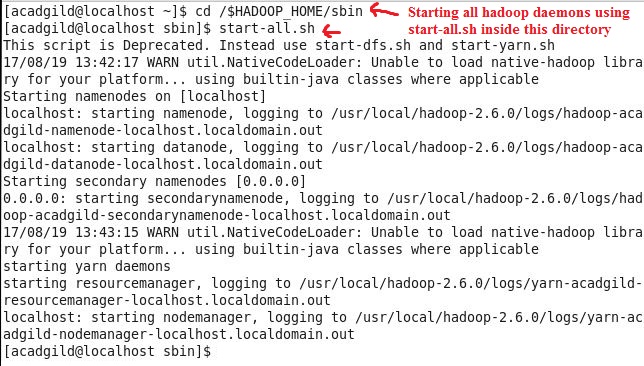
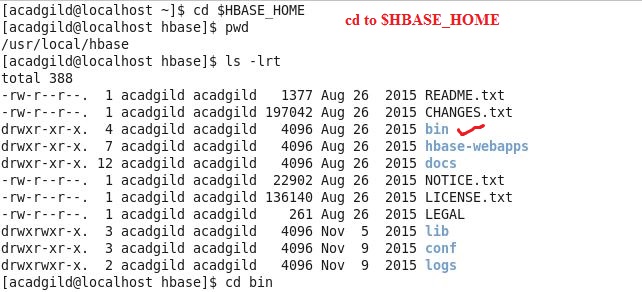
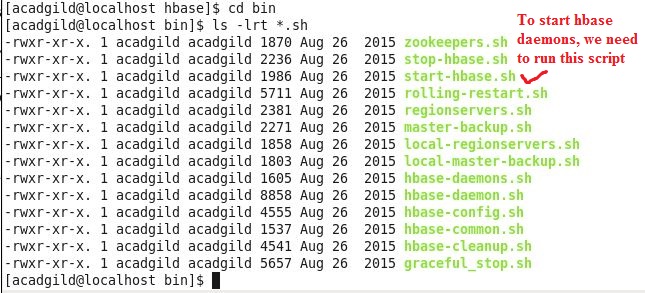
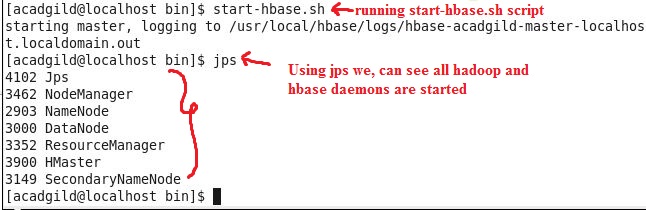
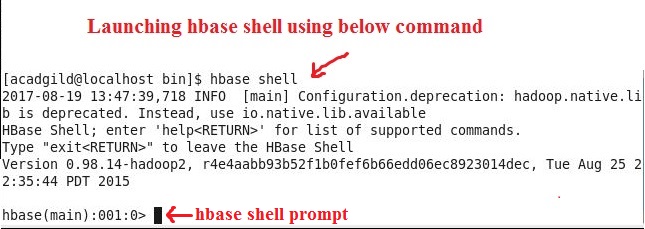
**HBase - Assignment**

To launch HBase Shell, below steps are performed:

1. Started all hadoop daemons, using start-all.sh command inside /$HADOOP\_HOME/sbin directory.
2. Started hbase daemon, using start-hbase.sh command inside /$HBASE\_HOME/bin
3. Using jps, we can see all daemons have started or not.
4. Since, all daemons started, then launched hbase shell using hbase shell command.

Refer below screenshots for above steps:





****Since hbase shell prompt has appeared, now we can proceed with operations on hbase tables

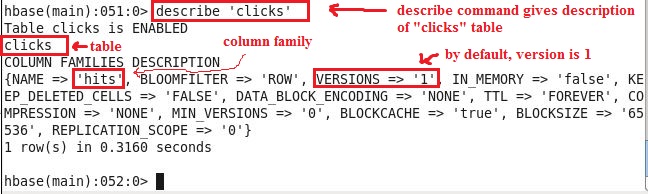
Problem 1:

Create an HBase table named 'clicks' with a column family 'hits' such that it should be able to  
store last 5 values of qualifiers inside 'hits' column family.

**Below steps are followed to accomplish above task:**

**Step 1: Create table “clicks” with single column family “hits”**

**Step 2: Describe “clicks” table**

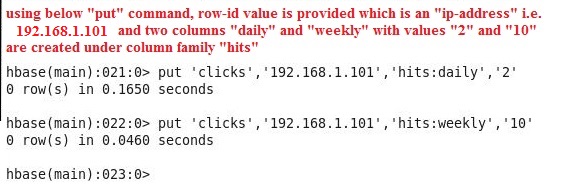


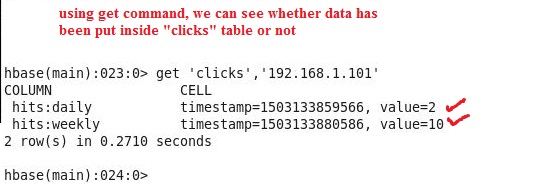
**NOTE: Version information is altered/ updated to 5, while accomplishing Task 2.**

Problem 2:

Add few records in the table and update some of them. Use IP Address as row-key. Scan the  
table to view if all the previous versions are getting displayed.

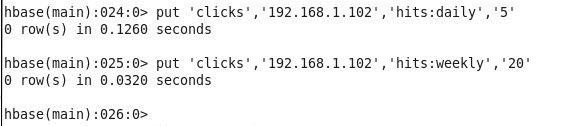
**Below steps are followed to accomplish above task:**

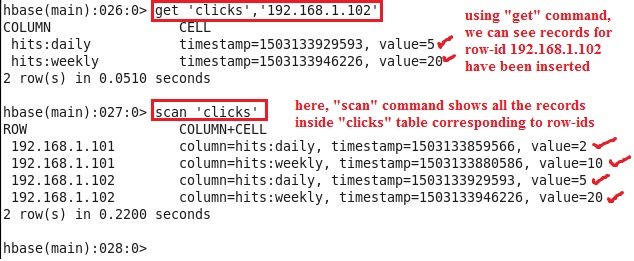
**Step 1: Insert data inside “click” table**

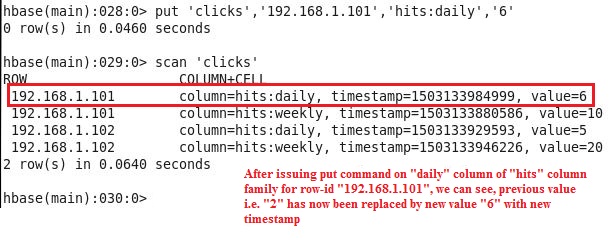
**Step 2: Check whether data has been put inside “clicks” table or not, using get command**

**Above “get” command displays column-family:columns, timestamp and values**

**NOTE: Atleast one column-family needs to be specified along with table name in get command, else error with suggestion would be returned.**

**Step 3: Putting new records for row-id “192.168.1.102”**

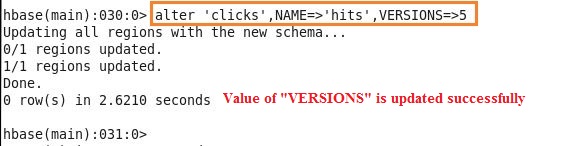
**Step 4: Till now we have created four records, two records for row-id “192.168.1.101” and two records for row-id “192.168.1.102”. Using get and scan command we can see the inserted records**

**Step 5: Updating value of column “daily” for column-family “hits” for row-id “192.168.1.101”**

**Explanation of above screenshot:**

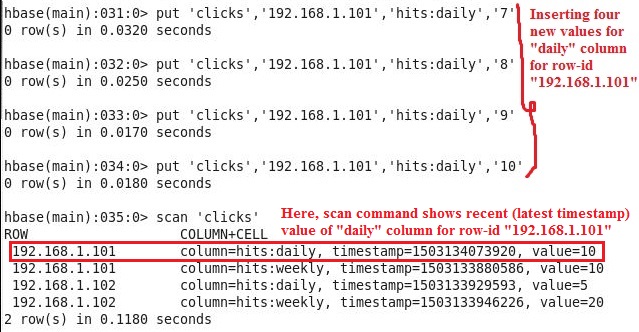
**Here, old value is replaced by new value, because by default “VERSIONS” is set to 1. So if we want that five versions of column-family “hits” must be maintained, then we need to change value of “VERSIONS” to 5 explicitly.**

**Refer next step for altering the value of VERSIONS to 5.**

**Step 6: Update value of VERSIONS to 5**

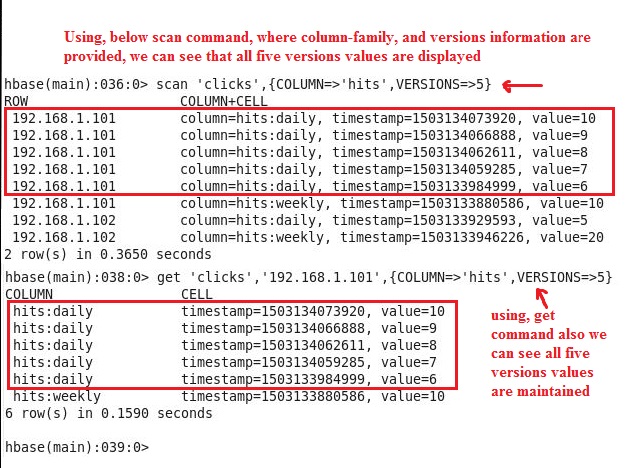
**Since, value of “VERSIONS” is updated for column-family “hits”, now by inserting new values for same column for same row-id, we can check whether different versions are maintained or not.**

**Refer next step for this.**

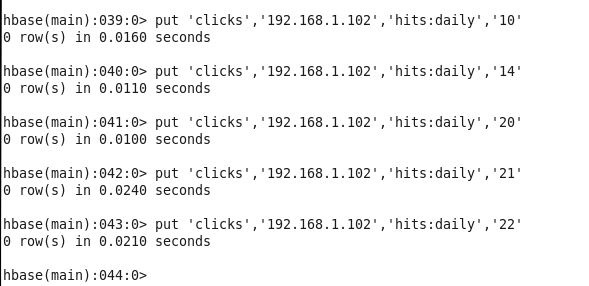
**Step 7: Insert new values for “daily” column for row-id “192.168.1.101”, and using “scan” check whether different versions for “daily” column of “hits” column-family are maintained or not.**

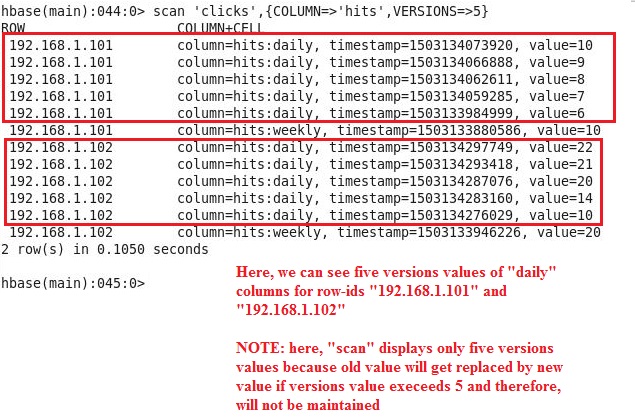
**In above screenshot, we can see, “scan” command displays only latest value 10 for “daily” column of “hits” column-family for row-id “192.168.1.101”.**

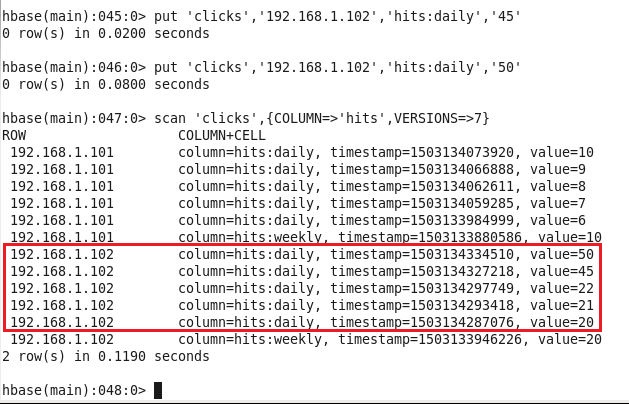
**Now to see five versions of column family, refer next step.**

**Step 8: Show all five versions values for “daily” column of “hit” column-family**

**Step 9: Now insert 5 new records for “daily” column of row-id “192.168.1.102”**

****

**Step 10: Check using scan, different versions values of “daily” column for both row-ids**

**Step 11: Again insert few records for daily column for row-id “192.168.1.102”, and try scan command with “VERSION=>7”**

**Explanation of above screenshot:**

**In above screenshot, we can see, after inserting two new values for “daily” column for row-id “192.168.1.102”, only 5 records are shown not 7, however in “scan” command VERSIONS=>7 is mentioned.**

**This is because, initially using “alter” command, we set VERSIONS=>5, so only 5 versions are maintained, therefore if new values come after five versions values, then old value is replaced by new value.**

**In Step 10 Screenshot, we can see values for “daily” column for row-id “192.168.1.102” are “22,21,20,14,10” and now values are “50,45,22,21,20”, values “14 and 10” are discarded because their timestamp values are old.**