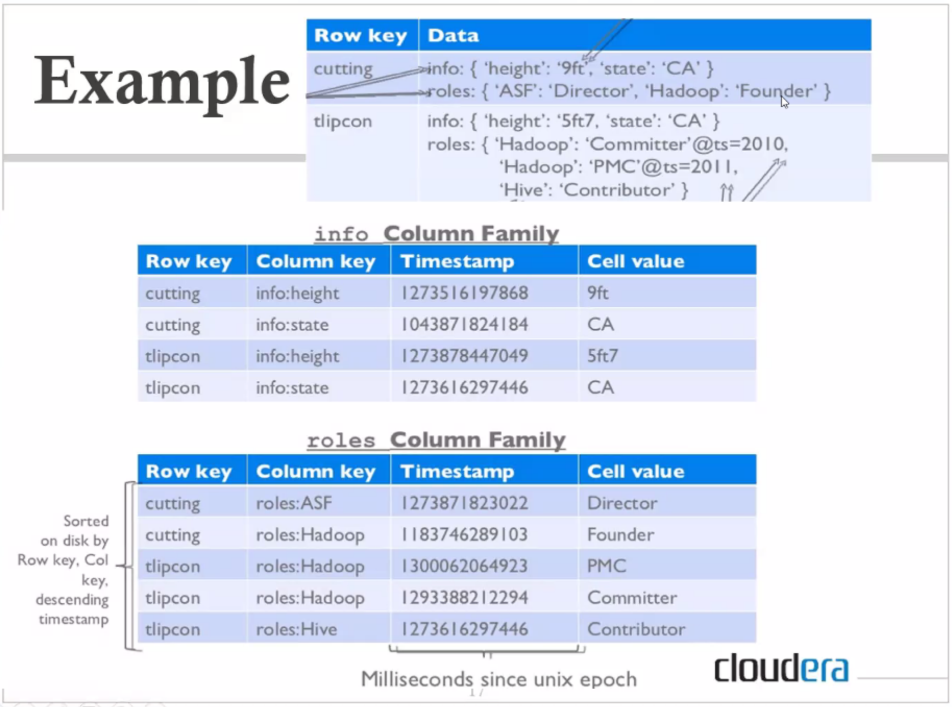
→



Example

Row key	Data

cutting	info: { 'height': '9ft', 'state': 'CA' }
	roles: { 'ASF': 'Director', 'Hadoop': 'Founder' }

tlipcon	info: { 'height': '5ft7, 'state': 'CA' }
	roles: { 'Hadoop': 'Committer'@ts=2010,
	'Hadoop': 'PMC'@ts=2011,
	'Hive': 'Contributor' }

info Column Family

Row key	Column key	Timestamp	Cell value
cutting	info:height	1273516197868	9ft
cutting	info:state	1043871824184	CA
tlipcon	info:height	1273878447049	5ft7
tlipcon	info:state	1273616297446	CA

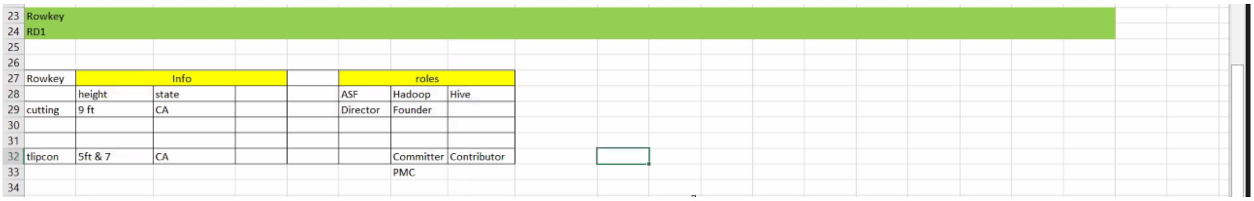
roles Column Family

Row key	Column key	Timestamp	Cell value
cutting	roles:ASF	1273871823022	Director
cutting	roles:Hadoop	1183746289103	Founder
tlipcon	roles:Hadoop	1300062064923	PMC
tlipcon	roles:Hadoop	1293388212294	Committer
tlipcon	roles:Hive	1273616297446	Contributor

Sorted on disk by Row key, Col key, descending timestamp

Milliseconds since unix epoch

cloudera



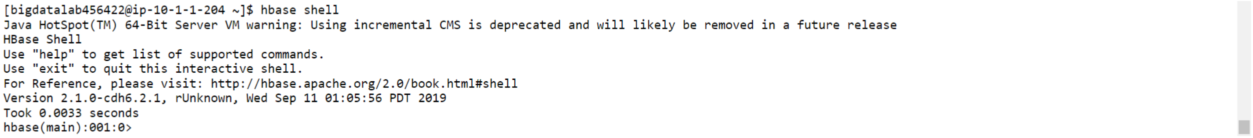
Rowkey

RD1

Rowkey	Info				roles		
	height	state			ASF	Hadoop	Hive
cutting	9 ft	CA			Director	Founder	
tlipcon	5ft & 7	CA				Committer	Contributor
						PMC	

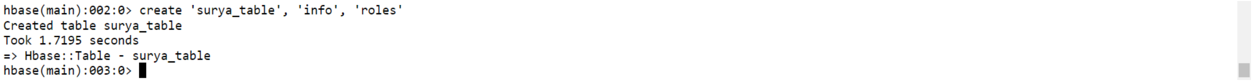
TO create table
CREATE 'tablename', 'cf1', 'cf2'

```
[bigdatalab456422@ip-10-1-1-204 ~]$ hbaseshell
```



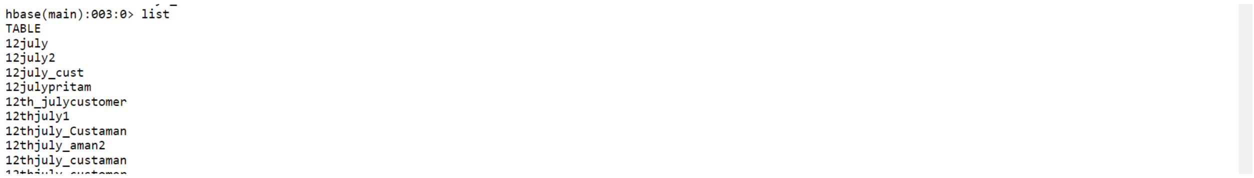
```
[bigdatalab456422@ip-10-1-1-204 ~]$ hbase shell
Java HotSpot(TM) 64-Bit Server VM warning: Using incremental CMS is deprecated and will likely be removed in a future release
HBase Shell
Use "help" to get list of supported commands.
Use "exit" to quit this interactive shell.
For Reference, please visit: http://hbase.apache.org/2.0/book.html#shell
Version 2.1.0-cdh6.2.1, rUnknown, Wed Sep 11 01:05:56 PDT 2019
Took 0.0033 seconds
hbase(main):001:0>
```

```
hbase(main):002:0> create 'surya_table', 'info', 'roles'
```

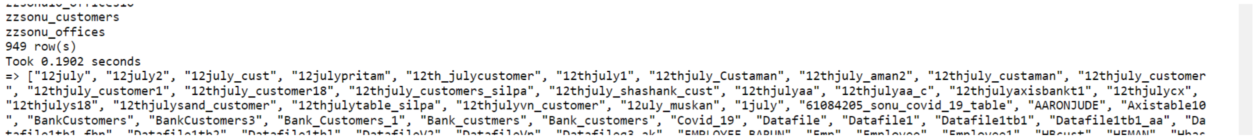


```
hbase(main):002:0> create 'surya_table', 'info', 'roles'
Created table surya_table
Took 1.7195 seconds
=> Hbase::Table - surya_table
hbase(main):003:0>
```

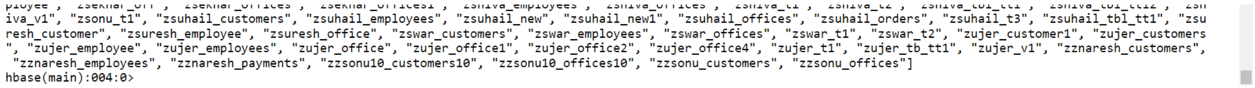
```
hbase(main):003:0> list
```



```
hbase(main):003:0> list
TABLE
12july
12july2
12july_cust
12julypritam
12th_julycustomer
12thjuly1
12thjuly_Custaman
12thjuly_aman2
12thjuly_custaman
```

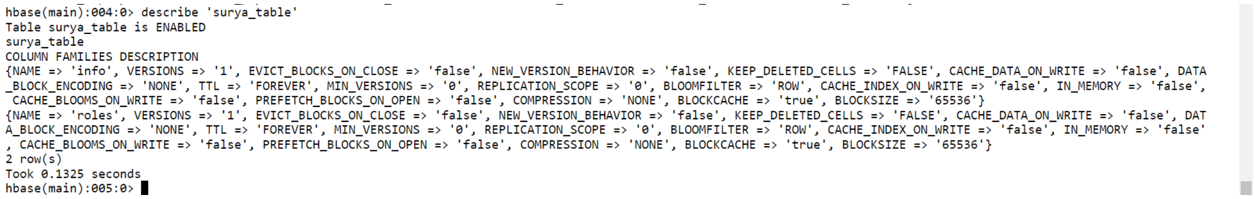


```
zzsonu_customers
zzsonu_offices
949 row(s)
Took 0.1902 seconds
=> ["12july", "12july2", "12july_cust", "12julypritam", "12th_julycustomer", "12thjuly1", "12thjuly_Custaman", "12thjuly_aman2", "12thjuly_custaman", "12thjuly_customer
", "12thjuly_customer1", "12thjuly_customer18", "12thjuly_customers_silpa", "12thjuly_shashank_cust", "12thjulyaa", "12thjulyaa_c", "12thjulyaxisbankt1", "12thjulycx",
"12thjulys18", "12thjulysand_customer", "12thjulytable_silpa", "12thjulyvn_customer", "12uly_muskan", "1july", "61084205_sonu_covid_19_table", "AARONJUDE", "Axistable10
", "BankCustomers", "BankCustomers3", "Bank_Customers_1", "Bank_custmers", "Bank_customers", "Covid_19", "Datafile", "Datafile1", "Datafile1tb1", "Datafile1tb1_aa", "Da
```



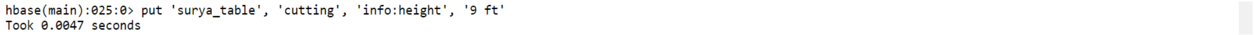
```
iva_v1", "zsonu_t1", "zsuhail_customers", "zsuhail_employees", "zsuhail_new", "zsuhail_new1", "zsuhail_offices", "zsuhail_orders", "zsuhail_t3", "zsuhail_tbl_tt1", "zsu
resh_customer", "zsuresh_employee", "zsuresh_office", "zswar_customers", "zswar_employees", "zswar_offices", "zswar_t1", "zswar_t2", "zujer_customer1", "zujer_customers
", "zujer_employee", "zujer_employees", "zujer_office", "zujer_office1", "zujer_office2", "zujer_office4", "zujer_t1", "zujer_tb_tt1", "zujer_v1", "zznaresh_customers",
 "zznaresh_employees", "zznaresh_payments", "zzsonu10_customers10", "zzsonu10_offices10", "zzsonu_customers", "zzsonu_offices"]
hbase(main):004:0>
```

```
hbase(main):004:0> describe 'surya_table'
```



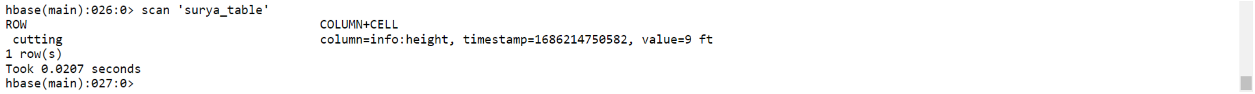
```
hbase(main):004:0> describe 'surya_table'
Table surya_table is ENABLED
surya_table
COLUMN FAMILIES DESCRIPTION
{NAME => 'info', VERSIONS => '1', EVICT_BLOCKS_ON_CLOSE => 'false', NEW_VERSION_BEHAVIOR => 'false', KEEP_DELETED_CELLS => 'FALSE', CACHE_DATA_ON_WRITE => 'false', DATA
_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', MIN_VERSIONS => '0', REPLICATION_SCOPE => '0', BLOOMFILTER => 'ROW', CACHE_INDEX_ON_WRITE => 'false', IN_MEMORY => 'false',
 CACHE_BLOOMS_ON_WRITE => 'false', PREFETCH_BLOCKS_ON_OPEN => 'false', COMPRESSION => 'NONE', BLOCKCACHE => 'true', BLOCKSIZE => '65536'}
{NAME => 'roles', VERSIONS => '1', EVICT_BLOCKS_ON_CLOSE => 'false', NEW_VERSION_BEHAVIOR => 'false', KEEP_DELETED_CELLS => 'FALSE', CACHE_DATA_ON_WRITE => 'false', DAT
A_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', MIN_VERSIONS => '0', REPLICATION_SCOPE => '0', BLOOMFILTER => 'ROW', CACHE_INDEX_ON_WRITE => 'false', IN_MEMORY => 'false'
, CACHE_BLOOMS_ON_WRITE => 'false', PREFETCH_BLOCKS_ON_OPEN => 'false', COMPRESSION => 'NONE', BLOCKCACHE => 'true', BLOCKSIZE => '65536'}
2 row(s)
Took 0.1325 seconds
hbase(main):005:0>
```

```
hbase(main):025:0> put 'surya_table', 'cutting', 'info:height', '9
ft'
```



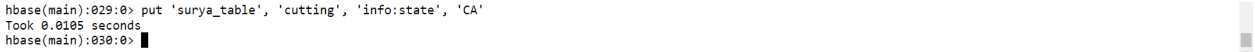
```
hbase(main):025:0> put 'surya_table', 'cutting', 'info:height', '9 ft'
Took 0.0047 seconds
```

```
hbase(main):026:0> scan 'surya_table'
```



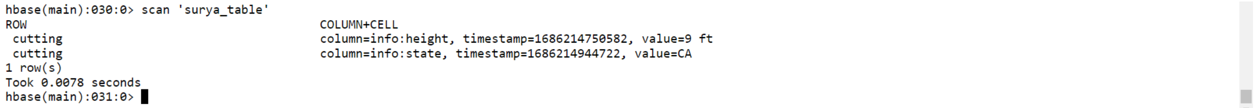
```
hbase(main):026:0> scan 'surya_table'
ROW                                      COLUMN+CELL
 cutting                                 column=info:height, timestamp=1686214750582, value=9 ft
1 row(s)
Took 0.0207 seconds
hbase(main):027:0>
```

```
hbase(main):029:0> put 'surya_table', 'cutting', 'info:state', 'CA'
```



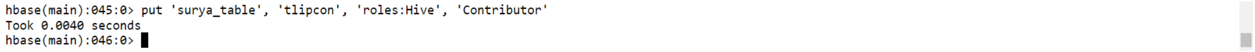
```
hbase(main):029:0> put 'surya_table', 'cutting', 'info:state', 'CA'
Took 0.0105 seconds
hbase(main):030:0>
```

```
hbase(main):030:0> scan 'surya_table'
```



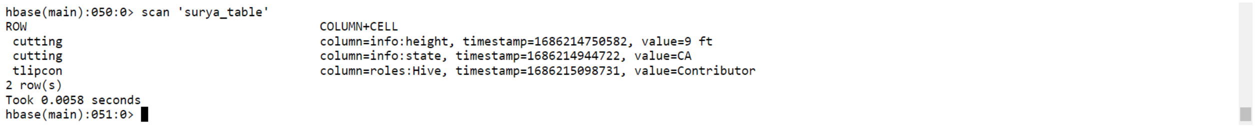
```
hbase(main):030:0> scan 'surya_table'
ROW                                      COLUMN+CELL
 cutting                                 column=info:height, timestamp=1686214750582, value=9 ft
 cutting                                 column=info:state, timestamp=1686214944722, value=CA
1 row(s)
Took 0.0078 seconds
hbase(main):031:0>
```

```
hbase(main):045:0> put 'surya_table', 'tlipcon', 'roles:Hive',
'Contributor'
```



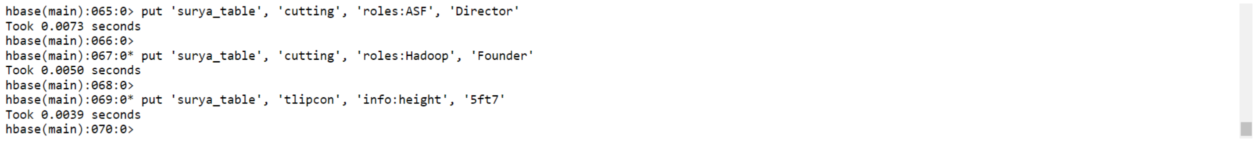
```
hbase(main):045:0> put 'surya_table', 'tlipcon', 'roles:Hive', 'Contributor'
Took 0.0040 seconds
hbase(main):046:0>
```

```
hbase(main):050:0> scan 'surya_table'
```



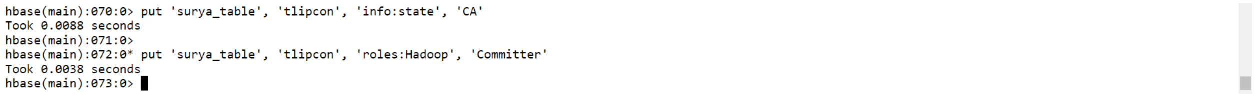
```
hbase(main):050:0> scan 'surya_table'
ROW                                         COLUMN+CELL
 cutting                                    column=info:height, timestamp=1686214750582, value=9 ft
 cutting                                    column=info:state, timestamp=1686214944722, value=CA
 tlipcon                                    column=roles:Hive, timestamp=1686215098731, value=Contributor
2 row(s)
Took 0.0058 seconds
hbase(main):051:0>
```

hbase(main):065:0> put 'surya_table', 'cutting', 'roles:ASF', 'Director'
hbase(main):067:0* put 'surya_table', 'cutting', 'roles:Hadoop', 'Founder'
hbase(main):069:0* put 'surya_table', 'tlipcon', 'info:height', '5ft7'



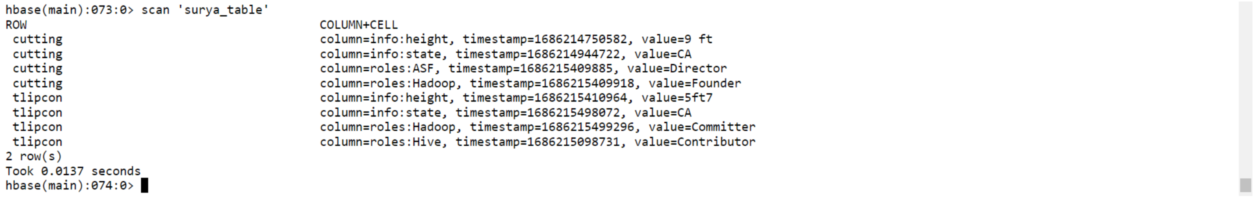
```
hbase(main):065:0> put 'surya_table', 'cutting', 'roles:ASF', 'Director'
Took 0.0073 seconds
hbase(main):066:0>
hbase(main):067:0* put 'surya_table', 'cutting', 'roles:Hadoop', 'Founder'
Took 0.0050 seconds
hbase(main):068:0>
hbase(main):069:0* put 'surya_table', 'tlipcon', 'info:height', '5ft7'
Took 0.0039 seconds
hbase(main):070:0>
```

hbase(main):070:0> put 'surya_table', 'tlipcon', 'info:state', 'CA'
hbase(main):072:0* put 'surya_table', 'tlipcon', 'roles:Hadoop', 'Committer'



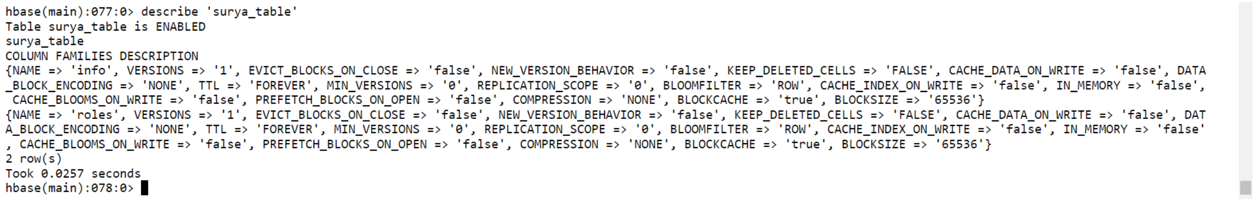
```
hbase(main):070:0> put 'surya_table', 'tlipcon', 'info:state', 'CA'
Took 0.0088 seconds
hbase(main):071:0>
hbase(main):072:0* put 'surya_table', 'tlipcon', 'roles:Hadoop', 'Committer'
Took 0.0038 seconds
hbase(main):073:0>
```

hbase(main):073:0> scan 'surya_table'



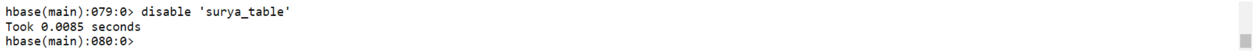
```
hbase(main):073:0> scan 'surya_table'
ROW                                         COLUMN+CELL
 cutting                                    column=info:height, timestamp=1686214750582, value=9 ft
 cutting                                    column=info:state, timestamp=1686214944722, value=CA
 cutting                                    column=roles:ASF, timestamp=1686215409885, value=Director
 cutting                                    column=roles:Hadoop, timestamp=1686215409918, value=Founder
 tlipcon                                    column=info:height, timestamp=1686215410964, value=5ft7
 tlipcon                                    column=info:state, timestamp=1686215498072, value=CA
 tlipcon                                    column=roles:Hadoop, timestamp=1686215499296, value=Committer
 tlipcon                                    column=roles:Hive, timestamp=1686215098731, value=Contributor
2 row(s)
Took 0.0137 seconds
hbase(main):074:0>
```

hbase(main):077:0> describe 'surya_table'



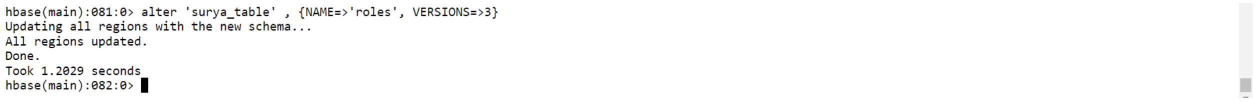
```
hbase(main):077:0> describe 'surya_table'
Table surya_table is ENABLED
surya_table
COLUMN FAMILIES DESCRIPTION
{NAME => 'info', VERSIONS => '1', EVICT_BLOCKS_ON_CLOSE => 'false', NEW_VERSION_BEHAVIOR => 'false', KEEP_DELETED_CELLS => 'FALSE', CACHE_DATA_ON_WRITE => 'false', DATA
_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', MIN_VERSIONS => '0', REPLICATION_SCOPE => '0', BLOOMFILTER => 'ROW', CACHE_INDEX_ON_WRITE => 'false', IN_MEMORY => 'false',
 CACHE_BLOOMS_ON_WRITE => 'false', PREFETCH_BLOCKS_ON_OPEN => 'false', COMPRESSION => 'NONE', BLOCKCACHE => 'true', BLOCKSIZE => '65536'}
{NAME => 'roles', VERSIONS => '1', EVICT_BLOCKS_ON_CLOSE => 'false', NEW_VERSION_BEHAVIOR => 'false', KEEP_DELETED_CELLS => 'FALSE', CACHE_DATA_ON_WRITE => 'false', DAT
A_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', MIN_VERSIONS => '0', REPLICATION_SCOPE => '0', BLOOMFILTER => 'ROW', CACHE_INDEX_ON_WRITE => 'false', IN_MEMORY => 'false'
, CACHE_BLOOMS_ON_WRITE => 'false', PREFETCH_BLOCKS_ON_OPEN => 'false', COMPRESSION => 'NONE', BLOCKCACHE => 'true', BLOCKSIZE => '65536'}
2 row(s)
Took 0.0257 seconds
hbase(main):078:0>
```

hbase(main):079:0> disable 'surya_table'



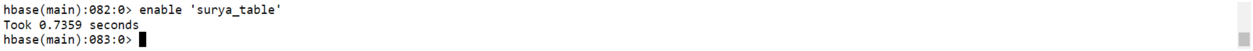
```
hbase(main):079:0> disable 'surya_table'
Took 0.0085 seconds
hbase(main):080:0>
```

hbase(main):081:0> alter 'surya_table' , {NAME=>'roles', VERSIONS=>3}



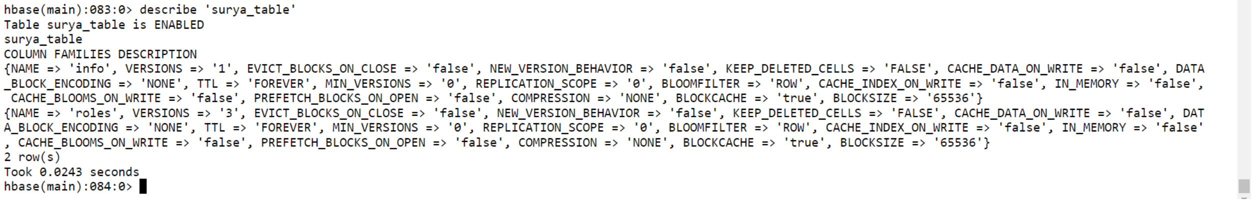
```
hbase(main):081:0> alter 'surya_table' , {NAME=>'roles', VERSIONS=>3}
Updating all regions with the new schema...
All regions updated.
Done.
Took 1.2029 seconds
hbase(main):082:0>
```

hbase(main):082:0> enable 'surya_table'



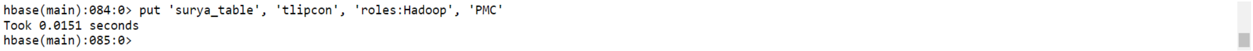
```
hbase(main):082:0> enable 'surya_table'
Took 0.7359 seconds
hbase(main):083:0> █
```

hbase(main):083:0> describe 'surya_table'



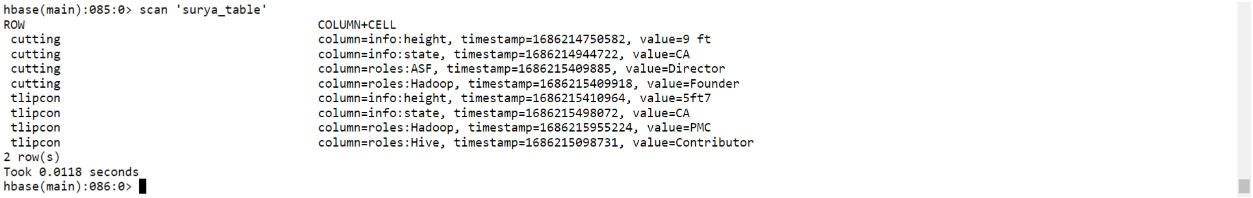
```
hbase(main):083:0> describe 'surya_table'
Table surya_table is ENABLED
surya_table
COLUMN FAMILIES DESCRIPTION
{NAME => 'info', VERSIONS => '1', EVICT_BLOCKS_ON_CLOSE => 'false', NEW_VERSION_BEHAVIOR => 'false', KEEP_DELETED_CELLS => 'FALSE', CACHE_DATA_ON_WRITE => 'false', DATA
_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', MIN_VERSIONS => '0', REPLICATION_SCOPE => '0', BLOOMFILTER => 'ROW', CACHE_INDEX_ON_WRITE => 'false', IN_MEMORY => 'false',
 CACHE_BLOOMS_ON_WRITE => 'false', PREFETCH_BLOCKS_ON_OPEN => 'false', COMPRESSION => 'NONE', BLOCKCACHE => 'true', BLOCKSIZE => '65536'}
{NAME => 'roles', VERSIONS => '3', EVICT_BLOCKS_ON_CLOSE => 'false', NEW_VERSION_BEHAVIOR => 'false', KEEP_DELETED_CELLS => 'FALSE', CACHE_DATA_ON_WRITE => 'false', DAT
A_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', MIN_VERSIONS => '0', REPLICATION_SCOPE => '0', BLOOMFILTER => 'ROW', CACHE_INDEX_ON_WRITE => 'false', IN_MEMORY => 'false'
, CACHE_BLOOMS_ON_WRITE => 'false', PREFETCH_BLOCKS_ON_OPEN => 'false', COMPRESSION => 'NONE', BLOCKCACHE => 'true', BLOCKSIZE => '65536'}
2 row(s)
Took 0.0243 seconds
hbase(main):084:0> █
```

hbase(main):084:0> put 'surya_table', 'tlipcon', 'roles:Hadoop', 'PMC'



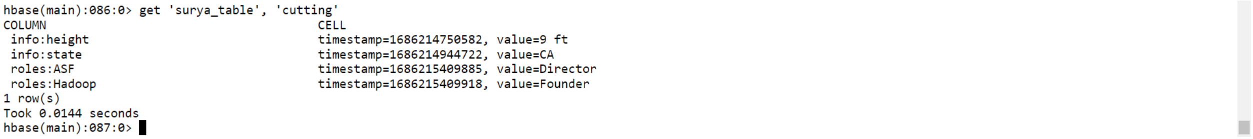
```
hbase(main):084:0> put 'surya_table', 'tlipcon', 'roles:Hadoop', 'PMC'
Took 0.0151 seconds
hbase(main):085:0>
```

hbase(main):085:0> scan 'surya_table'



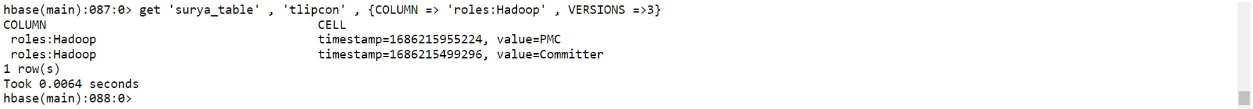
```
hbase(main):085:0> scan 'surya_table'
ROW                                      COLUMN+CELL
 cutting                                 column=info:height, timestamp=1686214750582, value=9 ft
 cutting                                 column=info:state, timestamp=1686214944722, value=CA
 cutting                                 column=roles:ASF, timestamp=1686215409885, value=Director
 cutting                                 column=roles:Hadoop, timestamp=1686215409918, value=Founder
 tlipcon                                 column=info:height, timestamp=1686215410964, value=5ft7
 tlipcon                                 column=info:state, timestamp=1686215498072, value=CA
 tlipcon                                 column=roles:Hadoop, timestamp=1686215955224, value=PMC
 tlipcon                                 column=roles:Hive, timestamp=1686215098731, value=Contributor
2 row(s)
Took 0.0118 seconds
hbase(main):086:0> █
```

hbase(main):086:0> get 'surya_table', 'cutting'



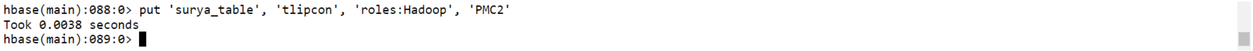
```
hbase(main):086:0> get 'surya_table', 'cutting'
COLUMN                                   CELL
 info:height                             timestamp=1686214750582, value=9 ft
 info:state                              timestamp=1686214944722, value=CA
 roles:ASF                               timestamp=1686215409885, value=Director
 roles:Hadoop                            timestamp=1686215409918, value=Founder
1 row(s)
Took 0.0144 seconds
hbase(main):087:0> █
```

hbase(main):087:0> get 'surya_table' , 'tlipcon' , {COLUMN => 'roles:Hadoop' , VERSIONS =>3}



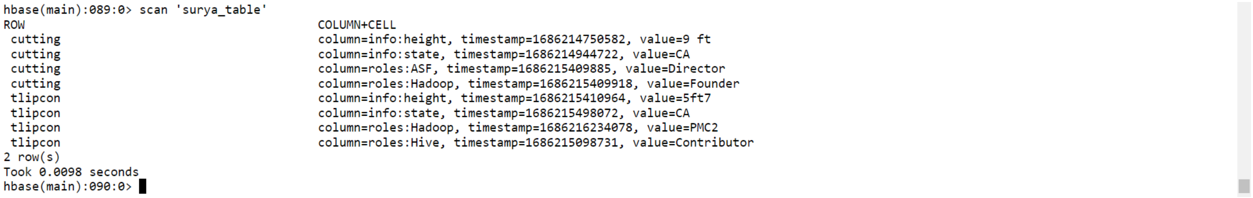
```
hbase(main):087:0> get 'surya_table' , 'tlipcon' , {COLUMN => 'roles:Hadoop' , VERSIONS =>3}
COLUMN                                   CELL
 roles:Hadoop                            timestamp=1686215955224, value=PMC
 roles:Hadoop                            timestamp=1686215499296, value=Committer
1 row(s)
Took 0.0064 seconds
hbase(main):088:0>
```

hbase(main):088:0> put 'surya_table', 'tlipcon', 'roles:Hadoop', 'PMC2'



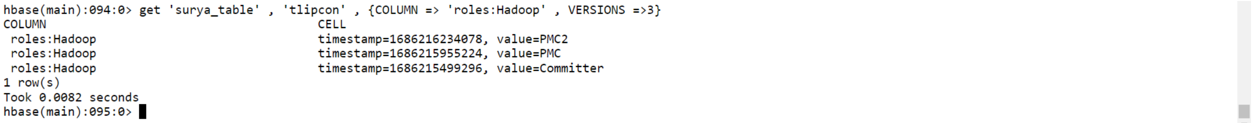
```
hbase(main):088:0> put 'surya_table', 'tlipcon', 'roles:Hadoop', 'PMC2'
Took 0.0038 seconds
hbase(main):089:0> █
```

```
scan 'surya_table'
```



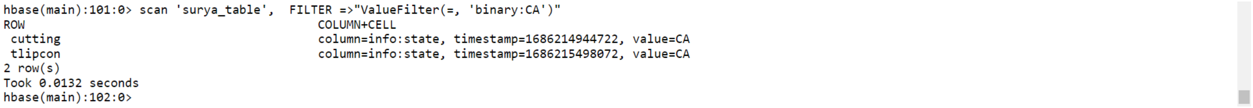
```
hbase(main):089:0> scan 'surya_table'
ROW                                          COLUMN+CELL
 cutting                                     column=info:height, timestamp=1686214750582, value=9 ft
 cutting                                     column=info:state, timestamp=1686214944722, value=CA
 cutting                                     column=roles:ASF, timestamp=1686215409885, value=Director
 cutting                                     column=roles:Hadoop, timestamp=1686215409918, value=Founder
 tlipcon                                     column=info:height, timestamp=1686215410964, value=5ft7
 tlipcon                                     column=info:state, timestamp=1686215498072, value=CA
 tlipcon                                     column=roles:Hadoop, timestamp=1686216234078, value=PMC2
 tlipcon                                     column=roles:Hive, timestamp=1686215098731, value=Contributor
2 row(s)
Took 0.0098 seconds
hbase(main):090:0> 
```

```
get 'surya_table' , 'tlipcon' , {COLUMN => 'roles:Hadoop' , VERSIONS
=>3}
```



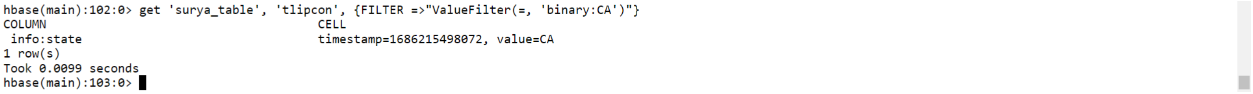
```
hbase(main):094:0> get 'surya_table' , 'tlipcon' , {COLUMN => 'roles:Hadoop' , VERSIONS =>3}
COLUMN                                       CELL
 roles:Hadoop                                timestamp=1686216234078, value=PMC2
 roles:Hadoop                                timestamp=1686215955224, value=PMC
 roles:Hadoop                                timestamp=1686215499296, value=Committer
1 row(s)
Took 0.0082 seconds
hbase(main):095:0> 
```

```
hbase(main):101:0> scan 'surya_table', FILTER =>"ValueFilter(=,
'binary:CA')"
```



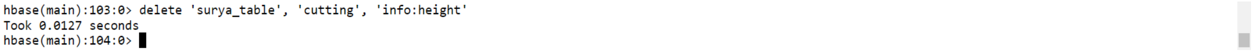
```
hbase(main):101:0> scan 'surya_table',  FILTER =>"ValueFilter(=, 'binary:CA')"
ROW                                          COLUMN+CELL
 cutting                                     column=info:state, timestamp=1686214944722, value=CA
 tlipcon                                     column=info:state, timestamp=1686215498072, value=CA
2 row(s)
Took 0.0132 seconds
hbase(main):102:0>
```

```
hbase(main):102:0> get 'surya_table', 'tlipcon', {FILTER
=>"ValueFilter(=, 'binary:CA')"}
```



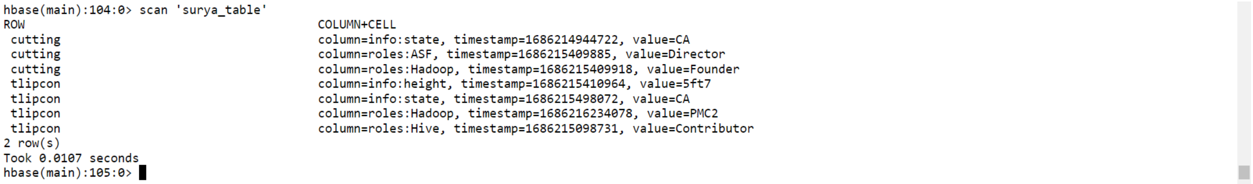
```
hbase(main):102:0> get 'surya_table', 'tlipcon', {FILTER =>"ValueFilter(=, 'binary:CA')"}
COLUMN                                       CELL
 info:state                                  timestamp=1686215498072, value=CA
1 row(s)
Took 0.0099 seconds
hbase(main):103:0> 
```

```
hbase(main):103:0> delete 'surya_table', 'cutting', 'info:height'
```



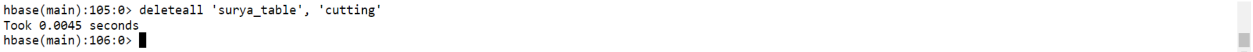
```
hbase(main):103:0> delete 'surya_table', 'cutting', 'info:height'
Took 0.0127 seconds
hbase(main):104:0> 
```

```
hbase(main):104:0> scan 'surya_table'
```



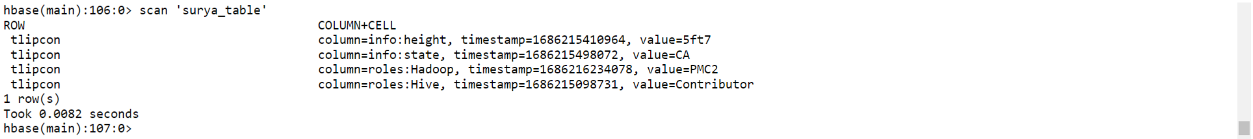
```
hbase(main):104:0> scan 'surya_table'
ROW                                          COLUMN+CELL
 cutting                                     column=info:state, timestamp=1686214944722, value=CA
 cutting                                     column=roles:ASF, timestamp=1686215409885, value=Director
 cutting                                     column=roles:Hadoop, timestamp=1686215409918, value=Founder
 tlipcon                                     column=info:height, timestamp=1686215410964, value=5ft7
 tlipcon                                     column=info:state, timestamp=1686215498072, value=CA
 tlipcon                                     column=roles:Hadoop, timestamp=1686216234078, value=PMC2
 tlipcon                                     column=roles:Hive, timestamp=1686215098731, value=Contributor
2 row(s)
Took 0.0107 seconds
hbase(main):105:0> 
```

```
hbase(main):105:0> deleteall 'surya_table', 'cutting'
```



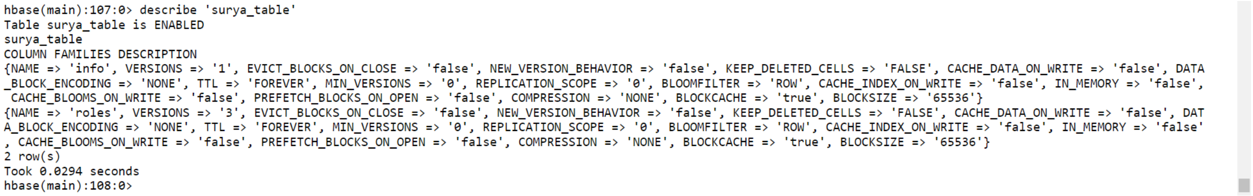
```
hbase(main):105:0> deleteall 'surya_table', 'cutting'
Took 0.0045 seconds
hbase(main):106:0> 
```

```
hbase(main):106:0> scan 'surya_table'
```



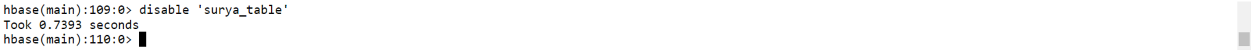
```
hbase(main):106:0> scan 'surya_table'
ROW                                          COLUMN+CELL
 tlipcon                                     column=info:height, timestamp=1686215410964, value=5ft7
 tlipcon                                     column=info:state, timestamp=1686215498072, value=CA
 tlipcon                                     column=roles:Hadoop, timestamp=1686216234078, value=PMC2
 tlipcon                                     column=roles:Hive, timestamp=1686215098731, value=Contributor
1 row(s)
Took 0.0082 seconds
hbase(main):107:0>
```

```
hbase(main):107:0> describe 'surya_table'
```



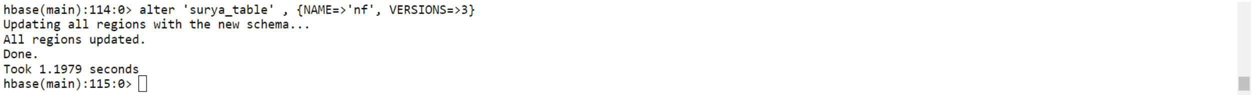
```
hbase(main):107:0> describe 'surya_table'
Table surya_table is ENABLED
surya_table
COLUMN FAMILIES DESCRIPTION
{NAME => 'info', VERSIONS => '1', EVICT_BLOCKS_ON_CLOSE => 'false', NEW_VERSION_BEHAVIOR => 'false', KEEP_DELETED_CELLS => 'FALSE', CACHE_DATA_ON_WRITE => 'false', DATA
_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', MIN_VERSIONS => '0', REPLICATION_SCOPE => '0', BLOOMFILTER => 'ROW', CACHE_INDEX_ON_WRITE => 'false', IN_MEMORY => 'false',
 CACHE_BLOOMS_ON_WRITE => 'false', PREFETCH_BLOCKS_ON_OPEN => 'false', COMPRESSION => 'NONE', BLOCKCACHE => 'true', BLOCKSIZE => '65536'}
{NAME => 'roles', VERSIONS => '3', EVICT_BLOCKS_ON_CLOSE => 'false', NEW_VERSION_BEHAVIOR => 'false', KEEP_DELETED_CELLS => 'FALSE', CACHE_DATA_ON_WRITE => 'false', DAT
A_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', MIN_VERSIONS => '0', REPLICATION_SCOPE => '0', BLOOMFILTER => 'ROW', CACHE_INDEX_ON_WRITE => 'false', IN_MEMORY => 'false'
, CACHE_BLOOMS_ON_WRITE => 'false', PREFETCH_BLOCKS_ON_OPEN => 'false', COMPRESSION => 'NONE', BLOCKCACHE => 'true', BLOCKSIZE => '65536'}
2 row(s)
Took 0.0294 seconds
hbase(main):108:0>
```

```
hbase(main):109:0> disable 'surya_table'
```



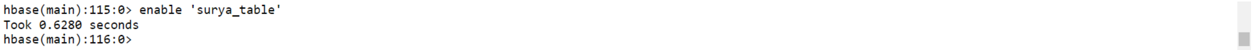
```
hbase(main):109:0> disable 'surya_table'
Took 0.7393 seconds
hbase(main):110:0>
```

```
hbase(main):114:0> alter 'surya_table' , {NAME=>'nf', VERSIONS=>3}
```



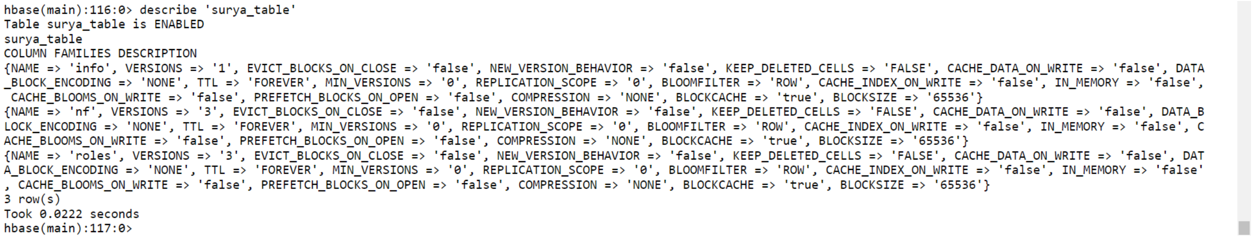
```
hbase(main):114:0> alter 'surya_table' , {NAME=>'nf', VERSIONS=>3}
Updating all regions with the new schema...
All regions updated.
Done.
Took 1.1979 seconds
hbase(main):115:0>
```

```
hbase(main):115:0> enable 'surya_table'
```



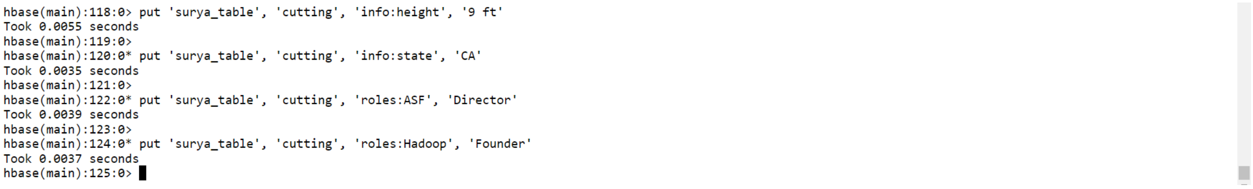
```
hbase(main):115:0> enable 'surya_table'
Took 0.6280 seconds
hbase(main):116:0>
```

```
hbase(main):116:0> describe 'surya_table'
```



```
hbase(main):116:0> describe 'surya_table'
Table surya_table is ENABLED
surya_table
COLUMN FAMILIES DESCRIPTION
{NAME => 'info', VERSIONS => '1', EVICT_BLOCKS_ON_CLOSE => 'false', NEW_VERSION_BEHAVIOR => 'false', KEEP_DELETED_CELLS => 'FALSE', CACHE_DATA_ON_WRITE => 'false', DATA
_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', MIN_VERSIONS => '0', REPLICATION_SCOPE => '0', BLOOMFILTER => 'ROW', CACHE_INDEX_ON_WRITE => 'false', IN_MEMORY => 'false',
 CACHE_BLOOMS_ON_WRITE => 'false', PREFETCH_BLOCKS_ON_OPEN => 'false', COMPRESSION => 'NONE', BLOCKCACHE => 'true', BLOCKSIZE => '65536'}
{NAME => 'nf', VERSIONS => '3', EVICT_BLOCKS_ON_CLOSE => 'false', NEW_VERSION_BEHAVIOR => 'false', KEEP_DELETED_CELLS => 'FALSE', CACHE_DATA_ON_WRITE => 'false', DATA_B
LOCK_ENCODING => 'NONE', TTL => 'FOREVER', MIN_VERSIONS => '0', REPLICATION_SCOPE => '0', BLOOMFILTER => 'ROW', CACHE_INDEX_ON_WRITE => 'false', IN_MEMORY => 'false', C
ACHE_BLOOMS_ON_WRITE => 'false', PREFETCH_BLOCKS_ON_OPEN => 'false', COMPRESSION => 'NONE', BLOCKCACHE => 'true', BLOCKSIZE => '65536'}
{NAME => 'roles', VERSIONS => '3', EVICT_BLOCKS_ON_CLOSE => 'false', NEW_VERSION_BEHAVIOR => 'false', KEEP_DELETED_CELLS => 'FALSE', CACHE_DATA_ON_WRITE => 'false', DAT
A_BLOCK_ENCODING => 'NONE', TTL => 'FOREVER', MIN_VERSIONS => '0', REPLICATION_SCOPE => '0', BLOOMFILTER => 'ROW', CACHE_INDEX_ON_WRITE => 'false', IN_MEMORY => 'false'
, CACHE_BLOOMS_ON_WRITE => 'false', PREFETCH_BLOCKS_ON_OPEN => 'false', COMPRESSION => 'NONE', BLOCKCACHE => 'true', BLOCKSIZE => '65536'}
3 row(s)
Took 0.0222 seconds
hbase(main):117:0>
```

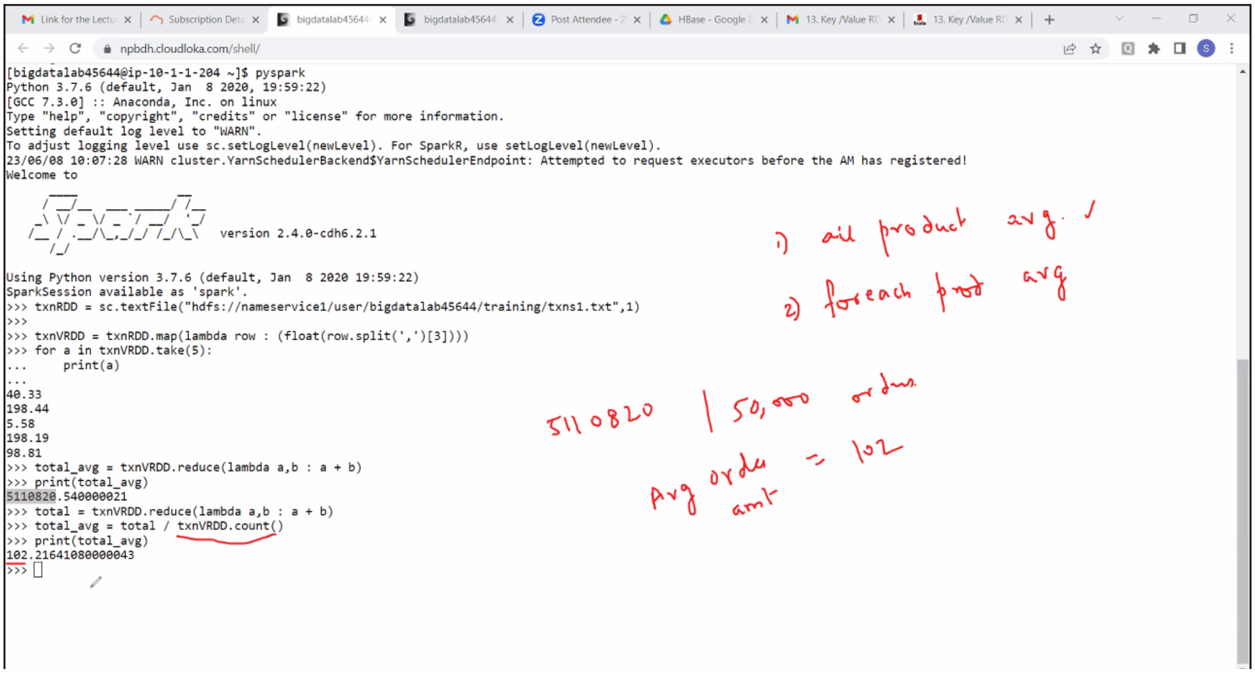
```
hbase(main):118:0> put 'surya_table', 'cutting', 'info:height', '9 ft'
hbase(main):120:0* put 'surya_table', 'cutting', 'info:state', 'CA'
hbase(main):122:0* put 'surya_table', 'cutting', 'roles:ASF', 'Director'
hbase(main):124:0* put 'surya_table', 'cutting', 'roles:Hadoop', 'Founder'
```



```
hbase(main):118:0> put 'surya_table', 'cutting', 'info:height', '9 ft'
Took 0.0055 seconds
hbase(main):119:0>
hbase(main):120:0* put 'surya_table', 'cutting', 'info:state', 'CA'
Took 0.0035 seconds
hbase(main):121:0>
hbase(main):122:0* put 'surya_table', 'cutting', 'roles:ASF', 'Director'
Took 0.0039 seconds
hbase(main):123:0>
hbase(main):124:0* put 'surya_table', 'cutting', 'roles:Hadoop', 'Founder'
Took 0.0037 seconds
hbase(main):125:0>
```

→

→



```
[bigdatalab45644@ip-10-1-1-204 ~]$ pyspark
Python 3.7.6 (default, Jan  8 2020, 19:59:22)
[GCC 7.3.0] :: Anaconda, Inc. on linux
Type "help", "copyright", "credits" or "license" for more information.
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
23/06/08 10:07:28 WARN cluster.YarnSchedulerBackend$YarnSchedulerEndpoint: Attempted to request executors before the AM has registered!
Welcome to
      ____              __
     / __/__  ___ _____/ /__
    _\ \/ _ \/ _ `/ __/  '_/
   /__ / .__/\_,_/_/ /_/\_\   version 2.4.0-cdh6.2.1
      /_/

Using Python version 3.7.6 (default, Jan  8 2020 19:59:22)
SparkSession available as 'spark'.
>>> txnRDD = sc.textFile("hdfs://nameservice1/user/bigdatalab45644/training/txns1.txt",1)
>>>
>>> txnVRDD = txnRDD.map(lambda row : (float(row.split(',')[3])))
>>> for a in txnVRDD.take(5):
...     print(a)
...
40.33
198.44
5.58
198.19
98.81
>>> total_avg = txnVRDD.reduce(lambda a,b : a + b)
>>> print(total_avg)
5110820.540000021
>>> total = txnVRDD.reduce(lambda a,b : a + b)
>>> total_avg = total / txnVRDD.count()
>>> print(total_avg)
102.21641080000043
>>> 
```

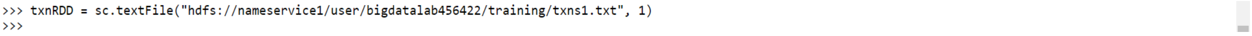
1) all product avg. ✓

2) foreach prod avg

5110820 | 50,000 orders

Avg order amt = 102

>>> txnRDD = sc.textFile("hdfs://nameservice1/user/bigdatalab456422/training/txns1.txt", 1)



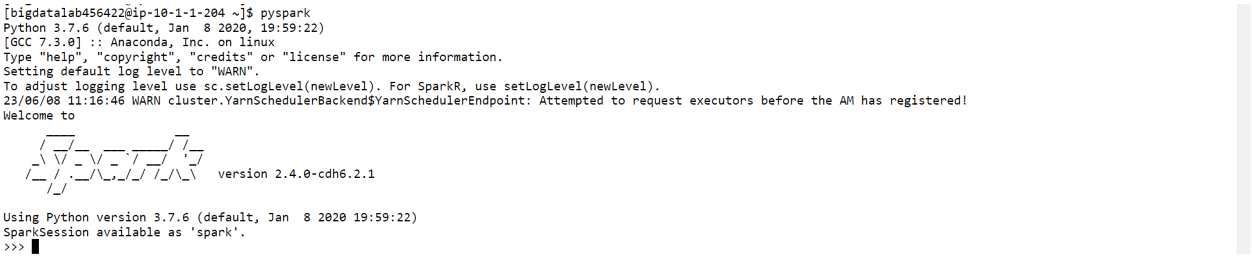
```
>>> txnRDD = sc.textFile("hdfs://nameservice1/user/bigdatalab456422/training/txns1.txt", 1)
>>>
```

>>> txnVRDD = txnRDD.map(lambda row: (float(row.split(','), [3])))

```
>>> txnVRDD = txnRDD.map(lambda row: (float(row.split(','), [3])))
>>>
```

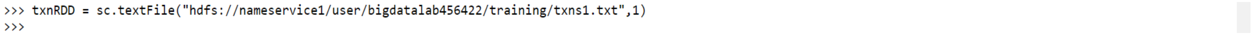
→

```
[bigdatalab456422@ip-10-1-1-204 ~]$ pyspark
```



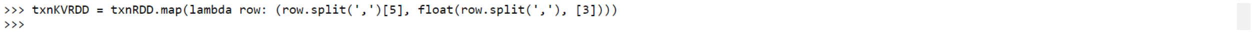
```
[bigdatalab456422@ip-10-1-1-204 ~]$ pyspark
Python 3.7.6 (default, Jan  8 2020, 19:59:22)
[GCC 7.3.0] :: Anaconda, Inc. on linux
Type "help", "copyright", "credits" or "license" for more information.
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
23/06/08 11:16:46 WARN cluster.YarnSchedulerBackend$YarnSchedulerEndpoint: Attempted to request executors before the AM has registered!
Welcome to
      ____              __
     / __/__  ___ _____/ /__
    _\ \/ _ \/ _ `/ __/  '_/
   /__ / .__/\_,_/_/ /_/\_\   version 2.4.0-cdh6.2.1
      /_/

Using Python version 3.7.6 (default, Jan  8 2020 19:59:22)
SparkSession available as 'spark'.
>>> 
```

```
>>> txnRDD = sc.textFile("hdfs://nameservice1/user/bigdatalab456422/training/txns1.txt",1)
```



```
>>> txnRDD = sc.textFile("hdfs://nameservice1/user/bigdatalab456422/training/txns1.txt",1)
>>>
```

```
>>> txnKVRDD = txnRDD.map(lambda row: (row.split(',')[5], float(row.split(','), [3])))
```



```
>>> txnKVRDD = txnRDD.map(lambda row: (row.split(',')[5], float(row.split(','), [3])))
>>>
```

```
>>> amtAndCount = txnKVRDD.mapValues(lambda a: (a, 1))
```

```
>>> amtAndCount = txnKVRDD.mapValues(lambda a: (a, 1))
>>>
```

```
>>> totalByProd = amtAndCount.reduceByKey(lambda a,b : (a[0]+b[0], a[1]+b[1]))
```

```
>>> totalByProd = amtAndCount.reduceByKey(lambda a,b : (a[0]+b[0], a[1]+b[1]))
>>>
```

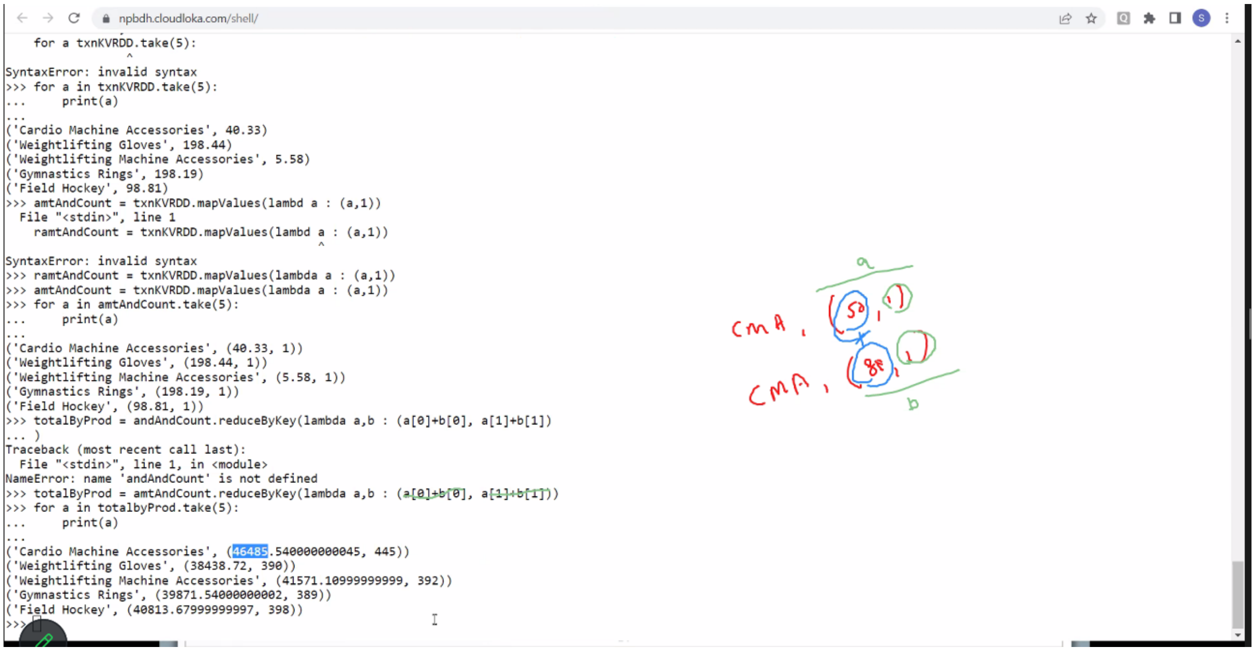
```
>>> avgforProd = totalByProd.mapValues(lambda a : a[0]/a[1])
```

```
>>> avgforProd = totalByProd.mapValues(lambda a : a[0]/a[1])
>>>
```

```
>>> maxAvg = avgforProd.sortBy(lambda a : -a[1])
```

```
>>> maxAvg = avgforProd.sortBy(lambda a : -a[1])
>>>
```

```
>>> for a in maxAvg.collect():
...     print(a)
...
```



npbdh.cloudloka.com/shell/

```
    for a txnKVRDD.take(5):
                 ^
SyntaxError: invalid syntax
>>> for a in txnKVRDD.take(5):
...     print(a)
...
('Cardio Machine Accessories', 40.33)
('Weightlifting Gloves', 198.44)
('Weightlifting Machine Accessories', 5.58)
('Gymnastics Rings', 198.19)
('Field Hockey', 98.81)
>>> amtAndCount = txnKVRDD.mapValues(lambd a : (a,1))
  File "<stdin>", line 1
    ramtAndCount = txnKVRDD.mapValues(lambd a : (a,1))
                                            ^
SyntaxError: invalid syntax
>>> ramtAndCount = txnKVRDD.mapValues(lambda a : (a,1))
>>> amtAndCount = txnKVRDD.mapValues(lambda a : (a,1))
>>> for a in amtAndCount.take(5):
...     print(a)
...
('Cardio Machine Accessories', (40.33, 1))
('Weightlifting Gloves', (198.44, 1))
('Weightlifting Machine Accessories', (5.58, 1))
('Gymnastics Rings', (198.19, 1))
('Field Hockey', (98.81, 1))
>>> totalByProd = andAndCount.reduceByKey(lambda a,b : (a[0]+b[0], a[1]+b[1])
... )
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
NameError: name 'andAndCount' is not defined
>>> totalByProd = amtAndCount.reduceByKey(lambda a,b : (a[0]+b[0], a[1]+b[1]))
>>> for a in totalbyProd.take(5):
...     print(a)
...
('Cardio Machine Accessories', (46485.540000000045, 445))
('Weightlifting Gloves', (38438.72, 390))
('Weightlifting Machine Accessories', (41571.10999999999, 392))
('Gymnastics Rings', (39871.54000000002, 389))
('Field Hockey', (40813.67999999997, 398))
>>>
```