HBASE Exercises

Problem definition:

1) Create a table named **EMP** in HBase with following columns.

   EMP_NO
   DEPT_NO
   FNAME
   LNAME
   SALARY
   DATE_OF_BIRTH

2) Populate at least 8 employee records in this table using **Java API.**

3) Use **Hbase-shell to verify** the data you have populated.

4) **Update salary** for one of the employee using **Java API.**

5) Use **Hbase-shell to verify** the data you have updated.

6) **Delete** one of the employee using **Java API.**

7) Use **Hbase-shell to verify** the data you have deleted.

8) Use **Java API to display remaining data** on screen which uses **GET operation.**

9) Use **Java API to display remaining data** on screen which uses **SCAN operation.**

Solution:

Start HBase Master and Region Server using following commands:

```
$sudo service hbase-master start
$sudo service hbase-regionserver start
```

1) Create a table named **EMP** in HBase with following columns.
EMP_NO, DEPT_NO, FNAME, LNAME, SALARY, DATE_OF_BIRTH

Compile and run **createEMPTable.java**:

```
$javac -cp `hbase classpath` creatEMPTable.java

$java -cp `hbase classpath` creatEMPTable
```

Verify creation of table in Hbase Shell:

```
hbase(main):007:0> list
TABLE
EMP
1 row(s) in 0.0070 seconds

=> ["EMP"]
hbase(main):008:0> desc 'EMP'
Table EMP is ENABLED
EMP
COLUMN FAMILIES DESCRIPTION
{NAME => 'ED', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE',
MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY
 => 'false', BLOCKCACHE => 'true'}
{NAME => 'EM', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE',
MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY
 => 'false', BLOCKCACHE => 'true'}
{NAME => 'ER', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE',
MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65536', IN_MEMORY
 => 'false', BLOCKCACHE => 'true'}
3 row(s) in 0.0370 seconds
```

2) Populate at least 8 employee records in this table using **Java API:**

Load following data from CSV file (empdata.csv) to Hbase table (EMP):

```
001,D001,Rohit,Saxena,1985-08-23,3000
002,D007,Naina,Ghanchi,1974-06-11,4600
003,D003,Abdul,Hameed,1980-01-30,3950
004,D001,Burt,Christian,1995-11-17,4100
005,D002,Indu,Rohilla,1967-12-09,5230
006,D001,Harman,Shoker,1997-06-09,2900
007,D002,Nola,Smith,1970-03-31,3400
008,D007,Mike,Layton,1977-07-29,4370
009,D001,Shailesh,Damodaran,1965-02-27,4750
010,D003,Billu,Hansra,1960-05-28,2800
```

Compile and run **insertEMPData.java**:

```
$javac -cp `hbase classpath` insertEMPData.java

$java -cp `hbase classpath` insertEMPData
```

3) Use **Hbase-shell to verify** the data you have populated:

```
hbase(main):009:0> scan 'EMP'
ROW                                                    COLUMN+CELL
 row0                                                  column=ED:DEPT_NO, timestamp=1489949895294, value=D001
 row0                                                  column=ED:EMP_NO, timestamp=1489949895294, value=001
 row0                                                  column=EM:DATE_OF_BIRTH, timestamp=1489949895294, value=1985-08-23
 row0                                                  column=EM:FNAME, timestamp=1489949895294, value=Rohit
 row0                                                  column=EM:LNAME, timestamp=1489949895294, value=Saxena
 row0                                                  column=ER:SALARY, timestamp=1489949895294, value=3000
 row1                                                  column=ED:DEPT_NO, timestamp=1489949895309, value=D007
 row1                                                  column=ED:EMP_NO, timestamp=1489949895309, value=002
 row1                                                  column=EM:DATE_OF_BIRTH, timestamp=1489949895309, value=1974-06-11
 row1                                                  column=EM:FNAME, timestamp=1489949895309, value=Naina
 row1                                                  column=EM:LNAME, timestamp=1489949895309, value=Ghanchi
 row1                                                  column=ER:SALARY, timestamp=1489949895309, value=4600
 row2                                                  column=ED:DEPT_NO, timestamp=1489949895313, value=D003
 row2                                                  column=ED:EMP_NO, timestamp=1489949895313, value=003
 row2                                                  column=EM:DATE_OF_BIRTH, timestamp=1489949895313, value=1980-01-30
 row2                                                  column=EM:FNAME, timestamp=1489949895313, value=Abdul
 row2                                                  column=EM:LNAME, timestamp=1489949895313, value=Hameed
 row2                                                  column=ER:SALARY, timestamp=1489949895313, value=3950
```

```
 row3                                                    column=ED:DEPT_NO, timestamp=1489949895317, value=D001
 row3                                                    column=ED:EMP_NO, timestamp=1489949895317, value=004
 row3                                                    column=EM:DATE_OF_BIRTH, timestamp=1489949895317, value=1995-11-17
 row3                                                    column=EM:FNAME, timestamp=1489949895317, value=Burt
 row3                                                    column=EM:LNAME, timestamp=1489949895317, value=Christian
 row3                                                    column=ER:SALARY, timestamp=1489949895317, value=4100
 row4                                                    column=ED:DEPT_NO, timestamp=1489949895320, value=D002
 row4                                                    column=ED:EMP_NO, timestamp=1489949895320, value=005
 row4                                                    column=EM:DATE_OF_BIRTH, timestamp=1489949895320, value=1967-12-09
 row4                                                    column=EM:FNAME, timestamp=1489949895320, value=Indu
 row4                                                    column=EM:LNAME, timestamp=1489949895320, value=Rohilla
 row4                                                    column=ER:SALARY, timestamp=1489949895320, value=5230
 row5                                                    column=ED:DEPT_NO, timestamp=1489949895323, value=D001
 row5                                                    column=ED:EMP_NO, timestamp=1489949895323, value=006
 row5                                                    column=EM:DATE_OF_BIRTH, timestamp=1489949895323, value=1997-06-09
 row5                                                    column=EM:FNAME, timestamp=1489949895323, value=Harman
 row5                                                    column=EM:LNAME, timestamp=1489949895323, value=Shoker
 row5                                                    column=ER:SALARY, timestamp=1489949895323, value=2900
 row6                                                    column=ED:DEPT_NO, timestamp=1489949895327, value=D002
 row6                                                    column=ED:EMP_NO, timestamp=1489949895327, value=007
 row6                                                    column=EM:DATE_OF_BIRTH, timestamp=1489949895327, value=1970-03-31
 row6                                                    column=EM:FNAME, timestamp=1489949895327, value=Nola
 row6                                                    column=EM:LNAME, timestamp=1489949895327, value=Smith
 row6                                                    column=ER:SALARY, timestamp=1489949895327, value=3400
 row7                                                    column=ED:DEPT_NO, timestamp=1489949895330, value=D007
 row7                                                    column=ED:EMP_NO, timestamp=1489949895330, value=008
 row7                                                    column=EM:DATE_OF_BIRTH, timestamp=1489949895330, value=1977-07-29
 row7                                                    column=EM:FNAME, timestamp=1489949895330, value=Mike
 row7                                                    column=EM:LNAME, timestamp=1489949895330, value=Layton
 row7                                                    column=ER:SALARY, timestamp=1489949895330, value=4370
 row8                                                    column=ED:DEPT_NO, timestamp=1489949895334, value=D001
 row8                                                    column=ED:EMP_NO, timestamp=1489949895334, value=009
 row8                                                    column=EM:DATE_OF_BIRTH, timestamp=1489949895334, value=1965-02-27
 row8                                                    column=EM:FNAME, timestamp=1489949895334, value=Shailesh
 row8                                                    column=EM:LNAME, timestamp=1489949895334, value=Damodaran
 row8                                                    column=ER:SALARY, timestamp=1489949895334, value=4750
 row9                                                    column=ED:DEPT_NO, timestamp=1489949895338, value=D003
 row9                                                    column=ED:EMP_NO, timestamp=1489949895338, value=010
 row9                                                    column=EM:DATE_OF_BIRTH, timestamp=1489949895338, value=1960-05-28
 row9                                                    column=EM:FNAME, timestamp=1489949895338, value=Billu
 row9                                                    column=EM:LNAME, timestamp=1489949895338, value=Hansra
 row9                                                    column=ER:SALARY, timestamp=1489949895338, value=2800
10 row(s) in 0.4480 seconds
```

4) **Update salary** for one of the employee using **Java API:**

Compile and run **updateSalary.java:**

```
$javac -cp `hbase classpath` updateSalary.java

$java -cp `hbase classpath` updateSalary
```

5) Use **Hbase-shell to verify** the data you have updated:

Verify updated Salary for row5:

```
hbase(main):019:0> get 'EMP','row5'
COLUMN                                                  CELL
 ED:DEPT_NO                                             timestamp=1489949895323, value=D001
 ED:EMP_NO                                              timestamp=1489949895323, value=006
 EM:DATE_OF_BIRTH                                       timestamp=1489949895323, value=1997-06-09
 EM:FNAME                                               timestamp=1489949895323, value=Harman
 EM:LNAME                                               timestamp=1489949895323, value=Shoker
 ER:SALARY                                              timestamp=1489967900612, value=3200
6 row(s) in 0.0160 seconds
```

6) **Delete** one of the employee using **Java API:**

Compile and run **deleteEMP.java:**

```
$javac -cp `hbase classpath` deleteEMP.java

$java -cp `hbase classpath` deleteEMP
```

7) Use **Hbase-shell to verify** the data you have deleted:

Verify that row3 is deleted:

```
hbase(main):023:0> scan 'EMP'
ROW                                                    COLUMN+CELL
 row0                                                  column=ED:DEPT_NO, timestamp=1489949895294, value=D001
 row0                                                  column=ED:EMP_NO, timestamp=1489949895294, value=001
 row0                                                  column=EM:DATE_OF_BIRTH, timestamp=1489949895294, value=1985-08-23
 row0                                                  column=EM:FNAME, timestamp=1489949895294, value=Rohit
 row0                                                  column=EM:LNAME, timestamp=1489949895294, value=Saxena
 row0                                                  column=ER:SALARY, timestamp=1489949895294, value=3000
 row1                                                  column=ED:DEPT_NO, timestamp=1489949895309, value=D007
 row1                                                  column=ED:EMP_NO, timestamp=1489949895309, value=002
 row1                                                  column=EM:DATE_OF_BIRTH, timestamp=1489949895309, value=1974-06-11
 row1                                                  column=EM:FNAME, timestamp=1489949895309, value=Naina
 row1                                                  column=EM:LNAME, timestamp=1489949895309, value=Ghanchi
 row1                                                  column=ER:SALARY, timestamp=1489949895309, value=4600
 row2                                                  column=ED:DEPT_NO, timestamp=1489949895313, value=D003
 row2                                                  column=ED:EMP_NO, timestamp=1489949895313, value=003
 row2                                                  column=EM:DATE_OF_BIRTH, timestamp=1489949895313, value=1980-01-30
 row2                                                  column=EM:FNAME, timestamp=1489949895313, value=Abdul
 row2                                                  column=EM:LNAME, timestamp=1489949895313, value=Hameed
 row2                                                  column=ER:SALARY, timestamp=1489949895313, value=3950
 row4                                                  column=ED:DEPT_NO, timestamp=1489949895320, value=D002
 row4                                                  column=ED:EMP_NO, timestamp=1489949895320, value=005
 row4                                                  column=EM:DATE_OF_BIRTH, timestamp=1489949895320, value=1967-12-09
 row4                                                  column=EM:FNAME, timestamp=1489949895320, value=Indu
 row4                                                  column=EM:LNAME, timestamp=1489949895320, value=Rohilla
 row4                                                  column=ER:SALARY, timestamp=1489949895320, value=5230
 row5                                                  column=ED:DEPT_NO, timestamp=1489949895323, value=D001
 row5                                                  column=ED:EMP_NO, timestamp=1489949895323, value=006
 row5                                                  column=EM:DATE_OF_BIRTH, timestamp=1489949895323, value=1997-06-09
 row5                                                  column=EM:FNAME, timestamp=1489949895323, value=Harman
 row5                                                  column=EM:LNAME, timestamp=1489949895323, value=Shoker
 row5                                                  column=ER:SALARY, timestamp=1489967900612, value=3200
 row6                                                  column=ED:DEPT_NO, timestamp=1489949895327, value=D002
 row6                                                  column=ED:EMP_NO, timestamp=1489949895327, value=007
 row6                                                  column=EM:DATE_OF_BIRTH, timestamp=1489949895327, value=1970-03-31
 row6                                                  column=EM:FNAME, timestamp=1489949895327, value=Nola
 row6                                                  column=EM:LNAME, timestamp=1489949895327, value=Smith
 row6                                                  column=ER:SALARY, timestamp=1489949895327, value=3400
 row7                                                  column=ED:DEPT_NO, timestamp=1489949895330, value=D007
 row7                                                  column=ED:EMP_NO, timestamp=1489949895330, value=008
 row7                                                  column=EM:DATE_OF_BIRTH, timestamp=1489949895330, value=1977-07-29
 row7                                                  column=EM:FNAME, timestamp=1489949895330, value=Mike
 row7                                                  column=EM:LNAME, timestamp=1489949895330, value=Layton
 row7                                                  column=ER:SALARY, timestamp=1489949895330, value=4370
 row8                                                  column=ED:DEPT_NO, timestamp=1489949895334, value=D001
```

```
 row8                                                    column=ED:EMP_NO, timestamp=1489949895334, value=009
 row8                                                    column=EM:DATE_OF_BIRTH, timestamp=1489949895334, value=1965-02-27
 row8                                                    column=EM:FNAME, timestamp=1489949895334, value=Shailesh
 row8                                                    column=EM:LNAME, timestamp=1489949895334, value=Damodaran
 row8                                                    column=ER:SALARY, timestamp=1489949895334, value=4750
 row9                                                    column=ED:DEPT_NO, timestamp=1489949895338, value=D003
 row9                                                    column=ED:EMP_NO, timestamp=1489949895338, value=010
 row9                                                    column=EM:DATE_OF_BIRTH, timestamp=1489949895338, value=1960-05-28
 row9                                                    column=EM:FNAME, timestamp=1489949895338, value=Billu
 row9                                                    column=EM:LNAME, timestamp=1489949895338, value=Hansra
 row9                                                    column=ER:SALARY, timestamp=1489949895338, value=2800
9 row(s) in 0.1090 seconds
```

8) Use **Java API to display remaining data** on screen which uses **GET operation:**

Compile and run **getDataUsingGET.java:**

```
$javac -cp `hbase classpath` getDataUsingGET.java

$java -cp `hbase classpath` getDataUsingGET
```

Output on the screen:

```
EMP_NO: 001 DEPT_NO: D001 FNAME: Rohit LNAME: Saxena DATE_OF_BIRTH: 1985-08-23 SALARY: 3000
EMP_NO: 002 DEPT_NO: D007 FNAME: Naina LNAME: Ghanchi DATE_OF_BIRTH: 1974-06-11 SALARY: 4600
EMP_NO: 003 DEPT_NO: D003 FNAME: Abdul LNAME: Hameed DATE_OF_BIRTH: 1980-01-30 SALARY: 3950
EMP_NO: 005 DEPT_NO: D002 FNAME: Indu LNAME: Rohilla DATE_OF_BIRTH: 1967-12-09 SALARY: 5230
EMP_NO: 006 DEPT_NO: D001 FNAME: Harman LNAME: Shoker DATE_OF_BIRTH: 1997-06-09 SALARY: 3200
EMP_NO: 007 DEPT_NO: D002 FNAME: Nola LNAME: Smith DATE_OF_BIRTH: 1970-03-31 SALARY: 3400
EMP_NO: 008 DEPT_NO: D007 FNAME: Mike LNAME: Layton DATE_OF_BIRTH: 1977-07-29 SALARY: 4370
EMP_NO: 009 DEPT_NO: D001 FNAME: Shailesh LNAME: Damodaran DATE_OF_BIRTH: 1965-02-27 SALARY: 4750
EMP_NO: 010 DEPT_NO: D003 FNAME: Billu LNAME: Hansra DATE_OF_BIRTH: 1960-05-28 SALARY: 2800
Table data retrieved
```

9) Use **Java API to display remaining data** on screen which uses **SCAN operation:**

Compile and run **getDataUsingSCAN.java:**

```
$javac -cp `hbase classpath` getDataUsingSCAN.java

$java -cp `hbase classpath` getDataUsingSCAN
```

Output on the screen:

```
EMP_NO: 001 DEPT_NO: D001 FNAME: Rohit LNAME: Saxena DATE_OF_BIRTH: 1985-08-23 SALARY: 3000
EMP_NO: 002 DEPT_NO: D007 FNAME: Naina LNAME: Ghanchi DATE_OF_BIRTH: 1974-06-11 SALARY: 4600
EMP_NO: 003 DEPT_NO: D003 FNAME: Abdul LNAME: Hameed DATE_OF_BIRTH: 1980-01-30 SALARY: 3950
EMP_NO: 005 DEPT_NO: D002 FNAME: Indu LNAME: Rohilla DATE_OF_BIRTH: 1967-12-09 SALARY: 5230
EMP_NO: 006 DEPT_NO: D001 FNAME: Harman LNAME: Shoker DATE_OF_BIRTH: 1997-06-09 SALARY: 3200
EMP_NO: 007 DEPT_NO: D002 FNAME: Nola LNAME: Smith DATE_OF_BIRTH: 1970-03-31 SALARY: 3400
EMP_NO: 008 DEPT_NO: D007 FNAME: Mike LNAME: Layton DATE_OF_BIRTH: 1977-07-29 SALARY: 4370
EMP_NO: 009 DEPT_NO: D001 FNAME: Shailesh LNAME: Damodaran DATE_OF_BIRTH: 1965-02-27 SALARY: 4750
EMP_NO: 010 DEPT_NO: D003 FNAME: Billu LNAME: Hansra DATE_OF_BIRTH: 1960-05-28 SALARY: 2800
Table data retrieved
```

NOTE: Just printing from Scanned row was displaying data in the **following** format. I used **Bytes.toString** method in combination with 'Get' to convert it into a readable form.

```
keyvalues={row0/ED:DEPT_NO/1489949895294/Put/vlen=4/seqid=0, row0/ED:EMP_NO/1489949895294/Put/vlen=3/seqid=0,
row0/EM:DATE_OF_BIRTH/1489949895294/Put/vlen=10/seqid=0, row0/EM:FNAME/1489949895294/Put/vlen=5/seqid=0,
row0/EM:LNAME/1489949895294/Put/vlen=6/seqid=0, row0/ER:SALARY/1489949895294/Put/vlen=4/seqid=0}

keyvalues={row1/ED:DEPT_NO/1489949895309/Put/vlen=4/seqid=0, row1/ED:EMP_NO/1489949895309/Put/vlen=3/seqid=0,
row1/EM:DATE_OF_BIRTH/1489949895309/Put/vlen=10/seqid=0, row1/EM:FNAME/1489949895309/Put/vlen=5/seqid=0,
row1/EM:LNAME/1489949895309/Put/vlen=7/seqid=0, row1/ER:SALARY/1489949895309/Put/vlen=4/seqid=0}

keyvalues={row2/ED:DEPT_NO/1489949895313/Put/vlen=4/seqid=0, row2/ED:EMP_NO/1489949895313/Put/vlen=3/seqid=0,
row2/EM:DATE_OF_BIRTH/1489949895313/Put/vlen=10/seqid=0, row2/EM:FNAME/1489949895313/Put/vlen=5/seqid=0,
row2/EM:LNAME/1489949895313/Put/vlen=6/seqid=0, row2/ER:SALARY/1489949895313/Put/vlen=4/seqid=0}

keyvalues={row4/ED:DEPT_NO/1489949895320/Put/vlen=4/seqid=0, row4/ED:EMP_NO/1489949895320/Put/vlen=3/seqid=0,
row4/EM:DATE_OF_BIRTH/1489949895320/Put/vlen=10/seqid=0, row4/EM:FNAME/1489949895320/Put/vlen=4/seqid=0,
row4/EM:LNAME/1489949895320/Put/vlen=7/seqid=0, row4/ER:SALARY/1489949895320/Put/vlen=4/seqid=0}
```

```
keyvalues={row5/ED:DEPT_NO/1489949895323/Put/vlen=4/seqid=0, row5/ED:EMP_NO/1489949895323/Put/vlen=3/seqid=0,
row5/EM:DATE_OF_BIRTH/1489949895323/Put/vlen=10/seqid=0, row5/EM:FNAME/1489949895323/Put/vlen=6/seqid=0,
row5/EM:LNAME/1489949895323/Put/vlen=6/seqid=0, row5/ER:SALARY/1489967900612/Put/vlen=4/seqid=0}

keyvalues={row6/ED:DEPT_NO/1489949895327/Put/vlen=4/seqid=0, row6/ED:EMP_NO/1489949895327/Put/vlen=3/seqid=0,
row6/EM:DATE_OF_BIRTH/1489949895327/Put/vlen=10/seqid=0, row6/EM:FNAME/1489949895327/Put/vlen=4/seqid=0,
row6/EM:LNAME/1489949895327/Put/vlen=5/seqid=0, row6/ER:SALARY/1489949895327/Put/vlen=4/seqid=0}

keyvalues={row7/ED:DEPT_NO/1489949895330/Put/vlen=4/seqid=0, row7/ED:EMP_NO/1489949895330/Put/vlen=3/seqid=0,
row7/EM:DATE_OF_BIRTH/1489949895330/Put/vlen=10/seqid=0, row7/EM:FNAME/1489949895330/Put/vlen=4/seqid=0,
row7/EM:LNAME/1489949895330/Put/vlen=6/seqid=0, row7/ER:SALARY/1489949895330/Put/vlen=4/seqid=0}

keyvalues={row8/ED:DEPT_NO/1489949895334/Put/vlen=4/seqid=0, row8/ED:EMP_NO/1489949895334/Put/vlen=3/seqid=0,
row8/EM:DATE_OF_BIRTH/1489949895334/Put/vlen=10/seqid=0, row8/EM:FNAME/1489949895334/Put/vlen=8/seqid=0,
row8/EM:LNAME/1489949895334/Put/vlen=9/seqid=0, row8/ER:SALARY/1489949895334/Put/vlen=4/seqid=0}

keyvalues={row9/ED:DEPT_NO/1489949895338/Put/vlen=4/seqid=0, row9/ED:EMP_NO/1489949895338/Put/vlen=3/seqid=0,
row9/EM:DATE_OF_BIRTH/1489949895338/Put/vlen=10/seqid=0, row9/EM:FNAME/1489949895338/Put/vlen=5/seqid=0,
row9/EM:LNAME/1489949895338/Put/vlen=6/seqid=0, row9/ER:SALARY/1489949895338/Put/vlen=4/seqid=0}

Table data retrieved
```