CS5560- KDM – Lab#3 Assignment

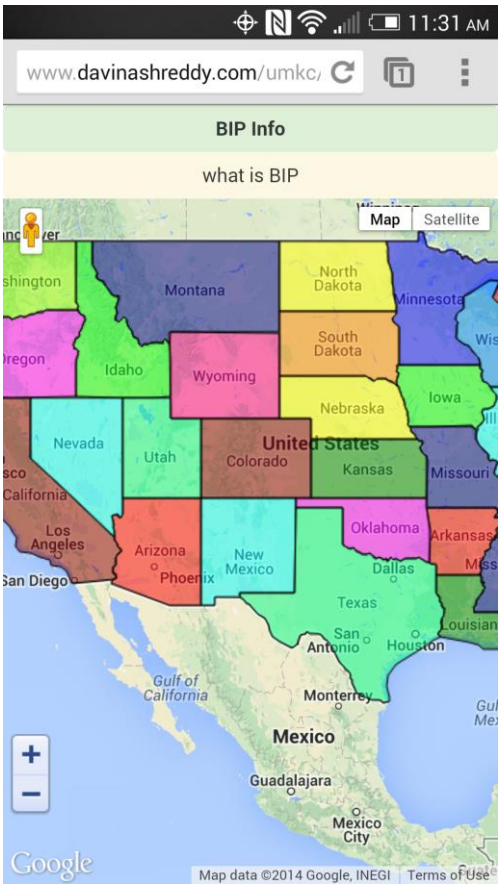
This document consists of following information.

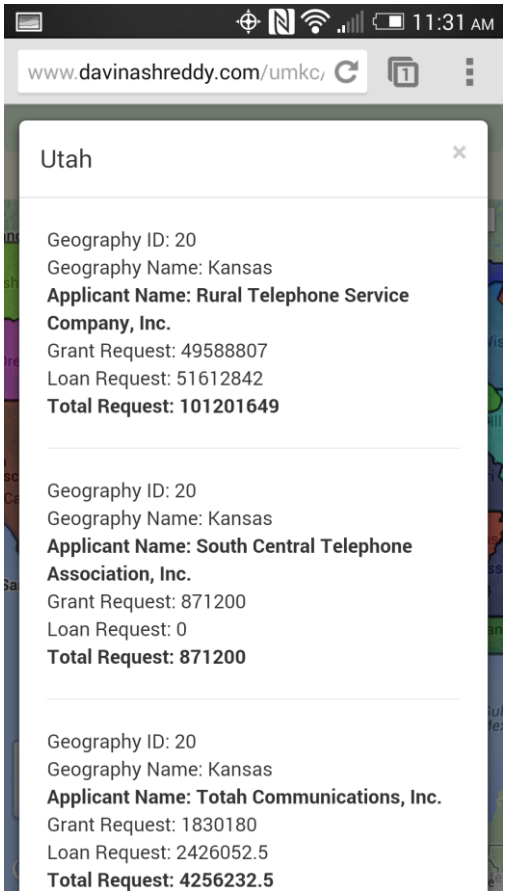
1. Make a **Mashup application** including various services (e.g., Google Map, Google Chart, Google Search, Yahoo, Amazon, Twitter, Facebook) Web Services (e.g., Google Map Services, Weather Services) using either (1) Mobile Web Technology with HMLT5 Local DB (Refer to Tutorial 3).
2. **Cloudera/MapReduce:** Download the Cloudera Image, implement the WordCount MapReduce and run it. (a bonus point for implementing a new MapReduce algorithm) The code and guidelines will be available in Tutorials/Tutorial 5.
3. **Cloudera/Mahout:** Configure your Cloudera with Mahout. Run Naive Bayes classifier with the input data (a bonus point for using your own data).

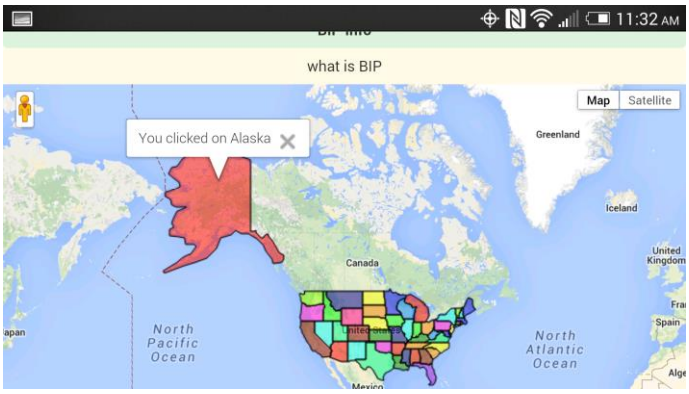
- GitHub Account Screenshots(username: davinashreddy)
- ScrumDo Account Screenshots(link: https://www.scrumdo.com/organization/umkc32/dashboard)

---

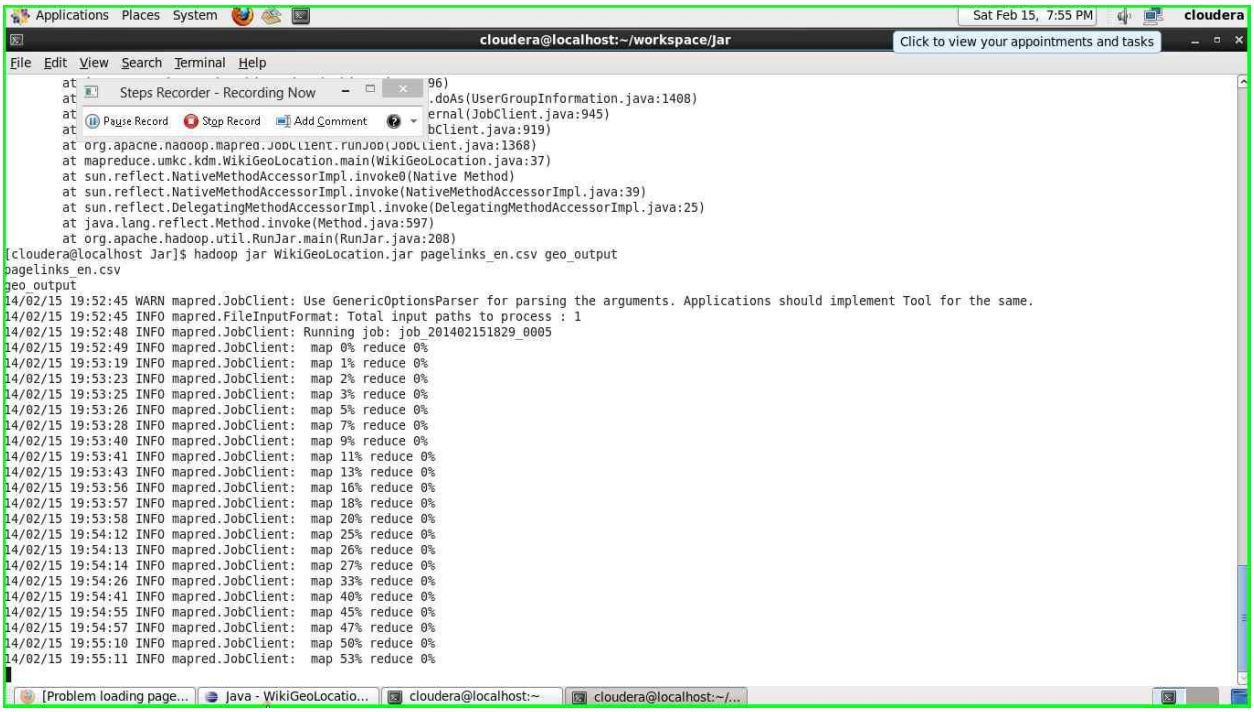
1. **Mashup Application:**
    - My application includes Google map service on which States in united states have an overlay with different color and name.
    - When you click on a specific state then a service call is made to the following url and the BIP information of the state is retrieved.
        - BIP's goal is to provide financial boost to the nation during the economic crisis [1].
        - http://www.broadbandmap.gov/broadbandmap/bip/states/kansas?format=json – Sample Service call, to Kansas statistics.
        - http://www.responsinator.com/?url=http%3A%2F%2Fwww.davinashreddy.com%2Fumkc%2Fkdm-lab3%2F
        - http://www.davinashreddy.com/umkc/kdm-lab3/
    - Following are some screenshots taken in HTC one device in chrome browser.

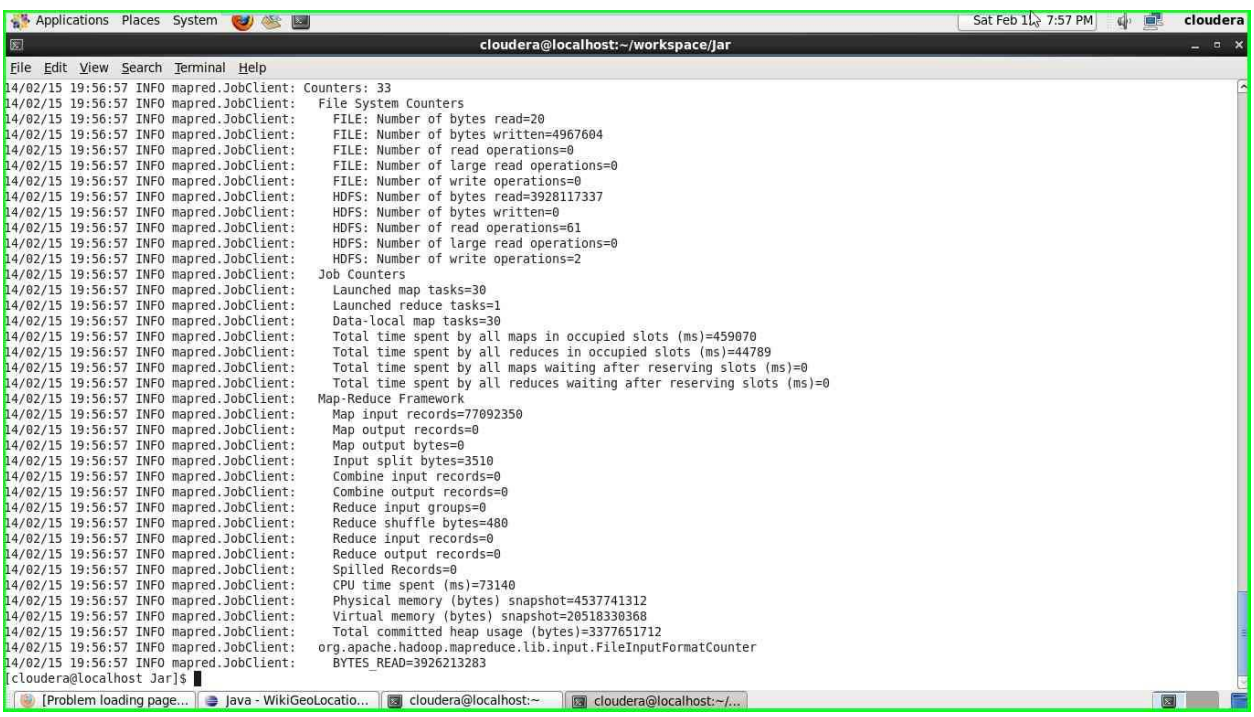

11:31 AM
www.davinashreddy.com/umkc/
BIP Info
what is BIP
Map
Satellite
Montana
North Dakota
Minnesota
South Dakota
Idaho
Wyoming
Nebraska
Iowa
Nevada
Utah
Colorado
Kansas
Missouri
California
Los Angeles
San Diego
Arizona
Phoenix
New Mexico
Oklahoma
Arkansas
Dallas
Texas
San Antonio
Houston
Gulf of California
Monterrey
Mexico
Guadalajara
Mexico City
Google
Map data ©2014 Google, INEGI
Terms of Use


11:31 AM
www.davinashreddy.com/umkc/
Utah
Geography ID: 20
Geography Name: Kansas
Applicant Name: Rural Telephone Service Company, Inc.
Grant Request: 49588807
Loan Request: 51612842
Total Request: 101201649
Geography ID: 20
Geography Name: Kansas
Applicant Name: South Central Telephone Association, Inc.
Grant Request: 871200
Loan Request: 0
Total Request: 871200
Geography ID: 20
Geography Name: Kansas
Applicant Name: Totah Communications, Inc.
Grant Request: 1830180
Loan Request: 2426052.5
Total Request: 4256232.5

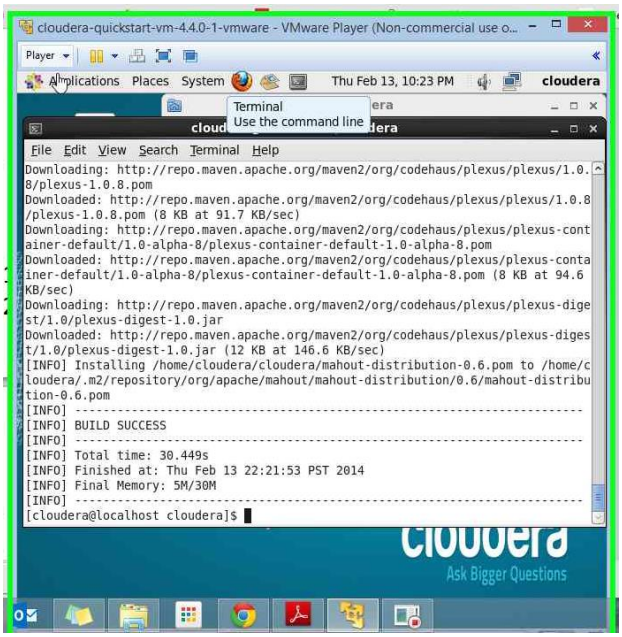

11:32 AM
what is BIP
Map
Satellite
You clicked on Alaska
Greenland
Iceland
Canada
North Pacific Ocean
North Atlantic Ocean

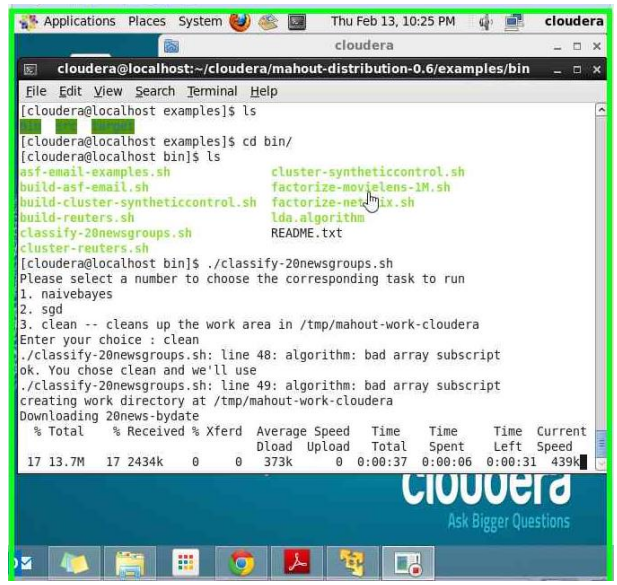
2. **Cloudera/Mapreduce**
    - My MapRecude program runs on DB-Pedia GeoLocation data where the goal is to extract geo-location of specific URL [2].
    - Source code is available in GitHub
    - Screenshots of my work as follows.

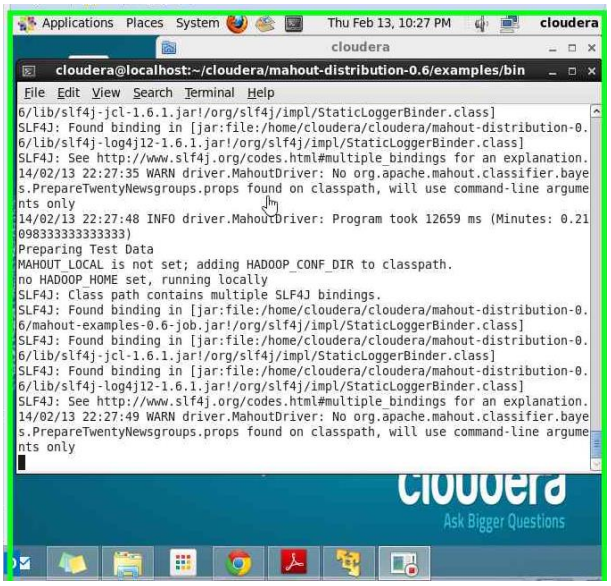


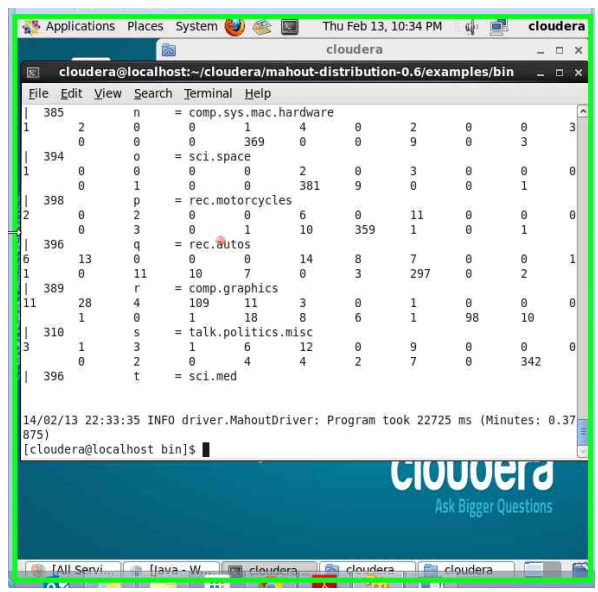


3. **Cloudera/Mahout:**
   To run mahout in Cloudera CentOS edition, Maven has to be installed. Steps to install Maven follows:
   - Download the maven and mahout repos along with .pom file which installs all the dependencies required.
   - Using mahout .pom file install mahout using maven by running
     - mvn –f filename.pom –DskipTests install
   - Set the permissions for Mahout and run Naviebayes.
   - Below are the screenshots of my work.









References

1. http://www.rurdev.usda.gov/utp_bip.html
2. Automatic IO Filtering for Optimizing Cloud Analytics Microsoft Technical Report MSR-TR-2012-3 by *Christos Gkantsidis at Microsoft Research, Cambridge, UK.*