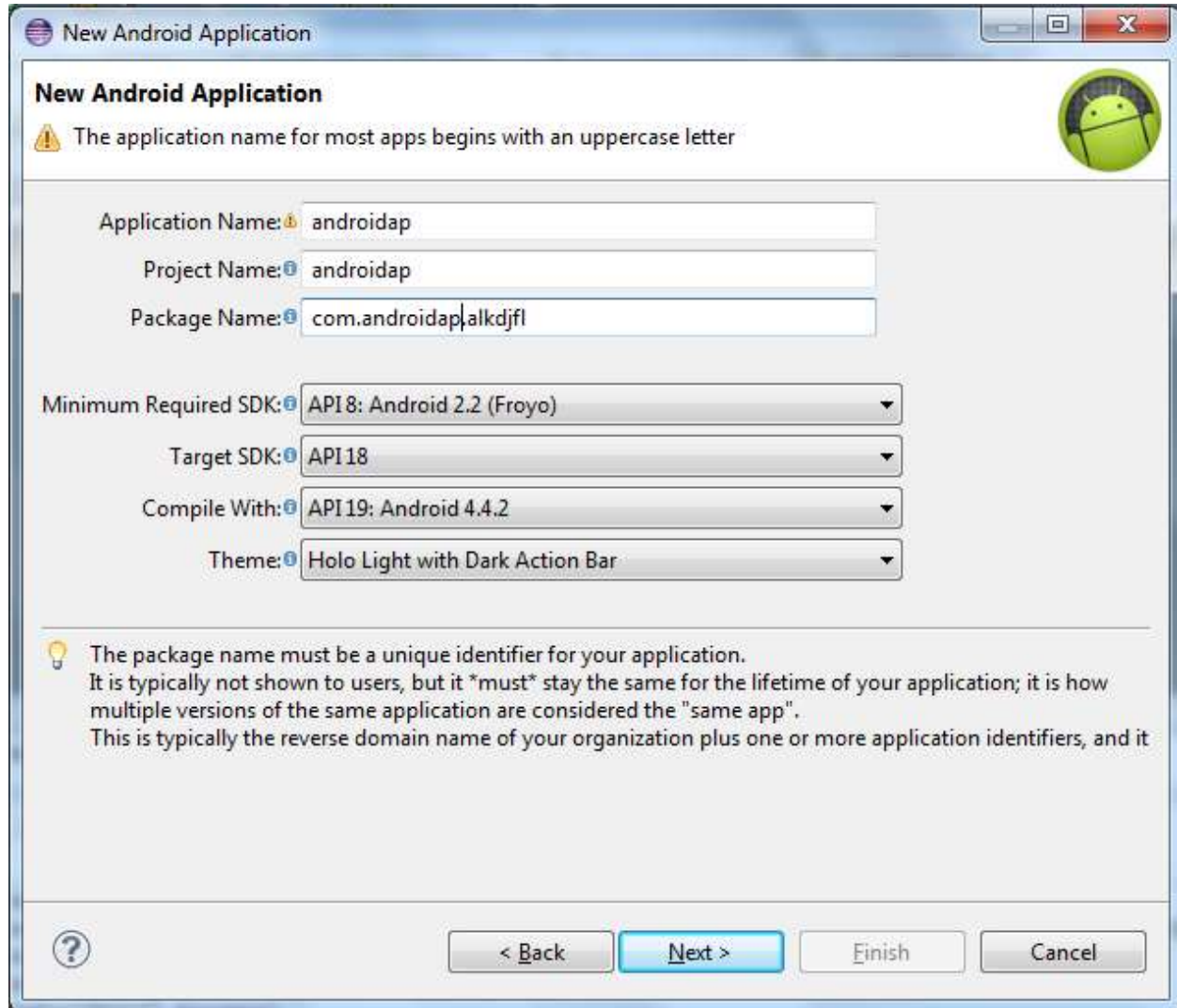# KDM

# LAB – III

Submitted by

Mahesh Vemula

16158759

1.Make a **Mashup application** including various services (e.g., Google Map, Google Chart, Google Search, Yahoo, Amazon, Twitter, Facebook) Web Services (e.g., Google Map Services, Weather Services) Android App Technology

Step 1: Create new android application using ECLIPSE

Step 2: Now create and modify the activities and layout in java and XML.

For creating this application files created : MainActivity.java

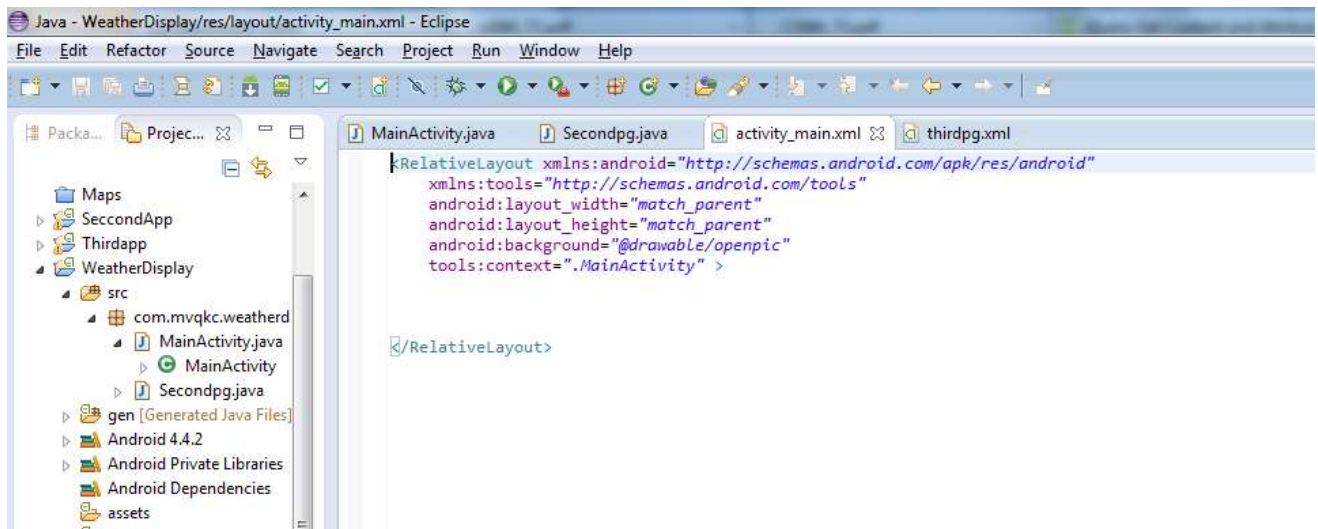Secondpage.java

Weatherpage.java

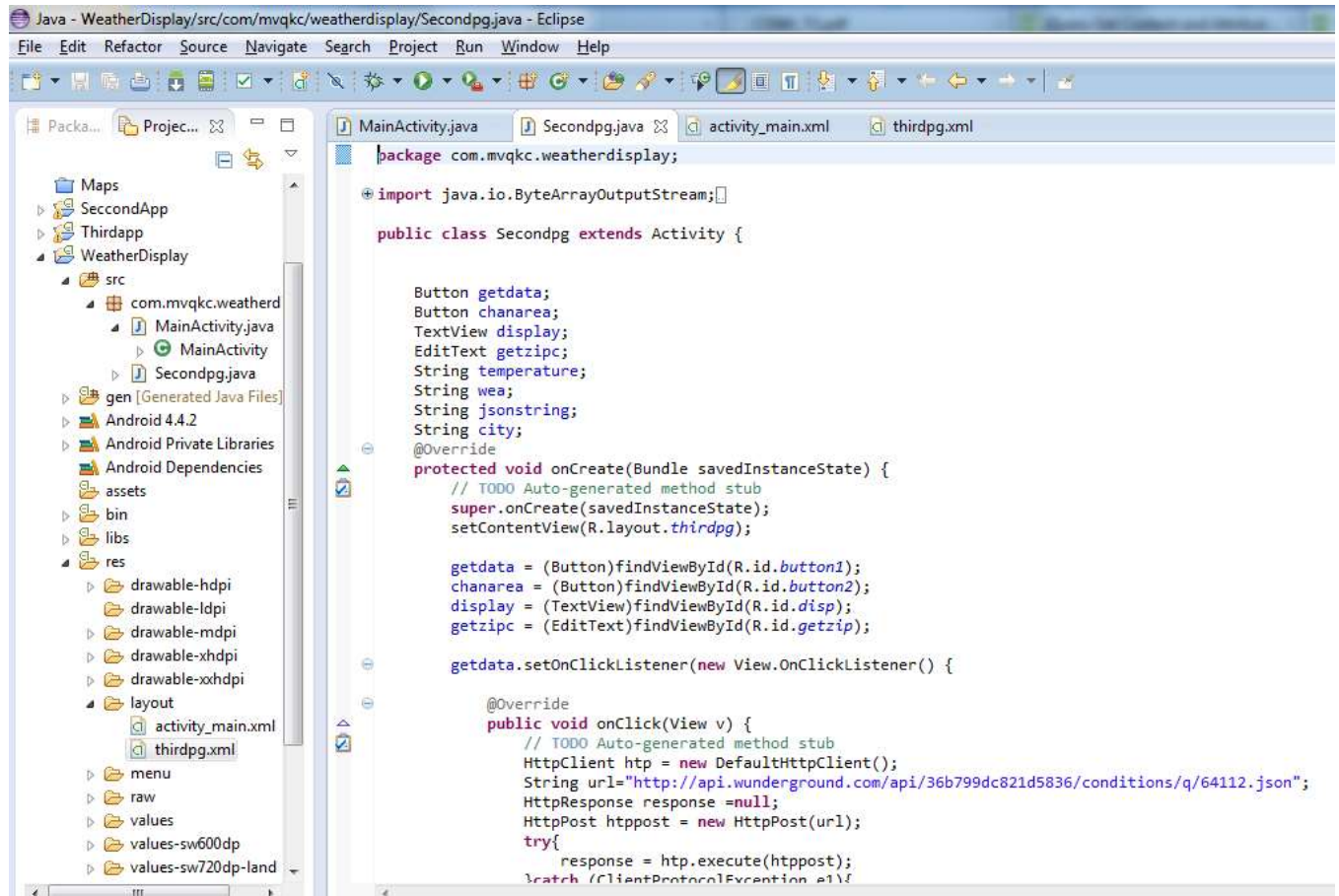Countrypage.java

Activity_main.xml

Secondpage.xml

Weatherpage.xml

Countrypage.xml

AppManifest.xml



Main Activity XML File

File   Edit   Refactor   Source   Navigate   Search   Project   Run   Window   Help

| Packa... | Projec... 23 | | | MainActivity.java | Secondpg.java 23 | activity_main.xml | thirdpg.xml |

```java
package com.mvqkc.weatherdisplay;

import java.io.ByteArrayOutputStream;

public class Secondpg extends Activity {

    Button getdata;
    Button chanarea;
    TextView display;
    EditText getzipc;
    String temperature;
    String wea;
    String jsonstring;
    String city;
    @Override
    protected void onCreate(Bundle savedInstanceState) {
        // TODO Auto-generated method stub
        super.onCreate(savedInstanceState);
        setContentView(R.layout.thirdpg);

        getdata = (Button)findViewById(R.id.button1);
        chanarea = (Button)findViewById(R.id.button2);
        display = (TextView)findViewById(R.id.disp);
        getzipc = (EditText)findViewById(R.id.getzip);

        getdata.setOnClickListener(new View.OnClickListener() {

            @Override
            public void onClick(View v) {
                // TODO Auto-generated method stub
                HttpClient htp = new DefaultHttpClient();
                String url="http://api.wunderground.com/api/36b799dc821d5836/conditions/q/64112.json";
                HttpResponse response =null;
                HttpPost httppost = new HttpPost(url);
                try{
                    response = htp.execute(httppost);
                }catch (ClientProtocolException e1){
```

- Maps
- SeccondApp
- Thirdapp
- WeatherDisplay
  - src
    - com.mvqkc.weatherd
      - MainActivity.java
        - MainActivity
      - Secondpg.java
  - gen [Generated Java Files]
  - Android 4.4.2
  - Android Private Libraries
  - Android Dependencies
  - assets
  - bin
  - libs
  - res
    - drawable-hdpi
    - drawable-ldpi
    - drawable-mdpi
    - drawable-xhdpi
    - drawable-xxhdpi
    - layout
      - activity_main.xml
      - thirdpg.xml
    - menu
    - raw
    - values
    - values-sw600dp
    - values-sw720dp-land

Main Activity Java File

Window of app XML

Second window editing in Graphical View

STEP 3: Now after editing the required file we test application on created Virtual AVD


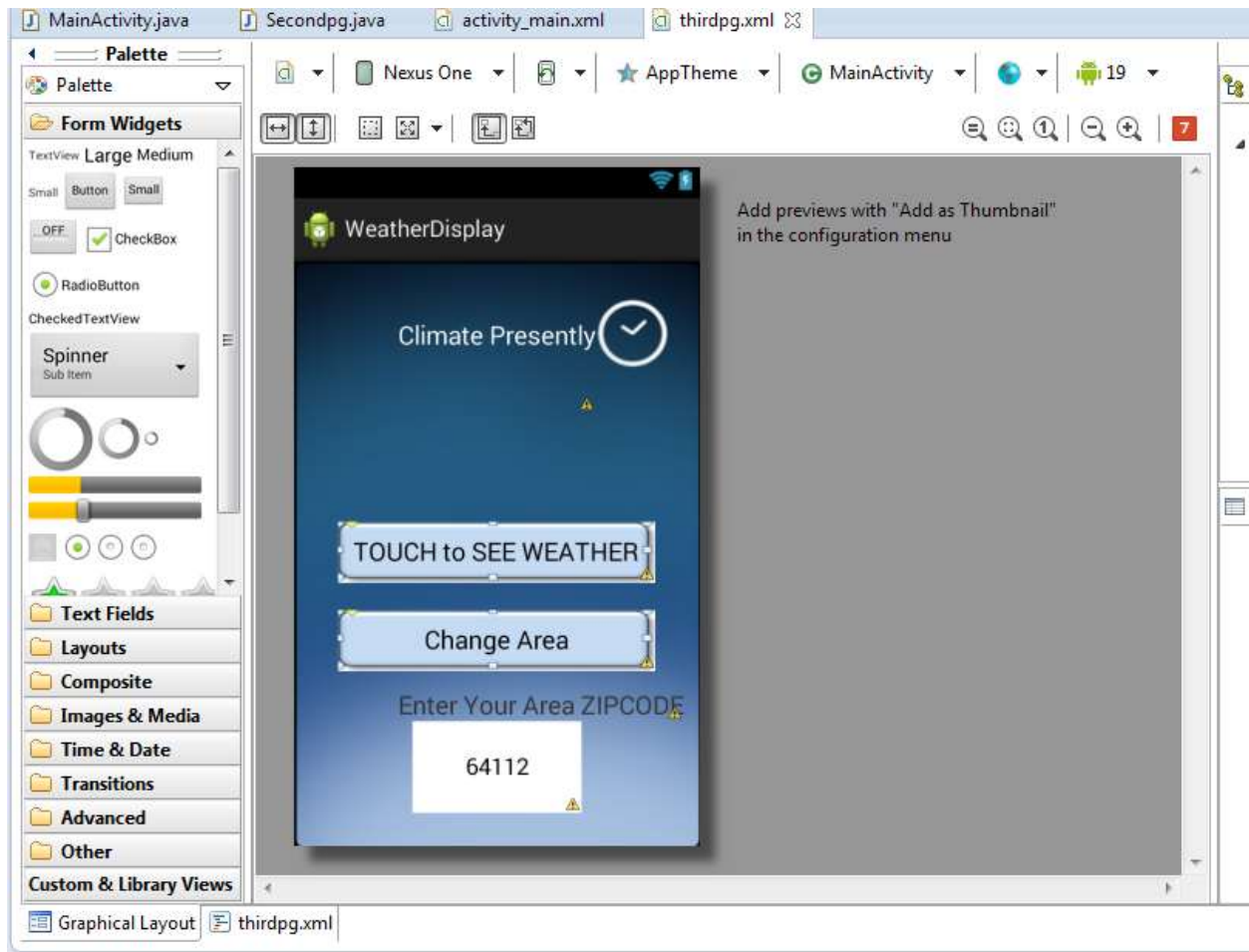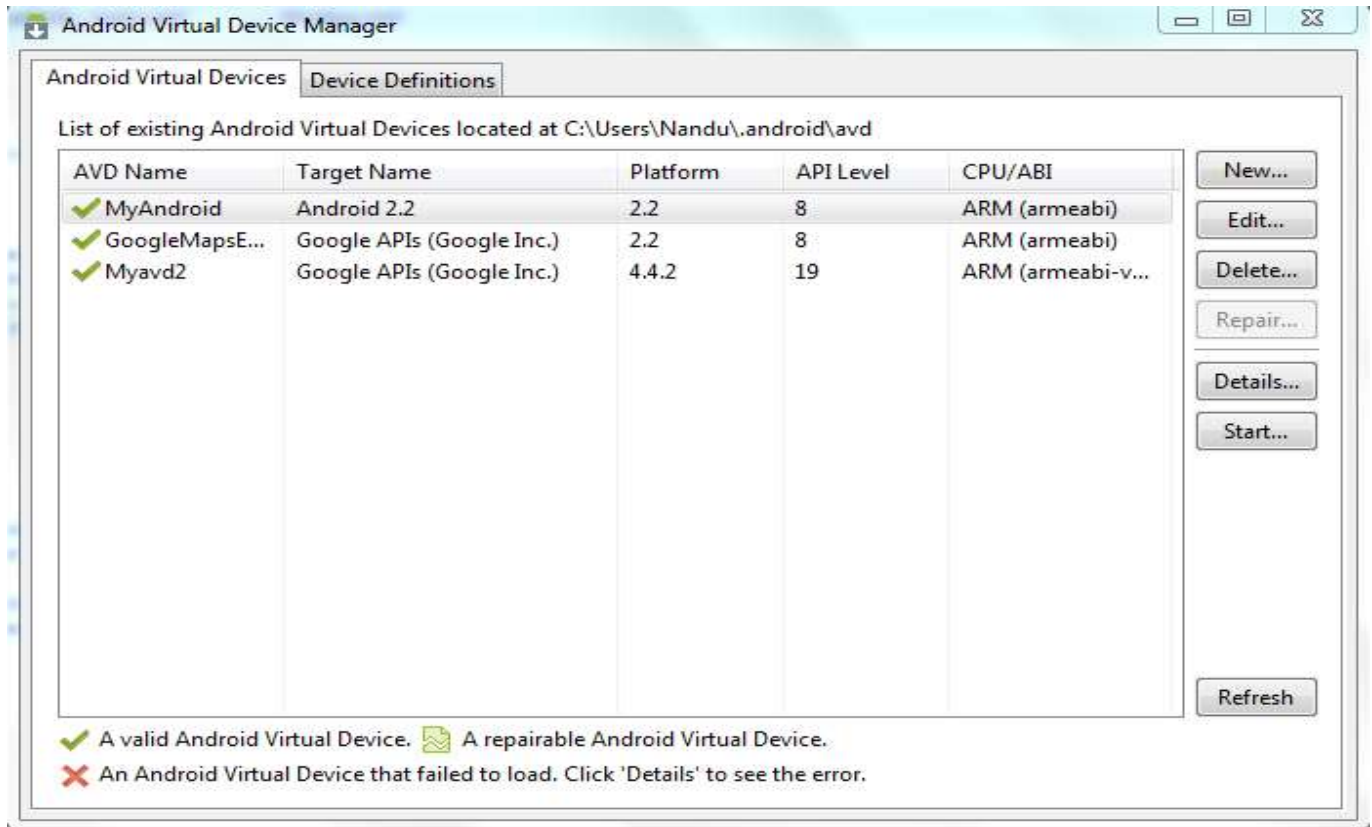
Select the required Virtual AVD to test app
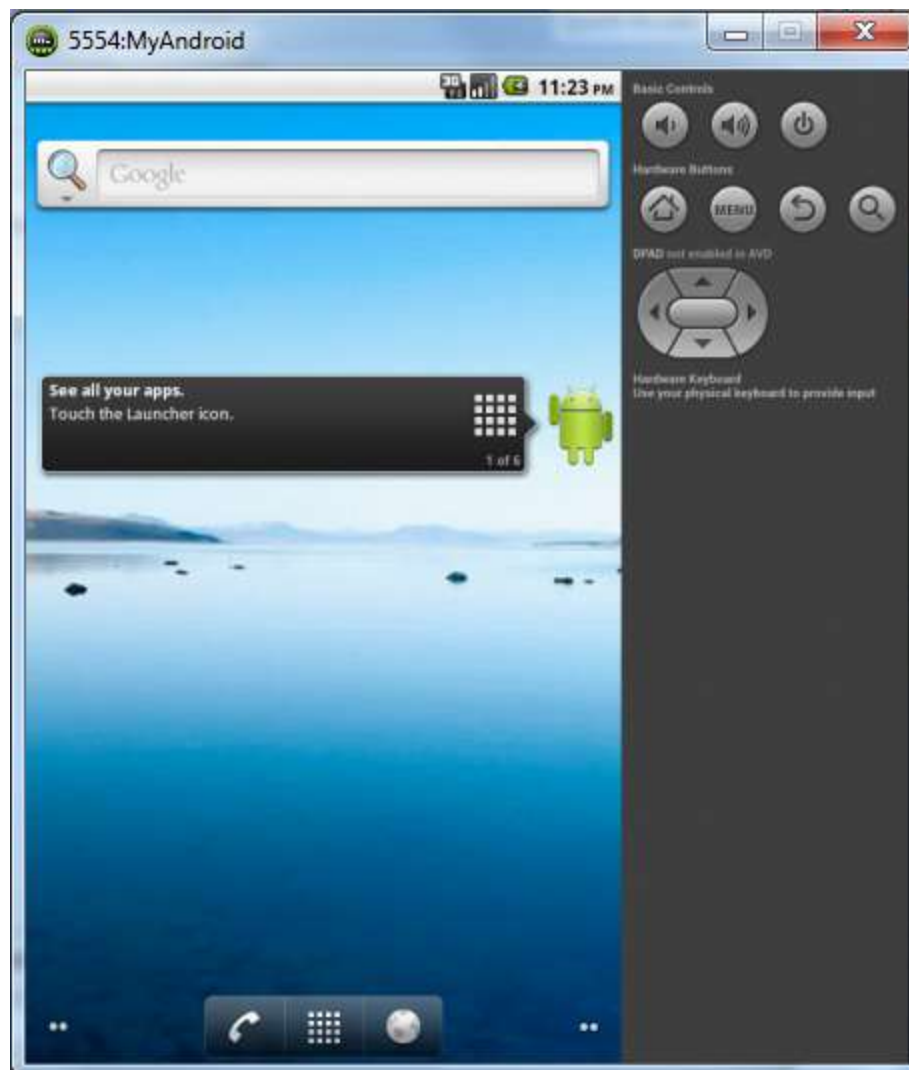
The first screen of virtual AVD

Application First Screen

Application second page view

The application weather page view
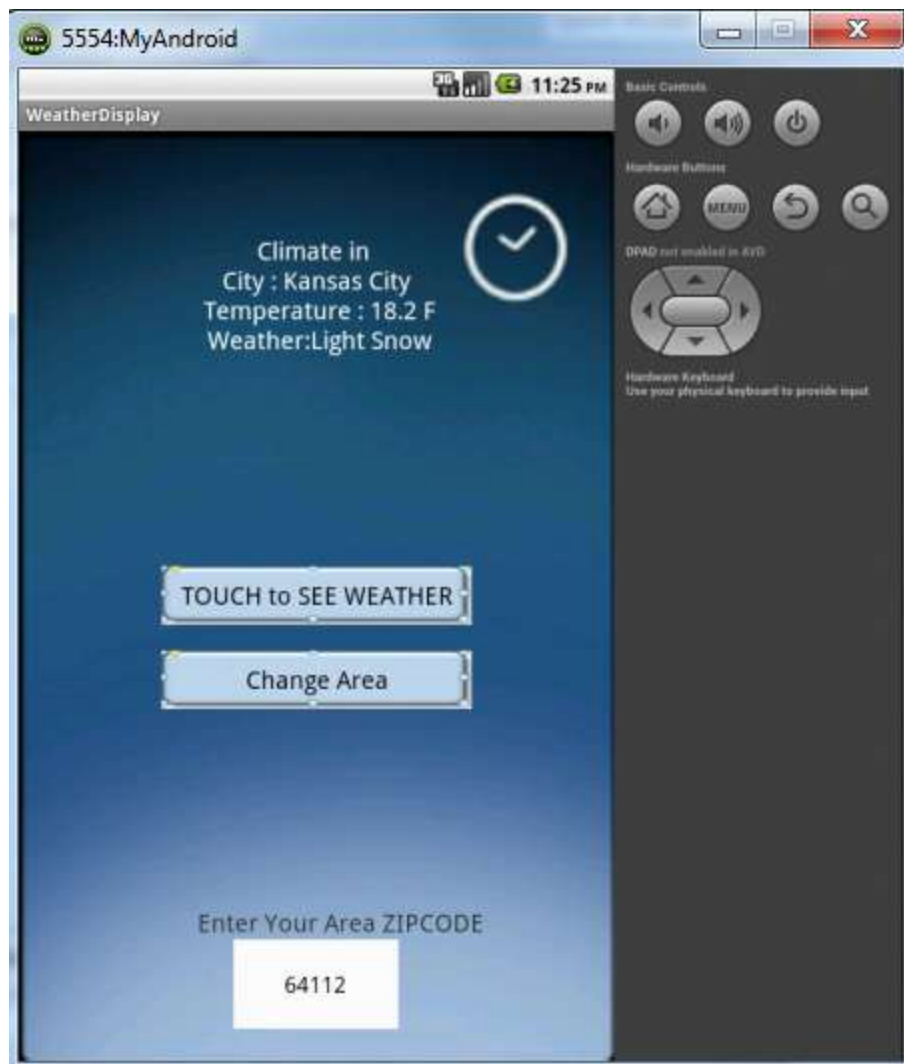
View of application after clicking on Touch to SEE WEATHER

View of application after entering zipcode and clicking on change area

Application Country page view

After entering country the application returning the number of states and teritories
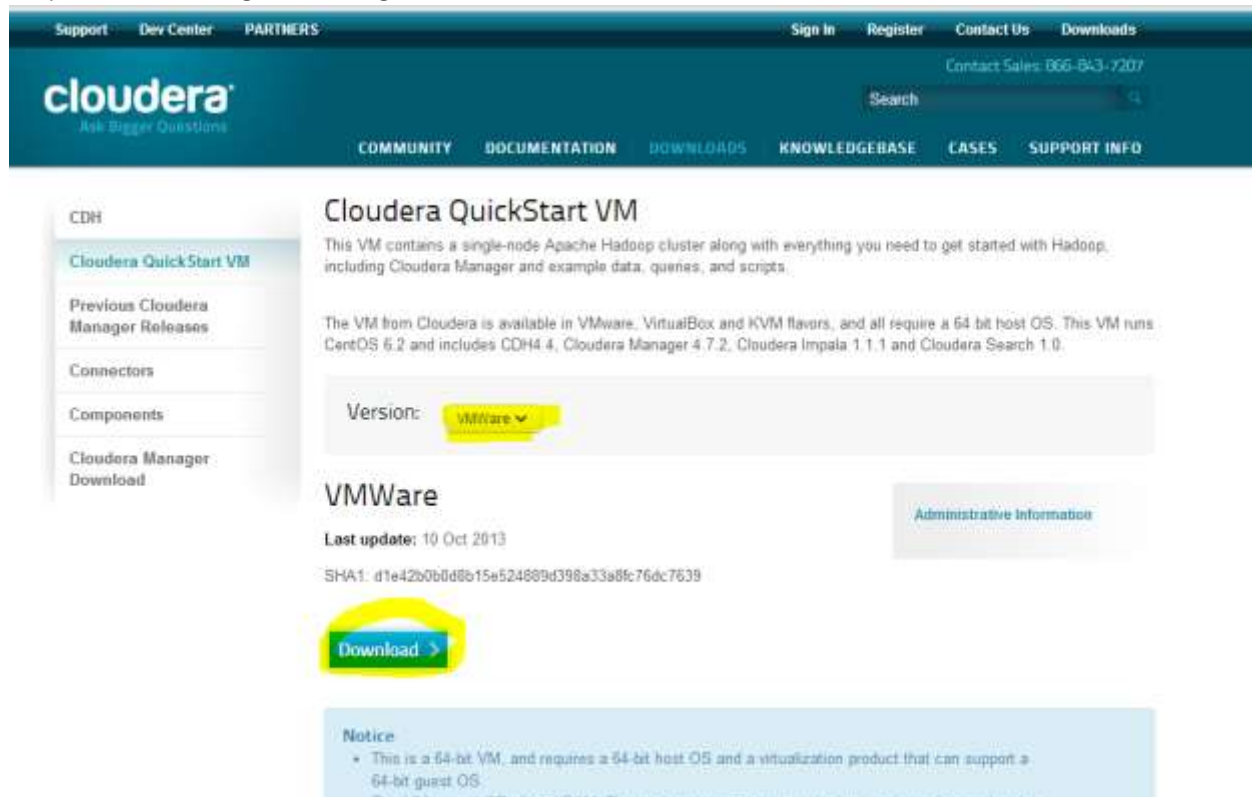
Finally an android application is successful designed to show weather by zipcode input and Country details with country name as input.

The additional files are uploaded in Github lab3 Folder

Link : https://github.com/Mahesh-Vemula/KDM

2. **Cloudera/MapReduce**: Download the Cloudera Image, implement the WordCount MapReduce and run it.

Step 1: Downloading cloud image



Step 2: open image using VM player

Step 3: open cloudera manager entering the credentials.



Step 4 : Downloading Word Count file and running using terminal. Download Page:
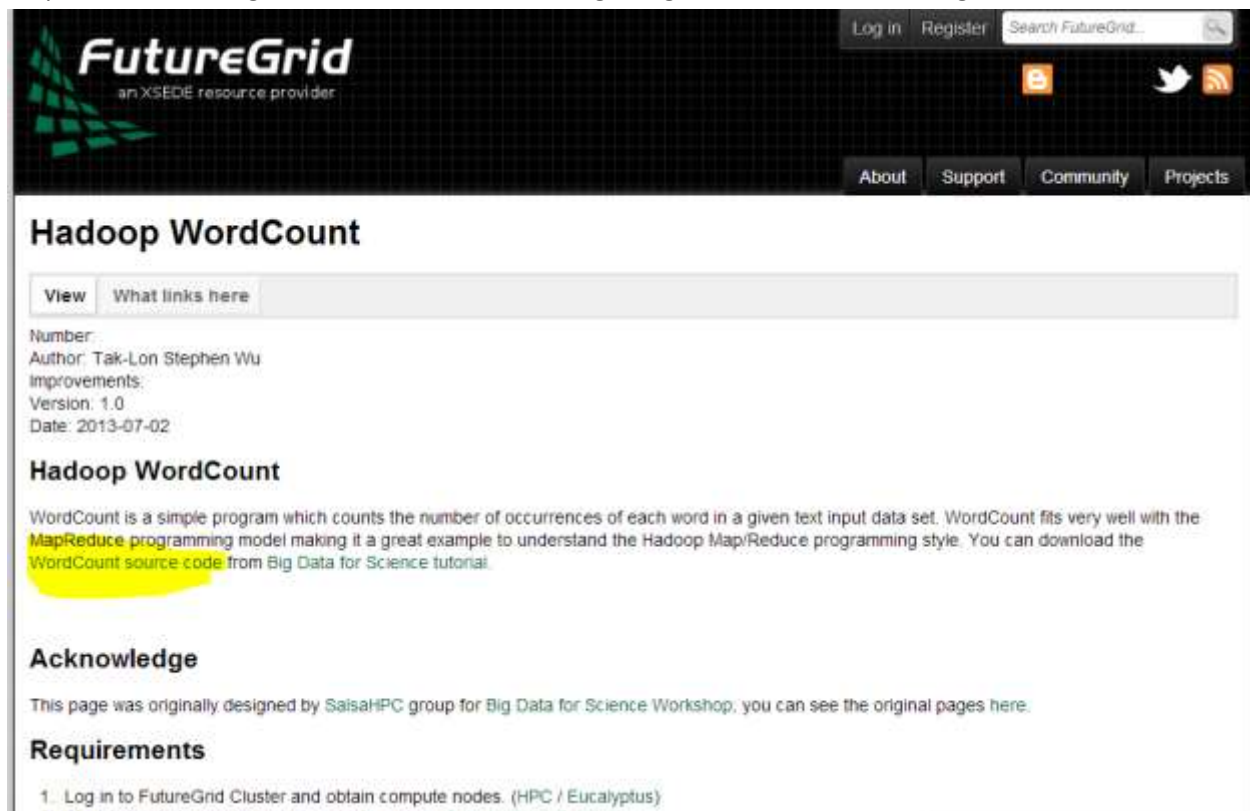
Step 5: Extracting the files

File  Edit  View  Search  Terminal  Help

```
[cloudera@localhost ~]$ dir
datasets  Documents  eclipse  Music     Public     Videos
Desktop   Downloads  lib      Pictures  Templates  workspace
[cloudera@localhost ~]$ cd Downloads
[cloudera@localhost Downloads]$ dir
hadoop-0.20.203.0-customized.tar.gz  Hadoop-WordCount  Hadoop-WordCount.zip
[cloudera@localhost Downloads]$ cd Hadoop-WordCount
```
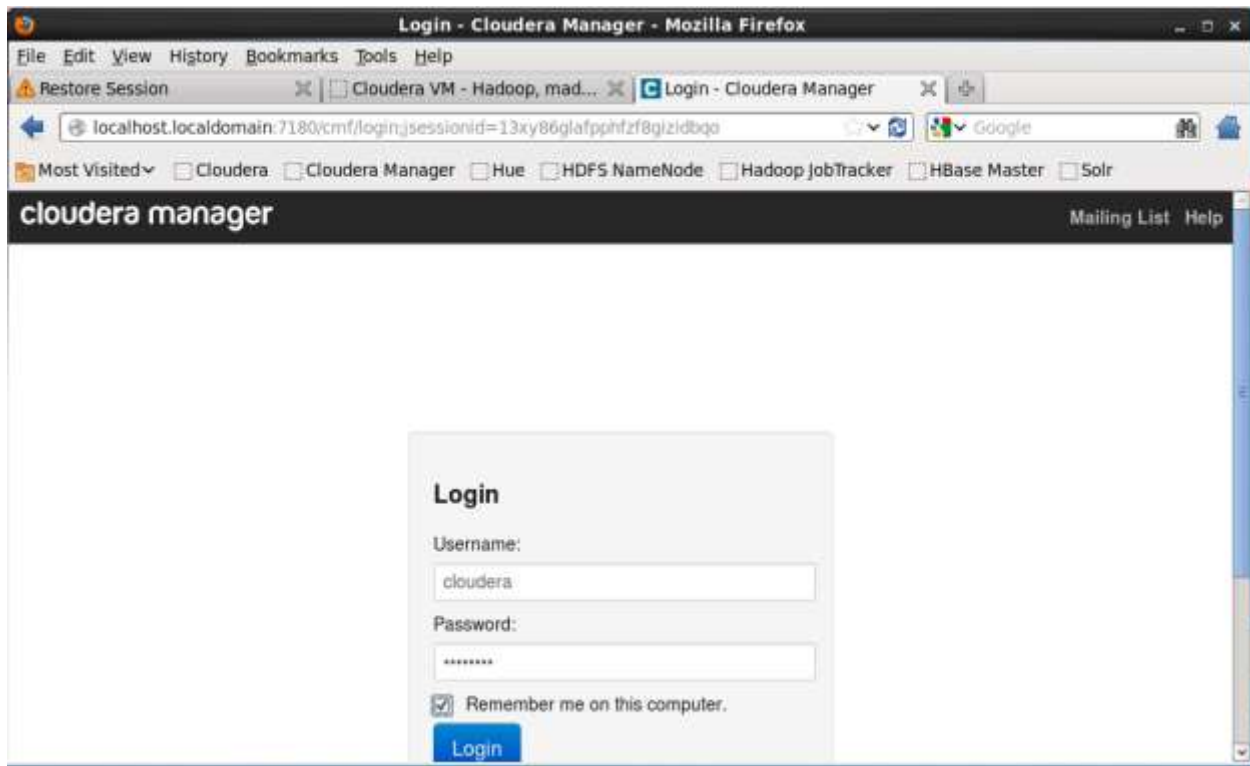
Step 5: Input text folder to processing folder

File  Edit  View  Search  Terminal  Help

```
[cloudera@localhost ~]$ cd Downloads/
[cloudera@localhost Downloads]$ cd Hadoop-WordCount
[cloudera@localhost Hadoop-WordCount]$ ls
build.sh  classes  clean.sh  input  wordcount.jar  WordCount.java
[cloudera@localhost Hadoop-WordCount]$ hadoop fs -put input/Word_Count_input.txt
 input2
put: `input2': File exists
[cloudera@localhost Hadoop-WordCount]$ hadoop fs -put input/Word_Count_input.txt
 input5
[cloudera@localhost Hadoop-WordCount]$ hadoop jar wordcount.jar WordCount input5
 output5
14/02/13 10:10:33 WARN mapred.JobClient: Use GenericOptionsParser for parsing th
e arguments. Applications should implement Tool for the same.
14/02/13 10:10:35 INFO input.FileInputFormat: Total input paths to process : 1
14/02/13 10:10:43 INFO mapred.JobClient: Running job: job_201402121956_0001
14/02/13 10:10:44 INFO mapred.JobClient:  map 0% reduce 0%
```

Step 6: now running the word count

```
cloudera@localhost:~/Downloads/Hadoop-WordCount                    _  □  ×
File  Edit  View  Search  Terminal  Help
14/02/13 10:11:32 INFO mapred.JobClient:      Total time spent by all maps in occ
upied slots (ms)=30503
14/02/13 10:11:32 INFO mapred.JobClient:      Total time spent by all reduces in
occupied slots (ms)=9825
14/02/13 10:11:32 INFO mapred.JobClient:      Total time spent by all maps waitin
g after reserving slots (ms)=0
14/02/13 10:11:32 INFO mapred.JobClient:      Total time spent by all reduces wai
ting after reserving slots (ms)=0
14/02/13 10:11:32 INFO mapred.JobClient:    Map-Reduce Framework
14/02/13 10:11:32 INFO mapred.JobClient:      Map input records=9488
14/02/13 10:11:32 INFO mapred.JobClient:      Map output records=67825
14/02/13 10:11:32 INFO mapred.JobClient:      Map output bytes=643386
14/02/13 10:11:32 INFO mapred.JobClient:      Input split bytes=119
14/02/13 10:11:32 INFO mapred.JobClient:      Combine input records=67825
14/02/13 10:11:32 INFO mapred.JobClient:      Combine output records=11900
14/02/13 10:11:32 INFO mapred.JobClient:      Reduce input groups=11900
14/02/13 10:11:32 INFO mapred.JobClient:      Reduce shuffle bytes=86987
14/02/13 10:11:32 INFO mapred.JobClient:      Reduce input records=11900
14/02/13 10:11:32 INFO mapred.JobClient:      Reduce output records=11900
14/02/13 10:11:32 INFO mapred.JobClient:      Spilled Records=23800
14/02/13 10:11:32 INFO mapred.JobClient:      CPU time spent (ms)=5380
14/02/13 10:11:32 INFO mapred.JobClient:      Physical memory (bytes) snapshot=24
0623616
14/02/13 10:11:32 INFO mapred.JobClient:      Virtual memory (bytes) snapshot=133
6045568
14/02/13 10:11:32 INFO mapred.JobClient:      Total committed heap usage (bytes)=
171315200
[cloudera@localhost Hadoop-WordCount]$
```

Step 7: output file
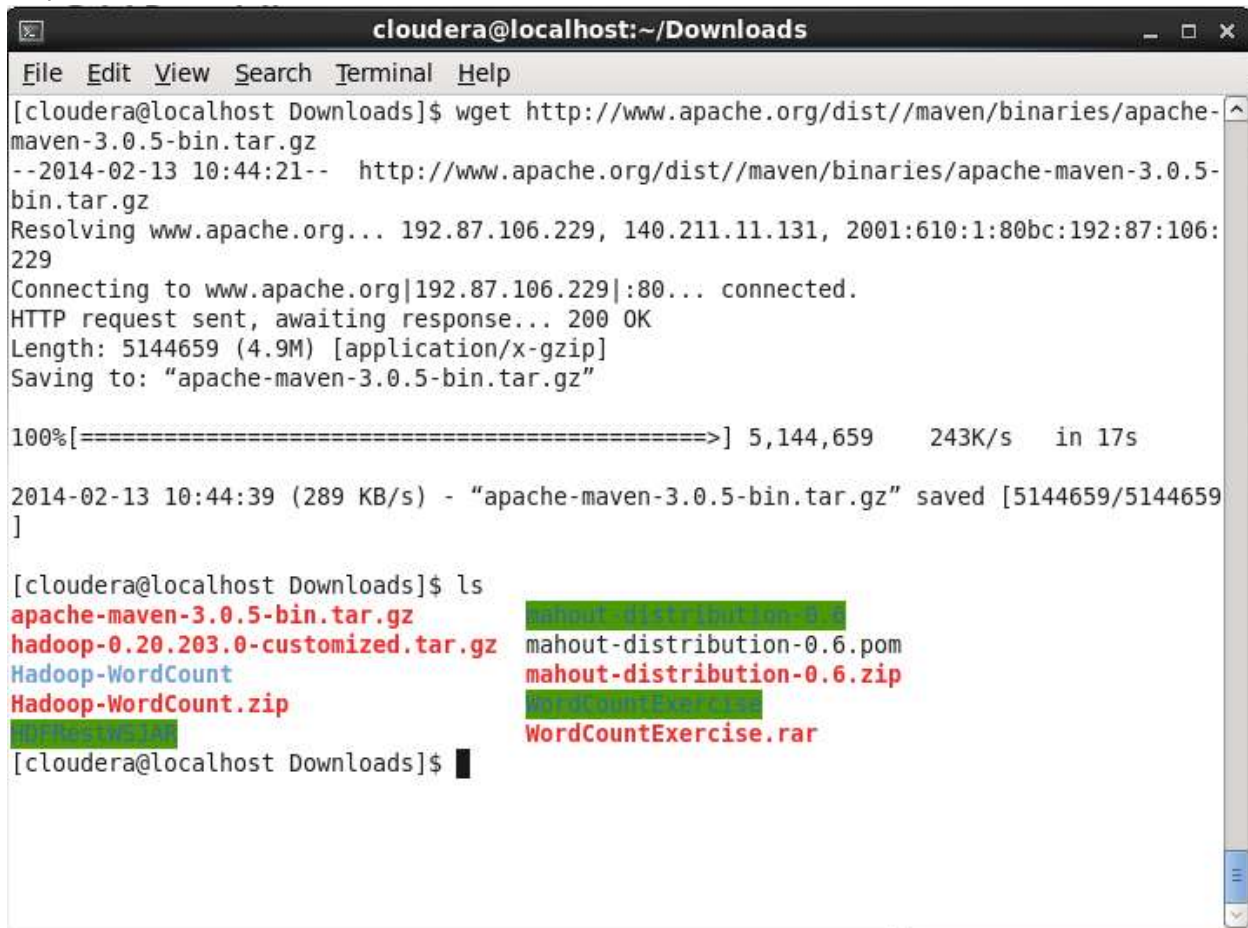


```
you.      14
you."     15
you;      1
you?      2
you?"     23
young     42
young,    7
young.    2
young."  1
younger  1
youngest          3
your      74
yours,"  1
yours?"  1
yourself          5
yourself,         1
yourself.         1
yourself."        1
yourself;         1
youth     7
youth,   1
youth.   1
youthful          1
[cloudera@localhost ~]$
```

3. **Cloudera/Mahout**: Configure your Cloudera with Mahout. Run Naive Bayes classifier with the input data

Step 1: Download maven

```
cloudera@localhost:~/Downloads                                    _ □ ✕

File  Edit  View  Search  Terminal  Help

[cloudera@localhost Downloads]$ wget http://www.apache.org/dist//maven/binaries/apache-
maven-3.0.5-bin.tar.gz
--2014-02-13 10:44:21--  http://www.apache.org/dist//maven/binaries/apache-maven-3.0.5-
bin.tar.gz
Resolving www.apache.org... 192.87.106.229, 140.211.11.131, 2001:610:1:80bc:192:87:106:
229
Connecting to www.apache.org|192.87.106.229|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 5144659 (4.9M) [application/x-gzip]
Saving to: "apache-maven-3.0.5-bin.tar.gz"

100%[====================================================>] 5,144,659    243K/s   in 17s

2014-02-13 10:44:39 (289 KB/s) - "apache-maven-3.0.5-bin.tar.gz" saved [5144659/5144659
]

[cloudera@localhost Downloads]$ ls
apache-maven-3.0.5-bin.tar.gz          mahout-distribution-0.6
hadoop-0.20.203.0-customized.tar.gz    mahout-distribution-0.6.pom
Hadoop-WordCount                       mahout-distribution-0.6.zip
Hadoop-WordCount.zip                   WordCountExercise
HDPRestWSJAR                           WordCountExercise.rar
[cloudera@localhost Downloads]$ █
```

Step 2: Extract the file

```
                              cloudera@localhost:~                       _ □ ✕
File   Edit   View   Search   Terminal   Help
apache-maven-3.0.5/lib/maven-artifact-3.0.5.jar
apache-maven-3.0.5/lib/maven-plugin-api-3.0.5.jar
apache-maven-3.0.5/lib/sisu-inject-plexus-2.3.0.jar
apache-maven-3.0.5/lib/sisu-inject-bean-2.3.0.jar
apache-maven-3.0.5/lib/sisu-guice-3.1.0-no_aop.jar
apache-maven-3.0.5/lib/sisu-guava-0.9.9.jar
apache-maven-3.0.5/lib/maven-model-builder-3.0.5.jar
apache-maven-3.0.5/lib/maven-aether-provider-3.0.5.jar
apache-maven-3.0.5/lib/aether-api-1.13.1.jar
apache-maven-3.0.5/lib/aether-spi-1.13.1.jar
apache-maven-3.0.5/lib/aether-util-1.13.1.jar
apache-maven-3.0.5/lib/aether-impl-1.13.1.jar
apache-maven-3.0.5/lib/maven-compat-3.0.5.jar
apache-maven-3.0.5/lib/wagon-provider-api-2.4.jar
apache-maven-3.0.5/lib/commons-cli-1.2.jar
apache-maven-3.0.5/lib/wagon-http-2.4-shaded.jar
apache-maven-3.0.5/lib/wagon-file-2.4.jar
apache-maven-3.0.5/lib/aether-connector-wagon-1.13.1.jar
[cloudera@localhost Downloads]$ ls
apache-maven-3.0.5                 Hadoop-WordCount.zip
apache-maven-3.0.5-bin.tar.gz     HDFRestWSJAR
hadoop-0.20.203.0-customized.tar.gz  WordCountExercise
Hadoop-WordCount                  WordCountExercise.rar
[cloudera@localhost Downloads]$ ca apache-maven-3.0.5
```
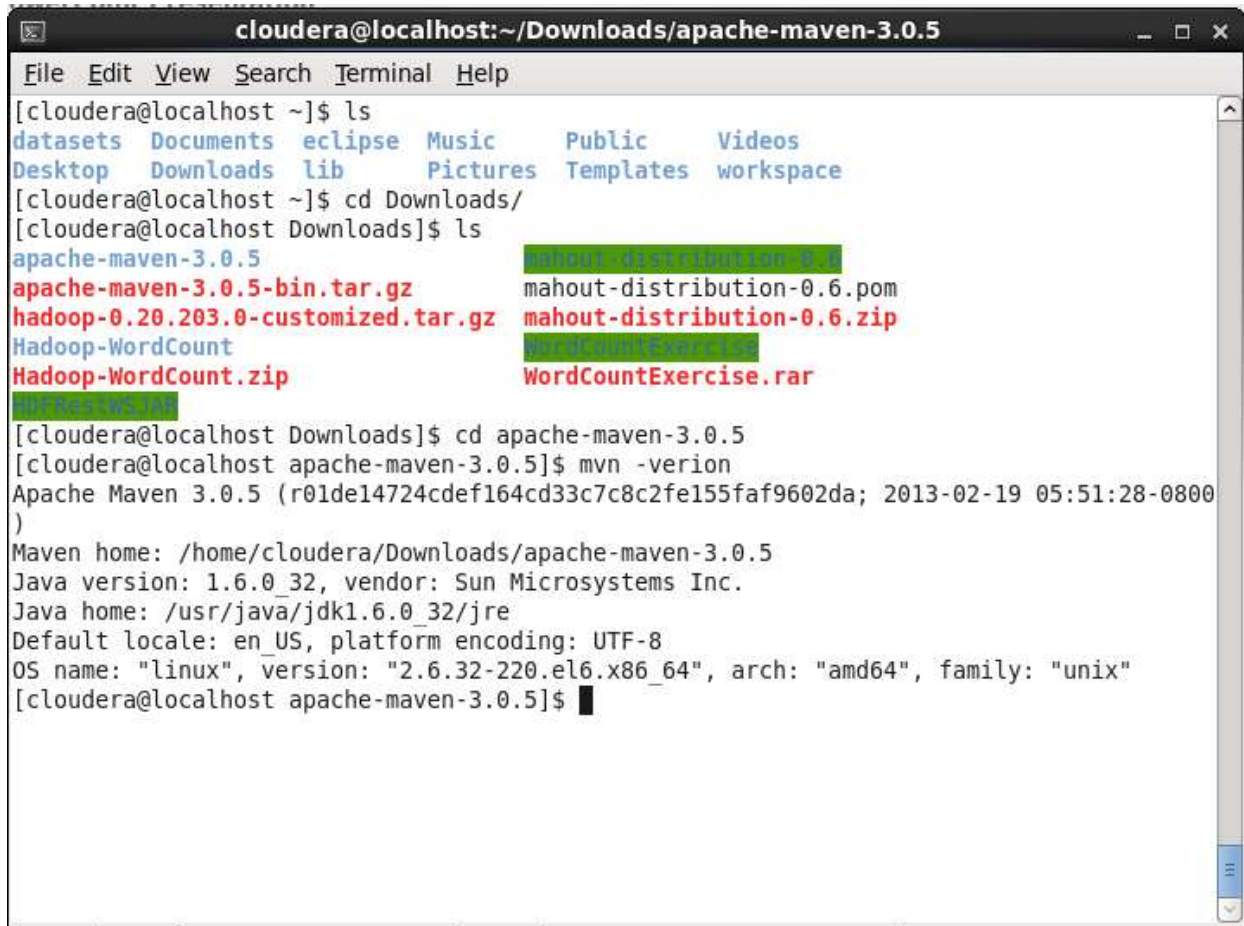
Step 3: Add path for maven

```
[cloudera@localhost ~]$ export M2_HOME=/home/cloudera/Downloads/apache-maven-3.0
.5
[cloudera@localhost ~]$ export PATH=${M2_HOME}/bin:${PATH}
[cloudera@localhost ~]$ mvn --version
Apache Maven 3.0.5 (r01de14724cdef164cd33c7c8c2fe155faf9602da; 2013-02-19 05:51:
28-0800)
Maven home: /home/cloudera/Downloads/apache-maven-3.0.5
Java version: 1.6.0_32, vendor: Sun Microsystems Inc.
Java home: /usr/java/jdk1.6.0_32/jre
Default locale: en_US, platform encoding: UTF-8
OS name: "linux", version: "2.6.32-220.el6.x86_64", arch: "amd64", family: "unix
"
[cloudera@localhost ~]$
```

Step 4: Download .pom and .zip files to install

File   Edit   View   History   Bookmarks   Tools   Help

about:sessionrestore          ✕      Cloudera VM - Hadoop, ma... ✕    Blackboard Learn                    ✕

repo2.**maven.org**/maven2/org/apache/mahout/mahout-distribution/0.6/

Most Visited ✓     Cloudera   Cloudera Manager   Hue   HDFS NameNode   Hadoop JobTracker

| | | |
|---|---|---|
| mahout-distribution-0.6-src.tar.gz.md5 | 01-Feb-2012 13:46 | 32 |
| mahout-distribution-0.6-src.tar.gz.sha1 | 01-Feb-2012 13:46 | 40 |
| mahout-distribution-0.6-src.zip | 01-Feb-2012 13:46 | 10570424 |
| mahout-distribution-0.6-src.zip.asc | 01-Feb-2012 13:46 | 881 |
| mahout-distribution-0.6-src.zip.asc.md5 | 01-Feb-2012 13:46 | 32 |
| mahout-distribution-0.6-src.zip.asc.sha1 | 01-Feb-2012 13:46 | 40 |
| mahout-distribution-0.6-src.zip.md5 | 01-Feb-2012 13:46 | 32 |
| mahout-distribution-0.6-src.zip.sha1 | 01-Feb-2012 13:46 | 40 |
| mahout-distribution-0.6.pom | 01-Feb-2012 13:38 | 3948 |
| mahout-distribution-0.6.pom.asc | 01-Feb-2012 13:46 | 881 |
| mahout-distribution-0.6.pom.asc.md5 | 01-Feb-2012 13:46 | 32 |
| mahout-distribution-0.6.pom.asc.sha1 | 01-Feb-2012 13:46 | 40 |
| mahout-distribution-0.6.pom.md5 | 01-Feb-2012 13:38 | 32 |
| mahout-distribution-0.6.pom.sha1 | 01-Feb-2012 13:38 | 40 |
| mahout-distribution-0.6.tar.bz2 | 01-Feb-2012 13:41 | 58676099 |
| mahout-distribution-0.6.tar.bz2.asc | 01-Feb-2012 13:46 | 881 |
| mahout-distribution-0.6.tar.bz2.asc.md5 | 01-Feb-2012 13:46 | 32 |
| mahout-distribution-0.6.tar.bz2.asc.sha1 | 01-Feb-2012 13:46 | 40 |
| mahout-distribution-0.6.tar.bz2.md5 | 01-Feb-2012 13:43 | 32 |
| mahout-distribution-0.6.tar.bz2.sha1 | 01-Feb-2012 13:43 | 40 |
| mahout-distribution-0.6.tar.gz | 01-Feb-2012 13:38 | 60072260 |
| mahout-distribution-0.6.tar.gz.asc | 01-Feb-2012 13:46 | 881 |
| mahout-distribution-0.6.tar.gz.asc.md5 | 01-Feb-2012 13:46 | 32 |
| mahout-distribution-0.6.tar.gz.asc.sha1 | 01-Feb-2012 13:46 | 40 |
| mahout-distribution-0.6.tar.gz.md5 | 01-Feb-2012 13:41 | 32 |
| mahout-distribution-0.6.tar.gz.sha1 | 01-Feb-2012 13:41 | 40 |
| mahout-distribution-0.6.zip | 01-Feb-2012 13:43 | 65070887 |
| mahout-distribution-0.6.zip.asc | 01-Feb-2012 13:46 | 881 |
| mahout-distribution-0.6.zip.asc.md5 | 01-Feb-2012 13:46 | 32 |
| apache/mahout/mahout-   .zip.asc.sha1 | 01-Feb-2012 13:46 | 40 |

illa Firefox

Step 5: Install and check installation

```
cloudera@localhost:~/Downloads/apache-maven-3.0.5

File  Edit  View  Search  Terminal  Help
[cloudera@localhost ~]$ ls
datasets   Documents   eclipse   Music      Public      Videos
Desktop    Downloads   lib       Pictures   Templates   workspace
[cloudera@localhost ~]$ cd Downloads/
[cloudera@localhost Downloads]$ ls
apache-maven-3.0.5                       mahout-distribution-0.6
apache-maven-3.0.5-bin.tar.gz            mahout-distribution-0.6.pom
hadoop-0.20.203.0-customized.tar.gz      mahout-distribution-0.6.zip
Hadoop-WordCount                         WordCountExercise
Hadoop-WordCount.zip                     WordCountExercise.rar
HDFRestWSJAR
[cloudera@localhost Downloads]$ cd apache-maven-3.0.5
[cloudera@localhost apache-maven-3.0.5]$ mvn -verion
Apache Maven 3.0.5 (r01de14724cdef164cd33c7c8c2fe155faf9602da; 2013-02-19 05:51:28-0800
)
Maven home: /home/cloudera/Downloads/apache-maven-3.0.5
Java version: 1.6.0_32, vendor: Sun Microsystems Inc.
Java home: /usr/java/jdk1.6.0_32/jre
Default locale: en_US, platform encoding: UTF-8
OS name: "linux", version: "2.6.32-220.el6.x86_64", arch: "amd64", family: "unix"
[cloudera@localhost apache-maven-3.0.5]$
```

Step 6: Now running Navie Bayes on data

```
cloudera@localhost:~/Downloads/mahout-distribution-0.6/examples/bin    _  □  ×

File  Edit  View  Search  Terminal  Help

apache-maven-3.0.5                      mahout-distribution-0.6
apache-maven-3.0.5-bin.tar.gz           mahout-distribution-0.6.pom
hadoop-0.20.203.0-customized.tar.gz     mahout-distribution-0.6.zip
Hadoop-WordCount                        WordCountExercise
Hadoop-WordCount.zip                    WordCountExercise.rar
HDFRestWSJAR
[cloudera@localhost Downloads]$ cd mahout-distribution-0.6
[cloudera@localhost mahout-distribution-0.6]$ ls
bin            docs          mahout-core-0.6.jar      mahout-integration-0.6.jar
buildtools     examples      mahout-core-0.6-job.jar  mahout-math-0.6.jar
conf           integration   mahout-distribution-0.6.pom  ...
core           lib           mahout-examples-0.6.jar  NOTICE.txt
distribution   LICENSE.txt   mahout-examples-0.6-job.jar  README.txt
[cloudera@localhost mahout-distribution-0.6]$ cd examples/
[cloudera@localhost examples]$ ls
bin   src   target
[cloudera@localhost examples]$ cd bin
[cloudera@localhost bin]$ ls
asf-email-examples.sh               classify-20newsgroups.sh      factorize-netflix.sh
build-asf-email.sh                  cluster-reuters.sh            lda.algorithm
build-cluster-syntheticcontrol.sh   cluster-syntheticcontrol.sh   README.txt
build-reuters.sh                    factorize-movielens-1M.sh
[cloudera@localhost bin]$ ./classify-20newsgroups.sh
Please select a number to choose the corresponding task to run
1. naivebayes
2. sgd
3. clean -- cleans up the work area in /tmp/mahout-work-cloudera
Enter your choice : █
```

Ouput of Navie bayes classification

File   Edit   View   Search   Terminal   Help

```
txt
14/02/13 10:25:03 INFO bayes.TestClassifier: ==========================================
=============
Summary
-------------------------------------------------------
Correctly Classified Instances          :     6019        79.9124%
Incorrectly Classified Instances        :     1513        20.0876%
Total Classified Instances              :     7532

=======================================================
Confusion Matrix
-------------------------------------------------------
a       b       c       d       e       f       g       h       i       j       k       l
        m       n       o       p       q       r       s       t       <--Classified a
s
383     0       9       0       0       0       0       3       0       0           0       0
        0       1       0       0       1       0       0       0       |   397         a
        = rec.sport.baseball
4       370     0       1       0       4       1       1       0       0       1       0
        1       3       0       0       1       7       0       2       |   396         b
        = sci.crypt
9       2       382     0       1       1       0       2       0       0           0       0
        0       0       0       1       0       1       0       0       |   399         c
        = rec.sport.hockey
2       12      0       327     1       5       0       2       0       0           0       1
        1       0       2       4       2       0       1       4       |   364         d
        = talk.politics.guns
4       0       1       0       372     1       2       1       0       2           0       0
        1       0       3       2       0       3       0       6       |   398         e
        = soc.religion.christian
2       14      0       0       0       324     5       8       0       0           0       1
        10      5       2       6       4       8       0       4       |   393         f
        = sci.electronics
```