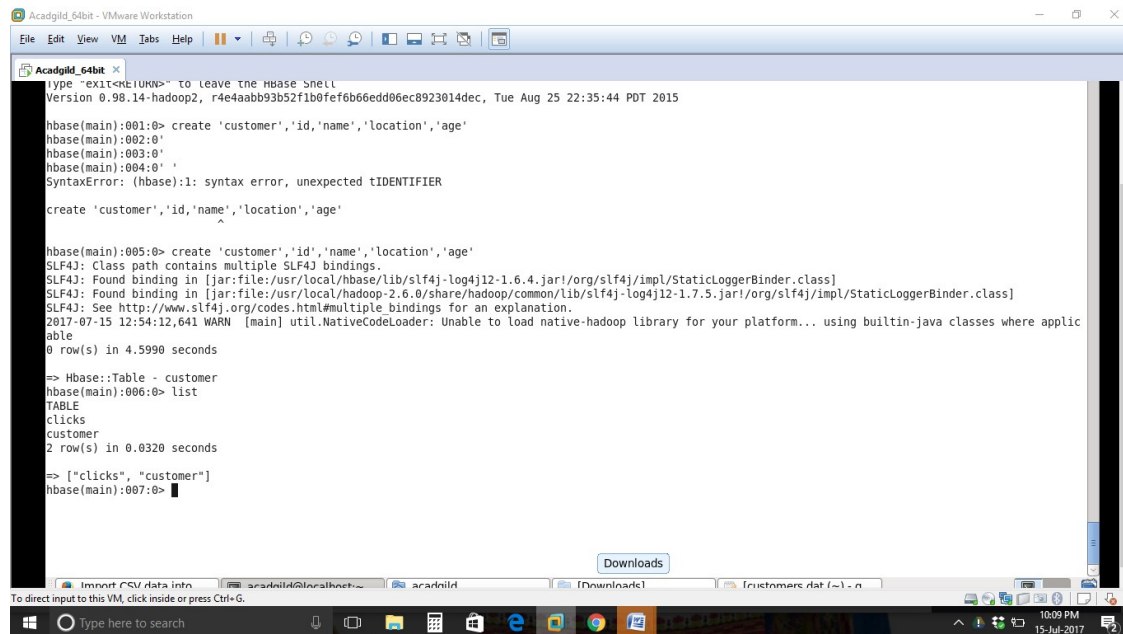


Assignment 18.1

Problem Statement

Create a HBase table 'customer' with one column family 'details'. Write a shells script that loads the content of the file customers.dat in the HBase table. The structure of file looks like id, name, location, age

Customer table created in hbase:



```
Acadgild_64bit - VMware Workstation
File Edit View VM Tabs Help
Acadgild_64bit x
Type "exit" or "Ctrl+D" to leave the HBase shell
Version 0.98.14-hadoop2, r4e4aabb93b52f1b0fef6b66edd06ec8923014dec, Tue Aug 25 22:35:44 PDT 2015

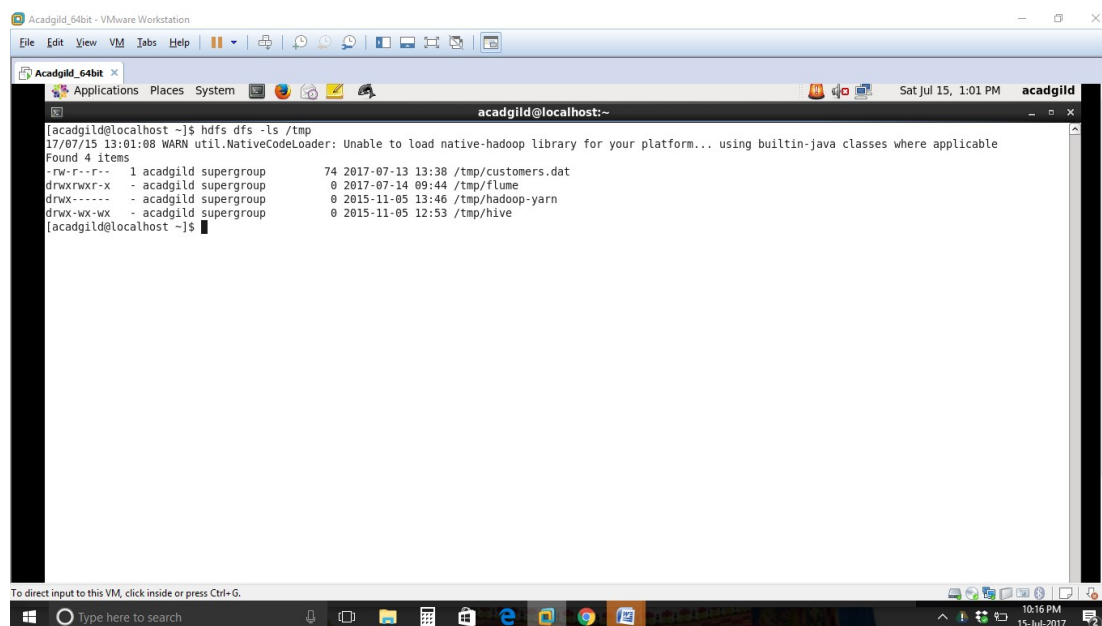
hbase(main):001:0> create 'customer','id','name','location','age'
hbase(main):002:0>
hbase(main):003:0>
hbase(main):004:0> '
SyntaxError: (hbase):1: syntax error, unexpected TIDENTIFIER

create 'customer','id','name','location','age'
^
hbase(main):005:0> create 'customer','id','name','location','age'
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.6.4.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hadoop-2.6.0/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
2017-07-15 12:54:12,641 WARN [main] util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
0 row(s) in 4.5990 seconds

=> Hbase:Table - customer
hbase(main):006:0> list
TABLE
clicks
customer
2 row(s) in 0.0320 seconds

=> ["clicks", "customer"]
hbase(main):007:0>
```

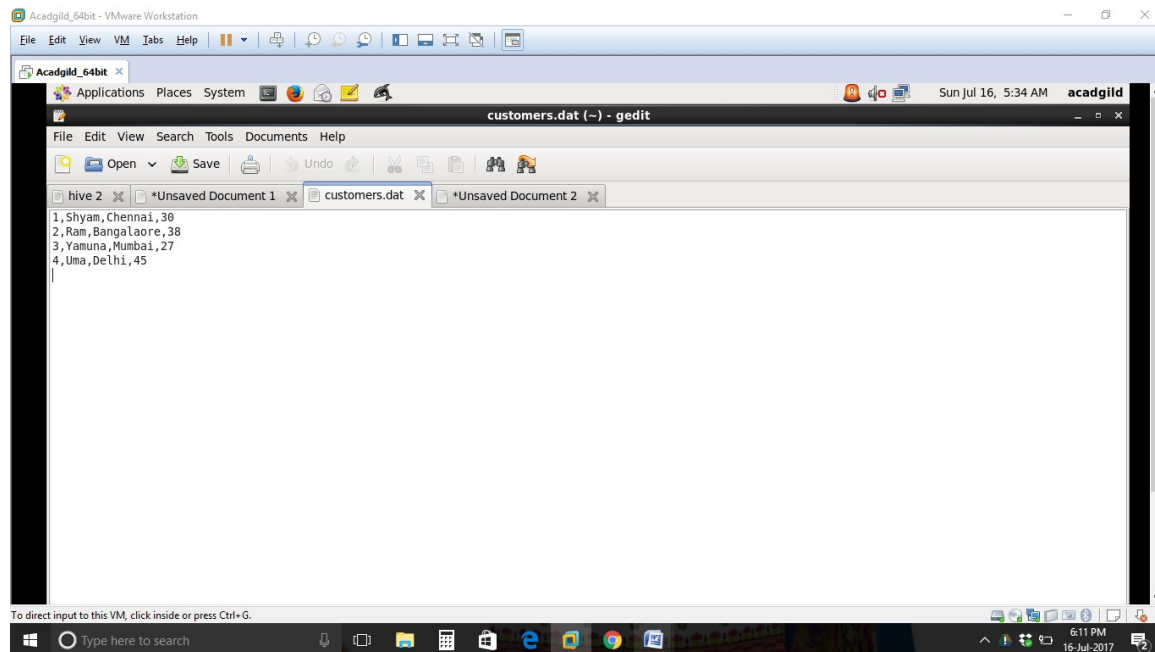
Customers.dat in hdfs



```
Acadgild_64bit - VMware Workstation
File Edit View VM Tabs Help
Acadgild_64bit x
Applications Places System
acacgild@localhost:~
Sat Jul 15, 1:01 PM acacgild

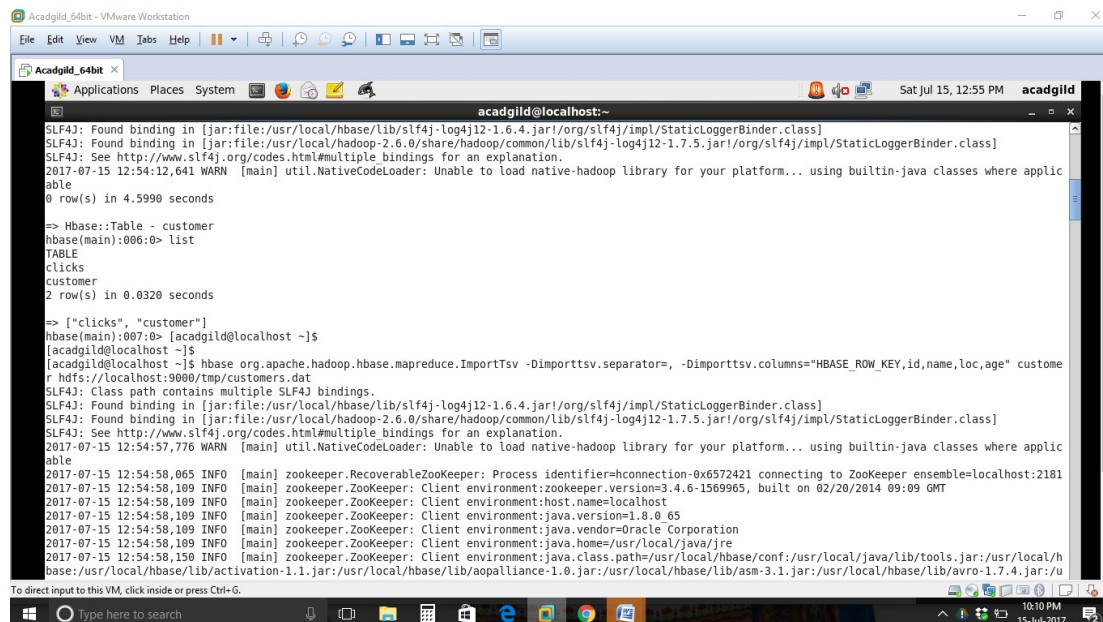
[acacgild@localhost ~]$ hdfs dfs -ls /tmp
17/07/15 13:01:08 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 4 items
-rw-r--r-- 1 acacgild supergroup 74 2017-07-13 13:38 /tmp/customers.dat
drwxrwxr-x - acacgild supergroup 0 2017-07-14 09:44 /tmp/flume
drwx----- - acacgild supergroup 0 2015-11-05 13:46 /tmp/hadoop-yarn
drwx-wx-wx - acacgild supergroup 0 2015-11-05 12:53 /tmp/hive
[acacgild@localhost ~]$
```

Contents of customers.dat file

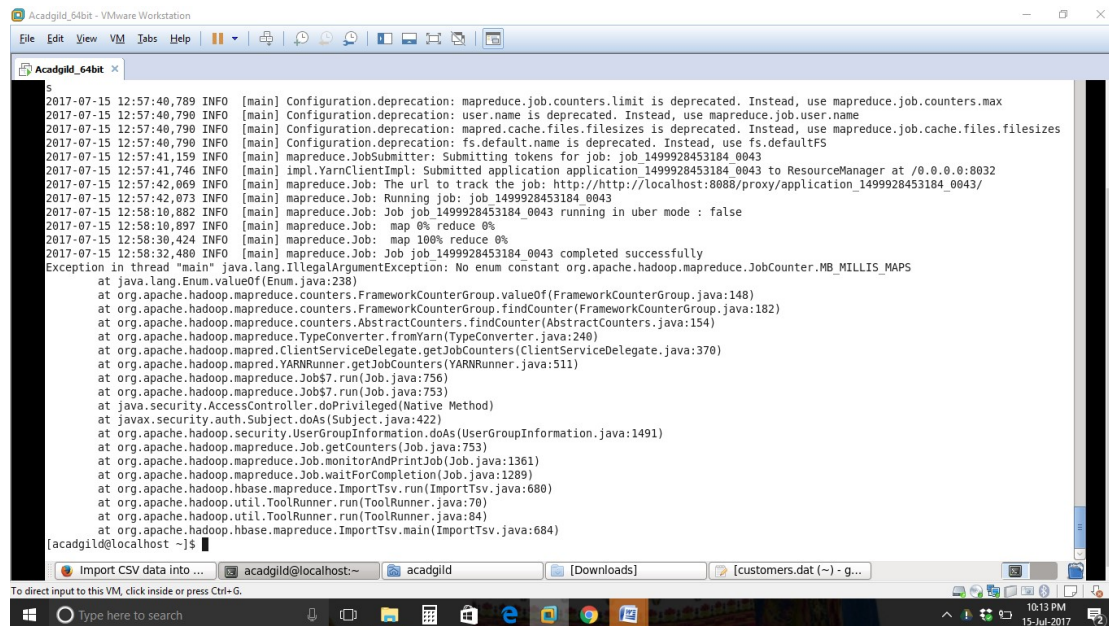


Shell script to load data from customer.dat file to hbase table 'customer'

Hbase org.apache.hadoop.hbase.mapreduce.ImportTsv -Dimporttsv.separator=, -
Dimporttsv.columns="HBASE_ROW_KEY,id,name,loc,age" customer
hdfs://localhost:9000/tmp/customers.dat

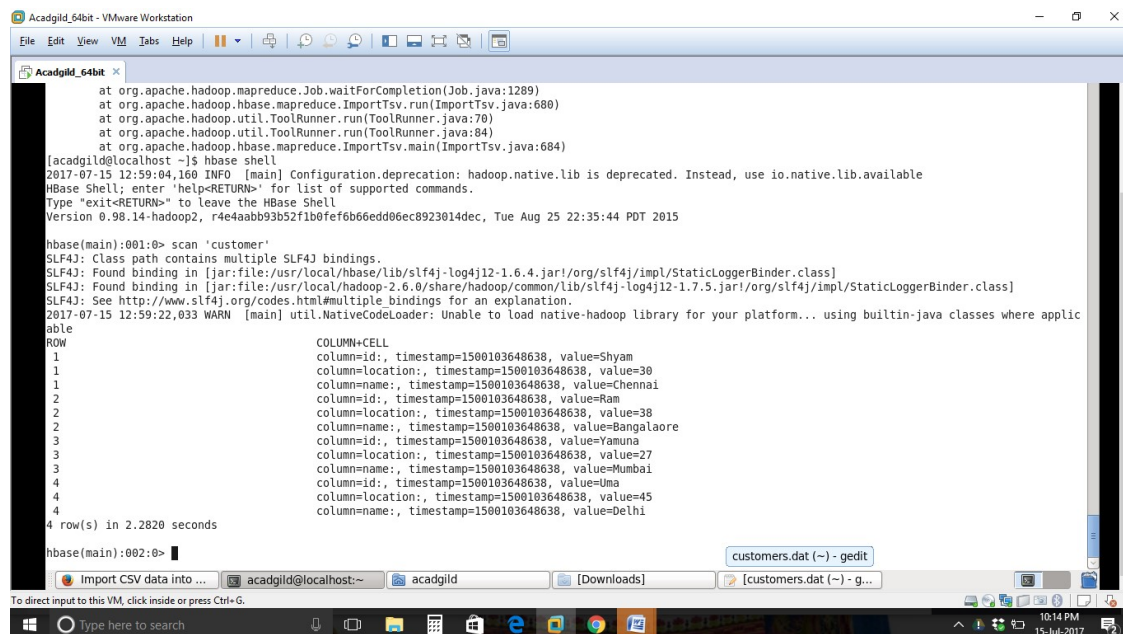


Data copied successfully



```
2017-07-15 12:57:40,789 INFO [main] Configuration.deprecation: mapreduce.job.counters.limit is deprecated. Instead, use mapreduce.job.counters.max
2017-07-15 12:57:40,790 INFO [main] Configuration.deprecation: user.name is deprecated. Instead, use mapreduce.job.user.name
2017-07-15 12:57:40,790 INFO [main] Configuration.deprecation: mapred.cache.files.filesizes is deprecated. Instead, use mapreduce.job.cache.files.filesizes
2017-07-15 12:57:40,790 INFO [main] Configuration.deprecation: fs.default.name is deprecated. Instead, use fs.defaultFS
2017-07-15 12:57:41,159 INFO [main] mapreduce.JobSubmitter: Submitting tokens for job: job_1499928453184_0043
2017-07-15 12:57:41,746 INFO [main] impl.YarnClientImpl: Submitted application application_1499928453184_0043 to ResourceManager at /0.0.0.0:8032
2017-07-15 12:57:42,069 INFO [main] mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1499928453184_0043/
2017-07-15 12:57:42,073 INFO [main] mapreduce.Job: Running job: job_1499928453184_0043
2017-07-15 12:58:10,882 INFO [main] mapreduce.Job: Job job_1499928453184_0043 running in uber mode : false
2017-07-15 12:58:10,897 INFO [main] mapreduce.Job: map 0% reduce 0%
2017-07-15 12:58:30,424 INFO [main] mapreduce.Job: map 100% reduce 0%
2017-07-15 12:58:32,480 INFO [main] mapreduce.Job: Job job_1499928453184_0043 completed successfully
Exception in thread "main" java.lang.IllegalArgumentException: No enum constant org.apache.hadoop.mapreduce.JobCounter.MB_MILLIS_MAPS
    at java.lang.Enum.valueOf(Enum.java:238)
    at org.apache.hadoop.mapreduce.counters.FrameworkCounterGroup.valueOf(FrameworkCounterGroup.java:148)
    at org.apache.hadoop.mapreduce.counters.FrameworkCounterGroup.findCounter(FrameworkCounterGroup.java:182)
    at org.apache.hadoop.mapreduce.counters.AbstractCounters.findCounter(AbstractCounters.java:154)
    at org.apache.hadoop.mapreduce.TypeConverter.fromYarn(TypeConverter.java:240)
    at org.apache.hadoop.mapred.ClientServiceDelegate.getJobCounters(ClientServiceDelegate.java:370)
    at org.apache.hadoop.mapred.YARNRunner.getJobCounters(YARNRunner.java:511)
    at org.apache.hadoop.mapreduce.JobS7.run(Job.java:756)
    at org.apache.hadoop.mapreduce.JobS7.run(Job.java:753)
    at java.security.AccessController.doPrivileged(Native Method)
    at javax.security.auth.Subject.doAs(Subject.java:422)
    at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1491)
    at org.apache.hadoop.mapreduce.Job.getCounters(Job.java:753)
    at org.apache.hadoop.mapreduce.Job.monitorAndPrintJob(Job.java:1361)
    at org.apache.hadoop.mapreduce.Job.waitForCompletion(Job.java:1289)
    at org.apache.hadoop.hbase.mapreduce.ImportTsv.run(ImportTsv.java:680)
    at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:70)
    at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:84)
    at org.apache.hadoop.hbase.mapreduce.ImportTsv.main(ImportTsv.java:684)
[acadgild@localhost ~]$
```

Scan 'customer' in hbase shows data successfully loaded into habse table



```
[acadgild@localhost ~]$ hbase shell
2017-07-15 12:59:04,160 INFO [main] Configuration.deprecation: hadoop.native.lib is deprecated. Instead, use io.native.lib.available
HBase Shell; enter 'help<RETURN>' for list of supported commands.
Type "exit<RETURN>" to leave the HBase Shell
Version 0.98.14-hadoop2, r4e4aabb93b52f1b0fef6b66edd06ec8923014dec, Tue Aug 25 22:35:44 PDT 2015

hbase(main):001:0> scan 'customer'
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.6.4.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hadoop-2.6.0/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
2017-07-15 12:59:22,033 WARN [main] util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applic
able
COLUMN=CELL
ROW
1
column=id:, timestamp=1500103648638, value=Shyam
column=location:, timestamp=1500103648638, value=30
1
column=name:, timestamp=1500103648638, value=Chennai
2
column=id:, timestamp=1500103648638, value=Ram
2
column=location:, timestamp=1500103648638, value=38
2
column=name:, timestamp=1500103648638, value=Bangalore
3
column=id:, timestamp=1500103648638, value=Yamuna
3
column=location:, timestamp=1500103648638, value=27
3
column=name:, timestamp=1500103648638, value=Mumbai
4
column=id:, timestamp=1500103648638, value=Uma
4
column=location:, timestamp=1500103648638, value=45
4
column=name:, timestamp=1500103648638, value=Delhi
4 row(s) in 2.2820 seconds
hbase(main):002:0>
```